NERC GEOPHYSICAL EQUIPMENT FACILITY

SCIENTIFIC REPORT - LOAN 819

USING MULTI-SCALE REMOTE SENSING TO STUDY HABITAT SELECTION BY CEREAL STEPPE BIRDS IN PORTUGAL

Osborne, P.E.[1], Leitão, P.J.[1,2] & Moreira, F.[2]

1 - Centre for Environmental Sciences, School of Civil Engineering and the Environment, University of Southampton, UK
2 - Centro de Ecologia Aplicada "Prof. Baeta Neves", Univ. Técnica de Lisboa, Portugal

Abstract

In May 2006, airborne remote sensing imagery was acquired on a pseudo-steppe area of south Portugal by NERC ARSF, as part of the Western Mediterranean Campaign. In this mission, multispectral (CASI) and laser altimetry (LiDAR) data were collected. After having been corrected for atmospheric, illumination and geometric effects, the spectral data can be used for describing the vegetation on the ground. The altimetry data were processed in order to generate a Digital Terrain Model and to map individual tree locations, and subsequently the distance to trees. These variables can be used to model the fine-scale habitats of the steppe bird community that occupies pseudo-steppe landscapes. The data will be incorporated into a multi-scale study (ranging from (this) local/patch scale to landscape and regional scales) to predict the distributions of pseudo-steppe birds.

Background

Portugal's cereal steppes hold many bird species of conservation concern. Their protection status is precarious, often relying on compensation payments to farmers for maintaining traditional practises. Although management prescriptions are based on information on habitat selection, it is often not appreciated that habitat selection and our knowledge of it are scale dependent. Policy is applied at large spatial scales whereas ecologists think and birds select habitats at (different) smaller scales.

The aim of this project is therefore to study habitat selection by steppe birds at multi-spatial scales to determine which features are important and how they link hierarchically from fine to coarse resolutions. Bird data are being modelled against habitat data derived from remotely sensed imagery and field studies on 1000m, 100m and 10m grids. SPOT VGT imagery provides the 1000m data, Landsat TM the 100m data, and airborne imagery the 10m data. The airborne data will also be used to calculate certain landscape variables at the 100m resolution. Models are built both separately for each spatial scale and hierarchically to resolve misclassifications of species occurrences at coarser resolutions. In the end this study will reveal which habitat features are important at different spatial scales, how the spatial resolution of habitat data limits our understanding of habitat selection, and the implications for conservation management.

For the fine scale component of this project airborne imagery (CASI, ATM and

LiDAR) was requested from NERC's Airborne Research & Survey Facility (ARSF). The CASI and ATM imagery can be used directly to describe the vegetation and soil moisture that compose the birds' habitats. Both these image sources are considered important because of CASI's high number of bands, specifically in the red edge areas (the CASI imagery requested were according to the default VEG band set), and ATM's information on the mid-infrared part of the spectrum, as well as the similarity of some of its spectral bands with the Landsat TM bands, this way facilitating a cross-scale interpretation. The LiDAR data is capable of producing information about the topography of the terrain and about the vegetation's vertical structure, both potentially good predictors for the studies species.

The study area is located in the south of Portugal, within the Castro Verde Special Protection Area (SPA) for birds, which is the main area of cereal steppes in the country, having national and international importance for populations of several steppe birds (Figure 1).

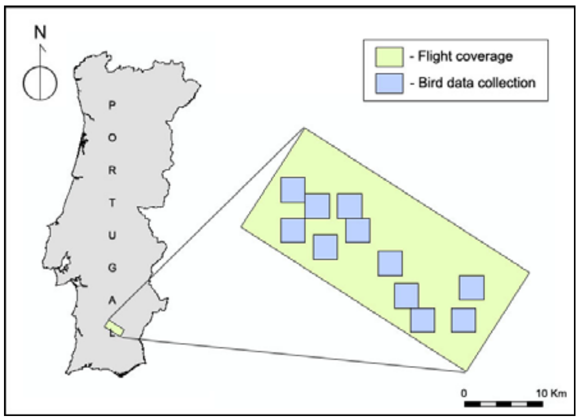

Figure 1 - Location of the study site and coverage of NERC ARSF's flight.

Field and airborne surveys

For the purpose of the bird data collection, 11 squares of 3x3km^2 were randomly located within the steppe areas of Castro Verde SPA. On each square a systematic grid of 100 locations was set where the birds' presence and absence was estimated using 5-minutes point counts.

All the 1100 point locations were recorded by use of a *Thales Mobile Mapper* DGPS receiver. At these locations, each individual bird observation within a 50m radius from the point count location was assigned to a 10x10m^2 pixel in a grid, to allow overlaying with the habitat layers.

An equipment loan was provided from NERC Geophysical Equipment Facility (GEF) in the form of a GPS receiver (*Leica System 1200 GPS* with antenna). This was mounted on top of a trig point (approximately at the centre of the flight area) and served as a reference GPS base-station to aid NERC ARSF airplane's navigation and posterior image processing.

The flight campaign took place on the 18th and 19th of May 2006 (Julian days 138 and 139). On day 138, 12 flight lines were flown between 10h30 and 13h00 covering all of the area with the CASI and the ATM sensors. Owing to the failure of the LiDAR sensor, only seven flight lines were covered by this sensor. To compensate for this, on day 139 the other five flight lines were re-flown between 12h30 and 14h00 covering only the southern part of the area with all three sensors. The average flight altitude for both days was 2250m above ground.

Field spectroscopy and aerosol optical thickness measurements were taken during the flights, through an equipment loan of a *GER1500* spectroradiometer and a *Microtops II* sunphotometer, provided by NERC Field Spectroscopy Facility (FSF). Field spectra were measured at three distinct surfaces (asphalt, bare soil, and green grass) on both days, to allow future calibration by using the Empirical Line method. Spectra measurements were also taken for characteristic land-uses in the area, such as cereal fields, fallows or shrublands, to aid in the imagery interpretation. Aerosol optical thickness measurements were taken at all of the previous locations.

Field vegetation height measurements were also taken in order to validate the LiDAR data. A total of 589 height measurements were taken on fallow and cereal fields and on shrub patches.

Data quality

The ATM sensor had numerous problems, resulting in data quality deficiency, as reported by the NERC ARSF:

- Problems with the optical encoder on the scan head main mirror motor shaft, resulting in occasional desynchronised scans;
- Electronic interference in the aircraft resulting in occasional noisy pixels (particularly on bands 9 and 11);
- A cooling issue that resulted in very poor quality data for the thermal bands (especially band 10);
- Existence of a small jitter in some of the data, where subsequent scan lines may be offset by 2 pixels;
- Existence of no overlap area between some flight lines resulting data gaps between successive lines in some areas.

For these reasons, the ATM data were considered unfit for the purpose intended and were not processed further.

On day 138 the weather was reported as overcast patchy mottled surface, which resulted in mottled shadow / bright sun imagery, with some band saturation on all multispectral data (Figure 2). The overall data quality reported on the flight log sheet was 5/10. On the following day (139) the weather was clear and the data were not affected by the above effects.

Nevertheless, only five flight lines were flown on this day (to complete the LiDAR coverage of the area) which means that only about half of the area was covered with good quality CASI data from this sensor, the imagery on day 138 being affected by the light conditions.

The LiDAR data obtained were of apparently good quality and covered all of the area. Height and Intensity information on both First and Last pulses were captured at an averaged point density of 1.9m and 1.8m (respectively for days 138 and 139).

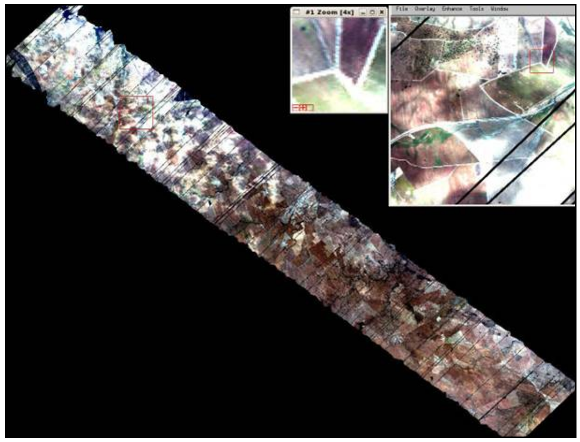

Figure 2 - ATM flight line 5 (day 138) showing the bad (desynchronised) scans as well as the mottled shadow / bright pattern resulting from the varying cloudy and sunny conditions.

Data Processing & Interpretation

Owing to the poor quality of the CASI imagery collected on day 138 (badly affected by cloud shadows), only data collected on day 139 has been processed.

The steps taken to pre-process this imagery were atmospheric correction, cross-track shading correction and geometric correction.

The atmospheric correction was implemented in ATCOR-4, using the Gaussian option and taking the upper and lower wavelength values in the HDF file as the half-power bandwidth. ATCOR-4 uses the radiance values in two near infra-red bands to retrieve the water vapour amount, which could be compared with the sunphotometer measurements collected at the time. The imagery was then processed using the following parameters:

Estimated visibility:	20 km
Mean ground elevation:	115 metres
Bands used for water vapour retrieval:	CASI bands 14 and 15
Aerosol model	Rural

At this stage the resulting reflectance values, even though seemly plausible, were not compared with the spectra measured in the field.

The cross-track shading corrections were done in ENVI, by applying a scaling to the data based on the pixel position relative to nadir. For this purpose, the multiplicative method was chosen and a fourth-order polynomial fitted to the column averages in each CASI band. The multiplicative method assumes that the cross-track variation is a function of the DN values, rather than simply an additive effect.

This method was generally successful in improving the visual appearance of the atmospherically-corrected data, with the exception of two flight lines (c13928 and c13929) where the sun-sensor geometry was such that the hot-spot peak was within the flight line.

AZGCORR (rel104pt) was used to geometrically correct the imagery. The corrected flight lines were then composed together into a mosaic in ENVI (Figure 3).

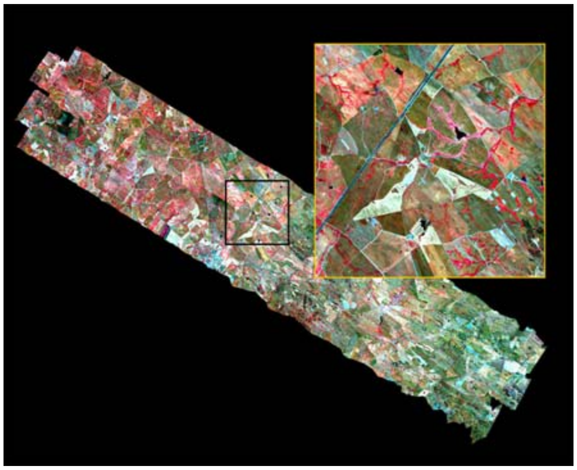

Figure 3 - The mosaic of the CASI data collected on 19th May (bands 14, 6, 3 [R,G,B]) after atmospheric correction and reduction of cross-track shading.

The data is now ready to be used for building the bird habitat models, which have yet to be done.

The LiDAR Height data was anticipated to be able to generate a DSM (Digital Surface Model - topography plus landscape features, vegetation, etc.) a DTM (Digital Terrain Model - topography), and a VHM (Vegetation Height Model - by subtracting the DSM with the DTM). The Intensity data should also have been able to depict different surfaces (such as different land uses). All of these could were to be used in the habitat models.

The LiDAR data was provided in point cloud form (for each flight line), which had to be split into the different signals (Last Pulse Height, Last Pulse Intensity, First Pulse Height and First Pulse Intensity), converted into grid (raster format) and finally mosaicked.

The gridding was done in the GEON points2grid Utility, provided by the Geosciences Network (GEON) Project (http://geongrid.org/). An output pixel resolution of 5m was selected.

Initially, both the First Pulse and Last Pulse Height signals were gridded by averaging the elevation values of the points that fell within the search radius, from the centre of the pixel. We have used the default search radius of *(square root of 2) / 2 times the resolution* of the grid.

When computing the VHM, as described above, the result was not satisfactory, as it was not capable of depicting known landscape features (such as shrub patches, etc.). It also returned unrealistic values (low heights) for trees. Additionally, the data showed no sensible pattern in relation to the ground (herbaceous) vegetation. By inspecting the original point cloud data it was observed that the two signals were too similar to be processed in this manner.

Thus, the First Pulse Height data was gridded by assigning the highest value (within the search radius) to the output pixel, in order to compute the DSM. In the same way, the Last Pulse Height data was gridded by assigning the pixel with the minimum value, to compute the DTM. The assumption behind this methodology is that the highest point within a grid pixel should correspond to the top of vegetation, and the lowest to the ground (Figure 4). The DTM was also used to calculate slope.

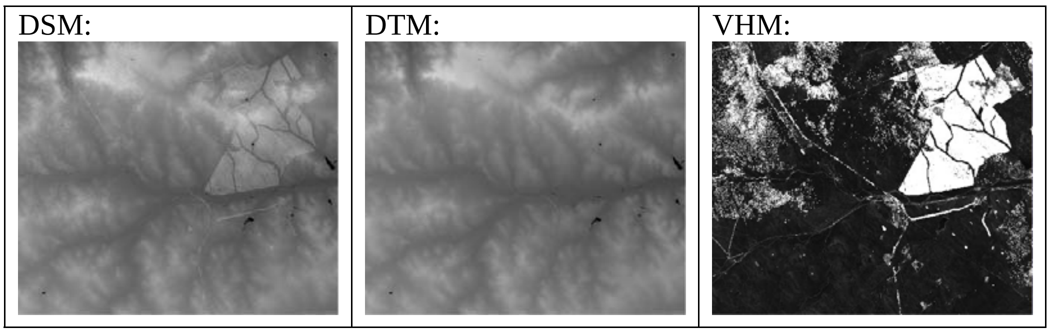

Figure 4 – Extract from the Digital Surface Model (DSM), the Digital Terrain Model (DTM) and the Vegetation Height Model (VHM), computed from the LiDAR point cloud data.

The VHM resulting from this second approach was suitable for identifying trees, with plausible heights. The gridding method used, however, is expected to be sensitive to the terrain variability within each pixel (as this will affect the maximum and minimum values existing within a pixel), which should be related to the terrain slope. In fact, this variable is highly correlated (Pearson $r = 0.844$) with slope.

When comparing these two images, the VHM showed some landscape features (such as trees) on top of the underlying slope pattern (Figure 5). Also, the resulting model could not reproduce the ground (herbaceous) vegetation measurements taken in the field campaigns.

Hence it was concluded that this method (applied to this dataset) is only capable of identifying vegetation features that are greater in height than the underlying terrain height differences. In the study area this corresponds to trees - eucalyptus, holm oaks and olive trees.

A *Tree* data layer was generated from the VHM by assuming every feature greater than 3 metres high to be a tree. This layer was used to calculate a *Distance-to-tree* variable, which was degraded to both 10m and 100m grids, in order to be used in the habitat models for these two scales. Also, the *Tree* layer was degraded to 100m grid this way generating a *Tree-density* variable (corresponding to the number of $5x5m^2$ pixels within the larger grid that have trees).

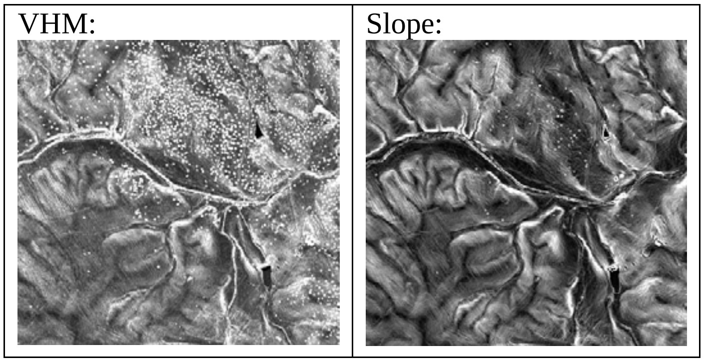

Figure 5 – Comparison between the VHM (stretched between 0 and 1m) and Slope: in the VHM it is visible the underlying slope pattern beneath the landscape features.

The First Pulse Intensity data was also gridded (to a $5x5m^2$ resolution) by averaging the values of the points within the grid search radius. Nevertheless, there seemed to be a decrease in the intensity of the signal from the centre to the edge of the data collection track (further from the flight path). This results in different Intensity values for the same land surfaces across the track on each flight line (Figure 6). This banding problem has not yet been solved.

Also, the Intensity signal was not equivalent on both flight days. The data collected on the second flight day (139) show intensity values greater than the data collected on the previous day (empirically estimated as $Dn_{139} = Dn_{138} + 6.5$).

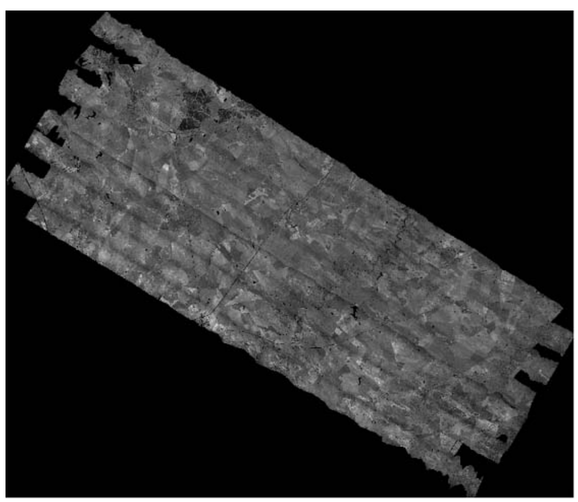

Figure 6 – LiDAR First Pulse Intensity signal mosaic (with correction from between- day differences), showing the banding resulting from the lower intensity values at the edges of the flight tracks.

Conclusions and recommendations

Overall, there were several problems relating to the data, from instrument failure (ATM) to poor quality of data retrieved due to weather conditions. All these have influenced the applications of the data, as well as the possible outcomes of the project.

Nevertheless, the CASI data collected on day 139 is of very good quality and will be very useful for modelling the presence of steppe birds in the study area at the desired scale.

The LiDAR data collected was unable to discriminate the herbaceous vegetation and it was concluded that for such a goal, in future missions, it should be collected at a much lower altitude (and perhaps in a smaller area), in order to capture the desired detail. However, it proved successful at generating a Digital Terrain Model, a subsequent Slope layer, as well as an accurate map of tree locations, this way enabling the creation of a distance-to-tree map. According to the literature, these layers are expected to be good predictors of occurrence for the bird species being studied.

At this stage, the bird models have not yet been concluded but they are anticipated to be capable of explaining the bird-habitat associations, and the data collected therefore able of describing the relevant habitats.

Publications

The habitat models are still being concluded, so there are no submitted publications.

An oral paper has been presented at the NERC ARSF Workshop, held during the 2007 Annual Conference of the Remote Sensing and Photogrammetry Society in Newcastle University, on the 11th September:

- Leitão, P.J., Milton, E.J., Mockridge, B., Osborne, P.E. & Moreira, F. *Pre-processing issues affecting the use of CASI and LiDAR data for steppe bird habitat monitoring and management in southern Portugal.*

Two other papers making use of these data are anticipated to be submitted for publication during 2008:

- One paper on the use of airborne remote sensing data for modelling fine scale bird-habitat associations; aimed at Remote Sensing of Environment.
- One paper on the multi-scale habitat-selection by steppe birds using remote sensing and map data; aimed at Journal of Applied Ecology.

Acknowledgements

Prof. Ted Milton and Bill Mockridge are thanked for their unstinting help in trying to resolve the considerable problems with the data and its processing. Funding for the flights was provided by EUFAR.