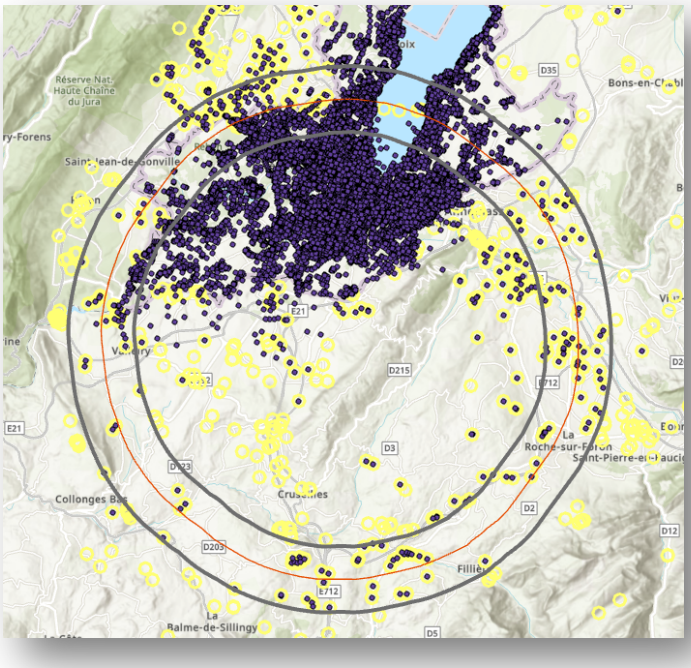
UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Department of Earth Sciences

Geo-Energy Group
Department of Earth Sciences
University of Geneva

ARHICTECTURE, CONTENT AND MANAGEMENT OF THE SUBSURFACE DATA BASE FOR THE FCC PROJECT



Report n.	GEG2022001
Date	23.03.2022
Author (s)	Emna Meftha
Approved by	Prof. Andrea Moscariello

UNIVERSITÉ DE GENÈVE
swissuniversity.ch

This page is intentionally left blank

ARHICTECTURE, CONTENT AND MANAGEMENT OF THE SUBSURFACE DATA BASE FOR THE FCC PROJECT

Rationale of borehole data collection:

This report concerns the collection of multiple subsurface boreholes data from neighboring France and Switzerland. The aim of this work is to integrate well data from several data bases such as SITG, BRGM, GADZ into a unique UNIGE Data base, for the purpose of revising the old geological model of GEOMOL project and to ameliorate the geological model of the FCC. Ultimately, the data collected by UNIGE will be transferred to the SOLSTIS platform operated by the GESDEC (State of Geneva) and become part of the larger and public data base available for the Geneva Basin and neighboring areas.
The collection of new data is needed in the neighboring France to improve the 3D geological model. For this task, therefore each single database were examined by carrying out the main following tasks :i) check the quality of existing information to identify data completeness and consistency, ii) check and add – where unavailable - the coordinate system, iii) fill-in information about the geological units (Top Molasse, Top Cretaceous); iv) provide higher level of details on lithological zone where possible.
The work of Data Collection and Data Base build-up was divided into five phases

- BRGM open data base (InfoTerre)
- Verification and integration of GADZ data base into UNIGE data base
- Verification and integration of M. Haas PhD's data base into UNIGE data base
- SITG, integration of GESDEC data base
- Adaptation to SOLSTISS data base architecture and data transfer from UNIGE to SOLSTISS

The SITG / SOLSTISS Data Base

The data processed by GESDEC are stored in different databases, geodatabases and file tree systems.
SOLSTISS (SOL-System Territorial Information of the Subsurface) is the new information system (IS) of the SITG. The application allows the systematic integration, conservation, processing and administration of all the results of prospecting, exploration and exploitation work carried out in the subsurface.
These data are the basis for the production of maps and geological models that facilitate the development of urban projects and the development of resources as well as the administrative procedures associated with these projects.
SOLSTISS allows a more dynamic update with an access to an updated state of knowledge on the cantonal subsurface according to the projects carried out on the Geneva territory. It constitutes the common

reference system on which the canton relies to improve knowledge of the subsurface and ensure sustainable management of the resources it contains.

The origin of data into this geodatabase are from different types and purposes. GESDEC manage geological, hydrogeological, geophysical, geothermal, polluted sites, gravel pits, soil and waste data. The basic geodata therefore consists mainly of borehole and piezometer data used for the geological model.

The main functions of the IS SOLSTISS can be summarized as follows, according to the figure 1:

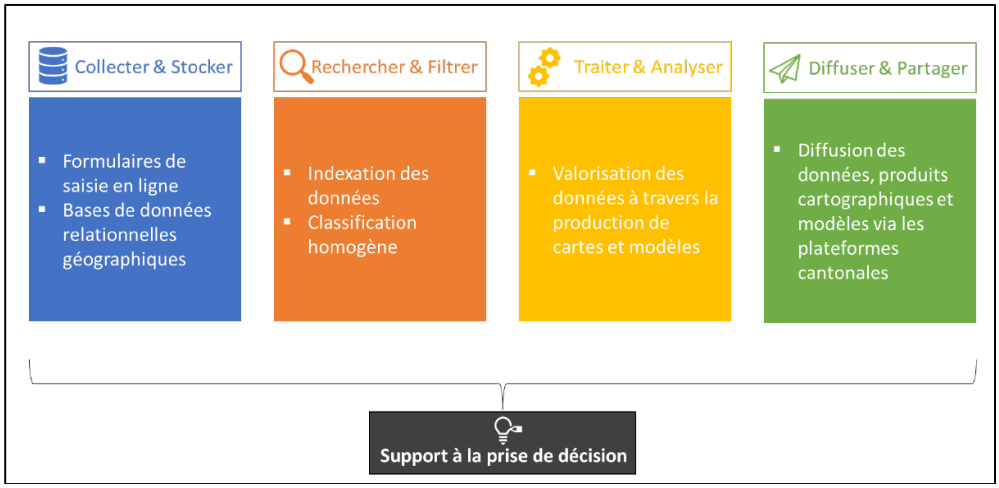


Figure 1: The main functions of the IS SOLSTISS (Source: SITG)

Geological project data are captured, centralized and stored in a relational database. The development of a data model allows to define the structure of the relational database through a modeling process (Figure 2).

It also gathers all the data on geological objects: equipped or not, seismic and electric lines, outcrops and geothermal installations of open and closed system type…

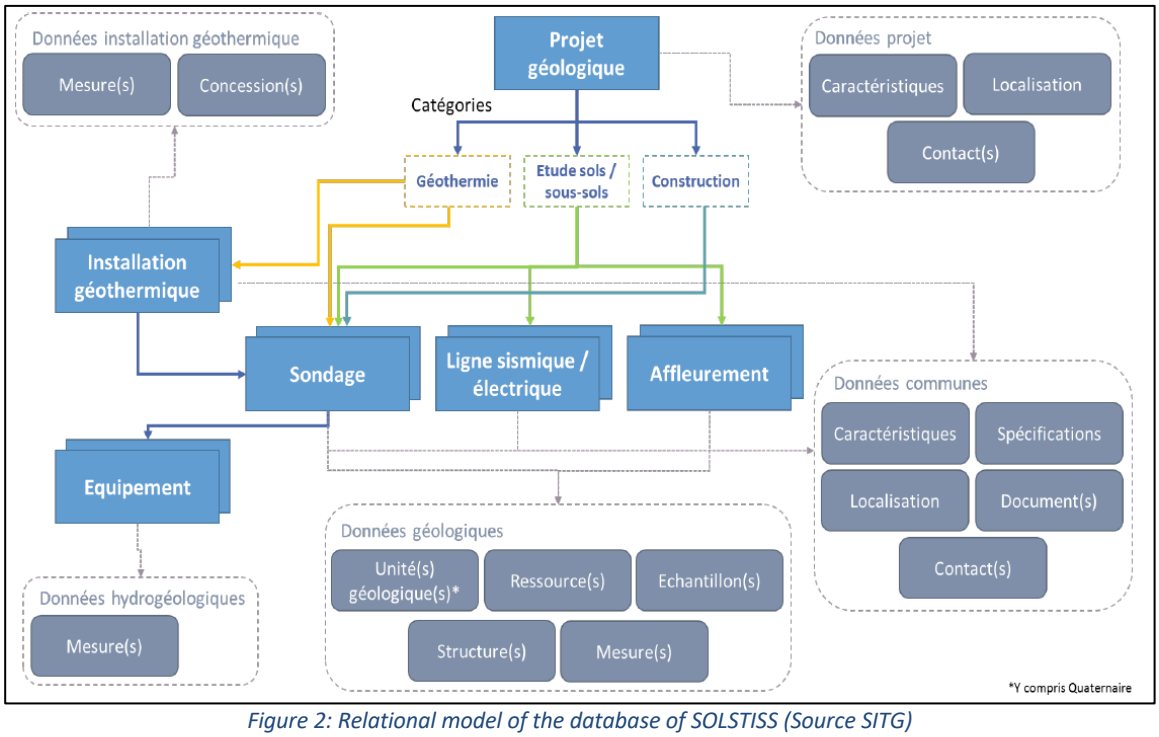


Figure 2: Relational model of the database of SOLSTISS (Source SITG)

In the figure 3, a series of objects grouped by attribute tables allow the item in question to be characterized. The links between the tables are relations, allowing the connection of information between them during specific queries. The exploitation of this data allows the creation of different cartographic products.

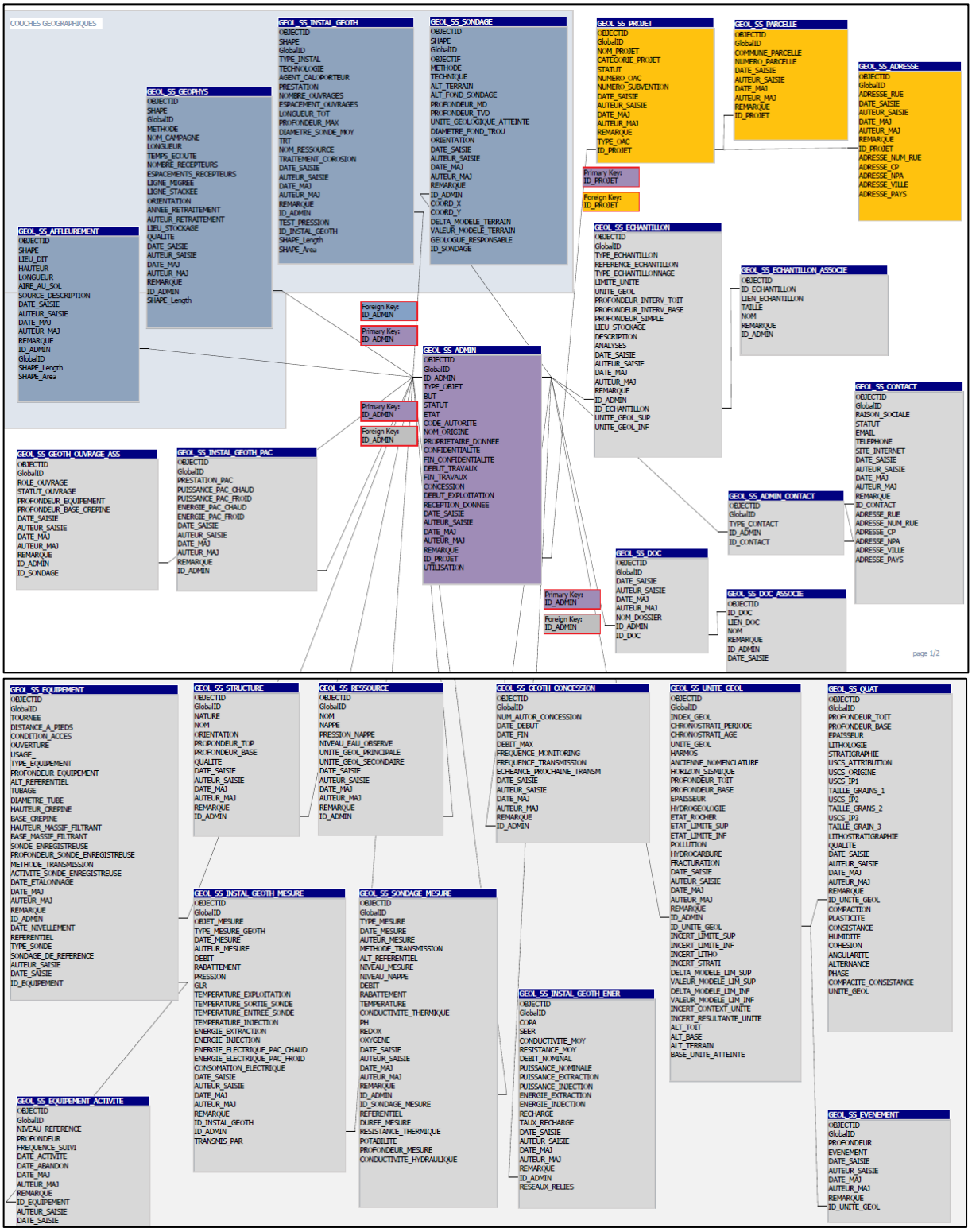


Figure 3: Global model of the database SOLSTISS

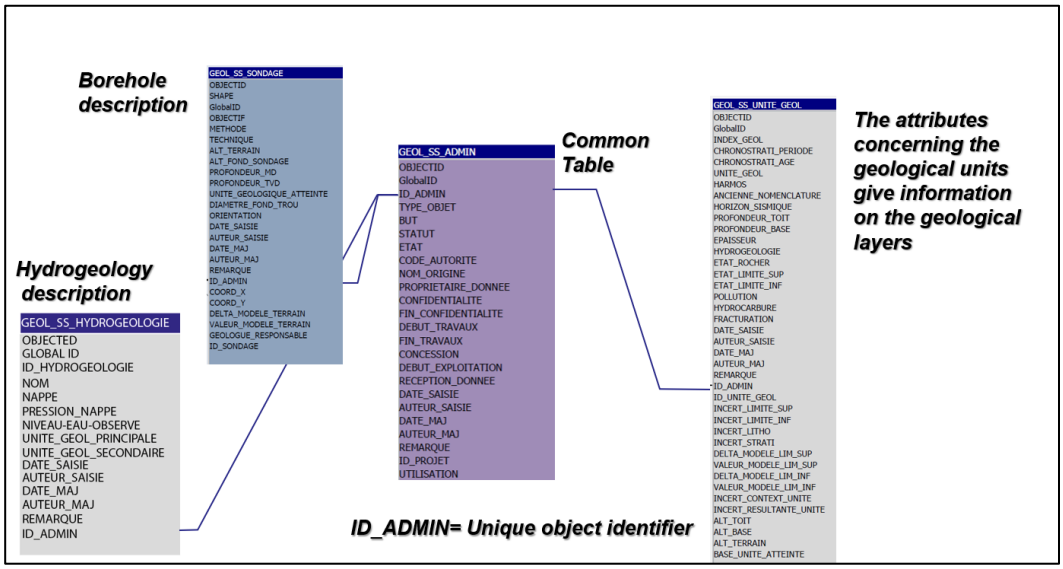


Figure 4: Table will be filled by the UNIGE

For the FCC project, the data filling of these attribute tables is the UNIGE team work. The attribute table required for the 3D model are presented in the Figure 4.
"Geol_ss_unite_geol" Table describes the different geological layers identified through the object. The information is based on the documentation provided.
This table is related directly to a common Table "geol_ss_admin", which is related to all the attribute table and it belongs to the SOLSTISS data base.
For the task of the UNIGE is to focus on these three-table related to the common table: GEOL_SS_UNITE_GEOL, GEOL_SS_SONDAGE and GEOL_SS_HYDROGEOLOGY.

The object catalog lists all the attributes developed in the SOLSTISS DB. For each table, the list of attributes is described with the name of the attribute, the type of data (ID, text, number, drop-down menu). This is a document created in Excel that accompanies the data model schema, and provides a quick description of the attributes that make up the data model.

The integration of the borehole collected from the BRGM data base is possible by:

- Exchanging information, and looking for correspondence between the attributes
- Adding new attribute (coordinate system, water table elevation, ….) by calculation, geolocalisation or filling from reports.

The SITG present a very large number of boreholes: 18'000 , 8969 public wells and 347 are deeper than 50m (information extracted from GOL_SONDAGE) see Figure 5. For the hydrogeological information, 464 piezometers are used to generate the hydrogeological map, see Figure 6.

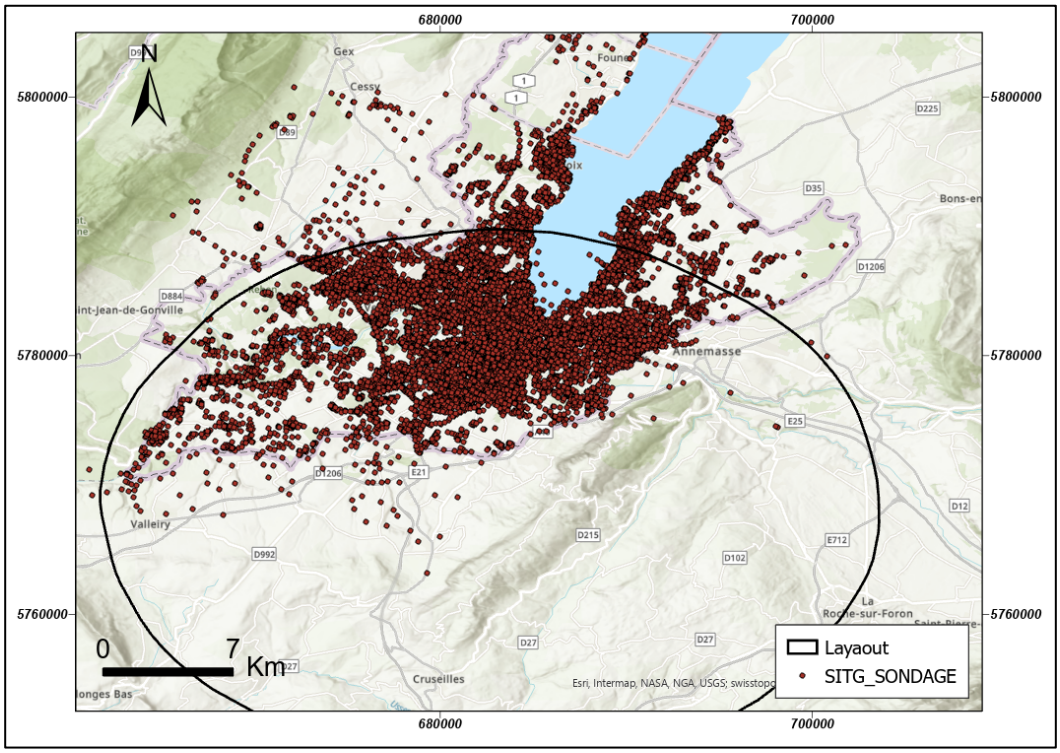


Figure 5: boreholes from GOL_SONDAGE shp file (Source SITG)

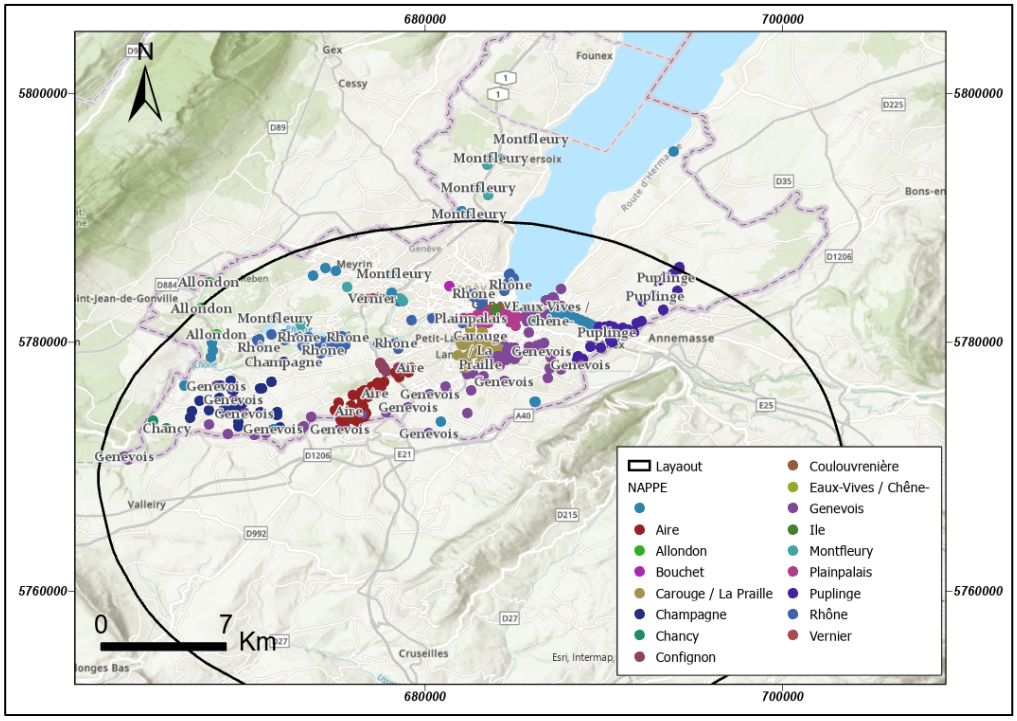


Figure 6: Piezometer database in the Geneva canton and the name of the associated water table, data modified from SITG (Source SITG)

Data and cartographic maps are available to the public through the SITG platform. As indicated, some data are available in Open Data, such as the isolines of the geological layers of the Riss and Molasse, the water protection areas and the cadaster of piezometers (Figures 7).

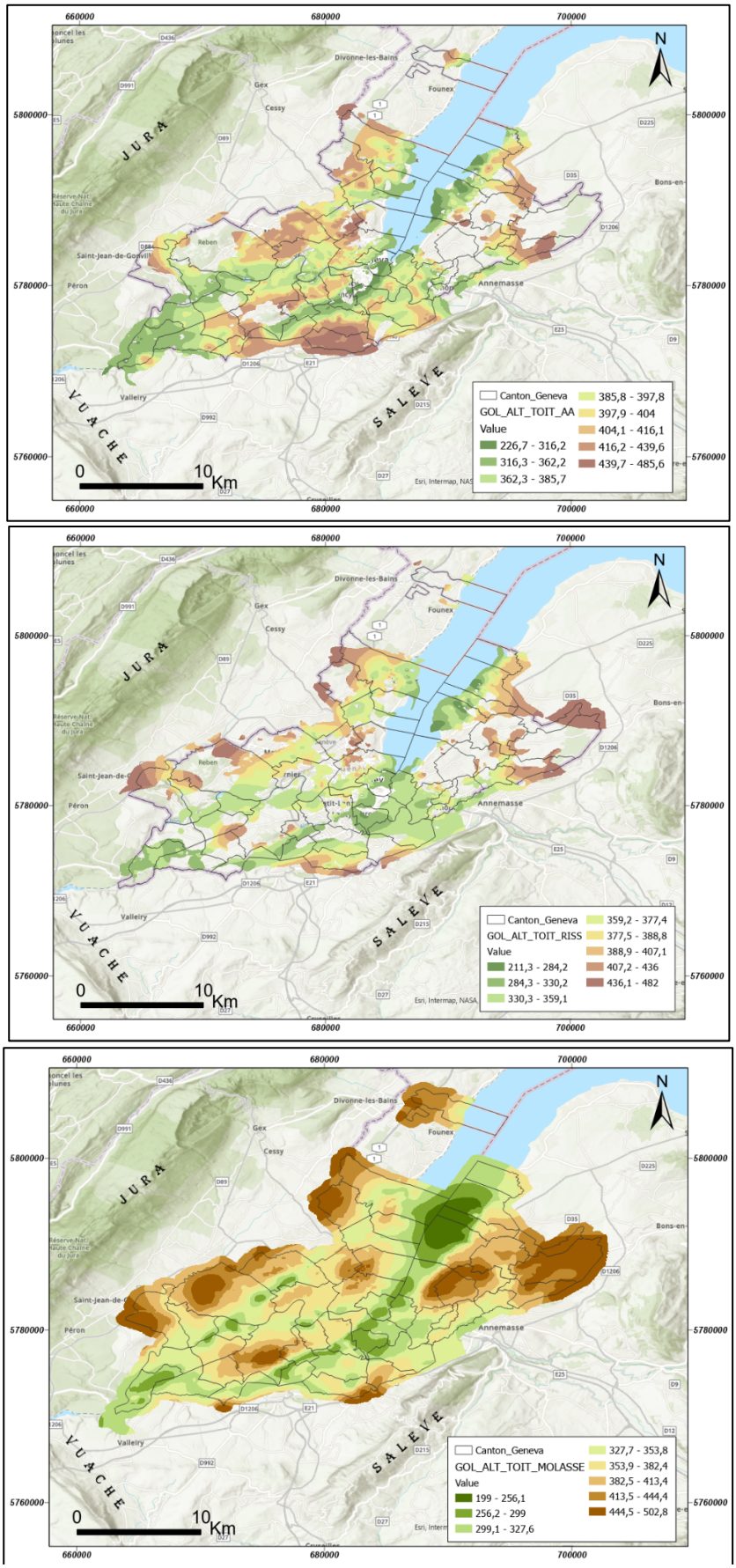


Figure 7: Elevation maps: respectively Top Alluvian Ancienne, Top Riss and Top Molasse (Source SITG)

The GADZ data

The data received from GADZ are in the format of Excel file. The files contain the identifier of the wells and the references, the X and Y coordinates within Lambert 93 projection system and the elevation of the Molasse in each well. In Some boreholes, the Molasse does not exist and /or it's deeper than the hole. All the wells received from GADZ are collected form the BRGM database (Figure 8).
The task of the UNIGE is to review this data, well by well and add the coordinate of each well within the swiss coordinate system and finally check the information about the Molasse with the support of the documents and reports found in info Terre open access.

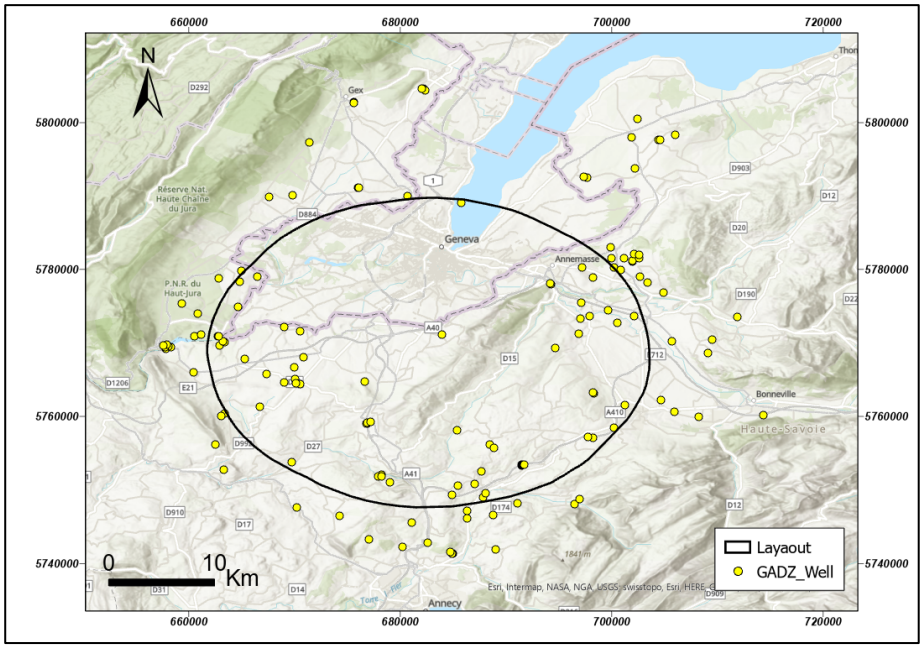


Figure 8: Boreholes from GADZ, where the molasse is recorded

Boreholes with No Molasse recorded	146
Boreholes with Molasse recorded	146
TOTAL	292
Wells used by UNIGE	283
Wells are existing in the data base of UNIGE	9

PhD student Maximillian data base:

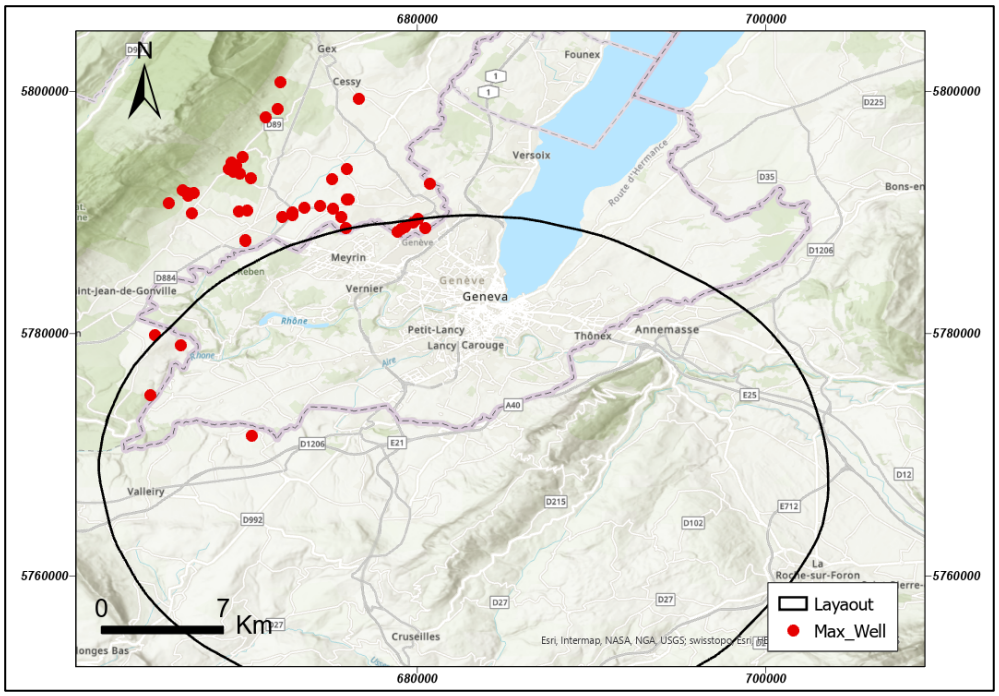


Figure 9: Boreholes collected by PhD student Maximilain,

In the Figure 9, all the wells that have been collected by PhD student Maximillian. These boreholes have been verified, checked and modified the coordinates system.

Wells are used in the data base of GADZ / Wells with No Molasse recorded/sallow wells	604
Wells used by UNIGE	57
Total	661

The BRGM Data base

The data collection was carried out over 7 months through the open data base InfoTerre [1] , it's an institutional website to BRGM scientific and technical information. InfoTerre and its map viewer are discovery tools for knowing and understanding the subsurface. InfoTerre offers planning professionals and the general public access to digital geological knowledge and interoperable geoscientific data: ***The La Banque du Sous-Sol : BSS of BRGM.***

BSS is based on legislative and administrative regulations and is one of the main missions of BRGM's National Geological Survey. It is a database that stores and manage the geological and technical descriptions of underground structures (drillings, boreholes, underground structures or excavation work) on French territory.

There are two possibilities to present data, the first one is to look at the availability of documents (Figure 10 and Table below).

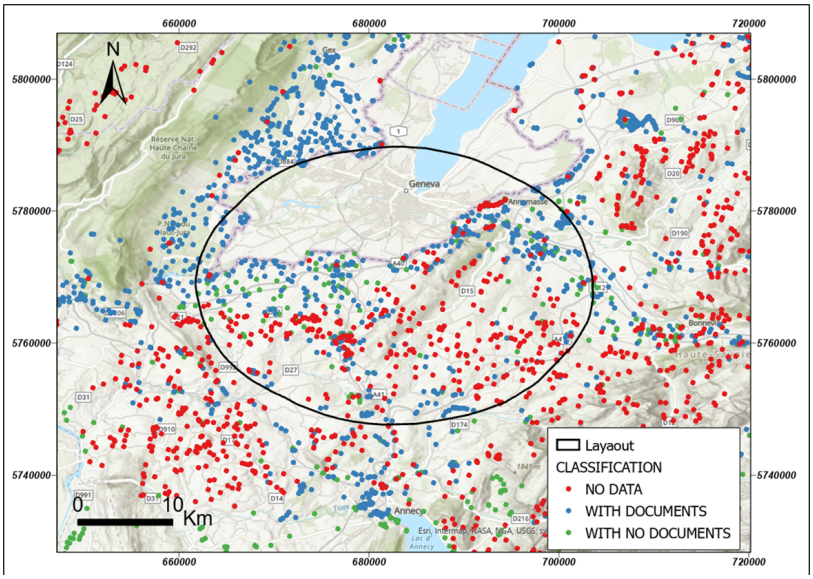


Figure 10: BRGM Boreholes classification (Source BRGM)

No data	1787
With documents	2280
Without documents but with digitalized information within info Terre	1007
Total	5074

The second way to see data are related to the classification of BRGM of these data, in order to facilitate the exploitation of the BSS data layers. there are six classifications, related the availability of document and the verification by BRGM (Figure 11).

[1] i-InfoTerre | InfoTerre (brgm.fr)

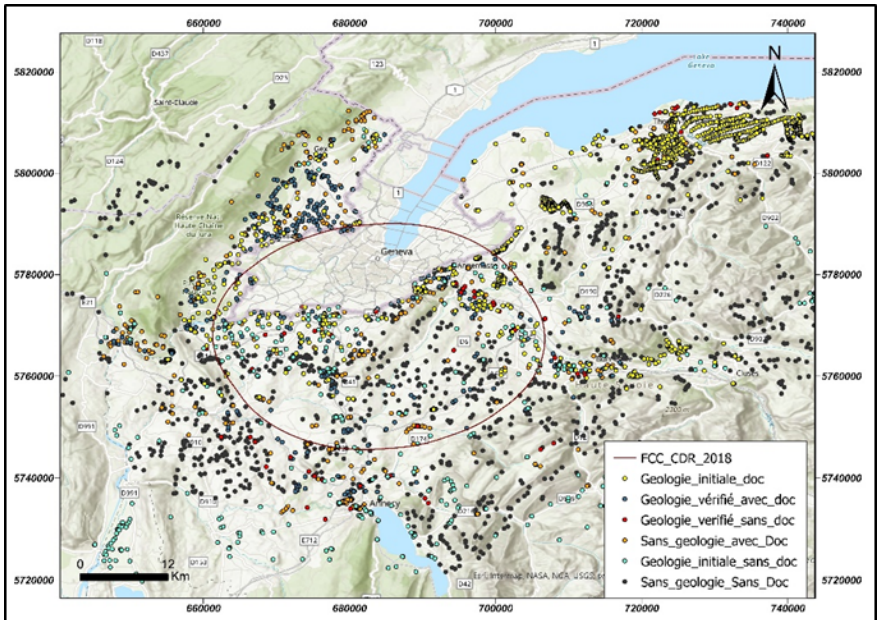


Figure 11: The BSS data layers (Source BRGM)

Verified geology with doc	484
Verified geology No doc	105
Geology Not verified with documents	621
Initial Geology with documents	1175
Initial Geology No documents	902
Geology Not verified & No documents	1787
BRGM DB	5078

The Unige team task was to collect the deepest wells that have information. For the wells that does not contain Molasse, they are used to equilibrate the surface of the base quaternary.
Some wells are shallow and recorded the top Molasse. In the Table below is the sum of the wells digitalized by UNIGE.

Wells are used in the data base of GADZ or Max	185
Wells used by UNIGE	541
Total	726

The UNIGE Data Base

It's basically an ArcGIS database, a collection of geographic datasets of different types stored in common file system folder. The database is intended for the management, storage, updating and consultation of entities of different natures, this data base is not relational data base. All the shape file is formed by flat attribute table. Furthermore, these entities could have link with each other by admitting and integrating

the SOLSTISS data base.
Well data are afterwards integrated in our GIS database within ArcGIS[2] software. These data projected in the Swiss coordinate system [3] using Projection Tool add the coordinate X and Y (see figure 12).

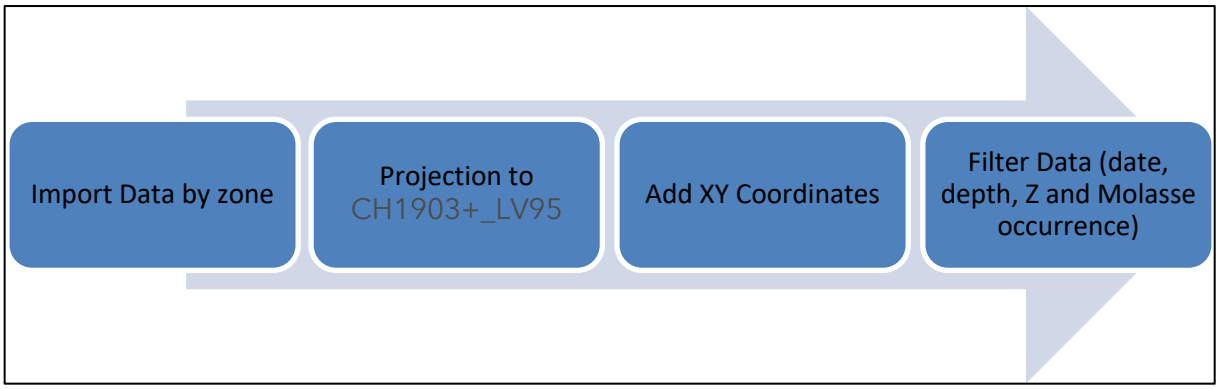


Figure 12: Geoprocessing steps

Project Tool: allow to specify the data's coordinate system without having to modify the input data. For the BRGM data base, the input feature class or dataset has a specified coordinate system Lambert 93 projection. Define Projection Another Tool of data management Tool, we used to intend for datasets that have an unknown or incorrect coordinate system defined.
There is the possibility to choose the coordinate system (geographic or projected) on the InfoTerre interface before downloading data.

Data integrated to UNIGE database:
We gathered all the data collected in this database, from the sources mentioned above. The total of points is 4213.

Source	Numbers of Borehole
GADZ	283
BRGM	541
MAX	57
SITG	3332
Total	4213

[2] ArcGIS Pro 2.7.0
[3] Projects spatial system CH1903+_LV95

All the data used on the 3D model are presented in the Tbale1:

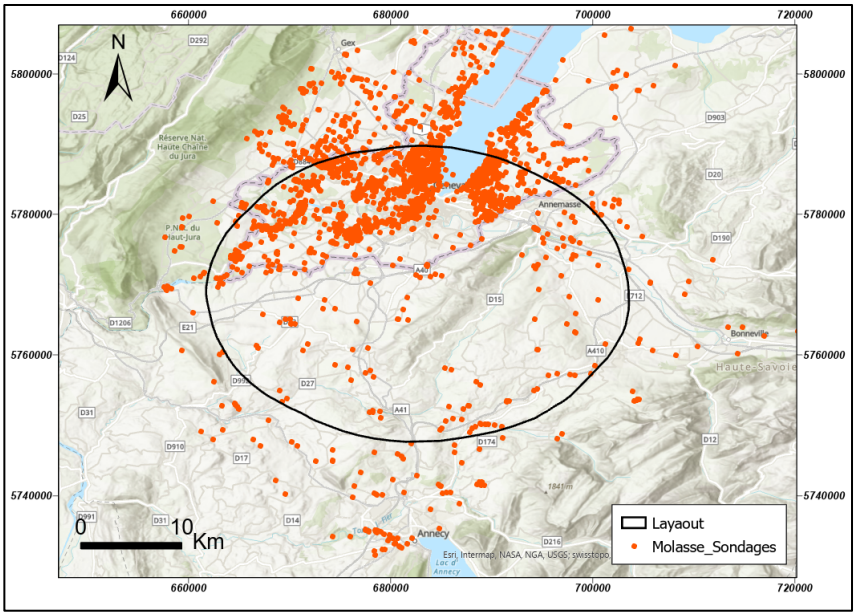
Table 1: UNIGE DATA BASE,

Source	Type	Format	Description	Spatial Reference	
				Before	After
BRGM[4]	BRGM_DB	Shp (point)	Geological survey *Processed Data*	Lambert 93 projection	Projected coordinate system CH1903+LV95
	BRGM_Piezometer	Shp (point)	boreholes equipped with a crenelated tube that allow to have access to the groundwater and to measure its altimetric level. *Processed Data*	Lambert 93 projection	
SITG[5]	GOL_SONDAGE_Public	Shp (point)	Geological survey *Processed Data*	CH1903+LV95	
	GOL_HYDRO_NAPPE	Shp (point)	hydrogeological map of the canton is the result of interpolations of borehole surveys *Row data*	CH1903+LV95	
	GOL_Piezometer	Shp (point)	boreholes equipped with a crenellated tube that allow to have access to the groundwater and to measure its altimetric level.	CH1903+LV95	
	DEM	Raster	Digital elevation model, values are different from 0 m. *Processed Data*	CH1903+LV95	
	Isoline: Molasse/Toit Riss/Toit Alluvion Ancienne	Shp (line)	isolines calculated from the numerical interpolation model *Row data*	GCS_CH1903	
	GOL_GEOPHYSIQUE_LIGNES, GEOL_EMPRISE_SISMIQUE_3D, etc.	Shp (line)	Geophysical data (2D-3D seismic, gravity, electric, etc)	CH1903+LV95	
	tdata_scpgeol_quat	Table	Geotechnical parameters	CH1903+LV95	
GADZ	GADZ_DB	Excel	Geology survey collected form GADZ data base	Lambert 93 projection	
PHD	Max_DB	Excel	Geology survey collected form Max Haas PhD data base	CH1903+LV95	
UNI	TopCret; TopMol	Shp (point)	Key geological surfaces (Top Carbonates, Top Molasse).	CH1903+LV95	

[4] Accueil | InfoTerre (brgm.fr)
[5] SITG | Le territoire genevois à la carte

GE	Faults	Shp (point)	Key structural features i.e. faults	CH1903+LV95	

The geological model of the top molasse surface is generated only with boreholes with molasse recorded.



the first figure presents all the well-integrated into the database of the UNIGE from all the sources.
The Figure presents the boreholes used to generate the geological model (Figure 13).

Figure 13: boreholes used to generate the 3D model

Requirements Filter: the main filters applied to this database:
Finally, and before integrating these data in the 3D Petrel project some requirements filters are applied:

1. The Molasse unit is recorded
2. Well is deeper than 50 m (for the purpose of the FCC tunnel project)
3. Wells between 20-40 m where deep wells are not available
4. Geological information and/or the geological reports are available

All the wells delivered to the CERN are digitalized and stored with the RGBA server.

The UNIGE Data Base architecture

In order to allow an effective transfer of subsurface data in to the SOLSTISS data base, the architecture of UNIGE Data Base has adopted the SOLSTISS structure ensuring a correspondence of attributes between the two data bases.
This work requires a careful examination of all BRGM data examining the all available original documents (pdf, jpg) and extracting the information required to fill up as much as possible the attributes included into the SOLSTISS data base.

In detail the attributes have been divided in 4 main categories (Figure 14):

1. Admin Common table
2. Borehole description
3. Hydrogeology Description
4. Geological Units

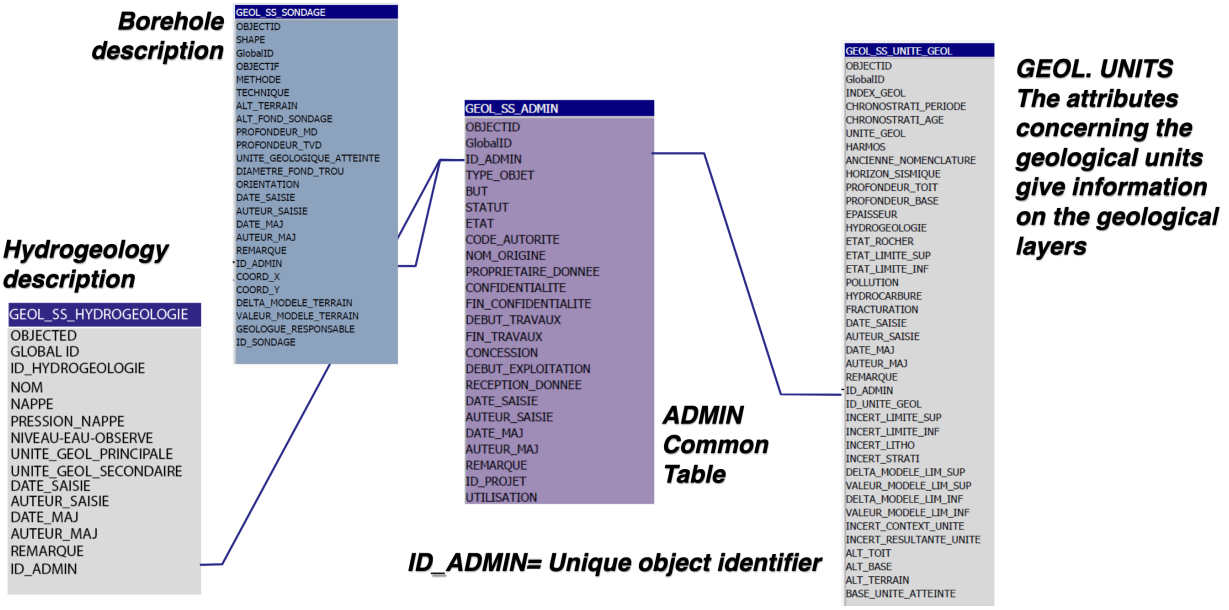


Figure 14: data base elements contained in UNIGE/SOLSTISS data base

More specifically Figures 15 and 16 provide an overview of the multiple excel worksheets used to capture the data for the selected subsurface boreholes.

Table	Nom attribut	Attribute name	Type de donnees	Valeurs de l'attribut	Description attribut	Unites
geol_ss_sondage	id_sondage	id_borehole	serial		Identifiant unique du sondage décrit.	
	objectif	objective	menu deroulant	Exploration; Exploitation; Autre; Inconnu	Renseigne sur l'objectif du sondage.	
	methode	methodology	menu deroulant	Carotté; Destructif; Non-foré (fouille/tranchée); Carotté et destructif; Autre; Inconnu	Indication sur la méhode de sondage majoritairement utilisée.	
	technique	drilling_type	menu deroulant	Forage à la boue; Forage à air comprimé (à sec); Forage à l'eau; Battage; Sonic; MFT - Marteau fond de trou; Pelle mécanique; CD - Double; CS -	Pour les sondages ayant plusieurs méthodes d'exploitation, se Renseigne sur la technique de sondage réalisée.	
	alt_terrain	surface_elevation	nombre decimal		Altitude du terrain à l'emplacement de la tête du sondage en mètres au-dessus du niveau de la mer.	m
	delta_modele_terrain	delta_surface_model	nombre decimal		Détermine la différence entre la valeur d'altitude de terrain saisie et la valeur du modèle existant pour le modèle	m
	valeur_modele_terrain	value_surface_model	nombre decimal		Détermine la valeur du modèle existant pour le modèle numérique de terrain	m
	alt_fond_sondage	down_drill_elevation	nombre decimal		Altitude mesurée au fond du trou du sondage. Champ auto-calculé : alt_terrain - profondeur_tvd	m
	profondeur_md	depth_md	nombre decimal		Profondeur effective du sondage en mètres. Contient aussi le rallongement lié à la courbure et aux déviations.	m
	profondeur_tvd	depth_tvd	nombre decimal		Profondeur correspondant à l'écart vertical en mètres entre la tête et le fond du sondage. Egalement appelée profondeur	m
	unite_geologique_atteinte	layer_reached	menu deroulant	Couverture et remblais; Ruisselement, Alluvions; Alluvions indifférenciées; Dépôts lacustres; Moraine würmienne; Alluvion ancienne; Interglaciaire Riss Würm; Moraine rissienne; Molasse; Molasse aquitanienne; Molasse indifférenciée; Retrait glaciaire indifférencié; Moraine indifférenciée; Retrait würmien; Retrait rissien; Grès et Marnes gris à gypse; Marnes et Grès bariolés; Calcaires eau douce inf; Gompholite; Grès sidérolithique; Calcaire crayeux à silex; Grès verts jurassiens; Rhodanien; Calcaire marneux de la Rivière; Urgonien blanc; Urgonien jaune; Complexe de la Pierre jaune de Neuchâtel; Complexe des Marnes d Hauterive; Calcaires à Alectryonia rectangularis; Calcaires roux; Chambotte sup; Marnes d Arzier; Chambotte inf; Vions; Pierre Châtel; Goldberg; Tidalites de Vouglans; Couches du Chailley; Calcaires de Landaize; Calcaires plaquetés; Calcaires récifaux; Calcaires de Tabalcon; Couches à Céphalopodes; Calcaires pseudolithographiques; Calcaires lités; Couches d Effingen Geissberg; Couches de Birmensdorf; Calcaires d Arnans; Calcaires terreux; Calcaires à entroques; Calcaires siliceux; Calcaires gréso micacé à Cancellophycus; Argiles et marnes de transition;Alternances micacés à bancs durs; Schistes en carton; Dalles échinodermiques; Marnes noires à nodules et à Tisoa; Marnes à amalthées; Marnes calcaires à bélémnites; Calcaires argileux à cassure conchoïdale; Calcaire à Gryphées; Calcaires gréseux à Chlamys; Grès blonds; Argiles à estheries; Intervalle marneux dolomitique; Alternances argileux gypseuses; Gypse massif; Keuper; Lettenkhole; Muschelkalk; Bundsandstein; Socle	Niveau effectivement atteint par le sondage. Peut ne pas correspondre à l'horizon cible. La liste des valeurs exhaustive de cet attribut correspondent à l'attribut "Chronostratigraphie_age" de la classe "Couche_geologique".	
	diametre_fond_trou	down_drill_diameter	nombre decimal		Il s'agit du plus petit diamètre mesuré dans le sondage. Ce dernier informe sur les équipements qu'il est possible d'installer.	m
	orientation	orientation	menu deroulant	Vertical vers le bas; Incliné; Horizontal; Montant; Vertical vers le haut; Dévié; Inconnu	Orientation à la tête du sondage.	
	coord_x	x_coord	nombre decimal		Coordonnée x du point de sondage.	
	coord_y	y_coord	nombre decimal		Coordonnée y du point de sondage.	
	geologue_responsable	responsible_geologist	texte		Précise le géologue responsable, en charge du relevé de sondage géologique. Renseigner le nom et prénom, ou les	
	date_saisie	input_date	date		Date à laquelle sont saisies les informations dans la base de données.	
	auteur_saisie	input_autor	texte		Auteur de la saisie des informations dans la table de la base de données.	
	date_maj	update_date	date		Date de la dernière mise-à-jour effectuée.	
	auteur_maj	update_autor	texte		Auteur de la dernière mise-à-jour effectuée.	
	remarque	note	texte		Remarques supplémentaires concernant d'éventuels	
	id_admin	id_admin	nombre entier		Identifiant unique de l'objet permettant la description de	

Figure 15: "Borehole" worksheet used in the UNIGE/SOLSTISS data base

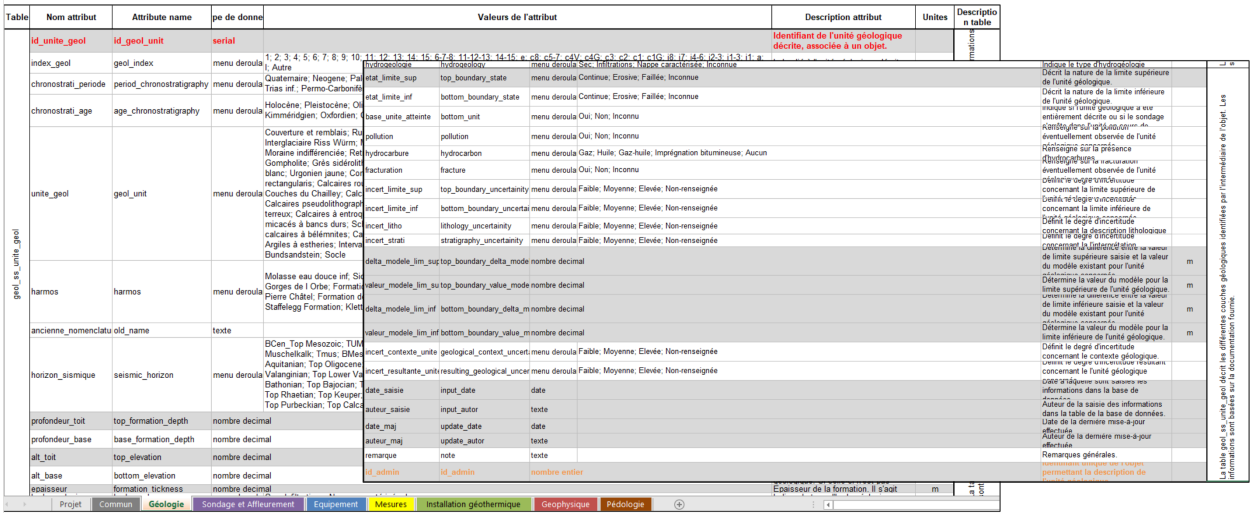


Figure 16: "Geology" worksheet used in the UNIGE/SOLSTISS data base

DATA BASE MANAGEMENT PLAN

The data base management plan refer to the phase of data base construction and organization managed by UNIGE following the mandate by CERN.

The following data base management plan does not include the tasks associated with the maintenance and update of the data base once the latter will be in place.

#	Task	Description	Task Owner
1	Data collection and QC along FCC trace within a 2 km buffer zone (4 km wide)	Collection, QC and harmonization of 300 borehole data from BRGM, GESDEC, SiTG.	UNIGE
2	Data architecture	Data are organized in a coherent format and list of attributes corresponding to the SOLSTISS catalogue	UNIGE
3	Data content	Data will include: 1. Admin Common table 2. Borehole description 3. Hydrogeology Description 4. Geological Units 5. Geotechnical parameters 6. Geophysical data (seismic, gravity, resistivity etc) 7. Key geological surfaces (Top Carbonates, Top Molasse).	UNIGE

		8. Key structural features i.e. faults	
4	Data uploading to FCC data base	Data are loaded locally at UNIGE in a ArcGIS data base/project	UNIGE
5	Data transfer / exchange	ArcGIS project is transferred to GESDEC to be incorporated into the SOLSTISS data base. Same will be transferred to CERN	UNIGE/CERN/GESDEC
6	Third party data access and data extraction	Data access will be ensured via the SOLSTISS web interface/GESDEC portal (GeoAnalyse) or/and FCC tools.	GESDEC/CERN ?

Following the data base construction, the GESDEC will ensure the updates, integration and management of new data in the Swiss territory. However, decisions on how to upload, QC and manage new data from the French territory (BRGM) need to be discussed.

CONCLUSIONS

- This document describes the workflow defined by UNGE and CERN on the context of the FCC program aimed at establishing a unique and consistent large regional subsurface data base covering the area along the proposed FCC trace in support of the FCC tunnelling project addressing both litho-geotechnical and hydrogeological aspects.
- This data base built by UNIGE is adopting the SOLSTISS architecture of the Canton of Geneva and results from the assemblage of multiple data existing in separated and non-homogenous data base such as the SITG (Canton of Geneva), the BRGM (France), unpublished data collected by the Geneva-based geotechnical company GADZ, CERN and by PhD and MSc research carried out at the University of Geneva.
- The large inconsistency in architecture and type of information (attributes) between the BRGM and SOLSTISS data base requires a detailed manual examination of individual borehole data (> 2000 records) which is currently in progress.
- The proper completion of the FCC subsurface data base requires an unexpected large work load which might need extra resources to be performed within the timeframe required by the FCC project.
- A data base management plan is proposed and needs to be discussed and endorsed by all parties.