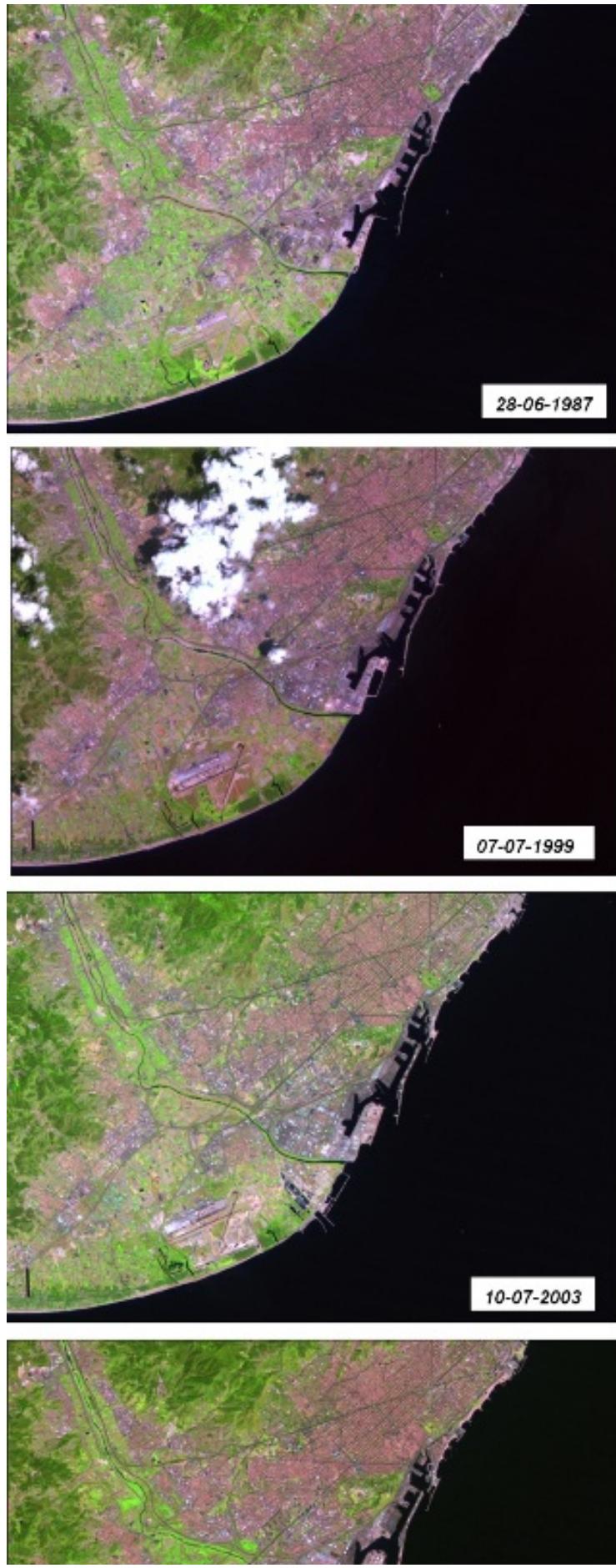




12/2014

Contributions to the Automated Analysis of Large Volumes of Data





This thesis is basically a study of the search for automatisms in different algorithms applied to Geographical Information Science. The works that form the thesis have an important common goal: to obtain the highest possible quality in results when replacing the usual (manual) method in the analysis of large volumes of data for the corresponding automatic method. This method shows several advantages over the former and in which metadata, i.e., information about data objects or resources, play an important role.

The thesis entitled "Contributions to the automatic analysis of big data in the context of GI Science" is essentially a study of the research in automation for several algorithms. The works that form the thesis have an important goal: the replacement of the usual method (manual) for the automatic one must pursue the highest possible quality of results.

The automatic method has several advantages over the manual: objectivity in decision making, repeatability, traceability, consistency and homogeneity (process quality, keep *human attention* for repetitive tasks), and a very important one for research goals: we can analyse the different responses of our model based on replications, we can change a few parameters or criteria, can replicate them in other regions or other dates, analyse different scenarios, etc. Moreover, from the operational point of view, the automation adds a powerful property (it is no possible by manual one): managing a large volume of data, the Big Data. This is a current problem (quite welcome, by the way) that is increasing fast. Today is widespread free distribution of data with a finer spatial and temporal resolution compared to only few years ago.

The main contributions following these specific objectives are:

- Automatic variogram analysis, distributed design and parallel solution for a geostatistical spatial interpolation method: kriging. It allows using this method in real-time applications with large volumes of data while maintaining the highest quality on the interpolated results.
- The integration of all processes in GIS environment: from geostatistical automatic modelling of various species of flora and fauna to the final map publishing (also automatic) to Internet of continuous variables.
- The geostatistical analysis of large series of remote sensing images has allowed to include research components on spatial pattern studies that improve other methodologies that do not take into account the spatial structure of the radiometric properties of images. Some

examples are: the selection of high quality images and the study of alterations caused by lossy compression methods.

- Automatic generation of reference values in areas with almost constant reflectance, and its applicability to a new model of automatic radiometric correction of Landsat images opens new challenges for studies where handling a large number of these images is needed.

Finally, as overall contribution of the thesis, we highlight the role of metadata on automatic processes. This new role adds geographic properties to more usual metadata applications as searches on catalogues, in this case, farer purposes from its geographical component.

Left figure: 4 scenes, selected from large series of Landsat images, showing the changes in the past 30 years in the mouth region of the Llobregat River and neighbouring infrastructures such as the airport and the harbour of Barcelona.

Lluís Pesquer

l.pesquer@creaf.uab.cat

References

“Contribucions a l'anàlisi automàtica de grans volums de dades en el context de la Ciència de la Informació Geogràfica”, Lluís Pesquer (CREAF) doctoral thesis, read at the Department of Geography and supervised by Dr. Xavier Pons.

[View low-bandwidth version](#)