

Tools and methods constructing 3D geological models in the urban environment. The Paris case.

Bernard Bourguine, Anne-Marie Leparmentier, Carole Lembezat, Pierre Thierry, Claire Luquet, Christian Robelin

► To cite this version:

Bernard Bourguine, Anne-Marie Leparmentier, Carole Lembezat, Pierre Thierry, Claire Luquet, et al.. Tools and methods constructing 3D geological models in the urban environment. The Paris case.. GEOSTATS 2008, Dec 2008, Santiago, Chile. pp.951-960. hal-00614827

HAL Id: hal-00614827

<https://hal-brgm.archives-ouvertes.fr/hal-00614827>

Submitted on 16 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TOOLS AND METHODS FOR CONSTRUCTING 3D GEOLOGICAL MODELS IN THE URBAN ENVIRONMENT. THE PARIS CASE.

BERNARD BOURGINE, ANNE-MARIE PRUNIER-LEPARMENTIER(*),
CAROLE LEMBEZAT, PIERRE THIERRY, CLAIRE LUQUET and
CHRISTIAN ROBELIN

BRGM, Department of Geology, Orléans, France; (*)IGC, Paris, France

ABSTRACT

Building a geological model for an urban environment is a relatively difficult task, because of: (1) the heterogeneous and commonly poor quality of the data; (2) the high spatial variability of the formations to be modelled, the relevant slice of ground for the planner being located at surficial formation level (fluvial alluvium, for example); (3) the necessity of handling a very large amount of data (several thousand drill holes); and (4) the fact that most of the data provide inequality constraints for the tops / bottoms of the layers to be modelled. Here we present a set of tools and methods designed and tested at BRGM over the past few years for (1) performing automatic consistency checks before and during modelling, and (2) facilitating the building of geological models that, in particular, take geological rules and inequality constraints into account.

INTRODUCTION

Many of the problems of today's major urban centres are directly or indirectly related to the geological, geotechnical and hydrogeological conditions beneath and around the city. It is thus desirable, before any work is undertaken, to be able to anticipate the geological conditions likely to be encountered at any given location, and this is possible with a geological model based on available information. Before building such a model, however, one must first control the quality and consistency of the available data, which is the subject of first part of this article.

We then present a method for modelling geological interfaces, which integrates the definition of depositional sequences / formation erosion, as well as the management of data inequalities. Finally, we discuss the limitations of the current methods and the developments needed to overcome them. The discussed tools and methods are illustrated with reference to Paris (France).

PARIS CASE STUDY

Geological context of Paris

Paris, which lies on the Seine River within a Tertiary and Mesozoic sedimentary basin, is underlain by backfill, three surficial Quaternary formations, sixteen Tertiary formations and the top of the Cretaceous substratum (Figure 1). A 3D geological model of these 21 formations was built in 2007 in order to evaluate the risk of ground movement and collapse due to gypsum dissolution (Thierry *et al.* 2007). Here we illustrate the results of this model on the "Brackish Marl and Limestone" formation (CAIL), which is one of two formations with significant thicknesses of gypsum (up to 15 m in areas where the formation exceeds 30 m).

Surficial deposits & Oligocene		Late Eocene (Priabonian)		Late + Middle Eocene (Bartonian + Lutetian)		Lower Eocene (Ypresian) & Paleocene	
TVRE	backfill	MAVE	Priabonian	OUEN	Bartonian	SYPR	Ypresian
EBOL	Quaternary deposits	MSGY		BCHA	Lutetian	FGLA	
ALUM		MGYP		CAIL		AUTE	
ALUA		MIGY		CGRO	APLA		
FONT	Oligocene	SVER			DAMO	Paleocene	
BRIE					CRAI	Cretaceous	

Figure 1: Reference lithostratigraphic succession

Lithostratigraphic codes: TVRE = backfill + topsoil; EBOL = scree, tableland loam and loess; ALUM = recent alluvium; ALUA = old alluvium; FONT = Fontainebleau Sand; BRIE = Brie Limestone and Oyster Marls; MAVE = Green marl; MSGY = Supra-gypseous marl; MGYP = gypsiferous formation; MIGY = Infra-gypseous marl; SVER = Greensand; OUEN = Saint-Ouen Limestone; BCHA = Beauchamps Sands; CAIL = Brackish marl and limestone; CGRO = Lutetian limestone; SYPR= Soissonais Sands; FGLA = False marl; AUTE = Auteuil Sands; APLA = Plastic clay; DAMO = Meudon Limestone and Marl; CRAI = Cretaceous chalk.

Arrows indicate the presence of significant gypsum thicknesses.

Data gathering

A preliminary evaluation of the number of drill holes in Paris recorded in databases and paper documents (originating mainly from IGC - City of Paris and BRGM - French Geological Survey) gave more than 9000. Having neither the time nor the budget to collect, input, homogenize, reinterpret, codify and process this large amount of data, it was decided to restrict the study to 3280 drill holes within the Paris limits, giving a maximum spacing of 200 m. The spacing was calculated beforehand on a subset of data according to the thickness variogram of the CAIL formation. An additional 700 drill holes around the city limits were

then added to control the boundary effects, bringing to 3980 the total number of drill holes to be processed.

The drill-hole log descriptions were homogenized and a lithostratigraphic code (Figure 1) assigned to each of the approximately 50,000 downhole runs. The initial full description of each run was kept for interpretation and checking purposes. The information was recorded in an Access© database, along with the drill-hole's identifier, X,Y,Z coordinates, and downhole depth of the run ends.

The IGC Geological Atlas at 1:5000 scale and the BRGM 1:25,000 scale geological map for areas not covered by the IGC maps were used for digitizing the boundaries of the modelled lithostratigraphic units (6068 points after sampling at 50 metre intervals). The boundaries provide passage points to constrain the interpolated surfaces and can also be used to limit the interpolations to zones where the different units are effectively present.

Finally, a 20 m grid digital elevation model (DEM) with an altitude precision of ~1 m was calculated from the levelling data provided by the City of Paris. This was used for calculating point altitudes of the geological boundaries and provided the model's upper limit. Also, by comparing the drill-hole collar altitudes with the DEM, it was possible to identify the drill holes whose given X,Y,Z coordinates were erroneous.

DATA CONTROL

Data control is a vital step. It is essential, before data are used for modelling (or other) purposes, to identify as systematically and predictably as possible all the errors that could have affected them, especially considering the large volume, the heterogeneous quality, the different phases of database input, and the successive phases of interpretation. Because reviewing all the data "by hand" is not envisageable due to the quantity involved, one must resort to automatic tests for identifying potential anomalies. We have incorporated such functions in a modelling tool (MultiLayer software, developed by BRGM; Bourguin 2007) that also automatically integrates the geological modelling context (deposition / erosion sequence) and features such as drill holes and geological maps.

Consistency check between drill-hole and geological-map data

In this consistency check the geological formations at outcrop, as indicated by the geological map, are compared with the topmost formation intersected by the drill holes. Figure 2a gives an example of the graphic output produced by this check. The software marks as red squares the drill holes in which the topmost intersected formation is younger than the "outcropping" formation shown on the geological map, and as blue circles the drill holes in which the reverse situation is observed; crosses represent the drill holes consistent with the geological map. The origin of these errors could be a wrong interpretation of the drill hole, an incorrectly positioned drill hole, or an imprecise geological boundary on the

map. In the Paris case, the erroneous drill holes were corrected where possible or eliminated. The geological map was also redrawn locally to be consistent with correct drill-hole data.

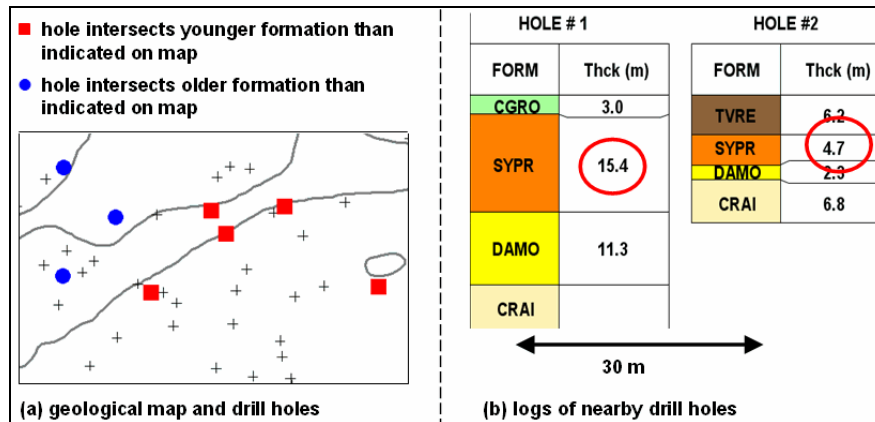


Figure 2: Controlling the data consistency

Consistency check between nearby drill holes or nearby data

Tests have been developed for comparing the logs of nearby drill holes. The input parameters for the tests are (1) the proximity radius within which the drill-hole logs are to be compared, and (2) the acceptable difference between logs in terms of maximum variation in formation thickness, maximum variation in the top or bottom elevations, and minimum similarity of logs. The tool provides a list of drill-hole pairs that do not satisfy the input criteria, with the description of the criterion that has not been met. Drill holes can also be compared to passage points given by the geological map. The example in Figure 2b shows two drill holes 30 m apart in which the difference in the SYPR formation thickness is 11 m. In such cases the drill holes have to be checked so as to determine whether this difference is due to an error of interpretation or input, or even to incorrect coordinates. This verification phase is "manual" and very often requires a return to the base data, i.e. to the original or scanned documents.

A geostatistical cross validation is another way of performing this test. It is less empirical because outliers in the geostatistical case are determined by analysing the normalized kriging error, itself based on the variogram.

The two tests are complementary, each with its own advantages. The empirical test that we used here is performed at an early stage of the study and only requires information that the geologist can easily provide. It processes all the geological formations in one step and at the same time analyses the elevations and thicknesses. It is also able to take inequalities into account (see later for the definition of inequalities). It is a rapid method for checking inconsistencies at an

early stage, and does not exclude a later geostatistical cross validation once the data have been “cleaned” and a variogram calculated.

BUILDING THE GEOLOGICAL MODEL

Here we only address the case where the geometry can be modelled in so-called 2.5D, i.e. when the altitude values z of geo-objects can be represented as a continuous function of the geographical coordinates $z = f(x,y)$. (Co)-kriging the tops of the geological layers probably remains the most efficient technology in this case, and it has the advantage of providing an estimation of the interpolation error, which is essential for risk analysis (Chilès & Blanchin 1995). The geologist, even in this "simple" case, is nevertheless faced with many problems when building a 3D model, and in particular how to manage the interpolated surfaces when they cross one another, and how to take into account the inequalities. The solution we propose consists in (1) modelling the stratigraphic surfaces and intersecting them correctly, and (2) providing tools for checking and dealing with inequality constraints.

Modelling and intersecting stratigraphic surfaces

Figure 3 shows three interpolated surfaces: the top of A (*TA*, from passage points *A1* and *A2*), the erosion surface *ES* (from passage points *E1*, *E2* and *E3*), and the top of C (*TC*, from passage points *C1* and *C2*). These surfaces are interpolated over a larger area than their present coverage in order to allow intersection with the other surfaces. For example, the removed branch of *TA* (on the right) corresponds to the top of A where A has been eroded by the erosion surface *ES*. The left part of *TC* has to be removed because C could never have been deposited when B was already in place. Finally, the upper part of *ES* does not exist because of the topographic erosion (*TOPO*).

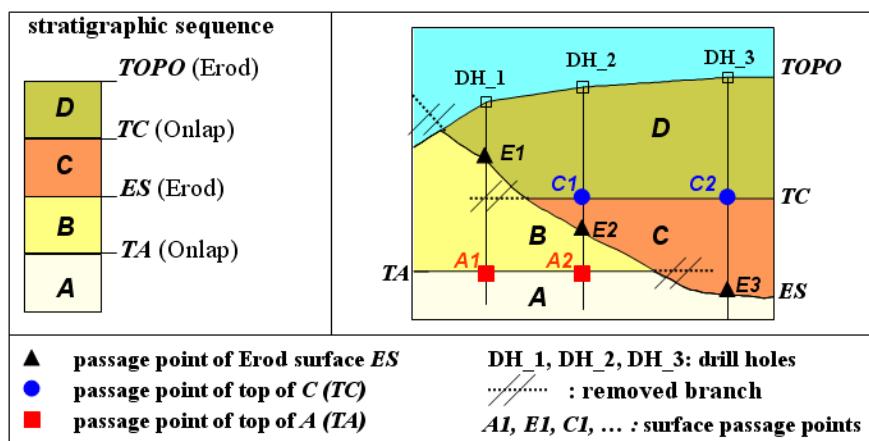


Figure 3: Modelling and intersecting stratigraphic surfaces

The simple example shown in Figure 3 can be generalized by the following procedure: (1) defining the stratigraphic sequence to be modelled, as given by the list of formations that have been successively emplaced and the type of relationship between two successive formations, i.e. "Onlap" when a formation is deposited following the previous formation without intermediate erosion, or "Erod" when there is an intermediate erosion phase; (2) constructing the erosion surfaces through interpolation from the passage points of these surfaces; (3) combining/intersecting the erosion surfaces, giving priority to the most recent erosion which obviously erodes everything that precedes it; (4) constructing "Onlap" surfaces through interpolation from the passage points of these surfaces; (5) combining/intersecting "Onlap" surfaces by first eliminating branches that have been eroded, and then introducing the successive "Onlap" surfaces in the model, from the bottom to the top, while eliminating any parts that could not have been deposited because an older formation was already in place. This mechanism was, to our knowledge, first implemented in 1990 in the EarthVision software through the "Streamline" model-building process (apparently never published, but mentioned in the website of Dynamics Graphics 2005). It was then, in particular, taken up by BRGM in 1999, incorporated in Geomodeller (Courrioux *et al.* 2003, Aug 2004) and MultiLayer software (Bourgine 2007) and applied to various cases (e.g. Thierry *et al.* 2000).

Defining passage points for the tops / bottoms of geological formations

In order to enable the method to be implemented, the passage points of the "Onlap" and "Erod" surfaces need to be correctly determined, which is possible using the following algorithm as a basis:

- let FZ/FA be a contact between two formations FZ and FA , observed in a drill hole or in outcrop, FZ being the younger,
- let N_Erod be the number of erosion surfaces in the stratigraphic sequence between FA and FZ , and let $Last_Erod$ be the last of them,
- if $N_Erod = 0$ (i.e. no erosion between the deposition of FA and that of FZ), the contact FZ/FA is a passage point of the top of FA , otherwise the contact FZ/FA is a passage point of the last erosion surface $Last_Erod$.

In Figure 3, in drill hole DH_2, the B/A contact is interpreted as a passage point of the top of A (i.e. no erosion between the deposition of A and that of B in the stratigraphic sequence); similarly the D/C contact corresponds to a passage point of the top of C . The C/B contact (point $E2$), however, is interpreted as a passage point of the erosion surface ES which lies between C and B in the stratigraphic sequence. In standard modelling, the C/B contact could have been interpreted as a passage point of the top of B , but the "true" top of B no longer exists since B has been eroded either by the topography or by the erosion surface ES . In our case we do not explicitly model formation B ; rather B is defined implicitly by the space remaining between the top of A and the erosion surface.

Determining inequality constraints

The drill holes and the geological map not only provide passage points for the intersected "Onlap" or "Erod" surfaces, they also impose inequality constraints on the position of other deposition or erosion surfaces. Figure 4 shows an example of the constraints generated at the top and bottom of a drill hole starting in an outcropping formation *FZ* and ending in a formation *FA*. Any *FZ/FA* interface in a drill hole will generate the same type of constraints with the *Z1* and *Z2* of Figure 4 being replaced by the elevation of the *FZ/FA* contact.

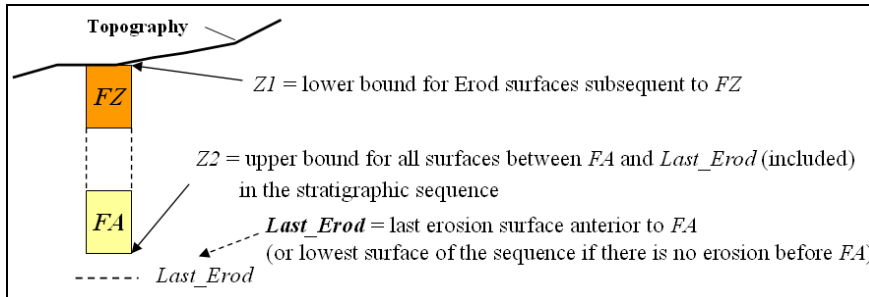


Figure 4: Some inequality constraints for a drill hole starting in formation *FZ* and ending in *FA*

The geological map provides a virtually infinite number of inequalities. Firstly, any point of the boundary between two formations not only defines a passage point for the formation top or for an erosion surface, but also defines inequality constraints for other formations. Secondly, any point lying inside a polygon of a formation *FZ* provides inequalities for the top of *FZ*, bottom of *FZ*, etc.

When interpreting available data using the previous algorithms, the data are divided into (1) exact data, corresponding to the passage points of the surfaces to be modelled, and (2) inequality data, corresponding to the upper and lower bounds of these surfaces. Table 1 summarizes the number of passage points and inequality data for the top of the "CAIL" formation in the Paris case. From a total of 10,004 points (drill holes + geological map boundaries), we obtained exact data for 2,056 points and inequality data (lower bound + upper bound) for 4,460 points, mainly from the geological map, giving 6,516 data to be processed for the top of "CAIL". Considering that this number of data represents but one out of 21 formations to be processed, it is easy to understand the need for automatic inconsistency-finding methods and adapted processing.

Table 1: Available number of defined data points for the top of the CAIL formation.

Type of data	Lower bound	Exact value	Upper bound	Total
Drill hole (3,936 pts)	797	1,100	749	2,646
Geological map (6,068 pts)	955	956	1,959	3,870
Total (10,004 pts)	1,752	2,056	2,708	6,516

Taking inequality constraints into account

Taking inequality constraints into account is a very difficult problem with configurations such as we are treating here, i.e. (1) a large number of formations and of exact and inequality data, and (2) geological formations that are not everywhere present, some having been locally eroded.

Co-kriging the top and bottom of a same formation gives a solution that does not respect all the inequalities due to drill-holes that intersect a formation without crossing it. Moreover, co-kriging does not always provide a satisfactory solution where formations wedge out and locally disappear.

Kriging under inequality constraints would appear to be better adapted. A set of constraint points meeting the inequality constraints can be generated automatically with a Gibbs sampler and added to the initial data (Freulon & De Fouquet 1993, Aug 2004). Unfortunately, this method is not always applicable and this for several reasons. First, it assumes that the data follow a Gaussian distribution, so that when this is not the case the data have to first be transformed by Gaussian anamorphosis techniques. Second, it only works in the stationary case, which is rare for formation tops and bottoms and so the drift must first be removed. Finally, the method only strictly works in unique neighbourhoods, which limits its practical use to configurations of few hundreds to one thousand data items at the most, which is rarely the case in urban environments. Large data sets require a long calculation time and the iterative process for constructing the model, verifying the results and possible data correction becomes very onerous and, in practice, impossible.

In view of the potential problems with the above methods, we adopted the following procedure. First, passage points ("hard data") are used to interpolate the different surfaces by kriging or co-kriging. The obtained surfaces are then compared to the inequality constraints ("soft data") taking the kriging error into account. The detected inconsistencies are ranked by order of significance. The software automatically plots plan and section graphics centred on the strongest anomalies (Figure 5). As these graphics are interactive, with a link to the data, the user can rapidly check if the anomaly is due to an error in the data and so, if necessary, correct them. This procedure enables one to detect errors that could not have been detected in the preceding phases (standard cross validation does not take inequality constraints into account). When there is no error in the data one can introduce one or several constraint points so as to force the model to meet the inequality data. These constraint points are introduced as supplementary passage points and the surfaces are then reinterpolated. With successive iterations one can obtain a geological model consistent with all the data. The geological model is then controlled by drawing isohypse or isopach maps of the different formations and checking, on cross-sections, that the geological features have been properly reconstructed.

In the Paris case the different control and modelling steps led to 1.5% of the drill holes being eliminated and 15% of the drill holes being modified or

reinterpreted. We had to introduce 1464 constraint points, including 149 for the "CAIL" formation.

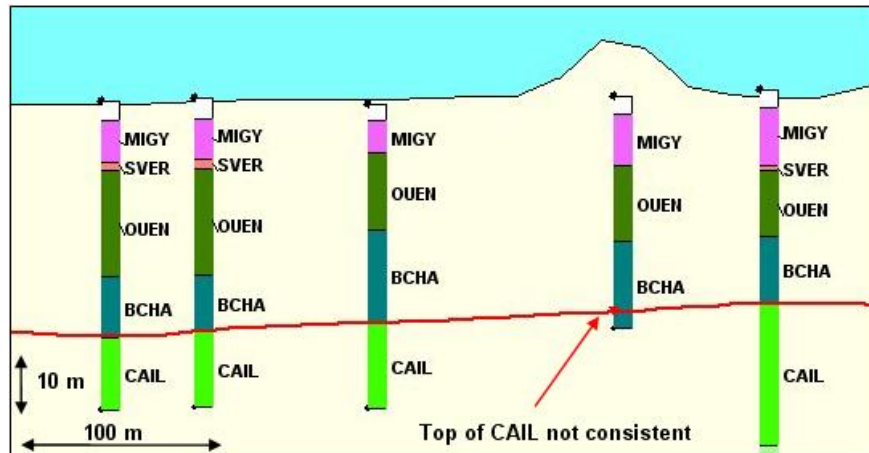


Figure 5: Automatic graphic showing a non-respected inequality constraint at the end of a drill hole

LIMITATIONS AND POSSIBLE IMPROVEMENTS

When kriging under inequality constraints is not possible, the constraint points in the above procedure are generally introduced manually. This can pose major problems for updating the model when new data are introduced, because the new data can make the introduced constraint points unnecessary, or even conflict with these points. One must therefore eliminate the corresponding constraint points and repeat the interpolation of the surfaces. So that this backtracking does not become too cumbersome, one has to devise algorithms that automatically search for the constraint points to be eliminated, then regenerate the model. A neater solution would be an automatic generation of the constraint points. This would require the kriging under inequality constraints to be adapted so that it could function (1) in moving neighbourhood (or in unique neighbourhood with larger amounts of data) and (2) with non-stationary cases. Methods like those proposed by Furrer *et al.* (2006) and Gribov & Krivoruchko (2004) could possibly be adapted. We appeal to the geostatistical community to work in this direction.

The way the geological map is considered can also be improved. At present, it is only the boundaries of the geological polygons that are correctly taken into account. Points inside a polygon are inequalities that can be introduced as a grid that discretizes the polygons. This, however, leads to a large number of points, most of which are useless. It is better only to retain the most pertinent points.

CONCLUSIONS

The methods used in the Paris project have demonstrated that it is possible to automatically and rapidly identify different sources of errors in the data, resulting in considerable time savings given the considerable volume of information that needs to be handled. By using geostatistical interpolation methods governed by the consideration of a stratigraphic sequence and the surface construction rules deduced from this, it is possible to construct a geological model of the principal formations quite quickly. Thanks to the rapid automatic detection of non-respected inequality constraints, it is possible to introduce a set of constraint points in order to obtain a model compatible with all the data, and in particular with the inequality data, which is the most abundant and the most difficult to take into consideration.

However, the limitations of kriging under inequality constraints resulting from the large amount of data in the urban environment, makes this kriging method difficult to use. Consequently, building the final model requires that most of the constraint points are manually introduced. This is an obstacle for automatic model generation and easy update, and calls for new developments in this field.

REFERENCES

- Aug, C. (2004). *Modélisation géologique 3D et caractérisation des incertitudes par la méthode des champs de potentiel*. Thesis, Paris School of Mines.
- Bourgine, B., (2007). *Modélisation géologique 3D à l'aide du programme Multilayer*. BRGM/RP-53111-FR.
- Chilès, J.P. and Blanchin, R. (1995). *Contribution of geostatistics to the control of the geological risk in civil-engineering projects: The example of the Channel Tunnel*. In Applications of Statistics and Probability - Civil Engineering Reliability and Risk Analysis, M. Lemaire, J.L. Favre, & A. Mébarki (Eds.), A.A. Balkema, Rotterdam, Netherlands, 2, 1213-1219.
- Courrioux, G., Bourgine, B., Bertrand, E., Bour, M., Guillen, A. and Mouroux, P. (2003). *Modèle complet 3D-4G pour la caractérisation de l'aléa sismique à Nice*. In SIRNAT-JRPN, Orléans 2003 - <http://www.brgm.fr/Fichiers/SIRNAT/sirnat29.pdf>.
- Dynamic Graphics (2005). *Web site for Earth Vision software*: <http://www.dgi.com/earthvision/index.shtml>
- Freulon, X. and De Fouquet, C. (1993). *Conditionning a gaussian model with inequalities*. In Geostatistics Tróia '92. A. Soares, ed. Kluwer, Dordrecht, Netherlands, 1, 201-212.
- Furrer, R., Genton, M.G. and Nychka, D. (2006). *Covariance tapering for interpolation of large spatial datasets*. Journal of Computational and Graphical Statistics, 15(3), 502-523.
- Gribov, A. and Krivoruchko K. (2004). *Geostatistical mapping with continuous moving neighbourhood*. Mathematical Geology, 36(2), 267.
- Thierry, P., Bourgine, B., Vairon, J. and Vinet, P. (2000). *Methods and tools for constructing a 3D geological model in an urban environment: an example from the west of Greater Lyons*. In Decision making in Urban and Civil Engineering - 2nd - Lyon - France - 20-22/11/2000, 1, 525-537.
- Thierry, P., Leparmentier, A.M., Lembezat, C., Nowak, C., Plat, C., Vanoudheusden, E., Vernoux, J.F., Vincent, M. (2007) *Cartographie de la susceptibilité à l'apparition de désordres liés à différents phénomènes géologiques sur la Ville de Paris – Rapport final*. BRGM Report RP-55895-FR.