

Very early prediction of wine yield based on satellite data from VEGETATION

MÁRIO CUNHA*†‡§, ANDRÉ R. S. MARÇAL‡¶ and LISA SILVA¶

†Departamento de Geociências, Ambiente e Ordenamento do Território, Faculdade de Ciências, Universidade do Porto, Rua Padre Armando Quintas, 4485-661 Vairão, Portugal

Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências do Porto §Academic Visitor, Centre for Agri-Environmental Research (CAER),

University of Reading, UK ¶Departamento de Matemática Aplicada, Faculdade de Ciências,

Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

(Received 7 May 2008; in final form 21 November 2008)

A forecast model for estimating the annual variation in regional wine yield based on remote sensing was developed for the main wine regions of Portugal. Normalized Difference Vegetation Index (NDVI) time-series obtained by the VEGETATION sensor, on board the most recent Satellite Pour l'Observation de la Terre (SPOT) satellite, over the period 1998-2008 were used for four test sites located in the main wine regions of Portugal: Douro (two sites), Vinhos Verdes and Alentejo. The CORINE (Coordination of Information on the Environment) Land Cover maps from 2000 were initially used to select the suitable regional test sites. The NDVI values of the second decade of April of the previous season to harvest were significantly correlated to the wine yield for all studied regions. The relation between the NDVI and grapevine induction and differentiation of the inflorescence primordial or bud fruitfulness during the previous season is discussed. This NDVI measurement can be made about 17 months before harvest and allows us to obtain very early forecasts of potential regional wine yield. Appropriate statistical tests indicated that the wine yield forecast model explains 77-88% of the inter-annual variability in wine yield. The comparison of official wine yield and the adjusted prediction models, based on 36 annual data records for all regions, shows an average spread deviation between 2.9% and 7.1% for the different regions. The dataset provided by the VEGETATION sensor proved to be a valuable tool for vineyard monitoring, mainly for inter-annual comparisons on a regional scale due to their high data acquisition rates and wide availability. The accuracy, very early indication and low-cost of the developed forecast model justify its use by the winery and viticulture industry.

1. Introduction

Grapevines are the number one planted fruit crop, with more than seven million ha of grapevines grown worldwide, ranging from 50° N, through the tropics, to 43° S, in all continents except Antarctica. Although grapevines grow from temperate to tropical regions, most vineyards are planted in areas with temperate climates, with the most

^{*}Corresponding author. Email: mcunha@mail.icav.up.pt

concentrated cultures occurring in Europe. Portugal is number five in the European wine producers' ranking and number eight worldwide (OIV 2008).

Portuguese wine production is characterized by a strong year to year variability, with adverse consequences for all operators related to the wine business. Despite this instability in production and the great importance of wine production in Portugal, as for many other countries, there is no timely, accurate, low-cost operational vintage forecast system (Cunha *et al.* 2003).

Crop forecasts are performed to improve the efficiency of vineyard and winery operations. It is essential for efficient harvest organization, regional pricing negotiations, crusher intake scheduling, tank space allocation for vintage, investment in new winery capital equipment and the development of marketing strategies for both domestic and export markets. The government can use forecast information to implement regulator mechanisms provided under the Common Organization of the Wine Market for moderating the effects of year to year crop variability (e.g. price policy, assigning economic aid, production quotas, stock management and other instruments) (Panigai and Moncomble 1988, Clingeleffer *et al.* 2001, Cunha *et al.* 2003).

Typically, crop forecasting systems are based on calculation of future wine yield (WY; 1 ha⁻¹) from estimates of its components, using a formula such as:

$$WY = VY(WYP) \tag{1}$$

where V (vines ha⁻¹) is the number of vines per hectare, Y (kg vine⁻¹) is the vine yield and WYP (l kg⁻¹) is the wine yield after grape fruit processing. The WYP depend both on the field environment and wine-making techniques at the winery (Cunha *et al.* 2003). The variable V can be expanded as:

$$V = NSCT \tag{2}$$

where N is the number of grapevine nodes retained at winter pruning, S (shoots N^{-1}) is the budburst, C (cluster shoot⁻¹) is the fruitfulness and T (kg) is cluster weight.

For forecasting purposes, observations of the yield component are generally collected through a field network. These field measurements are time consuming and expensive, especially when large regions are concerned. Thus, when seeking to develop a wine yield forecast system, it would be efficient to devote more effort to the measurement of yield components in proportion to their relative contribution to inter-annual variation (Dunn and Martin 2000). Similarly, there would be little point in devoting a great deal of time and effort to controlling a minor source of the inter-annual variation of yield, while neglecting a major one.

Previous work related to the relative importance of each yield component on future yield has consistently demonstrated that the component 'cluster vine⁻¹' explains 60 - 80% of the inter-annual variation in vine yields (Dunn and Martin 1998, May 2004). Cour and Van-Camp (1980) demonstrated a significant relationship ($R^2 = 0.8$) between pollen counts and both flower/cluster and cluster number. A wine forecast model based on airborne pollen concentration showed good results in crop prediction in the main wine regions of northern Portugal (Cunha *et al.* 2003). The main disadvantages of airborne pollen forecasts are the placement representivity of the airborne pollen sampling devices at regional level and the expensive and complex laboratory process involved.

Cluster per vine counts prior to bloom at a trial plot, then extrapolated to a regional level, are currently the basis of most classical methods for wine yield forecast (Huglin

and Schneider 1985, Panigai and Moncomble 1988, Clingeleffer *et al.* 2001). The accuracy of this method is greatly dependent on the quality of cluster sampling, which is generally limited by labour requirements and costs. According to these studies, while an accurate estimate of wine yield may be possible with these classic forecast systems (cluster count or airborne pollen), the precocity of the information as well as the labour and costs required to produce the estimates makes them unsatisfactory.

The payoff to the farmer of adopting a particular operational forecast system depends upon the timing and the accuracy of the information. Hence, part of our research has focused toward understanding the factors that regulate early grapevine bud fruitfulness (see 1-2) and how to estimate them.

As with most perennial and deciduous plants, grapevine induction and differentiation of the inflorescence primordial for next year's crop begins soon after bud-break of the current season. The first visible sign of the inflorescence evocation occurs during early spring of the previous season, by the apical meristem of an extra-lateral meristematic structure called the 'anlagen' (initials), a term introduced by Barnard (1932). In brief, the 'anlagen' are formed in the bud that is situated in the axil of every leaf of the young green shoot (e.g. May 2004). In the Mediterranean climate of South Australia the differentiation of 'anlagen' into the inflorescence primordial, for different grape varieties, was observed 4-6 weeks after budburst (Watt et al. 2008). In terms of grapevine phenology, bud initiation occurring during the Baggiolini (1952) stage G (more than six leaves unfolded and inflorescences of the current season separated and spaced along shoot) around two weeks prior to full bloom - Baggiolini (1952) stage I (Carneiro 1983, Clingeleffer et al. 2001, May 2004). In Portugal, as for many other Mediterranean countries, this physiological event occurs in April, about 17 months before the resultant fruit is harvested (Carneiro 1983). Since wine yield fluctuations from year to year are largely determined by the bud fruitfulness of a vine, a very earlyseason prediction of wine yield may be possible.

Forced single node cuttings (May 1961, May and Antcliff 1963, Buttrose 1974) or dissecting dormant latent buds under the microscope (May 1961, Sánchez and Dokoozlian 2005) to estimate annual variations in bud fruitfulness has been used to predict yield potential (May 1961, 1972, McMichael and Robinson 1998, Clingeleffer *et al.* 2001). According to Clingeleffer *et al.* (2001), while an accurate estimate may be possible with these techniques, the time resources and effort required to produce this estimate may not be operational and commercially practical for industry.

As one of the premises for crop yield, grapevine bud fruitfulness has been the focus of many studies, the majority dating back more than 30 years and reviewed by several authors (Buttrose 1974, Srinivasan and Mullins 1981, May 2000, Mullins *et al.* 2002, Sánchez and Dokoozlian 2005). Most of these studies have focused on agronomic and environmental regulation of fruitfulness and have consistently determined that light and temperature are the most important climatic factors for inflorescence induction and differentiation. The microclimate around and within the vine's canopy could be related to fruitfulness. A shoot's light exposure has a significant positive effect on bud fruitfulness and excessive shoot growth is associated with dense canopies, low bud fruitfulness and inferior fruit quality (Smart *et al.* 1982, Dry 2000, Sánchez and Dokoozilian 2005).

Vine canopy microclimate monitoring at field level is generally made sporadically and punctually using conventional technology, such as meteorological and/or ecophysiological devices. With the advent of remote sensing technology, data acquired by Earth Observation Satellites (EOS) provide a synoptic and repetitive coverage. A much richer and more timely dataset about microclimate canopy can thus be compiled based on EOS data, with which it may be possible to make more accurate and timely predictions of data related to future crop yield (Hall *et al.* 2002, Taylor 2004).

Potentially, one of the most powerful tools in precision viticulture is the use of remote sensing through its ability to provide a rapid synoptic view of grapevine shape, size and vigour over entire vineyards (Hall *et al.* 2003, 2008, Johnson *et al.* 2003). Its potential for improving viticultural practices is evident from the relationships that are known to exist between these canopy descriptors and grape quality and yield.

The most common use of remote sensing for viticulture is through high resolution airborne data, used as part of an integrated management tool for vineyards (Hall *et al.* 2003, Johnson *et al.* 2003). However, the use of high resolution satellite images for mapping and monitoring vineyards is somehow restricted due to limitations in spatial, spectral and temporal resolutions of the datasets available. Although the use of hyper-spectral sensors allows for vegetation indices based on several narrow bands to be used (Zarco-Tejada *et al.* 2005, Renzullo *et al.* 2006), most applications are based on conventional vegetation indices produced from only two spectral bands, such as the Normalized Difference Vegetation Index (NDVI) (Hall *et al.* 2002, Taylor 2004).

In the past few years a number of applications of remote sensing data for viticulture have been reported in the literature: leaf area index, vigour, absorbed photosynthetically active radiation, photosynthetic rate, soil properties, phenology, pests and diseases and other parameters related to vine yield and quality (Hall *et al.* 2002, 2003, 2008, Johnson *et al.* 2003, Zhang *et al.* 2003, Taylor 2004, Vall-Hossera *et al.* 2005, Renzullo *et al.* 2006, Schiavon *et al.* 2007). Accurate mapping of vineyard for wine-growing regions has frequently been used by wine-grower cooperatives to improve the monitoring of quality compliance in areas registered in Controlled Origin Denomination, as well as for management of pollution, erosion, flood risks and other social land managements purposes, particularly areas where vineyards are dominant (Delenne *et al.* 2008). Vine-plot mapping from high spatial resolutions and low spectral resolution images can be done automatically by using different classification techniques (e.g. Lanjeri *et al.* 2004, Chanussot *et al.* 2005, Trias-Sanz 2006, Schiavon *et al.* 2007, Delenne *et al.* 2008).

The main limitations for the use of EOS images in viticulture include the low spatial resolution of most sensors, insufficient revisiting rates and the difficulties or high costs in accessing the data. A number of EOS sensors currently offer low spatial resolution images with a high revisiting rate (or high temporal resolution), such as the VEGETATION sensor, on board the most recent Satellite Pour l'Observation de la Terre (SPOT) satellite (SPOT 2008). Although this sensor cannot provide detailed spatial information for vineyard mapping, it can be a valuable tool for monitoring, mainly for year to year comparisons on a regional scale, due to its high data acquisition rate and the wide availability of the datasets (Marçal *et al.* 2007).

In recent years, a variety of remote sensing products has been used to crop forecast the biomass with uniform cover crops, such as sugarcane (Rudorff and Batista 1990), wheat (Hochheim and Barber 1998), cotton (Mkhabela and Mkhabela 2000), rice (Xiao *et al.* 2005), oilseed rape (Piekarczyk *et al.* 2006) and maize (Rojas 2007, Salazar *et al.* 2008). For grapevines, as for most other fruit trees, no operational remote sensing-based wine yield forecast system was found in the literature.

The aim of this research was to demonstrate the applicability of the forecast model in estimating the inter-annual variability of the regional wine yield on the basis of VEGETATION satellite images in four regional test sites, which are very diverse in terms of weather, soils, vine-growing systems, crop-growing techniques and grape varieties. The model developed allows us to predict regional wine yield 17 months before harvest under a range of environmental and agronomic conditions. The accuracy and operation of this forecast model is presented and compared with other methods to predict regional wine yield reported in the literature. References are made to experimental data and methodological approaches that provide support to the hypotheses on the interaction of NDVI and grapevine differentiation inflorescence primordial or bud fruitfulness in the season prior to the one in which the harvest takes place.

2. Materials and methods

2.1 Study area

Our work was carried out during the last ten cropping seasons (1998–2007) in the three main wine regions of Portugal (figure 1): Vinhos Verdes (VVR), Alentejo (ALT) and Douro, which has two sub-regions – Western (DWR) and Eastern (DER). These regions represent about 42% of the vineyard surface and around 43% of the total wine production in Portugal (239 103 ha and 71 108 l). These regions are very diverse in terms of weather, soils, grape variety, vine-growing systems, crop-growing techniques, impact of diseases on crop size and wine yield.

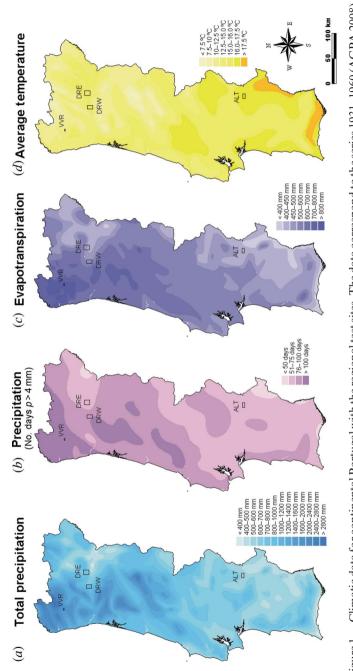
The climate of all four regions is Mediterranean, with evident continental influence and marked annual thermal contrast and water stress in summer, mostly in DWR, DER and ALT. There are, however, some considerable climatic differences between the regions, which can be seen in figure 1, where the climatic records of precipitation (total and number of days above 4 mm), evapotranspiration and average temperature are presented for Continental Portugal (AGPA 2008).

Vines grown in VVR have unique characteristics, namely, the form of guiding systems with wide vegetative expansion and growth high above the ground. In Douro, the Port wine region, vineyards are in stony soils and the large majority planted in hillsides with steep sloops. In ALT and Douro (mostly the eastern part), vineyards are in some of the most arid regions in Europe, with strong and consistent post-flowering vine water and thermal stress. The two test sites in Douro, although only about 50 km apart, have rather different climatic conditions, as can be observed in figure 1 and table 1. Table 1 shows the means and dispersion statistics of meteorological conditions and grape phenology for all (sub)regions studied. The average date of budburst was very close to Julian day 80 for all regions and the budburst-flowering period has a length between 60 to 67 days.

In the studied regions the long-term average rainfall ranges between 579 mm (DER) to 1497 mm (VVR) and the vines are grown without irrigation (table 1). In the VVR region, with a wetter post-flowering period, disease incidence-generated productions losses are frequent. In the other regions, being hotter and drier after flowering, the problems associated with diseases are less frequent (Cunha *et al.* 2003).

For all the test sites, the vineyard floor is managed with natural or sown grass. Due to the natural limitations on water of the Mediterranean climate, the growth of the grass cover is temporary (between November and the end of spring), due to the choice of species sown or, in the case of natural autumn and winter grasses, through mechanical or chemical control.

The WY ($1 ha^{-1}$) for the 1998 to 2007 period was calculated from the information of the annual regional wine production (1) and the vineyard area (ha) provided by the Instituto da Vinha e do Vinho (IVV 2008). Except for two sub-regions of Douro (DWR vs. DER; r = 0.77, p < 0.015), no statistically significant correlations were found between wine yields amongst the studied regions. For the four regional test





		Regional test sites							
		DWR		DER		VVR		ALT	
Variable	Units	avg	cv (%)	avg	cv (%)	avg	cv (%)	avg	cv (%)
Phenology*									
Budburst (C)	Julian	78.94	7.5	80.31	9.2	83.37	9.7	77.73	9.3
Flowering (I)	Julian	139.83	7.8	147.31	1.9	148.40	5.8	138.21	7.2
Climatic data (ar	nual)								
Precipitation	mm	942	27.7	578.9	25.5	1496.7	28.1	660.9	25.4
Temperature	°C	15.4	3.4	12.2	5.8	14.1	3.3	15.7	3.1
Climatic data du	ring budb	urst to ha	rvest						
Precipitation	mm	237.2	34.6	175.2	47.5	434.8	34.0	139.7	46.9
Temperature	°C	19.8	3.6	16.7	5.1	16.9	4.4	19.1	3.9

Table 1. Descriptive statistics of grape phenology and climatic data for the four regional test sites.

Avg, average; cv, coefficient of variation (standard error/average \times 100). Douro West (DWR), Douro East (DER), Vinhos Verdes (VVR) and Alentejo (ALT).

*Phenological stages correspond to the Baggiolini (1952) scale. The budburst (stage C) and flowering (stage I) events are considered to occur when 50% of the plants are exhibiting the physiological response at the field level.

sites the year-to-year variations in vineyard planted are small and wine yield has no trend during the 1998–2007 period.

To evaluate the overall effect of the flowering or grape maturing conditions on crop size, we used the annual data of WYP ($l kg^{-1}$) or equivalent volume, obtained by region or sub-region, from wineries known to maintain consistent wine-making techniques during the study period.

2.2 VEGETATION data

Agricultural remote sensing is frequently based on so-called vegetation indices that are combinations of spectral measurements in different wavelengths as recorded by a radiometric sensor. They aid in the analysis of multi-spectral image information by shrinking multi-dimensional data into a single value. The NDVI has been the most frequently used vegetation index for agrometeorological analysis (Rouse *et al.* 1973):

$$NDVI = (NIR - RED)/(NIR + RED)$$
(3)

NIR and RED are, respectively, the reflectance (%) in the near-infrared and in the red channels.

The VEGETATION sensor has provided daily coverage of the entire Earth since 1998 at a spatial resolution of 1 km (SPOT 2008). The sensor acquires data in four spectral bands in the visible and near-infrared, ranging from 0.43 to 1.75 μ m (VITO 2008). Ten-day synthesis VEGETATION products (S10) are obtained from the compilation of daily data from ten consecutive days, providing atmospherically corrected data (values corresponding to surface reflectance). The resulting surface reflectance value for each pixel corresponds to the date with maximum NDVI reflectance at the top of the atmosphere for that pixel (SPOT 2008). The VEGETATION S10 syntheses are provided on 10 possible regions of interest. One of these pre-defined

regions is 'Europe', covering an area between 25° N and 75° N, and between 11° W and 62° E (VITO 2008).

All VEGETATION S10 syntheses of Europe were transferred from the free VEGETATION distribution site (VITO 2008). The software CROP VGT (VITO 2008) was used to crop a small section from each of these images, with the region of interest corresponding to continental Portugal (only 404×617 pixels). The final image set covers a period of 10 years, 1998–2007, with 36 images each year. As there is a lack of data from the first three months of 1998, a total of 351 NDVI 10-days synthesis images were thus available, between April 1998 and December 2007.

The VEGETATION dataset for Portugal was used to produce temporal NDVI profiles for the test sites selected within each wine region (see below). As each NDVI image is obtained by merging data from 10 consecutive days, the whole site was considered as a unit, instead of using a pixel-by-pixel approach. This is done to prevent misregistration and other sources of errors contaminating the temporal profiles. There is still the problem of cloud cover, as occasionally there is no valid data for a full 10-day period for the whole site. An initial criterion was used to select the valid observation for each image and site. Pixels with very low NDVI values, corresponding to clouds, were rejected. The average median, standard deviation and upper quartile values were computed for each image/site, using only the valid pixels.

2.3 Regional test sites

The CORINE (Coordination of Information on the Environment) Land Cover maps from 2000 (Painho and Caetano 2005) were used to select a suitable test site for each wine region. Initially, all the 1×1 km pixels with 70% or more of the area occupied by vineyards were selected. There are a large number of pixels that verify this condition, but most are isolated pixels. As the VEGETATION images have a pixel of 1 km², and the 10-day syntheses are produced from several images, it is important to select only large contiguous areas, with at least 3×3 km. This last process was done manually, by a visual inspection of the image of valid pixels produced by the initial selection criteria.

For the VVR there are very few 1 km pixels with 70% or more covered by vineyard, and so the criterion was softened for this region.

A detailed representation of the vineyard coverage (%) of the four test sites is presented in figure 2. DWR and DER each have two separate compact groups of pixels all with 80% or more vineyard coverage (18 pixels for DER and 40 pixels for DWR). ALT fails in having a compact group of pixels, but the percentage of vineyard is high for all 20 pixels used. VVR is a very weak selection, but it was still considered as this is a very important wine region. Grapevines in VVR occupy an area of almost 35 103 ha (15% of the national viticultural area) and is the only region studied that produces mainly white wine.

Detailed information of NDVI temporal plots for regions DER, DRW and ALT for the period 1998–2005 are available in Marçal *et al.* (2007).

2.4 Wine forecast models

For each test site, the one to one correlation matrix was calculated between the wine yield of a current year and the full set of 10-day NDVI synthesis computed. The widely accepted concept that grapevines' reproductive structures for the next year's crop begins after bud-break – stage C (Baggiolini 1952) of the current season was the main reason why the 10-day NDVI of the previous season were also considered in this study.

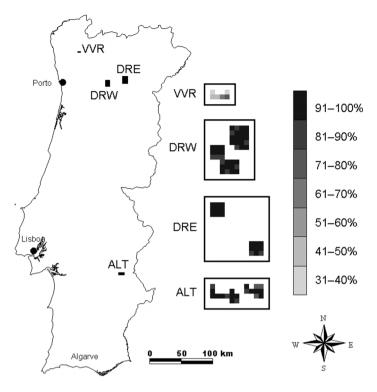


Figure 2. Location of the four regional test sites in continental Portugal, and details of vineyard coverage (%) for each pixel used from the four test sites: Douro West (DWR), Douro East (DER), Vinhos Verdes (VVR) and Alentejo (ALT).

For each region, the 10-day NDVI data were selected regarding the best significant correlation coefficient between each 10-day NDVI and regional wine yield, over the 10 years (1998–2007). The correlation between the NDVI of the previous season and wine yield were computed using only nine years of data, as no wine yield data are available yet for 2008.

For each wine region or sub-region (r) a linear regression was fitted:

$$WY_r = (\beta_{0r} + \beta_r)(NDVI)_r$$
(4)

For each test site a linear regression was fitted and *t*-tests were performed to evaluate the coefficients (β_0 and β) significance. The regression diagnostics of the most influential observations were tested according to the procedures described by Montgomery *et al.* (2006) and, unless otherwise stated, the parameters tested were found not to be significant. Leverage statistics were used to detect outliers among the predictor variables with a cutoff of 3(p + 1)/n, where *p* is the number of predictors and *n* the number of observations used to fit the model.

To assess the predictive accuracy of the developed wine yield remote sensing-based forecast model, the root mean square error (RMSE) and relative prediction error between observed values and values adjusted to the forecast models were calculated. The RMSE test assumes that larger forecasts errors are of greater importance than smaller errors, so it gives a more-than-proportionate penalty.

Since the database available for this study only comprises 10 or 9 (when the previous seasons were used) cropping seasons, the external validation of the forecasting models developed was performed using leave-one-out (LOO) cross-validation, which evaluates the model prediction performance for a year not considered in the estimation step, thus providing independent estimates of the predictive capability of the selected models. This technique consists of the removal of one year from the database used and the estimation of a new regression model with the remaining years. This new model will be used to predict the wine yield of the year withdrawn and to calculate the relative prediction error and RMSE between the real production data and the data predicted by the model.

3. Results

As previously stated, the wine yield in the regions selected for this study have great inter-annual variability, and this can be observed in the data presented in table 2 and figure 3. The crop seasons (1998–2007) used in this work are representative of the absolute maximum and minimum of the historical regional data of wine yield, which should allow for the formulation of forecasts with a wide validation interval.

The average, median and upper quartile NDVI values recorded over the 10-year period (1998–2007) were computed for each test site. The full correlation matrices between each NDVI and the wine yield for the period 1998–2007 are too large to be shown here. For all regional test sites the NDVI during the second decade of April (NDVI_{2AP}) of the previous season showed the best correlation with regional wine yield. This period is most likely related to the induction and initiation of inflorescence in the season before flowering and fruiting takes place. The strong consistency exhibited by all regions in the selection of the period in which the grape production system may be sensitive to NDVI indicates that wine yield could be modelled relatively simply and very early.

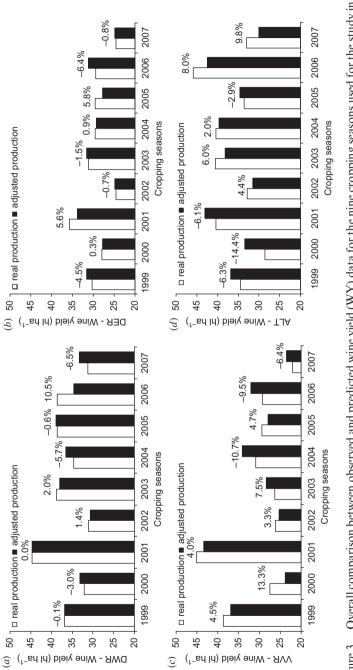
Table 2 show the regional results of $NDVI_{2AP}$, WY and WYP for the last 10 years (1998–2007). There is a strong variability in $NDVI_{2AP}$ over the years 1998 to 2007. These results suggested that variability could be related to overall regional climate. The $NDVI_{2AP}$ and the annual precipitation showed similar patterns among the regions. The higher $NDVI_{2AP}$ values for VVR coincided with high values in

		Regional test sites							
		DWR		DER		VVR		ALT	
Variables*	Units	avg	cv (%)	avg	cv (%)	avg	cv (%)	avg	cv (%)
WY NDVI _{2AP} WYP×10 ³	hl ha ⁻¹ Index l kg ⁻¹	36.3 0.376 759	12.3 12.5 1.5	27.85 0.321 696.5	20.3 17.% 1.1	30.6 0.571 680	23.1 13.2 6.4	34.7 0.354 750	22.7 17.5 2.9

Table 2. Descriptive statistics (1998–2007) of selected variables in the wine yield forecast models to the test sites regions.

Avg, average; cv, coefficient of variation (standard error/average \times 100). Douro West (DWR) Douro East (DER), Vinhos Verdes (VVR) and Alentejo (ALT).

*WY, wine yield; NDVI_{2AP}, NDVI for the second decade of April in the previous season to harvest; WYP, wine yield after grape processing.





precipitation in north Portugal, mostly during the bud-break to harvest period (table 1). In this region, soil water content is high at bud-break due to heavy winter rainfall. Thus, vine-water status does not restrict growth early in the growing season.

The Alentejo and Douro regions are characterized by low NDVI_{2AP}, being the regions that have the lowest precipitation levels. The high inter-annual variability of NDVI_{2AP} in the regions ALT and DRE could be related to extremely low precipitation (140–175 mm) with strong coefficient of variation (about 47%) during the bud-break-harvest period. In these regions, soil-water content is frequently low at bud-break, due to lack of winter rainfall, and could restrict growth in early stages of the growing season.

The inter-annual variability in regional wine yield (coefficient of variation, cv = 12.3-23.1%) and NDVI_{2AP} (cv = 12.5-17.5%) is much higher than that of WYP (cv < 6.5%) (table 2). In RVV the high inter-annual variation reflects the great heterogeneity of precipitation values during the grape maturing period (table 1). The small variation range of the WYP indicates the little importance this variable has in causing inter-annual variability in wine yield for these regions.

Table 3 shows the estimates of the regression coefficients and the significance of the *t*-test. The slope (β) and intercept (β_0) values for all regions were significantly different from zero, with a significant level of 0.000 for all cases, except one (0.004). The statistical analyses revealed negative linear correlation coefficients between NDVI_{2AP} and the regional yield production for all test sites.

The resulting statistical tests related to the model adequacy and validation are presented in table 4. Appropriate statistical tests indicated that the prediction model adjusted for each region describes between 73% and 88% of the regional wine yield variation over years. Leverage statistics calculated to detect influential observations do not exceed the cut-off limits defined by Montgomery *et al.* (2006).

For the four regional test sites, with 36 annual data records, the descriptive statistics show that in 53% of these cases the differences between actual production and production adjusted by the forecast model were below 5%. Only in three observations were the differences higher than 10%, with no difference higher than 14.4% registered (figures 3(a)-(d)).

The LOO was used to avoid any strong influence in the forecast model for a specific year. Table 4 shows the evolution of the relative prediction error and the RMSE after

Regional test sites	Coefficients*	Value	Std error	<i>t</i> -test	Sig.†
DWR	β_0	67.30	5.22	14.31	0.000
(n = 9)	β	-82.23	13.74	-6.44	0.000
DER	β_0	46.49	2.63	17.65	0.000
(n = 9)	β	-53.70	8.07	-6.66	0.000
VVR	β_0	79.08	6.72	11.76	0.000
(n = 9)	β	-85.72	11.80	-7.27	0.000
ÀLT	β_0	61.58	5.91	10.41	0.000
(<i>n</i> = 9)	β	-70.37	16.42	-4.29	0.004

 Table 3. Estimates and statistical analysis of the forecast model regression coefficients for the four regional test sites.

Douro West (DWR), Douro East (DER), Vinhos Verdes (VVR) and Alentejo (ALT) $WY_r = (\beta_{0r} + \beta_r)(NDVI)_{r2Ap}$.

*Significance level of t-test value for each regression coefficient.

	Regional test sites							
Statistics	DWR	DER	VVR	ALT				
Model adequacy								
R^2 (significance)	0.84 (p<0.001)	0.87 (p<0.000)	0.88 (p<0.000)	0.73 (p<0.004)				
Std estimate	1.93	1.35	2.59	3.04				
Leverage (min., max.)*	0.00, 0.52	0.00, 0.26	0.00, 0.46	0.00, 0.46				
Average D (%)	3.3	2.9	7.1	6.7				
RMSE (hl ha ⁻¹)	4.7	3.8	9.2	9.0				
Leave-one-out (LOO) cross-validation [†]								
D (%)	4.7	3.8	7.8	7.6				
RMSE (hl ha ⁻¹)	5.5	5.0	10.0	10.5				

Table 4. Statistics of forecast model adequacy and validation for the four test sites regions.

Douro West (DWR), Douro East (DER), Vinhos Verdes (VVR) and Alentejo (ALT). *Leverage statistics range for each model. The leverage ranges from 0 (no influence on the fit) to (n - 1)/n and the cut-off criteria used to detect outliers was 3(p+1)/n, where *p* is the number of predictors and *n* is the number of observations used to fit the model.

†Relative prediction error (D) and the RMSE after the LOO had been applied.

the LOO had been applied. When LOO was performed the relative prediction error and the RMSE for each region are similar to the predictions obtained with the observations used to fit the model.

At the present date, it is possible to make a prediction for the wine yield in 2008, based on the NDVI_{2AP} 2007 data. The predicted wine yield for 2008 is 37.51 ha^{-1} for DWR, 28.21 ha^{-1} for DER, 25.91 ha^{-1} for VVR and 37.31 ha^{-1} for ALT. This is truly an independent prediction, as the actual wine yield for 2008 will be available only in March 2009.

4. Discussion

The seasonal variation in wine yield exhibits random patterns in the studied test sites, as for many other regions, and this is why it is difficult to forecast wine yields from historical records. According to previous research, the yield from a vineyard fluctuates significantly from year to year, and is determined largely by the bud fruitfulness. If environmental and agronomic conditions during formation of inflorescence primordial (bud differentiation) determine the potential number of clusters that the vine will carry in the next season, then an excellent opportunity exists to perform a very early potential wine yield forecast.

The NDVI_{2AP} was found to be significantly correlated to the wine yield of the following year in all test sites. Since the second decade of April is generally characterized by the induction and differentiation of the inflorescence primordial (Carneiro 1983, Clingeleffer *et al.* 2001, Watt *et al.* 2008), which is under the influence of the canopy microclimate during this period (Smart *et al.* 1982, Sánchez and Dokoozilian 2005), from a physiological viewpoint it is reasonable to suppose that the NDVI_{2AP} is significantly correlated to bud differentiation.

The significant negative correlation between NDVI and wine yield could be interpreted as an indicator of a direct response of NDVI or vigour to changes in microclimate canopy (Hall *et al.* 2003, 2008, Johnson *et al.* 2003). High NDVI_{2AP} are associated with excessive vegetative growth of grapevines with dense canopies, which can cause shading and low bud fruitfulness. Low values of NDVI during this period are associated with reduced foliage density, which increases the illumination within the canopy and results in improved fertility of basal buds, a factor that leads to increased fruitfulness of the vine as a whole (Smart *et al.* 1982, Sánchez and Dokoozilian 2005).

The forecasting model developed showed a suitable prediction capability 17 months before harvest for all regions, which allows for a very early season forecast of potential wine yield to be made at a regional level. Appropriate statistical tests indicated that 77–88% of annual variability in regional wine yield can be explained by the NDVI_{2AP} in the previous season to harvest. The comparison of the official productions and the wine yield predicted by the model showed an average spread deviation between 2.9% and 7.1% (figure 3).

Our results are encouraging when compared with those reported by Panigai and Moncomble (1998) in France, Clingeleffer *et al.* (2001) in Australia and Cunha *et al.* (2003) in Portugal. The forecast model developed provides almost the same accuracy as these other methods, but is much more efficient in terms of precocity, cost and operation. These results also indicate that, whilst NDVI_{2AP} assesses the crop potential at the bud differentiation stage, other factors are probably also important in determining the final production levels. By updating estimates with real-time weather data, as well as information on diseases and grape crop management with field measurements as a season progresses, the predictions could be continuously updated. However, as some of these variables can be known only a short time before harvest, their inclusion in a forecast model would imply a loss of forecast opportunity, decreasing the model's usefulness.

In this work the evaluation of the overall possible impact of grape maturing conditions on limited crop production was made by the analysis of the wine yield after grape fruit processing. When, during a sampling period, the wineries maintain the wine-making techniques, the WYP can be used to evaluate the overall environmental (diseases, water and thermal stress) effects of grape maturing conditions on wine yield (Cunha *et al.* 2003). The WYP exhibits a variation coefficient lower than 6.4% in all regions tested, indicating a strong annual stability when compared with the regional wine yield (table 3). However, there seems to be an increase in the inter-annual variability of WYP in regions with higher summer rainfall.

We found not only higher coefficients of variation value (cv = 28%) in the VVR region (1496 mm annual rainfall) but also a wider coefficient of variation for the mean WYP (cv=6.4%). In this region, appropriate statistical tests of model adequacy indicated high values of both mean relative predictions errors (7.1%, with two values higher than 10%), and RMSE (9.21 ha⁻¹). This can be due to the fact that the test site selected for VVR is far from ideal, as there are not many 1 km pixels mostly covered by vineyard in this region.

In regions with severe plant water stress in the grape maturing period (Douro and ALT), the smallest variation range of the WYP (cv < 2.9%) indicates the greater interannual stability of this variable (table 3). In the Douro and ALT regions, where the water stress in summer is consistent over the years and the grape diseases are less frequent, it is difficult to consider grape maturing conditions as an important factor causing inter-annual fluctuations in wine yield. Low inter-annual variability of WYP under conditions of high frequency of water stress in summer were also reported by Cunha *et al.* (2003) in Portugal and Clingeleffer *et al.* (2001) in Australia. These results also show the very low influence on wine yield of grape maturing conditions relative to those that affect bud differentiation in the previous year. Thus, this model can be used in many other arid zones where large parts of the world's vineyard are planted.

One of the crucial aspects for the practical development of a vineyard monitoring tool based on EOS data is the spatial and temporal resolution of the images, as well as the availability and cost-related issues. High spatial resolution sensors, such as IKONOS, Quickbird and airborne-mounted sensors, can be used to produce vegetation indices, such as NDVI and leaf area index, at an adequate scale for vineyard monitoring (Hall et al. 2003, 2008, Johnson et al. 2003, Vall-Hossera et al. 2005). However, these images are still infrequent and very expensive, both factors limiting the practical implementation of a satellite-based monitoring system (Hall et al. 2008). Intermediate spatial resolution sensors, such as SPOT High Resolution Geometrical (HRG) and Landsat Thematic Mapper (TM), can partly overcome these problems, although the temporal resolution is not very good. As the prediction model developed is based on EOS data from a narrow time window (2nd decade of April), it might be possible to use high resolution EOS images to compute NDVI values at a higher scale. However, the use of data from IKONOS, SPOT or Landsat might be limited to a single image with favourable cloud cover conditions, in the relevant time window. However, a single image can be used to compute valid NDVI values only if the appropriate atmospheric corrections are carried out, which is not a straightforward task.

Low resolution sensors, such as VEGETATION, provide daily images (globally) which can be used to produce a maximum value composite for the period of interest. Although the images have low spatial resolution, the much higher acquisition rate provides a way of obtaining consistent year-to-year quantitative NDVI values.

5. Conclusion

A practical evaluation of the applicability of VEGETATION satellite images to provide meaningful information about regional wine yield forecast was carried out. The current study demonstrated that the NDVI derived from VEGETATION is a valuable tool for predicting the wine yield in regions with great differences in climate, soils, grape variety, wine yield, vine-growing systems, crop-growing techniques and impact of diseases on crop size.

A very early prediction of potential wine yield can be made about 17 months before harvest based on EOS data from the VEGETATION sensor. These results showed the low influence on wine yield of grape maturing conditions when compared to the influence of NDVI during bud differentiation in the season before the one in which harvest takes place. However, the interaction of NDVI and bud fruitfulness during the previous season is not yet fully understood and is therefore a matter for further physiological study to enhance control of crop development and crop estimation, and secure quality and production of major wine grape varieties.

The accuracy, very early capability of the forecast model and the comparison of marginal information costs with respect to the benefits justify its use for the winery and viticulture industry both for economic and technical reasons.

The results indicate that the VEGETATION sensor can provide useful information about wine yield forecast, with adequate spectral and temporal resolution, but there is clearly a limitation in terms of spatial resolution. Further work is required in order to evaluate the applicability of the forecast model developed with vegetation indices extracted for other EOS sensors to provide meaningful information about wine yield forecast in small vineyard parcels. An operational tool can be implemented in the future if the information provided by the EOS images proves to be effective for the viticulture industry.

Acknowledgments

The authors wish to thank the VEGETATION programme for providing the satellite data, Instituto do Ambiente for the CORINE Land Cover maps from 2000 and the following regional technical bodies for providing climate, wine yield and grapevine regional data: Associação para o Desenvolvimento da Viticultura Duriense (ADVID), Associação Técnica dos Viticultores do Alentejo (ATEVA) and Estação Vitivinícola Amândio Galhano (EVAG). The first author wishes to acknowledge the 'Fundação para a Ciência e a Tecnologia' for the SFRH/BSAB/ 813 /2008 grant.

References

- AGPA, 2008, Agência Portuguesa do Ambiente, Ministério do ambiente, do ordenamento do Território e do Desenvolvimento Regional. Available online at: http://www.iambiente. pt/atlas (accessed 1 April 2008).
- BAGGIOLINI, M., 1952, Les stades reperes dans le developpement annuel de la vigne et leur utilization pratique. *Revue Romande. Agriculture. Viticulture Arboriculture*, **8**, pp. 4–6.
- BARNARD, C., 1932, Fruit bud studies. I. The Sultana. An analysis of the distribution and behaviour of the buds of the Sultana vine, together with an account of the differentiation of development of the fruit buds. *Journal of the Council of Scientific and Industrial Research*, 5, pp. 47–52.
- BUTTROSE, M., 1974, Climatic factors and fruitfulness in grapevines. *Horticulture Abstracts*, 44, pp. 319–325.
- CARNEIRO, L., 1983 Estudo da organogenese dos gomos de videira (*Vitis vinifera* L. cv. Dona Maria). *Ciência e Técnica Vitivinícola*, **2**, pp. 5–13.
- CHANUSSOT, J., BAS, P. and BOMBRUN, L., 2005, Airborne remote sensing of vineyards for the detection of dead vine trees. In *Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS)*, 5, pp. 3090–3093.
- CLINGELEFFER, P., DUNN, G., KRSTIC, M. and MARTIN, S., 2001, Crop development, crop estimation and crop control to secure quality and production of major wine grape varieties: a national approach. Final report to Grape and Wine Research & Development Corporation (GWRDC), Project number CSH 96/1, Wayville, Australia.
- COUR, P. and VAN CAMP, M., 1980, Previsión de récolte à partir de l'analyse du contenu pollinique de l'atmosphére. Comptes rendus hebdomadaires des Séances de l'Académie des Sciences, Paris, Série D, Sciences Naturelles, 290, pp. 1043–1046.
- CUNHA, M., ABREU, I., PINTO P. and CASTRO, R., 2003, Airborne pollen samples for early-season estimates of wine production in a mediterranean climate of northern Portugal. *American Journal of Enology and Viticulture*, 54, pp. 189–194.
- DELENNE, C., DURRIEU, S., RABATEL, G., DESHAYES, M., BAILLY, J., LELONG, C. and COUTERON, P., 2008, Textural approaches for vineyard detection and characterization using very high spatial resolution remote sensing data. *International Journal of Remote Sensing*, 29, pp. 1153–1167.
- DRY, P., 2000, Canopy management for fruitfulness. *Australian Journal of Grape and Wine Research*, **6**, pp. 109–115.
- DUNN, G. and MARTIN, S., 1998, Optimising vineyard sampling to assess yield. In Proceedings of the 10th Australian Wine Industry Technical Conference, R. Blair, A.N. Sas, P.F. Hayes and P.B Hoj, (Eds), 2–5 August 1998, Sydney, Australia, pp. 270–271.
- DUNN, G. and MARTIN, S., 2000, Spatial and temporal variation in vineyard yields. In 5th International Symposium on Cool Climate Viticulture & Oenology, Precision Management Workshop, 16–20 January 2000, Melbourne, Australia, pp. 1–4.
- HALL, A., LAMB, D., HOLZAPFEL, B. and LOUIS, J., 2002, Optical remote sensing applications in viticulture a review. *Australian Journal of Grape and Wine Research*, **8**, pp. 36–47.
- HALL, A., LOUIS, J. and LAMB, D., 2003, Characterising and mapping vineyard canopy using highspatial-resolution aerial multispectral images. *Computers Geosciences*, 29, pp. 813–822.

- HALL, A., LOUIS, J. and LAMB, D., 2008, Low-resolution remotely sensed images of winegrape vineyards map spatial variability in planimetric canopy area instead of leaf area index. *Australian Journal of Grape and Wine Research*, 14, pp. 9–17.
- HOCHHEIM, P. and BARBER, G., 1998, Sring wheat yield estimation for western Canada using NOAA NDVI data. *Canadian Journal of Remote Sensing*, **24**, pp. 17–27.
- HUGLIN, P. and SCHNEIDER, C., 1985, Research de méthodes de prevision quantitative de la vendange. *Bulletin de l'Organisation Internationale de la Vigne et du Vin*, **58**, pp. 71–89.
- IVV, 2008, Instituto da Vinha e do Vinho, dados estatísticos sobre a produção de vinho em Portugal. Available online at: http://www.ivv.min-agricultura.pt/estatistica (accessed 1 April 2008).
- JOHNSON, L., ROCZEN, D., YOUKHANA, S., NEMANI, R. and BOSCH, D., 2003, Mapping vineyard leaf area with multispectral satellite imagery. *Computers and Electronics in Agriculture*, 38, pp. 33–44.
- LANJERI, S., SEGARRA, D. and MELIAACUTE, J., 2004, Interannual vineyard crop variability in the Castilla-La Mancha region during the period 1991-1996 with Landsat Thematic Mapper images. *International Journal of Remote Sensing*, 25, pp. 2441–2457.
- MARÇAL, A., GONÇALVES, J., GONÇALVES, H. and CUNHA, M., 2007, Analysis of the temporal signature of vineyards in Portugal using VEGETATION. In *Proceedings of the 26th EARSeL Symposium, New Developments and Challenges in Remote Sensing*, Z. Bochenek (Ed.) (Rotterdam: Millpress), pp. 377–384.
- MAY, P., 1961, The value of an estimate of fruiting potential in the Sultana, Vitis, 3, pp. 15–26.
- MAY, P., 1972, Forecasting the grape crop. Australian Wine, Brewing and Spirit Review, 90, pp. 46–48.
- MAY, P., 2000, From bud to berry, with special reference to inflorescence and bud morphology in *Vitis vinifera* L. *Australian Journal of Grape and Wine Research*, **6**, pp. 82–98.
- MAY, P., 2004, Flowering and Fruitset in Grapevines (Australia: Lythrum Press).
- MAY, P. and ANTCLIFF, A., 1963, The effect of shading on fruitfulness and yield in the Sultana. *Journal of Horticultural Science*, **38**, pp. 85–94.
- MCMICHAEL, P. and ROBINSON, B., 1998, Bud fruitfulness practical responses to bunch number determinations made during dormancy. *The Australian Grapegrower and Winemaker, Annual Technical Issue*, 414, pp. 7–8.
- MKHABELA, M. and MKHABELA, M, 2000, Exploring the possibilities of using NOAA, AVHRR data to forecast cotton yield in Swaziland. UNISWA Journal of Agriculture, 9, pp. 13–21.
- MONTGOMERY, D., PECK, E. and VINING, G., 2006, *Introduction to Linear Regression Analysis*, 4th edn (New York: Wiley-Interscience).
- MULLINS, M., BOUQUET, A. and WILLIAMS, L., 2002, *Biology of the Grapevine* (Cambridge, UK: Cambridge University Press).
- OIV, 2008, Organisation Internationale de la Vigne et du Vin Statistiques. Available online at: http://www.oiv.int/ (accessed 1 April 2008).
- PAINHO, M. and CAETANO, M., 2005, *CORINE Land Cover 2000 em Portugal* (Portugal: Relatório Técnico, Instituto do Ambiente).
- PANIGAI, L. and MONCOMBLE, D., 1988, La prevision de recoltes en Champagne. *Le vigneron Champanois*, **6**, pp. 359–367.
- PIEKARCZYK, J., WÓJTOWICZ, M. and WÓJTOWICZ A., 2006, Estimation of agronomic parameters of winter oilseed rape from field reflectance data. *Acta Agrophysica*, **8**, pp. 205–218.
- RENZULLO, L., BLANCHFIELD, A. and POWELL, K., 2006, A method of wavelength selection and spectral discrimination of hyperspectral reflectance spectrometry. *Geoscience and Remote Sensing, IEEE Transactions*, 44, pp. 1986–1994.
- ROJAS, O. 2007. Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenia. *International Journal of Remote Sensing*, 28, pp. 3775–3793.
- ROUSE, W., HAAS, R., SCHEEL, J. and DEERING, W., 1973, Monitoring vegetation systems in great plains with ERST. In *Proceedings of the Third ERTS Symposium*, NASA SP-3511 (Washington, DC: US Government Printing Office), pp. 309–317.

- RUDORFF, T. and BATISTA, T., 1990, Yield estimation of sugarcane based on agrometeorological-spectral models. *Remote Sensing of Environment*, **33**, pp. 183–192.
- SALAZAR, L., KOGAN, F. and ROYTMAN, L., 2008, Using vegetation health indices and partial least squares method for estimation of corn yield. *International Journal of Remote Sensing*, 29, pp. 175–189.
- SANCHEZ, L. and DOKOOZLIAN, N., 2005, Bud microclimate and fruitfulness in Vitis vinifera L. American Journal of Enology and Viticulture, 56, pp. 319–329.
- SCHIAVON, G., SOLIMINI, D. and BURINI, A., 2007, Sensitivity of multi-temporal high resolution polarimetric C and L-band SAR to grapes in vineyards. In *Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS)*, 23–27 July, Barcelona, Spain, pp. 3651–3654.
- SMART, R., SHAULIS, N. and LEMON, E., 1982, The effect of Concord vineyard microclimate on yield. II. The interrelations between microclimate and yield expression. *American Journal of Enology and Viticulture*, 33, pp. 109–116.
- SPOT, 2008, VEGETATION Programme. Available online at: http://www.spot-vegetation.com (accessed 1 April 2008).
- SRINIVASAN, C. and MULLINS, M., 1981, Physiology of flowering in the grapevine: A review. *American Journal of Enology and Viticulture*, **32**, pp. 47–63.
- TAYLOR, J., 2004, Digital terroirs and precision viticulture: Investigations into the application of information technology in Australian vineyards. PhD thesis, Sydney University, New South Wales, Australia.
- TRIAS-SANZ, R., 2006, Texture orientation and period estimator for discriminating between forests, orchards, vineyards, and tilled fields. *Geoscience and Remote Sensing, IEEE Transactions*, 44, pp. 2755–2760.
- VALL-HOSSERA, M., CAMPS, A., CORBELLA, I., TORRES, F., DUFFO, N., MONERRIS, A., SABIA, R., SELVA, D., ANTOLIN, C., LOPEZ-BAEZA, E., FERRER, F. and SALEH, K., 2005, SMOS REFLEX 2003: L-band emissivity characterization of vineyards. *Geoscience and Remote Sensing, IEEE Transactions*, 43, pp. 973–982.
- VITO, 2008, Free VEGETATION Products. Available online at: http://free.vgt.vito.be/ (accessed 1 April 2008).
- WATT, A., MAY, P., DUNN, G., CRAWFORD, S. and BARLOW, E., 2008, Development of inflorescence primordia in *Vitis vinifera* L. cv. Chardonnay from hot and cool climates. *Australian Journal of Grape and Wine Research*, 14, pp. 46–53.
- XIAO, X., BOLES, S., JIYUAN, L., DAFANG, Z., FROLKING, S., CHANGSHENG, L., SALAS, W. and MOORE, B., 2005, Mapping paddy rice agriculture in southern China using multitemporal MODIS images. *Remote Sensing of Environment*, 95, pp. 480–492.
- ZARCO-TEJADA, P., BERJON, A., LÓPEZ-LOZANO, R., MILLER, J., MARTIN, P., CACHORRO, V., GONZALEZ, M. and FRUTOS, A., 2005, Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, **99**, pp. 271–287.
- ZHANG, X., FRIEDL, M., SCHAAF, C., STRAHLER, A., HODGES, J., GAO, F., REED, B. and HUETE, A., 2003, Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84, pp. 471–475.