

# A Decision Support System for Spatial Analysis of Agricultural Production in Madagascar

Aimé Richard Hajalalaina<sup>1,2</sup>, Solofoson Georges Andriniaina<sup>1</sup>

<sup>1</sup>Ecole de Management et d'Innovation Technologique (EMIT), University of Fianarantsoa, Fianarantsoa, Madagascar

<sup>2</sup>Laboratoire d'Informatique et des Mathématiques Appliqués au Développement (LIMAD), University of Fianarantsoa, Fianarantsoa, Madagascar

Email: arhajalalaina@yahoo.fr

**How to cite this paper:** Hajalalaina, A.R. and Andriniaina, S.G. (2021) A Decision Support System for Spatial Analysis of Agricultural Production in Madagascar. *Journal of Data Analysis and Information Processing*, 9, 1-22.

<https://doi.org/10.4236/jdaip.2021.91001>

**Received:** January 3, 2021

**Accepted:** February 7, 2021

**Published:** February 10, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

In this article, our research aims to set up a geo-decisional system, more precisely we are particularly interested in the spatial analysis system of agricultural production in Madagascar. For this, we used the spatial data warehouse technique based on the SOLAP spatial analysis tool. After having defined the concepts underlying these systems, we propose to address the research issues related to them from four points of view: needs study of the Malagasy Ministry of Agriculture, modeling of a multidimensional conceptual model according to the MultiDim model and the implementation of the system studied using GeoKettle, PostGIS, GeoServer, SPAGO BI and Géomondrian technologies. This new system helps improve the decision-making process for agricultural production in Madagascar.

## Keywords

Geo-Decisional System, Agricultural Production, Decision-Making, Spatial Analysis, Data Warehouse, MultiDim Model, Business Intelligence, Madagascar

## 1. Introduction

Agriculture is one of the important issues for Madagascar's economy and it needs breakthrough techniques in this century. With today's computerized world, the processing and analysis of agricultural data are growing needs for farmers and policy makers. The current transactional systems of the Malagasy Ministry of Agriculture do not make it possible to support decision-making support processes concerning the spatial analysis of agricultural production. To solve this problem, the implementation of a new spatial analysis system based on a data warehouse is essential.

The need for storage of large quantities of data and the need for a Decision Support System (DSS) gave birth to business intelligence. By integrating data warehouse technology, the decision-making process provides a response to the problem of continued growth of data that may be of different formats [1]. Organizations have increasingly resorted to setting up data warehouses to collect and integrate all of their data. Therefore, many authors like [2] and [3] consider that this technology has become the most used way to obtain information and to meet the needs of analyzes in the majority of public institutions.

In addition to data warehouse method, decision-makers need another perspective on the visualization of spatially distributed data, hence the integration of the Geographic Information System (GIS). The combination of these two systems brings out a new concept: Spatial On-Line Analytical Processing (SOLAP), which is a DSS [4]. Thus, it is estimated that about 80% of the data stored in corporate databases have a spatial reference [5].

Regarding the use of this system, several studies have already been conducted on the implementation of SOLAP. The implementation of the first SOLAP tool, JMap, is carried out by the geomatics research team of the University of Laval, Quebec, Canada [6] that is the creator of the SOLAP acronym and the first to commercialize SOLAP technology. In his works, [7] proposes the use of Geobrowser Google Earth which implements the flexible SOLAP Web system, GoOlap, combining both OLAP functionalities and advanced geovisualization features. Another solution is also realized by [8] for the forest inventory. And there are also other alternatives proposed by another research team GeoSOA and several commercial companies.

Thus, our work concerns the setting up of a DSS within the Ministry of Agriculture in Madagascar, using the SOLAP system. This Ministry collects every year a large volume of data from these services that are distributed geographically throughout the island. However, it does not have any tools necessary to exploit, analyze and synthesize these data. The question is, can the implementation of this new system improve the decision-making process within the Ministry of Agriculture?

To answer these questions, we will design a data warehouse for agricultural production and then implement a spatial analysis tool for decision support.

## 2. Basic Concepts

### 2.1. Decision Support System

Generally, a DSS encompasses all the means by which an organization can monitor and evaluate its daily activities. The DSS is defined as “a set of Information Technology solutions that allow the analysis of company data, in order to derive new qualitative information, to detect macroscopic information hidden in large volumes of data [9]. This system also means “the means, the tools and the methods that make it possible to collect, consolidate, model and restore the data, material or immaterial, of a company in order to ‘provide decision support and

allow a decision maker to have an overview of the activity being handled”.

The main purpose of a DSS is to provide indicators about the life of an organization. As a result, it exploits a wealth already at its disposal, *i.e.* all the data that an organization generates in the context of its activity. In addition, it plays a more important role strategically because the data, after its exploitation, become a useful indicator. A DSS must, therefore, be able to transform a set of raw and scattered data into an organized system to provide, in an optimized manner, the data needed for decision-making [10].

Generally, the architecture of a DSS can be summed up in three distinct phases:

- The construction phase, which combines the extraction of multi-source and heterogeneous data, the transformation or cleaning of this data and then the loading into the data warehouse;
- The selection phase consists of specifying the data warehouse requirements;
- And finally the analysis phase to perform the reporting and various analyzes.

## 2.2. Spatial Data Warehouse

According to [11], a data warehouse is a collection of thematic (topic-oriented), integrated, nonvolatile, and historical data for support of a decision-support process. Thus, it makes it possible to store relevant data of the organization with a view to making a decision in a shorter period of time. Generally, a data warehouse is considered a centralized repository that consolidates in a consistent format data from different sources that can be transformed and stored in a historical manner for decision-making analysis [1].

A Spatial Data Warehouse (SDW) is a collection of spatial data and non-spatial: subject-oriented, integrated, time variant, and nonvolatile dedicated to spatial decision-making [12]. An SDW is a reformulation of a data warehouse with the sole difference of integrating spatial nature. The latter impose new operators of navigation and a reformulation of classical concepts of dimensions and measures to take into account the spatial component [13].

A dimension corresponds to the axis of the analysis; it is described by one or more hierarchies. A fact models the subject of the analysis; it is formed of corresponding measures to the information of the analyzed activity. The dimensions are in the form of a hierarchically organized list of elements. In spatial data warehouses, dimensions and measurements may contain spatial components. There are three types of spatial dimensions:

- Geometric: each element of the dimension is associated with a geometric object;
- Non-geometric: in this case, the dimension contains only data non-spatial. The data only have the nominal value of the spatial reference (name of the country, name of the department, name of the city);
- Mixed: this at the same time contains geometric and non-geometric data.

Measurements of a fact are usually numerical and provide a description quantitative of the fact. A fact is associated to one or more combinations of dimension members. Some measures can be calculated from other measures or mem-

ber properties. They are called derived measures [14]. Space measurements are the result of a calculation obtained by the application of a space operator, they can be in the following three forms: geometric objects (polygons, circles, points, etc.), metric values (area, distance, length) or spatial pointers. In the latter case, geometric shapes are stored in a separate file and this when the technology used does not allow storage in the data warehouse structure.

A relational SDW inherits several components of traditional data warehouses, such as fact tables, dimensions, numeric measures, and hierarchies which aggregate these measurements according to several levels of granularity. The dimensions must be organized according to several levels of granularity or according to different levels of detail to support the operations of the spatial analysis system that we will develop in the next section 2.3.

### 2.3. Spatial Analysis System

According to [15], spatial analysis is “the analysis of phenomena whose location and extent in this space is unknown”. In other words, it is the implication of the phenomena to be analyzed in space.

One of the concepts introduced by [16] on the spatial analysis system is the Spatial On-Line Analytical (SOLAP). This system is based on integrating GIS capabilities into the OLAP system [17]. Three categories of solutions for implementing this SOLAP system [16]:

- OLAP dominant integrates all the functionalities of an OLAP tool and only a very small subset of the functions of a GIS;
- GIS Dominant offers all the functionality of the GIS tool, but only a subset of the OLAP tool’s features;
- Hybrid solution integrates both GIS and OLAP functionality into an environment that allows to explore and to analyze data using spatial components.

Several geo-decisional systems have presented these different functionalities to perform spatial analysis. These systems highlight several technologies to implement the SOLAP system.

Camptocamp realized, in 2008, a SOLAP prototype called GeOLAP which proposes an interactive cartographic interface presenting sociodemographic indicators. This application allows the user to realize, on the map, space operators (slice, drill-down, drill-up, etc.). [7] proposes a web solution coupling the GIS and OLAP. GoOLAP combines both OLAP features and advanced geovisualization features

There is also Map 4 Decision, which is an achievement of the company Intelli in collaboration with the University of Laval Quebec and Sovar. It is an integral solution of geospatial dimensions in a decision-making environment for strategic planning [17].

### 2.4. Modeling a Spatial Data Warehouse

Data Warehouse data models have been designed to facilitate decisional analysis

[18]. As a result, the conceptual model of a Data Warehouse has its particularities in terms of modeling. This model is based on three main types of schemas: star schema, snowflake schema and constellation schema. The design of these schemas depends on the data warehouse to be modeled: temporal, spatial or both.

Spatially referenced data needs a well-defined model to design the structure of the future database. This is the reason why several authors have immersed their research in this field.

Most of this research is based on the Unified Modeling Language (UML). And they present themselves as a complement or an extension of this language. Reference [19] and his team, from the Geomatics Center in University of Laval Quebec, have developed a general method for extending any graphical modeling formalism for geographic design. Thus, they proposed a Plug-in for Visual Language (PVL) extension for UML formalism and developed a module called Perceptory that can integrate with Microsoft Visio [16] [19] [20].

[21] also focuses on the use of UML for conceptual models of spatio-temporal applications. They have introduced in his work a new technical modeling, based on minimal extensions of existing models, developed to take into account the particularities of the combined spatial and temporal information. In this regard, the Entity-Relation (ER) model and the UML have been extended as prototypical examples.

Another model describes the warehouse as a central repository of complex and temporal data is also proposed by [22] and [23]. Reference [22] defines data manipulation algebra of the warehouse. This algebra based on an extension of the object algebras proposes temporal operators and state set manipulation operators of the objects of the warehouse.

Reference [22] establishes a model integrating three concepts: the warehouse object, the warehouse class and the environment. The latter defines the temporal parts in the warehouse schema with a relevant granularity (attribute, class and graph).

Reference [24] presents a model-driven approach to the design of the Multi-dimensional Data Warehouse. The strength of its approach lies in the modeling of the warehouse data while including the associated ETL (Extract, Transform and Load) processes. This approach, defined in the methodological framework of Model Driven Engineering (MDI), covers the entire design cycle of the Data Warehouse.

[25] proposes a UML profile developed specifically for the conceptual modeling of geographic databases called GeoProfile. They see this work as the first step towards unification of the different existing models, aiming mainly at semantic interoperability.

The last work we will see is that of [26]. They introduce the MultiDim Entity Relation (ER) model, a conceptual model that can be used to represent multidimensional data for the Data Warehouse (temporal, spatial and spatio-temporal) and OLAP applications.

## 2.5. Synthesis on the Basic Concepts

The previously presented concepts allowed us to gather the necessary elements for the realization of a decisional system based on spatial analysis. We have described the concept of decision-making systems and spatial analysis. The establishment of a decision-making system combined with spatial analysis within a ministerial department is, therefore, paramount in view of its activities. This system can provide for this ministerial department a profit in terms of efficiency and performance. These various concepts allow the specification and adoption of suitable and flexible solutions.

The requirement of the ministerial department is to take charge of the geographical components. To meet this requirement, department officials use a newly developed spatial analysis system to better understand the characteristics of phenomena studied in space. And in the field of computing, this analysis is often associated with the use of SOLAP technologies.

For the design of our system which is based on the concepts of spatial data warehouse, we look to the use of the MultiDim model proposed by [26] for the following reasons:

- It introduces fundamental extensions for the development of spatial data warehouses;
- Its logical model remains easy to read by integrating spatial and temporal extensions into the model;
- The model also has extensions necessary for the design and implementation of the various phases and the improvement of the efficiency of these phases.

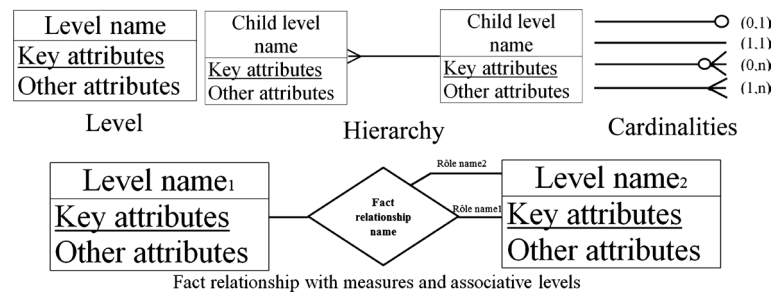
We have therefore collected all the elements necessary to carry out our work and which focuses on the analysis of agricultural production in Madagascar.

## 3. MultiDim Model

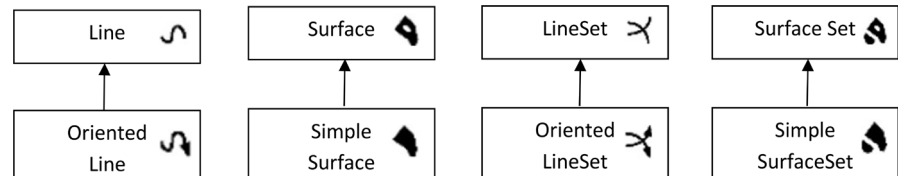
In this section, we define the MultiDim model [26], a multidimensional conceptual concept model. This model allows representing at the conceptual level all the elements required in the data warehouse and OLAP applications, *i.e.* dimensions, hierarchies, and the facts with associated measures. In that, the studied system has only one table of fact and this can be linked with several dimension tables. The graphic notation of the MultiDim model that we use is shown in **Figure 1**; it is similar to that of the entity relation model.

Several spatial data types can be used to represent the spatial extent of real world objects. In our model, we use the spatial data types defined by the conceptual spatiotemporal model show in **Figure 2**. These data types and their associated icons are shown in **Figure 2** and are organized in a hierarchy.

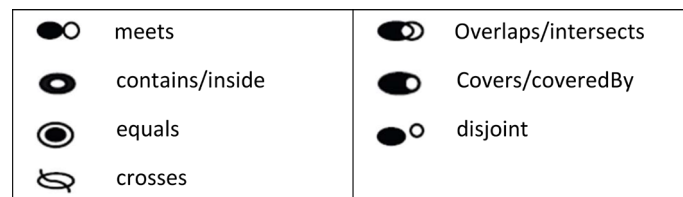
Topological relationships specify how two spatial values relate to each other. The existing topological relations are: meets, contains/inside, equals, crosses, overlaps or intersects covers or covered by and disjoint. To represent these topological relationships in our model, we use the associated icons given in **Figure 3**.



**Figure 1.** Some notation of the MultiDim model.



**Figure 2.** Some spatial data types.



**Figure 3.** Icons for the various topological relationships.

## 4. Decisional System for Spatial Analysis of Agricultural Production

### 4.1. Requirements Analysis

In this section, we will specify the needs that the application must meet. These needs are divided into two parts: functional needs and non-functional needs. The functional requirements are vital needs for a project because they specify not only the processes in its implementation, but especially the expected functionalities. For this project, these requirements are defined around the following features:

- Identify and analyze all agricultural production data available to the Ministry;
- put in place a decision-making system to exploit these data;
- Implement a data storage system for the decision-making system (data warehouse);
- Put in place an application to analyze these data in space (display on a map) and manipulate them in order to draw a synthesis for the future decision;
- Integrate in the application a functionality allowing to make a report.

For the non-functional requirements, they are the subject of the following points:

- Ensure the consistency and integrity of the data;
- Keep the stored data up-to-date to increase knowledge about the agricultural

production in Madagascar;

- Provide aids for the use of the application and assist the manager of the system in handling it;
- Ensure the security of the spatial data warehouse on agricultural production in Madagascar.

## 4.2. Analysis and Criticism of Existing System

The study of the existing is the point of passage that materializes the first contact of the designer with users. According to our interview with the Directorate of Information Systems, the Agricultural Statistics and Livestock Service (StatAgriEl) keeps the data on the situation of Madagascar in terms of agriculture and livestock. It establishes, every two or three years, depending on the availability of data, an agricultural statistical yearbook.

Following the interview, the general problem of the department lies in:

- The StatAgriEl Service does not yet have a database concerning its field of activity;
- They still manipulates Excel files to make reports and establishes a static cartography on the distribution of the productions;
- In addition, the Development Service (IDSR) has not yet put in place a decision-making system on the needs of decision-makers;
- Lack of decision support tool;
- Lack of a structured database or data warehouse.

Criticism of the existing is a crucial and important step in the realization of a project. It aims to make an objective judgment in order to identify any shortcomings encountered during the study of the existing in order to propose solutions to the problems posed [27].

In addition, two aspects are always revealed during this criticism, one of which is positive and the other negative. These two aspects deserve to be raised as the need for perfection will always be desired by users for proper operation.

At the end of the analysis of what exists, it must be admitted that the Ministry of Agriculture has at least one well defined organization system of the functional and organizational point. The existence of a Directorate occupying information systems is a reflection of this functioning. Thus, the merger of two Ministries into one strengthens the capacity of the crew members. The fact of having sufficient human resources easily makes it possible to carry out the missions conferred on it. For the good functioning of the Ministry, the Information Systems Department offers a better service to other departments and assists users in the field of information technology.

On the other hand, the Ministry of Agriculture, like most public institutions, is still lagging behind technological development because of lack of resources. As a result, even if the Information Systems Department wants to carry out a project, it cannot go beyond it.

The lack of a decision-making system slows down decision-making regarding,



for example, the import or export of agricultural products, especially at the right time. The absence of a data storage system makes it difficult to manipulate these data and it turns out that these data cannot be exploited well.

Another problem posed to the Information Systems Department is the lack of material resources because the computers used are less efficient. These performance gaps have an impact on the development time.

### 4.3. Conceptual Model for a Decisional System for Spatial Analysis

This section presents the modeling of the data warehouse for the analysis of agricultural production in Madagascar. This modeling is preceded by an inventory of requirements which is the subject of the section 4.1. According to [28], the analysis of requirements of a data warehouse system should aim to analyze, understand, and model the business context a data warehouse. So, our model is based on the Multidim notation (cf. 3).

Complementary to the approach introduced by Malinowski and Zimányi [26], we use the method of R. Laurin, 2015<sup>1</sup>, to identify the decision model from the target analysis. This method is both very practical and very effective. The principle is simple: highlight the analysis axes (dimensions) with their attributes, as well as the main elements to be analyzed (facts).

The modeling will be done using a table by “fact”. Each column denotes a dimension, each element of each column means the characteristics of the dimensions, and in the first column are placed the features built into the table of facts (measures, factors).

The analysis phase of the project allowed us to complete the **Table 1** that constituent elements of our spatial data warehouse, by asking the following series of questions to decision makers:

- What do you want to analyze? => The last line of the table (Fact);
- What are your criteria for analysis? => The first line of the table (Dimensions);
- Until what level of detail do you want to go through? => Measures in fact and dimensions Attributes.

From **Table 1**, we can now model our data warehouse. The adoption of the

**Table 1.** Identification of the constituent elements of data warehouse.

Fact Table	
Name	Element
Production	Index of production area cultivated, rate of production
Dimensions Tables	
Name	Table
Product	Type of product, designation of product
Communities territorial	Province, Region, District
Time	Year

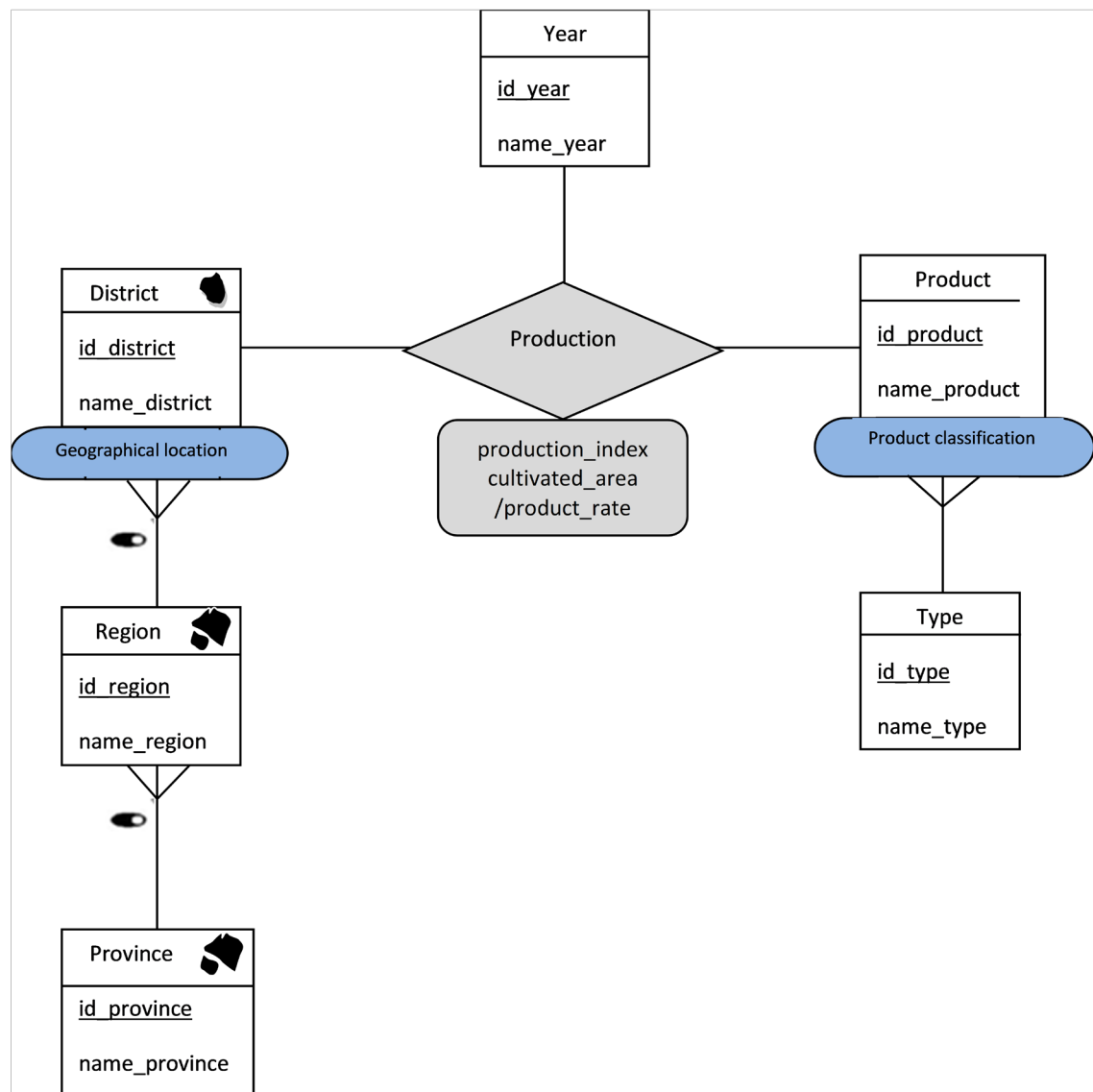
<sup>1</sup>Enseignant à l'Université de Sherbrooke (Québec, Canada).

scheme for the model therefore depends on the advantage or disadvantage of each. For our model, we chose the snowflake diagram for the following reasons [29]:

- Formalization a hierarchy within a dimension;
- Maintenance tables of simplified dimensions;
- Reduction redundancy.

The conceptual model of our data warehouse designed for agricultural production is shown in **Figure 4**.

In order to provide decision-makers with an efficient means of spatial analysis, in their process of analyzing phenomena as important as agricultural production, we have integrated geographic data into a data warehouse supporting the decision-making system to set up a geo-decisional system. In this section, we propose the conceptual model (cf. **Figure 4**) of the spatial data warehouse implemented. As



**Figure 4.** Conceptual model of the agricultural production data warehouse.

we have already mentioned, in this study we based ourselves on the needs analyzes of the Ministry of Agriculture which presented missing shortcomings (cf. 4.1). As a result, our proposed model takes care of these shortcomings, particularly in the integration of spatial data through the *production* dimension which constitutes a predominant factor in the analysis. Decision-makers can thus take advantage of the advantages of the map and can therefore use the map in an interactive way. Examples are presented in the section devoted to the implementation of our geo-decisional system (cf.5). In fact, current GIS focus on the topographic aspect, but our system adds to these aspect measurements taken from the warehouse and which will be displayed on the map. We are carrying out a design based on the MultiDim model of [22] (cf. 3).

The dimensions are organized in hierarchy like the dimension presenting the administrative limit *District, Region, Province*, the dimension presenting the time as *Year* and the dimension representing the product and its type *Product, Type*. Each member belongs to a particular hierarchical level (or level of granularity).

The time dimension is represented by the *Year* table having as attribute *id\_year* which is the identification of the year and *name\_year* its label. The spatial dimensions are represented by the *District, Region* and *Province* tables respectively having identifications *id\_District, id\_Region, id\_Province* attributes and as attributes names *name\_District, name\_Region, name\_Province*.

The Product and Type tables represent the respective dimensions of the products and their type.

The fact table is represented by the *Production* table containing the measures of production index, cultivated area and production rate which are represented respectively by the attributes *production\_index, cultivated\_area, production\_rate*.

## 5. Implementation of a Decisional System for Spatial Analysis

### 5.1. Extraction Transformation Load Tools

To load data in the data warehouse is often use an Extraction Transformation Load (ETL) tool to facilitate the selection and cleaning of the data needed for the analysis. Among the ETL tools standard and parameterizable developed [30], there are the commercial tools such Informatica Power Center, Microsoft SQL Server Integration Services, Warehouse Oracle Builder, SAS BusinessObject, etc. There is also the open source alternative: Talend Open Studio, Pentaho Data Integration (PDI) also known as Kettle and CloverETL from Javlin, which are the most complete and recognized.

However, the existence of these tools does not prevent the development of a tool ETL specific in case of need. This is the case in this project for the non-spatial data. Since the sources of our data are all Excel files, so we have decided to develop a program to read these files, extract useful data and load directly into the data warehouse. This program is written in JAVA, is based on the use of the libraries of “The Apache Software Foundation”, called Apache POI3

(Poor Obfuscation Implementation). For spatially referenced data, we used another ETL tool, called GeoKettle, which can manipulate data of this type.

## 5.2. Database Management Software

Several Database Management Systems (DBMS) software choices are available on the market, from free DBMS (MySQL, PostgreSQL) up to DBMS specifically for professionals (MS SQL server, Oracle), with more features, but more expensive [31]. But our first choice remains on the use of Open Source. PostgreSQL is selected in this project to provide a fast, stable and standards compliant database with an excellent geospatial extension PostGIS.

## 5.3. Map Server

The map server processes the user's request to display maps on his computer station [32]. The user, from its terminal, makes requests via the Internet communication protocol (TCP/IP), to request the display of a specific card; the map server interprets this request and returns the map in the form of a matrix image (png, jpg, ...) or vector (svg, swf, ...). The map engine can be controlled by languages of 54 script such as PHP, javascript, Python or Perl that allow it to dynamically generate a map in response to a request prepared by a user interface [32].

Many publishers offer Open Source solutions for the server MapServer, GeoServer, QGISServer, Mapnik, etc. We will be very clear about using GeoServer against others servers. Despite its weak points, GeoServer has many advantages and offers various features. It is a web server used to serve maps and data from a variety of formats, from standard web and desktop clients to GIS programs advances. In other words, we can store our spatial data in almost any of our favorite formats but also that our users do not need to know exactly what GIS data is. GeoServer is the reference implementation of the Open Geospatial Consortium (OGC) for Web Feature Service Standards (WFS) and Web Coverage Service (WCS), as well as a certified high performance server Web Map Service (WMS) compliant.

## 5.4. Spatial Analysis Tools

We use SpagoBI and GeoMondrian as analysis tools.

SpagoBI is entirely Open Source Business Intelligence suite. It covers all the analytical areas of Business Intelligence projects, with innovative themes and engines. SpagoBI offers a wide range of analytical tools, as reporting, multidimensional analysis, charts, ad-hoc reporting, location intelligence, data mining, network analysis, ETL, collaboration, masterdata management and external process.

GeoMondrian is an Open Source Spatial Online Analytical Processing Server, a spatially-enabled version of Pentaho Analysis Services (Mondrian). GeoMondrian is the first implementation of a true SOLAP server. It provides a consistent integration of spatial objects into the OLAP data cube structure, instead of fetching them from a separate spatial database, web service or GIS file. To make a simple

analogy, GeoMondrian brings to the Mondrian OLAP server what PostGIS brings to the PostgreSQL database management system. It implements a native geometry data type and provides spatial extensions to the Multidimensional Expressions (MDX) query language, allowing embedding spatial analysis capabilities into analytical queries. These geospatial extensions to the MDX query language provide many more possibilities, such as in-line geometry constructors, member filters based on topological predicates, spatial calculated members and measures and calculations based on scalar attributes derived from spatial features. GeoMondrian only supports PostGIS based data warehouses.

To manage the OLAP server, SpagoBI calls Mondrian (SpagoBIPivotEngine). It is already equipped with an OLAP client (JPivot) for server queries

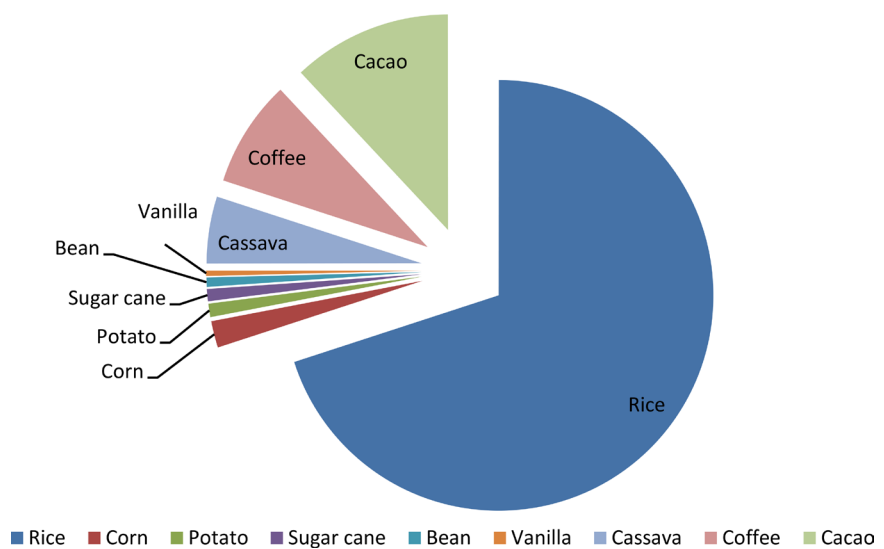
## 6. Presentation of a Decisional System for Spatial Analysis

### 6.1. Realization with SpagoBI Tool

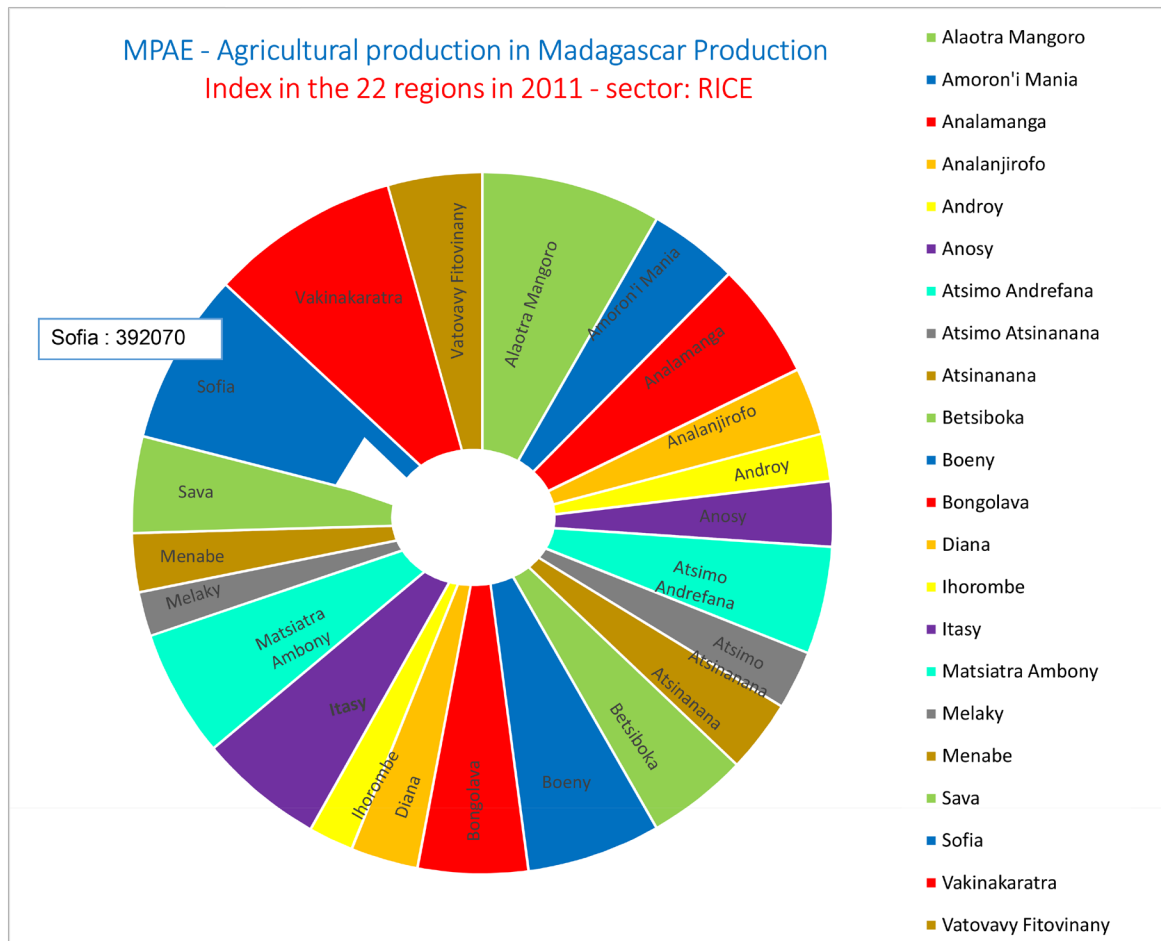
SpagoBI offers various functionalities to allow the effective exploration of data, whether in the field of reporting (circular diagram, bar chart, multidimensional analysis, Spatial analysis) or in rendering cartographic. It is, therefore, really necessary to have a geo-decisional tool in the functioning of a ministerial department.

The circular diagram allows us to visualize our data according to the chosen parameters previously. In our application, we have set as parameters the “District”, the “Year” and the “Product”, which are none other than the dimensions (cf. 4.3), according to the analysis at carry out. So, we have a dynamic diagram following the parameters. In **Figure 5**, we show a circular diagram to visualize the production index, which takes the District and the year as parameters. **Figure 6** shows the circular diagram of the distribution of rice production values for each region in 2011.

The bar graph shows a comparative analysis of the production index (yield) versus the area cultivated for each type of product. **Figure 7** shows the bar chart



**Figure 5.** Circular diagram showing the production index for each product.



**Figure 6.** Circular diagram showing rice production for each region.

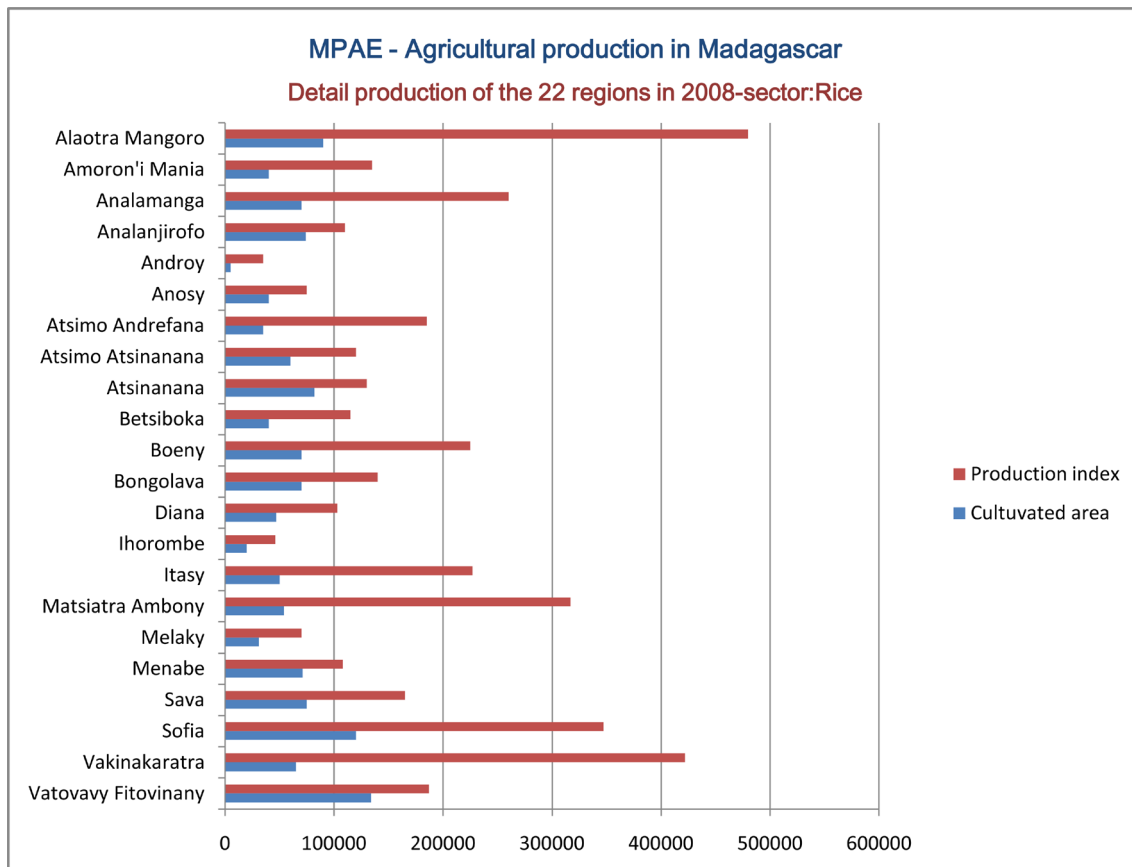
of the rice production details for each region (production index and area cultivated). This diagram allows you to know the productivity of a product in relation to the cultivated area. Knowing this information allows decision-makers to provide their aid and support farmers to increase yields through the use of fertilizer and good seed, for example, or learning a better technique.

SpagoBI also offers the functionality of OLAP which allows analyzing the different dimensions of multidimensional data that we modeled in section 4.3. SpagoBI integrates an OLAP server (Mondrian) in its engine. **Figure 8** shows an example of how this functionality is implemented.

SpagoBI use Location Intelligence application for spatial analysis in our data warehouse.

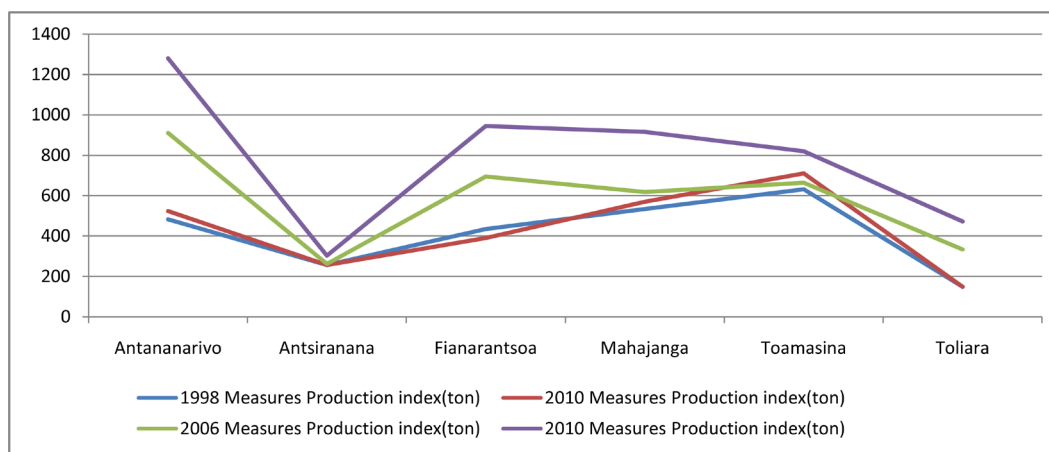
Regarding our analysis, we did it according to the geographical hierarchies existing in our model (cf. 4.3) province, region and district. For this spatial analysis, the principle is the same for these hierarchies. We choose the type of product and the desired year; and running this request, SpagoBI displays the data concerning it on the map. SpagoBI uses GeoServer to be able to overlay the data with its layer.

**Figures 9-11** respectively shows a spatial analysis of production by province, region, district depending on the product and year chosen.

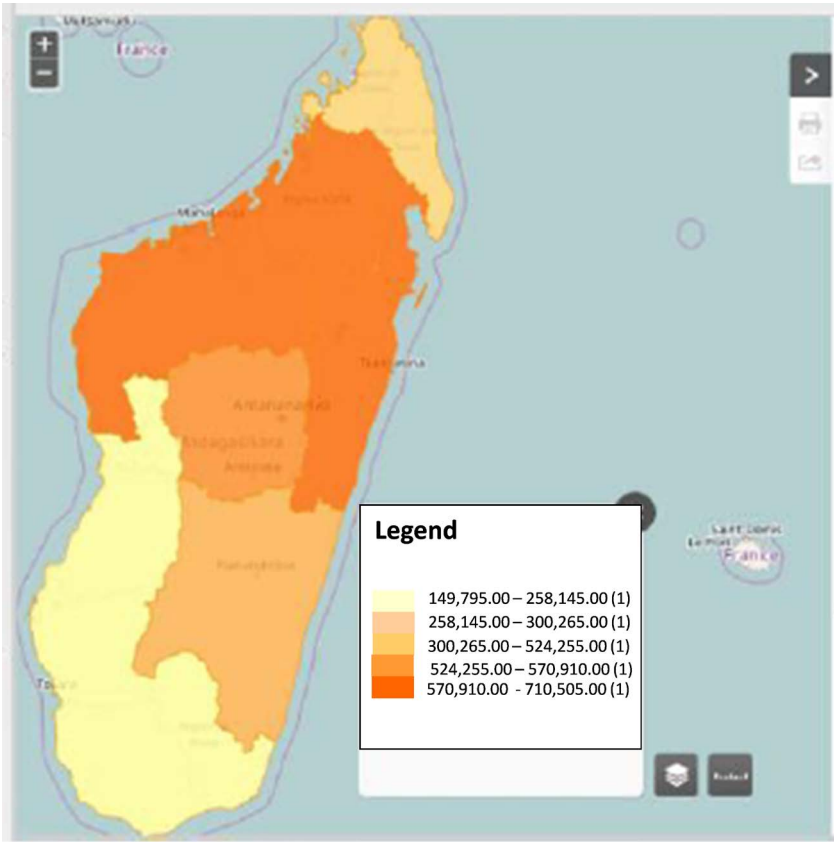


**Figure 7.** Bar graph detailing rice production index and cultivated area for each region in 2008.

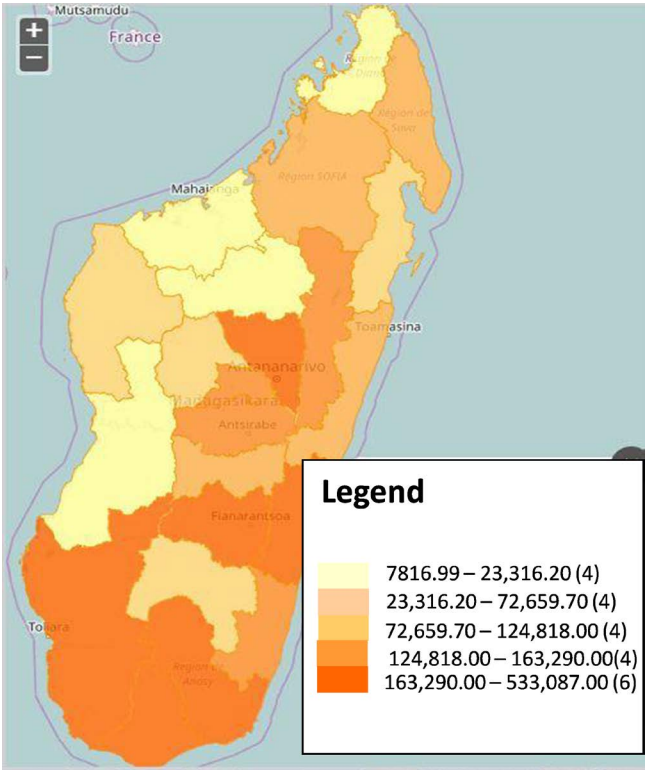
	Year			
	1998	2002	2006	2010
	Measures	Measures	Measures	Measures
Territorial collectivity	Production index(ton)	Production index(ton)	Production index(ton)	Production index(ton)
Antananarivo	482.555	524.255	911.41	1281.12
Antsiranana	255.97	258.145	262.69	302.835
Fianarantsoa	435.385	390.265	695.905	944.595
Mahajanga	533.405	570.91	618.16	915.865
Toamasina	631.81	710.595	664.675	820.95
Toliara	148.685	149.795	335.09	472.6



**Figure 8.** OLAP application with JPivot in SpagoBI.



**Figure 9.** Spatial analysis of rice production by province in 2002.



**Figure 10.** Spatial analysis of manioc production by region in 2005.

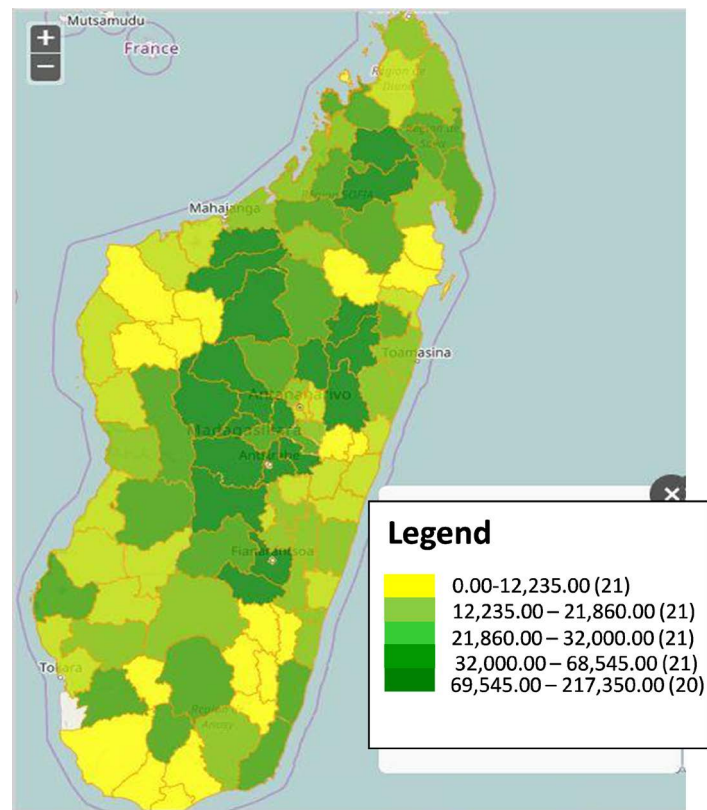


## 6.2. Realization with GeoMondrian and JPivot

The purpose of building a data warehouse is to be able to offer the user a interface, using an OLAP client, allowing it to perform interactive analysis and reporting. The main goal is to discover, from a single visualization, knowledge that can subsequently influence the decision of decision-makers. In our case, the client to run OLAP navigations is JPivot. This OLAP client comes with GeoMondrian and built from a Java Server Page (JSP) tag library. **Figure 12** shows our OLAP client main interface.

In **Figure 13**, we obtained, from the MDX query, a crosstab showing the production index and the growth rate for the selected years (1998, 2002, 2006, 2010). These values are distributed in the territorial collectivity (province), and we can now play on this table using OLAP operators.

From the graph toolbar, we can choose the different types of graph allowing to graphically representing the table. In **Figure 14**, we represent the values of the production index for each province by year using 3D diagram.



**Figure 11.** Spatial analysis of rice production by district in 2011.

	year							
	1998		2002		2006		2010	
	Measures		Measures		Measures		Measures	
Territorial collectivities	1998, Production Index(ton)	1998, Rate of growth(%)	2002, Production Index(ton)	2002, Rate of growth(%)	2006, Production Index(ton)	2006, Rate of growth(%)	2010, Production Index(ton)	2010, Rate of growth(%)
Antananarivo	1,482,980.00	-3.41%	1,526,060.25	-5.00%	2,054,595.00	1.67	2,454,062.50	1.1
Antsiranana	964,160.00	1.96%	962,790.00	-1.00%	565,565.00	1.20%	620,255.38	7.89
Fianarantsoa	2,166,725.00	-1.01%	2,096,175.00	-6.16%	2,241,845.00	-0.66	2,486,658.25	-0.68
Mahajanga	1,253,320.00	-3.93%	1,281,410.00	1.43%	952,300.31	2.90%	1,247,659.75	0.78
Toamasina	1,334,160.00	0.39%	1,448,215.00	0.47%	1,131,669.00	1.38%	1,310,164.63	2.5
Toliara	1,030,310.00	-1.52%	1,047,615.00	-3.85%	1,761,985.00	1.60%	1,862,098.50	-2.32

**Figure 12.** OLAP client main interface.

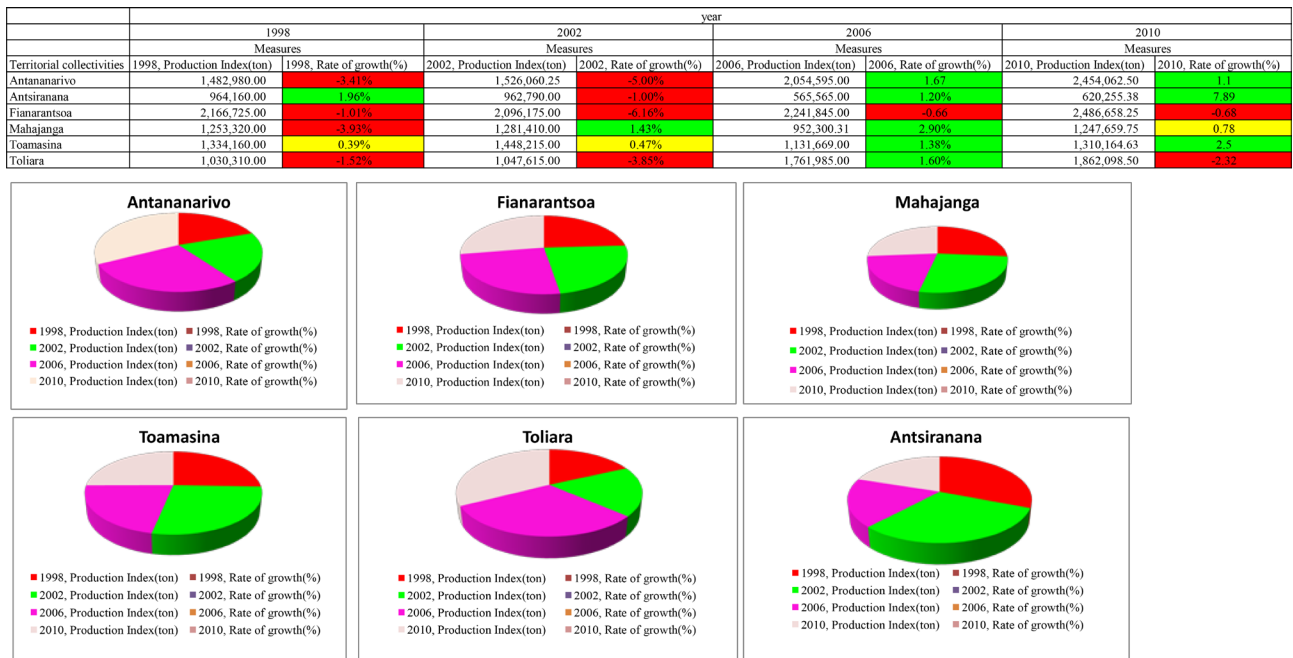


Figure 13. Cross table of dimensions and measures with diagram.

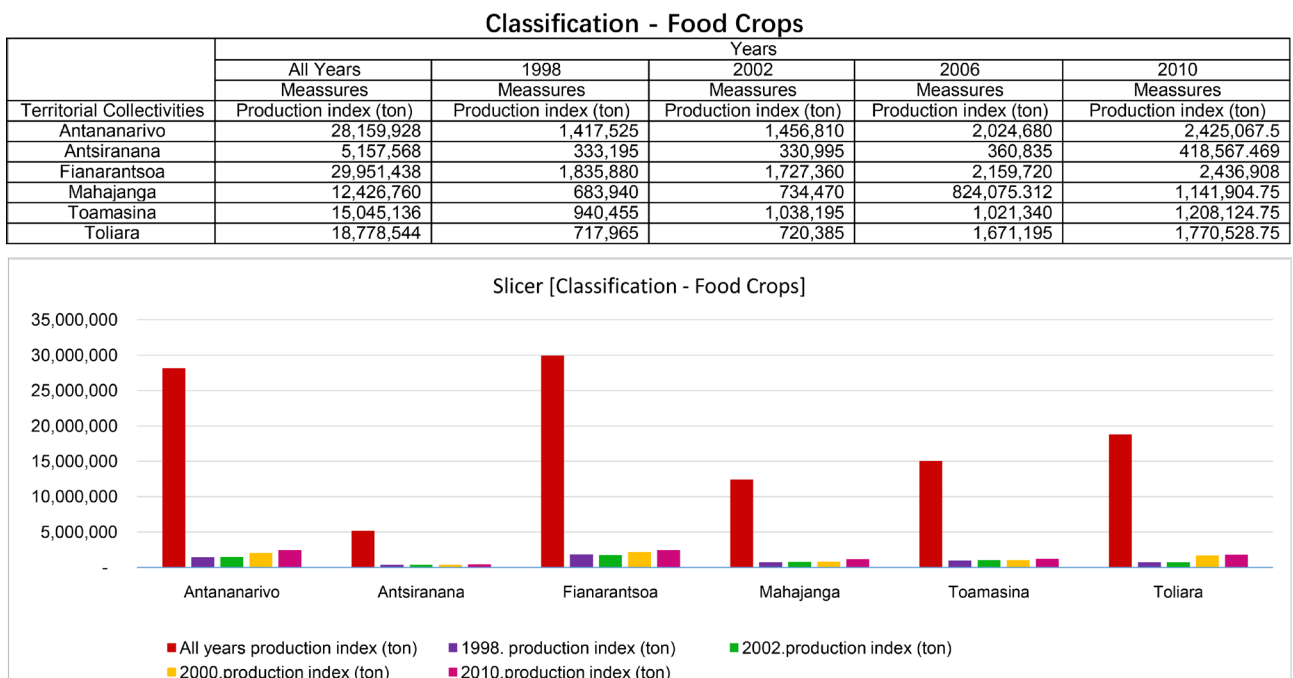


Figure 14. 3D Diagram of the cross table of production index values.

Figure 15 show the results of the nesting query. The lines of this table contain two dimensions, product and local authorities. But the geographic dimension is nested in the product dimension.

## 7. Conclusions and Perspectives

Although the Ministry of Agriculture has a large amount of data regarding its

		Year				
		1997	2001	2005	2009	2011
		Measures	Measures	Measures	Measures	Measures
Product	Territorial collectivities	Production Index (ton)	Production Index (ton)	Production Index (ton)	Production Index (ton)	Production Index (ton)
Food	Antananarivo	1 469 105,25	1,537,090	1 990	2 398 372	2 207 483,25
	Antsiranana	315,355	332,635	353 401,125	371 868,406	375 510,688
	Fianarantsoa	1 892 835	1 865 060	2 141 230	2 436 162,75	2 290 772,75
	Mahajanga	708 034,5	719,555	799 588,938	1 103	1 147
	Toamasina	940 185,125	1 032 580	1 003	1 170	956 809,25
	Toliara	744,435	767,830	1,646,036,50	1 814 910,5	2 433 104,75
Industrial cultures	Antananarivo	66,035	69,015	30 053,576	28,850	15,715
	Antsiranana	614,950	613,070	188 440,031	186 349,078	187 640,828
	Fianarantsoa	269,740	338,280	81 358,156	18,620	6,895
	Mahajanga	595,320	542,220	116 116,406	122,745	109,745
	Toamasina	357,270	376,595	92 824,508	93,445	86,640
	Toliara	301,200	321,020	87 090,891	91,250	76,650
Export cultures	Antananarivo	240	305	240.123	190	135
	Antsiranana	15,360	17,180	17 025,883	16,660	13,620
	Fianarantsoa	26,300	30,470	34 169,031	48,805	22,140
	Mahajanga	1,295	1535	9 767,187	11,510	10,495
	Toamasina	31,522	32,295	19 895,037	14,065	4,280
	Toliara	605	720	1 091,035	200	100

**Figure 15.** Nesting of the geographic dimension in the product dimension.

activities and problems arise when making decisions due to a lack of decision-making system to exploit and analyze this data.

Faced with these problems, we have put in place a decision-making system to perform a spatial analysis of the data. This solution began with the establishment of a data warehouse to bring all the data together. The realization of this warehouse must go through the design phase and we have chosen to use the MultiDim model introduced by Malinowski and Zimányi. And subsequently, its implementation was done in a ROLAP server which used relational database technology.

The implementation of this system, which allows the analysis of all of this data, required the use of a lot of free technologies. To do this, we used Open Source tools (SpagoBI and GeoMondrian) which have functionalities that meet the needs of decision-makers in terms of spatial analysis.

However, this system is not completely perfect, it lacks the functionality to perform production forecasting and agricultural statistics. We offer avenues for research on this problem. The first is the use of data mining techniques based on the machine learning method to construct a forecast map of agricultural production. The second is the use of statistical methods to establish a model of agricultural production. The third is the establishment of a system that takes into account the factors that influence production in Madagascar, such as climatological variables (precipitation, temperature, soil moisture, etc.), soil properties and socioeconomic variables of each decentralized community as well as geopolitical parameters.

## Acknowledgements

Writing this paper would not have been possible without the help of Thérèse Libou-

rel at Université Montpellier-UMR Espace-dev (IRD 228 Space for Development) France, PHD-HDR Computer Science. Her experiences in the Spatial Data Warehouse domain, her advice and kindness were proved irreplaceable. We would like to thank her sincerely.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Farve, C., Bentayeb, F., Boussaid, O., Darmont, J., Gavin, G., Harbi, N., Kabachi, N., and Loudcher, S. (2013) Les entrepôts de données pour les nulls... ou pas!. 2e Atelier aIde à laDécision à tous les Etages en conjonction avec les 13ièmes Journées Francophones d'Extraction et de Gestion des Connaissances (EGC/AIDE 2013). Extraction et de Gestion des Connaissances, Toulouse.
- [2] Meta Group (1999) Data Warehouse Scorecard: Cost of Ownership and Successes in Application of Data Warehouse Technology. META Group, Stamford.
- [3] Barquin, R.C. (1997) A Data Warehouse Manifesto. In: Barquin, R. and Edelstein, H., Eds., *Planning and Desingning the Data Warehouse*, Prentice-Hall, Upper Saddle River, 3-16.
- [4] Keenan, P. (1996) Using a GIS as a DSS Generator. In: Darzentas, J., Darzentas, J. S. And Spyrou, T. Eds., *Perspectives on Decision Support System*, Grèce University of the Aegean, Mytilene, 33-40.
- [5] Franklin, C. (1992) An Introduction to Geographic Information Systems: Linking Maps to Databases. *Database*, **15**, 12-21.
- [6] Bédard, Y. (1997) Spatial OLAP. *2nd Forum Annuel sur la R-D, Géomatique VI: Un monde Accessible*, Montréal, 13-14 November 1997, 13-14.
- [7] Bimonte, S., Di Martino, S., Bertolotto, M., Ferrucci, F., Leano, V. and Pinet, F. (2009) GoOlap: Vers l'Analyse Spatio-Multidimensionnelle à l'aide du Geobrowser Google Earth. *Actes du XXVII congrès Informatique des Organisations et Systèmes d'Information et de Décision*, Toulouse, 26-29 May 2009, 379-396.
- [8] Van Damme, M.D. (2010) Entrepôts de données dans le domaine spatial pour l'inventaire forestier. Mém. Ingéniorat: informatique. Conservatoire National des Arts et Métiers, Orléans, 90 p.
- [9] Richeton, N. and Chentouf, B. (2010) Livre blanc—Décisionnel Solutions Open Source. Disponible Sur Demande Sur, France.  
[http://www.smile.fr/content/smile/livreblanc/decisionnel\\_open\\_source.htm](http://www.smile.fr/content/smile/livreblanc/decisionnel_open_source.htm)
- [10] Direction Générale de l'Enseignement Scolaire (2012) Mise en place d'un système d'information décisionnel dans une entreprise [en ligne]. Direction Générale de l'Enseignement Scolaire, France.
- [11] Inmon, W.H. (1996) Building the Data Warehouse. 2nd Edition, John Wiley & Sons, New York, 401 p.
- [12] Stefanovic, N., Han, J. and Koperski, K. (2000) Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes. *IEEE Transactions on Knowledge and Data Engineering*, **12**, 938-958.  
<https://doi.org/10.1109/69.895803>
- [13] Ruiz, M.C. and Fernandez, L. (2009) Environmental assessment in construction us-

- ing a spatial Decision Support System. *Automation in Construction*, **18**, 1135-1143. <https://doi.org/10.1016/j.autcon.2009.07.005>
- [14] Blaschka, M., Sapia, C., Höfling, G. and Dinter, B. (1998) Finding your way through multidimensional data models. *9nd International Workshop on Database and Expert Systems Applications*, Vienne, 24-28 August 1998, 198-203. <https://doi.org/10.1109/DEXA.1998.707403>
- [15] Champoux, P. and Bédard, Y. (1992) Notions fondamentales d'analyse spatiale et d'opérateurs spatiaux. *Revue des Sciences de l'Information Géographique et de l'analyse spatiale*, **2**, 187-208.
- [16] Bédard, Y., Proulx, M.J., Larrivée, S. and Bernier, E. (2002) Modeling Multiple Representations into Spatial Data Warehouses: A UML-Based Approach. Symposium on Geospatial Theory, Processing and Applications, Ottawa, 9-12 July, 6.
- [17] Zaamoune, M. (2015) Intégration et optimisation des grilles régulières de points dans une architecture SOLAP relationnelle. Doctoral Dissertation, Informatique: Univ. Blaise Pascal—Clermont II, 152 p.
- [18] Kimball, R. (1996) The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. John Wiley & Sons, New York.
- [19] Bédard, Y. (1999) Visual Modelling of Spatial Databases: Towards Spatial PVL and UML. *Geomatica*, **53**, 169-186. <https://doi.org/10.5623/geomat-1999-0025>
- [20] Brodeur, J., Bédard, Y. and Proulx, M.J. (2000) Modelling Geospatial Application Databases Using UML-Based Repositories Aligned with International Standards in Geomatics. *Proceedings of the 8th ACM International Symposium on Advances in Geographic Information Systems*, Washington DC, 10-11 November 2000, 39-46. <https://doi.org/10.1145/355274.355280>
- [21] Tryfona, N., Price, R. and Jensen, C.S. (2003) Chapter 3: Conceptual Models for Spatio-Temporel Applications. In: Sellis, T.K., Koubarakis, M., Frank, A., Grumbach, S., Hartmut Güting, R., Jensen, C., et al., Eds., *Spatio-Temporal Databases*, Vol. 2520, Springer, Berlin, Heidelberg, 79-116. [https://doi.org/10.1007/978-3-540-45081-8\\_3](https://doi.org/10.1007/978-3-540-45081-8_3)
- [22] Teste, O. (2000) Elaboration d'entrepôts de données complexes. *18ème congrès INFORMATIQUE des ORGANISATIONS et Systèmes d'Information et de Décision*, Lyon, May 2000, 229-245.
- [23] Ravat, F. and Teste, O. (2001) Modélisation et manipulation de données historisées et archivées dans un entrepôt orienté objet. Communication dans un congrès: Bases de données avancées, Agadir, 243-256.
- [24] Atigui, F. (2013) Approche dirigée par les modèles pour l'implantation et la réduction d'entrepôts de données. Doctoral Dissertation, Université de Toulouse, Toulouse 158 p.
- [25] Lisboa-Filho, J., Sampaio, G.B., Nalon, F.R. and Borges Karla, A. de V. (2010) A UML Profile for Conceptual Modeling in GIS Domain. *CAiSE 2010 Workshop DE@CAiSE 10*, Hammamet, 7-9 June 2010, 18-31.
- [26] Malinowski, E. and Zimányi, E. (2008) Advanced Data Warehouse Design: From conventional to Spatial and Temporal Applications. Springer, Berlin, Heidelberg.
- [27] Kabongo, M.B. (2010) Mise en œuvre d'une solution informatique pour l'optimisation de la gestion des ouvrages dans une bibliothèque [en ligne]. Institut Supérieur de Statistique, Graduat.
- [28] Prakash, N. and Gosain, A. (2003) A Requirements Engineering for Data warehouse Development. *Proceedings of the 15th Conference on Advanced Information Systems Engineering*, Klagenfurt, 16-20 June 2003, 4.

- [29] Espinasse, B. (2013) Introduction aux entrepôts de données.  
<https://pageperso.lis-lab.fr/bernard.espinasse/Supports/DWDM/2-Intro-Entrepots-4p.pdf>
- [30] Vassiliadis, P. (2009) A survey of Extract-transform-Load technology. *International Journal of Data Warehousing and Mining*, **5**, 1-27.  
<https://doi.org/10.4018/jdwm.2009070101>
- [31] Dorival, A. (2012) Mise en place d'un outil de gestion pour le service éditique du RSI. Mém. Ingéniorat: Informatique. Paris: Conservatoire National des Arts et Métiers, 236 p.
- [32] Coutard, M., Klipfel, J.P. and Blanc, S. (2005) La cartographie SIG en ligne ou Web mapping: Les outils "libres" [en ligne]. *M@ppemonde*, **80**, 14.