**Universidad de Valladolid**

**Escuela Técnica Superior de Ingenieros de Telecomunicación**

**Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática**


TESIS DOCTORAL:

# Artificial Neural Networks applied to the resolution of Regression and Classification Multivariate Analysis problems in the Agricultural and the Industrial fields


Presentada por Víctor Martínez Martínez para optar al grado de doctor por la Universidad de Valladolid

Dirigida por:
Jaime Gómez Gil
Javier Manuel Aguiar Pérez

# Abstract

Artificial Neural Networks (ANN) are processing models inspired in the human brain that have been widely employed to solve regression and classification problems because of the input-output mapping capability of the ANNs. Agriculture and industry are two examples of fields in which there are many processes affected by multiple variables that need to be modelled, and the input-output mapping capability of the ANNs makes a good alternative to be used to build these models.

The main objective of this thesis is to design, implement, and evaluate specific ANN-based models to be applied to agricultural and industrial applications. These models will be designed considering a set of variables larger or different than the previously considered in the scientific literature, minimizing the previous knowledge about the application needed to design them, and optimizing its performance in terms of processing time and computation needed.

The methodology employed to achieve the objectives proposed in this thesis is composed of five stages: literature review, hypothesis formulation, developing and evaluation, result analysis, and result dissemination. The literature review was considered to learn about the ANN-based modelling techniques implemented and about the application to be modelled. After the literature review, a research hypothesis was formulated to be used as the basis of the research done on each application. Then, a scenario to develop and test the proposed ANN-based models was proposed and implemented in the developing and evaluation stage. Results obtained in the testing stage were compared with the obtained by other methods proposed in the scientific literature in the result analysis stage. This comparison determined if the research hypothesis was valid or if it needed to be redefined or rejected. Finally, if the hypothesis needed to be redefined or rejected, the methodology went back to the hypothesis formulation stage, and if the hypothesis was valid, the obtained results were disseminated.

The proposed methodology was applied to five agricultural and industrial applications with different types of regression and classification problems to be modelled: the tobacco drying process, the switchgrass (*Panicum virgatum*) drying process, the predictive maintenance of a machine, the evaluation of steel pieces in a production line, and the early detection of plant diseases. An ANN to model the tobacco drying process was proposed to estimate and predict the temperature and relative humidity inside the tobacco mass using temperature and relative humidity data from measurement points placed outside the tobacco mass (first article of the compendium). The switchgrass drying process was modelled to predict the future value of the moisture content of the switchgrass when it was being dried in the land with variable weather conditions (second article of the compendium). A vibration model of a machine was performed to be used in a predictive maintenance application, estimating the state of the rotary components of the monitored machine as a function of a vibration signal acquired in a single point of the machine structure (third article of the compendium). A non-destructive testing (NDT) model was proposed to evaluate steel pieces as a function of the eddy-current induced impedances measured applying the eddy-current at different frequencies (fourth article of the compendium). The early detection of plant diseases application was done by means of a reflectance-based disease model, which estimated the severity of a disease of a plant as a function of reflectance data acquired in previous stages of the plant development (fifth article of the compendium).

The experiments performed with each one of the five models developed in this thesis showed findings that can be grouped into three categories. The first one is that most of the models proposed were developed with a set of variables larger or different than those commonly considered in the scientific literature. The second group of findings is that the proposed models were developed avoiding previous knowledge about the process modelled. This characteristic made the proposed models more general and easier to be adapted to other similar processes compared to other models of scientific literature. The last group of results focuses on the optimization of the model, minimizing the processing capability or the training time needed to adjust or execute the proposed method. This characteristic of the developed models allows them to be used in real applications with requirements of low-latency or low computational load requirements.

The results obtained in this thesis suggest the capability of ANN-based models to solve regression or classifications problems of the agricultural and industrial fields. Three main conclusions can be extracted from these results. The first one is that models proposed in scientific literature can be improved considering a set of variables larger or different than the proposed by other authors: for example, the tobacco drying model considers a measurement point inside the product to be dried, the switchgrass drying model employs information about the rain events, the vibration model of a machine uses information from a measurement point placed far away from the rotary components considered, and the NDT model employs impedance data at different frequencies. The second conclusion is that the generalization capability of the ANNs allows researchers to develop ANN-based models that do not include previous knowledge about the application to be modelled: for example, the tobacco drying process model does not consider information about the control algorithm, both the tobacco and the switchgrass drying models do not employ any information about the product to be dried, the vibration model of a machine do not used any information about the rotary components monitored, and the models for the NDT of steel pieces and for the early detection of plant diseases can be implemented with acquisition systems with other characteristics than the used in the experiments of this thesis. The third conclusion is that the processing time and the computation needed depends on the way the ANN is configured and optimized, and because of that the models proposed in this thesis were optimized as follows: the tobacco and the switchgrass drying models and the early detection of plant diseases model were designed to be used in near real-time applications employing features with low processing requirements as inputs of these models, the vibration model of a machine and the NDT model do not implemented a feature extraction stage so they can be employed in real-time applications, and the vibration model of a machine uses a genetic algorithm to improve the performance of the model and to reduce the time needed to train the ANN.

**Keywords**

Artificial Neural Network (ANN), modelling, classification, regression, agricultural, industrial, multivariate analysis.

# Resumen

Las Redes Neuronales Artificiales (RNAs) son modelos de procesado inspirados en el cerebro humano que han sido ampliamente empleadas para resolver problemas de regresión y clasificación gracias a la capacidad de mapeo entrada-salida de las RNAs. Agricultura e industria son dos ejemplos de campos en los que hay muchos procesos afectados por múltiples variables que necesitan ser modelados, con lo que la capacidad de mapeo entrada-salida de las RNAs hace que éstas sean una buena alternativa para construir estos modelos.

El principal objetivo de esta tesis es diseñar, implementar y evaluar modelos específicos basados en RNAs que se utilizarán en aplicaciones agrícolas e industriales. Estos modelos se diseñarán considerando un conjunto de variables mayor o diferente a los que se han utilizado previamente en la literatura científica, minimizando el conocimiento previo sobre la aplicación para diseñar el modelo, y optimizando su desempeño en términos de tiempo de procesado y de capacidad de cómputo necesario.

La metodología empleada para alcanzar los objetivos propuestos en esta tesis se componen de cinco etapas: revisión bibliográfica, formulación de hipótesis, desarrollo y evaluación, análisis de los resultados y divulgación de los resultados. La revisión bibliográfica se realizó para aprender sobre las técnicas de modelado basadas en RNAs y sobre las aplicaciones que se han modelado durante esta tesis. Después de la revisión bibliográfica se formuló la hipótesis de investigación, que se empleó como punto de partida para la investigación realizada en cada aplicación agro-industrial. Posteriormente, se propuso e implementó un escenario para desarrollar y evaluar el modelo basado en RNAs. Los resultados obtenidos en la fase de desarrollo y evaluación se compararon en la fase de análisis de resultados con los resultados obtenidos por otros métodos propuestos en la literatura científica. Esta comparación determinó si la hipótesis de investigación fue válida o si necesitaba ser definida o rechazada. Por último, si la hipótesis necesitaba ser redefinida o rechazada, la metodología volvía a la etapa de formulación de hipótesis, mientras que si la hipótesis era válida los resultados obtenidos se divulgaron.

La metodología propuesta se implementó en cinco aplicaciones agrícolas e industriales con diferentes tipos de problemas de regresión y clasificación a ser modelados: el proceso de curado de tabaco, el proceso de secado del pasto varilla (*Panicum virgatum*), el mantenimiento predictivo de una máquina, la evaluación de piezas de acero en una línea de producción y la detección temprana de enfermedades en plantas. Una RNA para el modelado del proceso de curado de tabaco se propuso para estimar y predecir la temperatura y la humedad relativa dentro de la masa de tabaco utilizando datos de temperatura y humedad relativa adquiridos en puntos de medida situados fuera de la masa de tabaco (primer artículo del compendio). El proceso de secado del pasto varilla se modeló para predecir el valor futuro del contenido de humedad del pasto varilla cuando ésta se seca en el exterior con condiciones meteorológicas variables (segundo artículo del compendio). Un modelo de las vibraciones de una máquina se desarrolló para emplearse en una aplicación de mantenimiento predictivo, estimando el estado de componentes rotantes de la máquina monitorizada en función de la señal de vibración adquirida en un único punto de la estructura de la máquina (tercer artículo del compendio). Un modelo para ensayos no destructivos (END) se propuso para evaluar piezas de acero en función de la impedancia inducida por corrientes de Foucault (*eddy-current*) en la pieza, empleando corrientes de Foucault a diferentes frecuencias (cuarto artículo del compendio). La aplicación de detección temprana de enfermedades en plantas se abordó mediante un modelo de la enfermedad basado en la reflectancia, que estimaba la severidad de una planta en función de los datos de reflectancia adquiridos en etapas tempranas del desarrollo de la planta (quinto artículo del compendio).

Los experimentos realizados con cada uno de los cinco modelos desarrollados en esta tesis mostraron resultados que se pueden agrupar en tres categorías. La primera es que la mayor parte de los modelos propuestos se desarrollaron con un conjunto de variables mayor o diferente a los comúnmente propuestos previamente en la literatura científica. El segundo grupo de hallazgos se deriva de que los modelos propuestos se desarrollaron evitando el conocimiento previo sobre los procesos modelados. Esta característica hace que los modelos sean más generales y sea más sencilla su adaptación a otras aplicaciones similares, en comparación con otros modelos de la literatura científica. El último grupo de resultados se centra en la optimización de los modelos, minimizando la capacidad de proceso y el tiempo necesarios para ajustar o ejecutar el método propuesto. Esta característica de los modelos desarrollados permite que se empleen en aplicaciones en tiempo real con requerimientos de baja latencia o baja carga computacional.

Los resultados obtenidos en esta tesis sugieren que los modelos basados en RNAs son capaces de resolver problemas de regresión y clasificación del ámbito agrícola e industrial. Tres conclusiones se pueden extraer de estos resultados. La primera es que los modelos propuestos en la literatura científica se pueden mejorar empleando conjuntos de variables mayores o diferentes a los propuestos por otros autores: por ejemplo, el modelo de curado de tabaco emplea un punto de medida dentro del producto a secar, el modelo de secado del pasto varilla emplea información sobre los eventos de lluvia, el modelo de vibración de una máquina utiliza información adquirida en un punto de medida que está alejado de los elementos rotantes considerados, y el modelo de END emplea datos de impedancia a diferentes frecuencias. La segunda conclusión es que la capacidad de generalización de las RNAs permite a los investigadores desarrollar modelos basados en RNAs que no incluyen conocimiento previo sobre la aplicación a modelar: por ejemplo, el modelo de curado de tabaco no emplea información sobre el algoritmo de control de curado utilizado, tanto el modelo de curado de tabaco como el modelo de secado de pasto varilla no emplean información sobre el producto a secar, el modelo de vibraciones de una máquina no utiliza información sobre los elementos rotantes monitorizados y los modelos de END de piezas de acero y de detección temprana de enfermedades en plantas se pueden implementar con sistemas de adquisición con especificaciones diferentes a los utilizados en los experimentos de esta tesis. La tercera conclusión es que el tiempo de procesado y el procesamiento necesario depende del modo en el que la RNA se configura y optimiza, y por ello los modelos propuestos en esta tesis se optimizaron como se muestra a continuación: los modelos de curado de tabaco, de secado de pasto varilla y de detección temprana de enfermedades en plantas se diseñaron para su utilización en aplicaciones de cuasi tiempo real, empleando características que requerían de poco procesado como entrada de estos modelos; los modelos de vibración de una máquina y de END no implementaron una etapa de extracción de características, lo que permite su empleo en aplicaciones de tiempo real; y el modelo de vibraciones de una máquina utilizó un algoritmo genético para mejorar la precisión del modelo y para reducir el tiempo necesario para entrenar la RNA.

**Palabras clave:**

Red Neuronal Artificial (RNA), modelado, clasificación, regresión, agrícola, industrial, análisis multivariado.

# Acknowledgements/Agradecimientos

Esta tesis es el resultado de varios años de trabajo en el Departamento de Teoría de la Señal, Comunicaciones e Ingeniería Telemática de la Universidad de Valladolid. Este trabajo no habría sido posible sin la ayuda y el apoyo de muchas personas, a las cuales me gustaría agradecer su apoyo y ayuda durante todo este proceso.

En primer lugar quiero agradecer a mi familia, especialmente a mis padres Alberto y María del Carmen y a mi hermana Lidia. Sin su ayuda durante todos estos años y su apoyo en los malos momentos no habría sido capaz de llegar a buen fin este trabajo.

En segundo lugar quiero agradecer a mis tutores, Jaime y Javier, por toda su ayuda durante todos estos años. Al principio, por confiar en mí para comenzar mi carrera investigadora sin tener experiencia. Y durante todo este proceso, en el que vuestra experiencia y conocimientos me han servido de guía y han hecho que haya podido llegar al final del camino de manera satisfactoria.

También tengo que agradecer a las personas con las que he trabajado y compartido el día a día en el laboratorio. En primer lugar a Gonzalo y Jorge, con los que he coincidido durante todo este camino y con los que he trabajado y de los que he aprendido mucho. En segundo lugar a Ernesto, con el que coincidí al iniciar mi doctorado y compartí el principio de este camino, que me ayudó con su experiencia previa a afrontar algunos de los problemas que me fui encontrando en el camino. En tercer lugar a Rubén, con el que también he compartido horas de trabajo y discusión muy productivas. Y en cuarto lugar a toda la gente que ha estado por el laboratorio y con los que hemos compartido horas de trabajo de una y otra manera. Destacar en este grupo a Carlos, al que conozco desde mi primer año en Valladolid y con el que he compartido también parte del trabajo de mi doctorado con su Proyecto Fin de Carrera.

I want also to express my gratitude for the opportunity to do three research stays in two international research groups. First, I have to thank the Biosystems and Agriculture Department of the University of Kentucky (USA) and Tim Stombaugh for the opportunity to do two stays to work with his research group. It was very profitable because of the chance to work with a multidisciplinary group, to use their facilities to perform experiments, and to improve my English. Second, I have to thank the Departamento de Engenharia Agrícola of the Universidade Federal de Viçosa (Brazil) for their support, allowing me to research, to teach their students, and to learn a little bit of Portuguese. It was also a very good opportunity for me to get to know two new cultures and to meet a lot of people. Thank you to all of you: Francisco, Tim, Evandro, Michael, Sarvio, the people from BCM, and all the other people I met during my stays.

También quiero agradecer a todas las personas con las que he colaborado durante estos años en mi investigación, y que también son parte importante del trabajo realizado: Luis Manuel, Carlos, Francisco Javier, Javier y José Carlos. Y a todos ellos con los que he compartido esta experiencia de una u otra manera.

Por último, no quiero olvidarme de toda la gente con la que he compartido estos años de mi vida, a toda la gente que he conocido y a todos esos amigos que he tenido y con los que he compartido esta etapa de mi vida y con las que espero seguir contando en el futuro.

Muchas gracias a todos

Thank you

Obrigado

# Contents

**Abstract** ..... **i**

**Resumen** ..... **v**

**Acknowledgements/Agradecimientos** ..... **ix**

**Contents** ..... **xi**

**List of Figures** ..... **xiii**

**I Introduction and Summary** ..... **1**

1. Introduction ..... 1
2. Artificial Neural Networks overview ..... 2
   - 2.1. Neurons ..... 2
   - 2.2. Artificial Neural Networks ..... 4
3. Objectives ..... 5
4. Methodology ..... 6
5. Results ..... 7
   - 5.1. Agricultural and industrial applications modelled ..... 8
   - 5.2. Results obtained ..... 13
6. Original contributions to the state of the art ..... 16
   - 6.1. First original contribution to the state of the art ..... 16
   - 6.2. Second original contribution to the state of the art ..... 17
7. Conclusions ..... 18
8. Future work ..... 20
9. Merits and diffusion of the results ..... 22
   - 9.1. Publications ..... 22
   - 9.2. Projects ..... 24
   - 9.3. Research stays ..... 25
   - 9.4. Teaching ..... 26
   - 9.5. Other merits ..... 26

**II Articles** ..... **29**

Article 1: Tobacco drying process ..... 31

Article 2: Switchgrass drying process ..... 51

Article 3: Predictive maintenance of a machine ........ 71
Article 4: Evaluation of steel pieces in a production line ........ 83
Article 5: Early detection of plant diseases ........ 97
**Bibliography ........ 115**

# List of Figures


**Figure 1:** Block diagram of a perceptron neuron, where it can be seen the input values ($x_n$), the weights ($w_n$) and bias ($b$) values, the activation function ($\varphi$), and the output value ($y$)....3

**Figure 2:** Block diagram of a radial-basis function neuron, where it can be seen the input values ($x_n$), the weights ($w_n$) and bias ($b$) values, the radial basis function (*radbas*), and the output value ($y$)....3

**Figure 3:** Schematic representation of a generic *totally connected feedforward* ANN. This ANN has four input neurons, a single hidden layer with five neurons and a single output neuron....4

**Figure 4:** Block diagram of the methodology stages of the thesis and the connections among them....6

**Figure 5:** Tobacco drying curves graph, in which the target temperature and relative humidity along the time are represented. This curve was extracted from the book of Hawks (Hawks et al., 1986)....9

**Figure 6:** Image of the switchgrass harvest done to acquire the data to develop the ANN-based model. It can be seen the switchgrass before and after harvesting, a clipper to cut the switchgrass, and a box to place the switchgrass in the dryer. ....10

**Figure 7:** Acquisition system employed to acquire the data needed to implement and test the vibration model of a machine. It can be seen an accelerometer on the structure of the machine, an acquisition system connected to this accelerometer by means of a wire, and a laptop to acquire and visualize the data....11

**Figure 8:** Schema of the non-destructive testing method employed in this thesis, were an inductive coil was approached to a steel piece to calculate the impedance induced over the steel piece. ....12

**Figure 9:** Acquisition system employed to acquire the reflectance data from the bean crops, where the reflectometer is placed in the top of the structure in order to acquire the canopy data and store it in the computer....13

# I Introduction and Summary

The first part of this document contains the introduction and summary of this thesis as a compendium of publications. The first section presents a general overview and the research context of this PhD thesis. The second briefly introduces the Artificial Neural Networks as the main processing technique employed in this thesis. The third describes the objectives defined. The fourth presents the methodology employed to reach these objectives. The fifth shows the results obtained during the thesis. The sixth explains the original contributions obtained. The seventh presents the main conclusions of the work performed. The eighth suggests some future work to do after this thesis. Finally, the ninth briefly presents all the activities done during this thesis and the merits obtained.

## 1. Introduction

The advance of Information and Communication Technologies (ICTs) and the growing importance of Big Data and data mining techniques is a fact that is reflected in several fields: economics (Allen et al., 2012), medicine (Esfandiari et al., 2014), industry (Köksal et al., 2011), sports (Lewis, 2004; Oliver, 2004), and agriculture (Mucherino et al., 2009; Liao et al., 2012). Agriculture and industry are two important fields where data analysis techniques can be applied because of the large number of variables that are involved in processes of these fields. Classic models have not usually considered all these variables because of the difficulty in including multivariate a large number of variables on these models.

Classic models have been proposed in literature based on the “observation and experimentation” approach, designing models that describe the physical relationship between the dependent and independent variables that characterize the problem considered. Data analysis techniques, conversely, aim to solve this problem from another perspective: looking for mathematical and logical relationships among the variables involved in the model instead of physical relationships. The main advantages of changing from a physical perspective to a mathematical or logical perspective is that the second perspective (i) allows researchers to add variables to the model without a big increase of the model complexity; (ii) can easily consider variables of any type, such as boolean, discrete, categorical, or real among others; and (iii) does not require much knowledge about the process considered, because the “intelligence” is provided by the data analysis technique instead of by the researcher that designs the model. The main disadvantage of these techniques is that not taking into account the physical relations among the variables makes it more difficult to extract conclusions and knowledge from the developed model.

There are two classic problems in the scientific literature relating dependent and independent variables: regression problems (Draper et al., 1966), where the dependent variable is a numeric variable, and classification problems (Duda et al., 2001), where the dependent variable is a categorical variable. Classification problems can be solved by employing the methods proposed to solve regression methods by means of giving a numeric value to each category considered in the dependent variable. Nevertheless, these problems use to be solved employing different techniques that allow to obtain better results than the obtained solving them with methods designed for regression problems. One the one hand, some examples of techniques used to solve regression problems are linear regression, polynomial regression, or logistic regression. On the other hand, examples of techniques employed to solve classification problems are linear classifiers, Bayesian classifiers, or *k-means* classifiers.

Artificial Intelligence (AI) is an academic discipline that studies how to give intelligent behavior to machines. AI techniques have been employed in the scientific literature as an alternative to solve regression and classification problems using the mathematical or logical perspective mentioned above. There are many techniques in AI, most of them are inspired by the nature, attempting to imitate its intelligent behavior. For example, Neural Computation is a field of research inspired by the animal nervous system that tries to imitate the way they learn and process information. Another example is Genetic Algorithms (GAs), which are AI optimization techniques based on the gene evolution of a human or an animal population. GAs are characterized by combining the survival of the fittest individual and the diversity of the population. Ant Colony Optimization is a third example of nature-based AI techniques. This is a probabilistic technique proposed to solve the salesman problem, which consists of finding the best path to go from one point to another, based on the way a group of ants seek for a path from their colony to a source of food.

In this context, the issue covered in this thesis is the application of neural computation to solve regression and classification problems in the agricultural and industrial fields.

# 2. Artificial Neural Networks overview

Artificial Neural Networks (ANNs) are processing models inspired by how the brain works, and they are one of the main data analysis techniques in the scientific literature. An ANN is composed by a set of neurons and the connection that join these neurons, and it is characterized by having a good processing power despite of the simplicity of the neurons that compose the ANN (Haykin, 1999).

## 2.1. Neurons

Neurons are the simplest processing units that compound an ANN. One example of a neuron model is the *perceptron*, which was proposed by Rosenblatt on 1958 and it is considered as the neuron model that induced the development of the ANNs (Rosenblatt, 1958). The general structure of a *perceptron* is presented in Figure 1, where it can be seen that each *perceptron* is defined by its weights ($w_n$) and its bias ($b$) values, and its transfer or activation function ($\varphi$).

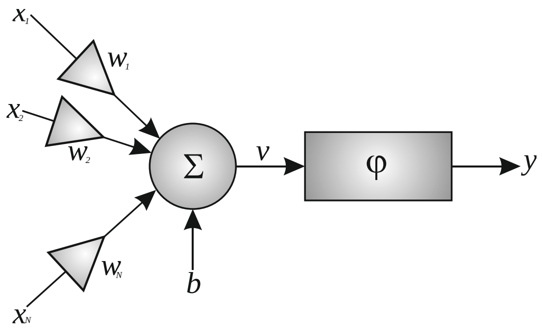



**Figure 1:** Block diagram of a perceptron neuron, where it can be seen the input values ($x_n$), the weights ($w_n$) and bias ($b$) values, the activation function ($\varphi$), and the output value ($y$).

The transfer function of a *perceptron* is shown in equation (1), where the *perceptron* first performs a linear combination of the input values ($x_n$), the weights, and the bias to obtain the $v$ value, and then applies the activation function ($\varphi$) to the result of the linear combination ($v$) to obtain the output value ($y$). The original *perceptron* proposed by Rosenblatt implemented a hard limiter function as the activation function. Nevertheless, nowadays there are a lot of different activation functions proposed to implement *perceptrons*, such as the sigmoid functions, linear functions, or saturating functions.

$$y = \varphi\left(\sum_{n=1}^{N} w_n \cdot x_n + b\right) \tag{1}$$

Another example of a neuron model is the radial-basis function neuron, whose structure is presented in Figure 2. Broomhead and Lowe were the first researchers that employed radial-basis functions in the design on ANNs (Broomhead and Lowe, 1988).

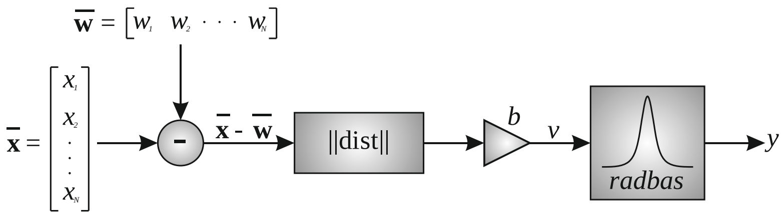



**Figure 2:** Block diagram of a radial-basis function neuron, where it can be seen the input values ($x_n$), the weights ($w_n$) and bias ($b$) values, the radial basis function (*radbas*), and the output value ($y$).

A radial-basis function is a function whose value depends on the distance between the input variables values and a point that characterizes this function, which is called centroid. Some examples of radial-basis functions are the Gaussian function or the multiquadric function. The transfer function of a radial-basis function neuron is presented in equation (2).

$$y = radbas(\|\bar{\boldsymbol{x}} - \bar{\boldsymbol{w}}\| \cdot b) \tag{2}$$

## 2.2. Artificial Neural Networks

The processing capability of a neuron is very limited, and for this reason it is not suitable to solve most of the scientific literature problems. The way the neurons are employed in the scientific literature to solve problems is combining them in ANNs, which give them a very good processing capability.

Neurons of an ANN are usually grouped in layers, having three different types of layers: the input layer, the output layer, and the hidden layers. Every ANN has one input layer, with as many neurons as input variables are in the ANN, and one output layer, with as many neurons as output variables are in the ANN. Moreover, an ANN can have zero or more hidden layers, being the number of hidden layers and the number of neurons of each hidden layer parameters that has to be chosen by the ANN designer. An example of an ANN is presented in Figure 3. Considering the *perceptron* and the radial-basis function neurons presented above, the most common ANNs that use these neurons are the MultiLayer Perceptron (MLP), which is composed by several layers of *perceptron* neurons, and the Radial-Basis Function ANN (RBF-ANN), which is composed by a hidden layer of radial-basis function neurons and an output layer of *perceptron* neurons.

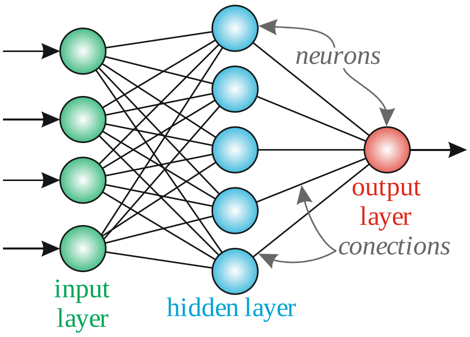



**Figure 3:** Schematic representation of a generic *totally connected feedforward* ANN. This ANN has four input neurons, a single hidden layer with five neurons and a single output neuron.

The connection among the neurons of an ANN can be viewed as connections among the layers of the ANN. This makes possible to classify the ANNs in terms of the connections among its layers: if the connections are only in the direction from the input to the output, the ANN is named *feedforward* ANN; and if there are connections from neurons of one layer to neurons of the same layer or to neurons from the previous layers, the ANN is named *feedback* ANN. Unless *feedback* ANNs are more powerful because of the greater number of connections they can have, *feedforward* ANNs are widely used because they do not have stability problems and they can solve a lot of problems that have been proposed in literature. *Feedforward* ANNs can be also classified as *totally connected* and *partially connected* depending on whether all the neurons of a layer are connected with all the neurons of the adjacent layers or there are neurons without connections with neurons of the adjacent layers.

After designing the structure of an ANN, it is necessary to adjust the weights and biases values of its neurons. This adjustment process is called training, and there are different strategies to perform this process. The training process is called *supervised training* when the adjustment is performed providing the input values and the output values related to each set of input values, and *unsupervised training* when only the input values are provided to the ANN. One of the training algorithms most commonly used in literature is the *backpropagation* algorithm (Rumelhart et al., 1988), which is employed to train the *feedforward* MLP. This algorithm is usually implemented using the Levenberg-Marquardt method, which combines the gradient descent method and the Gauss-Newton optimization method (Levenberg, 1944; Marquardt, 1963). Unless training algorithms can be deterministic, there are random conditions that make the adjustment of the weights and biases a non-deterministic procedure. First of all, the result of the adjustment process depends on the initialization of the weights and biases of the ANN. Secondly, the way the training samples are provided to the algorithm also can vary the final values of the weights and biases. Two examples related with this second condition are the order of the samples and the number of samples provided on each iteration of the training algorithm. Finally, training methods usually have several training parameters that modify its performance and therefore its final result. For example, some training methods need a finalization criterion that can modify the trained ANN obtained.

# 3. Objectives

The multivariate analysis problem is a common issue in agricultural and industrial fields, as it has been shown in the Introduction section. Taking this into consideration, the main objective of this thesis is **to design, implement, and evaluate ANN-based models with specific characteristics that allow them to be applied to agricultural and industrial applications**. In particular the three sub-objectives of this thesis are:

- **To design, implement, and evaluate ANN-based models that consider a set of variables larger and/or different than those previously considered in the scientific literature.**

  Classic models usually consider a small number of variables because of the difficulties to relate each input variable with each input variables and to model the interaction of the input and output variables among them. Moreover, the election of the variables that participate in the model is usually done considering the ease to acquire the variable and the ease to relate the variables within the classic model considered. The aim of this objective is to consider models with a larger number of variables and with variables different than the considered in the scientific literature and the state-of-the-art techniques. It gives more or better information to the model and, because of that, it is expected that the performance of the model will improve.

- **To design, implement, and evaluate ANN-based models that minimize the previous knowledge about the process employed to design the models.**

  Most of the models proposed in the scientific literature require previous knowledge of the process modelled and they are proposed to solve a specific problem, but they cannot be adapted to solve other problems. The aim of this objective is to propose models that can be applied to different processes or different conditions of the considered process without modifying them.

- **To design, implement, and evaluate ANN-based models with a good performance in terms of processing time and computation needed.**

  Generally, models that consider more variables have a greater complexity and, because of that, the computing and time requirements of these models increase. This objective aims to minimize the disadvantages of creating more complex models.

# 4. Methodology

A methodology to achieve the objectives presented in the previous section was proposed in this thesis. The aim of this methodology was to design and develop ANN-based models whose characteristics match the three objectives presented in the previous section. The proposed methodology is composed of five stages: literature review, hypothesis formulation, developing and evaluation, result analysis, and results dissemination. The relationships among these stages are represented in Figure 4.

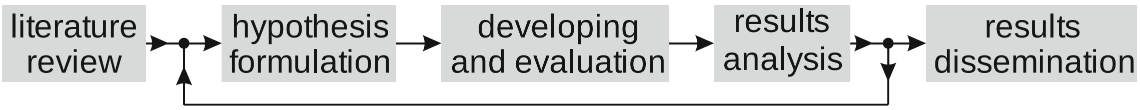



**Figure 4:** Block diagram of the methodology stages of the thesis and the connections among them.

The **first stage of the methodology**, which is the literature review, consists of a study of the state of the art. This stage is needed to acquire the knowledge needed to do the work proposed in this thesis. This stage of the methodology was divided in two sub-stages: the first one was about the modelling technique and the second one was about the process to be modelled.

The study about the modelling technique consists of learning about ANNs and the way they can be employed to build models. For this reason, the study focused only on supervised knowledge, analyzing both classification and regression problems. Moreover, the study also covered the optimization of this type of models, learning about the techniques employed to improve the performance of ANN-based models. Two categories of techniques to improve the models were considered: techniques based on the structure of the ANN and techniques based on the way the information is provided to the ANN. The reference bibliography consisted of the books of S. Haykin (Haykin, 1999), C. M.. Bishop (Bishop, 2007), and Hagan *et al.* (Hagan et al., 1996), among others.

The study about the process to be modelled consists of its observation, conversations with experts, and a literature review about the process. The observation and conversations with experts help to understand the aspects of the process that need to be modelled and allows the researcher to know how to apply an ANN-based model to this process. The literature review provides information about how other authors have dealt with similar problems in the scientific literature and allows the researcher to compare the results obtained using the developed model with the results obtained by other researchers. Section 5.1 presents and explains each one of the processes modelled in this thesis.

The **second stage of the methodology** consists of the formulation of the research hypothesis and the design of an ANN-based model. The research hypothesis formulation is done with the knowledge acquired in the first stage and taking into account the objectives of the thesis. Moreover, the design of an ANN-based model proposed that takes into account this hypothesis is also proposed in this stage.

The **third stage of the methodology** involves the proposal of a scenario to develop and evaluate the proposed ANN-based model, and the selection of parameters used to measure the performance of both the proposed ANN-based model and of other models commonly proposed in the scientific literature. In this stage, the experiments to acquire the data are defined and performed, the processing algorithms to analyze this data are designed and implemented, and the parameters to measure the performance of the models employed as processing algorithms are chosen. The processing algorithms consist of the ANN-based model designed in the second stage and other models that will be used to be compared with the proposed model.

The **fourth stage of the methodology** consists of analyzing the hypothesis proposed. This analysis will be based on comparing the results obtained by each proposed ANN-based method with other methods commonly proposed in similar works proposed in other research articles. This stage will allow the validation, redefinition, or reject of the research hypothesis formulated in the second stage. On the one hand, if the results of the experiments show that the research hypothesis is true, the only remaining work will be the diffusion of the results obtained, which is done in the last stage of the methodology. On the other hand, if the results show that the research hypothesis must be refined or reformulated, the next step of the work will go back to the second stage of the methodology, in order to propose a new research hypothesis.

The **fifth and last stage of the methodology** consists of the diffusion of the results obtained in the work performed. This diffusion is going to be done via an article that describes each work performed. The articles generated as a result of this thesis are going to be submitted to peer-reviewed scientific journals, aiming to publish them in good ranked journals indexed in the Journal Citation Report (JCR), which will ease the diffusion of this work to the scientific world and to validate the research work performed.

# 5. Results

The main results of this thesis are five ANN-based models presented in the five articles included in the compendium of this thesis. Each of these models is proposed to work on a different agricultural or industrial application, proposing each ANN to solve a regression or a classification problem related with the modelled process.

The first subsection of this section presents the five agricultural and industrial applications considered to develop the ANN-based models proposed in this thesis and the second subsection presents the results obtained with the proposed models.

## 5.1. Agricultural and industrial applications modelled

The five agricultural and industrial applications were considered in this thesis to develop the ANN-based models are the tobacco drying process, the switchgrass drying process, the predictive maintenance of a machine, the evaluation of steel pieces in a production line, and the early detection of plant diseases. The next paragraphs present a brief explanation of each one of these five applications.

### 5.1.1. Tobacco drying process

- Description: The tobacco drying process is a controlled drying process in which the temperature and relative humidity are kept under control to get a high quality dried product (Hawks et al., 1986). The control of the temperature and the relative humidity of the dryer is done by means of a heater, which is commonly implemented with a gas heater or a hot water heat exchanger, and an air hatchway, respectively. An example of a control algorithm employed to control this process is shown in Figure 5.

- Problem to be solved: The tobacco drying process requires a high quantity of energy to be done, so it is possible to significantly reduce the production cost of the final product by reducing the energy consumption. Moreover, the final quality of the product is important, so a precise value of the variables that describe the state of the plant is desirable. These precise values could be employed to optimize the drying process by means of both reducing the energy consumption and maximizing the final quality of the dried product.

- Model proposed: A model to predict the variables values in the future or to estimate the variables values inside the tobacco mass as a function of sensors outside the tobacco mass was proposed because it can be employed as a part of an advanced control system. This model was proposed in the first article of the compendium (Martínez-Martínez et al., 2012) and it will hereinafter be referred as **tobacco drying model**.

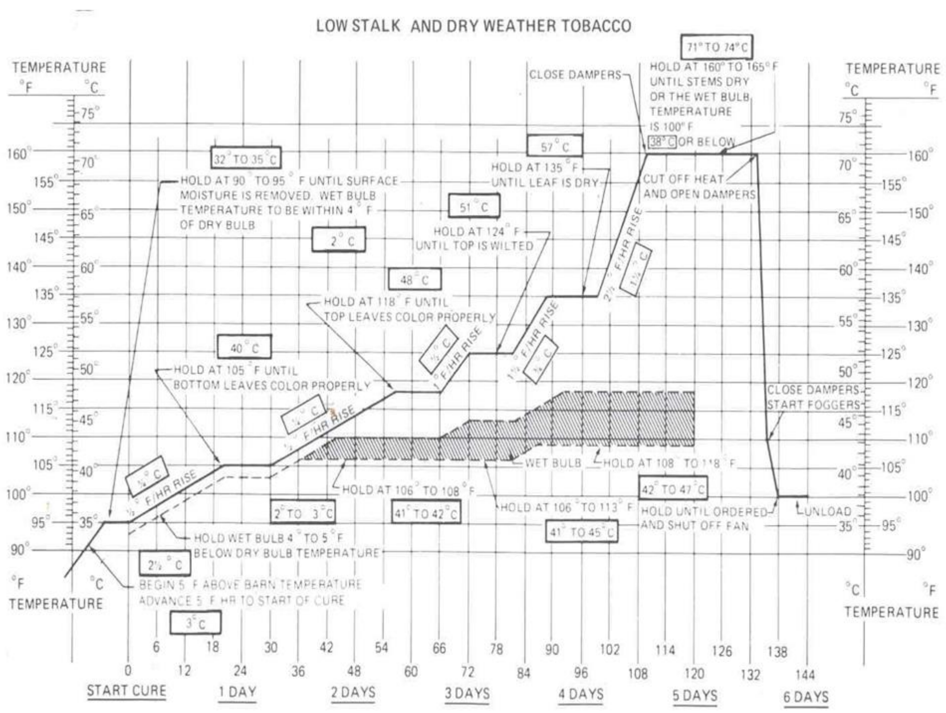



**Figure 5:** Tobacco drying curves graph, in which the target temperature and relative humidity along the time are represented. This curve was extracted from the book of Hawks (Hawks et al., 1986).

### 5.1.2. Switchgrass (*Panicum virgatum*) drying process

- Description: The switchgrass drying process is an uncontrolled outdoor drying process that consists of reducing the moisture content of the switchgrass. This process is affected by weather conditions, with a duration and a quality of the final product that depend on these conditions (Raghavan and Mujumdar, 1987). Figure 6 presents an image of the switchgrass harvested to perform the experiments of this thesis.

- Problem to be solved: The planning of the switchgrass harvesting is difficult because of the uncertainty about the duration of the process. Farmers need to plan the harvest in order to recollect the dried switchgrass in the desired day instead of, leaving the dried switchgrass in the field with the risk that the rain wets the grass again.

- Model proposed: A model to predict the switchgrass moisture content was proposed in the second article of this compendium (Martínez-Martínez et al., 2015b). This model estimates the future value of the moisture content as a function of present and predicted future values of weather variables related with the rain and the temperature and relative humidity of the air. It is expected to give an estimation of the moisture content more accurately than other models that do not consider so many input variables. This model will hereinafter be referred to as **switchgrass drying model**.

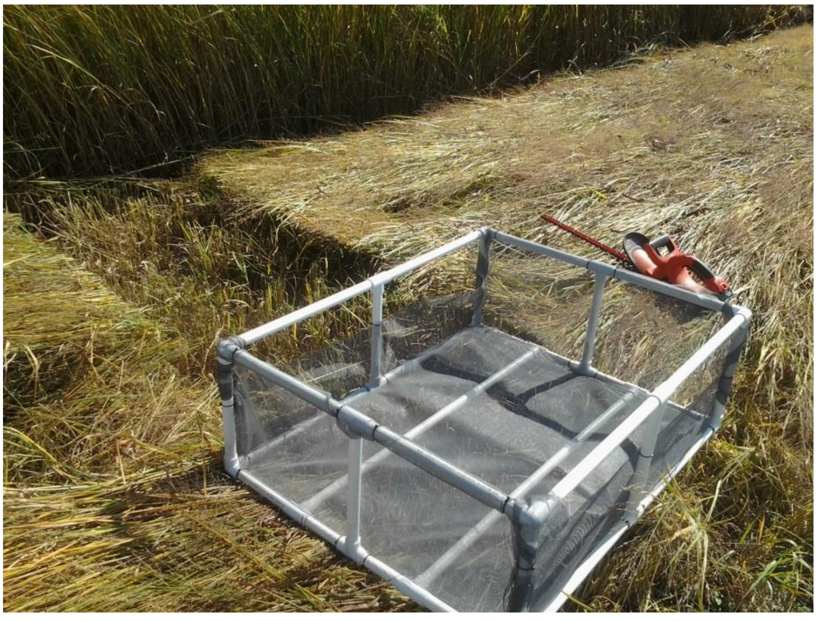


**Figure 6:** Image of the switchgrass harvest done to acquire the data to develop the ANN-based model. It can be seen the switchgrass before and after harvesting, a clipper to cut the switchgrass, and a box to place the switchgrass in the dryer.

### 5.1.3. Predictive maintenance of a machine

- Description of the application: Machinery rotary components generate vibration signals in the machine structure. For this reason it is expected that the working statuses of machinery rotary components will be reflected in these generated vibration signals (Li et al., 2010). Moreover, as the vibration signal can be propagated through the machine structure without a big attenuation, it is expected that the vibration signal acquired in a point far away from the rotary components will have information about the different rotary components that are working in the machine.

- Problem to be solved: The status of a rotary component of a machine can be measured with an intrusive sensor, that is, a sensor in contact with this rotary component. Nevertheless, this mechanical contact has some disadvantages such as the friction of the sensor with the rotary component, which can reduce the life span of the pieces in contact. For this reason, a *sensorless* technique, which is a technique to measure a variable of an element without contact between the sensor and the element, can be used to measure the state of the rotary component avoiding the disadvantages of the intrusive techniques.

- Model proposed: A model to predict the statuses of the rotary components of an agricultural or industrial machine using a single accelerometer was proposed in the third article of this compendium (Martínez-Martínez et al., 2015a). This model used only one accelerometer whatever number of rotary components is considered. Moreover, the model considered the absolute value of the coefficients of the Fourier transform of the vibration signal as the input variables of the model. Figure 7 shows the acquisition system employed to acquire the data needed to implement and test this model, which will hereinafter be referred to as **vibration model of a machine**.

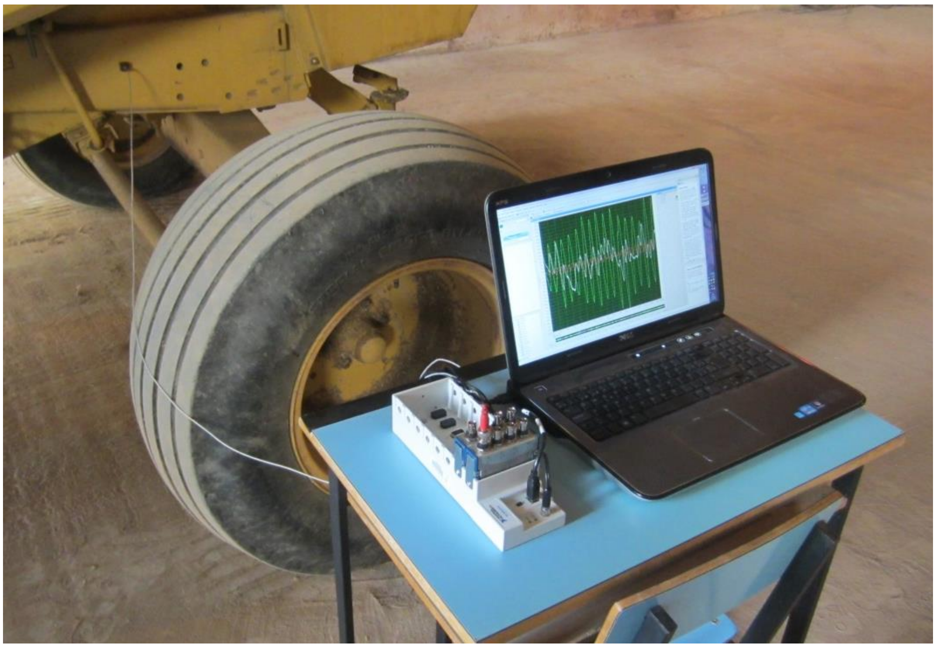


**Figure 7:** Acquisition system employed to acquire the data needed to implement and test the vibration model of a machine. It can be seen an accelerometer on the structure of the machine, an acquisition system connected to this accelerometer by means of a wire, and a laptop to acquire and visualize the data.

### 5.1.4. Evaluation of steel pieces in a production line

- Description of the application: Quality of metallic pieces in the industry must be tested in order to discard the pieces with problems. Traditionally, this task has been done by means of destructive tests, which causes the piece analyzed to be discarded for future use. For this reason, quality evaluation control processes performed with destructive test must be done over a small subset of the pieces production, in order to discard the batches of pieces associated to pieces with problems in the quality control tests. In order to solve the problems associated to destructive tests, non-destructive tests were proposed. These tests do not damage the evaluated piece, allowing the manufacturer to use the evaluated piece after performing the quality control tests and, because of that, also allowing the manufacturer to evaluate all the pieces of the production. The non-destructive test considered in this thesis was based on measuring the inducted impedance after applying a magnetic field to steel pieces, and it is presented in Figure 8.

- Problem to be solved: The problem considered consisted of classifying the analyzed steel pieces as correct or incorrect. This classification method must be non-destructive and suitable to be included in a production line of an industry, and it should have a good success rate.

- Model proposed: A model to classify steel pieces as correctly or incorrectly treated using the eddy-current induced impedance as input was proposed in the fourth article of the compendium. This model used the impedance eddy-current induced by a non-destructive test, considering for each piece several impedances induced by eddy currents generated at different frequencies. The real part and imaginary parts of each impedance value acquired were treated as two different input variables for this model. It will hereinafter be referred to as **ANN-based model for the NDT of steel pieces**.

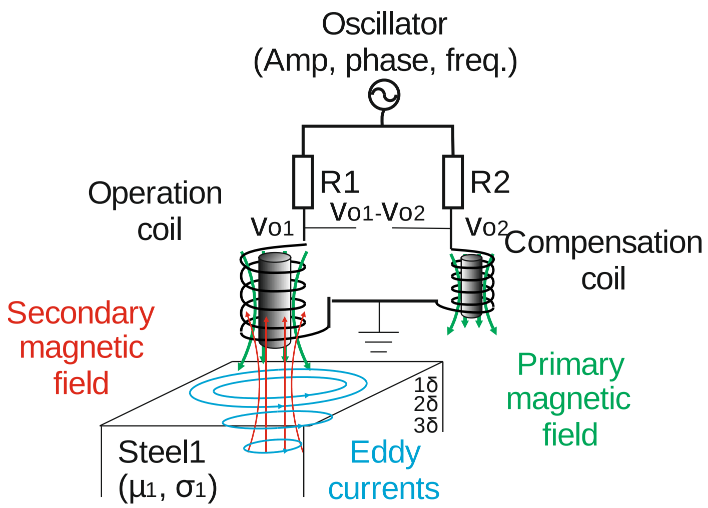



**Figure 8:** Schema of the non-destructive testing method employed in this thesis, were an inductive coil was approached to a steel piece to calculate the impedance induced over the steel piece.

### 5.1.5. Early detection of plant diseases

- Description of the application: Agricultural production is very uncertain because of the variable parameters that affect the crops. Some variables that affect the production of a crop are related with diseases, which can cause significant losses when they act on a crop. Diseases affect the agricultural production in two ways: on the one hand, reducing the final quantity and quality of the product, and on the second hand, requiring to spend money to treat the disease. One way to reduce these losses is the early detection of the diseases because it allows the farmer to treat the crop earlier, decreasing the loss in quantity and quality of the product and the treatment expense. One possible solution to detect diseases in crops earlier is analyzing the reflectivity of the crop, because reflectivity can reflect a disease of a crop before this disease is visible to the human eye.

- Problem to be solved: The problem to be solved is the early detection of the disease *Xanthomonas malvacearum* on common bean crops. To do that, the reflectivity of the bean plants canopy and leaves will be acquired an analyzed in order to select the best technique to estimate the disease severity. Different reflectance spectrum bands, data acquisition techniques, and processing techniques will be considered in order to design an optimum estimation method. The data acquisition employed in this application is presented in Figure 9.

- Model proposed: A model to estimate the disease severity as a function of the reflectance spectral data was proposed in the fifth article of the compendium. The input of the model were the most representative components obtained as a result of the Principal Components Analysis of the reflectance spectral data, and the output of the model was the disease severity. This model will hereinafter be referred to as **reflectance based disease model**.

**Figure 9:** Acquisition system employed to acquire the reflectance data from the bean crops, where the reflectometer is placed in the top of the structure in order to acquire the canopy data and store it in the computer.

## 5.2. Results obtained

This subsection presents the results obtained for the five agricultural and industrial applications described in the subsection 5.1. First, the **general results** of this thesis are grouped into three categories: results about the set of variables considered in the proposed model, results referred to the previous knowledge about the process considered in the models, and results about the optimization of their processing time and computational load. Second, the main numerical results of each application, which are referred as **particular results**, will be briefly presented in this subsection. The remaining results of each application are included in the five articles presented in the second part of this thesis.

The first category groups the **general results** concerning to **the set of variables included in the model**. Four of the five models presented in this thesis have results that can be included in this category. The tobacco drying model (Martínez-Martínez et al., 2012) considered three measurement points: two of them outside the tobacco mass and another one inside it. This contrasts with the use of a single measurement point proposed by most of the similar drying models proposed in the scientific literature. The switchgrass drying model (Martínez-Martínez et al., 2015b) characterized this outside drying process considering as many weather variables as necessary to describe this process. This need gave rise to a model with a number of variables larger than other comparable models found in the scientific literature. Moreover, this model included information about rain events, which to the best of my knowledge had not been considered in other similar models. The vibration model of a machine (Martínez-Martínez et al., 2015a) also considered different input variables compared with other similar models found in the scientific literature. The main difference is that a single measurement point placed far away from the rotary components was considered to acquire the input signal. Finally, the ANN-based model for the NDT of steel pieces also included the novelty of considering as the input variables the impedance at several frequencies.

The second category groups the **general results** related to the **previous knowledge necessary to build the proposed models and its generalization capability**. The models proposed in this thesis were designed with the minimum previous information possible about the process to be modelled. Moreover, the proposed models were designed as general as possible in order to not limit its functionalities for a specific application. For example, the tobacco drying model (Martínez-Martínez et al., 2012) does not have information about the set points of the drying controller. This fact makes this model adjustable to other drying processes with different configurations or to other similar products to be dried. Both the tobacco and the switchgrass drying models (Martínez-Martínez et al., 2012; Martínez-Martínez et al., 2015b) do not need information about the product to be dried. For this reason it is expected to yield good results with other similar biomass or grass drying processes. The vibration model of a machine (Martínez-Martínez et al., 2015a) is the main example in this thesis of a model that does not include previous knowledge of the process modelled in its design. This model uses the absolute value of the Fourier coefficients of the vibration signal as the input variables, instead of performing a feature extraction stage to obtain the input variables. Feature extraction stages usually need information of the rotary components considered, such as its angular speed, to configure the features to extract. Moreover, this model does not need information about the location of the measurement point, which suggests that this location is not going to be a problem if this model is used in other machines. The ANN-based model for the NDT of steel pieces was also designed without taking into account the specific process modelled. Therefore, the frequencies considered were chosen as representative values of the interest frequency range instead of choosing the frequencies with the best results for the problem analyzed. Finally, the design of the reflectance-based disease model was done in a similar way to the ANN-based model for the NDT of steel pieces. Thus, the wavelength ranges were selected among the available ranges of the RapidEye satellite instead of analyzing the wavelength ranges acquired and selecting the ranges with the best results.

The third and last category groups the **general results** concerning to the **optimization of the processing time and computational load of ANN-based models**. In general, the use of models based on ANNs does not require high processing requirements because of the high parallelism of ANNs and the simplicity of a single neuron. The drying processes modelled in this thesis (Martínez-Martínez et al., 2012; Martínez-Martínez et al., 2015b) does not have special processing or latency requirements in order to be implemented in a real application. For this reason, the use of an ANN-based model and the usage of features easy to be calculated as inputs of the model ensure that they do not have problems with the processing time. The vibration model of a machine (Martínez-Martínez et al., 2015a) needed to be optimized by means of a GA in order to improve its performance and to reduce the time needed to train the model. Moreover, as the proposed model does not have a feature extraction stage, this model could be implemented in a real time application because of the low latency of the operations performed: the FFT, the absolute value, and a 3 layers ANN. The ANN-based model for the NDT of steel pieces was also optimized and evaluated in order to check whether it could be used in a real production line, as can be seen in the work of Garcia-Martin *et al.* (Garcia-Martin et al., 2014). Finally, the reflectance-based disease model tried to optimize its operation choosing among multiple parameters. The first parameter was the number of PCA components, which is related with the number of inputs of the ANN. The second parameter was the reflectance wavelength ranges considered. The third parameter was choosing between using the FDR or using the reflectance data. The comparative work performed in this article aimed to propose an optimum method to estimate the disease severity. The method was optimized in terms of input information required by the model, the complexity of the acquisition system needed, and the processing operations required. Moreover, this model only needs simple operations to be implemented: a signal pruning, a signal subsampling, a PCA, and the ANN. For this reason, the latency of the model is expected to be low, which allows to implement it on a real-time device.

Finally, the key **particular results** of each one of the five applications proposed are presented below.

- Tobacco drying model: The experiments performed in the first article of the compendium (Martínez-Martínez et al., 2012) evaluated the estimation and prediction capabilities of this model. This model was able to estimate the temperature inside the tobacco mass using information of sensors placed outside the tobacco mass with an error under 2%. Moreover, it can predict the temperature inside the tobacco mass with a time horizon of 150 minutes with an error 1.5 times lower than the obtained by an interpolation method.

- Switchgrass drying model: Two experiments were performed in the second article of the compendium (Martínez-Martínez et al., 2015b). First, the proposed model was compared with other well-known models when drying processes in constant conditions were performed. This comparison showed that the proposed model estimated the moisture content with a correlation coefficient of 99.9%. Moreover, the Mean Squared Error (MSE) was $2.6 \cdot 10^{-4}$, which was more than 3 times better than the best of other six well-known models analyzed. Second, this article analyzed the performance of this model for processes with variable air conditions and rain. For these processes, the MSE was lower than $3.3 \cdot 10^{-3}$ and the correlation coefficient higher than 98.7%.

- Vibration model of a machine: The third article of the compendium (Martínez-Martínez et al., 2015a) analyzed the classification success rate of this model and the performance of the GA-based training method proposed to adjust the model. First, the mean success rate of the proposed method was 92.96%, being able to estimate four of the five components considered with more than 95% of success rate. Second, the GA-based training method was able to reduce 70% of the iterations needed to train the proposed ANN-based model.

- ANN-based model for the NDT of steel pieces: The fourth article of the compendium evaluated the proposed model in terms of precision and area under the Receiver Operating Characteristic (ROC). This evaluation was undertaken comparing the performance of the proposed method with the performance of an Euclidean Distance Classifier (EDC) and several monofrequency Linear Discriminant Analysis (LDA) classifiers. The precision of the proposed model was 95%, improving by 5% the performance of the EDC and the best LDA. The area under the ROC curve of this model was 0.98, which was 2% better than the EDC and 5% better than the best LDA.

- Reflectance based disease model: The fifth article of the compendium evaluated the performances of different configurations of the proposed model by means of the Root

MSE (RMSE) and the correlation coefficient. The best configurations of the proposed model were able to determine the severity of the disease with a RMSE lower than 2.4 and with a correlation coefficient greater than 92%.

# 6. Original contributions to the state of the art

The work performed in this thesis gave rise to two original contributions to the state of the art. These contributions were made to the state of the art of two of the five applications considered in this thesis. These original contributions are presented and explained below, explaining the differences between the work performed in this thesis and the work presented in other published articles.

## 6.1. First original contribution to the state of the art

The first original contribution is a **model** for **drying processes** performed outdoors that considers variable air conditions and **raining events information**. This model was proposed in the second article of this thesis (Martínez-Martínez et al., 2015b). The problem to be solved, the solution proposed, the advantages of the original contribution, and the comparison between this contribution and other similar models are presented below.

- Problem to be solved: Drying processes performed outdoors have a lot of factors that affect the drying process. Most scientific articles about drying processes modelling are applied to drying processes or with outdoor drying processes with very limited conditions.
- Solution proposed: The proposed model used information about the air temperature, the air relative humidity, and the rain events. Thus, the air temperature and the air relative humidity were characterized by their standard deviation and the $10^{th}$, $50^{th}$, and $90^{th}$ percentiles. Moreover, the initial time, the duration and the water precipitation were used to characterize the rain events.
- Advantages of the model proposed: This model is able to estimate the moisture content evolution in a more accurate way than other models because of its ability to model processes in outside conditions.
- Comparison: Several drying models proposed in the scientific literature model drying processes with constant air conditions (Movagharnejad and Nikzad, 2007; Erenturk and Erenturk, 2007; Wang, 2002). These drying models can be employed for drying processes performed in drying chambers, but not for drying processes performed on the land. There are other models in the scientific literature that consider uncontrolled air conditions (Barr and Brown, 1995; Diamante and Munro, 1993). Nevertheless, they only can model drying processes performed on the land when there are not rain events during the drying process. To the best of my knowledge, the first article of the scientific literature that models a land-drying process considering variable air temperature, variable relative humidity, and rain events as variables that affect the drying process was the proposed in the second article of this thesis (Martínez-Martínez et al., 2015b).

## 6.2. Second original contribution to the state of the art

The second original contribution is a **predictive maintenance** method that employed the vibration signal acquired from a **single measurement point** to monitor **several rotary components**. This model was proposed in the third article of this thesis (Martínez-Martínez et al., 2015a). The problem to be solved, the solution proposed, the advantages of the original contribution, and the comparison between this contribution and other similar models are presented below.

- <u>Problem to be solved:</u> Methods for the predictive maintenance of the rotary components of machines need to have information about all the rotary components considered. The acquisition systems of these methods proposed in the scientific literature need to be adapted for each machine taking into account the rotary components considered. This complicates the design of the acquisition system and makes it less scalable when more rotary components are considered.
- <u>Solution proposed:</u> This predictive maintenance method acquires the vibration signal from a point of the structure of the machine, and not very close to the rotary components monitored. As the machine has a rigid structure, the acquired vibration signal will be a sum of the vibration signals generated by the rotary components monitored, the vibration signals generated by other elements of the machine, and noise. It is expected that this signal will have enough information to determinate the state of the rotary components monitored.
- <u>Advantages of the model proposed</u>: The main advantage of this contribution is the reduction of the cost of the acquisition system. This reduction is because **it only needs a single sensor instead of needing a sensor for each rotary component monitored.** Moreover, it makes this method **more scalable** because it is possible to consider more rotary components with the same acquisition system.
- <u>Comparison:</u> There are a lot of works in the scientific literature that employ the vibration signal to estimate the statuses of the rotary components of machines. On the one hand, there are a lot of works that estimate the status of a single rotary component using one or more accelerometers (Chwan-Lu et al., 2014a; Chwan-Lu et al., 2014b; Liang et al., 2014; Yang et al., 2015) On the other hand, there are some works that estimate the statuses of several rotary components using a single accelerometer. Nevertheless, all of these works consider a group of rotary components placed close to each other (Bin et al., 2012; Li et al., 2013; Lu et al., 2015; Nembhard et al., 2015). To the best of my knowledge, the first articles that have proposed methods to estimate the state of rotary components placed far from each other were the published as a result of this thesis (Martínez-Martínez et al., 2015a; Ruiz-Gonzalez et al., 2014).

# 7. Conclusions

The results obtained from this thesis and presented in the previous sections suggest the next three conclusions, which are aligned with the three sub-objectives considered for this thesis.

- The first conclusion is that models proposed in the scientific literature can be improved by means of modifying the variables considered to describe the application. These variables can be modified considering other variables than the commonly proposed in the scientific literature or adding new variables to the commonly used by other authors. This conclusion is aligned with the first sub-objective of this thesis, and it was accomplished in the next models:
  - Tobacco drying model: The performance of this kind of model can be improved considering a measurement point inside the product to be dried, as it can be seen in the first article of the compendium (Martínez-Martínez et al., 2012).
  - Switchgrass drying model: It is possible to estimate the state of drying processes with rain events with good precision using information about the rain events as input. This conclusion was extracted from the second article of the compendium (Martínez-Martínez et al., 2015b).
  - Vibration model of a machine: The state of the rotary components of a machine can be estimated employing a single measurement point located far away from these rotary components. This conclusion can be seen in the third article of the compendium (Martínez-Martínez et al., 2015a).
  - ANN-based model for the NDT of steel pieces: The usage of impedance data at several frequencies instead of using data at a single frequency as the inputs of a NDT method improves significantly the performance of the NDT method, as it can be seen in the fourth article of the compendium.

- The second conclusion is that it is possible to create ANN-based models without previous knowledge of the application modelled because of its generalization capability. This conclusion is related with the second sub-objective of this thesis, and it can be seen in the five articles of the compendium:
  - Tobacco drying model: It is expected to obtain good results with this model when other control algorithms are employed in the drying process. This requires researchers to train the ANN with data acquired in all the conditions considered, because this model does not include information about the control algorithm. This conclusion was extracted from the first article of the compendium (Martínez-Martínez et al., 2012).
  - Tobacco drying model and switchgrass drying model: It is expected to get good results with these two drying models when they are applied to other agricultural products with similar drying processes. It is possible because these models do not use information about the product to be dried, as it can be seen in the first and the second articles of this compendium (Martínez-Martínez et al., 2012; Martínez-Martínez et al., 2015b).
  - Vibration model of a machine: The design of this ANN-based model suggests that the proposed ANN architecture can be easily adapted to other machines and to other rotary components. The proposed ANN-based model does not include a feature extraction stage and does not use information about the rotary

components monitored or the location of the measurement point. To adapt this model it is necessary to consider as many output neurons as rotary components are considered, and to train this ANN with data of the new process modelled. This conclusion can be seen in the third article of the compendium (Martínez-Martínez et al., 2015a).

- ANN-based model for the NDT of steel pieces and reflectance based disease model: The results obtained in the articles related to these models suggest that other acquisition systems can be employed to reply the proposed experiments. For example, acquisition systems that have different frequencies or wavelengths can be employed as input information for the proposed models. It is possible because the frequencies and wavelengths considered in the model were not selected using the previous knowledge about the two processes considered. The only requirements for the acquisition system are to have a similar frequency band or wavelength band respectively and a resolution similar to the systems employed in this thesis. This conclusion was extracted from the fourth and the fifth articles of this compendium.

- The third conclusion is that the processing time and the computation needed is highly dependent on the way the ANN is configured and optimized. This conclusion is aligned with the third objective of this thesis. This conclusion suggests that it is important to perform a good design and optimization of ANN-based models. It is especially important when the size of the model is large, because the improvement of the optimization is significant. The processing time and computation needed was taken into account in the design or the configuration of the five ANN-based models proposed:
  - Tobacco drying model, switchgrass drying model, and reflectance based disease model: The employment of features easy to be calculated and the avoidance of variables that require more processing time allows researchers to build ANN-based models with low latency. Examples of these recommended features are mean value, actual value, percentile, and standard deviation. This kind of variables makes the models suitable for near real-time applications such as the proposed for these models. It can be seen in the first, second, and fifth articles of this compendium (Martínez-Martínez et al., 2012; Martínez-Martínez et al., 2015b).
  - Vibration model of a machine and ANN-based model for the NDT of steel pieces: Results obtained in these articles suggest that it is possible to create precise ANN-based models for real-time applications by means of suppressing the feature extraction stage and using a massively parallel processor unit to take into advance the parallelism of the ANNs. The third and fourth articles of this compendium and the article of Garcia-Martin *et al.* present the works from which this conclusion was extracted (Martínez-Martínez et al., 2015a; Garcia-Martin et al., 2014).
  - Vibration model of a machine: The usage of genetic algorithms to train large ANNs allows to improve the performance of the ANN-based model and to reduce the time needed to adjust the model, as it can be seen in the third article of this compendium (Martínez-Martínez et al., 2015a).

In summary, the main conclusion of this thesis is that ANN-based models are a good alternative to the methods proposed in the scientific literature to solve both regression and classification problems of the agricultural and the industrial fields.

# 8. Future work

This thesis showed the ability of ANNs to build multivariate models for agricultural and industrial applications. During the work presented in this thesis several lines of future work emerged. These future work lines can be divided into two groups: **general future work lines**, which are associated to the overall content of this thesis, and **particular future work lines**, which are associated to each one of the five applications considered in this thesis.

The **general future work lines** can be divided into **future research lines** and **future development lines**. The work performed in this thesis suggests two **future research lines**. The first one is to design and implement models similar to the presented in this thesis for other agricultural or industrial applications. As the proposed models have been validated with five different agricultural and industrial applications, it is expected that they can be applied to other applications to solve regression and classification problems in multivariate processes. The second future research line is to improve the models proposed in this thesis. One example of an improvement related with these models is the proposal of a control system in which an estimation model acts as a part of the control logic, in order to use the output of the ANN-based model as one of the control variables. Another improvement is to design and implement training procedures that could be employed while the ANN is working and not only before it, in order to allow the models to learn from its errors during the execution time and improve its future performance. These two improvements will allow the system to act as an artificial intelligent system because it can perform three tasks. The first task is to make predictions and estimations as a function of several variables that it can measure. The second task is to use these predictions and estimations to make decisions and perform actions. The third task is to learn from its errors in order to improve its future performance. The **future development lines** that arise from this thesis consist of using the knowledge generated in this work to use them in the agricultural and industrial fields. These future lines can be divided into three groups of tasks. The first group consists of designing and implementing the hardware and software necessary to perform the acquisition of the variables considered. The second group consists of processing these variables with the ANN-based models presented in this thesis. The third group of tasks consists of performing an action with the output of the ANN-based model. Some examples of actions usually performed with the output of the model are representing it, storing it, or activating an actuator that interacts with the process.

The **particular future work lines** are associated with each one of the five applications considered in this thesis. First, some of these work lines are particularizations of the general future work lines presented above. Second, most of them are included in the articles attached to this thesis. For these reasons, the particular future work lines are going to be briefly explained in order to not repeat information included in other parts of this thesis.

- There are several future lines of work related with the tobacco drying process. The first one will be a model for the tobacco drying process considering the position of the measurement point inside the tobacco mass as a variable of the model. This future line will allow the researchers to know the variation of the interest variables inside the tobacco mass and in this way to choose the measurement point that describes better the drying process. There are some other future lines using the estimation of the interest variables in the future, such as using them to control the process or employing this information to deploy a predictive alarm system.

- The work done with the switchgrass drying process has suggested also some future work lines, some of them similar to the proposed for the tobacco drying process. One future work is to repeat the proposed experiments in a real drying process instead of in a simulated drying process. Another future line, which is related with the previous one, is to consider more variables in the model, such as the wind direction and the wind speed or some soil characteristics that may affect the drying process. Finally, the last future work line will be the application of this ANN-based model to plan the optimum time and date to mow the switchgrass.

- The work for the predictive maintenance of a machine also has some future lines that could be considered for its further implementation. One of them is repeating the experiments again in real conditions to check whether the proposed model can be adapted to real working conditions. Another future line is the usage of the proposed method applied to other agro-industrial machines to generalize it. Finally, some future lines to improve the functionalities of the proposed method emerged: the detection of the speed of the rotary components in addition to its state, the usage of more than one sensor to acquire the vibration signal and the usage of other sensors like microphones, and an analysis of the relationship between the position of the sensor(s) and the results of the method.

- The work about the NDT for steel pieces has some future work lines to be implemented. One of them is to consider more variables in this application in order to check whether it is possible to design a general model for this application or it is necessary to design several models. Some possible variations on the experiments proposed in this thesis are to consider different problems in the steel pieces, to evaluate pieces of different geometries, or to analyze pieces of different materials. Other future line is the usage of more frequencies to acquire the induced impedance. Another one is to consider other NDT input variables for the proposed method in order to have a method more robust and with better performance compared with the one presented in this thesis.

- The method to estimate diseases in common bean crops also has some future lines that could be considered. The first one is to recollect and consider more information about the crops in order to improve the method proposed. For example, information about the weather conditions of the crop could give information to not consider humidity problems in the leaves as problems with the disease analyzed. Another future line is to register the part of the plant where the data was acquired in order to analyze the variability along the plant. This analysis will show the best part of the plant to acquire data to implement this method.

# 9. Merits and diffusion of the results

This section presents a summary of the merits obtained during the research process of this thesis and the diffusion of the results obtained: the publications produced, the projects in which I have participated, the research stays done in other international research groups, the courses taught, and other merits.

## 9.1. Publications

This subsection presents the articles sent to peer-reviewed international journals, congresses or conferences, which has been published or accepted for publication.

### 9.1.1. Publications in JCR-indexed international journals

The articles published in JCR-indexed international journals are presented below. Moreover, Table 1 shows the main quality parameters of the journals where these articles were published:

- Martínez-Martínez V, Baladrón C, Gomez-Gil J, Ruiz-Ruiz G, Navas-Gracia LM, Aguiar JM, Carro B (2012). Temperature and Relative Humidity Estimation and Prediction in the Tobacco Drying Process Using Artificial Neural Networks. Sensors 12 (10):14004-14021.
- Martínez-Martínez V, Gomez-Gil J, Stombaugh TS, Montross MD, Aguiar JM (2015). Moisture Content Prediction in the Switchgrass (Panicum Virgatum) Drying Process Using Artificial Neural Networks. Drying Technology 33 (14):1708-1719.
- Martínez-Martínez V, Gomez-Gil FJ, 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 Systems with Applications 42 (17–18):6433-6441.
- Ruiz-Gonzalez R, Gomez-Gil J, Gomez-Gil F, Martínez-Martínez V (2014). An SVM-Based Classifier for Estimating the State of Various Rotating Components in Agro-Industrial Machinery with a Vibration Signal Acquired from a Single Point on the Machine Chassis. Sensors 14 (11):20713-20735
- Garcia-Martin J, Martínez-Martínez V, Gomez-Gil J (2014). Clasificación del tratamiento térmico de aceros con ensayos no destructivos por corrientes inducidas mediante redes neuronales. Dyna 89 (5):526-532.

<table>
<tr><th>Journal</th><th>ISSN</th><th>Year</th><th>Impact Factor</th><th>Journal Ranking</th></tr>
<tr><td rowspan="2">Sensors</td><td rowspan="2">1424-8220</td><td>2012</td><td>1.953</td><td>• Q1 on Instruments & Instrumentation (8/57)<br>• Q3 on Chemistry, Analytical (38/75)<br>• Q3 on Electrochemistry (15/26)</td></tr>
<tr><td>2014</td><td>2.245</td><td>• Q1 on Instruments & Instrumentation (10/56)<br>• Q2 on Chemistry, Analytical (31/74)<br>• Q3 on Electrochemistry (14/28)</td></tr>
<tr><td>Drying Technology</td><td>0737-3937</td><td>2014</td><td>1.518</td><td>• Q2 on Engineering, Mechanical (36/130)<br>• Q2 on Engineering, Chemical (62/134)</td></tr>
<tr><td>Expert Systems with Applications</td><td>0957-4174</td><td>2014</td><td>2.240</td><td>• Q1 on Operations research & Management Science (12/81)<br>• Q1 on Engineering, Electrical & Electronic (48/249)<br>• Q1 on Computer Science, Artificial Intelligence (29/123)</td></tr>
<tr><td>Dyna</td><td>0012-7361</td><td>2014</td><td>0.179</td><td>• Q4 on Engineering, Multidisciplinary (82/83)</td></tr>
</table>

**Table 1:** Quality parameters of the JCR-indexed journals where articles with the results of this thesis have been published.

### 9.1.2. Publications in other peer-reviewed international journals

This subsection presents the articles published in peer-reviewed international journals that have not been indexed in the JCR index:

- Martínez-Martínez V, Baladrón C, Gomez-Gil J, Ruiz-Ruiz G, Navas-Gracia LM, Aguiar JM, Carro B (2013) Modelo basado en redes neuronales artificiales para el cálculo de parámetros ambientales en el proceso de curado del tabaco. Informador Técnico 77 (1):35-46.

### 9.1.3. Publications in international conference proceedings

This subsection presents the works presented in international conferences and international workshops:

- Martínez-Martínez V, Ruiz-Ruiz G, Navas-Gracia LM, Gomez-Gil J, Miñón-Martínez J, Blanco I. Monitoring and Estimation of Tobacco Drying Variables for Process Modeling. International Conference of Agricultural Engineering CIGR-AgEng 2012, Valencia (Spain), July 7$^{th}$ 2012, ISBN 978-84-615-9928-8.

- Martínez-Martínez V, Ruiz-Ruiz G, Navas-Gracia LM, Pérez R, Correa Guimarães A, Gomez-Gil J. Definición de una arquitectura de sistema de adquisición y gestión de datos para la monitorización de redes de sensores en aplicaciones agroforestales. VI Congreso Ibérico de AgroIngeniería, Évora (Portugal), September 5$^{th}$ 2011.

- Martínez-Martínez V, Rodríguez-Pérez I, Vázquez-Sánchez E., Ruiz-Ruiz G, Navas-Gracia LM, Gomez-Gil J. Una revisión sobre el uso de la Transformada de Hough en la Agricultura. III Jornadas Ibero-Americanas de Agricultura de Precisão, Évora (Portugal), March 2$^{nd}$ 2010.

- Rodríguez-Pérez I, Martínez-Martínez V, Vázquez-Sánchez E., Ruiz-Ruiz G, Navas-Gracia LM, Gomez-Gil J. Una revisión del procesado de datos mediante redes neuronales en la agricultura. III Jornadas Ibero-Americanas de Agricultura de Precisão, Évora (Portugal), March 2$^{nd}$ 2010.

### 9.1.4. Publications in national conference proceedings

This subsection presents the works presented in national conferences:

- Miñón-Martínez J, Ruiz-Ruiz G, Navas-Gracia LM, Rad-Moradillo JC, Martínez-Martínez V. Sistema de Monitorización del crecimiento de algas filamentosas basado en análisis de imagen. I Symposium Nacional de Ingeniería Hortícola: “La Agromótica en la Horticultura”, Orihuela, Alicante (Spain), February 20$^{th}$ 2014.

- Ruiz-Ruiz G, Martínez-Martínez V, Navas-Gracia LM, Gomez-Gil J, Miñón-Martínez J. Sistema de control avanzado para el curado de tabaco Virginia. I Symposium Nacional de Ingeniería Hortícola: “La Agromótica en la Horticultura”, Orihuela, Alicante (Spain), February 20$^{th}$ 2014.

- Miñón-Martínez J, Martínez-Martínez V, Ruiz-Ruiz G, Navas-Gracia LM, Pérez R., Correo Guimarães A. Nueva arquitectura de adquisición y gestión de datos para la monitorización de redes de sensores en aplicaciones agrarias. XIII Congreso Nacional de Ciencias Hortícolas. “Convergencia a las Tecnologías Hortofrutícolas”, Aguadulce, Roquetas de Mar, Almería (Spain), April 16$^{th}$ 2014.

- Martínez-Martínez V, Casaseca de la Higuera P, Martínez Fernández M, Alberola C. Un algoritmo de compresión de señales de ECG basado en un modelo de síntesis. Análisis comparativo. XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica – CASEIB, Cáceres (Spain), November 16$^{th}$ 2011.

## 9.2. Projects

This subsection presents information about a project in which I have participated during my PhD thesis:

- Distribución inteligente de servicios multimedia utilizando redes cognitivas adaptativas definidas por software:
  - Expected date to start the project: January, 2016.
  - Duration: 36 months.
  - Funding entity: Ministerio de Economía y Competitividad (Ministry of Economy and Competitiveness, Spain).
  - Participant entities: Polytechnic University of Valencia, University of Valladolid

- Principal investigator: Belén Carro Martínez.
- Role: Investigator.

Moreover, I also collaborated with another proyect entitled “Mejora de la competitividad del sector del tabaco en Extremadura: nuevos procesos y productos”, with project number IDI-20100986. This project was financed by the Centre for Industrial Technological Development (Centro para el Desarrollo Tecnológico Industrial – CDTI) (Ministry of Economy and Competitiveness, Spain), and as a result of this collaboration I did the research presented in the first article of this compendium (Martínez-Martínez et al., 2012).

## 9.3. Research stays

This subsection presents information about the three research stays done in this thesis:

- First stay:
  - Organization: University of Kentucky.
  - Center: Biosystems & Agricultural Engineering Department.
  - Location: Lexington (Kentucky), USA.
  - Start of the stay: September, 2012.
  - Duration: 3 months.
  - Funding entity: University of Valladolid (Spain).
  - Stay purpose: Research about the switchgrass drying process using ANNs.

- Second stay:
  - Organization: University of Kentucky.
  - Center: Biosystems & Agricultural Engineering Department.
  - Location: Lexington (Kentucky), USA.
  - Start of the stay: September, 2014.
  - Duration: 3 months.
  - Funding entity: University of Valladolid (Spain).
  - Stay purpose: Research about the switchgrass drying process using ANNs and work for a CNH Industrial company project about estimating the flux of water through nozzles by means of the noise signal.

- Third stay:
  - Organization: Universidade Federal de Viçosa.
  - Center: Departamento de Engenharia Agrícola.
  - Location: Viçosa (Minas Gerais), Brazil.
  - Start of the stay: March, 2015.
  - Duration: 2 months.
  - Funding entity: Banco Santander, by means of the “Becas Iberoamérica. Jóvenes profesores e investigadores. Santander Universidades” grant.
  - Stay purpose: Research about the bean diseases detection by means of reflectance data using ANNs.

## 9.4. Teaching

This subsection presents information about two courses given during this thesis:

- Redes Neuronales Aplicadas al Ámbito Agrícola
  - Type of course: This course was one part of the subject "ENG 791 – Tópicos Especiais II", which belongs to the Post-Doc program entitled Programa de Pós-Graduação em Engenharia Agrícola.
  - Date: April-May 2015.
  - Place: Departamento de Engenharias Agrícola – Centro de Ciências Agrárias – Universidade Federal de Viçosa (UFV), Mina Gerais, Brazil.
  - Teaching load: 30 hours.

- Redes Neuronales Aplicadas a Problemas de Ingeniería
  - Type of course: Training course for under-graduate students.
  - Date: May 2015.
  - Place: Departamento das Engenharias de Telecomunicações e Mecatrónica – Universidade Federal de São João del-Rei (UFSJ), Mina Gerais, Brazil.
  - Teaching load: 16 hours.

## 9.5. Other merits

This subsection presents information about other merits obtained during this thesis:

- Registration of a patent with the next information:
  - Name: Secadero de biomasa de algas y proceso de secado.
  - Country: Spain.
  - Patent number: P201230037.
  - Authors: Jorge Miñón Martínez, Luis Manuel Navas Gracia, Gonzalo Ruiz Ruiz, Víctor Martínez Martínez.
  - Owner: Universidad de Valladolid (University of Valladolid).
  - Issue date: January 13$^{th}$ 2012.

- Obtained a prize with the next information:
  - Name of the prize: Premio ¿Investigamos? 2011.
  - Founding entity: Itagra.CT –Centro Tecnológico Agrario y Agroalimentario.
  - Date: June 2011.
  - Name of the project: Desarrollo de un sistema SCADA para el control del curado de tabaco.
  - Authors: Víctor Martínez Martínez.

- Obtained a prize with the next information:
  - Name of the prize: Premio Prometeo 2015.

- Founding entity: Fundación General Universidad de Valladolid.
- Date: June 2015.
- Name of the project: Sistema para la asistencia al conductor empleando sensores inerciales y de automoción.
- Authors: Carlos Meléndez Pastor, Rubén Ruiz González, Víctor Martínez Martínez.

* Co-directed a Final Degree Project with the next information:
  - Title of the project: Diseño, implementación y testeo de un Sistema de fusión de datos de posicionamiento GPS y datos de sensores, empleando filtro de Kalman extendido y sistema de comunicación OBDII, para la mejora de la precisión de posicionamiento de un vehículo.
  - Degree: Ingeniero Superior de Telecomunicación, Universidad de Valladolid.
  - Date: July 2015.
  - Author: Carlos Meléndez Pastor.
  - Directors: Rubén Ruiz González, Víctor Martínez Martínez, Jaime Gómez Gil.

# II Articles

The second part of the document presents a collection of articles that reveal the main results of this thesis. This part contains five sections, each one is related to one of the five articles included in the compendium. Each section will contain a brief introduction to the article and a copy of the article published or expected to publish. Three of the articles of this compendium have been accepted or published in JCR-indexed peer-reviewed journals of the first quartile of their category in the 2014 Journal Ranking. The other two are articles that have not been already published but are expected to be submitted in a JCR-indexed peer-reviewed journal soon. Each one of the next sections contains an introduction to the article, some bibliographic information, and a version of the article published or expected to be submitted.

# Article 1: Tobacco drying process

The first article presents the **tobacco drying model** described in the "Introduction and Summary" part of the document. The main bibliographic data about this article are presented below:

- Title: Temperature and Relative Humidity Estimation and Prediction in the Tobacco Drying Process Using Artificial Neural Networks.
- Authors: Martínez-Martínez V, Baladrón C, Gomez-Gil J, Ruiz-Ruiz G, Navas-Gracia LM, Aguiar JM, Carro B
- Journal: Sensors
- Editor: MDPI
- Impact factor: 1.953 (2012)
- Journal Ranking:
    - Q1 on *Instruments & Instrumentation* (8/57).
    - Q3 on *Chemistry, Analytical* (38/75).
    - Q3 on *Electrochemistry* (15/26).
- Date of publication: October 17$^{th}$, 2012.
- ISSN: 1424-8220.
- Volume (Issue): 12(10).
- Pages: 14004-14021.
- DOI: 10.3390/s121014004.
- URL: http://www.mdpi.com/1424-8220/12/10/14004.
- Number of cites: 5.

The article presented below is a version of the original article published in the previously mentioned journal. This version of the article has been authorized by the journal editor to be included in this PhD thesis as a compendium of publications.

*Sensors* **2012**, *12*, 14004-14021; doi:10.3390/s121014004

OPEN ACCESS

**sensors**

**ISSN 1424-8220**
www.mdpi.com/journal/sensors

*Article*

# Temperature and Relative Humidity Estimation and Prediction in the Tobacco Drying Process Using Artificial Neural Networks

**Víctor Martínez-Martínez [1,*], Carlos Baladrón [1], Jaime Gomez-Gil [1], Gonzalo Ruiz-Ruiz [2], Luis M. Navas-Gracia [2], Javier M. Aguiar [1] and Belén Carro [1]**

[1] Department of Signal Theory, Communications and Telematics Engineering, University of Valladolid, 47011 Valladolid, Spain; E-Mails: cbalzor@ribera.tel.uva.es (C.B.); jgomez@tel.uva.es (J.G.-G.); javagu@tel.uva.es (J.M.A.); belcar@tel.uva.es (B.C.)

[2] Department of Agricultural and Forestry Engineering, University of Valladolid, 34004 Palencia, Spain; E-Mails: gruiz@iaf.uva.es (G.R.-R.); lmnavas@iaf.uva.es (L.M.N.-G.)

* Author to whom correspondence should be addressed; E-Mail: vmarmar@ribera.tel.uva.es; Tel.: +34-636-797-528; Fax: +34-983-423-667.

*Received: 2 August 2012; in revised form: 28 September 2012 / Accepted: 6 October 2012 / Published: 17 October 2012*

**Abstract:** This paper presents a system based on an Artificial Neural Network (ANN) for estimating and predicting environmental variables related to tobacco drying processes. This system has been validated with temperature and relative humidity data obtained from a real tobacco dryer with a Wireless Sensor Network (WSN). A fitting ANN was used to estimate temperature and relative humidity in different locations inside the tobacco dryer and to predict them with different time horizons. An error under 2% can be achieved when estimating temperature as a function of temperature and relative humidity in other locations. Moreover, an error around 1.5 times lower than that obtained with an interpolation method can be achieved when predicting the temperature inside the tobacco mass as a function of its present and past values with time horizons over 150 minutes. These results show that the tobacco drying process can be improved taking into account the predicted future value of the monitored variables and the estimated actual value of other variables using a fitting ANN as proposed.

**Keywords:** estimation; prediction; Artificial Neural Networks (ANN); tobacco drying process; signal processing

## 1. Introduction

Tobacco farming is one of the main economic activities in agricultural areas of countries such as Brazil, China, and India [1]. During the last decades, a large number of technological innovations in this specific crop have been developed in diverse fields such as machinery, cultivation techniques, and drying techniques. Some advances in tobacco cultivation are the improvement of the health and quality of the tobacco plants by means of genetic modifications [2], the utilization of more effective fertilizers as a consequence of studies of their influence in tobacco growing and quality [3], or the employment of new machinery and techniques to irrigate and collect tobacco plants [4]. Furthermore, the tobacco drying process has been studied and improved by employing new drying strategies and new machinery in order to increase the quality of the dry tobacco and the overall energy efficiency of the process [5].

The tobacco drying process requires exhaustive monitoring of its representative variables to optimize it [6]. Monitoring systems are usually employed to acquire and store the most useful variables of a process. In the literature, these systems have been proposed to monitor atmospheric phenomena such as air pollution [7], livestock farms such as aviary production systems [8], or environments including rivers [9], oceans [10], and ecological systems [11]. In the field of agriculture, monitoring systems have been proposed for indoor locations such as greenhouses [12–14], and for outdoor agricultural processes including wine harvesting [15], olive growing [16], or any generic process [17].

Data stored by monitoring systems have been used to analyse the connection among the system's variables and to obtain a model of the monitored process [18]. These models can be defined by means of fuzzy-logic, probabilistic estimators, or Artificial Neural Networks (ANN) among others. Using fuzzy-logic, Schulz *et al.* [19] modelled water flow, Barros *et al.* [20] analysed different demographic and environmental models, Malins *et al.* [21] modelled the spatial extent of salinity on a farming land, and Papantoniou *et al.* [22] proposed a model of wind power for farm applications. Using probabilistic estimators, Cornford *et al.* [23] and Ng *et al.* [24] proposed different models to forecast precipitations, Holland *et al.* [25] modelled animal populations, and Trombe *et al.* [26] modelled offshore wind power fluctuations. Using ANN, Recknagel *et al.* [27] modelled algal growing, Smith *et al.* [28] predicted the air temperature, and Singh *et al.* [29] modelled water quality.

The aforementioned processing data tools have also been employed to control and predict variables in agricultural drying processes. Using fuzzy-logic, Yliniemi *et al.* [30–32] designed the control logic for a rotary dryer and Li *et al.* [33–35] developed and tested a control algorithm for microwave drying of vegetables. Employing probabilistic estimators, Banga *et al.* [36] designed a stochastic control algorithm using temperature and relative humidity as input variables. Using ANN, Yliniemi *et al.* [30,31] designed a system to identify the dynamics of the drying process in a rotary dryer, and Movagharnejad *et al.* [37] proposed a tomato drying model which fitted the experimental data more accurately than other mathematical models.

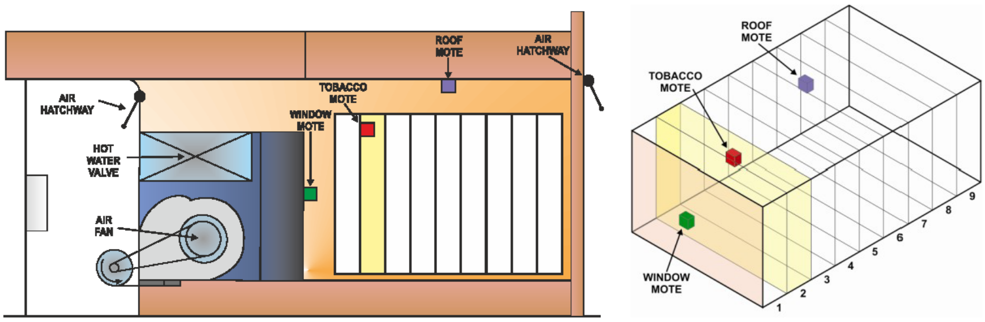
A system to estimate and predict the temperature and relative humidity in the tobacco drying process is proposed and tested in this work. This system acquires temperature and relative humidity data in different points of a tobacco dryer, and processes this data with an ANN in order to estimate and predict their value in other locations or in the future. The methodology of the article was composed by four stages: (i) a wireless sensor network (WSN)-based monitoring system was deployed in the drying installation to capture input variables, (ii) an artificial neural network was developed to estimate output variables using input variables, (iii) different experiments were performed to analyse the proposed methods, and (iv) experimental results were analysed to evaluate the proposed method.

## 2. Tobacco Drying

### *2.1. Materials*

A tobacco dryer from the Agrotex Company located in Rosalejo (Cáceres, Spain), was employed in this research. This tobacco dryer has a hot water valve to heat the air, a fan to move the air, and two air hatchways to remove the humidity. In this dryer, three measurement motes were deployed in its drying chamber to acquire the temperature and relative humidity in different points of the dryer. Figure 1(a) shows the schema of the cross section of this tobacco dryer, where the aforementioned equipment is represented. Figure 1(b) shows the spatial distribution of the three motes deployed inside the drying chamber: the first one was placed next to the supervision window (green), the second one was placed next to the roof (purple), and the third one was placed into the tobacco mass (red). The sensors of these measurement motes have been numbered as Sensors 1, 2, and 3, respectively, for easy reference.

**Figure 1.** (**a**) Cross section of the tobacco dryer, where an air fan, a water valve, and two air hatchways are represented in their real locations. (**b**) Spatial distribution of the measurement motes inside the drying chamber: next to the supervision window (green), in the second container in the tobacco mass (red), and next to the dryer roof (purple).




(**a**) (**b**)

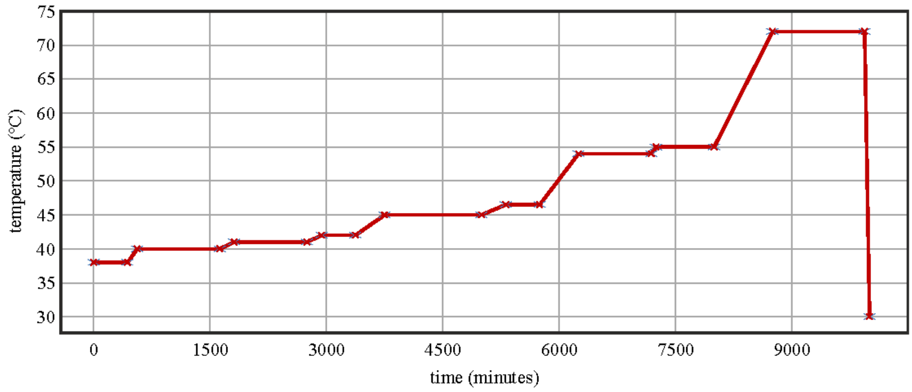
A WSN with MEP510 motes from the Crossbow Company was employed to take the measurements. This mote has a SHT11-Digital Humidity Sensor from the Sensirion Company to acquire both temperature and relative humidity. LabVIEW and Matlab language programming tools were used to develop applications to acquire data from the WSN, to store it in a structured database, and to develop an ANN to estimate and predict these data.

### *2.2. Description of the Tobacco Drying Process*

The specific tobacco drying process monitored in this research was based on increasing the tobacco dryer chamber temperature in order to extract the plants' moisture. To this end, a time-varying target temperature as that shown in Figure 2, which is a typical target temperature evolution in this type of dryer [5], was chosen.

The drying process was directed by a general purpose microcontroller, which used a temperature sensor to acquire the instantaneous temperature next to the window measurement mote [Figure 1(b)] and set the output values of the heater according to an ON/OFF algorithm: if the dryer temperature is greater than the target temperature, the heater is turned off, and, if the dryer temperature is lower than the target temperature, the heater is turned on. Moreover, a hysteresis of 2 °C was employed in the control algorithm in order to avoid activating and deactivating the heater repeatedly. The target temperature evolution consisted of a concatenation of two types of phases: phases with a constant temperature, where the target temperature is invariable, and phases with a constant slope increasing temperature, where the target temperature increases progressively. These two types of phases are represented in the target temperature evolution shown in Figure 2.

**Figure 2.** Target temperature evolution in the analysed drying processes.




*2.3. Data Set*

Nearly 900 hours (to be precise 53833 minutes) of drying process data were acquired. The temperature in each drying process varied between 20 °C and 75 °C, and the relative humidity range fluctuated between 10% and 100%. These data were acquired in the 2011 tobacco drying campaign, in the tobacco dryer presented in the Tobacco Drying section. An acquisition sample time of 3 minutes was used for every variable in all the drying processes, resulting in a total of 37 days containing 216,828 samples.

## 3. Artificial Neural Network for Data Completion

*3.1. Aims and Purpose*

The aim of this section was to design an ANN-based system for the estimation and prediction of temperature and relative humidity at different spatial points and with different time horizons inside a tobacco dryer, from the data sensed at a set of different meaningful points. This means that the system must carry out two different tasks:

- *Data Estimation*: from the readings of temperature and relative humidity obtained at a given moment *t* from a set of available sensors located at fixed positions inside the tobacco dryer (*input points*), the ANN must estimate the values of those magnitudes at a set of different spatial points (*target points*) at the same moment *t*. For that, the network was trained with the values of a single drying process for which the magnitudes at the target points are also known, but, once trained, it operates only with the sensors located at the input points.

- *Data Prediction*: from the readings of temperature and relative humidity obtained from a set of available sensors located at fixed positions inside the tobacco dryer at a set of *n* instants (*t-1, t-2, t-3, …, t-n)*, the ANN must predict the value of those magnitudes at the same spatial points at the moment $t + t_0$. For that, the network was trained with the values of a single drying process. Once trained, it operates predicting the temperature and/or relative humidity of a process in $t+t_0$ using as input the temperature and relative humidity data at (*t-1, …, t-n)* interval.

### *3.2. Artificial Neural Network Design*

Normally, when the data to be handled within an ANN is dependent on time, a straightforward solution is to use a time-series ANN [38,39], that is, a neural network with a delay architecture specifically designed to operate with an input which is a temporal sequence and capable of forecasting values of that sequence in the future.

However, in the case of this work, while the data to be handled is also a time series of temperature and relative humidity measures, another possible approach is to consider the drying procedure as a prototype pattern that will be replicated every other drying procedure. Using this approach, estimation and prediction are not made on the basis of the similarity and progression of values registered in moments near in time, but on the basis of the similarity of the input values with a canonical drying procedure pattern used for training. A fitting ANN is a good solution for implementing this approach. It is trained with the data available for one complete drying procedure and tested against the rest of the data.

Additionally, using a fitting network has the advantage of being more robust against failures in the sensors than a time-series ANN, especially for the *Data Estimation* task: if there is a failure in a sensor, in the case of the time-series network, all the predictions given by the ANN using that measure as an input fail. For instance, if the ANN uses the past $n = 20$ values of the input variables for its operation, a single missing or failing value will invalidate 20 predictions of the network. However, if a fitting ANN is employed for *Data Estimation*, one faulty input results in only one faulty output. This feature is highly desirable in environments like the one considered in this work where it is not unusual for sensors or sensor networks to fail.

Therefore, along this work, a fitting neural network was employed. The architecture of these networks was comprised by three neuron layers: an input layer, a hidden layer, and an output layer. The size of the input and output layers was determined by the number of inputs/outputs defined for each case, and this changes along this work according to the number of inputs/outputs employed in the different tests. The size of the hidden layer normally depends on the complexity of the data to be fitted. For this work, this size was set up at 20.

### *3.3. Trial Design*

For each of the two tasks considered in this work, different tests were conducted in order to determine the performance of the ANN and find the influence of the different parameters in the accuracy of the output. For the *Data Estimation* task, the varying parameters considered were the following:

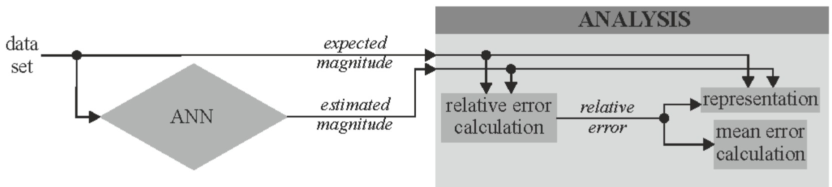
- Input magnitude: different tests were performed using only temperature, only humidity, or a combination of both as input data for the network.
- Number of inputs: different tests were performed using data from one sensor or combining data from two sensors as input.
- Output magnitude: two sets of tests were performed for *Data Estimation*, one of them estimating the temperature in one sensor, and the other estimating the humidity in one sensor.

For the *Data Prediction* task, the varying parameters considered are the following:

- Time horizon: different tests were performed to predict the values of a magnitude in different moments in the future: 15, 30, 150 and 300 minutes in advance.
- Length of input sequence: in order to perform the prediction, the length of the input sequence feed into the ANN was also varied.

For numerical characterization of the results, a validation stage of the system was conducted in which its operation data were measured. Validation trials were carried out according to the diagram shown in Figure 3.

**Figure 3.** Diagram of the validation trials.




The relative error between the estimation and the measured magnitude value was obtained by means of Equation (1). Since the expected temperature was always greater than 0 °C (in fact it was always greater than 15 °C) there was no problem with the fraction denominator. The same formula was also applied to the calculations made for humidity:

$$e[n](\%) = \frac{T_{expected}[n] - T_{estimated}[n]}{T_{expected}[n]} \cdot 100 \quad (1)$$

### *3.4. Results*

This section presents the numerical results of the test trials performed with the ANN. In all cases, the fitting ANN was trained with the data of a complete drying process and then tested against the rest of the data. With three sensors available and two magnitudes measured per sensor, data from Sensors 1 and 2 were always considered as potential inputs, and data from Sensor 3 was always considered as the target, that is, the data that the network predicts and estimates.

3.4.1. Data Estimation

Tables 1 and 2, and their respective graphical representations in Figures 4 and 5, depict the results obtained for temperature and humidity *Data Estimation* at Sensor 3. This means that along this set of tests, the ANN was estimating the temperature (or humidity) given by Sensor 3 at moment *t* by using as an input a combination of data from temperature and humidity given by Sensors 1 and/or 2, also at moment *t*. Each cell of the tables represents a specific combination of inputs, and for each combination of inputs, 20 different tests were performed. The mean errors of the output against the real data were measured and averaged.

**Table 1.** *Data Estimation* results (% error ANN/% error interpolation) for Temperature (T) in Sensor 3 (S3), combining Temperature (T) and Humidity (H) in Sensors 1 and 2 (S1 and S2) as input.

| Input T / Input H | No T used | T at S1 | T at S2 | T at S1 & S2 |
|---|---|---|---|---|
| No H used | | 2.30%/5.87% | 2.19%/3.99% | 2.12%/3.99% |
| H at S1 | 6.66%/NA | 1.96%/5.87% | 2.01%/3.99% | 1.89%/3.99% |
| H at S2 | 6.60%/NA | 2.45%/5.87% | 2.10%/3.99% | 2.03%/3.99% |
| H at S1 & S2 | 8.77%/NA | 2.23%/5.87% | 1.97%/3.99% | 1.80%/3.99% |

**Table 2.** Data Estimation results (% error ANN/% error interpolation) for Humidity (H) in Sensor 3 (S3), combining Temperature (T) and Humidity in Sensors 1 and 2 (S1 and S2) as input.

| Input T / Input H | No T used | T at S1 | T at S2 | T at S1 & S2 |
|---|---|---|---|---|
| No H used | | 6.48%/NA | 5.92%/NA | 5.78%/NA |
| H at S1 | 7.30%/33.43% | 4.95%/33.43% | 5.66%/33.43% | 4.46%/33.43% |
| H at S2 | 5.36%/28.72% | 5.19%/28.72% | 5.59%/28.72% | 5.32%/28.72% |
| H at S1 & S2 | 6.97%/28.72% | 4.80%/28.72% | 6.51%/28.72% | 4.98%/28.72% |

Additionally, for the cases when the input included a measure of the same magnitude to be estimated, a direct interpolation method was also implemented, in which the output (estimated value) was directly calculated as the input value of the same magnitude in a different sensor. This method assumed that points located in the vicinity of a sensor have similar measures for the same magnitude. The error of this interpolation method was also computed as a means to measure, by comparison, the improvement offered by the ANN estimation. In each of the cells of Tables 1 and 2, the leftmost value is the averaged mean error of the ANN's output, and the rightmost value is the averaged mean error of the simple interpolation method.

**Figure 4.** *Data Estimation* results (% error ANN/% error interpolation) for Temperature (T) in Sensor 3 (S3), combining Temperature (T) and Humidity (H) in Sensors 1 and 2 (S1 and S2) as inputs.

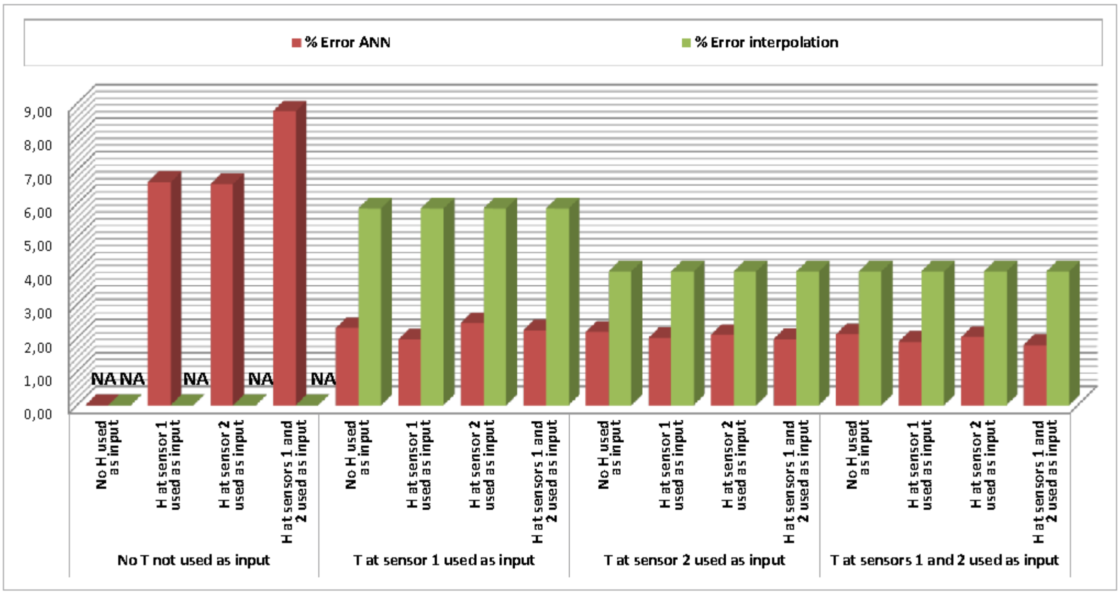
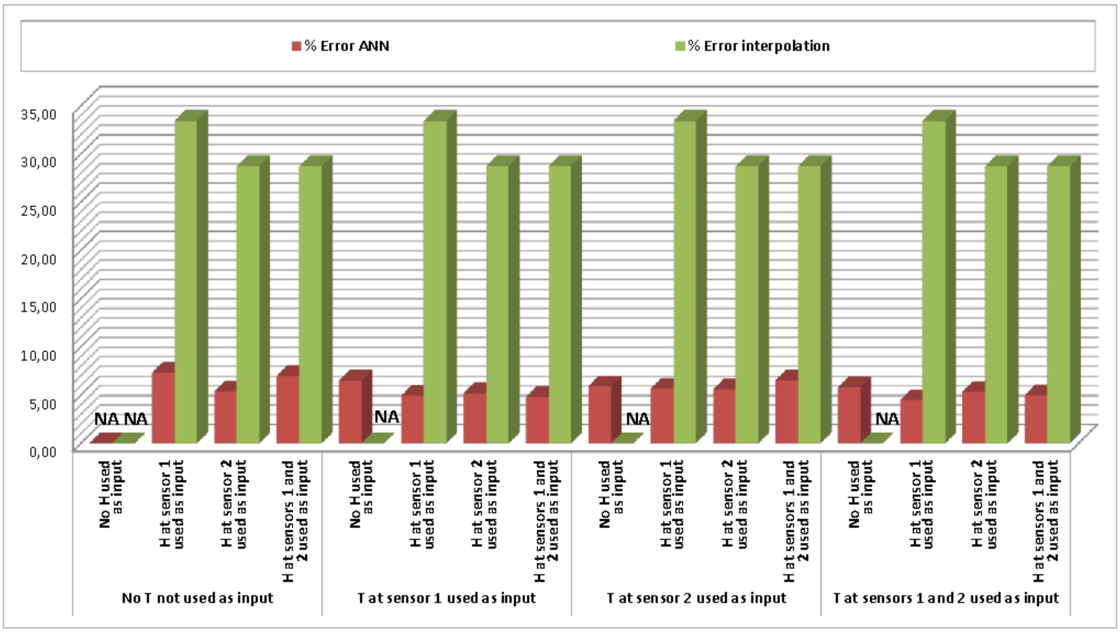



**Figure 5.** Data Estimation results (% error ANN/% error interpolation) for Humidity (H) in Sensor 3 (S3), combining Temperature (T) and Humidity (H) in Sensors 1 and 2 (S1 and S2) as input.



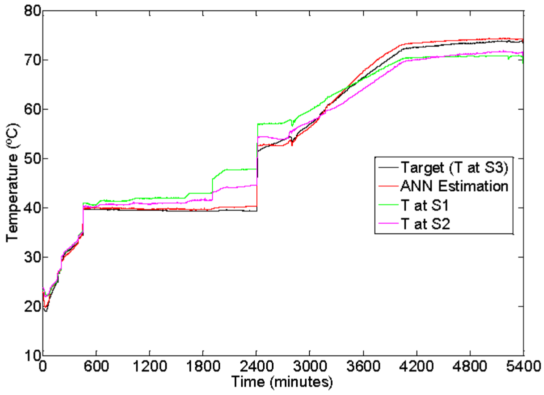
Looking at the results, it is easy to realize that when estimating temperature (Table 1, Figure 4), the ANN is roughly between two and three times more accurate (200%–300% improvement) than the simple interpolation method.

Another aspect to be noticed is that the results are not very accurate (mean errors around 6–9%) when only using humidity as an input (leftmost column of the table), but they demonstrate that a rough approximation to temperature can be made by the ANN when only humidity is available. However, even if humidity shows itself not to be sufficient to estimate temperature accurately, it contains information to improve the performance of the estimation when combined with temperature: the upper row shows results of cases when only temperature was considered as an input, and the lower rows, that combine temperature with humidity, generally show better results.

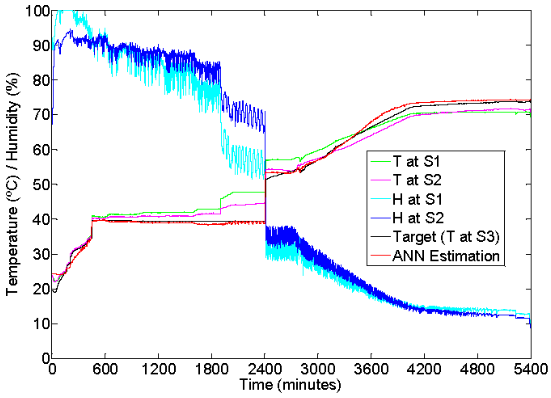
Looking at Table 2 and Figure 5 it is obvious that it is more difficult to estimate humidity than temperature. The best errors are around 5%, doubling those obtained for estimation of temperature. However, it is also obvious that an acceptable estimation of humidity (mean error of 5–6%) can be obtained only on the basis of temperature (upper row), and that the usage of the ANN in this case greatly improves the estimation based on interpolation, showing errors around six times smaller (errors around 5% against errors around 30%). Combined inputs of humidity and temperature render again the best results.

As an example, Figure 6 shows the results of a specific execution of the ANN estimation for temperature given by Sensor 3 during one drying session, using as input: (left) temperature at Sensors 1 and 2 and (right) temperature and relative humidity in Sensors 1 and 2. Figure 7 shows another specific execution of the ANN estimation for humidity, using as input (left) humidity at Sensors 1 and 2 and (right) temperature and humidity in Sensors 1 and 2. It is easy to see that for this specific session, there is a discontinuity in the data around 2,400 minutes as a consequence of a failure in the sensing hardware. However, the fitting ANN approach allows a successful estimation of values after the discontinuity, which is not the case when using a time-series based estimation.

**Figure 6.** Example of ANN estimation of Temperature (T) at Sensor 3 (S3) for one drying process using Sensors 1 and 2 (S1 and S2) as input. (**a**) Input is only Temperature (T). (**b**) Input combines Temperature (T) and Humidity (H).

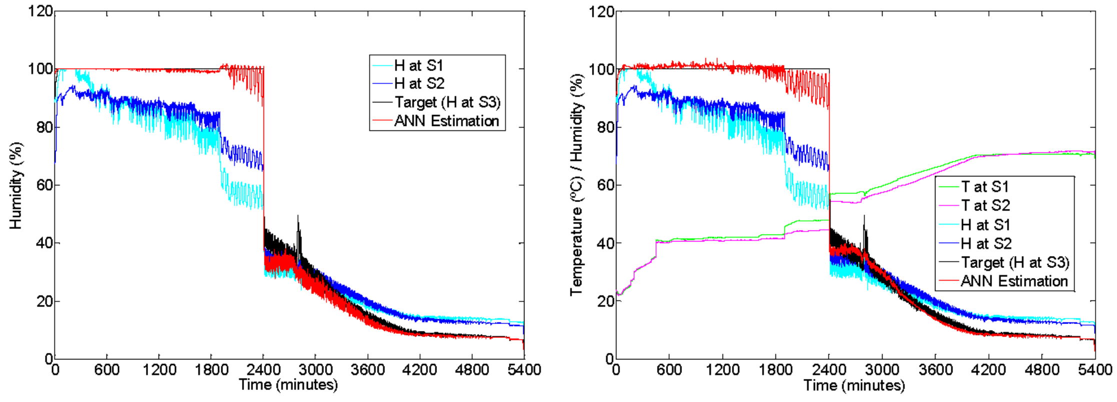



(**a**)




(**b**)

**Figure 7.** Example of ANN estimation of Humidity (H) at Sensor 3 (S3) for one drying process using Sensors 1 and 2 (S1 and S2) as input. (**a**) Input is only Humidity (H). (**b**) Input combines Temperature (T) and Humidity (H).




(**a**) (**b**)

3.4.2. Data Prediction

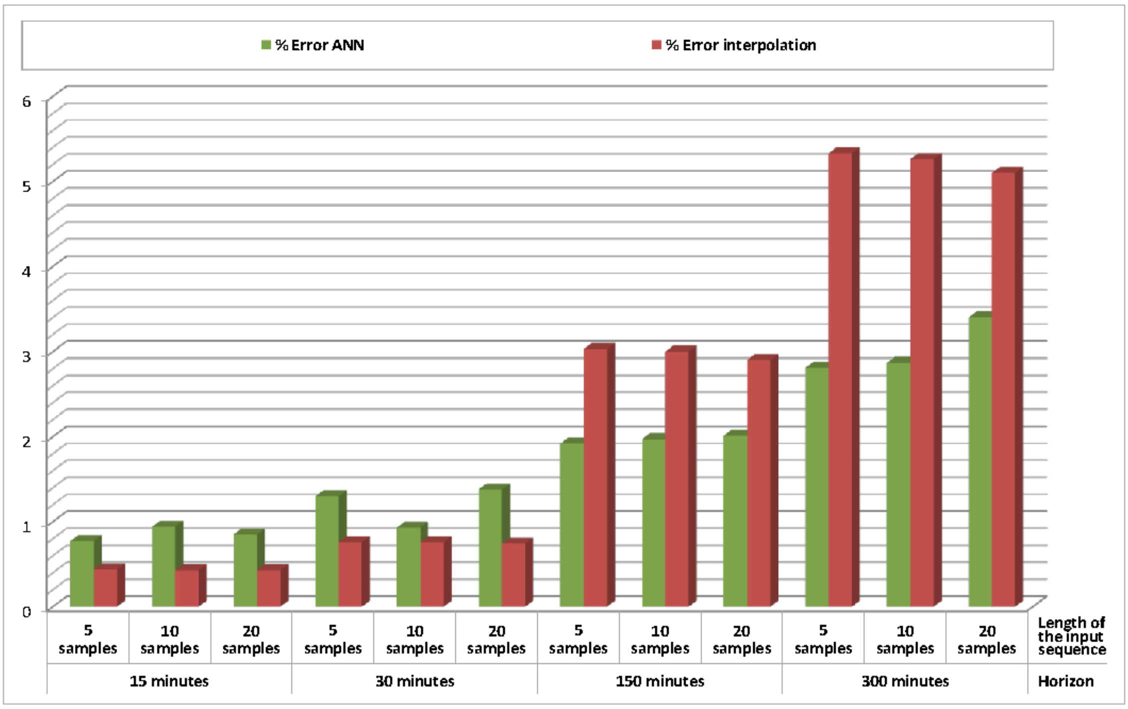
Table 3 and its respective graphical representation in Figure 8 present the results obtained for *Data Prediction* of temperature in Sensor 1, using just its past values as an input. Different lengths for the input sequence (which contains values from *t-1* to *t-length_of_sequence*) and different time horizons (the value to be predicted is *t* + *time_horizon*) were considered and combined. For each cell of the table, 20 different tests were performed and the mean error of the output was measured and averaged. Again, an interpolation method was also implemented as a reference, using the value of temperature at *t* as the predicted value for temperature at *t* + *time_horizon*.

An analysis of these results shows that when the time horizon is short, even if the ANN gives small errors, the simple interpolation method performs better, giving extremely small errors. This happens because temperature varies relatively slowly, and thus recent values are normally valid for a very good prediction. However, as the time horizon increases, interpolation performs much worse, and errors obtained by the ANN are smaller than those obtained by interpolation.

**Table 3.** *Data prediction* results (% error ANN/% error interpolation) for temperature in Sensor 1, using temperature in Sensor 1 as an input.

| Length of Input Sequence \ Horizon | 15 minutes | 30 minutes | 150 minutes | 300 minutes |
|---|---|---|---|---|
| **5 samples** | 0.77%/0.44% | 1.30%/0.76% | 1.92%/3.03% | 2.81%/5.33% |
| **10 samples** | 0.94%/0.43% | 0.93%/0.76% | 1.97%/3.00% | 2.87%/5.26% |
| **20 samples** | 0.85%/0.43% | 1.38%/0.74% | 2.01%/2.90% | 3.40%/5.10% |

**Figure 8.** *Data prediction* results (% error ANN/% error interpolation) for temperature in Sensor 1, using temperature in Sensor 1 as an input.



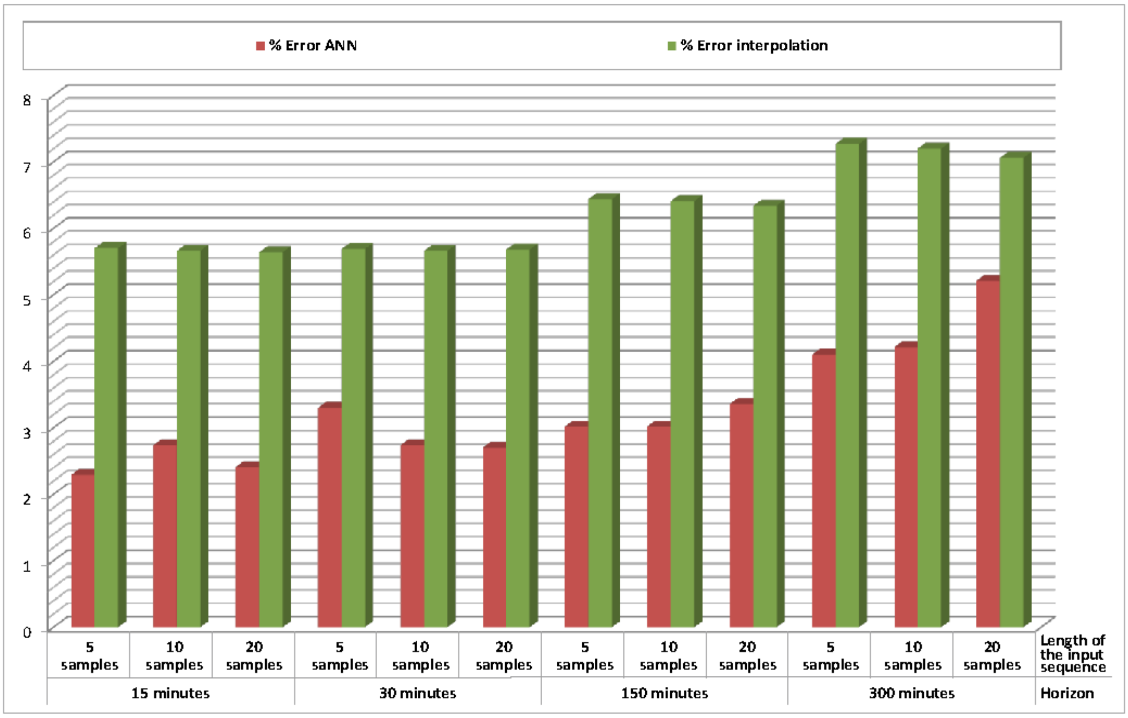

These numbers also reveal that increasing the length of the input sequence could distort the prediction, demonstrating that saturating the ANN with data is not always a guarantee for better results.

Finally, Table 4 and Figure 9 present the results for the combined operation of *Data Estimation* and *Data Prediction*. In this scenario, temperature in Sensor 3 at a moment *t+time_horizon* was predicted on the basis of temperature in Sensor 1. This test was done to show that it is possible to use the ANN to perform both tasks at the same time, showing again good results when compared to interpolation.

**Table 4.** *Data prediction* results (% error ANN/% error interpolation) for temperature in Sensor 3, using temperature in Sensor 1 as an input.

| Horizon / Length of input sequence | 5 samples | 10 samples | 50 samples | 100 samples |
|---|---|---|---|---|
| **5 samples** | 2.29%/5.70% | 3.30%/5.69% | 3.02%/6.44% | 4.10%/7.27% |
| **10 samples** | 2.74%/5.66% | 2.74%/5.66% | 3.02%/6.41% | 4.21%/7.20% |
| **20 samples** | 2.41%/5.64% | 2.70%/5.68% | 3.36%/6.34% | 5.21%/7.06% |

**Figure 9.** *Data prediction* results (% error ANN/% error interpolation) for temperature in Sensor 3, using temperature in Sensor 1 as an input.




## 4. Discussion

In light of the results, it is easy to realize that the use of the ANN results in a more efficient estimation and prediction task performance. For instance, for temperature estimation when there are no direct temperature readings in the system, the neural network offers a sufficient performance in the surroundings of 6–7% mean error, which is more than enough for an approximation; but even more, when there is temperature data, the ANN is capable of being between two and three times more precise than the interpolation method considered. When estimating the humidity, the results are even more meaningful, because even if the absolute mean error registered is around 5%, this error is between five and six times lower than the one obtained by interpolation.

The prediction task is slightly more complex, since, depending on the conditions, the performance of the ANN could greatly vary when compared to interpolation. Due to magnitudes varying slowly, it has to be considered that for small prediction horizons it might be better to directly interpolate the values. However, it has been shown that as the horizon increases, the neural network is capable of a more accurate prediction.

Estimations provided by the proposed system can be employed to detect sensors' failures and to recover their measurements when failures happen. Sensor failures can be detected comparing a threshold with the absolute difference between the estimated measurement obtained by the estimation method and the sensor measurement. If the absolute difference is higher than the threshold, then a failure has occurred. In this case, and until the sensor problem is solved, the measurements of this sensor can be replaced by the estimation provided by the ANN using the rest of the sensors measurements as input data. Moreover, this data recovering application can be used to recover data when sensors have other type of failures, such as sensor signal corruption or sensor signal loss, which are two typical problems in WSN [40,41].

Another application of the proposed estimation methods is the design optimization of data acquisition systems. To this end, a set of sensors must be deployed in several significant points of the monitored system. After that, the proposed estimation methods can be employed to discard the sensors whose measurements can be more accurately estimated using the rest of the deployed sensors with a low error. One advantage of these designed systems is their economy due to discarding the sensors whose measurements can be estimated with the other sensors of the system. Another advantage is that the measurements of sensors sited in a point with dangerous or hostile conditions that can quickly degrade the hardware can be replaced with the estimated measurements calculated from data of other sensors sited in safer places with better conditions.

Proposed estimation and prediction methods can be used to model processes. Models establish relationships among the variables of a process and predict the future evolution of these variables. They have been applied in the literature to plant growing processes [42–46], vegetal evapotranspiration processes [19,47,48], or drying processes [37,49,50] among others. One way to establish the relationships and to predict the evolution of processes' variables is using ANNs, which have been used to model algal growing [46] or drying processes as in the carrot drying process [51] and in the tomato drying process [37]. Process models are usually developed as a part of model-based predictive control (MPC) algorithms. MPC are multivariable control algorithms which use prediction methods to estimate the future state of the system and an optimization cost function in order to calculate the optimum control signals taking into account the predicted future state [52–63]. ANN-based prediction methods as that proposed in this article can be used in a MPC algorithm as the prediction method. ANN-based MPC algorithms have been proposed to control industrial applications [59–61], aircraft navigation [62], and laboratory processes [63].

Another application of the proposed data prediction methods is the improvement of alarm systems' performance. A typical problem in alarm systems is the high rate of false positives and false negatives. One cause of this problem is that alarm systems only take into account present data, or past and present data, so the incorporation of predicted future information about the system can improve the precision of the alarm system [64].

Finally, the proposed system's predictions can be easily employed to anticipate decisions about the drying process. The operator can see a forecast of the variables of the tobacco dryer in the future on the basis of the present conditions and also under a series of alternative conditions, making the process of taking the most appropriate decision to reach the desired values in the future much easier.

## 5. Conclusions

Tobacco drying has been traditionally a manually controlled activity, relying solely on experts' craftsmanship, and has been considered almost an art; however, as in many other areas of agriculture, new technologies have appeared as an easy solution to greatly improve the results and performance of the production processes while reducing costs at the same time. Specifically, introduction of sensor networks and intelligent automation systems represents a strong trend, as shown in the state of the art of drying processes. The performance of these systems depends heavily on the accuracy of the data provided by the sensor network, and as such, the better the data provided, the better the results.

The work presented in this paper has shown that by using a fitting ANN it is possible to accurately estimate the values of the different environmental magnitudes employed to control the drying process at different spatial points. The numerical results obtained present very low errors, which at less than 2% could even be below some sensors' accuracy.

Additionally, the system proposed is also capable of predicting the values of those environmental magnitudes with different time horizons. In this case, the results are dependent on the specific time horizon considered. For nearby time horizons, the mean errors given by interpolation are similar to those of the ANN system; however, for distant time horizons, the mean errors given by the ANN are much lower than those obtained by interpolation.

Finally, it is worth mentioning that both capabilities of the presented system are potentially very useful for improving the performance of automatic control of tobacco drying processes while at the same time reducing costs, by increasing the quality and quantity of available data.

**Acknowledgments**

This work was supported partially by two projects. The first one, entitled *Mejora de la competitividad del sector del tabaco en Extremadura: nuevos procesos y productos*, was financed by the Centre for Industrial Technological Development (CDTI) under project IDI-20100986. The second one was financed by the regional 2010 Research Project Plan of the *Junta de Castilla y León* (Spain), under project VA034A10-2.

The first author's work was also made possible thanks to a *Formación de Personal Investigador* program grant. This program was financed by the *Universidad de Valladolid* (Spain) and co-financed by the *Banco de Santander*. Moreover, the experiments were partially supported by another grant obtained by the first author in the *¿Investigamos? 2011* competition, by means of a project entitled *Desarrollo de un sistema SCADA para el control del curado del tabaco*. The competition was organized by the *Centro Tecnológico Agrario y Agroalimentario ITAGRA*.

**References**

1. Food Agriculture Organization of the United Nations. Commodities Trade Division Agriculture Organization of the United Nations. Raw Materials, Tropical Horticultural Products Service. *Issues in the Global Tobacco Economy: Selected Case Studies*; Raw Materials, Tropical and Horicultural [sic] Products Service, Commodities and Trade Division, Food and Agricultural Organization of the United Nations: Rome, Italy, 2003.
2. Hopkins, D.W.; Webster, E.A.; Chudek, J.A.; Halpin, C. Decomposition in soil of tobacco plants with genetic modifications to lignin biosynthesis. *Soil Biol. Biochem.* **2001**, *33*, 1455–1462.
3. Karaivazoglou, N.A.; Tsotsolis, N.C.; Tsadilas, C.D. Influence of liming and form of nitrogen fertilizer on nutrient uptake, growth, yield, and quality of Virginia (flue-cured) tobacco. *Field Crop. Res.* **2007**, *100*, 52–60.

4. Cakir, R.; Cebi, U. The effect of irrigation scheduling and water stress on the maturity and chemical composition of Virginia tobacco leaf. *Field Crop. Res.* **2010**, *119*, 269–276.
5. Hawks, S.N.; Collins, W.K.; Carbonell, H.P.; Martínez, E.C. *Tabaco Flue-Cured: Principios Básicos de su Cultivo y Curado*; Ministerio de Agricultura, Pesca y Alimentación: Madrid, Spain, 1986.
6. Frost, A.R.; Schofield, C.P.; Beaulah, S.A.; Mottram, T.T.; Lines, J.A.; Wathes, C.M. A review of livestock monitoring and the need for integrated systems. *Comput. Electron. Agric.* **1997**, *17*, 139–159.
7. Jung, Y.J.; Lee, Y.K.; Lee, D.G.; Lee, Y.; Nittel, S.; Beard, K.; Nam, K.W.; Ryu, K.H. Design of Sensor Data Processing Steps in an Air Pollution Monitoring System. *Sensors* **2011**, *11*, 11235–11250.
8. Lokhorst, C.; Lamaker, E.J.J. An expert system for monitoring the daily production process in aviary systems for laying hens. *Comput. Electron. Agric.* **1996**, *15*, 215–231.
9. Park, S.-Y.; Choi, J.H.; Wang, S.; Park, S.S. Design of a water quality monitoring network in a large river system using the genetic algorithm. *Ecol. Model.* **2006**, *199*, 289–297.
10. Baladrón, C.; Aguiar, J.M.; Calavia, L.; Carro, B.; Sánchez-Esguevillas, A.; Hernández, L. Performance study of the application of artificial neural networks to the completion and prediction of data retrieved by underwater sensors. *Sensors* **2012**, *12*, 1468–1481.
11. Suri, A.; Iyengar, S.S.; Cho, E. Ecoinformatics using wireless sensor networks: An overview. *Ecol. Inform.* **2006**, *1*, 287–293.
12. Pawlowski, A.; Guzman, J.; Rodríguez, F.; Berenguel, M.; Sánchez, J.; Dormido, S. Simulation of greenhouse climate monitoring and control with wireless sensor network and event-based control. *Sensors* **2009**, *9*, 232–252.
13. Park, D.-H.; Park, J.-W. Wireless sensor network-based greenhouse environment monitoring and automatic control system for dew condensation prevention. *Sensors* **2011**, *11*, 3640–3651.
14. Li, X.-H.; Cheng, X.; Yan, K.; Gong, P. A Monitoring system for vegetable greenhouses based on a wireless sensor network. *Sensors* **2010**, *10*, 8963–8980.
15. Cirilli, M.; Bellincontro, A.; De Santis, D.; Botondi, R.; Colao, M.C.; Muleo, R.; Mencarelli, F. Temperature and water loss affect ADH activity and gene expression in grape berry during postharvest dehydration. *Food Chem.* **2012**, *132*, 447–454.
16. Costantini, E.A.C.; Pellegrini, S.; Vignozzi, N.; Barbetti, R. Micromorphological characterization and monitoring of internal drainage in soils of vineyards and olive groves in central Italy. *Geoderma* **2006**, *131*, 388–403.
17. Hwang, J.; Shin, C.; Yoe, H. Study on an agricultural environment monitoring server system using wireless sensor networks. *Sensors* **2010**, *10*, 11189–11211.
18. Yue, T.-X.; Jorgensen, S.E.; Larocque, G.R. Progress in global ecological modelling. *Ecol. Model.* **2011**, *222*, 2172–2177.
19. Schulz, K.; Huwe, B. Water flow modeling in the unsaturated zone with imprecise parameters using a fuzzy approach. *J. Hydrol.* **1997**, *201*, 211–229.
20. Barros, L.C.; Bassanezi, R.C.; Tonelli, P.A. Fuzzy modelling in population dynamics. *Ecol. Model.* **2000**, *128*, 27–33.

21. Malins, D.; Metternicht, G. Assessing the spatial extent of dryland salinity through fuzzy modeling. *Ecol. Model.* **2006**, *193*, 387–411.
22. Papantoniou, A.; Coonick, A. Fuzzy Logic Control of a Unified Power Flow Controller for Wind Farm Applications. In *Proceedings of IEE Colloquium on Power Electronics for Renewable Energy (Digest No: 1997/170)*, London, UK, 16 June 1997; pp. 1–6.
23. Cornford, D. A Bayesian state space modelling approach to probabilistic quantitative precipitation forecasting. *J. Hydrol.* **2004**, *288*, 92–104.
24. Ng, W.W.; Panu, U.S. Comparisons of traditional and novel stochastic models for the generation of daily precipitation occurrences. *J. Hydrol.* **2010**, *380*, 222–236.
25. Holland, E.P.; Burrow, J.F.; Dytham, C.; Aegerter, J.N. Modelling with uncertainty: Introducing a probabilistic framework to predict animal population dynamics. *Ecol. Model.* **2009**, *220*, 1203–1217.
26. Trombe, P.-J.; Pinson, P.; Madsen, H. A General probabilistic forecasting framework for offshore wind power fluctuations. *Energies* **2012**, *5*, 621–657.
27. Recknagel, F.; French, M.; Harkonen, P.; Yabunaka, K.-I. Artificial neural network approach for modelling and prediction of algal blooms. *Ecol. Model.* **1997**, *96*, 11–28.
28. Smith, B.A.; Hoogenboom, G.; McClendon, R.W. Artificial neural networks for automated year-round temperature prediction. *Comput. Electron. Agric.* **2009**, *68*, 52–61.
29. Singh, K.P.; Basant, A.; Malik, A.; Jain, G. Artificial neural network modeling of the river water quality—A case study. *Ecol. Model.* **2009**, *220*, 888–895.
30. Yliniemi, L. *Advanced Control of a Rotary Dryer*; University of Oulu: Oulu, Finland, 1999.
31. Yliniemi, L.; Koskinen, J.; Leiviska, K. *Advanced Control of a Rotary Dryer*; Pergamon Press Ltd: Oxford, UK, 1999.
32. Koskinen, J.; Yliniemi, L.; Leiviska, K. *Fuzzy Modelling of a Pilot Plant Rotary Dryer*; Inst Electrical Engineers Inspec Inc: Edison, NJ, USA, 1998.
33. Li, Z.; Raghavan, G.S.V.; Wang, N.; Gariepy, Y. Real-time, volatile-detection-assisted control for microwave drying. *Comput. Electron. Agric.* **2009**, *69*, 177–184.
34. Li, Z.; Raghavan, G.S.V.; Wang, N. Carrot volatiles monitoring and control in microwave drying. *LWT - Food Sci. Technol.* **2010**, *43*, 291–297.
35. Li, Z.; Vijaya Raghavan, G.S.; Wang, N. Apple volatiles monitoring and control in microwave drying. *LWT - Food Sci. Technol.* **2010**, *43*, 684–689.
36. Banga, J.R.; Paul Singh, R. Optimization of air drying of foods. *J. Food Eng.* **1994**, *23*, 189–211.
37. Movagharnejad, K.; Nikzad, M. Modeling of tomato drying using artificial neural network. *Comput. Electron. Agric.* **2007**, *59*, 78–85.
38. Hill, T.; O'Connor, M.; Remus, W. Neural network models for time series forecasts. *Manag. Sci.* **1996**, *42*, 1082–1092.
39. Dorffner, G. Neural networks for time series processing. *Neural Netw. World* **1996**, *6*, 447–468.
40. Kwong, K.H.; Wu, T.-T.; Goh, H.G.; Sasloglou, K.; Stephen, B.; Glover, I.; Shen, C.; Du, W.; Michie, C.; Andonovic, I. Practical considerations for wireless sensor networks in cattle monitoring applications. *Comput. Electron. Agric.* **2012**, *81*, 33–44.
41. Tamboli, N.; Younis, M. Coverage-aware connectivity restoration in mobile sensor networks. *J. Netw. Comput. Appl.* **2010**, *33*, 363–374.

42. Yang, H.S.; Dobermann, A.; Lindquist, J.L.; Walters, D.T.; Arkebauer, T.J.; Cassman, K.G. Hybrid-maize—A maize simulation model that combines two crop modeling approaches. *Field Crop. Res.* **2004**, *87*, 131–154.
43. Subedi, N.; Sharma, M. Individual-tree diameter growth models for black spruce and jack pine plantations in northern Ontario. *For. Ecol. Manag.* **2011**, *261*, 2140–2148.
44. Lhotka, J.M.; Loewenstein, E.F. An individual-tree diameter growth model for managed uneven-aged oak-shortleaf pine stands in the Ozark Highlands of Missouri, USA. *For. Ecol. Manag.* **2011**, *261*, 770–778.
45. Noy-Meir, I. Rotational grazing in a continuously growing pasture: A simple model. *Agric. Syst.* **1976**, *1*, 87–112.
46. Millie, D.F.; Weckman, G.R.; Young Ii, W.A.; Ivey, J.E.; Carrick, H.J.; Fahnenstiel, G.L. Modeling microalgal abundance with artificial neural networks: Demonstration of a heuristic 'Grey-Box' to deconvolve and quantify environmental influences. *Environ. Model. Softw.* **2012**, *38*, 27–39.
47. Li, X.; Yang, P.; Ren, S.; Li, Y.; Liu, H.; Du, J.; Li, P.; Wang, C.; Ren, L. Modeling cherry orchard evapotranspiration based on an improved dual-source model. *Agric. Water Manag.* **2010**, *98*, 12–18.
48. Villagarcía, L.; Were, A.; García, M.; Domingo, F. Sensitivity of a clumped model of evapotranspiration to surface resistance parameterisations: Application in a semi-arid environment. *Agric. For. Meteorol.* **2010**, *150*, 1065–1078.
49. Toğrul, H. Suitable drying model for infrared drying of carrot. *J. Food Eng.* **2006**, *77*, 610–619.
50. Babalis, S.J.; Papanicolaou, E.; Kyriakis, N.; Belessiotis, V.G. Evaluation of thin-layer drying models for describing drying kinetics of figs (Ficus carica). *J. Food Eng.* **2006**, *75*, 205–214.
51. Erenturk, S.; Erenturk, K. Comparison of genetic algorithm and neural network approaches for the drying process of carrot. *J. Food Eng.* **2007**, *78*, 905–912.
52. Blasco, X.; Martínez, M.; Herrero, J.M.; Ramos, C.; Sanchis, J. Model-based predictive control of greenhouse climate for reducing energy and water consumption. *Comput. Electron. Agric.* **2007**, *55*, 49–70.
53. Coelho, J.P.; de Moura Oliveira, P.B.; Cunha, J.B. Greenhouse air temperature predictive control using the particle swarm optimisation algorithm. *Comput. Electron. Agric.* **2005**, *49*, 330–344.
54. Piñón, S.; Camacho, E.F.; Kuchen, B.; Peña, M. Constrained predictive control of a greenhouse. *Comput. Electron. Agric.* **2005**, *49*, 317–329.
55. van Straten, G.; Challa, H.; Buwalda, F. Towards user accepted optimal control of greenhouse climate. *Comput. Electron. Agric.* **2000**, *26*, 221–238.
56. Grondin-Perez, B.; Benne, M.; Bonnecaze, C.; Chabriat, J.-P. Industrial multi-step forward predictor of mother liquor purity of the final stage of a cane sugar crystallisation plant. *J. Food Eng.* **2005**, *66*, 361–367.
57. Coen, T.; Anthonis, J.; De Baerdemaeker, J. Cruise control using model predictive control with constraints. *Comput. Electron. Agric.* **2008**, *63*, 227–236.
58. Coen, T.; Saeys, W.; Missotten, B.; De Baerdemaeker, J. Cruise control on a combine harvester using model-based predictive control. *Biosyst. Eng.* **2008**, *99*, 47–55.

59. Benne, M.; Grondin-Perez, B.; Chabriat, J.P.; Hervé, P. Artificial neural networks for modelling and predictive control of an industrial evaporation process. *J. Food Eng.* **2000**, *46*, 227–234.
60. Damour, C.; Benne, M.; Grondin-Perez, B.; Chabriat, J.-P. Nonlinear predictive control based on artificial neural network model for industrial crystallization. *J. Food Eng.* **2010**, *99*, 225–231.
61. Al Seyab, R.K.; Cao, Y. Nonlinear system identification for predictive control using continuous time recurrent neural networks and automatic differentiation. *J. Process Control* **2008**, *18*, 568–581.
62. Akpan, V.A.; Hassapis, G.D. Nonlinear model identification and adaptive model predictive control using neural networks. *ISA Trans.* **2011**, *50*, 177–194.
63. Song, Y.; Chen, Z.; Yuan, Z. Neural network nonlinear predictive control based on tent-map chaos optimization. *Chin. J. Chem. Eng.* **2007**, *15*, 539–544.
64. Zhou, H.F.; Ni, Y.Q.; Ko, J.M. Structural damage alarming using auto-associative neural network technique: Exploration of environment-tolerant capacity and setup of alarming threshold. *Mech. Syst. Signal Proc.* **2011**, *25*, 1508–1526.

# Article 2: Switchgrass drying process

The second article presents the **switchgrass drying model** described in the “Introduction and Summary” part of the document. The main bibliographic data about this article are presented below:

- Title: Moisture Content Prediction in the Switchgrass (*Panicum Virgatum*) Drying Process Using Artificial Neural Networks.
- Authors: Martínez-Martínez V, Gomez-Gil J, Stombaugh TS, Montross MD, Aguiar JM.
- Journal: Drying Technology.
- Editor: Taylor & Francis.
- Impact factor: 1.518 (2014).
- Journal Ranking:
  - Q2 on *Engineering, Mechanical* (36/130).
  - Q2 on *Engineering, Chemical* (62/134).
- Date of publication: February 3rd, 2015.
- ISSN: 0737-3937.
- Volume (Issue): 33 (14).
- Pages: 1708-1719.
- DOI: 10.1080/07373937.2015.1005228.
- URL: http://www.tandfonline.com/eprint/6C99MTtGguMuwxREwR9u/full.
- Number of cites: 1.

The article presented below is a version of the original article published in the previously mentioned journal. This version of the article has been authorized by the journal editor to be included in this PhD thesis as a compendium of publications.

Drying Technology: An International Journal

# Moisture Content Prediction in the Switchgrass (Panicum virgatum) Drying Process Using Artificial Neural Networks

Víctor Martínez-Martínez[a], Jaime Gomez-Gil[a], Timothy S. Stombaugh[b], Michael D. Montross[b] & Javier M. Aguiar[a].

a Department of Signal Theory, Communications and Telematics Engineering, University of Valladolid, Valladolid, Spain

b Biosystems and Agricultural Engineering Department, University of Kentucky, Lexington, Kentucky, USA

Accepted author version posted online: 03 Feb 2015.Published online: 03 Feb 2015.

**Abstract:** This article proposes two artificial neural network (ANN)-based models to characterize the switchgrass drying process: The first one models processes with constant air temperature and relative humidity and the second one models processes with variable air conditions and rainfall. The two ANN-based models proposed estimated the moisture content (MC) as a function of temperature, relative humidity, previous MC, time, and precipitation information. The first ANN-based model describes MC evolution data more accurately than six mathematical empirical equations typically proposed in the literature. The second ANN-based model estimated the MC with a correlation coefficient greater than 98.8%.

**Keywords:** Artificial neural networks; Drying; Modeling; Prediction; Switchgrass.

## INTRODUCTION

Switchgrass is the common name of *Panicum virgatum*, a warm-season perennial bunchgrass native to North America. Switchgrass has been studied for use as fodder production during summer months because it is more productive than cool-season grasses [1]. Moreover, switchgrass has been selected as a model herbaceous energy crop by the Bioenergy Feedstock Development Program of the U.S. Department of Energy because this crop combines excellent conservation attributes, and good compatibility with conventional farming practices [2]. Compared with other bioenergy crops, switchgrass has a high productivity of energy, low nutrient requirements, low production costs, and adaptation to different geographic locations or to different marginal soils [2, 3].

Forage and herbaceous crops in general, and switchgrass in particular, have been widely studied because of their importance both to feed animals and to produce energy. Some researchers have focused on studying the influence of different irrigation methods or fertilizers in the productivity of the crop [4]. Another line of research has focused on improving the plant quality or productivity by means of genetic manipulations [5, 6]. Finally, many other authors have studied the drying and wetting processes of these types of crops: Brück and van Elderen researched the drying process of hay and wheat straw [7], Pitt and McGechan modeled the rewetting of grass [8], and Barr and Brown worked with forage [9].

Drying is the oldest method of preserving and increasing the storage life of food [10, 11]. Moreover, drying techniques make it possible to reduce the product weight, increase storage periods, reduce packaging requirements, and conserve the nutritional properties of the food [12]. Nowadays, most food is dried in industrial dryers because the drying process can be controlled properly, the length of the drying process can be reduced, and some problems such as weather uncertainty or insect infestation can be avoided [10, 13]. Nevertheless, some agricultural products like fodders are dried on the land, where the drying process cannot be controlled and the duration and product quality after drying is uncertain. In these cases, a tool to predict the impact of uncontrolled variables on the drying process will be useful to improve the overall result of the process.

Researchers usually develop mathematical models of processes to predict the unknown value of a process variable as a function of known values of variables of this process. Mathematical models have been proposed in biology and agriculture to explain the evolution of several processes such as animal population evolution [14], vegetal or plant growing [15], and weather condition evolution [16, 17]. Moreover, models and empirical equations have been proposed for drying processes in the field of agriculture and the food industry [7, 13, 18, 19]. These mathematical models and empirical equations give accurate results only in very specific conditions, but there is no way to obtain general equations with good results [10, 11]. To solve this, some authors suggest the application of artificial neural networks (ANN) because of their learning ability and their ability to approximate nonlinear functions. ANN-based models have been proposed to model algal growth [20], weather condition evolution [21, 22], and drying processes. ANN-based models show better global results than other mathematical models [10, 11, 23]. ANN-based models have been developed to describe drying in tomatoes [10], carrots [24], grapes [11], onions [25], potato [26], tobacco [23], hay [9], and kiwi [27]. However, there is little information known about the processes of drying windrows of switchgrass.

The goal of this article was to develop a model for predicting moisture content (MC) in switchgrass land conditions. To this end, two ANN-based models were proposed, developed, and tested. The first ANN was designed to predict the future MC in processes with constant air conditions and no rain, in order to compare it with other empirical equations proposed in the literature and analyze whether an ANN-based model is appropriated for this purpose. The second ANN was designed to predict the future MC in processes with variable air conditions and rain, which are processes more similar to the drying conditions of the switchgrass when it is dried in the land. The proposed work was completed in four steps: (1) a controlled environment chamber and a sprinkler were used to simulate the switchgrass drying processes and to acquire experimental data; (2) two ANN-based models were developed to estimate MC in the two drying scenarios proposed; (3) different experiments were performed to test the proposed models; and (4) experimental results were analyzed to evaluate the proposed models and compare them with other similar works reported in literature.

# MATERIALS AND METHODS

This section presents the materials and methods employed in this article: the switchgrass employed, the method to prepare the samples, the drying equipment, the drying procedures, and the data acquisition and data processing methods.

## Switchgrass

The switchgrass used in the experiments was cut by hand from a 2-ha experimental plot at the University of Kentucky located just North of Lexington, Kentucky. The plot was planted with the lowland 'Alamo' variety, established in early June of 2007. The experiments were conducted between October and December 2012 when plants were approximately 1.8m tall. A killing frost occurred on November 5, 2012, and samples were collected prior to and after the killing frost.

## Sample Preparation

Switchgrass samples were hand-harvested from the field approximately 2 h before the beginning of each drying test. The plants were cut 10 cm above the ground, which was a typical harvest mowing height. The samples were cut into approximately 100-cm lengths in order to put them inside the sample baskets.

Several baskets were built to place the switchgrass inside the drying chamber. The structure of the baskets was made of PVC plastic and the walls and floor were lined with a plastic mesh to allow air circulation while containing a smaller sample (Fig. 1). Basket dimensions were 106x84.5x37 cm. The amount of switchgrass deployed in the baskets varied between 0.770 and 4.905 kg in the experiments due to differences in bulk density and MC tested during the drying processes. The cut samples were arranged in the same general direction within each basket to simulate the orientation and distribution that would be found in a windrow after mechanical mowing.

Rainfall was simulated using a common turf sprinkler (Fig. 1d). Samples were removed from the drying chamber and placed under the sprinkler for the desired amount of time and then returned immediately to the drying chamber.

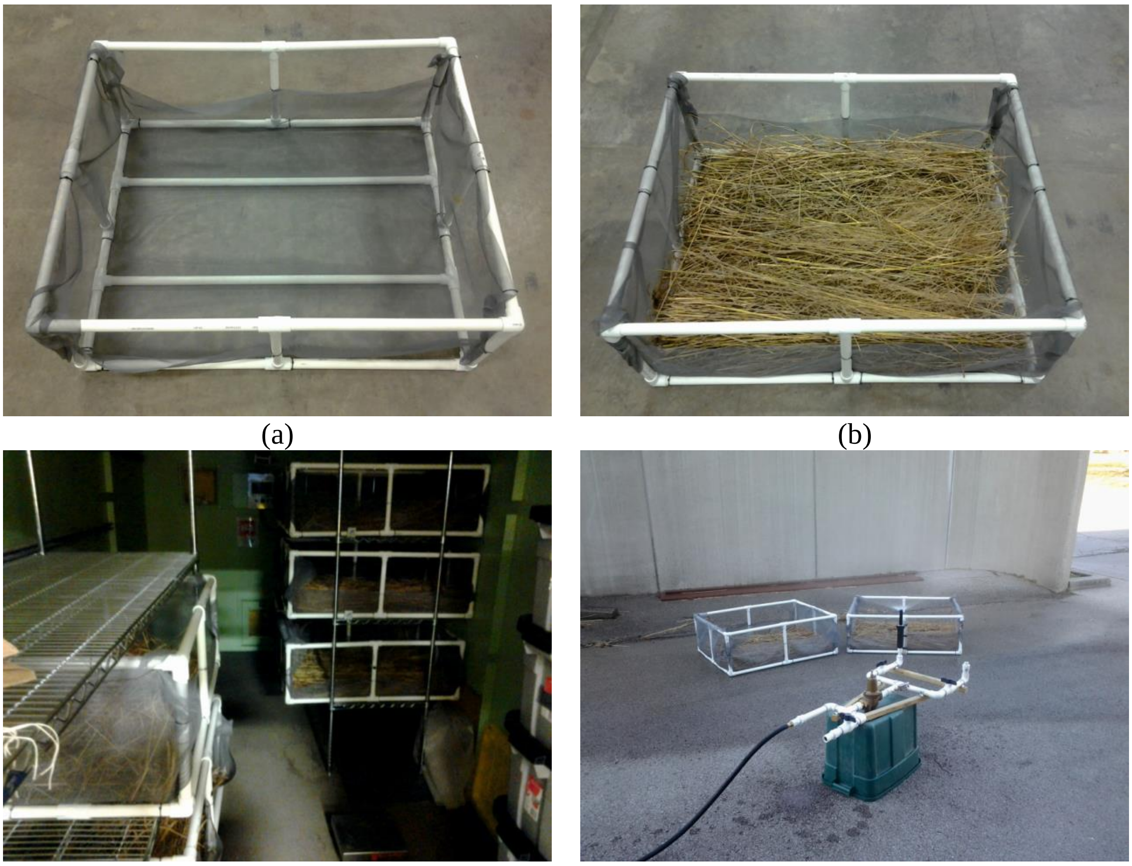
## Drying Equipment

The drying experiments were performed in an insulated ISO cargo container (6x2.5x2.5 m) located in an environmentally controlled laboratory in the Biosystems and Agricultural Engineering Department. Temperature and relative humidity inside the container were controlled based on the dew point temperature and the air was reheated to desired temperature using a series 9230 room conditioner of the Parameter Generation & Control company (Black Mountain, NC, USA), with an airflow rate of 0.5 $m^3$/s.

An industrial oven was employed to dry the biomass and calculate the MC according to the American Society of Agricultural and Biological Engineers (ASABE) Standard ANSI/ASAE S358.3 [28].

## Drying Procedures

Fifty drying tests were performed at various air temperature and relative humidity conditions and rain precipitation events. The duration of each test was approximately 5 or 6 days until equilibrium was reached, which depended on the environmental conditions simulated. The temperature was varied between 3 and 21 ºC, and the relative humidity was varied between 20 and 100% during the simulated drying processes. Rain precipitation was simulated in 20 of the 50 drying tests performed. The precipitation varied between 10.16 and 26.42mm total precipitation over durations between 45 and 60 min. The ranges of the drying variables were chosen as the most representative ranges for the months of October and November in the state of Kentucky based on historical data from 2001 to 2013 extracted from the National Climatic Data Center of the National Oceanic and Atmospheric Administration.



(a) (b)

(c) (d)

**Figure 1:** (a) Empty basket employed to carry switchgrass inside the drying chamber. (b) Basket filled with switchgrass before a drying process. (c) Baskets inside the dryer chamber. (d) Rain simulation equipment.

## Data Acquisition

Temperature and relative humidity conditions within the container, rain simulation conditions, and initial basket weights were measured in each process. MC was determined during the drying process by measuring the total basket weight; a total of 1,028 basket weights were determined. At the conclusion of each test, the biomass from each simulated drying process was completely dried in an oven at 103 ºC to calculate the final equilibrium MC according to the ASABE Standard ANSI/ASAE S358.3 [28]. Both biomass weight evolution and final MC were used to calculate other process variables. MC dry basis, which is also known as absolute humidity, was the MC definition employed in this article because it is the definition employed by some other authors in the literature [10, 11]. Equation (1) presents its formula:

$$MC = \frac{Mass\ of\ water}{Mass\ of\ dry\ sample} \cdot 100(\%) \tag{1}$$

where mass of wet sample is the weight of the sample prior to drying, mass of dry sample is the weight of the sample after being dried according to the ASABE Standard ANSI/ASAE S358.3 [28], and the mass of water is the difference between the mass of wet sample and the mass of dry sample.

Moisture ratio (MR) is another parameter related to MC that was calculated from the data acquired for its later analysis. Equation (2) presents its formula:

$$MR = \frac{MC - MC_e}{MC_0 - MC_e} \tag{2}$$

where MC is the moisture content at a specific time, $MC_0$ is the initial MC, and $MC_e$ is the equilibrium MC. $MC_e$ is the MC value after an infinite exposure time to a constant temperature and air relative humidity. In this study, $MC_e$ was numerically estimated as the final MC at the conclusion of the drying experiment, as it is done in other similar articles [24].

Temperature and relative humidity of the air were also acquired during the drying process by means of HOBO (ONSET Computer Corp., Bourne, MA, USA) temperature and relative humidity data loggers.

## Data Processing

The methodology followed in this article consists of three stages: the preprocessing stage, the processing stage, and the analysis stage. The procedure to develop these stages is presented in this subsection.

Firstly the acquired data were preprocessed to calculate the MC and MR taking into account Eq. (1) and Eq. (2), respectively.

Secondly, two ANN-based systems were designed to predict the biomass MC as a function of the time horizon considering different drying process conditions and using the acquired and preprocessed data. The time horizon is the difference between the actual time and the future time in which the MC is going to be predicted. The ANN considered in our work was a totally connected multilayer perceptron (MLP), which is a network composed by neurons called perceptrons, whose structure is shown in Fig. 2a. They have a set of m inputs ($x_1$, $x_2$, ... , $x_m$) multiplied by their respective weights ($w_1$, $w_2$, ... , $w_m$) and combined with a bias ($b$) to generate the signal $v = \sum_{i=1}^{m} w_i \cdot x_i + b$. The activation function $\varphi(\cdot)$ is then applied to generate the output $y = \varphi(v)$. MLPs have connections with all of the perceptrons of the previous layer and their outputs serve as an input of all of the neurons of the next layer [29, 30]. Figure 2b shows one example of a totally connected MLP with two inputs, one output, and one hidden layer with three neurons.

The MLP-ANNs proposed in this article had three layers: the input layer, one single hidden layer, and one output layer. The number of neurons in the input and output layers was determined respectively by the number of input and output variables considered in the model. The number of neurons in the hidden layer was chosen analyzing the mean squared error (MSE) of the trained ANN when a different number of hidden neurons were used in the ANN. To minimize the randomness of the training process, 100 repetitions of the experiment for each number of hidden neurons were performed. The activation function used in the proposed ANN is the tan-sigmoid function (Eq. (3)) for the neurons of the hidden layer and a linear function for the neurons of the output layer. Because the input layer does not receive signals from other neurons, their neurons do not have an activation function because they only send the input variables to the neurons of the hidden layer.

$$tan - sigmoid(k; x) = tanh(kx) = \frac{e^{kx} - e^{-kx}}{e^{kx} + e^{-kx}}, k \in \mathbb{R}, k > 0 \quad (3)$$

ANN training was performed by applying the backpropagation method with the Levenberg-Marquardt algorithm [31] employing the *trainlm* MATLAB function. The minimum value of the gradient was set to $10^{-10}$, setting a stop criterion when the gradient falls under it. Moreover, the initial value of $\mu$ was set to $10^{-3}$, the factor of increment and decrement of $\mu$ during the learning process ($\beta$) was set to 10, and the maximum value of $\mu$ was set to $10^{10}$, according to the nomenclature of Hagan and Menhaj [31].

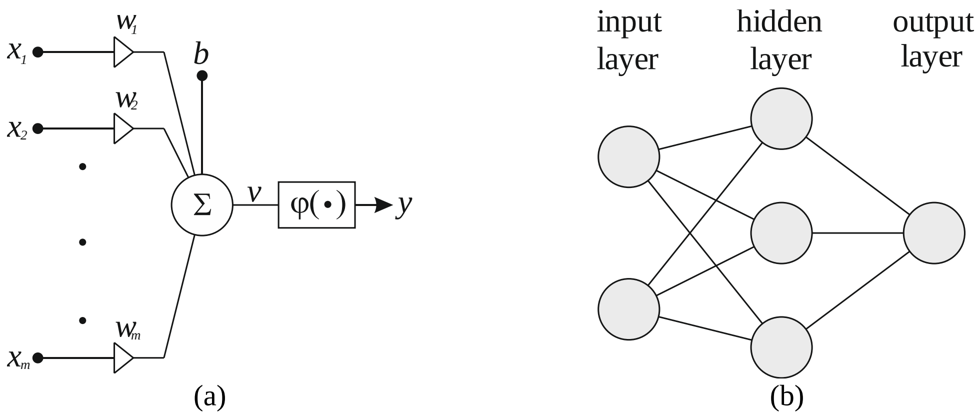



**Figure 2:** (a) Signal-flow graph of a perceptron, where ($x_1$, $x_2$, ... , $x_m$) are the inputs; ($w_1$, $w_2$, ... , $w_m$) are the weights; $b$ is the bias; $v$ is the linear combination of the inputs, weights, and bias $v = \sum_{i=1}^{m} w_i \cdot x_i + b$; $\varphi(\cdot)$ is the activation function; and $y = \varphi(v)$ is the output. (b) Example of a three-layer MLP with two inputs, one output, and three neurons in the hidden layer.

Moreover, some mathematical empirical equations commonly proposed in the literature to describe drying processes were employed to model the drying processes of the first experiment [32–37]. Table 1 shows the drying empirical equations proposed and their equations.

| Name | Equation | Reference |
|---|---|---|
| Lewis | $\mathrm{MR} = e^{-kt}$ | [32] |
| Page | $\mathrm{MR} = e^{-kt^n}$ | [33] |
| Henderson and Pabis | $\mathrm{MR} = a \cdot e^{-kt}$ | [34] |
| Logarithmic | $\mathrm{MR} = a \cdot e^{-kt} + c$ | [35] |
| Two-term | $\mathrm{MR} = a \cdot e^{-k_0 t} + b \cdot e^{-k_1 t}$ | [36] |
| Wang and Singh | $\mathrm{MR} = 1 + at + bt^2$ | [37] |

**Table 1.** Definition of the drying empirical equations used in this article.

Thirdly and finally, the MC prediction error was calculated for each experiment using three error parameters used in other similar articles in the literature [10, 11]: MSE, mean relative percentage error ($P$), and correlation coefficient ($\rho$), shown in Eqs. (4)–(6).

$$MSE = \frac{\sum_{i=1}^{N}\left(MR_{pred}(i) - MR_{exp}(i)\right)^2}{N - N_p} \tag{4}$$

$$P = \frac{100}{N} \cdot \sum_{i=1}^{N} \left| \frac{MR_{pred}(i) - MR_{exp}(i)}{MR_{exp}(i)} \right| \tag{5}$$

$$\rho(\%) = \left(1 - \frac{\sum_{i=1}^{N}\left[MR_{pred}(i) - MR_{exp}(i)\right]^2}{\sum_{i=1}^{N}\left[MR_{exp}(i)\right]^2}\right) \cdot 100 \tag{6}$$

where $MR_{pred}$ is the predicted MR, $MR_{exp}$ is the MR obtained in the experiments, $N$ is the number of samples, and $N_p$ is the number of parameters considered in the equations.

MATLAB programing environment was used to perform all of the tasks of the data processing stage using the Neural Network Toolbox to develop and analyze the ANN-based models.

## EXPERIMENTAL

Two different experiments were conducted in this study to evaluate the performance of two ANN-based models proposed in this article. The first experiment consisted of the MC estimation in processes with constant air conditions and without rain simulation, and it served to compare empirical equations with ANN-based models. The second experiment consisted of the MC estimation in processes simulating variable air conditions and rain simulations, and it served to analyze the capability of the proposed ANN-based model to estimate MC in these conditions.

The drying process conditions and the data processing–specific considerations for each experiment will be explained in the next two subsections.

### First Experiment: Comparison in Processes with Constant Air Conditions

Data from processes with constant air conditions and without rain events were used in this experiment to estimate the MC using six empirical equations and an ANN-based model. Temperature and relative humidity varied among different drying processes between 10 and 21 ºC and between 40 and 90%, respectively, but both variables were maintained constant for each drying process simulated. Data from processes in these conditions were used because empirical equations presented in Table 1 only have time as an input variable, so they are mainly used in processes with constant air conditions [10,11]. 445 measurements of MC were used in this experiment, dividing the data obtained from these samples randomly into three groups or data sets to be used for training, validation, and testing. Training and validation data sets were used to adjust the weights of the ANN in the ANN-based model and the constant values in the empirical equations. In the ANN training stage, the validation data set was used to evaluate the quality of the network during the training process, in order to continue the training process with the training data set until a target error was reached in the validation data set. The testing data set was used to calculate the performance of each MC estimator and obtain the results.

A variable time horizon was considered, obtaining a total of 4,249 input–output samples for this experiment. These samples were split into training, validation, and testing data sets, with 54.7, 27.7, and 17.6% of the samples, respectively. The empirical equation constants were adjusted with a gradient descent–based optimization method on each equation with the same training data used to train the ANN-based model. The performance of both the proposed ANN-based model and the empirical equations was measured using the testing data.

The ANN-based model proposed to characterize the first experiment drying processes has seven input variables and one output variable. The input variables are (1) MC at the beginning of the drying process, (2) MC at the time at which the estimation in going to be calculated, (3) date of the beginning of the drying process, (4) date in which the estimation is going to be calculated, (5) date for which the MC is going to be estimated, (6) temperature, and (7) relative humidity of the air during the drying process, and the output variable is the predicted MC. The number of hidden neurons is nine in order to minimize the MSE of the method, as can be seen in Fig. 3a.

The aim of this experiment is to compare the performance of the proposed ANN-based model with the classic empirical equations presented in Table 1, which have been proven to work properly in the drying conditions employed in this experiment.

### Second Experiment: Analysis of the Proposed ANN-Based Model for Processes with Variable Air Conditions and Rain

Data from processes with variable air conditions and with and without simulated precipitation were used in this experiment to estimate the MC using an ANN-based model. Temperature and relative humidity varied between 3 and 21 ºC and between 20 and 100%, respectively, in the simulated drying processes. Simulated precipitation treatments were evaluated in 40% of the tests used in this experiment, and their duration and precipitation varied between 45 and 60 min, and between 10.16 and 26.42 mm of water was added to the switchgrass. The 1,028 data points of MC collected in this experiment were divided into three groups for training, validation, and testing similar to how they were divided in the comparison experiment.

A variable time horizon was implemented as in the comparison experiment, obtaining a total of 10,869 input–output samples. These samples were split into training, validation, and testing data sets, with proportions of 50.9, 24.4, and 24.7% of the samples, respectively. The training process and the evaluation of the ANN-based model were performed with the same procedure that was used for the comparison experiment.

The ANN-based model proposed to characterize the second experiment drying processes has 16 input variables and one output variable. The 16 input variables consist of two variables related to the MC, three related to the date, four statistical parameters of the air temperature, four statistical parameters of the air relative humidity, and three variables to describe the rain event. The two variables related to the MC are (1) MC at the beginning of the drying process and (2) MC at the time on which the estimation is going to be calculated. The three variables related to the date are (1) the date of the beginning of the drying process, (2) the date in which the estimation is going to be calculated, and (3) the date for which the MC is going to be estimated. The four statistical parameters selected to describe the temperature and relative humidity were the median or 50$^{th}$ percentile, standard deviation, 90$^{th}$ percentile, and 10$^{th}$ percentile of the temperature and relative humidity, respectively. Median, 90$^{th}$ percentile, and

$10^{th}$ percentile were used instead of mean, maximum, and minimum values because percentile values reduce the influence of outlier values of data loggers. The three variables that describe the rain events are the hour in which the rain event starts, the duration, and the water precipitation in millimeters during this event. The output variable was the predicted MC. The number of hidden neurons is 28 in order to minimize the MSE of the method, as can be seen in Fig. 3b.

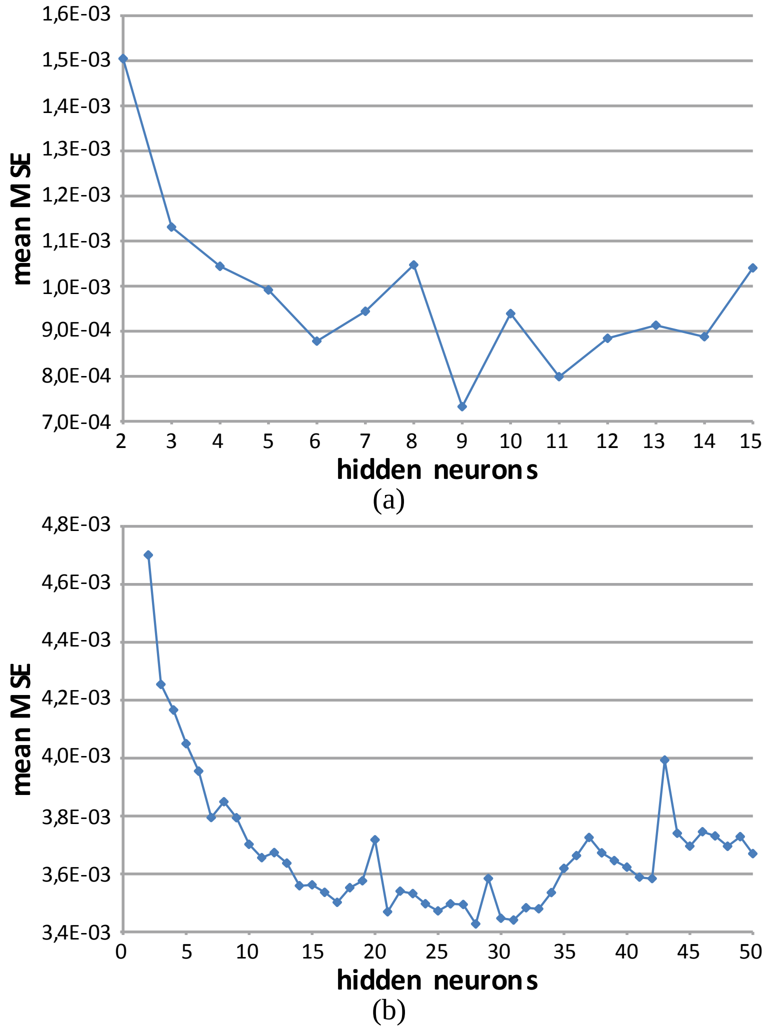



**Figure 3:** Results of the experiments performed to choose the number of hidden neurons for (a) the first ANN-based model proposed and (b) the second ANN-based model proposed.

The aim of this experiment was to analyze the capability of the proposed ANN-based model to estimate the MC in a real switchgrass drying process. To this end, variable air conditions and rain information of the simulated switchgrass drying processes were monitored to acquire the data used in this experiment. An ANN-based model was trained to predict the MC of the switchgrass along the process, and it was tested in two different conditions: the first condition consisted of predicting the MC as in the comparison experiment and the second condition consisted of predicting the final MC given the initial MC. The first testing condition will allow us to compare the results with results obtained by other authors and the second testing condition will show the performance of the proposed model in the worst case, that is, with the longest time horizon.

# RESULTS AND DISCUSSION

The numerical results of the proposed experiments are presented and analyzed in this section. Moreover, the results obtained are discussed, comparing them with other results in the literature and analyzing the advantages, limitations, and possible applications of the ANN-based models proposed.

## First ANN-Based Model Proposed Versus Six Mathematical Empirical Equations

Table 2 and Fig. 4 illustrate the results obtained in the first experiment. MSE, *P*, and $\rho$, were used in Table 2 to evaluate the prediction capability of both the ANN-based model and the six empirical equations.

| Equation/model | MSE | *P* (%) | Corr ($\rho$, %) |
|---|---|---|---|
| Lewis | $1.073 \cdot 10^{-3}$ | 5.583 | 99.719 |
| Page | $3.104 \cdot 10^{-3}$ | 11.573 | 99.186 |
| Henderson and Pabis | $9.796 \cdot 10^{-4}$ | 5.263 | 99.743 |
| Logarithmic | $1.512 \cdot 10^{-3}$ | 8.148 | 99.604 |
| Two-term | $1.721 \cdot 10^{-3}$ | 8.805 | 99.549 |
| Wang and Singh | $2.634 \cdot 10^{-3}$ | 8.979 | 99.309 |
| ANN-based model | $2.613 \cdot 10^{-4}$ | 2.442 | 99.897 |

**Table 2.** First experiment results with constant temperature and relative humidity and no precipitation, with classic empirical equations and the proposed ANN-based model to predict the future MC value.

Figure 4 shows the results for two drying processes of this experiment, where experimental data and predicted data from the empirical equations and the proposed ANN-based model are shown. Experimental data are represented by black asterisks, ANN data with black circles, and empirical equations data with colored circles.

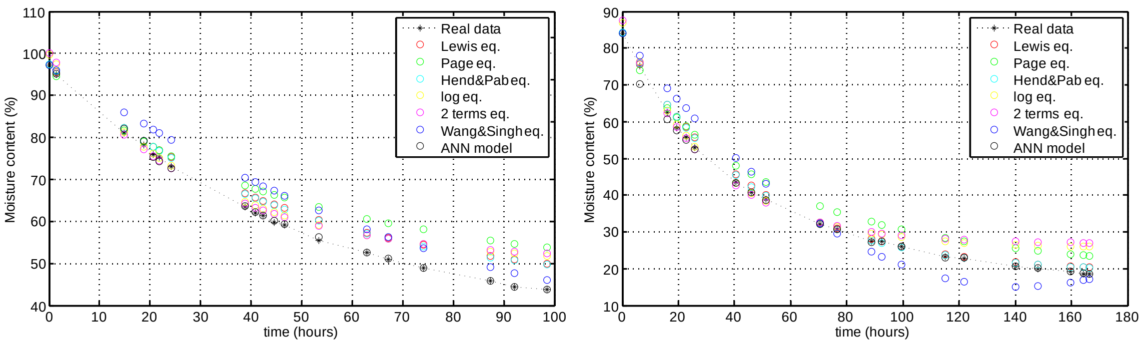



**Figure 4:** First experiment results for two drying tests under constant temperature and relative humidity, where the MC evolution over time is shown. Lewis, Page, Henderson and Pabis (Hend&Pab), logarithmic (log), two-term (2 terms), and Wang and Singh (Wang & Singh) empirical equations are represented by red, green, cyan, yellow, purple, and dark blue circles, respectively; the proposed ANN model data are represented by black circles and the experimental data by black asterisks.

Another interesting result is the relationship between the MC prediction error and the time horizon, which is represented in Fig. 5. The linear correlation coefficient between the prediction error and the time horizon was $\rho = -0.1411$.

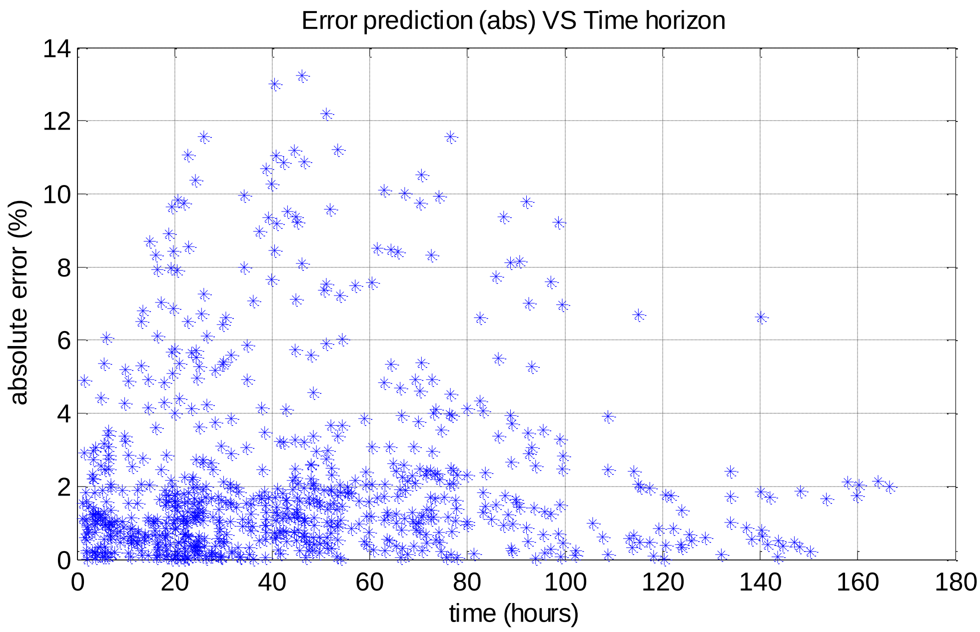



**Figure 5:** Representation of the absolute prediction error against the time horizon obtained in the first experiment using the first ANN-based model proposed.

### Second ANN-Based Model Proposed Performance Results

Results obtained in the second experiment are presented in this subsection. The prediction capability of the second ANN-based model proposed was evaluated with the MSE, *P*, and $\rho$ in the two conditions proposed in the last paragraph of the ''Sample Preparation'' subsection, and the results are shown in Table 3 and Fig. 6.

| Condition | MSE | *P* (%) | Corr ($\rho$, %) |
|---|---|---|---|
| First | $3.262 \cdot 10^{-3}$ | 5.609 | 98.782 |
| Second | $9.706 \cdot 10^{-4}$ | 5.530 | 99.275 |

**Table 3.** Second experiment results, where the second ANN-based model proposed prediction capability is measured in two conditions: variable time horizon and the longest time horizon for each drying process

Figure 6 shows graphical results of the second experiment for three different drying processes. Temperature (red) and relative humidity (green) of the air are represented for each drying process; experimental data (black) and the data estimated with the second ANN-based model proposed (blue) are also represented. In the first drying process (Figs. 6a.1 and 6a.2), the measured air conditions were almost constant, with a temperature of 22 ºC and a relative humidity of 55% and a rain event of 1-h duration, yielding a total precipitation of 13.72 mm in hour 140 of the process. In the second drying process (Figs. 6b.1 and 6b.2), the measured air conditions varied with a temperature between 10 and 18 ºC, relative humidity between 65 and 80%, and without rain events. In the third drying process (Figs. 6c.1 and 6c.2), the measured air conditions varied with a temperature between 3 and 21 ºC, relative humidity between 15 and

100%, and a rain event of 45-min duration, yielding a total precipitation of 18.29 mm in hour 95 of the process.

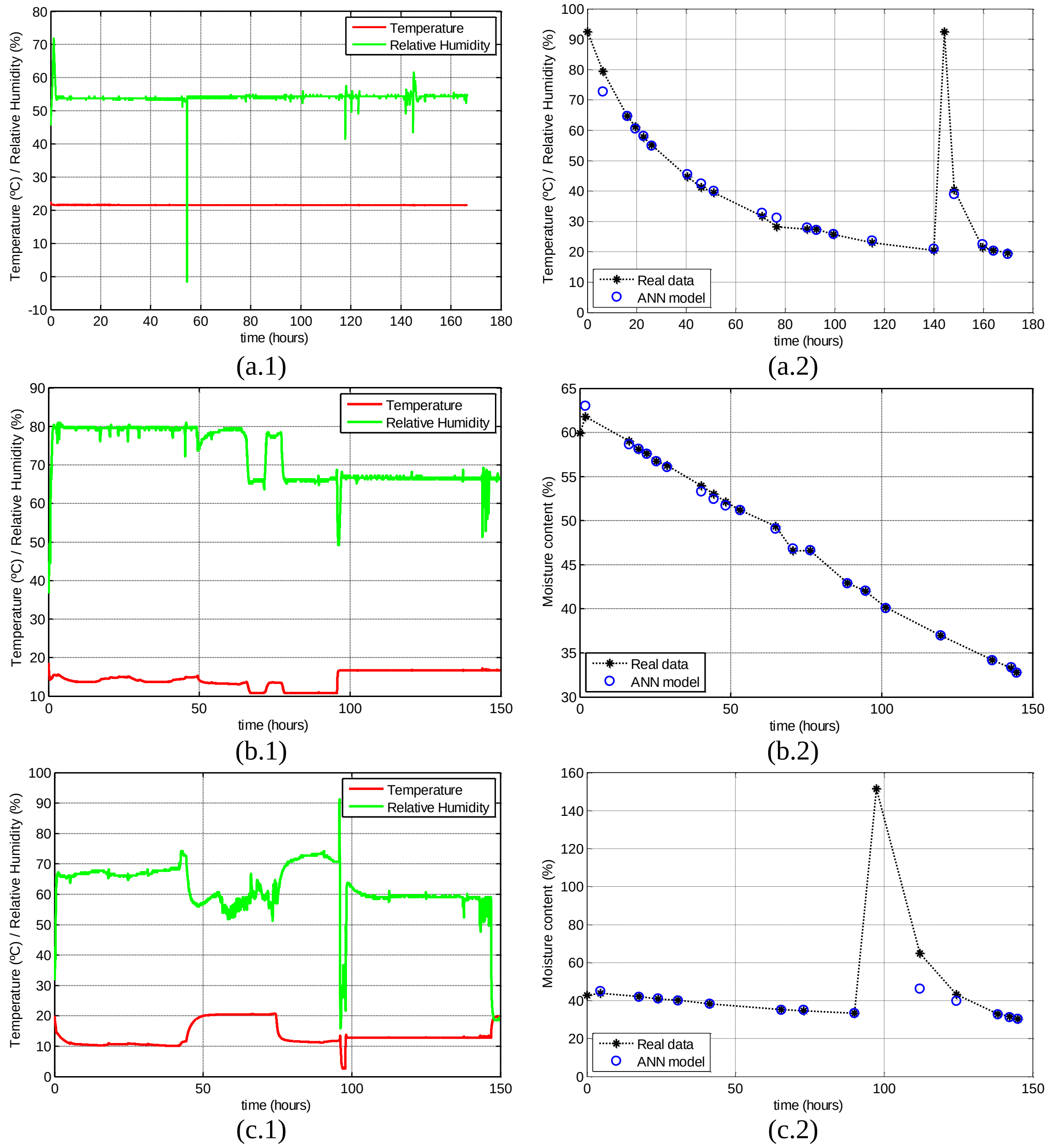



**Figure 6:** Second experiment results for drying processes under different conditions. Temperature and relative humidity of the air along the process are represented in (a.1), (b.1), and (c.1) (first column), and MC is represented in (a.2), (b.2), and (c.2) (second column) for three experiments: the first one, with constant air conditions (temperature and relative humidity) and with a rain event in hour 140 ((a.1) and (a.2)), the second one with variable air conditions and without rain ((b.1) and (b.2)), and the third one with variable air conditions and with a rain event in hour 95 ((c.1) and (c.2)).

## Discussion

The experiments conducted in this article show two main results: (1) the proposed ANN-based model could fit the MC data from drying processes with constant air temperature and relative humidity conditions more accurately than six empirical equations widely proposed in the literature and (2) data from the drying processes with variable air conditions and with rain events could be fitted with low errors using the second ANN-based model proposed. These results prove the suitability of ANN-based models to predict the MC in laboratory switchgrass windrow drying processes and presume the capability of ANN-based models to predict the future MC in windrow switchgrass drying processes.

The first ANN-based model proposed was compared with six empirical equations commonly proposed in the literature. The first experiment results showed that the overall performance of the ANN-based model was significantly better than the performance of the empirical equations. Figure 4 shows that some empirical equations, such as the logarithmic equation, have a low error in the first hours of the drying process, whereas other empirical equations have more accurate predictions toward the end of the drying process. However, only the proposed ANN-based model can adjust to the data during the drying process and low errors were observed during the entire drying process. This concurs with other researchers' results [10, 11], who also concluded that their proposed ANN-based models were more accurate than the traditional empirical equations. The less accurate modeling performance obtained by the empirical equations in relation to ANN-based models may be due to the limitation that they only take into account time as input variable, whereas ANN-based models can easily consider other variables such as temperature or relative humidity.

The second ANN-based model proposed was tested in drying processes with varying air conditions and rain. The results showed the capability of this ANN-based model to fit experimental data from drying processes that simulated field drying conditions of switchgrass with an acceptable error. This observation can be seen based on the performance of the second ANN-based model shown in Fig. 6 for processes with and without rain events and with constant and varying air conditions. The prediction performance of the ANN-based model just after rain events is noteworthy. The first sample after the rain event has a high prediction error, but this error decreases in the subsequent samples. This behavior of the model is acceptable because it only occurs for a short time after a rainfall event. Results of the second experiment cannot be compared with the results of other similar articles in the literature because some of these works consider constant air conditions without rain [10, 11] and other works consider variable air conditions and rain but do not analyze the MC evolution [7–9].

One common characteristic of the two ANN-based models proposed is the low dependence between the MC prediction error and the time horizon. There are three facts that support this assertion. First, results of the first experiment showed that the linear correlation coefficient between the MC prediction error and the time horizon was $\rho = -0.1411$, which indicated that these two variables are almost uncorrelated. Second, the highest errors for the first experiment were obtained for a time horizon lower than 80 h, whereas the mean duration of a drying process, and hence the maximum time horizon considered, was around 120 or 150 h. Third, the results from the second experiment presented in Fig. 6 showed that the largest error for processes with rain events happened in the first sample after the rain event but then this error decreased quickly in the next samples and the final MC is predicted with a low error. Moreover, the estimation error for the final MC as a function of the initial MC (second condition explained in the ‘‘Sample Preparation’’ subsection) is comparable to the estimation error for estimations with shorter time horizons. These three facts for both the first and second models indicated that there was a low dependence between the MC prediction error and the time horizon, and the third one suggests that other factors such as the rain events could affect the MC prediction even more than the time horizon.

Comparing both experiments, results obtained from the second experiment are not as strong as the results from the first experiment. This is because the drying processes modeled in the second experiment are more complex than the processes modeled in the first; more drying conditions were considered for the second model and more input variables were used in the ANN. Rain event information was included in the model because these events make a big change in the MC of the switchgrass, as observed in Figs. 6a.2 and 6c.2. In addition, variation in the air conditions was modeled and utilized the standard deviation and $10^{th}$, $50^{th}$, and $90^{th}$ percentiles of the air conditions. Analyzing other similar works proposed in the literature [10, 11], their results are comparable with the results of the first experiment proposed. Moreover, the complexity of these models, denoted by the number of factors that affect the models, is also comparable to the complexity of the first ANN-based model proposed.

One observation about the second ANN-based model proposed is that some of its input variables were future values of the air conditions or the rain events. In the proposed experiments, these input variables values were obtained without error because all of the data for the whole drying process were available. However, if this model is used as a real-time application in an uncontrolled environment, these variables must be predicted, thereby adding a prediction error to the model error. This variable prediction error is expected to be low because there are weather forecast methods in the literature accurate enough for the prediction horizons needed for the switchgrass drying process [22].

Analyzing the deployed model, it can be seen that the MC evolution depends on the temperature and relative humidity in the same way that basic drying theory indicates: the drying rate decreases as temperature falls and/or relative humidity rises. Taking into account the third variable considered –rain– shows some interesting observations about its influence in the drying process. The first is that the MC evolution is influenced by the time in which the rain event occurs: if it occurs in the first hours of the drying process, the number of hours needed to reach a target MC is higher than if it occurs in the last hours of the drying process. This may be because as switchgrass dries its absorption capacity is lower and the moisture increment due to the rain event can be removed easier.

The main contribution of this article is the proposal of an ANN-based model for the switchgrass drying process taking into account the usual weather conditions that occur. On the one hand, there are several works in the literature that model or describe the industrial drying process of agricultural or food products, in which the drying conditions are under control and maintained at constant values [10, 23, 24, 26]. On the other hand, there are other articles that consider drying with uncontrolled variable air conditions [9, 33]. Nevertheless, to the best of our knowledge, there are no articles in the literature that model land drying processes that consider variable air temperature and relative humidity and rain events as variables that affect the drying process.

One limitation of the proposed experiments is that they were performed with simulated conditions. These simulated conditions did not take into account factors such as soil conditions, solar radiation, atmospheric pressure, wind velocity, and wind direction, which affect the switchgrass drying process in real outdoor conditions. For this reason, a new model with more input variables that take into account these factors should be developed to estimate MC in real outdoor conditions. Future work on this topic could be conducted to model windrow drying processes of switchgrass or other grass crops. Due to the input–output mapping property of ANNs, windrow drying processes could be modeled using ANN-based models.

The methods proposed in this article could be applied in agricultural and industrial tasks. One example of the utilization of these methods is in planning optimal times for mowing. Switchgrass and other herbaceous crops are typically mowed and dried in windrows prior to baling and storage. The final quality of the product is highly dependent on the drying process, in which a rain event can cause quality degradation. Proposed ANN-based models could be used to plan the mowing and drying processes taking into account the weather forecast in order to avoid or minimize the possible quality loss of the dried product. The function of the models on this application will be the estimation of the duration of the drying process taking into account the initial MC of the switchgrass and the weather forecast information.

## CONCLUSIONS

This article shows the capability of two ANN-based models to describe the MC evolution in simulated switchgrass drying processes. The results show that the first model proposed fits MC data from drying processes with constant air temperature and relative humidity more accurately than six empirical equations commonly used in the agro-industrial drying processes literature. The second ANN-based model proposed for drying processes with variable air temperature and relative humidity conditions and with rain events fit the change in moisture content accurately, with a correlation coefficient between the experimental data and the predicted data greater than 98.8%. Considering these results, it can be concluded that an ANN-based model may be used to estimate MC in real switchgrass drying processes with an acceptable error.

## FUNDING

The first author's work was made possible thanks to a Formación de Personal Investigador program grant. This program was financed by the Universidad de Valladolid (Spain) and cofinanced by Banco de Santander. Moreover, the experiments and collaboration among the authors were possible thanks to a grant entitled *Ayudas para Estancias Breves en el Desarrollo de Tesis Doctorales* financed by the Universidad de Valladolid (Spain). This grant was obtained by the first author in 2012 and it financed the stay at the University of Kentucky during the months while the proposed experiments were performed.

## REFERENCES

[1] Moser, L.E.; Vogel, K.P. Switchgrass, Big Bluestem, and Indiangrass. In *Forages: Vol. I. An Introduction to Grassland Agriculture*; Barnes, R.F.; Miller, D.A.; Nelson, C.J., Eds.; Iowa University Press: Ames, IA, 1995; 409–420.

[2] McLaughlin, S.B. New switchgrass biofuels research program for the Southeast. In *Proceedings of the Annual Automotive Technology Development Contractors Meeting*, Dearborn, MI, November 2–5, 1992.

[3] Sanderson, M.A.; Reed, R.L.; McLaughlin, S.B.; Wullschleger, S.D.; Conger, B.V.; Parrish, D.J.; Wolf, D.D.; Taliaferro, C.; Hopkins, A.A.; Ocumpaugh, W.R.; Hussey,

M.A.; Read, J.C.; Tischler, C.R. Switchgrass as a sustainable bioenergy crop. *Bioresource Technology* **1996**, *56*(1), 83–93.

[4] Ercoli, L.; Mariotti, M.; Masoni, A.; Bonari, E. Effect of irrigation and nitrogen fertilization on biomass yield and efficiency of energy use in crop production of Miscanthus. *Field Crops Research* **1999**, *63*(1), 3–11.

[5] Somleva, M.N.; Tomaszewski, Z.; Conger, B.V. Agrobacteriummediated genetic transformation of switchgrass. *Crop Science* **2002**, *42*(6), 2080–2087.

[6] Rubin, E.M. Genomics of cellulosic biofuels. *Nature* **2008**, *454*(7206), 841–845.

[7] Brück, I.G.M.; van Elderen, E. Field drying of hay and wheat. *Journal of Agricultural Engineering Research* **1969**, *14*(2), 105–116.

[8] Pitt, R.E.; McGechan, M.B. The rewetting of partially dried grass swaths by rain: Part 1, lumped and distributed models of moisture fluctuation. *Journal of Agricultural Engineering Research* **1990**, *45*, 55–67.

[9] Barr, A.G.; Brown, D.M. Estimating forage yield and quality changes during field drying for hay. 2. Model of forage drying. *Agricultural and Forest Meteorology* **1995**, *76*(2), 107–127.

[10] Movagharnejad, K.; Nikzad, M. Modeling of tomato drying using artificial neural network. *Computers and Electronics in Agriculture* **2007**, *59*(1–2), 78–85.

[11] Çakmak, G.; Yıldız, C. The prediction of seedy grape drying rate using a neural network method. *Computers and Electronics in Agriculture* **2011**, *75*(1), 132–138.

[12] Simal, S.; Femenı´a, A.; Llull, P.; Rossello´ , C. Dehydration of aloe vera: Simulation of drying curves and evaluation of functional properties. *Journal of Food Engineering* **2000**, *43*(2), 109–114.

[13] Toğrul, I.T.; Pehlivan, D. Modelling of thin layer drying kinetics of some fruits under open-air sun drying process. *Journal of Food Engineering* **2004**, *65*(3), 413–425.

[14] Holland, E.P.; Burrow, J.F.; Dytham, C.; Aegerter, J.N. Modelling with uncertainty: Introducing a probabilistic framework to predict animal population dynamics. *Ecological Modelling* **2009**, *220*(9–10), 1203–1217.

[15] Vance, R.R.; Nevai, A.L. Plant population growth and competition in a light gradient: A mathematical model of canopy partitioning. *Journal of Theoretical Biology* **2007**, *245*(2), 210–219.

[16] Cornford, D. A Bayesian state space modelling approach to probabilistic quantitative precipitation forecasting. *Journal of Hydrology* **2004**, *288*(1–2), 92–104.

[17] Chamorro, L.P.; Porte´-Agel, F. Turbulent flow inside and above a wind farm: A wind-tunnel study. *Energies* **2011**, *4*(11), 1916–1936.

[18] Huang, B.; Mujumdar, A.S. Use of neural network to predict industrial dryer performance. *Drying Technology* **1993**, *11*(3), 525–541.

[19] Kamiński, W.; Stawczyk, J.; Tomczak, E. Presentation of drying kinetics in a fluidized bed by means of radial basis functions. *Drying Technology* **1997**, *15*(6–8), 1753–1762.

[20] Recknagel, F.; French, M.; Harkonen, P.; Yabunaka, K.-I. Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling* **1997**, *96*(1–3), 11–28.

[21] Smith, B.A.; Hoogenboom, G.; McClendon, R.W. Artificial neural networks for automated year-round temperature prediction. *Computers and Electronics in Agriculture* **2009**, *68*(1), 52–61.

[22] Maqsood, I.; Khan, M.R.; Abraham, A. An ensemble of neural networks for weather forecasting. *Neural Computing & Applications* **2004**, *13*, 112–122.

[23] Martínez-Martínez, V.; Baladrón, C.; Gomez-Gil, J.; Ruiz-Ruiz, G.; Navas-Gracia, L.M.; Aguiar, J.M.; Carro, B. Temperature and relative humidity estimation and prediction in the tobacco drying process using artificial neural networks. *Sensors* **2012**, *12*(10), 14004–14021.

[24] Erenturk, S.; Erenturk, K. Comparison of genetic algorithm and neural network approaches for the drying process of carrot. *Journal of Food Engineering* **2007**, *78*(3), 905–912.

[25] Wang, J. A single-layer model for far-infrared radiation drying of onion slices. *Drying Technology* **2002**, *20*(10), 1941–1953.

[26] Islam, M.R.; Sablani, S.S.; Mujumdar, A.S. An artificial neural network model for prediction of drying rates. *Drying Technology* **2003**, *21*(9), 1867–1884.

[27] Fathi, M.; Mohebbi, M.; Razavi, S.M.A. Application of fractal theory for prediction of shrinkage of dried kiwifruit using artificial neural network and genetic algorithm. *Drying Technology* **2011**, *29*(8), 918–925.

[28] American Society of Agricultural Engineers. *Moisture Measurement—Forages*, ASAE S358.3; American Society of Agricultural Engineers: St. Joseph, MI, 2012.

[29] Haykin, S.S. *Neural Networks: A Comprehensive Foundation*, 2nd ed.; Prentice Hall: Englewood Cliffs, NJ, 1999.

[30] Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 2007.

[31] Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* **1994**, *5*(6), 989–993.

[32] Lewis, W.K. The rate of drying of solid materials. *Journal of Industrial & Engineering Chemistry* **1921**, *13*(5), 427–432.

[33] Diamante, L.M.; Munro, P.A. Mathematical modelling of the thin layer solar drying of sweet potato slices. *Solar Energy* **1993**, *51*(4), 271–276.

[34] Zhang, Q.; Litchfield, J.B. An optimization of intermittent corn drying in a laboratory scale thin layer dryer. *Drying Technology* **1991**, *9*(2), 383–395.

[35] Yagcioglu, A.; Degirmencioglu, A.; Cagatay, F. Drying characteristic of laurel leaves under different conditions. In *Proceedings of the 7th International Congress on Agricultural Mechanization and Energy,* Adana, Turkey, May 26–27, 1999.

[36] Henderson, S.M. Progress in developing the thin layer drying equation. *Transactions of the ASABE* **1974**, *17*(6), 1167–1172.

[37] Wang, C.Y.; Singh, R.P. A single layer drying equation for rough rice. ASAE Paper No. 3001; 1978.

# Article 3: Predictive maintenance of a machine

The third article presents the **vibration model of a machine** described in the “Introduction and Summary” part of the document. The main bibliographic data about this article are presented below:

- Title: 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.

- Authors: Martínez-Martínez V, Gomez-Gil FJ, Gomez-Gil J, Ruiz-Gonzalez R
- Journal: Expert Systems with Applications.
- Editor: Elsevier.
- Impact factor: 2.240 (2014).
- Journal Ranking:
    - Q1 on *Operations research & Management Science* (12/81)
    - Q1 on *Engineering, Electrical & Electronic* (48/249)
    - Q1 on *Computer Science, Artificial Intelligence* (29/123)
- Date of publication: April 17th 2015.
- ISSN: 0957-4174.
- Volume (Issue): 42 (17).
- Pages: 6433-6441.
- DOI: 10.1016/j.eswa.2015.04.018.
- URL: http://www.sciencedirect.com/science/article/pii/S0957417415002481.
- Number of cites: 1.

The article presented below is a version of the original article published in the previously mentioned journal. This version of the article has been authorized by the journal editor to be included in this PhD thesis as a compendium of publications.

Expert Systems with Applications 42 (2015) 6433–6441

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

# 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

Víctor Martínez-Martínez [a,*], Francisco Javier Gomez-Gil [b], Jaime Gomez-Gil [a], Ruben Ruiz-Gonzalez [a]

[a] Department of Signal Theory, Communications and Telematics Engineering, University of Valladolid, 47011 Valladolid, Spain
[b] Department of Electromechanical Engineering, University of Burgos, 09006 Burgos, Spain

ARTICLE INFO

*Article history:*
Available online 17 April 2015

*Keywords:*
Artificial Neural Networks
Predictive maintenance
Genetic Algorithms
Vibration
Agricultural machinery
Digital signal processing

ABSTRACT

This article proposes (i) the estimation method of an expert system to predict the statuses of several agro-industrial machine rotary components by using a vibration signal acquired from a single point of the machine; and, (ii) a learning method to fit the estimation method. Both methods were evaluated in an agricultural harvester. Vibration signal data were acquired from a single point of the harvester under working conditions, by varying (1) the engine speed status (high speed/low speed), (2) the threshing operating status (on/off), (3) the threshing balance status (balanced/unbalanced), (4) the chopper operating status (on/off), and (5) the chopper balance status (balanced/unbalanced). Positive frequency spectrum coefficients of the vibration signal were used as the only inputs of an Artificial Neural Network (ANN) that predicts the five rotary component statuses. Four Genetic Algorithm (GA) based learning methods to fit the ANN weights and biases were implemented and its performance was compared to select the best one. The prediction system that is developed was able to estimate the rotary component status under consideration with a mean success rate of 92.96%. Moreover, the best GA-based learning method that was implemented reduced the number of generations by 70% in the best case, compared with a random learning method, allowing a similar reduction in the time needed to reach the expected success rate. The results obtained suggest that (i) an ANN-based expert system could estimate the status of the rotary components of an agro-industrial machine to a high degree of accuracy by processing a vibration signal acquired from a single point on its structure; and, (ii) by using the best implementation of the GA-based learning method proposed to fit the ANN weights and biases, it is possible to improve the success rate and by doing so to reduce the time needed to perform the adjustment. The main contribution of this work is the proposal of a classification method that estimates the status of several rotary elements placed each one far from the others employing the signal acquired from only one accelerometer and non-requiring a feature extraction stage.



## 1. Introduction

Vibration signals can reflect the status of rotary machinery (Li, Tse, Yang, & Yang, 2010). These signals sometimes propagate throughout the structure of the machine with low attenuation, which makes it possible to monitor rotating elements by positioning an accelerometer on the structure of the machine (Li et al., 2013). Nevertheless, this propagation capability has also the disadvantage of transmitting other vibration signals from other elements of the structure that can be added to the signal of interest, making it difficult to extract the relevant information (Albarbar, Gu, & Ball, 2010).

Vibration signals from rotary machinery are commonly analyzed in the frequency domain because of the relation between the rotation frequency of the element and their peaks in the spectrum signal. Some authors have used fast Fourier transform (Chwan-Lu, Shun-Yuan, Shou-Chuang, Jen-Hsiang, & Ke-Fan, 2014; Chwan-Lu et al., 2014; Taghizadeh-Alisaraei, Ghobadian, Tavakoli-Hashjin, & Mohtasebi, 2012), short-time Fourier transform (Vulli, Dunne, Potenza, Richardson, & King, 2009), wavelet transform (Bin, Gao, Li, & Dhillon, 2012; Chen, Tang, & Chen, 2013; Jayaswal, Verma, & Wadhwani, 2011; Lu, Xiao, & Malik, 2015; Rodriguez-Donate, Romero-Troncoso, Cabal-Yepez,

* Corresponding author. Tel.: +34 636 797 528.
*E-mail addresses:* vmarmar@ribera.tel.uva.es (V. Martínez-Martínez), fjggil@ubu.es (F.J. Gomez-Gil), jgomez@tel.uva.es (J. Gomez-Gil), rruigon@ribera.tel.uva.es (R. Ruiz-Gonzalez).

http://dx.doi.org/10.1016/j.eswa.2015.04.018
0957-4174/

Garcia-Perez, & Osornio-Rios, 2011; Wang, Makis, & Yang, 2010; Yan, Gao, & Chen, 2014), S-transform (McFadden, Cook, & Forster, 1999), and the Hilbert–Huang transform (Cheng, Cheng, Shen, Qiu, & Zhang, 2013; Lei, Lin, He, & Zuo, 2013; Li et al., 2010; Wang, Ma, Zhu, Liu, & Zhao, 2014) among others, to perform this analysis. Because of the relationship between the rotation frequencies of the machine elements and the peaks of the spectrum signal, experts can estimate the status of machine elements looking for patterns in the spectrum signal, when sufficient information is gathered on aspects such as the rotation frequency of the machine component. Nevertheless, to do so requires expert analysis of the vibration signals and a detailed knowledge of the working components of the machine, which are not always available. Automated systems have been proposed to estimate the status of machine elements using this strategy when no expert is available. These systems apply knowledge of the machine to extract characteristics from the spectrum signal and to estimate its status according to those characteristics.

Artificial Neural Networks (ANNs) are widely proposed in the literature as mathematical tools to implement the estimation methods needed in automated systems of this kind, because of its learning capability. There are some examples of ANNs implementing estimation methods in the literature (Bin et al., 2012; Cheng et al., 2013; Chwan-Lu et al., 2014; Chwan-Lu, Shun-Yuan, Shou-Chuang, et al., 2014; Jayaswal et al., 2011; Vulli et al., 2009; Yildirim, Erkaya, Eski, & Uzmay, 2009). Nevertheless, the authors have found no previous literature that applies to machine maintenance and vibration signal analysis in the context of ANN-based estimation systems that use previous knowledge to extract characteristics. In contrast, systems have been found with this feature for handwritten character recognition (Polat & Yıldırım, 2008a; Pradeep, Srinivasan, & Himavathi, 2011), hand geometry identification (Polat & Yıldırım, 2008b), sonar target classification (Erkmen & Yıldırım, 2008), and object recognition (Polat & Yıldırım, 2008a). The learning process of ANNs is typically called training, and it is a multivariable optimization problem which can involve hundreds or thousands of variables. To improve the performance of the ANN and to reduce the necessary training time, some authors propose the use of optimization heuristics such as Genetic Algorithms (GAs), particle swarm optimization, and evolutionary programming (Chia-Feng, 2004; Huang, Li, & Xiao, 2015; Leung, Lam, Ling, & Tam, 2003; Song, Hu, Xie, & Zhou, 2013; Xin & Yong, 1997).

Our objective here is to study the feasibility of estimating the status of the rotary machinery in agricultural and industrial machines by using a vibration signal acquired from a single point of the machine structure. To accomplish this general objective, three specific sub-objectives have been proposed. The first one is to design a prediction method to estimate the status of rotary machinery. This method must be able to perform the estimation using the vibration signal specified in the general objective. Moreover, this method must be generalizable to apply it to different rotary components and different agricultural or industrial machines. The second sub-objective is to design and implement four learning methods, to optimize the learning process for the prediction method that is designed for the first objective. The third sub-objective is to evaluate both the prediction method that is designed and the proposed learning methods by implementing an expert system which uses them in an agricultural harvester. To this end, five rotary component statuses of the agricultural harvester were considered: (1) the engine speed status (high speed/low speed), (2) the threshing operating status (on/off), (3) the threshing balance status (balanced/unbalanced), (4) the straw chopper operating status (on/off), and (5) the straw chopper balance status (balanced/unbalanced).

## 2. Background

This section introduces the theoretical basis of Artificial Neural Networks (ANNs) and Genetic Algorithms (GAs).

### 2.1. Artificial Neural Network

ANNs are massive parallel tools capable of relating a set of input and output variables. They are modeled on the animal nervous system, which is a highly complex, nonlinear, parallel processing system. They are composed of a set of neurons, commonly organized in layers, and by connections which connect the neurons.

An example of a perceptron, which is the most common neuron, is shown in Fig. 1(a), where $\{x_i\}_{i=1}^{N}$ are the input signals, $\{w_i\}_{i=1}^{N}$ the weights, $b$ the bias and $y$ the output signal. Perceptron applies a linear combination of its inputs, obtaining the signal $v = \sum_{i=1}^{N} x_i \cdot w_i + b$, and then applies a function f, which is called the transfer function, to this intermediate signal v to obtain the output signal. Sigmoid functions are commonly used as the transfer function to give the perceptron a nonlinear behavior.

One of the most commonly used ANN structures is the Multi-Layer Perceptron (MLP), because of its capability to solve non-linearly separable classification problems and to approximate continuous functions (Haykin, 1999). MLP topology comprises an input layer, one or more hidden layers, and an output layer, as shown in Fig. 1(b). The adjustment of its weights and biases is performed in a supervised way in the training stage, providing a set of pairs of input–output values, which allow the MLP to learn the relations between the input and the output variables. The training stage is usually done with the backpropagation (BP) method, in which the error of the MLP is calculated for each input–output pair and then this error is propagated from the output layer to the input layer, proportionally modifying the weights and biases of the MLP to the error committed by its neuron.

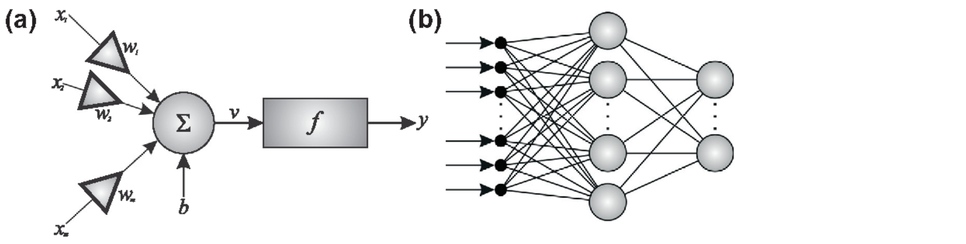



**Fig. 1.** (a) Example of a perceptron, where *x* are the input signals, *w* the weights, *b* the bias, *f* the transfer function and *y* the output signal. (b) Generic structure of a MLP with a single hidden layer.

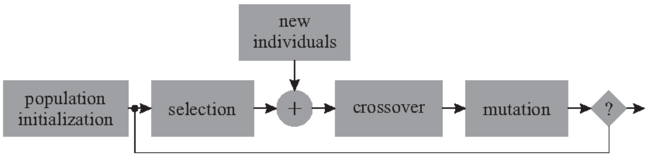



**Fig. 2.** Flux diagram of a generic GA.

*2.2. Genetic Algorithm*

A GA is an iterative search heuristic based on the process of natural selection and it is employed to solve optimization problems efficiently. GA belongs to the larger class of evolutionary algorithms and it combines randomized information exchange methods, which incorporate the survival of the fittest strategy, to find optimal solutions to the problem. The design of a GA involves the choice of the way the solutions should be represented, the design of a function which evaluates the fitness of the solution, the choice of the selection strategy to evolve the population, and the design of the information exchange methods (Arango, Cortés, Escudero, & Onieva, 2013; Goldberg, 1989; Michalewicz, 1996).

An example of typical GA is shown in Fig. 2. It contains one initialization stage, which executes itself once only the first time, and a set of stages that execute themselves iteratively once each generation. These stages are explained below.

The population initialization stage consists of the random generation of individuals, which will start the iterative process of the GA.

The fitness function must represent in terms of a single figure the profit, utility or goodness that the GA is going to maximize. As this function depends on the optimization problem, its design will be a different and specific task for each problem.

The selection stage consists of selecting some of the individuals of the population, discarding the remaining individuals for the next stages of the GA. The selection stage must consider two inversely related factors: the population diversity or exploration tradeoff and the selective pressure or exploitation tradeoff. One selection strategy, which does not take into account the exploitation tradeoff, consists of selecting the individuals from one generation to another randomly. Another strategy, which only takes into account the exploitation tradeoff, consists of selecting the individuals with the highest aptitude. An intermediate solution is the *roulette-wheel selection method*, in which each individual has a probability, which is directly proportional to its fitness, of selection for the next generation, thereby giving more opportunities to the individual with better fitness, but also giving some chances to the individuals with low fitness.

After the selection stage, new individuals are added. The purpose of doing so is to avoid searching only around the neighborhood of the best individuals, but of searching through all of the solution space.

The crossover stage consists of combining two individual chromosomes, which exchange certain information between each other. The mutation stage consists of the random modification of some of the elements of one chromosome. The crossover and the mutation stages generate new individuals, respectively, by combining two individuals or by generating a new one in the neighborhood of the initial individual respectively. As the initial individuals in both cases are usually good, these genetic operators are considered to be a suitable way to generate new solutions that are potentially effective at solving this problem.

Finally, the conditional finish condition is checked at the end of each generation, in order to decide whether the iterative process should continue. Typically, the fitness of the best individual, the number of iterations, and the execution time are the parameters that are analyzed, when deciding to finalize the GA.

## 3. Materials and methods

This section presents the materials and methods employed in the three experimental stages which comprise this article: data acquisition stage, preprocessing stage, and ANN processing stage. Fig. 3 presents the substages that comprise each stage, which are explained in detail in the subsequent subsections.

*3.1. Data acquisition stage*

Vibration data employed were recorded from a 3800 working hours New Holland TC56 harvester with the harvester stopped. A Kistler 8690C50 accelerometer was used to measure the vibration signal. It was positioned on the left-hand side of the chassis of the harvester, neither too close nor too far from the rotary machinery under study, as it can be seen in Fig. 3. The vibration signals were acquired using the NI Sound and Vibration Assistant software and a National Instruments (NI) data acquisition (DAQ) system, composed of a NI 9234 data-acquisition module for analog input

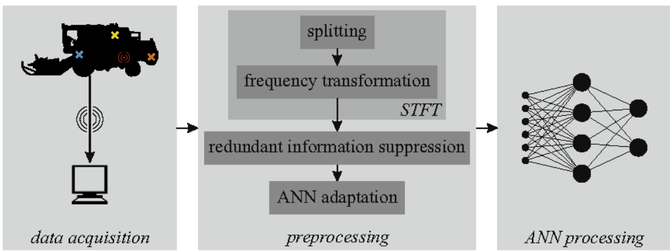



**Fig. 3.** Processing schema employed in this study, showing the three proposed processing stages: data acquisition stage, preprocessing stage, and ANN processing stage. The blue cross on the harvester represents the location of the threshing, the yellow one the location of the engine, the orange one the location of the straw, and the red symbol shows the location of the accelerometer sensor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

signals and a NI compact DAQ chassis NI cDAQ-9172 to connect the DAQ module with a laptop.

Several operation conditions were simulated in the harvester, in order to acquire data under all the specified conditions: (i) velocity of the engine varied between slow motion and maximum velocity; (ii) threshing and chopper varied between the on and off status; and, (iii) threshing and chopper worked balanced and unbalanced when they were in the on status. Thereby, a total of 18 different data acquisition processes were performed, recording 174.2238 s of machine operation, using a frequency sample of 1706.48 Hz to obtain a total of 297,360 samples.

### 3.2. Preprocessing stage

The time-series data were adapted in the preprocessing stage for the ANN processing stage. This process consisted of the 4 substages represented in Fig. 3: splitting, frequency transformation, redundant information suppression, and ANN adaptation.

The splitting substage created data sequences of length $N$ from the raw data. The $N$ parameter adjusts the frequency separation of two consecutive samples ($\Delta f = \frac{f_s}{N}$) of the discrete Fourier transform (DFT) of a sequence, having an inversely proportional relation between $N$ and $\Delta f$. A sequence length of 512 was chosen, obtaining a frequency resolution of $\Delta f$ = 3.3330 Hz.

The frequency transformation substage consisted of calculating the discrete Fourier transform (DFT) of the input sequences and refining the frequency data. In this case, DFT was computed using the fast Fourier transform (FFT) algorithm, obtaining an output signal of the same length as the input sequence. Both splitting and frequency transformation substages represent a particular case of a short-time Fourier transform (STFT), where the window is a rectangular function of length $N$ and the overlapping length is zero.

The redundant information suppression substage deleted the duplicated information of the sequence. As the vibration signal is a real signal, the DFT of the splitted sequence is a symmetric signal, due to the DFT properties. For this reason, the last $N/2 - 1$ points (where $N$ is an even number) of the output signal were discarded in this substage to remove the redundant information.

The ANN adaptation substage consisted of two tasks. The first one consisted of calculating the absolute value of the sequences, so as to use real numbers in the ANN processing stage. The second one consisted of applying a linear transformation of the input samples of the ANN, to produce variables with zero mean and unity standard deviation. This task was included to improve the performance of the fitted ANN, because ANN training algorithms are designed for normalized input signals.

### 3.3. ANN processing stage

The preprocessed data was processed with a three layer feedforward MLP. The input layer had 256 neurons, which is the length of the sequences obtained after the preprocessing stage, while the output layer had 5 neurons, each related to the 5 operational statuses under consideration: (i) the engine speed status (high speed/low speed), (ii) the threshing operating status (on/off), (iii) the threshing balance status (balanced/unbalanced), (iv) the straw chopper operating status (on/off), and (v) the straw chopper balance status (balanced/unbalanced). The number of neurons in the hidden layer was varied in the experiments, in order to choose the number which obtained good results without over-fitting the network. The transfer function for both the hidden layer and the output layer was the *tan-sigmoid* function, setting a threshold of zero in the output layer to generate the Boolean output.

Four learning algorithms were employed to fix the weights and biases of the ANN using the preprocessed data. To this end, data was divided into three datasets: the learning dataset, the validation dataset, and the testing dataset, which contained the 45%, 15%, and 40% of the samples of each of the 18 experiments performed respectively. The learning dataset was employed by the learning algorithms to adjust the weights and biases of the ANNs, and a validation dataset was used to verify if the learning process had reached the algorithm finish condition and to analyze the performance of the learning algorithm, and a testing dataset was employed to analyze the performance of the ANNs after the learning process.

Learning and testing procedures are explained below.

#### 3.3.1. Learning

Four learning methods were considered to adjust the weights and biases of the proposed ANNs. The first one randomly adjusted the weights and biases using a Monte Carlo (MC) sampling technique (Liu, 2008). The second one adjusted the weights and biases using a generic GA. The third one initialized the ANN weights and biases using a MC sampling technique and trained the initialized ANN with the BP method. The fourth one initialized the weights and biases using a GA and trained the initialized ANN with the BP algorithm. The purpose of using a GA in the second and fourth learning methods is to develop an efficient search through the multidimensional space of weight and biases, to find an optimal solution of the ANN or an optimal ANN initialization, respectively in less time than using a MC sampling technique. The four proposed learning methods (LMs) were respectively labeled *MC LM*, *GA LM*, *MC&BP LM*, and *GA&BP LM* for ease of reference. Table 1 summarizes the names of the four learning methods proposed and their respective characteristics.

Four GAs were developed to implement the four proposed learning methods to compare them under the same conditions. For this reason, the same chromosomes, fitness function, selection method, genetic operators, and finish conditions were used in the four GAs that were developed.

The chromosomes that represented the solutions consisted of two matrixes: the first one ($H$) contained the weights and biases of the hidden layer neurons, and the second one ($O$) contained the weights and biases of the output layer neurons. Each matrix has as many rows as neurons of the layer and one more column than the number of neurons of the previous layer. Eq. (1) represents a chromosome, where $w_H(i,j)$ and $w_O(i,j)$ are the weight of the connection between the $i$th neuron of the current layer and the $j$th neuron of the previous layer for the hidden layer ($H$) and the output layer ($O$), respectively.

$$H = \begin{bmatrix} w_H(1,1) & \dots & w_H(1,n_I) & b_H(1) \\ \vdots & \ddots & \vdots & \vdots \\ w_H(n_H,1) & \dots & w_H(n_H,n_I) & b_H(n_H) \end{bmatrix} \quad O = \begin{bmatrix} w_O(1,1) & \dots & w_O(1,n_H) & b_O(1) \\ \vdots & \ddots & \vdots & \vdots \\ w_O(n_O,1) & \dots & w_O(n_O,n_H) & b_O(n_O) \end{bmatrix} \tag{1}$$

The fitness of every individual was calculated as the mean success rate of the ANN built with the weights and biases of this individual for the validation dataset.

Selection of all but one of the individuals from one generation to the next generation was done using the *roulette-wheel selection*

**Table 1**
Characteristics and names summary of the four proposed learning methods (LMs).

| Learning method abbreviation | Weights and biases adjustment procedure |
|---|---|
| *MC LM* | MC sampling technique |
| *GA LM* | Generic GA |
| *MC&BP LM* | MC sampling technique (ANN initialization) + BP (ANN training) |
| *GA&BP LM* | GA (ANN initialization) + BP (ANN training) |

V. Martínez-Martínez et al. / Expert Systems with Applications 42 (2015) 6433–6441 

*method*, where the probability of selection for the next generation was directly proportional to the aptitude of the individual. The selected individual without taking this method into account was the best individual of the previous generation, which was always selected for the next generation, in order to ensure its survival.

Genetic operators considered for the proposed Genetic Algorithms were crossover and mutation. Crossover consisted of exchanging a row of one matrix ($H$ or $O$) from one individual to the same row of the same matrix of another individual of the generation. Mutation consisted of randomly modifying one element of one of the matrixes which define a chromosome of a concrete individual. From the ANN perspective, the crossover operation consisted of exchanging the weights and bias of one neuron from one individual to another and the mutation operation consisted of modifying the weight and/or bias value of one or more neurons of the ANN.

The finish condition for the proposed GAs was the maximum number of iterations, in order to compare the different performance of the GAs after the same number of generations of evolution.

The GAs designed to implement the *MC LM* and the *GA LM* are a particularization of the GA structure presented in Fig. 2, while the *MC&BP LM* and the *GA&BP LM* have the GA structure shown in Fig. 4.

The GA configuration and the parameters selected to implement the *GA LM* and the *GA&BP LM* were selected from among several options with a trial and error strategy. These GAs have a population of 5 individuals, with 40% of new random individuals in each generation, 40% of the generation individuals affected by crossover, and 20% affected by mutation with 5% probability of mutation for each weight and bias. The GAs which implemented the *MC LM* and the *MC&BP LM* also have a population of 5 individuals, with 100% of new random individuals in each generation and no one of the individuals affected by the genetic operators. The ANN training stage was implemented using the BP method with the Levenberg–Marquardt algorithm.

*3.3.2. Testing*

The performance of the ANNs after the learning process was evaluated with the testing dataset, calculating its success rate for each of the five parameters under consideration, which are the five outputs of the ANN. These five success rates will show the precision of the ANN estimating each of the five aforementioned statuses.

Moreover, the performance of the four AG-based ANN learning methods was analyzed using the evolution of the mean and the maximum fitness of each generation. It allows us to analyze the capability to find suboptimal solutions of the proposed methods within a limited number of iterations.

## 4. Results

The numerical results of the proposed GA-based learning method and the ANN-based classification system are presented in this section.

### *4.1. GA-based learning method results*

The performance of the learning method was analyzed for a MLP with 5 neurons in the hidden layer. This analysis consisted of measuring the mean and the maximum fitness in each generation for each of the four proposed GA-based learning methods. The learning process was repeated 20 times for each GA and the mean values were calculated, in order to minimize the randomness of the results. Moreover, the maximum number of generations used as finish conditions in the four GAs was 5000 for the *MC LM* and the *GA LM*, and 100 for the *MC&BP LM* and the *GA&BP LM*.

Fig. 5 shows the mean and maximum success rate for each generation versus the number of generations for the *MC LM* and the *GA LM*. It can be appreciated that the maximum success rate for the *GA LM* is greater than 2% better than the *MC LM* after the 500th generation. Moreover, the trend of the mean success rate for the *GA LM* is also around 2% better compared with the *MC LM*.

Fig. 6 shows the number of generations versus the maximum success rate for each generation for the *MC LM* and the *GA LM* and the relative difference between them. It can be seen that the number of generations needed to obtain a method with a success

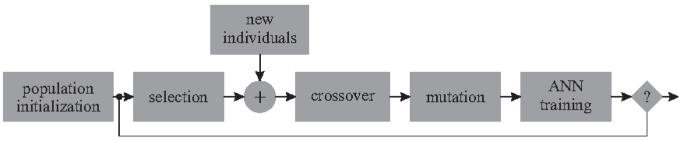



**Fig. 4.** Flux diagram of GA used to implement the *MC&BP LM* and the *GA&BP LM*, where the weights and biases of the ANN are initialized with a generic GA before the ANN training.

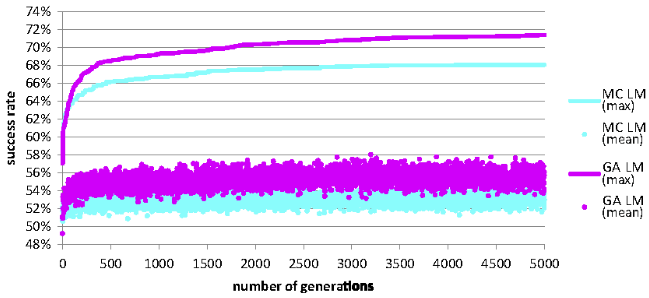



**Fig. 5.** Mean and maximum success rate versus the number of generations for the *MC LM* (cyan) and the *GA LM* (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

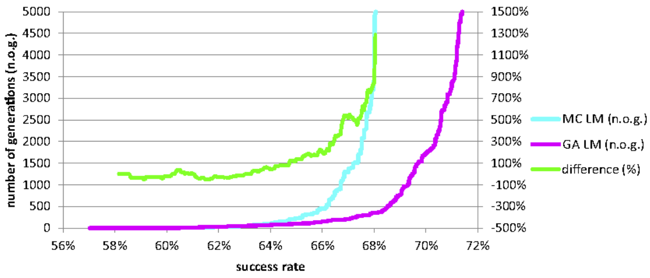



**Fig. 6.** Number of generations versus the maximum success rate for the *MC LM* (cyan) and the *GA LM* (magenta), and relative difference (green) between them. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

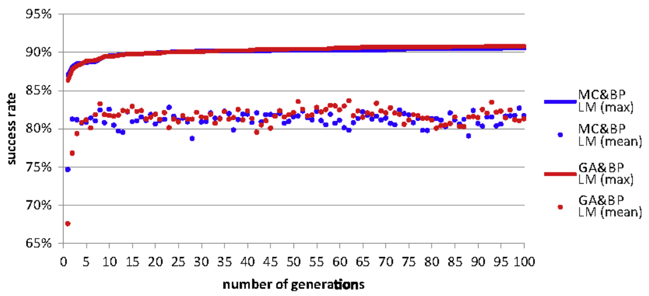



**Fig. 7.** Mean and maximum success rate versus the number of generations for the *MC&BP LM* (blue) and the *GA&BP LM* (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

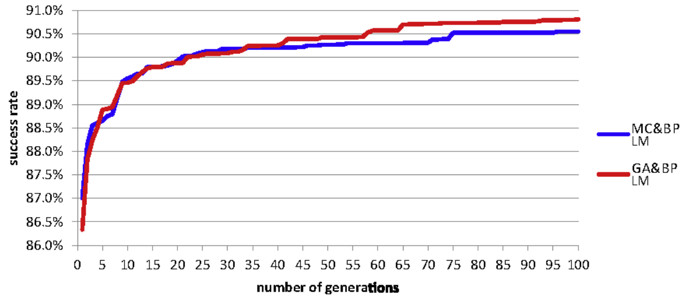



**Fig. 8.** Maximum success rate versus the number of generations for the *MC&BP LM* (blue) and the *GA&BP LM* (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

rate higher than 64% is visibly greater for the *MC LM* in comparison with the *GA LM*. Moreover, this relative difference increases as the target success rate rises to as high as 1300% for a success rate of 68%.

Fig. 7 shows the mean and maximum success rate for each generation versus the number of generations for the *MC&BP LM* and the *GA&BP LM*, while Fig. 8 shows in detail the maximum success rate for each generation for the same GAs. In both figures, the maximum success rate for the *GA&BP LM* is shown to be about 0.2% better than the *MC&BP LM* after the 40th generation, while there is no visible difference in the mean success rate.

Fig. 9 shows the number of generations versus the maximum success rate of each generation, for the *MC&BP LM* and the *GA&BP LM*, and the relative differences between them. It can be seen that the number of generations needed to obtain a method with a success rate higher than 90.25% is between 20% and 70% greater for the *MC&BP LM* in comparison with the *GA&BP LM*. Moreover, the maximum success rate reached with the *GA&BP LM* after 100 generations is higher than that reached with the *MC&BP LM*.

### *4.2. ANN-based classification system results*

Nine ANN-based classifiers with different number of neurons in the hidden layer were tested after adjusting their weights and biases using the *GA&BP LM*. The results obtained by each classifier are shown in Table 2, presenting the mean success rate of each ANN and the specific success rate for each operational status considered.

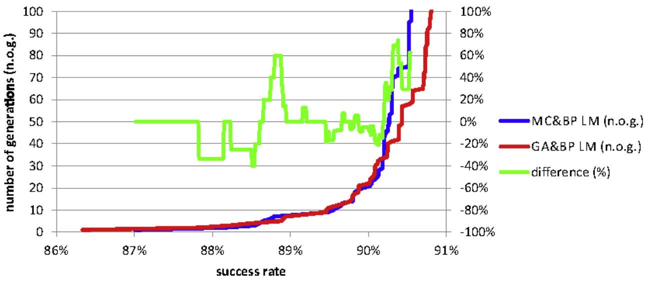



**Fig. 9.** Number of generations versus the maximum success rate for the *MC&BP LM* (blue) and the *GA&BP LM* (red), and relative differences (green) between them. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 2**
Classification results with different number of neurons in the hidden layer.

| Hidden neurons | Mean success rate (%) | Engine speed status success rate (%) | Threshing operating status success rate (%) | Threshing balance status success rate (%) | Straw chopper operating status success rate (%) | Straw chopper balance status success rate (%) |
|---|---|---|---|---|---|---|
| 2 | 80.52 | 96.52 | 88.70 | 57.83 | 77.39 | 82.17 |
| 3 | 86.70 | 98.26 | 87.83 | 72.17 | 90.87 | 84.35 |
| 4 | 91.13 | 100 | 91.30 | 76.09 | 95.65 | 92.61 |
| 5 | 90.87 | 98.70 | 92.17 | 76.09 | 94.35 | 93.04 |
| 6 | 92.00 | 99.57 | 94.78 | 77.39 | 92.17 | 96.09 |
| 7 | 90.78 | 100 | 93.91 | 76.52 | 92.61 | 90.87 |
| 8 | 92.96 | 100 | 95.22 | 80.00 | 96.52 | 93.04 |
| 9 | 91.57 | 100 | 93.91 | 73.48 | 95.22 | 95.22 |
| 10 | 92.52 | 100 | 92.61 | 77.83 | 96.96 | 95.22 |

## 5. Discussion

Previously presented results show the potential of the proposed ANN to accurately predict the operational statuses of the rotary components of a harvester and the ability of the proposed AG-based learning methods to adjust the weights and biases of a MLP ANN. The main application field for the proposed ANN-based expert system is to perform the prediction stage of a predictive maintenance system.

The ANN-based expert system to classify the status of the rotary machinery proposed in this article showed good accuracy. The mean success rate of the proposed ANN was greater than 90% when the number of hidden neurons was greater than three, with a maximum success rate of 92.96% for eight hidden neurons. Analyzing the success rate obtained for each rotary component, the suitability of the proposed ANN to estimate each separate status of the machinery is evident. On the one hand, the status of the rotary machinery with the best success rate was engine speed, with a success rate of between 98.70% and 100% when the number of hidden neurons was greater than three. On the other hand, the worst success rate was obtained by the threshing balance status, the success rate for which was between 73.48% and 80% when the number of hidden neurons was greater than three. Analyzing the vibration signal spectrum, there are differences in the signal, when the engine speed is varied between a high and a low speed, while there are not visible changes to the signal spectrum when the threshing is balanced or unbalanced. These results show that the proposed ANN is able to classify the status of rotary machinery to a high degree of accuracy when the difference between the spectrum signals is visible, such as the engine speed. They also show that it can obtain acceptable success rates for rotary machinery where there is no visible difference between the spectrum signals, such as the threshing balance status.

There is an aspect of the success rate that should be considered, if this ANN is to be applied in a real predictive maintenance application. As the proposed ANN has 5 outputs, the prediction method was optimized to obtain the best mean success rate, but not the best success rate for each status of the rotary machinery under consideration. This can be seen in Table 2, where the mean value of the best success rate for each status (93.65%) is 0.69% higher than the best mean success rate (92.96%). For this reason, a better success rate will be expected if independent prediction methods designed *ad-hoc* for each status are implemented. Nevertheless, implementing independent prediction methods has the disadvantage of requiring five prediction methods instead of one, which requires more processing and memory resources.

Four GA-based learning methods were proposed in this article. Two of them, *GA LM* and *GA&BP LM*, were designed to optimize the ANN weights and biases adjustment process. The other two learning methods, *MC LM* and *MC&BP LM*, implemented a Monte Carlo sampling technique in order to use their results as the reference to be compared with the other two learning methods. Two comparisons were performed in this study, obtaining several observations. The first comparison analyzed the GAs which simulated learning techniques without using the BP method to train the ANN: *GA LM* and *MC LM*. The results showed that *GA LM* noticeably improves the performance obtained by the random procedure implemented by *MC LM*. This is because the *GA LM* searches through a population of possible solutions with a better fitness than the *MC LM*, as other authors have shown in the literature (Lu et al., 2015; Polat & Yıldırım, 2008a; Soares, Antunes, & Araújo, 2013). The mean success rates presented in Fig. 5 show this improvement. The reason why the mean success rate remains constant after the 200th iteration is because of the high percentage of new individuals added to each generation, which limits the mean fitness trend. The second comparison analyzed the GAs which used the BP method to train the ANN after initializing it: *GA&BP LM* and *MC&BP LM*. In this comparison, the improvement of the GA-based method with respect to the MC-based method was lower than the one obtained in the first comparison. This may be because the BP training method contributes to the success rate more than the GA-based initialization of the ANN. Fig. 7 justifies this assertion, because there is not a significant difference between the mean success rate trend for *MC&BP LM* and *GA&BP LM*. Nevertheless, Fig. 9 shows that a saving of 70% in the number of generations can be obtained if *GA&BP LM* is used as the learning method, which makes its implementation worthwhile. Even though the number of generations is the parameter that is analyzed to compare the proposed GAs, it is worth mentioning the direct relationship between the number of generations and the execution time of a GA. Because of this direct relationship, the comparison results obtained in our article for the number of generations are expected to be similar, if the execution time rather than the number of generations is compared.

The main strength of the proposed prediction method is its generalization capability, which allows to apply it to different machines and to consider other rotary machinery, without needing an expert to design the prediction system and to supervise it. This generalization capability has been achieved because the spectrum coefficients of the vibration signal were used as the inputs of the ANN instead of features of the vibration signal. Not implementing a feature extraction stage implies that no previous knowledge on the signal or the elements which generated the signal was employed, which justifies the generalization capability of the prediction method. As mentioned in the Introduction section, there are few articles in the literature working with ANNs without a feature extraction stage (Erkmen & Yıldırım, 2008; Polat & Yıldırım, 2008a, 2008b; Pradeep et al., 2011), none of which are related to either the vibration signal analysis or to predictive maintenance. For this reason, the novelty of this prediction method is remarkable.

In addition to the generalization capability, another noteworthy strength of the proposed prediction method is that only one vibration measurement point is needed even though several statuses of rotary components placed far from each other are considered. Some recent works in the literature that evaluate the status of elements using a vibration signal obtained precision values between 90% and 98% (Chwan-Lu et al., 2014; Li et al., 2013; Paulraj, Yaacob, Majid, Kazim, & Krishnan, 2013), values that are numerically in the range of the results obtained in this article. Nevertheless, our method obtained the results presented in Table 2 having harder conditions to predict, because, from the data of a single sensor, either (i) our work estimated the status of some rotary components whereas other articles in the literature estimated the status of a single rotary component (Chwan-Lu et al., 2014; Chwan-Lu, Shun-Yuan, Shou-Chuang, et al., 2014; Liang et al., 2014; Yang, Dong, Peng, Zhang, & Meng, 2015) or (ii) our work estimated the statuses of several rotary components placed far from each other whereas other articles in the literature estimated the statuses of several rotary components placed close to each other (Bin et al., 2012; Li et al., 2013; Lu et al., 2015; Nembhard, Sinha, & Yunusa-Kaltungo, 2015). This makes our results promising because they are in the range of the results obtained in articles that consider easier estimation conditions.

There are some limitations which should be taken into account before implementing the proposed prediction method in a predictive maintenance system. One drawback of the proposed prediction method could be the high number of ANN inputs, which is higher than the needed if a feature extraction strategy had been used. This size increment makes the weight and biases adjustment process slower and more difficult to optimize. Nevertheless, this drawback is overcome by using *GA&BP LM* to reduce the time needed to adjust the weights and biases and to improve the performance of the resulting ANN, which is another notable aspect of this proposal. Another weakness of this article is related to the data acquisition performed to validate the proposed ANN-based expert system and the proposed learning methods. On the one hand, the vibration signal was acquired with the harvester stopped, to facilitate the acquisition procedure. If the proposed prediction method is used when the monitored machine is in motion, low-frequency interference signals could appear, but they are not expected to be a problem because the frequencies of interest in the rotary components of these machines are usually higher than the interference frequencies. On the other hand, rotary component breakdowns were simulated by adding an eccentric weight to the rotary axes to provoke unbalances in the rotary components that mimic real breakdowns. Nevertheless, the proposed ANN-based prediction method is also expected to achieve good results with other kinds of breakdowns that are not related to rotary machinery.

There are some aspects about this work which can be improved or extended in the future. The first one is to apply the proposed method to other machines or to other rotary components. The second one is to perform the same experiments with the machine in operating conditions in order to check if the rotary components statuses can be estimated when the machine is working in the land. The third one is to implement an independent method for each rotary component, in order to optimize the methods and obtain better results than the obtained with a single general method. Finally, the last aspect which will be interesting to analyze is the use of more than one sensor in different locations.

## 6. Conclusions

The results that have been obtained in this study suggest that, by processing a vibration signal acquired from a single point of an agro-industrial machine structure, it is possible to estimate the status of its rotary components with a mean success rate that is higher than 90% with a three layer MLP ANN-based expert system. Moreover, by using the best proposed GA-based learning method to adjust the ANN weights and biases, it is possible to improve the success rate and to reduce by 70% the time needed to perform the adjustment, compared to the typical backpropagation method for the ANN adjustment. The main contribution of this work to the Expert and Intelligent Systems field is the proposal of a classification method that estimates the status of several rotary elements placed each one far from the others employing the signal acquired from only one accelerometer and non-requiring a feature extraction stage.

## Acknowledgments

The first author's work was made possible thanks to a *Formación de Personal Investigador* program grant. This program was financed by the *Universidad de Valladolid* (Spain) and co-financed by *Banco Santander*. Fourth author's work was also financed during this article's research by a collaboration grant from the *Ministry of Education, Culture and Sport* of Spain, with reference number *BOE-A-2013-7736*.

## References

Albarbar, A., Gu, F., & Ball, A. D. (2010). Diesel engine fuel injection monitoring using acoustic measurements and independent component analysis. *Measurement, 43*, 1376–1386.

Arango, C., Cortés, P., Escudero, A., & Onieva, L. (2013). 17 – Genetic algorithm for the dynamic berth allocation problem in real time. In X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, & M. Karamanoglu (Eds.), *Swarm intelligence and bio-inspired computation* (pp. 367–383). Oxford: Elsevier.

Bin, G. F., Gao, J. J., Li, X. J., & Dhillon, B. S. (2012). Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature

extraction and neural network. *Mechanical Systems and Signal Processing, 27*, 696–711.
Chen, F., Tang, B., & Chen, R. (2013). A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm. *Measurement, 46*, 220–232.
Cheng, G., Cheng, Y.-L., Shen, L.-H., Qiu, J.-B., & Zhang, S. (2013). Gear fault identification based on Hilbert–Huang transform and SOM neural network. *Measurement, 46*, 1137–1146.
Chia-Feng, J. (2004). A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 34*, 997–1006.
Chwan-Lu, T., Shun-Yuan, W., Foun-Yuan, L., Jen-Hsiang, C., Yin-Hsien, S., & Ta-Peng, T. (2014). An intelligent motor rotary fault diagnosis system using Taguchi method. In *2014 IEEE international conference on systems, man and cybernetics (SMC)* (pp. 2311–2316).
Chwan-Lu, T., Shun-Yuan, W., Shou-Chuang, L., Jen-Hsiang, C., & Ke-Fan, C. (2014). A motor rotary fault diagnosis system using dynamic structural neural network. In *2014 international symposium on computer, consumer and control (IS3C)* (pp. 430–433).
Erkmen, B., & Yıldırım, T. (2008). Improving classification performance of sonar targets by applying general regression neural network with PCA. *Expert Systems with Applications, 35*, 472–475.
Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.
Haykin, S. S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). New Jersey, USA: Prentice Hall.
Huang, H.-X., Li, J.-C., & Xiao, C.-L. (2015). A proposed iteration optimization approach integrating backpropagation neural network with genetic algorithm. *Expert Systems with Applications, 42*, 146–155.
Jayaswal, P., Verma, S., & Wadhwani, A. (2011). Development of EBP-Artificial neural network expert system for rolling element bearing fault diagnosis. *Journal of Vibration and Control, 17*, 1131–1148.
Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing, 35*, 108–126.
Leung, F. H. F., Lam, H. K., Ling, S. H., & Tam, P. K. S. (2003). Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Transactions on Neural Networks, 14*, 79–88.
Li, Y., Tse, P. W., Yang, X., & Yang, J. (2010). EMD-based fault diagnosis for abnormal clearance between contacting components in a diesel engine. *Mechanical Systems and Signal Processing, 24*, 193–210.
Li, Z., Yan, X., Tian, Z., Yuan, C., Peng, Z., & Li, L. (2013). Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis. *Measurement, 46*, 259–271.
Liang, S. Y., Li, Y., Billington, S. A., Zhang, C., Shiroishi, J., Kurfess, T. R., & Danyluk, S. (2014). Adaptive prognostics for rotary machineries. *Procedia Engineering, 86*, 852–857.
Liu, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer.
Lu, N., Xiao, Z., & Malik, O. P. (2015). Feature extraction using adaptive multiwavelets and synthetic detection index for rotor fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing, 52–53*, 393–415.
McFadden, P. D., Cook, J. G., & Forster, L. M. (1999). Decomposition of gear vibration signals by the generalised S transform. *Mechanical Systems and Signal Processing, 13*, 691–707.
Michalewicz, Z. (1996). *Genetic algorithms + data structures = evolution programs*. Springer.
Nembhard, A. D., Sinha, J. K., & Yunusa-Kaltungo, A. (2015). Development of a generic rotating machinery fault diagnosis approach insensitive to machine speed and support type. *Journal of Sound and Vibration, 337*, 321–341.
Paulraj, M. P., Yaacob, S., Majid, M. S. A., Kazim, M. N. F. M., & Krishnan, P. (2013). Structural steel plate damage detection using non destructive testing, frame energy based statistical features and artificial neural networks. *Procedia Engineering, 53*, 376–386.
Polat, Ö., & Yıldırım, T. (2008a). Genetic optimization of GRNN for pattern recognition without feature extraction. *Expert Systems with Applications, 34*, 2444–2448.
Polat, Ö., & Yıldırım, T. (2008b). Hand geometry identification without feature extraction by general regression neural network. *Expert Systems with Applications, 34*, 845–849.
Pradeep, J., Srinivasan, E., & Himavathi, S. (2011). Neural network based handwritten character recognition system without feature extraction. In *2011 international conference on computer, communication and electrical technology (ICCCET)* (pp. 40–44).
Rodriguez-Donate, C., Romero-Troncoso, R., Cabal-Yepez, E., Garcia-Perez, A., & Osornio-Rios, R. (2011). Wavelet-based general methodology for multiple fault detection on induction motors at the startup vibration transient. *Journal of Vibration and Control, 17*, 1299–1309.
Soares, S., Antunes, C. H., & Araújo, R. (2013). Comparison of a genetic algorithm and simulated annealing for automatic neural network ensemble development. *Neurocomputing, 121*, 498–511.
Song, C.-S., Hu, Y., Xie, S., & Zhou, Z. (2013). Dynamic modeling of magnetic suspension isolator using artificial neural network: a modified genetic approach. *Journal of Vibration and Control, 19*, 847–856.
Taghizadeh-Alisaraei, A., Ghobadian, B., Tavakoli-Hashjin, T., & Mohtasebi, S. S. (2012). Vibration analysis of a diesel engine using biodiesel and petrodiesel fuel blends. *Fuel, 102*, 414–422.
Vulli, S., Dunne, J. F., Potenza, R., Richardson, D., & King, P. (2009). Time-frequency analysis of single-point engine-block vibration measurements for multiple excitation-event identification. *Journal of Sound and Vibration, 321*, 1129–1143.
Wang, X., Makis, V., & Yang, M. (2010). A wavelet approach to fault diagnosis of a gearbox under varying load conditions. *Journal of Sound and Vibration, 329*, 1570–1585.
Wang, Y. S., Ma, Q. H., Zhu, Q., Liu, X. T., & Zhao, L. H. (2014). An intelligent approach for engine fault diagnosis based on Hilbert–Huang transform and support vector machine. *Applied Acoustics, 75*, 1–9.
Xin, Y., & Yong, L. (1997). A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks, 8*, 694–713.
Yan, R., Gao, R. X., & Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: a review with applications. *Signal Processing, 96*(Part A), 1–15.
Yang, Y., Dong, X. J., Peng, Z. K., Zhang, W. M., & Meng, G. (2015). Vibration signal analysis using parameterized time–frequency method for features extraction of varying-speed rotary machinery. *Journal of Sound and Vibration, 335*, 350–366.
Yildirim, Ş., Erkaya, S., Eski, İ., & Uzmay, İ. (2009). Noise and vibration analysis of car engines using proposed neural network. *Journal of Vibration and Control, 15*, 24.

# Article 4: Evaluation of steel pieces in a production line

The fourth article presents the **ANN-based model for the NDT of steel pieces** described in the "Introduction and Summary" part of the document.

The article presented below is a version that is going to be submitted to a JCR-indexed peer-reviewed journal. For this reason, there is not bibliographic data about this article in this introduction.

# Radial Basis Function Neural Network applied to the classification of heat-treated steel pieces by multifrequency nondestructive eddy current testing

**Abstract:** This article proposes a Radial Basis Function Artificial Neural Network (RBF-ANN) to classify heat-treated steel pieces using multifrequency nondestructive testing data. Eddy current testing impedances at five frequencies between 10 and 300 kHz were employed to perform the classification after checking their suitability using the ANalysis Of VAriance (ANOVA) test. Afterwards, twelve classifiers were implemented and compared: one RBF-ANN classifier, ten linear discriminant analysis classifiers, and one Euclidean distance classifier. The proposed RBF-ANN achieved the best performance, with a precision of 95% and an area under the Receiver Operating Characteristic (ROC) curve of 0.98. The results obtained suggest RBF-ANN classifiers processing multifrequency impedance data to classify heat-treated steel pieces with a performance better than other classic classifiers.

**Keywords:** nondestructive testing; eddy currents; radial basis function neural networks; heat treatment; analysis of variance.

## 1. Introduction

Nondestructive testing (NDT) techniques are analysis methods employed to inspect pieces without causing them permanent damage or modifications. Eddy current testing has been one of the most popular NDT techniques, and has been implemented as an alternative to metallography [1] and indentation hardness tests [2]. Eddy current testing is based on the Faraday's electromagnetic induction law, which was proposed in 1831, and on the Hughes experiments, which were performed in 1879. Hughes found changes in the properties of a coil when it was approached to metals with different electric conductivities and magnetic permeabilities.

The usage of eddy current testing techniques has significantly increased since the 1950s in the aeronautical [3, 4] and nuclear [5, 6] industries, among others. In the late years, state of the art electronic components and processors have been applied to build computer-eddy current instrumentation [7, 8]. Moreover, eddy current testing techniques have been employed to analyze materials: Konoplyuk *et al.* [9] predicted the matrix microstructure in ductile cast irons, Mercier *et al.* [10] classified steel decarburizing pieces, and Wrzuszczak *et al.* [11] detected defects on conducting materials.

NDT techniques provide data that can be employed to decide if an analyzed piece of material has defects or damages. This decision can be made either by an expert person or by an expert system. Artificial Neural Networks (ANNs) are processing tools inspired in the human nervous system, which can implement expert systems using their input-output mapping capability [12]. Various authors have combined ANNs and NDT techniques in the literature: Wang *et al.* employed a back propagation ANN to monitor the stress and the temperature of steel using the Barkhausen noise theory [13], Junyan *et al.* employed a multilayer perceptron (MLP) to detect subsurface defects in different materials using thermography [14], Cao *et al.*

employed a Radial Basis Function ANN (RBF-ANN) to evaluate wire ropes with eddy current inspection [15], and Wrzuszczak *et al.* combined the eddy current information with ANNs to detect cracks on conducting layers and on ferrous tubes [11].

The objective of this article is to analyze the suitability of RBF-ANNs to classify heat-treated pieces of steel using multifrequency impedance data acquired from eddy current NDT [12] and to compare an RBF-ANN classifier with other classifiers. To this end: (i) multifrequency eddy current impedances were acquired from two DIN 100Cr6 steel cam sets with different heat treatments; (ii) ANalysis Of Variance (ANOVA) test was used to analyze the suitability of the acquired impedances to classify the steel pieces; (iii) classification was performed with RBF-ANN, Linear Discriminant Analisys (LDA), and Euclidean distance classifiers; and (iv) classification results were compared considering the precision and the area under the ROC curve ($A_{ROC}$).

## 2. Theoretical background

Eddy currents, which are also called Foucault currents, are electric currents induced by a variable magnetic field in a conductor. According to the Faraday's law, an eddy current flow is generated in a piece of steel when a coil fed with AC current is approached as Fig. 1 shows [16-18].

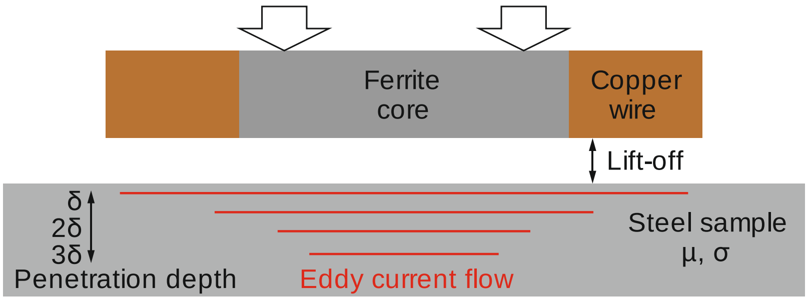



**Fig. 1.** Eddy current flow generated by approaching a ferrite-inductive coil probe to a steel sample. The main parameters involved are the lift-off, the penetration depth, the magnetic permeability ($\mu$), and the electrical conductivity ($\sigma$).

The magnitude and the phase of the induced eddy currents affect the loading of the coil and thus its global impedance $Z_{coil}$. The impedance of the coil is also affected by the lift-off, which is the separation between the coil and the steel sample. This impedance can be measured using an eddy-current instrumentation device [18, 19]. The coil excitation frequency $f$ permits the adjustment of the penetration depth $\delta$, which is proportional to $f^{0.5}$ according to the skin effect [20]. The acquired impedance of the coil $Z_{coil}$ is mainly related to the electrical conductivity $\sigma$, to the magnetic permeability $\mu$, to the lift-off, and to the applied and the residual stresses of the steel sample [2].

## 3. Materials

The work performed in this article was done using a coil probe, an instrumentation device, a kit of steel samples, and a computer (Fig. 2).

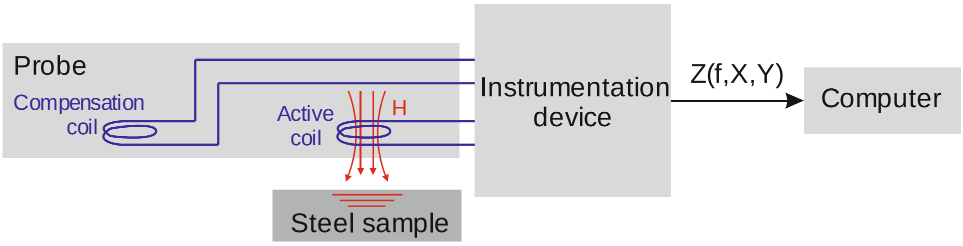



**Fig. 2.** Layout of the materials employed in this work.

The probe was composed by one active-coil (a-c) with impedance $Z_{a\text{-}c}$ and one compensation coil (c-c) with impedance $Z_{c\text{-}c}$. The operation mode chosen for the probe was the absolute with compensation mode in order to increase its dynamic range. A Wheatstone bridge as the one presented in Fig. 3 was employed to interconnect the probe with the instrumentation device. The resistance and inductance values at 100 kHz were respectively 12.893 Ω and 2.5888 mH for the active coil and 12.502 Ω and 2.5317 mH for the compensating coil. The coil probe was used to induce the eddy current flow in the steel samples and to acquire the $V_o$-$V_{ref}$ signal of Fig. 3.

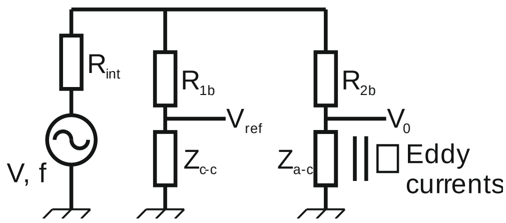



**Fig. 3.** Wheatstone bridge interconnection of the absolute with compensation coil probe with impedances $Z_{a\text{-}c}$ and $Z_{c\text{-}c}$ to the instrumentation device.

The instrumentation device was employed to feed the probe at the desired frequency *f* and amplitude *V*, and to acquire the eddy current signal, obtaining the real and imaginary parts of the impedance ($Re\{Z_{a\text{-}c}\}$ and $Im\{Z_{a\text{-}c}\}$ respectively) from the acquired the signal $V_o$-$V_{ref}$ (Fig. 3).

The kit of steel samples was composed by eight cams of steel DIN 100Cr6, being the first four cams forged in different conditions than the last four cams. The first four cams, which are referred to as Group 1, had a hardness greater than 55 HRC because they had been subjected to a rapid cooling in their manufacturing process. The predominant metallurgical structure of the pieces of Group 1 was martensite and bainite. The other four cams, which are referred to as Group 2, had a hardness lower than 45 HRC because they had been subjected to a cooling process slower than the cooling process employed for the pieces of Group 1. The predominant metallurgical structure of the pieces of Group 2 was perlite and bainite. This kit of samples was employed to extract the data employed to train and validate the classifiers proposed in this article.

The computer was a general purpose laptop with an Intel Core i3 M350 @ 2.27 GHz processor and 4 GB RAM. This computer was employed to implement and execute the data preprocessing stage, the statistical analysis stage, and the classification stage, presented in Methods section.

# 4. Methods

The methodology was comprised by five stages: the impedance acquisition stage, the data preprocessing stage, the statistical analysis stage, the classification stage, and the ROC analysis stage, as Fig. 4 shows.

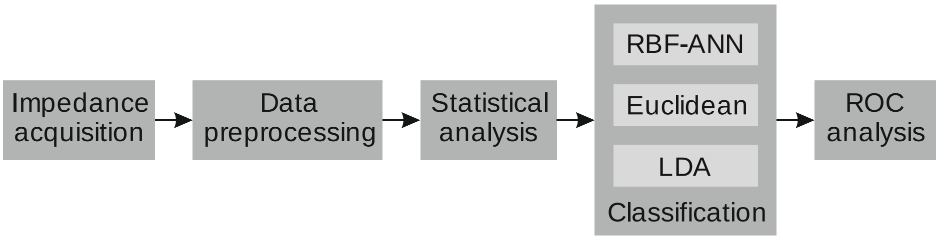



**Fig. 4.** Block diagram of methodology employed in this article.

## 4.1. Impedance acquisition stage

Data impedance was acquired at room temperature (23 ºC) by approaching the steel cams to the coil probe, which was powered at one of the five different operating frequencies considered: 300, 100, 50, 20, and 10 kHz. The impedance was acquired with a lift-off lower than 0.2 mm, acquiring 5 measures for each steel cam. These data were divided into two datasets: a training dataset, which contained the data of 50% of the cams, and the testing dataset, which contained the data of the other 50% of the cams. From this point forward, we will refer to these impedance components as 300R, 300X, 100R, 100X, 050R, 050X, 020R, 020X, 010R, and 010X, where the number refers to the acquisition frequency in kHz and the letter indicates if it refers to the real (R) or the imaginary part (X) of the impedance.

## 4.2. Data preprocessing stage

Acquired impedance components were normalized to adapt them to the classifiers implemented [21]. The normalization adjusted the mean value and the standard deviation of the data to zero and one respectively, using

$$Z_i' = R_i' + j \cdot X_i' \text{ where } \begin{cases} R_i' = \frac{R_i - \mu_{R_i}}{\sigma_{R_i}} \\ X_i' = \frac{X_i - \mu_{X_i}}{\sigma_{X_i}} \end{cases} \quad (1)$$

where $\mu_{R_i}$ and $\sigma_{R_i}$ are the mean and the standard deviation respectively of the real part of the impedance $R_i$, and $\mu_{X_i}$ and $\sigma_{X_i}$ are the mean and the standard deviation respectively of the imaginary part of the impedance $X_i$.

## 4.3. Statistical analysis stage

The one-way ANOVA test was performed to analyze the statistical significance of the preprocessed impedances. The p-value of each variable was obtained and compared with a significance level of α=0.01. According to the ANOVA test, p-values lower than α denote significant differences between the analyzed groups, while p-values greater than α and denote non-significant differences [22].

## 4.4. Classification stage

Twelve classifiers were implemented: one Radial Basis Function Artificial Neural Network (RBF-ANN), 10 LDA classifiers, and one Euclidean Distance Classifier (EDC). The RBF-ANN classifier proposed in this article was compared with the 10 LDA classifiers and with the EDC. LDA and EDC were chosen because they are, respectively, one-dimensional and multidimensional classifiers typically employed in the literature [23-25]. All the classifiers implemented employed impedance data as inputs, so none of them implemented a feature extraction stage before the classification stage.

### *4.4.1. RBF-ANN classifier*

A RBF-ANN with a single hidden layer was proposed to classify the steel pieces. The input layer was composed of 10 neurons, which get the real and imaginary values of the impedances acquired at the five different frequencies considered. The number of hidden neurons was set to the number of training samples, which was 20, and the number of output neurons was set to 1. An input sample was assigned to Group 1 or to Group 2 when the output was greater or lower than a threshold value respectively. Fig. 5 shows a schematic of the proposed RBF-ANN.

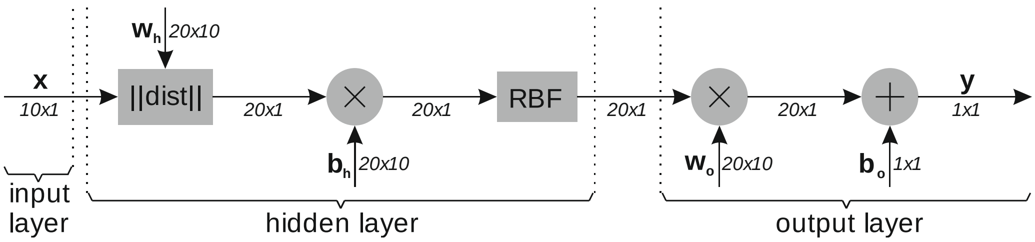



**Fig. 5.** Schematic of the RBF-ANN classifier proposed with 10 inputs, one hidden layer with 20 neurons, and one output. In this representation, *x* represents the input signal vector, *y* the output signal vector, *w* the weights vector, and *b* the biases vector. The *h* and *o* subindexes were used to refer to the hidden layer and the output layer respectively. The number next to each arrow indicates the size of the matrix data on this point of the ANN.

The training method of the RBF-ANN consisted of two steps: (1) setting the weights of the hidden layer as the values of the training samples; and (2) adjusting the output layer weights and biases by means of solving the linear expression that relates the output of the hidden layer and the target values of the training samples [26-28].

### *4.4.2. LDA classifiers*

Ten LDA classifiers were proposed to classify the steel samples using independently each one of the ten acquired and preprocessed impedance components. The training method of the LDA classifiers consisted of adjusting the threshold, selecting the one that obtained a better classification rate with the training samples [29]. Each classifier was named LDA-FFFC, where FFFC referred to the impedance component associated with the LDA classifier, using the nomenclature defined in Section 4.1.

### *4.4.3. Euclidean distance classifier*

EDC is a multidimensional classifier that assigns a sample to the group whose multifrequency descriptive point, which is called centroid, is closer. The training method of the EDC consisted of calculating the centroids of the two groups of samples, called $\overline{c^{I}}$ and $\overline{c^{II}}$, as the average of the training samples of each group.

The trained EDC computed the Euclidean distance from the sample point to each centroid, evaluated the difference between the distances from the point to the two centroids $\left\|\bar{x}-\overline{c^{I}}\right\|-\left\|\bar{x}-\overline{c^{II}}\right\|$, and assigned the sample to Group 1 or to Group 2 when the difference was higher than the threshold value.

### 4.5. ROC analysis stage

The ROC analysis consisted in analyzing the performance of the classifiers described in the previous subsection as a function of their thresholds [12, 30]. Three performance parameters were calculated in the ROC analysis: sensitivity, specificity, and precision. Sensitivity is the success rate considering only samples of Group 2, which is known as true positive rate. Specificity is the success rate considering only samples of Group 1, which is known as false positive rate. Precision is the success rate considering samples of both groups. The main graphical result of the ROC analysis is the ROC curve, which is a function that represents the sensitivity versus the specificity of the analyzed classifier for different values of the threshold. The main numerical result of the ROC analysis is the area under the ROC curve ($A_{ROC}$). The optimum precision, which is the maximum precision that can be obtained with the analyzed classifier, is another numerical result commonly employed to describe the classifier.

## 5. Results

The results of the ANOVA test and the performance comparison of the implemented classifiers are presented in the next two subsections.

### 5.1. ANOVA test results

Results obtained after applying the ANOVA test to each one of the 10 impedance components considered in this work are shown in Table 1, presenting the p-value for each impedance component.

| Impedance component | p-value |
|---|---|
| 300R | $3.177 \cdot 10^{-9}$ |
| 300X | $1.645 \cdot 10^{-7}$ |
| 100R | $4.087 \cdot 10^{-1}$ |
| 100X | $1.095 \cdot 10^{-8}$ |
| 050R | $2.030 \cdot 10^{-5}$ |
| 050X | $3.427 \cdot 10^{-5}$ |
| 020R | $1.399 \cdot 10^{-8}$ |
| 020X | $5.064 \cdot 10^{-2}$ |
| 010R | $3.699 \cdot 10^{-3}$ |
| 010X | $9.146 \cdot 10^{-1}$ |

**Table 1.** ANOVA p-value of each impedance component.

Considering a significance level of $\alpha$=0.01, Table 1 shows that only the impedance components 100R, 020X, and 010X had a p-value greater than $\alpha$, while the other seven components had a p-value lower than $\alpha$. According to the ANOVA test, the above-mentioned three impedance components did not have significant differences between the samples of the Groups 1 and 2 while the other seven impedance components had significant differences.

## 5.2. Classifiers performance results

The optimum precision, the ROC, and the $A_{ROC}$ of the implemented classifiers were calculated in order to compare the classifiers performance. They are shown in Table 2 and in Fig. 6.

| Classifier | Optimum precision | $A_{ROC}$ |
|---|---|---|
| RBF-ANN | 95% | 0.98 |
| LDA-300R | 90% | 0.93 |
| LDA-300X | 80% | 0.86 |
| LDA-100R | 70% | 0.64 |
| LDA-100X | 90% | 0.93 |
| LDA-050R | 75% | 0.82 |
| LDA-050X | 80% | 0.79 |
| LDA-020R | 85% | 0.91 |
| LDA-020X | 70% | 0.67 |
| LDA-010R | 75% | 0.74 |
| LDA-010X | 55% | 0.49 |
| EDC | 90% | 0.96 |

**Table 2.** Optimum precision and area under the ROC curve of the LDA, EDC, and RBF-ANN classifiers obtained for the cams classification performed.

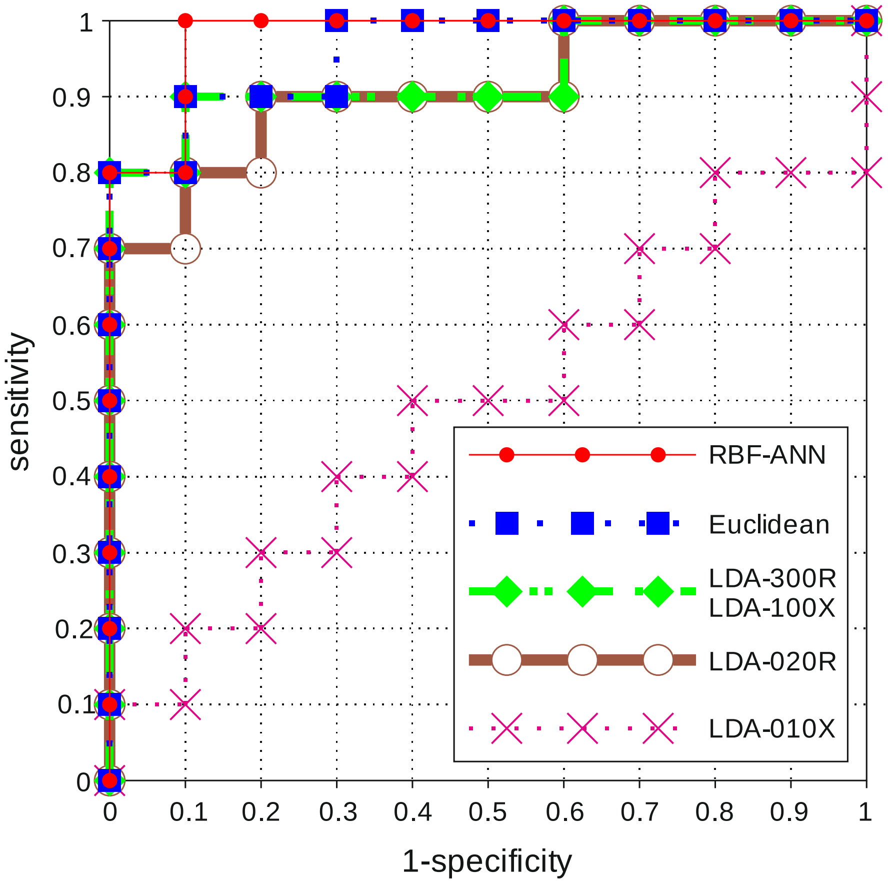



**Fig. 6.** ROC curves for the LDA-300R (green), LDA-100X (green), LDA-020R (brown), LDA-010X (pink), EDC (blue), and RBF-ANN (red) classifiers obtained for the cams classification performed, where the X-axis value (1-specificity) is the complementary percentage of the specificity expressed as a real number between 0 and 1.

According to Table 2, the RBF-ANN proposed in this article showed the best classification performance, with an optimum precision of 95% and an $A_{ROC}$ of 0.98. The second best classification performance was achieved by the EDC, with an optimum precision of 90% and an $A_{ROC}$ of 0.93.

## 6. Discussion

The results presented in the previous section show two main findings. The first one is that the proposed RBF-ANN can implement a method to classify heat-treated steel pieces with high precision and high $A_{ROC}$. This finding is supported by the above presented RBF-ANN classifier results, which show a precision of 95% in the steel samples classification. The second one is that the RBF-ANN had a better performance than the other classifiers implemented in this article. This is supported by the results presented in Table 2, which show that the RBF-ANN obtained a precision and an $A_{ROC}$ 5% and 2% higher than the best precision and $A_{ROC}$ obtained by the LDA classifiers and the EDC.

The frequency range of the impedances employed as input values for the classifiers was around 10 kHz and 300 kHz. The authors of this article consider the range of frequencies chosen to be suitable for two reasons. On the one hand, the ANOVA test performed on the data used in this article showed 70% of components had significant differences between the two groups considered. This result suggests that the frequency range chosen is suitable to differentiate groups of samples with different heat treatments received. On the other hand, the frequencies employed in this article are in line with the work of other authors who chose a similar frequency range in their research works. For example, Konoplyuk *et al.* [9] employed impedances at several kHz to estimate the perlite composition in ductile cast irons, and Mercier *et al.* [10] and Habiby *et al.* [31] evaluated steel properties employing data in the range of a few kHz.

There are three noteworthy observations when analyzing the performance of the classifiers. The first observation is that the best classifiers considering the precision are also the best classifiers considering the $A_{ROC}$, which indicates that the classification ranking considering the precision is going to be the same ranking when considering the $A_{ROC}$. The second observation is the performance improvement of the RBF-ANN and the EDC compared to the LDA classifiers. This improvement may be because the information necessary to perform the classification was not contained in a singular impedance component. Using multiple variables to improve the performance of univariate classifiers is common for both NDT field classifications [12] and general classification problems [27, 28], and this is one of the possible reasons why classifiers with multiple inputs are mostly employed in the literature [23]. The third observation is the relation between the p-values of the ANOVA test and the classification results obtained by the LDA classifiers: the LDA classifier implemented for the impedance components with the highest p-values obtained the worst classification results and vice versa.

ROC analysis shows the performance of the proposed classifiers when having the best precision is not the main requirement of the classifier. For example, quality control applications require an optimum sensitivity in order to not consider one bad sample as a good one. When analyzing Fig. 6, the specificity and precision of the classifiers with a 100% sensitivity is respectively 90% and 95% for the RBF-ANN classifier, 40% and 80% for the best LDA classifier, and 70% and 85% for the EDC. The higher specificity and precision of the proposed RBF-ANN classifier compared to the EDC and LDA classifiers show the strength of the proposed classifier when different design conditions are considered.

One limitation of the presented work is related to the conditions in which the experiments have been performed. Data acquisition was done in an experimental laboratory with a controlled ambient temperature and away from noise sources, while industrial data acquisition is usually performed with variable ambient temperature and external source noises. Nevertheless, the external noise influences could be minimized with a filtering stage, and the temperature changes influence in the measurements is expected to be reduced because of the self-compensating coil. Another limitation is that the implemented classifiers have not yet been implemented in a quality control line to test its performance in real conditions. Nevertheless Garcia-Martin *et al.* [32] found that an ANN without a feature extraction stage can classify steel pieces with an execution time small enough to not consider the processing time a limitation factor, which suggests that this limitation is not an important issue in the classification system proposed in this article.

## 7. Conclusions

This study suggests two conclusions about employing a RBF-ANN to classify heat-treated steel samples using multifrequency eddy current NDT impedances. The first one is that a classifier based on a RBF-ANN can perform the classification with good precision and $A_{ROC}$, 95% and 0.98 respectively in our test. The second one is that RBF-ANN based classifiers achieve higher performance than other classifiers such as LDA and EDC, at least a precision of 5% higher and an $A_{ROC}$ of 2% higher in our test.

## 8. Acknowledgements

The authors thank the company Ingeniería y Sistemas de Ensayos no Destructivos ISEND (www.isend.es/en) for the valuable help in the experimental tests. The first author's work was also possible thanks to a "Formación de Personal Investigador" program grant. This program was financed by the "Universidad de Valladolid" (Spain) and co-financed by "Banco Santander". The authors would also like to thank the Writing Center of the University of Kentucky (USA) their help reviewing this article.

## 9. References

[1] Khan S, Ali F, Khan AN, Iqbal M (2008) Pearlite determination in plain carbon steel by eddy current method. J Mat Process Technol 200:316-318.

[2] Zergoug M, Lebaili S, Boudjellal H, Benchaala A (2004). Relation between mechanical microhardness and impedance variations in eddy current testing. NDT & E Int 37:65-72.

[3] Le Diraison Y, Joubert P-Y, Placko D (2009) Characterization of subsurface defects in aeronautical riveted lap-joints using multi-frequency eddy current imaging. NDT & E Int 42:133-140.

[4] He Y, Luo F, Pan M, Weng F, Hu X, Gao J, et al (2010) Pulsed eddy current technique for defect detection in aircraft riveted structures. NDT & E Int 43:176-181.

[5] Yusa N, Janousek L, Rebican M, Chen Z, Miya K, Chigusa N, et al (2006) Detection of embedded fatigue cracks in Inconel weld overlay and the evaluation of the minimum thickness of the weld overlay using eddy current testing. Nucl Eng Des 236:1852-1859.

[6] Yusa N, Machida E, Janousek L, Rebican M, Chen Z, Miya K (2005) Application of eddy current inversion technique to the sizing of defects in Inconel welds with rough surfaces. Nucl Eng Des 235:1469-1480.

[7] Betta G, Ferrigno L, Laracca M (2012). GMR-Based ECT Instrument for Detection and Characterization of Crack on a Planar Specimen: A Hand-Held Solution. IEEE Trans Instrum Meas 61:505-512.

[8] Bernieri A, Ferrigno L, Laracca M, Molinara M (2008) Crack shape reconstruction in eddy current testing using machine learning systems for regression. IEEE Trans Instrum Meas 57:1958-1968.

[9] Konoplyuk S (2010) Estimation of pearlite fraction in ductile cast irons by eddy current method. NDT & E Int 43:360-364.

[10] Mercier D, Lesage J, Decoopman X, Chicot D (2006) Eddy currents and hardness testing for evaluation of steel decarburizing. NDT & E Int 39:652-660.

[11] Wrzuszczak M, Wrzuszczak J (2005) Eddy current flaw detection with neural network applications. Measurement 38:132-136.

[12] Gros X (1996) NDT data fusion. Elsevier.

[13] Wang P, Zhu L, Zhu Q, Ji X, Wang H, Tian G, et al (2013) An application of back propagation neural network for the steel stress detection based on Barkhausen noise theory. NDT & E Int 55:9-14.

[14] Junyan L, Qingju T, Xun L, Yang W (2012) Research on the quantitative analysis of subsurface defects for non-destructive testing by lock-in thermography. NDT & E Int 45:104-110.

[15] Cao Q, Liu D, He Y, Zhou J, Codrington J (2012) Nondestructive and quantitative evaluation of wire rope based on radial basis function neural network using eddy current inspection. NDT & E Int 46:7-13.

[16] Mix PE (2005) Introduction to nondestructive testing: a training guide. John Wiley & Sons.

[17] Handbook A (1989) Nondestructive evaluation and quality control. ASM International, Metals Park, OH. 241-246.

[18] Qu Z, Zhao Q, Meng Y (2014) Improvement of sensitivity of eddy current sensors for nano-scale thickness measurement of Cu films. NDT & E Int 61:53-57.

[19] Dodd C, Deeds W (1968) Analytical Solutions to Eddy-Current Probe-Coil Problems. J Appl Phys 39:2829-2838.

[20] Garcia-Martin J, Gomez-Gil J, Vazquez-Sanchez E (2011) Non-destructive techniques based on eddy current testing. Sensors 11:2525-2565.

[21] Deza MM, Deza E (2009) Encyclopedia of distances. Springer, Berlin Heidelberg.

[22] Scheffe H (1999) The analysis of variance. John Wiley & Sons.

[23] Duda RO, Hart PE, Stork DG (2012) Pattern classification: John Wiley & Sons.

[24] Marco VR, Young DM, Turner DW (1987) The Euclidean distance classifier: an alternative to the linear discriminant function. Commun Stat-Simul Comput 16:485-505.

[25] Zhao W, Chellappa R, Nandhakumar N (1998) Empirical performance analysis of linear discriminant classifiers. In Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on: IEEE; pp. 164-169.

[26] Broomhead DS, Lowe D (1988) Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal Signals and Radar Establishment Malvern, United Kingdom.

[27] Haykin S (1998) Neural Networks: A Comprehensive Foundation. Prentice Hall.

[28] Bishop CM (2007) Neural Networks for Pattern Recognition. 608 Oxford, UK.

[29] Webb AR (2003) Statistical pattern recognition: John Wiley & Sons.

[30] Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27:861-874.

[31] Habiby F, Siddiqui T, Khan S, Khan A (1992) Austenite determination by eddy current measurements in a maraging steel. NDT & E Int 25:145-146.

[32] Garcia-Martin J, Martínez-Martínez V, Gomez-Gil J (2014) Clasificación del tratamiento térmico de aceros con ensayos no destructivos por corrientes inducidas mediante redes neuronales. Dyna 89:526-532.

# Article 5: Early detection of plant diseases

The fifth article presents the **reflectance based disease model** described in the "Introduction and Summary" part of the document.

The article presented below is a version that is going to be submitted to a JCR-indexed peer-reviewed journal. For this reason, there is not bibliographic data about this article in this introduction.

# Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops

**Abstract:** This study is aimed (i) to estimate the *angular leaf spot* disease severity in common beans crops from leaf and canopy spectral reflectance data, (ii) to evaluate the informative spectral bands in the detection, and (ii) to compare the estimation accuracy when the reflectance or the first derivative reflectance (FDR) is employed. Three data sets of useful spectral reflectance measurements in the 440 to 850 nm range were employed, being them taken over the leaves and canopy of bean crops with different levels of disease. A system based in Principal Component Analysis (PCA) and in Artificial Neural Networks (ANN) was developed to estimate the disease severity from leaf and canopy hyperspectral reflectance spectra. Levels of disease to be taken as true reference were determined from percentages of area with necrosis of bean leaf RGB images. In the estimation of *angular leaf spot* disease severity in bean crops by hyperspectral reflectance spectrometry, this study suggest that (i) successfully estimations with coefficients of determination of around 0.8 can be achieved if the spectra is acquired by the spectroradiometer in contact with the leaves, (ii) unsuccessfully estimations are be obtained when the spectra are acquired by the spectroradiometer from one or more meters above the crop, (iii) the red to near-infrared bands (630-850 nm) offers the same precision in the estimation than the blue to near-infrared bands (440-850), and (iv) neither significant improvements, nor significant detriments, are achieved when the input data to the estimation processing system are first derivative reflectance spectra, instead of the reflectance spectra.

**Keywords:** Phaseolus vulgaris; remote sensing; artificial neural networks; hyperspectral; multispectral; reflectance; disease detection; spectroradiometry.

## 1. Introduction

*Phaseolus vulgaris*, the common bean, is a legume consumed worldwide. In the 2012 year, 24 million tons of dry common beans and 21 million tons of green beans were produced, being respectively in 11$^{th}$ and 78$^{th}$ position of production ranking list of crops worldwide [1, 2]. Beans are an excellent source of protein, carbohydrates, dietary fiber, vitamins, minerals, and phytochemicals [3]. Consumption of beans has healthy benefits because it can reduce the cholesterol levels [4], the diabetes [5], and the risk of prostate cancer [6] and mammary cancer [7]. Some studies suggest that diets with legumes such as beans are sometimes associate with longevity [8].

Several bacterial, fungal, and viral diseases attack aerial and underground parts of common beans [9-11]. Among fungal diseases, the Angular Leaf Spot (ALS), caused by the fungus *Phaeoisariopis griseola*, is responsible for important losses worldwide [12-14]. An early detection of ALS disease in a crop allows effective disease treatments.

Spectral reflection of light by the leaf and the canopy of crops in visible (VIS=400-700nm), near-infrared (NIR=700-1200 nm), and shortwave infrared (SWIR=1200-2400) spectral regions, can be measured to detect diseases. Healthy plants appear green since the green light band (515-600 nm) is reflected relatively efficiently compared with blue, and red bands, which are absorbed by photoactive pigments. Diseased plants usually exhibit discrete lesions on leaves, corresponding to necrotic or chlorotic regions, which increase the reflectance in the visible range [15]. A sharp transition from low to high reflectance usually occurs in the reflectance spectrum in the transition from VIS to NIR (around 700 nm), and this transition usually shifts to shorter wavelengths in diseased crops [16].

Multispectral reflectance spectra, characterized by having bands with more of 40 nm of width, measure energy usually over between 2 and 10 separated bands. Spectral Vegetation Indices (SVIs) can be computed from combinations of some spectral values of a multispectral reflectance spectrum. Disease estimation in crops employing a multispectral reflectance spectrum, in some cases computing indices, is common in the literature. For example Dammer *et al.* [17] only needed two spectral bands, the red and the infrared, to compute the NDVI index to estimate head blight on wheat crops, Cui *et al.* [18] detected rust disease in soybean crops employing different vegetation indices, Xiao and Mcpherson [19] classified the health of trees by means of the NDVI index, and Pietrzykowski *et al.* [20] detected *Mycosphaerella* in *Eucalyptus globulus* foliage with an index computed from two spectral bands.

Hyperspectral reflectance spectra, characterized by having bands with less than 10 nm in width, is typically composed by 200 or more bands, and enable the construction of an almost continuous reflectance spectra. This allows the in-depth examination of crop features, not being this possible with the relative coarse bandwidths acquired by multispectral spectrometers. Disease estimation in crops from its hyperspectral reflectance spectrum is common in the literature. For example Zhang *et al.* [21] detected *late blight* disease in tomatoes crops, Muhammed *et al.* [22] quantified disease severity in wheat crops, Wu *et al.* [23] detected *botytris cinerea* in eggplant crops, Liu *et al.* [24] classified the severity of *blight* disease in rice crops in levels, and Prabhakar *et al.* [25] estimated *yellow mosaic* disease in black gram.

Hyperspectral reflectance spectra contain high correlated information in a large number of wavelengths. Several methods have been employed to reduce this correlation, as the computation of SVIs [19, 26, 27], Principal Component Analysis (PCA) [24, 28, 29], Principal Component Regression (PCR) [30], and Partial Least Squares (PLS) regression [23, 27, 31]. Once that correlation is reduced, decision trees [32, 33], Artificial Neural Networks (ANN) [24, 34, 35], Support Vector Machines (SVM) [36-38], and clustering [32, 39, 40], among others, are technics usually employed to estimate crop features in remote sensing.

This study is aimed (i) to estimate the ALS disease severity in common beans crops from leaf and canopy spectral reflectance data, (ii) to evaluate the informative spectral bands in the detection of the disease, and (ii) to compare the estimation accuracy when the reflectance or the FDR is or not employed.

## 2. Material and Methods

### 2.1. Experimental area

Three experiments, named as UFV1, UFV2, and FEVP3 were conducted. UFV1 and UFV2 were performed in the *Diogo Alves de Melo* experimental plot of the *Federal University of Viçosa*, in Viçosa, MG, Brazil (20º 45’ S, 42º 52’ W, altitude 648 m). FEVP3 was sited in *Vale do Piranga* experimental plot in Oratorios, MG, Brazil (20º 24’ S, 42º 48’ W, altitude 450m).

The soil was prepared with a disc harrow. The crops were seeded on April 8, 2011, July 29, 2011, and July 26, 2011, respectively for UFV1, UFV2, and FEVP3. The variety employed was Ouro Vermelho beans, provided by the *Grupo Comercial Vermelho* Brazilian company. The seed density was 10 seeds by meter along a 5.5 meters row, having five rows in each experiment, with a row spacing of 0.5 m. Common insecticides as well as 8-28-16 fertilizer were applied along the crop growing. Manual weed control was done.

Straw contaminated with the disease was spread in the plots eight days after the emergence of the crop. The plot was divided in four parts, applying different doses of fungicide to each part in order to achieve different disease severities. The AMISTAR® fungicide, whose active ingredient is Azoxistrobina, was applied in doses of 0, 50, 100 and 150 g/Ha on each part. The fungicide application was applied three times along the crop development using a back sprayer.

## 2.2. Data acquisition

Three types of data were acquired: leaf reflectance spectra, canopy reflectance spectra, and leaf RGB images.

Leaf reflectance spectral data in the 440-850 nm band were acquired employing an *ASD Field Spec HandHeld-2* spectroradiometer, manufactured by the *Analytic Spectral Devices* company, Boulder, USA. A leaf clip accessory, also from the same company, was connected to the spectroradiometer (Figure 1a). Calibrations were done using a *spectralon* white panel, following the guidelines of the spectroradiometer user manual. All the acquisitions were performed between 10:00 and 14:00, according to Brasilia time. The acquisition software employed was *ViewSpecpro*, which was provided by the manufacturer of the spectroradiometer. Each leaf reflectance spectrum of the crop was considered as the mean of the reflectance spectra obtained with 27 leaves taken from different parts of the crop.

Canopy reflectance spectral data were acquired by the same spectroradiometer without the leaf clip accessory and employing a four wheels metallic platform and a four meters long optic fiber. In this way, it was possible to focus the spectroradiometer to the canopy, from three meters above the ground (Figure 1b).

Leaf RGB images were acquired employing an ordinary digital camera, an 0.8×0.8×0.8 m wooden box with an opening at the top, and six fluorescent lamps illuminating the inside of the box. A set of leaves of the crop were placed at the bottom of the box, and photos with the camera were taken by the opening at the top (Figure 1c). The effective resolution was 3.78 pixels/mm.

(a) (b) (c)

**Figure 1:** Representative images of the data acquisition: (a) spectroradiometer and leaf clip accessory acquiring a leaf reflectance spectrum, (b) spectroradiometer and platform acquiring a canopy reflectance, and (c) several leaves RGB image taken.

### 2.3. Processing overview

Figure 2 shows an overview of the processing carried out in this study. It presents a *RGB image processing* method and a *reflectance spectral processing* method.

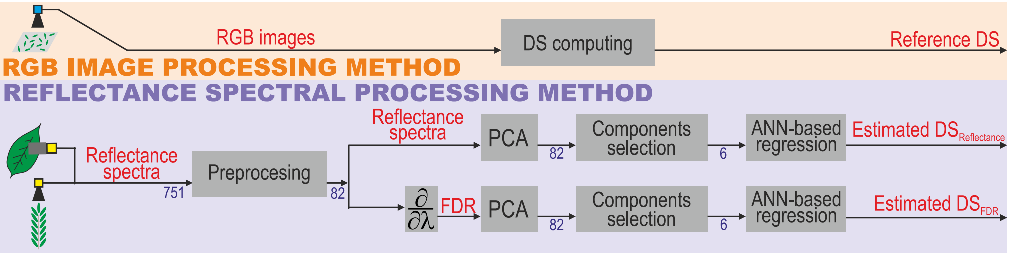



**Figure 2:** Overview of the processing method performed in this article.

The *RGB image processing* method estimates the severity of the crop by computing the percentage area with necrosis in RGB images of crop leaves. Values obtained were used as reference, to check the accuracy of the *reflectance spectral processing* method.

The *reflectance spectral processing* method estimates the disease severity of the crop, being its source a leaf or canopy spectrum. This method includes a computation of the FDR spectrum, a conversion to another space data by means of a PCA statistical procedure, a selection of the most relevant components in the new data space, and a regression process implemented by an ANN (Artificial Neural Networks) to estimate the disease severity.

The *RGB image processing* and the *reflectance spectral processing* methods include supervised learning techniques. Both methods implement training and testing stages. Blocks of both methods were implemented in Matlab®.

### 2.4. The RGB image processing method

Disease severity (DS) values of the crop obtained by the *DS computing* block were taken as reference in this study. The DS was computed from the RGB images of leaves as the fraction of the area that is infected [26]. The *DS computing* block obtained DS values of each experiment following the next three steps:

- Step 1: A representative leaf RGB image that contained parts with disease and parts without disease was selected. They were selected form the image (i) 50 pixels from healthy green areas, (ii) 50 pixels from light brow areas with chlorosis, (iii) 50 from dark brown areas with necrosis, and (iv) 50 pixels from the white background.
- Step 2: Pixels of all the RGB images were associated to a healthy class, to a chlorosis class, to a necrosis class, or to a background class, employing a *Quadratic Discriminant Analysis* classifier. This classifier was based on the observations corresponding to each one of 50 pixels previously selected. It assigned an observation x to the class k for which

$$\delta_k(x) = -\frac{1}{2}log|\Sigma_k| - \frac{1}{2}\left(x-\mu_k\right)'\Sigma_k^{-1}\left(x-\mu_k\right) + log(\pi_k) \qquad (1)$$

is largest, being $\mu_k$ the average of the training observations for the $k^{th}$ class, $\pi_k$ the proportion of the training observations that belong to the $k^{th}$ class, and $\Sigma_k$ the covariance matrix for the $k^{th}$ class.

- Step 3: For each RGB image collected in the experiment (Figure 1c), the DS was computed as the number of pixels of the area with necrosis, divided by the sum of pixels of areas healthy, with necrosis, and with chlorosis.

## 2.5. The reflectance spectral processing method

The spectra processing carried out in this study was performed by the *Preprocessing*, *Derivative*, *PCA*, *Components selection* and *ANN-based regression* blocks, according to the disposition shown in Figure 2.

### *2.5.1 Preprocessing block*

The preprocessing block performed three actions in each spectra captured by the spectroradiometer: it applied a 5-points moving average filter, it downsampled the spectrum in a 5 to 1 rate, and it removed data outside the 440-850 nm band.

### *2.5.2 Derivative block*

The *derivative* block computed the derivative spectrum from the reflectance spectrum. The derivative value at each wavelength was computed as

$$\left.\frac{ds}{d\lambda}\right|_i = \frac{s(\lambda_{i+1}) - s(\lambda_{i-1})}{2 \cdot \Delta\lambda} \quad (2)$$

where *s* represents the reflectance spectrum, $\lambda$ the wavelength, and *i* the index of both the spectrum and the derivative vectors.

### *2.5.3 PCA block*

The *PCA* block applied a PCA to both the reflectance and the FDR data. PCA is a technique which transforms the input reflectance data to another multidimensional orthogonal space. PCA defines the first principal component as the one with the highest possible variance, and the rest of the components as the component with the highest variance possible that is orthogonal to all the previous components. The input signals of the *PCA* block were normalized to calculate the PCA of a set of variables with zero mean and unit variance.

### *2.5.4 Components selection block*

The *Components selection* block selected the first components of the PCA analysis performed in the PCA block and removed the rest of the components because they contained only a few percentage of the variance of the original data. The number of components selected in this work, which is going to be referred as $N_{comp}$, varied in order to perform a comparison and choose the optimal option.

### *2.5.5 ANN-based regression block*

The *ANN-based regression* block implements a backpropagation multilayer perceptron (MLP) ANN with a single hidden layer. This ANN has $N_{comp}+1$ inputs, which are the $N_{comp}$ components selected in the previous block and the number of days elapsed between the day when the bean plants were planted and the day when the data was acquired. The hidden layer of the ANN has a variable number of neurons, which is going to be referred as $N_{hidden}$, with a *logsig* activation function. In addition, the MLP ANN has one output, which is the estimated DS, with a linear activation function. A different ANN was designed for each one of the six experiments performed, using the samples of the 30% of the plots for testing the ANN and the rest of the samples for training and validation. The 75% of these samples were employed to train the ANN and the other 25% to validate it.

## 3. Results

### 3.1. Data analysis results

The crop in the experiments grow successfully, being the average productivities 1621, 1725 and 1261 kg/ha, respectively for UFV1, UFV2 and FEVP3. The disease appeared at the flowering stage, being found the highest levels at the pod maturation stage.

Figure 3 shows the mean DS values in the experiments. The DS was estimated from the area of necrosis found in the images of leaves, following the procedure detailed in the Material and Methods section. As expected, the DS generally increased with time, and the slight decreases could be due to the controlled application of fungicides.

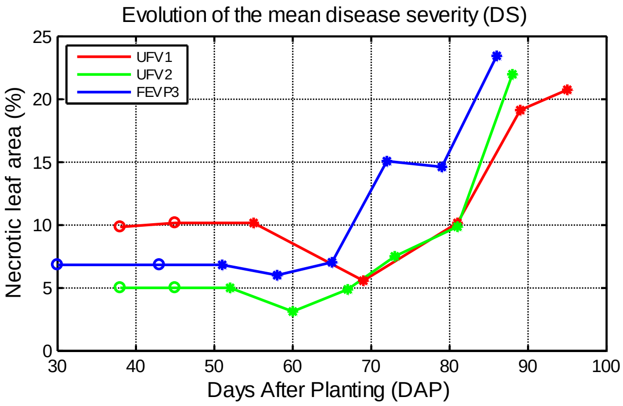



**Figure 3:** Evolution of the mean DS for each one of the three experiments conducted. The DS was estimated as the percentage of necrosed area of the leaves.

Figure 4 shows, for the leaf and canopy data acquired, the mean reflectance spectra and the mean FDR spectra in the 450-850 nm waveband. The mean reflectance spectra signature were as expected for bean crops [41]. The canopy reflectance spectral signal in the UFV2 experiment was a slightly lower above 700 nm wavelength, due to a fewer development of the crop, which was caused by a weed infestation.

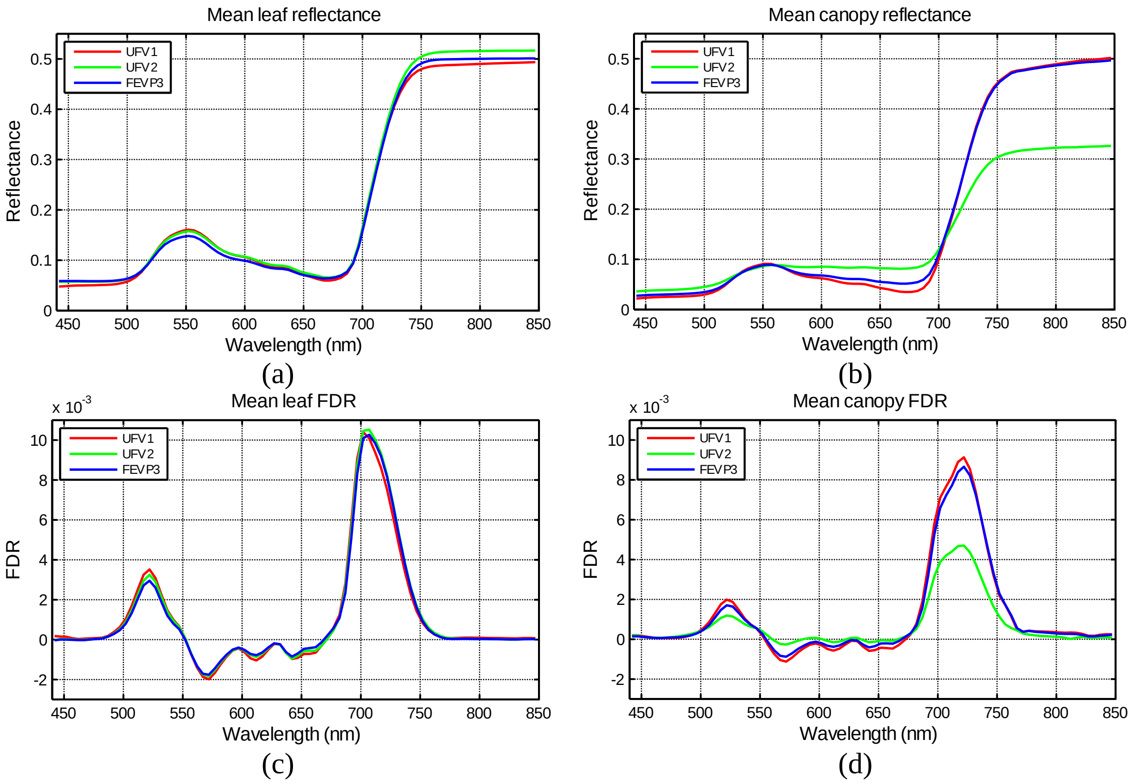



**Figure 4:** Mean reflectance spectra of the leaf (a) and the canopy (b), and mean FDR spectra of the leaf (c) and the canopy (d).

Correlation (r) graphs between the disease severity measurements and the mean reflectance spectra, the FDR spectra spectra are shown in Figure 5.

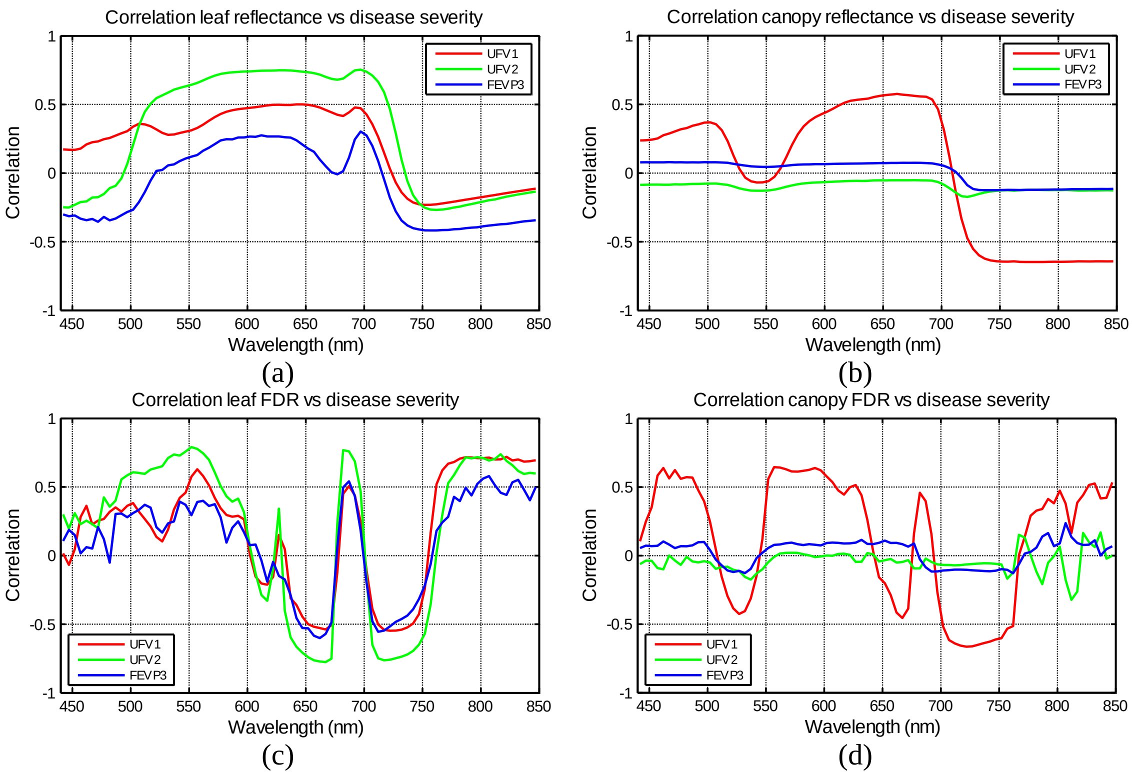



**Figure 5:** Correlation of spectral data between disease severity and reflectance in 450-900 nm.

### 3.2. Methodology evaluation results

This section presents the results associated to the experiments developed to evaluate the methodology proposed in this article. The first experiment consisted on analyzing the $N_{comp}$ paramether, that is, the number of components of the PCA that should be considered in the proposed methodology. To do so, the percentage of information contained in the first components of the PCA was calculated, obtaining the results presented on Figure 6.

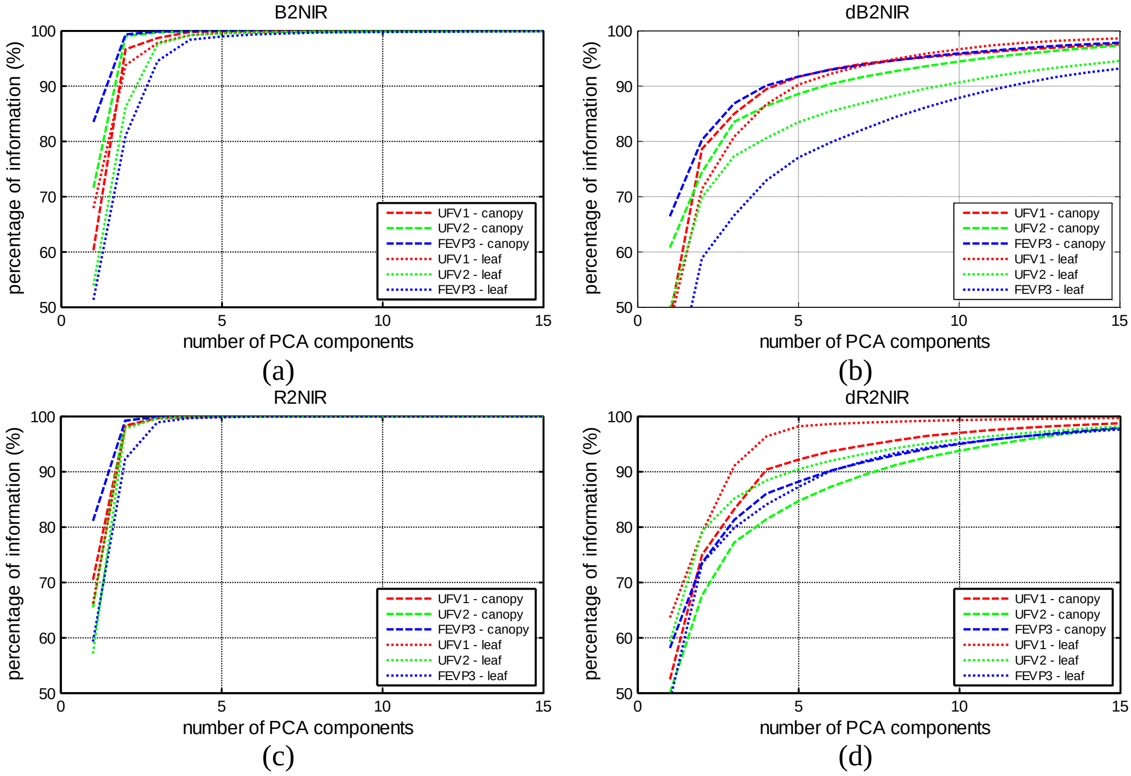



**Figure 6:** Percentage of information contained in the first components of the PCA.

The second experiment evaluated the performance of the proposed methodology in the different conditions considered in this article: using canopy reflectance spectral data or leaf reflectance spectral data, using data acquired from the using the reflectance or the FDR, and using the information from the blue band to the NIR band or from the red band to the NIR band. Moreover, each experiment considered the next values for the $N_{comp}$ paramether: 3, 5, 6, 8, 10, 12, and 15. Performance of the methodology was evaluated considering the root mean square error (RMSE) and the correlation coefficient. Results obtained in this experiment are presented in Table 1 and Table 2.

| | | ***N_comp*** | **3** | **5** | **6** | **8** | **10** | **12** | **15** |
|---|---|---|---|---|---|---|---|---|---|
| UFV1 | canopy | B-NIR | 7,00976 | 6,85969 | 6,68877 | 6,74070 | 6,87590 | 6,92869 | 7,05479 |
| | | dB-NIR | 6,86505 | 6,93099 | 6,56481 | 6,92614 | 6,84546 | 6,83529 | 6,86752 |
| | | R-NIR | 6,80645 | 7,02559 | 6,72468 | 6,80139 | 6,72523 | 6,96982 | 7,37304 |
| | | dR-NIR | 7,11346 | 7,06309 | 6,91847 | 7,04893 | 6,87113 | 6,92608 | 7,00584 |
| | leaf | B-NIR | 4,49312 | 4,93517 | 4,97555 | 4,74867 | 4,29175 | 4,60809 | 4,22537 |
| | | dB-NIR | 4,45709 | 4,21346 | 4,30137 | 4,04546 | 4,20474 | 4,17128 | 4,46979 |
| | | R-NIR | 4,73917 | 4,64860 | 4,12954 | 4,83252 | 4,22273 | 4,14769 | 4,81145 |
| | | dR-NIR | 4,03835 | 4,20344 | 4,28629 | 4,34264 | 4,13154 | 4,20475 | 4,57075 |
| UFV2 | canopy | B-NIR | 2,73193 | 2,46498 | 3,08349 | 2,39929 | 2,82308 | 2,89123 | 2,71306 |
| | | dB-NIR | 2,62997 | 2,64949 | 2,68307 | 2,73013 | 2,75957 | 2,47165 | 2,95301 |
| | | R-NIR | 2,43366 | 2,70749 | 2,68556 | 2,58002 | 2,69200 | 2,58916 | 2,84158 |
| | | dR-NIR | 2,76262 | 2,59683 | 2,63116 | 2,67225 | 2,96534 | 2,64264 | 2,79352 |
| | leaf | B-NIR | 3,88117 | 3,69829 | 3,15733 | 3,51994 | 3,63774 | 3,64676 | 3,48067 |
| | | dB-NIR | 3,07801 | 3,54331 | 3,56512 | 3,44123 | 3,53413 | 3,48449 | 3,80720 |
| | | R-NIR | 4,35607 | 3,59561 | 3,71764 | 3,78989 | 4,00755 | 4,50216 | 4,09669 |
| | | dR-NIR | 3,04899 | 3,36122 | 3,63634 | 3,63552 | 3,50566 | 3,81573 | 3,75746 |
| FEVP3 | canopy | B-NIR | 2,42630 | 2,41724 | 2,42604 | 2,31505 | 2,44415 | 2,20739 | 2,42745 |
| | | dB-NIR | 2,43444 | 2,45005 | 2,37226 | 2,42038 | 2,16277 | 2,42339 | 2,36823 |
| | | R-NIR | 2,43149 | 2,38523 | 2,40674 | 2,46940 | 2,43112 | 2,24629 | 2,45336 |
| | | dR-NIR | 2,47718 | 2,47231 | 2,41217 | 2,26789 | 2,28110 | 2,51494 | 2,30882 |
| | leaf | B-NIR | 3,71974 | 3,74536 | 3,78653 | 3,67351 | 3,39174 | 4,09360 | 3,34909 |
| | | dB-NIR | 4,19016 | 3,61421 | 3,55841 | 3,65243 | 3,88566 | 4,00238 | 3,34158 |
| | | R-NIR | 4,09353 | 3,81880 | 3,77409 | 4,08809 | 3,84880 | 3,97728 | 3,72713 |
| | | dR-NIR | 3,61475 | 3,67846 | 3,85353 | 3,62864 | 3,91553 | 3,95688 | 3,21183 |

**Table 1:** RMSE for the experiments proposed to evaluate the methodology.

| | | $N_{comp}$ | 3 | 5 | 6 | 8 | 10 | 12 | 15 |
|---|---|---|---|---|---|---|---|---|---|
| UFV1 | canopy | B-NIR | 75,14% | 74,078 % | 74,345 % | 73,586 % | 73,618 % | 73,709 % | 74,089 % |
| | | dB-NIR | 73,39% | 73,271 % | 76,641 % | 75,989 % | 75,304 % | 74,615 % | 74,209 % |
| | | R-NIR | 74,90% | 73,86% | 75,02% | 73,75% | 73,97% | 74,46% | 70,11% |
| | | dR-NIR | 72,49% | 71,83% | 74,24% | 71,77% | 75,22% | 74,24% | 72,74% |
| | leaf | B-NIR | 78,59% | 76,34% | 73,21% | 76,18% | 80,84% | 78,76% | 81,73% |
| | | dB-NIR | 79,70% | 81,47% | 80,80% | 83,27% | 82,23% | 82,01% | 82,25% |
| | | R-NIR | 75,98% | 78,26% | 82,84% | 77,28% | 81,61% | 82,24% | 76,31% |
| | | dR-NIR | 83,23% | 82,33% | 81,30% | 80,58% | 82,72% | 81,72% | 78,52% |
| UFV2 | canopy | B-NIR | 38,15% | 48,64% | 33,38% | 52,49% | 41,36% | 51,22% | 43,07% |
| | | dB-NIR | 42,47% | 30,76% | 39,56% | 30,28% | 57,76% | 45,41% | 53,50% |
| | | R-NIR | 46,54% | 38,03% | 35,82% | 42,44% | 41,49% | 49,04% | 49,89% |
| | | dR-NIR | 34,72% | 38,61% | 42,38% | 39,57% | 32,06% | 35,49% | 36,26% |
| | leaf | B-NIR | 89,86% | 90,28% | 92,67% | 91,12% | 89,93% | 91,01% | 91,14% |
| | | dB-NIR | 93,07% | 91,98% | 90,76% | 91,43% | 90,54% | 91,05% | 89,08% |
| | | R-NIR | 84,43% | 92,26% | 90,00% | 88,75% | 89,30% | 87,01% | 88,12% |
| | | dR-NIR | 92,97% | 92,13% | 91,69% | 90,59% | 91,16% | 90,72% | 89,86% |
| FEVP3 | canopy | B-NIR | 36,52% | 31,80% | 30,28% | 35,98% | 27,81% | 52,82% | 35,19% |
| | | dB-NIR | 29,84% | 26,81% | 37,50% | 38,53% | 51,88% | 33,59% | 31,12% |
| | | R-NIR | 28,60% | 28,69% | 34,98% | 28,90% | 27,39% | 43,79% | 37,28% |
| | | dR-NIR | 23,34% | 24,35% | 32,51% | 46,38% | 41,79% | 28,30% | 38,70% |
| | leaf | B-NIR | 82,86% | 83,56% | 82,25% | 81,79% | 84,97% | 81,02% | 86,60% |
| | | dB-NIR | 80,81% | 84,50% | 84,36% | 85,15% | 80,91% | 79,05% | 85,19% |
| | | R-NIR | 80,00% | 82,63% | 81,74% | 79,77% | 80,41% | 79,46% | 83,71% |
| | | dR-NIR | 83,74% | 82,91% | 83,94% | 84,50% | 80,24% | 81,72% | 87,49% |

**Table 2:** Correlation coefficient for the experiments proposed to evaluate the methodology.

## 4. Discussion

The experiments performed in this article allowed to analyze the influence of different paramethers of the proposed methodology estimating the ALS disease severity in common bean crops. The parameters analyzed were (i) the way the reflectance data is acquired (leaf or canopy), (ii) the variable considered in the analysis (reflectance or FDR data), (iii) the wavelength bands more interesting for the analysis (from blue to NIR or from red to NIR), and (iv) the number of inputs for the ANN-based regression block, which is related with $N_{comp}$, the number of components considered as relevant information.

The first observation about the results obtained is that the reflectance measured from the canopy does not obtain results good enough to consider it as a possible measurement to estimate the disease severity. It can be seen in Table 2, where mean the correlation values for all the experiments is 50,07% for the experiments with data measured from the canopy and 84,41% for the experiments with data measured from the leaf.

The second observation is that there are not significant differences between the results of the methods which employed the reflectance signal and the methods which employed the FDR signal. Analyzing the results presented in Table 1 and Table 2, it can be seen that reflectance is better than FDR for some experiments and vice-versa, without a consistent performance. Analyzing the mean correlation values for all experiments, experiments that used reflectance data obtained a 67,059% of mean correlation and experiments that used FDR data obtained a 67,417%, which is not a significant difference. There are studies in the literature where the use of the FDR present advantages [28, 42, 43] but there are other studies in the literature where the use of the FDR present advantages only in specific situation [44, 45]. Our study show neither advantages nor disadvantages using the reflectance data and the FDR data, which concurs with the literature review as it was said in the previous sentence.

The third observation is that the methods which employed the information of the wavelength band between blue and NIR obtained results similar to the obtained by the methods which employed the information of the wavelength band between red and NIR. This observation can be checked with the information presented in Table 2, where the mean correlation values for all the experiments is 67.91% and 66,57% for the experiments that considered the wavelength band between blue and NIR and between red and NIR respectively.

The fourth observation was obtained by the PCA, and it shown that most of the information of the different signals considered can be comprised with the first components of the PCA, as it can be seen in Figure 6. This observation suggests that most of the less significant components of the PCA could be discarded in order to minimize the complexity of the proposed method. Moreover, results presented in Table 1 and Table 2 support this observation because there is not a direct correlation between $N_{comp}$ and the RMSE or the correlation: for values of $N_{comp}$ greater than 10 the results obtained do not improve significantly.

Results obtained suggest that a method which uses the leaf reflectance for a wavelength band between the red and the NIR wavelengths (630-850 nm) can be employed to estimate the disease severity of common beans crops. Unless the reflectance and the FDR obtained similar results, the reflectance was chosen because it requires less processing, as FDR is obtained from the reflectance signal, and because the reflectance signal can include more information than the FDR in the first components of the PCA, as it can be seen in Figure 6. Moreover, using the wavelength band between red and NIR (630-850 nm) instead of the band between blue and NIR (440-850 nm) requires a simpler reflectance sensor, which can reduce its prize, and allow us to work with a smaller band (220 nm vs 410 nm), which reduces the processing load needed.

## 5. Conclusions

In the estimation of ALS disease severity in bean crops by hyperspectral reflectance spectrometry, this study suggest that (i) successfully estimations with coefficients of determination of around 0.8 can be achieved if the spectra is acquired by the spectroradiometer in contact with the leaves, (ii) unsuccessfully estimations are be obtained when the spectra are acquired by the spectroradiometer from one or more meters above the crop, (iii) the red to near-infrared bands (630-850 nm) offers the same precision in the estimation that the blue to near-infrared bands (440-850), and (iv) no significant improvements, neither significant detriments, are achieved when the input data to the estimation processing system are first derivative reflectance spectra, instead the reflectance spectra.

## Acknowledgments

The first author's work was made possible thanks to a *Formación de Personal Investigador* program grant. This program was financed by the Universidad de Valladolid (Spain) and cofinanced by Banco Santander. Moreover, the experiments and collaboration among the authors were possible thanks to a grant entitled *Becas Iberoamérica. Jóvenes Profesores e Investigadores. Santander Universidades* financed by Banco Santander. This grant was obtained by the first author in 2014 and it financed the stay at the Federal University of Viçosa (MG, Brasil) during the months while the proposed experiments were performed.

## References

[1] FAO The statistics division of the FAO. http://faostat.fao.org/ (accessed September 22, 2015).

[2] Broughton, W.J.; Hernández, G.; Blair, M.; Beebe, S.; Gepts, P.; Vanderleyden, J. Beans (Phaseolus spp.) – model food legumes. Plant Soil 2003, 252, 55-128.

[3] Tharanathan, R.N.; Mahadevamma, S. Grain legumes—a boon to human nutrition. Trends Food Sci. Technol. 2003, 14, 507-518.

[4] Bazzano, L.A.; Thompson, A.M.; Tees, M.T.; Nguyen, C.H.; Winham, D.M. Non-soy legume consumption lowers cholesterol levels: A meta-analysis of randomized controlled trials. Nutr. Metab. Carbiovasc. Dis. 2011, 21, 94-103.

[5] Messina, V. Nutritional and health benefits of dried beans. Am. J. Clin. Nutr. 2014, 100, 437S-442S.

[6] Deschasaux, M.; Pouchieu, C.; His, M.; Hercberg, S.; Latino-Martel, P.; Touvier, M. Dietary Total and Insoluble Fiber Intakes Are Inversely Associated with Prostate Cancer Risk. J. Nutr. 2014, 144, 504-510.

[7] Thompson, M.D.; Mensack, M.M.; Jiang, W.; Zhu, Z.; Lewis, M.R.; McGinley, J.N.; Brick, M.A.; Thompson, H.J. Cell signaling pathways associated with a reduction in mammary cancer burden by dietary common bean (Phaseolus vulgaris L.). Carcinogenesis 2012, 33, 226-232.

[8] Darmadi-Blackberry, I.; Wahlqvist, M.L.; Kouris-Blazos, A.; Steen, B.; Lukito, W.; Horie, Y.; Horie, K. Legumes: The most important dietary predictor of survival in older people of different ethnicities. Asia Pac. J. Clin. Nutr. 2004, 13, 217-220.

[9] Hagedorn, D.J.I., D.A. Handbook of Bean Diseases; University of Wisconsin: 1986.

[10] Achwartz, H.F.S., J.R.; Hall, R.; Foster, R.L. Compendium of bean diseases, 2nd ed.; APS Press, St. Paul, MN.: 2005.

[11] Singh, S.P.; Schwartz, H.F. Breeding Common Bean for Resistance to Diseases: A Review. Crop Sci. 2010, 50, 2199-2223.

[12] Wortmann, C.S.K., Roger Alexander; Eledu, Charles A.; Allen, David J. Atlas of common bean (Phaseolus vulgaris L.) production in Africa.; Centro Internacional de Agricultura Tropical (CIAT), Cali, CO: 1998.

[13] Schwartz, H.F.C., F.; Pineda, P.A.; Otoya, M.M.; Katherman, M.J. Dry Bean Yield Losses Caused by Ascochyta, Angular, and White Leaf Spots in Colombia. Plant Dis. 1981, 65, 494-496.

[14] de Jesus, W.C.; do Vale, F.X.R.; Coelho, R.R.; Hau, B.; Zambolim, L.; Costa, L.C.; Filho, A.B. Effects of Angular Leaf Spot and Rust on Yield Loss of Phaseolus vulgaris. Phytopathology 2001, 91, 1045-1053.

[15] Campbell, J.B. Introduction to Remote Sensing; 4th ed.; The guildford press: New York, NY, USA, 2007; pp. 447-492.

[16] Gazala, I.F.S.; Sahoo, R.N.; Pandey, R.; Mandal, B.; Gupta, V.K.; Singh, R.; Sinha, P. Spectral reflectance pattern in soybean for assessing yellow mosaic disease. Indian J. Virol. 2013, 24, 242-249.

[17] Dammer, K.-H.; Möller, B.; Rodemann, B.; Heppner, D. Detection of head blight (Fusarium ssp.) in winter wheat by color and multispectral image analyses. Crop Prot. 2011, 30, 420-428.

[18] Cui, D.; Zhang, Q.; Li, M.; Zhao, Y.; Hartman, G. Detection of soybean rust using a multispectral image sensor. Sens. & Instrumen. Food Qual. 2009, 3, 49-56.

[19] Xiao, Q.; McPherson, E.G. Tree health mapping with multispectral remote sensing data at UC Davis, California. Urban Ecosyst. 2005, 8, 349-361.

[20] Pietrzykowski, E.; Stone, C.; Pinkard, E.; Mohammed, C. Effects of Mycosphaerella leaf disease on the spectral reflectance properties of juvenile Eucalyptus globulus foliage. Forest Pathol. 2006, 36, 334-348.

[21] Zhang, M.; Qin, Z.; Liu, X.; Ustin, S.L. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. Int. J. Appl. Earth Obs. Geoinf. 2003, 4, 295-310.

[22] Hamid Muhammed, H.; Larsolle, A. Feature Vector Based Analysis of Hyperspectral Crop Reflectance Data for Discrimination and Quantification of Fungal Disease Severity in Wheat. Biosyst. Eng. 2003, 86, 125-134.

[23] Wu, D.F., L.; Zhang, C.; He, Y. Early Detection of Botrytis cinerea on Eggplant Leaves Based on Visible and Near-Infrared Spectroscopy. Trans. ASABE 2008, 51, 1133-1139.

[24] Liu, Z.-Y.; Wu, H.-F.; Huang, J.-F. Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis. Comput. Electron. Agric. 2010, 72, 99-106.

[25] Prabhakar, M.; Prasad, Y.G.; Desai, S.; Thirupathi, M.; Gopika, K.; Rao, G.R.; Venkateswarlu, B. Hyperspectral remote sensing of yellow mosaic severity and associated pigment losses in Vigna mungo using multinomial logistic regression models. Crop Prot. 2013, 45, 132-140.

[26] Ashourloo, D.; Mobasheri, M.; Huete, A. Developing Two Spectral Disease Indices for Detection of Wheat Leaf Rust (Pucciniatriticina). Remote Sens. 2014, 6, 4723-4740.

[27] Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 2003, 86, 542-553.

[28] Sankaran, S.; Mishra, A.; Maja, J.M.; Ehsani, R. Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. Comput. Electron. Agric. 2011, 77, 127-134.

[29] Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early detection of Fusarium infection in wheat using hyper-spectral imaging. Comput. Electron. Agric. 2011, 75, 304-312.

[30] Liu, Z.-y.; Huang, J.-f.; Shi, J.-j.; Tao, R.-x.; Zhou, W.; Zhang, L.-l. Characterizing and estimating rice brown spot disease severity using stepwise regression, principal component regression and partial least-square regression. J. Zhejiang Univ. - Sci. B 2007, 8, 738-744.

[31] Teófilo, R.F.; Martins, J.P.A.; Ferreira, M.M.C. Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression. J. Chemometr. 2009, 23, 32-48.

[32] Yang, C.-C.; Prasher, S.O.; Enright, P.; Madramootoo, C.; Burgess, M.; Goel, P.K.; Callum, I. Application of decision tree technology for image classification using remote sensing data. 2003, 76, 1101-1117.

[33] DeFries, R.S.; Chan, J.C.-W. Multiple Criteria for Evaluating Machine Learning Algorithms for Land Cover Classification from Satellite Data. Remote Sens. Environ. 2000, 74, 503-515.

[34] Zhang, H.; Hu, H.; Zhang, X.-b.; Zhu, L.-f.; Zheng, K.-f.; Jin, Q.-y.; Zeng, F.-p. Estimation of rice neck blasts severity using spectral reflectance based on BP-neural network. Acta Physiol. Plant. 2011, 33, 2461-2466.

[35] Monteiro, S.T.; Minekawa, Y.; Kosugi, Y.; Akazawa, T.; Oda, K. Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery. ISPRS-J. Photogramm. Remote Sens. 2007, 62, 2-12.

[36] Konduru, T.; Rains, G.C.; Li, C. Detecting sour skin infected onions using a customized gas sensor array. J. Food Eng. 2015, 160, 19-27.

[37] Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.C.; Dehne, H.W.; Plümer, L. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Comput. Electron. Agric. 2010, 74, 91-99.

[38] Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. ISPRS-J. Photogramm. Remote Sens. 2011, 66, 247-259.

[39] Dutta, S.; Singh, S.; Panigrahy, S. Assessment of Late Blight Induced Diseased Potato Crops: A Case Study for West Bengal District Using Temporal AWiFS and MODIS Data. J Indian Soc Remote Sens 2014, 42, 353-361.

[40] Zhang, M.; Qin, Z.; Liu, X. Remote Sensed Spectral Imagery to Detect Late Blight in Field Tomatoes. Precision Agric 2005, 6, 489-508.

[41] Blackburn, G.A. Hyperspectral remote sensing of plant pigments. J. Exp. Bot. 2006, 58, 855-867.

[42] Malthus, T.J.; Madeira, A.C. High resolution spectroradiometry: spectral reflectance of field bean leaves infected by Botrytis fabae. Remote Sens. Environ. 1993, 107-116.

[43] Yi, Q.-x.; Huang, J.-f.; Wang, F.-m.; Wang, X.-z. Quantifying biochemical variables of corn by hyperspectral reflectance at leaf scale. J. Zhejiang Univ. Sci. B 2008, 9, 378-384.

[44] Demetriades-Shah, T.H.; Steven, M.D.; Clark, J.A. High resolution derivative spectra in remote sensing. Remote Sens. Environ. 1990, 33, 55-64.

[45] Rundquist, D.C.H., L.; Schalles, J.F.; Peake, J.S. Remote Mesaurements of Algal Chlorophyll in Surface Waters: The Case for the First Derivative of Reflectance Near 690 nm. Photogramm. Eng. Remote Sens. 1996, 62, 195-200.

# Bibliography


Allen DE, Powell RJ and Singh AK. (2012) Chapter 32 - Machine Learning and Short Positions in Stock Trading Strategies. In: Gregoriou GN (ed) *Handbook of Short Selling.* San Diego: Academic Press, 467-478.

Barr AG and Brown DM. (1995) Estimating forage yield and quality changes during field drying for hay. 2. Model of forage drying. *Agricultural and Forest Meteorology* 76(2): 107-127.

Bin GF, Gao JJ, Li XJ, et al. (2012) Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network. *Mechanical Systems and Signal Processing* 27(0): 696-711.

Bishop CM. (2007) *Neural Networks for Pattern Recognition,* Oxford, UK: Oxford University Press.

Broomhead DS and Lowe D. (1988) Radial basis functions, multi-variable functional interpolation and adaptive networks. DTIC Document.

Chwan-Lu T, Shun-Yuan W, Foun-Yuan L, et al. (2014a) An intelligent motor rotary fault diagnosis system using Taguchi method. *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on.* 2311-2316.

Chwan-Lu T, Shun-Yuan W, Shou-Chuang L, et al. (2014b) A Motor Rotary Fault Diagnosis System Using Dynamic Structural Neural Network. *Computer, Consumer and Control (IS3C), 2014 International Symposium on.* 430-433.

Diamante LM and Munro PA. (1993) Mathematical modelling of the thin layer solar drying of sweet potato slices. *Solar Energy* 51(4): 271-276.

Draper NR, Smith H and Pownell E. (1966) *Applied regression analysis*: Wiley New York.

Duda RO, Hart PE and Stork DG. (2001) *Pattern classification*: John Wiley & Sons.

Erenturk S and Erenturk K. (2007) Comparison of genetic algorithm and neural network approaches for the drying process of carrot. *Journal of Food Engineering* 78(3): 905-912.

Esfandiari N, Babavalian MR, Moghadam A-ME, et al. (2014) Knowledge discovery in medicine: Current issue and future trend. *Expert Systems with Applications* 41(9): 4434-4463.

Garcia-Martin J, Martínez-Martínez V and Gomez-Gil J. (2014) Clasificación del tratamiento térmico de aceros con ensayos no destructivos por corrientes inducidas mediante redes neuronales. *Dyna* 89(5): 526-532.

Hagan MT, Demuth HB and Beale M. (1996) *Neural network design*: PWS Publishing Co.

Hawks SN, Collins WK, Carbonell HP, et al. (1986) *Tabaco flue-cured: principios básicos de su cultivo y curado*: Ministerio de Agricultura, Pesca y Alimentación.

Haykin SS. (1999) *Neural networks: a comprehensive foundation,* New Jersey, USA: Prentice Hall.

Köksal G, Batmaz İ and Testik MC. (2011) A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications* 38(10): 13448-13467.

Levenberg K. (1944) A method for the solution of certain non–linear problems in least squares. *Quarterly of Applied Mathematics* 2(2): 164-168.

Lewis M. (2004) *Moneyball: The art of winning an unfair game*: WW Norton & Company.

Li Y, Tse PW, Yang X, et al. (2010) EMD-based fault diagnosis for abnormal clearance between contacting components in a diesel engine. *Mechanical Systems and Signal Processing* 24(1): 193-210.

Li Z, Yan X, Tian Z, et al. (2013) Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis. *Measurement* 46(1): 259-271.

Liang SY, Li Y, Billington SA, et al. (2014) Adaptive Prognostics for Rotary Machineries. *Procedia Engineering* 86(0): 852-857.

Liao S-H, Chu P-H and Hsiao P-Y. (2012) Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications* 39(12): 11303-11311.

Lu N, Xiao Z and Malik OP. (2015) Feature extraction using adaptive multiwavelets and synthetic detection index for rotor fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing* 52–53(0): 393-415.

Marquardt DW. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics* 11(2): 431-441.

Martínez-Martínez V, Baladrón C, Gomez-Gil J, et al. (2012) Temperature and Relative Humidity Estimation and Prediction in the Tobacco Drying Process Using Artificial Neural Networks. *Sensors* 12(10): 14004-14021.

Martínez-Martínez V, Gomez-Gil FJ, Gomez-Gil J, et al. (2015a) 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 Systems with Applications* 42(17–18): 6433-6441.

Martínez-Martínez V, Gomez-Gil J, Stombaugh TS, et al. (2015b) Moisture Content Prediction in the Switchgrass (Panicum Virgatum) Drying Process Using Artificial Neural Networks. *Drying Technology* 33(14): 1708-1719.

Movagharnejad K and Nikzad M. (2007) Modeling of tomato drying using artificial neural network. *Computers and Electronics in Agriculture* 59(1–2): 78-85.

Mucherino A, Papajorgji PJ and Pardalos PM. (2009) *Data mining in agriculture*: Springer Science & Business Media.

Nembhard AD, Sinha JK and Yunusa-Kaltungo A. (2015) Development of a generic rotating machinery fault diagnosis approach insensitive to machine speed and support type. *Journal of Sound and Vibration* 337(0): 321-341.

Oliver D. (2004) *Basketball on paper: rules and tools for performance analysis*: Potomac Books, Inc.

Raghavan G and Mujumdar A. (1987) Drying of agricultural products. *Handbook of industrial drying*: 555-570.

Rosenblatt F. (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* 65(6): 386.

Ruiz-Gonzalez R, Gomez-Gil J, Gomez-Gil F, et al. (2014) An SVM-Based Classifier for Estimating the State of Various Rotating Components in Agro-Industrial Machinery with a Vibration Signal Acquired from a Single Point on the Machine Chassis. *Sensors* 14(11): 20713-20735.

Rumelhart DE, McClelland JL and Group PR. (1988) *Parallel distributed processing*: IEEE.

Wang J. (2002) A single-layer model for far-infrared radiation drying of onion slices. *Drying Technology* 20(10): 1941-1953.

Yang Y, Dong XJ, Peng ZK, et al. (2015) Vibration signal analysis using parameterized time–frequency method for features extraction of varying-speed rotary machinery. *Journal of Sound and Vibration* 335(0): 350-366.