



**UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH**

**Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona**

**Chirp-based Direct Phase Modulation of VCSELs managed
by Neural Networks**

A Master's Thesis

Submitted to the Faculty of the

**Escola Tècnica d'Enginyeria de Telecomunicació de
Barcelona**

Universitat Politècnica de Catalunya

by

Johanna Libertad Kais Carranza

In partial fulfilment

of the requirements for the degree of

MASTER IN TELECOMMUNICATIONS ENGINEERING

Advisor: José Antonio Lázaro Villa

Co Advisor: Samael Sarmiento Hernández

Barcelona, May 2019



Title of the Thesis: Chirp-based Direct Phase Modulation of VCSELS managed by Neural Networks

Author: Johanna Libertad Kais Carranza

Advisor: José Antonio Lázaro Villa

Co Advisor: Samael Sarmiento Hernández

Abstract

VCSEL's capacity of direct modulation and its low cost makes this device a feasible cost-effective transmitter for ultra-dense wavelength division multiplexing (uDWDM) metro-access networks using coherent detection. However, performing direct-phase modulation in semiconductors can be complex due to its nonlinear characteristics. This research presents Neural Network (NN) training techniques for Time-Series analysis in order to describe the correlation between the input current given to the device and its output optical phase, using a 1550nm RayCan SM-VCSEL. Main goal is training a NN capable of predicting an ideal optical power signal for a specific phase result achievable by inverse training, that is: optical phase is the neural network input while the optical power is the desired target. The experiment is done in three stages: (i) VCSEL's characterization, (ii) NN training to predict input current knowing optical power, and (iii) NN training to predict optical power from a known optical phase.

Keywords: *VCSEL, Neural Network, Time-Series Analysis.*



You leave home to seek your fortune and, when you
get it, you go home and share it with your family.

- *Anita Baker*

No hubiese podido hacerlo sin ustedes,
considérenlo suyo tanto como es mío.

JLKais.

Acknowledgements

To my parents being always through thick and thin with me, my sister and brother for being the reason I push myself harder every day. Yaya, for you the words are never enough.

Vero, my little partner, you have been the light in the darkness.

To May, you do Rock! Thanks for being here all the way.

I wish I could give enough credit to all of the people that had help in this process. To all of you, thanks for caring.

Special thanks for José Antonio Lázaro and Samael Sarmiento, you have been not only guidance but strength.

Revision History and Approval Record

Revision	Date	Purpose
0	22/05/2019	Document creation
1	23/05/2019	Document revision

Written by:		Reviewed and Approved by:	
Date	22/05/2019	Date	23/05/2019
Name	Johanna Libertad Kais Carranza	Name	José Antonio Lázaro Villa
Position	Project Author	Position	Project Supervisor

Table of Contents

Abstract.....	iii
Acknowledgements	v
Revision History and Approval Record	vi
Table of Contents.....	vii
List of Figures.....	ix
List of Tables.....	xii
1. Introduction.....	1
2. State of the Art.....	6
2.1. Timeline for Optical Technology.....	6
2.2. From LASERs to VCSELs.....	11
2.3. Neural Networks as Deep Machine Learning	20
2.3.1. The Artificial Neuron.....	23
2.3.2. Network's Basic Structure	24
2.3.3. Neural Network Learning Techniques	25
2.4. Time Series Prediction using Neural Networks.....	27
2.4.1. Prediction in One-time Step.....	28
2.4.2. Prediction on Multiple Steps.....	30
3. Methodology - Project Development	33
3.1. VCSEL's Characterization.....	33
3.1.1. New footprint and symbol for the VCSEL	33
3.1.2. Designing and manufacturing the board.....	35
3.1.3. Light-Current-Voltage (L-I-V) Curves.....	37
3.2. Creating and Training a Neural Network	41
3.2.1. MATLAB's Neural Networks Toolbox	41
3.2.2. Designing, Training and using a Basic Neural Network	42
3.3. Training a Neural Network to estimate a VCSEL's input current given its output optical power.....	45
3.4. Training a Neural Network to predict a VCSEL's optical power given its corresponding optical phase	47
4. Results	49

4.1. VCSEL's Characterization.....	49
4.2. Training a Neural Network to estimate a VCSEL's input current given its output optical power.....	51
4.3. Training a Neural Network to predict a VCSEL's optical power given its corresponding optical phase	56
5. Budget	61
5.1. Available Resources	61
5.2. Components List and Costs	62
6. Conclusions and Future Development	63
Bibliography	65
Appendices	68
Appendix A – VCSEL's Characteristics.....	68
Appendix B – VCSEL's First Board Design.....	69
Appendix C – Definitive VCSEL Board Design with Dimensions	70
Appendix D – VCSEL after inclusion on board with dissipation element	71
Appendix E – Resistance – Temperature Table for VCSEL's Thermistor	72
Appendix F – Front and Rear Panel for LDC200C	73
Appendix G – Front and Rear Panel for TED200C.....	75
Appendix H – Values obtained from VCSEL's Characterization	77
Appendix I – Diagrams for MATLABs Deep Learning Toolbox solvers in Dynamic Time-Series Analysis	82
Glossary.....	83

List of Figures

Figure 1.- Flexible 5G Metro-Access Network scenario. Inlet: considered flexible ultra-dense WDM frequency slot division [2]	1
Figure 2.- Modulation Signal at 1.25 Gb/s [1].....	3
Figure 3.- Transient and Adiabatic Chirp [1]	3
Figure 4.- Total, Transient and Adiabatic Optical Phase [1]	4
Figure 5.- Experimental results: Optical phase eye diagram and IQ diagram [1].....	4
Figure 6.- Illustration of the optical Telegraph and its inventor Claude Chappe [6].....	7
Figure 7.- BL product increment in different lightwave systems [8]	8
Figure 8.- First FLAG link, Europe-Asia (FEA) segment [12].....	9
Figure 9.- Submarine Optic Fibers Nowadays [14]	10
Figure 10.- Increase of capacity in Lightwave Systems after 1980 [6].....	10
Figure 11.-Block diagram of an optical transmitter [6].....	12
Figure 12.- The Emission-Absorption Principle [8]	12
Figure 13.- Coherent Laser Light (a) and Incoherent LED Light (b) [17]	13
Figure 14.- Characteristics for a 1300nm-band InGaAsP/InP Semiconductor Diode [21].....	14
Figure 15.- Schematic of a Semiconductor Laser [8]	15
Figure 16.- Outline of a DFB laser structure [23].....	16
Figure 17.- Schematic longitudinal cross-section of three-section DBR laser [23].....	16
Figure 18.- Lasing direction for edge and surface emitting devices [24].....	17
Figure 19.- Metallic Mirror VCSEL Structure [22]	17
Figure 20.- Epitaxially Bragg Reflectors VCSEL structure [25].....	18
Figure 21.- VCSEL with DBR mirror on a cantilever [23].....	19
Figure 22.- Example of Linear Regression [29].....	21
Figure 23.- A basic decision tree for the papaya example [30].....	21
Figure 24.- General Scheme of a McCulloch-Pitts Cell [4]	22
Figure 25.- Three-Layered feed-forward fully-connected NN schematic [31]	24
Figure 26.- Convolutional Neural Network Architecture [33]	26
Figure 27.- Recurrent Neural Network Structure [31].....	27
Figure 28.- Recursive Neural Network Scheme [34]	27

Figure 29.- One-Step Time Prediction [4]	29
Figure 30.- Multiple Time Step Prediction [4].....	30
Figure 31.- Delay-based Multiple-time-step Predictor Architecture [4]	31
Figure 32.- Side and bottom view of VCSEL. Dimensions in millimeters (mm) [35].....	34
Figure 33.- Transversal and bottom view of the VCSEL with dimensions (in mm) [35]	34
Figure 34.- Footprint and symbol created for the VCSEL.....	35
Figure 35.- VCSEL board design. Left: software view. Right: pre-manufacturing stage.	36
Figure 36.- VCSEL included in manufactured board.....	36
Figure 37.- Pin assignment of the “LD OUT” jack (female, rear panel view) [36].....	37
Figure 38.- Cable to connect VCSEL and LDC.....	38
Figure 39.- Pin assignment of the “TE OUTPUT” jack (female, rear panel view) [37].....	38
Figure 40.- Connection cable for VCSEL and Temperature Controller.....	39
Figure 41.- General experimental setup for VCSEL characterization.	40
Figure 42.- Experimental setup for VCSEL characterization (optical power)	40
Figure 43.- Architecture of an example Neural Network.....	42
Figure 44.- MSE and R results from different training processes; BR and SCG respectively. Notice that the algorithm BR doesn't include validation values, as is not used to stop the training process for this case.	43
Figure 45.- Performance results for the training process using the algorithm LM.....	44
Figure 46.- Regression results for a training process using the algorithm BR.	44
Figure 47.- Comparison between original bits, exponentially modulated signal and the result from filtering the exponential signal.	46
Figure 48.- Experimental setup to obtain the optical power response from the VCSEL.	46
Figure 49.- VCSEL wavelength in terms of the bias current and the temperature.....	50
Figure 50.- Optical Spectrum signal from the VCSEL output. $I=7\text{ mA}$, $T= 15^{\circ}\text{C}$	50
Figure 51.- Comparison between the original pattern $I(t)$ and the received signal $P(t)$	51
Figure 52.- Regression result for NN training. Using Algorithm LM and training specifications: $D=4$, $HL=15$	52
Figure 53.- Comparing the original pattern $I(t)$ and the resized received signal $P_{\text{SCALED}}(t)$	53
Figure 54.- Performance of NN trained with resized signals. Algorithm LM, $D=6$, $HL=25$. (S-NN-A)	54
Figure 55.- Regression values for the NN trained with resized signals. Algorithm LM, $D=6$, $HL=25$ (S-NN-A).....	54

Figure 56.- Comparison between the $I_{S-NN}(t)$ and the original $I(t)$	56
Figure 57.- Result of computing Mean Squared Error (MSE) between original signal $I(t)$ and the predictions: $I_{S-NN-A}(t)$, $I_{S-NN-B}(t)$ and $I_{S-NN-D}(t)$, respectively	56
Figure 58.- An optical power signal $P(t)$ compared against the corresponding optical phase $Q(t)$	57
Figure 59.- Comparison between original $P(t)$ and $P_{NNx}(t)$ from NNs.....	59
Figure 60.- Result of computing Mean Squared Error (MSE) between original signal $P(t)$ and the predictions: $P_{NNa}(t)$, $P_{NNb}(t)$, $P_{NNc}(t)$, $P_{NNd}(t)$, $P_{NNe}(t)$, $P_{NNf}(t)$, and $P_{NNg}(t)$	59
Figure 61.- Comparison between original $Q(t)$ and computed $Q_{NNx}(t)$ from ODE45.....	60
Figure 62.- Results for Mean Squared Error (MSE) between original signal $Q(t)$ and computed $Q_{NNx}(t)$ from ODE45.....	60

List of Tables

Table 1.- Training pattern. One-time-step prediction [4].....	29
Table 2.- Training patterns. For delay-based Multiple-time-step Prediction [4].....	31
Table 3.- VCSEL's pinout-function correspondence, width for each pin and XY coordinates 35	
Table 4.- Connections for LD OUT pins [36].....	38
Table 5.- Connections for TE OUTPUT pins.....	39
Table 6.- Parameters of Delay and Hidden Layers used to train different NNs	52
Table 7.- Results from the training of NNs with expressed parameters in Table 6	53
Table 8.- Characteristics and Results for NNs trained using the scaled signals	55
Table 9.- Algorithms and Parameters used to trained the following NNs	58
Table 10.- Components Lists and Costs	62

1. Introduction

Lightwave systems are called in this way to distinguish them from microwave systems, mainly due to their difference in carrier frequency. While microwaves work at frequencies around gigahertz (GHz) range, optical (lightwave) communication systems use high carrier frequencies (around 100 terahertz [THz]). Lightwave technology includes fiber-optic communications systems where the information transmission is done through optical fibers.

Since their first commercial generation in the mid-1970s, better modulation schemes are being developed continuously to take advantage of this field, thanks to new optical sources and receivers, and optic-fiber as the communication channel. Main breakthroughs include: (i) the construction of optical sources from semiconductors that not only are being used in telecommunications, but also in medicine and national security projects; (ii) the intercontinental submarine optical network that communicates all continents around the world; (iii) the exploiting of wavelength division multiplexing (WDM) technique in lightwave systems that considerably increased optic-fiber communication channel capacity, and many others.

From the start of the 21st century, the traffic demand is growing exponentially due to cloud computing based services, the Internet of Things (IoT), demanding the convergence between wireless and optical communications in the new 5G paradigm. One practical and promising solution is the use of ultra-dense wavelength division multiplexing (uDWDM) metro-access networks using coherent detection (see Figure 1). However, cost-effective transmitters need to be investigated and designed in order to make economically affordable these new networks [1].

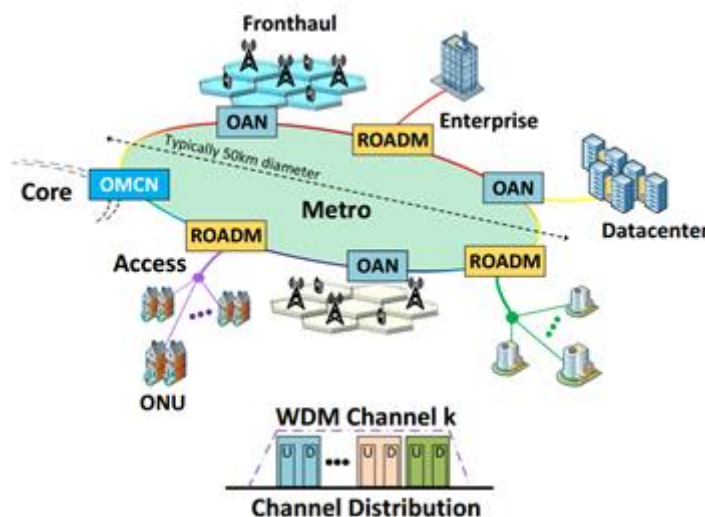


Figure 1.- Flexible 5G Metro-Access Network scenario. Inlet: considered flexible ultra-dense WDM frequency slot division [2]

Over the recent years, one of the proposed cost-effective devices to manage this issue is intensity modulated vertical-cavity surface-emitting lasers (VCSELs), in order to take advantage of both: his low cost, and its potential as low cost local oscillator for heterodyne receivers. Using these facts; Altabás, Izquierdo, Lázaro and Garces proposed in 2016 a 2.5

Gigabit per second (Gb/s) binary differential phase-shift keying (DPSK) transmitter based on direct-phase modulated VCSEL, achieving a -39.5 dBm sensitivity.

While higher data rates, e.g. 10 Gb/s, have been previously achieved by a VCSEL by more simple On-Off keying (OOK) modulation, the achieved 2.5 Gb/s DPSK with a direct-phase modulated VCSEL, was a first time demonstration of a VCSEL used as a phase modulation transmitter using its own chirp parameters, a complex process as the optical phase non-linear dependence with the input current of the VCSEL, as is further explained.

The obtained sensitivity: has only 5.5 dB power penalty compared to a 2.5 Gb/s DPSK implemented with a Mach-Zehnder Modulator (MZM), considered as nearly ideal transmitter, though at a cost 20-30 times higher, including the required MZM's drivers, taking into account that the cost of a used tuneable laser source (TLS) is above thirty thousand euros (€ 32,500) [39], while the VCSEL used in this experiment around two-hundred euros (€ 200, cost following purchasing invoice of UPC). Furthermore, has a sensitivity improvement of 2.75 dB in relation to a 2.5 Gb/s OOK modulation with the same VCSEL. All these good performances, using a low cost device, make this transmitter a promising cost-effective candidate for access networks [1].

The experiment first characterized the VCSEL in order to obtain the static and dynamic parameters. Measured static variables are: (i) lasing threshold, (ii) slope efficiency, (iii) wavelength and (iv) optical power, in terms of the bias current and temperature. On the other hand, the dynamic parameter is the device's frequency chirp and can be mathematically described as:

$$\Delta\nu(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \frac{\alpha}{4\pi} \left(\frac{1}{P(t)} \frac{dP(t)}{dt} + \kappa P(t) \right) \quad \text{Eq. 1}$$

where $\Delta\nu(t)$ is optical frequency shift, $\phi(t)$ instantaneous optical phase, and $P(t)$ is the instantaneous optical power. The *transient chirp* (α) is related to the variations of the emitting optical power, while the *adiabatic chirp* (κ) depends on the variations of the instantaneous emitted optical power. Measured values for each are 2.24 ± 0.1 and 7.6 ± 0.8 gigahertz per mill watts (GHz/mW), respectively, at a similar VCSEL than the one used for this TFM [1].

The used pulse was a sharp transition at the start of each symbol and an exponential decay after. Figure 2 shows the comparison between the ideal signal and the experimental sequence produced by the VCSEL noticing that the emitted signal follows better the expected one at higher current values, while at lower values the device response is slower. However, the "ideal" pattern for this signal was deducted from background knowledge on signal processing and the emitted signal doesn't quite follow the estimation; to take full advantage of the VCSEL capabilities for direct modulation, is necessary to find the signal that optimizes the symbols' transitions once the data is recovered.

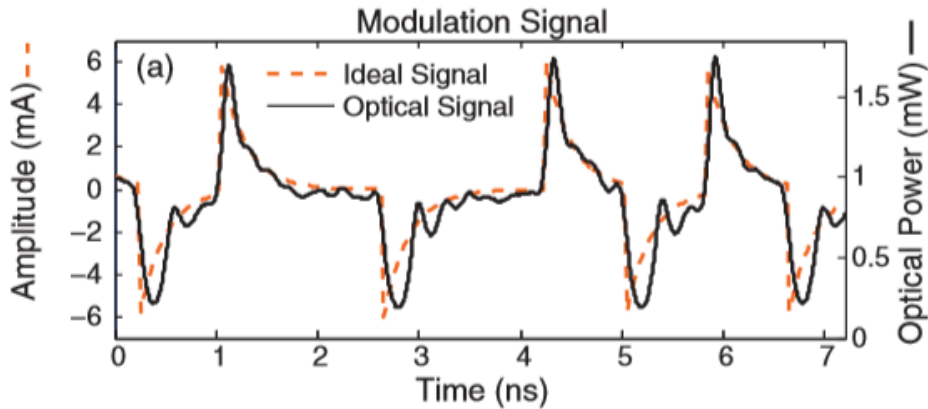


Figure 2.- Modulation Signal at 1.25 Gb/s [1]

The frequency response is directly affected by both terms in Equation 1. The first term is the transient chirp which shows strong peaks and a decay when the sharp transition occurs at the beginning of each symbol. The term that corresponds to the adiabatic chirp mainly follows the modulation signal. The simulated representation for the behaviour of both terms is depicted in Figure 3.

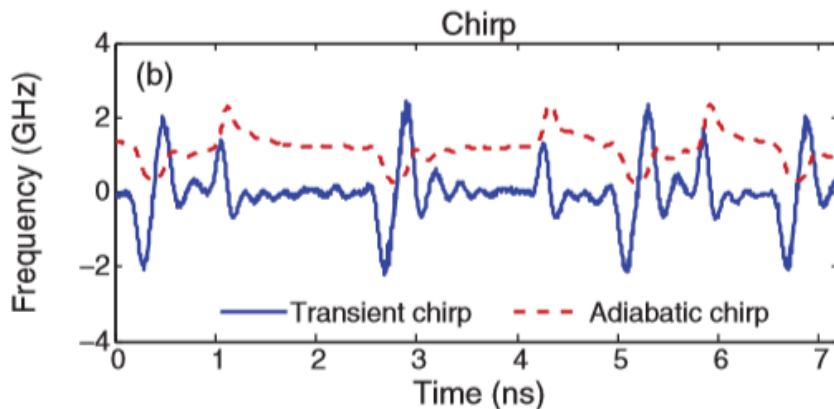


Figure 3.- Transient and Adiabatic Chirp [1]

Finally, Figure 4 displays the simulations for the optical phase variations, that are obtained by integrating the frequency chirp equation (Equation 1). The variations that correspond to the transient chirp describes sharp transitions at the start of each symbol followed by an exponential decay, similarly to the modulating signal, being noticeable that negative modulation pulses produce stronger and sharper variations. The phase change related to the adiabatic chirp, however, presents a slow slope while following the symbols. As both phase terms happen simultaneously, the total optical phase shows an asymmetrical behaviour.

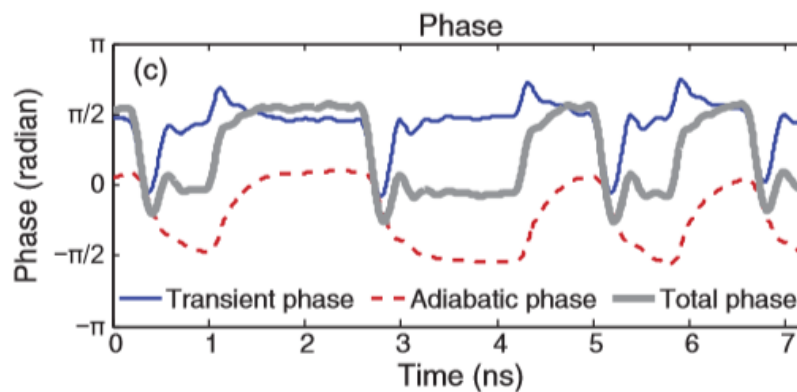


Figure 4.- Total, Transient and Adiabatic Optical Phase [1]

For the analysis of the optical phase eye diagram and the IQ diagram (Figure 5), the results validate that the symbols' transition does not quite match the unity circle due to the residual amplitude modulation. Hence, the modulating signal used is not optimal.

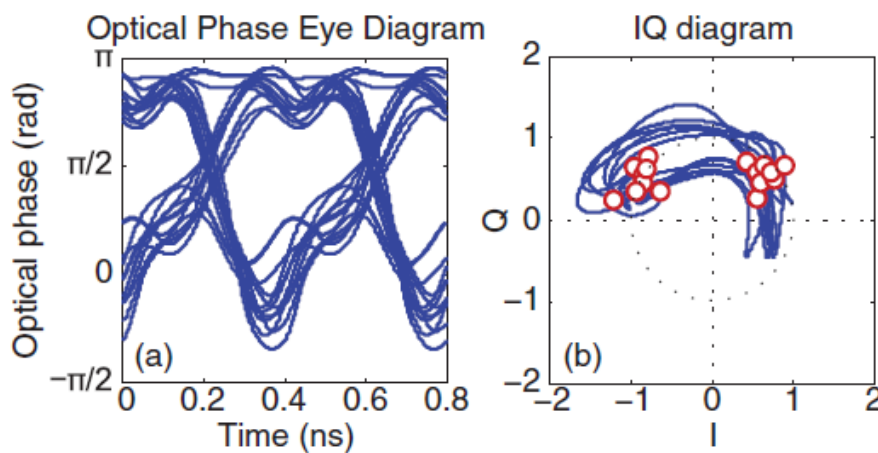


Figure 5.- Experimental results: Optical phase eye diagram and IQ diagram [1].

As explained before and given the results obtained by Altabás *et al.*, predicting the parameters for a modulation signal from an expected phase sequence can be optimized if a correct predictor is found. In data transmission, predictors are as simple as a system that includes a Channel State Information (CSI) properties. Usually the estimation is done at the receiver, quantized and fed back to the transmitter, the information that each element of the systems knows are called channel state information at the receiver or transmitter, CSIR and CSIT respectively.

A new prediction technology has taken the interest of many investigators and developers, that is Artificial Neural Networks (ANN), most commonly known as Neural Networks (NN). Its first implementation was a mechanical device that tried to simulate the working of the nervous system. The results from the machine were compared with the ones given by live nervous connections, this comparison proved that weak nervous connections may become strong if provided with more sensory points and outside excitors, a principle comparable with the learning process which is the foundation for NNs [3].

However, it wasn't until the 1940s decade, where the first neuron from McCulloch and Pitts was developed and allowed to define a neural network model as: a collection of McCulloch-Pitts neurons, all with the same time scale, where its outputs are connected to the inputs of other neurons [4]. Through the years, many experiments have proven the value of the technology by its capabilities of predicting figures, images and even handwritten symbols. Late studies have shown that artificial neural networks are useful even predicting time based sequences.

Based on this information, it is natural to inquire whether is feasible to train a NN, such that is able to predict the optimal modulating signal based on the knowledge and the expected result for the optical phase in Altabás, Izquierdo, Lázaro and Garces experiment from 2016.

Main idea is based on the fact that the phase of a VCSEL is, mathematically, the differential equation described previously. Thus, is possible to calculate a set of expected phases corresponding to a given batch of known signals. The calculated phases would act as training sets for the neural network in an inverse manner; that is: the expected prediction (phase) is the input signal and the training will provide a network that internally knows the relations between phase and current.

Once the network is trained, a phase whose input signal is known is given to the NN in order to evaluate its fidelity by comparing the network's output with the known signal. This validates the training and accuracy of the process. Neural networks managing VCSELs can reduce characterization time and estimate inputs for different expected signal parameters, giving aid by improving the modulation and reception process when using this type of devices. Additionally, these principles can be extrapolated into other optical devices.

2. State of the Art

Incoming section explains the framework in which this Master's Thesis is based. Optical systems became an interesting topic in the communications field since 1980 when it started revolutionizing the telecommunications' field and continues to do so until now. Not only is necessary to have certain knowledge of basic concepts and elements for the referenced topics, but also a background on the evolution of the technology and the necessities that lead to this progress.

As a main element in our discussion, it will be exposed the theoretical and practical analysis that led to the fabrication and commercialization of VCSELs (and others light emitting devices) since the first *solid-state* laser known as Ruby Laser, manufactured by Theodore Maiman in 1960 based on foundations laid by Albert Einstein. In addition, basic characteristics, advantages and drawbacks of the devices development will be described.

Moreover, a short review on *neural networks* as *deep machine learning* algorithms which allows the system to be trained with a number of pairs input-output type and later be able to predict different outcomes according to the given inputs. A system first modelled in 1943 that has come a long way, proven to be efficient in learning traces, images, shapes, letters and symbols.

A background and a comprehensive review of the literature is required. This is known as the Review of Literature and should include relevant recent research conducted on the subject matter.

2.1. Timeline for Optical Technology

The use of light for communication purposes dates back a long time ago if the term *optical communication* is used in a broad sense. A lot of civilizations have used mirrors, fire beacons and/or smoke signals to convey information. However, the first commercial use was done in late 1790, when an "optical telegraph" (also known as *semaphore telegraph*) was designed to be used between the cities of Paris and Lille (about 200 kilometers [km] of distance).

The system created by Claude Chappe pictured in Figure 6, expected to transmit coded messages over long distances (more than 100 km) by using *repeaters* or *regenerators*, known then as relay stations [5]. The role of light was to make the coded signals visible enough to be perceived by the next relay station. Even though it worked with very low effectivity (bit rate smaller than 1 bit per second), by 1830 the network had expanded throughout Europe.

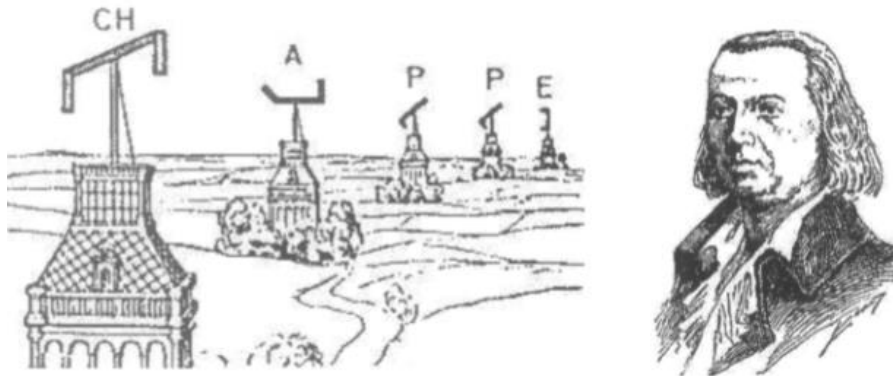


Figure 6.- Illustration of the optical Telegraph and its inventor Claude Chappe [6]

However, the birth of electricity was a game changer. It replaced light at the beginning of the electrical communications era. Later on, telegraphs, telephones and coaxial cables instead of twisted pair were followed by the first microwave system transmission with a carrier frequency of 4 GHz in 1948. During the next 25 years, microwaves and coaxial systems evolved considerably.

Around the 1950s, the idea of using optical waves instead of microwaves was expected to increase the system capacity considerably, but there wasn't either an optical source nor a transmission medium. The breakthrough came with the invention and demonstration of the laser in the 1960s.

Later, in 1966 it was suggested in that optical fibers might be the best choice, due to their capabilities of guiding the light in a comparable manner as the copper wires do with electrons. This insight and first ground-breaking measurements from Charles K. Kao, awarded the Nobel Prize in Physics in 2009 [7]. The principal issue was the high losses for former optical fibers (1000 dB/km). It wasn't until the 1970s where the fiber loss parameter was reduced to 20 dB/km in the wavelength region near one micrometer (μm) by using a new fabrication technique [6].

The enormous progress realized over the 25-year period extending from 1975 to 2000 can be grouped into several distinct generations (Figure 7). There is an evident increase in the bit rate by a factor of 100,000; while the transmission distances grew from 10 to 10,000 km, both in mentioned time period. As a result, the *bit rate-distance product* (BL , B is bit rate and L the distance) of modern lightwave systems can exceed by a factor of 10^7 compared with the first-generation lightwave systems [8]. Each generation manages to increase BL until it saturates as the technology matures, each one brings a fundamental change that helps to improve the system performance.

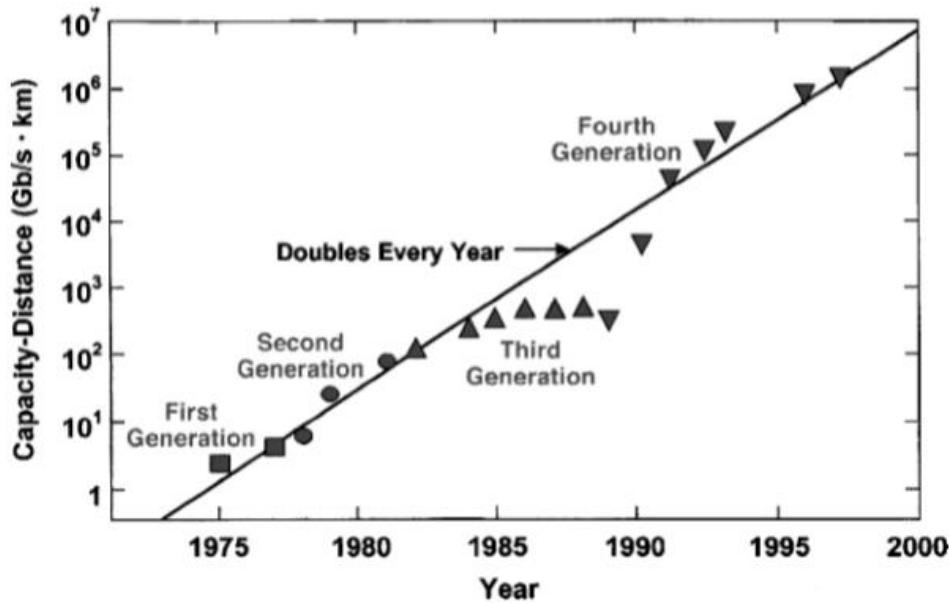


Figure 7.- BL product increment in different lightwave systems [8]

The first generation of lightwave systems operated near 0.8 μm and used gallium arsenide (GaAs) semiconductor lasers. After several field trials during the period 1977-79, such systems became available commercially in 1980 [6], obtaining up to 45 megabits per second (Mb/s) as bit rate with maximum distance between repeaters of 10 km. This distance improvement from the 1 km for coaxial systems was a big motivation for designers: it decreased both installation and maintenance costs associated with each repeater.

The second generation of fiber-optic communication systems was available in the early 1980s, but the bit rate was limited to below 100 Mb/s because of dispersion in multimode fibers. This limitation was overcome by the use of *single-mode* (SM) fibers, as an experiment in 1981 achieved to transmit at 2 Gb/s over 44 km of SM fiber [9]. By the end of the decade, second generation of lightwave systems was operating at 1.7 Gb/s with a repeater spacing of about 50 km and commercially available. The distance between the repeaters was limited due to the losses at the operating frequency of 1.3 μm (usually 0.5 dB/km).

The introduction of third-generation lightwave systems operating at 1.55 μm was considerably delayed by its large fiber dispersion near that wavelength. Conventional indium gallium arsenide phosphide (InGaAsP) semiconductor lasers could not be used because of pulse spreading occurring as a result of simultaneous oscillation of several longitudinal modes. This issue was thought to be overcome either by using dispersion-shifted fibers designed to have minimum dispersion near 1.55 μm or by limiting the laser spectrum to a single longitudinal mode. Both approaches were followed during the 1980s.

By 1985, laboratory experiments succeeded transmitting data at bit rates of up to 4 Gb/s over more than 100 km [10]. This generation became commercially available in 1990 with transmission up to 2.5 Gb/s, the best performance was achieved by using lasers oscillating in a single longitudinal mode with dispersion-shifted fibers.

A drawback of the third-generation 1.55 μm systems was the periodically signal regeneration process using electronic repeaters located 60-70 km apart from each other. This spacing was increased by including homodyne or heterodyne detection in the schemes which improved the receiver sensitivity. These systems are called as coherent lightwave systems, were developed around the globe in the 1980s with prove of the benefits in many experiments. However, they were commercially belated with the arrival of fiber amplifiers in 1989.

The fourth generation for lightwave systems took advantage of the newly known *optical amplification* by increasing the distance between repeaters and WDM technique to enhance the bit rate around the 1990s. Experiments in 1991 successfully transmitted at 2.5 Gb/s over 21,000 km and at 5 Gb/s over 14,300 km using a recirculating-loop scheme [6], proving that an amplifier-based, all-optical, submarine transmission system was feasible for intercontinental communication. By 1996, not only transmission over 11,300 km at a bit rate of 5 Gb/s had been demonstrated by using actual submarine cables [11], but commercial transatlantic and transpacific cable systems also became available.

Next big step was taken in November 1997, when the FLAG (Fiber-Optic Link Around the Globe) cable system was fully functional. It connected 12 countries from the U.K. to Japan with the FLAG Europe Asia (FEA) segment, pictured in Figure 8. This link was expected (and managed) to offer 10 Gb/s synchronous digital hierarchy communication using 27,000-km of an undersea fiber-optic network and EDFAs (*erbium-doped fiber amplifiers*) pumped at 1480 nanometer (nm) [12].

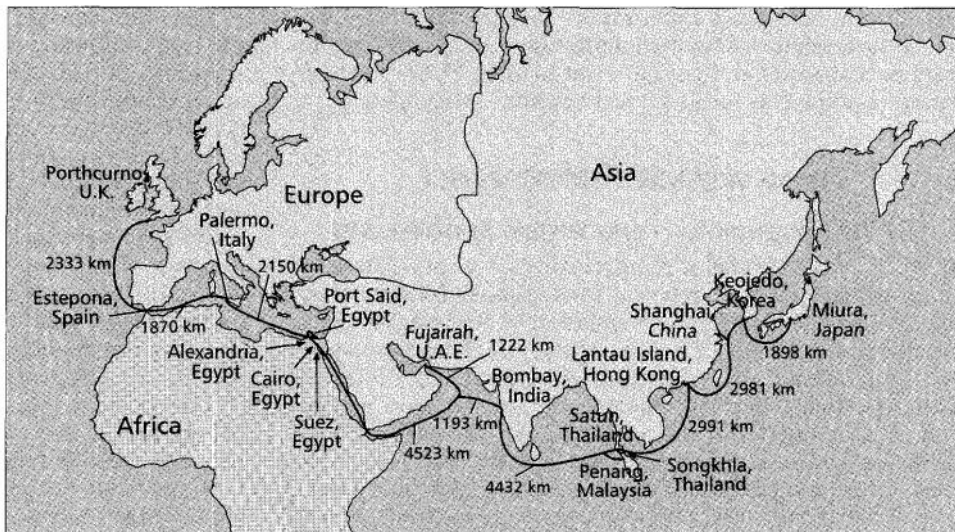


Figure 8.- First FLAG link, Europe-Asia (FEA) segment [12]

Since then, new links have been included in the submarine network, as shown in Figure 9. The total number of cables changes all the time due to replacement of older cables, by the beginning of 2018 approximately 448 submarine cables were already laid, which most are fiber-optic technology. It was calculated that there are over 1.2 million kilometers of submarine cables in service globally [13].

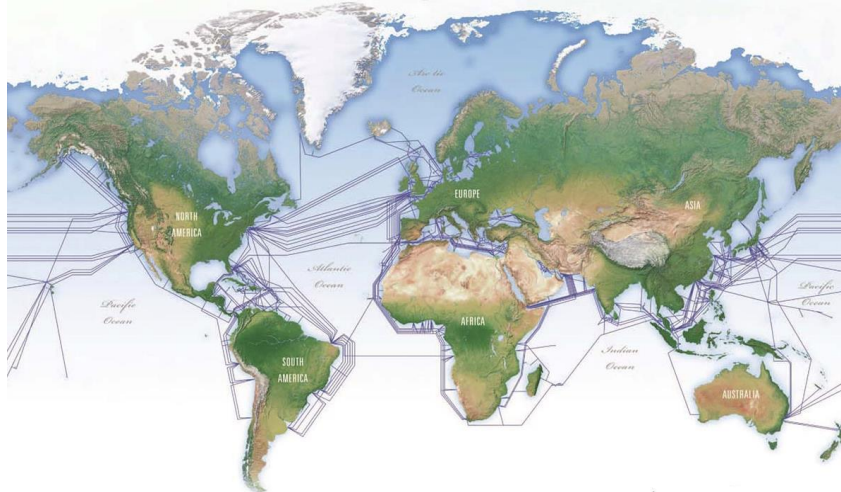


Figure 9.- Submarine Optic Fibers Nowadays [14]

Seen below, Figure 10 shows the increase of capacity for lightwave systems through several generations of development. Additionally, the rapid growth after 1992 due to the birth of WDM technology. This advent allowed to reach an operating bit rate of 10 terabits per second (Tb/s) by 2001 [6]. The more channels were included in the transmission schemes; it was often impossible to amplify all the channels using a single amplifier. As a result, new amplification schemes (such as *Distributed Raman amplification*) were developed to cover the spectral region from 1.45 to 1.62 μm [8].

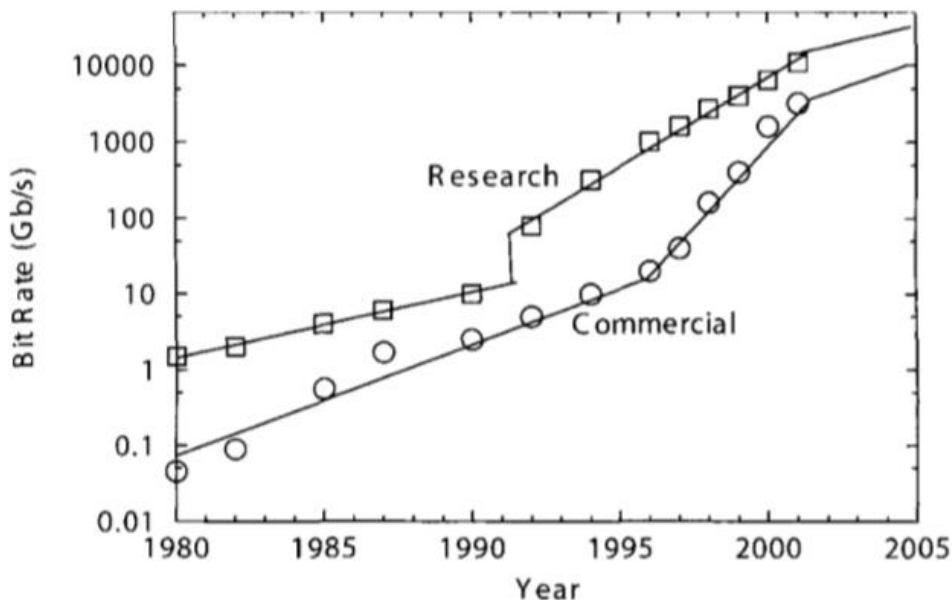


Figure 10.- Increase of capacity in Lightwave Systems after 1980 [6]

The new amplification schemes led to more experiments: in 2000, 82 channels were transmitted at 40 Gb/s over 3,000 km (which sums up to 3.28 Tb/s) and months later, the system capacity was able to increase to almost 11 Tb/s by using 279 WDM channels, but lowering the transmission distance to 117 km. Commercially, the first use of Raman

amplification was by 2003, when transmitting 80 channels, at 40 Gb/s each for a total of 3.2 Tb/s.

Finally, the fifth generation of fiber-optic communication systems is focusing on extending the wavelength range for simultaneous WDM operation. Until now, this technology works in the conventional C band (1.53-1.57 μm) but the extension (on both sides of the wavelength range) covers both L and S bands, both in which Raman amplification technique can be used. Moreover, a new kind of fiber known as *dry fiber* was developed with a unique property that its losses are small over the entire wavelength region from 1.30 to 1.65 μm [8]. It is expected that these new fibers and the new amplification schemes allowed to design lightwave systems using thousands of WDM channels.

The focus for this generation is mostly increasing the spectral efficiency for WDM systems. The next step was to include advanced modulation formats, even when typically, were formats developed for microwave technology. The idea was to code the information by using both amplitude and phase of the optical carrier [15], which allowed to increase the spectral efficiency, typically below 0.8 bits per second per Hertz (b/s/Hz) for the fourth-generation systems to more than 8 b/s/Hz.

The results were validated in a 2010 experiment [16], a new record was established when 64-Tb/s transmission was realized over 320 km using 640 WDM channels, spanning both the C and L bands with 12.5 GHz channel spacing. Each channel contained two polarization-multiplexed 107-Gb/s signals coded with a modulation format known as quadrature amplitude modulation (QAM).

At this point, optical communications have grown continuously since its beginning, specially taking into account that it was increased its capacity in a factor bigger than 70,000 over 25 years, knowing that the first generation (1980) achieved only 45 Mb/s. The expansion is not only driven by the velocity it can accomplish in the communication field, but also for its use in many other areas, such as: medicine, manufacture of scientific objects, photography, optical disk storage and more important inventions that impact our lives.

2.2. From LASERs to VCSELs

Since it was known that optic waves could be used for data transmission, a great amount of scientists, laboratories and big companies have tried to develop optical technology. The optical transmitter role was (and remains to be) to convert an electrical input into a corresponding optical signal that can be sent out through a channel; in this case, and optical fiber. It is composed by several elements: (i) optical source, (ii) optical modulator, (iii) electrical pulse generator, and (iv) an electronic circuit used to power and operate the device, an example is shown in Figure 11 [6].

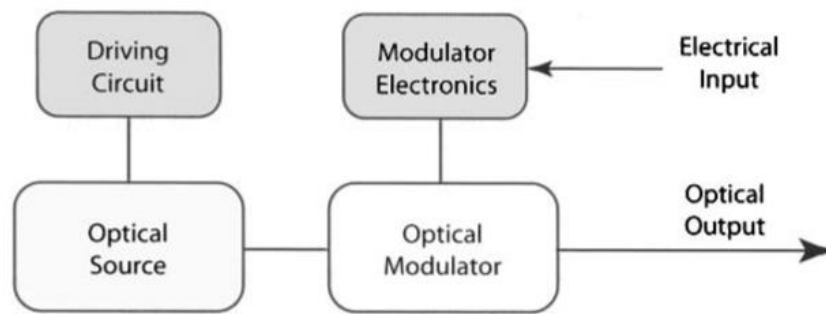


Figure 11.-Block diagram of an optical transmitter [6]

The main element (and our interest) is the optical source. Lightwave systems use mainly two types of semiconductor optical sources: Light-Emitting Diodes (LEDs) and semiconductor LASERs (Light Amplification by Stimulated Emission of Radiation) due to their several advantages, such as: compact size, reliability, high efficiency, small emissive area and more [8].

In order to understand the characteristics and working principles of both devices, is necessary to recall how the light emission process occurs in a semiconductor. Naturally, all materials are known to absorb light rather than emit it under normal conditions; the emission process is only possible for some elements, under specific conditions and according to the medium's physical properties. Every material is composed by positive and negative charges (also known as holes and electrons, respectively) and different energy states.

There are three fundamental processes that can occur when an electron moves between two energy levels, that is: (a) absorption, (b) spontaneous emission, and (c) stimulated emission (Figure 12). As mentioned before, *absorption* is the most common one in all materials, in this case, a photon's energy is taken by the electrons of the material, therefore the electron will gain energy and move to a higher energy state.

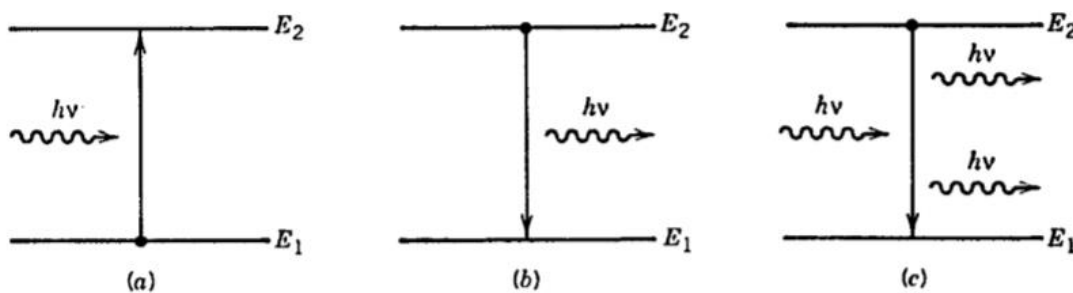


Figure 12.- The Emission-Absorption Principle [8]

For the emission process, it can occur naturally or provoked. Both process take place at any optical semiconductor source, though each one dominating as main process in LEDs and lasers, respectively. *Spontaneous emission* happens when an excited electron releases energy (lowering to ground state energy level) in the form of a photon with random frequency, phase and direction. In the case of *stimulated emission*, an existing photon forces an excited

electron to release its energy as a new photon with exactly the same frequency, phase and direction of the original one.

All lasers, including semiconductor lasers, emit light through the process of stimulated emission and are said to emit coherent light. In contrast, LEDs emit light through the incoherent process of spontaneous emission (Figure 13).

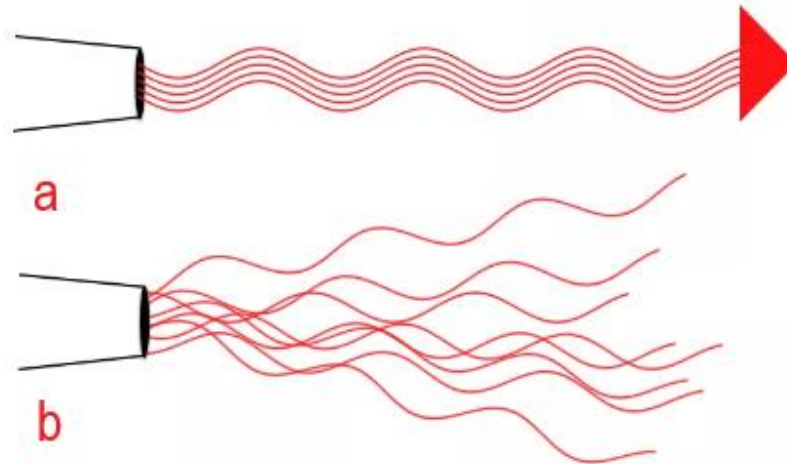


Figure 13.- Coherent Laser Light (a) and Incoherent LED Light (b) [17]

The viability of the stimulated emission in semiconductors was first considered in 1917 by Albert Einstein who described the theory between the energy states. Later on, engineers began to use this principle for practical use at the end of the 1940s, while trying to improve a radar-related vacuum tubes. Gas atoms enclosed in a special cavity were stimulated with an electromagnetic beam expecting to release energy exactly as the beam waves. The energy that exited the cavity was indeed a much more powerful beam, hence the device was an amplifier. It was called MASER, from Microwave Amplification by the Stimulated Emission of Radiation [18].

Its inventor, Charles Townes and others also believed that by using higher frequency energy they could create an optical MASER, but it was Theodore Maiman in 1960 who achieved it. The first working laser (named ruby laser) used synthetic ruby crystals as its gain medium and accomplished to produce pulses of coherent visible light at 694.3 nm [19]. The ruby rod was placed between mirrors that formed an optical cavity, when oscillating the light produced by the ruby fluorescence caused the stimulated emission. The device had to be pumped with very high energy.

Initially, lasers were conceived by using a Fabry-Perot resonator with no transverse dependencies and reflecting end facets that would bounce the light onto a single output window; a property called *feedback*, and it can be defined as: the ability of the laser to re-use back-reflected photons emitted by the active layer to exponentially increase the laser's stimulated emission, until the saturation is reached [20].

In 1962, several research groups realised that semiconductors were capable of emitting light. In fact, a single semiconductor diode can act both as a LED or a LASER; the breakpoint is called the threshold current: once the injected current goes above the margin, stimulated

emission is induced and coherent light will be emitted (laser mode), on the contrary if the input current remains below the threshold the device will emit spontaneous emission as a LED, as can be seen in Figure 14 [21].

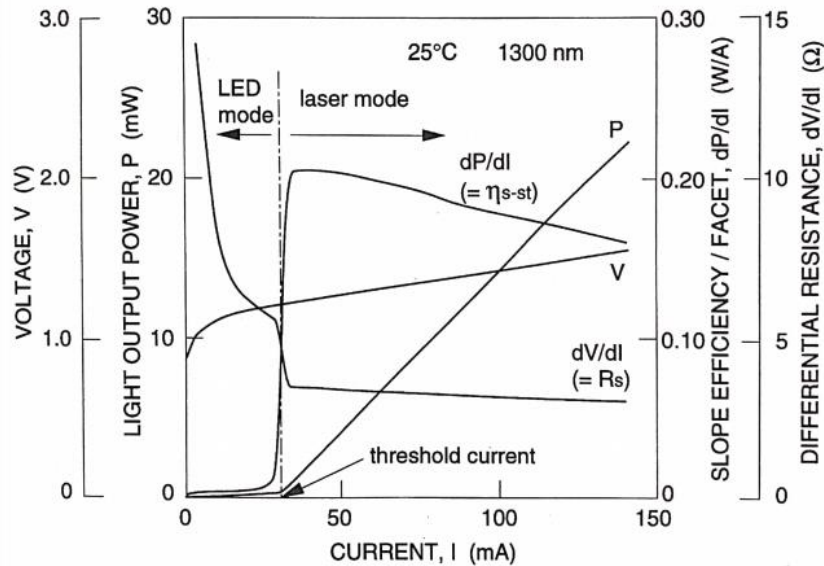


Figure 14.- Characteristics for a 1300nm-band InGaAsP/InP Semiconductor Diode [21]

Semiconductors operating in LED mode differ from LASERs in some parameters, which define their applications possibilities. On the positive side, LEDs are less expensive to fabricate and last longer. Major drawbacks are their lower output power and wider optical spectrum. LED technology has been widely used for several equipment and devices, especially in the electronics field, they are available across the visible, ultraviolet and infrared wavelengths, which to the human eye is perceived as different light colours.

However, for optical communications systems, LEDs can be used only for systems with low data rate requirements (around 100 Mb/s) or in local-area networks, where a coherent optical source is not required. Thus, the preferred optical source for telecommunications is the LASER, as a result of its monochromatic, coherent and highly directional radiation is feasible to be used in long-distance optical-fiber communications [22].

The first created light emitting device was a GaAs (Gallium Arsenide) p-n junction whose facets were polished to form a Fabry-Perot cavity (that means, both facets acted like mirrors causing the laser's feedback). The operation wavelength was of 0.88 μm. Soon enough, others semiconductor materials were used for p-n junctions that would work at other different wavelengths, such as: GaAsP, GaInAs, InAs and InP (Gallium Arsenide Phosphide, Gallium Indium Arsenide, Indium Arsenide and Indium Phosphide, respectively). However, their practical use was limited due to their high threshold current density, which prevent them to operate at room temperature [22].

By 1963 it was suggested that semiconductor lasers would improve considerably if designed with a heterostructure in which the semiconductor active layer is sandwiched between two cladding layers of a second semiconductor with a wider band gap, as seen in Figure 15. However, the concept couldn't be demonstrated until 1969 when LPE (liquid-phase

epitaxy) technique was perfected and used to successfully operate a semiconductor laser based on heterojunction at room temperature [22].

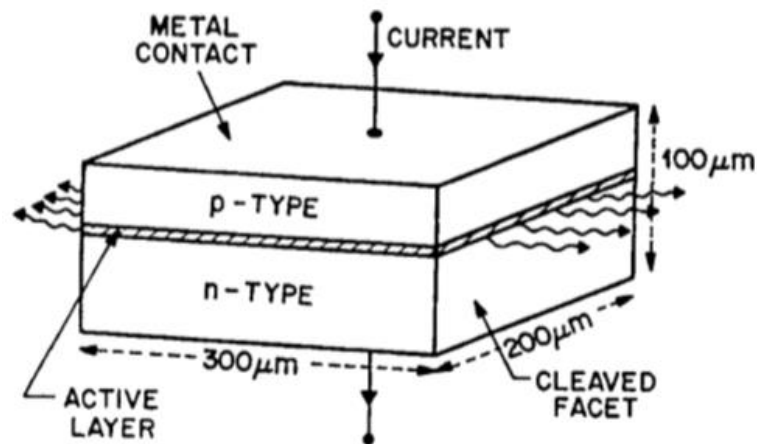


Figure 15.- Schematic of a Semiconductor Laser [8]

Around 1970, the high-performance AlGaAs/GaAs double-heterostructure laser became commercially available [23]. Since then, semiconductor lasers have been deployed extensively and used almost exclusively for fiber-optic communications.

In 1971, the idea of a new paradigm arrived by the studies of Kogelnik and Shank, who create a new device using dyed gelatin on a glass substrate. The intention was to achieve the laser's feedback with an internal grating structure distributed along the optical cavity [22]. The grating caused the feedback index to vary (as a result of the Bragg diffraction phenomenon) which translates to an amplification of the incident light beam. The use of this technique gave name to a new type of devices called Distributed-Feedback Lasers (DFB), which continued under development until their commercial availability by 1989.

The first semiconductor DFB laser oscillation was achieved by optical pumping into a corrugated GaAs surface, reported by Nakamura and others in 1973. Later, Nakamura accomplished the first continuous-wave (cw) oscillation at room temperature in 1975. Finally, room temperature cw operation at 1.5 μm in InGaAsP/InP DFB laser was obtained in 1981 by Utaka and fellow scientists [22].

Simultaneously, another device with gratings outside the active layer was being studied. For this device, the end regions will act as mirrors whose reflectivity is maximum to obtain a wavelength as close as possible as the laser's basic longitudinal mode, this special type of DFB laser is called Distributed Bragg Reflector (DBR) [8].

The first DBR laser oscillator was demonstrated by Reinhart, using GaAlAs/GaAs in 1975. As the grating for a DBR laser is fabricated without touching the active layer, this device is suitable to be integrated with outer optical devices, which can result as a multifunctional light source, as multielectrode wavelength-tunable lasers [22].

Figure 16 shows the structure of the DFB laser; it is worth mentioning that the refractive index in the grating (waveguide area) is higher than the one of InP (semiconductor), but lower than the refractive index of the active layer (InGaAsP material). In addition, antireflection

coatings (AR) are applied to both sides, in order to suppress any reflections from the end of the grating.

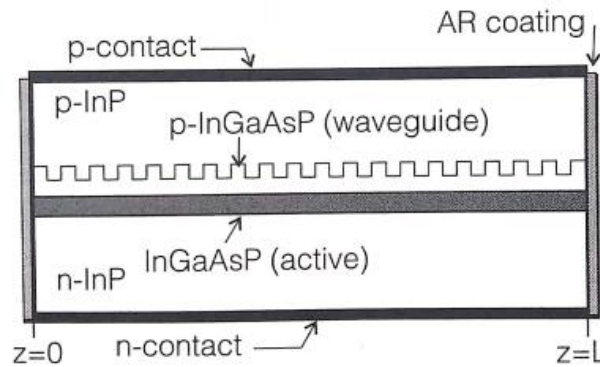


Figure 16.- Outline of a DFB laser structure [23]

DFBs are widely manufactured, despite their complexity to obtain the grating, but are reliable enough and they are being used since 1992 in all transoceanic lightwave systems. Nowadays, are used in nearly all 1.55- μm optical communications systems operating at 2.5 Gb/s or more.

After the successful development of SM laser diodes, around 1300 nm and 1550 nm in the early 1980s and the birth of WDM systems, it was required to have single-mode, narrow-linewidth lasers, capable of remaining at a fixed frequency over time. DFBs filled those requirements, but their stability compromised the tunability of the device. The increasing number of DFB lasers used in a WDM transmitter made the designs and its maintenance expensive and impractical. Therefore, it was necessary to create a laser that could be tuned [23].

By 1990s, multisection DFB and DBR were developed to meet both stability and tunability. The new design consisted in three sections, named: active, phase-control and Bragg sections, shown in Figure 17. Each one could be biased independently by injecting different currents in the Bragg and phase-control segments. By 1997, the tuning range had achieved 17 nm and output powers of 100 mW with high reliability [8].

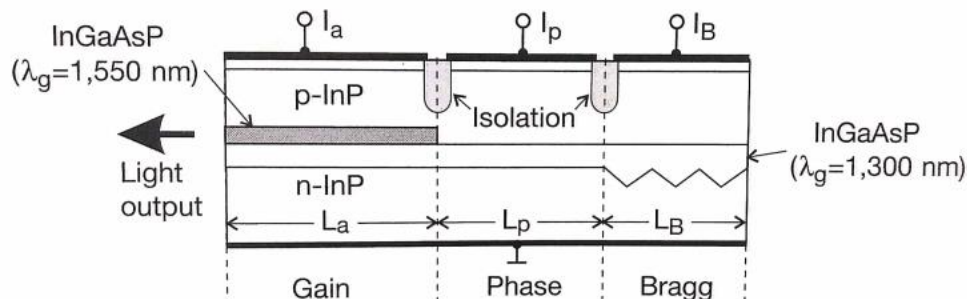


Figure 17.- Schematic longitudinal cross-section of three-section DBR laser [23]

The demonstrated performance of diode lasers suggested that they could be used for much more applications beyond fiber optics, compact optical discs, and optical recording.

These new possibilities caused an increased interested in *surface-emitting* geometries for semiconductors, in order to find the best device configuration for a given application [24]. In this new layout, the light beam would radiate perpendicularly to the junction plane, instead of the parallel way used by edge emitting lasers, as shown in Figure 18.

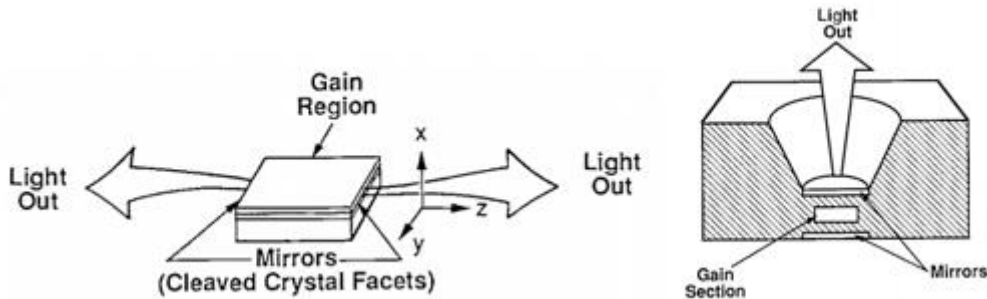


Figure 18.- Lasing direction for edge and surface emitting devices [24]

There are three basic types of surface-emitting lasers, (i) Vertical-Cavity Surface-emitting Laser (also known as VCSEL), (ii) Grating-outcoupled surface-emitting Laser (GSE laser), and (iii) Integrated Beam Deflector Surface-emitting Laser (or *folded cavity* laser).

The first VCSEL was pioneered in 1979 by Soda, Iga and co-workers. The designed is shown in Figure 19 [22]. It was a double heterostructure with an InGaAsP active layer emitting at 1.3 μm using metallic mirrors with high absorption rate and reflectivity.

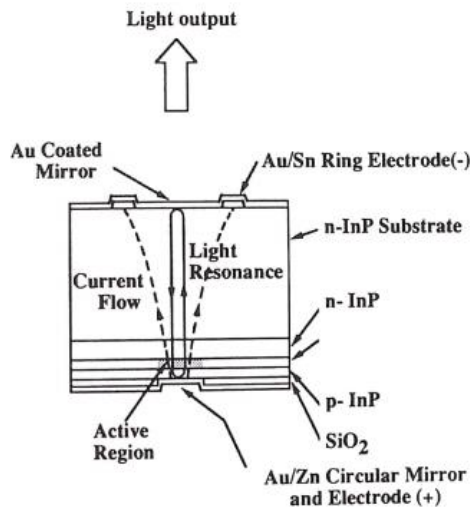


Figure 19.- Metallic Mirror VCSEL Structure [22]

Due to the short cavity length (around 1 μm), the width of the longitudinal-mode range is elevated compared with the gain curve, which means that if the centred (resonant) wavelength is closed to the gain peak, the single-longitudinal mode is assured. However, if the active layer diameter is large, multi-transverse mode operation can occur [23].

Mathematically, cavity length is inversely proportional to the threshold gain, hence, having such a small cavity increases the threshold current. This is solved by increasing the facets (or mirrors) reflectivity, by using metal coatings or distributed Bragg reflectors [21]. An

idea proposed by Burnham, Scifres and Streifer based on an epitaxial growth structure in 1982, however this technique was not available until after 1983.

Later on, using the epitaxial technology advances, specially control of thickness and epitaxial layers' composition, a new semiconductor Bragg reflector (with high reflectivity) was produced. Which lead, on 1988 to the first demonstration of a continuous-wave, room-temperature, low threshold current VCSEL [22].

Figure 20 presents the schematic of that type of VCSEL, which consisted in an active region enclosed between n and p type doped semiconductor Bragg reflectors. The DBR-mirrors consist of alternating layers of high and low refractive indices materials. Each layer is $\lambda/4$ thick, where λ is the wavelength the VCSEL emits, a technique that rises reflectivity above 99% [23].

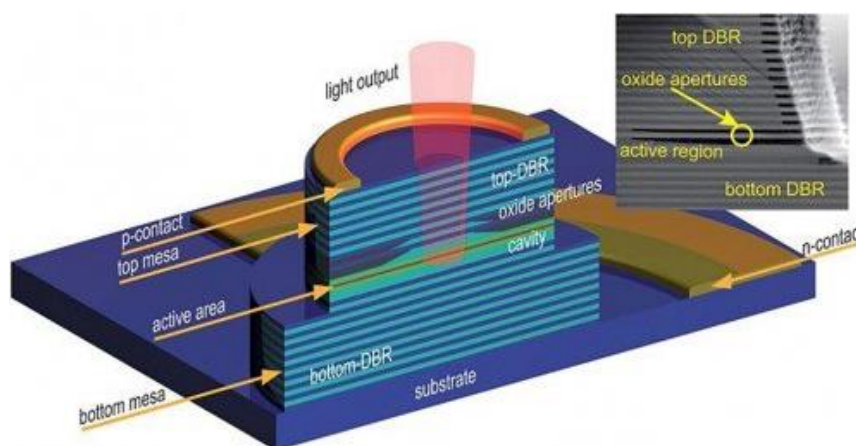


Figure 20.- Epitaxially Bragg Reflectors VCSEL structure [25]

The VCSEL structure provides a considerable number of advantages, as explained by Iga [26]:

- Wavelength and thresholds are relatively insensitive against temperature variation.
- Dynamic SM operation is possible.
- Is capable of high-speed modulations, as a result of its large relaxation frequency.
- Long device lifetime.
- High-power conversion efficiency, above 50%.
- Easy coupling to optical fibers thanks to the small spot size.
- Manufacturing uses full monolithic processes yielding to very low cost production.
- Bonding and mounting are relatively easy.
- Modules and package costs are cheap, which turns them useful for data transfer and local-loop applications.
- Well suited for WDM applications for two reasons:
 - ✓ Is possible to create two-dimensional VCSELs arrays where each laser emits at a different wavelength.
 - ✓ One VCSEL's wavelength can be tuned over a 50 nm range by using MEMS technology (micro-electro-mechanical system)

Additionally, VCSELs are capable of being versatile, proved by the possibility of fabricating a tuned device. The simplest way is to use the temperature dependence in the refractive index of the Bragg reflector. Thermal tuning can be achieved by resistive heating (adding an extra contact) or by means of the Peltier effect in combination with resistive heating [23]. However, thermal tuning produces a considerable increase in temperature and electronic VCSEL tuning is not possible due to their geometry.

An alternative is to tune directly the cavity length, by means of an external mirror close to the VCSEL surface, in a manner that the distance between the laser and the mirror can be adjusted. Figure 21 describes a design in which the length of the optical cavity is changed by applying a voltage in the tuning contact which moves the cantilever where a top mirror is placed. This device was able to achieve a tuning range of 19.1 nm, with a threshold current of 0.5 milliamperes (mA) and a power of 0.9 mW over 7 nm of the tuning range [23].

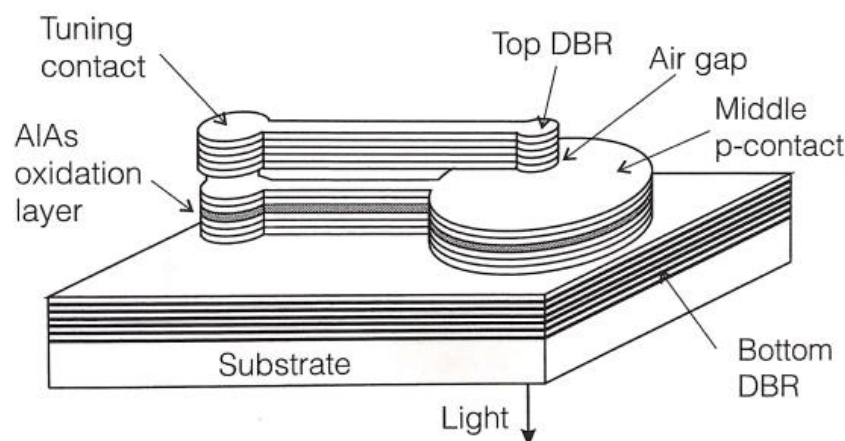


Figure 21.- VCSEL with DBR mirror on a cantilever [23]

Thanks to the VCSELs smaller footprint and the other numerous advantages, VCSELs are expected to make an important impact on future lightwave communications. One of the driving reasons, due to the potential much more reduced price than previous DFB or DBR lasers. Despite of the numerous advantages that VCSELs can provide and are expected to provide under further development, the physics involved at the fabrication and operation process of VCSELs, makes its performance to be less attractive than previous, and more mature, DFB and DBR lasers.

Nevertheless, current artificial intelligence solutions, as Neural Networks and Deep Machine Learning, are nowadays also more mature than in previous years. The main objective of this project is to develop this kind of AI tools, on one hand, to correct the imperfections of the performance of currently available VCSELs. On the other hand, to explore new frontiers for applications of lasers in general, and VCSELs in particular, to lightwave communications. Therefore, the new tools of Neural Networks as Deep Machine Learning are introduced in next section.

2.3. Neural Networks as Deep Machine Learning

As the science and the technology advances, one of the biggest challenges humanity faces is to be able to construct intelligent systems. By system, it is understood any physical or logical device with the ability to perform a required task. This is the main objective for the scientific field known as *Artificial Intelligence* (AI) [4].

Artificial intelligence includes two main categories: symbolic or non-symbolic approaches. Symbolic artificial intelligence is the construction of systems including selected characteristics that can be defined as “intelligent”. For example: a problem is selected and a system is designed to follow schemes that lead to the solution of that problem. It follows a *top-down* scheme, where an approximation to the solution is necessary.

On the other hand, non-symbolic artificial intelligence doesn’t need high-level problem solving design. Instead, the starting point is a generic system capable to adapt and construct itself as much as necessary until developing a system that is able to solve the problem using a *bottom-up* scheme [4]. The scheme is defined as such, as all systems are designed simple and identical, but they collect the main characteristics of the systems they try to imitate and automatically generate complex calculations, through predefined learning mechanisms. This class of AI includes the techniques called as *machine learning*.

Arthur Samuel, a pioneer in AI, defined machine learning as “the field of study that gives computers the ability to learn without being explicitly programmed”. The interest in *machine learning* has exploded over the past decade. It is seen in computer science programs, industry conferences, and even Wall Street Journal commonly. Fundamentally, machine learning uses algorithms to extract information from raw data and represent it in some form of model. The generated model (sometimes called *structural description*), is then used to infer things about other data that has not been processed [27].

Structural descriptions can take many forms, including: (i) linear regression, (ii) decision trees, and (iii) neural network weights. Each model has its own way for applying rules to the known data in order to predict the unknown data.

Linear regressions are perhaps the most common and understood algorithms for machine learning, in fact the technique was actually developed for the statistics field. The objective is to create a set of parameters to represent the input data; that is, to create a relationship between the input and output variables of a system, the most common model is $y = Ax + B$, where A and B are coefficients, and y is the output value for a given input x . There are various techniques to study this model, such as: Simple Linear Regression, Ordinary Least Squares and Gradient Descent [28].

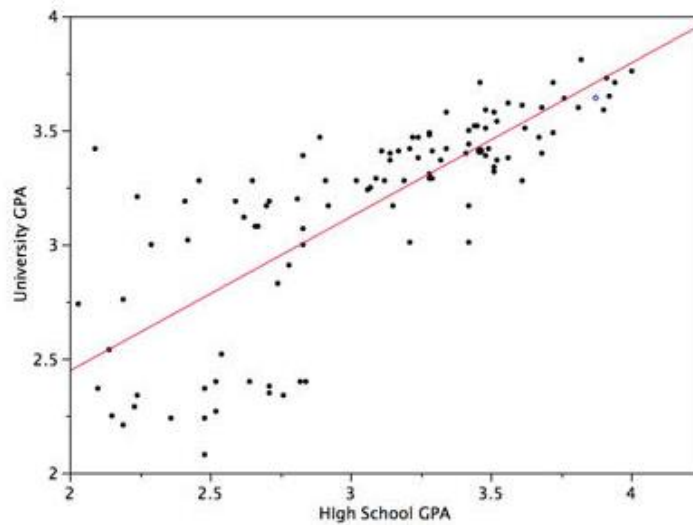


Figure 22.- Example of Linear Regression [29]

Figure 22 shows a basic example, where a set of plots define the relationship between High School and University *grade-point-averages* (GPAs) for a group of students. The dots represent the relation between both variables for a single student, it seems as the data almost follows an increasing slope. Using mathematical formulas, a linear function is deduced and it will be the linear regression that approximates to the actual values, it can be used even to predict new information.

A second type of model is the decision tree: a very simple predictor, $h : X \rightarrow Y$. It creates a set of rules in the form of a tree structure. Traveling from a root-node of the tree to a leaf, there is a label associated to an instance x . At each node on the root-to-leaf path, the input space (a question) is splitted into the successor path (the possible answers). Every leaf contains a specific label.

A basic example is given by Figure 23, a test to check if a papaya is tasty or not. The system first considers the colour of the fruit. If it is not pale green to pale yellow (a label), then the papaya is for sure not tasty. Otherwise, the papaya softness is evaluated as a second condition. This time, if the papaya gives slightly to palm pressure, the decision tree predicts that the papaya is tasty; if not, the prediction is “not-tasty” [30].

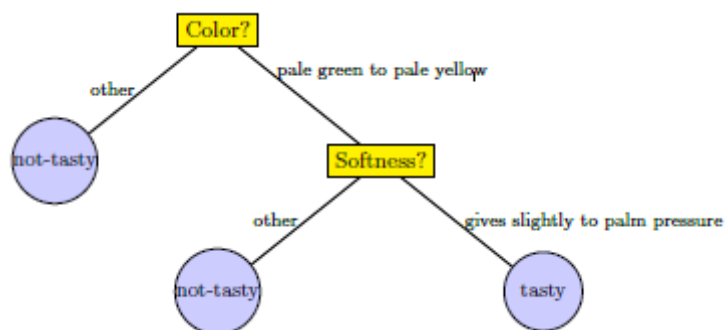


Figure 23.- A basic decision tree for the papaya example [30]

And finally neural networks weights, inspired by *biological neural networks*; that is, the brain and nervous system. The method uses a *parameter vector* that represents the weights on the connections between the nodes in the network [27]. Basically, an artificial neural network is a computational model consisting of a large numbers of basic computing devices (neurons) that are connected to each other in a complex communication network. However, the complexity of the system has evolved since the mid-20th century where simpler NN existed.

The first implementation for an Artificial Neural Network (ANN) was a hydraulic device described by Russell in 1913 [4]. However, it wasn't until the 40s decade where the field's study started to increase until today. McCulloch and Pitts (in 1943) realized the first mathematical model for an ANN, in fact, is considered the first neuron even though it was called a cell. The model presents the idea of a function crossing a threshold and it was used by many other ANN models.

Figure 24 shows a general scheme for a McCulloch-Pitts cell, it models a structured and a simplified brain neuron functioning with only two possible states: OFF (0) and ON (1). The cell will activate if the sum of the input values (x_i) multiplied by the weights of every connection (w_i) is higher than the threshold (θ). Otherwise, it will remain turned off.

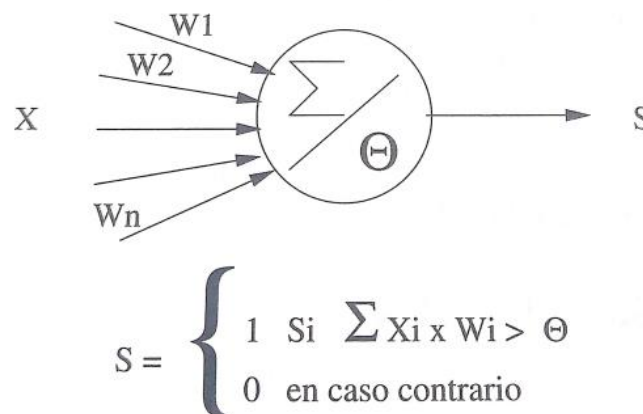


Figure 24.- General Scheme of a McCulloch-Pitts Cell [4]

Later on, in 1949 Donald Hebb developed a mathematical learning process based on his studies of neurons and learning conditions. It gave birth to a new paradigm named after him, the Hebbian theory. However, the first practical results were obtained by Minsky and Edmons in 1951, thanks to a designed machine with 40 neurons whose connections continuously adjust themselves according to different outcomes. The machine was constructed with tubes, engines and relays that simulated a rat searching food in a labyrinth [4].

Frank Rosenblatt achieved to generalize the McCulloch-Pitts cell model by giving it learning capabilities in 1957, the model was called the *Perceptron*. Initially, the design had two levels that adjusted the weights between the input and output connections, proportionally to the error between the expected and the obtained outcome. The model was conceived as a system capable of performing classification tasks automatically.

The idea of the learning process, was to have an example set of different classes as starting point, from which the system would determine the equations that made a distinction among those classes. The pre-existing classes are from now on called *patterns* or *training sets*. A new-three-level Perceptron case was studied, but no solid mathematical method was found to train the hidden connections layer.

A couple of years later, Bernard Widrow designed a new NN called *Adaptive Linear Element*, also known as ADALINE. It was very similar than the previous model, but the applications for each were very different. By 1960, this model was used to prove mathematically that in some cases is possible to minimize the error as much as desired [4].

By the 1960s, the upcoming designs were already closed to the possibilities known today. That year, Steinbuch was the first investigator to develop a *NN information coding method*, used to recognize distorted handwritten characters, machinery failure recognition and control several production lines.

In 1968, James Anderson, created a model for an *Associative Linear Memory* based on synapses and Hebb's theory. To obtain a significant performance, the model required independent linear vectors. The system included a new error correction method and replaced the linear threshold function for another one in ramp, the result was a new model called *Brain-state-in-a-box* (BSB).

Anderson's studies supported new discoveries: (i) Teuvo Kohonen managed to design a system that optimized linearly dependent vectors. While, (ii) Longuet and Higgins developed a system of coded equations that were able to store and recover signals sequences. A direct result is the creation of the *Holophone* in 1970, a device capable of storing input sequences and obtaining full signals from a reduced part of it [4].

New schematics were studied in 1975, when Kuniyiko Fukushima analyzed space and space-time models for vision systems. His first model, the *Cognitron* had an upgraded version in 1980, known as *Neocognitron*.

It wasn't until 1984, when Terence Sejnowski and Geoff Hinton developed the algorithm for Boltzmann's machine and its extension, the first ANN to recognize a learning algorithm for a three-level system, giving birth to *Deep Neural Networks* [27]. Also called *Deep Machine Learning*, are systems categorized as two-or-more-level systems, which caused the analysis, testing and design of the *Backpropagation algorithm*, usually used in voice recognition applications.

There are several models of ANN; the classification relays on a combination of different parameters: (i) the artificial neuron, (ii) the network's basic structure, and (iii) learning technique. Each one is an essential element of the networks, with its own functions and outcomes.

2.3.1. The Artificial Neuron

Based on the *biological neural cell*, the *artificial neuron* (or *node*) is the most basic unit of the network and are also stimulated by inputs. They pass some information they receive to

other nodes (often with transformations). As the brain cells, artificial neurons can be trained to pass forward only data that is useful for achieving a “larger goal” [27].

The overall input for each cell will be the sum of all the inputs multiplied by the weights of their connections, that is $E = x_1w_1 + x_2w_2 + \dots + x_nw_n$. Additionally, is called S to the set of all possible neurons' state, which can be either: binary, $S = \{0, 1\}$, a multiple set as $S = \{0, 1, 2, \dots, n\}$ or a continuous interval, $S = [0, 1]$. The value given by the cell is defined by the *activation* or *transition function*, when the total input is evaluated. According to it, there are different kinds of activation functions:

- Linear: $S = KE$, where K is a constant.
- Threshold: $S = 1$, if $E \geq \theta$; $S = 0$, if $E < \theta$; where θ is the threshold value.
- Any function: $S = F(E)$

2.3.2. Network's Basic Structure

For simplicity, it is assumed that the networks are organized in layers. Meaning, that there are a set of nodes that can be decomposed into a union of (nonempty) disjoint subsets, such that every node in $V(t - 1)$ connects to some node in $V(t)$, for some $t \in [T]$. Important elements on a NN structure are [30]:

- Bottom - or input layer: it contains the n neurons. Here n is the dimensionality of the input space.
- Depth of the network: is the number of layers in the network, excluding the input layer (V_0). The NN's depth of the depicted example in Figure 25 is 3.
- Hidden layers: includes from V_1 to V_{T-1} , that is all layers in the network except for the input and output ones. The example below has 2 hidden layers.
- The top layer is also known as output layer. In simple prediction problems this layer contains a single neuron whose outcome is the network's output.

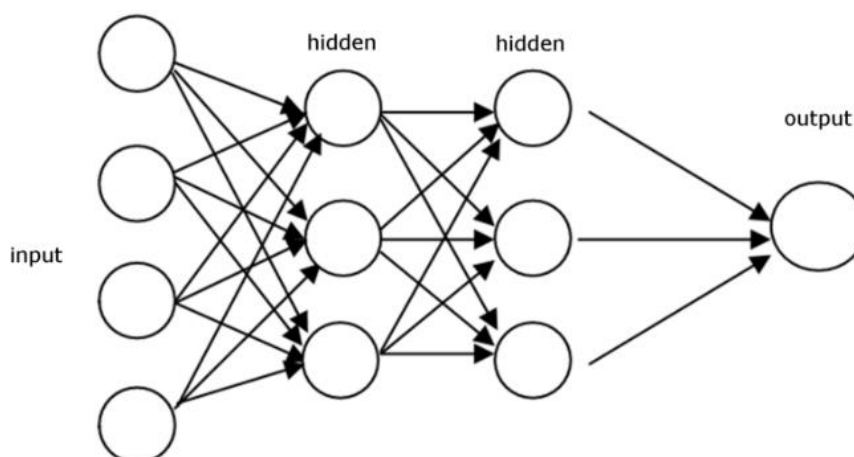


Figure 25.- Three-Layered feed-forward fully-connected NN schematic [31]

2.3.3. Neural Network Learning Techniques

The most important part in a NN, as the selected scheme for each network is the key definer on the kind of problems the system is capable of resolving. As Artificial Neural Networks are learning machines based on examples, the training set must fulfil two characteristics [4]:

- The amount of examples within a data set has to be enough. If the number of patterns is reduced, the networks won't be able to adapt the weights efficiently.
- The training set elements must be diverse. If a set has a larger number of examples from one class than other, the network will "specialize" in the first subset data. It is important that all the different categories are evenly represented in the whole learning set.

Generally, the training/learning process consists in receiving all the patterns while modifying the connections weights following a selected learning scheme, this modification can be done after each element is processed or once they are all inside the system. Once all the examples are computed, is time to validate if convergence criteria exist. Otherwise, the process is repeated by processing the whole data set one more time.

The convergence criteria depend on the type of network used or the kind of problem studied. The convergence criteria mean that the learning process will end once [4]:

- A fixed number of iterations: it is decided beforehand the number of times that the entire set is expected to be scanned. Once this value is reached, the process stops and the resulting network is accepted.
- Once the error is below a pre-established threshold: in this case, an error function is defined as a maximum error value. The learning process stops once the error function yields to a value below the threshold. It may happen that the expected threshold is never achieved. In that case, the solution hasn't been found.
- When the weights' modifications are irrelevant: for some models, the defined training scheme causes the connections to adjust less abruptly at each iteration. If the learning continues, the system reaches a point where there are no more variations between the connections' weights. Let's say the NN cannot learn more. If reached, then is said the network converges and a solution is found.

Once the training set is defined (following the required parameters) and the convergence criteria is chosen as best suits the problem type and the networks structure, is time to select a final crucial parameter, the learning scheme the ANN follows to attempt to resolve the problem. All learning methods can be classified in two major categories [32]:

- Supervised Learning: this scheme incorporates an external "teacher", so that for each input signal, it is said which is the expected correct answer or output. During the process, global information might be required. Some prototypes of supervised learning include: error-correction learning, reinforcement learning and stochastic learning. Main drawback for this technique is error-convergence, as the aim is to determine a group of weights which minimizes the error.
- Unsupervised Learning: it uses no external teacher and is based upon local information only. It is sometimes referred as Self-Organization Learning, in the sense that the

network organizes the data presented to it, and detects their emerging collective properties, regularities and/or redundancies. Some models for this scheme are Hebbian learning and competitive learning.

Is noticeable to mention, that some consider a third learning scheme to exist. Is called *Reinforcement Learning*, and it is a variation of the supervised learning where exact information on the error at the network output (at each learning iteration) is available [4]. The scheme simply determines whether the outcome for a given pattern is adequate or not, the result is achieved by the Stochastic Gradient Descent algorithm (SGD), which rewards the network for good guesses and penalizes it for bad ones [27].

Today, *Deep Neural Networks* are involved in many applications, and its development is increasing exponentially. Most used models and its main applications are [27]:

- Convolutional Neural Networks (CNNs): are in fact one of the main reasons the power of deep learning is recognized. They potential major advances for machine vision, that is the principal reason for its application to be self-driving cars, robotics, drones, and treatment for the visually impaired.

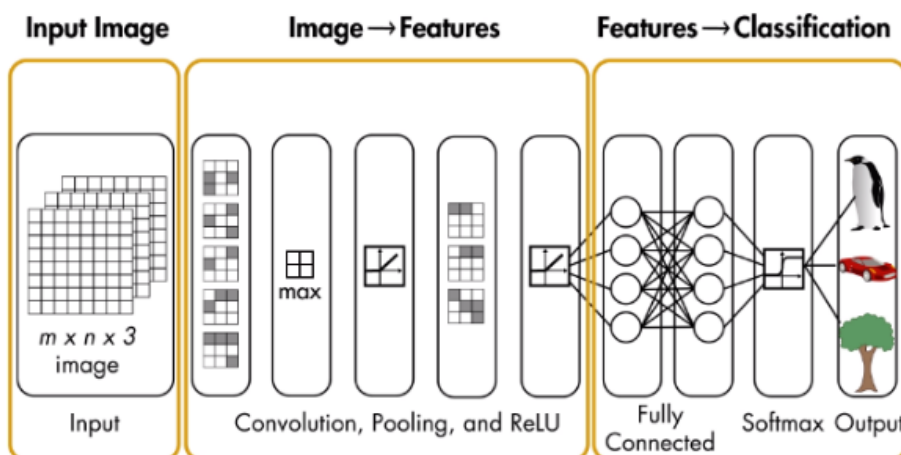


Figure 26.- Convolutional Neural Network Architecture [33]

- Recurrent Neural Networks (RNNs): it is contained in the feed-forward neural network family, whoever this kind of system sends the information over time-steps. Historically they have been harder to train but recent advances have made them more approachable. Their loop-connection, as shown in Figure 27, allows them to achieve higher accuracy in time-series prediction, speech-language synthesis or modelling, music generation, and music information retrieval.

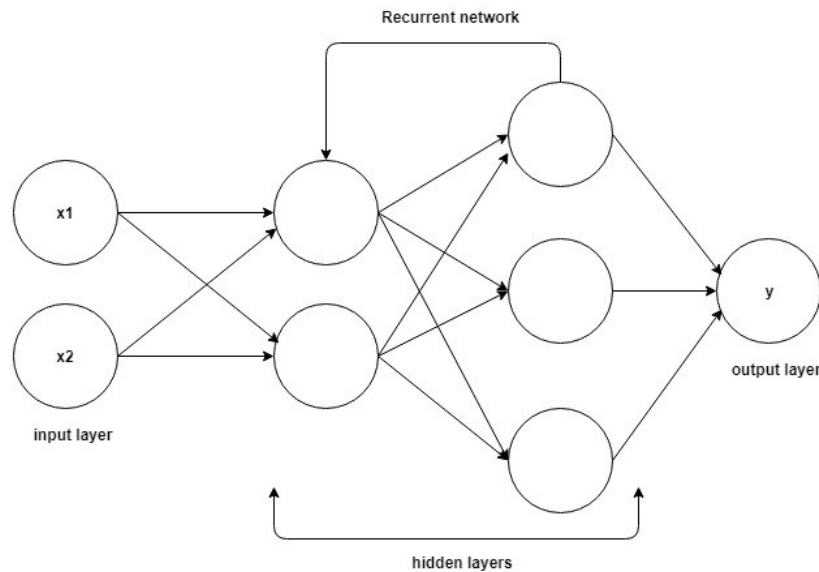


Figure 27.- Recurrent Neural Network Structure [31]

- Recursive Neural Networks: a generalization of RNNs. However, they can recover granular structure and higher-level hierarchical structure in datasets such as images or sentences. These advantages are achieved thanks to the fact that the connection weights are shared at every node of the system while in RNNs the sharing is performed along the whole length of the sequence. This means, that they are better at decomposing larger information for segmented processing. Applications cases for this type of network are: image scene decomposition. Audio-to-text transcription and natural language processing (NLP)

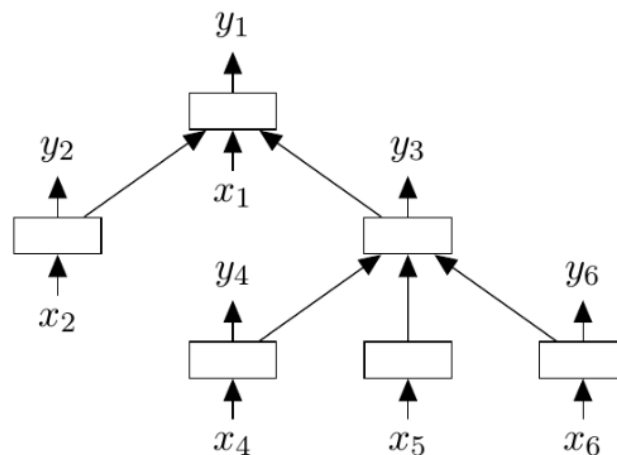


Figure 28.- Recursive Neural Network Scheme [34]

2.4. Time Series Prediction using Neural Networks

A time-series is a collection of data or values for a particular event across time, $\{x(t)\}_{t \in \mathbb{R}^+}$ where t represents the time variable and $x(t)$ the serie's values at that time instance. Time-series are characterized because their time evolution does not depend explicitly of the time variable, but also of the time-series values in previous time instants.

The capacity of predicting the future behaviour for a time-series is an important objective for various field study, such as: physics, biology, medicine, and more. Occasionally, it is possible to construct models that permit to describe the temporal evolution of a certain series, which are usually a group of differential equations. However, for most time series describing real phenomenon, such as temperature, power consumption, and more - is not easy to define a model that approximates to the real time-series evolution [4].

If the second is the case, is necessary to use approximation methods to build models in order to facilitate the prediction, the difficulty of achieving the prediction depends on the dynamic behaviour of the series. Following sections are mostly centred in defining time-series that can be described as NAR (nonlinear autoregressive) models.

These models collect the temporal behaviour from the series by expressing its value in a time instant $t + 1$ as a non-linear function of $r + 1$ values of the time series in previous instants of time, mathematically expressed as:

$$x(t + 1) = F(x(t), x(t - 1), \dots, x(t - r)) + \varepsilon(t) \quad \text{Eq. 2}$$

where t is the discrete time variable, $\varepsilon(t)$ a residual error assumed to be white gaussian noise (WGN), and F a non-linear and unknown function, which has to be an estimation from the observed time-series data, $\{x(t)\}_{t=0, \dots, N}$. Therefore, the function F construction is based on previous samples by means of approximation techniques, such as: neural networks [4].

Artificial Neural Networks have been widely used last years for prediction of time-series. Taking advantages of their own features, as:

- Ability of neural networks to approximate and capture connections from an example set, without the necessity of knowing additional information about the data distribution.
- The neural networks' capability to build non-linear relations. Even starting from incomplete data or noisy information.
- Models based on neural networks are easy to build and use.

In a generic sense, time-series prediction expects to anticipate future values of signal sequence from a set of samples from the original data. Predictions can be performed by evaluating the time series over one-time step or multiple steps.

2.4.1. Prediction in One-time Step

It consists in predicting the value for the time series in the time instant that immediately follows the current time instant t , using the values obtained until that moment. Static NNs can be used to approximate the function defined in Equation 2, obtaining the following model:

$$\tilde{x}(t + 1) = \tilde{F}(x(t), x(t - 1), \dots, x(t - r)) + \varepsilon(t) \quad \text{Eq. 3}$$

where \tilde{F} represents the neural approximation for function F and $\tilde{x}(t + 1)$ is the prediction given by the neural model in the time instant $t + 1$. Schematically, the network must have $r + 1$ input neurons that receive the values for the previous $r + 1$ time instants and a single output neuron that represents the prediction in the time step that follows the current instant. An example is depicted in Figure 29.

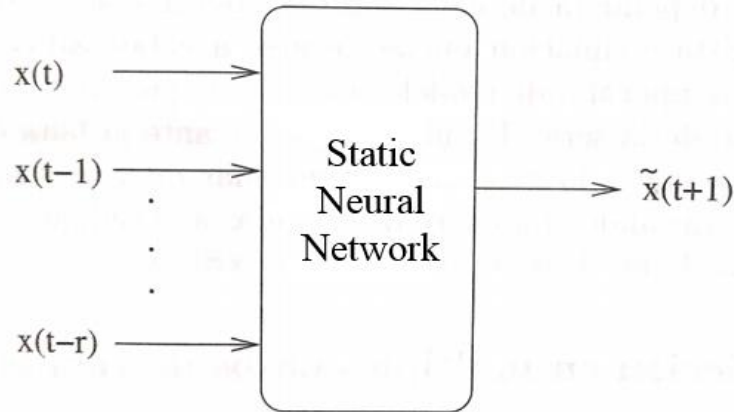


Figure 29.- One-Step Time Prediction [4]

The training sets are obtained from the available samples as follows: given the set $\{x(t)\}_{t=0,\dots,N}$ and the value of r , each pattern is constructed by using the $r + 1$ previous values, as expressed in Table 1. Thereby, each input section has the necessary information to predict the value of the series in the instant $t + 1, \forall t = r, \dots, N - 1$.

	Input	Expected Outcome
Sequence 1	$x(r), x(r - 1), \dots, x(1), x(0)$	$x(r + 1)$
Sequence 2	$x(r + 1), x(r), \dots, x(2), x(1)$	$x(r + 2)$
Sequence 3	$x(r + 2), x(r + 1), \dots, x(3), x(2)$	$x(r + 3)$
Sequence 4	$x(r + 3), x(r + 2), \dots, x(4), x(3)$	$x(r + 4)$
...
Sequence $N - r$	$x(N - 1), x(N - 2), \dots, x(N - r), x(N - (r + 1))$	$x(N)$

Table 1.- Training pattern. One-time-step prediction [4]

After obtaining the input-output pairs, the network is trained expecting to capture the relation between the time-series' past values and the current. The weights are adjusted to minimize the quadratic error in the output neuron, using the equation:

$$e(t + 1) = \frac{1}{2} (x(t + 1) - \tilde{x}(t + 1))^2, \forall t = r, \dots, N - 1 \quad \text{Eq. 4}$$

taking into account that $x(t + 1)$ should be the actual result of the time series for a time instant while $\tilde{x}(t + 1)$ is the prediction given by the NN. The model proves to have great results.

However, the key strategy is to define correctly the length of the temporal sequence presented as the network's input; that is, the value of r .

2.4.2. Prediction on Multiple Steps

This model consists in not only estimating the immediate next value, but also in a much further time instant, both values create the *prediction interval* $[t + 1, t + h + 1]$, where h is a natural number that represents the *prediction horizon*. As previous case, the prediction is performed based on the actual values obtained until the current moment. There are two type of schemes available:

- Prediction Scheme A: In this case, the neural network uses a recurrent model and the past values. The system is defined by the mathematical function: $\tilde{x}(t + h + 1) = \tilde{F}(x(t + h), x(t + h - 1), \dots, x(t + h - r))$ (Eq. 5). However, in time instant t , not all the input information is available, as the values for the series while $h > 1$ are unknown. To face this issue, the estimated values that the network predicts are used instead. Figure 30 shows an architecture of the network.

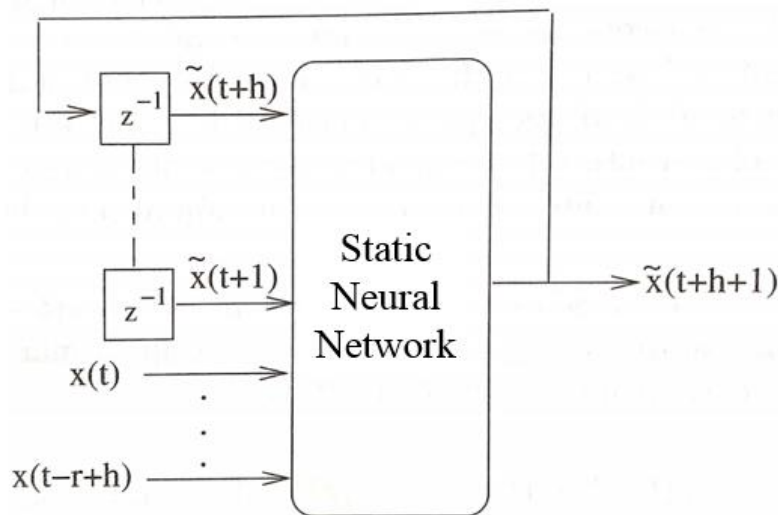


Figure 30.- Multiple Time Step Prediction [4]

The predictions are obtained following the next model:

$$\begin{aligned}\tilde{x}(t + 1) &= \tilde{F}(x(t), x(t - 1), \dots, x(t - r)) \\ \tilde{x}(t + 2) &= \tilde{F}(x(t + 1), x(t), x(t - 1), \dots, x(t - r + 1)) \\ \tilde{x}(t + 3) &= \tilde{F}(x(t + 2), x(t + 1), x(t), \dots, x(t - r + 2)) \\ &\dots \\ \tilde{x}(t + h + 1) &= \tilde{F}(x(t + h), x(t + h - 1), \dots, x(t + 1), x(t), \dots, x(t + h - r))\end{aligned}$$

Main drawback for this scheme is that when the model compares between the samples (to minimize the error), a group of the input values are predicted by the network in past time instants. Therefore, if an error was committed while predicting them, the error will

be dragged into future outputs, possibly lowering the quality of the predictions, thus damaging the network.

- Prediction Scheme B: This model uses a NAR (Non-linear autoregressive) predictor to directly estimate the value of the series in a time instant $t + h + 1$ using the information available in time instant t .

$$x(t + h + 1) = G(x(t), x(t - 1), \dots, x(t - d)) \tag{Eq. 6}$$

where G is a non-linear function that relates the time series value in time instant $t + h + 1$ with a limited sequence of values for the series available in the current instant. By using the input vector $(x(t), x(t - 1), \dots, x(t - d))$, static neural networks are used to approximate G , as:

$$\tilde{x}(t + h + 1) = \tilde{G}(x(t), x(t - 1), \dots, x(t - d)) \tag{Eq. 7}$$

The scheme, still expects to estimate the future value after the prediction horizon, while not using any predictions but actual past values. Figure 31 shows the schematic of the model and Table 2 how the training patterns are obtained.

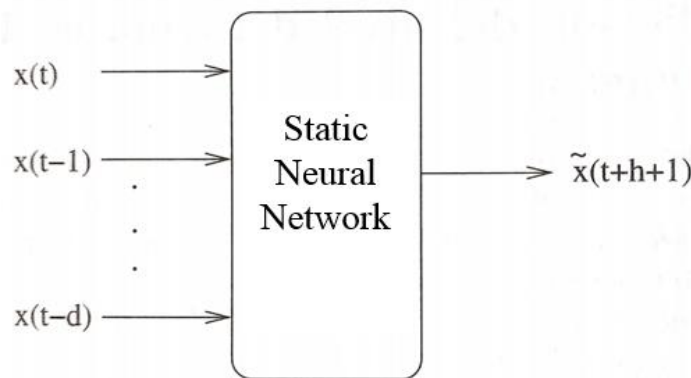


Figure 31.- Delay-based Multiple-time-step Predictor Architecture [4]

	Input	Expected Outcome
Sequence 1	$x(d), x(d - 1), \dots, x(1), x(0)$	$x(d + h + 1)$
Sequence 2	$x(d + 1), x(d), \dots, x(2), x(1)$	$x(d + h + 3)$
Sequence 3	$x(d + 2), x(d + 1), \dots, x(3), x(2)$	$x(d + h + 3)$
Sequence 4	$x(d + 3), x(d + 2), \dots, x(4), x(3)$	$x(d + h + 4)$
...
Sequence $N - d - h$	$x(N - h - 1), x(N - h - 2), \dots, x(N - 1 - h - d)$	$x(N)$

Table 2.- Training patterns. For delay-based Multiple-time-step Prediction [4]

After the training, the network is expected to lower the average error in the output, by following the equation:

$$e(t + h + 1) = \frac{1}{2}(x(t + h + 1) - \tilde{x}(t + h + 1))^2, \forall t = d, \dots, N - h - 1 \quad \text{Eq. 8}$$

As this predictor is based on real previous sequences, there is no possibilities to concatenate past errors, however is important to take into account that as the prediction horizon increases, the relationship between the temporal values tends to deteriorate, hence, there is no sense to design such prediction schemes.

After reading, analyzing and creating a more solid knowledge of neural networks, their capabilities, advantages and drawbacks, is time to test them in the laboratory environment and validate how close the theory comes with the practice. Next chapter explains the techniques used to perform the tests for this Master's Thesis.

3. Methodology - Project Development

The present chapter describes all the steps and analysis methods used in order to obtain the results for this Master's Thesis. The main idea is to design and use a Neural Network by means of available software that learns the patterns and behaviours of an optical source, for our purposes a SM-VCSEL. To achieve the objective, certain steps need to be placed in action:

1. Characterization and preparation of the VCSEL to be used in the laboratory.
2. Basic learning on a software oriented Neural Network solution: main drawbacks and advantages, available software, previous case studies in time-series problems.
3. Input current - Optical power study: creating a Neural Network that can predict a specific input current signal in order to obtain a desired optical power signal.
4. Optical power - Optical phase study: in a parallel manner as before, analyzing the possibility to design a neural network capable of learning patterns between an optical power signal for a VCSEL and the corresponding optical phase.

3.1. VCSEL's Characterization

The optical source to be used is a RayCan 1550 nm SM-VCSEL, designed for high-speed, high-performance communication applications, its electrical and optical characteristics as well as maximum reachable values are found in Appendix A. Basic features include:

- Low dependence of electrical and optical characteristics over temperature.
- Data rate up to 10 Gb/s.
- Cylindrical transistor outline (TO) package with *single-mode* fiber pigtail.
- Enclosed thermistor and thermoelectric cooler (TEC, also known as Peltier device)

The first step is to create a board that allows the VCSEL to be powered in order to characterize its static parameters: lasing threshold, wavelength and optical power. The design must include all connections required to connect the available pins to the expected temperature and current controllers as well as SMA connectors to feed the VCSEL with bias and high speed data.

The software to design the board is Eagle from Autodesk, since it's fairly intuitive and is possible to be used under free student licenses, plus it includes packages that create all GERBER files (different board layers) while validating that the design satisfies the minimum technical requirements for manufacturing. A footprint and symbol are needed for the VCSEL, considering the software doesn't include it.

3.1.1. New footprint and symbol for the VCSEL

The footprint must be based on the device's schematic given by its manufacturer, shown in Figure 32 and Figure 33. In addition, Table 3 defines the function for each of the device's five (5) pins.

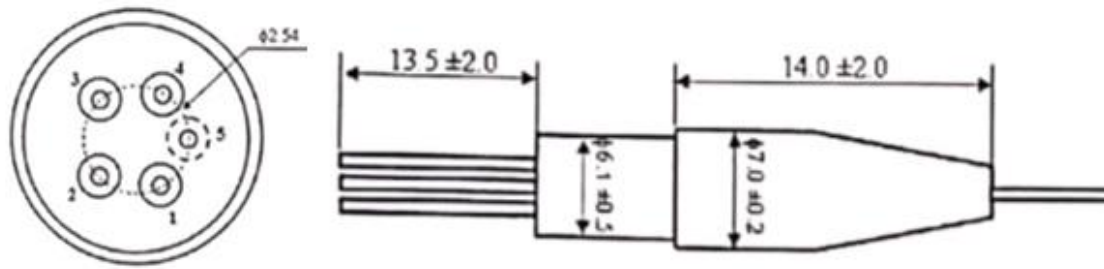


Figure 32.- Side and bottom view of VCSEL. Dimensions in millimeters (mm) [35]

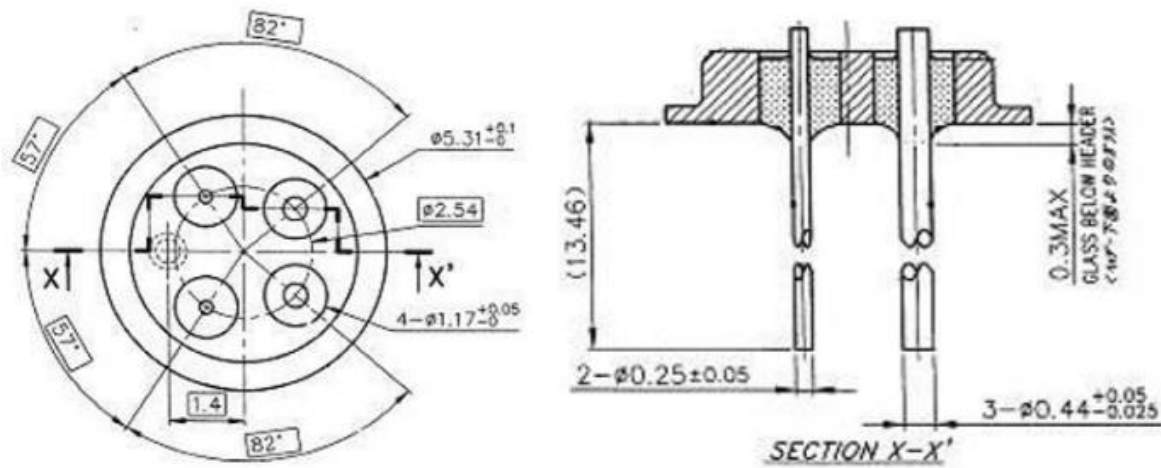


Figure 33.- Transversal and bottom view of the VCSEL with dimensions (in mm) [35]

With the previous schematic and section X-X' is known that pin number five (5) is located 1.4 mm from the centre while all others are distributed in a 1.27 mm radius from it in different angles:

- Between pairs 1-5 and 4-5 a 57° angle.
- Between pairs 1-2 and 3-4 an 82° angle.

By trigonometry, the horizontal and vertical coordinates for each pin in millimeters are computed taking as horizontal axis the one perpendicular to the X-X' section, furthermore the width of each pin is depicted in the transversal image. The results are given in Table 3, and the footprint and symbol generated in Figure 34.

Pin out and function	Pin width	XY Coordinates
1 – VCSEL cathode (-)	0.25 ± 0.05	(1.0651; 0.6916)
2 – TEC (-)	$0.44^{+0.050}_{-0.025}$	(0.8331; -0.9584)
3 – TEC (+) & Thermistor	$0.44^{+0.050}_{-0.025}$	(-0.8331; 0.9584)
4 – VCSEL anode (+)	0.25 ± 0.05	(-1.0651; 0.6916)
5 – Case & thermistor	$0.44^{+0.050}_{-0.025}$	(0; 1.21)

Table 3.- VCSEL's pinout-function correspondence, width for each pin and XY coordinates

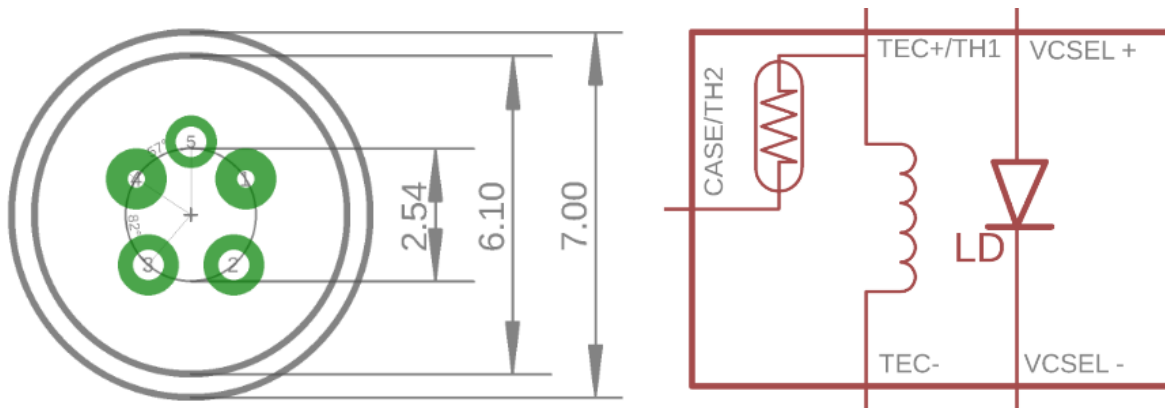


Figure 34.- Footprint and symbol created for the VCSEL.

3.1.2. Designing and manufacturing the board

Once the footprint and the symbol are created, next step is to design the board - as previously mentioned - that permits the VCSEL to be used, when the correct parameters are given to it. The design includes two SMA connectors in order to be able (if needed) to feed the device differentially, these connectors must be equally distanced from pins, to avoid data synchronization issues.

First design was dismissed (Appendix B), due to not having enough space to connect bias-tees (needed to feed the device with both current and data simultaneously) but also due to the fact that the paths to transmit data and current are too thin and can be easily damaged in higher currents.

Notice in the second and definitive version of the board, that the paths to the VCSEL's parameters are as wide as possible. Below is shown in software and preview of manufacturing (Figure 35). The design includes two holes to screw the board and VCSEL into a dissipation element constructed by the UPC's mechanical department. The final design with its dimensions can be seen in Appendix C.

Figure 36 depicts the VCSEL installed in the manufactured board and screwed to the dissipation element, which is not quite clear in the picture, nevertheless Appendix D shows more pictures from other angles.

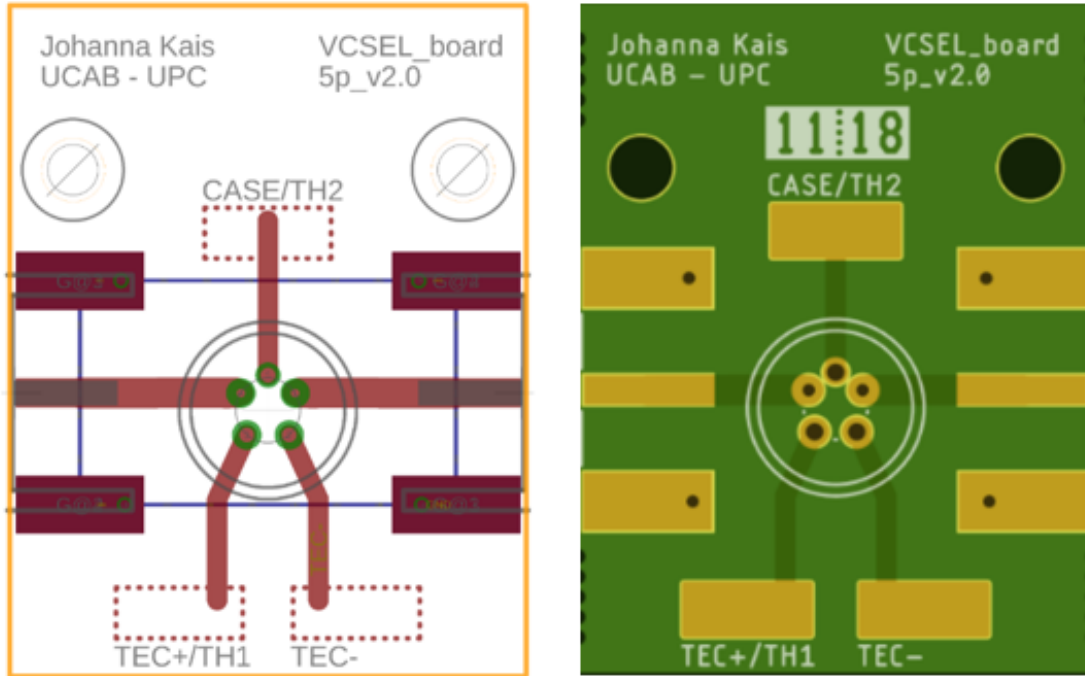


Figure 35.- VCSEL board design. Left: software view. Right: pre-manufacturing stage.

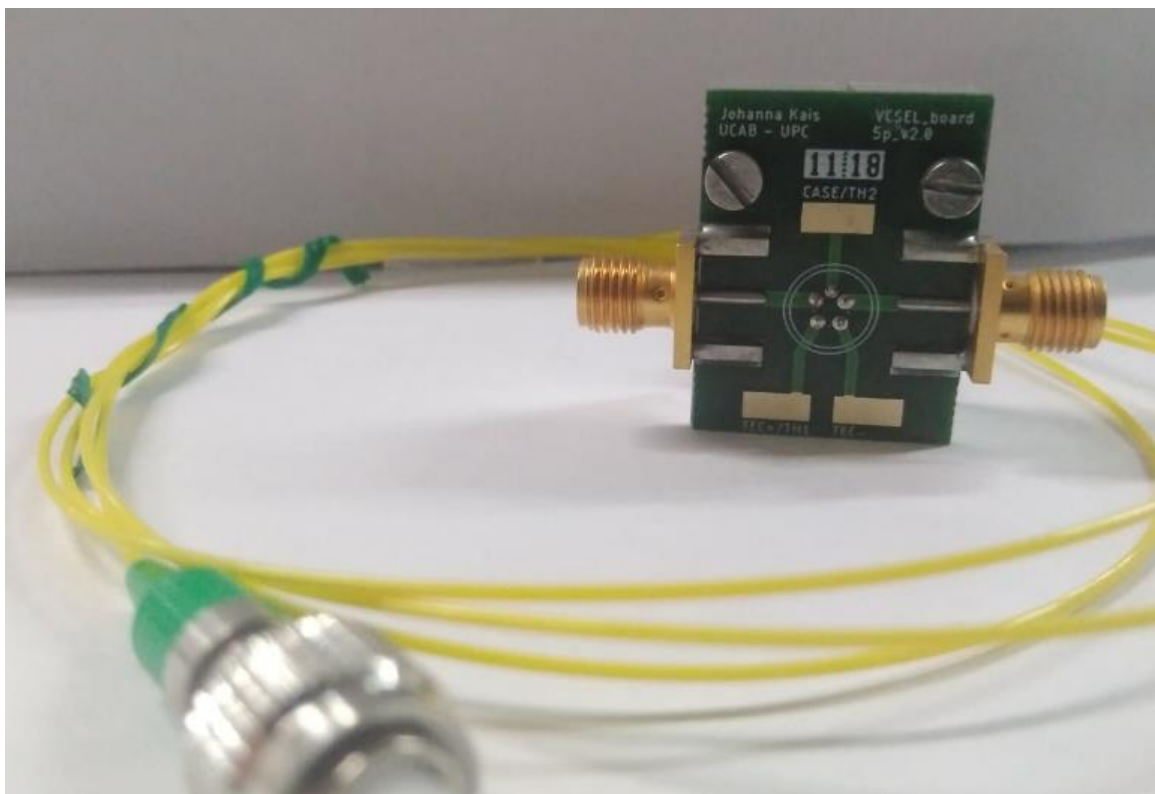


Figure 36.- VCSEL included in manufactured board.

3.1.3. Light-Current-Voltage (L-I-V) Curves

Once the board is ready and the VCSEL is included in it, is time to obtain the static parameters for the optical source. One of most important quantification in a light emitting device is known as LIV curve, which simultaneously measures the electrical and optical power characteristics of the device.

The properties of the semiconductor optical sources, as described in previous section, significantly depend on the temperature. Therefore, the characterization of the LIV curves may also significantly depend on the temperature, and this has to be investigated. Due to those facts, the measurement of the LIV curves was done by previously controlling the laser temperature as next detailed.

The curve is obtained by sweeping the thermistor value (according to the manufacturer datasheet - Appendix E) in a range from 10 to 50 degree Celsius ($^{\circ}\text{C}$), in 5 $^{\circ}\text{C}$ steps. This swipe is done at different input currents to acquire maximum data, making sure the TEC and VCSEL maximum current are respected (0.7 A and 15 mA, respectively).

Consequently, for each loop the input current will remain the same, as the thermistor value is modified. The used instruments are (more information on the devices is explained in chapter 5):

- THORLABS Laser Diode Controller (LDC200CV)
- THORLABS Thermoelectric Temperature Controller (TED200)
- EXFO Optical Spectrum Analyzer FTB-2.
- GO4FIBER Power meter GFHS-A.
- TENMA Digital multimeter 72-7780.

In order to be able to connect the VCSEL to the controllers (both temperature and current) is necessary to design cables by following the instructions given by THORLABS to use these devices and obtain the expected parameters from them. As can be seen in Appendix E, the LDC has a 9-pin D-SUB jack out such as shown in Figure 37:

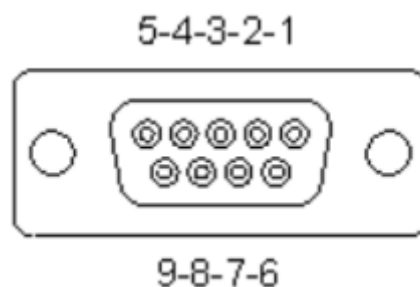


Figure 37.- Pin assignment of the “LD OUT” jack (female, rear panel view) [36]

The LDC is capable of control a laser in both polarities: cathode or anode grounded (CG and AG, respectively) according to the necessities of the experiment. In this case we are using a CG configuration, therefore and based on the operation manual of the laser diode controller, the cable for this device is a 9-pin connected as explained in Table 4:

Pin #	Connection
1	Interlock and status LASE ON / OFF
5	Ground for pin 1
3	Laser diode ground
8	Laser diode anode (in case of polarity CG)

Table 4.- Connections for LD OUT pins [36]

To connect the cable to the VCSEL board a cable with a SMA terminal (according to the VCSEL connector) is included between the laser diode pins. The final result of the cable is depicted in Figure 38:



Figure 38.- Cable to connect VCSEL and LDC.

Additionally, the VCSEL expects to be connected to the temperature controller so can measure and control this changes to force the experimental results to be current dependant only. According to the operating manual, the output for the TED is a 15-pin D-SUB jack, the pins distribution follows Figure 39 and Table 5:

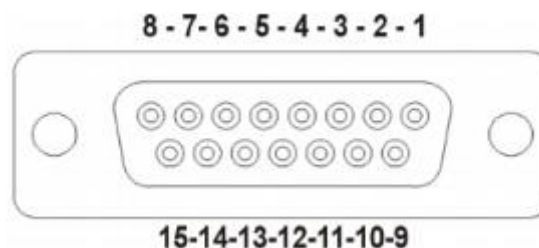


Figure 39.- Pin assignment of the "TE OUTPUT" jack (female, rear panel view) [37]

Pin #	Connection
1	Status LED (+) TEC ON / OFF indicator
3	Thermistor (-), ground
4	Thermistor (+)
5	TEC (+)
4	TEC (-)
5	Status LED (-)

Table 5.- Connections for TE OUTPUT pins.

In this case, the external LED will only indicate whether the controller is ON or OFF. The cable constructed is the jack with the four (4) wires and the LED. The final cable is shown in Figure 40:



Figure 40.- Connection cable for VCSEL and Temperature Controller.

The parameters measured in order to characterize the VCSEL are:

- Optical wavelength (λ), its optical power (P_λ) and the optical signal-to-noise ratio (OSNR) with an optical spectrum analyzer (OSA) in average mode using ten (10) measurements.
- Total optical power (P_{TOTAL}) measured with the optical power meter.
- VCSEL's voltage (V_{VCSEL}) is obtained by means of a voltmeter between anode (+) and cathode (-) of the device.

- Finally, TEC current is measured by the temperature controller connected between the thermistor's pins and the TEC + and TEC - pins of the VCSEL.

The setup used to perform the measurements is shown in Figure 41 and Figure 42:

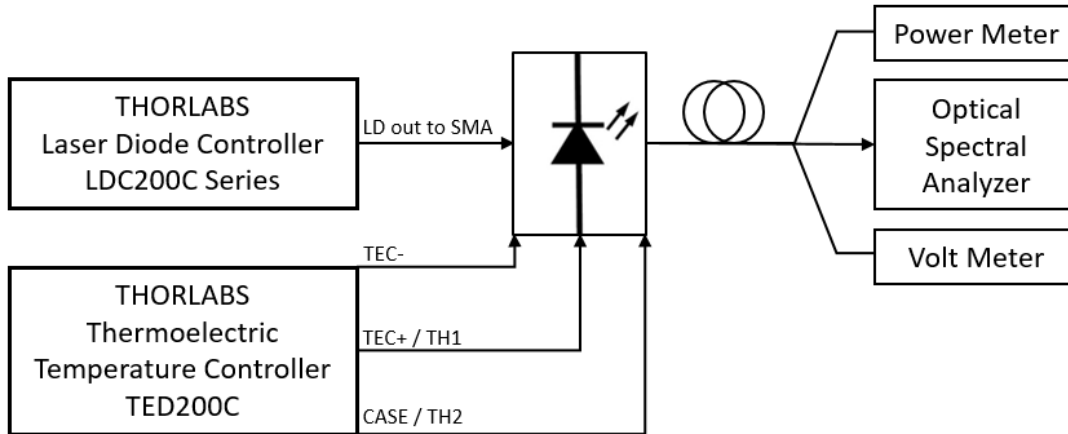


Figure 41.- General experimental setup for VCSEL characterization.



Figure 42.- Experimental setup for VCSEL characterization (optical power)

First characterization to be done using the layout previously described (Figure 41), connecting the LDC200C into the SMA corresponding to the VCSEL's anode while the cathode is short circuited. To obtain the measurements, a single current value (in unitary steps from 2mA to 11mA) is statically selected; then, temperature is swipec in the five Celsius degree steps in order to obtain all the experimental values mentioned above.

After the characterization is done, VCSEL parameters are known and the acquired results can be filled in tables, graphs or other visual aids that allow to analyze and compute the device and its behaviour. The tables with all the data collected at this stage can be found in Appendix H, while graphs and other visual aids, as well as comparison of the results are later described in chapter 4 (Results).

3.2. Creating and Training a Neural Network

As was mentioned in previous chapters, neural networks are being developed in a wide range of fields and each day another view and perspective comes into action with this topic. This Master's thesis used a MATLAB's toolbox called Deep Learning toolbox (previously called Neural Network Toolbox) that allows a test environment to design, implement, visualize and simulate deep NNs with algorithms, pre-trained models and applications [38]. It supports feedforward networks, radial basis networks and dynamic networks. In order to be able to use it, first it is necessary to get familiar with the instrument and its capabilities.

3.2.1. MATLAB's Neural Networks Toolbox

Basically, the Toolbox is enabled to compute different scenarios that are included in a range of four (4) principal tasks, such as: Function Fitting, Pattern Recognition, Data Clustering and, our focus: Time-Series Analysis.

Any of this options can be designed using the MATLAB's GUI (graphic user interface) by using the command `nnstart` or through command line operations and script coding. After the problem has been identified as one of the possible tasks the NN are capable to resolve, a certain numbers of steps are taken to use the toolbox, which are:

- Collect the Data to be analyzed: Neural Networks must have training sets (input-output pairs) that allow it to learn the patterns of the system.
- Create the Network: by selecting the correct kind of problem between the available options (mentioned above, in our case, selecting: Time-Series Analysis).
- Configure the Network: in the case of time-series analysis, is possible to choose between three different sub-problem solvers. Appendix I shows the diagrams and quick descriptions of each of the solvers.
- Initialize the Weights and Biases: a crucial step in the design process and later use of the NN, as it defines the number of hidden layers (HL), the initial delay of the training process (D), the bias applied at the layers in different stages of the training (if applicable) and the weights of this layers.
- Train the Network: after the basic characteristics for the algorithm are selected, is time for the actual training process. To do so, the total number of samples (data) is divided in three (3) groups according to the designer wishes: training, validation and testing.
 - ✓ Training set: is recommended to be more than 50% of the data samples (MATLAB's default value is 70%) as this is the amount of information that is used to learn the patterns of the problem in the hidden layers and adjust according to its errors.
 - ✓ Validation set: this input-output pairs are used to confirm that the network has learn general characteristics of the data and not specifics that can lead to bad predictions, by default in the toolbox is 15% of the data set. It also helps to stop the training to avoid overfitting, usually is done by including a validation stop.
 - ✓ Test set: is the remaining 15% of the input-output data pairs. Have no effect of training and provide an independent measure of the network's performance during and after the training process.

- Use the Network: once the network is generated can be used to approximate possible values of data samples of the same kind of the used to train the neural network.

3.2.2. Designing, Training and using a Basic Neural Network

Before using the toolbox in our main case studies, we tested a basic time-series to become familiar with the software abilities, trying all different problem solvers and different configurations. As mentioned before, our goal is to study the VCSEL’s response as a time-series problem, as shown in Appendix I there are three (3) different manners to attack this kind of task. This Master’s Thesis focuses on the Nonlinear Input-Output, as other models are defined for autoregressive outputs patterns that are recovered and then included again into the prediction system, which is not our case.

After the problem is chosen and according its capabilities, the input time series $[x(t)]$ and the target time series, defining the desired output $[y(t)]$ are selected; as well as the percentages for the use of the data samples as manner of: training, validation and testing. Afterwards, the number of hidden layers in the networks and the delays to compute the results are set.

Figure 43 shows an example of an architecture with 15 hidden layers and a five (5) time-steps delay. An example from MATLAB’s repository of Time-Series Analysis evaluates eight (8) different measurements and predicts between three (3) possible outcomes; defined as inputs $[x(t)]$ and outputs $[y(t)]$, respectively. This architecture is used to learn how the MATLAB’s Toolbox operates, in order to have a basic knowledge when designing the NNs in the experimental steps defined in the rest of this chapter.

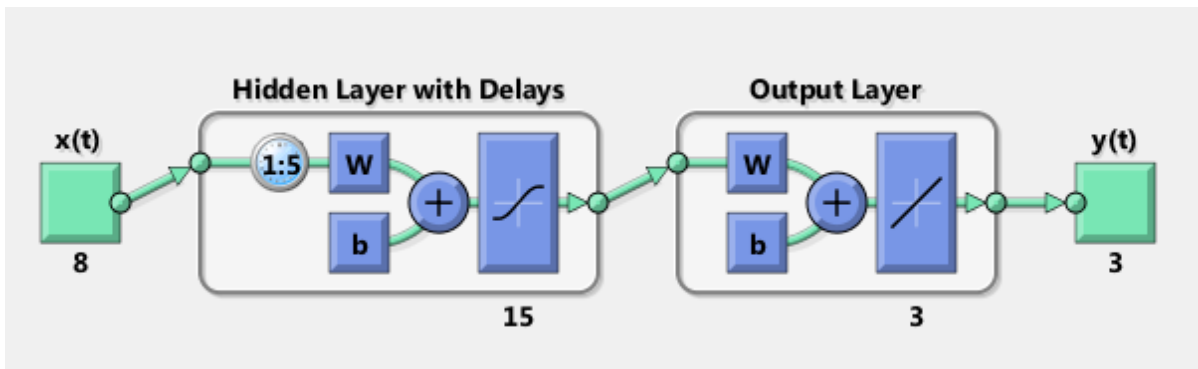


Figure 43.- Architecture of an example Neural Network.

Finally, before actually training the network an algorithm is defined to perform the learning and adjusting process. The possible algorithms are:

- Levenberg-Marquardt (LM): is an algorithm that typically requires more memory but less time. Training will stop when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.
- Bayesian Regularization (BR): it takes more time, but can result in good generalization for difficult, small or noisy datasets. Training stops according to adaptive weight minimization.

- Scaled Conjugate Gradient (SCG): is the one that uses less memory. Training stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.

If the training of the NN is done through MATLAB's GUI, the interface prompts step by step the user to select the necessary parameters. After the network is trained, a set of plots and information is delivered to the software user, the ones used in this Master's Thesis as valid information are the following:

- Target Values: defines the division of the whole data set into the three (3) components for the NN training previously stated: training, validation and testing.
- Mean Square Error (MSE): is the average mean squared difference between the outputs and targets in the stages of the NN generation: training, validation and testing. Zero means no error.
- Regression (R): this value measures the correlation between outputs and targets. A unitary R value means a close relationship, while a zero value represents a random relation between the values.

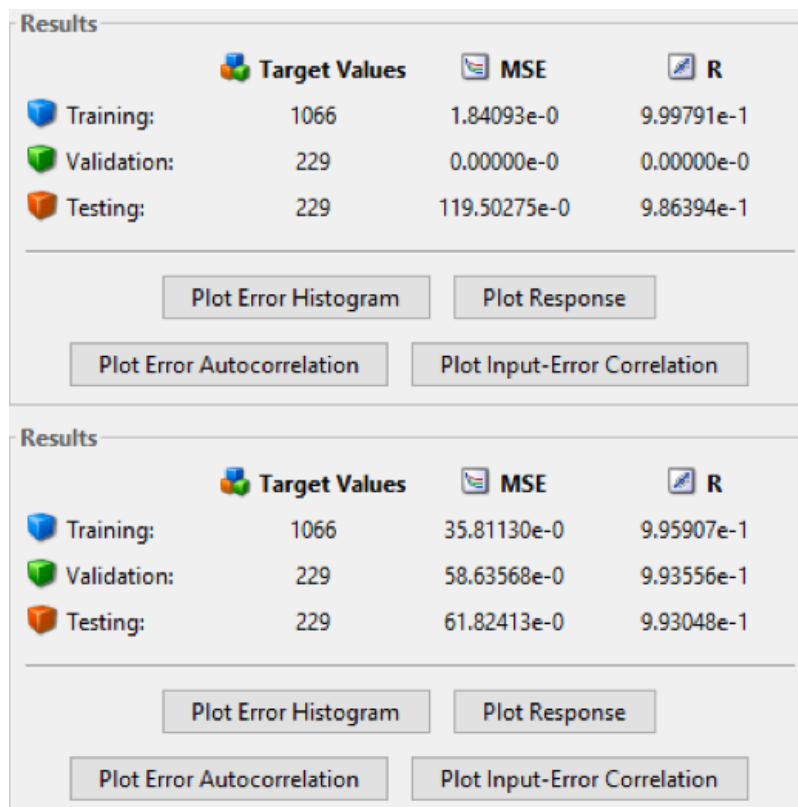


Figure 44.- MSE and R results from different training processes; BR and SCG respectively. Notice that the algorithm BR doesn't include validation values, as is not used to stop the training process for this case.

- Performance: is a relation between how the training is developing, it shows the state of the training, validation (if the algorithm uses it) and testing process for each iteration of the process. The next example (Figure 45) pictures a training that stopped at iteration number 7, as its next six (6) loops didn't showed any improvement, the end of training process was 13.

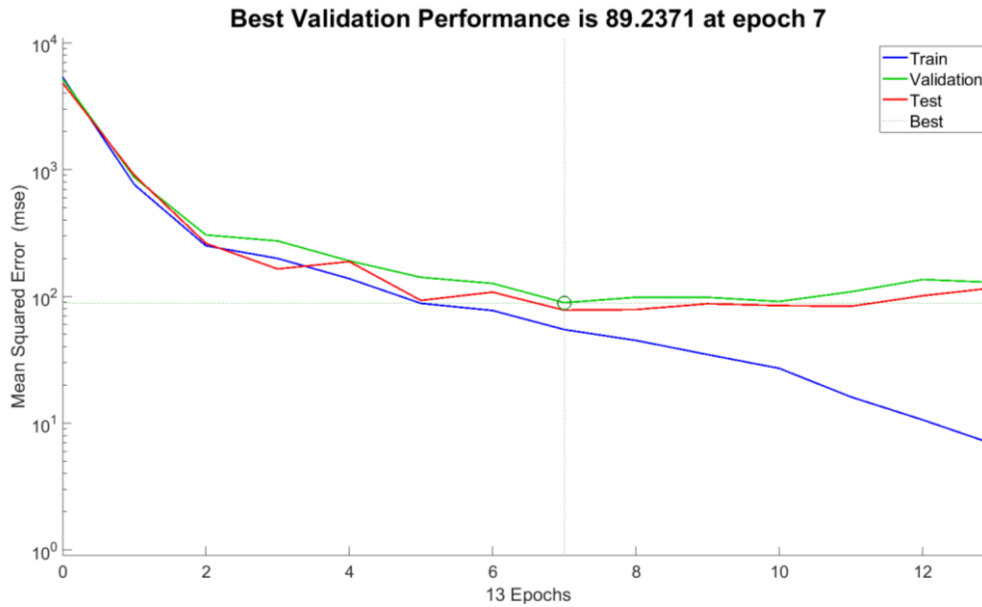


Figure 45.- Performance results for the training process using the algorithm LM.

- Regression: shows a relation between the target aimed during the training and the output obtained. The ideal graph is that the data samples (circles) and the “Fit” line match as much as possible, in this sense the target have been achieved. Figure 46 shows a training that has reached a close relation between the samples in the training state, but the overall regression is not as good due to a random relation in the testing stage.

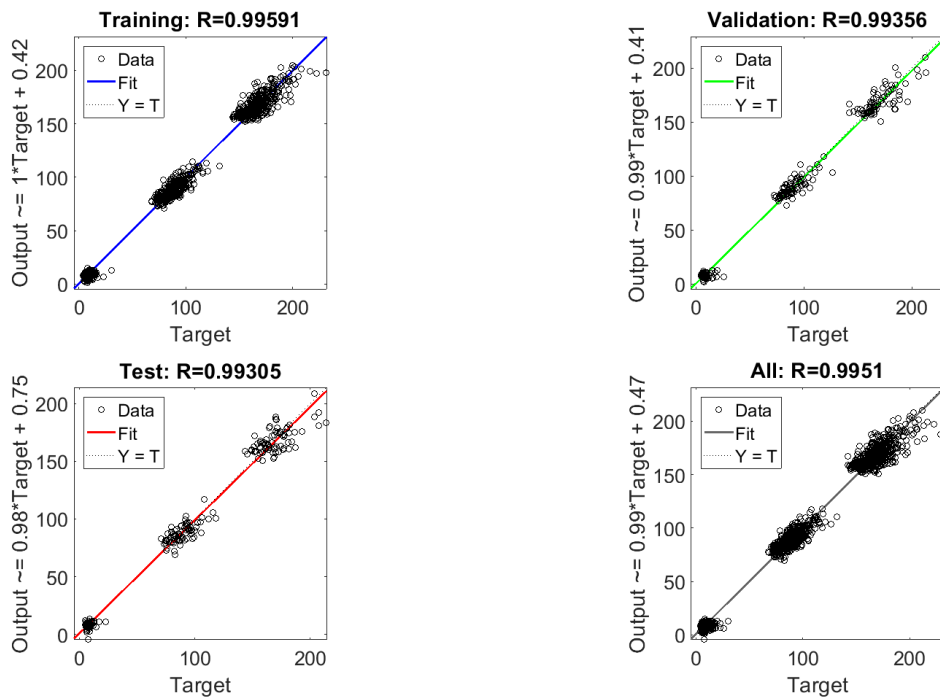


Figure 46.- Regression results for a training process using the algorithm BR.

Once the training process is described, is time to develop a technique that enables to unite the knowledge in optic communications and the neural networks, especially in a manner that allows to develop our main interest. That is, designing a neural network capable of learning the relationship between the input current given to the research device (VCSEL) and its output optical power, as initial stage; and the correspondence between the input current and the output optical phase in a more complex scenario.

3.3. Training a Neural Network to estimate a VCSEL's input current given its output optical power

The first phase of the experiment is to define where the neural network is capable of predicting a simpler relation, between the input current given to the VCSEL and the output optical power, as a matter of avoiding the complexity of solving a semiconductor's rate equations due to its nonlinear nature. Where an analytical solution can be obtained for cases of small-signal modulation, where laser is biased above threshold ($I_b > I_{th}$).

New devices are included in this experimental part, such as (more information on the devices is explained in chapter 5):

- TEKTRONIX Mixed Signal Oscilloscope (MSO 70804C)
- TEKTRONIX Arbitrary Waveform Generator (AWG 70002A)

To perform this first part, a set of 300 bits at 1 Gb/s was generated in MATLAB and modulated using an exponential signal, as the one used by Altabás *et al.* in their 2016 paper; mathematically, is: $y = a * e^{-tb}$, where a is the bias coefficient and, b the decay factor for the exponential function. Choosing of this signal is based on results depicted in Figure 2, where an ideal current signal doesn't correspond to the corresponding output optical power. The VCSEL gives faster responses for higher input currents than lower ones, resulting in an asymmetry of the laser's response, main reason to decide using NNs to predict ideal input current signals based on an optical power pattern objective.

Figure 47 depicts a sample of the bit stream, the exponential signal with coefficient values: $a=1$ and $b=4$; and the result from filtering the exponential signal, in order to avoid immediate transitions between states. For experimental purposes, we have created a bank of signals creating combinations between the coefficients values going from a 1.5 to 4.5 in 0.5 steps; that is, an array of coefficients, $C = [1.5, 2, 2.5, 3, 3.5, 4, 4.5]$; resulting in 49 different current signals, enough information to perform NN trainings.

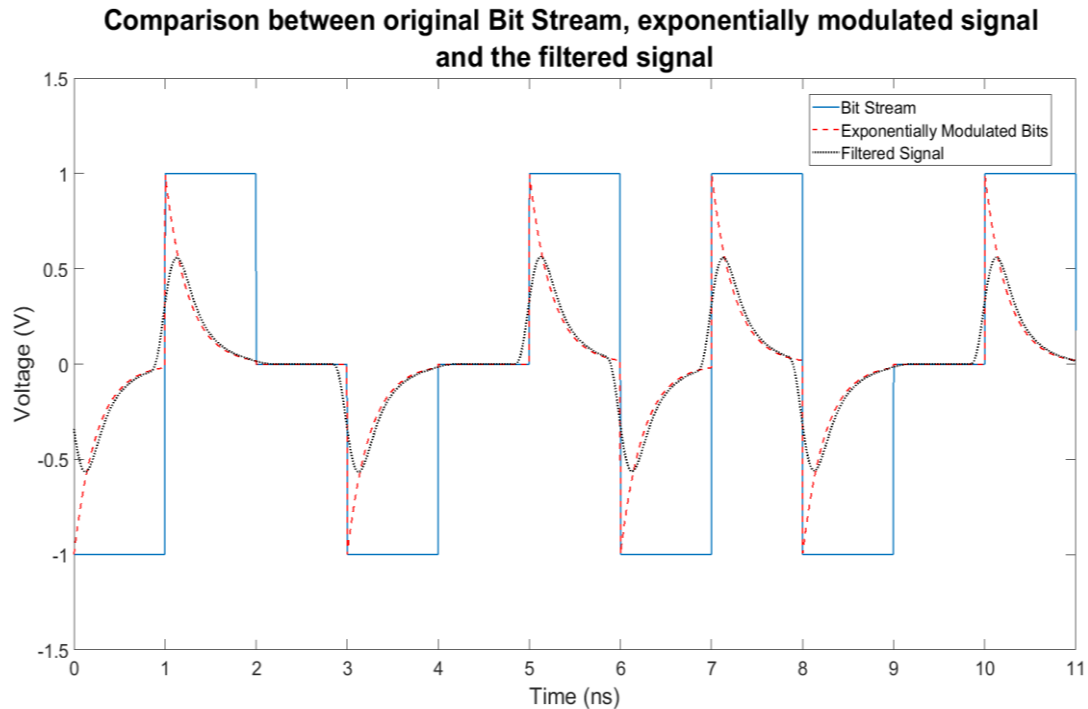


Figure 47.- Comparison between original bits, exponentially modulated signal and the result from filtering the exponential signal.

Moreover, a clock signal is included in each filtered wave to perform the synchronism in the processing. The resulting signal is processed and converted into a .txt file to feed the Arbitrary Waveform Generator (AWG), which permits to pass the total signal (exponential bits and clock) into the VCSEL by means of a bias-tee connector, one at the time.

This extra element receives the data through its radio frequency (RF) input port, while the current is given to its bias (DC) input port. The output from the bias-tee is a combined signal, includes the data (from the AWG) and the current (from the THORLABS LDC) and feeds the VCSEL using the SMA connector in the optical device's pin connected to its anode. The VCSEL's optical output then is received by an optical-electrical converter that is connected to the oscilloscope, in which we can visualize the power component of the signal. The setup used for this experiment is shown below in Figure 48.

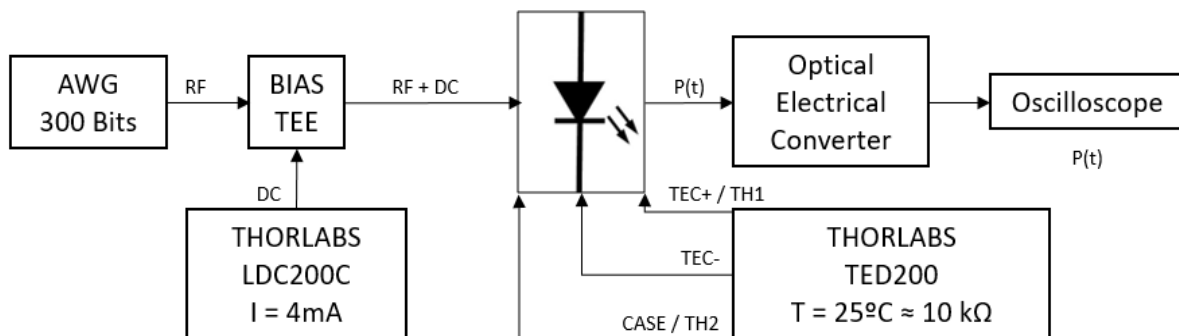


Figure 48.- Experimental setup to obtain the optical power response from the VCSEL.

The received signal [$P(t)$] is recovered by using a MATLAB routine and turned into a readable file by this software, the recovery process is also done individually. However, is necessary to find a correspondence between the obtained data and the set of bits originally generated, as the sample rate in the AWG for the system's input signal is 5 giga samples per second (GS/s), while the oscilloscope is set to 25 GS/s (enough to fulfil Nyquist Theorem), this operation is done with the overall set of signals, the current input and the optical power output, a pair for each combination.

Once all the signals are correlated in the correct numbers of samples, we have the 49 pairs: current [$I(t)$] and optical power [$P(t)$]. These sequences are then united into a single file each, a unique signal containing all forty-nines [$I_i(t)$] called InTotal and a unique signal for the forty-nines [$P_i(t)$] called OutTotal, both for $i = 1, 2, \dots, 49$.

The concatenation for both current and optical power are then used to train a Neural Network in an inverse manner; that is, the current wave is the target while the optical power is the input; as the goal is to predict the characteristics of the input signal to obtain a certain output power. In other words; the objective is to train a NN capable of learning the hidden relationship between the VCSELS mentioned parameters, however it behaves.

Finally, in order to validate the percentage of achievement in this stage, a known pair of signals [$I(t)$ and $P(t)$] are used to test the outputs of the designed NN. By knowing the training is performed backwards, the known [$P(t)$] signal is given as input to the NN, whose result is a signal [$I_{NN}(t)$], that is therefore compared with the original $I(t)$ that was given to the VCSEL experimentally and produces an output optical power [$P(t)$], being this signal the NN's input variable.

3.4. Training a Neural Network to predict a VCSEL's optical power given its corresponding optical phase

For the second experimental part, the use of the Neural Network and the analysis behind it becomes more complex. Just as the previous case the intention is to design, generate and train a NN capable of predicting the VCSEL's optical power given its optical phase.

Based on the mathematical knowledge that relates the optical power and phase (described below in Equation 1) is possible to create a MATLAB routine that given the power signal computes automatically the phase, using an Ordinary Differential Equation (ODE) solver. Previously defined variables: $\Delta\nu(t)$ stands for optical frequency shift, $\varphi(t)$ the instantaneous optical phase, and $P(t)$ is the instantaneous optical power.

Other significant variables are: *transient chirp* (α) and the *adiabatic chirp* (κ), whose values are taken from the characterization done with a similar optical device by José Altabás, and other in 2016; that is 2.24 ± 0.1 and 7.6 ± 0.8 gigahertz per milliwatts (GHz/mW), respectively [1].

$$\Delta\nu(t) = \frac{1}{2\pi} \frac{d\varphi(t)}{dt} = \frac{\alpha}{4\pi} \left(\frac{1}{P(t)} \frac{dP(t)}{dt} + \kappa P(t) \right)$$

Eq. 1

ODEs results are the integral of a differential equation, so if you applied the right side of previous equation to a ODE45 solver, its output signal is directly the optical phase $\varphi(t)$, to which you can later apply a derivative operation and obtain the optical frequency shift for the signal.

As previous stage, a bit stream is created, modulated with the exponential signal defined as: $y = a * e^{-tb}$ and filtered to avoid the software unrealistic transitions between symbols. This signal is given as input to the ODE45 solver due to its abilities of solving most of differential equations.

Once the ODE solver is given a signal output, this is called $\varphi(t)$; or optical phase. At this moment, the learning pair is obtained for each combination from the possible forty-nine permutations mentioned before. Analogically, once the signals $P_i(t)$ and $\varphi_i(t)$ are generated a single signal is formed to be used in the training of the Neural Network.

Is important to remember, that the goal is to produce a successful inverse training, where the optical power signal is the expected target given the optical phase, in order to have the capacity to deduce the necessary power (and therefore current) that the VCSEL needs to obtain a specific optical phase as a result. To achieve this, the Neural Network must analyze and learn inside its layers the correlation between the input and output given to it.

Once the NN is trained, the validation is performed by selecting a specific training set; that is: a single pair of optical phase $[\varphi(t)]$ and optical power $[P(t)]$; let's say for example, the couple: $P_{t_2.5_3.5}$ and $Q_{t_2.5_3.5}$, which means the optical power wave $[P(t)]$ generated with a exponential signal with the coefficients $a=2.5$, $b=3.5$; and the result from the ODE45 solver $[\varphi(t)]$, when this particular input signal is given to it.

Different Neural Networks are trained by varying the values for the hidden layers and delays as well as the training algorithm, in order to compare the obtained results in contrast with the real pair (training set mentioned above); and thus, evaluating the performance of the NNs and which is closest to reality.

The validation is done by giving the original phase signal $[\varphi_{ORIGINAL}(t)]$ to the Neural Network, whose output is a prediction on how the optical power wave should be, named $[P_{NN}(t)]$. This power signal is then applied as input to the ODE45 solver, resulting in a phase signal computed from the estimated power signal, that is $[\varphi_{NN}(t)]$. Finally, two comparison validations are possible:

- For the optical phase signals: $\varphi_{ORIGINAL}(t)$ and $\varphi_{NN}(t)$.
- For the optical power signals: $P_{ORIGINAL}(t)$ and $P_{NN}(t)$.

The obtained results for this contrasts are indicators on how the objectives for this project were achieved, and a measure on how good or bad the NN designed and trained is capable of predicting the VCSELS operation, and are presented in the following chapter.

4. Results

This chapter's presents the results, issues and solutions given through the development of the project, explaining the objectives that were achieved or failed. First section is the VCSEL's characterization where the static parameters of the optical source are measured. Specifically, the dependence of the VCSEL's optical output power (P_{OUT}) and wavelength (λ) is studied vs. the device's temperature and the input current.

Second section, describes the results obtained after training a NN in order to predict the current signal that has to be feed to the VCSEL for obtaining a particular optical power response. To make possible that training, a set of 49 signals is generated. In order to explore the NNs' performances, different values of delay and hidden layers have been considered. Once the NNs are generated, they are validated by using a training pair, and comparing the original signal [$I(t)$] with the estimations given by the networks.

Finally, to generalize the use of the NN-based solution previously proposed, a new training process has been done for estimating the relation between the VCSEL's optical power and the phase resulting for that power. As a result, the NN-based solution can estimate the optical power signal necessary to generate a specific optical phase modulation.

4.1. VCSEL's Characterization

As described in chapter 3, properties of a semiconductor optical source depend foremost on the temperature and current given to the device. For the characterization of the VCSEL, the experimental setup described in Figure 41 allows to obtain the static parameters of the laser, considering the dynamic characteristics to be the same as a similar optical device (commercially available RayCan VCSEL) used by Altabás *et al.*, in a 2016 journal paper [0]; that is: 2.24 ± 0.1 and 7.6 ± 0.8 gigahertz per milliwatts (GHz/mW) for the transient chirp and adiabatic chirp, respectively.

To perform this measurement, the temperature of the VCSEL is set to a single value from 10 to 50 °C and a current swipe is done from 2 to 11 mA. The parameters measured and its units are: wavelength (in nm), VCSEL's TEC current (in mA), maximum power at wavelength (in dBm), OSNR (in dB), total optical power (in dBm) and VCSEL's voltage (in volts). Appendix H includes tables with the gathered values from the experimental characterization procedure.

The obtained lasing threshold for this VCSEL is 1.427 mA at 25°C (temperature in which the electrical and optical characteristics of the VCSEL are given by the manufacturer, as shown in Appendix A). Figure 49 presents the variations on wavelength (λ) in nm, corresponding to different values of current and temperature.

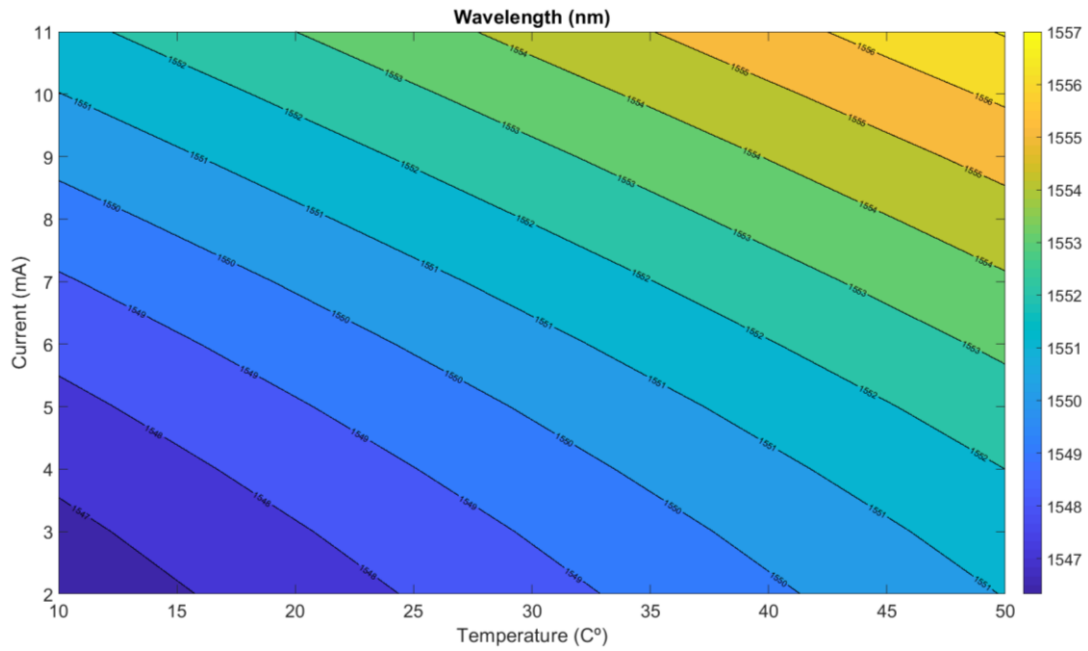


Figure 49.- VCSEL wavelength in terms of the bias current and the temperature

Optical power spectral density of the VCSEL when input current is 7 mA and temperature 15 °C is shown in Figure 50, validating previous results by obtaining a wavelength of 1549.50 nm for that values.

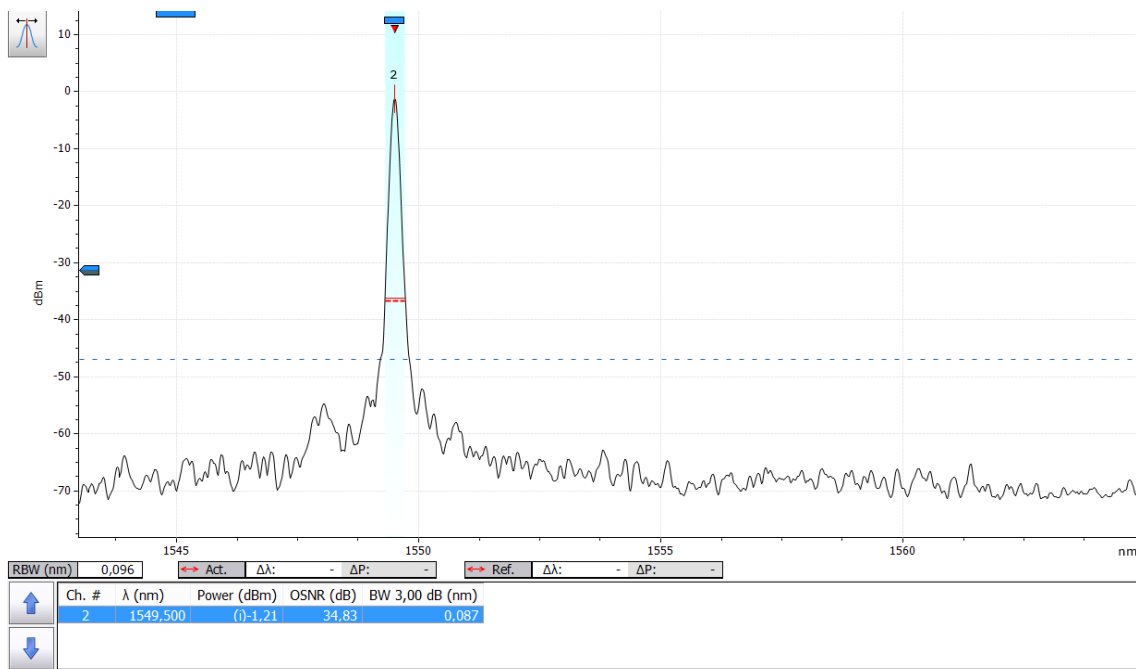


Figure 50.- Optical Spectrum signal from the VCSEL output. $I=7$ mA, $T=15^\circ\text{C}$

4.2. Training a Neural Network to estimate a VCSEL's input current given its output optical power

First experimental stage is to train a NN capable of learning the inverse transfer function between a VCSEL's input current and the optical power, therefore it can predict the required current signal to produce certain optical power pattern.

Initially a set of 300 bits is generated at 1 Gb/s, that is the original current signal $[I(t)]$. Moreover, a set of original signals is created with the 49 possible combinations of the coefficients for the exponential signal (fully described in section 3.3.). The group $I_i(t), \forall i = [1, 2, \dots, 49]$ is the set of signals that will feed the system described in Figure 48.

One by one each signal is feed to the VCSEL using an AWG set at a symbol rate of 5 GS/s, and the optical power is received (individually) at the system's output by an oscilloscope working at a sample rate of 25 GS/s. The set of this pairs $I_i(t) \rightarrow P_i(t), \forall i = [1, 2, \dots, 49]$ becomes the training set for the NNs. However, as the sampling rate and length of the received signal is much higher than the input signal, an extra processing is necessary to find the pattern where the obtained and original signals share the same bit stream. This is crucial step, as the input and output signals have to be correlated to become a useful training set, otherwise the NN is not learning from a connected input-output pair. To achieve this a synchronization function based on cross correlation has been performed. The comparison between one of the signal pairs is shown in Figure 51.

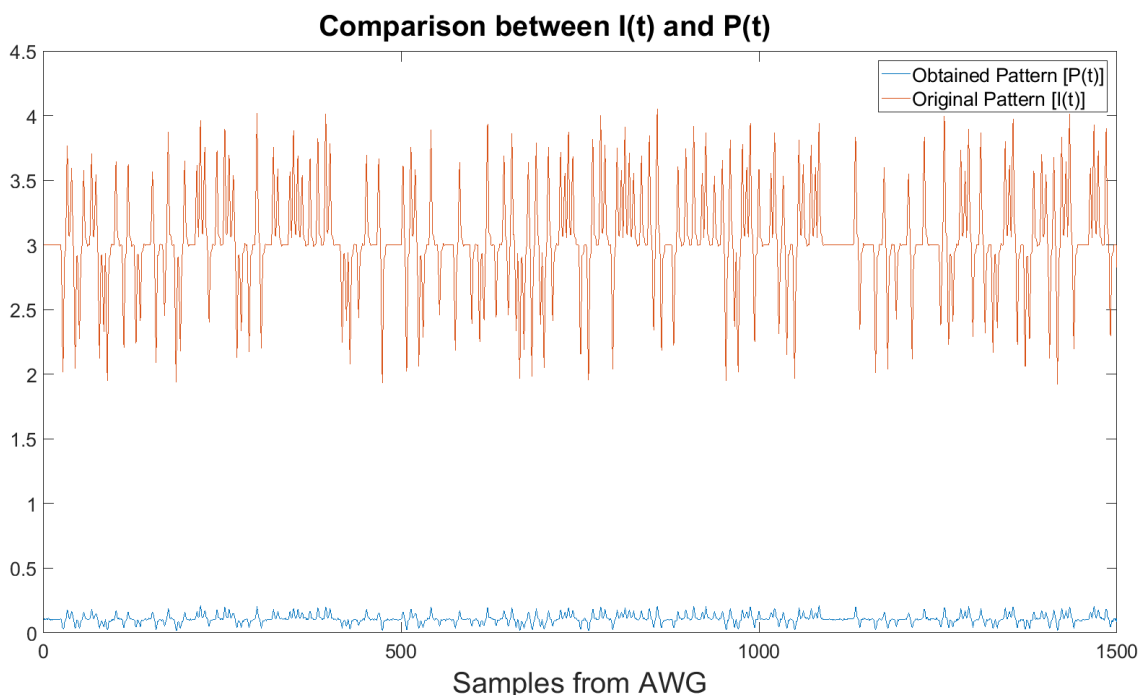


Figure 51.- Comparison between the original pattern $I(t)$ and the received signal $P(t)$

Once all the processing is done, and the signals are correlated this input-output pairs will aid the NN to learn the system's behaviour. In order to train the NN, a unique input and output pair is created, by concatenating the current signals and the power signals, each in a single variable $[I_{TOTAL}(t)$ and $P_{TOTAL}(t)]$

First training of a NN is done by using the LM algorithm, with a delay of 4 and 14 hidden layers. The result plots for the regression shown an unstable training, as this value measures the correlation between the targets and the resulting outputs; higher results indicate a closer relation between the two, which means a relation below 40% for the training, validation and testing process performed by the NN, as shown in Figure 52.

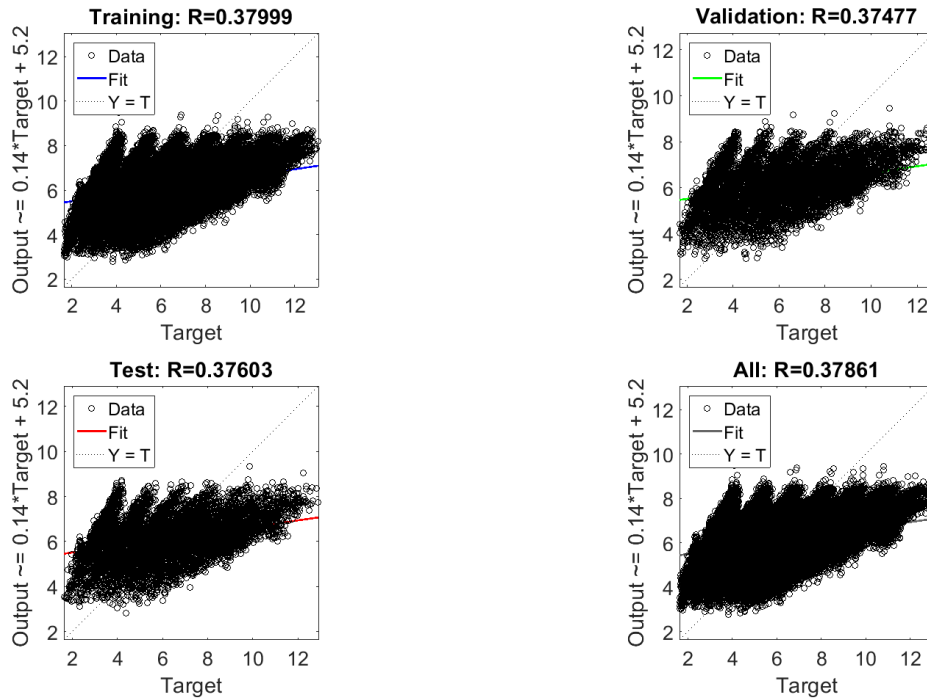


Figure 52.- Regression result for NN training. Using Algorithm LM and training specifications: $D=4$, $HL=15$

It can be deduced that the bad regression could be a result of the training algorithm or the given training specifications. Therefore, some other networks are trained while varying the parameters of the NN. This allows to obtain a global perspective of how the LM algorithm behaves by using the same training set with different inside variables. Table 6 presents the combinations done to train these new networks.

Parameters / NN	NN-A	NN-B	NN-C	NN-D
Delay	4	5	6	10
Hidden Layers	15	20	25	50

Table 6.- Parameters of Delay and Hidden Layers used to train different NNs

Table 7, presents the obtained value for number of iterations, MSE and values of R for the different trainings. Based on the results, the best performance is obtained when using NN-D since it presents the minimum MSE value and the highest regression. Those are the two main parameters that indicate whether a network has been trained good or bad (as explained in section 3.2.2.). Expected values are as close to zero for the MSE (as it measures the error of the training) and as high for the regression (it quantifies the correlation between targets and outputs)

	Best Performance @ Iteration	MSE	Training Regression	Validation Regression	Testing Regression
NN-A	56	4,3399	0,37999	0,37477	0,37603
NN-B	55	4,2726	0,38541	0,37426	0,36946
NN-C	20	4,304	0,38273	0,3822	0,38283
NN-D	28	4,189	0,39125	0,38667	0,03702

Table 7.- Results from the training of NNs with expressed parameters in Table 6

Next validation is to know if the pair signals are not adequate to perform a training, especially when the received signal shows very little power as depicted in Figure 51. To avoid this issue, a scaling of the received signal must be done (again each pair at the time) after the pattern of samples is obtained. The comparison between one original pattern $[I(t)]$ and the resized obtained pattern $[P_{SCALED}(t)]$ is shown in Figure 53, for visual aid only 500 samples instead of the 1500 corresponding to the whole signal are plotted.

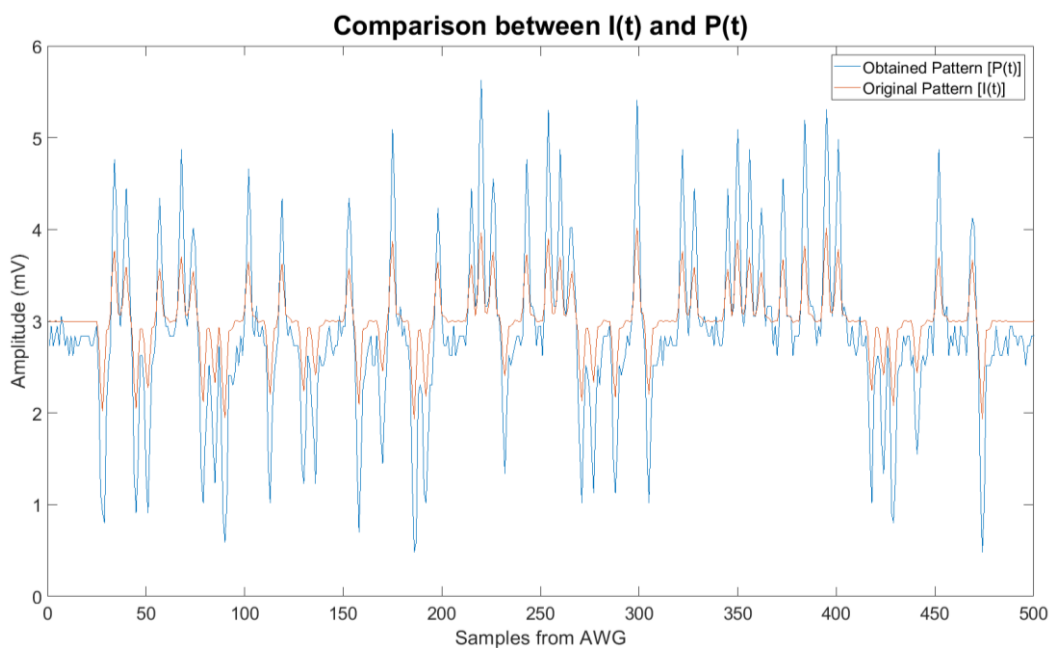


Figure 53.- Comparing the original pattern $I(t)$ and the resized received signal $P_{SCALED}(t)$

The received and scaled signals are now concatenated in a single variable named $[P_{TOTAL_SCALED}(t)]$ and paired with the corresponding current signal $[I_{TOTAL}(t)]$ for training the new networks.

In order to compare with previous networks, the first training is done using the LM algorithm and a delay value of 6 and hidden layers 25 (just as NN-C from previous trainings, that is, without scaling). The results for this particular scaled training (called S-NN-A) produces much better performance of the network, with a MSE of 0.39015 and a regression value of

more than 96% as can be seen in Figure 54 and 55, respectively. Even though the regression outputs are not straight lines, the majority of the samples are positioned on top of the fit line.

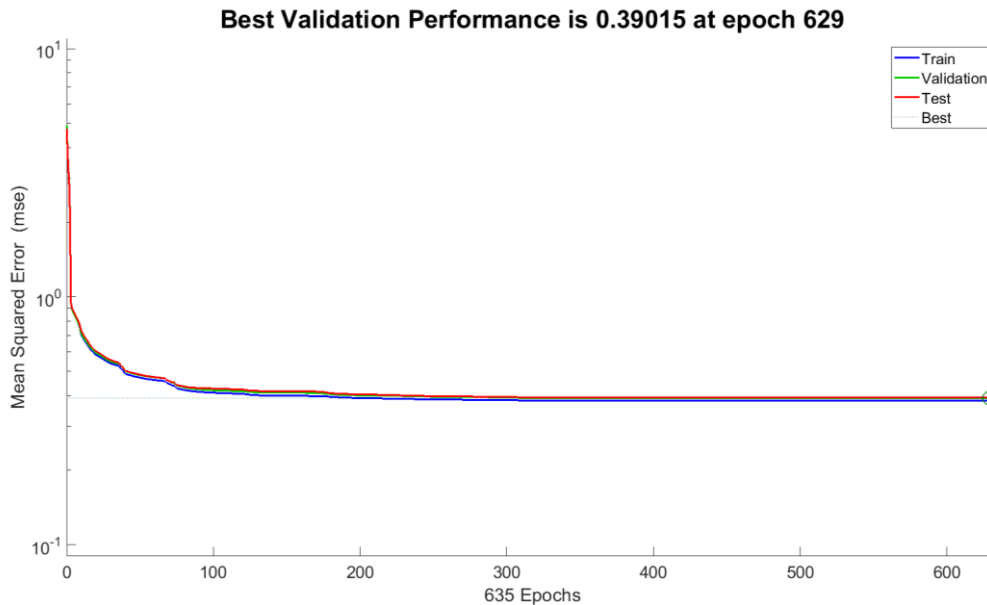


Figure 54.- Performance of NN trained with resized signals. Algorithm LM, D=6, HL=25. (S-NN-A)

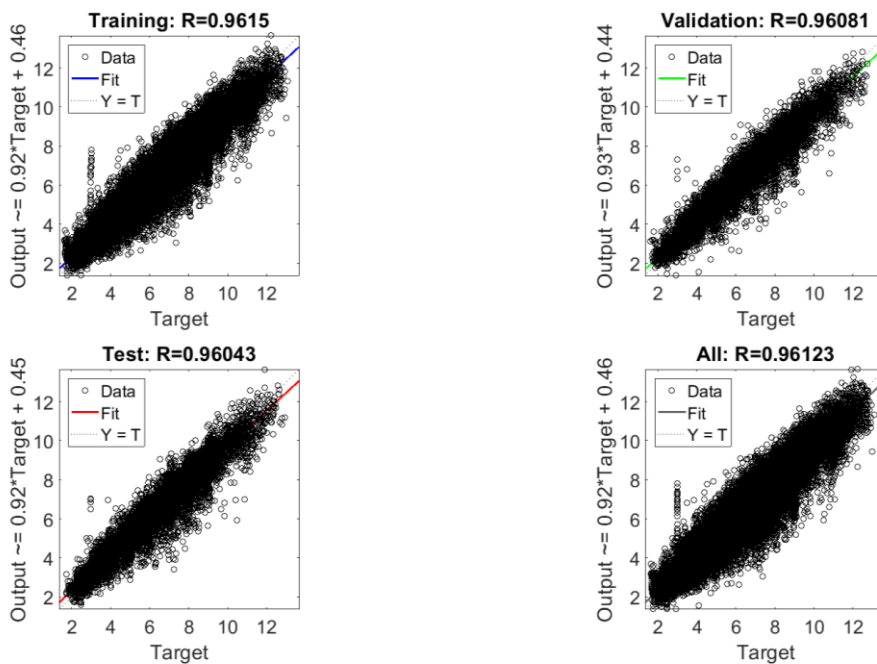


Figure 55.- Regression values for the NN trained with resized signals. Algorithm LM, D=6, HL=25 (S-NN-A)

To better analysis of the training of the NN with the resized signal and its general behaviour, a set of networks is trained with different algorithms and parameters, as described in Table 8.

	Algorithm & Parameters	Best Performance @ iteration	MSE	Training Regression	Validation Regression	Testing Regression
S-NN-A	LM D=6, HL=25	629	0,39015	0,96150	0,96081	0,96043
S-NN-B	BR D=6, HL=25	273	0,3673	0,96297	-	0,96357
S-NN-C	SCG D=6, HL=25	249	0,7130	0,92767	0,92729	0,92783
S-NN-D	LM D=10, HL=50	207	0,22938	0,97863	0,97686	0,9777
S-NN-E	BR D=10, HL=50	1000	0,19496	0,98055	-	0,97801
S-NN-F	SCG D=10, HL=50	391	0,53018	0,94619	0,94670	0,94430

Table 8.- Characteristics and Results for NNs trained using the scaled signals

From the previous results, the algorithm SCG is dismissed, as the MSE is not as closed to the desired value (zero) as the other two algorithms in both training processes. Additionally, is important to mention that the training for S-NN-E (BR algorithm, D=10 and HL=50) did not converge, therefore is not appropriate to say that this particular training found the best performance solution, as the training process stopped due to the fact that 1000 iterations were done.

This leaves us with 3 trained networks to perform the final validation for this experimental stage. By looking at the results, is easy to think that the best trained network is S-NN-D, as it achieves the lower MSE value and higher regressions. Nevertheless, for deciding the best type of NN, a mathematical comparison based on calculating the mean squared error between the input signal and each of the estimations obtained from the different networks must be performed. The computed error between a pattern and its estimation is calculated using the MSE formula:

$$\frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - \tilde{x}_i)^2} \quad \text{Eq. 9}$$

To measure the quality of the training of each network, a known training set is applied to the NNs, in the same manner as the training was performed; that is, backwards. An input signal

$[I(t)]$ from the system and the scaled received pattern $[P_{SCALED}(t)]$ associated to that input must be used. Then, the scaled optical power signal is given as input to the selected networks ($S-NN-A$, $S-NN-B$ and $S-NN-D$) and each one will be the output of its own estimation of the pattern $[I(t)]$; that is: $I_{S-NN-A}(t)$, $I_{S-NN-B}(t)$, $I_{S-NN-D}(t)$.

Figure 56 shows the comparison between the four signals in a 50 ns time window to appreciate the transitions the signals suffers as well as how close the predictions from the NNs follow the original signal.

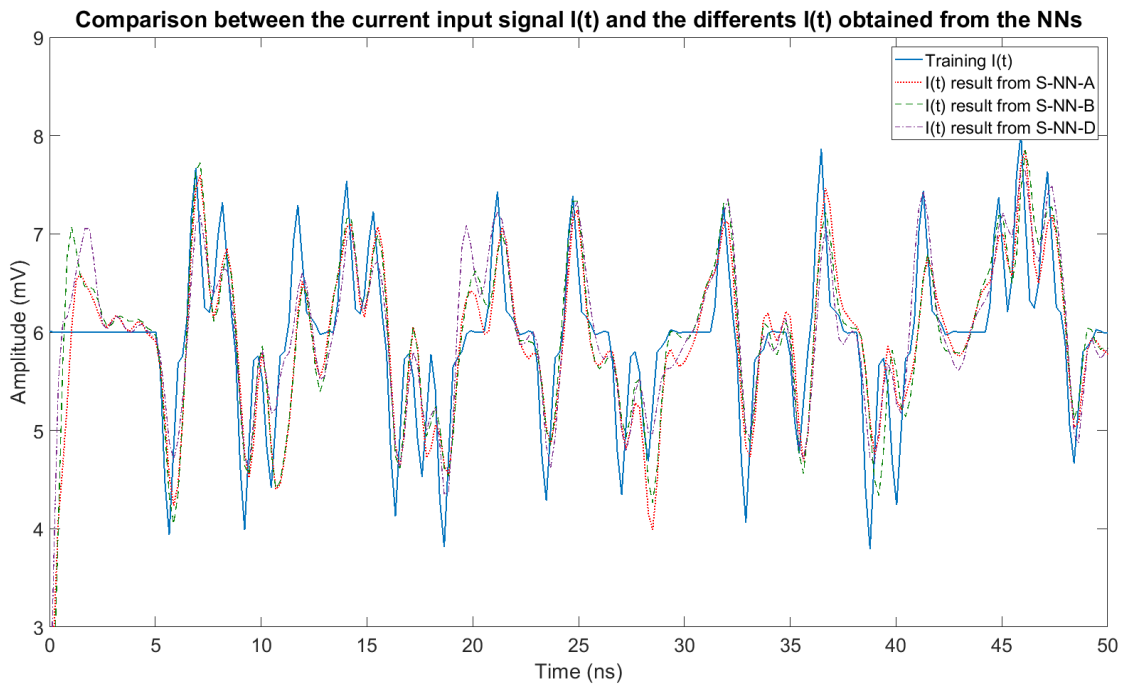


Figure 56.- Comparison between the $I_{S-NN}(t)$ and the original $I(t)$

Moreover, Figure 57 shows the error measurement for each prediction in contrast with the original, signal. Based on the numerical values, is possible to affirm that the Neural Network $S-NN-D$ does have a better performance that the other ones evaluated with it, as its error measurement is the lowest.

$$\text{Errors} = \begin{matrix} 8.6350 & 8.6668 & 8.6216 \end{matrix}$$

Figure 57.- Result of computing Mean Squared Error (MSE) between original signal $I(t)$ and the predictions: $I_{S-NN-A}(t)$, $I_{S-NN-B}(t)$ and $I_{S-NN-D}(t)$, respectively

4.3. Training a Neural Network to predict a VCSEL's optical power given its corresponding optical phase

Final stage for the experimental phase of the project is to create a NN such that is capable of learning the intrinsic relation for the VCSELs dynamic parameters, as is explained

in section 3.3., specially it should be capable to resolve the complications that can occur while studying the device due to its nonlinear nature.

As explained by Agrawal [8, pp. 98] the nonlinearity characteristic for semiconductors forces to solve the rate equations numerically. This section describes the results obtained by training and comparing different NNs in a similar manner as previous section.

Initially, an exponential signal is created to obtain a certain VCSEL's optical power $[P(t)]$. This signal is filtered to avoid unrealistic transitions between bit's changes and then is given as input to an Ordinary Differential Equation (ODE) solver routine in MATLAB. The result of this ODE represents the optical phase of the VCSEL given the optical power as input, following the mathematical expression described in Equation 1. It is important to mention, that this experiment only processed 25 bits at 1 Gb/s in 6001 samples, as the ODE solver would not converge for larger bit patterns and showed better results for higher sampling.

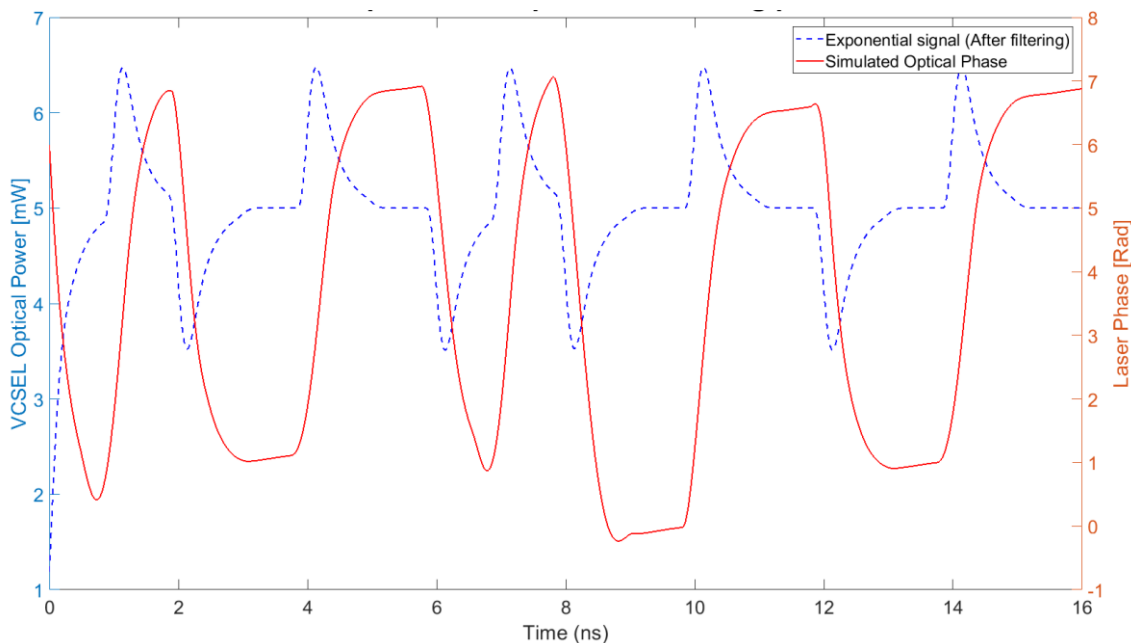


Figure 58.- An optical power signal $P(t)$ compared against the corresponding optical phase $Q(t)$

Compared to Figure 4, Figure 58 follows the experimental optical phase obtained by Altabás *et al.*, where a sharper transition occurs for positive pulses. However, the simulated optical phase presents a certain delay, probably due to the initiation of the ODE routine.

Once the result of the ODE is validated as a correct result for the differential equation, the training set is generated, by applying the 49 combinations of $[P(t)]$ generated by modifying the coefficients of the exponential equation: $y = a * e^{-tb}$. Therefore, a training set of 49 input-output pairs are gathered, being the signals in the form of: $P_i(t) \rightarrow \varphi_i(t), \forall i = [1, 2, \dots, 49]$. As previous case, a single vector for inputs and outputs containing the 49 combinations is necessary to train the network, which leads to the overall training set: $P_{TOTAL}(t)$ and $\varphi_{TOTAL}(t)$.

Since the objective of this stage is to find the inverse correspondence between the VCSEL's optical power and phase, it is necessary to determine if the relation between the optical power the phase can be estimated by the Neural Network. Therefore, the NN has to

learn the behaviour of the dynamic parameters of the device. In this case, the training of the NNs (following a similar manner as previous experiment) is done backwards; that is giving the optical phase signal $[\varphi_{TOTAL}(t)]$ as the networks input and selecting as targets the optical power function $[P_{TOTAL}(t)]$.

For this experiment the algorithm SCG was not used, not only due to its poor performance in previous experiment in contrast with the other algorithms, but also because training process would be invalid most of the tried times as it was attempted. An explanation of that could be the required amount of training samples, that rise up to 294,049 (49 combinations of 6001 sample each). Table 9 displays the characteristics of the performed training processes.

	Used Algorithm	Delays	Hidden Layers	Iterations
NNa	LM	4	15	960
NNb	BR	4	15	1000
NNc	LM	5	20	443
NNd	BR	5	20	1000
NNe	LM	6	25	286
NNf	BR	6	25	1000
NNg	LM	10	50	411

Table 9.- Algorithms and Parameters used to trained the following NNs

Once more, the BR training algorithm has failed to find a reasonable result, that is the main reason why the training for parameters $D=10$ and $HL=50$ was only done with the algorithm LM. After the networks are trained, is time to validate the accuracy of them. For this final results, a known training set of single input-output pair is used: $P_{1.5_2.5}$ and $Q_{1.5_2.5}$, called $P_{ORIGINAL}(t)$ and $\varphi_{ORIGINAL}(t)$.

The known $\varphi_{ORIGINAL}(t)$ signal is given as input for all the trained NNs, obtaining an estimation of the optical power signal $[P_{NNx}(t)]$ for each trained network. Figure 59 shows the comparison of the original signal and the corresponding predictions.

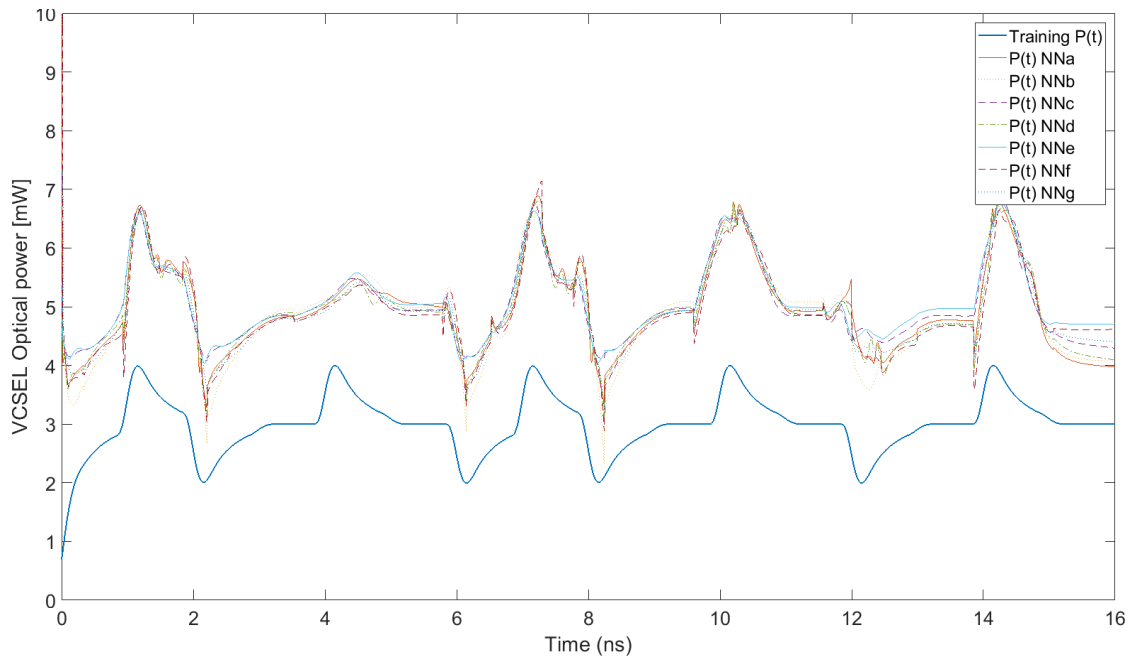


Figure 59.- Comparison between original $P(t)$ and $P_{NNx}(t)$ from NNs

Visually, the difference between the original $P(t)$ and the predictions performed by the NNs is negligible, however the MSE is mathematically computed using the previously defined Equation 9, the numerical results for each MSE calculation is given by Figure 60.

ErrorP =

5.7668 5.7352 5.8311 5.7530 5.8835 5.7509 5.7711

Figure 60.- Result of computing Mean Squared Error (MSE) between original signal $P(t)$ and the predictions: $P_{NNa}(t)$, $P_{NNb}(t)$, $P_{NNc}(t)$, $P_{NNd}(t)$, $P_{NNe}(t)$, $P_{NNf}(t)$, and $P_{NNg}(t)$

Measured MSE proves that the closest estimation to the original signal is the output given by trained set NNb with characteristics $D=4$ and $HL=15$; even though the training stopped not because the MSE value reached its minimum while performing the process but because the maximum number of iterations was reached.

All obtained $[P_{NNx}(t)]$ are then applied to the ODE45 solver, to contrast the original optical phase $\varphi_{ORIGINAL}(t)$ with the results from the differential equation for each estimated optical power signal $P_{NNx}(t)$, the relation is depicted in Figure 61.

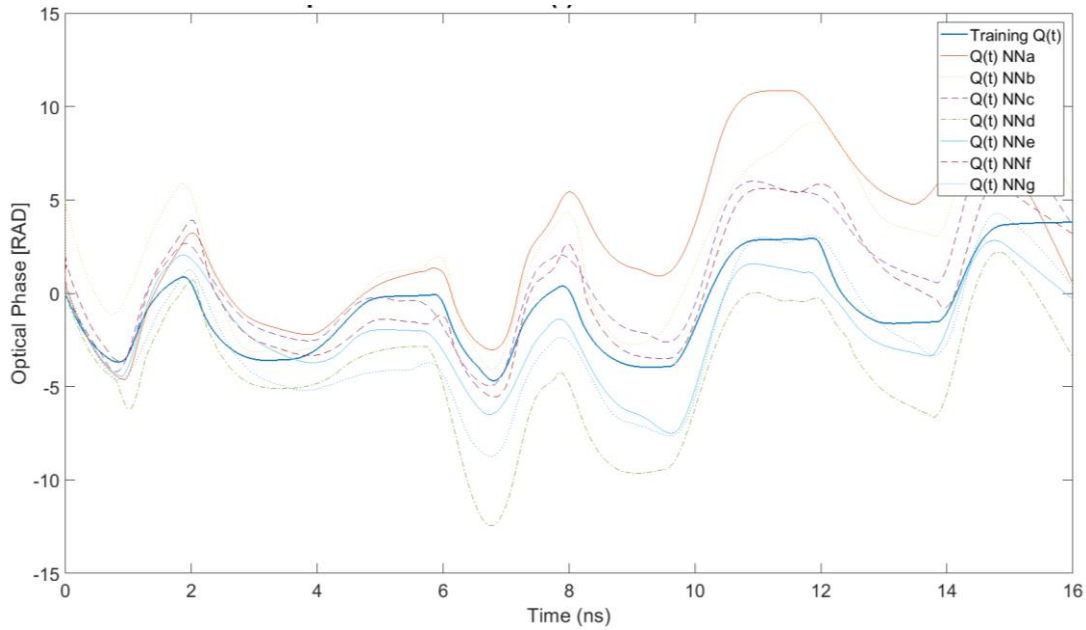


Figure 61.- Comparison between original $Q(t)$ and computed $Q_{NNx}(t)$ from ODE45

The MSE between these new patterns and the original signal $\varphi_{ORIGINAL}(t)$ is calculated once more by Equation 9, while the results are given later on in Figure 62.

ErrorQ =

5.1689 4.7343 3.7296 5.7415 3.7353 3.7278 4.5187

Figure 62.- Results for Mean Squared Error (MSE) between original signal $Q(t)$ and computed $Q_{NNx}(t)$ from ODE45

The closest relation between the original optical phase signal and the results of the ODE solver when inputs are the optical power estimations is NNf, the second closest estimation for the original $P(t)$; this inconsistency is due to the fact that the ODE45 is using estimation signals as inputs which carry errors from the corresponding NN.

5. Budget

This project was performed in the Optics Laboratory in UPC's Department of Signal Theory and Communications (TSC) and therefore most of the necessary elements are available for student's use.

However, delivering a budget for the project is a requirement for the Master's Thesis. Hence, the available devices given by the department are only mentioned and the quantification is done directly for the resources that had to be bought, fabricated or invested in order to fulfil the project.

5.1. Available Resources

Below are listed all the technical devices within the TSC Optical Laboratory in order to perform any experiment described in this memory. Additionally, a brief description of its specifications is provided:

- TEKTRONIX Mixed Signal Oscilloscope (MSO 70804C) with four (4) channels, and a maximum bandwidth of 8 GHz, typical rise time of 50 ps (picoseconds) and a sample rate up to 25 Giga samples per second (GS/s), capable of recording up to 250 million samples at once.
- TEKTRONIX Arbitrary Waveform Generator (AWG 70002A) is able to provide an analog bandwidth of 14 GHz and an output frequency of 20 GHz with a sample rate that can achieve up to 50 GS/s and a record length of 2 giga samples (GS) and an optical feature of including up to 16 GS.
- THORLABS Laser Diode Controller (LDC 200CV) provides the bias current and protects the VCSEL from high currents above the established limit. Is capable of giving up to 20 mA with a precision of 0.001 mA, allowing fine stability and control for a delicate device such as a VCSEL. Appendix F shows its front and rear connections and controls.
- THORLABS Thermoelectric Temperature Controller (TED 200) is able to measure devices with internal thermistors (20 and 200 kOhms) and TECs with maximum current up to 2 amperes. It includes a 5-digit display, up to three (3) decimals in the 200 kΩ range. Appendix G pictures front and back connections and controls.
- EXFO Optical Spectrum Analyzer FTB-2, capable of receiving wavelengths between 1255 nm and 1650 nm in three (3) different reception modes: single, real time (continuously) and average (by a count number selected by the user). The received signal can be zoomed in with vertical and horizontal markers to locate specific points.
- GO4FIBER Power meter GFHS-A, with six (6) calibrated wavelengths and an accuracy of $\pm 5\%$, capable of measuring up to +10 dBm and scales of dB and mW.
- TENMA Digital multimeter 72-7780, used to measure voltage in the VCSEL with a 0.001 volts precision and capable of quantifying 600 volts.

5.2. Components List and Costs

This section includes of the components and parts that are used to fabricate the prototype (VCSEL transmitter) as well as needed extra elements to perform the experimental analysis.

Component	Quantity	UOM*	Unit Cost	Extended Cost
SM1550 VCSEL pigtail, 10G	1	each	€ 200,00	€ 200,00
9-pin D-SUB jack	1	each	€ 0,98	€ 0,98
15-pin D-SUB jack	1	each	€ 1,19	€ 1,19
VCSEL board	1	each	€ 31,33	€ 31,33
Cable	3	mts	€ 0,75	€ 0,75
Total				€ 234,25

* UOM (Unit of Measurement)

Table 10.- Components Lists and Costs

In addition to the available resources and the extra components, is unavoidable to take into account the time invested into the following stages:

- Studying of the theoretical background.
- Developing a project plan.
- Design of the board prototype.
- Characterization of the VCSEL
- Experimental testing.

The amount of time invested into this project is around eight-hundred hours. Normally, an hour for a grade student is paid around €9 per hour, therefore:

$$Cost_{person} = hours_{person} * total_{hours}$$

$$Cost_{person} \approx 800 * € 9$$

$$Cost_{person} \approx € 7.200$$

Taking all factors into account, the overall cost for this project is around $Cost_{TOTAL} \approx € 7434.25$. Relying of the obtained results for the experimental stages, is expected that the research described in this project is a feasible project to continue, in order to achieve more robust results, as is suggested in next chapter.

6. Conclusions and Future Development

The main objective for this Master's Thesis was to develop a Neural Network training, capable of learning the inverse transfer function between the input current signal and the optical phase that corresponds to it for a semiconductor laser diode, in this case: a VCSEL. Principal reason for doing so, is the possibility of future uses of VCSELs as a cost-effective transmitter due to its many benefits, such as: fast and reliable performance, low cost and the possibility to be directly modulated.

The development of the project was structures in different stages. First, the characterization of the available optical source, a RayCan 1550nm SM-VCSEL with maximum transmission velocity of 10 Gb/s in order to find the dependence of the static parameters of the VCSEL with respect to the current and temperature. Measured values are: lasing threshold, wavelength, TEC current, optical power and voltage. The dynamic characteristics, are supposed to be the same as a similar VCSEL (also from RayCan) characterized by Altabás *et al.*, that is: 2.24 ± 0.1 for the transient chirp and 7.6 ± 0.8 gigahertz per milliwatts (GHz/mW) for adiabatic chirp.

For the first experiment involving Neural Networks, is focused on analyzing the correlation between the input current and the output optical power. To do that, it was used the layout pictured in Figure 48. Initially a set of bits modulated exponentially at 1 Gb/s was generated and feed to the VCSEL using an AWG at 5 GS/s and recovered by an oscilloscope sampling at 25 GS/s. A cross correlation between input and received signals we performed to get the same bit pattern in both as this correlated pair is then used to train the first NN.

Since our goal is to predict input signals for an expected output, the training of the network is done inversely, such that the optical power is the input given to the NN to learn the input current signal that generates the optical power pattern. This training performed poorly as proven by numerical results for regression (R) and mean squared error (MSE), both key parameters on NN's training process.

New networks were created to validate that the parameters given to the training process or the algorithm used were correct, however new results didn't improve either MSE or R values. To take further action, a scaling for the recovered signal from the oscilloscope was done as its amplitude was rather low which may cause the failing of the training while trying to create the correct association between input-output learning patterns. After a scaled training set is produced, six (6) new NNs are trained by using the original signals and its corresponding scales output.

After the training is done the NNs that obtained badly performance values (MSE and R) are discarded, while the rest are used to estimate a known current signal by giving the optical power pattern, and then comparing the results and computing the MSE between the original signal and the estimations. Performance results for the training of the different networks was numerically proven when it obtained the minor value for the MSE calculation.

Finally, taking into account that dynamic parameters for a semiconductor are difficult to measure directly due to its non-linear characteristics but can be numerically obtained by

differentially solving the device's rate equation; the last experiment aimed to train a Neural Network capable of learning the VCSELs behaviour in order to predict an ideal optical power signal that the device need to receive in order to emit a specific optical phase. This means, performing an inverse training process, where the input is the optical phase and the desired targets the optical power signal.

Initial step to achieve this is to generate an input-output training set. For this section an exponential signal representing 25 bits of optical power at 1 Gb/s was generated and processed using an ODE solver MATLAB's routine. The ODE solver output is the corresponding optical phase for given input, therefore a training set is generated. One different combinations are created (to give enough samples for the training process), different trainings are done with a same data set and different training variables.

Once the training is over, a known pair is used to validate the performance of each trained network: that is, giving the optical phase as input for the different trained NNs and recovering estimations on an ideal optical power pattern, in order to compare the results and conclude on the best training process. Moreover, a MSE is calculated between the original $[P(t)]$ used for the training and the different NNs results to obtain a quantification on the best training performance at the minimum MSE value.

Additionally, the results from the NNs are applied to the ODE45 solver to contrast the fidelity of the calculated optical phase signals from an estimated optical power. Similar to previous training, a MSE is computed. However, the best result for this part does not match the best performance for the MSE of $[P(t)]$, as the ODEs are not computing from an actual optical power signal, but from an estimation.

Based on the obtained results and having learned the strengths and drawbacks for different training algorithms and training variables, is suggested to continue this research by:

- Applying more data rate for the generated signals in both cases, specially taking into account that the used for this project can transmit up to 10 Gb/s.
- More powerful equipment in terms of processing capacity can help in order to train new NNs more robust that can minimize the training time (some of the NNs, especially BR algorithm took up to two hours to train).
- Validating optical phase correlation with optical power by taking the simulations into experimental field, where a high-resolution complex optical spectrum analyzer (HRCOSA) to capture the phase of the VCSEL and compare it with the simulated result from the ODE.

Bibliography

- [1] J. Altabás, D. Izquierdo, J. Lázaro, and I. Garces. “Chirp-based Direct Phase Modulation of VCSELs for Cost-effective Transceivers”. *Optics Letters*, vol. 42, no. 3, pp. 583-586, 2017. DOI: 10.1364/OL.42.000583.
- [2] S. Sarmiento, S. Spadaro and J. Lázaro. “Cost-effective ROADM design to maximize the Traffic Load Capacity of u-DWDM coherent metro-access networks”. *Optical Switching and Networking*, vol. 30, pp. 53-61, 2018. DOI: 10.1364/JOCN.9.001116
- [3] S. Russell. “A practical device to simulate the working of nervous discharges”. *Journal of Animal Behavior*, vol. 3, no. 1, pp. 15-35, January 1913. DOI: 10.1037/h0070584.
- [4] P. Viñuela, I. Galván. *Redes de Neuronas Artificiales. Un Enfoque Práctico*, 1st ed. Madrid, España: Pearson Educación, S.A., 2004. ISBN: 84-205-4025-0.
- [5] G. Holzmann, B. Pehrson. *The Early History of Data Networks*, 1st ed. Hoboken, USA: John Wiley & Sons, Inc., 2003. ISBN: 978-0-818-66782-4.
- [6] Agrawal. *Lightwave Technology: Telecommunication Systems*, 1st ed. Hoboken, USA: John Wiley & Sons, Inc., 2005. ISBN: 978-0-471-21572-1.
- [7] The Nobel Prize. “Charles K. Kao - Facts”. *Nobel Media AB 2019*. [Online] Available: <<https://www.nobelprize.org/prizes/physics/2009/kao/facts/>> [Accessed: January 25th, 2019]
- [8] Agrawal. *Fiber-Optic Communication Systems*, 4th ed. Hoboken, USA: John Wiley & Sons, Inc., 2010. ISBN 978-0-470-50511-3.
- [9] J. Yamada, S. Machida, T. Kimura. “2 Gbit/s Optical Transmission Experiments at 1.3 μ m with 44 km Single-mode Fiber”. *Electronics Letters*, vol. 17, no. 13, pp. 479-480, June 1981. Print ISSN: 0013-5194. DOI: 10.1049/el:19810334.
- [10] A. Gnauck, B. Kasper, R. Linke, R. Dawson, and others. “4-Gbit/s Transmission over 103 km of Optical Fiber using a novel electronic Multiplexer/Demultiplexer”. *Journal of Lightwave Technology*, vol. 3, no. 5, pp. 1032-1035, October 1985. Electronic ISSN: 1558-2213. DOI: 10.1109/JLT.1985.1074318.
- [11] T. Otani, K. Goto, H. Abe, M. Tanaka, and others. “5.3 Gbit/s 11300 km Data Transmission using actual Submarine cables and repeaters”. *Electronics Letters*, vol. 31, no. 5, pp. 380-381, March 1995. Print ISSN: 0013-5194. DOI: 10.1049/el:19950271.
- [12] T. Welsh, R. Smith, H. Azami, R. Chrisner. “The FLAG Cable System”. *IEEE Communications Magazine*, vol. 34, no. 2, pp. 30-35, February 1996. Electronic ISSN: 1558-1896. DOI: 10.1109/35.481241.
- [13] TeleGeography, Authoritative Telecom Data. “Submarine Cable 101”, 2018. [Online] Available: <<https://www2.telegeography.com/submarine-cable-fags-frequently-asked-questions>> [Accessed: January 17th, 2019]

- [14] V. Coffey. "Sea change: The Challenges Facing Submarine Optical communications". *Optics & Photonics News*, 2014. [Online] Available: <https://www.osa-opn.org/home/articles/volume_25/march_2014/features/sea_change_the_challenges_facing_submarine_optical/> [Accessed: January 17th, 2019]
- [15] P. Winsor, R. Essiambre. "Advanced Modulation Formats for High-Capacity Optical Transport Networks". *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4711-4728, December 2006. Electronic ISSN: 1558-2213. DOI: 10.1109/JLT.2006.885260.
- [16] X. Zhou, J. Yu, M. Huang, Y. Shao, and others. "64-Tb/s (640x107-Gb/s) PDM-36QAM Transmission over 320km Using Both Pre- and Post-Transmission Digital Equalization". *In Proceedings of the Optical Fiber Communication Conference*, 21-25 March 2010, San Diego, USA. Electronic ISBN: 978-1-55752-884-1.
- [17] A. Larsen. "LASER vs. LED: What's the Difference". *Acupuncture Technology News*, 2014. [Online] Available: <<https://www.miridiatech.com/news/2014/02/laser-vs-led-whats-the-difference/#comments>> [Accessed: January 18th, 2019]
- [18] Engineering and Technology History Wiki. "Laser". 2017. [Online] Available: <<https://ethw.org/Laser>> [Accessed: January 22nd, 2019]
- [19] Wikipedia, the free encyclopedia. "Ruby Laser". 2018. [Online] Available: <https://en.wikipedia.org/wiki/Ruby_laser> [Accessed: January 22nd, 2019]
- [20] E. Hardy. "Optical feedback in lasers". *My Laser Spectrum*, 2016. [Online] Available: <<http://mylaserspectrum.com/2016/09/optical-feedback-lasers/>> [Accessed: January 25th, 2019]
- [21] M. Fukuda. *Optical Semiconductors Devices*, 1st ed. USA: John Wiley & Sons, Inc., 1999. Print ISBN: 0-471-14959-4.
- [22] G. Agrawal, J. Bowers, C. Chang-Hasnain, and others. *Semiconductors Lasers. Past, Present, and Future*, 1st ed. Woodbury, USA: American Institute of Physics Press, 1995. Print ISBN 1-56396-21-X.
- [23] J. Buus, M. Amann, D. Blumenthal. *Tunable Laser Diodes and Related Optical Sources*, 2nd ed. Hoboken, USA: John Wiley & Sons, Inc., 2005. Print ISBN: 0-471-20816-7.
- [24] G. Evans, J. Hammer. *Surface Emitting Semiconductor Lasers and Arrays*, 1st ed. San Diego, USA: Academic Press, Inc., 1993. Print ISBN: 0-12-244070-6.
- [25] H. Li, P. Wolf, P. Moser, and others "Vertical-Cavity Surface-Emitting Lasers for Optical Interconnects". *SPIE Newsroom*, November 2014. DOI: 10.1117/2.1201411.005689.
- [26] K. Iga. "Surface-emitting Laser – Its Birthday and Generation of New Optoelectronics Fields". *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 6, no. 6, pp. 1201-1215, November/December 2000. DOI: 10.1109/2944.902168.
- [27] Patterson, A. Gibson. *Deep Learning*, 1st ed. Sebastopol, USA: O'Reilly Media, Inc., 2017. ISBN: 978-1-491-91425-0.

- [28] J. Brownlee. "Linear Regression for Machine Learning". *Machine Learning Mastery - Understand Machine Learning Algorithms*, 2016. [Online] Available: <<https://machinelearningmastery.com/linear-regression-for-machine-learning/>> [Accessed: January 27th, 2019]
- [29] D. Lane, D. Scott, M. Helb, R. Guerra, and others. "Introduction to Statistics". *Online Statistics Education: An Interactive Multimedia Course of Study*. [Online] Available: <http://onlinestatbook.com/Online_Statistics_Education.pdf> [Accessed: January 29th, 2019]
- [30] S. Shalev-Shwartz, S. Ben-David. *Understanding Machine Learning. From Theory to Algorithms*, 1st ed. New York, USA: Cambridge University Press, 2014. ISBN 978-1-107-05713-5.
- [31] D. Debarko. "RNN or Recurrent Neural Network for Noobs". *Hackernoon - A Medium Corporation*, 2018. [Online] Available: <<https://hackernoon.com/rnn-or-recurrent-neural-network-for-noobs-a9afbb00e860>> [Accessed: January 22nd, 2019]
- [32] C. Stergiou, D. Siganos. "Neural Networks". *Surprise 96 Journal - Imperial College of Science, Technology and Medicine*, 1996. [Online] Available: <https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Introduction%20to%20neural%20networks> [Accessed: January 26th, 2019]
- [33] Mathworks. "Deep Learning Onramp". *MATLAB and Simulink Training*, 2017. [Online] Available: <<https://matlabacademy.mathworks.com/R2017b/portal.html?course=deeplearning#chapter=1&lesson=2§ion=1>> [Accessed: September 18th, 2018]
- [34] Azrael. "Recurrent vs. Recursive Neural Networks: Which is better for NLP?". *Cross Validated - Stack Exchange Network*, 2015. [Online] Available: <<https://stats.stackexchange.com/questions/153599/recurrent-vs-recursive-neural-networks-which-is-better-for-nlp>> [Accessed: January 28th, 2019]
- [35] RayCan. "1550nm Vertical-Cavity Surface-Emitting Laser", RC34xxx1-Ft datasheet, July 2015 [Revision 1.0]
- [36] Thorlabs. "Laser Diode Controller", LDC200C series operation manual, September 2017 [Revision 3.5]
- [37] Thorlabs. "Thermoelectric Temperature Controller", TED200C operation manual, October 2017 [Revision 3.5]
- [38] Mathworks. "Deep Learning Toolbox". [Online] Available <<https://es.mathworks.com/products/deep-learning.html>> [Accessed: February 3rd, 2019]
- [39] Used-Line. "Keysight Technologies (Agilent HP) 81640A for Sale". Used & New Hi-Tech & Scientific Equipment [Online] Available <<https://www.used-line.com/optical/light-sources/agilent-hp-81640a/item-11991237>> [Accessed: May 15th, 2019]

Appendices

Appendix A – VCSEL’s Characteristics

Electrical and optical characteristics

(T = 25°C unless otherwise stated)

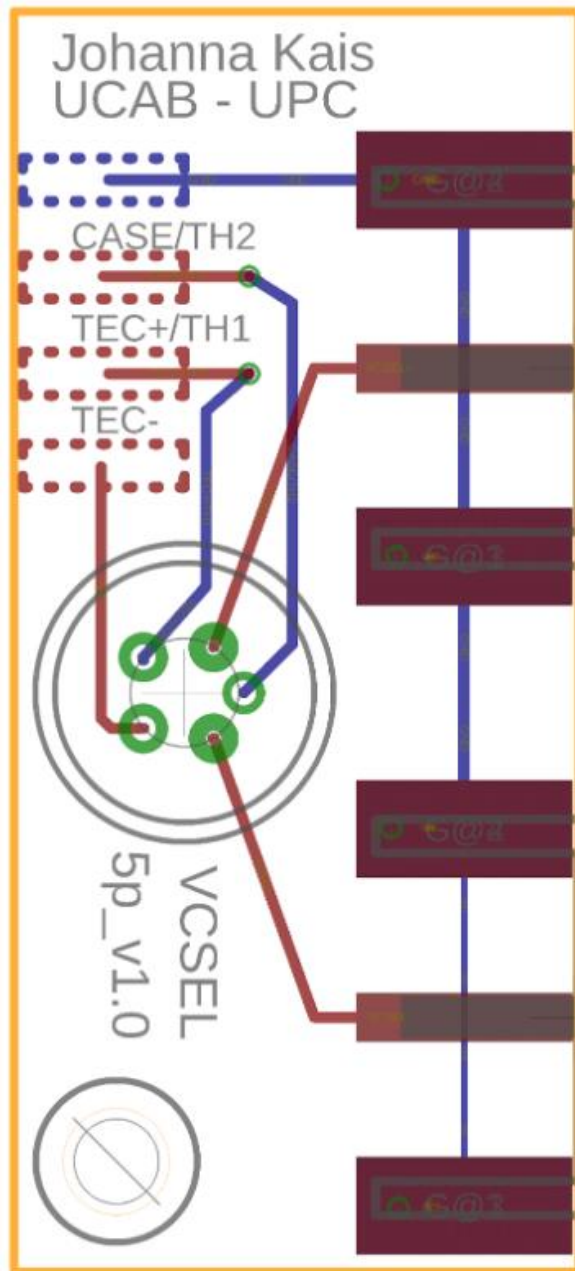
Parameter	Symbol	Min.	Typ.	Max.	Unit	Notes
Threshold current	I_{th}		2	4	<i>mA</i>	
Forward voltage	V_f			3	<i>V</i>	
Series resistance	R_s		100	200	Ω	
Output power	P_o	0.4	0.5		<i>mW</i>	
Wavelength	λ	1530	1550	1570	<i>nm</i>	
Side mode suppression	<i>SMSR</i>	30	35		<i>dB</i>	
Rise and fall time	t_r t_f		~ 60 ~ 60		<i>psec</i>	(20%-80%)

Absolute maximum ratings

(T = 25°C unless otherwise stated)

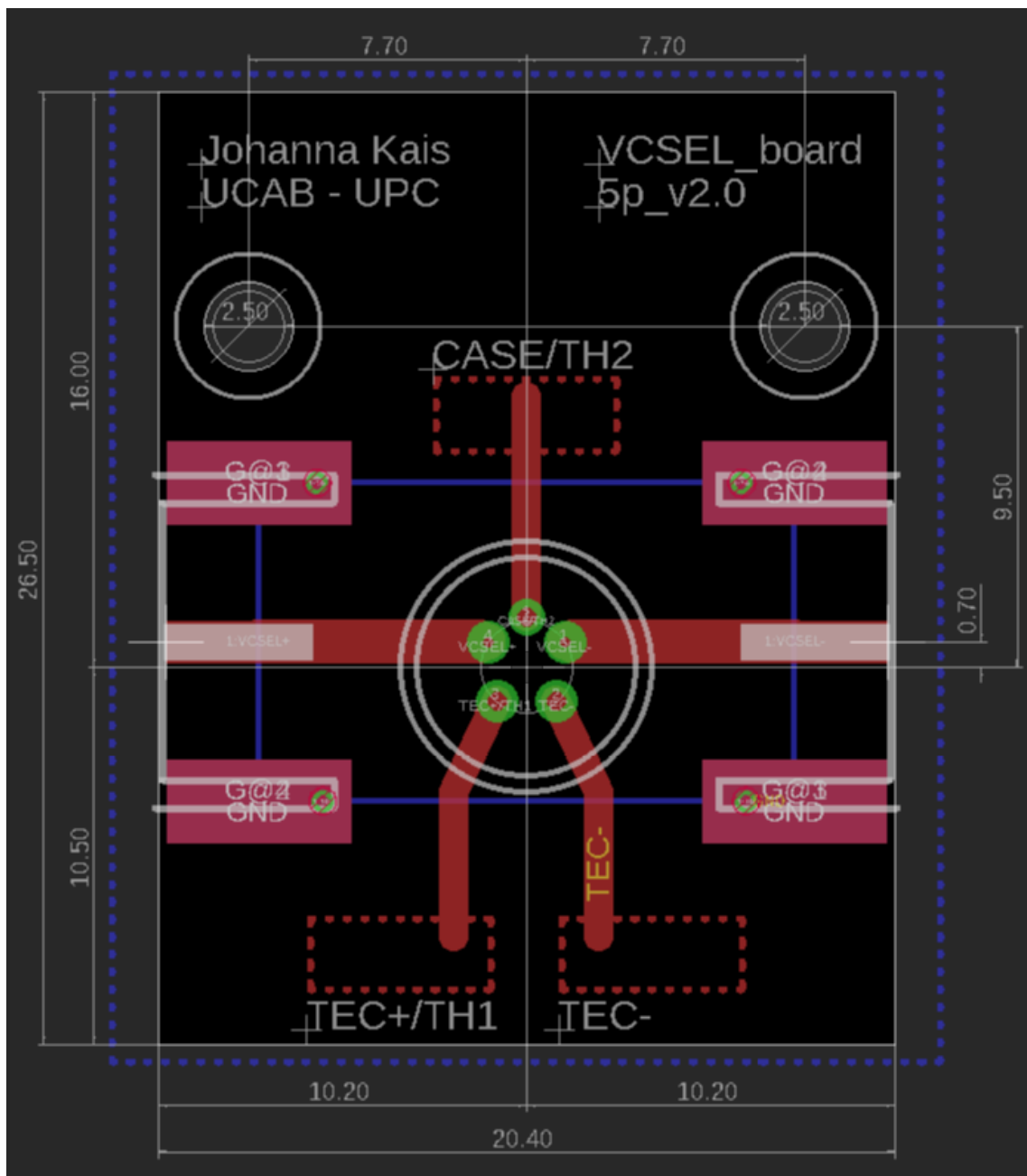
Parameter	Symbol	Rating	Unit	Notes
Forward current	I_f	15	<i>mA</i>	
Reverse voltage	V_r	5	<i>V</i>	
Operating temperature	T_{op}	70	$^{\circ}C$	
Storage Temperature	T_{stg}	0 – 100	$^{\circ}C$	
Reflow Temperature	T_{ref}	260	$^{\circ}C$	10 sec. 2 mm from case
TEC maximum current	I_{TMC}	0.7	<i>A</i>	

Appendix B – VCSEL’s First Board Design



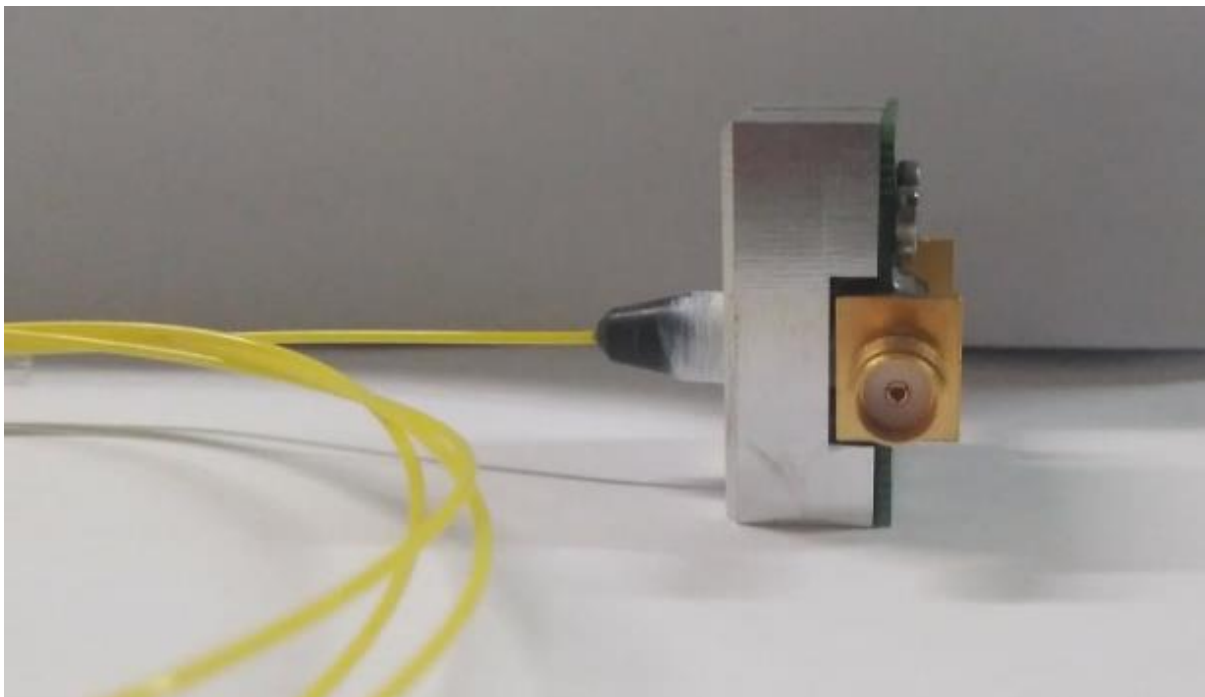
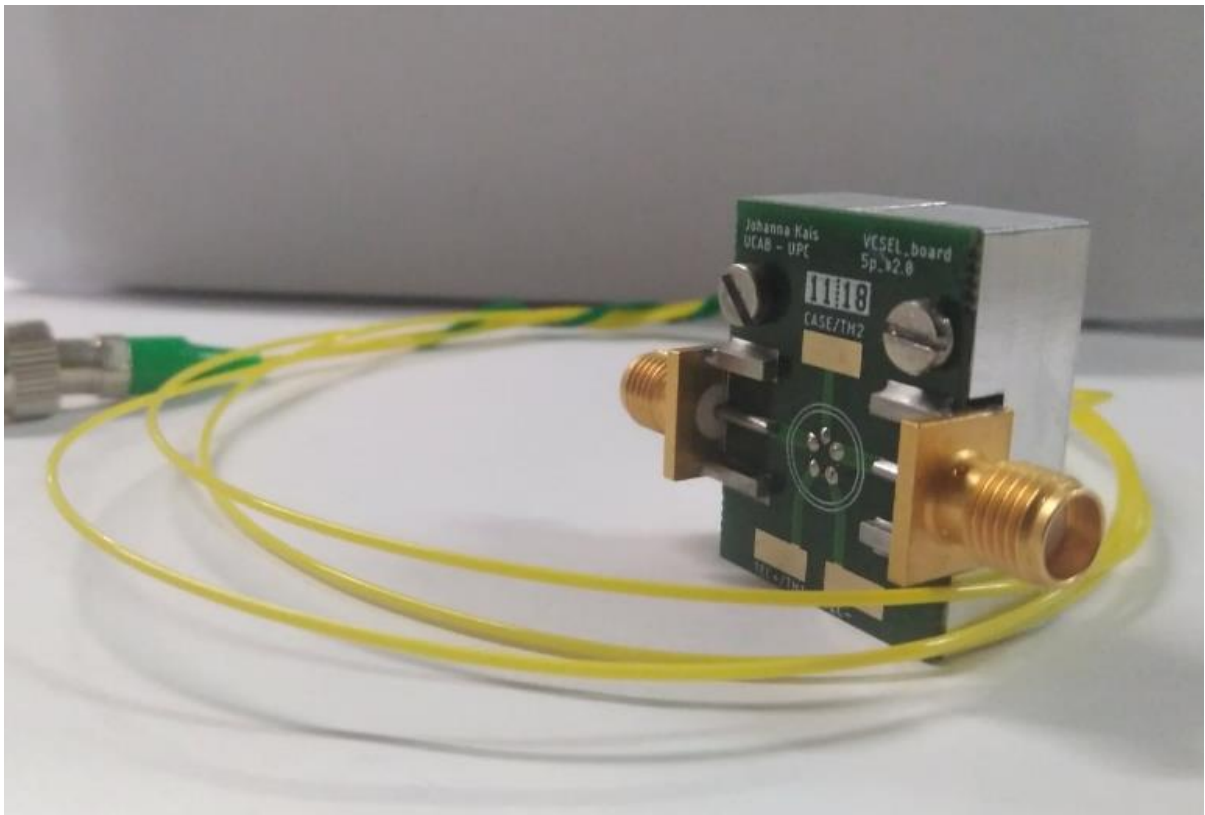
Notice that the SMA connectors are too close together to allow connections with two bias-tee at the same time, even if this task is managed, the connection and disconnection of cables are not easily maneuverer.

Appendix C – Definitive VCSEL Board Design with Dimensions



The resulting final design is 20.40 mm in width and 26.50 mm in height. SMA connectors are to be placed to the sides associated to the pins number one (1) and four (4), while the pins number two (2), three (3) and five (5) are linked to copper patches identified as corresponds. If the board is seen from the top (as above) the VCSELs anode is located to the left side above TEC+/TH1 patch while its cathode is located to the right.

Appendix D – VCSEL after inclusion on board with dissipation element

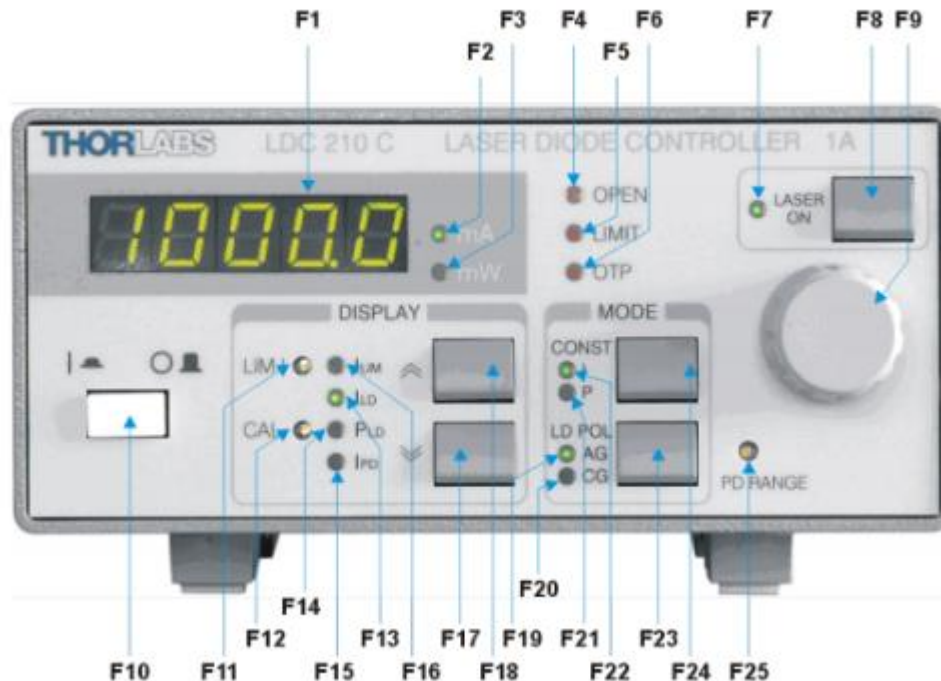


Appendix E – Resistance – Temperature Table for VCSEL’s Thermistor

		LNDG103HB				B25/85 = 3976K ± 2% R25 = 10.00kΩ ± 3%		Est Date	2005.07.01		
								Rev No.	0		
								Rev Date	-		
T(°C)	Rmin(kΩ)	Rcent(kΩ)	Rmax(kΩ)	DR(%)	DT(°C)	T(°C)	Rmin(kΩ)	Rcent(kΩ)	Rmax(kΩ)	DR(%)	DT(°C)
-40	231.4	254.5	279.7	9.89%	1.55	21	11.52	11.91	12.31	3.36%	0.76
-39	219.2	240.8	264.3	9.77%	1.54	22	11.03	11.40	11.77	3.27%	0.74
-38	207.7	227.9	249.9	9.65%	1.53	23	10.56	10.91	11.26	3.18%	0.73
-37	196.8	215.7	236.3	9.53%	1.52	24	10.12	10.44	10.77	3.09%	0.71
-36	186.5	204.2	223.4	9.41%	1.51	25	9.700	10.00	10.30	3.00%	0.69
-35	176.8	193.3	211.3	9.29%	1.50	26	9.281	9.577	9.873	3.09%	0.72
-34	167.6	183.1	199.9	9.17%	1.49	27	8.883	9.173	9.465	3.18%	0.74
-33	158.9	173.4	189.1	9.05%	1.48	28	8.503	8.789	9.076	3.27%	0.77
-32	150.6	164.2	178.9	8.93%	1.47	29	8.142	8.422	8.705	3.35%	0.79
-31	142.8	155.6	169.3	8.81%	1.46	30	7.797	8.073	8.351	3.44%	0.82
-30	135.5	147.4	160.2	8.69%	1.45	31	7.469	7.739	8.012	3.53%	0.84
-29	128.5	139.7	151.6	8.58%	1.44	32	7.156	7.421	7.690	3.62%	0.87
-28	121.9	132.4	143.6	8.46%	1.43	33	6.857	7.118	7.381	3.70%	0.89
-27	115.7	125.5	135.9	8.34%	1.42	34	6.573	6.828	7.087	3.79%	0.92
-26	109.8	118.9	128.7	8.23%	1.41	35	6.302	6.552	6.805	3.87%	0.95
-25	104.2	112.8	121.9	8.11%	1.39	36	6.043	6.288	6.537	3.96%	0.97
-24	98.93	106.9	115.5	8.00%	1.38	37	5.796	6.036	6.280	4.05%	1.00
-23	93.94	101.4	109.4	7.88%	1.37	38	5.560	5.795	6.034	4.13%	1.03
-22	89.20	96.22	103.7	7.77%	1.36	39	5.335	5.565	5.799	4.21%	1.05
-21	84.73	91.30	98.29	7.66%	1.35	40	5.120	5.345	5.575	4.30%	1.08
-20	80.48	86.64	93.17	7.55%	1.34	41	4.915	5.135	5.360	4.38%	1.11
-19	76.47	82.23	88.34	7.43%	1.32	42	4.719	4.934	5.155	4.47%	1.14
-18	72.67	78.06	83.77	7.32%	1.31	43	4.532	4.743	4.958	4.55%	1.16
-17	69.06	74.11	79.45	7.21%	1.30	44	4.353	4.559	4.770	4.63%	1.19
-16	65.65	70.38	75.37	7.10%	1.29	45	4.183	4.384	4.590	4.71%	1.22
-15	62.42	66.84	71.51	6.99%	1.27	46	4.019	4.216	4.418	4.79%	1.25
-14	59.36	63.50	67.86	6.88%	1.26	47	3.863	4.055	4.253	4.88%	1.28
-13	56.45	60.33	64.41	6.77%	1.25	48	3.714	3.901	4.095	4.96%	1.30
-12	53.70	57.33	61.15	6.66%	1.24	49	3.571	3.754	3.944	5.04%	1.33
-11	51.10	54.49	58.06	6.55%	1.22	50	3.435	3.614	3.799	5.12%	1.36
-10	48.62	51.80	55.14	6.44%	1.21	51	3.304	3.479	3.660	5.20%	1.39
-9	46.28	49.26	52.38	6.34%	1.20	52	3.179	3.350	3.526	5.28%	1.42
-8	44.06	46.84	49.76	6.23%	1.18	53	3.059	3.226	3.399	5.36%	1.45
-7	41.95	44.56	47.29	6.12%	1.17	54	2.945	3.108	3.276	5.44%	1.48
-6	39.95	42.39	44.94	6.02%	1.16	55	2.835	2.994	3.159	5.51%	1.51
-5	38.05	40.34	42.72	5.91%	1.14	56	2.730	2.885	3.047	5.59%	1.54
-4	36.25	38.39	40.62	5.81%	1.13	57	2.629	2.781	2.939	5.67%	1.57
-3	34.54	36.55	38.63	5.70%	1.12	58	2.533	2.681	2.835	5.75%	1.60
-2	32.92	34.80	36.75	5.60%	1.10	59	2.441	2.585	2.736	5.82%	1.63
-1	31.38	33.14	34.96	5.50%	1.09	60	2.352	2.493	2.640	5.90%	1.66
0	29.92	31.56	33.27	5.40%	1.08	61	2.267	2.405	2.549	5.98%	1.69
1	28.53	30.07	31.66	5.29%	1.06	62	2.186	2.320	2.461	6.05%	1.72
2	27.22	28.65	30.14	5.19%	1.05	63	2.108	2.239	2.376	6.13%	1.75
3	25.96	27.31	28.70	5.09%	1.03	64	2.033	2.161	2.295	6.20%	1.79
4	24.77	26.03	27.33	4.99%	1.02	65	1.961	2.086	2.217	6.28%	1.82
5	23.64	24.82	26.04	4.89%	1.00	66	1.892	2.014	2.142	6.35%	1.85
6	22.57	23.67	24.81	4.79%	0.99	67	1.826	1.945	2.070	6.43%	1.88
7	21.55	22.58	23.64	4.69%	0.97	68	1.762	1.879	2.001	6.50%	1.91
8	20.58	21.54	22.53	4.59%	0.96	69	1.701	1.815	1.934	6.58%	1.95
9	19.66	20.56	21.48	4.50%	0.94	70	1.643	1.754	1.870	6.65%	1.98
10	18.78	19.62	20.48	4.40%	0.93	71	1.587	1.695	1.809	6.72%	2.01
11	17.94	18.73	19.54	4.30%	0.91	72	1.533	1.638	1.750	6.79%	2.04
12	17.15	17.89	18.64	4.20%	0.90	73	1.481	1.584	1.693	6.87%	2.08
13	16.39	17.08	17.78	4.11%	0.88	74	1.431	1.532	1.638	6.94%	2.11
14	15.67	16.32	16.97	4.01%	0.87	75	1.383	1.481	1.585	7.01%	2.14
15	14.99	15.59	16.20	3.92%	0.85	76	1.337	1.433	1.534	7.08%	2.18
16	14.34	14.90	15.47	3.82%	0.84	77	1.293	1.386	1.485	7.15%	2.21
17	13.72	14.24	14.77	3.73%	0.82	78	1.250	1.341	1.438	7.22%	2.24
18	13.13	13.62	14.11	3.64%	0.81	79	1.209	1.298	1.393	7.29%	2.28
19	12.56	13.02	13.48	3.55%	0.79	80	1.170	1.257	1.349	7.36%	2.31
20	12.03	12.45	12.88	3.45%	0.77						

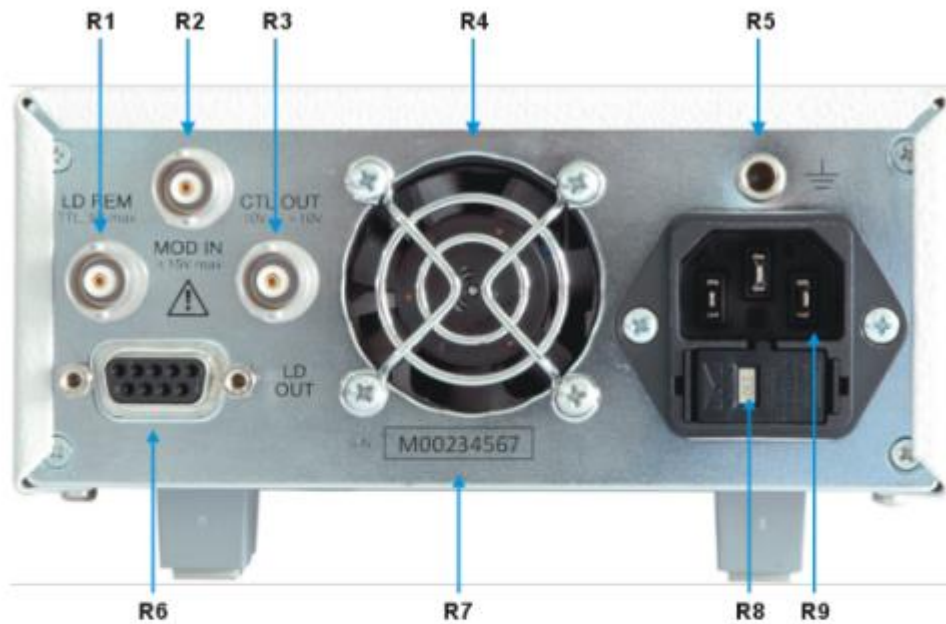
Appendix F – Front and Rear Panel for LDC200C

Front Panel



F1	-	5-digit LED display
F2	LED "mA"	Current display in mA
F3	LED "mW"	Power display in mW
F4	LED "OPEN"	No laser diode connected, or Interlock open
F5	LED "LIMIT"	Adjusted current limit reached
F6	LED "OTP"	Overtemperature protection is active
F7	LED "LASER ON"	Laser current is switched on
F8	Key "LASER ON"	On / Off switch for the laser current
F9	-	Knob for adjusting the current or power set value
F10	-	Line switch (ON / OFF)
F11	LIM I	Potentiometer for setting the current limit
F12	CAL	Potentiometer for calibrating the power display
F13	LED "I _{LD} "	Display shows the laser current
F14	LED "P _{LD} "	Display shows the optical power
F15	LED "I _{PD} "	Display shows the photodiode current
F16	LED "I _{LIM} "	Display shows the current limit
F17	Key "DOWN"	Select the parameter to be displayed
F18	Key "UP"	Select the parameter to be displayed
F19	LED "AG"	Selected laser polarity: anode grounded
F20	LED "CG"	Selected laser polarity: cathode grounded
F21	LED "P"	Constant power mode
F22	LED "I"	Constant current mode
F23	Key "LD POL"	Select laser polarity: anode grounded or cathode grounded
F24	Key "CONST"	Select constant current mode or constant power mode
F25	PD RANGE	Potentiometer for setting the photodiode current range

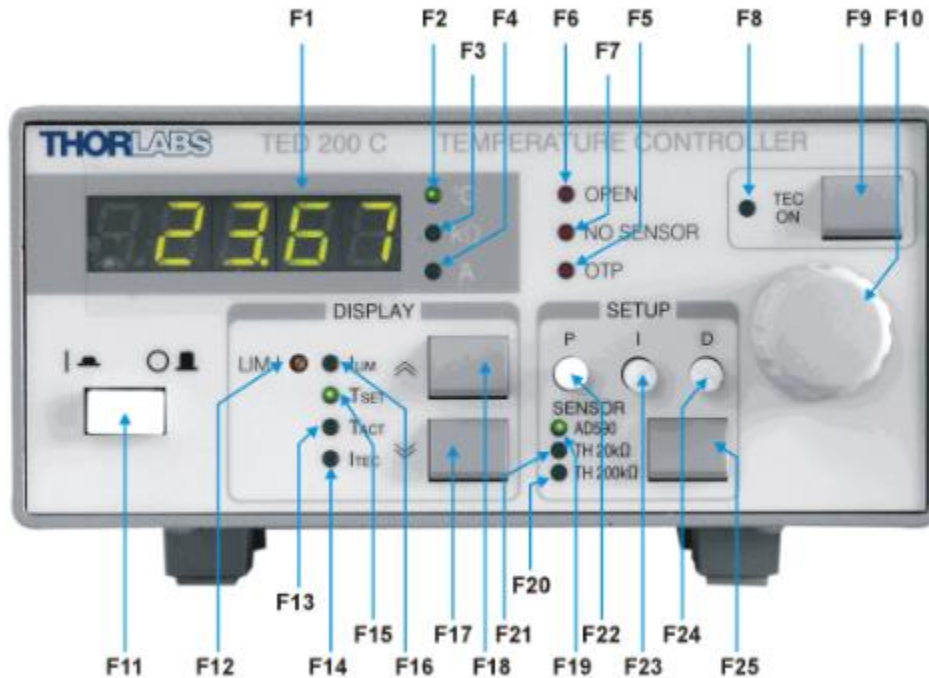
Rear Panel



- R1** TTL input "LD REM", 0 ... +5 V
- R2** Modulation input / analog control input "MOD IN", -10V ... +10 V
- R3** Analog monitoring output "CTL OUT", 0 ... ±10V
- R4** Fan
- R5** 4 mm banana jack for chassis ground
- R6** Connector "LD OUT" for laser diode, photodiode, interlock, status LED
- R7** Serial number of the unit
- R8** Indicator / switch for line voltage (included in fuse holder)
- R9** Mains connector and fuse holder

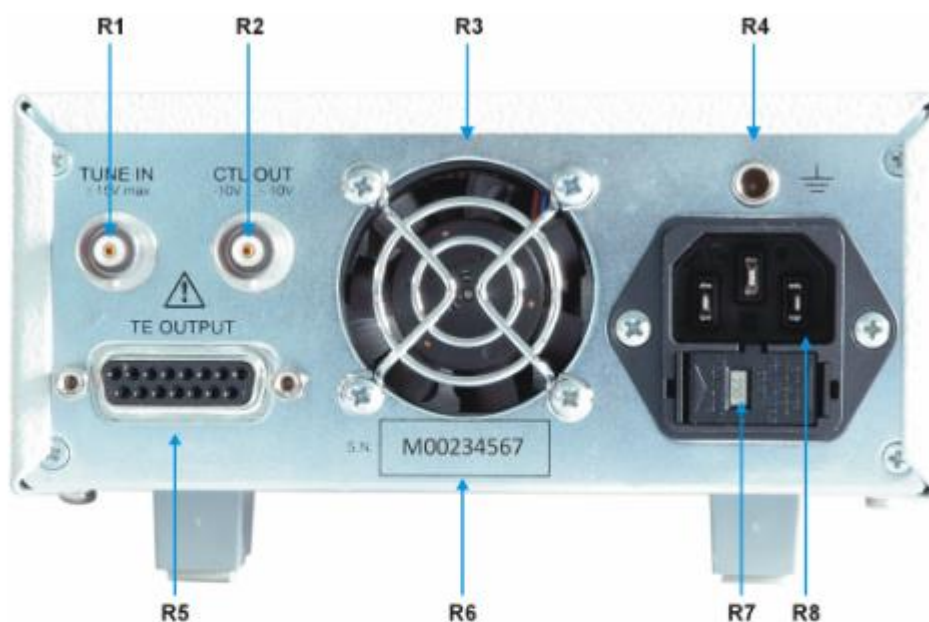
Appendix G – Front and Rear Panel for TED200C

Front Panel



F1	-	5-digit LED display
F2	LED "°C"	Temperature display in C
F3	LED "kΩ"	Resistance display in k
F4	LED "A"	Current display in A
F5	LED "OTP"	Over temperature protection is active
F6	LED "OPEN"	TEC element is not connected or too high resistance
F7	LED "NO SENSOR"	Temperature sensor is wrong or not connected
F8	LED "TEC ON"	TEC output is switched on
F9	Key "TEC ON"	On / Off switch for the TEC output
F10	-	Knob for adjusting the set temperature / resistance
F11	-	Line switch (ON / OFF)
F12	LIM I	Potentiometer for setting the TEC current limit
F13	LED "TACT"	Display shows the actual temperature / resistance
F14	LED "ITEC"	Display shows the TEC current
F15	LED "TSET"	Display shows the set temperature / resistance
F16	LED "ILIM"	Display shows the current limit
F17	Key "DOWN"	Select the parameter to be displayed
F18	Key "UP"	Select the parameter to be displayed
F19	LED "AD590"	Selected sensor is AD 590, AD 592, LM 135 or LM 335
F20	LED "TH 200kΩ"	Selected sensor is thermistor in the 200 kΩ range
F21	LED "TH 20kΩ"	Selected sensor is thermistor in the 20 kΩ range
F22	P	Potentiometer for setting P- (gain) share of control loop
F23	I	Potentiometer for setting I- (integral) share of control loop
F24	D	Potentiometer for setting D- (derivative) share of control loop
F25	Key "SENSOR"	Select sensor / disable I-share (press for more than 1 sec.)

Rear Panel



- R1** Analog temperature control input "TUNE IN", -10 ... +10 V
- R2** Analog temperature control output "CTL OUT", -10 ... +10 V
- R3** Fan
- R4** 4 mm banana jack for chassis ground
- R5** 15-pin D-SUB jack for the TEC element and the temperature sensor "TE OUTPUT"
- R6** Serial number of the unit
- R7** Indicator / switch for line voltage (included in fuse holder)
- R8** Mains connector and fuse holder

Appendix H – Values obtained from VCSEL's Characterization

The following tables include the different measures performed to the VCSEL while the characterization was performed. The used setup is shown in Figure XX from chapter 3, as well as a definition on how each value was obtained in each iteration.

Input current: 2 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.546,325	0,211	-8,49	33,32	-8,95	1,421
15.590 - 15 °C	1.546,910	0,136	-8,83	33,24	-9,31	1,413
12.450 - 20 °C	1.547,485	0,073	-9,37	33,29	-9,83	1,405
10.000 - 25 °C	1.548,071	0,015	-9,86	33,45	-10,42	1,397
8.073 - 30 °C	1.548,658	-0,036	-10,60	33,80	-10,97	1,390
6.552 - 35 °C	1.549,246	-0,082	-11,53	33,32	-11,90	1,384
5.345 - 40 °C	1.549,839	-0,124	-12,81	33,85	-13,17	1,378
4.384 - 45 °C	1.550,439	-0,162	-14,75	33,38	-15,20	1,372
3.614 - 50 °C	1.551,035	-0,198	-18,99	33,30	-19,34	1,368

Input current: 3 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.546,742	0,211	-5,53	33,96	-6,06	1,551
15.590 - 15 °C	1.547,324	0,137	-5,84	33,99	-6,33	1,540
12.450 - 20 °C	1.547,910	0,073	-6,10	33,91	-6,66	1,531
10.000 - 25 °C	1.548,497	0,017	-6,50	33,24	-7,00	1,522
8.073 - 30 °C	1.549,086	-0,033	-6,93	35,31	-7,49	1,514
6.552 - 35 °C	1.549,680	-0,078	-7,49	35,33	-7,94	1,506
5.345 - 40 °C	1.550,283	-0,119	-8,02	35,35	-8,59	1,498
4.384 - 45 °C	1.550,879	-0,157	-8,87	35,50	-9,39	1,492
3.614 - 50 °C	1.551,491	-0,194	-9,86	35,60	-10,30	1,485

Input current: 4 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.547,215	0,219	-3,73	34,06	-4,20	1,656
15.590 - 15 °C	1.547,799	0,144	-3,89	34,04	-4,48	1,645
12.450 - 20 °C	1.548,392	0,079	-4,23	34,43	-4,70	1,634
10.000 - 25 °C	1.548,981	0,022	-4,54	35,56	-5,05	1,624
8.073 - 30 °C	1.549,579	-0,029	-4,91	35,62	-5,38	1,615
6.552 - 35 °C	1.550,181	-0,075	-5,26	35,74	-5,80	1,606
5.345 - 40 °C	1.550,785	-0,117	-5,91	36,09	-6,36	1,598
4.384 - 45 °C	1.551,390	-0,156	-6,44	36,04	-6,92	1,590
3.614 - 50 °C	1.552,002	-0,191	-7,24	36,06	-7,70	1,583

Input current: 5 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.547,724	0,224	-2,33	34,01	-2,90	1,745
15.590 - 15 °C	1.548,316	0,147	-2,57	34,31	-3,10	1,734
12.450 - 20 °C	1.548,910	0,083	-2,90	35,33	-3,40	1,722
10.000 - 25 °C	1.549,509	0,026	-3,18	35,55	-3,70	1,712
8.073 - 30 °C	1.550,113	-0,025	-3,51	35,62	-4,05	1,702
6.552 - 35 °C	1.550,724	-0,071	-4,07	35,97	-4,52	1,693
5.345 - 40 °C	1.551,333	-0,112	-4,49	36,04	-4,98	1,683
4.384 - 45 °C	1.551,947	-0,151	-5,17	36,27	-5,62	1,675
3.614 - 50 °C	1.552,575	-0,187	-5,81	36,06	-6,31	1,667

Input current: 6 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.548,286	0,234	-1,39	33,89	-1,91	1,823
15.590 - 15 °C	1.548,880	0,157	-1,69	35,22	-2,18	1,811
12.450 - 20 °C	1.549,486	0,090	-2,00	35,43	-2,50	1,800
10.000 - 25 °C	1.550,096	0,032	-2,31	35,53	-2,84	1,788
8.073 - 30 °C	1.550,713	-0,021	-2,88	35,97	-3,33	1,778
6.552 - 35 °C	1.551,327	-0,068	-3,29	36,08	-3,75	1,768
5.345 - 40 °C	1.551,948	-0,111	-3,93	36,24	-4,37	1,758
4.384 - 45 °C	1.552,577	-0,150	-4,51	36,32	-5,00	1,749
3.614 - 50 °C	1.553,202	-0,185	-5,31	36,43	-5,75	1,738

Input current: 7 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.548,891	0,237	-0,91	34,59	-1,36	1,892
15.590 - 15 °C	1.549,500	0,160	-1,21	34,83	-1,72	1,879
12.450 - 20 °C	1.550,116	0,094	-1,57	35,27	-2,09	1,867
10.000 - 25 °C	1.550,735	0,037	-2,11	35,78	-2,56	1,856
8.073 - 30 °C	1.551,353	-0,014	-2,54	35,98	-3,00	1,845
6.552 - 35 °C	1.551,977	-0,060	-3,13	36,28	-3,58	1,834
5.345 - 40 °C	1.552,610	-0,103	-3,69	36,45	-4,14	1,825
4.384 - 45 °C	1.553,241	-0,144	-4,39	36,48	-4,81	1,815
3.614 - 50 °C	1.553,878	-0,181	-5,10	36,44	-5,57	1,807

Input current: 8 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.549,564	0,245	-0,74	34,93	-1,21	1,952
15.590 - 15 °C	1.550,183	0,170	-1,11	35,31	-1,59	1,939
12.450 - 20 °C	1.550,803	0,102	-1,60	35,91	-2,02	1,927
10.000 - 25 °C	1.551,424	0,043	-2,01	36,03	-2,47	1,915
8.073 - 30 °C	1.552,047	-0,009	-2,48	36,27	-2,93	1,905
6.552 - 35 °C	1.552,680	-0,057	-2,95	36,47	-3,44	1,894
5.345 - 40 °C	1.553,313	-0,100	-3,60	36,52	-4,00	1,884
4.384 - 45 °C	1.553,51	-0,139	-4,17	36,60	-4,63	1,875
3.614 - 50 °C	1.554,596	-0,176	-4,89	36,47	-5,38	1,866

Input current: 9 mA

TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.550,268	0,258	-0,63	35,60	-1,13	2,007
15.590 - 15 °C	1.550,883	0,178	-0,99	36,01	-1,44	1,994
12.450 - 20 °C	1.551,505	0,109	-1,38	36,03	-1,86	1,982
10.000 - 25 °C	1.552,128	0,049	-1,72	36,18	-2,21	1,970
8.073 - 30 °C	1.552,762	-0,003	-2,23	36,49	-2,67	1,959
6.552 - 35 °C	1.553,396	-0,051	-2,71	36,53	-3,15	1,949
5.345 - 40 °C	1.554,039	-0,094	-3,30	36,55	-3,71	1,939
4.384 - 45 °C	1.554,694	-0,134	-3,91	36,47	-4,37	1,930
3.614 - 50 °C	1.555,353	-0,171	-4,80	36,63	-5,23	1,919

Input current: 10 mA

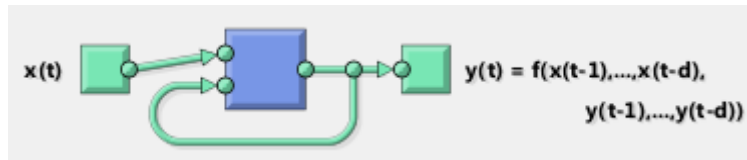
TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.550,979	0,268	-0,34	36,08	-0,77	2,056
15.590 - 15 °C	1.551,605	0,189	-0,69	36,15	-1,14	2,044
12.450 - 20 °C	1.552,233	0,119	-0,96	36,21	-1,46	2,032
10.000 - 25 °C	1.552,869	0,057	-1,48	36,43	-1,87	2,020
8.073 - 30 °C	1.553,511	0,003	-1,91	36,51	-2,35	2,009
6.552 - 35 °C	1.554,160	-0,045	-2,47	36,50	-2,87	1,999
5.345 - 40 °C	1.554,825	-0,090	-3,13	36,56	-3,56	1,988
4.384 - 45 °C	1.555,492	-0,130	-4,05	36,53	-4,39	1,979
3.614 - 50 °C	1.555,174	-0,167	-5,16	36,57	-5,52	1,970

Input current: 11 mA

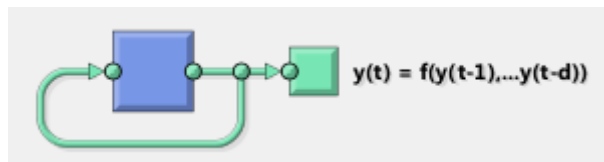
TH (Ω) - Temp	λ (nm)	I_{TEC} (Amp)	P_{λ} (dBm)	OSNR (dB)	P_{TOTAL} (dBm)	V_{VCSEL} (V)
19.620 - 10 °C	1.551,716	0,276	0	36,35	-0,41	2,102
15.590 - 15 °C	1.552,353	0,192	-0,31	36,23	-0,78	2,090
12.450 - 20 °C	1.552,997	0,122	-0,84	36,44	-1,19	2,077
10.000 - 25 °C	1.553,646	0,061	-1,29	36,40	-1,72	2,066
8.073 - 30 °C	1.554,307	0,008	-1,88	36,41	-2,32	2,055
6.552 - 35 °C	1.554,980	-0,039	-2,71	36,48	-3,08	2,041
5.345 - 40 °C	1.555,658	-0,083	-3,63	36,48	-4,00	2,033
4.384 - 45 °C	1.556,351	-0,123	-4,85	36,57	-5,19	2,024
3.614 - 50 °C	1.557,071	-0,162	-6,36	36,48	-6,77	2,014

Appendix I – Diagrams for MATLABs Deep Learning Toolbox solvers in Dynamic Time-Series Analysis

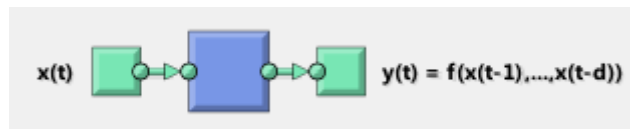
Nonlinear Autoregressive with External (Exogenous) Input (NARX): Predicts series $y(t)$ given d past values of $y(t)$ and another series $x(t)$



Nonlinear Autoregressive (NAR): Predicts series $y(t)$ given d past values of $y(t)$



Nonlinear Input-Output: Predicts series $y(t)$ given d past values of $x(t)$



NOTE: NARX solutions are more accurate than this solution. However, use this case if past values of $y(t)$ will not be available when the system is deployed.

Glossary

AI	Artificial intelligence
ANN	Artificial neural network
AR	Antireflection coatings
AWG	Arbitrary waveform generator
b/s/Hz	Bits per second per Hertz
BL	Bit rate-distance product
BR	Bayesian Regularization. NN training algorithm
CNN	Convolutional neural networks
CSI	Channel state information
CSIR	Channel state information at receiver
CSIT	Channel state information at transmitter
cw	Continuous wave
D	Delay, for a NN training
dB	Decibels
DBR	Distributed Bragg reflector
DFB	Distributed-Feedback lasers
DPSK	Differential phase-shift keying
EDFA	Erbium-doped fiber amplifiers
FEA	Flag Europe Asia
FLAG	Fiber-Optic Link Around the Globe
GaAs	Gallium arsenide
GaAsP	Gallium arsenide phosphide
Gb/s	Giga bit per second
GaInAs	Gallium indium arsenide
GHz	Giga Hertz
GS/s	Giga samples per second
GUI	Graphic user interface
HL	Hidden layers, for a NN training
InAs	Indium arsenide
InGaAsP	Indium gallium arsenide phosphide
InP	Indium phosphide

IoT	Internet of things
km	Kilometers
LASER	Light Amplification by Stimulated Emission of Radiation
LED	Light emitting diode
LIV	Light-current-voltage
LM	Levenberg-Marquardt. NN training algorithm
LPE	Liquid-phase epitaxy
mA	Milli ampere
MASER	Microwave Amplification by the Stimulated Emission of Radiation
Mb/s	Mega bit per second
MEMS	Micro-electro-mechanical system
MSE	Mean squared error. NN key parameter
mW	Milliwatts
MZM	Mach-Zehnder modulator
NAR	Nonlinear autoregressive
nm	Nanometer
NN	Neural network
ODE	Ordinary differential equation
OOK	On-off keying
QAM	Quadrature amplitude modulation
R	Regression. NN key parameter
RNN	Recurrent neural networks
SCG	Scaled Conjugate Gradient, NN training algorithm
SGD	Stochastic Gradient Descent. NN training algorithm
SM	Single mode
Tb/s	Tera bits per second
TEC	Thermoelectric cooler
THz	Tera Hertz
TLS	Tuneable laser source
uDWDM	Ultra-dense wavelength division multiplexing
μm	Micrometer
VCSEL	Vertical-cavity surface-emitting laser
WDM	Wavelength division multiplexing