Generation of $k$th-order random toposequences

Nathan P. Odgers*, Alex. B. McBratney, Budiman Minasny

Faculty of Agriculture, Food and Natural Resources, McMillan Building A05, The University of Sydney, Sydney, New South Wales 2006, Australia

Received 12 September 2006; received in revised form 7 May 2007; accepted 11 May 2007

Abstract

The model presented in this paper derives toposequences from a digital elevation model (DEM). It is written in ArcInfo Macro Language (AML). The toposequences are called $k$th-order random toposequences, because they take a random path uphill to the top of a hill and downhill to a stream or valley bottom from a randomly selected seed point, and they are located in a streamshed of order $k$ according to a particular stream-ordering system. We define a $k$th-order streamshed as the area of land that drains directly to a stream segment of stream order $k$. The model attempts to optimise the spatial configuration of a set of derived toposequences iteratively by using simulated annealing to maximise the total sum of distances between each toposequence hilltop in the set.

The user is able to select the order, $k$, of the derived toposequences. Toposequences are useful for determining soil sampling locations for use in collecting soil data for digital soil mapping applications. Sampling locations can be allocated according to equal elevation or equal-distance intervals along the length of the toposequence, for example.

We demonstrate the use of this model for a study area in the Hunter Valley of New South Wales, Australia. Of the 64 toposequences derived, 32 were first-order random toposequences according to Strahler’s stream-ordering system, and 32 were second-order random toposequences.

The model that we present in this paper is an efficient method for sampling soil along soil toposequences. The soils along a toposequence are related to each other by the topography they are found in, so soil data collected by this method is useful for establishing soil–landscape rules for the preparation of digital soil maps.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Toposequence; Digital elevation model; Stream order; Digital soil mapping; Soil survey; Soil sampling

1. Introduction

1.1. What is a catena?

In 1935, Milne proposed the term “catena” to describe a sequence or transect of soils that are related to each other by the topography that they are found in, though the soils that comprise the sequence may be of different types (Milne, 1935).
The word “catena” means “chain” in Latin, so each soil in the sequence can be likened to a link in the chain—it is related to the soils (links) on either side of it. Variation in topography influences the pedological processes that operate downhill between the members of the sequence (Huggett, 1975; Moss, 1965). These pedological processes include erosion, soil solute movement and drainage, for example (Conacher and Dalrymple, 1977).

Though Milne recognised catenary soil relationships in situations with both uniform and variable parent material (Bushnell, 1942; Williams, 1968), he realised that using the term “catena” in situations with variable parent material would spoil its “simplicity of connotation” (Bushnell, 1942). Hence, the term “catena” is correctly used when it is applied to situations where geology is constant along the length of the transect.

The term catena should not be confused with the term “soil association”, which is also found in the literature and implies merely a geographic association of soil types (“associates”) rather than a process-based linkage between soil types (Conacher and Dalrymple, 1977). Similarly, a toposequence can be defined as a relief profile along a transect drawn across a map (Gobin et al., 2000). Therefore it, too, has a morphological or topographic implication (Conacher and Dalrymple, 1977), but unlike the “catena” does not by necessity connote any process-based linkage between different soil types.

1.2. Why are toposequences useful?

The soil catena or toposequence is a useful concept for understanding soil–landscape relationships on a hillslope. Using the concept, we can study soil in the context of the landscape in which it is found (Huggett, 1975). One factor that affects soil formation is soil material movement across the hillslope via addition and loss of soil at individual sites. When soils on the same hillslope are linked using the toposequence concept, it is possible to elucidate the effects that will arise from movement of soil material from one location to another. For example, colluvium from upslope will add to surface horizons further down the slope, and so on. In this way, different soils can develop along one hillslope (Milne, 1936), though in reality these effects often prove to be quite subtle and therefore must be distinguished from soil development effects in situ (Watson, 1964). We can visualise a toposequence in the way that Milne did (Fig. 1):

Whilst the catena is a very useful concept, we realise that there seems to be much ambiguity surrounding use of the word in its classical sense. Toposequences and soil associations are closely related concepts, so we have decided to use the term “toposequence” for the remainder of this paper. In adopting this term, we aim to use it to encompass all kinds of catenary relationships, so as to avoid the confusion about different facets of what is, essentially, a line on a map.

1.3. Stream order

The drainage basin, or drainage area, is the area of land that contributes water to a channel or set of channels (Leopold et al., 1964). The drainage network is the pattern of streams (tributaries and main streams) in a particular drainage basin, and focuses the movement of water throughout the drainage area (Petts and Foster, 1985).

The drainage network can be thought of as a hierarchy of stream segments, or “links”. We can designate each stream segment an order according to its position in the hierarchy, and numerous stream-ordering systems have been developed. In Strahler’s (1957) system, first-order streams are those that have no tributaries. Where two first-order segments meet, the common, adjoining, segment becomes a second-order segment (Fig. 2). Where two second-order segments meet, the follow-

![Fig. 1. Two-dimensional profile view of a catena after Milne (1936). Numerals 1–7 represent different soils.](image-url)
ing downstream segment is a third-order segment, and so on.

We call the area of land that drains directly to a stream segment a *streamshed*. Based on the concept of stream order, above, we can give each streamshed an order. We call this a *kth-order streamshed*. Thus we can have first- and second-order streamsheds, and so on as required for a particular network (Fig. 2). We can characterise the soils of streamsheds using the toposequence sampling approach, and compare and investigate the relationships between streamsheds of different order.

The streamshed concept is different to a drainage basin in that second-order *basins*, for example, consist of the area of land that drains directly to them *as well as* the area of land taken up by their constituent first-order basins. First-order basins are a subset of second-order basins, whereas first- and second-order streamsheds do not intersect.

### 1.4. Digital soil mapping

Digital soil maps (DSMs) are essentially soil class and property maps whose method of production is entirely digital. In a DSM, soil classes or properties are predicted using any of a variety of statistical methods including co-kriging, regression kriging and other regression techniques, among others. An advantage of DSMs is that because they are predicted *quantitatively*, it is possible to also produce a map of *uncertainty* of the predictions (McBratney et al., 2003).

DSMs should not be confused with *digitised* soil maps, which may be scanned versions of pre-existing paper maps produced by traditional soil survey methods. These are usually qualitative, as the maps are based on an “expert-driven” approach which relies heavily on the surveyor’s experience and knowledge/interpretation of the situation at hand, among other factors (Behrens et al., 2005).

Soil is a function of the environment in which it is found. To make a DSM, it is essential that we establish a statistical relationship between soil and environment. In distilling previous work, McBratney et al. (2003) suggested that soil is a function of seven factors (the so-called *scorpan* factors), as follows:

\[
S = f(s, c, o, r, p, a, n),
\]

where *S* is soil (*S_c* for soil classes or *S_a* for soil attributes or properties) at a point; *s* refers to existing soil information; *c* represents climate; *o* represents organisms or biological activity; *r* represents topographic or landscape attributes; *p* represents parent material; *a* represents age; *n* represents spatial position.

Topography greatly influences soil characteristics and soil formation processes. Today, we can use digital raster grids to represent topographic and environmental variables relatively simply and cheaply. For example, we can derive slope, landscape curvature and flow direction, among others, from a digital elevation model (DEM). These variables are myriad, and are also known as “terrain parameters” or “topographic attributes”, among other names (Hengl et al., 2004). We can represent other environmental variables such as rainfall and temperature on raster grids as well.

We can employ various statistical methods to quantify the relationship between soil and environmental variables at known locations. Once a statistical relationship is known, the soil type or property of interest can be predicted onto a grid of “unknown” locations at a suitable resolution to produce a DSM.
1.5. How do we make toposequences?

A toposequence is a monotonically downhill line transect (not necessarily a straight one) that begins at a hilltop and ends at a stream or valley bottom. We can visualise it in a geographical information system (GIS) as a line linking these two points (Fig. 3), using a DEM to provide the elevation data.

Derivation of a digital toposequence from a DEM begins at a point on a landscape and a path uphill and/or downhill is found according to a set of rules: for example, we define the principal toposequence as one that follows the steepest path downhill from the top of a hill, but there are other kinds of toposequences. The toposequence that takes a random path downhill is called a random toposequence.

The grid nature of DEMs allows toposequence derivation to be simple, because the landscape is divided into a grid of spatially discrete units at a uniform resolution. It only involves the analysis of a single $3 \times 3$ cell neighbourhood of elevation data to determine the next move.

1.6. Toposequences and soil sampling

We can use toposequences to determine soil sampling locations for collecting soil data for use in the production of DSMs. They provide a defined corridor along the landscape from which to sample the soil (see Fig. 3), so the challenge is to determine where along the toposequence to sample. The methods that we suggest here are based on the point-intercept approach (Johnson, 1961), in which samples are taken at regular intervals along a transect. Care must be taken to specify the appropriate interval size (Webster and Cuanalo de la Cerda, 1975). This method is often used in the soil sciences, and examples include Steers and Hajek (1979), Buchter et al. (1991) and Davis et al. (2004).

The simple methods that we suggest are the equal distance interval method and the equal elevation interval method. Essentially, we use these methods to determine sampling locations at intervals of approximately equal distance and equal elevation along a toposequence, respectively. We use the endpoints of the toposequence as the first and last sampling points.

![Fig. 3. A digital toposequence is essentially a line linking a hilltop with a stream or valley bottom. Possible sampling locations are indicated using red dots.](image-url)
More methods can be defined. The preference, if any, of one method of sampling over another for different types of landscape needs to be investigated. A method’s suitability probably depends upon the modelling application at hand and the type of hillslope that is under study. For example, the equal elevation interval sampling method may be better suited to hillslopes with concave or convex profile, whereas the equal elevation or equal distance interval methods may both be useful for hillslopes with relatively constant gradient. Coram (1998) presented typical hillslope profiles found in Australia.

Other kinds of methods for sampling along toposequences can be devised, too. Johnson (1961) also defines a line-intercept sampling method in which he counted the number of paces between soil-type boundaries along a transect. In this case, the soil that exists along the transect is already known, or the type boundaries are readily discernible.

This kind of sampling raises questions: Do we sample at the type boundaries? Or do we sample in the middle of the type? What difference does it make to the data we will collect? Line intercept sampling is commonly used in ecological studies (for example, Butler and McDonald, 1983; Pontius, 1998; Ringvall and Ståhl, 1999), but we have not been able to find any examples of its use in the soil sciences besides Johnson’s description.

2. Methods

The model presented in this paper was written in ArcInfo Macro Language (AML), a procedural programming language for the automation of complex mapping tasks. AML was developed by the Environmental Systems Research Institute and implemented in its ArcInfo software.1

The advantage of generating toposequences using computer software is that the process is automated and, hence, requires human input only at the beginning of the procedure. The computer is able to process large numbers of toposequences much faster than a human could. Using a Pentium 4 3.20 GHz Windows PC with 1 Gb RAM, it takes approximately 2.5 h to process 1000 toposequences.

2.1. Generating a kth-order random toposequence

The toposequences that are produced by this model are called \textit{kth-order random toposequences}: that is, each toposequence originates from a randomly generated seed point in the landscape, finds a random path to the top and bottom of a hillslope and ends at a stream of known stream order, \( k \). We use the random toposequence and not the principal toposequence because we are interested in the path that a falling drop of water, say, could potentially take downhill, rather than the path it might preferentially take. The path will be more tortuous but less biased.

Each toposequence is essentially a series of points, which can be connected to form a line transect (Fig. 3). It is convenient to describe the operation of the model in two sections (a schematic outlining the model is presented in Fig. 4):

- **Uphill section**: The section of the toposequence drawn uphill from the randomly generated seed point.
- **Downhill section**: The section of the toposequence drawn downhill from the randomly generated seed point.

The model requires a DEM and a stream order grid. Raster resolution ought to be carefully chosen. This is particularly important for the DEM, so as not to eliminate important features in the study landscape if it is too coarse, or to introduce excessive “noise” if it is too fine. A disadvantage of modern, high-resolution DEMs is that there is often difficulty in distinguishing between actual depressions and artefacts caused by DEM production (Lindsay and Creed, 2006). These issues must be considered when selecting a DEM resolution.

It is important that the stream order grid accurately reflects the location of actual streams in the study area—i.e. if it is derived from a DEM using an automated process, it must be sufficiently “ground-truthed” via a map or some other means. The model does not distinguish between Strahler stream order, Shreve magnitude (Shreve, 1966) or any other ordering system.

A seed point is randomly generated from within the extent of the study area. We could align it with the closest DEM cell centre, but this would not be totally random. However, the points that comprise a toposequence have the same resolution as the DEM they are derived from, and individual points

---

inherit the elevation and stream order value of the cell they are located in. Fig. 5 demonstrates this: in the figure, the toposequence points are located to the northeast of the centre of each grid cell.

Hence, if the elevation of the cell whose coordinate is (340300, 6370975) in Fig. 5 is 500 m, the elevation of toposequence point (340306, 6370985) will be 500 m. For a given toposequence
there can only be one toposequence point within a single DEM/stream order grid cell.

2.1.1. Uphill section

Once the data are loaded, the model generates a seed point. As mentioned previously, seed points are generated by randomly selecting easting and northing values from within the extent of the DEM using the minimum and maximum values of easting and northing ascertained when the DEM was loaded. A number of checks are performed on the generated seed point to make sure it meets the appropriate requirements. A seed point is rejected if it meets any of the following conditions:

- **If it occurs in a cell on the edge of the DEM**: The objective of the modelling exercise is to produce complete toposequences—i.e., so that we have toposequences that start at hilltops and end at streams or valley bottoms.
- **If its elevation value is NODATA**: This can occur if the DEM is irregularly shaped (i.e., not rectangular): a point can be selected that is within the overall extent of the DEM, yet is not within the elevation surface.
- **If its value in the stream order grid is NODATA**: Similar to the previous condition, or if the stream order grid has a smaller spatial extent than the DEM.

Once we have a suitable seed point, the process of drawing a toposequence begins. The seed point becomes what is known as the focus point: the point that is currently being worked on. If the focus point has an elevation less than or equal to all its eight immediate neighbour DEM cells, a number of paths can be taken.

The first option is to follow the first condition in the list above: we do not accept focus points that occur at the edge of the DEM. If we accept a toposequence that extends past the edge of the DEM, we do not have a complete toposequence.

Secondly, a flat area is designated if the maximum elevation of the focus point’s neighbours is equal to that of the focus point. Flat areas represent lakes or perhaps are anomalies in the elevation surface, and given this uncertainty we decided to reject toposequences that reach these areas.

Otherwise, a random uphill point is chosen from the current focus point’s immediate neighbours in the DEM by finding all the immediate uphill neighbour cells. If an uphill neighbour cell is a stream cell, it will not be considered a candidate for a new focus point. An uphill neighbour cell is randomly chosen from those that are available, and it becomes the location of the new focus point. If the new focus point is at the edge of the grid, the toposequence is discarded.

The toposequence reaches the top of a hill when the elevation of the focus point is greater than the elevation of the surrounding neighbour DEM cells. Once the uphill section is complete, the downhill section of the toposequence is drawn from the seed point.

2.1.2. Downhill path

If the seed point of the current toposequence is located in a stream cell, a check is performed to determine if the stream order matches the desired stream order specified by the user at the beginning of the model run. If the stream order matches the desired stream order, the toposequence is accepted. If it does not match, the toposequence is rejected.

If the current toposequence is acceptable, the DEM neighbour cells of the current focus point that are downhill are determined. Two options are available here.

Firstly, if the minimum elevation in the neighbourhood is greater than or equal to the focus point elevation, the downhill path has reached a depression. We reject toposequences that end in a depression so that ambiguity regarding anomalous depressions is eliminated.

Otherwise, the toposequence can proceed further downhill if the minimum DEM elevation in the current focus point’s neighbourhood is less than the elevation of the focus point. A downhill neighbour cell is selected to become the location of the new focus point. It cannot be a stream cell or an edge cell. The toposequence is rejected if the stream cell occurs at the edge of the DEM.

The downhill path continues in this way until the randomly selected downhill neighbour cell is a stream cell. If the stream cell is not of the desired stream order (specified by the user), the toposequence is rejected. If the stream cell is of the desired order, the toposequence is accepted and saved.

Once the desired number of toposequences is derived, information about each toposequence’s points is stored in a text file. The file contains information about each point’s coordinate, elevation and other attributes.
2.2. Optimisation of spatial distribution of toposequences

The configuration of a set of toposequences is the spatial arrangement of the toposequences in a particular set. We want a final set of toposequences to be as far apart in space from each other as possible so that we can sample across as much of the study area as we can. One way of keeping track of this is by summing the total distance, $E$, between every toposequence hilltop in the set. Ideally, we would like each toposequence to finish at a unique hilltop.

We could maximise the total sum of hilltop distances by replacing individual toposequences in the set at random and accept only configurations whose total sum of distances ($E_{\text{current}}$) is greater than the previous accepted sum ($E_{\text{previous}}$), but this can lead to local clustering of toposequences within the overall spatial distribution. If toposequences are clustered in regions, we will inevitably not capture as much of the natural variation in landscape characteristics as possible. This will lead to weaker statistical soil prediction models later on.

Another method of optimisation is to use simulated annealing as described in Press et al. (1992), which is based on the Metropolis algorithm (Metropolis et al., 1953). In our case it attempts to reduce or eliminate local clustering by periodically accepting a configuration of toposequences with a lesser sum of hilltop distances (a “worse” configuration) than the previous configuration. The model begins by asking the user for the number, $n$, of toposequences they wish to derive from the DEM. A set of $n$ acceptable toposequences is derived before the optimisation process begins so that there are sufficient toposequences for $E_{\text{current}}$ to be calculated for the first time. Therefore, if 50 toposequences are desired, there will be at least 50 iterations of the model before optimisation can begin. Depending on the size, shape and hydrological characteristics of the study area, there may be many more than $n$ pre-optimisation iterations once rejected toposequences are taken into account.

Once optimisation commences, the model iteratively tests different configurations of toposequences and accepts or rejects a particular configuration based on a set of rules, outlined below. A new configuration of toposequences is created by selecting a toposequence at random from the previously accepted set and replacing it with a new one. Only one configuration is tested in each iteration.

All configurations of toposequences with a positive change in sum of hilltop distances (positive $\Delta E$; “better”) are accepted. Acceptance of “worse” (negative or zero $\Delta E$) configurations is controlled by a probability, $P$, based on the change in the sum of hilltop distances ($\Delta E$) and the “temperature” of the model, $T$. Temperature ranges between 1 and 0, and is incrementally decreased from 1 after each iteration of the model during the annealing process. Probability is calculated using the Boltzmann factor:

$$P = e^{(\Delta E/kT)}.$$  \hspace{1cm} (2)

The constant $k$ (Boltzmann’s constant) relates $E$ to temperature, $T$. Once $P$ is calculated, it is compared to a randomly generated decimal number between 0 and 1. If $P$ is greater than the randomly generated decimal, the current “worse” configuration is accepted and the total sum of hilltop distances for this configuration will become the $E_{\text{previous}}$ value that is compared in the next iteration of the annealing schedule. If, however, the $P$ for the configuration under consideration is less than or equal to the randomly generated decimal, the configuration is rejected and the next iteration of the model begins.

3. Results

3.1. Study area

We demonstrate use of the toposequence model on a 16 km $\times$ 16 km area of the lower Hunter Valley in New South Wales, Australia. The study area is located to the north of the town of Cessnock, and consists mostly of undulating hills that ascend to low mountains in the southwest corner. We used a 25 m DEM from the New South Wales Land Information Centre. The 25 m resolution is suitable for this study area as it preserves the main landscape features of the area. We used the Terrain Analysis System\textsuperscript{2} to remove depressions and to create the stream network grid.

---

3.2. Model run

Sixty-four toposequences were derived. Two runs of the toposequence model were made: one to produce 32 first-order toposequences and one to produce 32 second-order toposequences according to Strahler's stream-ordering system (Strahler, 1957). A map of the toposequences and their streamsheds are displayed in Fig. 6. The results show that the toposequences are spread fairly evenly across the study area. Local clustering has been minimised, with only one pair of first-order and four pairs of second-order toposequences sharing common hilltops.

4. Discussion

4.1. The model

The model presented in this paper derives toposequences from a DEM using ESRI’s ArcInfo software. It uses an iterative approach so that large numbers of toposequences can be processed easily in two parts: the uphill section is drawn first, followed by the downhill section. An optimisation procedure is integrated with the model so that local clustering of toposequences can be minimised. The procedure uses simulated annealing to maximise the total sum of distances between all toposequence hilltops. Many configurations of toposequences are compared by randomly substituting one toposequence after another, so not every toposequence that is produced in a single model run is accepted.

The model seems to work well, and the major bugs have been ironed out; however, patience is a requirement when creating large sets of toposequences. ArcInfo does not process AML as fast as we would like it to. For this reason we hope to develop a Java-based version of this model in the future. We anticipate this will also make the model more flexible.

4.2. Stream network considerations

The model is able, through trial and error, to produce toposequences that drain only to a stream...
with specified stream order. No work has been done so far to investigate relationships between toposequences and stream order: it is not known whether higher-order stream segments are associated with longer toposequences, for example.

We do know that it is improbable that an entire study area can be covered evenly using only toposequences of one particular stream order. Consider this situation: the drainage network of a particular study area consists mostly of first-order stream segments that drain into only a few second-order stream segments. If we do not create any second-order toposequences, the portion of the study area contained in second-order streamsheds will not be sampled.

In the current version of the model, the user specifies only one desired stream order in a single model run. If the user desires toposequences that end at streams of \( n \) different orders, \( n \) runs of the model are required, each producing toposequences of unique order. The advantage is that the exact number of toposequences of order \( k \) can be specified. It also results in simpler code, and, hence, faster processing time.

### 4.3. Sampling considerations

Using the sampling methods we have suggested, sampling points will be distributed at roughly equal intervals along a toposequence. However, if a given toposequence has a short, steep slope that extends for only 10 m, it may only be sampled once or not at all if, say, equal-distance interval sampling with an interval of 20 m is used.

To avoid the possibility of having to collect soil samples from the middle of streams, it may be useful to take the penultimate toposequence point in the downhill direction as the final point in the toposequence (instead of the toposequence point that occupies the stream cell). If most of the streams in a study are not perennial, this step is probably not necessary. This illustrates one of the pitfalls of determining stream networks from DEMs: estimation of the location of streams is based solely on elevation data (Yin and Wang, 1999), and not photographic interpretation or expert knowledge of which streams are actually “wet”.

The stream network grid must be as accurate as possible so that errors due to anomalous stream segments are minimised. An incorrect stream network grid gives rise to toposequences of incorrect length, especially when the stream network shows streams that are not present in reality. In this case, toposequences will terminate earlier than in reality because a possible path for the toposequence is occupied by a “stream” that should not be there.

The presence of anomalous tributaries poses problems for the allocation of stream orders, too. For example, stream segments can be misclassified when a “real” headwater tributary and an anomalous headwater tributary converge to form a second-order tributary.

### 4.4. Possible uses of the model

In this paper, we have only described the toposequence model, how it works and the rationale behind it, but there are a number of possible uses for it. Of course the primary purpose of its creation is to determine sampling locations to collect soil data for DSM applications. For example, following the framework proposed by McBratney et al. (2003), we can use this soil data as the soil (“s”) variable in the scorpam model to spatially predict soil classes or properties.

Perhaps a very valuable use is to validate the accuracy of DSMs. In this case, the model is run to produce toposequences (and, hence, sampling locations) for an initial soil survey. The soil data collected in the initial survey is used to create a DSM then the model is run again to produce a second set of toposequences from which a second set of sampling locations is derived. From these we assemble a second set of soil data and use it to validate the original DSM.

We also ought to consider not only how soil varies from location to location down a hillslope, but also how soil varies at or near a single location. Soil spatial variability is a complex phenomenon and can affect spatial prediction (Heuvelink and Webster, 2001; Lin et al., 2005). We can investigate it by sampling supplemental points at either random or fixed distances from toposequence sampling points.

### 5. Conclusions

The model presented in this paper derives toposequences from a digital elevation model using ESRI’s ArcInfo software. It uses an iterative approach so that large number of toposequences can be processed easily in two parts: first the uphill section is drawn, followed by the downhill section. An optimisation procedure is integrated with the
model so that local clustering of toposequences can be avoided as much as possible. It uses simulated annealing to maximise the total sum of distances between all toposequence hilltops in a set. Many configurations of toposequences are compared by substituting one toposequence after the other, so not every toposequence that is produced in a single run will be accepted. The optimisation procedure provides satisfactory results, which we have demonstrated in our study in the lower Hunter Valley.

We have suggested a number of possible uses for the toposequence model, and it is likely that there are more. Work needs to be done in this regard. The model is, above all, a method for sampling soil along soil toposequences. Using this soil data, we can establish soil–landscape rules by understanding the pedological processes operating on a hillslope. All of this information can be incorporated into predictive models for creating digital soil maps.

Acknowledgements

This work is supported by an Australian Research Council Discovery Project on Digital Soil Mapping. We thank Dr. Hannes Reuter (DG Joint Research Centre, European Commission) and Dr. Robert MacMillan (LandMapper Environmental Solutions, Alberta, Canada) for their comments and suggestions on improving the model.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2007.05.018.

References


