Modelling Three-dimensional Geoscientific Fields with the Voronoi Diagram and its Dual*

Hugo Ledoux
Delft University of Technology
h.ledoux@tudelft.nl

Christopher M. Gold
University of Glamorgan
cmgold@glam.ac.uk

March 25, 2008

Fields as found in the geosciences have properties that are not usually found in other disciplines: the phenomena studied are often three-dimensional, they tend to change continuously over time, and the collection of samples to study the phenomena is problematic, which often results in highly anisotropic distributions of samples. In the GIS community, raster structures (voxels or octrees) are the most popular solutions, but, as we show in this paper, they have shortcomings for modelling and analysing 3D geoscientific fields. As an alternative to using rasters, we propose a new spatial model based on the Voronoi diagram (VD) and its dual the Delaunay tetrahedralization (DT), and argue that they have many advantages over other tessellations. We discuss the main properties of the 3D VD/DT, present some GIS operations that are greatly simplified when the VD/DT is used, and, to analyse two or more fields, we also present a variant of the map algebra framework where all the operations are performed directly on VDs. The usefulness of this Voronoi-based spatial model is demonstrated with a series of potential applications.

Keywords: Three-dimensional GIS, Fields, Spatial modelling, Voronoi diagram, Delaunay triangulation.

1 Introduction

The objects studied in geoscience (e.g. meteorology, geology, oceanography and geophysics) are often not man-made objects, but rather the spatial distribution of three-dimensional continuous geographical phenomena such as the salinity of a body of water, the humidity of the air or the percentage of gold in the rock. These are referred to as *fields*, and are in contrast with the object-view of the world that is popular in geographical information systems (GISs) and related standards (e.g. ISO/TC 211). The object-view approach considers the space as being 'empty' and populated with discrete entities (e.g. a building or a road) embedded in space and having their own properties, while the field-view approach considers space as continuous, and every location in space has a certain property (Peuquet, 1984; Couclelis, 1992; Goodchild, 1992a). Fields are further defined in Section 2.

The modelling of fields as found in geoscience is problematic because they have properties that are usually not present in fields in other disciplines. The special properties come from the fact that it is usually impossible to measure continuous phenomena everywhere, and that we have to resort to

*Published in *International Journal of Geographical Information Science*, vol. 22(5):547–574.

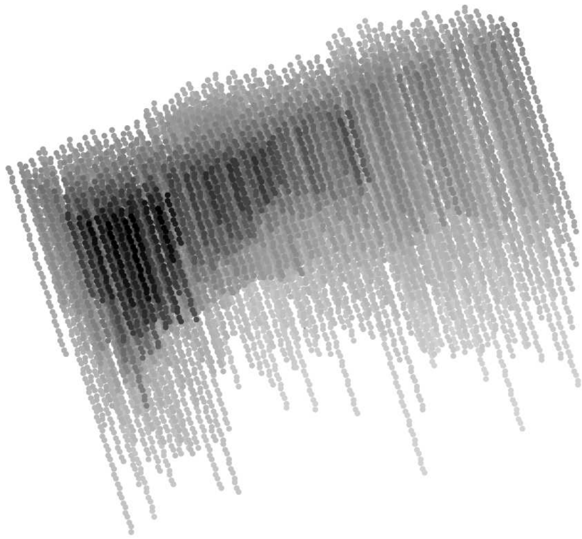


Figure 1: Perspective view of an oceanographic dataset, in which samples are distributed along water columns. Each sample has a location in 3D space ($x-y-z$ coordinates) and one or more attributes attached to it.

collecting samples at some finite locations and reconstructing fields from these samples. Each sample indicates the value of the attribute studied at a certain location $x-y-z$ in space and at a certain time t. In engineering, medicine or computer-aided design for instance, the studied object is usually easily sampled because we have direct access to it; on the other hand, in geoscience, samples can be very hard and expensive to collect because of the difficulties encountered and the technologies involved. To collect samples in the ground we must dig holes or use other devices (e.g. ultrasound penetrating the ground); underwater samples are collected by instruments moved vertically under a boat, or by automated vehicles; and samples of the atmosphere must be collected by devices attached to balloons or aircrafts. Because of the way they are collected, three-dimensional geoscientific datasets often have a highly anisotropic distribution: as shown in Figure 1, the distribution can be abundant vertically but sparse horizontally. Another peculiar property of geoscientific datasets is that the samples they contain represent the spatial variability of a given attribute only at time t, and many geographical phenomena in geoscience change and evolve relatively quickly over time.

Traditional GISs are not suitable to model geoscientific datasets since they have been mostly designed for static and two-dimensional objects. This problem has been identified by many researchers in the geosciences. In oceanography and marine applications, Davis and Davis (1988) state that systems capable of handling three dimensions, plus time and attributes are needed to adequately represent marine phenomena. Li and Saxena (1993) notice that very few artificial objects are found at sea and hence phenomena are mostly represented by points (not lines and polygons), and they also propose a 'wish list' of features that GISs should have, one of them being the possibility to do simulation of oceanographic processes. In meteorology, Bernard *et al.* (1998) and Nativi *et al.* (2004), among others, come to similar conclusions: the representation of the atmosphere requires three spatial dimensions, and a way to integrate time with the spatial dimensions. In geology, the fact that the community has developed its own tools for modelling and analysing data (Houlding, 1994) highlights the fact that current GISs have shortcomings.

Within the GIS community, the raster is the most popular and used structure to represent fields, although in 2D the triangulated irregular network (TIN) is also popular for modelling and analysing elevation data. The popularity of raster structures is probably due to the fact that they are naturally managed in a computer in the form of an array of numbers and also because no complex data structures are required for the location of objects: the position of an object (a pixel) is implicitly known from the sequential position of pixels in the array. However, as explained in Section 3, they have been severely criticised by many because of their theoretical (e.g. unnatural discretisation) and practical (e.g. amount of memory used) shortcomings.

As an alternative to using rasters for representing, modelling and analysing 3D geoscientific fields, we propose in this paper a new spatial model based on both the Voronoi diagram (VD) and its dual the Delaunay tetrahedralization (DT); the two structures, and their relationships, are formally defined in Section 4. We argue that constructing the VD/DT of the samples that were collected to study the field can be beneficial for reconstructing the field and extracting meaningful information from it. A tessellation of space into Voronoi regions offers a natural discretisation where phenomena can be freely represented, and not enforced by a rigid structure like grids. Also, as explained in Section 5, the duality between the VD and the DT can be exploited to help in the manipulation, analysis and visualisation of a field. We also present in Section 5.2 a variant of the *map algebra* framework—which is commonly used in GIS to manipulate several different fields and extract information from them—where every field and every operation is based on the VD. Finally, the usefulness of this Voronoi-based spatial model is demonstrated in Section 6 with a series of potential applications in geoscience.

Note that the ideas of using the VD and a Voronoi-based map algebra to model 3D geoscientific fields were briefly presented in Ledoux and Gold (2006a). The 3D VD was also used in Ledoux and Gold (2006b) for representing and analysing oceanographic datasets, but only one dataset could be modelled at a time and the concept of fields was not investigated. We present in the following a substantial extension to these papers. This extension includes the theoretical background, the motivations behind the use of a Voronoi-based spatial model, the 3D spatial analysis operations possible (including the ones where the DT is explicitly used), the possibility of analysing two or more fields, and a list of concrete geoscientific applications. It is also worth mentioning here that the 3D VD has been briefly discussed as a potential spatial model in Gold and Edwards (1992), and that it has also been mentioned in the past as a solution to model geographical data. However, to our knowledge, its properties and its use to model geoscientific datasets have never been investigated (especially when both the VD and the DT are used to analyse data), and no attempts were made to build one. Although the implementation of the spatial model—admittedly a rather important issue—is not discussed in the following (because of space constraints), all the details can be found in PhD thesis of the first author (Ledoux, 2006), and a summarised version is also available in Ledoux and Gold (2006b).

2 Fields

A field is a concept rather difficult to define because it is not tangible and not part of our intuitive knowledge. It is easy for us to see and describe entities such as houses or chairs, but, although we can imagine fields, they are somewhat an abstract concept. The consequences of that are firstly that formalising a field is difficult, and secondly that many definitions exist in different disciplines (Peuquet *et al.*, 1999). The definition usually used in a GIS context, and the one we use in the following, is borrowed and adapted from physics. Physicists in the 19th century developed the concept of a *force field* to model the magnetic or the gravitational force, where a force (a vector with an orientation and a length) has a value *everywhere* in space, and changes from location to location, and also over time. In this paper, the vector assigned to each point of the Euclidean space is replaced by a scalar value, and we obtain *scalar fields*; we assume in the following that all fields are scalar fields.

Because each point in space possesses a value, a field can be represented mathematically. It is a model of the spatial variation of a given attribute a over a spatial domain, and it is modelled by a function, from $\mathbb{R}^d$ to $\mathbb{R}$ in a d-dimensional Euclidean space, mapping the location to the value of a, thus:

$$a = f(location). \tag{1}$$

The function can theoretically have any number of independent variables (i.e. the spatial domain can have any dimensions), but in the context of geographical phenomena the function is usually bivariate (x, y) or trivariate (x, y, z). Note that the domain can also incorporate time as a new dimension, and *dynamic fields*, such that $a = f(location, time)$, are thus obtained (Kemp, 1993).

2.1 Two Different Types of Fields

Depending on the scale of measurement used for the values of the attribute, two different types of fields are possible:

1. **Continuous scale:** the value of an attribute can have any value. Temperature, precipitation or salinity are examples because they can be measured precisely. The *interval* and *ratio* scales commonly used in GIS, as defined by Stevens (1946), fall into this category. We refer to this type of field as a *continuous field*.

2. **Discrete scale:** the values of an attribute are simply labels. Stevens's *nominal* and *ordinal* scales fall into this category. Nominal values are meaningless, in the sense that they are just labels: an example is a map of Europe where each location contains the name of the country. Ordinal values are labels that can be ordered, e.g. a certain region can be categorised according to its suitability to agriculture from 1 to 5: 1 being poor, and 5 very good. We refer to this type of field as a *discrete field*. Observe that they are actually fields that also act as objects, because each regions can also be modelled as a unique object (Peuquet *et al.*, 1999). Although they are less common, they are still useful for some application domains, as demonstrated in Section 6.

Notice that here the terms "continuous" and "discrete" refer to the scale of attribute measurement, and not to the spatial continuity of a field. Indeed, both types of fields are spatially continuous, as they are represented by a function.

2.2 Digital Representations of Fields

Since fields are continuous functions, they must be *discretised*—broken into finite parts—to be represented in a computer, and to be analysed and visualised. The space covered by a field can be partitioned, or tessellated, *regularly*, *hierarchically*, or *irregularly*. In a regular tessellation, regions all have the same shape and size. The most common regular tessellation in GIS is by far the grid, where the elements are squares in 2D (usually called *pixels* as an analogy to digital images) and cubes in 3D (called *voxels*, a portmanteau of the words 'volume' and 'pixel'). Notice that other regular shapes are possible, such as hexagons or rectangles. To solve the problem with massive grid files, hierarchical tessellations can be used. A commonly used structure in 3D is the *octree*, which is a generic term for a family of tessellations that recursively subdivide the space into eight octants (Samet, 1984). It is easily implemented in a computer because it is a tree where each node has exactly eight children, if any. By contrast, the elements of an irregular tessellation can be of any shape and size; in three dimensions, each element is a polyhedron. Irregular tessellations usually follow the data points (the samples that were collected to study the field for instance), albeit this is not a requirement. Subdividing the space based on the samples has the main advantage of producing a tessellation that is adaptive to the sample distribution and to the complexity of the phenomenon studied.

Once the space is tessellated, the field function becomes a *piecewise function*: to each region is assigned a function describing the spatial variation in its interior. The function within each element is usually a simple mathematical function, and as Goodchild (1992a) points out, this function can be constant, linear, or of a higher order. A constant function means that the value of the attribute is constant within one region, for example to represent a discrete field, as in a choropleth map. An example of the use of a linear function is a TIN: the spatial variation within each region (a triangle) is described by the linear function (a plane) defined by the three vertices lifted to their respective elevation. The value of the attribute of a field at an unsampled location x is thus obtained by linearly interpolating on the plane passing through the three vertices of the triangle containing x. Akima (1978) shows the advantages of using higher order functions in each region of a TIN—the main one being that the slope of the terrain is continuous everywhere, i.e. there are no discontinuities at the border of two triangles. The same functions can obviously be used within each volume element in three dimensions.

Also, an important consideration is that if one possesses a set of samples, then a valid representation of a field is this set of samples together with a spatial function that will permit us to reconstruct the field. Examples of such functions would be the parameters of inverse distance weighted (IDW) interpolation (e.g. the size of the 'searching sphere' in 3D and how the distance influences the weight), the natural neighbour interpolation methods (as described in Section 5.1), or the Kriging parameters, although that would be somewhat problematic in a real-world environment because of the complexity of the method (see Matheron (1971) and Oliver and Webster (1990) for more details).

For the two-dimensional case, some other representations have also been mentioned and used, notably contour lines and irregularly spaced points (the samples to which attributes are attached). In our opinion, the latter representation is incomplete if the spatial function used to reconstruct the field is not explicitly defined, and therefore should not be considered a valid representation of a field.

3 Related Work

Within the GIS community, regular tessellations have been for years more or less synonymous with continuous fields (Goodchild, 1992b), and that in two and three dimensions. A few implementations of a voxel-based GIS exist, notably the open-source system GRASS[1] in which diverse operations such as interpolation and visualisation are possible. Most of the earlier attempts at building 3D GISs were also using voxels, as explained in Raper (1989). Furthermore, examples of the use of octrees in GIS projects are plentiful: Bak and Mill (1989) and Jones (1989) use it for modelling geological structures; Li and Saxena (1993) use it for modelling the three-dimensional objects in their prototype marine GIS; and O'Conaill *et al.* (1992) and Mason *et al.* (1994) show its advantages for many applications in the geosciences.

This wide popularity of regular tessellations is probably due to the fact that they permit us to integrate easily datasets from different disciplines, and also due to simplicity. Indeed, a grid is naturally stored in a computer as an array (each grid cell is addressed by its position in the array, and only the value of the grid cell is stored), and thus the spatial relationships between cells are implicit. The algorithms to analyse and manipulate (boolean operations such as intersection or containment) are also trivially implemented in a computer (assuming that the grids have the same orientation and resolution). On the other hand, raster representations have been severely criticised, both from a theoretical and a technical point of view. The issues involved when using pixels as the main element for geographical analysis are summarised in Fisher (1997). One problem is that a grid arbitrarily tessellates the *space*, without taking into consideration the objects embedded in it or the phenomena modelled. Kemp (1993) states that "[a raster representation] requires us to enforce a structure on reality rather than allowing reality to suggest a more appropriate structure for our analysis". On the technical side, the problems most often cited are: (1) the meaning of a grid is unclear (are the values at the centre of each grid cell, or at the intersections of grid lines?), (2) the size of a grid (if a fine

[1]Geographic Resources Analysis Support System (`http://grass.itc.it`).

resolution is wished, then the size of a grid can become huge, especially in three dimensions), (3) grids scale badly and are not rotationally invariant, (4) an entity is not represented explicitly, but by a group of grid cells with similar values, which means that its boundaries will be jagged. Also, unless a grid comes from an image sensor (remote sensing or 3D images as obtained in medicine), we can assume that it was constructed from a set of samples. Converting samples to a raster structure is dangerous because the original samples—which could be meaningful points, e.g. the boundary between a warm and a cold front in meteorology—are not present in the resulting structure. The samples are thus in a way lost. Also, when a user only has access to a grid, he often does not know how it was constructed and what interpolation method was used, unless meta-data are available. Note also that the shortcomings of quadtrees/octrees are similar to those of rasters: the rotation and scaling operations are difficult to handle, and only a rough approximation of the boundaries of an object is possible.

As in the 2D case with the use of TINs for terrain modelling, the disadvantages of raster representations in 3D have also led to the use of 3D triangulations, called tetrahedralizations since the space is tessellated into tetrahedra, for modelling sub-surface phenomena, see for instance Lattuada (1998) and Xue *et al.* (2004).

4 The 3D Voronoi Diagram and the Delaunay Tetrahedralization

Let S be a set of n points in a d-dimensional Euclidean space $\mathbb{R}^d$. The Voronoi cell of a point $p \in S$, defined $\mathcal{V}_p$, is the set of points $x \in \mathbb{R}^d$ that are closer, or equal, to p than to any other point in S. From a mathematical point of view:

$$\mathcal{V}_p = \{x \in \mathbb{R}^d \mid \|x - p\| \leq \|x - q\|, \; \forall\, q \in S\}. \tag{2}$$

The union of the Voronoi cells of all generating points $p \in S$ form the Voronoi diagram of S, defined VD(S). In two dimensions, $\mathcal{V}_p$ is a convex polygon (see Figure 2(a)), and in 3D it is a convex polyhedron (see Figure 2(b)).

The VD has a geometric dual structure called the Delaunay triangulation, and that in any dimensions. That means that both structures represent the same thing, just from a different viewpoint. In 2D, this structure is defined by the partitioning of the plane into triangles—where the vertices of the triangles are the points generating each Voronoi cell—that satisfy the *empty circumcircle* test. That is a triangle is Delaunay if the unique circle on which lie its three vertices does not contain any other point in the set S; Figure 2(c) illustrates a Delaunay triangulation. The generalisation to 3D of the Delaunay triangulation is the Delaunay tetrahedralization: each triangle becomes a tetrahedron that satisfies the *empty circumsphere* rule.

Let S be a set of points in $\mathbb{R}^d$, and r one of its points. $\mathcal{V}_r$ is unbounded if r bounds the convex hull of S. Observe for instance in Figure 2(a) that p has a bounded Voronoi cell (dark grey cell), but that some points do not (e.g. point r). Although the VD of a set of points will theoretically cover the whole space, the spatial extent for which we can reasonably reconstruct the field is the convex hull of the set of points (outside the convex hull, we are *extrapolating* the values).

The duality between the VD and the DT in $\mathbb{R}^3$ is simple as each element of a structure corresponds to one and only one element in the dual: each polyhedron becomes a point and each line becomes a face, and vice-versa. For example, as shown in Figure 3: a Delaunay vertex p becomes a Voronoi cell (Figure 3(a)), a Delaunay edge α becomes a Voronoi face (Figure 3(b)), a Delaunay triangular face κ becomes a Voronoi edge (Figure 3(c)), and a Delaunay tetrahedron τ becomes a Voronoi vertex (Figure 3(d)). A Voronoi vertex is located at the centre of the sphere circumscribed to its dual tetrahedron, and two vertices in S have a Delaunay edge connecting them if and only if their two respective dual Voronoi cells are adjacent.

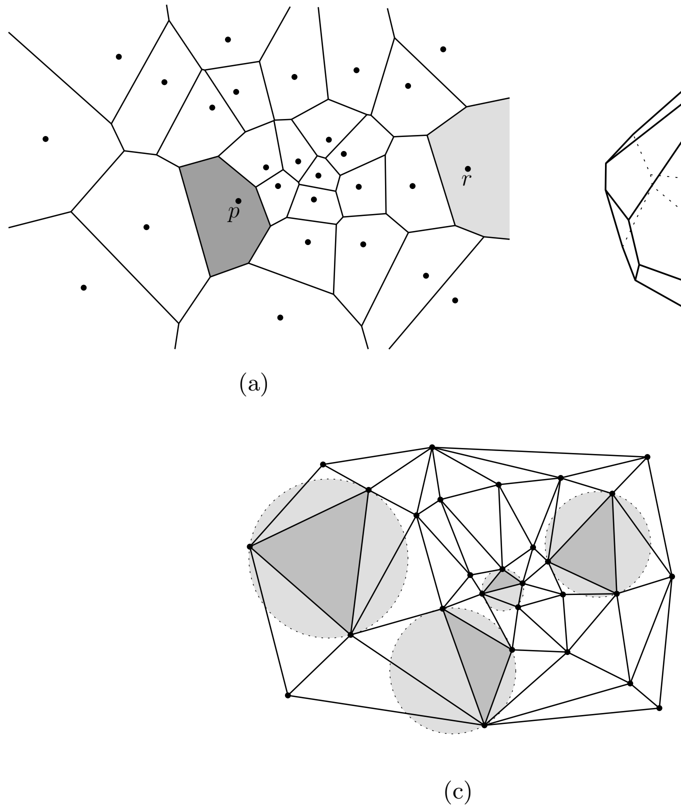


Figure 2: **(a)** The VD for a set of points in the plane. **(b)** Two Voronoi cells adjacent to each other in 3D (they share the grey face). (Figure from Ledoux and Gold (2006b)) **(c)** The DT of a set of points in the plane (same point set as (a)). Every triangle has an empty circumcircle, and for some (grey triangles) they are shown.

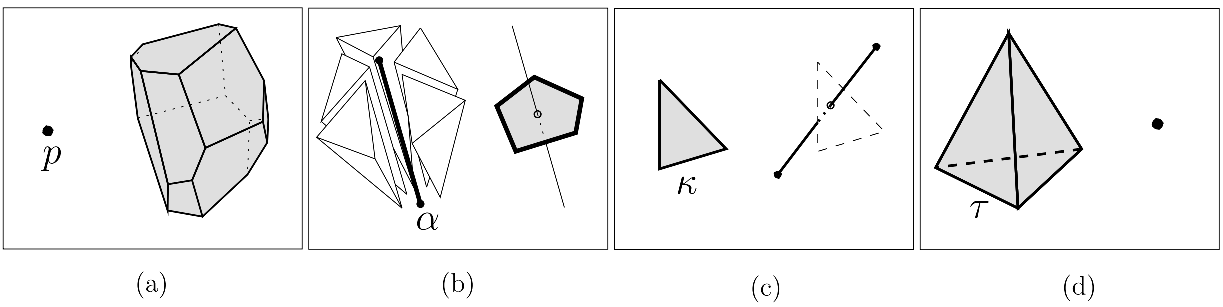


Figure 3: Duality in 3D between the elements of the VD and the DT. (Figure from Ledoux and Gold (2006b))

We can exploit this duality to our advantage to use one subdivision for some purposes, and the other one for some others. For instance, the manipulation operations can be based uniquely on the DT because it is simpler to manipulate triangles/tetrahedra over arbitrary polygons/polyhedra, but the reconstruction can be VD-based because of its neighbourhood properties. Only one of the two tessellations need to be stored as the knowledge of one tessellation always implies the knowledge of the dual tessellation.

4.1 Construction and Manipulation of a 3D VD/DT

Only three operations are necessary to reconstruct a field, to manage the temporal dimension, and to interactively analyse a field: the insertion, the deletion, and the movement of one sample (Mioc, 2002). *Interactive modelling* is an approach where the user does not only use standard operations (e.g. GIS spatial analysis or statistical methods) on a dataset, but actually interacts with it by manipulating, editing and transforming it and then looks at the consequences of his/her manipulations (Anselin, 1999; Bailey and Gatrell, 1995; Gold, 1993). The emphasis is put on the interaction between the user and the dataset through modification and visualisation tools. A key factor for these three operations is the speed at which they are performed, to ensure that the user gets an (almost) instantaneous result from a query or an operation, so that he/she is not disturbed by long waits. One must therefore implement *dynamic* and *kinetic* algorithms, that is algorithms that perform without recomputing from scratch the whole structure.

In 2D, such algorithms are known and relatively easy to implement, but it is not the case in 3D. Indeed, even the simplest of the three operations (the insertion of a single point in order to construct the 3D VD/DT) is far from being straightforward, as Field (1986) and Sugihara and Inagaki (1995) explain. This is mostly caused by the many degenerate (special) cases that arise in 3D geometric computing (Hoffmann, 1989). The description of the algorithms for the three operations is out of scope for this paper, but one should keep in mind that there exist optimal (or near-optimal) algorithms for all these operations, see for instance Edelsbrunner and Shah (1996), Devillers and Teillaud (2003) and Albers (1991). For implementation details, the reader is referred to Ledoux (2006). He shows that the three operations can be implemented with the same set of atomic operations—called the bistellar flips, see Edelsbrunner and Shah (1996) for details—and that this simplifies greatly the process since the maintenance of adjacency relationships in the VD/DT is encapsulated in the atomic operations. A strong emphasis is put on *robustness* of the algorithms, so that they output a correct solution, regardless of the spatial distribution of the points in the input; the many degeneracies that can occur with real-world datasets are described and handled. The practical performances of such an implementation are also shown to be efficient since they are comparable to what is arguably the *de facto* standard in computational geometry implementations, the Computational Geometry Algorithms Library (`CGAL`) (Boissonnat *et al.*, 2002).

4.2 Advantages of Modelling 3D Fields with the VD/DT

Since most fields in geography must first be sampled to be studied, we argue in this paper that the tessellations obtained with VD and the DT have many advantages over other tessellations for representing fields. The main advantages are the following:

1. the VD offers a natural discretisation of space, which is based on the samples that were collected to study the field. The phenomena studied can thus be represented freely, and not enforced by a rigid structure like voxels. The shape and the size of the cells in the VD are adaptive to the spatial distribution of the samples, which is crucial when dealing with highly anisotropic distributions such as the ones found in the geosciences. Observe in Figure 2(a) that where the data distribution is dense the cells are smaller.

2. it gives a clear and consistent definition of the spatial relationships between unconnected points. The one-to-one mapping between the points and the Voronoi cells ensure that the original samples—the "meta-field" according to Kemp and Včkovski (1998)—are kept and not 'lost', as is the case when gridding.

3. the reconstruction of a field can be entirely based on the VD, as explained in the next section. Such a reconstruction is efficient and entirely automatic, and that for discrete and continuous fields.

4. the knowledge of the VD implies the knowledge of its dual the DT, which helps greatly for manipulation operations, and also for some spatial analysis operations.

5. local updates to the spatial model are possible, i.e. insertions, deletions or movements of points in the VD can be made without recomputing the structure from scratch. This is fundamental for the interactive exploration of a dataset, to model dynamic fields, and to maintain large datasets.

6. many GIS and spatial analysis operations in three dimensions, such as the extraction of isosurfaces, are possible and even optimised when the VD/DT of a set of points is constructed, as demonstrated in the next section.

5 GIS Operations for Trivariate Fields

This section describes a few GIS and spatial analysis operations, for trivariate fields, which are based on the VD and/or the DT. Besides the obvious advantages that the indexing of a set of unconnected points brings (e.g. the closest-point queries are efficiently performed), many spatial analysis operations use the properties of the VD/DT to help the user have a better understanding of a dataset. While the VD (or the DT) is not mandatory for all the operations presented, it greatly simplifies and improves performances for many of them. The four operations described are: (1) spatial interpolation, which permits us to reconstruct a field from the samples; (2) analysis of different fields with map algebra; (3) visualisation tools; (4) maintenance of temporal data.

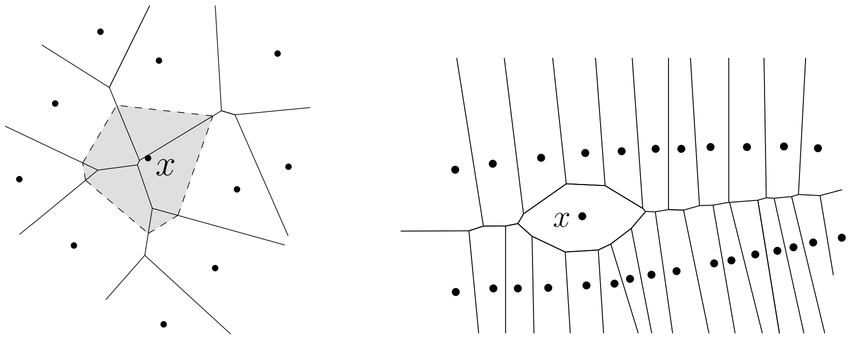
The list of operations presented is not exhaustive, as many other operations based on the VD are possible, especially when other metrics are used. For instance, Okabe *et al.* (1994) present 35 GIS-related operations in two dimensions based on the generalised VD, operations that are theoretically possible in three dimensions. It should also be said that besides the section concerning the map algebra (Section 5.2), the other sections simply reports on methods and algorithms that already exist and that have been published. We report on them for the sake of completedness of the paper.

5.1 Spatial Interpolation

Given a set of samples to which an attribute a is attached, spatial interpolation is the procedure used to estimate the value of the attribute at an unsampled location x. To achieve that, it creates a function f, called the interpolant, that tries to fit the samples as well as possible. Interpolation methods are an essential operation in this paper because they permit us to reconstruct a field from the set of samples that were collected to study it. They are moreover required to model, visualise and better understand a dataset, and also to convert a dataset to another format (e.g. from scattered data to voxels).

If a constant value function is assigned to each Voronoi cell, the VD permits us to elegantly represent discrete fields. To know the value of a given attribute at a location x, one simply has to find the cell containing x—Mücke *et al.* (1999) describe an efficient way to achieving that.

To reconstruct a continuous field from a set of samples, more elaborate interpolation methods are needed since the VD creates discontinuities at the border of each cell. An interesting one in the context



(a) (b)

Figure 4: **(a)** Natural neighbour coordinates of x in the plane. The insertion of x creates the new grey Voronoi cell. **(b)** The VD permits us to select the neighbours that will influence location x.

of this paper is the natural neighbour interpolation method (Sibson, 1981), because it has been shown by different researchers to have many advantages over other methods when the distribution of samples is highly anisotropic: it behaves consistently whatever the distribution of the samples is, and it is an automatic method that does not require user-defined parameters (Gold, 1989; Sambridge *et al.*, 1995; Watson, 1992). This is a method entirely based on the VD for both selecting the samples involved in the interpolation process, and to assign them a weight (an importance). The neighbours used in an estimation are selected using the adjacency relationships of the VD, which results in the selection of neighbours that both surround and are close to x. That helps greatly when the distribution of samples is highly anisotropic; Figure 4(b) shows an example in 2D. The weight of each neighbour is based on the volume (area in 2D) that the Voronoi cell of x 'steals' from the Voronoi cell of the neighbours in the absence of x; Figure 4(a) shows a 2D example, and in 3D the idea is the same. The resulting function is exact (the samples are honoured), and also smooth and continuous everywhere except at the samples themselves. See Gold (1989) and Watson (1992) for further discussion of the properties of the method.

Although the concepts behind natural neighbour interpolation are simple and easy to understand, its implementation is far from being straightforward, especially in three and higher dimensions. The main reasons are that the method requires the computation of two VDs—one with and one without the interpolation point—and also the computation of volumes of Voronoi cells. Ledoux and Gold (2004) present a simple algorithm that works directly on the DT and uses the concept of flipping for inserting and deleting the interpolation point. The Voronoi cells are extracted from the DT and their volumes are calculated by decomposing them. The algorithm is computationally efficient and considerably simpler to implement than other known methods, as it is based on the algorithm for the construction of the VD/DT.

5.2 Map Algebra

Map algebra refers to the framework, first developed and formalised by Tomlin (1983), to model and manipulate fields stored as grids. It is called an algebra because each field (also called a map) is treated as a variable, and complex operations on fields are formed by a sequence of primitive operations, like in an equation (Berry, 1993). A field operation always takes a field (or many fields) as input and returns a new field as output (the values of the new field are computed location by location). Operations can be unary (input is a single field), binary (two fields) or n-ary (n fields); I describe here only the unary and binary cases because n-ary operations can be obtained with a series of binary operations. For the sake of simplicity, operations are mostly described and illustrated in 2D, but the framework and all

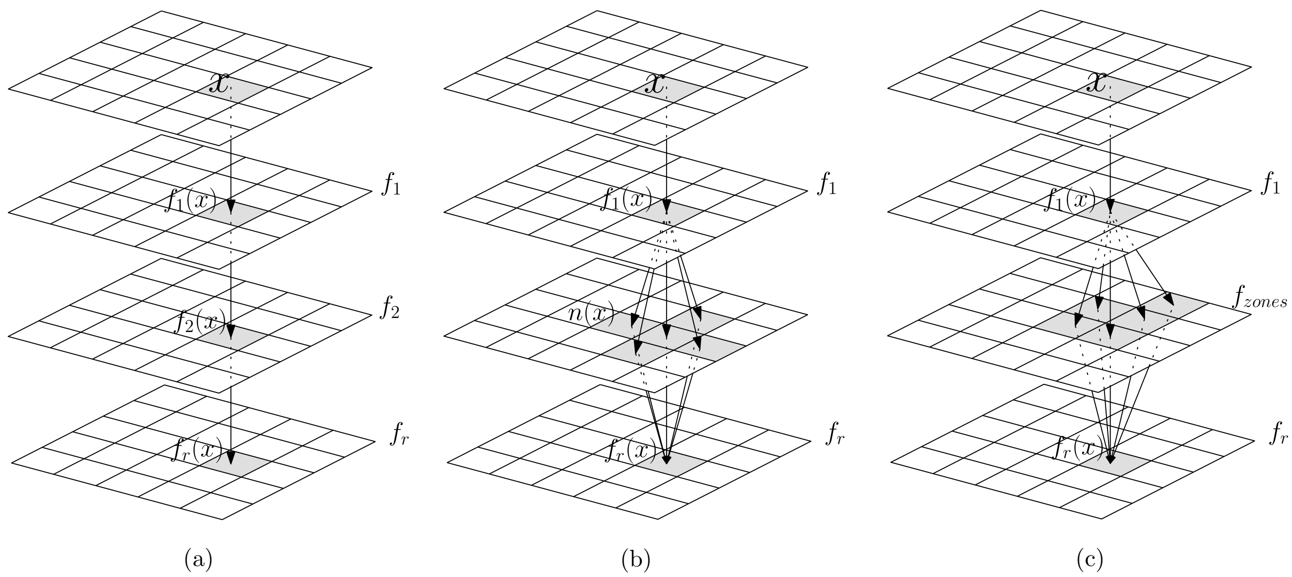


Figure 5: The map algebra operations with a raster structure. **(a)** A binary local operation. **(b)** An unary focal operation. **(c)** A zonal operation that uses a set of zones (f_{zones}) stored as a grid.

the operations are valid in any dimensions; Mennis *et al.* (2005) and Karssenberg and de Jong (2005) have proposed recently voxel implementations. Tomlin (1983) describes three categories of operations:

1. **Local operation:** (Figure 5(a)) the value of the new field at location x is based on the value(s) of the input field(s) at location x. An unary example is the conversion of a field representing the elevation of a terrain from feet to metres. For the binary case, the operation is based on the overlay in GIS: the two fields f_1 and f_2 are superimposed, and the result field f_r is pointwise constructed. Its value at location x, defined $f_r(x)$, is based on only $f_1(x)$ and $f_2(x)$. An example is when the maximum, the average or the sum of the values at each location x is sought.

2. **Focal operation:** (Figure 5(b)) the value of the new field at location x is computed as a function of the values in the input field(s) in the neighbourhood of x. As Worboys and Duckham (2004) describe, the neighbourhood function $n(x)$ at location x associates with each x a set of locations that are 'near' to x. The function $n(x)$ can be based on distance and/or direction, and in the case of raster it is usually the four or eight adjacent pixels. An unary example is the derivation of a field representing the slope of a terrain, from an elevation field.

3. **Zonal operation:** (Figure 5(c)) given a field f_1 and a set of *zones*, a zonal operation creates a new field f_r for which every location x summarises or aggregates the values in f_1 that are in a given zone. The set of zones is usually also represented as a field, and a zone is a collection of locations that have the same value (e.g. in a grid file, all the adjacent cells having the same attribute). For example, given a field representing the temperature of a given day across Europe and a map of all the countries (each country is a zone), a zonal operation constructs a new field such that each location contains the average temperature for the country.

Although the operations are arguably simple, the combination of many makes map algebra a rather powerful tool. It is indeed being used in many commercial GISs, albeit with slight variations in the implementations and the user interfaces (Bruns and Egenhofer, 1997).

Despite its popularity, the biggest handicap to the use of map algebra is arguably that is was developed for regular tessellations only, although the concepts are theoretically valid with any tessellation of space (Takeyama, 1996; Worboys and Duckham, 2004). The shortcomings of raster structures to represent fields, as described earlier, all apply to map algebra. A further important consideration is that in order to perform binary operations, the two grids must 'correspond', i.e. the spatial extent, the resolution and the orientation of the two grids must be the same, so that when they are overlaid each pixel corresponds to one and only one pixel in the other grid. If the grids do not correspond, then *resampling* of one grid (or both) is needed. This involves the interpolation of values at regularly distributed locations with different methods such as nearest neighbour or bilinear interpolation, but unfortunately each resampling degrades the information represented by the grid (Gold and Edwards, 1992).

5.2.1 A Voronoi-based Map Algebra

Because the map algebra framework and its use of regular tessellations have been severely criticised, many have proposed improvements. For example, Takeyama (1996) and Pullar (2001) both propose to extend the concept of neighbourhood of a location x with the use of templates, and they show how they can help to solve several practical GIS-related problems. Recognising that grids are the weakest part of map algebra, Kemp (1993) shows that alternative representations (e.g. TINs and contour lines) are viable solutions. She proposes to have operations—similar to map algebra's—for modelling fields, which are not all stored under the same representation. She therefore defines a set of rules to convert the different types of fields to other ones when binary operations are applied. For example, if two fields, one stored as a TIN and the other as contour lines, are analysed then the contours must first be converted to TIN before any manipulation is done. Haklay (2004), also to avoid the drawbacks of raster structures, proposes a system where only the data points (samples) and the spatial interpolation function used to reconstruct the field are stored. Each field is thus defined mathematically, which allows the manipulation of different fields in a formulaic form.

We present here a variant of map algebra where every field and every operation is based on the VD. This eliminates the need to first convert to grids all the datasets involved in an operation (and moreover to grids that have the same orientation and resolution), as the VD can be used directly to reconstruct the fields.

Unary Operation When a field is represented by the VD, unary operations are simple and robust. Indeed, as shown in the previous section, discrete fields can be elegantly represented directly with the VD where a constant value function is assigned to each cell, and continuous fields can be efficiently reconstructed with natural neighbour interpolation. Also, the neighbouring function needed for focal operations is simply the natural neighbours of every location x, as defined in the previous section. Figure 6(a) shows a focal operation performed on a field f_1. Since at location x there are no samples, a data point is temporarily inserted in the VD to extract the natural neighbours of x (the generators of the shaded cells). The result, $f_r(x)$, is for example the average of the values of the samples; notice that the value at location x is involved in the process and can be obtained easily with natural neighbour interpolation.

Binary Operation Although Kemp (1993) claims that "in order to manipulate two fields simultaneously (as in addition or multiplication), the locations for which there are simple finite numbers representing the value of the field must correspond", we argue that there is no need for two VDs to correspond in order to perform a binary operation because the values at any locations can be obtained readily with interpolation functions. Moreover, since the VD is rotationally invariant (like a vector

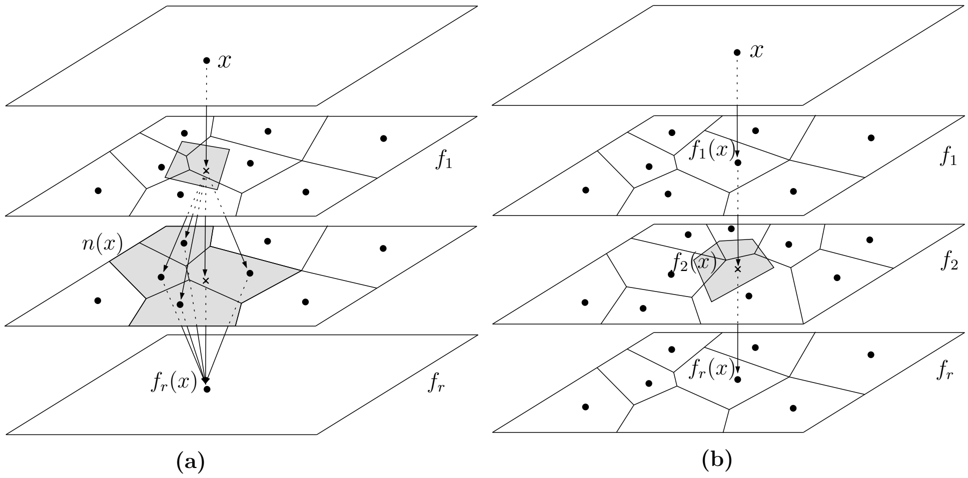


Figure 6: Two Voronoi-based map algebra operations. The top layer represents the spatial extent of the fields, and x is a location for which the value in the resulting field f_r (bottom layer) is sought. **(a)** A unary focal operation performed on the field f_1. The third layer represents the neighbourhood function $n(x)$. **(b)** A binary local operation performed on the fields f_1 and f_2. The resulting field f_r has points at the same locations as in f_1.

map), we are relieved from the burden of resampling datasets to be able to perform operations on them.

When performing a binary operation, if the two VDs do not correspond—and in practice they rarely will do!—the trickiest part is to decide where the 'output' data points will be located. Let two fields f_1 and f_2 be involved in one operation, then several options are possible. First, the output data points can be located at the sampled locations of f_1, or f_2, or even both. An example where the output data points have the same locations as the samples in f_1 is shown in Figure 6(b). Since there are no samples at location x in f_2, the value is estimated with natural neighbour interpolation. The result, $f_r(x)$, could for example be the average of the two values $f_1(x)$ and $f_2(x)$.

It is also possible to randomly generate a 'normal' distribution of data points in space (e.g. a Poisson distribution) and to output these. But one should keep in mind that in many applications the samples can be meaningful, and it is therefore recommended to always keep the original samples and if needed to densify them by randomly adding some data points. The VD also permits us to vary the distribution of data points across space, for example having more data points where the terrain is rugged, and less for flat areas.

The influence of the distribution of output points on the resulting field is difficult to assess and very much application- and case- dependent. We can however state that in the worst case (outputting points only at locations where no samples existed in the input fields), the VD would give conceptually the same results as when regular tessellations are used. Indeed, the grids that are input in an operation also had to be created, which probably means that the original samples were 'lost'.

Zonal Operation Since the local and focal operations are rather straightforward when the fields are stored with the VD, the only operation left to discuss is the zonal operation. As with the other map algebra operations, a zonal operation must also output a field because its result might be used subsequently as the input in another operation. With a Voronoi-based map algebra, the output has to be a VD, and the major difficulty in this case is that we must find a VD that conforms to (or approximates) the set of zones. Since zones come from many sources, different cases will arise. A first example is a voxel file representing for instance the soil type. Such a dataset can easily be converted

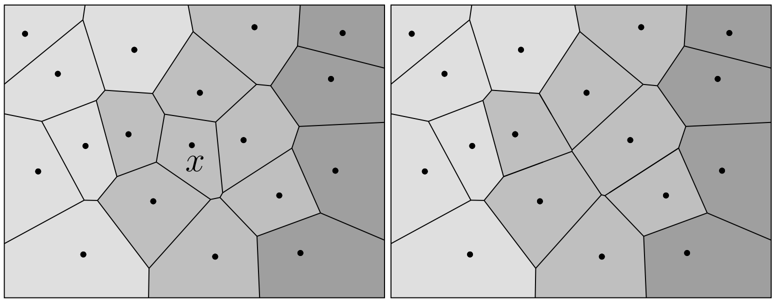


Figure 7: Simplification of a discrete field represented with the VD. The natural neighbours of the data point x all have the same attribute as x (here the value is defined by the colour), and deleting it does not change the field.

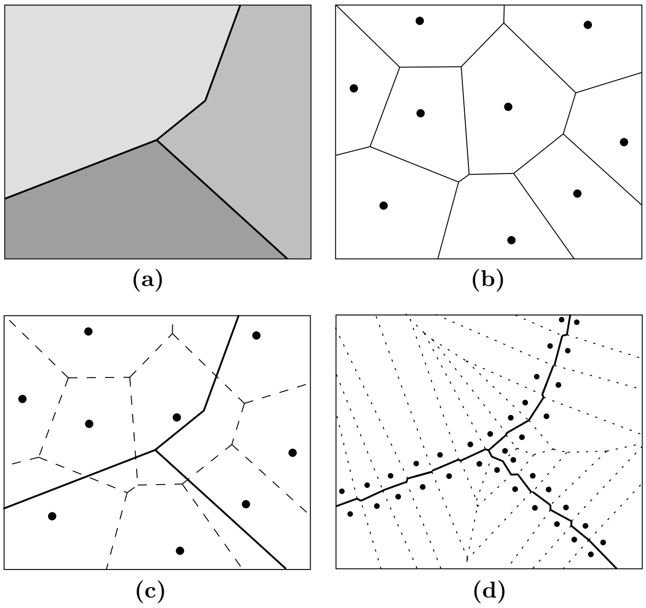


Figure 8: Zonal operations with a Voronoi-based map algebra. **(a)** A vector map of three zones. **(b)** A continuous field represented with the VD. **(c)** When overlaid, notice many Voronoi cells overlap the zones. **(d)** Approximation of the borders of the zones with the VD.

to a VD: simply construct the VD of the centre of every voxel. Although this results in a huge VD, it can easily be simplified by deleting all data points whose natural neighbours have the same value. Notice in Figure 7 that the deletion of a single point is a local operation, and that the adjacent cells will simply merge and fill up the space taken by the original cell (the same is true in 3D). A second example is a set of arbitrary zones, such as a vector map of Europe. In two dimensions, it is possible to approximate the zones with a VD (where each zone is mapped to one Voronoi cell), but the algorithm is complex and the results are not always satisfactory (Suzuki and Iri, 1986); to our knowledge, no 3D counterpart exists. A simpler option is to define a set of "fringe" points on each side of a line segment (randomly insert points close to and on each side of a line), and label each point with the value associated to the zone. Gold *et al.* (1996) show that the original line segment can be reconstructed automatically from the VD of fringe points: the Voronoi edges having two neighbours with different labels approximate the line segment. An example is shown in Figure 8: a set of three zones appears in Figure 8(a), and in Figure 8(d) the Voronoi edges for which the values on the left and right are different are used to approximate the boundaries of the zones. Since each location x in the output field of a zonal operation summarises the values of a field in a given zone, we must make sure that the locations used for the operation are sufficient and distributed all over the zone. Let us go back to the example of the temperature across Europe to find the average in each country. Figure 8(a) shows a vector map with three countries, and the temperature field f_1 is represented by a VD in Figure 8(b). Notice that when the two datasets are overlaid (Figure 8(c)), many Voronoi cells cover more than one zone. Thus, simply using the original samples (with a point-in-polygon operation) will clearly yield inaccurate results. The output field f_r, which would contain the average temperature for each country, must be a VD, and it can be created with the fringe method (Figure 8(d)). Because the value assigned to each data point correspond to the temperature for the whole zone, we suggest estimating, with the natural neighbour interpolation, the value at many randomly distributed locations all over each zone.

5.3 Visualisation

It is notoriously difficult to visualise trivariate fields, even if a three-dimensional computer environment offering translation, rotation and zoom is available. The major problem is that unlike bivariate fields, where the attribute a can be treated as another dimension, we can not 'lift' every location x_i by its value a_i to create a surface in one more dimension and visualise/analyse it—we can not see in four dimensions!

We therefore have to resort to other techniques to be able to visualise the spatial variation of trivariate fields and to extract meaningful information from them. Within the field of *volume visualisation* (Kaufman, 1996), there are mainly three approaches: (1) direct volume rendering; (2) extraction of isosurfaces; (3) slicing.

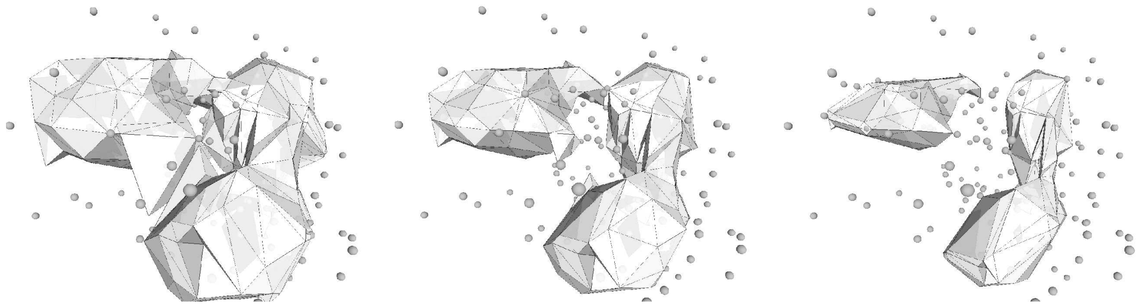
The first approach permits us to visualise the whole dataset at once by projecting the information onto a single plane (the computer screen). The mapping is done by assigning at every location in the field a colour and an *opacity*, and then creating a single image where the colour of a given pixel is obtained by accumulating the colour attributes of the locations that project to this pixel. This approach yields good results, but is also complex, the rendered image takes long to compute, and algorithms are usually based on voxel data. Hence, since the VD/DT is not relevant to this technique, only the approaches of isosurfaces and slicing are discussed in the following.

5.3.1 Extraction of Isosurfaces

Given a trivariate field $f(x, y, z)$, an isosurface is the set of points in space where $f(x, y, z) = a_0$, where a_0 is a constant. Isosurfaces, also called *level sets*, are the 3D analogous concept to isolines (also called contour lines), which have been traditionally used to represent the elevation in topographic maps. Of course, in practice, isolines and isosurfaces are only approximated from the computer representation of a field.

In 2D, isolines are usually extracted directly from a TIN or a regular grid. The idea is to compute the intersection between the level value (e.g. 200m) and the terrain, represented for instance with a TIN. Each triangle is scanned and segment lines are extracted to form an approximation of an isoline.

The same idea can be used in three dimensions to extract surfaces from a field, see the three examples in Figure 9. The principal and most known algorithm for extracting an isosurface is the *Marching Cubes* (Lorensen and Cline, 1987). The isosurface is computed by finding the intersections between the isosurface and each cell of the representation of a field, which is assumed to be voxel. Linear interpolation is used along the edges of each cube of a voxel representation to extract 'polygonal patches' of the isosurface. There exist 256 different cases for the intersection with a cube (considering that the value of each of the eight vertices of a cube is 'above' of 'under' the threshold), although if we consider the symmetry in a cube that comes down to only 15 cases. The major problem with



(a) (b) (c)

Figure 9: An example of an oceanographic dataset where each point has the temperature of the water, and three isosurface extracted (for a value of respectively 7°C, 8°C and 9°C) from this dataset.

Marching Cubes is that the isosurface may contain 'holes' or 'cracks' when a cube is formed by certain configurations of above and under vertices (Wilhems and van Gelder, 1990).

Although it is possible to fix the ambiguities, the simplest solution is to subdivide the cubes into tetrahedra (Cignoni *et al.*, 1998). The so-called *Marching Tetrahedra* algorithm is very simple: each tetrahedron is tested for the intersection with the isosurface, and triangular faces are extracted from the tetrahedra by linear interpolation on the edges. The resulting isosurface is guaranteed to be topologically consistent, i.e. will not contain holes. The nice thing about the algorithm is that only three cases for the intersection of the isosurface and a tetrahedron can arise: (1) the four vertices have a higher (or lower) value—no intersection; (2) one vertex has a higher (or lower) value, hence the three others have a lower (or higher) value. Three intersections are thus defined, and a triangular face is extracted; (3) two vertices have a higher (or lower) value and the others have a lower (or higher) value. Four intersections are thus defined. To ensure that triangular faces are extracted (better output for graphics cards), the polygon can be split into two triangles, with an arbitrary diagonal.

5.3.2 Slicing a Dataset

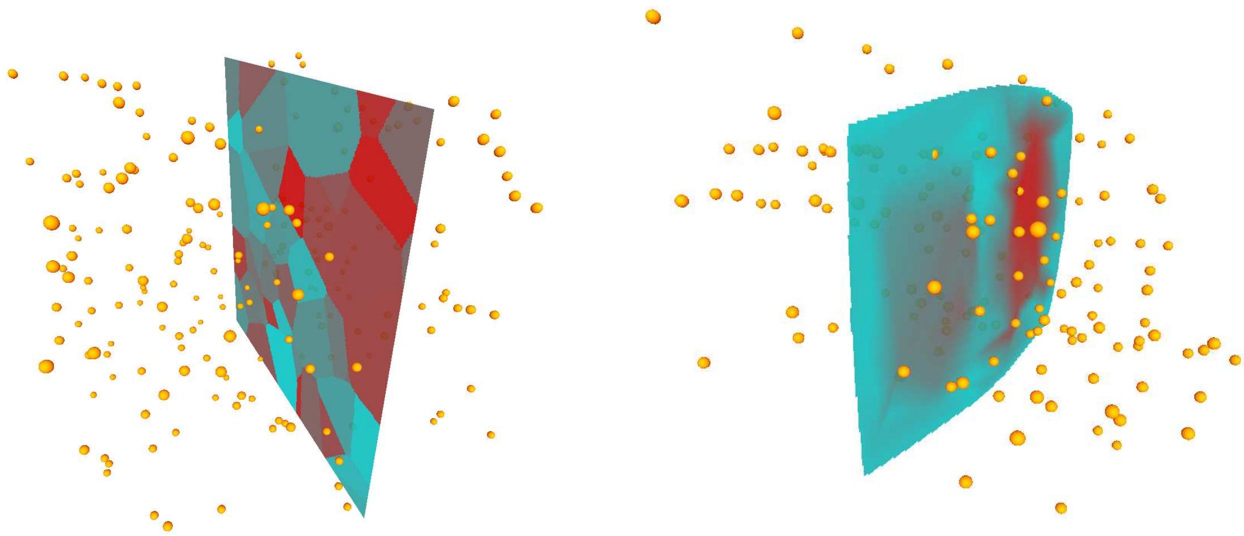
Although the slicing method is rather simple, it can still help users to have a greater understanding of a dataset, and therefore reduce the analysis time. Slicing involves 'cutting' a dataset according to a certain plane, and then visualising the spatial variation of a given attribute on that plane, for instance with the help of colours.

For a discrete field represented with the VD, it would for example mean slicing the Voronoi cells and outputting on the screen the intersections with a colour that is based on the value of the attribute assigned to each cell. Figure 10(a) illustrates one example; notice that even if the cells might look like Voronoi cells in two dimensions, they are not.

Figure 10(b) shows an example where a continuous field (represented by its samples), is sliced according to a certain plane, but instead of just using the Voronoi cells, interpolation is performed at regular intervals on the plane with natural neighbour interpolation (but any other method could also be used). In the case here, red means that the value of the attribute is higher.

5.4 Temporal Data / History Maintenance

The VD permits insertion, deletion and movement of points with local modifications only, thus every operation is reversible. As shown in Gold (1996) and Mioc (2002), by simply keeping a 'log file' of every operation done it is possible to rebuild each 'topological state' of a VD, at any time. This represents one way of solving the problems GISs have with temporal data. There is no need to keep various 'snapshots' of the data at different times for further analysis: when a representation of a field at a



(a) (b)

Figure 10: **(a)** A dataset, representing a discrete field, sliced according to a plane. Each colour represents one Voronoi cell. **(b)** A dataset, representing a continuous field, sliced according to a plane.

specific time is required, it is reconstructed from the original data and from the log file. A dynamic field can also be viewed like a 'movie' of the changes that occurred during a certain period of time, provided of course that such information is available.

6 Applications in Geoscience

A spatial model based on the VD and the DT, with the manipulation and analytical functions mentioned in Sections 4.1 and 5, is a generic tool that can be used in a wide variety of domains and applications. Indeed, any applications where trivariate fields—represented by their samples—are involved can benefit from the properties of the VD and the DT. Moreover, many application domains where the datasets are point-based can also be modelled and analysed with the VD/DT. The major impediment to the use of the VD to model three-dimensional geoscientific datasets is that faults and/or discontinuities in the data and phenomena are not allowed. As a consequence, many geological applications can not use the spatial model directly, although some have used tetrahedralizations for specific problems in geology, see for instance Lattuada (1998) and Xue *et al.* (2004).

It should be noted that most scientific papers related to the use of GIS to model and analyse three-dimensional geoscientific datasets focus on methods to circumvent the static and 2D structures of traditional systems. The analytical methods they present are in two dimensions, as they almost in every case slice the datasets and analyse each slice separately. Examples of such papers are plentiful in the surveys of Chapman and Thornes (2003), Sui and Maggio (1999) and Valavanis (2002) concerning the use of GIS in the geosciences. Consequently, discussions concerning 'real' three-dimensional modelling of datasets, and of three-dimensional spatial analysis operations, are scarce in the literature.

This section presents some real and potential applications where the properties of the VD/DT are useful to model fields, both continuous and discrete. This is by no means an exhaustive survey, and we are confident that many more applications of the spatial model will appear in the future.

6.1 Simulations

Different techniques can be used to perform the simulation of a real-world process. Most simulations of processes in the ocean (e.g. tracking of pollution plumes) and in the atmosphere (e.g. dispersion models) are made with the finite difference method (FDM) or the finite element method (FEM), both usually performed on voxel structures. The FDM performed on grids, as used by systems for weather forecasting, is well-known, efficient and mostly accurate. However, the use of grids can sometimes lead to unreliable results (see Augenbaum (1985) for a few examples), and some other technical problems also arise (for instance the curvature of the Earth is problematic for large datasets).

An interesting alternative to FDM is the Free-Lagrange method (FLM) (Fritts *et al.*, 1985). With this method, the flow being simulated is approximated by a set of points (called particles) that are allowed to move freely and interact, and a tessellation of space is kept up-to-date as points are moving (this is fundamental in order to discretise the continuous flow). Each point has a fixed mass and a velocity. Any tessellation is theoretically possible, but the VD has desirable properties because of the shapes of its cells, and also because, as points are moving, the changes in the Voronoi cells are 'smooth'. In other words, topological events arise gradually by the addition/deletion of a face to a cell (Fritts *et al.*, 1985); by contrast, if a triangulation is used to represent the flow, abrupt changes occur (the topological operations required, the flips, change drastically a triangulation, see for instance Roos (1991)). Mostafavi and Gold (2004) propose using the VD for the FLM because of the shapes of the cells obtained, but also because of the kinetic properties of the VD. Indeed, earlier implementations of the FLM were very slow because the adjacency relationships between cells had to be rebuilt at each step of the process. With the kinetic VD, all the topological events are managed locally, and the time steps that were previously used (which could lead to overshoots and unwanted collisions) can be avoided as topological events are used. They show the advantages of the kinetic VD with the simulation of global tides on the Earth (thus using the VD on a sphere).

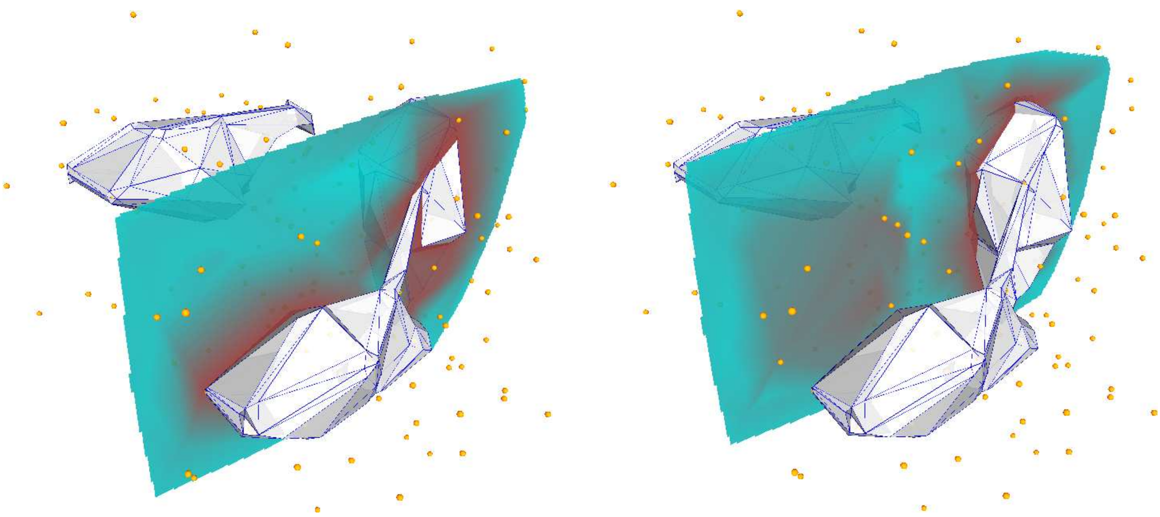
The FLM based on the VD could obviously be used in three dimensions, provided that we can formalise the physical forces applying to every location in space. Because the movement of points in a VD is rather computationally expensive (when all the points are moving simultaneously), the simulation of atmospheric or oceanographic phenomena on a large scale might not be the most suitable examples right now—we want to obtain the weather forecast for tomorrow today! A representative example is the simulation of underground water, for instance for a city. Questions such as "where does the ground water come from?", "how does it travel?", and "where do water contaminants come from, and where are they going?", can all be answered if we can adequately model the phenomenon.

The spatial model presented here has already been used, by Dr Mostafavi at the Université Laval, Québec City, Canada (Mostafavi, 2006), for the development of a prototype GIS modelling underground water. His team is currently working on defining the governing equations to obtain the vector and the velocity of every point in three dimensions, and it is hoped that the kinetic VD will yield results that are more accurate than the ones with methods currently used.

6.2 Visualisation of Attributes

Tools that help with the visualisation, as described previously, can be combined, and with the help of new techniques developed in recent years in computer graphics, it is possible for instance to draw many isosurfaces and view them all by using 'transparency' techniques, assigning different colours to each, 'peeling off' surfaces and navigating inside and outside to see the shape. For instance, Figure 11 shows one case where a cutting plane and an isosurface for a high value of the attribute are displayed. With the help of such tools and visualisation tools to 'navigate' in the scene, a user can understand better the spatial variation of the attribute being studied.

Head *et al.* (1997) describes several visualisation operations useful in the context of oceanographic datasets represented with grids, but the same operations are possible for any type of geoscientific



(a) (b)

Figure 11: Two slices taken at different locations in the same field (red indicates a high value). For both cases, an isosurface for a high value is also shown.

datasets and are even optimised when the VD/DT is used, as shown in Section 5.3.

6.3 Oceanography

The most common type of volumetric data in the marine environment are CTD data (**c**onductivity, **t**emperature and **d**epth of the water), which were traditionally gathered through water columns. Volumetric information about the seawater can also be collected with remotely operated vehicles and autonomous underwater vehicles (Nichols *et al.*, 2003). In both cases the samples collected have highly anisotropic distributions.

Examples of applications in oceanography have already been covered in a previous paper (Ledoux and Gold, 2006b), and we present here only a summary. Briefly, three main applications were presented. The first one is the concept of **real-time applications**, i.e. when data are collected at sea, quickly processed and then added directly to the GIS. This concept permits us to obtain directly at sea a first check of the quality of a survey, and to correct mistakes or collect new data when some are missing (Nichols *et al.*, 2003). The second one is **biogeography**, which is the science that deals with the spatial distribution of species. It helps us understand where animals and plants live, and tries to understand why they live there; its link with GIS is obvious. The third one is the **study of upwellings**, which refers to the movement of subsurface water to the surface. Su and Sheng (1999) studied this process by visualising water properties over a certain period of time. Different isosurfaces, for each property, were created and then animated; the location and pattern of these isosurfaces helped scientists to determine upwelling characteristics. For all these examples, the dynamic and kinetic properties of the VD as presented in this article could be very useful. The gridding of data (as it is usually performed) could be circumvented, and the data visualised and analysed directly from the VD/DT.

6.4 Meteorology

Datasets collected to study the properties of the atmosphere are closely related to those in oceanography. Although collecting data in the atmosphere is somewhat simpler than underwater, obtaining a complete representation of the atmosphere is still difficult. The vast majority of datasets in meteorology have a highly anisotropic distribution similar to that in oceanography (Betancourt, 2004).

As is the case in oceanography, three-dimensional datasets are usually broken into many two-dimensional datasets to be analysed and visualised (Chapman and Thornes, 2003), or converted to voxels (Bernard *et al.*, 1998; Nativi *et al.*, 2004). A spatial model based on the VD/DT to represent and analyse atmospheric data would basically bring the same advantages as in oceanography: the gridding of the samples could be omitted and the data could be analysed directly. The dynamic/kinetic properties of the spatial model are particularly useful here as the atmosphere tends to change quickly over a short period of time—the analysis and weather forecasting must usually be performed quickly.

Trajectory of Air Parcels An early concrete example of application where the VD would be beneficial is given in Barjenbruch *et al.* (2002). They claim that the study of the trajectories of air parcels—represented by points in 3D space—is crucial in the weather forecasting process. These trajectories are computed from scientific models, and their visualisation and analysis can inform the operator about significant coming events, such as heavy rain. The visualisation of parcels has traditionally been done in two dimensions, but they show a computer program where the trajectories are viewed in a three-dimensional environment. They also state that understanding where and why the parcel travelled is important (with respect to the attributes of the atmosphere, such as temperature or humidity) so they colour the trajectories based on certain attributes. It is important to state here that all their datasets are first converted to grids. The air parcels moving in the atmosphere can be modelled with the kinetic VD, as explained previously for other application domains.

Simulation of Winds Another example of a meteorological application is presented in Ciolli *et al.* (2004): the simulation of the temperature and the wind strength near the surface of the Earth, for instance in a valley. That project is actually a good example of the integration of environmental models and GIS. The model in this case is local, i.e. the attributes of the air and the wind are obtained from the slope of the terrain (which is calculated from a digital terrain model) and solar radiation at different times of the day. Ciolli *et al.* used GRASS to build their model, and output 3D grids of the attributes. The grids at different times can be visualised, and also analysed; they use for example three-dimensional map algebra functions to derive grids of some attributes of the air, and they also extract isosurfaces. In a subsequent paper, they admit themselves that grids have shortcomings: "the (3D) raster approach implies that, while it is possible to choose the resolution along the three axes, the variables must be estimated over the whole domain. Since heavy geometric calculation is involved [...] this can be needless burdensome when the values in only a small set of points are needed" (Vitti *et al.*, 2004). They propose instead a variant of their method where only scattered points are output by the model, using the new 3D vector capabilities of GRASS. However, they fail to explain how the fields could be reconstructed from the data points created. To visualise and analyse the results, they output the points at regular intervals, in fact recreating the grid they wanted to avoid in the first place. The VD/DT could help solve this problem.

6.5 Geological Microstructures

Although the spatial model proposed in this paper can not be used directly for all the geological applications (because of overfolds and discontinuities), some geology-related problems can clearly benefit from the VD. One of them is the study of the processes of deformation in rocks, which is usually done at the micro-structure level. We describe briefly in the following an ongoing collaborative project with the Department of Geosciences, Basel University, Switzerland.

The aim of the research project is to study the deformation mechanisms in crystalline rocks, which are formed by different 'phases', i.e. a rock is an amalgam of different minerals. Crystalline rocks

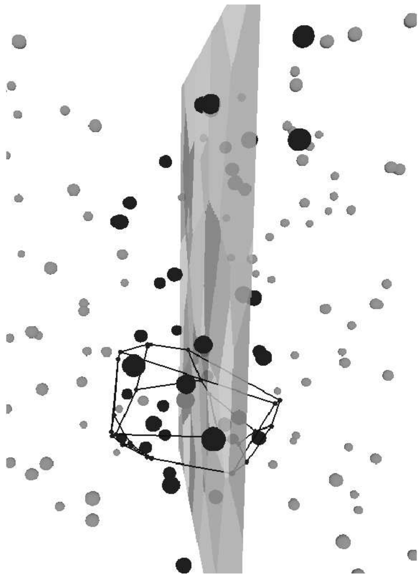


Figure 12: Visualisation of one slice in a discrete field, combined with the display of one Voronoi cell that is intersected by the slice. The dark black data points are the ones whose Voronoi cell is sliced.

can therefore be seen as discrete fields, where a value is assigned to each phase in the rock. The claim of Mackenzie *et al.* (2006) is that the spatial distribution of phases in a rock (the adjacency relationships between phases, and the fact that phases are clustered or not) is a robust indicator of the deformation mechanisms in the rock, because the usual geological processes are not likely to destroy them. Unfortunately, there is a major obstacle to the study of the distribution of phases in 3D: only 2D thin sections of the rock (sliced through the rock) can be collected on-site.

This problem has been tackled in different ways. Kretz (1969) quantifies the spatial distribution of two phases by comparing the boundary fractions (with two phases A and B there will be three types of boundaries between phases: AA, BB and AB) on the thin sections with the expected values if the two phases were randomly distributed in three dimensions. With statistical tests, he can assess if the two phases are clustered or not. His work is based on the fact that the area of the surface boundary between two phases can be estimated to be equal to the length of the boundary between the same two phases on a 2D slice (Underwood, 1970). Also, Jerram *et al.* (1996) build a three-dimensional model of a rock, where the different parts of the rock are represented by spheres (of the same size), which serves as a reference to compare the pattern obtained in the thin sections. This approach has the major shortcoming of not creating a space-filling model, and thus adjacency relationships between phases can not be studied.

The idea of Mackenzie and Heilbronner is to build a computer generated model of the rock in 3D, and to offer the possibility to vary the ratio of phases, and their spatial distribution in the volume. Then 2D slices from the model could be obtained, and they would serve as reference when they are compared to the thin sections of real rocks (Mackenzie *et al.*, 2006).

This is where the VD is useful. By generating the centre of different phases as points in 3D space, the cells of the VD creates a space-filling model, different phases can be easily generated (simply by changing attributes of the generators), and different spatial distributions can be generated also easily. Moreover, the spatial analysis operations described previously can be used to analyse the model. It was found that visualising the model and slices was essential, and that the possibility to view which Voronoi cells are sliced was useful to understand and assess the results; Figure 12 shows one slice, one Voronoi cell that was sliced, and the data points whose Voronoi cells were sliced (the ones in dark black). The surface of contact between two phases in the 3D model can furthermore be computed easily with simple and efficient functions (it is simply the area of the Voronoi faces).

7 Discussion

The choice of a spatial model to represent geographical phenomena is a crucial one that is more than often overlooked by GIS practitioners. As Goodchild (1992b) states: "there exists a multiplicity of possible conceptual data models [spatial models] for spatial data, and the choice between them for a given phenomenon is one of the more fundamental issues of spatial data handling". For a given task, practitioners tend to simply use the spatial models and data structures available in commercial GISs, without assessing the consequences of their choice.

The representation of geoscientific fields in three dimensions and time is particularly problematic because, regardless of a few notable models developed in academia, the only solutions available are raster-based models. The wide popularity of raster structures is probably due to their conceptual simplicity and to the fact that they are easily and naturally stored in computers. However, as argued in this paper, they have also many shortcomings, both conceptually and technically. They force an unnatural discretisation of continuous phenomena and imply a fair amount of preprocessing of datasets, which is usually hidden to the user.

We have presented a new spatial model to represent and analyse trivariate fields as found in the geosciences, i.e. the datasets collected to study fields are point-based. Notice that when a field is already represented with a given tessellation, it is nevertheless always possible to extract points from it. The new spatial model is based on the 3D Voronoi diagram and its dual the Delaunay tetrahedralization. As we have argued in this paper, the tessellations obtained with the VD and the DT have many advantages over other tessellations, particularly over the widely used raster structures. The spatial continuity of fields can be reconstructed with interpolation methods, and many fields analysed with a variant of the map algebra framework where gridding and resampling processes are circumvented.

Although the algorithms to manipulate and analyse a 3D VD/DT are admittedly more complex (and usually also slower) than the ones for raster structures, they are nevertheless readily available—all the algorithms described have been implemented, see Ledoux (2006)—and we believe that the benefits that the VD/DT brings compensates largely (just as modelling elevation with TINs has arguably many advantages over rasters).

Acknowledgements

We would like to thank the financial support of the EU Marie Curie Chair in GIS at the University of Glamorgan.

References

Akima, H., 1978, A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Transactions on Mathematical Software*, **4**, 148–159.

Albers, G., 1991, Three-dimensional dynamic Voronoi diagrams *(in German)*. PhD thesis, Universität Würzburg, Würzburg, Germany.

Anselin, L., 1999, Interactive techniques and exploratory spatial data analysis. In *Geographical Information Systems* (second edn), P.A. Longley, M.F. Goodchild, D.J. Maguire and D.W. Rhind (Eds) (John Wiley & Sons), pp. 253–266.

Augenbaum, J.M., 1985, A Lagrangian method for the shallow water equations based on the Voronoi mesh-flows on a rotating sphere. In *Free-Lagrange method*, M.J. Fritts, W.P. Crowley and H.E. Trease (Eds), 238 (Berlin: Springer-Verlag), pp. 54–87.

Bailey, T.C. and Gatrell, A.C., 1995, *Interactive spatial data analysis* (New York: Wiley).

Bak, P.R.G. and Mill, A.J.B., 1989, Three dimensional representation in a Geoscientific Resource Management System for the minerals industry. In *Three Dimensional Applications in Geographic Information Systems*, J. Raper (Ed.) (Taylor & Francis), pp. 155–182.

Barjenbruch, D.B., Thaler, E. and Szoke, E.J., 2002, Operational applications of three dimensional air parcel trajectories using AWIPS D3D. In *Proceedings of the Interactive Symposium on AWIPS*, Orlando, Florida, USA.

Bernard, L., Schmidt, B. and Streit, U., 1998, ATMOGIS—integration of atmospheric models and GIS. In *Proceedings of the Proceedings 8th International Symposium on Spatial Data Handling*, Vancouver, B.C., Canada, pp. 267–276.

Berry, J.K., 1993, Cartographic modeling: The analytical capabilities of GIS. In *Environmental Modeling with GIS*, M.F. Goodchild, B.O. Parks and L.T. Steyaert (Eds) (New York: Oxford University Press), pp. 58–74.

Betancourt, T., 2004, Some concepts in atmospheric data. In *Proceedings of the The ArcGIS Atmospheric Data Modeling Workshop*, Seattle, USA.

Boissonnat, J.D., Devillers, O., Pion, S., Teillaud, M. and Yvinec, M., 2002, Triangulations in CGAL. *Computational Geometry—Theory and Applications*, **22**, 5–19.

Bruns, H.T. and Egenhofer, M.J., 1997, User interfaces for Map Algebra. *Journal of the Urban and Regional Information Systems Association*, **9**, 44–54.

Chapman, L. and Thornes, J.E., 2003, The use of geographical information systems in climatology and meteorology. *Progress in Physical Geography*, **27**, 313–330.

Cignoni, P., Montani, C. and Scopigno, R., 1998, Tetrahedra based volume visualization. In *Mathematical Visualization—Algorithms, Applications, and Numerics*, H.C. Hege and K. Polthier (Eds) (Spinger-Verlag), pp. 3–18.

Ciolli, M., de Franceschi, M., Rea, R., Vitti, A., Zardi, D. and Zatelli, P., 2004, Development and application of 2D and 3D GRASS modules for simulation of thermally driven slope winds. *Transactions in GIS*, **8**, 191–209.

Couclelis, H., 1992, People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS. In *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, A.U. Frank, I. Campari and U. Formentini (Eds), 639 of *Lecture Notes in Computer Science* (Springer-Verlag), pp. 65–77.

Davis, B.E. and Davis, P.E., 1988, Marine GIS: Concepts and considerations. In *Proceedings of the Proceedings GIS/LIS '88*, Falls Church, VA, USA.

Devillers, O. and Teillaud, M., 2003, Perturbations and vertex removal in a 3D Delaunay triangulation. In *Proceedings of the Proceedings 14th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Baltimore, MD, USA, pp. 313–319.

Edelsbrunner, H. and Shah, N.R., 1996, Incremental topological flipping works for regular triangulations. *Algorithmica*, **15**, 223–241.

Field, D.A., 1986, Implementing Watson's algorithm in three dimensions. In *Proceedings of the Proceedings 2nd Annual Symposium on Computational Geometry*, 246–259 (Yorktown Heights, New York, USA: ACM Press).

Fisher, P.F., 1997, The pixel: A snare and a delusion. *International Journal of Remote Sensing*, **18**, 679–685.

Fritts, M.J., Crowley, W.P. and Trease, H.E., 1985, *The Free-Langrange method*, Vol. 238 (Berlin: Springler-Verlag).

Gold, C.M., 1989, Surface interpolation, spatial adjacency and GIS. In *Three Dimensional Applications in Geographic Information Systems*, J. Raper (Ed.) (Taylor & Francis), pp. 21–35.

Gold, C.M., 1993, Forestry spatial decision support system classification, and the 'flight simulator' approach. In *Proceedings of the Proceedings GIS'93: Eyes on the Future*, Vancouver, Canada, pp. 797–802.

Gold, C.M., 1996, An event-driven approach to spatio-temporal mapping. *Geomatica, Journal of the Canadian Institute of Geomatics*, **50**, 415–424.

Gold, C.M. and Edwards, G., 1992, The Voronoi spatial model: Two- and three-dimensional applications in image analysis. *ITC Journal*, **1**, 11–19.

Gold, C.M., Nantel, J. and Yang, W., 1996, Outside-in: An alternative approach to forest map digitizing. *International Journal of Geographical Information Science*, **10**, 291–310.

Goodchild, M.F., 1992a, Geographical data modeling. *Computers & Geosciences*, **18**, 401–408.

Goodchild, M.F., 1992b, Geographical information science. *International Journal of Geographical Information Systems*, **6**, 31–45.

Haklay, M., 2004, Map Calculus in GIS: A proposal and demonstration. *International Journal of Geographical Information Science*, **18**, 107–125.

Head, M.E.M., Luong, P., Costolo, J.H., Countryman, K. and Szczechowski, C., 1997, Applications of 3-D visualizations of oceanographic data bases. In *Proceedings of the Proceedings Oceans '97—MTS/IEEE*, 2, pp. 1210–1215.

Hoffmann, C.M., 1989, The problems of accuracy and robustness in geometric computation. *Computer—IEEE Computer Society Press*, **22**, 31–42.

Houlding, S.W., 1994, *3D geoscience modeling: Computer techniques for geological characterization* (Berlin: Springer-Verlag).

Jerram, D.A., Cheadle, M.J., Hunter, R.H. and Elliot, M.T., 1996, The spatial distribution of grains and crystals in rocks. *Contributions to Mineralogy and Petrology*, **125**, 60–74.

Jones, C.B., 1989, Data structures for three-dimensional spatial information systems in geology. *International Journal of Geographical Information Systems*, **3**, 15–31.

Karssenberg, D. and de Jong, K., 2005, Dynamic environmental modelling in GIS: 1. Modelling in three spatial dimensions. *International Journal of Geographical Information Science*, **19**, 559–579.

KAUFMAN, A.E., 1996, Volume visualization. *ACM Computing Surveys*, **28**, 165–167.

KEMP, K.K., 1993, Environmental modeling with GIS: A strategy for dealing with spatial continuity. Technical report 93-3, National Center for Geographic Information and Analysis, University of California, Santa Barbara, USA.

KEMP, K.K. and VČKOVSKI, A., 1998, Towards an ontology of fields. In *Proceedings of the Proceedings 3rd International Conference on GeoComputation*, Bristol, UK.

KRETZ, R., 1969, On the spatial distribution of crystals in rocks. *Lithos*, **2**, 39–66.

LATTUADA, R., 1998, A triangulation based approach to three dimensional geoscientific modelling. PhD thesis, Department of Geography, Birkbeck College, University of London, UK.

LEDOUX, H., 2006, Modelling three-dimensional fields in geoscience with the Voronoi diagram and its dual. PhD thesis, School of Computing, University of Glamorgan, Pontypridd, Wales, UK.

LEDOUX, H. and GOLD, C.M., 2004, An efficient natural neighbour interpolation algorithm for geoscientific modelling. In *Developments in Spatial Data Handling—11th International Symposium on Spatial Data Handling*, P.F. Fisher (Ed.) (Springer), pp. 97–108.

LEDOUX, H. and GOLD, C.M., 2006a, A Voronoi-based map algebra. In *Progress in Spatial Data Handling—12th International Symposium on Spatial Data Handling*, A. Reidl, W. Kainz and G. Elmes (Eds) (Springer), pp. 117–131.

LEDOUX, H. and GOLD, C.M., 2006b, La modélisation de données océanographiques à l'aide du diagramme de Voronoï tridimensionnel *(in French). Revue internationale de géomatique*, **16**, 51–70.

LI, R. and SAXENA, N.K., 1993, Development of an integrated Marine Geographic Information System. *Marine Geodesy*, **16**, 293–307.

LORENSEN, W.E. and CLINE, H.E., 1987, Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, **4**, 163–168.

MACKENZIE, J.R., HEILBRONNER, R. and STÜNITZ, H., 2006, Simulating crystalline microstructures—an aid in quantifying the spatial distribution of phases in two-phases crystalline rocks. In *Proceedings of the Proceedings General Assembly 2006 of the European Geosciences Union (EGU 2006)*, Vienna, Austria.

MASON, N.C., O'CONAILL, M.A. and BELL, S.B.M., 1994, Handling four-dimensional georeferenced data in environmental GIS. *International Journal of Geographical Information Systems*, **8**, 191–215.

MATHERON, G., 1971, The theory of regionalised variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleu*, **5**.

MENNIS, J., VIGER, R. and TOMLIN, C.D., 2005, Cubic map algebra functions for spatio-temporal analysis. *Cartography and Geographic Information Science*, **32**, 17–32.

MIOC, D., 2002, The Voronoi spatio-temporal data structure. PhD thesis, Département des Sciences Géomatiques, Université Laval, Québec, Canada.

MOSTAFAVI, M.A., "Personal communication", 2006.

Mostafavi, M.A. and Gold, C.M., 2004, A global kinetic spatial data structure for a marine simulation. *International Journal of Geographical Information Science*, **18**, 211–228.

Mücke, E.P., Saias, I. and Zhu, B., 1999, Fast randomized point location without preprocessing in two- and three-dimensional Delaunay triangulations. *Computational Geometry—Theory and Applications*, **12**, 63–83.

Nativi, S., Blumenthal, M.B., Caron, J., Domenico, B., Habermann, T., Hertzmann, D., Ho, Y., Raskin, R. and Weber, J., 2004, Differences among the data models used by the geographic information systems and atmospheric science communities. In *Proceedings of the Proceedings 84th Annual Meeting of the American Meteorological Society*, Seattle, USA.

Nichols, C.R., Porter, D.L. and Williams, R.G., 2003, *Recent advances and issues in oceanography* (Wesport, USA: Greenwood Press).

O'Conaill, M.A., Bell, S.B.M. and Mason, N.C., 1992, Developing a prototype 4D GIS on a transputer array. *ITC Journal*, **1992**, 47–54.

Okabe, A., Boots, B. and Sugihara, K., 1994, Nearest neighbourhood operations with generalized Voronoi diagrams: A review. *International Journal of Geographical Information Systems*, **8**, 43–71.

Oliver, M.A. and Webster, R., 1990, Kriging: A method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, **4**, 313–332.

Peuquet, D.J., 1984, A conceptual framework and comparison of spatial data models. *Cartographica*, **21**, 66–113.

Peuquet, D.J., Smith, B. and Brogaard, B., 1999, The ontology of fields: Report of a specialist meeting held under the auspices of the VARENIUS project. Technical report, National Center for Geographic Information and Analysis, Santa Barbara, USA.

Pullar, D., 2001, Mapscript: A map algebra programming language incorporating neighborhood analysis. *GeoInformatica*, **5**, 145–163.

Raper, J. (Ed.) , 1989, *Three dimensional applications in Geographic Information Systems* (London: Taylor & Francis).

Roos, T., 1991, Dynamic Voronoi diagrams. PhD thesis, Universität Würzburg, Germany.

Sambridge, M., Braun, J. and McQueen, H., 1995, Geophysical parameterization and interpolation of irregular data using natural neighbours. *Geophysical Journal International*, **122**, 837–857.

Samet, H., 1984, The quadtree and related hierarchical data structures. *ACM Computing Surveys*, **16**, 187–260.

Sibson, R., 1981, A brief description of natural neighbour interpolation. In *Interpreting Multivariate Data*, V. Barnett (Ed.) (New York, USA: Wiley), pp. 21–36.

Stevens, S.S., 1946, On the theory of scales and measurement. *Science*, **103**, 677–680.

Su, Y. and Sheng, Y., 1999, Visualizing upwelling at Monterey Bay in an integrated environment of GIS and scientific visualization. *Marine Geodesy*, **22**, 93–103.

Sugihara, K. and Inagaki, H., 1995, Why is the 3D Delaunay triangulation difficult to construct?. *Information Processing Letters*, **54**, 275–280.

Sui, D.Z. and Maggio, R.C., 1999, Integrating GIS with hydrological modeling: Practices, problems, and prospects. *Computers, Environment and Urban Systems*, **23**, 33–51.

Suzuki, A. and Iri, M., 1986, Approximation of a tesselation of the plane by a Voronoi diagram. *Journal of the Operations Research Society of Japan*, **29**, 69–96.

Takeyama, M., 1996, Geo-Algebra: A mathematical approach to integrating spatial modeling and GIS. PhD thesis, Department of Geography, University of California at Santa Barbara, USA.

Tomlin, C.D., 1983, A map algebra. In *Proceedings of the Proceedings of the 1983 Harvard Computer Graphics Conference*, Cambridge, MA, USA, pp. 127–150.

Underwood, E., 1970, *Quatitative stereology* (Reading, USA: Addison-Wesley Publishing).

Valavanis, V.D., 2002, *Geographic information systems in oceanography and fisheries* (Taylor & Francis).

Vitti, A., Zatelli, P. and Zottele, F., 2004, 3D vector approach to local thermally driven slope winds modeling. In *Proceedings of the Proceedings FOSS/GRASS Users Conference 2004*, Bangkok, Thailand.

Watson, D.F., 1992, *Contouring: A guide to the analysis and display of spatial data* (Oxford, UK: Pergamon Press).

Wilhems, J. and van Gelder, A., 1990, Topological considerations in isosurface generation. *Computer Graphics*, **24**, 79–86.

Worboys, M.F. and Duckham, M., 2004, *GIS: A computing perspective*, Second edition (CRC Press).

Xue, Y., Sun, M. and Ma, A., 2004, On the reconstruction of three-dimensional complex geological objects using Delaunay triangulation. *Future Generation Computer Systems*, **20**, 1227–1234.