

## Stakeholder interpretation of probabilistic representations of uncertainty in spatial information: an example on the nutritional quality of staple crops

Christopher Chagumaira, Patson C. Nalivata, Joseph G. Chimungu, Dawd Gashu, Martin R. Broadley, Alice E. Milne & R. Murray Lark

To cite this article: Christopher Chagumaira, Patson C. Nalivata, Joseph G. Chimungu, Dawd Gashu, Martin R. Broadley, Alice E. Milne & R. Murray Lark (2022): Stakeholder interpretation of probabilistic representations of uncertainty in spatial information: an example on the nutritional quality of staple crops, International Journal of Geographical Information Science, DOI: [10.1080/13658816.2021.2020278](https://doi.org/10.1080/13658816.2021.2020278)

To link to this article: <https://doi.org/10.1080/13658816.2021.2020278>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 01 Mar 2022.



[Submit your article to this journal](#)










[View related articles](#)



[View Crossmark data](#)

## Stakeholder interpretation of probabilistic representations of uncertainty in spatial information: an example on the nutritional quality of staple crops

Christopher Chagumaira<sup>a,b,c</sup> , Patson C. Nalivata<sup>c</sup> , Joseph G. Chimungu<sup>c</sup> , Dawd Gashu<sup>d</sup> , Martin R. Broadley<sup>a,b</sup> , Alice E. Milne<sup>b</sup>  and R. Murray Lark<sup>a</sup> 

<sup>a</sup>Future Food Beacon of Excellence and School of Biosciences, University of Nottingham, Loughborough, UK; <sup>b</sup>Sustainable Agriculture Sciences Department, Rothamsted Research, Hertfordshire, UK; <sup>c</sup>Crop and Soil Sciences Department, Bunda College, Lilongwe University of Agriculture and Natural Resources, Lilongwe, Malawi; <sup>d</sup>Center for Food Science and Nutrition, Addis Ababa University, Addis Ababa, Ethiopia

### ABSTRACT

Spatial information, inferred from samples, is needed for decision-making, but is uncertain. One way to convey uncertain information is with probabilities (e.g. that a value falls below a critical threshold). We examined how different professional groups (agricultural scientists or health and nutrition experts) interpret information, presented this way, when making a decision about interventions to address human selenium (Se) deficiency. The information provided was a map, either of the probability that Se concentration in local staple grain falls below a nutritionally-significant threshold (negative framing) or of the probability that grain Se concentration is above the threshold (positive framing). There was evidence for an effect of professional group and of framing on the decision process. Negative framing led to more conservative decisions; intervention was recommended at a smaller probability that the grain Se is inadequate than if the question were framed positively, and the decisions were more comparable between professional groups under negative framing. Our results show the importance of framing in probabilistic presentations of uncertainty, and of the background of the interpreter. Our experimental approach could be used to elicit threshold probabilities which represent the preferences of stakeholder communities to support them in the interpretation of uncertain information.

### ARTICLE HISTORY


Received 7 April 2021

Accepted 15 December 2021

### KEYWORDS

Decision-making; framing effect; Interpret probability; values mean probability value; Uncertainty

**CONTACT** Christopher Chagumaira  [christopher.chagumaira@nottingham.ac.uk](mailto:christopher.chagumaira@nottingham.ac.uk), [chris.chagumaira@gmail.com](mailto:chris.chagumaira@gmail.com)

 Supplemental data for this article can be accessed [here](#).

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

### 1.1. The problem

There is increasing awareness that, while much progress has been made to address malnutrition with respect to energy and protein supply, micronutrients (such as zinc, iron, iodine and selenium) may remain deficient among populations of many countries (Ligowe *et al.* 2020). This micronutrient deficiency (MND) or ‘hidden hunger’ has implications for human health, growth and cognitive function. In the GeoNutrition project, funded by the Bill and Melinda Gates Foundation, micronutrient studies in soil, crops and the human population are being conducted in Malawi and Ethiopia (Gashu *et al.* 2020, 2021). There is interest in how MND problems may vary spatially due to variation in soil and other environmental conditions. If this occurs, then interventions might be more effectively targeted where particular MND are prevalent.

Through the GeoNutrition project, a large dataset has been collected on soil and crop micronutrient status in Malawi and Ethiopia. This allows the micronutrient concentration in soil and staple crops to be mapped. The spatial predictions are uncertain, but the statistical models on which they are based allow us to compute the probability that a particular micronutrient concentration falls below or above a nutritionally relevant threshold at some unsampled location. It is often suggested that mapping this probability will help interpret the information while allowing for its uncertainty in the spatial data. However, it remains unclear how various stakeholders, for whom such information is required to support decisions on interventions to address MND, would use the probabilities in order to account for uncertainty.

In this paper we describe a study to examine how stakeholders interpret probability that local grain micronutrient concentration falls below a threshold. Groups of stakeholders were provided with different scenarios, in which this probability took different values, and were asked to indicate in which they would recommend an intervention (such as campaign to promote fertiliser to increase crop micronutrient concentration, or the deployment of nutrient supplements or fortified food). We used these responses to estimate and compare the mean probability value at which different stakeholder groups chose to recommend an intervention. We also examined how the framing of the question affected the responses. That is to say, whether the responses of stakeholders presented with a positive framing (probability that the grain Se content is sufficient) would be different to those who were presented information with negative framing (probability that the grain Se content is inadequate). On this basis we aimed to assess the feasibility of using formal elicitation to estimate the threshold probability at which groups of stakeholders would recommend an intervention, as a basis both for examining critically how they interpret probabilistic information and developing rules for interpretation which reflect stakeholder opinion and assumptions.

### 1.2. The general context

Spatial information has uncertainty, which arises from error (location error, measurement error), environmental heterogeneity, and our uncertainty about the interpretation of information (e.g. the vagueness of concepts such as a ‘deep soil’, which play

a part in data interpretation) (Li *et al.* 2018). For this reason it is widely recognized in geographical information science (GIScience) that the uncertainty about spatial information must be communicated to its end-users if they are to apply it effectively (Li *et al.* 2012, Greiner *et al.* 2018). Heuvelink and Burrough (2002) suggested that it is necessary to address how stakeholders deal with problems of uncertainty in spatial information as part of a decision making process. The study reported here fits into that research agenda, and is concerned with how stakeholders make decisions based on comparison of spatial variables to threshold values when the uncertainty about the true value of the variable relative to the threshold is expressed in terms of probability.

A simple and common decision model is where some action is taken at a location if the value of a variable there exceeds (or falls below) a threshold. For example, action must be taken to remediate soil where the concentration of a contaminant exceeds a soil guideline value (Cole and Jeffries 2009) or legislative thresholds (Marchant *et al.* 2017). Fertilisers might be recommended where the measured concentration of a nutrient in soil is smaller than an index value and liming might be recommended where soil pH is less than a threshold. For example, in Malawi, it is recommended that liming should be done when soil pH is below 5.0 (Chilimba *et al.* 2013); whereas, in the UK if soil pH falls below 6.0 in pasture land, then liming is recommended to maintain yield and forage quality (DEFRA 2010). Interventions to address micronutrient deficiencies in human populations can be recommended where measurements of a biomarker (such as concentration of the nutrient in blood serum or urine) falls below a threshold (e.g. Likoswe *et al.* 2020, Phiri *et al.* 2020) or where inferred intake is less than a quantity such as the recommended daily allowance (RDA) or estimated average requirements (EAR) (e.g. Joy *et al.* 2014, 2015).

Such management decisions are usually made in the face of uncertainty because the variable concerned is estimated or predicted from partial data or a model (Goovaerts 1997). Spatial uncertainty can be quantified in a number of ways. In geostatistical mapping, the spatial uncertainty of the predictions is quantified directly by the prediction error variance or the kriging variance. The kriging variance varies spatially, and its values are small in the neighbourhood of sample points and larger further away. The kriging variance is the variance of the prediction distribution at an unsampled site of interest, or the conditional distribution given the data and the geostatistical model. The width of this prediction distribution (indicated by its variance) represents the uncertainty of the predicted value there (Heuvelink 2018). The kriging variance might be mapped directly as an indicator of uncertainty (e.g. Hatvani *et al.* 2021). Alternatively, it might be more accessible to compute prediction intervals from the prediction distribution, that is to say an interval of values which contains the true value at the location with some specified probability (e.g. Karl 2010). These methods are useful to experts familiar with the underlying concepts, but may be inaccessible for decision makers who do not necessarily understand kriging variance. Prediction intervals and kriging variance were the methods of communicating and quantifying uncertainty least-preferred by end-users (Chagumaira *et al.* 2021).

When there are decisions to be made relative to thresholds, spatial uncertainty can be quantified by using probabilities. This uncertainty can be quantified by the probability that the threshold is exceeded or not. Ideally this probability can be obtained

from the prediction distribution of the variable from data and an appropriate statistical model (conditional probability). Marchant *et al.* (2017) took this approach to compute probabilities that arsenic and mercury concentration exceeds soil guidance values and to map this across France. Lark *et al.* (2014) similarly computed the probability that local soil conditions indicate a risk of cobalt deficiency in grazing sheep across part of the north of Ireland. Approaches such as disjunctive kriging (DK) and indicator kriging (IK) are commonly used to compute conditional probabilities (Webster and Oliver 2007). Ordinary kriging may be used, along with an assumption of normal errors. However, indicator kriging is more robust to any failures of this assumption, and is also more resistant to local outliers. Lark *et al.* (2016) used DK to map the probability that soil pH under pasture in the north of Ireland is below 6.0, to indicate where liming would be advised. Goovaerts *et al.* (1997) used IK to map the probability that cadmium concentration exceeds a regulatory threshold at sites across the Swiss Jura, to indicate where remediation might be necessary. Other approaches have been used to compute local probabilities that variables exceed thresholds of environmental significance. These include copulas, conditional simulation and Bayesian methods to compute or sample from a local posterior distribution (Goovaerts 2001, Marchant *et al.* 2011, Greiner *et al.* 2018).

Much work has focused on computing the conditional probability that a variable exceeds a threshold, and there is an implicit assumption that if the stakeholder has been given the probability they will be able to use it to make decisions with the uncertain information (Lark *et al.* 2016). Little attention has been given to how stakeholders might use such information and how they might be helped to do so more consistently and effectively. The use of probability to communicate uncertainty is not straightforward (Milne *et al.* 2015) and probabilities are not always easily interpreted by stakeholders who have to make the decision (Spiegelhalter *et al.* 2011). Because of this, verbal interpretations of probability based on ‘calibrated phrases’ (e.g. ‘unlikely’) have been proposed — e.g. the Intergovernmental Panel for Climate Change (IPCC) scale due to Mastrandrea *et al.* (2010). Although calibrated phrases have been widely used, Budescu *et al.* (2009) showed that they may be interpreted regressively (i.e. any phrase indicating uncertainty about an outcome is thought to indicate that its probability is around 0.5). Furthermore, calibrated phrases may be subject to severity bias, depending on how the outcome of interest is expressed (e.g. if it is stated that ‘severe flooding is very unlikely’ the adjective ‘severe’ influences the assessment of risk more than does the phrase indicating the uncertainty). However, Jenkins *et al.* (2019) showed that stakeholders regard probabilities expressed in numerical form as more credible than calibrated phrases. Chagumaira *et al.* (2021) found that, despite these challenges in interpretation of probabilities, varied stakeholders preferred statements of uncertainty expressed as probabilities to more general measures such as prediction intervals or a prediction error variance.

Spatial uncertainty is an important subject in GIScience (Heuvelink and Burroughs 2002, Li *et al.* 2012, 2018) and presenting spatial datasets together with their uncertainties is necessary because it adds to the quality of spatial information used in decision making. As we have noted, a common approach to presenting uncertain information about the value of a variable relative to a threshold is to compute the

probability that the variable exceeds (or falls below) that threshold. However, we contend that insufficient attention has been given to how stakeholders incorporate such uncertain information into decision-making processes.

A stakeholder, using uncertain information to support a decision, must in effect decide on the probability threshold at or above which they would choose to act as if the threshold was exceeded/not exceeded. Taking the concentration of Se in staple grain as an example, would a stakeholder approve an intervention at a certain location where there were a 50% probability that the concentration of Se falls below the threshold? Would they make the same decision if the probability were 25%, or 75%?

A stakeholder deals with an unknown state, the true value of the environmental variable either indicates that the action should be taken or it does not. They also have a choice of two actions to intervene or not. We might expect that the threshold probability at which a stakeholder would choose to intervene will reflect their assessment of the loss attached to each possible outcomes—the intervention was necessary or not, as determined by the unknown state, under each decision (intervene or not). These losses may reflect factors such as the social, economic, individual and political consequences of failing to address a problem, and the opportunity costs of resources expended on unnecessary intervention. In some cases these losses may be quantified, and used in a formal analysis e.g. Ramsey *et al.* (2002) who considered the losses associated with different decisions and outcomes in the management of contaminated land. However, for many applications the different losses under decisions and outcomes may be complex and hard to quantify. The question that we address in this paper is how and whether one might identify a threshold probability that consistently reflects the perception of the losses by a stakeholder group, and how they weight these, tacitly if not explicitly. Before refining this question, we consider a theoretical framework.

### 1.3. Theory

Let  $L_1$  be the loss incurred if we intervene unnecessarily, where with perfect knowledge we would intervene only if the variable (nutrient concentration)  $z < z_t$ , where  $z$  is the unknown true value and  $z_t$  is the threshold of interest. In this treatment we regard the loss as zero if we intervene appropriately. Let  $L_2$  be the loss incurred if we choose not to intervene, but should have done so. Again, we regard loss as zero if we correctly choose not to intervene. If  $P$  is the probability that the concentration is below the threshold,  $z < z_t$ , then expected loss if we choose to intervene is

$$(1-P)L_1. \quad (1)$$

If we choose not to intervene then the expected loss is

$$PL_2. \quad (2)$$

If we wish to make the decision with the smaller expected loss, a rational assumption, then it follows that we should intervene if  $P$  takes a value such that

$$(1-P)L_1 \leq PL_2, \quad (3)$$

and not intervene otherwise. By simple algebraic rearrangement of Equation (3) we can show that we should intervene if

$$P \geq \frac{L_1}{(L_1 + L_2)}, \quad (4)$$

and not otherwise, that is to say if  $P$  exceeds or equals a threshold value,  $P_t$  where,

$$P_t = \frac{L_1}{L_1 + L_2}. \quad (5)$$

The larger the loss from an unnecessary intervention relative to a failure to intervene where necessary, the larger  $P_t$  must be.

In a situation where  $L_1$  and  $L_2$  can be quantified directly,  $P_t$  could be computed from Equation (5). However, complex real-world problems components of the loss associated with outcomes maybe difficult to quantify (e.g. the political cost of a failure to address a public health problem) and controversial (e.g. do disability adjustment life years, DALYs, lost really capture all the social loss from a failure to act where a nutritional deficiency pertains?) and may not be commensurable. The value of  $P_t$  at which an agent chooses to act therefore reflects a complex judgement.

This study is based on two principles. First, while the provision of conditional probabilities is a natural way to communicate the uncertainty associated with the information of a variable which users of that information will interpret relative to threshold values, the problem of decision-making is not solved by those probabilities. As we have seen, a judgement must still be made. Second, we suggest that one approach to this problem is to elicit a threshold probability from members of relevant stakeholder communities. We assume that an individual stakeholder has a least a tacit sense of the values of  $L_1$  and  $L_2$  that they would assume in making a judgement from conditional probabilities. In principle, then, a suitable process might be used to elicit a value of  $P_t$  from individuals or groups of stakeholders that represent an individual opinion or a group consensus. Such an elicitation would be analogous to the process by which probabilities of unknown states or distribution for uncertainty quantification are formally elicited from expert panels (O'Hagan *et al.* 2006).

The aim of the study reported here was to address the following:

- *Can a consistent (i.e. reasonably precise) estimate of  $P_t$  be elicited from a stakeholder group?*
- *Does the estimated  $P_t$  depend on the specific interests of the group (e.g. does it differ between nutritionists and agronomists)?*
- *Is the estimated  $P_t$  prone to framing effects (i.e. does the estimate depend on how the question is posed)?*

These are practical and useful questions to address. If decisions are to be based on uncertain information then a value of  $P_t$  is required for a decision making and should be obtained by some transparent process in which the underlying questions are examined. The findings of this study should provide a basis for designing a formal procedure to elicit a value of  $P_t$  for this and similar problems. In this study we address these questions, considering a core study concerned with decision on interventions to

improve micronutrient supply based on estimates of the amount provided locally by staple crops. We asked two stakeholder groups individually to identify a threshold probability at which an intervention would be recommended, and used these to estimate an underlying mean value for each group. Furthermore, we investigated whether the framing of the question influenced the responses.

## 2. Method

### 2.1. Basic approach

The approach was to offer respondents a set of scenarios for which the probability that concentration of Se in staple crop is less than the threshold Se concentration ( $Se_{\text{grain}} < t_{\text{Se}}$ ) took a series of values over the range 0–1. For each one they were invited to respond as to whether the intervention would be recommended or not.

The respondents were asked to self-identify as either (i) A public health and nutrition specialist, or (ii) an agronomists and soil scientist. Each respondent was also allocated at random to one of two groups. The first group was presented with a positive framing of the question (i.e. to select a probability that  $Se_{\text{grain}} > t_{\text{Se}}$  below which an intervention would be recommended). The second group was presented with a negative framing of the question (i.e. to select a probability that  $Se_{\text{grain}} < t_{\text{Se}}$  above which an intervention would be recommended).

More detail on the practical organization of the experiment is given in section 2.2. The threshold Se concentration,  $t_{\text{Se}}$ , in grain to which we referred is  $38 \mu\text{g kg}^{-1}$ , such that a serving of 330 g of grain flour provides a third of the daily EAR of Se for an adult woman. We used EAR because it is one of the commonly-used measure of intake when assessing nutritional status and planning intervention.

The respondents were presented with probabilities that Se concentration in grain falls below or above a threshold from specific locations on maps of Amhara, Ethiopia or Malawi dependent on the location of the particular session. These maps were derived by indicator kriging (see Webster and Oliver 2007) from data collected in the GeoNutrition project (Gashu *et al.* 2021). Indicator kriging was used because it requires no specific assumption that the kriging errors are normally distributed (Rivoirard 1994). More detail on this is provided by Chagumaira *et al.* (2021). Note that the grain samples in this project, in both Ethiopia and Malawi, were collected on a consistent sample support: a 0.1-ha circular plot in the centre of the sampled field. The probabilities therefore relate to mean values of grain concentration across such a support within a field at a specified location.

### 2.2. Organization of the experiment

The experiment was done in two sessions at Lilongwe, Malawi (November 2019) and Addis Ababa, Ethiopia (January 2020). Ethical approval to conduct this study was granted by the University of Nottingham School of Sociology and Social Policy Research Ethics Committees (BIO-1920-004 for Malawi, and BIO-1920-007 for Ethiopia), as approved by Lilongwe University of Agriculture and Natural Resources (LUANAR), and Addis Ababa University (AAU).



**Table 1.** Composition of different professional groups during the experiment in Ethiopia and Malawi.

Professional group	Location		Total
	Ethiopia	Malawi	
Agronomist	4	5	9
Soil scientist	12	13	25
Public health and nutrition specialist	12	5	17
Total	28	23	51

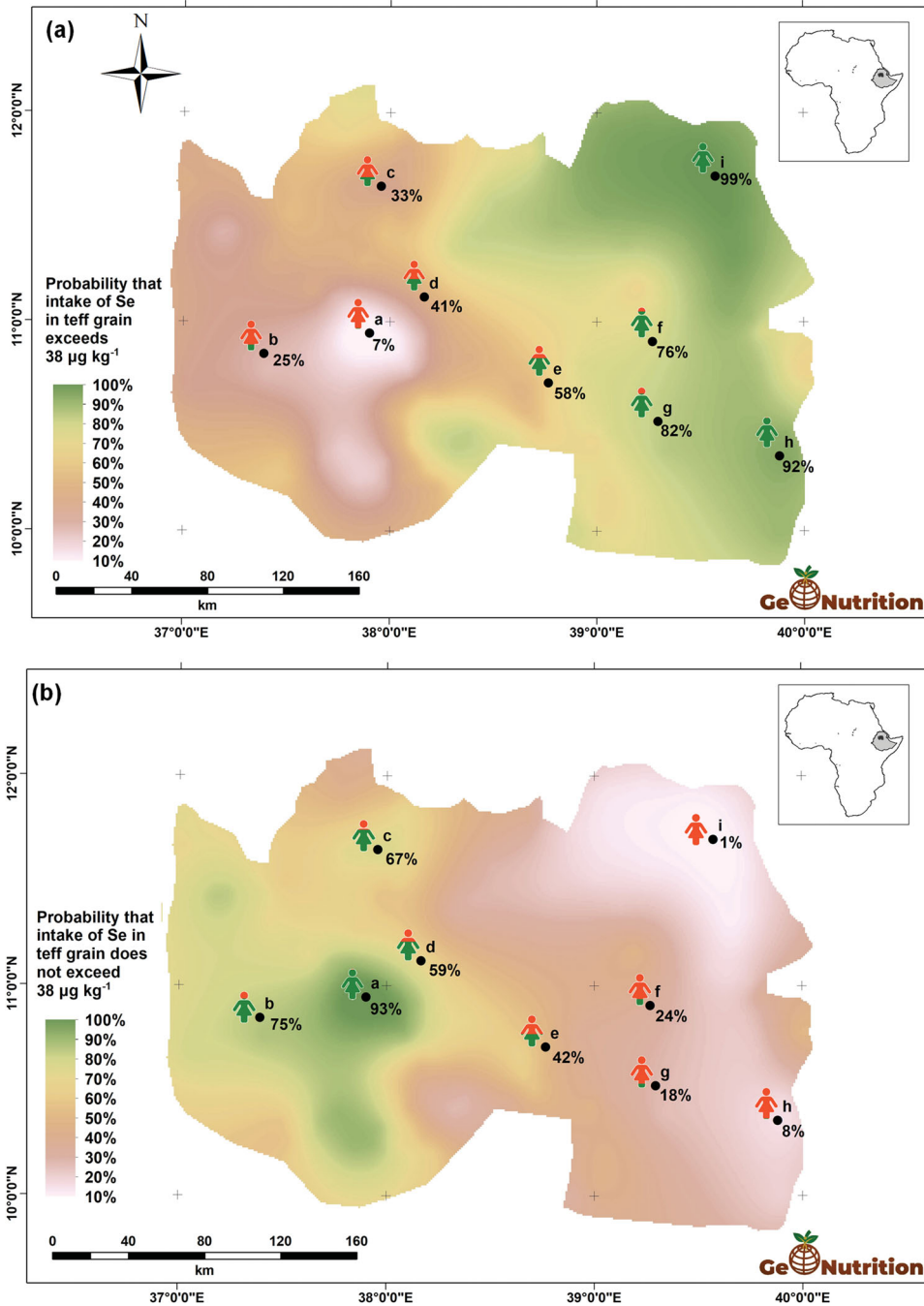
We invited participants from among professionals working in agriculture, nutrition and health, at NGOs, universities and government departments from Ethiopia, Malawi and in the wider GeoNutrition project. Recruitment was undertaken by the local GeoNutrition Project team. In total we had 51 participants, 34 were agronomists and soil scientists and 17 were public health and nutrition specialists, see [Table 1](#).

In each workshop, we started by randomly allocating participants to one of two groups one for positive framing and the other for negative. This was done by asking each participant to draw a shuffled card from a pot of cards bearing group labels. Cards were not replaced. We did not explain why we were grouping them until after the exercise had been completed.

We presented the first group with a map of probability that  $Se_{\text{grain}} > t_{Se}$ . The locations were identified on the map, and at each probability that  $Se_{\text{grain}} > t_{Se}$  was also illustrated by a pictograph (see [Figure 1\(a\)](#)). The questions were targeted to their areas of expertise. Specifically, agronomists and soil scientists were asked to decide whether or not they would recommend an intervention to provide and promote Se-fortified fertiliser. The public health and nutrition specialists to decide whether or not they would recommend a programme to provide Se-fortified food at that site. In both cases we asked the participants to assume that checks would be undertaken before the intervention took effect to ensure that no one was exposed to toxic levels of Se. The map showed nine locations, labelled **a, b, c, d, e, f, g, h** and **i**, at which probability that  $Se_{\text{grain}} > t_{Se}$  was 7%, 25%, 33%, 41%, 58%, 76%, 82%, 92%, 99%, respectively.

For each location in turn and by referring to the probability (as shown on map with pictograph, and explicitly stated in words), each participant recorded in a questionnaire whether or not they would recommend an intervention at the site given the probability. Using location **a** as an example, we phrased our question as follows: 'At site **a** there is 7% probability that the concentration of grain Se concentration exceeds the threshold, would you approve this intervention?' We chose a range of probabilities giving coverage of the interval [0,1] so as not to limit the responses participants could give.

When the first group had completed filling in the questionnaires we invited participants from the second group into the room. To this group we presented a map of probability that  $Se_{\text{grain}} < t_{Se}$ . At each location, probability that  $Se_{\text{grain}} < t_{Se}$  was also illustrated by a pictograph (see [Figure 1\(b\)](#)). The map showed the same nine locations but with 93%, 75%, 67%, 59%, 42%, 24%, 18%, 8%, 1% probability that  $Se_{\text{grain}} < t_{Se}$ . The participants answered the same questions as the first group, for the same location, but with a negative framing. For example, we asked them, 'At site **a** there is 93%



**Figure 1.** (a) Probability that concentration of Se in teff grain is greater than  $38 \mu\text{g kg}^{-1}$  ( $\text{Se}_{\text{grain}} > t_{\text{Se}}$ ) in Amhara region, Ethiopia. This was presented to the first group, with a positive framing of the question. The locations labelled **a, b, c, d, e, f, g, h,** and **i** at which probability that  $\text{Se}_{\text{grain}} > t_{\text{Se}}$  is also illustrated with a pictograph. (b) Probability that the concentration of Se in teff grain is less than  $38 \mu\text{g kg}^{-1}$  ( $\text{Se}_{\text{grain}} < t_{\text{Se}}$ ) in Amhara region, Ethiopia. This was presented to the second group of participants, with a negative framing of the question. The locations labelled **a, b, c, d, e, f, g, h,** and **i** at which probability that  $\text{Se}_{\text{grain}} < t_{\text{Se}}$  is also illustrated with a pictograph.

*probability that the concentration of grain Se concentration does not exceed the threshold, would you approve this intervention?*

Participants did this exercise independently, and were asked not discuss the questions with each other until they had completed the exercise. In the introduction to this exercise, it was pointed out to the participants that errors could go on both directions, resulting in an intervention where it was not needed (error of commission), or failing to intervene where the nutritional supply from staple foods was inadequate (error of omission). We encouraged participants to consider the sources of losses under errors of commission or omission. For example, the agronomists and soil scientists group should consider the costs of buying Se-enriched fertilisers especially given that Se does not improve crop yield. For public health and nutrition specialists, there would be costs associated with failing to intervene when there is need because of increased risk of health complications and mortality especially with people with compromised immunity due Se deficiency (e.g. thyroid dysfunction and suppressed immune response), but that unnecessary interventions are likely to represent a loss as resources are used which could address other public health initiatives. However, we did not ask the participants to attempt to calculate any of these costs. Rather, the aim was that having considered the possible outcomes, they should make a judgement in the light of their experience. This would be expected to reduce any framing effect (Almashat *et al.* 2008). When both groups had completed the exercise, we brought them together and we then explained the objectives of the exercise and background of the loss functions.

### **2.3. Model and analysis**

The following sections describe the statistical methodology used in this paper to analyse the data from the experiment. We summarize the methods briefly here for the benefit of readers for whom the mathematical content is of limited interest. We propose a statistical model for a set of responses to the questionnaires. Under the model any individual respondent is assumed to advocate intervention once the probability that grain Se concentration is less than  $38 \mu\text{g kg}^{-1}$  exceeds some value  $p'$ . We assume that the values of  $p'$  for a set of respondents can be treated as a random variable with a Beta distribution, a distribution particularly suited to modelling values which are constrained on an interval, and able to accommodate a wide range of behaviours. The two parameters of the Beta distribution can be estimated for a set of observations by a maximum likelihood method. Of interest is an estimate of the mean of the distribution, which we refer to as  $P_i$ , the expected value of,  $p'$  for an individual from the population of which the set of respondents is a sample. The maximum likelihood estimation allows us to evaluate evidence that, for example, it is necessary to model the responses from positive or negative framing with different parameter sets. This is done by means of the log-likelihood ratio test to compare a null model (in which responses with the two framings are pooled) with an alternative (in which distinct parameters are estimated for each framing). We used this approach to test the effect of framing, location (Ethiopia or Malawi), and professional group (agronomists and soil scientists or public health and nutrition specialists).

Having explored the data by modelling we decided that we wished to estimate the mean value  $P_t$  for all professional groups and locations pooled, for the responses to the negatively framed question. We did this by Bayesian estimation, using very uninformative prior distributions for the Beta parameters (that is, priors that have very little influence on the posterior distribution, which is dominated by the data).

### 2.3.1. Form of the data and their interpretation

Our data are a set of responses to questions, asking whether an intervention would be recommended in a situation given the probability that Se concentration in grain exceeds a nutritionally-significant threshold (positive framing) or is below the threshold (negative framing). The probabilities were expressed as percentages. Let the ordered set of percent probabilities (negatively framed) be  $\{P_1, P_2, \dots, P_m\}$ . The positively-framed question set was directly equivalent, referring to the same scenarios, and so the percent probabilities presented with the positively framed questions were  $\{100 - P_m, \dots, 100 - P_2, 100 - P_1\}$ .

For purposes of analysis the probabilities were scaled to  $[0, 1]$ , and the positively-framed probabilities were converted to the equivalent probability that  $Se_{\text{grain}} < t_{Se}$ . We denote these probabilities by  $\{p_1, p_2, \dots, p_m\}$ .

A response to the question is deemed to be consistent only if the respondent indicated that, for some  $i \in \{1, 2, \dots, m\}$ , an intervention should be considered for all scenarios where the probability that  $Se_{\text{grain}} < t_{Se}$  was greater than or equal to  $p_i$ , and that the intervention should not be considered otherwise. If a response was not consistent in this sense, then it was discarded. Our data therefore comprise a set of  $n$  index values,  $q_j$ , where  $q_j = i$  if the  $j^{\text{th}}$  respondent stated that interventions would be recommended in all cases where  $P(Se_{\text{grain}} < t_{Se}) \geq p_i$ . Of the 51 responses five were inconsistent (for example, the respondent recommended an intervention in a case where the probability of deficiency took some value, but did not recommend it in cases with both larger and smaller probabilities of deficiency). Three of the responses were anomalous, the respondent advocated an intervention for cases with a small probability of deficiency, and did not recommend intervention in cases with a large probability of deficiency. These 8 returns were discarded, leaving 43 for analysis, but they do illustrate the difficulties that stakeholders can have with the interpretation of probabilities.

We assume that each respondent has a latent 'personal' probability,  $p'$  such that, given all available information, they would advocate an intervention at a site where  $P(Se_{\text{grain}} < t_{Se}) \geq p'$ . Furthermore, we assume that, if the respondent indicates that an intervention should be recommended for all scenarios in the set for which the probability equals or exceeds  $p_i$ , then the lower and upper bounds on  $p'$  are given by

$$l_i = \begin{cases} \frac{p_i + p_{i-1}}{2} & i \neq 1, \\ 0 & i = 1, \end{cases} \quad (6)$$

and

$$u_i = \begin{cases} \frac{p_i + p_{i+1}}{2} & i \neq m, \\ 1 & i = m. \end{cases} \quad (7)$$

### 2.3.2. The statistical model and its estimation

We assume that the distribution of  $p'$  within any group of respondents has a Beta distribution, such that the probability density function for some value  $x \in [0, 1]$  is given by

$$f_{\beta}(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

and  $\Gamma(\cdot)$  denotes the gamma function. The Beta distribution is particularly appropriate for modelling probabilities as random variables, because a Beta random variable is continuous but constrained to a fixed interval (here  $[0,1]$ ), and it is very flexible, accommodating a wide range of behaviours: bell-shaped, symmetrical with large or small kurtosis, uniform, strongly positively or negatively skew, straight-line or U-shaped (Tjims 2018).

The parameters of the gamma distribution are  $\alpha$  and  $\beta$  but a convenient reparameterization (because of the correlation of these parameters) is to the mean  $U$  and a precision parameter  $V$  which is smaller the more dispersed the distribution of  $x$ ;

$$U = \frac{\alpha}{\alpha + \beta}, \quad (9)$$

and

$$V = \alpha + \beta. \quad (10)$$

We denote the probability density function for some set of parameters  $\theta = \{U, V\}$  by  $f_{\beta}(x|\theta)$  (McDonald and Xu 1995).

If the value of  $p'$  for the  $j$ th respondent can be regarded as a Beta random variable with probability density function (PDF)  $f_{\beta}(x|\theta_k)$  then the probability of observing  $q[j] = i$  can be obtained as the integral of the Beta PDF over the limits  $l_i$  and  $u_i$ :

$$\text{Prob}\{q[j] = i\} = \int_{l_i}^{u_i} f_{\beta}(x|\theta_k) dx. \quad (11)$$

If we treat all our respondents as members of a single population of interest, then the log-likelihood for a proposed set of parameters  $\theta$  for that population can be obtained by computing, for each entry in  $q$  the probability for the observed value of  $i$  by evaluating Equation (11). The sum of the logarithms of these probabilities gives the log likelihood. A maximum likelihood estimate of  $\theta$  can be found numerically, as described below.

For our purposes we want to estimate models for our observations which assumes that there are different sub-populations from which they are drawn, and that different values of the Beta parameters may be estimated for such a sub-population. For example, we might choose to fit a model in which we assume that all responses from individuals who were presented with information with positive framing are drawn from a sub-population with a set of Beta parameters, and that those responses where

the framing was negative constitute a second sub-population. The likelihood, as described above, must be extended to this more complex model.

Consider a set of responses from a group of  $n$  subjects. The subjects can each be assigned to one of  $Q$  sub-populations, and our hypothesis is that a particular set of values of the parameters  $\theta_k = \{U_k, V_k\}$  can be proposed for the  $k^{\text{th}}$  sub-population. We denote the full set of  $Q$  parameters by  $\Theta = [\theta_1^T, \theta_2^T, \dots, \theta_Q^T]$ .

Given the assumptions set out in Equations (6) and (7) above, the log-likelihood for proposed values of the parameters  $\Theta$ , given a set of  $n$  responses can be obtained as

$$\uparrow(\varrho; \Theta) = \sum_{k=1}^Q \sum_{j=1}^n \sum_{i=1}^m I_{k,j,i} \log \int_{l_i}^{u_i} f_{\beta}(x|\theta_k) dx, \quad (12)$$

where  $I_{k,j,i}$  is an indicator variable which takes the value 1 if  $\varrho[j] = i$  and the  $j^{\text{th}}$  respondent belongs to the  $k^{\text{th}}$  sub-population of respondents. In all other cases  $I_{k,j,i} = 0$ . This indicator variable allows us to simplify the notation. The three nested summations implies that we compute the log of the probability for every sub-population parameter set over every set of bounds for each observation, but the indicator takes the value zero for any combination where the  $j^{\text{th}}$  respondent is not in the  $k^{\text{th}}$  sub-population, and  $\varrho[j] \neq i$ . Equation (12) therefore allows us to compute the log likelihood for a proposed set of Beta parameters,  $\Theta$  for a corresponding model of a set of responses.

In this study we found maximum likelihood estimates of the parameters  $\hat{\theta}_k, k \in \{1, 2, \dots, P\}$  which minimized  $-\ell(\varrho; \Theta)$  given the data in  $\varrho$ . This was done using the *optim* function in base R (R Core Team 2020), using the default optimizer which is the simplex algorithm of Nelder and Mead (1965).

A series of nested models were fitted to the data. In the first, model  $M_0$ , all respondents were considered as a single population. In the second, model  $M_1$ , respondents who were presented with a negative framing were treated as a distinct sub-population from respondents presented with a positive framing. These two models were compared by computing the log-likelihood ratio statistic:

$$L = 2(\ell_{M_1} - \ell_{M_0}), \quad (13)$$

where  $\ell_{M_1}$  and  $\ell_{M_0}$  denote the maximized log-likelihood for models  $M_1$  and  $M_0$  respectively. Under a null-hypothesis where the parameters for the two sub-populations can be regarded as equal (as in  $M_0$ , termed the 'null model')  $L$  is asymptotically distributed as  $\chi^2(2)$ , the degrees of freedom being equal to the number of additional parameters in  $M_1$  relative to  $M_0$ .

Further models were considered in which sub-populations were defined by (i) the location of the experiment and (ii) the broad professional group, both tested with the groups with positive and negative framing. The first of these was considered in case there were some differences in the way the meetings in two locations were conducted. Differences could also be due to composition of the participants group (see Table 1), we had fewer public health and nutrition specialists in the Malawi meeting. For familiarity and engagement, we used a probability map from Ethiopia's Amhara region in the experiment in Ethiopia and a map of Malawi in the experiment in Malawi. The comparison between the groups (agronomists and soil scientists or public health and nutrition specialists), was considered to test the hypothesis that cultural

differences between the two professional groups contribute to differences in sensitivity to the framing effect, and in the relative weighting of the cost of errors of commission and omission.

### 2.3.3 Bayesian estimation

After examining the alternative models described in the previous section, it was decided to make a final estimate of the mean value of  $U$  for all respondents (both locations and professional groups) within the sub-sets presented with negative framing. A Bayesian approach was taken for this final step so as to quantify uncertainty in the parameter estimates without the assumptions of linearity required in methods based on the information matrix or the assumption that estimation errors are normal (Spiegelhalter and Rice 2009).

The Bayesian approach requires prior distributions for the parameters  $U$  and  $V$ . A uniform prior distribution over  $(0, 1)$  was assumed for  $U$ . This is entirely uninformative about the parameter. The prior for  $V$  was a gamma distribution with parameters  $(1, 20)$ . This is a weakly informative prior, so the posterior distribution is dominated by the data.

The prior predictive density for the data was obtained by integrating out the parameters, this was done with the *adaptIntegrate* function from the cubature library for the R platform (Narasimhan *et al.* 2020). The posterior joint density of  $U$  and  $V$  is then straightforward to evaluate. The posterior density of  $U$  was then evaluated at a fine set of locations by integrating out  $V$  with the *integrate* function of base R. The highest posterior density credible interval for  $U$  (95%) was then evaluated by applying the *hdi* function from the *HDInterval* library for R (Meredith and Kruschke 2018) to the set of density values. Finally the mean of  $U$  was obtained by integration over its posterior density.

## 3. Results

We had similar numbers of attendees whose professional background was agronomy and soil science in both workshops (see Table 1). However, we had more professionals who were public health and nutrition specialists in the Ethiopian experiment.

**Table 2.** Fitted models for respondent data and maximized log-likelihood.

	Model	Number of parameters	$\ell$
$M_0$	All respondents pooled	2	-81.58
$M_1$	Respondents separated by framing	4	-73.22
$M_2$	Respondents separated by framing within location	8	-69.15
$M_3$	Respondents separated by framing within professional group	8	-67.35

**Table 3.** Log-likelihood ratio tests to compare models.

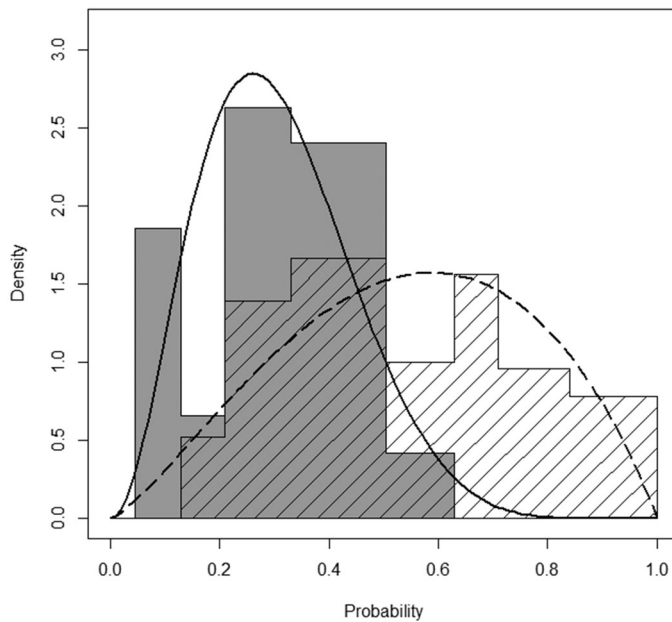
Null model	Model	$L$	Degrees of freedom	$p$
$M_0$	$M_1$	16.71	2	0.0002
$M_1$	$M_2$	8.13	4	0.087
$M_1$	$M_3$	11.74	4	0.019

### 3.1. Nested model analysis

Table 2 shows the fitted models for the combined respondent data and their maximised log-likelihood.

Table 3 shows log-likelihood tests to compare the models. There is strong evidence to reject the model with all respondents pooled ( $M_0$ ) and to accept an overall difference between the groups with different framing ( $M_1$ ) ( $p = 0.0002$ ). However, there is no strong evidence to reject  $M_1$  by comparison to the more complex model with locations ( $M_2$ ) ( $p = 0.087$ ).

When comparing a more complex model with professional group ( $M_3$ ) with model with respondents separated only by framing, there is some evidence ( $p = 0.019$ ) to reject  $M_1$ . Therefore, further analysis of the respondent data was based on  $M_1$  and  $M_3$ .

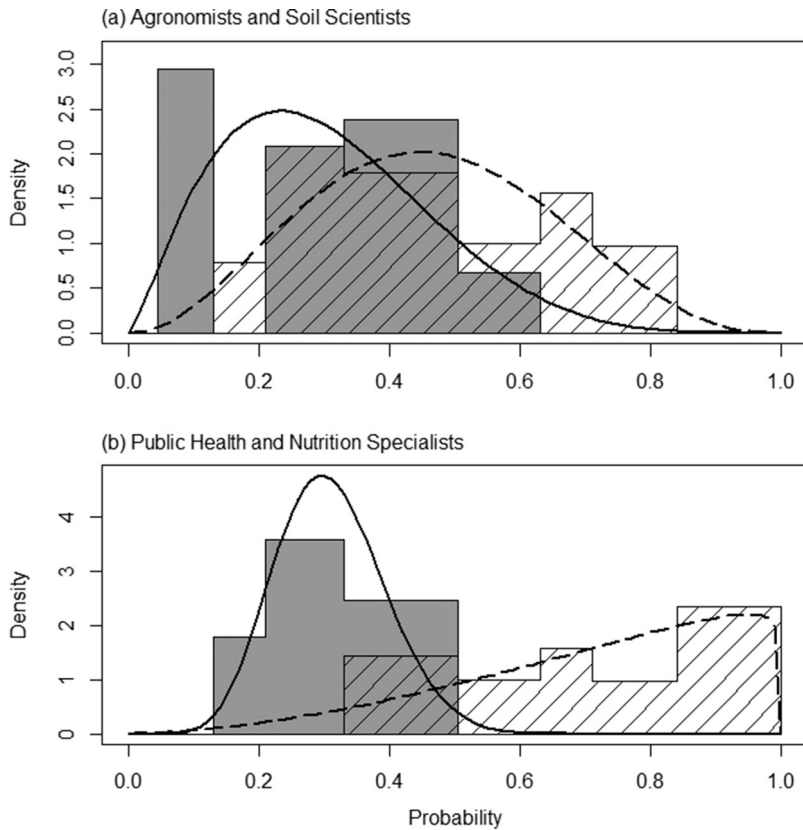


**Figure 2.** Fitted beta distributions for model  $M_1$  (negative or positive framing, professional groups and locations pooled) superimposed on histograms for the results. The solid line and dark grey histogram corresponds to the respondents with negative framing. The broken line and hatched histogram are for respondents with positive framing.

**Table 4.** Maximum likelihood estimates of parameters  $U$  and  $V$  for models  $M_1$  and  $M_3$ .

Model	Sub-group	Parameters	
		$U$	$V$
$M_1$	All respondents with negative framing	0.307	10.55
$M_1$	All respondents with positive framing	0.547	4.26
$M_3$	All public health and nutrition specialists with negative framing	0.310	30.19
$M_3$	All public health and nutrition specialists with positive framing	0.712	3.80
$M_3$	All agronomists and soil scientists with negative framing	0.303	7.69
$M_3$	All agronomists and soil scientists with positive framing	0.462	6.81





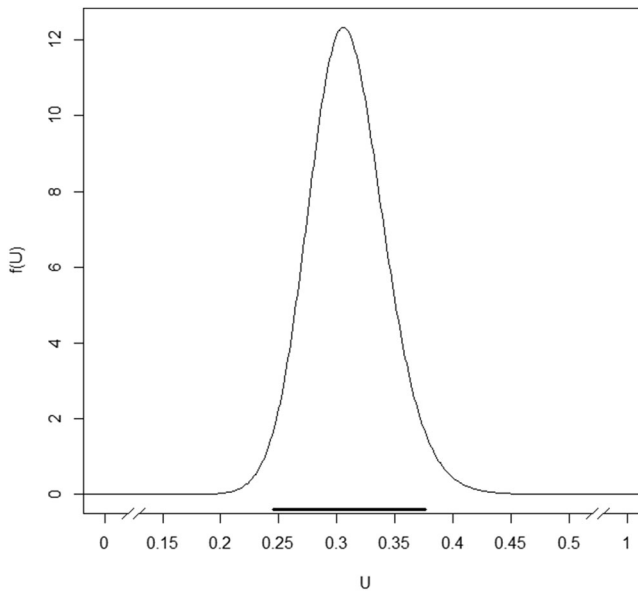
**Figure 3.** Fitted beta distributions for model  $M_3$  (negative or positive framing, locations pooled) superimposed on histograms for the results for professionals from (a) agronomists and soil scientists and (b) public health and nutrition specialists. The solid line and dark grey histogram corresponds to the respondents with negative framing. The broken line and hachured histogram are for respondents with positive framing.

### 3.2. Model fitting

Figure 2, shows the beta probability densities for positive and negative framing under model  $M_1$ . The histograms show empirical densities from the responses over the probability ranges in each group. The solid line and dark grey histogram corresponds to the respondents with negative framing. The broken line and hachured histogram are respondents with positive framing. The figure also shows that negative framing results in a decision to intervene at a smaller probability that the threshold is not exceeded than does the positive framing.

Table 4, shows the estimated parameters for  $M_1$ . Figure 3 shows fitted beta distributions for model  $M_3$ . Here again, decisions to intervene are at a smaller probability for the respondents with negative framing in both professional groups, although the difference is most marked for the public health and nutrition specialists.

Table 4, shows the estimated parameters for  $M_3$ . The mean values for  $U$  are very similar in both professional groups with negative framing. The estimates of  $U$  under positive framing in the public health and nutrition specialists group is close to the



**Figure 4.** Posterior density for  $U$  and (solid bar) the highest posterior density credible interval (95%) estimated from pooled data for all respondents with negative framing.

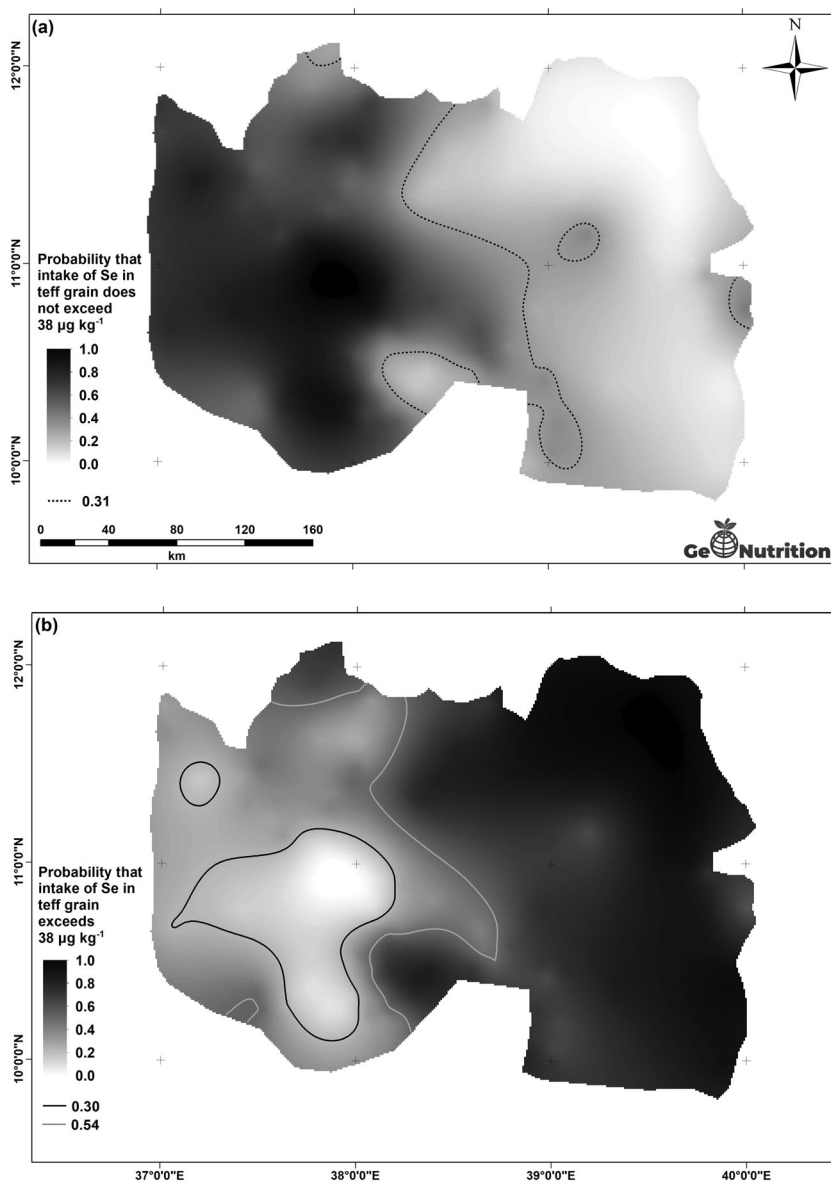
complement of this value under negative framing, and the dispersion is large. It is possible that this reflects some misunderstandings of the probabilities with this group. On this basis we pooled the negatively framed responses for further analysis.

The mean of  $U$  from the posterior distribution for the pooled (over professional group) responses to the negatively-framed question was 0.31 (similar to the ML estimate). The posterior density is shown in Figure 4.

Close to symmetrical, the highest-posterior density credible interval for  $U$ , is [0.25–0.38], so comfortably below 0.5. For positive framing, further analysis was based on the separate professional groups. The mean of  $U$  from the posterior distribution for the public health and nutrition specialists group to the positively framed question was 0.70 (very close to the ML estimate 0.71) with a highest-posterior density credible interval for  $U$ , is [0.55–0.85]. Whilst for the agronomists and soil scientists group it was 0.46 (similar to the ML estimate) with a highest-posterior density credible interval for  $U$ , is [0.37–0.55].

Figure 5(a) shows a map of the probability that the concentration of Se in teff grain less than the threshold,  $38 \mu\text{g kg}^{-1}$  in Amhara region, Ethiopia. The dashed line is the probability isoline or contour at which the probability is equal to the estimated mean value of  $P_t$  for the pooled (over professional group) responses to the negatively- framed question. If this value is used as a guide to decisions, then interventions would be recommended where probabilities mapped on this figure exceed the specified isoline. In these circumstances intervention would be recommended over 50% of the mapped area ( $34,672 \text{ km}^2$ ).

Figure 5(b) shows the same probabilities as 5a, but this time with two probability isolines, one (black) is the estimated mean value of  $P_t$  for the response of the public health and nutrition specialists group to the positively framed to the positively-framed



**Figure 5.** (a) Probability that the concentration of Se in teff grain less than  $38 \mu\text{g kg}^{-1}$  in Amhara region, Ethiopia. The dashed probability isoline is the mean probability value,  $P_t$ , at which a stakeholder would judge that an intervention should be made. This is the probability at which either professional group would recommend an intervention in Amhara region, Ethiopia which the question was framed negatively. (b) Probability that the concentration of Se in teff grain exceeds  $38 \mu\text{g kg}^{-1}$  in Amhara region, Ethiopia. The grey probability isoline is the mean probability value,  $P_t$ , at which agronomists and soil scientists would judge that an intervention should be made which the question was framed positively. The black probability isoline is the mean probability value,  $P_t$ , at which public health and nutrition specialists would judge that an intervention should be made which the question was framed positively.

question, this encloses an area where an intervention would be recommended corresponding to proportion, 12% of the mapped area (7,792 km<sup>2</sup>). The second isoline (grey) is the estimated mean value of  $P_t$  for the response of the agronomist and soil scientist group to the same question. Decisions based on this value of  $P_t$  would see interventions over proportion, 40% of the mapped area (26,596 km<sup>2</sup>).

## 4. Discussion

### 4.1. Our findings

Our results have shown (Figure 4) that a reasonably precise estimate of the mean probability value,  $P_t$ , at which a stakeholder would judge that an intervention should be made, can be elicited from a stakeholder group. The estimated mean value of  $P_t$  from a group of stakeholders in Malawi and Ethiopia, 0.31, is shown visually as a contour on the map of probabilities for Amhara region in Ethiopia (Figure 5(a)). This is the estimated mean probability at which either professional group would recommend an intervention in Amhara, Ethiopia and Malawi, if the question were framed negatively (i.e. in terms of deficiency). This  $P_t$  should not be interpreted as an objective optimal threshold value for the decision. Rather, it reflects the judgement of some group of stakeholders and their tacit assessment of losses and costs associated with making a choice with uncertain information. The methodology provided here to elicit this quantity from a stakeholder group allows us to identify a threshold  $P_t$  to use so as to present uncertain information with an interpretation which reflects the assumptions and decision-making of a particular stakeholder group. The elicitation method may also help to make that tacit process of judgement more explicit.

We also examined whether the elicited  $P_t$  depended on the specific interests of the group, and whether it is prone to framing effects (i.e. how the question is posed). With or without the effects of professional group (both  $M_1$  and  $M_3$ ), our results show that the negative framing resulted in a decision to intervene at a much smaller probability than positive framing. We also observed similar estimates of  $U$  for both professional groups within the negative framing. With the public health and nutrition specialists group positive framing resulted in a much larger threshold probability of deficiency for intervention than was the case with the agronomists and soil scientists group.

Framing effects are well known in the psychology of decision-making. Decisions are influenced by irrelevant aspects of the way information is presented, even though the same information is presented with different framings (Tversky and Kahneman 1981). In this example, a negative framing of the question draws the participant's attention to deficiency, rather than to sufficiency, and hence to a more conservative decision. We see such an effect despite preparatory activities in the experiment to draw the attention of participants to the possibility, and the implications, of interpretative errors in both directions, as suggested by Almashat *et al.* (2008). The greater consistency of responses across professional groups with negative framing may indicate that stakeholders find this easier to interpret. This maybe because stakeholders are accustomed to think about the specific problem in terms of nutrient deficiency. This shows the importance of framing spatial information, and statements of its uncertainty, in terms with which the user of the information is familiar.

We noted above that our samples and predictions, with associated probabilities were on a consistent, fixed support. A change of support (e.g. to predict a mean value across a ward or other small region, or a cell in raster GIS) will reduce the local uncertainty of the prediction. It would be interesting to see whether awareness that a probability refers to a mean across a local administrative unit, rather than a small bulk sample from within a field (which is particularly relevant to the nutrient supply to subsistence farmers) changes stakeholder's interpretation, and whether any such effect interacts with framing.

#### **4.2. Generalizability, and topics for further work**

The probability threshold which we estimated here is for a very specific problem, micronutrient concentration in staple crops, and is unlikely to serve as a general one for interpretation of spatial information. We would expect the threshold probability to differ between settings depending on the particular stakeholder perspective on the costs entailed if an intervention is not recommended where it should be, or is implemented unnecessarily. The approach which we have used could be applied to different groups and different problems and settings where decisions are based on uncertain information.

The framing effect which we have seen has been identified in other studies on decision-making under uncertainty (e.g. Chen *et al.* 2014), and so is likely to apply in other cases where probabilities are used to indicate whether the state of affairs at a location requires an intervention. In our case negative framing led to a more conservative outcome because the stakeholders are directed to think in terms of nutrient deficiency. This cannot be generalised for different problems and settings. For example, in the case of assessing concentrations of a potentially harmful element in soil against soil guideline values, a positive framing (probability that the threshold is exceeded) might be expected to result in more conservative decisions.

It would be interesting to see whether the interaction of professional group and framing holds more generally for other problems (e.g. the interpretation of information on environmental contaminants). In particular our finding in this instance, that the interpretation of probabilities was more consistent between professional groups under the framing which led to more conservative decisions, would be of practical significance if it is found to hold consistently.

Probabilities are not straightforward to interpret. As noted above, our experimental procedure included presentations to participants about uncertainty and its implications for decision making prior to their completing the exercise. However, it would have been possible to spend more time in 'priming' participants before the exercise. This could be achieved by discussion of probability problems from everyday life, like weather forecasts, when decisions are made. This might reduce the framing effect, as well as the rate of rejection due to inconsistent or anomalous interpretations. However, the responses based on minimal priming are perhaps of more practical interest, because they may better represent how a stakeholder approaches probabilistic information in the course of their ordinary working life. The fact that eight returns received from our experiment had to be discarded because they were inconsistent or

anomalous underlines the difficulties that stakeholders with professional expertise in their own fields may have with the interpretation of probability. This has already been recognized (e.g. Spiegelhalter *et al.* 2011), although paradoxically, Jenkins *et al.* (2019) found that stakeholders seem to attach greater authority to numerical statements of probability than to calibrated phrases.

Some professional groups may have been able to handle and interpret probabilities better than others because of the content of education and training programmes which they typically complete. Further work to assess this, with a more varied range of professional groups, would be interesting, and might help to show how professional skills in the interpretation of uncertain spatial information could be best developed, either in higher education curricula or in particular professional training.

When decisions are made, stakeholders weigh up the pros and cons for the decision they make. We suggest that this process might be better-emulated in an experiment such as ours if more time could be spent in engagement with stakeholder groups to co-create scenarios for decision-making, and outcomes which are possible given the uncertainty in the spatial information which is used and the stakeholders' professional experience.

### **4.3. Implications for practice in GIScience**

The mean value of  $P_t$  obtained in this experiment will be used for practical purposes to aid interpretation of maps of nutrient supply from staple crops produced in the GeoNutrition project. We shall add a contour line to probability maps (for negative framing), as in Figure 5(a), annotating the legend to indicate that the mean threshold value applied by our stakeholder group means that interventions would be recommended where the probability takes larger values. The value can also be used as a starting point for discussion with other stakeholder groups, at national and local level, about the implications of the spatial information provided by the project.

In GIScience, it is common to validate prediction distributions by assessing the coverage of prediction intervals for validation data at different probabilities. Lark *et al.* (2019) provide an example from the study of soil nutrients. The coverage of the prediction intervals may be consistent with their probability over some ranges of values but not others. One value of this study for practical purposes in the GeoNutrition project is that we shall be able to focus our assessments of methods for spatial mapping on the validity of prediction intervals for probabilities close to  $P_t$ .

If decisions are based on uncertain information, presented in terms of the probability that a variable exceeds or falls below a threshold, then, other factors being equal, the decision process is equivalent to selecting a value of  $P_t$ . We suggest that this be done through a transparent process in which the underlying questions are examined by relevant stakeholders. Our experimental procedure, supplemented by standardized processes to co-create scenarios and to set the scene on uncertainties, could provide the basis for a formal elicitation methodology to achieve this. There is increasing interest in the use of elicitation methods to formalize the decision processes and conceptual models which individuals and communities of stakeholders may hold and use, at least tacitly, when forming expert judgements. Methods for expert elicitation have

been applied to problems in medical diagnosis, the interpretation of data on natural hazards and engineering design (e.g. O'Hagan *et al.* 2006).

The development of an elicitation procedure should take account of our findings with respect to framing effects, differences between professional groups and the interaction of professional group with framing. In our particular study there was greater consistency between the two professional groups with negative framing, and a more conservative outcome. These would be reasons for using negative framing when eliciting  $P_t$  for this particular problem, but as we note above further work is needed to see how far this finding can be generalized. At the very least it is important to ensure that framing is done consistently (i.e. we do not use mix positive and negative framing for the same problem) and that framing is coherent with standard terminology in the relevant stakeholder community, e.g. whether nutrient supply is generally described in terms of deficiency (deficient or not) or sufficiency (sufficient or not).

In the theoretical framework for this study we noted that a threshold probability,  $P_t$ , can be expressed in terms of the relative losses of contrasting decisions relative to those made with perfect information. We also noted that these losses, in general, are not accessible as they may be complex and have multiple components including actual costs (e.g. money required for interventions, the economic value of disability-adjusted life years saved or not saved) but also losses which are less tangible, and which may not be directly commensurable, (the value of public health, political and reputational losses). It is possible that the elicitation of a value of  $P_t$  could help to make public or community discussions of these losses more explicit. For example, if a stakeholder group decides that interventions to address micronutrient deficiency be recommended if probability of deficiency is  $\geq 0.1$  then it could be pointed out that this implies that the losses arising from a failure to intervene where intervention is required are nine times larger than the losses arising from an unnecessary intervention. Stakeholders might then reflect on whether this undervalues the opportunities to apply resources to other better-focussed interventions. This discussion could be built into a group elicitation process on the lines of the behavioural elicitation methods proposed by Reagan-Cirincione (1994) under which, after initial modelling of values returned by individuals, a group works together to arrive at a consensus.

We note one further development of our approach, which could be of practical relevance. In our conceptual framework we assume discrete states: an intervention happens or does not in response to whether or not a spatial variable exceeds a threshold. In practice spatial information might be used to set a continuous value at which some intervention is applied (e.g. a rate of fortification of a foodstuff, or a rate for a fertilizer or other agronomic input). In such a case, rather than discrete losses, there may be a continuous loss function of the error of the prediction, which is zero at zero error and increases with both under- and over-estimation of the target variable. If we assume that the loss function is piece-wise linear with error in the target variable, and that  $\alpha_1$  is the loss per unit of error of overestimation and  $\alpha_2$  is the loss per unit of error of underestimation, then the expected loss is minimized at a location with some particular prediction distribution for the target variable if we use as our estimate of the target variable the value  $\tilde{X}$

$$\tilde{X} = F^{-1}(P_o), \quad (14)$$

where  $F^{-1}(p)$  denotes the quantile of the prediction distribution corresponding to probability  $P_o$  and

$$P_o = \frac{\alpha_2}{\alpha_1 + \alpha_2}, \quad (15)$$

(Journal 1984). The formal similarity with our conceptual model for  $P_t$  in the case of discrete decisions (intervene or not) is apparent. Lark and Knights (2015) showed how the continuous loss-function model could be used to compute an implicit loss function, the loss function implied by a particular level of effort to obtain spatial information, and suggested that this could be used to support decision making about sampling effort. However, it requires a value for the ratio of  $\alpha_1$  and  $\alpha_2$ . One approach to obtaining this would be to provide stakeholders with scenarios in which the predicted value of the target variable is at the threshold for intervention, and to elicit a value of  $P_t$  which, under negative framing, could be regarded as an approximation to  $P_o$  in Equation (15) above.

Visualization of spatial uncertainty is important in GIScience. It is important to use appropriate colour scales to visualize spatial information, including uncertainty (Kunz *et al.* 2011, Kinkeldey *et al.* 2014). Uneven colour scales, such as rainbows, can distract from the information content of the image, and even generate artefacts (Cramer *et al.* 2020). Probabilities are ordered, continuous quantities, and we have no particular interest in values relative to a centric value (as we might for a variable on a scale from  $-1$  to  $+1$ ). For this reason, following Cramer *et al.* (2020), we decided that a sequential colour scale was appropriate. Because we wish to have good discrimination across the range of probabilities, a two-hue sequential scale is preferred. We therefore selected the ‘terrain’ HCL (hue-chroma-luminance) colour scale (Zeileis *et al.* 2020) to present probabilities to participants.

## 5. Conclusions

Much effort in GIScience and spatial statistics has focused on how to obtain prediction distributions, and probabilities from these (disjunctive kriging, indicator kriging, Bayesian methods), but it is clear (e.g. Chagumaira *et al.* 2021) that the task of communicating the uncertainty in spatial information is not complete when that is achieved, at least if the objective is that a general range of stakeholders should be able to use the information. This paper is a step towards that development. In our study we have shown we can go beyond just computing probabilities, and consider how uncertainty can be communicated to a diverse group of end-users for decision making for interventions. We also have shown that a reasonably precise estimate of the mean probability value at which a stakeholder would judge that an intervention should be made, can be elicited from a stakeholder group with particular expertise and interests.

There were more consistent estimates of the mean probability value under negative framing. This might not apply generally, whether it should be a matter for further research. Note that ‘negative’ framing relative to a threshold in this setting gives rise to a conservative response, but that in other contexts (e.g. if the threshold is a pollutant), the positive framing might be expected to do so. Hence the framing effect can



be pronounced in the interpretation of probabilistic representation of uncertainty presented as maps, and that this effect interacts with professional group.

## Acknowledgements

The authors gratefully acknowledge the contributions made to this research by the participating farmers and field sampling teams. In Ethiopia, field sampling teams were from the Amhara National Regional Bureau of Agriculture. In Malawi, field sampling teams were from the Department of Agricultural Research Services, and Lilongwe University of Agriculture and Natural Resources.

## Data and codes availability statement

The data and code that support the research are available at <https://doi.org/10.6084/m9.figshare.14339987.v4>

## Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

This work was supported by GeoNutrition projects funded by the Bill & Melinda Gates Foundation (BMGF) [INV-009129], and the Nottingham-Rothamsted Future Food Beacon Studentships in International Agricultural Development. The funders were not involved in the study design, the collection, management, analysis, and interpretation of data, the writing of the report or the decision to submit the report for publication.

## Notes on contributors

*Christopher Chagumaira* is a PhD candidate at the Future Food Beacon of Excellence International Agriculture Doctoral Training Programme at the University of Nottingham, Rothamsted Research and Lilongwe University of Agriculture and Natural Resources. His research focuses on spatial statistics, understanding and communicating uncertainty in spatial information of micronutrients to design site-specific interventions to address 'hidden hunger' in sub-Saharan Africa regions.

*Patson Nalivata* is an Associate Professor of Soil Science and head of Crop and Soil Science Department at Lilongwe University of Agriculture and Natural Resources, Malawi. His research focuses on understanding biogeochemical dynamics of selenium and iodine in tropical soils; doubled-up legume technologies; optimization of fertiliser recommendations; biochar influences on soil fertility; effects of water stress on root architecture among Malawi maize germplasm; composting systems and processes in Malawi and the quality of human compost (ecosan toilets).

*Joseph Chimungu* is a Senior Lecturer in Soil Science at Lilongwe University of Agriculture and Natural Resources, Malawi. His research interest is on understanding the genetic, physiological, and ecological basis of plant adaptation to drought and low soil fertility.

**Dawd Gashu** is an Associate Professor of Food Science and Nutrition at Addis Ababa University. His research focuses on micronutrient nutrition and its link to human health and finding feasible ways of addressing micronutrient deficiency.

**Martin Broadley** is Professor of Plant Nutrition at the University of Nottingham. His research seeks to increase our understanding of the movement of micronutrients and trace elements in food systems. Potential outcomes of the research include improving the nutritional quality of soils and crops for human and livestock diets.

**Alice Milne** is a mathematical modeller in the Sustainable Agricultural Sciences department at Rothamsted Research. Her research focuses on the mathematical analysis and modelling of agro-ecological systems using various statistical and geostatistical techniques. She also has keen interest in quantifying and communicating uncertainty in model-based predictions.

**Murray Lark** is Professor of Environmetrics at the University of Nottingham. His research is concerned with spatial statistics, sampling and experimental design, and capacity strengthening in agricultural and environmental research. Most of his current work addresses issues of sustainability and human health in relation to food and agricultural systems in Africa.

## ORCID

Christopher Chagumaira  <http://orcid.org/0000-0003-0590-0350>

Patson C. Nalivata  <http://orcid.org/0000-0001-8429-2829>

Joseph G. Chimungu  <http://orcid.org/0000-0002-0330-0872>

Dawd Gashu  <http://orcid.org/0000-0002-6067-4432>

Martin R. Broadley  <http://orcid.org/0000-0003-3964-7226>

Alice E. Milne  <http://orcid.org/0000-0002-4509-0578>

R. Murray Lark  <http://orcid.org/0000-0003-2571-8521>

## References

- Almashat, S., et al., 2008. Framing effect debiasing in medical decision making. *Patient Education and Counseling*, 71 (1), 102–107.
- Budescu, D.V., Broomell, S., and Por, H.H., 2009. Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, 20 (3), 299–308.
- Chagumaira, C., et al., 2021. Communicating uncertainties in spatial predictions of grain micronutrient concentration. *Geoscience Communication*, 4 (2), 245–265.
- Chen, S.Y., Ross, B.H., and Murphy, G.L., 2014. Decision making under uncertain categorization. *Frontiers in Psychology*, 5 (991), 991–913.
- Chilimba, A.D.C., et al., 2013. *Agricultural lime application for improved soil and crop production in Malawi*. Malawi: Soil Health Consortium of Malawi (SOHCOM).
- Cole, S., and Jeffries, J., 2009. *Using soil guideline values*. Bristol, United Kingdom: Environmental Agency.
- Cramer, F., Shephard, G.E., and Heron, P.J., 2020. The misuse of colour in science communication. *Nature Communications*, 11 (1), 7.
- DEFRA 2010., *Fertiliser Manual (RB209)*. Norwich, United Kingdom. Available from: <https://ahdb.org.uk/rb209>.
- Gashu, D., et al., 2020. Spatial prediction of the concentration of selenium (Se) in grain across part of Amhara Region, Ethiopia. *The Science of the Total Environment*, 733, 139231.
- Gashu, D., et al., 2021. The nutritional quality of cereals varies geospatially in Ethiopia and Malawi. *Nature*, 594 (7861), 71–76.

- Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. New York, Oxford: Oxford University Press.
- Goovaerts, P., 2001. Geostatistical modelling of uncertainty in soil science. *Geoderma*, 103 (1–2), 3–26.
- Goovaerts, P., Webster, R., and Dubois, J.P., 1997. Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics. *Environmental and Ecological Statistics*, 4 (1), 49–48.
- Greiner, L., et al., 2018. Uncertainty indication in soil function maps – transparent and easy-to-use information to support sustainable use of soil resources. *SOIL*, 4 (2), 123–139.
- Hatvani, I., et al., 2021. Geostatistical evaluation of the design of the precipitation stable isotope monitoring network for Slovenia and Hungary. *Environment International*, 146, 106263.
- Heuvelink, G.B.M., 2018. Uncertainty and uncertainty propagation in soil mapping and modeling. In: A.B. McBratney, B. Minasny and U. Stockmann, eds. *Pedometrics*. Cham: Springer International Publishing, 439–461.
- Heuvelink, G.B.M., and Burrough, P.A., 2002. Developments in statistical approaches to spatial uncertainty and its propagation. *International Journal of Geographical Information Science*, 16 (2), 111–113.
- Jenkins, S.C., Harris, A.J.L., and Lark, R.M., 2019. When unlikely outcomes occur: the role of communication format in maintaining communicator credibility. *Journal of Risk Research*, 22 (5), 537–554.
- Journel, A.G., 1984. mad and conditional quantile estimators. In: G. Verly, M. David, A.G. Journel and A. Marechal, eds. *Geostatistics for natural resource characterisation part 1*. Dordrecht. Netherlands: Springer Netherlands, 261–270.
- Joy, E.J.M., et al., 2014. Dietary mineral supplies in Africa. *Physiologia Plantarum*, 151 (3), 208–229.
- Joy, E.J.M., et al., 2015. Dietary mineral supplies in Malawi: spatial and socioeconomic assessment. *BMC Nutrition*, 1 (1), 25.
- Karl, J.W., 2010. Spatial predictions of cover attributes of rangeland ecosystems using regression kriging and remote sensing. *Rangeland Ecology & Management*, 63 (3), 335–349.
- Kinkeldey, C., MacEachren, A.M., and Schiewe, J., 2014. How to assess visual communication of uncertainty? a systematic review of geospatial uncertainty visualisation user studies. *The Cartographic Journal*, 51 (4), 372–386.
- Kunz, M., Grêt-Regamey, A., and Hurni, L., 2011. Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system. *Natural Hazards*, 59 (3), 1735–1751.
- Lark, R.M., et al., 2014. 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., and Broadley, M.R., 2019. Combining two national-scale datasets to map soil properties, the case of available magnesium in England and Wales. *European Journal of Soil Science*, 70 (2), 361–377.
- Lark, R.M., Ander, L., and Knights, K., 2016. Information for agriculture from regional geochemical surveys: the example of soil pH in the Tellus and Tellus Border data. In: M. Young, ed. *Unearthed: impacts of the Tellus surveys of the North of Ireland*. Dublin, Ireland: Royal Irish Academy, 193–204.
- Lark, R.M., and Knights, K.V., 2015. The implicit loss function for errors in soil information. *Geoderma*, 251–252, 24–32.
- Li, D.R., Zhang, J.X., and Wu, H.Y., 2012. Spatial data quality and beyond. *International Journal of Geographical Information Science*, 26 (12), 2277–2290.
- Li, L., et al., 2018., 1.22 - spatial data uncertainty. In: B. Huang, ed. *Comprehensive geographic information systems*. Amsterdam: Elsevier, 313–340.
- Ligowe, I.S., et al., 2020. Selenium deficiency risks in sub-Saharan African food systems and their geospatial linkages. *Proceedings of the Nutrition Society*, 79 (4), 457–467.
- Likoswe, B.H., et al., 2020. Inflammation adjustment by two methods decreases the estimated prevalence of zinc deficiency in Malawi. *Nutrients*, 12 (6), 1563.

- Marchant, B.P., et al., 2011. Spatial prediction of soil properties with copulas. *Geoderma*, 162 (3–4), 327–334.
- Marchant, B.P., Saby, N.P.A., and Arrouays, D., 2017. A survey of topsoil arsenic and mercury concentrations across France. *Chemosphere*, 181, 635–644.
- Mastrandrea, M.D., et al., 2010. *Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties*. Intergovernmental Panel on Climate Change (IPCC), Report October. Available from: <https://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf>.
- McDonald, J.B., and Xu, Y.X., 1995. A generalization of the beta distribution with applications. *Journal of Econometrics*, 66 (1–2), 133–152.
- Meredith, M., and Kruschke, D., 2018. *HDInterval: Highest (Posterior) Density Intervals*. R package version 0.2.0. Available from: <https://CRAN.R-project.org/package=HDInterval>.
- Milne, A.E., et al., 2015. Communicating the uncertainty in estimated greenhouse gas emissions from agriculture. *Journal of Environmental Management*, 160, 139–153.
- Narasimhan, B., et al., 2020., *cubature: Adaptive Multivariate Integration over Hypercubes*. R package version 2.0.4. Available from: <https://CRAN.R-project.org/package=cubature>.
- Nelder, J.A., and Mead, R., 1965. A simplex method for function minimization. *The Computer Journal*, 7 (4), 308–313.
- O’Hagan, A., et al., 2006., *Uncertain judgements: eliciting experts’ probabilities*. Chichester, UK: John Wiley & Sons.
- Phiri, F.P., et al., 2020. Urine selenium concentration is a useful biomarker for assessing population level selenium status. *Environment International*, 134, 105218.
- R Core Team. 2020., *R: A language and environment for statistical computing*. Available from: <https://www.r-project.org/>.
- Ramsey, M., Taylor, P., and Lee, J., 2002. Optimized contaminated land investigation at minimum overall cost to achieve fitness-for-purpose. *Journal of Environmental Monitoring*, 4 (5), 809–814.
- Reagan-Cirincione, P., 1994. Improving the accuracy of group judgement: a process intervention combining group facilitation, social judgement analysis and information technology. *Organizational Behavior and Human Decision Processes*, 58 (2), 246–270.
- Rivoirard, J., 1994. *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford, UK: Oxford University Press.
- Spiegelhalter, D., and Rice, K., 2009. Bayesian statistics. *Scholarpedia*, 4 (8), 5230.
- Spiegelhalter, D., Pearson, M., and Short, I., 2011. Visualizing uncertainty about the future. *Science*, 333 (6048), 1393–1400.
- Tjims, H., 2018. *Probability: a lively introduction*. Cambridge: Cambridge University Press.
- Tversky, A., and Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*, 211 (4481), 453–458.
- Webster, R., and Oliver, M.A., 2007. *Geostatistics for natural environmental scientists*. 2nd ed. West Sussex: John Wiley & Sons Chichester.
- Zeileis, A., et al., 2020. Colorspace: A toolbox for manipulating and assessing colors and palettes. *Journal of Statistical Software*, 96 (1), 1–49.