

Intégration des données champs continus incomplets dans l'OLAP : de la modélisation conceptuelle à l'implémentation

Mehdi Zaamoune*, Sandro Bimonte*, François Pinet*, Philippe Beaune**

*IRSTEA, TSCF, 24 Av. des Landais, Aubière, France
mehdi.zaamoune,sandro.bimonte,francois.pinet@irstea.fr

**Agaetis, 10 allée Evariste Galois, 63 000 Clermont-Ferrand
pbeaune@agaetis.fr

Résumé. Les champs continus sont des types de représentations utilisés pour modéliser des phénomènes tels que la température, la pollution ou l'élévation. Différents travaux se sont intéressés à l'intégration de ce type de représentations dans un système SOLAP. Ainsi, dans ce papier, nous proposons une modélisation ainsi qu'une implémentation des dimensions spatiales représentant des champs continus incomplets sous forme de grilles régulières à des résolutions différentes, à fin de les intégrer dans un système SOLAP.

1 Introduction et motivations

Les Entrepôts de Données Spatiaux (EDS) et les systèmes OLAP spatiaux (SOLAP) jouent un rôle important en aidant les décideurs à obtenir le maximum d'avantages de ces grandes quantités de données géographiques (Yvan Bédard et Han (2001)). Ces technologies étendent les entrepôts de données (ED) et les systèmes OLAP pour intégrer les données spatiales avec les données classiques entreposées pour réaliser l'analyse en ligne de grands ensembles de données géoréférencées. Les systèmes SOLAP intègrent le système OLAP et les systèmes d'information géographique (SIG) dans un cadre unique généralement basé sur le stockage relationnel des données spatiales selon le modèle vectoriel, leur analyse via des opérateurs SOLAP (Spatial Roll-Up, Slice spatiale, etc) et leur visualisation à l'aide d'affichages tabulaires, graphiques et cartographiques. Les EDS sont modélisés en fonction du modèle spatio-multidimensionnel qui étend le modèle traditionnel multidimensionnel pour définir les dimensions spatiales (c-à-d axes d'analyse avec les attributs spatiaux) et des mesures spatiales (c-à-d sujets d'analyse) qui intègrent les informations géographiques en utilisant le modèle vectoriel.

L'information géographique peut être représentée selon deux modèles, en fonction de la nature des données : champ discret (vecteur) et champ continue (Tomlin (1990)). Les champs continus (également appelés "données spatiales continues" ou "Continuous fields") représentent des phénomènes physiques qui changent continuellement dans l'espace (Luca Paolino et Vitiello (2010)), par exemple la température, la densité de la couverture végétale, etc (Gintautas (2009)). Puisque les représentations continues ne peuvent pas être implémentées dans les systèmes existants, deux catégories de représentations des données spatiales continues ont été proposées : la représentation incomplète et la représentation complète (Luca Paolino et Vi-

Champs continus incomplets et OLAP

tiello (2010)). Les représentations incomplètes stockent un échantillon de points et ont besoin de fonctions supplémentaires pour calculer les valeurs dans les zones non échantillonnées (grille de points, TIN, etc) (figure 1). Les représentations complètes associent des valeurs à des régions et supposent que ces valeurs sont valables pour chaque point dans ces régions (e.g raster). Pour les données continues, des opérateurs d'analyse ad-hoc (map algèbre), différents des opérateurs topologiques pris en charge par le modèle vectoriel ont été définis, et qui permettent une analyse point par point (Tomlin (1990)) pour les champs complets et (Ledoux (2008)) pour les champs incomplets).

En revanche, la représentation des données géographiques à différentes échelles ou résolutions (figure 1-d) est souhaitable pour une analyse efficace des phénomènes spatiaux complexes. Puisque le raisonnement humain est généralement de haut en bas, les phénomènes spatiaux présentent des résultats différents lorsqu'on les analyse à différentes échelles ou résolutions, ce qui représente une méthode efficace d'extraction des connaissances (méthodes de géovisualisation). Par conséquent, ces résolutions ou échelles représentent les besoins d'analyse des décideurs qui doivent être explicitement représentés dans les données et dans le modèle de requêtes.

Nous proposons donc dans cet article un profil UML pour la représentation multi-dimensionnelle des champs continus. Nous implémentons ce profil dans une architecture relationnelle SOLAP selon deux stratégies : la première utilise une extension du modèle logique classique en étoile (SEAF) et le langage multidimensionnel MDX (defacto standard des systèmes OLAP) et la 2^{ème} stratégie, en utilisant la continuité intrinsèque des données géographiques, utilise un nouveau modèle logique, appelé SEIF, et un ensemble de nouvelles fonctions MDX. Nous validons notre proposition à l'aide d'un cas d'études réel portant sur le contrôle des odeurs, et nous fournissons quelques expérimentations montrant sa faisabilité. Le cas d'étude que nous proposons est basé sur des données issues de la surveillance des odeurs en milieu urbain. Pour chaque 15 minutes et chaque type d'odeur (par exemple NO₂) une grille régulière (champs continu incomplet) est produite au moyen de quelques points d'échantillonnage et d'un modèle de simulation (ADMS5). Le modèle de simulation permet d'estimer les odeurs pour toute une zone urbaine et produit des grilles thématiques à une résolution de 100*100. Des exemples des points de la grille sont représentés dans la figure 1-a (les valeurs des odeurs sont représentées par les couleurs : vert, jaune, rouge,...) pour 10h00 et 10h15 du 19-2-2012. Supposons maintenant que l'utilisateur veut agréger des données le long d'une dimension temporelle (minute, heure, jour, mois, année) en utilisant la moyenne pour obtenir une carte d'odeur agrégée. Il s'agit d'une opération RollUp de OLAP sur la dimension temporelle qui correspond à une opération map algebra locale (Figure 1-a).

En outre, puisque l'espace est représenté d'une manière continue, les décideurs devraient être en mesure de demander le résultat d'une requête OLAP en tout point de la dimension spatiale (par exemple, ils peuvent être intéressés par la valeur de l'odeur à 10 : 00 dans une zone située derrière un bâtiment) (Figure 1-b). Il est également possible d'appliquer un opérateur "slice" spatial sur la dimension spatiale (c-à-d en utilisant un prédicat spatial pour sélectionner un sous-ensemble de l'espace) (Figure 1-c).

Afin de répondre à ces deux dernières requêtes, des méthodes d'interpolation spatiale sont nécessaires, puisque dans le champ continu incomplet seules les valeurs fournies par le modèle de simulation sont stockées. L'interpolation spatiale est le processus d'estimation presque exact des valeurs des attributs à des endroits non échantillonnés à partir des mesures effectuées à

des points de contrôle dans la même zone (O'Sullivan et J.Unwin (2002)). Dans notre cas, la fonction d'interpolation utilisée est l'interpolation bilinéaire, qui est une méthode déterministe locale. Elle utilise un échantillon de points de la grille (les 4 voisins les plus proches du point inconnu) et calcule une moyenne pondérée, en utilisant la distance qui détermine dans quelle proportion la valeur d'un voisin influence sur la valeur du point à estimer.

Enfin, comme indiqué dans la section précédente, puisque la visualisation de données spatiales à différentes résolutions est fondamentale dans le processus d'exploration/analyse, les décideurs devraient être en mesure d'interroger les données spatiales entreposées à différentes résolutions. Dans notre cas d'étude, par exemple, les concepteurs devraient être en mesure d'analyser les données à des résolutions plus détaillées (ex : 400 * 400), afin d'avoir une vision plus fine du phénomène (Figure 1-d).

Il est très important de noter que, généralement, pour chaque phénomène spatial un ensemble de résolutions utiles et connues existent, afin qu'elles puissent être prédéfinies en fonction des données et des besoins des utilisateurs.

Pour résumer, l'analyse spatio-multidimensionnelle des données de type champs continus doit supporter : (i) les opérateurs OLAP classiques comme *map algebra*, (ii) la *vue continue des données spatiales*, (iii) les opérateurs spatiaux (*slice*), et (iv) l'interrogation des données à différentes résolutions prédéfinies.

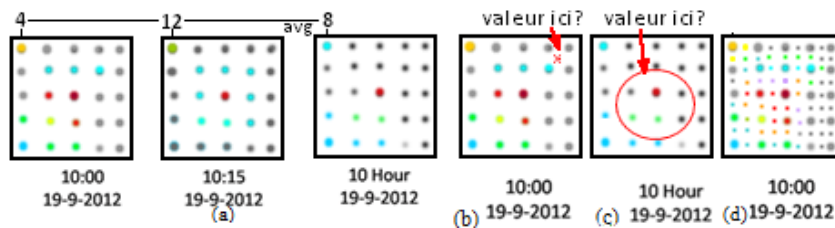


FIG. 1 – (a) *map algebra* (b) *la vue continue des données spatiales* (c) *les opérateurs spatiaux (slice)* (d) *l'interrogation des données à différentes résolutions prédéfinies*

2 Profil UML pour les entrepôt de données intégrant les champs continus incomplets

Les profils UML constituent un moyen pour étendre et adapter UML à des domaines ou des plateformes particulières. Les avantages de l'utilisation de UML pour la modélisation des entrepôts de données complexes, ont déjà été démontrées dans plusieurs travaux (Bédard Yvan et Bernier (2002)).

Dans cette section, nous décrivons notre modèle spatio-multidimensionnel intégrant les champs continus incomplets à différentes résolutions. Notre modèle étend le profil proposé par (Boulil (2012)) à fin d'y intégrer les données de type champs continus à différentes résolutions. (Boulil (2012)) proposent un profil UML pour la modélisation des entrepôts de données spatiales qui permet de représenter conceptuellement tous les aspects avancés de la modélisation

Champs continus incomplets et OLAP

spatio-multidimensionnelle. Le profil présente des stéréotypes pour chaque élément spatio-multidimensionnel, par exemple : "Fact" pour le fait, "SpatialAggLevel" pour les niveaux des dimensions spatiales, "AggRelationship" pour les relation hiérarchiques, etc.

Un fait "Fact" est composé de mesures "Measure" et d'une relation "DimRelationship" qui le relie aux niveaux des dimensions "AggLevel". Un "AggLevel" est composé d'attributs "Dimensional attribute" et peut être thématique, spatial ou temporel.

Dans notre extension nous avons défini un niveau "FieldIncomplete" (champs continu incomplet) comme un type spécial du niveau spatial où chaque membre présente une fonction d'interpolation pour représenter la continuité, le niveau de résolution auquel il appartient et une relation de voisinage ("NeighborhoodRelationship") sur laquelle se base la fonction d'interpolation pour son calcul. En effet, les fonctions d'interpolation spatiale utilisent un ensemble de points, qui dépendent de la fonction (par exemple une fonction bilinéaire utilise une grille $2 * 2$), pour estimer une valeur à un point donné. Dans notre approche, la navigation d'un niveau field à un autre niveau field (multi-résolution) n'implique pas forcément l'utilisation d'une opération d'agrégation, mais plutôt une interpolation ("FieldInterpolation"). Pour cela nous avons redéfini les règles de navigation dans les niveaux d'une hiérarchie, proposée par (Bouilil (2012)), en définissant le stéréotype "LevelToFieldRule". En effet, selon le modèle de (Bouilil (2012)), l'agrégation d'une mesure est définie grâce au stéréotype "BasicIndicator" qui présente une valeur marquée ("AggregatedAttribute") représentant la mesure, une valeur marquée représentant la fonction d'agrégation ("Agregator") et la dimension sur laquelle s'applique l'agrégation avec le stéréotype "DimensionAggRule". Nous avons étendu ce modèle tel que définit précédemment pour y ajouter une nouvelle opération "FieldInterpolation" qui fait office de règle lors d'une navigation entre des niveaux de type "FieldIncomplete" au lieu de l'agrégation.

Une instance du profil pour le cas d'études est illustrée dans la figure 2. Le cube est composé de quatre dimensions : la dimension "Tracer" qui représente le type du polluant (e.g NO₂), la dimension "Source" qui représente la source de l'odeur (e.g usine), la dimension temporelle "TimeD" constituée 4 niveaux de granularité (année, mois, jour et minute) et la dimension spatiale "Location" qui représente un champs continu incomplet, et qui est composée de 3 niveau de résolution (Grid100, Grid200 et Grid400) de type "fieldIncomplete" qui sont composés chacun d'une géométrie "Point", de leur niveau de résolution, et de leur fonction d'interpolation "ContinuityInterpolation".

Dans l'indicateur "OdorIndicator" est formulée la règle d'agrégation utilisée pour la navigation dans les résolutions (Bilinear) ou la navigation dans les hiérarchies non spatiales (avg).

3 Nouveaux modèles relationnels OLAP pour les champs continus incomplets

Dans cette section, nous présentons l'implémentation de notre modèle multidimensionnel dans une architecture typique SOLAP relationnelle basée sur SQL (Relational DBMS standard language) et MDX, qui est la norme standard des serveurs OLAP. Cela donne un caractère général pour notre approche, étant donné qu'il est possible de l'implémenter dans toute architecture de ce type.

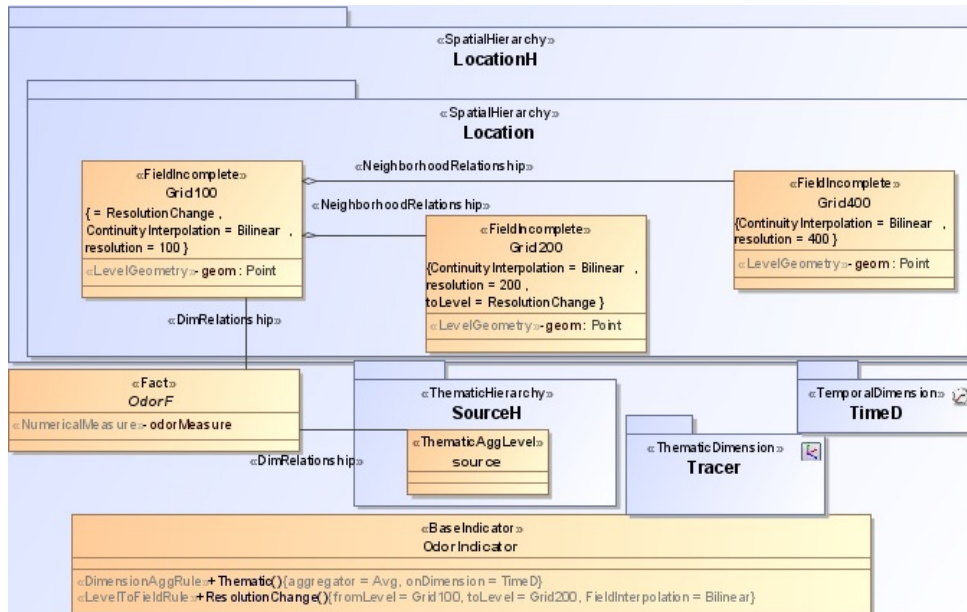


FIG. 2 – Cas d'étude : instance du profil

Supposons que nous avons un "niveau field" représentant une grille de points à la résolution $100 * 100$, le modèle logique dans ce cas là est un schéma en étoile classique. Ce modèle est composé d'une table de faits contenant des mesures et les clés étrangères des tables de dimensions. Chaque table de dimension est dé-normalisée, et possède des attributs qui représentent ses niveaux. La "dimension Field" représente une grille régulière de points et se compose d'un niveau représentant la grille régulière à la résolution $100*100$ (10000 points) qu'on appellera "Res100", et qui est composé à son tour d'un identifiant et d'une géométrie qui représente un point.

La représentation des champs continus incomplets tel que présentée dans ce modèle logique permet d'effectuer des requêtes MDX via des opérateurs map algebra (agrégation point par point) comme ce qui suit :

Requête 1 : la valeur moyenne de l'odeur pour chaque membre field pour l'année 2012.

```
SELECT [Field].[Res100].Members ON ROWS, [time].[2012] ON
COLUMNS FROM [odorCube] WHERE [Measures].[value]
```

3.1 Champs continus incomplets

Afin d'avoir une représentation continue du niveau field, nous avons défini une fonction GeoMDX (extension de MDX avec des opérateurs spatiaux) définie par l'utilisateur qui représente une interpolation spatiale sous la forme :

```
NumericType InterpolatePoint (géométrie)
```

Champs continus incomplets et OLAP

La fonction "InterpolatePoint" prend en entrée une géométrie (point) et retourne une valeur numérique, qui est une mesure dérivée dans le modèle OLAP, et qui représente une valeur estimée calculée en utilisant les valeurs du voisinage du point donné en entrée. Notons que le voisinage est défini différemment selon l'implémentation de la fonction abstraite "InterpolatePoint" (bilinéaire, bicubique, krieging,...). Ainsi, supposons que nous voulons récupérer une valeur du field à une localisation dont la géométrie est définie par le point de coordonnées géométriques (42.3521 -72.1235), pour répondre à ce besoin en utilisant la fonction d'interpolation bilinéaire, les décideurs doivent simplement utiliser la fonction GeoMDX de la manière suivante : `InterpolatePoint (POINT (-72.1235 42.3521))`. Ce qui permet à l'utilisateur de retrouver une valeur à chaque point du domaine spatiale d'un façon transparente, comme s'il existait déjà dans le EDS.

Ainsi, la fonction va récupérer les voisins du point donné comme paramètre dans le niveau field ([Field] . [Res100]), sur la base de la distance (st-distance de GeoMDX), et trouver leurs valeurs respectives dans la table de faits, puis évaluer la valeur du point à estimer en utilisant ces valeurs, et enfin renvoyer la mesure estimée obtenue. Voici un exemple de requête qui utilise une implémentation de la fonction "InterpolatePoint" :

Requête 2 : la valeur de l'odeur au point avec les coordonnées (721148 3140020) pour l'année 2012.

```
With member [Measures].[value] as 'InterpolatePointBilinear(  
ST_GeomFromText("POINT (721148 3140020)") )'  
SELECT [Measures].[value] ON ROWS, [time].[2012] ON COLUMNS  
FROM [odorCube]
```

De cette façon, nous obtenons une vision d'analyse et d'affichage continue des données de type champs continus incomplets comme indiqué dans la section 2.

3.2 Multi-résolution

La multi-résolution est une approche qui consiste à définir des niveaux de résolution susceptibles d'améliorer le rendu des requêtes faites par l'utilisateur. Pour modéliser un champ continu incomplet à plusieurs résolutions dans un modèle multidimensionnel, nous proposons deux approches fondées sur le "schéma en étoile classique" : L'approche "schéma en étoile d'agrégation field" (SEAF) et l'approche "schéma en étoile d'interpolation field" (SEIF).

3.2.1 Schéma en étoile d'agrégation field (SEAF)

Basé sur le schéma en étoile décrit précédemment, nous proposons un schéma logique où les mesures sont des valeurs du champ continu et où la dimension spatiale représente des niveaux field à différentes résolutions (fig. 3-a). Ce modèle étend la dimension field avec 2 niveaux supplémentaires représentant chacun un niveau différent de résolution ([Field] . [res200] et [Field] . [res400]). Chaque niveau de la dimension field est composé d'un identifiant et d'une géométrie qui représente un point. La table de faits est associée, classiquement, au niveau de la dimension field le plus détaillé. De cette façon, le décideur peut analyser les données spatiales à différentes résolutions lors de la même session d'analyse MDX.

Il suffit de changer le niveau de résolution dans la requête pour modifier le niveau de détails du résultat. En utilisant cette approche, il suffit de préciser dans une requête MDX, le niveau approprié de résolution de la dimension "field" comme dans ce qui suit, où la requête 1 devient :

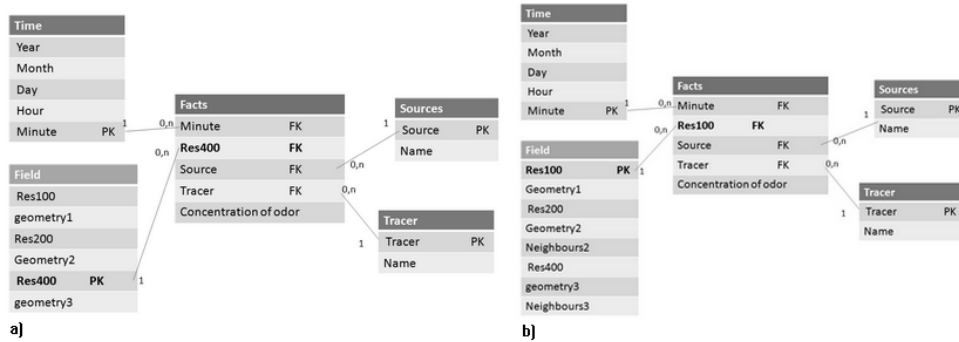


FIG. 3 – a) schéma en étoile d'agrégation field, b) schéma en étoile d'interpolation field

Requête 3 : la valeur moyenne de l'odeur pour les membre de la dimension field à la résolution 400*400 pour l'année 2012.

```
SELECT [Field].[res400]. Members ON ROWS, [time].[2012] ON
COLUMNS FROM [odorCube] WHERE [Measures].[value]
```

3.2.2 Schéma en étoile d'interpolation field (SEIF)

Comme indiqué dans la section 2, afin d'obtenir des données de type champ continu incomplet à des résolutions plus fines, les méthodes d'interpolation spatiale peuvent être utilisées. Nous proposons ici une variante du schéma proposé précédemment pour gérer plusieurs niveaux de résolution sur le champs continu incomplet, en associant la table de faits au niveau field à la résolution la moins détaillée, comme indiqué sur la figure 3-b. Notons que cette approche n'est possible que lorsqu'il s'agit de données spatiales, parce que, selon la loi de Tobler, la position géographique des données peut être utilisée pour estimer les valeurs manquantes à des locations non échantillonnées .

Pour implémenter le modèle que nous proposons (figure 3-b), nous définissons une fonction GeoMDX de la même manière que la fonction définie dans la section 4.1 :
Numeric-type InterpolateResolution (Field Member)

Toutefois, cette fonction "InterpolateResolution" est préparée à recevoir en entrée un élément d'un niveau du field au lieu d'une géométrie, et à retourner une valeur interpolée de ce membre. Nous pouvons également voir que dans ce cas, les voisins de chaque membre d'une résolution plus élevée que celle d'origine sont également stockés comme attributs (Neighbors2 et Neighbors3), puisque les membres de chaque résolution sont pré-définis à l'avance.

L'appel de l'implémentation (bilinéaire) de cette fonction comme ce qui suit :
InterpolateResolutionBilinear ([Field].[res400].CurrentMember)
dans la formule d'une mesure dérivée, permet de trouver les valeurs de tous les membres du niveau "res400" (champ incomplet à une résolution 400 * 400) à l'aide de leurs voisins dans le niveau de résolution 100*100 (Neighbors3). Ainsi, la **requête 3** peut être réalisée comme suit :

```
With member [Measures].[EstimatedValue] as
```

Champs continus incomplets et OLAP

```
' InterpolateResolutionBilinear ([Field].[res400].CurrentMember) '  
SELECT [Field].[res400].Members ON ROWS, [time].[2012] ON  
COLUMNS FROM [odorCube] Where [Measures].[EstimatedValue]
```

Comme on peut le voir dans la requête précédente, l'appel de la mesure calculée avec la fonction `InterpolateBilinear()` permet de retrouver les valeurs à une échelle donnée de manière transparente pour le décideur comme pour une agrégation classique (SQL).

4 Expérimentations

Dans cette section, nous détaillons les performances des deux approches proposées dans la section 4.2 en termes de stockage et de temps de calcul.

L'ordinateur utilisé pour les essais suivants a la configuration suivante : processeur Intel Core i3 2,20 GHz, 4 Go RAM, Système d'exploitation Windows 7 professionnel, système OS 64 bits. Les données spatiales sont stockées dans le SGBD spatial PostGIS. PostGIS est un logiciel open source qui ajoute le support d'objets géographique à la base de données PostgreSQL. GeoMondrian fournit une intégration cohérente des données spatiales vectorielles dans la structure de données du cube OLAP pour définir des niveaux spatiaux.

Afin de tester notre proposition, nous avons défini différents cas où les dimensions spatiales contiennent : un niveau field de résolution $100 * 100$, deux niveaux de résolution $100 * 100$ et $200 * 200$, et enfin trois niveaux de résolution $100 * 100$, $200 * 200$ et $400 * 400$. Nous avons également fait varier la taille de la dimension temporelle afin de comprendre l'impact de la dimension spatiale et non spatiale sur les performances.

En ce qui concerne les performances de stockage, en utilisant l'approche SEIF la taille de la base de donnée a été réduite de 1,176 Go à 87,628 Mo. Nous pouvons facilement voir deux différences importantes : i) l'approche SEAF est plus coûteuse en termes de stockage que l'approche SEIF puisque cette dernière stocke seulement les valeurs des faits à la granularité spatiale la moins détaillée, ii) dans l'approche SEIF la taille de la table de faits varie uniquement en fonction de la taille des dimensions non spatiales. Ainsi, même en augmentant la taille de la dimension spatiale en ajoutant de nouvelles résolutions, la table de faits ne change pas, car elle ne contient que les mesures relatives au premier niveau de résolution.

Afin d'évaluer les performances en temps de calcul, nous avons exécuté les requêtes précédemment citées, où nous avons combiné des opération d'agrégation (roll-up) sur les dimensions non spatiales et les opérateurs spatiaux (slice) sur le champ continu incomplet à différentes résolutions. On a noté un certain degré d'approximation en temps d'exécution entre les deux approches avant l'utilisation du niveau de résolution $400*400$. Au-delà de ce niveau, on constate que l'écart se creuse considérablement. Ainsi, minimiser le stockage et ainsi les relations entre la table de fait et les dimensions, a permis à la méthode SEIF, que nous proposons, d'avoir un meilleur temps d'exécution que l'approche SEAF à tous les niveaux de résolution ($100 * 100$, $200 * 200$ et $400 * 400$).

5 État de l'art

Afin d'intégrer les données de type champs continus dans un modèle SOLAP, (Ahmed et Miquel (2005)) proposent un modèle multidimensionnel pour le traitement des champs conti-

nus incomplets, en stockant un échantillon de points comme membres spatiaux, puis créent un cube discret qui est interpolé du côté du client pour simuler la continuité. Une requête qui implique des opérateurs spatiaux (i.e spatial slice), ne peut pas être exécutée sur les points créés à travers le processus d'interpolation. (McHugh (2008)) définit de nouveaux types de dimensions qui représentent des grilles régulières de carrés (raster) : "dimension hybride", "mixte dimension hybride", "dimension de la matrice mixte" et "dimension de la matrice géométrique". Elle définit aussi le "cube matriciel" où les faits sont des cellules (carrés). L'approche SEAF, présentée dans la section 4.2.1 est basé sur ce travail. (Leticia I. Gomez et Vaisman (2012)) présentent un modèle de données discret pour représenter des champs continus et une algèbre pour utiliser les opérateurs OLAP (Dice, Slice, Roll-up, drill-down ...). Cependant, ce modèle discret que les auteurs proposent ne supporte pas l'aspect continu du champ, qui consiste à extraire une valeur pour chaque point de coordonnées x et y dans le domaine spatial.

(Bédard Yvan et Bernier (2002)) définit un modèle conceptuel basé sur UML qui intègre plusieurs propriétés géométriques et sémantiques des représentations de niveaux spatiaux. Toutefois, ce travail ne présente pas un modèle complet et multidimensionnel pour les faits et les hiérarchies.

En outre, (Maryvonne Miquel et al. (2002)) suggèrent (sans fournir de détails) d'utiliser un entrepôt de données spatiale pour chaque représentation de la donnée spatiale. Par conséquent, changer la représentation correspond à passer d'un entrepôt de données spatiales à un autre, ce qui est compliqué à faire dans une même session d'analyse. (Gascueña et Guadalupe (2009)) proposent un modèle conceptuel pour la multi-représentation des membres spatiaux. Ils proposent aussi un schéma physique, mais aucune implémentation dans une architecture classique ROLAP n'est présentée.

6 Conclusion

Dans cet article, nous présentons un modèle multidimensionnel pour les champs continus incomplets à plusieurs résolutions et son implémentation dans une architecture SOLAP basées sur les normes (SQL et MDX). Nous travaillons sur l'utilisation des méthodes de forage de données spatiales pour accélérer les opérations de map algèbre et sur l'implémentation d'un client de visualisation SOLAP.

7 Remerciements

Les auteurs tiennent à remercier le FEDER et la région Auvergne pour le financement de ce projet et les partenaires du projet qui nous ont fourni les données pour les expérimentations.

Références

Ahmed, T. O. et M. Miquel (2005). "multidimensional structures dedicated to continuous spatiotemporal phenomena". *In proceeding of : Database : Enterprise, Skills and Innovation, 22nd British National Conference on Databases, BNCOD 22, Sunderland, UK, July 5-7, 2005, Proceedings.*

Champs continus incomplets et OLAP

- Bédard Yvan, Marie Josée Proulx, S. L. et E. Bernier (2002). "modeling multirepresentations into spatial data warehouses : A uml-based approach". *Symposium on Geospatial Theory, Processing and Applications, Ottawa*.
- Boulil, K. (2012). "*Une Approche Automatisée basée sur des Contraintes d'Intégrité définies en UML et OCL pour la Vérification de la Cohérence Logique dans les Systèmes SOLAP*". Ph. D. thesis, Université Blaise Pascal.
- Gascueña, C. M. et R. Guadalupe (2009). "a multidimensional methodology with support for spatio-temporal multigranularity in the conceptual and logical phases". *IGI Global*.
- Gintautas, M. (2009). "the continuous field view of representing forest geographically : from cartographic representation towards improved management planning". *Published in S.A.P.I.E.N.S., 2.2 | 2009*.
- Ledoux, H. (2008). "fieldgml : An alternative representation for fields". *the Proceedings of SDH 2008 - The 13th International Symposium on Spatial Data Handling*.
- Leticia I. Gomez, S. A. G. et A. A. Vaisman (2012). "a generic data model and query language for spatiotemporal olap cube analysis". *EDBT '12 Proceedings of the 15th International Conference on Extending Database Technology, Pages 300-311 ACM New York, NY, USA l'2012*.
- Luca Paolino, Monica Sebillio, G. T. et G. Vitiello (2010). "integrating discrete and continuous data in an opengeospatial-compliant specification". *Transactions in GIS Volume 14, Issue 6, pages 731-753, December 2010*.
- Maryvonne Miquel, Y. B., A. Brisebois, J. Pouliot, P. Marchand, et J. Brodeur (2002). "modeling multi-dimensional spatio-temporal data warehouses in a context of evolving specifications". *Symposium sur la théorie, les traitements et les applications des données Géospatiales, Ottawa*.
- McHugh, R. (2008). "*Intégration de la structure matricielle dans les cubes spatiaux*". Ph. D. thesis, Département des sciences géomatiques faculté de foresterie et géomatique université laval.
- O'Sullivan, D. et D. J. Unwin (2002). "geographic information analysis - john wiley and sons, 15 nov. 2002 - 448 pages".
- Tomlin, C. D. (1990). "geographic information systems and cartographic modeling". *Prentice-Hall, Englewood-Cliffs, NJ, 1990*.
- Yvan Bédard, T. M. et J. Han (2001). "fundamentals of spatial data warehousing for geographic knowledge discovery". *Data Mining and Knowledge Discovery*.

Summary

Continuous fields are types of representations used to model phenomena such as temperature, pollution or elevation. Various studies have focused on the integration of such representations in a SOLAP system. Thus, in this paper, we propose a model and an implementation of spatial dimensions representing incomplete continuous fields as regular grids at different resolutions in order to integrate them into a SOLAP system.