



Universidad de Valladolid

**PROGRAMA DE DOCTORADO EN CIENCIA E INGENIERÍA
AGROALIMENTARIA Y DE BIOSISTEMAS**

TESIS DOCTORAL

**Applications of artificial neural networks in
three agro-environmental systems: microalgae
production, nutritional characterization of soils
and meteorological variables management**

Presentada por Blas Manuel Franco Ortellado
para optar al grado de
Doctor por la Universidad de Valladolid

Dirigida por:
Dr. Luis Manuel Navas Gracia
Dr. Luis Hernández Callejo

Passion is what gets you through the hardest times that might otherwise make strong men weak, or make you give up.

Neil deGrasse Tyson

Acknowledgments

This research and the thesis development were possible thanks to the funding from the Spanish Ministry of Education and Science via the predoctoral scholarships “*Formación de Profesorado Universitario*” (FPU) that was conceived to me [grant number FPU15/01707] and the project LIFE+ Integral Carbon, as part of the LIFE program and financial instrument of the European Union [grant number ENV/ES/001251].

This thesis is the result of a long journey and effort of many people on both sides of the world. Thanks to all and every single person from Paraguay (ñane reta) and Spain for this collective achievement, without the spiritual support during these four years this work would not be possible.

Firstly, thanks to all my family, parents, Aurea and Blas, brothers, Paulo, José and Celeste, grandparents, uncles, cousins, and those families and friends that the blood does not relate us, but the souls do it in such a strong bond. Thanks for all you shared with me, happy moments (and sad moments, I’m sure there are, but I do not even remember them), knowledge, wisdom, love, and faith in God.

Thanks to the University of Valladolid for being the house where the knowledge and creativity expressed in the past few years. Thanks to the directors of this work, Luis Manuel Navas and Luis Hernández for the guide and support to accomplish this thesis. Thanks to Matilde and Oscar for the support in the last months for the writing of the thesis.

Thanks to all the friends who were the closest during the stay in this strange but familiar country of what is known as the old continent. Thanks to Anais, M^a. Ángeles, Charo, Ángela R., Alberto, Carmen, Bea, Laura, Marta S., Ángela B., Vicky, Felipe, Montse, and Maria; to the newest friends in the last part of this journey Ángel, José, Marta J., Shabnam, Miguel Angel, Alfredo, Jorge and Mario.

A special acknowledgment for those who were not just good friends but also terrific teachers in these years: Jorge, Gonzalo, Víctor and Michel. Once again, it was shown that the humblest people are always the ones whose more good things have to share with you. Thanks for being on the train in which we shared fragments of the trip but had different arrival stations.

For all the people in Spain, maybe I am wrong with this but, a song from The Beatles comes into my head when I think about all this — the lyrics sing “... *you and I have memories longer than the road that stretches out ahead...*”. If we are lucky enough, life will give us other chances to laugh together again.

Ipahápe, Ñandéjárape ñande jajerovia ha mba'eichapa ñanembo'e Ñandéjára Ñe'ẽ: Upéicharõ, peñemomirĩkena ha peñemoĩ Ñandéjára ipu'akapáva poguýpe, ikatu hañuáicha ha'e penembotuicha oñuahẽ vove pe tiémpo (1 Peter 5:6). Aguije Tupã.

INDEX OF CONTENTS

Index of Contents

| | |
|--|-------------|
| Acknowledgments | v |
| Index of Contents | ix |
| List of Tables | xiii |
| List of Figures..... | xv |
| List of Equations | xix |
| List of Acronyms | xxi |
| Abstract/Resumen..... | xxv |
| 1. Introduction..... | 3 |
| <i>1.1 Research problems and hypothesis.....</i> | <i>4</i> |
| <i>1.2 Objectives.....</i> | <i>4</i> |
| 1.2.1 General objective | 4 |
| 1.2.2 Specific objectives | 5 |
| <i>1.3 Methodology overview</i> | <i>6</i> |
| <i>1.4 Motivation for the research</i> | <i>6</i> |
| <i>1.5 Innovative aspects of the research</i> | <i>7</i> |
| <i>1.6 Thesis structure</i> | <i>8</i> |
| 2. Theoretical Background | 11 |
| <i>2.1 Challenges in agriculture and new technologies.....</i> | <i>11</i> |
| <i>2.2 Microalgae culture.....</i> | <i>19</i> |
| 2.2.1 Microalgae overview | 19 |
| 2.2.2 Specificity in microalgae production systems | 21 |
| <i>2.3 Agricultural soil fertility.....</i> | <i>24</i> |
| 2.3.1 Machine vision and color theory..... | 24 |
| 2.3.1.1 Computer vision algorithms | 25 |
| 2.3.1.2 Color theory and systems..... | 27 |
| 2.3.2 Soil color and fertility properties | 30 |
| <i>2.4 Agricultural meteorology data.....</i> | <i>33</i> |

| | |
|--|-----------|
| 2.4.1 Meteorological information and agricultural importance | 33 |
| 2.4.2 Meteorological data estimation..... | 34 |
| 2.5 Artificial neural networks..... | 36 |
| 2.5.1 The artificial neuron | 36 |
| 2.5.2 The artificial neural network | 39 |
| 2.5.3 Activation functions | 43 |
| 2.5.3.1 Hardlim function | 43 |
| 2.5.3.2 Sigmoid function | 45 |
| 2.5.3.3 Hyperbolic tangent function | 46 |
| 2.5.3.4 Softsign function | 47 |
| 2.5.3.5 Softmax function | 48 |
| 2.5.3.6 Rectified linear unit function..... | 48 |
| 2.5.4 Phases of artificial neural networks | 49 |
| 2.5.4.1 Data preprocessing | 50 |
| 2.5.4.2 Learning phase | 51 |
| 2.5.4.3 Validation phase..... | 56 |
| 3. Materials and Methods..... | 63 |
| 3.1 <i>Monoalgal and mixed algal cultures discrimination by using artificial neural network</i> | 63 |
| 3.1.1 Microalgae species and growth conditions..... | 64 |
| 3.1.2 Light absorption measurements | 65 |
| 3.1.3 Artificial neural network and data analysis..... | 66 |
| 3.2 <i>Artificial neural network for soil color analysis and characterization.....</i> | 68 |
| 3.2.1 Study area and soil sampling procedure | 69 |
| 3.2.2 Image acquisition system..... | 71 |
| 3.2.3 Computer vision software for color conversion..... | 72 |
| 3.2.4 Munsell soil color classification and description..... | 74 |
| 3.2.5 Soil properties characterization by artificial neural networks | 74 |
| 3.3 <i>Virtual weather stations for meteorological data estimations.....</i> | 76 |
| 3.3.1 Meteorological dataset | 77 |
| 3.3.2 Dataset statistical analysis | 78 |
| 3.3.3 Virtual weather station development..... | 78 |
| 3.3.4 Interpolation methods | 79 |
| 4. Results and Discussion | 83 |
| 4.1 <i>Monoalgal and mixed algal cultures discrimination by using artificial neural network</i> | 83 |
| 4.1.1 Microalgae light absorption | 83 |

| | |
|--|------------|
| 4.1.2 Artificial neural network culture discrimination | 86 |
| 4.1.3 Final Remarks | 91 |
| <i>4.2 Artificial neural network for soil color analysis and characterization</i> | <i>92</i> |
| 4.2.1 Computer vision software for color conversion..... | 92 |
| 4.2.2 Soil samples description and color classification | 96 |
| 4.2.3 Characterization of soil fertility parameters through artificial neural networks . | 100 |
| 4.2.4 Final remarks..... | 103 |
| <i>4.3 Virtual weather stations for meteorological data estimations.....</i> | <i>104</i> |
| 4.3.1 Statistical summary of the dataset | 104 |
| 4.3.2 Interpolation methods comparison | 106 |
| 4.3.3 Season effect in data interpolation quality..... | 111 |
| 4.3.4 Virtual weather station | 116 |
| 4.3.5 Final Remarks | 116 |
| 5. Conclusions | 121 |
| 5.1 General conclusions | 121 |
| 5.2 Monoalgal and mixed algal cultures discrimination by using artificial neural network conclusions | 122 |
| 5.3 Artificial neural network for soil color analysis and characterization conclusions | 122 |
| 5.4 Virtual weather stations for meteorological data estimations conclusions..... | 123 |
| 6. Future work | 127 |
| 7. References | 131 |
| Appendix | 167 |
| Appendix A | 169 |
| Appendix B | 171 |
| Appendix C | 173 |

List of Tables

| | |
|---|-----|
| Table 1. Comparison of the results from the experimental and the artificial neural network output of the validation experiment. | 87 |
| Table 2. Statistical summary of fertility parameters in soil dataset. | 97 |
| Table 3. Means comparison of fertility parameters by Munsell color hue. Means with the same letter are not significantly different from each other (Kruskal-Wallis H test followed by the Mann-Whitney U test). | 99 |
| Table 4. Regression analysis between actual and predicted values of soil fertility parameters using RGB and L*a*b* for ANNs. | 100 |
| Table 5. Statistical summary of meteorological observations dataset (n = 18,234). | 104 |
| Table 6. Statistical summary of meteorological observations during the summer months and the winter months. | 105 |
| Table 7. Analysis of meteorological data interpolations results for the ANNs approach. | 108 |
| Table 8. Analysis of meteorological data interpolation results for alternative methods. | 109 |
| Table 9. Analysis of meteorological data interpolations results for the ANNs approach for summer months. | 112 |
| Table 10. Analysis of meteorological data interpolation results for the alternative methods for summer months. | 113 |
| Table 11. Analysis of meteorological data interpolations results for the ANNs approach for winter months. | 114 |
| Table 12. Analysis of meteorological data interpolation results for the alternative methods for winter months. | 115 |
| Table 13. X-Rite ColorChecker Classic L*a*b* color information following the enumeration presented in Figure 35. | 173 |

List of Figures

| | |
|---|----|
| Figure 1. CO ₂ emissions estimations for the 1990s, 2000s and 2010 and projections for 2030 and 2050 for agricultural related activities (elaborate from data of Tubiello <i>et al.</i> , 2014). | 12 |
| Figure 2. Climate-smart agriculture foundations for the development of sustainable agriculture. | 13 |
| Figure 3. Microalgae production system in an open-culture system (a) and a closed-culture system (b). | 20 |
| Figure 4. RGB color system diagram. Purple, amber and light blue color RGB values are described. | 27 |
| Figure 5. CIELAB color space representation. Purple, amber and light blue L*a*b* colors are described. | 29 |
| Figure 6. Schematic diagram of the Munsell color system for a given value (5). The hues are placed circumferentially and the chroma scale extends outward perpendicular to the center. | 31 |
| Figure 7. Biological neuron cell scheme. The signal travels from left to right, from the dendrites (input), passing through the body or soma and going out through the axon and its terminals. | 37 |
| Figure 8. McCulloch and Pitts artificial neuron. Where x_n = inputs and w_n = weights. The sum of products of inputs and weights plus the bias is passed to the activation function (σ) to generate the neuron output (\hat{y}). | 38 |
| Figure 9. ANN scheme. Where x_n are the input variables that enter in the input layer and \hat{y}_n are the outputs from the output layer. The hidden layer contains the activation function. | 40 |
| Figure 10. Hardlim activation function plot. | 44 |
| Figure 11. Sigmoid activation function plot. | 45 |
| Figure 12. Hyperbolic tangent activation function plot. | 46 |
| Figure 13. Softsign activation function plot. | 47 |
| Figure 14. Rectified linear unit (relu) activation function plot. | 49 |
| Figure 15. Example of the error function with respect to the weight gradient during the backpropagation algorithm execution. | 55 |
| Figure 16. Microalgal and mixed algal cultures discrimination by ANN experiment scheme. | 63 |
| Figure 17. Microalgae cultures collection of “Las Palmerillas” Research Center. ... | 64 |

| | |
|--|-----|
| Figure 18. Microalgae suspension for monoalgal and mixed algal cultures (a); and sample for light absorption measure in the colorimeter (b). | 65 |
| Figure 19. ANN architecture schema, with 31 neurons in the input layer, 45 in the hidden layer and 4 in the output layer. | 67 |
| Figure 20. Soil color analysis and characterization by ANN experiment scheme.... | 68 |
| Figure 21. ECa raster map from the Veris data acquisition system at 36 cm (a) and 90 cm (b) depth; Segments for soil sampling generated through the processed raster data (c). | 70 |
| Figure 22. Image acquisition system; light and camera setup (a) and shooter camera app for smartphone (b)..... | 71 |
| Figure 23. X-Rite ColorChecker Classic, color calibration card. | 72 |
| Figure 24. ANN architecture schema, with 3 neurons in the input layer, 8 in the hidden layer and 3 in the output layer. | 73 |
| Figure 25. VWS development methodology scheme. | 76 |
| Figure 26. Meteorological station network of the InfoRiego program of the Agrarian Technology Institute of Castilla y León. | 77 |
| Figure 27. Light absorption spectrum of monoalgal cultures as obtained from the colorimeter (left), and relative absorbance conversion (right). | 85 |
| Figure 28. Correlation between the experimental composition of the samples and that predicted by the artificial neural network developed for the samples contained in Table 1..... | 90 |
| Figure 29. The user interface of DigiCIELAB CV software with the “Color Calibration” option (a) to generate the ANN model for the color analysis, and the “Colorimeter” option (b) for the color measurement of samples with the calibrated model..... | 93 |
| Figure 30. DigiCIELAB software workflow scheme. The part "a" of the scheme corresponds to the “Color Calibration” option, and the part "b" is regarding the “Colorimeter” option. | 94 |
| Figure 31. CV process to detect the object of study and ROI in the digital photograph: Original image (a), hue plane extraction (b), binary image (c), filled image (d), Petri plate detection (e), ROI for soil color extraction in the image center (f). | 95 |
| Figure 32. Munsell color classification for soil samples dataset. | 98 |
| Figure 33. Soils samples with their respective lightness (L*) and OM (organic matter) contents (a); and samples with their respective red component (a* and R) and Fe (iron) contents (b)..... | 102 |
| Figure 34. Boxplots and mean comparison of meteorological observations grouped by the season of the year..... | 107 |

Figure 35. X-Rite ColorChecker Classic color charts enumerated..... 173

List of Equations

| | |
|--|----|
| Eq. 1. Perceptron mathematical model..... | 38 |
| Eq. 2. Hardlim activation function equation..... | 44 |
| Eq. 3. Sigmoid activation function equation..... | 45 |
| Eq. 4. Hyperbolic tangent activation function equation..... | 46 |
| Eq. 5. Softsign activation function equation..... | 47 |
| Eq. 6. Softmax activation function equation. | 48 |
| Eq. 7. Rectified linear unit activation function equation..... | 48 |
| Eq. 8. Z-score normalization formula..... | 51 |
| Eq. 9. Min-Max normalization formula..... | 51 |
| Eq. 10. Chain rule for the backpropagation algorithm. | 53 |
| Eq. 11. Mean square error formula. | 57 |
| Eq. 12. Root mean square of error formula..... | 57 |
| Eq. 13. Mean absolute error formula. | 57 |
| Eq. 14. Mean absolute percentage error formula. | 58 |
| Eq. 15. Cross-entropy error formula..... | 58 |
| Eq. 16. Correlation coefficient formula. | 58 |
| Eq. 17. Determination coefficient formula. | 59 |
| Eq. 18. Relative frequencies conversion formula. | 66 |
| Eq. 19. Minimum Euclidean distance match algorithm. | 74 |
| Eq. 20. Inverse distance weighted interpolation method formula..... | 80 |

List of Acronyms

| | |
|-----------------|--|
| ACE | Average cross-entropy |
| Al | Aluminum |
| ANN | Artificial neural networks |
| Ca | Calcium |
| CE | Cross-entropy error |
| CH ₄ | Methane |
| CIE | Commission Internationale de l'Eclairage (International Commission on Illumination) |
| CIELAB | Device independent color system; where L* represents lightness; a* the redness/greenness axis; and value b* the yellowness/blueness axis |
| CMOS | Complementary metal-oxide semiconductor |
| CO ₂ | Carbon dioxide |
| CPU | Central processing unit |
| CSA | Climate-smart agriculture |
| Cu | Copper |
| CV | Computer vision |
| DNN | Deep neural network |
| EC | Electrical conductivity |
| ECa | Apparent electrical conductivity |
| EDA | Exploratory data analysis |
| ETo | Reference evapotranspiration |
| Fe | Iron |
| FTP | File transfer protocol |
| GHG | Greenhouse gases |
| GPU | Graphics processing unit |
| HSI | Hue, saturation and intensity |
| HSL | Hue, saturation and luminance |
| HSV | Hue, saturation and value |
| IDW | Inverse distance weighted |
| ISDW | Inverse square distance weighted |
| ISO | International organization of standardization |
| K | Potassium |
| L*a*b* | Lightness, redness/greenness and yellowness/blueness axis in the CIELAB color system |
| MAE | Mean absolute error |
| MAPE | Mean absolute percentage error |

| | |
|------------------|---|
| Mg | Magnesium |
| ML | Machine learning |
| MLP | Multilayer perceptron |
| Mn | Manganese |
| MRL | Multiple linear regression |
| MSE | Mean square error |
| MV | Machine vision |
| N | Nitrogen |
| Na | Sodium |
| NH ₃ | Ammonia |
| NO ₂ | Nitrous oxide |
| NO ⁻³ | Nitrate |
| OM | Organic matter |
| P | Phosphorous |
| pH | Potential of hydrogen |
| Precip | Precipitation |
| R | Correlation coefficient |
| R ² | Determination coefficient |
| RAM | Random access memory |
| relu | Rectified linear unit |
| RFR | Random forest regression |
| RGB | Device dependent color system based on the red, green and blue components |
| RH | Relative humidity |
| RMSE | Root mean squared error |
| RNN | Recurrent neural network |
| ROI | Region of interest |
| ROI | Regions of interest |
| SOC | Soil organic carbon |
| St. dev. | Standard deviation |
| tanh | Hyperbolic tangent |
| Temp | Temperature |
| TSI | Total solar irradiation |
| vis-NIR | Visible–near infrared |
| VWS | Virtual weather stations |
| w/v | The weight volume ratio in g of soil per mL of water |
| WS | Wind speed |
| XOR | Exclusive OR logic problem |

ABSTRACT

Abstract

Agriculture is an essential human activity, highly dependent on meteorological conditions and focus of research and innovation to confront several challenges. Climate change, global warming, and the degradation of agricultural ecosystems are just a few of the problems that humans are facing for continuing the essential food production. Seeking the innovation in the agricultural sector, three main research topics were considered for this thesis; such as microalgae production, soil color and fertility, and meteorological data acquisition.

These subjects have increasing roles in agriculture, specifically under the uncertainty in the future of food production. Microalgae are a healthy alternative for crops fertilization and soil sustainability; while the soil fertility parameters need to be more studied to aim lower cost and faster analysis methods to help the management. Agriculture, as a highly weather-dependent activity, needs meteorological data to anticipate events, planning, and management crops in an efficient mode. These topics were selected with the purpose to improve the current state of the art, propose new alternatives based, mainly, in the application of artificial neural networks (ANNs) as a novel manner to solve the problems and generate knowledge of direct application in crop systems. ANNs are a useful tool to modeling and solve complex nonlinear problems; they are a mathematical model of the animal brain and their ability to deal with complicated issues drive the scientific community to use them to find solutions hardly found with other techniques.

The main objective of this thesis was to generate ANN models capable of addressing agricultural related problems as an alternative to traditional and more expensive methods for management, analysis, and data acquisition in the crop systems. For the microalgae culture experiments, monoalgal and mixed algal culture spectral signatures from light absorption measurements were analyzed. Additionally, an ANN was used alongside the spectral signature in order to create a model capable of classifying the microalgae cultures and determine the species present in suspension. The results show that the ANN was capable of distinguishing between monoalgal and mixed algal cultures, identifying the microalgae species in the monoalgal cultures and providing the approximate composition of mixed algal cultures.

Regarding the soil study, soil samples were classified according to the Munsell color notation, and the obtained hues were used to group soils and perform statistical analysis over their fertility attributes. In addition, RGB and $L^*a^*b^*$ colors

were used as input in ANNs to create models to describe the soil fertility parameters. The RGB and the L*a*b* colors were obtained from digital color photographs and computer software programmed to perform quick and accurate measurements. The results showed that the soil aggrupation based on the Munsell color hue resulted not efficient to separate samples according to fertility levels and to regard the ANNs approach to describe soils, the obtained models were not capable of providing accurate results.

Finally, for the meteorological data acquisition study, daily data from a weather station network were used to perform interpolations with several methods, including traditional techniques and ANNs, and the results were evaluated to compare the accuracy of the methods. With these interpolations algorithms, the development of virtual weather stations (VWS) was proposed. A script code to automatically acquire, process and interpolate meteorological data was made to perform the estimations in different locations where weather stations are not available. The results showed that the interpolation methods have high accuracies — decreasing slightly during summer and winter. From a practical perspective, the methods here described could be useful to produce meteorological time series data with the VWS.

In conclusion, ANNs have proved to be a powerful tool to solve classifications, estimations and predictions problems. The use of ANN in agricultural related issues is a critical step to find solutions for problems and help the workers to make faster and better decisions in the productive chain. For instance, in microalgae commercial production, the development method would provide a fast and reliable tool for managing. In soil fertility, ANNs still have a road to follow to find accurate models; this research gave light for new directions to take in this matter. Moreover, regarding the meteorological data acquisition, the VWS can help farmers to obtain data of importance for cultures and to plan more efficient irrigations.

Resumen

La agricultura es una actividad esencial para los humanos, es altamente dependiente de las condiciones meteorológicas y foco de investigación e innovación con el objetivo de enfrentar diversos desafíos. El cambio climático, calentamiento global y la degradación de los ecosistemas agrícolas son sólo algunos de los problemas que los humanos enfrentamos para continuar con la esencial producción de alimentos. Buscando la innovación en el sector agrícola, se consideraron tres tópicos principales de investigación para esta tesis; la producción de microalgas, el color del suelo y la fertilidad, y la adquisición de datos meteorológicos.

Estos temas tienen roles cada vez más importantes en la agricultura, especialmente bajo la incertidumbre del futuro de la producción de alimentos. Las microalgas son una interesante alternativa para la fertilización de cultivos y la sostenibilidad del suelo; mientras que los parámetros de fertilidad del suelo necesitan ser más estudiados para desarrollar métodos de análisis de menor costo y más rápidos para ayudar al manejo. La agricultura, como actividad altamente dependiente del clima, necesita de datos meteorológicos para anticipar eventos, planificar y manejar los cultivos eficientemente. Estos temas se seleccionaron con el propósito de mejorar el estado actual de la técnica, proponer nuevas alternativas basadas, principalmente, en la aplicación de redes neuronales artificiales (ANN) como una manera novedosa de resolver los problemas y generar conocimiento de aplicación directa en sistemas de cultivos. Las ANNs son una herramienta útil para modelar y resolver problemas complejos no lineales; son un modelo matemático del cerebro de los animales y su capacidad para tratar problemas complicados lleva a la comunidad científica a usarlas para encontrar soluciones que difícilmente otras técnicas pueden resolver.

El objetivo principal de esta tesis fue generar modelos de ANNs capaces de abordar problemas relacionados con la agricultura, como una alternativa a los métodos tradicionales y más costosos empleados en el manejo, análisis y adquisición de datos en los sistemas agrarios. Para los experimentos con cultivos de microalgas, se analizaron las huellas espectrales provenientes de mediciones de la absorción de luz de suspensiones monoalgales y cultivos mixtos. Además, se utilizó una ANN junto con las huellas espectrales para crear un modelo capaz de clasificar los cultivos de microalgas y determinar las especies presentes en suspensión. Los resultados muestran que la ANN fue capaz de distinguir entre cultivos monoalgales

y mixtos, identificando las especies de microalgas en los cultivos de monoalgales y proporcionando una composición aproximada en los cultivos mixtos.

Con respecto al estudio del suelo, muestras de suelo se clasificaron según la notación de color Munsell y los tonos obtenidos se utilizaron para agrupar los suelos y realizar análisis estadísticos sobre sus atributos de fertilidad. Además, los colores RGB y $L^*a^*b^*$ se utilizaron como entrada en ANNs para crear modelos que describan los parámetros de fertilidad del suelo. Los colores RGB y $L^*a^*b^*$ se obtuvieron a partir de fotografías digitales a color y una aplicación informática programada para realizar mediciones rápidas y precisas. Los resultados mostraron que la agrupación de suelos basada en el tono del color de Munsell no resultó eficiente para separar las muestras según niveles de fertilidad y respecto al enfoque de las ANNs para describir los suelos, los modelos obtenidos no fueron capaces de proporcionar resultados precisos.

Finalmente, para el estudio de adquisición de datos meteorológicos, los datos diarios de una red de estaciones meteorológicas se utilizaron para realizar interpolaciones con varios métodos, incluidas las técnicas tradicionales y ANNs, y los resultados fueron evaluados para comparar la precisión de los métodos. Con estos algoritmos de interpolación, se propuso el desarrollo de estaciones meteorológicas virtuales (VWS). Se programó un código para adquirir, procesar e interpolar automáticamente datos meteorológicos para realizar estimaciones en diferentes lugares sin disponibilidad de estaciones meteorológicas. Los resultados mostraron que los métodos de interpolación tienen altas precisiones — estas disminuyen ligeramente durante el verano y el invierno. Desde una perspectiva práctica, los métodos aquí descritos podrían ser útiles para producir series de datos meteorológicos con las VWS.

En conclusión, las ANNs han demostrado ser una herramienta poderosa para resolver problemas de clasificación, estimación y predicción. El uso de ANNs en temas relacionados con la agricultura es un paso crítico para encontrar soluciones a los problemas y ayudar a los productores a tomar decisiones más rápidas y mejores durante los procesos productivos. Por ejemplo, en la producción comercial de microalgas, el método desarrollado proporcionaría una herramienta rápida y confiable para la gestión. En la fertilidad del suelo, las ANNs aún tienen un camino que seguir para encontrar modelos precisos; esta investigación dio luz a nuevas direcciones a tomar en este asunto. Y con respecto a la adquisición de datos meteorológicos, las VWS pueden ayudar a los agricultores a obtener datos de importancia para los cultivos y para planificar un riego más eficiente.

INTRODUCTION

1. Introduction

Artificial neural networks (ANNs) have emerged in the last decades as a wide area of new opportunities for academic research and applications in many fields such as remote sensing, image classification, speech recognition, economics, robotics, automotive engineering, web-informatics, bio-informatics, intelligent medical diagnosing and treatment and so on. All these applications aim to solve plenty of real-world problems employing the ANNs high capacity to find solutions to non-linear and complex problems (Cheng *et al.*, 2018; Kakuda *et al.*, 2019; Li *et al.*, 2018). ANNs create new chances for traditional issues and new technologies related matters by giving new approaches to improve techniques and model accuracies through the capacity of ANNs of mapping input variables with a given target value, this ability has been used in several studies to make predictions, classifications and estimations with higher precisions in comparison with traditional techniques (Abrishami *et al.*, 2018; Alanis, 2018; Amrouche and Le Pivert, 2014; Tealab *et al.*, 2017; Tkáč and Verner, 2016).

The present work is a compendium of research in agricultural related issues applying ANNs to solve problems. Therefore, three main research topics were considered, such as microalgae production, soil color and fertility, and meteorological data. These represent essential roles in crops system, specifically in the future of food production, global warming, and sustainability. Microalgae are a healthy alternative for crops fertilization and soil fertility sustainability, while the soil fertility parameters need to be more studied to aim lower cost and faster analysis methods to help the management. Last but not least, agriculture is a highly weather-dependent activity and needs meteorological data to anticipate events, planning and management crops in an efficient mode.

These topics were selected with the purpose to improve the current state of the art, propose new alternatives based, mainly, in the application of ANNs as a novel manner to solve the problems and generate knowledge of direct application in crop systems. In the following sections, the research problems and the hypothesis (section 1.1), objectives (section 1.2) and a summary of the methodology used (section 1.3) on this thesis will be detailed. Furthermore, the motivation (section 1.4) and innovative aspect (section 1.5) of this research, and an overview of the thesis structure (section 1.6) are also presented in this chapter.

1.1 Research problems and hypothesis

The research topics, as mentioned earlier, microalgae production, soil color and fertility, and meteorological data are a niche for innovation, as other fields in the agricultural sector. This innovation is a crucial concern due to the fact of the challenges for the agricultural sector in the present and future scenarios.

Considering this, ANNs were applied in this research to solve issues in the territory of these topics; in the microalgae assessment, the issue was to develop a fast and reliable method to analyze microalgae cultures; in the soil color and fertility study, the issue was to relate color and fertility in order to find relationships between both. Lastly, in the meteorological data assessment, the matter was to generate a method for data acquisition in locations with no availability of weather stations.

In this thesis, the hypothesis formulated for the research topics were:

- Microalgae light absorbance analysis through ANNs can elucidate the biological culture composition and be used as an alternative for the assessment of microalgae cultures.
- Soil color analysis can provide information about fertility parameters, and ANNs can be applied to map this relationship and conceive a model for rapid soil evaluation.
- Meteorological data acquisition can be performed by using interpolation techniques and ANNs, alongside real data from weather stations to interpolate values with high accuracy.

1.2 Objectives

1.2.1 General objective

The general objective of the research is to generate ANN models capable of addressing agricultural related problems as an alternative to traditional and more expensive methods for management, analysis, and data acquisition in the crop systems.

1.2.2 Specific objectives

The specific objectives were separated for each of the main research topics of the thesis. The established objectives are presented below.

In the microalgae culture study:

- Study light absorption spectra of the different microalgae species.
- Evaluate the ability of ANN to differentiate between monoalgal and mixed algal cultures.
- Determine the feasibility of using ANN to estimate the biological composition of mixed microalgae cultures.

Regarding the soil color and fertility study:

- Analyze soil color in order to determinate its capability to group soils according to fertility levels.
- Develop a computer tool for soil color measurement and classification.
- Describe soils fertility parameters using color analysis through ANNs.

Concerning the meteorological data acquisition study:

- Evaluate and compare the accuracy of several interpolation algorithms — including the traditionally used and ANNs.
- Study the seasons and the extreme phenomena effects in the performance of interpolation methods.
- Devise a set of algorithms to acquire real meteorological data, process them, and generate accurate estimations in distinct locations, economically and straightforwardly.

1.3 Methodology overview

To conduct the experimental phase of the thesis, three experiments were carried out. For the microalgae experience, monoalgal and mixed algal culture spectral signatures from light absorption measurements were analyzed. Additionally, an ANN was used alongside the spectral signature in order to create a model capable of classifying the microalgae cultures and determine the species present in suspension.

For the soil study, soil samples were classified according to the Munsell color notation, and the obtained hues were used to group soils and perform statistical analysis over the fertility parameters. In addition, RGB and L*a*b* colors were used as input in ANNs to create models to describe the soil fertility parameters. The RGB and the L*a*b* colors were obtained from digital color photographs and computer software programmed to perform quick and accurate measurements.

Finally, for the meteorological data acquisition study, daily data from a weather station network were used to perform interpolations with several methods, including ANNs and traditional techniques. With these interpolations algorithms, the development of virtual weather stations (VWS) was proposed, scripts automatically acquire, and process meteorological data were coded, and afterward, perform estimations in different locations without weather stations availability.

1.4 Motivation for the research

ANNs studies are nowadays a multidisciplinary science, a crucial factor for that is their application to solving problems in different areas, pursuing the modernization of techniques, better prediction of phenomena, classification, and estimation of everyday and important matters. The room for applications of ANN are infinitely potential, and the scope of this tool is changing the manner scientific process data and conduct experiments.

The agriculture is a niche for infinite research since with new days, new problems and challenges appear and need solutions, and at the same time, better solutions for older problems are also needed. With the aim of solving agricultural related subjects, this research was motivated and focused on microalgae culture analysis, soil color and fertility studies and meteorological data acquisition through interpolation methods.

Microalgae are an interesting group of microorganisms, and their cultivation is a promising tool for carbon sequestration and biomass utilization, for example, as an organic soil amendment — therefore their agricultural and environmental importance will arise in the next few years. Regarding soil fertility, this parameter is one of the most important in the production ecosystem, and its knowledge is vital for the correct and sustainable management of soils. Concerning the meteorological data, agriculture as an activity that is highly dependent on the weather requires the availability of these type of data can be useful for improving, for instance, the irrigation calculus and planning, crop phenology, pest studies and modeling, among other aspects.

1.5 Innovative aspects of the research

This thesis was designed with the practical application of technologies and techniques in mind, to solve specific problems in agricultural disciplines. The innovative aspects and contributions of this research are methodologies and tools that can be used directly in culture systems, been microalgae or conventional crops. The main contributions are cited below:

- An ANN to elucidate microalgae species in suspensions was made. The model provides a fast and powerful tool for microalgae culture management at the commercial scale, provides information regarding the biological composition of cultures, and approximates the relative proportions of species in the suspensions. The developed method means a cheaper and faster analysis method in comparison with the typically used.
- A scientific paper with the title “Monoalgal and mixed algal cultures discrimination by using an artificial neural network” (Appendix A) was published in the “Algal Research” journal. A Q1 journal in Biotechnology and applied microbiology journal with an impact factor of 3.75.
- The development of computer software, the DigiCIELAB, which is a powerful tool to perform the color measurement in a faster and more affordable manner in comparison with traditional colorimeters. The software was selected by the University of Valladolid in the “Prometeo” 2017 program and intellectually protected as a result of this selection (Appendix B).

- The development of the VWS algorithms to acquire and process meteorological data from weather station networks with the ultimate purpose of estimate data where no weather station is available. The VWS is a feasible alternative for the acquisition of meteorological data of importance for agricultural activities.

1.6 Thesis structure

After this overview of the thesis research problems, hypothesis, objectives, and other introductory aspects treated in this chapter. The remaining of the manuscript is organized as follows: a theoretical background with the literature of the current estate of the art in agriculture challenges, ANNs, brief microalgae production concepts, soil fertility and color theory and meteorological data importance — with the descriptions of techniques used in each one of these topics — is presented in Chapter 2.

The materials used and the methodology applied in the experimental phase for the microalgae, soil color, and meteorological data studies are detailed in Chapter 3. The results of these experiments alongside with discussion with the most relevant literature to contrast the obtained results are presented, separately by research topic, in Chapter 4. The conclusion, the general from the overall research experience and the specifics regarding each topic, are illustrated in Chapter 5. Finally, the cited bibliography is referenced (Chapter 7), and the additional information is shown in the Appendix section.

THEORETICAL BACKGROUND

2. Theoretical Background

In this chapter, the theoretical background that supports the thesis development will be presented, including the basic concepts for the methodology applied to seek the research objectives. This section will describe a brief agriculture panorama, future challenges and possible manners to address the issues with technologies and alternative agricultural practices (section 2.1); the microalgae production and its agriculture importance (section 2.2), soil fertility and color analysis (section 2.3) and meteorological data for agriculture (section 2.4) are treated as well. Finally, an overview of ANNs, concepts, modeling and training process, and applications of this technique (section 2.5).

2.1 Challenges in agriculture and new technologies

Achieving maximum crop yield at minimum cost is one of the goals of agricultural production from an economic point of view. However, since 1950, a climatological transformation is taking place as a consequence of several phenomena like deforestation, the emissions of greenhouse gases (GHG) increment, ozone loss, the increment of global temperature, change in precipitations regimes and other climate change effects, with a higher acceleration from the 90s (Smith *et al.*, 2014; Steffen *et al.*, 2015). These transformations, specifically regarding the anthropogenic climate change, are considered as one of the most significant environmental, social and economic threats to the future world (Ghosh *et al.*, 2019) and agriculture is one of the most exposed to climatic impacts (Martins *et al.*, 2019; Tran *et al.*, 2019).

Global food demand is expected to increase considerably in the near future as a result of the growing population. Agricultural research needs to step up to meet a more sustainable development for food production, human nutrition, climate change and environmental protection in a world with 9.7 billion people by 2050 (Thornton *et al.*, 2018). Considering that, with the modernization of agriculture, high inputs of fertilizers, pesticides, and mechanical energy are demanded for labors (Erb *et al.*, 2008; Wu *et al.*, 2017), GHG emissions such as carbon dioxide (CO₂) (Li *et al.*, 2019), nitrous oxide (NO₂) which is one of the most harmful GHG (Liu *et al.*, 2019; Rowlings *et al.*, 2013; Wolff *et al.*, 2017) and others

such as methane (CH₄), nitrate (NO₃⁻), ammonia (NH₃) will increase (Sanz-Cobena *et al.*, 2017a).

It is estimated that agriculture contributes 11% to the anthropogenic GHG emissions, approximately 5.3 Gt CO₂ equivalents in 2010 (**Figure 1**), the trend indicates that it will increase 9% in 2030 and 18% in 2050 with respect to 2010, or up to 37% more considering the 90s as base reference (Tubiello *et al.*, 2014).

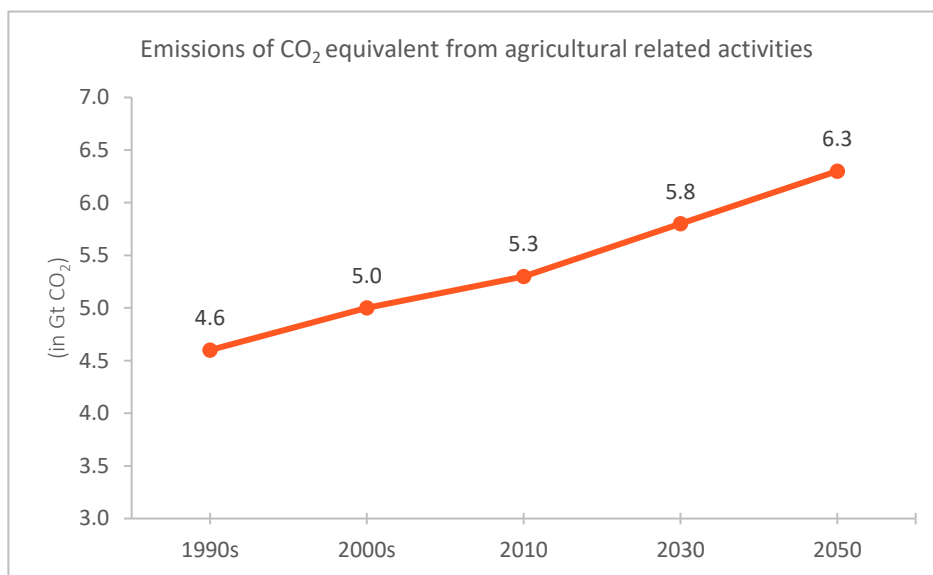


Figure 1. CO₂ emissions estimations for the 1990s, 2000s and 2010 and projections for 2030 and 2050 for agricultural related activities (elaborate from data of Tubiello *et al.*, 2014).

Climate change is a danger for crops and food production (Fitzgerald *et al.*, 2019; Manners and van Etten, 2018) and risk for smallholder farmers and livestock sector, particularly in dryland regions (Hansen *et al.*, 2019; Herrero *et al.*, 2015). Globally, annual climate variability accounts for roughly a third (32–39%) of the observed crop yield variability (Ray *et al.*, 2015); increasing the yield losses in warmer years; it is estimated that for each degree Celsius (°C) of mean global temperature, there is a reduction in global wheat grain production of about 6% (Asseng *et al.*, 2015).

Minimizing and preventing environmental degradation is a critical sustainability challenge of the present and coming decades (Bais-Moleman *et al.*,

2019). Climate-smart agriculture (CSA) is widely promoted as an approach for reorienting agricultural development under the realities of climate change (**Figure 2**), seeks to meet three challenges: improve the adaptation capacity of farming systems to climate change, reduce the greenhouse gas emissions of these systems, and enhance agricultural productivity (Acosta-Alba *et al.*, 2019; Jagustović *et al.*, 2019; Thornton *et al.*, 2018).

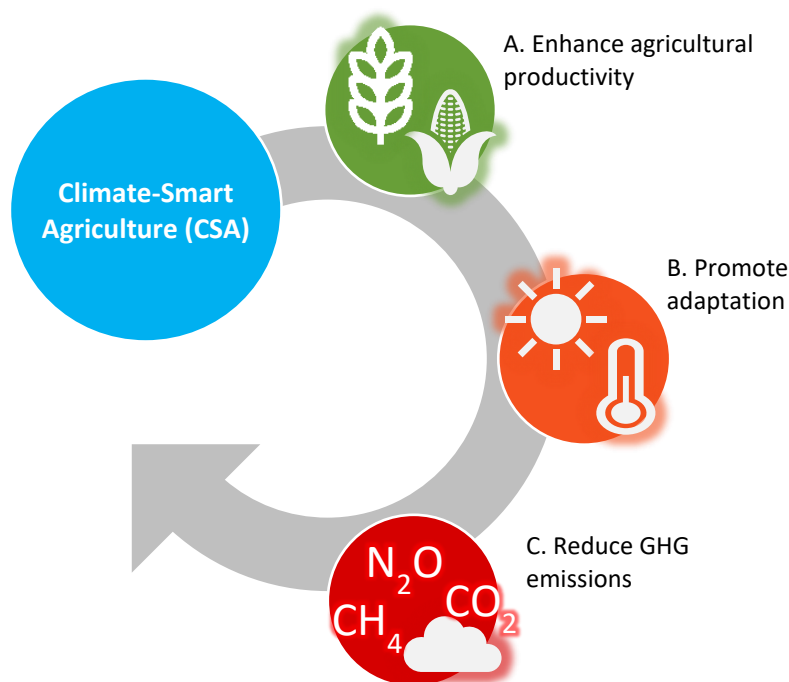


Figure 2. Climate-smart agriculture foundations for the development of sustainable agriculture.

A key reason for the emergence of the CSA concept is the recognition that agriculture, and related food security issues, require a synthesized approach which may not be achieved by tackling climate mitigation and adaptation objectives separately (Long *et al.*, 2016). CSA is the combined strategies to respond to the challenges of making food security, providing public goods and ecosystem services to society (Bais-Moleman *et al.*, 2019).

Some strategies to achieve the CSA objectives regarding enhance agricultural productivity (A); promote adaptation (B); and, reduce GHG emissions (C) are described below:

A. Enhance agricultural productivity

The agricultural force labor has globally declined from the 19th century from 75% of the total population to levels underneath the 5% in the current Century (Tomich *et al.*, 2019). Less agriculturist produced food for more people; this was possible using the technological improvements in the sector. However, the unsustainability of agricultural ecosystems under the current techniques have created the need for actions to not only to increase the production but rather to maintain or recovery degraded fields (Ahmad *et al.*, 2017).

On the one hand, the progressive substitution of mineral fertilizers for organic fertilizers can improve soil fertility and recover poor soils (Sanz-Cobena *et al.*, 2017b; Zhang *et al.*, 2019). On the other hand, practices such as no-till of lands significantly increased yields and improved soil quality, also shifting the microbial community and activity associated with an increase in organic matter, carbon, nitrogen and phosphorus cycling (Mbuthia *et al.*, 2015). Understanding of trade-offs over temporal and spatial scales between the instantaneous short-term gain, such as incomes, and the long-term negative impact, such as reduction of crop yields, is a crucial factor to incentivize agriculturist to adopt new techniques (Jagustović *et al.*, 2019).

In terms of alternatives for chemical fertilizers, microalgae have been considered potential producers of organic fertilizers for many decades and their potential is still under scope and development today (Benemann, 1979; Morales-Amaral *et al.*, 2015). Microalgae are photosynthetic microorganism adapted into several aquatic and terrestrial ecosystems (Dashkova *et al.*, 2016; Feng *et al.*, 2016) and with a strong presence in agricultural fields (Morowvat and Ghasemi, 2016).

Recent research activities have demonstrated that microalgae extracts have a stimulation effect when used as a foliar product on wheat (Shah *et al.*, 2013) and maize (Zermeño-González *et al.*, 2015). In the same way, the application of microalgae on wheat seeds at the time of sowing provides benefits for the plant and crop development. The causes can be related to the generation of phytohormones such as indoleacetic acid (Hussain and Hasnain, 2011; Jäger *et*

al., 2005; Swarnalakshmi *et al.*, 2013) or the activity of nitrogenase, which increases the binding of atmospheric N₂ (Abd-Alla *et al.*, 1994). In the current market, microalgae extracts are available in compounds for organic farming labeled as bio-stimulators.

Microalgae biomass application in soils is beneficial for crops and yields (Shaaban, 2001), the organic matter, carbon content and biological activity in soils increase with these microorganisms (Castro *et al.*, 2017; Miralles *et al.*, 2012; Xie *et al.*, 2007). Microalgae, specifically when they are obtained from wastewater treatments (De-Bashan and Bashan, 2004), contains several nutrients removed from the treated sewage. The recycling of nutrients by microalgae is a matter that has been more studied recently and tested in agricultural applications (Egle *et al.*, 2016; Mukherjee *et al.*, 2015).

B. Promote adaptation

Augmenting climate resiliency of agricultural systems with improved varieties of cultivars is one of the most suitable action to adapt agriculture for future scenarios (Dey *et al.*, 2019). Adaptation to climate change is necessary to ensure food security and protect the livelihoods of farmers (Sapkota *et al.*, 2019). However, assigning a label of “climate adaptation” to any promising innovation in agriculture is a common flaw. Many of these crop varieties do not always arise in minimizing the climate change impacts, other than maximizing the profits for a given technology (Lobell, 2014).

The major global studies are focused in different adaptation strategies such as changes in planting dates, dynamic irrigation, nutrient management or their combination and also included carbon fertilization effects (Aggarwal *et al.*, 2019). Other adaptation strategies include the calendar days of a crop and the growing degree days, cultivars with more extended growth periods and higher thermal requirements could partially mitigate the adverse effects of global warming (Martins *et al.*, 2019; Wei *et al.*, 2014).

Adopting other productive practices and technologies, such as more organic farming, can be a powerful mechanism for reducing farm vulnerability (Bouttes *et al.*, 2019). The adoption of alternative techniques requires interdisciplinary efforts to identify real options (Bommarco *et al.*, 2012). For instance, more information is required for better irrigation management and crop calendar planning. This information includes meteorological data for the

estimation of water requirements (Allen *et al.*, 1998) or taking actions to prevent extreme temperatures damage, stress, yield losses, germination problems, pest and disease affections (Johkan *et al.*, 2011; Pearce, 2001).

C. GHG emission reductions

Reducing GHG emissions is a cross effort in distinct areas. New technologies are consigned to study GHG causes, increasing the efficiency of machinery and proposing alternative practices are conventional approaches nowadays (Echevarria and Xiu, 2014; Nejat *et al.*, 2015).

Carbon sequestration in agricultural soils is an exciting alternative against GHG (Poeplau and Don, 2015). Increasing soil organic carbon (SOC) to store CO₂ amounts into grounds that otherwise will reach the atmosphere (Meyer *et al.*, 2018), also reduces external inputs in terms of fertilizers and makes a better scenario for crops under variable and extreme climate conditions (Pittelkow *et al.*, 2015). The changes in SOC stocks are a result of the carbon inputs, mainly in the form of dead plant material or manure and outputs, mainly caused by decomposition, leaching, and erosion (Poeplau and Don, 2015).

Tillage is a significant contributor to energy consumption and carbon emissary in the sector. Different management practices such as non-tillage, rotations and cover crops can mitigate GHG emissions, especially the ones related to CO₂ and NO₂ from savings in fuel and nitrogen volatility (Bais-Moleman *et al.*, 2019; Pratibha *et al.*, 2019; Sanz-Cobena *et al.*, 2017b). Non-tillage contributes to organic matter (OM) accumulation, increasing the soil water retention capacity, biological activity, increase the efficiency of fertilization and reduces its use (Rusco *et al.*, 2001; Sanz-Cobena *et al.*, 2017b). At the same time, no-till produced the highest revenues, despite having higher initial and production costs than conventional systems, but the overall balance is positive in favor of non-tillage systems (Hansen *et al.*, 2019; TerAvest *et al.*, 2019).

There is an increasing concern about the negative impacts associated with the release of reactive nitrogen from fertilizers (Recio *et al.*, 2018). N-source chemical fertilizers are the main source of NO₂ GHG in agriculture (Wolff *et al.*, 2017). Replacing a part of chemical N fertilizer is an excellent strategy to mitigate the negative impacts (Jat *et al.*, 2019). However, there are still several limitations in the economic and source of alternative fertilizers (Zhang *et al.*, 2019). Evidently, chemical fertilization cannot and should not be totally removed from

crop management tasks, instead, better fertilization programs have been carefully elaborated since the efficiency of fertilization is a critical issue to reduce environmental impacts of agricultural activities and ensure food production (Recio *et al.*, 2018; Zhang *et al.*, 2015). Variable application rates of nitrogen achieve lower expenses on supplies, higher profits and reduce nitrate leaching (Basso *et al.*, 2016); better timing for applications (Rowlings *et al.*, 2013) or fertigation with nitrate-based fertilizers (Abalos *et al.*, 2014) are others alternatives for a more efficient use of nitrogen fertilizers and address the challenges of environmental degradation and climate change.

CSA priorities encompass the adaptation of agricultural activities to the climate change for the 1.5 future while reducing the impact on the environment. The 1.5 scenario is the estimation where the mean global temperature will increase by 1.5 °C (Martinez-Baron *et al.*, 2018). As a summary, farmers adopt practices and technologies such as improved varieties, planting at optimal times, and enhanced water and fertilizer management (Aggarwal *et al.*, 2019) and researchers have to develop, test and validate those techniques to drive present and future situations.

There are a bunch of promising technology innovations; for instance, nanotechnology has proven to improve seed plant germination, growth, yields, and quality of the harvest. In post-harvest, nanotechnology increases the storage period for vegetables and fruits (Huang *et al.*, 2015; Prasad *et al.*, 2017). Information and communication technologies, as agro-advisory services, could be highly useful for decision making. However, to achieve this objective, more rigorous information is needed, and this warrants further work for researchers (Westermann *et al.*, 2018). The implementation of these tools at various scales and cases and the amount and quality of the information needed for making decisions and extract conclusions propose a big challenge for this approach (Thornton *et al.*, 2018).

Regarding technologies that contribute to achieving CSA challenges, novel techniques to analyze data, solving problems, create models, and predict high-quality datasets can be useful to face CSA objectives. In these terms, artificial neural networks (ANNs) are a promising tool for analyzing data and solve complex and nonlinear problems, with several potential application in the agricultural sector. ANNs are mathematical models that were developed from studying how the animal brain works. They consist of interconnected units called artificial neurons, which have adaptive parameters that can be used for learning, classifications, and predictions (Olatunji and Arif, 2016).

ANNs adoption to solve diverse kinds of problems have been taken enormous attention in the last two decades (Tkáč and Verner, 2016). ANNs can be used for classifications problems, approximation functions, image analysis, patterns analysis, and others (Zurada, 1992). To mention some examples, ANNs are capable of estimating the status of rotary elements of agricultural machinery by processing the signal from accelerometers without other variable analysis (Martínez-Martínez *et al.*, 2015).

Crop yields predictions based on soil fertility parameters and metrological data was possible by using ANNs, thanks to its abstraction capability (Dahikar and Rode, 2014). In plant diseases, an ANN was able to estimate the severity of angular leaf spot in bean cultures, using leaf and canopy hyperspectral reflectance spectra (Martínez-Martínez *et al.*, 2018). The use of ANNs have a widespread future in the development of new techniques, and there is more room for future research and applications of ANNs (Darji *et al.*, 2015; Ghanbarzadeh *et al.*, 2009; Laaboudi *et al.*, 2012).

2.2 Microalgae culture

Microalgae include prokaryotic and eukaryotic microscopic unicellular and multicellular photosynthetic organisms (Soni *et al.*, 2017). These organisms are a promising tool to fight climate change by capturing CO₂ (Bai *et al.*, 2017) and also in agriculture as fertilizers (Abdel-Raouf N, 2012). Microalgae contain several nutrients, especially those removed from the treated sewage for their recycling (De-Bashan and Bashan, 2004) and generating biomass in a fast and efficient way compared to other photosynthetic organisms, being up to 10 times more photosynthetic efficient than plants (Singh and Ahluwalia, 2013).

Considering that CSA is oriented to reduce GHG emissions and enhance the productivity of agricultural ecosystems, microalgae culture and production can be a decisive factor for future scenarios. Several experiments were conducted to study the benefits of microalgae in agriculture, from the fertility perspective (Mulbry *et al.*, 2005; Raposo and Morais, 2011; Renuka *et al.*, 2016; Shaaban, 2001) to the environmental approach (Castro *et al.*, 2017; Egle *et al.*, 2016; Renuka *et al.*, 2016).

Microalgae technology for production is under continuous development, and there is a place for improvements. This section of the text is focused on describing the microalgae and its production, emphasizing the management and the need for controlling the production systems in terms of biological composition.

2.2.1 Microalgae overview

Microalgae are diverse, unicellular microorganisms that grow in multiple environments. The three most important classes of microalgae in terms of abundance are the diatoms (Bacillariophyceae), the green algae (Chlorophyceae), and the golden algae (Chrysophyceae). All these microalgae are eukaryotes distinguished by the presence of a nucleus and separate organelles for photosynthesis, the chloroplasts, and for respiration, mitochondria. Likewise, some prokaryote organisms are also considered as microalgae for their photosynthesis capacity, the cyanobacteria, or as commonly known, blue-green algae (Cyanophyceae). The cyanobacteria are part of the eubacteria, and as a prokaryote, is lacking a membrane-bounded nucleus. There are thousands of different species of microalgae considering this classification (Enzing *et al.*, 2014).

From all these thousands of species, only a few are being produced commercially today, primarily for high-value products (Rickman *et al.*, 2013). Microalgae have been proposed for a wide range of applications, from the production of foods and animal feed, cosmetics, biofuels and wastewater treatment processes (Borowitzka, 2013; Chisti, 2007; Enzing *et al.*, 2014; Olguín, 2012).

Microalgae are cultivated at an industrial scale in two widespread systems, the open-culture systems and the closed-culture systems (**Figure 3**). Open-culture systems, for example, open ponds and raceways, are the simplest and less expensive in comparison to closed-culture systems. Open-culture systems are almost always located outdoors and rely on natural light for illumination. Unfortunately, the microalgae culture can be easily contaminated, and little control of the operating conditions can be made on such systems (De Andrade *et al.*, 2016). Closed-culture systems, such as tubular photobioreactors, allow a certain control level of operating conditions and to avoid contamination, being possible to obtain high-value algal products. Closed photobioreactors may be located indoors or outdoors, but the outdoor location is more common because it can make use of free sunlight (Molina Grima *et al.*, 2003).

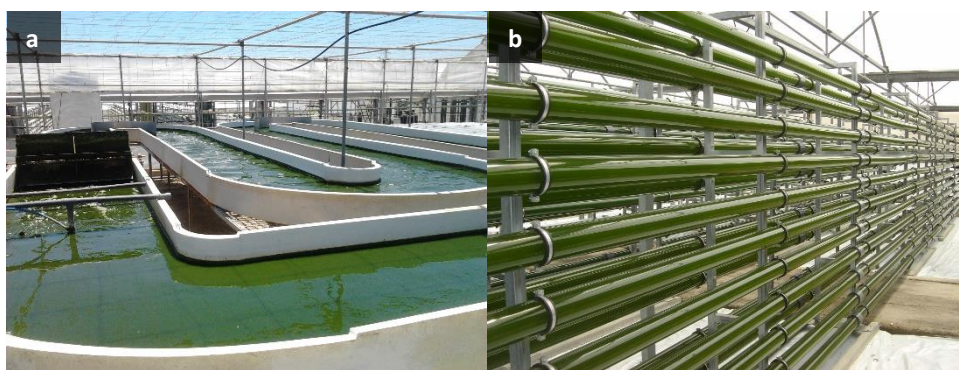


Figure 3. Microalgae production system in an open-culture system (a) and a closed-culture system (b).

Despite the large variety of applications proposed, only a few are presently performed at the commercial scale using a limited number of algal strains. Examples of this are the production of carotenoids, beta-carotene and astaxanthin from *Dunaliella salina* and *Haematococcus pluvialis* (Forján *et al.*, 2014), biomass for foods from *Chlorella vulgaris* and *Spirulina platensis* (Fradique *et al.*, 2010), and

biomass for aquaculture from *Nannochloropsis gaditana*, *Tetraselmis suecica* and *Isochrysis galbana* T-ISO (Shields and Lupatsch, 2012).

Microalgae have other applications such as carbon sequestration by transferring the environmental CO₂ into the microalgae mass (Acién *et al.*, 2012; Sobczuk *et al.*, 2002), wastewater treatment (Acién *et al.*, 2016; Ledda *et al.*, 2015), and microalgae biomass fertilizers (Mulbry *et al.*, 2005; Raposo and Morais, 2011; Shaaban, 2001). For fertilization, microalgae are a source of one of the scarcest nutrient, the phosphorus, which is abundant in the algal mass (Egle *et al.*, 2016; Melia *et al.*, 2017; Mukherjee *et al.*, 2015). Despite being effective, these alternative uses are not implemented at commercial scale yet.

However, recent advances allow the scientific community to be optimistic in the development of production technologies which are more economically viable, and within the next 10 to 15 years these technologies will allow applications that are not viable at present (Acién *et al.*, 2014). New technologies are in development to decrease the cost of microalgae production such as species selection, supplies optimization, and better condition controls (Collet *et al.*, 2011; Madkour *et al.*, 2012). These advances make the microalgae biomass one of the most attractive alternatives for fertilizers in agriculture and CO₂ sequestration to achieve the CSA objectives of mitigating and reduce the GHG emissions and improve agricultural sustainability.

2.2.2 Specificity in microalgae production systems

For microalgae cultures and the harvested product, it is often required the maintenance of monoalgal cultures when a specific compound is desired. Whereas when focusing on biofuel production or wastewater treatment, the utilization of mixed cultures is usually acceptable. In the case of mixed algal cultures, the relative composition is gradually modified according to changes in environmental or operational conditions (Godos *et al.*, 2009).

Monitoring the biological composition of microalgae cultures is a necessary task, generally performed through routine microscopic examination. By means of light microscopy, an expert can distinguish the presence of a “contaminating” microorganism and whether the current algal strain is close to the expected value. However, only highly skilled taxonomists are capable of correctly recognizing algal strains (their species and genera) using light microscopy

observation solely based on morphology – this is because most of the strains are small round cells with similar features, only a few have easily recognizable morphology. Some microalgae can be classified by a computer using high-resolution images for the morphological characterization (Walker and Kumagai, 2000). This classification is performed automatically only with simple forms such as circle, ellipse or cell sizes; complex forms or similar species require the intervention of an operator, but the process is under human error that can lead to confusion (Mirto *et al.*, 2015).

Alternative methods, based on omics allow to accurately identify the microalgae strains in cultures (Godhe *et al.*, 2002), but these are expensive and require much time (reducing time can be useful for making operation process decisions). As a standard method, biochemical analyses, such as the chlorophyll to carotenoid ratio and the fatty-acid profile have also been used as tools for verifying the biological composition of microalgae cultures; however, their precision is limited. These methods are unable to identify the presence of low-level contamination; likewise, they take time, although their cost is much lower than that for the omics methods (Serive *et al.*, 2017; Sydney *et al.*, 2011).

Microalgae species identification up to the phylum or class levels has been achieved based on the fluorescence properties of photosynthetic pigments using flow cytometry (Cellamare *et al.*, 2010). Employing light-emitting diode induced fluorescence analysis is possible to differentiate between *Anabaena* sp. and *Cylindrospermum* sp. cells by comparing the fluorescence spectra (Ng *et al.*, 2017). Microalgae have photosynthetic pigments providing different spectral signatures for different species; thus, it is possible to build classes based on the presence of pigments. As a result, the relative content of chlorophyll, carotenoids and other pigments can be used to differentiate between diatoms, red/green/brown microalgae, and cyanobacteria groups through their light absorption spectra (Serive *et al.*, 2017).

Microalgae species can be distinguished by their spectral signature. As an example, *Rivularia* M-216 exhibits a different absorbance signature to that of *Anabaena variabilis* – the heterocyst absorbance from *Rivularia* is more than double than that from *A. variabilis* at wavelengths between 540 and 620 nm; this variation is a result of the different phycocyanin and chlorophyll contents (Nozue *et al.*, 2017). The spectral signatures of *Botryococcus braunii*, *Chlorella* sp. and *Chlorococcum littorale* allow identifying them by comparing their absorption indexes (Lee *et al.*, 2013). Moreover, the absorption spectrum in the 400 to 700 nm

range is used to determine the extinction coefficient of the biomass, a specific microalgae strain and culture conditions indicator (Rubio Camacho *et al.*, 2003). Based on this, absorption properties could be a possible approach to distinguish between microalgae species (Coltelli *et al.*, 2017).

ANNs are a powerful tool for finding relationships between experimental data and the phenomena behind these data. ANNs have been used to predict harmful algal blooms in lakes (Recknagel, 1997; Tian *et al.*, 2017), as well as microalgae growth and biomass concentration under laboratory conditions and outdoor environments (García-Camacho *et al.*, 2016; Sharon Mano Pappu *et al.*, 2013). In microalgae identification using ANN, extracted features such as the perimeter, shape, area and Fourier Transform of microalgae micrographics were used to train a model capable of identifying the genera *Navicula*, *Scenedesmus*, *Microcystis*, *Oscillatoria* and *Chroococcus* (Mosleh *et al.*, 2012). Microalgae micrographic image processing and color analysis were used with ANN to achieve taxonomic accuracy of up to 99% by first detecting a cell in the image and subsequently extracting the detected cell color (Coltelli *et al.*, 2017).

Considering the capacity of ANNs and the light spectral signature properties of microalgae, it is possible to consider a fast and economical method derived from these two elements for the biological monitoring of cultures, especially those that need a specific species composition, pure or in proportions.

2.3 Agricultural soil fertility

Fertility is the soil ability to supply essential plant nutrients in adequate amounts, proportions, and moments for the growth and reproduction of plants (Bünemann *et al.*, 2018). In the last few decades, intensive agricultural management practices in European agriculture have resulted in soil degradation (Freibauer *et al.*, 2004; Virto *et al.*, 2015). OM and N content in soils are now 30% to 60% lower than their undisturbed (virgin) equivalents, about 44% of southern Europe lands exhibits low OM content as a result of intensive cultivation (Rusco *et al.*, 2001).

In the framework of sustainable agriculture, better soil management must be addressed to preserve and increase soil fertility in agricultural lands. Among the environmentally friendly labors such as non-tillage, pesticides reduction and administration of fertilizers, these can be used more effectively to minimize supplies spending and contamination. A fast method for estimating soil nutritional content can help in the process of decision-making.

Soil color contents some information about fertility (Barrios and Trejo, 2003; Fleming *et al.*, 2004; Gray and Morant, 2003). The analysis of this property through machine vision can help to characterize soils. In consideration, the description of machine vision, color theory, and soil color properties will be reviewed in the section below.

2.3.1 Machine vision and color theory

Computer vision (CV) is the science of the design and operation of the software for the analysis and processing of image, whereas Machine Vision (MV), means a more global concept: the study of the hardware environment, the image acquisition techniques and the software, CV, needed for the development of applications (Davies, 2018). Following its origin in the 1960s, MV has experienced significant growth, and its applications started to expand until the present in diverse fields like medical diagnostic imaging, factory automation, remote sensing, forensics, autonomous vehicle and robot guidance (Brosnan and Sun, 2004).

Regarding the hardware for MV, the acquisition system is usually composed of three main components: a color digital camera, an illumination source, and an image processing software. The lighting source must provide uniform and consistent illumination across the sample to photograph, color temperature is also

considered and usually fixed around 5,000 K to 5,500 K. To ensure uniform illumination conditions, multiple light sources can be employed as long as they are homogeneous (Tarlak *et al.*, 2016; Valous *et al.*, 2009). The camera is usually located at a certain distance so that the measurement does not interfere with the illumination since the resulting images are highly affected by the light (ten Bosch and Coops, 1995).

Besides a digital color camera in a single setup, the most used for computer vision, dual cameras setups for stereoscopic image analysis can be used in controlled conditions as was previously described, the stereoscopic setups are used for objects dimension calculation inside the images (Zion, 2012). Multiple color cameras mounted in drones are usually used in outdoors studies (Rumpler *et al.*, 2017), or infrared cameras both indoors and outdoors (Celenk, 1990) are some acquisition systems employed in MV.

In the software aspect, the CV is composed of several algorithms and a vast number of variations of them that help the processing of a digital image. Between the most common algorithms are segmentation, thresholding, shape and edge detections, morphological operations, color extraction, among others (Chen, 2015).

2.3.1.1 Computer vision algorithms

In CV processing, several algorithms intervene; the main ones are segmentation, thresholding, shape and edge detections. A brief description of these algorithms is presented in the paragraphs below:

A. Segmentation

Segmentation consists of separate uniform and homogeneous regions of a digital image, generally objects, with respect to some background or other heterogeneous segments of the image (He *et al.*, 1985). Segmentation, for instance, is used to differentiate landscape sections in remote sensing or different tissues in biomedical images and to extract these objects/parts from the background, identifying blood cells in biomedical pictures, detecting machinery parts, among others (Prats-Montalbán *et al.*, 2011).

B. Thresholding

Thresholding is a type of segmentation. A given threshold value is applied to the pixels in a gray-scaled image, and pixels are classified according to this threshold in a binary image that results from this operation (Sudarsan *et al.*, 2016). A binary image indicates that the pixel value of “1” is where the pixel past the threshold and the pixels with value “0” indicate the segment did not pass the threshold (Promdaen *et al.*, 2014). The binary image is a mask that can be used, which contains less information and is easier to process by the machine (Zhang *et al.*, 2017). For instance, after the thresholding, a morphological operation can be performed in order to fill the possible holes presented in the binary image (Mery and Pedreschi, 2005).

C. Shape detections

Shape detections can be performed through several algorithms that compare matrices corresponding to shapes, such as circles, squares, triangles, and other more complex figures, with a sample image in which these matrices are compared. When the matrix patterns are found in the image, it means that a given shape is present in the image (Allili and Corriveau, 2007). Object detection is an essential and challenging task; this becomes particularly tricky in images with a strongly cluttered background, and with objects subjected to scale changes, rotation changes, and substantial intra-class variations (Wei *et al.*, 2017).

D. Edge detection

Edge detection methods are commonly applied to the image to assess any change in the intensity profile of neighboring pixels; therefore, a substantial intensity change between an object and the background corresponds to accurate edge information and yields excellent results (Williams *et al.*, 2014). The main challenges for edge detection are due to pixels noise in the background or another object interfering with the edge of the object of interest (Lu *et al.*, 2010). In addition, there is the concept of regions of interest (ROI). The ROI is assigned on the image in specific coordinates and has a given size, for instance, 60×60 pixels at the position $x = 200$ and $y = 200$ in a digital image (Sumriddetchkajorn *et al.*, 2014). From this ROI, multiple operations and analysis can be performed inside this region, including edge detection, among others (Losson *et al.*, 2013).

2.3.1.2 Color theory and systems

The color is the result of the interaction of certain light wavelengths with an object; this is an attempt to relegate color to the purely physical domain. Instead, it is proper to state that those stimuli are perceived to be of a particular color when viewed under specified conditions (Fairchild, 2013). In terms of computer vision, color is given by the R (red), G (green), and B (blue) tristimulus values. A color image is usually provided by the three values, RGB, at each pixel. In a typical 8-bit image, each component can take values from 0 to 255 (Celenk, 1990). Values of RGB show the total amounts of the three primaries that are required to make a color inside the RGB color system space (Gershon, 2005). In this system, the pure red color is given by (255, 0, 0), white by (255, 255, 255), black by (0, 0, 0) and so on combining the components. A scheme of the RGB system is shown in **Figure 4**.

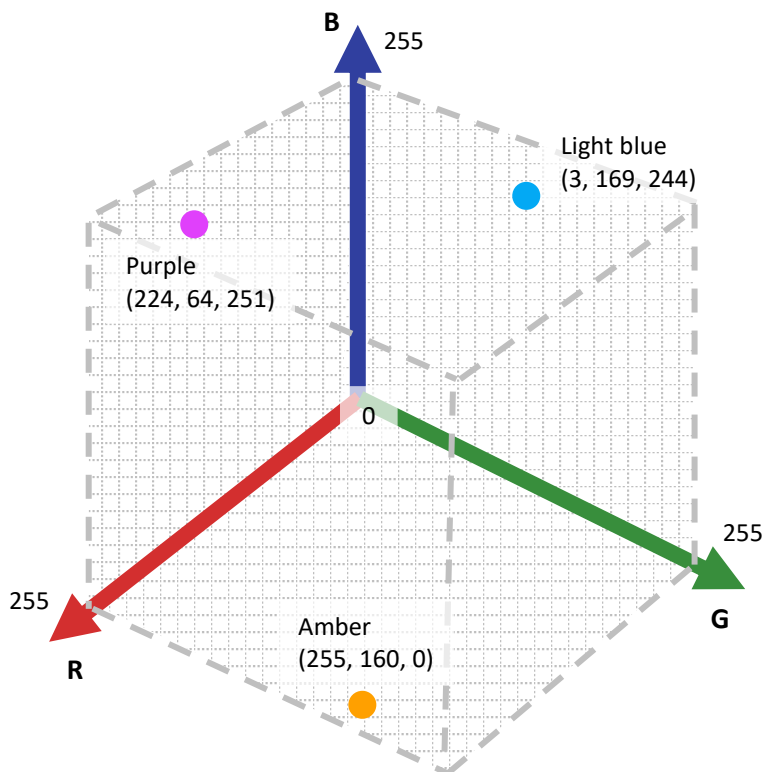


Figure 4. RGB color system diagram. Purple, amber and light blue color RGB values are described.

In CV, RGB color features are important as a mechanism to a rapid and non-destructive inspection of objects (Xia *et al.*, 2016). The color can be used to estimate the quality and maturity of tomatoes, citrus, cranberries, mangos and estimate chemical components associated with indices of quality, such as Brix degrees (Francis, 1995; Jha *et al.*, 2007). Others studies with color images were conducted for meat color and quality (Trinderup *et al.*, 2015), water purity analysis (Andrade *et al.*, 2013), and Earth surface color and texture assessments (Zhao *et al.*, 2016).

In these types of studies, a colorimeter is typically the first option for color measurements. However, the colorimeter shows some limitations because of the non-homogeneous color in the surface of the object to analyze, especially in food engineering and research — colorimeters analyzes points, not the whole surface (Barbin *et al.*, 2016; Girolami *et al.*, 2013; Yam and Papadakis, 2004). In this aspect, the color analysis of the digital images through CV is an advantage.

In research, color is frequently represented using the $L^*a^*b^*$ coordinates of the CIELAB color space. This color model is considered of uniform proximity, i.e., the distance between two colors in a linear color space corresponds to the perceived differences between them (Mendoza *et al.*, 2006). The CIELAB system was proposed by the International Commission on Illumination (CIE - by its initials in French, the Commission Internationale de l'Eclairage).

In this color space, the positive a^* axis points in the direction of red color, the negative axis in the direction of green stimuli; positive b^* points in the direction of yellow stimuli; negative b^* in the direction of blue stimuli. L^* is the luminance; thus a value “0” of lightness indicates the black or non-lightness and “100” represent the white in saturated light conditions (Schanda, 2007), a scheme is presented in **Figure 5**.

The $L^*a^*b^*$ color is device independent, providing consistent color regardless of the input or output device such as digital camera, scanner, monitor, and printer (Yam and Papadakis, 2004). This is the crucial difference between RGB and $L^*a^*b^*$, RGB is device dependent and the color of a given picture, under equal illumination conditions, leads different cameras to obtain different colors, in fact, vastly different in colors are usually obtained among cameras (Ilie and Welch, 2005; Kim *et al.*, 2012). There are other device dependent color systems such as HSL (for hue, saturation and luminance), HSI (for hue, saturation and intensity) and HSV (for hue, saturation and value), which are mathematically similar and can be calculated

from the RGB; these color systems are commonly used in image processing (Saravanan *et al.*, 2016).

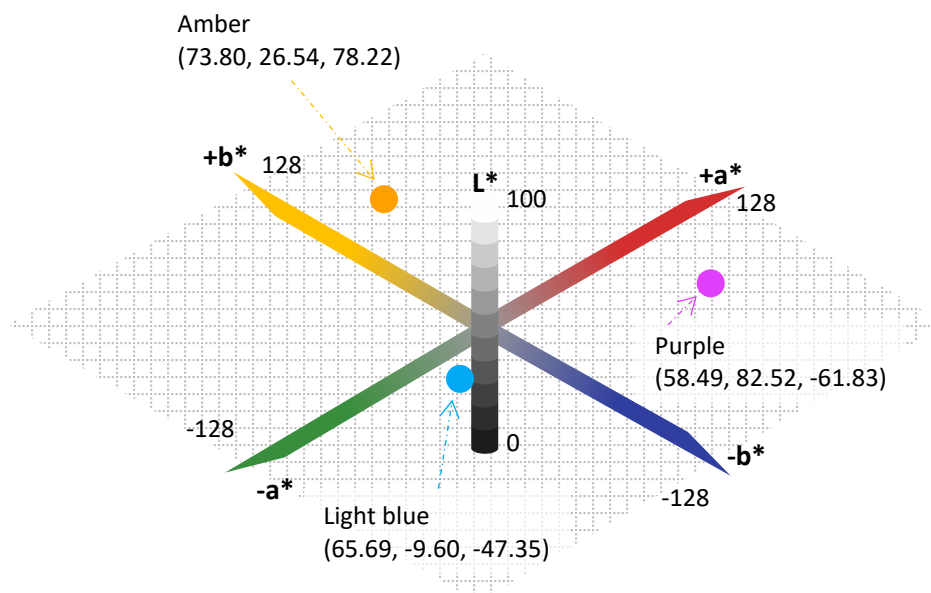


Figure 5. CIELAB color space representation. Purple, amber and light blue $L^*a^*b^*$ colors are described.

To address the non-uniformity results of digital cameras, a colorchecker is often used. A colorchecker is a color calibrated array of cards that help to tune cameras and counteract the variability among them. Colorchecker have different natural colors (red, blue, green, yellow, and others), grey tones, black and white; and with computer software the image color is modified according to the color cards to balance the image (Girolami *et al.*, 2013; Kirillova *et al.*, 2017; Potočník *et al.*, 2015). Additionally, colorchecker can also be used to convert the RGB values of the image into CIELAB standards if the manufacturer provides the $L^*a^*b^*$ data of the cards.

The RGB to $L^*a^*b^*$ color transformation can be performed through several methods such as equation systems (Barbin *et al.*, 2016), quadratic models, (Tarlak *et al.*, 2016), linear models and ANN (Afshari-Jouybari and Farahnaky, 2011). Between these methods, ANN present remarkable results for this conversion (Afshari-Jouybari and Farahnaky, 2011; León *et al.*, 2006; Pedreschi *et al.*, 2006;

Pothula *et al.*, 2015). This fact makes the CV color analysis possible using L*a*b* color in non-homogeneous surfaces such agricultural soils, and as in the fruit examples; examine possible relations between color and chemical properties, principally for nutrients and other fertility parameters.

2.3.2 Soil color and fertility properties

The soil is described as a complex and heterogeneous system with interactions with physical, chemical, and biological components (Awiti *et al.*, 2008). Soil management techniques are crucial for minimizing the impact in the use of this limited resource and have a strong influence in sustainable food production, biofuels, nutrient depletion issues and environmental concerns related to soils (Brevik and Hartemink, 2010). In agricultural soil management, soil fertility properties are one of the main concerns; usually, laboratory analyses are used to determinate soil properties to plan the management. Nevertheless, under the current situations — in which a large number of samples are required to develop a variable fertilization plan, conventional methods of soil analyses take a long time, in addition to their expensive costs (Mohamed *et al.*, 2016).

The complexity and cost of conventional soil analyses methods and field surveys lead to the development of indirect estimation methods based on proximal and remote sensing (Curcio *et al.*, 2013) such as visible–near infrared (vis-NIR) spectroscopy (Conforti *et al.*, 2015; Stenberg, 2010; Viscarra Rossel *et al.*, 2009) and soil color sensing (Baumann *et al.*, 2016; Castañeda and Moret-Fernández, 2013; Hausmann *et al.*, 2016), through these studies it has been possible to establish the relationship of these properties with soil OM content, organic carbon, clay, silt, and sand and other minerals such as content, kaolinite, haematite and goethite.

Color is an easily measurable physical attribute of soil through colorimeters, color sensors, computer scanners and smartphone cameras (Han *et al.*, 2016; Kirillova *et al.*, 2017; Stiglitz *et al.*, 2016; Zhang *et al.*, 2016). The difference in soil color has often been used in the soil classification (Maejima *et al.*, 2000), different studies indicate that the OM influences the soil color, darker soils have more OM (Kweon and Maxton, 2013). Iron oxides show both redness and yellowness (Vodyanitskii and Kirillova, 2016), other properties such as clay and humus contents also is related with the redness of the soil (Hu *et al.*, 2014).

The Munsell color system is one of the most widespread among soil scientists (Vodyanitskii and Savichev, 2017). In the Munsell System, the color is described by three factors such as hue, value and chroma. Hue is the color name such as Red (R), Yellow-Red (YR), Blue-Green (BG), among others. Value is a number; usually, from 8 to 2, that represents the lightness or darkness, a higher number stands for a lighter color and vice-versa. The chroma represents how weak or strong is a color; the minimum of 2 is a grayish color and maximum of 14 is a pure saturated color. Every color is alpha-numerically labeled with a letter for the hue and numbers for the value and chroma, for instance, 5Y 5/6 — the number before the hue letter indicates a subdivision of hue (Cochrane, 2014; Tokumaru *et al.*, 2002). An example of the Munsell color system is shown in **Figure 6**.

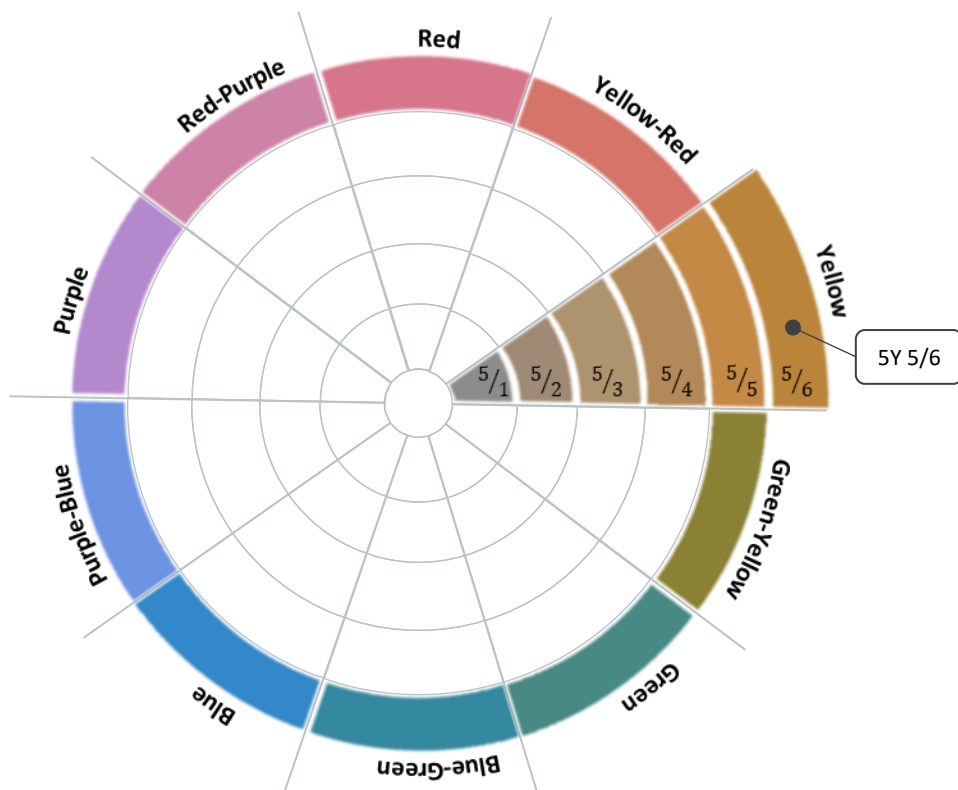


Figure 6. Schematic diagram of the Munsell color system for a given value (5). The hues are placed circumferentially and the chroma scale extends outward perpendicular to the center.

However, the Munsell color system has a severe drawback – the cylindrical coordinates complicate its use in statistical calculations (Vodyanitskii and Kirillova, 2016). The CIELAB system is more convenient and suitable for soil scientists, since it represents a universal color space in cartesian coordinates, for instance, Fe content contribution in soil color is more significant in the L*a*b* coordinate system than in the Munsell system, because this latter is divided into discrete and rough intervals (Kirillova *et al.*, 2015).

A suitable method to analyze and correlate color with fertility parameters of soil is the utilization of ANNs. ANNs can learn and generalize based on previous observations, and they can often correctly predict with high accuracies when the data interaction is complex for other techniques (Khashei and Bijari, 2010). ANNs for digital soil mapping could reduce costs and time in semi-detailed soil surveys with limited input data through data interpolation and extrapolation (Bagheri Bodaghabadi *et al.*, 2016). An ANN approach was used to determine relationships between soil color and a range of physical and chemical characteristics, using the database of National Soil Inventory of Scotland in an extensive scale survey (Aitkenhead *et al.*, 2013).

2.4 Agricultural meteorology data

The meteorology is the part of the atmospheric science responsible for providing weather forecast and communicates to the users who will use this data. Other than that, another important task of meteorology is to increase the knowledge of the atmosphere, with constant endeavor through research and development to improve the accuracy of the weather forecast (Rao, 2008).

Regarding the agricultural meteorology, the data generated in the last 20 years were useful to study plants growth, pests and pathogens appearance and propagation, water cycle, and protection of the crops from extreme conditions (Waggoner, 2016). The meteorological conditions vary in microscales and are subjected to several local elements and phenomena (Robitu *et al.*, 2006), for that, research is needed to provide additional information to farmers and the microclimate in regions where no source of meteorological data is emplaced. In this section, the importance of meteorological data for agricultural development will be reviewed and possible manners to generate an accurate estimation of these data will be presented as an alternative to acquiring information for better crop management.

2.4.1 Meteorological information and agricultural importance

Agriculture is one of the most sensitive activities to weather conditions, and the climate change has several impacts in this sector, from food production to economic problems in agriculture and related industries, making food security a tangible problem in the next few decades (Decker, 1994; Mason d’Croz *et al.*, 2013; Springmann *et al.*, 2016). Meteorological conditions are essential for cultures, and certain conditions must be met for their development (Griffiths, 1994).

Nonetheless, some changes have occurred, climatic anomalies and the modification of precipitations regimens are striking the agriculture (Luo *et al.*, 2019). In recent years, weather and climate have become an issue of interest for scientists, media and population because of the harmful effects associated with global warming (Ahrens, 2012; Webster, 2013), and humans activities have been reported as one of the leading causes (Barry and Chorley, 2009).

The knowledge of weather conditions helps to make better decisions in crops management, elaborate sowing calendars, pest population models (Naylor *et*

al., 2018), and precision irrigation dosage (Teshome *et al.*, 2018). For instance, irrigated agriculture is the primary water user, accounting for nearly 70% of total water consumption in the world (Conforti, 2011) and that consumption can be reduced using evapotranspiration models such as the proposed by the FAO (Allen *et al.*, 1998), which require prior knowledge of meteorological variables.

Meteorological observations are obtained through automatic weather stations, which are increasing throughout the world in the form of station networks (Estévez *et al.*, 2011). Nonetheless, the cost of an automatic weather station is still expensive and are not an attractive investment for most farmers. Considering the complex nature of the atmosphere and the massive computational power required to solve all of the equations that describe the atmosphere, the forecasts is also an expensive technique and new manners to obtain meteorological information more efficiently and less costly is becoming an interesting matter (Akram and El, 2016).

2.4.2 Meteorological data estimation

With new techniques for data processing and analysis, such as ANNs, novel methods have been developed for forecasting; precipitations (Hung *et al.*, 2009; Partal *et al.*, 2015) wind speed (Cao *et al.*, 2012), forecasting sea level in coastal areas (de Oliveira *et al.*, 2009; Filippo *et al.*, 2012) and air quality forecasting (Feng *et al.*, 2015; Hrust *et al.*, 2009). ANNs were also used for daily estimation of solar radiation (Amrouche and Le Pivert, 2014; Hasni *et al.*, 2012; Şenkal and Kuleli, 2009), power generation in wind energy systems (Fan *et al.*, 2009; Olaofe, 2014; Sideratos and Hatzigaryriou, 2007) and the estimation of evapotranspiration (Abrishami *et al.*, 2018; Ballesteros *et al.*, 2016; Chowdhury *et al.*, 2017; Cooper, 2010). However, all these studies worked based on measured data, i.e., were driven by previous known meteorological data to use them as the input to create models, there is a lack of studies focused on generating the first data and validating its quality using ANN.

Interpolation is a way to generate data in geographical locations where no data is acquired. Although deterministic and parametric statistical models have been the traditional approaches for data interpolation, such as inverse distance weighting, inverse squared distance weighting (Li and Heap, 2011; Lu and Wong, 2008) and multilinear regression (Jin *et al.*, 2016; Nalder and Wein, 1998); there are studies with new approaches such as random forest regression (RFR) machine learning method and ANN (Li *et al.*, 2011; Li and Heap, 2014; Wang *et al.*, 2009).

Regarding interpolation studies of meteorological variables, methods such as Thiessen polygons, kriging, regressions, nearest neighbor and inverse to distance weighted were employed to mainly interpolate precipitations and temperatures with different levels of accuracy (Berndt and Haberlandt, 2018; Jeffrey *et al.*, 2001; Mendez and Calvo-Valverde, 2016; Wagner *et al.*, 2012; Wu and Li, 2013).

Considering these facts, an ANN approach for data interpolation can be an exciting alternative given its capability in modeling and solving complex nonlinear problems (Voyant *et al.*, 2017). Generating more data and time series in various locations can be the basis for other studies, for instance, pathogen infestation in farms (Tripathy *et al.*, 2011), planning water management (Chowdhury *et al.*, 2017; Valipour, 2016) or checking the performance of solar energy based equipment and building in terms of energy consumption, especially for heating and cooling (Kumar *et al.*, 2015; Shu Fan *et al.*, 2009), using good quality interpolated data.

Meteorological data is an essential class of temporal record that can be useful for agricultural, energy, scientific, financial, and other applications. Since it is not possible to emplace a weather station in every location where data is needed, more research and information are needed to estimate meteorological data economically and accurately — especially considering the importance of this type of information in agriculture for crop management (Dahikar and Rode, 2014; Iizumi and Ramankutty, 2015; Kang *et al.*, 2009; O. Rauff and Bello, 2015).

2.5 Artificial neural networks

An ANN is a mathematical/computational model that is inspired by the structure and functional aspects of biological neural networks. It is a highly accepted technology to alternatively solve complex problems (Bouselham *et al.*, 2017). Comparative surveys show that accuracy of ANN methods is superior to that of traditional statistical methods in dealing with problems, especially regarding nonlinear patterns (Bahrammirzaee, 2010).

ANN belongs to machine learning (ML) set of techniques. ML can be described as such algorithms capable of producing learning by the mathematics in the models, and the fine-tuning of several parameters responsible for fitting the input data to the target. The optimization of such parameters could be understood as an unknown black-box function that invokes algorithms developed for such problems (Snoek *et al.*, 2012). Besides ANN, other remarkable ML algorithms are decision trees, random forest, support vector machine, k-nearest neighbors, among others (Maglogiannis, 2007).

In the following sections, an overview of ANN is presented. The ANN fundamentals, components, structures, and mathematics behind this technique are described. The review includes the basic units of ANNs, the artificial neuron or perceptron, how the learning is achieved through this technique, to the validation of the models for posterior inferences, classifications, and predictions.

2.5.1 The artificial neuron

Animals have complex brains composed of hundreds or thousands of millions of elements called neurons; the number varies according to species. In humans, the brain consists of approximately 10^{11} neurons – in this text animal, and consequently, human neurons will be referred to as biological neurons. These units are highly interconnected (approximately 10^4 synapses or connections per neuron), and an electromagnetic signal is passed through them, this signal is information that roves the brain, being processed by the neurons and derives into the diverse forms of actions or reactions (Zurada, 1992).

The biological neuron is a highly specialized cell type. Morphologically three major regions can be defined (**Figure 7**): a cell body (or soma), which contains the nucleus; a variable number of dendrites, which emanate from the soma and

ramify; and a single axon, which extends far from the soma and has numerous terminals to interconnect to other neurons dendrites (Squire *et al.*, 2008).

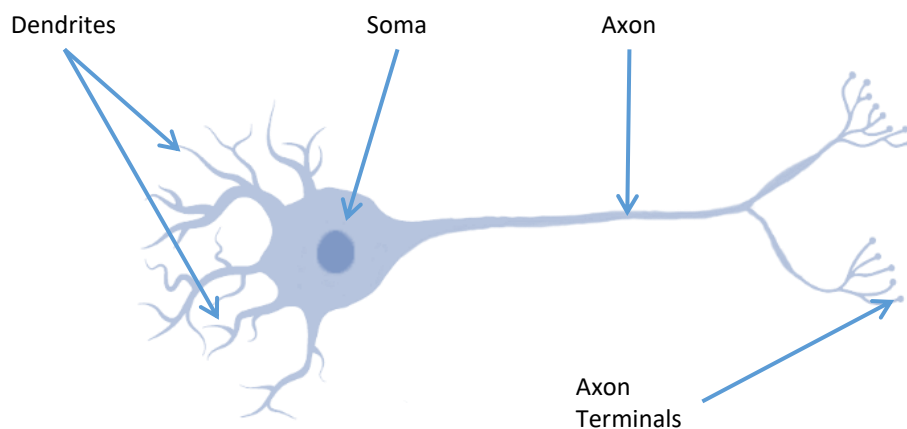


Figure 7. Biological neuron cell scheme. The signal travels from left to right, from the dendrites (input), passing through the body or soma and going out through the axon and its terminals.

An artificial neuron is a simple mathematical model inspired in its biological counterpart. Artificial neurons produce an output value in two steps. First, the neuron computes a weighted sum of its signals, input variables for the case, and, in a second stage, applies an activation function to this sum to derive the product as the output (Russell and Norvig, 2016).

In 1943, McCulloch and Pitts proposed the first artificial neuron model – today is known as the perceptron, composed of binary threshold activation function (**Figure 8**). This mathematical neuron computes a weighted sum of its input signals and generates an output of “1” if this sum is above a certain threshold, “0” or another positive number, otherwise, the function returns “0” as a result, this logic is also known as hardlim function (Jain *et al.*, 1996).

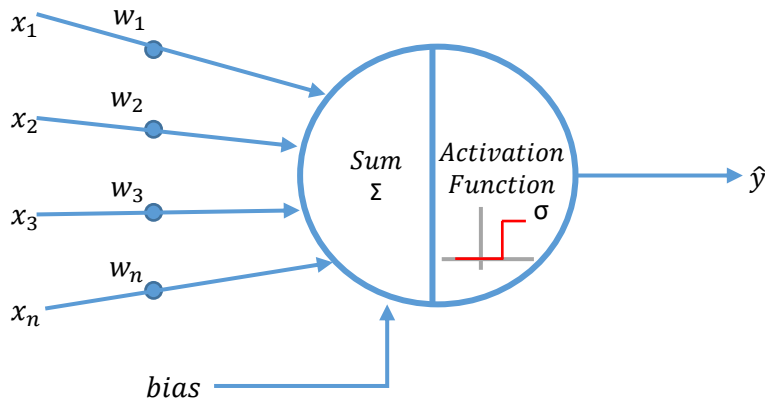


Figure 8. McCulloch and Pitts artificial neuron. Where x_n = inputs and w_n = weights. The sum of products of inputs and weights plus the bias is passed to the activation function (σ) to generate the neuron output (\hat{y}).

The perceptron applies a linear combination of its inputs, obtaining the signal in which is applied an activation function to obtain the output signal. Nonlinear functions are the most used ones to give the perceptron nonlinear behavior (Martínez-Martínez *et al.*, 2015). The mathematical model of the perceptron is presented in **Eq. 1**:

$$\hat{y} = \sigma \left(\sum_{i=1}^n x_i w_i + bias \right) \quad \text{Eq. 1}$$

Were \hat{y} is the neuron output; σ is the activation function; x is the input vector of n elements; w is the weight vector and $bias$ is a value that allows the shift of the activation function. The bias is somehow similar to the constant b of a linear function $y = ax + b$.

The biological neuron and the artificial neuron, the perceptron, are comparable in certain manners, a parallelism from an operational point of view can be described as follows: the weight (w) corresponds to the strength of a synapse, the neuron body is represented by the summation and the activation function, and the neuron output (\hat{y}) is represented by the signal on the axon (Mindiola *et al.*, 2015). In the perceptron, the analogous of the of electromagnetic signals in the processing of brain are, in deep, mathematical operations that take place in processor and random access memory (RAM) of a computer (Misra and Saha, 2010).

2.5.2 The artificial neural network

The perceptron by itself does not have a lot of processing power, especially when it is compared to an ANN, which is composed of several layers of perceptrons. ANN can be defined as a highly connected array of neurons, which are interconnected with each other; a widely used model called the multi-layered perceptron (MLP) is the most common type of ANN (Haykin, 1999; Park *et al.*, 1991). The MLP consists of one input layer, one or more hidden layers, and one output layer. Each layer employs several neurons, and each neuron in a layer is connected to the neurons in the adjacent layer with different weights, which is used to determine how much one unit will affect the other (Chen *et al.*, 2005).

Neural networks are typically represented by a network diagram, which is composed of nodes connected by directed links. Nodes are arranged in layers, and the structure of the most used neural network consists of three layers: an input, a hidden and an output layer of nodes (**Figure 9**), signals flow into the input layer, pass through the hidden layers, and arrive at the output layer in a unidirectional path in the MLP (Hastie *et al.*, 2009).

ANN is a kind of array which can realize a nonlinear mapping from the inputs to the outputs (Rui and El-Keib, 1995). Thereby, the input neurons receive the data, and the inputs and weights products are computed; the signals enter the hidden layer where a sum is performed, a bias is added, and an activation function is applied in each neuron. Then, the signals leaving the neurons in the hidden layer, are multiplied again by weights and enter the neurons in the output layer, where a sum is performed plus the bias addition to generate the output of the ANN.

Initially, the McCulloch and Pitts perceptron model presented limitations, and the nonlinear capability of ANN was not a reality until long after McCulloch and Pitts pioneer work. It was not until 1962, 20 years after the first proposal, that Rosenblatt implemented hardware with enough computing power to process an array of artificial neurons and the term “perceptron” was coined. He proposed single-layer networks composed of a few processing units. These networks were applied to classification problems, in which the inputs were usually binary images of characters or simple shapes (Bishop, 1995).

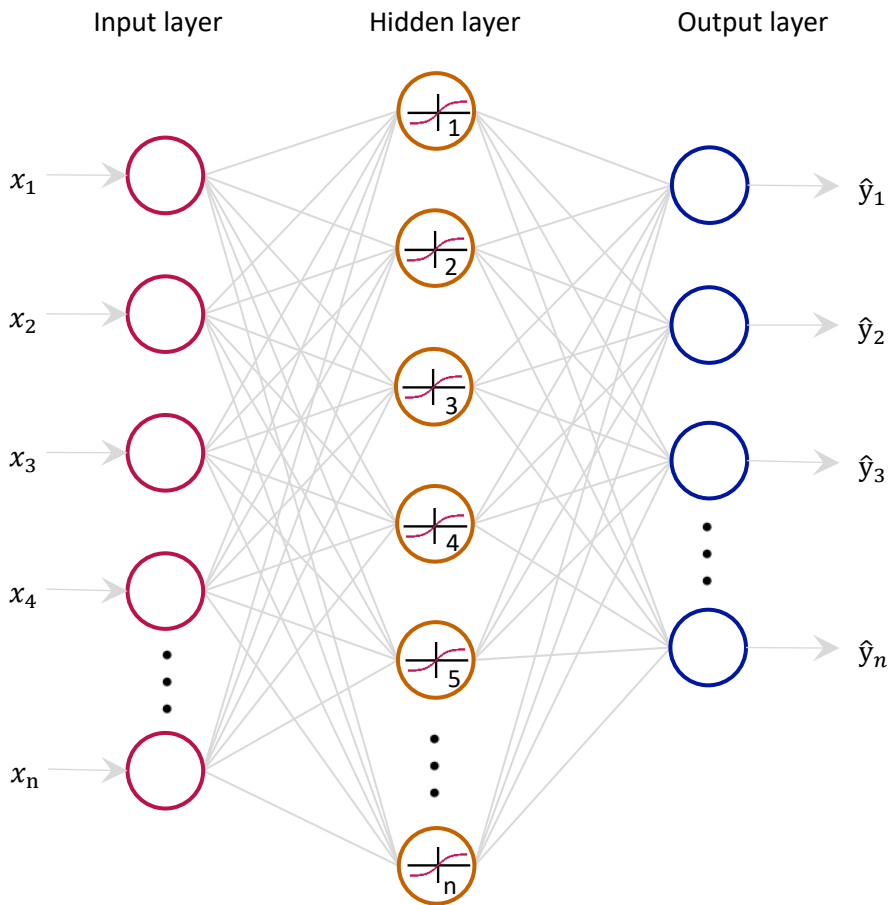


Figure 9. ANN scheme. Where x_n are the input variables that enter in the input layer and \hat{y}_n are the outputs from the output layer. The hidden layer contains the activation function.

However, in a large number of interesting cases, the neurons were not capable of solving problems. A classic example is the “exclusive OR” (XOR) problem, which was reported by Minsky and Papert in 1969 (Özbay *et al.*, 2007). The report resulted in a lack of enthusiasm and research in ANNs in the computer science community for almost 20 years (Basheer and Hajmeer, 2000). The recession was followed by regeneration of ANNs with the introduction of the ANN models of Hopfield during the 1980s, who popularized the MLP by employing the training algorithm of backpropagation proposed by Werbos in 1974 in his doctoral thesis at Harvard University, and the interest of the scientific community returned (Nastos *et al.*, 2013). This interest reborn was coupled with the rapid growth of computing

capabilities. Nowadays, the MLP can solve different nonlinear problems, including the XOR case with a quite simple network structure (Singh, 2016).

The task of an MLP is to get the desired output to the inputs of certain types. With this goal in mind, the learning of perceptron is performed. The inputs and the desired outputs are loaded to the scheme, and the error of the response is determined. The parameters of the system, the weights in the inter-connections, are changed during the learning phase of the ANN to diminish the difference between the desired and the real output (Negrov *et al.*, 2017).

The neural networks can be classified into two classes according to its architecture and interconnection between neurons: feed-forward networks and feed-back (recurrent) networks (Tealab *et al.*, 2017). The most popular ANN in the studies is the multilayer feed-forward network, in which the neurons are organized into series of layers, and information signal flows through the network solely in one direction, from the input layer to the output layer (Tkáč and Verner, 2016). Meanwhile, in an ANN where the outputs of some neurons are feed back to the same neurons or neurons in preceding layers are called recurrent neural networks (RNN). This feed-back enables a flow of information in both forward and backward directions (Basheer and Hajmeer, 2000). This type of neural network is more focused in time series predictions thanks to using as the input data part of the output data from the previous element in the series, such as forecast financial time series (Cavalcante *et al.*, 2016; Li *et al.*, 2018).

While MLP is more focused on solving nonlinearly separable problems, classification, and approximate continuous functions (Haykin, 1999), there is another architecture of neural networks based on the number of hidden layers; when the ANN consists of several of perceptrons layers, the term deep neural network (DNN) is employed (Li *et al.*, 2018). These types of neural networks are employed by excellence for image analysis and classifications, including videos (Antipov *et al.*, 2016; Cheng *et al.*, 2018). However, for most of the classification and predictions problems, there is no need to use more than one hidden layer (Heaton, 2008), in some cases augmenting the number of hidden layers can lead to the detriment of performance (de Villiers and Barnard, 1993).

In an MLP, adding hidden layers and neurons are generally limited for a few hidden layers. Adding more layers not only increases the complexity of the model and the computational cost, but it also does not ensure more accuracy or better overall performance of the model. Experimental results indicate that ANNs with two hidden layers are prone to perform more inaccurately in comparison to

one hidden layer ANNs (de Villiers and Barnard, 1993). Increasing the number of hidden neurons, in forms of a neuron inside the hidden layer or increasing the raw number of hidden layers itself, can dismiss the learning process leading to less accurate models (Karlik, 2011).

ANN, especially in complex models with several neurons and layers, have the cost of the training time (Moraes *et al.*, 2013). Although new central processing units (CPUs) are faster than ever, have multicore structures (2 or 4 cores in general) and instructions for high-performance parallel processing (Bergstra *et al.*, 2010; Vanhoucke *et al.*, 2011). Moreover, graphics processing units (GPUs) can be used for faster ANN implementations, due to higher capacity of parallel computing by the multicore system with hundreds of cores (Oh and Jung, 2004; Yang *et al.*, 2011). GPU implementation of ANN can go up to 50 or 60 times faster than standard CPU implementations (Cires *et al.*, 2003; Ramachandran *et al.*, 2015).

Regarding ANN frameworks, there are free and open source libraries such as Caffe, Theano, Torch, and TensorFlow, among others, which are widely spread (Abadi *et al.*, 2016; Rampasek and Goldenberg, 2016). The importance of being free and open source lies in the affordability for use in academics and research. TensorFlow, one of the frameworks that most interest has caught in developers in the last few years, is developed by the Google Brain Group, part of the Google Machine Intelligence Research Institute (Qin *et al.*, 2019), and the use of the framework is increasing research since it was released in 2015 (Abadi *et al.*, 2016; Cabañas *et al.*, 2019; Hazan *et al.*, 2018; Kulkarni *et al.*, 2018; Vázquez-Canteli *et al.*, 2019; Zhang and Kagen, 2017).

Several ANN studies have been conducted in distinct disciplines with overwhelming results; in the photovoltaic energy industry to forecast the power generation of system under diverse climatic scenarios (Bouselham *et al.*, 2017; Ding *et al.*, 2011; Veerachary and Yadaiah, 2000); in the industry sector to control and monitoring of production process, such as distillation (Singh *et al.*, 2007), lifetime prediction of machinery (Tian, 2012), fault diagnosis in elements of a power installation (Bi *et al.*, 2000) and controlling robotics machinery (Zhao *et al.*, 2014). In medicine, ANNs are useful for the diagnosis of cardiovascular diseases, diabetes, cancer and tumors (Jiang *et al.*, 2010; Singh *et al.*, 2015). In the finance sector, to schedule the energy loads for a more economical consumption and better energy prices (Alanis, 2018; Chen *et al.*, 2001), bankruptcy prediction (Li *et al.*, 2018; Tkáč and Verner, 2016; Zhang *et al.*, 1999) and prices and time series prediction (Abhishek *et al.*, 2012; Göçken *et al.*, 2016; Kaastra and Boyd, 1996; Moghaddam *et al.*, 2016;

Patel *et al.*, 2015); in the telecommunication sector, neural networks can treat and recover digital information in channels when a data loss occurs (Das *et al.*, 2014; Panda *et al.*, 2015), to cite some of the multiple examples of ANN applications.

2.5.3 Activation functions

Activations functions are an essential component of ANNs. In a perceptron, the weighted inputs are summed and passed through a limiting function, which scales the output to a fixed range of values. The output of the limiter is then broadcasted out of the neuron to the next layer (Ozturk and Karaboga, 2011). In general, S-shaped functions, such as sigmoid and hyperbolic tangent, are adopted for the activation. Once the activation function is performed, a neuron sends its activated value to the other neurons through the connections (Tsai and Wang, 2001; Zamanlooy and Mirhassani, 2014).

The activation function is typically nonlinear, and it ensures that the entire network can estimate a nonlinear function which is learned from the input/output data pair (Moraes *et al.*, 2013). Nonlinear activation function implementation achieve higher accuracy and improves the learning and generalization capabilities of ANNs in contrast to the linear primitive functions (Zamanlooy and Mirhassani, 2014).

Usually, neurons in the hidden layer are the ones that perform an activation function; output neurons, in general terms, return a weighted summation of the previous layer output without any transformation. It has been reported that a nonlinear activation function in the output layer failed to improve the performance and accuracy of an ANN (Yonaba *et al.*, 2010). There are several activation functions, with a diverse range of bounded outputs; according to the problem to solve, one or other activation function performs better (Gautam and Ravi, 2015). Some of the main activation functions are described in the subsections below.

2.5.3.1 Hardlim function

The hardlim function, also known as a binary step function, was the first activation function presented by McCulloch and Pitts, fathers of the artificial neuron. The function gives binary output values, "0" or "1", "0" when the signal

value is less than “0” and “1” when the signal value is equal or greater than “0”. The formula is presented in **Eq. 2**.

$$f(x) = \text{hardlim}(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases} \quad \text{Eq. 2}$$

The function plot is shown in **Figure 10**. The hardlim function presents two possible stages according to the input value into the activation function, thus the name of the binary step. Regardless of how negative or positive the input value is, the logic of this function will always return “0” or “1”.

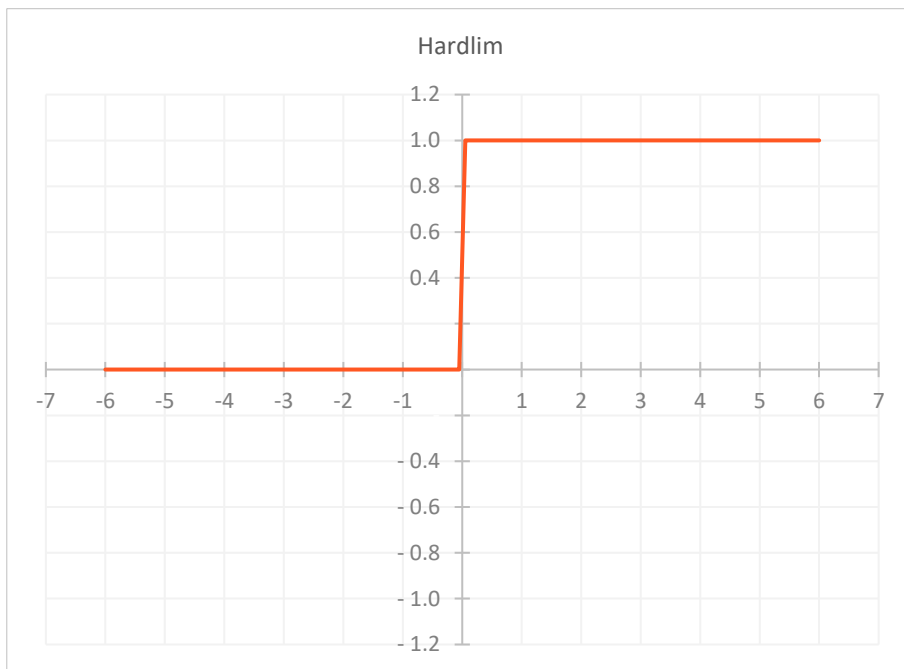


Figure 10. Hardlim activation function plot.

Hardlim is rarely used nowadays since there is no linear relationship between inputs and output patterns in most modern problems of interest (Abhishek *et al.*, 2012). Hardlim activation function was not very effective to solve several problems in comparison with the S-shaped functions in a comparative study using several datasets from the University of California Irvine machine learning repository (Zhang and Suganthan, 2016).

2.5.3.2 Sigmoid function

The sigmoid activation function is one of the most used in ANNs, and it rapidly replaces the hardlim function. There are some variations of the sigmoid functions, but the one known as logistic sigmoid is often used in studies and applications (Menon *et al.*, 1996). The sigmoid activation function produces positive numbers only between “0” and “1”, and the formula is presented in **Eq. 3**. This function is also known as the S-shaped function, and its bounded output values make it one of the most useful in training ANNs (Sibi *et al.*, 2013).

$$f(x) = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad \text{Eq. 3}$$

ANN with sigmoid activation functions was used to solve several issues such as inference of rivers water quality (Palani *et al.*, 2008); electricity power demand prediction (Manohar and Reddy, 2008); and solar air-heater modeling (Ghritlahre and Prasad, 2017), among others. The typical S-shaped graph of this function is shown in **Figure 11**.

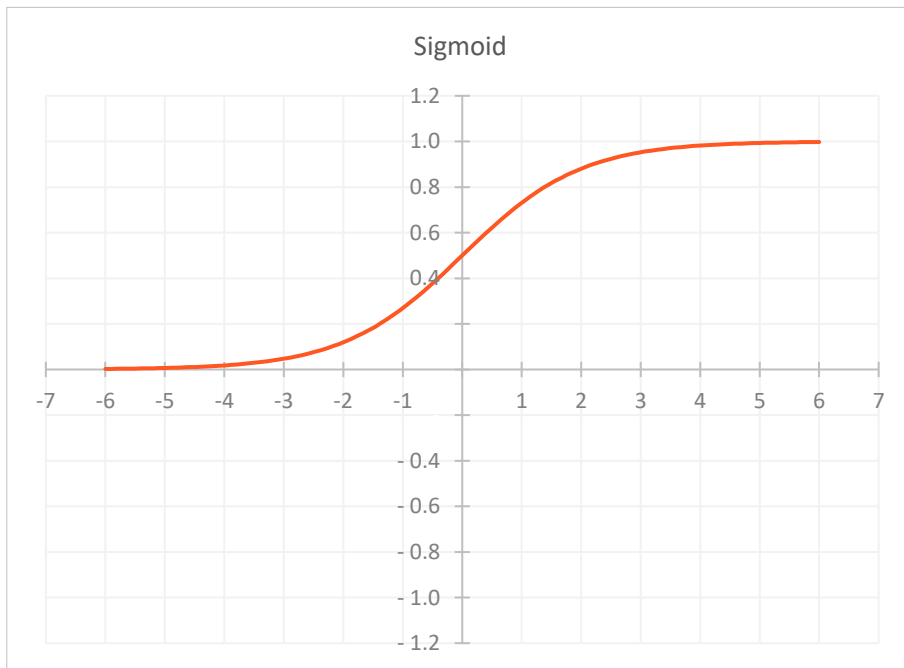


Figure 11. Sigmoid activation function plot.

2.5.3.3 Hyperbolic tangent function

The hyperbolic tangent function (\tanh) is similar to a sigmoid function, it is also an S-shaped function but with a subtle difference; its outputs bounds are between “-1” and “1” (Karlik, 2011). The \tanh formula is presented in **Eq. 4**.

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad \text{Eq. 4}$$

The function gives output values from “-1” to “1”, returns “0” when the input is “0” and as more positive or more negative the input is, the proportionality of the output decreases creating rough asymptotes. The function plot is shown in **Figure 12**.

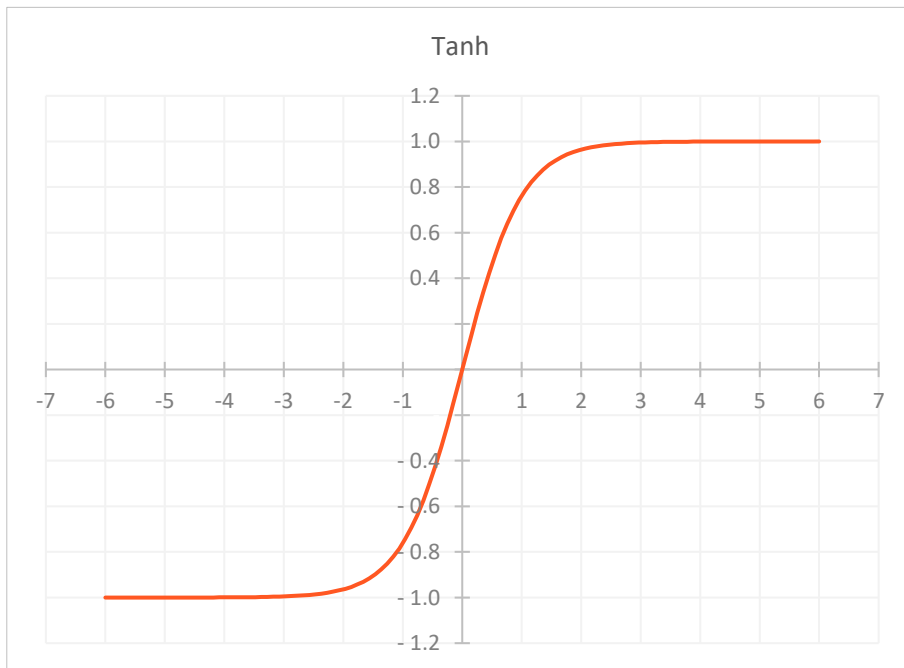


Figure 12. Hyperbolic tangent activation function plot.

Some examples of ANN with \tanh activation functions are the tracking of maximum power point for an automatic adjustable photovoltaic system (Punitha *et al.*, 2013) and markets index predictions (Guresen *et al.*, 2011).

2.5.3.4 Softsign function

The softsign is similar to the tanh function but might behave differently in terms of saturation because of its smoother asymptotes, polynomial instead of exponential (Glorot and Bengio, 2010). The softsign formula is presented in **Eq. 5**.

$$f(x) = \text{softsign}(x) = \frac{x}{1 + |x|} \quad \text{Eq. 5}$$

The softsign function is similar to the tanh function, giving output values from “-1” to “1”. The function plot is shown in **Figure 13**. This function is relatively novel, first introduced by Bergstra *et al.* (2009), proposing this function for better learning of image features. The applications of this activation function are mainly for image classification problems (Ertam, 2017; Kakuda *et al.*, 2019) but it can also be employed for inference problems such as indoor temperature forecasting (Romeu *et al.*, 2013).

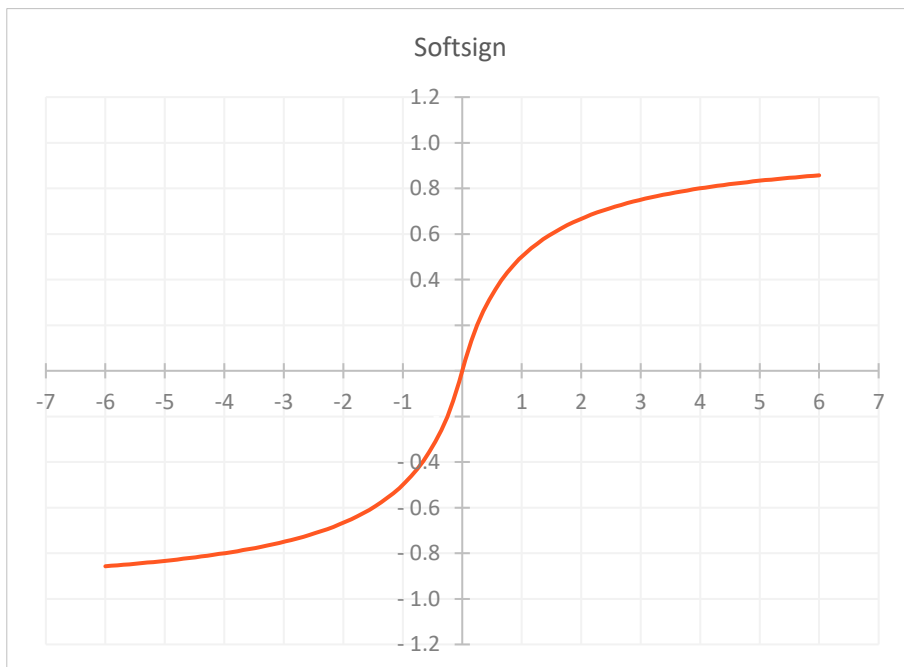


Figure 13. Softsign activation function plot.

2.5.3.5 Softmax function

The softmax function has been extensively used in ANN, especially in DNN architectures, is one of the few activation functions placed in the output layer due to its characteristics (Zheng *et al.*, 2015). The softmax has a classifier character since the function returns a number that can be interpreted as the probability of a particular class (Dong *et al.*, 2015). The softmax formula is shown in **Eq. 6**.

$$f(\vec{x}) = \text{softmax}(\vec{x}) = \frac{e^{x_i}}{\sum_{j=1}^J e^{x_j}} \quad \text{for } i = 1, \dots, J \quad \text{Eq. 6}$$

Softmax ensures a probability distribution, and it has the properties that every single output (\hat{y}_i) for a set of observations (\vec{x}) fit the condition $0 \leq \hat{y}_i \leq 1$, and the sum of outputs ($\sum_{i=1}^J \hat{y}_i$) is equal to “1”, as it required for probabilities (Bishop, 1996). Softmax activation functions in the last layer of a text analysis ANN gives the probability of appearance of words in a text string (Mikolov *et al.*, 2010; Tran *et al.*, 2014). In the same way, it computes the probable labels in sentiment analysis of tweets (Severyn and Moschitti, 2015) and gives a class probability of given images (Nguyen *et al.*, 2015).

2.5.3.6 Rectified linear unit function

Sigmoid neurons were for a long time the biologically more plausible model for neurons. However, the recent work in ANNs shows that rectified linear unit (relu) function is an even better model of biological neurons and yield equal or better performance than other activation functions despite the hard nonlinearity and non-differentiability at zero. Rectifier networks are not only beneficial to image classification tasks, in which it is the principal function but also might yield powerful tools in the future (Glorot *et al.*, 2011; Maas *et al.*, 2013). The relu activation function formula is presented in **Eq. 7**.

$$f(x) = \text{relu}(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad \text{Eq. 7}$$

Since it was first proposed by Nair and Hinton (2010), relu functions proved to be efficient compared with other alternatives in image analysis. This function was able to learn specific features that are better for object recognition. The relu

function is more accurate than tanh, sigmoid, and softsign activation functions for image classification (Ertam, 2017). Therefore, relu layers are preferred for image analysis (Zagoruyko and Komodakis, 2015), although, they can also be applied for video analysis (Ebrahimi Kahou *et al.*, 2015) and automatic speech recognition (Zhang and Woodland, 2016). The relu function plot is shown in **Figure 14**.

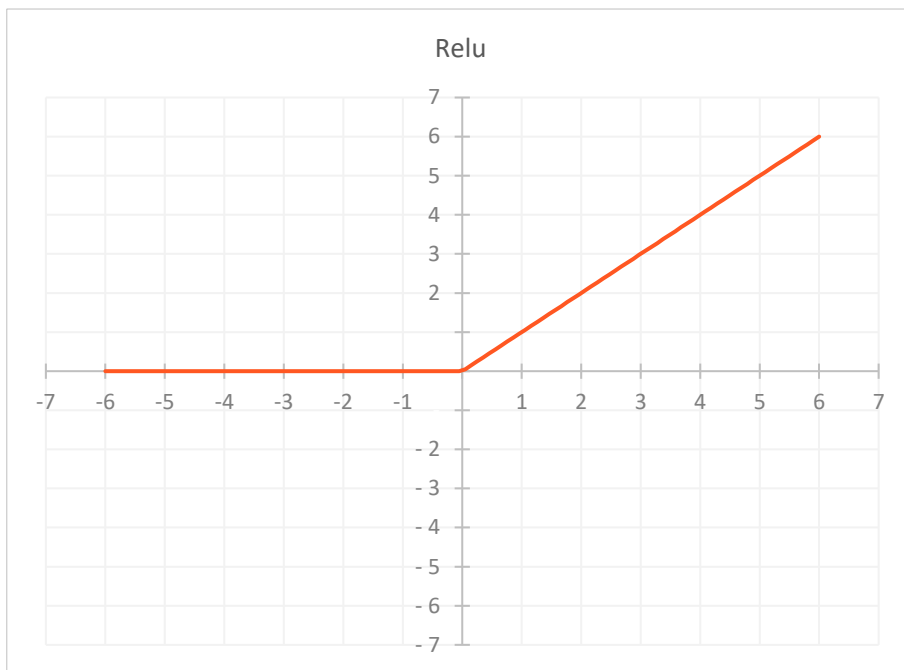


Figure 14. Rectified linear unit (relu) activation function plot.

2.5.4 Phases of artificial neural networks

ANNs have rapidly entered into the panorama as a powerful tool to solve problems, especially those highly nonlinear ones that traditional models are not able to solve or do not achieve the accuracy needed (Bahrammirzaee, 2010; Huang and Babri, 1998). The correct implementation of ANN is a meticulous process and can be accomplished by following the appropriate methodology, which ensures the maximum performance of the algorithms and the ability of the ANN to solve the problems that arise.

The process can be divided into mainly 3 phases. In the first phase, generally, the data is preprocessed before entering the ANN input layer. The second

phase is called the learning phase; here, the training and testing are performed. These two different processes are performed in parallel and should not be separated; the abstraction capability of the model is developed during this phase. Finally, in the third phase, the trained model is validated in terms of goodness and fit obtained in new inferences of ANN. In the following subsections, each of these phases will be detailed.

2.5.4.1 Data preprocessing

Before the learning phase of an ANN takes place, the dataset needs to be conditioned. A correctly partitioning of the available input-output data, predictors and response variables, into three different datasets such as the training, test and validation sets must be done. These sets have specific roles for the successful implementation of ANNs (Altinay *et al.*, 1997). By now, this subsection will be focused on the preprocessing of the data, and the roles of each dataset will be detailed in the following subsections.

Regarding the input variables, if they are measured in different scales or if the standard deviations for the variables are different among them, a variable might dominate the following calculations due to its dimensionality (Martinez *et al.*, 2017). Hence, normalization or standardization procedures are the most viable alternative to deal with these types of bias. In the same manner, normalization of variables increases the computational efficiency of the ANN training process, which is also highly enhanced in terms of accurate results when it is coupled with preprocessing techniques (Nawi *et al.*, 2013; Specht, 1991).

Normalization of input data has the advantage of scaling the data in the same range of values for each input feature in order to minimize bias within the neural network for one feature to another, helping the system a better weight adjustment for each element (Jayalakshmi and Santhakumaran, 2011).

In neural networks, there are fundamentally two types of preprocessing methods. The first is to standardize the data by subtracting the mean and dividing by the standard deviation to make the data have a mean of “0” and variance of “1”, referred as “Z-score normalization” as is shown in **Eq. 8** (Priddy and Keller, 2005).

$$z_i = \frac{x_i - \mu}{\sigma} \quad \text{Eq. 8}$$

Where x_i is a given input value from a specific feature; μ is the mean of the feature; σ is standard deviation; and z_i is the normalized value for the given input.

The second (Eq. 9), is to rescale the data to a small interval from “0” to “1” using the “Min-Max normalization” (Eberhart, 2014).

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad \text{Eq. 9}$$

Where x_i is a given input value from a specific feature; x_{min} is the minimum observed value for the feature; x_{max} is the maximum observed value; and x_i' is the normalized or re-scaled value for the given input.

Input data normalization prior to a training process is crucial to speed up the subsequent computational calculations significantly and obtain better ANN output. Data normalization affects the performance error of parameter estimators trained to predict the value of several variables (Gnana Sheela and Deepa, 2013; Sola and Sevilla, 1997). The attempt to implement ANN without input normalization would cause the data dot product, between input data and weights, propagation to be driven by the larger input cases. Normalization provides the system with a realizable manner to assign the right weights to the input, based on the real merit of the input data and not just its absolute size or unit of measurement (Badiru and Sieger, 1998).

2.5.4.2 Learning phase

The learning process of ANNs is carried out by two processes running complementary one to each other, the training and testing phases. The learning process is a multivariable optimization problem which can involve hundreds or thousands of variables (Martínez-Martínez *et al.*, 2015). During the learning phase, the train and test sets are used to tune the model until it reaches the wanted output precision.

The training set is the largest dataset, and it is used by the neural network to learn the patterns present in the data. The testing set, ranging in size from 10% to 30% of the training set, is used to evaluate the generalization ability of a supposedly trained network (Kaastra and Boyd, 1996). The data in the test set is used to supervise the training process and avoid memorization or “overfitting”, in other terms, giving good outcomes only for trained data (Dündar and Şahin, 2013). Hence, the training phase can be stopped at the point where the network has the smallest testing error; if the testing error starts to increase while the training error decreases, overfitting is taking place (Zhang and Morris, 1998).

The “best” network on training and testing data is the network that best passes the learning phase. However, its performance is finally qualified in the validation data (Zhang *et al.*, 2003). In other words, ANN must be trained, tested, and subsequently validated (Belgrano *et al.*, 2002; Patanaik *et al.*, 2018). Several studies use the test set as a validation set for the final performance assessment of the model, not in the testing phase rightly said (Khashman, 2010; Trejo-Perea *et al.*, 2009; Yang *et al.*, 2009). Other studies do not use all datasets or the test or the validation dataset (Drucker *et al.*, 1993; Kong *et al.*, 2016; Leu *et al.*, 2001). This confusion may be due that ANN is a particular technique that needs these three datasets, for train, test and validate the model, meanwhile, for other predictions methods, just two datasets are usually demanded, a train and a test set (James *et al.*, 2013).

In the training phase, the mathematics behind the abstraction capability of ANNs is called the learning algorithm. The most widely used learning algorithm is the backpropagation gradient descent, its popularity is presumably caused by simplicity, universality, and good availability in libraries (Phansalkar and Sastry, 1994; Tkáč and Verner, 2016). There are other learning algorithms such as genetic algorithms and particle swarm algorithms, these algorithms are not as widely implemented by machine learning libraries as the gradient descent but still an alternative method to study (Rios and Sahinidis, 2013).

The purpose of the backpropagation training is to iteratively change some parameters of the neurons in a direction that minimizes the error, or as it is defined, the error function — which basically performs the difference between the desired output and the actual outcome of the ANN across all the training and testing patterns (Örkücü and Bal, 2011). The ANN has unknown parameters, weights, and bias, just as the perceptron does, and these parameters are one of the responsible for the abstraction capability of the model. The objective of the learning phase is to

seek values for the unknown parameters of the weights and bias to make the model fit the training data well. Usually, starting values for weights are chosen to be random values near zero, for instance from “-1” to “1”. Hence the model starts nearly linear and becomes nonlinear as the weights change (Cao *et al.*, 2015; Hastie *et al.*, 2009).

The backpropagation procedure computes the gradient of an objective function, the error function, concerning the weights and bias of a multilayer stack of modules. The backpropagation is nothing more than a practical application of the chain rule for derivatives, and it can be defined by **Eq. 10**:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial \hat{y}_{ij}} \times \frac{\partial \hat{y}_{ij}}{\partial z_{ij}} \times \frac{\partial z_{ij}}{\partial w_{ij}} \quad \text{Eq. 10}$$

Where:

E is the error or cost function.

w_{ij} is a certain weight of a certain synapsis.

y_{ij} is a certain output of a neuron.

z_{ij} is the product of the weight and a input variable.

The chain rule is applied in functional dependence relationship — when a function depends on other function. For instance, the error function (E) depends on the output of a neuron (\hat{y}), that at the same time depends on the product of the input variables (z), making the link to the final parameter of this chain the weights and bias (w). The above equation describes the error gradient respect a specific weight correction, but it is also valid for a particular bias value since they are adjusted in the training phase in the same way. For the backpropagation, all the individual weights, bias, and errors are computed and added to obtain the cost function gradient of the ANN and adjust the parameters according to the calculus (Rosenbaum and Johnson, 1984).

The key insight is that the derivative or gradient of the error function of ANN can be computed by working backward from the gradient concerning the output layer to subsequent layers. The output of the initial network model and corresponding error is computed, then, at each iteration, new weights and biases are determined to add or to subtract a value to them, this value is called learning rate. Again, new outputs and errors are determined, and if the new error is above the previous one, the new weights and biases are rejected, and the fixed value is again added or subtracted according to the gradient (Rusk, 2015; Singh *et al.*, 2015).

Training the network involves adjusting the connection weights to correctly map the training set to obtain a desirable output, at least to within some defined error limit, it is basically an optimization problem (Palani *et al.*, 2008). In effect, the network learns from the training set; if the training set is satisfactory and the training algorithm is effective, the network should then be able to correctly estimate the output even for the inputs not belonging to the training set — this phenomenon is termed as “generalization” (Bishop, 1995; Punitha *et al.*, 2013; Wang *et al.*, 2006).

The training phase is a critical part of the use of neural networks. For a given problem, the network training is supervised, given the fact that the target for each input pattern is always known a priori. During the training process, the patterns or examples are presented to the input layer of a network (Zhang *et al.*, 1999). With the gradient descent backpropagation algorithm, a gradient descent search is performed — it measures the output error and calculates the gradient of the error by adjusting the weights in the descending gradient direction (Behrang *et al.*, 2010; Sharma and K. Venugopalan, 2014).

The randomly generated weights and the bias are adjusted by a magnitude called learning rate. The effectiveness and convergence of the training algorithm depend significantly on the value of the learning step, the optimum value of the learning step is system-dependent and varies according to the problem and other ANNs related features. For systems that possess broad minima, a large learning rate value will result in a more rapid convergence. Meanwhile, in a system with a narrow minimum, a small learning rate value is more suitable. There are no general rules to obtain an optimal learning step; typically used values are 0.9, 0.25, and 0.05 (Rui and El-Keib, 1995).

Randomly initialized weights achieve much faster learning speeds; that is why the weights and bias are usually randomly generated by the computer (Cao *et al.*, 2018). The backpropagation gradient descent is the rule by which the weights are modified, the reason why it is called the stochastic descent method is that the modifications are performed in an “a priori unknown” sets of weights and bias (Amari, 1993). The backpropagation algorithm is one of the most powerful supervised learning algorithms but if it is not well employed, the ANN will not converge to the minimum error point or will fall into a local minimum, not the global minimum (**Figure 15**) and the convergence speed can be reduced (Bi *et al.*, 2005; Ding *et al.*, 2011; Örkücü and Bal, 2011).

Other strategies to ensure better performance of the training algorithms are the stopping criteria and the dataset number. The stopping criteria can involve stopping after a certain amount of runs through all of the training data and also stopping when the total target error reaches some low level, for example, less than 2% during the learning phase (Palani *et al.*, 2008). Monitoring the error in the train and the test sets is also a good strategy for early stopping; if the testing error starts to increase in comparison to the training error, early stopping the training prevent the overfitting (Tetko *et al.*, 1995; Zhang and Morris, 1998).

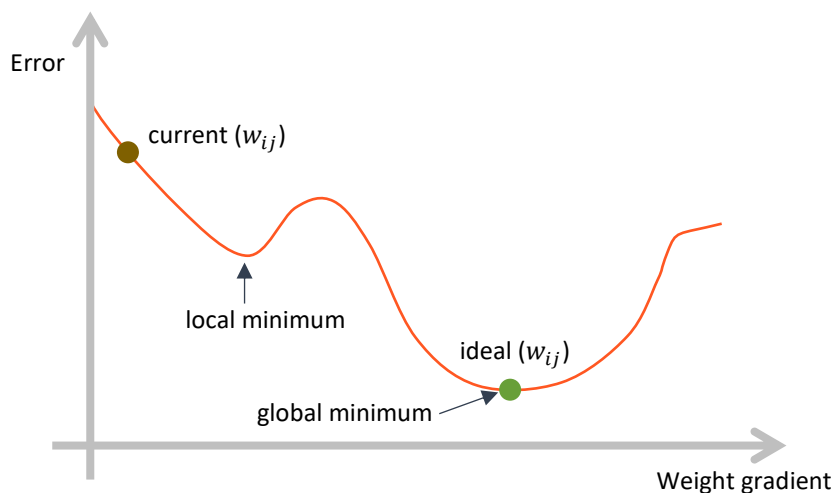


Figure 15. Example of the error function with respect to the weight gradient during the backpropagation algorithm execution.

Regarding the number of samples in the datasets, generally speaking, a large number of input samples is necessary to avoid overfitting (Ramachandran *et al.*, 2015; Wang *et al.*, 2006; Yosinski *et al.*, 2014). There is not a recommended fixed number of training samples, there is a direct relation between large datasets and proper learning levels of ANNs (Kourou *et al.*, 2015; Krizhevsky *et al.*, 2012), a good indicator can be that the number of training samples should be larger than the number of hidden neurons or neurons in the hidden layers (Huang *et al.*, 2004).

Performing proper training is a key for the ANN; avoiding the overfitting of the network is crucial for that purpose. Overfitting affects the generalization capability of ANN, and thus affects the prediction accuracy (Tian, 2012), and in the other hand, early stopping is better in terms of performance-to-cost ratio than

conventional stop mechanisms (Iyer and Rhinehart, 2000). The number of samples in the dataset, early stopping by reaching a certain number of iterations, error tolerance on the training set, and monitoring the testing error should be considered together to overcome the overfitting problem and not in isolation (Chen *et al.*, 2005).

2.5.4.3 Validation phase

After the training phase has been finalized, ANN performance is evaluated over the validation set (Ahmadi, 2011; Altinay *et al.*, 1997; Jafar *et al.*, 2010). The validation set must be composed by samples not included in the learning phase, i.e., into the train and test sets, for appropriate model evaluation (Basma and Kallas, 2004; Tian, 2012). In terms of the proportion of data needed for each dataset, the division into three parts follows the following relation: 70% for the learning process, 15% for test phase and 15% for validation phase (Bal and Buyle-Bodin, 2013), or 50%, 25%, 25% respectively, can be also an alternative (El Tabach *et al.*, 2007) as much as other proportions in similar ranges.

The lack of fit between the observed phenomena, the real values, and estimated values by the ANN during the validation phase indicates that the model should be re-trained, with the proper learning procedure and probably with more massive datasets to ensure more accurate results (Palani *et al.*, 2008; Yang *et al.*, 2009). The cost function, the error algorithm, is used to measure the model accuracy, and it is one crucial algorithm to the network development.

In ANNs, the most used cost functions are mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and cross-entropy (CE), among others. These functions can be used during the learning and the validation phases, meanwhile, functions such as correlation coefficient (R) and determination coefficient (R^2) are most appropriate exclusively for the validation phase (Ali *et al.*, 2017; Ay and Kişi, 2017; Cheng *et al.*, 2018; Moghaddam *et al.*, 2016; Sarkar and Pandey, 2015).

The MSE algorithm (**Eq. 11**) is generally selected because of its excellent performance, relative to the backpropagation algorithm (Zanetti *et al.*, 2007) and it is one the most popular cost function in ANNs (Adya and Collopy, 1998). MSE function appeals to be adequate when dealing with a single variable in a study, since

using multiple variables and dimensions for each one make the error comparison more difficult (Köksoy, 2006).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{Eq. 11}$$

Where n is the number of samples in the dataset; y_i is the real data; and \hat{y}_i is the estimation of the ANN. The same symbology will be applied to the remaining equations of this section.

The RMSE is calculated easily, in the same manner than MSE, and then taking the square root of the same (Eq. 12). The RMSE hence summarizes the overall error of the model, i.e., the precision of the model (Aptula *et al.*, 2005); the main difference between MSE and RMSE is that the RMSE is most useful when large errors are particularly undesirable (Saigal and Mehrotra, 2012) and the small errors tend to be less penalized.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad \text{Eq. 12}$$

The MAE, shown in Eq. 13, is an intuitively appealing measure. The MAE penalizes under and over-prediction respecting to the actual outcome and is the most natural measure of average error magnitude, is an unambiguous measure of average error magnitude and useful for inter-comparisons of model performance with different magnitudes and measure units (Willmott and Matsuura, 2005).

While the MAE gives the same weight to all errors, the RMSE penalizes variance — as it provides errors with larger absolute values more overall weight than errors with smaller absolute values. When both metrics are calculated, the RMSE is, by definition, never lower than the MAE (Chai and Draxler, 2014).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad \text{Eq. 13}$$

A modification of the MAE algorithm is the MAPE. This algorithm (Eq. 14) is computed through a term-by-term comparison of the relative error in the prediction concerning the actual value of the variable. Consequently, the MAPE is an unbiased statistic to measure the predictive capability of models (Wang *et al.*, 2009). MAPE is often used because of its intuitive interpretation in terms of relative

error, but just viable when the quantity to predict is above zero (de Myttenaere *et al.*, 2016); otherwise, a zero division is produced, and non-result can derive from this formula.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100 \quad \text{Eq. 14}$$

Regarding the CE function, this error method reflects the similarity between variables from the perspective of probability (Men *et al.*, 2016). This cost function (**Eq. 15**) is frequently with ANN with softmax activation function in the output layer (Dahl *et al.*, 2013; Liew *et al.*, 2016; Maas *et al.*, 2013).

CE has significant advantages over other error functions for probabilities problems, performs better on large and small target values because they tend to result in similar relative errors for both cases. The CE function performs better at estimating low posterior probabilities than squared-error functions (Kline and Berardi, 2005). The output of the softmax layer returns a series of probabilities, zero or positive values, and the CE penalizes all the results which do not match the current probability of the target.

$$CE = - \sum_{i=1}^n y_i \times \log \hat{y}_i \quad \text{Eq. 15}$$

The cost functions reviewed so far are eligible for both the learning and the validation phases. Nonetheless, some algorithms are exclusively used for validation of the model, such as the R and R². The R, as is shown in **Eq. 16**, investigates the degree of association between two variables, that is, it defines how much a given relationship is fitted by a straight line (Tripepi *et al.*, 2008).

$$R = \frac{n(\sum y_i \hat{y}_i) - (\sum y_i)(\sum \hat{y}_i)}{\sqrt{[n \sum y_i^2 - (\sum y_i)^2][n \sum \hat{y}_i^2 - (\sum \hat{y}_i)^2]}} \quad \text{Eq. 16}$$

The R² is an algorithm (**Eq. 17**) that is widely used in ANNs to assess the performance of models (Gholami and Fakhari, 2017), including approaches in regression, classification and predictions problems (Ghritlahre and Prasad, 2017; Moghaddam *et al.*, 2016; Palani *et al.*, 2008; Tetko *et al.*, 1995; Turan *et al.*, 2011).

$$R^2 = 1 - \frac{\sum_{i=1}^n [y_i - \hat{y}_i]^2}{\sum_{i=1}^n [y_i - \hat{y}_{mean}]^2} \quad \text{Eq. 17}$$

The cost function and the model evaluation are crucial tasks in the ANN assays. There is not a universal solution for better performance, and most of the work in creating good quality models is to test not only different network architectures, but also cost functions, and the performance in the validation samples. Once that several models are evaluated, it can be settled that the selected ANN is the most pertinent for solving the problem.

MATERIALS & METHODS

3. Materials and Methods

In this chapter, the materials used, and the methodology applied to the experimental phase of the thesis will be described. Following the established objectives, three main research topics were based on the application of ANNs to solve problems in the context of agriculture.

The first topic, regarding the microalgae production and management by using an ANN and light absorbance spectral signature to a rapid inspection of microalgae suspensions, is described in section 3.1. Concerning the second topic, soil fertility characterization through color assays, the methodology for inquiry the relationship between both variables using ANNs are detailed in section 3.2. Finally, the third topic, the procedure to the developing of virtual weather stations (VWS), for meteorological variables acquisition system, through interpolations of real data acquired from weather stations is treated in section 3.3.

3.1 Monoalgal and mixed algal cultures discrimination by using artificial neural network

In this section, the ANN assay to differentiate between monoalgal and mixed algal cultures will be described. For this experiment, only the light absorption spectra of the different microalgae cultures were used, without any additional equipment or image analysis algorithms. A scheme of the experiment is shown in **Figure 16**.

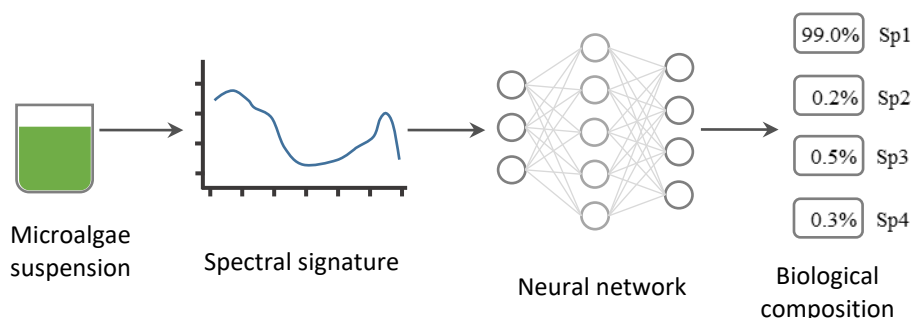


Figure 16. Microalgal and mixed algal cultures discrimination by ANN experiment scheme.

The objective of this work was to determine the feasibility of using such methodology to elucidate the presence of contaminants in monoalgal microalgae cultures and to approximate the microalgae species contaminating the cultures. The remaining subsections are organized as follows: microalgae species and growth conditions for the experiments (section 3.1.1), light absorption measure of monoalgal and mixed algal cultures (section 3.1.2), and the ANN used to elucidate the microalgae suspensions (section 3.1.3).

3.1.1 Microalgae species and growth conditions

The microalgae *Nostoc* sp., *Scenedesmus almeriensis*, *Spirulina platensis* and *Chlorella vulgaris* were obtained from the culture collection of “Las Palmerillas” Research Center (Almería, Spain) (Figure 17). These species were selected because they include two fast-growing green microalgae — *Chlorella vulgaris* and *Scenedesmus almeriensis*, and two distinct cyanobacteria — *Nostoc* sp. and *Spirulina platensis*; therefore, they cover a wide range of possible microalgae species.



Figure 17. Microalgae cultures collection of “Las Palmerillas” Research Center.

Continuous cultures of the different microalgae were cultivated under laboratory conditions in 2 L bubble-column photobioreactors at an irradiance of $750 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, adapting the described methodology (Gómez *et al.*, 2013). Arnon medium was used for the four microalgae, although in the case of *S. platensis*, it was enriched with $16 \text{ g}\cdot\text{L}^{-1}$ of sodium bicarbonate. The temperature of the cultures was

maintained at 25 °C, while the pH was controlled at 8.0 by the on-demand injection of CO₂. To perform the absorption measurements, samples from the reactors were collected and placed in batch mode inside 0.2 L flasks for ten days, with no CO₂ injection and at an irradiance of 250 μE·m⁻²·s⁻¹ to simulate variations in pigments and composition that can be found in batch cultures of each of the selected microalgae.

3.1.2 Light absorption measurements

Samples of *Nostoc* sp., *S. almeriensis*, *S. platensis* and *C. vulgaris* were taken from the reactors over five consecutive days. Light absorption measurements were performed on the monoalgal cultures. Additionally, light absorption measurements were carried out on paired cultures in proportions of 50%/50% and 75%/25% (**Figure 18a**). In total, 22 suspensions were prepared daily; these included both monoalgal and mixed algal cultures. Each suspension was measured 25 times. Over the five days, a total of 550 samples were measured to obtain sufficient data to train the ANN; the more significant the amount of data, the lower the model error and the higher the accuracy of the developed ANN (Ying-Jin *et al.*, 2004).

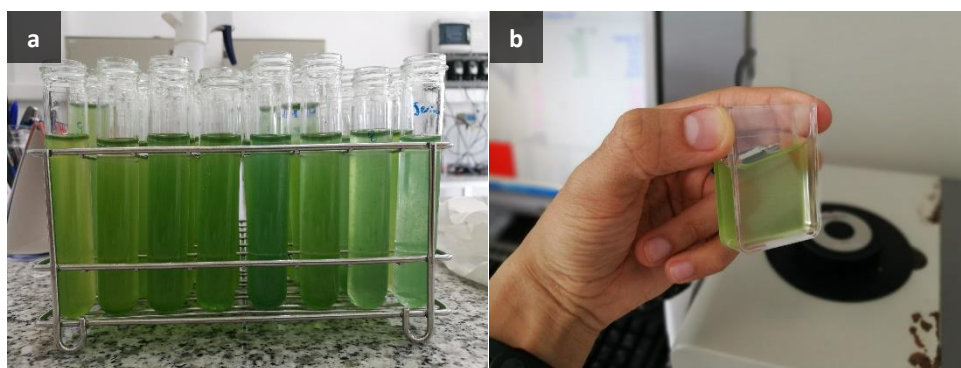


Figure 18. Microalgae suspension for monoalgal and mixed algal cultures (a); and sample for light absorption measure in the colorimeter (b).

In order to validate the ANN, additional measurements were taken over two further days for monoalgal and mixed algal suspensions. Mixed algal suspensions were of the same and different proportions to those used in the training phase; in addition, mixtures of 3 algal species were used. Measurements were

performed using a Minolta CM-3500d colorimeter - 31 light absorption data points were acquired per sample in the 400 nm to 700 nm range, in bandwidths of 10 nm. Each sample consisted of a 12 mL microalgae suspension, which was stirred and homogenized (**Figure 18b**).

3.1.3 Artificial neural network and data analysis

The absorbance values were converted to their relative frequencies using **Eq. 18**, where x_i is the relative absorbance for a specific bandwidth, a_i is the raw absorbance in a bandwidth, and $\sum a$ is the total light absorption of a microalgae suspension across the spectrum. Consequently, the relative absorbance represents the quantity of light absorption in each bandwidth without the cell concentration effect.

$$x_i = \frac{a_i}{\sum a} \quad \text{Eq. 18}$$

For ANN development, data were normalized in order to improve network performance, increase output accuracy, and reduce computational cost (Gnana Sheela and Deepa, 2013; Sola and Sevilla, 1997). The statistical, or Z-score, normalization technique as is shown in **Eq. 8** was applied to the relative absorbance data; where x_i is the relative light absorption datum (for each absorbance band per sample), μ is the mean of the whole dataset, σ is the standard deviation and z_i is the normalized datum.

The ANN was a fully connected feed-forward neural network, programmed in Python 3.6.5 language using the TensorFlow 1.8.0 machine learning library. After various configuration tests, the most straightforward structure — with less hidden layers and fewer neurons per hidden layers — and the activation function with the least training and testing errors was selected. Hence, the neural network architecture was 31 neurons in the input layer, one neuron for each absorbance datum acquired per sample, 45 neurons in the hidden layer with the hyperbolic tangent activation function and 4 neurons in the output layer with the softmax activation function to weight each microalgae species in suspension; a schema is presented in **Figure 19**.

The optimizer was the gradient descent algorithm with the CE cost function (**Eq. 15**); the CE was calculated for all samples in each dataset, and the error

was expressed as average cross-entropy (ACE), dividing the sum of CE by the number of cases in each set of data. In the training and testing phase, 10% of the input dataset was used for testing. The early stop, used to prevent overfitting, was applied before the testing error started to increase (Piotrowski and Napiorkowski, 2013). In order to validate the model, the ANN output was analyzed by performing regression using the expected microalgae concentration in suspension.

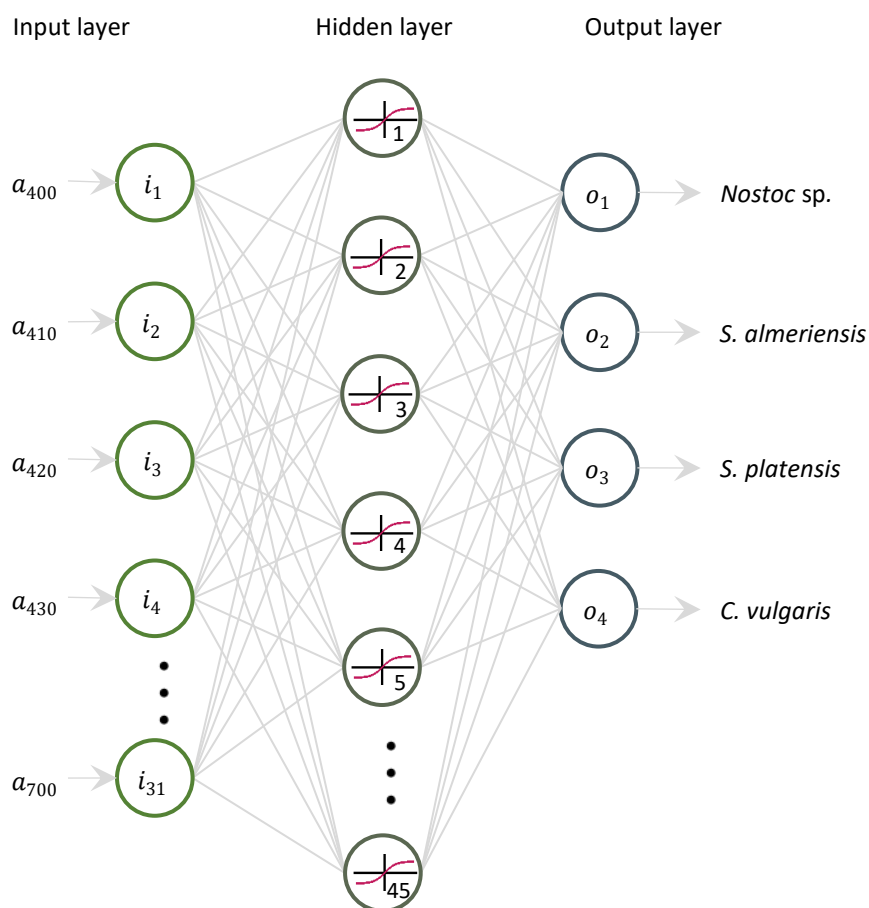


Figure 19. ANN architecture schema, with 31 neurons in the input layer, 45 in the hidden layer and 4 in the output layer.

3.2 Artificial neural network for soil color analysis and characterization

In this section, the soil characterization and color analysis by ANNs methodology will be detailed. For the experimental work, soil samples and digital color photographs were analyzed through ANNs to find possible relations between color and nutritional contents. The main objective was to develop a tool for soil color classification and analysis in order to develop a fast and inexpensive tool for soil management in agricultural lands. A scheme of the experiments is shown in **Figure 20**.

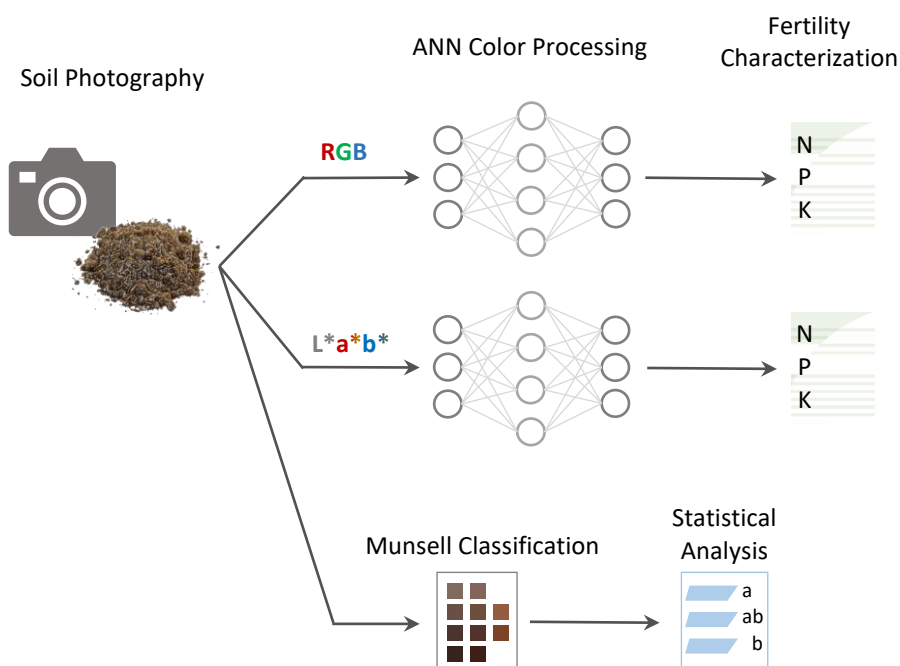


Figure 20. Soil color analysis and characterization by ANN experiment scheme.

The remaining subsections are organized as follows: study area and soil sample methodology (section 3.2.1), image acquisition system for soil sample photographs description (section 3.2.2), the CV software for soil color processing and RGB to L*a*b* conversion (section 3.2.3), Munsell soil color classification and statistical analysis (section 3.2.4), and the ANNs for the soil color characterization (section 3.2.5).

3.2.1 Study area and soil sampling procedure

The study area consisted in a 43.01 ha vineyard located between 40°00'55.25" and 40°01'24.99" N to 2°56'55.05" and 2°55'34.83" W in Tarancón, province of Cuenca (Castilla La Mancha, Spain). Previous mapping of apparent electrical conductivity (ECa) was performed for a representative soil sampling. The ECa was measured with a Veris Q2800 at depths of 36 and 90 cm in consecutive parallel paths with 10 m of separation between lines throughout the parcel, the ECa tend to be useful to sort information and identify sets of attributes with similar trends (Officer *et al.*, 2004).

The obtained information, one raster file for each depth (**Figure 21 a and b**), were reclassified using the Jenks natural breaks classification method (Chen *et al.*, 2013) — 3 classes were defined in each raster. The intersection of the classes in both depths was used to create new segments, and each soil sample was extracted from these new segments (**Figure 21c**). Hence, 174 sampling points were defined, 87 for each depth in order to get samples of each homogeneous segment of the parcel. The soil samples were analyzed in the laboratory for 13 parameters, such as the OM and total nitrogen (N), expressed in percentage, are measured using combustion autoanalyzer LECO TruSpec; phosphorous (P), potassium (K), calcium (Ca), magnesium (Mg), iron (Fe), copper (Cu), manganese (Mn), sodium (Na) and aluminum (Al), are measured by ICP-OES after acid digestion in nitric acid using microwave assisted digestion, and are expressed in mg/kg of soil; pH and electrical conductivity (EC) were measured in water (1:5 w/v) and are in dS/m.

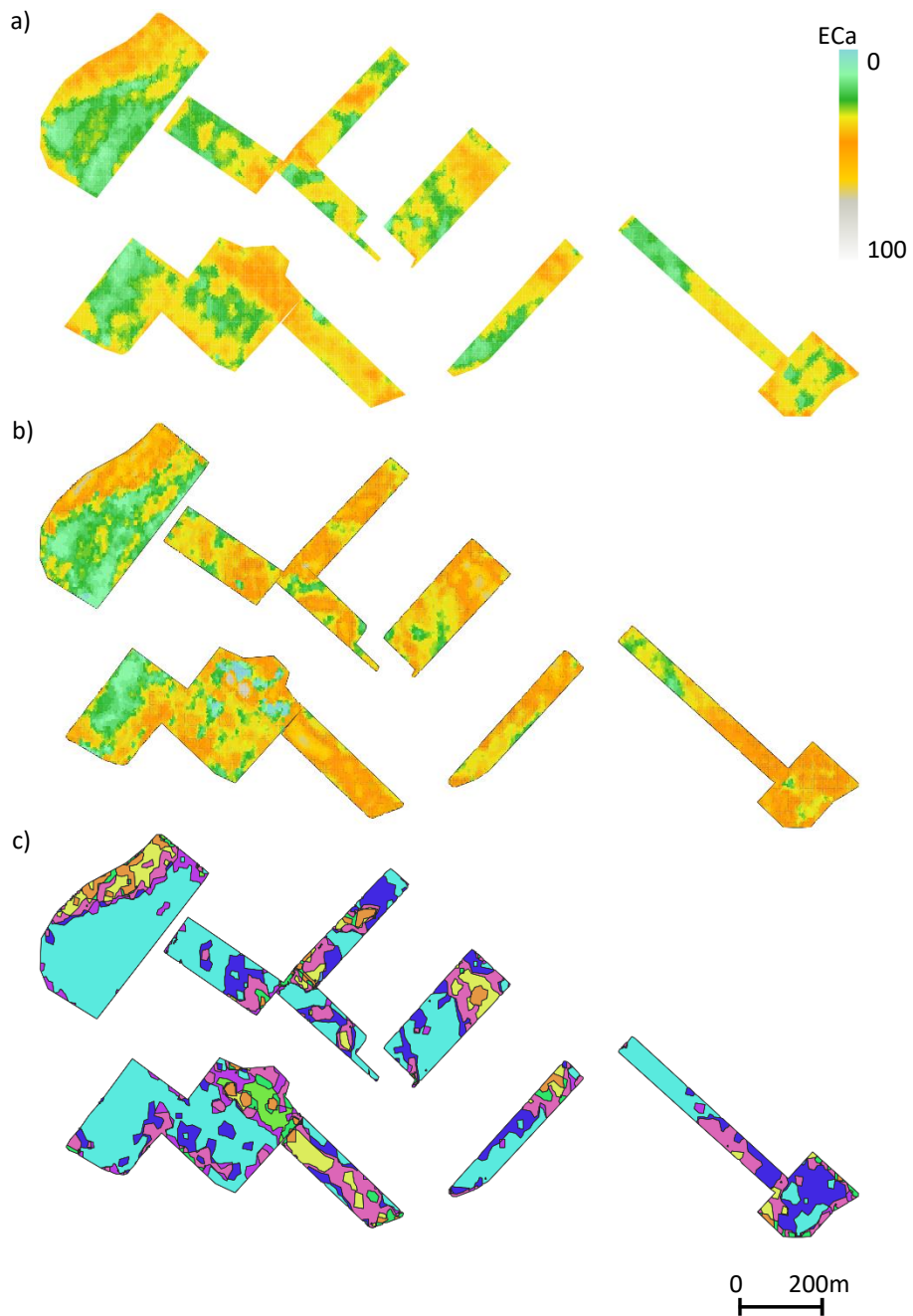


Figure 21. ECa raster map from the Veris data acquisition system at 36 cm (a) and 90 cm (b) depth; Segments for soil sampling generated through the processed raster data (c).

3.2.2 Image acquisition system

Soil digital photos were taken with an image acquisition system consisted of a Panasonic DMC-TZ70EG-S compact digital camera, with 12.1 Mpx, CMOS sensor and Leica DC Vario Elmar 24 mm lens. The illumination set was arranged by a pair of 6,500 K professional photoshoot light sources to provide uniform illumination to the sample (Mendoza *et al.*, 2006) (**Figure 22a**). The image acquisition system in controlled conditions is the most appropriate manner to take digital photographs, under non-controlled conditions the color can result less accurate and less appropriate for this experiment (Lapins *et al.*, 2013; Potočnik *et al.*, 2015).

The camera was mounted on a stand and remotely operated via the “Image App” smartphone application (**Figure 22b**). The camera aperture speed was set at 1/30s, ISO-80, luminosity f/3.3, and focal distance 4 mm. Soil samples were placed in Petri plates of 55 mm diameter, uniformly filled; 5 repetitions (plates) per soil sample were prepared, a total of 870 experimental units were photographed. Soil samples were dried at ambient temperature to avoid the interference of moisture content in the soil color expression (Sánchez-Marañón *et al.*, 2007).

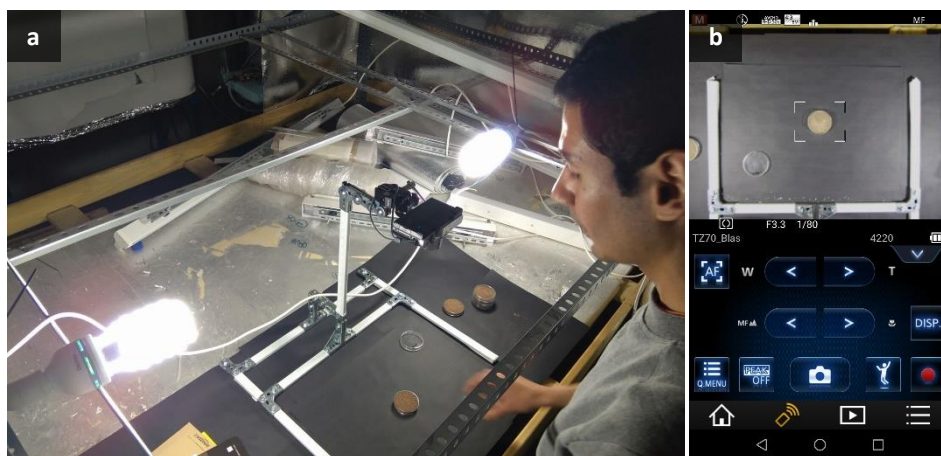


Figure 22. Image acquisition system; light and camera setup (a) and shooter camera app for smartphone (b).

3.2.3 Computer vision software for color conversion

In order to accelerate the process of color measurement of soil samples, a CV application has been programmed to analyze the color in digital photographs and perform the conversion from RGB system to the international standard CIELAB color system. The software and user interface were developed in NI LabVIEW 2018 using the Vision Module. This programming environment has the versatility to work with scripts from other programming languages, such as Python, and it is able to execute ANN's scripts with the advantage that it allows designing a user interface which enables the user to use an ANN without writing a single line of code.

The soil color conversion from RGB to $L^*a^*b^*$ was performed using an ANN following the methodology proposed (León *et al.*, 2006). To generate the model, the "X-Rite ColorChecker Classic" (**Figure 23**) was used. This colorchecker includes 24 color charts with natural, chromatic, and primaries colors in addition to a greyscale. The color charts are scientifically prepared and verified, see **Appendix C** for technical specs of the colorchecker. The RGB values from color chambers in digital color photographs were extracted and used as input in the ANN, and the $L^*a^*b^*$ values provided by the manufacturer were the output in the training and testing phases.

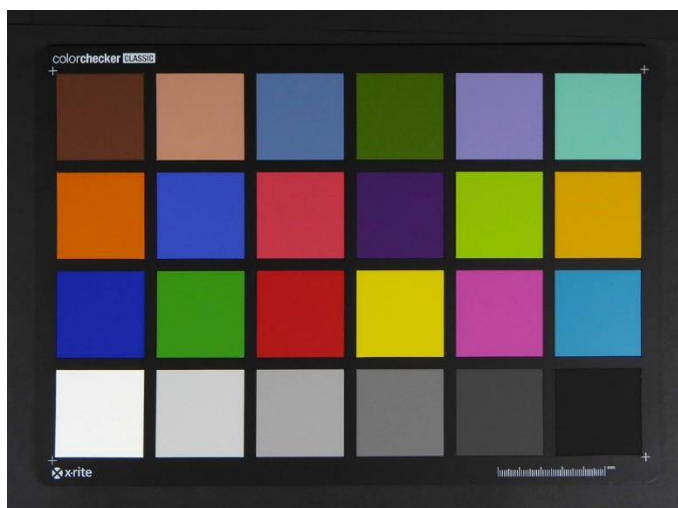


Figure 23. X-Rite ColorChecker Classic, color calibration card.

The ANN was composed of 3 neurons in the input layer, one for each RGB color element; 8 neurons in the hidden layer with the hyperbolic tangent activation

function; and 3 neurons in the output layer, one for each $L^*a^*b^*$ color element — a schema is shown in **Figure 24**. The early stop was applied to prevent overfitting. The ANN was programmed in Python 3.6.5 using TensorFlow 1.8 machine-learning library. The obtained model was applied to convert the soil RGB dataset into $L^*a^*b^*$ color.

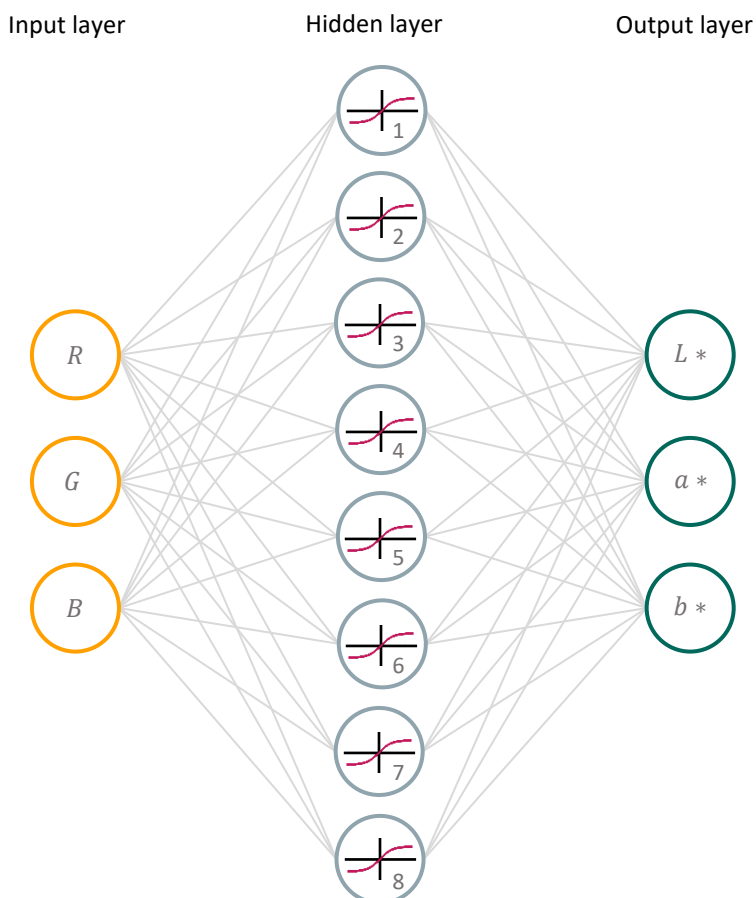


Figure 24. ANN architecture schema, with 3 neurons in the input layer, 8 in the hidden layer and 3 in the output layer.

3.2.4 Munsell soil color classification and description

To perform the Munsell color soil classification, a Munsell soil color charts book was photographed in the same image acquisition system described in section 3.2.2. A CV software was programmed in LabVIEW 2018 using the Vision Module to process the color Chart Book and extract RGB median values of each notation color in the book.

The RGB mean values of the 5 repetitions of each soil sample were computed to compare them with the RGB obtained from each Munsell color notation. The minimum Euclidean distance match algorithm was performed to compare the colors, as is shown in **Eq. 19**; where R_s , G_s and B_s are the mean of the RGB color values of the soil samples and R_M , G_M and B_M are soil chart RGB color of each Munsell notation. Each mean of the RGB values from the soil samples was compared with the RGB values from the Munsell colors, and the notation with the absolute minimum difference was taken as the soil Munsell color of the sample (Stanco *et al.*, 2011).

$$d = \sqrt{(R_s - R_M)^2 + (G_s - G_M)^2 + (B_s - B_M)^2} \quad \text{Eq. 19}$$

Before using the dataset, outlier points were removed based on fertility parameters, the soil samples classified according to their Munsell color notation were analyzed statistically; the Kolmogorov–Smirnov and Levene tests were performed, data normality and equality of variances were discarded according to the tests. The means were compared with the Kruskal-Wallis H test to verify differences between groups and the Mann-Whitney U test were used to compare the groups based on their Munsell hue color classification; all statistical tests were performed in R programming language version 3.4.4.

3.2.5 Soil properties characterization by artificial neural networks

For the ANN analysis, two datasets were prepared with the soil color values, one for the RGB values and one for the L*a*b* values, and the laboratory analysis for soil fertility parameters. Each dataset was used to train, test and validate ANNs. The data were normalized to improve the network performance (Gnana Sheela and Deepa, 2013), the Z-score normalization was applied to the datasets as

is shown in **Eq. 8**; where x_i is the datum of the variable, μ is the mean of the variable, σ is its standard deviation and z_i is the normalized datum.

The ANNs were fully connected feed-forward neural networks, with the RMSE cost function, programmed in Python 3.6.5 language using the TensorFlow 1.8.0 machine learning library. After various configuration attempts, the most straightforward structure — with less hidden layers and fewer neurons per hidden layers — and the activation function with the least training and testing errors was selected. The ANNs structures were set with 3 neurons in the input layer, one neuron for each color component; 1 hidden layer with 18 neurons, with hyperbolic tangent activation function; 13 neurons in the output layer, one neuron for each soil property to predict. For training and testing phases, 554 and 62 experimental units were used, respectively. The validation phase was carried out with 25 soil samples not previously shown to the ANN in training or testing phases, the mean color of the 5 repetitions of each soil sample was used to make the 25 data to the validation dataset.

3.3 Virtual weather stations for meteorological data estimations

In this section, the methodology to develop the VWS is described. This development was planned with the primary objective of evaluate and compare the accuracy of several interpolation algorithms, ANN approaches and the most used interpolation methods, to generate meteorological data and develop the VWS as an option to acquiring accurate data economically and straightforwardly. A scheme of the procedure is shown in **Figure 25**.

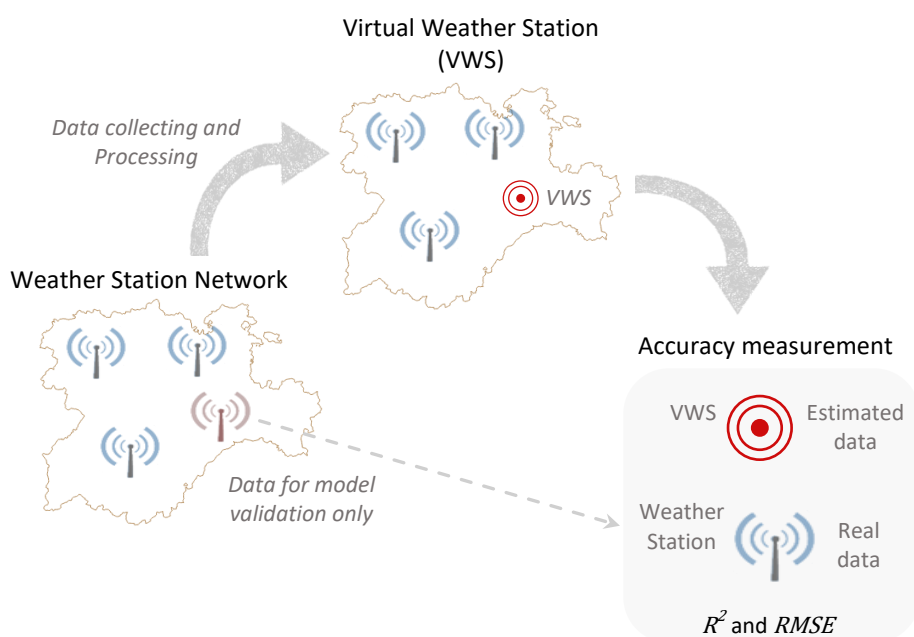


Figure 25. VWS development methodology scheme.

The remaining subsections of the VWS methodology are organized as follows: the meteorological data source (section 3.3.1), statistical analysis of the dataset (section 3.3.2), the development of the VWS and validation procedure (section 3.3.3), and the description of the interpolations algorithms employed (section 3.3.4).

3.3.1 Meteorological dataset

The data was obtained from the InfoRiego metrological station network records of the Agrarian Technological Institute of Castilla y León (Spain), the locations are represented in **Figure 26**. The meteorological records were downloaded from the file transfer protocol (FTP) server available; a Python script automates the download of the data by dates. The registers consisted of daily data summaries of 53 meteorological stations distributed throughout the territory of the autonomous community of Castilla y León. The records date from July 1st, 2017 to June 30th, 2018 to compile information for a whole year.

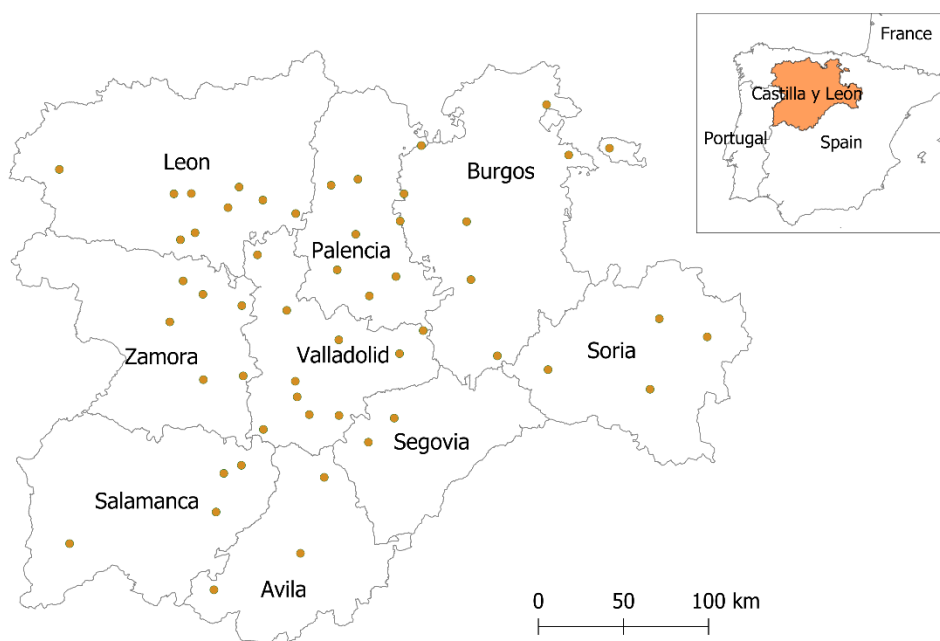


Figure 26. Meteorological station network of the InfoRiego program of the Agrarian Technology Institute of Castilla y León.

The meteorological observation records were grouped into a set of daily precipitation (Precip) in mm of rainfall, reference evapotranspiration (ET_o) in mm of water, daily mean air temperature (Mean Temp), maximum registered temperature (Max. Temp) and minimum registered temperature (Min. Temp) in °C, daily mean air relative humidity (Mean RH), maximum registered air relative humidity (Max. RH)

and minimum registered air relative humidity (Min. RH) in percentage (%), mean wind speed (Mean WS) in m/s and total solar irradiation (TSI) in MJ/m² from all the 53 weather stations. Additionally, the UTM X and Y geographical coordinates, in ETRS89 geodetic system, and the Z coordinate of the meteorological stations were added to the dataset to perform the interpolations.

3.3.2 Dataset statistical analysis

An exploratory data analysis (EDA) was performed to the whole dataset, the weather stations with reading errors and the days with less than 90% of the wheater station in the records were removed to ensure a minimum of observations per day. The dataset resulted in 18,234 sets of observations from 355 days. A statistical summary was generated with the purpose of describing the behavior of the studied meteorological variables in the period from July 2017 to June 2018. In the same manner, data from the summer season months (July, August and September) and from the winter months (December, January and February) were filtrated to analyze the registers in periods with higher peaks phenomena such as high or low temperatures, rainy periods and other seasonal meteorological events.

The means of the meteorological events during summer months, winter months, and the rest of the year were compared. The Kolmogorov-Smirnov and Levene tests were performed, data normality and equality of variances were discarded according to the tests. The means were compared with non-parametric tests; the Kruskal-Wallis H test to verify differences between groups and the Mann-Whitney U test was used to compare the groups based on the season of occurrence; all statistical analyses were performed in R programing language version 3.4.4.

3.3.3 Virtual weather station development

The VWS is composed of algorithms capable of getting and processing data from weather station networks and predict the possible meteorological state in a given location. For that, records from the InfoRiego station network were massively downloaded from the FTP server of the ITACYL through a Python script and used to interpolate values of daily Precip (mm), ETo (mm), Mean Temp (°C), Max. Temp (°C), Min. Temp (°C), Mean RH (%), Max. RH (%), Min. RH (%), Mean WS (m/s) and TSI (MJ/m²).

To evaluate the quality of interpolated data, a randomly selected weather station was daily removed from the dataset, and the results of the interpolations were compared with the actual records from these stations. The interpolations method approaches were through ANNs, deterministic and statistical interpolations methods, and random forest regression. All the interpolation algorithms were evaluated with the same validation dataset, and the accuracy was measured with the R^2 between meteorological observations and interpolated values, and the RMSE. All statistical tests were performed in R programming language version 3.4.4.

3.3.4 Interpolation methods

For the ANN approach in data interpolation, the performance of 5 ANNs with different activation functions such as hardlim, sigmoid, tanh, softsign and relu were programmed. The ANNs were multilayer perceptron, fully connected feed-forward neural networks, with the RMSE cost function — programmed in Python 3.6.5 language using the TensorFlow 1.8.0 machine learning library. After several attempts, the most straightforward structure — with less hidden layers and fewer neurons per hidden layers — and with the least overall training and testing errors was selected. The ANNs structures were set with 2 neurons in the input layer, one neuron for each geographical position coordinate (X and Y); 1 hidden layer with 15 neurons; and 10 neurons in the output layer, one neuron for each meteorological variable to interpolate; using this protocol for modeling, with the XY coordinates is possible to estimate the meteorological observations. The Z geographical coordinate was not used after verifying that its inclusion leads to less accurate results; selecting the appropriate input variables is a crucial step in ANN modeling (López *et al.*, 2005).

The dataset was filtered by day to create a daily re-trained model to each of the 355 days of the study. The daily subset of meteorological observations was divided into a train, test and validation sets; a randomly selected weather station, with its own set of measurements, was separated for validation and the remaining stations were used in the training and testing phases, with 10% of stations reserved for testing. The ANNs were re-trained and validated for interpolations every day with their datasets. The dataset with the XY geographical position and the meteorological observations were normalized to obtain better ANN performance (Gnana Sheela and Deepa, 2013), the Z-score normalization was applied to the

dataset following the **Eq. 8**; where x_i is the datum of the variable, μ is the mean of the variable, σ is its standard deviation and z_i is the normalized datum.

For the alternative approaches, the following methods were applied; the deterministic methods inverse distance weighted (IDW) and the inverse square distance weighted (ISDW) were performed using the **Eq. 20**; the statistic method multiple linear regression (MRL); and another machine learning method, the random forest regression (RFR) to contrast the ANNs performance. For the RFR, the Scikit-Learn (version 0.20) machine learning for Python was used, the number of trees or estimators were 15, the same train, test and validation datasets used for the ANNs were used for the RFR. For the deterministic and statistic methods, the train and test sets were used to generate the models, and the validation set remains for the same aim.

$$\hat{v} = \frac{\sum_{i=1}^n \left(\frac{v_i}{d^p} \right)}{\sum_{i=1}^n \left(\frac{1}{d^p} \right)} \quad \text{Eq. 20}$$

Eq. 20, were v_i is the known value; d^p is the distance to the point to be estimated to the power of $p = 1$ for IDW and $p = 2$ for the ISDW; and \hat{v} is the estimated value.

Additionally, all interpolations methods were tested in summer and winter months by separate to check the capability of the models during periods with more extreme phenomena, and the results of the interpolation of the entire year were compared to the results in these two seasons.

RESULTS & DISCUSSION

4. Results and Discussion

In this chapter, the obtained results from the experiments are presented and discussed with relevance literature to contrast these outcomes. The results will be offered divided into three sections, one for each of the main topic described in the methodology. Regarding the first topic, the microalgae suspensions light absorbance spectral signature and the ANN to elucidate the biology composition of monoalgal and mixed algal cultures are presented in section 4.1. The next topic of research, the color analysis and fertility characterization of soils results are detailed in section 4.2. Finally, concerning the third topic, the results and validation of the VWS for meteorological data acquisition are exposed in section 4.3.

4.1 Monoalgal and mixed algal cultures discrimination by using artificial neural network

In this section, the results of the ANN assay to elucidate microalgae cultures are presented. The structure presented is the following: the microalgae light absorption spectral signature result (section 4.1.1), the ANN used and its capability to examine microalgae suspensions (section 4.1.2), and final remarks about the research topic (section 4.1.3).

4.1.1 Microalgae light absorption

The light absorption spectrum of a microalga is a function of the biomass concentration in the culture and the pigment content of the biomass. The light absorption spectrum measurement is presented in **Figure 27** (left). The results showed that no substantial variations in the absorption spectrum occurred for any of the microalgae over the time course measurements, the slight variations observed being due to differences in the biomass content — the higher the cell density, the greater the absorbance (Helena *et al.*, 2011; Santos-Ballardo *et al.*, 2015). These variations were ruled out after the absorbance was converted to the relative absorbance for each measurement (**Eq. 18**).

When normalizing to relative absorption, the spectral signature is a function of the pigment content; this varies depending on the algal species. One can

observe how the lines overlap each other, thus making the spectral signature of the different microalgae used more comparable (**Figure 27** right). The results show that each microalga species had a unique spectral signature: *S. almeriensis* and *C. vulgaris* had similar spectral signatures while *Nostoc* sp. and *S. platensis* had particular shapes that allow distinguishing one from the other. In the case of *Nostoc* sp., slight variations in spectral signature were observed over time, especially on days 4 and 5. This variation, especially in the chlorophyll absorbance region, was not due to changes in environmental conditions because these were kept constant under laboratory conditions so that it might be related to changes in the cells' biochemical composition.

Nutrient supply and light conditions are two significant factors that modify the pigment content of a microalga species. Thus, the pigment content is related to the physiological status of the cells (Pancha *et al.*, 2015). Furthermore, the growth-cycle phase also modifies the pigment content (Lubián *et al.*, 2000). Other pigment content variations are related to changes in environmental and operational conditions (Pancha *et al.*, 2014; Solovchenko *et al.*, 2013). Despite the difference in days between sample measurements, other factors remained constant, and the absorption spectrum remained uniform. Under different culture conditions, the spectral signature could present variations within the same microalgae species across the measurements.

The variation in a microalga species' spectral signature over time was minor compared to the inter-species signature variation. Similar light absorption spectra, with peaks at 450 nm and 680 nm, have been obtained for other microalgae, such as *Chlamydomonas reinhardtii* (Isono *et al.*, 2015), *Thalassiosira pseudonana* (Hewes, 2016) and *C. vulgaris* (Myers *et al.*, 2013). Absorbance peaks are caused by photosynthetic pigments. Thus, while chlorophylls present two distinct absorption maxima, one between 400 and 500 nm and the other between 600 and 700 nm, the maximum absorption for carotenoids can increase above 500 nm due to spectral shifts caused by different contributing pigments (Holzinger *et al.*, 2016).

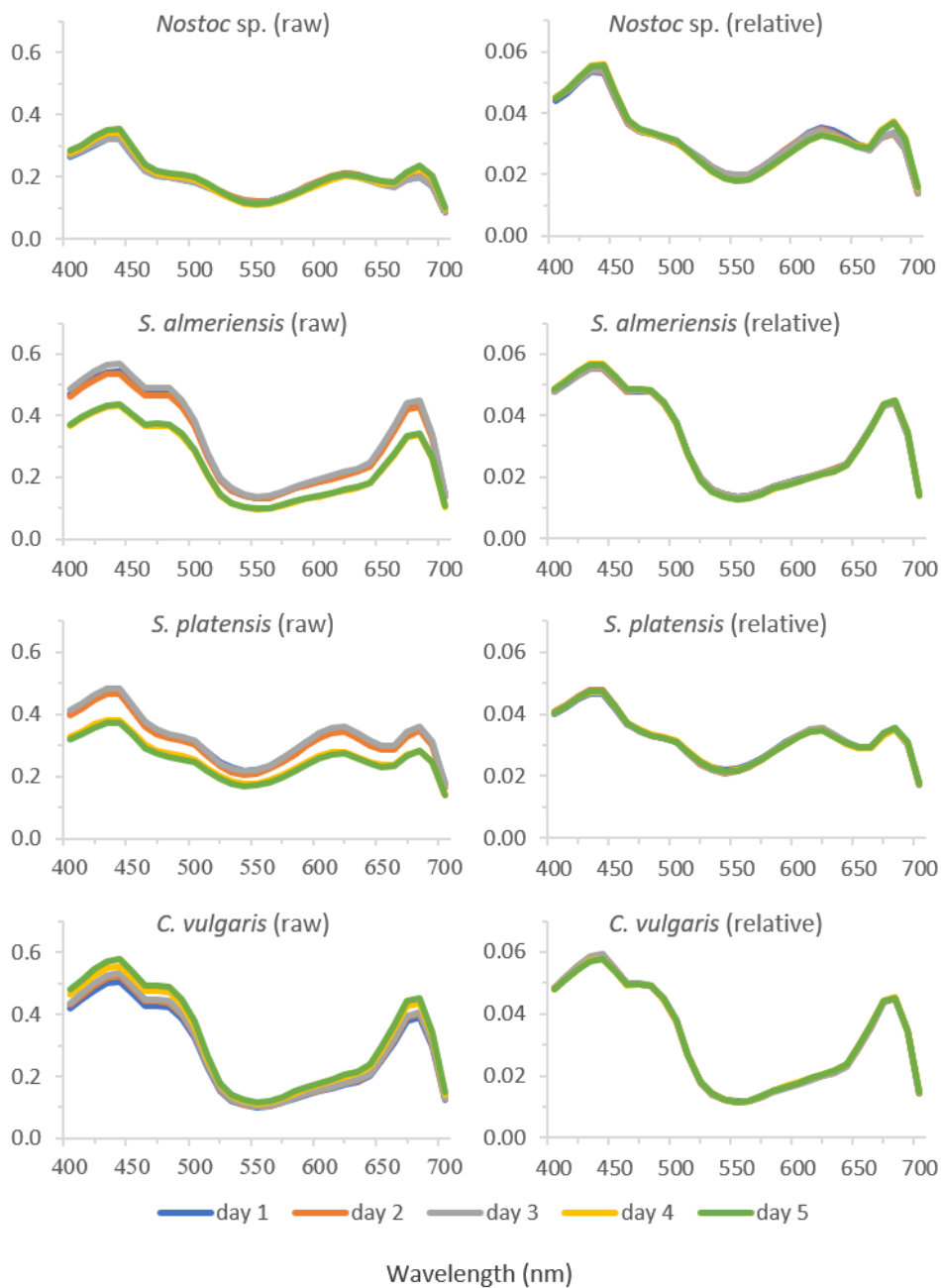


Figure 27. Light absorption spectrum of monoalgal cultures as obtained from the colorimeter (left), and relative absorbance conversion (right).

4.1.2 Artificial neural network culture discrimination

Data from 550 samples were measured to train and test the ANN. Once the training and testing phases were concluded, the obtained model was used to determine whether the validation samples were monoalgal or mixed algal cultures. The number of samples met the requirement of having more training samples than hidden neuron units (Huang *et al.*, 2004). The final ACE during the training and testing phases were 0.520 and 0.487, respectively. For the validation phase, the ACE was 0.527. The comparison between the experimental data and the ANN output is shown in **Table 1**.

Monoalgal cultures (S101 to S104) from the first validation experiment day were correctly identified by the ANN, providing results above 98% purity. Regarding the mixed algal samples, S105 and S106 correspond to not-previously-performed combinations; thus, they are “unknown combinations” for the model. Despite this, the model was capable of approximating the composition of these samples. Therefore, these samples were close to the monoalgal cultures and the model was able to identify the most abundant species and its relative proportion. The ANN considered that small amounts of other species were also present in these samples; nevertheless, it was always capable of identifying the most prevalent species and percentage composition.

Regarding the samples with 50% of two different species (S107 and S108), the artificial neural network was, likewise, capable of approximating the composition of these mixed algal samples, identifying the two predominant species and their approximate percentage. In no case, these samples were identified as monoalgal cultures. The same validation protocol was repeated on the second day. In this case, when using monoalgal cultures (S201 to S204), the ANN provided results above 99% purity for these samples.

For samples containing 90-95% of one species and 10-5% of another (S205 to S209), these samples were “unknown combinations” for the model. Nonetheless, it was able to identify the predominant algal species present in each sample, with percentages similar to monoalgal cultures. The one exception was the 90% *S. platensis* and 10% *C. vulgaris* sample, for which an accurate estimate was given. Similar mixed algal samples to those used for the training process (S210 to S227) provided better results, with none of them being classified as monoalgal. The last sample (S228) also corresponded to an “unknown combination” for the model, and it was, likewise, not classified as monoalgal.

Table 1. Comparison of the results from the experimental and the artificial neural network output of the validation experiment.

| Day | Sample | Experimental | | | | Model | | | |
|-----|--------|--------------|----------------|--------------|-------------|------------|----------------|--------------|-------------|
| | | Nostoc sp. | S. almeriensis | S. platensis | C. vulgaris | Nostoc sp. | S. almeriensis | S. platensis | C. vulgaris |
| 1 | S101 | 100.00% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% |
| | S102 | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% | 98.70% | 0.00% | 1.30% |
| | S103 | 0.00% | 0.00% | 100.00% | 0.00% | 0.02% | 0.00% | 99.98% | 0.00% |
| | S104 | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% | 0.14% | 0.00% | 99.86% |
| | S105 | 95.00% | 5.00% | 0.00% | 0.00% | 99.93% | 0.05% | 0.00% | 0.02% |
| | S106 | 90.00% | 10.00% | 0.00% | 0.00% | 98.71% | 1.19% | 0.00% | 0.09% |
| | S107 | 50.00% | 0.00% | 0.00% | 50.00% | 60.16% | 0.91% | 0.00% | 38.93% |
| | S108 | 0.00% | 50.00% | 50.00% | 0.00% | 0.01% | 30.01% | 69.36% | 0.62% |
| | S109 | 33.33% | 33.33% | 33.33% | 0.00% | 47.63% | 36.42% | 15.29% | 0.66% |
| | S110 | 12.50% | 50.00% | 37.50% | 0.00% | 5.76% | 42.75% | 16.39% | 35.10% |
| 2 | S201 | 100.00% | 0.00% | 0.00% | 0.00% | 99.83% | 0.00% | 0.17% | 0.00% |
| | S202 | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% | 99.99% | 0.00% | 0.01% |
| | S203 | 0.00% | 0.00% | 100.00% | 0.00% | 0.10% | 0.00% | 99.89% | 0.01% |
| | S204 | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% | 0.26% | 0.00% | 99.74% |
| | S205 | 90.00% | 10.00% | 0.00% | 0.00% | 99.68% | 0.25% | 0.01% | 0.06% |
| | S206 | 0.00% | 90.00% | 10.00% | 0.00% | 0.02% | 99.70% | 0.21% | 0.06% |
| | S207 | 0.00% | 0.00% | 90.00% | 10.00% | 0.01% | 0.51% | 89.19% | 10.29% |
| | S208 | 0.00% | 5.00% | 0.00% | 95.00% | 0.00% | 0.02% | 0.00% | 99.98% |
| | S209 | 10.00% | 0.00% | 0.00% | 90.00% | 0.01% | 0.12% | 0.00% | 99.87% |

Table 1. Comparison of the results from the experimental and the artificial neural network output of the validation experiment (continuation).

| Day | Sample | Experimental | | | | Model | | | |
|------|--------|--------------|----------------|--------------|-------------|------------|----------------|--------------|-------------|
| | | Nostoc sp. | S. almeriensis | S. platensis | C. vulgaris | Nostoc sp. | S. almeriensis | S. platensis | C. vulgaris |
| 2 | S210 | 75.00% | 25.00% | 0.00% | 0.00% | 87.42% | 12.42% | 0.00% | 0.16% |
| | S211 | 75.00% | 0.00% | 25.00% | 0.00% | 73.57% | 0.00% | 26.43% | 0.00% |
| | S212 | 75.00% | 0.00% | 0.00% | 25.00% | 68.98% | 0.87% | 0.00% | 30.15% |
| | S213 | 0.00% | 75.00% | 25.00% | 0.00% | 0.09% | 82.03% | 17.85% | 0.03% |
| | S214 | 0.00% | 75.00% | 0.00% | 25.00% | 0.00% | 79.11% | 0.00% | 20.88% |
| | S215 | 25.00% | 75.00% | 0.00% | 0.00% | 32.55% | 67.43% | 0.02% | 0.00% |
| | S216 | 0.00% | 0.00% | 75.00% | 25.00% | 0.00% | 0.28% | 60.20% | 39.52% |
| | S217 | 25.00% | 0.00% | 75.00% | 0.00% | 19.39% | 0.00% | 80.59% | 0.01% |
| | S218 | 0.00% | 25.00% | 75.00% | 0.00% | 0.01% | 19.11% | 80.44% | 0.44% |
| | S219 | 25.00% | 0.00% | 0.00% | 75.00% | 7.92% | 0.03% | 0.01% | 92.04% |
| | S220 | 0.00% | 25.00% | 0.00% | 75.00% | 0.00% | 17.98% | 0.00% | 82.02% |
| S221 | 0.00% | 0.00% | 25.00% | 75.00% | 0.06% | 0.02% | 9.61% | 90.31% | |
| S222 | 50.00% | 50.00% | 0.00% | 0.00% | 62.62% | 37.37% | 0.01% | 0.00% | |
| S223 | 50.00% | 0.00% | 50.00% | 0.00% | 49.55% | 0.00% | 50.45% | 0.00% | |
| S224 | 50.00% | 0.00% | 0.00% | 50.00% | 48.69% | 0.59% | 0.02% | 50.70% | |
| S225 | 0.00% | 50.00% | 50.00% | 0.00% | 0.04% | 38.49% | 61.39% | 0.09% | |
| S226 | 0.00% | 50.00% | 0.00% | 50.00% | 0.01% | 53.80% | 0.00% | 46.20% | |
| S227 | 0.00% | 0.00% | 50.00% | 50.00% | 0.00% | 0.01% | 56.01% | 43.98% | |
| S228 | 20.00% | 0.00% | 40.00% | 40.00% | 0.79% | 0.56% | 36.22% | 62.44% | |

According to these results, mixed algal samples with 10% or less contamination were weighted as monoalgal suspensions. Mixed algal samples composition were less accurately predicted in relation to monoalgal suspensions. In a study using flow cytometry with the SYTO9 stain, it was possible to classify *C. vulgaris*, *Scenedesmus obliquus*, *Chlamydomonas reinhardtii*, and *Navicula pelliculosa*, with errors ranging from 5-10%. However, the method misidentified microalgae cells in mixed algal samples (Peniuk *et al.*, 2016). Given that contamination by non-target microalgae is a severe problem to microalgae cultivation (Wen *et al.*, 2016), our method could be a powerful alternative for supervising algal cultures. Additionally, the method can provide information about the relative composition of a sample, an advantage over traditional methods in which more steps and time are needed to reach similar conclusions.

The ANN gives the approximate composition of a sample based on data input processing; each absorbance bandwidth is weighted during the training phase. Monoalgal samples composition were slightly under 100%, probably due to the manner each absorbance bandwidth influences the ANN output. For instance, in S102, a *S. almeriensis* monoalgal sample, *C. Vulgaris* obtained a 1.30% prediction compared to 98.70% for *S. almeriensis* — both microalgae species exhibited similar spectral signatures (**Figure 27** right). Furthermore, *C. vulgaris* was overestimated in mixed algal samples, whereas *S. platensis* was underestimated (S110, S216 and S221). This confusion may be due to the manner that the resulting spectral signature bandwidths, in the mixed algal samples, are weighted by the ANN.

To better show the accuracy of the developed ANN, a regression analysis of the experimental and modeled values was performed (**Figure 28**). The results show that, regardless of the microalga species, the ANN fitted the experimental values, with the R^2 ranging from 0.951 to 0.970.

Microalgae identification using ANN has been previously reported with taxonomic accuracy of up to 99%, using micrographic image analysis (Coltelli *et al.*, 2017), similar accuracy to that achieved in the present study for monoalgal cultures. The method reported here demonstrated its efficiency in discriminating mixed algal cultures, whereas it was less efficient when there were smaller percentages of another species in the samples. If the ANN output indicated 90% or less, there was a high probability that the examined sample was derived from a mixed algal culture; conversely, if it was more than 90%, the culture should be examined to confirm that it was monoalgal. Therefore, this decreases the number of chemical analyses required to monitor the biological composition of microalgae cultures.

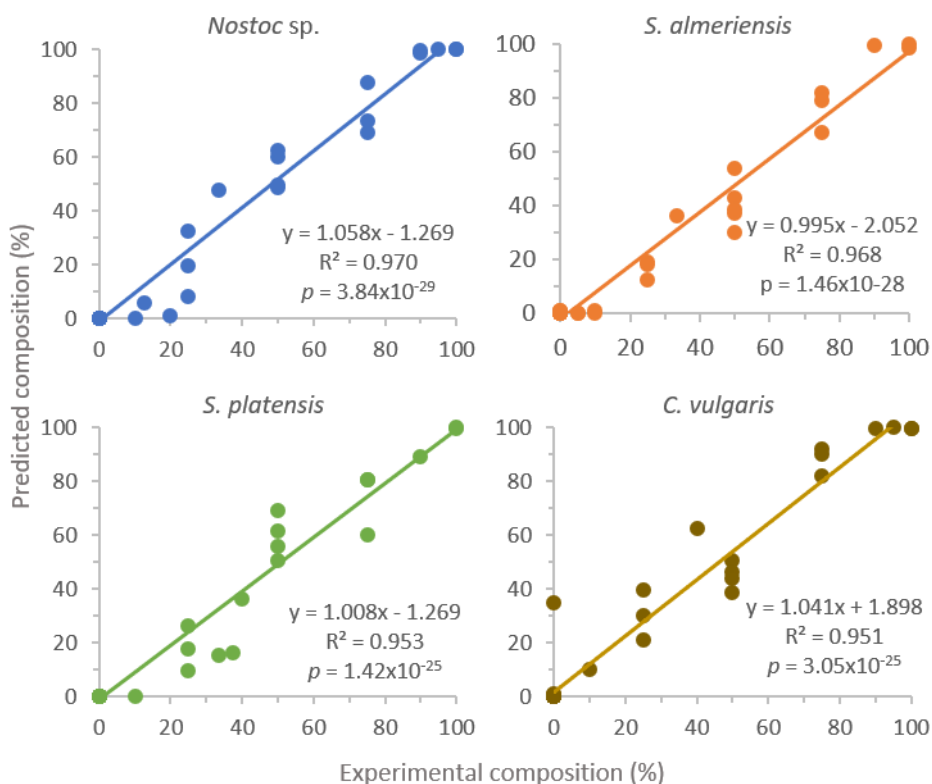


Figure 28. Correlation between the experimental composition of the samples and that predicted by the artificial neural network developed for the samples contained in Table 1.

Re-training the ANN with more samples of mixed algal cultures in a wider variety of relative composition or using a higher spectral resolution could improve the ANN precision. The re-training could also be applied to incorporate more microalgae species into the model and test the model capability to differentiate a more significant number of microalgae. Re-training is a relatively fast process and can take anywhere between 5 minutes to about 1 hour depending on the computer hardware, training the ANN in a GPU is significantly faster than in a CPU only system (Ramachandran *et al.*, 2015).

4.1.3 Final Remarks

A rapid methodology to elucidate microalgae species in suspensions has been developed and validated. To do this, microalgae spectral signatures from light absorption measurements of different microalgae species were analyzed through an ANN in order to describe and classify them. Four important species were used: *Nostoc* sp., *Scenedesmus almeriensis*, *Spirulina platensis* and *Chlorella vulgaris*. Absorbance from monoalgal and mixed algal cultures was the input data for training, testing and validating the ANN. The results show that the ANN was capable of distinguishing between monoalgal and mixed algal cultures, identifying the microalgae species in the monoalgal cultures and providing the approximate composition of mixed algal cultures. These results confirm that the application of spectral signatures with ANN is a suitable method for approximating the biological composition of microalgae cultures.

4.2 Artificial neural network for soil color analysis and characterization

In this section, the soil characterization and color analysis by ANNs results will be detailed. The remaining subsections are organized as follows: the computer software developed for the color conversion from RGB color system to L*a*b* color system (section 4.2.1), soil samples Munsell color classification and description results (section 4.2.2), the characterization of soil fertility parameters through ANNs (section 4.2.3), and final remarks about the research topic (section 4.2.4).

4.2.1 Computer vision software for color conversion

The DigiCIELAB is the CV software developed to perform the soil color experiments. Although the software was initially used for the soil color analysis, it is able to analyze the color of distinct objects such as fruits, plastic, color paints, and other objects. The DigiCIELAB was selected by the University of Valladolid in the “Prometeo” program, 2017 edition, and protected intellectually by the university in consequence of the award (**Appendix B**).

Colorimeters are the most commonly used instrument to perform color measurements in scientific researches; however, this instrument has limitations for assessing some materials, especially those with the irregular color surface (Barbin *et al.*, 2016; Girolami *et al.*, 2013; Yam and Papadakis, 2004). A more economical and versatile alternative was developed using CV and an ANN to assess distinct types of surfaces, including irregularly colored. This application allows accelerating the process of color measurement through digital photographs analysis and the RGB color system conversion to the L*a*b* system, to obtain the CIELAB international standard for color researches. The user interface is shown in **Figure 29**.

In the main screen of the software, two options are displayed; “Color Calibration” and “Colorimeter” as shown in **Figure 29a** and **Figure 29b** respectively. The “Color Calibration” option is used to generate the model capable of converting RGB to L*a*b* with a specific camera in a particular light condition. The “Colorimeter” option is used to perform the color measurement of new photographed samples.

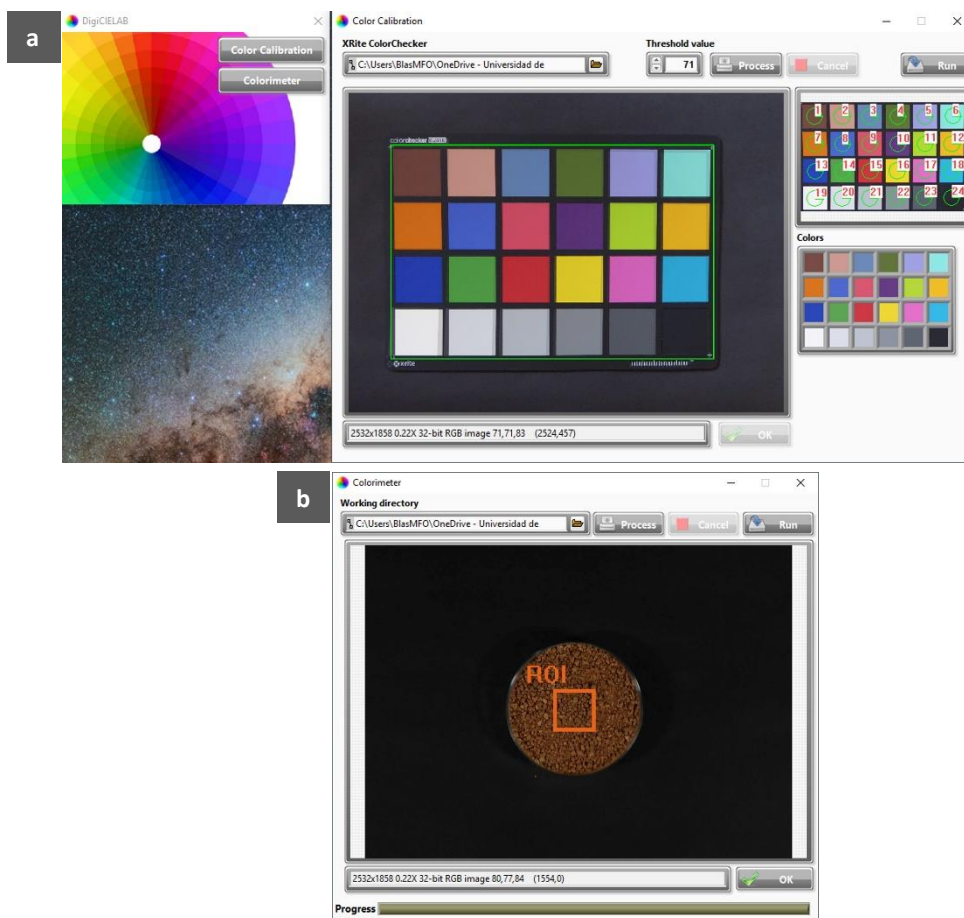


Figure 29. The user interface of DigiCIELAB CV software with the “Color Calibration” option (a) to generate the ANN model for the color analysis, and the “Colorimeter” option (b) for the color measurement of samples with the calibrated model.

The application automatically detects the colorchecker in a digital color photograph, detects each color sub-frame from the card and extracts the RGB color components. These RGB values and the $L^*a^*b^*$ color values provided by the manufacturer are employed to generate an ANN model capable of converting color systems, from RGB to $L^*a^*b^*$. To perform the color measurement of new objects, photographed samples must be taken in the same photographic set that the colorchecker used to calibrate the ANN, i.e., same illumination conditions and camera configuration. The workflow diagram of DigiCIELAB is shown below in **Figure 30**.

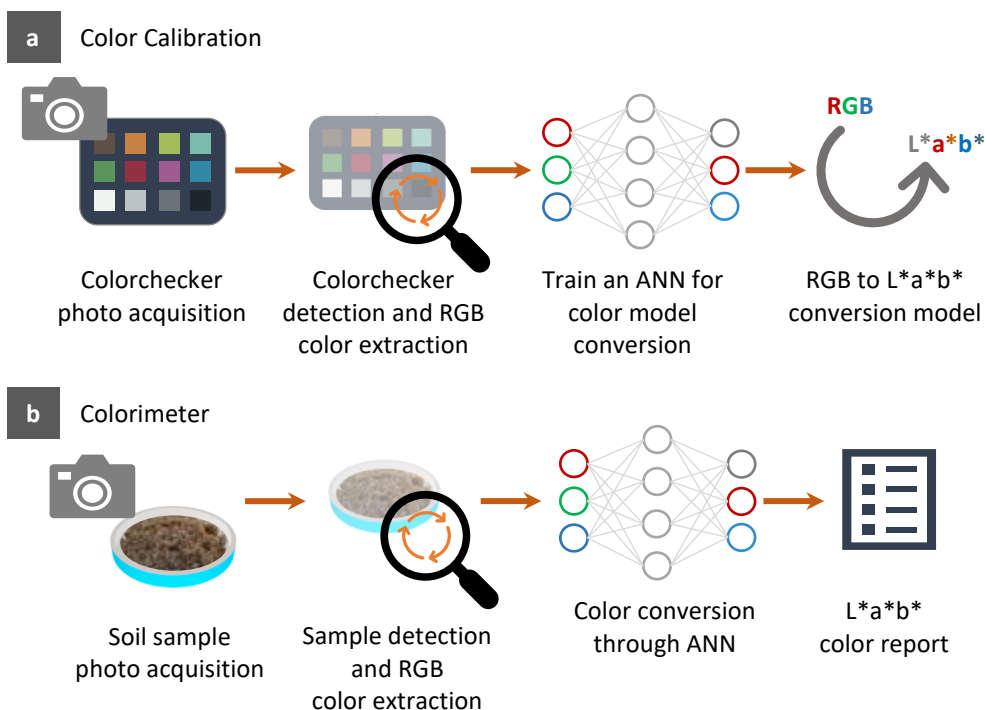


Figure 30. DigiCIELAB software workflow scheme. The part "a" of the scheme corresponds to the "Color Calibration" option, and the part "b" is regarding the "Colorimeter" option.

In part "a" of the scheme (**Figure 30a**), it is represented how the "Color Calibration" option of the software is executed. The software searches the colorchecker, separates all the 24 color sub-frames and extracts the RGB values for each one of them. Subsequently, the RGB data of the colorchecker photograph and the L*a*b* values provided by the manufacturer are used to generate an ANN model. Once the model is obtained, the color calibration phase is completed and can be used to measure the color in new photographed samples.

In part "b" of the scheme (**Figure 30b**), the "Colorimeter" option execution is schematized. In this window, the user selects the photographs to analyze and obtain the L*a*b* color. The user chooses an ROI in the photo — the area where the RGB color information will be extracted — or let the CV algorithms determine automatically the ROI based on the shape of the photographed object. The ANN trained in the previous phase ("Color Calibration" option) is used here to estimate

the color in the CIELAB system for the new samples. Finally, the software prepares a report with the sample name, RGB color, and L*a*b* color values for each case.

Each time the user wants to measure the color of images in different scenarios — different lighting conditions and camera configuration, it is necessary to take a picture of the colorchecker in each different scenario. Otherwise, errors in the L*a*b* color conversion will take place due to the differences caused by light variations (Tarlak *et al.*, 2016; ten Bosch and Coops, 1995; Valous *et al.*, 2009) and cameras color acquisition (Ilie and Welch, 2005; Kim *et al.*, 2012). The recommended colorchecker to use with the DigiCIELAB software is the “X-Rite ColorChecker Classic” (**Figure 23, Appendix C**).

In the DigiCIELAB, the process to autodetect the colorchecker and the sample to assess the color is similar. For the case, the process to detect a soil sample in a Petri plate will be detailed below as an example and the step results of this process are shown in **Figure 31**.

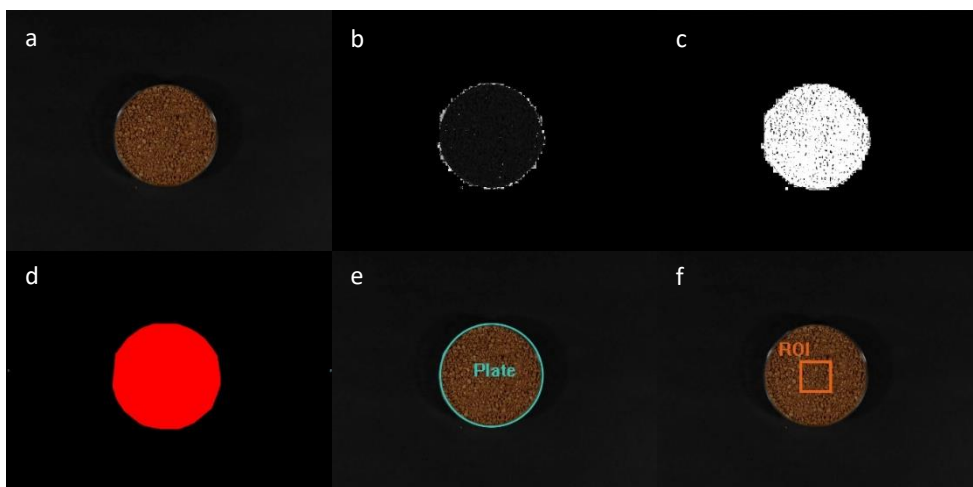


Figure 31. CV process to detect the object of study and ROI in the digital photograph: Original image (a), hue plane extraction (b), binary image (c), filled image (d), Petri plate detection (e), ROI for soil color extraction in the image center (f).

In the first step, from the original image (**Figure 31a**), the CV software performs a hue plane extraction (**Figure 31b**), i.e., separate the hue plane from the saturation and value planes of the HSV system. Then, a binary image is obtained (**Figure 31c**) by thresholding the image and leaving just the pixels of the object with values “1”; all other pixels, which are mainly the background in the image, are set to

“0”. The object is filled to eliminate inner “0” value pixels and obtain the skull of the object of interest (**Figure 31d**). Subsequently, the borders are detected, and the Petri plate is located in the image (**Figure 31e**). Finally, an ROI in the center of the detected plate is set, and the RGB values are extracted (**Figure 31f**).

ANNs present remarkable results for color conversion from RGB to L*a*b* color (Afshari-Jouybari and Farahnaky, 2011; León *et al.*, 2006; Pedreschi *et al.*, 2006; Pothula *et al.*, 2015). However, no previous work was conducted with the objective to make the color conversion through ANN more user-friendly. Instead, all the procedures from previous works were manually performed, from the color extraction from colorchecker, data pre-treatment, code and tuning the ANN, extract the color from objects to assess and run the model; this activity is highly time-consuming and not at the hands of all users since code ANN is not one of the most common skills between researches. With the DigiCIELAB, the user does not need to know about code or ANN architectures and algorithms to be able to use this tool, performing faster and reliable color assessments.

4.2.2 Soil samples description and color classification

The statistical summary of soil parameters is detailed in **Table 2**. The data range indicates that the parcel was not uniform in terms of soil characteristics, considering the relatively small area of study. The colors varied from light yellows to dark browns, going through different shades of red with an amplitude of lightness levels as the RGB and L*a*b* indicate.

Variables such as the OM went from 0.34 to 3.45%, with a mean of 1.58% and a standard deviation of 0.71. The contents of N, pH and Al were more uniform variables in the dataset. Meanwhile, EC, Ca, Mg and Na presented wider ranges and standard deviations.

The Munsell system classification resulted in 34 color notations from 6 hues (**Figure 32**). The yellow-red hues (YR) represented 74.32% of the soil samples, being the 10YR (30.00%) the most abundant among them. In previous studies, similar results for soil Munsell color were obtained in other regions of Spain, where YR was the prevailing hue; red soils are typical in Mediterranean areas (Castañeda and Moret-Fernández, 2013; Torrent *et al.*, 1980).

Table 2. Statistical summary of fertility parameters in soil dataset.

| Parameter | Range | Min. | Max. | Mean | St. dev. |
|------------|----------|----------|----------|----------|----------|
| R | 92.00 | 42.00 | 134.00 | 74.47 | 22.05 |
| G | 102.00 | 20.00 | 122.00 | 49.57 | 18.99 |
| B | 87.00 | 8.00 | 95.00 | 27.83 | 14.65 |
| L* | 39.99 | 36.79 | 76.77 | 46.45 | 8.42 |
| a* | 25.93 | 2.46 | 28.40 | 13.07 | 7.21 |
| b* | 17.30 | 11.80 | 29.10 | 18.85 | 3.81 |
| ECa | 127.53 | 1.45 | 128.98 | 25.13 | 20.28 |
| OM (%) | 3.10 | 0.34 | 3.45 | 1.58 | 0.71 |
| N (%) | 0.14 | 0.00 | 0.14 | 0.07 | 0.03 |
| P (mg/kg) | 24.31 | 0.04 | 24.35 | 6.07 | 4.35 |
| K (mg/kg) | 228.00 | 21.60 | 249.60 | 58.02 | 45.06 |
| Ca (mg/kg) | 2,577.40 | 1,735.60 | 4,313.00 | 3,045.97 | 698.20 |
| Mg (mg/kg) | 467.20 | 47.20 | 514.40 | 147.68 | 96.03 |
| Fe (mg/kg) | 4.31 | 0.25 | 4.57 | 1.80 | 0.99 |
| Cu (mg/kg) | 1.44 | 0.13 | 1.56 | 0.66 | 0.29 |
| Mn (mg/kg) | 6.37 | 0.25 | 6.63 | 2.58 | 1.39 |
| Na (mg/kg) | 219.80 | 294.20 | 514.00 | 341.96 | 41.36 |
| Al (mg/kg) | 0.35 | 0.06 | 0.41 | 0.17 | 0.07 |
| pH | 1.33 | 7.82 | 9.15 | 8.36 | 0.35 |
| EC (dS/m) | 2.81 | 0.23 | 3.04 | 1.41 | 1.04 |

The statistical analysis found significant differences in 5 of the 13 soil fertility parameters, N, Ca, Mg, pH and EC (**Table 3**). Inside these variables' analysis, not more than 3 different groups were detected among the 6 Munsell hues groupings. For OM, P, K, Fe, Cu, Mn, Na and Al, the Munsell hue parameter was not able to generate heterogeneous groups. In consequence, the hue parameter was not enough to produce a classification that leads to an estimation of the relative content of the studied fertility parameters; the analysis using the hue, value and chroma of the Munsell system was not feasible due to the limited availability of experimental units for each particular category.

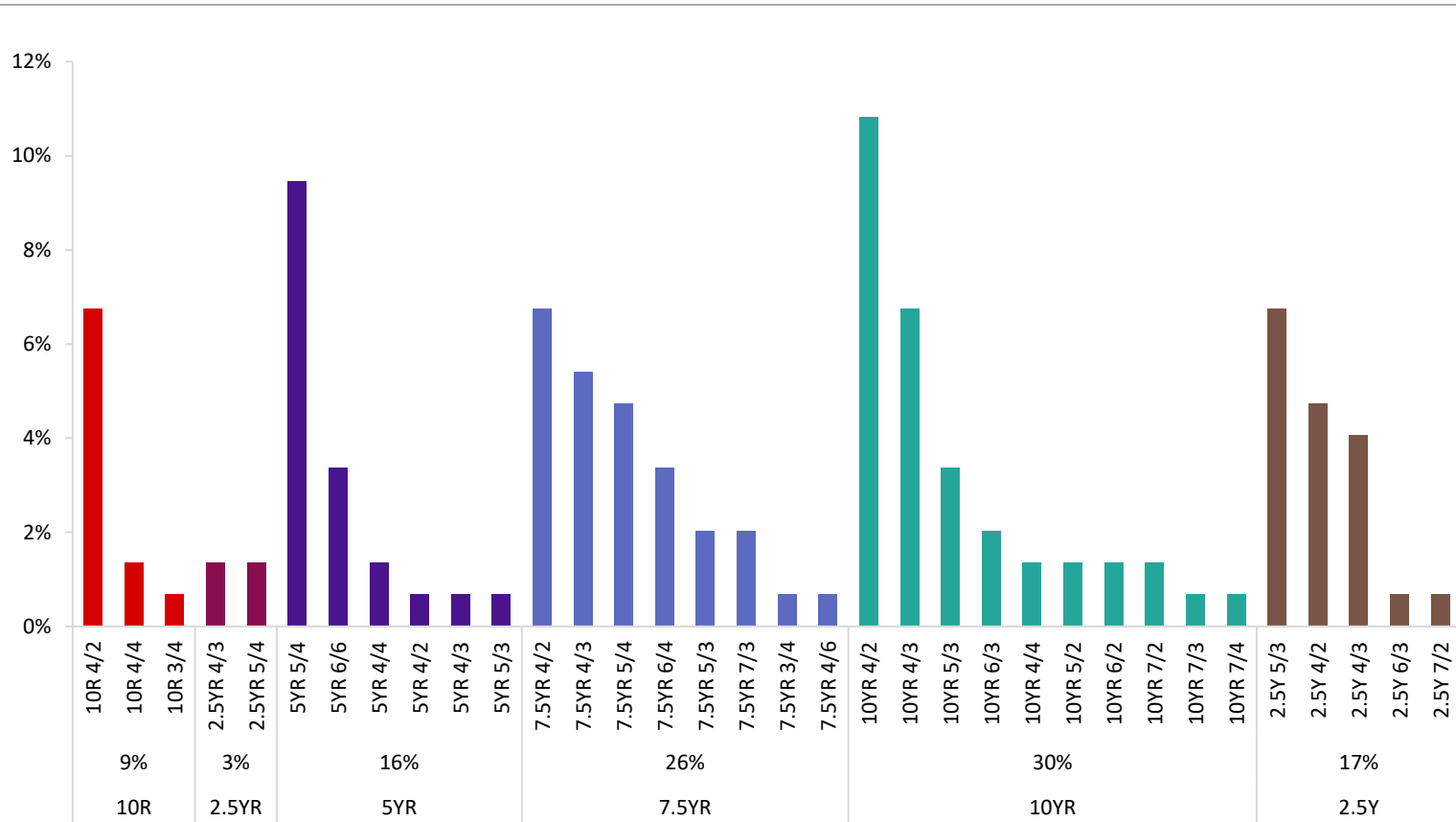


Figure 32. Munsell color classification for soil samples dataset.

Table 3. Means comparison of fertility parameters by Munsell color hue. Means with the same letter are not significantly different from each other (Kruskal-Wallis H test followed by the Mann-Whitney U test).

| Hue | OM (%) | N (%) | P (mg/kg) | K (mg/kg) | Ca (mg/kg) | Mg (mg/kg) | Fe (mg/kg) |
|-------|--------|---------|-----------|-----------|-------------|------------|------------|
| 10R | 1.39 a | 0.08 b | 4.82 a | 46.32 a | 3,270.28 ab | 178.15 ab | 2.09 a |
| 2.5YR | 1.98 a | 0.05 ab | 8.35 a | 44.00 a | 2,928.85 ab | 106.45 ab | 1.94 a |
| 5YR | 1.85 a | 0.04 a | 5.78 a | 36.03 a | 2,722.82 a | 93.59 a | 1.42 a |
| 7.5YR | 1.52 a | 0.07 b | 7.52 a | 65.48 a | 2,949.19 ab | 149.42 ab | 1.91 a |
| 10YR | 1.54 a | 0.07 b | 5.70 a | 65.75 a | 3,304.19 b | 170.59 b | 1.94 a |
| 2.5Y | 1.52 a | 0.08 b | 5.05 a | 62.54 a | 2,950.94 ab | 147.38 ab | 1.60 a |

| Hue | Cu (mg/kg) | Mn (mg/kg) | Na (mg/kg) | Al (mg/kg) | pH | EC (dS/m) |
|-------|------------|------------|------------|------------|---------|-----------|
| 10R | 0.71 a | 2.70 a | 319.75 a | 0.18 a | 8.25 a | 1.75 ab |
| 2.5YR | 0.72 a | 2.87 a | 327.85 a | 0.12 a | 8.49 ab | 0.87 ab |
| 5YR | 0.50 a | 2.18 a | 350.23 a | 0.18 a | 8.62 b | 0.83 a |
| 7.5YR | 0.73 a | 3.03 a | 348.57 a | 0.16 a | 8.42 ab | 1.24 ab |
| 10YR | 0.67 a | 2.44 a | 342.60 a | 0.16 a | 8.21 a | 1.86 b |
| 2.5Y | 0.66 a | 2.39 a | 336.62 a | 0.18 a | 8.34 a | 1.36 ab |

4.2.3 Characterization of soil fertility parameters through artificial neural networks

The characterization of soil fertility parameters using soil color and ANNs are shown in **Table 4**. For the RGB color system, there are significant p values for parameters such as N, P, Ca, Fe, Cu, Mn, pH, and EC. However, just N and EC had an R^2 higher than 0.5, with 0.62 and 0.76 respectively. The L*a*b* color system presented similar results; significant p values were obtained for the same variables as in the RGB system; showing R^2 higher than 0.5 the N, pH and EC with of 0.57, 0.54 and 0.71 respectively.

Table 4. Regression analysis between actual and predicted values of soil fertility parameters using RGB and L*a*b* for ANNs.

| Parameter | RGB | | L*a*b* | |
|-----------|-------|-------|--------|-------|
| | R^2 | p | R^2 | p |
| OM | 0.03 | 0.40 | 0.01 | 0.60 |
| N | 0.62 | 0.00* | 0.57 | 0.00* |
| P | 0.24 | 0.01* | 0.19 | 0.03* |
| K | 0.11 | 0.10 | 0.10 | 0.13 |
| Ca | 0.39 | 0.00* | 0.39 | 0.00* |
| Mg | 0.02 | 0.53 | 0.03 | 0.40 |
| Fe | 0.36 | 0.00* | 0.36 | 0.00* |
| Cu | 0.48 | 0.00* | 0.41 | 0.00* |
| Mn | 0.31 | 0.00* | 0.23 | 0.02* |
| Na | 0.14 | 0.07 | 0.04 | 0.31 |
| Al | 0.13 | 0.08 | 0.09 | 0.16 |
| pH | 0.48 | 0.00* | 0.54 | 0.00* |
| EC | 0.76 | 0.00* | 0.71 | 0.00* |

* significant for $p < 0.05$

The color assessment has no response to OM content. Other studies showed higher correlations between color and SOC, R^2 between 0.52 to 0.80 using multilinear regression models (Stiglitz *et al.*, 2017; Viscarra Rossel *et al.*, 2006). In contrast, a study of soil color relationship with N, SOC and clay with regression analysis techniques showed R^2 values of less than 0.50 for soil samples in a 60 km²

area with homogenous climate characteristics, but great diversity in soil parent material, aspect, topography, vegetation and land management (Ibáñez-Asensio *et al.*, 2013).

Correlations between soil organic components can be affected due to nature of the OM; soils with a high content of more humified substances show lower lightness (L^*) than soils in which fulvic acids predominates upon the same content of SOC (Vodyanitskii and Kirillova, 2016). In the same manner, there is a non-proportionality relationship between the dark soil and the OM, for instance, when the humus content exceeds 6%, the soil color varies only slightly in comparison with soil color variation under this OM threshold (Valeeva *et al.*, 2016). The organic substance present in Fe-coated particles also neutralizes its effect on the soil color (Vodyanitskii and Savichev, 2017).

In **Figure 33a**, it is observed how soils with similar L^* values had different OM content – twice as much in comparison; and a lighter soil with more OM than a darker soil as well. In such wise, the association of higher OM content in darker soils (Castañeda and Moret-Fernández, 2013) is not always in this line. Regarding the iron, another standard assumption is that red color is associated with this element in soils (Torrent *et al.*, 2006; Vodyanitskii and Kirillova, 2016). However, the similar phenomena to those that took place with OM were observed for iron (**Figure 33b**). The color caused by iron concentrations can also be masked by hydrologic conditions and the weathering that can modify the iron color by changing the oxidation state (Maejima *et al.*, 2000).

In a study of ANN and soil color to predict 44 parameters, physical and chemical, in a national scale study (Aitkenhead *et al.*, 2013), higher R^2 were found (in average 0.40 for predictions to the RGB color system and 0.41 to the $L^*a^*b^*$ color system) compared to those obtained here (in average 0.31 to the RGB and 0.28 to the $L^*a^*b^*$). In another study of ANN, using Landsat-8 satellite images, clay and OM for the prediction of soil cation exchange capacity, the obtained R^2 was 0.80. However, both studies were not clear regarding the dataset used for the validation phase; training, testing and validating, all three phases, must be performed to generate a proper ANN model (de Oliveira *et al.*, 2009), not using all phases can lead to an “apparently good” result product of overfitting or reduced generalization capacity (Piotrowski and Napiorkowski, 2013).

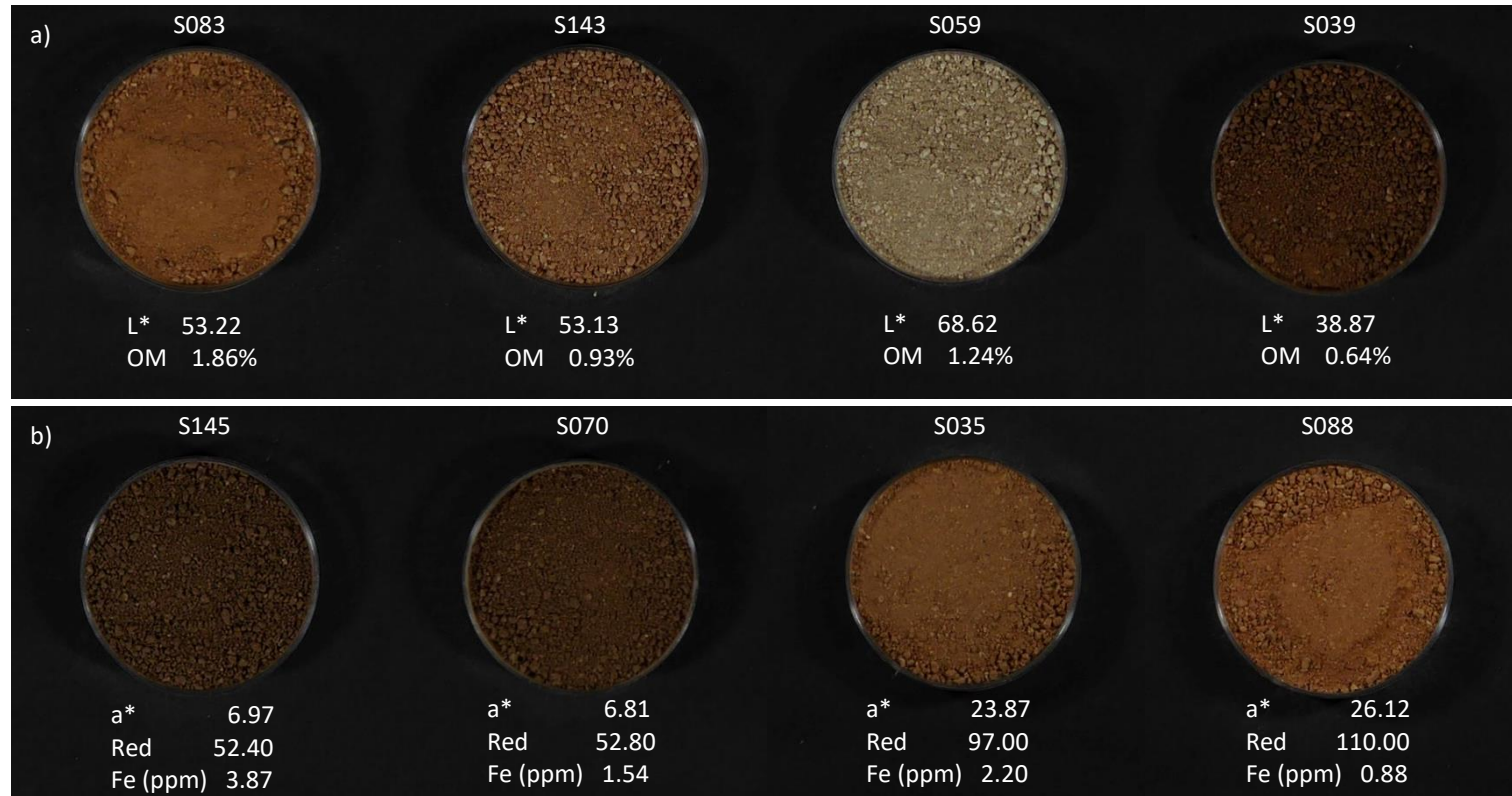


Figure 33. Soils samples with their respective lightness (L^*) and OM (organic matter) contents (a); and samples with their respective red component (a^* and R) and Fe (iron) contents (b).

Despite that ANN has been used in plenty of studies, validation data is frequently not used. In fact, test and validation data is often assumed to be the same such as it is reported (Bagheri Bodaghabadi *et al.*, 2016). The same authors working with ANN for interpolation and extrapolation of soil classes obtained results with errors of 30% for interpolations and 80% or higher for extrapolations in the validation phases.

4.2.4 Final remarks

The color is a widely studied attribute in soils and it is routinely used as a parameter for soil description. In this research, the soil color was analyzed to determinate its capability to group soils according to levels of fertility parameters such as OM, N, P, K, Ca, Mg, Fe, Cu, Mn, Na, Al, pH and EC and perform soil properties characterization from color data. For that, soils samples were classified according to the Munsell color notation and the obtained hues were used for statistical analysis; in addition, with the RGB and L*a*b* color, two ANNs were trained, tested and validated to describe the soils. Soil aggrupation based on the Munsell color hue resulted not efficient to separate samples according to fertility levels of the studied variables, and regarding the ANNs approach to describe soils, the obtained models were not capable of providing accurate results. The experiments suggest that soil color does not contain enough information to predict soil fertility parameters.

4.3 Virtual weather stations for meteorological data estimations

In this section, the VWS result and validation will be described. The interpolations algorithms were evaluated, and contrast to each other for assessing their performance in different seasons of the year. The following subsections are organized as: a statistical summary of the meteorological dataset, including a data breakdown by summer and winter seasons (section 4.3.1), the interpolation methods results comparison (section 4.3.2), seasons effects in the interpolated data quality (section 4.3.3), the VWS operation description (section 4.3.4), and final remarks about the research topic (section 4.3.5).

4.3.1 Statistical summary of the dataset

After EDA for removing lecture errors and outliers, the statistical summary of the whole dataset was obtained (**Table 5**). During the studied period (from July 2017 to June 2018), 18,234 sets of observations were recorded for the 53 meteorological stations. The maximum registered precipitation was 42.85 mm in a day; the mean was 1.24 mm per day during the period. The mean ETo was 2.82 mm, which mean that the overall water balance is negative. Temperatures were 11.01 °C on average, 42.97 °C and -19.75 °C for maximum and minimum registered. RH was 71.06 % on average, mean WS was 1.92 m/s² and mean TSI was 15.99 MJ/m².

Table 5. Statistical summary of meteorological observations dataset (n = 18,234).

| Parameter | Range | Min. | Max. | Mean | St. dev. |
|--------------------------|-------|--------|--------|-------|----------|
| Precip (mm) | 42.85 | 0.00 | 42.85 | 1.24 | 3.52 |
| ETo (mm) | 11.02 | 0.11 | 11.13 | 2.82 | 1.93 |
| Mean Temp (°C) | 36.96 | -8.95 | 28.01 | 11.01 | 7.19 |
| Max. Temp (°C) | 42.97 | -2.50 | 40.47 | 18.25 | 8.81 |
| Min. Temp (°C) | 40.50 | -19.75 | 20.75 | 4.31 | 6.04 |
| Mean RH (%) | 78.74 | 21.26 | 100.00 | 71.06 | 15.50 |
| Max. RH (%) | 63.05 | 36.95 | 100.00 | 92.68 | 8.55 |
| Min. RH (%) | 99.01 | 0.99 | 100.00 | 43.87 | 20.73 |
| Mean WS (m/s) | 17.72 | 0.01 | 17.73 | 1.92 | 1.22 |
| TSI (MJ/m ²) | 33.56 | 0.46 | 34.02 | 15.99 | 8.44 |

The mean temperature was similar to the previously described in the same region (del Río *et al.*, 2005), 11.17 °C for a 37 years dataset (from 1961 to 1997) in comparisons with the 11.01 °C obtained here in the 2017-2018 period. Regarding the precipitations, the same authors postulated an average of 664 mm for the same period, the average rainfall registered in the present study was 441.12 mm, less than the minimum of 480 mm registered in 1996 by the authors in the driest year. In a study from 1981 to 2010 in Castilla y León, the mean temperature was 11 °C likewise (Nafría *et al.*, 2013).

The statistical summary of the meteorological variables during the summer and winter months are presented below in **Table 6**.

Table 6. Statistical summary of meteorological observations during the summer months and the winter months.

| Parameter | Range | Min. | Max. | Mean | St. dev. |
|--------------------------|-------|--------|--------|-------|----------|
| Summer (n = 4,744) | | | | | |
| Precip (mm) | 42.85 | 0.00 | 42.85 | 0.50 | 2.69 |
| ETo (mm) | 10.19 | 0.94 | 11.13 | 4.89 | 1.47 |
| Mean Temp (°C) | 20.76 | 7.25 | 28.01 | 18.98 | 3.75 |
| Max. Temp (°C) | 26.25 | 14.22 | 40.47 | 27.94 | 4.56 |
| Min. Temp (°C) | 23.36 | -2.61 | 20.75 | 10.12 | 3.63 |
| Mean RH (%) | 73.13 | 22.27 | 95.40 | 56.63 | 12.50 |
| Max. RH (%) | 60.96 | 39.04 | 100.00 | 87.58 | 10.59 |
| Min. RH (%) | 86.03 | 1.17 | 87.20 | 26.45 | 11.61 |
| Mean WS (m/s) | 17.34 | 0.39 | 17.73 | 1.79 | 0.95 |
| TSI (MJ/m ²) | 32.01 | 1.98 | 33.99 | 23.16 | 5.51 |
| Winter (n = 4,538) | | | | | |
| Precip (mm) | 39.76 | 0.00 | 39.76 | 1.29 | 3.63 |
| ETo (mm) | 3.11 | 0.11 | 3.22 | 0.86 | 0.43 |
| Mean Temp (°C) | 23.36 | -8.95 | 14.41 | 3.24 | 3.32 |
| Max. Temp (°C) | 22.67 | -2.50 | 20.17 | 8.62 | 3.66 |
| Min. Temp (°C) | 32.63 | -19.75 | 12.88 | -1.46 | 4.17 |
| Mean RH (%) | 67.79 | 32.21 | 100.00 | 83.61 | 10.37 |
| Max. RH (%) | 37.44 | 62.56 | 100.00 | 95.99 | 4.36 |
| Min. RH (%) | 98.93 | 1.07 | 100.00 | 62.31 | 19.48 |
| Mean WS (m/s) | 11.18 | 0.01 | 11.19 | 2.09 | 1.37 |
| TSI (MJ/m ²) | 18.62 | 0.46 | 19.08 | 7.24 | 3.97 |

The average precipitation during the winter is more than twice that in summer, 1.29 mm and 0.50 mm respectively; in this location, the summer is the driest season and the rainiest season is autumn or winter, depending on the specific site. (Nafría *et al.*, 2013). Other variables such as temperatures, ETo, TSI are higher in summer; meanwhile, the RH is generally higher in winter in concordance with precipitations. The extreme temperatures, the minimum, and maximum registered in the entire year occurred during these two seasons. The standard deviation in these two seasons are lower in comparison with those seen in the whole year summary (**Table 5**), the range measurements of the phenomena were also narrow because of the uniformity of conditions during specific seasons in contrast to ranges of an entire year. The general behavior of the variables during the summer, winter, and the rest of the year is presented in **Figure 34**.

It is observable that for ETo, temperatures, RH and TSI, more extreme lectures were obtained during summer and winter. Meanwhile, the rest of the year presented more intermedium values. The letters above the boxplots show the differences between groups by the Mann-Whitney U test. The variations of the means for precipitation, mean WS and max. RH in the periods were not as notorious as the previously mentioned variables; nonetheless, the statistical test found significant differences in these variables as well.

4.3.2 Interpolation methods comparison

Different ANN models as well other interpolation methods were compared in terms results accuracy for meteorological variable estimations. Generally, good agreement between the estimated values and the actual records from the weather stations were observed for the interpolation methods. Results from ANNs approach are shown in **Table 7**, and results from alternative approaches are shown in **Table 8**.

Between the 5 activation functions in the ANNs, the softsign had the higher R^2 , 0.84, following by the sigmoid and tanh with an R^2 of 0.83, the relu with an R^2 of 0.82 and the hardlim with the lowest R^2 result, 0.79. The mean and maximum Temp were the most accurate variable to predict, with R^2 in the range from 0.96 to 0.98; following by the ETo, TSI and minimum Temp with R^2 higher than 0.91. Lower R^2 were obtained for mean WS, maximum RH and Prep., with ranges from 0.43 to 0.75. In general rule, the temperature is a more precise variable to interpolate in comparison to the precipitations (Jeffrey *et al.*, 2001).

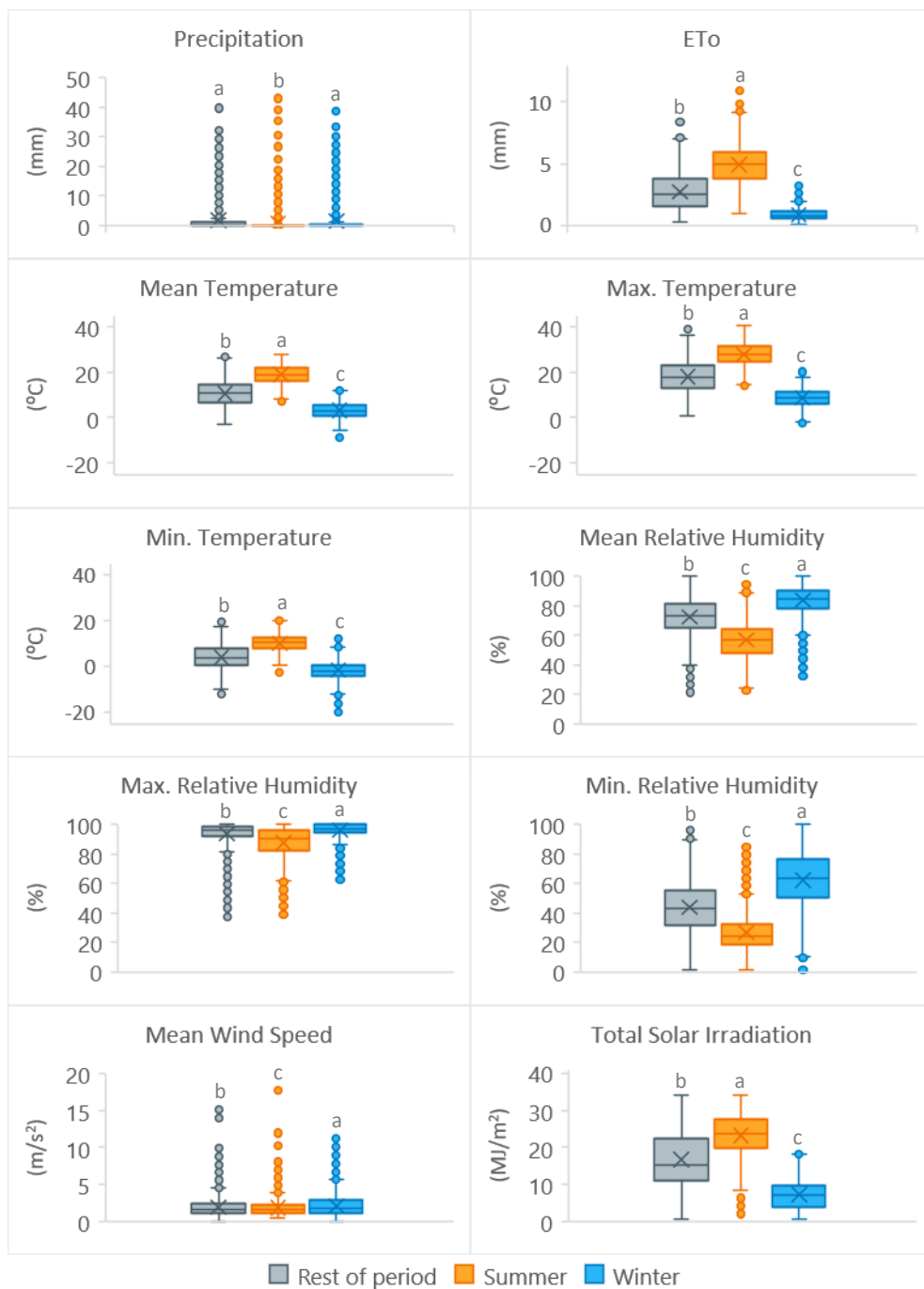


Figure 34. Boxplots and mean comparison of meteorological observations grouped by the season of the year.

Table 7. Analysis of meteorological data interpolations results for the ANNs approach.

| Parameter | Hardlim | | | Sigmoid | | | Tanh | | | Softsign | | | Relu | | |
|-----------|----------------|--------|------|----------------|--------|------|----------------|--------|------|----------------|--------|------|----------------|--------|------|
| | R ² | ρ | RMSE | R ² | ρ | RMSE | R ² | ρ | RMSE | R ² | ρ | RMSE | R ² | ρ | RMSE |
| Prep. | 0.62 | 0.00* | 2.22 | 0.73 | 0.00* | 1.88 | 0.71 | 0.00* | 2.04 | 0.75 | 0.00* | 1.82 | 0.73 | 0.00* | 1.87 |
| ETo | 0.92 | 0.00* | 0.55 | 0.94 | 0.00* | 0.48 | 0.94 | 0.00* | 0.46 | 0.94 | 0.00* | 0.48 | 0.93 | 0.00* | 0.51 |
| Mean Temp | 0.96 | 0.00* | 1.45 | 0.98 | 0.00* | 1.03 | 0.98 | 0.00* | 1.10 | 0.98 | 0.00* | 1.04 | 0.97 | 0.00* | 1.15 |
| Max. Temp | 0.96 | 0.00* | 1.86 | 0.98 | 0.00* | 1.23 | 0.98 | 0.00* | 1.28 | 0.98 | 0.00* | 1.28 | 0.97 | 0.00* | 1.49 |
| Min. Temp | 0.91 | 0.00* | 1.83 | 0.93 | 0.00* | 1.55 | 0.94 | 0.00* | 1.52 | 0.93 | 0.00* | 1.57 | 0.92 | 0.00* | 1.71 |
| Mean RH | 0.83 | 0.00* | 6.64 | 0.87 | 0.00* | 5.87 | 0.88 | 0.00* | 5.63 | 0.87 | 0.00* | 5.68 | 0.86 | 0.00* | 6.05 |
| Max. RH | 0.51 | 0.00* | 6.24 | 0.57 | 0.00* | 5.71 | 0.60 | 0.00* | 5.50 | 0.58 | 0.00* | 5.62 | 0.58 | 0.00* | 5.66 |
| Min. RH | 0.81 | 0.00* | 9.18 | 0.84 | 0.00* | 8.40 | 0.85 | 0.00* | 8.08 | 0.85 | 0.00* | 8.06 | 0.82 | 0.00* | 8.86 |
| Mean WS | 0.43 | 0.00* | 1.06 | 0.47 | 0.00* | 1.01 | 0.51 | 0.00* | 0.96 | 0.51 | 0.00* | 0.97 | 0.50 | 0.00* | 0.98 |
| TSI | 0.93 | 0.00* | 2.25 | 0.96 | 0.00* | 1.81 | 0.96 | 0.00* | 1.75 | 0.96 | 0.00* | 1.70 | 0.96 | 0.00* | 1.79 |
| Mean | 0.79 | | 3.33 | 0.83 | | 2.90 | 0.83 | | 2.83 | 0.84 | | 2.82 | 0.82 | | 3.01 |

* significant for $p < 0.05$

Table 8. Analysis of meteorological data interpolation results for alternative methods.

| Parameter | IDW | | | ISDW | | | MLR | | | RFR | | |
|-----------|----------------|----------|------|----------------|----------|------|----------------|----------|------|----------------|----------|------|
| | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE |
| Prep. | 0.63 | 0.00* | 2.16 | 0.71 | 0.00* | 1.89 | 0.65 | 0.00* | 2.11 | 0.72 | 0.00* | 1.95 |
| ETo | 0.94 | 0.00* | 0.48 | 0.95 | 0.00* | 0.44 | 0.93 | 0.00* | 0.50 | 0.94 | 0.00* | 0.47 |
| Mean Temp | 0.97 | 0.00* | 1.14 | 0.98 | 0.00* | 0.92 | 0.98 | 0.00* | 1.02 | 0.98 | 0.00* | 0.92 |
| Max. Temp | 0.97 | 0.00* | 1.47 | 0.98 | 0.00* | 1.16 | 0.98 | 0.00* | 1.27 | 0.98 | 0.00* | 1.12 |
| Min. Temp | 0.93 | 0.00* | 1.61 | 0.94 | 0.00* | 1.43 | 0.93 | 0.00* | 1.62 | 0.94 | 0.00* | 1.50 |
| Mean RH | 0.83 | 0.00* | 6.52 | 0.88 | 0.00* | 5.57 | 0.84 | 0.00* | 6.41 | 0.86 | 0.00* | 5.97 |
| Max. RH | 0.58 | 0.00* | 5.63 | 0.64 | 0.00* | 5.24 | 0.52 | 0.00* | 6.02 | 0.53 | 0.00* | 5.98 |
| Min. RH | 0.81 | 0.00* | 9.09 | 0.85 | 0.00* | 8.07 | 0.80 | 0.00* | 9.42 | 0.84 | 0.00* | 8.40 |
| Mean WS | 0.50 | 0.00* | 0.98 | 0.53 | 0.00* | 0.94 | 0.48 | 0.00* | 1.00 | 0.52 | 0.00* | 0.96 |
| TSI | 0.95 | 0.00* | 1.90 | 0.97 | 0.00* | 1.54 | 0.95 | 0.00* | 1.89 | 0.97 | 0.00* | 1.60 |
| Mean | 0.81 | | 3.10 | 0.84 | | 2.72 | 0.81 | | 3.13 | 0.83 | | 2.89 |

* significant for $p < 0.05$

The alternative methods presented similar performance, the best performing was the ISDW with an R^2 of 0.84, following by the RFR with an R^2 of 0.83 and IDW and MLR with an R^2 of 0.81. The mean and maximum Temp were the most accurate variable to predict, as in the ANN approach, with R^2 in the range from 0.97 to 0.98; following by the ETo, TSI and minimum Temp with R^2 higher than 0.93. Lower R^2 were obtained for mean WS, maximum RH and Prep., range from 0.48 to 0.72. Although the performances of the methods were similar, the ANN with softsign activation function and the ISDW were slightly better considering the R^2 . However, the deterministic method had the lowest RMSE, 2.72 in contrast to 2.82 of the machine learning method.

In a study of daily rainfall interpolation with different spatial methods such as several kriging variations, IDW and Thiessen polygons, the resulted RMSE were between 9.7 to 12.3; been 10.1 for the IDW method (Wagner *et al.*, 2012) — which are higher than the registered in this assessment. Hourly interpolations of rainfall seem to perform better, with R^2 up to 0.75 using IDW, nearest neighbor and kriging (Mendez and Calvo-Valverde, 2016). Another study with daily interpolations of precipitations and temperatures registered RMSE of 0.623 and 0.106 for rainfalls and temperatures respectively using IDW (Berndt and Haberlandt, 2018); these results are better than the obtained, backed by the use of a more significant number of weather stations up to 200 in 0.05 km² in contrast with 53 for 94.226 km², an overwhelming disproportion.

In general, climate and rainfall are highly nonlinear and complicated phenomena, which require advanced computer modeling for an accurate prediction (Nayak *et al.*, 2013). Others studies used meteorological variables as inputs in ANNs to estimate other meteorological variables. For instance, the ETo estimation with R of 0.89 and RMSE of 2.77 using temperatures (minimum and maximum) and WS as inputs in a period of 3 years (Falamarzi *et al.*, 2014), the RMSE obtained for the ETo in the current assay are lower for any interpolation method in comparison to that work. Using only the temperatures (minimum and maximum) to estimate the ETo, the obtained RMSE were in the range from 0.54 to 0.66 (Diamantopoulou *et al.*, 2011), compared to 0.44 to 0.55 obtained here. Local estimation of crop water requirements with ANN can be as accurate as of the evapotranspiration models calculations (Abrishami *et al.*, 2018).

Global solar irradiation forecasting with ANN using temperature, RH, and dates parameters (month, day and hour) achieved an R^2 of 0.99 and an RMSE of 0.17 (Hasni *et al.*, 2012), this prediction has more accuracy than the obtained here, using

more input data and historical data of site for the forecast. More precision in models, including interpolations, could be achieved with increasing amounts of observation data (Valipour, 2016), in this case, increasing the number of weather stations to have more input data for to generate the models.

The ANN approach, except for the hardlim ANN, outperforms the MLR interpolation with higher R^2 and lower RMSE. ANN outperforming MLR is a tendency and have been reported previously in meteorological studies with wind speed forecasting (Bilgili and Sahin, 2010), seasonal rain (Mekanik *et al.*, 2013) and ETo estimations (Laaboudi *et al.*, 2012).

4.3.3 Season effect in data interpolation quality

Results from summer season for the ANNs approach are shown in **Table 9**, and for the alternative approaches are shown in **Table 10**. Regarding the results from the winter season, for the ANNs approach are shown in **Table 11**, and for the alternative approaches are shown in **Table 12**.

In comparison with the interpolations of an entire year, during summer and winter seasons the accuracy for all interpolations methods was lower, R^2 presented values in an average of 0.76 and 0.71 respectively in contrast to 0.82 of the overall year. The RMSE, on the contrary, exhibited a slightly higher average value throughout the year, 2.97, in comparison to 2.91 and 2.84 for summer and winter respectively. The higher RMSE could be due that this error measurement penalizes the peaks, which presented larger absolute values more weight than errors with smaller absolute values (Shao *et al.*, 2014).

In a study of temperature interpolations during the months of January and July, selected as representatives of winter and summer months respectively; the precision in terms of RMSE resulted in values between 1.35 to 2.62 in January and 1.20 to 3.47 in July for MLR and kriging algorithms, been the lower RMSE to the MLR algorithm (Wu and Li, 2013). During the summer and the winter months, the intensities of meteorological phenomena had more peaks and when modeling these non-common events, are more difficult to abstract and create models able to predict these values in contrast to periods with less extreme values. Forecast the peaks of intensity still a tough task for models, including ANNs, since those peaks appear with low frequency (Nastos *et al.*, 2013).

Table 9. Analysis of meteorological data interpolations results for the ANNs approach for summer months.

| Parameter | Hardlim | | | Sigmoid | | | Tanh | | | Softsign | | | Relu | | |
|-----------|----------------|----------|------|----------------|----------|------|----------------|----------|------|----------------|----------|------|----------------|----------|------|
| | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE |
| Prep. | 0.91 | 0.00* | 1.19 | 0.96 | 0.00* | 0.82 | 0.95 | 0.00* | 1.53 | 0.92 | 0.00* | 1.47 | 0.90 | 0.00* | 1.55 |
| ETo | 0.72 | 0.00* | 0.83 | 0.81 | 0.00* | 0.68 | 0.81 | 0.00* | 0.67 | 0.80 | 0.00* | 0.70 | 0.79 | 0.00* | 0.71 |
| Mean Temp | 0.85 | 0.00* | 1.56 | 0.93 | 0.00* | 1.04 | 0.94 | 0.00* | 0.99 | 0.92 | 0.00* | 1.10 | 0.92 | 0.00* | 1.12 |
| Max. Temp | 0.81 | 0.00* | 2.16 | 0.93 | 0.00* | 1.28 | 0.91 | 0.00* | 1.44 | 0.92 | 0.00* | 1.31 | 0.90 | 0.00* | 1.47 |
| Min. Temp | 0.77 | 0.00* | 1.86 | 0.78 | 0.00* | 1.81 | 0.84 | 0.00* | 1.55 | 0.79 | 0.00* | 1.75 | 0.81 | 0.00* | 1.66 |
| Mean RH | 0.68 | 0.00* | 7.82 | 0.78 | 0.00* | 6.40 | 0.80 | 0.00* | 6.01 | 0.81 | 0.00* | 5.93 | 0.79 | 0.00* | 6.17 |
| Max. RH | 0.48 | 0.00* | 7.93 | 0.50 | 0.00* | 7.76 | 0.60 | 0.00* | 6.95 | 0.55 | 0.00* | 7.33 | 0.55 | 0.00* | 7.38 |
| Min. RH | 0.63 | 0.00* | 8.01 | 0.76 | 0.00* | 6.14 | 0.78 | 0.00* | 5.77 | 0.82 | 0.00* | 5.32 | 0.77 | 0.00* | 6.05 |
| Mean WS | 0.29 | 0.00* | 0.77 | 0.27 | 0.00* | 0.72 | 0.31 | 0.00* | 0.70 | 0.34 | 0.00* | 0.68 | 0.28 | 0.00* | 0.72 |
| TSI | 0.87 | 0.00* | 2.06 | 0.91 | 0.00* | 1.69 | 0.90 | 0.00* | 1.78 | 0.91 | 0.00* | 1.70 | 0.92 | 0.00* | 1.63 |
| Mean | 0.70 | | 3.42 | 0.76 | | 2.83 | 0.78 | | 2.74 | 0.78 | | 2.73 | 0.76 | | 2.84 |

* significant for $p < 0.05$

Table 10. Analysis of meteorological data interpolation results for the alternative methods for summer months.

| Parameter | IDW | | | ISDW | | | MLR | | | RFR | | |
|-----------|----------------|----------|------|----------------|----------|------|----------------|----------|------|----------------|----------|------|
| | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE |
| Prep. | 0.83 | 0.00* | 2.10 | 0.96 | 0.00* | 1.00 | 0.88 | 0.00* | 2.09 | 0.95 | 0.00* | 0.96 |
| ETo | 0.80 | 0.00* | 0.71 | 0.84 | 0.00* | 0.63 | 0.80 | 0.00* | 0.70 | 0.80 | 0.00* | 0.69 |
| Mean Temp | 0.91 | 0.00* | 1.21 | 0.94 | 0.00* | 0.98 | 0.93 | 0.00* | 1.06 | 0.95 | 0.00* | 0.92 |
| Max. Temp | 0.90 | 0.00* | 1.53 | 0.93 | 0.00* | 1.23 | 0.91 | 0.00* | 1.41 | 0.94 | 0.00* | 1.11 |
| Min. Temp | 0.83 | 0.00* | 1.61 | 0.85 | 0.00* | 1.51 | 0.83 | 0.00* | 1.61 | 0.82 | 0.00* | 1.61 |
| Mean RH | 0.72 | 0.00* | 7.36 | 0.81 | 0.00* | 6.00 | 0.73 | 0.00* | 7.03 | 0.78 | 0.00* | 6.30 |
| Max. RH | 0.57 | 0.00* | 7.47 | 0.65 | 0.00* | 6.64 | 0.47 | 0.00* | 8.03 | 0.46 | 0.00* | 8.20 |
| Min. RH | 0.69 | 0.00* | 6.82 | 0.75 | 0.00* | 6.10 | 0.68 | 0.00* | 6.99 | 0.79 | 0.00* | 5.69 |
| Mean WS | 0.38 | 0.00* | 0.64 | 0.49 | 0.00* | 0.59 | 0.28 | 0.00* | 0.72 | 0.34 | 0.00* | 0.68 |
| TSI | 0.89 | 0.00* | 1.89 | 0.94 | 0.00* | 1.39 | 0.91 | 0.00* | 1.74 | 0.94 | 0.00* | 1.41 |
| Mean | 0.75 | | 3.13 | 0.82 | | 2.61 | 0.74 | | 3.14 | 0.78 | | 2.76 |

* significant for $p < 0.05$

Table 11. Analysis of meteorological data interpolations results for the ANNs approach for winter months.

| Parameter | Hardlim | | | Sigmoid | | | Tanh | | | Softsign | | | Relu | | |
|-----------|----------------|----------|-------|----------------|----------|-------|----------------|----------|------|----------------|----------|-------|----------------|----------|-------|
| | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE |
| Prep. | 0.72 | 0.00* | 1.76 | 0.83 | 0.00* | 1.43 | 0.78 | 0.00* | 1.56 | 0.82 | 0.00* | 1.41 | 0.71 | 0.00* | 1.80 |
| ETo | 0.46 | 0.00* | 0.33 | 0.72 | 0.00* | 0.20 | 0.51 | 0.00* | 0.30 | 0.64 | 0.00* | 0.23 | 0.71 | 0.00* | 0.21 |
| Mean Temp | 0.89 | 0.00* | 1.22 | 0.92 | 0.00* | 0.93 | 0.90 | 0.00* | 1.03 | 0.92 | 0.00* | 0.98 | 0.94 | 0.00* | 0.85 |
| Max. Temp | 0.82 | 0.00* | 1.60 | 0.87 | 0.00* | 1.30 | 0.89 | 0.00* | 1.19 | 0.84 | 0.00* | 1.48 | 0.88 | 0.00* | 1.29 |
| Min. Temp | 0.84 | 0.00* | 1.78 | 0.87 | 0.00* | 1.49 | 0.87 | 0.00* | 1.53 | 0.89 | 0.00* | 1.38 | 0.90 | 0.00* | 1.35 |
| Mean RH | 0.56 | 0.00* | 6.49 | 0.65 | 0.00* | 5.73 | 0.64 | 0.00* | 5.85 | 0.67 | 0.00* | 5.56 | 0.66 | 0.00* | 5.62 |
| Max. RH | 0.05 | 0.03* | 4.21 | 0.08 | 0.01* | 3.84 | 0.12 | 0.00* | 3.73 | 0.14 | 0.00* | 3.64 | 0.15 | 0.00* | 3.63 |
| Min. RH | 0.63 | 0.00* | 11.82 | 0.73 | 0.00* | 10.19 | 0.73 | 0.00* | 9.99 | 0.69 | 0.00* | 10.74 | 0.70 | 0.00* | 10.62 |
| Mean WS | 0.76 | 0.00* | 0.68 | 0.70 | 0.00* | 0.72 | 0.79 | 0.00* | 0.60 | 0.76 | 0.00* | 0.65 | 0.72 | 0.00* | 0.71 |
| TSI | 0.79 | 0.00* | 1.85 | 0.83 | 0.00* | 1.64 | 0.87 | 0.00* | 1.43 | 0.87 | 0.00* | 1.47 | 0.80 | 0.00* | 1.81 |
| Mean | 0.65 | | 3.17 | 0.72 | | 2.75 | 0.71 | | 2.72 | 0.72 | | 2.75 | 0.72 | | 2.79 |

* significant for $p < 0.05$

Table 12. Analysis of meteorological data interpolation results for the alternative methods for winter months.

| Parameter | IDW | | | ISDW | | | MLR | | | RFR | | |
|-----------|----------------|----------|-------|----------------|----------|------|----------------|----------|-------|----------------|----------|------|
| | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE | R ² | <i>p</i> | RMSE |
| Prep. | 0.49 | 0.00* | 2.38 | 0.71 | 0.00* | 1.80 | 0.76 | 0.00* | 1.61 | 0.80 | 0.00* | 1.50 |
| ETo | 0.71 | 0.00* | 0.21 | 0.80 | 0.00* | 0.17 | 0.69 | 0.00* | 0.22 | 0.80 | 0.00* | 0.17 |
| Mean Temp | 0.90 | 0.00* | 1.06 | 0.94 | 0.00* | 0.83 | 0.94 | 0.00* | 0.84 | 0.94 | 0.00* | 0.86 |
| Max. Temp | 0.82 | 0.00* | 1.55 | 0.89 | 0.00* | 1.19 | 0.89 | 0.00* | 1.23 | 0.89 | 0.00* | 1.21 |
| Min. Temp | 0.83 | 0.00* | 1.75 | 0.88 | 0.00* | 1.46 | 0.86 | 0.00* | 1.60 | 0.88 | 0.00* | 1.50 |
| Mean RH | 0.59 | 0.00* | 6.27 | 0.68 | 0.00* | 5.48 | 0.62 | 0.00* | 6.07 | 0.66 | 0.00* | 5.66 |
| Max. RH | 0.13 | 0.00* | 3.60 | 0.16 | 0.00* | 3.52 | 0.08 | 0.01* | 4.01 | 0.08 | 0.01* | 3.94 |
| Min. RH | 0.65 | 0.00* | 11.46 | 0.74 | 0.00* | 9.93 | 0.66 | 0.00* | 11.36 | 0.74 | 0.00* | 9.92 |
| Mean WS | 0.69 | 0.00* | 0.72 | 0.75 | 0.00* | 0.64 | 0.74 | 0.00* | 0.66 | 0.80 | 0.00* | 0.60 |
| TSI | 0.76 | 0.00* | 1.98 | 0.86 | 0.00* | 1.55 | 0.84 | 0.00* | 1.63 | 0.89 | 0.00* | 1.32 |
| Mean | 0.66 | | 3.10 | 0.74 | | 2.66 | 0.71 | | 2.92 | 0.75 | | 2.67 |

* significant for $p < 0.05$

Other intrinsic factors that can alter the quality of the interpolated data are episodes of high-intensity rains, densely cloudy days and frost in winters (Thorsen and Höglind, 2010), temperature inversions (Bailey *et al.*, 2011), and heat waves in summers (Luber and McGeehin, 2008; Meehl and Tebaldi, 2004). In a study using ANN to forecast the TSI, the mean error was different according to the month and season of the year in which the predictions were made, with higher errors in autumn and winter and lower in spring and summer (Kemmoku *et al.*, 1999). The estimation of TSI by ANN using Meteosat-9 images as input was better in clear-periods than rainy or overcast ones; the RMSE was 21.20% against 5.13% for rainy and clear-sky days months respectively (Linares-Rodriguez *et al.*, 2013).

4.3.4 Virtual weather station

The algorithms composing the VWS are capable of access the InfoRiego FTP server, other servers address can be configurable into the script in order to access different FTP servers or similar protocols. Once the access is done, the user performs a filtered selection of the files to download information of a specified period. Thereupon, once the data is available in the user's computer, the interpolations algorithms can be carried out executing the preferred one by the user, introducing as input the XY UTM coordinates in the ETRS89 geodetic system of the weather stations and paring this data with the station lectures.

According to the results, the most appropriate methods to perform the data estimation thought interpolations in a given location are the ISDW and the ANN with the softsign function. The interpolations can be made to any given coordinate inside Castilla y León or other areas with a weather station network and accessible data to generate models. The innovative aspect of the VWS lies in the possibility of the user to choose a specific location, and estimated temperatures, RH, ETo, precipitations, TSI, WS with just one method, other studies of interpolations are focused just in a couple of variables using non-automated data access and processing.

4.3.5 Final Remarks

Meteorological data is an important series of observations for agricultural activities. The data is generally obtained from automatic weather stations; however,

the data can also be acquired from VWS. A VWS is an integration of algorithms to estimated meteorological data from nearby weather stations observations to other locations with no available stations. To develop the VWS, the performance of different interpolation methods were evaluated to test their accuracy. Daily data from an automatic weather station network were used to perform the interpolations. ANNs with the hardlim, sigmoid, tanh, softsign and relu activations functions were employed, as well as IDW, ISDW, MLR, and RFR to interpolate the daily observations.

Additionally, interpolations in the summer and winter months were performed to check the capability of the models during periods with more extreme phenomena registers. The results showed that the interpolation methods have an R^2 up to 0.98 for variables such as temperatures for the period of one year. Meanwhile, during the summer and winter, the models presented lower accuracy. From a practical perspective, the methods here described can be an alternative to meteorological data acquisition.

CONCLUSIONS

5. Conclusions

In this thesis, the application of ANN to solve distinct problems associated with agricultural activities were tested, in the microalgae production, soil fertility analysis, and meteorological variables acquisition. In the present section, the general conclusion (section 5.1) and the specific conclusions for each of the three main topics of research (sections 5.2, 5.3, and 5.4) will be detailed. Finally, future research niches based on the obtained result will be proposed (section 6).

5.1 General conclusions

While this thesis explores three different topics — microalgae, soil fertility, and meteorological data — the ANN approach to solving problems associated with these subjects in the agricultural plane is the common ground in which the present work is substantiated. From the experiments performed, the following general conclusions are proposed:

- Firstly, ANNs have proved to be a powerful tool to solve classifications, estimations and predictions problems. The use of ANNs in agricultural related issues is a critical step to find solutions for problems and help the users of this technology to make faster and better decisions in the productive chain, for instance, in the monitoring of microalgae cultures and crop management. By contrast, ANNs are not always capable of mapping the input variables with a target. In these cases, the network configuration must be rethought, dataset size increased and considered the fact that the input variables may not have a relation with the desired output.

- Secondly, ANNs can perform and adapt to multiple problems. There is not strictly ANN architecture, activation function, error measure technique, and other parameters to ensure the best performance. Several models for each case were evaluated, and the selected model was the most accurate for the validation set; the dataset in which the model evaluation must be performed to conclude the inquiry.

5.2 Monoalgal and mixed algal cultures discrimination by using artificial neural network conclusions

- It was demonstrated that microalgae light absorption spectra vary mainly as a function of the microalga species, although minor variations due to environmental and operational conditions can also take place.

- When maintaining the cultures under similar conditions, the light absorption spectra can be used to develop an ANN that differentiates monoalgal from mixed algal cultures and identifies the predominant species. In addition, it is useful to be able to approximate the percentage of each species in mixed cultures.

- A major advantage of this method is that it does not require much time or chemical analysis; a single light absorption spectrum is enough to quantify the biological composition of the cultures. It can provide a fast and powerful tool for microalgae culture management at the commercial scales.

5.3 Artificial neural network for soil color analysis and characterization conclusions

- Although the soil color classification is a frequently used technique in the current days and is it helpful for preliminary classification of soils; in the present assessment, the color did not contain enough information to classify soils according to levels of fertility parameters.

- The Munsell color hues analysis was not able to separate heterogeneous soil groups for all the fertility variables studied. For N, Ca, Mg, pH and EC statistical difference were detected, but no differences between all hues for a given variable. For the other eight fertility parameters (OM, P, K, Fe, Cu, Mn, Na and Al), no significative difference was found between Munsell hues.

- The ANN approach to predict the fertility parameters in soils based on the color did provide an inaccurate model. The neural network was not able to generate a consistent model from the soil samples due to the nature of the dataset; soils samples did not exhibit a consistent pattern regarding the color and fertility attributes, soils with similar color values drastically differed in the fertility parameter and soils with different colors exhibited identical fertility levels.

5.4 Virtual weather stations for meteorological data estimations conclusions

- In the section, the concept of VWS was introduced. The current state of the weather station networks and the online availability of the records makes possible to get the data and process them to perform estimations of the meteorological observations from the measurements of weather stations to other locations with no stations.

- The overall performances of the interpolation methods were accurate to estimate the meteorological variables; the ISDW and the ANN with the softsign activation functions were the most precise approach to perform the interpolations and to use in the VWS as preferred methods.

- The occurrence of extreme meteorological events, such as peaks of high and low temperatures or intense rainfalls, had negative impacts on the accuracy of the interpolation models. Therefore, these events should be considered in the model generation process.

- The success of the models suggests that they could be used in situations where no weather stations are present, but meteorological data series is needed for various purposes, for instance, the register of the ETo and calculus of crop water requirements.

FUTURE WORK

6. Future work

To conclude this thesis, there are several aspects to consider as future work that will be highlighted in this section. As is usually appreciated in research, the experience raises more questions than answer. Everything since the improvement of the ANN algorithms to the extension or modification of the research approaches or apply these techniques to other topics. In the paragraphs below, possible research niches for each of the thesis topics will be presented.

In the microalgae culture subject, the elucidation through ANN can be tested with more species to expand the network capability to identify microalgae. In the same manner, more similar species in terms of spectral footprints, such as species of the same genera or different strains of the same species, can be analyzed to test the scope of the ML technique.

Regarding the soil color and fertility analysis, different predictor variables should be studied since the color did not contain enough information to create an accurate model. Spectrometry and vis-NIR can be alternative approaches to try to predict fertility parameters with ANNs seeking better accuracies.

For the VWS, more research is needed to amplify the possible geographical application of this technology. To determine the appropriate minimum number of nearby weather stations to use as a source of data, distance to the stations and the effects on the accuracy of the interpolation and the inclusion of more meteorological variables or interpolation methods are some exciting possibilities. In terms of the accuracy of interpolations during different seasons of the year, more statistics techniques are necessary to reduce the negative impact of these “outliers” phenomena registers in the modeling process. Finally, the user interface for VWS building is also a key factor to bring this technology to more users, who do not necessarily want to interact with source code but can be interested in assays with this technology.

ANN applications are continually innovating in distinct disciplines of science and agriculture should not be alien to all this revolution. The possibilities are countless; phenology of crops, varieties and cultivars, yield estimation, pest models and pesticides application advisors, and machinery supervision to cite some examples. The improvements in this sector are crucial to maintaining food production, convert agriculture to more sustainable, improve competitiveness, and to minimize the effects of global warming.

REFERENCES

7. References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. Tensorflow: A system for large-scale machine learning, in: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16). pp. 265–283.
- Abalos, D., Sanchez-Martin, L., Garcia-Torres, L., van Groenigen, J.W., Vallejo, A., 2014. Management of irrigation frequency and nitrogen fertilization to mitigate GHG and NO emissions from drip-fertigated crops. *Sci. Total Environ.* 490, 880–888. <https://doi.org/10.1016/j.scitotenv.2014.05.065>
- Abd-Alla, M.H., Mahmoud, A.L.E., Issa, A.A., 1994. Cyanobacterial biofertilizer improved growth of wheat. *Phyton (B. Aires)*. 34, 11–18.
- Abdel-Raouf N, 2012. Agricultural importance of algae. *African J. Biotechnol.* 11, 11648–11658. <https://doi.org/10.5897/AJB11.3983>
- Abhishek, K., Khairwa, A., Pratap, T., Prakash, S., 2012. A stock market prediction model using Artificial Neural Network, in: 2012 3rd International Conference on Computing, Communication and Networking Technologies, ICCCNT 2012. pp. 1–5. <https://doi.org/10.1109/ICCCNT.2012.6396089>
- Abrishami, N., Sepaskhah, A.R., Shahrokhnia, M.H., 2018. Estimating wheat and maize daily evapotranspiration using artificial neural network. *Theor. Appl. Climatol.* 1–14. <https://doi.org/10.1007/s00704-018-2418-4>
- Ación, F.G., Fernández, J.M., Molina-Grima, E., 2014. Economics of microalgae biomass production, in: *Biofuels from Algae*. Elsevier, pp. 313–325.
- Ación, F.G., Gómez-Serrano, C., Morales-Amaral, M.M., Fernández-Sevilla, J.M., Molina-Grima, E., 2016. Wastewater treatment using microalgae: how realistic a contribution might it be to significant urban wastewater treatment? *Appl. Microbiol. Biotechnol.* 100, 9013–9022. <https://doi.org/10.1007/s00253-016-7835-7>
- Ación, F.G., González-López, C. V, Fernández, J.M., Molina, E., 2012. Conversion of CO₂ into biomass by microalgae: How realistic a contribution may it be to significant CO₂ removal? *Appl. Microbiol. Biotechnol.* 96, 577–586. <https://doi.org/10.1007/s00253-012-4362-z>
- Acosta-Alba, I., Chia, E., Andrieu, N., 2019. The LCA4CSA framework: Using life cycle assessment to strengthen environmental sustainability analysis of climate smart agriculture options at farm and crop system levels. *Agric. Syst.* 171, 155–170. <https://doi.org/10.1016/j.agsy.2019.02.001>
- Adya, M., Collopy, F., 1998. How effective are neural networks at forecasting and prediction? A review and evaluation. *J. Forecast.* 17, 481–495.

- Afshari-Jouybari, H., Farahnaky, A., 2011. Evaluation of photoshop software potential for food colorimetry. *J. Food Eng.* 106, 170–175. <https://doi.org/10.1016/j.jfoodeng.2011.02.034>
- Aggarwal, P., Vyas, S., Thornton, P.K., Campbell, B., 2019. How much does climate change add to the challenge of feeding the planet this century? *Environ. Res. Lett.* 14, 1–10. <https://doi.org/10.1088/1748-9326/aafa3e>
- Ahmad, F., Goparaju, L., Qayum, A., 2017. FAO guidelines and geospatial application for agroforestry suitability mapping: case study of Ranchi, Jharkhand state of India. *Agrofor. Syst.* 93, 1–14. <https://doi.org/10.1007/s10457-017-0145-y>
- Ahmadi, M.A., 2011. Prediction of asphaltene precipitation using artificial neural network optimized by imperialist competitive algorithm. *J. Pet. Explor. Prod. Technol.* 1, 99–106. <https://doi.org/10.1007/s13202-011-0013-7>
- Ahrens, C.D., 2012. *Meteorology Today: An Introduction to Weather, Climate, and the Environment*. Cengage Learning.
- Aitkenhead, M.J., Coull, M., Towers, W., Hudson, G., Black, H.I.J., 2013. Prediction of soil characteristics and colour using data from the National Soils Inventory of Scotland. *Geoderma* 200–201, 99–107. <https://doi.org/10.1016/j.geoderma.2013.02.013>
- Akram, M., El, C., 2016. Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks. *Int. J. Comput. Appl.* 143, 7–11. <https://doi.org/10.5120/ijca2016910497>
- Alanis, A.Y., 2018. Electricity Prices Forecasting using Artificial Neural Networks. *IEEE Lat. Am. Trans.* 16, 105–111. <https://doi.org/10.1109/TLA.2018.8291461>
- Ali, Z., Hussain, I., Faisal, M., Nazir, H.M., Hussain, T., Shad, M.Y., Mohamd Shoukry, A., Hussain Gani, S., 2017. Forecasting Drought Using Multilayer Perceptron Artificial Neural Network Model. *Adv. Meteorol.* 2017, 1–9. <https://doi.org/10.1155/2017/5681308>
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., others, 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome* 300, D05109.
- Allili, M., Coriveau, D., 2007. Topological analysis of shapes using Morse theory. *Comput. Vis. Image Underst.* 105, 188–199. <https://doi.org/10.1016/j.cviu.2006.10.004>
- Altinay, O., Tulunay, E., Tulunay, Y., 1997. Forecasting of ionospheric critical frequency using neural networks. *Geophys. Res. Lett.* 24, 1467–1470. <https://doi.org/10.1029/97GL01381>
- Amari, S. ichi, 1993. Backpropagation and stochastic gradient descent method. *Neurocomputing* 5, 185–196. [https://doi.org/10.1016/0925-2312\(93\)90006-O](https://doi.org/10.1016/0925-2312(93)90006-O)
- Amrouche, B., Le Pivert, X., 2014. Artificial neural network based daily local forecasting for global solar radiation. *Appl. Energy* 130, 333–341. <https://doi.org/10.1016/j.apenergy.2014.05.055>

- Andrade, S.I.E., Lima, M.B., Barreto, I.S., Lyra, W.S., Almeida, L.F., Araújo, M.C.U., Silva, E.C., 2013. A digital image-based flow-batch analyzer for determining Al(III) and Cr(VI) in water. *Microchem. J.* 109, 106–111. <https://doi.org/10.1016/j.microc.2012.03.029>
- Antipov, G., Berrani, S.A., Dugelay, J.L., 2016. Minimalistic CNN-based ensemble model for gender prediction from face images. *Pattern Recognit. Lett.* 70, 59–65. <https://doi.org/10.1016/j.patrec.2015.11.011>
- Aptula, A.O., Jeliaskova, N.G., Schultz, T.W., Cronin, M.T.D., 2005. The better predictive model: High q^2 for the training set or low root mean square error of prediction for the test set? *QSAR Comb. Sci.* 24, 385–396. <https://doi.org/10.1002/qsar.200430909>
- Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurrealde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.K., Müller, C., Naresh Kumar, S., Nendel, C., O’leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* 5, 143–147. <https://doi.org/10.1038/nclimate2470>
- Awiti, A.O., Walsh, M.G., Shepherd, K.D., Kinyamario, J., 2008. Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence. *Geoderma* 143, 73–84. <https://doi.org/10.1016/j.geoderma.2007.08.021>
- Ay, M., Kişi, Ö., 2017. Estimation of dissolved oxygen by using neural networks and neuro fuzzy computing techniques. *KSCE J. Civ. Eng.* 21, 1631–1639. <https://doi.org/10.1007/s12205-016-0728-6>
- Badiru, A.B., Sieger, D.B., 1998. Neural network as a simulation metamodel in economic analysis of risky projects. *Eur. J. Oper. Res.* 150, 130–142.
- Bagheri Bodaghabadi, M., Martínez-Casasnovas, J.A., Esfandiarpour Borujeni, I., Salehi, M.H., Mohammadi, J., Toomanian, N., 2016. Database extension for digital soil mapping using artificial neural networks. *Arab. J. Geosci.* 9, 701. <https://doi.org/10.1007/s12517-016-2732-z>
- Bahrammirzaee, A., 2010. A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems. *Neural Comput. Appl.* 19, 1165–1195. <https://doi.org/10.1007/s00521-010-0362-z>
- Bai, A., Popp, J., Petó, K., Szóke, I., Harangi-Rákos, M., Gabnai, Z., 2017. The significance of forests and algae in CO₂ balance: A Hungarian case study. *Sustain.* 9, 1–24. <https://doi.org/10.3390/su9050857>

- Bailey, A., Chase, T.N., Cassano, J.J., Noone, D., 2011. Changing temperature inversion characteristics in the U.S. southwest and relationships to large-scale atmospheric circulation. *J. Appl. Meteorol. Climatol.* 50, 1307–1323. <https://doi.org/10.1175/2011JAMC2584.1>
- Bais-Moleman, A.L., Schulp, C.J.E., Verburg, P.H., 2019. Assessing the environmental impacts of production- and consumption-side measures in sustainable agriculture intensification in the European Union. *Geoderma* 338, 555–567. <https://doi.org/10.1016/j.geoderma.2018.11.042>
- Bal, L., Buyle-Bodin, F., 2013. Artificial neural network for predicting drying shrinkage of concrete. *Constr. Build. Mater.* 38, 248–254. <https://doi.org/10.1016/j.conbuildmat.2012.08.043>
- Ballesteros, R., Ortega, J.F., Moreno, M.Á., 2016. FORETo: New software for reference evapotranspiration forecasting. *J. Arid Environ.* 124, 128–141. <https://doi.org/10.1016/j.jaridenv.2015.08.006>
- Barbin, D.F., Mastelini, S.M., Barbon, S., Campos, G.F.C., Barbon, A.P.A.C., Shimokomaki, M., 2016. Digital image analyses as an alternative tool for chicken quality assessment. *Biosyst. Eng.* 144, 85–93. <https://doi.org/10.1016/j.biosystemseng.2016.01.015>
- Barrios, E., Trejo, M.T., 2003. Implications of local soil knowledge for integrated soil management in Latin America. *Geoderma* 111, 217–231. [https://doi.org/10.1016/S0016-7061\(02\)00265-3](https://doi.org/10.1016/S0016-7061(02)00265-3)
- Barry, R.G., Chorley, R.J., 2009. *Atmosphere, weather and climate*. Routledge.
- Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. *J. Microbiol. Methods* 43, 3–31.
- Basma, A.A., Kallas, N., 2004. Modeling soil collapse by artificial neural networks. *Geotech. Geol. Eng.* 22, 427–438. <https://doi.org/10.1023/B:GEGE.0000025044.72718.db>
- Basso, B., Dumont, B., Cammarano, D., Pezzuolo, A., Marinello, F., Sartori, L., 2016. Environmental and economic benefits of variable rate nitrogen fertilization in a nitrate vulnerable zone. *Sci. Total Environ.* 545–546, 227–235. <https://doi.org/10.1016/j.scitotenv.2015.12.104>
- Baumann, K., Schöning, I., Schrupf, M., Ellerbrock, R.H., Leinweber, P., 2016. Rapid assessment of soil organic matter: Soil color analysis and Fourier transform infrared spectroscopy. *Geoderma* 278, 49–57. <https://doi.org/10.1016/j.geoderma.2016.05.012>
- Behrang, M.A., Assareh, E., Ghanbarzadeh, A., Noghrehabadi, A.R., 2010. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. *Sol. Energy* 84, 1468–1480. <https://doi.org/10.1016/j.solener.2010.05.009>
- Belgrano, A., Malmgren, B.A., Lindahl, O., 2002. Application of Artificial Neural Networks (ANN) to Primary Production Time-series Data. *J. Plankton Res.* 23, 651–658. <https://doi.org/10.1093/plankt/23.6.651>

- Benemann, J.R., 1979. Production of nitrogen fertilizer with nitrogen-fixing blue-green algae. *Enzyme Microb. Technol.* 1, 83–90.
- Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J., Warde-Farley, D., Bengio, Y., 2010. Theano: a CPU and GPU Math Expression Compiler, in: SciPy.
- Bergstra, J., Desjardins, G., Lamblin, P., Bengio, Y., 2009. Quadratic polynomials learn better image features. Tech. report, Tech. Rep. 1337, Département d'Informatique Rech. Oper.
- Berndt, C., Haberlandt, U., 2018. Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. *J. Hydrol. Reg. Stud.* 15, 184–202. <https://doi.org/10.1016/j.ejrh.2018.02.002>
- Bi, T.S., Ni, Y.X., Shen, C.M., Wu, F.F., 2000. A novel ANN fault diagnosis system for power systems using dual GA loops in ANN training, in: 2000 Power Engineering Society Summer Meeting (Cat. No. 00CH37134). pp. 425–430. <https://doi.org/10.1109/pess.2000.867624>
- Bi, W., Wang, X., Tang, Z., Tamura, H., 2005. Avoiding the local minima problem in backpropagation algorithm with modified error function. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* 88, 3645–3653. <https://doi.org/10.1093/ietfec/e88-a.12.3645>
- Bilgili, M., Sahin, B., 2010. Comparative analysis of regression and artificial neural network models for wind speed prediction. *Meteorol. Atmos. Phys.* 109, 61–72. <https://doi.org/10.1007/s00703-010-0093-9>
- Bishop, C.M., 1996. Neural networks: a pattern recognition perspective, in: *Handbook of Neural Computation*. pp. 1–23. <https://doi.org/10.1.1.46.8742>
- Bishop, C.M., 1995. *Neural networks for pattern recognition*. Clarendon Press.
- Bommarco, R., Marini, L., Vaissière, B.E., 2012. Insect pollination enhances seed yield, quality, and market value in oilseed rape. *Oecologia* 169, 1025–1032. <https://doi.org/10.1007/s00442-012-2271-6>
- Borowitzka, M.A., 2013. High-value products from microalgae-their development and commercialisation. *J. Appl. Phycol.* 25, 743–756. <https://doi.org/10.1007/s10811-013-9983-9>
- Bouselham, L., Hajji, M., Hajji, B., Bouali, H., 2017. A New MPPT-based ANN for Photovoltaic System under Partial Shading Conditions. *Energy Procedia* 111, 924–933. <https://doi.org/10.1016/j.egypro.2017.03.255>
- Bouttes, M., Bize, N., Maréchal, G., Michel, G., Cristobal, M.S., Martin, G., 2019. Conversion to organic farming decreases the vulnerability of dairy farms. *Agron. Sustain. Dev.* 39. <https://doi.org/10.1007/s13593-019-0565-3>
- Brevik, E.C., Hartemink, A.E., 2010. Early soil knowledge and the birth and development of soil science. *Catena* 83, 23–33. <https://doi.org/10.1016/j.catena.2010.06.011>

- Brosnan, T., Sun, D.W., 2004. Improving quality inspection of food products by computer vision - A review. *J. Food Eng.* 61, 3–16. [https://doi.org/10.1016/S0260-8774\(03\)00183-3](https://doi.org/10.1016/S0260-8774(03)00183-3)
- Bünemann, E.K., Bongiorno, G., Bai, Z., Creamer, R.E., De Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T.W., Mäder, P., Pulleman, M., Sukkel, W., van Groenigen, J.W., Brussaard, L., 2018. Soil quality – A critical review. *Soil Biol. Biochem.* 120, 105–125. <https://doi.org/10.1016/j.soilbio.2018.01.030>
- Cabañas, R., Salmerón, A., Masegosa, A.R., 2019. InferPy: Probabilistic modeling with Tensorflow made easy. *Knowledge-Based Syst.* 168, 25–27. <https://doi.org/10.1016/j.knosys.2018.12.030>
- Cao, F., Ye, H., Wang, D., 2015. A probabilistic learning algorithm for robust modeling using neural networks with random weights. *Inf. Sci. (Ny)*. 313, 62–78. <https://doi.org/10.1016/j.ins.2015.03.039>
- Cao, Q., Ewing, B.T., Thompson, M.A., 2012. Forecasting wind speed with recurrent neural networks. *Eur. J. Oper. Res.* 221, 148–154. <https://doi.org/10.1016/j.ejor.2012.02.042>
- Cao, W., Wang, X., Ming, Z., Gao, J., 2018. A review on neural networks with random weights. *Neurocomputing* 275, 278–287. <https://doi.org/10.1016/j.neucom.2017.08.040>
- Castañeda, C., Moret-Fernández, D., 2013. Superficial Color Patches as a Visual Diagnostic Criterion for Agricultural Management. *Pedosphere* 23, 740–751. [https://doi.org/10.1016/S1002-0160\(13\)60066-1](https://doi.org/10.1016/S1002-0160(13)60066-1)
- Castro, J. de S., Calijuri, M.L., Assemany, P.P., Cecon, P.R., de Assis, I.R., Ribeiro, V.J., 2017. Microalgae biofilm in soil: Greenhouse gas emissions, ammonia volatilization and plant growth. *Sci. Total Environ.* 574, 1640–1648. <https://doi.org/10.1016/j.scitotenv.2016.08.205>
- Cavalcante, R.C., Brasileiro, R.C., Souza, V.L.F., Nobrega, J.P., Oliveira, A.L.I., 2016. Computational Intelligence and Financial Markets: A Survey and Future Directions. *Expert Syst. Appl.* 55, 194–211. <https://doi.org/10.1016/j.eswa.2016.02.006>
- Celenk, M., 1990. A color clustering technique for image segmentation. *Comput. Vision, Graph. Image Process.* 52, 145–170. [https://doi.org/10.1016/0734-189X\(90\)90052-W](https://doi.org/10.1016/0734-189X(90)90052-W)
- Cellamare, M., Rolland, A., Jacquet, S., 2010. Flow cytometry sorting of freshwater phytoplankton. *J. Appl. Phycol.* 22, 87–100. <https://doi.org/10.1007/s10811-009-9439-4>
- Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* 7, 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chen, C., 2015. Handbook of pattern recognition and computer vision. World Scientific Publishing Company.

- Chen, H., Cañizares, C.A., Singh, A., 2001. ANN-based short-term load forecasting in electricity markets. *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.* 2, 411–415. <https://doi.org/10.1109/PESW.2001.916876>
- Chen, J., Yang, S., Li, H., Zhang, B., Lv, J., 2013. Research on geographical environment unit division based on the method of natural breaks (Jenks). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 40, 47–50. <https://doi.org/10.5194/isprsarchives-XL-4-W3-47-2013>
- Chen, W.H., Hsu, S.H., Shen, H.P., 2005. Application of SVM and ANN for intrusion detection. *Comput. Oper. Res.* 32, 2617–2634. <https://doi.org/10.1016/j.cor.2004.03.019>
- Cheng, G., Yang, C., Yao, X., Guo, L., Han, J., 2018. When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs, in: *IEEE Transactions on Geoscience and Remote Sensing*. pp. 2811–2821. <https://doi.org/10.1109/TGRS.2017.2783902>
- Chisti, Y., 2007. Biodiesel from microalgae. *Biotechnol. Adv.* 25, 294–306. <https://doi.org/10.1016/j.biotechadv.2007.02.001>
- Chowdhury, A., Gupta, D., Paswan-Das, D., Bhowmick, A., 2017. ESTIMATION OF REFERENCE EVAPOTRANSPIRATION USING ARTIFICIAL NEURAL NETWORK FOR MOHANPUR, NADIA DISTRICT, WEST BENGAL: A CASE STUDY. *Int. J. Res. Eng. Technol.* 06, 125–130. <https://doi.org/10.15623/ijret.2017.0607021>
- Cires, D.C., Meier, U., Masci, J., Gambardella, L.M., 2003. Flexible, High Performance Convolutional Neural Networks for Image Classification, in: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence Flexible*. pp. 1237–1242. <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-210>
- Cochrane, S., 2014. The Munsell Color System: A scientific compromise from the world of art. *Stud. Hist. Philos. Sci. Part A* 47, 26–41. <https://doi.org/10.1016/j.shpsa.2014.03.004>
- Collet, P., Hélias Arnaud, A., Lardon, L., Ras, M., Goy, R.A., Steyer, J.P., 2011. Life-cycle assessment of microalgae culture coupled to biogas production. *Bioresour. Technol.* 102, 207–214. <https://doi.org/10.1016/j.biortech.2010.06.154>
- Coltelli, P., Barsanti, L., Evangelista, V., Gualtieri, P., 2017. Algae through the looking glass. *Microsc. Res. Tech.* 80, 486–494. <https://doi.org/10.1002/jemt.22820>
- Conforti, M., Castrignanò, A., Robustelli, G., Scarciglia, F., Stelluti, M., Buttafuoco, G., 2015. Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content. *Catena* 124, 60–67. <https://doi.org/10.1016/j.catena.2014.09.004>
- Conforti, P., 2011. Looking ahead in world food and agriculture: perspectives to 2050. Food and Agriculture Organization of the United Nations (FAO), Rome.
- Cooper, M., 2010. Advanced Bash-Scripting Guide An in-depth exploration of the art of shell scripting Table of Contents. Okt 2005 Abrufbar über <http://www.tldp.org/LDP/abs/guide.pdf> Zugriff 1112 2005 2274, 2267–2274. <https://doi.org/10.1002/hyp>

- Curcio, D., Ciraolo, G., D'Asaro, F., Minacapilli, M., 2013. Prediction of Soil Texture Distributions Using VNIR-SWIR Reflectance Spectroscopy. *Procedia Environ. Sci.* 19, 494–503. <https://doi.org/10.1016/j.proenv.2013.06.056>
- Dahikar, S., Rode, S., 2014. Agricultural Crop Yield Prediction Using Artificial Neural Network Approach. *Int. J. Innov. Res. Electr.* 2, 2321–5526.
- Dahl, G., Sainath, T., Hinton, G., 2013. Improving deep neural networks for LVCSR using rectified linear units and dropout, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*. pp. 8609–8613.
- Darji, M.P., Dabhi, V.K., Prajapati, H.B., 2015. Rainfall forecasting using neural network: A survey. *Conf. Proceeding - 2015 Int. Conf. Adv. Comput. Eng. Appl. ICACEA 2015* 706–713. <https://doi.org/10.1109/ICACEA.2015.7164782>
- Das, G., Pattnaik, P.K., Padhy, S.K., 2014. Artificial Neural Network trained by Particle Swarm Optimization for non-linear channel equalization. *Expert Syst. Appl.* 41, 3491–3496. <https://doi.org/10.1016/j.eswa.2013.10.053>
- Dashkova, V., Segev, E., Malashenkov, D., Kolter, R., Vorobjev, I., Barteneva, N.S., 2016. Microalgal cytometric analysis in the presence of endogenous autofluorescent pigments. *Algal Res.* 19, 370–380. <https://doi.org/10.1016/j.algal.2016.05.013>
- Davies, E.R., 2018. *Computer Vision: Principles, Algorithms, Applications, Learning*. Academic Press. <https://doi.org/https://doi.org/10.1016/C2015-0-05563-0>
- De-Bashan, L.E., Bashan, Y., 2004. Recent advances in removing phosphorus from wastewater and its future use as fertilizer (1997-2003). *Water Res.* 38, 4222–4246. <https://doi.org/10.1016/j.watres.2004.07.014>
- De Andrade, G.A., Berenguel, M., Guzmán, J.L., Pagano, D.J., Acién, F.G., 2016. Optimization of biomass production in outdoor tubular photobioreactors. *J. Process Control* 37, 58–69. <https://doi.org/10.1016/j.jprocont.2015.10.001>
- de Myttenaere, A., Golden, B., Le Grand, B., Rossi, F., 2016. Mean Absolute Percentage Error for regression models. *Neurocomputing* 192, 38–48. <https://doi.org/10.1016/j.neucom.2015.12.114>
- de Oliveira, M.M.F., Ebecken, F.F., de Oliveira, J.L.F., de Azevedo Santos, I., 2009. Neural network model to predict a storm surge. *J. Appl. Meteorol. Climatol.* 48, 143–155. <https://doi.org/10.1175/2008JAMC1907.1>
- de Villiers, J., Barnard, E., 1993. Backpropagation neural nets with one and two hidden layers, in: *IEEE Transactions on Neural Networks*. pp. 136–141.
- Decker, W.L., 1994. Developments in agricultural meteorology as a guide to its potential for the twenty-first century. *Agric. For. Meteorol.* 69, 9–25. [https://doi.org/10.1016/0168-1923\(94\)90076-0](https://doi.org/10.1016/0168-1923(94)90076-0)
- del Río, S., Penas, Á., Fraile, R., 2005. Analysis of recent climatic variations in Castile and Leon (Spain). *Atmos. Res.* 73, 69–85. <https://doi.org/10.1016/j.atmosres.2004.06.005>
- Dey, A., Gupta, A.K., Singh, G., 2019. Innovation, investment and enterprise: Climate resilient entrepreneurial pathways for overcoming poverty. *Agric. Syst.* 172, 83–90. <https://doi.org/10.1016/j.agsy.2018.07.013>

- Diamantopoulou, M.J., Georgiou, P.E., Papamichail, D.M., 2011. Performance Evaluation of Artificial Neural Networks in Estimating Reference Evapotranspiration With Minimal Meteorological Data. *Glob. NST J.* 13, 18–27.
- Ding, M., Wang, L., Bi, R., 2011. An ANN-based approach for forecasting the power output of photovoltaic system. *Procedia Environ. Sci.* 11, 1308–1315. <https://doi.org/10.1016/j.proenv.2011.12.196>
- Dong, L., Wei, F., Tan, C., Tang, D., Zhou, M., Xu, K., 2015. Adaptive Recursive Neural Network for Target-dependent Twitter Sentiment Classification 49–54. <https://doi.org/10.3115/v1/p14-2009>
- Drucker, H., Schapire, R.E., Simard, P., 1993. Improving Performance in Neural Networks Using a Boosting Algorithm, in: *Advances in Neural Information Processing Systems*. pp. 42–49.
- Dündar, S., Şahin, I., 2013. Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways. *Transp. Res. Part C Emerg. Technol.* 27, 1–15. <https://doi.org/10.1016/j.trc.2012.11.001>
- Eberhart, R.C., 2014. *Neural Network PC Tools: A Practical Guide*. Academic Press.
- Ebrahimi Kahou, S., Michalski, V., Konda, K., Memisevic, R., Pal, C., 2015. Recurrent Neural Networks for Emotion Recognition in Video, in: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction - ICMI '15*. pp. 467–474. <https://doi.org/10.1145/2818346.2830596>
- Echevarria, R.N., Xiu, T., 2014. Energy related CO2 emissions and the progress on CCS projects: A review. *Renew. Sustain. Energy Rev.* 31, 368–385. <https://doi.org/10.1016/j.rser.2013.12.002>
- Egle, L., Rechberger, H., Krampe, J., Zessner, M., 2016. Phosphorus recovery from municipal wastewater: An integrated comparative technological, environmental and economic assessment of P recovery technologies. *Sci. Total Environ.* 571, 522–542. <https://doi.org/10.1016/j.scitotenv.2016.07.019>
- El Tabach, E., Lancelot, L., Shahrou, I., Najjar, Y., 2007. Use of artificial neural network simulation metamodelling to assess groundwater contamination in a road project. *Math. Comput. Model.* 45, 766–776. <https://doi.org/10.1016/j.mcm.2006.07.020>
- Enzing, C., Ploeg, M., Barbosa, M., Sijtsma, L., 2014. *Microalgae-based products for the food and feed sector: an outlook for Europe*, JRC Scientific and Policy Reports. European Commission. <https://doi.org/10.2791/3339>
- Erb, K.-H., Gingrich, S., Krausmann, F., Haberl, H., 2008. Industrialization, Fossil Fuels, and the Transformation of Land Use. *J. Ind. Ecol.* 12, 686–703. <https://doi.org/10.1111/j.1530-9290.2008.00076.x>
- Ertam, F., 2017. Data classification with deep learning using tensorflow, in: *2nd International Conference on Computer Science and Engineering, UBMK 2017*. pp. 755–758. <https://doi.org/10.1109/UBMK.2017.8093521>

- Estévez, J., Gavilán, P., Giráldez, J. V., 2011. Guidelines on validation procedures for meteorological data from automatic weather stations. *J. Hydrol.* 402, 144–154. <https://doi.org/10.1016/j.jhydrol.2011.02.031>
- Fairchild, M.D., 2013. *Color Appearance Models*, The Wiley-IS&T Series in Imaging Science and Technology. Wiley.
- Falamarzi, Y., Palizdan, N., Huang, Y.F., Lee, T.S., 2014. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). *Agric. Water Manag.* 140, 26–36. <https://doi.org/10.1016/j.agwat.2014.03.014>
- Fan, S., Liao, J.R., Yokoyama, R., Chen, L., Lee, W.J., 2009. Forecasting the wind generation using a two-stage network based on meteorological information. *IEEE Trans. Energy Convers.* 24, 474–482. <https://doi.org/10.1109/TEC.2008.2001457>
- Feng, J., Guo, Y., Zhang, X., Wang, G., Lv, J., Liu, Q., Xie, S., 2016. Identification and characterization of a symbiotic alga from soil bryophyte for lipid profiles. *Biol. Open* 5, 1317–1323. <https://doi.org/10.1242/bio.019992>
- Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., Wang, J., 2015. Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmos. Environ.* 107, 118–128. <https://doi.org/10.1016/j.atmosenv.2015.02.030>
- Filippo, A., Rebelo Torres, A., Kjerfve, B., Monat, A., 2012. Application of Artificial Neural Network (ANN) to improve forecasting of sea level. *Ocean Coast. Manag.* 55, 101–110. <https://doi.org/10.1016/j.ocecoaman.2011.09.007>
- Fitzgerald, H., Palmé, A., Asdal, Å., Endresen, D., Kiviharju, E., Lund, B., Rasmussen, M., Thorbjörnsson, H., Weibull, J., 2019. A regional approach to Nordic crop wild relative in situ conservation planning. *Plant Genet. Resour. Characterisation Util.* 17, 196–207. <https://doi.org/10.1017/S147926211800059X>
- Fleming, K.L., Heermann, D.F., Westfall, D.G., 2004. Evaluating soil color with farmer input and apparent soil electrical conductivity for management zone delineation. *Agron. J.* 96, 1581–1587.
- Forján, E., Navarro, F., Cuasmas, M., Vaquero, I., Ruíz-Domínguez, M.C., Gojkovic, Ž., Vázquez, M., Márquez, M., Mogedas, B., Bermejo, E., Girlich, S., Domínguez, M.J., Vílchez, C., Vega, J.M., Garbayo, I., 2014. Microalgae: Fast-Growth Sustainable Green Factories. *Crit. Rev. Environ. Sci. Technol.* 45, 1705–1755. <https://doi.org/10.1080/10643389.2014.966426>
- Fradique, M., Batista, A.P., Nunes, M.C., Gouveia, L., Bandarra, N.M., Raymundo, A., 2010. Incorporation of *Chlorella vulgaris* and *Spirulina maxima* biomass in pasta products. Part 1: Preparation and evaluation. *J. Sci. Food Agric.* 90, 1656–1664. <https://doi.org/10.1002/jsfa.3999>
- Francis, F.J., 1995. Quality as influenced by color. *Food Qual. Prefer.* 6, 149–155. [https://doi.org/10.1016/0950-3293\(94\)00026-R](https://doi.org/10.1016/0950-3293(94)00026-R)

- Freibauer, A., Rounsevell, M.D.A., Smith, P., Verhagen, J., 2004. Carbon sequestration in the agricultural soils of Europe. *Geoderma* 122, 1–23. <https://doi.org/10.1016/j.geoderma.2004.01.021>
- García-Camacho, F., López-Rosales, L., Sánchez-Mirón, A., Belarbi, E.H., Chisti, Y., Molina-Grima, E., 2016. Artificial neural network modeling for predicting the growth of the microalga *Karodinium veneticum*. *Algal Res.* 14, 58–64. <https://doi.org/10.1016/j.algal.2016.01.002>
- Gautam, C., Ravi, V., 2015. Data imputation via evolutionary computation, clustering and a neural network. *Neurocomputing* 156, 134–142. <https://doi.org/10.1016/j.neucom.2014.12.073>
- Gershon, R.O.N., 2005. Aspects of perception and computation in color vision. *Comput. Vision, Graph. Image Process.* 30, 372. [https://doi.org/10.1016/0734-189x\(85\)90184-7](https://doi.org/10.1016/0734-189x(85)90184-7)
- Ghanbarzadeh, A., Noghrehabadi, A.R., Assareh, E., Behrang, M.A., 2009. Solar radiation forecasting based on meteorological data using artificial neural networks, in: *IEEE International Conference on Industrial Informatics (INDIN)*. IEEE, pp. 227–231. <https://doi.org/10.1109/INDIN.2009.5195808>
- Gholami, R., Fakhari, N., 2017. Support Vector Machine: Principles, Parameters, and Applications, 1st ed, *Handbook of Neural Computation*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-811318-9.00027-2>
- Ghosh, A., Das, S., Ghosh, T., Hazra, S., 2019. Risk of extreme events in delta environment: A case study of the Mahanadi delta. *Sci. Total Environ.* 664, 713–723. <https://doi.org/10.1016/j.scitotenv.2019.01.390>
- Ghritlahre, H.K., Prasad, R.K., 2017. Prediction of Thermal Performance of Unidirectional Flow Porous Bed Solar Air Heater with Optimal Training Function Using Artificial Neural Network. *Energy Procedia* 109, 369–376. <https://doi.org/10.1016/j.egypro.2017.03.033>
- Girolami, A., Napolitano, F., Faraone, D., Braghieri, A., 2013. Measurement of meat color using a computer vision system. *Meat Sci.* 93, 111–118. <https://doi.org/10.1016/j.meatsci.2012.08.010>
- Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks, in: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. pp. 249–256.
- Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks, in: *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*. pp. 315–323. <https://doi.org/10.1.1.208.6449>
- Gnana Sheela, K., Deepa, S.N., 2013. Neural network based hybrid computing model for wind speed prediction. *Neurocomputing* 122, 425–429. <https://doi.org/10.1016/j.neucom.2013.06.008>

- Göçken, M., Özçalici, M., Boru, A., Dosdoğru, A.T., 2016. Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Syst. Appl.* 44, 320–331. <https://doi.org/10.1016/j.eswa.2015.09.029>
- Godhe, A., Anderson, D.M., Rehnstam-Holm, A.S., 2002. PCR amplification of microalgal DNA for sequencing and species identification: Studies on fixatives and algal growth stages. *Harmful Algae* 1, 375–382. [https://doi.org/10.1016/S1568-9883\(02\)00049-5](https://doi.org/10.1016/S1568-9883(02)00049-5)
- Godos, I. de, Blanco, S., García-Encina, P.A., Becares, E., Muñoz, R., 2009. Long-term operation of high rate algal ponds for the bioremediation of piggery wastewaters at high loading rates. *Bioresour. Technol.* 100, 4332–4339. <https://doi.org/10.1016/j.biortech.2009.04.016>
- Gómez, C., Escudero, R., Morales, M.M., Figueroa, F.L., Fernández-Sevilla, J.M., Acién, F.G., 2013. Use of secondary-treated wastewater for the production of *Muriellopsis* sp. *Appl. Microbiol. Biotechnol.* 97, 2239–2249. <https://doi.org/10.1007/s00253-012-4634-7>
- Gray, L.C., Morant, P., 2003. Reconciling indigenous knowledge with scientific assessment of soil fertility changes in southwestern Burkina Faso. *Geoderma* 111, 425–437. [https://doi.org/10.1016/S0016-7061\(02\)00275-6](https://doi.org/10.1016/S0016-7061(02)00275-6)
- Griffiths, J.F., 1994. *Handbook of agricultural meteorology*.
- Guresen, E., Kayakutlu, G., Daim, T.U., 2011. Using artificial neural network models in stock market index prediction. *Expert Syst. Appl.* 38, 10389–10397. <https://doi.org/10.1016/j.eswa.2011.02.068>
- Han, P., Dong, D., Zhao, X., Jiao, L., Lang, Y., 2016. A smartphone-based soil color sensor: For soil type classification. *Comput. Electron. Agric.* 123, 232–241. <https://doi.org/10.1016/j.compag.2016.02.024>
- Hansen, J., Hellin, J., Rosenstock, T., Fisher, E., Cairns, J., Stirling, C., Lamanna, C., van Etten, J., Rose, A., Campbell, B., 2019. Climate risk management and rural poverty reduction. *Agric. Syst.* 172, 28–46. <https://doi.org/10.1016/j.agsy.2018.01.019>
- Hasni, A., Sehli, A., Draoui, B., Bassou, A., Amieur, B., 2012. Estimating global solar radiation using artificial neural network and climate data in the south-western region of Algeria. *Energy Procedia* 18, 531–537. <https://doi.org/10.1016/j.egypro.2012.05.064>
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The elements of statistical learning: data mining, inference and prediction*, 2nd ed. Springer.
- Hausmann, J., Dietrich, P., Vienken, T., Werban, U., 2016. Technique, analysis routines, and application of direct push-driven in situ color logging. *Environ. Earth Sci.* 75. <https://doi.org/10.1007/s12665-016-5515-7>
- Haykin, S.S., 1999. *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, Ontario.

- Hazan, H., Saunders, D.J., Khan, H., Sanghavi, D.T., Siegelmann, H.T., Kozma, R., 2018. BindsNET: A machine learning-oriented spiking neural networks library in Python. *Front. Neuroinform.* 12, 1–18. <https://doi.org/10.3389/fninf.2018.00089>
- He, H.J., Zheng, C., Sun, D.W., 1985. Image Segmentation Techniques. *Comput. Vision, Graph. Image Process.* 29, 100–132. <https://doi.org/10.1016/B978-0-12-802232-0.00002-5>
- Heaton, J., 2008. Introduction to Neural Networks with Java. Heaton Research.
- Helena, L., Rodrigues, R., Arenzon, A., Raya-rodriguez, M.T., Fontoura, N.F., 2011. Algal density assessed by spectrophotometry : A calibration curve for the unicellular algae *Pseudokirchneriella subcapitata*. *J. Environ. Chem. Ecotoxicol.* 3, 225–228.
- Herrero, M., Lotze-Campen, H., Weindl, I., Popp, A., Müller, C., Schmitz, C., Havlík, P., Rolinski, S., 2015. Livestock in a changing climate: production system transitions as an adaptation strategy for agriculture. *Environ. Res. Lett.* 10, 1–12. <https://doi.org/10.1088/1748-9326/10/9/094021>
- Hewes, C.D., 2016. The color of mass culture: Spectral characteristics of a shallow water column through shade-limited algal growth dynamics1. *J. Phycol.* 52, 252–259. <https://doi.org/10.1111/jpy.12393>
- Holzinger, A., Allen, M.C., Deheyn, D.D., 2016. Hyperspectral imaging of snow algae and green algae from aeroterrestrial habitats. *J. Photochem. Photobiol. B Biol.* 162, 412–420. <https://doi.org/10.1016/j.jphotobiol.2016.07.001>
- Hrust, L., Klaić, Z.B., Križan, J., Antonić, O., Hercog, P., 2009. Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. *Atmos. Environ.* 43, 5588–5596. <https://doi.org/10.1016/j.atmosenv.2009.07.048>
- Hu, X.F., Du, Y., Guan, C.L., Xue, Y., Zhang, G.L., 2014. Color variations of the Quaternary Red Clay in southern China and its paleoclimatic implications. *Sediment. Geol.* 303, 15–25. <https://doi.org/10.1016/j.sedgeo.2014.01.006>
- Huang, G. Bin, Babri, H.A., 1998. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. *IEEE Trans. Neural Networks* 9, 224–229. <https://doi.org/10.1109/72.655045>
- Huang, G., Zhu, Q., Siew, C., 2004. Extreme Learning Machine : A New Learning Scheme of Feedforward Neural Networks, in: *IEEE International Joint Conference on Neural Networks*. pp. 985–990. <https://doi.org/10.1109/IJCNN.2004.1380068>
- Huang, S., Wang, L., Liu, L., Hou, Y., Li, L., 2015. Nanotechnology in agriculture, livestock, and aquaculture in China. A review, *Agronomy for Sustainable Development*. <https://doi.org/10.1007/s13593-014-0274-x>
- Hung, N.Q., Babel, M.S., Weesakul, S., Tripathi, N.K., 2009. Hydrology and Earth System Sciences An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrol. Earth Syst. Sci* 13, 1413–1416.
- Hussain, A., Hasnain, S., 2011. Phytostimulation and biofertilization in wheat by cyanobacteria. *J. Ind. Microbiol. Biotechnol.* 38, 85–92.

- Ibáñez-Asensio, S., Marqués-Mateu, A., Moreno-Ramón, H., Balasch, S., 2013. Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosyst. Eng.* 116, 120–129. <https://doi.org/10.1016/j.biosystemseng.2013.07.013>
- Iizumi, T., Ramankutty, N., 2015. How do weather and climate influence cropping area and intensity? *Glob. Food Sec.* 4, 46–50. <https://doi.org/10.1016/j.gfs.2014.11.003>
- Ilie, A., Welch, G., 2005. Ensuring color consistency across multiple cameras, in: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. pp. 1268–1275.
- Isono, T., Yamashita, K., Momose, D., Kobayashi, H., Kitamura, M., Nishiyama, Y., Hosoya, T., Kanda, H., Kudo, A., Okada, N., Yagi, T., Nakata, K., Mineki, S., Tokunaga, E., 2015. Scan-free absorbance spectral imaging $A(x, y, \lambda)$ of single live algal cells for quantifying absorbance of cell suspensions. *PLoS One* 10, 1–16. <https://doi.org/10.1371/journal.pone.0128002>
- Iyer, M.S., Rhinehart, R.R., 2000. A novel method to stop neural network training, in: Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No. 00CH36334). pp. 929–933.
- Jafar, R., Shahrour, I., Juran, I., 2010. Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. *Math. Comput. Model.* 51, 1170–1180. <https://doi.org/10.1016/j.mcm.2009.12.033>
- Jäger, K., Ördög, V., Barnabás, B., 2005. Effect of cyanobacterial and microalgal biomass on anther culture response of wheat (*Triticum aestivum* L.). *Acta Agron. Hungarica* 53, 99–107.
- Jagustović, R., Zougmore, R.B., Kessler, A., Ritsema, C.J., Keesstra, S., Reynolds, M., 2019. Contribution of systems thinking and complex adaptive system attributes to sustainable food production: Example from a climate-smart village. *Agric. Syst.* 171, 65–75. <https://doi.org/10.1016/j.agsy.2018.12.008>
- Jain, A.K., Mao, J., Mohiuddin, K.M., 1996. Artificial neural networks: A tutorial. *Computer (Long. Beach. Calif.)* 29, 31–44. <https://doi.org/10.1109/2.485891>
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning: with Applications in R. Springer.
- Jat, S.L., Parihar, C.M., Singh, A.K., Kumar, B., Choudhary, M., Nayak, H.S., Parihar, M.D., Parihar, N., Meena, B.R., 2019. Energy auditing and carbon footprint under long-term conservation agriculture-based intensive maize systems with diverse inorganic nitrogen management options. *Sci. Total Environ.* 664, 659–668. <https://doi.org/10.1016/j.scitotenv.2019.01.425>
- Jayalakshmi, T., Santhakumaran, A., 2011. Statistical normalization and back propagation for classification. *Int. J. Comput. Theory Eng.* 3, 89–93.
- Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ. Model. Softw.* 16, 309–330. [https://doi.org/10.1016/S1364-8152\(01\)00008-1](https://doi.org/10.1016/S1364-8152(01)00008-1)

- Jha, S.N., Chopra, S., Kingsly, A.R.P., 2007. Modeling of color values for nondestructive evaluation of maturity of mango. *J. Food Eng.* 78, 22–26. <https://doi.org/10.1016/j.jfoodeng.2005.08.048>
- Jiang, J., Trundle, P., Ren, J., 2010. Medical image analysis with artificial neural networks. *Comput. Med. Imaging Graph.* 34, 617–631. <https://doi.org/10.1016/j.compmedimag.2010.07.003>
- Jin, Q., Zhang, J., Shi, M., Huang, J., 2016. Estimating loess plateau average annual precipitation with multiple linear regression kriging and geographically weighted regression kriging. *Water (Switzerland)* 8, 266. <https://doi.org/10.3390/W8060266>
- Johkan, M., Oda, M., Maruo, T., Shinohara, Y., 2011. Crop production and global warming, in: *Global Warming Impacts-Case Studies on the Economy, Human Health, and on Urban and Natural Environments*. IntechOpen.
- Kaastra, I., Boyd, M., 1996. Designing a neural network for forecasting financial time series. *Neurocomputing* 10, 215–236.
- Kakuda, K., Enomoto, T., Miura, S., 2019. Nonlinear activation functions in CNN based on fluid dynamics and its applications. *C. - Comput. Model. Eng. Sci.* 118, 1–14. <https://doi.org/10.31614/cmcs.2019.04676>
- Kang, Y., Khan, S., Ma, X., 2009. Climate change impacts on crop yield, crop water productivity and food security - A review. *Prog. Nat. Sci.* 19, 1665–1674. <https://doi.org/10.1016/j.pnsc.2009.08.001>
- Karlik, B., 2011. Performance analysis of various activation functions in generalized MLP architectures of neural networks. *Int. J. Artif. Intell. Expert Syst.* 1, 111–122.
- Kemmoku, Y., Orita, S., Nakagawa, S., Sakakibara, T., 1999. Daily insolation forecasting using a multi-stage neural network. *Sol. Energy* 66, 193–199. [https://doi.org/10.1016/S0038-092X\(99\)00017-1](https://doi.org/10.1016/S0038-092X(99)00017-1)
- Khashei, M., Bijari, M., 2010. An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Syst. Appl.* 37, 479–489. <https://doi.org/10.1016/j.eswa.2009.05.044>
- Khashman, A., 2010. Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Syst. Appl.* 37, 6233–6239. <https://doi.org/10.1016/j.eswa.2010.02.101>
- Kim, S.J., Lin, H.T., Lu, Z., Süsstrunk, S., Lin, S., Brown, M.S., 2012. A new in-camera imaging model for color computer vision and its application. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 2289–2302. <https://doi.org/10.1109/TPAMI.2012.58>
- Kirillova, N.P., Kemp, D.B., Artemyeva, Z.S., 2017. Colorimetric analysis of soil with flatbed scanners. *Eur. J. Soil Sci.* 68, 420–433. <https://doi.org/10.1111/ejss.12442>
- Kirillova, N.P., Vodyanitskii, Y.N., Sileva, T.M., 2015. Conversion of soil color parameters from the Munsell system to the CIE-L*a*b* system. *Eurasian Soil Sci.* 48, 468–475. <https://doi.org/10.1134/S1064229315050026>

- Kline, D.M., Berardi, V.L., 2005. Revisiting squared-error and cross-entropy functions for training neural network classifiers. *Neural Comput. Appl.* 14, 310–318. <https://doi.org/10.1007/s00521-005-0467-y>
- Köksoy, O., 2006. Multiresponse robust design: Mean square error (MSE) criterion. *Appl. Math. Comput.* 175, 1716–1729. <https://doi.org/10.1016/j.amc.2005.09.016>
- Kong, S., Shen, X., Lin, Z., Mech, R., Fowlkes, C., 2016. Photo Aesthetics Ranking Network with Attributes and Content Adaptation, in: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), *Computer Vision -- ECCV 2016*. Springer International Publishing, Cham, pp. 662–679.
- Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M. V., Fotiadis, D.I., 2015. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* 13, 8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet Classification with Deep Convolutional Neural Networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., pp. 1097–1105.
- Kulkarni, P.A., Dhoble, A.S., Padole, P.M., 2018. Deep neural network-based wind speed forecasting and fatigue analysis of a large composite wind turbine blade. *J. Mech. Eng. Sci.* 233, 2794–2812. <https://doi.org/10.1177/0954406218797972>
- Kumar, R., Aggarwal, R.K., Sharma, J.D., 2015. Comparison of regression and artificial neural network models for estimation of global solar radiations. *Renew. Sustain. Energy Rev.* 52, 1294–1299. <https://doi.org/10.1016/j.rser.2015.08.021>
- Kweon, G., Maxton, C., 2013. Soil organic matter sensing with an on-the-go optical sensor. *Biosyst. Eng.* 115, 66–81.
- Laaboudi, A., Mouhouche, B., Draoui, B., 2012. Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions. *Int. J. Biometeorol.* 56, 831–841. <https://doi.org/10.1007/s00484-011-0485-7>
- Lapins, D., Paskausks, P., Putniece, G., Putnieks, A., 2013. Soil colour spectral analysis. *Eng. Rural Dev.* 93–96.
- Ledda, C., Romero Villegas, G.I., Adani, F., Ación Fernández, F.G., Molina Grima, E., 2015. Utilization of centrate from wastewater treatment for the outdoor production of *Nannochloropsis gaditana* biomass at pilot-scale. *Algal Res.* 12, 17–25. <https://doi.org/10.1016/j.algal.2015.08.002>
- Lee, E., Heng, R.L., Pilon, L., 2013. Spectral optical properties of selected photosynthetic microalgae producing biofuels. *J. Quant. Spectrosc. Radiat. Transf.* 114, 122–135. <https://doi.org/10.1016/j.jqsrt.2012.08.012>
- León, K., Mery, D., Pedreschi, F., León, J., 2006. Color measurement in L*a*b* units from RGB digital images. *Food Res. Int.* 39, 1084–1091. <https://doi.org/10.1016/j.foodres.2006.03.006>

- Leu, S., Sen, C.N., Chang, S.L., 2001. Data mining for tunnel support stability: Neural network approach. *Autom. Constr.* 10, 429–441. [https://doi.org/10.1016/S0926-5805\(00\)00078-9](https://doi.org/10.1016/S0926-5805(00)00078-9)
- Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Model. Softw.* 53, 173–189. <https://doi.org/10.1016/j.envsoft.2013.12.008>
- Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecol. Inform.* 6, 228–241. <https://doi.org/10.1016/j.ecoinf.2010.12.003>
- Li, J., Heap, A.D., Potter, A., Daniell, J.J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. *Environ. Model. Softw.* 26, 1647–1659. <https://doi.org/10.1016/j.envsoft.2011.07.004>
- Li, S., Zhou, C., Wang, S., 2019. Does modernization affect carbon dioxide emissions? A panel data analysis. *Sci. Total Environ.* 663, 426–435. <https://doi.org/10.1016/j.scitotenv.2019.01.373>
- Li, Y., Jiang, W., Yang, L., Wu, T., 2018. On neural networks and learning systems for business computing. *Neurocomputing* 275, 1150–1159. <https://doi.org/10.1016/j.neucom.2017.09.054>
- Liew, S.S., Khalil-Hani, M., Bakhteri, R., 2016. Bounded activation functions for enhanced training stability of deep neural networks on visual pattern recognition problems. *Neurocomputing* 216, 718–734. <https://doi.org/10.1016/j.neucom.2016.08.037>
- Linares-Rodriguez, A., Ruiz-Arias, J.A., Pozo-Vazquez, D., Tovar-Pescador, J., 2013. An artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images. *Energy* 61, 636–645. <https://doi.org/10.1016/j.energy.2013.09.008>
- Liu, L., Huang, G., Baetz, B., Huang, C.Z., Zhang, K., 2019. Integrated GHG emissions and emission relationships analysis through a disaggregated ecologically-extended input-output model; A case study for Saskatchewan, Canada. *Renew. Sustain. Energy Rev.* 106, 97–109. <https://doi.org/10.1016/j.rser.2019.03.001>
- Lobell, D.B., 2014. Climate change adaptation in crop production: Beware of illusions. *Glob. Food Sec.* 3, 72–76. <https://doi.org/10.1016/j.gfs.2014.05.002>
- Long, T.B., Blok, V., Coninx, I., 2016. Barriers to the adoption and diffusion of technological innovations for climate-smart agriculture in Europe: Evidence from the Netherlands, France, Switzerland and Italy. *J. Clean. Prod.* 112, 9–21. <https://doi.org/10.1016/j.jclepro.2015.06.044>
- López, G., Batlles, F.J., Tovar-Pescador, J., 2005. Selection of input parameters to model direct solar irradiance by using artificial neural networks. *Energy* 30, 1675–1684. <https://doi.org/10.1016/j.energy.2004.04.035>
- Losson, O., Porebski, A., Vandenbroucke, N., Macaire, L., 2013. Color texture analysis using CFA chromatic co-occurrence matrices. *Comput. Vis. Image Underst.* 117, 747–763. <https://doi.org/10.1016/j.cviu.2013.03.001>

- Lu, C.E., Adluru, N., Ling, H., Zhu, G., Latecki, L.J., 2010. Contour based object detection using part bundles. *Comput. Vis. Image Underst.* 114, 827–834. <https://doi.org/10.1016/j.cviu.2010.03.009>
- Lu, G.Y., Wong, D.W., 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* 34, 1044–1055. <https://doi.org/10.1016/j.cageo.2007.07.010>
- Luber, G., McGeehin, M., 2008. Climate Change and Extreme Heat Events. *Am. J. Prev. Med.* 35, 429–435. <https://doi.org/10.1016/j.amepre.2008.08.021>
- Lubián, L., Montero, O., Moreno-Garrido, I., Huertas, I.E., Sobrino, C., González-del Valle, M., Pares, G., 2000. Nannochloropsis (Eustigmatophyceae) as source of commercially valuable pigments. *J. Appl. Phycol.* 12, 249–255. <https://doi.org/10.1023/A:1008170915932>
- Luo, M., Liu, T., Meng, F., Duan, Y., Bao, A., Frankl, A., De Maeyer, P., 2019. Spatiotemporal characteristics of future changes in precipitation and temperature in Central Asia. *Int. J. Climatol.* 39, 1571–1588. <https://doi.org/10.1002/joc.5901>
- Maas, A.L., Hannun, A.Y., Ng, A.Y., 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models, in: *Proceedings of the 30 Th International Conference on Machine Learning*. p. 6.
- Madkour, F.F., Kamil, A.E.W., Nasr, H.S., 2012. Production and nutritive value of *Spirulina platensis* in reduced cost media. *Egypt. J. Aquat. Res.* 38, 51–57. <https://doi.org/10.1016/j.ejar.2012.09.003>
- Maejima, Y., Nagatsuka, S., Higashi, T., 2000. Mineralogical composition of iron oxides in red- and yellow-colored soils from Southern Japan and Yunnan, China. *Soil Sci. Plant Nutr.* 46, 571–580. <https://doi.org/10.1080/00380768.2000.10409122>
- Maglogiannis, I.G., 2007. *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies*, *Frontiers in artificial intelligence and applications*. IOS Press.
- Manners, R., van Etten, J., 2018. Are agricultural researchers working on the right crops to enable food and nutrition security under future climates? *Glob. Environ. Chang.* 53, 182–194. <https://doi.org/10.1016/j.gloenvcha.2018.09.010>
- Manohar, T.G., Reddy, V.C.V., 2008. Load Forecasting by a Novel Technique Using ANN. *ARPN J. Eng. Appl. Sci.* 3, 19–25.
- Martinez-Baron, D., Orjuela, G., Renzoni, G., Loboguerrero Rodríguez, A.M., Prager, S.D., 2018. Small-scale farmers in a 1.5°C future: The importance of local social dynamics as an enabling factor for implementation and scaling of climate-smart agriculture. *Curr. Opin. Environ. Sustain.* 31, 112–119. <https://doi.org/10.1016/j.cosust.2018.02.013>

- Martínez-Martínez, V., Gomez-Gil, F.J., Gomez-Gil, J., Ruiz-Gonzalez, R., 2015. An Artificial Neural Network based expert system fitted with Genetic Algorithms for detecting the status of several rotary components in agro-industrial machines using a single vibration signal. *Expert Syst. Appl.* 42, 6433–6441. <https://doi.org/10.1016/j.eswa.2015.04.018>
- Martínez-Martínez, V., Gomez-Gil, J., Machado, M.L., Pinto, F.A.C., 2018. Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops. *PLoS One.* <https://doi.org/10.1371/journal.pone.0196072>
- Martinez, W.L., Martinez, A.R., Solka, J., 2017. *Exploratory data analysis with MATLAB*, 3rd ed. Chapman and Hall/CRC.
- Martins, M.A., Tomasella, J., Dias, C.G., 2019. Maize yield under a changing climate in the Brazilian Northeast: Impacts and adaptation. *Agric. Water Manag.* 216, 339–350. <https://doi.org/10.1016/j.agwat.2019.02.011>
- Mason d’Croz, D., Deryng, D., Elliott, J., Tabeau, A., Von Lampe, M., Schmitz, C., van der Mensbrugghe, D., Heyhoe, E., Kyle, P., Schmid, E., van Meijl, H., Robertson, R., Sands, R.D., Popp, A., Müller, C., Lotze-Campen, H., Nelson, G.C., Ahammad, H., Valin, H., Havlík, P., Robinson, S., Hasegawa, T., Willenbockel, D., Fujimori, S., 2013. Climate change effects on agriculture: Economic responses to biophysical shocks. *Proc. Natl. Acad. Sci.* 111, 3274–3279. <https://doi.org/10.1073/pnas.1222465110>
- Mbuthia, L.W., Acosta-Martínez, V., DeBryun, J., Schaeffer, S., Tyler, D., Odoi, E., Mpheshea, M., Walker, F., Eash, N., 2015. Long term tillage, cover crop, and fertilization effects on microbial community structure, activity: Implications for soil quality. *Soil Biol. Biochem.* 89, 24–34. <https://doi.org/10.1016/j.soilbio.2015.06.016>
- Meehl, G.A., Tebaldi, C., 2004. More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century. *Science* (80-). 305, 994–997. <https://doi.org/10.1126/science.1098704>
- Mekanik, F., Imteaz, M.A., Gato-Trinidad, S., Elmahdi, A., 2013. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *J. Hydrol.* 503, 11–21. <https://doi.org/10.1016/j.jhydrol.2013.08.035>
- Melia, P.M., Cundy, A.B., Sohi, S.P., Hooda, P.S., Busquets, R., 2017. Trends in the recovery of phosphorus in bioavailable forms from wastewater. *Chemosphere* 186, 381–395. <https://doi.org/10.1016/j.chemosphere.2017.07.089>
- Men, B., Long, R., Zhang, J., 2016. Combined forecasting of streamflow based on cross entropy. *Entropy* 18, 1–12. <https://doi.org/10.3390/e18090336>
- Mendez, M., Calvo-Valverde, L., 2016. Assessing the Performance of Several Rainfall Interpolation Methods as Evaluated by a Conceptual Hydrological Model. *Procedia Eng.* 154, 1050–1057. <https://doi.org/10.1016/j.proeng.2016.07.595>
- Mendoza, F., Dejmek, P., Aguilera, J.M., 2006. Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biol. Technol.* 41, 285–295. <https://doi.org/10.1016/j.postharvbio.2006.04.004>

- Menon, A., Mehrotra, K., Mohan, C.K., Ranka, S., 1996. Characterization of a Class of Sigmoid Functions with Applications to Neural Networks. *Neural Networks* 9, 819–835. [https://doi.org/https://doi.org/10.1016/0893-6080\(95\)00107-7](https://doi.org/https://doi.org/10.1016/0893-6080(95)00107-7)
- Mery, D., Pedreschi, F., 2005. Segmentation of colour food images using a robust algorithm. *J. Food Eng.* 66, 353–360. <https://doi.org/10.1016/j.jfoodeng.2004.04.001>
- Meyer, R.S., Cullen, B.R., Whetton, P.H., Robertson, F.A., Eckard, R.J., 2018. Potential impacts of climate change on soil organic carbon and productivity in pastures of south eastern Australia. *Agric. Syst.* 167, 34–46. <https://doi.org/10.1016/j.agsy.2018.08.010>
- Mikolov, T., Karafiát, M., Burget, L., Černocký, J., Khudanpur, S., 2010. Recurrent neural network based language model, in: Eleventh Annual Conference of the International Speech Communication Association. pp. 1045–1048.
- Mindiola, L.P., Freile, G.M., Bertiz, C.S., 2015. Biological Inspiration—Theoretical Framework Mitosis Artificial Neural Networks Unsupervised Algorithm. *Int. J. Commun. Netw. Syst. Sci.* 08, 374–398. <https://doi.org/10.4236/ijcns.2015.89036>
- Miralles, I., Domingo, F., Garcí\`ia-Campos, E., Trasar-Cepeda, C., Leirós, M.C., Gil-Sotres, F., 2012. Biological and microbial activity in biological soil crusts from the Tabernas desert, a sub-arid zone in SE Spain. *Soil Biol. Biochem.* 55, 113–121.
- Mirto, M., Conte, L., Aloisio, G., Distante, C., Vecchio, P., De Giovanni, A., 2015. Measuring cells in phytoplankton images. *Proc. 2015 Int. Conf. High Perform. Comput. Simulation, HPCS 2015* 525–530. <https://doi.org/10.1109/HPCSim.2015.7237085>
- Misra, J., Saha, I., 2010. Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing* 74, 239–255. <https://doi.org/10.1016/j.neucom.2010.03.021>
- Moghaddam, A.H., Moghaddam, M.H., Esfandyari, M., 2016. Stock market index prediction using artificial neural network. *J. Econ. Financ. Adm. Sci.* 21, 89–93. <https://doi.org/10.1016/j.jefas.2016.07.002>
- Mohamed, E.S., Saleh, A.M., Belal, A.B., Gad, A., 2016. Application of near-infrared reflectance for quantitative assessment of soil properties. *Egypt. J. Remote Sens. Sp. Sci.* 1–14. <https://doi.org/10.1016/j.ejrs.2017.02.001>
- Molina Grima, E., Belarbi, E.-H., Ación Fernández, F.G., Robles Medina, A., Chisti, Y., 2003. Recovery of microalgal biomass and metabolites: process options and economics. *Biotechnol. Adv.* 20, 491–515.
- Moraes, R., Valiati, J.F., Gavião Neto, W.P., 2013. Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Syst. Appl.* 40, 621–633. <https://doi.org/10.1016/j.eswa.2012.07.059>
- Morales-Amaral, M. del M., Gómez-Serrano, C., Ación, F.G., Fernández-Sevilla, J.M., Molina-Grima, E., 2015. Production of microalgae using centrate from anaerobic digestion as the nutrient source. *Algal Res.* 9, 297–305. <https://doi.org/10.1016/j.algal.2015.03.018>

- Morowvat, M.H., Ghasemi, Y., 2016. Evaluation of antioxidant properties of some naturally isolated microalgae: Identification and characterization of the most efficient strain. *Biocatal. Agric. Biotechnol.* 8, 263–269. <https://doi.org/10.1016/j.bcab.2016.09.010>
- Mosleh, M. a a, Manssor, H., Malek, S., Milow, P., Salleh, A., 2012. A preliminary study on automated freshwater algae recognition and classification system. *BMC Bioinformatics* 13 Suppl 1, S25. <https://doi.org/10.1186/1471-2105-13-S17-S25>
- Mukherjee, C., Chowdhury, R., Ray, K., 2015. Phosphorus recycling from an unexplored source by polyphosphate accumulating microalgae and cyanobacteria-a step to phosphorus security in agriculture. *Front. Microbiol.* 6, 1–7. <https://doi.org/10.3389/fmicb.2015.01421>
- Mulbry, W., Westhead, E.K., Pizarro, C., Sikora, L., 2005. Recycling of manure nutrients: Use of algal biomass from dairy manure treatment as a slow release fertilizer. *Bioresour. Technol.* 96, 451–458. <https://doi.org/10.1016/j.biortech.2004.05.026>
- Myers, J.A., Curtis, B.S., Curtis, W.R., 2013. Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophys.* 6, 4. <https://doi.org/10.1186/2046-1682-6-4>
- Nafría, D.A., Garrido, N., Álvarez, M. V, Cubero, D., Fernández, M., Villarino, I., Gutiérrez, A., Abia, I., 2013. Atlas Agroclimático de Castilla y León, Junta de Castilla y León. Instituto Tecnológico Agrario de Castilla y León. Ministerio de Agricultura, Alimentación y Medio Ambiente.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines, in: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. pp. 807–814.
- Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic Normals: Test of a new method in the Canadian boreal forest. *Agric. For. Meteorol.* 92, 211–225. [https://doi.org/10.1016/S0168-1923\(98\)00102-6](https://doi.org/10.1016/S0168-1923(98)00102-6)
- Nastos, P.T., Moustris, K.P., Larissi, I.K., Paliatsos, A.G., 2013. Rain intensity forecast using Artificial Neural Networks in Athens, Greece. *Atmos. Res.* 119, 153–160. <https://doi.org/10.1016/j.atmosres.2011.07.020>
- Nawi, N.M., Atomi, W.H., Rehman, M.Z., 2013. The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. *Procedia Technol.* 11, 32–39. <https://doi.org/10.1016/j.protcy.2013.12.159>
- Nayak, D.R., Mahapatra, A., Mishra, P., 2013. A Survey on Rainfall Prediction using Artificial Neural Network. *Int. J. Comput. Appl.* 72, 32–40.
- Naylor, R.L., Battisti, D.S., Tewksbury, J.J., Tigchelaar, M., Deutsch, C.A., Merrill, S.C., Huey, R.B., 2018. Increase in crop losses to insect pests in a warming climate. *Science* (80-). 361, 916–919. <https://doi.org/10.1126/science.aat3466>

- Negrov, D., Karandashev, I., Shakirov, V., Matveyev, Y., Dunin-Barkowski, W., Zenkevich, A., 2017. An approximate backpropagation learning rule for memristor based neural networks using synaptic plasticity. *Neurocomputing* 237, 193–199. <https://doi.org/10.1016/j.neucom.2016.10.061>
- Nejat, P., Jomehzadeh, F., Taheri, M.M., Gohari, M., Muhd, M.Z., 2015. A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). *Renew. Sustain. Energy Rev.* 43, 843–862. <https://doi.org/10.1016/j.rser.2014.11.066>
- Ng, C.L., Chen, Q.Q., Chua, J.J., Hemond, H.F., 2017. A multi-platform optical sensor for in vivo and in vitro algae classification. *Sensors (Switzerland)* 17, 1–14. <https://doi.org/10.3390/s17040912>
- Nguyen, A., Yosinski, J., Clune, J., 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 427–436.
- Nozue, S., Katayama, M., Terazima, M., Kumazaki, S., 2017. Comparative study of thylakoid membranes in terminal heterocysts and vegetative cells from two cyanobacteria, *Rivularia M-261* and *Anabaena variabilis*, by fluorescence and absorption spectral microscopy. *Biochim. Biophys. Acta - Bioenerg.* 1858, 742–749. <https://doi.org/10.1016/j.bbabi.2017.05.007>
- O. Rauff, K., Bello, R., 2015. A Review of Crop Growth Simulation Models as Tools for Agricultural Meteorology. *Agric. Sci.* 06, 1098–1105. <https://doi.org/10.4236/as.2015.69105>
- Officer, S.J., Kravchenko, A., Bollero, G.A., Sudduth, K.A., Kitchen, N.R., Wiebold, W.J., Palm, H.L., Bullock, D.G., 2004. Relationships between soil bulk electrical conductivity and the principal component analysis of topography and soil fertility values. *Plant Soil* 258, 269–280.
- Oh, K.S., Jung, K., 2004. GPU implementation of neural networks. *Pattern Recognit.* 37, 1311–1314. <https://doi.org/10.1016/j.patcog.2004.01.013>
- Olaofe, Z.O., 2014. A 5-day wind speed & power forecasts using a layer recurrent neural network (LRNN). *Sustain. Energy Technol. Assessments* 6, 1–24. <https://doi.org/10.1016/j.seta.2013.12.001>
- Olatunji, S.O., Arif, H., 2016. Identification of Erythematous-Squamous Skin Diseases Using Extreme Learning Machine and Artificial Neural Network. *ICTACT J. Soft Comput.* 4, 627–632. <https://doi.org/10.21917/ijsc.2013.0090>
- Oluín, E.J., 2012. Dual purpose microalgae – bacteria-based systems that treat wastewater and produce biodiesel and chemical products within a Biore fi nery. *Biotechnol. Adv.* 30, 1031–1046. <https://doi.org/10.1016/j.biotechadv.2012.05.001>
- Örkcü, H.H., Bal, H., 2011. Comparing performances of backpropagation and genetic algorithms in the data classification. *Expert Syst. Appl.* 38, 3703–3709. <https://doi.org/10.1016/j.eswa.2010.09.028>

- Özbay, Y., Kara, S., Latifoğlu, F., Ceylan, R., Ceylan, M., 2007. Complex-valued wavelet artificial neural network for Doppler signals classifying. *Artif. Intell. Med.* 40, 143–156. <https://doi.org/10.1016/j.artmed.2007.02.001>
- Ozturk, C., Karaboga, D., 2011. Hybrid Artificial Bee Colony algorithm for neural network training. 2011 IEEE Congr. Evol. Comput. CEC 2011 84–88. <https://doi.org/10.1109/CEC.2011.5949602>
- Palani, S., Liong, S.Y., Tkalich, P., 2008. An ANN application for water quality forecasting. *Mar. Pollut. Bull.* 56, 1586–1597. <https://doi.org/10.1016/j.marpolbul.2008.05.021>
- Pancha, I., Chokshi, K., George, B., Ghosh, T., Paliwal, C., Maurya, R., Mishra, S., 2014. Nitrogen stress triggered biochemical and morphological changes in the microalgae *Scenedesmus* sp. *CCNM 1077. Bioresour. Technol.* 156, 146–154. <https://doi.org/10.1016/j.biortech.2014.01.025>
- Pancha, I., Chokshi, K., Mishra, S., 2015. Enhanced biofuel production potential with nutritional stress amelioration through optimization of carbon source and light intensity in *Scenedesmus* sp. *CCNM 1077. Bioresour. Technol.* 179, 565–572. <https://doi.org/10.1016/j.biortech.2014.12.079>
- Panda, S., Mohapatra, P.K., Prasada Panigrahi, S., 2015. A new training scheme for neural networks and application in non-linear channel equalization. *Appl. Soft Comput. J.* 27, 47–52. <https://doi.org/10.1016/j.asoc.2014.10.040>
- Park, D.C., El-Sharkawi, M., Marks, R., Atlas, L., Damborg, M., 1991. Electric load forecasting using an artificial neural network. *IEEE Trans. Power Syst.* 6, 442–449. <https://doi.org/10.1109/59.76685>
- Partal, T., Cigizoglu, H.K., Kahya, E., 2015. Daily precipitation predictions using three different wavelet neural network algorithms by meteorological data. *Stoch. Environ. Res. Risk Assess.* 29, 1317–1329. <https://doi.org/10.1007/s00477-015-1061-1>
- Patanaik, A., Ong, J.L., Gooley, J.J., Ancoli-Israel, S., Chee, M.W.L., 2018. An end-to-end framework for real-time automatic sleep stage classification. *Sleep* 41, 1–11. <https://doi.org/10.1093/sleep/zsy041>
- Patel, J., Shah, S., Thakkar, P., Kotecha, K., 2015. Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Syst. Appl.* 42, 259–268. <https://doi.org/10.1016/j.eswa.2014.07.040>
- Pearce, R.S., 2001. Plant freezing and damage. *Ann. Bot.* 87, 417–424. <https://doi.org/10.1006/anbo.2000.1352>
- Pedreschi, F., León, J., Mery, D., Moyano, P., 2006. Development of a computer vision system to measure the color of potato chips. *Food Res. Int.* 39, 1092–1098. <https://doi.org/10.1016/j.foodres.2006.03.009>
- Peniuk, G.T., Schnurr, P.J., Allen, D.G., 2016. Identification and quantification of suspended algae and bacteria populations using flow cytometry: applications for algae biofuel and biochemical growth systems. *J. Appl. Phycol.* 28, 95–104. <https://doi.org/10.1007/s10811-015-0569-6>

- Phansalkar, V. V., Sastry, P., 1994. Analysis of the Back-Propagation Algorithm with Momentum. *IEEE Trans. Neural Networks* 5, 505–506.
- Piotrowski, A.P., Napiorkowski, J.J., 2013. A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *J. Hydrol.* 476, 97–111. <https://doi.org/10.1016/j.jhydrol.2012.10.019>
- Pittelkow, C.M., Liang, X., Linqvist, B.A., Groenigen, L.J. Van, Lee, J., Lundy, M.E., Gestel, N. Van, Six, J., Venterea, R.T., Kessel, C. Van, 2015. Productivity limits and potentials of the principles of conservation agriculture. *Nature* 517, 365–368. <https://doi.org/10.1038/nature13809>
- Poeplau, C., Don, A., 2015. Carbon sequestration in agricultural soils via cultivation of cover crops - A meta-analysis. *Agric. Ecosyst. Environ.* 200, 33–41. <https://doi.org/10.1016/j.agee.2014.10.024>
- Pothula, A.K., Igathinathane, C., Shen, J., Nichols, K., Archer, D., 2015. Milled industrial beet color kinetics and total soluble solid contents by image analysis. *Ind. Crops Prod.* 65, 159–169. <https://doi.org/10.1016/j.indcrop.2014.12.001>
- Potočnik, M., Klemenc, B., Solina, F., Herlec, U., 2015. Computer aided method for colour calibration and analysis of digital rock photographs. *Geologija* 58, 247–260. <https://doi.org/10.5474/geologija.2015.020>
- Prasad, R., Bhattacharyya, A., Nguyen, Q.D., 2017. Nanotechnology in sustainable agriculture: Recent developments, challenges, and perspectives. *Front. Microbiol.* 8, 1–13. <https://doi.org/10.3389/fmicb.2017.01014>
- Pratibha, G., Srinivas, I., V. Rao, K., M.K. Raju, B., Shanker, A.K., Jha, A., Uday Kumar, M., Srinivasa Rao, K., Sammi Reddy, K., 2019. Identification of environment friendly tillage implement as a strategy for energy efficiency and mitigation of climate change in semiarid rainfed agro ecosystems. *J. Clean. Prod.* 214, 524–535. <https://doi.org/10.1016/j.jclepro.2018.12.251>
- Prats-Montalbán, J.M., de Juan, A., Ferrer, A., 2011. Multivariate image analysis: A review with applications. *Chemom. Intell. Lab. Syst.* 107, 1–23. <https://doi.org/10.1016/j.chemolab.2011.03.002>
- Priddy, K.L., Keller, P.E., 2005. *Artificial Neural Networks: An Introduction*, SPIE tutorial texts. SPIE Press.
- Promdaen, S., Wattuya, P., Sanevas, N., 2014. Automated microalgae image classification. *Procedia Comput. Sci.* 29, 1981–1992. <https://doi.org/10.1016/j.procs.2014.05.182>
- Punitha, K., Devaraj, D., Sakthivel, S., 2013. Artificial neural network based modified incremental conductance algorithm for maximum power point tracking in photovoltaic system under partial shading conditions. *Energy* 62, 330–340. <https://doi.org/10.1016/j.energy.2013.08.022>
- Qin, J., Liang, J., Chen, T., Lei, X., Kang, A., 2019. Simulating and predicting of hydrological time series based on tensorflow deep learning. *Polish J. Environ. Stud.* 28, 795–802. <https://doi.org/10.15244/pjoes/81557>

- Ramachandran, R., Rajeev, D.C., Krishnan, S.G., Subathra, P., 2015. Deep learning in neural networks: An overview. *Int. J. Appl. Eng. Res.* 10, 25433–25448. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Rampasek, L., Goldenberg, A., 2016. TensorFlow: Biology's Gateway to Deep Learning? *Cell Syst.* 2, 12–14. <https://doi.org/10.1016/j.cels.2016.01.009>
- Rao, G.S.L.H.V.P., 2008. *Agricultural Meteorology*. PHI Learning.
- Raposo, M.F.D.J., Morais, R.M.S.C. De, 2011. *Chlorella vulgaris* as soil amendment: Influence of encapsulation and enrichment with rhizobacteria. *Int. J. Agric. Biol.* 13, 719–724.
- Ray, D.K., Gerber, J.S., Macdonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. *Nat. Commun.* 6, 1–9. <https://doi.org/10.1038/ncomms6989>
- Recio, J., Vallejo, A., Le-Noë, J., Garnier, J., García-Marco, S., Álvarez, J.M., Sanz-Cobena, A., 2018. The effect of nitrification inhibitors on NH₃ and N₂O emissions in highly N fertilized irrigated Mediterranean cropping systems. *Sci. Total Environ.* 636, 427–436. <https://doi.org/10.1016/j.scitotenv.2018.04.294>
- Recknagel, F., 1997. ANNA – Artificial Neural Network model for predicting species abundance and succession of blue-green algae. *Hydrobiologia* 349, 47–57. <https://doi.org/10.1023/a:1003041427672>
- Renuka, N., Prasanna, R., Sood, A., Ahluwalia, A.S., Bansal, R., Babu, S., Singh, R., Shivay, Y.S., Nain, L., 2016. Exploring the efficacy of wastewater-grown microalgal biomass as a biofertilizer for wheat. *Environ. Sci. Pollut. Res.* 23, 6608–6620. <https://doi.org/10.1007/s11356-015-5884-6>
- Rickman, M., Pellegrino, J., Hock, J., Shaw, S., Freeman, B., 2013. Life-cycle and techno-economic analysis of utility-connected algae systems. *Algal Res.* 2, 59–65. <https://doi.org/10.1016/j.algal.2012.11.003>
- Rios, L.M., Sahinidis, N. V., 2013. Derivative-free optimization: A review of algorithms and comparison of software implementations. *J. Glob. Optim.* 56, 1247–1293. <https://doi.org/10.1007/s10898-012-9951-y>
- Robitu, M., Musy, M., Inard, C., Groleau, D., 2006. Modeling the influence of vegetation and water pond on urban microclimate. *Sol. Energy* 80, 435–447. <https://doi.org/10.1016/j.solener.2005.06.015>
- Romeu, P., Zamora-Martínez, F., Botella-Rocamora, P., Pardo, J., 2013. Time-series forecasting of indoor temperature using pre-trained deep neural networks, in: *International Conference on Artificial Neural Networks*. pp. 451–458.
- Rosenbaum, R.A., Johnson, G.P., 1984. *Calculus: Basic Concepts and Applications*. Cambridge University Press.
- Rowlings, D.W., Grace, P.R., Scheer, C., Kiese, R., 2013. Influence of nitrogen fertiliser application and timing on greenhouse gas emissions from a lychee (*Litchi chinensis*) orchard in humid subtropical Australia. *Agric. Ecosyst. Environ.* 179, 168–178. <https://doi.org/10.1016/j.agee.2013.08.013>

- Rubio Camacho, F., García Camacho, F., Fernández Sevilla, J.M., Chisti, Y., Molina Grima, E., 2003. A mechanistic model of photosynthesis in microalgae. *Biotechnol. Bioeng.* 81, 459–473. <https://doi.org/10.1002/bit.10492>
- Rui, Y., El-Keib, A.A., 1995. A review of ANN-based short-term load forecasting models, in: *Proceedings of the Twenty-Seventh Southeastern Symposium on System Theory*. pp. 78–82. <https://doi.org/10.1109/ssst.1995.390613>
- Rumpler, M., Tscharf, A., Mostegel, C., Daftry, S., Hoppe, C., Pretenthaler, R., Fraundorfer, F., Mayer, G., Bischof, H., 2017. Evaluations on multi-scale camera networks for precise and geo-accurate reconstructions from aerial and terrestrial images with user guidance. *Comput. Vis. Image Underst.* 157, 255–273. <https://doi.org/10.1016/j.cviu.2016.04.008>
- Rusco, E., Jones, R., Bidoglio, G., 2001. *Organic Matter in the soils of Europe: Present status and future trends*, Institute for Environment and Sustainability European Soil Bureau.
- Rusk, N., 2015. Deep learning. *Nat. Methods* 13, 35. <https://doi.org/10.1038/nmeth.3707>
- Russell, S.J., Norvig, P., 2016. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.
- Saigal, S., Mehrotra, D., 2012. Performance comparison of time series data using predictive data mining techniques. *Adv. Inf. Min.* 4, 57–66.
- Sánchez-Marañón, M., Ortega, R., Miralles, I., Soriano, M., 2007. Estimating the mass wetness of Spanish arid soils from lightness measurements. *Geoderma* 141, 397–406. <https://doi.org/10.1016/j.geoderma.2007.07.005>
- Santos-Ballardo, D.U., Rossi, S., Hernández, V., Gómez, R.V., del Carmen Rendón-Unceta, M., Caro-Corrales, J., Valdez-Ortiz, A., 2015. A simple spectrophotometric method for biomass measurement of important microalgae species in aquaculture. *Aquaculture* 448, 87–92. <https://doi.org/10.1016/j.aquaculture.2015.05.044>
- Sanz-Cobena, A., Lassaletta, L., Aguilera, E., Del Prado, A., Garnier, J., Billen, G., Iglesias, A., Sanchez, B., Guardia, G., Abalos, D., Plaza-Bonilla, D., Puigdueta-Bartolomé, I., Moral, R., Galán, E., Arriaga, H., Merino, P., Infante-Amate, J., Mejjide, A., Pardo, G., Álvaro-Fuentes, J., Gilsanz, C., Báez, D., Doltra, J., González-Ubierna, S., Cayuela, M., Menéndez, S., Díaz-Pinés, E., Le-Noë, J., Quemada, M., Estellés, F., Calvet, S., van Grinsven, H.J., Westhoek, H., Sanz, M., Gimeno, B., Vallejo, A., Smith, P., 2017a. Strategies for greenhouse gas emissions mitigation in Mediterranean agriculture: A review. *Agric. Ecosyst. Environ.* 238, 5–24. <https://doi.org/10.1016/j.agee.2016.09.038>
- Sanz-Cobena, A., Lassaletta, L., Garnier, J., Smith, P., 2017b. Mitigation and quantification of greenhouse gas emissions in Mediterranean cropping systems. *Agric. Ecosyst. Environ.* 238, 1–4. <https://doi.org/10.1016/j.agee.2016.12.032>
- Sapkota, T.B., Vetter, S.H., Jat, M.L., Sirohi, S., Shirsath, P.B., Singh, R., Jat, H.S., Smith, P., Hillier, J., Stirling, C.M., 2019. Cost-effective opportunities for climate change mitigation in Indian agriculture. *Sci. Total Environ.* 655, 1342–1354. <https://doi.org/10.1016/j.scitotenv.2018.11.225>

- Saravanan, G., Yamuna, G., Nandhini, S., 2016. Real time implementation of RGB to HSV/HSI/HSL and its reverse color space models, in: International Conference on Communication and Signal Processing. IEEE, pp. 462–466. <https://doi.org/10.1109/ICCSP.2016.7754179>
- Sarkar, A., Pandey, P., 2015. River Water Quality Modelling Using Artificial Neural Network Technique. *Aquat. Procedia* 4, 1070–1077. <https://doi.org/10.1016/j.aqpro.2015.02.135>
- Schanda, J., 2007. Colorimetry: Understanding the CIE System. Wiley, Vienna.
- Şenkal, O., Kuleli, T., 2009. Estimation of solar radiation over Turkey using artificial neural network and satellite data. *Appl. Energy* 86, 1222–1228. <https://doi.org/10.1016/j.apenergy.2008.06.003>
- Serive, B., Nicolau, E., Bérard, J.B., Kaas, R., Pasquet, V., Picot, L., Cadoret, J.P., 2017. Community analysis of pigment patterns from 37 microalgae strains reveals new carotenoids and porphyrins characteristic of distinct strains and taxonomic groups. *PLoS One* 12, 1–35. <https://doi.org/10.1371/journal.pone.0171872>
- Severyn, A., Moschitti, A., 2015. UNITN: Training Deep Convolutional Neural Network for Twitter Sentiment Classification, in: Proceedings of the 9th International Workshop on Semantic Evaluation. pp. 464–469. <https://doi.org/10.18653/v1/s15-2079>
- Shaaban, M.M., 2001. Nutritional Status and Growth of Maize Plants as Affected by Green Microalgae as Soil Additives. *J. Biol. Sci.* 1, 475–479. <https://doi.org/10.3923/jbs.2001.475.479>
- Shah, M.T., Zodape, S.T., Chaudhary, D.R., Eswaran, K., Chikara, J., 2013. Seaweed sap as an alternative liquid fertilizer for yield and quality improvement of wheat. *J. Plant Nutr.* 36, 192–200.
- Shao, Y., Pan, J., Zhang, C., Jiang, L., He, Y., 2014. Detection in situ of carotenoid in microalgae by transmission spectroscopy. *Comput. Electron. Agric.* 112, 121–127. <https://doi.org/10.1016/j.compag.2014.10.008>
- Sharma, B., K. Venugopalan, P., 2014. Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images. *IOSR J. Comput. Eng.* 16, 31–35. <https://doi.org/10.9790/0661-16123135>
- Sharon Mano Pappu, J., Vijayakumar, G.K., Ramamurthy, V., 2013. Artificial neural network model for predicting production of *Spirulina platensis* in outdoor culture. *Bioresour. Technol.* 130, 224–230. <https://doi.org/10.1016/j.biortech.2012.12.082>
- Shields, R.J., Lupatsch, I., 2012. Algae for Aquaculture and Animal Feeds. *Tech. – Theor. und Prax.* 21, 23–37.
- Shu Fan, Luonan Chen, Wei-Jen Lee, 2009. Short-Term Load Forecasting Using Comprehensive Combination Based on Multimeteorological Information. *IEEE Trans. Ind. Appl.* 45, 1460–1466. <https://doi.org/10.1109/tia.2009.2023571>
- Sibi, P., Allwyn Jones, S., Siddarth, P., 2013. Analysis of different activation functions using back propagation neural networks. *J. Theor. Appl. Inf. Technol.* 47, 1264–1268.

- Sideratos, G., Hatziargyriou, N.D., 2007. An advanced statistical method for wind power forecasting. *IEEE Trans. Power Syst.* 22, 258–265. <https://doi.org/10.1109/TPWRS.2006.889078>
- Singh, B.K., Verma, K., Thoke, A.S., 2015. Adaptive gradient descent backpropagation for classification of breast tumors in ultrasound imaging. *Procedia Comput. Sci.* 46, 1601–1609. <https://doi.org/10.1016/j.procs.2015.02.091>
- Singh, U.B., Ahluwalia, A.S., 2013. Microalgae: A promising tool for carbon sequestration. *Mitig. Adapt. Strateg. Glob. Chang.* 18, 73–95. <https://doi.org/10.1007/s11027-012-9393-3>
- Singh, V., Gupta, I., Gupta, H.O., 2007. ANN-based estimator for distillation using Levenberg-Marquardt approach. *Eng. Appl. Artif. Intell.* 20, 249–259. <https://doi.org/10.1016/j.engappai.2006.06.017>
- Singh, V.K., 2016. Proposing Solution to XOR Problem Using Minimum Configuration MLP. *Procedia Comput. Sci.* 85, 263–270. <https://doi.org/10.1016/j.procs.2016.05.231>
- Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E.A., Haberl, H., Harper, R., House, J., Jafari, M., others, 2014. Agriculture, forestry and other land use (AFOLU), in: *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms, in: *Advances in Neural Information Processing Systems*. pp. 2951–2959. [https://doi.org/10.1016/s2468-2667\(17\)30214-1](https://doi.org/10.1016/s2468-2667(17)30214-1)
- Sobczuk, T.M., Camacho, F.G., Rubio, F.C., Fernández, F.G.A., Grima, E.M., 2002. Carbon dioxide uptake efficiency by outdoor microalgal cultures in tubular airlift photobioreactors. *Biotechnol. Bioeng.* 67, 465–475. [https://doi.org/10.1002/\(sici\)1097-0290\(20000220\)67:4<465::aid-bit10>3.0.co;2-9](https://doi.org/10.1002/(sici)1097-0290(20000220)67:4<465::aid-bit10>3.0.co;2-9)
- Sola, J., Sevilla, J., 1997. Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. Nucl. Sci.* 44, 1464–1468. <https://doi.org/10.1109/23.589532>
- Solovchenko, A., Solovchenko, O., Khozin-Goldberg, I., Didi-Cohen, S., Pal, D., Cohen, Z., Boussiba, S., 2013. Probing the effects of high-light stress on pigment and lipid metabolism in nitrogen-starving microalgae by measuring chlorophyll fluorescence transients: Studies with a $\delta 5$ desaturase mutant of *Parietochloris incisa* (Chlorophyta, Trebouxiophyceae). *Algal Res.* 2, 175–182. <https://doi.org/10.1016/j.algal.2013.01.010>
- Soni, R.A., Sudhakar, K., Rana, R.S., 2017. Spirulina – From growth to nutritional product: A review. *Trends Food Sci. Technol.* 69, 157–171. <https://doi.org/10.1016/j.tifs.2017.09.010>
- Specht, D.F., 1991. A general regression neural network. *Neural Networks, IEEE Trans.* 2, 568–576. <https://doi.org/10.1109/72.97934>

- Springmann, M., Mason-D’Croz, D., Robinson, S., Garnett, T., Godfray, H.C.J., Gollin, D., Rayner, M., Ballon, P., Scarborough, P., 2016. Global and regional health effects of future food production under climate change: A modelling study. *Lancet* 387, 1937–1946. [https://doi.org/10.1016/S0140-6736\(15\)01156-3](https://doi.org/10.1016/S0140-6736(15)01156-3)
- Squire, L., Berg, D., Bloom, F.E., Du Lac, S., Ghosh, A., Spitzer, N.C., 2008. Fundamental neuroscience, *Journal of Neuropathology and Experimental Neurology*. Academic Press. <https://doi.org/10.1097/00005072-199712000-00013>
- Stanco, F., Tanasi, D., Bruna, A., Maugeri, V., 2011. Automatic color detection of archaeological pottery with Munsell system. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 6978 LNCS, 337–346. https://doi.org/10.1007/978-3-642-24085-0_35
- Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., Ludwig, C., 2015. The trajectory of the Anthropocene: The Great Acceleration. *Anthr. Rev.* 2, 81–98. <https://doi.org/10.1177/2053019614564785>
- Stenberg, B., 2010. Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma* 158, 15–22. <https://doi.org/10.1016/j.geoderma.2010.04.008>
- Stiglitz, R., Mikhailova, E., Post, C., Schlautman, M., Sharp, J., 2017. Using an inexpensive color sensor for rapid assessment of soil organic carbon. *Geoderma* 286, 98–103. <https://doi.org/10.1016/j.geoderma.2016.10.027>
- Stiglitz, R., Mikhailova, E., Post, C., Schlautman, M., Sharp, J., 2016. Evaluation of an inexpensive sensor to measure soil color. *Comput. Electron. Agric.* 121, 141–148. <https://doi.org/10.1016/j.compag.2015.11.014>
- Sudarsan, B., Ji, W., Biswas, A., Adamchuk, V., 2016. Microscope-based computer vision to characterize soil texture and soil organic matter. *Biosyst. Eng.* 152, 41–50. <https://doi.org/10.1016/j.biosystemseng.2016.06.006>
- Sumriddetchkajorn, S., Chaitavon, K., Intaravanne, Y., 2014. Mobile-platform based colorimeter for monitoring chlorine concentration in water. *Sensors Actuators, B Chem.* 191, 561–566. <https://doi.org/10.1016/j.snb.2013.10.024>
- Swarnalakshmi, K., Prasanna, R., Kumar, A., Pattnaik, S., Chakravarty, K., Shivay, Y.S., Singh, R., Saxena, A.K., 2013. Evaluating the influence of novel cyanobacterial biofilmed biofertilizers on soil fertility and plant nutrition in wheat. *Eur. J. Soil Biol.* 55, 107–116.
- Sydney, E.B., da Silva, T.E., Tokarski, A., Novak, A.C., de Carvalho, J.C., Woiciechowski, A.L., Larroche, C., Soccol, C.R., 2011. Screening of microalgae with potential for biodiesel production and nutrient removal from treated domestic sewage. *Appl. Energy* 88, 3291–3294. <https://doi.org/10.1016/j.apenergy.2010.11.024>
- Tarlak, F., Ozdemir, M., Melikoglu, M., 2016. Computer vision system approach in colour measurements of foods: Part I. development of methodology. *Food Sci. Technol.* 36, 382–388. <https://doi.org/10.1590/1678-457x.11615>

- Tealab, A., Hefny, H., Badr, A., 2017. Forecasting of nonlinear time series using ANN. *Futur. Comput. Informatics J.* 2, 39–47. <https://doi.org/10.1016/j.fcij.2017.05.001>
- ten Bosch, J.J., Coops, J.C., 1995. Tooth color and reflectance as related to light scattering and enamel hardness. *J Dent Res* 74, 374–380. <https://doi.org/10.1177/00220345950740011401>
- TerAvest, D., Wandschneider, P.R., Thierfelder, C., Reganold, J.P., 2019. Diversifying conservation agriculture and conventional tillage cropping systems to improve the wellbeing of smallholder farmers in Malawi. *Agric. Syst.* 171, 23–35. <https://doi.org/10.1016/j.agsy.2019.01.004>
- Teshome, Y., Biazin, B., Wolka, K., Burka, A., 2018. Evaluating performance of traditional surface irrigation techniques in Cheleleka watershed in Central Rift Valley, Ethiopia. *Appl. Water Sci.* 8, 1–14. <https://doi.org/10.1007/s13201-018-0862-z>
- Tetko, I. V., Livingstone, D.J., Luik, A.I., 1995. Neural Network Studies. 1. Comparison of Overfitting and Overtraining. *J. Chem. Inf. Comput. Sci.* 35, 826–833. <https://doi.org/10.1021/ci00027a006>
- Thornton, P.K., Whitbread, A., Baedeker, T., Cairns, J., Claessens, L., Baethgen, W., Bunn, C., Friedmann, M., Giller, K.E., Herrero, M., Howden, M., Kilcline, K., Nangia, V., Ramirez-Villegas, J., Kumar, S., West, P.C., Keating, B., 2018. A framework for priority-setting in climate smart agriculture research. *Agric. Syst.* 167, 161–175. <https://doi.org/10.1016/j.agsy.2018.09.009>
- Thorsen, S.M., Höglind, M., 2010. Assessing winter survival of forage grasses in Norway under future climate scenarios by simulating potential frost tolerance in combination with simple agroclimatic indices. *Agric. For. Meteorol.* 150, 1272–1282. <https://doi.org/10.1016/j.agrformet.2010.05.010>
- Tian, W., Liao, Z., Zhang, J., 2017. An optimization of artificial neural network model for predicting chlorophyll dynamics. *Ecol. Modell.* 364, 42–52. <https://doi.org/10.1016/j.ecolmodel.2017.09.013>
- Tian, Z., 2012. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *J. Intell. Manuf.* 23, 227–237. <https://doi.org/10.1007/s10845-009-0356-9>
- Tkáč, M., Verner, R., 2016. Artificial neural networks in business: Two decades of research. *Appl. Soft Comput. J.* 38, 788–804. <https://doi.org/10.1016/j.asoc.2015.09.040>
- Tokumaru, M., Muranaka, N., Imanishi, S., 2002. Color design support system considering color harmony, in: 2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. pp. 378–383. <https://doi.org/10.1109/FUZZ.2002.1005020>
- Tomich, T.P., Lidder, P., Coley, M., Gollin, D., Meinen-Dick, R., Webb, P., Carberry, P., 2019. Food and agricultural innovation pathways for prosperity. *Agric. Syst.* 172, 1–15. <https://doi.org/10.1016/j.agsy.2018.01.002>

- Torrent, J., Barrón, V., Liu, Q., 2006. Magnetic enhancement is linked to and precedes hematite formation in aerobic soil. *Geophys. Res. Lett.* 33, 4–7. <https://doi.org/10.1029/2005GL024818>
- Torrent, J., Schwertmann, U., Schulze, D.G., 1980. Iron oxide mineralogy of some soils of two river terrace sequences in Spain. *Geoderma* 23, 191–208. [https://doi.org/10.1016/0016-7061\(80\)90002-6](https://doi.org/10.1016/0016-7061(80)90002-6)
- Tran, H.N., Huynh, T., Do, T., 2014. Author name disambiguation by using deep neural network, in: *Asian Conference on Intelligent Information and Database Systems*. pp. 123–132.
- Tran, T.A., Nguyen, T.H., Vo, T.T., 2019. Adaptation to flood and salinity environments in the Vietnamese Mekong Delta: Empirical analysis of farmer-led innovations. *Agric. Water Manag.* 216, 89–97. <https://doi.org/10.1016/j.agwat.2019.01.020>
- Trejo-Perea, M., Herrera-Ruiz, G., Rios-Moreno, J., Miranda, R.C., Rivas-Araiza, E., 2009. Greenhouse energy consumption prediction using neural networks models. *Int. J. Agric. Biol.* 11, 1–6.
- Trinderup, C.H., Dahl, A., Jensen, K., Carstensen, J.M., Conradsen, K., 2015. Comparison of a multispectral vision system and a colorimeter for the assessment of meat color. *Meat Sci.* 102, 1–7. <https://doi.org/10.1016/j.meatsci.2014.11.012>
- Tripathy, A.K., Adinarayana, J., Sudharsan, D., Merchant, S.N., Desai, U.B., Vijayalakshmi, K., Raji Reddy, D., Sreenivas, G., Ninomiya, S., Hirafuji, M., Kiura, T., Tanaka, K., 2011. Data mining and wireless sensor network for agriculture pest/disease predictions. *Proc. 2011 World Congr. Inf. Commun. Technol. WICT 2011* 1229–1234. <https://doi.org/10.1109/WICT.2011.6141424>
- Tripepi, G., Jager, K.J., Dekker, F.W., Zoccali, C., 2008. Linear and logistic regression analysis. *Kidney Int.* 73, 806–810. <https://doi.org/10.1038/sj.ki.5002787>
- Tsai, K.M., Wang, P.J., 2001. Predictions on surface finish in electrical discharge machining based upon neural network models. *Int. J. Mach. Tools Manuf.* 41, 1385–1403. [https://doi.org/10.1016/S0890-6955\(01\)00028-1](https://doi.org/10.1016/S0890-6955(01)00028-1)
- Tubiello, F.N., Salvatore, M., Córdor Golec, R.D., Ferrara, A., Rossi, S., Biancalani, R., Federici, S., Jacobs, H., Flammini, A., 2014. *Agriculture, Forestry and Other Land Use Emissions by Sources and Removals by Sinks*, Statistics Division, Food and Agriculture Organization, Rome. <https://doi.org/10.13140/2.1.4143.4245>
- Turan, N.G., Mesci, B., Ozgonenel, O., 2011. The use of artificial neural networks (ANN) for modeling of adsorption of Cu(II) from industrial leachate by pumice. *Chem. Eng. J.* 171, 1091–1097. <https://doi.org/10.1016/j.cej.2011.05.005>
- Valeeva, A.A., Aleksandrova, A.B., Kuposov, G.F., 2016. Color estimation of forest-steppe soils by digital photography under laboratory conditions. *Eurasian Soil Sci.* 49, 1033–1037. <https://doi.org/10.1134/S1064229316090131>
- Valipour, M., 2016. How Much Meteorological Information Is Necessary to Achieve Reliable Accuracy for Rainfall Estimations? *Agriculture* 6, 53. <https://doi.org/10.3390/agriculture6040053>

- Valous, N.A., Mendoza, F., Sun, D.W., Allen, P., 2009. Colour calibration of a laboratory computer vision system for quality evaluation of pre-sliced hams. *Meat Sci.* 81, 132–141. <https://doi.org/10.1016/j.meatsci.2008.07.009>
- Vanhoucke, V., Senior, A., Mao, M.Z., 2011. Improving the speed of neural networks on CPUs, in: *Proc Deep Learning and* pp. 611–620. <https://doi.org/10.1186/1472-6750-10-60>
- Vázquez-Canteli, J.R., Ulyanin, S., Kämpf, J., Nagy, Z., 2019. Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities. *Sustain. Cities Soc.* 45, 243–257. <https://doi.org/10.1016/j.scs.2018.11.021>
- Veerachary, M., Yadaiah, N., 2000. ANN based peak power tracking for PV supplied dc motors. *Sol. Energy* 69, 343–350. [https://doi.org/10.1016/S0038-092X\(00\)00085-2](https://doi.org/10.1016/S0038-092X(00)00085-2)
- Virto, I., Imaz, M.J., Fernández-Ugalde, O., Gartzia-Bengoetxea, N., Enrique, A., Bescansa, P., 2015. Soil degradation and soil quality in Western Europe: Current situation and future perspectives. *Sustain.* 7, 313–365. <https://doi.org/10.3390/su7010313>
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* 150, 253–266. <https://doi.org/10.1016/j.geoderma.2009.01.025>
- Viscarra Rossel, R.A., Minasny, B., Roudier, P., McBratney, A.B., 2006. Colour space models for soil science. *Geoderma* 133, 320–337. <https://doi.org/10.1016/j.geoderma.2005.07.017>
- Vodyanitskii, Y.N., Kirillova, N.P., 2016. Application of the CIE-L*a*b* system to characterize soil color. *Eurasian Soil Sci.* 49, 1259–1268. <https://doi.org/10.1134/S1064229316110107>
- Vodyanitskii, Y.N., Savichev, A.T., 2017. The influence of organic matter on soil color using the regression equations of optical parameters in the system CIE- L*a*b*. *Ann. Agrar. Sci.* 15, 380–385. <https://doi.org/10.1016/j.aasci.2017.05.023>
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* 105, 569–582. <https://doi.org/10.1016/j.renene.2016.12.095>
- Waggoner, P., 2016. *Agricultural Meteorology, Meteorological Monographs.* American Meteorological Society.
- Wagner, P.D., Fiener, P., Wilken, F., Kumar, S., Schneider, K., 2012. Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *J. Hydrol.* 464–465, 388–400. <https://doi.org/10.1016/j.jhydrol.2012.07.026>
- Walker, R.F., Kumagai, M., 2000. Image analysis as a tool for quantitative phycology: a computational approach to cyanobacterial taxa identification. *Limnology* 1, 107–115. <https://doi.org/10.1007/s102010070016>
- Wang, W., Gelder, P.H.A.J.M. Van, Vrijling, J.K., Ma, J., 2006. Forecasting daily streamflow using hybrid ANN models. *J. Hydrol.* 324, 383–399. <https://doi.org/10.1016/j.jhydrol.2005.09.032>

- Wang, W.C., Chau, K.W., Cheng, C.T., Qiu, L., 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol.* 374, 294–306. <https://doi.org/10.1016/j.jhydrol.2009.06.019>
- Webster, P.J., 2013. Improve weather forecasts for the developing world. *Nature* 493, 17–19.
- Wei, H., Yang, C., Yu, Q., 2017. Contour segment grouping for object detection. *J. Vis. Commun. Image Represent.* 48, 292–309. <https://doi.org/10.1016/j.jvcir.2017.07.003>
- Wei, T., Cherry, T.L., Glomrød, S., Zhang, T., 2014. Climate change impacts on crop yield: Evidence from China. *Sci. Total Environ.* 499, 133–140. <https://doi.org/10.1016/j.scitotenv.2014.08.035>
- Wen, X., Du, K., Wang, Z., Peng, X., Luo, L., Tao, H., Xu, Y., Zhang, D., Geng, Y., Li, Y., 2016. Effective cultivation of microalgae for biofuel production: A pilot-scale evaluation of a novel oleaginous microalga *Graesiella* sp. WBG-1. *Biotechnol. Biofuels* 9, 1–12. <https://doi.org/10.1186/s13068-016-0541-y>
- Westermann, O., Förch, W., Thornton, P., Körner, J., Cramer, L., Campbell, B., 2018. Scaling up agricultural interventions: Case studies of climate-smart agriculture. *Agric. Syst.* 165, 283–293. <https://doi.org/10.1016/j.agsy.2018.07.007>
- Williams, I., Bowring, N., Svoboda, D., 2014. A performance evaluation of statistical tests for edge detection in textured images. *Comput. Vis. Image Underst.* 122, 115–130. <https://doi.org/10.1016/j.cviu.2014.02.009>
- Willmott, C., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* 30, 79–82.
- Wolff, M.W., Hopmans, J.W., Stockert, C.M., Burger, M., Sanden, B.L., Smart, D.R., 2017. Effects of drip fertigation frequency and N-source on soil N₂O production in almonds. *Agric. Ecosyst. Environ.* 238, 67–77. <https://doi.org/10.1016/j.agee.2016.08.001>
- Wu, H., Yuan, Z., Geng, Y., Ren, J., Jiang, S., Sheng, H., Gao, L., 2017. Temporal trends and spatial patterns of energy use efficiency and greenhouse gas emissions in crop production of Anhui Province, China. *Energy* 133, 955–968. <https://doi.org/10.1016/j.energy.2017.05.173>
- Wu, T., Li, Y., 2013. Spatial interpolation of temperature in the United States using residual kriging. *Appl. Geogr.* 44, 112–120. <https://doi.org/10.1016/j.apgeog.2013.07.012>
- Xia, Z., Wu, D., Nie, P., He, Y., 2016. Non-invasive measurement of soluble solid content and pH in Kyoho grapes using a computer vision technique. *Anal. Methods* 8, 3242–3248. <https://doi.org/10.1039/C5AY02694F>
- Xie, Z., Liu, Y., Hu, C., Chen, L., Li, D., 2007. Relationships between the biomass of algal crusts in fields and their compressive strength. *Soil Biol. Biochem.* 39, 567–572.

- Yam, K.L., Papadakis, S.E., 2004. A simple digital imaging method for measuring and analyzing color of food surfaces. *J. Food Eng.* 61, 137–142. [https://doi.org/10.1016/S0260-8774\(03\)00195-X](https://doi.org/10.1016/S0260-8774(03)00195-X)
- Yang, C.T., Huang, C.L., Lin, C.F., 2011. Hybrid CUDA, OpenMP, and MPI parallel programming on multicore GPU clusters. *Comput. Phys. Commun.* 182, 266–269. <https://doi.org/10.1016/j.cpc.2010.06.035>
- Yang, C.T., Marsooli, R., Aalami, M.T., 2009. Evaluation of total load sediment transport formulas using ANN. *Int. J. Sediment Res.* 24, 274–286. [https://doi.org/10.1016/S1001-6279\(10\)60003-0](https://doi.org/10.1016/S1001-6279(10)60003-0)
- Ying-Jin, C., Davis, S., Chao-Kun, C., Xue, B., 2004. A study of sample size with neural network. *Mach. Learn. Cybern. 2004. Proc. 2004 Int. Conf.* 6, 3444–3448 vol.6. <https://doi.org/10.1109/ICMLC.2004.1380382>
- Yonaba, H., Antcil, F., Fortin, V., 2010. Comparing Sigmoid Transfer Functions for Neural Network Multistep Ahead Streamflow Forecasting. *J. Hydrol. Eng.* 15, 275–283. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000188](https://doi.org/10.1061/(asce)he.1943-5584.0000188)
- Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks?, in: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q. (Eds.), *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc., pp. 3320–3328.
- Zagoruyko, S., Komodakis, N., 2015. Learning to compare image patches via convolutional neural networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4353–4361. <https://doi.org/10.2320/matertrans.maw200919>
- Zamanlooy, B., Mirhassani, M., 2014. Efficient VLSI implementation of neural networks with hyperbolic tangent activation function. *IEEE Trans. Very Large Scale Integr. Syst.* 22, 39–48. <https://doi.org/10.1109/TVLSI.2012.2232321>
- Zanetti, S., Sousa, E., Oliveira, V., Almeida, F., Bernardo, S., 2007. Estimating evapotranspiration using artificial neural network and minimum climatological data. *J. Irrig. Drain. Eng.* 133, 83–89. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2007\)133](https://doi.org/10.1061/(ASCE)0733-9437(2007)133)
- Zermeño-González, A., Cárdenas-Palomo, J.O., Ramírez-Rodríguez, H., Benavides-Mendoza, A., Cadena-Zapata, M., Campos-Magaña, S.G., 2015. Fertilización biológica del cultivo de maíz. *Rev. Mex. Ciencias Agrícolas* 2399–2408.
- Zhang, C., Liu, S., Wu, S., Jin, S., Reis, S., Liu, H., Gu, B., 2019. Rebuilding the linkage between livestock and cropland to mitigate agricultural pollution in China. *Resour. Conserv. Recycl.* 144, 65–73. <https://doi.org/10.1016/j.resconrec.2019.01.011>
- Zhang, C., Woodland, P.C., 2016. DNN speaker adaptation using parameterised sigmoid and ReLU hidden activation functions. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* 2016-May, 5300–5304. <https://doi.org/10.1109/ICASSP.2016.7472689>

- Zhang, G., Hu, M.Y., Patuwo, B.E., Indro, D.C., 1999. Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis. *Eur. J. Oper. Res.* 116, 16–32. [https://doi.org/10.1016/S0377-2217\(98\)00051-4](https://doi.org/10.1016/S0377-2217(98)00051-4)
- Zhang, J., Martin, E.B., Morris, A.J., Kiparissides, C., 2003. Inferential estimation of polymer quality using stacked neural networks. *Comput. Chem. Eng.* 21, S1025–S1030. [https://doi.org/10.1016/S0098-1354\(97\)87637-5](https://doi.org/10.1016/S0098-1354(97)87637-5)
- Zhang, J., Morris, A.J., 1998. A sequential learning approach for single hidden layer neural networks. *Neural Networks* 11, 65–80. [https://doi.org/10.1016/S0893-6080\(97\)00111-1](https://doi.org/10.1016/S0893-6080(97)00111-1)
- Zhang, L., Suganthan, P.N., 2016. A comprehensive evaluation of random vector functional link networks. *Inf. Sci. (Ny)*. 367–368, 1094–1105. <https://doi.org/10.1016/j.ins.2015.09.025>
- Zhang, X., Davidson, E.A., Mauzerall, D.L., Searchinger, T.D., Dumas, P., Shen, Y., 2015. Managing nitrogen for sustainable development. *Nature* 528, 51–59. <https://doi.org/10.1038/nature15743>
- Zhang, Y.C., Kagen, A.C., 2017. Machine Learning Interface for Medical Image Analysis. *J. Digit. Imaging* 30, 615–621. <https://doi.org/10.1007/s10278-016-9910-0>
- Zhang, Z., Trivedi, C., Liu, X., 2017. Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data. *Saf. Sci.* 0–1. <https://doi.org/10.1016/j.ssci.2017.11.023>
- Zhang, Z., Wang, B., Hu, Q., Sommerfeld, M., Li, Y., Han, D., 2016. A new paradigm for producing astaxanthin from the unicellular green alga *Haematococcus pluvialis*. *Biotechnol. Bioeng.* 113, 2088–2099. <https://doi.org/10.1002/bit.25976>
- Zhao, D., Ni, W., Zhu, Q., 2014. A framework of neural networks based consensus control for multiple robotic manipulators. *Neurocomputing* 140, 8–18. <https://doi.org/10.1016/j.neucom.2014.03.041>
- Zhao, G., Di Girolamo, L., Diner, D.J., Bruegge, C.J., Mueller, K.J., Wu, D.L., 2016. Regional Changes in Earth's Color and Texture as Observed From Space Over a 15-Year Period. *IEEE Trans. Geosci. Remote Sens.* 54, 4240–4249. <https://doi.org/10.1109/TGRS.2016.2538723>
- Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., Torr, P.H.S., 2015. Conditional random fields as recurrent neural networks, in: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 1529–1537. <https://doi.org/10.1109/ICCV.2015.179>
- Zion, B., 2012. The use of computer vision technologies in aquaculture - A review. *Comput. Electron. Agric.* 88, 125–132. <https://doi.org/10.1016/j.compag.2012.07.010>
- Zurada, J.M., 1992. *Introduction to artificial neural systems*. West publishing company St. Paul.

APPENDIX

Appendix A

Journal article present as academic merit to present the doctoral thesis

Title: Monoalgal and mixed algal cultures discrimination by using an artificial neural network

Authors: B.M. Franco, L.M. Navas, C. Gómez, C. Sepúlveda, F.G. Acién

Corresponding author: B.M. Franco

Journal: Algal Research

Editor: Elsevier

ISSN: 211-9264

Date of publication: March 2019

Volume: 38

DOI: <https://doi.org/10.1016/j.algal.2019.101419>

URL: www.sciencedirect.com/science/article/pii/S2211926418306799

Journal impact factor: 3.75 (2017)

Journal ranking:

Q1 in Biotechnology & applied microbiology (38/161)

Appendix B

DigiCIELAB computer software intellectual property register

REGISTRO CENTRAL DE LA PROPIEDAD INTELECTUAL

 MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE

 GOBIERNO DE ESPAÑA

MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE

REGISTRO GENERAL DE LA PROPIEDAD INTELECTUAL

Según lo dispuesto en la Ley de Propiedad Intelectual (Real Decreto Legislativo 1/1996, de 12 de abril), quedan inscritos en este Registro los derechos de propiedad intelectual en la forma que se determina seguidamente:

NÚMERO DE ASIENTO REGISTRAL 00 / 2018 / 16

Título: DigiCIELAB

Objeto de propiedad intelectual: Programa de ordenador

Clase de obra: Programa de ordenador

PRIMERA INSCRIPCIÓN

Autor/es y titular/es originarios de derechos

- **Apellidos y nombre:** NAVAS GRACIA, Luis Manuel
Nacionalidad: España **D.N.I./N.I.F./Pasaporte:** 50067027Y
- **Apellidos y nombre:** FRANCO ORTELLADO, Blas Manuel
Nacionalidad: Paraguay **D.N.I./N.I.F./Pasaporte:** Y3090374Q

Transmisión de derechos

- **Titular cesionario:** UNIVERSIDAD DE VALLADOLID
Nacionalidad: España **D.N.I./N.I.F./Pasaporte:** Q4718001C

Es titular en exclusiva de los derechos de explotación de esta obra por mediación de relación laboral y estar creada en el ejercicio de las funciones asignadas en la empresa al autor/es o siguiendo las instrucciones de su empresario.

- **Cedente:** FRANCO ORTELLADO, Blas Manuel
Nacionalidad: Paraguay **D.N.I./N.I.F./Pasaporte:** Y3090374Q
- **Cedente:** NAVAS GRACIA, Luis Manuel
Nacionalidad: España **D.N.I./N.I.F./Pasaporte:** 50067027Y

Datos de la solicitud

Núm. solicitud: VA-317-17





(Continúa la inscripción)

00/2018/16

**REGISTRO CENTRAL
DE LA PROPIEDAD INTELECTUAL**

MINISTERIO
DE EDUCACIÓN, CULTURA
Y DEPORTE



Fecha de presentación y efectos: 14/09/2017

Hora: 12:30

En Madrid, a cuatro de enero de dos mil dieciocho



LA REGISTRADORA CENTRAL

Lucía Gutiérrez García

00/2018/16

Appendix C

X-Rite ColorChecker Classic technical specifications

In this appendix, the technical specifications of the color calibration chart used in the experimental phase of the thesis are described. The “X-Rite ColorChecker Classic” is a twenty-four color chart card, as is shown in **Figure 35**.



Figure 35. X-Rite ColorChecker Classic color charts enumerated.

The data below presented (**Table 13**) is the CIELAB measurement of all colorchecker charts provided by the manufacturer. The $L^*a^*b^*$ values are an average of each color chart.

Table 13. X-Rite ColorChecker Classic $L^*a^*b^*$ color information following the enumeration presented in Figure 35.








| Number | Color | CIELAB | | | |
|--------|--------------|---|---------|----------|----------|
| | | L^* | a^* | b^* | |
| 1 | Dark skin |  | 37.9860 | 13.5550 | 14.0590 |
| 2 | Light skin |  | 65.7110 | 18.1300 | 17.8100 |
| 3 | Blue sky |  | 49.9270 | -4.8800 | -21.9250 |
| 4 | Foliage |  | 43.1390 | -13.0950 | 21.9050 |
| 5 | Blue flower |  | 55.1120 | 8.8440 | -25.3990 |
| 6 | Bluish green |  | 70.7190 | -33.3970 | -0.1990 |
| 7 | Orange |  | 62.6610 | 36.0670 | 57.0960 |

Table 13. X-Rite ColorChecker Classic L*a*b* color information following the enumeration presented in Figure 35 (continuation).

| Number | Color | CIELAB | | |
|--------|---------------|---------|----------|----------|
| | | L* | a* | b* |
| 8 | Purplish blue | 40.0200 | 10.4100 | -45.9640 |
| 9 | Moderate red | 51.1240 | 48.2390 | 16.2480 |
| 10 | Purple | 30.3250 | 22.9760 | -21.5870 |
| 11 | Yellow green | 72.5320 | -23.7090 | 57.2550 |
| 12 | Orange yellow | 71.9410 | 19.3630 | 67.8570 |
| 13 | Blue | 28.7780 | 14.1790 | -50.2970 |
| 14 | Green | 55.2610 | -38.3420 | 31.3700 |
| 15 | Red | 42.1010 | 53.3780 | 28.1900 |
| 16 | Yellow | 81.7330 | 4.0390 | 79.8190 |
| 17 | Magenta | 51.9350 | 49.9860 | -14.5740 |
| 18 | Cyan | 51.0380 | -28.6310 | -28.6380 |
| 19 | White | 96.5390 | -0.4250 | 1.1860 |
| 20 | Neutral 8 | 81.2570 | -0.6380 | -0.3350 |
| 21 | Neutral 6.5 | 66.7660 | -0.7340 | -0.5040 |
| 22 | Neutral 5 | 50.8670 | -0.1530 | -0.2700 |
| 23 | Neutral 3.5 | 35.6560 | -0.4210 | -1.2310 |
| 24 | Black | 20.4610 | -0.0790 | -0.9730 |