Characterising an Environmental Problem in Terms of its Scale Requirements

Linda Lilburne

Landcare Research
P.O. Box 69, Lincoln, New Zealand
Ph +64 3 325-6700, Fax +64 3 325-2418
Email: LilburneL@Landcare.cri.nz

Presented at SIRC 2000 – The 12th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 10-13th 2000

ABSTRACT

Environmental and natural resource management problems are often resolved or studied by combining spatial data with computerised simulation models. However, scale compatibility of problems, data and models is rarely examined despite the possibility of significant scale effects and consequent risk of making decisions with invalid information. This paper presents a problem classification which organises scale and goodness-of-fit metrics by problem-type and data-type. It is suggested that this will assist a problem-solver critically assess scale compatibility. A case study is used to illustrate how a decision-maker can analyse whether information provided by a soil carbon model is suitable for three hypothetical problems at national, regional and farm scales. It is concluded that the classification encourages a systematic analysis of scale compatibility.

Keywords and phrases: scale, accuracy, precision, decision, fitness of use, environmental problem

1.0 INTRODUCTION

Environmental and natural resource management *problems* are often resolved or studied by combining spatial *data* with *models* (e.g., statistical, empirical, logical, or mechanistic models). Such problems, data, and models are all scale-dependent, where scale is described by *extent*, *accuracy*, and *precision* (Lilburne and Benwell, 2000). For example, models have an area or extent over which model assumptions are valid, data resolution (precision) can alter model accuracy, and a problem may be solvable at one level of spatial detail but not another. GIS practitioners need to assess scale compatibility of data and model with a given problem to ensure the validity of this information for decision-making.

The need to assess scale compatibility is becoming more critical, as the increasing availability of data and models on the Internet makes it easier for users to combine data and models inappropriately in an attempt to solve a specific problem. However, assessing scale compatibility is not easy as scale limitations of data and models are rarely described explicitly, and there is little information on how to approach the assessment.

A procedure called the “Scale Matcher” is being developed that involves defining the scale requirements of a problem, which can then be matched to the scale limitations of data and models (Lilburne and Benwell, 2000). By first classifying these scale requirements, the GIS practitioner can be more objective and systematically critical about scale compatibility. This paper presents a classification for identifying and organising the scale requirement metrics of a spatial problem. A case study illustrates how this classification and associated scale metrics can be used to assess the scale compatibility of one model of soil carbon distribution with three problems, each of different scales.

2.0 BACKGROUND

The difficulties posed by scale dependency of data and models have been highlighted by numerous authors (e.g., Bian (1997), Blöschl (1999), Heuvelink (1999), Wolock and McCabe (2000)). However, few, if any, examples in the literature explicitly and comprehensively state the minimum scale requirements of any information that might be used to solve the problem being considered. The need to do this is recognised in one of the USEPA Data Quality Objective guidelines (1994), in which they recommend that the spatial and temporal scales of a decision be specified in the decision rule. This research addresses this need.

Part of the complexity in assessing scale lies in the ambiguity of the term "scale" itself . Scale describes the level of detail or *precision* in a map, as well as its *extent*. Scale-related *inaccuracy* arises from imprecision (spatial or attribute) at the collection or representation stage where data were not recorded as being imprecise (Lilburne and Benwell, 2000). Another part of the complexity is the way these components interrelate. By categorising the application or problem and by breaking scale into its components, assessing scale compatibility can be made more tractable.

Other related areas of research provide useful insights, for example, research into spatial error propagation, and quality of spatial data. Quality of spatial data comprises the following properties: positional accuracy, attribute accuracy, semantic accuracy, completeness, logical consistency, temporal information and lineage (Guptill and Morrison, 1995). Various standards for measures of these properties have been drafted, e.g. CEN/TC287, NCDCDS (Duckham, 1999), SDTS, ISO/TC 211 (Aalders, 1999). Of these, the accuracy and completeness related measures are useful in assessing scale compatibility.

Data quality focuses on how well a dataset satisfies its associated data specification, whereas the concept of *fitness of use* focuses on how well the data satisfies the requirements of an application (Brassel *et al.*, 1995). Agumya and Hunter (1999), Brimicombe (1997), and Veregin and Hargitai (1995), amongst others, have researched practical aspects of assessing fitness of use. In a similar way, assessing scale compatibility could be called *fitness of scale*. This is not the same as fitness of use as scale is a part of the data specification: both broad- and fine-scale data could be regarded as having very high quality. Fitness of scale is how well the data specification satisfies the requirements of an application or problem.

3.0 CATEGORISING PROBLEMS AND DERIVING SCALE METRICS

3.1 Problem Characterisation

Beard (in Griffin, 1995) classifies GIS applications into six types: siting, logistic, routing, navigation, inventory, monitoring/analysis. This classification has been adapted by simplifying it into four GIS-user tasks: *locate, report, decision,* and *learn* (Table 1).

User tasks	Description
Locate	Locate where something of interest is (e.g., areas of risk, high-yielding areas); determine how to navigate to somewhere
Report	Provide information for monitoring, reporting or inventory purposes
Decision	Choose between options, e.g., routes, sites
Learn	Exploring causal relationships through analysis in order to learn about a phenomenon of interest

Table 1: Types of GIS-user tasks.

Each of these may use a variety of simple or complex transformations and analyses, which may or may not be scale dependent. So this classification does not directly help with identification of scale requirements. However, the underlying problem or question that is being asked in any given GIS application can be typified as *where, what,* *when*, *which*, and *why* problems. This taxonomy is similar to Kepner and Tregoe's description of decisions as being ones of identity (what), location (where), timing (when) and magnitude (how much) (Watson and Buede, 1989, p.165) (Table 2).

Problem type	**Description**
where	information on the spatial position, distribution or variation of a phenomenon of interest
what	summary information, e.g., totals, averages; information on magnitude, (aspatial information)
when	information on the temporal distribution (which may also be spatial)
which	location-allocation problems; optimal site or route
why	exploratory data analysis; determining cause/effect relationships

Table 2: Types of GIS problems

Table 3 relates the problem type to each GIS-user task.

	Locate	**Report**	**Decision**	**Learn**
where	✓	✓	✓	
what		✓	✓	
when	✓	✓	✓	
which			✓	
why				✓

Table 3: Problem type vs GIS-user task

Many spatial problems fall into the *where* category. This category seeks to locate or spatially identify locations or areas of interest. It includes the following variations:

- threshold: a binary classification of land that meets or exceeds some tolerance level
- site suitability or vulnerability: a classification of levels of suitability or vulnerability
- specific query: identification of one or more specified features
- general variability: a map of the general distribution of a phenomenon of interest
- trend: determining a spatial trend.

Note that all *where* and *what* problems may have limiting spatial or aspatial conditions imposed, e.g., information may only be required in a specified county, or in all catchments larger than a specified size, or within all reserves within 10 km of a specific location.

Where and *what* problems may be concerned with nominal, average, total, dominant, quantile, or extreme values. The use of average values usually smooths variability. *Which* problems and some *where/what* problems require relative values rather than absolute values, e.g., vulnerability ranking, site suitability, optimal routes.

Information used to solve *where/what* problems is either quantitative (integer or real number) or qualitative. Quantitative data may be a surface of the value of interest, a surface of probability or fuzzy membership values, or vector data with a quantitative attribute. Qualitative data may be Boolean, ordinal, interval or nominal. The latter includes features (point, line, polygon) which satisfy some categorical identification criteria (e.g., name = "High Street"), as well as classified surfaces (e.g., landcover data derived from satellite imagery). Interval data is where quantitative data is grouped into ranges or intervals (e.g., 0–10, 10–20, etc.).

3.2 Scale Requirement Metrics

Defining the scale requirements of a problem involves considering its minimum extent, accuracy and precision (Lilburne and Benwell, 2000). A *what* problem implicitly defines the level of spatial precision, whereas in a *where* problem, the spatial precision is often undefined. For example, "what is the expected yield of each paddock of Green Farm", or "what is the estimated population of each region of New Zealand", define the spatial precision to be that of a paddock and region respectively. Extent is often a part of the problem statement (i.e., Green Farm and New Zealand in the previous example), but attribute precision, and spatial/attribute accuracy are rarely defined.

Even though *what* problems are aspatial, they are likely to have spatial as well as aspatial scaling issues, especially where spatial operations like overlay, adjacency and connectivity are involved in deriving the output. Trend problems typically have an extent requirement, which for logistical/computation issues drives the spatial precision, e.g., those examining a global trend use a coarse spatial precision of 1x1 degree (or 12,100 sq km at the equator) (DeFries *et al.*, 1997). Spatial and/or temporal precision may also be decreased to remove "noise". Accordingly users interested in trends will often tolerate a lower level of accuracy; they may also consider relative accuracy rather than absolute accuracy, which is usually less demanding, as just the ranking of the results needs to be accurate, not the actual values.

Sub-mapunit (i.e., polygon/pixel) variability can be expected to be more of an issue in problems concerned with extreme values (i.e. min, max, 75, 80 or 90 quantiles) than ones concerned with dominant, average or total values.

The format of accuracy and precision requirements varies according to whether the information supplied to solve the problem is quantitative or qualitative. The problem definition often, but not necessarily, determines the data-type. Note that scale requirements are independent of the data representation (vector vs raster).

Metrics for scale requirements of quantitative data include root mean square error (RMSE), mean error (ME), mean absolute error (MAE), standard error (SE), confidence interval (CI), percent within a given confidence interval, coefficient of determination (R^2), modelling efficiency (EF), coefficient of variability (CoV), PRESS (or GCV; $\mathrm{RGCV} = \mathrm{GCV}^{\frac{1}{2}}$) statistic. Scale requirement metrics for qualitative data include the confusion matrix, percent correctly classified (PCC), user's and producer's error (UCE and PCE), and the kappa statistic (KP). These metrics may be calculated for the whole map or for a subset(s) of it. Some metrics are precision related, others accuracy related, so more than one metric should be used. Note, however, that accuracy and precision are interrelated: outliers will increase measures of both error magnitude (i.e., accuracy) and error spread (which can be treated as the precision of an estimate).

These metrics can be used to specify the scale requirements of a problem. Table 4 lists which metrics are suitable for which problem and data-type combinations.

Type	**Data-type**	**Accuracy metric**	**Precision metric**
where	quantitative	• specified % of data correctly identified within defined tolerance, error band, or confidence interval (spatial and/or attribute) • ME, MAE, SE, RGCV, RMSE < x • EF, $R^2 > x$ • maximum cost function	• tolerance bands: (± range, σ, RMSE, CoV, CI) • located to nearest x metres • mapping unit • temporal unit
where	probability (of classes)	• probability or fuzzy membership is correct within ± y, x% of the time • maximum cost function	• mapping unit • temporal unit • thematic unit

where	qualitative - nominal	• maximum cost function
		• specified maximum acceptable proportion of land incorrectly identified or minimum correctly identified & located (i.e., PCC, KP, or producer/user classification error limits for specific categories). Proportions might be weighted by perceived importance. $\Sigma(acc + w(100-acc))/n$
			• mapping unit
			• temporal unit
			• class or thematic unit
			  • class intervals are sufficiently narrow
			  • categories are at a sufficient taxonomic level
			• located to nearest x metres

where	qualitative - object	• x % of objects correctly identified	• object located to nearest x metres • x % of object correctly located (polygon) • mapping unit • temporal unit
where	qualitative - ordinal	• order or rank is maintained (rank correlation) • proportion correctly ranked > x%	
where	qualitative - dominant	• dominant portion of mapunit is correctly identified x% of the time	
what	quantitative	• x% accurate within attribute precision limits • ME, MAE < x (for averages)	• spatial aggregation unit • temporal unit • ± range

what	qualitative	• x% correctly identified & described
		• y% incorrectly, not identified
		• PCC, UCE, PCE, KP > x
			• spatial aggregation unit
			• temporal unit
			• Class or thematic unit
			  • class ranges are of sufficient numerical precision
			  • categories are at a sufficient taxonomic level

why		• minimum goodness-of-fit test • magnitude of relationship is not changed • order of importance of key variables is not changed • goodness of fit of relationship is not substantially changed	• mapping unit • temporal unit
which	line or route area	• choice is optimum or near optimum x% of time	• located to nearest x metres • epsilon band width • x% of area correctly located NB schematic output would have no spatial precision requirement

Table 4: Accuracy/precision metrics.

If the information used to solve a problem is derived using spatial relationships, some additional metrics may be needed in order to attain a specified output accuracy requirement. For example, in a *which* problem that requires an optimal route to be identified, connectivity of the route segments is essential. Any problem dependent on spatial relations such as spatial adjacency, connectivity or co-incidence will need to check the likelihood and impact of inaccuracy in any relevant spatial relationships.

Spatial precision requirements may be relaxed where visual clues are available to help the person on the ground identify the required location. For example, spatial precision requirements for locating underground pipes or cables may be very narrow (e.g., ± 5 cm), whereas, the location of an overhead power line may only need to be spatially precise to the nearest 20 m.

It is difficult to define accuracy requirements for *why* problems as a causal relationship is supported by evidence rather than proved. Consequently accuracy is difficult to ascertain. A goodness-of-fit test merely indicates the potential reliability of any results. A number of authors believe that analyses should be done at multiple scales, e.g., Gardner (1998), and Goodchild and Quattrochi (1997). If a multi-scale analysis shows results are sensitive to the scale of inputs, then analysis results may be unreliable. A scale requirement that the magnitude of a relationship, or that the order of importance of key variables, is not changed when inputs are perturbed within their defined precision/accuracy limits can be used to identify unreliable relationships.

4.0 CASE STUDIES

Several hypothetical problems based on soil data are described in this section to illustrate the scale requirement metrics. Problem A belongs to an analyst charged with calculating the expected carbon content in soils throughout New Zealand for international reporting purposes. Problem B is that of the regional policy analyst interested in maintaining soil quality in the MacKenzie Country of the South Island. Finally problem C belongs to a landowner wanting to improve the productive capacity of his or her sheep station.

Lynn *et al.* (submitted) describe testing a model that was developed to predict seven key soil properties, including carbon content, in the dry steeplands of New Zealand. Output from this model can potentially be used to answer each of the three problems, **if** the output's scale limitations match the problem's scale requirements. Test results from soil samples are used to estimate model accuracy.

4.1 Problem A - Carbon Inventory

The analyst is interested in the total value of soil carbon, i.e., problem-type = *what*, data-type = quantitative.

	Problem scale requirement	**Information scale limitation**	**Comment**
Extent:	New Zealand	Dry steeplands of the South Island	Information from the model will only satisfy a subset of the problem requirement.
Spatial precision:	mapping unit = whole of NZ	100-m pixels. Attribute value is assumed to be mean of pixel.	Information has more detail but this can be easily aggregated to a mapunit covering the whole of the South Island dry steeplands.
Attribute precision:	± 10% equates to approx. ± 2,000,000 tonnes for dry steeplands.		Under the central limit theorem, the sample data gives a 95% CI = [19.2, 25.1 million] (i.e. ± 2,950,000). However, the sample is not random.
Spatial accuracy:	n/a		
Attribute accuracy:	100%, i.e., the single output value is believed to be within the specified attribute precision.	ME(samples) = 3.65 t/ha. Therefore model is predicted to underestimate total carbon by ≈1,720,000 tonnes	

Table 5: Problem A - carbon inventory metrics

It is difficult to estimate a confidence interval because the sample data are not random, the support of the samples (10 m) is not equal to the support of the model (100 m), pixel values are not estimates of the average carbon content in the pixel, and the spatial correlation structure of the error is unknown. Scatter plots show that model residuals are not correlated with model inputs, therefore, it is assumed that the mean model error (ME) can be applied to the whole of the dry steeplands area. This indicates that the model will underestimate the total amount of carbon by 1.7 million tonnes, which is still within the attribute precision specified by the analyst.

4.2 Problem B - Soil Quality

The analyst is interested in identifying those areas that are at risk of soil degradation. Soil experts have determined that the threshold value of soil carbon below which soil can be considered degraded is 2.5%. Consequently the analyst wishes to locate areas in which soils are likely to be < 2% (seriously degraded), and areas with carbon levels between 2% and 3% (of concern). The problem-type is *where*, and the data-type = qualitative (three classes: seriously degraded, of concern, and not degraded).

	Problem scale requirement	**Information scale limitation**	**Comment**
Extent:	Dry steeplands in the MacKenzie Country	Dry steeplands of the South Island	Information satisfies problem requirement.
Spatial precision:	hillslope size	100-m pixels	Information has more detail but contiguous blocks of interest of sufficient dimensions can be easily identified
Attribute precision:	Three classes (0–2%, 2–3%, 3+%)	quantitative with a numerical precision of approx ± 0.1%	Quantitative data can be easily classified into the three required classes
Spatial accuracy:	areas approximately located to nearest 200 m	100-m pixels	Above-ground visual indicators (e.g., vegetation vs bare ground) can help identify land once in the field, therefore the spatial accuracy requirement is quite coarse.
Attribute accuracy:	max. of 20% of degraded/of concern area incorrectly classified as non-degraded; max of 40% of non-degraded area incorrectly identified as being degraded/of concern.	quantitative with an average numerical accuracy (ME) of approx ± 1%. Classified error rates are: PCC = 59%, UCE(degraded+of concern)= 28%, UCE(non-degraded) = 32% (Lynn *et al.*, submitted)	As the analyst is interested in monitoring potentially degraded areas, s/he is less concerned with non-degraded areas identified as being degraded as monitoring will reveal their true status, whereas areas incorrectly identified as being degraded are not monitored at all. PCC was higher than the accuracy of the quantitative data suggests, as larger errors tended to be on non-degraded land and these errors did not often result in a change to the classification.

Table 6: Problem B - soil quality metrics

The data meet the analyst’s spatial requirements but are not sufficiently accurate (due to model error) to meet the attribute accuracy requirement. However, an analyst may reluctantly opt to relax the requirement of accurately identifying degraded or “of concern” land to 30%, as information from a finer scale is simply not available. Another option is to change the classes so that they are more conservative (e.g., 0–2%, 2–4%, 4+%). This gives a PCC of 63%, UCE (degraded + of concern) = 10%, but UCE(non-degraded) = 52%. This equates to unnecessarily monitoring some land that has 52% probability of being of no risk at all. The analyst will have to choose between the original requirement in which a significant amount of degraded land may be missed, and a lower probability of missing potentially degraded sites but with the additional cost of monitoring land in the 3–4% range.

4.3 Problem C - Increasing Productivity

The landowner wishes to locate land that is suitable for aerial oversowing and topdressing (OSTD). This has been modelled as a combination of suitable levels of pH, P retention, elevation and carbon (Lynn *et al.*, submitted). The problem-type is *where*, and the data-type = qualitative (two classes: suitable and unsuitable).

	Problem scale requirement	**Information scale limitation**	**Comment**
Extent:	Glencairn Station	Dry steeplands of the South Island	Information satisfies problem requirement.
Spatial precision:	min. OSTD area is approx. 20 ha. Width of aerial application (single strip) = 20 m	100-m pixels	Insufficient spatial precision but within precision of management blocks
Attribute precision:	1 class (Suitable)	2 classes (suitable and unsuitable)	Level of classification is appropriate
Spatial accuracy:	High accuracy (due to use of GPS); however, this is dependent on roughness of terrain	100-m pixels	Visual indicators can be used for better spatial accuracy (e.g., vegetation, ridge lines)
Attribute accuracy:	High probability that OSTD results in a profit	UCE (suitable)=29%, UCE (unsuitable) = 39%.	Effect of incorrect OSTD data can be estimated from test soil samples (Lynn *et al.*, submitted) which are assumed to be representative. High level of inaccuracy results in a marginal investment.

Table 7: Problem C - land productivity metrics

Tither (1992) provides an analysis of the benefits over 30 years of developing Longslip Station. It is assumed that these figures are based on a careful and conservative identification of suitable land (i.e., UCE(suitable) = 0%). If 29% of this land were in fact unsuitable, so the new seedlings did not grow, then the benefit would be reduced by approximately 29%, but the costs would remain the same. This would result in a benefit over 30 years of approximately $420,000 based on 1992 figures. According to Tither (1992) , this is marginal compared with putting the money in the bank. Considering that Glencairn Station is in a drier area than Longslip Station and therefore the benefit is likely to be lower (Floate and Cossens, 1992), this makes the benefit of OSTD based on the inaccurate suitability information even more dubious. However, if the UCE(suitable) can be reduced through use of additional information, or the UCE(suitable) for Longslip were in fact comparable with that of Glencairn Station, then the data may be of sufficient scale, especially with the recent high wool prices.

Global positioning systems (GPS) are now used in planes to ensure accurate placement of OSTD. This results in efficient and accurate application of OSTD, thus minimising costs (B. Aubrey, Glencairn Station, pers. comm.) although the precision of application is very dependent on the roughness of the terrain and the constraints this places on flight paths (I. Lynn, Landcare Research, pers. comm.). However, there is little advantage in OSTD small areas if these areas are not managed separately, as stock will graze improved areas within a block and ignore the unimproved remainder. At a resolution of 100 m (i.e., 1 ha) the information is less precise than that obtained by precision application but is well within the precision of the management blocks (30–150 ha).

The information is insufficiently accurate to be used *per se;* however, it may be of use to help land managers identify blocks *potentially* suitable for OSTD, and areas which it *may* be appropriate to fence so that they can be grazed independently.

Thus the scale of the available soil information is marginally acceptable for problems A and B (after modification of the problem scale requirement) but is not suitable (without additional verification by the landowner) for problem C.

5.0 DISCUSSION AND CONCLUSION

There are many different metrics which each measure different aspects of scale, i.e., extent, accuracy and precision. The metrics identified in this paper are ones that are frequently cited in the literature. There are other metrics, especially measures of goodness of fit, which are not widely used or understood, for example, the index of agreement (Willmott, 1984), coefficient of residual mass (Loague and Green, 1991), and model performance (Kumar and Heatwole, 1995). The coefficient of determination R^2 and correlation ρ, although widely used, are in fact unsuitable for making comparisons between observed and predicted data (Altman and Bland, 1983; Willmott, 1984). Further research is needed to identify a useful and comprehensive set of metrics, and their advantages/disadvantages. The metrics in table 4 also need to be verified against a wider range of problems.

While it may be possible to specify a minimum accuracy requirement of the problem, it is not always possible to obtain an estimate of the accuracy of the available information. Accuracy is the closeness of an prediction to its true value, therefore, the only way to estimate accuracy is to have access to ground-truth data (Heuvelink, 1998). But ground-truth data may be too expensive or logistically difficult to collect (e.g. pesticide leachate in soil), or it may be a subjective value (e.g., vulnerability) which is not directly measurable. Even if ground-truth data is measurable, it may not be practical for the sample support (i.e., area of sample) to be the same as that of the model. In these cases, estimates of confidence in the information will have to be used rather than measures of accuracy.

The classification presented in this paper organises scale metrics according to problem-type (where, what, why, which and when) and information data-type (quantitative and qualitative). This helps a user to identify appropriate metrics for defining the scale requirements of a given problem. The case study of a model of soil carbon showed how these requirements could then be matched with data and model scale limitations. It is believed that this systematic approach enables a user to gain a better understanding of scale effects, and consequently to recognise when information is unsuitable for use in decision-making. More research will be undertaken to integrate these results with other work on the Scale Matcher (Lilburne and Benwell, 2000; Lilburne *et al.*, 1999).

ACKNOWLEDGEMENTS

This work was supported by non-specific output funding allocated to Landcare Research Ltd by the Foundation for Research, Science and Technology. James Barringer, George Benwell, and Ian Lynn kindly reviewed it, and Christine Bezar edited it. Joanna Orwin’s advice in how to improve its clarity is much appreciated.

REFERENCES

Aalders, H. J. G. L. (1999) The registration of quality in a GIS. *Proceedings of International symposium on spatial data quality*, Hong Kong, pp. 23-32.

Agumya, A. and G. J. Hunter. (1999) Assessing "fitness for use" of geopgraphic information: what risk are we prepared to accept in our decisions? *In* Spatial accuracy assessment: Land information uncertainty in natural resources, K. Lowell and A. Jaton, eds., Ann Arbor Press, Michigan, pp. 35-43.

Altman, D. G. and J. M. Bland. (1983) Measurement in medicine: the analysis of method comparison studies. *The Statistician*, 32, pp. 307-317.

Bian, L. (1997) Multiscale nature of spatial data in scaling up environmental models. *In* Scale in remote sensing and GIS, D. A. Quattrochi and M. F. Goodchild, eds., Lewis Publishers, pp. 13-26.

Blöschl, G. (1999) Scaling issues in snow hydrology. *Hydrological processes*, 13, pp. 2149-2175.

Brassel, K., F. Bucher, E.-M. Stephan and A. Vckovski. (1995) Completeness. *In* Elements of spatial data quality, S. C. Guptill and J. L. Morrison, eds., Elsevier, pp. 81-107.

Brimicombe, A. (1997) A universal translator of linguistic hedges for the handling of uncertainty and fitness-for-use in GIS. *In* Innovations in GIS 4, Z. Kemp, ed., Taylor & Francis, pp. 115-126.

DeFries, R. S., J. R. Townshend and S. O. Los. (1997) Scaling land cover heterogeneity for global atmosphere-biosphere models. *In* Scale in remote sensing and GIS, D. A. Quattrochi and M. F. Goodchild, eds., Lewis Publishers, pp. 231-246.

Duckham, M. (1999) Implementing an object-oriented error-sensitive GIS. *In* Spatial accuracy assessment: Land information uncertainty in natural resources, K. Lowell and A. Jaton, eds., Ann Arbor Press, Michigan, pp. 209-215.

Floate, M. and G. Cossens. (1992) Land resources. *In* Guide to tussock grassland farming, M. Floate, ed., AgResearch, Invermay, pp. 23-37.

Gardner, R. H. (1998) Pattern, process, and the analysis of spatial scales. *In* Ecological scale : theory and applications, D. L. Peterson and V. T. Parker, eds., Columbia University Press, New York.

Goodchild, M. F. and D. A. Quattrochi. (1997) Introduction: Scale, multiscaling, remote sensing and GIS. *In* Scale in remote sensing and GIS, D. A. Quattrochi and M. F. Goodchild, eds., Lewis Publishers, pp. 1-11.

Griffin, C. B. (1995) Data quality issues affecting GIS use for environmental problem-solving. *Proceedings of Environmental problem solving with geographic information systems*.

Guptill, S. C. and J. L. Morrison. (1995) Elements of spatial data quality, Elsevier.

Heuvelink, G. B. M. (1998) Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient cycling in agroecosystems*, 50, pp. 255-264.

Heuvelink, G. B. M. (1999) Aggregation and error propagation in GIS. *In* Spatial accuracy assessment: Land information uncertainty in natural resources, K. Lowell and A. Jaton, eds., Ann Arbor Press, Michigan, pp. 219-225.

Kumar, D. and C. D. Heatwole. (1995) Probabilistic procedures for evaluating nonpoint source pollution models. *Proceedings of International symposium on water quality modeling*, Orlando, Florida, pp. 164-176.

Lilburne, L. R. and G. Benwell. (2000) The Scale Matcher: determining scale compatibility of environmental data and models. *Proceedings of Accuracy 2000*, Amsterdam.

Lilburne, L. R., A. E. Hewitt, I. Lynn and G. L. Benwell. (1999) Some scale issues in spatial modelling of soil properties. *Proceedings of Third conference of the working group on pedometrics of the international union of soil science*, Sydney, pp. 53-55.

Loague, K. and R. E. Green. (1991) Statistical and graphical methods for evaluating solute tranpsort models: overview and application. *Journal of contaminant hydrology*, 7, pp. 51-73.

Lynn, I. H., L. R. Lilburne and P. D. McIntosh. (submitted) Testing a soil-landscape model for dry greywacke steeplands on three mountain ranges in the South Island, New Zealand. *Australian journal of soil research*.

Tither, P. (1992) Economic decision making - how to evaluate and plan sustainable high country development. *In* Guide to tussock grassland farming, M. Floate, ed., AgResearch, Invermay, pp. 111-115.

USEPA. (1994) Guidance for the data quality objectives process. EPA/600/R-96/055, EPA, Washington DC.

Veregin, H. and P. Hargitai. (1995) An evaluation matrix for geographical data quality. *In* Elements of spatial data quality, S. C. Guptill and J. L. Morrison, eds., Elsevier, pp. 167-188.

Watson, S. R. and D. M. Buede. (1989) Decision synthesis. The principles and practice of decision analysis, Cambridge University Press.

Willmott, C. J. (1984) On the evaluation of model performance in physical geography. *In* Spatial statistics and models, G. L. Gaile and C. J. Willmott, eds., D. Reidel Publishing Company, pp. 443-460.

Wolock, D. M. and G. J. McCabe. (2000) Differences in topographic characteristics computed from 100- and 1000-m resolution digital elevation model data. *Hydrological processes*, 14, pp. 987-1002.