
DATABASE AND GIS MANAGEMENT FOR THE INVESTIGATION OF HEAVY POLLUTED INDUSTRIAL SITES

Cédric PITOUT*, Claude KERGOMARD**

*National Research Center for Contaminated Sites and Soils (CNRSSP), France
pitout@cnrssp.org

** Sciences and Technology University of Lille, France
Laboratoire de Géographie des Milieux Anthropisés
Claude.Kergomard@univ-lille1.fr

Working Group IV/ 3

KEY WORDS: Geographical Information System (GIS), Database, Geo-referencing, Aerial Photograph, Industrial Contaminated Site.

ABSTRACT

Geographical Information Systems (GIS) are commonly used for territorial and environmental management. Recently, they have been applied to the topic of polluted industrial sites. Specific difficulties directly related to industrial site characteristics (heterogeneity, lack of data, confidentiality, ...) make it difficult for use and especially for analysing results. Based on our experience in the treatment of data from two sites located in northern France, three main concepts are to be taken into account for improvement in this framework:

- Database architecture : data structure and organisation has to be coherent by taking into account all aspects of site description and physico-chemical analysis, but has also to remain open by allowing to append new fields at any time. A particular attention has to be paid: non-redundancy of the records, classification in regards to the concerned medium (soil, water, air...), and structuring of data groups in a hierarchical system.
- Data geo-referencing : precise data geo-referencing appears essential for integrating field measurements, maps and older data within a GIS. Presently easy to use and not much expensive, the DGPS (Differential Global Positioning System) is a good way to resolve the question of geo-referencing. DGPS is particularly useful for comparison and overlay different kinds of maps, drawn at different times, etc. It gives, in common use, a metric to sub-metric precision (x, y, z), enough for our purpose.
- Aerial photographs : the study of site history is achieved with the treatment and the interpretation of aerial photographs. General site evolution during time can be assessed using a set of views since the oldest to the more recent one. Better still, it permits to visualise the soil movements and/or the waste zone location. Therefore, the aerial photograph collection constitutes an interesting support to locate primary pollution sources.

As a conclusion, this study should allow industrial site owners and administrators to better define investigation techniques and environmental management tools.

1 INTRODUCTION

Since a few years in France, there is an increasing awareness of the necessity of contaminated sites management and rehabilitation. Simultaneously, there is an increase in environmental GIS applications and needs. It thus appears necessary to adapt GIS for managing data specifics to industrial sites. After a short presentation of the data model in a cartographic database, we shall stress two major points that are necessary for a correct GIS utilization : data precise geo-referencing and sites historical reconstitution. These points have been assessed from two cases-studies of former industrial sites in the Nord-Pas-de-Calais region (northern France). The first site is a former coke plant with main pollution by coal tar, the second one is a former ore treatment factory with a pollution by heavy metals. All the figures presented in this paper correspond to the coke plant.

2 DATA MODEL

2.1 Context

First of all, data are divided in two classes. The first one is geometrical data generated by the GIS proper (points, lines, polygons with their coordinates), whereas a DBMS (DataBase Management System) manages environmental data. This paragraph focuses only on the last type of data. This approach can be applied no matter what GIS and DBMS software is used. For this work, ESRI ArcView3.0a© GIS and Microsoft ACCESS97© have been used.

Before to define a data model, it is necessary to synthesize the available information. The interest is to store a large amount of qualitative and quantitative parameters to describe the site. By another way, it does not seem advisable to consider an exhaustive list of pollutants at the risk of weigh down the database. For this reason only parameters specifics to our case-studies are indexed. Nevertheless, it is important to think of the universality of the GIS since each industrial site is particular with very heterogeneous data sets. The database model must also allow later append of new fields without important modification of the initial structure.

2.2 Environmental data and data format

Data result essentially from piezometers, cores, pits, cuttings, packings. For the two last examples, some problems are highlighted like reworked soil samples and the geo-referencing of these samples. Whatever the data source, investigations are often the same, nevertheless separated databases are built according to the kind of data : soil samples and water samples in our case. To resume, the table 1 proposes the environmental data collating. For the chemical databases, the last column indicates the number of pollutants for each medium.

Some databases have qualitative fields, i.e. the data format is not numeric but text. In this case, data information extraction with Syntax Query Language (SQL) could be problematic. Thus, qualitative data are integrated only for information, except if a data coding lexicon is created.

Database	Water	Soil	Number of pollutants / parameters
Geologic profile	X	X	
Major element analyses	X	X	20 (with 10 only in soil)
Alkaline / earth alkali analyses	X	X	3
Metal / metalloids analyses	X	X	16
Mineral analyses	X	X	10
Halogen analyses	X	X	2
Organic analyses	X	X	25 (with 7 only in water)
Physical-chemical parameters	X	X	18 (with 9 in water and 6 in soil)
Groundwater level	X		
Pump test, hydrodynamic characteristics	X		
Piezometer equipment	X		

Table 1. Environmental databases

A major point to take into account is the non-redundancy of records. To avoid this, fields of the table 2 are indexed, and used for the common model of all databases except saturation level, pump test and piezometer equipment. For each database, the common model is completed by specific parameters or pollutants.

Indexed fields	Format	Signification
Nom	String [10]	Name of the object
Date	Date [mm/dd/yy]	Date
De_NGF	Real	Absolute top elevation
A_NGF	Real	Absolute bottom elevation
De	Real	Relative top elevation
A	Real	Relative bottom elevation
Code_Ech	String [30]	Name of the sample

Table 2. Common databases model

2.3 Analytical detection limits

Managing chemical compound analysis within a database generates several problems, among them coding the detection limits is an example. On analytical sheets the detection limits are commonly preceded by the "<" sign ; if using this symbolic, we work with strings and not either with numbers. As a chemical analysis references only positive values, it is possible to replace by convention the "<" by "-" sign. If this artifice allows the data query, a new problem arises for the computation or graphic data representation. Even if we lost information, the most interesting way is to take into account the absolute value of results. Nevertheless in all cases, there is a loss of information, since we can not define a specific value ; indeed for a same chemical compound, we find sometimes different detection limits according to the laboratory analysis method, the dilution, etc.

3 DATA REFERENCING

3.1 Generalities

It is known that the specificity of spatial information systems is the notion of geo-reference / location. The location data may be obtained from digitized maps or, to have a better reliability, from a DGPS (Differential Global Positioning System) survey. Anyway, it is necessary and essential to define a cartographic projection : in our study, we use metric Lambert I coordinates (north of France).

3.2 Spatial references reliability

3.2.1 Initial documents. It is very difficult to obtain informations on industrial sites history. Most data come from studies perform on already brown fields. Thus, there is often a lack of information, and when data are available the precision of referencing is usually low. In our cases, three kinds of maps were used:

- Sketch-plans;
- Some photocopies of CAO plans without information on reduction/enlargement scale factor.

Apart from these imprecisions, the coordinates are relative without true geographic reference; hence the necessity to use a geo-referencing tool to get calibration data.

- Manual transfer of objects between two geo-referenced maps has generated a series of imprecision and incertitude causing finally a translation with the real positioning.

For accurate geographic reference, such documents can not be used without treatment and require the use of a reliable tool for geographic referencing.

3.2.2 Field works. Except when realized by surveyors, positioning on site often involves the use of approximate methods as mentioned by Chaussier *et al.* (1981). These methods are not able to provided a resolution as good as metric or sub-metric that is required for large scale applications. Thereby, the positioning by DGPS has been used here.

3.3 DGPS application in the study

Quickly, DGPS performances are highlighted with regard to initial map reliability. The use of this system has been optimized with the realization of several referencing surveys. DGPS has been used to plot point objects (piezometers, cores, etc.), linear objects (limits of site, roads, etc.) and polygonal objects (factory buildings, decanting zones, etc.). . Moreover, even if they are brown fields, these sites know an activity especially with the temporary waste dumping of soils. These soil transfers generate some modifications of the site topography and occupation. It is the reason why another DGPS use has been to re-locate cores from an old investigation.

Later, a new survey at a larger scale has been realized to obtain reference marks on an aerial photograph.

3.4 Geo-referencing results

To compare the DGPS data precision and a geo-referenced and ortho-rectified aerial view (precision 1 pixel # 50 cm) several tests have been realized. Figure 1 shows the overlay of the two kind of data:

- Building location (A) is hidden exactly on the aerial photograph since only the building shadow is visible;
- A difference is highlighted between limits of site (B) and the road network;
- Roads (C) overlies road network. If the road legend is more wide, the network view should be totally hidden.
- The waste dump (D) location is also exactly defined.

This survey has also shown a translation of 29 m (cf. 3.2.1) between an initial plan and the DGPS survey. It is the reason why we question the initial map reliability and propose systematic geo-referencing by DGPS.

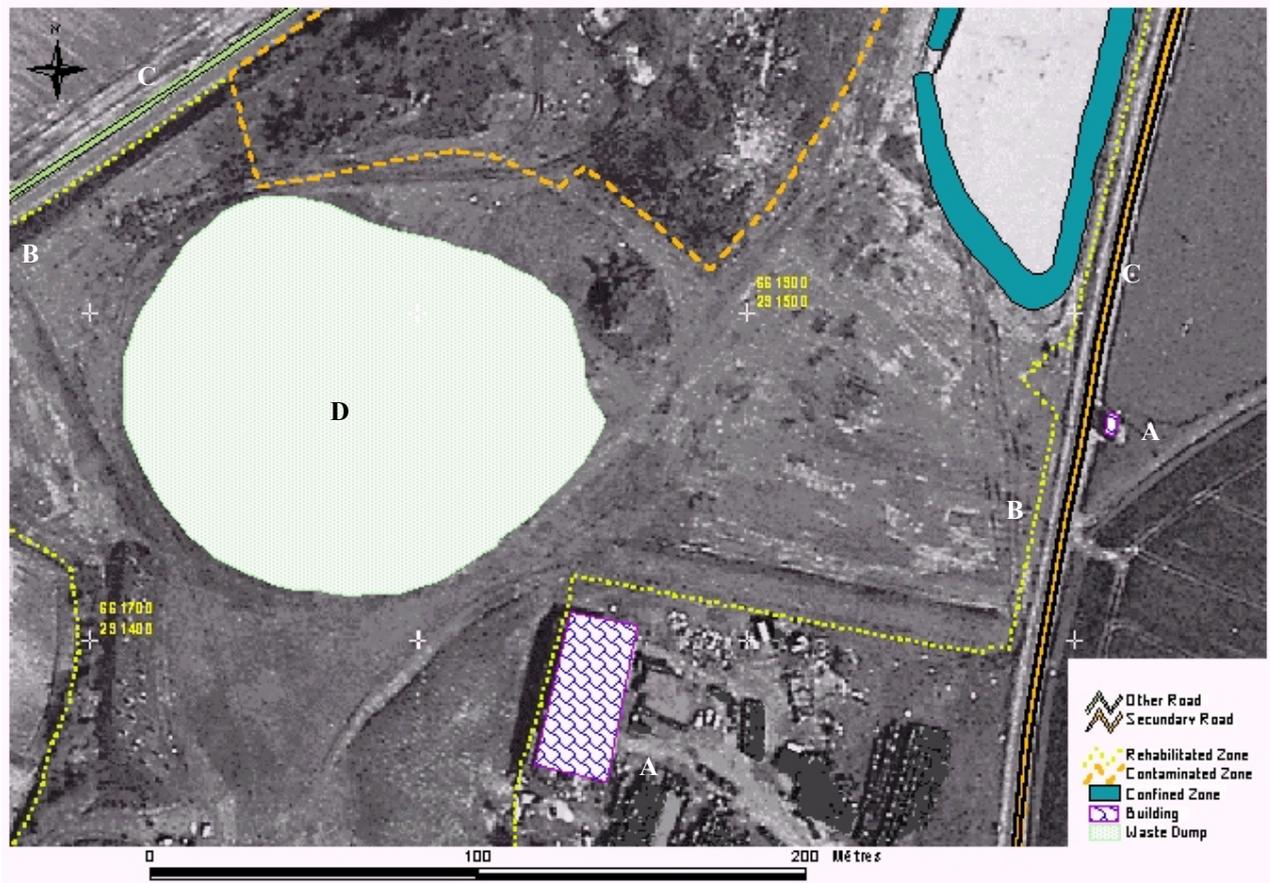


Figure 1. Overlay of a DGPS geo-referencing survey and a geo-referenced and ortho-rectified aerial view(ArcView©)

4 AERIAL PHOTOGRAPH USE AND INTEREST

The only accurate way to know the evolution of site occupation is to use time series of aerial photographs.

4.1 Photographic surveys availability and characteristics

Until now in France, the National Geographic Institute (IGN) is quite in position of monopoly for the aerial photographic survey and management. In general, the regular civil surveys started in the 50's, linked to the development of photography and aviation progress. For the Nord-Pas-de-Calais territory, the frequency of photographic surveys is 33% greater than for the rest of the French territory. This is related to the coal field presence and the intensive past industrial activity.

Except for the nineties, all the available views are in black & white / panchromatic. As the studied sites stopped their activities in the end of the seventies, only panchromatic photograph have thus been acquired. The photograph scales are between 1 / 30000th and 1 / 4000th.

4.2 Photograph from paper to numeric

Scanning is the more adapted method to integrate raster data in a GIS (Star & Dickinson, 1996). After scanning, an aerial photograph is composed by an XY array of pixels. The site dimensions are typically inscribed within a square of 2500 m², allowing a high scanning resolution (e.g. 1000 dpi for a 1 / 20000th photograph for 1 pixel correspond to 50 cm).

After scanning, an image has only a relative referencing and can not be overlaid with other geo-referenced objects. For Basquiat (1998), the principle of geo-referencing is to attribute geographic coordinates for each pixel of a raster picture. Thus, localization data are coded into a cartesian coordinates system (Star & Dickinson, 1996). In order to geo-reference scanned photographs, it is necessary to use an already geo-referenced document.

For this purpose, two kinds of document have been used. The first one is a numerical geo-referenced orthophoto (resolution = 50 cm x 50 cm) from a private authority. For the second site, no support was available so all the geo-referencing was performed using DGPS. The reference marks must be visible on the aerial view like surfaces (e.g. fields, stadium, etc.) but especially cross-lines (e.g. cross-roads, bridges, etc.).

The studied-sites have a small surface and a very flat topography consequently deformations are minimized except for waste dumps. Thereby, for geographical corrections, "register" and "rectify" ArcInfo© tools are used with good results.

4.3 Aerial photograph working

4.3.1 Evolution of factory building during time. Although the first aerial photographs appeared about thirty years after the beginning of the industrial activity, they allow to follow the evolution of factory building (Figure 2) and to locate the production places with the help of general site plan.

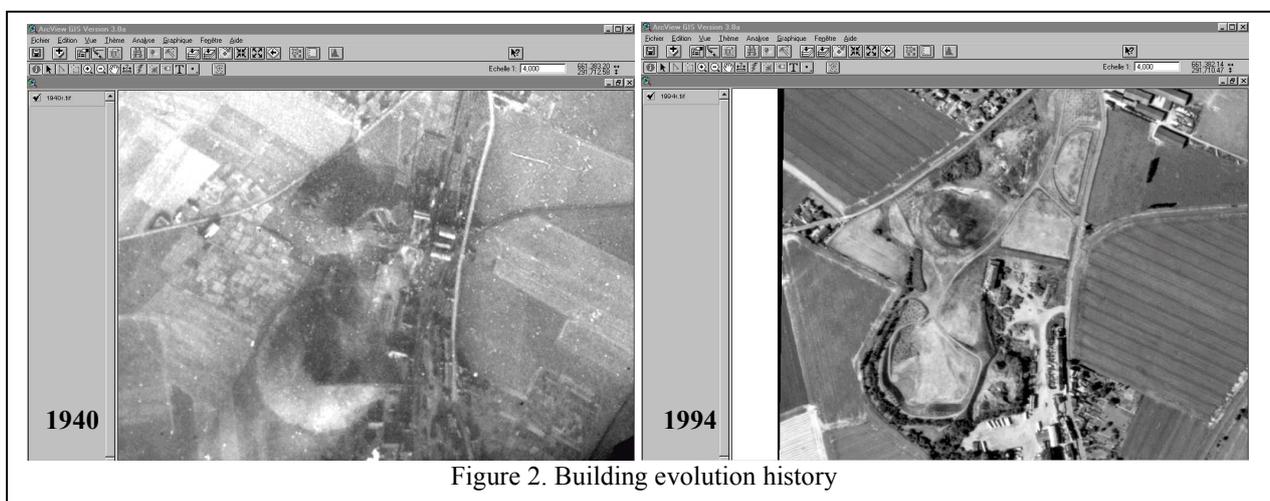


Figure 2. Building evolution history

From each photograph, the outline of buildings have been redrawn. At the end, the overlay of these limits shows that there differences in size and overlap. The main causes are the different conditions of aerial surveys (scale, angle of incidence, brightness, etc.). To minimize this problem, center coordinates (X, Y), perimeter (P), surface (S) are extracted for each building, then the mean X, Y, P, S are computed. From these data, the mean building implantation is created (Figure 3). To check accuracy, we note that the digitizing corresponds exactly with the existing substructure referenced by DGPS.

4.3.2 Evolution of decanting zones during time.

First, it is possible to follow the history of decanting zones during industrial activities. Second, aerial photographs are used to study the brown field itself. Indeed, even if the factory have stopped their activities since twenty years ago, site managers use them to excavate or to stock contaminated soils.

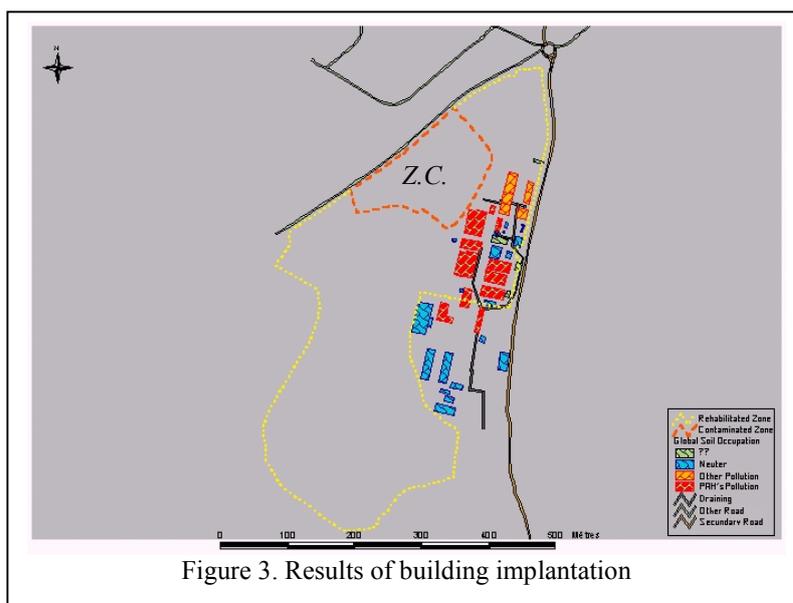


Figure 3. Results of building implantation

to the following aspects : singleness of the records, classification of the data in agreement with their environment (soil, water, air...), and structuring of data groups in a hierarchical system.

Independently of difficulties to get access to existing data, or to rebuild the history of sites, we underline many imprecision especially for geo-referencing the objects. Now, geographical data referencing is primordial for site investigations and necessary for the GIS conception. To solve this problem, the DGPS use is the best solution since the implementation is easy, not much expensive, and providing with a precision from metric to sub-metric. At last, the user can get locations but also to re-locate cores from older investigation.

One part of the site history reconstitution is possible from the use, the treatment and the interpretation of aerial photograph. From the oldest view to the most recent, we can study the global evolution from site to brown field state, and for example identify and locate the primary sources of pollution.

As a conclusion, these aspects have been included in a methodological guide for the building of a GIS dedicated to industrial site pollution diagnosis. This state-of-the-art guide should allow industrial site owners and administrators to better define investigation techniques and environmental management tools.

ACKNOWLEDGMENTS

We thank Philippe Branchu* for the reading of this paper.

REFERENCES

Basquiat J.P., 1998. Utiliser les SIG : le dossier des Systèmes d'Information Géographique. <http://www.admiroutes.asso.fr/action/THEME/nti/sig/utisig.htm>, 7 p.

Chaussier J.-B.; Morer J., 1981. Manuel du prospecteur minier. Manuels & Méthodes n°2. Orléans, France, Eds BRGM, 273 p.

IGN, <http://www.ign.fr/>

Star J.L., Dickinson H., 1996. Chapitre 5 : La saisie et l'entrée des données. In : Systèmes d'information géographique : Notes de cours, notions de base. Québec, Université de Laval, LATIG, Département de géographie. Goodchild M.F., Kemp K.K., Thériault M., Roche Y. (Eds), 1^{ère} édition, Vol. 1, pp 42-51.