Issue related to vegetation characterization and mapping using fieldwork, G.I.S. and remote sensing high resolution data

Casalegno S.

in

Rozakis S. (ed.), Sourie J.-C. (ed.).
Comprehensive economic and spatial bio-energy modelling

Chania : CIHEAM / INRA
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 48

2002
pages 139-147

Issue related to vegetation characterization and mapping using fieldwork, G.I.S. and Remote Sensing high resolution data

Stefano Casalegno
UMR Economie Publique - INRA

Abstract: Integrating remotely sensed high resolution data, Digital Elevation Model, agro-ecological data and fieldwork using G.I.S can help in estimating biomass yield. The present paper shows how vegetation mapping and characterization carried out in Southern Baja California, can be used as an example to propose some wider applicable principles and procedures. General problems related to fieldwork and to the acquisition and analysis of satellite images are also discussed.

Keywords: Digital elevation model, remote sensing, biomass yield, Sierra de la Laguna - Mexico.

Introduction

As suggested elsewhere in this volume, geographical detail is essential for the modelling of energy systems. Geographic Information Systems (GIS) allow the integration of Remotely Sensed High Resolution data (RS) with spatially referenced data as field data on the ecology of the biota, agronomic data, soil type data, numerical topographic data, climatic data and hydrologic data. The characterization and cartography of vegetation units, resulting from the analysis and synthesis of these geographically referenced data, can be introduced into a biomass yield model and lead to map estimation of biomass yield. The procedure for obtaining such a result needs to be adapted to the specific crop and particular geographical features, while proposals and generalization are at the same time possible. Considering the scientific research concerning cartography and agro-ecological characterization of a mountain forest landscape in Baja California, this paper aims to discuss the general problems related to the collection and analysis of geo-referenced data.

General procedure of mapping biomass yield using GIS, remotely sensed high resolution data and agro-ecological data

The principal steps in biomass yield mapping are as follows: (1) characterization of biota in the field and in remotely sensed data; (2) mapping and validation of mapping; (3) construction of a biomass yield model for the specific biota to relate the characterization of vegetation types in high resolution remotely sensed images to green biomass production under specific agro-ecological conditions. A comprehensive review of agrometeorological models was proposed by Vossen (1994), whose paper focused on point 3. In this paper we focus on points 1 and 2 of the procedure (see figure 1).

A large amount of georeferenced data is provided by Remote Sensing. Remote Sensing is defined as the science, technology and art of obtaining reliable information about the earth and its environment through the process of recording, measuring, interpreting and displaying imagery and digital representations thereof derived from non-contact sensor systems (Shunji MURAI, ISPRS President, November 1993). These images bring into instant view one or many phenomena; they have access to spectral bands not seen by the human eye, provide multi-temporal data running from a few days to many decades, and also a global synoptic vision of a large area (from regional to the whole Earth).

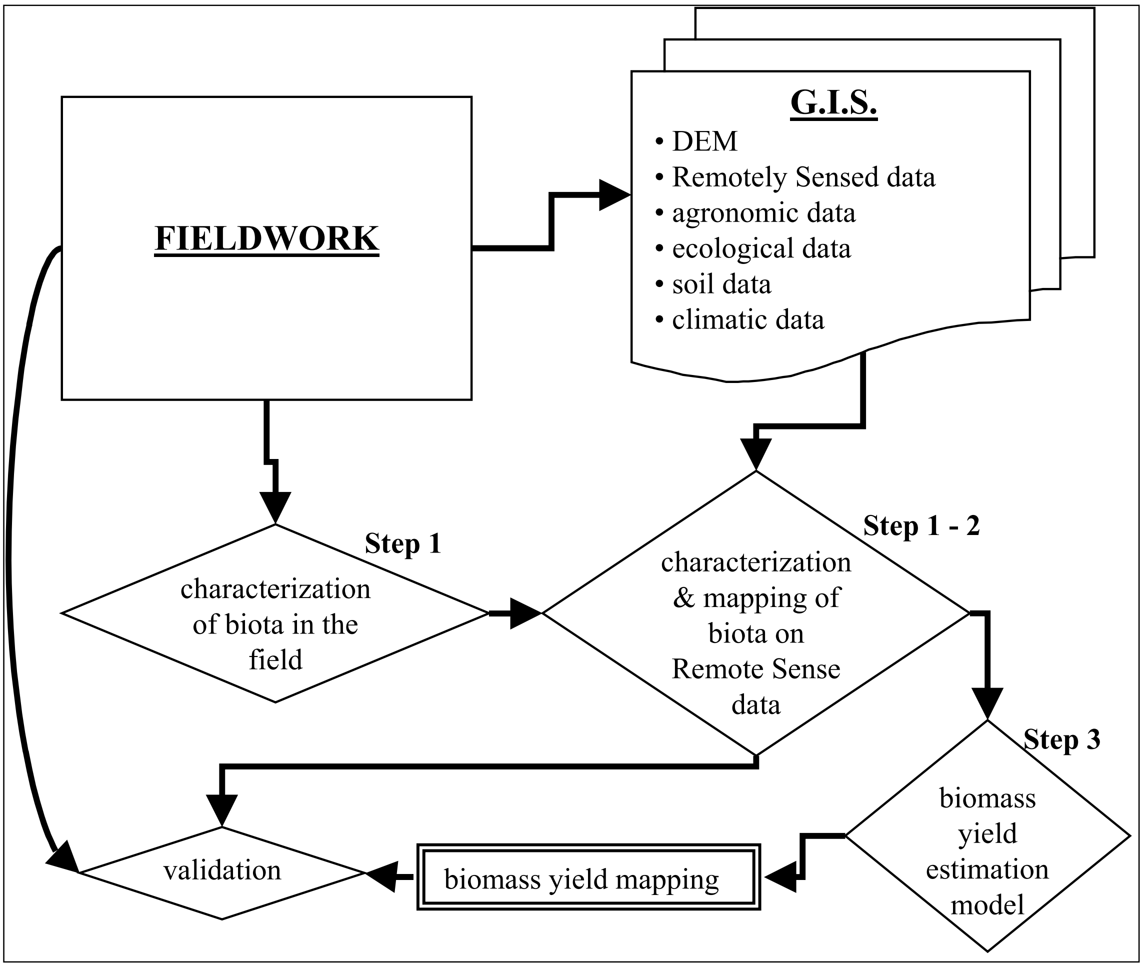

Figure 1. General procedure of mapping biomass yields

It has been observed in laboratory research (Allen *et al.* 1969; Knipling, 1970) that RS data can detect vegetation biomass using the following properties of the interaction between reflectance spectrum and vegetation covers:

> Between 380 and 700nm (visible), spectral behavior of reflected fluxes is related to chlorophyllous pigments; between 700 and 1300nm (near infrared), spectral behavior is related to vegetation tissue structure and between 1300 and 2500nm (middle infrared) spectral behavior is related to the water content of plants. Physiological changes of plants (pigment content, internal tissue structure, water content) are related to the growth, ripening and senescence and are responsible for important changes in the spectral behaviour of reflected fluxes in the visible and infra-red electromagnatic waves. Furthermore, the spectral behaviour of objects in the field can be close to that observed in the laboratory, if vegetation cover is close to 100%, but canopy structure and related shadow induced phenomena may cause a decreased reflectance.

Low vegetation cover (<20%) can be interpreted using soil properties of spectral reflectance, and intermediary vegetation cover can be interpreted using both soil and vegetation properties at the same time.

A combination (sum, difference, ratio or linear combination) of visible and infrared spectral band allow the calculation of different vegetation indices (Knipling, 1970; Rouse *et al.*, 1974;

Deering *et al.*, 1975; Richardson *et* Wiegand, 1977; Kauth and Thomas, 1976; Jackson, 1983; Huete, 1988; Kaufman and Tarne, 1992) that are useful for distinguishing between soil and vegetation and for evaluating plant biomass production.

Once the different land units have been defined by the rates of biomass production and a relationship between biomass in the field and satellite recorded data has been established, the RS data analysis then allows mapping to be performed. Mapping can be referred to studies, sciences, arts and techniques, used for elaboration and drawing up, from the results of direct observations or exploitation of documents, of maps, plans and other expression modes, as well as their use. (*International Mapping Association, UNESCO, Paris, 1966*). To make reliable the information recorded by RS sensors and mapping, a correlation has to be made between object units in the field defined at a spatial temporal dimension and RS numerical data obtained and analyzed at a reliable coherent spatial temporal dimension. Once field objects and numerical data have been correlated, an extrapolation is possible and maps are drawn up following this correlation.

An interesting option offered by mapping is the possible use of Digital Elevations Models (DEM), climatic and agro-ecological data to define and map potential yield estimate and successive map yield estimates using RS. Among the advantages of G.IS. is the ability to find the best location for the power plant that may be necessary in bio-energy systems modelling. G.I.S. software permits the integration of these different sources of data.

General problems and limitations related to the acquisition and analysis of remotely sensed data

The technology limitations of remotely sensed data are the choice of spectral and geometric resolution. In effect, some devices simply do not exist. One of the abilities of the amplitude and geometric resolution of spectral band for Landsat TM, SPOT and Ikonos images is to observe and quantify green biomass at a defined level of precision of analysis concerning forest and agricultural plot. Confusion between different types of chlorophyllian plants can be attributed to the extremely wide amplitude of satellite captors. Data derived from the MIVIS hyperspectral sensor (AIT 2001), carried by airborne support, prevent such a problem, but the cost of images are best adapted to research projects than to operative applications.

If the information required as input for the biomass yield model needs to be more detailed, RS may not be useful or may need to be supplemented by other sources of data.

Atmospheric limitations (rain, clouds, dust) may prevent data capture from satellite limiting specific data scene capture or revisiting capacity.

Another limitation of RS data is that sensing concerns only tall objects and surface phenomena or processes. Even when in agriculture, in general, this is not the case, if the information needed concerns soil properties or a lower stratum of vegetation, the object of the detection may be masked.

Signal saturation may prevent culture differentiation, especially with regard to agricultural parcels when dense vegetation is studied. Beginning with leaf index equal to two and eight, reflectance saturation is observed on visible and near infrared spectrum, respectively. Saturation is reached at a lower leaf index when the cover concerns horizontal habit plants such as sugar beet and at a higher leaf index when calculated over perpendicular habit plants such as graminaceous (Girard and Girard, 1999).

Study case of characterization and mapping of vegetation using GIS and Remotely sensed high resolution data: problems related

The case-study (Casalegno, 2001) took place in the mountainous region “Sierra de la Laguna”, in Baja California. The combination of a field study and numerical data (Spot XS, Landsat TM, DEM climatic and agro-ecological data) resulted in the characterisation and mapping of the transition zone between the tropical deciduous forest and the oak woodland. Problems related to the study procedure and results regarding fieldwork and the data analysis are presented in the next two sections.

Fieldwork: characterizing biota on the field

Problems related to the definition of spatial dimension of sampling concern the representativity of the sampled facts that we want to map. Facts that are directly sensed using RS data may concern : individuals, stands, or formations. Differentiation of unit can be done in function of vegetation structure, physiognomy, biomass, or landscape type. The object of the case-study concerned entities (more or less abstract) or concepts : species, sub-species, regions, land use, plant communities, ecological groups, dynamics of vegetation. In the case-study, mapping concerned vegetation structure - ecological groups – regions – landscapes. In a biomass yield mapping context, the facts sensed may be different, but in all cases, the choices regarding the definition of the data resolution (i.e. spatial and measurement type) and the fieldwork sampling technique have to be made. The quality of the scale-up process will depend on the resolution of the reference data in the field plot. The choice has to be made taking into account the final result of the object mapping, so that the sampling unit in the field is adapted to input in the biomass yield model and is able to detected in RS numerical data.

In the study-case, once defined the objective of mapping (the transition between two oro-biomes), we used the following protocol:

- Analysis of bibliographic data (spatial and semantic)
- Pilot sampling in order to:
  - Define the sampling technique (ex.: transect – stratified; stratified; aleatory) tacking into account of the geographical distribution and the extent of the mapping area estimated by bibliographic data.
  - Define the spatial resolution of the data: the minimum area of sampling that may be representative for the sample unit and detectable on RS data.
  - Define the resolution of the data in terms of measurement type (ex.: density, height of plants, trunk diameter, canopy cover, fine cartography)
- Sampling
- Data syntesis and analysis and characterization of the vegetation on the field

Pilot fieldwork was necessary to define minimum area (figure 2) as well as the resolution in measurement type, technique of sampling and extent of the sampling area.

In this case-study, the minimum sampling area of 400 m^2 allows description of the previously mentioned object defined in the objective of the study : the forest. Indeed, a parcel of 400 m^2 is representative of a larger forest zone with a sufficiently stable density of trees. This parcel dimension is also detectable by the Spot and Landsat TM satellite images that have a pixel resolution equal to 400 m^2 and 900 m^2, respectively. The dimension of the magnifying glass in the case-study was adapted to the object observed. A larger pixel resolution (ex.: NOAA captors = 1 km^2) would not be used in this study for the minimum fieldwork sampling area unless homogeneity from 400 m^2 to 1 km^2 could be investigated and tested or unless the minimum field sampling area were wider or repeated several times to relate the information in the field and with RS data : scale-up.

In the fieldwork, the type of measurement might vary, as well as the sampling area dimension from low resolution (ex.: density of species) to higher resolution (density and plant dimension such as diameter, height and canopy cover for each single plant) and intermediate resolution (density and plant dimension for a number of representative species or specimens).

Fieldwork costs occupy a large part of the budget in research work, as well as in case-study and biomass yield estimation projects. The opportunity cost of fieldwork has to be investigated, taking into account its efficiency and precision in describing field objects. A choice has to be made between high resolution measurement type and the number of repetitions of sampling plots necessary to describe the panoply of land unit type required. In the case-study, for example, we chose to apply various fieldwork sampling resolutions, adapted to different types of sub-objectives : (1) High resolution in measurement type (density, height, diameter and canopy cover for all woody plants) and low repetition of sampling plots (at least one for vegetation type for geographical location) allowed a precise characterization of forest types from a botanical and structural point of view. (2) Low resolution measurement type (density of plants) and several plot repetitions (several for each representative geographic location) allowed a description of the size limits for each forest type. (3) Intermediate resolution type (density of plants and height, diameter and canopy cover for few more representative species) and intermediate number of plot repetitions allowed a comprehensive functional interaction among forest types. Additional data concerning vegetation phenology, soil type, herbaceous stratum, and signs of human induced disturbance or fire were recorded in all sample plots. The three types of techniques have equivalent time cost expenditure and high-resolution measurement type, therefore several plot repetitions would be an undesirable opportunity cost option.

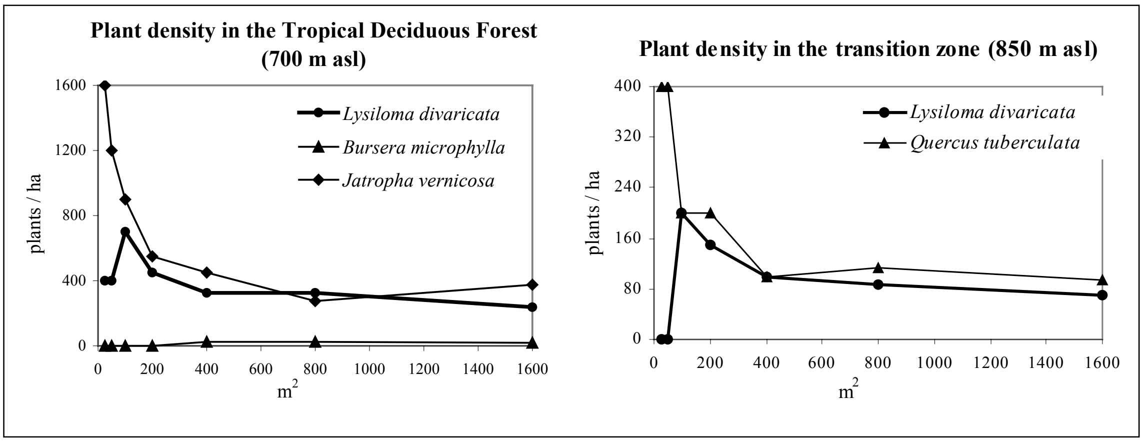

Figure 2. Definition of the minimum fieldwork sampling area. Density of dominant woody plant is recorded in plots of surface area running from 5 m^2 to 1600 m^2 doubling field plots dimension. Minimum sampling area is 400m^2. (starting from this value, density variation is lower of 3 % of the total variation).

In a biomass yield estimation study, sampling strategy (1) is the best option. The number of land units corresponding to different biomass yield levels are not numerous (three or four are satisfactory), but the precision needed for each land unit plot is high. A pilot study should be carried out to test the efficiency and precision of the fieldwork method to eventually modify and adopt a new procedure or technique to accomplish the objective. In reality, biomass estimation might be indirectly related to measurements recorded by RS data (ex.: detection of foliage biomass or physiological stage of foliage development to estimate sugar beet biomass), and field data needs to be collected that takes into account the special features of biota and its environment.

Numerical data analysis: issue related to match remote sense data and objects on the field and to use GIS technology for biomass yield estimate mapping.

The second part of the study-case was devoted to the RS data analysis in order to characterize and map forest types on wide and continuous geographic information covering the whole area of interest.

Principal steps of data analysis where:

1) Spectral characterization of forest types and definition of image processing procedure
2) Definition of best adapted date for image shooting

3) Image processing and vegetation mapping

4) Spatial modelling for potential distribution of natural vegetation and of areas subjected to ancient or recent anthropological pressure.

1)
In order to match remotely sensed data with objects in the field, a spectral characterization of forest types was attempted by means of definition of a classification procedure that would consider the problems related to image processing in rugged areas. Spot and Landsat TM images were analyzed using basic image processing methods (see Lillestand and Kieffer, 1994; Gomarasca, 1997; Girard and Girard, 1999). Unsupervised classification methods (Ascending Hierarchical Classification by Terravue software and K-Means Clustering by ER Mapper Silicon Graphics Computer Systems software) distinguished the spectral behavior and spatial distribution of each forest type. Confusion between forest type and no-forest land occupation still persisted (Casalegno and Girard, 1999). Further refined image processing using DEM consisted of splitting the original image, excluding the altimetry zone of non-interest, and classifying separately sub-images corresponding to the different slope expositions (Figure 2a). In this way, non-forest area where excluded from classification and problems related to shade were reduced.

In biomass yield estimation using RS data, slight differences in vegetation cover are sought. As in the case-study, non-interest areas had to be masked. Normalization models for image processing are available (Yang, 1992; Yang et al., 1993; Fashi et al., 2000) to ameliorate problems regarding signal disturbance induced by topography, atmospheric factors, illumination conditions and angle sighting. These methods need comparable spatial resolution between RS data and DEM.

2)
The choice of the most appropriate date of scene shot is fundamental. Image processing and field data on phenology allow selection of the best temporal scene adapted to the distinction of forest type. It was found in this study that the contrast among tropical deciduous forest / transition zone / oak woodland, was more pronounced at the end of the winter and early spring, while the contrast between well- preserved natural vegetation and disturbed vegetation was more visible on images at the end of the summer in the post-pluvial season (Casalegno *et al.*, 2000).

In the case-study as in any biomass yield estimation project, knowledge of the biological cycle of the selected culture and of other agricultural products or natural vegetation occurring in the same landscape is an essential condition for image analysis. It might happen as was the case in the case-study that images captured during different seasons are necessary to map the different levels of biomass yield for the same area and culture.

3)
The cartography of forest type and disturbed areas required two images (Spot scene of winter season and Landsat TM scene of post-pluvial season) processing and integration in one map of the whole data results (Casalegno, 2001). Validation on field points was carried out separately for each image. What is interesting to observate regarding a future biomass yield estimation project is the time of realization of such a RS analysis. Time of realization should consider pilot fieldwork, a sampling campaign, image processing and an second fieldwork to validate map results. The duration time of the biological cycle for a specific culture will be the factor that determines the time of the project's realization.

4)
GIS allowed the construction of a spatial model of potential distribution for natural vegetation and spatial distribution of areas subjected to past or recent anthropological pressure. Models were used to validate hypothesis about factors that determine vegetation distribution and forest ecology functioning.

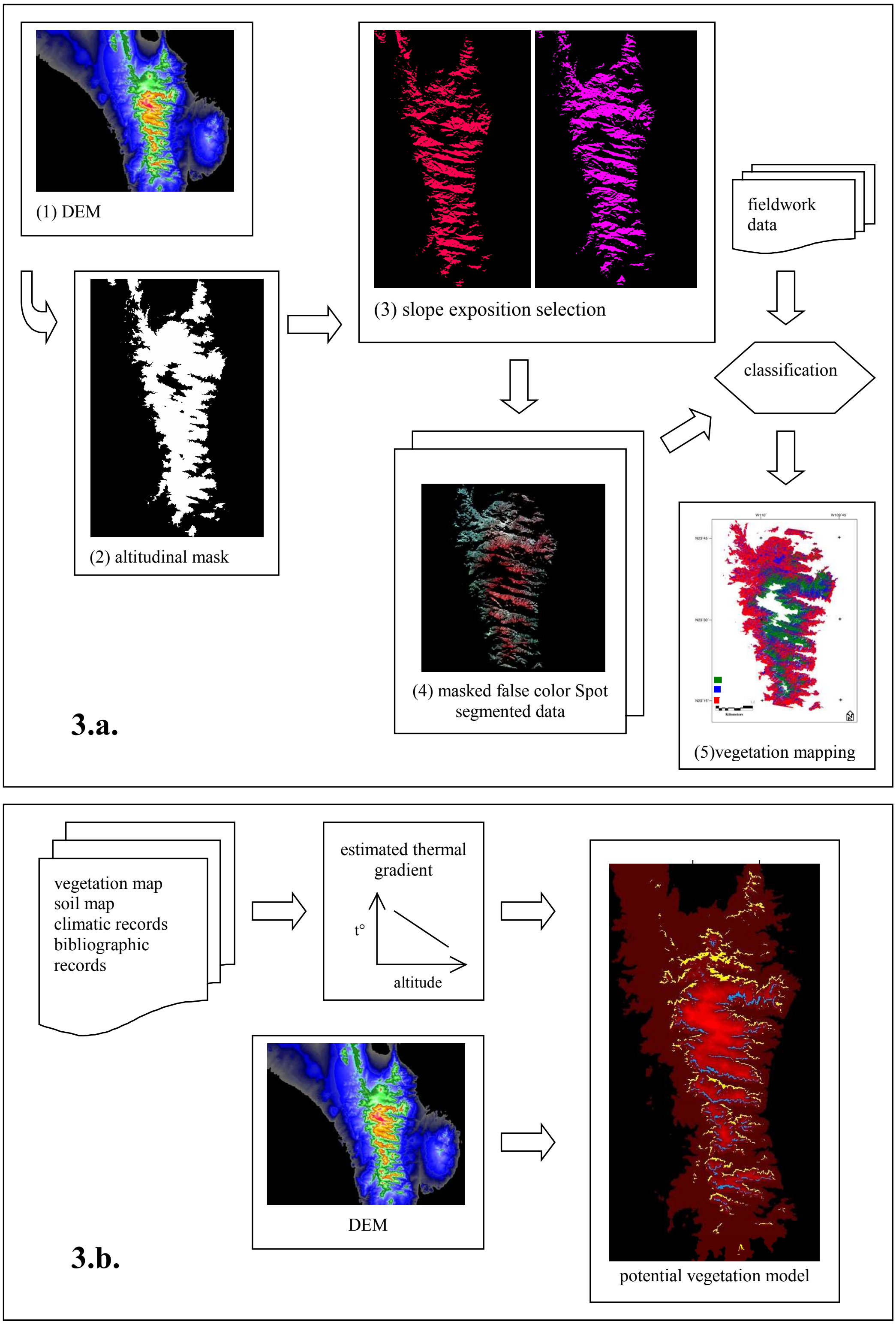

Figure 3a. Use of G.I.S. in DEM-Remote Sensing-fieldwork data analysis procedure.
3b. Spatial modelling procedure for potential vegetation distribution.

Boundaries existing in forest (as in the case-study) or in natural areas can be gradual. The potential distribution of natural forest type (Figure 2b) was proposed using the existing DEM and climatic data. This potential forest distribution map was traced using a soil and vegetation map, literature information about thermal limits for tree species distribution and the thermal gradient that was calculated using 48 year and 10 year-climatic station records. GIS was used to integrate human pressure activities such as livestock production, road and dust road network, past mining activity and distinguished potential degraded zones. Anthropogenic/historical factors and biotic/abiotic factors generating different types of land occupation in the studied area interacted. When a comparison was made of field observations, the RS generated map and spatial models, it was clearly indicated that spatial models agreed with other sources of data but the precision of the spatial model is very low and improvement in the input data is required. The modelling procedure was efficient but the parameters needed to be adjusted (e.g.: ecological limits for tree species) and refined (e.g.: considering mid-exposure not just Northern and Southern exposure. Another possible source of discrepancy was found to be an equal level of biomass detection in RS data for different types of forest (degraded /well-preserved) due to the evolution from a forest type ecosystem to that of a scrub type without consistent loss of green cover.

In biomass yield mapping, the special traits of agronomical culture allow a precise differentiation of vegetation units. Boundaries between agricultural parcels are defined by farmers in terms of their personal preferences, national or local politics, productivity and yield of the land submitted to a specific biotic and abiotic environment. In the present case, boundaries were sharp and defined. Confusion among parcels and their different biomass production levels might persist, as was noticed in the case-study where disturbed or natural forest output levels differed. For this reason, a clear differentiation among parcels in the field and an eventual aggregation of biomass production level might be proposed to avoid such confusion.

In regions where agronomical land registry and land use information are available, potential biomass yield in agriculture can be estimated by integrating georeferenced agronomical information in GIS and by calculating both the yield by means of DEM data and climatic data annually. A similar procedure was used for potential vegetation mapping in the present case-study and in estimating foliage biomass in Japan (Tatsuhara and Kurashige, 2001). This procedure would be more reliable for areas with extensive climatic and agro-ecological data.

Conclusion

The present spatial analysis application case-study, lies in its ability to provide general information about vegetation characterization and mapping through fieldwork, Remote Sensing and DEM using G.I.S.

Remote Sensing does not hinder fieldwork, but supplements it and makes it more efficient, allowing a scaling-up process and mapping to be carried out. The integration of climatic and agro-ecological data with RS data in a G.I.S. improves and clarifies the modelling process. Special features of the biota, its environment and modelling objectives should be taken into consideration when the fieldwork sampling decision and the RS data choice are made. The quality and precision of the results obtained are directly determined by the data source and by the data source compatibility.

Acknowledgements

The study case was partially funded by the French - Mexican Ecos M94B02 research project. I would like to thank Dr. L. Arriaga, Dr Gilliot, Dr. C-M. Girard, Dr. M-F. Passini, researchers and technicians from *CIBNOR* La Paz - Mexico and from *INA PG - DMOS* France. We are grateful to Sierra de la Laguna farmers and guides for their participation and help in fieldwork.

References

[1] AIT Associazione Italiana di Telerilevamento. 2001. *I rilevamenti iperspettrali MIVIS: ricerca e applicazioni*. Rivista Italiana di Telerilevamento : 20/21 – Sett.200 / Genn. 2001.

[2] Allen, W.A., Gausman A.J., Richardson A.J., Thomas J.W. 1969. Interaction of isotopic light with a compact plant leaf. *J. Opt. Soc. Amer*. **59** : 1376-1379.

[3] Casalegno S., Girard C. 1999. *Spectral Characterization of a Mountain Forest Ecotone*. (Caratterizzazione spettrale di un ecotono forestale di montagna) in ASITA, ed. 3a Conferenza Nazionale ASITA. Informazioni territoriali e rischi ambientali, Napoli, Italia.

[4] Casalegno, S., Arriaga L., Girard C.M., Passsini, M-F. 2000. Utilisation de données de télédétection pour la caractérisation d'un écotone en région tropicale de montagne: Exemple dans la Sierra de la Laguna, Basse Californie Sud Mexique. *Bulletin de la Société Française de Photogrammétrie et Télédétection* **160** (4): 3-12.

[5] Casalegno, S. 2001. Caractérisation et cartographie d'un écotone forestier de montagne afin de fournir des indicateurs de perturbation d'origine anthropique. Analyse à différents niveaux d'appréciation et spatialisation de la transition chênaie - forêt tropicale caducifoliée sur le massif Sierra de la Laguna au Mexique*).* PhD Thesis, Institut National Agronomique Paris - Grignon.

[6] Deering D.W., Rouse J.W., Haas R.H., Schell J.A. 1975. *Measuring forage production of grazeing units from Landsat MSS data*. Proceedings 10th Intern. Sympos. *Remote Sensing of Environment*, vol II: 1169-1178.

[7] Fahsi, A., Tsegaye, T. Tadesse, W. Coleman, T. 2000. Incorporation of Digital elevation models with Landsat-Tm data to improve land cover classification accuracy. Forest Ecology and Management 128: 57 - 64.

[8] Girard M.C, Girard C.M. 1999. *Traitement des données de télédétection*. Ed. DUNOD, Paris.

[9] Gomarasca M. 1997. *Introduzione a Telerilevamento e GIS per la gestione delle risorse ambientali*. AIT. 246p.

[10] Huete A.R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of the Environment* **25** : 295-309.

[11] Jackson R.D. 1983. Spectral indices in n-space. *Remote Sensing of the Environment* **13** : 409-421.

[12] Kaufman Y.J., Tarne D. 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS, IEEE. *Transactions Geoscience Remote Sensing* **30** : 261-270.

[13] Kauth R.J., Thomas G.S. 1976. The Tassel Cap, a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat*. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, IEEE Catalogue 76 1103-1 MPRSD, LARS Pruudue University, West Lafayette, Indiana.

[14] Knipling E. B. 1970. Physical and Physiological bases for the reflectance ; of visible and near infrared radiation from vegetation. *Remote Sensing of the Environment*, **1** : 155-159.

[15] Lillestand T. M., Kieffer R.W. 1994. Remote sensing and image interpretation. Thoird edition. John Wiley & Sons Inc. 750p.

[16] Richardson A.J., Wiegand C. 1977. Distinguishing vegetation from soil background information. *Photogrammetric Engeneering & Remote Sensing* **43** : 1541-1552.

[17] Rouse J.W., Haas R.H., Schell J.A., Deering D.W., Harlan J.C. 1974. Monitoring the vernal advancement of natural vegetation, NASA/GSFC Final Report, Greenbelt, MD, 371p.

[18] Tatsuhara S., Kurashige H. 2001. Estimating foliage biomass in a natural deciduous broad-leaved forest in a mountainous district. *Forest Ecology and Management* 152: 141-148.

[19] Vossen, P. 1994. An overview of methods for national crop yield forecasting. In : Proceedings of the Workshop for Central and Eastern Europe on Agrometeorological models : theory and applications in the Mars Project. Ispra (Italy) 21-25 November 1994. J.F. Dallemand and P. Vossi editors. JRC-European Commission.

[20] Yang, C. 1992. Correction radiométrique des effets topographiques sur les images satellitaires SPOT. Les Dossiers de la revue de géographie alpine 8: 15 - 27.

[21] Yang, C., Urua,W. Christian, P. 1993. Correction radiométrique sur les données satellitaires en zone de relief : Modelisation et application. Pages 65 - 80 in Aupelf-Uref, ed. Télédetection et Cartographie. Les Presses de l'Université de Québec, Québec, Canada.