

# **GIS & REMOTE SENSING FUNDAMENTALS**

45 h – 3 VNC – 6 ECTS

### Contact lecturers:

- Dr. Thi Nhat Thanh NGUYEN, Vietnam National University, Vietnam Coordinator
- Prof. LAFFLY Dominique, University of Toulouse Jean Jaurès, France
- Dr. Thi Minh Phuong LE, Hanoi Architectural University, Vietnam
- Dr. Quang Hung BUI, Vietnam National University, Vietnam
- Dr. Serge FARAUT, École Nationale Supérieure d'Architecture de Toulouse, France

### Objectives:

- Knowledge:
  - Introduction to the fundamentals of remote sensing and GIS, privileged tools of integrated analysis in environmental science. Presentation of theoretical concepts, quantitative methods mobilized and major domains of application.
- Skills:
  - Know how to apply the basic knowledge about GIS and Remote sensing; build GIS database; image processing.
- Attitude:
  - Active participation in group discussion, have the spirit to gain knowledge about GIS.

### Content of the courses:

- Chapter 1: Geographic information system
  - Lecture 1: a Brief History of GIS: the tool of integrated geographical analysis (component and concepts)
  - Lecture 2: Geographical and Map projection coordinate system; georeferencing and geocoding; GNSS input
  - o Lecture 3: Attribute data: data input, relationship, statistics, relation and joint
  - Lecture 4: Vector spatial data analysis: statistical map, graph theory and topology
  - Lecture 5: Raster spatial data analysis
  - Lecture 6: Vectorial multiscale analysis
  - Lecture 7: 3D analysis and data visualization
  - Lecture 8: Geographical PDF, layout and GIS report
- Chapter 2: Remote sensing
  - Lecture 1: History of aerial photography and remote sensing: From above we see the Earth better
  - Lecture 2: Physics of remote sensing
  - Lecture 3: Image quality:
  - Lecture 4: Sensors and application domain: land, water, atmosphere
  - Lecture 5: Remote sensing mathematical analysis
  - Lecture 6: Land use and land cover change (LULCC)
  - Lecture 7: Atmospheric remote sensing



## References

- Paul Bolstad, 2016, GIS Fundamentals: A First Text on Geographic Information Systems, Published by XanEdu Publishing Inc
- Báuden Bhatta, 2011, Remote Sensing and GIS, published by OUP Higher Education Division.
- Le Thi Minh Phuong, 2019, Mapping and GIS, Construction Publisher
- MIT <u>https://ocw.mit.edu/resources/res-str-001-geographic-information-system-gis-tutorial-january-iap-2016/introduction-to-gis/</u>
- IHE Delft https://www.youtube.com/watch?v=AfKp5mTa3XY&list=PLeuKJkIxCDj34vA8vQi\_S4ltQDoQb Wo72, https://ocw.un-ihe.org/course/view.php?id=11&section=0
- QGIS documentation: https://www.ggis.org/en/docs/index.html
- ArcGIS documentation: <u>https://learn.arcgis.com/en/</u>
- opencourseware PennState college: Cartography and Visualization (https://www.e-education.psu.edu/geog486/), Environmental Applications of GIS (https://www.e-education.psu.edu/geog487/node/11), Exploring Imagery and Elevation data in GIS Applications (https://www.e-education.psu.edu/geog480/node/11)
- Physics of Remote Sensing, ESA: <u>https://earth.esa.int/documents/973910/2642313/JG1to3.pdf</u>
- Atmospheric Radiation, MIT: <u>https://ocw.mit.edu/courses/earth-atmospheric-and-planetary-sciences/12-815-atmospheric-radiation-fall-2006/</u>
- Introduction to Remote Sensing and Visible RS, TU Delft: https://ocw.tudelft.nl/course-lectures/introduction-to-remote-sensing-and-visible-rs/



# Why here - and now - but not there?

Dominique LAFFLY

# The map as a principle of representation of geographic information

The map is the preferred medium for talking about geographic space, territory, environment and sustainable development. The map offers a synthetic vision - more or less detailed, we will see - of what is happening at a given moment in a portion of space. With the map we "see" how the information is distributed in relation to each other to derive interpretations and hypotheses:

- If there is no geographical distinction in the distribution of information: a priori there is no link between position and information.
- If the information is distributed spatially according to a spatial dichotomy (north-south opposition, for example) or in concentric areas around nuclei or even gradually in a privileged direction: a priori the position determines or at least influences the information.

In the second case it is decisive to understand why position plays on information? Is it a mountain / sea opposition that explains the dichotomy? Is it a climate variant that plays on the gradation of information? Is it the densities of populations which decrease in a halo around the urban cores which explain the central places? Are these administrative boundaries that define the spatial partition of information?

By repeating this analysis at different time steps - historically long or short - we will also observe the influence of the moment on information:

- If the maps are roughly identical at all time steps: the moment does not influence the card;
- If the maps are very different depending on the time step: the moment influences the map.

In this second case, as before, it is decisive to understand why the moment plays on information and its spatial distribution? Is it the climatic season that determines the differences? Are these climatic differences also part of longer term cycles / trends? Are it daily variations that seem to determine the spatial variations in information? It is easy to understand that this is the case for urban pollution linked to mobility, for example.

Spatial analysis therefore consists of understanding how time and place influence information? In the 1980s a French geographer, Roger BRUNET, proposed a graphic language for spatial modeling of geographic phenomena. Based on 4 forms (point, line, surface and network) and 7 strategies (mesh, grid, attraction, contact, tropism, territorial dynamics and hierarchy) BRUNET identifies 28 graphic modeling scenarios of geographic space, choremes (figure 1). This approach - partly decried in quarrels between schools of thought - has been taken up in numerous international publications applied to various themes.



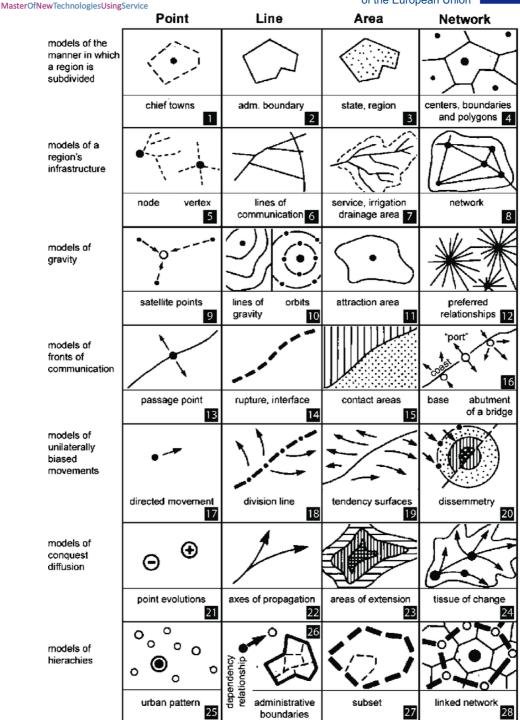
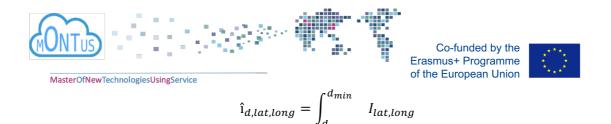


Figure 1: Table of Brunet's choremes with English labels, in A. Reimer, « Understanding Chorematic Diagrams: Towards a Taxonomy », November 2010, Cartographic Journal The, 47(4):330-350. Note that the subtitles are examples, not necessarily exhaustive descriptors.

Beyond this chorematic, a sort of quintessence of the analysis of spatiotemporal phenomena, let us retain that, if the moment and the position influence the information then we will be able to determine - one will say according to the case to extrapolate or model - for any point of the space provided with its latitude (lat) and longitude (long) coordinates, value of the information I am knowing the moment m (given not the date d) such as:



The information I itself being the result of the integration of n geographic information i\_g:

$$I = \int_{n}^{1} i_{g}$$

Maps make it possible to visualize information in geographic space at different time steps, spatial analysis makes it possible to integrate time and space to produce an explanatory model of information - geographic information. What are the paradigms and methods associated with this approach?

# Integrated analysis as a model for thinking about geographic information

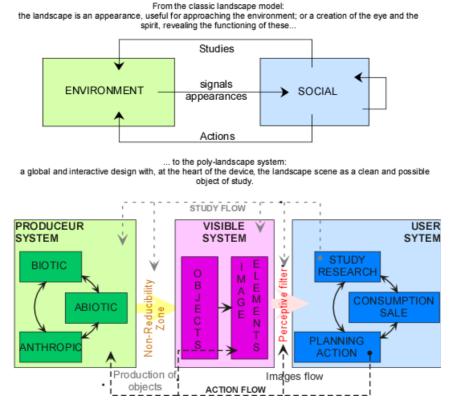


Figure 2: In the early days of the "Systemic Landscape" (modified from Brossard and Wieber, 1984)



Co-funded by the Erasmus+ Programme of the European Union



Geographical space can be considered as a system within which flows between constituent elements lead to feedback loops in an ever-changing temporal dynamic. T. Brossard and J.-C. Wieber<sup>1</sup> (1981), 2 french geographers, proposed a conceptual diagram of a systemic definition of geographical space (figure 2). Between production - the "physical" producing system - and consumption - the "social" user system - the geographical space is expressed by what is visible and non-reducible - "visible landscape" system - to one or the other previous subsystems. This specificity of the geographer to apprehend the landscape so as to make sense of space places it at the crossroads of multidisciplinary scientific paths, "The specialists of other disciplines now know that" nature "is never quite" natural ", or, conversely, that the analysis of social systems can no longer be considered detached from the environments in which they are located." (A. Frémont<sup>2</sup>).

Geographical space is a key concept, it's an entry point for themes that have a practical impact. This concept is linked to an analysis method specific to the geographer and his needs to spatialize - in the sense of continuously covering the space. The landscape's "signs" - information - allow for a quantitative approach that relies on the use of statistical and computer tools in search of the fundamental structures to, in a way, "replace the "visible complicated" perceived landscapes by "the invisible simple" spatial structures ". D. Laffly also explains that "Using measures, observations and systems of knowledge inevitably means introducing the notion of representativity in various ways. It includes questions about sampling strategies, the nature of the data, their disaggregation and aggregation, the equations used to model, extrapolate or interpolate information (the same mathematical function can be used for one or the other of these methods)... Any reasoned approach in information analysis tries to integrate at best these different aspects of the measurement and their spatiotemporal representability". We are also confronted with the difficulty of linking thematically specialized - endogenous punctual information - specific descriptions with general - exogenous areal information - the only ones that are amenable to taking into account the spatial continuum (figure 3). To do this, in the framework of the systemic formulation of the landscape and the mode of analysis related to it, we present a formalization based on four key elements: point, trace, order and inference.

Point: It is the basic spatial unit of endogenous observations made in situ. It is the subject of a precise location (differential GNNS and/or geocoding of addresses) and a standard description. Surveys are conducted according to a cybernetic logic and a systematic protocol so as to lend themselves to quantitative analyzes that describe and parameterize information structures. Sampling strategies are based on thematic and spatial criteria. For example, for biogeographic facies surveys, stratified nonaligned systematic sampling is commonly used at two levels. (De Keersmaecker<sup>3</sup>, 1989). The first to define the overall sampling plan of the points to observe in the field, the second to stop the in situ observation strategy for each previously defined

<sup>&</sup>lt;sup>1</sup> « Essai de formalisation systémique d'un mode d'approche du paysage », Bulletin de l'association des géographes français, n° 468, pp. 103-111.

<sup>&</sup>lt;sup>2</sup> « La télédétection spatiale et la géographie en France aujourd'hui », l'Espace géographique, n° 3, pp. 285-287.

<sup>&</sup>lt;sup>3</sup> Marie-Laurence de Keersmaecker, « Stratégie d'échantillonnage des données de terrain intégrées dans l'analyse des images satellitaires », L'Espace géographique, 1987, Vol. 16-3, pp. 195-205

This document was produced within the MONTUS project financed by the European Union in the framework of Erasmus + Capacity Building 598264-EPP-1-2018-1-FR-EPPKA2-CBHE-JP.



entity (Laffly & Mercier<sup>4</sup>, 2001; Moreau *et al.*<sup>5</sup>, 2004). Here we find the notion of integrated or holistic analysis.

- Trace: This is the message, the sign that reflects the links between the structures identified from the analysis of endogenous data and the exogenous information that will serve as a reference for spatialization. In this case, images of satellites and other geographical information such as altitude, slope, orientation, age of surfaces, distance to objects... Any information likely to describe the landscapes and available under continuous blanket form of space. It is the extension via the geographical coordinates of the description of the point in the exogenous information base. Beyond the pixels of images that are ideally suited to our approach, it can nevertheless be generalized to socio-economic data identified by reference administrative unit, *i.e.* the finest level available: IRIS<sup>6</sup> in France, NUTS<sup>7</sup> for GADM<sup>8</sup>. It is still necessary that these data exist, that they are validated, updated and accessible. The point data observed in situ will first be summarized (pivot table) by reference administrative unit to then be confronted with the identification of links the trace here potential.
- Order: Refers essentially to the spatial structuring of data, the arrangement of landscape elements relative to each other that induces differentiated spatial constraints and practices. In image analysis, order refers to the notions of textures and texture mosaics, spatial autocorrelation and opens the perspective of the frequency analysis of Fourier transforms and wavelets... From vector objects typically reference administrative entities, the analysis of spatial structuring uses topological operators of graph theory: shape descriptor (perimeter, surface, width, length ...); contiguity, inclusion, neighborhood; connection; of smaller distance... (see landscape ecology).
- Inference: This is inference in the statistical sense of the term, that is, the application of the rules developed in the previous steps to ensure the link between endogenous and exogenous information. It is an ergodic approach "which makes it possible to statistically determine all the achievements of a random process from an isolated realization of this process<sup>9</sup>" based on probabilistic models, which makes it possible to restore the continuity of geographical space from partial knowledge. We think in particular of Bayesian probability models Bayes, the way! as well as the Metropolis-Hastings algorithm, "Today, the entire field of MCMCs, the Monte-Carlo Markow-Chain, whose unreasonable efficiency in physics, chemistry and biology has not yet been explained. It is not a deterministic exploration, nor is it a completely random

<sup>&</sup>lt;sup>4</sup> D. Laffly & D. Mercier, "Global change and paraglacial morphodynamic modification in Svalbard", International Journal of Remote Sensing, Volume 23, 2002 - Issue 21.

<sup>&</sup>lt;sup>5</sup> M. Moreau, D. Mercier, D. Laffly, « Un siècle de dynamiques paraglaciaires et végétales au Svalbard (Midre Lovénbreen, Spitsberg nordoccidental) », 2004, Géomorphologie, 2 : 157-168.

<sup>&</sup>lt;sup>6</sup> "In order to prepare for the dissemination of the 1999 population census, INSEE has developed a division of the homogenous size territory called IRIS2000. An acronym that stands for "Grouped Islands for Statistical Information" and referred to the target size of 2000 inhabitants per elementary mesh". (https://www.insee.fr/fr/metadonnees/definition/c1523).

<sup>&</sup>lt;sup>7</sup> Nomenclature of territorial units for statistics, 6 levels (https://gadm.org/metadata.html)

<sup>&</sup>lt;sup>8</sup> https://gadm.org

<sup>&</sup>lt;sup>9</sup> http://www.cnrtl.fr/definition/ergodique

This document was produced within the MONTUS project financed by the European Union in the framework of Erasmus + Capacity Building 598264-EPP-1-2018-1-FR-EPPKA2-CBHE-JP.



exploration, it is a random walk exploration. But basically it's not new, it's the same in life: going a bit at random from one situation to another, we explore so many more possibilities, like a researcher who changes scientific continent at will meetings" (C. Villani<sup>10</sup>, 2012).

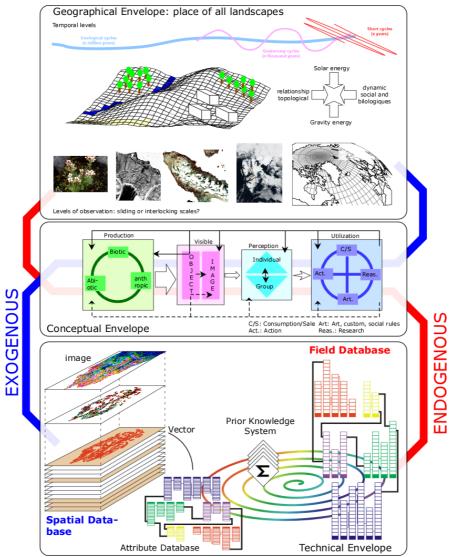


Figure 3: Elements of formalization of the geographical space system

From an operational point of view, the proposed formalization consists in measuring the degree of connection between endogenous and exogenous information, it's the way of integrated analysis.

It is fundamental to integrate here the distinction between scales of restitution and level of observation of information. The restitution scale refers to the dimension ratio between a measurement in the field and its representation on a map: 1 cm of the map represents 25,000 cm in

 $<sup>^{10}</sup>$  « Théorème vivant », Gallimard NRF, 350 p.

This document was produced within the MONTUS project financed by the European Union in the framework of Erasmus + Capacity Building 598264-EPP-1-2018-1-FR-EPPKA2-CBHE-JP.



the field on a scale of 1 / 25,000, for example. The level of observation of the information refers to the spatial properties of the measurement:

- i. The measurement can be point, linear, surface while knowing that a surface can just as easily be considered as a point, a line as a surface ...
- What minimum area does the information relate to? The pixel of a satellite image varies from a few meters to a few kilometers; a temperature measurement is theoretically located at the sensor; a river flow is for a local section of the stream...

Strictly speaking, all levels of observations can be rendered at any scale, except that there are limits to common sense: what does a scale of 1 / 10,000 mean for an image with pixels of 1 km on a side, or a 10 cm square on the map? Even with data of better spatial resolutions, there are restitution scale thresholds below and beyond which the data is not suitable (figure 4).

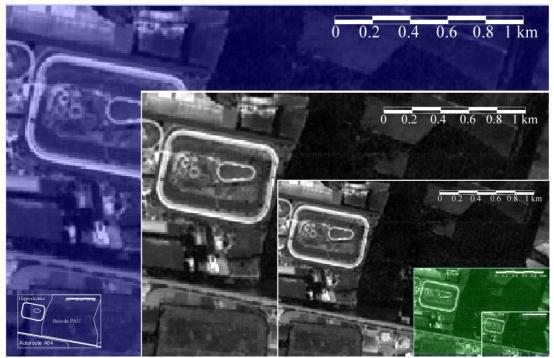


Figure 4: Observation level versus restitution scales, SPOT P image at 10 m (blue and green, scales not suitable)

Finally, the analyzed phenomenon (s) will not be geographically relevant depending on the level of observation, a fine local structure will not be highlighted with an excessively higher level of observation. For example, a study which aims to map all the water points - census, pollution measurement, diachronic evolution... - will not give the same results depending on whether the reference images have pixels of 2.5 m or 10 m or 100 m or more... and that the geographical space studied is made up of a multitude of small water points and rivers or, on the contrary of large lakes and large rivers.

Geographic space, level of observation, scales of restitution, types of data and methods of analysis must be considered together, this is the purpose of integrated analysis or the conceptual dimension *This document was produced within the MONTUS project financed by the European Union in the framework of Erasmus + Capacity Building 598264-EPP-1-2018-1-FR-EPPKA2-CBHE-JP.* 



of the study and / or research . Figure 5 illustrates this type of intellectual approach to research on sleeping sickness in Africa with all possible levels - from the parasite to endemic areas - where we distinguish the key elements (bioclimatic zones, demography, occupation of the soil, agricultural practices and social ties) and the data to be used according to the scales of restitution envisaged.

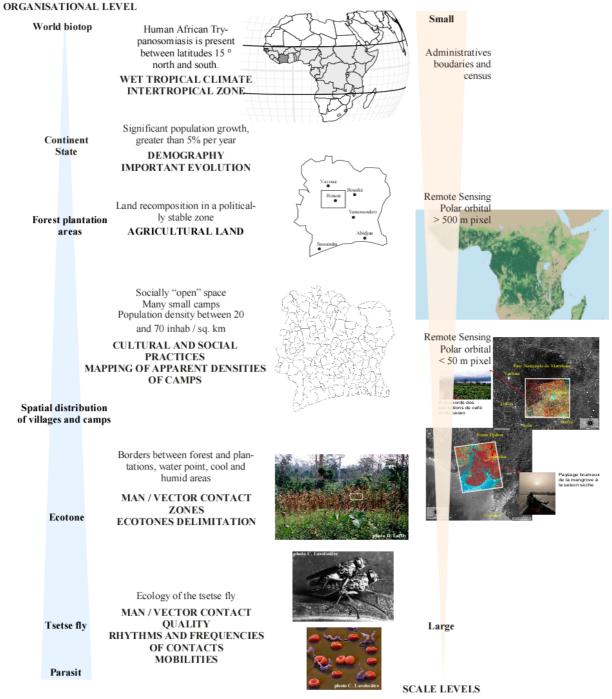


Figure 5 : How to think a sleeping sickness research according to the geographical environment in Africa



Co-funded by the Erasmus+ Programme of the European Union



This example applied to the environmental approach to sleeping sickness has its roots deeply in what we now call Geographic Information Systems as recalled by the authors of the Wikipedia page devoted to this subject: "One of the first known instances in which spatial analysis was used, came from the field of epidemiology in the, "Report on the march and effects of cholera in Paris and the department of the Seine" (1832). French geographer and cartographer, Charles Picquet, created a map outlining the forty-eight Districts in Paris, using halftone color gradients, to provide a visual representation for the number of reported deaths due to cholera, per every 1,000 inhabitants. In 1854, John Snow, an epidemiologist and physician, was able to determine the source of a cholera outbreak in London through the use of spatial analysis. Snow achieved this through plotting the residence of each casualty on a map of the area, as well as the nearby water sources. Once these points were marked, he was able to identify the water source within the cluster that was responsible for the outbreak. This was one of the earliest successful uses of a geographic methodology in pinpointing the source of an outbreak in epidemiology. While the basic elements of topography and theme existed previously in cartography, Snow's map was unique due to his use of cartographic methods, not only to depict, but also to analyze clusters of geographically dependent phenomena." (Figure 6).

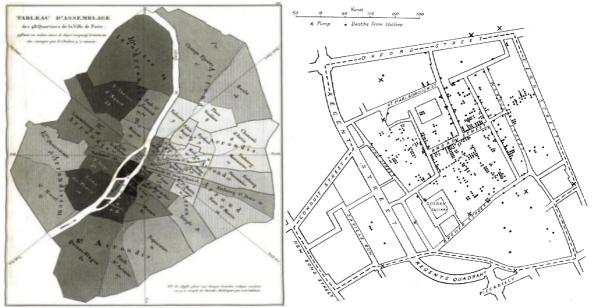


Figure 6: The origins of GIS, maps by Charles PICQUET (1832) and John SNOW (1854) to study cholera in Paris and London

How then to operationalize the concept developed to think about the geographic space in relation to a given problem? Geographic information systems and very often satellite images will be used for this purpose. This is the technical solution applied to integrated spatial analysis, the two go hand in hand, always.

# GIS & Remote Sensing - Fundamentals

### GIS





Co-funded by the Erasmus+ Programme of the European Union



To put it simply, a Geographic Information System (GIS) is a computer application which makes it possible to "superimpose" and "combine" vector-type data layers - provided with attributes - and / or others in the form of grid - we also say raster or image. The "superposition" is ensured by the consistency of the georeferencing of the data. The "combination" comes from the analysis which takes forms as varied as visual rendering, logical criteria, calculations between layers and spatial modeling... Figure 7 is a classic representation of a GIS where the geographic space is cut into successive layers which overlap perfectly... (source: USGS & NOAA).

Historically (lecture 1), GIS were first SI - a system for managing databases of vector data attributes with the Canadian Geographical System or CGIS: "The phrase," geographic information system ", was coined by Roger Tomlinson in 1963, when he published the scientific paper, "A Geographic Information System for Regional Planning". Ultimately, Tomlinson created a framework for a database that was capable of storing and analyzing huge amounts of data; leading to the Canadian government being able to implement its National Land-Use Management Program » (Wikipedia)

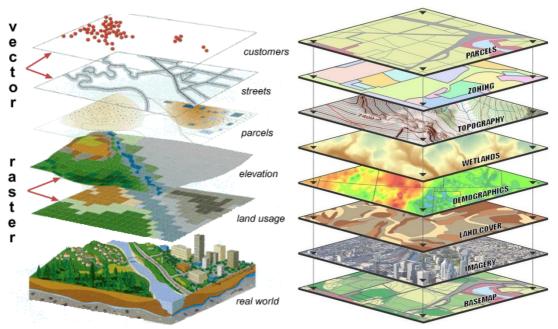


Figure 7: A GIS is made up of several layers of superimposed geographic information

In practice, the first difficulty consists in managing the georeferencing (lecture 2) and its configuration (we speak of Coordinate Reference Systen - RCS), all the data are not directly georeferenced and / or are not provided with CRS.

The realization of a map in a GIS must respect the rules of different graphic semiologies depending on the nature of the data (lecture 3) to produce either choropleth representations or other from the statistical description of the information. In addition, the GIS makes it possible to join the tables between them from common fields. Vector data can also be provided with a topological dictionary (lecture 4) - to be created if the sources are in the so-called spaghetti format - to have access to the topological operators of graph theory (continuity, contiguity, inclusion, exclusion, etc.).



Co-funded by the Erasmus+ Programme of the European Union



Raster data is distinguished from vector data (lecture 5) in that it is in the form of an image of one or more pieces of information. In all cases each raster layer (from a single file or from different layers within the same file) is the source of the information, there is no attribute table. It is important to distinguish a vector layer with 50 m square meshes (with attributes) from a raster layer with a resolution of 50 m, even if they overlap perfectly! The logic underlying the two are totally different and technically revert to different treatments.

It is easily understood that in a GIS the spatial "resolutions" can be different between the layers, the good control of the CRS allows a coherent rendering at the same scale of restitution. This is not to be confused with the multiscalar analysis of the data (lecture 6) where the same phenomenon can be apprehended at different levels of organization: typically the municipality, the department, the region ... If the higher levels come from the lower levels care must be taken to respect the nature of the data during the change: the population of the department is the sum of the population of all the municipalities in the department; the temperature on February 4, 1966 in the department is not the sum of the temperatures of all the municipalities in the department on that day!

When the data are provided with a height "Z" in addition to their X and Y or Latitude and Longitude coordinates (respectively data projected in meters, geographic data in angular coordinates - see CRS) we will be able to use the analysis. 3-dimensional so-called "3D" (lecture 7): Either for visualizations of volume and / or surfaces (digital terrain model - DTM) or to integrate 3D into the calculations (intervisibility, for example).

Finally, the production of reports that integrate geographic properties (graticules, legends, CRS, etc.) is an added value of GIS (lecture 8).

### **REMOTE SENSING**

Satellite images are essential to the analysis of geographical space, today the number and diversity of satellites allows systematic - and often free - use of this information. We usually talk about remote sensing.

Remote sensing consists - etymologically - in observing from a distance without contact between what is being measured and the device which performs the measurement. Defined as remote sensing applies to many fields, the one that interests us here is the observation of the Earth (surface and atmosphere) from the sky or from space. It being understood that the notion of sky corresponds to the atmosphere in which planes or drones fly and that the notion of space corresponds to the orbits of satellites. How to set a limit to the sky and the space? Do not imagine that this is pure fantasy because, for the very official International Aeronautical Federation "The Kármán line extends, according to international standards, 100 km above the surface of the Earth". (Wikipedia)



Co-funded by the Erasmus+ Programme of the European Union



Figure 8 : F. Nadar, aerial view of the district of Etoile, draw from negative glass in wet colledion 24 x 30 cm, July 16, 1868. *Ibid.* 

From the sky or space, therefore, the Earth and the atmosphere are observed by devices that measure electromagnetic radiation either from an echo of the signal sent by the device itself (we speak of radar remote sensing) or from the re-emitted solar radiation and that specific to the Earth (we speak of optical remote sensing).

Historically (**lecture 1**) everything began in Paris at the end of the 19th century: "Before the winter of 1858, Nadar chartered a balloon and rented the services of a member of the Godard family, the famous dynasty of aeronauts. The laboratory is installed in the nacelle, the fixed camera and lens oriented towards the ground. Nadar rises, triggers his "horizontal guillotine" acting as a shutter, immediately develops the image [...] We can distinguish perfectly the roof tiles and on the road a tapestry whose carter stopped short in front of the balloon [figure 8] " (*in* Études photographiques, n° 9, mais 2001).

Today sky-view images - with incredible precision - are part of everyday life for citizens of the world with applications such as Google Earth downloaded more than a billion times (in 2011, Wikipedia). However few people - including within the scientific community - know in detail all aspects of remote sensing a; when it is fundamental for the scientific use of data.

In the first place, as we have said, remote sensing measures an electromagnetic wave and it is necessary to be interested in the physics of the measurement (**lecture 2**) to understand the laws which govern this signal, the properties of the signal and its interactions during of its journey through the escape and the atmosphere as well as with the elements of the geographical space.

Satellite images are incredibly precise, as has also been said, whereas they should not be when we know all the constraints that penalize data acquisition: attitude deviation of the satellite on its sound. orbit; celestial mechanics - everything is in motion; optical and quantization aberration ... And yet the images are incredibly precise as the image quality processing is carried out and controlled (**lecture 3**).





The surface of the Earth mentioned here also includes continental surfaces, water surfaces (including ice) and the atmosphere (**lecture 4**). So which remote sensing for which application: radar or optical; continental or oceanic surface; air pollution or water vapor ...?

How do you analyze satellite images, from visualization in grayscale or in colored compositions to physical data modeled from the raw signal (for example, atmospheric depth) or even to land use maps ? All these analyzes mobilize mathematical methods to be taken into consideration (**lecture 5**).

Speaking of land use, as the first images go back to the early 1970s, as there are more and more compatible sensors providing free data - and we will see that yet remote sensing is very expensive - a lot of studies now provide analyzes of changes in land use to integrate the impact on the geographical space in the future (**lecture 6**).

Finally, in the same logic of sustainable development and taking into account the health impact of human activities, remote sensing is also used to measure air pollution (**lecture 7**).