HAL
archives-ouvertes.fr

Towards a model for the multidimensional analysis of field data

S. Bimonte, M. Kang

▶ To cite this version:

S. Bimonte, M. Kang. Towards a model for the multidimensional analysis of field data. 14th East-European Conference on Advances in Databases and Information Systems, ADBIS 2010, Sep 2010, Novi Sad, Serbia. 15 p., 2010. <hal-00527729>

HAL Id: hal-00527729
https://hal.archives-ouvertes.fr/hal-00527729

Submitted on 14 Dec 2010

Towards a Model for the Multidimensional Analysis of Field Data

Sandro Bimonte[1], Myoung-Ah Kang[2]

[1]CEMAGREF, Campus des CEZEAUX,
63173 AUBIERE, France
[2]LIMOS-UMR CNRS 6158,ISIMA, Blaise Pascal University,Campus des CEZEAUX,
63173 AUBIERE, France
[1]Sandro.bimonte@cemagref.fr, [2]kang@isima.fr

Abstract. Integration of spatial data into multidimensional models leads to the concept of Spatial OLAP (SOLAP). Usually, SOLAP models exploit discrete spatial data. Few works integrate continuous field data into dimensions and measures. In this paper, we provide a multidimensional model that supports measures and dimension as continuous field data, independently of their implementation.

Keywords: Spatial OLAP, Field data, Spatial Data Warehouses, Multidimensional models

1 Introduction

It has been estimated that about 80% of the data stored in corporate databases integrates geographic information [7]. This information is typically represented according two models, depending on the nature of data: *discrete (vector)* and *field* [22]. The latter model represents the space as a continuous field. Fields have to be discretized to be represented into computers according to data input, and data analysis. These representations can be grouped into two categories: *incomplete* and *complete*. Incomplete representations store only some points and need supplementary functions to calculate the field in non-sampled areas. Complete representations associate estimated values to regions and assume that this value is valid for each point in the regions (raster). Fields are very adapted for modeling spatial phenomena such as pollution, temperature, etc. They allow a point by point analysis through the Map Algebra operators [16] [22].

In order to benefit from Data warehousing and OLAP decision support technologies [11] also in the context of spatial data, some works extended them leading to the concept of Spatial OLAP (SOLAP), which integrates OLAP and Geographic Information Systems (GIS) functionalities into a unique framework [2]. As for the model underlying SOLAP systems, several research issues from theoretical and implementation point of view have risen. Indeed, several works focus on indexing [20] and visualization techniques [6]. Motivated by the relevance of a formal

representation of SOLAP data and operators, some spatio-multidimensional models based on vector data have been defined (a review can be found in [4]). Few works consider field data into multidimensional models [1], [23] and [15]. These models present some limitations that do not allow fully exploiting continuous field data into OLAP from these points of view: *aggregation functions*, *hierarchies based on field*, and *independence of field data implementation*. Thus, in this paper we propose a spatio-multidimensional model integrating field data.

The reminder of this paper is organized as follows: Section 2 recalls fundamentals of spatial analysis techniques, and existing SOLAP models for field data. Our model is proposed in the Section 3. Finally, in Section 4 conclusions and future work are presented.

2 Related work

The term Map Algebra was first introduced in [22] to describe operators on raster data (complete field data). Map Algebra operators are classified according to the number of grids and cells involved. *Local operators* apply a mathematical or logical function to the input grids, cell by cell. *Focal operators* calculate the new value of each cell using neighboring cells values of the input grid. *Zonal operators* calculate new values as a function of the values of the input grid which are associated with the zone of another grid, called zone layer. An extension of the Map Algebra (Cubic Map Algebra) to the temporal dimension is presented in [16]. The authors redefine Map Algebra operators on a cube of cells whose coordinates are three-dimensional. Then, the sets of operators are modified accordingly as shown on Figure 1. In [5] the authors define an algebraic model for field data. This framework provides a formal definition of Map Algebra operators independent of their implementation. Along this line, the work of [8] formally describes how Map Algebra fundamentals can be used for image manipulation. Finally, Map Algebra has been defined also for incomplete field data such as Voronoi tessellations [13]. Indeed, "*Map Algebra obliges analysts to organize reality according to a particular data structure (raster) instead of allowing making the reality suggest the most adequate data structure for the analysis*" [10].

On the other hand, the integration of spatial data into data warehousing and OLAP systems leads to the concept of Spatial OLAP (SOLAP). SOLAP is based on the spatial multidimensional model that defines *spatial dimensions* and *spatial measures*. In particular, a spatial dimension is a classical dimension whose levels contain spatial attributes. This allows for visualizing measures on maps to discover (spatial) relations and unknown patterns. According to the vector model, a spatial measure is defined as a set of geometries and/or the result of spatial operators [2] [14]. Some spatial multidimensional models, based on the vector model, have been proposed [4]. Despite of the significant analysis power of field data and its associated spatial analysis operators, only few works address this issue. [15] extends the concepts of spatial dimension to informally define "spatial matrix (raster) dimension" where at least one level represents raster data. Then, a member is a cell of the raster. Moreover, she introduces also the concept of "matrix cube" where each fact is associated to a cell of the raster data with its attributes. Aggregation functions are Map Algebra operators

(local, focal and zonal). The author limits field data to raster. By this way, the work loses in terms of abstraction for field data. Indeed, Map Algebra has been defined also on other forms of representation of field data such as Voronoi tessellation [13], and terrain representations, such as TIN, that could be very useful for understanding and analyzing multidimensional data [6].
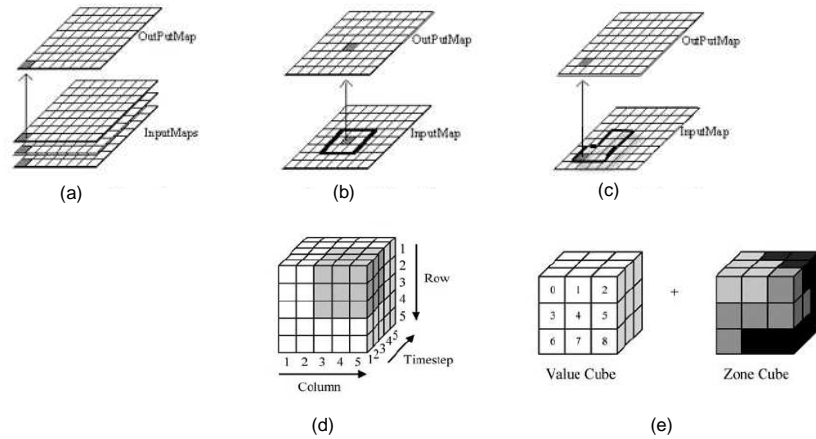

Figure 1. Map Algebra: a) local, b) focal, c) zonal [22], Cubic Map Algebra: d) focal, e) zonal [16].

Hence, [23] proposes a conceptual multidimensional model for taking into account field data independently of their implementation. They define a "field measure" as a value that changes in space and/or time. The model limits aggregation functions for field measure to local Map Algebra functions. They also introduce a "field hierarchy" when a level is a field data that is not linked to facts. Consequently, it is not possible to have field measures at different granularities, which are mandatory for spatial analysis [21]. Finally [1] defines the concept of "continuous cube" when spatial dimensions are composed of infinite spatial members whose associated numerical measures are calculated using interpolation functions, as they use an incomplete representation of spatial members. This model does not allow introduce field data as measures and in hierarchies.

Therefore, a formal model is necessary to create the foundation for a framework for the multidimensional analysis of field data as dimensions (field hierarchies) and measures. Table 1 shows the requirements for this model and how existing works address them. As shown in table 1, there is no existing model that supports all of these requirements. This is the reason why we propose a new SOLAP model in this paper.

Table 1: Requirements for spatial multidimensional model for continuous field data.

Requirements	[1]	[23]	[15]
Measures as continuous field data	NO	YES	Partially (only for raster data)
Hierarchy on continuous field data	NO	NO	Partially (only for raster data)
Aggregation functions as Map Algebra functions	NO	Partially (only local functions)	Partially (only for raster data)
Independence of implementation	YES	YES	NO

3 Spatio-multidimensional model for field data

Before introducing our model we present a SOLAP application for monitoring earthquakes in Italian regions. Spatial analyst wants answer to queries like this: "*Where were earthquakes, and what was their intensity per region at different scales (resolutions)?*". This SOLAP application presents a measure that is a field object representing the earthquakes, a temporal dimension, and a spatial dimension that represents terrain models of Italian regions at different scales.

3.1 Geographic data model

In this section, we provide a uniform representation for field and vector data, which are used by the multidimensional model to define measures and dimensions members.

An Object represents an object of the real world described by some alphanumeric attributes. It is used in the model to represent levels and members (Sec. 3.2).

Definition 1. Object
An Object Structure S_e is a tuple $\langle a_1, \ldots a_n \rangle$ where $\forall i \in [1,\ldots n]$ a_i is an attribute defined on a domain $dom(a_i)$
An Instance of an Object Structure S_e is a tuple $\langle val(a_1),\ldots val(a_n) \rangle$ where $\forall i \in [1,\ldots n]$ $val(a_i) \in dom(a_i)$
We denote by '$I(S_e)$' the set of instances of S_e

A Geographic Object extends an Object to represent geographic information according to the vector model. Indeed, a Geographic Object [3] is a geometry (*geom*) and an optional set of alphanumeric attributes (*[a_1, ...a_n]*) whose values are associated to the whole geometry according to the vector model (Figure 2a).

Definition 2. Geographic Object
Let $g \subset R^2$ i.e. a subset of the Euclidian space. An Object Structure $S_e = \langle geom, [a_1, \ldots a_n] \rangle$ is a Geographic Object Structure if the domain of the attribute geom is a set of geometries: $dom(geom) \in 2^g$
geom is called 'geometric support'

Example 1.

The geographic object structure representing Italian regions is $S_{region}=\langle geom, name \rangle$ where 'geom' is the geometric support, and 'name' is the name of the region. An instance of S_{region} is $t_2 = \langle p_{lo}, Lombardia \rangle$ where 'p_{lo}' is the geometry of the region Lombardia (Figure 2a).

(a)   (b)

Figure 2: a) Italian regions: Instances of S_{region}, b) Earthquakes: an instance of $S_{earthq.}$

According to [13] fields are geographic objects with a function that maps each point of their geometry to an alphanumeric value. This definition allows for representing field data independently of their implementation (complete/incomplete) (Field Object). Thus, A Field Object extends a Geographic Object with a function that associates each point of the geometry to an alphanumeric value. In this way, a Field Object allows for representing geographic data according to the field and the vector model at the same time ("Independence of implementation" requirement of Table 1).

Definition 3. Field Object

Let $S_e = \langle geom, field, [a_1, \dots a_n] \rangle$ a Geographic Object Structure. S_e is a Field Object Structure if the domain of the attribute field is a set of functions defined on m sub-sets of points of geom having values in an alphanumeric domain dom_{field} : $dom(field)= \{f_1 \dots f_m\}$

An Instance of an Field Object Structure S_e is a tuple $\langle g, f_j, val(a_1), \dots val(a_n) \rangle$ where:

− $\forall i \in [1, \dots n]$ $val(a_i) \in dom(a_i)$, $g \in dom(geom)$

− $f_j : g \rightarrow dom_{field}$ and $f_j \in \{f_1, \dots, f_m\}$

We note 'field support' the input domain of f_j

Example 2.

The field object structure representing earthquakes is $S_{earthq}=\langle geom, intensity \rangle$ where 'geom' is the geometric support, and 'intensity' is a set of functions defined on 'geom' with values in R. 'intensity' represents the intensity of the earthquake. An instance of S_{earthq} is $t_2 = \langle p_2, f_2 \rangle$ where p_2 is a geometry and f_2 represents the intensity of the earthquake on the 11-1999 in Lombardia. f_2 is defined on each point of p_2 with values in R, for example $f_2 (x;y) = 12$ (Figure 2b).

By the same way, we can define a field object structure to represent terrain models of Italian regions at different scales by adding to the geographic object S_{region} the field attribute representing terrain elevation.

3.2 Spatio-multidimensional model for field data

A spatio-multidimensional model organizes data using the concepts of dimensions composed of hierarchies, and facts described by measures. An instance of the spatio-multidimensional model is a hypercube. Section 3.2.1 presents the concepts of dimensions, facts, and measures, and Section 3.2.2 formalizes cuboids.

3.2.1 Hierarchies and facts

According to [3] a spatial hierarchy organizes vector objects in a hierarchical way. Formally, a Spatial Hierarchy organizes the Geographic Objects [3] (i.e. vector objects) into a hierarchy structure using a partial order $\leq_h$ where $S_i \leq_h S_j$ means that S_i is a less detailed level than S_j. An instance of a hierarchy is a tree ($<_h$) of instances of Geographic Objects (spatial members). Then, measures are aggregated according to the groups of spatial members defined by the tree $<_h$.

Hence, in this work we define a *Field Hierarchy* as a hierarchy of field objects. For that, we extend the spatial hierarchy by defining a tree ($<_f$) on the geometric coordinates (field supports) of the spatial members represented by field objects ("Hierarchy on continuous field data" requirement of Table 1). By this way it is possible to visualize field objects at different scales or resolutions (Figure 4). Moreover, the alphanumeric values associated to each point of the field measures are aggregated according to the groups of coordinates of spatial members defined by the tree $<_f$. By this way, our model uses the continuous representation of spatial members (Field Objects) to aggregate measure, allowing visualizing field measures at different scales or resolutions (see Figure 6b).

Definition 4. Field Hierarchy
A Field Hierarchy Structure, $\mathcal{H}_h$, is a tuple $\langle \mathcal{L}_h, \lfloor_h, \lceil_h, \leq_h \rangle$ where:
- *$\lfloor_h, \lceil_h$ are of Field Object Structures, and $\mathcal{L}_h$ is a set of Field Object Structures*
- *$\leq_h$ is a partial order defined on $\mathcal{L}_h, \lfloor_h, \lceil_h$ as defined in [3]*

An Instance of a Field Hierarchy Structure $\mathcal{H}_h$ is two partial orders: $<_h$ and $<_f$ such that:
- *$<_h$ is defined on the instances of $\mathcal{L}_h, \lfloor_h, \lceil_h$ as defined in [3]*
  We note $<_h$ 'geographic objects order'

- *$<_f$ is defined on the field supports of the instances of $\mathcal{L}_h, \lfloor_h, \lceil_h$ such that:*
    - *if $cood_i <_f cood_j$ then $S_i \leq_h S_j$, where $cood_i$ belongs to a field support of an instance of S_i, and $cood_j$ belongs to a field support of an instance of S_j ($cood_i$ and $cood_j$ are geometric coordinates)*

    - *$\forall cood_i$ which does not belong to the field supports of the instances of $\lceil_h, \exists$ one $cood_j$ belonging to the field support of an instance of S_j such that $cood_i <_f cood_j$*

    - $\forall cood_i$ which does not belong to the field supports of the instances of $\lfloor_h, \exists cood_j$ belonging to the field support of an instance of S_j such that $cood_j <_f cood_i$

We note $<_f$ field objects order'

The set of leafs of the tree represented by $<_h$ with root t_i are denoted as leafs($\mathcal{H}_{fi}$, t_i). The set of leafs of the tree represented by $<_f$ with root $cood_i$ are denoted as leafsFieldSupport($\mathcal{H}_{fi}$, $cood_i$).

Example 3.
The field hierarchy structure representing the administrative dimension that groups regions into zones is $H_{location} = \langle \mathcal{L}_{location}, S_{region}, S_{all\_location}, \leq_{location} \rangle$ where $\mathcal{L}_{location} = \{S_{zone}\}$ and ($S_{region} \leq_{location} S_{zone}$). S_{region} and S_{zone} are the spatial levels of the hierarchy (Figure 3a). An example of instance of $H_{location}$ is shown on Figure 3b and 3c. We can notice two trees: the geographic objects order that is represented by black lines (Figure 3b), and the field objects order, which is represented by dashed lines (Figure 3c).
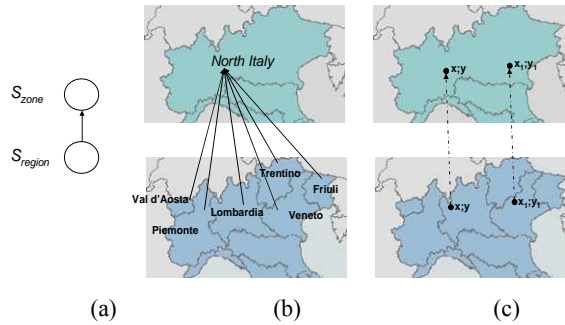

(a)   (b)   (c)

Figure 3: Field hierarchy grouping regions into zones, a) Schema, b) Hierarchical relationships between geographic objects, c) Hierarchical relationships between geometric coordinates.

Example 4.
The hierarchy structure representing the terrain models of Italian regions at different resolutions is $H_{regres} = \langle \mathcal{L}_{regres}, S_{region}, S_{all\_regres}, \leq_{regres} \rangle$ where $\mathcal{L}_{regres} = \{S_{regres}\}$ and ($S_{region} \leq_{regres} S_{regres}$) (Figure 4a). Its instance is shown on Figure 4b and Figure 4c. Note that a geometric coordinate at the coarser resolution is associated with a set of geometric coordinates at the most detailed resolution (Figure 4c). For example, this hierarchy can be defined using the bilinear interpolation algorithms used for changing resolution for raster data at different scales.
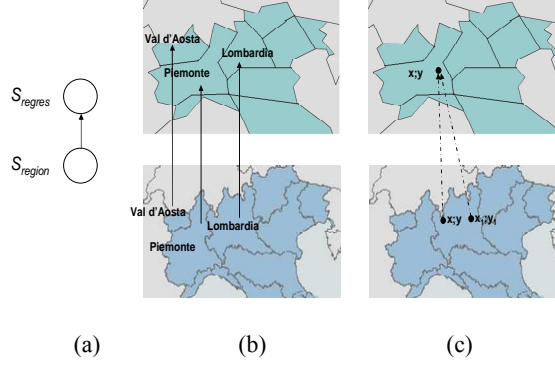
Figure 4: Field hierarchy representing regions at different scales, a) Schema, b) Hierarchical relationships between geographic objects (geographic objects order), c) Hierarchical relationships between geometric coordinates (field objects order).

Once we have introduced the concepts of measures, and hierarchies for field data, we present the concept of Field Cube. A *Field Cube Structure* represents the spatio-multidimensional model schema where a field object is used as measure that is analyzed according classical hierarchies and one field hierarchy ("Measures as continuous field data" requirement of Table 1). Note that without losing in generalization, we suppose to have only one spatial dimension and one field measure to simplify the formalism of the model. In the same way, we do not introduce numerical measures as they can be simply represented using numeric attributes as defined in [3]. An *instance of a field cube structure* represents the fact table or the basic cuboid of the lattice cuboids.

Definition 5. Field Cube
A Field Cube Structure, FC_c, is a tuple $\langle \mathcal{H}_1,\ldots\mathcal{H}_n, FieldObject \rangle$ where:
– *$\mathcal{H}_1$ is a Field Hierarchy Structure* (Spatial dimension)
– *$\forall i \in [2,\ldots n]\ \mathcal{H}_i$ is a Hierarchy Structure* (Dimensions)
– *FieldObject is Field Object Structure* (Field measure)

An Instance of a Field Cube Structure FC_c, $I(FC_c)$, is a set of tuples $\{\langle t\textit{6}_1,\ldots t\textit{6}_n, t\textit{6}_f \rangle\}$ where:
– *$\forall i \in [1,\ldots n]\ t\textit{6}_i$ is an instance of the bottom level of $\mathcal{H}_i$ ($\underaccent{\bar}{L}_i$)* (Most detailed levels members)
– *$t\textit{6}_f$ is an instance of FieldObject* (Field measure value)

Example 5.
The field cube structure of our case study is $FC_{earthq} = \langle H_{regtres}, H_{time}, S_{earthq} \rangle$. $H_{regtres}$ is the field hierarchy (the spatial dimension), H_{time} is the temporal dimension, and S_{earthq} is the field measure. It allows answering previous formulated query "*Where were earthquakes, and what was their intensity per region at different scale (resolutions)?*". Table 2 shows the instance of FC_{earthq}. Its cartographic representation is shown on Figure 5.

Table 2: Instance of FC_{earthq}.

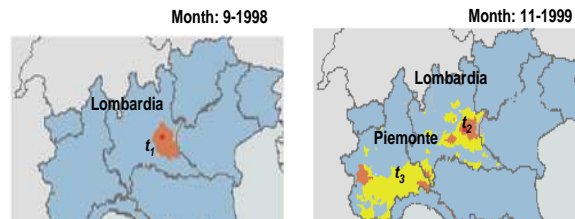| Reg	Month	Earthq
Lombardia	9-1998	t_1
Lombardia	11-1999	t_2
Piemonte	11-1999	t_6

Figure 5: Cartographic representation of the instance of $FC_{earthq.}$

3.2.2   Hypercube

The instance of the spatio-multidimensional model is a hypercube. A hypercube can be represented as a *hierarchical lattice of cuboids* [9]. The most detailed cuboid contains detailed measures (*basic cuboid*). Other cuboids contain aggregated measures. Then, cuboids are represented by levels and (aggregated) measures values (Sec. 3.2.2.2). How field measures are aggregated from fact table data (basic cuboid) to represent non-basic cuboids is presented in Sec. 3.2.2.1.

3.2.2.1 Aggregation of field measures

The aggregation of field measures is defined by means of:
- For the geometric support: spatial aggregation,
- For alphanumeric attributes: alphanumeric aggregations

- For the field attribute:

  - Local Map Algebra, or focal/zonal map cubic algebra operator *when aggregating on the non-field hierarchies*

  - Alphanumeric aggregation *when aggregating on the field hierarchy*

   Indeed aggregation of field measures is done in two steps. In the first step we aggregate along the non-field hierarchies, and then along the field hierarchy.

3.2.2.1.1   Aggregation on non-field hierarchies

   In this section we formalize the geometric and alphanumeric aggregations.

Definition 6. Spatial aggregation
Let G the geometric attribute. Its aggregation is defined by means of a function O_G that has as input n geometries of the attribute G, and that returns one geometry:

$\quad O_G : dom(G) \times ... \times dom(G) \rightarrow 2^g$ *where g is a subset of the Euclidian Space R^2*

Definition 7. Alphanumeric aggregation
Let A be an alphanumeric attribute. Its aggregation is defined by means of a function O_A that has in input n values of the attribute A, and that returns one value of the attribute A:

$\quad O_A : dom(A) \times ... \times dom(A) \rightarrow dom(A)$

On non-field hierarchies, the aggregation of the field attribute is defined by means of a function (O_F) that takes as input a set of functions representing the field attributes values ($f_1...f_n$), and it returns a new function (f_{1n}). This function is defined on the field support of $f_1...f_n$, and the value of each point ($f_{1n}(x,y)$) is calculated by applying a alphanumeric function O_A to the values of the other functions ($O_F (f_1(x,y)...f_n(x,y))$). Then, O_F represents a map/map cubic algebra operator that is specialized in local, focal or zonal by means of the O_A function. Indeed, O_A is applied point by point for local map function, or to sets of coordinates defined by the functions *Neighborhood(x,y)* and *Zone(FieldObjects, (x,y))* for focal map cubic and zonal map cubic operators respectively (("Aggregation functions as Map/Map Cubic Algebra functions " requirement of Table 1)).

Definition 8. Aggregation of the Field attribute on non-field hierarchies using Map Algebra and Cubic Map algebra functions
Let F be a field attribute, and $f_1...f_n$ functions of the domain of F with g as field support (without loss of generality, we suppose the $f_1...f_n$ have the same field support)

$\quad f_1 : g \rightarrow dom_F ... f_n : g \rightarrow dom_F,$ *and* $f_1...f_n \in dom(F)$.
Let O_A be an alphanumeric aggregation

Then, the aggregation of F is defined by means of a function O_F that takes as input $f_1...f_n$, and that returns a function f_{1n} defined on g and having values in dom_F ($f_{1n} : g \rightarrow dom_F$) ($f_{1n} = O_F (f_1...f_n)$) such that:
– *using Local operator:*

$\quad f_{1n} (x, y) = O_A(f_1(x,y), ..., f_n(x,y))$ *for each point (x,y) of g*

– *using Cubic Focal operator:*

$\quad f_{1n} (x, y) = O_A(f_1 (Neighborhood(x,y))...f_n(Neighborhood(x,y)))$ *for each point (x,y) of g where:*

$\quad$ *Neighborhood(x,y) is a function that returns the neighbourhood points of (x,y),*

– *using Cubic Zonal operator:*

$\quad f_{1n} (x, y) = O_A(f_1(Zone(FieldObjects,(x,y)))...f_n(Zone(FieldObjects,(x,y)))$ *for each point (x,y) of g where:*

Zone(FieldObjects, (x;y)) is a function that takes as input a set of Field Objects and a point, and it returns the neighbourhood points of (x;y) that belong to the zone indentified by the FieldObjects on this point.

Definition 9. built non-field (Figure 7)

Let $tb_{f1},\ldots tb_{fk}$ and t_{nf} instances of the field object structure $S_e = \langle geom, field, [a_1, \ldots a_m] \rangle$

Let ONF, called non-field aggregation mode, a set of aggregation functions:

- *O_G the spatial aggregation for geom*
- *$O_{1\ldots} O_m$ the alphanumeric aggregations for $a_1, \ldots a_m$*
- *O_F the Map Algebra/Map Cubic Algebra aggregation for field*

We say that t_{nf} is built non-field from $tb_{f1},\ldots tb_{fk}$ using ONF if:

- *$t_{nf}.geom = O_G (tb_{f1}.\, geom,\ldots, tb_{fk}.geom)$*
- *$\forall i \in [1,\ldots m]\ t_{nf}.a_i. = O_i\ (tb_{f1}.\, a_i,\ldots, tb_{fk}.\, a_i)$*
- *$t_{nf}.field = O_F(tb_{f1}.field,\ldots, tb_{fk}.\, field)$*

Example 6.

Let an instance of S_{earthq} $t_1 = \langle p_1, f_1 \rangle$ where p_1 is a geometry and f_1 represents the intensity of the earthquake on the 9-1998 in Lombardia. It is defined on each point of p_1 with values in *R*. For example $f_1 (x;y) = 10$ (Figure 6a). To aggregate the field attribute intensity on the temporal dimension, we use a cubic focal operator AVG. Therefore the result of the aggregation of f_1 and f_2 on (x;y) by taking into account neighbours of (x;y) is $f_3 (x;y)= ((13*4+10)+(11*4+12))/10=11.7$ (we suppose that the values of neighbourhood points of (x;y) of t_1 and t_2 are 10 and 12 respectively) (Figure 6a). We suppose that we apply the geometric union for geometry.

Then, t_3 is built non-field from t_1 and t_2. t_3 is an aggregated measure of the cuboid defined by the century level of the temporal dimension (see Table 3).

Table 3: Instances of the cuboid defined by the *century* level of the temporal dimension.

Region	Century	Earthq
Lombardia	900	t_3
Piemonte	900	t_6

Then, as a field measures is mapped also on spatial dimensions, then a particular aggregation must be provided taking into account the field hierarchy in order to allow the visualization of field measures at different resolutions or scales.

3.2.2.1.2 Aggregation on the field hierarchy

The aggregated measures of a cuboid defined by coarser spatial levels are the aggregation of the (aggregated) measures of the cuboids defined by non-spatial levels.

Definition 10. built field (Figure 7)

Let $t^1_{nf},\ldots, t^v_{nf}$ and t_f be instances of the field object structure $S_e \langle geom, field, [a_1, \ldots a_m] \rangle$.

Let OF, called field aggregation mode, a set of aggregation functions:

- *O_G the spatial aggregation for geom*

- $O_1 \ldots O_m$ *the alphanumeric aggregations for* $a_1, \ldots a_m$
- O_A *the alphanumeric aggregation for field*

We say that t_f *is built field from* $t^1{}_{nf}, \ldots, t^v{}_{nf}$ *using OF if:*
- $t_f.geom = O_G \, (t^1{}_{nf}.\,geom, \ldots, t^v{}_{nf}.geom)$
- $\forall\, i \in [1, \ldots m] \; t_f.a_i = O_G \, (t^1{}_{nf}.\,a_i, \ldots, t^v{}_{nf}.\,a_i)$
- $t_f.field \, (x;\, y) = O_A \, (f^1(x^1;y^1), \ldots, f^m(x^m;y^m))$ *where* $f^1, \ldots, f^m$ *belong to* $t^1{}_{nf}.field, \ldots t^v{}_{nf}.field$, *for each point* $(x;y)$ *of the field support of field*
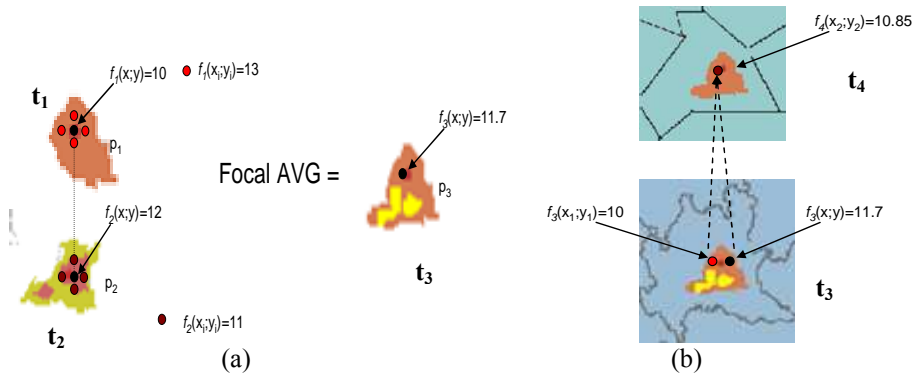

Figure 6: a) Aggregation on the "intensity" field attribute on the temporal dimension, b) Aggregation on the "intensity" field attribute on the field hierarchy.

Example 7.
In order to visualize the measure at different resolutions, we aggregate on the Field Hierarchy H_{regres} applying the average. Then $f_4(x;y) = $ AVG(leavesFieldSupport($H_{deptres}$, $(x_2;y_2)$))) = AVG($f_3(x;y)$, $f_3(x_1;y_1)$) = (10+11.7)/2 = 10.85 (Figure 6b). *Then, t_4 is built field from t_3. t_4 is an aggregated measure of the* cuboid defined by the *century* and *regres* levels (see Table 4).

Table 4: Instance of FC_{earthq}.

Regres (region at scale 1:1.000)	Century	Earthq
Lombardia	900	t_4
Piemonte	900	t_5

3.2.2.2 Cuboids of field data

Once described how the measures of the different cuboids are related by aggregation functions, in this section we formalize the concept of cuboid. In particular, a cuboid schema, noted *Field View Structure*, is composed by a set of levels, a non-field aggregation mode, and a field aggregation mode. An *instance of a field view structure* is a set of tuples composed of a member for each level and a (aggregated) field measure value. The aggregated field measure value on the spatial

dimension (t_f) is obtained aggregating measures (t^1_{nf} ..., t^v_{nf}) obtained after the aggregation on the non-spatial dimensions of detailed measures (tb_{f1} ..., tb_{fk}) as shown on Figure 7.

Definition 11. Field View

A Field View Structure $\mathcal{V}_v$ is a tuple $\langle FC_c, \mathcal{L}, O\mathcal{NF}, O\mathcal{F}\rangle$ where:

- *$FC_c=\langle\mathcal{H}_1,...\mathcal{H}_n, FieldObject\rangle$ is a Field Cube Structure* (Spatio-multidimensional model schema)
- *$\mathcal{L}$ is a tuple $\langle S_1,...S_n\rangle$ where $\forall\, i \in [1,...n]\, S_i$ is a level of $\mathcal{H}_i$* (Levels that define the cuboid)
- *$O\mathcal{NF}$ is a non field aggregation mode* (Aggregation functions used on non-spatial dimensions)
- *$O\mathcal{F}$ is a field aggregation mode* (Aggregation functions used on the spatial dimension)

An Instance of a field view Structure is a set of tuples $\{\langle t_1,...t_n, t_f\rangle\}$ where:

- *$\forall\, i \in [1,...n]\, t_i$ is an instance of S_i* (Dimensions members)
- *t_f is:* ((aggregated) field measure on spatial dimension, Figure 7 - see Table 4 for an example)
    - *an instance of FieldObject*
    - *built field from $t^1_{nf},..., t^v_{nf}$ using $O\mathcal{F}$ where* ((aggregated) field measures on non-spatial dimensions, Figure 7 - see Table 3 for an example)*:*
        - *$t_f.field\,(x;y)= O\mathcal{F}.O_{\mathcal{A}}\,(f^1(x^1;y^1),\ ...,\ f^m(x^m;y^m))$ for each point $(x;y)$ of its field support where:*
            - *$(x^1;y^1),\ ...,\ (x^m;y^m)$ belong to leafsFieldSupport$(\mathcal{H}_1, (x;y))$*
            - *$f^1,\ ...,f^m$ belong to $t^1_{nf}.field,..., t^v_{nf}.field$*
        - *Each t^i_{nf} is built non field from $tb^i_{f1},\ ...,\ tb^i_{fk}$ using $O\mathcal{NF}$ where* (Non aggregated field measures, Figure 7 - see Table 2 for an example)*:*
            - *$tb^i_{f1},\ ...,\ tb^i_{fk}$ are the measure values of the tuples of $I(FC_c)$ $\langle tb^i_1,$ $tb^1_2... tb^1_n, tb^i_{f1}\rangle,\ ...,\langle tb^i_1, tb^k_2... tb^k_n, tb^i_{fk}\rangle$ where:*
                - *$\forall\, i \in [2,...n]\, tb^1_i\, ...\, tb^k_i = leafs(\mathcal{H}_i, t_i)$*
                - *tb^i_1 belongs to leafs$(\mathcal{H}_1, t_1)$*

Example 8.

The Field View Structure representing earthquakes per region at the scale 1:10000 and per century is $V_{earthq} = \langle FC_{earth}, \langle S_{century}, S_{regres}\rangle, \langle Union, Focal\text{-}Avg\rangle, \langle Union, Avg\rangle\rangle$. Table 4 shows its instance.

4 Conclusion and future work

Integration of spatial data into multidimensional models leads to the concept of SOLAP. SOLAP models exploit the discrete representation of spatial data. Few works integrate continuous field data into dimensions and measures. In this paper, motivated by the relevance of a formal representation of SOLAP data, we provide a multidimensional model that considers field data independently form their

implementation, as measures and dimensions. In particular we provide a unique data model for vector and field data (Geographic and Field Objects). We provide a formal representation of the spatio-multidimensional model schema (Field Cube: Field Hierarchy and Field Measures) and the associated hypercube's cuboids (Field View). Actually, we are working on the formal definition of SOLAP operators that allows the navigation between the cuboids (roll-up/drill-down), and slicing the cuboids (slice). We plan to work on the implementation of the model in a ROLAP architecture. This implies the definition of: (i) query languages for OLAP server [18] for field data [12], (ii) indexes [20] and pre-aggregation techniques [19] for spatial data warehouses using field dimensions and measures, and (iii) interactive field maps [17] for SOLAP clients.
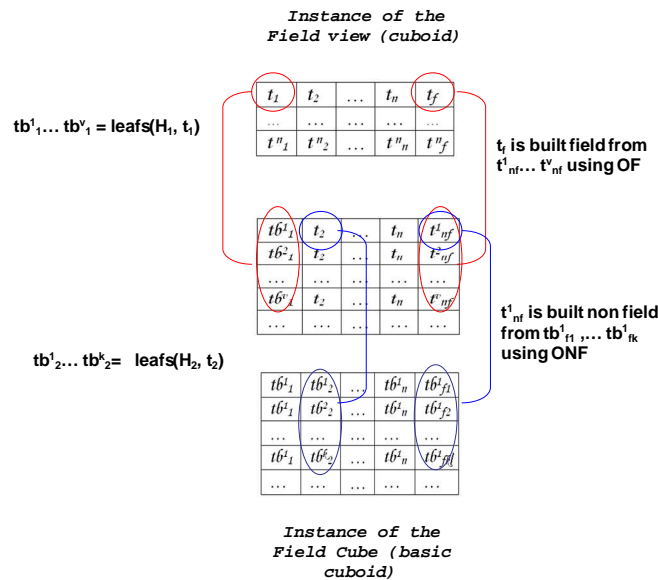

Figure 7: Instance of a Field View Structure

References

[1] Ahmed, T., Miquel, M.: Multidimensional Structures Dedicated to Continuous Spatiotemporal Phenomena. In: 22th British National Conference on Databases. LNCS, vol. 3567 pp. 29-40. Springer, Heidelberg (2005)

[2] Bédard, Y., Han, J.: Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery. Geographic Data Mining and Knowledge Discovery. Taylor & Francis, New York, USA (2009)

[3] Bimonte, S., Gensel, J., Bertolotto, M.: Enriching Spatial OLAP with Map Generalization: a Conceptual Multidimensional Model. In: IEEE International Workshop on Spatial and Spatiotemporal Data Minin, pp. 332 – 34. IEEE CS Press (2008)

[4] Bimonte, S., Tchounikine, A., Miquel, M., Pinet, F.: When Spatial Analysis Meets OLAP: Multidimensional Model and Operators. International Journal of DataWarehousing and Mining (to appear)

[5] Câmara, G., De Freitas, U., Cordeiro, J.: Towards an algebra of geographical fields. In: Brazilian Symp. on computer graphics and image processing, Anais, Curitiba, pp. 205–212 (1994)

[6] Di Martino, S., Bimonte, S., Bertolotto, M., Ferrucci; F.: Integrating Google Earth within OLAP Tools for Multidimensional Exploration and Analysis of Spatial Data. In: ICEIS 2009. LNBIP, vol. 24, pp. 940-951. Springer, Heidelberg (2009)

[7] Franklin, C.: An Introduction to Geographic Information Systems: Linking Maps to databases. Database, 15(2), 13-21 (1992)

[8] Gutierrez, A., Baumann, P.: Modeling Fundamental Geo-Raster Operations with Array Algebra. In: IEEE International Workshop on Spatial and Spatiotemporal Data Mining, pp. 607-612. IEEE CS Press (2007)

[9] Harinarayan, V., Rajaraman, A., Ullman, J. D.: Implementing Data Cubes Efficiently. In: ACM SIGMOD International Conference on Management of Data, pp. 205-216. ACM Press, New York (1996)

[10] Kemp, K.: 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 (1993)

[11] Kimball, R.: The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. John Wiley & Sons, New York, USA (1996).

[12] Laurini, R., Gordillo, S.: Field Orientation for Continuous Spatio-temporal Phenomena. In: International Workshop on Emerging Technologies for Geo-based Applicatons. pp. 77-101 Swiss Federal Institute of Technology, Lausanne (2000)

[13] Ledoux, H., Gold., C.M.: A Voronoi-based Map Algebra. In: 12th International Symp. on Spatial Data Handling, pp. 117-131. Springer, Heidelberg (2006)

[14] Malinowski, E., Zimányi, E.: Advanced Data Warehouse Design From Conventional to Spatial and Temporal Applications. Springer, Heidelberg (2008)

[15] McHugh, R.: Intégration De La Structure Matricielle Dans Les Cubes Spatiaux. Université Laval (2008)

[16] Mennis, J., Viger, R., Tomlin, C.D.: Cubic Map Algebra functions for spatio-temporal analysis. Cartography and Geographic Information Systems, 30(1), 17-30 (2005)

[17] Plumejeaud, C., Vincent, J., Grasland, C., Bimonte, S., Mathian, H., Guelton, S., Boulier, J., Gensel, J.: HyperSmooth, a system for Interactive Spatial Analysis via Potential Maps. In: International Symposium on Web and Wireless Geographical Information Systems. LNCS, vol. 5373, pp. 4-16. Springer, Heidelberg (2008)

[18] Silva, J., Castro Vera, A.S., Oliveira, A.G., Fidalgo, R., Salgado, A.C., Times, V.C.: Querying geographical data warehouses with GeoMDQL. In: Brazilian Symposium on Databases, pp. 223–237 (2007)

[19] Stefanovic, N., Han, J., Koperski, K.: Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes. IEEE Transactions on Knowledge and Data Engineering, 12 (6), 938-958 (2000)

[20] Tao, Y., Papadias, D.: Historical spatio-temporal aggregation. ACM Trans. Inf. Syst., 23(1), 61-102 (2005)

[21] Timpf, S., Frank, A. U.: Using hierarchical spatial data structures for hierarchical spatial reasoning. In: Spatial Information Theory - A Theoretical Basis for GIS.LNCS, vol. 1329, pp 69-83. Springer, Heidelberg (1992)

[22] Tomlin, C.D.: Geographic Information Systems and Cartographic Modeling. Prentice Hall, Englewood Cliffs, NJ (1990)

[23] Vaisman, A., Zimányi, E.: A multidimensional model representing continuous fields in spatial data warehouses. In: 17th ACM SIGSPATIAL International Symp. on Advances in Geographic Information Systems, pp. 168-177. ACM Press, New York, USA (2009)