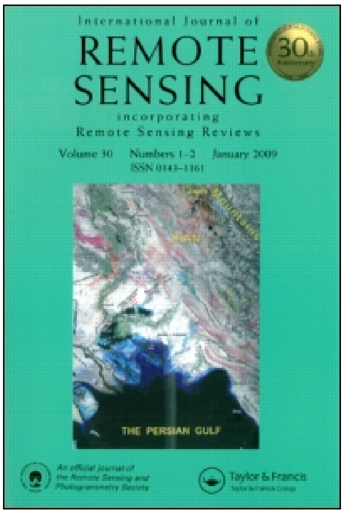
This article was downloaded by: *[Rocchini, Duccio]*
On: *26 May 2010*
Access details: *Access Details: [subscription number 922546915]*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

What is up? Testing spectral heterogeneity versus NDVI relationship using quantile regression

Duccio Rocchini[a]; Arianna Vannini[a]
[a] IASMA Research and Innovation Centre, Foundazione Edmund Mach, TN, Italy

Online publication date: 26 May 2010

To cite this Article Rocchini, Duccio and Vannini, Arianna(2010) 'What is up? Testing spectral heterogeneity versus NDVI relationship using quantile regression', International Journal of Remote Sensing, 31: 10, 2745 — 2756

To link to this Article: DOI: 10.1080/01431160903085651

URL: http://dx.doi.org/10.1080/01431160903085651

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

International Journal of Remote Sensing
Vol. 31, No. 10, 20 May 2010, 2745–2756

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

What is up? Testing spectral heterogeneity versus NDVI relationship using quantile regression

DUCCIO ROCCHINI*†‡ and ARIANNA VANNINI†
†IASMA Research and Innovation Centre, Fondazione Edmund
Mach, Environment and Natural Resources area Via E. Mach 1, 38010
S. Michele all 'Adige, TN, Italy
‡Via Pio La Torre 7, 52037, Sansepolcro, AR, Italy

(*Received 21 November 2007; in final form 10 July 2008*)

Environmental diversity and net primary productivity (NPP) are powerful indicators of local plant species richness (α-diversity). Remote sensing proxies of environmental diversity, such as spectral heterogeneity and NPP, are often used in modelling species richness variability, usually through regression analysis. As multicollinearity may affect analysis of species diversity, the interdependence of such proxies should be a major concern in their use. However, few attempts have been made to examine the interdependence between spectral heterogeneity and NPP proxies such as the Normalized Difference Vegetation Index (NDVI), in most cases using Ordinary Least Square (OLS) regression or Pearson correlations. We test the possible dependence of Landsat Enhanced Thematic Mapper (ETM+) local spectral heterogeneity versus NDVI using quantile regression and rejecting the main assumption of OLS regression, i.e. the symmetry of model residuals. A second-order polynomial function was fitted to the data and both OLS and quantile regression led to a humped-back relationship between spectral heterogeneity and biomass. Nonetheless while for most of the quantiles the humped-back curve was significant (with a negative and significant quadratic slope), for quantiles higher than 0.90, the parabola opened up until it reached an almost linear shape, showing that, at very low values of biomass, pixels may show high levels of local heterogeneity. Hence, patterns of spectral heterogeneity versus NDVI are possible when considering maximum potential spectral variability. We show that the investigation of all possible subsets within a scatter plot may lead to identification of patterns that remain hidden in OLS regression.

1. Introduction

The determination of biodiversity hotspots and species richness is often based on the measurement of environmental heterogeneity and net primary productivity (NPP) (Fairbanks and McGwire 2004). The use of these two variables is linked to two basic relationships between species richness and the environment. On the one hand, ecological variability in environmental properties should increase heterogeneity in habitat characteristics by allowing higher local species richness (Palmer and White 1994). On the other hand, energy availability, i.e. the availability of resources such as nutrient and water (strictly related to NPP), should shape richness gradients (Hawkins *et al.* 2003).

---

*Corresponding author. Email: duccio.rocchini@iasma.it

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2010 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903085651

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

Remote sensing has long been used to predict species-rich sites at a local scale (α-diversity) based on environmental diversity as derived from spectral heterogeneity (Gould 2000, Palmer *et al.* 2002, Foody and Cutler 2003, Rocchini *et al.* 2004) and NPP as derived from the Normalized Difference Vegetation Index (NDVI) (Whittaker and Heegaard 2003, Fairbanks and McGwire 2004, Gillespie 2006).

However, a major issue in such analyses is the interdependence between these potential explanatory variables, since problems of multicollinearity can affect the prediction of species diversity (Graham 2003, Meynard and Quinn 2007, Rocchini 2007b). To date, few attempts have been made to model the interdependence between spectral heterogeneity and NDVI (e.g. Gillespie 2006). Most attempts have been limited to Pearson coefficients or Ordinary Least Square (OLS) regression, considering the whole set of input data (hereafter referred to as the universe of points *U*). On the other hand, regression functions based on *U* could hide patterns that may be present for grouped data (Koenker and Hallock 2001).

In regression analysis, a parametric class of models is generally expected to describe a phenomenon by considering the whole distribution of input data within a regression model (Koenker and Bassett 1978). The best model is fitted by minimizing the residual sum of squares within a regression model for the mean of the dependent-variable distribution (Sokal and Rohlf 1995). However, many datasets are characterized by high variability within the scatter plot (Schröder *et al.* 2005, Rocchini 2007a), which can add noise to the regression model. This is particularly true if we consider the potentially very large number of available pixel values, which is a function of the grain (spatial resolution) and the extent (e.g. in terms of number of rows and columns) of the image being considered. In these cases, quantile regressions can be used to evaluate trends at different quantiles, rather than the whole distribution of input points, by constraining the regression model on a given quantile threshold τ (Koenker and Bassett 1978, Cade and Guo 2000, Cade and Noon 2003). Quantile regression may allow for a more comprehensive description of response patterns across a range of regression quantiles (Cade *et al.* 1999).

The present study examines the possible interdependence between local spectral heterogeneity and NDVI by rejecting one of the basic assumptions of OLS regression, i.e. the symmetry of model residuals (Koenker and Hallock 2001). We apply a quantile regression to investigate all possible trends of data considering different subsets *S*.

2. Methods

2.1 *Study area*

The Tuscany Region, in central Italy, is located between 9° and 12° E and between 42° and 44° N (WGS84). The region is characterized by a complex physical geography covering about 21 000 km^2 (figure 1). Following CORINE Land Cover data, 44% of the total area is covered by forests (broad-leaved forests), while crops cover about 46% (Bossard *et al.* 2000). Forested areas are characterized by different plant communities, varying from coastline evergreen Mediterranean forests dominated by *Quercus ilex*, to mountain beech and fir forests dominated by *Fagus sylvatica* and *Abies alba*. The climate ranges from typically Mediterranean to temperate warm and temperate cool in relation to altitudinal and latitudinal gradients (Rapetti and Vittorini 1995).

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

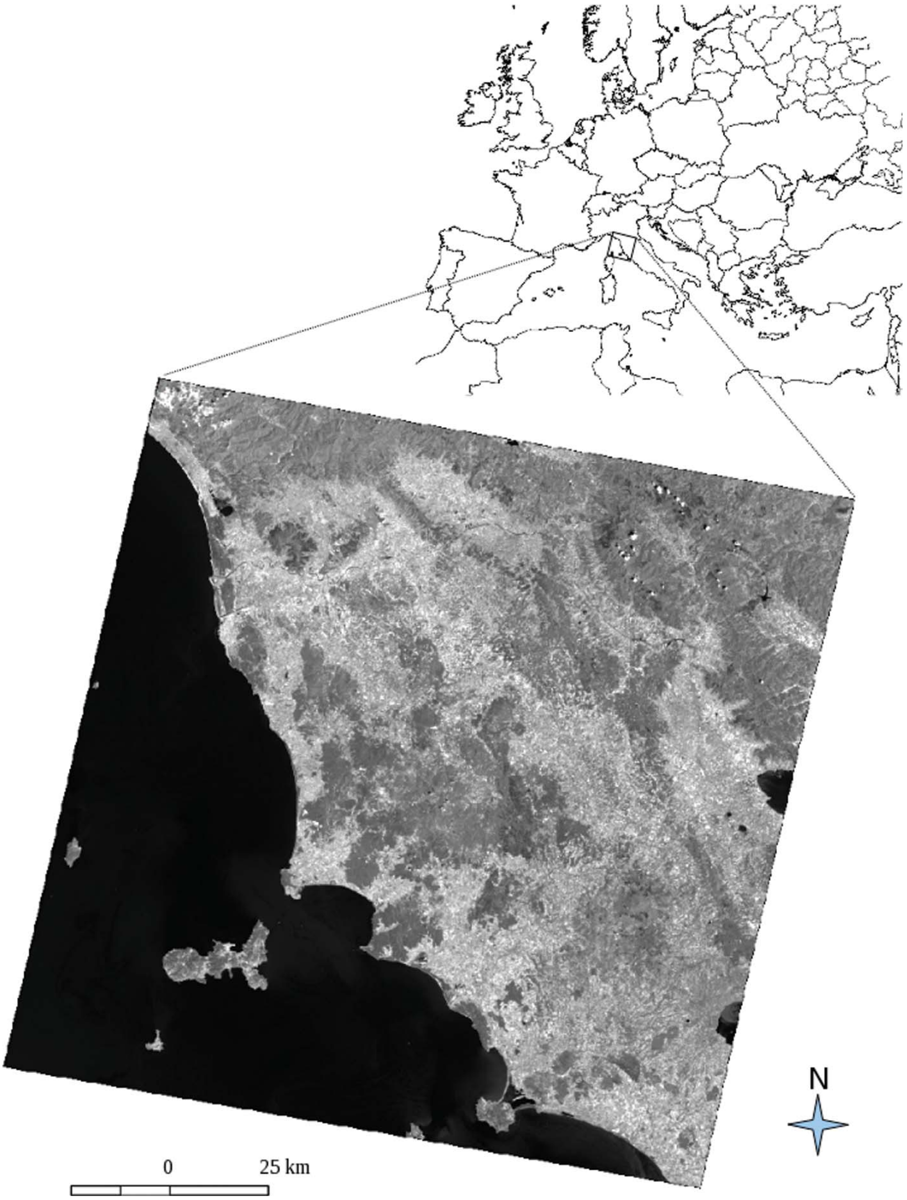


Figure 1. The ortho-Landsat ETM+ image (path 192, row 030) of Tuscany, Italy. PC1, explaining 86.53% of the total variance, is shown.

2.2 *Image acquisition and calculation of spectral variables*

An ortho-Landsat Enhanced Thematic Mapper (ETM+) image (path 192, row 030, acquisition date 20 June 2000; bands 1–5 and 7 spatial resolution 28.5 m) was acquired from the Global Land Cover Facility site hosted by the University of Maryland (htpp://glcfapp.umiacs.umd.edu, figure 1). We refer to Tucker *et al.* (2004) for complete information about image pre-processing.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

Spectral heterogeneity was derived as the spectral standard deviation within a moving window of 3 × 3 cells (Albani *et al.* 2004). Rather than directly using the whole multispectral dataset, an unstandardized Principal Component Analysis (PCA) was applied. PCA has long been used as a data compression tool that reduces the number of dimensions to be analysed, (see Jensen (1996) or Duda *et al.* (2001)). In this paper, the first PCA component (PC1), explaining 86.53%, was retained for the local (3 × 3 cells) spectral standard-deviation calculation.

In order to calculate the NDVI, band 4 (0.76–0.90 μm, near-infrared (NIR)) and band 3 (0.63–0.69 μm, red) were linearly combined as

$$\mathrm{NDVI} = \frac{\lambda_{\mathrm{NIR}} - \lambda_{\mathrm{R}}}{\lambda_{\mathrm{NIR}} + \lambda_{\mathrm{R}}}, \tag{1}$$

where λ_{NIR} is the reflectance in the NIR part of the spectrum and λ_{R} is the reflectance in the red part of the spectrum. The NDVI can vary from a theoretical minimum of –1 (minimum reflectance in the NIR and maximum in the red, low biomass) and a theoretical maximum of 1 (maximum reflectance in the NIR and minimum in the red, high biomass). Strictly speaking, the NDVI increases at high biomass due to (i) the high reflectance by vegetation in the NIR, which is linked to scattering processes at the leaf scale, and (ii) the low reflectance in the red spectrum due to the absorption by chloroplasts for photosynthesis.

In order to maintain the same scale of analysis with respect to the previously described PC1, mean NDVI within 3 × 3 cells was calculated and will be referred to as NDVI.

2.3 *Quantile regression fitting procedure*

Figure 2 shows a scatter plot of two variables ϖ_x and ϖ_y that appear to be related by a positive exponential model. However, when fitting an OLS-based model, the best fit is found to be a linear model (adjusted coefficient of determination $R^2 = 0.046$, slope = 0.37, probability threshold of regression models $p < 0.05$), rather than an exponential one (adjusted $R^2 = 0.010$, slope = 0.006, $p = 0.16$). Conclusions from the use of the 'mean' of the input data would therefore be that: (i) variables are weakly related to each other with a very low slope and (ii) variables are related to each other by a linear rather than an exponential function.

However, when considering the upper trend of values, the exponential relation between ϖ_x and ϖ_y becomes apparent (slope = 0.013, $p < 0.05$, $\tau = 0.75$), reaching higher values at $\tau = 0.90$ and 0.95, with slopes equalling 0.025 and 0.030 ($p < 0.01$), respectively. The linear pattern found by OLS regression, which obscures other possible patterns, is seen when the quantiles approach 0.5 (not shown in figure 2). The quantile at 0.5 (i.e. $\tau = 0.5$) represents the median of the input data.

Thus, a variable ϖ_y may show different response patterns with respect to ϖ_x when considering different subsets $s_{\omega y}$. A method considering different responses with respect to each $s_{\omega y}$ is strongly recommended to fully understand the phenomenon under study. Obviously, the same reasoning may be applied by considering lower trends in the data.

Formally, let $\{\varpi_{y1}, \varpi_{y2}, \ldots, \varpi_{yn}\}$ denote the n values of a set of points of the variable ϖ_y. OLS regression minimizes residuals by solving

$$\text{residual} = \min \sum \left(\varpi_{yi} - \hat{\varpi}_{yi}\right)^2, \tag{2}$$

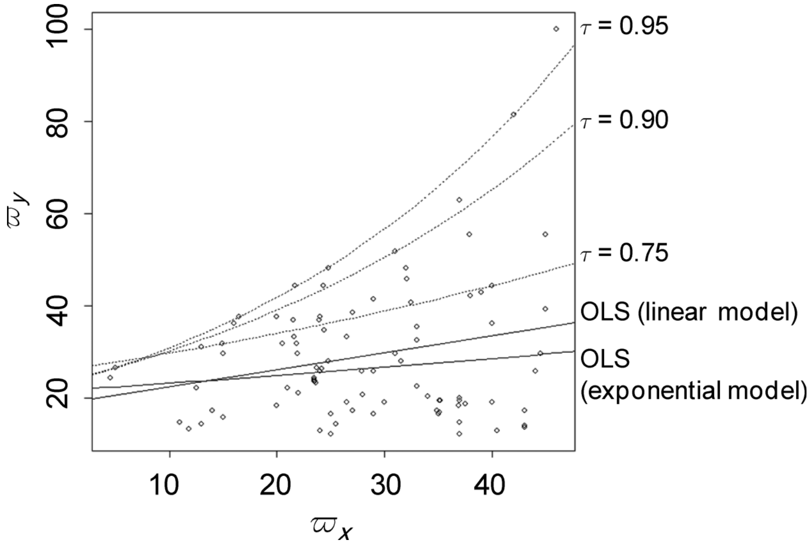


Figure 2. Worked example of OLS versus quantile regression. A linear pattern is found by fitting an OLS-based model. On the contrary, quantile regression may reveal previously hidden patterns such as an upper exponential relation (considering higher quantile values, τ) between the considered variables (ϖ_y and ϖ_x). See Cade and Noon (2003) for a similar example.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

where $\hat{\varpi}_{yi}$ is the estimated value for each *i*th point. Once the residual of each ϖ_{yi} value with its corresponding $\hat{\varpi}_{yi}$ value is calculated, the model is fitted using a symmetrical minimization of the achieved residuals.

Giving different weights to positive and negative residuals leads to an asymmetric minimization of residuals such that

$$\text{residual} = \min \sum \left|\varpi_{yi} - \hat{\varpi}_{yi}\right| T, \qquad (3)$$

where T is a multiplier term equalling τ (the quantile value) for positive deviations $\left(\varpi_{yi} - \hat{\varpi}_{yi}\right)$ and $(1 - \tau)$ for negative deviations. This asymmetric minimization fits a regression model through the upper data for high τ and through the lower data for small τ.

It should be noted that the quantile minimization of residuals (equation (3)) is based on absolute values rather than on squared deviations as in OLS regression, thus reducing outlier effects. For a more detailed description of quantile-based fitting, see Koenker and Hallock (2001), Cade and Noon (2003) and Gotelli and Ellison (2004).

2.4 *Building the model: OLS and quantile regression of local spectral variability versus NDVI*

We considered a subset of 1000 cloud-free random pixels. Pixels located in the sea were removed from the analysis, leaving a total number of 604 pixels for the analysis. The variability of local spectral heterogeneity (PC1 standard deviation within 3 × 3 cells) with respect to the NDVI (mean NDVI within 3 × 3 cells) was then tested using regression analysis.

The NDVI is expected to increase with increasing vegetation cover up to about 80–100%. Local spatial heterogeneity of NDVI, or any spectral measure, is expected to increase up to about 40–60% cover and then begin to decrease towards 100% (Woodcock and Strahler 1987). Consequently, a second-order polynomial function is expected to give the best fit. Second-order polynomial models at various quantile thresholds ($0.10 \leq \tau \leq 0.99$) were fitted to the data by the quantreg package (Koenker 2007) within R software (R Development Core Team 2007) following

$$h = \beta_1(\mathrm{NDVI})^2 + \beta_2(\mathrm{NDVI}) + h_0, \tag{4}$$

where h is the spectral heterogeneity, β_1 is the slope of the two sides of the parabola and β_2 is the coefficient basically regulating the coordinates of the centre of the parabola (together with β_1 and intercept h_0).

The comparison with the OLS linear regression was made by only considering the quadratic slope term β_1, which is expected to highly impact the slope of the parabola. In this paper $\beta_1 < 0$, i.e. as expected from the theoretical assumptions discussed above, a humped-back parabola fits the data.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

3. Results and discussion

OLS regression based on a parabola model led to a significant negative slope ($\beta_1 = -52$, $p < 0.01$, figures 3 and 4), indicating that a humped-back curve fits the relationship between spectral heterogeneity versus NDVI.

Considering quantile regression, β_1 showed low but significant values at lower quantiles. For quantiles approaching 0.65, β_1 reached its minimum ($\beta_1 = -70$), thus

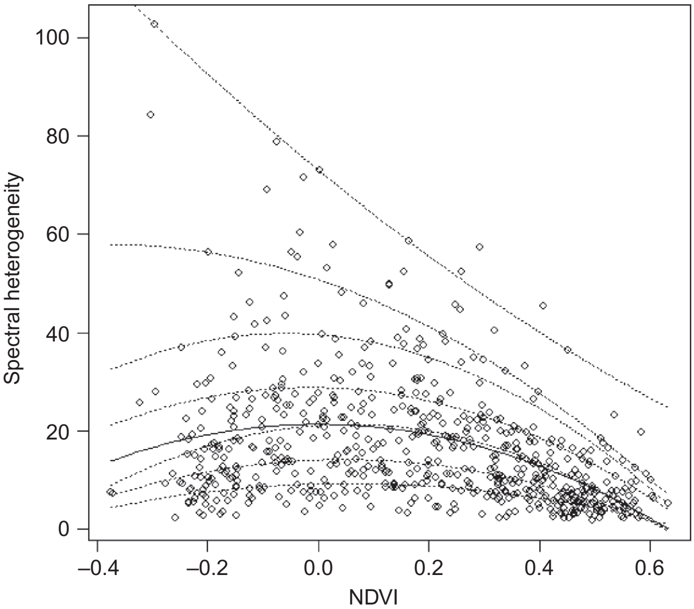


Figure 3. Scatter plot of spectral heterogeneity as measured by PC1 standard deviation (3 × 3 cells) versus NDVI (mean NDVI within 3 x 3 cells). Dotted lines: regression curves derived from quantile regression at different thresholds τ, The following quantiles are shown, from the top to the bottom: 0.99, 0.95, 0.90, 0.80, 0.60, 0.40, 0.20. Solid line: regression curve derived from OLS regression.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

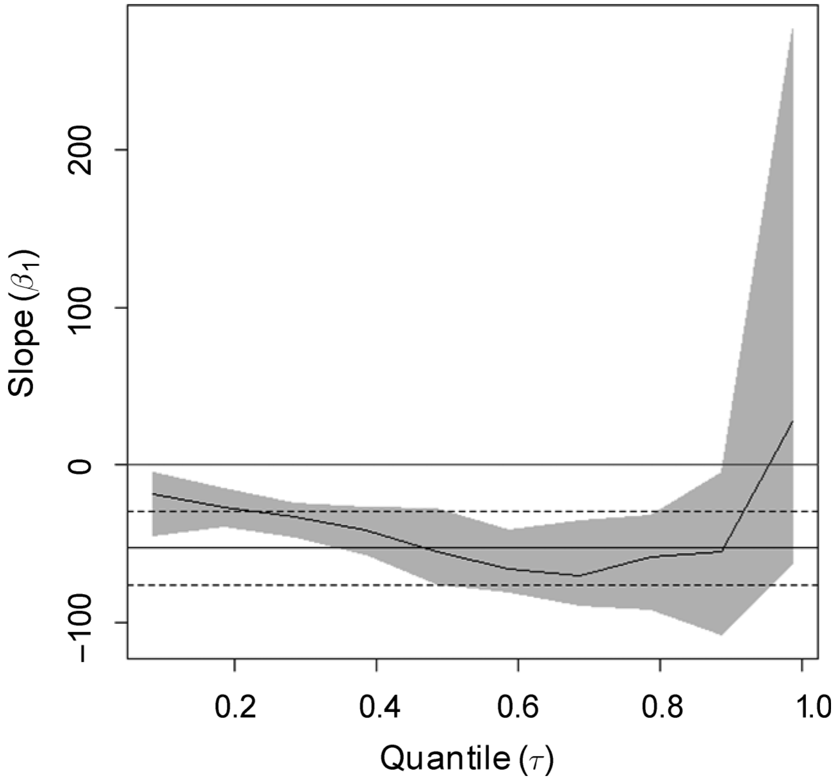


Figure 4. Variation (solid black line) of the quadratic slope coefficient (β_1) of the considered second-order polynomial models (equation (4)). The grey area represents 99% confidence intervals; thus, when they include zero (solid grey line) the coefficient is not significant at $p < 0.01$. The horizontal solid black line represents the slope (β_1) value with 99% confidence intervals (dashed black lines) obtained by OLS regression. A similar representation of quantile and OLS parameters is provided by Koenker and Hallock (2001).

increasing once again until its non-significance at $\tau \geq 0.90$. While, for most of the quantiles, the humped-back curve was significant (β_1 being negative and different from zero at $p < 0.01$), as τ equalled *ca* 0.90, the parabola opened up, reaching an almost linear shape (figures 3 and 4). In fact, fitting linear models (not shown in figure 3) at higher quantiles led to significant negative slopes (e.g. the slope was –48 at $\tau = 0.9$, –68 at $\tau = 0.95$ and –80 at $\tau = 0.99$, $p < 0.01$).

Of course, both OLS and quantile regression models carry the same conceptual information, i.e. a humped-back relationship between spectral (and thus spatial) heterogeneity and biomass. From a statistical point of view, Woodcock and Strahler (1987) found that the local variability depends on the match between the dimension of the scene objects and that of the pixels. In particular, local spectral variability should reach its maximum when the pixel spatial resolution is approximately the same as the scene objects being considered. Local spectral variability should decrease as the pixel dimension becomes larger or smaller (Woodcock and Strahler 1987, Song and Woodcock 2002), as the correlation among the nearest pixels decreases and the local variability rises (Ricotta *et al.* 1999). In this paper, the scene objects are considerably larger than the pixel size at the two extremes of the humped-back curve model (low and high NDVI), mostly related to bare soil and dense-canopy cover, respectively.

From an ecological point of view, the chief source of variation within the dataset is caused by differences among different land-use classes (Woodcock and Strahler 1987, Ricotta *et al.* 1999). In fact, considering points falling in ecosystems with a high degree of canopy cover (right-hand side of figure 3), it is expected that a denser canopy cover will lead to lower variability of spectral reflectance. When dense canopy occurs in forested environments, the spatial heterogeneity perceived by the remote sensor should be low (Nagendra 2001). On the other hand, when the canopy opens up (central part of figure 3), understorey vegetation and bare soil should increase spectral contrast between isolated tree crowns, shrubs and herbaceous/soil background. Finally, completely bare soil or water (left-hand side of figure 3) is expected to show a very low amount of variability for medium-spatial-resolution images. The spectral consequences of such reduced variability on high and low NDVI values results in the humped-back relationship shown here.

The pattern reported in this paper is not novel and has been reported elsewhere in the Mediterranean biogeographical region. A number of studies have found high levels of heterogeneity, even at very low NDVI values in a range of ecosystems, including wetland areas (Rocchini *et al.* 2005, Laba *et al.* 2008), urban areas (Small 2005, Nichol and Wong 2007) and bare soil areas such as the Mediterranean badlands (Phillips 1998, Marignani *et al.* 2008). Rocchini *et al.* (2007) demonstrated that wetland areas might show a high local variability due to the presence of local clearings within marshes, despite the very low NDVI values enhanced by the presence of water, which severely affects the spectral response in the NIR. More generally, Small (2005) highlighted the extremely high variability of urban reflectance at a variety of spatial scales, which is suspected to show lower values of NDVI when compared with shrubland and woodland areas. If an OLS model is used to fit the whole universe U of points, such patterns are not discernible.

The noise present in the examined scatter plots is largely due to the statistical distribution of the data and the effects of the spatial and spectral resolution of the satellite sensor. Statistical distributions of ecological data often show unequal variation due to complex interactions between the factors being accounted for (Cade and Noon 2003). In other words, most of the actual response variability is lost when only one factor is considered as an explanatory variable (Gotelli and Ellison 2004). Univariate models are expected to reach only a part of the total variance of the system (Orlóci 1978, Orlóci and Kenkel 1985, Legendre and Legendre 1998). Therefore, the NDVI cannot, by itself, explain the statistical variability in spectral heterogeneity, as the NDVI is not necessarily a limiting factor of spectral heterogeneity.

With regard to spatial resolution, ecological processes provide different information at different scales (Dungan *et al.* 2002). Therefore, inference is strictly related to the spatial scale of analysis and some patterns or processes can be recognized only at specific resolutions (Palmer and White 1994, Stohlgren *et al.* 1997a,b, Hortal and Lobo 2005, Legendre *et al.* 2005). Since pixels represent spatially explicit entities containing several objects (Fisher 1997, Cracknell 1998), mixed pixels increase as resolution decreases (Schiewe 2005, Small 2005). A medium-resolution sensor, such as Landsat ETM+, may fail in detecting NDVI and spectral heterogeneity. Rocchini (2007b) demonstrated that spectral resolution can compensate where spatial resolution is limiting. In fact, high spectral resolution should bring higher information content by minimizing low spatial-resolution effects (see Nagendra and Rocchini (2008) for

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

a review on the matter). Here, the spectral information content of the whole multispectral dataset was reduced to PC1. This may partially account for the noise found in the presented scatterplot. However, it should be stressed that the main aim of this paper was not to test for absolute models of spectral heterogeneity versus NDVI, but to examine relative differences among regression models applied to the same input variables.

4. Conclusion

As stressed by Woodcock *et al.* (1988a,b), understanding the causes of spatial variation in images can provide a basis for developing new streamlined image analysis procedures. Here, we demonstrate that quantile regression may represent a powerful technique for the analysis of remotely sensed data, as it may reveal patterns that OLS regression may hide. On the one hand, regression-based analysis represents an important tool in the first stages of any ecological study, where little is known *a priori* about the ecosystem being studied and researchers are limited to unstructured surveys based on correlative patterns between variables (Anderson and Gribble 1998). On the other hand, OLS regression has both statistical (e.g. Schröder *et al.* 2005, Cade and Richards 2006, Cade *et al.* 2006) and ecological limitations (e.g. Cade *et al.* 2005, Griffith and Peres-Neto 2006, Buhk *et al.* 2007, Cho *et al.* 2007). Such drawbacks are reviewed by Koenker and Hallock (2001, p. 154), where the authors state:

> 'what the regression curve does is give a grand summary for the averages of the distributions corresponding to the set of *x*'s. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a more complete picture of the set. Ordinarily this is not done and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a corresponding incomplete picture for a set of distributions.'

In this paper, we argue that caution needs to be taken with this 'mean concept' (Schröder *et al.* 2005) when performing regression analysis on reflectance data, while a more detailed examination of all possible subsets within a scatter plot may increase the understanding of patterns that are otherwise lost by commonly used regression techniques.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

Acknowledgements

We greatly thank the Editor Arthur P. Cracknell for efforts made in improving a previous draft of the paper. The final message of this paper has been developed in a more efficient way thanks to two anonymous referees who provided useful insights about the model presented. Brian S. Cade provided R functions and enormous insights about quantile regression. The description of the study area was rearranged from a complete dissertation provided by Francesco Geri. Daniel J. McGlinn, Steven A. Loiselle and George L.W. Perry provided useful insights.

References

Albani, M., Klinkenberg, B., Andison, D.W. and Kimmins, J.P., 2004, The choice of window size in approximating topographic surfaces from digital elevation models. *International Journal of Geographical Information Science*, **18**, pp. 577–593.

ANDERSON, M.J. and GRIBBLE, N.A., 1998, Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Australian Journal of Ecology*, **23**, pp. 158–167.

BOSSARD, M., FERANEC, J. and OTAHEL, J., 2000, *CORINE Land Cover Technical Guide – Addendum 2000*. Technical Report No.40, European Environment Agency.

BUHK, C., RETZER, V., BEIERKUHNLEIN, C. and JENTSCH, A., 2007, Predicting plant species richness and vegetation patterns in cultural landscapes using disturbance parameters. *Agriculture, Ecosystems & Environment*, **122**, pp. 446–452.

CADE, B.S. and GUO, Q., 2000, Estimating effects of constraints on plant performance with regression quantiles. *Oikos*, **91**, pp. 245–254.

CADE, B.S. and NOON, B.R., 2003, A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, **1**, pp. 412–420.

CADE, B.S. and RICHARDS, J.D., 2006, A permutation test for quantile regression. *Journal of Agricultural, Biological, and Environmental Statistics*, **11**, pp. 106–126.

CADE, B.S., TERRELL, J.W. and SCHROEDER, R.L., 1999, Estimating effects of limiting factors with regression quantiles. *Ecology*, **80**, pp. 311–323.

CADE, B.S., NOON, B.R. and FLATHER, C.H., 2005, Quantile regression reveals hidden bias and uncertainty in habitat models. *Ecology*, **86**, pp. 786–800.

CADE, B.S., RICHARDS, J.D. and MIELKE JR, P.W., 2006, Rank score and permutation testing alternatives for regression quantile estimates. *Journal of Statistical Computation and Simulation*, **76**, pp. 331–355.

CHO, M.A., SKIDMORE, A., CORSI, F., VAN WIEREN, S.E. and SOBHAN, I., 2007, Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation*, **9**, pp. 414–424.

CRACKNELL, A.P., 1998, Synergy in remote sensing: what's in a pixel? *International Journal of Remote Sensing*, **19**, pp. 2025–2047.

DUDA, R., HART, P. and STORK, D., 2001, *Pattern Classification* (New York, NY: Wiley Interscience).

DUNGAN, J., PERRY, J.N., DALE, M.R.T., LEGENDRE, P., CITRON-POUSTY, S., FORTIN, M.-J., JAKOMULSKA, A., MIRITI, M. and ROSENBERG, M.S., 2002, A balanced view of scale in spatial statistical analysis. *Ecography*, **25**, pp. 626–640.

FAIRBANKS, D.H.K. and MCGWIRE, K.C., 2004, Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi-temporal NDVI. *Global Ecology and Biogeography*, **13**, pp. 221–235.

FISHER, P., 1997, The pixel: a snare and a delusion. *International Journal of Remote Sensing*, **18**, pp. 679–685.

FOODY, G.M. and CUTLER, M.E.J., 2003, Tree biodiversity in protected and logged Bornean tropical rain forests and its measurement by satellite remote sensing. *Journal of Biogeography*, **30**, pp. 1053–1066.

GILLESPIE, T.W., 2006, Predicting woody-plant species richness in tropical dry forests: a case study from South Florida, USA. *Ecological Applications*, **15**, pp. 27–37.

GOTELLI, N. and ELLISON, A., 2004, *A Primer Of Ecological Statistics* (Sunderland, USA: Sinauer Associates).

GOULD, W., 2000, Remote sensing of vegetation, plant species richness, and regional biodiversity hot spots. *Ecological Applications*, **10**, pp. 1861–1870.

GRAHAM, M.H., 2003, Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, pp. 2809–2815.

GRIFFITH, D.A. and PERES-NETO, P.R., 2006, Spatial modeling in ecology: the flexibility of eigenfunction spatial analysis. *Ecology*, **87**, pp. 2603–2613.

HAWKINS, B.A., FIELD, R., CORNELL, H.W., CURRIE, D.J., GUÉGAN, J.-F., KAUFMAN, D.M., KERR, J.T., MITTELBACH, G.G., OBERDORFF, T., O'BRIEN, E.M., PORTER, E.E. and

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010

TURNER, J.R.G, 2003, Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, pp. 3105–3117.

HORTAL, J. and LOBO, J., 2005, An ED-based protocol for optimal sampling of biodiversity. *Biodiversity and Conservation*, **14**, pp. 2913–2947.

JENSEN, J.R., 1996, *Introductory Digital Image Processing: a Remote Sensing Perspective* (New York, NY: Prentice Hall).

KOENKER, R., 2007, *quantreg: Quantile Regression. R Package Version 4.10*. Available online at: http://www.r-project.org.

KOENKER, R. and BASSETT JR, G., 1978, Regression quantiles. *Econometrica*, **46**, pp. 33–50.

KOENKER, R. and HALLOCK, K., 2001, Quantile regression. *Journal of Economic Perspectives*, **15**, pp. 143–156.

LABA, M., DOWNS, R., SMITH, S., WELSH, S., NEIDER, C., WHITE, S., RICHMOND, M., PHILPOT, W. and BAVEYE, P., 2008, Mapping invasive wetland plants in the Hudson River National Estuarine Research Reserve using quickbird satellite imagery. *Remote Sensing of Environment*, **112**, pp. 286–300.

LEGENDRE, P. and LEGENDRE, L., 1998, *Numerical Ecology* (Amsterdam, The Netherlands: Elsevier).

LEGENDRE, P., BORCARD, D. and PERES-NETO, P.R., 2005, Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs*, **75**, pp. 435–450.

MARIGNANI, M., ROCCHINI, D., TORRI, D., CHIARUCCI, A. and MACCHERINI, S., 2008, Planning restoration in a cultural landscape in Italy using an object-based approach and historical analysis. *Landscape and Urban Planning*, **84**, pp. 28–37.

MEYNARD, C.N. and QUINN, J.F., 2007, Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, **34**, pp. 1455–1469.

NAGENDRA, H., 2001, Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, **22**, pp. 2377–2400.

NAGENDRA, H., and ROCCHINI, D., 2008, Satellite imagery applied to biodiversity study in the tropics: the devil is in the detail. *Biodiversity and Conservation*, **17**, pp. 3431–3442.

NICHOL, J. and WONG, M.S., 2007, Remote sensing of urban vegetation life form by spectral mixture analysis of high-resolution IKONOS satellite images. *International Journal of Remote Sensing*, **28**, pp. 985–1000.

ORLÓCI, L., 1978, *Multivariate Analysis in Vegetation Research* (The Hague, The Netherlands: Dr W. Junk).

ORLÓCI, L. and KENKEL, N.C., 1985, *Introduction to Data Analysis, With Examples From Population and Community Ecology* (Fairland, USA: International Cooperative Publishing House).

PALMER, M.W. and WHITE, P.S., 1994, Scale dependence and species-area relationship. *The American Naturalist*, **144**, pp. 717–740.

PALMER, M.W., EARLS, P., HOAGLAND, B.W., WHITE, P.S. and WOHLGEMUTH, T., 2002, Quantitative tools for perfecting species lists. *Environmetrics*, **13**, pp. 121–137.

PHILLIPS, C.P., 1998, The badlands of Italy: a vanishing landscape? *Applied Geography*, **18**, pp. 243–257.

R DEVELOPMENT CORE TEAM, 2007, *R: a Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing).

RAPETTI, F. and VITTORINI, S., 1995, *Carta Climatica Della Toscana* (Pisa, Italy: Pacini Editore).

RICOTTA, C., AVENA, G.C. and VOLPE, F., 1999, The influence of principal component analysis on the spatial structure of a multispectral dataset. *International Journal of Remote Sensing*, **20**, pp. 3367–3376.

ROCCHINI, D., 2007a, Distance decay in spectral space in analysing ecosystem β-diversity. *International Journal of Remote Sensing*, **28**, pp. 2635–2644.

ROCCHINI, D., 2007b, Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. *Remote Sensing of Environment*, **111**, pp. 423–434.

ROCCHINI, D., CHIARUCCI, A. and LOISELLE, S.A., 2004, Testing the spectral variation hypothesis by using satellite multispectral images. *Acta Oecologica*, **26**, pp. 117–120.

ROCCHINI, D., ANDREINI BUTINI, S. and CHIARUCCI, A., 2005, Maximizing plant species inventory efficiency by means of remotely sensed spectral distances. *Global Ecology and Biogeography*, **14**, pp. 431–437.

ROCCHINI, D., RICOTTA, C. and CHIARUCCI, A., 2007, Using remote sensing to assess plant species richness: the role of multispectral systems. *Applied Vegetation Science*, **10**, pp. 325–332.

SCHIEWE, J., 2005, Status and future perspectives of the application potential of digital airborne sensor systems. *International Journal of Applied Earth Observation and Geoinformation*, **6**, pp. 215–228.

SCHRÖDER, H.K., ANDERSEN, H.E. and KIEHL, K., 2005, Rejecting the mean: estimating the response of fen plant species to environmental factors by non-linear quantile regression. *Journal of Vegetation Science*, **16**, pp. 373–382.

SMALL, C., 2005, A global analysis of urban reflectance. *International Journal of Remote Sensing*, **26**, pp. 661–681.

SOKAL, R.R. and ROHLF, F.J., 1995, *Biometry: the Principles and Practice of Statistics in Biological Research* (New York, NY: Freeman and Co.).

SONG, C. and WOODCOCK, C.E., 2002, The spatial manifestation of forest succession in optical imagery: the potential of multiresolution imagery. *Remote Sensing of Environment*, **82**, pp. 271–284.

STOHLGREN, T.J., CHONG, G.W., KALKHAN, M.A. and SCHELL, L.D., 1997a, Multiscale sampling of plant diversity: effects of minimum mapping unit size. *Ecological Applications*, **7**, pp. 1064–1074.

STOHLGREN, T.J., COUGHENOUR, M.B., CHONG, G.W., BINKLEY, D., KALKHAN, M.A., SCHELL, L.D., BUCKLEY, D.J. and BERRY, J.K., 1997b, Landscape analysis of plant diversity. *Landscape Ecology*, **12**, pp. 155–170.

TUCKER, C.J., GRANT, D.M. and DYKSTRA, J.D., 2004, NASA's global orthorectified Landsat data set. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 313–322.

WHITTAKER, R.J. and HEEGAARD, E., 2003, What is the observed relationship between species richness and productivity? Comment. *Ecology*, **84**, pp. 3384–3390.

WOODCOCK, C.E. and STRAHLER, A.H., 1987, The factor of scale in remote sensing. *Remote Sensing of Environment*, **21**, pp. 311–332.

WOODCOCK, C.E., STRAHLER, A.H. and JUPP, D.L.B., 1988a, The use of variograms in remote sensing: I. Scene models and simulated images. *Remote Sensing of Environment*, **25**, pp. 349–379.

WOODCOCK, C.E., STRAHLER, A.H. and JUPP, D.L.B., 1988b, The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, **25**, pp. 349–379.

Downloaded By: [Rocchini, Duccio] At: 13:23 26 May 2010