Title
Unit 187 - Managing Uncertainty in GIS

Permalink
https://escholarship.org/uc/item/31w369dp

Authors
187, CC in GIScience
Hunter, Gary J.

Publication Date
2000

Peer reviewed
Unit 187 - Managing Uncertainty in GIS

Written and edited by Gary J Hunter, Department of Geomatics, The University of Melbourne, Parkville, Victoria, Australia. Peer reviewed by Kate Beard, Peter Fisher, Gerard Heuvelink, and Howard Veregin.

This unit is part of the NCGIA Core Curriculum in Geographic Information Science. These materials may be used for study, research, and education, but please credit the authors Gary J. Hunter, and the project, NCGIA Core Curriculum in GIScience. All commercial rights reserved. Copyright 1998 by Gary J. Hunter.

Advanced Organizer

Unit Topics

- this unit outlines
  - the issues behind the uncertainty debate
  - strategy for managing uncertainty in GIS
  - approaches to uncertainty reduction and absorption
  - future directions in uncertainty management

Intended Learning Outcomes

- after reading this unit, you should be able to
  - explain why managing uncertainty in GIS has now become a major concern within the geographic information industry
  - describe the key components of a strategy for managing uncertainty
  - discuss the various methods of uncertainty reduction and absorption in geographic information products

Full Table of Contents

Metadata and Revision History
Managing Uncertainty in GIS

1. Introduction

- Increasingly, concern is being expressed within the geographic information industry at our inability to effectively deal with uncertainty and manage the quality of information outputs.

- This concern has resulted from:
  - the requirement in some jurisdictions for mandatory data quality reports when transferring data
  - the need to protect individual and agency reputations, particularly when geographic information is used to support administrative decisions subject to appeal
  - the need to safeguard against possible litigation by those who allege to have suffered harm through the use of products that were of insufficient quality to meet their needs
  - the basic scientific requirement of being able to describe how close their information is to the truth it represents

- We often forget that traditional hardcopy maps contained valuable forms of accuracy statements such as reliability diagrams and estimates of positional error:
  - Although these descriptors were imperfect, they at least represented an attempt by map makers to convey product limitations to map users, however
    - this approach assumed a knowledge on the part of users as to how far the maps could be trusted, and
    - new users of this information are often unaware of the potential traps that can lie in misuse of their data and the associated technology

- The lack of accuracy estimates for digital data has the potential to harm reputations of both individuals and agencies - particularly where administrative decisions are subject to judicial review

- The era of consumer protection also has an impact upon the issue:
  - While we would not think of purchasing a microwave oven or video recorder without an instruction booklet and a warranty against defects, it is still common for organisations to spend thousands of dollars purchasing geographic data without receiving any quality documentation

- Finally, if the collection, manipulation, analysis and presentation of geographic information is to be recognised as a valid field of scientific endeavour, then it is inappropriate that GIS users remain unable to describe how close their information is to the truth it represents

Figure 1. Participants in the uncertainty debate

- The obligation to resolve the issues associated with uncertainty rests equally with data
2. A Strategy for Managing Uncertainty

- In developing a strategy for managing uncertainty (Figure 2) we need to take into consideration the core components of the uncertainty research agenda, viz.
  - developing formal, rigorous models of uncertainty
  - understanding how uncertainty propagates through spatial processing and decision making
  - communicating uncertainty to different levels of users in more meaningful ways
  - designing techniques to assess the fitness for use of geographic information and reducing uncertainty to manageable levels for any given application
  - learning how to make decisions when uncertainty is present in geographic information, i.e. being able to absorb uncertainty and cope with it in our everyday lives

  Figure 2. A strategy for managing uncertainty

- In applying the strategy, consideration is initially given to
  - the type of application
  - the nature of the decision to be made
    - low risk vs high risk
    - non-controversial vs controversial
    - non-political vs political
  - the degree to which system outputs are utilised within the decision making process

- Ideally, this prior knowledge permits an assessment of the final product quality specifications to be made before a project is undertaken, however this may have to be decided later when the level of uncertainty becomes known

- Data, software, hardware and spatial processes are combined to provide the necessary information products

- Assuming that uncertainty in a product is able to be detected and modeled, the next consideration is how the various uncertainties may best be communicated to the user

- Finally, the user must decide what product quality is acceptable for the application and whether the uncertainty present is appropriate for the given task. There are two choices available here
  - either reject the product as unsuitable and select uncertainty reduction techniques to create a more accurate product
  - or, absorb (accept) the uncertainty present and use the product for its intended purpose
3. Determining Product Quality

- In determining the significant forms of uncertainty in a product, trade-offs in one area may have to be made at the expense of others to achieve better accuracy, for example:
  - it may be acceptable that adjacent objects - such as a road, a railway and a river - are shown in their correct relative positions even though they have been displaced for cartographic enhancement
  - attribute accuracy or logical consistency may be the dominant considerations, such as in emergency dispatch applications where missing street addresses or non-intersecting road segments can cost lives

- A well-known case described by Blakemore (1985) provides a good example of the lack of understanding of geographic data accuracy requirements:
  - A file of boundary coordinates for British administrative districts was collected for a government department as a background for its thematic maps
  - The agency did not require highly accurate positional recording of the boundaries and emphasis was placed more on attribute accuracy
  - However, when the data became extremely popular because of its extent and convenient format, secondary users soon started to experience problems with point-in-polygon searches after some locations '... seemed suddenly to be 2 or 3 km out into the North Sea'
  - Clearly, convenience can sometimes override data quality concerns

- Where the data quality requirements are unknown, it can be a useful exercise to estimate costs for each combination of data collection and conversion technologies and procedures
  - users often take a pragmatic approach to the 'cost vs accuracy' issue when forced do so

- Having decided which forms of uncertainty will have significant impact upon the quality of a product, some form of uncertainty assessment and communication must be undertaken

- Ultimately, however, it is the user who must decide whether the level of uncertainty present in a product is acceptable or not:
  - which implies there is really no such thing as bad data - just data that does not meet a specific need. For example
    - a road centreline product digitised from source material at a scale of 1:25000 would probably have poor positional accuracy for an urban utility manager, yet may be quite acceptable for a marketing agency planning advertising catchments
    - similarly, street addresses associated with the road centreline segments would probably need to be error free for an emergency dispatch system, whereas the marketeer would probably be content if 80% were correct

- So quality relates to the purpose for which the product is intended to be used - the
essence of the 'fitness for use' concept

4. Uncertainty Reduction

- Uncertainty is reduced by acquiring more information and/or improving the quality of the information available
  - however, it needs only to be reduced to a level tolerable to the decision maker

- Methods for reducing uncertainty include
  - defining and standardising technical procedures
  - improving education and training
  - collecting more or different data
  - increasing the spatial/temporal data resolution
  - field checking of observations
  - better data processing methods
  - using better models and improving procedures for model calibration

- A good example of uncertainty reduction in relation to forest resource management is given in Prisley and Smith (1991)

- by developing an understanding of error propagation in the algorithms used to calculate timber volumes and areas, knowledge was gained as to when inventory methods could be improved to reduce uncertainty and, conversely, when they could be relaxed yet still achieve the desired results

5. Uncertainty Absorption

- Sometimes we have to accept it may be either too costly, impossible or impractical to reduce uncertainty any further

- Openshaw (1989) suggests this step might not necessarily require hard quantitative assessments, but instead may involve a more pragmatic approach on the part of users

"... what many applications seem to need is not precise estimates of error but some confidence that the error and uncertainty levels are not so high as to render in doubt the validity of the results in a particular data specific situation"

- The amount of uncertainty absorbed is considered to be the risk associated with using the data

- Sometimes there may be institutional uncertainty absorption applied
  - for example, when a government agency takes responsibility for guaranteeing land title records to be correct
  - or else, when a government agency authorises a particular dataset as being the
'official' version

- Another form of absorption involves limiting the extent to which GIS is used in the decision making process
  - for example, Laws et al. (1989) describe a case study linking land use planning decisions to the uncertainty in the datasets acquired for the task
    - rather than use uncertainty reduction techniques, they analysed the uncertainty in their data and held it as a constraint upon their decision making
      - adopting the attitude they simply had to work with the data available
  - they then examined the types of decisions to be made and determined limits for which the data could be used in each case
    - at the State planning level, they decided the GIS results were appropriate for non-binding advisory and management decisions
    - but for regulatory and land purchasing decisions (subject to judicial challenge) the data were considered suitable only for initial screening to give an indication of areas worthy of more detailed field inspection and evaluation

- A traditional means of absorbing uncertainty on the part of data producers has been to issue legal disclaimers with datasets
  - however the benefit of disclaimers would seem to be diminishing with stronger consumer protection laws
  - many data producers also prepare detailed data quality statements which may help lessen the financial impact of legal action against them

- Self-insurance is a common means for governments to protect themselves against various forms of uncertainty, for example
  - when a government is prepared to pay from general revenues for harm that results from decisions based on negligently prepared maps and charts
  - or when a special fund is developed to cover losses associated with a specific agency activity

- A new technique for absorbing uncertainty is to take private insurance cover as protection against potential liability claims
  - this is attractive to private sector producers of datasets specifically tailored for use by clients in high risk areas, such as emergency dispatch
  - discussions with producers who have taken out such policies reveal the assessment process is laborious, requiring
    - proof of recognised quality assurance accreditation (including all sub-contractors)
    - production of detailed data quality statements (which must be kept up to date)
    - adherence to industry standards and accepted practices
    - placement of copies of the data supplied to customers into locked vaults maintained by independent escrow companies

- Finally, some environmental risk managers prefer to adopt the average or worst case
scenarios when evaluating their data

6. Future Directions

- One vision for the future is the application of 'intelligent' systems to handle uncertainty
  - Burrough (1991) suggests such systems could help decision makers evaluate the consequences of employing different combinations of data, technology, processes and products, to gain an estimate of the uncertainty expected in their analyses before they start
  - Nijkamp and Scholten (1991) suggest such systems should be able to answer questions like
    - "What are the optimum uses of a given set of data inputs?"
    - "What are the optimum data inputs for a given set of uses?"

- Stoms (1987) discusses knowledge-based approaches which employ various methods of reasoning under uncertainty for specific applications
  - He foresees GIS being embedded in decision support systems of the future to provide decision makers with measures of reliability of the evidence set before them, and the conclusions they might reasonably draw from that information

- Elmes and Cai (1992) have investigated incorporation of a data quality module in a decision support system to advise on management of forest pest infestations
  - By examining user needs in an introductory on-line question and answer tutorial, the system would help determine whether those needs can be met by examining the lineage of the data to be used, the spatial processes to be employed, and the nature of the outputs to be provided
  - Users would then be presented with a range of measures to portray the uncertainty of their results, including sensitivity analyses, summary statistics, and the range of values that any pixel may possess at any time in the overall process

- From a different perspective, Beard (1989) has examined the usage of geographic information and suggests that databases might be re-designed to help prevent misuse
  - Systems could be structured so that validity of mathematical operations may be verified before processing, whereby data resolution and type would be automatically assessed to see if they are appropriate for a given operation
  - In circumstances deemed to be possible misuse, users would be given explanatory warnings prior to executing their instructions which, if they choose to override them, would be added as notations to the product lineage report
    - This approach has the advantage of catering for novice users by acting as an educational tool

- Finally, Agumya and Hunter (1996) believe the uncertainty debate must now advance from its present emphasis on the effect of uncertainty in the information, to considering the effect of uncertainty on the decisions which rely on it
  - in other words, users should be asking "How good are their decisions?" rather than "How good is their information?"
they propose a method of assessing fitness for use of geographic information which aims to determine acceptable levels of uncertainty in geographic information by analysing the risks associated with decisions based upon use of that information

7. Summary of Important Points

- There are four key reasons why the uncertainty debate has grown in importance
  - mandatory data quality reporting in some jurisdictions
  - the need to protect reputations
  - as a means of safeguarding against litigation
  - and as part of the basic scientific quest for knowledge

- A strategy for handling uncertainty has been presented, involving
  - consideration of the type of application and the nature of the decision to be made
  - determining the error in the information product and comparing it with the error specifications for the task at hand
  - methods for either reducing uncertainty or else absorbing it
    - the former usually involves technical approaches, while the latter usually involves institutional methods

- Future directions in uncertainty management include
  - 'intelligent' systems
  - knowledge-based approaches to reasoning under terms of uncertainty
  - re-design of systems to prevent misuse
  - using risk management techniques to understand how uncertainty in the information translates into risk in the final decision

8. Reference Materials


---

**Citation**

To reference this material use the appropriate variation of the following format:

Unit 187 - Managing Uncertainty in GIS

Table of Contents

Advanced Organizer
  Unit Topics
  Intended Learning Outcomes
  Metadata and revision history

Body of unit
  1. Introduction
  2. A Strategy for Managing Uncertainty
  3. Determining Product Quality
  4. Uncertainty Reduction
  5. Uncertainty Absorption
  6. Future Directions
  7. Summary of Important Points
  8. Reference Materials

Citation

Back to the Unit
Unit 187 - Managing Uncertainty in GIS

Metadata and Revision History

1. About the main contributors
   - Written and edited by Gary J Hunter, Department of Geomatics, The University of Melbourne, Parkville, Victoria, Australia.
   - Peer reviewed by Kate Beard, Peter Fisher, Gerard Heuvelink, and Howard Veregin.

2. Details about the file
   - unit title
     - Managing Uncertainty in GIS
   - unit key number
     - 187

3. Key words

4. Index words

5. Prerequisite units

6. Subsequent units

7. Other contributors to this unit

8. Revision history
   - Created: February 27, 1997

Back to the Unit.
Figure 1. Participants in the uncertainty debate
Figure 2. A strategy for managing uncertainty