

A framework for the modelling of uncertainty between remote sensing and geographic information systems

Mark Gahegan^{a,*}, Manfred Ehlers^b

^a *Department of Geography, The Pennsylvania State University, University Park, PA 16802, USA*

^b *Institute for Environmental Sciences, University of Vechta, Germany*

Received 29 March 1999; accepted 4 March 2000

Abstract

This paper addresses the modelling of uncertainty in an integrated geographic information system (GIS), specifically focused on the fusion of activities between GIS and remote sensing. As data is abstracted from its 'raw' form to the higher representations used by GIS, it passes through a number of different conceptual data models via a series of transformations. Each model and each transformation process contributes to the overall uncertainty present within the data. The issues that this paper addresses are threefold. Firstly, a description of various models of geographic space is given in terms of the inherent uncertainty characteristics that apply; this is then worked into a simple formalism. Secondly, the various transformation processes that are used to form geographic classes or objects from image data are described, and their effects on the uncertainty properties of data are stated. Thirdly, using the formalism to describe the transformation processes, a framework for the propagation of uncertainty through an integrated GIS is derived. By way of a summary, a table describing sources of accumulated uncertainty across four underlying models of geographic space is derived. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: uncertainty; GIS integration; data modelling; transformation description

1. Introduction

Good science requires statements of accuracy by which the reliability of results can be understood and communicated. Where accuracy is known objectively, then it can be expressed as error, where it is not, the term uncertainty applies (Hunter and Goodchild, 1993). Thus, uncertainty covers a broader range of doubt or inconsistency, and in the context of this paper, includes error as a component. The under-

standing of uncertainty as it exists in geographic data remains a problem that is only partly solved (Story and Congleton, 1986; Goodchild and Gopal, 1989; Veregin, 1995; Ruiz, 1997; Worboys, 1998). However, without quantification, the reliability of any results produced remains problematic to assess and difficult to communicate to the user. A geographic information system (GIS) provides a whole series of tools with which data can be manipulated, without offering any control over misuse. To quote Openshaw et al. (1991):

A GIS gives the user complete freedom to combine, overlay and analyse data from many different sources, regardless of scale, accuracy, resolu-

* Corresponding author. Tel.: +1-814-865-2612; fax: +1-814-863-7943 <http://www.geog.psu.edu/~mark>.

E-mail address: mark@geog.psu.edu (M. Gahegan).

tion and quality of the original map documents and without any regard for the accuracy characteristics of the data themselves.

This is a serious issue; without quantification of uncertainty, the results themselves may only be considered as qualitative information, and this greatly devalues their merit in both a scientific and a practical sense. To compound the problem, in the fusion of activities from remote sensing and GIS, an integrated approach to managing geographic information is required. This must necessarily support many different types of data (Ehlers et al., 1991), gathered according to different models of geographic space (Goodchild, 1992), each possessing different types of inherent errors and uncertainties (Chrisman, 1991). As well as providing individual support for these different models of space, it is necessary to explicitly include methods to keep track of uncertainty, as data is changed from low level forms (such as remotely sensed image data) to the higher level abstractions required by cartography and GIS (such as distinct objects and themes). It is particularly toward the propagation of uncertainty through different conceptual models that this paper is oriented.

Whether a particular data set can be considered suitable for a given task depends on many different criteria, and despite the fact that various aspects of uncertainty can be measured objectively, their importance will be largely determined by the current task. Our overall goal with modelling uncertainty is therefore threefold: (i) to produce a statement of uncertainty to be associated with each data set so that an objective statement of reliability may be reported, (ii) to develop methods to propagate uncertainty as the data is processed and transformed, and (iii) to ultimately determine the suitability of a data set for a given task ('fitness for use'). Another valid goal, not considered here, is to communicate uncertainty information to the user (e.g. Hunter and Goodchild, 1996).

1.1. Current approaches

A good deal of research effort has been directed towards the study of uncertainty in geographic data. A useful framework, recognising the separate error components of value, space, time, consistency and completeness was proposed by Sinton (1978) and

later embellished by Chrisman (1991). However, uncertainty in geographic data can be described in a variety of alternative ways, such as those provided by Bedard (1987), Miller et al. (1989) and Veregin (1989). Although different, these approaches all have a number of aspects in common, including the observation that uncertainty itself occurs at different levels of abstraction. For example, positional and temporal errors describe uncertainty in a metric sense within a space–time framework, whereas completeness and consistency represent more abstract concepts describing coverage and reliability, and are consequently more problematic to describe.

Work to date on uncertainty addresses the inherent errors present within specific types of data structure (e.g. raster or vector) or data models (e.g. field, object). The affects of combining data layers together within these various paradigms have been studied by Veregin (1989, 1995), Openshaw et al. (1991), Goodchild et al. (1992), Heuvelink and Burrough (1993), Ehlers and Shi (1997) and Leung and Yan (1998). Scant attention has so far been given to the problem of modelling uncertainty as the data is transformed through different models of geographic space, a notable exception is the work of Lunetta et al. (1991). A typical path taken by data captured by satellite, then abstracted into a suitable form for GIS is shown in Fig. 1, and involves four such models. Continuously varying fields are quantised by the sensing device into image form, then classified and finally transformed into discrete mapping objects. The overall object extraction process is sometimes referred to as semantic abstraction (Waterfeld and Schek, 1992) due to the increasing semantic content of the data as it is manipulated into forms that are easier for people to work with.

1.2. Problem definition

When transforming data between different conceptual models of geographic space, the uncertainty

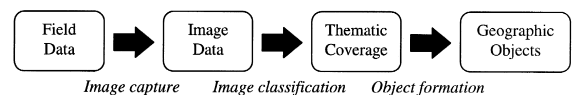


Fig. 1. Continuum of abstraction from field model to feature model.

characteristics in the data may change; in that techniques used to transform the data also alter the inherent uncertainty and in addition may introduce further uncertainty of their own. Furthermore, many of the abstraction techniques employed combine data with different uncertainty characteristics, for example multi-temporal image classification (Jeon and Landgrebe, 1992) or knowledge-based feature extraction (McKeown, 1987). Consequently, two inter-related problems must be addressed, namely:

1. How do the uncertainty characteristics of data change as data are transformed between models?
2. How do the transformation methods used affect and combine the uncertainty present in the data (Lodwick et al., 1990)?

This paper concentrates mainly on the first question, proposing a framework within which the second question can be tackled.

One of the consequences of the separation of GIS and RS activities into separate communities and separate software environments is that there is an artificial barrier between the two disciplines. Thus, the integration of these two branches of science is to some extent an artificial problem. As a result, there is no easy flow of meta-data between systems, interoperability is often restricted to the exchange of image files or object geometry and the problem of managing uncertainty is compounded.

The four stages shown in Fig. 1 represent four models of geographic space that are considered here. They are termed the field (F), image (I), thematic (T) and object (or feature) (O) models and are typical (though not exhaustive) of those used in the integration of GIS and remote sensing activities. These models represent the conceptual properties of the data only and are considered here as independent from any particular data structure that might be used to encode and organise the data.

The description of uncertainty used here follows that proposed by Sinton (1978). It covers the sources of error as they occur in remote sensing and GIS integration (although other approaches may be equally valid). Uncertainty is restricted to the following properties: (i) value (including measurement and label errors), (ii) spatial, (iii) temporal, (iv) consistency and (v) completeness. These are symbolised

below using the following five letters of the Greek alphabet (α , β , χ , δ and ε , respectively). Of these, α , β and χ can be applied either individually to a single datum or to any set of data. The latter two properties of consistency and completeness can only apply to a defined data set since they are comparative (either internally among data or to some external framework). Issues regarding scale and resolution of the data (e.g. Bruegger, 1995) are postponed; these are not objective uncertainty criteria, but instead become important when assessing the 'fitness for use' of data for a specific task. Also, we do not concern ourselves here with structures for the provision of lineage information (e.g. Lanter, 1991), although these are certainly required to support the propagation of uncertainty.

The techniques proposed for quantifying error and uncertainty referred to in this paper are not in themselves new. The purpose of the formal notation developed later is to describe succinctly and unambiguously the forms of uncertainty that exist in the data, where they originate from and what they affect. Ongoing work by the authors and others is attempting to quantify some of these uncertainty terms as they relate to the integration of GIS and remote sensing, for example ISPRS Commissions 2, 3 and 4 span the full range of integration activities described here (see: http://www.isprs.org/technical_commissions.html for more details).

2. Description of data model properties and their uncertainty characteristics

More formally, a single geographic datum (a) at any level of abstraction can be described by its value (d), spatial extents (s) and temporal extents (t) (Gahegan, 1996) each of which have associated uncertainty u ; where u has three components, α , β and χ

$$a(d, s, t, \alpha, \beta, \chi). \quad (1)$$

The above expression assumes that α , β and χ may be orthogonalised. As will be shown later, it is often difficult to treat these components separately, since each may affect the others. A similar formula-

tion, using upper case letters, describes a data set, but now also includes terms for consistency (δ) and completeness (ε)

$$A(D, S, T, \alpha, \beta, \chi, \delta, \varepsilon). \quad (2)$$

Each datum exists only within the geographic model under which it was captured or made, so a description of a datum or data set includes a reference to a conceptual model, C_x , as shown above, where C_O = object, C_T = thematic, C_I = image and C_F = field

$$A(D, S, T, \alpha, \beta, \chi, \delta, \varepsilon):(C_x). \quad (3)$$

The transformations shown in Fig. 1 may be described generally by

$$A(C_F) \rightarrow A'(C_I) \rightarrow A''(C_T) \rightarrow A'''(C_O) \quad (4)$$

where each transformation modifies the data characteristics described above. At each stage in the process, a new data set is created, having modified values for some or all of the above properties. The following section describes these properties as they apply to the four conceptual models considered here. Following from this, the outcome of the transformation process on these properties is described, providing one way to specify the effects of these transformations on all the properties of a data set given above.

Veregin (1989) defines two distinct categories of uncertainty as: (i) ‘error propagation’, describing the manner in which error accumulates during processing and (ii) ‘error production’, defined as the introduction of error by the use of incorrect procedures or data. Here we note that ‘error production’ must be widened to include errors induced by the process of semantic abstraction. Each of the transformations shown in Fig. 1 have a quantising or generalising effect on the data, changing the nature of the data value and its inherent uncertainty. Specifically, these include gridding error, classification error, object recognition and object definition error.

As described in Section 1, the aim of this research is to account for all the types of uncertainty that are present or have risen in data sets as they undergo processing and transformation. Since uncertainty does not just ‘happen’, each uncertain term must originate from some data source or process, and thus may be

accounted for. Therefore, if an uncertainty term appears on the right hand side of any expression (Eq. 5), then a corresponding term must occur within at least one of: (i) the input data sets $A()$ (left hand side), (ii) as a consequence of the transformation Ψ or (iii) as additional data (Q) used by a transformation (such as spectral training pixels or ground control points (GCPs))

$$A() \xrightarrow[\quad]{\Psi} A'(). \quad (5)$$

The origins of each uncertainty term are thus described in a straightforward manner, to which simple rules may be applied to advise of potential mistakes, inconsistencies and invalid assumptions. Section 3 attempts to describe these sources of uncertainty, where they originate from and what effect they have.

3. The transformation process between geographic models

This section describes the three transformation processes between geographic models, field \rightarrow image, image \rightarrow theme and theme \rightarrow object and their effects on the data characteristics introduced above.

3.1. From field model to image model (image capture)

Remote sensing devices sample the physical fields produced by different forms of electro-magnetic radiation from which an image ‘snapshot’ is captured. Hence, the act of image capture is actually a transformation from the field model to the image model.

$$A(D, S, T):(C_F) \rightarrow A'(D', S', T, \alpha, \beta, \delta):(C_I). \quad (6)$$

The above expression implies the following.

(1) The objective reality of the sensed radiation is assumed to contain no errors, so the left side of the expression has no uncertainty terms.

(2) A change in the set of data values, D , occurs on image capture. This is due to measurement error and quantisation effects, which are determined by

sensor design and operation. Consequently, the right side gains a term for value uncertainty, α . (Inconsistent response across the sensor due to manufacturing or operating discrepancy and sensor drift may also add to α , but an attempt is usually made to remove these effects by a calibration correction).

(3) A change in S is enforced by the gridding process, i.e. spatial quantisation, and is also subject to sensor design and operation. Spatial uncertainty (β) is introduced as a result of gridding since the geometry of the sensing process usually produces pixels which represent different sized regions, due to the combination of platform position and attitude, sensing angle and landscape topography. That portion of β due to the sensor can be described by a function applied across all pixels, according to their position, so that it is consistent. Distortion due to relief is inconsistent and can be significant in mountainous or built up areas and requires an elevation model for correction.

(4) No term for temporal uncertainty is required. In most practical cases, image capture may be assumed to be instantaneous (denoted by T) and known within the accuracy required for analysis. A single instantaneous value is, therefore, sufficient. However, there are examples where the sensing cycle is not negligible, such as when a mosaic scene is being produced from collections of aerial photographs.

(5) The image is assumed to be complete but not necessarily consistent, hence the introduction of δ in the result. Completeness is achieved since each pixel has a value, and consistency may be achieved if all pixels are sensed in the same manner. In practice, the image may not be gathered consistently due to partial cloud or some other atmospheric condition (Kaufman, 1988). For example, the reflected signal is effectively filtered by the intervening space. The filter itself is inconsistent and may vary spatially, hence the introduction of the uncertainty term for value and consistency. For satellites such as Landsat TM, it is not possible to be certain of the extent of atmospheric effects since no part of the sensor can quantify their presence. However, it is a usual practice to try to work with images that appear to be consistent and the processing techniques that are applied often assume image consistency.

After being captured, the image data often undergoes a number of additional processes in order to

prepare it for further analysis. Each of these transformations may be described with a similar expression to that given above, except that they all retain the image as a basis so no model transformation occurs. Three common examples follow.

Registration, by the use of GCPs, introduces an additional form of spatial uncertainty due to the positional error in the GCPs themselves and the algorithm used to perform the image resampling. The nearest neighbour resampling method does not change pixel values (D) but does affect the uncertainty of these values (α) due to translation

$$A(D, S, T, \alpha, \beta, \delta) \xrightarrow[\text{GCPs}(S, \beta, \varepsilon)]{\Psi_{\text{registration}}} A'(D, S', T, \alpha', \beta', \delta, \varepsilon). \quad (7)$$

Interpolation using bilinear or cubic methods additionally modifies data values as well as adding to value uncertainty, by changing the spatial basis (S) and then re-apportioning the data to it

$$A(D, S, T, \alpha, \beta, \delta, \varepsilon) \xrightarrow[\text{Grid}(S')]{\Psi_{\text{resampling}}} A'(D', S', T, \alpha', \beta', \delta, \varepsilon). \quad (8)$$

Certain image enhancement methods (such as contrast stretching) may also dramatically change the data values, although they usually preserve the ordering of any trends.

$$A(D, S, T, \alpha, \beta, \delta, \varepsilon) \xrightarrow{\Psi_{\text{contrast_stretch}}} A'(D', S, T, \alpha', \beta, \delta, \varepsilon). \quad (9)$$

The above expressions show that apart from uncertainty inherent in the image acquisition process itself, associated pre-processing activities introduce a variety of data and spatial uncertainty that accumulates at each successive transformation.

3.2. From image model to thematic model (image classification)

Thematic interpretations help us to understand the trends, structure or content of imagery in higher level

terms by replacing a (multi-valued) sensor reading with a single class label (i.e. the value element in the data description changes type). Measurement errors no longer apply but are replaced with measures of confidence in the labelling process. It is common to use some form of supervised classification as the transformation process, in which case each theme or class is described in terms of examples or average characteristics.

Individual image pixels will have a different response to the classifier (for a maximum likelihood classifier (MLC) this can be quantified in terms of posterior probability values). So, the thematic value error applies on a per-pixel basis, but may be summarised by class for the entire coverage. If we regard the classification process simply in terms of changing the value associated with a pixel, then we may divorce α and β , since we need never be concerned as to the overall shape of a class or theme. However, this does not help the user who wishes to understand the likely accuracy of a boundary, or the chances that an object is incorrectly segmented or labelled. This issue is thus carried over from the thematic model into the object model and is described in Section 3.3.

Many different approaches to classification are possible, employing a wide range of additional data sources. Some distinct scenarios are considered below.

3.2.1. Single image classification

For each pixel, the image data (D) are replaced with class labels (D'). The expression shown below is of a MLC, an example of a supervised classifier. Supervised classification uses data examples for training from which k class descriptions ($Q_{1...k}$) are created as probability density functions. These descriptions act as a controlling input to the classification process Ψ (but are not changed by it), and are shown under the transformation arrow.

The spatial extents of the class descriptions used for training, $Q(S^*)$ are assumed to be the entire coverage; S^* is given by the spatial extents of $A(S)$. In other words, we assume the class descriptions can be applied to any pixel in the image. It is interesting to note that each class description has, in theory at least, completeness and consistency properties that describe the validity of this assumption. Since the

class descriptions are used comparatively for labelling, any uncertainty they possess is likely to increase the final labelling uncertainty, $A'(\alpha)$

$$A(D, S, T, \alpha, \beta, \delta) : (C_1)$$

$$\xrightarrow{\Psi_{\text{MLC}}} \left. \begin{array}{l} Q_1(D_1, S^*, T, \alpha_1, \delta_1, \varepsilon_1) \\ Q_2(D_2, S^*, T, \alpha_2, \delta_2, \varepsilon_2) \dots \\ Q_k(D_k, S^*, T, \alpha_k, \delta_k, \varepsilon_k) \end{array} \right\}$$

$$A'(D', S, T, \alpha', \beta, \delta', \varepsilon) : (C_T). \quad (10)$$

The following points also apply.

(1) D properties change from a vector of quantitative values to a single class label (or possibly a probabilistic label vector if the classification is 'soft').

(2) S and β properties remain unchanged. The spatial basis of the image is retained, so S and β are not affected.

(3) T properties remain unchanged; the resulting classes represent an interpretation of a single image in isolation so no temporal uncertainty is introduced by the classification process. This and the previous point assume that $A(S) = Q(S_{1...k})$ and $A(T) = Q(T_{1...k})$, i.e. the training values are drawn from A . Major problems may occur if $A(S) \neq Q(S_{1...k})$ or $A(T) \neq Q(T_{1...k})$, since the training data are representatives of different place or time, giving rise to spatial or temporal inconsistency (Section 3.2.3).

(4) α must either expand or change to include the uncertainty in the class labelling process, on a per-pixel basis, to represent our confidence in the classifier.

(5) Completeness may be an issue since the defined classes themselves may not cover all possibilities. Pixels may be left unassigned or lumped together into an 'unknown' class; the partitioning of pixels into classes is thus incomplete. Any such problems originate with the training data in terms of how well each individual class is described. This is shown by the value and completeness uncertainty terms in the description of $Q_{1...k}$, and is carried over into the result, A' . In the case of completeness, a new error term is introduced into the result as a consequence. Of course, these uncertainties in A' are not the same as those in A or Q , but instead

represent the consequences of uncertainty in A and Q as manifested in the thematic model.

(6) It is normal for a consistent approach to classification to be taken, where all pixels are treated exactly the same by the classifying algorithm; that being the case, no further consistency error is introduced by Ψ_{MLC} .

The accumulation of uncertainty from one model to another is certainly not straightforward; large but consistent ‘errors’ in recorded pixel reflectance values may have little influence on classification accuracy. For example, reflectance values in Landsat TM may be attenuated by up to 40% by non-obvious atmospheric effects (i.e. $A(\alpha) = 40\%$). However, this results only in a reduction of the dynamic range of the data; the ability of a classifier to separate out the themes will be impaired by a much smaller amount — provided, of course, that the training data values are taken from the affected image. Inconsistencies in the attenuation are likely to have a much greater effect than the attenuation itself. Another related problem may arise from effects introduced by terrain variations, e.g. shadows within otherwise consistent classes that cause misclassifications.

3.2.2. Multi-temporal image classification

Now we consider the case where the same sensor provides two or more temporally separated images that are used together as input to the classifier (Jeon and Landgrebe, 1992), shown by

$$\left. \begin{array}{l} A_1(D_1, S, T_1, \alpha_1, \beta_1, \delta_1):(C_1) \\ A_2(D_2, S, T_2, \alpha_2, \beta_2, \delta_2):(C_1) \dots \\ A_n(D_n, S, T_n, \alpha_n, \beta_n, \delta_n):(C_1) \end{array} \right\} \xrightarrow{\Psi_{MLC}} \left. \begin{array}{l} Q_1(D_1, S^*, T, \alpha_1, \delta_1, \varepsilon_1) \\ Q_2(D_2, S^*, T, \alpha_2, \delta_2, \varepsilon_2) \dots \\ Q_k(D_k, S^*, T, \alpha_k, \delta_k, \varepsilon_k) \end{array} \right\} A'(D', S, T', \alpha', \beta', \delta', \varepsilon):(C_T). \quad (11)$$

It is assumed that all images are co-registered and therefore share the same spatial basis (S) with known uncertainty $\beta_1 \dots \beta_n$. As before, temporal uncertainty is not an issue, since the capture date and time for each image should be known accurately. However, the

time represented by the resulting classified overlay poses a philosophical problem. The contributing images each represent different dates, possibly from different seasons (perhaps to improve crop differentiation) or from different years but the same season (to study landcover change). In either case, the input data represents several discrete points in time, between which we know nothing but may infer much due to knowledge of plant growth, cropping patterns and so forth. In the former case, we are inferring a continuation in landcover type, using evidence from several epochs to differentiate between classes that may be largely inseparable using only a single scene. In the latter case, the classes created represent types of change, rather than types of landcover. So, in both cases the resulting temporal property should represent the (convex) temporal interval covering all observed times, as described by Ladkin (1986). That being the case, the resulting temporal interval is given by first calculating the union of all contributing times and then forming the interval that exactly spans the resulting set T^* as

$$T = \text{Convex Interval} \left(T^* = \bigcup_{i=1}^n T_i \right). \quad (12)$$

As Fig. 2 shows, each of the input data sets can legitimately take any location within a positional uncertainty band (Shi, 1994), and each possible combination of locations produces an equally viable thematic coverage. Such perturbations can be simulated using Monte Carlo techniques (e.g. Openshaw et al., 1991). Of course, for a realistic simulation, the other

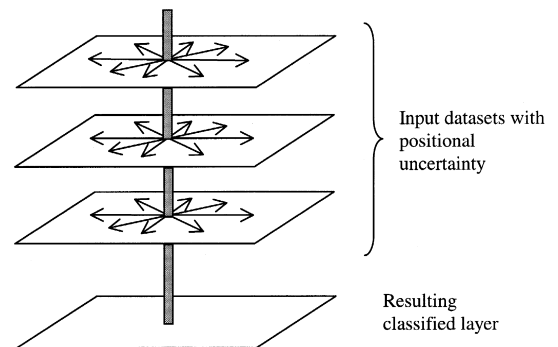


Fig. 2. Three layers of data combined via classification, each having positional uncertainty.

forms of uncertainty (such as value error and grid distortion) must also be varied concurrently.

In the absence of any bias, the results of any such simulation will show central tendency since the error measures represent distributions where values at the extremes are rarer. Consequently, the straightforward combination of the data, ignoring the uncertainty, is likely to produce the best estimate. However, knowledge of the uncertainty can be useful in ascertaining the reliability of this result. From the perspective of classification, what is most interesting is not the uncertainty envelopes per se, but rather where the conjunction of uncertainty measures can give rise to a different interpretation — i.e. a different class label assigned to a pixel. A practical goal is therefore to be able to report, for each pixel, whether there is scope within the accumulated uncertainty for a different outcome to have occurred, and its likelihood.

3.2.3. Single image classification using ancillary data sources

Two scenarios are considered. In the first, an existing spatial framework is provided (for example, a set of field boundaries from a GIS coverage) to which labels are to be added (for example, the crop type). The spatial framework in data set A_2 imposes geometric constraints (object shapes) onto the classifier. As before, the labelling, uncertainty originates entirely from the image data (D_1) but now the spatial uncertainty (β) of the result is determined by the boundary data set (S_2) and by the positional error of the image (S_1):

$$\left. \begin{array}{l} A_1(D_1, S_1, T_1, \alpha_1, \beta_1, \delta_1):(C_1) \\ A_2(D_2, S_2, T_2, \alpha_2, \beta_2, \chi_2, \delta_2, \varepsilon_2):(C_0) \end{array} \right\} \xrightarrow{\Psi_{\text{LABELLING}}} \left. \begin{array}{l} Q_1(D_1, S^*, T, \alpha_1, \delta_1, \varepsilon_1) \\ Q_2(D_2, S^*, T, \alpha_2, \delta_2, \varepsilon_2) \dots \\ Q_k(D_k, S^*, T, \alpha_k, \delta_k, \varepsilon_k) \end{array} \right\} A_R(D_R, S_R, T_R, \alpha_R, \beta_R, \chi_R, \delta_R, \varepsilon_R):(C_0). \quad (13)$$

Notice that this expression contains a mix of conceptual models. The labelling method must resolve this

difference somehow before the image and object data may be successfully merged.

The second scenario is so-called ‘contextual’ classification as described by Gurney (1983), where many different ancillary data layers (such as soil type and elevation) are added into the process with the aim of improving class differentiation

$$\left. \begin{array}{l} A_1(D_1, S, T_1, \alpha_1, \beta, \delta_1):(C_1) \\ A_2(D_2, S, T_2, \alpha_2, \beta_2, \chi_2, \delta_2, \varepsilon_2):(C_1) \dots \\ A_n(D_n, S, T_n, \alpha_n, \beta_n, \chi_n, \delta_n, \varepsilon_n):(C_1) \end{array} \right\} \xrightarrow{\Psi_{\text{MLC}}} \left. \begin{array}{l} Q_1(D_1, S^*, T, \alpha_1, \delta_1, \varepsilon_1) \\ Q_2(D_2, S^*, T, \alpha_2, \delta_2, \varepsilon_2) \dots \\ Q_k(D_k, S^*, T, \alpha_k, \delta_k, \varepsilon_k) \end{array} \right\} A_R(D_R, S_R, T_R, \alpha_R, \beta_R, \chi_R, \delta_R, \varepsilon_R):(C_T). \quad (14)$$

In addition to the uncertainty described in the previous example, further uncertainties may arise as a result of the possible mismatch between the S and T properties of all input data sets $A_1 \dots A_n$. These data sets must each provide data values for a set of given spatial locations, defined by each pixel in the image, i.e., the value of d and α must be determined in A_2, \dots, A_n for each pixel location in all input data sets; a task which may require registration if these data sets do not share the same spatial structure as the image. As previously described (Section 3.1), registration carries its own data uncertainty legacy.

For a given pixel or spatial location(s), the class label produced is typically a function of the data values in each data set (Heuvelink et al., 1989)

$$A'(d', s) = f(A_1(d_1, s), A_2(d_2, s), \dots A_n(d_n, s)), \quad (15)$$

where f might represent a MLC. In a similar manner, the resulting uncertainty term in α represents some (usually unknown) function applied across the value uncertainty of each contributing layer at the same location, described as

$$A'(s, \alpha') = f(A_1(s, \alpha_1), A_2(s, \alpha_2), \dots A_n(s, \alpha_n)). \quad (16)$$

However, in practice α is defined not by the input data uncertainty, but by the posterior probabilities of the classification process, and most commonly by the sum of the posterior probabilities excluding the dominant class (q), i.e.

$$A'(s, \alpha') = \sum_{i=1}^k P(A'(d', s) = i) \quad i \neq q. \quad (17)$$

The above expression ignores the uncertainty already present in the image data, which in turn may lead to an overestimate of the reliability of the result.

Temporal uncertainty may arise when data is used to represent a time other than that at which it was gathered. This does not apply to the previous cases (described in Section 3.2.2), since the data is used according to the time it was captured. However, where different ancillary data sets are included, then temporal uncertainty may creep in. Examples include the use of historically averaged rainfall data, the assumption being that the same values apply over an instantaneous image, or where data from previous times (e.g. years) are used by necessity. The ‘temporal distance’ between two data sets is difficult to quantify and depends on the nature of the data concerned and the rate at which it changes. For example, a geology map is unlikely to become out of date within the lifetime of a system, but a remotely sensed image might become obsolete in a matter of days. Such data is cyclic in its usefulness, but not perfectly so, since two images 1 year apart will not represent the same point in the season — the environmental conditions may well be different.

3.2.4. Multi-sensor classification

Where images are acquired from different platforms and combined (fused) in the classification process, then the data will initially exist in two different spatial bases, hence $S_1 \neq S_2$. The resulting grid is usually that of the data set with the highest resolution, so $S_R = \max(S_1, S_2)$. Spatial uncertainty becomes dependent on both β_1 and β_2 . As in the collaborative labelling example (Eq. 13) above, the method employed must be more sophisticated than a conventional classifier, to manage the conflicting spatial bases.

$$\left. \begin{array}{l} A_1(D_1, S_1, T_1, \alpha_1, \beta_1, \delta_1):(C_1) \\ A_2(D_2, S_2, T_2, \alpha_2, \beta_2, \delta_2):(C_1) \end{array} \right\} \xrightarrow{\Psi_{\text{FUSION}}} \left. \begin{array}{l} Q_1(D_1, S^*, T, \alpha_1, \delta_1, \varepsilon_1) \\ Q_2(D_2, S^*, T, \alpha_2, \delta_2, \varepsilon_2) \dots \\ Q_k(D_k, S^*, T, \alpha_k, \delta_k, \varepsilon_k) \end{array} \right\} A_R(D_R, S_R, T_R, \alpha_R, \beta_R, \chi_R, \delta_R, \varepsilon_R):(C_T). \quad (18)$$

3.3. From thematic model to object model

An object is considered here as an independent entity with a defined shape; any pixels incorrectly included or excluded from the entity will cause the shape to be in error. This gives rise to a new dimension of spatial uncertainty in the object model due to the need to define the shape property (unlike a pixel, whose shape is given). The object model thus has shape error, which is treated here as an extension of β into the object model. Any positional error inherited from the image data remains, and becomes positional error in objects derived from this data. It is typical in remote sensing applications that classification errors tend to occur around the edges of objects (Shi, 1994), due to transition zones and mixed pixel effects. It is also typical of some types of objects that they do not possess discrete boundaries to begin with (Burrough, 1989; Edwards and Lowell, 1996). Fig. 3 shows two different types of spatial error for a given area object. The arrows represent positional error, while the different shadings represent possible shapes for the object.

In a similar manner, α must expand to include identity uncertainty, representing the chances that the object is incorrectly ‘typed’, perhaps as a result of classification uncertainty

$$A(D, S, T, \alpha, \beta, \chi, \delta, \varepsilon):(C_T) \xrightarrow[\text{Object Formation Rules}]{\Psi_{\text{Object Formation}}} A'(D', S', T, \alpha', \beta', \chi, \delta', \varepsilon):(C_o). \quad (19)$$

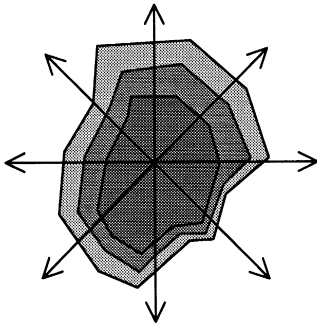


Fig. 3. Two forms of spatial uncertainty in an object, inherited from the image from which it was formed and due to uncertainty in boundary delineation and position.

Gahegan and Flack (1999) provide one possible approach to combine attribute and positional uncertainty encountered during object formation using Dempster–Shafer evidential reasoning (Shafer, 1976).

Another source of uncertainty within the object model is that of defining the objects themselves. These are typically imprecise, if defined in terms of image properties, or difficult to quantify, if defined in high-level terms such as those suggested by Kuhn (1994). Consequently, object types (which are imprecise) pass on uncertainty to object instantiations as an additional facet to object identity uncertainty described by α .

Most object formation strategies function basically the same way, although they differ in complexity of reasoning — an object acquires its label due to passing a threshold of similar or identically labelled pixels within a hypothesised region. Where uncertainty is high, it is possible that a predominance of different values or labels might arise under different instantiations of the uncertainty components, leading to a different object being formed.

This situation is most likely to occur where: (i) the object itself is small, (ii) positional error is large and (iii) the values or class labels are ‘close’ or overlap (in other words, differentiation is poor or the classification scheme is inappropriate). Reporting methods exist in remote sensing to identify this latter problem, such as the classifier accuracy table showing errors of omission and commission, but there is no counterpart in the object domain.

3.3.1. From image model directly to object model

Some object extraction strategies bypass the classification stage altogether. These strategies differ greatly in terms of their complexity, their data requirements and their organisation. At one extreme, objects may be formed by directly interpreting the image data manually, as with ‘heads up’ digitising from an image on the display screen. At the other extreme, extraction strategies may employ sophisticated pattern recognition tools and high-level reasoning (McKeown, 1987; Matsuyama and Hwang, 1990; Gahegan and Flack, 1996; 1999). Such strategies make use of a variety of transformations that are not covered here. A useful overview of the general approach to object recognition taken in computer vision is given by Sonka et al. (1993, p. 43).

4. Discussion, conclusions and future work

The expressions described above seem to offer some insight into the complex transformation processes that occur within integrated GIS, particularly with reference to keeping track of the sources of errors and uncertainties and their effects. Table 1 categorises these errors and uncertainties as they apply to the four geographic models considered here.

Uncertainty propagates from left to right across the table as the higher models inherit uncertainty properties from the lower models and add to this their own specific measures. Of course, the increase in uncertainty is accompanied by a corresponding information ‘gain’ as a result of semantic abstraction — the identity of the data becomes more certain as the associated properties become less certain. For example, a thematic map may inherit data uncertainty from the imaging process as well as acquiring further label uncertainty through the imperfections of the classification process, but the result is a higher semantic entity. So, for example a value taken by α for a thematic data set must account for the data error in the input data set(s) as well as the classifier labelling uncertainty, L

$$\alpha_T = \alpha_I \oplus L. \quad (20)$$

Table 1

Types of uncertainty in four models of geographic space (α = data or value, β = space, χ = time, δ = consistency, ε = completeness)

	Field	Image	Thematic	Object
α	Measurement error and precision	Quantisation of value in terms of spectral bands and dynamic range	Labelling uncertainty (classification error)	Identity error (incorrect assignment of object type), object definition uncertainty
β	Locational error and precision	Registration error, sampling precision	Combination effects when data represented by different spatial properties are combined	Object shape error, topological inconsistency, 'split and merge' errors
χ	Temporal error and precision	(Temporal error and precision are usually negligible for image data)	Combination effects when data representing different times is combined	Combination effects when data representing different times is combined
δ	Samples/readings collected or measured in an identical manner	Image is captured identically for each pixel, but medium between satellite and ground is not consistent; inconsistent sensing, light falloff; shadows	Classifier strategies are usually consistent in their treatment of a data set	Methods for object formation may be consistent, but often are not. Depends on extraction strategy
ε	Sampling strategy covers space, time and attribute domains adequately	Image is complete, but parts of ground may be obscured	Completeness depends on the classification strategy. (Is all the data set classified or are only some classes extracted?)	Depends on extraction strategy. Spatial and topological inconsistencies may arise as a result of object formation

The definition of the combination operator, \oplus , is itself problematic. As Veregin (1989) shows, it is rarely appropriate to simply add the uncertainty terms together.

As semantic abstraction progresses, the 'fitness for use' of the data also changes; it becomes more readily used as an identifiable object as opposed to a continuous field. A question remains as to whether the associated loss of data (i.e. variance) should cause an uncertainty penalisation. We can be certain of the identity of the model to which the datum belongs, so there should be no mistakes in interpreting what the value represents. Therefore, abstraction should not necessarily, in itself, incur an uncertainty penalty. Classes are defined as collections with variance (e.g. MLC) so they already represent variance in a summarised form. However, definitions for objects in terms of variance and uncertainty are yet to be fully developed and represent a real barrier to progress at present. Quantifying the performance of object formation has become a lively area of debate among the computer vision community (e.g. Haralick, 1994), but a good deal more research is required before this final stage of model transformation can be fully understood and effectively communicated.

Perhaps the most difficult aspect of managing uncertainty in an integrated GIS is the inter-relationship between the various types of uncertainty described. For example, α and β strongly influence each other; an incorrect data value for a pixel may ultimately result in an object being assigned the 'wrong' shape. Conversely, positional errors in image data may lead to mis-classification. In data modelling terms, we say that α and β are functionally dependent on each other

$$\alpha \Rightarrow \beta \text{ and } \beta \Rightarrow \alpha. \quad (21)$$

Such circular referencing is not easily resolvable, but occurs as a consequence of moving from one model of space to another. Simulation and experimentation offer a way of understanding these inter-relationships, but new forms of metrics will be required to describe them.

Consistency and completeness are extremely problematic to describe, since they relate to the nature of variance of error over a coverage, and are thus a summary of the effects of α , β and χ . In many cases, acceptable metrics for δ and ε are simply not available so these terms cannot be easily

defined. Furthermore, their combination through transformation has not been addressed, as yet.

The approach described in this paper is currently being used as the basis for a simulation package, developed in an attempt to model these various forms of uncertainty in order to test their inter-relationships and to evolve better ways of assessing and reporting their significance.

References

- Bedard, Y., 1987. Uncertainties in land information databases. *Proc. Auto-Carto 8*, Baltimore, pp. 175–184.
- Bruegger, B.P., 1995. Theory for the integration of scale and representation formats: major concepts and practical implications. In: Frank, A.U., Kuhn, W. (Eds.), *Spatial Information Theory, Lecture Notes in Computer Science* vol. 988 Springer Verlag, Berlin, pp. 297–310.
- Burrough, P.A., 1989. Fuzzy mathematical methods for soil survey and land evaluation. *J. Soil Sci.* 40 (3), 477–492.
- Chrisman, N.R., 1991. The error component in spatial data. In: Maguire, D.J., Goodchild, M.F., Rhind, D.W. (Eds.), *Geographical Information Systems* vol. 1 Longman, pp. 165–174, Principles.
- Edwards, G., Lowell, K.E., 1996. Modelling uncertainty in photo-interpreted boundaries. *Photogramm. Eng. Remote Sens.* 62 (4), 377–391.
- Ehlers, M., Greenlee, D., Smith, T., Star, J., 1991. Integration of remote sensing and GIS: data and data access. *Photogramm. Eng. Remote Sens.* 57 (6), 669–675.
- Ehlers, M., Shi, W., 1997. Error modelling for integrated GIS. *Cartographica* 33 (1), 11–21.
- Gahegan, M.N., 1996. Specifying the transformations within and between geographic data models. *Trans. GIS* 1 (2), 137–152.
- Gahegan, M.N., Flack, J.C., 1996. A model to support the integration of image understanding techniques within a GIS. *Photogramm. Eng. Remote Sens.* 62 (5), 483–490.
- Gahegan, M., Flack, J.C., 1999. Recent developments towards integrating scene understanding within a geographic information system for agricultural applications. *Trans. GIS* 3 (1), 31–50.
- Goodchild, M.F., Gopal, S. (Eds.), *The Accuracy of Spatial Databases*. Taylor & Francis, London.
- Goodchild, M.F., 1992. Geographical data modelling. *Comput. Geosci.* 18 (4), 401–408.
- Goodchild, M.F., Guoqing, S., Shiren, Y., 1992. Development and test of an error model for categorical data. *Int. J. Geogr. Inf. Syst.* 6 (2), 87–104.
- Gurney, C.M., 1983. The use of contextual information in the classification of remotely sensed data. *Photogramm. Eng. Remote Sens.* 49 (1), 55–64.
- Haralick, R.M., 1994. Performance characterisation in computer vision. *CVGIP-Image Understanding* 60 (2), 245–249.
- Heuvelink, G.B.M., Burrough, P.A., Stein, A., 1989. Propagation of errors in spatial modelling with GIS. *Int. J. Geogr. Inf. Syst.* 3 (4), 303–322.
- Heuvelink, G.B.M., Burrough, P.A., 1993. Error propagation in cartographic modelling using Boolean logic and continuous classification. *Int. J. Geogr. Inf. Syst.* 7 (3), 231–246.
- Hunter, G., Goodchild, M.F., 1993. Mapping uncertainty in spatial databases, putting theory into practice. *J. Urban Reg. Inf. Syst. Assoc.* 5 (2), 55–62.
- Hunter, G., Goodchild, M.F., 1996. Communicating uncertainty in spatial databases. *Trans. GIS* 1 (1), 13–24.
- Jeon, B., Landgrebe, D.A., 1992. Classification with spatio-temporal interpixel class dependency contexts. *IEEE Trans. Geosci. Remote Sens.* 30 (4), 663–672.
- Kaufman, Y., 1988. Atmospheric effect on spectral signature — measurements and corrections. *IEEE Trans. Geosci. Remote Sens.* 26 (4), 441–450.
- Kuhn, W., 1994. Defining semantics for spatial data transfers. In: Waugh, T.C., Healey, R.G. (Eds.), *Proc. 6th International Symposium on Spatial Data Handling*. University of Edinburgh, Scotland, pp. 973–987.
- Ladkin, P., 1986. Time representation: a taxonomy of interval relations. *AAAI-86-5th National Conference on Artificial Intelligence* vol. 1 Morgan Kaufmann, Philadelphia PA, USA, pp. 360–366.
- Lanter, D.P., 1991. Design of a lineage-based meta-data base for GIS. *Cartography Geogr. Inf. Syst.* 18 (4), 255–261.
- Leung, Y., Yan, J., 1998. A locational error model for spatial features. *Int. J. Geogr. Inf. Sci.* 12 (6), 607–620.
- Lodwick, W.A., Monson, W., Svoboda, L., 1990. Attribute error and sensitivity analysis of map operations in geographical information systems: suitability analysis. *Int. J. Geogr. Inf. Syst.* 4 (4), 413–428.
- Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGwire, K.C., Tinney, L.R., 1991. Remote sensing and geographic information system data integration: error sources and research issues. *Photogramm. Eng. Remote Sens.* 57 (6), 677–687.
- Matsuyama, T., Hwang, V.S., 1990. *Sigma: A Knowledge-Based Aerial Image Understanding System*. Plenum, New York.
- McKeown, D.M. Jr., 1987. The role of artificial intelligence in the integration of remotely sensed data with geographic information systems. *IEEE Trans. Geosci. Remote Sens.* GE-25 (3), 330–348.
- Miller, R., Karimi, H., Feuchtwanger, M., 1989. Uncertainty and its management in geographical information systems. *Proc. CISM '89*. Canadian Institute of Surveying and Mapping, Ottawa, Canada, pp. 252–259.
- Openshaw, S., Charlton, M., Carver, S., 1991. Error propagation: a Monte Carlo simulation. In: Masser, I., Blakemore, M. (Eds.), *Handling Geographic Information*. Longman, England, pp. 78–101.
- Ruiz, M.O., 1997. A causal analysis of viewshed error. *Trans. GIS* 2 (1), 85–94.
- Shafer, G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton NJ, USA.
- Shi, W., 1994. Modelling positional and thematic uncertainties in

- the integration of remote sensing and geographic information systems. PhD Thesis, ITC (publication number 22), Enschede, The Netherlands.
- Sinton, D., 1978. The inherent structure of information as a constraint to analysis: mapped thematic data as a case study. In: Dutton, G. (Ed.), *Harvard Papers on Geographic Information Systems* vol. 6 Addison-Wesley, Reading MA, USA.
- Sonka, M., Hlavac, V., Boyle, R., 1993. *Image Processing, Analysis and Machine Vision*. Chapman & Hall.
- Story, M., Congleton, R.G., 1986. Accuracy assessment: a users perspective. *Photogramm. Eng. Remote Sens.* 52 (3), 397–399.
- Veregin, H., 1989. Error modelling for the map overlay operation. In: Goodchild, M.F., Gopal, S. (Eds.), *Accuracy of Spatial Databases*. Taylor and Francis, pp. 3–19.
- Veregin, H., 1995. Developing and testing of an error propagation model for GIS overlay operations. *Int. J. Geogr. Inf. Syst.* 9 (6), 595–619.
- Waterfeld, W., Schek, H.-J., 1992. The DASBDS Geokernel — an extensible database system for GIS. In: Turner, A.K. (Ed.), *Three-Dimensional Modelling with Geoscientific Information Systems*. Kluwer Academic Publishing, Netherlands, pp. 45–55.
- Worboys, M., 1998. Computation with imprecise geospatial data. *Comput., Environ. Urban Syst.* 22 (2), 85–106.