Monitoring of biophysical parameters in vineyards through hyperspectral reflectance techniques

Renan Tosin

Mestrado em Engenharia Agronómica
Departamento de Geociências, Ambiente e Ordenamento do Território
2018

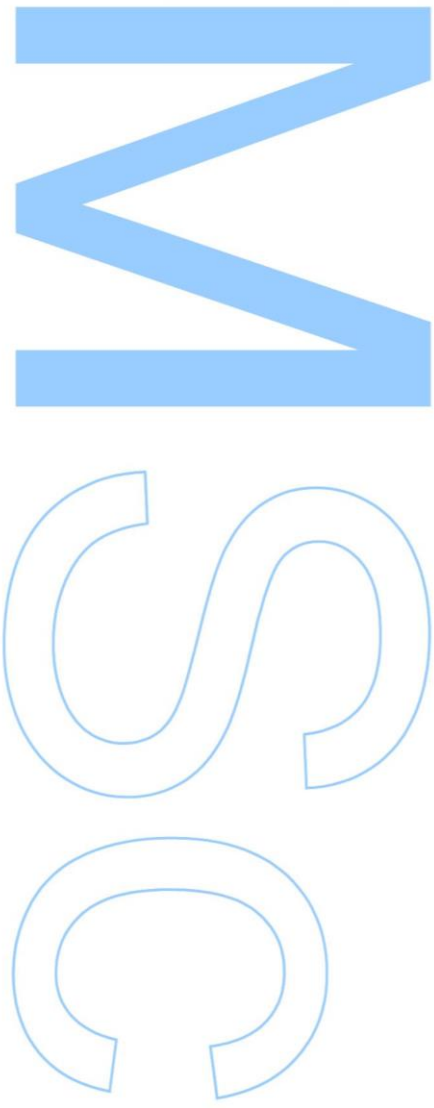
Orientador
Doutor Mário Manuel de Miranda Furtado Campos Cunha, Professor Auxiliar, FCUP

Coorientadora
Doutora Isabel Maria Valgôde Alves Pôças, Investigadora, FCUP

Todas as correções determinadas
pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____



Previous Note

In the elaboration of this dissertation, we opted for an integral incorporation of a set of submitted research works or in preparation for publication in international indexed journals with cientific arbitration, which are included in the Section 3 (“Case Studies”) of this dissertation. In the presented articles, the candidate participates as first author in two of them, being co-author of another. The candidate clarifies that in the two articles tha appears as first author, despite the collaboration of the other authors, participated in the data collection, he was responsible for its design, as well as for its obtaition, analysis and discussion of the results, and for the elaboration of the published version. The article wich the candidate collaborated as co-author, gathers data of three years (2014, 2015 and 2017), being the candiated participated in the data collection of 2017 and collaborated in the data analysis and in the writing of the final version of the article.

Part of the activities of this dissertation derives from the “Wine-spectra” project that aims to assess the hydric condition of grapevines in the Douro with remote sensing techniques. This project results from a protocol between the Faculty of Sciences of the University of Porto (FCUP), the Association for the Development of Viticulture Duriense (ADVID), and the companies Real Companhia Velha and Symington Family Estates. The responsability of “Wine-spectra” is given by Doctor Mário Cunha, Assistant Professor of the Department of Geosciences, Environment and Spatial Planning of the Faculty of Sciences of the University of Porto and Researcher at INESCTEC and Doctor Isabel Pôças, researcher at the Center for Research in Geo-Spatial Sciences of the University of Porto, who are also the tutors of this dissertation. The author of this dissertation is part of the reseracher’s team of the “Wine-spectra” project since 2017.

Nota Prévia

Na elaboração desta dissertação, optamos pela incorporação integral de um conjunto coerente de trabalhos de investigação submetidos ou em preparação para publicação em revistas internacionais indexadas e com arbitragem científica, os quais integram a Secção 3 ("Case Studies") da presente dissertação. Nos artigos apresentados, o candidato participa como primeiro autor em 2 deles, sendo co-autor de outro. O candidato esclarece que nos dois artigos em que aparece como primeiro autor, não obstante a colaboração de outros autores, participou na colheita dos dados, foi o responsável pela sua conceção, bem como pela obtenção, análise e discussão de resultados, e ainda pela elaboração da sua forma publicada. O artigo em que o candidato colaborou como co-autor, integra dados de 3 campanhas (2014, 2015 e 2017), tendo o candidato participado, na colheita dos dados da campanha de 2017 e colaborou na análise dos dados e, ainda, na redação da versão final deste artigo.

Parte das actividades desta dissertação enquadram-se nas linhas de trabalho do projecto "Wine-spectra" que tem como objectivo avaliar a condição hídrica da videira na região do Douro com base em técnicas de detecção remota. Este projeto resulta do protocolo de colaboração entre a Faculdade de Ciências da Universidade do Porto (FCUP), a Associação para o Desenvolvimento da Viticultura Duriense (ADVID), e as empresas Real Companhia Velha e *Symington Family Estates*. A responsabilidade do "Wine-spectra" está a cargo do Doutor Mário Cunha, Professor Auxiliar do Departamento de Geociências, Ambiente e Ordenamento do Território da Faculdade de Ciências da Universidade do Porto e Investigador no INESCTEC e da Doutora Isabel Pôças, Investigadora no Centro de Investigação em Ciências Geo-Espaciais da Universidade do Porto, que são também os orientadores desta dissertação. O autor desta dissertação integra a equipa de investigação do projecto "Wine-spectra" desde o ano de 2017.

Acknowledgments

I would like to thank everyone who has following me since I started my master's degree. My family and friends have always supporting me in each decision that I took during my student's life. Also, the Faculty of Sciences of The University of Porto (FCUP) for welcomed me with the best wishes and for has always dared me to achieve a new step in my student career. I have enjoyed each minute that I passed at FCUP. The friends that I made at FCUP and the teachers who helped me to enrich my knowledge will never be forgotten.

For these reasons, I express my special grateful to:

- My beloved mother Lourdes, my stepfather Marcos, my dead godfather Renato, my godmother Meire, my cousin Cibele and my dead "Australian mother" Margaret for have always supporting me in all my academic decisions. Since my first adventure out home when I moved to a different city to study, until when I decided to go overseas. I also thank them for dealing well with all my stress and anxiety crisis during the development of this dissertation.
- My dear long-date friends Luíz, Leandro, Lucas and Ana for showing their truly friendship since we were studying in the same classroom many years ago and for always supporting my decisions in personal and academic life. Also, my friends Érika, Tainá, Leandro and Rui for their advices and support during my stressful moments.
- I also give a special thank to my tutor, Professor Doctor Mário Cunha for his brilliant orientation, patience and serenity to explain to me an area of knowledge that were not well-known by me. Professor Mário has always pushed me to accept new challenges, showing alternatives for problems that seemed impossible to solve and always tried to motivate me to go further during the development of this dissertation.
- Similarly, I thank my co-tutor Doctor Isabel Pôças for all the knowledge transmitted to me. Isabel has always showed available to teach me how to use the equipments and softwares required during the realization of this work. She has helped me in the interpretation of the results and showed me possible paths to discuss the results obtained. Moreover, I thank her for the excellent text reviewing during the elaboration of the multiple documents originated form this thesis.

- Lastly, I thank the opportunity to participate in the cooperation protocol "WineSpectra" that involves FCUP, Associação para o Desenvolvimento da Viticultura Duriense (ADVID), and two wine companies in the Douro Valley: Real Companhia Velha and Symington Family Estates. Without these supports this work could not be done.

"Choose to focus your time, energy and conversation around people who inspire you, support you and help you to grow you into your happiest, strongest, wisest self."

Karen Salmansohn

Abstract

The Mediterranean climate, characterized by warm and drought summers, gathers ideal conditions for the practice of viticulture. Such conditions are found in Douro region, in the Northeast of Portugal, where viticulture represents the most important agricultural and economic activity. The application of remote sensing (RS) tools in precision agriculture (PA) is revolutionizing the input management in agriculture, including in viticulture. This dissertation presents three case studies of RS techniques application in Portuguese vineyards, based on physiological assumptions and supported by modelling approaches. The first case study is related to the use of hyperspectral and thermal data to assess the application of kaolin in grapevines. According to the bibliography, kaolin is reported as a potential tool to minimize the effect of high temperatures over the vegetation during the Mediterranean summer. The main findings presented in this case study report the wavelength 535 nm, related to foliar pigments responsible for the heat dissipation (i.e. xanthophyll), and the wavelength 733 nm as possible indicators of thermal stress promoted by the solar radiation. Moreover, a principal component analysis (PCA) showed that the thermal data also highly contribute for explaining the vines response to kaolin application throughout the grapevine cycle, while considering the time-losing effect after kaolin application. The other two case studies present different modelling approaches for the estimation of predawn leaf water potential (ψ_{pd}) in vineyards of two sub-regions in Douro wine region using hyperspectral data: (i) the second case study presents a generalized predictive model of (ψ_{pd}) based on machine learning techniques applied in regression mode – Random Forest (RF), Bagging Trees (BT), Gaussian Process Regression (GPR), and Variational Heteroscedastic Gaussian Process Regression (VH-GPR); (ii) in the third case study a multisite predictive model of (ψ_{pd}) were developed based on a machine learning technique applied in the classification mode – ordinal logistic regression (OLR) using the data-set from 2017. A large set of spectral vegetation indices (VIs) were tested, optimized, and validated as models' predictors. In both case studies, the VIs selected encompassed the visible and near-infrared (NIR) zones of the electromagnetic spectrum. Analyzing model's performance, the second case study provided overall accuracy between 82-83% and the third case study an overall accuracy of 73.2% in the estimation of classes of ψ_{pd} values. The results obtained from both case studies showed potential for the models' operational use to estimate the ψ_{pd} in support to irrigation management.

Keywords: grapevine; hyperspectral; machine learning; predawn leaf water potential; vegetation indices.

Resumo

O clima mediterrânico, caracterizado por verões quentes e secos, reúne condições ideais para a prática da viticultura. Tais condições são encontradas na região do Douro, no nordeste de Portugal, onde a viticultura representa a cultura mais importante e a principal atividade económica. A aplicação de técnicas de deteção remota em agricultura de precisão está a revolucionar o gerenciamento de *inputs* na agricultura, incluindo a viticultura. Esta dissertação apresta três casos de estudo de aplicação de deteção remota em vinhas Portuguesas, abordando parâmetros fisiológicos através de modelos matemáticos. O primeiro caso de estudo é relacionado com o uso de informação hiperespectral e térmica para avaliar a aplicação de caulino em videiras. De acordo com a bibliografia, o caulino é relatado como uma potencial técnica para minimizar os efeitos das altas temperaturas na vegetação durante o verão mediterrânico. Os principais resultados apresentados neste caso de estudo reportam o comprimento de onda 535 nm, relacionado com pigmentos foliares responsáveis pela dissipação do calor (p.e. xantofilas), e o comprimento de onda 733 nm como possíveis indicadores de estresse térmico promovidos pela radiação solar. Além disso, uma análise de componentes principais (PCA) mostrou que a informação térmica também contribui fortemente para a explicação da resposta espectral da vinha sobre a aplicação do caulino no ciclo da videira, quando considerado o efeito de perda do caulino após a sua aplicação. Os outros dois casos de estudo apresentam diferentes abordagens de modelação para a estimação do potencial hídrico de base na folha ao amanhecer (ψ_{pd}) em vinhas de duas sub-regiões da região vinícola do Douro usando dados hiperespectrais: (i) o segundo caso de estudo apresenta um modelo generalista para a previsão do (ψ_{pd}) baseado em técnicas de *machine learning* aplicadas em modo de regressão - *Random Forest* (RF), *Bagging Trees* (BT), *Gaussian Process Regression* (GPR), e *Variational Heteroscedastic Gaussian Process Regression* (VH-GPR); (ii) no terceiro caso de estudo é apresentado um modelo multi-local para a previsão do (ψ_{pd}), baseado em técnicas de *machine learning* em modo de classificação – regressão ordinal logística (OLR) utilizando os dados do ano de 2017. Foi testada uma grande variedade de índices de vegetação, otimizada, e validada como preditores dos modelos. Em ambos os casos de estudo, os índices de vegetação selecionados abrangeram as zonas do espectro eletromagnético do visível e do infravermelho próximo. Analisando a performance dos modelos, o segundo caso de estudo obteve uma exatidão média entre 82-83% e o terceiro caso de estudo uma exatidão média de 73.2 % na estimação dos valores das classes de ψ_{pd}. Os resultados obtidos de ambos casos de estudo mostraram potencial operacionalização dos modelos para estimar o ψ_{pd} como suporte do maneio da rega.

Palavras chave: videira, hiperspectral, *machine learning*, potencial hídrico de base da folha ao amanhecer, índices de vegetação.

Summary

List of tables

List of figures

List of abbreviations

3BSI: Three-Band Spectral Indices

AIS: Airborne Hyperspectral Imager

ARI: Anthocyanin Reflectance Index

ARTMO: Automated Radiative Transfer Models Operator

ATS: Applications Technology Satellite

AVHRR: Advanced Very High-Resolution Radiometer

CARI: Chlorophyll Absorption Ratio Index

CI: Curvature Index

ClAInt: Chlorophyll Absorption Integral

CRI: Carotenoid Concentration Index

D: Derivative Reflectance Indices

DD: Double Difference Index

DDn: New Double Difference Index

DODGE: Department of Defense Gravity Experiment

DPI: Double Peak Index

DVI: Derivate Vegetation Index

DWSI: Disease–Water Stress Indices

EGFN: Edge-Green First Derivative Normalised Difference

EGFR: Edge-Green First Derivative Ratio

EO: Earth Observation

EO-1: Earth Observing-1

EOS-AM: Earth Observing System Morning Crossing (Ascending) Mission

EROS: Earth Resources Observation Satellite

ERTS-1: Earth Resources Technology Satellite - 1

EVI: Enhanced Vegetation Index

G: Soil Heat

GDVI: Difference NIR/Green Green Difference Vegetation Index

GI: Greenness Index

GMI: Gitelson and Merzlyak Index

GPS: Global Position System

H: Sensible Heat

LAI: Leaf Area Index

LASSO: Least Absolute Shrinkage and Selection Absolute Shrinkage and Selection Operator

LNC: Leaf Nitrogen Concentration

MCARI: Modified Chlorophyll Absorption in Reflectance Index

mND: Modified Normalized Difference

mND705: Modified Normalized Difference

mNDVI: Modified Normalized Difference Vegetation Index

MODIS: Moderate Resolution Imaging Spectroradiometer

MPRI: Modified Photochemical Reflectance Index

mREIP: Modified Red-Edge Inflection Point

MSAVI: Modified Soil Adjusted Vegetation Index

mSR: Modified Simple Ratio

MTCI: MERIS (Medium Resolution Imaging Spectrometer) Terrestrial Chlorophyll Index

MTVI: Modified Triangular Vegetation Index

NASA: National Aeronautics and Space Administration

ND: Normalized Differences

NDVI: Normalized Difference Vegetation Index

NIR: Near Infrared

NPCI: Normalized Pigment Chlorophyll Index

OSAVI: Optimized Soil Adjusted Vegetation Index

PA: Precision Agriculture

PARS: Photosynthetically Active of Reflectance Spectra

PCR: Principal Components Regression

PLSR: Partial Least Squares Regression

PRI: Photochemical Reflectance Index

PRI_norm: Normalized Photochemical Reflectance Index

PSND: Pigment Specific Normalised Difference

PSRI: Plant Senescence Reflectance Index

PSSR: Pigment Specific Simple Ratio

PWI: Plant Water Index

R^2: Determination Coefficient

RADAR: Radio Detection and Ranging

RDVI: Renormalized Difference Vegetation Index

REP_LE: Red-Edge Position Linear Interpolation

REP_Li: Red-Edge Position Linear Interpolation

RMSE: Root Mean Square Error

Rn: Net Radiation

RR: Ridge (Regulated) Regression

RS: Remote Sensing

SAVI: Soil-Adjusted Vegetation Index

SIPI: Structure Intensive Pigment Index

SMLR: Stepwise Multiple Linear Regression

SPVI: Spectral Polygon Vegetation Index

SR: Simple Ratio

SRPI: Simple Ratio Pigment Index

SRWI: Simple Ratio Water Index

Sum_Dr: Sum of First Derivative Reflectance

SWIR: Shortwave Infrared

TB: Tinta Barroca

TCARI: Transformed Chlorophyll Absorption Ratio

TF: Touriga Franca

TGI: Triangular Greenness Index

TIR: Thermal Infrared

TIROS: Television Infrared Observation Satellite

TIROS-N: Advanced Television Infrared Observation Satellite

TN: Touriga Nacional

TVI: Triangular Vegetation Index

US: United States

USSR: Union of Soviet Socialist Republics

VI: vegetation index

λe: Latent Heat

ρ: Reflectance

ψ_{pd}: Predawn Leaf Water Potential

Scope of the dissertation

This dissertation presents an overall review of the main concepts related with RS techniques and their application for agricultural purposes.

Specifically, RS applications related to the monitoring of crop water status and leaf protective energy dissipation in vineyards are considered. The use of machine learning techniques in the modelling of biophysical parameters based on hyperspectral and thermal data is tested. Three case studies are presented, corresponding to three articles submitted or prepared for publication in peer reviewed scientific journals.

The dissertation involved training in fieldwork skills associated with vineyard surveys and remote sensing data collection using field hyperspectral spectroradiometer and thermocam; technical skills associated with the collecting of plant spectra, plant predawn leaf water potential; statistical, and analytical skills associated with hyperspectral and thermal remote sensing data processing (e.g. spectral vegetation index optimization) which included computer programming skills through the application of R and Matlab specific applications.

Dissertation structure

The dissertation is organized in five sections. First, an introduction section presents a contextualization of the the topics covered throughout the dissertation, followed by a state of the art section, providing the most relevant information on agricultural remote sensing applications. Afterwards, three case studies are presented in a specific section where the main dissertation' findings are explained: i) Spectral and thermal data as a proxy for leaf protective energy dissipation under kaolin application in grapevine cultivars; ii) Toward a generalized predictive model of grapevine water status in Douro region from hyperspectral data; iii) Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region. An overall discussion about the results of the three case studies is presented in the following section, comparing with the findings in the bibliography. Finally, the main gaps existing in the use of spectral information in precision agriculture are presented in a section of conclusions and perspectives.

1. Introduction

The modelling of the data collected by RS optical sensors has been increasingly used in the agrarian sciences domain aimed at estimating different biophysical parameters from crops. The spectral reflectance allows monitoring crop parameters such as: crop water condition (Pôças et al., 2017), incidence of plagues and diseases (Ray et al., 2017), crop nutrients content (Samborski et al., 2016), leaf area index (Cheng et al., 2014), and foliar pigments content (e.g. xanthophyll) used as proxy of plant physiological processes (Middleton et al., 2012).

The spectral response from the crops can be obtained through proximity sensors (e.g., handheld spectroradiometers) and sensors onboard drones, aeroplanes or satellites, with different spatial, temporal and spectral resolutions. The quantity and width of bands captured from the electromagnetic spectrum characterizes the spectral resolution and allow distinguishing between multispectral or hyperspectral data. The hyperspectral data differ from multispectral due to the higher number of spectral bands with narrow width, allowing more easily detecting subtle changes in the energy reflected by the objects surface (Jones and Vaughan, 2010).

Often, the implementation of spectrally-based methodologies to estimate plant biophysical parameters follows statistical approaches. The large amount of data generated through RS tools and the often complex pattern of biophysical parameters data demands the selection of appropriate statistical techniques (e.g. machine learning algorithms) and predictors adjusted to the study conditions (Cheng et al., 2014; Gorsevski and Gessler, 2009). Nevertheless, these statistical approaches represent semi-empirical methodologies, hindering models transferability when applied to conditions differing from those tested in their development (Rivera et al., 2014).

Therefore, despite the wide potential of hyperspectral data for crops monitoring, their operational use for assessing crop biophysical parameters with complex data patterns, like those related with crop water status, still needs more/better knowledge. From this limitation, this work aims to assess the potential of hyperspectral reflectance collected from proximity sensors (handheld spectroradiometer and thermal camera) for estimating vines biophysical parameters. The study focuses in vineyards of two sub-regions within Douro wine region and data collected in different years, thus encompassing variability conditions regarding agronomic, environmental and climatic features. Various statistically-based methodologies focusing on machine learning algorithms are tested for training the big data obtained from the proximity sensors aiming to estimate the target parameters.

2. State of the art

2.1. Precision agriculture (PA)

PA, also known as precision farming, represents the application of geospatial techniques and tools, including remote sensing, to identify in-field soil and crop variability and to deal with them using alternative strategies aiming for improving farming practices and optimizing agronomic inputs (Khanal et al., 2017; Zhang and Kovacs, 2012), with the aim to improve the agronomic output while reducing the input, that is, producing 'more with less'.

The PA has begun with the use of global positioning system (GPS) in the agriculture in the 1980's following mainly crop mechanization (Zhang et al., 2002). In the early 2010s, PA was boosted by sophisticated technology such as smart sensors, remote sensed data, actuators and micro-processors, high bandwidth cellular communication, nanotechnology, cloud-based Information and communication technology systems as well as data sciences. As such, data is no longer only sourced from the farm equipment, but new services coupled with new algorithms are being available to transform data into actionable intelligence (Zarco-Tejada et al., 2014).

In Portugal, the viticulture is one of the most important perennial crops, with the grapes produced for wine representing the second large area of permanent crops, totalizing approximately 176805 hectares with a total of 6558 tons of wine produced (INE, 2018). The use and optimization of geospatial techniques and tools in this activity possibly will bring a smart approach on how this crop is conducted. Precision viticulture can be defined by monitoring and managing spatial variation linked to physical and biochemical parameters related to yield and quality (Hall et al., 2002). According to these authors, the relationship between the yield and quality are complex and the cultivars, climates and season are affected by the management, being the soils one of the factors with high variability. With the development of precision viticulture, it is possible to map the within vineyards variability regarding multiple parameters (Thenkabail et al., 2000), such as water status (Borgogno-Mondino et al., 2018), chlorophyll content (Zarco-Tejada et al., 2013b) and grape ripening (Orlandi et al., 2018).

2.2. Platforms of obtaition of spectral data

Different types of RS sensors and platforms can be used to collect spectral data (Figure 1). As an example, proximity sensors like handheld spectroradiometers allow obtaining

hyperspectral data from the energy reflected by the vegetation (Verrelst et al., 2018), as represented in Figure 1 by the device closest to the vegetation. Other examples are the spectral and thermal cameras on-board drones (Gago et al., 2017), airborne platforms (Serrano et al., 2002), and satellites (Huang et al., 2018) as illustrated in Figure 1.

Figure 1. Different platforms for Earth observation data collection (senseFly, 2015).

2.3. The electromagnetic spectrum and spectral signatures

The radiation usually refers to the electromagnetic radiation, which is the form of energy that travels at the speed of light and have wave-like properties (e.g., wavelength and frequency). The electromagnetic spectrum describes the range of frequencies of the electromagnetic radiation and it is possible to define wavelength as the distance of wave crests in a specific frequency that hits a determined point in one second (Hungate et al., 2008). The electromagnetic spectrum encompasses radio waves, microwaves, infrared, visible light, ultraviolet, x-ray and gama-ray (Schmitt, 2002).

The Figure 2 represents the electromagnetic spectrum and characterizes its energy and frequency according to the wavelengths. As the wavelength increases, the energy and frequency decreases. This kinetics can be explained by the following formula:

$$C = \lambda . V$$

Where, C is the speed of light in the vacuum (299.792,458 km/s), λ the wavelength (nm) and V the frequency of the wavelength (Hz).

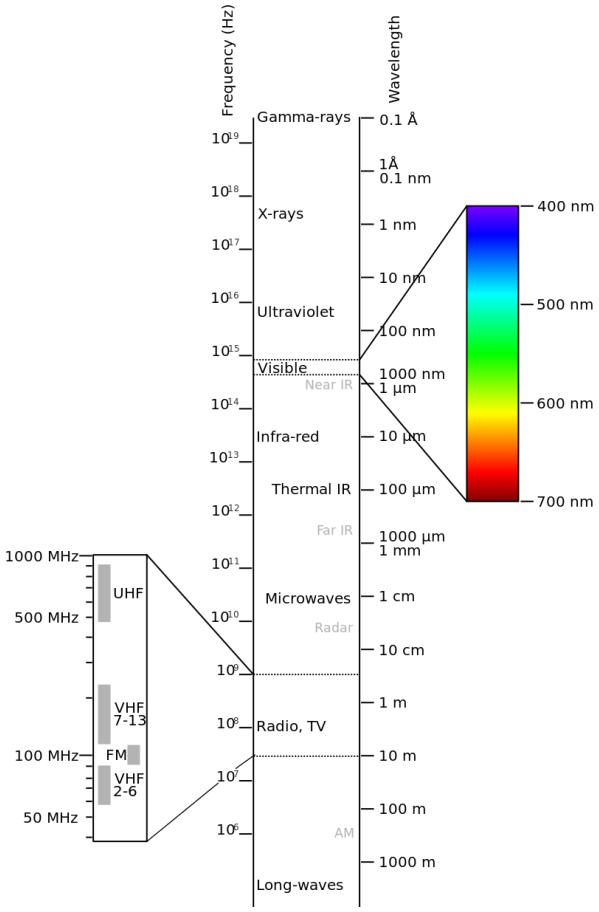


Figure 2. Representation of the electromagnetic spectrum, encompassing the various domains and their corresponding energy and frequency (Penubag, 2012).

All the radiant energy received from the sun that is not absorbed or dissipated, is transformed in other form of non-radiant energy such as soil heat (G), biochemical energy (P, photosynthesis), sensible (H) and latent heat (λE) as represented in Figure 3 (except P flow). The relationship between radiant and non radiante energy flows, which represents the energy balance in the Earth surface is presented in equation 1.

The net radiation (Rn) integrates the total amount of incoming and outgoing radiation components, accounting for the shortwave and longwave radiation portions (Petropoulus, 2013). The shortwave radiation corresponds to the radiation with short wavelengths (0.4 – 2.5 µm) that usually refers to the part of the electromagnetic spectrum dominated by the incoming solar radiation; the longwave radiation is the radiation with longer wavelengths (2.5 – 14 µm) that typically corresponds to the part of the electromagnetic spectrum dominated by the radiation emitted by the surface of Earth and atmosphere (Campbell and Norman, 1998). Through the energy balance equation, it is possible to estimate the crops evapotranspiration that broadly represents crop water status and, consequently, its water

needs (Anapalli et al., 2018; Ortega-Farías et al., 2016; Pôças et al., 2013). The Rn is partitioned into sensible heat (H), soil heat (G) and latent heat (λE) fluxes.

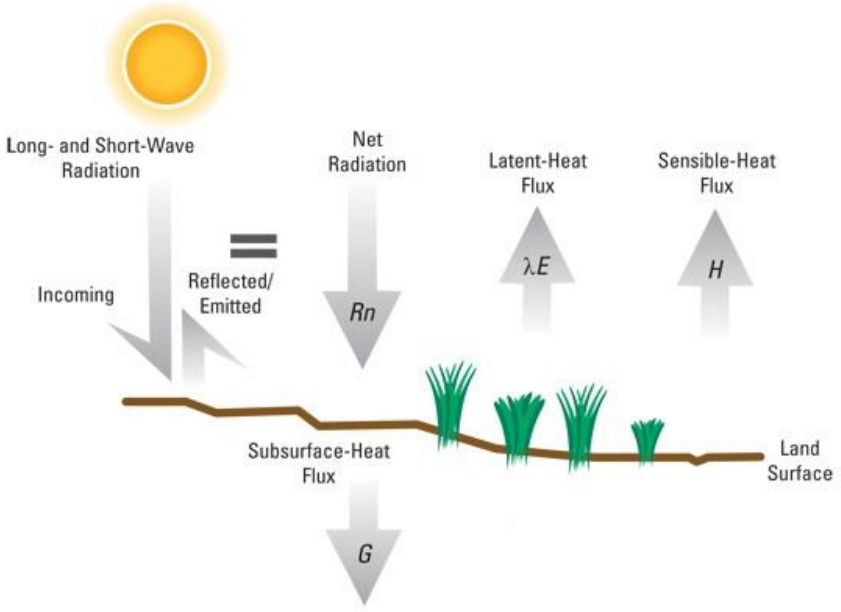


Figure 3. The energy balance diagram in Earth adapted from DeMeo et al. (2003)

$$Rn = H + G + \lambda E \quad (1)$$

The Rn and G are components that can be measured or estimated based on climatic parameters, while determining the H component is more challenging and requires the accurate measurement of temperature gradients above the surface (Allen et al., 1998). λE represents the evapotranspiration fraction and is derived from all the others components of the energy balance equation (Allen et al., 2011; Allen et al., 1998).

The incident radiation reflected is unique for each material/object due to its biophysical composition (Hungate et al., 2008), guaranteeing a spectral signature for each object that is sampled. Figure 4 illustrates the different spectral signatures of green vegetation, soil and water.

Green vegetation leaves are characterized by a higher absorption (low reflectance) in the visible wavelengths and lower absorption (high reflectance) in the infra-red zone. When specifically focusing the visible domain, there is a lower absorption (higher reflectance) of the radiation in the green zone when compared to blue and red zones This reflectance profile of vegetation in the visible domain is mainly characterized by the pigments presence (Hall et al., 2002). For example, chlorophylls absorption of light is mainly associated with the red and blue zones, while carotenoids and xanthophylls are reported to the blue-green

zones (Jones and Vaughan, 2010). The lower absorption of radiation and higher reflectance in the infra-red zone is associated to the water content. The reported water absorption bands corresponding to 0.97 µm, 1.20 µm, 1.40 µm and 1.94 µm wavelengths (Curran, 1989), support that the infrared zone is potentially useful as indicator for water status in plants (Peñuelas et al., 1997; Rodríguez-Pérez et al., 2018).

The soil spectral signature is characterized by the increment of the reflectance from the visible to the near-infrared zone. However, there are many factors that influences the soil reflectance, including the soil water content (Borgogno-Mondino et al., 2018; Jones and Vaughan, 2010; Roosjen et al., 2015), the level of organic matter (Jones and Vaughan, 2010), the soil texture (Jones and Vaughan, 2010) and roughness (Eshel et al., 2004; Jones and Vaughan, 2010), and the iron content (Chicati et al., 2017; Jones and Vaughan, 2010).

Conversely, the spectral signature of the water has a small proportion of reflectance of the incident radiation when compared to the green vegetation and soil spectral signatures (Jones and Vaughan, 2010). There are some factors that affect the water reflectance including the water turbidity, chlorophyll content, surface and below surface content related with sediments and phytoplankton (Jones and Vaughan, 2010; Zeng et al., 2017).

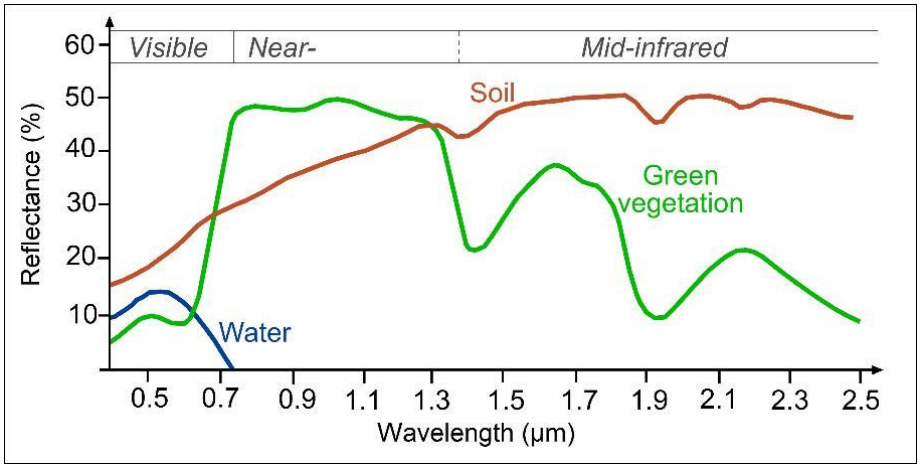


Figure 4. Curve of reflectance of soil, water and green vegetation from the visible until the mid-infrared zone (SEOS, 2018).

Towards an agricultural point of view, the reflectance signature of the crops under specific agronomic conditions provides an understanding of the different crops response due to the environmental conditions (Mulla, 2013).

2.4. Spectral information for sensing vegetation properties

The spectral information has a high potential to describe vegetation properties (Hall et al., 2002; Jones and Vaughan, 2010) due to its power to depict the effects of the

environment conditions over the plants (Féret et al., 2017). The changes in the crop reflectance in different regions of the electromagnetic spectrum can be associated to a wide range of biophysical parameters, including crop water status (Pôças et al., 2017; Rodríguez-Pérez et al., 2018), fungal disease (Ray et al., 2017), fertilization content (Li et al., 2018), fruit maturation (Orlandi et al., 2018; zhang et al., 2016), chlorophyll fluorescence (Rapaport et al., 2015) and canopy temperature (Verrelst et al., 2015b).

The various zones of the electromagnetic spectrum can be explored to assess specific properties of the vegetation (and soil): (i) visible and near infrared (NIR) are related with plant structural characteristics and pigments content (Chen et al., 2015; Datt, 1999), (ii) short wave infrared (SWIR) relates with plant moisture (Lobell et al., 2001; Rodríguez-Pérez et al., 2018), (iii) thermal infrared (TIR) provides thermal information from the crops and can be further related with crop water conditions (Poblete et al., 2018), and (iv) microwaves allows determining soil moisture (Schellberg et al., 2008). In the context of this dissertation, a particular focus will be devoted to the visible, NIR and thermal domains.

The visible zone provides information about changes in the concentration of xanthophyll and other plant pigments and are associated to the biochemical absorption in the mesophyll (Blackburn, 1998; Knipling, 1970). The different pigments present in the leaves permit to assess physiological plant condition through the photosynthesis rate. Among the leaf's pigments, it is important to highlight the family of chlorophylls, carotenoids and anthocyanins. The combination of pigments in the leaves will control the sunlight that is absorbed and used in the photosynthesis, even though, the pigments try to control the excess of light and heating that is received from the sun (Féret et al., 2017). Under stress conditions of temperature, nutrition, and water, the radiation absorbed will decrease due to the defence mechanism of plants that is controlled by the xanthophyll that promote the dissipation of the heating (Middleton et al., 2012). Additionally, under stress, the canopy can present other symptoms related to its vigour, such as: changes in the angle of the leaves, foliar area, concentration of pigments in the epidermis, shape of the plant and depth of the root (Middleton et al., 2012). Regarding the NIR zone, various studies report its relation with plant water content and plant water stress in liaison with the water absorption bands in this spectral zone as referred in section 2.3 (Peñuelas et al., 1993; Peñuelas et al., 1997; Rodríguez-Pérez et al., 2018). Thus, spectral information from the visible and NIR domains, with relation with plant structural characteristics and pigments content, can be used as proxy of plants response to stress conditions, including those from water stress (Suárez et al., 2008).

The thermal data has provided positive evidences that environmental conditions influences the canopy (Buitrago Acevedo et al., 2017; Sepulcre-Cantó et al., 2006). Thermal data is reported to influence in pigments concentration (Shellie and King, 2013), water status (Zarco-Tejada et al., 2013a), lignin and cellulose concentrations and leaf area (Buitrago Acevedo et al., 2017).

The combination of optical and thermal data is able to improve the crop monitoring due to the detailed information provided (Khanal et al., 2017).

Furthermore, the leaf pigments play an important rule when pretends to create biophysical models. The models PROSPECT (Jacquemoud and Baret, 1990) and PROSAIL (Verhoef, 1984, 1985) have been largely used since their creation (Jacquemoud et al., 2009). These models use as input the leaf reflectance, leaf transmittance and soil reflectance (Jacquemoud et al., 2009). These models are basically applied to assess the canopy architecture (e.g. leaf area, angle distribution) and leaf biochemical content (e.g. chlorophyll content) (Féret et al., 2017; Jacquemoud et al., 2009).

2.5. Hyperspectral and multispectral data in precision agriculture

As mentioned in the introduction section in RS, the spectral resolution represents the ability of a sensor to define fine wavelength intervals and is determined by the number and width of spectral bands (Jones and Vaughan, 2010). The terms multispectral and hyperspectral are defined according to the number and width of spectral bands recorded by a sensor. While multispectral refers to instruments that record information in a small number of spectral bands (usually two to ten) and with a broad width, hyperspectral is associated to sensors recording information from a large amount of bands (more than ten) with fine width (Hall et al., 2002). Handheld spectroradiometers and cameras onboard satellites and drones are examples of devices that collect multispectral and hyperspectral data. Although multispectral data are more easily available, often associated to satellite missions (e.g., Landsat and Sentinel missions), there is an increasing interest over hyperspectral data for purposes related with the assessment and monitoring of vegetation biophysical parameters.

The main advantage of hyperspectral data reports to the high detailed information provided, allowing the differentiation of multiple surfaces due to its high sensibility to detect differences owing to the narrow band information (Blackburn, 2007; Jones and Vaughan, 2010; Mariotto et al., 2013). However, hyperspectral data has a high level of bands

correlation, in other words, many wavelengths provide similar information, creating dimensionality problems and making hard the bands selection (Jones and Vaughan, 2010). The optimization of VIs and the application of machine learning techniques are used to reduce the dimensionality of the bands. Additionally, the use of hyperspectral data is still limited due to the high cost of data acquisition and processing (Hamzeh et al., 2016).

2.6. Vegetation indices

The VIs are mathematical combinations of two or more spectral bands. There are different typologies of VIs, as summarized in Table 1. Most VIs consider two spectral bands. However, two-bands indices are more likely to suffer saturation problems mostly because of the environmental conditions of the reflectance (Wu et al., 2008). To overcome the saturation problem, VIs with three or more bands have been used in the construction of VI (Wang et al., 2012).

Table 1. Examples of types of VI. Adapted from: Verrelst et al. (2015c).

Type	Formula	Example
ρ (reflectance)	ρ_a	Boochs (Boochs et al., 1990)
SR (simple ratio)	ρ_a/ρ_b	GI (Smith et al., 1995)
ND (normalized differences)	$(\rho_a - \rho_b)/(\rho_a + \rho_b)$	NDVI (Tucker, 1979)
mSR (modified simple ratio)	$(\rho_a - \rho_c)/(\rho_b - \rho_c)$	SIPI (Peñuelas et al., 1995b)
mND (modified normalized difference)	$(\rho_a - \rho_b)/(\rho_a + \rho_b - 2\rho_c)$	mNDVI (Sims and Gamon, 2002)
3BSI (three-band spectral indices)	$(\rho_a - \rho_c)/(\rho_b + \rho_c)$	Datt (Datt, 1999)
3BSI Wang (three-band spectral indices Wang)	$(\rho_a - \rho_b + 2\rho_c)/(\rho_a + \rho_b - 2\rho_c)$	LNC Wang (Wang et al., 2012)
3BSI Tian (three-band spectral indices Tian)	$(\rho_a - \rho_b - \rho_c)/(\rho_a + \rho_b + \rho_c)$	LNC Tian (Tian et al., 2014)
DVI (derivate vegetation index)	$\rho_a - \rho_b$	DVI (Jordan, 1969)

VI can be used to highlight changes in vegetation condition (Hall et al., 2002) and to estimate a vast number of parameters such as plant water content, stomatal conductivity, chlorophyll fluorescence, soil water content and others (Katsoulas et al., 2016). The Attachment 1 reports multiple formulation of VI, for example, the curvature index (CI), related to the chlorophyll a fluorescence (Zarco-Tejada et al., 2003), the spectral polygon vegetation index (SPVI) associated to the pant nutrition (Vincini et al., 2006) and the plant senescence reflectance index (PSRI), which tells information about the plant senescence and fruit ripening (Merzlyak et al., 1999). Additionally, it is important to highlight the most popular VIs, such as the normalized difference vegetation index (NDVI) (Tucker, 1979) related to the plant biomass content, the soil-adjusted vegetation index (SAVI) (Huete, 1988) that describes the soil-vegetation system, and the photochemical reflectance index (PRI) (Gamon et al., 1992) used to assess the photosynthesis efficiency.

The VIs can be either broadband or narrowband depending if they are built with bands with a large range of wavelengths (e.g., encompassing the full range of the red and NIR zones) or with bands corresponding to specific wavelengths. A large diversity of VIs is currently available in the bibliography (Attachment 1). Additionally, several studies based on hyperspectral data have aimed at optimizing VIs, through the selection of the best combination of bands (Verrelst et al., 2015a; Wang et al., 2017). This optimization of VIs is performed to obtain the best band combination consonant with the target variable. The selection of the band combination considers statistical parameters such as root mean square error (RMSE) and determination coefficient (R^2) (Li et al., 2018; Rodríguez-Pérez et al., 2018). For example, Pôças et al. (2017) optimized normalized indices to find the best combination of wavelengths to explain the water status in vines in the Douro Wine Region. Similarly, Elvanidi et al. (2018) searched for the best combination of bands to explain the water stress in soilless tomato through VI in a controlled growth chamber located in Velestino, Central Greece and Sims and Gamon (2002) developed the modified simple ratio index (mSR705) VI to better explain the chlorophyll content in diverse species leaves collected from plants growing on the California State University, Los Angeles campus, or from natural habitats in Southern California and Southern Nevada (Nevada Test Site, Mercury, NV).

2.7. The assessment of RS data in agriculture

To provide a very rich information about the crop condition, the (hyper)spectral data needs to be extracted, stored and analysed (Huang et al., 2018). Figure 5 broadly

summarizes the steps for the spectral data processing and analysis when applied for PA purposes. However, this type of data brings the challenge of data analysis (Heikkinen, 2018; Katsoulas et al., 2016) and the methods of data processing (Verrelst et al., 2018).

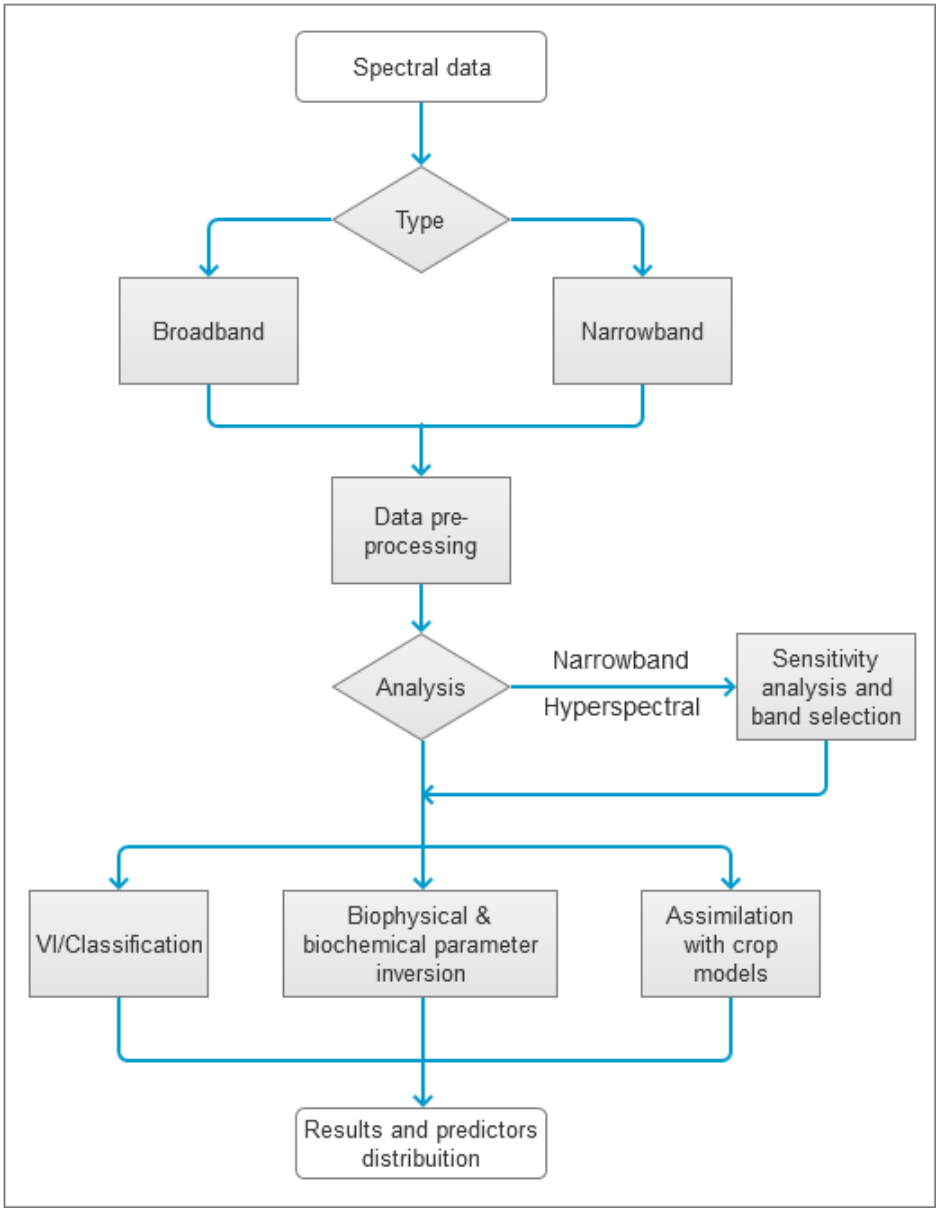


Figure 5. Chartflow of the analysis of spectral images in precision agriculture.

The processing and analysis of spectral data can be complex when a large amount of data, with redundant information, is captured along with the target information (Huang et al., 2018). That is the case when hyperspectral data is assessed as a high correlation between bands can occur and thus the band selection constitutes an important challenge (Feng et al., 2017; Huang and He, 2005; Wang et al., 2017). Another issue related with the large amount of information collected with either proximity sensors or sensors onboard aerial and satellite

platforms (imagery), generating the so called big data (Huang et al., 2018), concerns the data processing demands. A high computational performance and robust modelling tools are required to analyse large amounts of data (Huang et al., 2018; Verrelst et al., 2018). Nevertheless, the rising improvements in both sensors technology and computation associated with the increasing efforts to develop robust modelling approaches positively contribute to overcome the mentioned challenges and limitations.

Techniques of machine learning have been widely used to improve the modelling in PA (Verrelst et al., 2018; Wang et al., 2017). Parametric and non-parametric models have been used in modelling regressions. In a very simple way, what differs the non-parametric methods from the parametric method is the use of weights in the training data and the fact that non-parametric methods do not assume a specific distribution of the data, in other words, it is possible to say that the non-parametric methods learn with the training of the data (Verrelst et al., 2018). Also, the non-parametric methods are subclassified in linear and non-linear methods. While the linear non-parametric methods englobe methods such as: stepwise multiple linear regression (SMLR) (Draper and Smith, 2014); principal components regression (PCR) (Wold et al., 1987); partial least squares regression (PLSR) (Geladi and Kowalski, 1986), ridge (regulated) regression (RR) (Geladi and Kowalski, 1986) and least absolute shrinkage and selection absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996), the non-linear non-parametric methods are focused in methods such as decision-trees, artificial neural-networks and kernel-based machine learning methods (Patrício and Rieder, 2018; Verrelst et al., 2018). Authors have been using these techniques to optimize the estimation of biophysical parameters that will help in decision making processes related with crop production (Elvanidi et al., 2018; Orlandi et al., 2018; Pôças et al., 2017; Rodríguez-Pérez et al., 2018; Veraverbeke et al., 2018; Verrelst et al., 2018).

The modelling of biophysical parameters requires high computational requirements due to their exhaustive computational process and long-time processing (Huang et al., 2018; Verrelst et al., 2018). Techniques of modelling are applied in software such as MATLAB (The Math Works, Inc., USA) and R (R Core Team, 2017) due to their versatility to use different toolboxes in the case of MATLAB, and multiple packages in R. Both ARTMO (Automated Radiative Transfer Models Operator) (Verrelst et al., 2011), which is a software that runs in MATLAB environment, and the package HSDAR (Lehnert et al., 2017) for R include features of machine learning able to handle multispectral and hyperspectral data. Also, the package caret (Kuhn, 2008) for R includes multiples algorithms to create non-parametric and parametric models that can be applied to RS data.

Moreover, through these tools, it is possible to create physically-based models such as PROSPECT (Jacquemoud and Baret, 1990), as well as create and optimize VI and construct nonparametric and parametric models, which have a statistical basis (Caicedo et al., 2014; Lehnert et al., 2017; Verrelst et al., 2011), aimed at retrieving biophysical parameters of vegetation. Examples of the use of these features are given by Verrelst et al. (2015c) who used ARTMO to estimate the leaf area index (LAI) by choosing the band combination of bands from multispectral imaging. Also, Dechant et al. (2017) used the HSDAR package to estimate the chlorophyll, carotenoid and water contents by PROSPECT model and Pôças et al. (2017) used the same HSDAR package to find the best combination of bands to optimize normalized indices, and then used the caret package to construct predictive models for estimating the water status in vine.

3. Case studies

The articles presented follow the same editorial pattern of the dissertation. However, the number of figures, tables and references follow the manuscript requirements of journal that they were submitted, and they are not included in the table of contents.

3.1. Spectral and thermal data as a proxy for leaf protective energy dissipation under kaolin application in grapevine cultivars

Journal: The Open Agriculture Journal
Date of submission: 2/10/2018
Manuscript number: OPAG-D-18-00143
Status: under review

Open Agriculture

Spectral and thermal data as a proxy for leaf protective energy dissipation under kaolin application in grapevine cultivars

--Manuscript Draft--

Manuscript Number:	OPAG-D-18-00143
Full Title:	Spectral and thermal data as a proxy for leaf protective energy dissipation under kaolin application in grapevine cultivars
Short Title:	
Article Type:	Research Article
Keywords:	handheld spectroradiometer; leaves reflectance; leaves temperature; thermal camera; xanthophyll
Corresponding Author:	Mário Cunha PORTUGAL
Corresponding Author's Institution:	
First Author:	Renan Tosin, Licenciatura
First Author Secondary Information:	
Order of Authors:	Renan Tosin, Licenciatura
	Isabel Pôças
	Mário Cunha

Figure 6. Receipt of submission of manuscript number 1.

Note: this paper is based on the “long abstract” previously presented by the same authors in the “I Symposium Ibérico de Ingeniería Hortícola”, Lugo, Spain, 21-23 February, 2018.

Spectral and thermal data as a proxy for leaf protective energy dissipation under kaolin application in grapevine cultivars

Renan Tosin[1], Isabel Pôças[1,2,3,4] and Mário Cunha[1,3,4,*]

[1] Faculdade de Ciências da Universidade do Porto (FCUP), Rua do Campo Alegre sn, 4169-007 Porto, Portugal; renantosin@gmail.com

[2] Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Portugal; isabel.pocas@fc.up.pt

[3] Centro de Investigação em Ciências Geo-espaciais (CICGE), FCUP

[4] Institute for Systems and Computer Engineering, Technology and Science (INESCTEC)

[*]Corresponding author: mccunha@fc.up.pt

Abstract

The dynamic effects of kaolin clay particle film application on the temperature and spectral reflectance (325 to 1075 nm) of leaves of two autochthonous varieties (Touriga Nacional and Touriga Franca) were studied in Douro wine region. The study was implemented during the summer of 2017, in conditions prone to multiple environmental stresses that include excessive light and temperature as well as water shortage. Light reflectance from kaolin-sprayed leaves was higher than the control (leaves without kaolin) in all dates. The kaolin protective effect over leaves temperatures was low on the 20 days after application and ceased about 60 days after its application. At this latter date, the leaves protected with kaolin recorded higher temperatures than leaves without kaolin. Differences between leaves with and without kaolin were explained by the normalized maximum leaf temperature, reflectance at 400 nm, at 535 nm, and at 733 nm. The wavelengths of 535 nm and 733 nm are associated with plant physiological processes, which support the selection of these variables for assessing kaolin effects on leaves. The application of a principal component analysis based on these four variables allowed explaining 85% of data variability, obtaining a clear differentiation between leaves with and without kaolin treatment. The normalized maximum leaf temperature and the reflectance at 535 nm were the variables with a greater contribution for explaining data variability. The results improve the understanding of the vines response

to kaolin throughout the grapevine cycle and support decisions about the re-application timing.

Keywords: handheld spectroradiometer, leaves reflectance, leaves temperature, thermal camera, xanthophyll.

1. Introduction

Temperature and solar radiation coupled with water shortage are central to how climate influences the growth, yield and quality of grapevine. Despite the advances in technology, breeding included, as well as the incremental use of irrigation in vineyards, grapevine remains highly dependent on weather, which can affect both the quantity and quality of the harvest. In semi-arid environments, typical of Mediterranean climate, excessive radiant heat load combined with excessive temperature, often limits the plants physiological processes and growth during the rainless summer, mainly in no irrigation crop system such as vineyard. According to Palliotti et al. (2015), these climate factors can have synergetic effects on CO_2 assimilation related with stomata conductance or damage leaves pigments responsible for protective energy dissipation which can limit growth, yield, and fruit composition.

In Douro wine region, potential multiple abiotic stresses are very frequent during summer as a consequence of the high solar radiation, excessive temperature and the elevated gradients of the water vapour pressure between the leaves and the air (Chaves and Rodrigues, 1987). This may be further exacerbated by the foreseen drier and warmer climate scenarios, despite the noticeable advances in vineyard technologies (Cunha and Richter, 2016; Giorgi and Lionello, 2008; IPCC, 2007; Moriondo et al., 2015). Therefore, adaptive cultural practices are needed to mitigate the negative impact of actual and future climate scenarios in order to keep the competitiveness of the wine industry.

Modifications of the solar radiation balance in the plant leaves can be obtained by spraying foliage with a white suspension of inert reflective material such as a kaolin-particle film. The kaolin is a product originated from clay that transmits visible light and gases necessary for photosynthesis, while reflecting ultraviolet and infrared bands. According to (Glenn et al., 2002), the kaolin application reduces temperatures of treated leaves through increment of reflectivity radiations (e.g. ultraviolet), promoting the activation of enzymes responsible for the impairment of protective leaf pigments related with energy dissipation (Shellie and King, 2013).

Among the protective pigments presented in the leaves, it is important to highlight the family of chlorophylls, carotenoids and anthocyanins, which are associated with light harvesting and energy dissipation and varies greatly according to the environmental factors.

The combination of pigments in the leaves will control the sunlight that can be absorbed and used in the photosynthesis, even though, the pigments try to control the excess of light and heating that is received from the sun in several horticultural crops, orchards, and grapevines (Dinis et al., 2016; Moutinho-Pereira et al., 2007; Palliotti et al., 2015; Shellie and King, 2013; Zarco-Tejada et al., 2013), tobacco (Falcioni et al., 2017), Persian walnut (Gharaghani et al., 2018) and others species (Féret et al., 2017; Gitelson et al., 2003; Middleton et al., 2012; Müller et al., 2001). Under stress conditions of temperature, humidity, nutrition, and water, the radiation absorbed will decrease due to the defence mechanism of plants controlled by the xanthophylls, which promotes the dissipation of energy (Middleton et al., 2012). Nevertheless, under stress, the canopy can present other symptoms related to its vigour, such as: changes in the angle of the leaves, foliar area, concentration of pigments in the epidermis, shape of the plant and depth of the root (Middleton et al., 2012).

The application of a kaolin-particle film has been reported to protect canopy against excessive light and temperature as well as pest and diseases pressure (Ferrari et al., 2017; Glenn and Puterka, 2010) in different crop systems. Many studies demonstrate that kaolin-film application can be an important practice for reduced midday leaf temperature (AbdAllah et al., 2018; Jifon and Syvertsen, 2003), leaf to air vapour pressure differences net gas exchange (Glenn et al., 2010), chlorophyll fluorescence (Shellie and King, 2013) as well as for regulating plant water status overcoming the negative impact of water stress (AbdAllah, 2017), preserving plant growth, yield and quality (Djurović et al., 2016; Gharaghani et al., 2018). The application of a kaolin-particle film over the grapevines in Douro region has shown good results to protect the plants from the effects of heating and radiation (Dinis et al., 2016) and its application in commercial vineyards is often considered due to its low cost for application. These authors have also demonstrated that kaolin is able to improve the quality of the grape under stress conditions by improving its concentration of phenols, flavonoids, anthocyanins and vitamins C.

Remote sensing techniques have been applied in crop production to provide a better understanding of the crop's spectral response under different agronomic and environmental (e.g., soil, climate) conditions (Féret et al., 2017). Spectral-derived data provide information about plant status. These data reflect the changes in leaf pigments, which vary seasonally in pigment composition, enabling them to be used as a proxy of physiological processes (Hall et al., 2002; Jones and Vaughan, 2010), including the epoxidation state of xanthophyll cycle pigments and chlorophyll fluorescence emission (Gamon et al., 1997; Moya et al., 2004; Ustin et al., 2009).

Thermal imagery can also be used to detect canopy temperature differences associated to plant responses (Sepulcre-Cantó et al., 2006), e.g., to different pigments

concentration (Shellie and King, 2013) and to assess water status of the canopy (Zarco-Tejada et al., 2013). The combination of the optical spectral data and thermal imagery opens up new avenues for ecophysiological observations, potentially providing insights into the remote detection/monitoring of the viticulture practices impact on the abiotic multiple stresses.

Although the effects of kaolin applications on diseases control and leaf and canopy physiology, plant growth, yield and fruit composition have been deeply studied, their time-dynamics impacts on leaf spectral reflectance modification are not well understood.

The main goal of this study was to evaluate the effect of kaolin on the leaf microclimate, throughout the grapevine growth cycle. We hypothesize that kaolin application would reduce the temperature through the increase reflectivity of leaves and the analysis of the leaf's reflectance pattern could be a proxy for the physiological process of protective energy dissipation through leaf pigments.

2. Materials and methods

The study was conducted in a commercial vineyard (Quinta dos Aciprestes, Real Companhia Velha) located in Soutelo do Douro (latitude 41.24ºN; longitude 7.43ºW) in Douro wine region (Northeast of Portugal). The region is characterized by a climate of Mediterranean type, with high temperatures and low precipitation during the summer period as shown in the Figure 1.

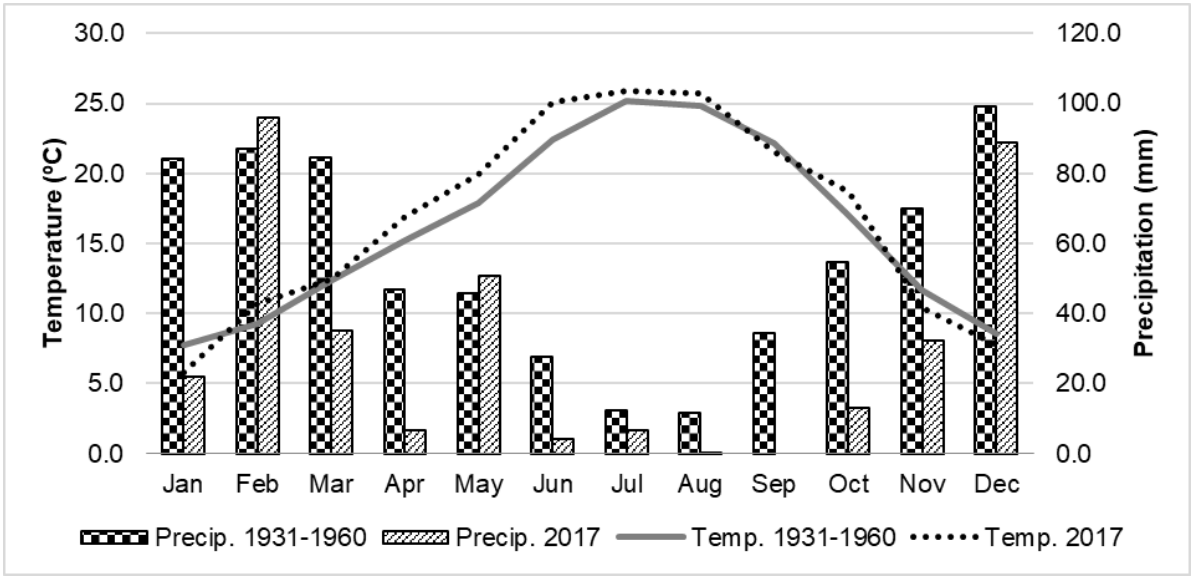


Figure 1. Temperature and precipitation characterization of the Douro wine region for the 30-years period between 1931-1960 Ferreira (1965) and comparison with the temperature and precipitation records in 2017 during study period.

The Douro valley is one of the most non-irrigated arid wine regions of the world, with strong and consistent post-flowering water, thermal and solar radiation stresses. These multiple stresses occur mainly during the ripening period (20th July to 1st September), where the precipitation is typically lower than 20 mm, accompanied by high values of solar irradiation and temperature and also elevated gradients of the water vapour pressure between the leaves and the air (Reis and Lamelas, 1988).

The study was conducted in the year 2017, which was characterized by low precipitation levels from April to September (Figure 1).

The vineyard has a total area of 1.17 ha, with an undulating terrain and an average slope of 25%. The vines were planted according to a bilateral Royat system, following the orientation Northeast-Southwest, with 2.2 m × 1 m plant spacing and a maximum plant height of 1.5 m. Two cultivars were considered: Touriga Nacional (TN) and Touriga Franca (TF). A porous kaolin-particle film (a “SunProtect” by Epagro) was applied over both cultivars leaves during the summer period. The concentration followed the instructions of the manufacture (5 kg /100 L of water). The kaolin-particle film was applied in June 6 (DOY 156) in the whole vine. The application was performed by spraying over the vines canopy, covering a large percentage of the leaves but also leaving unsprayed leaves.

Ground measurements of spectral reflectance data and thermal data were collected in four dates of 2017: July 5, 29 days after kaolin application (DAA 29), July 20 (DAA 44), August 3 (DAA 57), and August 31 (DAA 85). Four leaves with kaolin-particle film treatment and another four leaves without kaolin-particle film treatment, of each one of the two cultivars, were randomly selected and collected for data acquisition in each date. An exception was considered for DAA 29, when no data were collected for TF. A total of 56 observations of spectral reflectance data and thermal data were collected.

The spectral reflectance data were collected using a portable spectroradiometer (Handheld 2, ASD Instruments) recording spectral signatures between 325 nm and 1075 nm of the electromagnetic spectrum, with a wavelength interval of 1 nm. The measurements were done in nadir position, in cloud free conditions, between 11 and 14 h (local time) aiming to minimize changes in solar zenith angle. Ten repetitions per leaf were performed and later averaged to minimize the effect of noise. Only data between 400 nm and 1010 nm were considered for further analysis due to noise at the limits of the spectrum.

Thermal image data were collected using a portable thermal camera (Flir Systems, Inc.) in the same leaves where spectral reflectance data were measured. Thermal images were processed using Flir Tools 6.3.17227.1001 software (Flir Systems). The maximum temperature of the leaves was obtained by the analysis of the thermal images. A normalization of the maximum temperature of the leaves (T_max_f_N) was performed,

dividing by the hourly air temperature at the time of measurements, in order to minimize the effect of the air temperature increase during the 3-hours period of data acquisition in each date.

Leaf temperature differences between means of kaolin treatment within grape varieties for each date of measurements were assessed by one-way ANOVA and these mean differences were considered statistically significant if the F-test from ANOVA gives a significance level P (P value) below 0.05. The Duncan test was performed to do the separation of means between kaolin treatments and cultivars. These statistical analyses were performed in R (R Core Team, 2017) through the packages “car” (Fox and Weisberg, 2011) for ANOVA and “agricolae” (Mendiburu, 2017) for the Duncan test.

An assessment of the spectral and thermal behaviour of the vineyard leaves following the kaolin-particle film treatment was performed through a principal component analysis (PCA). The reflectance data at each wavelength between 400 nm and 1010 nm, as well as the T_max_f_N were considered as variables. Previously to the PCA, a correlation analysis between variables was performed to reduce collinearity among variables. A threshold of 0.70 for the correlation value was selected based on Kuhn and Johnson (2013). Following the correlation analysis four variables were selected for the application of PCA: T_max_f_N, reflectance at 400 nm, reflectance at 535 nm, and reflectance at 733 nm. The software R (R Core Team, 2017), combined with the package “factoextra” (Kassambara and Mundt, 2017), was used for computing the correlation, cluster analysis and the PCA.

3. Results

3.1. Effects of kaolin on leaves temperature

The temperature measurements in the leaves of TN and TF cultivars in both kaolin treatments, as expected, were roughly 10 to 20ºC (DAA 57) higher than the air temperature, which was close to 30ºC in all dates (Figure 2).

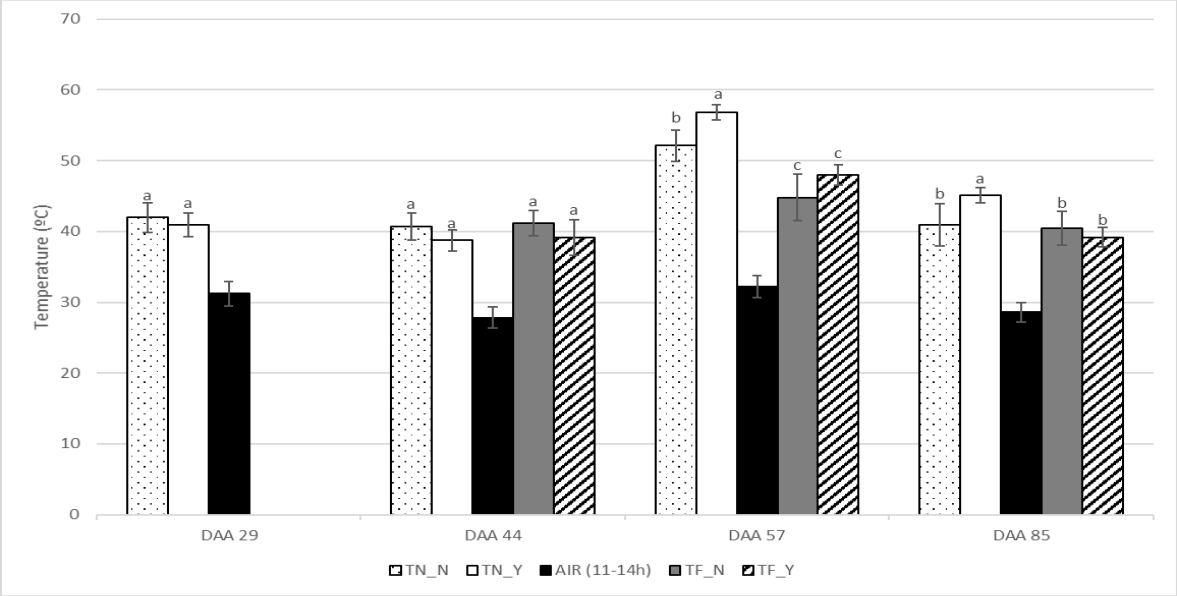


Figure 2. Average maximum temperatures of the leaves in Touriga Nacional without kaolin (TN_N), Touriga Nacional with Kaolin (TN_Y), Touriga Franca without kaolin (TF_N), Touriga Franca with kaolin (TF_Y) and the average maximum air temperature from 11 AM to 2 PM [AIR (11-14H)] in the field for the four measurement dates. Vertical bars are the standard deviation and the letters show the mean comparation through Duncan test with a confidence level of 95%.

In the two first dates of data acquisition, carried out in DAA 29 and in DAA 44, the temperatures in leaves with kaolin are always lower than those without kaolin, but these temperature differences are not statistically significant ($p>0.468$ DAA 29; $p>0.275$ DAA 44). The temperature of the leaves in the third measurement (DAA 57), in both treatments and varieties, reached values higher than 50ºC, being much higher than the previous measurements (DAA 29 and DAA 44). Also, in this date (DAA 57), the temperature in both grapevine cultivars is significantly ($p<0.000$) higher in the treatments with kaolin than without kaolin and the cultivar TN reached higher temperatures than TF in both treatments (Figure 2).

3.2. Effects of kaolin on leaves reflectance

Figure 3 presents the leaves spectral reflectance patterns grouped by varieties and kaolin treatments within each measurement date. Broadly, the light reflectance from kaolin-sprayed leaves were at least 2 times higher in the visible domain (<700 nm) and very similar in the near infrared domain (>700 nm) than in the control leaves for all dates. In both treatments, the cultivar TF, when compared with the TN, presents consistently higher levels of reflectance in the visible region. Consistently, the higher reflectance obtained in the leaves of TF suggest a correlation, but not statistically different, with lower temperature of leaves without kaolin in the DAA 29 and DAA 44.

The leaves partially lost the kaolin film as observed during field measurements in the DAA 57. In this third measurements date, the leaves with kaolin presented higher temperatures than leaves without kaolin (Figure 2) and the reflectance for both cultivars and treatments were very alike (Figure 3). This effect may be related with a long-term reduction of leaves ability to dissipate energy under conditions of heat and radiation stresses when they were treated with kaolin at an early stage.

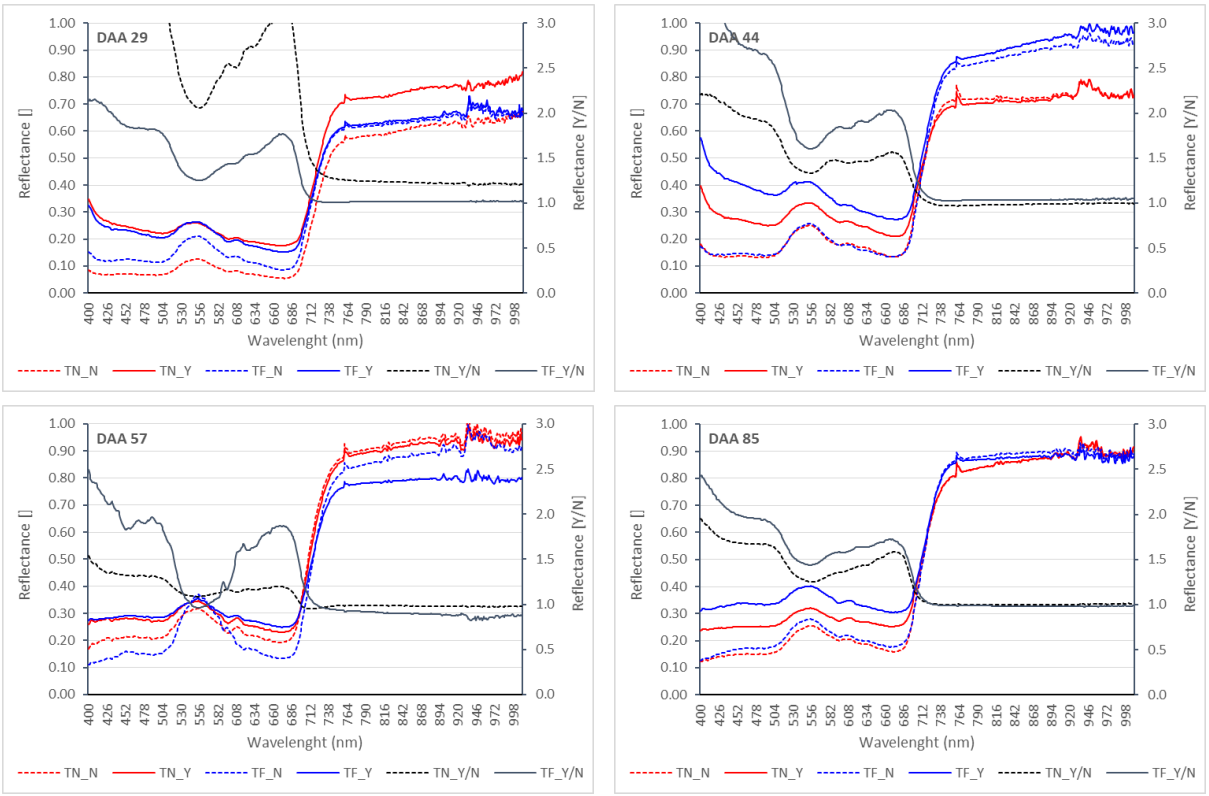


Figure 3. Average reflectance recorded in leaves of Touriga Nacional without kaolin (TN_N), Touriga Nacional with Kaolin (TN_Y), Touriga Franca without kaolin (TF_N), Touriga Franca with kaolin (TF_Y) during the four-measurement time. The secondary Y axis represents the ratio of reflectance of leaves with kaolin per leaves without kaolin for TN (TN_Y/N) and TF (TF_Y/N).

3.3. Effects of kaolin assessed through multivariate analyses

The Figure 4 presents the results for the principal component analysis (PCA) of leaf spectral and leaf temperature in axis ordination planes, with kaolin treatment and grapevine cultivars as response variables. The PCA results show the separation of leaves with (Y) and without kaolin (N), especially in the first (DAA 29) and second (DAA 44) dates. However, in the third (DAA57) and fourth (DAA85) dates, data of leaves with and without kaolin are closer.

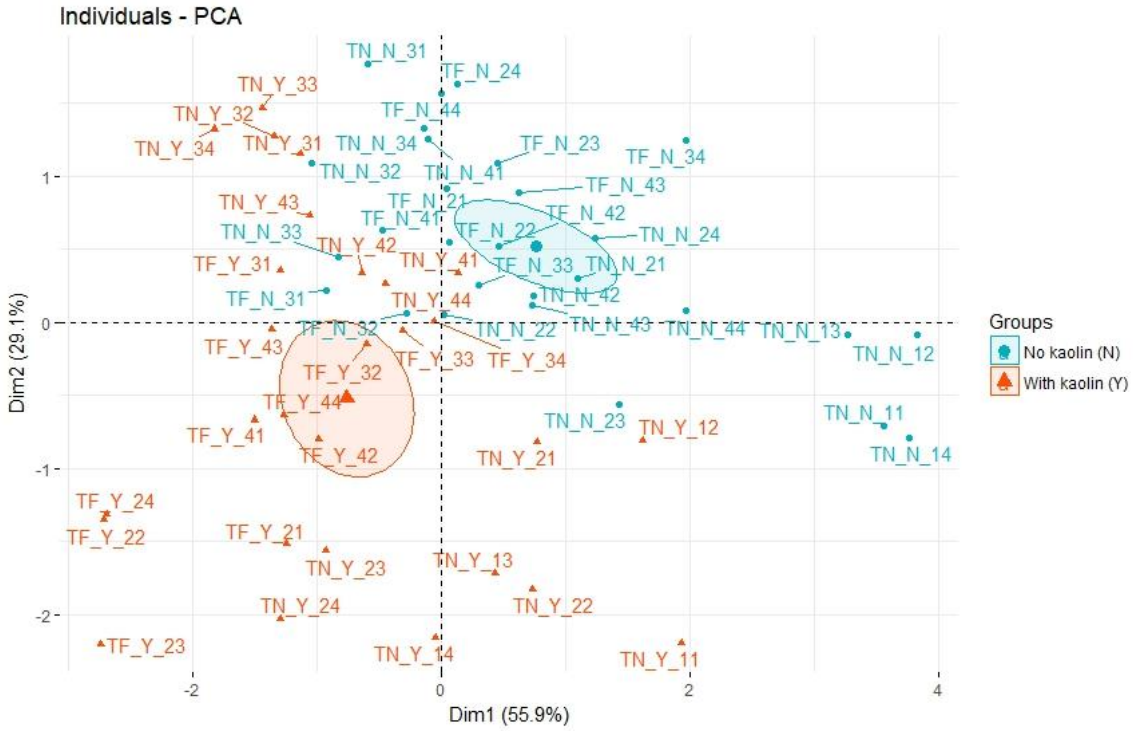


Figure 4. Principal Component Analysis for kaolin treatment in grapevine leaves (cultivars Touriga Franca – TF and Touriga Nacional – TN). The first digit refers to DAA: 1: DAA 29; 2: DAA 44; 3: DAA 57; 4: DAA 85 and the second digit is the repetition number.

The Table 1 presents the eigen values obtained from PCA and the contribution of the principal components. The first two principal components (Dim 1 and Dim 2) are enough to explain the data variability due to their high level of explanation (85%). The R535 and R733 present the higher contribution in Dim 1 (Y axis), while the T_max_f_N and the R400 show higher contribution in Dim 2 (X axis).

Figure 5 shows the data aggregation in six groups according to a cluster analysis. It is possible to see the combination of data from the first and second dates into close groups (G1-G3), and a larger dispersion among G4, G5 and G6 groups, encompassing leaves with and without kaolin during the third and fourth dates. Also, the first groups (G1-G3) are distinguished according to the cultivar, treatment and measurement dates. The first group (G1) includes observations of TN_N in the first date, while the second group (G2) principally groups data of TN_Y also in the first date. The third group (G3) mainly groups leaves of TN_N from the second data collection. The fourth group (G4) mostly gathers leaves of TF_Y for the second and fourth collecting dates. In the fifth group (G5), although only leaves without kaolin have been aggregated, there is an almost equal distribution of both cultivars during the third and fourth dates. In the sixth group (G6) it is observed the mix of leaves from both cultivars and treatments during the third and fourth measurements dates.

Table 1. Variance explained by PCA and the contribution of each principal component (Dim) as well as the contribution of each variable to the principal component.

	Eigen value	Variance %	Cumulative variance %	Contribution of each Principal Component			
				T_max_f_N	R400	R535	R733
Dim. 1	2.24	55.88	55.88	12.01	18.44	39.09	30.46
Dim. 2	1.16	29.10	84.98	44.54	40.13	4.86	10.47
Dim. 3	0.49	12.30	97.28	43.30	19.77	2.25	34.68
Dim. 4	0.11	2.72	100.00	0.16	21.67	53.79	24.38

PCA – Principal component analysis; T_max_f_N – maximum temperature of the leaves; R400 – reflectance at 400 nm; R535 – reflectance at 535 nm; R733 – reflectance at 733 nm.

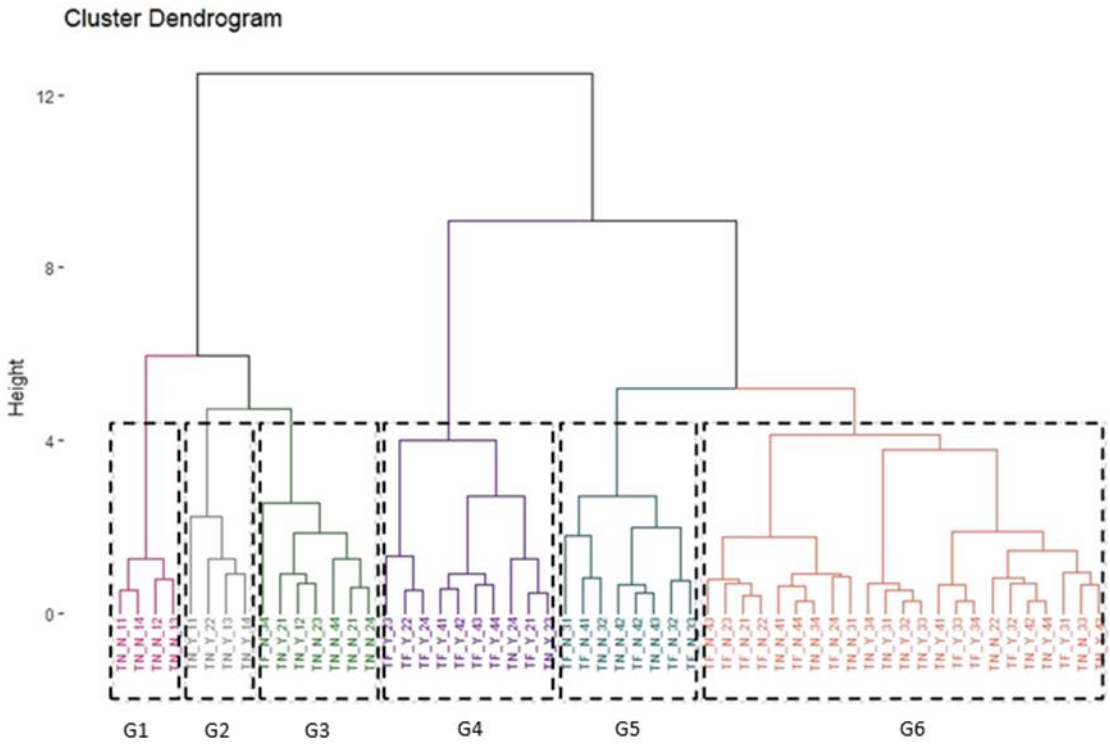


Figure 5. Cluster analysis grouped in six groups. Cultivars Touriga Franca without kaolin – TF_N, Touriga Franca with kaolin – TF_Y, Touriga Nacional without kaolin – TN_N and Touriga Nacional with kaolin – TN_Y. The first digit refers to DAA: 1: DAA 2

4. Discussion

The results presented in this paper show a decrease, although not statistically different, of the leaves temperature sprayed with kaolin particle film until DAA 44 (Figure 2). On the contrary, in the DAA 57, leaves with kaolin for both varieties achieved temperature

significantly higher ($p<0.000$) than leaves without kaolin. Equally, the effect of reflectance of kaolin-sprayed leaves (Figure 3) was consistently higher until DAA 44 than those of without kaolin. Additionally, in the spectral zone related to dissipation of heat by the leaves, around 531-535 nm (Middleton et al., 2012), the DAA 57 showed no reflectance differences between leaves from both treatments.

These findings suggest that the effect of kaolin is lower in the third and fourth dates of measurements due to the environmental factors that are responsible for the loose of the product in the leaves (Cantore et al., 2009), e.g., wind, precipitation occurrence. When the leaves loose the kaolin particle film and the stomates close, the plants will be more likely to be exposed under stress conditions and will reduce the leaf transpiration that would relief the leaves from the high heating (Glenn et al., 2010; Shellie and Glenn, 2008; Shellie and King, 2013). Consequently, the leaves with less kaolin would increase the temperature as shown in the Figure 2 and behave physiologically more similarly to leaves without kaolin.

The separation of leaves with and without kaolin shown in the Figure 5 can be explained by the variances found in different levels of both leaves reflectance and temperature. As shown by Shellie and King (2013) and observed in the Figure 3, the reflectance of leaves with kaolin is higher than leaves without kaolin. As the kaolin particles film indirectly induces the reduction of leaf temperature (Glenn et al., 2010; Ou et al., 2010; Shellie and King, 2013), the segregation into groups with and without kaolin appoints this circumstance. In addition, the group G6, which encompasses both leaves with and without kaolin from DAA 57 and DAA 85, suggest that kaolin loses its effect with time as discussed above and a new application is required to guarantee the same effect, as the last application of kaolin is usually at the beginning of grape ripening (Brillante et al., 2016).

The plant physiology can provide insights for explaining the differences on leaves temperature and reflectance between the cultivars observed. In a study conducted in the Douro Valley where three distinct cultivars were analysed, it was found that cultivars with low chlorophyll concentration are brighter and present a high rate of chlorophylls a and b (Chl a/b) that will promote a low photo absorbance and increase the photochemical rate (Moutinho-Pereira et al., 2007). This effect may be related with the differences in reflectance between TN and TF in DAA 57 (Figure 3), when TF reflectance was higher than TN. Additionally, kaolin will increase the whiteness of leaves that will promote high levels of reflectance and probably decrease the temperature (AbdAllah et al., 2018; Jifon and Syvertsen, 2003). Also, TN does not present high rate of photosynthesis per stomata conductance (Moutinho-Pereira et al., 2007), and the application of kaolin would promote high levels of internal CO2 concentration due to the lower temperature in the leaves and the increase of the photosynthesis per stomata conductance rate (Jifon and Syvertsen, 2003).

While many studies have shown that kaolin film application is a powerful tool to improve the crop development under stress conditions, others did not observe significant results with this practice. In a study with Persian walnut, the kaolin was positively efficient to reduce the leaf temperature, gas exchange rate, sunburn and it has also improved the quality of the nut and kernel (Gharaghani et al., 2018). Dinis et al. (2016) and Shellie and King (2013) have demonstrated that kaolin is able to improve the quality of the grape under stress conditions by improving its pigments concentration and also a better water use efficiency in some cultivars. However, some of these positive physiological effects of kaolin were recorded few days after kaolin application and/or under intensive treatments, namely with applications up to twice a week for three weeks (Brillante et al., 2016; Gharaghani et al., 2018; Glenn et al., 2010; Jifon and Syvertsen, 2003; Ou et al., 2010; Shellie and King, 2013). On the other hand, Dinis et al. (2016) sprayed kaolin in vines twice in a day to ensure the efficient adhesion of kaolin in the leaves and showed the good efficiency of kaolin in grapevine. Contrarily, Russo and Diaz-Perez (2005) did not obtain any improvement on physiological measurements, leaf temperature (also observed in the present study in the Figure 2), disease incidence or yields in two cultivars of peppers following the use of several kaolin applications under stress conditions. Also, Ou et al. (2010) did not observed significant differences in water potential, ºBrix, pH and titratable acidity with the application of kaolin in Merlot grapevine cultivar, although the kaolin impact on the studied parameters varied with different levels of irrigation. Thus, the meaningful differences found in this study are related to the reflectance of kaolin.

The variables selected to explain the separation of leaves with and without kaolin are related to physiological processes of the energy dissipation in leaves. It is notable the strong weight of the wavelength 535 nm in Dim. 1. The leaves with kaolin are mainly distributed on the side of the axis where R535 is projected, while the leaves without kaolin are on the opposite side. This area of the electromagnetic spectrum allows detecting subtle changes in the xanthophyll cycle pigment activity resulting from stress conditions, including thermal stress (Middleton et al., 2012). In the Figure 3 it is notable the visual difference of reflectance between the treatments during the four-measurement time. When the kaolin was highly present in the leaves, the reflectance at the wavelength 535 nm was always high when compared to leaves without kaolin, suggesting that the level of oxidated xanthophylls is lower in leaves with kaolin. Nevertheless, in the third measurement-time, when the leaves have lost part of the kaolin particle film, the reflectance at this wavelength is alike to the leaves without kaolin film. This observation may be explained by the comfort level that kaolin promotes on leaves. The leaves that were previously protected with the particle film might not have developed enough pigments to protect from the high incidence of solar radiation

when the particles film has lost its adherence to the leaves. This argument might support the not successful impact of the application of kaolin in the physiological parameters analysed by Russo and Diaz-Perez (2005).

The second variable ranked in Dim 1 is R733, which corresponds to the red edge region of the electromagnetic spectrum (sharp transition of vegetation's reflectance between red and near-infrared spectral ranges: 680 – 740 nm). The red edge has been often considered as an indicator of plant stress responses (Behmann et al., 2014; Filella and Penuelas, 1994; Zarco-Tejada et al., 2013). The Figure 3 shows the differences of reflectance in this zone of the spectrum between the leaves of both treatments, mainly in the first two measurement dates (DAA 29 and DAA 44). During the first measurement, the reflectance of leaves with kaolin is higher, suggesting that these leaves are more likely to support stress conditions. As shown by Glenn et al. (2010), the kaolin particle film can increase the water potential in well-watered vines minimizing the heating stress. Though, from the second until the last measurement-time of our study, there is no visual difference from the reflectance in this zone. Nevertheless, it is noteworthy that all grapevines were under (increasing) moderate to high water stress throughout our study period.

In Dim 2, the T_max_f_N presents a dominant role in the distribution of the observations, which agrees with the effect of kaolin to control leaf temperature in vines under drought conditions (Dinis et al., 2016). Likewise, the lower temperature is associated to the better water use efficiency in vines with kaolin (Glenn et al., 2010). Yet, when the reflectance become alike to the leaves with and without kaolin, we suggest that there is no difference in the temperature in the leaves from both treatments and the leaves that were previously protected with particle film can also increase its temperature due to the lack of protection mechanisms presented in the leaves, such as the xanthophylls.

5. Conclusion

These preliminary results improve the understanding of the vines response to kaolin throughout the grapevine cycle and support decisions about the re-application timing. The results show that the kaolin film protects the leaves against the thermal stress promoted by the solar radiation, while the product is well dispersed in the leaves. The effect of the kaolin film on the leaves may be explained by the concentration of foliar pigments such as the xanthophyll, which is related to the dissipation of heat in the leaf. Even though this work has shown the potential of spectral and thermal data to explain the effect of kaolin in different cultivars of grapevines, physiological and biochemical analysis should be further tested in future work to complement and strengthen the findings presented in this paper. Additionally,

future studies should consider the effect of kaolin (e.g. potential increase in pH) in the wine quality of Portuguese varieties, in the particular climate conditions of Douro Valley.

Acknowledgements

I.P. acknowledges Portuguese Foundation for Science and Technology (FCT) for the Post-Doc research grant (SFRH/BPD/79767/2011) and for the Post-Doctoral grant of the project ENGAGE-SKA POCI-01-0145-FEDER-022217, co-funded by FEDER through COMPETE (POCI-01-0145-FEDER-022217). The authors also thank the wine company Real Companhia Velha (and its Coordinator for Viticulture Rui Soares) and Associação para o Desenvolvimento da Viticultura Duriense (ADVID) for the meteorological data and funding part of the field-work missions. Special thanks to Miguel Carretero for allowing the use of the Thermocam.

References

AbdAllah, A., 2017. Impacts of Kaolin and Pinoline foliar application on growth, yield and water use efficiency of tomato (Solanum lycopersicum L .) grown under water deficit: A comparative study. Journal of the Saudi Society of Agricultural Sciences. 10.1016/j.jssas.2017.08.001

AbdAllah, A.M., Burkey, K.O., Mashaheet, A.M., 2018. Reduction of plant water consumption through anti-transpirants foliar application in tomato plants (Solanum lycopersicum L). Scientia Horticulturae 235, 373-381. 10.1016/j.scienta.2018.03.005

Behmann, J., Steinrücken, J., Plümer, L., 2014. Detection of early plant stress responses in hyperspectral images. ISPRS Journal of Photogrammetry and Remote Sensing 93, 98-111.

Brillante, L., Belfiore, N., Gaiotti, F., Lovat, L., Sansone, L., Poni, S., Tomasi, D., 2016. Comparing Kaolin and Pinolene to Improve Sustainable Grapevine Production during Drought. PLoS One 11, e0156631. 10.1371/journal.pone.0156631

Cantore, V., Pace, B., Albrizio, R., 2009. Kaolin-based particle film technology affects tomato physiology, yield and quality. Environmental and Experimental Botany 66, 279-288. 10.1016/j.envexpbot.2009.03.008

Chaves, M., Rodrigues, L., 1987. Photosynthesis and water relations in grapevines response to environmental factors, In: Tenhunen, J.D.e.a. (Ed.), Plant Response to Stress-functional analises in Mediterranean Ecosystems. Springer Verlag, Berlin, pp. 279-290.

Cunha, M., Richter, C., 2016. The impact of climate change on the winegrape vineyards of the Portuguese Douro region. Climatic Change 138, 239-251. 10.1007/s10584-016-1719-9

Dinis, L.T., Bernardo, S., Conde, A., Pimentel, D., Ferreira, H., Felix, L., Geros, H., Correia, C.M., Moutinho-Pereira, J., 2016. Kaolin exogenous application boosts antioxidant capacity and phenolic content in berries and leaves of grapevine under summer stress. J Plant Physiol 191, 45-53. 10.1016/j.jplph.2015.12.005

Djurović, N., Ćosić, M., Stričević, R., Savić, S., Domazet, M., 2016. Effect of irrigation regime and application of kaolin on yield, quality and water use efficiency of tomato. Scientia Horticulturae 201, 271-278. 10.1016/j.scienta.2016.02.017

Falcioni, R., Moriwaki, T., Bonato, C.M., de Souza, L.A., Nanni, M.R., Antunes, W.C., 2017. Distinct growth light and gibberellin regimes alter leaf anatomy and reveal their influence on leaf optical properties. Environmental and Experimental Botany 140, 86-95. 10.1016/j.envexpbot.2017.06.001

Féret, J.B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sensing of Environment 193, 204-215. 10.1016/j.rse.2017.03.004

Ferrari, V., Disegna, E., Dellacassa, E., Coniberti, A., 2017. Influence of timing and intensity of fruit zone leaf removal and kaolin applications on bunch rot control and quality improvement of Sauvignon blanc grapes, and wines, in a temperate humid climate. Scientia Horticulturae 223, 62-71. 10.1016/j.scienta.2017.05.034

Ferreira, H.A., 1965. Normais climatológicas do continente, Açores e Madeira correspondentes a 1931-1960. Serviço Meteorológico Nacional, Lisboa.

Filella, I., Penuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. International Journal of Remote Sensing 15, 1459-1470. 10.1080/01431169408954177

Fox, J., Weisberg, S., 2011. An (R) Companion to Applied Regression, Second. Thousand Oaks (CA): Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Gamon, J.A., Serrano, L., Surfus, J.S., 1997. The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. Oecologia 112, 492-501. 10.1007/s004420050337

Gharaghani, A., Mohammadi Javarzari, A., Vahdati, K., 2018. Kaolin particle film alleviates adverse effects of light and heat stresses and improves nut and kernel quality in Persian walnut. Scientia Horticulturae 239, 35-40. 10.1016/j.scienta.2018.05.024

Giorgi, F., Lionello, P., 2008. Climate change projections for the Mediterranean region. Global and Planetary Change 63, 90-104. 10.1016/j.gloplacha.2007.09.005

Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J Plant Physiol 160, 271-282. 10.1078/0176-1617-00887

Glenn, D.M., Cooley, N., Walker, R., Clingeleffer, P., Shellie, K., 2010. Impact of kaolin particle film and water deficit on wine grape water use efficiency and plant water relations. HortScience 45, 1178-1187.

Glenn, D.M., Prado, E., Erez, A., McFerson, J., Puterka, G.J., 2002. A reflective, processed-kaolin particle film affects fruit temperature, radiation reflection, and solar injury in apple. Journal of the American Society for Horticultural Science 127, 188-193.

Glenn, D.M., Puterka, G., 2010. Particle Films: A New Technology for Agriculture, Horticultural Reviews. doi:10.1002/9780470650882.ch1

Hall, A., Lamb, D.W., Holzapfel, B., Louis, J., 2002. Optical remote sensing applications in viticulture - a review. Australian Journal of Grape and Wine Research 8, 36-47. 10.1111/j.1755-0238.2002.tb00209.x

IPCC, 2007. Climate Change 2007: the AR4 Synthesis Report. Edited by Rajendra K. Pachauri, IPCC Chairman, Andy Resinger, Head of Technical Support Unit, The Core Writing Team. Published by IPCC, Geneva, Switzerland, 114p, Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Jifon, J.L., Syvertsen, J.P., 2003. Kaolin Particle Film Applications Can Increase Photosynthesis and Water Use Efficiency of `Ruby Red' Grapefruit Leaves. Journal of the American Society for Horticultural Science 128, 107-112.

Jones, H.G., Vaughan, R.A., 2010. Remote sensing of vegetation: principles, techniques, and applications. Oxford University Press Inc., New York, USA.

Kassambara, A., Mundt, F., 2017. factoextra: Extract and Visualize the Results of Multivariate Data Analyses, 2017. R package version 1.0.5. https://CRAN.R-project.org/package=factoextra

Kuhn, M., Johnson, K., 2013. Applied predictive modeling. Springer Science+Business Media, New York. 10.1007/978-1-4614-6849-3

Mendiburu, F.d., 2017. agricolae: Statistical Procedures for Agricultural Research, R package version 1.2-8. https://CRAN.R-project.org/package=agricolae

Middleton, E.M., Huemmrich, K.F., Cheng, Y.-B., Margolis, H.A., 2012. Spectral Bioindicators of Photosynthetic Efficiency and Vegetation Stress, In: Thenkabail, P., Lyon, J., Huete, A. (Eds.), Hyperspectral Remote Sensing of Vegetation. Taylor & Francis Group, LLC, Boca Raton, pp. 265-288.

Moriondo, M., Ferrise, R., Trombi, G., Brilli, L., Dibari, C., Bindi, M., 2015. Modelling olive trees and grapevines in a changing climate. Environmental Modelling & Software 72, 387-401. 10.1016/j.envsoft.2014.12.016

Moutinho-Pereira, J., Magalhães, N., Gonçalves, B., Bacelar, E., Brito, M., Correia, C., 2007. Gas exchange and water relations of three Vitis vinifera L. cultivars growing under Mediterranean climate. Photosynthetica 45, 202-207. 10.1007/s11099-007-0033-1

Moya, I., Camenen, L., Evain, S., Goulas, Y., Cerovic, Z., Latouche, G., Flexas, J., Ounis, A., 2004. A new instrument for passive remote sensing: 1. Measurements of sunlight-induced chlorophyll fluorescence. Remote Sensing of Environment 91, 186-197.

Müller, P., Li, X.-P., Niyogi, K.K., 2001. Non-photochemical quenching. A response to excess light energy. Plant physiology 125, 1558-1566.

Ou, C., Du, X., Shellie, K., Ross, C., Qian, M.C., 2010. Volatile compounds and sensory attributes of wine from Cv. Merlot (Vitis vinifera L.) grown under differential levels of water deficit with or without a kaolin-based, foliar reflectant particle film. J Agric Food Chem 58, 12890-12898. 10.1021/jf102587x

Palliotti, A., Tombesi, S., Frioni, T., Silvestroni, O., Lanari, V., D'Onofrio, C., Matarese, F., Bellincontro, A., Poni, S., 2015. Physiological parameters and protective energy dissipation mechanisms expressed in the leaves of two Vitis vinifera L. genotypes under multiple summer stresses. J Plant Physiol 185, 84-92. 10.1016/j.jplph.2015.07.007

R Core Team, 2017. R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Reis, R., Lamelas, H., 1988. Statistical study of decade series of water balance and its components of potencial evapotranspiration calculated by Penman's method. Instituto Nacional de Meteorologia e Geofisica, Lisbon.

Russo, V., Diaz-Perez, J., 2005. Kaolin-based particle film has no effect on physiological measurements, disease incidence or yield in peppers. HortScience 40, 98-101.

Sepulcre-Cantó, G., Zarco-Tejada, P., Jiménez-Muñoz, J., Sobrino, J., De Miguel, E., Villalobos, F., 2006. Detection of water stress in an olive orchard with thermal remote sensing imagery. Agricultural and Forest Meteorology 136, 31-44.

Shellie, K., Glenn, D.M., 2008. Wine Grape Response to Foliar Particle Film under Differing Levels of Preveraison Water Stress. HORTSCIENCE 43, 1392–1397.

Shellie, K.C., King, B.A., 2013. Kaolin Particle Film and Water Deficit Influence Malbec Leaf and Berry Temperature, Pigments, and Photosynthesis. American Journal of Enology and Viticulture 64, 223-230. 10.5344/ajev.2012.12115

Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A., Zarco-Tejada, P., 2009. Retrieval of foliar information about plant pigment systems

from high resolution spectroscopy. Remote Sensing of Environment 113, S67-S77. https://doi.org/10.1016/j.rse.2008.10.019

Zarco-Tejada, P.J., González-Dugo, V., Williams, L.E., Suárez, L., Berni, J.A.J., Goldhamer, D., Fereres, E., 2013. A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. Remote Sensing of Environment 138, 38-50. 10.1016/j.rse.2013.07.024

3.2. Toward a generalized hyperspectral based predictive model of grapevine water status in Douro region

Journal: Agricultural and Forest Meteorology

Date of submission: 25/10/2018

Manuscript number: AGRFORMET-D-18-01041

Status: under review

Elsevier Editorial System(tm) for
Agricultural and Forest Meteorology
Manuscript Draft

Manuscript Number:

Title: Toward a generalized predictive model of grapevine water status in Douro region from hyperspectral data

Article Type: Research Paper

Section/Category: Plant physiology, Crop Modelling, water relations including evapotranspiration, WUE, interception

Keywords: handheld spectroradiometer; machine learning; predawn leaf water potential; reflectance; vegetation indices; water deficit.

Corresponding Author: Dr. Isabel Pôças, Ph.D

Corresponding Author's Institution: Faculty of Sciences, University of Porto (FCUP); Linking Landscape, Environment, Agriculture and Food - LEAF, Instituto Superior de Agronomia, Universidade de Lisboa; Geo-Space Sciences Research Centre (CICGE)

First Author: Isabel Pôças, Ph.D

Order of Authors: Isabel Pôças, Ph.D; Renan Tosin; Igor Gonçalves; Mario Cunha, PhD

Figure 7. Receipt of submission of manuscript number 2.

Toward a generalized predictive model of grapevine water status in Douro region from hyperspectral data

Isabel Pôças[1,2,3,4,*], Renan Tosin[1], Igor Gonçalves[5], Mario Cunha[1,2,4]

[1] Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, Porto 4169-007, Portugal

[2] Geo-Space Sciences Research Centre, (CICGE), Rua do Campo Alegre, Porto 4169-007, Portugal

[3] Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisboa 1349-017, Portugal

[4] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Campus da Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, Porto 4200-465, Portugal

[5] Associação para o Desenvolvimento da Viticultura Duriense, Edifício Centro de Excelência da Vinha e do Vinho Parque de Ciência e Tecnologia de Vila Real, Régia Douro Park, Portugal

* Corresponding author: isabel.pocas@fc.up.pt

Abstract

The predawn leaf water potential (ψ_{pd}) is an eco-physiological indicator widely used for assessing vines water status and thus supporting irrigation management in several wine regions worldwide. However, the ψ_{pd} is measured in a short time period before de sunrise and the collection of a large sample of points is necessary to adequately represent a vineyard, which constitute operational constraints. In the present study, an alternative method based on hyperspectral data derived from a handheld spectroradiometer and machine learning algorithms was tested and validated for assessing grapevine water status. Two test sites in Douro wine region, integrating three grapevine cultivars, were studied for

the years of 2014, 2015, and 2017. Four machine learning regression algorithms were tested for predicting the ψ_{pd} as a continuous variable, namely Random Forest (RF), Bagging Trees (BT), Gaussian Process Regression (GPR), and Variational Heteroscedastic Gaussian Process Regression (VH-GPR). Three predicting variables, including two vegetation indices ($NRI_{554,561}$ and $WI_{900,970}$) and a time-dynamic variable based on the ψ_{pd} (ψ_{pd_0}), were applied for modelling the response variable (ψ_{pd}). Additionally, the predicted values of ψ_{pd} were aggregated into three classes representing different levels of water deficit (low, moderate, and high) and compared with the corresponding classes of ψ_{pd} observed values. A root mean square error (RMSE) and a mean absolute error (MAE) lower or equal than 0.15 MPa and 0.12 MPa, respectively, were obtained with an independent validation data set (n=72 observations) for the various algorithms. When the modelling results were assessed through classes of values, a high overall accuracy was obtained for all the algorithms (82 – 83 %), with prediction accuracy by class ranging between 79 % and 100 %. These results show a good performance of the predictive models, which considered a large variability of climatic, environmental, grape cultivars and agronomic conditions. By predicting both continuous values ψ_{pd} and classes of ψ_{pd}, the approach presented in this study allowed obtaining 2-levels of accurate information about vines water status, which can be used to feed management decisions of different types of stakeholders.

Keywords: handheld spectroradiometer; machine learning; predawn leaf water potential; reflectance; vegetation indices; water deficit.

1. Introduction

The vineyard is traditionally rainfed in most parts of the world. However, many of the wine-producing regions undergo seasonal drought coincident with the grapevine growing season, and an increase in aridity is additionally foreseen related to climate change scenarios (Chaves et al., 2010; Flexas et al., 2010; Fraga et al., 2018). Consequently, deficit irrigation strategies, in critical phenological stages, have been adopted in the last decades aiming to stabilize wine production and quality (Costa et al., 2016; Pisciotta et al., 2018; Serrano et al., 2010).

Nevertheless, water resources are increasingly under pressure as a result of the population growth and changes in lifestyle and diets, along with climate change and climate variability, with increasing occurrence of drought events (HLPE, 2015; Pereira, 2017). Therefore, the improved water management is now one of the most important challenges in

agriculture, including in viticulture, and the enhancement of irrigation management is paramount within this context.

A key aspect for the improved irrigation management relies on the accurate monitoring of the vines water status. Often, the irrigation scheduling is based on eco-physiological indicators of the vines response to water deficit, integrating both the plant and the environment influence (De Bei et al., 2011; Rodrigues et al., 2012; Williams et al., 2012). The predawn leaf water potential (ψ_{pd}) is one of these eco-physiological indicators, being considered adequate for assessing vines water status (Alves et al., 2012; Lopes et al., 2011; Williams and Araujo, 2002) and thus widely used to support irrigation management in several wine regions worldwide. The ψ_{pd} approximates the water potential of the soil before the sunrise and correspondingly the measurement period is restricted to a short daily time window, with operational constraints when the collection of a large sample of points is needed to adequately represent a vineyard. Additionally, some studies have shown a significant variability of ψ_{pd} measurements at the block and vineyard level, particularly in conditions of high water restriction, resulting in the need of an increased sampling density (e.g., Ojeda et al., 2005; Taylor et al., 2010). Therefore, it is of most importance setting alternative methods suitable for reliably and operationally assessing the plant water status of large vineyards in order to support precision viticulture production systems.

The spectral data has been increasingly used for retrieving information about vegetation, including greenness (e.g., leaf area index), dynamics (e.g., phenology), physiology (e.g., pigments content), and plant conditions (e.g., water status) (e.g., Marshall et al., 2016; Ustin et al., 2009; Verrelst et al., 2015; Zarco-Tejada et al., 2013). The wavelengths of 970 nm, 1200 nm, 1450 nm, 1930 nm, and 2500 nm, from the near infrared (NIR) and mid infrared regions of the electromagnetic spectrum, are well recognized by their strong water absorption of radiation and thus represent natural regions for assessing crop water status (Roberto et al., 2012). The thermal infrared spectra has also been considered for assessing plants water status ant water stress signs through the canopy temperatures (e.g., Bellvert et al., 2014; Buitrago et al., 2016; Neinavaz et al., 2017). However, the availability of spectral data in these domains is still limited, particularly concerning hyperspectral data. Sensing the vegetation through hyperspectral data has proven advantages compared to broadband (multispectral) data, including for detecting plant stress, as discussed by Thenkabail et al. (2012). Therefore, several studies have focused on using hyperspectral data from the visible and NIR, which are more easily available and allow analyzing stress indicators that are considered proxy for the crop water status, e.g., the xanthophyll pigment cycle or the chlorophyll fluorescence (Hernández-Clemente et al., 2011; Middleton et al., 2012; Pôças et al., 2017; Zarco-Tejada et al., 2013).

Several approaches have been considered for estimating biophysical parameters using hyperspectral data, including physically-based and statistically-based methods (e.g., Clevers et al., 2010; Lázaro-Gredilla et al., 2014). The physically-based approaches (inversion of radiative transfer models) establish a cause-effect relationship grounded on physical knowledge and are considered sound methodologies, although its application is rather challenging (Berger et al., 2018; Verrelst et al., 2015). Alternatively, the statistically-based approaches are less input demanding than physically-based models, while producing good results for retrieving several biophysical parameters, including plant water status (Pôças et al., 2017; Pôças et al., 2015; Rossini et al., 2013; Zarco-Tejada et al., 2013), and thus are more suitable for operational applications. Hence, the present study focuses on statistical approaches.

The statistically-based methodologies can be either parametric or non-parametric. The parametric models assume an explicit relationship between the spectral data (predictors) and the biophysical parameter (response variable), while non-parametric models assume a flexible approach to exploit the data and do not rely on a predefined relationship (Kuhn and Johnson, 2013). The parametric models are often based on spectral vegetation indices (VIs) that are mathematical combinations of a few selected spectral bands to describe the biophysical parameter (Jones and Vaughan, 2010). The VIs have been considered for approximating a large set of parameters related with plant water status, including plant water content, leaf water potential, and equivalent water thickness (Casas et al., 2014; González-Fernández et al., 2015b; Pôças et al., 2015; Rallo et al., 2014)

The predictive modelling techniques are often considered for the retrieval of biophysical parameters. Specifically, the machine learning algorithms, which represent non-parametric methods, are often applied for such purpose due to their potential to generate robust relationships between the predictors and the response variables with complex and non-linear patterns (e.g., crop water status). In machine learning, a training data set is used to learn the data patterns and train the model, which is further tested and assessed using a validation dataset. A large set of machine learning algorithms have been developed and applied for predicting biophysical parameters using hyperspectral data, e.g., random forest (e.g., Doktor et al., 2014; Pôças et al., 2017), bagging trees (e.g., Verrelst et al., 2015), partial least squares regression (e.g., González-Fernández et al., 2015a; Rapaport et al., 2015), Gaussian process regression (e.g., Lázaro-Gredilla et al., 2014; Verrelst et al., 2012a; Verrelst et al., 2015). The ultimate goal of the different types of machine learning algorithms is to develop a model that makes an accurate prediction. Several studies present predictive models specifically developed for estimating the crop water status based on spectral reflectance data. However, such models often rely on small sample sizes,

corresponding to a few dates of field measurements (e.g., Romero et al., 2018), or a large number of predicting variables (e.g., González-Fernández et al., 2015a; Rallo et al., 2014), and thus their applicability is limited. Additionally, an important issue for the generalization of predictive models is associated with the validation process. Often, the validation of the predictive models is performed with small size independent data sets or through cross validation procedures (e.g., Pôças et al., 2017; Rallo et al., 2014), limiting its generalized application on an operational basis.

The present study aims at develop and validate a simple generalized model for predicting vines water status based on hyperspectral data in Douro wine region. Specific goals include: (i) testing and validating multiyear modelling approaches on a diversified set of climatic, environmental, grape varieties and agronomic conditions; (ii) comparing the performance of various modelling approaches based on regression mode considering different levels of water deficit; (ii) predicting two levels of vines water status data, i.e., numeric continuous values vs. classes of values, to assist different types of stakeholders.

2. Material and Methods

2.1. Study area

The study area is located in Douro Wine Region, Northeast of Portugal (Figure 1), where vineyards are mainly built over terraces and slopes with soils mostly derived from shale. The vineyards represent one of the most important features of Douro landscape, counting for 18.3 % of the region total area. The Douro Wine Region covers around 250 000 ha, divided into three sub-regions: Baixo Corgo, Cima Corgo, and Douro Superior, distributed from the western up to eastern part of the region, and representing 32.2%, 22.0%, and 9.3% of land cover by vineyard, respectively (Fig. 1).

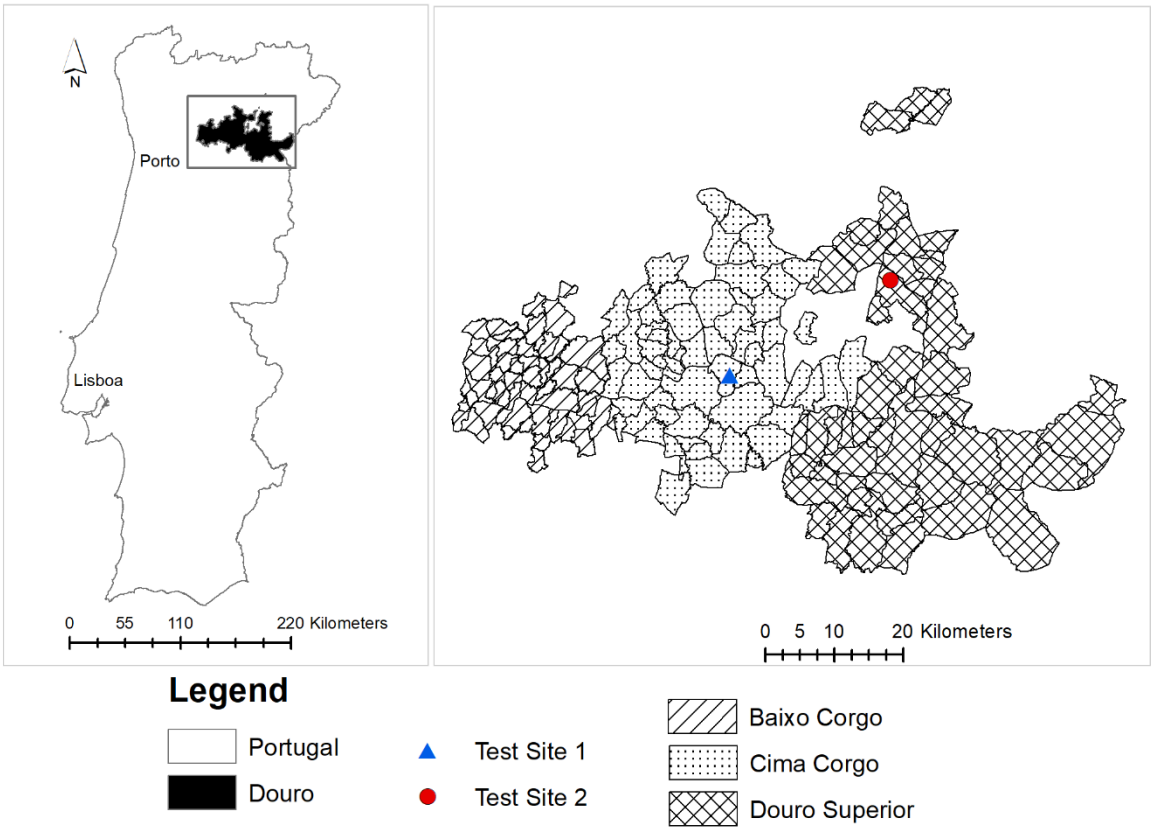


Figure 1 – Location of the study area, identifying the test site 1 and test site 2 in the Douro Wine Region, Northeast Portugal, and the sub-regions: Baixo Corgo, Cima Corgo and Douro Superior.

Overall, the region is characterized by a Mediterranean climate, with high temperatures and low precipitation values during summer period, resulting in frequent water deficit occurrence. The annual precipitation amount is 856 mm in Baixo Corgo, 658 mm in Cima Corgo and 539 mm in Douro Superior, with precipitation amounts during summer period corresponding to 10.3 %, 8.8 % and 7.4 % of the annual precipitation, respectively (INMG, 1965). The average temperature in the summer is higher in Douro Superior and Cima Corgo (24.1ºC) compared to Baixo Corgo (22.4ºC) (INMG, 1991). A more detailed characterization of the region and sub-regions climate is presented by Pôças et al. (2017).

Two test sites, integrated in commercial vineyards, were considered for the study (Figure 1): (i) Quinta dos Aciprestes (wine company Real Companhia Velha) located in Cima Corgo sub-region (Test site 1; Latitude 41.21º N; Longitude 7.43º W; 145 m a.s.l.), and (ii) Quinta do Ataíde (wine company Symington Family Estates), in Douro Superior (Test site 2; Latitude 41.25º N; Longitude 7.11º W; 161 m a.s.l.). Table 1 summarizes the vineyards overall characteristics in the test sites.

Table 1 – Overall characteristics of the vineyards in the two test sites in Douro Wine Region

Vineyard characteristics	Test site 1	Test site 2
Vines planting system	Bilateral Royat	Bilateral Royat
Plants spacing	2.2 m x 1 m	2.1 m x 1.1 m
Vines maximum height	1.5 m	1.8 m
Irrigation system	Drip irrigation	Drip irrigation
Spacing between emitters	1 m	0.5 m
Emitters discharge	2 Lh^{-1}	1.6 Lh^{-1}

Two cultivars were studied in test site 1: (i) Touriga Nacional (TN; years 2014, 2015, and 2017), and (ii) Touriga Franca (TF; years 2015 and 2017). Two plots, with two replicate areas, were sampled for TN, each plot including three irrigation treatments: non-irrigated (TN_NI), irrigation treatment 1 (TN_IT1), and irrigation treatment 2 (TN_IT2). A single plot and a single irrigation treatment (TF_IT) were considered for TF.

In test site 2, four plots, covering three cultivars, were studied in the years 2015 and 2017: two plots of TN (TN1 and TN2), one plot of TF, and one plot of Tinta Barroca (TB). Two irrigation treatments were sampled for TN and TF: (i) irrigated treatment: TN1_IT, TN2_IT, TF_IT, and (ii) non-irrigated treatment: TN1_NI, TN2_NI, TF_NI; and for TB only an irrigated treatment was sampled (TB_IT).

The irrigation amounts/dates per teste site, plot, and irrigation treatment (Table 2) were managed by each wine company, following the ψ_{pd} regular measurements and aiming to adjust for quality criteria.

Table 2 – Irrigation period and irrigation amounts (L/Plant/day) per test site and irrigation treatment.

Test site		Irrigation period			Irrigation events			Irrigation amount		
		2014	2015	2017	2014	2015	2017	2014	2015	2017
1	TN_IT1	26/7-8/8	3/7 - 25/8	19/6 - 11/8	3	7	6	32	62	92
	TN_IT2	26/7-8/8	3/7 - 25/8	19/6 - 11/8	4	9	8	36	130	124
	TF_IT	26/7-8/8	3/7 - 25/8	19/6 - 11/8	-	7	6	-	62	92
2	TN1_IT	-	30/6 - 25/8	20/6 - 22/8	-	5	10	-	95	157
	TN2_IT	-	30/6 - 25/8	20/6 - 22/8	-	5	10	-	95	157
	TF_IT	-	30/6 - 25/8	24/6 - 14/8	-	5	8	-	95	125
	TB_IT	-	30/6 - 25/8	20/6 - 14/8	-	5	9	-	95	154

Test site 1: TN_IT1– Touriga Nacional – irrigation treatment 1; TN_IT2– Touriga Nacional – irrigation treatment 2; TF_IT – Touriga Franca – irrigation treatment. Test site 2: TN1_IT –Touriga Nacional, plot 1 – irrigation treatment; TN2_IT – Touriga Nacional, plot 2 – irrigation treatment; TF_IT – Touriga Franca – irrigation treatment; TB_IT – Tinta Barroca– irrigation treatment.

2.2. Ground measurements

The study was implemented in the years 2014, 2015, and 2017, with ground data being collected in the vineyards of both test sites. In the year 2014, no data were collected in the test site 2, while in the test site 1 only data of cultivar TN were collected.

The ground measurements were done between post-flowering and harvest, roughly between June and September. The climatic conditions during the ground measurements period presented large variability among years. The year 2017, in both test sites, was

characterized by high temperatures and very low precipitation during the summer period, while 2014 was the coldest and wet year (Table 3). This climatic pattern among years can explain the irrigation amounts presented in table 2 that increase from 2014 to 2017.

Predawn leaf water potential (ψ_{pd})

The ψ_{pd} measured with a pressure chamber (Scholander et al., 1965) (PMS600, Albany, OR, USA) was used as a reference for the plants water status. A minimum of six plants per plot were sampled in each test site. A total of 21 measurement dates were considered, six in 2014 (Test site 1) and eight both in 2015 and 2017 (four in test site 1 and four in test site 2), as shown in Table 3.

Table 3 – Ground measurement dates, number of observations in each grape cultivar and climatic conditions during the study period per test site and year

Teste Site	2014	2015	2017
Test site 1			
Measurement dates	June 16	June 25	July 5
	July 10	July 16	July 20
	July 26	August 6	August 3
	August 19	September 3	August 31
	September 9		
*Cultivar**	TN (120)	TN (96), TF (40)	TN (96), TF (40)
Average Temperature	22.7ºC	23.8ºC	24.6ºC
Precipitation amount	69.8 mm	115.6mm	11.2 mm
Test site 2			
Measurement dates	-	June 26	July 4
		July 15	July 21
		August 7	August 4
		September 2	September 1
*Cultivar**	-	TN (208), TF (48) and TB (24)	TN (196), TF (36) and TB (18)
Average Temperature	21.6ºC	22.9ºC	25.0ºC
Precipitation amount	146.2mm	105.9mm	19.6mm

*TN, TF and TB represent the cultivars Touriga Nacional, Touriga Franca and Tinta Barroca, respectively, with the number of plants sampled in parenthesis. The temperature and precipitation refer to the means or sum for measurements period.

The ψ_{pd} data set presents a large intra and inter annual variability (Figure 2a), covering in each studied year, all the vineyard water deficit conditions defined by Deloire et al. (2005): (i) none up to mild water deficit (0 MPa > ψ_{pd} > −0.2 MPa); (ii) mild to moderate (−0.2 MPa > ψ_{pd} > −0.4 MPa); (iii) moderate to high (−0.4 MPa > ψ_{pd} > −0.6 MPa); and (iv) high (−0.6 MPa > ψ_{pd}). The year 2014 presented lower water deficit conditions while the higher water deficit values were recorded in the year 2017 and intermediate values were observed in 2015 (Figure 2a). Thus, the years 2014 and 2017 encompassed the largest variability of water deficit conditions within the overall data set.

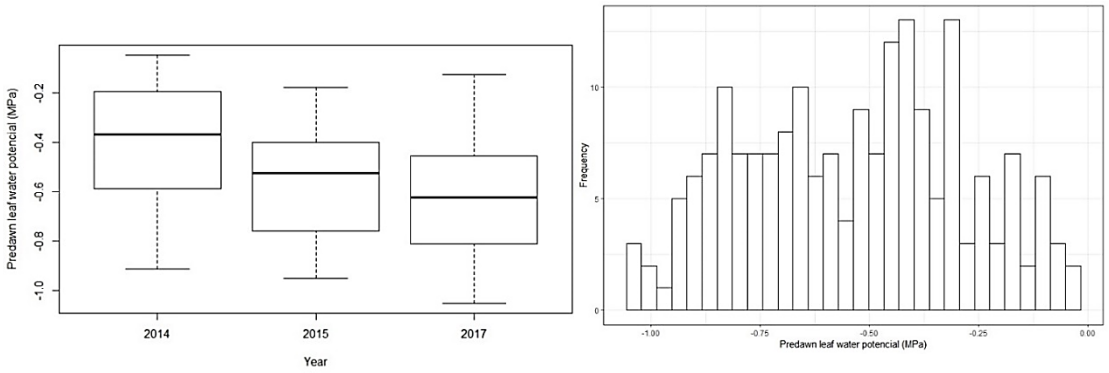


Figure 2 – Dispersion of predawn leaf water potential (MPa), represented by (a) Boxplot for each year (2014, 2015, and 2017) and the comparison between years, and (b) Histogram of the data with the class frequency distribution.

An one-way analysis of variance (ANOVA) with p-value associated to the Fischer test was performed to compare the means of ψ_{pd} between years regarding the test sites location, the irrigation treatment, and the cultivars.

Hyperspectral data

The hyperspectral data were measured in the same dates and plants of ψ_{pd} measurements (Table 3) using a portable spectroradiometer (Handheld 2, ASD Instruments, Boulder, CO, USA) maintained approximately 30 cm above canopy and directed vertically downward. The spectroradiometer records reflectance signatures between 325 nm and 1075 nm of the electromagnetic spectrum (corresponding to visible and NIR), with a wavelength interval of 1 nm. However, reflectance data below 400 nm and above 1010 nm were discarded due to noise occurrence in the spectroradiometer spectral limits.

Measurements were done between 11 h and 13 h (local time) to minimize changes in solar zenith angle, in cloud free days. Prior to canopy spectral data acquisition, a dark current correction was performed and the reflectance of a spectralon (white reference panel) was measured for directly obtaining a reflectance output. Ten repetitions per plant were collected and later averaged to minimize the effect of noise.

2.3. Data processing

The hyperspectral and ψ_{pd} data were averaged per plot and treatment aiming for minimizing the noise effects. Following this procedure, 218 observations of hyperspectral profiles were obtained and similarly for the ψ_{pd} data. An analysis of outliers was performed resulting in the identification of 28 outliers, which were removed from the study. Therefore, a total of 190 observations were considered, covering different conditions regarding year, location, cultivar, and irrigation. The 190 observations accounted for 54 observations of the year 2014, 72 of 2015, and 64 of 2017.

Hyperspectral data were further processed into spectral VIs. Two VIs tested in a previous study in Douro vineyards (Pôças et al., 2017) were computed: (i) $NRI_{554,561}$, which was optimized by Pôças et al. (2017); and (ii) $WI_{900,970}$, which follows the original formulation by Peñuelas et al. (1997). The $NRI_{554,561}$ is a normalized index (eq. 1), while the $WI_{900,970}$ is a simple ratio of bands (eq. 2).

$$NRI_{i,j} = (b_i - b_j)/(b_i + b_j) \quad \text{(eq. 1)}$$

$$SR_{i,j} = b_i/b_j \quad \text{(eq. 2)}$$

Where b_i and b_j correspond to the reflectance in the band wavelengths i and j, respectively.

A time-dynamic variable based on the ψ_{pd} was also computed by integrating, in each ψ_{pd} date to be predicted, the information of previous ψ_{pd} measurements (ψ_{pd_0}), aimed at assimilating all the available information regarding the crop water status dynamics in the post-flowering – harvest period. Additionally, this ψ_{pd_0} aimed at minimizing the spurious association between ψ_{pd} and hyperspectral data resulting from a common (downward) time trend inherent to each one of these two types of data. This time trend effect was discussed by Pôças et al. (2017). The ψ_{pd_0} was defined for each measurement point and measurement date as the ψ_{pd} value corresponding to the previous measurement.

2.4. Modelling approaches

For adequately train the predictive models, the training data set should be representative of the entire sample population, thus representing a broad range of the response variable (ψ_{pd}) conditions (Kuhn and Johnson, 2013). As shown in Figure 2a, the year 2017 recorded

high ψ_{pd} values, which were not observed in the previous years (2014 and 2015) and thus could not be learnt by the models if the data of 2014 and 2015 were used for training the models. Therefore, the data from the years 2014 and 2017 were used as training data set (118 observations) aiming to adequately learn the data patterns and train the machine learning algorithms with the larger possible variability of the ψ_{pd} conditions.

Various machine learning techniques were tested, in the regression mode, for modelling the ψ_{pd} using a training data set integrating the pairs of concurrent measurements of the ψ_{pd} and the corresponding values of the predicting variables. Four state-of-the-art machine learning algorithms were tested for predicting the ψ_{pd} as a continuous variable: (i) Random Forest, RF (Breiman, 2001); (ii) Bagging Trees, BT (Breiman, 1996); (iii) Gaussian Process Regression, GPR (Rasmussen and Williams, 2006); and (iv) Variational Heteroscedastic Gaussian Process Regression, VH-GPR (Lázaro-Gredilla et al., 2014). These machine learning algorithms have been successfully applied in previous studies related with the retrieval of vegetation biophysical parameters (e.g., Lázaro-Gredilla et al., 2014; Pôças et al., 2017; Verrelst et al., 2012a; Verrelst et al., 2015).

RF and BT algorithms are tree-based models, while GPR and VH-GPR algorithms are Bayesian statistical inference models (Kuhn and Johnson, 2013; Verrelst et al., 2015).

In the tree-based models, the data are progressively split into smaller groups, more homogenous regarding the response variable, and learning decision rules inferred from the training data are used to predict the values of the response variable (Kuhn and Johnson, 2013). In the BT, multiple versions of the predictive model are generated (ensemble technique) by making bootstrap of the training set and each model is then used for predicting a new sample (Breiman, 1996). The multiple versions of the predictive model are then averaged to obtain an aggregated model prediction. Contrarily to the BT, where all the original predictors are considered at every split of every tree, in the RF, the algorithm randomly selects predictors at each split, consequently reducing the trees correlation and potentially impacting on models performance (Breiman, 2001; Kuhn and Johnson, 2013).

In the Bayesian statistical inference models, a process of assigning and refining probability statements about unknown quantities is applied, which incorporates and updates prior knowledge and accounts for all sources of uncertainty (Link and Barker, 2010). The GPR algorithm provides a probabilistic approach for learning generic regression problems using flexible kernels functions (Verrelst et al., 2012a). An assumption of homoscedasticity is considered in GPR algorithm, i.e. assumes a constant noise power (error) in the relationship between the predicting variables and the response variable; such assumption is often not verified in biophysical retrieval studies because the noise can affect differently the

acquisition process depending on the range of the measured variable (Lázaro-Gredilla et al., 2014). The VH-GPR overcomes this issue by allowing the noise power to vary throughout the input space, i.e. allowing a nonstandard variational approximation, which is more adjusted to signal-dependent noise scenarios (Lázaro-Gredilla et al., 2014).

In the application of the various machine learning regression algorithms, the $NRI_{554,561}$, the $WI_{900,970}$ and the ψ_{pd_0} were used as predicting variables for modelling the response variable (ψ_{pd}). The selection of the predicting variables $NRI_{554,561}$ and $WI_{900,970}$ was knowledge-assisted, following previous results in the same vineyards (Pôças et al., 2017), and the ψ_{pd_0} was newly-added, allowing the model(s) to learn, in each moment, from prior grapevine water status data.

The software ARTMO (Automated Radiative Transfer Models Operator) (Verrelst et al., 2012b), through the machine learning regression algorithm toolbox (Rivera et al., 2014), was used in the implementation of the four algorithms.

2.5. Models performance assessment

In the machine learning approaches, the relationship between the predicting variables and the response variable is learnt by fitting a flexible model from the data (training data) and adjusted to minimize the prediction error of an independent data set (validation data). Therefore, a validation procedure through an independent data set, corresponding to the data set built with data collected in the year 2015, was considered.

Several goodness-of-fit indicators were used to evaluate the prediction error of the regression-mode models (RF, BT, GPR, and VH-GPR), including the root mean squared error (RMSE) and the mean absolute error as suggested by Kuhn and Johnson (2013).

When applying the ψ_{pd} for supporting irrigation scheduling, often data are aggregated into classes of vines water deficit (Deloire et al., 2005; Ojeda et al., 2001), which allows an easier data use by stakeholders, while minimizing the inherent variability of ψ_{pd}. Therefore, the predicted values of ψ_{pd} obtained by each machine learning algorithm were further aggregated into classes of water deficit and compared with the corresponding classes of ψ_{pd} observed values. The definition of ψ_{pd} classes was based on the analysis of the ψ_{pd} distribution (Figure 2b), resulting three class labels: (i) class low water deficit: 0 MPa > ψ_{pd} > −0.25 MPa; (ii) class moderate water deficit: −0.25 MPa > ψ_{pd} > −0.50 MPa; and (iii) class high water deficit: −0.50 MPa > ψ_{pd}. This class definition was further supported by the fact that, in Mediterranean regions, the irrigation (under deficit irrigation strategies) most often

starts when plants are under ψ_{pd} values below -0.5 MPa to stabilize the wine quality conditions (Lopes et al., 2011; Van Leeuwen et al., 2009).

The 72 observations of ψ_{pd} of the validation dataset were distributed into the three classes of water deficit as: (i) 5 observations in class low, (ii) 24 observations in class moderate, (iii) and 42 observations in class high. For evaluating the results of the comparison between observed and predicted classes following the data aggregation, a confusion matrix was built and the percentages of the overall model accuracy as well as the positive prediction value by class were obtained. The overall model accuracy reflects the agreement between observed and predicted classes and corresponded to the ratio of the number of cases correctly predicted (represented in the diagonal position of the confusion matrix) and the total number of cases, expressed as a percentage (Kuhn and Johnson, 2013). The overall accuracy makes no distinction about the type of errors being made. The positive prediction value by class was also computed considering the ratio of the number of cases correctly assigned in a class (true positives) and the total cases of that class, expressed as a percentage, thus taking in consideration the prevalence of the event (Kuhn and Johnson, 2013).

The validation procedures were implemented in ARTMO (Verrelst et al., 2012b), through the machine learning regression algorithm toolbox (Rivera et al., 2014).

3. Results

3.1. Comparison between study years

The average results of ψ_{pd} for each year, test site, irrigation treatment, and cultivar show the large variability of climate, agronomic, and environmental conditions (Table 5). Nevertheless, no statistically significant differences between the years 2015 and 2017 were observed for the test site 1, while the opposite was observed for the test site 2. Regarding the irrigation conditions, there were statistically significant differences between years for the non-irrigated treatment but not for the irrigated treatment. For the cultivars, statistically significant differences between years were observed only for TN. Overall, the mean ψ_{pd} of the three years (2014, 2015, and 2017) was statistically different between test sites and between irrigation treatments, but not between cultivars. Nevertheless, it is noteworthy that the number of observations for cultivars TF and TB is much lower when compared to TN. The mean values of the ψ_{pd} are significantly different between 2014 and the years 2015 and 2017, and the absolute value increases from 2014 to 2017.

Table 5. Statistical results of predawn leaf water potential (ψ_{pd}, MPa) for the different structural conditions and comparison among the three studied years

Structural Parameters	Nobs	Year (ψ_{pd}, MPa) 2014	2015	2017	Mean ψ_{pd} (MPa)	ANOVA F*
Location						
- Test site 1	127	-0.404[a]	-0.613[b]	-0.576[b]	-0.512	0.002
- Test site 2	63	Na	-0.527[a]	-0.667[b]	-0.589	0.029
ANOVA F*	---	---	*0.099*	*0.132*	*0.047*	---
Irrigation						
- No Irrigation	61	-0.386[a]	-0.585[b]	-0.833[c]	-0.605	0.000
- Irrigation	129	-0.414[a]	-0.564[a]	-0.517[a]	-0.506	0.178
ANOVA F*	---	*0.712*	*0.709*	*0.000*	*0.012*	---
Cultivars						
T. Nacional (TN)	162	-0.404[a]	-0.590[b]	-0.626[b]	-0.538	0.000
T. Franca (TF)	21	Na	-0.505[a]	-0.582[a]	-0.538	0.478
T. Barroca (TB)	7	Na	-0.511[a]	-0.541[a]	-0.524	0.889
ANOVA F*	---	---	*0.418*	*0.760*	*0.990*	---
Overall mean	190	-0.404[a]	-0.571[b]	-0.616 [b]	-0.538	0.000

ANOVA F*: is the p-value associated to the Fischer test performed in the ANOVA; means with p-value less than 0.05 is considered statistically different. Within lines, means followed by the same letter are not significantly different according to Duncan test (α=5%).

Na: No data available.

3.2. Models performance

The predicting variables used in the application of the machine learning regression algorithms were $NRI_{554,561}$, $WI_{900,970}$, and ψ_{pd_0}. The performance of the machine learning regression algorithms based on these predictors was assessed through indicators of the predicting errors for the validation data set (2015) (Table 6).

Table 6. Goodness-of-fit indicators of the machine learning regression algorithms performance

Machine learning algorithms	Calibration dataset (2014 & 2017; n=118 obs.)		Validation dataset (2015; n=72 obs.)	
	RMSE (MPa)	MAE (MPa)	RMSE (MPa)	MAE (MPa)
Gaussian Processes Regression (GPR)	0.126	0.099	0.142	0.111
VH. Gaussian Processes Regression (VH_GPR)	0.127	0.100	0.142	0.111
Bagging trees (BT)	0.127	0.100	0.146	0.119
Random Forest (RF)	0.130	0.100	0.151	0.120

The results show RMSE and MAE lower or equal than 0.151 MPa and 0.120 MPa, respectively, for the application of all the algorithms in the independent validation dataset (Table 6). These results are close to those obtained with the calibration dataset, indicating a good robustness of the models.

The predicted values of ψ_{pd} were further grouped into three classes (low, moderate, and high water deficit) and the comparison with the corresponding classes of observed ψ_{pd} is shown in a confusion matrix (Table 7). The overall accuracy (percentage), obtained by the ratio of the cases correctly assigned (in the matrix diagonal) and the total number of cases, was 82 % for GPR and RF and 83% for VH_GPR and BT. For all the algorithms, the prediction accuracy in the class low water deficit was 100% and for classes moderate and high water deficit was always higher or equal than 81% and 79%, respectively.

Through this approach it is possible to predict 2-levels of information about ψ_{pd} (values and classes of values) that can assist different types of stakeholders.

Table 7. Confusion matrices for the comparison between observed and predicted predawn leaf water potential (ψ_{pd}) in the validation dataset using different predictive models.

ψ_{pd} predicted	ψ_{pd} observed			Positive prediction by class (%)
	Low	Moderate	High	
Gaussian Process Regression				
Low	1	0	0	100.0
Moderate	2	17	2	81.0
High	2	7	40	81.6
VH. Gaussian Process Regression				
Low	2	0	0	100.0
Moderate	1	17	2	85.0
High	2	7	40	81.6
Bagging Trees				
Low	1	0	0	100.0
Moderate	2	17	1	85.0
High	2	7	41	82.0
Random Forest				
Low	2	0	0	100.0
Moderate	1	15	1	88.2
High	2	9	41	78.8

Predawn leaf water potential (ψ_{pd}) classes: low water deficit, MPa > ψ_{pd} > −0.25 MPa; moderate water deficit, −0.25 MPa > ψ_{pd} > −0.50 MPa; high water deficit, −0.50 MPa > ψ_{pd}.

4. Discussion

The predicting variables $NRI_{554,561}$, and $WI_{900,970}$ were also used for assessing grapevine water status in Douro wine region in the years 2014 and 2015 (Pôças et al., 2017). However, in this latter study, the variability regarding water deficit conditions was smaller, as 2014 and 2015 presented a shorter range of water deficit, without ψ_{pd} values exceeding - 1 MPa, when compared to the year 2017, as shown in Figure 2a. The $NRI_{554,561}$ integrates spectral information of the green domain (520 – 570 nm), similarly to the Photochemical Reflectance Index (PRI; Gamon et al., 1992) that combines wavelengths of 531 nm and 570 nm and is often used for assessing crop water status (Hernández-Clemente et al., 2011; Pôças et al., 2015; Zarco-Tejada et al., 2013). The green spectral domain has been used as a proxy of stress conditions, including water deficit conditions, due to its ability for detecting subtle changes in the xanthophyll cycle pigment activity resulting from stress conditions (Gamon et al., 1992; Middleton et al., 2012). The $WI_{900,970}$ integrates wavelengths of the near infrared domain, specifically including the wavelength 970 nm corresponding to a water absorption peak (Peñuelas et al., 1993; Peñuelas et al., 1997; Roberto et al., 2012). A novelty was introduced in the model by integrating the ψ_{pd_0}, which accounts for the natural dynamics of ψ_{pd} along the grapevine cycle and thus strengthening the learning process by the machine learning algorithms.

The results of the goodness-of-fit indicators of the various algorithms for the validation dataset (Table 6) were similar or better than the results of previous studies. The RMSE results were similar to those obtained for assessing water status in grapevine (Rapaport et al., 2015) and significantly better than those obtained for olive orchard (Rallo et al., 2014) using a partial least squares regression algorithm. Nevertheless, in these studies, the spectral range included from the visible up to the shortwave infrared (SWIR) domains, thus larger than the range considered in the present study and encompassing several SWIR wavelengths typical for water absorption (Jones and Vaughan, 2010). The RMSE results were also similar to those obtained by Pôças et al. (2017) when assessing grapevines water status in Douro region in 2014 and 2015 when conditions of lower variability of ψ_{pd} were observed (Figure 2a), which suggests a robust performance of the models in the present study. Additionally, it is important to highlight that the results of RMSE and MAE were obtained for an independent data set with 72 observations, instead of considering a cross validation procedure as in most of the previous studies (Pôças et al., 2017; Rallo et al., 2014).

The good performance of the four algorithms is likely related with the high adaptability of machine learning regression methods, with a non-parametric structure, to cope with nonlinear relations of biophysical parameters and complex environment/agronomic conditions (e.g., Doktor et al., 2014; Im et al., 2009; Lázaro-Gredilla et al., 2014).

The models based on Bayesian statistical inference, the GPR and the VH_GPR algorithms, showed slightly better performance than the tree-based models, with lower RMSE and MAE (Table 6). Also, other authors have reported good performance of GPR and VH_GPR for the retrieval of biophysical parameters, e.g., leaf area index and leaf chlorophyll content (Lázaro-Gredilla et al., 2014; Verrelst et al., 2012a; Verrelst et al., 2015). Such results are likely related with the ability of Bayesian inference to provide more realistic models, when compared to classical machine learning techniques (e.g., RF), by capturing uncertainties related to parameters and models (Zhou et al., 2018).

The assessment of the predicted values of ψ_{pd} through classes provided good results regarding the overall accuracy (equal or higher than 82%) and the predictive positive value by class showed good performance for all classes of water deficit (equal or higher than 79%; Table 7). This ψ_{pd} data aggregation by classes facilitates the application of the methodology proposed in operational contexts, by providing a more user-friendly output, while minimizing the inherent variability of ψ_{pd}.

5. Conclusions

Predictive models applied in the regression modes were used for retrieving crop water status in vineyards of Douro wine region. A large set of climatic, environmental, and agronomic conditions were sampled to test model's robustness. Two years of data, representing the largest variability of ψ_{pd}, were used for calibrating the models, and a third year was used for an external validation of the models generated.

Overall, the models showed a good performance, including for the class of high water deficit, which represents the class of ψ_{pd} values that are considered for irrigation management decisions. When the results were further aggregated according to water deficit classes, continuous ψ_{pd} and classes of ψ_{pd} values were predicted. These two types of information can be used to feed management decisions of different types of stakeholders, including wine producers, water managers, and scientists.

Acknowledgments

Isabel Pôças acknowledges FCT by the postdoctoral grant SFRH/BPD/79767/2011 and the Post-Doctoral grant of the project ENGAGE-SKA POCI-01-0145-FEDER-022217, co-funded by FEDER through COMPETE (POCI-01-0145-FEDER-022217). The authors also thank the wine company Real Companhia Velha (and its Coordinator for Viticulture Rui Soares) and the wine company Symington Family Estates (and its R&D Viticulture Manager Fernando Alves) for the facilities provided for fieldwork, as well as the Associação para o Desenvolvimento da Viticultura Duriense (ADVID) for funding part of the field-work missions. This study was implemented under a cooperation protocol integrating Faculty of Sciences, University of Porto, ADVID, Real Companhia Velha, and Symington Family Estates.

References

Alves, F. et al., 2012. Influence of climate and deficit irrigation on grapevine physiology, yield and quality attributes, of the cv. Touriga Nacional at Douro Region. , IXe International erroirs Congress, Dijon - Reims, France, pp. 20 - 24.

Bellvert, J., Zarco-Tejada, P.J., Girona, J. and Fereres, E., 2014. Mapping crop water stress index in a 'Pinot-noir' vineyard: comparing ground measurements with thermal remote sensing imagery from an unmanned aerial vehicle. Precision Agric, 15(4): 361-376.

Berger, K. et al., 2018. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. Remote Sensing, 10(1): 85.

Breiman, L., 1996. Bagging predictors. Machine Learning, 24(2): 123-140.

Breiman, L., 2001. Random Forests. Machine Learning, 45: 5-32.

Buitrago, M.F., Groen, T.A., Hecker, C.A. and Skidmore, A.K., 2016. Changes in thermal infrared spectra of plants caused by temperature and water stress. ISPRS Journal of Photogrammetry and Remote Sensing, 111(Supplement C): 22-31.

Casas, A., Riaño, D., Ustin, S.L., Dennison, P. and Salas, J., 2014. Estimation of water-related biochemical and biophysical vegetation properties using multitemporal airborne hyperspectral data and its comparison to MODIS spectral response. Remote Sensing of Environment, 148: 28-41.

Chaves, M.M. et al., 2010. Grapevine under deficit irrigation: hints from physiological and molecular data. Annals of Botany, 105(5): 661-676.

Clevers, J.G.P.W., Kooistra, L. and Schaepman, M.E., 2010. Estimating canopy water content using hyperspectral remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 12(2): 119-125.

Costa, J.M. et al., 2016. Modern viticulture in southern Europe: Vulnerabilities and strategies for adaptation to water scarcity. Agricultural Water Management, 164: 5-18.

De Bei, R. et al., 2011. Non-destructive measurement of grapevine water potential using near infrared spectroscopy. Australian Journal of Grape and Wine Research, 17(1): 62-71.

Deloire, A. et al., 2005. Influence de l´état hydrique de la vigne sur le style de vin. Progrès agricole Viticole, 122(21): 455-462.

Doktor, D., Lausch, A., Spengler, D. and Thurner, M., 2014. Extraction of Plant Physiological Status from Hyperspectral Signatures Using Machine Learning Methods. Remote sensing, 6(12): 12247.

Flexas, J. et al., 2010. Improving water use efficiency in grapevines: potential physiological targets for biotechnological improvement. Australian Journal of Grape and Wine Research, 16: 106-121.

Fraga, H., García de Cortázar Atauri, I. and Santos, J.A., 2018. Viticultural irrigation demands under climate change scenarios in Portugal. Agricultural Water Management, 196: 66-74.

Gamon, J.A., Penuelas, J. and Field, C.B., 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment, 41: 35-44.

González-Fernández, A.B., Rodríguez-Pérez, J.R., Marabel, M. and Álvarez-Taboada, F., 2015a. Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. Scientia Horticulturae, 188: 15-22.

González-Fernández, A.B., Rodríguez-Pérez, J.R., Marcelo, V. and Valenciano, J.B., 2015b. Using field spectrometry and a plant probe accessory to determine leaf water content in commercial vineyards. Agricultural Water Management, 156: 43-50.

Hernández-Clemente, R., Navarro-Cerrillo, R.M., Suárez, L., Morales, F. and Zarco-Tejada, P.J., 2011. Assessing structural effects on PRI for stress detection in conifer forests. Remote Sensing of Environment, 115(9): 2360-2375.

HLPE, 2015. Water for food security and nutrition. A Report by the High Level Panel of Experts on Food

Security and Nutrition of the Committee on World Food Security, Rome.

INMG, 1965. O Clima de Portugal. Fascículo XIII. Normais climatológicas do Continente, Açores e Madeira correspondentes a 1931 – 1960. INMG, Lisboa.

INMG, 1991. O Clima de Portugal. Fascículo XLIX. Normais climatológicas da região de Trás-os-Montes e Alto Douro e Beira Interior correspondentes a 1951-1980. INMG, Lisboa.

Jones, H.G. and Vaughan, R.A., 2010. Remote sensing of vegetation. Principles, techniques, and applications. Oxford University Press, Oxford, 353 pp.

Kuhn, M. and Johnson, K., 2013. Applied predictive modeling. Springer Science+Business Media, New York, 600 pp.

Lázaro-Gredilla , M., Titsias, M.K., Verrelst, J. and G., C.-V., 2014. Retrieval of Biophysical Parameters With Heteroscedastic Gaussian Processes. IEEE Geoscience and Remote Sensing Letters, 11(4): 838-842.

Link, W.A. and Barker, R.J., 2010. Bayesian Inference with ecological applications. Academic Press, London.

Lopes, C.M. et al., 2011. Combining cover cropping with deficit irrigation in a Mediterranean low vigor vineyard. Scientia Horticulturae, 129(4): 603-612.

Marshall, M., Thenkabail, P., Biggs, T. and Post, K., 2016. Hyperspectral narrowband and multispectral broadband indices for remote sensing of crop evapotranspiration and its components (transpiration and soil evaporation). Agricultural and Forest Meteorology, 218-219: 122-134.

Middleton, E.M., Huemmrich, K.F., Cheng, Y.-B. and Margolis, H.A., 2012. Spectral bioindicators of photosinthetic efficiency and vegetation stress. In: P.S. Thenkabail, J.G. Lyon and A. Huete (Editors), Hyperspectral remote sensing of vegetation. CRC Press. Taylor & Francis Group, Boca Raton, pp. 265-288.

Neinavaz, E., Skidmore, A.K., Darvishzadeh, R. and Groen, T.A., 2017. Retrieving vegetation canopy water content from hyperspectral thermal measurements. Agricultural and Forest Meteorology, 247(Supplement C): 365-375.

Ojeda, H. et al., 2005. Precision viticulture and water status II: Quantitative and qualitative performance of different within field zones, defined from water potential mapping.

Groupe d'Etude des Systemes de COnduite de la vigne (GESCO), Geisenheim, pp. 741-748.

Ojeda, H., Deloire, A. and Carbonneau, A., 2001. Influence of water deficit on grape berry growth. Vitis, 40(3): 141-145.

Peñuelas, J., Filella, I., Biel, C., Serrano, L. and Savé, T., 1993. The reflectance at the 950-970 nm as an indicator of plant water status. International Journal of Remote Sensing, 14(10): 1887-1905.

Peñuelas, J., Pinol, J., Ogaya, R. and Filella, I., 1997. Estimation of plant water concentration by the reflectance water index WI (R900/R970). International Journal of Remote Sensing, 18(13): 2869-2875.

Pereira, L.S., 2017. Water, Agriculture and Food: Challenges and Issues. Water Resources Management, 31(10): 2985-2999.

Pisciotta, A., Di Lorenzo, R., Santalucia, G. and Barbagallo, M.G., 2018. Response of grapevine (Cabernet Sauvignon cv) to above ground and subsurface drip irrigation under arid conditions. Agricultural Water Management, 197: 122-131.

Pôças, I. et al., 2017. Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region. International Journal of Applied Earth Observations and Geoinformation, 58: 177-190.

Pôças, I. et al., 2015. Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices. Remote Sensing, 7(12): 16460–16479.

Rallo, G., Minacapilli, M., Ciraolo, G. and Provenzano, G., 2014. Detecting crop water status in mature olive groves using vegetation spectral measurements. Biosystems Engineering, 128(0): 52-68.

Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A. and Rachmilevitch, S., 2015. Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. ISPRS Journal of Photogrammetry and Remote Sensing, 109: 88-97.

Rasmussen, C.E. and Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning. MIT Press, 266 pp.

Rivera, J.C., Verrelst, J., Munoz-Mari, J., Moreno, J. and Camps-Valls, G., 2014. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(4): 1249-1259.

Roberto, C., Lorenzo, B., Michele, M., Micol, R. and Cinzia, P., 2012. Optical remote sensing of vegetation water content. In: P.S. Thenkabail, J.G. Lyon and A. Huete (Editors), Hyperspectral remote sensing of vegetation. CRC Press. Taylor & Francis Group, Boca Raton, pp. 227-244.

Rodrigues, P. et al., 2012. Influence of soil water content and atmospheric conditions on leaf water potential in cv. “Touriga Nacional” deep-rooted vineyards. Irrigation Science, 30(5): 407-417.

Romero, M., Luo, Y., Su, B. and Fuentes, S., 2018. Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. Computers and Electronics in Agriculture, 147: 109-117.

Rossini, M. et al., 2013. Assessing canopy PRI from airborne imagery to map water stress in maize. ISPRS Journal of Photogrammetry and Remote Sensing, 86(0): 168-177.

Scholander, P.F., Hammel, H.T., Bradstreet, E.D. and Hemmingsen, E.A., 1965. Sap pressure in vascular plants. Science 148: 339-346.

Serrano, L., González-Flor, C. and Gorchs, G., 2010. Assessing vineyard water status using the reflectance based Water Index. Agriculture, Ecosystems & Environment, 139(4): 490-499.

Taylor, J.A., Acevedo-Opazo, C., Ojeda, H. and Tisseyre, B., 2010. Identification and significance of sources of spatial variation in grapevine water status. Australian Journal of Grape and Wine Research, 16(1): 218-226.

Thenkabail, P.S., Lyon, J.G. and Huete, A., 2012. Advances in hyperspectral remote sensing of vegetation and agricultural croplands. In: P.S. Thenkabail, J.G. Lyon and A. Huete (Editors), Hyperspectral remote sensing of vegetation. CRC Press. Taylor & Francis Group, Boca Raton, pp. 4-35.

Ustin, S.L. et al., 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment, 113: S67-S77.

Van Leeuwen, C., Tregoat, O. and Choné, X., 2009. Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes? Sonoma County Wine Library. Journal International des Sciences de la Vigne et du Vin.

Verrelst, J. et al., 2012a. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. Remote Sensing of Environment, 118: 127-139.

Verrelst, J. et al., 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison. ISPRS Journal of Photogrammetry and Remote Sensing, 108: 260-272.

Verrelst, J., Romijn, E. and Kooistra, L., 2012b. Mapping Vegetation Density in a Heterogeneous River Floodplain Ecosystem Using Pointable CHRIS/PROBA Data, 4, 2866-2889 pp.

Williams, L.E. and Araujo, F.J., 2002. Correlations among predawn leaf, midday leaf, and midday stem water potential and their correlations with other measures of soil and plant water status in Vitis vinifera. Journal of American Society for Horticultural Science, 127(3): 448-454.

Williams, L.E., Baeza, P. and Vaughn, P., 2012. Midday measurements of leaf water potential and stomatal conductance are highly correlated with daily water use of Thompson Seedless grapevines. Irrigation Science, 30(3): 201-212.

Zarco-Tejada, P.J. et al., 2013. A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. Remote Sensing of Environment, 138(0): 38-50.

3.3. Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region

Journal: Journal of Applied Remote Sensing

Date of submission: 12/11/2018

Manuscript number: JARS 180901

Status: under review

Manuscript #	180901
Journal	Journal of Applied Remote Sensing
Current Revision #	0
Submission Date	2018-11-12 05:37:15
Current Stage	Submission Check - Editorial Office
Title	Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region
Manuscript Type	Regular Paper
Special Section	N/A
Category	Remote Sensing for Engineering and Science Applications (45)
Corresponding Author	Mário Cunha (Faculdade de Ciências da Universidade do Porto)
Contributing Authors	Renan Tosin , Isabel Pôças , Igor Gonçalves , Mário Cunha (corr-auth)
Abstract	Hyperspectral data collected through a handheld spectroradiometer (400 – 1010 nm) were tested for assessing the grapevine predawn leaf water potential (ψpd) in two test ... View full abstract
Associate Editor	Not Assigned
Index Terms	handheld spectroradiometer, ordinal logistic regression, vegetation indices, vineyard, grapevine water status
Multimedia Files	Manuscript has NO multimedia files
Funding Information	Funding Summary
Prior Publication	Are any of the major results (data/figures/etc.) in your manuscript previously published or currently under consideration for publication? (Check one below): No. The work I am submitting has never been published before, and it is not currently under consideration for publication in another journal.

Figure 8. Receipt of submission of manuscript number 3.

Estimation of grapevine predawn leaf water potential based on hyperspectral reflectance data in Douro wine region

Renan Tosin[1], Isabel Pôças[1,2,3,4], Igor Gonçalves[5], Mario Cunha[1,2,4*]

[1] Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, Porto 4169-007, Portugal

[2] Geo-Space Sciences Research Centre, (CICGE), Rua do Campo Alegre, Porto 4169-007, Portugal

[3] Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisboa 1349-017, Portugal

[4] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Campus da Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, Porto 4200-465, Portugal

[5] Associação para o Desenvolvimento da Viticultura Duriense, Edifício Centro de Excelência da Vinha e do Vinho Parque de Ciência e Tecnologia de Vila Real, Régia Douro Park, Portugal

[*] Corresponding author: mccunha@fc.up.pt

Abstract

Hyperspectral data collected through a handheld spectroradiometer (400 – 1010 nm) were tested for assessing the grapevine predawn leaf water potential (ψ_{pd}) in two test sites of Douro Wine region. The study was implemented in 2017, a year with climatic conditions prone to water shortage. Three grapevine cultivars, Touriga Nacional, Touriga Franca and Tinta Barroca, were sampled both in rainfed and irrigate vineyards, totalizing 325 plants assessed. An ordinal logistic regression model was applied to the hyperspectral data to estimate the ψ_{pd}. A large set of vegetation indices computed with the hyperspectral data and optimized for the ψ_{pd} values, as well as structural variables, were used as predictors in the

model. From 631 possible predictors, four were selected through a forward stepwise procedure and the Wald statistics: irrigation treatment, test site, $ARI_{opt\_656,647}$ and $NRI_{711,700}$. 70 % of the dataset was randomly selected for model calibration and 30 % for model validation. The model accuracy with the validation dataset was 73.2 %, with the class of ψ_{pd} for high-water deficit presenting a positive prediction value of 79.3 %. The accuracy and operability of the predictive model indicates good perspectives for its use in the grapevine water status monitoring and to support irrigation scheduling.

Keywords: handheld spectroradiometer, ordinal logistic regression, vegetation indices, vineyard, grapevine water status.

1. Introduction

The severe hydric water stress affects the quantity and quality of wine grapes. Therefore, in regions where precipitation is scarce and concentrated in a short period of the year, as in the Mediterranean regions, irrigation has been increasingly considered to regulate the grapevines yield and quality (e.g. Chaves, Zarrouk et al. 2010). However, in the context of foreseen warming and dry climate scenarios and the increasing competition for water among different economic sectors, a correct irrigation management is essential to ensure the sustainability of Mediterranean irrigated areas (Medrano, Tomás et al. 2015, Cunha and Richter 2016).

The Douro wine region is one of the most arid regions of Europe where a strong water deficit occurs in summer as a consequence of the low soil water content, associated to the low annual rainfall and high gradients of the water vapour pressure between the leaves and the air (Jones and Alves 2012, Alves, J et al. 2013, Prata-Sena, Castro-Carvalho et al. 2018).

The grapevines irrigation scheduling is often based on ecophysiological measures of vines water status. One of the most widely used measure is referred to the predawn leaf water potential (ψ_{pd}) using a Scholander chamber (Scholander, Bradstreet et al. 1965). Despite being a very reliable technique (Moutinho-Pereira, Magalhães et al. 2007, Alves, Costa et al. 2012, Merli, Gatti et al. 2015), it is a destructive method (Rodríguez-Pérez, Ordóñez et al. 2018) and depends on the collection of a large set of measurement points to get a very accurate assessment of the target area due to the variability in soil conditions (Oumar and Mutanga 2010). Thus, efforts have been made to find alternative methods

capable of providing accurate information about vines water status, while being easy-to use and non-destructive.

The contribution of remote sensing to improve water management has increased in the last years. Spectral reflectance obtained through proximity sensors (e.g. handheld spectroradiometers), cameras mounted in drones or satellite imagery has been widely used to estimate crop biophysical parameters (Blackburn 2007, Zarco-Tejada, González-Dugo et al. 2013), including for estimating and monitoring water status in vineyards (Pôças, Rodrigues et al. 2015, Pôças, Gonçalves et al. 2017, Rodríguez-Pérez, Ordóñez et al. 2018). Different zones of the electromagnetic spectrum have been studied for the monitoring of plant water status, including the near and mid infrared, which present wavelengths of strong water absorption by the radiation, and the shortwave infrared due to the relation between canopy temperature and crop water status (Clevers, Kooistra et al. 2010, Bellvert, Zarco-Tejada et al. 2014). Additionally, the spectral zones of visible and near infrared (NIR) are potentially useful to estimate crop water status (Suárez, Zarco-Tejada et al. 2008, De Bei, Cozzolino et al. 2011). However, the spectral data from the visible and NIR domains are more easily accessible from commonly available handheld spectroradiometers, as well as from satellite sensors and unmanned aerial vehicles (Zarco-Tejada, González-Dugo et al. 2013).

A particular focus has been given to the use of hyperspectral data, which are characterized from numerous narrow bands continuously distributed across the electromagnetic spectrum, for assessing crop water status. These hyperspectral data are sensitive to subtle variations in the energy reflected and thus have great potential for detecting differences between surface characteristics (Blackburn 2007, Jones and Vaughan 2010, Mariotto, Thenkabail et al. 2013). Nevertheless, the large amount of data generated from hyperspectral sensors can result in redundancy of the information captured (Blackburn 2007, Wu, Niu et al. 2008, Caicedo, Verrelst et al. 2014, Rivera, Verrelst et al. 2014, Feng, Itoh et al. 2017). Thus, the adequate data processing, coping with dimensionality issues, and modelling tools are required for its efficient use.

The hyperspectral data can be combined into vegetation indices, which can be specifically optimized for vines water status, and thus only using a small portion of the spectrum (Suárez, Zarco-Tejada et al. 2008, Zarco-Tejada, González-Dugo et al. 2013, Pôças, Rodrigues et al. 2015).

Also, the techniques of machine learning are often used to cope with the high dimensionality of hyperspectral datasets. Diverse studies have applied non-parametric regression models to estimate the water status in grapevines, (Pôças, Gonçalves et al.

2017, Rodríguez-Pérez, Ordóñez et al. 2018). The machine learning classification methods have also been applied to hyperspectral data for estimating biophysical and biochemical crop parameters (Im, Jensen et al. 2009). One of such classification methods is the ordinal logistic regression (OLR), which is used to explain a ranking variable (Harrell 2015) and has been employed in many environmental studies (Brant 1990, Rutherford, Guisan et al. 2007, Coppock 2011, Notario del Pino and Ruiz-Gallardo 2015). The OLR algorithm has been used for modelling the relationship between an ordinal response variable and one or several continuous independent variables, while considering the inherent ordering of the response variable, thus making full use of the ordinal information (Kleinbaum and Klein 2010). Often, the OLR is fitted through a proportional-odds logit model (McCullagh 1980), applied to obtain an ordinal response (Verwaeren, Waegeman et al. 2012). The proportional-odds logit model assumes that identical feature variables might result in different values for the underlying response variable and therefore the model contains a deterministic component and an error term, which is assumed to follow a logistic distribution (Verwaeren, Waegeman et al. 2012).

We argue that machine learning classification methods based on hyperspectral data could be an alternative to estimate grapevine ψ_{pd}, resulting classes of water status. In fact, although the ψ_{pd} is recorded as a continuous variable, farmers often use classes of ψ_{pd} to characterize the vines stress conditions (Deloire, Ojeda et al. (2005). Thus, the main goal of this work is modelling the water status in grapevines through a classification predictive regression model based on hyperspectral data. Specific goals include testing and validating the model in two different zones of Douro Wine region and considering three cultivars growing in to irrigation regimes.

2. Material and methods

2.1. Study area

This study was conducted in the Douro Wine Region, Northeast of Portugal (Figure 1), where the vineyards dominate the landscape and are established mainly over terraces and slopes with shale-derived soils. The region is divided into three sub-regions: Baixo Corgo, Cima Corgo, and Douro Superior, distributed from the western up to eastern part of the region (Figure 1), all with rigorous climate conditions.

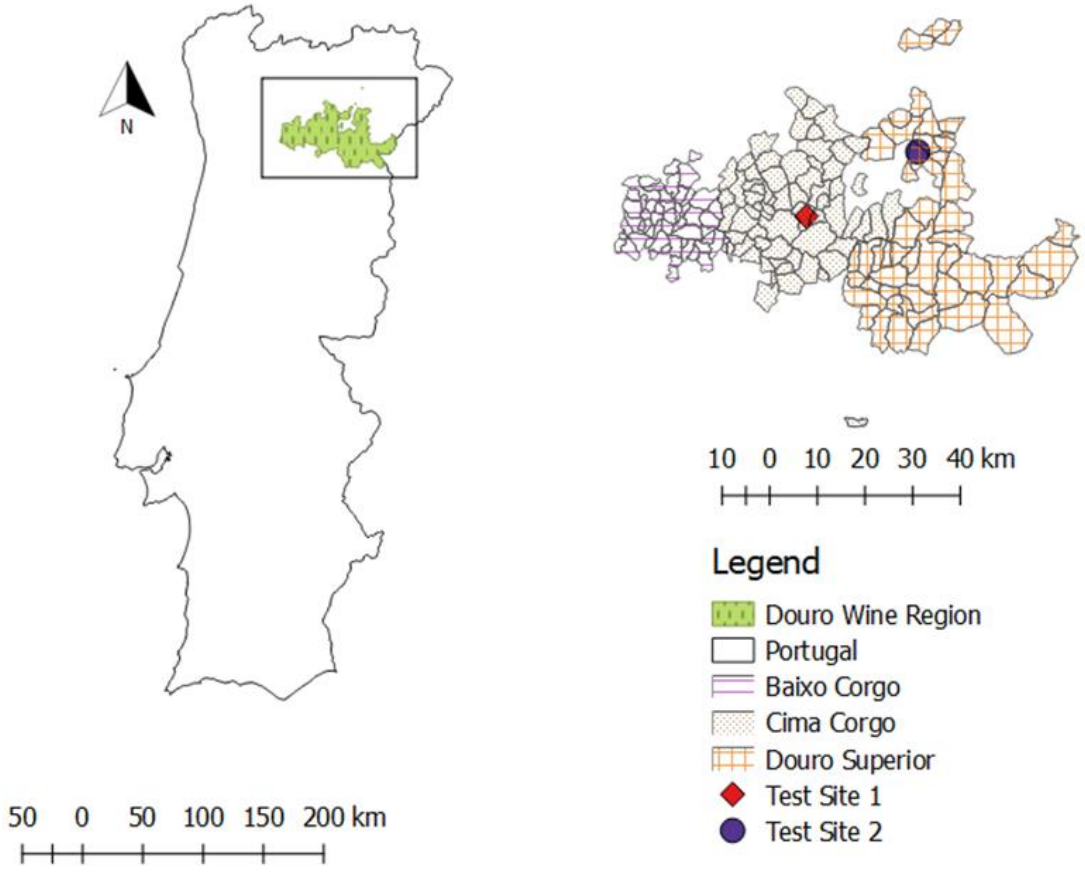


Figure 1. Location of the study area, identifying the test sites 1 and 2 in the Douro Wine Region, Northeast Portugal, and respective the sub-regions: Baixo Corgo, Cima Corgo and Douro Superior.

The region presents a Mediterranean climate, with high average temperatures during the summer period, ranging between 22.4ºC in Baixo Corgo and 24.1ºC in Douro Superior and Cima Corgo (INMG 1991). In Baixo Corgo the annual precipitation is 856 mm, while in Cima Corgo is 658 mm and in Douro Superior is 539 m, with summer precipitation representing 10.3 %, 8.8 % and 7.4 % of the annual precipitation, respectively. A detailed characterization of the region and sub-regions climate is presented by Pôças, Gonçalves et al. (2017). The Figure 2 compares the climate characterization of the Douro Wine Region with both test sites in the year of 2017. The Figure 2 compares the climate characterization of the Douro Wine Region in both test sites in the year of 2017.

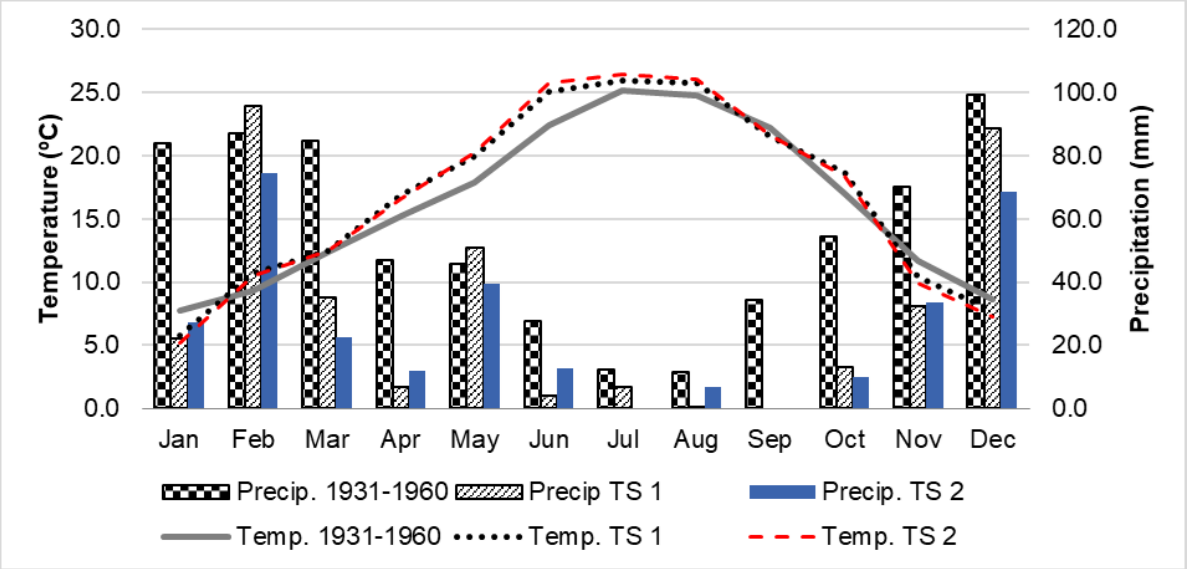


Figure 2. Temperature and precipitation characterization of the Douro wine region for the reference period 1931-1960 (Ferreira 1965) and comparison with the temperature and precipitation records in 2017 during study period in the test site 1 (TS 1) and test site 2 (TS 2).

Two commercial vineyards were considered for the study (Figure 1): (i) Test site 1 (TS1) – Quinta dos Aciprestes (wine company Real Companhia Velha) located in Cima Corgo sub-region (Latitude 41.21º N; Longitude 7.43º W; 145 m a.s.l.), and (ii) Test site 2 (TS2) – Quinta do Ataíde (wine company Symington Family Estates), in Douro Superior (Latitude 41.25º N; Longitude 7.11º W; 161 m a.s.l.).

In test site 1 the cultivars studied were: (i) Touriga Nacional (TN) – two plots, with two replicate areas, including three irrigation treatments: non-irrigated (TN_NI), irrigation treatment 1 (TN_IT1), and irrigation treatment 2 (TN_IT2), and (ii) Touriga Franca (TF) – a single plot and a single treatment (TF_IT). Three cultivars were studied in test site 2: (i) TN – two plots (TN1 and TN2) with two irrigation treatments: irrigated (TN1_IT, TN2_IT) and non-irrigated (TN1_NI, TN2_NI), (ii) TF – one plot with an irrigated treatment (TF_IT) and a non-irrigated treatment (TF_NI), and (iii) Tinta Barroca (TB) – one plot with an irrigated treatment (TB_IT).

The vines pruning system is Bilateral *Royat* in both test sites with planting spacing and vines maximum height respectively of 2.2 m × 1 m and 1.5 m in test site 1 and 2.1 m × 1.1 m and 1.8 m in test site 2.

The irrigation was managed by each wine company, according to the ψ_{pd} regular measurements and aiming to adjust for quality criteria. Table 1 summarizes the irrigation dates and amounts in each test site.

Table 1. Irrigation dates and irrigation amounts (L/Plant/day) per test site and irrigation treatment. Test site 1: TN_IT1 – Touriga Nacional – irrigation treatment 1; TN_IT2 – Touriga Nacional – irrigation treatment 2; TF_IT – Touriga Franca – irrigation treatment. Test site 2: TN1_IT – Touriga Nacional, plot 1 – irrigation treatment; TN2_IT – Touriga Nacional, plot 2 – irrigation treatment; TF_IT – Touriga Franca – irrigation treatment; TB_IT – Tinta Barroca– irrigation treatment.

Date	Test site 1*			Date	Test site 2*			
	TN_IT1	TN_IT2	TF_IT		TN1_IT	TN2_IT	TF_IT	TB_IT
19 June	16	16	16	20 June	9.6	9.6		16
23 June	0	16	0	24 June	16	16	16	16
29 June	16	16	16	1 July	19.2	19.2	16	19.2
13 July	16	16	16	7 July	19.2	19.2	16	19.2
21 July	16	16	16	15 July	16	16	16	19.2
28 July	16	16	16	25 July	16	16	16	16
4 August	0	16	0	31 July	16	16	16	16
11 August	12	12	12	7 August	16	16	16	16
				14 August	16	16	13	16
				22 August	13	13		
Total	92	124	92		128	128	112	137.6

*Drip irrigation system with emitters discharge: Test site 1: 2 Lh-1; Test site 2: 1.6 Lh-1 with Spacing between emitters of 1 m (test site 1) and 0.5 m (test site 2).

2.2. Ground measurements

The study was conducted in 2017 between the post-flowering (end of June) and the harvest (early September). During this period, the average temperature was 24.6 ºC in the test site 1 and 25 ºC in the test site 2 and the precipitation in test site 1 was 11.2 mm while for test site 2 was 19.6 mm (Fig. 2).

Ground measurements of ψ_{pd} data and spectral reflectance data were collected in four dates in both test sites: test site 1 (July 5, July 20, August 3, and August 31) and test site 2 (July 4, July 21, August 4, and September 1). A minimum of six plants per irrigation treatment and plot were sampled in each test site for ground measurements, resulting 325 observations (grapevines), 135 in test site 1 and 190 in test site 2 (Table 2).

Table 2. Number of observations (gapevines) per test site, cultivar and irrigation conditions.

Test site	Cultivar	Irrigated	Non Irrigated	Total
Test site 1	Touriga Nacional	64	31	95
	Touriga Franca	40	0	40
Total test site 1		104	31	135
Test site 2	Touriga Nacional	68	68	136
	Touriga Franca	18	18	36
	Tinta Barroca	18	0	18
Total test site 2		104	86	190
Total		208	117	325

A pressure chamber (Scholander, Bradstreet et al. 1965) (PMS600, Albany, OR, USA) was used for measuring the $\psi_{pd.}$ A large variability both between test sites and within each test site was recorded in the ψ_{pd} data set (Figure 3).

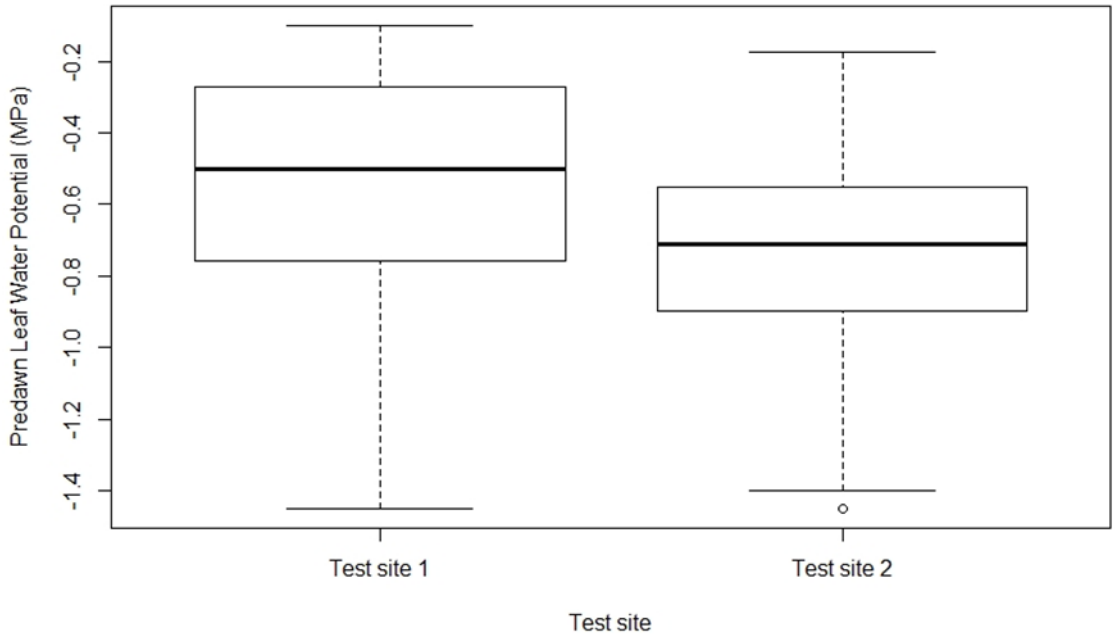


Figure 3. Dispersion of predawn leaf water potential (MPa), represented by a Boxplot for the test site.

The hyperspectral data were measured using a portable spectroradiometer (Handheld 2, ASD Instruments, Boulder, CO, USA) maintained approximately 30 cm above canopy and directed vertically downward. The spectroradiometer records reflectance signatures between 325 nm and 1075 nm of the electromagnetic spectrum (corresponding to visible and NIR), with a wavelength interval of 1 nm. However, only reflectance data between 400 nm to 1010 nm were considered due to noise occurrence outside of these spectral limits. The measurements were done in cloud free days between 11 h to 14 h (local time) to minimize changes in solar zenith angle. Prior to canopy spectral data acquisition, a dark current correction was performed and the reflectance of a spectralon (white reference panel) was measured for directly obtaining a reflectance output. Ten repetitions per plant were collected and later averaged to minimize the effect of noise.

2.3. Data processing

The hyperspectral and ψ_{pd} data were analyzed by each plant (Table 2).

An one-way analysis of variance (ANOVA) with p-value associated to the Fischer test was performed to compare the means of ψ_{pd} between the test sites regarding the irrigation treatment and the cultivars. These statistical analyses were computed in R (R Core Team 2017) combined with car package (Fox and Weisberg 2011) and agricolae package (Mendiburu 2017).

The hyperspectral data were processed into spectral vegetation indices. A large diversity of vegetation indices, including two-band indices, represented by simple ratios (SR), normalized indices (NRI) and also other formulations defined in the literature were computed (Table 3).

Following previous studies (Pôças, Rodrigues et al. 2015, Pôças, Gonçalves et al. 2017), a band selection procedure for the two-band vegetation indices optimization was considered, testing all two-band combinations (for simple ratio indices and normalized indices) within the spectral range of 400 nm and 1010 nm. Additionally, all combinations of broad bands within specific combinations of the spectral domains of blue, green, red, red edge, and near infrared were tested for the normalized difference vegetation index formulation. The range considered for each spectral domain was 451–520 nm for blue, 521–570 nm, for green, 571–700 nm for red, 681–740 nm for red edge, and 701–950 nm for near infrared (NIR).

A linear fitting function was used for the band selection optimization, having the ψ_{pd} as the dependent variable. A calibration dataset, corresponding to 70% of the total

observations, and a validation dataset, with the 30% remaining observations, were used for assessing the best combination of bands. The bands selected for each vegetation index corresponded to the best combinations obtained for both the calibration and validation datasets, expressed through the determination coefficient (R^2). A total of fifteen vegetation indices were selected following the optimization procedure (Table 3).

The vegetation indices computation and bands optimization were performed in the HSDAR package (Lehnert, Meyer et al. 2017), implemented in software R (R Core Team 2017) and in the spectral indices toolbox of ARTMO software (Verrelst, Rivera et al. 2011, Rivera, Verrelst et al. 2014).

Table 3. Vegetation indices formulations with bands (b) optimized according to grapevines predawn leaf water potential.

Vegetation index[a]	Formulation	Original reference
2-bands – Normalized indices		
$NRI_{515,523}$	$NRI_{515,523} = (b_{523} - b_{515})/(b_{523} + b_{515})$	-
$NRI_{520,701}$	$NRI_{520,701} = (b_{701} - b_{520})/(b_{701} + b_{520})$	-
$NRI_{520,615}$	$NRI_{615,520} = (b_{615} - b_{520})/(b_{615} + b_{520})$	-
$NRI_{520,694}$	$NRI_{520,694} = (b_{694} - b_{520})/(b_{694} + b_{520})$	-
$NRI_{524,615}$	$NRI_{524,615} = (b_{615} - b_{524})/(b_{615} + b_{524})$	-
$NRI_{535,701}$	$NRI_{535,701} = (b_{701} - b_{535})/(b_{701} + b_{535})$	-
$NRI_{529,694}$	$NRI_{520,694} = (b_{694} - b_{529})/(b_{694} + b_{529})$	-
$NRI_{711,700}$	$NRI_{711,700} = (b_{700} - b_{711})/(b_{700} + b_{711})$	-
$NRI_{718,723}$	$NRI_{718,723} = (b_{723} - b_{718})/(b_{723} + b_{718})$	-
2-bands – Simple ratios		
$SR_{718,723}$ [a]	$SR_{718,723} = b_{723}/b_{718}$	-
$WI_{900,970}$	$WI_{900,970} = b_{900}/b_{970}$	(Peñuelas, Filella

		et al. 1993)

2-bands – Other formulations

$ARI_{opt\_665,647}$	$ARI_{opt\_665,647} = (1/b_{647}) - (1/b_{665})$	(Gitelson, Merzlyak et al. 2001)
$MSAVI_{opt\_701,587}$	$\mathrm{MSAVIopt}_{701,587} = \left[2 * b_{701} + 1 - [(2 * b_{701} + 1)^2 - 8(b_{701} - b_{587})]^{1/2}\right]/2$	(Qi, Chehbouni et al. 1994)
$OSAVI_{opt\_745,700}$	$OSAVI_{opt\_745,700} = (b_{745} - b_{700})/(b_{745} + b_{700} + 0.16)$	(Rondeaux, Steven et al. 1996)
$RDVI_{opt\_745,700}$	$\mathrm{RDVI}_{\mathrm{opt\_745,700}} = (b_{745} - b_{700})/\left[(b_{745} + b_{700})^{1/2}\right]$	(Roujean and Breon 1995)

[a] NRI – Normalized Reflectance Index; SR – simple ratio; WI – Water Index; ARI_{opt} – Anthocyanin Reflectance Index optimized; $OSAVI_{opt}$ - Optimal Soil Adjusted Vegetation Index optimized; $MSAVI_{opt}$ - Modified Soil Adjusted Vegetation Index optimized; $RDVI_{opt}$ - Renormalized Difference Vegetation Index optimized.

A time-dynamic variable based on the ψ_{pd} (ψ_{pd_0}) was also used as predictor aimed at representing crop water status dynamics in the post-flowering - harvest period. This ψ_{pd_0} was computed by integrating, in each ψ_{pd} date to be predicted, the information of previous ψ_{pd} measurements. Additionally, this ψ_{pd_0} aimed at minimizing the spurious association between ψ_{pd} and hyperspectral data resulting from a common (downward) time trend inherent to each one of these two types of data. The ψ_{pd_0} was defined for each measurement point and measurement date as the ψ_{pd} value corresponding to the previous measurement.

2.4. Modelling approaches

In the modelling approaches the ψ_{pd} was used as response variable and 631 predictors candidates from both hyperspectral (626) and structural (5) data were tested: i) 611 predictors in the form of spectral reflectance in the range between 400 nm and 1010 nm (wavelength interval of 1 nm), ii) 15 predictors in the form of vegetation indices (Table 3), iii) three qualitative variables related with structural parameters, including, irrigation conditions (two levels: IT_I – irrigated, IT_NI – non-irrigated), cultivar (tree levels: TN, TF, and TB) and test site (two levels: TS_1 and TS_2), iv) the time-dynamic predictor ψ_{pd_0} (three levels of the classification: $\psi_{pd\_0\text{-}1:}$ low $\psi_{pd\_0\text{-}2:}$ moderate, and $\psi_{pd\_0\text{-}3:}$ high) and v) the days after flowering (DAF).

To run the statistical model, the dataset was split into training data (70% of random observations) and validation data (30% of the remains observations) (Kuhn and Johnson 2013). The training and validation datasets integrate the pairs of concurrent measurements of the ψ_{pd} and the corresponding values of the predicting variables.

In the modelling approach, the ψ_{pd} was used as categorical variable. The definition of classes of ψ_{pd} values was based on the analysis of the ψ_{pd} dispersion and on the ψ_{pd} threshold of -0.5 MPa often considered by farmers in Mediterranean regions for irrigation decisions under deficit irrigation strategies (Van Leeuwen, Tregoat et al. 2009, Lopes, Santos et al. 2011). Three class labels were defined: (i) class 1 (low water deficit): 0 MPa > ψ_{pd} > −0.25 MPa; (ii) class 2 (moderate water deficit): −0.25 MPa > ψ_{pd} > −0.5 MPa; and (iii) class 3 (high water deficit): −0.5 MPa > ψ_{pd}.

2.4.1. Predictive modelling applied in classification mode

The ORL was selected for modeling the ordinal response variables ψ_{pd}. The OLR allows building a predictive model on a probabilistic basis. In the present study, the OLR was fitted through a proportional-odds logit model (McCullagh 1980), which is widely applied to represent ordinal responses (Verwaeren, Waegeman et al. 2012). The proportional-odds logit model defines a probability density function over the class labels for a given feature vector x, which belongs to the input space X (McCullagh 1980, Verwaeren, Waegeman et al. 2012).

The “polr” function from the MASS library in software R (Venables and Ripley 2002) was used for (Harrell Jr 2018) applying this methodological approach.

A stepwise regression procedure was used for variables selection. A forward strategy was applied for selecting within the initial 631 predictors candidates. Such selection is based on a saturated model and a null model defined using the “multinom” function from the nnet library in software R (Venables and Ripley 2002). The saturated model uses all possible predictors and assumes that each data point has its own parameters and thus fully explaining individual observations; contrarily, the null model adopts one for the intercept-only model (Venables and Ripley 2002). In each step of the forward stepwise regression, the predicting variable that most contributes for the model improvement compared to the model in the previous step is chosen, based on the lowest value for Akaike information criterion (AIC; Akaike (1974).

2.5. Model performance assessment

The residual deviance (McCullagh 1980, Kleinbaum and Klein 2010) and the Akaike information criterion (AIC; Akaike (1974)) were computed for assessing the model’s quality. The residual deviance represents the ratio of the likelihood of the current model with the likelihood of a model that perfectly predicts the response variable, thus the smaller the deviation the better the fit of the model (Kleinbaum and Klein 2010). The AIC statistics allows comparing between model’s performance, with lower AIC values corresponding to simpler models with fewer predictors (Kuhn and Johnson 2013). The Wald statistic was also used to assess the significance of the predictors selected for the model (Peng, Lee et al. 2002). The odds ratios were calculated to analyse the weight of each predictor. Additionally, the positive prediction value by class and the overall model accuracy were computed and organized in a confusion matrix. The overall model accuracy corresponds to the agreement between predict and observed values in each categorical class, making no distinction in the types of misclassification (Kuhn and Johnson 2013). Differently, the calculated positive prediction value reports the percentage of correct classification cases by class considering the prevalence of the event (Kuhn and Johnson 2013). The caret package (Kuhn 2018), implemented in software R (R Core Team 2017), was used to automatically compute the confusion matrices resulting from the predictive modelling approach in the classification mode.

3. Results

3.1. Analysis of the variability between and within test sites

The impact of irrigation treatments and cultivars in each test site on the ψ_{pd} is presented in Table 4. The test site 2 consistently presented lower ψ_{pd} values, evidencing that vineyards in this test site of sub-region of Douro Superior (Figure 2) are more likely to present water deficit conditions. In both test sites, the irrigation and the grapes varieties presented a significant impact on ψ_{pd} values. When the cultivars are considered, ANOVA suggests that TN is more susceptible to be under stress (lower ψ_{pd}) when compared to TB and TF.

Table 4. Statistical results of predawn leaf water potential (ψpd, MPa) for the different irrigation regimes and grape varieties in the test sites studied.

Structural Parameters	N. Obs	Location (ψ_{pd}, MPa)		Mean ψ_{pd} (MPa)	ANOVA F*
		Test site 1	Test site 2		
Irrigation					
- No Irrigation	117	-0.822	-0.976	-0.936	0.002
- Irrigation	208	-0.434	-0.559	-0.496	0.000
ANOVA F*		*0.000*	*0.000*	*0.000*	---
Grape cultivars					
T. Nacional (TN)	231	-0.566[b]	-0.781[b]	-0.692[b]	0.000
T. Franca (TF)	76	-0.423[a]	-0.726[b]	-0.566[ab]	0.000
T. Barroca (TB)	18	na	-0.541[a]	-0.541[a]	---
ANOVA F*		*0.012*	*0.003*	*0.003*	---
Overall mean	325	-0.523	-0.748	-0.655	0.000

ANOVA F*: is the p-value associated to the Fischer' test performed in the ANOVA; means with p-value less than 0.05 is considered statistically different. Within columns, means followed by the same letter are not significantly different according to Duncan test ($\alpha < 5\%$).

na: No data available.

3.2. Model performance

From the 631 predictors initially considered, nine variables were selected through the stepwise procedure: i) two qualitative variables: Irrigation treatment (IT) and Test site (TS); ii) four vegetation indices: $ARI_{opt\_656,647}$, $NRI_{745,700}$, $NRI_{711,700}$, $WI_{900,970}$; iii) days after flowering (DAF); iv) ψ_{pd_0}, and iv) one wavelength: R996. Then, the OLR model was applied to these nine predicting variables selected by the stepwise method (model 1). This model 1 presented a residual deviance of 246.88 and an AIC of 270.88.

The results of the individual regression coefficients of Wald statistics for each predictor, showed that only the variables "IT_NI", "TS_2", "$ARI_{opt\_656,647}$" and "$NRI_{711,700}$" were statistically significant ($p < 0.01$) (Table 5). These results suggest that an alternative model (model 2) solely including these four statistically significant predictors could be applied to the data. The model 2, combining the four selected variables, presents residual deviance of 250.06 and AIC of 262.06. This AIC value improved when compared to model 1.

For the model 2 the results of the odds ratio indicate that the non-irrigation treatment (IT_NI) has the biggest influence on the assignment of the class, followed by the $ARI_{opt\_656,647}$, test site 2 (TS_2), and the $NRI_{711,700}$.

Table 5. Coefficients determinate by Wald Statistics and odd ratios of the predictors in the models created to estimated ψpd.

	Model 1		Model 2	
Predictors	Coefficient	Odd ratio	Coefficient	Odd ratio
IT_NI	3.06 x 10***	1.95 x 10^{13}***	1.89 x 10***	1.68 x 10^{8}***
Test site_2	1.10***	3.01***	1.35***	3.86***
$ARI_{opt\_656,647}$	1.92***	6.83***	1.77***	5.86***
DAF	-4 x 10^{-3}	9.96 x 10^{-1}	-	-
$NRI_{745,700}$	-1.43	2.40 x 10^{-1}	-	-
$NRI_{711,700}$	-2.28 x 10***	0.00***	-1.80 x 10***	0.00***
$WI_{900,970}$	9.13	9,26 x 10^{3}	-	-
ψpd_0	3.63 x 10^{-1}	1.44	-	-
R996	-2.05	1.28*10^{-1}	-	-

IT_NI: non-irrigation treatment; TS_2: Test site 2; $ARI_{opt\_656,647}$: Anthocyanin Reflectance Index optimized; DAF: days after flowering; NRI: Normalized reflectance index; WI: Water index; ψ_{pd_0}: time-dynamic variable based on ψpd; R996: reflectance at wavelength 966 nm. *p<0.1; **p<0.05; ***p<0.01

The model 2 overall accuracy, assessed through the validation dataset, was 73.2 % and the positive prediction value per class was: 33.3%, 44.4% and 79.3% for the classes low (1), moderate (2), and high (3) water deficit, respectively (Table 6). The variability in the prediction value per class is likely due to the impact of the imbalanced data set used for validation (10, 20 and 67 cases in class 1, class 2, and class 3, respectively).

Additionally, considering the training dataset, it was calculated the probability to estimate the class of ψ_{pd} when all the model's predictors are at their mean values. From the observations analyzed, there are more chances of the ψ_{pd} be classified in the class 3 (68.62%), followed by the class 2 (20.86%) and then class 1 (10.52%).

Table 6. Confusion matrix for the comparison between observed and predicted predawn leaf water potential (ψpd) in the validation dataset.

ψ_{pd} predicted	ψ_{pd} observed			Positive prediction by class (%)
	Low	Moderate	High	
Low	2	3	0	33.3%
Moderate	5	6	2	44.4%
High	3	11	65	79.3%

Predawn leaf water potential (ψ_{pd}) classes: low water deficit, 0 MPa > ψ_{pd} > −0.25 MPa; moderate water deficit, −0.25 MPa > ψ_{pd} > −0.50 MPa; high water deficit, <−0.50 MPa.

4. Discussion

The hyperspectral based predictive model developed in our study is a valuable tool for predicting grapevine ψ_{pd} because of the integration of biophysically sound predictors (model 2): Irrigation treatment (IT), Test Site (TS), $ARI_{opt\_656,647}$ and $NRI_{711,700}$.

The selection of a predictor related with the IT is consistent with the differences between irrigation treatment (Table 4) resulting in differences on the reflectance of irrigated and non-irrigated plants (Pôças, Gonçalves et al. 2017). The selection of a variable relative to the test site is likely associated to the different climatic conditions between sub-regions of Douro Wine region (Figure 2), being the test site 2 located in a warmer and drier zone than test site 1.

Following the vegetation indices optimization for bands selection, the $ARI_{opt\_656,647}$ integrates wavelengths of the red domain, close to the red edge domain, instead of the wavelengths of the original formulation in the green and the red edge domains (Gitelson, Merzlyak et al. 2001). The wavelengths of $ARI_{opt\_656,647}$ are close to those studied by Blackburn (2010), the wavelengths of 680 nm and 635 nm, which are related to the chlorophyll a and chlorophyll b concentrations, respectively. Also, Sonobe, Sano et al. (2018) studied similar bands range to estimate the chlorophyll content. The content of chlorophylls a and b is a potential indicator of vegetation stress (Zarco-Tejada, Miller et al. 2002, Wu, Niu et al. 2008), which includes water stress. As discussed by these authors, several physiological perturbations in the light-dependent reactions of photosynthesis that occur in plants under stress can be related with changes in chlorophylls a and b and assessed through differences in spectral reflectance.

The $NRI_{711,700}$ combines wavelengths of red and red edge. The red-edge zone is reported as a potential indicator of water stress in plants and thus the construction of vegetation indices using this zone of the spectrum can provide information about the water status of the plant (Zarco-Tejada, González-Dugo et al. 2013, Fang, Ju et al. 2017, Rodríguez-Pérez, Ordóñez et al. 2018).

Although it was obtained a good accuracy of prediction (73.2%), the model was able to better classify the classes of higher stress, which may be due to the imbalanced number of observations of each class in the data set. As discussed by Brodersen, Ong et al. (2010), an imbalanced data set may lead to misleading conclusions about the performance of a classification predictive model when using an average accuracy measure.

The analysis of the positive prediction value per class, which is a more robust measure as it takes in consideration the prevalence of each class, shows a good performance for the class of high-water deficit (class 3 (ψ_{pd} > -0.5 MPa) = 79.27% of cases correctly classified). In Mediterranean regions, the irrigation (under deficit irrigation strategies) most often stars when plants are under ψ_{pd} values below -0.5 MPa to stabilize the wine quality conditions (Van Leeuwen, Tregoat et al. 2009, Lopes, Santos et al. 2011) and thus the results obtained for this specific class are particularly interesting.

5. Conclusion

In this study we presented how the predawn leaf water potential in vineyards of Douro wine region could be predicted by a classification model based on hyperspectral reflectance

data. A large set of climatic, environmental, and agronomic conditions were sampled to test model's accuracy and robustness.

The developed predictive model presented an overall accuracy of 73.2%. The variables selected provide information of plant physiology relevant for the prediction of the water status in grapevines. Nevertheless, the modelling could be improved if a higher number of samples were assessed in the field to avoid problems related to imbalanced observations in the classes.

The use of the classification model to estimate ψ_{pd} brings a potential application to support irrigation decision in viticulture. Usually, the operational decisions about the vine's irrigation scheduling are done for ψ_{pd} values associated to the class 3 of this study, where the model obtained performed better. Therefore, the results of the proposed model have potential to be used in support to irrigation tasks. Moreover, the use of classes of ψ_{pd} instead of continuous values provides easier-to-use information for farmers. The accuracy and operability of this predictive model justify their use to support decision-making process related with improvement of water productivity in vineyards. This work analyses data obtained on the ground level, while these results are the first step towards applications using other sensors mounted on aerial platforms (e.g. drones or satellites). This is in line, with the high number of forthcoming hyperspectral sensors mounted in aerial platforms, which will allow for the generation of hyperspectral time-series, giving access to spatial and temporal dynamics of crop biophysical parameters. Thus, the results presented in this work can be used in support to the development of new technologies for vineyards water status monitoring based on hyperspectral data.

Also, these results could be used with other sensors mounted in drones or satellite in order to mapping this information in vineyards.

Acknowledgments

Isabel Pôças acknowledges the Post-Doctoral grant of the project ENGAGE-SKA POCI-01-0145-FEDER-022217, co-funded by FEDER through COMPETE (POCI-01-0145-FEDER-022217). The authors also thank the wine company Real Companhia Velha (and its Coordinator for Viticulture Rui Soares) and the wine company Symington Family Estates (and its R&D Viticulture Manager Fernando Alves) for the facilities provided for fieldwork, as well as the Associação para o Desenvolvimento da Viticultura Duriense (ADVID) for funding part of the field-work missions. This study was implemented under a cooperation protocol

integrating Faculty of Sciences, University of Porto, ADVID, Real Companhia Velha, and Symington Family Estates.

6. References

Akaike, H. (1974). "A new look at the statistical model identification." IEEE Transactions on Automatic Control 19(6): 716-723.

Alves, F., J. Costa, P. Costa, C. Correia, B. Gonçalves, R. Soares and J. Moutinho-Pereira (2012). Influence of climate and deficit irrigation on grapevine physiology, yield and quality attributes, of the cv. Touriga Nacional at Douro Region. Proceeding of the IX international Terroirs congress.

Alves, F., C. J, P. Costa, C. Correia, B. Gonçalves, S. R and J. Moutinho Pereira (2013). Grapevine water stress management in Douro Region: Long-term physiology, yield and quality studies in cv. Touriga Nacional.

Bellvert, J., P. J. Zarco-Tejada, J. Girona and E. Fereres (2014). "Mapping crop water stress index in a 'Pinot-noir' vineyard: comparing ground measurements with thermal remote sensing imagery from an unmanned aerial vehicle." Precision Agriculture 15(4): 361-376.

Blackburn, G. A. (2007). "Hyperspectral remote sensing of plant pigments." J Exp Bot 58(4): 855-867.

Blackburn, G. A. (2010). "Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves." International Journal of Remote Sensing 19(4): 657-675.

Brant, R. (1990). "Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression." Biometrics 46(4): 1171-1178.

Brodersen, K. H., C. S. Ong, K. E. Stephan and J. M. Buhmann (2010). The Balanced Accuracy and Its Posterior Distribution. 2010 20th International Conference on Pattern Recognition.

Caicedo, J. P. R., J. Verrelst, J. Muñoz-Marí, J. Moreno and G. Camps-Valls (2014). "Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7(4): 1249-1259.

Chaves, M. M., O. Zarrouk, R. Francisco, J. M. Costa, T. Santos, A. P. Regalado, M. L. Rodrigues and C. M. Lopes (2010). "Grapevine under deficit irrigation: hints from physiological and molecular data." Ann Bot 105(5): 661-676.

Clevers, J. G. P. W., L. Kooistra and M. E. Schaepman (2010). "Estimating canopy water content using hyperspectral remote sensing data." International Journal of Applied Earth Observation and Geoinformation 12(2): 119-125.

Coppock, D. L. (2011). "Ranching and Multiyear Droughts in Utah: Production Impacts, Risk Perceptions, and Changes in Preparedness." Rangeland Ecology & Management 64(6): 607-618.

Cunha, M. and C. Richter (2016). "The impact of climate change on the winegrape vineyards of the Portuguese Douro region." Climatic Change 138(1): 239-251.

De Bei, R., D. Cozzolino, W. Sullivan, W. Cynkar, S. Fuentes, R. Dambergs, J. Pech and S. Tyerman (2011). "Non-destructive measurement of grapevine water potential using near infrared spectroscopy." Australian Journal of Grape and Wine Research 17(1): 62-71.

Deloire, A., H. Ojeda, O. Zebic, N. Bernard, J. J. Hunter and A. Carbonneau (2005). "Influence de l'état hydrique de la vigne sur le style de vin." Progres Agricole et Viticole 122(21): 455-462.

Fang, M., W. Ju, W. Zhan, T. Cheng, F. Qiu and J. Wang (2017). "A new spectral similarity water index for the estimation of leaf water content from hyperspectral data of leaves." Remote Sensing of Environment 196: 13-27.

Feng, S., Y. Itoh, M. Parente and M. F. Duarte (2017). "Hyperspectral Band Selection From Statistical Wavelet Models." IEEE Transactions on Geoscience and Remote Sensing 55(4): 2111-2123.

Ferreira, H. A. (1965). Normais climatológicas do continente, Açores e Madeira correspondentes a 1931-1960. Lisboa, Serviço Meteorológico Nacional.

Fox, J. and S. Weisberg (2011). An (R) Companion to Applied Regression. Thousand Oaks (CA), Sage. Second.

Gitelson, A. A., M. Merzlyak, Y. Zur, R. Stark and U. Gritz (2001). "Non-destructive and remote sensing techniques for estimation of vegetation status."

Gitelson, A. A., M. N. Merzlyak and O. B. Chivkunova (2001). "Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves¶." Photochemistry and Photobiology 74(1): 38-45.

Harrell, F. E. (2015). Ordinal Logistic Regression. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Springer International Publishing: 311-325.

Harrell Jr, F. E. (2018). rms: Regression Modeling Strategies. R package version 5.1-2.

Im, J., J. R. Jensen, M. Coleman and E. Nelson (2009). "Hyperspectral remote sensing analysis of short rotation woody crops grown with controlled nutrient and irrigation treatments." Geocarto International 24(4): 293-312.

INMG (1991). O Clima de Portugal. Fascículo XLIX. Normais climatológicas da região de Trás-os-Montes e Alto Douro e Beira Interior correspondentes a 1951-1980. Lisboa, INMG.

Jones, G. V. and F. Alves (2012). "Impact of climate change on wine production: a global overview and regional assessment in the Douro Valley of Portugal." International Journal of Global Warming 4(3-4): 383-406.

Jones, H. G. and R. A. Vaughan (2010). Remote sensing of vegetation: principles, techniques, and applications. New York, USA, Oxford University Press Inc.

Kleinbaum, D. G. and M. Klein (2010). Ordinal Logistic Regression. Logistic Regression. Statistics for Biology and Health. M. Gail, K. Krickeberg, J. M. Samet, A. Tsiatis and W. Wong. New York, NY, Springer.

Kuhn, M. (2018). Caret: Classification and Regression Training. R package version 6.0-80.

Kuhn, M. and K. Johnson (2013). Applied predictive modeling. New York, Springer Science+Business Media.

Lehnert, L. W., H. Meyer and J. Bendix (2017). hsdar: Manage, analyse and simulate hyperspectral data in R.

Lopes, C. M., T. P. Santos, A. Monteiro, M. L. Rodrigues, J. M. Costa and M. M. Chaves (2011). "Combining cover cropping with deficit irrigation in a Mediterranean low vigor vineyard." Scientia Horticulturae 129(4): 603-612.

Mariotto, I., P. S. Thenkabail, A. Huete, E. T. Slonecker and A. Platonov (2013). "Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission." Remote Sensing of Environment 139: 291-305.

McCullagh, P. (1980). "Regression Models for Ordinal Data." Journal of the Royal Statistical Society. Series B (Methodological) 42(2): 109-142.

Medrano, H., M. Tomás, S. Martorell, J.-M. Escalona, A. Pou, S. Fuentes, J. Flexas and J. Bota (2015). "Improving water use efficiency of vineyards in semi-arid regions. A review." Agronomy for Sustainable Development 35(2): 499-517.

Mendiburu, F. d. (2017). agricolae: Statistical Procedures for Agricultural Research.

Merli, M. C., M. Gatti, M. Galbignani, F. Bernizzoni, E. Magnanini and S. Poni (2015). "Comparison of whole-canopy water use efficiency and vine performance of cv. Sangiovese (Vitis vinifera L.) vines subjected to a post-veraison water deficit." Scientia Horticulturae 185: 113-120.

Moutinho-Pereira, J., N. Magalhães, B. Gonçalves, E. Bacelar, M. Brito and C. Correia (2007). "Gas exchange and water relations of three Vitis vinifera L. cultivars growing under Mediterranean climate." Photosynthetica 45(2): 202-207.

Notario del Pino, J. S. and J.-R. Ruiz-Gallardo (2015). "Modelling post-fire soil erosion hazard using ordinal logistic regression: A case study in South-eastern Spain." Geomorphology 232: 117-124.

Oumar, Z. and O. Mutanga (2010). "Predicting plant water content in Eucalyptus grandis forest stands in KwaZulu-Natal, South Africa using field spectra resampled to the Sumbandila Satellite Sensor." International Journal of Applied Earth Observation and Geoinformation 12(3): 158-164.

Peng, C.-Y. J., K. L. Lee and G. M. Ingersoll (2002). "An Introduction to Logistic Regression Analysis and Reporting." The Journal of Educational Research 96(1): 3-14.

Peñuelas, J., I. Filella, C. Biel, L. Serrano and T. Savé (1993). "The reflectance at the 950-970 nm as an indicator of plant water status." International Journal of Remote Sensing 14(10): 1887-1905.

Pôças, I., J. Gonçalves, P. M. Costa, I. Gonçalves, L. S. Pereira and M. Cunha (2017). "Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region." International Journal of Applied Earth Observation and Geoinformation 58: 177-190.

Pôças, I., A. Rodrigues, S. Gonçalves, P. Costa, I. Gonçalves, L. Pereira and M. Cunha (2015). "Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices." Remote Sensing 7(12): 16460-16479.

Prata-Sena, M., B. M. Castro-Carvalho, S. Nunes, B. Amaral and P. Silva (2018). "The terroir of Port wine: Two hundred and sixty years of history." Food Chem 257: 388-398.

Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr and S. Sorooshian (1994). "A modified soil adjusted vegetation index." Remote Sensing of Environment 48(2): 119-126.

R Core Team (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Rivera, J., J. Verrelst, J. Delegido, F. Veroustraete and J. Moreno (2014). "On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization." Remote Sensing 6(6): 4927-4951.

Rodríguez-Pérez, J. R., C. Ordóñez, A. B. González-Fernández, E. Sanz-Ablanedo, J. B. Valenciano and V. Marcelo (2018). "Leaf water content estimation by functional linear regression of field spectroscopy data." Biosystems Engineering 165: 36-46.

Rondeaux, G., M. Steven and F. Baret (1996). "Optimization of soil adjusted vegetation indices." Remote Sensing of Environment 55: 95-107.

Roujean, J.-L. and F.-M. Breon (1995). "Estimating PAR absorbed by vegetation from bidirectional reflectance measurements." Remote Sensing of Environment 51(3): 375-384.

Rutherford, G. N., A. Guisan and N. E. Zimmermann (2007). "Evaluating sampling strategies and logistic regression methods for modelling complex land cover changes." Journal of Applied Ecology 44(2): 414-424.

Scholander, P. F., E. D. Bradstreet, E. A. Hemmingsen and H. T. Hammel (1965). "Sap Pressure in Vascular Plants." Negative hydrostatic pressure can be measured in plants 148(3668): 339-346.

Sonobe, R., T. Sano and H. Horie (2018). "Using spectral reflectance to estimate leaf chlorophyll content of tea with shading treatments." Biosystems Engineering 175: 168-182.

Suárez, L., P. J. Zarco-Tejada, G. Sepulcre-Cantó, O. Pérez-Priego, J. R. Miller, J. C. Jiménez-Muñoz and J. Sobrino (2008). "Assessing canopy PRI for water stress detection with diurnal airborne imagery." Remote Sensing of Environment 112(2): 560-575.

Van Leeuwen, C., O. Tregoat and X. Choné (2009). Vine water status is a key factor in grape ripening and vintage quality for red Bordeaux wine. How can it be assessed for vineyard management purposes?, Journal International des Sciences de la Vigne et du Vin.

Venables, W. N. and B. D. Ripley (2002). Modern Applied Statistics with S. Fourth Edition. New York, Springer.

Verrelst, J., J. Rivera, L. Alonso and J. Moreno (2011). ARTMO: an Automated Radiative Transfer Models Operator toolbox for automated retrieval of biophysical parameters through model inversion. Proc. EARSeL 7th SIG-Imag. Spectrosc. Workshop.

Verwaeren, J., W. Waegeman and B. De Baets (2012). "Learning partial ordinal class memberships with kernel-based proportional odds models." Computational Statistics & Data Analysis 56(4): 928-942.

Wu, C., Z. Niu, Q. Tang and W. Huang (2008). "Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation." Agricultural and Forest Meteorology 148(8-9): 1230-1241.

Zarco-Tejada, P. J., V. González-Dugo, L. E. Williams, L. Suárez, J. A. J. Berni, D. Goldhamer and E. Fereres (2013). "A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index." Remote Sensing of Environment 138: 38-50.

Zarco-Tejada, P. J., J. R. Miller, G. H. Mohammed, T. L. Noland and P. H. Sampson (2002). "Vegetation Stress Detection through Chlorophyll + Estimation and Fluorescence Effects on Hyperspectral Imagery." Journal of Environment Quality 31(5).

Renan Tosin is currently a student of master's degree in agricultural engineering at Faculty of Sciences of University of Porto, Portugal. His main research field involves remote sensing techniques applied in precision agriculture, with focus on vineyards inputs optimization.

Isabel Pôças obtained a PhD degree in Agrarian Sciences in 2010. She currently holds a post-doctoral position at Geo-Space Sciences Research Centre, in liaison with Faculty of Sciences, University of Porto, Portugal. Her specialization domains include the development and application of methodologies based on remote sensing data and tools to support the agriculture water management and the monitoring of vegetation and landscape dynamics. Additionally, she has teaching experience in Agrarian Sciences domains.

Igor Gonçalves holds a degree in Biology and is currently a student of master's degree in Oenology and Viticulture at UTAD, Vila Real. Works at ADVID since 2012 and has developed work in several areas of viticulture, namely in research and development projects. Is currently responsible for the Agro-Environmental Measures Processes and Grapevine Water Relations, conducting several essays in the scope of the short-term adaptation measures regarding climate change in the Douro Wine Region.

Mario Cunha PhD in Agrarian Sciences, is currently an Assistant Professor with the Dep. Geosciences, Environment and Spatial Planning at Sciences Faculty - University of Porto and senior researcher at Institute for Systems and Computer Engineering, Technology and Science (INESC-TEC), CRIIS. His research interests include various topics in agronomy and agricultural engineering mostly focused on crop modelling (statistical or process based), crop yield forecast (climate change included) and remote sensing agricultural applications.

4. General discussion

The PA has been increasingly attracting the attention from researchers and commercial companies working with different crop systems (Hall et al., 2002). The application of PA techniques considers the between and within crop fields variability aiming for the specific management for each area (Huang et al., 2018). Thus, the PA can be fed by RS-derived data that is processed and analysed to support decision making in relation to diverse agronomic purposes (Huang et al., 2018; Hungate et al., 2008; Jones and Vaughan, 2010; Mulla, 2013). Accurate quantitative estimation of vegetation biochemical and biophysical characteristics is necessary for a large variety of agricultural applications. RS, because of its global coverage, repetitiveness, and non-destructive and relatively cheap characterization of land surfaces, has been recognized as a reliable method and a practical means of estimating various biophysical and biochemical vegetation variables as reported by Rapaport et al. (2015) and Orlandi et al. (2018). The advent of hyperspectral remote sensing has offered possibilities for measuring specific vegetation variables that were difficult to measure using conventional multi-spectral sensors. Considering the present-day technology and the available literature, the cornerstone for maximizing the potential use of future hyperspectral data is additional research on spatial and temporal enhancement approaches through synergies with other sensors such as proxima sensors. Such investigations could help to overcome the expensive acquisition of airborne hyperspectral images, which are spatially and temporally limited.

Viticulture represents the second most important perennial crop in Portugal (INE, 2018) and technologies that help in decision making, including those based on RS-derived data and products, will bring benefits in the optimization of the vines production system (Usha and Singh, 2013).

This dissertation presents three case studies that show in what extent RS techniques can be applied to optimize the viticulture in Portugal. One of the case studies presents the application of thermal and spectral data to assess the effect of kaolin application throughout the grapevines cycle (article 3.1), while the other case studies present the application of spectral data for predicting the ψ_{pd} using different modelling approaches (articles 3.2 and 3.3). The experimental units in the article 3.2 are groups of plants, while in the article 3.3 individual plants are used. Overall, a large set of agronomic, environmental, and climatic conditions were sampled in the various case studies, including various grapevine cultivars (articles 3.1, 3.2 and 3.3), two test sites in different sub-regions of the Douro wine region (articles 3.2 and 3.3), and multi-years (article 3.2). Hyperspectral data from the visible and

NIR zones of the electromagnetic spectrum were considered in all the case studies. These reflectance data were directly applied in articles 3.1 and 3.3, while in the articles 3.2 and 3.3 were applied in the form of VIs, combining data from different zones of the spectrum (articles 3.1, 3.2, and 3.3). Additionally, thermal data were considered in the case study presented in article 3.1.

4.1. Application of kaolin in vineyards

In article 3.1, statistical differences were observed in the reflectance in the visible (wavelengths of 400 nm and 535 nm) and red edge (wavelength of 733 nm) zones of the spectrum, which can be related to the pigmentation of the leaves and water absorption, respectively. The selection of a predicting variable of the green zone (R535 nm) by the modelling approach developed suggests that kaolin is able to protect the canopy by reflecting the light from the sun, in the green zone of the spectrum. Such canopy protection is obtained by avoiding the de-epoxidation of xanthophylls, responsible for the heating dissipation in the leaves (Middleton et al., 2012; Shellie and King, 2013). Additionally, the red-edge zone, which was also considered in the modelling approach (R733), can be related to plant stress (Peñuelas et al., 1994; Zarco-Tejada et al., 2013a), including stress related with plant water and heat conditions. Moreover, the thermal data allowed assessing the impact of the loosing effect of kaolin throughout the grapevines cycle. This effect was shown by the increment of temperatures and of reflectance in the zone related to xanthophyll in leaves that were previously covered by the particle-film. Thus, the combined use of spectral data and thermal data in our case study indicated that kaolin has positive effects during the first days after the application as also shown by other authors (Shellie and Glenn, 2008); however, to keep the positive benefit in the vineyards, more application of the product is required throughout the grapevine cycle.

4.2. Assessment of grapevine water status

In the articles 3.2 and 3.3, the spectral data in the form of VIs were used as predictors for a predictive model of ψ_{pd}. The visible and NIR zones that were used to build the VIs are reported in the bibliography as potentially useful to assess the water status in plants, either by relating with water absorption bands or by acting as proxy of physiological processes related with water status (Jones and Vaughan, 2010).

The zones of the electromagnetic spectrum that were considered in the VIs selected for modelling the ψ_{pd} included the green ($NRI_{554.561}$) and the NIR ($WI_{900,970}$) in article 3.2, and the red ($ARI_{opt\_656,647}$) and red edge ($NRI_{711,703}$) in article 3.3. As discussed in the previous case study (article 3.1), the green (represented by the $NRI_{554.561}$) and the red edge (represented by the $NRI_{711,703}$) zones of the spectrum can be used as proxy for stress conditions, including those related with plant water stress (Middleton et al., 2012; Pôças et al., 2017; Zarco-Tejada et al., 2013a). Also, the red zone, close to the red-edge, expressed by $ARI_{opt\_656,647}$ can be related to the Chlorophyll a and Chlorophyll b content (Blackburn, 2010). The Chlorophylls a and b are pigments associated with stress in plants, which includes water stress (Wu et al., 2008; Zarco-Tejada et al., 2002). Differently, the $WI_{900,970}$ (Peñuelas et al., 1997) integrates in its formulation the wavelength 970 nm, reported as one of the wavelengths associated to the water absorption in leaves (Jones and Vaughan, 2010).

Additionally to the VIs, an innovative predicting variable was included in the modelling approaches of both article 3.2 and 3.3, representing a time variable based on ψ_{pd} (ψ_{pd} _0). This time-dynamic variable allowed the models to learn, in each moment, from previous ψ_{pd} dynamics.

Two different modelling approaches were considered for estimating the ψ_{pd}, one based on continuous values of ψ_{pd} (article 3.2) and another based on classes of ψ_{pd} values (article 3.3). Therefore, two types of data can be predicted (continuous values and classes of values), which allow covering the needs of different types of stakeholders, including farmers, water managers, and researchers. Various machine learning techniques have been successfully applied in modelling water status in canopy when spectral data is used (Pôças et al., 2017; Pôças et al., 2015; Rodríguez-Pérez et al., 2007). The application of different algorithms aims to improve the performance of each predictor in the model (Gitelson et al., 2003; Huang et al., 2018; Pôças et al., 2017). In article 3.2, four algorithms using non-linear non-parametric methods were considered for predicting ψ_{pd} (continuous) values, while article 3.3 tested an ordinal logistic regression (OLR) algorithm to classify the ψ_{pd}. The good performance of the methodologies presented in both case studies indicate the potential of these modelling tools for application in operational conditions. In fact, it is important to highlight that the case studies were implemented in commercial vineyards, under actual operation conditions.

Different validation procedures were considered in articles 3.2 and 3.3. In article 3.2, where data of three cultivars, two study areas and three years of data collection were used, an external validation was considered. Differently, in article 3.3, where the data of three cultivars and two study areas relative to a single year were used, an internal validation was

performed by considering a random sample corresponding a 30% of the full dataset. Both validation procedures can be found in studies relative to the application of spectral data for assessing biophysical parameters of vegetation (Caicedo et al., 2014; Pôças et al., 2017; Verrelst et al., 2012). Nevertheless, the external validation brings more robustness to the model (Harrell, 2015).

The model presented in the second case study provided and overall accuracy (83%) that is higher than the model developed in the case study 3 (73.2%). The lower performances of the proposed classification model could be explained by the use of data from individual grapevine (case study 3) instead a group of grapevines (case study 2), as well as the exceptionally high temperature and very low precipitation during the summer period of the year 2017.

4.3. Crop monitoring

As shown by the case studies presented, the use of RS portable instruments, including handheld spectroradiometers and thermal cameras, provides spectral and thermal information that can be used to assess the grapevine status and to help in decision making processes, e.g., those related to kaolin application and irrigation management.

The hyperspectral data provided at field level both by handheld spectroradiometers and thermal cameras constitute a very important asset for better understanding the relationship between different zones of the electromagnetic spectrum and different biophysical parameters of the vegetation. Additionally, these data can be used for calibrating the information to be collected by cameras onboard aerial platforms, e.g. drones. The high cost of data collected by hyperspectral cameras attached to drones and satellites limits its use in operational applications. Nevertheless, the information obtained with portable instruments can be used for establishing correlations with multispectral data derived from drones and satellites, which are more easily available, and also for adjusting filters of bands to be applied in cameras onboard aerial platforms aiming for the data collection in the spectral zones of interest.

5. General conclusions and perspectives

This dissertation contributes to the field of information extraction from hyperspectral measurements and enhances our understanding of grapevine biophysical characteristics estimation. Several achievements have been registered in exploiting spectral information for the retrieval of vegetation biophysical parameters using different statistical approaches such as principal components and machine learning.

The hypothesis that RS is a valid tool for PA is well established, as mentioned along the dissertation. In order to provide potential RS applications for PA, there is a need to be more specific, therefore, three cases studies exploring new RS methods for potential use in conjunction with PA was presented in this work. These involve the derivation of optimal spectral vegetation indices coupled with the successful modelling of the impact of kaolin applications and crop water status (with extensive validation).

An important contribute of RS in viticulture crops settled in the Mediterranean climate can be related with applications associated with irrigation management, considering the increasing events of droughts in summers periods. The (deficit) irrigation in specific stages of grapevines cycle can positively contribute to increase the yield and improve the grape quality and thus easy-to-use and rigorous methodologies for assessing vines water status are of upmost relevance. In this dissertation, there were presented two different models to estimate the water status in grapevines based on hyperspectral data. Both methods were successfully tested through different approaches to estimate water status based on ψ_{pd}.

Also, one of the techniques reported to minimize the impacts of high temperatures over vineyards during summer in Mediterranean conditions is the application of kaolin. In this dissertation, the assessment of hyperspectral and thermal data allowed evaluation the effects of the kaolin application throughout the grapevine cycle, following physiological assumptions in the vineyards, as well as to make recommendations about the kaolin application.

The decision making in PA is usually better understood by mapping the results. The results provided in this dissertation can be further used to support the mapping of the parameters assessed, i.e., the leaf protective energy dissipation under kaolin application and the vines water status. In the case of kaolin application, the mapping could improve the application of the product by dosing the adequate quantity in each area and when the application should be done. The same thought is valid for the vines water status assessment, where a ψ_{pd} mapping could show the differences of water requirements in the vineyard and when the irrigation system should be used.

Notwithstanding this work has given valuable information regarding potential application of RS in PA, specifically focusing vineyards in Douro wine region, these findings should be further tested and comprehended before achieving a commercial utilization. The domain of the techniques by the users is fundamental to apply these finding in commercial utilization.

This dissertation analyses data obtained on the ground level, while these results are the first step towards applications using other sensors mounted on aerial platforms (e.g. drones or satellites). On the other hand, the forthcoming hyperspectral sensors that are planned for very soon, will allow for the generation of hyperspectral time-series, giving access to spatial and temporal dynamics of crop biophysical parameters. Thus, the results presented in this work can be used in support to the development of new technologies for crop monitoring based on hyperspectral data.

References

Allen, R, Irmak, A, Trezza, R, Hendrickx, J M H, Bastiaanssen, W, and Kjaersgaard, J (2011). Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrological Processes* **25**, 4011-4027. doi:10.1002/hyp.8408

Allen, R G, Pereira, L S, Raes, D, and Smith, M (1998). FAO Irrigation and Drainage paper No. 56, Crop Evapotranspiration. Rome, Italy: FAO.

Anapalli, S S, Green, T R, Reddy, K N, Gowda, P H, Sui, R, Fisher, D K, Moorhead, J E, and Marek, G W (2018). Application of an energy balance method for estimating evapotranspiration in cropping systems. *Agricultural Water Management* **204**, 107-117. doi:10.1016/j.agwat.2018.04.005

Apan, A, Held, A, Phinn, S, and Markley, J (2004). Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery. *International Journal of Remote Sensing* **25**, 489-498. doi:10.1080/01431160310001618031

Blackburn, G A (1998). Quantifying Chlorophylls and Caroteniods at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches. *Remote Sensing of Environment* **66**, 273-285. doi:https://doi.org/10.1016/S0034-4257(98)00059-5

Blackburn, G A (2007). Hyperspectral remote sensing of plant pigments. *J Exp Bot* **58**, 855-67. doi:10.1093/jxb/erl123

Blackburn, G A (2010). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing* **19**, 657-675. doi:10.1080/014311698215919

Boochs, F, Kupfer, G, Dockter, K, and Kühbauch, W (1990). Shape of the red edge as vitality indicator for plants. *Remote Sensing* **11**, 1741-1753.

Borgogno-Mondino, E, Novello, V, Lessio, A, and de Palma, L (2018). Describing the spatio-temporal variability of vines and soil by satellite-based spectral indices: A case study in Apulia (South Italy). *International Journal of Applied Earth Observation and Geoinformation* **68**, 42-50. doi:10.1016/j.jag.2018.01.013

Broge, N H, and Leblanc, E (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment* **76**, 156-172. doi:https://doi.org/10.1016/S0034-4257(00)00197-8

Buitrago Acevedo, M F, Groen, T A, Hecker, C A, and Skidmore, A K (2017). Identifying leaf traits that signal stress in TIR spectra. *ISPRS Journal of Photogrammetry and Remote Sensing* **125**, 132-145. doi:10.1016/j.isprsjprs.2017.01.014

Caicedo, J P R, Verrelst, J, Muñoz-Marí, J, Moreno, J, and Camps-Valls, G (2014). Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**, 1249-1259. doi:10.1109/JSTARS.2014.2298752

Campbell, G S, and Norman, J M (1998). The Light Environment of Plant Canopies. *In* "An Introduction to Environmental Biophysics", pp. 247-278. Springer New York, New York, NY. doi:10.1007/978-1-4612-1626-1_15

Carter, G A (1994). Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote sensing* **15**, 697-703.

Chappelle, E W, Kim, M S, and McMurtrey, J E (1992). Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sensing of Environment* **39**, 239-247. doi:https://doi.org/10.1016/0034-4257(92)90089-3

Chen, J M (1996). Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Canadian Journal of Remote Sensing* **22**, 229-242. doi:10.1080/07038992.1996.10855178

Chen, S, Zhang, F, Ning, J, Liu, X, Zhang, Z, and Yang, S (2015). Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging. *Food Chem* **172**, 788-93. doi:10.1016/j.foodchem.2014.09.119

Cheng, T, Rivard, B, Sánchez-Azofeifa, A G, Féret, J-B, Jacquemoud, S, and Ustin, S L (2014). Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* **87**, 28-38. doi:10.1016/j.isprsjprs.2013.10.009

Chicati, M L, Nanni, M R, Cezar, E, de Oliveira, R B, and Chicati, M S (2017). Spectral classification of soils: A case study of Brazilian flooded soils. *Remote Sensing Applications: Society and Environment* **6**, 39-45. doi:10.1016/j.rsase.2017.04.002

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

Dash, J, and Curran, P J (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing* **25**, 5403-5413. doi:10.1080/0143116042000274015

Datt, B (1998). Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a+b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sensing of Environment* **66**, 111-121. doi:https://doi.org/10.1016/S0034-4257(98)00046-7

Datt, B (1999). Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing* **20**, 2741-2759. doi:10.1080/014311699211778

Daughtry, C S T, Walthall, C L, Kim, M S, de Colstoun, E B, and McMurtrey, J E (2000). Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment* **74**, 229-239. doi:https://doi.org/10.1016/S0034-4257(00)00113-9

Dechant, B, Cuntz, M, Vohland, M, Schulz, E, and Doktor, D (2017). Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism. *Remote Sensing of Environment* **196**, 279-292. doi:10.1016/j.rse.2017.05.019

DeMeo, G A, Laczniak, R J, Boyd, R A, Smith, J, and Nylund, W E (2003). "Estimated ground-water discharge by evapotranspiration from Death Valley, California, 1997–2001."

Draper, N R, and Smith, H (2014). "Applied regression analysis," John Wiley & Sons.

Elvanidi, A, Katsoulas, N, Ferentinos, K P, Bartzanas, T, and Kittas, C (2018). Hyperspectral machine vision as a tool for water stress severity assessment in soilless tomato crop. *Biosystems Engineering* **165**, 25-35. doi:10.1016/j.biosystemseng.2017.11.002

Elvidge, C D, and Chen, Z (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sensing of Environment* **54**, 38-48. doi:https://doi.org/10.1016/0034-4257(95)00132-K

Eshel, G, Levy, G J, and Singer, M J (2004). Spectral Reflectance Properties of Crusted Soils under Solar Illumination. *Soil Science Society of America Journal* **68**. doi:10.2136/sssaj2004.1982

Feng, S, Itoh, Y, Parente, M, and Duarte, M F (2017). Hyperspectral Band Selection From Statistical Wavelet Models. *IEEE Transactions on Geoscience and Remote Sensing* **55**, 2111-2123. doi:10.1109/tgrs.2016.2636850

Féret, J B, Gitelson, A A, Noble, S D, and Jacquemoud, S (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment* **193**, 204-215. doi:10.1016/j.rse.2017.03.004

Filella, I, and Peñuelas, J (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing* **15**, 1459-1470. doi:10.1080/01431169408954177

Gago, J, Fernie, A R, Nikoloski, Z, Tohge, T, Martorell, S, Escalona, J M, Ribas-Carbo, M, Flexas, J, and Medrano, H (2017). Integrative field scale phenotyping for investigating metabolic components of water stress within a vineyard. *Plant Methods* **13**, 90. doi:10.1186/s13007-017-0241-z

Gamon, J A, Peñuelas, J, and Field, C B (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment* **41**, 35-44. doi:https://doi.org/10.1016/0034-4257(92)90059-S

Gandia, S, Fernández, G, García, J, and Moreno, J (2004). Retrieval of vegetation biophysical variables from CHRIS/PROBA data in the SPARC campaign. *Esa Sp* **578**, 40-48.

Garrity, S R, Eitel, J U H, and Vierling, L A (2011). Disentangling the relationships between plant pigments and the photochemical reflectance index reveals a new approach for remote estimation of carotenoid content. *Remote Sensing of Environment* **115**, 628-635. doi:10.1016/j.rse.2010.10.007

Geladi, P, and Kowalski, B R (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta* **185**, 1-17. doi:https://doi.org/10.1016/0003-2670(86)80028-9

Gitelson, A, and Merzlyak, M N (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology* **22**, 247-252. doi:https://doi.org/10.1016/1011-1344(93)06963-4

Gitelson, A A, Buschmann, C, and Lichtenthaler, H K (1999). The Chlorophyll Fluorescence Ratio F735/F700 as an Accurate Measure of the Chlorophyll Content in Plants. *Remote Sensing of Environment* **69**, 296-302. doi:https://doi.org/10.1016/S0034-4257(99)00023-1

Gitelson, A A, Gritz, Y, and Merzlyak, M N (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J Plant Physiol* **160**, 271-82. doi:10.1078/0176-1617-00887

Gitelson, A A, Kaufman, Y J, and Merzlyak, M N (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment* **58**, 289-298. doi:https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A A, Merzlyak, M, Zur, Y, Stark, R, and Gritz, U (2001). Non-destructive and remote sensing techniques for estimation of vegetation status.

Gitelson, A A, and Merzlyak, M N (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing* **18**, 2691-2697. doi:10.1080/014311697217558

Gorsevski, P V, and Gessler, P E (2009). The design and the development of a hyperspectral and multispectral airborne mapping system. *ISPRS Journal of Photogrammetry and Remote Sensing* **64**, 184-192. doi:10.1016/j.isprsjprs.2008.09.002

Guyot, G, and Baret, F (1988). Utilisation de la haute resolution spectrale pour suivre l'etat des couverts vegetaux. *In* "Spectral Signatures of Objects in Remote Sensing", Vol. 287, pp. 279.

Haboudane, D, Miller, J R, Tremblay, N, Zarco-Tejada, P J, and Dextraze, L (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment* **81**, 416-426. doi:https://doi.org/10.1016/S0034-4257(02)00018-4

Hall, A, Lamb, D W, Holzapfel, B, and Louis, J (2002). Optical remote sensing applications in viticulture - a review. *Australian Journal of Grape and Wine Research* **8**, 36-47. doi:10.1111/j.1755-0238.2002.tb00209.x

Hamzeh, S, Naseri, A A, AlaviPanah, S K, Bartholomeus, H, and Herold, M (2016). Assessing the accuracy of hyperspectral and multispectral satellite imagery for categorical and Quantitative mapping of salinity stress in sugarcane fields. *International Journal of Applied Earth Observation and Geoinformation* **52**, 412-421. doi:10.1016/j.jag.2016.06.024

Harrell, F E (2015). Describing, Resampling, Validating, and Simplifying the Model. *In* "Regression Modeling Strategies", pp. 103-126. Springer.

Heikkinen, V (2018). Spectral Reflectance Estimation Using Gaussian Processes and Combination Kernels. *IEEE Trans Image Process* **27**, 3358-3373. doi:10.1109/TIP.2018.2820839

Hernández-Clemente, R, Navarro-Cerrillo, R M, Suárez, L, Morales, F, and Zarco-Tejada, P J (2011). Assessing structural effects on PRI for stress detection in conifer forests. *Remote Sensing of Environment* **115**, 2360-2375. doi:10.1016/j.rse.2011.04.036

Hernández-Clemente, R, Navarro-Cerrillo, R M, and Zarco-Tejada, P J (2012). Carotenoid content estimation in a heterogeneous conifer forest using narrow-band indices and PROSPECT+DART simulations. *Remote Sensing of Environment* **127**, 298-315. doi:10.1016/j.rse.2012.09.014

Huang, R, and He, M (2005). Band Selection Based on Feature Weighting for Classification of Hyperspectral Data. *IEEE Geoscience and Remote Sensing Letters* **2**, 156-159. doi:10.1109/lgrs.2005.844658

Huang, Y, Chen, Z-x, Yu, T, Huang, X-z, and Gu, X-f (2018). Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture* **17**, 1915-1931. doi:10.1016/s2095-3119(17)61859-8

Huete, A R (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment* **25**, 295-309. doi:https://doi.org/10.1016/0034-4257(88)90106-X

Huete, A R, Liu, H Q, Batchily, K, and van Leeuwen, W (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment* **59**, 440-451. doi:https://doi.org/10.1016/S0034-4257(96)00112-5

Hungate, W, Watkins, R, and Borengasser, M (2008). Hyperspectral Remote Sensing; Principles and Applications. CRC Press.

Hunt, E R, Doraiswamy, P C, McMurtrey, J E, Daughtry, C S T, Perry, E M, and Akhmedov, B (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International Journal of Applied Earth Observation and Geoinformation* **21**, 103-112. doi:10.1016/j.jag.2012.07.020

INE (2018). "Estatísticas Agrícolas 2017," Lisboa, Portugal.

Jacquemoud, S, and Baret, F (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment* **34**, 75-91. doi:https://doi.org/10.1016/0034-4257(90)90100-Z

Jacquemoud, S, Verhoef, W, Baret, F, Bacour, C, Zarco-Tejada, P J, Asner, G P, François, C, and Ustin, S L (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment* **113**, S56-S66. doi:10.1016/j.rse.2008.01.026

Jones, H G, and Vaughan, R A (2010). "Remote sensing of vegetation: principles, techniques, and applications," Oxford University Press Inc., New York, USA.

Jordan, C F (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **50**, 663-666. doi:doi:10.2307/1936256

Katsoulas, N, Elvanidi, A, Ferentinos, K P, Kacira, M, Bartzanas, T, and Kittas, C (2016). Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review. *Biosystems Engineering* **151**, 374-398. doi:10.1016/j.biosystemseng.2016.10.003

Khanal, S, Fulton, J, and Shearer, S (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture* **139**, 22-32. doi:10.1016/j.compag.2017.05.001

Kim, M S, Daughtry, C, Chappelle, E, McMurtrey, J, and Walthall, C (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (A par).

Knipling, E B (1970). "Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation." doi:10.1016/S0034-4257(70)80021-9

Kuhn, M (2008). Building Predictive Models in R Using the caret Package. *2008* **28**, 26. doi:10.18637/jss.v028.i05

le Maire, G, François, C, and Dufrêne, E (2004). Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment* **89**, 1-28. doi:10.1016/j.rse.2003.09.004

le Maire, G, Francois, C, Soudani, K, Berveiller, D, Pontailler, J, Breda, N, Genet, H, Davi, H, and Dufrene, E (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment* **112**, 3846-3864. doi:10.1016/j.rse.2008.06.005

Lehnert, L W, Meyer, H, and Bendix, J (2017). hsdar: Manage, analyse and simulate hyperspectral data in R.

Li, D, Wang, C, Jiang, H, Peng, Z, Yang, J, Su, Y, Song, J, and Chen, S (2018). Monitoring litchi canopy foliar phosphorus content using hyperspectral data. *Computers and Electronics in Agriculture* **154**, 176-186. doi:10.1016/j.compag.2018.09.007

Lichtenthaler, H K, Lang, M, Sowinska, M, Heisel, F, and Miehé, J A (1996). Detection of Vegetation Stress Via a New High Resolution Fluorescence Imaging System. *Journal of Plant Physiology* **148**, 599-612. doi:10.1016/s0176-1617(96)80081-2

Lobell, D B, Asner, G P, Law, B E, and Treuhaft, R N (2001). Subpixel canopy cover estimation of coniferous forests in Oregon using SWIR imaging spectrometry. *Journal of Geophysical Research: Atmospheres* **106**, 5151-5160. doi:10.1029/2000jd900739

Maccioni, A, Agati, G, and Mazzinghi, P (2001). New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *Journal of Photochemistry and Photobiology B: Biology* **61**, 52-61. doi:https://doi.org/10.1016/S1011-1344(01)00145-2

Mariotto, I, Thenkabail, P S, Huete, A, Slonecker, E T, and Platonov, A (2013). Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. *Remote Sensing of Environment* **139**, 291-305. doi:10.1016/j.rse.2013.08.002

Merzlyak, M N, Gitelson, A A, Chivkunova, O B, and Rakitin, V Y (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum* **106**, 135-141. doi:doi:10.1034/j.1399-3054.1999.106119.x

Middleton, E M, Huemmrich, K F, Cheng, Y-B, and Margolis, H A (2012). Spectral Bioindicators of Photosynthetic Efficiency and Vegetation Stress. *In* "Hyperspectral Remote Sensing of Vegetation" (P. Thenkabail, J. Lyon and A. Huete, eds.), pp. 265-288. Taylor & Francis Group, LLC, Boca Raton.

Mulla, D J (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* **114**, 358-371. doi:10.1016/j.biosystemseng.2012.08.009

Oppelt, N, and Mauser, W (2004). Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *International Journal of Remote Sensing* **25**, 145-159. doi:10.1080/0143116031000115300

Orlandi, G, Calvini, R, Pigani, L, Foca, G, Vasile Simone, G, Antonelli, A, and Ulrici, A (2018). Electronic eye for the prediction of parameters related to grape ripening. *Talanta* **186**, 381-388. doi:10.1016/j.talanta.2018.04.076

Ortega-Farías, S, Ortega-Salazar, S, Poblete, T, Kilic, A, Allen, R, Poblete-Echeverría, C, Ahumada-Orellana, L, Zuñiga, M, and Sepúlveda, D (2016). Estimation of Energy Balance Components over a Drip-Irrigated Olive Orchard Using Thermal and Multispectral Cameras Placed on a Helicopter-Based Unmanned Aerial Vehicle (UAV). *Remote Sensing* **8**. doi:10.3390/rs8080638

Patrício, D I, and Rieder, R (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture* **153**, 69-81. doi:10.1016/j.compag.2018.08.001

Penubag (2012). Electromagnetic spectrum. *In* "Secondary Electromagnetic spectrum" (Secondary Penubag, ed.). Accessed in 19/09/2018 at https://et.wikipedia.org/wiki/Fail:Electromagnetic-Spectrum.svg

Peñuelas, J, Baret, F, and Filella, I (1995a). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* **31**, 221-230.

Peñuelas, J, Filella, I, Biel, C, Serrano, L, and Savé, T (1993). The reflectance at the 950-970 nm as an indicator of plant water status. *International Journal of Remote Sensing* **14**, 1887-1905.

Peñuelas, J, Filella, I, Lloret, P, Muñoz, F, and Vilajeliu, M (1995b). Reflectance assessment of mite effects on apple trees. *International Journal of Remote Sensing* **16**, 2727-2733. doi:10.1080/01431169508954588

Peñuelas, J, Gamon, J A, Fredeen, A L, Merino, J, and Field, C B (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment* **48**, 135-146. doi:https://doi.org/10.1016/0034-4257(94)90136-8

Peñuelas, J, Pinol, J, Ogaya, R, and Filella, I (1997). Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). *International Journal of Remote Sensing* **18**, 2869-2875. doi:10.1080/014311697217396

Petropoulus, G P (2013). Remote Sensing of Surface Turbulent Energy Fluxes. *In* "Remote Sensing of Energy Fluxes and Soil Moisture Content" (G. P. Petropoulus, ed.), pp. 49-84. CRC Press, Boca Raton, FL.

Poblete, T, Ortega-Farias, S, and Ryu, D (2018). Automatic Coregistration Algorithm to Remove Canopy Shaded Pixels in UAV-Borne Thermal Images to Improve the Estimation of Crop Water Stress Index of a Drip-Irrigated Cabernet Sauvignon Vineyard. *Sensors (Basel)* **18**. doi:10.3390/s18020397

Pôças, I, Cunha, M, Pereira, L S, and Allen, R G (2013). Using remote sensing energy balance and evapotranspiration to characterize montane landscape vegetation with focus on grass and pasture lands. *International Journal of Applied Earth Observation and Geoinformation* **21**, 159-172. doi:10.1016/j.jag.2012.08.017

Pôças, I, Gonçalves, J, Costa, P M, Gonçalves, I, Pereira, L S, and Cunha, M (2017). Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region. *International Journal of Applied Earth Observation and Geoinformation* **58**, 177-190. doi:10.1016/j.jag.2017.02.013

Pôças, I, Rodrigues, A, Gonçalves, S, Costa, P, Gonçalves, I, Pereira, L, and Cunha, M (2015). Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices. *Remote Sensing* **7**, 16460-16479. doi:10.3390/rs71215835

Qi, J, Chehbouni, A, Huete, A R, Kerr, Y H, and Sorooshian, S (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment* **48**, 119-126. doi:https://doi.org/10.1016/0034-4257(94)90134-1

R Core Team (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rapaport, T, Hochberg, U, Shoshany, M, Karnieli, A, and Rachmilevitch, S (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. *ISPRS Journal of Photogrammetry and Remote Sensing* **109**, 88-97. doi:10.1016/j.isprsjprs.2015.09.003

Ray, M, Ray, A, Dash, S, Mishra, A, Achary, K G, Nayak, S, and Singh, S (2017). Fungal disease detection in plants: Traditional assays, novel diagnostic techniques and biosensors. *Biosens Bioelectron* **87**, 708-723. doi:10.1016/j.bios.2016.09.032

Rivera, J, Verrelst, J, Delegido, J, Veroustraete, F, and Moreno, J (2014). On the Semi-Automatic Retrieval of Biophysical Parameters Based on Spectral Index Optimization. *Remote Sensing* **6**, 4927-4951. doi:10.3390/rs6064927

Rodríguez-Pérez, J R, Ordóñez, C, González-Fernández, A B, Sanz-Ablanedo, E, Valenciano, J B, and Marcelo, V (2018). Leaf water content estimation by functional

linear regression of field spectroscopy data. *Biosystems Engineering* **165**, 36-46. doi:10.1016/j.biosystemseng.2017.08.017

Rodríguez-Pérez, J R, Riaño, D, Carlisle, E, Ustin, S, and Smart, D R (2007). Evaluation of hyperspectral reflectance indexes to detect grapevine water status in vineyards. *American Journal of Enology and Viticulture* **58**, 302-317.

Rondeaux, G, Steven, M, and Baret, F (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment* **55**, 95-107. doi:https://doi.org/10.1016/0034-4257(95)00186-7

Roosjen, P P J, Bartholomeus, H M, and Clevers, J G P W (2015). Effects of soil moisture content on reflectance anisotropy — Laboratory goniometer measurements and RPV model inversions. *Remote Sensing of Environment* **170**, 229-238. doi:10.1016/j.rse.2015.09.022

Roujean, J-L, and Breon, F-M (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment* **51**, 375-384. doi:https://doi.org/10.1016/0034-4257(94)00114-3

Samborski, S M, Gozdowski, D, Stępień, M, Walsh, O S, and Leszczyńska, E (2016). On-farm evaluation of an active optical sensor performance for variable nitrogen application in winter wheat. *European Journal of Agronomy* **74**, 56-67. doi:10.1016/j.eja.2015.11.020

Schellberg, J, Hill, M J, Gerhards, R, Rothmund, M, and Braun, M (2008). Precision agriculture on grassland: Applications, perspectives and constraints. *European Journal of Agronomy* **29**, 59-71. doi:10.1016/j.eja.2008.05.005

Schmitt, R (2002). 1 - Introduction and Survey of the Electromagnetic Spectrum. *In* "Electromagnetics Explained" (R. Schmitt, ed.), pp. 1-24. Newnes, Burlington. doi:https://doi.org/10.1016/B978-075067403-4/50002-1

senseFly (2015). Filling The Gap (Infographic) – Drones vs Other Geospatial Data Sources. *In* "Secondary Filling The Gap (Infographic) – Drones vs Other Geospatial Data Sources" (Secondary senseFly, ed.). Accessed in 30/08/2018 at https://waypoint.sensefly.com/infographic-geospatial-data-collection-drones-satellite-manned/

SEOS (2018). Classification Algorithms and Methods. *In* "Secondary Classification Algorithms and Methods" (Secondary SEOS, ed.). Accessed in 10/09/2018 at http://www.seos-project.eu/modules/classification/classification-c01-p05.html

Sepulcre-Cantó, G, Zarco-Tejada, P, Jiménez-Muñoz, J, Sobrino, J, De Miguel, E, and Villalobos, F (2006). Detection of water stress in an olive orchard with thermal remote sensing imagery. *Agricultural and Forest Meteorology* **136**, 31-44.

Serrano, L, Peñuelas, J, and Ustin, S L (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment* **81**, 355-364. doi:https://doi.org/10.1016/S0034-4257(02)00011-1

Shellie, K, and Glenn, D M (2008). Wine Grape Response to Foliar Particle Film under Differing Levels of Preveraison Water Stress. *HORTSCIENCE* **43**, 1392–1397.

Shellie, K C, and King, B A (2013). Kaolin Particle Film and Water Deficit Influence Malbec Leaf and Berry Temperature, Pigments, and Photosynthesis. *American Journal of Enology and Viticulture* **64**, 223-230. doi:10.5344/ajev.2012.12115

Sims, D A, and Gamon, J A (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **81**, 337-354. doi:https://doi.org/10.1016/S0034-4257(02)00010-X

Smith, R, Adams, J, Stephens, D, and Hick, P (1995). Forecasting wheat yield in a Mediterranean-type environment from the NOAA satellite. *Australian Journal of Agricultural Research* **46**, 113-125. doi:https://doi.org/10.1071/AR9950113

Suárez, L, Zarco-Tejada, P J, Sepulcre-Cantó, G, Pérez-Priego, O, Miller, J R, Jiménez-Muñoz, J C, and Sobrino, J (2008). Assessing canopy PRI for water stress detection with diurnal airborne imagery. *Remote Sensing of Environment* **112**, 560-575. doi:10.1016/j.rse.2007.05.009

Thenkabail, P S, Smith, R B, and De Pauw, E (2000). Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sensing of Environment* **71**, 158-182. doi:https://doi.org/10.1016/S0034-4257(99)00067-X

Tian, Y-C, Gu, K-J, Chu, X, Yao, X, Cao, W-X, and Zhu, Y (2014). Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. *Plant and Soil* **376**, 193-209. doi:10.1007/s11104-013-1937-0

Tibshirani, R (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267-288.

Tucker, C J (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* **8**, 127-150. doi:https://doi.org/10.1016/0034-4257(79)90013-0

Usha, K, and Singh, B (2013). Potential applications of remote sensing in horticulture—A review. *Scientia Horticulturae* **153**, 71-83. doi:10.1016/j.scienta.2013.01.008

Veraverbeke, S, Dennison, P, Gitas, I, Hulley, G, Kalashnikova, O, Katagis, T, Kuai, L, Meng, R, Roberts, D, and Stavros, N (2018). Hyperspectral remote sensing of fire:

State-of-the-art and future perspectives. *Remote Sensing of Environment* **216**, 105-121. doi:10.1016/j.rse.2018.06.020

Verhoef, W (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment* **16**, 125-141. doi:https://doi.org/10.1016/0034-4257(84)90057-9

Verhoef, W (1985). Earth observation modeling based on layer scattering matrices. *Remote Sensing of Environment* **17**, 165-178. doi:https://doi.org/10.1016/0034-4257(85)90072-0

Verrelst, J, Camps-Valls, G, Muñoz-Marí, J, Rivera, J P, Veroustraete, F, Clevers, J G P W, and Moreno, J (2015a). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. *ISPRS Journal of Photogrammetry and Remote Sensing* **108**, 273-290. doi:10.1016/j.isprsjprs.2015.05.005

Verrelst, J, Malenovský, Z, Van der Tol, C, Camps-Valls, G, Gastellu-Etchegorry, J-P, Lewis, P, North, P, and Moreno, J (2018). Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surveys in Geophysics*. doi:10.1007/s10712-018-9478-y

Verrelst, J, Muñoz, J, Alonso, L, Delegido, J, Rivera, J P, Camps-Valls, G, and Moreno, J (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment* **118**, 127-139. doi:http://dx.doi.org/10.1016/j.rse.2011.11.002

Verrelst, J, Rivera, J, Alonso, L, and Moreno, J (2011). ARTMO: an Automated Radiative Transfer Models Operator toolbox for automated retrieval of biophysical parameters through model inversion. *In* "Proc. EARSeL 7th SIG-Imag. Spectrosc. Workshop", pp. 11-13.

Verrelst, J, Rivera, J P, van der Tol, C, Magnani, F, Mohammed, G, and Moreno, J (2015b). Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence? *Remote Sensing of Environment* **166**, 8-21. doi:10.1016/j.rse.2015.06.002

Verrelst, J, Rivera, J P, Veroustraete, F, Muñoz-Marí, J, Clevers, J G P W, Camps-Valls, G, and Moreno, J (2015c). Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison. *ISPRS Journal of Photogrammetry and Remote Sensing* **108**, 260-272. doi:10.1016/j.isprsjprs.2015.04.013

Vincini, M, Frazzi, E, and D'Alessio, P (2006). Angular dependence of maize and sugar beet VIs from directional CHRIS/Proba data. *In* "Proc. 4th ESA CHRIS PROBA Workshop", Vol. 2006, pp. 19-21.

Vogelmann, J E, Rock, B N, and Moss, D M (1993). Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing* **14**, 1563-1575. doi:10.1080/01431169308953986

Wang, M, Wan, Y, Ye, Z, Gao, X, and Lai, X (2017). A band selection method for airborne hyperspectral image based on chaotic binary coded gravitational search algorithm. *Neurocomputing*. doi:10.1016/j.neucom.2017.07.059

Wang, W, Yao, X, Yao, X, Tian, Y, Liu, X, Ni, J, Cao, W, and Zhu, Y (2012). Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crops Research* **129**, 90-98. doi:10.1016/j.fcr.2012.01.014

Wold, S, Esbensen, K, and Geladi, P (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* **2**, 37-52. doi:https://doi.org/10.1016/0169-7439(87)80084-9

Wu, C, Niu, Z, Tang, Q, and Huang, W (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology* **148**, 1230-1241. doi:10.1016/j.agrformet.2008.03.005

Wu, W (2014). The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sensing* **6**, 1211-1233. doi:10.3390/rs6021211

Zarco-Tejada, P, Hubbard, N, and Loudjani, P (2014). Precision Agriculture: An Opportunity for EU Farmers—Potential Support with the CAP 2014-2020. *Joint Research Centre (JRC) of the European Commission*.

Zarco-Tejada, P J, González-Dugo, V, Williams, L E, Suárez, L, Berni, J A J, Goldhamer, D, and Fereres, E (2013a). A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sensing of Environment* **138**, 38-50. doi:10.1016/j.rse.2013.07.024

Zarco-Tejada, P J, Guillén-Climent, M L, Hernández-Clemente, R, Catalina, A, González, M R, and Martín, P (2013b). Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). *Agricultural and Forest Meteorology* **171-172**, 281-294. doi:10.1016/j.agrformet.2012.12.013

Zarco-Tejada, P J, and Miller, J R (1999). Land cover mapping at BOREAS using red edge spectral parameters from CASI imagery. *Journal of Geophysical Research: Atmospheres* **104**, 27921-27933. doi:10.1029/1999jd900161

Zarco-Tejada, P J, Miller, J R, Mohammed, G H, Noland, T L, and Sampson, P H (2002). Vegetation Stress Detection through Chlorophyll + Estimation and Fluorescence

Effects on Hyperspectral Imagery. *Journal of Environment Quality* **31**. doi:10.2134/jeq2002.1433

Zarco-Tejada, P J, Pushnik, J C, Dobrowski, S, and Ustin, S L (2003). Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sensing of Environment* **84**, 283-294. doi:https://doi.org/10.1016/S0034-4257(02)00113-X

Zeng, C, Richardson, M, and King, D J (2017). The impacts of environmental variables on water reflectance measured using a lightweight unmanned aerial vehicle (UAV)-based spectrometer system. *ISPRS Journal of Photogrammetry and Remote Sensing* **130**, 217-230. doi:10.1016/j.isprsjprs.2017.06.004

zhang, C, Guo, C, Liu, F, Kong, W, He, Y, and Lou, B (2016). Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine. *Journal of Food Engineering* **179**, 11-18. doi:10.1016/j.jfoodeng.2016.01.002

Zhang, C, and Kovacs, J M (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture* **13**, 693-712. doi:10.1007/s11119-012-9274-5

Zhang, N, Wang, M, and Wang, N (2002). Precision agriculture—a worldwide overview. *Computers and electronics in agriculture* **36**, 113-132.

Attachments

Attachment 1. Vegetation indices

Name	Formula	Reference
ARI	$\frac{1}{R_{550}} - \frac{1}{R_{700}}$	(Gitelson et al., 2001)
Boochs	D_{703}	(Boochs et al., 1990)
Boochs2	D_{720}	(Boochs et al., 1990)
CARI	$a = (R_{700} - R_{550})/150$ $b = R_{550} - (a * 550)$ $R_{700} * \frac{abs(a * 670 + R_{670} + b)}{R_{670} * (a^2 + 1)^{0.5}}$	(Kim et al., 1994)
Carter	R_{695}/R_{420}	(Carter, 1994)
Carter2	R_{695}/R_{760}	(Carter, 1994)
Carter3	R_{605}/R_{760}	(Carter, 1994)
Carter4	R_{710}/R_{760}	(Carter, 1994)
Carter5	R_{695}/R_{670}	(Carter, 1994)
Carter6	R_{550}	(Carter, 1994)
CI	$R_{675} * R_{690}/R_{683}^2$	(Zarco-Tejada et al., 2003)
CI2	$\frac{R_{760}}{R_{700}} - 1$	(Gitelson et al., 2003)
ClAInt	$\int_{600nm}^{735nm} R$	(Oppelt and Mauser, 2004)
CRI1	$1/R_{515} - 1/R_{550}$	(Gitelson et al., 2003)
CRI2	$1/R_{515} - 1/R_{770}$	(Gitelson et al., 2003)
CRI3	$1/R_{515} - 1/R_{550} * R_{770}$	(Gitelson et al., 2003)
CRI4	$1/R_{515} - 1/R_{700} * R_{770}$	(Gitelson et al., 2003)
D1	D_{730}/D_{706}	(Zarco-Tejada et al., 2003)
D2	D_{705}/D_{722}	(Zarco-Tejada et al., 2003)
Datt	$(R_{850} - R_{710})/(R_{850} - R_{680})$	(Datt, 1999)
Datt2	R_{850}/R_{710}	(Datt, 1999)
Datt3	R_{754}/R_{704}	(Datt, 1999)
Datt4	$R_{672}/(R_{550} * R_{708})$	(Datt, 1998)
Datt5	R_{672}/R_{550}	(Datt, 1998)
Datt6	$(R_{860})/(R_{550} * R_{708})$	(Datt, 1998)

DD	$(R_{749} - R_{720}) - (R_{701} - R_{672})$	(le Maire et al., 2004)
DDn	$2 * (R_{710} - R_{660} - R_{760})$	(le Maire et al., 2008)
DPI	$(D_{688} * D_{710})/D_{697}^2$	(Zarco-Tejada et al., 2003)
DVI	$R_{800} - R_{675}$	(Jordan, 1969)
DWSI4	R_{550}/R_{680}	(Apan et al., 2004)
EGFN	$\frac{(\max(D_{650:750}) - \max(D_{500:550}))}{(\max(D_{650:750}) + \max(D_{500:550}))}$	(Peñuelas et al., 1994)
EGFR	$\max(D_{650:750}) / \max(D_{500:550})$	(Peñuelas et al., 1994)
EVI	$2.5 * \left(\frac{(R_{800} - R_{670})}{(R_{800} - (6 * R_{670}) - (7.5 * R_{475}) + 1)}\right)$	(Huete et al., 1997)
GDVI	$(R_{800}^n - R_{680}^n)/(R_{800}^n + R_{680}^n)^*$	(Wu, 2014)
GI	R_{554}/R_{667}	(Smith et al., 1995)
Gitelson	$1/R_{700}$	(Gitelson et al., 1999)
Gitelson2	$(R_{750} - R_{800}/R_{695} - R_{740}) - 1$	(Gitelson et al., 2003)
GMI1	R_{750}/R_{550}	(Gitelson et al., 2003)
GMI2	R_{750}/R_{700}	(Gitelson et al., 2003)
Green NDVI	$(R_{800} - R_{550})/(R_{800} + R_{550})$	(Gitelson et al., 1996)
LNC Tian	$(R_{605} - R_{521} - R_{682})/(R_{605} + R_{521} + R_{682})$	(Tian et al., 2014)
LNC Wang	$(R_{924} - R_{703} + 2 * R_{423})/(R_{924} + R_{703} - 2 * R_{423})$	(Wang et al., 2012)
Maccioni	$(R_{780} - R_{710})/(R_{780} + R_{710})$	(Maccioni et al., 2001)
MCARI	$\big((R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550})\big) * (R_{700}/R_{670})$	(Daughtry et al., 2000)
MCARI/ OSAVI	$MCARI/OSAVI$	(Daughtry et al., 2000)
MCARI2	$\big((R_{750} - R_{705}) - 0.2 * (R_{750} - R_{550})\big) * (R_{750}/R_{705})$	(Wu et al., 2008)
MCARI2/OS AVI2	$MCARI2/OSAVI2$	(Wu et al., 2008)
mND705	$(R_{750} - R_{705})/(R_{750} + R_{705} - 2 * R_{445})$	(Sims and Gamon, 2002)
mNDVI	$(R_{800} - R_{680})/(R_{800} + R_{680} - 2 * R_{445})$	(Sims and Gamon, 2002)
MPRI	$(R_{515} - R_{530})/(R_{515} + R_{530})$	(Hernández-Clemente et al., 2011)
mREIP	Red-edge inflection point using Gaussain fit	
MSAVI	$0.5 * (2 * R_{800} + 1 - ((2 * R_{800} + 1)^2 - 8 * (R_{800} - R_{670}))^2)$	(Qi et al., 1994)
mSR	$(R_{800} - R_{445})/(R_{680} - R_{445})$	(Sims and Gamon, 2002)

mSR2	$(R_{750}/R_{705}) - 1/(R_{750}/R_{705} + 1)^{0.5}$	(Chen, 1996)
mSR705	$(R_{750} - R_{445})/(R_{705} - R_{445})$	(Sims and Gamon, 2002)
MTCI	$(R_{754} - R_{709})/(R_{709} - R_{681})$	(Dash and Curran, 2004)
MTVI	$1.2 * (1.2 * (R_{800} - R_{550}) - 2.5 + (R_{670} - R_{550}))$	
NDVI	$(R_{800} - R_{680})/(R_{800} + R_{680})$	(Tucker, 1979)
NDVI2	$(R_{750} - R_{705})/(R_{750} + R_{705})$	(Gitelson and Merzlyak, 1994)
NDVI3	$(R_{682} - R_{553})/(R_{682} + R_{553})$	(Gandia et al., 2004)
NPCI	$(R_{680} - R_{430})/(R_{680} + R_{430})$	(Peñuelas et al., 1994)
OSAVI	$(1 + 0.16) * (R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	(Rondeaux et al., 1996)
OSAVI2	$(1 + 0.16) * (R_{750} - R_{705})/(R_{750} + R_{705} + 0.16)$	(Wu et al., 2008)
PARS	R_{746}/R_{513}	(Chappelle et al., 1992)
PRI	$(R_{531} - R_{570})/(R_{531} + R_{570})$	(Gamon et al., 1992)
PRI*CI2	$PRI * CI2$	(Garrity et al., 2011)
PRI_norm	$PRI * (-1)/(RDVI * R_{700}/R_{670})$	(Zarco-Tejada et al., 2013a)
PSND	$(R_{800} - R_{470})/(R_{800} - R_{470})$	(Blackburn, 1998)
PSRI	$(R_{678} - R_{500})/R_{750}$	(Merzlyak et al., 1999)
PSSR	R_{800}/R_{635}	(Blackburn, 1998)
RDVI	$(R_{800} - R_{670})/\sqrt{R_{800} + R_{670}}$	(Roujean and Breon, 1995)
REP_LE	Red-edge position through linear extrapolation	
REP_Li	$R_{re} = (R_{670} + R_{780})/2$ $700 + 40 * (\frac{R_{re} - R_{700}}{R_{740} - R_{700}})$	(Guyot and Baret, 1988)
SAVI	$(1 + L) * (R_{800} - R_{670})/(R_{800} + R_{670} + L)$	(Huete, 1988)
SIPI	$(R_{800} - R_{445})/(R_{800} - R_{680})$	(Peñuelas et al., 1995a; Peñuelas et al., 1995b)
SPVI	$0.4 * 3.7 * (R_{800} - R_{670}) - 1.2 * ((R_{530} - R_{670})^2)^{0.5}$	(Vincini et al., 2006)
SR	R_{800}/R_{680}	(Jordan, 1969)
SR1	R_{750}/R_{700}	(Gitelson and Merzlyak, 1997)
SR2	R_{752}/R_{690}	(Gitelson and Merzlyak, 1997)
SR3	R_{750}/R_{550}	(Gitelson and Merzlyak, 1997)

SR4	R_{700}/R_{670}	
SR5	R_{675}/R_{700}	(Chappelle et al., 1992)
SR6	R_{750}/R_{710}	(Zarco-Tejada and Miller, 1999)
SR7	R_{440}/R_{690}	(Lichtenthaler et al., 1996)
SR8	R_{515}/R_{550}	(Hernández-Clemente et al., 2012)
SRPI	R_{430}/R_{680}	(Peñuelas et al., 1995a)
SRWI	R_{850}/R_{1240}	(Zarco-Tejada et al., 2003)
Sum_Dr1	$\sum_{i=626}^{795} D1_i$	(Elvidge and Chen, 1995)
Sum_Dr2	$\sum_{i=680}^{780} D1_i$	(Filella and Peñuelas, 1994)
TCARI	$3 * ((R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550}) * (R_{700}/R_{670}))$	(Haboudane et al., 2002)
TCARI/OSAVI	$TCARI/OSAVI$	(Haboudane et al., 2002)
TCARI2	$3 * ((R_{750} - R_{705}) - 0.2 * (R_{750} - R_{550}) * (R_{750}/R_{705}))$	(Wu et al., 2008)
TCARI2/OSAVI2	$TCARI2/OSAVI2$	(Wu et al., 2008)
TGI	$-0.5 * (190 * (R_{670} - R_{550}) - 120 * (R_{670} - R_{480}))$	(Hunt et al., 2013)
TVI	$0.5 * (120 * (R_{750} - R_{550}) - 200 * (R_{670} - R_{550}))$	(Broge and Leblanc, 2001)
Vogelmann	R_{740}/R_{720}	(Vogelmann et al., 1993)
Vogelmann2	$(R_{734} - R_{747})/(R_{715} + R_{726})$	(Vogelmann et al., 1993)
Vogelmann3	D_{715}/D_{705}	(Vogelmann et al., 1993)
Vogelmann4	$(R_{734} - R_{747})/(R_{715} + R_{720})$	(Vogelmann et al., 1993)
WI	R_{900}/R_{970}	(Peñuelas et al., 1997)

*For GDVI n must be defined appending an underscore and the intended exponent to the index