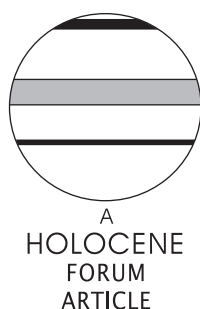


Equifinality and uncertainty in the interpretation of pollen data: the Multiple Scenario Approach to reconstruction of past vegetation mosaics

M.J. Bunting* and R. Middleton

(Department of Geography, University of Hull, Cottingham Road, Hull HU6 7RX, UK)

Received 28 August 2008; revised manuscript accepted 9 February 2009



Abstract: The long-term goal of pollen analysis has always been the reconstruction of past vegetation mosaics. However, the pollen signal is spatially integrated, and ecologically distinct vegetation structures can produce identical pollen signals. This paper proposes an iterative approach to reconstruction, the Multiple Scenario Approach, which involves creating a large number of hypothetical landscapes using a combination of deterministic and probabilistic ‘rules’ for plant placement coupled with environmental parameters such as topography, simulating the pollen assemblage produced by each, and comparing them statistically with a sedimentary pollen signal to identify possible reconstructions. The method expressly aims to find multiple possible reconstructions, and also gives some insight into the spatial aspects of the vegetation being reconstructed in addition to its composition.

Key words: Multiple hypotheses, palaeoecology, pollen dispersal and deposition, Prentice–Sugita model, pollen dispersal and deposition, vegetation reconstruction.

Introduction

Quantitative reconstruction of past vegetation has long been the goal of pollen analysis. The potential audience for such reconstructions includes both academic (eg, archaeologists, climatologists, ecologists) and non-academic (eg, museums, local communities, schools, educationalists, policy makers, nature conservancy planners) groups. In order to better communicate with these groups, and to enable pollen analytical data to contribute fully to key contemporary debates on topics such as protection of biodiversity and restoration of ‘natural’ ecosystems, pollen data need to be translated not just into a quantitative measure of past vegetation, but into maps of past vegetation distribution.

Various approaches to achieving this goal can be identified in the literature, including the use of the modern analogue technique (MAT: eg, Overpeck *et al.*, 1985; Calcote, 1998) to match past pollen assemblages to present assemblages and thus vegetation formations, or using algebraic descriptions of the pollen–vegetation relationship to translate pollen counts into a numerical estimate of past vegetation (eg, Davis, 1963; Prentice and Parsons, 1983; Sugita, 2007a, b). Both methods have enabled advances in the interpretation of pollen signals and thus the understanding of past vegetation, but both have their limitations.

The MAT is limited by the availability of modern surface samples from locations representing possible past communities. Finding representative communities is problematic, especially in heavily human-impacted parts of the world such as northwest Europe, where semi-natural or traditional cultural landscape fragments are typically embedded in a landscape dominated by industrial agriculture. In addition, the sampling context (eg, basin size, surface vegetation) can have a marked effect on the pollen signature of a landscape (see, eg, Jacobson and Bradshaw, 1981; Sugita, 1993, 1994), and therefore past and modern analogue samples need to come from matched contexts to ensure a robust reconstruction.

The algebraic description approach is limited by the assumptions of the pollen–vegetation model and the requirements of the reconstruction approach chosen. For example, the Landscape Reconstruction Algorithm (LRA: Sugita, 2007a, b) requires that large (>750 ha) basins are available for reconstruction of the regional vegetation and thus the past background pollen rain component (Sugita, 2007a), and assumes that the region (defined as a radius of 100–200 km around the large site) has uniform vegetation composition. These reconstructions of regional vegetation are coarse-grained, but sufficient and appropriate for continental-scale research questions, such as determining the changing albedo of the land-mass of Europe as an input to regional climate models (Anderson *et al.*, 2006). Where suitable large sites are available,

*Author for correspondence (e-mail: m.j.bunting@hull.ac.uk)

and the assumption of broad vegetation uniformity can be made, this approach has considerable potential. However, many research questions concern areas where either large sites are not available or the regional uniformity assumption cannot be made, or they require finer-scale reconstructions.

Within the LRA, small basin pollen records are used to reconstruct vegetation at a finer scale. The regional background pollen component is derived from the reconstructed regional vegetation from the large basin data, allowing separation of the local component of the signal from the small site, therefore both sites need a robust chronological framework to support correlation. The extended R-value approach (Prentice and Parsons, 1983) is then applied to this local component to calculate the distance-weighted plant abundance within the Relevant Source Area of Pollen (*sensu* Sugita, 1994) of the basin (Sugita, 2007b). However, obtaining a single distance-weighted measure of the vegetation around one or more basins in a past landscape does not fully translate the data into a form which can be readily understood and used by 'end-users' of pollen analysis research, and interpolation between the scattered points is a non-trivial problem.

In this paper we present an alternative method for quantitative reconstruction of past vegetation mosaics, using a combination of modelling and analogue matching. We argue that this 'Multiple Scenario Approach' (MSA), whilst having its own limitations, provides a useful alternative to the MAT and LRA approaches, especially in complex geographical settings where vegetation composition is expected to vary markedly over relatively short distances, or for some types of non-analogue landscape mosaics.

The Multiple Scenario Approach

All interpretation of pollen diagrams makes certain assumptions, whether these are explicitly spelt out or not, about factors such as the relative pollen productivity of different plants and their ecological requirements. Assumptions about pollen dispersal and deposition are overtly spelt out in models of the process such as the Prentice–Sugita model (Prentice, 1985; Sugita, 1993). With rapid increases in desktop computer power and speed, running simulation models is becoming simpler and quicker. The development of easily available user-friendly software for simulating pollen dispersal and deposition such as MOSAIC (Middleton and Bunting, 2004) and HUMPOL (Bunting and Middleton, 2005) and their dissemination particularly through the workshops and other outreach activities of the PollLandCal Network (see <http://www.geog.ucl.ac.uk/ecrc/pollandcal>) has led to a recent increase in the use of models by pollen analysts. Using simulation approaches has proved useful not just for better understanding the behaviour of the pollen dispersal and deposition model but also for explaining observations on the pollen signal in present-day vegetation mosaics (eg, Sugita, 1994; Sugita *et al.*, 1999; Bunting *et al.*, 2004; Broström *et al.*, 2005).

Simulations are also beginning to be used to support the interpretation of past pollen records. Alternative past vegetation mosaics for a specific landscape can be created in a GIS by, for example, varying tree line position, cereal cultivation limits or community composition. Pollen assemblages can then be simulated for a specified location, and simulated values compared with actual assemblages from a core from that point to identify which of the possible alternatives is the most likely reconstruction of the landscape mosaic at the time when the core assemblage was forming (eg, Fyfe, 2006; Caseldine and Fyfe, 2006; Caseldine *et al.*, 2007; Bunting *et al.*, 2008).

However, similar pollen signals can be generated by ecologically distinct landscapes. For example, woodland made up of monospecific stands of oak and elm will produce the same signal in a local

lake as randomly mixed oak-elm woodland, or a single cereal field adjacent to the basin may produce as strong a cereal pollen signal as a large area of arable cultivation several hundred metres distance from the shore. The local vegetation on the surface of a sampling site can also profoundly affect the pollen signal; *Calluna* growing on the mire surface and *Calluna* growing on surrounding hill slopes both produce the same pollen type, therefore reconstructing the landscape represented by a pollen record from the mire can be challenging (eg, Bunting, 2008). If palynologists are to produce meaningful interpretations of pollen data, they (we) have to embrace these ambiguities. Equifinality, different landscapes producing the same pollen signal, is innate to the system being studied, and we should not just mention alternative interpretations in passing but actively seek out the multiple possibilities. As scientists, this seems obvious, but as human beings identifying possible alternatives can be challenging, and a single definitive interpretation is highly attractive. The simulation approach outlined here offers tools for seeking out equifinal solutions in an objective manner, albeit within the innate limitations of the models used.

The MSA methodology is specifically designed to generate multiple ecologically distinct reconstructions of past vegetation from pollen signals, forming multiple hypotheses for subsequent testing. It is hoped that this approach will help palaeoecology as a discipline achieve a shift in emphasis from the collection of data and the construction of narratives about that data towards an overtly analytical, hypothesis-testing approach which is probably long overdue (see, eg, Ball, 1975; Birks, 1993, 1996).

The basic approach of the MSA is the same as that used in earlier reconstruction-via-simulation studies (eg, Fyfe, 2006; Caseldine and Fyfe, 2006; Caseldine *et al.*, 2007). Plant requirements, ecological processes and random factors are combined with known landscape parameters (eg, altitude) in a GIS to generate possible vegetation maps. A model of pollen dispersal and deposition is used to simulate the pollen assemblage at a known sampling point in the landscape for each possible map. These simulated assemblages are compared with an actual past assemblage from the same location to select which of those maps can be considered a likely reconstruction of past vegetation. The main differences lie in the number of maps generated (10^4 – 10^5), the increased emphasis on random factors and on the expectation of multiple outcomes, and in the formalisation of methods for identifying likely scenarios, all of which are discussed in more detail below.

At present the pollen dispersal and deposition model used is the Prentice–Sugita model, which requires that relative pollen productivity and fall speed are specified for all of the pollen types modelled, and makes other explicit assumptions (see, eg, Sugita, 1994). However, alternative dispersal and deposition models would be equally suitable (for example, a form which includes water-borne as well as aerial pollen transport routes), and the software mentioned below has been designed so that alternative algorithms can be added with minimal additional work.

An MSA reconstruction

The reconstruction process is carried out first for individual points in time. The reconstruction(s) produced for one point in time are referred to as 'time-slices', since each reconstruction is a spatially explicit two-dimensional vegetation map for a single time period. It is possible to carry out MSA reconstructions with a single pollen record (eg, Bunting *et al.*, 2007, 2008), although stronger results will be obtained where multiple sites with a strong chronological framework allow each time-slice to be constrained by multiple data points (eg, Pratt, 2006). Time-slices are then stacked chronologically at each sampling location to generate a reconstruction of probable landscape dynamics.

Reconstructing a single timeslice

Step 1: define the boundaries of the initial scenario swarm

The area to be reconstructed is identified, and data on invariant aspects of the physical landscape such as altitude, slope angle and geology collated. In some cases aspects of palaeogeography (eg, altered sea level) can also be specified. The taxa to be included in the reconstruction are then chosen. These are generally the main types in the pollen diagram, although some limits are also imposed by whether pollen productivity estimates are available to allow simulation of the pollen signal of the taxon. Once taxa are selected, their broad autecological and synecological requirements can be identified, bearing in mind that several plant taxa are often represented by one palynomorph. In some situations it may be easier to specify ecological controls on the distribution of particular communities rather than individual taxa, but if this approach is used then past community composition also has to be defined.

Ecological requirements are then combined with the physical data in a GIS to generate a vegetation map. Because the proportion of suitable landscape areas (as defined by ecological requirements) occupied by each taxon can vary, many possible coverage combinations can be generated. For example, a simple British upland landscape might be modelled as heathland scattered with patches of *Betula* woodland on steeper slopes, valley mire communities in some depressions, blanket mire covering parts of flatter areas and distinct pasture communities on alluvial soils. The proportion of possible landscape (eg, steep slopes) occupied by the patch (eg, *Betula*) is not known, but different possibilities can be tried. If three different coverage proportions were tried for each of the four communities independently, then 3^4 (81) possible configurations need to be tested. Variations in other factors might also be tested. For example, the maximum size of areas of pasture or the slope angle limits on *Betula* growth or blanket bog formation could also be varied. Even if only three options were tried for each of these parameters, 3^7 (2187) possible combinations now need to be tested. Unless all possible landscape areas are covered 100% with the patch community, random placement factors will mean that multiple maps can be generated from each combination of landscape rules. A group of maps generated from one such combination is referred to here as a family of scenarios; this simple scenario would generate 2187 families.

These rule-derived vegetation maps can be generated using a standard GIS (eg, Fyfe, 2006; Caseldine *et al.*, 2007), but handling the large number of parameters and combinations is time-consuming, therefore LANDSCAPE (Middleton, unpublished data, 2006) has been developed for this task.

Step 2: identify likely families

In step 2, the pollen assemblage at a specified location in the physical landscape is simulated for multiple members of each family (typically 10–100, depending on the variability of the outputs). These pollen assemblages are then compared with an actual pollen assemblage from that location using the approach developed for the MAT, for example by calculating squared chord distance (Maher, 2000; Wahl, 2004), and the results for each family are summarised by measures of central tendency and dispersion (eg, median and quartiles, mean and s.d.) – the MSA software suite (HUMMSA) includes CHORD for this purpose. Likely families are defined as those in which the squared-chord distance for at least a proportion of members falls below a specified cut-off value (eg, 0.20). The cut-off thresholds identifying a ‘good match’ are debated in the modern analogue research community (eg, Lytle and Wahl, 2005), but values in the range 0.15–0.25 are typically used for squared-chord distance analogue matching (Wahl, 2004), although values as low as 0.05 may be appropriate when trying to discriminate between communities within a single biome (Calcote, 1998).

Where more than one palaeoecological record is available within the landscape being reconstructed, step 2 is slightly more complex since multiple pollen assemblages have to be simulated and compared for each scenario, and likely families have to contain at least some members with squared-chord distances below the cutoff for every site.

Step 3: identify possible scenarios

In step 3, a larger number of replicates are run for each likely family to identify one or more ‘best fit’ members, those with the lowest distance measure. Comparison of possible members (those below say a 0.05 cutoff) of each likely family could also be used to assess how robust the reconstruction is, and how secure the reconstruction is for different parts of the landscape. A likely family can be thought of as a reconstruction of the vegetation patterns in the broader landscape, whilst the ‘best-fit’ members give insight into patterns within the immediate vicinity of each site. A Bayesian mathematical approach may be an appropriate alternative means of choosing between reconstructions.

Developing a time series of reconstructions

Stacks of time-slices are easily generated, and landscape dynamics reconstructed by linking sequential time-slices. However, since multiple possible time-slices are expected for each point in time using this method, a large number of possible pathways can be traced. Caseldine and Fyfe (2006) propose a plausibility requirement which can be used to reduce the multiplicity of possible dynamic pathways, defining a plausible time-slice reconstruction as one which can be linked to the preceding and following time-slice reconstructions by an ecologically sensible process-based explanation. A plausible pathway could therefore be defined as one in which all time-slices can be linked by such an explanation.

An alternative approach to time series reconstruction would be to take possible scenarios from one time-slice and ‘evolve’ them forwards (or backwards) in time by modelling the dynamic effects of ecological processes. Since models of ecological processes include chance elements, repeating the evolution process multiple times with the same starting conditions would generate a family of related but distinct outcome scenarios. The same matching approach outlined above could then be used to identify which (if any) members of the family are probable scenarios for the future time slice and therefore to construct a plausible dynamic pathway.

Discussion

The MSA has several advantages over other methods for quantitative reconstruction of past vegetation from pollen records. It produces vegetation reconstructions in the form of maps, which are both well suited to dissemination to a range of end-users and present clear hypotheses for further testing. The MSA is not model-dependent, and the software mentioned here has been developed so that alternative models of pollen dispersal and deposition can be incorporated relatively simply as they become available.

The MSA allows reconstruction to be carried out with, at a minimum, a single pollen record, some physical parameters (eg, a digital elevation model) for a landscape and some minimal ecological rules or boundaries on vegetation proportions. Therefore it can be used in situations where the requirements of the LRA for multiple sites both large and small linked by a robust chronological framework are not available (Figure 1 compares the MSA and LRA approaches to landscape reconstruction). A highly variable wider region can be included in the model, and nesting grids of different resolution (see, eg, Bunting and Middleton, 2005) enable large areas to be modelled without slowing the process excessively.

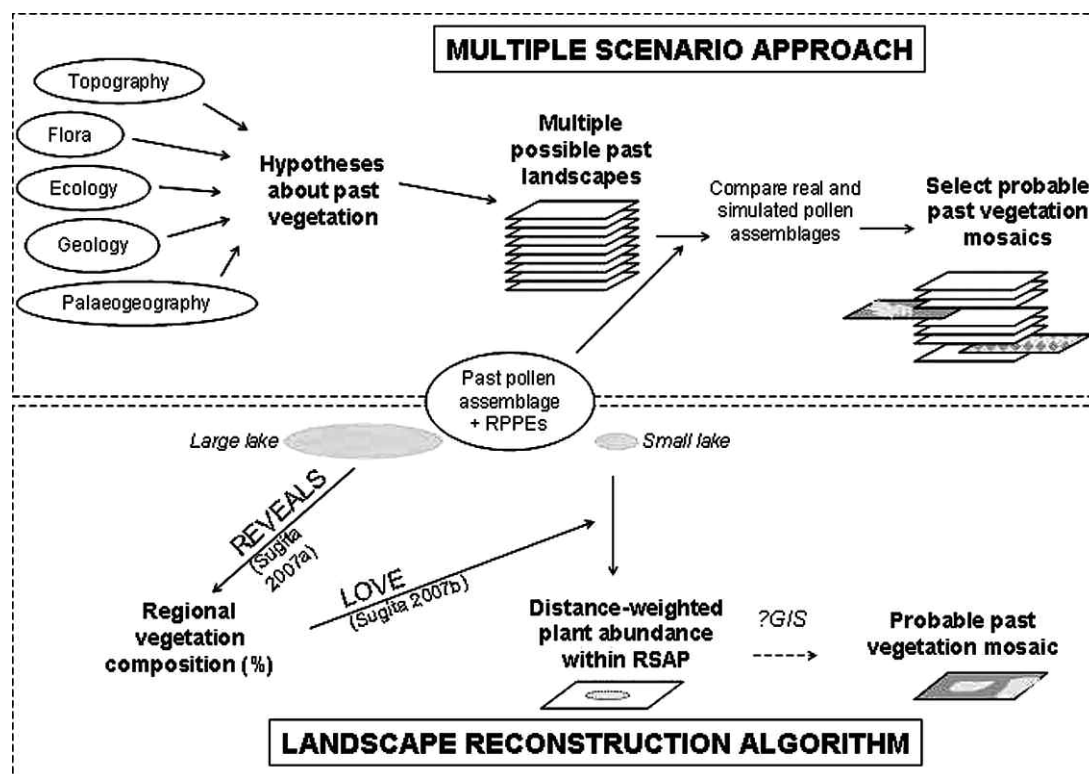


Figure 1 Schematic comparison of the Multiple Scenario Approach and the Landscape Reconstruction Algorithm. Oval boxes contain information to be input into the process, and labels along arrows (REVEALS, LOVE) indicate the software used during particular steps of the LRA

Therefore the MSA can be used in areas where the LRA assumption of regional uniformity is not reasonable. The combinations of plants modelled is only restricted by the availability of taxon-specific parameters, therefore non-analogue communities can easily be created and their pollen signal tested, avoiding the lack of analogues problem that occurs when using the MAT. However, the assumptions of the MSA also limit its applicability.

MSA assumptions

- That a pollen morphotype represents either one or several sets of plant species which are ecologically consistent over time (eg, *Alnus glutinosa* represents alder, and the ecological requirements of alder are consistent over time. Poaceae represents several groups of grasses associated with particular habitats (eg, upland, pasture, wetland); each group is assumed to have consistent, albeit different, ecological requirements).
- That the autecology and synecology of each morphotype's species set can be defined, and that they are uniform over the time period studied.
- That the model of pollen dispersal and deposition chosen (and its parameters such as relative pollen productivity) is correct and invariant through time.

Palaeoecology relies heavily on the principle of uniformitarianism, and assumptions 1–3 are essentially uniformitarian. However, these assumptions become less valid as past conditions become increasingly different from the present day. Published ecological requirements of plants are largely based on studies of the modern distribution, rather than on knowledge of the underlying tolerances of the species. A species might behave quite unexpectedly under non-analogue environmental conditions, such as lateglacial conditions when permafrost environments were widespread under temperate-zone light regimes, without any evolutionary changes. Whilst non-analogue environments can be studied using the MSA, the approach is better justified for interglacial situations similar to those under

which the ecological behaviours of the species have been defined. Simulating the pollen assemblages from deciduous woodlands with composition and coverage levels not found today seems a reasonable application of the approach, but using the MSA to reconstruct last glacial maximum environments is more of a stretch.

All reconstruction methods assume that the pollen-vegetation relationship is consistent and comprehensible. The MSA and LRA use a mathematical model of this relationship which enables them to explore non-analogue situations. The MAT, which is limited by requiring samples from modern analogues of communities and therefore can only identify changes in distribution of these communities in the past, overtly identifies 'non-analogue' situations when the method cannot produce a reconstruction, therefore might be considered to be potentially less misleading than the other methods. However, the MAT makes no allowance for equifinality, the possibility that a non-analogue community might produce the same pollen assemblage as an analogue community.

Currently the Prentice–Sugita algorithm appears to be the best available model of pollen dispersal and deposition, and is certainly the most developed and tested model. Its use involves making the following additional assumptions, which also limit the applicability of the MSA:

- that pollen transport is dominated by the above-canopy component
- that the vegetation surface over which this transport takes place is flat
- that past wind speed, wind rose distribution and atmospheric conditions of pollen release can be specified (in most cases, neutral atmospheric conditions, wind speed of 3 m/s and an even distribution of wind around the compass during the flowering season are assumed)
- that fall speed for a palynomorph can be sensibly specified by a single number
- that fall speed can be estimated from the geometry of a palynomorph (or that direct measurements made under appropriate conditions of grain hydration are available)

- that relative pollen productivity values for any pollen morpho-type are constant in space and time, and therefore that empirical estimates obtained today can be used in the past.

A major advantage of the MSA is that it seeks out uncertainty in reconstructions. It is designed to throw up multiple possible answers, ecologically distinct and valid landscape arrangements which produce very similar pollen signals, and therefore to generate multiple hypotheses for further testing. Such testing might involve further palynological work (eg, identifying a location within the landscape where the predicted pollen signal differs between the hypotheses being tested and where a sediment core can be collected) or the use of other palaeoecological methods (eg, Coleopteran analysis, since Coleoptera often have very specific associations with individual plant species and habitats). On this basis, it presents a very real step towards shifting the emphasis of palaeoecological research from data collection and narrative construction towards an overtly analytical, hypothesis-testing approach to knowledge generation.

Note

This paper presents a conceptual proposal for a method of vegetation reconstruction, the MSA. A software suite (HUMMSA) has been developed for carrying out the key steps (Middleton, unpublished data, 2008) and some initial reconstructions have been made (Pratt, 2006; Bunting *et al.*, 2007, 2008; Karlsson, 2008). Further development, validation and application will occur as research funding is obtained. Collaborations leading to alternative pollen dispersal and deposition models, and suggestions about the statistical thresholds and methods for selecting families and best fit scenarios, are currently sought.

Acknowledgements

We gratefully acknowledge the PolLandCal Network for many inspiring conversations and questions, and particularly for providing the stimulus to write the MOSAIC software which started off our work in this area. We especially thank Ralph Fyfe and Chris Caseldine for discussions about reconstruction via simulation, Kathy Willis and John Birks for positive feedback on earlier versions of these ideas at a seminar in Oxford. We also acknowledge Claire Twiddle, Sandra Marks, Hanna Karlsson, Simon Kitcher and other students who tested various versions of the software and asked good questions which made us think.

References

- Anderson, N.J., Bugmann, H., Dearing, J.A. and Gaillard, M.-J. 2006: Linking palaeoenvironmental data and models to understand the past and to predict the future. *Trends in Ecology & Evolution* 21, 696–704.
- Ball, I.R. 1975: Nature and formulation of biogeographical hypotheses. *Systematic Zoology* 24, 407–30.
- Birks, H.J.B. 1993: Quaternary paleoecology and vegetation science – current contributions and possible future developments. *Review of Palaeobotany and Palynology* 79, 153–77.
- 1996: Contributions of Quaternary palaeoecology to nature conservation. *Journal of Vegetation Science* 7, 89–98.
- Broström, A., Sugita, S., Gaillard, M.-J. and Piljesö, P. 2005: Estimating the spatial scale of pollen dispersal in the cultural landscape of southern Sweden. *The Holocene* 15, 252–62.
- Bunting, M.J. 2008: Pollen in wetlands: using simulations of pollen dispersal and deposition to better interpret the pollen signal. *Biodiversity and Conservation* 17, 2079–96.
- Bunting, M.J. and Middleton, R. 2005: Modelling pollen dispersal and deposition using HUMPOL software: simulating wind roses and irregular lakes. *Review of Palaeobotany and Palynology* 134, 185–96.
- Bunting, M.J., Gaillard, M.-J., Sugita, S., Middleton, R. and Broström, A. 2004: Vegetation structure and pollen source area. *The Holocene* 14, 651–60.
- Bunting, M.J., Middleton, R. and Twiddle, C.L. 2007: Wetland records of a coastal cultural landscape in north-west Scotland: an application of the Multiple Scenario Approach to landscape reconstruction. In Barber, B., Clark, C., Cressey, M., Crone, A., Hale, A., Henderson, J., Housley, R., Sands, R. and Sheridan, A., editors, *Archaeology from the wetlands: recent perspectives*. WARP Occasional Paper 18, Society of Antiquaries of Scotland, 109–17.
- Bunting, M.J., Twiddle, C.L. and Middleton, R. 2008: Using models of pollen dispersal and deposition in hilly landscapes: some possible approaches. *Palaeogeography Palaeoclimatology Palaeoecology* 259, 77–91.
- Calcote, R. 1998: Identifying forest stand types using pollen from forest hollows. *The Holocene* 8, 423–32.
- Caseldine, C. and Fyfe, R. 2006: A modelling approach to locating and characterising elm decline landnam landscapes. *Quaternary Science Reviews* 25, 632–44.
- Caseldine, C., Fyfe, R., Langdon, C. and Thompson, G. 2007: Simulating the nature of vegetation communities at the opening of the Neolithic on Achil Island, Co. Mayo – the potential role of models of pollen dispersal and deposition. *Review of Palaeobotany and Palynology* 144, 135–44.
- Davis, M.B. 1963: On the theory of pollen analysis. *American Journal of Science* 261, 897–912.
- Fyfe, R. 2006: GIS and the application of a model of pollen deposition and dispersal: a new approach to testing landscape hypotheses using the PolLandCal models. *Journal of Archaeological Science* 33, 483–93.
- Jacobson, G.L., Jr and Bradshaw, R.H.W. 1981: The selection of sites for paleovegetational studies. *Quaternary Research* 16, 80–96.
- Karlsson, H. 2008: *Vegetation changes and forest-line positions in the Swedish Scandes during late Holocene: anthropogenic impact vs. climate*. Acta Universitatis agriculturae Sueciae no 2008:31.
- Lytle, D.E. and Wahl, E.R. 2005: Palaeoenvironmental reconstructions using the modern analogue technique: the effects of sample size and decision rules. *The Holocene* 15, 554–66.
- Maher, L.J., Jr 2000: Modpol.exe: a tool for searching for modern analogs of Pleistocene pollen data. *Inqua sub-commission on data handling methods Newsletter* 20 (available online at <http://www.chrono.qub.ac.uk/inqua/index.htm>)
- Middleton, R. and Bunting, M.J. 2004: Mosaic v1.1: landscape scenario creation software for simulation of pollen dispersal and deposition. *Review of Palaeobotany and Palynology* 132, 61–66.
- Overpeck, J.T., Webb, T., III and Prentice, I.C. 1985: Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs for pollen data. *Quaternary Research* 23, 87–108.
- Pratt, S. 2006: Long term landscape dynamics in Caledonian pine forest. Unpublished PhD thesis, University of Edinburgh.
- Prentice, I.C. 1985: Pollen representation, source area, and basin size – toward a unified theory of pollen analysis. *Quaternary Research* 23, 76–86.
- Prentice, I.C. and Parsons, R.W. 1983: Maximum-likelihood linear calibration of pollen spectra in terms of forest composition. *Biometrics* 39, 1051–57.
- Sugita, S. 1993: A model of pollen source area for an entire lake surface. *Quaternary Research* 39, 239–44.
- 1994: Pollen representation of vegetation in Quaternary sediments: theory and method in patchy vegetation. *Journal of Ecology* 82, 881–97.
- 2007a: Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. *The Holocene* 17, 229–41.
- 2007b: Theory of quantitative reconstruction of vegetation II: all you need is LOVE. *The Holocene* 17, 243–57.
- Sugita, S., Gaillard, M.-J. and Broström, A. 1999: Landscape openness and pollen records: a simulation approach. *The Holocene* 9, 409–21.
- Wahl, E.R. 2004: A general framework for determining cutoff values to select pollen analogs with dissimilarity metrics in the modern analogue technique. *Review of Palaeobotany and Palynology* 128, 263–80.