Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
Colture Arboree ed Agrosistemi Forestali, Ornamentali e Paesaggistici

Ciclo XXV

Settore Concorsuale di afferenza: 07/B2
Settore Scientifico disciplinare: AGR/05

NEW TECHNIQUES FOR THE REMOTE SENSING OF FOLIAR NITROGEN CONCENTRATION IN FOREST ECOSYSTEMS

Presentata da: Dott.ssa Elena Mezzini

Coordinatore Dottorato

Prof. Luca Corelli Grappadelli

Relatore

Prof. Federico Magnani

Esame finale anno 2013

Alla mia famiglia,
passata, presente e futura.

“The clearest way into the Universe is through a forest wilderness.”
(John Muir)
“We can go anywhere we want, this is satellite.”
(Howard Stern)

INDEX

1. INTRODUCTION

This work is structured as a sequence of articles produced by my PhD activity at the University of Bologna, in the area of Forest Management and Silviculture.

The first chapter is a general introduction about the importance and the methods to analyze N concentration and N content in forest environments, both by analytical methodologies and from remote sensing imagery.

The second, third and fourth chapters describe three different studies: one in S. Rossore (Pisa) in 2000, another one also in S. Rossore but with samples collected in 2009 and the last one in Catalonia with data acquired between 1988 and 2001. Part of these experiences have taken place abroad: a piece of S. Rossore 2009 work was realized during a stay of three months at the University of New Hampshire, NH, USA; while almost all the study concerning Catalonia was carried out during a stay of three months at the Universitad Autonoma de Barcelona, Barcelona, Spain.

The aim of this thesis is to test new methodologies to estimate foliar N concentration from remote sensing data, using both aerial and satellite (MODIS - MODerate resolution Imaging Spectroradiometer, Landsat) imagery, also trying to understand the real significance of results and their applicability as parameterization of eco-physiological models.

1.1. Relevance of leaf and canopy N concentration

Foliar nitrogen concentration is a key parameter for plant eco-physiological processes and forest increments studies (Serrano et al. 2002). Nitrogen is a primary component of chlorophyll and plant enzymes, and has therefore a central role as photosynthesis and leaf respiration regulator (and therefore controls in the end net ecosystem productivity); moreover it is correlated to soil characteristics such as C/N ratio and soil fertility, and to forest canopy proprieties, in particular albedo. Also its temporal and spatial variations are full of information because of their dependence not just from site characteristics, but also from external perturbation such as climate change, N deposition and increases in atmospheric CO_2.

Particularly interesting for our studies is the correlation between N and GPP (Gross Primary Production), as a result of the effects of leaf N on light use efficiency. In fact a high N availability generates a high photosynthetic capacity and a high photosynthetic rate (Serrano et al. 2002).

Higher levels of foliar N are also often associated with a greater LAI (Leaf Area Index) and a higher fraction of absorbed photosynthetically active radiation (fPAR). Finally, the higher N availability associated with high leaf N concentrations reduces belowground growth allocation and increases

even more the ANPP (Aboveground Net Primary Production) of the forest. Eventually, a higher ANPP is associated with an increase in aboveground wood production, and therefore in current annual increment (CAI).

Recent studies (Ollinger et al. 2008) have highlighted the positive influence of N concentration on the maximum canopy-level photosynthetic capacity (Fig. 1.1). This regression is true both if the photosynthetic capacity is per unit of ground area or per unit of canopy leaf mass, although a better determination coefficient is generally observed when considering the latter parameter. Comparable results are also obtained when considering leaf N content (g m^{-2}) instead of concentration (%) (Peterson et al. 1999).

Other studies (Smith et al. 2002) contribute in demonstrating of the link between leaf N concentration, ANPP and wood production. In fact, these two parameters are positively related also to LAI and leaf mass, but with different trends depending on the vegetation type considered (evergreen or deciduous forests), so enhancing the possibility of discerning these two categories through the analysis of this relations (Fig. 1.2).

Further researches (Ollinger et al. 2002) studied the relation that exists between N cycle and leaf chemical characteristics; a significant relationship has been highlighted between foliar N concentration and N mineralization for broadleaf species (Fig. 1.3a), different depending on the species; this is in contrast with the same analysis performed for coniferous (Fig. 1.3b).

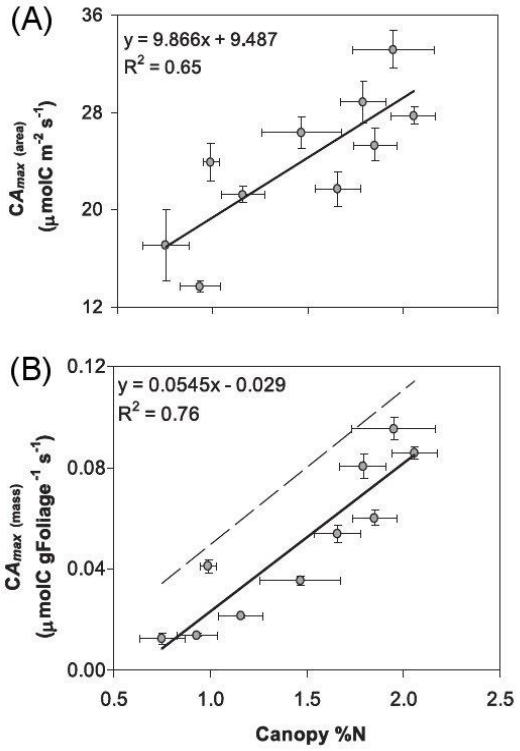


Fig. 1. 1 Relations between canopy N concentration and photosynthetic capacity. Linear regressions between canopy N concentration and photosynthetic capacity per unit of ground area (A) and per unit of mass (B) (Ollinger et al. 2008).

No significant pattern was observed in species belonging to this functional group, whose leaf characteristics thus seem more independent from N availability. Therefore only for broadleaves it could be possible to reverse the relation and estimate N mineralization, an important biogeochemical process, from foliar nitrogen concentration, a parameter much more easily measured.

The same research also showed the existing link between soil C/N ratio and leaf lignin/N ratio. Since leaf lignin/N could be estimate from remote sensing (in Ollinger study through AVIRIS imagery), it would be possible to utilize remote sensing data to estimate soil C/N ratio (both organic an mineral) through the relation that link it to lignin/N ratio, so as to generate useful C/N maps at landscape and regional level.

Leaf N concentration is not only important for the analysis of forest canopy characteristics, but its spatial and temporal evolution gives information for the study of global change and the influence of human activities on the environment. Recent reviews (Galloway et al. 2004) have detailed the evolution of atmospheric nitrogen deposition over the last century. Galloway's maps displays not only a global increase in nitrogen deposition, but also a higher localization in developing countries

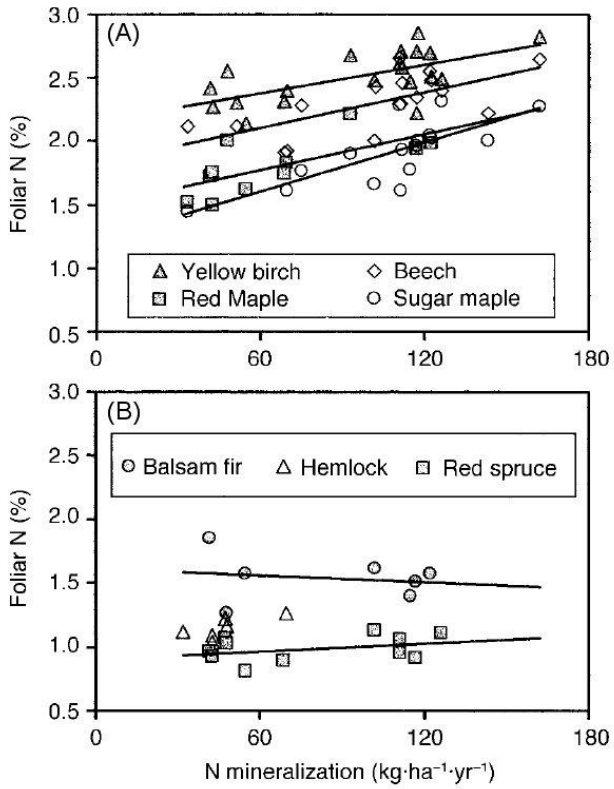


Fig. 1. 2 Relation between N mineralization and foliar N concentration. Foliar N concentration for individual species in deciduous (A) and coniferous (B) forests. Relations are significant at $P<0.05$ for all deciduous species (*Betula alleghaniensis*, *Fagus* spp., *Acer rubrum*, *Acer saccharum*). Trends are not significant for coniferous species (*Abies balsamea*, *Tsuga* spp., *Picea rubens*) (Ollinger et al. 2002).

(India, China, Latin America, etc.) in recent years, in contrast with the stabilization and reduction achieved in Europe and North America.

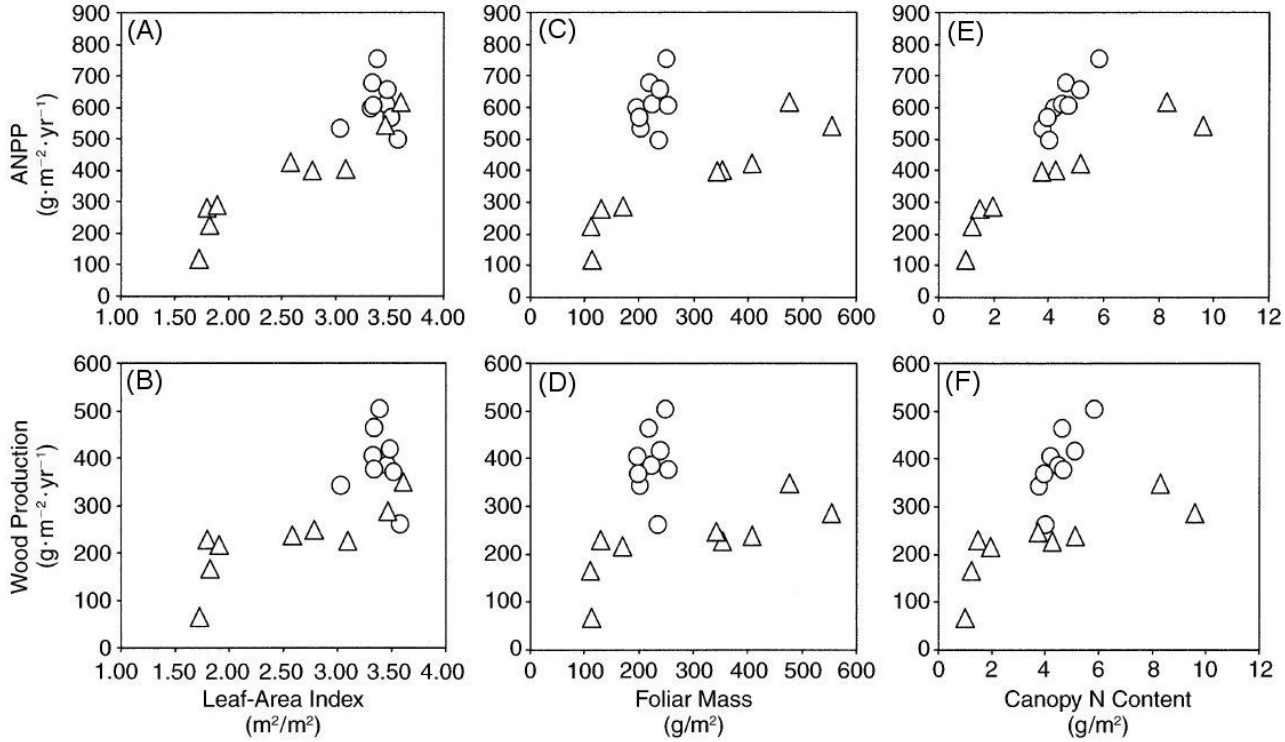


Fig. 1. 3 Relations between forest productivity and different canopy characteristics. Relations between forest productivity (both ANPP and Wood production) and Leaf Area Index (A and B), foliar mass (C e D), canopy N content (E e F). All these variables measured both for broad-leaved deciduous (open circles) and needle-leaved evergreen (open triangles) forest types of White Mountains area, New Hampshire, USA (Smith et al. 2002).

Further studies demonstrate that the most of nitrogen emission of the last century originate from human activities, above all from agricultural fertilizations (Holland et al. 2005).

These emissions and the consequent nitrogen deposition are one of the main causes of several environmental damages: acid rains (although the most damaging effects are mainly caused by S deposition)and , forest decline, N leaching and eutrophication. Actually there are also positive effects of the increase in N depositions, above all a productivity increase, demonstrated also by the greater annual increment commonly observed as a result of long-term fertilization (Hogberg 2007, Hogberg et al. 2006).

Foliar nitrogen is mainly related to forest canopy functional characteristics; recent work across a number of FLUXNET sites (Ollinger et al. 2008) shows also a really good correlation between canopy N concentration and shortwave albedo, as estimated from MODIS imagery (Fig. 1.4) in addition to the one with photosynthetic capacity already shown in Fig. 1.1.

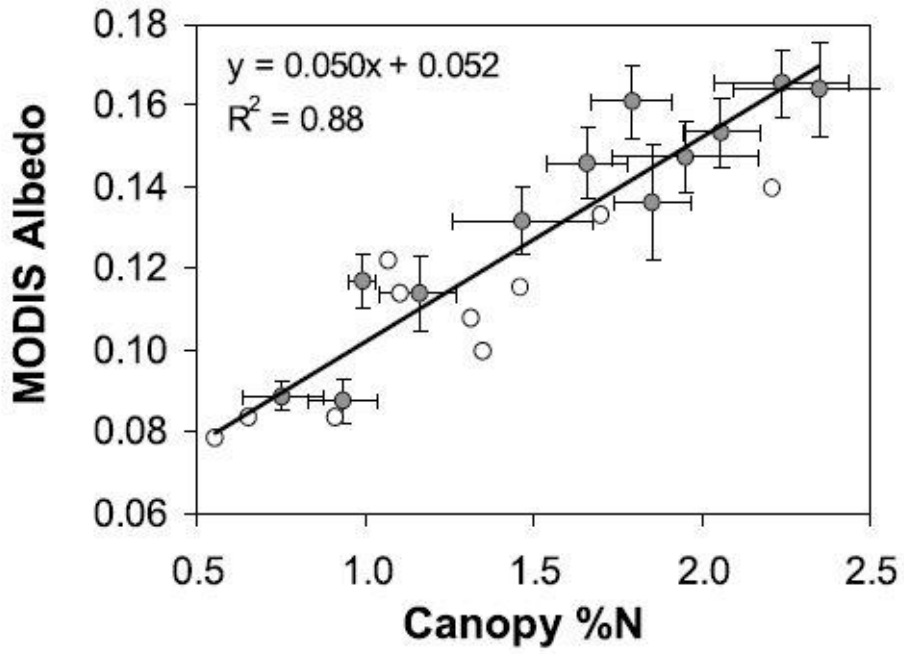


Fig. 1. 4 Relation between forest canopy N concentration and shortwave albedo from MODIS instrument.
N has been measured form FLUXNET eddy covariance towers, while MODIS albedo was calculated for the peak portion of the growing season. Open circles are about Canadian sites from a previous study. Albedo error bars are obtained from standard deviation of all growing season samples between 2000 and 2006 (Ollinger et al. 2008).

If confirmed, this relationship would imply a two-fold effect of canopy N concentration on the global climate, through its effects not only on C exchange but also on the surface energy balance through (canopy albedo). This has been the subject of an intense debate over the last few years, as the relationship between N concentration and albedo has been suggested to be an artifact, mainly related to the parallel but independent effects of forest composition on both variables (Knyazikhin et al., 2012). Evergreen species in particular are characterized by low values of albedo, %N and photosynthetic capacity, whereas broadleaves are characterized by high values. An increase in forest albedo even in the absence of any changes in stand composition has been recently documented, however, in response to experimental long-term N fertilization (Wicklein et al., 2012), although the effects appears to be related to changes in canopy structure rather than in leaf spectroscopic characteristics.

Moreover, also the possibility that the relationship could be mediated by species composition does not subtract to its potential significance, as an increase in N availability (because either of atmospheric N deposition or of increasing decomposition rates and N mineralization in response to global warming) could produce a change in the forest composition toward an higher presence of N-depending species or of plants with an high foliar N content.

1.2. Spectroscopic assessment of N concentration

Traditionally foliar N was analyzed in laboratory through analytical methods, scientifically excellent, but really expensive in terms of both time and cost.

Over the years researchers have developed several spectroscopic methodologies based on different leaves physical state: dried and grounded leaves, dried but not grounded leaves, whole and fresh leaves. The most utilized is the first method, because of the capacity to bypass usual problems associated with water content and the lack of homogeneity in the sample. These two issues mask analysis responses, since the first one causes deep absorption bands (especially in the Short Wave InfraRed, SWIR, region where also N-associated features are often found) typical of water presence and generates a loss of data (Jacquemoud et al. 1996; Kokaly and Clark 1999); the second one causes higher noise in the response because of the scattering due to different-sized particles. The dried and grounded leaves method is the most commonly applied in foliar spectroscopic analysis and gives good results, with a strong correlation between N measured by wet chemistry vs spectroscopic prediction of N (Kokaly and Clark 1999). However this technique presents another problem, as it works really well for similar sites, but not so well in different regions, and therefore needs site-specific calibrations that cause a further increase in time and resources required.

Recently a new method has been developed to reduce the time required without any loss in accuracy. Also this one is based on spectroscopy, but does not measure directly leaf N concentration, but utilizes chlorophyll (as measured by portable instruments, e.g. SPAD (Chang and Robinson 2003)) as proxy to foliar N. In fact chlorophyll is made mostly by N-containing enzymes and other organic compounds, thus it is positively linked to foliar N (Lamb et al. 2002). SPAD is a portable and fast chlorophyll meter that measures the differential transmittance of light through the leaf non-destructively, and is based only on two wavelengths in the red (at 650 nm) and in the NIR (Near InfraRed) region (at 940 nm). The value displayed by SPAD is a non-dimensional index between 0 and 100. Chang and Robinson (Chang and Robinson 2003) verified the existence of relations between SPAD values and foliar N concentration in several hardwood species (*Liquidambar styraciflua* L., *Platanus occidentalis* L., *Populus heterophylla* L. e *Fraxinus pennsylvanica* Marsh.). These relations show that high N concentrations correspond to high SPAD indexes, but these regressions are different depending on the specie involved. Although this method is characterized by several advantages, above all the speed of the analysis, it could be applied only in little areas because of survey time and the large amount of samples.

1.3. Remote sensing assessment of canopy N content and concentration

Over the last few years researchers dealing with N concentration at stand and regional scale have focused their attention on remote sensing techniques, in order to avoid the problem of survey expensiveness, both in time and money. These consist in analysing object characteristics through images from aerial or satellite sensors without physical contact with the object itself. Because fo differences in instrument conformation (satellite/aeroplane, sensor, flight altitude, etc.) a wide variety of imagery exists in terms of spatial, temporal and spectral resolutions, giving to this technique a great flexibility and a huge range of applications.

The physical process is based on solar light, that after scattering processes reaches the object of detection; here it could be absorbed, transmitted and reflected by it. Reflected sunlight is the most common source of information; after interacting with the atmosphere as a result of scattering effects, light is detected by the sensor and is transmitted to a ground station, where the acquisition and corrections are performed. After elaboration, remote sensing products could also be assimilated by mathematical models of ecosystem function, which when applied at regional scale require a huge amount of data that are difficult to obtain from field sample campaigns.

As demonstrated by spectroscopic studies, concerning vegetation and particularly hardwood plants, the most representative wavelengths are IR and visible bands (Gates et al., 1965; Woolley, 1971). Visible leaf reflectance is due to the absorbance by photosynthetic pigments in red and blue wavelengths: chlorophyll b and carotenoids mainly absorb blue light, while chlorophyll a absorbs both blue and red light. On the other hand, green light is less strongly absorbed, but is partly reflected and transmitted; this generates the typical leaf color.

Infrared and above all NIR reflectance allows the discrimination between broadleaves and evergreens, as the first are characterized by an higher reflectance than coniferous species. This is understood to be due to differences in leaf structure: light is reflected where there is the transition between liquid and gas, as in intercellular spaces and cytoplasm. These structures are copious in broadleaves spongy mesophyll, while in conifers are almost missing (Ollinger 2011).

Further increase in NIR reflectance is caused by the amount of leaf layers. In fact these wavelengths are little absorbed, but a large fraction of the light that reaches the top leaves is transmitted to lower canopy layers, where it is largely reflected back, so contributing to canopy reflectance.

Since on the contrary red light absorption increases asymptotically with the number of leaf layers in the canopy (Asner 1998) the comparison of red and NIR reflectance is the basis of LAI determination from a satellite. To estimate it, a normalized index is utilized: NDVI (Normalized Difference Vegetation Index), that is obtained as the ratio between the difference of NIR and red reflectance and their sum. This index is positively related to LAI, and as a result also to total canopy

chlorophyll and N content (Gamon et al., 1995), although the relationship is curvilinear and saturates at high values of LAI, making it less useful for dense canopies, a quasi-linear relationship is commonly observed with the fraction of absorbed radiation (Asrar et al., 1989), which is directly related to canopy GPP.

Remote sensing has been utilized to estimate several forest canopy characteristics, mainly chemical proprieties (Curran 1989; Martin et al. 2008). Several different methods have been applied to estimate N content by remote sensing techniques. Initially they were based on SWIR bands (Martin and Aber, 1997; Wessman et al., 1998; Serrano et al. 2002), but also in remote sensing these wavelengths presented the same problem as in laboratory methodologies, as water absorption bands masked the real vegetation response.

Recent studies (Ollinger et al. 2008) have also suggested the possibility to estimate N concentration from MODIS NIR reflectance bands, with very good determination coefficients across a number of forest sites.

1.4. Considerations and conclusion

Estimating foliar N is useful and necessary to scientific purposes, but also for the definition of sound environmental policies at regional and global scale. As discussed, canopy N is directly involved in several ecological processes that need to be quantified for the calculation of forest C sink, as required in Kyoto Protocol, and possibly in other biophysical processes central to the climate change debate. The availability of methods to estimate foliar N at regional to global scale is becoming urgent, as the parameter is required as an input for biogeochemical and climate models.

Remote sensing could help in this research because of its speed, but unfortunately focused techniques are not available. All the methodologies cited have to face with problems, some associated with spectral characteristics and overlaps, as those about SWIR bands, and some concerning the technical characteristics and limitations of available instrumentation. To obtain a good estimate multi- or hyperspectral satellite sensors are needed, together with studies to determine the most suitable bands for estimating N.

This measure is not so easy, as all available techniques have their limitations.

The recently reported correlation with canopy NIR reflectance would open very interesting opportunities. It is not yet clear, however, if the signal obtained from remote sensing is directly related to leaf N content or is mainly determined by leaf or canopy structure.

Recent studies (Wright et al. 2004) have demonstrated a general relationship between leaf N and LMA (Leaf Mass per Area) which is itself correlated to leaf thickness; leaf thickness could be

related to the amount of intercellular spaces, which are seen as principal factor that drives NIR reflectance.

Also other studies (Ollinger 2011, Knyazikhin et al. 2012) highlight the difficult in understanding reflectance in NIR region, because of the presence of multiple combinations of independent proprieties.

To better know these processes and to understand the cause of the responses obtained from remote sensing it's necessary to develop new studies to distinguish N from LMA effects.

Concerning SPAD methodology, the relation between chlorophyll and foliar N has not been demonstrated to be true for all forest species. A similar problem also affects the albedo technique, in which the relation is empirically based, so that it could be necessary to investigate if the relation is due to species diversity or foliar N diversity.

In the end, all these methodologies need further research, as it is necessary to test these relations both for different species and in different sites. For this reason, to estimate N from remote sensing, it is useful to study areas with a good species diversity, so as to increment the range of N content values explored and obtain a solid and general relationship.

1.5. References

Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment 64, 234-253.

Chang, S. X. and Robinson, D. J. (2003). Nondestructive and rapid estimation of hardwood foliar nitrogen status using the SPAD-502 chlorophyll meter. For. Ecol. Manag. 181, 331-338.

Curran, P. J. (1989). Remote sensing of foliar chemistry. Remote Sensing of Environment 30, 271-278.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., and Vorosmarty, C. J. (2004). Nitrogen cycles: past, present, and future. Biogeochemistry 70, 153-226.

Gamon, J. a., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., Peñuelas, J., et al. (1995). Relationships between NDVI canopy structure and photosynthesis in three Californian vegetation types.pdf. *Ecological Applications*, *4*(1), 11–20.

Gates, D. M., Keegan, H. J., Schleter, J. C., & Weidner, V. R. (1965). Spectral Properties of Plants. *Applied Optics*, *4*(1), 11. doi:10.1364/AO.4.000011

Hogberg, P. (2007). Nitrogen impacts on forest carbon. Nature, 447, 781–782.

Hogberg, P., Fan, H., Quist, M., Binkley, D., & Tamm, C. O. (2006). Tree growth and soil acidification in response to 30 years of experimental nitrogen loading on boreal forest. Global Change Biology, 12(3), 489–499. doi:10.1111/j.1365-2486.2006.01102.x

Holland, E.A., Lee-Taylor, J., Nevison, C., and Sulzman, J. (2005). Global N Cycle: Fluxes and N_2O Mixing Ratios Originating from Human Activity. Data set. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G., and Hosgood, B. (1996). Estimating leaf biochemistry using the PROSPECT leaf optical properties model. Remote Sensing of Environment 56, 194-202.

Kokaly, R. F., Asner, G. P., Ollinger, S. V., Martin, M. E., & Wessman, C. a. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment*, *113*, S78–S91. doi:10.1016/j.rse.2008.10.018

Kokaly, R. F. and Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. Remote Sensing of Environment 67, 267-287.

Knyazikhin, Y., Schull, M. a, Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., et al. (2012). Hyperspectral remote sensing of foliar nitrogen content. Proceedings of the National Academy of Sciences of the United States of America, 1–8. doi:10.1073/pnas.1210196109

Lamb, D. W., Steyn-Ross, M., Schaare, P., Hanna, M. M., Silvester, W., and Steyn-Ross, A. (2002). Estimating leaf nitrogen concentration in ryegrass (*Lolium* spp.) pasture using chlorophyll red-edge: theoretical modelling and experimental observations. International Journal of Remote Sensing 23, 3619-3648.

Martin, M. E., Plourde, L. C., Ollinger, S. V., Smith, M. L., and Mcneil, B. E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. Remote Sensing of Environment 112, 3511-3519.

Martin, M., & Aber, J. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications*, *7*(2), 431–443.

Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. The New phytologist, 189(2), 375–94. doi:10.1111/j.1469-8137.2010.03536.x

Ollinger, S. V., Richardson, A. D., Martin, M. E., Hollinger, D. Y., Frolking, S., Reich, P. B., Plourde, L. C., Katul, G. G., Munger, J. W., Oren, R., Smith, M. L., Paw U, K. T., Bolstad, P. V., Cook, B. D., Day, M. C., Martin, T. A., Monson, R. K., and Schmid, H. P. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. PNAS 105, 19335-19340.

Ollinger, S. V., Smith, M. L., Martin, M. E., Hallett, R. A., Goodale, C. L., and Aber, J. D. (2002). Regional variation in foliar chemistry and N cycling among forests of diverse history and composition. Ecology 83, 339-355.

Peterson, A. G., Ball, J. T., Luo, Y. Q., Field, C. B., Reich, P. B., Curtis, P. S., Griffin, K. L., Gunderson, C. A., Norby, R. J., Tissue, D. T., Forstreuter, M., Rey, A., and Vogel, C. S. (1999). The photosynthesis leaf nitrogen relationship at ambient and elevated atmospheric carbon dioxide: a meta-analysis. Global Change Biol 5, 331-346.

Serrano, L., Peñuelas, J., and Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation form AVIRIS data: Decomposing biochemical from structural signals. Remote Sensing of Environment 82, 355-364.

Smith, M. L., Ollinger, S. V., Martin, M. E., Aber, J. D., Hallett, R. A., and Goodale, C. L. (2002). Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecological Applications 12, 1286-1302.

Wessman, C., Aber, J., Peterson, D., & Melillo, J. (1988). Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature*, *335*, 154–156. Retrieved from http://www.nature.com/nature/journal/v335/n6186/abs/335154a0.html

Woolley, J. T. (1971). Reflectance and transmittance of light by leaves. *Plant physiology*, *47*(5), 656–62. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=396745&tool=pmcentrez&rendertype=abstract

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M. L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., and Villar, R. (2004). The worldwide leaf economics spectrum. Nature 428, 821-827.

Wicklein, H. F., Ollinger, S. V, Martin, M. E., Hollinger, D. Y., Lepine, L. C., Day, M. C., Bartlett, M. K., et al. (2012). Variation in foliar nitrogen and albedo in response to nitrogen fertilization and elevated CO(2). *Oecologia*, (2008). doi:10.1007/s00442-012-2263-6.

2. Remote-sensing of foliar nitrogen concentration in coastal ecosystems. (i) San Rossore 2000

Published as Mezzini, E., Raddi, S., Pippi, I., & Magnani, F. (2010). Telerilevamento della concentrazione di azoto fogliare in ecosistemi forestali litoranei. Atti XIV Conferenza Nazionale ASITA – Brescia 9-12 novembre 2010, pp. 1335–1338. Available from http://www.attiasita.it/ASITA2010/Pdf/223.pdf

2.1. Introduction

Foliar nitrogen concentration is one of the most important parameters in the study of forest growth and eco-physiology. Nitrogen is a regulator of photosynthesis and foliar respiration; moreover it is correlated with soil and forest canopy characteristics (Ollinger, 2010;Ollinger et al., 2002;Smith et al., 2002). In particular the correlation between N and Net Primary Production (NPP) is really interesting due to the effects of foliar N concentration on photosynthetic light use efficiency. Also nitrogen temporal and spatial variations are full of information. These variations are influenced by area characteristics and by disturbances, like climatic change, nitrogen deposition and increase in atmospheric CO_2.

Foliar nitrogen measurements are customarily made by destructive techniques, which are however very expensive, both in terms of time and costs. Therefore in the last decades remote sensing techniques have become increasingly common for the analysis at landscape and regional scale.

These new techniques involve two main areas of the electromagnetic spectrum: SWIR and NIR bands. The first one is the traditional technique, but it has some problems mainly due to the presence of water absorption features that generate a loss of information. Recently some studies suggest the possibility of estimating forest foliar N concentration by the means of canopy NIR reflectance (Ollinger, 2010;Ollinger et al., 2008). This possibility could be due to the link between foliar nitrogen concentration and foliar and canopy structure and to the effect of the structure on foliar and canopy reflectance (Hollinger et al., 2009;Jacquemoud et al., 1996). A validation of these relations could open new prospects in the remote sensing of forest canopy nitrogen.

The aim of the present study is to test these new techniques in a different ecosystem. In particular, previous studies focused on a rather homogeneous and extensive ecosystem (boreal forests of North America) and analyzed nitrogen concentrations by means of MODIS NIR reflectance and indexes.

In this study we analyze a coastal forest ecosystem, characterized by its high heterogeneity with a small patch scale. In this case MODIS is not the best instrument, because of its rather coarse spatial

resolution (250-1000 m, depending on the band considered); moreover we would like to drastically decrease the cost of image acquisition by means of different sensors.

2.2. Materials and methods

As already mentioned in the introduction, MODIS spatial resolution (with a pixel footprint 250x250 m^2 in the red and near-infrared) is not suitable for the high spatial variability commonly found in the Italian territory. Italy is characterized by high landscape diversity even over small areas. This Italian ecosystems characteristic determine our choice of remote-sensing sensor.

The study area of San Rossore Natural Park (Pisa, Italy) is characterized by Mediterranean climate and vegetation. It's covered by extended woods both of pine plantations (*Pinus pinea* and *Pinus pinaster*) and mixed hardwoods (*Quercus robur*, *Acer campestre*, *Carpinus betulus*, *Populus alba*, *Fraxinus oxycarpa*, *Alnus glutinosa*).

2.2.1. Field data collection

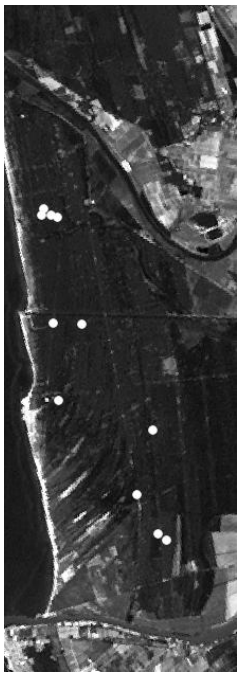


Fig. 2. 1 Landsat 7 ETM+ image with field sample areas.

To cover the main forest types of the Natural Park and the widest possible range of forest structures, eleven stands belonging to three chronosequences were chosen (4 stands of *Pinus pinea*, 4 stands of *P. pinaster* and 3 stands of hardwood species) (Fig. 2.1). For each stand, various biophysical parameters were measured: stand density, basal area, green canopy depth, height and LAI. Sample plots are composite-samples that are made of five points per stand, located at the vertices and at the centre of a cross, at a distance of 30 m from each other. At each point, samples were collected with a shotgun from the upper portion of the canopy. Nitrogen analyses were carried out in the laboratory by elemental analyzer CHNS-O (EA 1110, Thermo Fisher, Milan, Italy). Field data were collected within a few days of the MIVIS flight.

2.2.2. MIVIS dataset

The dataset used in this work was acquired on 21st of June 2000 by MIVIS (Multispectral Infrared and Visible Imaging Spectrometer), mounted on a Casa 212 aircraft. The flight was performed at an altitude of 1500 m, in south to north direction.

The MIVIS sensor is characterized by 102 channels (92 in the visible and near infra-red wavelength and 10 channels in the thermal infra-red), its NIR channels spectral coverage is from 813 to 833 nm (from ch. 17 to ch. 20) with 20 nm FWHM (Full-Width Half-Maximum). It's also characterized by

a spatial resolution (pixel footprint) of about 2.5 m from an altitude of 1500 m, with an IFOV (Instantaneous Field of View) of 2 mrad.

MIVIS characteristics make it more suitable for Italian territory analysis than MODIS.

As a first step of the analysis, MIVIS frames were geometrically and atmospherically corrected by CNR-IFAC (Barducci and Pippi, 2001) using the MODTRAN 6 software.

Then the best frame was georeferenced by the ENVI software (Exelis, Inc. Visual Information Solutions, McLean, Virginia). The co-registration was based on vector-graphics CTR (Carta Tecnica Regionale – Regional Technical Map) 1:10000 of Tuscany, UTM, zone 32 N, WGS84. In total, 100 GCPs (Ground Control Points) were selected, obtaining a total RMS (Root Mean Square) error of 16.40 m with a first degree polynomial warp methodology. The triangulation warp method was eventually applied, because of the high deformation of the frame, followed by a nearest neighbour resampling.

These same data had already been analyzed (Bernasconi et al., 2001); in that study authors had calculated the coefficients of determination R^2 (Table 2.1) for different wavelength combinations and different kind of indices (simple ratio and normalized index). They also highlighted a very good correlation between N concentration and normalized index [Eq. 1] or simple ratio [Eq. 2] and between LAI and normalized index [Eq. 3] (green peak normalized with chlorophyll) or simple ratio [Eq. 4].

$$NI_N = \frac{\rho_{2243} - \rho_{2048}}{\rho_{2243} + \rho_{2048}} \qquad [1]$$

$$SR_N = \frac{\rho_{820}}{\rho_{540}} \qquad [2]$$

$$NI_{LAI} = \frac{\rho_{640} - \rho_{660}}{\rho_{640} + \rho_{660}} \qquad [3]$$

$$SR_{LAI} = \frac{\rho_{675}}{\rho_{550}} \qquad [4]$$

Based on this preliminary work, the value of both normalized indices for N and LAI was computed by ENVI Band Math.

Table 2. 1 R^2 coefficients of normalized indexes and simple ratios against N and LAI.

	R^2 with N	**R^2 with LAI**
NI_N	0.995	
SR_N	0.97	
NI_{LAI}		0.9
SR_{LAI}		0.86

2.2.3. Landsat 7 ETM+ dataset

To test the proposed correlation between canopy N concentration and canopy NIR reflectance (Ollinger et al., 2008) on San Rossore Reserve, it could be possible to use MODIS, MIVIS or other satellite images. The most suitable satellite sensor for this kind of study is Landsat 7 ETM+. This sensor is characterized by a spatial resolution of about 30 m (pixel footprint); moreover it has 5 channels in visible and NIR wavelengths, 1 channel in SWIR and 1 channel in thermal infra-red (TIR) with a lower spatial resolution (60 m); finally, it has also a panchromatic image product with a higher spatial resolution (15 m). So Landsat spectral and spatial characteristics can be related and compared to MODIS ones, and also to MIVIS ones; moreover Landsat frames are free and more appropriate to Italian reality.

Fortunately Landsat flew over San Rossore on the 20th of June 2000, so it could produce a frame comparable to the one available for MIVIS.

To compare these two sources, it was necessary to correct and georeference also the Landsat image.

As a first step, the Landsat image was co-registered with MIVIS frame by means of GCPs and nearest neighbour resampling. The image was then stacked over MIVIS frame.

Consistent with standard practice (Chander et al., 2009) we use a three-steps approach to correct ETM+ sensor frame:

a) From calibrated digital number (Q_{cal}) to radiance (L_λ);

b) From L_λ to planetary reflectance (ρ_λ) Top of Atmosphere (TOA);

c) From ρ_λ TOA to ρ_λ Top of Canopy (TOC).

Conversion to at-sensor spectral radiance (L_λ). Spectral radiance was computed from calibrated digital numbers as:

$$L_\lambda = G_\lambda \cdot Q_{cal} + B_\lambda \qquad [5]$$

where

L_λ= Spectral radiance at the sensor's aperture [W/(m^2 sr μm)]

Q_{cal}= Quantized calibrated pixel value [DN]

G_λ= Band-specific rescaling gain factor [(W/(m^2 sr μm))/DN]

B_λ= Band-specific rescaling bias factor (offset) [W/(m^2 sr μm)]

This conversion was carried out using the ENVI pre-processing tool, Apply Gain and Offset.

Gain and offset values are provided by (Chander et al., 2009).

Conversion to TOA reflectance, Top-of-atmosphere (TOA) reflectance at each wavelength was computed as:

$$\rho_\lambda = \frac{\pi \cdot d^2}{ESUN_\lambda \cdot \cos\theta_s} \cdot L_\lambda \qquad [6]$$

where

ρ_λ= Planetary TOA reflectance [unitless]

d= Earth-Sun distance [astronomical units]

$ESUN_\lambda$= Mean exoatmospheric solar irradiance [W/ (m^2 μm)]

θ_s= Solar zenith angle [degrees]

Also in this case the equation was computed by ENVI pre-processing tool, Apply Gain and Offset. In this case, the offset would be null and the gain would be equal to the first multiply factor.

Earth-Sun distance and solar irradiance values were derived from the literature (Chander et al., 2009), whereas solar zenith angle was calculated from sun elevation at the time of the overpass.

Conversion to TOC reflectance.

Top-of-canopy (TOC) reflectance was computed by the dark pixel subtraction technique. Using the ENVI pre-processing Dark Subtraction routine, the NIR reflectance of a water region of interest (ROI) was subtracted from the entire image; oligotrophic waters are characterized by a null reflectance in the NIR band of interest for the present study. This operation was applied automatically to the whole frame.

2.2.4. ROIs sample

In order to be sure to analyze the correct tile of the frame, we imported in ENVI software the location of the georeferenced stands. Three different kinds of ROIs were then created. The first level was a one-pixel (30 m, like a Landsat pixel) ROI centred into the point that represent the stand. The second level was a 3x3 pixel ROI square centred into the stand point, paying attention to the presence of streets or other atrophic elements. In the last level, each ROI covered an entire forest plot, as derived from the 2009 Forest Plan of the San Rossore Natural Park, where the plot

was the one containing the stand. In MIVIS analyses the number of sample decreases from 11 to 10 stands because of the dimension of the frame mainly due to a different IFOV of the sensor (2 mrad, FOV 71.06 °).

2.2.5. Statistical analysis

Thanks to these three levels of ROIs we could apply the statistic tool to each ROI and then analyze them by OLS (Ordinary Least Squares) in Excel. The ENVI statistic tool provided for each ROI minimum and maximum values, mean value and standard deviation value for each wavelength and each kind of sensor. These results were also plotted as a spectral plot or as a scatter plot.

2.3. Results and discussion

The procedure by which it was possible to verify that Landsat ETM+ could be used as an estimator of forest canopy nitrogen concentration was quite complex. MODIS reflectances could not be compared directly to field measurements of canopy N concentration, because of the spatial resolution of the sensor and the heterogeneity of the scene. Instead, the applicability of the approach was tested in three related steps, diagrammatically presented in Fig.2.2.

Table 2. 2 Normalized indexes based on MIVIS reflectance, and regression coefficients against nitrogen concentration at the 10 ROIs in the San Rossore Natural Park. Based on Bernasconi et al. (2001).

Normalized index	**R^2**
(760-620)/(760+620)	0.9015
(780-620)/(780+620)	0.9089
(800-620)/(800+620)	0.9057
(820-620)/(820+620)	0.8993
(1175-620)/(1175+620)	0.9296
(1175-680)/(1175+680)	0.9067
(2243-2048)/(2243+2048)	0.8992

As a first step, the correlation between leaf N concentrations and canopy NIR reflectance was tested using MIVIS channel 20, which most closely matches the MODIS band proposed by Ollinger et al. (2008) as the best candidate (Fig. 2.2, step 1). A rather good coefficient of determination (R^2 = 0.5197) was observed between MIVIS NIR reflectance (channel n. 20) and field values of canopy nitrogen concentration. This is consistent with results previously obtained with a wide set of normalized indexes (Bernasconi et al., 2001), as many of those showing the best regression coefficients with nitrogen concentration were based on wavelengths in the NIR region (Table 2.2).

In a more recent study (Ollinger, 2010) a novel Difference Vegetation Index (DVI) [Eq. 7] is also presented as a good estimator of nitrogen concentration.

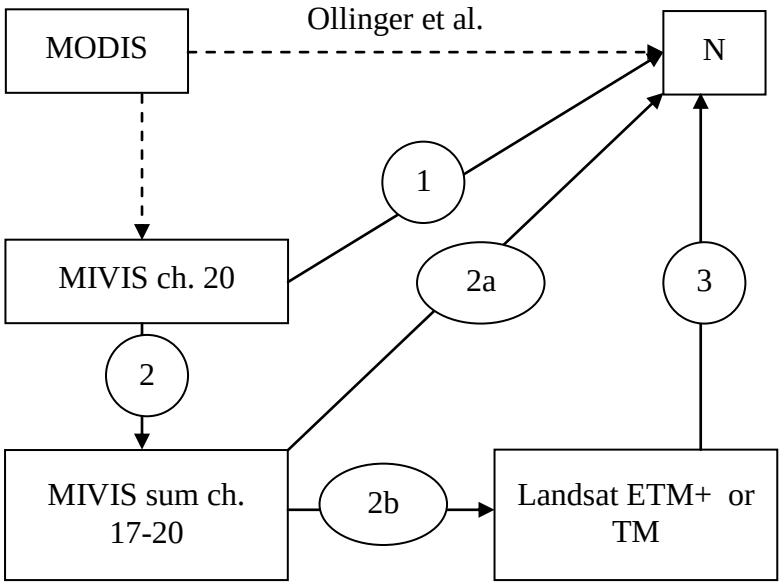


Fig. 2. 2 Diagram of the whole procedure. Numbers refer to the individual steps of the analysis, as described in the text.

$$DVI = \rho_{810} - \frac{\rho_{610} + \rho_{661}}{2} \qquad [7]$$

The bands involved were not available in MIVIS, so this index was changed into other three possible indexes (Table 2.3) which generated different regression coefficients against nitrogen. The DVI that involves the bands at 800, 620 and 680 nm, which match most closely the original bands, seemed to be the best.

As a second step, the relation between channel n. 20 of MIVIS and the mean reflectance of MIVIS channels 17,18,19,20 (i.e. all MIVIS NIR bands) was verified (Fig. 2.2, step 2). This step was important because the sum of these channels broadly covers the spectral region corresponding to the band n. 4 of Landsat ETM+ (NIR). The narrow- and broad-band NIR reflectances were found to be very strongly correlated, with a R^2 of 0.9995. As a result, a similar relationship was found between canopy N concentration and MIVIS broadband NIR reflectance (Fig. 2.2, step 2a) as the one already observed with narrowband reflectance, yielding at a R^2 of 0.5063.

The relationship between MIVIS (broadband: sum of channels n. 17-20 or narrowband: channel n. 20) and Landsat 7 ETM+ NIR reflectances was also tested (Fig. 2.2, step 2b). This analysis highlighted that the relation between channel n. 20 and NIR band of Landsat was stronger than the one that involved the sum of all the NIR channels.

Table 2. 3 Difference Vegetation Indexes and regression coefficients against nitrogen concentration.

Difference Vegetation Index	**R^2**
800-[(620+680)/2]	0.6995
820-[(620+680)/2]	0.6783
[(800+820)/2]-[(620+680)/2]	0.6881

The last step of the whole procedure involved the comparison of Landsat and nitrogen field analyses (Fig. 2.2, step 3). The relation between N concentration and Landsat data could be demonstrated both by Landsat 7 ETM+ NIR TOC (Top of Canopy) reflectance and Landsat 7 ETM+ DVI, defined as (Coops and Waring, 2001):

$$DVI = \rho_{835} - \rho_{662} \qquad [8]$$

The regression between simple TOC NIR reflectance and N resulted in a R^2 of 0.9314, while the one between DVI and N gave us a regression coefficient of 0.9519 (Fig. 2.3). Both results are far

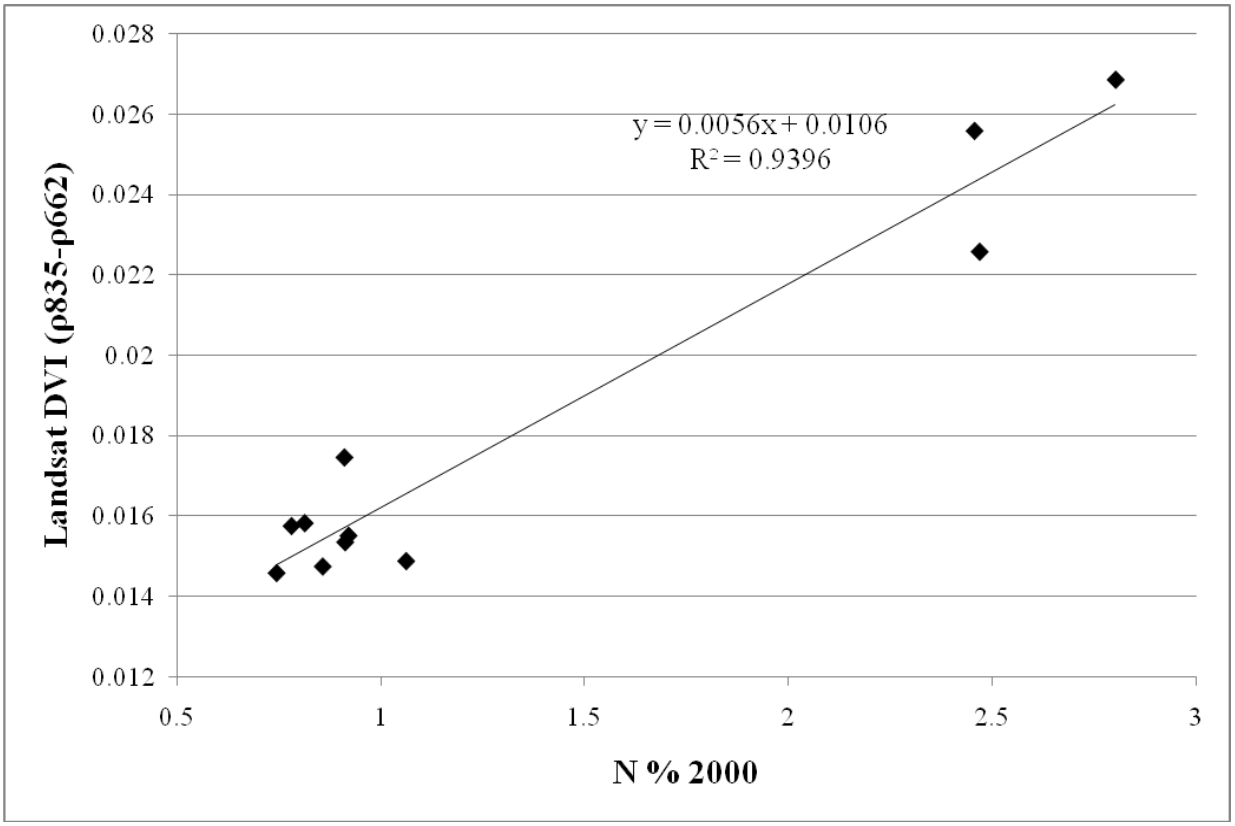


Fig. 2. 3 Chart of the regression between foliar N concentration of the 11 samples and ROIs DVI calculated from Landsat 7 ETM+ image.

superior to the ones obtained with narrowband MIVIS reflectances and indices.

As already mentioned in field sampling paragraph, MIVIS give us some problems. We tried to identify the cause, and the main problem that surfaced was likely due to the acquisition geometry of the sensor, which like most airborne sensors is characterized by a rather wide field-of-view (FOV = 71.06°), and by the reflectance anisotropy of vegetation. Vegetation reflectance is characterized by considerable anisotropy, i.e. it changes depending on the viewing angle (and up to a point on solar angle), with a bi-directional reflectance distribution function (BRDF) which depends on the wavelength under consideration. The shape of the BRDF is known to be largely a function of canopy structure (Asner, 2000), and can be therefore exploited in order to gain additional information on canopy biophysical characteristics (Raddi et al., 2005; Stagakis et al., 2010). Reflectance anisotropy is a well known issue in the analysis of airborne imagery, because of the across-track brightness gradient resulting from the wide viewing angle of airborne sensors (Schiefer et al., 2006); the problem is reduced when working with normalized indices based on close wavelengths, but introduces a further source of error when dealing with reflectance itself. Images can be partly corrected for the effects of reflectance anisotropy through the application of semi-empirical BRDF models (Hikler et al., 2008), but this is beyond the scope of the current study. Although the images used in the current analysis were acquired with a flight path aligned at a small angle with the sun, reflectance anisotropy would appear to make it impossible to make use of reflectance (or reflectance difference) data acquired from an aerial platform. Therefore we made a secondary analysis, by which we assessed the relevance in regression against N concentration of ROIs row and column position in the MIVIS frame. This procedure was applied through a stepwise analysis, looking for a multiple regression and including row and column number together with canopy N concentration as possible explanatory variables of MIVIS reflectance. At the end of this step we found that in ROIs which were near the edge of the image the regression between reflectance (or indexes) and N was widely affected by the position of the ROI in the frame (Table 2.4). Considering MIVIS normalized indexes, they were clearly not dependent from the column position but only from the row position. On the contrary, MIVIS DVI and raw TOC reflectance were influenced both by row and column.

The same kind of analysis was carried out on the regression between MIVIS NIR reflectance and Landsat NIR reflectance, proving MIVIS to be dependent on Landsat reflectance and column position, but independent from row position (Table 2.5).

Table 2. 4 Correlation indexes: MIVIS normalized indexes, MIVIS DVI, MIVIS reflectance in relation with N concentration, N % and ROI position in MIVIS frame. n/s = no significant.

Independent variables / **Dependent variables**	N%	N% and column	N% and column and row	N% and row
Normalized index				
(760-620)/(760+620)	0.901	n/s	n/s	0.945
(780-620)/(780+620)	0.909	n/s	n/s	0.945
(800-620)/(800+620)	0.906	n/s	n/s	0.945
(820-620)/(820+620)	0.899	n/s	n/s	0.948
(1175-620)/(1175+620)	0.93	n/s	n/s	0.965
(1175-680)/(1175+680)	0.907	n/s	n/s	n/s
(2243-2048)/(2243+2048)	0.899	n/s	n/s	n/s
DVI				
800-[(620+680)/2]	0.7	0.902	0.96	n/s
Reflectance				
Sum of ch. 17-20	0.506	0.869	0.934	n/s
Ch. 20	0.52	0.864	0.933	n/s

Table 2. 5 Correlation indexes: MIVIS NIR reflectance in relation with Landsat NIR reflectance and ROI position in MIVIS frame. n/s = no significant.

Independent variables / **Dependent variables**	Landsat band 4	Landsat band 4 and column	Landsat band 4 and column and row	Landsat band 4 and row
MIVIS sum of ch. 17-20	0.683	0.972	n/s	n/s
MIVIS ch. 20	0.668	0.974	n/s	n/s

2.4. Conclusions

The results of this study are promising, especially the regression between Landsat DVI and nitrogen concentration. The observed good correlation between Landsat DVI or NIR TOC reflectance and

nitrogen concentration could open new outlooks in the remote sensing of forest canopy nitrogen, also due to the simple techniques and little cost associated. Obviously it is important to keep on testing these new techniques. But both this regression and the one between Landsat NIR TOC reflectance and N concentration are characterized by a split-up in two groups, easy identifiable as coniferous and hardwood forests, with a strong leverage effect on the relationship. Probably this phenomenon is mainly due to the small sample size available for the analysis. To try to solve this problem we are pursuing new studies which are involving 33 stands to increase the diversity of our samples. Moreover these researches will utilize also other sensors, mainly airborne CASI and Landsat 5 TM satellite imagery.

The use of images from the CASI sensor, with a smaller FOV=38.4° across-track, will make it possible to test the relevance of the problem discussed above. A validation of this thesis could really change ideas related to the utilization of airborne sensors.

2.5. Acknowledgements

This study was possible thanks to the economic support from the EU FP7 project GHG-Europe. We also thank all the CNR- IFAC team for their help and preliminary analyses.

2.6. References

1. Asner, G. P. 2000. Contributions of multi-view angle remote sensing to land-surface and biogeochemical research. Remote Sensing Reviews, 18(2-4):137–162. doi:10.1080/02757250009532388

2. Barducci, A. and I. Pippi. 2001. Analysis and rejection of systematic disturbances in hyperspectral remotely sensed images of the Earth. Applied Optics 40:1464-1477.

3. Bernasconi, L., Pippi, I., and Raddi, S. Predicting stand structure in mediterranean conifer and hardwoods by airborne hyperspectral imaging spectrometers. 2001. Collecting and analyzing information for sustainable forest management and biodiversity monitoring with special reference to Mediterranean ecosystems. European Commission.

4. Chander, G., B. L. Markham, and D. L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of Environment 113:893-903.

5. Coops, N. C. and R. H. Waring. 2001. The use of multiscale remote sensing imagery to derive regional estimates of forest growth capacity using 3-PGS. Remote Sensing of Environment 75:324-334.

6. Hilker, T., Coops, N. C., Hall, F. G., Black, T. A., Wulder, M. a., Nesic, Z., & Krishnan, P. 2008. Separating physiologically and directionally induced changes in PRI using BRDF models. Remote Sensing of Environment, 112(6):2777–2788. doi:10.1016/j.rse.2008.01.011

7. Hollinger, D. Y., S. V. Ollinger, A. D. Richardson, T. P. Meyers, D. B. Dail, M. E. Martin, N. A. Scott, T. J. Arkebauer, D. D. Baldocchi, K. L. Clark, P. S. Curtis, K. J. Davis, A. R. Desai, D. Dragoni, M. L. Goulden, L. Gu, G. G. Katul, S. G. Pallardy, K. T. Paw, H. P. Schmid, P. C. Stoy, A. E. Suyker, and S. B. Verma. 2009. Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. Global Change Biol 16:696-710.

8. Jacquemoud, S., S. L. Ustin, J. Verdebout, G. Schmuck, G. Andreoli, and B. Hosgood. 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. Remote Sensing of Environment 56:194-202.

9. Ollinger, S. V. 2010. Sources of variability in canopy reflectance and the convergent properties of plants. New Phytologist1-20.

10. Ollinger, S. V., A. D. Richardson, M. E. Martin, D. Y. Hollinger, S. Frolking, P. B. Reich, L. C. Plourde, G. G. Katul, J. W. Munger, R. Oren, M. L. Smith, K. T. Paw U, P. V. Bolstad, B. D. Cook, M. C. Day, T. A. Martin, R. K. Monson, and H. P. Schmid. 2008. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. PNAS 105:19335-19340.

11. Ollinger, S. V., M. L. Smith, M. E. Martin, R. A. Hallett, C. L. Goodale, and J. D. Aber. 2002. Regional variation in foliar chemistry and N cycling among forests of diverse history and composition. Ecology 83:339-355.

12. Raddi, S., Cortes, S., Pippi, I., & Magnani, F. 2005. Estimation of vegetation photochemical processes: an application of the photochemical reflectance index. Proceedings of the 3rd ESA CHRIS PROBA Workshop. Esrin 21-23 March 2005. Esrin.

13. Schiefer, S., P. Hostert, and A. Damm. 2006. Correcting brightness gradients in hyperspectral data from urban areas. Remote Sensing of Environment 101:25-37.

14. Smith, M. L., S. V. Ollinger, M. E. Martin, J. D. Aber, R. A. Hallett, and C. L. Goodale. 2002. Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecological Applications 12:1286-1302.

15. Stagakis, S., Markos, N., Sykioti, O., & Kyparissis, A. 2010. Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. Remote Sensing of Environment, 114(5):977–994. doi:10.1016/j.rse.2009.12.006

3. Remote-sensing of foliar nitrogen concentration in coastal ecosystems. (ii) San Rossore 2009

This article is being submitted to the ISI journal European Journal of Remote Sensing.

3.1. Introduction

As well known, the foliar nitrogen concentration is one of the most important parameters in the study of forest growth and eco-physiology (Heimann & Reichstein, 2008; LeBauer & Treseder, 2008; Magnani et al., 2007; Serrano, Peņuelas, & Ustin, 2002). Its interactions with carbon at foliar level is a fundamental mechanism controlling terrestrial carbon cycle (Wright et al., 2004). Besides being a regulator of photosynthesis and foliar respiration, nitrogen is correlated to forest canopy and soil characteristics (Ollinger, 2011; Ollinger et al., 2002; Smith et al., 2002). Of particular interest is the correlation between N and Net Primary Production (NPP) (LeBauer & Treseder, 2008), resulting from the effect of leaf N concentration on light use efficiency. As a result, atmospheric N deposition are understood play a major role in the evolution of NPP observed over recent decades, since an increase in N availability is known to result in a stimulation of NPP in many boreal and temperate forests, delaying the accumulation of CO_2 in the atmosphere (LeBauer & Treseder, 2008).

For all these reasons, in the years several methodologies had been developed to measure or estimate N concentration in forests. Some of the methods in current use are based on laboratory analyses (Kokaly, 1999), while others are based on field spectroscopy (Chang, 2003; Johnson, 2001) and on remote sensing techniques (Ollinger et al., 2008). Because of its convenience in terms both of time and costs, and its applicability at regional to global scale, the last category of techniques has attracted considerable attention over recent years (M. Martin & Aber, 1997; M. E. Martin, Plourde, Ollinger, Smith, & McNeil, 2008; Wessman, Aber, Peterson, & Melillo, 1988). In particular, the possibility that foliar N concentration could be estimated from satellite remote sensing of canopy reflectance at NIR (Near InfraRed) wavelengths has been recently suggested (Ollinger et al., 2008). Whatever the biochemical or biophysical basis of the relationship (Knyazikhin et al., 2012; Ollinger, 2011), a validation of this methodology could provide new opportunities in N analyses.

We had already tested the relation between N concentration and NIR reflectances through different sensors, both airborne and satellite, and our previous work (Mezzini, Raddi, Pippi, & Magnani, 2010) highlighted some unresolved problems, mainly related to viewing geometry, sample size and the strong effects of forest type (broadleaf vs coniferous) on both leaf biochemistry and reflectance.

Thus we decided to exploit a new larger database recently acquired over the same area already explored, in order to extend the initial field sample database (from 11 to 30 ground samples), so avoiding the strong grouping previously observed in the relationship between N concentration and Landsat NIR reflectance, and demonstrate the possibility of estimating N concentration by satellite remote sensing techniques. Moreover, the use in the new study of the CASI hyperspectral airborne sensor, with a smaller field-of-view (FOV) than in the MIVIS sensor used in the earlier analysis, will make it possible to avoid the effects of reflectance anisotropy and view geometry (Mezzini et al., 2010), which appeared in the previous study to hinder the utilization of airborne sensors.

3.2. Materials and methods

As already discussed in our previous work (Mezzini et al., 2010) MODIS spatial resolution is not suited for the high spatial variability which is typical of the Italian territory. This characteristic of Italian ecosystems and landscapes determines our choice in remote-sensing sensor.

Our study area is locate in San Rossore Natural Park (Pisa, Italy) and is characterized by Mediterranean climate and vegetation. This site is largely covered by a forest cover comprising pine plantations (*Pinus pinea* and *P. pinaster*), mixed hardwoods (*Quercus robur*, *Acer campestre*, *Carpinus betulus*, *Populus alba*, *Fraxinus oxycarpa*, *Alnus glutinosa*) and evergreen hardwoods, especially *Quercus ilex*. Other land-use types are also found in the San Rossore field site, including agricultural, shrubland and natural vegetation.

The procedure developed for the current study utilized three different kinds of data: field sample data, aerial images and satellite images (Fig. 3.1). The study was carried out both in Italy and in New Hampshire (University of New Hampshire, USA) during an extended study period.

3.2.1. Field data collection

In order to extend the 2000 database already used in the previous study, we utilized a newer database collected in July 2009 as part of a campaign funded by the European Space Agency (ESA). The dataset consists of 30 multiple sample plots distributed over the whole area and covering the main forest and vegetation types of the San Rossore Natural Park (Fig. 3.2). The samples were collected on the 29^{th}-30^{th} June, 2^{nd}-4^{th} and 6^{th} July, 2009, and consist of 18 pine stands, 6 other forests and 6 agricultural and shrubland areas, in addition to two transects on a reference bright field for the validation of airborne reflectance. Both spectroradiometric and destructive measurements were carried out at each sample plot.

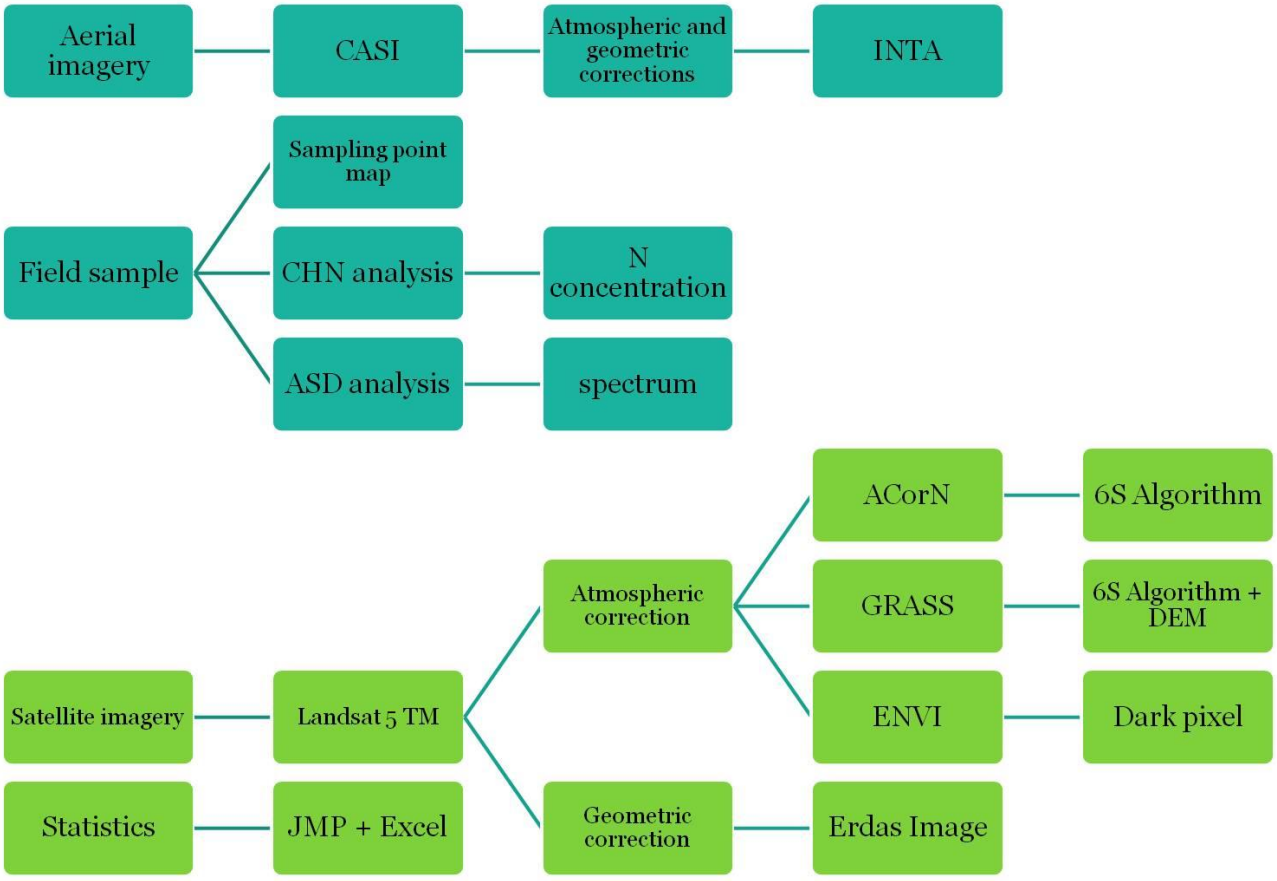


Fig. 3. 1 Procedure scheme. The blue part was realized in Italy, while the green sections were realized at University of New Hampshire, USA.

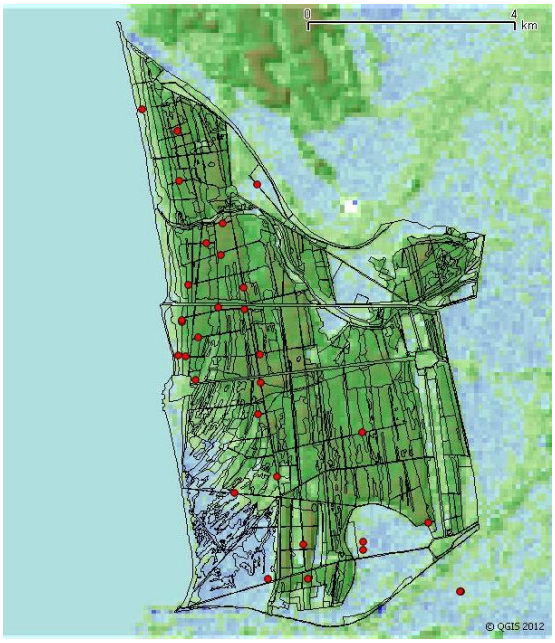


Fig. 3. 2 Field sample locations draped on the SRTM digital elevation model and forest plan map.

Two ASD FieldSpec Pro portable spectroradiometers were used for the measurement of nadir-view hyperspectral reflectance over the 350-2500 nm spectral range, with a spectral resolution (FWHM of a single emission line) of approximately 3 nm at around 700 nm. Light was collected in nadir view with a bare fiber optic with a field of view (FOV) of 25°. Measurements on a white reflective panel just prior of the canopy readings provided a reference for the estimation of canopy reflectance.

Because of the high heterogeneity of forest canopies, a complex sampling strategy was applied in order to have a correspondence between the field measurements and images that allows proper comparison and validation.

Access to the forest canopy was possible through a 32

m high mobile platform ('cherry-picker') on a lorry. Because of the dimensions of the lorry, measurements could only be taken in proximity of forest roads and paths, and could not be arranged along regular transects. In order to exclude the lorry and the road from the field of view, the measuring platform was offset a few meters from the centre. Measurements were taken using a nested clustered design, at four points around the cherry-picker and – at each point – in four directions around the platform. Since measurements had to avoid the inclusion of the road or the platform itself in the FOV, the sampling direction in the nested design was therefore not fixed.

Foliar biochemical characteristics were determined on a sample of leaves from each of the points included in canopy reflectance measurements. At each point, a branch was collected from each of 4 trees adjacent to the small forest gap accessed by the cherry-picker.

A sample of leaves from each branch was quickly frozen in liquid N and brought to the laboratory, then it was grounded and the measurement of leaf N concentration was carried out using the elemental analyzer CHNS-O (EA 1110, Thermo Fisher, Milan, Italy).

Field data were collected in the same days as the aircraft flight.

Because of the presence of some areas with sparse vegetation, resulting in contamination of acquired spectra with soil or litter reflectance, NDVI (Normalized Difference Vegetation Index) was computed from ground spectra for each plot and only plots with an NDVI greater than 0.7 were retained in some of the analyses; considering only forest stands, the inclusion of an NDVI threshold restricted the sample size to a total of 17 ground plots.

3.2.2. CASI dataset

Airborne images used in this work were acquired in two flights on 2nd and 4th of July 2009 by a CASI-1500i (S/N 2516) (Compact Airborne Spectrographic Imager) hyperspectral imager provided and operated by INTA (Instituto Nacional de Tecnica Aeroespacial, Madrid) as part of the ESA campaign. CASI is characterized by 288 spectral bands, with an IFOV of 0.49 mrad (FOV 40°). The survey included 7 flight lines parallel to the coast, with a spatial resolution of 1.35 x 3.46 m.

Ancillary measurements required for atmospheric corrections were carried out in coincidence with the overflight; air pressure and humidity were measured 3 m above the vegetation with a 10-Hz frequency by a LiCor Li-7500 open-path gas-analyzer; aerosol optical thickness (AOT) was estimated using a Cimel CE-317 sun photometer. Both atmospheric and geometric corrections of CASI imagery were performed by ESA SEN3EXP remote sensing team, in collaboration with INTA and ITRES Research Ltd, Canada.

In order to obtain a sample consistent with the footprint of ground spectroradiometric measurements, in the analyses of airborne imagery we utilized ROIs (Regions Of Interest) of 16 x

16 m. Because of the amount of CASI images (7 flight lines), each sample plot could be represented in different frames; for each ROI, canopy reflectance was derived from the image where it was nearest to the centre of the frame, so as to try to prevent the reflectance anisotropy problem faced in Mezzini et al. (2010). Finally, in each ROI only pure vegetation pixels were selected based on the ENVI ASCII export option and Excel analysis of their spectra, in order to minimize soil effects in the spectrum, avoiding road pixels or clearings.

3.2.3. Landsat 5 TM dataset

Using the same approach already described in the previous study (Mezzini et al., 2010) we decided to utilize Landsat imagery in San Rossore Reserve, since the spatial resolution of MODIS images used in the original study (Ollinger et al., 2008) would not be appropriate for the heterogeneity of the landscape under study. Unfortunately, good images could not be obtained from Landsat 7 ETM+, because of a problem in the sensor since 2007. Thus, we utilized images from Landsat 5 TM (Thematic Mapper), characterized by a spatial resolution of about 30 m (pixel footprint), 7 spectral bands (from visible to thermal wavelengths) and a FOV of 15°.

Out of the Landsat images available from the USGS web site, a good image was available for the 23rd of July 2009, almost coincident with all other data sources. The Landsat product had to be corrected both atmospherically and geometrically.

As a first step, the Landsat image was co-registered on a vector-graphics CTR (Carta Tecnica Regionale – Regional Technical Map) at 1:10000 scale of Tuscany (UTM, zone 32 N, WGS84) by means of GCPs and nearest neighbour resampling.

Several different software packages were compared for the atmospheric corrections of TM images: ENVI (Exelis, Inc. Visual Information Solutions, McLean, Virginia), GRASS GIS (GRASS Development Team, 2012) and ACOrN (Atmospheric Correction Now, vers. 6lx, ImSpec LLC, Seattle). The whole process was realized thanks to the collaboration with University of New Hampshire.

Firstly we tested Linux ENVI 4.7 pre-processing tools, beginning from Landsat TM calibration utilities, converting digital numbers in TOA (Top of Atmosphere) reflectance directly. Then, as in Mezzini et al. (2010), we performed Dark Pixel Subtraction by selecting a big ROI in the sea, so as to obtain TOC (Top of Canopy) reflectance. This procedure gave us good results, but appeared to overestimate vegetation reflectance in the red region, which was too high compared with a standard vegetation spectrum and ground sepctroradiometric measurements.

The Open Source solution of GRASS GIS (GRASS Development Team, 2012) was subsequently tested. Initially we utilized version 6.4 i.landsat.toar extension to convert digital numbers to TOA

reflectance, then in version 6.5 we launched a little script based on i.atcorr module to convert TOA to TOC reflectance. The last module was based on the 6S Algorithm (Vermote, Tanré, Deuzé, Herman, & Morcrette, 1997) and required also a DEM (Digital Elevation Model) of the area, that was derived from SRTM (Shuttle Radar Topography Mission) FTP site (NASA, 2009) using the r.in.srtm GRASS module. This procedure appeared to yield good results, but a problem with the TOC reflectance and DEM was detected. Hopefully new versions of the software will solve this problem.

Finally we applied another 6S Algorithm based software, ACOrN, thanks to the collaboration with the University of New Hampshire. In this case the result was fully satisfactory, and it was decided to utilize this product to select our ROIs of 1 Landsat pixel, covering an area of 30 x 30 m, by overlapping it to GPS coordinates of our field samples.

3.2.4. Statistical analysis

Minimum and maximum reflectance values in the ROI were extracted using the ENVI statistical tool, together with mean reflectance values and standard deviation for each wavelength and each kind of sensor. These results can also be plotted into spectral plot or into scatter plot using standard ENVI routines.

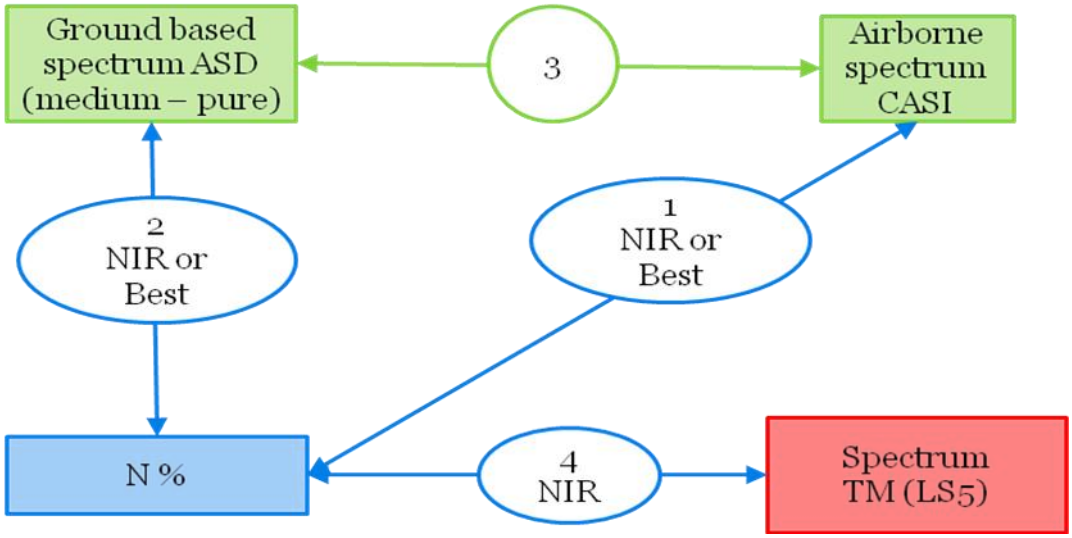


Fig. 3. 3 Scheme of statistical analyses.

Using results extracted with ENVI statistics routines, we analyzed the relationship between N concentration and ASD, CASI and Landsat vegetation reflectance. Moreover we tested also the relation between ASD field spectrum and CASI airborne spectrum (Fig. 3.3). In contrast with our previous study (Mezzini et al., 2010), we tried to statistically analyze data through multiple regressions techniques; both JMP 9 software (SAS Institute Inc.) and Excel were used for the analysis.

3.3. Results and discussion

We initially analyzed the simple regression between nitrogen concentration and NIR CASI reflectance (857 nm) as in Ollinger et al. (2008) (Fig. 3.3, step1), obtaining a quite good regression, with a coefficient of determination of 0.63 and an RMSE (Root Mean Square Error) of 0.44. We also tested a multiple stepwise regression using all CASI bands as explanatory variables, obtaining a better result for the whole database with R^2=0.66, and a also better regression for forest sub-set with a coefficient of determination of 0.77 and an RMSE of 0.36 (Fig. 3.4). This time, the bands involved were 934, 936, 939 nm, thus still in the NIR interval, but at slightly longer wavelengths than the one suggested by Ollinger et al. (2008). In both cases, the relationship was highly significant, as also demonstrated by the confidence curves in the figure (SAS Institute Inc, 2006). It should be noted, however, that also in this enlarged dataset a split of the samples is apparent between coniferous and deciduous forests, as seen in previous studies (Mezzini et al., 2010; Ollinger et al., 2008), resulting in a strong leverage effect of the two resulting clusters and leaving open the possibility of a chance correlation between leaf biochemistry and reflectance features induced by life history characteristics.

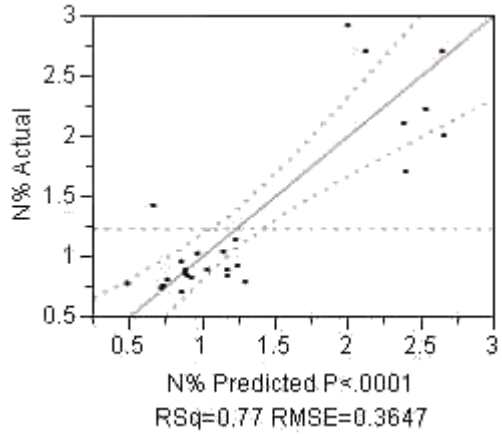


Fig. 3. 4 Multiple regression model of CASI bands in a leverage plot, considering forest sub-set (n=25). Horizontal dotted line is the response mean, while the others are the confidence curves.

As second step (Fig. 3.3, step 2) we studied the regressions between nitrogen concentration and ASD NIR reflectance. These analyses were performed both for the whole dataset and for the sub-set with NDVI>0.7 (only 17 plots). NDVI selection didn't give us better results, probably because of the little number of samples selected; the R^2 for the whole dataset was 0.59, while with the selection it declined to 0.55. Then, also in this case, we applied a stepwise regression, obtaining significantly better results (Fig. 3.5) with a coefficient of determination of 0.85 and an RMSE of 0.49; comparable results were obtained after selecting points according to NDVI as described. The bands involved were located at 703, 1202 and 1261 nm: the first wavelength is located in the red edge region and in literature was usually associated with chlorophyll and carotenoid contents (Asner & Martin, 2008),

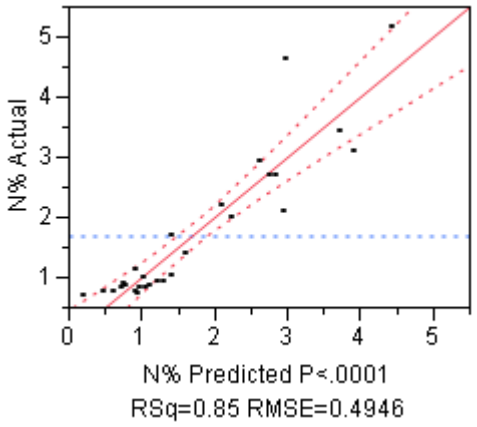


Fig. 3. 5 Multiple stepwise regression model of ASD bands in a leverage plot, considering the whole dataset (NDVI>0, n=30).

while the others are in the SWIR (Short Wave InfraRed) region historically associated to water and starch (Curran, Dungan, & Peterson, 2001), but also to cellulose and lignin absorption features (Curran, 1989). Results for the analysis of the entire dataset didn't show the clustering between coniferous and broadleaf species mentioned before, but a more even distribution across the entire range of leaf N concentrations.

In order to understand the observed differences between CASI and ASD results, despite the similarity of their spectral range, we compared for each area the CASI spectrum with ASD spectra

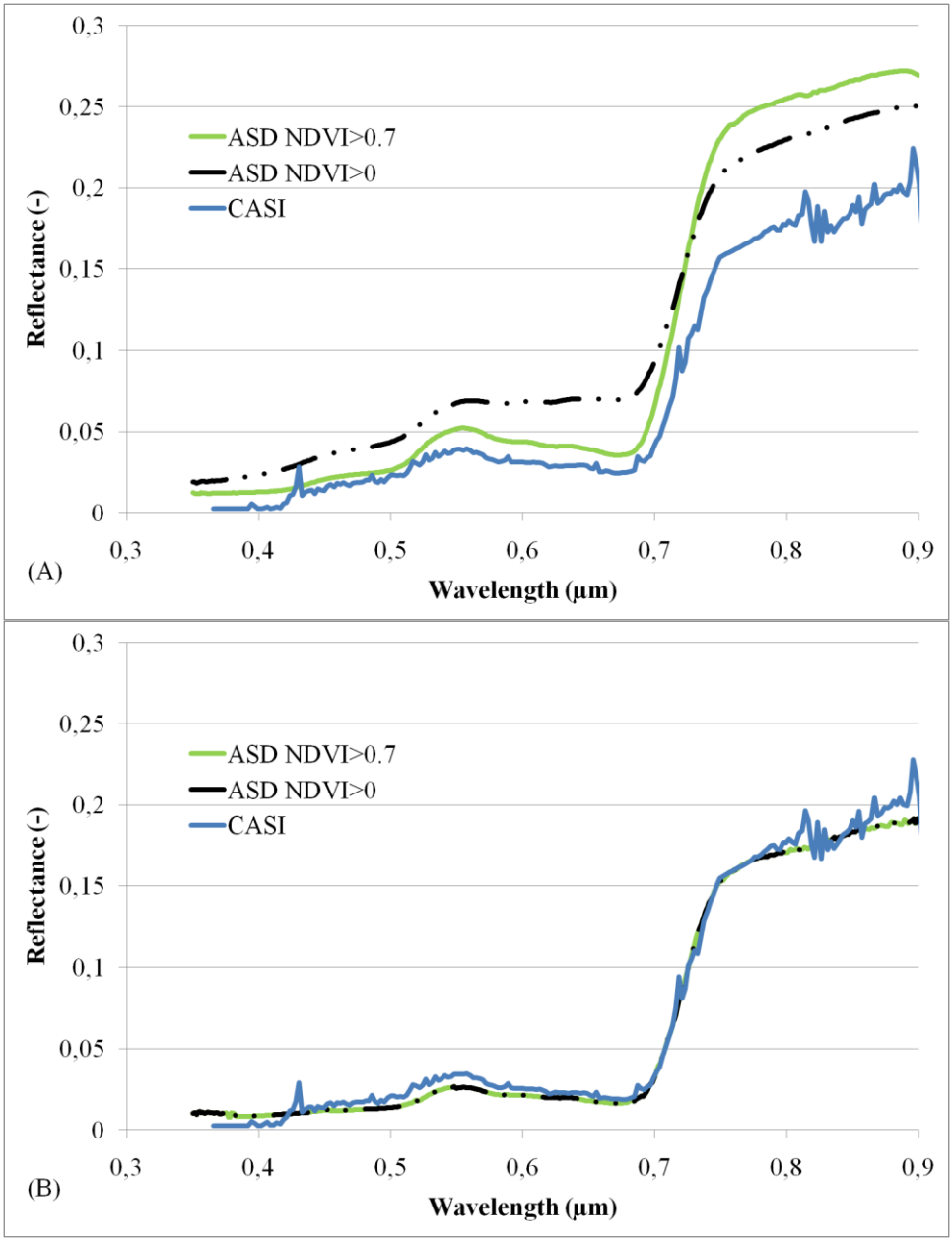


Fig. 3. 6 ASD and CASI spectra of two different areas: (A) conifer plot of *Pinus pinaster* (B) evergreen plot of *Quercus ilex*; deciduous broadleaf plots showed discrepancies similar to the one apparent in coniferous plots.

with or without selection for NDVI threshold. As shown in Fig. 3.6 results were really similar in broadleaf evergreen canopies, but in conifers and deciduous broadleaves a marked discrepancy was apparent, especially in the NIR region and green peak, where CASI underestimated canopy reflectance as directly meadured by ASD spectrometers. This underestimation of NIR reflectance in broadleaf species, which commonly show the highest value of canopy N concentration, could help explaining the partial clustering of data observed in the CASI analysis. To date we could not find an explanation for this difference, possibly related to problems with the CASI instrument or to atmospheric corrections. It is difficult to see how the problem could be related to ground measurements of canopy reflectance, as two calibrated ASD specrometers were used, yielding very similar results.

As a final step, we analyzed the relationship between nitrogen concentration and Landsat NIR reflectance (centered at 840 nm). Multiple regression techniques were not applied in this case, because of the small number of bands available outside the visible range. Despite this, a good correlation was observed, characterized by a coefficient of determination of 0.84 and an RMSE of 0.50, without any clustering in the distribution (Fig. 3.7).

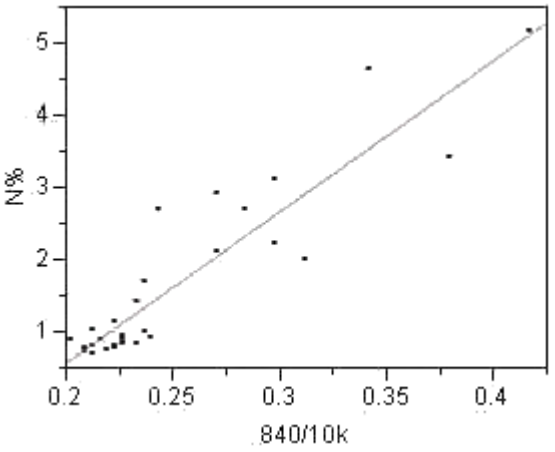


Fig. 3. 7 Regression between nitrogen concentration and Landsat NIR reflectance; R^2=0.84, RMSE=0.5, n=30.

3.4. Conclusions

The results of this work appear to be rather promising. Thanks to the larger sample size and to the wider interval of conditions and functional types explored in the new dataset, the clustering in the distribution between coniferous and broadleaf forests observed by Mezzini et al. (2010) was solved and the regression between Landsat 5 TM and canopy nitrogen concentration was really good. A problem with imagery from the CASI airborne sensor, on the contrary, prevented us from using these more detailed hyperspectral images to estimate N concentration. The differences between ground measurements with ASD spectrometers and CASI spectra are significant, and colud not be attributed to canopy anisotropy, as in the case of MIVIS images discussed in the previous chapter, since the rather narrow FOV and the availability of several overlapping overpasses made it possible to select only spectra acquired at near-nadir view. Possibly associated with this, CASI yielded the worst regression with N data.

Obviously additional studies will be required before these new techniques can be more widely applied, exploring also other sensors and environments. We therefore decided to test this novel satellite remote sensing approach also in a reality different from Italy, and with a larger database, and, this time, utilizing also MODIS imagery.

3.5. Acknowledgements

This work was supported by the EU FP7 project GHG-Europe and by the Marco Polo Program of the University of Bologna. We thank all the Sen3Exp team for their help and preliminary analysis of CASI imagery and Fondazione Edmund Mach, above all Markus Neteler's team, for their help in GRASS atmospheric corrections procedure. Many thanks also to Complex Systems Research Center of the Institute for the Study of Earth, Oceans and Space at University of New Hampshire, USA, above all Prof. Scott Ollinger and Dr. Lucie Lepine for their help, support and hospitality. We also gratefully thanks Dr. Paola Gioacchini for CHN analyses.

3.6. References

Asner, G., & Martin, R. (2008). Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment*, *112*(10), 3958–3970. doi:10.1016/j.rse.2008.07.003

Chang, S. (2003). Nondestructive and rapid estimation of hardwood foliar nitrogen status using the SPAD-502 chlorophyll meter. *Forest Ecology and Management*, *181*(3), 331–338. doi:10.1016/S0378-1127(03)00004-5

Curran, P.J., Dungan, J. L., & Peterson, D. L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry:: Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment*, *76*(3), 349–359. Retrieved from http://www.sciencedirect.com/science/article/pii/S0034425701001821

Curran, Paul J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, *30*(3), 271–278. doi:10.1016/0034-4257(89)90069-2

GRASS Development Team. (2012). Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.3. Open Source Geospatial Foundation. Retrieved from http://grass.osgeo.org

Heimann, M., & Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, *451*(7176), 289–92. doi:10.1038/nature06591

Johnson, L. F. (2001). Nitrogen influence on fresh-leaf NIR spectra. *Remote Sensing of Environment*, *78*(3), 314–320. doi:10.1016/S0034-4257(01)00226-7

Knyazikhin, Y., Schull, M. a, Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., et al. (2012). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences of the United States of America*, 1–8. doi:10.1073/pnas.1210196109

Kokaly, R. (1999). Spectroscopic Determination of Leaf Biochemistry Using Band-Depth Analysis of Absorption Features and Stepwise Multiple Linear Regression. *Remote Sensing of Environment*, *67*(3), 267–287. doi:10.1016/S0034-4257(98)00084-4

LeBauer, D. S., & Treseder, K. K. (2008). Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, *89*(2), 371–379. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18409427

Magnani, F., Mencuccini, M., Borghetti, M., Berbigier, P., Berninger, F., Delzon, S., Grelle, A., et al. (2007). The human footprint in the carbon cycle of temperate and boreal forests. *Nature*, *447*(7146), 848–50. doi:10.1038/nature05847

Martin, M., & Aber, J. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications*, *7*(2), 431–443.

Martin, M. E., Plourde, L. C., Ollinger, S. V., Smith, M.-L., & McNeil, B. E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment*, *112*(9), 3511–3519. doi:10.1016/j.rse.2008.04.008

Mezzini, E., Raddi, S., Pippi, I., & Magnani, F. (2010). Telerilevamento della concentrazione di azoto fogliare in ecosistemi forestali litoranei. *attiasita.it* (pp. 1335–1338). Retrieved from http://www.attiasita.it/ASITA2010/Pdf/223.pdf

NASA. (2009). SRTM Ftp site. Retrieved from http://dds.cr.usgs.gov/srtm/

Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *The New phytologist*, *189*(2), 375–94. doi:10.1111/j.1469-8137.2010.03536.x

Ollinger, S. V., Richardson, A. D., Martin, M. E., Hollinger, D. Y., Frolking, S. E., Reich, P. B., Plourde, L. C., et al. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19336–41. doi:10.1073/pnas.0810021105

Ollinger, S. V., Smith, M. L., Martin, M. E., Hallett, R. a., Goodale, C. L., & Aber, J. D. (2002). Regional Variation in Foliar Chemistry and N Cycling among Forests of Diverse History and Composition. *Ecology*, *83*(2), 339. doi:10.2307/2680018

SAS Institute Inc. (2006). JMP Points of interest. Retrieved from http://www.jmp.com/software/pdf/102565_pointsofinterest.pdf

Serrano, L., Peņuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data:: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2-3), 355–364. Retrieved from http://www.sciencedirect.com/science/article/pii/S0034425702000111

Smith, M.-L., Ollinger, S. V., Martin, M. E., Aber, J. D., Hallett, R. a., & Goodale, C. L. (2002). Direct Estimation of Aboveground Forest Productivity Through Hyperspectral Remote Sensing of Canopy Nitrogen. *Ecological Applications*, *12*(5), 1286–1302. doi:10.1890/1051-0761(2002)012[1286:DEOAFP]2.0.CO;2

Vermote, E., Tanré, D., Deuzé, J. L., Herman, M., & Morcrette, J.-J. (1997). Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, *35*(3), 675–686. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=581987

Wessman, C., Aber, J., Peterson, D., & Melillo, J. (1988). Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems. *Nature*, *335*, 154–156. Retrieved from http://www.nature.com/nature/journal/v335/n6186/abs/335154a0.html

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–7. doi:10.1038/nature02403

4. Remote sensing of foliar N concentration at regional scale

This article is being submitted to the interdisciplinary ISI journal "Remote Sensing of Envrionment" with coauthors: Federico Magnani, Jordi Sardans and Josep Peñuelas.

4.1. Introduction

Nitrogen availability is one of the main factor that limits plant growth and C fixation in boreal and temperate forests. Leaf N concentration could be therefore used for the estimation of site fertility and forest potential productivity. Estimating leaf N concentration by aerial and satellite remote sensing techniques is increasingly seen as an interesting opportunity in the monitoring of forest canopies and in the parameterization of biogeochemical models (Pan, Hom, & Jenkins, 2004; Turner, Ollinger, & Kimball, 2004). Recently, the possibility has been suggested of estimating forest foliar N concentration from canopy MODIS (MODerate resolution Imaging Spectroradiometer) NIR reflectance in the USA (Hollinger et al., 2010; Ollinger & Smith, 2005; Ollinger et al., 2008; Smith et al., 2002). In this study we will test this methodology at regional scale in a different context, characterized by a more patchy vegetation distribution, such as the temperate and Mediterranean forest of Catalonia, Spain.

4.2. Material and methods

4.2.1. Study site and field data

The study area covers the whole of Catalonia; Catalonia is located in the north-eastern part of Spain, and covers different climatic regions: from the Pyrenees to the Mediterranean Sea, from Alpine climatic region to the semiarid-Mediterranean region (Sardans & Peñuelas, 2012). We chose this particular area because of its landscape heterogeneity, similar to the one commonly found in Italy, and because of the availability of a good forest inventory, comprising detailed information not just on forest biomass and yield, but also on ancillary ecological characteristics such as leaf N concentration.

In this study we utilized data from the Ecological Forest Inventory of Catalonia (IEFC, Gracia et al., 2004) and the Third Spanish National Forest Inventory (Villaescusa & Díaz, 1998; Villanueva, 2005) . The dataset contains 3482 sampling plots (10 m radius), randomly distributed throughout 32114 km^2 of forested lands in the whole of Catalonia. Field samples were collected from 1988 to 2001: from 1988 to 1998 as far as the IEFC is concerned, and in 2000-2001 as part of IFN-III. This georeferenced database (WG84 Datum, UTM zone 31 Projection) contains typical forest inventory data but also other metrics of more biogeochemical relevance. Sampling position and date,

biological characterization of the plot (main species and ecotype), foliage biomass, leaf N concentration and total foliage N content are the most important variables for the current study.

4.2.2. Sampling and chemical analyses

With regards to foliar samples, in each of the 3482 plots a combination of foliar samples from at least three trees of the predominant species was analyzed. Samples were collected in all canopy directions in the upper middle part of the crown by telescopic loppers, therefore the final foliar sample should have included mainly sunlit (Gracia et al., 2004; Vilà et al., 2003). In contrast with N content, leaf N concentration is however known to vary little across the canopy (Niinemets, Kull, & Tenhunen, 1998).

To obtain C and N concentration, dried and grounded samples were analyzed by combustion coupled to gas chromatography (Thermo Electron Gas Chromatograph, model NA 2100, C.E. instruments – Thermo Electron, Milan, Italy) (Gracia et al., 2004; Peñuelas et al., 2010; Sardans & Peñuelas, 2012).

4.2.3. Satellite data and preprocessing

4.2.3.1. Surface reflectance imagery

To test the existence of a relationship between nitrogen and remote sensing reflectance we chose NASA's MODIS products as satellite imagery dataset, as suggested by Ollinger et al. (2008).

In particular we chose the MOD09 product (Terra MODIS, Terra Land Level 2 Products, MOD09 - Level 2, 5-min Land Surface Reflectance, Collection 5), that is characterized by three levels of spatial resolution (250 m, 500 m and 1 km) depending on the band considered and contains 17 reflectance bands from 405 nm to 2155 nm, and 3 thermal emissivity bands from 3.660 µm to 12.270 µm. Some reflectance bands are replicated with different spatial resolution (mainly bands 1-7). Departing from Ollinger et al. (2008) we utilized MOD09 instead of MOD43B because of the better spatial resolution, more appropriate to the landscape heterogeneity typical of Spain. Also the temporal resolution is different, MOD43B is a 16-day composite imagery, whereas MOD09 is a daily based imagery. Both products consist of atmospherically corrected surface reflectance data. Because of the different map projection system applied in the inventory and MODIS (Sinusoidal projection) we had to reproject all satellite images trough MODIS Reprojection Tool Swath (MRTSwath, Release 2.2, December 2010). This application needs an associated Geolocation file for any MODIS swath image, called MOD03 for MODIS Terra and freely available.

Since MODIS had been launched in 1999, satellite images for the period 2000-2001 were used, corresponding to the overlapping period with the inventory sampling period; moreover, because of

the presence of deciduous species, we preferred to analyze only images and data for the period May-July, corresponding to the peak of the growing season, so as to avoid any residual phenological effects. Four images MOD09 with the best weather and visibility conditions on Catalonia were therefore selected in the period May-July of 2000-2001 (2 images per year: one at the end of May- beginning of June, the other in the middle of July).

4.2.3.2. Land cover classification imagery

Both leaf N and canopy reflectance are known to be significantly affected by plant functional type. A vegetation type classification is obviously provided for each plot in the forest inventory, but htis source of information would not be available for the extrapolation of results to regional level; an independent estimate of vegetation type was therefore derived from MODIS dataset itself, using the Vegetation Continuous Fields Yearly L3 Global 250 m (MOD44B) and Land Cover Type Yearly L3 Global 500 m SIN Grid (MCD12Q1) products. The last layer contains five different land cover classification bands, but also this product is provided in Sinusoidal projection. Therefore, as for MOD09 we had to reproject MOD44B and MCD12Q1; since these are not swath products, we had to utilize for this purpose the freely available MODIS Reprojection Tool (MRT, Release 4.1, April 2011).

4.2.4. Remote sensing and GIS analyses

In order to extract the surface reflectance of pixels associated with each inventory plot, we sampled MODIS Surface Reflectance corrected imagery by the open source software Quantum GIS - Point Sampling Tool (QGIS, Quantum GIS Development Team, 2012), using IEFC plot locations as a point layer. By so doing, we obtained a new layer with all inventory information and surface reflectance of all MOD09 reflectance and MCD12Q1 vegetation type bands, as well as the ancillary MOD09 bands containing information on data quality (cloud presence, shadows, snow, land/water flag, etc.).

In order to better characterize each inventory plot, we also analyzed the morphology of Catalonia. From the MiraMon Map Server (Pons, 2009) we obtained the Digital Elevation Model (DEM) of the whole of Catalonia, with a resolution of 30 m, from which aspect and slope could be computed with QGIS; the elevation, aspect and slope of each inventory point were finally extracted.

4.2.5. Statistical analyses

All statistical analyses were completed using the software packages SAS v.9.0 (SAS Institute Inc., Berkeley, CA, USA) and Microsoft Excel for plotting results.

Initially a general analysis of inventory data was performed, so as to explore the characteristics and composition of Catalan forests. A preliminary analysis of the inventory dataset demonstrated the absence of a temporal pattern in the inventory data; it was therefore decided to include in the analysis data of all years (1988-2001) rather than only data collected in years corresponding to MODIS imagery (2000-2001).

In remote sensing analyses it is quite usual to face problems related to mixed pixels. On the ground these pixels represent areas characterized by different land cover types or characteristics (species, fertility, etc.). The spectral features of the pixel will be the average of the signatures of all the elements included, each weighted by its relative area contribution. This contrasts with the nature of forest inventory plots, each referring to a point, or rather to a small and homogeneous sample area (of 10 m radius). The problem will be particularly sever when dealing with satellite products with a low spatial resolution (250-1000 m in the case of MODIS) and when the signature of the pixel is affected by a large proportion of bare soil or urban areas. The extent of latter problem, however, could be highlighted by the value of satellite vegetation indices commonly used to quantify vegetation density. To bypass it we calculated for each plot the mean Normalized Difference Vegetation Index (NDVI) and the mean Enhanced Vegetation Index (EVI) by means of all MODIS surface reflectance bands at the best spatial resolution (Justice et al., 1998).

$$NDVI = (\rho_{NIR} - \rho_{red})/(\rho_{NIR} + \rho_{red}) \quad (1)$$

where ρ_{NIR} and ρ_{red} are reflectance at a wavelength of 859 and 645 nm, both available at 250 m spatial resolution (MODIS bands 1 and 2). EVI was computed as:

$$EVI = 2.5(\rho_{NIR} - \rho_{red})/(L + \rho_{NIR} + C_1\rho_{red} + C_2\rho_{blue}) \quad (2)$$

where ρ_{NIR}, ρ_{red}, ρ_{blue} are surface reflectance at a central wavelength of 859, 645 and 469 nm, respectively, and refer to band 1, 2 and 3 of MOD09, with a spatial resolution of 500 m; L is the canopy background and snow correction and is assumed to be equal to 1; C_1 and C_2 are coefficients related to aerosol and assumed values of 6 and 7.5 respectively, as suggested in Justice et al. (1998).

These indexes were used to filter inventory data; in particular, in the following analyses we selected only plots with NDVI values above 0.7, so as to ensure that the corresponding pixel was a predominantly vegetation pixel.

Another filter was applied, related to the quality of remote sensed data. In particular, we selected only plots with good data quality, as obtained from the band “1 km Reflectance Data State QA” of MOD09 product, as suggested by Neteler (2010). We rejected pixels classified as “cloudy”, “mixed”, “ocean coastlines and lake shorelines”, “cloud shadows”, “cirrus” and “snow”.

As a result of data selection, the total number of samples used on the analyses was 736, or 21.1 % of the original inventory plots.

For each plot, we calculated means and standard error of surface reflectance for the four dates considered.

4.2.5.1. Simple and multiple regressions

Regression analyses were used to determine relationships between surface reflectances and foliar nitrogen; a simple linear regression between NIR reflectance (band 2, central wavelength of 859 nm) and nitrogen was initially tested, followed by a more thorough multiple regressions against all available reflectances. Model selection in multiple regression was based on the adjusted R^2 method, so as not to penalize the model by too large number of parameters; the RMSE (Root Mean Square Error) was also computed for each model selected. A stepwise multiple regression analysis was also carried out; all variables left in the model are significant at the 0.15 level.

4.2.5.2. Partial Least Squares regression model

As in most remote sensing studies, problems of model instability could arise because of the large number of predictors, which could cause a possible over-fit of the observed data, whereby the model captures very well the calibration dataset, but fails in the prediction of independent validation data. To avoid this problem we analyzed data by the Partial Least Squares regression (PLS) technique (SAS Institute Inc, 1999; SAS Support, 2012; Tobias, 1995) which works also with collinear, noisy and numerous predictors (Haenlein & Kaplan, 2004; Wold, Sjöström, & Eriksson, 2001). We chose this method also because it is one of the most commonly used in remote sensing and ecosystem modeling (Asner & Martin, 2008; Keithley, Heien, & Wightman, 2009).

In order to overcome the problem of model over-fitting, it is important to choose the correct number of factors (or components) used in the PLS analysis by a cross-validation process. There are different kinds of cross-validation procedures; in this case we chose the SPLIT option available in the SAS software, as there was a real possibility that our observed predictors (i.e. reflectances at different wavelengths) were correlated with each other. This procedure divides the input dataset into two or more groups (in this study 7) and fits the model to all the groups except one, then validates the prediction capability of the model for the omitted group. All this procedure is repeated until each group is used for validation. The best number of components to be considered is computed iteratively. For each number of factors, the Predicted Residual Sum of Squares (PRESS) is computed, which is based on the residuals generated by this process. As suggested by Van Der Voet (1994),we selected the smallest number of factors with an approximate significance level that is greater than 0.10.

The comparison between results was based on the cross-validated coefficient of determination (R^2) as a measure of variability, and on cross-validated RMSE as a measure of precision of the prediction.

At the end of this analysis, the procedure (by the SOLUTION option) provided also the coefficients of the model (normalized by default), which could be finally applied to our remote sensing images to obtain predicted N maps.
The first iteration of the PLS analysis highlighted some outliers, that were rejected before repeating all these statistical analyses for the new dataset of 707 samples.
As a last step of the analysis, we tested the possibility to utilize PLS results obtained from MODIS reflectances to estimate nitrogen concentration of different vegetation macro-categories (evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, shrubland, temperate grassland, cereal crops). This could prove really important for the implementation of these data into ecological models.

4.2.6. Nitrogen map

The ultimate aim of this work was to create a regional map of modelled foliar nitrogen concentration. The PLS regression model was therefore utilized to estimate nitrogen in the whole of Catalonia, starting from MODIS mean reflectances images. As a first step, we created a mask layer starting from the reflectance quality band of the MOD09 product (1 km resolution); in particular, we masked out pixels flagged in layers "cloud state", "internal cloud algorithm", "land water", "cloud shadows", "cirrus detected" and "internal snow mask", using the GRASS module "i.modis.qc" and a reclassification process. Then we intersected this mask with Catalonia's boundary. Next we computed temporal mean reflectances of all resolutions and bands to finally apply the coefficients of the model and obtain the desired nitrogen map of the whole of Catalonia.
All processes were carried out in GRASS 6.4.3svn (GRASS Development Team, 2012).

Further details of all Material and methods procedures are presented in Appendix B.

4.3. Results and discussion

In this section we will present all the results obtained from the analyses of the reduced dataset, without outliers data, used on the comparison with MODIS reflectance, comprising a total of 707 samples.

4.3.1. Inventory distribution

The union of the two inventory databases (IEFC, Ecological Forest Inventory of Catalonia; IFN-III, Third Spanish National Forest Inventory) resulted in a large number of samples distributed over the whole of Catalonia. Only the area in the centre-west of Catalonia was not represented in the dataset,

because the province of Lleida is a mostly agricultural area, almost without forests. Because of the two different sources of inventory data, samples were collected over the course of several years (Fig 4.1). In particular in the IEFC rather small areas were covered each year from 1988 to 1998; while in the National Forest Inventory samples were collected in only two years and each year covered a great part of the area.

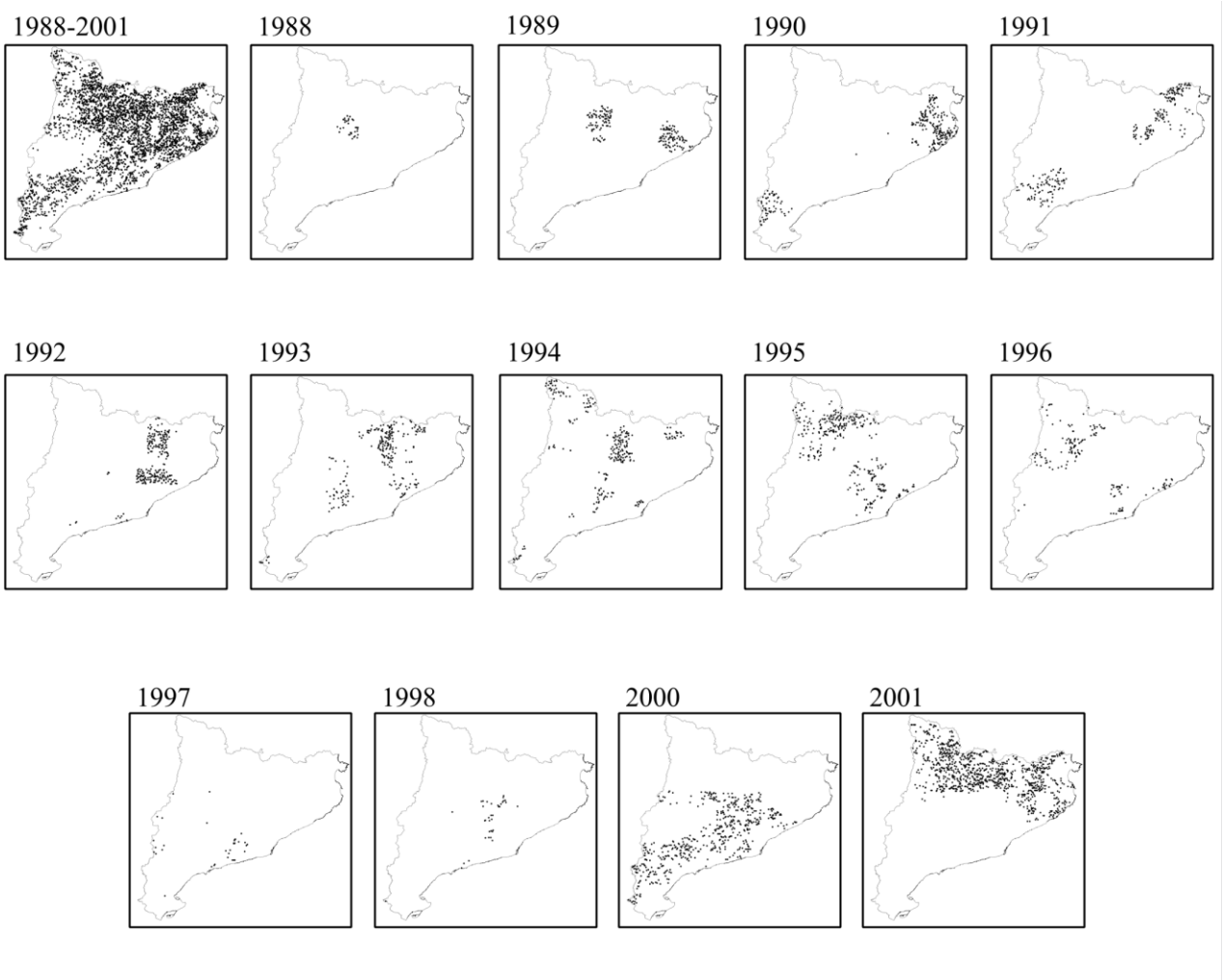


Fig. 4. 1 Temporal distribution of the whole dataset, both IEFC (1988-1998) and Spanish National Forest Inventory (2000-2001).

As already mentioned, such a scalarity of sample collection did not appear to result in any significant bias.

The main species represented in the dataset were *Pinus halepensis* (22.14 % of total plots), *Quercus ilex* (20.65 %), typical of Mediterranean climate and *Pinus sylvestris* (20.30%) typical of continental climate, localized in the Pyrenees. The relative contribution and mean characteristics of each species are presented in Table 4.1.

The spatial distribution of the three broad vegetation types detected in the inventory (conifers, deciduous and evergreen broadleaves) is presented in Fig. 4.2; as expected, most of the evergreen

broadleaf plots are near the Mediterranean coast, as they are mainly represented by *Quercus ilex* and *Quercus suber* stands. On the contrary, coniferous plots are more evenly distributed over the whole region, mainly due to the variety of ecological types in the genus *Pinus*: the main species are *P.halepensis*, typical of the Mediterranean area, and *P.nigra* and *P.sylvestris*, mainly observed in the Pyrenees.

Table 4. 1 Relative contribution and mean nitrogen concentration of each database species. The * indicates the number of plots of each species in the reduced database of 707samples.

Inventory species	n. of plots	n. of plots *	mean N %	st. error
Abies alba	23	7	1.06	0.06
Arbustus unedo	25	11	1.38	0.06
Acer campestre	2	-	3.16	1.18
Alnus glutinosa	1	-	2.66	-
Betula pendula	6	3	2.34	0.15
Castanea sativa	21	12	2.22	0.14
Cedrus deodara	1	-	1.49	-
Eucalyptus globulus	4	3	1.41	0.15
Fagus sylvatica	38	26	2.32	0.01
Fraxinus angustifolia	2	-	2.61	0.02
Fraxinus excelsior	5	2	2.30	0.01
Pinus halepensis	771	25	1.07	0.01
Pinus nigra	431	20	0.95	0.01
Pinus pinaster	34	6	0.99	0.05
Pinus pinea	65	7	1.17	0.03
Pinus radiata	2	1	1.51	0.17
Pinus sylvestris	707	209	1.19	0.01
Pinus uncinata	235	68	0.97	0.01
Populus hybrides	1	-	2.55	-
Populus nigra	3	-	2.34	0.43
Populus tremula	6	3	2.40	0.40
Prunus avium	2	2	2.53	0.02
Pseudotsuga maziesii	1	-	1.10	-
Quercus canariensis x humilis	13	3	2.09	0.20
Quercus cerroides	18	6	1.71	0.17
Quercus faginea	23	3	1.84	0.16
Quercus humilis	61	25	1.85	0.09
Quercus humilis/cerroides	29	10	2.04	0.09
Quercus ilex	719	182	1.37	0.01
Quercus petraea	30	17	2.29	0.12
Quercus pyrenaica	1	-	1.55	-
Quercus robur	2	2	2.28	0.12
Quercus suber	192	53	1.64	0.02
Robinia pseudacacia	2	-	4.04	0.46
Tilia cordata	3	3	2.97	0.17
Ulmus minor	2	-	2.49	0.76

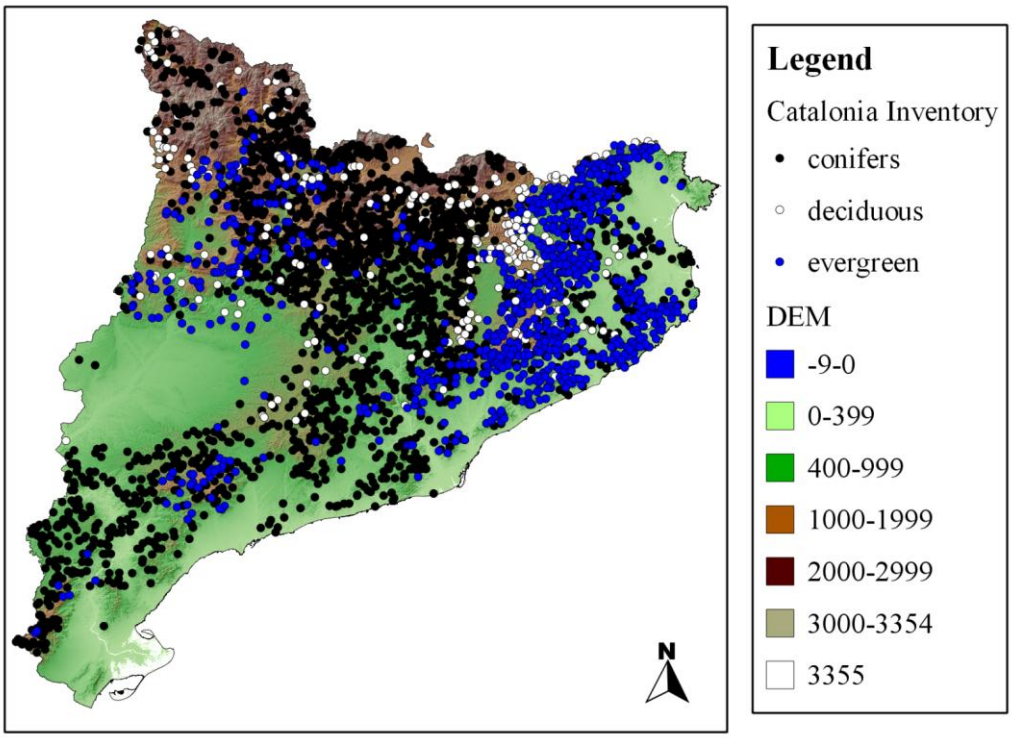


Fig. 4. 2 Spatial distribution of inventory vegetation types covered in the dataset, over a relief map of Catalonia.

4.3.2. Selection of NDVI threshold

As already discussed, the coarse spatial resolution of the MODIS is likely to result in a substantial error related to mixed pixels. A procedure was devised in order to select only pure pixels to be retained for the analysis. After computing NDVI and EVI for each pixel as already described, the regression between leaf N concentration and MODIS reflectance was tested in a recursive process, restricting the sample to meet increasing NDVI thresholds and looking at the resulting sample size and adjusted R^2; the procedure was applied for both multiple regression and stepwise regression.

A gradual increase in the NDVI threshold for inclusion in the analysis (Fig.4.3) resulted in sigmoidal decrease in sample size and an almost exponential increase of model adjusted R^2, so clearly demonstrating the significance of the error introduced as a result of MODIS coarse spatial resolution. If on the one hand the choice of an NDVI threshold of 0.9 would have resulted in the best R^2, but on the other hand this would have restricted the dataset to only 35 samples, which would not have been a significant dataset because of the great number of predicting variables involved (26). So, in order to obtain a significant dataset and a good regression model, we chose to impose for our analyses an NDVI threshold of 0.7, resulting in a sample size of 736 plots, and yielding a R^2 of about 0.37.

We repeated this analyses also using EVI as an index of vegetation cover and pixel homogeneity, obtaining for the same sample size (736), an EVI threshold of 0.3 and a R^2 of 0.33, lower than the

one obtained with NDVI. It was therefore decided to utilize NDVI as a proxy criterion to select pure vegetation pixels.

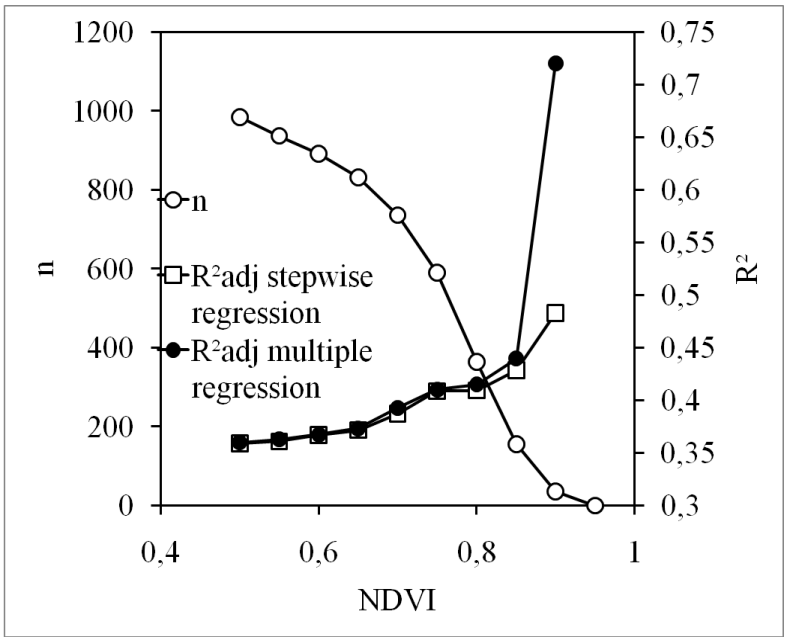


Fig. 4. 3 Interative procedure for the selection of an NDVI threshold and the reduction of mixed pixel effects. The results of multiple regression of leaf N concentration on MODIS reflectances are presented, when the dataset was restricted to meet increasing thresholds of NDVI; resulting sample size (white circles) is presented, together with the associated adjusted R^2 for multiple regression (multiple regression, black dots; stepwise multiple regression, white squares).

4.3.3. Regression models results

The relationship between leaf N concentration and MODIS reflectance was analyzed by (i) simple regression against NIR reflectance in band 2 (841-876 nm) (as in Ollinger et al., 2008), (ii) multiple regression, (iii) stepwise multiple regression and (iv) PLS regression against all possible MODIS bands. Results presented in Table 4.2 refer to the analysis of the restricted database of 707 samples, obtained after removing some outliers highlighted by the PLS procedure. Further details of the procedure and the associated results are presented in Appendix A. A more detailed analysis of results obtained from each methodology follows.

4.3.3.1. Regression against NIR reflectance

As a first step, a simple regression between leaf N concentration and MODIS NIR reflectance (band 2) was tested, as originally suggested by Ollinger (Ollinger et al., 2008).

This simple approach yielded the worst results (R^2=0.33; RMSE=0.48), although highly significant mainly as a result of the very large sample size. Moreover, the analysis suggested that most of correlation between N and remote sensing data was indeed due to NIR band.

Table 4. 2 Summary of results obtained when modeling foliar N concentration as a linear function of MODIS reflectance by simple regression, multiple regression, stepwise regression and PLS (Partial Least Squares) regression. The * indicates the R^2 is an adjusted R^2. The § indicates the coefficients are cross-validated.

Model	Response variable	Variables involved	Components	R^2	RMSE	P value	PRESS
Simple regression	NIR reflectance	1		0.33*	0.48	<0.0001	
Multiple regression	Full reflectance	11		0.38*	0.48	<0.0001	
Stepwise multiple regression	Full reflectance	6		0.38*	0.47	<0.0001	
PLS	Full reflectance	26	6	0.37§	0.47§		0.81
	Full reflectance	12	4	0.37§	0.47§		0.80

4.3.3.2. Multiple regressions

Both simple and stepwise multiple regressions resulted in better R^2 and RMSE compared to simple regression (R^2=0.38 in both cases; RMSE=0.48 and 0.47, respectively). Also in these analyses the importance of the NIR band was apparent but, as already observed by Wicklein et al. (2012), reflectance bands included in the best selected model covered almost all of the available reflectance (Appendix A). All the wavelengths involved were already known in the literature, some of them only in remote sensing and others in laboratory studies; a list of MODIS bands selected is presented in Table 4.3, together with a comparison with adjacent bands reported in the literature and of the associated features; in many cases, literature studies reported a relationship with other biochemical compounds (e.g. chlorophyll, carotenoids, lignin) which have been proposed to be more or less strongly related to N content, and sometimes even used as a proxy for its estimation.

4.3.3.3. PLS regressions

The final results of this more robust test were comparable with previous analyses; the greater robustness of PLS, which amounts to a proper validation of the explanatory model rather than a mere calibration, did not result in a substantial loss of predictive power, as also shown by Asner & Martin (2008). In fact the RMSE was the same as in multiple regression and the coefficient of determination was only slightly smaller (R^2=0.37, as opposed to 0.38 for multiple regression). Thanks to PLS analyses of matrix regression coefficients (B(PLS), B1 in Appendix A) and variable importance for the prediction (VIP in Appendix A), we could also reduce the number of variables

Table 4. 3 Wavelengths of MODIS bands involved in the best explanatory model selected by stepwise regression, compared with adjacent bands already proposed in the literature and their importance for estimating nitrogen or its proxies, at leaf or canopy level.

Wavelengths (nm)	Wavelengths in literature (nm)	Associated compound	Reference
405-420	418,430	Total chlorophyll	Curran et al. (2001), Curran (1989)
459-479	460, 470, 480	N (canopy level), carotenoids and chlorophyll	Serrano et al. (2002)
545-565	550-707, 505-550	Carotenoids and chlorophyll	Asner & Martin (2008), Gitelson et al. (2002)
841-876	841-876, 867, 876, 800-850	Lignin (canopy level), N (canopy level), leaf structure, canopy structure influence	Serrano et al. (2002), Ollinger (2008, 2011), Knyazikhin et al. (2012)
1628-1652	1649	N (canopy level)	Serrano et al. (2002)

Table 4. 4 Wavelengths of MODIS bands involved the best explanatory model selected in PLS regression, compared with adjacent bands already proposed in the literature and their importance for estimating nitrogen or its proxies, at leaf or canopy level.

Wavelengths (nm)	Wavelengths in literature (nm)	Associated compound	Reference
545-565	550-707, 505-550	Carotenoids and chlorophyll	Asner & Martin (2008), Gitelson et al. (2002)
841-876	841-876, 867, 876, 800-850	Lignin (canopy level), N (canopy level), leaf structure, canopy structure influence	Serrano et al. (2002), Ollinger (2008, 2011), Knyazikhin et al. (2012)
1230-1250	1200, 1208, 1216, 1241, 1252	Water, Lignin, Starch, Cellulose, N (canopy level)	Curran et al. (2001), Curran (1989), Serrano et al. (2002)
1360-1390	1400, 1404	Water, Chlorophyll b	Curran et al. (2001), Curran (1989)
1628-1652	1649	N (canopy level)	Serrano et al. (2002)
2105-2155	2130, 2157, 2172	N, N (canopy level)	Serrano et al. (2002), Curran et al. (2001)

involved, so as to simplify the model. Variables were dropped from the model when characterised by low B(PLS) and low VIP (SAS Institute Inc, 1999); only 12 variables were involved in the model, covering the spectral range explained in Table 4.4.

Looking at the variables of the PLS model, the importance of the visible area of the spectrum was reduced, compared to stepwise regression, while the NIR and SWIR bands became crucial. Most of the bands selected had already been proposed in remote sensing detection of N, while others were known to be tightly linked to leaf structure and chemistry. This particular fact suggested a questionable indirect effect of the relation between N and IR wavelengths (Knyazikhin et al., 2012; Ollinger, 2011).

Because of the greater robustness and equal predictive power of the PLS analysis, and considering that it is one of the most popular statistical procedures utilized in remote sensing studies, we decided to choose this model for the estimation of canopy N concentration at regional scale.

4.3.4. Nitrogen map

The maps generated by applying PLS coefficients to the bands involved in the model and overlapping it with the quality mask, gave a solid and rather consistent estimate of canopy N concentration also for pixels not involved in the model generation. Concerning the quality mask, obtained as explained in the Materials and Methods section (white areas in Fig. 4.4A), a regular pattern can be noticed in the central area of Catalonia. This was due to MODIS quality bands, in particular it was also seen in the "cloud state" band, and it could probably be due to a problem of the sensor in a particular acquisition date. The problem colud be eliminated by relaxing the quality requirements of the analysis, i.e. accepting also pixels deemed not reliable for some dates, although this would introduce an additional source of error in the estimate.

A visual inspection of Fig. 4.4 further demonstrates the problems stemming from MODIS spatial resolution. Where Fig. 4.4A shows areas of high predicted N concentration, the corresponding sample plots were also characterized by high observed values; while where Fig. 4.4B displayed samples with high N concentration, not always these correspond to a high concentration area in the spatialization map. Thus, the spatialization model worked fine at the low spatial resolution required for a large scale analysis, but at the finer scale required when studies need to look in more detail, (as in forest inventories, with 10 m radius plots), MODIS resolution appears not to be enough.

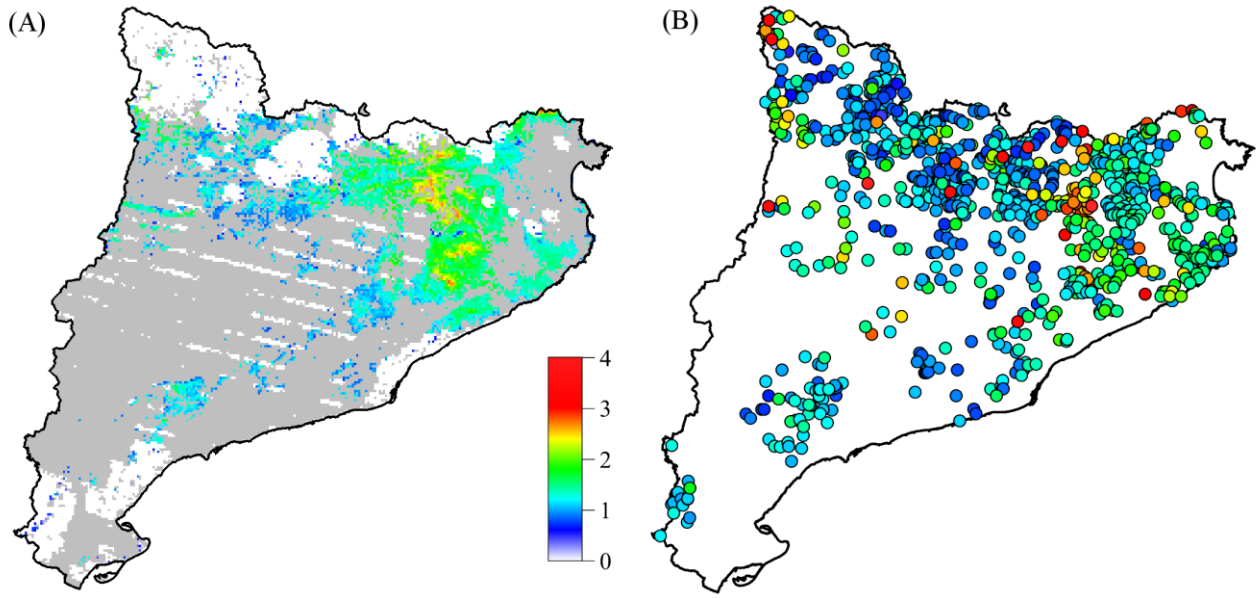


Fig. 4. 4 Maps of (A) spatialized N concentration, as predicted by the PLS procedure from May-July MODIS images, and (B) measured N concentration in May-July plots. Both are represented with the same legend in percent. White areas in map A correspond to quality masked pixels, selected as explained in Materials and methods section, while the grey areas refer to not-forest land-use types, obtained from the MODIS PFT classification.

4.3.5. Addressing the role of forest type

As already said before, the multiple regression and PLS models just discussed estimated N concentration without including as ancillary explanatory variable inventory vegetation types.

In general including also inventory vegetation types, as recorded in the inventory itself, all three regression models increased their R^2 to 0.71, while the RMSE declined to 0.31 (see Appendix A). Although the aim of the present study was to test the possibility to estimate N concentration from remote sensing data alone, these side results suggested that including a reliable estimate of vegetation types in the regression could really improve the estimate. Inclusion in the analysis of MODIS plant functional type (PFT) as an additional explanatory variable was not deemed appropriate, as MODIS PFTs are themselves estimated from pixel reflectance, so introducing an element of redundancy.

The overwhelming role of plant functional type, both on actual N concentration and on its prediction by remote sensing techniques, is demonstrated by a close analysis of PLS result (Fig. 4.5); visualizing the dataset with a graphical division between conifers, deciduous and evergreen broadleaves, differences between the three vegetation types were apparent. Without this division, the regression over the entire dataset seemed quite good, and with an homogeneous distribution of samples along the trend line. When discriminating between types, however, there was almost a clustering into three groups, consistent with what already reported by Mezzini et al. (2010) and discussed in the two previous chapters: coniferous plots were characterized by low N concentration

(both measured and estimated) as in Ollinger et al. (2008), deciduous samples were characterized by high N concentration and evergreen plots were in an intermediated position, with a strong leverage effect on the overall relationship.

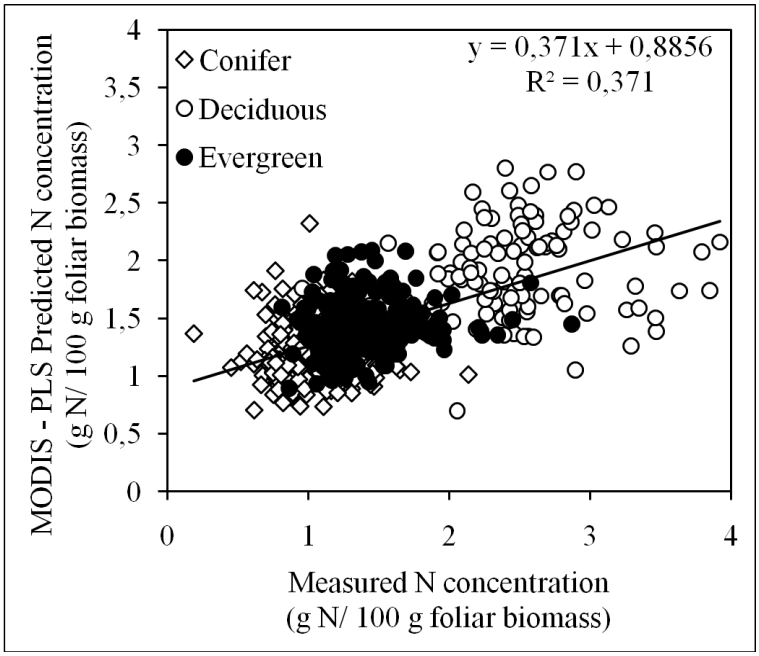


Fig. 4. 5 Plot of measured N concentration and PLS predicted N concentration, classified by inventory vegetation types.

This distribution suggested a possible effect of vegetation type on the relationship, as suggested by Knyazikihin et al. (2013): in practice, MODIS reflectance would estimate vegetation type, but because of the markedly different values of canopy N in conifers, deciduous and evergreen broadleaves this would also translate in and indirect prediction of this biochemical trait.

In order to test this possibility, data were further grouped by species rather than vegetation type (Fig. 4.6); only species comprising at least 10 inventory plots were considered in the analysis, for a total of 11 species. By averaging data within species, the white noise error associated with mixed pixel effects should be effectively removed, as forests should be randomly associated in MODIS pixels with areas of higher or lower reflectance. The existence of a direct relationship between canopy N and MODIS reflectance could be expected to result in a statistically significant correlation between measured and predicted N values across species within each vegetation type. A rather good relationship was observed for deciduous broadleaf species (R^2=0.81), but not for evergreen broadleaves (R^2=0.41) or conifers (R^2=0.07); the relationship, however, was in no case not statistically significant at the 5% level. A very good relationship (R^2=0.91***) was observed on the contrary when comparing all species, although this was largely the result of species clustering by vegetation type.

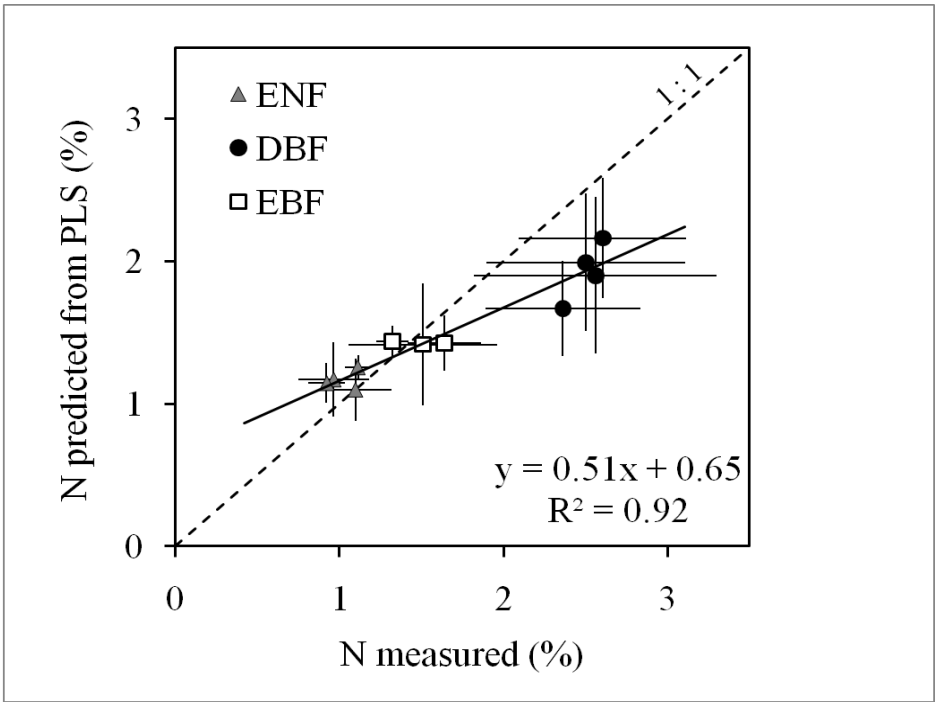


Fig. 4. 6 Relationship between canopy N concentration measured and predicted from MODIS reflectance by PLS for individual coniferous (ENF, evergreen needleleaf forest; grey triangles), deciduous (DBF, deciduous broadleaf forest; black dots) and evergreen (EBF, evergreen broadleaf forest; white squares) species; means ± SE. The regression refers to the entire dataset.

4.4. Conclusions

Beginning from methodological issues, the study demonstrated the possibility to use vegetation indexes, to select pure vegetation pixels and reduce the confounding effects of mixed pixels. Although the procedure would not remove problems associated with spectral mixing, of different vegetation types within a pixel, it would at least avoid the more serious effects of contamination with non-vegetated areas (urban, soil, etc.), which would be quite frequent during the summer in the Spanish landscape. The best procedure appeared to be the use of the classic NDVI instead of EVI, and with the application of a threshold of NDVI=0.7.

Coming to results, all three regression models (multiple, stepwise and PLS) gave approximately the same results, but because of the greater robustness of PLS, which lies behind its popularity in remote sensing analyses, we decided to apply this model. We applied it without inventory vegetation type variables, so as to demonstrate the potential of remote sensing *per se*, but we also saw that results could significantly improve when introducing also these ancillary variables. So in a real situation, whenever good inventory data concerning vegetation types are available, we suggest to utilize them.

Concerning the bands involved, NIR reflectance was indeed the most relevant band for the explanation of canopy N concentration; when used in isolation in simple regression, the R^2 was 0.33

compared with the final value of 0.37 when considering all available MODIS bands. In addition to NIR reflectance the model highlighted the importance of SWIR bands, known from the literature to be good predictors for chlorophyll and carotenoids contents (Curran et al., 2001; Curran, 1989; Serrano et al., 2002), and visible bands also widely applied for chlorophyll and carotenoids contents (Asner & Martin, 2008; Curran et al., 2001; Curran, 1989; Gitelson et al., 2002; Serrano et al., 2002).

The importance of NIR reflectance is consistent with its use for canopy N and lignin estimation from remote sensing in previous studies (Ollinger, 2011; Ollinger et al., 2008; Serrano et al., 2002). The question of the causal effects of the observed correlation, however, is still unresolved. Recent studies have demonstrated that the effect is not apparent at leaf scale, but is presumably the result of associated differences in canopy structure (Ollinger 2011; Wicklein et al., 2012; Knyazikhin et al., 2012). Knyazikhin et al. (2012) suggested that the observed correlation between canopy N and reflectance could be largely explained by differences between plant functional types, and by the relative contribution of coniferous species to the stand, generating as an indirect effect the positive correlation between NIR reflectance and N concentration. This appears to be largely consistent with the results of the present analysis.

Despite these reservations, the model was a good estimator of N concentration at landscape level (large scale), and after an accurate validation for the study region it could be therefore utilized as an input for ecological models, giving reference values for N concentration of different plant functional types in areas where inventory data are not available.

This work also highlighted the spatial resolution problems induced by the use of MODIS imagery as input data; for this reason, the analyses should be further refined with different sensors, with a better spatial resolution, like Landsat Thematic Mapper. NASA has just launched its new Landsat Data Continuity Mission (Landsat DCM or Landsat 8) which will cover our planet every 16 days, with spatial resolution of 30 m and some new bands (433-453 nm, 630-680 nm, 845-885 nm, 1360-1390 nm), all potentially useful for N estimation.

4.5. Acknowledgments

The PhD scholarship was funded by Italian MIUR, Bologna University and European Union project GHG Europe. We thank Universitad Autonoma de Barcelona, in particular CREAF center, for providing us inventory data with also nitrogen concentrations, and for hosting Elena Mezzini in their laboratories for three months during the year 2011; she also gratefully thanks Dr. Enrico Muzzi for introducing her to PLS analysis.

4.6. References

Asner, G., & Martin, R. (2008). Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment*, *112*(10), 3958–3970. doi:10.1016/j.rse.2008.07.003

Curran, P.J., Dungan, J. L., & Peterson, D. L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry:: Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment*, *76*(3), 349–359. Retrieved from http://www.sciencedirect.com/science/article/pii/S0034425701001821

Curran, Paul J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, *30*(3), 271–278. doi:10.1016/0034-4257(89)90069-2

Gitelson, A. A., Zur, Y., Chivkunova, O. B., & Merzlyak, M. N. (2002). Assessing Carotenoid Content in Plant Leaves with Reflectance. *Photochemistry and Photobiology*, *75*(3), 272–281. Retrieved from http://www.calmit.unl.edu/people/agitelson2/pdf/08_2002_P%26P_carotenoid.pdf

Gracia, C., Burriel, J. A., Ibañez, J., Mata, T., & Vayreda, J. (2004). *Inventari Ecològic i Forestal de Catalunya. Mètodes.* (p. 112). Bellaterra: CREAF. Retrieved from http://www.creaf.uab.es/iefc/pub/Metodes/index.htm

GRASS Development Team. (2012). Geographic Resources Analysis Support System (GRASS) Software, Version 6.4.3. Open Source Geospatial Foundation. Retrieved from http://grass.osgeo.org

Haenlein, M., & Kaplan, A. M. (2004). A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics*, *3*(4), 283–297. doi:10.1207/s15328031us0304_4

Hollinger, D. Y., Ollinger, S. V., Richardson, a. D., Meyers, T. P., Dail, D. B., Martin, M. E., Scott, N. a., et al. (2010). Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. *Global Change Biology*, *16*(2), 696–710. doi:10.1111/j.1365-2486.2009.02028.x

Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1228–1249. doi:10.1109/36.701075

Keithley, R. B., Heien, M. L., & Wightman, R. M. (2009). Multivariate concentration determination using principal component regression with residual analysis. *Trends in analytical chemistry TRAC*, *28*(9), 1127–1136. doi:10.1016/j.trac.2009.07.002

Knyazikhin, Y., Schull, M. a, Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., Marshak, A., et al. (2012). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences of the United States of America*, 1–8. doi:10.1073/pnas.1210196109

Mezzini, E., Raddi, S., Pippi, I., & Magnani, F. (2010). Telerilevamento della concentrazione di azoto fogliare in ecosistemi forestali litoranei. *attiasita.it* (pp. 1335–1338). Retrieved from http://www.attiasita.it/ASITA2010/Pdf/223.pdf

Neteler, M. (2010). Estimating Daily Land Surface Temperatures in Mountainous Environments by Reconstructed MODIS LST Data. *Remote Sensing*, *2*(1), 333–351. doi:10.3390/rs1020333

Niinemets, U., Kull, O., & Tenhunen, J. D. (1998). An analysis of light effects on foliar morphology, physiology, and light interception in temperate deciduous woody species of contrasting shade tolerance. *Tree physiology*, *18*(10), 681–696. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12651418

Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *The New phytologist*, *189*(2), 375–94. doi:10.1111/j.1469-8137.2010.03536.x

Ollinger, S. V., Richardson, A. D., Martin, M. E., Hollinger, D. Y., Frolking, S. E., Reich, P. B., Plourde, L. C., et al. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19336–41. doi:10.1073/pnas.0810021105

Ollinger, S. V., & Smith, M.-L. (2005). Net Primary Production and Canopy Nitrogen in a Temperate Forest Landscape: An Analysis Using Imaging Spectroscopy, Modeling and Field Data. *Ecosystems*, *8*(7), 760–778. doi:10.1007/s10021-005-0079-5

Pan, Y., Hom, J., & Jenkins, J. (2004). Importance of foliar nitrogen concentration to predict forest productivity in the Mid-Atlantic Region. *Forest Science*, *50*(3), 279–289. Retrieved from http://www.ingentaconnect.com/content/saf/fs/2004/00000050/00000003/art00002

Peñuelas, J., Sardans, J., Llusià, J., Owen, S. M., Carnicer, J., Giambelluca, T. W., Rezende, E. L., et al. (2010). Faster returns on "leaf economics" and different biogeochemical niche in invasive compared with native plant species. *Global Change Biology*, *16*(8), 2171–2185. doi:10.1111/j.1365-2486.2009.02054.x

Pons, X. (2009). *MiraMon. Geographic Information System and Remote Sensing software.* Barcelona, Spain: Centre de Recerca Ecològica i Aplicacions Forestals, CREAF.

Quantum GIS Development Team. (2012). Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved from http://qgis.osgeo.org

Sardans, J., & Peñuelas, J. (2012). Tree growth changes with climate and forest type are associated to relative allocation of nutrients, especially P, to leaves and wood. *Global Ecology and Biogeography*, n–a/n–a. In press.

SAS Institute Inc. (1999). The PLS Procedure. In SAS Institute Inc. (Ed.), *SAS/STAT ® User's Guide, Version 8*. Cary, NC.

SAS Support. (2012). Examples Using the PLS Procedure. SAS Support. Retrieved from http://support.sas.com/rnd/app/stat/papers/plsex.pdf

Serrano, L., Peņuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data:: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2-3), 355–364. Retrieved from http://www.sciencedirect.com/science/article/pii/S0034425702000111

Smith, M.-L., Ollinger, S. V., Martin, M. E., Aber, J. D., Hallett, R. a., & Goodale, C. L. (2002). Direct Estimation of Aboveground Forest Productivity Through Hyperspectral Remote Sensing of Canopy Nitrogen. *Ecological Applications*, *12*(5), 1286–1302. doi:10.1890/1051-0761(2002)012[1286:DEOAFP]2.0.CO;2

Tobias, R. D. (1995). An Introduction to Partial Least Squares Regression. *Analysis*, *2003*(16/12/2003), 1–8. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:An+Introduction+to+Partial+Least+Squares+Regression#0

Turner, D. P., Ollinger, S. V., & Kimball, J. S. (2004). Integrating Remote Sensing and Ecosystem Process Models for Landscape- to Regional-Scale Analysis of the Carbon Cycle. *BioScience*, *54*(6), 573. doi:10.1641/0006-3568(2004)054[0573:IRSAEP]2.0.CO;2

Van Der Voet, H. (1994). Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, *25*(2), 313–323. doi:10.1016/0169-7439(94)85050-X

Villaescusa, R., & Díaz, R. (Eds.). (1998). *Segundo Inventario Forestal Nacional (1986-1996). España*. Madrid: Ministerio de Medio Ambiente, ICONA.

Villanueva, J. A. (Ed.). (2005). *Tercer Inventario Forestal Nacional (1997-2007)*. Madrid: Ministerio de Medio Ambiente.

Vilà, M., Vayreda, J., Gracia, C., & Ibáñez, J. J. (2003). Does tree diversity increase wood production in pine forests? *Oecologia*, *135*(2), 299–303. doi:10.1007/s00442-003-1182-y

Wicklein, H. F., Ollinger, S. V, Martin, M. E., Hollinger, D. Y., Lepine, L. C., Day, M. C., Bartlett, M. K., et al. (2012). Variation in foliar nitrogen and albedo in response to nitrogen fertilization and elevated CO(2). *Oecologia*, (2008). doi:10.1007/s00442-012-2263-6

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory …*, *58*(2), 109–130. doi:10.1016/S0169-7439(01)00155-1

4.7. Appendix A

4.7.1. SAS Results of simple regression with NIR

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Adjusted R-Square Selection Method

Number in Model	Adjusted R-Square	R-Square	Root MSE	Variables in Model
1	0.3326	0.3335	0.48196	b2_02

4.7.2. SAS Results of multiple regression

4.7.2.1. Only remote sensing variables

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Adjusted R-Square Selection Method

Number in Model	Adjusted R-Square	R-Square	Root MSE	Variables in Model
11	0.3782	0.3879	0.46518	aspect slope altit ndvi evi b1_02 b2_02 b3_05 b4_05 b6_05 b8_10

4.7.2.2. Also inventory vegetation types variables

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Adjusted R-Square Selection Method

Number in Model	Adjusted R-Square	R-Square	Root MSE	Variables in Model
12	0.7126	0.7175	0.31628	evergreen deciduous aspect slope altit ndvi b1_02 b2_02 b3_05 b4_05 b6_05 b8_10

4.7.3. SAS Results of stepwise multiple regression

4.7.3.1. Only remote sensing variables

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Stepwise Selection: Step 1

Variable b2_02 Entered: R-Square = 0.3335 and C(p) = 48.8807

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	81.94878	81.94878	352.79	<.0001
Error	705	163.76108	0.23229		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-0.19625	0.08731	1.17366	5.05	0.0249
b2_02	7.39510	0.39372	81.94878	352.79	<.0001

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable b4_05 Entered: R-Square = 0.3510 and C(p) = 31.1983

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	86.23564	43.11782	190.34	<.0001
Error	704	159.47422	0.22653		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.10729	0.11092	0.21197	0.94	0.3337
b2_02	8.20022	0.43061	82.14996	362.65	<.0001
b4_05	-14.02865	3.22482	4.28686	18.92	<.0001

Bounds on condition number: 1.2266, 4.9063

Stepwise Selection: Step 3

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Stepwise Selection: Step 3

Variable b6_05 Entered: R-Square = 0.3649 and C(p) = 17.4719

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	89.66089	29.88696	134.64	<.0001
Error	703	156.04897	0.22198		

```
Corrected Total         706      245.70986


                    Parameter     Standard
         Variable    Estimate        Error   Type II SS  F Value  Pr > F

         Intercept    0.05847      0.11050      0.06215     0.28  0.5969
         b2_02        6.39156      0.62745     23.03360   103.77  <.0001
         b4_05       -35.71180     6.37649      6.96251    31.37  <.0001
         b6_05        9.40359      2.39387      3.42524    15.43  <.0001

                    Bounds on condition number: 8.0678, 46.858
-----------------------------------------------------------------------------------------

                              Stepwise Selection: Step 4


          Variable b3_05 Entered: R-Square = 0.3694 and C(p) = 14.3869


                                 Analysis of Variance

                                      Sum of           Mean
         Source                DF       Squares         Square    F Value    Pr > F

         Model                  4      90.76840       22.69210     102.81    <.0001
         Error                702     154.94146        0.22071
         Corrected Total      706     245.70986


                    Parameter     Standard
         Variable    Estimate        Error   Type II SS  F Value  Pr > F

         Intercept    0.18793      0.12442      0.50355     2.28  0.1314
         b2_02        7.56201      0.81515     18.99461    86.06  <.0001
         b3_05       38.27216     17.08532      1.10752     5.02  0.0254
         b4_05       -60.84456    12.89611      4.91311    22.26  <.0001
         b6_05        9.26390      2.38787      3.32197    15.05  0.0001

                                The SAS System   16:59 Saturday, December 11, 2012  67

                              The REG Procedure
                                Model: MODEL1
                           Dependent Variable: Nitrogen

                              Stepwise Selection: Step 4

                    Bounds on condition number: 20.132, 186.39
-----------------------------------------------------------------------------------------

                              Stepwise Selection: Step 5


          Variable altit Entered: R-Square = 0.3739 and C(p) = 11.3322


                                 Analysis of Variance

                                      Sum of           Mean
         Source                DF       Squares         Square    F Value    Pr > F

         Model                  5      91.86933       18.37387      83.72    <.0001
         Error                701     153.84053        0.21946
```

```
Corrected Total         706      245.70986

                    Parameter     Standard
      Variable       Estimate        Error   Type II SS  F Value  Pr > F

      Intercept       0.32092      0.13754      1.19472     5.44  0.0199
      altit      -0.00008209   0.00003665      1.10092     5.02  0.0254
      b2_02          7.62421      0.81330     19.28589    87.88  <.0001
      b3_05         38.34138     17.03668      1.11152     5.06  0.0247
      b4_05        -59.28475     12.87822      4.65079    21.19  <.0001
      b6_05          8.26796      2.42223      2.55692    11.65  0.0007

                   Bounds on condition number: 20.191, 240.14
-----------------------------------------------------------------------------------------

                              Stepwise Selection: Step 6

            Variable b8_10 Entered: R-Square = 0.3773 and C(p) = 9.5442

                                 Analysis of Variance

                                      Sum of           Mean
Source                     DF        Squares         Square    F Value    Pr > F

Model                       6       92.69437       15.44906      70.67    <.0001
Error                     700      153.01549        0.21859
Corrected Total           706      245.70986

                                  The SAS System   16:59 Saturday, December 11, 2012  68

                                The REG Procedure
                                  Model: MODEL1
                            Dependent Variable: Nitrogen

                              Stepwise Selection: Step 6

                    Parameter     Standard
      Variable       Estimate        Error   Type II SS  F Value  Pr > F

      Intercept       0.43439      0.14918      1.85343     8.48  0.0037
      altit      -0.00009382   0.00003707      1.39994     6.40  0.0116
      b2_02          7.58519      0.81195     19.07729    87.27  <.0001
      b3_05         51.84385     18.36869      1.74131     7.97  0.0049
      b4_05        -64.83467     13.16646      5.30047    24.25  <.0001
      b6_05          8.92832      2.44124      2.92387    13.38  0.0003
      b8_10        -13.74286      7.07388      0.82504     3.77  0.0524

                   Bounds on condition number: 21.189, 320.66
-----------------------------------------------------------------------------------------

          All variables left in the model are significant at the 0.1500 level.

       No other variable met the 0.1500 significance level for entry into the model.

                              Summary of Stepwise Selection

        Variable    Variable    Number    Partial     Model
```

Step	Entered	Removed	Vars In	R-Square	R-Square	C(p)	F Value	Pr > F
1	b2_02		1	0.3335	0.3335	48.8807	352.79	<.0001
2	b4_05		2	0.0174	0.3510	31.1983	18.92	<.0001
3	b6_05		3	0.0139	0.3649	17.4719	15.43	<.0001
4	b3_05		4	0.0045	0.3694	14.3869	5.02	0.0254
5	altit		5	0.0045	0.3739	11.3322	5.02	0.0254
6	b8_10		6	0.0034	0.3773	9.5442	3.77	0.0524

4.7.3.2. Also inventory vegetation types variables

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Stepwise Selection: Step 1

Variable deciduous Entered: R-Square = 0.6351 and C(p) = 186.7766

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	156.04035	156.04035	1226.82	<.0001
Error	705	89.66951	0.12719		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	1.19976	0.01467	850.69683	6688.35	<.0001
deciduous	1.26854	0.03622	156.04035	1226.82	<.0001

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable evergreen Entered: R-Square = 0.7038 and C(p) = 21.2024

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	172.92807	86.46403	836.34	<.0001
Error	704	72.78179	0.10338		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	1.05552	0.01739	381.03022	3685.61	<.0001
evergreen	0.34235	0.02679	16.88772	163.35	<.0001
deciduous	1.41278	0.03455	172.88944	1672.32	<.0001

Bounds on condition number: 1.1195, 4.4778

Stepwise Selection: Step 3

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Stepwise Selection: Step 3

Variable altit Entered: R-Square = 0.7061 and C(p) = 17.5576

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	173.49694	57.83231	563.00	<.0001
Error	703	72.21292	0.10272		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	1.14265	0.04088	80.25194	781.26	<.0001
evergreen	0.29262	0.03405	7.58655	73.86	<.0001
deciduous	1.39387	0.03536	159.60536	1553.77	<.0001
altit	-0.00007229	0.00003072	0.56887	5.54	0.0189

Bounds on condition number: 1.8206, 13.884

Stepwise Selection: Step 4

Variable ndvi Entered: R-Square = 0.7081 and C(p) = 14.6733

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	173.98917	43.49729	425.75	<.0001
Error	702	71.72069	0.10217		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.74002	0.18791	1.58457	15.51	<.0001
evergreen	0.27769	0.03463	6.56842	64.29	<.0001
deciduous	1.36186	0.03816	130.10790	1273.49	<.0001
altit	-0.00007770	0.00003073	0.65297	6.39	0.0117
ndvi	0.52110	0.23741	0.49223	4.82	0.0285

The REG Procedure
Model: MODEL1
Dependent Variable: Nitrogen

Stepwise Selection: Step 4

Bounds on condition number: 1.8937, 24.354

Stepwise Selection: Step 5

Variable slope Entered: R-Square = 0.7093 and C(p) = 13.6628

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	174.29256	34.85851	342.16	<.0001
Error	701	71.41730	0.10188		
Corrected Total	706	245.70986			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.72165	0.18795	1.50201	14.74	0.0001
evergreen	0.28759	0.03506	6.85643	67.30	<.0001
deciduous	1.37209	0.03857	128.94926	1265.71	<.0001
slope	-0.00114	0.00065944	0.30339	2.98	0.0848
altit	-0.00006000	0.00003236	0.35021	3.44	0.0642
ndvi	0.57804	0.23936	0.59417	5.83	0.0160

Bounds on condition number: 1.9457, 37.74

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection

Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	deciduous		1	0.6351	0.6351	186.777	1226.82	<.0001
2	evergreen		2	0.0687	0.7038	21.2024	163.35	<.0001
3	altit		3	0.0023	0.7061	17.5576	5.54	0.0189
4	ndvi		4	0.0020	0.7081	14.6733	4.82	0.0285
5	slope		5	0.0012	0.7093	13.6628	2.98	0.0848

4.7.4. SAS Results of PLS regression

The PLS Procedure

Split-sample Validation for the Number of Extracted Factors

Number of Extracted Factors	Root Mean PRESS	Prob > PRESS
0	1.000206	<.0001
1	0.845842	<.0001
2	0.832659	0.0070
3	0.825523	0.0110
4	0.817377	0.0450
5	0.815602	0.0440
6	0.809876	0.2530
7	0.808732	0.3520
8	0.807951	1.0000
9	0.808537	0.3990
10	0.808182	0.4700
11	0.810105	0.2260
12	0.810357	0.1930
13	0.810928	0.1540
14	0.811739	0.1020
15	0.811744	0.1020

Minimum root mean PRESS	0.8080
Minimizing number of factors	8
Smallest number of factors with p > 0.1	6

The PLS Procedure

Percent Variation Accounted for by Partial Least Squares Factors

Number of Extracted Factors	Model Effects Current	Model Effects Total	Dependent Variables Current	Dependent Variables Total
1	25.2018	25.2018	29.4979	29.4979
2	34.7502	59.9520	2.2452	31.7431
3	6.5158	66.4678	2.1539	33.8970
4	6.6768	73.1447	1.3724	35.2694
5	7.3224	80.4671	0.9744	36.2438
6	3.4016	83.8687	0.9716	37.2154

Parameter Estimates for Centered and Scaled Data

	Nitrogen
Intercept	0.0000000000
b1_02	0.0217626744
b2_02	0.4727211937
b1_05	-.0504236399
b2_05	0.0479333802
b3_05	-.0606766863
b4_05	-.1829528464
b5_05	0.0334544361
b6_05	0.1255105971
b7_05	0.1101642110
b1_10	-.0465692667
b2_10	-.0189980196
b3_10	-.0622625513
b4_10	-.1734302534

b5_10	-.0299190811
b6_10	0.0930569862
b7_10	0.1159697321
b8_10	0.0455282432
b9_10	-.0271843098
b10_10	0.0364548791
b11_10	-.0396242797
b12_10	0.0096285744
b13_10	-.0889911005
b14_10	0.0452335468
b15_10	-.0241034876
b16_10	-.0241034876
b26_10	-.0162397846

The PLS Procedure

Parameter Estimates

	Nitrogen
Intercept	-4523.376099
b1_02	1.849167
b2_02	6.053260
b1_05	-4.236964
b2_05	0.646390
b3_05	-9.284011
b4_05	-17.544792
b5_05	0.522218
b6_05	3.519343
b7_05	5.329300
b1_10	-3.283126
b2_10	-0.276630
b3_10	-8.105792
b4_10	-15.527864
b5_10	-0.516630
b6_10	2.690455
b7_10	5.113512
b8_10	7.829172
b9_10	-0.416704
b10_10	0.455336
b11_10	-0.381462
b12_10	0.070173
b13_10	-0.097314
b14_10	0.045624
b15_10	-788.862786
b16_10	-788.862786
b26_10	-11.615641

4.7.4.1. Selection of superfluous bands

Obs	X_VAR	B1	VIP
1	B1_02	0.02176	0.50840
2	B2_02	0.47272	2.10088
3	B1_05	-0.05042	0.46884
4	B2_05	0.04793	1.82940
5	B3_05	-0.06068	0.42423
6	B4_05	-0.18295	0.68041
7	B5_05	0.03345	1.75633
8	B6_05	0.12551	1.30105
9	B7_05	0.11016	0.51477
10	B1_10	-0.04657	0.69354

11	B2_10	-0.01900	1.71918
12	B3_10	-0.06226	0.71315
13	B4_10	-0.17343	0.54823
14	B5_10	-0.02992	1.63690
15	B6_10	0.09306	1.03490
16	B7_10	0.11597	0.48963
17	B8_10	0.04553	0.48557
18	B9_10	-0.02718	0.39865
19	B10_10	0.03645	0.44928
20	B11_10	-0.03962	0.45766
21	B12_10	0.00963	0.42912
22	B13_10	-0.08899	0.71274
23	B14_10	0.04523	0.67690
24	B15_10	-0.02410	0.62979
25	B16_10	-0.02410	0.62979
26	B26_10	-0.01624	1.04424

The PLS Procedure

Split-sample Validation for the Number of Extracted Factors

Number of Extracted Factors	Root Mean PRESS	Prob > PRESS
0	1.000206	<.0001
1	0.860542	<.0001
2	0.828109	0.0010
3	0.813737	0.0140
4	0.803789	0.4910
5	0.803649	1.0000
6	0.804237	0.3790
7	0.804916	0.2620
8	0.806351	0.1030
9	0.807798	0.0540
10	0.806329	0.1510
11	0.806577	0.1270
12	0.806468	0.1400

Minimum root mean PRESS	0.8036
Minimizing number of factors	5
Smallest number of factors with p > 0.1	4

The PLS Procedure

Percent Variation Accounted for by Partial Least Squares Factors

Number of Extracted Factors	Model Effects Current	Model Effects Total	Dependent Variables Current	Dependent Variables Total
1	57.8740	57.8740	26.6333	26.6333
2	27.3759	85.2499	5.6712	32.3045
3	3.4456	88.6955	3.3439	35.6484
4	3.2652	91.9608	1.4497	37.0981

Parameter Estimates for Centered and Scaled Data

	Nitrogen
Intercept	0.0000000000
b2_02	0.5066403990
b2_05	0.0440762370
b4_05	-.2829084975
b5_05	0.0231127976
b6_05	0.1175404627
b7_05	0.0894064798
b2_10	-.0140944172
b4_10	-.2467652779
b5_10	-.0223830461
b6_10	0.1149406172
b7_10	0.1380553696
b26_10	-.0438945878

Parameter Estimates

	Nitrogen
Intercept	0.27409985
b2_02	6.48760007
b2_05	0.59437597
b4_05	-27.13032734
b5_05	0.36078709
b6_05	3.29585878
b7_05	4.32512443
b2_10	-0.20522862
b4_10	-22.09382532
b5_10	-0.38650099
b6_10	3.32315259
b7_10	6.08734504
b26_10	-31.39596859

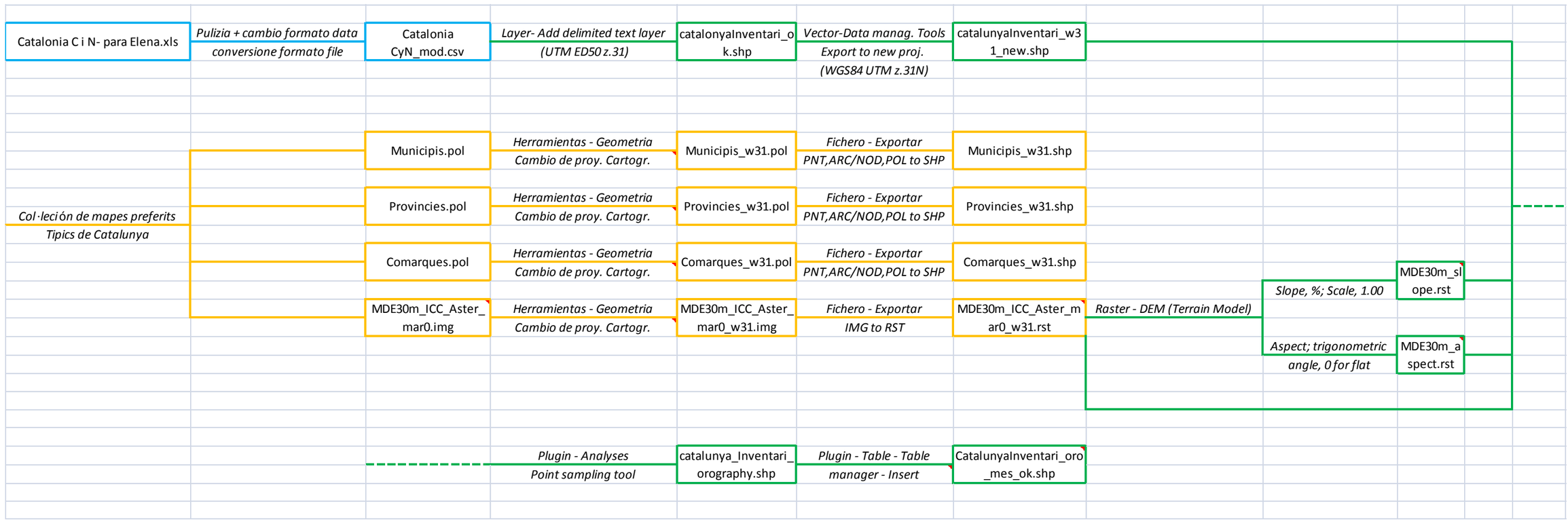
Obs	X_VAR	B1	VIP
1	B2_02	0.50664	1.48287
2	B2_05	0.04408	1.23025
3	B4_05	-0.28291	0.75343
4	B5_05	0.02311	1.17238
5	B6_05	0.11754	0.90687
6	B7_05	0.08941	0.65034
7	B2_10	-0.01409	1.15303
8	B4_10	-0.24677	0.76865
9	B5_10	-0.02238	1.09969
10	B6_10	0.11494	0.79648
11	B7_10	0.13806	0.72377
12	B26_10	-0.04389	0.90126

4.8. Appendix B

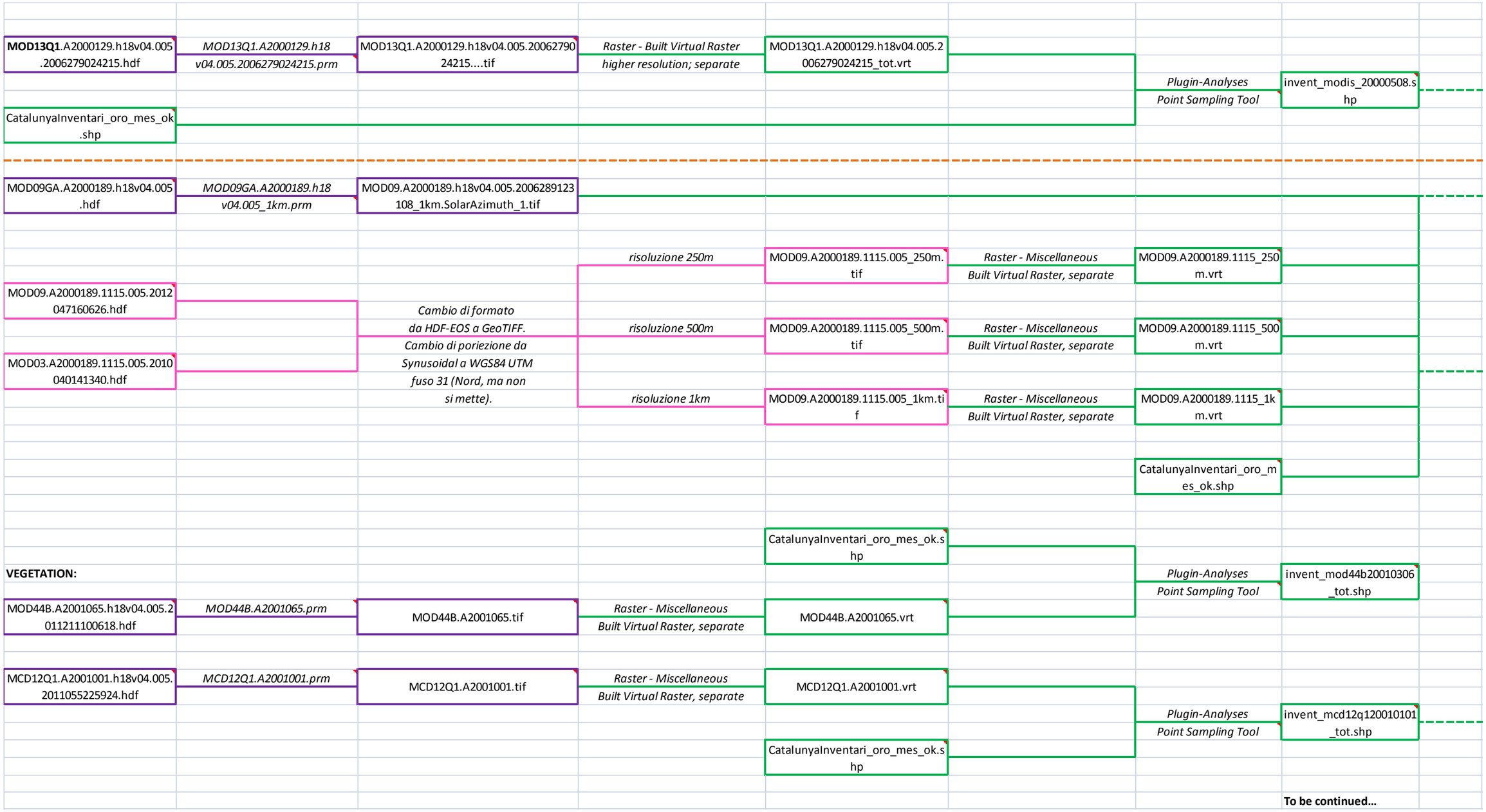
4.8.1. Software legend of following work-flow schemes

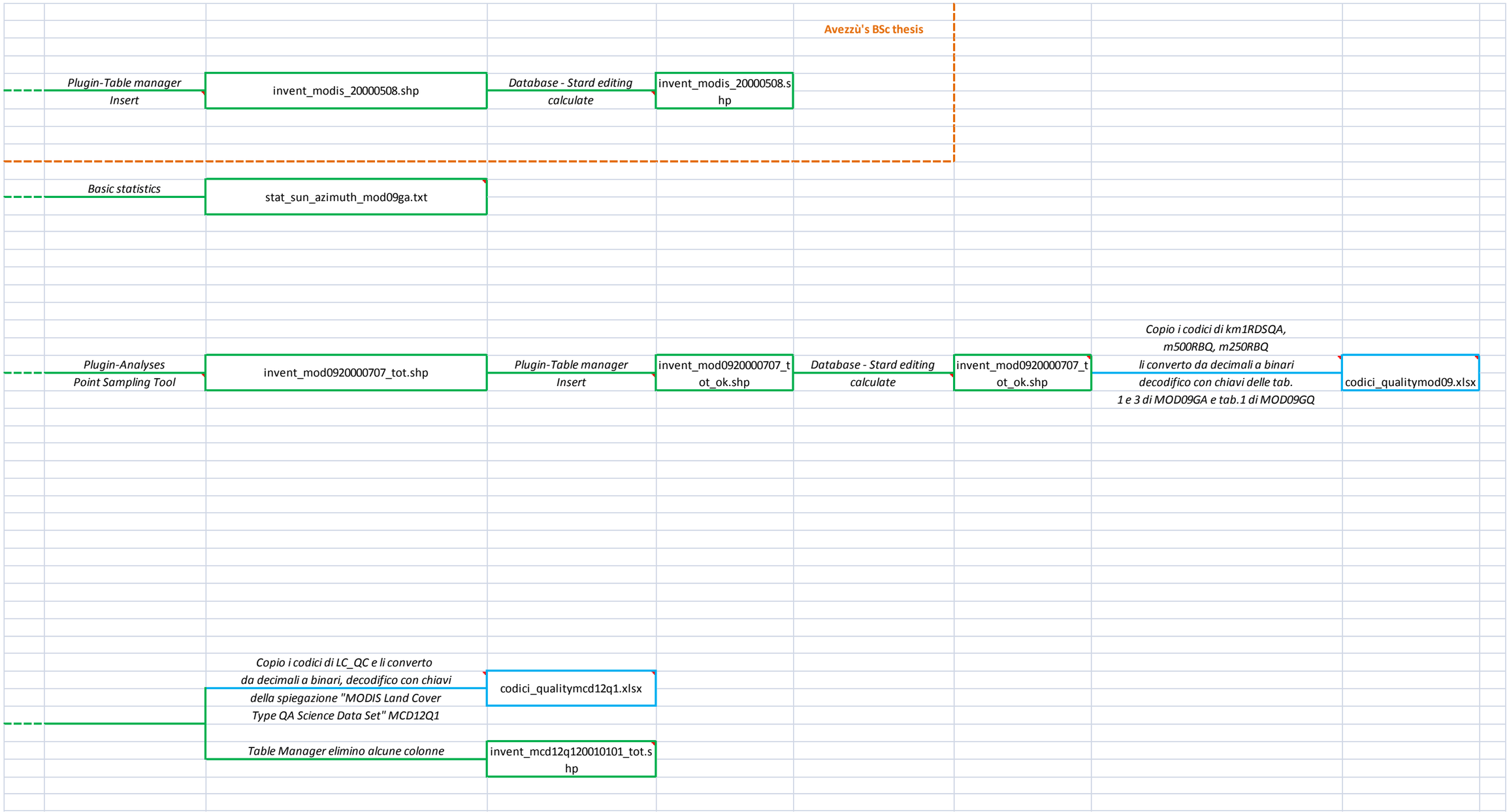
Quantum GIS	Open Source GIS Software		
MiraMon	CREAF GIS Software	GRASS	Open Source GIS Software
MRT MODIS Reprojection Tool	MODIS level L3 software	SAS	Statistical software
MRTSwath MODIS Reprojection Tool Swath	MODIS level L2 or lower software	Excel	
ENVI	GIS Software, mainly for remote sensing		

4.8.2. Elaboration work-flow of field samples

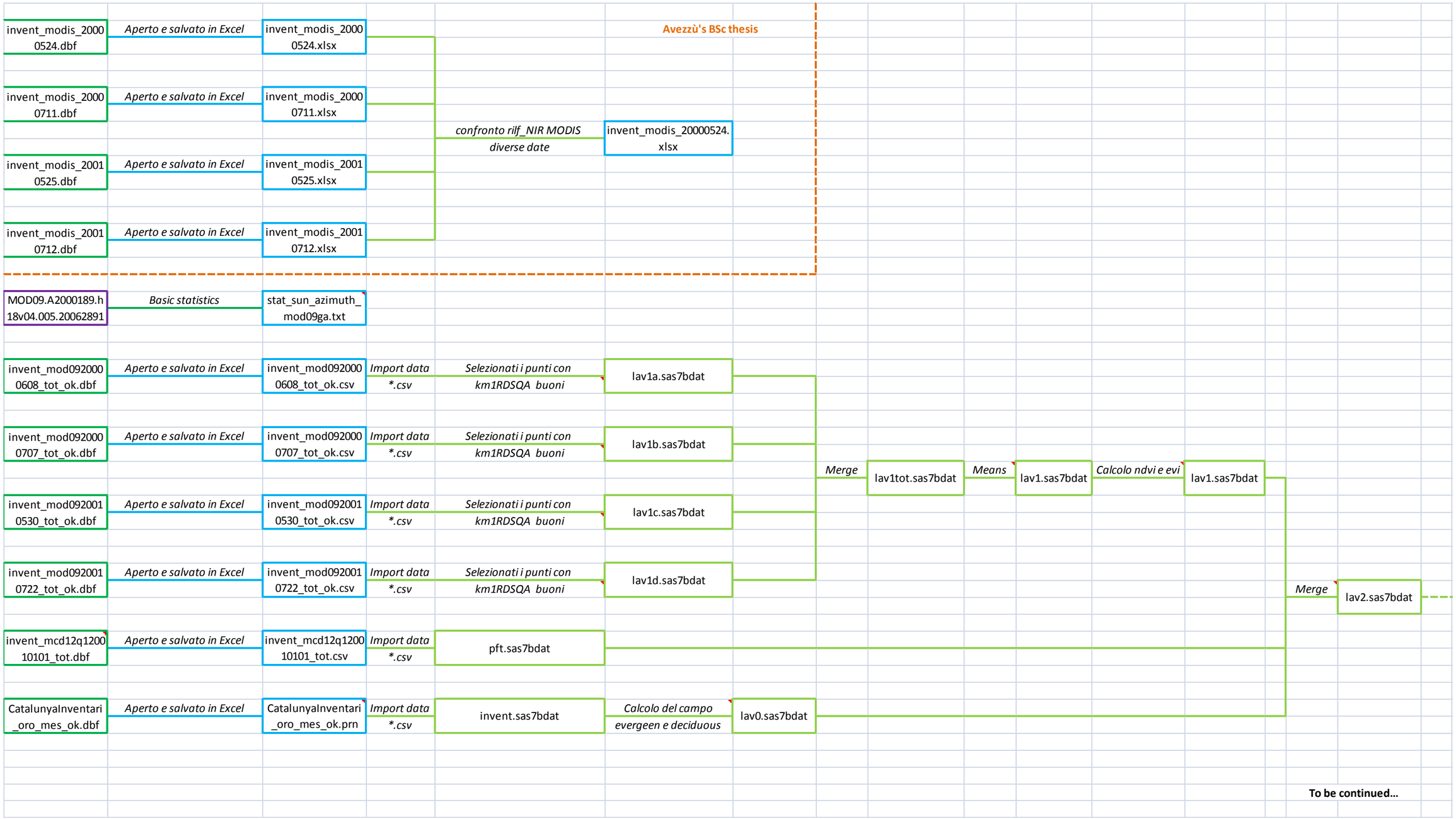


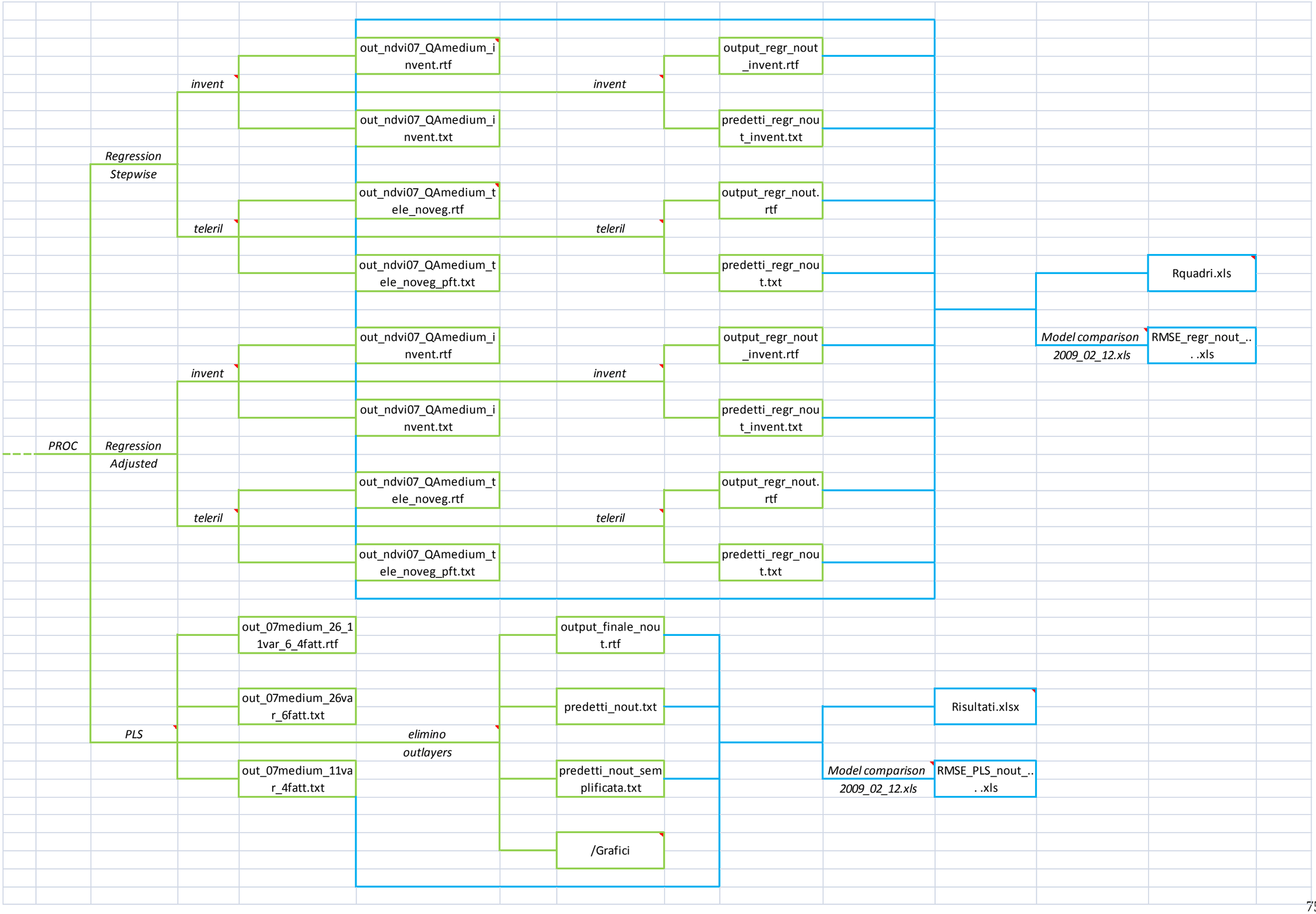
4.8.3. Work-flow of MODIS imgery preprocessing and sampling


MOD13Q1.A2000129.h18v04.005.2006279024215.hdf
MOD13Q1.A2000129.h18v04.005.prm
MOD13Q1.A2000129.h18v04.005.2006279024215....tif
Raster - Built Virtual Raster
higher resolution; separate
MOD13Q1.A2000129.h18v04.005.2006279024215_tot.vrt
Plugin-Analyses
Point Sampling Tool
invent_modis_20000508.shp
CatalunyaInventari_oro_mes_ok.shp
MOD09GA.A2000189.h18v04.005.hdf
MOD09GA.A2000189.h18v04.prm
MOD09.A2000189.h18v04.005.2006289123108_1km.SolarAzimuth_1.tif
risoluzione 250m
MOD09.A2000189.1115.005_250m.tif
Raster - Miscellaneous
Built Virtual Raster, separate
MOD09.A2000189.1115_250m.vrt
MOD09.A2000189.1115.005.2012047160626.hdf
Cambio di formato
da HDF-EOS a GeoTIFF.
Cambio di poriezione da
Synusoidal a WGS84 UTM
fuso 31 (Nord, ma non
si mette).
risoluzione 500m
MOD09.A2000189.1115.005_500m.tif
Raster - Miscellaneous
Built Virtual Raster, separate
MOD09.A2000189.1115_500m.vrt
MOD03.A2000189.1115.005.2010040141340.hdf
risoluzione 1km
MOD09.A2000189.1115.005_1km.tif
Raster - Miscellaneous
Built Virtual Raster, separate
MOD09.A2000189.1115_1km.vrt
CatalunyaInventari_oro_mes_ok.shp
CatalunyaInventari_oro_mes_ok.shp
VEGETATION:
Plugin-Analyses
Point Sampling Tool
invent_mod44b20010306_tot.shp
MOD44B.A2001065.h18v04.005.2011211100618.hdf
MOD44B.A2001065.prm
MOD44B.A2001065.tif
Raster - Miscellaneous
Built Virtual Raster, separate
MOD44B.A2001065.vrt
MCD12Q1.A2001001.h18v04.005.2011055225924.hdf
MCD12Q1.A2001001.prm
MCD12Q1.A2001001.tif
Raster - Miscellaneous
Built Virtual Raster, separate
MCD12Q1.A2001001.vrt
Plugin-Analyses
Point Sampling Tool
invent_mcd12q120010101_tot.shp
CatalunyaInventari_oro_mes_ok.shp
To be continued...

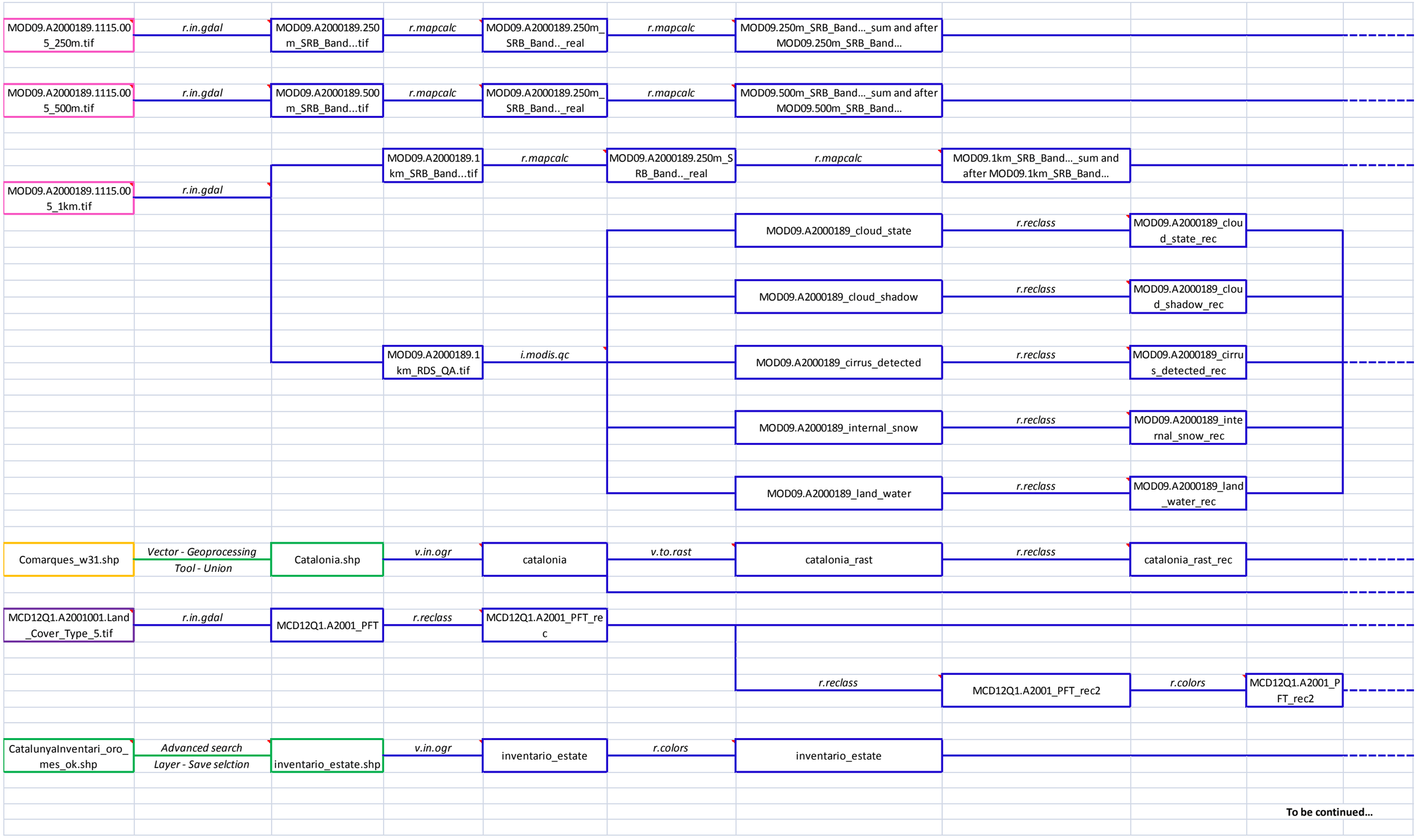

Avezzù's BSc thesis
Plugin-Table manager
Insert
invent_modis_20000508.shp
Database - Stard editing
calculate
invent_modis_20000508.shp
Basic statistics
stat_sun_azimuth_mod09ga.txt
Plugin-Analyses
Point Sampling Tool
invent_mod0920000707_tot.shp
Plugin-Table manager
Insert
invent_mod0920000707_tot_ok.shp
Database - Stard editing
calculate
invent_mod0920000707_tot_ok.shp
Copio i codici di km1RDSQA, m500RBQ, m250RBQ li converto da decimali a binari decodifico con chiavi delle tab. 1 e 3 di MOD09GA e tab.1 di MOD09GQ
codici_qualitymod09.xlsx
Copio i codici di LC_QC e li converto da decimali a binari, decodifico con chiavi della spiegazione "MODIS Land Cover Type QA Science Data Set" MCD12Q1
codici_qualitymcd12q1.xlsx
Table Manager elimino alcune colonne
invent_mcd12q120010101_tot.shp

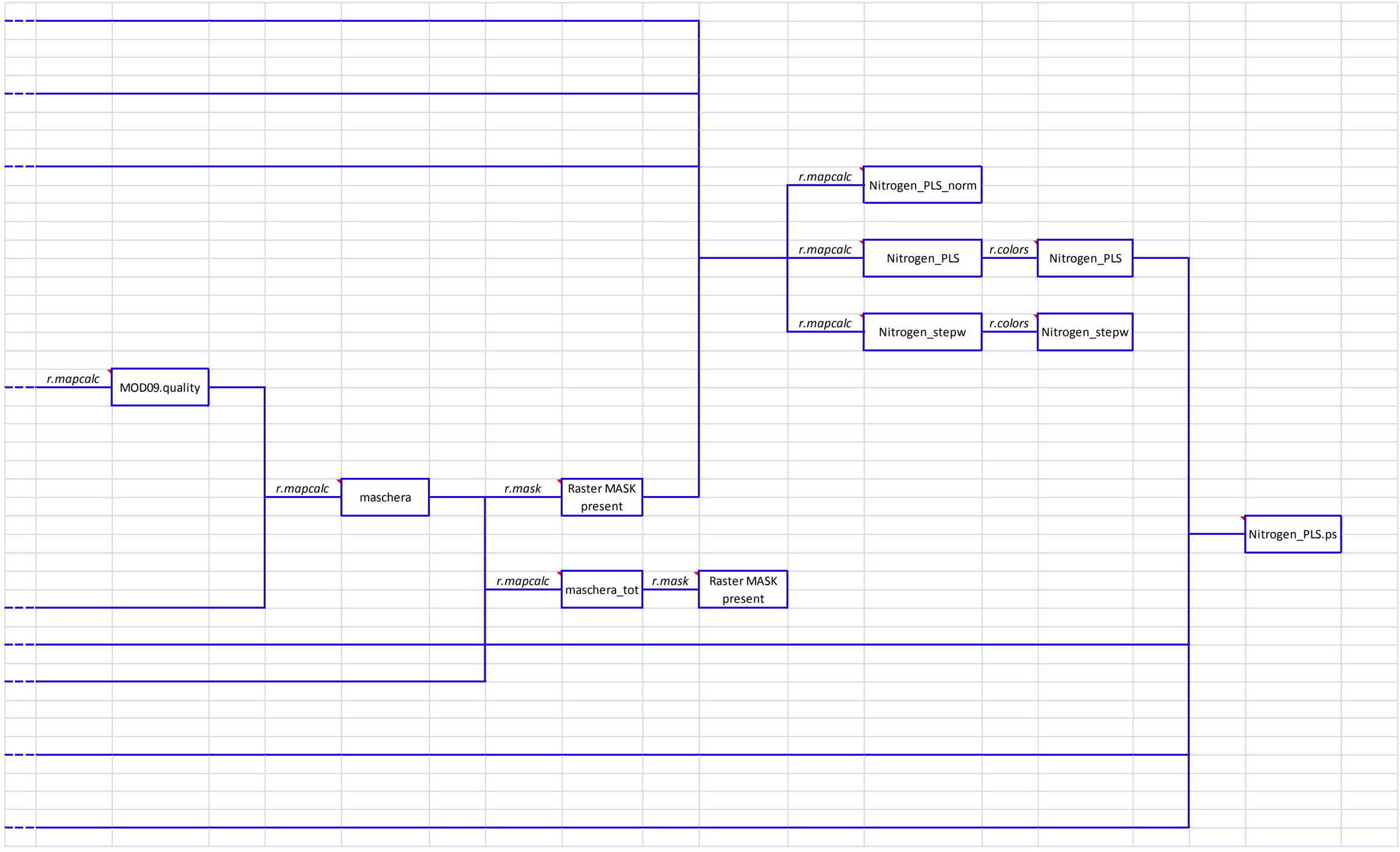
4.8.4. Work-flow of statistical analyses




out_ndvi07_QAmedium_invent.rtf
invent
out_ndvi07_QAmedium_invent.txt
Regression Stepwise
out_ndvi07_QAmedium_tele_noveg.rtf
teleril
out_ndvi07_QAmedium_tele_noveg_pft.txt
out_ndvi07_QAmedium_invent.rtf
invent
out_ndvi07_QAmedium_invent.txt
PROC
Regression Adjusted
out_ndvi07_QAmedium_tele_noveg.rtf
teleril
out_ndvi07_QAmedium_tele_noveg_pft.txt
out_07medium_26_11var_6_4fatt.rtf
out_07medium_26var_6fatt.txt
PLS
out_07medium_11var_4fatt.txt
elimino outlayers
output_regr_nout_invent.rtf
invent
predetti_regr_nout_invent.txt
output_regr_nout.rtf
teleril
predetti_regr_nout.txt
output_regr_nout_invent.rtf
invent
predetti_regr_nout_invent.txt
output_regr_nout.rtf
teleril
predetti_regr_nout.txt
output_finale_nout.rtf
predetti_nout.txt
predetti_nout_semplificata.txt
/Grafici
Model comparison 2009_02_12.xls
Rquadri.xls
RMSE_regr_nout_...xls
Model comparison 2009_02_12.xls
Risultati.xlsx
RMSE_PLS_nout_...xls

4.8.5. Nitrogen maps




r.mapcalc
Nitrogen_PLS_norm
r.mapcalc
Nitrogen_PLS
r.colors
Nitrogen_PLS
r.mapcalc
Nitrogen_stepw
r.colors
Nitrogen_stepw
r.mapcalc
MOD09.quality
r.mapcalc
maschera
r.mask
Raster MASK present
Nitrogen_PLS.ps
r.mapcalc
maschera_tot
r.mask
Raster MASK present

5. GENERAL CONCLUSIONS

As part of my Doctoral study I have been able to address the issue of the remote detection of canopy N concentration in forest stands, its potentials and problems, under many overlapping perspectives, exploiting the availability of three independent datasets acquired in Italian and Spanish forests at different scales and from different platforms. The three parts of this study, described in detail in Chapters 2, 3 and 4, made it possible to obtain a rather coherent picture, and to draw some general conclusions.

Starting from the analysis of S. Rossore 2000 dataset (Chapter 2), I tested several regressions between N concentration and NIR reflectances as derived from different sources (field samples, airborne sensors and satellite sensors). The analysis was further expanded using a larger dataset acquired in year 2009 as part of a new campaign funded by the European Space Agency (Chapter 3).

In both cases, a good correlation was observed, in particular, between Landsat NIR, using both Thematic Mapper (2009 data) and ETM+ (2000 data) imagery, and N concentration measured by a CHN elemental analyzer. In the S.Rossore 2000 dataset a problem of clustering in the distribution of data was unfortunately apparent. This problem was solved through the increase of the sample size (and the inclusion of a larger number of species and vegetation types) in the S. Rossore 2009 dataset, giving us a much more solid and promising result, because it confirmed the possibility to estimate N concentration from multi-spectral remote sensing at moderate spatial resolution (30 m).

Unfortunatly I did not obtain the same good results with the airborne remote sensing experience. In this case some problems were apparent both in MIVIS (2000 campaign; Chapter 2) and in CASI imagery (2009 campaign; Chapter 3), mainly due to the large FOV of these airborne sensors, and to the anisotropy of vegetation reflectance; the problem was addressed in the S. Rossore 2009 experience by selecting ROIs only near the centre of CASI's frame; the analysis was impaired by other problems still being investigated, highlighted by a comparison with ground-based canopy reflectance measurements. We also tested the relation between ground based ASD measures and nitrogen concentration, obtaining really good results. Based on results obtained at the S. Rossore site, I decided to expand my study to the regional level, focusing only on field and satellite measures.

For this purpose, I was able to analyze a large dataset for the whole of Catalonia, Spain; MODIS imagery was used in this new analysis, in consideration of its spectral characteristics and despite its rather poor spatial resolution (Chapter 4). Also in this case a regression between nitrogen concentration and reflectances was found, but not so good as in the previous experiences, and not

exclusively with NIR wavelengths. Moreover, vegetation type was found to play an important role in the observed relationship; this is consistent with modelling results and hypotheses recently proposed in the literature.

In particular, a recent study by (Knyazikhin et al., 2012) considered the BRF (Bidirectional Reflectance Factor) and DASF (Directional Area Scattering Factor) of vegetation, their relationship with vegetation type and their effect on canopy N reflectance and albedo. The hypothesis was put forward that the observed correlation between canopy N concentration and reflectance (Ollinger et al., 2008; Hollinger et al., 2009) could be the indirect result of these relationships, and of the wide differences in N concentration and structure between vegetation types, in particular conifers vs. other species.

A few general conclusions can be drawn from the study. Airborne sensors have problems associated with their large FOV and is necessary to take this into account before they can be used for estimation of nitrogen from airborne measurements of reflectance (as opposed to the use of normalized indexes).

Moreover we concluded that MODIS is not the most suitable satellite sensor in realities like Italy and Catalonia, which present a patchy and inhomogeneous vegetation cover; because of mixed pixel effects, average results could be utilized for the parameterization of eco-physiological and biogeochemical models, but not for really local nitrogen estimate. Thus multispectral sensors similar to Landsat Thematic Mapper, with better spatial resolution, could be the most appropriate sensors to estimate N concentration. For this reason, in the immediate future we would like to test also imagery from Landsat TM against the Catalonia dataset.

Finally, results were found to be partly associated with differences between vegetation types, in particular in their canopy structure. Differences in leaf and, more crucially, canopy structure have been suggested to be the main driver behind the observed variability in forest albedo, implying that the association with N concentration could be an indirect (albeit real; see Wicklein et al., 2012) effect.

Although this lack of an understanding of the causal mechanism does not subtract to the heuristic value of the results presented above, it is hoped that the issue will be resolved in coming years, so as to provide a more solid foundation for the remote detection of vegetation N concentrations.