# Bayesian Spatial Uncertainty Analysis

S. ZIMERAS[1] AND Y G. MATSINOS[2]

[1]Department of Statistics and Actuarial Science, University of the Aegean, Karlovassi, 832 00 Samos, Greece, E-mail: zimste@aegean.gr
[2]Biodiversity Conservation Laboratory, Department of Environmental Studies,University of the Aegean, 81100 Mytilene, Greece, e-mail: matsinos@aegean.gr

**Abstract:** The term uncertainty originates from imperfect knowledge of the processes under question. These errors are sometimes associated with questionable data quality and scarcity; complexity of the phenomena that are treated by models as simplified systems. Important components of uncertainty analysis include (i) qualitative analysis that identifies the uncertainties, (ii) quantitative analysis of the effects of the uncertainties on the decision process, and (iii) communication of the uncertainty (Funtowwicz and Ravetz 1990, Petersen 2000, Regan et a1.2002, Katz 2002). In landscape simulations (which are widely used in landscape ecology) a generation of landscape patterns is taking place for the investigation of the local or global connectivity between regions. Neighbourhood structures are used for that purpose to explain and analyse the spatial connectivity between smaller to bigger regions including investigation of the spatial homogeneity. Modelling of that spatiality involves conditional probabilities which are explained by Markov random fields models. Estimation of these particular models could be introduce a Bayesian statistics analysis based on these conditional probabilities which are explained the spatial variability into the regions especially when hidden information's are involved (Zimeras and Matsinos, 2011). In this work a spatial analysis methodology based on Bayesian analysis was introduced and procedures to solve the problem with spatial variability are described based on point estimation techniques.

Keywords: spatial modelling, uncertainty, Markov random fields, MCMC, Spatial point analysis, hierarchical models. Bayesian statistics.

## INTRODUCTION

Model evaluation assumes a certain general structure (e.g. multiple linear) and the model is built through adding terms (variables) which are significant or which aid in prediction (hierarchical modelling). Parameter uncertainty is defined as a problem of estimation. These models are similar to models common to risk analysis, which often leads to deterministic approaches for the evaluation of the parameters. Stochasticity is often introduced through stochastic functions (e.g. weather) or random effects in parameter values.

Although the uncertainty analysis problem is similar to statistical problems, in application the uncertainty analysis problem is often more complex than many statistical problems. Models are used due to the complexity of the problem. The choice of the model is a choice that leads to structural uncertainty. Uncertainty affects every aspect of modelling (Jansen 1998, Katz 2002, Stott and Kettleborough 2002). Data may contain errors that result from problems with sampling, measurement, or estimation procedures (O'Neill and Gardner 1979, Regan et al. 2002). Incomplete data are a common problem, especially in spatial modelling. (Reckhow 1994, Clark et al. 2001, Rypdal and Winiwarter 2001, Katz 2002).

Uncertainties arise from measurement errors. These errors are results of models when hidden information's are appears especially when repeated measures are involved as result of spatial variability of the landscape simulations. Uncertainty techniques should be used to decrease the variability effect from observations. Outliers are the result of data inhomogeneity and robust procedures must be applied to deal with that problem.

## SPATIAL UNCERTIANTY ANALYSIS

Uncertainty can become a major problem analyzing incomplete data especially in spatial modelling at different spatial and temporal scales. Models are sometimes imperfect and incomplete, especially in grain scaling data where biotic or other information of large regions need to be predicted by smaller ones. Uncertainty analysis in that perspective is the process of assessing uncertainty in modelling or scaling to identify its major sources, quantify their degree and relative importance, examine their effects on model output under different scenarios, and determine prediction accuracy. Involvement of modelling introduces statistical errors (especially in spatial cases where homogeneity of the neighbours is one of the measures to justify differences

ISBN: 978-1-61804-110-4

between regions). For that reason (incomplete data, spatial differences and statistical modelling) analysis of uncertainties in regions is a complex problem. Although the uncertainty analysis problem can be thought of in an analogous way to usual statistical problems, in application the uncertainty analysis problem is often more complex than many statistical problems. Thus specific models are used due to the complexity of the problem. The choice of the model is a choice that leads to structural uncertainty. Uncertainty affects every aspect of modelling (Jansen 1998, Katz 2002, Stott and Kettleborough 2002). Data may contain errors that result from problems with sampling, measurement, or estimation procedures (O'Neill and Gardner 1979, Regan et al. 2002). Incomplete data are also a common problem, especially in spatial modelling. (Reckhow 1994, Clark et al. 2001, Rypdal and Winiwarter 2001, Katz 2002).

Uncertainties in ecological modelling arise from either systematic or random errors. Systematic errors in measurements are mainly caused by imperfect calibration of measurement equipment, technically caused limits in data quality/sampling frequency (coarse grained vs. fine grained satellite data), limits in capacity (CPU power, hard disk space, memory) or data type (raster vs. vector, wrong classification). Random errors are mainly caused by unpredictable situations, natural and spatial variation, and appear when measurements are repeated. The random errors are similar to the standard deviations. The variation of the reported models can be explained by both systematic and random errors. Uncertainty in models can be divided in a similar way by statistical and systematic uncertainty. The statistical uncertainties arise from the variability of input variables and parameters where the variability is known. This variability can be described by probability density functions (PDFs) describing the variability of the input variables and the parameters. Systematic uncertainties arise from variability in input variables and parameters when variability is unknown. Also unknown processes in the model (e.g. incorrect model structure) contribute to the systematic uncertainties. Examples of different models of uncertainty are shown in Figure 1.

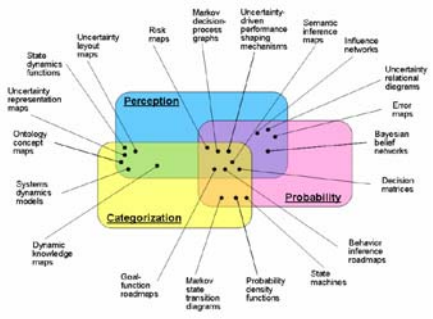


Fig. 1. Different models of uncertainty resulting from applications of probability theory, category theory, and perception theory (from Linkov et al. 2006).

The issue of uncertainty is of particular importance dealing with scaling in ecology; the extrapolation of information from one scale to another in time or space or both (Sivapalan 1995, Wu 1999). In the process of scaling, errors in data and models are critical for the results. Thus, uncertainty analysis is an essential part of scaling analysis because it provides critical information about the accuracy of scaling results (Katz 2002).

Landscape simulators have been widely applied in landscape ecology for generating landscape patterns. These models can be divided into two groups: mathematical (or pattern-based), which generate spatial patterns irrespective of the processes that shape these patterns, or models that attempt to generate patterns by mimicking the processes that shape them. A mathematical or pattern-based approach is used based on mathematical algorithms to generate patterns, regardless of the underlying processes. The simplest approach is the production of simple random maps ("neutral models"). These models lack any systematic processes that may structure a landscape. Even so, a remarkable degree of change in pattern occurs as p, the proportion of land-cover, changes. Another way to simulate landscape patterns is to use conditional probabilities models, based on exponential family models (Zimeras and Matsinos, 2011), defined as Markov random fields models based on the neighbourhood structure of the lattice system (Figure 1).

ISBN: 978-1-61804-110-4

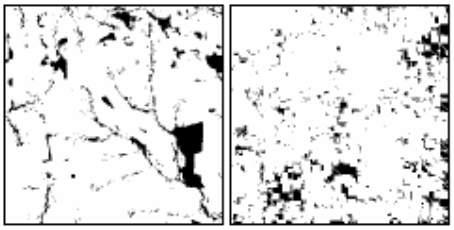

Fig 1. Landscape simulations

Interactions between regions at different scales are characterized by their local dynamics and the emergent spatial patterns are the outcome of different processes. Spatial process is focused on the information extraction where useful information about spatial pattern could be used in order to answer questions regarding the underlying processes (e.g. competition) (Turner, 1989; Wiens, 1989). However, the sometimes hidden spatial dependence in data can lead to violations of the spatial homogeneity assumptions (Legendre, 1993). Uncertainty in models can be divided in a similar way by statistical and systematic uncertainty. The statistical uncertainties arise from the variability of input variables and parameters where the variability is known. This variability can be described by probability density functions (PDFs) describing the variability of the input variables and the parameters. Systematic uncertainties arise from variability in input variables and parameters when variability is unknown. Also unknown processes in the model e.g. incorrect model structure contribute to the systematic uncertainties.

Scaling is the extrapolation of information from one scale to another in time or space or both (Sivapalan 1995, Wu 1999). In the process of scaling, errors in data and models may be effects the results. Thus, uncertainty analysis is an essential part of scaling because it provides critical information about the accuracy of scaling results (Katz 2002).

A methodology is introduced by considering grain scaling data. In landscape ecology, grain describes the size of the smallest homogeneous unit of study and determines the resolution at which a landscape is studied. It is equivalent to minimum mapping unit in cartography. A pixel in a digital image is analogous to a grain in a landscape. The choice of scale is important in determining the resolution of a study and constrains any inferences or extrapolations from the study. In studies of landscape, no pattern can be detected beyond the extent or below the grain of the study. For spatial data or images, spatial grain size corresponds to maximum resolution ratio or pixel size (Figure 2).

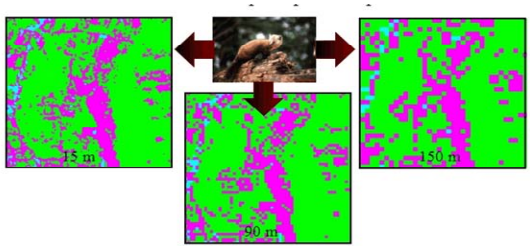


Fig. 2. Grain data based on different scaling

Model uncertainty has two basic components, model structure and model parameters (O'Neill and Gardner 1979, Jansen 1998, Katz 2002). Model structure uncertainty is caused by the modelling processes of simplification and formulation. Model simplification is essential to modelling and is the identification and selection of the relationships, and variables that are the most important to the modelling analysis.

This kind of uncertainty is fully attributed to the use of models (mathematical, statistical and simulation) that are developed in order to represent an ecological system. One possible way of introducing model uncertainty is the exclusion of processes thought not to be as important as the key model processes. This is quite often the case of the conceptual models. Another way has to do with the mathematical or statistical representations of variables that sometimes are hard to assess. For example the choice of a discrete or a continuous modelling approach depends on the nature of the variables. The only way to deal with this uncertainty type is model validation and verification, apart from analytical model investigation.

Model evaluation assumes a certain general structure (e.g. multiple linear) and the model is built through adding terms (variables) which are significant or which aid in prediction (hierarchical modelling). Parameter uncertainty is defined as a problem of estimation. These models are similar to models common to risk analysis, which often leads to deterministic approaches for the evaluation of the parameters. Stochasticity is often introduced through stochastic functions (e.g. weather) or random effects in parameter values. Grain scaling data introduces a pyramid modification where large regions are attempted to be modelled by smaller ones. In that point of view connectivity between different levels in the pyramid must be introduced leading us to spatial connectivity between levels' regions. Spatial connectivity includes neighbourhood structure between regions, where investigation of that structure includes modelling of the spatial homogeneity. The last could be

ISBN: 978-1-61804-110-4

illustrated by using spatial modelling techniques (like spatial autocorrelation, partition functions, and multilevel statistical models). Spatial investigation involves stochastic modelling especially in cases where the incomplete data involves hidden information. Stochasticity also could be introduced as departing from the random nature of data. So incompleteness (hidden information) leads us to the use of stochasticity where incomplete information could be explained (and modelled) by adding particular functions, which explain a large amount of information introducing a new modelling (Morgan and Hemion 1990, Klepper 1997, Katz 2002). Uncertainty analysis requires that statistical distributions of parameters be known. However, a common problem in uncertainty analysis is that the accuracy of measurements and/or estimates of parameters are unknown (O'Neill and Gardner 1979, Jansen 1998, Regan et al. 2002).

The issue of uncertainty is of particular importance dealing with scaling in ecology; the extrapolation of information from one scale to another in time or space or both (Sivapalan 1995, Wu 1999). In the process of scaling, errors in data and models are critical for the results. Thus, uncertainty analysis is an essential part of scaling analysis because it provides critical information about the accuracy of scaling results (Katz 2002).

Analyzing uncertainty with probabilistic models is thus important. The simplest situation is represented by most 1-dimensional (1D) models in which the distributions are taken to represent variability, and where there are adequate data to characterize the distributions. More complicated situations may involve the inclusion of uncertainties in 2-dimensional (2D) models. For distributions that represent variability, initial decisions may relate to the selection of data on which to base distributions. The data are a random sample from the populations of interest. In addition, data should represent a spatiotemporal scale appropriate for the model.

Variability in ecology systems may result from spatial heterogeneity of conditions or from randomness in interactions of different processes. Natural variability of data is a critical factor in uncertainty analysis because modelling only with average values can produce bias in predictions, especially for nonlinear models (O'Neill 1979, Scherm and van Bruggen 1994). Spatial variability in systems attributes and driving variables need to be effectively incorporated into simulation modelling. The variability in the variables is then used as a measure of the output uncertainty. Many techniques can be used to analyze uncertainties in model parameters and input data. These include applications of probability theory, Taylor series expansion, Monte Carlo simulation, generalized likelihood uncertainty estimation, Bayesian statistics, and sequential partitioning (Gardner and O'Neill 1983, Gardner et al. 1990, Rastetter et al. 1992, Heuvelink 1998b, Jansen 1998, Wiwatenadate and Claycamp 2000, Katz 2002).

## Bayesian analysis of uncertainty

Bayesian statistical methods quantify uncertainty by calculating probabilistic predictions. The procedure has three stages: (1) determination of the prior probability distribution for model parameters, (2) construction of a likelihood function for the statistical model, and (3) derivation of the posterior probability distribution for the parameters by using the Bayes rule to adjust the prior distribution based on the observed data (Katz 2002). Based on Bayesian analysis, uncertainty is modeled based on the randomness of the data, which can be explained by distribution of the measurements. These measurements could be used (combined with probability modeling) to estimate the amount of the uncertainty effects of the measurements. Statistical estimation techniques can be used to determine the parameters of the distribution. These techniques are useful to estimate probability distributions from available data or by collecting a large amount of theme (Figure 3) (Jansen et. al., 1998)

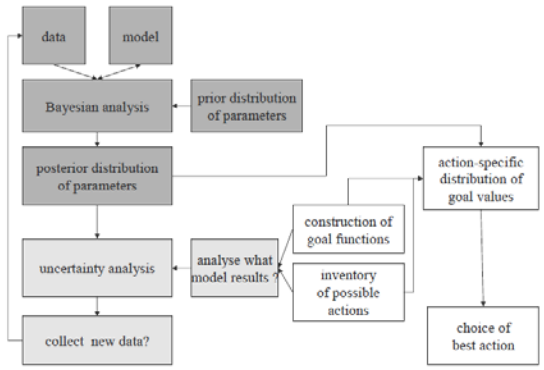


Fig. 3. Bayesian statistic uncertainty analysis (Jansen et. al., 1998)

Traditional hypothesis testing is based on calculating the probability of observing specific data (**Y**) given the null hypothesis, $\theta_{null}$, that is, $P(\mathbf{Y} < \mathbf{Y} \mid \theta_{null})$

So main interest is the likelihood ($L$) of competing hypotheses given the data, which is proportionate to the probability of the data given the hypothesis

ISBN: 978-1-61804-110-4

$$\mathcal{L}(\theta \mid \mathbf{Y}) \propto P(\mathbf{Y} \mid \theta).$$

where θ represents a set of parameter values specifying a particular model or models and that we are interested in the probability of observing the data **Y** given those parameter values.

Ecological models make predictions about states and processes of interest as functions of parameters (θ). For a single observation, the likelihood of the prediction of a model is proportional to the probability of that observation conditional on the model's parameters. More formally, assuming we have a model *f* (θ) that makes predictions on a variable of interest, for example, population density. We have a data set **Y** composed of *n* individual observations on that variable, $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$. Likelihood of the value of the parameter (or parameters) in our model given a single observation $y_i$ is

$$\mathcal{L}(\theta \mid y_i = g(y_i \mid \theta)$$

where the function $g(.)$ is a probability function (if the $y_i$ are discrete) or a probability density function (if the $y_i$ are continuous). If it is assumed that the deviations are independent of one another, then

$$\mathcal{L}(\theta \mid \mathbf{Y}) = \prod_{i=1}^{n} g(y_i \mid \theta).$$

For computational simplicity, it is often more useful to maximize the log likelihoods, in which case we have the following:

$$\ln[\mathcal{L}(\theta \mid \mathbf{Y})] = \sum_{i=1}^{n} \ln[g(y_i \mid \theta)].$$

Computing the likelihood or the log likelihood as a function of parameter values or model predictions provides a likelihood profile, which allows us to see how the model's likelihood changes as parameter values are changed. Likelihood profiles can be used to calculate confidence intervals on model parameters

Bayesian model averaging is an alternative approach to model selection and prediction (Hoeting et al. 1999). The idea of Bayesian model averaging is to average across several models instead of selecting one model. In computing the average, each model is weighted by its posterior model probability, a measure of the degree of model support in the data. Empirical and theoretical results over a broad range of model classes indicate that Bayesian model averaging can provide improved out-of-sample predictive performance as compared to single models.

**Geostatistical analysis**

Geostatistics is based upon the spatial pattern recognition that in the Earth sciences there is usually a lack of sufficient knowledge concerning how properties vary in space. Therefore, a deterministic model may not be appropriate. If we wish to make predictions at locations for which we have no observations, we must allow for uncertainty in our description as a result of our lack in knowledge. So, the uncertainty inherent in predictions of any property we cannot describe deterministically is accounted for through the use of probabilistic models. The most effective way to visualize spatial pattern data is to plot them as a dot map. A dot map is a region over which the events are observed as points. The dot map has long been one of the most popular cartographic tools of geographers. Figure 4 represents the locations of the data represents the positions of the centers of a region field. The data region is split into a grid with width 0.025m and the frequencies of the points in the pixels are calculated. Obviously the pattern appears clusters.

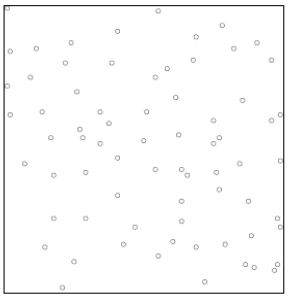

Fig. 4. Spatial locations of the epicenters of the region firld

One way to summarize the events in a spatial point pattern is to divide the regions into sub regions of equal areas (quadrants). By counting the number of events inside each quadrant, we end with a measure (frequency or histogram) that summarizes the spatial pattern. The intensity of a point pattern is the mean number of points per unit area. Intensity plots display a smooth estimate of intensity for a spatial point pattern. One can see how dependent the visual display is on some of the options when one considers the intensity image via the binning method. Figure 5 illustrates two cases for simulating cluster process using different fraction of the area values.

ISBN: 978-1-61804-110-4

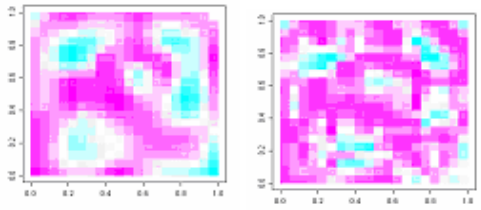

Fig. 5. Simulation of a cluster process with different fractions of area values.

Binning via nonparametric smoothing is a tool for intensity estimation. Rectangular bins are formed within the region. A moving window or 'span' is used to estimate the local intensity through the local nonparametric regression function loess. With the concentrations of locations in just two areas, changing the fraction of region would not accomplish much here visually. Widening the fraction will result in a much smoother/coarser image (Figure 6).

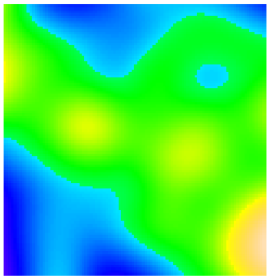

Fig. 6. Binning smoothing

Geostatistics provides the practitioner with a methodology to quantify spatial uncertainty. Statistics come into play because probability distributions are the meaningful way to represent the range of possible values of a parameter of interest. In addition, a statistical model is well suited to the apparent randomness of spatial variations.. It must be noted that there is considerable variety of statistical methods that have been applied in the analysis of spatial variation in ecological data, summarized by (Dale, 1999). These include dispersal analysis, spectral analysis, wavelet analysis, kriging and spatial Monte Carlo simulations and many geostatistics methods. Kriging was developed for estimating thresholds of continuous variables. It has been used for interpolation and simulation of categorical variables and for spatial uncertainty analysis. Auto-correlation of the spatial variables was represented by indicator auto-variogram models. The kriging estimator for the occurrence probability of a class $C_i$ at the location $x_o$ is defined as a linear combination of the surrounding indicator data $I_a$ with

$$P(x_o \mid x_{i_a}) = \sum_{i=1}^{a} \lambda_{i_a} I_{i_a}$$

where the set of weights $\lambda_{ia}$ are obtained by solving a system of linear equations

$$\begin{cases} \sum_{a=1}^{n} \lambda_{i_a} \gamma_{ii}(h_{a\beta}) + \mu_i = \gamma_{ii}(h_{\beta 0}) \\ \sum_{i=1}^{a} \lambda_{i_a} = 1 \end{cases}$$

where $\gamma_{ii}$ ($h_{a\beta}$) denotes the indicator auto-variogram between data locations $x_a$ and $x_\beta$ and $\gamma_{ii}$ ($h_{\beta 0}$) denotes the indicator auto-variogram between data locations $x_\beta$ and $x_0$; $\mu$ is the Lagrange parameter.

Markov Chain Geostatistics is a new non-kriging geostatistics. The basic idea of this geostatistics is to use Markov chains to perform multidimensional interpolation and simulation. Compared with the covariance-based (or variogram-based) geostatsitics, MCG is transition probability-based. Compared with the kriging-based geostatistics, MCG is Markov chain-based. MCG directly uses Markov chains to accomplish conditional simulation. The basic idea of MCG is that an unknown location is related on its nearest known neighbours in different directions. With a Markov chain moving around in a space, its conditional probability distribution at any unknown point is entirely dependent on its nearest known neighbors in different directions. The interaction between each nearest known neighbour and the unknown location is expressed by a transition probability at the corresponding distance. Therefore, transiograms are the explicit components of the conditional probability function.

The MCG estimator for the occurrence probability of a class C at the location x is defined as a non-linear combination of the surrounding data with

$$P(x = k \mid X_1 = l_1, \ldots, X_m = l_m) = \frac{\prod_{i=1}^{m} p_{kl}^{i}(h_i) p_{lk}^{l}(h_l)}{\sum_{j=1}^{n} \left[ \prod_{i=1}^{m} p_{jl}^{i}(h_i) p_{lj}^{l}(h_l) \right]}$$

where $p_{lk}^{i}(h)$ denotes a transiogram from state l to state k over the lag h in the ith direction and m is the number of data considered in different directions. Refering to conditional probabilities, due to the largness of the configuration space it is impractical to sample from it by direct computation of the probabilities. **Markov chains Monte Carlo**

ISBN: 978-1-61804-110-4

**(MCMC)** methods have been investigated by various researchers as an alternative to exact probability computation. The general method is to simulate a Markov chain with the required probability distribution as its equilibrium distribution. If the chain is aperiodic and irreducible, the convergence is guaranteed.
Then realisations of the Markov chain form a pseudo-sample from the required distribution. This pseudo-sample can then be used to estimate various statistical measures of the image. In particular our interest is concentrated on the distribution p(x) using a particular realisation $X^1, X^2, \ldots, X^N$ on the Markov chain with transition probability $p(x \to x')$. Typically asymptotic results include:

$$X^t \underset{t\to\infty}{\overset{d}{\to}} X \sim p(x)\ ;\ \ \frac{1}{t}\sum_{i=1}^{t} f(x^i) \xrightarrow[t\to\infty]{} E_p\{f(x)\}$$

where the expectation $E_p\{f(x)\}$ is to be estimated. The corresponding empirical average will be used namely:

$$\bar{f}_N = \frac{1}{N}\sum_{t+1}^{N} f\left(x^{(t)}\right)$$

Then realiations of the Markov chain form a pseudo-sample from the required distribution. This pseudo-sample can be used to estimate various statistical measures of the images. A special case of MCMC methods is the Gibbs sampler (Geman and Geman, 1984).
At each iterations, the value of only one pixel can change. The Gibbs sampler algorithm is defined as follows: Pick an arbitrary starting value $\mathbf{x}^0 = \left(x_1^0, x_2^0, \ldots, x_n^0\right)$. Simulate a new intensity for each piel from the corresponding conditional distribution $p\left(x_i \mid x_1^{t+1}, x_2^{t+1}, \ldots, x_{i-1}^{t+1}, x_{i+1}^{t}, \ldots, x_n^t\right)$.The iterative scheme is continued to produce a sequence $\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^t, \ldots$, which is a realisation of a Markov chain. After an initial period we obtain a realisation from the required probability distribution. Convergence is studied analytically in Green and Han (1992) and Sokal (1989).
Ripley's K is usually estimated as:

$$\hat{\mathrm{K}}(d) = \frac{\sum_{i\neq j}\sum I_d(d_{ij}) \big/ w_{ij}}{N^2 A}$$

where: $d_{ij}$ is the distance between the i-th and j-th observed event locations, which can be viewed as the radius of a circle centered at event point i and passing through j; $w_{ij}$ is a weighting term equal to this circle's proportion of the entire area A For specific distance/lag value d, $Id(d_{ij})$ is an indicator function which is 1 if $d_{ij} = d$.
The above estimate of K(d) is 'edge–corrected'. The weight wij accounts for the fact that an event j may not be in A given that it is a distance $d_{ij}$ from the point i. A valuable investigation for a point pattern is a plot of $K\hat{}\,(d)$ versus $\pi d2$ . If $K\hat{}\,(d) > \pi d2$ then there is evidence of clustering. Figure 7, represent the graph between Ripley K-function and theoretical one, where clearly there is evidence of regularity for landscape data.

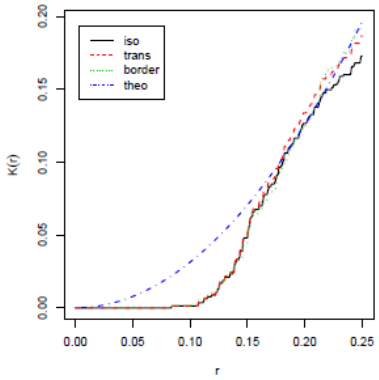


Fig. 7. Graph between Ripley K-function and theoretical one

(Ripley, 1981) also refers to a more general K-function as the reduced second-order measure. If we assume the process is completely random then the extra number of events within a distance d will be uniform on a disc. From this we see that

$$K(d) = \int_0^2 \int_0^{\pi d} \frac{\lambda_2(x)}{\lambda} x\,dx\,d\theta = \frac{2\pi}{\lambda}\int_0^d \lambda_2(x)x\,dx$$

with $\lambda_2(x) = \dfrac{\lambda^2}{2\pi I}\dfrac{\partial K(d)}{\partial I}$ .

## COMCLUSIONS

Uncertainty analysis is the part of risk analysis that focuses on the uncertainties in the data characteristics. Although the uncertainty analysis problem is similar to statistical problems, in application the uncertainty analysis problem is often more complex than many statistical problems. Models are used due to the complexity of the problem. The choice of the model is a choice that leads to structural uncertainty. A tool related to uncertainty analysis is sensitivity analysis. Sensitivity analysis is used to determine the

ISBN: 978-1-61804-110-4

importance of different parameters and components of the model on the output of the model. If the response variable depends on several variables, then the sensitivity of the response with respect to the variable or parameter is measured by the derivative of the response with respect to the variable or parameter.

Uncertainties in the biological modelling arise from either systematic or random errors. Variability in biological systems may result from spatial heterogeneity of environmental conditions or from randomness in interactions of different processes. Spatial variability in systems attributes and driving variables need to be effectively incorporated into simulation modelling. Model comparison can incorporate with uncertainty into model, especially when no data are available for model testing. The variability in the variables is then used as a measure of the output uncertainty. In this work, a review of various statistical techniques is represented, to measure the uncertainty introducing models. For the parameters of these models, techniques like Bayesian analysis, hierarchical modelling and geostatistical modelling are introduced to estimate the appropriate values.

ACKNOWLEDGMENT

This work is supported by an FP7 European Project, SCALES

## *References*

1. Smith, E.P. & Cairns, J.J. (1993). Extrapolation methods for setting ecological standards for water quality: statistical and ecological concerns, *Ecotoxicology* **2**, 203–219.
2. Linkov, I., Satterstrom, K., Seager, T.P., Kiker, G., Bridges, T., D. Belluck, and Meyer, A. (2006). "Multi-Criteria Decision Analysis: Comprehensive Decision Analysis Tool for Risk Management of Contaminated Sediments". *Risk Analysis* 26:61-78.
3. Allen, M. R., P. A. Stott, J. F. B. Mitchell, R. Schnur, and T. L. Delworth. 2000. Quantifying the uncertainty in forecasts of anthropogenic climate change. Nature 407:617-620.
4. Armstrong, J. S., and F. Collopy. 1992. Error measures for generalizing about forecasting methods: empirical comparisons. International Journal of Forecasting 8:69-80.
5. Beck, M. B. 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research 23(8):1393-1442.
6. Berk, R. A., P. Bickel, K. Campbell, S. Keller-McNutly, E. Kelly, and J. Sacks. 2002. Workshop on Statistical Approaches for the Evaluation of Complex Computer Models. Statistical Science 17: 173-192.
7. Beldring, S. 2002. Multi-criteria validation of a precipitation-runoff model. Journal of Hydrology 257:189-211.
8. Bliischl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling: a review. Hydrological Processes 9:251-290.
9. Bubier, J. L., S. Frolking, P. M. Crill, and E. Linder. 1999. Net ecosystem productivity and its uncertainty in a diverse boreal peatland. Journal of Geophysical Research 104:27683-27692.
10. Cale, W. G., R. V. O'Neill, and H. H. Shugart. 1983. Development and application of desirable ecological models. Ecological Modelling 18:171-186.
11. Clark, J. S., S. Carpenter, M. Berber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. E. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657-660.
12. Forest, C. E., P. H. Stone, A. P. Sokolov, M. R. Allen, and M. D. Webster. 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations. Science 295:113-117.
13. Franks, S. W., and K. J. Beven. 1997. Estimation of evapotranspiration at the landscape scale: a fuzzy disaggregation approach. Water Resources Research 33(12):2929-2938.
14. Funtowwicz, S. 0., and J. R. Ravetz. 1990. Uncertainty and Quality in Science for Policy. Kluwer Academic Publishers, Dordrecht.
15. Gardner, R. H., and R. V. O'Neill. 1983. Parameter uncertainty and model predictions: a review of Monte Carlo results. Pages 245-257 *in* M. B. Berk and G. V. Straten, editors. Uncertainty and Forecasting of Water Quality. Springer-Verlag, New York.
16. Gardner, R. H., V. H. Dale, and R. V. O'Neill. 1990. Error Propagation and uncertainty in process modeling. Pages 208-219 *in* R. K. Dixon, R. S. Meldahl, G. A. Ruark, and W. G. Warren, editors. Forest Growth: Process Modeling of Response to Environmental Stress. Timber Press, Portland.
17. Hansen, S., M. Thorsen, E. J. Pebesma, S. Klesschulte, and H. Svendsen. 1999. Uncertainty in simulated nitrate leaching due to uncertainty in input data. A case study. Soil Use and Management 15:167-175.
18. Hargrove, W. W., R. H. Gardner, M. G. Turner, W. H. Romme, and D. G. Despain. 2000. Sirnulating fire patterns in heterogeneous landscapes. Ecological Modelling 135:243-263.
19. Heuvelink, G. B. M. 1998a. Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems 50:255-264.
20. Heuvelink, G. B. M. 1998b. Error Propagation in Environmental Modeling with GIS. Taylor & Francis, London.
21. Hunsaker, C. T., M. F. Goodchild, M. A. Friedl, and T. J. Case, editors. 2001. Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications. Springer-Verlag, New York.
22. Jain, A. K., H. S. Kheshgi, and D. J. Wobbles. 1997. Is there an imbalance in the global budget of bombproduced radiocarbon? Journal of Geophysical Research 102:1327-1333.
23. Jenkins, J. C., D. W. Kicklighter, S. V. Ollinger, J. D. Aber, and J. M. Melillo. 1999. Sources of variability in net primary production predictions at a regional scale: a comparison using PnET-II and TEM4.0 in Northeastern US forests. Ecosystem 2:555-570.
24. Katz, R. W. 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Research 20:167-185.
25. Klepper, O. 1997. Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration. Ecological Modelling 101:1-13.
26. Kros, J., E. J. Pebesma, G. J. Reinds, and P. A. Finke. 1999. Uncertainty assessment in modeling soil

acidification at the European scale: a case study. Journal of Environmental Quality 28:366-377.
27. Krzysztofowicz, R. 1999a. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resources Research 35:2739-2750.
28. Krzysztofowicz, R. 1999b. Bayesian Forecasting via Deterministic Model. Risk Analysis 19(4):739-749.
29. Kvalseth, T. O. 1985. Cautionary note about R2 The American Statistician 39(4):279-285.
30. Lowell, K., and A. Jaton, editors. 1999. Spatial Accuracy Assessment - Land Information Uncertainty in Natural Resources. Ann Arbor Press, Chelsea.
31. Mayer, D. G., and D. G. Butler. 1993. Statistical validation. Ecological Modelling 68:21-32.
32. McKay, M. D., W. J. Conover, and R. J. Beckman. 1979. A comparison of three methods for selection values of input variables in the analysis of output from a computer code. Technometrics 2:239-245.
33. Melching, C. S., and W. Bauwens. 2001. Uncertainty in coupled nonpoint source and stream water quality models. Journal of Water Resources Planning and Management 127:403-413.
34. Mitchell, T. D., and M. Hulme. 1999. Predicting regional climate change: living with uncertainty. Progress in Physical Geography 23:57-78.
35. Morgan, M. G., and M. Henrion. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge.
36. Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models. Part 1 – a discussion of principle. Journal ofHydrology 10:282-290.
37. Nauta, M. J. 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. International Journal of Food Microbiology 57:9-18.
38. O'Neill, R. V. 1979. Natural variability as a source of error in model predictions. Pages 23-32 *in* G. S. Innis and R. V. O'Neill, editors. Systems Analysis of Ecosystems. International Co-operative, Fairland, MD.
39. O'Neill, R. V., and R. H. Gardner. 1979. Sources of uncertainty in ecological models. Pages 447-463 *in* B. P. M. Zeigles, M. S. Elzas, G. J. Klir, and T. 1. Oren, editors. Methodology in Systems Modelling and Simulation. North Holland, Amsterdam, Holland.
40. Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263:641-646.
41. Pate-Cornell, M. E. 1996. Uncertainties in risk analysis: six levels of treatment. Reliability Engineering and System Safety 54:95-111.
42. Petersen, A. C. 2000. Philosophy of climate science. Bulletin of the American Meteorological Society 81(2):265-271.
43. Rastetter, E. B., A. W. King, B. J. Cosby, G. M. Hornberger, R. V. O'Neill, and J. E. Hobbie, 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecological Applications 2:55-70.
44. Reckhow, K. H. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. Ecological Modelling 72:1-20.
45. Regan, H., M. Colyvan, and M. A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications 12:618-628.
46. Reynolds, J. F., D. W. Hilbert, and P. R. Kemp. 1993. Scaling ecophysiology from the plant to the ecosystem: a conceptual framework. Pages 127-140 *in* J. R. Ehleringer and C. B. Field, editors. Scaling Physiological Processes: Leaf to Globe. Academic Press, San Diego, CA.
47. Rykiel, E. J. 1996. Testing ecological models: the meaning of validation. Ecological Modelling 90: 229-244.
48. Rypdal, K., and W. Winiwarter. 2001. Uncertainties in greenhouse gas emission inventories ~ evaluation, comparability and implications. Environmental Science & Policy 4: 107-116.
49. Saltelli, A., K. Chan, and E. M. Scott, editors. 2000. Sensitivity Analysis. Wiley, Chichester, UK.
50. Scherm, H. 2000. Simulating uncertainty in climate-pest models with fuzzy numbers. Environmental Pollution 108:373-379.
51. Scherm, H., and A. H. C. van Bruggen. 1994. Global warming and nonlinear growth: how important are changes in average temperature? American Phytopathological Scoiety 84:1380-1384.
52. Schulze, R. 2000. Transcending scales of space and time in impact studies of climate and climate change on agrohydrological responses. Agriculture, Ecosystem and Environment 82: 185-212.
53. Stott, P. A., and J. A. Kettleborough. 2002. Origins and estimates of uncertainty in predictions of twentyfirst century temperature rise. Nature 416:723-726.
54. Wiwatenadate, P., and H. G. Claycamp. 2000. Error propagation of uncertainties in multiplicative models. Human and Ecological Risk Assessment 6:355-368.
55. Wobbles, D. J., A. Jain, 1. Edmonds, D. Harvey, and K. Hayhoe. 1999. Global change: state of the science. Environmental Pollution 100:57-86.
56. Wu, J. 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25:367-380.
57. Zhang, J., and C. T. Haan. 1996. Evaluation of uncertainty in estimated flow and phosphorus loads by FHANTM. Applied Engineering in Agriculture 12(6):663-669.
58. Dale, M. R. T., 1999. Spatial pattern analysis in plant ecology, Cambridge Univ. Press.
59. Green P. J., Han X-L., 1992. Metropolis methods, Gaussian proposals, and antithetic variables, Lect. Notes Stat. 74: 142-164.
60. Sokal A. D., 1989. Monte Carlo methods in statistical mechanics: foundations and new algorithms. Cours de Troisieme Cycle da la Physique en Suisse Romande, Lausanne.
61. Geman S., Geman D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell., 6: 721-741.
62. P.J. Diggle, J.A. Tawn and R.A. Moyeed. Model-based geostatistics (with discussions). In Applied Statistics, volume 47, issue3, pp. 299-350, 1998.
63. C.K. Wikle. Hierarchical Bayesian models for predicting the spread of ecological processes. In Ecology, volume 84, pp. 1382-1394, 2003.
64. Zimeras S and Matsinos Y. Spatial Uncertainty, Recent Researches in Geography, Geology, Energy, Environment and Biomedicine, 203-208, 2011
65. J. A. Wiens Spatial Scaling in Ecology Functional Ecology, Vol. 3, No. 4. (1989), pp. 385-397
66. Turner, M. G., 1989. Landscape ecology: the effect of pattern on process, Annu. Rev. Ecol. Syst. 20: 171–197.
67. Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm?, Ecology 74: 1659–1673.
68. Lennon, J. J., 2000. Red-shifts and red herrings in geographical ecology, Ecography 23: 101–113.
69. Jansen, M.J.W., 1998. Uncertainty analysis of food-chain models. *In:* Tijskens, L. M. M. and Hertog, M. L. A. T. M. eds. *Proceedings of the international symposium on applications of modelling as an innovative technology in the agri-food-chain, Model-It, Wageningen, Netherlands, 29 November-2 December, 1998. Acta Horticulturae ISHS*, Vol. 476. 33-40.