

Number 23

Redalyc 3.0

ESCI Web of science

UNIVERSIDAD POLITÉCNICA SALESIANA ECUADOR

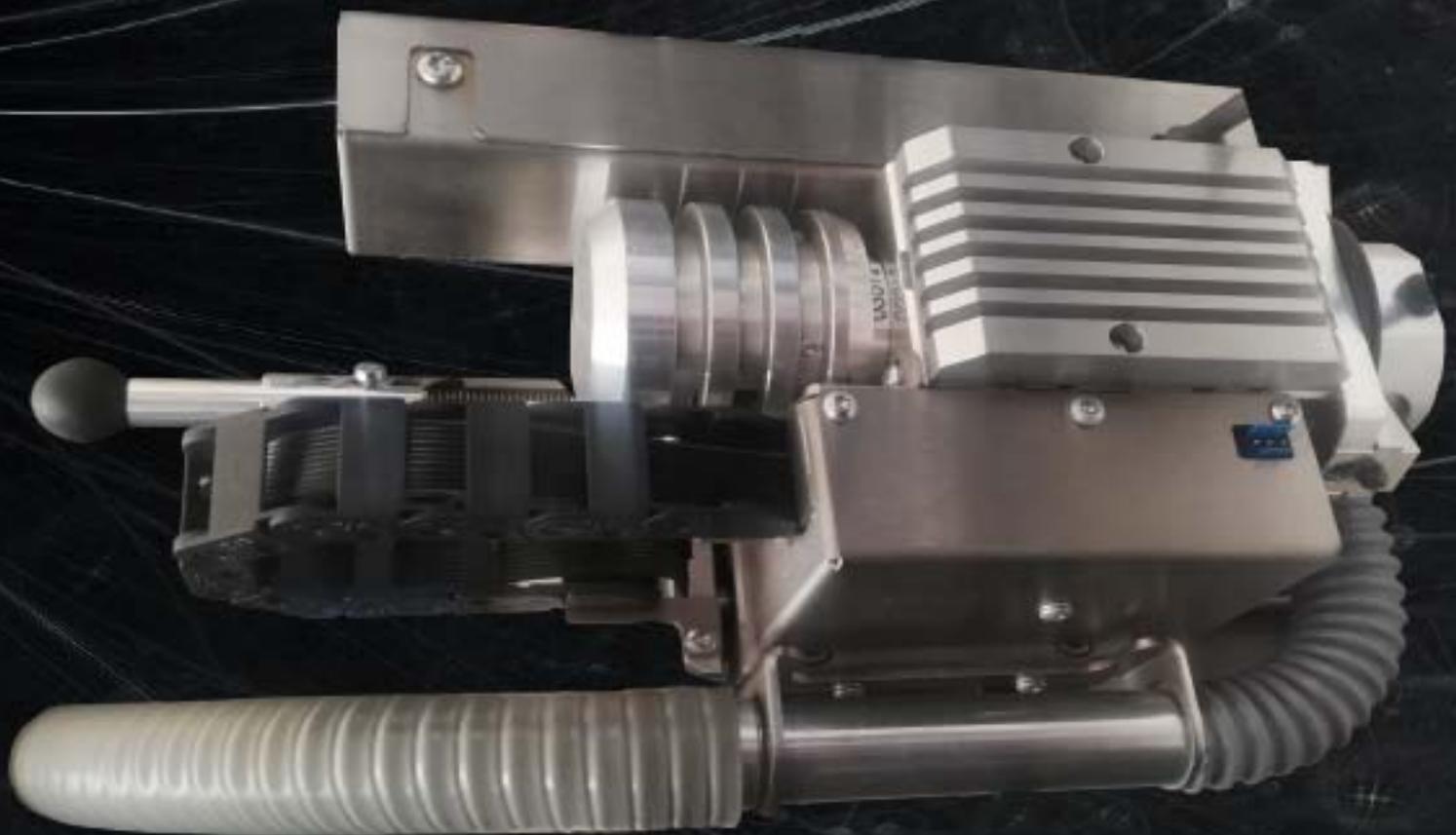
pISSN : 1390-650X

eISSN : 1390-860X

January/June 2020

# INGENIUS

Revista de Ciencia y Tecnología



- Experimental and numerical study of the pressure of the water flow in a Venturi Tube

Pag. 9

- Number of subcarrier filter coefficients in GFDM System: effect on performance

Pag. 53

- Parametric study for optimization of the ice tube generator laboratory

Pag. 86

- Robust filtering of weak signals from real phenomena

Pag. 109



# EXPERIMENTAL AND NUMERICAL STUDY OF THE PRESSURE OF THE WATER FLOW IN A VENTURI TUBE

## ESTUDIO EXPERIMENTAL Y NUMÉRICO DE LA PRESIÓN DEL FLUJO DE AGUA EN UN TUBO VENTURI

San Luis B. Tolentino Masgo<sup>1,2</sup>

### Abstract

The Venturi tube is a device used to measure the flow rate in different industrial processes. In the present work, a study is carried out for two cases, one experimental and another numerical of the pressure exerted by the flow of water on the walls of a Venturi tube. In the first case, five experiments with different flow rates are carried out. In the second, the flow is simulated for two types of meshes and two turbulence models, using the code COMSOL Multphysics 4.3. The experimental and numerical results showed that the pressures of the flow on the walls in two references identified as C and G keep their magnitude constant; in addition, the numerical profiles showed that the lowest pressure drop occurs in the wall at the inlet and outlet of the throat section. It is concluded that, the distribution of the flow pressure in the wall of the throat section has a convex profile, and the results of pressures obtained for the standard  $k - e$  turbulence model are more adjusted to the experimental data.

**Keywords:** Water flow, Turbulence model, Pressure, Simulation, Venturi tube.

### Resumen

El tubo Venturi es un dispositivo utilizado para medir el caudal en diferentes procesos de la industria. En el presente trabajo, se realiza un estudio para dos casos, uno experimental y otro numérico de la presión ejercida por el flujo de agua en las paredes de un tubo Venturi. En el primer caso, se realizan cinco experimentos con diferentes caudales. En el segundo, el flujo se simula para dos tipos de mallas y dos modelos de turbulencia, utilizando el código COMSOL Multphysics 4.3. Los resultados experimentales y numéricos mostraron que las presiones del flujo sobre las paredes en dos referencias identificadas C y G mantienen constante su magnitud; además, los perfiles numéricos mostraron que la menor caída de presión se presenta en la pared a la entrada y salida de la sección de la garganta. Se concluye que, la distribución de la presión del flujo en la pared de la sección de la garganta tiene un perfil convexo, y los resultados de presiones obtenidos para el modelo de turbulencia  $k-e$  estándar, se ajustan más a los datos experimentales.

**Palabras clave:** flujo de agua, modelo de turbulencia, presión, simulación, tubo Venturi.

<sup>1,\*</sup>Department of Mechanical Engineering, Universidad Nacional Experimental Politécnica “Antonio José de Sucre” Vice-Rectorado Puerto Ordaz, Bolívar, Venezuela.

<sup>2</sup>Group of Mathematical Modelling and Numerical Simulation, Universidad Nacional de Ingeniería, Lima, Perú.  
Corresponding author ✉: sanluis@gmail.com <http://orcid.org/0000-0001-6320-6864>

Received: 28-06-2019, accepted after review: 29-10-2019

Suggested citation: Tolentino Masgo, San Luis B. (2020). «Experimental and numerical study of the pressure of the water flow in a venturi tube». INGENIUS. N.º 23, (january-june). pp. 9-. DOI: <https://doi.org/10.17163/ings.n23.2020.01>.

## 1. Introduction

In the field of engineering, the instruments used to measure the flow rate of a fluid are classified in two types, mechanical instruments and pressure loss instruments [1, 2], and are described in detail in a book by ASME [3]. The Venturi tube is among the latter as the pressure loss occurs in a constricted section, and was invented by Clemens Herschel (1842-1930) [4], and named after Giovanni Battista Venturi (1746-1822) for his pioneering works about flow in conic sections [2].

The flows in such instruments are, mostly, turbulent, and characterized by having random and rapid fluctuations of eddies, which transport mass, energy and momentum to other regions of the flow, where the fluctuations add movement and energy transfer and, besides, are related to high values of friction coefficients, heat transfer and mass transfer [1, 2].

By means of the dimensionless Reynolds number [5, 6] it is determined if the flow is laminar or turbulent, where the Reynolds number is the parameter that express the relationship between the inertia and viscous forces in a fluid, being the inertia force the product of the average speed and the internal diameter, and of the viscous force the kinematic viscosity. The flow is considered as laminar for a value of the Reynolds number smaller than 2300, as transitional for a value of the Reynolds number in the range 2300-4000, and as turbulent for a value of the Reynolds number greater than 4000, thus resulting in the graphical representation known as Moody diagram [1, 2]. Other studies carried out on conditioned surfaces report that the Reynolds number takes different values for laminar and turbulent flow.

In order to reproduce the behavior of the flow by means of computational fluid dynamics (CFD) [7], experimental data are required for its calibration and validation, such as: pressure, temperature, velocity, which are parameters that can be measured in the laboratory or in open spaces. Besides, the CFD requires utilizing a turbulence model that solves the closure problem in the averaged equations of the Navier-Stokes general equations. Among the diversity of turbulence models for the flow of fluids, the profiles of the numerical solutions that better fit the experimental results should be chosen. Once the turbulence model has been evaluated, it is possible to continue developing the desired computational simulations to determine the behavior of the flow field under certain considerations, whether for internal or external flow. It is worth noting that, in certain critical regions, the numerical results evaluated with different turbulence models [8] are distinct, and in less critical regions the results are similar, due to the boundary layer and the interaction of the flow separation [9, 10].

An experimental research was carried out by Lindley [11] about the transit of flow in a classical Venturi

tube, where at the beginning of the throat section, the water flow experienced pressure drops in a section of the wall. Afterwards, for the same geometry, Sattery and Reader [12] simulated the flow through CFD, where in the throat section the numerical profiles showed pressure drops. A similar result of the pressure drop was also obtained by Tamhakar *et al.* [13], applying CFD, for a Venturi tube of different dimensions. Whereby, this constitutes a motivation for continuing investigating about the pressure of the flow in Venturi meters.

In the present work, an experimental and numerical study is carried out about the pressure exerted by the flow of water on the walls of a Venturi tube. In the experimental case the study was conducted for five experiments with different flow rates, and in the numerical case the study is conducted for two turbulence models: standard  $k - \epsilon$  of Launder and Spalding [14] and standard  $k - \omega$  of Wilcox [15], and for two domains discretized by the finite elements method (FEM), one with quadrilateral cells and another with triangular cells, in order to determine what regions of the flow experience the minimum pressure drops. Similarly, validate the two turbulence models with the experimental data of pressure. It is worth noting that the experimental values of pressure are measured at specific points that are separated from each other by a defined distance along the wall of the Venturi tube, and does not provide information about pressure in critical sections; while, the numerical profiles are continuous trajectories and provide more information about the pressure along the entire wall.

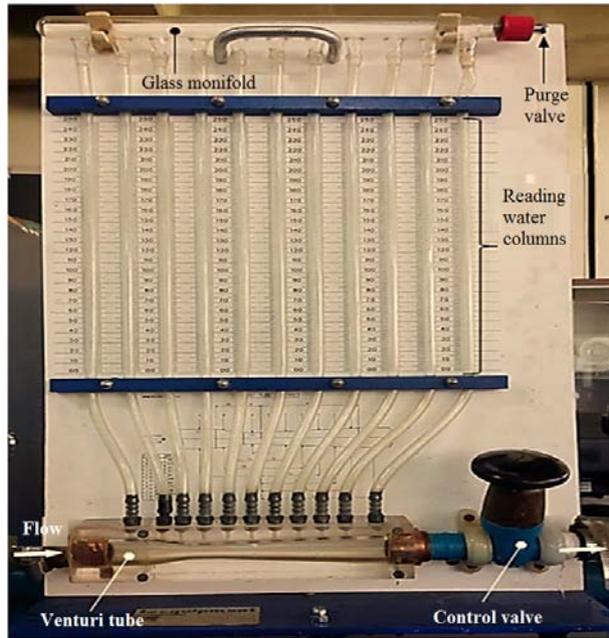
The mathematical foundation is presented, and the procedures and results of the experiments and of the numerical simulations are exposed. Afterwards, the conclusions of the analysis are presented.

## 2. Materials and methods

### 2.1. Experimental facility

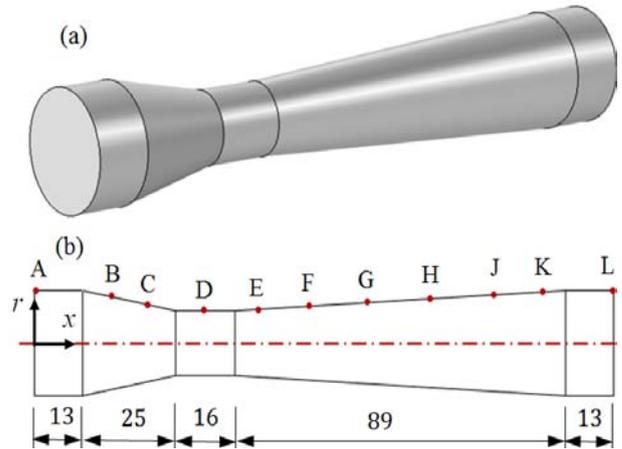
The Venturi tube under study in this work is shown in Figure 1. It is a device manufactured by the company TecQuipment, which is utilized in experimental tests to determine the water flow rate by the pressure difference. The experimental equipment is installed in the laboratory of the Section of Thermal Fluids of the Mechanical Engineering Department of the National Experimental Polytechnic University “Antonio José de Sucre”, Puerto Ordaz Vice-Rectorate, Bolívar, Venezuela. The Venturi tube comprises five sections: a straight section, a convergent section, a straight section which is the throat, a divergent section and another straight section. The image shows a total of eleven plastic hoses, called piezometric pipes, which measure the readings of piezometric heights of the water columns, and are connected to the Venturi tube and to the glass

manifold. The control valve is located at the outlet of the Venturi tube, and the purge valve is located at the right end of the collector. The valve for regulating the flow rate is ahead of the Venturi tube and after the pump, which is not shown in the figure.



**Figure 1.** Venturi tube experimental equipment. The readings of the water columns are measured in millimeters.

The 3D geometry and the projection on the plane of the Venturi tube are illustrated in Figure 2, which also shows the location of the references A, B, C, D, E, F, G, H, J, K and L, places to where the eleven plastic hoses are connected. Reference A is located at the beginning of the straight section; B and C are located in the convergent section; D in the middle of the straight section of the throat; E, F, G, H, J and K are located in the divergent section; and L is located at the end of the straight section, at the outlet of the Venturi tube. The greater internal diameter of the two straight sections is 26 mm, the internal diameter of the throat is 16 mm, and the total length of the Venturi tube is 156 mm. The internal diameters of the cross sections and of the locations of the references are shown in Table 1.

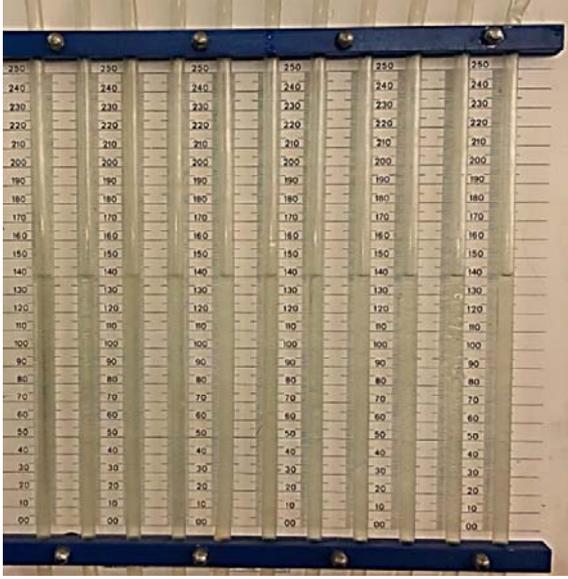


**Figure 2.** (a) 3D geometry of the Venturi tube. (b) Location of the references and dimensions of the longitudinal sections in millimeters.

**Table 1.** Internal diameter for each reference, and axial distance where the references are located

Ref.	Internal diameter (mm)	Axial distance, x axis (mm)
A	26	0
B	23,2	20
C	18,4	32
D	16	46
E	16,8	61
F	18,47	76
G	20,16	91
H	21,84	106
J	23,53	121
K	25,24	136
L	26	156

The experimental test was carried out through the following steps: initially, both valves, the valves for flow rate regulation and for flow rate control, were opened at 100 %. Once the pump of the test bench was in operation, the control valve was closed 100 %, the air trapped in the hoses and in the collector was let out through the purge valve leaving it totally full of water; subsequently, the regulation valve was closed 100 %. Then the control valve was opened 100 %, and through the purge valve air from the local atmosphere was let in, thus allowing the formation of water columns at the established height of 140 mm as initial position, all at the same level, for the eleven hoses. The reading of 140 mm remained inside the range of 0.0-200 mm of the panel, as shown in Figure 3. Afterwards, five experimental tests were carried out increasing the opening of the regulation valve, for visually taking the piezometric readings for the range of flow rate  $2,244 \times 10^{-4} - 3,7 \times 10^{-4} \text{ (m}^3/\text{s)}$ .



**Figure 3.** Initial position of the piezometric heights of 140 mm of the levels of the eleven water columns. From left to right, the first piezometric tube is connected in the reference A of the wall of the Venturi tube, the second in the reference B, and similarly, the remaining are located up to the reference L (see Figure 2).

## 2.2. Numerical simulation

### 2.2.1. Governing equations

The governing equations applied to the CFD, for an incompressible flow, in stationary conditions, and simulated for a 2D computational domain with axial symmetry, in their differential form are expressed as:

Equation of conservation of mass.

$$\frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial r}(\rho v_r) + \frac{\rho v_r}{r} = 0 \quad (1)$$

Equation of conservation of linear momentum, in the axial direction.

$$\begin{aligned} \frac{1}{r} \frac{\partial}{\partial x}(r \rho v_x v_x) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r v_x) &= -\frac{\partial P}{\partial x} + \\ \frac{1}{r} \frac{\partial}{\partial x} \left[ r \mu \left( 2 \frac{\partial v_x}{\partial x} - \frac{2}{3} (\nabla \cdot \vec{v}) \right) \right] + & \\ \frac{1}{r} \frac{\partial}{\partial r} \left[ r \mu \left( \frac{\partial v_x}{\partial r} + \frac{\partial v_r}{\partial x} \right) \right] + F_x & \end{aligned} \quad (2)$$

in the radial direction.

$$\begin{aligned} \frac{1}{r} \frac{\partial}{\partial x}(r \rho v_x v_r) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r v_r) &= -\frac{\partial P}{\partial r} + \\ \frac{1}{r} \frac{\partial}{\partial r} \left[ r \mu \left( \frac{\partial v_r}{\partial x} - \frac{\partial v_x}{\partial r} \right) \right] - 2 \mu \frac{v_r}{r^2} + \frac{2}{3} \frac{\mu}{r} (\nabla \cdot \vec{v}) + & \\ \frac{1}{r} \frac{\partial}{\partial r} \left[ r \mu \left( 2 \frac{\partial v_r}{\partial r} - \frac{2}{3} (\nabla \cdot \vec{v}) \right) \right] + F_r & \end{aligned} \quad (3)$$

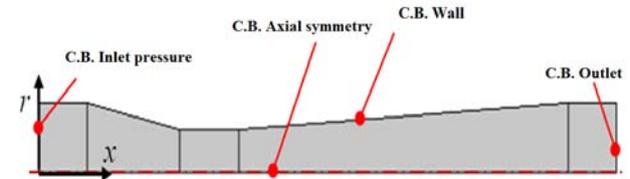
with  $\nabla \cdot \vec{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_r}{\partial r} + \frac{v_r}{r}$

where the parameters are: the density  $\rho$ , the axial velocity  $v_x$  and radial velocity  $v_r$ , the radius  $r$ , the viscosity  $\mu$ , the pressure gradients  $\frac{\partial P}{\partial x}$  and  $\frac{\partial P}{\partial r}$ , and the forces in the axial direction  $F_x$  and in the radial direction  $F_r$ .

The turbulence model is coupled to the equation of linear momentum, and are semi-empirical transport equations that model the mixing and diffusion that increase because of the turbulent eddies, and are solved through the Reynolds average number Navier-Stokes equation (RANS) [16]. The initial research studies about turbulence were conducted by Kolmogorov (1941), based on the results obtained by Reynolds (1883). It is worth noting that the turbulence models standard k-e of Launder and Spalding [14] and standard k- $\omega$  of Wilcox [15], are employed in the present work for the simulation of the flow.

### 2.2.2. Computational domain and mesh

The 2D computational domain with axial symmetry shown in Figure 4 is considered, due to the symmetry of the geometry of the Venturi tube. This simplification of the geometry from 3D to 2D contributes to reduce the number of cells in the mesh, the processing time, and the computational cost; the simplifications are very common for solids of revolution and symmetric primitive geometries. Besides, in the same figure of the 2D domain, the places where the boundary conditions are applied have been marked.



**Figure 4.** 2D computational domain with axial symmetry in the x axis, of the Venturi tube.

Figure 5 shows the 2D meshed domain, in which two types of cells are used, quadrilateral and triangular. The domain meshed with quadrilateral cells has 19600 elements, and the domain meshed with triangular cells 18047 elements. For both cases, the mesh was refined in the regions adjacent to the walls, due to the presence of shear stresses in those regions of flow. The throat section is also shown in detail in the same figure, where it is observed how the quadrilateral and triangular cells are distributed.

As part of a study of the numerical convergence, before obtaining the final mesh which is shown in Figure 5, the throat section was refined five times until obtaining an optimum mesh density. Such refinement in the throat was because this is a critical section due to pressure drop in the flow. The reference D is located in the middle of the throat length

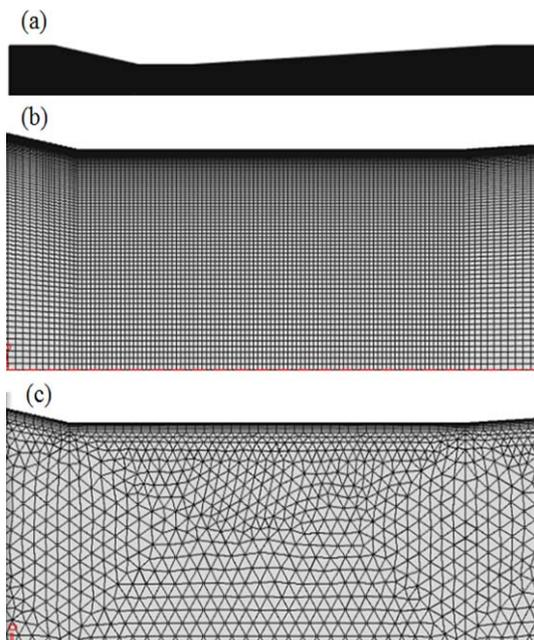
(see Figure 2), where the final numerical result of the pressure was 44.79 (mmH<sub>2</sub>O) for the mesh with quadrilateral cells and 51.38 (mmH<sub>2</sub>O) for the mesh with triangular cells, evaluated with the standard  $k - e$  turbulence model; and the pressure was 48.53 (mmH<sub>2</sub>O) for the mesh with quadrilateral cells and 55.75 (mmH<sub>2</sub>O) for the triangular cells, evaluated with the standard  $k - \omega$  turbulence model; obtaining for both cases numerical convergence errors smaller than 0.01 %.

The quality of the mesh was evaluated for two-dimensional cells, where for the case of quadrilateral cells it was obtained a maximum element size of 0.0105 mm, a minimum element size of  $4.68 \times 10^{-5}$  mm, a curvature of 0.3, and a rate of increase of 1.3; similarly, for the case of triangular cells it was obtained a maximum element size of  $3.64 \times 10^{-4}$  mm, a minimum element size of  $5.2 \times 10^{-5}$  mm, a curvature of 0.25, and a rate of increase of 1.15. These final results indicate that the two domains meshed with quadrilateral and triangular cells are of good quality.

The computational domains were discretized in the mesh platform of the code COMSOL Multiphysics version 4.3, which applies the finite element method (FEM).

The boundary conditions for the pressures of the water flow applied at the inlet (reference A) and at the outlet (reference L) of the Venturi tube, are shown in Table 2.

The walls of the Venturi tube are considered adiabatic. The velocity of the flow at the walls in the radial and axial direction is zero due to the presence of shear stresses. In the axial symmetry in the x axis, the velocity of the flow in the radial direction is zero.



**Figure 5.** (a) Meshed 2D computational domain. Detail of the throat section, (b) Structured mesh with quadrilateral cells, (c) Mesh with triangular cells.

Isothermal flow is considered along the entire computational domain, and for a water temperature of 24 °C the density was 997.1015 kg/m<sup>3</sup> and the dynamic viscosity 0.00091135 Pa.s, where both physical parameters are set as constants for the simulation of the flow.

**Table 2.** Boundary conditions: inlet pressure (reference A), outlet pressure (reference L)

Exp.	1	2	3	4	5
<b>Water column (mmH<sub>2</sub>O)</b>					
<b>Ref. A</b>	160	170	179	190	199,5
<b>Ref. L</b>	150	155,5	161	167	173

### 2.2.3. Method of computational solution and equipment

A 2D geometry with axial symmetry and stationary flow conditions, was chosen as the option for the simulation of the isothermal flow in the COMSOL Multiphysics code. The turbulence models standard  $k - e$  and standard  $k - \omega$  were applied for the turbulent flow, for the domains meshed with both quadrilateral and triangular cells. A fixed value of 0.001 was determined for the relative tolerance. For the solution, the maximum number of iterations was established as 100, and the solution method parallel sparse direct solver (PARDISO) was employed.

For data processing, an equipment with the following characteristics was utilized: Siragon Laptop, model M54R, Intel Core 2 Duo, two 1.8 GHz processors, and a RAM memory of 3 GB.

## 3. Results and discussion

### 3.1. Experimental results

For each experiment that was carried out, the flow rates of the water were obtained by means of the volumetric method, and the results are shown in Table 3. The Reynolds number was determined with these values of flow rate.

**Table 3.** Experimental data of flow rates

Exp.	1	2	3	4	5
<b>Caudal <math>\times 10^4</math> (m<sup>3</sup>/s)</b>					
	2,244	2,583	2,991	3,382	3,704

Table 4 presents the magnitudes of the Reynolds number obtained in the references A, L and D. In reference A the Reynolds number has the same magnitude than in reference L, and this is because the Venturi tube has the same diameter. For the five experiments,

the Reynolds number obtained in the references A and L was in the range  $12000 < Re < 20000$ , and in the throat section, in the reference D, was in the range  $19500 < Re < 32300$ . Whereby, it is observed that as the flow rate increases, the magnitude of the Reynolds number also increases, and this is due to the increase in the fluid inertial forces.

**Table 4.** Experimental data of the Reynolds number

Exp.	1	2	3	4	5
Reynolds number					
Ref. A and L	12021	13839	116028	18118	19844
Ref. D	19534	22489	26045	29442	32246

Table 5 shows the five experimental results of the piezometric heights obtained in the references A, B, C, D, E, F, G, H, J, K and L, Figure 6 shows the graph of the gauge pressure in (mmH<sub>2</sub>O), and Figure 7 shows an image of the piezometric heights of the eleven water columns corresponding to experiment 3, and it is observed the concave shape of the menisci due to the effect of the surface tension of the water that is formed in each column.

Taking as reference the initial position of the eleven columns with the water level at the height of 140 mm, it is observed that as the flow rate increases, the level of the water columns in the references A and B increases; in the reference C remains constant; in the references D, E and F decreases; in the reference G it also remains constant; and in the references H, J, K and L increases. Taking into account the differences of the water columns between A and D, the smallest pressure difference occurs for experiment 1, and the largest difference for experiment 5. The losses in the fluid pressure in the reference L with respect to the reference A are also observed in Figure 6. Among all experiments that were carried out, the one with the largest pressure loss is experiment 5.

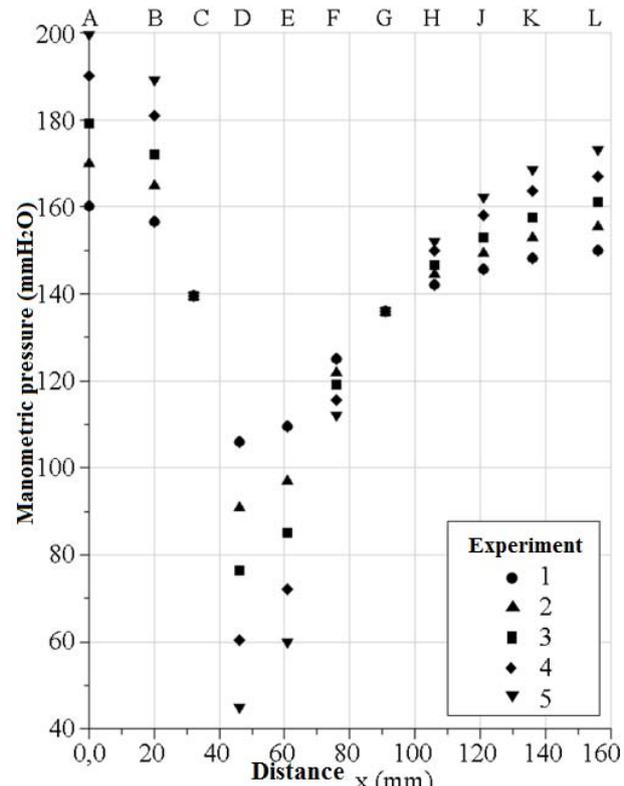
**Table 5.** Experimental data of gauge pressure in each reference of the wall of the Venturi tube Temperature of the water: 24 °C

Exp.	1	2	3	4	5
Ref. Water column (mmH <sub>2</sub> O)					
A	160	170	179	190	199,5
B	156,5	165	172	181	189
C	139,5	139,5	139,5	139,5	139,5
D	106	91	76,5	60,5	45
E	109,5	97	85	72	60
F	125	122	119	115,5	112
G	136	136	136	136	136
H	142	144,5	146,5	150	152
J	145,5	149,5	153	158	162
K	148	153	157,5	163,5	168,5
L	150	155,5	161	167	173

For each experiment, the pressure drop appears in the throat section, as it is shown in the reference D, where the fluid is forced to have a pressure drop, which is known as Venturi effect, and the smallest pressure drop occurs for experiment 5.

Among the results obtained in the references C, where the internal diameter is 18.4 mm, and G where the diameter is 20.16 mm, even though the flow velocity increases in those references due to the increase of the flow rate, the magnitude of the pressures do not vary coming together to a unique place for each reference. This occurs because the control valve remains open at 100 % and the flow is discharged in a reservoir at the local atmospheric pressure, and simultaneously the pressure of the air trapped in the collector, which is smaller than the local atmospheric pressure, holds the formation of the water columns.

It is worth to mention that, if the opening of the regulation valve is kept constant and the control valve is manipulated, there would be no interception in the references C and G, because the level of the water columns would move up or down due to the increase or decrease of the pressure of the flow in all the system of the Venturi tube.



**Figure 6.** Experimental data of readings of the piezometric heights of the water columns taken at different reference points of the walls of the Venturi tube Temperature of the water 24 °C.



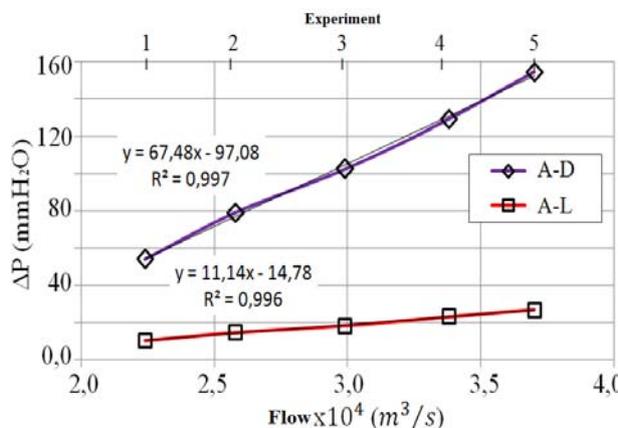
**Figure 7.** Piezometric heights for different levels of the water columns, corresponding to experiment 3.

On the other hand, if the pressure loss from the reference C to the reference G is analyzed, the pressure drop remains invariant even though the kinetic energy of the fluid is increased in such regions because, as shown in Table 5 and plotted in Figure 6, there is a fixed piezometric height of 139.5 mm in the reference C and of 136.0 mm in the reference G, where the pressure difference is 3.5 (mmH<sub>2</sub>O) for the five experiments that were conducted.

The pressure differences between the references A-D and A-L are presented in Table 6, and Figure 8 shows the behavior of both straight lines by means of the linear trend line, which has a determination coefficient  $R^2=0.997$  for the A-D pressure difference, and a value  $R^2=0.996$  for the A-L pressure difference. Both results show that there is a proportionality of pressure difference with respect to the flow rate.

**Table 6.** Pressure difference between the references A and D, and the references A and L, for each experiment

Exp.	1	2	3	4	5
$\Delta P$ mmH <sub>2</sub> O					
Ref. A-D	54	79	102,5	129,5	154,5
Ref. A-L	10	14,5	18	23	26,5



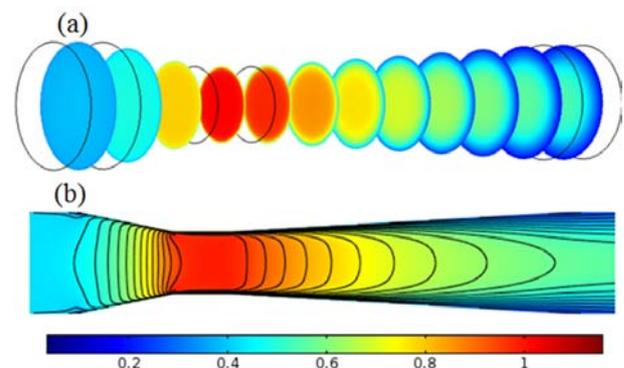
**Figure 8.** Trend lines and coefficient of determination  $R^2$ .

During the operation of the pump, for taking the experimental readings there were vibrations in the test bench. Therefore, after an estimated time between four and six minutes, when the disturbances were minimum, it was proceeded to take the readings by direct observation of the level of each of the eleven water columns in the measurement unit of one millimeter, and a magnifying glass was used for amplifying the image when taking a reading of the water level located in the middle of the unit of one millimeter.

The obtained experimental results do not quantify the magnitude of the pressure between each reference, because they have a separation distance. Therefore, it is of interest to quantify and know the behavior in a continuous manner of the trajectory of the pressure profile along the walls of the Venturi tube, being of greatest interest between the references C and E, because this is the place where the largest pressure drops occur. For this purpose, the flow should be simulated through CFD, and thus know what could really happen.

### 3.2. Numerical results and comparison with the experimental data

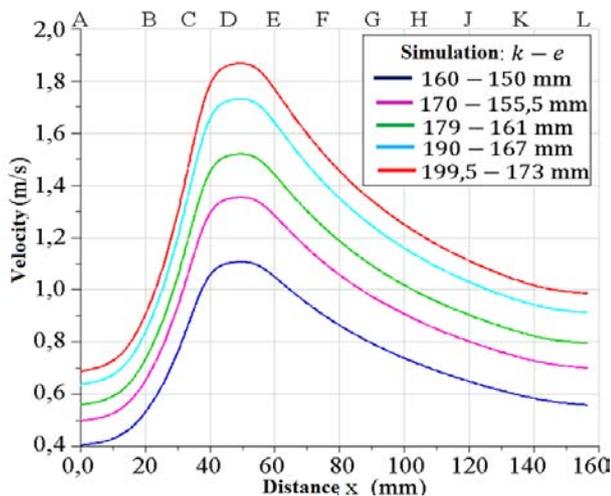
The simulation of the velocity distribution of the isothermal flow in the Venturi tube is shown in Figure 9, both in the cross section and in the plane, where the increase of the velocity of the flow occurs in the throat, and the decrease in the divergent section. In the latter the contour lines of velocity acquire a parabolic profile in direction to the x axis, due to the effect of the boundary layer. The velocity of the flow is maximum in the axial symmetry of the x axis, and its magnitude decreases toward the walls of the Venturi tube, thus having a gradient of velocity in the flow field. It is worth noting that the domain of the flow shown in Figure 9, was simulated with structured mesh and quadrilateral cells, employing the standard  $k - \epsilon$  turbulence model; since they are similar, other figures of the contour lines of velocity for the results of the standard  $k - \omega$  turbulence model are not shown.



**Figure 9.** Distribution of velocity (m/s). (a) In the cross section, and (b) Projected on the plane.

Figure 10 shows the behavior of the velocity profiles evaluated in the axial symmetry of the x axis. According to all the trajectories of the profiles, toward the end of the convergent section the flow increases its velocity, in the throat section reaches a maximum velocity in the reference D, and decreases its velocity in the divergent section. The magnitude of the velocity of the flow in the reference A is smaller with respect to the reference L, thus it is understood that the behavior of the profile of the flow velocity in the radial direction has smaller curvature in the reference A, and larger curvature in the reference L.

The numerical results of the flow rates are shown in Table 7, for each turbulence model and type of mesh. The largest magnitude of flow rate was for the standard k-e turbulence model and domain meshed with quadrilateral cells, and the smallest flow rate for the standard k- $\omega$  turbulence model and domain meshed with triangular cells. It should be noted that the flow rate was determined with the average velocity of the flow, using the numerical integration method.



**Figure 10.** Profiles of velocity evaluated in the axial symmetry of the x axis, for the standard k-e turbulence model.

When the numerical flow rates presented in Table 7 are compared with the experimental values presented in Table 3, it can be seen that the greatest percentage error was 9.68 % for the standard k-e turbulence model and the domain meshed with quadrilateral cells, and 8.68 % for the mesh with triangular cells; the minimum percentage error was 1.48 % for the standard k-e and the mesh with quadrilateral cells, and 1.01 % for the standard k- $\omega$  and the mesh with triangular cells, as shown in Table 8. Based on the results, it is evident that there is an influence on the numerical results of the type of mesh applied to the computational domain.

**Table 7.** Flow rates obtained for two turbulence models and two types of applied meshes

Simulation (mm H <sub>2</sub> O)	Flow rate $\times 10^4 (m^3/s)$			
	Quadrilateral c.		Triangular c.	
	k-e	k- $\omega$	k-e	k- $\omega$
160-150	2,29	2,27	2,28	2,26
170-155,5	2,83	2,8	2,8	2,78
179-161	3,19	3,16	3,15	3,12
190-167	3,65	3,61	3,6	3,56
199,5-173	3,94	3,89	3,88	3,84

**Table 8.** Percentage error of the flow rates

Exp	Simulation (mm H <sub>2</sub> O)	Percentage error (%)			
		Quadrilateral C.		Triangular C.	
		k-e	k- $\omega$	k-e	k- $\omega$
1	160-150	2,33	1,48	1,73	1,01
2	170-155,5	9,68	8,64	8,68	7,72
3	179-161	6,74	5,68	5,56	4,57
4	190-167	7,93	6,76	6,5	5,46
5	199,5-173	6,47	5,14	4,9	3,85

The numerical results of Reynolds number obtained for the greatest diameter in the references A and L are shown in Table 9, and for the smallest diameter in the reference D are shown in Table 10. In the references A and L, for the five simulations, the Reynolds number was in the range  $12100 < Re < 21200$ ; and in the reference D, in the range  $19700 < Re < 34400$ . After comparing the numerical results of the Reynolds number presented in Tables 9 and 10, with respect to the experimental values of the Reynolds number presented in Table 4, the percentage error was in the range 1.02-9.68 %.

**Table 9.** Reynolds numbers obtained in the greatest diameter, in the references A and L, for two turbulence models and two types of applied meshes

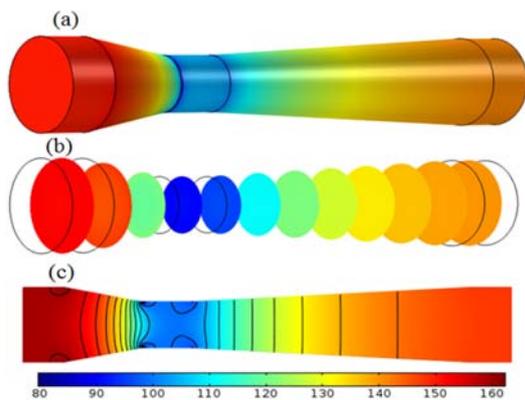
Simulation (mm H <sub>2</sub> O)	Reynolds Numbers in the references A y L			
	Quadrilateral C.		Triangular C.	
	k-e	k- $\omega$	k-e	k- $\omega$
160-150	12303	12201	12231	12144
170-155,5	15179	15034	15041	14908
179-161	17106	16935	16917	16758
190-167	19558	19345	19298	19109
199,5-173	21130	20866	20818	20610

**Table 10.** Reynolds numbers obtained in the smallest diameter, in the reference D, for two turbulence models and two types of applied meshes

Simulación (mm H <sub>2</sub> O)	Reynolds Numbers in the reference D			
	Cuadrilateral C.		Triangular C.	
	<i>k-e</i>	<i>k-ω</i>	<i>k-e</i>	<i>k-ω</i>
160-150	19992	19827	19875	19735
170-155,5	24666	24431	24442	24226
179-161	27797	27519	27490	27231
190-167	31781	31436	31359	31053
199,5-173	34336	33907	33829	33492

Figure 11 shows the pressure distributions on the surface of the walls of the convergent, throat and divergent sections, and also the pressure distribution of the internal flow projected on the plane, with the contour lines which are the isobar, and it can be observed how the pressures are distributed in a manner perpendicular to the x axis and toward the walls.

The profiles of the gauge pressures obtained at the walls of the Venturi tube are shown in Figures 12 to 16. The unit of the pressure is millimeters of water column (mmH<sub>2</sub>O), and the same graphs include the pressure data of the experiments that were carried out. The pressure profiles correspond to a domain meshed with quadrilateral cells, and for another domain meshed with triangular cells, with both cases simulated for isothermal flow with the turbulence models standard *k-e* and standard *k-ω*.

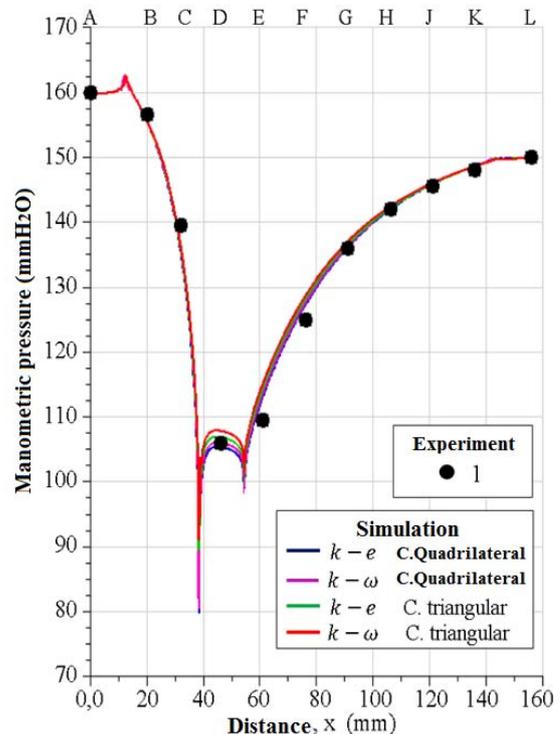


**Figure 11.** Pressure distribution (mm<sub>2</sub>O). (a) On the surface of the wall, (b) On the cross section, and (c) Projected on the plane.

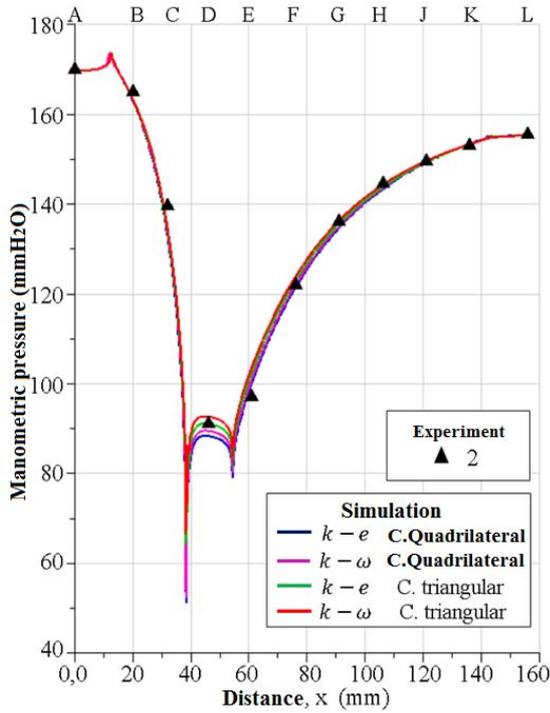
A simultaneous observation of Figures 12 to 16 shows that all pressure profiles have a similar behavior along the walls of the Venturi tube, the experimental data are intercepted and bordered. At the end of the straight section of the inlet of the Venturi tube, the profiles show a pressure increase to a magnitude greater than the inlet pressure. In the vertex between the convergent section and the throat, which is located

at a distance of 38 mm, an abrupt pressure drop occurs, even yielding a negative pressure for the profile *k-e* with quadrilateral cells for the simulation of experiment 4, and the same happens for the profiles *k-e* and *k-ω* with quadrilateral cells for the simulation of experiment 5. Similarly, it occurs in the region of the other vertex located at the right end of the throat at a distance of 54 mm, with a pressure drop of lower intensity but without the presence of negative pressure.

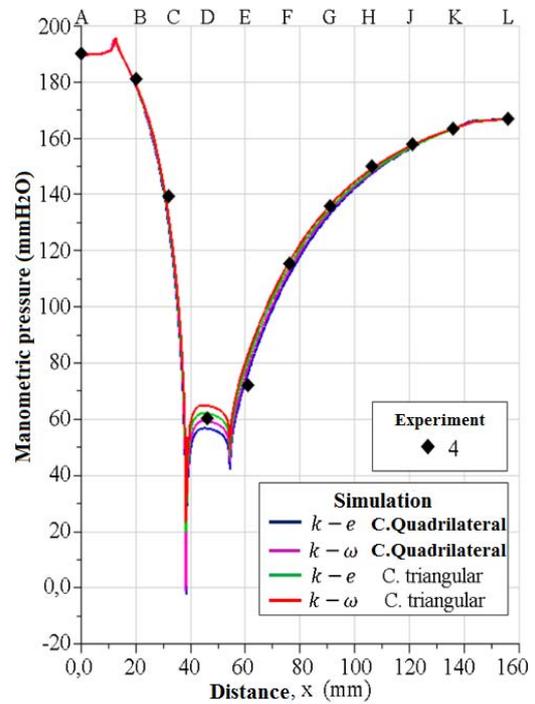
In the throat section, a stretch of the profiles exhibits a convex behavior, separated one from the other, ending up in the following order from the bottom up: the profile *k-e* and the profile *k-ω*, both simulated for the domain of mesh with quadrilateral cells, followed by the profile *k-e* and the profile *k-ω*, both simulated for the domain of mesh with triangular cells. It is observed how the five experimental data of pressure in the reference D were ranked. The stretches of the trajectories of the profiles *k-e* with quadrilateral cells have closer approximation to the experimental data. It should be pointed out that the pressure drop in the throat occurs because the flow passes through the narrow section at a higher velocity, due to the large difference of the flow pressure between the inlet and the outlet of the Venturi tube.



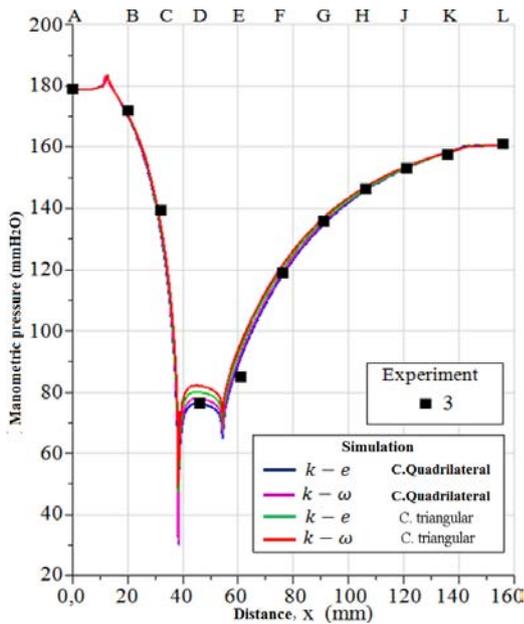
**Figure 12.** Experiment 1 of pressures of water columns, and pressure profiles evaluated at the wall of the Venturi tube. Pressure in (mmH<sub>2</sub>O): 160 mm at the inlet and 150 mm at the outlet.



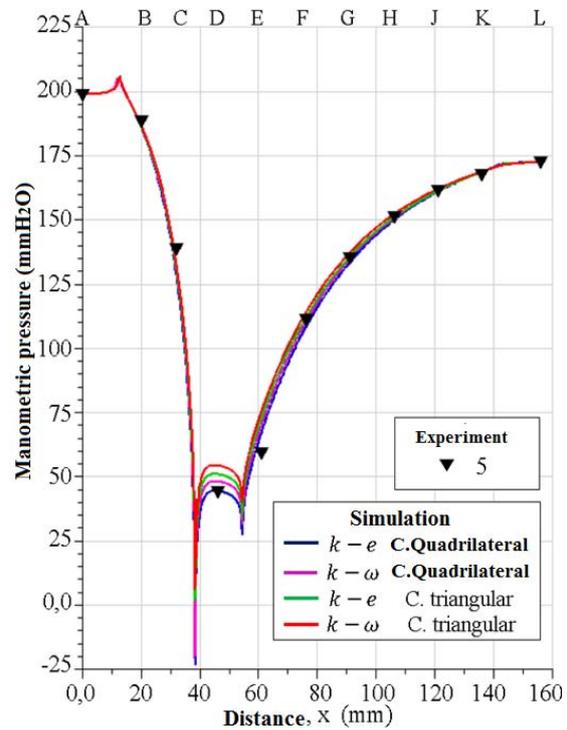
**Figure 13.** Experiment 2 of pressures of water columns, and pressure profiles evaluated at the wall of the Venturi tube. Pressure in (mmH<sub>2</sub>O): 170 mm at the inlet and 155.5 mm at the outlet.



**Figure 15.** Experiment 4 of pressures of water columns, and pressure profiles evaluated at the wall of the Venturi tube. Pressure in (mmH<sub>2</sub>O): 190 mm at the inlet and 167 mm at the outlet.



**Figure 14.** Experiment 3 of pressures of water columns, and pressure profiles evaluated at the wall of the Venturi tube. Pressure in (mmH<sub>2</sub>O): 179 mm at the inlet and 161 mm at the outlet.



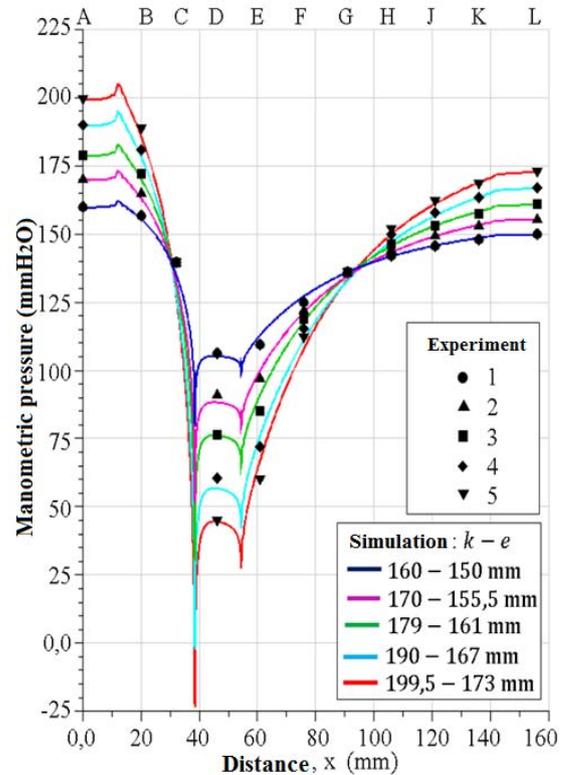
**Figure 16.** Experiment 5 of pressures of water columns, and pressure profiles evaluated at the wall of the Venturi tube. Pressure in (mmH<sub>2</sub>O): 199.5 mm at the inlet and 173 mm at the outlet.

The numerical results and the experimental data in the reference D are shown in Table 11. The standard  $k-e$  turbulence model showed the greatest percentage error of 5.88 % for experiment 4, for the standard  $k-\omega$  it was 7.84 % for experiment 5, and values smaller than these for the remaining results, i.e. for the mesh with quadrilateral cells. On the other hand, for the mesh with triangular cells, corresponding to experiment 5, the standard  $k-e$  had a percentage error of 14.17 %, for the standard  $k-\omega$  it was 21.66 %, and smaller for the remaining results.

**Table 11.** Experimental and numerical data for the reference D, for two turbulence models and two types of applied meshes

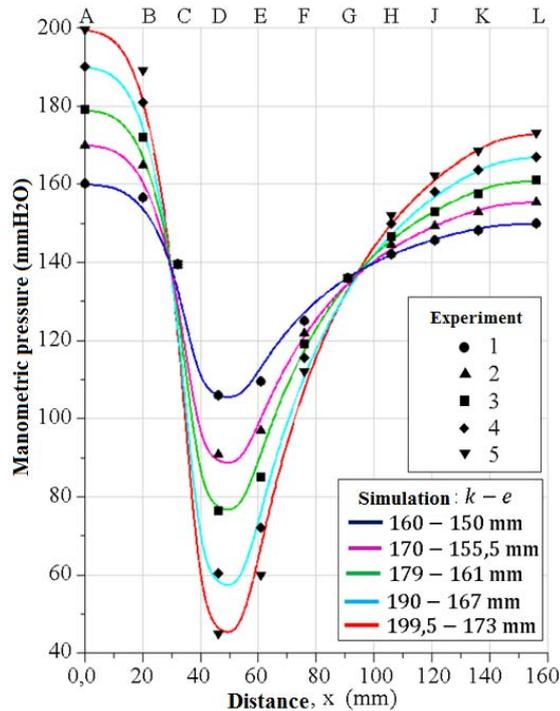
Data Exp. Ref. D mmH <sub>2</sub> O	Simulation: Reference D (mmH <sub>2</sub> O)				
	Cell Quadrilateral		Cell Triangular		
	$k-e$	$k-\omega$	$k-e$	$k-\omega$	
1	106	105,47	106,13	106,99	107,97
2	91	88,51	89,66	91,28	92,86
3	76,5	76,4	78,05	80,22	82,26
4	60,5	56,94	59,57	62,32	65,11
5	45	44,79	48,53	51,38	54,75

The pressure profiles for the standard  $k-e$  turbulence model, simulating the flow with the domain meshed with quadrilateral cells, and the experimental data of pressure presented in Table 5 and plotted in Figure 6, are presented in Figure 17, where it is observed that the trajectories of the profiles satisfy the validation with the experimental data. In the reference D located at the throat section, the trajectories of the profiles are convex. Although a graph similar to Figure 17 is not presented for the standard  $k-\omega$  turbulence model, this model also shows validity but with slightly varied margins of numerical results with respect to the standard  $k-e$  turbulence model, as shown in the results of Table 12 previously presented. From the analysis that was carried out, the numerical results are influenced by the type of mesh applied to the computational domain, and the structured mesh with quadrilateral cells provides more accurate numerical results compared to the mesh with triangular cells.



**Figure 17.** Experimental data of pressure of water columns and pressure profiles evaluated at the walls of the Venturi tube with the standard  $k-e$  turbulence model

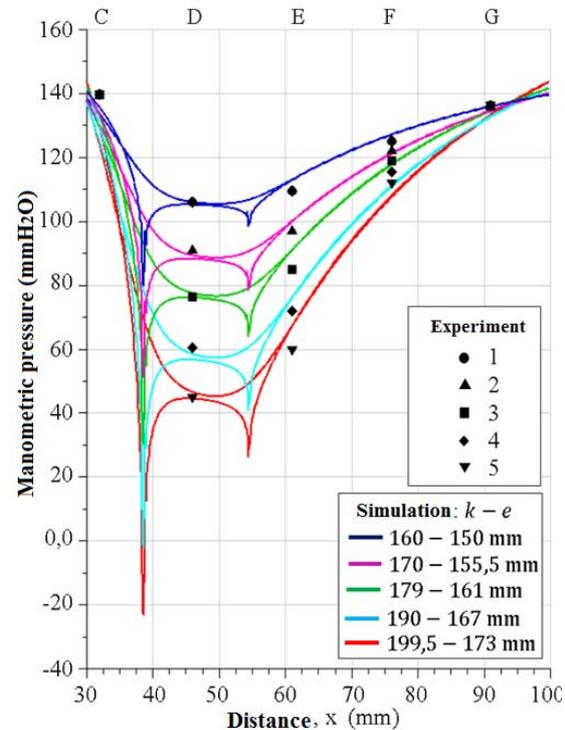
The pressures of the flow along the axial symmetry of the  $x$  axis, compared with the experimental data of the pressures on the walls of the Venturi tube, are shown in Figure 18. The profiles intercept with one another in the references C and G, and it is observed the evolution of the pressure drop trajectories in the left end of the reference D and of the pressure increase in the right end of the same reference; the trajectories of the curves are concave along the throat section. The small separations of the profiles of the experimental data are also shown, thus the magnitudes of the pressures in each of the references are some slightly greater and others slightly smaller with respect to the experimental data of the pressures on the walls. The profiles were obtained for the domain meshed with quadrilateral cells and the standard  $k-e$  turbulence model.



**Figure 18.** Experimental data of pressure of water columns and pressure profiles evaluated at the axial symmetry (x axis) of the Venturi tube with the standard  $k - e$  turbulence model.

When the numerical results of Figure 18 are compared with Figure 17, it is evident that for the locations of the references A, B, C, D, E, F, G, H, J, K and L, the pressures tend to be perpendicular to the x axis and to the walls, forming a trajectory of curves known as isobars. However, in the places where the sections come together, in the vertices, the configuration of the trajectories of the curves has a different behavior due to the sharp variations of pressures, induced by the geometrical profile of the section of gradual contraction and gradual expansion of the Venturi tube.

Figure 19, as a detail, unifies Figures 17 and 18, for the distance range 30-100 mm, showing superimposed profiles stretches for the pressures on the walls and on the x axis, which are compared with the experimental data for the references C, D, E, F and G. It is observed the pressure drops at the ends of the throat and how the curves intercept and border the experimental data. It is shown that the abrupt drops of the numerical pressures on the wall occur for the 38.67 mm position, at the beginning of the throat; and the other pressure drops occurs at the 54.49 mm position, which is located at the beginning of the divergent section.



**Figure 19.** Experimental data for the references C, D, E, F and G, and stretches of pressure profiles evaluated at the walls and at the x axis of the Venturi tube with the standard  $k - e$  turbulence model

Table 12 shows the numerical values of the pressure drops at the wall at the ends of the throat section. It is shown that the smallest drop of numerical pressure occurs at the 38.67 mm position, for the curve 199.5-173 mm which corresponds to experiment 5, being the magnitude of the pressure drop  $-23.23$  (mmH<sub>2</sub>O), and since it is a negative pressure, it is evident that it is a suction pressure; similarly, for the curve 190-167 mm corresponding to experiment 4 and for the same position, the pressure was  $-2.08$  (mmH<sub>2</sub>O).

The vertex, where the negative pressure occurs, corresponds to a very small part of an estimated radius of action of 0.2 mm, where the negative pressure is an unexpected result, because the hydrodynamic profile of the internal wall of the Venturi tube that measures the flow rate, the convergent section has an average angle of design with the purpose of avoiding negative pressures.

**Table 12.** Pressure drops at the ends of the throat

Numerical profile $k - e$	Position 38,67 mm $mmH_2O$	Position 54,49 mm $mmH_2O$
160-150 mm	79,75	98,94
170-155,5 mm	51,23	79,08
179-161 mm	30,17	64,73
190-167 mm	-2,08	42,07
199,5-173 mm	-23,23	27,63

Similarly, it is remarked that during experiments 4 and 5 no air bubbles were observed in the region around the vertex located at the inlet of the throat section and downstream, as a sign of cavitation. Therefore, the numerical result of the negative pressure, induces to investigate with sensitive instruments to capture the possible air bubbles with dimensions imperceptible to the human eye that might be present. Therefore, it should be verified simulating the flow with different turbulence models in a future work, to determine if negative pressures appear or not, thus obtaining conclusions close to the reality of the physical phenomenon.

From a comparison of the numerical results with the experimental data, it is evident that the simulation yields satisfactory results sustained by the ranges of error which are acceptable according to engineering criteria; thus the standard  $k - e$  and standard  $k - \omega$  turbulence models are validated. These two validated turbulence models strengthen their application in the computational fluid dynamics in the simulation of the flow in computational domains with simple or complex geometries in the field of engineering, and allow to determine the magnitude of some physical parameter that cannot be obtained through measuring instruments and analytical equations.

#### 4. Conclutions

Based on the analyses that were carried out, for the cases of experimental and numerical study, it is concluded that:

The obtained numerical flow rates for the standard  $k - e$  and standard  $k - \omega$  turbulence models, when compared to the five experimental data of flow rates, yielded percentage errors in the range 1.01-9.68 %. Similarly, it was determined the percentage error in the range 1.01-9.68 % for the Reynolds number. For the five experiments, the Reynolds number is in the range  $12000 < Re < 32300$ ; and for the numerical simulations, in the range  $12100 < Re < 34400$ .

The five experimental results of the pressures obtained in the references C and G, where the diameters of the Venturi tube are different, show that the magnitude of the pressures in that place do not vary, even

though the velocity of the flow increases in such references due to the increase in the flow rate. Besides, the smallest pressure drop occurs in the middle part of the throat section because of the Venturi effect.

Regarding the profiles obtained with the standard  $k - e$  and standard  $k - \omega$  turbulence models, which were compared with the experimental data of pressure, the standard  $k - e$  turbulence model with the domain meshed with quadrilateral cells gave the more accurate result, where the numerical results fitted better the experimental data in the middle part of the throat section, which is a critical section due to the abrupt drop in the fluid pressure. This evidences that the type of mesh influences the numerical results.

At the ends of the throat section, at the 38.67 mm and 54.49 mm positions, the pressure drops were more abrupt than in the middle part of the throat, appearing negative pressures at the 38.67 mm position for the numerical curves corresponding to experiments 4 and 5.

The smallest pressure drop at the wall does not occur in the middle part of the throat section, but at the ends of it, for the Venturi tube under study.

In future works, for experimental studies of the same type of Venturi tube, it is recommended to manipulate the control valve, and with the obtained experimental results, carry out the corresponding comparisons with the experimental results of the present work. Similarly, for cases of numerical studies, it is recommended to utilize other codes of CFD to simulate the flow of water employing other turbulence models, and compare with the numerical results of the present work.

#### Acknowledgements

My gratitude to Jehovah, my almighty God, my source of wisdom and inspiration. To the Mechanical Engineering Department of the National Experimental Polytechnic University “AJS”, Puerto Ordaz Vice-Rectorate, Bolívar, Venezuela. To the Group of Mathematical Modeling and Numerical Simulation (GMMNS) of the National University of Engineering (NUE), Lima, Peru.

#### References

- [1] F. M. White, *Mecánica de fluidos*, M.-H. I. de España S.L., Ed., 2004. [Online]. Available: <https://bit.ly/2MYFv6r>
- [2] Y. A. Çengel and J. M. Cimbala, *Mecánica de fluidos, fundamentos y aplicaciones*, M.-H. I. de España S.L., Ed., 2006. [Online]. Available: <https://bit.ly/2BUeJWD>
- [3] H. S. Bean, “Fluid meters their theory and application- sixth edition,” The American Society

- of Mechanical Engineers (ASME), Tech. Rep., 1971. [Online]. Available: <https://bit.ly/33c2OiM>
- [4] “Invention of the venturi mete,” *Nature*, no. 136, p. 254, 1935. [Online]. Available: <https://doi.org/10.1038/136254a0>
- [5] O. Reynolds, “An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels author(s): Osborne reynolds,” *Philosophical Transactions of the Royal Society of London*, vol. 174, pp. 935–982, 1883. [Online]. Available: <https://bit.ly/3211iyT>
- [6] N. Rott, “Note on the history of the reynolds number,” *Annual Review of Fluid Mechanics*, vol. 22, no. 1, pp. 1–12, 1990. [Online]. Available: <https://doi.org/10.1146/annurev.fl.22.010190.000245>
- [7] J. H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*, 3rd ed., S.-V. B. H. N. York, Ed., 2002. [Online]. Available: <https://doi.org/10.1007/978-3-642-56026-2>
- [8] D. C. Wilcox, *Turbulence modeling for CFD*, I. DCW Industries, Ed., 1993. [Online]. Available: <https://bit.ly/2Pxxkdf>
- [9] T. V. Karman, “The fundamentals of the statistical theory of turbulence,” *Journal of the Aeronautical Sciences*, vol. 4, no. 4, pp. 131–138, 1937. [Online]. Available: <https://doi.org/10.2514/8.350>
- [10] H. Schlichting and K. Gersten, *Boundary-layer theory*, S.-V. B. Heidelberg, Ed., 2017. [Online]. Available: <https://doi.org/10.1007/978-3-662-52919-5>
- [11] D. Lindley, “An experimental investigation of the flow in a classical venturimeter,” *Proceedings of the Institution of Mechanical Engineers*, vol. 184, no. 1, pp. 133–160, 1969. [Online]. Available: [https://doi.org/10.1243/PIME\\_PROC\\_1969\\_184\\_015\\_02](https://doi.org/10.1243/PIME_PROC_1969_184_015_02)
- [12] J. A. Sattery and M. J. Reader-Harris, “Computation of flow through venturi meters,” *North Sea Flow Measurement Workshop*, 1997. [Online]. Available: <https://bit.ly/2q2oZcT>
- [13] N. Tamhankar, A. Pandhare, A. Joglekar, and V. Bansode, “Experimental and cfd analysis of flow through venturimeter to determine the coefficient of discharge,” *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, vol. 3, no. 4, pp. 194–200, 2014. [Online]. Available: <https://bit.ly/2Nmabh6>
- [14] B. E. Launder and D. B. Spalding, *Lectures in mathematical models of turbulence*, A. P. London, New York, Ed., 1972. [Online]. Available: <https://bit.ly/2qWgZud>
- [15] D. C. Wilcox, “Reassessment of the scale-determining equation for advanced turbulence models,” *AIAA Journal*, vol. 26, no. 11, pp. 1299–1310, 1988. [Online]. Available: <https://doi.org/10.2514/3.10041>
- [16] G. Alfonsi, “Reynolds-averaged navier-stokes equations for turbulence modeling,” *Applied Mechanics Reviews*, vol. 62, no. 4, p. 040802, 2009. [Online]. Available: <https://doi.org/10.1115/1.3124648>



# ANALYSIS OF BEHAVIOR OF CO<sub>2</sub> EMISSIONS, CO AND THE LAMBDA FACTOR OF A VEHICLE WITH A CONVENTIONAL INJECTION SYSTEM WITH CATALYST AND WITHOUT CATALYST

## ANÁLISIS DEL COMPORTAMIENTO DE LAS EMISIONES DE CO<sub>2</sub>, CO Y DEL FACTOR LAMBDA DE UN VEHÍCULO CON SISTEMA DE INYECCIÓN CONVENCIONAL CON CATALIZADOR Y SIN CATALIZADOR

Edgar Vicente Rojas Reinoso<sup>1,\*</sup>, Vicente Javier Romero Hidalgo<sup>2</sup>,  
Johnny Marcelo Pancha Ramos<sup>2</sup>

### Abstract

The analysis of the behavior of the CO and CO<sub>2</sub> emissions is used to determine the performance of the engine work cycle, and in addition to verify the graph of the lambda factor, for which a study of the importance of a catalyst is carried out because sometimes the owners of the vehicles decide to eliminate the catalytic converter from the outlet lines of the engine burnt gases, and circulate along the roads of Ecuador ignoring the impact on the health of citizens and direct pollution towards the environment. With the analysis of the operation and characteristics of the pollutant emissions of an Otto cycle internal combustion engine, control models are generated for the projection of the amount of pollutant gases that are emitted when eliminating the catalytic converter and thereby establishing the levels of emission generated by a vehicle without a catalyst, despite the fact that the engine is in optimal operating conditions at different running speeds.

**Keywords:** catalyst, environmental mitigation, Otto cycle, Pollution.

### Resumen

El análisis del comportamiento de emisiones de CO, CO<sub>2</sub> sirve para determinar el comportamiento del ciclo de trabajo del motor, además de la verificación de la gráfica del factor lambda, para lo cual se realiza el estudio de la importancia de un catalizador porque en algunas ocasiones los propietarios de los vehículos deciden eliminar el convertidor catalítico de la línea de salida de los gases combustionados del motor y así circulan por las vías del Ecuador desconociendo la afectación hacia la salud de los ciudadanos y la contaminación directa hacia el medioambiente. Con el análisis del funcionamiento y características de las emisiones contaminantes de un motor de combustión interna ciclo Otto se generan modelos de control para la proyección de la cantidad de gases contaminantes que se emiten al eliminar el convertidor catalítico y de tal forma establecer los niveles de emisiones que un vehículo sin catalizador genera, a pesar de que el motor se encuentre en óptimas condiciones de funcionamiento en diferentes regímenes de giro.

**Palabras clave:** catalizador, ciclo Otto, gases de combustión, mitigación ambiental.

<sup>1,\*</sup>Researcher, Automotive engineering, Universidad Politécnica Salesiana, Ecuador.

Corresponding author ✉: [erojas@ups.edu.ec](mailto:erojas@ups.edu.ec), <http://orcid.org/0000-0001-5658-3055>

<sup>2</sup>Researcher, Automotive engineering, Escuela Superior Politecnica de Chimborazo, Ecuador.

<http://orcid.org/0000-0003-2317-7071>, <http://orcid.org/0000-0001-7320-2154>

Received: 30-05-2019, accepted after review: 31-10-2019

Suggested citation: Rojas Reinoso, E. V.; Romero Hidalgo, V. J. and Pancha Ramos, J. M. (2020). «Analysis of behavior of CO<sub>2</sub> emissions, CO and the lambda factor of a vehicle with a conventional injection system with catalyst and without catalyst». INGENIUS. N.º 23, (january-june). pp. 23-29. DOI: <https://doi.org/10.17163/ings.n23.2020.02>.

## 1. Introduction

At present, the polluting emissions of a vehicle are under constant analysis and study, in the search for more efficient engines and with low level of such emissions. For several years, vehicles have been considered an important source of polluting emissions to the environment due to the use of internal combustion engines. During the operation cycle of the engine and accomplishing an ideal combustion, molecular nitrogen ( $N_2$ ), water ( $H_2O$ ) and carbon dioxide ( $CO_2$ ) would be obtained. Nevertheless, as a result of the operating cycles of a thermal engine the combustion is not perfect, thus generating additional elements such as volatile organic compounds (VOC), carbon monoxide (CO), sulphur oxides, black smokes, lead compound and nitrous oxides (NO and  $NO_2$ ) [1]. Among some of the strategies to reduce the level of polluting gases it can be mentioned the use of catalytic converters at the outlet of the burnt gases of the engine, through chemical reactions and influenced by conditions such as the temperature, pressure and the application of materials that interact with the exhaust gases [2]. The use and application of these catalytic converters have been developed by means of studies conducted by each of the manufacturers, in order for their vehicles to be more environmental friendly; thus, it is necessary to keep them installed.

The importance of the study has been defined through the analysis of the polluting gases when the catalytic converter is eliminated; it is considered the initial values of the exhaust gases with the converter installed and the values obtained after it is eliminated, to define a mathematical model to forecast the importance of not suspending or uninstalling a catalytic converter from the exhaust line.

### 1.1. Polluting emissions

The polluting gases that originate in the vehicles act as irritants in the airways, harm the tissues altering their permeability, making them more vulnerable to develop illnesses, and that it possibly may appear viral and bacterial infection.

The nitrogen monoxide (NO) is formed due to the reaction of nitrogen and oxygen, at high temperatures in the combustion chamber. The nitrogen dioxide ( $NO_2$ ) is a reddish and irritant gas, which when inhaled adheres to the nasal mucosa forming nitric acid. The generation of this acid causes an immediate reaction: the irritation of the airways jointly with an eye discomfort; the lungs are affected causing respiratory problems and bronchopulmonary reactions.

The carbon monoxide (CO) is generated by the incomplete combustion of the fuel due to the presence of low levels of oxygen; it should be considered that the carbon monoxide increases with the variation of

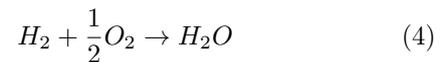
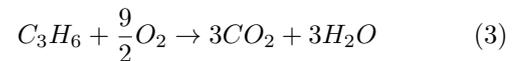
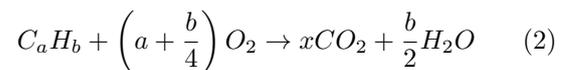
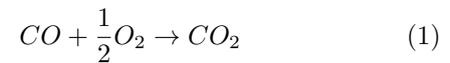
the air-fuel ratio during the mixing. The unburned hydrocarbons (UH) produce irritation in the eyes and directly affect the mucous in the respiratory tract, besides, they can cause a narcotic effect and are carcinogenic compounds. The hydrocarbons cause the presence of acid rains, and jointly with the ultraviolet rays produce the photochemical smoke [3].

#### 1.1.1. Catalytic converter

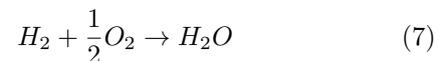
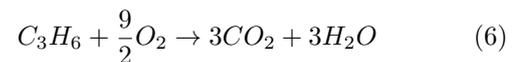
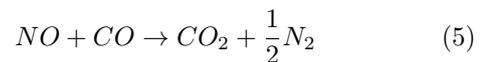
A solution of precious metals, used with different alloys ( $Al_2O_3$ ) is integrated to the catalytic converter, which is installed at the outlet of the exhaust gases [4].

Other elements that complement the structure of a catalytic converter are the platinum (Pt), rhodium (Rh) and the palladium (Pd); these catalytic materials integrate various types of catalysts; for example, the use of the platinum and palladium form the two-ways catalytic converter, also known as oxidation catalyst, while these elements together with the rhodium are used in the three-ways catalysts or reduction and oxidation catalysts [4].

Hereunder, the reduction chemical processes in a catalytic converter are shown.



The reactions generated in an oxidation catalytic converter are:



For a better performance of the catalysts it is necessary that the air-gasoline mixture is dosed to the engine; in other words, that it possesses a proportionate composition of one kilogram of gasoline per 14.7 kilograms of air. The element that registers the composition of the mixture is a monitoring device called lambda probe [5]. This device verifies and allows that a control unit carries out continuous adjustments on the air and fuel mixture, taking as reference the percentage of oxygen that exists in the burnt gases that exit through the exhaust pipe, to inform the unit that manages the injection of the engine about the amount of fuel; this characteristic is called as lambda factor, and

on this will depend the operation of the catalyst [5]. This is why in their design some catalytic converters use some type of material to reduce the levels of oxygen [6]. The elements usually applied such as the cerium (Ce) and the zirconium (Zr) store the oxygen and then release it according to the operating conditions when its presence decreases in the combustion gases [6].

## 2. Materials and methods

For the present research it has been chosen an experimentation strategy based on a Deming cycle, which is extended to the planning, implementation, verification and actuation founded on a spiral toward the continuous improvement [7].

The established operating parameters of an Otto cycle combustion engine (which uses a two-way catalyst in the exhaust line) are compared with the emission values of the same engine under the same characteristics, but eliminating the catalytic converter from the outlet line of the exhaust gases.

### 2.1. Experimental unit

For carrying out this study, a Sedan vehicle with a FS-ZM engine and a two-way catalyst in the exhaust line is utilized. Table 1 shows the characteristics of this engine.

**Table 1.** Characteristics of the FS-ZM with two-way catalyst

<b>Engine</b>	FS-ZM
<b>Cylindrical</b>	1600 cm <sup>3</sup>
<b>Maximum power</b>	97 kW
<b>Torque</b>	120 Nm
<b>Number of cylinders</b>	4
<b>Compression ratio</b>	9:01
<b>Fuel system</b>	Multipoint injection
<b>Type of catalyst</b>	Oxidation – two-ways

A gas analyzer, brand MAHA, model Met 6.3, was utilized for obtaining the values of emission of polluting gases. Table 2 shows the characteristics of the measuring equipment.

**Table 2.** Characteristics of the gas analyzer MAHA Met 6.3

Measurable Gases	HC, CO, CO <sub>2</sub> , O <sub>2</sub>
Measuring principle infrared spectrometry	HC, CO, CO <sub>2</sub>
Measuring principle electrochemical detection	O <sub>2</sub>
Flow index	3.5 l/min
Precision class	O (OIML)
CO - Measuring range/ Measuring accuracy (max.)	-15 % Vol. / 0,01
CO <sub>2</sub> - Measuring range/ Measuring accuracy (max.)	-20 % Vol. / 0,01
HC - Measuring range/ Measuring accuracy (max.)	-9999 ppm / 0,1
O <sub>2</sub> - Measuring range/ Measuring accuracy (max.)	-25 % Vol. / 0,01
Lambda (calculated)	0,5 - 9,99 / 0,01
Measuring principle	Extinktionsmessung
Measuring range concentration of particles	-1100 mg/m <sup>3</sup>
Resolution concentration of particles	1 mg/m <sup>3</sup>
Measuring interval opacity	-100%
Measuring area absorption coefficient	-9.99 m <sup>-1</sup>
Resolution absorption coefficient	0.01 m <sup>-1</sup>

### 2.2. Experimental design

For the development of this work, it has been applied an experimental design based on obtaining a mathematical model that forecasts the difference that exists after eliminating the catalytic converter from the outlet line of the exhaust gases, considering as output variables of the study the values of emission of exhaust gases in a four- cylinder engine whose design uses an oxidation catalyst [8,9].

### 2.3. Response variables

The response variables have been selected based on other research studies that have been conducted; taking into account the determination of the concentration of exhaust emissions at tick over conditions or Idle Static Test, 2000 [7,10,11]. Table 3 shows the response variables.

**Table 3.** Response variables

Variable	Symbol	Unit
Carbon monoxide	CO	%
Hydrocarbons	HC	ppm
Carbon dioxide	CO <sub>2</sub>	%
Lambda	λ	-
Oxygen	O <sub>2</sub>	%

## 2.4. Running speed

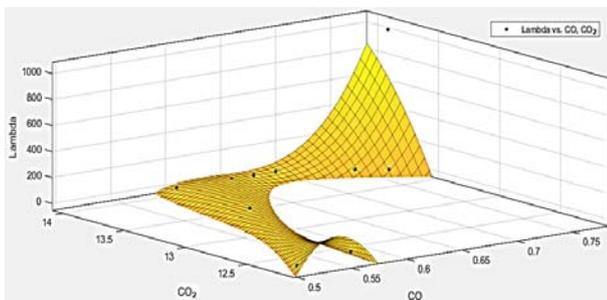
For the present study it has been considered as running speed what is established in the INEN standard, Environmental Administration. Air. Motor Vehicles. Determination of the Concentration of Exhaust Emissions at Tick Over Conditions or Idle Static Test, 2000, and the INEN, Vehicular Technical Revision. Procedures, 2003; procedures used in other similar emissions studies [7]. The first running speed considered was the idle condition at 700 rpm, and the second was at 2500 rpm [7].

## 3. Results and discussion

The emissions control equations define the prediction model of the emissions data, but the more relevant datum to stabilize such model is the comparison with the lambda factor, thereby finding the smallest amount of polluting emissions of CO<sub>2</sub>, CO and HC.

The catalyst is the control element for reducing the pollutants; for the analysis they are of two-ways, in this case for the same model of vehicle, which will enable obtaining the control data for vehicles that have the same style or type of catalyst. The tests were carried out at normal environmental conditions at a height of 2850 m.a.s.l., and instantaneously to ensure the truthfulness and inherency of the data.

The control prediction model is represented in Equation 1 that adapts to the behavior of the control surface. Figure 1 indicates the behavior of the data dispersion, which demonstrates that at low operating conditions (idle) the vehicle generates non-stable alteration peaks (without catalyst), which produce that the lambda soars, with which the CO<sub>2</sub> and CO progressively increase and the emissions destabilize, as can be checked in Equation 2 of Figure 2.

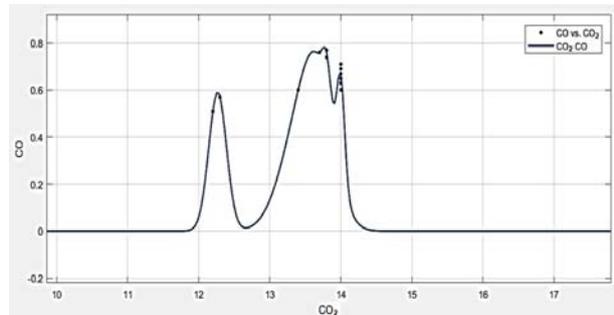


**Figure 1.** Surface of the behavior of the lambda vs. CO<sub>2</sub> vs. CO at idle condition without catalyst.

$$f(x, y) = p00 + p10x + p01y + p20x^2 + p11xy + p02y^2 + p30x^3 + p21x^2y + p12xy^2$$

Coefficients:

$$\begin{aligned} p00 &= -127.4(-891.9, 637) \\ p10 &= -164.1(-671.2, 343) \\ p01 &= 321.7(-343.1, 986.4) \\ p20 &= 114.3(-497.1, 725.7) \\ p11 &= 386.1(-731.3, 1504) \\ p02 &= -185.9(-2177, 1805) \\ p30 &= 130.4(-213.5, 474.2) \\ p21 &= 84.29(-912.1, 1081) \\ p12 &= -108.7(-1627, 1409) \end{aligned}$$



**Figure 2.** Behavior of the CO and CO<sub>2</sub>

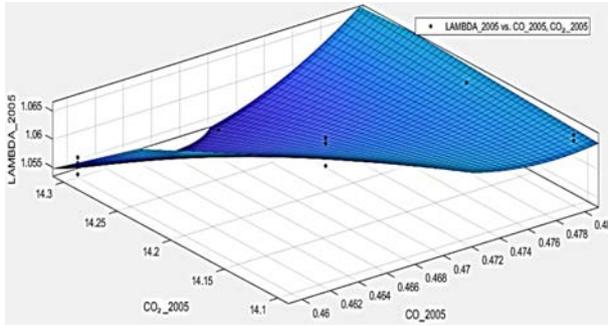
$$\begin{aligned} f(x) &= a1 \cdot \exp\left(-\left(\frac{x-b1}{c1}\right)^2\right) + \\ & a2 \cdot \exp\left(-\left(\frac{x-b2}{c2}\right)^2\right) + \\ & a3 \cdot \exp\left(-\left(\frac{x-b3}{c3}\right)^2\right) + \\ & a4 \cdot \exp\left(-\left(\frac{x-b4}{c4}\right)^2\right) + \\ & a5 \cdot \exp\left(-\left(\frac{x-b5}{c5}\right)^2\right) \end{aligned}$$

Coefficients:

$$\begin{aligned} a1 &= 0.1629(-1.753e^{+42}, 1.753e^{+42}) \\ b1 &= 13.8(-4.279e^{+41}, 4.279e^{+41}) \\ c1 &= 0.08264(-6.414e^{+41}, 6.414e^{+41}) \\ a2 &= 0.4192(-1.066e^{+41}, 1.066e^{+41}) \\ b2 &= 14(-8.979e^{+43}, 8.979e^{+43}) \\ c2 &= 0.07769(-1.604e^{+43}, 1.604e^{+43}) \\ a3 &= 0.3826(-1.532e^{+35}, 1.532e^{+35}) \\ b3 &= 13.41(-6.356e^{+34}, 6.356e^{+34}) \\ c3 &= 0.3917(-3.071e^{+34}, 3.071e^{+34}) \\ a4 &= 0.5899(-9.371e^{+28}, 9.371e^{+28}) \end{aligned}$$

$$\begin{aligned}
 b4 &= 12.27(-4.638e^{+30}, 4.638e^{+30}) \\
 c4 &= 0.1762(-9.03e^{+30}, 9.03e^{+30}) \\
 a5 &= 0.5032(-1.75e^{+34}, 1.75e^{+34}) \\
 b5 &= 13.7(-1.107e^{+35}, 1.107e^{+35}) \\
 c5 &= 0.3243(-9.532e^{+33}, 9.532e^{+33})
 \end{aligned}$$

The control prediction model is Equation 3 that adapts to the behavior of the control surface; and Figure 3 indicates the behavior of the data dispersion, which demonstrates that at low operating regime (idle) the vehicle generates a stable condition (with catalyst), which produces that the lambda generates an invariable sinusoidal curve, causing that the CO<sub>2</sub>, CO and the emissions stabilize as can be checked in Equation 4 of Figure 4; each value of CO<sub>2</sub> stabilizes with respect to a value of CO.

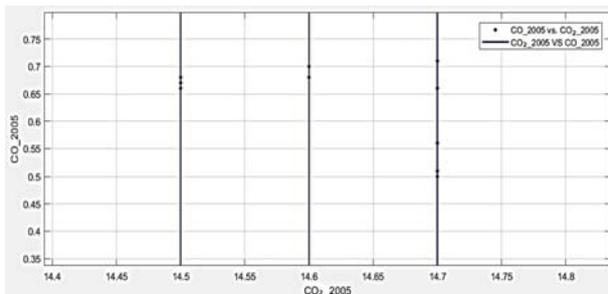


**Figure 3.** Surface of the behavior of the lambda vs. CO<sub>2</sub> vs. CO at idle condition with catalyst.

$$f(x, y) = p00 + p10x + p01y + p20x^2 + p11xy$$

Coefficients:

$$\begin{aligned}
 p00 &= 63.85(34.8, 92.9) \\
 p10 &= -161.1(-235.9, -86.2) \\
 p01 &= -3.437(-5.027, -1.846) \\
 p20 &= 62.83(32.05, 93.61) \\
 p11 &= 7.167(3.839, 10.49)
 \end{aligned}$$



**Figure 4.** Behavior of the CO and CO<sub>2</sub>

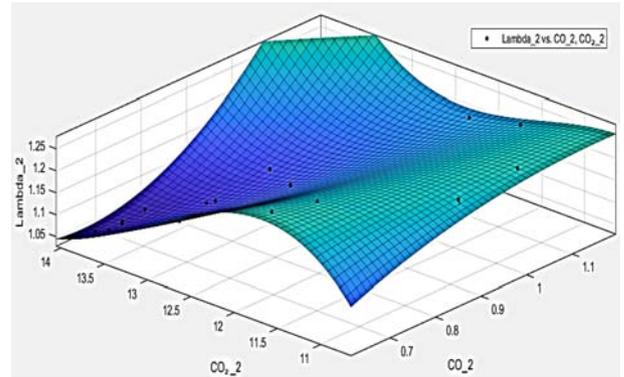
$$f(x) = p1x^3 + p2x^2 + p3x + p4$$

Coefficients:

$$\begin{aligned}
 p1 &= 1.257e^{+10}(-1.901e^{+09}, 2.705e^{+10}) \\
 p2 &= -5.508e^{+11}(-1.185e^{+12}, 8.326e^{+10}) \\
 p3 &= 8.041e^{+12}(-1.216e^{+12}, 1.73e^{+13}) \\
 p4 &= -3.913e^{+13}(-8.418e^{+13}, 5.916e^{+12})
 \end{aligned}$$

The stability of the emissions based on the use of the catalyst at low conditions is inherent in the behavior of the injection, moreover regarding fuel consumption, generating a greater amount of emissions. As a consequence, a similar analysis can be made for the high regime conditions at 4000 rpm which is the optimum range of operation with respect to the manufacturer of the model under study (Mazda allegro).

In high regime, the engine without catalyst generates the stability wave, but the values of CO and CO<sub>2</sub> are still very oscillating, which generates an excessively rich control lambda, as can be observed in Figure 5 of Equation 5.



**Figure 5.** Surface of the behavior of the lambda vs. CO<sub>2</sub> vs. CO at high regime (4500 rpm) without catalyst

$$\begin{aligned}
 f(x, y) &= p00 + p10x + p01y + p20x^2 + p11xy + \\
 &+ p02y^2 + p30x^3 + p21x^2y + p12xy^2 + p03y^3
 \end{aligned}$$

Coefficients:

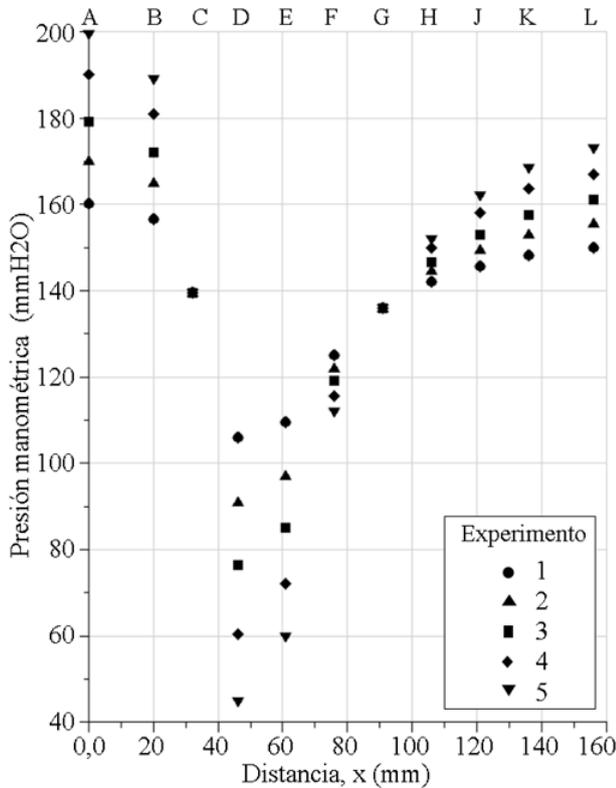
$$\begin{aligned}
 p00 &= -66.85(-111.5, -22.22) \\
 p10 &= 37.06(2.801, 71.32) \\
 p01 &= 14.25(5.461, 23.03) \\
 p20 &= -6.703(-17.69, 4.28) \\
 p11 &= -5.262(-9.398, -1.126) \\
 p02 &= -0.9794(-1.572, -0.3871) \\
 p30 &= 0.1117(-1.636, 1.86)
 \end{aligned}$$

$$p_{21} = 0.5634(-0.04962, 1.177)$$

$$p_{12} = 0.1778(0.05104, 0.3045)$$

$$p_{03} = 0.02223(0.008537, 0.03593)$$

In the engine with catalyst, it is noted that the sinusoidal wave is of greater period, generating a linearity in the emissions of CO and CO<sub>2</sub>, because the control lambda is more stable, as can be observed in Figure 6 of Equation 6, indicating that the injection system generates a constant stoichiometry in the work phases.



**Figure 6.** Surface of the behavior of the lambda vs. CO<sub>2</sub> vs. CO at high regime (4500 rpm) with catalyst

$$f(x, y) = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2 + p_{30}x^3 + p_{21}x^2y + p_{12}xy^2$$

Coefficients:

$$p_{00} = -653.5(-2064, 756.7)$$

$$p_{10} = 899.3(-979.2, 2778)$$

$$p_{01} = 92.64(-109.8, 295.1)$$

$$p_{20} = 99.1(-229.7, 427.9)$$

$$p_{11} = -132.2(-416.8, 152.4)$$

$$p_{02} = -3.273(-10.52, 3.974)$$

$$p_{30} = -0.3921(-1.442, 0.6578)$$

$$p_{21} = -6.694(-29.06, 15.67)$$

$$p_{12} = 4.832(-5.854, 15.52)$$

The fundamental difference in the surface of the control lambda is the stable sinusoidal and the linearity, indicating again that in a range of 4000 rpm the working velocity per cycle is larger and thus the emissions are much larger, as indicated in Figure 5, while in Figure 6 the surface is stable and thus the emissions are of lower level of variation.

## 4. Conclusions

The stability of the emissions based on the use of the catalyst at low conditions is inherent with the behavior of the injection system, moreover regarding fuel consumption, generating a greater amount of emissions.

At high regime, the engine without catalyst generates the stability wave, but the values of CO and CO<sub>2</sub> remain very oscillatory, which generates an excessively rich control lambda

In the engine with catalyst, it is noted that the sinusoidal wave is of a larger period, generating a linearity in the emissions of CO and CO<sub>2</sub>, because the control lambda is more stable.

In the analysis of the surface of the control lambda the sinusoidal and the linearity are stable, indicating again that in a range of 4000 rpm the working velocity per cycle is larger, and thus the emissions are much larger.

## References

- [1] I. Asensio, J. Rincón, R. Camarillo, and A. Martín, "Reciclado de catalizadores de automóviles. análisis de las técnicas actuales y propuestas de futuro," in *I Simposio Iberoamericano de Ingeniería de Residuos*, 2008. [Online]. Available: <https://bit.ly/33NJadl>
- [2] S. Bhattacharyya and R. K. Das, "Catalytic control of automotive nox: a review," *International Journal of Energy Research*, vol. 23, no. 4, pp. 351–369, 1999. [Online]. Available: [https://doi.org/10.1002/\(SICI\)1099-114X\(19990325\)23:4%3C351::AID-ER497%3E3.0.CO;2-T](https://doi.org/10.1002/(SICI)1099-114X(19990325)23:4%3C351::AID-ER497%3E3.0.CO;2-T)
- [3] G. D'Amato, G. Liccardi, and M. Cazzola, "Environment and development of respiratory allergy: I. outdoors," *Monaldi archives for chest disease = Archivio Monaldi per le malattie del torace*, vol. 49, no. 5, pp. 406–411, December 1994. [Online]. Available: <https://bit.ly/353T8aR>
- [4] K. C. R. Martins, F. Soto Pau, J. A. Silva, A. M. Santos, and R. F. dosSantos, "rtt," *Ingeniería Mecánica*, vol. 8, no. 1, pp. 39–46, 2005. [Online]. Available: <https://bit.ly/2NKOOHJ>

- [5] L. Martín, J. L. Arranz, O. Prieto, R. Trujillano, M. J. Holgado, M. A. Galán, and V. Rives, "Simulation three-way catalyst ageing: Analysis of two conventional catalyst," *Applied Catalysis B: Environmental*, vol. 44, no. 1, pp. 41–52, 2003. [Online]. Available: [https://doi.org/10.1016/S0926-3373\(03\)00008-0](https://doi.org/10.1016/S0926-3373(03)00008-0)
- [6] R. A. Daley, S. Y. Christou, A. M. Efstathiou, and J. A. Anderson, "Influence of oxychlorination treatments on the redox and oxygen storage and release properties of thermally aged pd-rh/cexzr1-xo2/al2o3 model three-way catalysts," *Applied Catalysis B: Environmental*, vol. 60, no. 1, pp. 117–127, 2005. [Online]. Available: <https://doi.org/10.1016/j.apcatb.2005.03.002>
- [7] N. Rivera, J. Chica, I. Zambrano, and C. García, "Estudio Del Comportamiento De Un Motor Ciclo Otto De Inyección Electrónica Respecto De La Estequiometría de la mezcla y del adelanto al encendido para la ciudad de Cuenca," *Revista Politécnica*, vol. 40, pp. 59–67, 10 2017. [Online]. Available: <https://bit.ly/376fV7F>
- [8] INEN, "Inen 2 203:2000. gestión ambiental. aire. vehículos automotores. determinación de la concentración de emisiones de escape en condiciones de marcha mínima o ralenti prueba estática," in *Servicio Ecuatoriano de Normalización*, 2000. [Online]. Available: <https://bit.ly/2CLVbnG>
- [9] D. D. Montgomery, *Design and Analysis of Experiments*, J. W. . Sons, Ed., 2017. [Online]. Available: <https://bit.ly/2qSdxkj>
- [10] INEN, "Inen 2 204:2002 gestión ambiental aire vehículos automotores límites permitidos de emisiones producidas por fuentes móviles terrestres de gasolina." in *Servicio Ecuatoriano de Normalización*, 2002. [Online]. Available: <https://bit.ly/375VPua>
- [11] —, "Inen 2 349:2003. revisión técnica vehicular. procedimientos," in *Servicio Ecuatoriano de Normalización*, 2003. [Online]. Available: <https://bit.ly/33LUpTC>



# PREDICTION OF CO AND HC EMISSIONS IN OTTO MOTORS THROUGH NEURAL NETWORKS

## PREDICCIÓN DE EMISIONES DE CO Y HC EN MOTORES OTTO MEDIANTE REDES NEURONALES

Rogelio Santiago León Japa<sup>1</sup>, José Luis Maldonado Ortega<sup>1</sup>,  
 Wilmer Rafael Contreras Urgilés<sup>1,\*</sup>

### Abstract

This paper explains the application of artificial neural networks (ANN) for the prediction of pollutant emissions generated by mechanical failures in ignition engines, from which the percentage of CO (% carbon monoxide) and the particulate in parts per millions of HC (ppm of unburned hydrocarbons) can be quantified, through the study of the Otto cycle intake phase, which is recorded through the physical implementation of a Manifold Absolute Pressure (MAP) sensor. A rigorous protocol of sampling and further statistical analysis is applied. The selection and reduction of attributes of the MAP sensor signal is made based on the greater contribution of information and significant difference with the application of three statistical methods (ANOVA, correlation matrix and Random Forest), from which a database that enables training two backpropagation feedforward neural networks, with which a classification error of  $5.4061e^{-09}$  and  $9.7587e^{-05}$  for CO and HC, respectively, can be obtained.

**Keywords:** prediction, pollutant emissions, carbon monoxide (CO), non-combustion hydrocarbons (HC), diagnostics, neural networks.

### Resumen

En el presente trabajo se explica la aplicación de RNA (redes neuronales artificiales) para la predicción de emisiones contaminantes generadas por fallas mecánicas en motores de encendido provocado, de la cual se puede cuantificar el porcentaje de CO (% monóxido de carbono) y el particulado por millón HC (ppm hidrocarburos sin quemar), a través del estudio de la fase de admisión del ciclo Otto, la cual es registrada por medio de la implementación física de un sensor MAP (*Manifold Absolute Pressure*). Se aplica un riguroso protocolo de muestreo y consecuente análisis estadístico. La selección y reducción de atributos de la señal del sensor MAP se realiza en función del mayor aporte de información y diferencia significativa con la aplicación de tres métodos estadísticos (ANOVA, matriz de correlación y Random Forest), de la cual se obtiene una base de datos que permite el entrenamiento de dos redes neuronales *feed-forward backpropagation*, con las cuales se obtiene un error de clasificación de  $5.4061e^{-9}$  y de  $9.7587e^{-5}$  para la red neuronal de CO y HC respectivamente.

**Palabras clave:** predicción, emisiones contaminantes, monóxido de carbono (CO), hidrocarburos no combustionados (HC), diagnóstico, redes neuronales artificiales.

<sup>1,\*</sup>Research Group of Transportation Engineering (GIIT), Automotive mechanical engineering, Universidad Politécnica Salesiana, Cuenca-Ecuador. Corresponding author ✉: [rcontreras@ups.edu.ec](mailto:rcontreras@ups.edu.ec)

<http://orcid.org/0000-0003-2142-3769>, <http://orcid.org/0000-0002-3846-2599>,

<http://orcid.org/0000-0003-2300-9457>

Received: 15-05-2019, accepted after review: 05-11-2019

Suggested citation: León Japa, R. S.; Maldonado Ortega, J. L. and Contreras Urgilés, W. R. (2020). «Prediction of CO and HC emissions in Otto motors through neural networks». INGENIUS. N.º 23, (january-june). pp. 30-39. DOI: <https://doi.org/10.17163/ings.n23.2020.04>.

## 1. Introduction

At present, the area of automotive transportation represents one of the main sources of air pollution; indeed, the discharge of pollutants to the environment has its origin in the accelerated population growth and the development of different urban centers, thus the deterioration of the quality of the air is due to mobile (vehicles), stationary (industry) and areal (domestic activities and services) sources.

The vehicular area (gasoline and diesel vehicles) is one of the main emitters of burnt fossil fuels to the environment, due to the pollutant gases generated during the operation of the automotive transportation, with the primary emissions being: carbon monoxide (CO), carbon dioxide (CO<sub>2</sub>), unburnt hydrocarbons (HC) and nitrous oxides (NO<sub>x</sub>), such that they affect public health and the equilibrium of the different ecosystems.

In the light of the previous paragraphs, it is necessary to develop new specialized and methodological techniques to obtain assertive diagnoses of mechanical failures; simultaneously, artificial neural networks and computational mathematics are used, due to the complexity of analyzing and interpreting the operational parameters of the ignition engine, and in this way determining the mechanical failures and the emissions that they produce, in short diagnosis times and optimizing resources. The air pollution has harmful effects on the health of all people, as demonstrated by the research study carried out by Ballester [1], who shows that in France, Switzerland and Austria 6 % of the mortality and an important number of new cases of respiratory diseases can be attributed to the air pollution, with half of this impact caused by the pollution emitted by motor vehicles.

The research by Restrepo *et al.* [2] estimates alarming contributions of polluting emissions that generate smog and contribute to the greenhouse effect in the city of Pereira. The study indicates the contribution to pollution of each vehicular category according to a software of the international model of emissions and an extrapolation, and results indicate that particular vehicles contribute more than 80 % of CO emissions, 60 % of CO<sub>2</sub>, 65 % of NO<sub>x</sub>, 40 % of SO<sub>x</sub>, and motorcycles contribute around 65 % of the particulate material (PM).

The use of neural networks is considered a technique of great contribution in the analysis of internal parameters of the ignition engines, as demonstrated by Li *et al.* [3] through the application of a neural network for predicting NO<sub>x</sub> emissions; the study utilizes intensity relations of flame radicals, together with flame temperature and NO<sub>x</sub> emissions, to train the neural network.

Cortina [4] proposes a model for predicting the concentration of the pollutants in the city of Salamanca (Mexico), with the most critical pollutants being SO<sub>2</sub>

and PM<sub>10</sub>; the model uses artificial neural networks (ANN) in combination with clustering algorithms, and the study uses particular meteorological variables as factors that influence the concentration of pollutants.

It is relevant to reduce the emissions of CO, HC and NO<sub>x</sub> of internal combustion engines with ignition start, since they produce different environmental problems, such as air pollution and global warming. Martínez *et al.* [5] used artificial neural networks (ANN) to predict the exhaust emissions of a 1.6 L ignition engine, with the purpose of optimizing such engine, by reducing the CO, HC and NO<sub>x</sub> emissions; the inputs of the ANN were six operating parameters of the engine, and the outputs were the three resulting exhaust emissions.

Similarly, Fontes *et al.* [6] apply multilayer perceptron (MLP) neural networks with a hidden layer as a classifier of the impact of the air quality on human health, using as only inputs traffic and meteorological data. Parallel and combined strategies can be used for determining the concentrations of emissions, for example, a hybrid learning of the artificial neural network (ANN) with the non-dominated sorting genetic algorithm – II (NSGAI) to improve the precision and predict the exhaust emissions of an ignition start gasoline four-stroke engine [7].

Different methods can be applied to analyze and predict emissions, such as the multivariate linear regression model to analyze the relationship between atmospheric pollutants and meteorological factors. Lopez and Pacheco [8] show that the source of benzene is the smoke from tobacco, gas stations, industrial emissions, and exhaust pipes of motor vehicles in the urban zone of the city of Cuenca (Ecuador), which has generated a raise in the number of clinical cases, such as asthma (36.34 %), bronchopneumonia (12.19 %), bronchiolitis (16.89 %), bronchitis (6.29 %), pharyngitis (12.41 %), pneumonia (11.73 %) and rhinitis (3.67 %), all this due to the increase of the concentration of benzene in one unit; on the other hand, the PM<sub>10</sub> exhibits a positive relationship with the venous thrombosis, causing an increment of 3.56 % of the clinical cases per each unit of increase in its concentration. Guadalupe [9] applies a new methodology for modeling the pollutant emissions from terrestrial moving sources in Quito (Ecuador), the international model of vehicular emissions (IMVE), which encompasses a bottom-up type of methodology that gathers a large amount of information to make up the inventory of emissions.

At present, the strategies for predicting the concentrations of gases are diverse. Leon and Piña [10] present a model for predicting emissions (NO<sub>x</sub>, CO, CO<sub>2</sub>, HC and O<sub>2</sub>) applied to gasoline powered vehicles, with the use of artificial neural networks (ANN); the input variables to the ANN are the mean effective pressure (MEP), RPM, load and MAP, and the model also predicts the load of the engine. Similarly, Contreras *et al.* [11] have proposed a diagnosis system that

can detect mechanical failures in Otto cycle engines with ignition start, by means of artificial neural networks (ANN); the system is based on the use of the signals from the MAP and CMP sensors, and has a classification error of  $1.89e^{-11}$ .

The prediction system proposed in this work can determine pollutant emissions and mechanical failures caused by particular emissions, such as carbon monoxide (CO) and unburnt hydrocarbons (HC); the Engine Control Unit (ECU) does not carry out the diagnosis performed by the prediction system. The system is based on the pressure of the intake manifold, which is recorded through the physical implementation of a MAP sensor, thus the system has the capability of reducing to a minimum the diagnosis time; in addition, the system does not utilize variables related to the quality of the air of the city neither meteorological variables for training the ANN, thus constituting a significant advance for predicting the emissions of exhaust gases and determining mechanical failures; activating this system in centers of automobile service and vehicular technical revision (VTR) is reliable and accessible.

## 2. Methods and materials

The main topics are developed in this section. These include experimental configuration and minimally invasive instrumentation, conditions for acquiring the samples, methodology for data acquisition, obtaining the matrix for the analysis and reduction of attributes, selection of the attributes for training the ANN, and neural network algorithm in MATLAB for diagnosing and predicting emissions.

### 2.1. Experimental configuration and minimally invasive instrumentation

The main consideration of the study is to avoid disassembling the elements and systems of the engine of a vehicle to diagnose the failures and predict the polluting emissions; consequently, the depression of the engine is measured through the installation of a MAP sensor in a vacuum connection of the intake manifold, placing it after the throttle valve such that the connection does not affect the operation of the ignition engine.

Table 1 summarizes the characteristics of the engine under test, and Table 2 includes the applied instrumentation.

**Table 1.** Characteristics of the experimental unit

Characteristic	Value
Model	Hyundai
Number of cylinders	4
Valve train	DOHC
Injection system	MFI
Cylinder	2000 cc
Power	175 CV @ 6000 rpm
Torque	168 N.m @ 4000 rpm
Fuel	Gasoline (RON 95)
Compression index	10.5: 1

**Table 2.** Applied instrumentation

Characteristic	Value
Personal computer	PC
Automotive scanner	Multiscan (Hanatech) OBDI, OBDII, EOBD y CAN.
Connection	Motor, Transmisión, ABS, Airbag. Ni DAQ-6009
Diagnosis	8 AI de 14 bits, a 48 KS/s 2 AO de 12 bits
Data acquisition card	
Inputs	
Outputs	
Intake pressure	MAP
Pressure range	20-117 KPa
Voltage	3.2V @ 1600 m.s.n.m Piezoelectric
Sensor type	
Gas analyzer for vehicles	QROTECH / QGA 6000
Measuring range CO	0.0 9.99 %
Resolution	0.01 %
Measuring range HC	0.0 9999 ppm
Resolution	1 ppm

The identification of each cylinder of the engine is carried out using the record of the signal of the camshaft position (CP) sensor.

Figure 1 shows the experimental unit under test, constituted by a Hyundai Sonata 2.0 DOHC engine, a gas analyzer, a personal computer (PC) and an automotive scanner. Figure 2 shows the connection of the MAP type sensor, the vacuum connection in the intake manifold and the data acquisition tool Ni DAQ-6009.



**Figure 1.** Instrumentation in the engine.



Figure 2. Connection of the MAP sensor.

**2.2. Conditions for acquisition of samples**

The installation of the intake pressure sensor is carried out, placing it in the intake manifold of the ignition engine, consequently acquiring samples of the emissions of NO<sub>x</sub>, CO, CO<sub>2</sub> and HC through the gas analyzer, and recording samples of the signal of the MAP sensor with a Ni DAQ-6009 card together with the software LabVIEW.

The samples of pressure and polluting emissions are acquired at idle condition, at approximately 850 RPM with an engine temperature range between 92 and 99 °C and engine load 35 %; an automotive scanner is used to confirm these conditions. Based on a pre-experimental study carried out in the research, it was determined that the signal of the MAP sensor has peaks of higher frequency, thus the sampling frequency is 10 KHz during a time period of 5 seconds for each of the signals; such frequency is larger than the one established according to Nyquist criterion (1.416 KHz) [11].

**2.3. Methodology for data acquisition**

Figure 3 presents the physical elements necessary for the corresponding diagnosis of the mechanical failures and prediction of polluting emissions.



Figure 3. Elements necessary for data acquisition.

The procedure represented in the flow diagram of Figure 4, is applied to obtain the signals of the MAP and CP sensors.

The process of obtaining the data starts with the revision of correct operation of the engine or supervised failure, and subsequently the connection of the sensors is inspected. If the connection is correct, the signal is saved with the software LabVIEW and registered in an Excel file, if not, the connection of the sensors

is verified [11]. The previously described procedure is applied to record the signals, both for the cases of engine in good operating condition (Figure 4(a)) and of engine with supervised failure (Figure 4(b)) [11]. The data acquisition process is performed 20 times for each of the engine conditions.

Table 3 indicates the total of six failures that are generated in the ignition engine experimental unit, each with the corresponding identification code; the condition of the engine in optimal operation is also indicated.

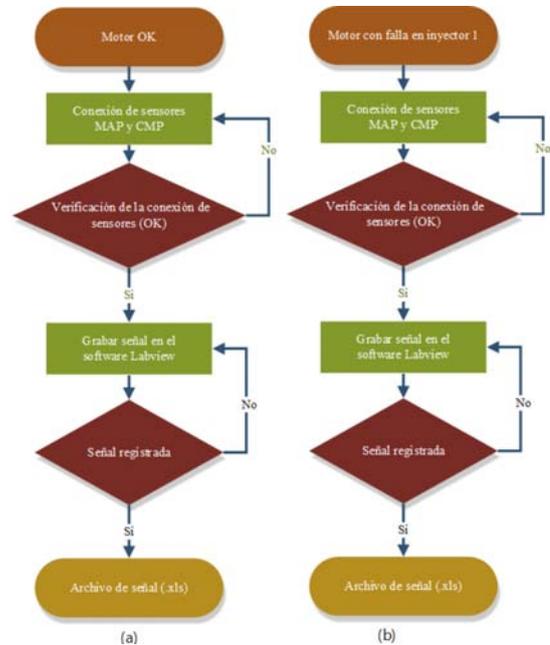


Figure 4. Flow diagram of the procedure for data acquisition (a) engine ok, (b) engine with failure.

Table 3. Operating conditions of the ignition engine experimental unit

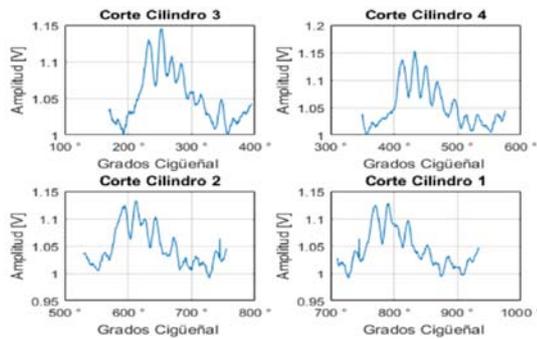
N.º	Type of mechanical condition	Identification code
1	Engine in optimal operating condition	100
2	Failure in Injector 1	200
3	Failure in Injector 2	300
4	Failure in Injector 3	400
5	Failure in Injector 4	500
6	Failure in Coil 1-4	1000
7	Failure in Coil 2-3	1100

**2.4. Obtaining the matrix of attributes analysis and reduction**

It is considered a complete segment of the signal of the MAP sensor, which corresponds to a cycle of the engine (720°±180°) taking into account the timing distribution of advance to intake opening (AIO) and delay

of intake closure (DIC), for each of the cylinders [11]. A windowing of the signal of the MAP sensor is carried out for each cylinder, as can be observed in Figure 5.

Once the time signals have been taken, an algorithm is developed in the software Matlab for reading and obtaining the general matrix with 18 attributes, namely: geometric mean, maximum, minimum, median, covariance, variance, standard deviation, mode, kurtosis factor, skewness coefficient, energy, power, area under the curve, entropy, coefficient of variation, range, mean square root and crest factor [11].



**Figure 5.** Windowing of the MAP sensor signal for each cylinder.

For the selection and reduction of the number of attributes, the general matrix is analyzed through 3 statistical methods: ANOVA, correlation matrix and Random Forest.

The application of the single factor ANOVA statistical method, allows to determine the best attributes that enter to the general matrix, through the analysis of the 18 attributes considering the greatest value of  $R^2$ , since values close to 100 % indicate that there is a correct fit of the model to the data, in other words, the variation between the attributes is determined. In addition, values of  $p$  close to 0 are considered, which determine if the attributes are statistically significant [11].

With respect to the correlation matrix, the attributes with coefficients close to -1 or 1 were discarded, since with them there is a strong relationship between the variables, negative or positive, respectively. Indeed, the attributes with coefficients close to zero were chosen, because with those attributes there is no strong correlation between the variables [11].

Regarding the Random Forest method, it allows to obtain the estimation of the importance of the attributes using the Curvature test, Standard CART and Interaction test methods. Afterwards, a Pareto analysis was applied to choose the attributes with greatest priority, considering only the top 95 % of the accumulated distribution [11].

## 2.5. Selection of attributes for training the ANN

In order to select the attributes that will be considered as inputs of the neural network, it was performed a match analysis of the general matrix from which the most often repeated attributes are chosen among the results of each statistical method applied [11]. The most often repeated attributes are shown in Table 4.

**Table 4.** Attributes utilized for training the artificial neural network

Statistical attributes	Number of repetitions
Área ( $v^2$ )	5
Energy (J)	5
Entropy (J)	4
Maximum (V)	4
Mean (V)	5
Minimum (V)	5
Power (mW)	5
RMS (V)	5

## 2.6. Matlab algorithm of the neural network for the diagnosis and prediction of emissions

Figure 6 presents the flow diagram of the procedure for creating the artificial neural networks corresponding to CO and HC.

The algorithm initiates reading the matrix of inputs and corresponding responses of the ANN. Then, the input and response vectors are normalized using the maximum value of each matrix, with the purpose of optimizing the creation of the ANN. Once the matrix of attributes was normalized, the ANN was created [11].

The neural networks are established based on the characteristics indicated in Table 5.

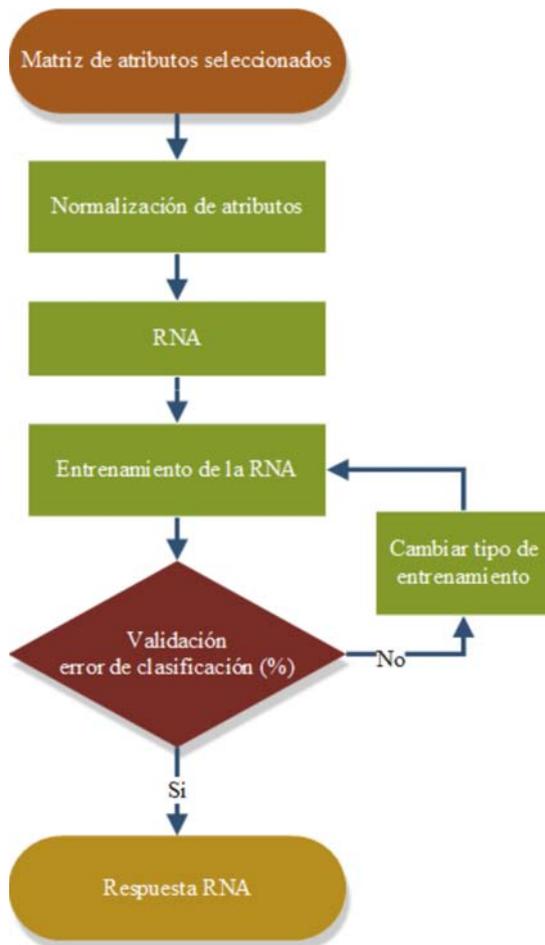


Figure 6. Flow diagram of the procedure for creating the ANN.

Table 5. Characteristics of the neural networks models

Characteristic	Tipo
Model	Backpropagation
Typology	4 layers Feedforward
On line/Off line	Off
Learning	Unsupervised
Learning rule	Error correction
Association	Hetero-association
Output	Analogous

Figure 7 includes the parameters of the creation of the feedforward backpropagation ANN for predicting the emission of the CO pollutant.

Similarly, Figure 8 includes the parameters of the creation of the feedforward backpropagation ANN for predicting the emission of the HC pollutant.

Once created, the networks were trained considering parameters such as: type of algorithm, number of epochs and maximum error.

The steps and formulas utilized to train the neural network are presented in the following:

1. The weights of the neural network are initialized with small random values.
2. An input pattern is entered to the network with the different conditions of the engine  $X_p$  ( $X_{p1}, X_{p2}, \dots X_{pn}$ ) and the target output of the network is specified as  $Y_m$ , which would be the value of emissions.
3. The actual output of the network is calculated.

The architecture of the network is shown in Figure 9, where subscript p indicates the p-th training vector, j is the number of hidden unit and the index i varies from 1 to the number of units of the input layer.

The classification error was verified for the CO and HC networks previously trained; if the error is greater than 5%, the parameters are changed to reduce such error.

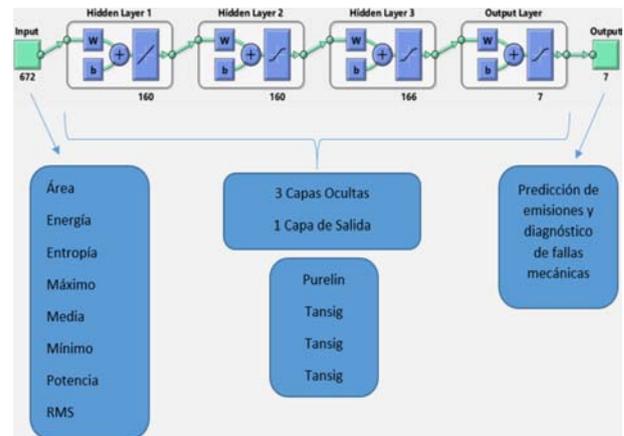


Figure 7. Structure of the CO neural network.

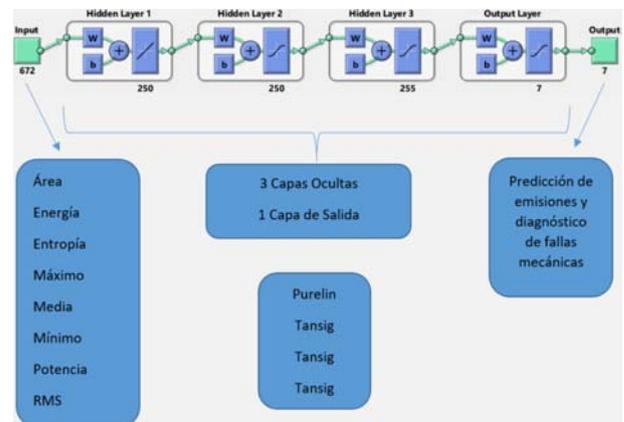


Figure 8. Structure of the HC neural network.

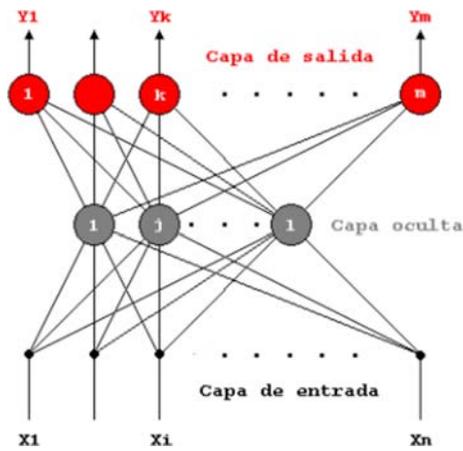


Figure 9. Architecture of the feedforward network.

The CO neural network utilizes the `trainscg` (Scaled Conjugate Gradient) training function for the system that predicts emissions and diagnoses mechanical failures, which presented an error of  $5.4061e^{-9}$ .

Similarly, the HC neural network utilizes the `trainscg` training function for predicting emissions and diagnosing mechanical failures, which presented an error of  $9.7587e^{-5}$ .

Figures 10 and 11 present the results of the Pearson correlation coefficient  $R$  of the CO and HC neural networks, respectively, which is provided by the training code `red.trainFcn='trainscg'` of the Matlab software.

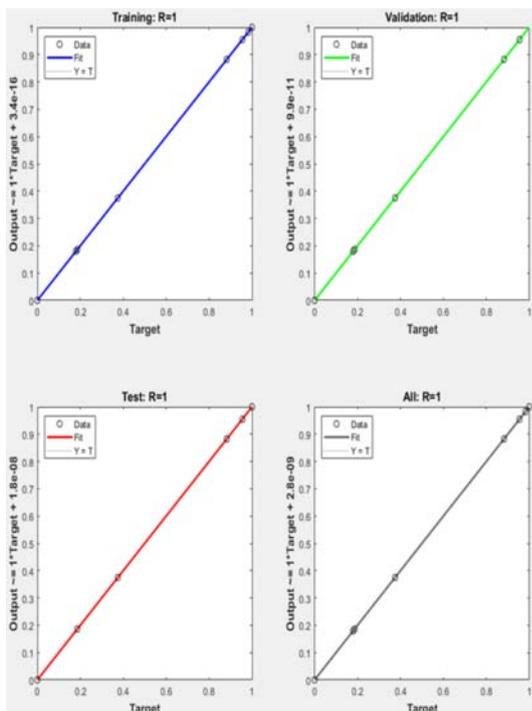


Figure 10. Correlation between the target values and the values predicted by the CO neural network.

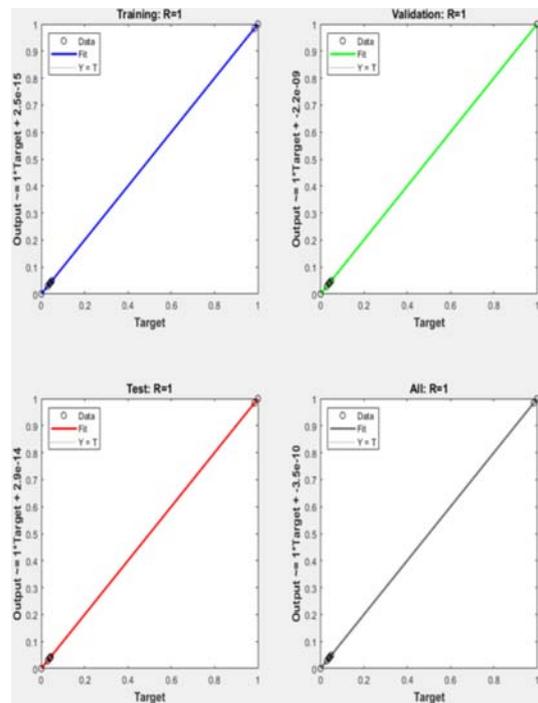


Figure 11. Correlation between the target values and the values predicted by the HC neural network.

The lines indicate the target values and the black circles represent the values predicted by the ANN. The prediction of the neural network is efficient and verifies a good performance, since it yields a global index of 1 in training, validation and testing, which indicates a strong positive linear relationship between the real conditions of the ignition engine and the results provided by the neural network [11].

Figures 12 and 13 show a comparison between the responses of the CO and HC neural network, respectively, and the corresponding target values; observe the seven actual mechanical conditions of the engine identified by the neural networks.

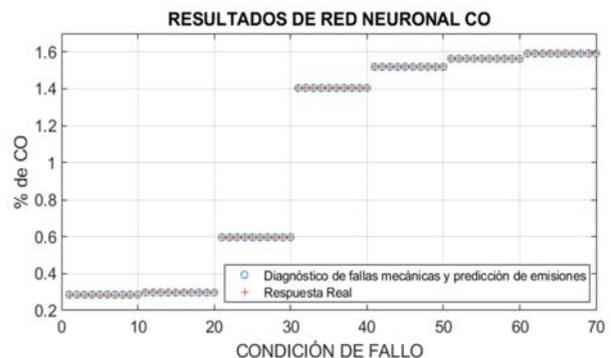


Figure 12. CO neural network with percentage of error  $5.4061e^{-9}$ , with training function `trainscg`.

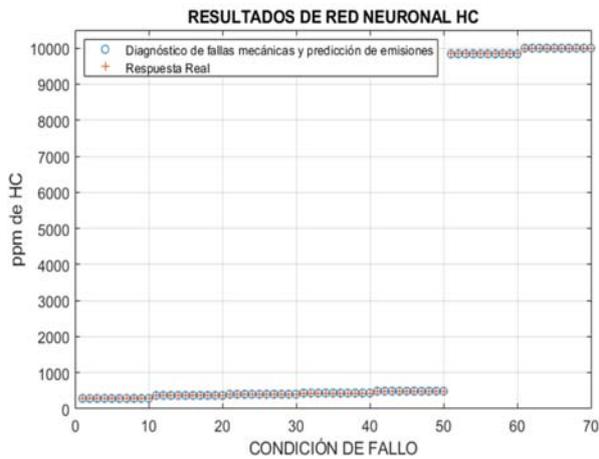


Figure 13. HC neural network with percentage of error  $9.7587e^{-5}$ , with training function trainscg.

### 3. Results and discussion

Various tests were carried out under diverse operating conditions, for the purpose of comparing the correct performance of the system that predicts emissions and diagnoses mechanical failures.

Two specific failure conditions are presented in this section: injector 2 (300) and coil 1-4 (1000).

Figure 14 shows the results of the values obtained by the CO neural network for the operating states when injector 2 fails.

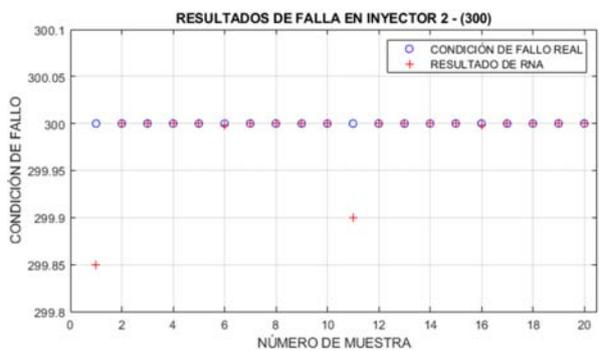


Figure 14. Result of the operating condition of injector 2.

Figure 15 shows the results of the values obtained by the HC ANN for operating conditions when coil 1-4 fails.

After obtaining the results of the operating conditions of the ignition engine, it may be remarked that the differences between the actual responses and the responses given by the CO and HC neural networks have a value close to zero. Therefore, the application of the system that diagnoses mechanical failures and predicts the pollutant emissions is capable of detecting the operational condition of mechanical failure and predicting the pollutant emission.



Figure 15. Result of the operating condition of the high voltage ignition coil 1-4.

Indeed, Figure 16 shows that after grouping the data corresponding to the actual engine condition and the responses obtained by the CO ANN, and in Figure 17 the corresponding to the HC ANN, using the Tukey statistical method with a confidence interval (CI) of 95 %, it is determined that the means are equivalent and there is no statistically significant difference, because the means of each of the responses coincide in a value close to zero.

In addition, Figures 18 and 19, which show the intervals of the CO and HC ANNs, respectively, indicate that there is no difference between the averages of the tests in the different operational conditions of the ignition engine.

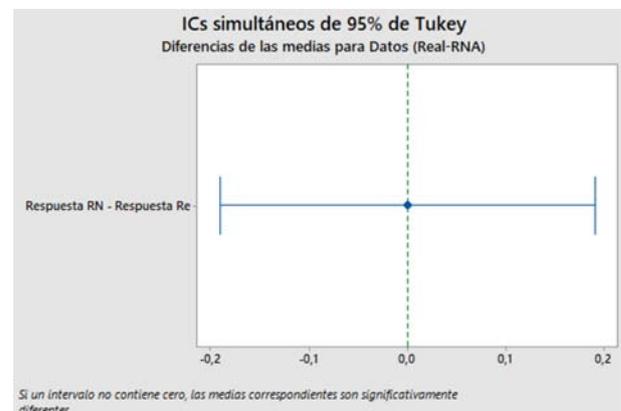
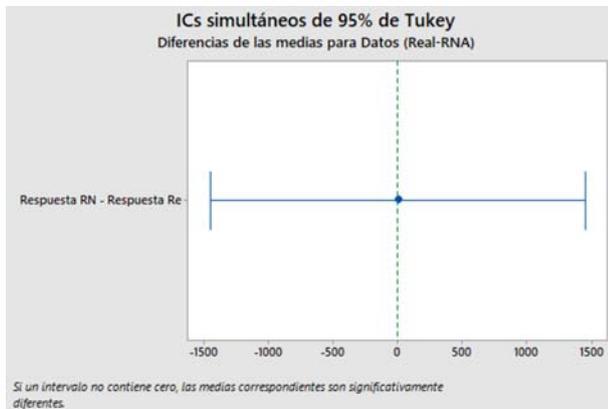
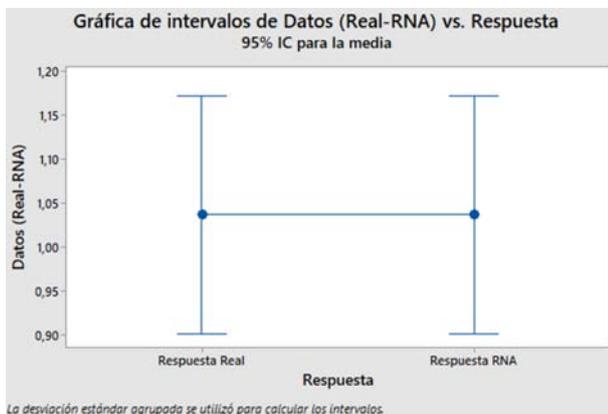


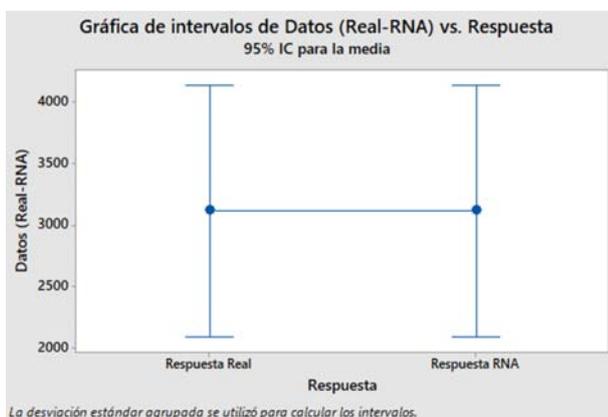
Figure 16. Graph of the differences of the means for data of actual response vs. the CO neural network.



**Figure 17.** Graph of the differences of the means for data of actual response vs. the HC neural network.



**Figure 18.** Graph of the data intervals of actual response vs. CO neural network.



**Figure 19.** Graph of the data intervals of actual response vs. HC neural network.

Similarly, Figures 20 and 21 corroborate that there is a relationship between the actual response and the response of the neural network, because they share the same grouping letter (A) and the p-value is equal to 1. This results in a reliability value of approximately

100.00 %, which is acceptable for issues of diagnosis of mechanical failures and prediction of polluting emissions of internal combustion ignition engines.

#### Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Respuesta	1	0,0000	0,000000	0,00	1,000
Error	138	45,1161	0,326928		
Total	139	45,1161			

#### Comparaciones en parejas de Tukey

##### Agrupar información utilizando el método de Tukey y una confianza de 95%

Respuesta	N	Media	Agrupación
Respuesta Real	70	1,0367	A
Respuesta RNA	70	1,0367	A

*Las medias que no comparten una letra son significativamente diferentes.*

**Figure 20.** Results of the analysis of variance and comparisons of Tukey pairs of the CO ANN.

#### Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Respuesta	1	0	0	0,00	1,000
Error	138	2590798996	18773906		
Total	139	2590798996			

#### Comparaciones en parejas de Tukey

##### Agrupar información utilizando el método de Tukey y una confianza de 95%

Respuesta	N	Media	Agrupación
Respuesta Real	70	3118	A
Respuesta RNA	70	3118	A

*Las medias que no comparten una letra son significativamente diferentes.*

**Figure 21.** Results of the analysis of variance and comparisons of Tukey pairs of the HC ANN.

## 4. Conclusions

The developed neural network models for the diagnosis and prediction of polluting emissions of both CO and HC, yield classification errors of  $5.4061e^{-9}$  and  $9.7587e^{-5}$ , respectively.

The traincg training function allows the precise identification of different types of mechanical conditions of the ignition engine and prediction of emissions, thus constituting a viable alternative to be integrated in a diagnosing system such as an automotive scanner or gas analyzer of gasoline powered vehicles, due to the computational speed offered by the artificial neural networks.

By means of the single factor analysis of variance a p-value equal to 1 was obtained, thus demonstrating that the actual response of classification of mechanical failures and prediction of emissions is equivalent to the result obtained through the developed neural networks, such that this value confirms that there is no statistically significant difference.

This work shows that the application of backpropagation feedforward neural networks is valid for the detection of mechanical failure conditions, as well as for the prediction of polluting emissions in gasoline powered vehicles; besides, the applied diagnosis technique has the advantage of avoiding disassembling elements and systems of the engine, by offering a technique which is minimally invasive, reliable and of great precision.

Results show that backpropagation feedforward neural networks with 160 or 250 hidden units and trained with the trainscg (Scaled Conjugate Gradient) function, may yield an average error of  $4.87962e^{-5}$ , which demonstrates that the emissions of gasoline powered vehicles can be predicted with high precision.

## References

- [1] F. Ballester, “Contaminación atmosférica, cambio climático y salud,” *Revista Española de Salud Pública*, vol. 79, pp. 159–175, 04 2005. [Online]. Available: <https://bit.ly/2piHMQZ>
- [2] A. Restrepo, S. Izquierdo, and R. López, “Estimación de factores que inciden sobre la contaminación ambiental generada por fuentes móviles en pereira,” *Scientia et Technica*, vol. 1, no. 37, pp. 267–272, 2007. [Online]. Available: <http://dx.doi.org/10.22517/23447214.4053>
- [3] X. Li, D. Sun, G. Lu, J. Krabicka, and Y. Yan, “Prediction of nox emissions through flame radical imaging and neural network based soft computing,” in *2012 IEEE International Conference on Imaging Systems and Techniques Proceedings*, July 2012, pp. 502–505. [Online]. Available: <https://doi.org/10.1109/IST.2012.6295594>
- [4] M. Cortina, “Aplicación de técnicas de inteligencia artificial a la predicción de contaminantes atmosféricos,” Ph.D. dissertation, 2012.
- [5] J. D. Martínez-Morales, E. R. Palacios-Hernández, and G. A. Velázquez-Carrillo, “Modeling and multi-objective optimization of a gasoline engine using neural networks and evolutionary algorithms,” *Journal of Zhejiang University SCIENCE A*, vol. 14, no. 9, pp. 657–670, Sep 2013. [Online]. Available: <https://doi.org/10.1631/jzus.A1300010>
- [6] T. Fontes, L. M. Silva, S. R. Pereira, and M. C. Coelho, “Application of artificial neural networks to predict the impact of traffic emissions on human health,” in *Progress in Artificial Intelligence*, L. Correia, L. P. Reis, and J. Cascalho, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 21–29. [Online]. Available: [https://doi.org/10.1007/978-3-642-40669-0\\_3](https://doi.org/10.1007/978-3-642-40669-0_3)
- [7] J. D. Martínez-Morales, E. R. Palacios-Hernández, and G. A. Velázquez-Carrillo, “Artificial neural network based on genetic algorithm for emissions prediction of a si gasoline engine,” *Journal of Mechanical Science and Technology*, vol. 28, no. 6, pp. 2417–2427, Jun 2014. [Online]. Available: <https://doi.org/10.1007/s12206-014-0535-2>
- [8] T. López Ortiz and A. Pacheco González, “Efectos de la contaminación atmosférica en la salud de las personas en la ciudad de cuenca,” 2015. [Online]. Available: <https://bit.ly/32M8jnz>
- [9] J. Guadalupe Almeida, “Modelación de emisiones contaminantes de fuentes móviles terrestres en Quito, Ecuador,” 2016. [Online]. Available: <https://bit.ly/2qdvTG8>
- [10] P. León Bacuilima and C. Piña Orellana, “Predicción de emisiones contaminantes de gases de escape a través de la presión media efectiva empleando redes neuronales en motores de encendido provocado, cuenca,” 2018. [Online]. Available: <https://bit.ly/2q2oTIT>
- [11] W. Contreras, J. Maldonado, and R. León, “Aplicación de una red neuronal feed-forward backpropagation para el diagnóstico de fallas mecánicas en motores de encendido provocado,” *INGENIUS*, 2019. [Online]. Available: <https://doi.org/10.17163/ings.n21.2019.03>



# EFFECT OF CHEMICAL BATH COMPOSITION ON MICROSTRUCTURE AND CORROSION RESISTANCE OF ZINC COATINGS BY HOT DIP: A REVIEW

## EFEECTO DE LA COMPOSICIÓN QUÍMICA DEL BAÑO EN LA MICROESTRUCTURA Y RESISTENCIA A LA CORROSIÓN DE LOS RECUBRIMIENTOS DE ZINC POR INMERSIÓN EN CALIENTE: UNA REVISIÓN

Jeanette Hernández<sup>1,\*</sup>, Maribel Suárez<sup>2</sup>

### Abstract

Metal coatings are widely used methods for protection of metal alloys, with the hot dip galvanizing process being one of the most industrially evolved. The objective of this work is to carry out a literature review about the influence of the addition of alloying elements in the bathroom, on the microstructure and corrosion behavior of zinc coatings obtained by the hot dipping technique. It was determined that the chemical composition of the galvanized baths influences the microstructural characteristics of the coatings and their corrosion behavior. The improvement of the corrosion resistance of zinc coatings is produced by the addition to the baths of elements generally more active than zinc, such as magnesium or aluminum that allow the formation of passive layers that delay the corrosive process.

**Keywords:** Zinc, Corrosion, Microstructure, Coating, Galvanized, hot dip.

### Resumen

Los recubrimientos metálicos son métodos ampliamente utilizados para la protección contra la corrosión de aleaciones metálicas, siendo el proceso de cincado por inmersión en caliente uno de los que presenta mayor evolución a nivel industrial. El objetivo de este trabajo es realizar una revisión bibliográfica sobre la influencia de la adición de elementos aleantes en el baño, en la microestructura y en el comportamiento a la corrosión de recubrimientos de zinc obtenidos por la técnica de inmersión en caliente. Se estableció que la composición química de los baños galvanizados influye en las características microestructurales de los recubrimientos y en su comportamiento a la corrosión. La mejora de la resistencia a la corrosión de los recubrimientos de zinc se produce por la adición a los baños de elementos generalmente más activos que el zinc, tales como el magnesio o el aluminio que permiten la formación de capas pasivas que retardan el proceso corrosivo.

**Palabras clave:** zinc, corrosión, microestructura, recubrimiento, galvanizado, inmersión en caliente.

<sup>1,\*</sup>Department of Metallurgy Engineering, Universidad Nacional Experimental Politécnica “Antonio José de Sucre” Vice Rectorate Barquisimeto – Venezuela. Corresponding author ✉: [jeanette\\_hernandez2000@yahoo.com](mailto:jeanette_hernandez2000@yahoo.com)

<http://orcid.org/0000-0002-8359-1058>

<sup>2</sup>School of Metallurgy Engineering and Science of Materials, Engineering Faculty, Universidad Central de Venezuela, Caracas – Venezuela. <http://orcid.org/0000-0003-0449-0417>

Received: 21-09-2019, accepted after review: 08-11-2019

Suggested citation: Hernández, J. and Suárez, M. (2020). «Effect of Chemical Bath Composition on Microstructure and Corrosion Resistance of Zinc Coatings by Hot Dip: A Review». INGENIUS. N.º 23, (january-june). pp. 40-52. DOI: <https://doi.org/10.17163/ings.n23.2020.04>.

## 1. Introducción

The processes for obtaining coatings through hot dip technologies, and in particular pure zinc coatings, have been used as protection against corrosion in different environments, especially urban, in which their behavior results very satisfactory in general. Nevertheless, in marine environments, in which the humidity and the concentration of sodium chloride are usually high, the resistance to corrosion of steels coated with pure zinc can significantly diminish [1].

Some alternatives proposed for overcoming this limitation include the generation of zinc-based coatings with the addition of some alloying elements, among which the most studied are aluminum, magnesium, silicon, copper, tin, nickel and other elements in a smaller proportion [1, 2].

From a technological point of view, the operational parameters to conduct the process of zinc coating on steel by the hot dip technique have remained practically invariable over time, but the new applications in the automotive and construction industries have caused that, at present, there exists a considerable number of research works about the aspects of the process and new types of zinc coatings [3, 4], which have increased the studies of the influence on the microstructure and the obtained coatings, of the addition of alloying elements to the zinc bath [3, 4].

In view of the above, the objective of this work is to conduct a literature review about the influence of adding alloying elements to the bath, on the microstructure and on the behavior against the corrosion of zinc coatings obtained through the hot dip technique.

## 2. Development of zinc coatings on steel

The American Galvanizers Association [5] indicates that the recorded history of galvanization dates back to the year 1742, when the French chemist Melouin, in a presentation to the Royal French Academy, described a method of iron coating through dipping in liquid zinc.

In 1836, Sorel, another French chemist, obtained a patent to manufacture a zinc coating on iron, after performing a cleaning with 9 % sulfuric acid and using a flux with ammonium chloride.

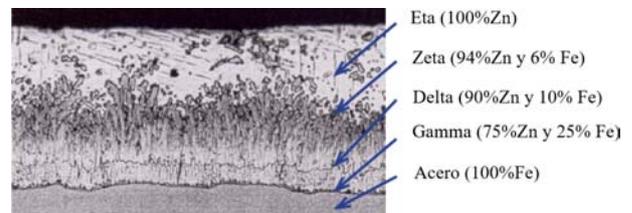
By 1850, the British galvanizing industry used yearly 10,000 tons of zinc for protection of the steel, and the first galvanizing plant in the United States was installed in 1870.

The coatings obtained through the galvanizing technique are widely used in diverse industries. Companies such as public services, chemical processes, pulp and paper, automotive and transportation, have extensively used galvanization for corrosion control, and

keep doing it nowadays. For more than 150 years, galvanization have had a proven history of commercial success as a method for protecting against corrosion in a great variety of applications worldwide.

Research studies conducted in recent years, have focused on the influence of the chemical composition of the bath on the microstructure and properties of the galvanized coatings. The changes in the chemical composition of the zinc baths search, undoubtedly, a significant improvement in the properties of the galvanized coating, either in its resistance to corrosion, in the weldability processes, and/or in its mechanical properties.

Figure 1 presents a typical coating of galvanized steel, where it can be observed three well-defined layers that are formed during the galvanized process. Each layer corresponds to the formation of intermetallic products, which according to their composition are known as Gamma, Delta and Zeta, and the top layer of pure zinc is Eta. Thru the galvanization process, the layers are naturally developed during a metallurgic reaction between the iron of the steel and the zinc of the bath [5].



**Figure 1.** Microstructure of a typical Zn coating, showing the formed layers with their respective chemical composition [5].

The research studies have been oriented to studying the metallurgy of the galvanized steel from a scientific point of view, to develop relations to the practical applications. These studies involve the systematic evaluation of the hot dip process, together with the diagrams of equilibrium or of phases Fe-Zn and the kinetics of formation of the layers of the Zn coating [3].

It should be noted that the aluminum is one of the elements mostly used in the zinc bath. The commercial development of these coatings, for instance Galvalume® (Zn-55Al) in 1970 and Galfan® (Zn-5Al) during 1980, successfully demonstrated that incorporating aluminum significantly improves the resistance to corrosion of the zinc base coating. Since then, numerous research studies have been carried out about the Zn-Al coatings applied on the steel by hot dip.

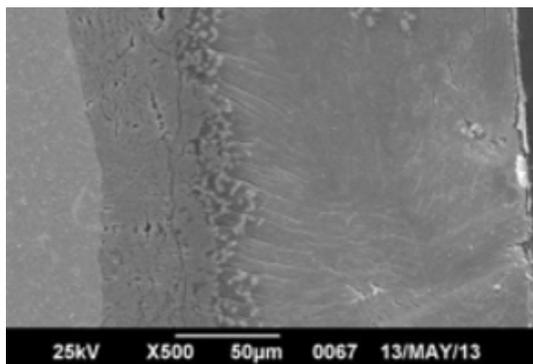
Galvalume® is a coating with approximately 55 % of aluminum, and also contains approximately 1.5 % of silicon which is added with the purpose of preventing an exothermic reaction in the coating/steel substrate interface. During the process, Fe-Zn-Al intermetallic

compounds are generated in the interface. The surface of the Galvalume® has characteristics of sequins, which consist in aluminum dendrites with a clearly measurable distance of the dendritic arm.

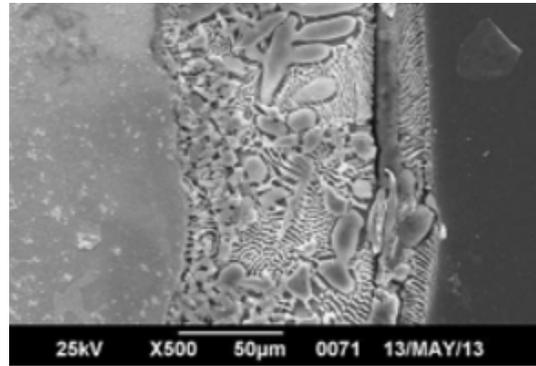
Galfan® is a zinc coating with 5 % of aluminum, which is close to the eutectic point, and has an excellent conformability and protection against corrosion. Two compositions have been reported based on the additions to the eutectic composition: small additions that contain lanthanum and cerium up to around 0.5 %, and additions of 0.5 % of magnesium. These additions are made to improve the wettability of the molten bath, without reducing the resistance of the coating to corrosion. The microstructure of the Galfan® is characterized by a two-phase structure, the eta ( $\eta$ ) pro-eutectoid phase rich in zinc, surrounded by a eutectoid type phase constituted by beta ( $\beta$ ) and eta ( $\eta$ ). Nevertheless, the structure may vary depending on the cooling velocity [3].

Figure 2 shows the latter effect, where Rico and Hernandez [6] studied the influence of the cooling velocity in the morphology and distribution of the phases of the Zn-7Al coatings on steel. The microstructures of the coating cooled in water shows the eutectic mixture segregated in the coating/steel interface, and the cooled in the furnace have a monotectoid micro-constituent of dendritic morphology.

According to the ASM Handbook: Corrosion [7], there is good evidence that small quantities of alloying elements may affect the resistance against corrosion of the galvanized coatings. The mechanisms for these effects are complex and not completely understood, but the literature suggests that the alloy elements segregate to the grain boundaries, where they can accelerate or delay the intergranular attack. Considering the alloy Zn/0.1-0.2Al typically used in hot galvanizing, it was found that adding it in the magnesium and copper bath reduces the intergranular corrosion mechanism, while the presence of bismuth, cadmium, lead and tin increases the localized attack. There are also some evidences that the addition of nickel to the bath reduces the hydrogen evolution reaction, and increases the resistance against the corrosion in environments that contain chlorides [8,9].



(a)



(b)

**Figure 2.** Micrographs through scanning electron microscopy (SEM) of Zn-7Al coatings, with different cooling conditions. A) water. B) furnace [6].

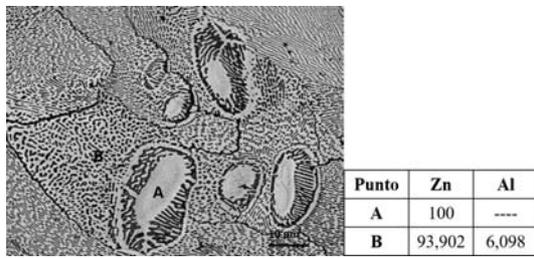
Recently, studies about the Zn-Al-Mg coatings have attracted attention. The reason for their systematic massive study is their performance in saline environments, where better behaviors against corrosion have been obtained, specifically in the order of 10 to 20 times with respect to Zn coatings, and 2 to 5 times with respect to Zn-Al coatings. Other important properties have also been found, such as the self-healing capacity (thus the area of the incision can be closed and protected again), scratch and wear resistance and lighter coating density, among other advantages [10].

Japanese, European and South African companies have already developed Zn-Al-Mg coatings, such as the Nisshin Steel Co., which has patented the ZAM® brand. Nevertheless, detailed information is still lacking in the open literature about the anticorrosive mechanism of the Zn-Mg-Al coating, and also little is specifically known about its electrochemical performance. On the other hand, the operating conditions of the hot dip process have not been optimized in order to obtain a smooth surface of such coating, and this is a challenge in this field [10].

Li *et al.* [2] evaluated the coating Zn-5Al-0.5Mg-0.08Si on alloy steel, which was compared with the coating Zn-5Al. They found that the resistance against corrosion of the Zn-5Al is not very high, compared to the resistance provided by magnesium and silicon.

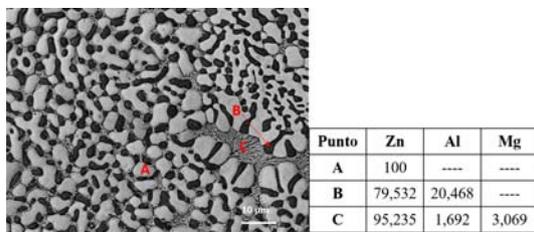
Figure 3 shows the morphology of the surface of the coating Zn-5Al, and the results corresponding to the chemical composition at points A and B. The coating consists in thick  $\beta$  (Zn) and binary eutectic phases of Zn-Al.

Due to the difference of potential between the two phases (Figure 3), when the coating is placed in the corrosive environment the phase with more electronegative potential will accelerate the corrosion process, while the phase with less electronegative potential will be protected from being corroded.



**Figure 3.** Micrograph through scanning electron microscopy (SEM) of a Zn-5Al coating and chemical composition in percentage of weight by energy dispersive Xray spectroscopy (EDS) at points A and B [2].

Figure 4 shows the image through scanning electron microscopy (SEM) of a Zn-5Al-0.5Mg-0.08Si coating. It can be seen that the separation of the dendritic arm of Zn-5Al of the coating significantly reduces from  $\sim 10 \mu m$  to  $2-5 \mu m$ , after adding the alloying elements Mg and Si. The phases  $\beta$  (Zn) (A), another rich in Al (B) and a tertiary eutectic MgZn<sub>2</sub>-Zn-Al (C) are formed [2]. The magnesium is only distributed in the grain boundaries, and this can cause that such grains become finer, thus preventing the intergranular corrosion.



**Figure 4.** Micrograph through scanning electron microscopy (SEM) of a Zn-5Al-0.5Mg-0.08Si coating and chemical composition in percentage of weight by energy dispersive Xray spectroscopy (EDS) at points A and B [2].

In this same sense, Schuerz *et al.* [11] carried out a study on steel sheets conventionally coated by hot dip of Zn (Z) and of Zn-Al-Mg (ZM) alloy, with the chemical composition indicated in Table 1.

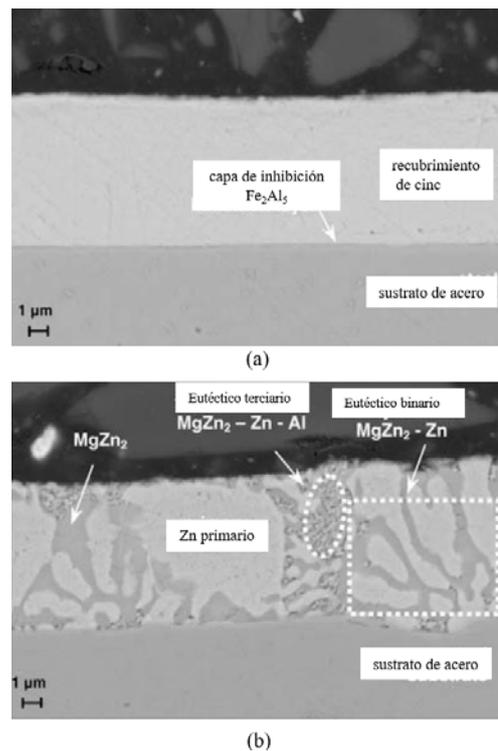
**Table 1.** Chemical composition, in percentage of weight, of the coatings evaluated [11]

Coating	Zn	Al	Mg
Z	99,8	0,2	—
ZM	96	2	2

The microstructure of the conventional coatings of zinc (Z) and zinc-aluminum-magnesium (ZM) differ considerably, as illustrated in Figure 5.

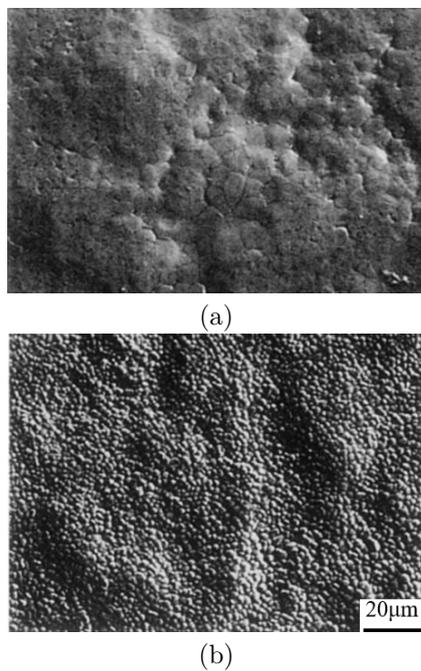
The Z coating consists of a layer of homogeneous material of pure zinc. The aluminum is enriched in the steel/coating interface in the form of a very thin layer of the intermetallic compound Al-Fe, which ensures a correct adherence of the coating on the steel surface.

On the other hand, the ZM coating consists of different phases: a zinc phase, which also includes small quantities of aluminum, an aluminum phase and the phase of MgZn<sub>2</sub>. The latter and the zinc also appear in the form of a binary eutectic (marked with a rectangle), as well as a tertiary eutectic of zinc, aluminum and the phase MgZn<sub>2</sub> (marked with an oval). The metallic coating ZM becomes a stable and adherent material, and a layer of oxide rich in aluminum that protects the steel substrate in corrosive environments. This layer is the main reason of the improved resistance against corrosion of the ZM coating in an environment that contains sodium chloride.



**Figure 5.** Micrograph through scanning electron microscopy (SEM) of (a) a conventional coating Z of Zn, and of (b) an alloy coating ZM of Zn-Al-Mg [11].

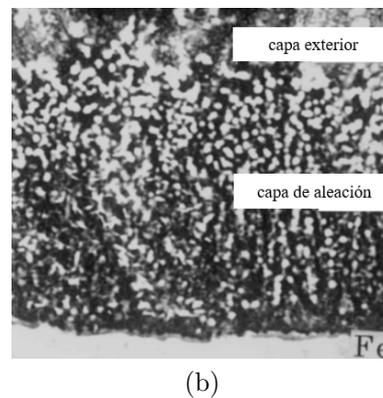
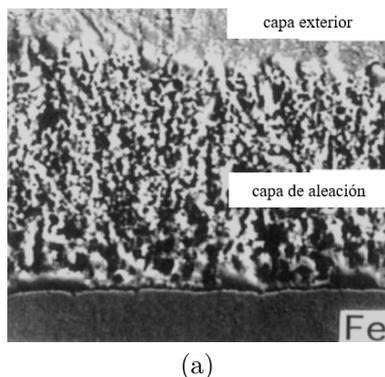
Other developments that have emerged as innovation to the processes of galvanizing by hot dip are the double dip processes, which were raised based on the operational problems during the hot dip of the Zn-Al coatings, since they do not have a good compatibility with the typical flux systems, and using others do not make it practical in the galvanization industry. Besides, the quantity of aluminum that gets in the coating layer represents an additional problem in the composition of the bath, since the concentration of aluminum is considerably reduced; this generates a continuous monitoring [4].



**Figure 6.** Superficial morphology by SEM of the coatings formed on steels of low carbon, by hot dip in two stages (a) Z6A and (b) Z6AMS [12].

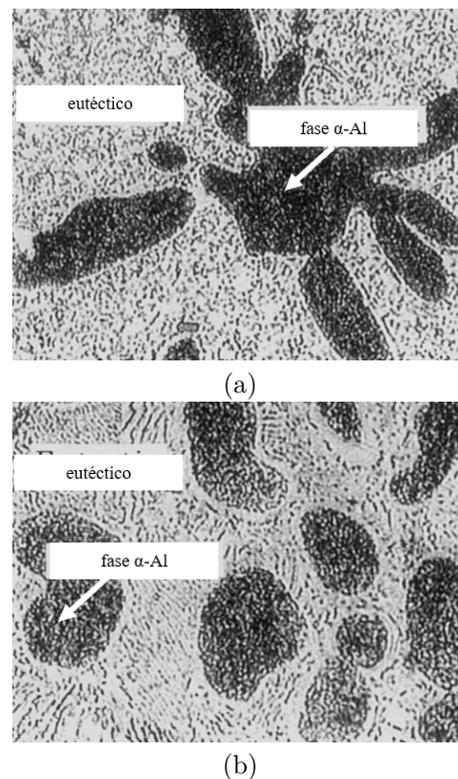
In this sense, Tanaka *et al.* [12] evaluated coatings on steel with low carbon content by hot dip in two stages (double dip), mainly in a bath of Zn and then in a bath of Zn-6Al (Z6A), with or without the addition of 0.5 % in weight of Mg and 0.1 % in weight of Si (Z6AMS). Figure 6 shows the morphology of the surface of the coatings formed by the hot dip of two stages with secondary baths of (a) Z6A and (b) Z6AMS. The surface of the coating formed in the Z6AMS showed very fine grains and few defects in comparison with the Z6A.

The layers of coating consisted in a duplex structure; an external layer and an internal alloy layer (Figure 7). They identified that the alloy layer was mainly constituted by an intermetallic phase of  $\text{Fe}_4\text{Al}_{13}\text{-Zn}$ . They observed that the crystals of this phase are randomly distributed in the Z6A coating, while the Z6AMS coating has a columnar structure with growth perpendicular to the steel substrate.



**Figure 7.** Microstructures of the cross sections of the coatings (a) Z6A and (b) Z6AMS [12].

Figures 8(a) and 8(b) show the external layer for the Z6A and Z6AMS coatings, respectively. The dark area is mainly a phase  $\alpha\text{-Al}$  and the clear area is the structure of a eutectic of Zn (Al) and  $\alpha\text{-Al}$ . It was found that the  $\alpha\text{-Al}$  in the Z6A coating in Figure 8(a) is much thicker than the one of the Z6AMS coating in Figure 8(b). The coating Z6AMS corroded slowly and in a homogeneous manner, while the coating of Z6A locally degrades due to a preferential corrosion along defects. The coating Z6AMS possesses Mg and Si in the external layer, and Si in the internal alloy layer, which makes it more anticorrosive.



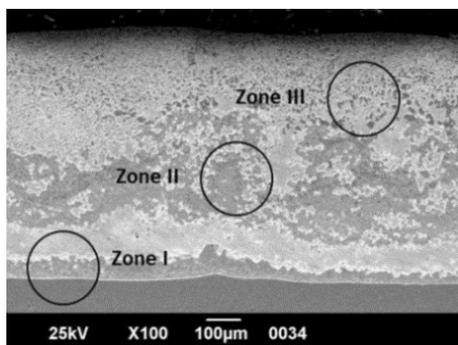
**Figure 8.** Microstructures of the external layer of the coatings (a) Z6A and (b) Z6AMS [12].

Among the double dip processes it is also high-

lighted the study conducted by Tachibana *et al.* [13], who evaluated a new coating by hot dip of Zn<sub>7</sub>Al on structural steel, which was applied through a double layer Zn-Zn<sub>7</sub>Al to prevent severe corrosion in coastal zones. The double coating was significantly effective in the prevention of the corrosion, particularly in the seaside. It was estimated that in the seaside, the life of the Zn-Zn<sub>7</sub>Al coating on the steel was four times the life of the steel coated with Zn. It was revealed that the excellent adhesiveness of the double coating Zn-Zn<sub>7</sub>Al on steel was due to the formation of the interface region that consists of a fine heterogeneous phase of a mixture of zinc, aluminum and iron.

Likewise, Rico and Carrasquero [8] carried out a microstructural characterization of zinc coatings on steel by the double dip method. The temperature of the dip bath was 550 °C, and the coatings were made using different immersion times. All the coatings showed three defined zones. The first zone, which corresponds to the steel/coating interface, presents properties which are typical of a traditional galvanization process at high temperatures. In zone II it was observed a very heterogeneous microstructure, while zone III showed a matrix of Zn-Fe with particles of compounds of Fe<sub>2</sub>Al<sub>5</sub>Zn<sub>0,4</sub>, as observed in Figure 9. The microstructural properties of the coatings were similar during the three immersion times. They determined the presence of Fe and Al along all the coating, due to the diffusion of the iron from the base metal and the aluminum provided by the second dip bath.

In such a way, the development of zinc coatings by hot dip has been based in modifying the chemical composition of the bath, which implies microstructural changes that improve the properties for a specific application. Aluminum [3, 6, 13–17], magnesium [2, 10–12, 16–19] and silicon [2] have been the alloying elements which have been more studied, with the purpose of establishing their influence in the general properties of these coatings. Other elements that have been studied in smaller proportion are copper [1, 20], titanium and nickel [21, 22].



**Figure 9.** Micrograph through SEM of the Zn/Zn-5Al coating with a total immersion time of 120 s, in which the three zones of the coating are shown [8].

### 3. Behavior against corrosion of the zinc coatings

The zinc coatings add resistance against corrosion in various ways; one as a barrier layer, since the continuous coating separates the steel from the corrosive environment; other by galvanic protection, since the zinc acts as a sacrificial anode that protects the steel in areas adjacent to the holes, scratches and edges of the coating. The behavior of the zinc can be seen in a galvanic series where zinc is less noble than steel in the majority of the media at ambient temperature [3].

A porous superficial layer of zinc oxide is formed in the surface due to a mechanism of dissolution/re-precipitation, which leads to preferential corrosion routes in the high porosity area or discontinuities; this explains the linear corrosion rate. Besides, after the dissolution of the metallic zinc, the zinc hydroxide may precipitate to the cathodic areas of the exposed steel, forming a secondary barrier layer. As a consequence, the zinc coating will corrode at a slower velocity than the steel substrate, even though the zinc corrosion rate will vary depending on the exposition environment [3].

In the case of Zn-Al coatings such as Galfan®, the corrosion in non-marine environments occurs in two stages: temporary protection due to the formation of passive layer of aluminum oxide, and the galvanic action of the zinc where the zinc sulfate is formed on the oxide layer. The kinetics is controlled by the second stage due to the diffusing need through the oxide layer, which explains the decrease of the corrosion velocity of the Galfan® with respect to conventional zinc coatings. It has been also demonstrated that the phase rich in zinc ( $\eta$  - *eta*) is preferentially corroded respect to the pro-eutectic and eutectic. The additions of Mg (up to 0.8 % in weight) increase the quantity of pro-eutectic of phase Al, which improves the resistance against corrosion. The Galfan® corrodes in a parabolic manner, while conventional coatings corrode in a linear manner [3].

The atmospheric corrosion of the Galvalume® coating occurs in the inter-dendritic regions rich in zinc, which enables a galvanic protection. Then, the products of corrosion remain trapped in those spaces, and act as a barrier against corrosion. The corrosive process is of parabolic type [3].

In view of the importance of systematically evaluating the behavior against corrosion of these coatings in different environments, multiple researchers have embarked in conducting accelerated corrosion tests, the traditional that involve exposition at atmospheric conditions, by dip and in saline mist chamber, to determine the velocity of corrosion, as well as the electrochemical techniques that allow to obtain more information about the mechanism and kinetics of the corrosion. These studies are normally carried out as a comparison of different coatings, in order to estab-

lish analogies and differences that allow industry to decide about the applicability of each of the coatings evaluated.

Table 2 shows a summary of the utilized techniques and the environments in which the different zinc base coatings are evaluated.

On the other hand, the study of the behavior against corrosion of the zinc coatings depends on many factors, but the most common orientation of the research studies is to relate the microstructural changes of the coatings with their resistance to corrosion; the addition of alloying elements to traditional zinc baths allow to obtain these changes. Also, in these research studies becomes important to establish the mechanisms that govern the corrosive processes, based on techniques that enable the characterization of the products of corrosion, to relate them with the species present in the environment and the elements that constitute the material [9, 15, 18, 19, 23, 24].

In this sense, Hosking *et al.* [25] studied the behavior against corrosion of Zn-Mg coatings, in comparison with traditional zinc coatings. Specifically, they followed an experimental program that included conducting corrosion tests in an acidified solution of NaCl, characterizing the products of the corrosion and proposing a mechanism of the corrosive process. Results revealed that the resistance against corrosion of the Zn-Mg coatings was three times the resistance of conventional

zinc coatings, and that this improvement was caused by the precipitation on the corrosive surface of insoluble Simonkolleite ( $\text{Zn}_5\text{Cl}_2(\text{OH})_8 \cdot \text{H}_2\text{O}$ ) possibly modified with absorbed magnesium, magnesium hydroxide ( $\text{Mg}(\text{OH})_2$ ) and a specie of hydroxycarbonate, thus delaying the corrosion reactions.

Similarly, Yao *et al.* [10] comparatively evaluated the behavior against corrosion of three types of zinc coatings with different combinations of Zn, Al, Mg and Cu (Zn-0.1 Cu (G), Zn-5Al-0.1Cu (ZA) and Zn-5Al-1Mg-0.1Cu (ZAM)). They conducted saline mist chamber and electrochemical tests in an aqueous solution of 5 and 3.5 %, respectively, of NaCl, to study the behavior against corrosion of the coatings.

Figures 10(a) and 10(b) show the polarization curves of the coatings, superficially prepared and after five days of tests in the saline mist chamber, respectively. Some changes in the parameters of the corrosion process were observed. There is a region in the polarization curve of the ZAM coating, in the range from  $-1.5$  V to  $-1.2$  V (Figure 9(b)), in which the current density is kept relatively constant independently of the variation of the potential; this indicates that the cathodic reaction of the ZAM coating was inhibited in this range of potential, due to the precipitation of corrosion products in the cathodic area that act as inhibitors, thus increasing the impedance of the surface of the coating.

**Table 2.** Techniques and environments to evaluate the behavior against corrosion of the zinc coatings

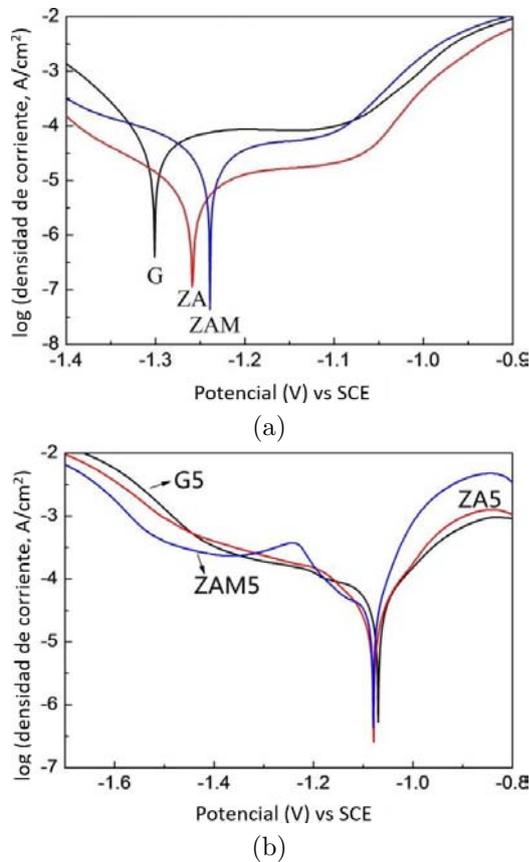
Type of coating	Corrosive environment	Evaluation technique	Ref.
Zn Zn-4,9Al	5 % NaCl in laboratory 3 % NaCl in laboratory	Saline mist chamber Potentiodynamic polarization	[26]
Zn-5Al Zn-55Al	Real Seawater (78 % NaCl) Synthetic Seawater (68 % NaCl)	Measurement of potentials	[27]
Zn Zn-5Al	3 % NaCl in laboratory	Saline mist chamber Potentiodynamic polarization	[28]
Zn-0,2Al Zn-2Al-2Mg	5 % NaCl in laboratory	Saline mist chamber	[11]
Zn Zn-5Al Zn-55Al	Atmospheric in field	Field tests	[29]
Zn Zn-55Al	Atmospheric in field Saline mist chamber and of SO <sub>2</sub> in the laboratory	Electrochemical impedance spectroscopy	[30]
Zn-55Al	0.5 M NaCl 0,5 M Na <sub>2</sub> SO <sub>4</sub> 0.5 M NaClO <sub>4</sub>	Electrochemical impedance spectroscopy	[31]
Zn Zn-5Al	5 % ammoniacal solution	Test of cyclic corrosion	[32]
Zn/Zn-6Al Zn/Zn-6Al-0,5Mg-0,1Si	5 % NaCl in laboratory	Dip	[12]
Zn Zn/Zn-7A	Real Seawater	Dip	[13]

Figures 11(a) and 11(b) show the EIS spectra for the coatings G, ZA and ZAM. The shapes of the (a) Nyquist diagrams and of the (b) phase angle Bode diagrams for the three coatings are similar. The Nyquist diagrams consist of a semicircle at high frequency and a circle at low frequency. The phase angle Bode diagrams exhibit two crests of wave, which indicates that the three coatings may have similar corrosion mechanisms. While the ZA coating has a larger impedance than the other two coatings in the high frequency range (10 Hz ~ 100000 Hz), the G coating has a relatively larger impedance than the ZA and ZM coatings at low frequency (0.01 Hz ~ 10 Hz).

A greater value of impedance normally indicates a lower corrosion velocity of the material. The impedance

of the ZAM coating is lower than other at this stage, which may be due to the addition of less noble Mg; this will make it react first when the ZAM coating is in a NaCl solution, contributing to a high density of current of corrosion at the initial stage.

The protective nature of the ZAM coating can be attributed to the initial corrosion of the phases rich in Mg. The corrosion products of Zn, Al and Mg are agglomerated in the cathodic area and act as inhibitors, blocking the corrosion routes (the micro-routes for the diffusion of O<sub>2</sub> and H<sub>2</sub>O) along the grain boundaries of the Zn crystals and increasing the impedance of the coating, thus the general corrosion process of the ZAM coating is delayed.

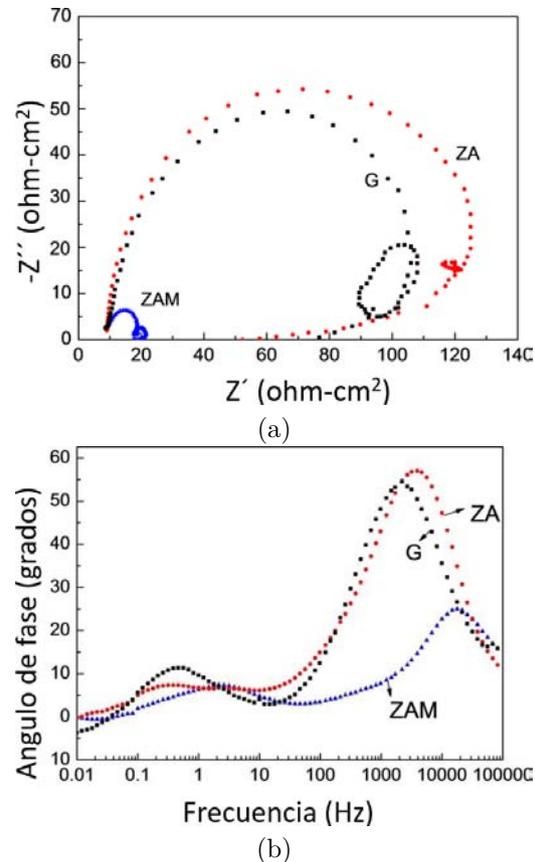


**Figure 10.** Potentiodynamic polarization curves for the zinc coatings; (a) superficially prepared, (b) after 5 days of saline mist chamber test [10].

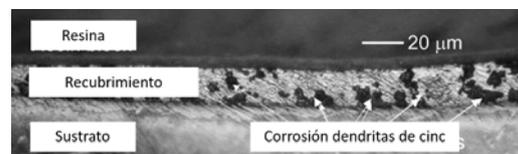
As was previously indicated, the behavior against corrosion depends on many factors, and among the most studied are the referred to the operating parameters of the hot dip process. For instance, Elvins *et al.* [14] studied the effect of the cooling velocity on the microstructure and the resistance against corrosion of the Galfan® (Zn-4,5Al) coating on the steel. They applied three different cooling velocities.

The increase in the cooling velocity did not significantly alter the volume of the fraction of primary zinc; nevertheless, the size and number of primary dendrites of zinc were altered. The specimens cooled quickly contained small but numerous dendrites, and on the ones of slow cooling there were less dendrites of bigger size.

The quantity of zinc dendrites influence the corrosion process since it is the phase that is preferentially corroded, thus resulting in a larger anodic area and consequently a larger corrosion velocity as the cooling velocity increases. Figure 12 shows the preferential corrosion of the zinc dendrites with respect to the eutectic.



**Figure 11.** Electrochemical impedance spectroscopy of the zinc coatings (a) Nyquist diagram, (b) Bode diagram [10].



**Figure 12.** Microstructure of the Zn-4,5Al coating, where it can be observed the preferential attack of the zinc dendrites [25].

Another way to study mechanisms and understand in a better manner the behavior against corrosion of the zinc coatings, is utilizing techniques of analysis of the corrosion products to propose process mechanisms, such as the work by Carbuicchio *et al.* [33], who studied the corrosion products formed in galvanized steel pipes exposed to hot water for two years, through metallographic techniques of X ray diffraction, as well as the Mössbauer and Raman spectroscopic techniques. The results obtained allowed to establish that the corrosive process was caused by an excess in the service temperature, occurring an inversion of the electrochemical polarity between the zinc and the iron, growing tubers of corrosion products in the dezincified zones of the pipe walls until perforation. The compounds present in the corrosion products allowed to suggest a mechanism of the process.

On the other hand, it is also important to mention that the corrosive processes are affected by some mechanical stresses, or some specific element such as hydrogen that contributes to enhance the corrosive process. As a consequence, there are also research works that evaluate synergic behaviors of the corrosion with other types of wear or mechanical stresses. In this sense, Reumont *et al.* [34], studied the influence of the microstructure in the susceptibility to assisted stress corrosion cracking (SCC) of a galvanized steel in an aqueous solution of sodium chloride. Deformation tests were carried out in an aqueous solution of NaCl at 3 %. The general behavior of the galvanized steel is ductile. There is a transition from ductile to fragile for a particular velocity of deformation in this environment. Little mechanical resistance of the coating and the hydrolysis reaction of the zinc is necessary to lead to the fragility, which may be associated with a fragility process by hydrogen. The formation of intermetallic components may produce a coating with microstructure resistant to the SCC.

A similar work was carried out by Sacco *et al.* [31], who evaluated the behavior of steel sheets with conventional zinc coatings and with Zn-55Al coatings uniaxially tractioned and exposed to solutions 0.5 M NaCl, 0.5 M Na<sub>2</sub>SO<sub>4</sub>, and 0.5 M NaClO<sub>4</sub>. The specimens were deformed until reaching elongations of 1.3, 3.75, 6.25, 7.5 and 12.5 %. These and other specimens without deformation were subjected to electrochemical tests and measures of electrochemical impedance. The results obtained allowed to infer that there are notable differences in the kinetics of the corrosion reactions that take place on the surface of both types of coatings, and that for both steel/metallic coating systems such kinetics is a function not only of the degree of applied deformation, but also of the composition of the aggressive medium under consideration.

Rico and Carrasquero [4] carried out a revision where it is established that the chemical composition of the galvanized baths, strongly influence the microstructural characteristics of the coatings and their mechanical behavior. This assertion together with the studies reported in this work, establish that the behavior against corrosion is also affected by the addition of elements to the zinc bath; besides, this is supported by other works that studied the synergic effects of the corrosion with the mechanical behavior [31, 34–38].

As was mentioned in the first section of this paper, other research works highlight because of their purpose of improving the properties against corrosion of zinc coatings by carrying out double dip processes, which consist of consecutively immersing the steel in two liquid baths with different chemical compositions; the first immersion is in a conventional galvanizing bath, and the second in a bath of a zinc base alloy with other elements, such as aluminum, magnesium and silicon. In this sense, Tanaka *et al.* [12] demonstrated that the

double dip has advantages such as good adherence of the coating and the viability of better controlling the thickness and the microstructure of the coating, thus also improving its resistance to corrosion. Standardized laboratory tests were carried out to establish the improvement in the behavior against corrosion.

Tachibana *et al.* [13], applying a double dip process as well, determined that the coatings obtained by this method exhibit a better performance against corrosion than the one manufactured by conventional methods. These results were obtained in tests carried out in a real saline environment.

It is important to remark that although some research works have been conducted about the application of double dip processes, there is no accurate information about the mechanisms that enable to establish the improvement in the resistance to corrosion of these processes compared with other conventional methods; nevertheless, it is estimated that the addition to the baths of elements generally more active than zinc, such as magnesium and aluminum, promote the formation of passive layers that delay the corrosive process.

## 4. Conclusions

The chemical composition of the galvanized baths influence the microstructural characteristics of the coatings.

The increase in the resistance against corrosion of the zinc coatings by hot dip is oriented to change the chemical composition of the baths, with the purpose of expanding the applicability of these coatings on more aggressive environments such as those that contain chlorides.

The improvement in the resistance against corrosion of the zinc coatings is caused by the addition to the baths of elements generally more active than zinc, such as magnesium or aluminum, that allow the formation of passive layers that delay the corrosive process.

The studies of the Zn-Al-Mg coatings have reported behaviors against corrosion of the order of 10 to 20 times with respect to the Zn coating, and of 2 to 5 times with respect to the Zn-Al coating.

## References

- [1] J. Cervantes, A. Barba, M. Hernández, J. Salas, J. Espinoza, C. Dénova, G. T.-V. nor, A. Conde, A. Covelo, and R. Valdez, “Obtención y caracterización de recubrimientos Zn-Al-Cu por inmersión en caliente sobre aceros de bajo carbono,” *Revista de Metalurgia*, vol. 49, no. 5, pp. 351–359, 2013. [Online]. Available: <https://doi.org/10.3989/revmetalm.1316>

- [2] L. Shiwei, G. Bo, T. Ganfeng, H. Yi, H. Liang, and Y. Shaohua, "Study on the corrosion mechanism of zn-5al-0.5mg-0.08si coating," *Journal of Metallurgy*, vol. 2011, Article ID 917469, pp. 1–6, 2011. [Online]. Available: <https://doi.org/10.1155/2011/917469>
- [3] A. R. Marder, "The metallurgy of zinc-coated steel," *Progress in Materials Science*, vol. 45, no. 3, pp. 191–271, 2000. [Online]. Available: [https://doi.org/10.1016/S0079-6425\(98\)00006-1](https://doi.org/10.1016/S0079-6425(98)00006-1)
- [4] Y. Rico and E. Carrasquero, "Efecto de la composición química en el comportamiento mecánico de recubrimientos galvanizados por inmersión en caliente: una revisión," *Ingenius*, no. 18, pp. 30–39, 2017. [Online]. Available: <http://dx.doi.org/10.17163/ings.n18.2017.04>
- [5] American Galvanizers Association, "Hot-dip galvanizing for corrosion protection a specifier's guide," 2012. [Online]. Available: <https://bit.ly/2pGEzeb>
- [6] Y. Rico and J. Hernández, "Influencia de la velocidad de enfriamiento sobre la microestructura y comportamiento a la corrosión de recubrimientos Zn-7Al por inmersión en caliente sobre acero," *Revista Latinoamericana de Metalurgia y Materiales*, vol. 35, pp. 269–275, 12 2015. [Online]. Available: <https://bit.ly/2OaV0sy>
- [7] S. D. Cramer and J. Bernard S. Covino, "Corrosion: Fundamentals, testing, and protection," ASM International. The Materials Information Society, Tech. Rep., 2003. [Online]. Available: <https://bit.ly/2pGH6Fd>
- [8] Y. Rico and E. J. Carrasquero, "Microstructural evaluation of double-dip galvanized coatings on carbon steel," *MRS Advances*, vol. 2, no. 62, pp. 3917–3923, 2017. [Online]. Available: <https://doi.org/10.1557/adv.2017.608>
- [9] P. Volovitch, T. N. Vu, C. Allély, A. A. Aal, and K. Ogle, "Understanding corrosion via corrosion product characterization: II. role of alloying elements in improving the corrosion resistance of Zn-Al-Mg coatings on steel," *Corrosion Science*, vol. 53, no. 8, pp. 2437–2445, 2011. [Online]. Available: <https://doi.org/10.1016/j.corsci.2011.03.016>
- [10] C. Yao, S. L. Tay, J. H. Yang, T. Zhu, and W. Gao, "Hot dipped zn-al-mg-cu coating with improved mechanical and anticorrosion properties," *International Journal of Electrochemical Science*, vol. 9, pp. 7083–7096, 2014. [Online]. Available: <https://bit.ly/2XChPZq>
- [11] S. Schuerz, M. Fleischanderl, G. H. Luckeneder, K. Preis, T. Haunschmied, G. Mori, and A. C. Kneissl, "Corrosion behaviour of zn-al-mg coated steel sheet in sodium chloride-containing environment," *Corrosion Science*, vol. 51, no. 10, pp. 2355–2363, 2009. [Online]. Available: <https://doi.org/10.1016/j.corsci.2009.06.019>
- [12] J. Tanaka, K. Ono, S. Hayashi, K. Ohsasa, and T. Narita, "Effect of mg and si on the microstructure and corrosion behavior of zn-al hot dip coatings on low carbon steel," *ISIJ International*, vol. 42, no. 1, pp. 80–85, 2002. [Online]. Available: <https://doi.org/10.2355/isijinternational.42.80>
- [13] K. Tachibana, Y. Morinaga, and M. Mayuzumi, "Hot dip fine zn and zn-al alloy double coating for corrosion resistance at coastal area," *Corrosion Science*, vol. 49, no. 1, pp. 149–157, 2007, progress in Corrosion Research in Commemoration of Centenary of Birth of Professor Go Okamoto. [Online]. Available: <https://doi.org/10.1016/j.corsci.2006.05.015>
- [14] J. Elvins, J. A. Spittle, and D. A. Worsley, "Microstructural changes in zinc aluminium alloy galvanising as a function of processing parameters and their influence on corrosion," *Corrosion Science*, vol. 47, no. 11, pp. 2740–2759, 2005. [Online]. Available: <https://doi.org/10.1016/j.corsci.2004.11.011>
- [15] D. Thierry and N. LeBozec, "Corrosion products formed on confined hot-dip galvanized steel in accelerated cyclic corrosion tests," *CORROSION*, vol. 65, no. 11, pp. 718–725, 2009. [Online]. Available: <https://doi.org/10.5006/1.3319098>
- [16] M. Dutta, A. K. Halder, and S. B. Singh, "Morphology and properties of hot dip zn-mg and zn-mg-al alloy coatings on steel sheet," *Surface and Coatings Technology*, vol. 205, no. 7, pp. 2578–2584, 2010. [Online]. Available: <https://doi.org/10.1016/j.surfcoat.2010.10.006>
- [17] N. LeBozec, D. Thierry, D. Persson, C. K. Riemer, and G. Luckeneder, "Influence of microstructure of zinc-aluminium-magnesium alloy coated steel on the corrosion behavior in outdoor marine atmosphere," *Surface and Coatings Technology*, vol. 374, pp. 897–909, 2019. [Online]. Available: <https://doi.org/10.1016/j.surfcoat.2019.06.052>
- [18] P. Volovitch, C. Allely, and K. Ogle, "Understanding corrosion via corrosion product characterization: I. case study of the role of mg alloying in zn-mg coating on steel," *Corrosion Science*, vol. 51, no. 6, pp. 1251–1262, 2009. [Online]. Available: <https://doi.org/10.1016/j.corsci.2009.03.005>

- [19] S. Schürz, G. H. Luckeneder, M. Fleischanderl, P. Mack, H. Gsaller, A. C. Kneissl, and G. Mori, "Chemistry of corrosion products on zn-al-mg alloy coated steel," *Corrosion Science*, vol. 52, no. 10, pp. 3271–3279, 2010. [Online]. Available: <https://doi.org/10.1016/j.corsci.2010.05.044>
- [20] P. Sixtos, J. Cervantes, A. Barba, J. Roviroza, E. Garduño, R. Valdez, M. Trujillo, J. González, J. Salas, and G. Torres, "Evaluación de recubrimientos de aleaciones Zn-Al-Cu obtenidos por inmersión en caliente y por vía termoquímica," in *XV Congreso Internacional Anual de la SOMIM, del 23 al 25 de septiembre del 2009. Cd. Obregón, Sonora, México*, 2009, pp. 678–682. [Online]. Available: <https://bit.ly/34gcWaQ>
- [21] A. Barba, J. Cervantes, R. González, D. Hernández, V. Hernández, R. Valdez, A. Covelo, and M. Hernández, "Influencia de la adición de pequeñas cantidades de Ni a baños de galvanizado por inmersión en caliente, en las propiedades de aceros al carbono galvanizados," in *XXIV Congreso Internacional Anual de la SOMIM 19 al 21 de septiembre de 2018, Campeche, México*, 2018. [Online]. Available: <https://bit.ly/37xBQVt>
- [22] C. M. Cotell, J. A. Sprague, and J. F.A. Smidt, *ASM Handbook, Volume 5: Surface Engineering*, ASM International, Ed., 1994. [Online]. Available: <https://bit.ly/35qcdUQ>
- [23] Y. Li, "Formation of nano-crystalline corrosion products on zn-al alloy coating exposed to seawater," *Corrosion Science*, vol. 43, no. 9, pp. 1793–1800, 2001. [Online]. Available: [https://doi.org/10.1016/S0010-938X\(00\)00169-4](https://doi.org/10.1016/S0010-938X(00)00169-4)
- [24] M. S. Azevedo, C. Allély, K. Ogle, and P. Volovitch, "Corrosion mechanisms of zn(mg, al) coated steel in accelerated tests and natural exposure: 1. the role of electrolyte composition in the nature of corrosion products and relative corrosion rate," *Corrosion Science*, vol. 90, pp. 472–481, 2015. [Online]. Available: <https://doi.org/10.1016/j.corsci.2014.05.014>
- [25] N. C. Hosking, M. A. Ström, P. H. Shipway, and C. D. Rudd, "Corrosion resistance of zinc-magnesium coated steel," *Corrosion Science*, vol. 49, no. 9, pp. 3669–3695, 2007. [Online]. Available: <https://doi.org/10.1016/j.corsci.2007.03.032>
- [26] M. Manna, G. Naidu, N. Rani, and N. Bandyopadhyay, "Characterisation of coating on rebar surface using hot-dip zn and zn-4.9al-0.1 misch metal bath," *Surface and Coatings Technology*, vol. 202, no. 8, pp. 1510–1516, 2008. [Online]. Available: <https://doi.org/10.1016/j.surfcoat.2007.07.001>
- [27] A. P. Yadav, H. Katayama, K. Noda, H. Masuda, A. Nishikata, and T. Tsuru, "Effect of al on the galvanic ability of zn-al coating under thin layer of electrolyte," *Electrochimica Acta*, vol. 52, no. 7, pp. 2411–2422, 2007. [Online]. Available: <https://doi.org/10.1016/j.electacta.2006.08.050>
- [28] H. C. Shih, J. W. Hsu, C. N. Sun, and S. C. Chung, "The lifetime assessment of hot-dip 5% al-zn coatings in chloride environments," *Surface and Coatings Technology*, vol. 150, no. 1, pp. 70–75, 2002. [Online]. Available: [https://doi.org/10.1016/S0257-8972\(01\)01508-0](https://doi.org/10.1016/S0257-8972(01)01508-0)
- [29] O. de Rincón, A. Rincón, M. Sánchez, N. Romero, O. Salas, R. Delgado, B. López, J. Uruchurtu, M. Marroco, and Z. Panosian, "Evaluating Zn, Al and Al-Zn coatings on carbon steel in a special atmosphere," *Construction and Building Materials*, vol. 23, no. 3, pp. 1465–1471, 2009. [Online]. Available: <https://doi.org/10.1016/j.conbuildmat.2008.07.002>
- [30] S. Flores and K. Paucar, "Evaluación de sistemas de recubrimientos mediante espectroscopía de impedancia electroquímica," *Jornadas SAM - CONAMET - Simposio Materia 2003*, 2003. [Online]. Available: <https://bit.ly/2KIjsPR>
- [31] E. Sacco, J. Culcasi, C. Eisner, and A. D. Darli, "Efecto de la deformación por tracción sobre la orientación cristalográfica del recubrimiento y la resistencia a la corrosión en CNS de chapas de acero con recubrimientos base cinc," *Revista de Metalurgia*, vol. 38, no. 6, pp. 403–409, 2002. [Online]. Available: <https://doi.org/10.3989/revmetal.2002.v38.i6.425>
- [32] H. Fujisawa, R. Kaneko, and H. Ishikawa, "Hot-Dip Zn-5 % Al alloy-coated Steel Sheets JFE ECOGAL ®," *JFE Technical Report*, no. 14, pp. 41–45, 2009. [Online]. Available: <https://bit.ly/335wY6I>
- [33] M. Carbucicchio, R. Ciprian, F. Ospitali, and G. Palombarini, "Morphology and phase composition of corrosion products formed at the zinc-iron interface of a galvanized steel," *Corrosion Science*, vol. 50, no. 9, pp. 2605–2613, 2008. [Online]. Available: <https://doi.org/10.1016/j.corsci.2008.06.007>
- [34] G. Reumont, J. B. Vogt, A. Iost, and J. Foct, "The effects of an Fe-Zn intermetallic-containing coating on the stress corrosion cracking behavior of a hot-dip galvanized steel," *Surface and Coatings Technology*, vol. 139, no. 2, pp. 265–271, 2001. [Online]. Available: [https://doi.org/10.1016/S0257-8972\(01\)01017-9](https://doi.org/10.1016/S0257-8972(01)01017-9)

- [35] N. Michailidis, F. Stergioudi, G. Maliaris, and A. Tsouknidas, "Influence of galvanization on the corrosion fatigue performance of high-strength steel," *Surface and Coatings Technology*, vol. 259, pp. 456–464, 2014. [Online]. Available: <https://doi.org/10.1016/j.surfcoat.2014.10.049>
- [36] K. A. Yasakau, S. Kallip, A. Lisenkov, M. G. S. Ferreira, and M. L. Zheludkevich, "Initial stages of localized corrosion at cut-edges of adhesively bonded zn and zn-al-mg galvanized steel," *Electrochimica Acta*, vol. 211, pp. 126–141, 2016. [Online]. Available: <https://doi.org/10.1016/j.electacta.2016.06.045>
- [37] E. Carrasquero, Y. Rico, and J. Minchala, "Evaluación de propiedades mecánicas en recubrimientos galvanizados por doble inmersión en caliente sobre acero al carbono," *INGENIUS*, no. 22, pp. 80–89, 2019. [Online]. Available: <https://doi.org/10.17163/ings.n22.2019.08>
- [38] K. Zhang, R. Song, and Y. Gao, "Corrosion behavior of hot-dip galvanized advanced high strength steel sheet in a simulated marine atmospheric environment," *International Journal of Electrochemical Science*, no. 14, pp. 1488–1499, 2019. [Online]. Available: <http://doi.org/10.20964/2019.02.13>



# NUMBER OF SUBCARRIER FILTER COEFFICIENTS IN GFDM SYSTEM: EFFECT ON PERFORMANCE

## NÚMERO DE COEFICIENTES DEL FILTRO DE LAS SUBPORTADORAS EN EL SISTEMA GFDM: EFECTO EN EL DESEMPEÑO

Randy Verdecia Peña<sup>1,\*</sup>, Humberto Millán Vega<sup>2</sup>

### Abstract

Generalized Frequency Division Multiplexing (GFDM) is a non-orthogonal multicarrier transmission scheme proposed for fifth (5G) and future generation wireless networks. Due to its attractive properties, it has been recently discussed as a candidate waveform for the future wireless communication systems. GFDM is introduced as a generalized form of the widely used Orthogonal Frequency Division Multiplexing (OFDM) modulation scheme and it uses only one cyclic prefix (CP) for a group of symbols. The main focus of this work is to present like impact on the system performance the coefficient quantity of the subcarrier filter. A simple method for the computation of the coefficients of the prototype filter is employed. Besides, it is presented a structure for the GFDM by taking advantage of the arrangement in the modulation matrix. We evaluated the Bit Error Rate (BER) using the receiver models presented in this work. The results showed that the BER is affected according to the coefficients quantity of the prototype filter. Based on the obtained results, the coefficients quantity has a relation with the number of time slots of the GFDM system.

**Keywords:** GFDM, number of coefficients, prototype filter, BER.

### Resumen

El GFDM (*Generalized Frequency Division Multiplexing*) es un esquema de transmisión multiportadora no ortogonal propuesta para la quinta (5G) y futura generación de redes inalámbricas. Por sus atractivas propiedades, está siendo investigada como una forma de onda a ser considerada para los futuros sistemas de redes de comunicaciones. La GFDM es introducida como una generalización del ampliamente utilizado esquema de modulación OFDM (*Orthogonal Frequency Division Multiplexing*) y usa un único prefijo cíclico (*Cyclic Prefix, CP*) para un grupo de símbolos. El objetivo principal de este trabajo es presentar cómo impacta la cantidad de coeficientes del filtro de las subportadoras en el desempeño del sistema. Se emplea un método simple para el cálculo de los coeficientes del filtro prototipo. Además, se presenta una estructura para la GFDM aprovechando la estructura de modulación matricial. Se evaluó la tasa de error de bit (*Bit Error Rate, BER*) usando los modelos de receptores presentados en este trabajo. Los resultados muestran que el BER es afectado según la cantidad de coeficientes del filtro prototipo. Basado en los resultados obtenidos, la cantidad de coeficientes tiene relación con el número de intervalos de tiempo del sistema GFDM.

**Palabras clave:** GFDM, número de coeficientes, filtro prototipo, BER.

<sup>1,\*</sup>Signals, Systems and Radiocommunications Department, Escuela Técnica Superior de Ingenieros de Telecomunicaciones (ETSIT), Universidad Politécnica de Madrid, España. Corresponding author ✉: randy.verdecia@upm.es

<http://orcid.org/0000-0003-4798-2681>

<sup>2</sup>Physical Department (Retired), Universidad de Granma – Cuba. <http://orcid.org/0000-0001-9421-7494>

Received: 25-06-2019, accepted after review: 25-11-2019

Suggested citation: Verdecia Peña, R. and Millán Vega, H. (2020). «Number of subcarrier filter coefficients in GFDM system: effect on performance». INGENIUS. N.º 23, (january-june). pp. 53-61. DOI: <https://doi.org/10.17163/ings.n23.2020.05>.

## 1. Introduction

Wireless and Mobile communication have become essential tools for the life and modern society. The future wireless networks of telecommunication need higher throughput based on very high spectral and energy efficiencies, very low latency and very high data rate. That requires a more effective physical layer (PHY) [1–3]. The core of the physical layer of fourth generation (4G) is the Orthogonal Frequency Division Multiplexing (OFDM). These systems allow high data throughput. OFDM modulation is widely adopted due to its favorable features like a simple implementation built on the Fast Fourier Transform (FFT) and robustness against fading channels [2, 4]. However, the application scenario previewed for fifth generation (5G) networks have challenges where OFDM could have limitations. The low latency needed for Vehicle to Vehicle Communications and Tactile Internet applications require a data cutoff where OFDM packet with one cyclic prefix (CP) per symbol have a low spectral efficient [1, 4–6]. The requirement of OFDM to preserve the orthogonality between individual subcarriers is essential for the machine-to machine (M2M) communication. Due to the need of low power consumption which influence the negative form on the synchronization process, this procedure is not possible by OFDM modulation [4, 7]. Other disadvantage of the OFDM system is the high out-of-band (OOB) radiation resulting from rectangular pulse shaping [8]. OFDM can fulfill the requirements of 5G in a limited way, due to these shortcomings.

In recent years, several waveforms have been proposed to overcome the above limitations of OFDM, this is the case of FBMC, UFMC, GFDM in references [9–14] are suggested many waveforms. Filter Bank Multicarrier (FBMC) the subcarriers are pulse shaped individually to reduce the OOB emissions, this is caused because the subcarriers have narrow bandwidth and the length of the transmit filter impulse response is long. The applications that to need a number of transmit of large symbols are benefit with this modulation. But it is clear, this modulation scheme is not suitable for low latency scenarios, where high efficiency must be achieved with short burst transmissions [1, 5–7]. Universal Filtered Multicarrier (UFMC) a group of subcarriers is filtered to reduce the OOB emission. A principal characteristic of this modulation is the impulse response can be short obtaining high spectral efficiency in short transmissions [1].

The disadvantage of UFMC does not require a CP, then is more sensitive to small time misalignment than CP-OFDM and might not be suitable for applications that need loose time synchronization to save energy [1, 5, 6]. In this context, the Generalized Frequency Division Multiplexing (GFDM) is one alternative multicarrier scheme that is currently under

evaluation as a candidate of the PHY layer for the next generation of mobile communication systems. It is interesting that one of the main relevance of the GFDM is that its generalized form of OFDM preserves most of the valuable properties of OFDM while addressing its limitations. The GFDM can provide a very low OOB radiation. It is more bandwidth efficient than OFDM as it uses only CP for group of symbols in its block rather than a CP per symbol as for the case of OFDM [9, 15].

The GFDM modulation is foreseen for the modulation of independent blocks where each block consists of a number of subcarriers and symbols. The data symbols belonging to the subcarriers are filters with a prototype that is circularly shifted in time and frequency domains. The subcarrier filtering results in non-orthogonal subcarriers, then inter-symbol (ISI) and inter-carrier (ICI) might arise. Filter Impulse Response (FIR) can be employed for filtering the subcarriers and this choice has a negative impact on the Bit Error Rate (BER) performance and the OOB emissions as shown in [1]. In this work, we present BER curves to compare the influence that to have the selection of the total number of coefficient's filter and is shown to exist a relationship between the number of time slots and the coefficients of the filter in the GFDM systems. It is necessary to present this aspect because performance degrades when the total coefficient is not chosen correctly.

A GFDM symbol consists of a block structure of MN samples, where each N subcarrier carries M time-slots. In a GFDM block, the overhead is kept small by adding a single CP for an entire block that contains multiple subcarriers. Thus period that benefit can be used to improve the spectral efficiency of the system. The remaining sections are organized as follows. The systems model and properties of the GFDM transmitter are presented in Section 2. Section 3 presents different receiver structures. Section 4 shows the expression of the prototype filter as obtained from the subcarrier filter coefficients. Section 5 analyzes the BER performance of GFDM including the theoretical equations assuming Zero-Forcing (ZF), Matched Filter (MF) and Matched Filter–Parallel Interference Cancellation (MF-PIC) receivers. We used the coefficients obtained in Section 4. Finally, Section 6 presents some conclusions. The main objective of this work is to present a structure for the GFDM by taking advantage of the arrangement in the modulation matrix.

**Notation:** Bold lower case is used for column vectors and bold upper case for matrices. All vectors are in column form. The vector and matrix transpose and Hermitian are indicated by the superscripts 'T' and 'H', respectively. We use  $\mathbf{W}_{MN}$  to denote the discrete Fourier transform (DFT) matrix of size MN. We also assume that  $\mathbf{W}_{MN}$  is normalized, such that  $\mathbf{W}_{MN} \mathbf{W}_{MN}^H = \mathbf{I}_{MN}$ , where  $\mathbf{I}_{MN}$  denotes the identity

matrix of size  $MN$ . Hence,  $\mathbf{W}_{MN}^H = \mathbf{W}_{MN}^{-1}$ . The terms FFT and iFFT refer to the fast implementation of DFT and inverse DFT (iDFT), respectively.

## 2. Materials and Methods

### 2.1. System model and properties of GFDM

The Generalized Frequency Division Multiplexing is a multicarrier system. The data packet in GFDM is such that only one CP per block of transmitted symbols is required [10]. Figure 1 presents the structure of a GFDM data packet. In the system GFDM the data symbols over each subcarrier are filtered through a well-localized band-pass filter with the aim of limiting the Inter-Carrier Interference (ICI) [16]. The GFDM data packet is organized in  $M$  time-slots and  $N$  subcarriers.



Figure 1. GFDM data packet.

The OFDM system can provide a high Out-Of-Band (OOB) radiation and a least bandwidth efficiency in comparison with GFDM [1, 8] due to the fact that OFDM system uses a CP per symbol as is presents in Figure 2.

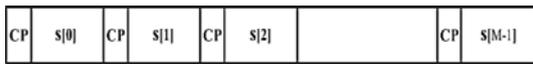


Figure 2. OFDM data packet.

Consider the block diagram of the transceiver depicted in Figure 3. A mapper, e.g QAM [7], maps the encoded bits to symbols from a  $2^\alpha$  valued complex constellation, where  $\alpha$  is the modulation order. The  $\mathbf{s}$  vector denotes a data block that contains  $MN$  symbols, which can be decomposed into  $M$  time-slots and  $N$  subcarrier each according to  $\mathbf{s} = [\mathbf{s}^T[0]\mathbf{s}^T[1], \dots, \mathbf{s}^T[M-1]]^T$  and  $\mathbf{s}[m] = [s_0[m]s_1[m], \dots, s_{N-1}[m]]^T, m = 1, 2, \dots, M-1$ .

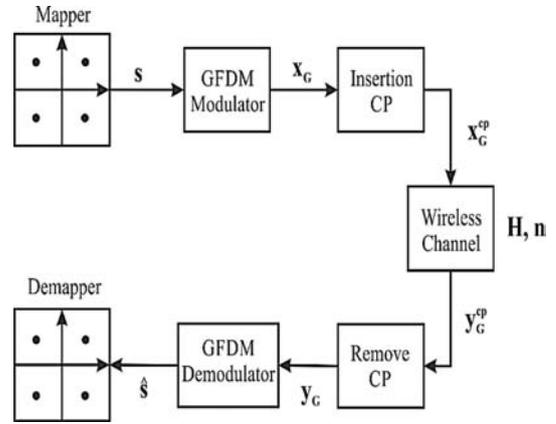


Figure 3. Block diagram of the transceiver for GFDM.

The data symbols are taken from a zero mean independent and identically distributed (i.i.d) process with unit variance. The expression that relates the input data symbols  $\mathbf{s}[m]$  and GFDM transmitter output  $\mathbf{x}[m]$ , may be expressed as [16].

$$\mathbf{x}[m] = \mathbf{W}_{MN}^{-1} \mathbf{C}_f \mathbf{s}_{ex}[m], \quad m = 0, 1, 2, \dots, M-1 \quad (1)$$

where  $\mathbf{W}_{MN}^{-1}$  is the iDFT matrix of size  $(MN \times MN)$ ,  $\mathbf{C}_f$  is circular matrix of the size  $(MN \times MN)$ , with the first column composed by the vector  $\mathbf{c} = [c_0 c_1 \dots c_{F-1} 0 \dots 0 c_{F-1} \dots c_1]^T$ . The coefficients  $c_f$  are the components of the discrete spectrum of the formatter pulse, with  $f \in (0, 1, \dots, F-1)$  and  $(1 \leq F \leq \frac{MN}{2})$  [17, 18]. It will be shown in this work that the coefficients quantity influences the GFDM system performance.  $\mathbf{s}_{ex}[m]$  is the expanded vector of the data symbols  $\mathbf{s}[m]$  that can be organized as  $\mathbf{s}_{ex}[m] = [s_0[m] \mathbf{z}_{M-1}^T s_1[m] \mathbf{z}_{M-1}^T s_2[m] \mathbf{z}_{M-1}^T \dots \mathbf{z}_{M-1}^T s_{N-1}[m]]^T$ , where  $\mathbf{z}_{M-1}$  represents the column vector of the size  $(M-1 \times 1)$  [17, 18].

The expression (1) is performed in two steps. First of all it is performed as the circular convolution of  $\mathbf{c}$  and  $\mathbf{s}_{ex}[m]$  for obtaining  $\mathbf{C}_f \mathbf{s}_{ex}[m]$ . Subsequently, it is applied an iFFT of size  $(MN \times MN)$  to the result of the first step for obtaining the vector  $\mathbf{x}[m]$  of size  $(MN \times 1)$ . It is useful to comment that the computational complexity represented in (1) is dominantly determined by an iFFT of dimension  $(MN \times MN)$ . The  $\mathbf{C}_f \mathbf{s}_{ex}[m]$  procedure can be calculated by:

$$\mathbf{C}_f \mathbf{s}_{ex}[m] = \mathbf{W}_{MN} [(\mathbf{W}_{MN}^{-1} \mathbf{c}) \odot (\mathbf{W}_{MN}^{-1} \mathbf{s}_{ex}[m])], \quad (2)$$

where  $\odot$  is an operator denoting the point-wise multiplication, and the circular convolution of the vectors  $\mathbf{c}$  and  $\mathbf{s}_{ex}[m]$  are performed through point-wise multiplication of their respective iDFTs and later it is applied applying a DFT to the result. If one considers the expressions (1) and (2) the vector  $\mathbf{x}[m]$  reduces to:

$$\mathbf{x}[m] = \mathbf{k} \odot (\mathbf{W}_{MN}^{-1} \mathbf{s}_{ex}[m]) \quad (3)$$

where  $\mathbf{k} = \mathbf{W}_{MN}^{-1}\mathbf{c}$  is the vector of the prototype filter coefficients that to influencing on the GFDM performance. The computational complexity in (3) can be reduced significantly by taken into consideration that the vector  $\mathbf{s}_{ex}[m]$  is the expanded version of the vector symbol  $\mathbf{s}[m]$ . However, the  $\mathbf{W}_{MN}^{-1}\mathbf{s}_{ex}[m]$  product can be obtained by  $M$  repetitions of  $\mathbf{W}_N^{-1}\mathbf{s}[m]$  in a column. Then the computational complexity in (3) can be calculated through an iFFT of dimension  $(N \times N)$ .

The data symbols  $\mathbf{s}$  of the GFDM packet to transmit in Figure 2 can be obtained by superposition of the  $M$  vectors  $\mathbf{x}[m]$ . We can describe this operation mathematically as:

$$\mathbf{x}_G = \sum_m^{M-1} \text{circshift}(\mathbf{x}[m], mN) \quad (4)$$

where  $\text{circshift}(\cdot)$  means downward circular shift. The packet to transmit is completed by adding the CP samples to obtain the vector  $\mathbf{x}_G^{cp}$ . An interesting model in GFDM system is a matrix model with the aim to have likeness with the OFDM system. The model presented in [17–22] and the vector  $\mathbf{x}_G$  can be expressed in matrix form as:

$$\mathbf{x}_G = \mathbf{A}\mathbf{s} \quad (5)$$

where  $\mathbf{s}$  is the column vector that contains all the data symbols of the GFDM packet of  $M$  time slots and  $N$  subcarrier as is illustrated in Figure 2.  $\mathbf{A}$  is the matrix of the GFDM system that is composed by the coefficient of the prototype filter  $c_f$  that affect the performance of the system. The  $c_f$  coefficient will be calculated in other section.

## 2.2. Receiver Implementation

The vector  $\mathbf{x}_G$  is the output of the GFDM modulator (see Figure 3),  $\mathbf{x}_G$  contains the transmitted samples that correspond to the GFDM data block  $\mathbf{s}$  of size  $(MN \times 1)$ . Finally, we added on the transmitter side a cyclic prefix of  $L_{CP}$  samples to produce  $\mathbf{x}_G^{cp}$ . After that, the signal is affected by the Additive Gaussian White Noise (AWGN),  $n \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{MN})$ , where  $\sigma_n^2$  is the noise variance. The receiver signal after CP samples removal can be expressed as:

$$\mathbf{y}_G = \mathbf{H}\mathbf{A}\mathbf{s} + \mathbf{n} \quad (6)$$

where  $\mathbf{H}$  represents the circular matrix of the channel of size  $(MN \times MN)$ . The first column is shaped by the vector  $\mathbf{h}_{ch} = [h_0 h_1 h_2 \dots h_{ch-1}]^T$  that corresponds to the impulse response of the discrete low-pass filter equivalent to the channel of size  $ch$  (completed with zeros). The circular matrix,  $\mathbf{H}$ , can be expressed as:

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & \dots & 0 \\ \vdots & h_0 & \ddots & 0 \\ h_{ch-1} & \vdots & \ddots & 0 \\ 0 & h_{ch-1} & \ddots & \vdots \\ 0 & 0 & \dots & h_0 \end{bmatrix} \quad (7)$$

From the matrix as represented by equation (6), we can use two standard GFDM receiver types, i.e. Zero Forcing (ZF) and Matched Filter (MF) receiver [4, 22, 23]. We has defined the  $\mathbf{B}$  matrix as the product of the  $\mathbf{H}$  and  $\mathbf{A}$ . Then equation (6) can be rewritten as:

$$\mathbf{y}_G = \mathbf{B}\mathbf{s} + \mathbf{n} \quad (8)$$

The equalization scheme employed in this work is presented in Figure 4. In the block scheme  $\mathbf{Q}(\cdot)$  is a function that maps each component of the transmitted signal vector to the symbol nearest to the signal constellation of the modulation employed and  $\mathbf{D}(\cdot)$  determines the minimum distance of the estimative first that is employed like a metric in the PIC detector. The switch in the figure defines the receiver employed in each state to obtain the final estimate. Here ZF and MF are linear detectors and PIC is the Parallel Interference Cancellation detector, respectively. Hence, the PIC is employed as a first estimation of the output signal of the MF block.

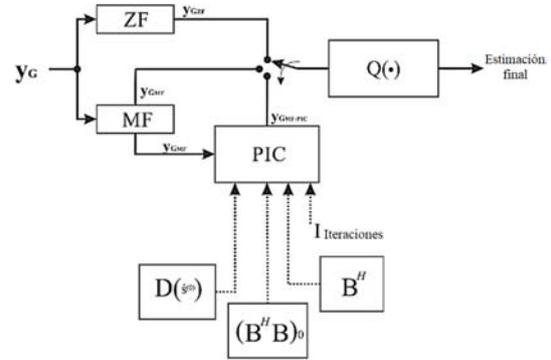


Figure 4. Block diagram of the receiver for GFDM.

The ZF receiver is characterized by the  $\mathbf{B}_{ZF}$  matrix that represents the  $\mathbf{B}$  inverse matrix. After obtaining the ZF equalization, the linear demodulation of the received signal can be expressed as:

$$\mathbf{y}_{G_{ZF}} = \mathbf{B}_{ZF}(\mathbf{B}\mathbf{s} + \mathbf{n}) = \mathbf{s} + \mathbf{n}_{ZF} \quad (9)$$

where  $\mathbf{n}_{ZF} = \mathbf{B}_{ZF}\mathbf{n}$  is the noise after the ZF equalization that affects the received signal of size  $(MN \times 1)$ .

The second type receiver, i.e. the MF, is described by the  $\mathbf{B}_{MF} = \mathbf{B}^H$  matrix. When it is applied the received vector in (8), the received signal can be expressed as:

$$\mathbf{y}_{G_{MF}} = \mathbf{B}_{MF}(\mathbf{B}\mathbf{s} + \mathbf{n}) = \mathbf{B}_{MF}\mathbf{B}\mathbf{s} + \mathbf{n}_{MF} \quad (10)$$

where  $\mathbf{n}_{MF} = \mathbf{B}_{MF}\mathbf{n}$  is the noise after the MF equalization of size  $(MN \times 1)$ .

The PIC detector implementation presents the least computational complexity as compared with other cancellation detectors as SIC [17–20]. The first estimation of the data symbols to the PIC detector is obtained as the output signal of the MF detector. This receiver can be implemented by the equations:

$$\hat{\mathbf{s}}^{(j)} = Q(\mathbf{r}^{(j-1)}), j = 1, 2, \dots, \quad (11)$$

$$\mathbf{r}^{(j-1)} = \mathbf{B}^H \mathbf{y}_G - (\mathbf{B}^H \mathbf{B})_z \hat{\mathbf{s}}^{(j-1)}, j = 1, 2, \dots, \quad (12)$$

where  $(\mathbf{B}^H \mathbf{B})_z$  corresponds to the matrix  $\mathbf{B}^H \mathbf{B}$  with zeros in the main diagonal.

Symbols estimation using equations (11) and (12) are sequentially generated up to a maximum number,  $J$ , of iterations. In the present work it was considered that the process can be interrupted after  $j$ -th iterations ( $1 \leq j \leq J$ ) depending on the quality of the generated estimates. The Maximum Likelihood (ML) metric employed here corresponds to the Minimum Distance (MD) metric. It can be computed as:

$$MD(\hat{\mathbf{s}}^{(j)}) = \|\mathbf{y}_G - \mathbf{B}\hat{\mathbf{s}}^{(j)}\|^2 \quad (13)$$

If one detects a reduction in the quality of a given estimate, that is,  $MD(\hat{\mathbf{s}}^{(j)}) \geq MD(\hat{\mathbf{s}}^{(j-1)})$ , the estimate  $MD(\hat{\mathbf{s}}^{(j-1)})$  is adopted as the final one.

### 2.3. Calculating Filter Coefficients of the Subcarriers

The filtered of the subcarriers in the GFDM modulator block presented in Figure 3 is essential to the performance of the system. In this section presented like determine its coefficients. It is presented in [16, 24–26] the prototype filter corresponding to a class of real low-pass filters whose impulse response can be express as:

$$f_p[n] = \begin{cases} c_0 + 2 \sum_{l=1}^{F-1} c_l \cos\left(\frac{2\pi ln}{P}\right), & 0 \leq n \leq P \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where  $P = FK$ , and  $c_l (0 \leq l < F)$  are real coefficients, the overlap factor  $F$  is a positive integer and  $K$  is the number of channels in the TMUX system.

The requirements for the coefficients  $c_l (0 \leq l < F)$ , after Mirabbasi and Martin [25] should meet the following conditions:

$$\begin{cases} c_0 = 1 \\ c_l^2 + c_{F-l}^2, & l = 1, 2, \dots, F/2 \end{cases} \quad (15)$$

If coefficients  $c_l$  are chosen such that expression in (15) hold, then the  $-3$  dB frequency of the prototype filter would be approximately  $\frac{\pi}{P}$ , when  $F$  is even. The minimum stopband attenuation (MSA) and the approximate rate of fall-off (ARF) of the sidelobes depend of the overlap factor  $F$  and independent of the filter order [25].

It is required to find the  $F$  coefficients  $c_l$ , and to solve a system for determining  $F$  coefficients  $c_l$ . It was obtained in [24] the auxiliary equation:

$$c_0 + 2 \sum_{l=1}^{F-1} c_l = 0 \quad (16)$$

If equation (16) is satisfied, then the side lobes of the discrete Fourier transform in equation (14) have the approximate fall-off rate of  $|\omega|^{-3}$ , with  $\omega$  defining the uniformly-spaced frequencies around the unit circle. It can be written as:

$$\omega_l = \frac{2\pi l}{P} \quad (l \text{ is an integer number}) \quad (17)$$

By using equations (15) and (16) it is possible to construct a system of equations with the same number of unknowns. Furthermore, equation (18) can be used to construct the remaining equations necessary to have a system of  $F$  equations.

$$\sum_{l=1}^{F-1} l^q c_l = 0, \quad q \geq 2 \quad (18)$$

with the above equations it is possible to obtain the values of the prototype filter coefficients for  $F = 2$ ,  $F = 3$  and  $F = 15$ . These are shown in the Table 1.

**Table 1.** Coefficients of the prototype Filter  $\mathbf{F}$

Coefficients	$F = 2$	$F = 3$	$F = 15$
$c_0$	1	1	1
$c_1$	-0,7071	-0,9114	-0,9999
$c_2$		0,4114	0,9992
$c_3$			-0,988
$c_4$			0,9435
$c_5$			-0,8797
$c_6$			0,8328
$c_7$			-0,7628
$c_8$			0,6466
$c_9$			-0,5536
$c_{10}$			0,4755
$c_{11}$			-0,3313
$c_{12}$			0,1543
$c_{13}$			-0,0412
$c_{14}$			0,0048

### 3. Results and Discussions

The simulation results along with the derived theoretical obtained expressions are presented in this section. In order to study the effect of the filter coefficients quantity of the subcarrier on the BER in the system GFDM, we have considered the case of the ZF, MF and MF-PIC receivers.

#### 3.1. BIT Error Rate Analysis

In this subsection we analyze the performance of the GFDM system in terms of BER versus  $E_b/N_0$  assuming that ZF, MF and MF-PIC are employed. The ZF is able to remove self-generation interference at the cost of introducing noise enhancement [1]. The MF-PIC receiver was the most flexible and adaptable to different configurations of the data package GFDM [17, 18] as described in Figure 1. The system parameters used for the simulations are presented in Table 2, while Table 3 shows the channel impulse response used in the BER performance evaluation. The impulse response of the multipath channel is normalized to unitary energy and the length of the CP guard band is  $G = ch$ .

**Table 2.** System Parameters

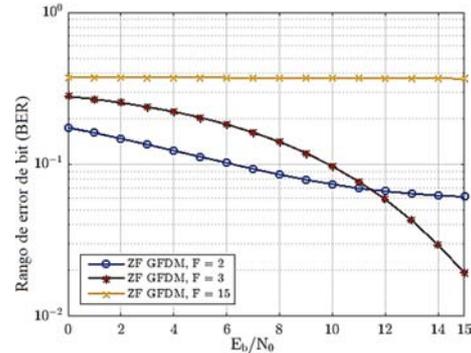
Parameters	GFDM (I)	GFDM (II)
Modulation	4-QAM	4-QAM
Channel ( $ch$ )	6	8
Times-slots $\times$ Subcarrier ( $M \times N$ )	$4 \times 64$	$16 \times 32$
Quantify of transmitted symbols	$76,8 \times 10^6$	$153,6 \times 10^6$
Quantify of filter coefficients ( $F$ )	2,3,15	-, -, 15

**Table 3.** Channel Model

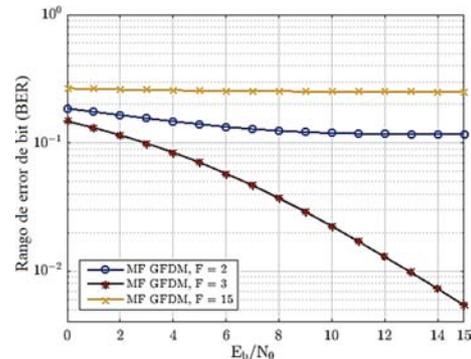
Channel	Discrete impulse response (I)	Discrete impulse response (II)
Multipath Channel	0,7774	0,6961
	0,4905	0,501
	0,3095	0,3605
	0,1953	0,2595
	0,1232	0,1867
	0,0777	0,1343
	-	0,097
	-	0,0696

Figure 5 compares the BER performance of the classical ZF in the system GFDM with different quantity of filter coefficients of the subcarriers considering the system parameters from Table 2 and the multipath channel from Table 3. The results presented in Figure 5 suggest that the system GFDM achieved the best

performance when  $F = 3$ . In this case the Bit Error Rate was in the order of  $2 \times 10^{-2}$ , when compared to the results presented for  $F = 2$  while  $F = 15$  had more than 3 dB of advantage.

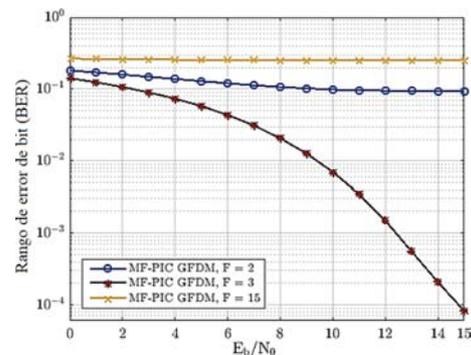


**Figure 5.** BER simulation result for ZF receiver in GFDM (I), channel I.



**Figure 6.** BER simulation result for MF receiver in GFDM (I), channel I.

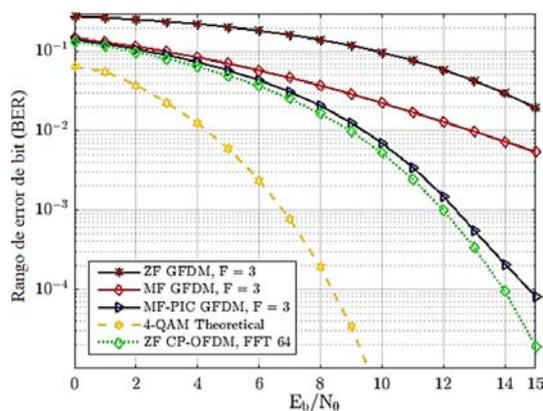
Figures 6 and 7 illustrate the BER performance for the MF and MF-PIC receivers considering the three  $F$  cases. The figures showed that performance of the system GFDM depends strongly on the quantity of coefficients of the prototype filter of the subcarriers. The case  $F = M - 1$  rendered the best choice of the quantity of filter coefficients in the system GFDM. Here,  $M$  represents the time slots of the system as depicted in Figure 2.



**Figure 7.** BER simulation result for MF-PIC receiver in GFDM (I), channel I.

The results shown in Figure 8 suggest that employing  $F = 3$  for the different receivers the MF-PIC detector had the best performance of the GFDM system. We found that the MF receiver is 4 dB more efficient than ZF with less computational complexity. The MF-PIC is more complex than ZF and MF due to the number of iterations [18]. Furthermore, in Figure 8 we present (as a comparison) two curves of the performance, a 4-QAM theoretical and other ZF CP-OFDM with 64 FFT. It is found that BER performance of the MF-PIC scheme is approximately the same with ZF CP-OFDM FFT 64, where the difference in the performance is 0,5 dB in favor of the CP-OFDM system. The cause behind this small difference is that the GFDM system is affected by the transmission matrix that depends on the coefficients quantity. However, both systems have the same computational complexity in the signal generation as they need FFT 64 but GFDM is more efficient than OFDM in terms of spectrum because the need of just only one CP to transmit a data packet of 256 symbols. On the other hand, the great difference in the BER performance of the 4-QAM is produced because it is considered as a system with AWGN.

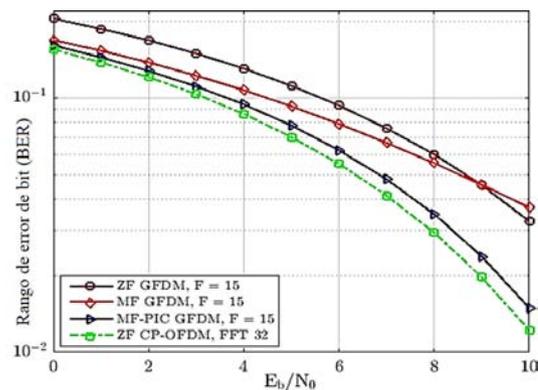
In the simulations both systems have the same computational complexity in the signal generation as they need FFT 64 but GFDM is more efficient than OFDM in terms of spectrum because the need of just only one CP to transmit a data packet of 256 symbols.



**Figure 8.** Comparison of the simulation result for ZF, MF and MF-PIC receiver in GFDM (I) with  $F=3$ , 4-QAM Theoretical and ZF CP-OFDM FFT 64, channel I.

Other simulations are presented in Figure 9 considering that the GFDM system has dimension matrices ( $512 \times 512$ ) and the properties described in Table 2 for GFDM (II). The impulse responses of the channel have 8 taps like exhibited in Table 3. Figure 9 is shown that performance of the MF-PIC scheme detector is approximately the same with ZF CP-OFDM FFT 32, where the difference in the performance is 0,5 dB in favor of the CP-OFDM system. Here, it is possible to verify again that there is an intrinsic relationship with

the total coefficients of the subcarrier filter, because the best performance that can reach the GFDM system is equal to the OFDM system. The degradation of the performance in Figure 9 in comparison with Figure 8 is given by channel effect by increase the number of taps.



**Figure 9.** Comparison of the simulation result for ZF, MF and MF-PIC receiver in GFDM (I) with  $F=15$  and ZF CP-OFDM FFT 32, channel II.

## 4. Conclusions

The expected implementation scenarios for the 5G wireless networks have challenges as the available physical layer technologies show a limited performance due to their shortcoming. The GFDM system seems a useful candidate by its rendering with the OFDM system. The key property of the GFDM system is the flexibility such that different applications can have a simple solution. This way, it is important to guarantee the coexistence with other technologies, as the current 4G.

We produced modulation and demodulation schemes for GFDM system. The presented schemes have a matrix structure that reduces the computational complexity without incurring in any performance loss penalty. By employing the matrix structure of the transmitter and receiver GFDM systems, we analyzed and compared the BER performance for the different calculated coefficients. It was shown that the BER performance of the GFDM system depends on the coefficients quantity of the filter and prototype filter. In the GFDM system, to increase the total number of the coefficient's filter not improve the performance in the GFDM system. The coefficient total depends on the number of subcarriers because it might filter symbols of others packets and generates interference. The performance of the system is conditional on accurate the coefficient total.

## References

- [1] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3045–3061, Sep. 2014. [Online]. Available: <https://doi.org/10.1109/TCOMM.2014.2345566>
- [2] E. Öztürk, E. Basar, and H. A. Çirpan, "Generalized frequency division multiplexing with flexible index modulation," *IEEE Access*, vol. 5, pp. 24 727–24 746, 2017. [Online]. Available: <https://doi.org/10.1109/ACCESS.2017.2768401>
- [3] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. T. Brink, I. Gaspar, N. Michailow, A. Festag, L. Mendes, N. Casiau, D. Ktenas, M. Dryjanski, S. Pietrzyk, B. Eged, P. Vago, and F. Wiedmann, "5gnow: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, February 2014. [Online]. Available: <https://doi.org/10.1109/MCOM.2014.6736749>
- [4] L. Sendrei and S. Marchevský, "On the performance of gfdm systems undergoing nonlinear amplification," *Acta Electrotechnica et Informatica*, vol. 15, no. 1, pp. 9–14, 2015. [Online]. Available: <http://doi.org/10.15546/aei-2015-0002>
- [5] S. K. Bandari, V. M. Vakamulla, and A. Drosopoulos, "Gfdm/oqam performance analysis under nakagami fading channels," *Physical Communication*, vol. 26, pp. 162–169, 2018. [Online]. Available: <https://doi.org/10.1016/j.phycom.2017.12.008>
- [6] F. Schaich and T. Wild, "Waveform contenders for 5g – ofdm vs. fbmc vs. ufmc," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, May 2014, pp. 457–460. [Online]. Available: <https://doi.org/10.1109/ISCCSP.2014.6877912> <https://doi.org/10.1109/ISCCSP.2014.6877912>
- [7] A. N. Ibrahim and M. F. L. Abdullah, "The potential of fbmc over ofdm for the future 5g mobile communication technology," *AIP Conference Proceedings*, vol. 1883, no. 1, p. 020001, 2017. [Online]. Available: <https://doi.org/10.1063/1.5002019>
- [8] E. Öztürk, E. Basar, and H. A. Çirpan, "Spatial modulation gfdm: A low complexity mimo-gfdm system for 5g wireless networks," in *2016 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2016, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/BlackSeaCom.2016.7901544>
- [9] A. Farhang, N. Marchetti, and L. E. Doyle, "Low-complexity modem design for gfdm," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1507–1518, March 2016. [Online]. Available: <https://doi.org/10.1109/TSP.2015.2502546>
- [10] A. M. Tonello and M. Girotto, "Cyclic block fnt modulation for broadband power line communications," in *2013 IEEE 17th International Symposium on Power Line Communications and Its Applications*, March 2013, pp. 247–251. [Online]. Available: <https://doi.org/10.1109/ISPLC.2013.6525858>
- [11] G. Fettweis, M. Krondorf, and S. Bittner, "Gfdm - generalized frequency division multiplexing," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, April 2009, pp. 1–4. [Online]. Available: <https://doi.org/10.1109/VETECS.2009.5073571>
- [12] H. Lin and P. Siohan, "An advanced multi-carrier modulation for future radio systems," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 8097–8101. [Online]. Available: <https://doi.org/10.1109/ICASSP.2014.6855178>
- [13] M. Renfors, J. Yli-Kaakinen, and F. J. Harris, "Analysis and design of efficient and flexible fast-convolution based multirate filter banks," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3768–3783, Aug 2014. [Online]. Available: <https://doi.org/10.1109/TSP.2014.2330331>
- [14] A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Filter bank multicarrier for massive mimo," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Sep. 2014, pp. 1–7. [Online]. Available: <https://doi.org/10.1109/VTCSFall.2014.6965986>
- [15] R. Datta and G. Fettweis, "Improved aclr by cancellation carrier insertion in gfdm based cognitive radios," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/VTCSpring.2014.7022943>
- [16] B. Farhang-Boroujeny and H. Moradi, "Derivation of gfdm based on ofdm principles," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 2680–2685. [Online]. Available: <https://doi.org/10.1109/ICC.2015.7248730>
- [17] R. Verdecia Peña, R. Pereira David, and R. Sampaio-Neto, "Detecção de sinais e estimação de canal em sistemas gfdm," in *XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinal SBrT2019, At: Petrópolis, RJ*, 2019. [Online]. Available: <https://bit.ly/34VPa17>

- [18] R. Verdecia Peña, “Análise espectral, detecção de sinais e estimação de canal em sistemas GFDM,” Master’s thesis, 2019. [Online]. Available: <https://bit.ly/34LM96N>
- [19] J. P. Mayoral Arteaga, “Detecção de sinais em sistemas com transmissão gfdm,” Master’s thesis, 2017. [Online]. Available: <https://bit.ly/2DEbbc2>
- [20] J. P. Mayoral Arteaga, R. Pereira David, and R. Sampaio Neto, “Simultaneous detection and parallel interference cancellation in GFDM for 5G,” in *XXXV Simposio brasileiro de telecomunicações e processamento de sinais - SBRT2017, 3-6 de setembro de 2017, São Pedro, SP*, 2017, pp. 220–224. [Online]. Available: <https://bit.ly/387wqRF>
- [21] R. Verdecia Peña, “Análisis del desempeño de los esquemas de modulación BPSK y QPSK para diferentes condiciones de canales en sistema GFDM,” *MASKAY*, vol. 8, no. 1, pp. 7–112, 2018. [Online]. Available: <http://dx.doi.org/10.24133/maskay.v8i1.506>
- [22] —, “Desempeño de los métodos de detección de señales con modulación QPSK en sistema GFDM para 5G.” *Revista Cubana de Ciencias Informáticas*, vol. 12, pp. 104–120, 09 2018. [Online]. Available: <https://bit.ly/33GWbo9>
- [23] N. Michailow, R. Datta, S. Krone, M. Lentmaier, and G. Fettweis, “Generalized frequency division multiplexing: A flexible multi-carrier modulation scheme for 5th generation cellular networks,” 2012. [Online]. Available: <https://bit.ly/2rQep9B>
- [24] N. Michailow, S. Krone, M. Lentmaier, and G. Fettweis, “Bit error rate performance of generalized frequency division multiplexing,” in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2012, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/VTCFall.2012.6399305>
- [25] S. Mirabbasi and K. Martin, “Overlapped complex-modulated transmultiplexer filters with simplified design and superior stopbands,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 50, no. 8, pp. 456–469, Aug 2003. [Online]. Available: <https://doi.org/10.1109/TCSII.2003.813592>
- [26] K. W. Martin, “Small side-lobe filter design for multitone data-communication applications,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no. 8, pp. 1155–1161, Aug 1998.



# MANUFACTURING OF COMPOSITE MATERIALS WITH HIGH ENVIRONMENTAL EFFICIENCY USING EPOXY RESIN OF RENEWABLE ORIGIN AND PERMEABLE LIGHT CORES FOR VACUUM-ASSISTED INFUSION MOLDING

# FABRICACIÓN DE MATERIALES COMPUESTOS DE ALTO RENDIMIENTO MEDIOAMBIENTAL CON RESINA EPOXI DE ORIGEN RENOVABLE Y NÚCLEOS LIGEROS PERMEABLES PARA INFUSIÓN ASISTIDA POR VACÍO

Diego Lascano<sup>1,2</sup>, Jorge Valcárcel<sup>1</sup>, Rafael Balart<sup>1</sup>, Luís Quiles-Carrillo<sup>1</sup>,  
Teodomiro Boronat<sup>1</sup>

## Abstract

This work focuses on the manufacturing and characterization of novel and lightweight hybrid sandwich-type structures, using different stacking sequences of flax and basalt fabrics as reinforcement fibers, both of them previously silanized. To reduce the overall weight and facilitate the manufacturing process, a polyester non-woven core, was used which, besides reducing the weight of the composite it also acts as a media to spread the resin. These composites were manufactured with a partially bio-based epoxy resin with a reactive diluent derived from epoxidized vegetable oils that contributes to a 31 % of biobased content. The hybrid composites were obtained by vacuum-assisted resin infusion moulding (VARIM), where the core was used as a media to spread the resin. The mechanical properties were evaluated in flexural and impact conditions.

## Resumen

Este trabajo se centra en la fabricación y caracterización de nuevos materiales tipo sándwich híbridos de bajo peso, con diferentes configuraciones de apilamiento de refuerzo de basalto y lino tratadas previamente con silanos. Para aligerar el peso y facilitar la fabricación, se empleó un núcleo de poliéster no tejido que, además de aligerar el peso del compuesto también actuó como medio de difusión de la resina. Se empleó una resina epoxi de origen parcialmente renovable con un diluyente reactivo derivado de aceites vegetales epoxidados que contribuye a un 31 % de origen renovable. Los compuestos híbridos se fabricaron mediante moldeado por infusión de resina asistida por vacío (VARIM), donde el núcleo se utilizó como medio de infusión de la resina. Las propiedades mecánicas se evaluaron en condiciones de impacto y de flexión.

<sup>1,\*</sup>Technological Institute of Materials (ITM), Universitat Politècnica de València (UPV), Plaza Ferrándiz y Carbonell 1, 03801 Alcoy, España. Corresponding author ✉: dielas@epsa.upv.es <http://orcid.org/0000-0002-0996-1946>

<http://orcid.org/0000-0000-0000-0000>, <http://orcid.org/0000-0001-5670-7126>

<http://orcid.org/0000-0001-8037-2215>, <http://orcid.org/0000-0002-2144-2874>

<sup>2</sup>Escuela Politécnica Nacional, 17-01-2759, Quito, Ecuador

Received: 07-11-2019, accepted after review: 09-12-2019

Suggested citation: Lascano, D.; Valcácer, J.; Balart, R.; Quiles-Carrillo, L. and Boronat, T. (2020). «Manufacturing of composite materials with high environmental efficiency using epoxy resin of renewable origin and permeable light cores for vacuum-assisted infusion molding». INGENIUS. N.º 23, (january-june). pp. 62-73. DOI: <https://doi.org/10.17163/ings.n23.2020.06>.

The interactions in the fiber-matrix interface were studied through field emission scanning electron microscopy (FESEM). The obtained data revealed that the silane (coupling agent) treatment works better on basalt fibers than on flax fibers, resulting in superior flexural properties on structures where these fibers are present. It is noteworthy to mention that the stacking sequence of plies directly influences the flexural properties, but it does not significantly affect the energy absorbed when these composites work on impact conditions.

**Keywords:** Hybrid composite materials; non-woven cores; silane coupling agents; VARIM process, basalt fibers; flax fibers.

La interacción en la interfaz fibra-matriz se evaluó por medio de microscopía electrónica de barrido de emisión de campo (FESEM). Los datos revelaron que el tratamiento de silanos funciona mejor en las fibras de basalto que en las fibras de lino, resultando en propiedades a flexión superiores en las estructuras donde estas fibras están presentes. Cabe mencionar que la distribución de apilamiento influye directamente en las propiedades a flexión, pero no afecta en la absorción de energía en condiciones de impacto.

**Palabras clave:** materiales compuestos híbridos, núcleo no tejidos, agentes de acoplamiento de silano, VARIM, fibras de basalto, fibras de lino.

## 1. Introduction

The use of composite polymeric materials has increased considerably in our lives not only in technical applications but in everyday applications. This is due to their ability to tailor the desired properties [1]. Thermosetting matrices such as unsaturated polyester resins (UP), epoxies (EP), phenolics (PF) [2–4], among others, and reinforcing fibers such as carbon (CF), aramid (AF), and glass (GF) are the most used due to the high mechanical properties they can provide together with very low weights that make them suitable for automotive, sports, ballistic, civil construction applications among others [5–9]. Despite this, the conventional high-performance fibers offer some disadvantages. Among others, a critical drawback of these fibers is related to the production costs and conditions as, for example, CF requires extremely high temperatures to obtain highly purified fibers. On the other hand, these conventional reinforcing fibers have a substantial environmental impact, related to the production stages and the problematics after their useful life; this means that their mass production is practically unfeasible [10, 11].

Matrices based on epoxy resins are the most used, due to excellent mechanical, thermal, and coating properties that provide to composite materials. This is generally achieved by the functional groups that are found in its structure, namely epoxy/oxirane rings which can polymerize to form 3D net structures. Most of the currently available epoxy resins are based on diglycidyl ether of bisphenol A (DGEBA), which comes from the reaction of epichlorohydrin and bisphenol A (BPA) [12]. These components are petroleum-derived materials and, like synthetic fibers have a considerable impact on the increase in the carbon footprint. In order to develop composite materials with low or restricted environmental impact, research studies have been carried out in both components, i.e. the thermosetting matrix and the reinforcing fibers [13]. Traditional epoxies are bioderived as above-mentioned, but in the last decades, new epoxies have been industrially manufactured from renewable resources thus contributing to lowering the overall carbon footprint

This type of resins is characterized by the fact that part of its content is obtained from renewable resources, unlike conventional resins obtained from petroleum derivatives. Vegetable oils (VOs) have gained high relevance due to their potential in manufacturing bio-based materials [14, 15]. The basic structure of these vegetable oils is based on a triglyceride structure composed of three different fatty acids chemically attached to a glycerol skeleton. Some of these fatty acids, such as oleic, linoleic, and linolenic acids, contain one, two and three unsaturated carbons, respectively, and this allows selective modification of VOs with different chemical functionalities. VO's can be epoxidized, maleinized, hydroxylated, acrylated, and so on to give partially

bio-based resins. These “eco-friendly” resins have been developed in such a way that their mechanical and thermal properties are similar to their petrochemical counterparts so that they can compete with petroleum-derived thermosetting resins with the additional feature of the renewable origin [13].

An alternative to the use of synthetic and rock-wool fibers in the manufacturing of high-performance composite materials is to replace fully or partially some of these fibers and/or fabrics with natural fibers. Several researchers have focused their research on developing hybrid composite materials with conventional and natural fibers. Natural fibers such as hemp, flax, palm leaf are the most used and offer interesting, balanced properties [16, 17]. One fiber that stands out for its excellent mechanical properties, due to its composition and structure is flax [9]. Nevertheless, all-flax composites cannot compete with high-performance composite materials. That is why the manufacturing of hybrid composite structures has become an interesting alternative to find a balance between environmental concerns and technical properties.

As natural fibers cannot compete with high-tech fibers such as carbon or aramids, it is quite common to combine natural fibers with mineral fibers, which are cost-effective compared to high tech fibers. Basalt fibers have been used in thermoplastics and thermosetting composite materials as an alternative to glass fiber. These fibers are obtained from basalt rocks, which is one of the most abundant on the earth's surface. Its structure is quite similar to glass fiber but with different silica and alumina content. Unlike glass fiber, the production process of basalt fiber is simpler, more efficient and does not generate waste due to the simple structure. Barouni and Dhakal [18] have developed flax/glass hybrid composites with improvement in the impact damage characteristics due to the capability to absorb the impact energy of flax fibers. On the other hand, Mazur *et al.* [19] studied the hybrid effect of composites based on basalt and carbon fibers, where an increase in the strength and tensile modulus was reported when the shared mass of the fibers was about 7 wt%.

Among the hybrid composite materials, it is worthy to note the interest on sandwich panels. These have been used in areas where larger thicknesses, lightweights, high rigidity, and insulation capacity are required. They are generally used in the construction, automotive, aircraft areas, and so on, due to their extraordinary ability to support flexural (out-of-the-plane) forces [20]. Sandwich panels are composed of two outer sheets (face sheets) and a core that generally has lower properties than the face sheets but it contributes to support the out-of-plane forces by shear with the outer sheets.

The face sheets are usually composite laminates with several layers with fabrics oriented in different di-

rections to improve the isotropy behaviour [21]. These face sheets often consist of unsaturated polyester (UP) sheets with glass fibers (GF) for applications with moderate mechanical performance, or epoxy (EP) plies with carbon fibers (CF) and/or aramid (AF) for applications with high mechanical responsibility [22, 23]. The core is usually made of materials with low density and relatively poor mechanical characteristics, since their purpose is not to support mechanical resistance. Core materials offer high lightness and a very good capacity of shear stresses absorption. Some of the widely used cores include balsa wood, polyurethane, or honeycomb structures [24–26]. Manufacturing of composite sandwich panels usually is carried out by hand layup and vacuum bagging.

In the recent years some research has been focused on the development of cores which are not only a part of the sandwich panel but also, can contribute to improve manufacturing processes. Vacuum assisted resin infusion moulding (VARIM) is a very widely used technique alternative to the industrial resin transfer moulding (RTM) to obtain composite materials with excellent balanced properties in a simple way. Nevertheless, the VARIM process requires different consumables (infusion nozzle, bleeding web, spiral tube, and so on), and it is not the best manufacturing process for sandwich panels since the typical cores do not transfer the infused resin. Recently, new cores have been developed with multifunctional features.

On the one hand, they can act as core materials in sandwich structures, but on the other hand, they offer enough porosity to enhance resin infusion through it. Some of these cores are composed of a cell structure separated by channels with synthetic microspheres, which do not absorb resin; this channel helps the resin flow through the composite material [27]. Chatys *et al.* [28] have used Lantor Soric® in carbon fiber sandwich composites to manufacture car safety bumpers due to its flexibility and the ability to absorb impact energy compared to conventional metal parts. Eum *et al.* [29] have reported that using the core as an infusion media in a conventional VARTM process, it is possible to provide similar properties to similar composites since this does not catch resin and helps to the correct its flow.

Hybrid composite materials, due to the diverse nature of their components, do not usually have the required synergy, since the stresses are not correctly transferred to the reinforcement fibers, which are generally the most resistant material. In general, to improve the interface interaction between fibers and the resin, several methods have been developed. The most used method is chemical treatment with tailored silanes. Silanes offer dual functionality, thus allowing them to react firstly with the fiber surface, and during the crosslinking, they can react with the resin, thus leading to real chemical bridges between the thermosetting

matrix and the reinforcement fibers [30, 31].

The objective of this work is the development of novel, highly lightweight sandwich structures, using basalt and flax fibers as reinforcement fibers. As a core material, a polyester non-woven with hexagonal cells has been used to assess the potential of manufacturing high-performance composite sandwich panels by using a conventional vacuum assisted resin infusion moulding (VARIM) using the core as infusion media. The obtained sandwich panels have been characterized in flexural and impact conditions to understand the strength properties and the effect of silane treatment. FESEM has been carried out to analyze the interface interaction between the reinforcement fibers and the partially bio-based epoxy resin.

## 2. Experimental

### 2.1. Materials

A partially bio-based (31% renewable content according to with ASTM D6866-12) commercial-grade epoxy resin was used as the thermosetting matrix. This consisted on a base epoxy resin Resoltech® 1070 ECO and an amine-based hardener grade Resoltech® 1074 ECO from Castro Composites (Pontevedra, Spain). The resin to hardener ratio was 100/35 parts by weight as recommended by the manufacturer. The system was deeply mixed by manual methods until a uniform mix was obtained.

Two different reinforcement fabrics were used. Basalt fabric BAS 940.1270 from Basaltex® (Wevelgem, Belgium) made of 100% continuous basalt filaments. This fabric shows a specific surface weight of 940 g/cm<sup>2</sup> and a thickness of 0.54 mm. Flax fabrics Biotex® Flax were obtained from Composites evolution (Chesterfield, United Kingdom) with a specific surface weight of 400 g/cm<sup>2</sup> with a thickness of 0.7 mm.

The core material was a nonwoven Lantor Soric® XF supplied by LANTOR® (Veenendaal, The Netherlands) with a specific surface weight of 250 g/cm<sup>2</sup>. This was used as core material and infusion media. To improve the interface interaction between the selected fibers/fabrics and the epoxy matrix, a glycidyl-functional silane (3-glycidylloxypropyl) trimethoxysilane was used as a coupling agent. This was supplied by Sigma- Aldrich (Madrid, Spain).

#### 2.1.1. Pre-treatment of fabrics

Generally, basalt fibers are coated by a silane-based sizing to impart strands that can interfere with the panels production process. To remove it and any external impurities, basalt fibers were initially subjected to a thermal treatment at 300 °C for 3 h.

The interface interaction between the fabrics and the matrix was enhanced by a silanization treatment. Both basalt and flax fibers were immersed in an aqueous solution with 1 wt% silane for 2 h at room temperature; then the solution was stirred with a magnetic stirrer to obtain a uniform solution. During this stage, hydrolysis of silane occurs and the subsequent hydroxyl groups move to the surface of the fabrics.

To complete the silanization process of chemical anchoring of the silanes through condensation with the hydroxyls on the surface of both fibers, an air circulating oven was used to dry the functionalized fabrics for 12 h at 80 °C. This stage provides strong links between the hydrolyzed silane groups and the hydroxyl group in both basalt and flax fabrics, through a condensation process with the release of water that is removed by evaporation.

### 2.1.2. Manufacturing of hybrid basalt/flax composite laminates

The process used to manufacture the hybrid sandwich laminates was the vacuum assisted resin infusion moulding (VARIM), instead of using the usual process with bleeding fabric, absorption mesh, and so on, which are not the best selection to manufacture composite sandwich panels.

Sandwich panels were obtained with Soric® core that acted as porous media for infusion. As in other infusion methods, the vacuum is responsible for spreading the resin throughout the geometry of the composite sandwich and avoids agglomeration of resin in fabrics.

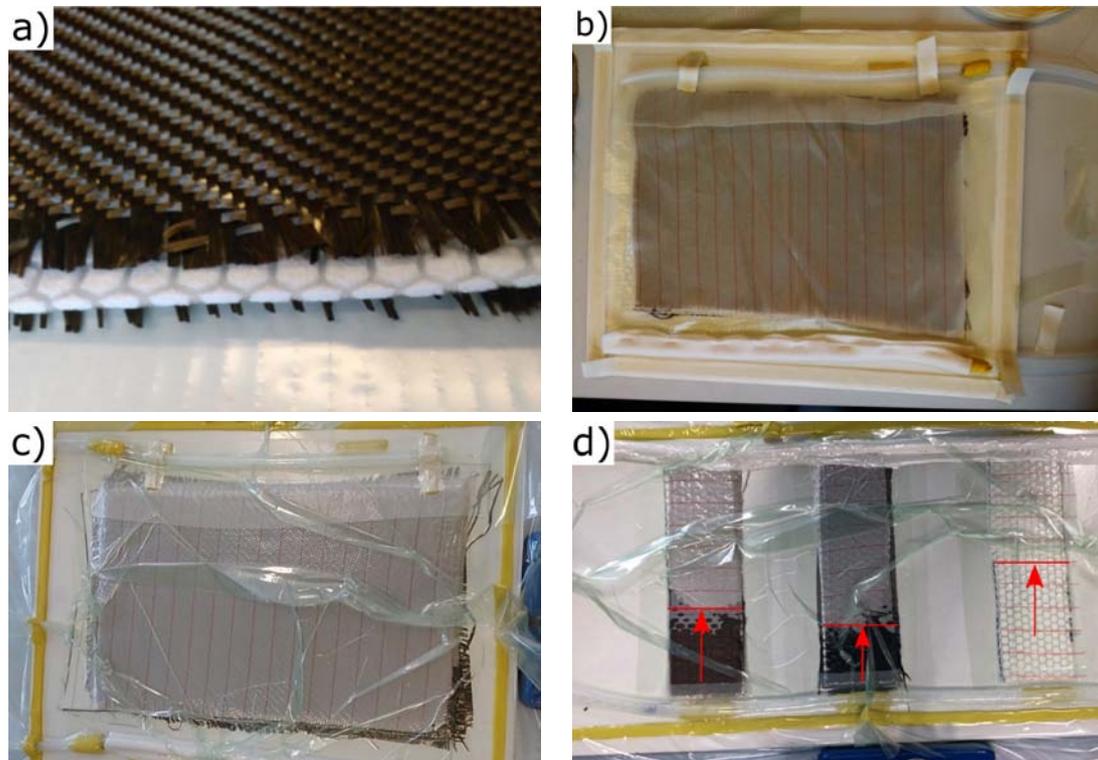
Different stacking sequences were manufactured, as can be seen in Table 1. The procedure was the follow-

ing. First, a flat surface was cleaned beforehand, then coated by a thin layer of a release agent, poly(vinyl alcohol) - PVA. Second, a peel-ply sheet was placed on the PVA thin layer to make more accessible the unmolding process. Then the fabrics and the core were stacked as indicated in Table 1 and Figure 1a). From this step, the process is slightly different from a conventional resin infusion (VARIM). As can be seen in Figure 1b), the bleeding mesh instead of covering the whole stacking, it is only placed on one side where the resin inlet will be placed. This is responsible for ensuring the flow of the resin from the resin inlet to the core. The resin distribution tube was placed over the mesh meanwhile, the vacuum tube was coated with a felt sheet that allowed the vacuum but restricted the resin flow. Finally, all the elements were sealed with a plastic bag and a double side sealing tape (Figure 1c). To ensure no leaking resin, the vacuum was tested beforehand. Then, the resin was allowed to flow until all the sheets were completely soaked. After this, the resin supply was cut off, but the vacuum was maintained for 8 h until the resin was completely cured at room temperature. With these curing conditions, no additional post-curing process is needed. As can be seen in Figure 1d) the core helps the resin to flow through the sandwich panel. Flax fibers, due to their porous structure, were the ones that better absorb the resin, followed by the basalt fiber. The Soric® Lantor core helps the resin to be spread homogeneously through the face sheets, thus leading to a complete wetting of the composite panel.

The obtained panels were machined by a computer numerical control milling machine to obtain specimens following international standards guidelines

**Table 1.** Composition and coding of basalt/flax/sandwich composite panels

Code	Ply number ratio (basalt/flax)	Sandwich upper face	Core	Sandwich bottom base	Volume fraction resin/material
BBSBB	4/0	basalt – basalt	Soric® XF	basalt – basalt	52.20/47.80
BLSLB	2-feb	basalt – flax	Soric® XF	flax – basalt	62.62/37.38
LBSBL	2-feb	flax – basalt	Soric® XF	basalt – flax	63.10/36.90
LLSLL	0/4	flax – flax	Soric® XF	flax – flax	74.76/25.24
BSB	2/0	basalt	Soric® XF	basalt	58.86/41.14
LSL	0/2	flax	Soric® XF	flax	77.64/22.36



**Figure 1.** Scheme of the manufacturing process of basalt/flax/sandwich composites by vacuum-assisted resin infusion moulding (VARIM) with the nonwoven core as infusion media. a) stacking configuration, b) placement of resin inlet and outlet tubes, c) sealing process and vacuum test procedure, d) resin infusion process with different stacking sequences.

## 2.2. Mechanical properties

The mechanical properties of sandwich panels were determined in flexural and impact conditions. Flexural tests were carried out on a universal testing machine ELIB 30 from S.A. E. Ibertest (Madrid, Spain). The flexural test was performed following the ISO 14125:1998 standard. In this test, the specimen is supported on two points separated from each other, and the increasing load applied in the center with a crosshead rate of 1 mm/min. The machine was equipped with a cell load of 5 kN.

To evaluate the impact strength of the sandwich panels the Charpy test was carried out in a Charpy pendulum supplied by Metrotec (San Sebastián, Spain), using a 6-J pendulum on “U” type notched samples (radius of 0.5 mm and 2 mm depth) following the guidelines of ISO 179 standard. At least five samples of every material were tested at room temperature; the results of all tests were collected and averaged. In addition, the standard deviation was obtained to estimate the error.

## 2.3. Interface interaction analysis by field emission scanning electron microscopy (FESEM)

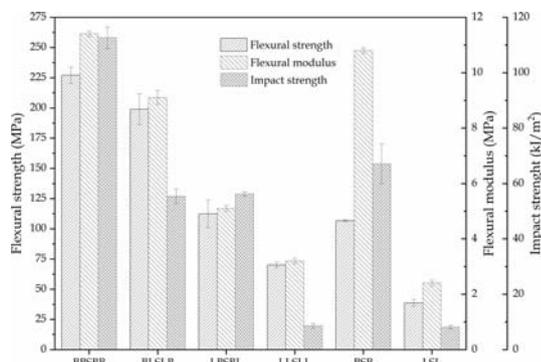
To evaluate the interface interaction between the reinforcing fibers and the epoxy matrix, samples were cryofractured and then, observed by field emission scanning electron microscopy (FESEM), in a ZEISS ULTRA 55 FESEM microscope supplied by Oxford Instruments (Abingdon, United Kingdom) working at an acceleration voltage of 2 kV. To provide electrical conducting properties to the sandwich panels, they were precoated with a gold-palladium layer using a high vacuum sputter coater EM MED20 supplied by Leica Microsystem (Milton Keynes, United Kingdom).

## 3. Results and discussion

### 3.1. Mechanical properties of sandwich panels based on basalt and flax fibers

Subjecting basalt/flax sandwich panels to flexural and impact conditions (Charpy test) gives interesting data about resistant properties as well as the ability to absorb energy in impact conditions. Figure 2 shows the values obtained from the two described tests. In sandwich materials, the core does not provide any mechanical strength to the material (in tensile conditions),

so final properties are strictly based on the reinforcing fibers. Nevertheless, core materials contribute to support out-of-plane stresses by shear with the face sheets.



**Figure 2.** Mechanical properties of basalt/flax sandwich hybrid composite with different stacking sequences obtained from flexural and impact tests.

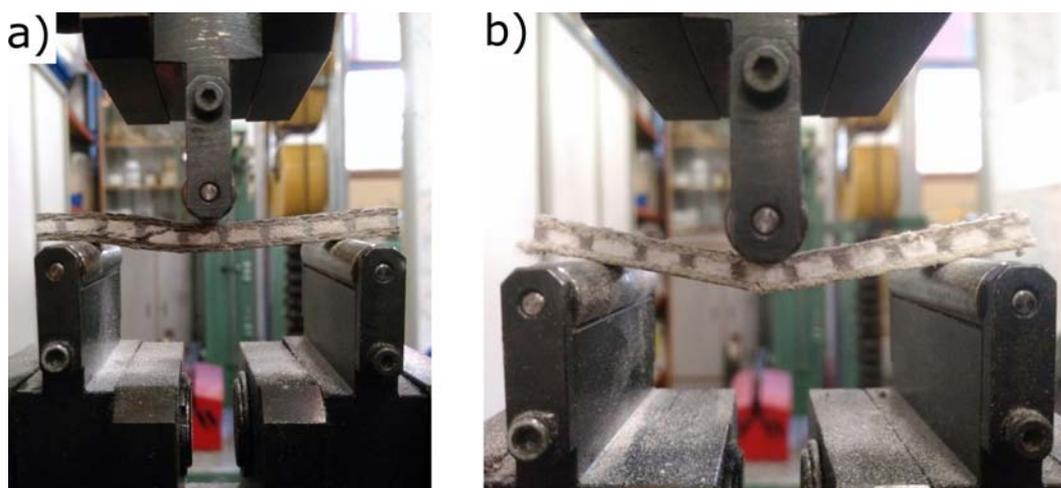
As expected, the BBSBB material with the highest number of basalt plies in its structure (4 basalt sheets, two at the top of the core and two at the bottom), shows a flexural strength,  $\sigma_f$ , and flexural modulus,  $E_f$  of  $227 \pm 6.79$  MPa and  $11.35 \pm 0.037$  GPa, respectively. These values are similar to other glass-fiber composite materials. The obtained values are relatively high, which means that besides the fact that basalt fibers have high resistance, the interaction of these fibers and the epoxy matrix is quite good, allowing excellent load transfer between the matrix to the fibers. These results are in accordance with the literature about basalt/epoxy systems [32].

Analyzing materials such as BLSLB and LBSBL, which have the same number of basalt and flax fibers in the face sheets (one basalt ply and one flax ply on

each face sheet), but with different stacking configuration, gives impressive results. The panel with the basalt sheets in the outer side of the face sheets offers higher values of flexural strength and modulus being only 12% and 19% less, respectively, compared with the material with four basalt plies. BLSLB stacking sequence leads to interesting mechanical properties of  $199 \pm 12.7$  MPa and  $9.14 \pm 0.25$  GPa for the flexural strength and modulus, respectively.

Concerning the LBSBL material, it is worthy to remark a notorious decrease in its flexural properties both in strength and stiffness, having a decrease of 43% and 44% in the flexural strength and the flexural modulus, respectively, compared with the BBSBB composite panel. This is clearly due to the nature of the fibers that are supporting the tensile and compression stresses. In the LBSBL composite panel, flax fibers are located at the outer face. As flax fibers are less resistant than basalt fibers, the result is a noticeable decrease in its properties. Similar results were obtained by Dhakal *et al.* [33] in hybrid reinforced composites based on basalt/hemp fibers, the presence of basalt fibers in the outside of the face sheets improved the flexural strength and modulus. As expected, panels composed entirely by flax fibers are those with the more inferior flexural properties.

Figure 3 shows an example of the type of failure suffered by the panels. Figure 3a) shows the BSB material that is composed only of a basalt ply in each of the faces, as mentioned before the core used, does not provide any extra strength to the final material. It can be seen that the failure of the material is caused by the core and not by the reinforcing fibers. On the other hand, the failure of the LSL material in figure 3b) indicates this failure is due to the low resistance that the flax fibers in the bottom face sheet, which are working in tensile conditions.



**Figure 3.** Forms of failure in flexion tests of a) BSB sandwich panel and b) LSL sandwich panel.

By using the Charpy impact test, one can estimate the energy that such materials can withstand under impact conditions, thus giving an estimation of the toughness.

As expected just as it happened in flexural characterization, the material composed entirely of basalt fibers (and a Soric® Lantor core), i.e. BBSBB is the one that absorbs the highest energy, obtaining values of  $112.6 \pm 3.9$  kJ/m<sup>2</sup>, which is indicating good stress transfer between the matrix to the reinforcing fabrics.

Basalt fibers, due to their composition, provide high stiffness to the composite but, in contrast, basalt composites cannot support large deformations [34]. By analyzing the composite panels with the same number of basalt and flax fibers (BLSLB and LBSBL), it can be seen that the distribution of the fibers does not affect the impact energy (see Figure 2), with absorbed-energy values of 55 – 56 kJ/m<sup>2</sup>, that are in accordance with the results reported by Fiore *et al.* [35]. They observed very slight changes in the impact strength by placing basalt or flax fibers in the outer face sheet.

As expected, just by replacing one-ply from basalt to flax, it leads to a decrease in the impact strength of about 50%. This is consistent because if one observes the BSB material made up by only one basalt layer in the face sheets, it has higher impact strength than materials that have flax fiber in their stacking sequence. It can be concluded that the impact stress is absorbed almost entirely by basalt fibers, since analyzing materials that are made entirely of flax fibers, either by four layers (LLSLL) or two layers (LSL) they have very low values and practically the same, around 8 kJ/m<sup>2</sup>. As mentioned above, the core does not improve the final properties of composite materials.

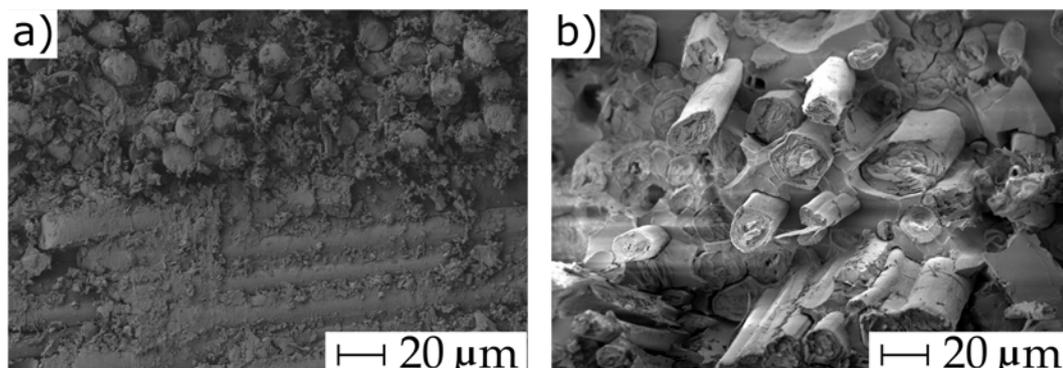
### 3.2. Matrix/reinforcement fiber interface interaction

The final properties of the hybrid composite materials are related by the strength of the fibers; therefore, a good synergy between them and the matrix is necessary since a functional interaction between the fibers

and the matrix will result in superior mechanical properties. This is because good fiber-matrix interactions allow stresses transfer from the matrix (with no reinforcing properties) to the reinforcing fiber [36]. So, a poor synergy between these components would result in composite materials with low mechanical properties.

For an accurate analysis of the interaction between the reinforcement fibers and the surrounding matrix, cryofractured surfaces from the face sheets were observed by FESEM (see Figure 4). Figure 4a) shows the interaction of the basalt fibers after the coupling agent treatment based on silane and the surrounding matrix. As can be seen, there is a lack of the gap between the fiber and the surrounding resin, therefore indicating that the silane treatment was successful. Similar findings were reported by Gao *et al.* [37], in silica-based fibers (glass fiber) and an epoxy matrix. They show the effectiveness of a silanization treatment on fibers which is directly related to improved mechanical properties. Figure 4b) shows the flax fibers after being subjected to the same chemical treatment with the silane coupling agent. In this case, it is shown that the interaction of the fibers was improved since the gap is very small. This improvement has also been reported by Sepe and Caputo [38], with hemp fibers subjected to a silanization process with (3-glycidyloxypropyl) trimethoxysilane as a coupling agent.

This coupling agent also decreases the hydrophilic behaviour that natural fibers intrinsically have, and this enhances the compatibility with hydrophobic polymeric matrices. Despite this, the presence of the gap does not allow a perfect stress transfer so that the final mechanical properties of panels will decline. Because the basalt fibers structure is based on silica, the effectiveness of the silanization with hydrolyzed silane, works better. The coupling agent with glycidyl-silane functionality has epoxy functional groups that can readily react with both the epoxy resin during the curing process (cross-linking) and with the hydroxyl functional groups of the basalt fibers, leading to a strong bridge between these two components.



**Figure 4.** Field emission scanning electron microscopy (FESEM) images at  $\times 500$  magnifications corresponding to cryofracture surfaces from impact tests of a) BSB sandwich panel and b) LSL sandwich panel.

This corroborates the results obtained in the mechanical tests since the improvements of the flexural properties of the panels reinforced by basalt fibers is the result of the good synergy between these fibers and the polymeric matrix, which allows an excellent distribution of stress.

#### 4. Conclusions

Through this research, it has been confirmed that the use of a porous core as a diffusion media in a conventional vacuum assisted resin infusion moulding (VARIM) is a successful alternative process to manufacture sandwich-type lightweight composite structures using a partially biobased epoxy resin (with 31 wt% biobased content derived from epoxidized vegetable oils). The use of an infusion core material allows reducing weight and costs of manufacturing high-tech composite parts as the typical consumable materials in a VARIM process can be reduced.

Silanes play a key role in improving mechanical performance as they provide strong links between the epoxy resin (through the glycidyl functional group) and the fiber (through a condensation process of the hydrolyzed silane).

Regarding the mechanical properties of sandwich panels, the best performance is obtained in composite panels with the stacking sequence BBSBB. However, the substitution of one of the basalt fabrics in the face sheets by a flax fabric has an essential effect on overall properties, mainly on flexural strength and modulus, while impact strength remains almost invariable. Although these flax-based composite panels offer inferior properties than all-basalt composites, they represent an interesting alternative from both technical and environmental points of view.

#### References

- [1] P. Cousin, M. Hassan, P. V. Vijay, M. Robert, and B. Benmokrane, "Chemical resistance of carbon, basalt, and glass fibers used in frp reinforcing bars," *Journal of Composite Materials*, vol. 53, no. 26–27, pp. 3651–3670, 2019. [Online]. Available: <https://doi.org/10.1177/0021998319844306>
- [2] B. Wang, Q. Fu, H. Li, L. Qi, Q. Song, and Y. Fu, "In situ growth of graphene on carbon fabrics with enhanced mechanical and thermal properties for tribological applications of carbon fabric-phenolic composites," *Tribology Transactions*, vol. 62, no. 5, pp. 850–858, 2019. [Online]. Available: <https://doi.org/10.1080/10402004.2019.1626519>
- [3] K. Liu, S. He, Y. Qian, Q. An, A. Stein, and C. W. Macosko, "Nanoparticles in glass fiber-reinforced polyester composites: Comparing toughening effects of modified graphene oxide and core-shell rubber," *Polymer Composites*, vol. 40, no. S2, pp. E1512–E1524, 2019. [Online]. Available: <https://doi.org/10.1002/pc.25065>
- [4] M. S. Z. Abidin, T. Herceg, E. S. Greenhalgh, M. Shaffer, and A. Bismarck, "Enhanced fracture toughness of hierarchical carbon nanotube reinforced carbon fibre epoxy composites with engineered matrix microstructure," *Composites Science and Technology*, vol. 170, pp. 85–92, 2019. [Online]. Available: <https://doi.org/10.1016/j.compscitech.2018.11.017>
- [5] M. Flanagan, A. Doyle, K. Doyle, M. Ward, M. Bizeul, R. Canavan, B. Weafer, C. M. Brádaigh, N. M. Harrison, and J. Goggins, "Comparative manufacture and testing of induction welded and adhesively bonded carbon fibre peek stiffened panels," *Journal of Thermoplastic Composite Materials*, vol. 32, no. 12, pp. 1622–1649, 2019. [Online]. Available: <https://doi.org/10.1177/0892705718792362>
- [6] M. R. Ricciardi, I. Papa, V. Antonucci, V. Lopresto, and A. Langella, "Impact behavior of polyester gfrp for naval applications: Influence of the clamping device and fluid-material interaction," *Journal of Materials Engineering and Performance*, vol. 28, no. 6, pp. 3196–3202, Jun 2019. [Online]. Available: <https://doi.org/10.1007/s11665-019-03902-w>
- [7] J. Naveen, M. Jawaaid, E. S. Zainudin, M. T. H. Sultan, and R. Yahaya, "Effect of graphene nanoplatelets on the ballistic performance of hybrid kevlar/cocos nucifera sheath-reinforced epoxy composites," *Textile Research Journal*, vol. 89, no. 21–22, pp. 4349–4362, 2019. [Online]. Available: <https://doi.org/10.1177/0040517519833970>
- [8] J. P. Arenas, J. L. C. no, L. Troncoso, and M. L. Auad, "Thermoplastic polyurethane/laponite nanocomposite for reducing impact sound in a floating floor," *Applied Acoustics*, vol. 155, pp. 401–406, 2019. [Online]. Available: <https://doi.org/10.1016/j.apacoust.2019.06.012>
- [9] M. Zhang, O. Bareille, and M. Salvia, "Cure and damage monitoring of flax fiber-reinforced epoxy composite repairs for civil engineering structures using embedded piezo micro-patches," *Construction and Building Materials*, vol. 225, pp. 196–203, 2019. [Online]. Available: <https://doi.org/10.1016/j.conbuildmat.2019.07.179>
- [10] M. Mudhukrishnan, P. Hariharan, K. Palanikumar, and B. Latha, "Optimization and sensitivity analysis of drilling parameters

- for sustainable machining of carbon fiber-reinforced polypropylene composites,” *Journal of Thermoplastic Composite Materials*, vol. 32, no. 11, pp. 1485–1508, 2019. [Online]. Available: <https://doi.org/10.1177/0892705718799816>
- [11] R. M. Gandia, F. C. Gomes, A. A. R. Corrêa, M. C. Rodrigues, and R. F. Mendes, “Physical, mechanical and thermal behavior of adobe stabilized with glass fiber reinforced polymer waste,” *Construction and Building Materials*, vol. 222, pp. 168–182, 2019. [Online]. Available: <https://doi.org/10.1016/j.conbuildmat.2019.06.107>
- [12] A.-S. Mora, R. Tayouo, B. Boutevin, G. David, and S. Caillol, “Synthesis of biobased reactive hydroxyl amines by amination reaction of cardanol-based epoxy monomers,” *European Polymer Journal*, vol. 118, pp. 429–436, 2019. [Online]. Available: <https://doi.org/10.1016/j.eurpolymj.2019.06.020>
- [13] C. François, S. Pourchet, G. Boni, S. Rautainen, J. Samec, L. Fournier, C. Robert, C. M. Thomas, S. Fontaine, Y. Gaillard, V. Placet, and L. Plasseraud, “Design and synthesis of biobased epoxy thermosets from biorenewable resources,” *Comptes Rendus Chimie*, vol. 20, no. 11, pp. 1006–1016, 2017. [Online]. Available: <https://doi.org/10.1016/j.crci.2017.10.005>
- [14] S. Torron, S. Semlitsch, M. Martinelle, and M. Johansson, “Biocatalytic synthesis of epoxy resins from fatty acids as a versatile route for the formation of polymer thermosets with tunable properties,” *Biomacromolecules*, vol. 17, no. 12, pp. 4003–4010, 2016, pMID: 27809488. [Online]. Available: <https://doi.org/10.1021/acs.biomac.6b01383>
- [15] M. Jebrane, S. Cai, C. Sandström, and N. Terziev, “The reactivity of linseed and soybean oil with different epoxidation degree towards vinyl acetate and impact of the resulting copolymer on the wood durability,” 2017. [Online]. Available: <https://doi.org/10.3144/expresspolymlett.2017.37>
- [16] R. Singh, B. Singh, and H. Tarannum, “Mechanical properties of jute fiber-reinforced up/pu hybrid network composites,” *Polymers and Polymer Composites*, vol. 27, no. 9, pp. 546–556, 2019. [Online]. Available: <https://doi.org/10.1177/0967391119852882>
- [17] E. Mahdi, D. Ochoa, A. Vaziri, and E. Eltai, “Energy absorption capability of date palm leaf fiber reinforced epoxy composites rectangular tubes,” *Composite Structures*, vol. 224, p. 111004, 2019. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2019.111004>
- [18] A. K. Barouni and H. N. Dhakal, “Damage investigation and assessment due to low-velocity impact on flax/glass hybrid composite plates,” *Composite Structures*, vol. 226, p. 111224, 2019. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2019.111224>
- [19] K. Mazur, S. Kuciel, and K. Salasinska, “Mechanical, fire, and smoke behaviour of hybrid composites based on polyamide 6 with basalt/carbon fibres,” *Journal of Composite Materials*, vol. 53, no. 28-30, pp. 3979–3991, 2019. [Online]. Available: <https://doi.org/10.1177/0021998319853015>
- [20] J. R. Vinson, “Sandwich structures: Past, present, and future,” in *Sandwich Structures 7: Advancing with Sandwich Structures and Materials*. Springer Netherlands, 2005, pp. 3–12. [Online]. Available: [https://doi.org/10.1007/1-4020-3848-8\\_1](https://doi.org/10.1007/1-4020-3848-8_1)
- [21] A. S. Herrmann, P. C. Zahlen, and I. Zuardy, “Sandwich structures technology in commercial aviation,” in *Sandwich Structures 7: Advancing with Sandwich Structures and Materials*, O. Thomsen, E. Bozhevolnaya, and A. Lyckegaard, Eds. Dordrecht: Springer Netherlands, 2005, pp. 13–26. [Online]. Available: [https://doi.org/10.1007/1-4020-3848-8\\_2](https://doi.org/10.1007/1-4020-3848-8_2)
- [22] A. Atiqah, M. T. Mastura, B. A. Ahmed Ali, M. Jawaid, and S. M. Sapuan, “A review on polyurethane and its polymer composites,” *Current Organic Synthesis*, vol. 14, pp. 233–248, 2017. [Online]. Available: <https://doi.org/10.2174/1570179413666160831124749>
- [23] M. Yaqoob, “Fabrication and structural equivalency analysis of cfrp nomex core sandwiched panels for fsae race car,” *The UNSW Canberra at ADFA Journal of undergraduate engineering research*, vol. 12, no. 2, 2017. [Online]. Available: <https://bit.ly/2LGQHUy>
- [24] C. Atas and C. Sevim, “On the impact response of sandwich composites with cores of balsa wood and pvc foam,” *Composite Structures*, vol. 93, no. 1, pp. 40–48, 2010. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2010.06.018>
- [25] G. Zhang, B. Wang, L. Ma, L. Wu, S. Pan, and J. Yang, “Energy absorption and low velocity impact response of polyurethane foam filled pyramidal lattice core sandwich panels,” *Composite Structures*, vol. 108, pp. 304–310, 2014. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2013.09.040X>
- [26] L. He, Y.-S. Cheng, and J. Liu, “Precise bending stress analysis of corrugated-core, honeycomb-core and x-core sandwich panels,”

- Composite Structures*, vol. 94, no. 5, pp. 1656–1668, 2012. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2011.12.033>
- [27] R. Stewart. (2011) Sandwich composites excel at cost-effective, lightweight structures. [Online]. Available: <https://bit.ly/2YAF3Q5>
- [28] R. Chatys, A. Panich, R. S. Jurecki, and M. Kleinhofs, “Composite materials having a layer structure of “sandwich” construction as above used in car safety bumpers,” in *2018 XI International Science-Technical Conference Automotive Safety*, April 2018, pp. 1–8. [Online]. Available: <https://doi.org/10.1109/AUTOSAFE.2018.8373320>
- [29] S. H. Eum, Y. H. Kim, J. W. Han, K. J. Kim, D. H. Shin, C. Yim, and R. Murakami, “A study on the mechanical properties of the honeycomb sandwich composites made by vartm,” in *Advances in Fracture and Strength*, ser. Key Engineering Materials, vol. 297. Trans Tech Publications Ltd, 8 2005, pp. 2746–2751. [Online]. Available: <https://doi.org/10.4028/www.scientific.net/KEM.297-300.2746>
- [30] J. M. España, M. D. Samper, E. Fages, L. Sánchez-Nácher, and R. Balart, “Investigation of the effect of different silane coupling agents on mechanical performance of basalt fiber composite laminates with biobased epoxy matrices,” *Polymer Composites*, vol. 34, no. 3, pp. 376–381, 2013. [Online]. Available: <https://doi.org/10.1002/pc.22421>
- [31] M. D. Samper, R. Petrucci, L. Sánchez-Nacher, R. Balart, and J. M. Kenny, “New environmentally friendly composite laminates with epoxidized linseed oil (elo) and slate fiber fabrics,” *Composites Part B: Engineering*, vol. 71, pp. 203–209, 2015. [Online]. Available: <https://doi.org/10.1016/j.compositesb.2014.11.034>
- [32] R. Petrucci, C. Santulli, D. Puglia, F. Sarasini, L. Torre, and J. M. Kenny, “Mechanical characterisation of hybrid composite laminates based on basalt fibres in combination with flax, hemp and glass fibres manufactured by vacuum infusion,” *Materials & Design*, vol. 49, pp. 728–735, 2013. [Online]. Available: <https://doi.org/10.1016/j.matdes.2013.02.014>
- [33] H. N. Dhakal, F. Sarasini, C. Santulli, J. Tirilló, Z. Zhang, and V. Arumugam, “Effect of basalt fibre hybridisation on post-impact mechanical behaviour of hemp fibre reinforced composites,” *Composites Part A: Applied Science and Manufacturing*, vol. 75, pp. 54–67, 2015. [Online]. Available: <https://doi.org/10.1016/j.compositesa.2015.04.020>
- [34] S. Boccardi, N. D. Boffa, G. M. Carlomagno, G. D. Core, C. Meola, P. Russo, and G. Simeoli, “Inline monitoring of basalt-based composites under impact tests,” *Composite Structures*, vol. 210, pp. 152–158, 2019. [Online]. Available: <https://doi.org/10.1016/j.compstruct.2018.11.038>
- [35] V. Fiore, T. Scalici, L. Calabrese, A. Valenza, and E. Proverbio, “Effect of external basalt layers on durability behaviour of flax reinforced composites,” *Composites Part B: Engineering*, vol. 84, pp. 258–265, 2016. [Online]. Available: <https://doi.org/10.1016/j.compositesb.2015.08.087>
- [36] D. Lascano, L. Quiles-Carrillo, S. Torres-Giner, T. Boronat, and N. Montanes, “Optimization of the curing and post-curing conditions for the manufacturing of partially bio-based epoxy resins with improved toughness,” *Polymers*, vol. 11, no. 8, 2019. [Online]. Available: <https://doi.org/10.3390/polym11081354>
- [37] X. Gao, R. E. Jensen, W. Li, J. Deitzel, S. H. McKnight, and J. R. J.W. Gillespie, “Effect of fiber surface texture created from silane blends on the strength and energy absorption of the glass fiber/epoxy interphase,” *Journal of Composite Materials*, vol. 42, no. 5, pp. 513–534, 2008. [Online]. Available: <https://doi.org/10.1177/0021998307086203>
- [38] P. K. Kushwaha and R. Kumar, “Effect of silanes on mechanical properties of bamboo fiber-epoxy composites,” *Journal of Reinforced Plastics and Composites*, vol. 29, no. 5, pp. 718–724, 2010. [Online]. Available: <https://doi.org/10.1177/0731684408100691>



# COMPRESSIVE SENSING BASED CHANNEL ESTIMATOR AND LDPC THEORY FOR OFDM USING SDR

## ESTIMADOR DE CANAL BASADO EN SENSADO COMPRESIVO Y LDPC PARA OFDM USANDO SDR

Anthony Yanza-Verdugo<sup>1</sup>, Christian Pucha-Cabrera<sup>1</sup> and Juan Inga-Ortega<sup>2,\*</sup>

### Abstract

This work proposes the application of a channel estimator based on Compressive Sensing (CS) on a system that employs Orthogonal Frequency Division Multiplexing (OFDM), utilizing Software Defined Radio (SDR) devices. The application of the CS theory is given through the use of sparse reconstruction algorithms such as Orthogonal Matching Pursuit (OMP) and Compressive Sampling Matching Pursuit (CoSaMP), in order to take advantage of the sparse nature of the pilot subcarriers used in OFDM, optimizing the bandwidth of system. In addition, to improve the performance of these algorithms, the concept of sparse parity checking matrix is used, which is implemented in the deployment of low density parity check codes (LDPC) to obtain a sensing matrix that improves the isometric restriction property (IRP) belonging to the CS paradigm. The document shows the model implemented in the SDR equipment, analyzing the bit error rate and the number of pilot symbols used.

**Keywords:** Channel Estimation, LDPC, OFDM, SDR, Compressive Sensing.

### Resumen

Este trabajo propone la aplicación de un estimador de canal basado en sensado compresivo (CS, del inglés *Compressive Sensing*) sobre un sistema que usa multiplexación por división de frecuencias ortogonales (OFDM, del inglés *Orthogonal Frequency Division Multiplexing*) usando dispositivos de radio definido por *software* (SDR, del inglés *Software Defined Radio*). La aplicación de la teoría de CS se da a través del uso de algoritmos de reconstrucción dispersa como *Orthogonal Matching Pursuit* (OMP) y *Compressive Sampling Matching Pursuit* (CoSaMP) con el fin de aprovechar la naturaleza dispersa de las subportadoras piloto usadas en OFDM optimizando el ancho de banda del sistema. Además, para mejorar el rendimiento de estos algoritmos, se utiliza el concepto de la matriz de comprobación de paridad dispersa que se implementa en el despliegue de códigos de comprobación de paridad de baja densidad (LDPC, del inglés *Low Density Parity Check*) para obtener una matriz de sensado que mejore la propiedad de restricción isométrica (RIP, del inglés *Isometric Restriction Property*) perteneciente al paradigma de CS. El documento muestra el modelo implementado en los equipos SDR analizando la tasa de error de bit y la cantidad de símbolos piloto usados.

**Palabras clave:** estimación de canal, LDPC, OFDM, SDR, sensado compresivo.

<sup>1,\*</sup>Electronics Engineering / GITEL, Universidad Politécnica Salesiana – Ecuador.

<http://orcid.org/0000-0002-1710-3052>, <http://orcid.org/0000-0002-4734-7218>

<sup>2</sup>Telecommunications / GITEL, Universidad Politécnica Salesiana – Ecuador.

Correspondence author ✉: [jinga@ups.edu.ec](mailto:jinga@ups.edu.ec) <http://orcid.org/0000-0003-2580-9677>

Received: 31-10-2019, accepted after review: 09-12-2019

Suggested citation: Yanza-Verdugo, A.; Pucha-Cabrera, C. and Inga-Ortega, J. (2020). «Compressive Sensing Based Channel Estimator and LDPC Theory for OFDM using SDR». INGENIUS. N.º 23, (january-june). pp. . DOI: <https://doi.org/10.17163/ings.n23.2020.07>.

## 1. Introduction

The orthogonal frequency division multiplexing (OFDM) is currently the transmission technique mostly implemented in wireless networks, due to its advantages in high speed data transmission through frequency selective channels. Thereby, the use of OFDM has had a high performance because of its high efficiency in the use of the radio spectrum and its robustness to multipath delay [1, 2]. In addition, the inter symbol interference (ISI) and inter carrier interference (ICI) are reduced through the use of a cyclic prefix, enabling the replacement of complex time-domain equalizers by a simple frequency-domain equalizer [3]. In this sense, for frequency selective and time-varying radio channels in broadband mobile communication systems, it is necessary to apply a reliable channel estimation to implement coherent detection [1].

There are three types of channel estimators: blind estimation, semi-blind estimation and pilot symbol assisted modulation (PSAM). PSAM uses pilot symbols that are known by both receiver and transmitter at different positions, to obtain an approximation of the channel in exchange for losing spectral efficiency. On the other hand, blind estimation is focused on statistical processes with high mathematical level exhibiting improvements regarding spectral efficiency, even though it is not commonly used nowadays due to its high complexity, slow speed of convergence and low performance [4]. At last, the semi-blind estimation combines the two previous approaches to yield a compromise between complexity and efficiency; for this matter, both training sequences and statistical models of the channel are utilized [5].

On the basis of the above, this work proposes the use of PSAM for channel estimation, where the gain of the channel and the phase distortion are obtained from the received signal at the positions of the pilot symbols [1]. These pilot symbols are also utilized to improve the synchronization of time and frequency in the communication system. Nevertheless, it is necessary to use a greater percentage of pilot symbols to increase the performance, which implies that certain subcarriers do not transmit information [2].

In addition, in this work the number of pilot subcarriers is reduced through the use of compressive sensing (CS) in the estimation of the channel, in order to improve the utilization of the bandwidth. This can be applied since it is possible to estimate the channel taking advantage of the CS theory, considering that the number of pilot symbols is sparse with respect to the total number of symbols that constitute an OFDM message, thus obtaining a compressed version of the channel corresponding to its effects on the pilot subcarriers [6].

The process is also optimized for implementation on software defined radio (SDR) equipment, through

the use of the low density parity check (LDPC) matrix in combination with the CS paradigm.

The rest of the paper is organized as follows. Section 2 discusses the mathematical modeling of OFDM, CS and the application of LDPC matrices with CS. Section 3 describes the implementation in the SDR devices using LDPC matrices with CS. Section 4 presents the results obtained, and the analysis of the bandwidth and the bit error rate (BER). At last, section 5 gives the conclusions and recommendations for future works.

## 2. Mathematical modeling

The implementation is developed from the mathematical model of OFDM, where an OFDM symbol is represented by  $S(q) \in \mathbb{R}^Q$ , such that a set of  $Q$  information symbols is transmitted by symbol  $j$  of OFDM; each with a subcarrier  $q$  is denoted as  $S_j(q)$ , thus that  $[S_j(0), \dots, S_j(q), \dots, S_j(Q-1)]^T$  represents the vector of information symbols transmitted in the  $j$ -th OFDM symbol [2, 7].

Each OFDM symbol uses the serial flow of information symbols, converting them into  $Q$  parallel flows and later inserting the pilot symbols necessary in the estimation of the channel according to the proposed transmission scheme. Besides, subcarriers in zero are added to establish guard bands between each OFDM symbol, thus avoiding the interference of adjacent channels. The number of subcarriers of an OFDM symbol is given by

$$Q = I + P + Z \tag{1}$$

where  $I$  corresponds to the number of subcarriers with information,  $P$  is the number of pilot subcarriers and  $Z$  is the number of subcarriers in zero, for an OFDM symbol. According to this, equation (2) describes the format of an OFDM symbol corresponding to the frequency domain, and then the inverse discrete Fourier transform is applied to obtain the symbol in the time domain, as described in equation (3). The distribution of the zero-padding is detailed in [8].

$$S_z(q) \triangleq \begin{cases} S(q), & \frac{Q-(I+P)}{2} \leq q \leq \frac{Q-(I+P)}{2} - 1 \\ 0, & \text{any other case} \end{cases} \tag{2}$$

$$S_z = F^H S \tag{3}$$

The reduction of ISI and ICI is due to the use of the so-called cyclic prefix (CP), which consists of a cyclic extension of the OFDM symbol. The CP is as long as the expected propagation delay, and the effects of both are found in [9]. On the other hand, the signal received for OFDM may be described in equation (4).

$$y_j = H_j s_j + z_j \tag{4}$$

The vector of received information symbols is  $y_j = [y_j(0), \dots, y_j(Q-1)]^T$  in the  $j$ -th OFDM symbol,  $z_j$  is the Gaussian noise and  $H_j$  denotes the value of the channel for the  $j$ -th OFDM symbol, which is obtained by means of the channel estimation block [7].

The receiver uses the Van de Beek algorithm, which takes advantage of the cyclic prefix to establish the beginning and the end of an OFDM symbol. This algorithm will also allow correcting the effects of the channel in possible phase shifts, and avoid losing the orthogonality in the subcarriers [10].

In the following the cyclic prefix is removed and the discrete Fourier transform (DFT) is applied to recover the information symbols. Prior to the recovery of the information symbols, the channel is estimated after removing the zero-padding and extracting the pilot symbols. The equalized data are obtained through the operation indicated in equation (5) [11], where  $\hat{H}_e(q)$  is the estimated channel. Table 1 presents a list of all the variables used along this work.

$$S_e = \frac{Y(q)}{\hat{H}_e(q)} \quad q = 0, 1, \dots, Q-1 \quad (5)$$

## 2.1. Compressive sensing

From another perspective, the CS paradigm considers a «sparse» signal  $x$  that possesses only  $K$  elements different than zero, with  $K \ll N$  and where  $x \in \mathbb{R}^N$ . Then, through a matrix  $\Phi$  of dimension  $M \times N$ , with  $N > M$ , known as sensing matrix, it is sought to concentrate the more relevant information in  $x$  through the use of another vector  $y \in \mathbb{R}^M$ . If this is not fulfilled, the ability of reconstructing  $x$  from  $y$  is reduced; this restriction is known as null space property (NSP) [11, 12].

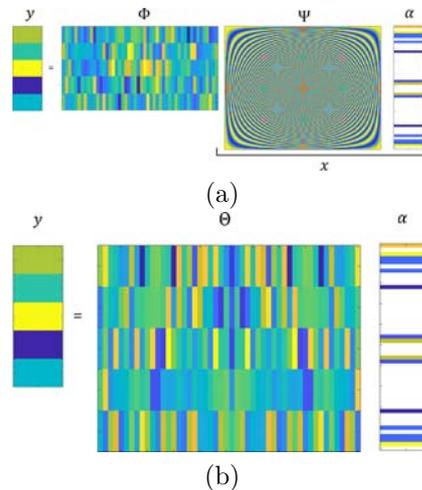
It is not easy to find signals considered «sparse», however, it is possible to find an approximation to this concept which appears when the signal  $x$  may be «compressible» in some vector basis different to the original [11–13]. Therefore,  $x = \Psi\alpha$ , where  $\Psi$  (also called dictionary matrix of  $x$ ) corresponds to the vector basis in which  $x$  is projected, and  $\alpha$  corresponds to the atoms of  $x$  inside of the domain of  $\Psi$ . Thereby, the original signal may be represented through the linear combination proposed in (6) [14]. For instance, the dictionary may be of Fourier, DCT or Wavelet [11, 15, 16].

$$x = \Psi\alpha = \sum_{i=1}^N \alpha_i \psi_i \quad (6)$$

The compression of the signal  $x$  in the signal  $y$  may be obtained through (7), where  $\Theta = \Phi\Psi$ . Figures 1(a) and 1(b) graphically show the concept of CS [13, 17].

$$y = \Phi x = \Phi\Psi\alpha = \Theta\alpha \quad (7)$$

The literature suggests that  $\Phi$  should be random to fulfill the isometric restriction property (IRP), and in this way  $x$  may be reconstructed from  $y$  [16, 18].



**Figure 1.** Principle of operation of compressive sensing: (a) measuring process in CS with random sensing matrix  $\Phi$  and measuring matrix  $\Psi$ . (b) measuring process with  $\Theta = \Phi\Psi$ .

**Table 1.** List of variables

Variables	Description
$Q$	Total of subcarriers in the OFDM symbol.
$S_j(q)$	Subcarrier $q$ in the $j$ -th OFDM symbol.
$I$	Total of subcarriers with information.
$P$	Total of pilot subcarriers.
$Z$	Total of subcarriers in zero.
$S_z(q)$	OFDM symbol with zero-padding in subcarrier $q$ .
$F$	Fourier transform matrix.
$s_z$	OFDM symbol with zero-padding in time.
$y_j$	Vector of received symbols in the $j$ -th OFDM symbol in time.
$H_j$	Channel vector in the $j$ -th OFDM symbol.
$s_j$	$j$ -th OFDM symbol transmitted in time.
$Z_j$	Complex Gaussian noise.
$Y(q)$	OFDM symbol in frequency, after removing the cyclic prefix, zero-padding and without pilots.
$S_e$	Symbols equalized in frequency.
$H_e(q)$	Estimated channel in frequency.
$x$	Data vector of length $N$ .
$y$	Compressed vector of length $M$ .
$\Phi$	Compressing or sensing matrix, Phi.
$N$	Total length of the vector of data to be compressed.
$K$	Total elements different than zero.
$M$	Total length of the vector of compressed data.
$\Psi$	Measuring or dictionary matrix, Psi.
$\alpha$	Vector of sparse coefficients, alpha.
$\Theta$	Sensing matrix in its complete form, Theta.
$W_c$	Number of logic "1"s per column in $H_{LDPC}$
$W_r$	Number of logic "1"s per row $H_{LDPC}$
$H_{LDPC}$	Parity matrix of LDPC.
$A$	Matrix of $D \times E$ elements.
$B$	Matrix of $D \times D$ elements.
$D$	Number of parity bits in the coding.
$E$	Number of information bits in the coding.
$G$	Total bits of information + parity (codeword).
$H_l$	Left side irregular parity matrix of LDPC.
$SC$	Right side staircase parity matrix of LDPC.
$Tr$	Right side triangle parity matrix of LDPC.
$\hat{H}_p$	Channel of pilots.
$Y_p$	Received pilot symbols.
$S_p$	Known pilot symbols.
$N_f$	Frequency separation between pilots.
$\Phi_{LDPC}$	Sensing matrix constructed with LDPC base.

In general, the sensing matrices may be obtained from a random process with normal distribution, achieving a good performance during the reconstruction of the signal, even though these distributions may demand many resources [15, 16].

Since  $N > M$  and  $N \gg K$ , the process of reconstructing  $x$  becomes the optimization problem formulated in equation (8), where it is sought to find the atoms  $\alpha$  that minimize the error  $\|y - \Theta\alpha\|_0$ .

$$\alpha = \arg \min \|\alpha\|_0 \quad s.t. \Phi\Psi\alpha \quad (8)$$

Then,  $x$  may be reconstructed from  $\alpha$  using equation (6). Nevertheless, the optimization problem has no solution since it is of type NP Hard [11–13]. In this respect, the sparse reconstruction algorithms seek to relax the optimization problem to obtain a pseudo-optimal solution. In addition, their performance can be improved if it is used an appropriate sensing matrix which enables preserving the information and guarantees the reconstruction of the original signal in a unique manner [15, 16, 19–22].

The Orthogonal Matching Pursuit (OMP) can be highlighted, among the most popular sparse reconstruction algorithms used in the application of CS for channel estimation. OMP belongs to the types of algorithms based in greedy search, thus it is based on successive approximations of the  $\alpha$  coefficients, identifying the support of the signal in an iterative manner until the convergence criterion is reached [19, 23]. The OMP algorithm is described in Table 2.

Another algorithm of sparse reconstruction employed in this case of application of CS, is the Compressive Sampling Matching Pursuit (CoSaMP) described in [24]. The algorithm approximates the objective signal in an iterative manner, and in each iteration, the current approximation obtains a residual updating the samples such that they reflect the current residue. These samples are utilized to identify the large components, to estimate the approximation using least squares. This process is repeated until the recoverable energy of the signal is found. The CoSaMP algorithm is described in Table 3 [25].

These algorithms are suitable to be used in applications or reconstructions where they operate in real time, due to their low computational cost.

Since the sensing matrix should fulfill the isometric restriction property (IRP), the use of parity check matrices is considered in the design of channel coders in the LDPC codes, since they fulfill the IRP criterion and are deterministic; therefore they consume less resources, thus being a good choice for the measuring matrix in CS [20–22]. This will enable to avoid the use of a pattern of pilots of the comb or block types, as it is commonly utilized in OFDM with PSAM. On this basis, it is sought to use a sparse pseudorandom pattern of pilots employing an LDPC parity check

matrix [5].

**Table 2.** OMP algorithm

<b>Algorithm 1:</b> Algorithm OMP (Orthogonal Matching Pursuit)	
Step 1	<b>Input:</b> $Y = \{y_1, \dots, y_M\}$ ; Compressed Input $\Theta_{N,N}$ : Complete sensing matrix $k$ : Number of repetitions
Step 2	<b>Output:</b> $\alpha_N$
Step 3	<b>Initialization:</b> $res_M = Y$ ; $indx = \phi$ ;
Step 4	<b>for</b> $iter = 1$ to $k$ : $\lambda = \arg \max  \Theta^T \times res_M$ ; $indx = indx \cup \lambda_M$ ; $\alpha_N = \phi$ ; $\alpha_N(indx) = pinv(\Theta(:, indx)) \times Y$ ; $res = Y - \Theta \times \alpha_N$ ; <b>endfor</b>
Step 5	<b>Return:</b> $\alpha_N$

**Table 3.** CoSaMP Algorithm

<b>Algorithm 2:</b> Algorithm CoSaMP (Compressive Sampling Matching Pursuit)	
Step 1	<b>Input:</b> $x \in R^N$ ; Compressed Input $\Theta \in R^{m \times N}$ ; con $N > m$ ; Complete sensing matrix $k$ : Number of repetitions
Step 2	<b>Output:</b> $S_{est}$
Step 3	<b>Initialization:</b> $u = \Theta x$ , Measuring vector; $\Omega = \{1, 2, \dots, N\}$ Column index of $\Theta$ ;
Step 4	<b>for</b> $iter = 1$ to $k$ : $\Omega_k \subset \Omega \rightarrow \Theta_{\Omega_k}$ ; $\Omega_{k+1} = J_s^*$ where $J_s^* \subset \Omega$ ; $J = Supp\{T_1( \Theta \times r_k ; \tau_{k,1})\}$ with $ J  \leq 2s$ ; where: $T_1$ is a threshold function such that: $\tau_{k,1} \geq 0$ ; $\tau_k$ : residual; with $\Omega_k \cup J$ : $\ u - \Theta I_{\Omega_k \cup J} b\ _2$ minimized; where: $I_{\Omega_k \cup J}$ is an $N \times N$ diagonal matrix; $i \in \Omega_k \cup J$ ; $J_s^* = Supp\{T_2( b ; \tau_{k,2})\}$ ; where: $T_2$ is a threshold function; $\tau_{k,2} \geq 0$ maximum of $s$ elements of $b$ that are retained; $S_{est} = b(J_s^*)$ ; $r_{k+1} = u - \Theta I_{\Omega_{k+1}} b$ ; <b>endfor</b>
Step 5	<b>Return:</b> $S_{est}$

## 2.2. LDPC Parity check matrix

The parity check matrix  $H_{LDPC}$  defines the relationships between the different codification symbols (source symbols and parity symbols). There are two types, the regular matrices which have a fixed number  $Wc$  of logic «1»s per column and a fixed number  $Wr$  of logic «1»s per row. The second is the case of the irregular matrices whose number of logic «1»s per row is  $Wr$  and per column  $Wc$ , with  $Wr \neq Wc$ . The matrix is constituted by elements with values «0» and «1», and is sparse since most of the elements are null [26]. This matrix is described in equation (9).

$$H_{LDPC} = [A|B] \quad (9)$$

Matrix  $A$  has dimension  $D \times E$  and matrix  $B$  is of dimension  $D \times D$ , where  $E$  corresponds to the number of information bits,  $D$  is the number of parity bits in

the LDPC codification, and  $G = E + D$  would correspond to the length of the codeword. The regular LDPC matrices are divided in Gallager, characterized by a structure of horizontal bands [27], and in Mackay-Neal, characterized by a random construction [28]. These type of matrices do not guarantee the independence between columns required by the IRP in CS, thus they will not be used in this implementation [13].

The irregular matrices are divided in staircase and triangle [29]. They are matrices that contain two submatrices, and each of them uses a different method of creation [29,30]. The left submatrix  $H_i$  of size  $D \times E$  is constructed as follows:

- Use a vector which contains a random list of possible positions with «1»s in the rows. This vector is utilized to guarantee a homogeneous distribution.
- Place in each column  $Wc$  nonzero elements, choosing the positions of the vector of possible localizations. If the condition of a maximum number  $Wr$  of nonzero elements per row is not fulfilled now, choose another available random position.
- Add nonzero elements in the rows with less than two elements, to avoid decoding problems.

The right submatrix makes the difference in the formation of an irregular LDPC matrix, since it can be formed in staircase or in triangular shape, and in any case is a matrix of dimension  $D \times D$ . In the case of matrix  $Sc$ , it is an identity matrix in which nonzero elements are later placed in the lower diagonal part. Equation (10) shows a matrix  $H_{LDPC}$  constructed using this method.

$$(H_i|SC) = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & \vdots & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & \vdots & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & \vdots & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & \vdots & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & \vdots & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (10)$$

The right triangle submatrix  $Tr$  is constituted by an identity matrix of dimension  $D \times D$  as base, and a sparse triangular lower matrix placed later. Equation (11) shows a  $H_{LDPC}$  matrix constructed by means of this method.

$$(H_i|Tr) = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & \vdots & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & \vdots & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & \vdots & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & \vdots & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & \vdots & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (11)$$

### 3. Implementation of CS and LDPC for channel estimation

This work is implemented in universal software radio peripheral (USRP) equipment. The block diagram of the implemented system is shown in Figure 2, where each stage of the communication system is presented.

The equipment used are the USRP Ettus X310, equivalent to the NI-USRP 2940R with the following specifications:

- Bandwidth up to 40 MHz per channel (2 channels).
- The image loaded in the FPGA enables a 1 Gbps Ethernet connection for 25 MS/s Full Duplex transmission.
- The used UHD architecture is the available for the «LABVIEW Communication Design Suite».
- Flexible clock architecture with configurable sampling frequency.

The bandwidth of the system is a function of the utilized IQ index, which determines the available bandwidth [31].

It should be indicated that with the purpose of working with a communication system closer to reality, a LDPC channel encoder has been used considering what was worked in [32,33]. The selection of the  $H_{LDPC}$  matrices of this implementation was established by validation through the analysis of the BER. The selection of the matrix  $H_{LDPC}$  applied simulations with a varying sparsity of «1»s of the matrix considering an AWGN channel with Rayleigh fading. These matrices have a size of G=40 and E=20, due to their efficiency and velocity according to Figure 3. For this reason, the implementation of the channel encoder uses parity check matrices of staircase type.



### 3.2. Channel estimation with CS and LDPC

As indicated in section 2, the sensing matrix  $\Phi$  should fulfill the IRP criterion to be able to appropriately recover the sparse signal; however, the matrix used in equation (13) fulfills this criterion at the restriction limits, since it is not random. Therefore, it is developed a matrix  $\Phi_{LDPC}$  of pilot positions that improves this requirement. In a form similar to equation (13), each nonzero element corresponds to a position of each subpilot carrier in equation (15) [6].

$$\Phi_{LDPC} = \begin{bmatrix} \overbrace{1 \ 0 \ 0 \ 0 \ 0 \ 0 \ \dots \ 0 \ 0}^{N_f} \\ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ \dots \ 0 \ 0 \\ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ \dots \ 0 \ 0 \\ \vdots \ \vdots \ \vdots \ \vdots \ \vdots \ \vdots \ \ddots \ \vdots \ \vdots \\ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \dots \ 0 \ 0 \end{bmatrix}_{P,N} \quad (15)$$

Equation (15) maintains the distance between the pilot subcarriers  $N_f$  as variable and pseudorandom, due to the use of the LDPC algorithms with which matrix  $H_{LDPC}$  is constructed. For the design of  $H_{LDPC}$ , the number of «1»s per row is equal to 1, taking into account the number of pilot subcarriers. In addition, the total percentage of nonzero elements should be equal to the number of pilots  $P$ .

After obtaining matrix  $\Phi_{LDPC}$ , the process continues with the estimation of the channel, using the sparse reconstruction algorithms OMP and CoSaMP, considering the distribution of pilots obtained in  $\Phi_{LDPC}$ .

### 3.3. Implemented scenarios

This work applied different scenarios, keeping a focus for Long Term Evolution (LTE) systems. For this purpose, characteristics of LTE indicated in [31,34] such as number of subcarriers, reference symbols, null subcarriers and length of the CP, were taken into account. A separation of 7.5 KHz between subcarriers was used, since the performance decays with the prolonged use of 15 KHz. Then, the implemented bandwidth was 1.92 MHz. It is necessary to clarify that a greater number of subcarriers was not used, as it is allowed by LTE, because the processing of the source code was carried out from the computer, and not on the FPGA of the equipment.

The first scenario implemented does not use the channel encoder and has the following characteristics:

- IQ sample rate: 1.92 MS/s.
- Frequency of carrier: 1.99 GHz.
- Modulation: 4 QAM.
- Iterations of the OMP algorithm: 5.

- Length of the transmitted message (bits train): 1600.
- Number of samples in the receiver:  $3 \times$  total of transmitted data.
- Total number of OFDM subcarriers: 256.
- Total of subcarriers with data + pilot subcarriers: 150.
- Gain of the transmitter: 15.5 - 21 dB.
- Sensitivity of the receiver: 0 dB.

In order to obtain the results of the behavior of the channel estimators for the different sparse reconstruction algorithms, OMP and CoSaMP, as well as of the linear estimator, two separate SDR located in the test laboratory (indoor) at a considerable distance, as shown in Figure 2, were used. The distance was considered as constant, and the power of the equipment were modified to obtain variations of the signal to noise ratio (SNR), and subject the different channel estimators to the BER analysis.

For the second scenario the previous conditions were repeated, adding the LDPC channel encoder using the following configuration:

- Iterations of the propagation algorithm for the LDPC decoding: 100.
- Matrix  $H_{LDPC}$  with G=40 and E=20.
- Type of LDPC matrix in the encoder: staircase with sparsity of «1»s between 5 and 10%.

The final scenario of analysis is implemented with variable distance between transmitter and receiver, also in a closed environment to evaluate the channel in more rigorous conditions. The distance of separation between the SDRs considered the length, the physical distribution of the laboratory with distances between 4 and 10 meters, and a constant power which does not saturate the channel.

The variation of distance enabled to recreate channels with great fading, generating an increase of the multipath effect to evaluate the behavior of the estimators. The frequency of transmission is 2.4 Ghz, thus coexisting with the Wi-Fi network of the laboratory; this causes a continuous variation in the channel.

This scenario was worked in the presence of students, using an LDPC encoder with the following configuration:

- IQ sample rate: 1.92 MS/s.
- Frequency of carrier: 2.4 GHz.
- Modulation: 4 QAM.
- Iterations of the OMP algorithm: 5.

- Iterations of the propagation algorithm for the LDPC decoding: 100.
- Matrix  $H_{LDPC}$  with  $G=40$  and  $E=20$ .
- Length of the transmitted message (bits train): 1240.
- Number of samples in the receiver:  $2 \times$  total of transmitted data.
- Total number of OFDM subcarriers: 256.
- Total of subcarriers with data + pilot subcarriers: 150.
- Type of LDPC matrix in the encoder: staircase with sparsity of «1»s between 5 and 10%, 10 and 20% and 20 and 30%.
- Gain of the transmitter: 20 dB.
- Sensitivity of the receiver: 20 dB.

## 4. Analysis of the results

### 4.1. Scenarios 1 and 2 – BER Analysis

In the implementation of the OMP algorithm, the number of iterations should be chosen such that  $K$  corresponds to the number of nonzero elements of the signal to be recovered. In order to define this value, multiple tests were conducted seeking to minimize the error in the estimation of the channel, obtaining that the appropriate value for this algorithm to converge is  $K = 5$ , corresponding to the minimum number of subcarriers used.

Figure 4 shows the behavior of the BER with 5 and 25 pilot subcarriers, for the estimators based on OMP, CoSaMP and the linear estimator for the first scenario.

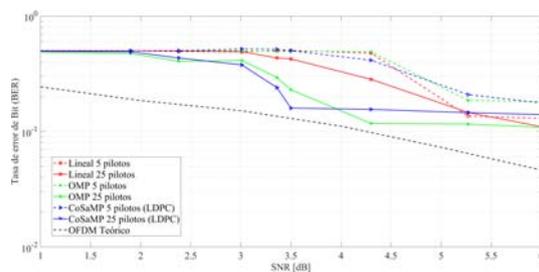


Figure 4. BER analysis without channel encoder.

It can be observed in Figure 4 that at low levels of SNR, the linear estimator exhibits a lower performance with respect to the employed estimators based on the CS paradigm. With 25 pilot subcarriers, the channel estimators achieve a considerable reduction in the probability of bits loss. In these tests, the algorithms CoSaMP and OMP worked with the  $\Phi_{LDPC}$

matrix and in its absence (comb-type pilot subcarriers distribution).

Figure 5 shows the behavior of the BER for 5 channel estimators, using the configurations of the second scenario. Then, according to what is expected with the use of a channel encoder, it can be observed in Figure 5(a) a clear improvement with respect to the previous case. The linear estimator exhibits the worst results for a low SNR. On the other hand, the CoSaMP estimators have a similar performance using matrix  $\Phi_{LDPC}$  or not using it. However, the version with matrix  $\Phi_{LDPC}$  exhibits a better performance. Also, as the SNR improves, the CoSaMP estimator without matrix  $\Phi_{LDPC}$  has a slight improvement, but both continuing with the same trend in their behavior. On the other hand, the OMP estimators exhibit a lower performance compared with the CoSaMP, taking into account that the version with matrix  $\Phi_{LDPC}$  is the one that offers the worst results for this particular algorithm.

Figure 5(b) shows the results of the performance of the estimators regarding the BER, using 25 pilot subcarriers for each of them. It can be observed that the linear estimator has the greatest probability of bit error for a low SNR, showing the worst performance among the estimators analyzed. Nevertheless, when the SNR a priori exceeds the value of 3.5 dB its performance improves, obtaining the smallest probability of bit error compared to the other estimators.

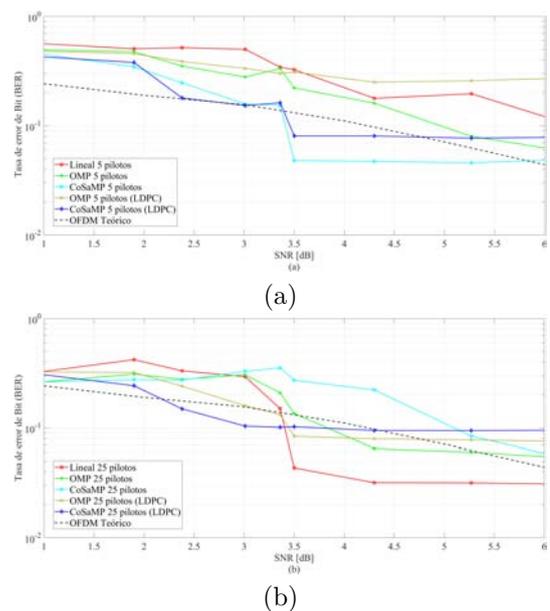


Figure 5. BER analysis with LDPC channel encoder: (a) BER with 5 pilots (b) BER with 25 pilots.

### 4.2. Scenario 3 - Analysis for variable distances

Figures 6 and 7 correspond to the behavior of the linear estimator, and of the estimators based on CS with and without matrix  $\Phi_{LDPC}$ .

Figure 6 shows that the linear estimator has a better performance at a very short distance, which corresponds with the previous results in BER analysis. Nevertheless, as the distance increases, the performance of this channel estimator is drastically reduced. The CoSaMP estimator with  $\Phi_{LDPC}$  is better compared with the estimator that uses OMP. It should be indicated that this figure takes into account the results using from 5 to 25 pilot subcarriers.

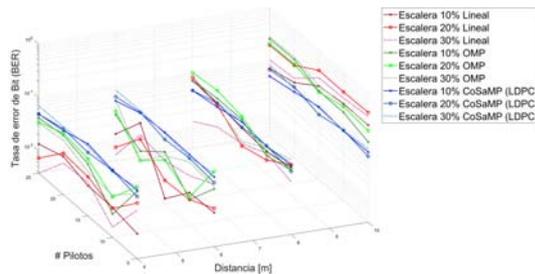


Figure 6. BER Analysis for different distances.

Since the channel decoder uses as input the responses of the channel estimators, the percentage of values that such decoder could not resolve was analyzed in this scenario. In this way, Figure 7 shows the percentage of null values (NaN) in contrast with the distance variations for each channel estimator.

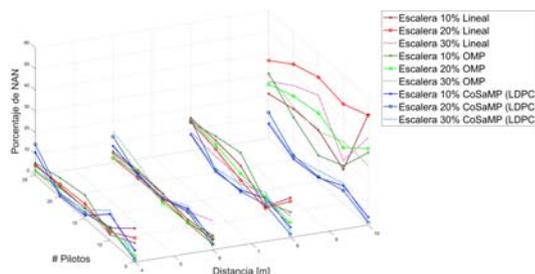


Figure 7. Analysis of the convergence of the channel decoder according to the response of the estimators at different distances.

Figure 7 also confirms the previous results, remarking the same trend in the estimators. The linear estimator works better at short distances, and the CS estimators have better performance than the linear ones for noisy channels, which is verified when the distance is increased.

### 4.3. Analysis of the bandwidth

Figure 8 shows the relationship between the bandwidth used by the pilot subcarriers and the bandwidth for sending information in an OFDM symbol. In this way,

using 5 pilot subcarriers instead of 25 corresponds to an improvement of 13.33% in spectral efficiency. This improvement corresponds to the use of the channel estimators based on CS, as seen in Figure 8.

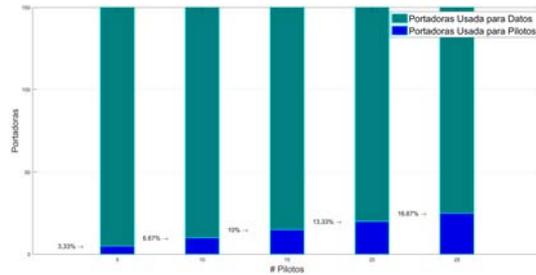


Figure 8. Bandwidth occupied at different number of pilots.

## 5. Conclusions

According to the results obtained, the CoSaMP estimators exhibit a better performance for low levels of SNR with the use of matrix  $\Phi_{LDPC}$ . The OMP estimators show a behavior similar to the CoSaMP, even though with smaller performance. Thereby, with the results obtained it is deduced that, a linear estimator works with the channel estimation problem with very small computational complexity. Nevertheless, when the channel is very noisy, due mainly to the multipath, it demands more pilot subcarriers to try to maintain the performance, which reduces the effective bandwidth. In contrast, the estimators based on CS used in this work, besides maintaining low computational complexity, show an adequate performance for noisy channels enabling data transfer in such channels, considering that an indoor environment was evaluated.

The use of the CS paradigm improves while the IRP and null space properties are guaranteed; in this case, the channel may be considered as sparse because the number of pilot subcarriers is sparse.

In addition, implementing a sensing matrix designed based on the concept of LDPC matrices has enabled to improve the performance of the estimators based on CS, and this improvement is clearer for the estimator that uses OMP. The CoSaMP estimator with  $\Phi_{LDPC}$  showed to be the best estimator in this working situation together with the channel encoder used.

Among the future works, it will be considered to implement channel estimation and prior processing procedures in the FPGA device, to enable that the communication between the equipment and the data source can use a higher data transmission rate.

## References

- [1] J. Sterba and D. Kocur, "Pilot symbol aided channel estimation for ofdm system in frequency selective rayleigh fading channel," in *2009 19th International Conference Radioelektronika*, April 2009, pp. 77–80. [Online]. Available: <https://doi.org/10.1109/RADIOELEK.2009.5158729>
- [2] C. Tzi-Dar and T. Pei-Yun, *OFDM Baseband Receiver Design OFDM Baseband Receiver Design for Wireless Communications*, J. W. . S. A. P. Ltd, Ed., 2007. [Online]. Available: <http://doi.org/10.1002/9780470822500>
- [3] M. R. Raghavendra, S. Bhashyam, and K. Giridhar, "Exploiting hopping pilots for parametric channel estimation in ofdm systems," *IEEE Signal Processing Letters*, vol. 12, no. 11, pp. 737–740, Nov 2005.
- [4] N. N. Jimenez Castro Martínez, "Identificación ciega de canal disperso basado en algoritmos de sub-espacio," Master's thesis, 2016. [Online]. Available: <http://inaoe.repositorioinstitucional.mx/jspui/handle/1009/40>
- [5] M. Cordero Limón, "Técnicas de estimación de canal en la capa física wirelessman-ofdm de la norma ieee 802.16e," 2009. [Online]. Available: <https://bit.ly/2P9Q1co>
- [6] S. Zhang, J. Kang, Y. Song, and N. Wang, "An optimization for channel estimation based on compressed channel sensing," in *2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, Aug 2012, pp. 597–602. [Online]. Available: <https://doi.org/10.1109/SNPD.2012.128>
- [7] Y. Liao, G. Sun, X. Shen, S. Zhang, X. Yang, X. Zhang, H. Yao, and N. Zhang, "Bem-based channel estimation and interpolation methods for doubly-selective ofdm channel," in *2018 IEEE International Conference on Smart Internet of Things (SmartIoT)*, Aug 2018, pp. 70–75. [Online]. Available: <https://doi.org/10.1109/SmartIoT.2018.00022>
- [8] Y. A. Al-Jawhar, K. N. Ramli, M. A. Taher, N. S. Mohd Shah, L. Audah, and M. S. Ahmed, "Zero-padding techniques in ofdm systems," *International Journal on Electrical Engineering and Informatics*, vol. 10, no. 4, pp. 704–725, 2018. [Online]. Available: <https://bit.ly/2YAQGqc>
- [9] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, no. 10, pp. 2908–2914, Oct 1994. [Online]. Available: <https://doi.org/10.1109/26.328961>
- [10] J. van de Beek, M. Sandell, and P. O. Börjesson, *ML estimation of timing and frequency offset in multicarrier systems*, div. of signal processing lulea university of technology s-971 87 lulea, sweden ed. Luleå tekniska universitet, 1996. [Online]. Available: <https://bit.ly/38qaO2W>
- [11] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006. [Online]. Available: <https://doi.org/10.1109/TIT.2006.871582>
- [12] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, July 2007. [Online]. Available: <https://doi.org/10.1109/MSP.2007.4286571>
- [13] R. Baraniuk, M. A. Davenport, M. F. Duarte, and C. Hegde, *An introduction to compressive sensing*, 2011. [Online]. Available: <https://bit.ly/38uCOCG>
- [14] Milliarde. (2016) Compressed sensing intro & tutorial w/ matlab. CODE PROJECT for those who code. [Online]. Available: <https://bit.ly/2YC5Ewa>
- [15] J. Inga-Ortega, E. Inga-Ortega, C. Gómez, and R. Hincapié, "Electrical load curve reconstruction required for demand response using compressed sensing techniques," in *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, Sep. 2017, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/ISGT-LA.2017.8126731>
- [16] M. Pinos and J. Inga, "Predicción de consumo eléctrico en la UPS de Cuenca usando P1P y censado comprimido," 2018.
- [17] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, March 2008. [Online]. Available: <https://doi.org/10.1109/MSP.2007.914731>
- [18] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, vol. 346, no. 9, pp. 589–592, 2008. [Online]. Available: <https://doi.org/10.1016/j.crma.2008.03.014>
- [19] E.-G. Astaiza-Hoyos and H. F. Pablo Emilio Bermúdez-Orozco, "Compressive sensing: A methodological approach to an efficient signal processing," *Dyna*, 2015. [Online]. Available: <https://bit.ly/2qLx8TI>

- [20] H. Yuan, H. Song, X. Sun, K. Guo, and Z. Ju, "Compressive sensing measurement matrix construction based on improved size compatible array ldpc code," *IET Image Processing*, vol. 9, no. 11, pp. 993–1001, 2015. [Online]. Available: <https://doi.org/10.1049/iet-ipr.2015.0117>
- [21] A. G. Dimakis, R. Smarandache, and P. O. Vontobel, "Ldpc codes for compressed sensing," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3093–3114, May 2012. [Online]. Available: <https://doi.org/10.1109/TIT.2011.2181819>
- [22] S. Pawar and K. Ramchandran, "A hybrid dft-ldpc framework for fast, efficient and robust compressive sensing," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2012, pp. 1943–1950. [Online]. Available: <https://doi.org/10.1109/Allerton.2012.6483460>
- [23] H. Wang, W. Du, and Y. Bai, "Compressed sensing based channel estimation for OFDM transmission under 3 GPP channels," *International Journal of Future Generation Communication and Networking*, vol. 9, no. 4, pp. 85–94, 2016. [Online]. Available: <http://dx.doi.org/10.14257/ijfgcn.2016.9.4.08>
- [24] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009. [Online]. Available: <https://doi.org/10.1016/j.acha.2008.07.002>
- [25] B. L. Sturn. (2011) Algorithm power hour: Compressive sampling matching pursuit (cosamp). [Online]. Available: <https://bit.ly/2RS0DOX>
- [26] I. D. F. Lava, "Implementación y evaluación de la codificación ldpc para la transmisión de ficheros en entornos unidireccionales," in *Computer Science*, 2013. [Online]. Available: <https://bit.ly/2YCauJQ>
- [27] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, January 1962. [Online]. Available: <https://doi.org/10.1109/TIT.1962.1057683>
- [28] D. J. C. Mackay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003. [Online]. Available: <https://bit.ly/2E5yfQX>
- [29] G. G. Andrade Salinas, "Implementación de la codificación y decodificación del código LDPC (low density parity check) en MATLAB," 2017. [Online]. Available: <https://bit.ly/34ahKxG>
- [30] V. Roca, C. Neumann, and D. Furodet. (2008) Low density parity check (ldpc) staircase and triangle forward error correction (fec) schemes. [Online]. Available: <https://bit.ly/2RH8PBq>
- [31] E. Seidel, "Overview lte phy: Part 1 - principles and numerology etc," *NOMOR Research*, 2007. [Online]. Available: <https://bit.ly/35cxH7R>
- [32] R. Prieto, A. Abril, and A. Ortega, "Experimental alamouti-stbc using ldpc codes for mimo channels over sdr systems," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, April 2017, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/CCECE.2017.7946842>
- [33] B. Peñafiel and A. Ortega, "A joint multilevel modulator and robust ldpc codes processing over optical systems by providing solutions for holistic 5g backhaul," in *2017 IEEE 13th Malaysia International Conference on Communications (MICC)*, Nov 2017, pp. 197–202. [Online]. Available: <https://doi.org/10.1109/MICC.2017.8311758>
- [34] J. Zyren, "Overview of the 3gpp long term evolution physical layer," NXP Semiconductors, Tech. Rep., 2007. [Online]. Available: <https://bit.ly/35gg8E0>



# PARAMETRIC STUDY FOR OPTIMIZATION OF THE ICE TUBE GENERATOR OF LABORATORY

## ESTUDIO PARAMÉTRICO PARA OPTIMIZACIÓN DE UN GENERADOR DE HIELO TUBULAR DE LABORATORIO

Rocío Guapulema-Maygualema<sup>1,\*</sup>, Víctor Hidalgo-Díaz<sup>2,3</sup>

### Abstract

In Ecuador, there is a large energy consumption by air conditioning and refrigeration in the industrial, commercial and residential sectors. A maximum electricity demand reduction method is to incorporate an optimal cooling thermal energy storage system. The main objective of this work was to develop a parametric study of laboratory ice tube generator for an after optimization. For that, the main parameters were studied, which are: water storage temperature, refrigerant temperature in the evaporator and condenser, ice subcooling temperature and ice formation speed. Two outstanding parameters that intervened in the ice formation process were the place environmental conditions and the water temperature used, when the environmental temperature decreased, thermal load also decreased and condenser efficiency improved, which directly influenced the efficiency of the equipment. The instability observed in the first hour of test intervened in the final water temperature, final temperature ranged from 1.1 °C to -0.4 °C in three hours.

**Keywords:** Energy efficiency, ice, bank, optimization, refrigeration, storage.

### Resumen

En Ecuador existe un gran consumo energético por climatización y refrigeración en los sectores industrial, comercial y residencial. Un método para reducir la demanda eléctrica máxima es incorporar un sistema óptimo de almacenamiento de energía térmica de refrigeración. Este trabajo tiene por objetivo realizar el estudio paramétrico de un generador de hielo tubular de laboratorio para su posterior optimización. Se estudiaron los principales parámetros que intervienen en la formación de hielo como temperatura del agua en el reservorio, temperaturas del refrigerante en el evaporador y condensador, temperatura de subenfriamiento del hielo y velocidad de formación de hielo. Los parámetros destacados que intervinieron en el proceso fueron las condiciones ambientales del lugar y la temperatura del agua utilizada. Al disminuir la temperatura ambiental, disminuye la carga térmica y mejora la eficiencia del condensador, el cual influye directamente sobre la eficiencia del equipo. La inestabilidad observada en la primera hora de prueba intervino en la temperatura final del agua, la misma que varió en el rango de 1,1 °C a -0,4 °C en 3 horas.

**Palabras clave:** eficiencia energética, generador, hielo, optimización, refrigeración, almacenamiento.

<sup>1,\*</sup>Faculty of Land Sciences and Water, Universidad Regional Amazónica Ikiam, Tena 150150, Ecuador. Corresponding author ✉: [rocio.guapulema@ikiam.edu.ec](mailto:rocio.guapulema@ikiam.edu.ec), [r.guapulema@hotmail.com](mailto:r.guapulema@hotmail.com)

<http://orcid.org/0000-0002-2768-9018>

<sup>2</sup>Department of Mechanical Engineering, Escuela Politécnica Nacional, Quito 170525, Ecuador.

<http://orcid.org/0000-0003-4732-2421>

<sup>3</sup>Laboratory of Informatics-Mechanics, Escuela Politécnica Nacional, Quito 170525, Ecuador.

Received: 24-10-2019, accepted after review: 09-12-2019

Suggested citation: Guapulema-Maygualema, R. and Hidalgo-Díaz, V. (2020). «Parametric Study for Optimization of the Ice Tube Generator of Laboratory». INGENIUS. N.º 23, (january-june). pp. 86-96. DOI: <https://doi.org/10.17163/ings.n23.2020.08>.

## 1. Introduction

Nowadays, saving energy and protecting the environment are important issues worldwide [1]. On the other hand, the storage of thermal energy is considered one of the most important energy technologies [2] for reducing electrical consumption. Studies on ice generators have been carried out with the purpose of evaluating and improving its operation for reducing electrical consumption; parameters evaluated such as thermal conductivity of the material of the condenser, tank capacity, longitudinal and cross pitches of the condenser tube and coolant temperatures [3], directly affect the design of the equipment, space requirement and efficiency of the system. Similarly, energy losses due to ice storage, cost structure and electricity tariffs, periods of tariffs, storage capacity of ice generators and impact of load prevision, have been also evaluated [4].

A model for the optimization of air conditioning systems with thermal accumulation and system simulation results are presented in [5]. Considering that the growing demand of air conditioning systems has led to a greater energy consumption during peak hours [6], thermal energy storage systems are becoming increasingly popular [7], enabling to conduct studies about physical, technical, economic, environmental [8,9] and raw material consumption [10] aspects of thermal energy storage refrigeration systems and their applications [2], with the purpose of finding a method to improve their efficiency [11].

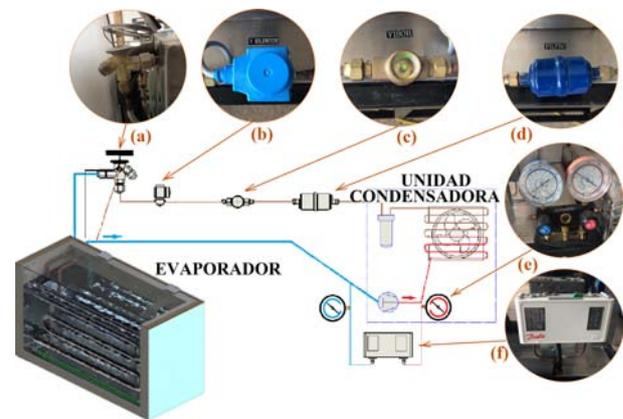
In Ecuador, the residential, industrial and commercial sectors represent the 31 %, 25.9 % and 15.9 %, respectively, of the energy demand [12]; approximately 50 % of the energy consumption in the residential and commercial sectors of the Costa and Sierra regions is due to cooling and conditioning [13]. The National Balance of Electric Energy presented in June 2019 by the Agency for Electricity Regulation and Control (in Spanish ARCONEL, Agencia de Regulación y Control de Electricidad), indicates that the electricity consumption by these sectors was 17736.87 GWh [12]. A method for reducing the maximum demand of electricity is incorporating storage of thermal energy for refrigeration [14], in order to transfer the energy consumption at peak hours to the hours of low consumption [7, 10, 15]. The thermal energy may be stored using sensible or latent heat [16, 17], namely, cold water or ice, the latter being advantageous since it requires smaller storage volumes [18, 19].

With the purpose of saving energy and optimizing processes which use storage of refrigeration energy, namely conditioning, dairy processing, breweries, plastics manufacturing and chemical processes, among others, it is important to simulate the efficiency that an equipment for generating and storing ice would have. A parametric study was carried out for the optimization of the tubular ice generator constructed in the

project entitled Design and construction of a laboratory tubular ice generator [20]. This generator, which corresponds to a system of ice on evaporator, consists of an arrangement of tubes submerged in a tank and through which cold coolant is pumped to freeze the water around them [21]. The parameters that intervene in the process of ice formation and storage of thermal energy for cooling, were studied by means of the mathematical modeling of the thermal and heat transfer cycle. A program was created in the Guide environment of Matlab for numerically simulating the operation of the equipment in different conditions, while an experimental study was conducted to assess the numerical simulation. The most important parameters that modify the operation and efficiency of the generator were defined through this study.

The parameters under consideration for the study include the water storage temperatures in the generator, temperatures of the coolant in the evaporator and condenser, ice subcooling temperature, ice thickness and speed of releasing and extracting the heat in a natural manner. It was important to study the generation of ice, taking into account that as the thickness increases, the resistance of heat transfer also increases [22, 23].

## 2. Materials and methods



**Figure 1.** Laboratory tubular ice generator and its constituting elements: a) expansion valve, b) solenoid valve, c) liquid viewer, d) drying filter, e) pressures gauges, f) pressure switch

The equipment used for the parametric study is constituted by a condensing unit L'UNIQUE HERMETIQUE CAJ2428ZBR of capacity 1/2 HP, which uses R404a as coolant, a TES 2-0.45 expansion valve, a A-TD -132 SAE filter, a SGI-10 liquid viewer, a TR ¼ 032F8107 solenoid valve, pressure switch, pressure gauges, control board, tank-reservoir set, air system for water agitation and an evaporator housed in the tank (Figure 1), which is constituted by two type L copper coils with a diameter of 15.88 mm (1/2 in.),

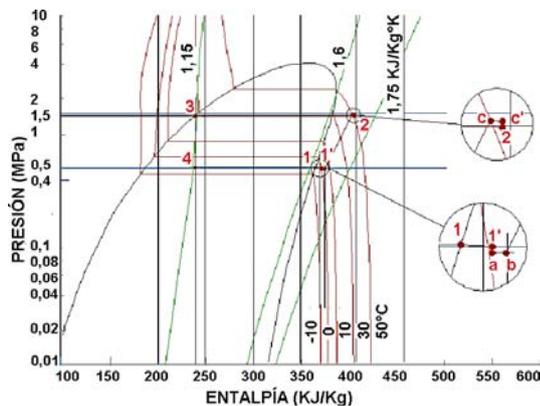
and each with a length of 3 m. The spacing between the coils is 120 mm.

In addition, the elements in the data acquisition system include the control interface, LM 35 sensors, Arduino UNO boards, computer and software for data acquisition.

## 2.1. Mathematical modeling of the ice generation

### 2.1.1. analysis of the thermal cycle

The process for the study and evaluation of the cooling cycle and its efficiency is indicated in [24]. This information was considered as the base for the mathematical modeling of the cooling cycle.



**Figure 2.** Pressure-Enthalpy Diagram of the tubular ice generator.

Four processes can be observed in the Pressure-Enthalpy diagram (Figure 2). The coolant that ideally can be found as saturated steam at a temperature  $T_1 = -0.5^\circ\text{C}$ , is compressed and reheated until it reaches condensation conditions at a pressure  $P \approx 1,416 \text{ MPa}$  and at a temperature  $T \approx 28,9^\circ\text{C}$ , and subsequently condensed without temperature variation. Afterwards it passes through an expansion valve, where the pressure and temperature decrease and the enthalpy is maintained, later absorbing heat in the reservoir and changing its state from liquid to gas.

The real enthalpy  $h'_1$  is calculated using the equation of the change of enthalpy at constant pressure shown in Equation 1.

$$h'_1 = h(T'_1, P_1) = h(T_1, P_1) = h(T_1, P_1) + \int_{T_1}^{T'_1} C_{p_{ref}} dT \quad (1)$$

where  $T'_1$  and  $P_1$  represent the real temperature and the pressure at point 1, respectively, and  $C_{p_{ref}}$  is the specific heat of the coolant.

The pressure drop in the suction valve towards the compressor is evaluated at point a, and the enthalpy is calculated by means of the equation of change

of enthalpy at constant temperature, as indicated in Equation 2.

$$h_a = h(T'_1, P'_1) + \int_{P'_1}^{P_a} \left[ v_a - T_a \frac{R_{ref}}{P'_1} \right] dT \quad (2)$$

where  $R_{ref}$  is the constant of the coolant,  $P'_1$  is the real pressure at point 1, and  $v_a$  and  $T_a$  are the specific volume of the coolant and the temperature at point a, respectively.

It is considered a pressure drop at the outlet of the condenser, due to the effect of the friction of the coolant as it passes through this component. Equation 3 is used for calculating the enthalpy of the coolant on the line of constant entropy, considering an isentropic compression process.

$$h_c = c_1 + c_2(T_c - T_b) + c_3(T_c - T_b)^2 + c_4(T_c - T_b)^3 + h_b \quad (3)$$

where,  $c_1, c_2, c_3, c_4$  are coefficients calculated according to the ASHRAE tables of the R404a coolant [25],  $T_b$  and  $T_c$  are saturation temperatures of the coolant as vapor at pressures and , and is the enthalpy evaluated through the contact of the coolant with the internal surface of the compressor. For points 1, b, c', 2, 3 and 4 in Figure 2, the calculation of the enthalpies is carried out according to the table of properties of the coolant.

The refrigeration efficiency of the system, the performance coefficient as a function of the useful cooling effect and the net energy supplied by the compressor are evaluated, while the amount of coolant supplied in the equipment is calculated as a function of the capacity  $\dot{Q}$  of the condensing unit, as observed in Equation 4.

$$\dot{m}_{refrigerante} = \frac{\dot{Q}}{h_b - h_4} \quad (4)$$

### 2.1.2. Analysis of heat transfer

Three basic situations were considered for the model:

- 1) Flow of environmental heat into the tank: The heat flows from a hot source to a cold source [26, 27], in other words, if the tank contains water at low temperature and ice on the evaporator, the heat will be transferred by free convection of the air towards the external walls of the tank, this being the first stage of the phenomenon. In the second stage, heat is transferred by conduction from the external walls of the tank to the isolation layer, and subsequently to the internal walls until reaching the water; here the transfer is studied as free convection.

- 2) Cooling of the water without formation of ice: Heat is transferred from the water to the coolant in two stages: transfer of free heat between the water and the outer surface of the evaporator, and heat transfer from the outer surface to the coolant.
- 3) Formation of ice on the evaporator: Two stages are supposed in the freezing phase; the first describes the heat transfer from the water to the ice layer, while the second stage considers the transfer of heat from the superficial layer of ice to the coolant fluid.

The data necessary for evaluating the heat transfer are taken based on the filmic temperature calculated as a function of the surface temperature  $T_s$ , and the temperature of the fluid  $T_\infty$ .

The increase of heat due to the environment is deduced as a function of the global coefficient of heat transfer, the area  $A$  of heat transfer and the temperature variations between the outside air and the water,  $T_{\infty outside\_air}$  and  $T_{mW}$ , respectively.

$$\dot{Q}_{pared} = \frac{T_{\infty outside\_air} - T_{mW}}{\frac{1}{h_{\infty a}} + \frac{2e_{ac}}{k_{ac}A} + \frac{e_p}{k_p A} + \frac{1}{h_{\infty W}A}} \quad (5)$$

The flow of heat that goes through the wall of the tank is defined by Equation 5, where  $e_{ac}$  is the thickness of the 304 stainless steel sheet,  $e_p$  is the thickness of the polyurethane,  $k_{ac}$  is the conductivity coefficient of the 304 stainless steel,  $k_p$  represents the conductivity coefficient of the polyurethane, and  $h_{\infty a}$  and  $h_{\infty w}$  are the convective coefficients of the outside air and the water, respectively.

For studying the cooling of water without formation of ice, it was evaluated the coefficient  $h_{\infty ref}$  of boiling forced convection, which is equal to the maximum value between the boiling coefficient  $h_{NBD}$  of the regions of dominant nucleate boiling, and the boiling coefficient of the regions of dominant convective boiling, calculated through Equations 6 and 7.

$$h_{NBD} = [0,6683Co^{-0,2}f_2(Fr_{lo}) + 1058Bo^{0,7}F_{fl}] \cdot (1-x)^{0,8}h_{lo} \quad (6)$$

$$h_{CBD} = [1,136Co^{-0,9}f_2(Fr_{lo}) + 667,2Bo^{0,7}F_{fl}] \cdot (1-x)^{0,8}h_{lo} \quad (7)$$

It was taken into account the convection number  $Co$ , the coolant friction factor  $f_2$ , the Froude number  $Fr_{lo}$ , the boiling point  $Bo$ , the surface-fluid association parameter  $F_{fl}$ , the quality  $x$  of the coolant fluid,

the heat transfer coefficient  $h_{lo}$ ; the way to calculate these coefficients is presented in [28].

Two thermal loads were determined,  $\dot{Q}_{S1}$  that takes into account only the thermal resistance of the water, and  $\dot{Q}_{S2}$  which considers the thermal resistances of the evaporator walls and of the coolant. The two thermal loads are defined by Equations 8 and 9.

$$\dot{Q}_{s1} = 2\pi r_e L h_{\infty w} (T_{mw} - T_{s1}) \quad (8)$$

$$\dot{Q}_{s2} = \frac{T_{s1} - T_1}{\frac{\ln\left(\frac{r_e}{r_i}\right)}{2\pi k_c L} + \frac{1}{2\pi r_i L h_{\infty ref}}} \quad (9)$$

where  $L$  represents the length of the evaporator tube,  $r_i$  and  $r_e$  are the internal and external radii of the evaporator,  $k_c$  is the thermal conductivity coefficient of the copper and  $T_{s1}$  is the temperature of the water-evaporator surface layer.

The thermal load  $\dot{Q}_{s3}$  due to the heat transfer from the water to the ice surface layer (Equation 10), and the thermal load  $\dot{Q}_{s4}$  due to the heat transfer from the ice surface to the coolant (Equation 11), are calculated in the phase of formation of ice on the evaporator.

$$\dot{Q}_{s3} = 2\pi(r_e + 2e)Lh_{\infty w2}(T_{mw} - T_o) \quad (10)$$

$$\dot{Q}_{s4} = \frac{T_0 - T_1}{\frac{\ln\left(\frac{r_e+e}{r_e}\right)}{2\pi k_h L} + \frac{\ln\left(\frac{r_e}{r_i}\right)}{2\pi k_c L} + \frac{1}{2\pi r_i L h_{\infty ref}}} \quad (11)$$

where  $e$  represents the thickness of the ice formed in a time  $t$ ,  $h_{\infty w2}$  is the convective coefficient of the water used in the heat transfer water-evaporator,  $T_{mw}$  is the arithmetic mean of the temperatures as a function of the initial and final temperatures of the water,  $T_{0\_water}$  and  $T_{f\_agua}$ , respectively, and  $k_h$  is the conductivity coefficient of the ice.

The time necessary to cool the water of the reservoir from an ambient temperature to a required temperature, varies directly with respect to  $\dot{Q}_{s1}$ . Equation 12 enables the calculation of the cooling time of the water in hours, considering the mass  $m_{water}$  of the water in the tank in kg, and the specific heat of the water  $Cp_{agua}$  in  $kJ/kg \text{ } ^\circ C$ . The net load  $\dot{Q}_{neto}$  that the system has for water cooling and ice generation, is equal to the difference between the capacity of the condensing unit and the total load lost.

$$t_{cool\_water} = \frac{1000m_{water}Cp_{water}(T_{0\_water} - T_{f\_water})}{3600(\dot{Q}_{net} - \dot{Q}_{s1})} \quad (12)$$

The total thickness of the ice generated in a time  $t$  (Equation 13) is determined as a function of the flow  $Q_{lat}$  of latent heat. This value results from the difference between the thermal load due to the heat transfer from the water to the surface layer of ice, and the thermal load from the ice surface to the coolant.

$$e = \frac{\sqrt{\frac{4t(Q_{lat})}{1000\pi\rho_h L h_{sf}} + D_e^2} - D_e}{2} \quad (13)$$

where  $D_e$  is the external diameter of the evaporator in meters,  $\rho_h$  is the ice density in  $kg/m^3$ ,  $h_{sf}$  is the latent heat of water fusion in  $kJ/kg$ . Meanwhile the mass  $m_h$  of ice generated is determined through Equation 14.

$$m_h = \frac{\pi}{4} ((D_e + 2e)^2 - D_e^2) L \rho_h \quad (14)$$

At an initial time the water in the tank is at ambient temperature, and with the operation of the equipment the temperature decreases until reaching a value equal or close to  $0^\circ C$ . The final temperature that the water will reach in a particular time, is calculated as a function of the difference between the net load and the load estimated by the thermal resistance  $\dot{Q}_{s1}$  of the water only (Equation 15).

$$T_{f\_water} = T_{0\_water} - \frac{t(\dot{Q}_{net} - \dot{Q}_{s1})}{m_{water} C_{p\_water}} \quad (15)$$

$$T_{f2\_water} = T_{0\_water} - \frac{t(\dot{Q}_{net} - \dot{Q}_{s3})}{m_{water} C_{p\_water}} \quad (16)$$

On the other hand, Equation 16 is used for evaluating the water temperature considering the existence of ice on the evaporator, for which it is utilized the load  $\dot{Q}_{s3}$  calculated due to the thermal resistance of the water in the freezing process.

## 2.2. Numerical simulation of the tubular ice generation

The program created in the Guide environment of Matlab enables the continuous modification of the data of ambient temperature and pressure, and water and coolant necessary for the thermodynamic and heat transfer analysis. Guide enabled the graphical design of the interface of the design editor, while the Matlab program to build it was automatically generated. When the changes carried out on the interface are saved, two files are created, one .fig and the other .m; the lines of code contained in the second file are the ones that create the interface in the .fig file. The program to be executed was inserted in the command zone of the push button «Calculate».

The programming sequence involves the input of type text data, which then will be converted to type double to use them in the solution of equations present in the programming, the calculation of parameters to be evaluated and the visualization of results in tables and in a graph showing the Pressure-Enthalpy diagram defined in the thermodynamic cycle. The equations

evaluated to simulate the operation of the ice generator were defined in the mathematical modeling. The interface created for the process comprises iterative cycles, for approximation of results and evaluation of the phenomenon at different values of fluid quality and of thickness of the generated ice.

Three operating scenarios were posed for the simulation, based on the morning, afternoon and night hours. This information enabled the creation of thickness variation curves, velocity of ice formation, temperature of the water in the tank and temperature of the surface layer in the evaporator.

## 2.3. Experimental analysis

The procedure utilized for starting-up the equipment and the data acquisition are detailed in Figure 3. The initial operating conditions of the equipment were posed when the working pressure was stabilized.

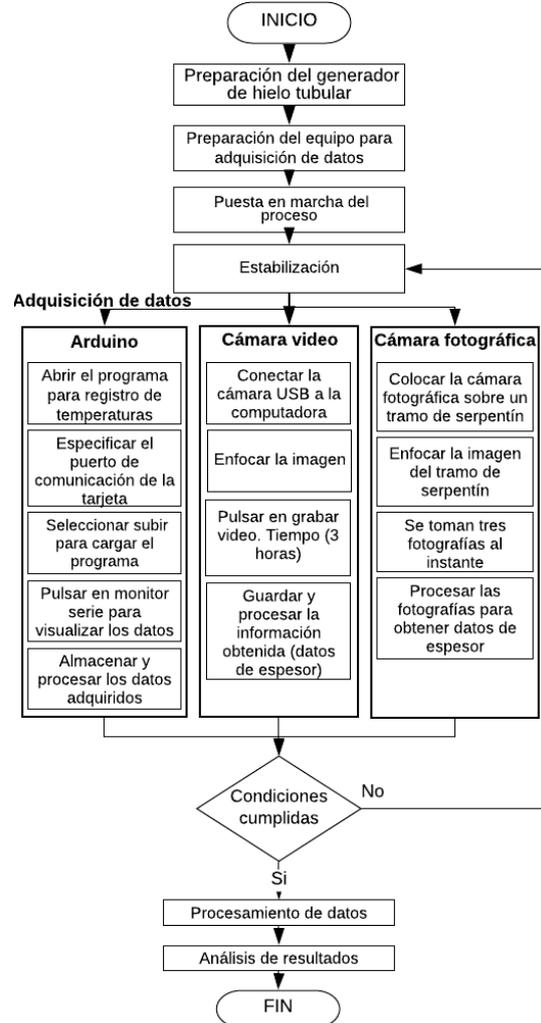


Figure 3. Process diagram of the experimental analysis.

The values of temperature at the inlet and outlet of the evaporator, at the outlet of the condenser, at the center of the reservoir and at distances of 3, 12

and 25 mm of the water-evaporator surface layer were registered. The growth of the thickness of ice was also registered using a photographic camera.

### 3. Results and discussion

Technical data of manufacturers and registered operating values were necessary for calculation and discussion of results. The low and high operating pressures of the system were 0.5 and 1.4 MPa (75 and 203 psig), respectively. It was determined that the coolant reaches the evaporator as a mixture of liquid and saturated steam, with a quality of 29 %. The working temperatures were  $-5\text{ }^{\circ}\text{C}$  in the evaporator and  $28.9\text{ }^{\circ}\text{C}$  in the condenser.

Through simulation the operating parameters were evaluated, indicating the following:

- The generation of ice in a particular time for the different experiences, started when the layer of fluid closest to the evaporator cooled until reaching the temperature of phase change, which was observed as a displacement greater than 0.2 h at the origin of each curve.
- The accelerated increase in the velocity of formation of ice and its further stabilization, represents the fast cooling that occurs in the water layers closest to the evaporator tube and its resulting freezing; lower velocities were registered as the thickness of ice was increased, with this velocity being greater in the first 1.5 h.

- The temperature of the water decreased at a rate of  $7\text{ }^{\circ}\text{C}/\text{h}$  during the first two hours; beyond that time, this value decreased considerably until reaching  $0.5\text{ }^{\circ}\text{C}/\text{h}$  after three hours of operation. With low ambient temperatures, the time of ice generation was smaller.
- The change in the temperature of the surface layer in the evaporator showed two variations, the first corresponding to the rate of change of the sensible heat, until reaching a temperature of  $0\text{ }^{\circ}\text{C}$ . Afterwards, the variation tends to stop for a range of 0.25 h; this phenomenon is due to a change of phase of the water in the zones closer to the wall of the tube. The subcooling of ice starts past the time of change of phase, until finally stabilizing at  $-7\text{ }^{\circ}\text{C}$ .
- An average cooling efficiency of 70 % and an average Coefficient of Performance (COP) of 5.5 were registered through numerical simulation.

Values of temperatures and thickness of ice along time were collected experimentally. Main data of the representative experiences considered for the study are detailed in Table 1.

Practices 1, 2, 3 and 4 were carried out considering natural convection of the water. On the other hand, practice 5 was carried out considering forced convection. Figure 4 shows the process of formation of ice observed in practice 2. The experimental results obtained are summarized in the following.

**Table 1.** Main data of the experiences under study

N.º Practice	Environment temperature $^{\circ}\text{C}$	Initial temperature of water $^{\circ}\text{C}$	Variation of temperature of condenser		Variation of pressure of condenser	
			$^{\circ}\text{C}$	$^{\circ}\text{F}$	MPa	psi
1	20,5	16	11,3	20,34	0,028	4
2	22,5	21,5	11	19,8	0,034	5
3	20	18	8,6	15,48	0,028	4
4	21,5	19,4	8,8	15,84	0,034	5
5	20	17	10,3	18,54	0,034	5

- No ice was formed in the time interval between 0 and 8 minutes. During this period, the layer of fluid closest to the evaporator cooled and then changed phase.
- The low temperature conditions resulted in a larger velocity of formation of ice, obtaining thicknesses of ice between 22 and 25 mm in 3 hours of operation.
- In an operation time greater than 3 hours, the generation of ice decreased considerably; after

5 hours of operation, the thickness of ice was approximately 35 mm, as observed in Figure5.

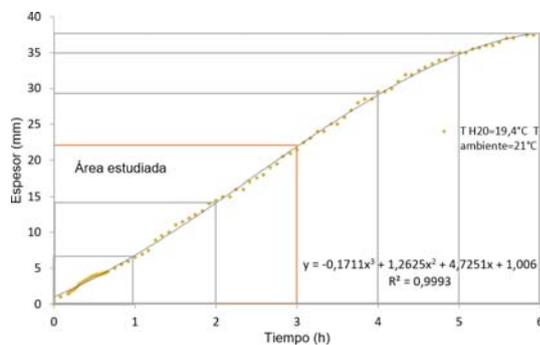
- The operation of the equipment in forced convection showed a faster water cooling and smaller growth of ice, compared with the results in practices in free convection.
- In a period of 5 to 10 minutes after the practice started, with a temperature  $2\text{ }^{\circ}\text{C}$  smaller than the initial temperature, begins the change of phase of the fluid film closest to the evaporator. The

velocity of ice generation varied between 3 and 9 mm/h.

- The temperature of the coolant in the evaporator descends from  $-3\text{ }^{\circ}\text{C}$  to  $-5\text{ }^{\circ}\text{C}$ , indicating the variation of the refrigeration capacity of the equipment according to the demand at that instant.



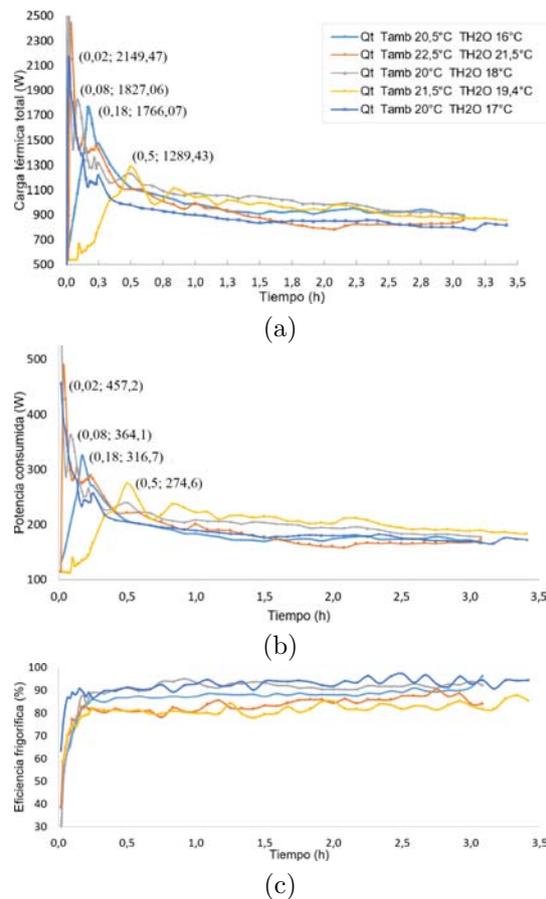
**Figure 4.** Thickness of ice at different time intervals. Practice 2.



**Figure 5.** Thickness of ice vs. Time T h20=19.4 °C.

- The temperature of the coolant at the outlet of the condenser stabilizes after 15 minutes of turning on the equipment, in a range from  $28\text{ }^{\circ}\text{C}$  to  $34\text{ }^{\circ}\text{C}$ . This value is a function mainly of the ambient temperature, temperature of the water and high pressure of the system considering the existing losses.

- The value of the total thermal load was obtained as the sum of the sensible heat and the latent heat for a particular time instant, considering the data of temperature of the water in the tank, the mass of ice and the mass of water; this value reaches an equilibrium point when the slopes of the curves of flow of sensible and latent heat have approximate values, namely, when the thermal load necessary to cool the water decreases in the same proportion in which the thermal load increases for ice generation and subcooling. It was no observed a significant variation of the thermal load past 1.5 h of operation. After 3 hours, the total thermal load varies in the range from 858 W to 892 W. In Figure 6 a), the two highest peaks (greatest thermal loads) are the result of experiences carried out in the afternoon, while the two following peaks are the result of experiences carried out at night, indicating a direct relationship with the ambient temperature and the temperature of the water.



**Figure 6.** Experimental results. a) Total thermal load vs. time. b) Power consumed by the compressor vs. time. c) Refrigeration efficiency vs. time.

- The values of thermal load obtained between 0 and 0.5 hours are greater than the obtained 3

hours later. The peaks observed in the curves are due to the requirement of more refrigeration capacity for cooling water. At 10 minutes of operation of the equipment, 90 % of the total required refrigeration capacity corresponds to the heat necessary for water cooling, but at 3 hours of equipment operation, 40 % of the required refrigeration capacity is used for water cooling.

- The consumption of the compressor in 3 hours of operation varied in the range from 165 W to 173 W, to remove a thermal load between 859 W and 892 W. The curves shown in Figure 6 b) indicate the variation of the power consumed by the compressor according to the required refrigeration capacity.
- For each watt consumed by the compressor, approximately 5 W of thermal energy were removed from the water in the reservoir tank.
- The COP observed for different ambient temperatures and temperatures of the water varies in the range from 5.8 to 6.1, while the refrigeration efficiency fluctuates between 78 % and 95 %.
- The smallest thermal load removed was 794.11 W, obtained for the experiment carried out at an initial temperature of the water of 17 °C and ambient temperature of 20 °C in forced convection; for this experiment, the final temperature of the water was 0 °C, the COP obtained was 5.92 and the refrigeration efficiency was 90.7 % (Figure 6c). If these values are compared with the results obtained in the experiment carried out with the lowest initial temperature of the water (natural convection), the power absorbed, COP and refrigeration efficiency are very similar, whilst, when setting equal the values of final temperature of water and thickness of ice, there are important differences of -0,4 °C and 5 mm, respectively. As a consequence, using forced convection in the water improved the refrigeration efficiency of the equipment, as indicated in [11], but operating the equipment in conditions such as the observed in the practice carried out at an initial temperature of water of 16 °C resulted the best option.

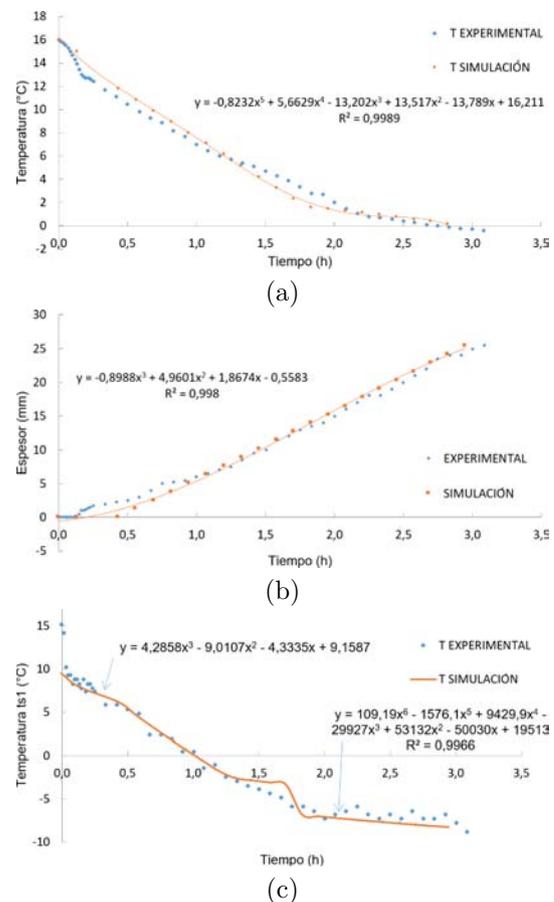
A validation of the experimental results was carried out with the purpose of observing the percentage of error of the data generated with the model developed. Data from practice 1 were considered.

Figure 7a presents the variation of the water temperature water obtained by simulation and experimentally. The curve obtained by simulation is very similar to the temperature data obtained in the experiment. A water temperature of 12.8 °C was obtained 10 minutes

after the experiment was initiated, while by simulation it was 14 °C; this is the largest variation observed when comparing the data. The curve obtained by simulation follows a polynomial trend of fifth order, with an R2 coefficient of 0.99.

Curves of data of thickness of ice obtained by simulation and experimentally are shown in Figure 7b. It can be observed a variation in the first hour of operation, with a maximum difference of thickness of ice of 2 mm; this is because the generation of ice started 10 minutes after the equipment was turned on, while by simulation it started half hour later. The curve follows a polynomial trend of third order, with an R2 coefficient of 0.998.

On the other hand, Figure 7c presents data of variation of temperature of the surface layer between the evaporator and the water ( $T_{s1}$ ), obtained by simulation and experimentally. Two equations need to be used to define the change in this temperature, the first defining the variation of temperature in the cooling and change of phase, while the second equation indicates subcooling. The curves presented follow polynomial trends of third and sixth order, with an R2 coefficient of 0.996.



**Figure 7.** Data of temperature and thickness of ice obtained by simulation and experimentally: a) Temperature of the water vs. time. b) Thickness of ice vs. time c) Temperature of the surface layer vs. time.

A particular situation occurs with the plot of velocity of formation of ice in time shown in Figure 8. Experimentally, the thickness of ice increases at larger velocity and stabilizes after one hour, while by simulation a velocity increasing in time was expected, starting after half hour of equipment operation. This difference is due to instabilities in the process of formation of ice. After one hour has elapsed, the two curves get closer indicating equality. From the third hour of equipment operation, the velocity of formation of ice tends to remain with minimum variations (approximately 1 mm per hour)

Table 2 summarizes the results obtained in the different experiments during 3 hours. It was determined the error in the simulation data (S) with respect to the obtained experimentally (E), remarking the following:

The differences observed when comparing the data of final temperature of water and thickness of ice obtained by simulation and experimentally, are due to the instability occurred in the process of formation of ice during the first hour of experiment. For the final temperature of the water, the percentage of error calculated is in the range from 1.78 % to 80 %, which is

large because the values obtained by simulation indicated smaller temperatures; the mathematical model considered a constant variation of temperature in the condenser, which is very similar to the variation of temperature observed in the tests carried out during the morning and the afternoon, but in tests performed at night the variation of temperature decreased significantly obtaining final temperatures of water higher than expected. For the data of thickness of ice generated, the percentage of error calculated varies in the range from 0.39 % to 15.24 %.

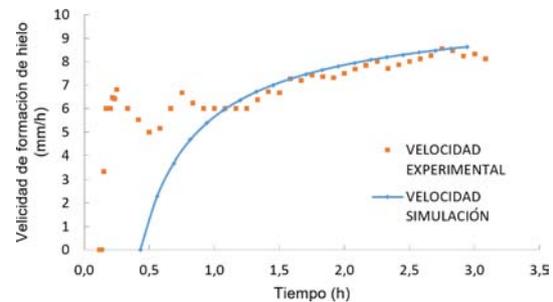


Figure 8. Velocity of formation of ice vs. time.

Table 2. Results of calculations carried out for the different practices in 3 hour

Parameters	Values											
	E	S	% Error	E	S	% Error	E	S	% Error	E	S	% Error
Initial temperature of water (°C)	16 °C			21,5 °C			18 °C			19,4 °C		
Final T of water (°C)	-0,4	-0,41	1,78	1,1	0,93	15,45	-0,5	-0,1	80	1	1,78	78
Ts1 (°C)	-7,82	-8,27	5,75	-7,33	-5,87	19,99	-6,35	-7,59	19,45	-7,82	-7,67	1,92
Thickness (mm)	25,5	25,4	0,39	21	17,8	15,24	23,5	22,9	2,55	23	22,9	0,43
Velocity of formation of ice (mm/h)	8,27	8,63	4,32	6,81	5,99	12,12	7,62	7,59	0,41	7,26	7,52	3,54
Total Q (W)	905,38	1139,28	25,83	858,55	1139,93	32,77	901,05	1140,45	26,57	869,93	1139,93	31,04
Work by compressor (W)	168,64	163,93	2,79	172,69	204,2	18,25	178,58	163,93	8,21	185,95	163,93	11,84
Energy consumed (Wh)	519,99	491,8	5,42	532,45	612,6	15,05	550,64	491,79	10,69	588,85	491,79	16,48
COP	5,97	7,02	17,7	5,85	5,63	3,7	5,94	7,02	18,26	5,85	7,02	20,04
Cooling efficiency (%)	91,39	74,22	18,79	84,02	71,27	15,17	91,87	74,22	19,21	86,19	74,22	13,88
Time at 1 °C (h)	2,17	2,2	1,7	3,08	3,17	2,81	2,5	2,36	5,6	3,08	2,66	13,73
Thickness at 1 °C (mm)	17	17,8	4,71	21	21	0	19	18,1	4,74	22,5	19,77	12,13

A relevant difference is observed in the percentage of error of the refrigeration capacity to be removed, obtained by simulation and experimentally. Since the value given in the product catalog was used in the simulation, this is an optimal refrigeration capacity for selection, but it is not the real one; an important difference is observed when compared with the real value, since the percentage of error varies from 25.83 % to 32.77 %.

The percentages of error evaluated for the velocity of formation of ice, work of the compressor, energy consumed, COP and refrigeration efficiency did not exceed 20 %. For the conducted parametric study, the margin of error obtained is acceptable.

## 4. Conclusions

It was developed a mathematical model based on the study of the thermal cycle and existent heat transfer, and a programming code for simulation in Guide of MATLAB. The program created enables varying the data of the equipment, operation, ambient and working fluid, with the purpose of observing the change in efficiency, power consumed and its interaction in time. The curves obtained by simulation were validated comparing them with the curves obtained with experimental data, and evaluating the percentage of error between these results.

The main parameters that intervened in the process of formation of ice and storage of thermal energy, were the ambient conditions of the place where the

equipment was installed and the temperature of the water used. The equipment installed in Riobamba decreased its refrigeration capacity to the one presented in the catalog, due to the barometric pressure of the place and the ambient temperature in sunny days.

The instability observed in the process of formation of ice during the first hour of the experiment, directly intervenes in the final temperature of the water. As the formation of ice starts 15 minutes before the time observed by simulation, more heat than the calculated was removed, obtaining a final temperature of water smaller than the observed by simulation. The final temperature of the water varied in the range from 1.1 °C to -0.4 °C in 3 hours.

The greatest efficiency found and the smallest energy consumption were observed in experiments carried out at night, when the ambient temperatures decrease, thus decreasing the thermal load and improving the efficiency of the condenser, which directly influences on the efficiency of the equipment. For operation at night with the water temperature equal to 16 °C.

## References

- [1] A. Saito, "Recent advances in research on cold thermal energy storage," *International Journal of Refrigeration*, vol. 25, no. 2, pp. 177–189, 2002. [Online]. Available: [https://doi.org/10.1016/S0140-7007\(01\)00078-0](https://doi.org/10.1016/S0140-7007(01)00078-0)
- [2] I. Dincer, "On thermal energy storage systems and applications in buildings," *Energy and Buildings*, vol. 34, no. 4, pp. 377–388, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778801001268>
- [3] J. H. M. Neto and M. Krarti, "Parametric analysis of an internal-melt ice-on-coil tank," *ASHRAE*, vol. 103, no. 2, pp. 322–333, 1997. [Online]. Available: <https://bit.ly/2PyxJ3r>
- [4] S. Sanaye and A. Shirazi, "Thermo-economic optimization of an ice thermal energy storage system for air-conditioning applications," *Energy and Buildings*, vol. 60, pp. 100–109, 2013. [Online]. Available: <https://doi.org/10.1016/j.enbuild.2012.12.040>
- [5] O. J. Venturini, M. d. S. Valente de Almeida, and E. Silva, "Optimización de un sistema de termoacumulación en un tanque de hielo con expansión directa," *Asociacion Brasileña de Ingeniería y Ciencia Mecánicas*, 1999. [Online]. Available: <http://bit.ly/2WaawaN>
- [6] M. H. Rahdar, A. Emamzadeh, and A. Ataei, "A comparative study on pcm and ice thermal energy storage tank for air-conditioning systems in office buildings," *Applied Thermal Engineering*, vol. 96, pp. 391–399, 2016. [Online]. Available: <https://doi.org/10.1016/j.applthermaleng.2015.11.107>
- [7] Z. Kang, R. Wang, X. Zhou, and G. Feng, "Research status of ice-storage air-conditioning system," *Procedia Engineering*, vol. 205, pp. 1741–1747, 2017, 10th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC2017, 19-22 October 2017, Jinan, China. [Online]. Available: <https://doi.org/10.1016/j.proeng.2017.10.020>
- [8] P. A. Intemann and M. Kazmierczak, "Heat transfer and ice formations deposited upon cold tube bundles immersed in flowing water. i. convection analysis," *International Journal of Heat and Mass Transfer*, vol. 40, no. 3, pp. 557–572, 1997. [Online]. Available: [https://doi.org/10.1016/0017-9310\(96\)00121-4](https://doi.org/10.1016/0017-9310(96)00121-4)
- [9] M. H. Rahdar, M. Heidari, A. Ataei, and J.-K. Choi, "Modeling and optimization of r-717 and r-134a ice thermal energy storage air conditioning systems using nsga-ii and mopso algorithms," *Applied Thermal Engineering*, vol. 96, pp. 217–227, 2016. [Online]. Available: <https://doi.org/10.1016/j.applthermaleng.2015.11.068>
- [10] J. Pu, G. Liu, and X. Feng, "Cumulative exergy analysis of ice thermal storage air conditioning system," *Applied Energy*, vol. 93, pp. 564–569, 2012, (1) Green Energy; (2) Special Section from papers presented at the 2nd International Energy 2030 Conf. [Online]. Available: <https://doi.org/10.1016/j.apenergy.2011.12.003>
- [11] Y. Li, C. Yang, Z. Yan, B. Guo, H. Yuan, J. Zhao, and N. Mei, "Analysis of the icing and melting process in a coil heat exchanger," *Energy Procedia*, vol. 136, pp. 450–455, 2017, 4th International Conference on Energy and Environment Research ICEER 2017. [Online]. Available: <https://doi.org/10.1016/j.egypro.2017.10.302>
- [12] ARCONEL. (2019) Balance nacional de energía eléctrica. [Online]. Available: <http://bit.ly/2PiU2eS>
- [13] CONELEC. (2013) Plan maestro de electrificación 2013-2022. aspectos de sustentabilidad y sostenibilidad social y ambiental. [Online]. Available: <http://bit.ly/33UgUFA>
- [14] S. Sanaye and A. Shirazi, "Four e analysis and multi-objective optimization of an ice thermal energy storage for air-conditioning applications," *International Journal of Refrigeration*, vol. 36, no. 3, pp. 828–841, 2013. [Online]. Available: <https://doi.org/10.1016/j.ijrefrig.2012.10.014>

- [15] X. Song, T. Zhu, L. Liu, and Z. Cao, "Study on optimal ice storage capacity of ice thermal storage system and its influence factors," *Energy Conversion and Management*, vol. 164, pp. 288–300, 2018. [Online]. Available: <https://doi.org/10.1016/j.enconman.2018.03.007>
- [16] H. H. Sait, "Experimental study of water solidification phenomenon for ice-on-coil thermal energy storage application utilizing falling film," *Applied Thermal Engineering*, vol. 146, pp. 135–145, 2019. [Online]. Available: <https://doi.org/10.1016/j.applthermaleng.2018.09.116>
- [17] G. D. Shinde and P. R. Suresh, "A review on influence of geometry and other initial conditions on the performance of a pcm based energy storage system," *International Journal of Thermal Technologies*, vol. 4, no. 3, pp. 214–222, 2014. [Online]. Available: <http://bit.ly/31KAL8A>
- [18] R. A. Jordan, L. A. B. Cortez, V. Silveira Jr., M. E. R. M. Cavalcanti-Mata, and F. D. de Oliveira, "Modeling and testing of an ice bank for milk cooling after milking," *Engenharia Agrícola*, vol. 38, pp. 510–517, 08 2018. [Online]. Available: <https://bit.ly/2E6ZACD>
- [19] K. A. R. Ismail, M. M. Gonçalves, and F. A. M. Lino, "Solidification of pcm around a finned tube: Modeling and experimental validation," *Journal of Basic and Applied Research International*, vol. 12, no. 2, pp. 115–128, 2015. [Online]. Available: <http://bit.ly/3410fjU>
- [20] R. P. Guapulema Maygualema and E. A. Jácome Domínguez, "Diseño y construcción de un generador de hielo tubular para laboratorio," 2013. [Online]. Available: <https://bit.ly/2RETbq9>
- [21] H. H. Sait, A. Hussain, and A. M. Selim, "Experimental investigation on freezing of water falling film on vertical bank of cold horizontal tubes," *Journal of thermal science and engineering applications*, vol. 4, no. 4, p. 041006, 2012. [Online]. Available: <https://doi.org/10.1115/1.4006314>
- [22] F. Wang, C. Liang, M. Yang, C. Fan, and X. Zhang, "Effects of surface characteristic on frosting and defrosting behaviors of fin-tube heat exchangers," *Applied Thermal Engineering*, vol. 75, pp. 1126–1132, 2015. [Online]. Available: <https://doi.org/10.1016/j.applthermaleng.2014.10.090>
- [23] H. G. Ramírez-Hernández, F. A. Sánchez-Cruz, F. J. Solorio-Ordaz, and S. Martínez-Martínez, "An experimental study of heat transfer on a tube bank under frost formation conditions," *International Journal of Refrigeration*, vol. 102, pp. 35–46, 2019. [Online]. Available: <https://doi.org/10.1016/j.ijrefrig.2019.01.031>
- [24] S. K. Wang, *Handbook of Air Conditioning and Refrigeration*, mcgraw-hill ed., 2000. [Online]. Available: <http://bit.ly/344CVlj>
- [25] ASHRAE, *ASHRAE handbook : fundamentals*, 2001. [Online]. Available: <https://bit.ly/2LMExtf>
- [26] Y. Çengel and M. Boles, *Termodinámica*, 2015. [Online]. Available: <https://bit.ly/2E5v8J3>
- [27] J. P. Holman, *Transferencia de calor*, 1998. [Online]. Available: <https://bit.ly/2PdyO1N>
- [28] F. Taboas Touceda, "Estudio del proceso de ebullición forzada de la mezcla amoniaco/agua en intercambiadores de placas para equipos de refrigeración por absorción," Ph.D. dissertation, 2006. [Online]. Available: <https://bit.ly/34b58Xk>



# SPEED CONTROL OF A PERMANENT MAGNET SYNCHRONOUS MOTOR ACTUATED BY A THREE-PHASE MULTI LEVEL INVERTER

## CONTROL DE VELOCIDAD DE UN MOTOR SÍNCRONO DE IMANES PERMANENTES ACCIONADO POR UN INVERSOR TRIFÁSICO MULTINIVEL

Rosalino Mayoral Lagunes<sup>1</sup>, José Antonio Juárez Abad<sup>2</sup>, Beatriz Angélica Aguilar López<sup>1</sup>, Jesús Linares Flores<sup>2</sup>, Jorge Luis Barahona Ávalos<sup>2</sup>

### Abstract

This paper presents the design and FPGA embedded implementation of robust controller design to speed tracking problem for a Permanent Magnet Synchronous Motor (PMSM). Then, a linear controller based on the exact static error dynamics passive output feedback (ESEDPOF) is proposed, where the uncertainty estimation is taken into account. The technique of passivity requires knowing the load torque, this is estimated with a traditional reduced-order observer. PMSM is driven by a five levels Three-Phase Cascaded Cell Multi-Level Inverter (3 $\Phi$ -CCMLI). The medium-scale field-programmable gate array (FPGA) Spartan-6 XC6SLX9 is used for implementing the ESEDPOF controller, the reduced-order observer, and the multilevel pulse width modulator. The parallel processing provided by these devices allowed to obtain a sampling time of 10 $\mu$ s. Simulation and Experimental validation shows an excellent dynamical performance.

**Keywords:** PMSM, Passivity Based Control, Multi Level Inverter, FPGA.

### Resumen

Este trabajo presenta el diseño e implementación de un controlador robusto para el seguimiento de velocidad de un motor síncrono de imanes permanentes (MSIP). Se propone un controlador lineal basado en la retroalimentación dinámica de la salida pasiva estática del error exacto. El controlador pasivo propuesto requiere del conocimiento del par de carga, por lo que el mismo es estimado con un observador tradicional de orden reducido. El MSIP es impulsado por medio de un inversor multinivel trifásico de celdas en cascada de cinco niveles. Para la implementación del controlador, estimador y modulador multinivel se emplea un arreglo de compuertas programable en campo (FPGA) de la familia Spartan-6 XC6SLX9. El procesamiento en paralelo que provee este dispositivo permite obtener un tiempo de muestreo de 10  $\mu$ s. Los resultados de simulación y experimentales muestran que el controlador propuesto tiene un excelente desempeño.

**Palabras clave:** PMSM, control por pasividad, inversor multinivel, FPGA.

<sup>1,\*</sup>Postgraduate Studies Department, Universidad Tecnológica de la Mixteca – México.

<http://orcid.org/0000-0003-3420-7416>, <http://orcid.org/0000-0002-2769-4992>.

<sup>2</sup>Institute of Electronics and Mechanics, Universidad Tecnológica de la Mixteca – México,

Corresponding author ✉: [jbarahona@mixteco.utm.mx](mailto:jbarahona@mixteco.utm.mx). <http://orcid.org/0000-0001-7440-0849>,

<http://orcid.org/0000-0002-5723-4786>, <http://orcid.org/0000-0002-5502-6692>.

Received: 22-11-2019, accepted after review: 13-12-2019

Suggested citation: Mayoral Lagunes, R.; Juárez Abad, J. A.; Aguilar López, B. A.; Linares Flores, J. and Barahona Ávalos, J. L. (2020). «Speed control of a permanent magnet synchronous motor actuated by a three-phase multi level inverter». INGENIUS. N.º 23, (january-june). pp. 97-108. DOI: <https://doi.org/10.17163/ings.n23.2020.09>.

## 1. Introduction

Electrical motors are actuators that are used in most of the industrial machinery and household appliances; their position and/or speed are controlled so that they can fulfill specific tasks, such as: Computer Numerical Control (CNC) machines, belt conveyors, robotics and other industrial processes. Among these applications it is common to find permanent magnet synchronous motors (PMSM), which have better precision in terms of speed and position due to their property of synchrony. Recently, controllers that enable modifying those magnitudes are implemented using reprogrammable and reconfigurable devices, such as digital signal processors (DSP) and field-programmable gate arrays (FPGA) [1, 2].

Different control schemes have been proposed for these type of motors, for instance the passivity-based control shown in [3], and a controller based on the Backstepping technique for position tracking, which is shown in [3].

The PMSM must be driven by a three-phase inverter, which must preferably provide a voltage wave with the lowest possible amount of harmonics, because otherwise some well-known damages may occur in both the electrical and mechanical parts of the motor [4]. A three-phase multilevel inverter of five levels is constructed in this work, with which the harmonics content is reduced thus improving the quality of the voltage supplied to the motor [5]. An additional feature of this type of converter is that it handles more power, since such power is distributed in the series arrays that constitute the commutation devices [6–8].

The rest of the work is structured in the following manner. Section 2 briefly describes the system to be controlled, and then section 3 describes the centered methodology employed for the implementation of processing systems in FPGA, taken from [9–13]; such methodology possesses two phases which are covered in sections 4 and 5. Afterwards, section 6 describes the construction of the power converter employed, in this case the three-phase multilevel inverter of five levels, and at last section 7 presents the experimental results obtained.

## 2. Description of the system

A passivity-based control is implemented for soft tracking the speed of a permanent magnet synchronous motor, driven by a three-phase inverter of cascade topology of 5 levels. This is carried out using a device of reconfigurable logic.

Figure 1 shows the scheme of the proposed system. The PMSM is driving an unknown mechanical load, and the motor power is supplied by the three-phase inverter of 5 levels. The passivity-based controller, the

signal processing and the PS-PWM modulator are implemented in a FPGA.

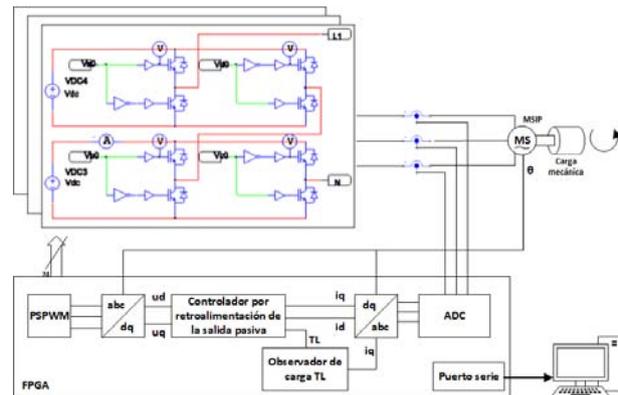


Figure 1. General scheme of the system.

## 3. Methodology

According to the methodology employed, taken from [7], this work is divided in 2 phases. In general, the first phase has to do with the modeling of the system, and the analysis of the mathematical properties of this model to formulate the controller of the system. The second phase deals with the implementation of the control algorithm formulated in phase 1, and the complementary modules which will be implemented in the reconfigurable logic device. Each of these phases is described with more detail in the following.

Phase 1 is graphically summarized in Figure 2. This phase is independent of the device to be used in the implementation, and in turn is divided in three stages:

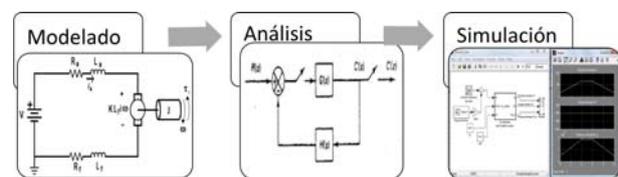


Figure 2. Phase 1 of the methodology.

- 1. Modeling:** In this stage the PMSM and the three-phase multilevel are mathematically modelled, using the laws and theorems necessary to obtain the differential equations of the system.
- 2. Analysis:** The dynamic models of the combination inverter-motor are analyzed, the proposed controller is designed by means of algebraic tools, and the feasibility of the complete system is determined.
- 3. Simulation:** In this stage the obtained mathematical models are simulated, to confirm the

approximation to the real system; the MATLAB-Simulink software together with the software PSIM are utilized for this purpose.

Phase 2 of the proposed design methodology is summarized in Figure 3, and comprises 7 stages. This proposal is a balanced solution between two opposing requirements: 1) a friendly method that perfectly adapts to a control engineer who is not an expert in digital design and 2) obtain a good performance of the control system [9]. The aforementioned stages are described in the following:

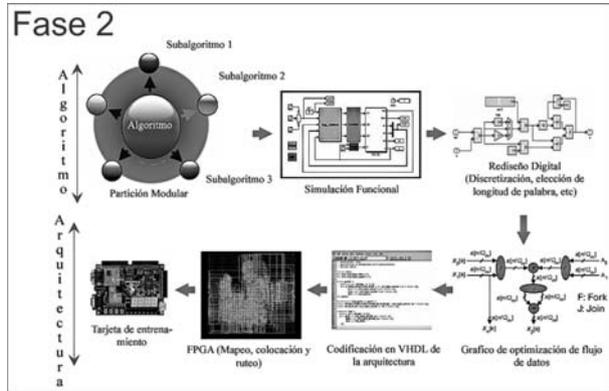


Figure 3. Phase 2 of the methodology. [9]

1. **Modular partition:** Consists of dividing the algorithm in reusable blocks that make sense from the functional point of view, in other words, the control algorithm and the modulator of the inverter are divided in modules, and those that may be reusable such as operations, coders, search tables, among others, are identified. It is sought to divide the design in smaller parts that may be manageable.
2. **Functional simulation:** In this stage the controller is validated as a function of continuous-time blocks using, for instance, the MATLAB-Simulink software.
3. **Digital redesign:** This stage is very important, since it includes the realization of the digital filter, the selection of the sampling time, the selection of the format of the coefficients and variables according to the control requirements and constraints, such as: word length, and fixed or floating point, among others.

The final four stages are for the implementation based on FPGA, and consist of the following:

4. **Optimization graph of the data flow:** In this stage, the data flow of the algorithm is modeled to have the best balance in terms of

time/area; for this purpose, a graphical representation of every module of the algorithm is obtained. The graphical representation is constituted by lines and nodes, each line corresponding to a data transfer and the nodes representing arithmetical operations, or logic or mathematical functions.

5. **HDL coding:** the graphic of data flow of the previous stage is transcribed to VHDL code; regularly, the data transfers are controlled by finite states machines, synchronized at the clock signal of the FPGA.
6. **Implementation in FPGA:** This is an automatic process, regularly carried out with the tool provided by the manufacturer of the device. In this stage a mapping of the resulting functions, placement and routing is performed.
7. **Experimental validation:** Consists of integrating the whole system, carrying out the performance tests, and adjusting all the necessary parameters until obtaining the desired results.

## 4. Phase 1 of the methodology

### 4.1. Control by DFSP0EE, modeling and analysis

Starting with the  $d - q$  model obtained from the equations of the original mathematical model presented in [14], the input voltages  $u = [u_d u_q]$  and input currents  $i = [i_d i_q]$  in the coordinate axes  $d - q$  are calculated. For the design of the controller it is supposed that the angular speed and the load torque are known. To start with the design of the controller by passive output feedback, it is considered the following model in the coordinate axes  $d - q$  [15]:

$$L \frac{di_d}{dt} = -r_e i_d + Li_q n_p \omega + u_d$$

$$L \frac{di_q}{dt} = -r_e i_q + Li_d n_p \omega - K_m \omega + u_q \quad (1)$$

$$J \frac{d\omega}{dt} = \frac{3}{2} K_m i_q - D\omega - T_L$$

Representing the system of equations (1) in matrix form and rewriting it in passive form, yields [16]:

$$A\dot{x} = [J(y) - R]x + Bu - N\eta, x \in R^3, u \in R^2$$

$$y_1 = i_d \quad (2)$$

$$y_2 = i_q$$

where:

$$A = \text{diag} \left( L \quad L \quad \frac{2}{3} J \right); N = \text{diag} (1 \quad 1 \quad 1) \quad (3)$$

$$J(y) = \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -k_m \\ 0 & k_m & 0 \end{pmatrix}}_{=J_0} + y \underbrace{\begin{pmatrix} 0 & n_p L & 0 \\ -n_p L & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=yJ_1} \quad (4)$$

$$R = \begin{pmatrix} r_e & 0 & 0 \\ 0 & r_e & 0 \\ 0 & 0 & \frac{2}{3}B \end{pmatrix}; B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}; \eta = \begin{pmatrix} 0 \\ 0 \\ \frac{2}{3}T_L \end{pmatrix} \quad (5)$$

The state vector is constituted by:

$$x^T(t) = (i_d, i_q, \omega) \quad (6)$$

$J_0$  and  $J_1$  are constant and antisymmetric matrices,  $y$  is a scalar representing the output  $y$  of the system which, in addition, is one of the variables of the state vector  $x$ ;  $R$  is a positive semidefinite symmetric matrix,  $B$  is a constant matrix of the control inputs, and finally  $\eta$  is a vector of load torques. The definitions of passivity and dissipativity discussed in [15] are used to verify that the average nonlinear system of the PMSM given in (1) is passive.

Let  $u^* \in \mathbb{R}^2$  denote the nominal input trajectory corresponding to the trajectory of the nominal state vector  $x^*(t) \in \mathbb{R}^3$ . Starting from this, the reference dynamics of the system is constructed, which must satisfy the following expression:

$$A\dot{x}^* = [J(y^*) - R]x^* + Bu^* - N\eta^* \quad (7)$$

It is defined the tracking error  $e = x - x^*$ , the error of the control input  $e_u = u - u^*$ , and the error of the disturbance input  $e_\eta = \eta - \eta^*$ , and considering that  $J(y) = J_0 + yJ_1$ , then:

$$A\dot{e} = J(y)e - Re + Be_u + J_1x^*l_3e - Ne_\eta$$

Defining  $M_3 = J_1x^*l_3$  as the  $3 \times 3$  matrix given

$$M_3 = \begin{pmatrix} 0 & n_p L & 0 \\ -n_p L & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} i_d^* \\ i_q^* \\ \omega \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \\ = n_p L \begin{pmatrix} 0 & 0 & i_d^* \\ 0 & 0 & -i_q^* \\ 0 & 0 & 0 \end{pmatrix}$$

Afterwards, matrix  $M_3$  is decomposed in a symmetric and an antisymmetric matrix using matrix algebra, having as result

$$M_3 = \frac{1}{2}n_p L \underbrace{\begin{pmatrix} 0 & 0 & i_d^* \\ 0 & 0 & i_q^* \\ -i_d^* & i_q^* & 0 \end{pmatrix}}_{W_3} + \frac{1}{2}n_p L \underbrace{\begin{pmatrix} 0 & 0 & i_d^* \\ 0 & 0 & i_q^* \\ -i_d^* & i_q^* & 0 \end{pmatrix}}_{S_3}$$

Subsequently  $W_3$  is added to the conservative matrix  $J(y)$ , and  $S_3$  is added to the dissipative part  $R$ . With this operation it is found the dynamic equation of the tracking error, expressed as:

$$A\dot{e} = J^*e - R^*e + Be_u \quad (8)$$

If it is considered that  $R^* > 0$ , in other words, positive definite, then the following condition is satisfied:

$$R^* = \begin{pmatrix} r_e & 0 & -\frac{1}{2}n_p Li_q^* \\ 0 & r_e & \frac{1}{2}n_p Li_d^* \\ -\frac{1}{2}n_p Li_q^* & \frac{1}{2}n_p Li_d^* & \frac{2}{3}B \end{pmatrix} \quad (9)$$

Using the Sylvester criterion it is verified that  $R^* > 0$ , namely, that it is a positive definite matrix, it must hold that

$$\det(R^*) = \frac{2}{3}r_e B - \frac{1}{4}n_p^2 L^2 [(i_d^*)^2 + (i_q^*)^2] > 0 \quad (10)$$

Following with the design methodology of the controller, using Lyapunov, it is proposed that  $e_u$  como  $e_u = -\delta B^T e$ , where:

$$\delta = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} > 2 \quad (11) \\ \delta_1, \delta_2 > 0$$

with  $\delta$  a positive matrix, such that

$$\begin{pmatrix} u_d \\ u_q \end{pmatrix} = \begin{pmatrix} u_d^* \\ u_q^* \end{pmatrix} - \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x^* \quad (12)$$

From equation (12) the control laws that will act on the system are obtained, yielding:

$$u_d = \widetilde{u}_d - \delta_1(i_d - \widetilde{i}_d) \\ u_q = \widetilde{u}_q - \delta_2(i_q - \widetilde{i}_q) \quad (13)$$

#### 4.1.1. Equilibrium points

The equilibrium points will be used to calculate the desired reference variables  $i_d^*$ ,  $i_q^*$ ,  $u_d^*$ ,  $u_q^*$ ,  $w^*$ . For this case,  $w^*$  is the desired value of the speed which will be measured in the motor shaft. To obtain these references, the derivatives of equations (1) are set equal to zero, thus obtaining the following expressions:

$$0 = L\widetilde{i}_q n_p \omega_d + u_d \quad (14)$$

$$0 = -r_e \widetilde{i}_q - K_m \omega_d + u_q \quad (15)$$

$$0 = \frac{3}{2}K_m \widetilde{i}_q - D\omega_d - T_L \quad (16)$$

from which it can be solved:

$$\begin{aligned} \tilde{i}_d &= 0 \\ \tilde{i}_q &= \frac{2}{3} \frac{D}{K_m} \omega_d + \frac{2}{3} \frac{T_L}{K_m} \\ \tilde{u}_d &= -L \tilde{i}_q n_p \omega_d \\ \tilde{u}_q &= r_e \tilde{i}_q + K_m \omega_d \end{aligned} \quad (17)$$

#### 4.1.2. Design of the observer of the load torque parameter

To design and develop an efficient and robust controller, it is necessary to know all the variables that disturb the system. Since the controller by feedback of the error passive output is not robust in the presence of disturbances of constant load torque, a reduced order observer is designed to estimate the load torque parameter together with the unmolded friction terms [17]. The quadrature current  $i_q$  is defined as the control input,  $u_1 = i_q$ , and the estimated angular speed is defined as the measured output,  $y_1 = \omega$ . The nominal values of the parameters  $J$  and  $K_m$  are considered known. Substituting variables  $u_1$  and  $y_1$  in the mechanical equation given in (1), yields the following expression:

$$J \frac{dy_1}{dt} = \frac{3}{2} k_m u_1 - B y_1 - T_L \quad (18)$$

The following observer of load torque is proposed as:

$$\frac{d\tilde{T}_L}{dt} = \lambda(T_L - \tilde{T}_L) \quad (19)$$

where  $\lambda > 0$ . The dynamics of the estimation error is given by:

$$e_{T_L} = T_L - \tilde{T}_L \quad (20)$$

$$\frac{de_{T_L}}{dt} = \frac{dT_L}{dt} - \frac{d\tilde{T}_L}{dt} \quad (21)$$

If  $T_L$  is considered as constant,  $\frac{dT_L}{dt} = 0$ , and substituting (19) in (21) results in the following dynamics of the estimation error:

$$\frac{de_{T_L}}{dt} = -\lambda e_{T_L} \quad (22)$$

Choosing a gain  $\lambda > 0$ , the observing error  $e_{T_L}$  converges exponentially to zero as  $t \rightarrow \infty$ . Now, substituting equation (18) in (19) gives:

$$\frac{d\tilde{T}_L}{dt} = \lambda \left[ \frac{3}{2} K_m i_q - \left( B y_1 + J \frac{dy_1}{dt} \right) \tilde{T}_L \right] \quad (23)$$

Solving and rearranging the previous equation yields the following expression:

$$\frac{d\tilde{T}_L}{dt} + \lambda J \frac{dy_1}{dt} = \frac{3}{2} K_m u_1 - \lambda B y_1 - \lambda \tilde{T}_L \quad (24)$$

Making the change of variable  $\Psi = \tilde{T}_L + \lambda J y_1$ , results in the following reduced order observer:

$$\frac{d\Psi}{dt} = -\lambda \Psi + (J\lambda - B)\lambda y_1 + \frac{3}{2} \lambda K_m i_q \quad (25)$$

$$\tilde{T}_L = \Psi - \lambda J y_1 \quad (26)$$

The estimated value of the load torque parameter together with the unknown friction terms  $\tilde{T}_L$ , adapt online to the control law for soft tracking the angular speed. The purpose of this adaptation is to reduce the effects produced by the load applied to the shaft of the PMSM.

#### 4.2. Simulations of the controller in MATLAB

With the use of the MATLAB-Simulink software, it is proceeded to design the whole system in block diagram. Figure 4 shows the main block diagram.

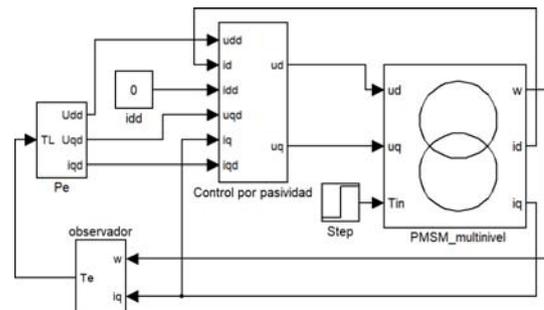


Figure 4. Main block diagram.

The passivity-based controller shown in Figure 4 is based on the set of equations (12); Figure 5 shows the implementation in blocks.

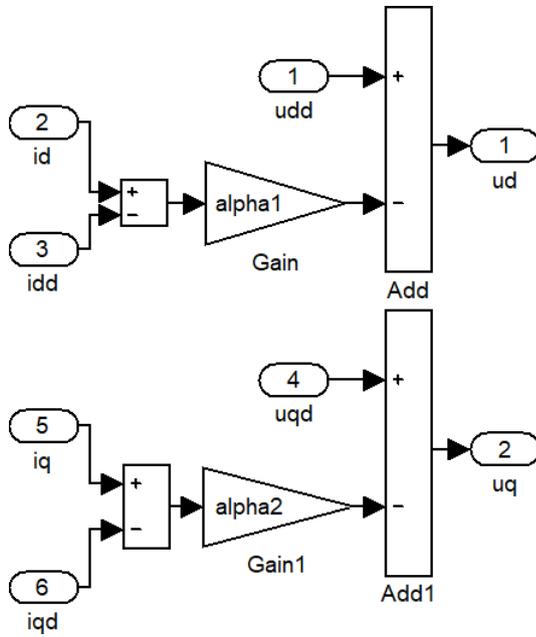


Figure 5. Passivity-based controller.

The load observer of equations (25) and (26) is shown in blocks in Figure 6.

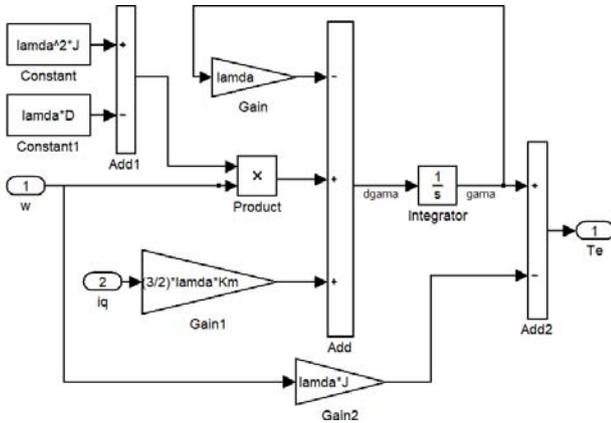


Figure 6. Reduced order load observer.

Equations (17) of the equilibrium points of the system, which are utilized to find the references and desired trajectories, are shown in blocks in Figure 7.

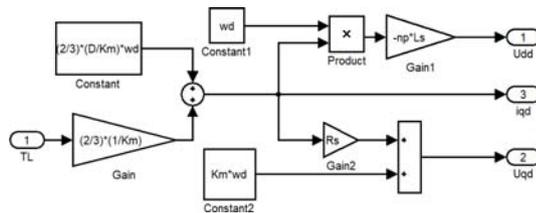


Figure 7. Reference signals.

A cosimulation with Simulink and Psim is carried out to validate the operation of the system. The electric diagram in Psim of the multilevel inverter in conjunction with the PMSM, is shown in Figure 8. It can be

observed that the multilevel inverter is constituted by three sections, equal to the one shown in the square at the left of Figure 8.

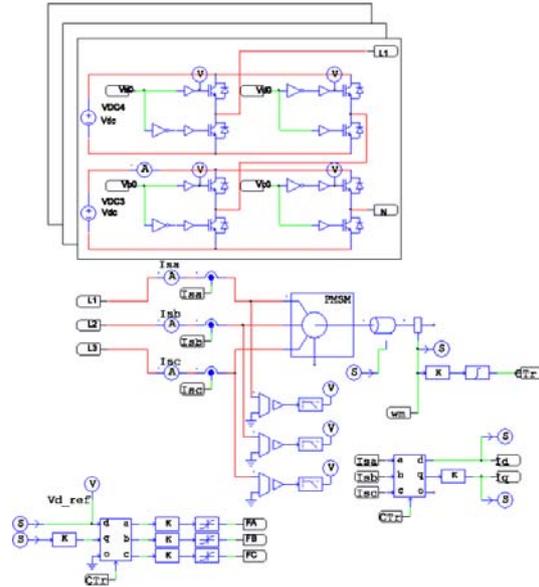


Figure 8. Electric diagram in Psim.

### 4.3. Multilevel inverter

For this work, a multilevel inverter with cascaded cells is utilized. This configuration is based on cells connected in series to add up the voltages, and obtain the 5 desired levels as shown in Figure 9.

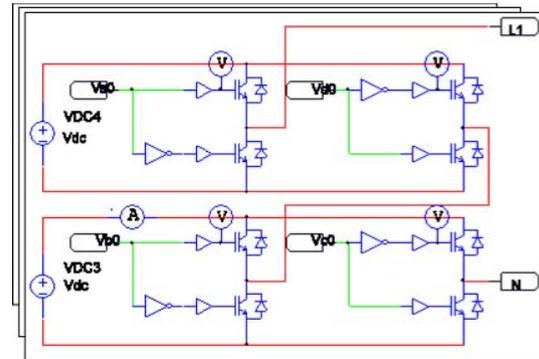


Figure 9. Topology of the selected cascade inverter of five levels.

The modulation technique employed is PS-PWM, which for the case of the inverter of 5 levels, consists of four triangular signals called carriers, 90° out of phase between them, i.e., 0°, 90°, 180° and 270°, respectively; in this case the carriers are generated at a frequency of 12 kHz. These signals are compared with a sinusoidal signal called modulating signal; for the case of interest, such modulating signal is generated by the transformation dq-abc. In general, its operation consists in obtaining logic ones and zeros by means of the comparison between the carrier and modulating signals: a

logic one is obtained if the carrier signal is greater than or equal to the modulating signal, otherwise a logic zero arises. In this way, PWM signals that activate the commutation devices are generated. It should be pointed out that each phase contains a modulating signal  $120^\circ$  out of phase with respect to the other phases; therefore, there is a total of 24 PWM signals, 12 main channels with their corresponding 12 complementary signals. Figure 10 shows the diagram to generate the PS-PWM modulation for a single phase, with its four main signals.

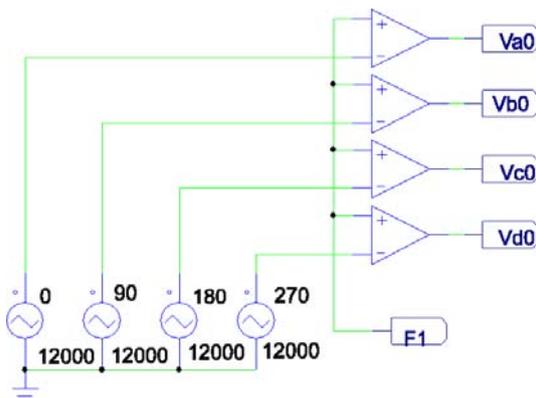


Figure 10. PS-PWM modulator for one phase.

#### 4.4. Simulation results

Hereunder, the results obtained from the cosimulation of the system in different scenarios are shown, with the purpose of performing a comparison. These scenarios are: the Psim model of the PMSM driven by a conventional three-phase inverter, and finally driven by the three-phase multilevel inverter of five levels. The designed passivity-based controller is applied in both cases. As a result of the simulation, the plot in Figure 11 shows the angular speed of the motor driven by both types of inverters. It can be observed in such plot that when a change of the load torque is applied on the motor shaft, at time  $t = 1$  s, the disturbance caused by such change is effectively counteracted by the controller. There is no significant difference between one type of inverter and the other.

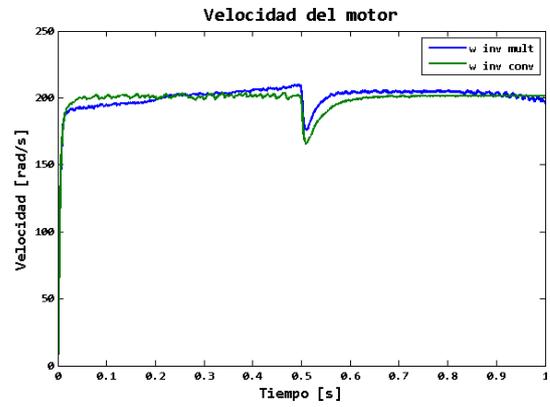


Figure 11. Motor speed.

Figures 12 and 13 show plots of the currents  $i_d$  and  $i_q$ , respectively, for the aforementioned conditions. It is important to point out that with the use of the multilevel inverter the ripple in the currents decreases considerably, due to the combined use of such inverter and the designed controller.

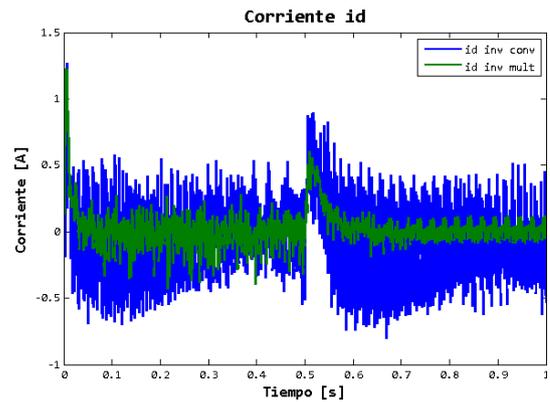


Figure 12.  $i_d$  current.

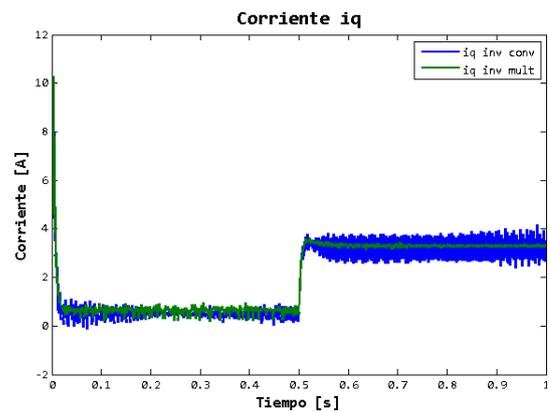


Figure 13.  $i_q$  current.

At last, Figure 14 shows the estimated load torque. It can be seen that the designed observer equally works better in the case of the multilevel inverter, which verifies its correct operation.

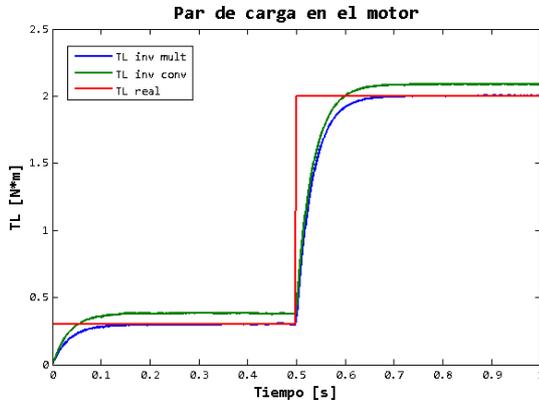


Figure 14. Estimated load torque of the motor.

## 5. Phase 2

The programming of the FPGA is carried out in this phase. It is worth mentioning that in this work, the coding is directly performed in the VHDL language, without employing any code generator; this approach enables making optimal use of the device resources.

### 5.1. Implementation of the PS-PWM modulator module

The modulator was programmed in VHDL language in the ISE DESIGN software of XILINX. For simplicity, Figure 15 shows the design by means of a schematic block diagram. The main programmed blocks are: block of carrier signals, in which the triangular signals are generated; block of modulating signal, which generates the modulating signal; at last, the comparators generate the PWM signals to perform the comparison between the carrier signals and the modulating signal.

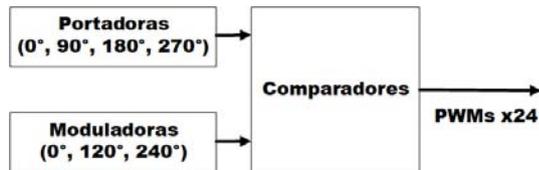


Figure 15. Block diagram of the modulator.

### 5.2. Implementation of the controller module

The controller is divided in various sub-modules: abc-dq transformation, dq-abc transformation, load observer, equilibrium points, passivity-based speed control, reading of current sensors, reading of position/speed sensor. These modules work with the IEEE 754 32-bit simple precision standard, which is the numerical format mostly used in hardware [18].

#### 5.2.1. dq – abc transformation module

The architecture shown in Figure 16 is implemented for the development of this module, which enables on-line solving equation (27) and optimizing resources. A multiplier, an adder and a register are utilized in the case of this module; although all these elements are inside the FPGA, their use should be optimized because they are limited. As observed in Figure 16, it requires 8 inputs: d, q and the 6 trigonometric functions. The outputs are the values a, b and c, which correspond to the value of voltage of the phases. Afterwards, Figure 17 shows the hardware implementation of equations (27); this methodology was taken from [19]. Since the complexity of the system makes difficult to exemplify the development, only this module is shown; however, the construction of the remaining modules is carried out in a similar manner. Table 1 shows the consumption of logic resources in the FPGA after performing the implementation of all modules involved in the design, and Figure 18 presents the highest hierarchy diagram of the system.

$$\begin{aligned}
 V_a &= V_d \cos(\theta) + V_q \sin(\theta) \\
 V_b &= V_d \cos(\theta - \frac{2}{3}\pi) + V_q \sin(\theta - \frac{2}{3}\pi) \\
 V_c &= V_d \cos(\theta + \frac{2}{3}\pi) + V_q \sin(\theta + \frac{2}{3}\pi)
 \end{aligned}
 \tag{27}$$

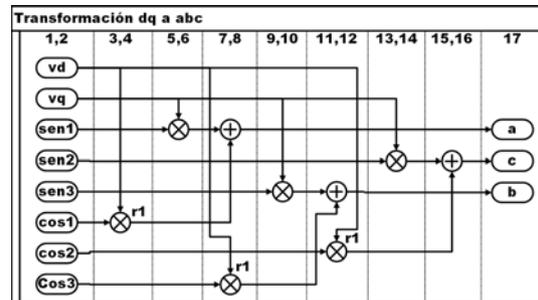


Figure 16. Sequence diagram of the transformation dq to abc.

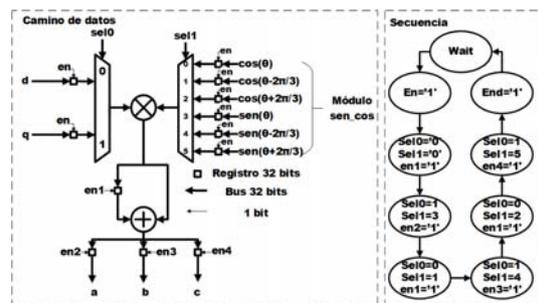


Figure 17. Flow diagram of data and machine of states.

Table 1. Level of utilization of the FPGA

Device Utilization Summary			
Slice Logic Utilization	Used	Available	Utilization
Number of Slice Registers	4,038	11,440	35%
Number of Slice LUTs	4,834	5,720	84%
Number used as logic	4,599	5,720	80%
Number used as Memory	179	1,440	12%
Number of occupied Slices	1,408	1,430	98%
Number of MUXCYs used	1,716	2,860	60%
Number with an unused Flip Flop	1,808	5,107	35%
Number with an unused LUT	273	5,107	5%
Number of fully used LUT-FF pairs	3,026	5,107	59%
Number of unique control sets	84		
Number of slice register sites lost to control set restrictions	245	11,440	2%
Number of bonded IOBs	56	186	30%
Number of LOCed IOBs	55	56	98%
Number of RAMB16BWERS	17	32	53%
Number of RAMB8BWERS	10	64	15%
Number of BUFIO2/BUFIO2_2CLKs	1	32	3%
Number of BUFIO2FB/BUFIO2FB_2CLKs	1	32	3%
Number of BUFG/BUFGMUXs	5	16	31%
Number of DCM/DCM_CLKGENs	1	4	25%
Number of DSP48A1s	16	16	100%
Average Fanout of Non-Clock Nets	3.65		

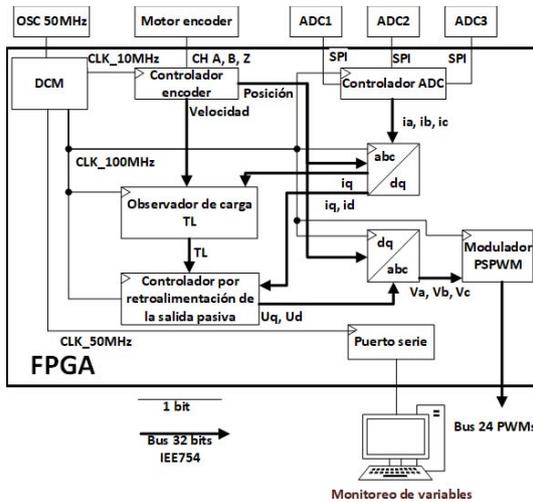


Figure 18. General schematic diagram of the system in the FPGA.

## 6. Experimental platform

The electrical configuration of the platform of the multilevel inverter is shown through the one-line diagram in Figure 19.

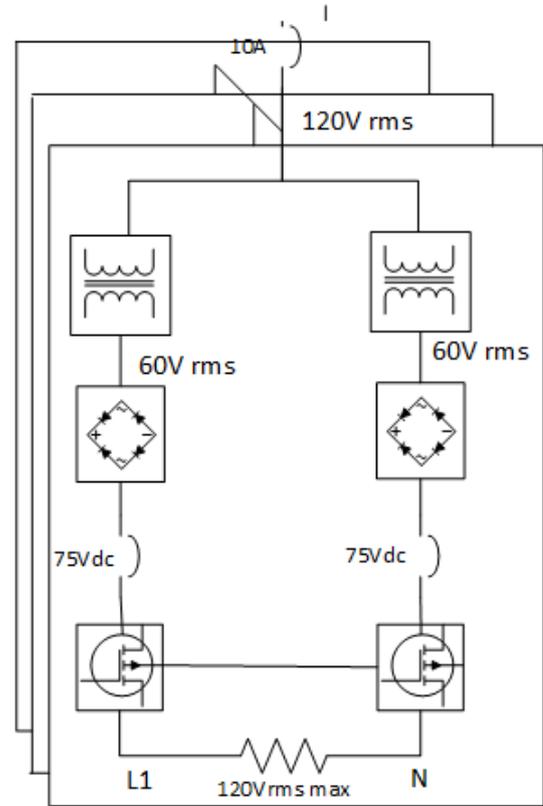


Figure 19. One-line diagram of the multilevel inverter.

Figure 20 shows the experimental platform constructed and used for the laboratory tests. The control signals are generated by the FPGA, which are used to switch the multilevel inverter and generate the power necessary to drive the PMSM according to the desired reference trajectory. The experimental platform is constituted by the following elements:

- Isolated DC sources
- H-bridges
- Main charge center
- Transformer of the DC source
- Diode bridge
- Capacitor of the DC source
- DIN rail
- Connection screw terminals
- Protection breakers

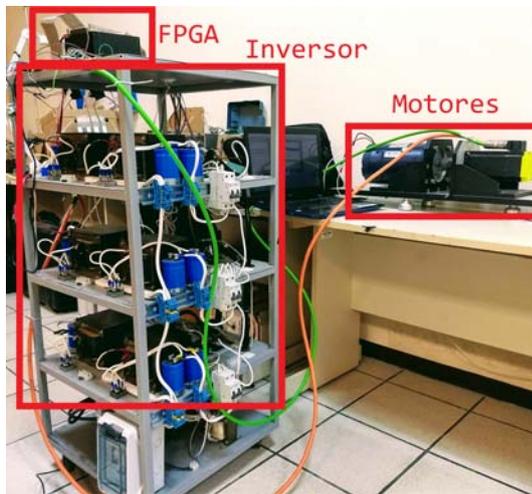


Figure 20. Experimental platform.

### 6.1. ALINX development kit

The development kit utilized is ALINX brand, AX309 model, shown in Figure 21. It comprises a FPGA SPARTAN-6 XC6SLX9, which is a low-cost development platform, with the following characteristics:



Figure 21. Development kit.

- FPGA model: XC6SLX9
- Size: 90x130 mm
- Crystal oscillator: 50 MHz
- Number of expansion I/O: 68
- SDRAM: 256 Mbit
- FLASH: 16 Mbit
- IO port level: 3.3 V
- Serial port: onboard U transfer serial port

### 6.2. Permanent magnet synchronous motor

The motor utilized is BALDOR Brand, whose main characteristics are 1.23 kW at 4000 rpm; Figure 22 shows the motor, while Table 2 shows its parameters.

Table 2. Motor parameters

Parameters	Value
CAT.NO.	BSM80N-275AF
SPEC.	S2P141W042G1
TORQ CONT STALL	3.2 NM
CUR CONT STALL/A RMS	4
POWER	1.23KW
RATED SPEED/RPM	4000
RATED BUS VOLTS	300
PEAK CUR/A RMS	14.4
MAX SPEED/RPM	7000
CLASS	F
AMB.	25



Figure 22. PMSM employed (left) and DC motor used as load (right).

## 7. Experimental results

Since the processing inside the FPGA device is numerical, it is necessary to have a mean to read the different signals that are processed. For this purpose, readings of the signals are taken and sent through a serial port/UART at a speed of 921600 bps; once received in the computer, they are processed and presented by means of a GUI designed in LabView. The measured parameters include the speed ( $\omega$ ), the currents  $i_q$ ,  $i_d$  and the estimated load torque (LT); each plot has a total of 70000 samples in 5 seconds.

On the other hand, the shaft of the PMSM is coupled to a direct current (DC) motor that operates as an external mechanical load, as can be seen in Figure 22.

Then, Figure 23 shows the angular speed  $\omega$  measured by the motor encoder. The desired speed  $\omega_d$  is determined by tracking a curve obtained from a 6th-order Bezier polynomial, which is used as reference. The duration of the starting curve is 5 s, after which it can be observed that the motor speed settles at 200

rad/s, as desired. Nevertheless, note that there is a ripple in the real speed, which is mainly due to the numerical differentiation algorithm used to obtain the speed from the readings of the encoder.

During the same period of time, Figure 24 shows the tracking of the  $i_q$  current at the start of the motor, which settles at 0.5 amperes; this is the active current, which is transformed by the motor into mechanical torque. Similarly, Figure 25 shows the regulation of the  $i_d$  current, which should remain equal to 0 amperes because it is the reactive current that cannot be utilized. Both currents exhibit ripple, due to the resolution of the employed sensor currents. At last, Figure 26 shows the estimation of the load torque at the motor shaft during the same period of time, which settles at 0.4 Nm when the desired speed is reached.

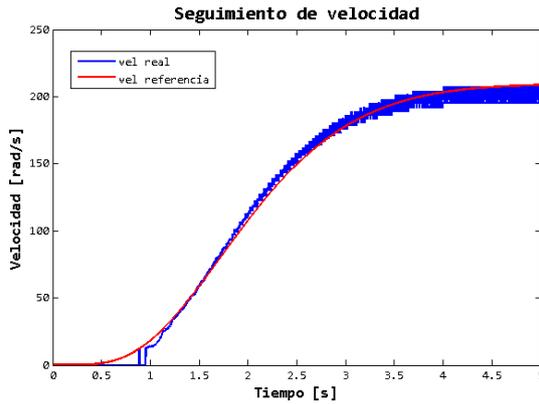


Figure 23. Soft tracking of velocity at start.

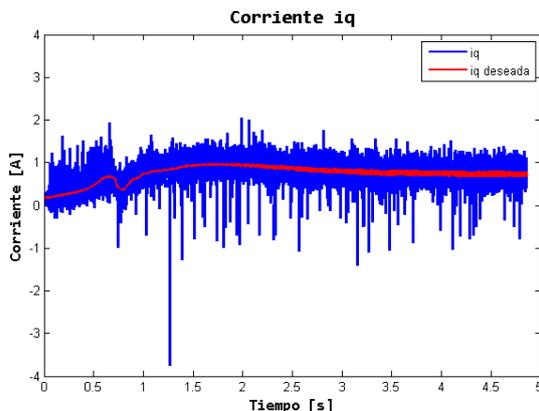


Figure 24. Tracking of the  $i_q$  current at start.

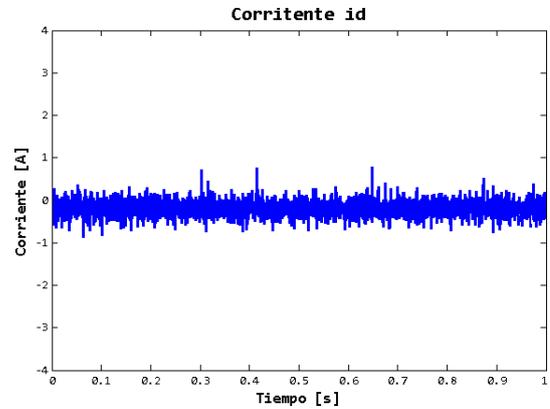


Figure 25. Tracking of the  $i_d$  current at start.

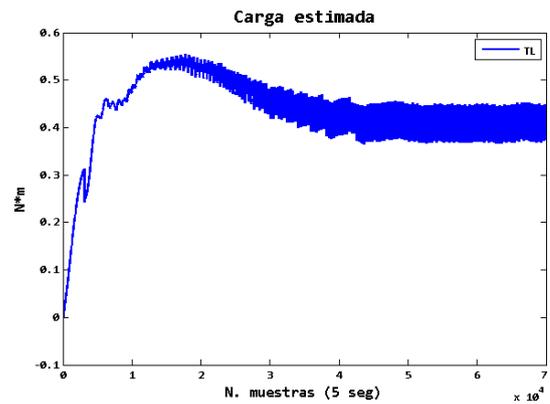


Figure 26. Estimated load at the motor shaft.

## 8. Conclusions

Based on the simulation and experimental results, it can be argued that the designed and implemented passive output feedback controller exhibits an adequate performance in the tracking and regulation of the angular speed of the PMSM. On the other hand, it is verified that the reduced-order load torque estimator, correctly estimates this parameter. In addition, in terms of hardware, it is observed that the multilevel inverter considerably reduces the harmonic distortion of the voltage, which in general results in less noise in the system and avoids the need of using filters. Due to the particularities of the digital device (FPGA) where the implementation was carried out, a sampling time of  $10 \mu\text{s}$  is achieved, which results in the benefit of having larger speed of response in the presence of sudden changes in the load.

As a future work, it is intended to implement a more complex algorithm, such as an algebraic estimator, to accomplish a better estimation of the mechanical load torque. The above would be possible taking advantage of the parallelism, which is the main characteristic of the FPGA.

## References

- [1] T. Wildi, *Máquinas eléctricas y sistemas de potencia*, 2007. [Online]. Available: <https://bit.ly/35aPtZm>
- [2] L. Blanco Rubio, “Diseño electromagnético de un motor síncrono de imanes permanentes para el accionamiento directo de la hélice de un barco,” 2017. [Online]. Available: <https://bit.ly/2PAJZAJ>
- [3] J. Linares-Flores, C. Garcia-Rodriguez, H. Sira-Ramirez, and O. D. Ramirez-Cárdenas, “Robust backstepping tracking controller for low speed pmsm positioning system: Design, analysis, and implementation,” in *2015 IEEE International Conference on Industrial Technology (ICIT)*, March 2015, pp. 2131–2138. [Online]. Available: <https://doi.org/10.1109/ICIT.2015.7125411>
- [4] IEEE, “Ieee recommended practices and requirements for harmonic control in electrical power systems,” *IEEE Std 519-1992*, pp. 1–112, April 1993. [Online]. Available: <https://doi.org/10.1109/IEEESTD.1993.114370>
- [5] M. H. Rashid, *Electrónica de potencia: circuitos, dispositivos y aplicaciones*, 2004. [Online]. Available: <https://bit.ly/2t9mZ42>
- [6] J. A. Juárez-Abad, J. Linares-Flores, E. Guzmán-Ramírez, and H. Sira-Ramírez, “Generalized proportional integral tracking controller for a single-phase multilevel cascade inverter: An fpga implementation,” *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 256–266, Feb 2014. [Online]. Available: <https://doi.org/10.1109/TII.2013.2242085>
- [7] F. Chauca Llusca, F. Llerena Rengel, and P. Chico Hidalgo, “Diseño y construcción de un inversor multinivel,” *Revista Politécnica*, vol. 33, no. 1, 2014. [Online]. Available: <https://bit.ly/36rAxpI>
- [8] L. G. Franquelo, J. Rodriguez, J. I. Leon, S. Kouro, R. Portillo, and M. A. M. Prats, “The age of multilevel converters arrives,” *IEEE Industrial Electronics Magazine*, vol. 2, no. 2, pp. 28–39, June 2008. [Online]. Available: <https://doi.org/10.1109/MIE.2008.923519>
- [9] E. Monmasson, L. Idkhajine, and M. W. Naouar, “Fpga-based controllers,” *IEEE Industrial Electronics Magazine*, vol. 5, no. 1, pp. 14–26, March 2011. [Online]. Available: <https://doi.org/10.1109/MIE.2011.940250>
- [10] D. G. Maxinez and J. Alcalá Jara, *VHDL: el arte de programar sistemas digitales*, 2002. [Online]. Available: <https://bit.ly/2PCE8dL>
- [11] J. J. Rodríguez-Andina, M. D. Valdés-Peña, and M. J. Moure, “Advanced features and industrial applications of FPGAs—a review,” *IEEE Transactions on Industrial Informatics*, vol. 11, no. 4, pp. 853–864, Aug 2015. [Online]. Available: <https://doi.org/10.1109/TII.2015.2431223>
- [12] W. Zhu, “Fpga logic devices for precision control: An application to large friction actuators with payloads,” *IEEE Control Systems Magazine*, vol. 34, no. 3, pp. 54–75, June 2014. [Online]. Available: <https://doi.org/10.1109/MCS.2014.2308691>
- [13] E. Mandado, L. J. Álvarez, and M. D. Valdés, *Dispositivos Lógicos Programables*, 2002. [Online]. Available: <https://bit.ly/38uCiEq>
- [14] R. Krishnan, *Permanent Magnet Synchronous and Brushless DC Motor Drives*, 2017.
- [15] J. Linares-Flores, C. García-Rodríguez, O. D. Ramírez-Cárdenas, C. Escobar-Noriega, and M. A. Contreras-Ordaz, “Control robusto de seguimiento suave de posición angular del motor síncrono de imanes permanentes,” in *Memorias del XVI Congreso Latinoamericano de Control Automático, Octubre 14-17, 2014. Cancún, Quintana Roo, México*, 2014, pp. 1113–1118. [Online]. Available: <http://doi.org/10.13140/2.1.2760.9607>
- [16] H. J. Marquez, *Nonlinear control systems: analysis and design*. Wiley-Interscience Hoboken, 2003, vol. 1. [Online]. Available: <https://bit.ly/2YMdMKM>
- [17] Y. E. Gliklikh, “Necessary and sufficient conditions for global-in-time existence of solutions of ordinary, stochastic, and parabolic differential equations,” *Abstract and Applied Analysis*, vol. 2006, Special Issue, p. 17, 2006. [Online]. Available: <https://doi.org/10.1155/AAA/2006/39786>
- [18] IEEE, *IEEE 754-2019 - IEEE Standard for Floating-Point Arithmetic*, 2019. [Online]. Available: <https://bit.ly/2E5tnvo>
- [19] M. Naouar, E. Monmasson, A. A. Naasani, I. Slama-Belkhodja, and N. Patin, “FPGA-based current controllers for AC machine drives—a review,” *IEEE Transactions on Industrial Electronics*, vol. 54, no. 4, pp. 1907–1925, Aug 2007. [Online]. Available: <https://doi.org/10.1109/TIE.2007.898302>



# ROBUST FILTERING OF WEAK SIGNALS FROM REAL PHENOMENA

# FILTRAJE ROBUSTO DE SEÑALES DÉBILES DE FENÓMENOS REALES

Valeri Kontorovich<sup>1,\*</sup>, Fernando Ramos-Alarcón<sup>1</sup>

## Abstract

In a large number of real-life scenarios it is required to process desired signals that are significantly immersed into background noise: tectonic signals from the entrails of the earth, signals coming from the far away cosmos, biometric telemetry signals, distant acoustic signals, noninvasive neural interfaces and so on. The purpose of this paper is to present the description of a robust and efficient platform for the real time filtering of signals deeply immersed in noise (rather weak signals) with rather different nature. The proposed strategy is based on two principles: the chaotic modelling of the signals describing the physical phenomena and the application of filtering strategies based on the theory of non-linear dynamical systems. Considering as a study case seismic signals, fetal electrocardiogram signals, voice-like signals and radio frequency interference signals, this experimental work shows that the proposed methodology is efficient (with mean squared error values less than 1%) and robust (the filtering structure remains the same although the phenomenological signals are drastically different). It turns out that the presented methodology is very attractive for the real time detection of weak signals in practical applications because it offers a high filtering precision with a minimum computational complexity and short processing times.

**Keywords:** Chaos, Nonlinear Filtering, Dynamic Systems, Kalman Filter, Weak signals, Real Signals.

## Resumen

En un gran número de escenarios de la vida real se requiere procesar señales de interés que se encuentran muy inmersas en medio de ruido de fondo: señales tectónicas de las entrañas de la Tierra, otras provenientes del lejano cosmos, de telemetría biomédica, acústicas lejanas, interfaces neuronales no invasivas, etc. El propósito de este trabajo es presentar la descripción de una plataforma robusta y eficiente para hacer filtraje en tiempo real de señales muy inmersas en ruido (bastante débiles) y de naturaleza muy diferente. La estrategia propuesta se basa en dos principios: el modelado de las señales de los fenómenos físicos mediante procesos caóticos y la aplicación de estrategias de filtraje basadas en la teoría de sistemas dinámicos no lineales. Tomando como caso de estudio señales sísmicas, señales de electrocardiogramas fetales, señales de tipo voz y señales de interferencias de radiofrecuencia, este trabajo experimental muestra que la metodología es eficiente (error cuadrático medio menor al 1 %) y robusta (la estructura de filtraje, basada en filtro de Kalman, es invariante ante diferentes señales fenomenológicas). La metodología presentada resulta ser muy atractiva para aplicaciones prácticas para la detección de señales débiles en tiempo real por su alta precisión de filtraje con una mínima complejidad computacional y tiempos de procesamiento muy cortos.

**Palabras clave:** caos, filtraje no lineal, sistemas dinámicos, filtro de Kalman, señales débiles, señales reales.

<sup>1,\*</sup> Department of Electric Engineering, Area of Communication, Research Center and Advanced Studies of the Instituto Politécnico Nacional (CINVESTAV-IPN), México. Corresponding author ✉: [valeri@cinvestav.mx](mailto:valeri@cinvestav.mx)

<http://orcid.org/0000-0002-1307-3001>, <http://orcid.org/0000-0002-8235-4886>

Received: 11-09-2019, accepted after review: 11-12-2019

Suggested citation: Kontorovich, V. and Ramos-Alarcón, F. (2020). «Robust filtering of weak signals from real phenomena ». INGENIUS. N.º 23, (january-june). pp. 109-119. DOI: <https://doi.org/10.17163/ings.n23.2020.10>.

## 1. Introduction

Signal processing is basic for many areas of science and engineering. One of the fundamental stages for any signal processing algorithm is filtering, in other words, eliminating (filtering) background noise that accompanies the signal under study, prior to the corresponding signal processing.

There are various strategies for signal filtering, to counteract the effects of noise. The conventional filtering methodologies (based on statistical processing) have proven to be very rather effective when the signal-to-noise ratio (SNR) is greater than or equal to 1, i.e.  $\text{SNR} \geq 1$ , or equivalently  $\text{SNR} \geq 0$  dB if it is expressed in decibels (namely, the magnitude of the noise less than or equal to the magnitude of the signal of interest).

In many practical applications, signals with noise are processed when the magnitude of the latter is greater than the magnitude of the signal of interest, i.e.,  $\text{SNR} < 1$ . In order to filter signals in these situations, either conventional strategies have been adapted or novel methodologies have been created (based on iterative procedures, wavelets, etc.). The cost of successfully filtering signals under these conditions is that real time processing is slightly affected.

The filtering process is a great challenge when it is required to detect very weak signals (magnitude of the noise much greater than the magnitude of the desired signal, in other words,  $\text{SNR} \ll 1$  or equivalently  $\text{SNR} \ll 0$  dB), such as signals from distant solar systems, fetal cardiac activity, small seisms precursors of earthquakes, voice signals immerse in background noise, information signals of radiofrequency, noninvasive neural interfaces, among others. The issue of detecting weak signals is not really new, and it is possible to find a great number of publications (see [1–6] for citing only some references) that address this issue using various methodologies and for different phenomena. Techniques ranging from different filtering schemes (adaptive, time-frequency, FIR, IIR, among others), fuzzy logic, chaotic systems, stochastic resonance, to different decomposition strategies (empirical mode, wavelets, orthogonal signals, among others), are employed. The novelty for the case of the present work is the application of chaotic signals as models of real phenomena, based on the theory of deterministic nonlinear dynamic systems.

With a history of more than 50 years, the theory of dynamic systems [7–9] is one of the pillars of many scientific areas such as physics, automatic control, communications, etcetera; in particular, it is very relevant for filtering the famous Kalman filter (proposed by Rudolph E. Kalman in 1960), which enables a very precise filtering considering that the desired signal is a linear dynamic system under the influence of additive white Gaussian noise (AGWN), and from its invention

until today there is a multitude of practical applications, both recent and old (for instance, [10, 11]).

How to achieve, based on the theory of nonlinear dynamic systems, an effective filtering strategy for weak signals from entirely different physical phenomena? One of the options that is explored in this work is modeling the phenomenological signals as signals generated by chaotic attractors, i.e., by nonlinear but deterministic dynamic systems. In this case the application of the concept of dynamic systems enables two very useful things: modeling real signals as formally deterministic processes, and applying all results about filtering based on the theory of dynamic systems.

Using chaotic models for real signals results original and rather effective, as shown later, even though the modeling of real phenomena by means of chaotic signals have been used during more than 50 years in areas such as seismology [1, 2, 12], statistical communication theory [13, 14], biomedical telemetry [15, 16], processing of submarine signals [3], and also in many areas related to applied physics [17].

This work presents an effective filtering (whose theoretical aspects were developed in [18, 19]) in the sense of a high precision in terms of very small values of the normalized mean square error ( $\text{NMSE} < 1\%$ ). The normalization of the mean square error (MSE) is considered related to the variance of the phenomenological signal. On the other hand, filtering is also robust in the sense that, for input signals coming from different physical phenomena, both the structure and the precision (values of the NMSE) of the filter are practically invariant.

Chaotic modeling is very useful because nearly all algorithms of quasi-optimal filtering are characterized by having a high precision (very small value of the NMSE), and a very low computational complexity. Theoretical details and demonstration of these properties were developed in [18, 19], where the interested reader may review them. This work applies and extends to the practical field the nonlinear filtering ideas (published in [19]), presenting only the experimental details of physical scenarios with different phenomenology, specifically seismic signals, cardiac signals (ECG), vocal tract signals and radiofrequency interfering signals.

## 2. Materials and methods

The development of this work is centered in two basic elements. On one side, previously developed methods that provide a rigorous theoretical foundation and, on the other, MatLab code programming to create a test bench. The MatLab code is developed from the filtering equations described later. In the theoretical aspect, chaotic signals are employed both for filtering nonlinear dynamic systems and for modeling phenomenological

signals. Since the idea is to obtain algorithms that may be implemented on a computational platform, it is important to establish the complexity of the nonlinear dynamic filtering in terms of operations and computational calculations.

## 2.1. Chaotic filtering and modeling

A chaotic process is defined from a set of ordinary differential equations and their corresponding parameters, i.e., is a deterministic process [17]. In the phase space, a chaotic process shapes an orbital trajectory, with the peculiar characteristic that none of all possible trajectories passes twice for exactly the same place [17]. In addition, a chaotic process is sensitive to changes in the initial conditions, i.e., two realizations of the same chaotic process whose initial conditions differ in an arbitrarily small value, are completely uncorrelated in the medium and long term [17].

Although being deterministic, a chaotic process generates realizations of processes that are described as stochastic, and it is precisely this deterministic-stochastic nature what is exploited to generate filtering strategies (with deterministic equations) and model physical processes (with stochastic realizations).

A vector chaotic process  $\mathbf{x}(t)$  may be generated by means of the following ordinary differential equation (ODE):

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t) \quad (1)$$

with initial condition  $\mathbf{x}(t_0) = \mathbf{x}_0$ ;  $\mathbf{F}(\cdot)$  is a time-varying vector function (represents the chaotic equations). In this case,  $\mathbf{F}(\mathbf{x}, t)$  is considered taking as example the chaotic attractors of Rossler, Lorenz and Chua:

Rossler

$$\begin{aligned} x_{k+1} &= x_k + T_s(-y_k - z_k) \\ y_{k+1} &= y_k + T_s(x_k - 0.2y_k) \\ z_{k+1} &= z_k + T_s(0.2 - z_k(5.7 - x_k)) \end{aligned} \quad (2)$$

Lorenz

$$\begin{aligned} x_{k+1} &= x_k + T_s(10(x_k - y_k)) \\ y_{k+1} &= y_k + T_s(28x_k - y_k + x_k \cdot z_k) \\ z_{k+1} &= z_k + T_s\left(-\frac{8}{3}z_k + x_k \cdot y_k\right) \end{aligned} \quad (3)$$

Chua

$$\begin{aligned} x_{k+1} &= x_{k+1} + T_s[9.205(y_k - U(x_k))] \\ y_{k+1} &= y_{k+1} + T_s[x_k - y_k + z_3] \\ z_{k+1} &= z_{k+1} + T_s[-14.3y_k] \end{aligned} \quad (4)$$

where  $U(x_k) = m_1 x_k + \frac{1}{2}(m_0 - m_1)(|x_k + 1| - |x_k - 1|)$ ,  $m_0 = -\frac{1}{7}$ ,  $m_1 = \frac{2}{7}$  and  $T_s$  is the sampling time. The discrete Kalman filter (described a little further) is the filtering algorithm used, for this reason, (2)-(4) are presented in discrete form.

The initial conditions of the signal to be processed are unknown in the filtering block; this produces uncertainty (divergence) effects that can be mitigated including an additive «process noise» in Equation (1). Thus, the ODE becomes a stochastic differential equation (SDE), which gives rise to an n-dimensional Markovian stochastic process:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t)) + \varepsilon \xi(t) \quad (5)$$

where  $\mathbf{f}(\mathbf{x}(t))$  is analogous to  $\mathbf{F}(\mathbf{x}, t)$  in (1). The influence of an external weak source of white noise is denoted as  $\xi(t)$ ; the intensities of the noise are given in form of the matrix  $\varepsilon = [\varepsilon_{ij}]^{n \times n}$ .

When using the SDE (5) as a model for the chaos, the first strategy that immediately comes to mind is the nonlinear filtering of chaotic signals developed rigorously in [18, 19], which is based on the Stratonovich-Kushner equations (SKE) [7, 8] that can be used to describe the dynamic equation of the a posteriori probability density function (PDF) of the chaos  $\mathbf{x}(t)$ . Note that the time evolution of the a posteriori PDF for  $x(t)$  is completely characterized by the SKE; however it does not have an exact analytical solution. One of the few exceptions is using a linear SDE, in other words, the well-known Kalman filter algorithm. Precisely, this is the reason why the nonlinear filtering algorithms are usually simplified making them quasi-optimal, or even quasi-linear.

The following question may arise: what are the advantages offered by the chaotic modeling for filtering weak signals? It turns out that the solution of the SKE for the dynamic ODE of the chaos (1), exhibits singularity properties when the solution is practically tuned with the deterministic chaos in (1), independently of the value of the SNR [18].

An important set of quasi-linear filtering algorithms apply the local Gaussian approximation (LGA) for the a posteriori PDF [8, 19], which results suitable for real time applications. Some of these algorithms are:

- Extended Kalman filter (EKF)
- Unscented Kalman filter (UKF)
- Quadrature Kalman filter (QKF)
- Gauss-Hermite quadrature filter (GHF)
- Conditionally optimal filter among others

Note that the difference between the algorithms based on the LGA, depends only on how the location of the instantaneous estimate of  $x(t)$  is chosen. For the case of a high filtering precision, all algorithms that apply LGA [18] may be successfully approximated by means of the EKF, because the correct value of the filtering process and the reference point for the application of the Gaussian approximation are obviously very close.

Given a certain SNR, all these filtering algorithms have different precision and a completely different computational complexity, for a pre-established filtering quality. When selecting any specific filtering algorithm for a particular scenario, it is necessary to consider the NMSE in conjunction with the computational complexity, as possible criteria.

## 2.2. Computational complexity

In real world applications, the computational complexity of the quasi-linear algorithms is essential. For the particular case of the EKF, UKF, QKF and GHF algorithms, the computational complexity may be analyzed in terms of additions and multiplications, Cholesky decomposition, nonlinear propagation and Jacobian calculation. The evaluation with respect to these terms is shown in Table 1.

**Table 1.** Computational complexity

	EKF	UKF	GHF	QKF
<b>Additions</b>	8	50	25	25
<b>Multiplications</b>	15	77	33	40
<b>Cholesky decomposition</b>	1	2	2	2
<b>Nonlinear Propagation</b>	0	15	21	6
<b>Jacobian calculation</b>	1	0	0	0

It can be noted that the UKF exhibits a greater complexity, while the EKF is the less complex. The EKF might degrade due to the Jacobian calculation (evaluation of the partial derivatives), if the equations of the attractor are sophisticated. However, for the models of formulas (2)-(4), the filtering structure based on the EKF is the best choice. As will be shown later, when detecting real weak signals, a rather acceptable fidelity may be achieved in all practical cases using a filtering structure based on the EKF, which internally uses chaotic models of the type given by (2)-(4).

As an alternative to the quasi-linear algorithms of the EKF, where the linearization is updated instantaneously, a robust and low computational complexity solution may be searched for using a «fixed linearization» (with a linearization matrix predefined according to the specific problem under study) instead of instantaneous linearization. In fact, this means that the standard Kalman filter (SKF) [7–11] would be used

instead of the EKF, and therefore, even though there would be a lower complexity, there would also be losses in the filtering precision. Nevertheless, it should be taken into account that the LGA of the a posteriori PDF assumes that all the components of the model are almost linear, and therefore the losses in precision may be reasonable.

Note that when the input data are variant, it is very common that the quasi-optimal filtering algorithms apply linearization strategies.

## 2.3. Multi-moment processing

To improve the filtering fidelity, it is required to take advantage of all the available information in the signal being processed. For this purpose, it makes sense to incorporate in the filtering methodology, applying the SKE equations, additional information (in different sequential time instants) of the received composite signal; in other words, it should be considered information in the form of blocks (at different time instants, i.e., multi-moment processing). The multi-moment algorithms are implemented through the generalization of the SKE using multi-moment data.

The multi-moment filtering algorithms are slightly practical for real time implementations, since the delay due to the processing of samples from different time instants is significant. To reach a compromise between complexity and increase of filtering precision, it is reasonable to consider processing only two adjacent samples. This processing is known as two-moments regime (2MM), which is a special case of the multi-moment filtering and may be reviewed in detail in [18, 19].

In the 2MM regime, two samples coming from two time instants (not instantaneous processing) are processed during every filtering cycle, and consequently the correlation coefficient between the two adjacent samples (denoted as  $\rho$ ) is a design parameter. The advantage of considering the 2MM regime, is that the benefits of the multi-moment processing can be obtained, practically without significant delays.

In the one-moment (1MM) regime, in which one sample from a single time instant is processed during each cycle (instantaneous processing), there is no increase in the filtering precision, and this is precisely the type of processing that characterizes the EKF and its variants previously listed.

## 2.4. Simulation model

The practical implementation of the proposed design methodology, was carried out developing a simulation test bench based on MatLab.

The methodology is constituted by two elements:

- 1) A nonlinear dynamic filtering structure based on the Kalman filter (EKF or SKF according to the case).

- 2) Appropriate modeling of the real weak signal, congruent with the filtering structure to be used (EKF or SKF).

A discrete version of the Kalman filter is used for simulation by means of MatLab. Such discrete version is now described. The state dynamics in a discrete system is given by:

$$\begin{aligned} x_{k+1} &= f(x_k) + \varepsilon_k \\ y_k &= s(x_k) + n_{0ks} \end{aligned} \quad (6)$$

where  $x_k$  represents the state of the system, and  $y_k$  is a measurement of the state of the system,  $\{n_{0k}\}$  and  $\{\varepsilon_k\}$  are independent Gaussian white noise processes with zero mean and covariance matrices

$$E[n_{0k}n_{0k}^T] = N_{0k} \quad \text{and} \quad E[\varepsilon_k\varepsilon_k^T] = Q_k$$

respectively,  $Q_k$  denotes process noise and measuring noise. The prediction and correction cycles of the Kalman filter are given by:

$$\begin{aligned} \text{Prediction} & \begin{cases} \hat{x}_{k+1}^- = f(x_k^+) \\ P_k^- = A_k P_k^+ A_k^T + Q_k \end{cases} \\ \text{Correction} & \begin{cases} G_k = P_k^- H_k^T [H_k P_k^- H_k^T + N_{0k}]^{-1} \\ \hat{x}_k^+ = \hat{x}_k^- + G_k [y_k - s(x_k^-)] \\ P_k^+ = P_k^- + G_k H_k P_k^- \end{cases} \end{aligned} \quad (7)$$

where  $G_k$  is the Kalman gain,  $\hat{x}_k^-$  is the a priori estimate of the state at the  $k$ -th update cycle,  $\hat{x}_k^+$  is the a posteriori estimate of the state at the  $k$ -th update cycle,  $P_k^-$ ,  $P_k^+$  are, respectively, the a priori and the a posteriori estimations of the covariance matrices at the  $k$ -th stage,  $A_k$  is the linearization matrix (or state-transition matrix) and  $H_k$  is the matrix that indicates the relationship between the measurement and the state vector at the  $k$ -th cycle, assuming absence of noise. For the case of the SKF,  $A_k$  is a fixed matrix in every cycle, while for the EKF the matrix is updated in each cycle by means of the Jacobian calculation:

$$A_k = \frac{\partial f(x_k)}{\partial x_k} \quad (8)$$

When using the EKF, the structure of the filter is given by (7); (2), (3) or (4) are used as the function  $f(x_k)$ , and the linearization is given by (8). Here, the real weak signal is modeled using any of the chaotic processes (2)-(4), i.e., it is analyzed which of the components ( $x$ ,  $y$  or  $z$ ) of (2), (3) or (4) is the most appropriate to be used as model. For this purpose, the sampling time ( $T_s$ ) of the discrete chaotic equations is first modified, until reaching a coincidence between the time variations of the component of the selected

chaotic attractor and the real signal (make the time scales as close as possible). Second, the desired signal is normalized with respect to the mean and the variance of the component of the attractor. It has been demonstrated in [18, 19], that the  $x$  component of the three chaotic attractors (2)-(4) is the more appropriate to model experimental signals. After carrying out a similar modeling analysis, it was found that the same criterion is applicable for the phenomenological signals of this work.

When using the SKF, the matrix  $A_k$  is fixed and the modeling of the signal should be reflected precisely in  $A_k$ . With this purpose, the MatLab System Identification Toolbox (SIT) [20] is used. Such tool is based on the theory of systems identification [21]. To identify the real signal (seismic, ECG, voice-type, RFI), the MatLab SIT analyzes its spectral properties, and gives a constant matrix as the model.

The experimental part of the next section shows that it is possible to use such matrix precisely as a fixed linearization matrix in the filtering structure given by (7), which really is only an approximation of the instantaneous linearization procedure required for the quasi-optimal filtering when the LGA is used.

To make a fair comparison with the EKF of dimension 3 according to formulas (2)-(4), a tridimensional SKF is designed. The «ident» command is used to obtain matrix  $A_k$  by means of the MatLab SIT. The identification (of the signal without noise) is made selecting the option «state space models» [7, 22] for the three-dimensional case. The program provides three estimation options (subspace method, regularized reduction and minimization of the prediction error), and at the end indicates the confidence percentage of each option. It was found experimentally that the option of minimization of the prediction error offers the best confidence percentage, for the estimation of the matrix.

In the methodology proposed here, the systems identification is precisely a modeling of the real signal. The identification is made for a vector (the largest possible) of the real signal (without noise). Assuming that the signals under study are stationary, the system identification matrix may be considered as the fixed linearization matrix  $A_k$  and, therefore, be a signal model for any other vector of the same phenomenological signal. When using the identification matrix in the structure of the SKF, a processing with a priori information and experimental data is achieved.

The 2MM regime is employed to achieve a greater filtering fidelity. As it is commented in [19], the quasi-optimal solutions (in this case for multi-moment algorithms) are based in some heuristic that may incorporate previous knowledge and/or structures. If this is the case, the 2MM regime utilized in this work has the form:

$$\begin{aligned}
& \hat{x}_{k+1}^- = f(\hat{x}_k^+) \\
& P_k^- = A_k P_k^+ A_k^T + Q_k \\
& G_k = P_k^- H_k^T [H_k P_k^- H_k^T + N_{0k}]^{-1} \\
& \hat{x}_k^+ = \hat{x}_k^- + G_k [y_k - s(x_k^-)] \\
& P_k^+ = P_k^- + G_k H_k P_k^-
\end{aligned}
\quad
\begin{aligned}
& {}_2\hat{x}_{k+1}^- = f({}_2\hat{x}_k^+) \\
& {}_2P_k^- = {}_2A_k {}_2P_k^+ (1 - \rho^2) {}_2A_k^T + Q_k \\
& {}_2G_k = {}_2P_k^- {}_2H_k^T [{}_2H_k {}_2P_k^- {}_2H_k^T + N_{0k}]^{-1} \\
& {}_2\hat{x}_k^+ = {}_2\hat{x}_k^- + {}_2G_k [y_k - s({}_2x_k^-)]
\end{aligned}
\quad (9)$$

where the subscript 2 of the left side of each variable denotes a 2MM variable. For this filtering algorithm, the only difference between the left and right columns of the structure given by (7), is that the last operation of the correction cycle does not exist; both columns operate in parallel. In (7), the output of the filter is  $\hat{x}_k^+$ , while in (9) is  ${}_2\hat{x}_{k+1}^-$ . In the 2MM regime, in the column of the right side in the last operation of the prediction cycle, the a priori estimate of the covariance matrix is calculated taking into account the correlation coefficient  $\rho$  of the two samples. Observe that the structure (9) may operate both for the EKF and the SKF, following the observations described in previous paragraphs.

Next section experimentally shows the efficiency of the filtering proposed here, considering signals of significantly different nature, such as, seismic signals, fetal electrocardiographic (FECG) signals, voice-type signals and radiofrequency interfering signals (RFI). Such signals may be considered chaotic [1, 2, 12, 15, 23, 24]

### 3. Results and discussion

The following results show the filtering by means of SKF and EKF, both for the 1MM and 2MM regimes. In the 2MM regime, the  $\rho$  parameter determines a different fidelity in the filtering. If  $\rho = 0$ , this corresponds to the 1MM regime (without increase in the fidelity). If  $\rho = 1$ , see formula (9), this corresponds to a singularity condition with a covariance matrix equal to zero (fidelity tends to  $+\infty$ ). It was chosen  $\rho = 0.85$  for a homogeneous analysis of the results.

The figures show the overlapped curves of the original signal (without noise), and the signal after being

filtered with the filtering scheme and the corresponding regime indicated in each figure. The figures only show a threshold case of weak signal when the  $SNR = \sim 3dB$ . Manipulating  $N_0$  to analyze different thresholds of weak signals ( $SNR \leq 0dB$ ), Tables 2-5 show the performance of the SKF and EKF under the 1MM and 2MM regimes, in terms of the NMSE (described in the paragraphs of the introduction). The average times (in seconds) required to process 5000 samples for each of the phenomena under study and their corresponding filtering, are also shown.

For both the SKF and EKF, an imperfect modeling of the phenomenological signal is carried out (there is a certain degree of uncertainty in the initial conditions for the filtering); as a consequence, a noise value of weak process, see  $Q$  in (7) and (9), should be included in the filtering structure (value of  $Q$  indicated in tables 2-5).

No results of the SKF are presented for the seismic signals, since it was not possible to obtain the corresponding fixed linearization matrix (systems identification matrix) because such signals have a very limited duration for an adequate spectral analysis by means of the SIT.

#### 3.1. Fetal electrocardiographic (FECG) signals

The experimental data were obtained from the PhysioNet [25] database. The signal for this experiment corresponds to the heart of a fetal product, at week 36 of pregnancy. For a  $SNR = \sim 3dB$ , Figure 1 shows the original signal and the signal filtered using the EKF 1MM with the x component of Rossler as model. Table 2 shows the complete results.

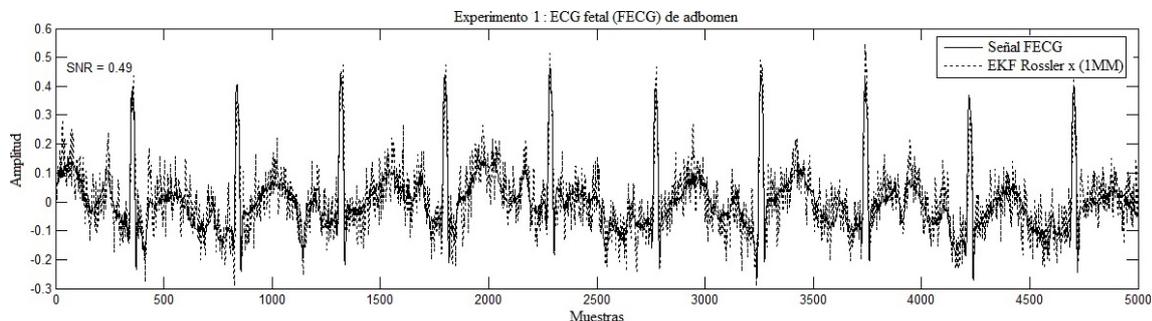


Figure 1. Signals in experiment 1

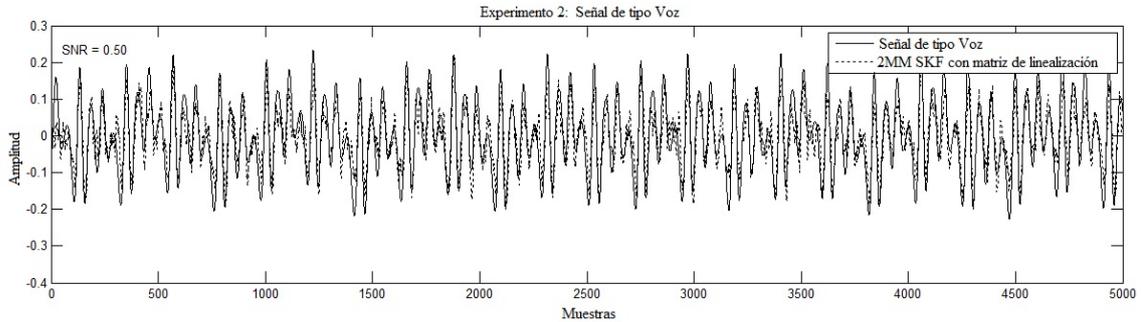
**Table 2.** Results of the NMSE for experiment 1

SNR	0 dB	-3 dB	-10 dB	Procesamiento time
SKF $Q = 0.04$ (with linearization matrix $A_k$ )				
1MM	0.0025	0.0037	0.0078	0.43 s
2MM	0.0021	0.0032	0.0065	0.89 s
EKF Rossler x $Q = 0.21$				
1MM	0.0026	0.0040	0.0098	1.825 s
2MM	0.0023	0.0036	0.0079	3.503 s
EKF Lorenz x $Q = 0.42$				
1MM	0.0029	0.0042	0.010	1.782 s
2MM	0.0023	0.0034	0.0083	3.59 s
EKF Chua x $Q = 0.075$				
1MM	0.0034	0.0053	0.015	1.812 s
2MM	0.0026	0.0042	0.012	3.61 s

### 3.2. Experiment 2. Voice-type signals

Sustained sounds of vowels (vowel «o») were used for this experiment. This type of signals are utilized in voice synthesis procedures [23]. Figure 2 shows with solid line the sustained sound of the vowel «o» (recorded during 5 seconds at 22050 Hz), and with

dotted line the signal filtered using the SKF 2MM with its matrix evaluated by means of SIT. The results (very similar to the previous experiment) are shown in Table 3. For this experiment none of the components of the Lorenz attractor resulted adequate for modeling voice-type signals.

**Figure 2.** Signals in experiment 2**Table 3.** Results of the NMSE for experiment 2

SNR	0 dB	-3 dB	-10 dB	Processing time
SKF $Q = 0.0081$ (with linearization matrix $A_k$ )				
1MM	0.0025	0.0037	0.0079	0.47 s
2MM	0.0015	0.0024	0.0053	0.95 s
EKF Rossler x $Q = 0.23$				
1MM	0.0029	0.0044	0.0124	1.792 s
2MM	0.0027	0.0039	0.011	3.611 s
EKF Chua x $Q = 0.76$				
1MM	0.0031	0.0048	0.0137	1.81 s
2MM	0.0025	0.0043	0.0130	3.58 s

### 3.3. Experiment 3. Seismic signals

A MatLab simulator based on the seismic models reported in [26] was used for this experiment. For an  $SNR = \sim 3dB$ , Figure 3 shows a seismic signal and

its filtered version obtained using the EKF 2MM with the x component of the Rossler model. The complete results are shown in Table 4. For the seismic signal it was not possible to obtain the linearization matrix, and consequently the SKF is not reported for this case.

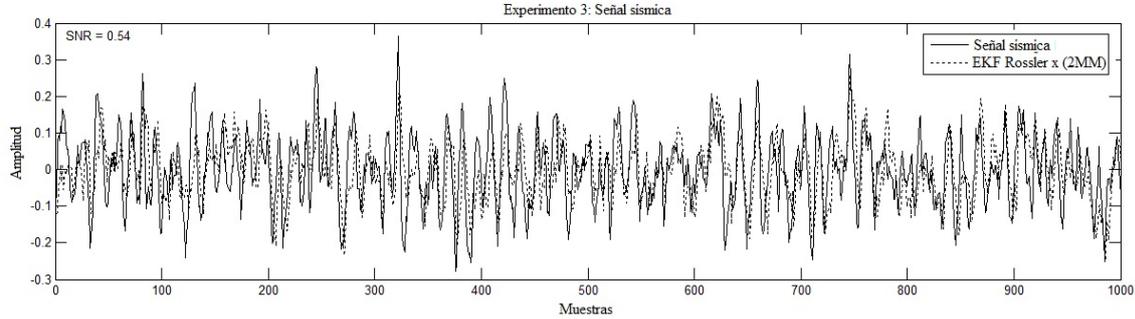


Figure 3. Signals in experiment 3

Table 4. Results of the NMSE for experiment 3

SNR	0 dB	-3 dB	-10 dB	Processing time
EKF Rossler x Q = 0.35				
1MM	0.0048	0.0074	0.0178	1.79 s
2MM	0.0047	0.0073	0.0135	3.53 s
EKF Lorenz x Q = 0.135				
1MM	0.0058	0.0093	0.0245	1.807 s
2MM	0.0054	0.0081	0.0187	3.62 s
EKF Chua x Q = 0.135				
1MM	0.0057	0.0095	0.029	1.816 s
2MM	0.0051	0.0084	0.023	3.65 s

### 3.4. Experiment 4: RFI signals

This experiment considers the RFI generated by computer equipment [24,27], which affects the transmission of desired information signals. For an  $SNR = \sim 3dB$ ,

Figure 4 shows the RFI signal and its filtered version obtained using the SKF 1MM, with the linearization matrix evaluated by means of the SIT. Table 5 shows the complete results.

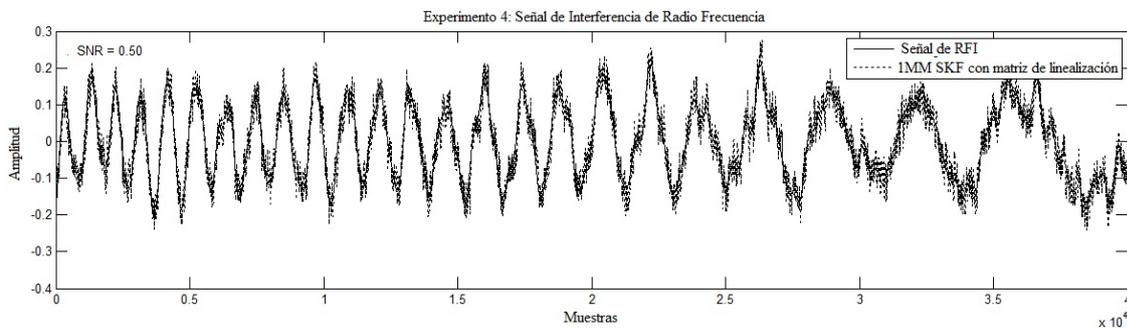


Figure 4. Signals in experiment 4.

**Table 5.** Results of the NMSE for experiment 4

SNR	0 dB	-3 dB	-10 dB	Processing time
SKF $Q = 0.02$ (with linearization matrix $A_k$ )				
1MM	0.0018	0.003	0.0098	0.51 s
2MM	0.0015	0.0025	0.0085	0.92 s
EKF Rossler x $Q = 0.2$				
1MM	0.0026	0.005	0.019	1.872 s
2MM	0.0023	0.0036	0.011	3.9 s
EKF Lorenz x $Q = 0.6$				
1MM	0.0023	0.0032	0.04	1.76 s
2MM	0.0016	0.0027	0.0083	3.81 s
EKF Chua x $Q = 0.4$				
1MM	0.0034	0.0053	0.035	1.86 s
2MM	0.0026	0.0042	0.029	3.77 s

From the tables it is observed that the 2MM method shows a slightly better NMSE. All the filtering methodologies presented are rather effective, because they are characterized by a very low value of the NMSE. In the scenario of  $SNR = \sim 10dB$  (an extremely weak signal), it is virtually impossible to visually distinguish (it is not shown in figures due to space issues) the desired signals from the noise; nevertheless, the NMSE has values around 1 % for the filtering by means of SKF and EKF, both for the 1MM and 2MM methodologies.

It should be taken into account that the 2MM methodology consumes more time, compared to the 1MM methodology, although it is not more than twice the time. The filtering by means of SKF is (almost 3 times) faster, because there is no linearization process. The processing times together with the complexity and the filtering fidelity, should be the selection criteria to choose the adequate filtering algorithm for each particular implementation.

The SKF with fixed linearization matrix (modeling the signals of interest by means of a system identification matrix) shows the best results, which indicates that for processing with quasi-linear filtering algorithms, the influence of the spectral properties of the input data prevails over the influence of the non-Gaussian statistics. The values of NMSE obtained in the simulations are very similar for filtering the different signals, and consequently, it is not really determinant in practice which particular model of chaotic attractor or linearization matrix (obtained from the SIT) is applied.

Why does this occur? All the chaotic attractors that have been employed to model the real signals, generate chaos as a quasi-deterministic stochastic process. For this reason, all the aforementioned quasi-optimal

filtering algorithms (including the EKF and its modifications) which apply the idea of chaotic modeling, operate in a regime very close to the singularity, i.e., the shape of the a posteriori PDF is concentrated around the a priori PDF of the desired signal, independently of the SNR value [18, 19], and that precisely enables obtaining rather low values of the NMSE for very weak signals (SNR smaller than 0 dB and up to -10 dB). Therefore, for high fidelity filtering, the linear term of the Taylor series expansion for the quasi-linear algorithm [8, 9, 22] is significantly more determinant than the terms related to the nonlinearities (Jacobian matrix, etc.), i.e., the linear approximation is sufficient.

## 4. Conclusions

It has been proposed a simple and robust filtering structure based on the Kalman filter, for processing weak signals. This structure enables the incorporation of the 2MM regime, which improves the filtering precision.

For non-Gaussian phenomenological signals, depending on the specific scenario, requirements of computational complexity and acceptable error, it may be employed as model for the real signal either the EKF using chaotic signals, or the SKF with a fixed linearization matrix; in other words, the MatLab SIT may be used to evaluate the systems identification matrix, which is then used as model of the real signal. This enables a «significant degree of freedom» for the design of the filtering block.

The experimental results show great precision in filtering weak signals for all scenarios considered here, and given the rather diverse nature of such scenarios, most likely it may be applied in other scenarios (future work).

## References

- [1] Y. Li, B. Yang, Y. Yuan, X. Zhao, and H. Lin, "Ability to detect weak effective seismic signals by utilizing chaotic vibrator system," *Chinese Science Bulletin*, vol. 51, pp. 3010–3017, 2006. [Online]. Available: <https://doi.org/10.1007/s11434-006-2191-0>
- [2] Y. Li, B. J. Yang, J. Badal, X. P. Zhao, H. B. Lin, and R. L. Li, "Chaotic system detection of weak seismic signals," *Geophysical Journal International*, vol. 178, no. 3, pp. 1493–1522, 2009. [Online]. Available: <https://doi.org/10.1111/j.1365-246X.2009.04232.x>
- [3] J. Shu-Yao, Y. Fei, C. Ke-Yu, and C. En, "Application of stochastic resonance technology in underwater acoustic weak signal detection," in *OCEANS 2016 - Shanghai*, April 2016, pp. 1–5. [Online]. Available: <https://doi.org/10.1109/OCEANSAP.2016.7485567>
- [4] S. L. Joshi, R. A. Vatti, and R. V. Tornekar, "A survey on ecg signal denoising techniques," in *2013 International Conference on Communication Systems and Network Technologies*, April 2013, pp. 60–64. [Online]. Available: <https://doi.org/10.1109/CSNT.2013.22>
- [5] H. Li, R. Wang, S. Cao, Y. Chen, N. Tian, and X. Chen, "Weak signal detection using multiscale morphology in microseismic monitoring," *Journal of Applied Geophysics*, vol. 133, pp. 39–49, 2016. [Online]. Available: <https://doi.org/10.1016/j.jappgeo.2016.07.015>
- [6] R. Han, J. Li, G. Cui, X. Wang, W. Wang, and X. Li, "Seismic signal detection algorithm based on gs transform filtering and emd denoising," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec 2018, pp. 1213–1217. [Online]. Available: <https://doi.org/10.1109/CompComm.2018.8781035>
- [7] H. Van Trees, *Detection, Estimation, and Modulation Theory: Detection, Estimation, and Linear Modulation Theory*, 2001. [Online]. Available: <https://doi.org/10.1002/0471221082>
- [8] A. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970. [Online]. Available: <https://bit.ly/34iDlnz>
- [9] R. L. Stratonovich, *Topics of the Theory of Random Noise*, 1967. [Online]. Available: <https://bit.ly/2PcK2DB>
- [10] W.-R. Wu and A. Kundu, "Image estimation using fast modified reduced update kalman filter," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 915–926, April 1992. [Online]. Available: <https://doi.org/10.1109/78.127963>
- [11] R. Parseh, K. Kansanen, and D. Slock, "Distortion outage analysis for joint space-time coding and kalman filtering," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2291–2305, May 2017. [Online]. Available: <https://doi.org/10.1109/TSP.2017.2655493>
- [12] M. de Sousa Vieira, "Chaos and synchronized chaos in an earthquake model," *Physical Review Letters*, vol. 82, pp. 201–204, Jan 1999. [Online]. Available: <http://doi.org/10.1103/PhysRevLett.82.201>
- [13] G. Kolumban, M. P. Kennedy, and L. O. Chua, "The role of synchronization in digital communications using chaos. i . fundamentals of digital communications," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 10, pp. 927–936, Oct 1997. [Online]. Available: <https://doi.org/10.1109/81.633882>
- [14] G. Kaddoum, "Design and performance analysis of a multiuser ofdm based differential chaos shift keying communication system," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 249–260, Jan 2016. [Online]. Available: <https://doi.org/10.1109/TCOMM.2015.2502259>
- [15] N. V. Thakor, "Chaos in the heart: signals and models," in *Proceedings of the 1998 2nd International Conference Biomedical Engineering Days*, May 1998, pp. 11–18. [Online]. Available: <https://doi.org/10.1109/IBED.1998.710541>
- [16] P. J. García-Laencina and G. Rodríguez-Bermudez, "Analysis of eeg signals using nonlinear dynamics and chaos: A review," *Applied Mathematics & Information Sciences*, pp. 2309–2321, 2015. [Online]. Available: <https://bit.ly/2Pg2NGa>
- [17] V. S. Anishchenko, V. Astakhov, A. Neiman, T. Vadivasova, and L. Schimansky-Geier, *Nonlinear Dynamics of Chaotic and Stochastic Systems*. Springer-Verlag Berlin Heidelberg, 2007. [Online]. Available: <http://doi.org/10.1007/978-3-540-38168-6>
- [18] V. Kontorovich and Z. Lovtchikova, "Nonlinear filtering of chaos for real time applications," in *Selected Topics in Nonlinear Dynamics and Theoretical Electrical Engineering*, K. Kyamakya, W. A. Halang, W. Mathis, J. C. Chedjou, and Z. Li, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 41–59. [Online]. Available: [https://doi.org/10.1007/978-3-642-37781-5\\_3](https://doi.org/10.1007/978-3-642-37781-5_3)

- [19] V. Kontorovich, Z. Lovtchikova, and F. Ramos-Alarcón, *Nonlinear Filtering of Weak Chaotic Signals*, 2017. [Online]. Available: <http://doi.org/10.5772/intechopen.70717>
- [20] E. Nikulchev, “Simulation of robust chaotic signal with given properties,” *Advanced Studies in Theoretical Physics*, vol. 8, no. 21, pp. 939–944, 2014. [Online]. Available: <http://dx.doi.org/10.12988/astp.2014.48106>
- [21] L. Ljung, *System Identification: Theory for the user*. Prentice Hall PTR, 1999. [Online]. Available: <https://bit.ly/2YIDEaj>
- [22] V. Pugachev and I. Sinitsyn, *Stochastic Differential systems Analysis and Filtering*. John Wiley & Sons, 1987. [Online]. Available: <https://bit.ly/2qKkrZc>
- [23] M. Banbrook, S. McLaughlin, and I. Mann, “Speech characterization and synthesis by nonlinear methods,” *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 1, pp. 1–17, Jan 1999. [Online]. Available: <https://doi.org/10.1109/89.736326>
- [24] D. Iltzky Arditti, A. Alcocer Ochoa, V. Kontorovich Mazover, and F. Ramos Alarcón Barroso, “Adaptive mitigation of platform-generated radio-frequency interference,” Patent US 2015.0051880A1, 2015. [Online]. Available: <https://bit.ly/2Pesh6G>
- [25] Physionet. (2019) Physiobank atm. [Online]. Available: <https://bit.ly/2LMxcKf>
- [26] F. R. Rofooei, A. Mobarake, and G. Ahmadi, “Generation of artificial earthquake records with a nonstationary kanai-tajimi model,” *Engineering Structures*, vol. 23, no. 7, pp. 827–837, 2001. [Online]. Available: [https://doi.org/10.1016/S0141-0296\(00\)00093-6](https://doi.org/10.1016/S0141-0296(00)00093-6)
- [27] E. X. Alban, M. E. Magana, H. G. Skinner, and K. P. Slattery, “Statistical modeling of the interference noise generated by computing platforms,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 3, pp. 574–584, June 2012. [Online]. Available: <https://doi.org/10.1109/TEM.2011.2164255>

OTHER PERIODIC PUBLICATIONS OF THE UNIVERSITY  
UNIVERSITAS, Journal of Social and Human Sciences.

LA GRANJA, Journal of Life Sciences.

ALTERIDAD, Journal of Education.

RETOS Journal of Administration Sciences and Economics.

UTOPIA, University Youth Ministry Magazine.

SOPHIA, Collection of Philosophy of Education.

