Daniel André Pires Duarte

Plataforma Ótica Multiparametro para a Monitorização da Qualidade de Líquidos

Optical Multiparameter Platform for the Quality Monitoring of Liquids

Universidade de Aveiro Departamento de Física
2020

Daniel André Pires Duarte

Plataforma Ótica Multiparametro para a Monitorização da Qualidade de Líquidos

Optical Multiparameter Platform for the Quality Monitoring of Liquids

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Física, realizada sob a orientação científica da Doutora Lúcia Maria Botas Bilro, Investigadora Auxiliar do Instituto de Telecomunicações e do Doutor Rogério Nunes Nogueira, Investigador Principal do Instituto de Telecomunicações.

"The progressive development of man is vitally dependent on invention. It is the most important product of his creative brain. Its ultimate purpose is the complete mastery of mind over the material world, the harnessing of the forces of nature to human needs"

Nicola Tesla

Aos meus pais, irmão e familia.

o júri / the jury

presidente / president

Prof. Doutor Fernando Joaquim Fernandes Tavares Rocha
Professor Catedrático, Universidade de Aveiro

vogais / examiners committee

Prof. Doutor João de Lemos Pinto
Professor Catedrático, Universidade de Aveiro

Prof. Doutor Joseba Andoni Zubia Zaballa
Professor Catedrático, Universidade do País Basco

Prof. Doutor Emanuel Soares Peres Correia
Professor Auxiliar, Universidade de Trás-Os-Montes e Alto Douro

Prof. Doutor Orlando José dos Reis Frazão
Professor Auxiliar Convidado, Universidade do Porto

Doutora Lúcia Maria Botas Bilro
Investigadora Auxiliar do Instituto de Telecomunicações de Aveiro

**agradecimentos/
acknowledgments**

Este trabalho, é não só o resultado da pesquisa, experimentação e reflexão realizada pelo autor, mas também pelo esforço e apoio incondicional de outras pessoas cujo contributo foi imprescindível para a sua realização.

Em particular gostaria de agradecer ao Doutor Rogério Nunes Nogueira e Doutora Lúcia Maria Botas Bilro pela orientação, apoio, permanente disponibilidade e o constante encorajamento, motivação e amizade durante a realização desta tese, e tudo o que isso contribuiu para o meu enriquecimento profissional e pessoal.

À Watgrid, pela colaboração prestada neste trabalho e sem a qual este não seria possível. Agradeço por me receber nas suas instalações e aos seus funcionários e amigos pelo fantástico apoio, troca de conhecimento e amizade criada e prestada, tanto nos bons como nos momentos de maior frustração. Ao Pedro Costa, Fábio Gonçalves, Nuno Oliveira, Ricardo Pereira e Mariana Figueira, pelo contributo direto proporcionado, e a todos os outros que passaram pela Watgrid durante o período deste trabalho, o meu muitíssimo obrigado.

A todos os colegas e amigos do Instituto de Telecomunicações, com menção especial à Mestre Filipa Sequeira e ao Doutor Ricardo Oliveira pelo constante companheirismo e discussões produtivas e ao Eng. João Prata pelo apoio laboratorial prestado.

Quero agradecer ao Instituto de Telecomunicações – Polo de Aveiro pelo acolhimento e excelentes condições proporcionadas para a realização deste trabalho, assim como ao Departamento de Física pela disponibilização de equipamento laboratorial imprescindível para a conclusão do mesmo.

À Fundação para a Ciência e a Tecnologia e ao Fundo Europeu de Desenvolvimento Regional, por tornarem viável financeiramente a realização deste Trabalho.

Por fim, não menos importante, quero agradecer ao meus amigos e à minha família, especialmente aos meus pais pelo amor, apoio, abnegação e encorajamento sempre prestados e ao meu irmão, cunhada e sobrinho pelos momentos agradáveis e de humor vividos durante este período.

A todos, o meu mais sincero agradecimento!

palavras-chave

Sensores óticos, sensores multiparametro, sensores para liquidos, sensor de cor e turbidez, sensor de indíce de refracção, aprendizagem de máquina, fusão de dados, internet das coisas.

resumo

A próxima década irá estabelecer-se com a maturação e massificação da internet das coisas, conhecida por IoT. A IoT não só servirá o consumidor doméstico com aparelhos e objetos que se encontram habilitados a estarem permanentemente ligados à Internet, como servirá de espinha dorsal para a nova revolução industrial 4.0, já em curso. Diversos dispositivos instalados ao longo da cadeia de produção irão promover uma grande interação com o utilizador, em permanente comunicação. É neste contexto que este trabalho de insere.

Neste trabalho foram desenvolvidos sensores óticos para a monitorização de líquidos. nomeadamente na monitorização dos parâmetros de cor, turbidez e índice de refração. Estes sensores são desenvolvidos numa perspetiva de baixo custo e possuem capacidade de comunicação sem fios para envio e visualização de dados numa plataforma localizada num servidor de internet.

O trabalho inicia-se com a pesquisa do estado da arte dos sensores para líquidos. Feito o levantamento, foram investigadas e desenvolvidas as tecnologias que melhor satisfazem os objetivos iniciais do sensor, isto é, capacidade de medição em linha, baixo custo, baixa complexidade e máxima resolução por custo possível.

Para a cor e turbidez foram desenvolvidos sensores baseados na transmissão e dispersão de luz em 90º entre fibras óticas de plástico numa configuração *air-gap*. Nesta configuração o liquido atravessa a zona de medição que fica na separação entre as fibras e consoante o perfil de absorção/dispersão de luz, os dois parâmetros são medidos. Devido à sua natureza multiváriavel, algoritmos de aprendizagem automática e fusão de dados foram desenvolvidos e aplicados aos valores em bruto para obter medições mais precisas e robustas dos parâmetros.

Para o índice de refração, foi desenvolvido um sensor baseado na reflexão total de luz num prisma onde o índice de refração é medido através do perfil de luz refletida, detetado por um dispositivo de carga acoplada.

Os sensores foram aplicados em casos de estudo. No caso do sensor de cor e turbidez este foi aplicado na monitorização da prensagem, fermentação e maturação de vinho. No caso do sensor de índice de refração, foi efetuado um teste para identificação da possível monitorização de soluções de limpeza para maquinaria industrial.

keywords

Optic sensors, multiparameter sensors, liquid sensors, color and turbidity sensor, refractive index sensor, machine learning, data fusion, internet of things.

abstract

The next decade will be established with the maturation and massification of the internet of things, known as IoT. The IoT not only will be used by the domestic consumer, with machines and objects having the ability to be permanently connected to the internet, but also will be the backbone for the new industrial revolution 4.0, already underway, where several installed devices in the production chains will promote a great interaction with the user, in permanent communication. It is within this context that this work was done.

In this work optical sensors for the monitoring of liquids were developed, mainly for the parameters monitoring of color, turbidity and refractive index. These sensors are developed in a low-cost perspective, having wireless communication capability to send and visualize data to an internet server based platform.

This work is initiated with the search of the state of art sensors for liquids. With the survey concluded, the best technologies that suffice the initial sensor objectives are selected, this is, in-line measurement, low-cost, low complexity and best cost-resolution ratio possible.

For the color and turbidity, sensors were developed based in the transmission and scattering at 90º of light between plastic optic fibers in an air-gap configuration. In this configuration, the liquid goes through the measurement zone that is situated between the fibers. Depending on the light absorption/scattering profile, the two parameters are measured. Due to the multivariate sensor nature, machine learning and data fusion algorithms were developed and applied to the raw data to achieve more precise and robust measurements of the parameters.

For the refractive index, it was developed a sensor based in the total reflection of light in a prism where the refractive index is measured by the profile of the reflected light detected by a charge-coupled device.

The sensors were applied in case studies. The color and turbidity sensor was tested in the monitoring of wine pressing, fermentation and maturation. The refractive index sensor was tested to identify the possibility to be used in the monitoring of cleaning solutions for industrial machinery.

CONTENTS

List of Acronyms

Abs	*Absorption Units*
AI	*Artificial Intelligence*
ANN	*Artificial Neural Network*
ATR	*Attenuated Total Reflection*
BC	*Before Christ*
CCD	*Charge-Coupled Device*
CO_2	*Carbon Dioxide*
EM	*Expectation Maximization*
EMGM	*Expectation Maximization Semi-Supervised Gaussian Mixture*
EPA	*United States Of America Environmental Protection Agency*
FBG	*Fiber Bragg Grating*
FFPI	*Fabry-Pérot Interferometer*
FLR	*Fuzzy Logic Rules*
FWHM	*Full Width At Half Maximum*
GLI	*Great Lakes Instrument*
IOT	*Internet Of Things*
IR	*Infra-Red Light*
ISO	*International Organization For Standardization*
LCD	*Liquid Crystal Display*
LDA	*Linear Discriminant Analysis*
LED	*Light Emitting Diode*
LPG	*Long Period Grating*
LSPR	*Localized Surface Plasmon Resonance*

MIP	*Molecular Imprinted Polymer*
MMF	*Multimode Fiber*
MMI	*Fiber Multimode Interference*
PCF	*Photonic Crystal Fiber*
PMT	*Photomultiplier Tube*
POF	*Plastic Optic Fibers*
RGB	*Red-Blue-Green*
RI	*Refractive Index*
RIU	*Refractive Index Units*
RMSE	*Root Mean Square Error*
rwA	*Red Wine A*
rwB	*Red Wine B*
SMF	*Single Mode Fiber*
SPR	*Surface Plasmon Resonance*
TFBG	*Tilted Fiber Bragg Grating*
TiO$_2$	*Titanium Dioxide*
TM	*Transverse Magnetically*
UV	*Ultra-Violet Light*
WAM	*Weighed Arithmetic Mean*
wwA	*White Wine A*
wwB	*White Wine B*

List of Symbols

A	Light Attenuation
$A1$	Ratio Of The Receiving Light Area In Fiber 1
$A1$	Sigmoid Positive Asymptote
$A2$	Ratio Of The Receiving Light Area In Fiber 2
$A2$	Sigmoid Negative Asymptote
A_{col}	Color Absorption
Act_1	Active Signal From The Detector 1 Of The Ratiometric Turbidity Sensor
Act_2	Active Signal From The Detector 2 Of The Ratiometric Turbidity Sensor
a_n	Mie Coefficient Of The Amplitudes Of The Scattered Field
A_{tb}	Turbidity Attenuation
A_{tot}	Total Attenuation With Two Species
A_{TSS}	Turbidity To Total Suspended Solids Correlation Coefficient A
A_v	Solved Matrix Of The Generalized Eigenvalue Problem
b_n	Mie Coefficient Of The Amplitudes Of The Scattered Field
B_{TSS}	Turbidity To Total Suspended Solids Correlation Coefficient B
c	Velocity In Vacuum
C_0	Calibration Coefficient C_0 Of The Ratiometric Calculation For Turbidity
C_i	Sellmeier's Equation Coefficients
CI	Color Intensity
C_s	Calibration Coefficient C_s Of The Ratiometric Calculation For Turbidity
D	Number Of Dimensions
d	Number Of Data Dimensions
D_j	Artificial Neural Network Desired Output Value

n_1		Refractive Index Of Medium 1
n_2		Refractive Index Of Medium 2
NA		Numeric Aperture
n_{cl}		Refractive Index Of The Fiber Cladding
n_{co}		Refractive Index Of The Fiber Core
N_i		Sample Size
N_{next}		Refractive Index Of External Medium
NTU		Turbidity In Nephelometric Turbidity Units
OD		Optical Density
$p(Z/X,\theta)$		Posterior Distribution
$P1$		Ratio Of Received Light Power In Fiber 1
$P2$		Ratio Of Received Light Power In Fiber 2
P_V		Light Power Variation
P_W		Light Power Obtained In Water
Q		Expectation Of The Complete-Data Likelihood To A General Set Of Model
Q_{sca}		Light Scattered Coefficient
R		Light Reflection
R_c		Radius Of The Fiber Core
Ref_1		Reference Signal From The Detector 1 Of The Ratiometric Turbidity Sensor
Ref_2		Reference Signal From The Detector 2 Of The Ratiometric Turbidity Sensor
r_s		Sphere Radius
S		Labeled Elements
S_B		Between-Class Scatter Matrix
S_i		Scatter Matrix
S_W		Within-Class Scatter Matrix
T_c		Data Fusion Dimensional Transmittance Value

T_{exp}	Data Fusion Expected Turbidity
TSS	Total Suspended Solids
U	Unlabeled Elements
v_n	Propagation Velocity
w_n	Associated Weight
X	Observed Data
x_0	Sigmoid Central Position
x_n	Artificial Neural Network Inputs
y	Artificial Neural Network Output Signal
y_n	Spherical Bessel Functions Of The Second Kind
Z	Latent Variables
z	Argument Of The Bessel Functions
β	Bias Variable
$\gamma_{i,j}$	Data Fusion Dimensional Error
δ_j	Artificial Neural Network Margin Of Error
ε	Artificial Neural Network And Data Fusion Error Defined Criterion
ε_a	Absorption Coefficient
ε_S	Scattering Coefficient
η	Artificial Neural Network Rate Constant
θ	Model Parameters
θ_c	Critical Angle
θ_{cc}	Complementary Angle Of θ_c.
θ_i	Angle Of Incident Light
θ_{max}	Maximum Incident Light Angle Of Light Propagating In A Fiber
θ^{new}	New Model Parameters
θ^{old}	Old Model Parameters

θ_t		*Angle Of Refracted Light*
λ		*Light Wavelength*
λ_v		*Eigenvalues*
μ		*Mean Vector*
v		*Eigenvectors*
π_k		*Mixing Coefficient*
Σ		*Covariance Matrix*
T		*Data Fusion Weighed Turbidity Vector*
τ_i		*Data Fusion Dimensional Turbidity*
τ_{nk}		*Posterior Probability*
ς		*Particles Concentration*
$|\Sigma|$		*The Determinant Covariance Matrix*

1. Introduction

1.1. Framework

The society of the todays developed world has created over the last years a market for real-time online information that has an explosive growth and demand, mainly because of a new generation information technology emergence, including internet, mobile internet, wearable computing, and sensor networks. Companies that do not embrace this market and technology will have the tendency to slowly become outdated and less competitive against its peers, resulting in its ultimate insolvency.

The usage of sensor networks, mainly based in wireless connectivity over traditional methods of production and environmental parameter analysis have its advantages [Tiwari, 2015]:

- A fixed infrastructure for network setup is unnecessary;
- Suitable for non-reachable places;
- Cheap implementation pricing;
- Avoids wiring;
- Scalable;
- Information accessed by using a centralized monitor.

This type of technology can give a company's product add value and/or the possibility of timely react to any production flaws that can occur. Storing the historical data obtained by this sensor network in the perspective of Big Data, could also give the opportunity of a company to compare its actual production to the previous ones, developing new approaches by identifying improvements that can be implemented definitively in a production line. It also could help identify systematic or regular flows that could not be easily perceived by a control engineering or even predict machinery damage or needed maintenance based on events of the past.

It is based on this premise of Smart Sensing that multiparameter sensors, an important role as the fundamental key element of a sensor network, can uplift traditional manufacturing companies to the competitiveness of the XXI century.

1.2.   Motivation

Liquid analysis and assessment of its physical properties and/or its chemical and biological constituents have an important society role, which is transversal to many scientific research areas. The ultimate example is water in its natural state which its quality guarantee for healthy consume is a global concern. Similarly, the drinks and beverages industries have also the caution to guarantee that the production of its water based products are between the desirable quality parameters, not only in its final process, but also during its production. Other non-water based liquids like oils and its derivatives need also regular quality analysis processes.

The most common methods for chemical and biologic analysis are chromatography based systems [Plumb, 2001; Žuvela, 2019], liquid chromatography with mass spectroscopy [Hernández, 2005; Alvarez-Rivera, 2019] and/or optical spectroscopy [Soloman, 2010; Pasquini, 2018], which generally require collecting a sample and transport it to a laboratory. These systems are usually of big dimensions and/or have high response times, incompatible with the necessity of immediate and remote analysis of liquids. In the field of the physical commercial sensing solutions, capable of measuring parameters as *e.g.* pressure, refractive index, turbidity or temperature, are usually electronic devices that can be more or less portable but are limited to a single parameter measurement of a collected and disposable sample. As alternative, the use of fiber optic based sensing is a rising possibility, which uses the fiber as a light waveguide for the interaction with the phenomena to detect. There is an extensive literature dedicated to the development of optical sensors for the detection of chemical/biologic [Wang, 2013; Yin, 2018; Gandhi, 2019] and physical [Joe, 2018; Kadhum Hisham, 2018] parameters. These fiber optic based sensors have the advantages of miniaturization and higher transportability; low cost, mainly the polymeric based optical fiber (POF); fast response; remote and real-time monitoring; multiplexing; immunity to electromagnetic interference and passive operation, being possible to be used in flammable liquids.

Additionally, they are also a viable option for smart sensing applications, where the permanent and systematic collecting of data and surrounding information is a requirement. This will then allow to generate well-founded knowledge which could serve as support for decision making process. However, if the collected data is not well fused, processed and analyzed, a loss can occur because of the non-timely decision making.

Computational methods of pattern recognition and anomaly detection are being applied with success in different scientific fields [Rudin, 2012; Ben Mabrouk, 2018; Dhiman, 2019] being effective

in the data management and analysis processes. This field of knowledge is designated as Machine Learning and its development begun since the start of computation, being increasingly relevant with the higher available computation power [Alpaydın, 2010; Portugal, 2018]. Machine Learning algorithms can be categorized in two distinct ways depending if there is or not previously labeled data. A supervised algorithm determines relational models of cause-effect based in labeled data. The learned model is then used to catalog new data whose cause is unknown [Alpaydın, 2010; Schrider, 2018]. The most common algorithms of this category are: Decision trees learning and Random forests [Kotsiantis, 2013; Torizuka, 2018]; Artificial neural networks [Georgieva, 2013; Cao, 2018]; Bayesian statistics [Parslow, 2013; Salmerón, 2018]; k-nearest neighbors [Suominen, 2013; Triguero, 2019] and Support Vector machines [Orrù, 2012; Zendehboudi, 2018]. The unsupervised learning algorithms are used for non-labeled data and goes by clustering these unlabeled data, searching for patterns or anomalies and associate them with the provoked effects [Alpaydın, 2010; Fan, 2018]. Examples of this algorithms are clustering analysis [Sridhar, 2012; Lakshmanaprabu, 2018] and Hidden Markov models [Fink, 2014; Calvo-Zaragoza, 2019]. Algorithms of both categories can also be combined to create more powerful prediction and analysis tools [Alpaydın, 2010].

Data Fusion techniques are also important tools used to combine and analyze data and information from multiple sensors in order to form a more accurate measurement which can be applied in different processing levels [Azimirad, 2015; Ghamisi, 2019; Nweke, 2019]. Several algorithms can be used for this propose mostly based in Probabilistic Bayesian rules, Fuzzy Logic and filtering models [Raol, 2016; Ghamisi, 2019].

Variations of these algorithms with constant improvements can be found in literature depending of the application from which they have been developed. Each of these algorithms will have its advantages and disadvantages relative to the data to be analyzed, without existing one that universally supplant the others. The development of this algorithms and their application on data obtained from optical sensors can be an important tool for detection and monitoring of different liquids in a smart sensing perspective and an important key element for a sensor network.

1.3. Objectives

The global objective of this thesis is the development of an innovative multiparameter optical based sensing platform for liquid monitoring. Although there are several good developed optical devices to measure different parameters individually, there are still problems to overcome in this field: there isn't a lot of solutions for in-line production measurements, the developed sensors that

could be adapted for an in-line measurement have low specificity when subject to a non-controlled conditions outside a laboratory or very specific environment and, because of this, the analysis of the collected data is of difficult interpretation and correlation leading to erroneous decisions.

What is addressed in this thesis is the suppression of these problems through the development of the multiparameter platform together with a data analysis system based in algorithms of data fusion and/or machine learning.

The specific objectives can be divided in the following main topics:

• Conceptual development of a low-cost optical platform that can measure simultaneously, remotely and in real time, different parameters of interest to quantify the quality of liquids. A broad study to the necessary physical concepts of the target parameter sensors will be made and studied its integration to the whole platform. The parameters to be addressed are: turbidity, refractive index (RI) and color. The specification of the sensors will be established considering the parameters to measure, sensitivity, selectivity, detection limits and the operation range and type of liquid to measure. The optical components and mechanisms to measure the desired parameter will be selected based on its potential performance and low-cost implementation, in both fabrication and detection phases. This will cut out the use of expensive implementation optical components like Bragg gratings or plasmonic resonance based surfaces. Simple phenomena like absorption, scattering, refraction and reflection of light will be used. Small versions of the setup as well as bigger versions will be made to understand the scaling likelihood. The source of light and its application to the platform will also be studied with the use of LED and/or lasers. Different liquids will have different optical signatures which will be able to evaluate the proposed parameters.

• Assembly of the prototype platforms and its experimental characterization. The conceptual designs will be developed to real prototypes with different configurations ready to be characterized. Exhaustive laboratory tests will be performed using different liquid solutions that will simulate different optical signatures which includes diverse colors, turbidities with different compounds and refractive index, as well as combinations of them at different temperatures. The conceptual sensibility, selectivity, detection limits and operation range will be tested in practice. This information will be essential to the following algorithm development.

• Development and implementation of data fusion and machine learning algorithms based on the physical modulation of the system, with the capacity of discrete and/or continuous classification in the detection of the physical quantities and/or anomalies and with innovative

approaches to data processing techniques for the optimization of the selectivity and detection limits of the sensor. The algorithms must be able to classify or assess parameters from previously learned signatures. They will be developed and optimized based on the previous studied physical models (Figure 1.1). Overall, the algorithm must run fast or in real time so it can provide objective information with easy interpretation for decision making and enable that a single data server can handle Big Data processing for several sensing platforms running simultaneously. On the other hand, it must be flexible so it has the capacity of incremental data addition in the event of the expansion of sensor elements. Machine Learning techniques will be tested for measurements based on signatures from previously learned data (Figure 1.1). The selection of the best algorithm will be based from both supervised and unsupervised learning using regression models and clustering techniques, linear discriminant analysis (LDA) and artificial neural network (ANN). Iterative data fusion algorithm will be also implemented using the information obtained from the sensor characterization (Figure 1.1) and will be based in a combination of weighed arithmetic mean (WAM) and fuzzy logic rules (FLR). All will be tested and compared by cross validation error evaluation, processing time and identified their advantages or disadvantages in relation to the traditional direct measuring methodology. Two programing languages will be used for benchmarking, Matlab and Python.
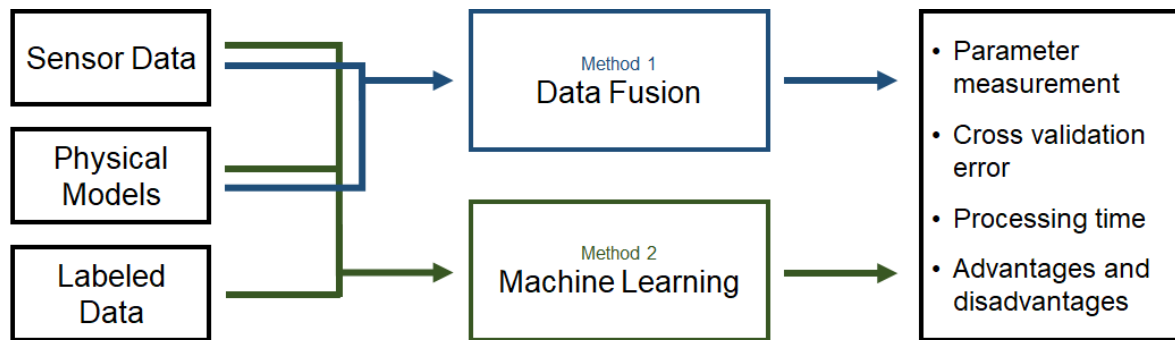
Figure 1.1 – General scheme of the methodologies to be applied to the sensing data obtained from the sensor and its topics of comparison.

• Application of the developed platform and algorithms in case studies involving the monitoring of wine fermentation/production and the assessment of cleaning solutions in industrial cleaning machines. An evaluation of potential application of the sensing platform and developed algorithms will be made to two distinct case studies. The first case is based on wine analysis with the study of red, *rosé* and white wine fermentations, pressing of grapes and ageing where the monitoring of color and turbidity is tested. This objective will have the collaboration of Watgrid by adding the developed prototype as a complement to their temperature, level and density sensors

and by using their online platform for a cloud based communication and data visualization. The sensor will be placed in real wine vats provided by the collaboration with several wine producers in Portugal and abroad. The second case will assess the quality of a cleaning solution used in the maintenance of industrial machinery by washing process. It will try to detect the optimal threshold from where the cleaning solution is not efficient anymore, that can be used for triggering an alarm indicating the necessity to change the liquid.

1.4.   Thesis Outline

The structure of this thesis is based on six distinct chapters where the prior specified objectives will be discussed:

Chapter 1 discusses the motivations, objectives and structure of the thesis, where the reasons that lead to the study of this subject are discussed. It is also presented in this chapter the main contributions that this study provided to the scientific and industrial community.

Chapter 2 reviews the state-of-art of liquid evaluation sensors comparing its advantages and disadvantages for each measurement parameter of interest, as well as its physical principle of measurement. The parameters of interest will be turbidity, color and refractive index.

Chapter 3 presents the design of the proposed sensor and its physical principle with theoretical expectations described. Concepts of reflection, refraction, transmission, dispersion and fiber optics are discussed. Preliminary tests of the sensor with different elements and configurations are tested in laboratory with controlled environments containing variation of the parameters of interest as well as variations of other sources that can influence the same parameters. Error measurement analysis will also be made.

Chapter 4 focus on the sensor multivariate data analysis where a state-of-art review of machine learning and data fusion algorithms will be made. A broad look to the pretended data analysis objectives will be discussed. A deeper approach of each technique will be made where dimensional reduction, machine learning algorithms and data fusion techniques (LDA, ANN, clustering, regressions) are implemented and compared with each other.

Chapter 5 presents the results of the real industrial application of the different optic platforms to a wine fermentation and cleaning solution case of study. Here an introduction of the both environments and obstacle measurement peculiarities is made and the final measurement results are presented.

Chapter 6 overviews the developed work and presents the main conclusions with future perspectives of the monitoring platforms.

1.5.   List of publications

All the work produced during the elaboration of this thesis has resulted in a book chapter, 8 articles submitted and published in international peer review journals and 10 conference proceedings. The list of these publications is presented below and its association with the main achievements of this thesis is presented in chapter 1.6. All the remaining publications do not cover exclusively the work here presented.

1.5.1.   Book chapter

(BC1)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2019), "Optical Fiber Sensing Principles" (chapter 3) in Plastic Optical Fiber Sensors: Science, Technology and Applications, *CRC Press,* 1st Edition, pp. 67 - 91*,* ISBN: 9781138298538

1.5.2.   Journal articles

(JA8)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2020), "Turbidity and RI dependency of a POF based chromatic sensor", *Sensors*, vol. 20, num. 19. DOI: 10.3390/s20010019

(JA7)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2020), "Low cost color assessment of turbid liquids using supervised learning data analysis – Proof of concept", *Sensors and actuator A: Physical*, vol. 305, pp. 111936. DOI: 10.1016/j.sna.2020.111936

(JA6)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2020), "A low-cost liquid refractive index sensor based on plastic optical fibre and CCD array", *Measurement Science and Technology*, vol.31, num. 4, pp. 047001. DOI: 10.1088/1361-6501/ab46ff

(JA5)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2019), "Semi-supervised Gaussian and *t*-distribution hybrid mixture model for water leak detection", *Measurement Science and Technology*, vol. 30, num. 12. DOI: 10.1088/1361-6501/ab3b48

(JA4)   F. Sequeira; Daniel P. Duarte; R. Nogueira; L. Bilro (2019), "Low-Cost Sensing with Plastic Optical Fibers—From Turbidity and Refractive Index to Chemical Sensing", *Proceedings*, vol. 15, num. 12. DOI: 10.3390/proceedings2019015012

(JA3)   F. Sequeira, Daniel P. Duarte, L. Bilro, A. Rudnitskaya, M. Pesavento, L. Zeni, N. Cennamo (2016), "Refractive Index Sensing with D-Shaped Plastic Optical Fibers for Chemical and Biochemical Applications", *Sensors*, Vol. 16, No. 12, pp. 2119 - 2119. DOI: 10.3390/s16122119

(JA2)   Daniel P. Duarte, N. Alberto, L. Bilro, R.N. Nogueira (2015), "Theoretical design of a high sensitivity SPR based optical fiber pressure sensor", *Journal of Lightwave Technology*, Vol. 33, No. 22, pp. 4606 – 4611. DOI: 10.1109/JLT.2015.2477353

(JA1)   N. Alberto, C. Vigário, Daniel P. Duarte, N. Almeida Almeida, G. Gonçalves, J. L. Pinto, Paula A.A.P. Marques, R.N. Nogueira, V. Neto (2015), "Characterization of graphene oxide coatings onto optical fibers for sensing applications", *Materials Today: Proceedings*, Vol. 2, No. 1, pp. 171 - 177, June, 2015. DOI: 10.1016/j.matpr.2015.04.019

1.5.3.   Conferences

(C10)   R. N. Nogueira, L. B. Bilro, F. Sequeira, R. Oliveira, Daniel P. Duarte (2019), "Liquid quality assessment with optical fibre sensors" (invited speaker), *28th International Conference on Plastic Optical Fibers,* Yokohama, Japan.

(C9)   L. Bilro, N. Oliveira, Daniel P. Duarte, F. Goncalves, P. Costa, R. Pereira, R. Nogueira (2019), "Wineplus 1110: the remote and real-time wine fermentation process monitoring system", *42nd World Congress of Wine and Vine*, Geneve, Switzerland, PO-261. ISBN: 978-285-038-0105

(C8)   Daniel P. Duarte, Rogério N. Nogueira, Lúcia Bilro (2017), "Portable, in-line and low cost refractive index sensor for liquids", *International Conf. on Plastic Optical Fibers - POF*, Aveiro, Portugal.

(C7)   N. Oliveira, Daniel P. Duarte, F. Gonçalves, P. Costa, S. Vieira, N. Fontes, P. Prior, A. Graça, M. Figueira, L. Bilro, R. Nogueira (2017), "Winegrid®: the remote and real-time wine production process monitoring system", *40th World Congress of Vine and Wine*, Sofia, Bulgária. ISBN: 979-10-91799-76-8

(C6)   F. Sequeira, Daniel P. Duarte, R.N. Nogueira, A. Rudnitskaya, N. Cennamo, L. Zeni, L. Bilro (2016), "Analysis of the roughness in a sensing region on D-shaped POFs", *International*

Conf. on Plastic Optical Fibers - POF, Birmingham, United Kingdom, pp. 148 - 156. ISBN: 978-1-85449-408-5

(C5)   N. D. Oliveira, Daniel P. Duarte, C. Ferreira, P. Silva, R. N. Nogueira, L. Bilro (2016), "Development and characterization of a sediment concentration low cost optical sensor", *25th International Conference on Plastic Optical Fibers POF 2016*, Birmingham, United Kingdom, pp. 287 - 290. ISBN: 978-1-85449-408-5

(C4)   F. Sequeira, Daniel P. Duarte, A. Rudnitskaya, M. T. S. R. S. R. Gomes, R. N. Nogueira, L. Bilro (2016), "Ammonium sensing in aqueous solutions with plastic optical fiber modified by molecular imprinting", *Sixth European Workshop on Optical Fibre Sensors EWOFS 2016*, Limerick, Ireland, Vol. 9916, pp. 99161I-1 - 99161I-4. DOI: 10.1117/12.2236941

(C3)   Daniel P. Duarte, N. D. Oliveira, P. Georgieva, R.N. Nogueira, L. Bilro (2015), "Wine classification and turbidity measurement by clustering and regression models", *10th Conference on Telecommunications Conftele*, Aveiro, Portugal.

(C2)   Daniel P. Duarte, S. Prats, J. J. Keizer, P. Georgieva, R.N. Nogueira, L. Bilro (2015), "Novel approach for simultaneous sediment classification and concentration determination of water turbidity", I*nternational Conference on Optical Fibre Sensors 24th International Conference on Optical Fibre Sensors OFS24*, Curitiba, Brazil, Vol. 9634, pp. 96342U-1 - 96342U-4. DOI: 10.1117/12.2194578

(C1)   Daniel P. Duarte, N. Alberto, L. Bilro, R.N. Nogueira (2014), "Theoretical modeling of an U-shaped SPR fiber sensor in 1550-nm spectral range for sensing applications", *II International Conf. on Applications of Optics and Photonics - AOP*, Aveiro, Portugal, Vol. 9286, pp. 928640-1 - 928640-5. DOI: 10.1117/12.2063724

1.6.   Main achievements

In the opinion of the author, the important contributions reported and developed in this work with both scientific and industrial significance are the following:

- Extensive review of the state of the art regarding the color, turbidity and refractive index sensors in both commercial and academic literature (BC1).

- Low-cost prototypes capable of measuring simultaneously turbidity and color of a solution in in-line and on-line applications with wireless connectivity for data visualization in cloud services (JA4, JA8, C5, C10).

- Deep prototype characterization and physical simulation to simultaneous external parameters' variations necessary to understand and develop posterior effective auto-compensation routines for parameter discrimination (JA4, JA6, JA8).

- Multivariable machine learning and data fusion algorithms capable of a better discrimination of the measuring parameters (color and turbidity) and measurement error reduction, presenting each one its advantages and disadvantages according to the final application (JA5, JA7, C2, C3).

- A low-cost refractive index sensor for in-line and on-line applications with wireless connectivity for data visualization, based on the prism total internal reflection (JA6, C8, C10).

- Validation of the developed prototype sensors in industrial applications such the real-time monitoring production and fermentation of wines and the assessment of the quality of cleaning solutions (C7).

2. Optical sensor concepts

2.1.   Introduction

Quality control is the fundamental principle of production that any company that manufactures for a developed country need to have to ensure, not only that all requisites of safety are conformed with the high regulated laws of the country, but also that the manufactured product is strictly within the parameters expected to be sold for the consumer. This is true not only for the general manufacturing world but, more importantly, to the food, drink and beverages industry. Production flaws can cause bad reputation or, even worse, health issues, which could lead to the total shutdown of that company. Liquid properties can be indirect indicators of the chemical or microbiological condition of the product. Product quality is often related to the concentration of an important ingredient or an unwanted chemical or biological contaminant like toxins or bacteria [Figura, 2007]. Nonfood industries of water based solutions, like cleaning products, liquid detergents, leach and oil derivatives have also their specific quality control policies. It is also because of safety consume and environmental concerns that any location containing a source of water or water flowing, needs to have its quality regularly controlled, principally if it's the source of where the water distribution system is capturing water for the nearby populations. Even indirect indications of changes to the surrounding environment morphology like river margins and shores could be assessed by monitoring the water streams [Gippel, 1995; Fondriest Environmental Inc., 2014]. In general, these liquids present suspended particles of solid matter with different sizes at some degree. If the liquid is constantly in a stirring flow, the particles will never settle down, which will make the liquid appear not clear or turbid. Turbidity is one of the physical parameter to assess the quality of a liquid and can be related to the Total Suspended Solids (TSS). Other physical parameters are color, density, refractive index, temperature, viscosity and conductivity which can be directly or indirectly relevant for quality monitoring, this is, two or more parameters can be used to assess a related parameter that would be more difficult to directly analyze [Figura, 2007].

2.1.1.   Turbidity

The turbidity parameter is based on the amount of light scattered by particles in a determined liquid column. More presence of particles means more light that will be scattered. While turbidity

and TSS are related, it is not a direct measurement of the total suspended materials in a liquid but instead a measure of relative clarity [Anderson, 2005]. For nearly direct relation, the size and shape of the suspended solids would need to be uniform, which in almost all cases it will not be. Instead turbidity is often used to indicate changes in the total suspended solids concentration without providing an exact measurement of solids. Nevertheless, a correlation between turbidity and TSS can be made by a linear regression modeled by two coefficients denoted by A_{TSS} and B_{TSS}:

$$NTU = A_{TSS}(TSS)^{B_{TSS}}$$ *(Eq. 2-1)*

The most accurate method of determining TSS is, however, by filtering and measure the mass of the particles, which is reported in milligrams of solids per liter (mg/L). Turbidity, on the other hand, is mainly reported in units designed as Nephelometric Turbidity Units (NTU). Nephelometric refers to the technology method used for measuring and that requires a photodetector to be placed at a 90° angle from an illumination source. As light bounces off the suspended particles, the photodetector can measure the scattered light. Other common used unity is Formazin Nephelometric Unit (FNU) if a turbidity meter has an infrared output, as opposed to the broadband light output. NTU and FNU will behave similarly when calibrated with formazine because of its common nephelometric technology, but in the field, they may operate differently due to the different light sources used and particle sizes present. FNU based sensors will be colorblind for liquids that do not absorb infrared, while NTU based sensors, using its broadband output, will have light absorbance phenomena for colored liquids [O'Dell, 1993; International Organization for Standardization, 1999; Anderson, 2005; Fondriest Environmental Inc., 2014]. Different units will be associated with this technique if other methods are used which are different of the nephelometric approach, like light attenuation, surface scatter and backscatter (Table 2.1).

Table 2.1 – Turbidity units used when measured with different design instrumentation.

Instrument Design	Unit
Nephelometric non-ratio turbidimeters	NTU
Ratio white light turbidimeters	NTRU
Nephelometric, near-IR turbidimeter, non-ratiometric	FNU
Nephelometric, near-IR turbidimeter, ratio metric	FNRU
Surface scatter turbidimeters	SSU
Near IR or monochrome light back scatter turbidimeter	FBU
Backscatter turbidimeter	BU
Near IR or monochrome light attenuation turbidimeter	FAU
Light attenuation turbidimeter	AU
Formazine nephelometric multibeam turbidimeter	FNMU
Nephelometric laser-diode turbidimeter	mNTU

Commercial turbidity meters are usually based in nine standards used for monitoring drinking water that are approved by the United States of America (USA) Environmental Protection Agency or by the International Organization for Standardization (ISO). The most common standard in USA is the EPA Method 180.1 that uses NTU units [O'Dell, 1993]. This standard uses nephelometric technology with the light scatter detector centered at 90° angle without extending more than 30° from that center point and with the light path lower than 10 cm (Figure 2.1). Other photodetectors can be used on other angles but the 90° must be the primary one. The light source used is a polychromatic tungsten lamp with a color temperature between 2000 K and 3000 K with a spectral peak response between 400-600 nm. This type of turbidity meter will measure turbidity levels between 0-40 NTU and have a resolution of 0.02 NTU, being only adequate for treated water monitoring. The Standard Method 2130B is nearly identical as the EPA Method 180.1 with differences only on the calibration methodology and in the maximum range of measurement where dilution of samples must be done for turbidity values over 40 NTU [Greenberg, 1992].
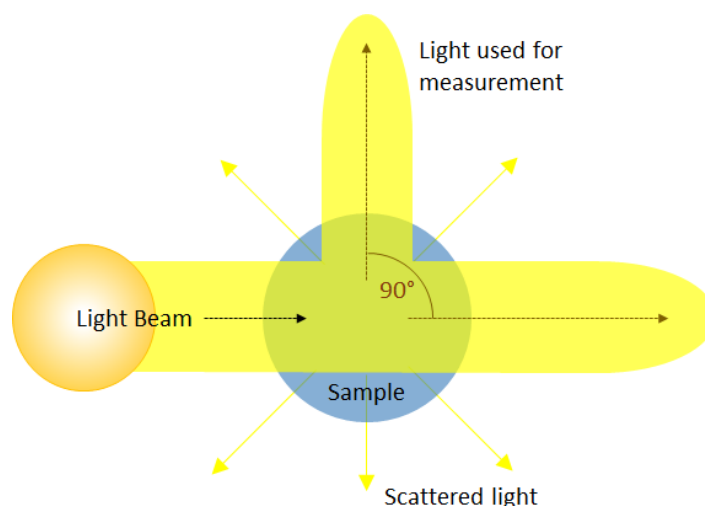
Figure 2.1 – Instrumental schematic of the EPA Method 180.1 measurement method with its polychromatic light source.

The ISO 7072 standard is the most used in Europe and was developed by the International Organization for Standardization. This standard is intended to have a high repeatability and comparability degree. The most notorious requirement is its monochromatic light source with a wavelength of 860 nm and a spectral bandwidth of 60 nm which eliminates most color interferences. Most instruments in compliance with this method use a LED light source and a single nephelometric 90° angle photodetector (±15° margin) with maximum 10 cm path, but additional detection angles are allowed provided that the 90° is the primary (Figure 2.2). Turbidity level between 0-40 FNU will be measured with this method, with dilutions needed for higher values. When multiple detectors are used (ratiometric), units in Formazin Nephelometric Ratio Units (FNRU) should be used [International Organization for Standardization, 1999].
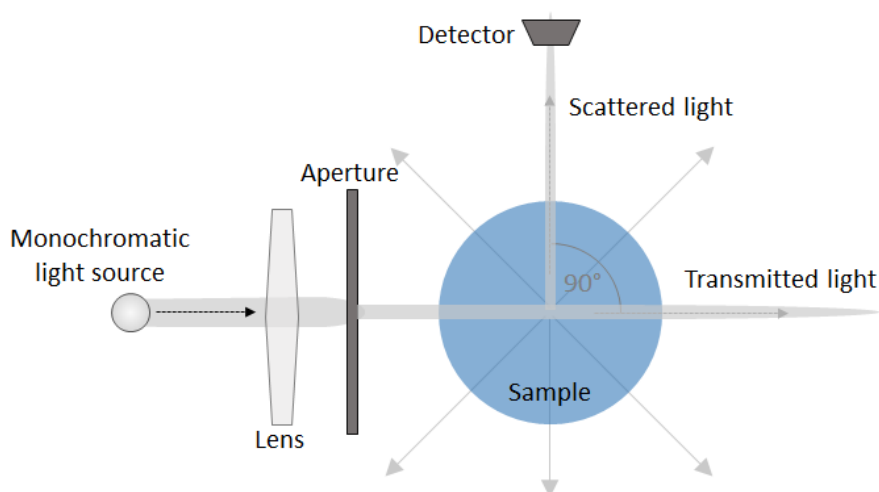
Figure 2.2 – Instrumental schematic of the ISO 7072 standard measurement method with its monochromatic light source.

Other standard is the Great Lakes Instrument Method 2 (GLI Method 2) which uses the double of photodiodes and photodetectors to calculate the average turbidity value of four measurements, which are the reason for this design to also been known as a modulated four-beam turbidity meter. Because it still uses 90° nephelometric technology it is yet considered a nephelometric turbidimeter. Two 860 nm LEDs are the light sources required for this standard. The LEDs alternate light pulses every half second. A simultaneous "active" and "reference" signals are taken by the detector directly across from the active LED and the detector at a 90° angle respectively (Figure 2.3). In a second cycle, two active and reference signal are taken [Gli international, 1992]. A ratiometric calculation is performed to determine turbidity since the light input and the output are directly proportional, ensuring that any errors that may appear are mathematically canceled out [Hydrolab, 2018]. The ratiometric calculation for turbidity determination is also dependent on the calibration coefficients C_s and C_0:

$$FNMU = C_s * \sqrt{\frac{Act_1 * Act_2}{Ref_1 * Ref_2}} - C_0 \qquad \textit{(Eq. 2-2)}$$

With Act_1, Act_2, Ref_1 and Ref_2 the active and reference signals from the detector 1 and 2 respectively. FNMU stands for the Formazin Nephelometric Multibeam Units which are the units recommended for this standard. The GLI2 method has increased sensitivity and error cancellation in the 0-100 FNMU range but loses accuracy as turbidity levels rise above 40 FNMU. It is however especially accurate for measurements in the 0-1 FNMU range.
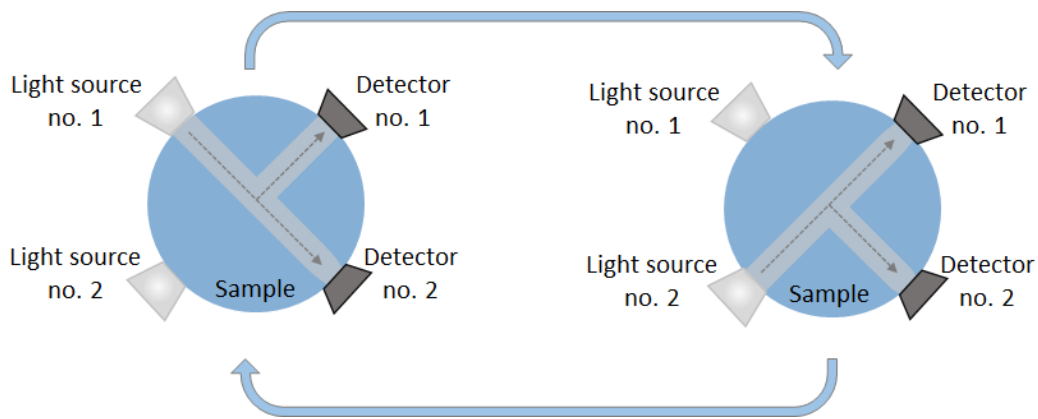
Figure 2.3 – Instrumental schematic of the GLI Method 2 standard method using its two-step cycle for the ratiometric measurement calculation.

Hach Method 10133 is a standard that also uses nephelometric technology but has a laser light source with wavelengths between 630 nm and 690 nm. The 90° light receiver is connected to a photomultiplier tube (PMT) via a fiber-optic cable (Figure 2.4). The PMT will have an increase in

sensitivity and is not recommended to measure values above 5 NTU. This setup is an in-line or *in-situ* process stream method as it is used in the environment to be measured, instead of collecting a sample cell. The units for this method are expressed as milli-Nephelometric Turbidity Units (mNTU) [Hach Company, 2005].
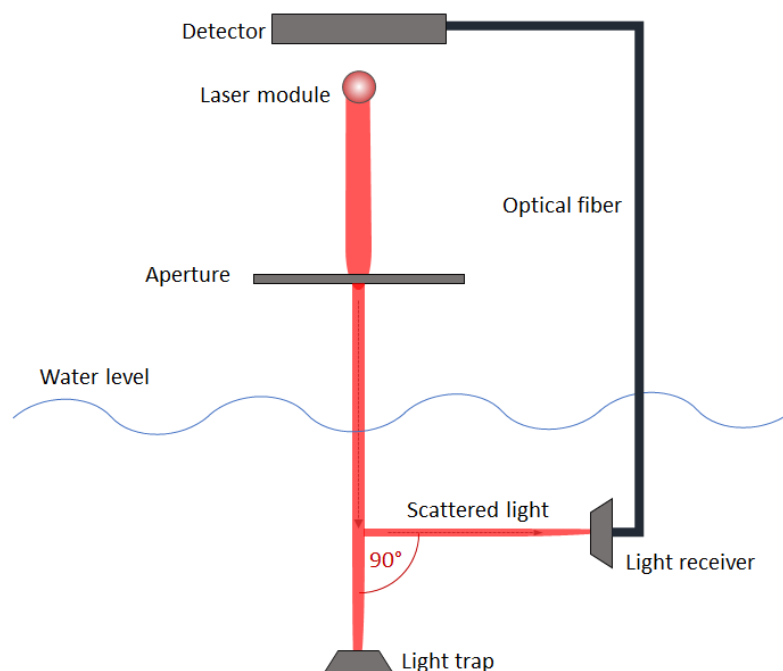
Figure 2.4 – Instrumental schematic of Hach Method 10133 method that can be used as an in-line or in-situ process stream measurement.

More uncommon are the standards Mitchell Methods M5271 and M5331 [Mitchell, 2009a, 2009b; Lambrou, 2010] which are based on the Hach Method 10133. M5271 requires a laser with 650 ± 30 nm, a maximum 10 cm light path and a nephelometric detector. The biggest difference is in the demanding technology for compliance with the standard, as it requires a bubble trap and anti-fog windows. M5331 has the same requirements but instead it uses a LED source with wavelength of 525 ± 15 nm. Both are designed as on-line continuous turbidity monitoring.

The Orion Method AQ4500 is another uncommon standard since it was developed by Thermo Scientific and is based on the use of their turbidity meter, Thermo Orion AQUAfast Turbidimeter Model AQ4500. It is similar with the EPA Method 180.1 except for the specified light source that uses a phosphorus coated blue LED that has a broad spectrum instead of the polychromatic tungsten lamp. The LED is rapid pulsing operation which permits synchronous detection where any stray light or electronic-induced errors can be reduced and nearly cancelled out. It also uses a

photodetector in 0° of the light source (direct transmission) which tries to reduce color errors due to color absorption using its signal as reference [Scientific, 2009] (Figure 2.5).
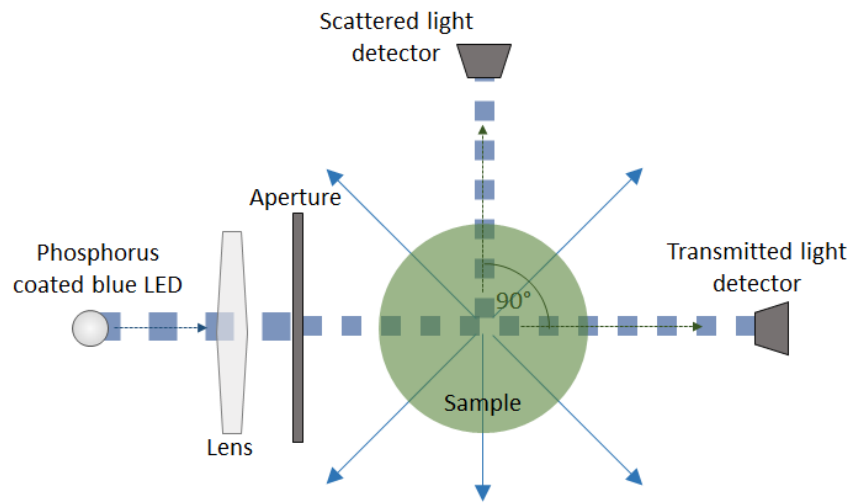
Figure 2.5 – Instrumental schematic of the Orion Method AQ4500 standard using its light rapid pulsing operation for synchronous detection.

The last standard is the AMI Turbiwell which is a non-contact turbidity meter. This design has a LED light source with a spectral response between 400 and 600 nm. The focused light beam has a 45 ± 5° angle with water surface. A beam splitter is used to deflect a portion the light beam before it hits the water which is used as reference signal to the monitoring light intensity. The primary photodetector is set at a right angle to the light source and an algorithm then determines the turbidity levels based on the light intensity of the scattered and reference signals. Again the total distance traveled by the light beam should not exceed 10 cm with turbidities up to 200 NTU being detected [SWAN Analytische Instrumente AG, 2009].

While commercial turbidity meters are mostly handheld solutions that do not permit on-line *in-situ* measurements, the existing real-time probes that has these capability, can be really expensive in comparison to the handheld counterpart, with prices ranging from around 700 € in the handheld version, to the 1500 € of the in-line probes. In academia, low cost methods to measure turbidity in water samples are constantly being developed. A first attempt to create a portable low cost and *in-situ* suspended solids sensor was reported by Campbell et al. in 2005 [Campbell, 2005]and Tran et al. in 2006 [Tran, 2006] where a fiber optic in-stream transmissometer was developed, consisting of linear and circular fiber bundles to transmit and receive light from a 640 nm wavelength LED source and a photodiode as detector (Figure 2.6). The bundles permit to capture more light than with a single fiber, especially when short diameter core fibers are used, and even have spatial

resolutions if the linear bundle is used. The produced electric signal from the photodiode was still dependable on a personal computer for data processing.
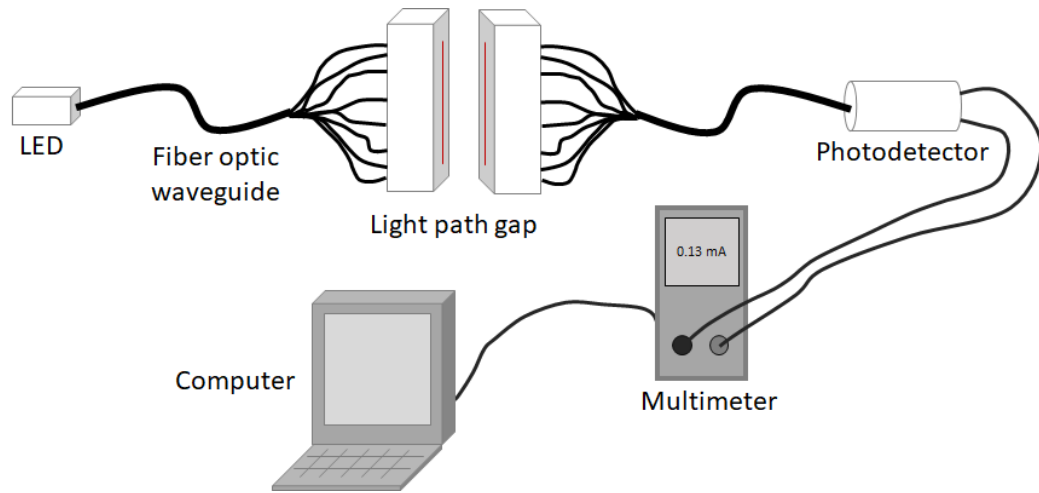
Figure 2.6 – Schematic diagram of a setup using the fiber bundle and measurement components (adapted from [Campbell, 2005]*).*

In 2008 it is also reported by Omar and MatJafri [Omar, 2008] four different turbidity sensor systems using a single plastic optic fibers (POF) with 1 mm core diameter to guide light from 470 or a 633 nm LED to a measurement cell where a water sample is placed. Depending on the system used, light from nephelometric 90° or 0° is captured to photodetectors by other fiber in which light intensity will vary with the turbidity level. In 2010, Bilro et al. [Bilro, 2010] uses these concepts to create a sensor with both 90° or 0° angles for light detection and takes the advantage of POF miniaturization and transportability to develop a potential in-line solution that does not requires sample holder. This system uses a red 660 nm LED (Figure 2.7 a). In the same year, Lambrou et al. [Lambrou, 2010], develops a fully *in-situ* low component cost turbidity sensor based in the Mitchell Method M5271 having a red 670 nm laser source and a LCD reader and wireless transceiver. A range of 0 - 100 NTU with precision of 0.2 NTU was reported. Both sensors are color sensitive and though unfitted for colored liquids which is the major drawback for these light sources, mainly when the desired liquid to measure is not water. A true color insensitive *in-situ* turbidity sensor was developed by Garcia et al. [Garcia, 2008] and Crespo et al. [Crespo, 2010] in 2008 and 2010 respectively, having the same principle as Bilro's proposal but in this case an Infra-red (IR) LED was used so it can be color insensitive. In 2015 a low-cost on-line and *in-situ* sensor proposal was made by Murphy et al. [Murphy, 2015], where an LED array light source was used together with two photodiode detectors for 90° and 0° detection angles with electronic control and GSM communications (Figure 2.7 b). The sensor LED array have five wavelengths: IR (850 nm, FWHM

45 nm), red (627 nm, FWHM 45 nm), amber (583 nm, FWHM 36 nm), green (515 nm, FWHM 30 nm) and blue (430 nm, FWHM 60 nm) and each of light sources are tested for different turbid solutions independently. Although this have the potential to have higher resolution because of the use of different LED light information, it was not made any correlation between the obtained data of the different wavelengths from the study.
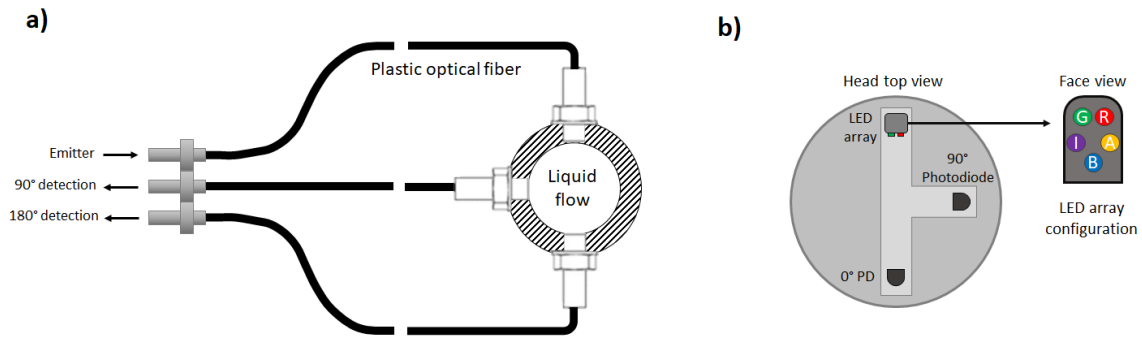
Figure 2.7 – a) Schematic diagram of a sensor head based on plastic optical fiber for the light waveguide (adapted from [Bilro, 2010]*). b) Design of a turbidity sensor head using a LED array of 5 different wavelengths (adapted from* [Murphy, 2015]*).*

A more recent proposal using multimode fibers was presented in 2019 by Yeoh et al. [Yeoh, 2019] It attached side by side two fibers with their beveled tips mounted vertically and due to evanescent light transference between the two fibers in the turbid medium, light intensity variations were observed. However low accuracy values were obtained.

Recently, other proposals are also appearing in academia with the same low-cost objective goal, but without the remotely on-line sensing perspective. One of this proposals is the affordable open-source handheld turbidity sensor presented by Kelley et al. [Kelley, 2014]. The sensor used an 860 nm LED light source and a light-to-frequency detector in a 90° placement. It presented measurement results comparable to the commercial handheld sensors using only 4% of the building cost. Other proposals take advantage of our daily used smartphone capacities to determine turbidity, mainly the use of the integrated cameras and communication system. Näykki et al. reported a measuring system comprising a container that is filled with the water to be measured [Näykki, 2014]. This system is based on the comparison of intensities of light measured over black, white and grey target areas at two depths of the container. The basic assumption is that the target areas at both depths have the same downwelling irradiance at the surface. The smartphone takes a photo of this target areas through a hole in the lid and sends it to a server for image processing which sends back the results to the smartphone (Figure 2.8 a). The range of

operation is between 1.7 – 7.0 FNU with 4% uncertainty. Other smartphone approach was proposed by Hussain et al. where a system with a sample holder and IR LED is coupled with the smartphone [Hussain, 2016]. This system uses the smartphone battery for powering the IR LED that illuminates the sample holder for a nephelometric 90° measurement that is done with the smartphone IR sensor (ambient light and proximity sensor). An application then reads the detected scattered light and displays its value (Figure 2.8 b). This proposal has an operation range of 0 – 400 NTU and 0.1 NTU sensitivity is achieved.
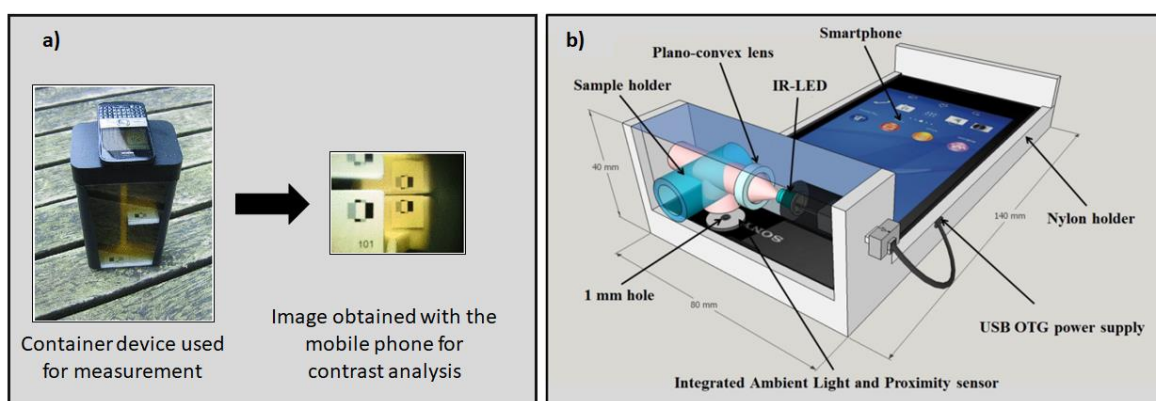
Figure 2.8 – a) Measurement setup of the proposed low cost turbidity sensor using a smartphone and a comprising container for contrast analysis (reproduced with permission of the rights holder, Elsevier [Näykki, 2014]*). b) Integrated sample holder turbidity sensor approach for smartphone low cost measurement (reproduced with permission of the rights holder, The Royal Society of Chemistry* [Hussain, 2016]*).*

2.1.2. Color

When color determination is necessary for quality control, the most common method resorts to laboratory analysis consisting of pre-centrifugation of the sample, for suspended solids removal, followed by spectroscopy measurements [Fang, 2006; Gandul-Rojas, 2013; Burns, 2015]. This type of analysis requires sample extraction from the line of production, which will be ultimately wasted. Depending on where the laboratory is and the type of sample, its transportation and analysis could be costly and time consuming. Spectrophotometers are also costly equipment because of its complex mechanism to measure light intensity through a vast range of light wavelengths. In liquids, spectrophotometers measure the transmittance or absorptance as a function of the wavelength of the incident light and needs several common components to accomplish this task, which are shared between the various spectrophotometer designs and that are analogous to the vision system's perception of color (Figure 2.9). The first of this component is the light source which requires to have stability, brightness and the spectral content desired to be measured. For measurements in the visible region of the spectrum, tungsten and halogen light sources are usually used. Other

component necessary is a scattering element that separates the transmitted light into its constituent wavelengths. Started to be prisms, usually nowadays these elements are diffraction gratings. The evaluation of the quality of the scattering is assessed by the efficiency of the system to reject undesirable wavelengths and for the degree of dispersed light that is distributed around the wavelength of interest. This dispersed light then needs to be detected which is usually done by a photomultiplier tube detector for higher resolution of a single wavelength or an array of photodetectors such as a charge-coupled device (CCD) to measure the whole spectrum at once. Ranges up to 5 absorption units (Abs) can be obtained before the need of dilutions to be performed [Owen, 2000; Tkachenko, 2006].
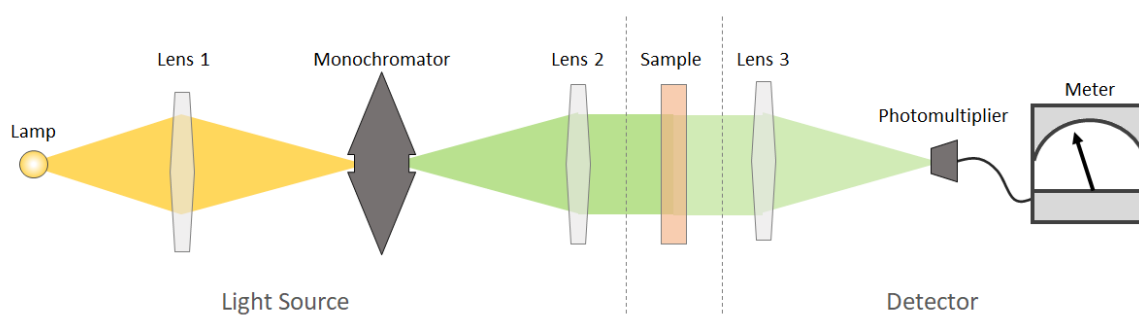
Figure 2.9 – Typical spectrophotometers design scheme with the necessary light source and detection components for transmittance or absorptance measurements.

Simple or more complex designs as the examples of single and double channel schemes, can be used whereas if more stability, resolution or speed of measurement is necessary, but overall, they are complex and costly, with prices reaching the 5000 €. On the other hand, all the information regarding the color spectrum of the liquid may be not necessary but only the transmittance or absorbance of a particular wavelength that could identify or monitor some chemical or biological sub products within the production line, as the example of the growth of a bacterial or yeast culture.

Colorimeters for liquid applications are portable transmission and absorption meters that differentiate from the spectrometers with its fixed light wavelength analysis. They are suitable for *in-situ* fast monitoring, requiring, as well, the sample extraction to a sample holder, being only adequate for clear liquids. Because of its fixed wavelengths it will not have the accuracy to determinate a general color like the spectrophotometer, but will be better adequate for assessment of color differences and quality control [Marcus, 2014]. The more common fixed wavelengths are situated in the red, green and blue region of the spectrum to simulate the human eye perception of color, with the light sources being LEDs and the detection performed by a photodiode. Since it uses lower cost components and have a simpler measurement scheme, this device is way cheaper

than the spectrophotometer and have maximum ranges on average to 2 Abs units. The only in-line commercial application for color measurement observed is a photometer with similar principle as a colorimeter which uses pre-selected LEDs from the manufacturer as light sources. The light is then guided with optical fiber to the side of a steal ring that functions as a measurement cell and where the solution for monitoring will flow. The transmitted light is then captured in the other side of the ring with other optical fiber which will guide it to a detector. A range of 5 Abs units is possible to be obtainable depending on the light chosen path length which is related to the diameter of the steel ring. Turbid liquids will affect the measurements. This solution can be more costly than a spectrophotometer [Kemtrak, 2016].

In academia, there are some approaches related to the assessment of liquid color that tries to create low-cost, portable and ease of use devices that were thought to be used for education proposes, but that could be used in practical applications as well. This is the case of a very simple sample holder based colorimeter, constructed using LEGO bricks that has a single wavelength LED as light source [Asheim, 2014] or a more complex colorimeter for chemistry laboratories that uses three LEDs of red (640 nm), green (524 nm), and blue (470 nm) wavelengths. It has an integrated LCD to present the transmittance values of each LED [Clippard, 2016]. This was the approach also done by Jiménez-Márquez et al. for a low-cost solution of color determination sensor to be used for the monitoring the maceration of red wine (Figure 2.10), where 420, 520 and 620 nm wavelengths are used in this industry to assess the anthocyanins and their flavylium combinations and grape types by using the Glories method [Jiménez-Márquez, 2015]. The transmitted light of the three LEDs of the wavelength of interest are disposed along the measurement cell and the measurement of the transmittance values are independently determined by a row of three photodiodes disposed in the same way as the LEDs. A range up to 1 Abs unit and an accuracy of ±0.003 Abs units was obtained. Other approach presented by Zhou et al. [Zhou, 2014] uses filters in the light path between the white LEDs sources and the sample holder instead of different color LEDs. A miniaturized On-Chip colorimeter is presented by Liu et al. [Liu, 2016] which costs two orders of magnitude less than the commercial solutions and proves that this devices can be very compact with dimensions of 6 cm x 4 cm x 6 cm. Smartphone based color determination sensor was also developed by Sumriddetchkajorn et al. [Sumriddetchkajorn, 2014] for chlorine concentration assessment where a self-referencing analysis is done for converting the color level of water to its corresponding chlorine concentration. A portable closed chamber is used as the support structure for the water sample and smartphone, with the concentration calculated from the ratio between

the color intensity of the water sample and the empty sample holder. This process requires prior reagent mixing for chlorine activation.
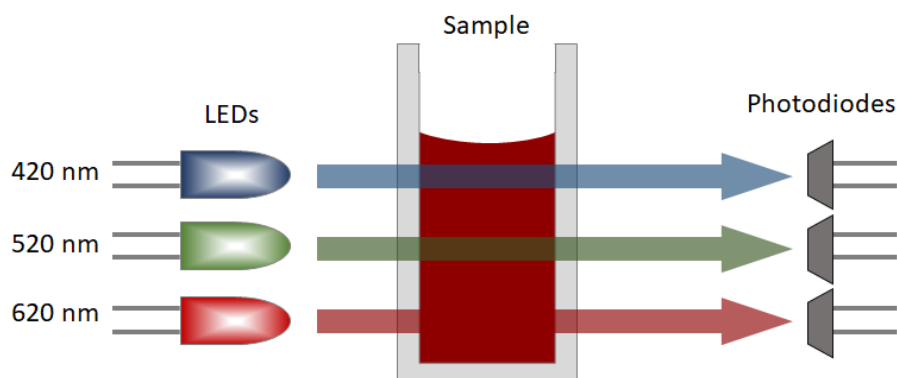
Figure 2.10 – Schematic approach for a color determination sensor using an LED array of different colors and light intensity detection with an array of photodiodes after passing through the sample to measure (adapted from [Jiménez-Márquez, 2015]*).*

Both commercial and academia solutions presented here have mostly in common its unique suitability for sample holder measurements. With the exception of the previous presented in-line commercial photometer based on the steel ring sensing, the only *in-situ* and on-line measurement device found in literature ready for color measurement was presented by Novo et al. [Novo, 2013, 2014] and Shrake et al. [Shrake, 2014] for wine monitoring. Novo uses the advantages of POF miniaturization and maneuverability to create four separate air-gap regions between two fiber extremities in a sensor head structure and where the wine to be measured will flow (Figure 2.11 a). One of the fiber will be connected to an LED and the other to a photodetector. The transmitted light will then be converted to digital signal in a main controller board with Bluetooth connection. Blue (470 nm), green (530 nm), red (660 nm) and IR (870 nm) LEDs are used for Glories method calculation. Shrake also uses POF for light guiding of 420, 470, 503, 525, 565 and 630 nm LEDs but this light is directly captured by photodiodes instead of a fiber. UV 280 and 320 nm LEDs are also used for phenolic compounds assessment but, because POF have high absorbance in the UV light, the LED are directly mounted on the flow cell instead of using intermediary fibers (Figure 2.11 b). This in-line system has a prior 2.0 µm pore size membrane filter to restrict the influence of the suspended particles and all the measured samples are then wasted like in the sample holder methods.
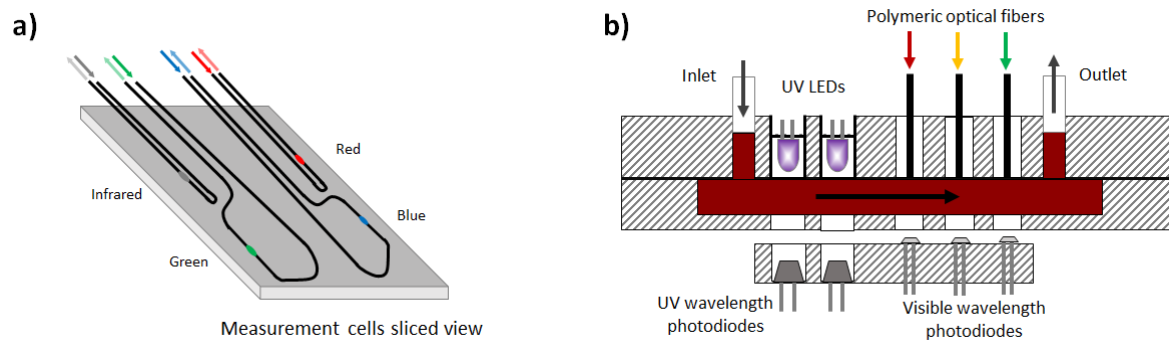
Figure 2.11 – a) Scheme of an in-situ on-line measurement device for wine monitoring using air-gap regions and polymeric fiber optic (adapted from [Novo, 2014]*). b) Alternative approach for an in-line measurement device which includes UV light (adapted from* [Shrake, 2014]*).*

Although there are developed solutions for color monitoring like the ones here presented, these devices have a major handicap of needing cleared liquids for color determination and, for that reason, they are unfitted to a direct measurement of turbid liquids without a prior centrifugation or filtration. There is no solution that addresses this problem in the literature.

2.1.3. Refractive Index (RI)

Refractive index (RI) is represented as a dimensionless number that describes how light propagates through a medium and thus is an important parameter of characterization of that material. The refractive index can be seen as the factor by which the speed of radiation with a certain wavelength are reduced with respect to their vacuum values. This reduction of velocities will have an effect on the refraction or bent and on the reflection of light between material interfaces depending on the angle of light incidence. In complex fluids, such as drinks or beverages, the refractive index is normally measured for assessment of dissolved or submicronic materials like sugar concentration which has a related scale called Brix. In other industries, it is also an important propriety to measure oil/water ratios, glycol/water ratios like in antifreeze, and inaccessible liquids such as the electrolyte of rechargeable cells [Meeten, 2014].

2.1.3.1. Commercial prism based sensors

The first refractometer industrialized and sold commercially in 1881 was developed by Ernst Abbe with a design that is still in use today, known as Abbe refractometer. The Abbe refractometer uses two prisms where the sample is placed between. One of the prisms will act as illuminating path and the other as the measuring device. Light from a polychromatic source enters the illuminating prism until it reaches the sample where refraction and total reflection will occur at critical angle. The light will then exit the measurement prims to a scaled telescope that rotates, with the objective

of searching for a dark and illuminated boundary that will be observed in a position that depends on the sample refractive index (Figure 2.12) [Tarrant, 2010]. Because it uses a light transmission design, measurements will be affected by the liquids properties like strong optically dense absorbing samples or samples containing bubbles or suspended solids which can lead to low contrasts of light and high difficult to measure. The commercial refractometers using this design are also prepared for solid measurements and have temperature control, increasing its costs compared to handheld devices, presenting prices around 650 €. This refractometer usually covers a refractive index wide range between 1.3 to 1.7 refractive index units (RIU) and a resolution in the order of magnitude of 10^{-4} RIU.
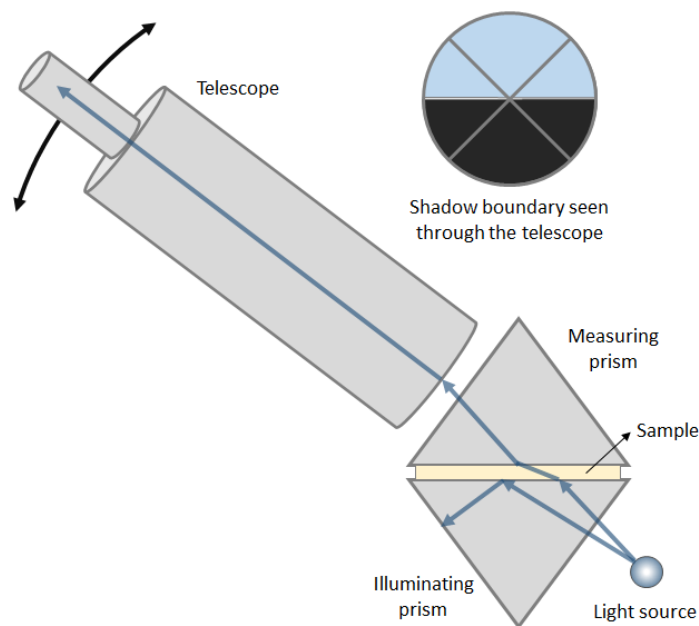
Figure 2.12 – Schematic setup of a basic Abbe refractometer under operation. The telescope is adjusted to find the position of the shadow boundary observed.

For high resolution measurements, with precision in the order of magnitude of 10^{-4} to 10^{-5} RIU, laboratory digital refractometers are usually used as the preferred commercial application. This highly cost refractometers (around 6000 €) cover a range of refractive index, typically between 1.3 – 1.6 RIU. The principle of operation of this type of refractometers are based on the determination of the critical angle of where total reflection of light occurs in the interface of two media. These two media are composed by the liquid of which RI we wanted to determinate and a prism that will act as sample holder. A light source, usually a long-life LED, is focused onto a prism surface via a lens system. Depending on its refractive index, the incoming light below the critical angle of total reflection is partly transmitted into the sample, whereas for higher angles the light is totally

reflected. This difference in reflected light intensity is detected by a high-resolution CCD sensor array. Depending on the position of the totally reflected light intensity, the refractive index of the sample can be determined (Figure 2.13). In contrast to the Abbe refractometer design, it is possible to measure the refractive index of optically dense strongly absorbing samples or samples containing air bubbles or solid particles. Mostly of these laboratory refractometers also have temperature control option by resistors and Peltier coolers [Meeten, 2014; Anton Paar, 2016b; Reichert Tecnologies Analitical Instruments, 2019].
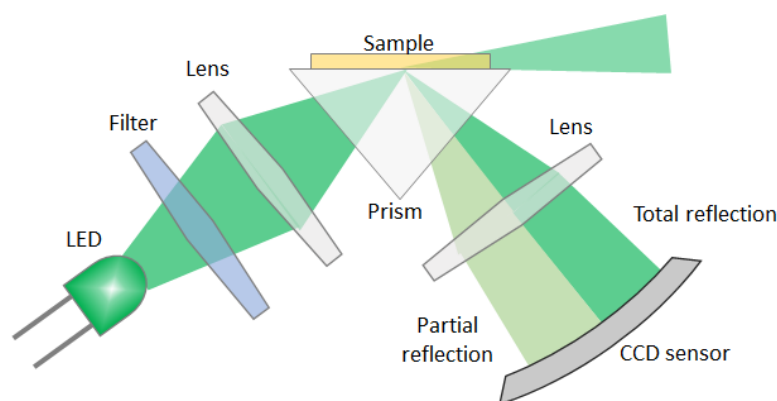
Figure 2.13 – Schematic diagram of a common bench refractometer based on the prism total reflection of light.

Commercial digital handheld devices are also based on the same principle by their bench counterparts but, because of its lower dimensions, it has less resolution (10^{-3} to 10^{-4} RIU) and do not have temperature control, but usually presents temperature compensation. They also may require a slightly larger amount of sample to read from, since the sample is not spread thinly against the prism. On the other hand, they are cheaper, with prices around 200 €, and portable which can make sample measurements on-site [Mettler-Toledo GmbH, 2014] (Figure 2.14 a). The cheaper handheld refractometer although is the traditional analogue one (20 €), which projects a shadow line onto a small glass reticle inside the instrument and is then viewed by the user through a magnifying eyepiece. The sample is placed between a small cover plate and the measuring prism. Light traveling through the sample is either passed through to the reticle or totally internally reflected. A shadow line appears between the illuminated area and the dark area and it is where it crosses the scale that a reading is taken (Figure 2.14 b). Depending on the liquids color and turbidity intended to measure, this line can be difficult to resolve. Most of these handheld devices are calibrated for indirect measurements like sugar content (Brix, Klosterneuburger scale, Oechsle scale, etc.), alcoholic content and salinity [A.KRÜSS Optronic GmbH, 2016].
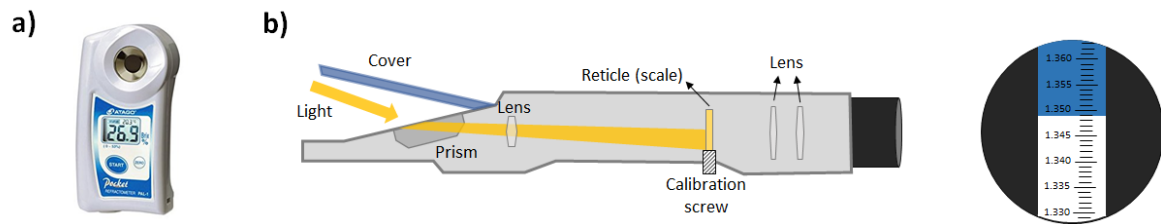
Figure 2.14 – a) Typical commercial digital handheld refractive index sensor device. b) Schematic diagram of the physical principal of an analog refractometer and its visualization through the eyepiece in the right.

In-line refractometers are also available in the market based on the same principle of digital bench refractometers and therefore insensible to color, turbidity and bubbles. This type of refractometers is usually manufactured ready to be installed in pipes. They have also high resolution in the magnitude order of 10^{-5} but measuring ranges with a maximum at 1.53 RIU which is the 100% Brix equivalent. To achieve this features, this type of refractometers present the highest cost price of the market with a value around 7000 € [Maselli Misure, 2017; Anton Paar, 2019].

2.1.3.2. Fiber extrinsic and intrinsic based sensing

In the quest for alternative low-cost ways of refractive index measurement, different approaches can be find in the academia literature, mostly using glass and plastic optic fiber as the transducer base element. The most common method of detecting is based on intensity variation which requires simple operation principles and instrumentation. In general, only a light source, a fiber and a photodetector is needed, which can be low cost elements like LEDs and photodiodes. This enables the manufacture of easy implementations devices with high portability. Refractive index POF based sensors can be divided in two broad sensing categories which are characterized by light leaving the fiber to an external medium (extrinsic sensing) or by staying within the fiber (intrinsic sensing) [Bilro, 2012] (Figure 2.15).
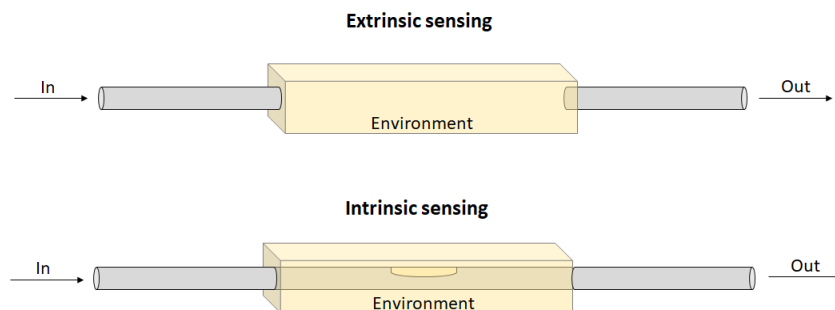
Figure 2.15 – Extrinsic and intrinsic sensing scheme where in the former light needs to leave the fiber to interact with the environment before it is captured again and in the latter the light interacts with the environment without leaving the fiber.

One of the extrinsic sensing techniques employed is the transmission configuration. Here the intensity variation is related to the optical signal coupling between two optical fibers. This technique was employed by Noiseux et al. to non-absorbing liquids and named as Air-Gap design [Noiseux, 2004]. It takes advantage of the light opening cone generated by the emission fiber that will have dependency on the medium to measure. The higher the refractive index, the small the cone will be, and by consequence, the higher intensity light will be captured for the same fiber area (Figure 2.16 a). Noiseux used a 0.47 mm gap between the fibers and detected a variation of light intensity of 5% between 1.3326 to 1.3803 RIU. For absorbing liquids like the ones with suspended particles, this design is unfitted to do reliable measurements. A quasi-intrinsic sensing technique was also reported by Shin and Park with its cascaded in-line holes POF for liquid refractive index measurement [Shin, 2015]. The holes are created by micro drilling. While the sensing mechanism is done in-line, light propagating in the fiber will technically leave the fiber to interact with the solution flowing in the holes and therefore being considered has an extrinsic sensing mechanism. Shin and Park obtained a sensitivity of 62.9 dB/RIU between the 1.33–1.42 range in a 3-hole structure.

Other type of design is the displacement based sensors that use the reflection of light as the main phenomena (Figure 2.16 b). This design proposed by Binu et al. [Binu, 2009] and Govindan et al. [Govindan, 2009] consists of a two-fiber element submerged in the liquid with the RI to be assessed. One of the fibers will emit light that will reflect in a reflective surface placed at some distance to the fiber tip. Then the second fiber, that can have a displacement of a certain distance to the first fiber, will receive the reflected light which intensity will depend on the refractive index of the liquid medium. A 13.5% light intensity variation was obtained between a range of 1.3322 to 1.3627 RIU. This design will suffer the same flaws of absorbing liquids as the obtained by the Air-gap design. If only one fiber tip is used to emit and to receive the light that is reflected in the tip itself and in contact with the medium (intrinsic sensing) (Figure 2.16 c), then it is considered a Fresnel based sensor [Kim, 2004; Su, 2007; Nath, 2008]. To achieve high resolutions, this design will have special needs as pulse generators and subsequent computational analysis to separate the components of emitting light and the received light. This will increase the costs in comparison to the simpler designs but can achieve 2.5×10^{-5} RIU of resolution [Kim, 2004].
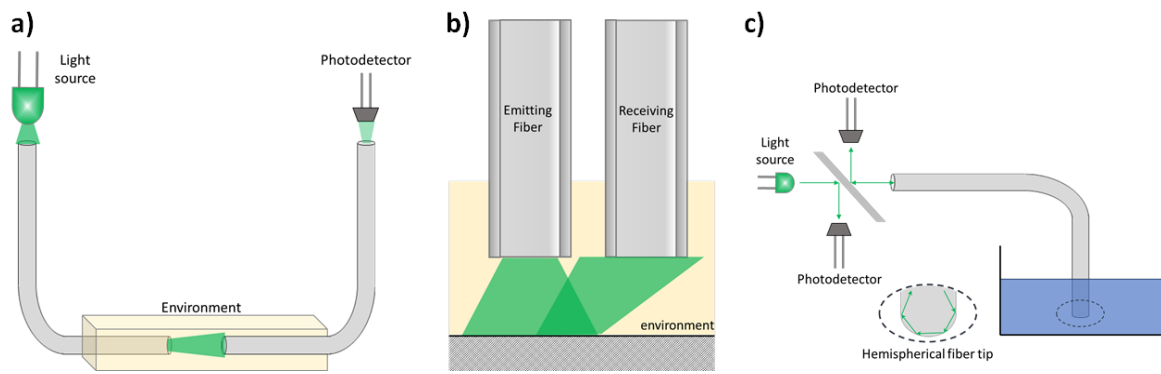
Figure 2.16 – Schematic setup of the fiber based sensors with: a) air-gap design; b) displacement fiber and c) Fresnel based hemispherical fiber tip (adapted from [Nath, 2008]*).*

Other intrinsic sensing technique also reported in the literature to measure refractive index uses the physical fiber geometry manipulation, this is, a macrobending fiber system to force light to interact with the outside medium and that can, or not, be assisted by side polishing of fiber. Early experimental testing of this concept can be found in the work of Zubia et al. where a step-index POF was laterally polished with a slight curvature until the core were exposed (Figure 2.17 a) [Zubia, 2000]. When submersed, a resolution of 5×10^{-3} RIU was reported for indices of refraction from 1.30 to 1.59. This concept was further studied by Ana Cao-Paz et al. with a high variation in the curvature of the fiber with radius from 0.1 to 1 cm, but without polishing the fiber (Figure 2.17 b) [Paz, 2007, 2010]. Their work was implemented in the measurement of wine fermentation and electrolyte density in lead-acid batteries. Different geometries of side polishing were also used to try to improve results. Bilro et al. has developed an analytical model, verified by experimental evaluation, with a concave shape side polishing in stretched fiber, similarly obtained by Zubia et al. [Bilro, 2011]. The same side polishing configuration was used by Jing et al. [Jing, 2015] but maintaining the curvature of the fiber. Sensitivity of 154 dB/RIU power loss were obtained in a range from 1.33 to 1.44 RIU. In other hand, De-Jun et al. reported a study using a D-shape format with different groove depth and fiber radius and its optimized values in relation to sensitivity (Figure 2.17 c) [De-Jun, 2014]. Sequeira et al. has also studied different lengths of the sensitive D-shaped fiber region obtained by mechanical polishing, obtaining a resolution of 6.48×10^{-3} RIU for the higher length of 6 cm [Sequeira, 2016]. They have also concluded in a parallel work that the sensitivity of the D-shape morphology was also strongly dependent on the roughness of the sensing surface which could be adjusted by polishing the surface with sandpaper having diverse grit sizes. Higher sizes produces higher losses for lower refractive index which produces higher sensitivities when the measured refractive index approaches the refractive index of the fiber cladding [F. Sequeira, 2016]. A similar concept and results were developed and obtained by Teng et al. with its

multi-notched structure imprinted like a LPG with die-press-print mechanism [Teng, 2017]. To increase sensitivity to refractive index measurements, the structure was used in a U-shaped configuration. Sensitivities of 1130%/RIU were obtained in changes of transmittance (resolution of 8.44×10^{-4} RIU) in a range of 1.33 to 1.41 RIU.
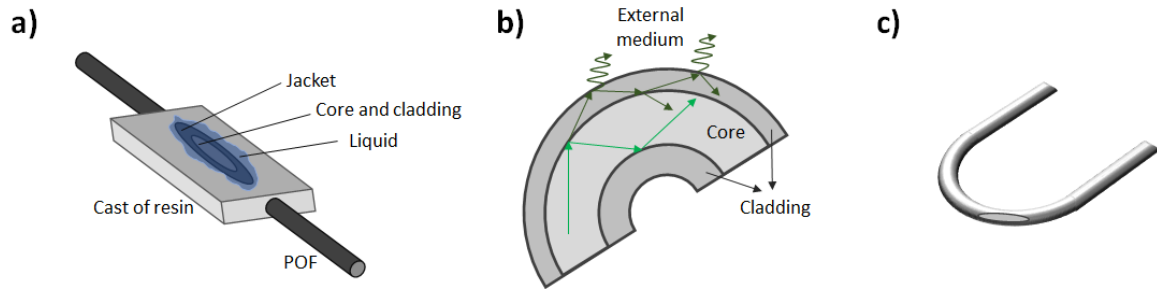
Figure 2.17 – a) Setup of the macrobending polished fiber intrinsic sensing in a cast of resin (adapted from [Zubia, 2000]). b) Bending fiber based sensor without the need of cladding removal or etching (adapted from [Paz, 2010]). c) D-shape fiber based structure for sensing that can be designed with different groove depth and radii.

Other fiber morphology changing, also commonly used as intrinsic sensing, is based in tapering the fiber to promote the interaction of the evanescent light wave with the external medium. This technique can be found applied to glass optical fiber from the early years of fiber sensing [Leung, 2007], but direct contact with the external medium using glass based materials makes this technique quite fragile. To overcome this issue, Polynkin et al. reported a simple measuring submicrometer taper glass fiber immersed in a transparent curable soft polymer, to guarantee higher robustness and measurement as a microfluidic channels (Figure 2.18 a) [Polynkin, 2005]. A resolution of 5×10^{-4} RIU was obtained. Plastic optical fiber is a more flexible material but has low ductility, which can be a difficulty factor to create tapers on this type of fibers, using the standard methods as used in glass fibers. A theoretical, numerical, and experimental, analysis of optical fiber tapering in POF was reported by Xue et al. [Xue, 2007], which become the foundation of where other authors developed further research in this sensing technique. Arrue et al. have reported the analysis of tapered graded-index polymer optical fibers for refractive index sensing and proven this concept for different narrowing ratios highlighting the advantages compared to other sensing mechanisms and materials as glass fiber (Figure 2.18 b) [Arrue, 2008]. A clear description how to fabricate taper POF was reported by Gravina et al. [Gravina, 2009] and several sensors for refractive index measurements were created with this procedure. As examples there is a salinity detection based sensor [Rahman, 2011], with a resolution of 2×10^{-3} RIU for a variation from 1.334 to 1.3576 RIU; a biosensor testing [Beres, 2011], using taper fiber and bending the taper in a U-shaped format to improve sensitivity (Figure 2.18 d), where resolution in the magnitude 10^{-3} RIU was obtained; a

double tapper sensor (Figure 2.18 c) [De-Jun, 2014] with sensitivity of 950 µW/RIU up to 1.42 RIU; and a cladded removed taper in a u-shaped format (Figure 2.18 e) [Teng, 2016], that has a range measurement to 1.45 RIU and a resolution of 5 x 10^{-4} with a waist diameter of 250 µm. An unconventional away of tapering a POF fiber was also developed by Shimada et al. [Shimada, 2017] by ultrasonically crushing it in a cost-effectively method using a press a horn connected to an ultrasonic transducer. When submerged in a solution for refractive index measurements up to 1.36 RIU, a sensitivity of -62 dB/RIU was obtained. These resolutions, although near, have still higher value number (lower resolution) than the ones obtained by the more traditional methods using a prism.
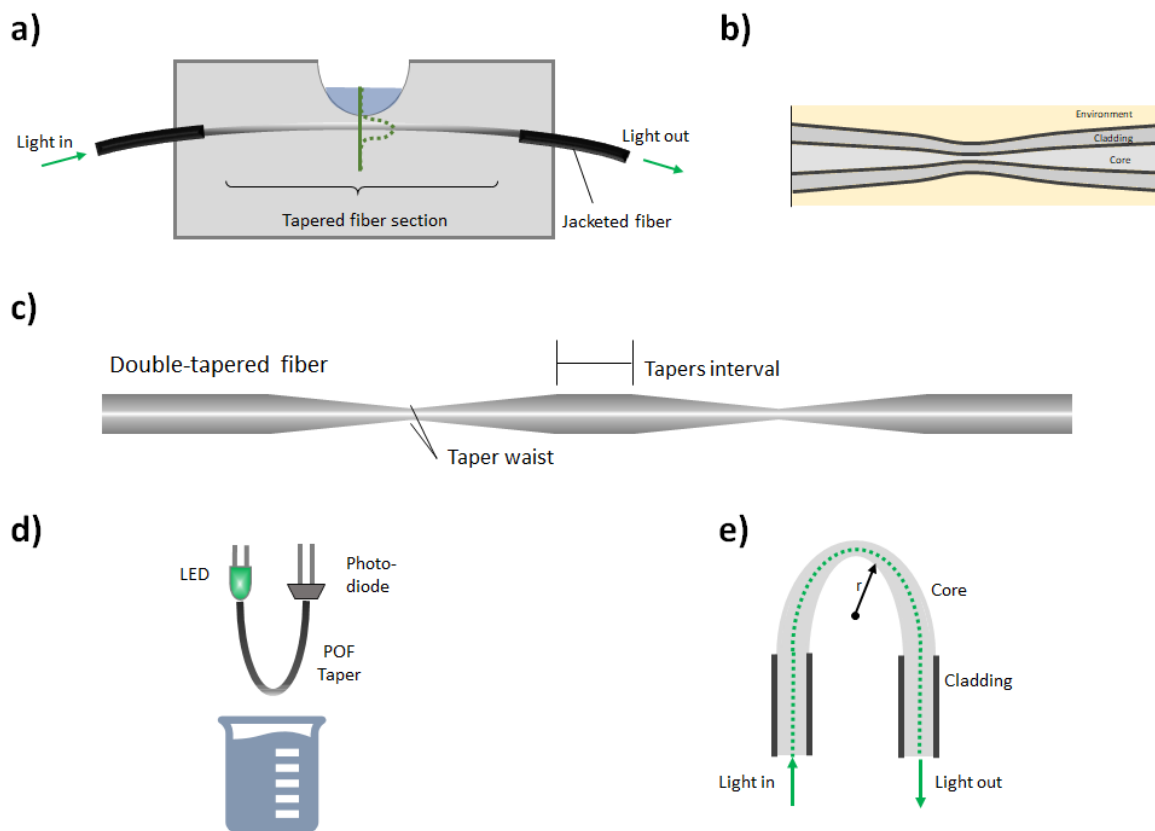
Figure 2.18 – a) Schematic setup of a taper glass fiber immersed in a transparent curable soft polymer for refractive index sensing (adapted from [Polynkin, 2005]). b) Straight plastic optical fiber taper used for RI sensing with its cladding present. c) Double taper fiber approach for increased sensitivity (adapted from [De-Jun, 2014]. d) U-shaped bended POF fiber with a taper for improved sensitivity (adapted from [Beres, 2011]). e) Cladding removed POF in a U-shaped format with taper (adapted from [Teng, 2016]).

2.1.3.3. Gratings and grating based interferometric sensors

Intrinsic sensing can also be done not only by changing the morphology of the fiber but also by modifying the optic properties of an unspoiled fiber. This can be done mainly by altering the

refractive index of the fiber waveguide in certain regions by UV or Infrared-light exposure, usually aided by a phase mask or interferometric methods for patterning. Fiber Bragg grating (FBG) and long period grating (LPG) are technologies that uses this principle by creating a periodic refractive index change in the core of the fiber. The principle of operation of the FBG relies on the dependence of the Bragg resonance with the effective refractive index and the grating pitch of the modified area. This will function as dielectric-mirror for some very specific wavelengths, mainly the Bragg wavelength. Since the effective refractive index is not influenced by the external medium for standard optical fibers, no sensitivity to external refractive index is expected. However, if fiber cladding diameter is reduced along the grating region by usually using chemical etching, the effective refractive index is significantly affected by external refractive index and therefore, shifts in the Bragg wavelength are expected (Figure 2.19 a). This technology usually uses frequency modulation and presents high sensitivity and multiplexing capabilities. An example of application of this technology to glass optical fiber was reported by Iadicicco et al. obtaining resolutions around 10^{-4} to 10^{-5} RIU for refractive index between 1.33 and 1.450 [Iadicicco, 2005, 2005]. When two gratings are consecutively placed at a certain effective cavity length, a fiber Fabry-Pérot interferometer (FFPI) is created which can also be used as refractive index sensor if the cavity is etched to promote interaction with the external medium (Figure 2.19 b). Compared to a single FBG, the FFPI sensors possess narrower resonance peaks and are more desirable for high accuracy wavelength measurement. This approach was reported by Liang et al. [Liang, 2005] and done in glass fiber optic obtaining a resolution of 1.4 x 10^{-5} RIU to a range up to 1.4 RIU. Tapering a fiber is an alternative method to have a low cavity waist for the FFPI. Shao et al. used this approach in multimode glass fiber and obtained a resolution of 10^{-5} RIU (1.33 to 1.42 RIU), observing that the peak intensity decreased with the external refractive index medium [Shao, 2007]. This some approach was reported using LPG instead of FBG by Ding et al. (Figure 2.19 c) [Jin-Fei Ding, 2005]. A peak wavelength shift was noticed and a resolution of 5.8×10^{-6} RIU can be achieved between 1.333 and 1.362 RIU if the measurement system has a normal spectral resolution of 1 pm. This reveal that these systems can achieve higher resolutions than the prism base systems.
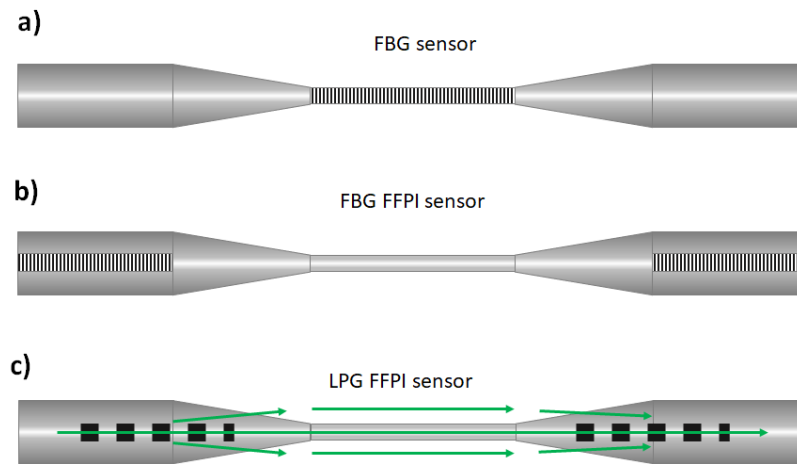
Figure 2.19 – a) Schematic design of a fiber Bragg grating in an etched section of the fiber for refractive index sensing. b) Fabry-Pérot interferometer made with two FBG with an etched cavity to promote interaction with the external medium. c) Fabry-Pérot interferometer made with two long period gratings with an etched cavity (adapted from [Jin-Fei Ding, 2005; Liang, 2005]*).*

A different spectral signature can be obtained if the grating is created with some tilt in relation to the axis of the fiber. Because of this tilt, the light that is reflected in the grating can access the cladding layer, creating new modes of propagation and interact with the external medium (Figure 2.20 a). Miao et al. [Miao, 2009] used this principle in a glass fiber and observed the decrease of the transmitted intensity power with the increase of the external medium refractive index up to 1.4532 RIU, which is expected as it reaches the refractive index of the cladding and thus the modes stop propagating within the cladding. A resolution of 10^{-4} is achieved with this intensity based sensor. While the cladding modes are insensitive to temperature, the core mode is sensitive. This enhances the use of this system as a dual refractive index and temperature sensor as proposed by Alberto et al. [Alberto, 2010]. With its optimization, a resolution of 3 x 10^{-5} RIU was achieved in a range between 1.415 to 1.455 RIU. A more exotic approach is proposed by Cai et al. [Cai, 2015, 2015] by coupling a fiber with a tilted fiber grating with a D-shaped fiber (Figure 2.20 b). The tiled fiber will reflect and couple the light to the D-shaped fiber through the surrounding medium being highly sensitive to its refractive index. A sensitivity of 13000 nW/RIU was obtained in a range of 1.38 to 1.40 RIU.
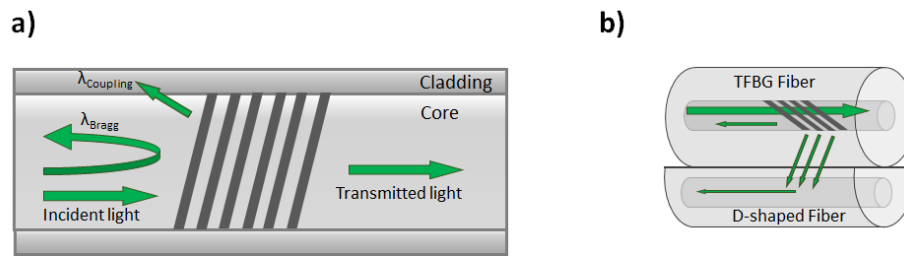
Figure 2.20 – a) Schematic operation of the tilted fiber Bragg grating for RI measuring. The light coupled in the cladding will interact with the external medium (adapted from [Miao, 2009]). b) Double fiber RI sensor design with the light coming from the first fiber TFBG being then coupled in the D-shape second fiber, interacting with the external medium in its path (adapted from [Cai, 2015]).

Although the most common grating inscription is made in glass fiber optics, an increasing trend over the last 4 years is the application of this technology to plastic based optical fibers, which tries to bond the advantages of both technologies. The system that inscribes a FBG into a POF is the same as observed in glass fiber although it must be noticed that low diameter fibers (preferably single-core POF) must be used for lower time UV-light exposure and many several research groups have explored the use of dopants to improve the photosensitivity [Webb, 2015]. All the grating technologies observed in glass based fiber can be applied in POF. Zhang et al. [Zhang, 2012] and Hu et al. [Hu, 2014a] reported an etched FBG refractive index sensor using plastic fiber obtaining a sensitivity of 44 nm/RIU in a range up to 1.372 with a resolution in the magnitude of 10^{-5} RIU if the measurement system has a normal spectral resolution of 1 pm. Fiber Fabry-Perot cavity interferometer was also reported in POF by Ferreira et al. but using the end tip face of a fiber as one of the extremities of the cavity (Figure 2.21) [Ferreira, 2017]. This system presented a resolution of 1×10^{-3} RIU up to 1.47 RIU. Tilted FBG in POF was also reported by Hu et al. and a test in refractive change medium shows a sensitivity of about 13 nm/RIU (7×10^{-5} RIU resolution) in the range of 1.42 to 1.49 RIU [Hu, 2014b].
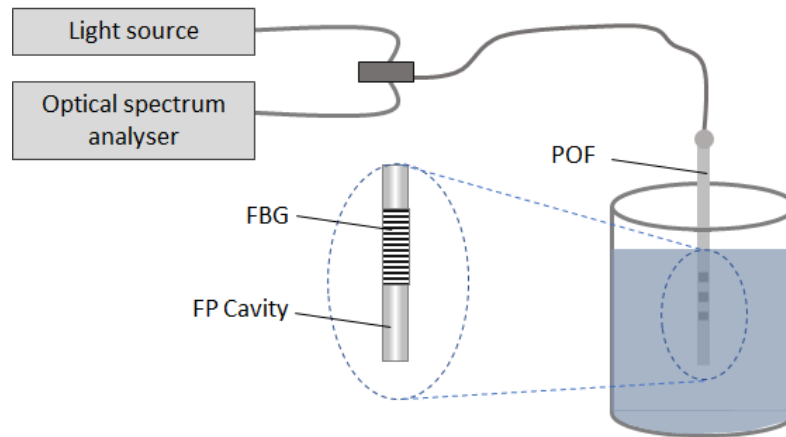
Figure 2.21 – Scheme of a fiber Fabry-Perot cavity interferometer using a FBG and the tip of the plastic optical fiber for refractive index measurement (adapted from [Ferreira, 2017]*).*

As can be seen by the state of the art literature, fiber based grating systems can present considerably higher resolutions for refractive index sensing in comparison to the prism based sensors. Unfortunately, this technology presents majors drawn backs that are preventing higher massification as commercial products. The main problem is the economic factor associated to the inscription of the gratings and detection of light. The inscription system requires a high power and costly laser with temperature control, working with very precise optical components. On the other hand, because most of sensing techniques uses wavelength shift detection, a high precision optical spectrum analyzer is needed. Even the intensity based sensor uses a very narrowband and defined light peaks which will require very precise filters. Overall, the advantages in resolution that these systems provide do not pay off the higher cost that the inscription and detection systems require.

2.1.3.4. Non-grating interferometric based sensors

Other morphology change that can be done to the fiber instead of the inscription of a refractive index change grating, is the introduction of microcavities in the core. This can be achieved by using a femtosecond laser that drills a micro-hole in the center of the single mode glass fiber followed by the splicing of it with another fiber, creating a Fabry-Pérot interferometer. Liao et al. used this method to report an extrinsic refractive index sensor [Liao, 2012]. In addition to the Fabry-Pérot cavity, they also mechanically drilled a hole vertically across the cavity, creating a micro channel where liquids can flow (Figure 2.22 a). By analyzing the wavelength shift of the reflection spectrum with the increase of the refractive index of the liquid, a sensor sensitivity of about 994 nm/RIU was achieved (10^{-6} RIU resolution). Other cost effective solution interferometer was reported by Domingues et al. whom produced a refractive index intrinsic sensor, based in this premise, by using the recycled fibers damaged by the catastrophic fuse effect [Domingues, 2016]. This effect is

characterized by the optical fiber continuous self-destruction produced by the vaporization of the fiber due to high optical power density. The RI sensor head was produced by cleaving the fiber along the voids cross-section and spliced to a standard single mode fiber, which allows the creation of a larger cavity than the initial void and thus not requiring an expensive femtosecond laser during the process (Figure 2.22 b). The other side of the fiber will act as the second mirror of the Fabry-Pérot interferometer. This sensor head creates an interferometric pattern that when submerged in different refractive index liquids have a modulation on its reflection spectra that presents a notorious period decrease as the solution RI increases. By determining the modulation period, the estimated proposed resolution of this sensor is about 3×10^{-4} RIU. Other cheap method to create spherical microcavities can also be obtained by splicing a single mode fiber with a photonic crystal fiber (PCF) which is a fiber that is produced with hollow cores (Figure 2.22 c). A Fabry–Pérot interferometer fiber sensor was created by Wang and Wang using this technique [Wang, 2012], having at one extremity, the microcavity, and at the other, a polished side that will be in contact with the solutions to measure. Two peaks in reflection spectrum were detected where one of the peaks is sensitive with the refractive index of the medium, but the other does not change and consequently used as reference. By calculating the amplitude ratio between the peaks in a range of 1.33 to 1.45 RIU, a resolution of about 1.2×10^{-5} RIU was obtained. Milenko et al. used the same approach but calculated the refractive index change by wavelength shifting, obtaining a resolution of 2×10^{-3} RIU in a range of 1.33 to 1.38 RIU [Mileńko, 2012].
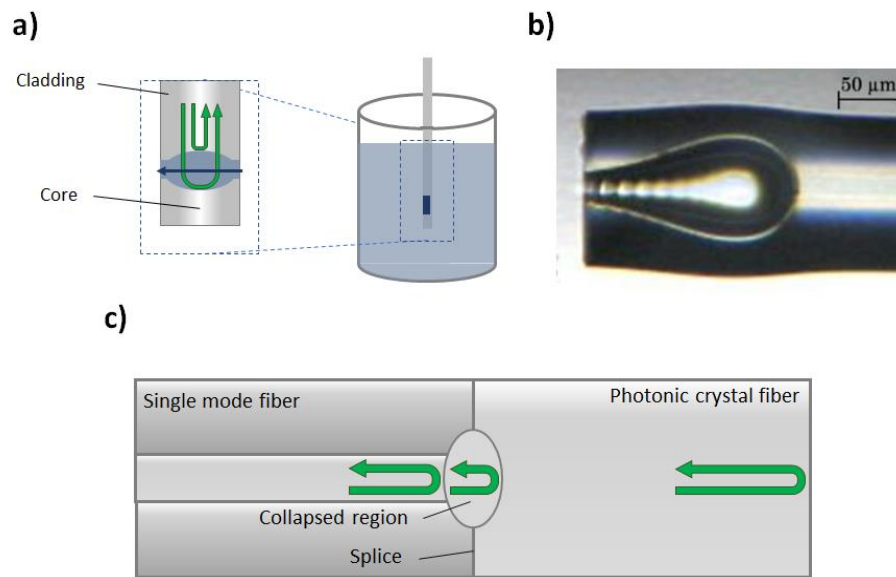
Figure 2.22 – a) Fabry-Pérot cavity mechanically drilled for a micro channel liquid flow RI measurement (adapted from [Liao, 2012]*). b) Image of a fiber damaged by the catastrophic fuse effect to create a Fabry-Pérot cavity (reproduced with permission of the rights holder, Elsevier* [Domingues, 2016]*). c) Splicing of a single mode fiber with a photonic crystal fiber (PCF) to create Fabry-Pérot cavity easily (adapted from* [Wang, 2012]*).*

The special morphology of the hollow cores of the photonic crystal fiber allows it to be used in ways that normal fibers cannot, mainly by using the hallow core as a microfluidic channel (Figure 2.23). This can be seen in the work reported by Wu et al. [Wu, 2009] and Wu et al. [Wu, 2014] where a PCF is coupled with standard glass fibers and one or more of the hollow cores are filled with different refractive index liquids. By measuring the obtained spectrum wavelength shift from a Sagnac loop interferometer, a resolution of 4.0×10^{-6} RIU was obtained for RI around 1.333 RIU. The main problem with this system in addition to the spectral analyzer is the requirement of a parallel liquid insertion system to the hallow cores, which makes it an overall complex apparatus. It is also unclear if this system is not affected by color absorbing liquids.
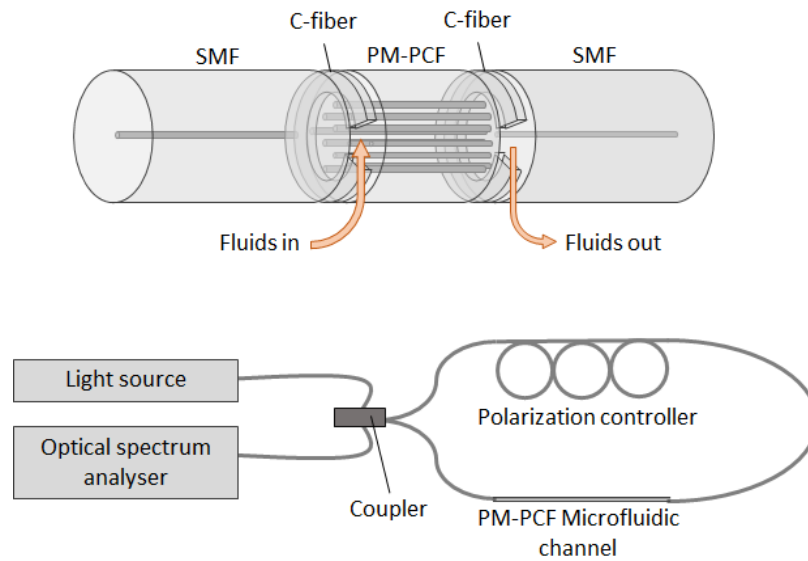
Figure 2.23 – Setup of a hallow core polarized maintaining photonic crystal fiber as a microfluidic channel for RI sensing in a Sagnac loop interferometer (adapted from [Wu, 2014]*).*

As seen before, the fiber Fabry-Pérot interferometer is a common used wavelength transducer sensor that can comprise gratings and microcavities elements to be used as mirrors. An example of Sagnac loop interferometer was also presented using PCF (Figure 2.23). There are however other interferometric systems that can be used for refractive index measurement. An example of Michelson interferometer sensor is reported by Tian et al. [Tian, 2008]. This type of interferometers uses a taper next to a fiber tip to couple part energy of the core mode into the cladding modes that will interact with the external medium. In the fiber tip, a thin gold coat layer is placed to create a perfect mirror. When light travels a second time through the tapper in the reverse order, the cladding cores will be coupled again in the core and an interference pattern will be created. A circulator is then used to guide the light to a spectral analyzer (Figure 2.24). In their sensor, with a tapper diameter of 40 μm, a wavelength shift created by the refractive index variation could lead to a resolution of 5.1 x 10^{-4} RIU in a range up to 1.36 RIU. Bin Ji et al. using exactly the same structure as Tian et al. but with a taper with only 4.6 μm in diameter, achieved a resolution of 5.35×10^{-7} RIU which demonstrates the higher interaction with the outside medium by lower diameters [Ji, 2012].
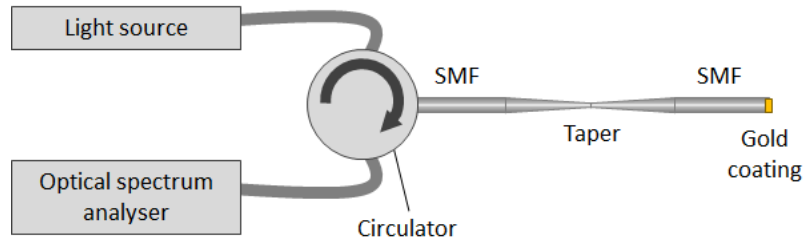
Figure 2.24 – Taper based interferometry RI sensor using a gold coating fiber tip for light reflection.

Much similar to the Michelson interferometer concept is the Mach-Zehnder interferometer. The main difference between them is the method of recoupling the cladding modes that were fomented to appear. While the Michelson method uses a mirror to force the reflected light to recouple in the same fiber element used to couple the cladding modes, the Mach-Zehnder method uses more than one fiber element in-line to couple and recouple the several cladding modes. With this methodology, a circulator is not needed and only the transmission spectra will be analyzed. Examples of Mach-Zehnder interferometer can be seen by the core-offset attenuators that can also be used as a Michelson interferometer if using a mirror (Figure 2.25 a). Tian et al. demonstrated and compared both methodologies using this technique [Tian, 2008]. It is based on the offsetting of two single mode fiber cores by several micrometers. Light from the main fiber will be split into two paths, with a fraction of light remaining in the core of the offset fiber, while the remainder is transferred to the cladding as cladding modes. In the Mach-Zehnder configuration, these modes are then recoupled in the next joint of fibers, located at some distance, creating a pattern when transmission spectrum is analyzed. In the Michelson configuration, a metal mirror is deposited in the offset fiber tip and the interferometer spectrum is obtained as described before for other techniques. Tian el al. only conducted refractive index experiments for the Michelson interferometer obtaining a sensibility for wavelength deep shift of 33,3 nm/RIU (resolution 3 x 10^{-5} RIU) in a range of 1.315 to 1.3618 RIU. Duan et al. [Duan, 2011], using the same Mach-Zehnder configuration of core offsetting, performed the refractive index changing analysis in gas sensing. In a range of 1.0 to 1.0022 RIU a resolution of 6.2×10^{-6} RIU was achieved by wavelength valley shift. Offsetting the fiber core is not the only way to couple cladding modes. It can be also achieved by fusing different types of fiber. Shao et al. used a thin core fiber after the standard single mode fiber and before a multicore fiber, which couples energy to the cladding modes followed for its recoupling to a single mode fiber (Figure 2.25 b) [Shao, 2014]. This Mach-Zehnder configuration reported a resolution of 1.34×10^{-4} RIU in a range of 1.333 to 1.403 RIU. As expected from the Michelson interferometer mentioned above, tapers can also be used for cladding coupling

however, in the Mach-Zehnder configuration, instead of using a single taper with a mirror, it is necessary the use of at least two tapers. When light reaches the fiber taper with larger taper angles, it is split into two optical paths, the core and the cladding, which can be combined again with a similar taper located forward in the fiber waveguide. To increase the interaction of the cladding modes with the external medium, a third taper with lower taper angle is used between the two previous ones (Figure 2.25 c). Zhu et al. used this strategy for its refractive index Mach-Zehnder interferometer sensor [Wu, 2011; Zhu, 2011]. In the range of 1.3320 to 1.3890 RIU, sensitivity of 28.6 nm/RIU (resolution of 4×10^{-5} RIU) was obtained with this configuration. A POF version using only two tapers was also presented by Jasim et al. where refractive index sensitivity of 3.44 nm/RIU was obtained (resolution of 2.9×10^{-4} RIU) in a range from 1.33 to 1.74 [Jasim, 2014]. Using a photonic crystal fiber between two tapers was also a configuration tested and reported by Wang et al. (Figure 2.25 d) [Wang, 2016]. A wavelength shift sensitivity of 260.8 nm/RIU (resolution of 3.8×10^{-6} RIU) between 1.3333 to 1.3737 RIU was achieved. Other type of interference phenomena starting to be used in a higher degree by the community is the fiber multimode interference (MMI) which is usually used with a structure comprising single-mode-multimode-single-mode configurations. The MMI theory can be understood by the phenomenon of self-imaging where an input single mode field profile is periodically reproduced along the multimode fiber by exciting multiple modes with different propagation constants, resulting in constructive interference patterns with maxima or minima in its transmission spectrum. With the cladded removed in the multimode section, usually a POF, this can be used to measure refractive index of an external medium. Oliveira et al. used this structure with single mode silica fibers and a multimode POF obtaining a sensitivity of 193.9 nm/RIU (resolution of 5×10^{-6} RIU) in a range from 1.333 to 1.354 and 515.3 nm/RIU (resolution of 2×10^{-6} RIU) from 1.354 to 1.368 RIU [R. Oliveira, 2017]. Again, the main problem concerning fiber interferometers as refractive index sensors comparing to other sensing technologies, is the necessity to use expensive spectrum analyzers.
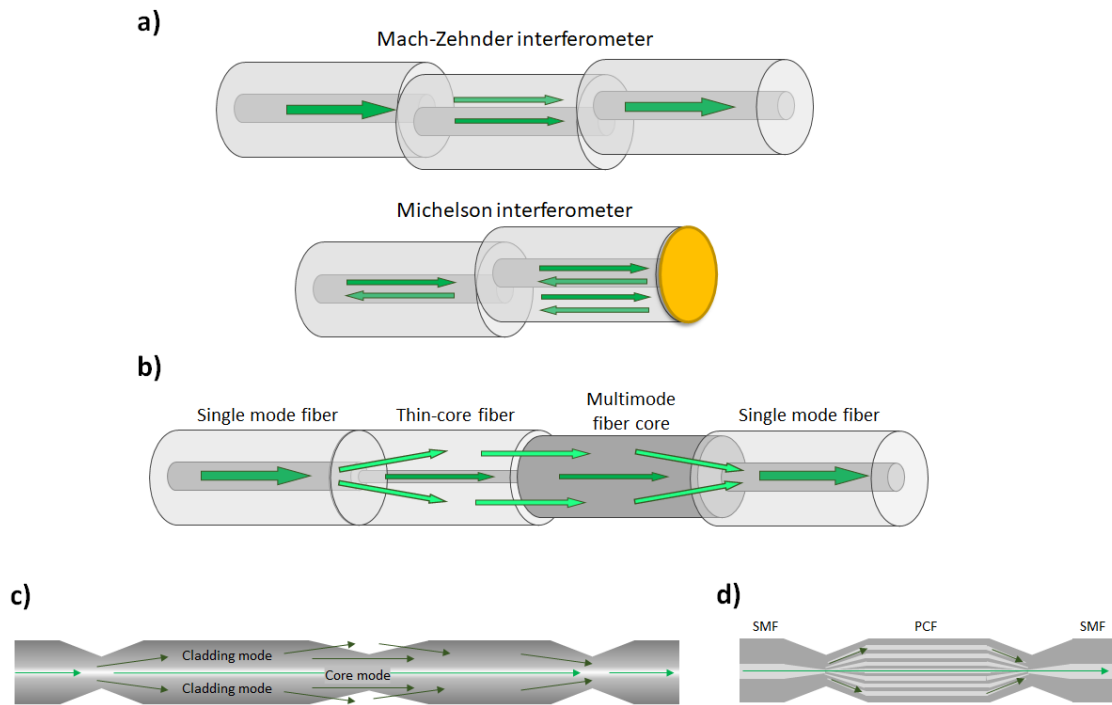
Figure 2.25 – a) Schematic setup of the Mach-Zehnder and Michelson interferometry using core offset technique for RI sensing (adapted from [Tian, 2008]*). b) Michelson configuration based on the fusion of fibers with different core diameters (adapted from* [Shao, 2014]*). c) Triple taper Mach-Zehnder baser RI sensor to create a double optical path (adapted from* [Zhu, 2011]*). d) Photonic crystal fiber between a double taper for a Mach-Zehnder baser RI sensor (adapted from* [Wang, 2016]*).*

2.1.3.5.   Surface Plasmon non-fiber based sensing

Another technique that generally requires spectrum analyzers and that in the last years become a hot topic in the scientific community, is the development of sensor mechanisms based in plasmon resonance phenomena. Surface plasmon resonance (SPR) occurs when collective electron oscillations propagate in metal-dielectric interface by interacting with light. This propagation will depend on the incident light frequency and on the natural resonant frequency of the electrons. These waves are also transverse magnetically (TM) polarized and get excited by only TM polarized light. By analyzing the Maxwell's equations for electromagnetism in this application, the coupling of light to form SPR cannot happened by direct incidence of light in the metal-dielectric interface. For this to be true, a condition must be met which is the equality between the propagation constant of the light in the dielectric medium and that of the surface plasmons at the interface. With direct lightning, the propagation constant of the light in the dielectric medium will always be smaller, therefore techniques that increase this value must be employed [Srivastava, 2013]. Although this phenomenon was expected in the beginning of the 20th century, it was only 1968 that Otto reported the first mechanism to excite SPR [Otto, 1968]. Otto used a prism to create an electromagnetic

evanescent wave based on the idea that the properties of both type of waves are similar, hence a strong possibility that SPR could be excited by coupling with the electromagnetic evanescence wave. An evanescence wave is created when TM light is totally reflected in the interface of two media with distinct refractive index. A metallic surface in the vicinity of this interface (≈200 nm) will interact with the created evanescent wave and excite SPR by energy transference (Figure 2.26 a). The reflected light will have a loss in its intensity due to the energy transference, which become named as the attenuated total reflection (ATR) technique. Since total reflection is angle dependent, SPR excitation is also angle dependent. This configuration presents however the big disadvantage of having a very short distance between the metal surface and prism, and thus difficult to apply in practice. This practical problem was solved by Kretschmann e Raether [Kretschmann, 1968] also in 1968 by using the same principle as Otto but instead of having a metal surface at some distance, the metal is coated at the prism surface and, by guaranteeing that this layer is thin enough (about 50 nm), the evanescent light can propagate thought the metal layer until it reaches the metal-dielectric interface (Figure 2.26 b).
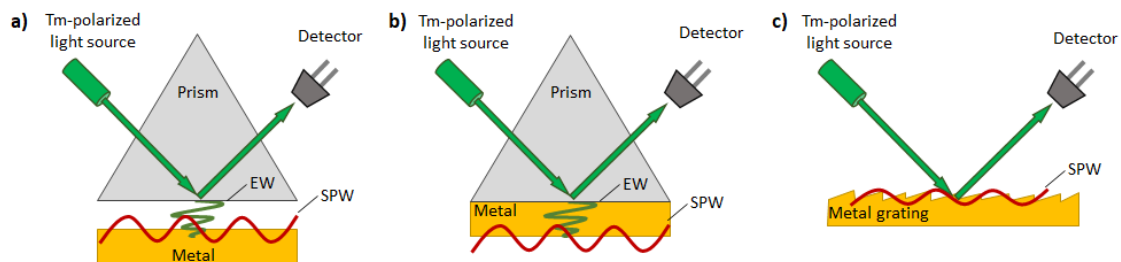
Figure 2.26 – a) Otto's prism configuration to create a evanescent wave (EW) that will promote a surface plasmon wave (SPW) in the vicinity of a metal surface b) Kretschmann's prism configuration where the metal layer is in contact with the prism and the evanescent wave crosses it, creating a SPW in the other side c) Creation of SPW with a metal grating where an evanescent will be created with direct incidence of TM light.

Several metals can be used as a coating for SPR coupling but the most common used are gold and silver due to its better coupling properties [Homola, 1999]. Other optical elements as metal gratings (Figure 2.26 c) and coated optical fibers can also be used to excite SPR. Because SPR excitation depends on wavelength of the light and on its incident angle, three ways of refractive index sensing are reported: intensity variation of the optical wave near the resonance wavelength, measurement of the resonant momentum of the optical wave in angular interrogation and the same momentum in wavelength interrogation. Homola et al. have published a deep review where comparison of sensitivities and resolutions of these three sensing mechanism are reported when using both prim and grating architectures [Homola, 1999]. In the prism, a 50-nm gold layer was

used, while in the grating a pitch and depth of 800 and 70 nm were used. These values can be seen in Table 2.2.

Table 2.2 – Summary resolutions obtained for both prism and grating coupling mechanisms using the angular and wavelength interrogation and intensity measurement methods [Homola, 1999].

Optical system	Angular interrogation (RIU)	Wavelength interrogation (RIU)	Intensity measurement (RIU)
Prism coupler-based	5×10^{-7}	10^{-6}	10^{-5}
Grating coupler-based	2×10^{-7}	3×10^{-5}	5×10^{-5}

*10^{-4} deg. resolution; 5 pm wavelength resolution; 0.2% optical power resolution

Prism coupler-based SPR presents overall higher sensitivity and resolutions compared to the grating coupler-based sensors, independently of the interrogation and measuring method. Angular interrogation presents the highest resolution with 5×10^{-7} RIU but it is a complex method of interrogation compared with the other described methods above, where mechanical precise engineering is mandatory and thus can be costly. The same disadvantage observed in other wavelength interrogation methods, this is, an expensive spectrum analyzer is also prevalent with SPR technique, despite its resolution of 10^{-6} RIU. Intensity measurement can be seen at first glance a costly effective option but it is important to notice that intensity variation needs to be measured near the resonance wavelength requiring very narrow band optical filters which can also be expensive. Nevertheless, a resolution of 10^{-5} RIU is obtained.

Thin metal films are not the only way to promote plasmonic resonance. This can be achieved also by metal nanoparticles with lower sizes than the wavelength of the incident light. When incident light interacts with these metal nanoparticles, the electromagnetic field of light induces a collective coherent oscillation of the surface electrons with the frequency of light. Free electrons in the nanoparticle are attracted due to the electromagnetic field, leading to a charge separation between the free electrons and the ionic metal core. In turn, the Coulomb repulsion forces among the free electrons pushes them to the opposite direction acting as a restoring force, which results in the collective oscillation of electrons (Figure 2.27). Because of this highly localized oscillations, the physical phenomenon is named as localized surface plasmon resonance (LSPR) [Willets, 2007; Cao, 2014].
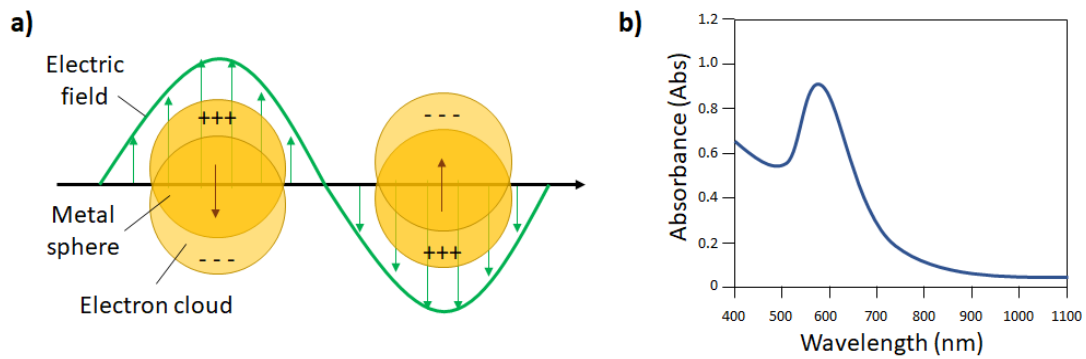
Figure 2.27 – a) Metal particle interaction with light creating localized surface plasmon resonance (LSPR). b) Typical absorbance spectrum of a solution containing metal nanoparticles with the peak obtained due to the LSPR energy transference.

The resonance wavelength has dependence in the size, shape, the type material of metal nanoparticles and, like in SPR, with the external refractive medium. It is however possible to be excited directly by incident light, contrary to thin film SPR. This gives it highly flexibility in sensor concepts since it can be used in simple glass surfaces or even liquid suspension.

2.1.3.6.   Fiber based SPR sensing

Both SPR and LSPR become very common with its integration in optic fiber in the last years because of the advantages of its miniaturization, simpler designs, fast response and remote sensing. Optical fiber acts just like a prism in the Kreshman's configuration and a very specific deep is detected in the transmitted spectrum related to the attenuated total reflection. Most of the fiber components discussed before can be used as a base for SPR and LSPR sensors. The first application totally in fiber was reported in 1993 by Jorgenson and Yee where a thin silver film was applied to an etched fiber (Figure 2.28 a) [Jorgenson, 1993]. A resolution of 1.5×10^{-6} RIU was reported for a spectral resolution of 0.01 nm in a range of 1.33 to 1.39 RIU. Other structures of fiber with its cladding removed by etching were also reported using other metals and nanoparticles. Some examples of these structures will be here presented also with the values of RIU resolutions all based in a spectral resolution of 0.01 nm. One of the examples can be found in the usage of gold nanoparticles (Figure 2.28 b) obtaining a resolution of 2×10^{-5} in a range of 1.33 to 1.41 RIU [Shao, 2010]; etched fiber tips with a gold thin film coat for back reflection spectral analysis (Figure 2.28 c) reporting a resolution of 7×10^{-6} RIU for a range between 1.33 to 1.3549 RIU [Hlubina, 2014]; etched fiber tip with silver nanoparticles and a silver coated mirror for back reflection (Figure 2.28 d) reporting a resolution of 2.5×10^{-5} RIU for a 1.33 to 1.40 RIU range [Chen, 2015]; a U-shaped etched plastic optical fiber with gold nanoparticles (Figure 2.28 e) reporting a resolution of 8.5×10^{-5} RIU

for a range of 1.33 to 1.37 [Gowri, 2016]. Versions of these configurations can also be seen using POF as waveguides which include: thin gold layer with sensitivities of 2185 nm/RIU in the 1.3353 to 1.3653 RIU range [Al-Qazwini, 2016]; deposition of 19 nm diameter gold nanoparticles by chemical modification of fibers and buffer immobilization of the nanoparticles with sensitivity of 42.9 nm/RIU from 1.33 to 1.44 RIU [Jin, 2016]; U-bent using a thin film gold layer with sensitivity of 1040 nm/RIU in a range 1.33 to 1.361 RIU [Christopher, 2018]; U-bent configuration, using graphene and silver nanoparticle hybrid structures in top of a silver layer for measurements of RI from 1.330 to 1.3657 RIU with sensitivity of 700.3 nm/RIU [Jiang, 2017]. A theoretical study of a U-shaped fiber SPR sensor without the need of cladding removing was also presented by the author of this dissertation [Duarte, 2014] which predicted a refractive index resolution of 10^{-8} RIU in the C-band spectral band but with a short range between 1.4122 and 1.4132 RIU.
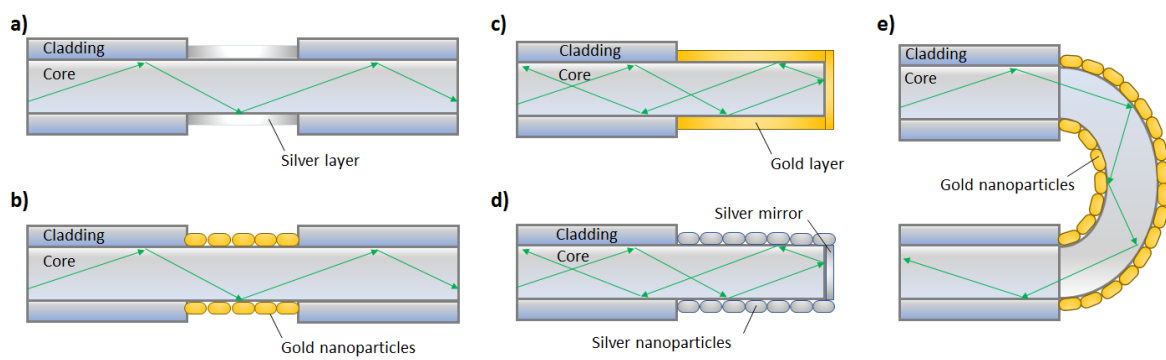
Figure 2.28 – Schematic designs of SPR based RI sensor using: a) silver layer in an etched fiber; b) gold nanoparticles in an etched fiber; c) gold film layer in the fiber tip; d) silver nanoparticles and mirror in the fiber tip; e) gold nanoparticles in an U-shaped etched fiber.

Lateral polishing of fiber can also be used for posterior metal coating, obtaining similar results as fiber etching. Examples of this technique can be observed in silica D-shaped structures where gold metal film is deposited (Figure 2.29 a) [Lin, 2013], obtaining a resolution of 6.3 X 10^{-6} RIU in a range between 1.33 to 1.39 RIU; curved side polished gold coated sensor (Figure 2.29 b) with different bending radii revealing its best resolution with 6 x 10^{-6} RIU for a curvature with 10 mm radius in a range from 1.33 to 1.42 RIU [Gasior, 2014]; D-shape gold nanoparticle deposited sensor (Figure 2.29 c) with a resolution of 4.1×10^{-4} RIU by transmission power interrogation in a range between 1.33 to 1.40 RIU [Chen, 2010]. These structures were also used with the potential lower cost polymeric fiber having as examples: a gold layer thin film deposited over a photoresist buffer layer with sensitivity of 1.325×10^{3} nm/RIU (0.001 RIU resolution) under refractive index characterization tests from 1.333 to 1.371 RIU [Cennamo, 2013]; metal bilayer composed by palladium and gold with a sensitivity of 3 x 10^{3} nm/RIU (3.3 x 10^{-6} RIU resolution) in a 1.38 to

1.42 RIU range [Cennamo, 2016]; deposition and immobilization of five-branched gold nanostars with measurements in a range between 1.333 to 1.371 RIU revealing a sensitivity of 84 nm/RIU (1.1×10^{-4} RIU resolution) [Cennamo, 2013].
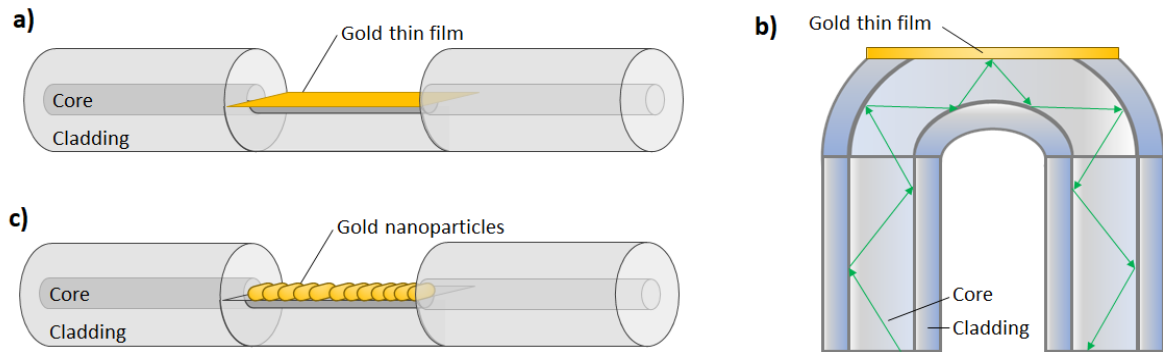
Figure 2.29 – a) D-shaped fiber with gold thin film deposition for SPR RI sensing. b) Side polished U-shaped fiber with a gold thin film deposited for SPR RI sensing. c) D-shaped fiber with gold nanoparticles deposited for SPR RI sensing.

To increase the interaction of the electromagnetic evanescent wave without the need of cladding removing, the fabrication of fiber tapers can be a good alternative. Metal thin films or nanoparticles can also be deposited in this structure. An example of a gold thin film deposition on a taper (Figure 2.30 a) is reported by Monzón-Hernández et al. where a resolution of 4×10^{-7} RIU was obtained but only in a very limited range of 1.435 to 1.445 RIU [Monzón-Hernández, 2004]. A taper-like structure made by etching a fiber was also reported with a double metal layer deposition, more specifically a gold thin layer followed with a TiO_2 thin layer (Figure 2.30 b) for resonances close to the C-band spectral band (1530 - 1565 nm) [Coelho, 2015]. A resolution of 2×10^{-6} RIU was obtained in a range from 1.33 to 1.36 RIU. As expected the use of metal nanoparticles were also reported for taper structures (Figure 2.30 c) [Lin, 2012]. In this work, gold nanoparticles were deposited in the single mode fiber taper structure and the transmission intensity integration measurements were made where a resolution of 3.2×10^{-5} RIU were obtained in the 1.33 to 1.40 RIU interval. Reflective intensity interrogation can also be performed if the taper finishes in a tip and is full coated with a metal layer (Figure 2.30 d). This is exactly what was reported by Chang et al. [Chang, 2006] by using a 40 nm thick gold layer obtaining a resolution of 2×10^{-5} RIU in the range from 1.33 to 1.375 RIU.
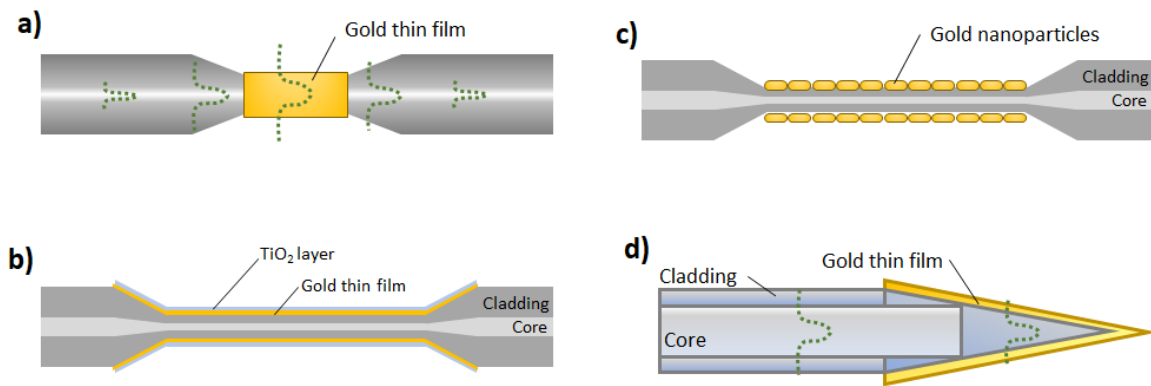
Figure 2.30 – a) Example schematic of a gold thin layer deposited over a taper structure. The evanescent wave of the light interacts between the layer and the external medium. b) Double layer deposition over a taper structure with gold and TiO₂ thin layers. c) Taper structure with gold nanoparticles deposited. d) Fiber tip with a gold thin layer deposited.

Tilted fiber Bragg grating are also good fiber elements to couple energy to SPR when a thin metal layer is coating this element (Figure 2.31). As previously seen, TFBG will reflect light in cladding modes of propagation and these modes, when in the presence of a metal layer, will excite SPR by attenuated total reflection, creating a specific spectral signature with the extinction of some modes which will be sensible to the external refractive index medium. Caucheteur et al. reported this study obtaining a resolution of 1×10^{-5} RIU between a refractive index range of 1.31 to 1.38 RIU [Caucheteur, 2011]. The only similar application in POF was reported by Hu et al. [Hu, 2015] with refractive index measurements between 1.408 to 1.428 RIU and revealing a sensitivity of around 550 nm/RIU (1.8×10^{-5} RIU resolution).
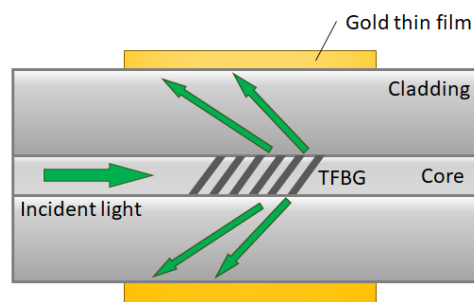
Figure 2.31 – Usage of a tilted fiber Bragg grating to promote the formation of plasmonic surface resonant wave in the metal-external medium boundary of a deposited thin film metal layer.

Plasmonic sensing reveals itself to be a very precise method for refractive index measurement, but most applications reported need spectral interrogation which increases substantially the costs for a commercial application. Even the intensity interrogation method needs expensive specific filter. In addition to this, the metal thin layer deposition needs a good deposition system with a high

precise control to guarantee layer uniformity for sensor repeatability creation, which is difficult and increase higher the costs.

2.1.3.7. RI summary sensing

A summary table with all the refractive index sensors here described can be observed in Table 2.3. If not mentioned as contrary, the methods presented based on optical fiber, uses standard single mode glass fibers.

Table 2.3 – Summary table with most of the refractive index sensors or devices found in literature.

Commercial prism based sensors			
	Range (RIU)	**Resolution (RIU)**	**Ref.**
Abbe Refractometer	1.3 – 1.7	10^{-4}	[Tarrant, 2010]
Lab. Digital Refractometer	1.3 – 1.6	$10^{-5} – 10^{-4}$	[Anton Paar, 2016b]
Digital Handheld Refractometer	1.32 – 1.5	5×10^{-3}	[Mettler-Toledo GmbH, 2014]
Analog handheld Refractometer	1.3 – 1.5	5×10^{-3}	[A.KRÜSS Optronic GmbH, 2016]
In-line Refractometer	1.3 – 1.53	10^{-5}	[Anton Paar, 2016a]
Fiber extrinsic based sensing			
	Range (RIU)	**Resolution (RIU)**	**Ref.**
Air-gap design	1.3326 – 1.3803	---	[Noiseux, 2004]
Cascaded in-line holes (POF)	1.33 – 1.42	(62.9 dB/RIU)	[Shin, 2015]
Displacement based (POF)	1.3322 – 1.3627	---	[Binu, 2009]
Intrinsic Intensity based sensors			
	Range (RIU)	**Resolution (RIU)**	**Ref.**
Fresnel based ratio metric	---	2.5×10^{-5}	[Kim, 2004]
Macrobending (POF)	1.3 – 1.59	5×10^{-3}	[Zubia, 2000]
Side Polishing (POF)	1.33 – 1.44	(154 dB/RIU)	[Jing, 2015]
D-shape (POF)	1.33 – 1.44	6.48×10^{-3}	[Filipa Sequeira, 2016]
Multi-notched LPG (POF)	1.33 - 1.41	8.44×10^{-4}	[Teng, 2017]
Taper	1.37 – 1.40	5×10^{-4}	[Polynkin, 2005]
Taper (POF)	1.334 – 1.3576	2×10^{-3}	[Rahman, 2011]
U-shaped Taper (POF)	1.33 – 1.39	10^{-3}	[Beres, 2011]
Double Taper (POF)	1.33 – 1.42	(950 μW/RIU)	[De-Jun, 2014]
U-shaped clad removed Taper (POF)	1.33 – 1.45	5×10^{-4}	[Teng, 2016]
Ultrasonic crushed fiber (POF)	1.33 – 1.36	(62 dB/RIU)	[Shimada, 2017]
Gratings and grating based interferometric sensors			
	Range (RIU)	**Resolution (RIU)**	**Ref.**
Etched FBG Fabry-Pérot	1.33 – 1.4	1.4×10^{-5} (a)	[Liang, 2005]

	1.33 – 1.42	10^{-5} (a)	[Shao, 2007]
FBG Fabry-Pérot Taper (Multimode Fiber)	1.33 – 1.42	10^{-5} (a)	[Shao, 2007]
LPG Fabry-Pérot Taper	1.333 – 1.362	5.8×10^{-6} (a)	[Jin-Fei Ding, 2005]
TFBG (intensity based)	1.33 – 1.4532	10^{-4}	[Miao, 2009]
TFBG	1.415 – 1.455	3×10^{-5} (a)	[Alberto, 2010]
TFFB to D-shaped coupling (int. based)	1.38 – 1.40	(13 µW/RIU)	[Cai, 2015]
Etched FBG (POF)	1.33 – 1.372	10^{-5} (a)	[Zhang, 2012]
Semi-FBG Fabry-Pérot (POF) (int. based)	1.33 - 1.47	1×10^{-3}	[Ferreira, 2017]
TFBG (POF)	1.42 – 1.49	7×10^{-5} (a)	[Hu, 2014b]

Non-grating interferometric based sensors

	Range (RIU)	Resolution (RIU)	Ref.
Fabry-Pérot catastrophic fused effect	1.33 – 1.43	3×10^{-4} (c)	[Domingues, 2016]
Fabry-Pérot PCF based (intensity based)	1.33 – 1.45	1.2×10^{-5}	[Wang, 2012]
Microcavity PCF tip interferometer	1.33 – 1.38	2×10^{-3} (a)	[Mileńko, 2012]
Sagnac loop PCF	≈ 1.333	4×10^{-6} (a)	[Wu, 2014]
Taper Michelson interferometer	1.33 – 1.36	5.1×10^{-4} (a)	[Tian, 2008]
Taper Michelson interf. (lower waist)	1.33 – 1.37	5.35×10^{-7} (a)	[Ji, 2012]
Core offset Michelson Interferometer	1.315 – 1.3618	3×10^{-5} (a)	[Tian, 2008]
Core offset Mach Zehnder	1 – 1.0022	6.2×10^{-6} (a)	[Duan, 2011]
Mach Zehnder Thin Core – SMF - MMF	1.333 – 1.403	1.34×10^{-4} (a)	[Shao, 2014]
Triple Taper Mach Zehnder	1.333 – 1.389	4×10^{-5} (a)	[Zhu, 2011]
Double Taper Mach Zehnder (POF)	1.33 - 1.74	2.9×10^{-4} (a)	[Jasim, 2014]
Mach Zehnder SMF Taper – PCF – Taper	1.333 – 1.3737	3.8×10^{-6} (a)	[Wang, 2016]
Etched MMI (POF)	1.333 - 1.354	5×10-6 RIU	[R. Oliveira, 2017]

Surface Plasmon non-fiber based sensing

	Range (RIU)	Resolution (RIU)	Ref.
Angular Interrogation Prism	---	5×10^{-7} (d)	
Wavelength Interrogation Prism	---	10^{-6} (a)	
Intensity measurement Prism	---	10^{-5}	[Homola, 1999]
Angular Interrogation Grating	---	2×10^{-6} (d)	
Wavelength Interrogation Grating	---	3×10^{-5} (a)	
Intensity measurement Grating	---	5×10^{-5}	

Fiber based SPR sensing

- **Etched fiber**	Range (RIU)	Resolution (RIU)	Ref.
Silver Film	1.33 – 1.39	1.5×10^{-6} (b)	[Jorgenson, 1993]
Gold Nanoparticles	1.33 – 1.41	2×10^{-5} (b)	[Shao, 2010]
Gold Film fiber tip	1.33 – 1.3549	7×10^{-6} (b)	[Hlubina, 2014]
Silver Nanoparticle fiber tip	1.33 – 1.46	2.5×10^{-5} (b)	[Chen, 2015]
U-shaped Gold Nanoparticles	1.33 – 1.37	8.5×10^{-5} (b)	[Gowri, 2016]
Gold Film (POF)	1.335 – 1.365	4.5×10^{-6} (b)	[Al-Qazwini, 2016]
Gold Nanoparticles (POF)	1.33 – 1.44	2.3×10^{-4} (b)	[Jin, 2016]

	U-bent Gold Film (POF)	1.33 – 1.36	9.6×10^{-6} (b)	[Christopher, 2018]
	U-bent Silver Film and Nanoparticles (POF)	1.330 – 1.366	1.4×10^{-5} (b)	[Jiang, 2017]
-	**Lateral polished**			
	D-shaped Gold Film	1.33 – 1.39	6.3×10^{-6} (b)	[Lin, 2013]
	Curved Gold Film	1.33 – 1.42	6×10^{-6} (b)	[Gasior, 2014]
	D-shaped Gold Nanoparticles (int. based)	1.33 – 1.40	4.1×10^{-4}	[Chen, 2010]
	D-shaped Gold Film (POF)	1.333 – 1.371	1×10^{-3} (b)	[Cennamo, 2013]
	D-shaped Palladium-Gold Bilayer (POF)	1.38 – 1.42	3.3×10^{-6} (b)	[Cennamo, 2016]
	D-shaped Gold Nanostars (POF)	1.333 – 1.371	1.1×10^{-4} (b)	[Cennamo, 2013]
-	**Taper**			
	Gold Film	1.435 – 1.445	4×10^{-7} (b)	[Monzón-Hernández, 2004]
	Gold-TiO$_2$ double layer	1.33 – 1.36	2×10^{-6} (b)	[Coelho, 2015]
	Gold Nanoparticles	1.33 – 1.40	3.2×10^{-5} (b)	[Lin, 2012]
	Fiber tip Gold Film	1.33 – 1.375	2×10^{-5} (b)	[Chang, 2006]
-	**Tilted fiber Bragg grating**			
	Gold Film	1.33 – 1.38	1×10^{-5} (b)	[Caucheteur, 2011]
	Gold Film (POF)	1.408 – 1.428	1.8×10^{-5} (b)	[Hu, 2015]

[a]1 pm spectral analyzer resolution; [b]10 pm spectral analyzer resolution; [c]Modulation period analysis; [d]10^{-4} deg. resolution;

3. Optical sensor design and characterization

3.1. Technologies used

For the development of the multiparameter optical based sensing platforms, the several technologies previously described were considered, however there are specific criteria that need to be fulfilled. The technologies to be used should be balanced and reasoned in the following conditions:

- Can be implemented as in-line sensing;
- Present low cost in its development and implementation to be commercial competitive;
- Present low complexity and high reproducibility in its development and implementation;
- Have the highest possible resolution by cost rate.

With these criteria in consideration, a plastic fiber based sensing mechanism for turbidity based in the technology reported by Garcia et al. [Garcia, 2008] and Crespo et al. [Crespo, 2010] was chosen by its in-line capabilities, low cost and ease of implementation, by requiring simple elements like IR LED and two photodiodes that can be implemented by basic electronic. To measure color, this technology can be expanded by adding several color LEDs or other form of light sources and couple them to the same transmitted light fiber maintaining the same in-line properties without a high increase of complexity and costs. Overall, this sensor can be categorized as a simple nephelometric turbidity sensor or in a multibeam nephelometric turbidity sensor, depending on the information analyzed by using all the color light sources or just the IR source. Having the information of turbidity intrinsically obtained with the same technology as measuring color, gives it the advantage to be able to easily estimate color by simple removing the effect of turbidity using specific algorithms.

Although the use of fiber optics as sensing mechanism for refractive index sensing seems straightforward since it is also used as turbidity/color sensing, some aspects of this technology should be considered, mainly because it presents high disadvantages to meet the specified criteria above stated. Most of the discussed technology using optical fiber that presents high resolutions is based in spectral interrogation. As discussed, this method of interrogation requires complex optical and electronic components and present high costs. The mechanisms of intensity measurement also

present issues concerning the costly methods of fabrication and the necessity of very narrow optical filters. Because of this, gratings, interferometric and SPR based technology cannot be considered. In other hand, intensity fiber based sensor using tapers, d-shaped and polished surfaces present lower resolutions in comparison with prism based sensing. The latter is equally cost effective and is not that complex to implement even in an in-line sensor. For this reason, the prism based sensing technology will be chosen to be integrated in the multiparameter optical platform.

3.2. Turbidity and Color Sensor

3.2.1. Physical working principle

The turbidity and color sensing mechanisms shares the same basic architecture that uses two major physical phenomena: light transmission and light scattering between two waveguides, this is, two plastic optical fibers. To understand the behavior of light exiting an optical fiber, one needs to understand how light propagates in the fiber in the first instance. Here, light in optical fiber propagation will be analyzed and both transmission and scattering phenomena will be looked in more detail.

3.2.1.1. Optical fiber waveguide

Light propagating in a different medium than vacuum will, in general, have its velocity lowered. By relating this propagation velocity (v_n) in a ratio with the light velocity in vacuum (c) the notion of refractive index (n) is introduced:

$$n = \frac{c}{v_n}$$

(Eq. 3-1)

Different light velocities will have a major impact when light propagates between different media. As light passes the border between media, depending upon the relative refractive indices of the two media, the light will either be refracted to a lesser angle, or a greater one (Figure 3.1).
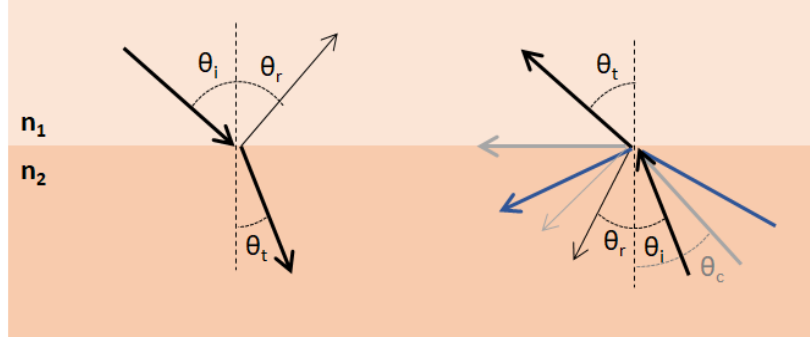
Figure 3.1 – Refraction and reflection of light between two media (with n1 < n2).

This is ruled by the Snell law of refraction:

$$\frac{\sin(\theta_i)}{\sin(\theta_t)} = \frac{n_2}{n_1} \qquad \text{(Eq. 3-2)}$$

where θ_i is the angle of incident light to the boundary and θ_t the angle of refracted light, measured with respect to the normal line, represented perpendicular to the boundary. n_1 and n_2 are the refractive indices of each medium. When light comes from a medium with higher to a lower refractive index, it is refracted with higher angles. If the incident angle is increased, a critical angle (θ_c) is achieved when the refracted angle equals the 90°, and total reflection of light starts to occur for higher angles of incidence. For $\theta_t = 90°$ we have,

$$\sin(\theta_c) = \frac{n_2}{n_1} \qquad \text{(Eq. 3-3)}$$

Total internal reflection is the principle behind the usage of the optical fiber. An optical fiber is constituted by two materials with different refractive indices. The core with a refractive index denoted by n_{co}, is enfolded by a cladding with a refractive index denoted by n_{cl}, with $n_{cl} < n_{co}$. When light is coupled by a light source to the fiber tip, only certain range of angle rays can propagate through the fiber (Figure 3.2).
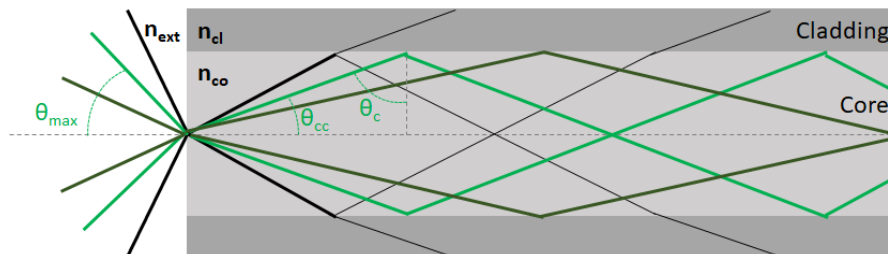
Figure 3.2 – Light propagation with total internal reflections in an optical fiber structure.

The maximum incident light angle that a fiber will be able to fully propagate, which is symbolized by θ_{max}, can be calculated by the following relation,

$$\frac{\sin(\theta_{max})}{\sin(\theta_{cc})} = \frac{n_{co}}{n_{ext}} \qquad \text{(Eq. 3-4)}$$

where $\sin(\theta_{cc}) = \cos(\theta_c)$, being θ_{cc} the complementary angle of θ_c. The quantity of light received by this optical system is characterized by the numeric aperture (NA),

$$NA = \sin(\theta_{max}) = \frac{n_{co}\cos(\theta_c)}{n_{ext}} \qquad \text{(Eq. 3-5)}$$

$$= \frac{n_{co}\sqrt{1 - \sin^2(\theta_c)}}{n_{ext}}$$

$$= \frac{n_{co}\sqrt{1 - (n_{cl}/n_{co})^2}}{n_{ext}}$$

$$= \frac{\sqrt{n_{co}^2 - n_{cl}^2}}{n_{ext}}$$

NA values are in general considered when $n_{ext} = 1$. In this case,

$$NA = \sqrt{n_{co}^2 - n_{cl}^2} \qquad \text{(Eq. 3-6)}$$

3.2.1.2.   Air gap fiber to fiber transmission

A simple geometrical model scheme of air gap transmission of light between two perfectly aligned multimode fibers is presented in Figure 3.3.
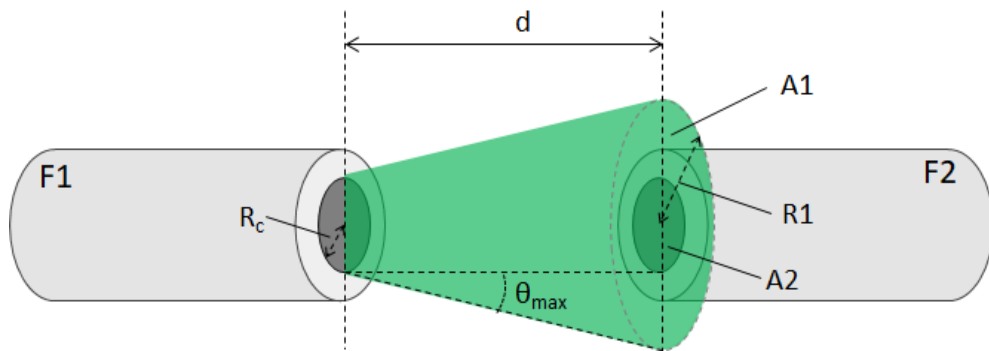
Figure 3.3 – Air-gap transmission of light between two perfectly aligned fibers.

The light emitted by one of the fibers, identified as *F1*, will propagate in a conical shape until it reaches the second fiber, *F2*, at a distance *d*. Because of its increasing conical shape, only a fraction of light reaches *F2*. The ratio of received light power *P2*, with the one emitted, *P1,* will be equal to

the ratio of the receiving area *A2*, with the light emitting area *A1* when it reaches the fiber *F2*, considering that the light energy is radially even through the fiber [Kuang, 2004],

$$\frac{A2}{A1} = \frac{P2}{P1} \qquad \text{(Eq. 3-7)}$$

A1 will be directly dependent of the numeric aperture of *F1* fiber mainly because of its θ_{max} angle. With *R1* as the radius of area section *A1*, and R_c as the radius of the fiber core, one can make some deductions,

$$R1 = R_c + d * \tan(\theta_{max}) \qquad \text{(Eq. 3-8)}$$

$$A1 = \pi * (R1)^2 = \pi * [R_c + d * \tan(\theta_{max})]^2 \qquad \text{(Eq. 3-9)}$$

$$A2 = \pi * (R_c)^2 \qquad \text{(Eq. 3-10)}$$

$$P_2 = P_1 \frac{A_2}{A_1} = P_1 \frac{\pi*(R_c)^2}{\pi*[R_c+d*\tan(\theta_{max})]^2} = P_1 \frac{R_c}{R_c+d*\tan(\theta_{max})} \qquad \text{(Eq. 3-11)}$$

As can be seen by (Eq. 3-11), the fraction of light power that reaches *F2* will be mostly dependent in geometric factors like the distance between fibers and the fiber core radius. θ_{max} is however an important component that depends, not only, on the characteristics of the fiber, by the refractive index of their core and cladding, which by themselves are characteristics of the material that they were made, but also it depends on the refractive index of the external medium. We can rewrite θ_{max} as,

$$\theta_{max} = \arcsin(NA) = \arcsin\left(\frac{\sqrt{n_{co}^2 - n_{cl}^2}}{n_{ext}}\right) \qquad \text{(Eq. 3-12)}$$

n_{co} will also be dependent on the used light wavelength (λ). Considering that polymethyl-methacrylate is the material of the polymeric fibers used in this work, the Sellmeier's equation [Ishigure, 1996] can be used to calculate n_{co},

$$n_{co}^2 - 1 = \sum_{i=1}^{3} \frac{C_i \lambda^2}{\lambda^2 - l_i^2} \qquad \text{(Eq. 3-13)}$$

with C_i and l_i the respective polymer material coefficients presented in Table 3.1.

Table 3.1 – Sellmeier's equation coefficients for the polymethyl-methacrylate [Ishigure, 1996].

Coefficient[*]	i=1	i=2	i=3
C_i	0.4963	0.6965	0.3223
l_i	0.0718	0.1174	9.2370

[*]for wavelength in μm

A theoretical representation of the percentage of light in relation to the initial emitted light power, with the variation of distance for different wavelengths of light, considering a coupling from an air medium into the *F2* fiber, is presented in Figure 3.4 a). Although the wavelength of the light can have some effect in the transmitted light power, mainly for lower wavelengths values, in relation to the effect obtained by the distance, it can be neglected. Other interesting analysis inside the scope of this work can be made by calculating the power variation (P_v) obtained with the increase of refractive index, mainly in liquids, and normalize its value with the power obtained in water (P_w) (Figure 3.4 b). Higher sensibilities are obtained for higher fiber distances but this effect rapidly fades with high distance value as expected by the less efficient light coupling.
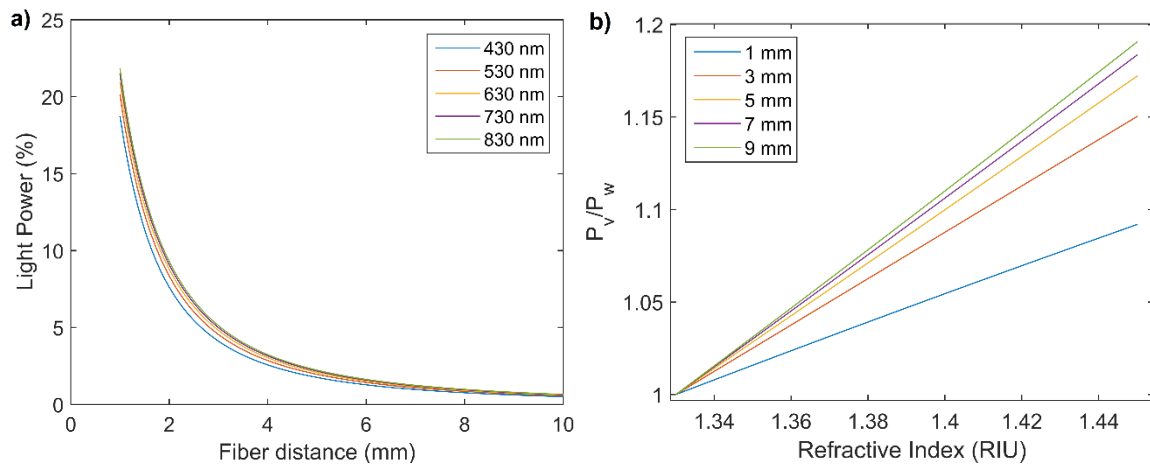
Figure 3.4 – a) Light power percentage coupled with the air-gap fiber in respect with the initial emitted light power and with the variation of distance for different wavelength of light. b) Power variation obtained with the increase of refractive index of the gap medium, normalized to water refractive index obtained power.

When submerged in solutions where there is the existence of colored or suspended particles, phenomena of scattering or absorption of light will occur. This is described by the Beer-Lambert law that states that attenuation (A) will be directly proportional to the light path length (L) and to the concentration (ς) of the attenuating species in the solution,

$$A = [\varepsilon_s(\lambda) + \varepsilon_a(\lambda)]L\varsigma \qquad \text{(Eq. 3-14)}$$

where ε_s and ε_a are the scattering and absorption coefficients that are intrinsic properties of the attenuating species and are light wavelength (λ) dependent. These coefficients can be very different even if the particles are made of the same material. Different values will be obtained by changing sizes and geometric shapes. This is also an additive law if two or more species are present in the solution, where each material coefficients are just added. Attenuation of light have a logarithmic relation with transmission,

$$A = -log_{10}\left(\frac{I}{I_0}\right) \qquad \text{(Eq. 3-15)}$$

with I and I_0 as the transmitted light intensity and the initial light intensity respectively. The variation of light intensity in relation to the properties of the attenuating species in the solution can then expressed by,

$$I = I_0 10^{[\varepsilon_s(\lambda)+\varepsilon_a(\lambda)]L\varsigma} \qquad \text{(Eq. 3-16)}$$

3.2.1.3. Particle scattering of light

Light scattering by particles occurs by the interaction of the light when it collides with a particle. Depending on the size of the particle, two main scattering formalisms are used to describe mathematically the phenomena. For particles much smaller than the wavelength of the radiation, the Rayleigh scattering formalism is used, while for particles with the same magnitude of the wavelength of the radiation, the Mie solution to Maxwell's equations is more adequate. In the scope of the work here presented, Mie solution to Maxwell's equations is the formalism that better represents the scattering of light of particles that are expected to have diameters in a range of 1 nm to 100 nm. This complex formalism can be found in [Bohren, 1983] and will not be here presented, but the numerical solution will be discussed instead.

The general solution for the light scattered coefficient (Q_{sca}) of a particle is determined by the following numerical expression,

$$Q_{sca} = \frac{2}{(kr_s)^2}\sum_{n=1}^{\infty}(2n+1)(|a_n|^2 + |b_n|^2) \qquad \text{(Eq. 3-17)}$$

where r_s is the sphere radius, $k = 2\pi/\lambda$ with a_n and b_n being the Mie coefficients of the amplitudes of the scattered field to compute. This coefficient values are calculated based in spherical Bessel and Hankel functions as,

$$a_n = \frac{m^2 j_n(mx)[x j_n(x)]' - j_n(x)[mx j_n(mx)]'}{m^2 j_n(mx)\left[x h_n^{(1)}(x)\right]' - h_n^{(1)}(x)[mx j_n(mx)]'}$$

(Eq. 3-18)

$$b_n = \frac{j_n(mx)[x j_n(x)]' - j_n(x)[mx j_n(mx)]'}{j_n(mx)\left[x h_n^{(1)}(x)\right]' - h_n^{(1)}(x)[mx j_n(mx)]'}$$

(Eq. 3-19)

with $x = k r_s$, m is the refractive index of the sphere relative to the external medium, j_n the spherical Bessel functions of the first kind. $h_n^{(1)}$ is the spherical Hankel functions which are a linear combinations of spherical Bessel functions of the first kind and of the second kind (y_n). Table 3.2 presents the formulation how to recursive calculate the necessary spherical Bessel functions.

Table 3.2 – Summary table of the recursive calculation of the spherical Bessel functions. Here z is the function argument.

n=0	n=1	n>1
$j_0 = \dfrac{\sin(z)}{z}$	$j_1 = \dfrac{\sin(z)}{z^2} - \dfrac{\cos(z)}{z}$	$j_{n+1}(z) = \dfrac{2n+1}{z} j_n(z) - j_{n-1}(z)$
$y_0 = -\dfrac{\cos(z)}{z}$	$y_1 = -\dfrac{\cos(z)}{z^2} - \dfrac{\sin(z)}{z}$	$y_{n+1}(z) = \dfrac{2n+1}{z} y_n(z) - y_{n-1}(z)$

For the calculation of the derivative functions, one can use the following rules (with z as the argument) based on the spherical Bessel functions themselves,

$$[z j_n(z)]' = z j_{n-1}(z) + n j_n(z); \quad \left[z h_n^{(1)}(z)\right]' = z h_{n-1}^{(1)}(z) + n h_n^{(1)}(z);$$

(Eq. 3-20)

The Mie solution was calculated numerically to compare it with real data obtained from solutions with non-absorbing particles (Figure 3.5). The particles have around 1 μm of diameter but unknown refractive index, which demands that this value must be guessed for a best fit. To ensure this, and because scattering is wavelength dependent, four different LEDs with central peak light wavelengths of 470, 530, 660 and 870 nm were used to make a correlation with real data.
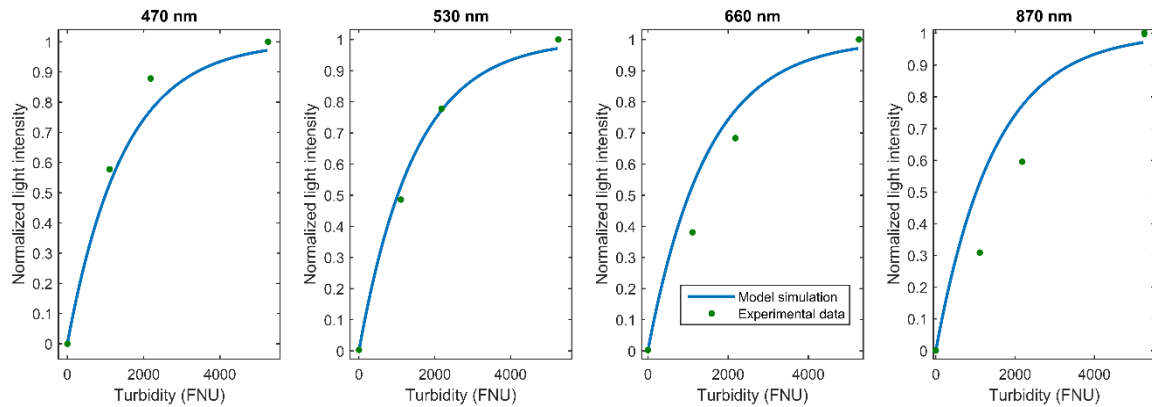
Figure 3.5 – Mie solution simulation comparison with experimental data obtained for four different central peak wavelengths LEDs.

The simulated data presents the expected behavior that was obtained experimentally with the solutions. There isn't however a perfect match between the two, which can be explained by use of a single wavelength of light in the simulation while the LEDs have a broader spectrum in addition to the unknown standard deviation of the particle sizes present. Nevertheless, the scattered light of turbid solutions can be well understood with the Mie solution approach.

3.2.2. Sensor Design

The research on the sensor and the prototypes for turbidity and color measurements were developed over time with sequential improvements regarding its components and mode of operation, mainly in the light sources and photodetectors. Nevertheless, all share the same electronic control board and physical principles of measurement. Here it will be presented all the used components followed by the different versions of sensors assembled and tested.

3.2.2.1. Electronic control board and fibers

The sensors will have the capabilities to detect both light from 0° angle (direct transmission) as well as from the 90° angle (nephelometric measurement). This requires at least two photodetectors, but a third photodetector will be used as a reference measurement to numerically minimize the influence of any electronic fluctuation, as well as thermal drift fluctuations that can produce light intensity peaks. Four light sources will be used that can be both LEDs or laser depending on the sensor version assembled. The light sources must be in the infra-red, red, green and blue region of the spectrum. The electronic board developed in collaboration with Watgrid (see chapter 5.1.3) was designed to be electric stable in the power delivery to the optical elements. A microcontroller is used for a controlled light emission where electronic currents of the light sources

and amplifying value detention can be settled digitally. The acquisition method starts by turning on and off sequentially each light source for a second, registering the values of each photodetector. The board have also connectivity expansibility where other type of sensors can be used. This feature is used to introduce external electronic temperature sensor in addition to the internal temperature sensor available. Besides of having a Universal Serial Bus (USB) connector for local communication with a computer, a Wi-Fi module can be used with the board to communicate the sensing measurements to a router for internet connection, which is ideal for wireless in-line sensing. The general design of the board with their main optical components can be seen in Figure 3.6.
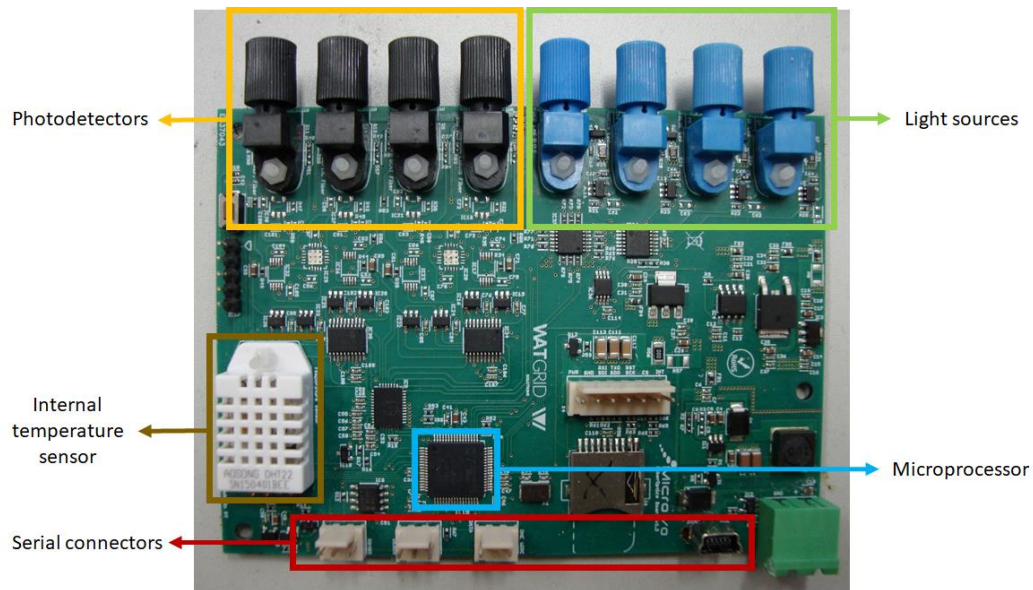
Figure 3.6 – Photograph of the used Watgrid's electronic board to control the optical elements, sensor expansions and external connections.

To couple light from all the light sources to a single emitting fiber, a 4x4 POF coupler from Industrial Fiber Optics (IF 544) was used with a splitting ratio of 25:25:25:25 %. All the four coupler inputs are used for each light source, while one of the outputs is used as the source light to the sensing head and a second is directly connected to the reference photodetector for numerically optical drifts compensation. In the measuring mode, the board will sequentially turn on and off each light source for a second and register the data obtained from the photodetectors. A full measurement will have information from the three photodetectors and from the four light sources, counting a total of 12 values. The plastic optical fiber used was the ESKA[tm] series model GHCP4001 from Mitsubishi Rayon Co., LTD. The characteristics of this fiber are presented in Table 3.3 and its typical wavelength dependence observed in Figure 3.7.

Table 3.3 – Summary characteristics of the ESKA^{tm} series model GHCP4001 fiber used in this work.

Parameter	Min.	Typ.	Max.
Core material		Polymethyl Methacrylate	
Cladding Material		Fluorinated Polymer	
Core refractive Index		1.49	
Refractive index profile		Step-index	
Numerical Aperture (NA)		0.5	
Core diameter (μm)	920	980	1040
Cladding diameter (μm)	940	1000	1060
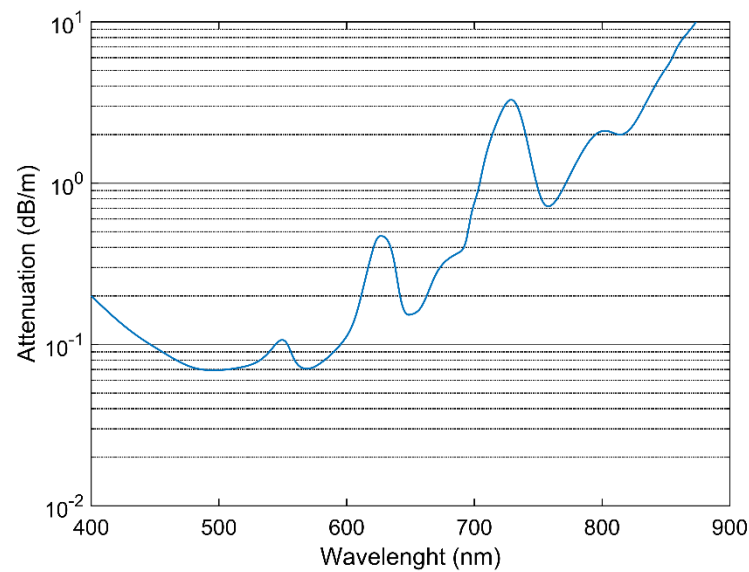
Figure 3.7 – Typical wavelength attenuation of light obtained with the used GHCP4001 model fiber.

3.2.2.2.   Photodetectors

Two photodetectors were used in the testing phase, both obtained from Industrial Fiber Optics. These photodetectors were chosen because of its easy implementation in the electronic control board, but mainly due to its "connector-less" style plastic fiber optic package that facilitates the connection of the fiber to the optic element, just by using a screw. The first photodetector (IF D91) is a photodiode, has lower responsivity and was used in the first prototypes tested. Because higher responsivities seemed to be required after the first testing phase, a second photodetector (IF D93), a photodarlighton, with higher responsivity started to be used. A summary of the important optical characteristics of both photodiodes is presented in Table 3.4. A representation of the typical photodetector response to wavelength can also be seen in Figure 3.8.

Table 3.4 – *Characteristics table of the used photodetectors.*

Parameter	IF D91	IF D93
Wavelength for Maximum Photosensitivity (nm)	880	850
Spectral Bandwidth (nm)	450 - 1050	400 - 1100
Responsivity min. ($\mu A/\mu W$) @ 870 nm	0.5	5600
@ 650 nm	0.4	5300
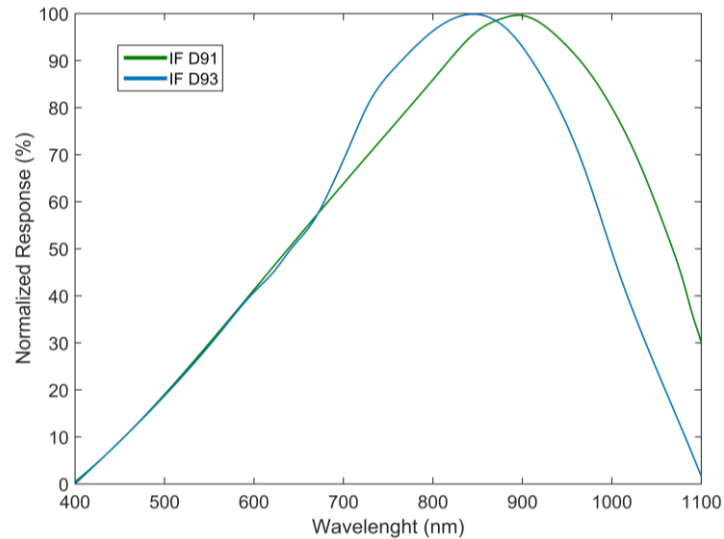
Figure 3.8 – *Normalized response of the used photodiodes with wavelength.*

3.2.2.3.   Light Emitting Diodes (LEDs)

Four main spectral regions were used as light sources: blue, green, red and infra-red. In the same way as stated with the upgrade necessary to be performed to the photodetectors for improvement of the sensitivity of the sensor after the first prototypes tests, higher power LEDs were also used for blue color in posterior prototypes since it revealed to be a necessity. A total of 5 LEDs were used in the different versions of the tested sensors. In the blue region, LEDs with the central peak wavelength of 430 (IF E92A) and 470 nm (IF E92B) were chosen. In the green region, the 522 nm (IF E93), in red the 660 nm (IF E97) and in the infra-red the 870-nm peak wavelength (IF E91D). A more detailed description of each LED optical characteristics can be observed in Table 3.5.

Table 3.5 – *Characteristics table of the used LEDs.*

Parameter	IF E92A	IF E92B	IF E93	IF E97	IF E91D
Peak Wavelength (nm)	430	470	522	660	870
Spectral Bandwidth (nm)	65	25	40	40	50
Typical Output Power (μW)	25	75	465	325	150

The typical spectral output by wavelength for each LED is displayed in Figure 3.9 a). Since the used photodetectors do not present a linear responsivity, the real output that the optical system will detect, in ideal conditions, needs to take this responsivity in consideration as observed in Figure 3.9 b). Notice that light from LEDs with higher wavelength will have better detection, but the fiber waveguide will have its lower attenuation values around 500 nm, while its high attenuation values are located in the infra-red region. In conclusion, the change of the 430 to 470 nm blue LED will improve detection, while the light from the IR LED will present lower detection values when using long optical fibers to guide the light, despite the high responsivity of the photodetector in this region.
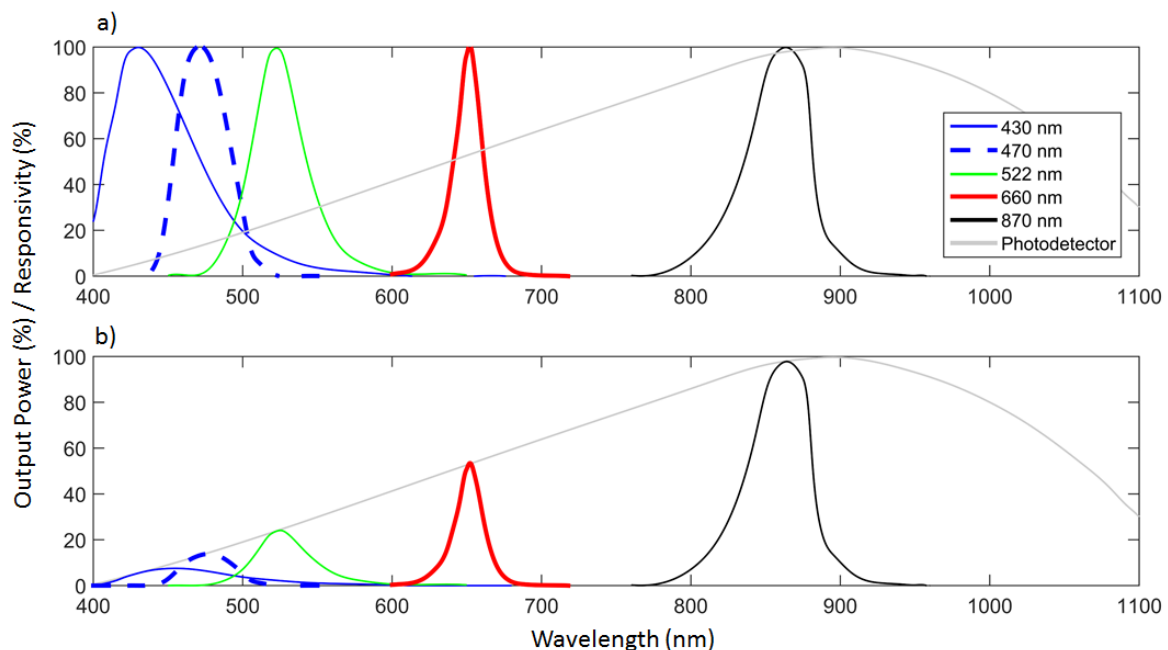
Figure 3.9 – a) Individual responsivity and spectral output power of the photodiodes and LEDs used. b) Real output power detected by the photodiode based on its responsivity.

3.2.2.4.   Light Amplification by Stimulated Emission of Radiation (Laser)

Laser light sources were also tested with the objective to increase color detection range in the first testing phase, mainly in the blue and green spectral region. Two laser diodes from Osram GmbH. were used with the identified model as PL 450B for central peak wavelength of 450 nm and model PL 520 B1 for central peak wavelength of 520 nm. The laser diode differentiates from the LED light source with their higher light output, narrow spectral bandwidth and coherent light. A more detailed information is presented in Table 3.6.

Table 3.6 – Characteristics table of the used laser light sources.

Parameter	PL 450B	PL 520 B1
Peak Wavelength (nm)	450	520
Spectral Width (FWHM) (nm)	2	2
Typical Output Power (mW)	80	50

An important thing to notice by using this type of light sources is its high dependence of output power with temperature variations, which can be bring stabilization issues in detection. This variation can be observed in Figure 3.10 for both lasers. Since the developed sensor using this technology had not a good heat sink mechanism, the sensor was at the mercy of the ambient temperature that, when in direct exposition with sun, can lead to very high variations in a daily cycle.
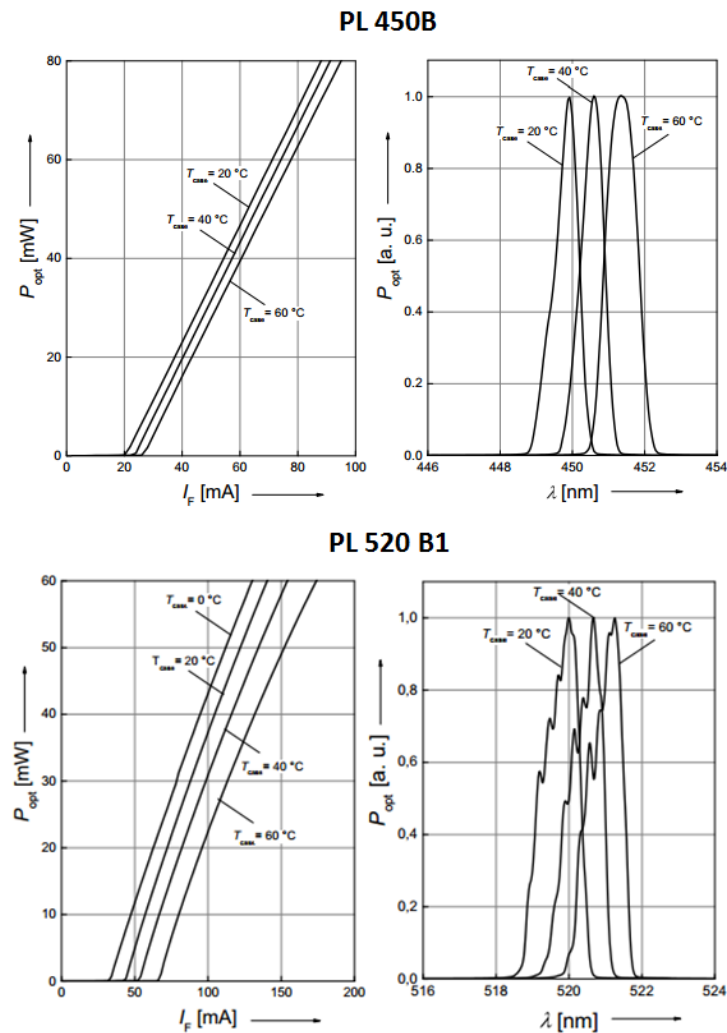
Figure 3.10 – Output power variation with temperature and respective spectra of the used laser light sources [Osram GmbH., 2015a, 2015b]*.*

3.2.2.5.   In-line submerged sensing head

The developed in-line turbidity and color sensors requires a direct interaction with the external medium from which it pretends to measure. This interaction can occur in two ways, by directly submerging the sensing head into the solution to assess or by continuously withdraw the solution from its environment/container in a closed circuit to a measurement cell where the measurement is carried on before it is returned to its original environment. The first approach was applied to 4 sensors that differentiate among them by using different optical components. They have however the same sensing principle with the same structure and sensing measurement head. The sensor head is made of 316 stainless steel, a food compatible material, containing an internal groove structure to allocate three plastic optical fibers coming from the electronic control board: the emitter from a light source; the direct transmission receiving fiber and the 90° receiving fiber that will guide light to the photodetectors (Figure 3.11). The aligned emitter and the direct transmission fiber are separated by a 5 mm gap which is a good compromise between broader range (shorter distances) and sensitivity (higher distances) [Bilro, 2010]. This also ensures that particle agglomerations do not obstruct with ease the sensing cavity which could happen for high turbid solutions as is the case of the environment intended to be tested with this sensor. The 90° scattering light fiber is placed longitudinal 2 mm from the emitter fiber to prevent internal reflections of light in the sensing cavity, but not so much distant that it is affected by the light cone. It is also because of the light cone that this fiber is placed 2.5 mm transversal to the emitter fiber so the effect is minimized in the absence of particles.
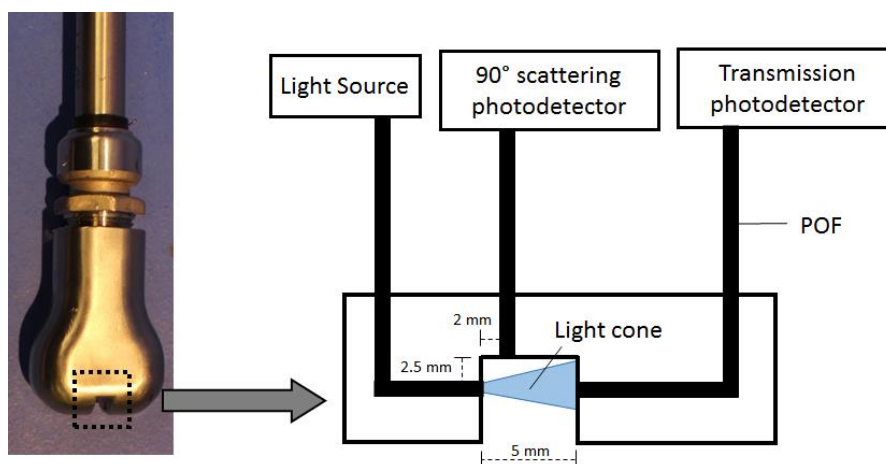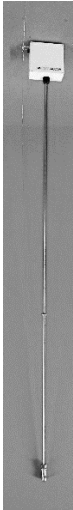
Figure 3.11 – Sensing head photograph and schematic of its operation principle.

The four developed sensors have also a connection between the sensor head and the electronic control board by a steel tubular protection element where the fibers go through. The first version

of the sensor used a 50-cm malleable steel hose that was connected to a polymeric case to ensure basic protection of the electronic board. This sensor version used the low responsivity photodetectors (IF D91) and 430, 520, 660 and 870 nm LEDs and is suited to measure in small liquid containers. For deep containers, a long 2-meter connector version were also tested. In this case the connector was rigid but the optic components used were the same as version 1, except for the required longer POF used. The electronic control board were also better incased in a water resistance IP67 polystyrene box. This sensor test was useful to understand the effects of dimension changes in the sensor. Because of the lack of color sensitivity verified to sensors version 1 and 2, more power light sources were used in version 3, mainly the blue and green laser. Every other optical component was the same as version 2 as well as the structure and the length of the sensing head connector. Later, during testing phase, the new photodetector IF D93 was available which had about 1000x higher responsivity and so a fourth version of the sensor using these photodetectors were used, returning solely to the use of LEDs, but now using the more detectable 470 nm LED instead of the 430 LED. A summary table with all the sensor versions using a sensing head are detailed in Table 3.7.

Table 3.7 – Summary of the developed and characterized sensors using the sensing head.

Sensor Version	Head connector length (cm)	Blue Source	Green Source	Red Source	IR Source	Photo detector
"Short LED sensor" (Version 1)	50	IF E92A (430 nm)	IF E93 (522 nm)	IF E97 (660 nm)	IF E91D (870 nm)	IF D91
"Long LED sensor" (Version 2)	200	IF E92A (430 nm)	IF E93 (522 nm)	IF E97 (660 nm)	IF E91D (870 nm)	IF D91
"Laser Sensor" (Version 3)	200	PL 450B (450 nm)	PL 520 B1 (520 nm)	IF E97 (660 nm)	IF E91D (870 nm)	IF D91
"High Resp. LED sensor" (Version 4)	200	IF E92B (470 nm)	IF E93 (522 nm)	IF E97 (660 nm)	IF E91D (870 nm)	IF D93

3.2.2.6.   Peristaltic pump sensor

A sensing mechanism based on the continuing withdraw of solution from its environment, in a closed-circuit system, to a flowing cell where the measurement occurs, was also developed to be compared to its sensing head version counterparts. This sensor was designed with the principle that an easy regular cleaning by an operator could be made without taking the sensor from the sensing liquid container or site, which have to be done in the sensors with the sensing head versions. To create a closing flow circuit, a peristaltic pump is used to suck the liquid to be measured into the measurement cell located inside the protection box, where is also located the electronic board and the optical components. A long silicon tube connected to the pump is used to be submerged in the liquid. The use of a peristaltic pump was chosen to prevent the creation of bubbles that are observed in other traditional pump systems with turbulent liquid flowing. The pumped liquid passes through the measurement cell made of a food compatible polypropylene plastic block with a 5-mm diameter hole. This hole is crossed in the center by other 1 mm diameter hole that is pierced along all the polypropylene block that is used to place the optical fibers in position of emitter and receiving transmission. Transversal to this hole and centered with the 5-mm cavity, a second hole is pierced to introduce the scattering fiber (Figure 3.12). Overall the sensing mechanism is the same as the one used with the sensing head, but with different arrangement. Because of its more enclosed design, it is expected that some light will be reflected in the walls of the measurement cell. The flowing cycle is then finished by the withdraw of the liquid from the measurement cell back to its original environment.
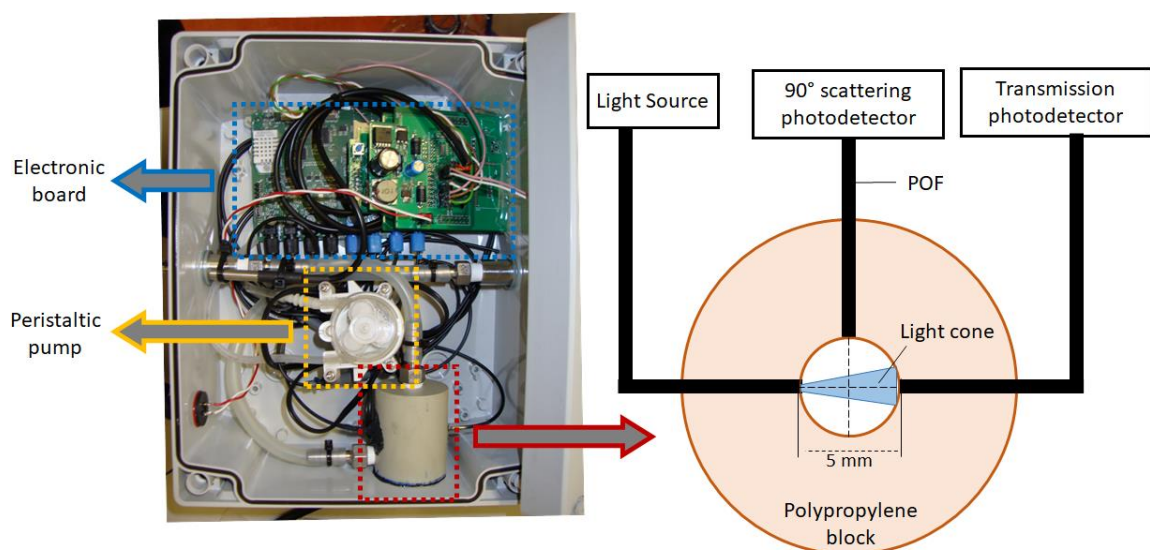
Figure 3.12 – Photograph of the box containing all the peristaltic flowing system and the electronic necessary to analyze the solution. In the right, a detailed scheme of the measurement cell made from a polypropylene block.

The biggest disadvantage of this system is the increase of complexity and higher failure risk associated with it, mainly to possible leakages on the flowing system. The costs-benefit are however increased by the unnecessary use of the costly machined steel sensor head. A summary table of the characteristics of this sensor can be observed in Table 3.8.

Table 3.8 – Summary of the developed and characterized sensor based on a peristaltic pump system.

Sensor Version		Tubing length (cm)	Blue Source	Green Source	Red Source	IR Source	Photo detector
"Peristaltic pump sensor" (Version 5)		200	IF E92B (470 nm)	IF E93 (522 nm)	IF E97 (660 nm)	IF E91D (870 nm)	IF D93

3.2.3. Turbidity measurement

Turbidity tests were performed for each version of the developed sensors. The testing was made by using corn starch particles in a water solution, with its turbidity values assessed by a commercial turbidimeter provided by Libelium S.L. During measurements, the solution was in constant stirring and protected from external light. Each sensor will have its performance assessed by repeatability error analysis made to each individual measurement dimensions, this is, for each LED color transmission and scattered detection channel. The values here presented for each channel are already numerical compensated by its reference values that is obtained in the measurements with the reference photodetector.

3.2.3.1.   Short LED sensor version

The variation of light intensity with the increase of turbidity using formazine based solutions can be observed at Figure 3.13. Each individual experimental point is the result of the mean value of 10 different measurements and the error bar its standard deviation. The intensity of light for each dimension is not correlated since it is numerical compensated by the reference values and have associated an individual electronic amplification. A 2nd order polynomial fitting was implemented to the scattering experimental data based on its approximate behavior to this type of function, which also resembles the Mie solution, but being simpler to implement. To the transmission values, a Beer-Lambert law based in a double exponential fitting was performed, having better fitting

results to the data than a single exponential. The obtained fitted parameters are displayed in Table 3.9.

Figure 3.13 – Short LED sensor light intensity variation with the increase of turbidity in red dots and its fitting curves in blue. The first row is related to the scattering 90° light and the second row to the transmitted light. Each column represents the peak wavelength of the light source measured.

Table 3.9 – Summary table of the fitting parameters curves observed in Figure 3.13.

LED channel		Linear model: $y = p1 * x^2 + p2 * x + p3$			R^2	RMSE
		$p1$	$p2$	$p3$		
Scattering	430 nm	-3.452×10^{-9}	6.703×10^{-5}	3.340×10^{-1}	0.9953	0.0093
	522 nm	-3.940×10^{-9}	7.207×10^{-5}	6.246×10^{-1}	0.9956	0.0094
	660 nm	-2.503×10^{-9}	4.441×10^{-5}	7.597×10^{-1}	0.9937	0.0068
	870 nm	-1.107×10^{-9}	2.182×10^{-5}	10.680×10^{-1}	0.9881	0.0049
		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				
		a	b	c	d	
Transmission	430 nm	5.636×10^{-1}	-5.682×10^{-4}	3.540×10^{-1}	-6.390×10^{-5}	0.9945	0.0273
	522 nm	5.407×10^{-1}	-5.773×10^{-4}	3.363×10^{-1}	-6.628×10^{-5}	0.9947	0.0259
	660 nm	8.514×10^{-1}	-5.386×10^{-4}	4.169×10^{-1}	-5.076×10^{-5}	0.9941	0.0400
	870 nm	11.280×10^{-1}	-5.536×10^{-4}	5.436×10^{-1}	-5.324×10^{-5}	0.9942	0.0527

As expected from the obtained fitted curves, the infrared scattering values have presented lower variation with the increase of turbidity. It also presented higher measurement error due to its absolute low light intensity emission and detection, which requires a higher electronic amplification in comparison with the other measurement channels. A study of the accuracy variation with the increase of turbidity was performed based on the standard deviation of each measurement point and the calculated turbidity error that this standard deviation would produce (Figure 3.14). A double exponential fitting regression was made to predict the measured accuracy variation for the whole range of turbidity values. This model was chosen because it presents the best correlation with the behavior of the experimental data. Note that different scales in the measurement error are used for a better data presentation. The obtained parameters are presented in Table 3.10.

Figure 3.14 – Short LED sensor turbidity measurement error expected for both scattering and transmitted light and for each light source with the variation of turbidity itself.

Table 3.10 – Summary table of the fitting parameters curves observed in Figure 3.14.

LED channel		Linear model: $y = a*exp(b*x) + c*exp(d*x)$				R^2	RMSE
		a	b	c	d		
Scattering	430 nm	1.775×10^{-1}	-0.941×10^{-3}	0.1221×10^{-2}	5.749×10^{-4}	0.9605	0.0177
	522 nm	3.141×10^{-1}	-1.499×10^{-3}	0.6483×10^{-2}	4.027×10^{-4}	0.9769	0.0243
	660 nm	18.760×10^{-1}	-2.604×10^{-3}	0.6832×10^{-2}	4.723×10^{-4}	0.9737	0.1671
	870 nm	25.960×10^{-1}	-1.706×10^{-3}	6.9430×10^{-2}	3.091×10^{-4}	0.9560	0.2429
Transmission	430 nm	1.379×10^{-1}	-1.944×10^{-3}	5.210×10^{-2}	1.932×10^{-5}	0.9829	0.0100
	522 nm	1.305×10^{-1}	-1.681×10^{-3}	4.855×10^{-2}	3.719×10^{-5}	0.9748	0.0112
	660 nm	3.656×10^{-1}	-2.248×10^{-3}	4.101×10^{-2}	7.812×10^{-5}	0.9977	0.0146
	870 nm	1.646×10^{-1}	-2.926×10^{-3}	5.753×10^{-2}	1.378×10^{-5}	0.9534	0.0139

The behavior of the measurement accuracy with the variation of turbidity shows that the scattering channel presents, overall, a lower measurement error for both 430 and 522 nm LEDs channel, mainly in intermediary turbidity values between 1000 to 5000 FNU. Outside this range, transmission values present lower error. It is also lower for almost all range of the 660 and 870 nm LEDs channels. The lower error for the 430 and 522 nm LEDs scattering channel can be explained by the fact that the absolute light power to reach the sensing head is considerable higher, which improves the collection of the scattering light obtaining higher rate variations with turbidity, without the need of higher electric amplifications that, in general, creates lower measurement error. Overall, the obtained measurement error is higher than the ones observed for commercial turbidimeters. Only with the scattered light from 430 and 522 nm LEDs, in turbidity values from 2000 to 5000 FNU, values with 5% accuracy measurement or lower are obtained, which shows that higher light intensity detection is needed for higher resolutions. This can be proven by the development of the sensor with a longer connection to the sensing head, which will require longer optical fibers and consequently higher absorption and lower detection will be observed.

3.2.3.2. Long LED sensor version

The results of the light intensity variation with turbidity for the longer version of the LED sensor are presented in Figure 3.15. The electronic configurations required the use of higher values of amplification in all detectors due to the extra attenuation obtained by the longer optic fiber. In Table 3.11 are presented the parameters of the same fitted models used before in the shorted sensor version.
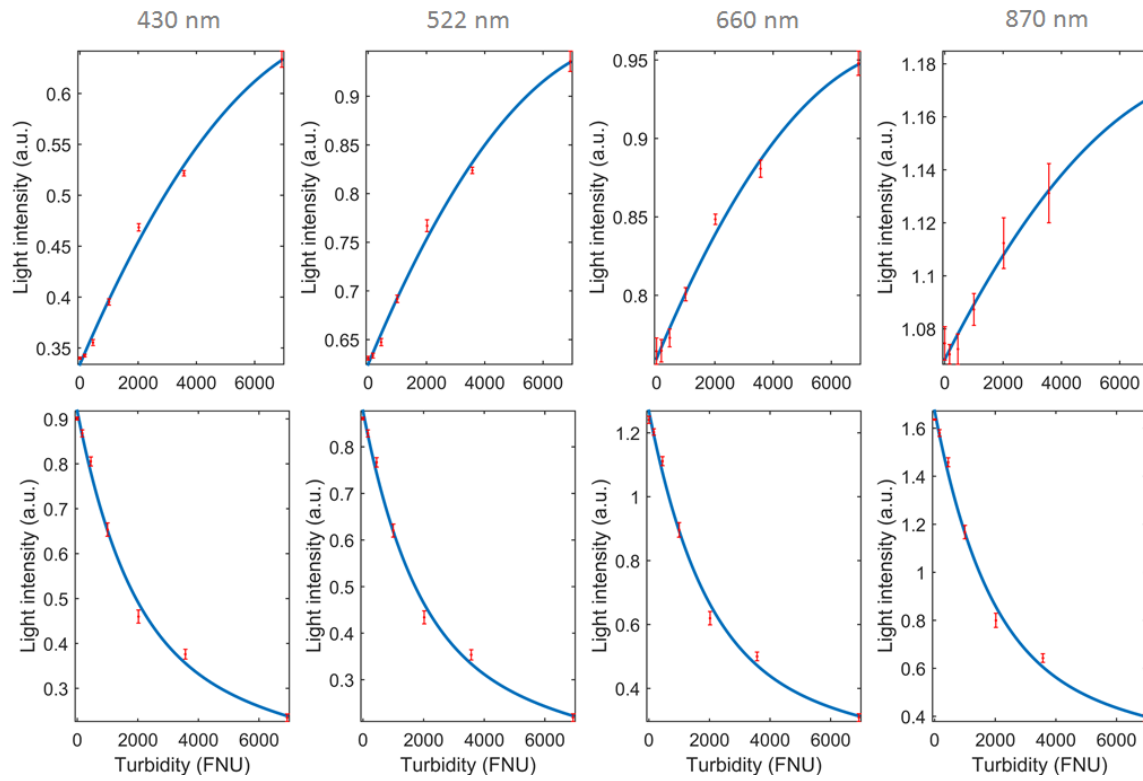
Figure 3.15 – Long LED sensor light intensity variation with the increase of turbidity in red dots and its fitting curves in blue. The first row is related to the scattering 90° light and the second row to the transmitted light. Each column represents the peak wavelength of the light source measured.

Table 3.11 – Summary table of the fitting parameters curves observed in Figure 3.15

LED channel		Linear model: $y = p1 * x^2 + p2 * x + p3$			R^2	RMSE
		$p1$	$p2$	$p3$		
Scattering	430 nm	-6.995 x 10^{-9}	1.371 x 10^{-4}	1.577 x 10^{-1}	0.9957	0.0160
	522 nm	-1.011 x 10^{-8}	2.339 x 10^{-4}	4.156 x 10^{-1}	0.9976	0.0217
	660 nm	-3.995 x 10^{-9}	0.735 x 10^{-4}	4.376 x 10^{-1}	0.9979	0.0059
	870 nm	-5.527 x 10^{-10}	0.152 x 10^{-4}	8.471 x 10^{-1}	0.9772	0.0046
		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				
		a	b	c	d	
Transmission	430 nm	4.623 x 10^{-1}	-6.125 x 10^{-4}	2.686 x 10^{-1}	-4.546 x 10^{-5}	0.9983	0.0119
	522 nm	9.271 x 10^{-1}	-6.360 x 10^{-4}	5.503 x 10^{-1}	-4.985 x 10^{-5}	0.9983	0.0238
	660 nm	7.974 x 10^{-1}	-6.162 x 10^{-4}	4.235 x 10^{-1}	-3.884 x 10^{-5}	0.9977	0.0230
	870 nm	5.292 x 10^{-1}	-5.587 x 10^{-4}	2.170 x 10^{-1}	-1.902 x 10^{-5}	0.9981	0.0131

It is possible to observe that generally, in comparison, the obtained results for this variation presents higher standard deviations than the shorter sensor version, which is a characteristic

associated with the measurement error. This is mainly evident in the scattering channel of the infrared LED where the amplification given to the photodetector was at its maximum, having a variation of light intensity considerably small, with a measurement error as high as its own light intensity variation range with turbidity. A more detailed evolution of the accuracy of measurement within the range of turbidity variation can be seen in Figure 3.16 with its curve fitting parameters in Table 3.12 using the same double exponential model as in the short sensor version to maintain consistency. Again different scales in the measurement error are used for better data presentation however, the measurement error of turbidities lower than 2000 NTU for the 870 nm wavelength is so high, that they are not represented for a better visualization and comparison.
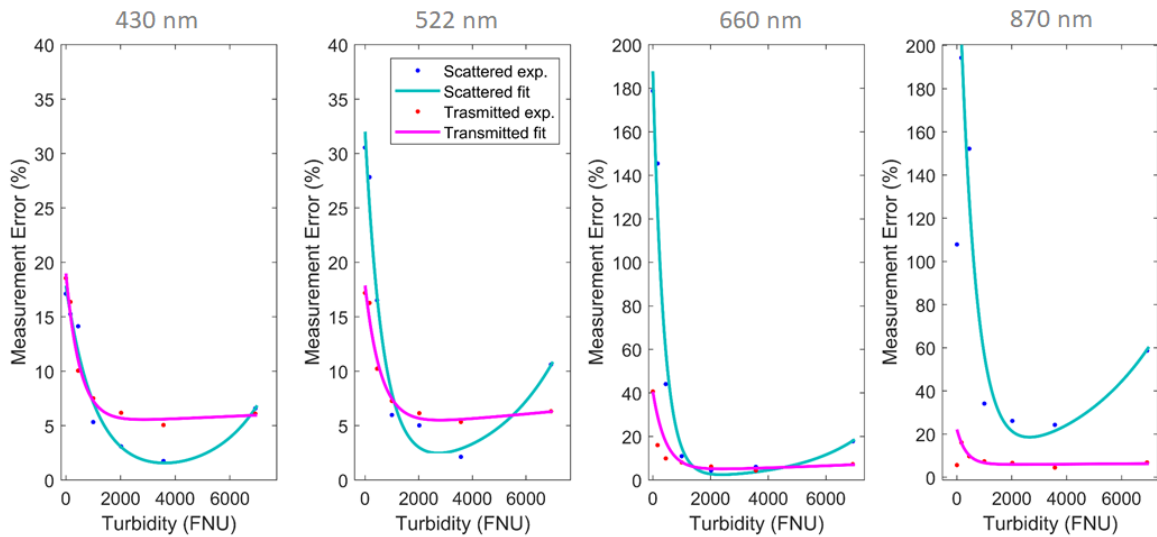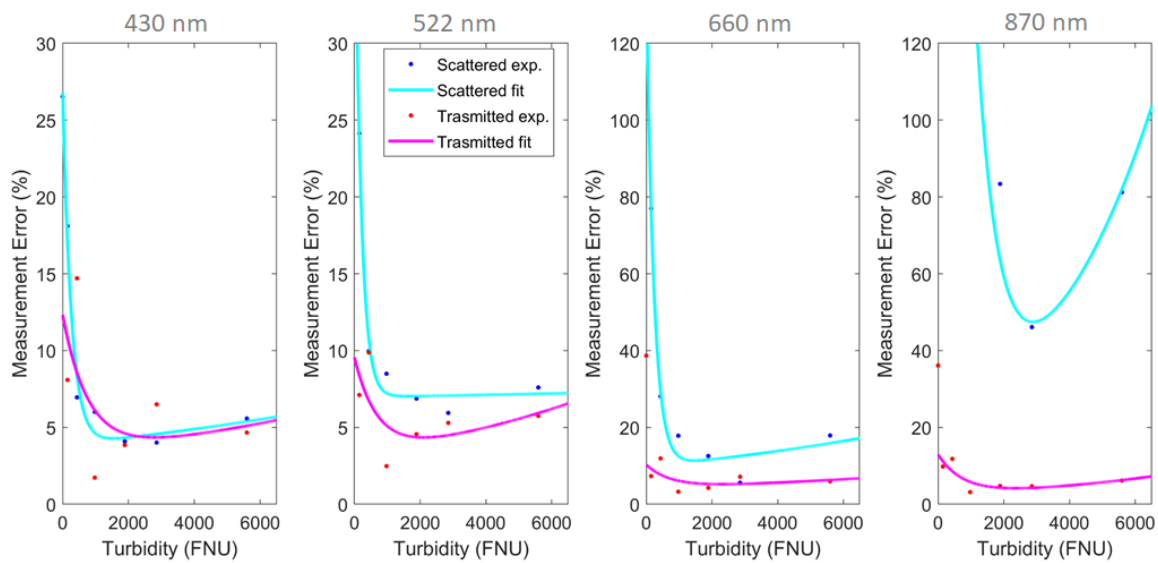

Figure 3.16 – Long LED sensor turbidity measurement error expected for both scattering and transmitted light and for each light source with the variation of turbidity itself.

Table 3.12 – Summary table of the fitting parameters curves observed in Figure 3.16.

LED channel		Linear model: $y = a*exp(b*x) + c*exp(d*x)$				R^2	RMSE
		a	b	c	d		
Scattering	430 nm	0.2304	-3.781×10^{-3}	0.384×10^{-1}	6.06×10^{-5}	0.9877	0.0137
	522 nm	0.3566	-5.120×10^{-3}	0.695×10^{-1}	6.11×10^{-6}	0.9961	0.0142
	660 nm	1.2670	-4.328×10^{-3}	0.983×10^{-1}	8.59×10^{-5}	0.9749	0.0646
	870 nm	6.1150	-1.561×10^{-3}	1.930×10^{-1}	2.58×10^{-4}	0.9429	0.8149
Transmission	430 nm	0.903×10^{-1}	-1.294×10^{-3}	0.331×10^{-1}	0.77×10^{-4}	0.3228	0.0590
	522 nm	0.636×10^{-1}	-1.439×10^{-3}	0.320×10^{-1}	1.10×10^{-4}	0.3877	0.0309
	660 nm	0.612×10^{-1}	-1.351×10^{-3}	0.415×10^{-1}	0.74×10^{-4}	0.2594	0.0414
	870 nm	1.044×10^{-1}	-1.326×10^{-3}	0.252×10^{-1}	1.62×10^{-4}	0.6485	0.0315

As expected, with the lower light intensity detected, higher measurement error was obtained, mainly in the scattering channel of all LEDs. Transmission, in the other hand, presented almost the same accuracy as the shorter LED sensor version, showing that the loss of light that was obtained in this channel was not significant. It is however important to notice that the error evolution obtained by the transmission channel presented a more inconstant behavior that can be seen reflected by the low R^2 value, which is an indication that the model used is poor to describe the data observed. To be competitive as a commercial alternative, a sensor with this length needs higher resolutions which can be obtained by increase light output power or higher resolution photodetectors. The next sensors developed and here presented try to address this issue by testing laser light sources and better photodetectors.

3.2.3.3.   Laser Sensor version

The big difference of this sensor in comparison to the former ones is the change of the blue and green LEDs to two laser diodes with a peak wavelength of 450 and 520 nm. This will not only improve the resolution of turbidity, but will also increase the color range measurement. With the other components being the same, only these two channels will be here analyzed. The obtained results of the variation of the light intensity with turbidity can be observed in Figure 3.17 and the parameters of the fitted regressions on Table 3.13.

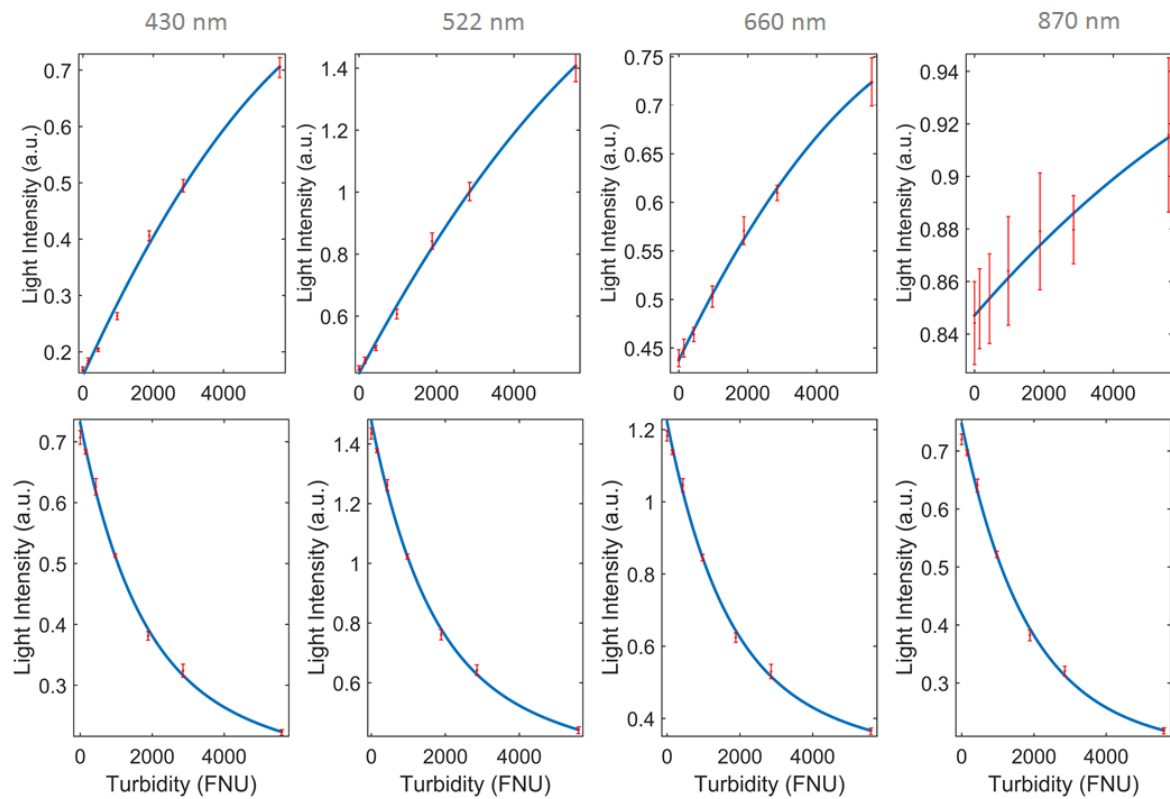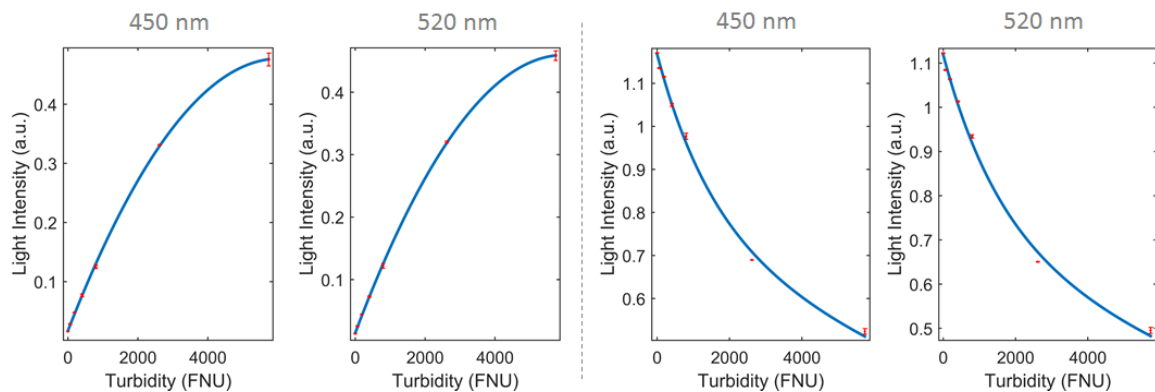*Figure 3.17 – Laser sensor light intensity variation with the increase of turbidity in red and its fitting curves in blue. Only the 450 nm and 520 nm peak central wavelength lasers were characterized since the other components are the same as observed in the long LED sensor. The scattering 90° light is represented in the left while the transmitted light is in the right.*

Table 3.13 – Summary table of the fitting parameters curves observed in Figure 3.17.

LED channel		Linear model: $y = p1*x^2 + p2*x + p3$			R^2	RMSE
		$p1$	$p2$	$p3$		
Scatt.	450 nm	-1.257×10^{-8}	1.521×10^{-4}	1.698×10^{-2}	0.9997	0.0034
	520 nm	-1.240×10^{-8}	1.485×10^{-4}	1.464×10^{-2}	0.9999	0.0025
		Linear model: $y = a*exp(b*x) + c*exp(d*x)$				
		a	b	c	d	
Trans.	450 nm	8.003×10^{-1}	-8.03×10^{-5}	3.671×10^{-1}	-6.92×10^{-4}	0.9990	0.0167
	520 nm	7.463×10^{-1}	-7.89×10^{-5}	3.720×10^{-1}	-6.66×10^{-4}	0.9988	0.0183

As expected, the higher light intensity given by the laser diode source originated a higher light intensity variation with turbidity in the scattering channel, in contrast with the ones obtained by the LEDs. This higher intensity requires lower detection amplification in the respective channels of the sensor. By consequence, it will present lower measurement error value which is seen by the almost imperceptible error bars in the plots of Figure 3.17. The assessment of this measurement accuracy with the increase of turbidity is presented in Figure 3.18 and the respective fitting parameter in Table 3.14.
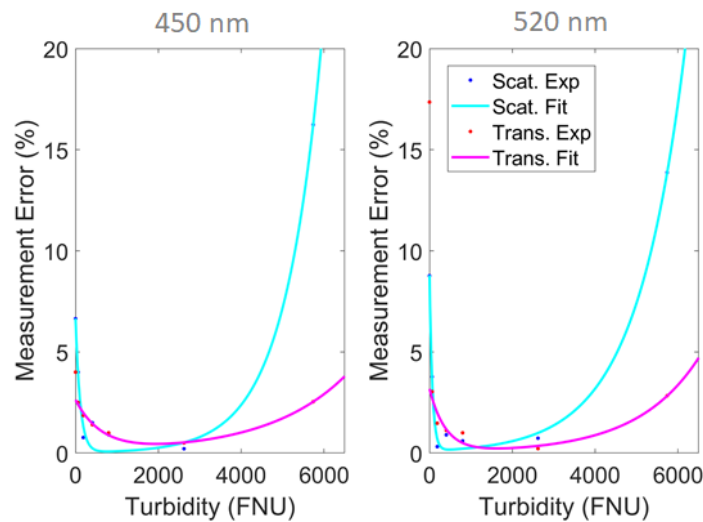
Figure 3.18 – Laser sensor turbidity measurement error expected for both scattering and transmitted light and for the 450 and 520 nm light source with the variation of turbidity itself.

Table 3.14 – Summary table of the fitting parameters curves observed in Figure 3.18.

LED channel		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				R^2	RMSE
		a	b	c	d		
Scatteri	430 nm	6.623	-8.714×10^{-3}	0.288×10^{-1}	1.104×10^{-3}	0.9862	0.9495
	522 nm	8.696	-1.446×10^{-2}	1.081×10^{-1}	8.455×10^{-4}	0.9944	0.5536
Transm	430 nm	2.506	-1.601×10^{-3}	1.216×10^{-1}	5.291×10^{-4}	0.9858	0.1542
	522 nm	3.113	-2.578×10^{-3}	0.606×10^{-1}	6.695×10^{-4}	0.8940	0.5689

It is noticeable that with the higher intensity of light using these lasers, the scattering channels present higher resolution (low error value) for lower turbidity values while the transmission channel is more appropriate for higher turbidity values. It is also noticeable that almost all the absolute error values are lower than 5% with exception of lower turbidities. Considering the error of both turbidity channels, this is, scattering and transmission, an error behavior lower of the 5% of the measurement value in all the range is possible to be obtained, which is usually the reference value of the commercial turbidimeters that measure in this range.

3.2.3.4. High Responsivity LED sensor version

The high responsivity photodetector sensor version returned to the usage of LEDs as source of light instead of the use of laser sources which would easily saturate the new high responsivity photodetectors. It used the same LEDs as the other first two versions, with the exception of the blue color that used a LED with a peak at the 470 nm as mentioned before. With this kind of photodetector, higher variation of signal is expected with the increase of turbidity while the error will be very low due to the lack of necessary electronic amplifications. The evolution of the light intensity with turbidity is presented in Figure 3.19 and the respective curve fitting parameters in Table 3.15.
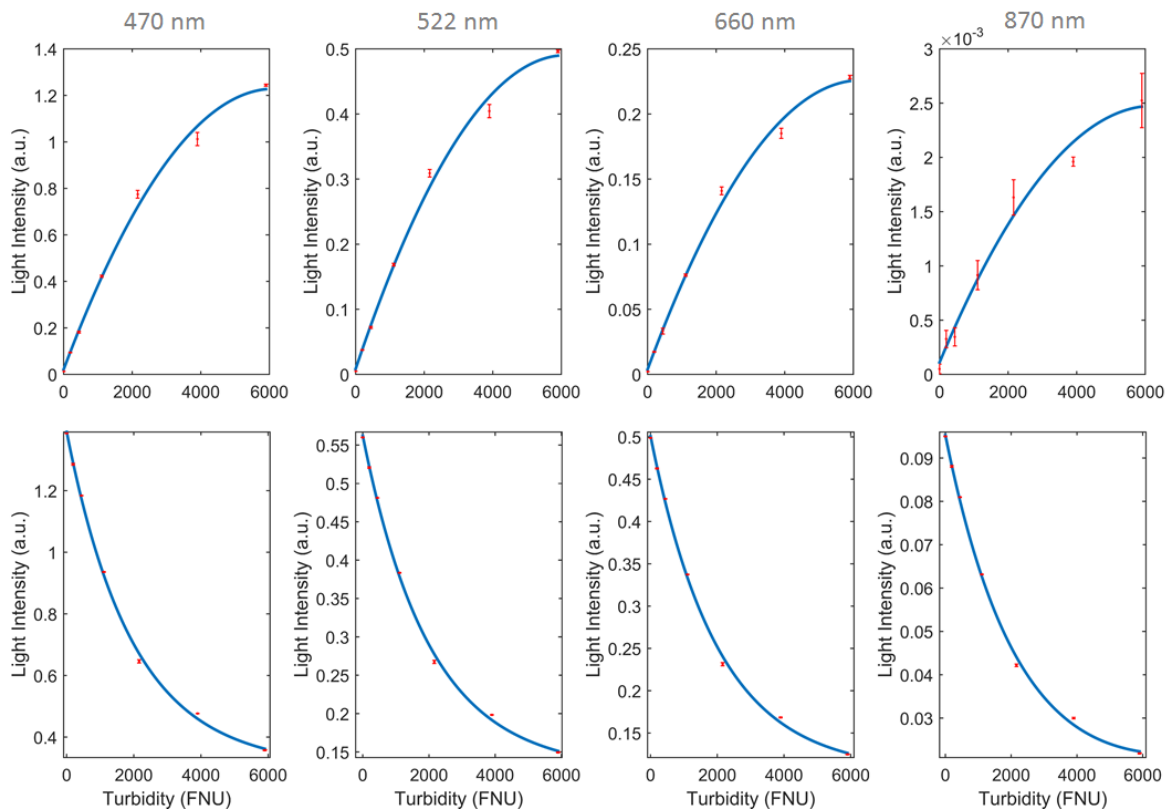
Figure 3.19 – High responsivity LED sensor light intensity variation with the increase of turbidity in red dots and its fitting curves in blue. The first row is related to the scattering 90° light and the second row to the transmitted light. Each column represents the peak wavelength of the light source measured.

Table 3.15 – Summary table of the fitting parameters curves observed in Figure 3.19.

LED channel		Linear model: $y = p1 * x^2 + p2 * x + p3$			R^2	RMSE
		$p1$	$p2$	$p3$		
Scattering	470 nm	-3.199×10^{-8}	3.931×10^{-4}	2.112×10^{-2}	0.9954	0.0400
	522 nm	-1.281×10^{-8}	1.571×10^{-4}	0.822×10^{-2}	0.9955	0.0157
	660 nm	-5.735×10^{-9}	7.138×10^{-5}	0.383×10^{-2}	0.9956	0.0072
	870 nm	-6.067×10^{-11}	7.576×10^{-7}	0.011×10^{-2}	0.9847	0.0001
		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				
		a	b	c	d	
Transmission	470 nm	1.0590	-5.181×10^{-4}	3.3330×10^{-1}	-1.111×10^{-5}	0.9988	0.0204
	522 nm	0.4262	-5.021×10^{-4}	1.3610×10^{-1}	-8.762×10^{-6}	0.9987	0.0083
	660 nm	0.3940	-4.984×10^{-4}	1.0690×10^{-1}	-2.151×10^{-6}	0.9987	0.0075
	870 nm	0.0856	-4.501×10^{-4}	0.0967×10^{-1}	8.846×10^{-5}	0.9986	0.0016

The analyzed data show the expected higher light intensity variation increase in the scattering channel in comparison with the two first sensor versions, due to the higher responsivity of the

photodetector. It is also visible the low measurement error obtained by the standard deviation of the measurements. Again, because of its low light power reaching the sensing cavity, the IR channel presents the lower variation of light with turbidity in the scattering channel. The evolution of the measurement accuracy with the turbidity is presented in Figure 3.20 and its curve fitting parameters in Table 3.19.
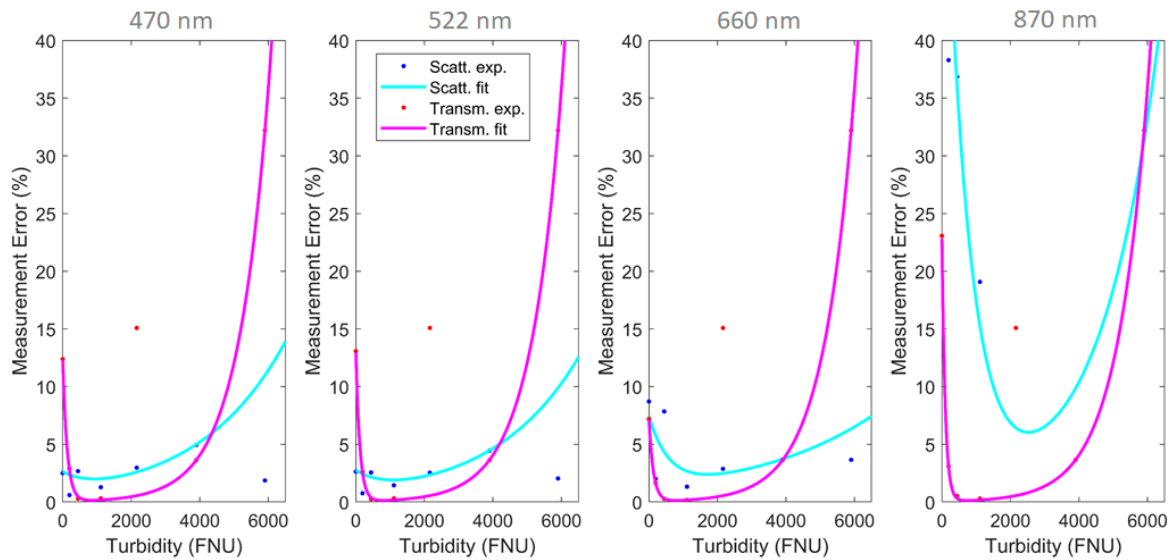
Figure 3.20 – High responsivity LED sensor turbidity measurement error expected for both scattering and transmitted light and for each light source with the variation of turbidity itself.

Table 3.19 – Summary table of the fitting parameters curves observed in Figure 3.20.

LED channel		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				R^2	RMSE
		a	b	c	d		
Scattering	470 nm	0.158×10^{-1}	-0.135×10^{-2}	1.072×10^{-2}	3.940×10^{-4}	0.8604	0.0099
	522 nm	0.182×10^{-1}	-0.126×10^{-2}	0.940×10^{-2}	3.990×10^{-4}	0.9081	0.0065
	660 nm	0.615×10^{-1}	-0.172×10^{-2}	1.308×10^{-2}	2.680×10^{-4}	0.3790	0.0388
	870 nm	6.558×10^{-1}	-0.143×10^{-2}	0.958×10^{-2}	5.910×10^{-4}	0.8893	0.1071
Transmission	470 nm	1.235×10^{-1}	-0.768×10^{-2}	5.200×10^{-4}	1.090×10^{-3}	0.9999	0.0018
	522 nm	1.302×10^{-1}	-0.843×10^{-2}	5.210×10^{-4}	1.089×10^{-3}	0.9999	0.0014
	660 nm	0.714×10^{-1}	-0.774×10^{-2}	5.180×10^{-4}	1.090×10^{-3}	0.9999	0.0001
	870 nm	2.301×10^{-1}	-1.039×10^{-2}	5.230×10^{-4}	1.089×10^{-3}	0.9999	0.0012

The higher responsivity of these photodetectors has proven to be a good cheap solution to obtain measurement error under the 5% reference of commercial applications without the use of laser sources of light. This is verified in at least the transmission channel for all LEDs, while in the scattering channel this is true only using the higher power light sources, particularly with the blue

and green LEDs, obtaining better results than the transmission channel for low turbidity values (<100 FNU). Using the information of all channels, a measurement error lower than 5% is obtainable, at least up to the values of 4000 FNU. Note that because of the high deviation to the expected behavior of the 2000 FNU data point, it was not considered in the transmitted light fitting curve. This isolated error with higher value could be due to reading fluctuations during the transmitted light measurements affecting all the channels or due to temporary instability of the system and not related to turbidity. Still the general behavior of the measurement error is obtained.

3.2.3.5. Peristaltic pump sensor version

This sensor is the most technically complex that was tested since it pumps the liquid from its environment to a sensing cell located in the same location as the electronic and optical components. Due to this proximity to the control board, less fiber optic and less attenuation is expected which will increment a higher radiant intensity that reaches the sensing cavity and thus increases the performance of the scattering channel for all the LEDs wavelength. The optical components used are the same as the previous high responsivity LED sensor. The higher intensity can be verified by observing Figure 3.21 with its curve fitting parameters in Table 3.17.
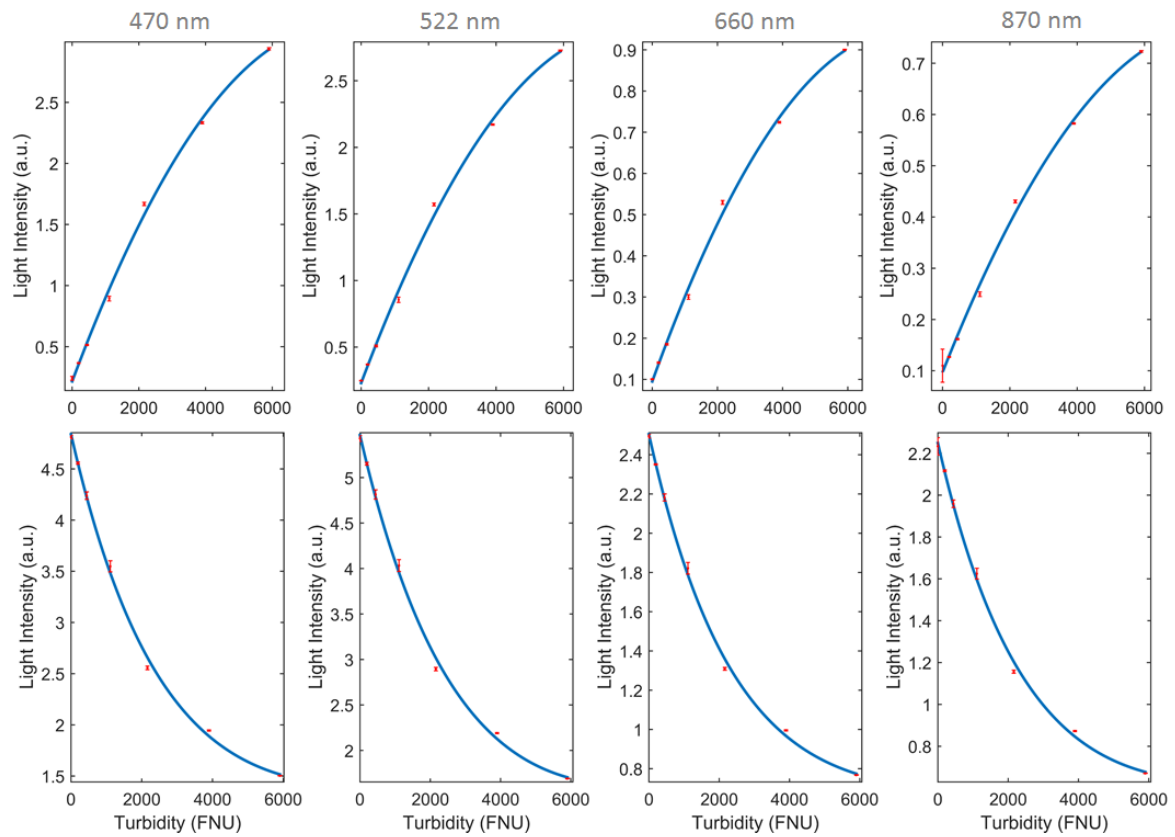
Figure 3.21 – Peristaltic pump sensor light intensity variation with the increase of turbidity in red dots and its fitting curves in blue. The first row is related to the scattering 90° light and the second row to the transmitted light. Each column represents the peak wavelength of the light source measured.

Table 3.17 – Summary table of the fitting parameters curves observed in Figure 3.21.

LED channel		Linear model: $y = p1*x^2 + p2*x + p3$			R^2	RMSE
		$p1$	$p2$	$p3$		
Scattering	470 nm	-4.574×10^{-8}	7.302×10^{-4}	2.157×10^{-1}	0.9976	0.0633
	522 nm	-4.272×10^{-8}	6.738×10^{-4}	2.301×10^{-1}	0.9975	0.0595
	660 nm	-1.416×10^{-8}	2.196×10^{-4}	0.950×10^{-1}	0.9977	0.0185
	870 nm	-1.019×10^{-8}	1.661×10^{-4}	0.970×10^{-1}	0.9967	0.0169
		Linear model: $y = a*exp(b*x) + c*exp(d*x)$				
		a	b	c	d	
Transmission	470 nm	3.996	-3.89×10^{-4}	8.421×10^{-1}	4.744×10^{-5}	0.9982	0.0791
	522 nm	4.762	-3.63×10^{-4}	7.041×10^{-1}	8.236×10^{-5}	0.9980	0.0956
	660 nm	1.979	-4.16×10^{-4}	5.252×10^{-1}	2.374×10^{-5}	0.9983	0.0395
	870 nm	1.865	-4.04×10^{-4}	3.838×10^{-1}	4.657×10^{-5}	0.9982	0.0375

A high variation in the light intensity scattered channel was obtained having, however, a standard error deviation lower that the one obtained using the sensor head, which is expected. The

measurement accuracy evolution with the increase of turbidity is presented in Figure 3.22 and Table 3.18.
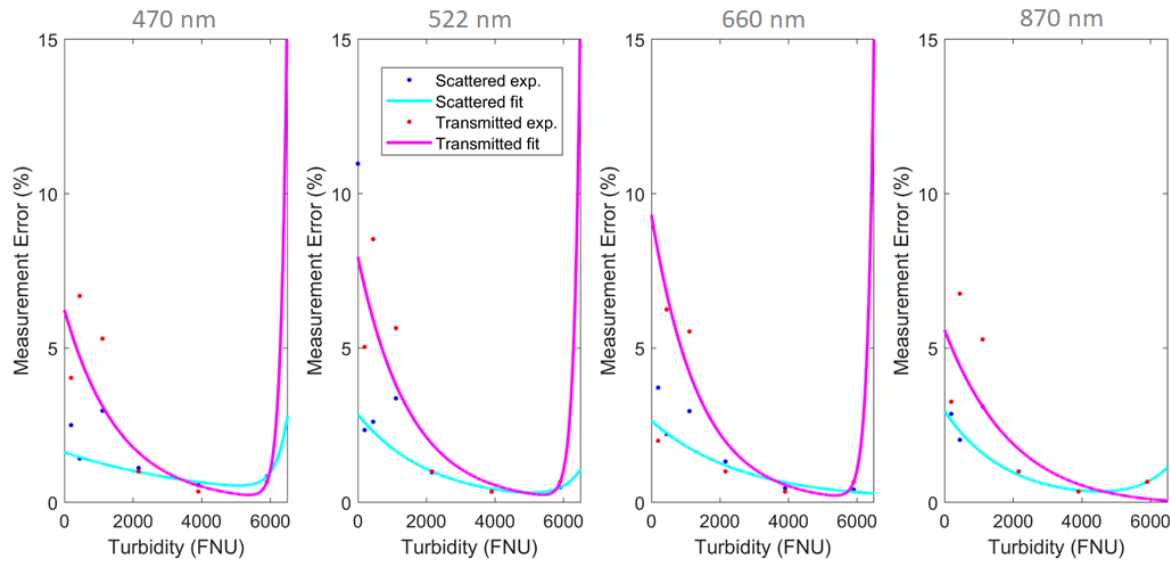
Figure 3.22 – Peristaltic pump sensor turbidity measurement error expected for both scattering and transmitted light and for each light source with the variation of turbidity itself.

Table 3.18 – Summary table of the fitting parameters curves observed in Figure 3.22.

LED channel		Linear model: $y = a * exp(b * x) + c * exp(d * x)$				R^2	RMSE
		a	b	c	d		
Scattering	470 nm	1.622×10^{-2}	-2.296×10^{-4}	1.55×10^{-10}	2.905×10^{-3}	0.9313	0.0065
	522 nm	2.853×10^{-2}	-4.843×10^{-4}	7.39×10^{-8}	1.810×10^{-3}	0.9620	0.0041
	660 nm	2.626×10^{-2}	-3.726×10^{-4}	1.524×10^{-4}	0.198×10^{-3}	0.9735	0.0039
	870 nm	2.963×10^{-2}	-5.625×10^{-4}	8.839×10^{-6}	1.093×10^{-3}	0.9678	0.0037
Transmission	470 nm	0.0624	-6.255×10^{-4}	5.078×10^{-18}	5.855×10^{-3}	0.7846	0.0255
	522 nm	0.0795	-6.637×10^{-4}	5.723×10^{-18}	5.834×10^{-3}	0.7756	0.0339
	660 nm	0.0932	-7.176×10^{-4}	1.847×10^{-17}	5.643×10^{-3}	0.9081	0.0175
	870 nm	1.4780	-3.857×10^{-4}	-1.422	-0.381×10^{-3}	0.6708	0.0308

Due to the increase of light reaching the sensing cell in comparison with the sensor head versions, all of LEDs scattering channels have error values lower than 5%. In comparison with the previous sensors, the transmission error values presented better results in high turbidity solutions. This can be explained by the higher internal light reflections that this enclosed sensing cell is more propitious to have, which promotes higher path length that is better suited for transmitted light.

For a better comparison between all the sensors measurement errors and to better analyze the evolution of the sensor development with each version increment, a summarized diagram is presented in Figure 3.23.
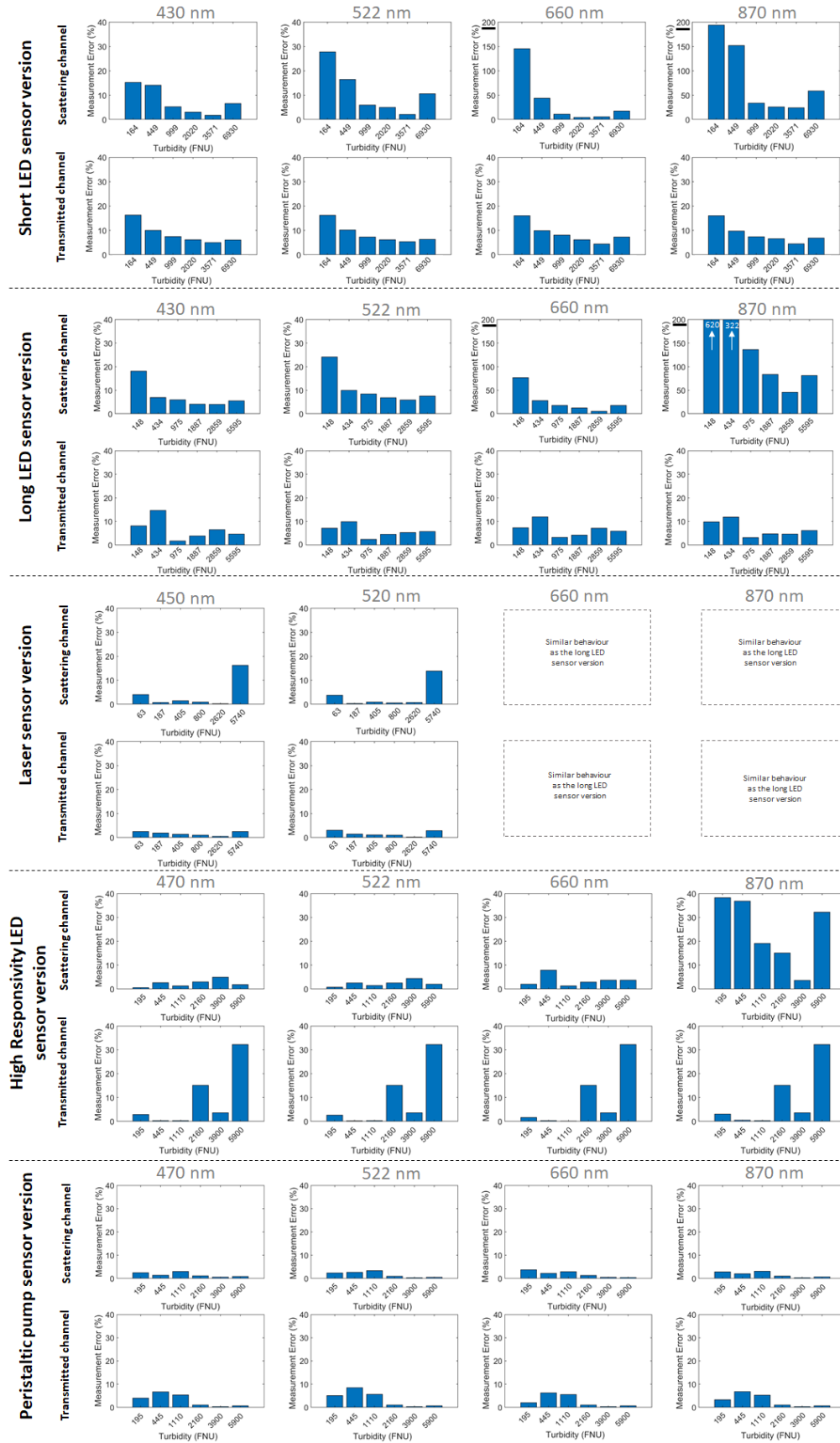
Figure 3.23 – Measurement error summary diagram for all the developed sensor versions. Notice that some of the graphics are not in the same scale and are underlined in its measurement error axis for a better identification.

3.2.4. Color measurement

The determination of color follows the principle of the Beer-Lambert law described in chapter 3.2.1.2 The problem associated with color determination resides with the presence of turbid solutions where the direct calculation of absorption is not possible. In colored turbid solutions, two different elements are participating in the decrease of light along the path of propagation. While the suspended particles are scattering light, the color is absorbing it. In Figure 3.24 this can be clearly seen by experimental data, where both turbidity and the color concentration of the solution are increased. Here, red dye was used as an example of liquid absorber agent where volumes of 0.5, 1, 1.5 and 2 ml were added to 500 ml of water. For increased turbidity, corn starch was added to the all the solutions to promote scattering centers. Nevertheless, while color attenuation is described by the Beer-Lambert law model, turbidity attenuation will present a similar behavior.
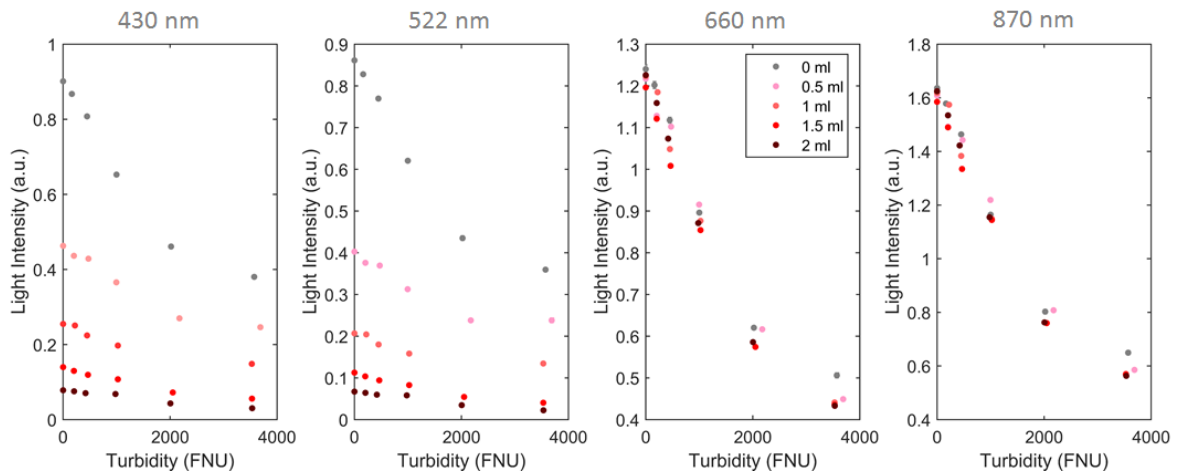
Figure 3.24 – Light intensity variation with the increase of turbidity (0 – 4000 NTU) in red dye solutions (concentration from 0 to 2 ml diluted in water). Transmission detection of light in the: 430 nm; 522 nm; 660 nm and 870 nm wavelengths.

The color absorption can be calculated separately of turbidity if the variation of light from the latter is already known and if the turbidity value is calculated by a mean that is unaffected by the color. As can be seen in the case of Figure 3.24 both red and infrared LEDs are unaffected by the red colored solution. These curves can be used to calculate turbidity separately from color. Knowing that attenuation of two different chemical species is the summation of the species attenuation separately, one can state in this case that the total attenuation (A_{tot}) is the sum of the turbidity attenuation (A_{tb}) and color absorption (A_{col}),

$$A_{tot} = A_{tb} + A_{col} \qquad \text{(Eq. 3-21)}$$

with A_{tb} being a previous known value by the light intensity variation curves (I_{tb}/I_0),

$$A_{tb} = -log_{10}\left(\frac{I_{tb}}{I_0}\right) \qquad \text{(Eq. 3-22)}$$

From $A_{tot} = A_{tb} + A_{col}$ we can correlate it with the total light intensity detected (I_{tot}) when submerged in this two-chemical species solution as:

$$A_{tb} + A_{col} = -log_{10}\left(\frac{I_{tot}}{I_0}\right) \qquad \text{(Eq. 3-23)}$$

With some mathematical logarithmic principles, one can reformulate (Eq. 3-23) and obtain the necessary methodology to calculate the color absorption in turbid liquids:

$$10^{-A_{col}} = \frac{I_{tot}}{I_{tb}} \Leftrightarrow A_{col} = -log_{10}\left(\frac{I_{tot}}{I_{tb}}\right) \qquad \text{(Eq. 3-24)}$$

This calculation was performed in the experimental data presented in Figure 3.24. To compare the results with the theoretical expected value from the increasing red dye concentration in a water solution, spectrographic measurements were taken and presented in Figure 3.25 a). Considering the spectra given by 430 and 522 nm LEDs and the responsivity of the photodetector with wavelength (Figure 3.9), the real detected transmission expected was calculated and presented in Figure 3.25 b).
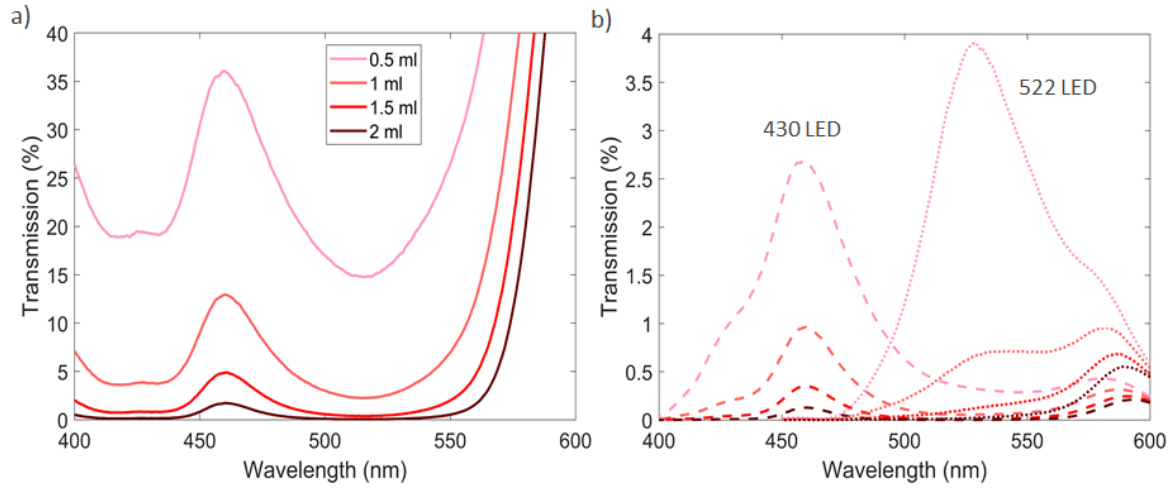
Figure 3.25 – a) Spectrographic measurements of red dye solution with different dye concentrations. b) Real transmission detected by the sensor photodetector based on its responsivity at the measured wavelengths.

By integrating the transmitted expected light variation with wavelength, the total expected normalized transmission can be obtained and the absorbance values deduced for each dye concentration. Using the methodology presented in (Eq. 3-24), a comparison between the experimental color absorbance and the expected theoretical ones is presented in Table 3.19.

Table 3.19 – Comparison between the expected color absorbance and the detected absorbance.

		430 nm LED				522 nm LED			
Dye concentration (ml)		0.5	1.0	1.5	2.0	0.5	1.0	1.5	2.0
Expected absorbance		0.61	1.13	1.55	1.86	0.63	1.32	1.72	1.90
Sensor absorbance	0 FNU	0.58	1.10	1.62	2.12	0.66	1.24	1.77	2.21
	164 FNU	0.60	1.08	1.65	2.12	0.69	1.22	1.81	2.22
	449 FNU	0.55	1.11	1.66	2.12	0.64	1.26	1.83	2.22
	999 FNU	0.50	1.03	1.57	1.96	0.60	1.19	1.75	2.06
	2020 FNU	0.46	0.99	1.61	2.06	0.52	0.95	1.80	2.20
	3571 FNU	0.38	0.82	1.67	2.20	0.36	0.85	1.89	2.41

The obtained values for the color sensor absorbance are close to what was expected from the deduced spectrophotometer color. It is important to notice that for low color concentrations, but high turbidity values (turbidity dominance), the error deviation of the calculated color is higher than when high color concentration is present (color dominance). In this region of measurement, color absorbance presents very stable calculated values.

3.2.4.1. LED wavelength color approximation

While lasers present a very narrow spectral band, LEDs have a broad band that encompasses wide range of wavelengths. In certain applications, only a specific wavelength is of interest. With a broad spectral band and without any dispersive optic element to select or filter the specified wavelength, a numerical approximation is needed based on the expected application. As we will see afterwards in chapter 5, this will be performed to approximately obtain absorbance values of the 420, 520 and 620 nm wavelengths from light of the LEDs with 470, 522 and 660 nm respectively, where the wavelengths of interest are in the range of each LED spectral band. To perform this approximation, it was used colored solutions that will simulate the conditions of which the sensor will encounter. The solution will have its absorption values of interest previously assessed by a spectrophotometer. Each solution will then be measured by the developed sensor and the obtained absorption values will then be analyzed against the expected ones from the single wavelength. A compensation regression is then performed and used from then on. Figure 3.26 present an example of compensation performed to a sensor to have approximate wavelength values of the 420, 520 and 620 nm.
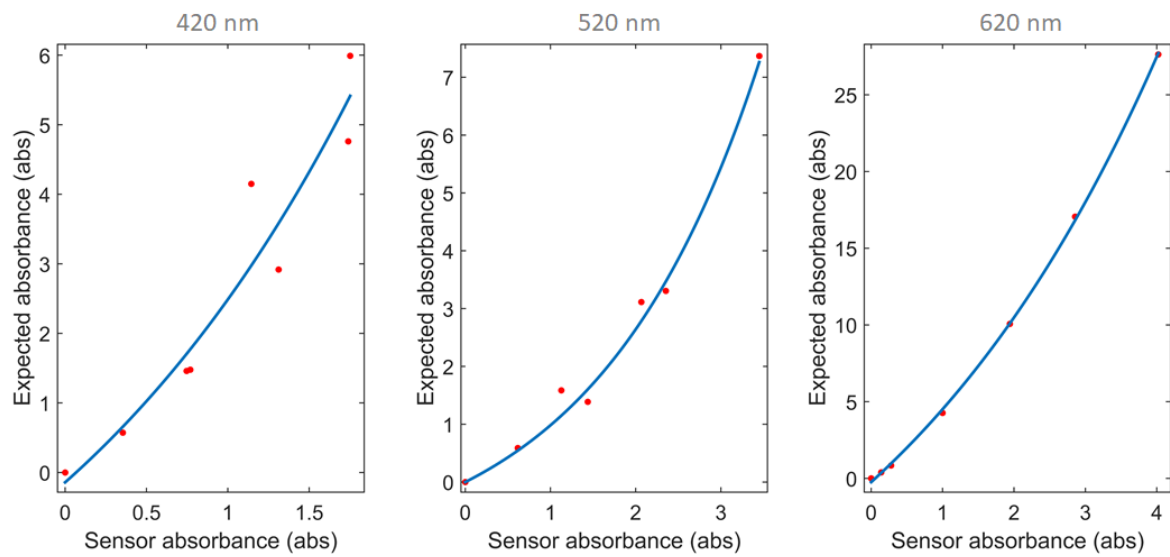
Figure 3.26 – Color compensation regression example of the absorbance obtained from the 470, 522 and 660 nm light sources to the absorbance of interest of the 420, 520 and 620 nm.

Table 3.20 – Parameters of the obtained regressions from Figure 3.26.

LED	Regression model: $y = a * exp(b * x) + c$			R^2	RMSE
	a	b	c		
470 nm	4.603	0.4519	-4.745	0.9161	0.7335
522 nm	1.428	0.5234	-1.428	0.9905	0.2949
660 nm	18.35	0.2301	-18.59	0.9997	0.2249

3.2.5. Refractive Index dependence

As discussed in chapter 3.2.1.2, the developed sensor will have variations in its light intensity when in presence of changes in the solutions' refractive index parameter. To corroborate the simulation, experimental results were carried out for clear solutions containing dissolved sucrose. A comparison with the simulation was made and presented in Figure 3.27. The refractive index of the solutions was measured by a bench refractometer model Abbemat 200 from Anton Paar GmbH (589 nm) with a temperature of 20 °C. Each solution was measured three times and the mean value calculated. As observed, the normalized light intensity experimental results are very close to what it was expected from the simulation results when using clear solutions. This variation is independent of the type of light source used and only geometry and waveguide type of material will have a changing effect.
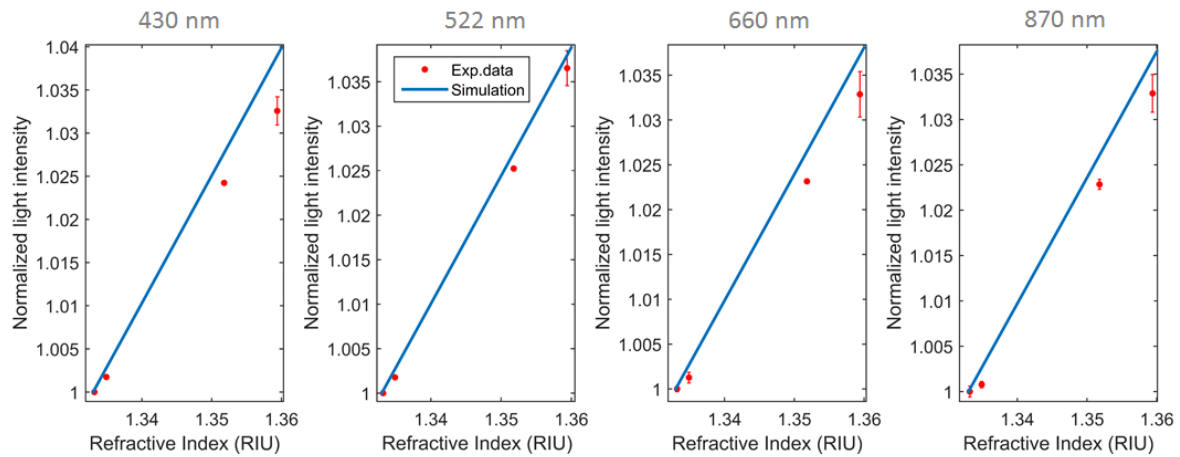
Figure 3.27 – Comparison between the experimentally obtained light intensity with the increase of the refractive index and the simulations discussed in chapter 3.2.1.2.

It is important to notice that this variation in light intensity, although it has influence in the transmission channel, it will not affect the scattering channel when in presence of clear solutions. When the refractive index is changed or is different than the one expected from water in turbid solutions, the variations in the transmitted channel will have different rates depending on the concentration of particles present in the solution. In Figure 3.28 a representation of this variation measured with the sensor is visible for solutions prepared with turbidities of 0, 1250 and 6010 FNU using corn starch. A linear fitting was also applied to the experimental values of each turbid solution variation observed for all the light sources and these values are presented in Table 3.21.
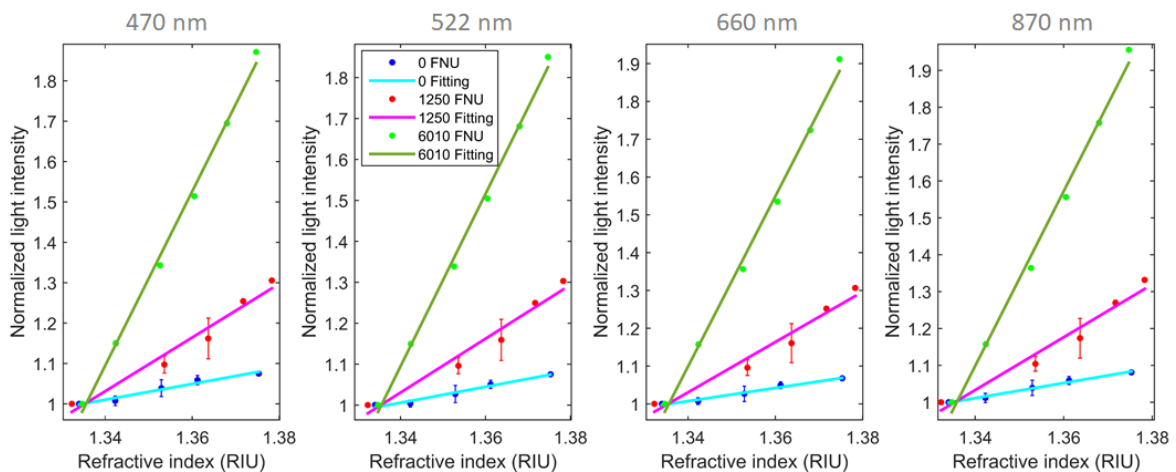
Figure 3.28 – Experimental light intensity variation obtained with the increase of the refractive index and normalized to the refractive index of water for solutions with 0, 1250 and 6010 FNU for all the transmission sensor channels light sources.

Table 3.21 –Linear regression parameters obtained from the experimental values observed in Figure 3.28.

Turbidity value (FNU)		Linear model: $y = ax + b$		R^2	RMSE
		a	b		
470 nm	0	1.962	-1.618	0.9668	0.0068
	1250	6.649	-7.879	0.9625	0.0273
	6010	17.29	-22.07	0.9952	0.0217
522 nm	0	1.950	-1.607	0.9741	0.0059
	1250	6.575	-7.780	0.9613	0.0274
	6010	17.01	-21.71	0.9957	0.0201
660 nm	0	1.741	-1.325	0.9799	0.0046
	1250	6.653	-7.885	0.9600	0.0282
	6010	18.10	-23.16	0.9956	0.0218
870 nm	0	2.052	-1.739	0.9894	0.0040
	1250	7.160	-8.562	0.9608	0.0300
	6010	19.42	-24.93	0.9941	0.0270

The different increasing variation rates in the detected light observed, in contrast to the one obtained from clear solutions, needs to be explained not only considering the geometrical factors but also the effect of the suspended particles in the solution. The light cone created in the exit of the emitting fiber will decrease its aperture with the increase of the refractive index of the solution and consequently focus more light into the receiving transmitted fiber, increasing its value of light intensity. Adding to this effect, the low aperture cone will also interact with less particles, this is, with less scattering centers and therefore, effectively, it will detect less turbidity than with a lower refractive solution by increasing the captured light that suffered less scattering (Figure 3.29).
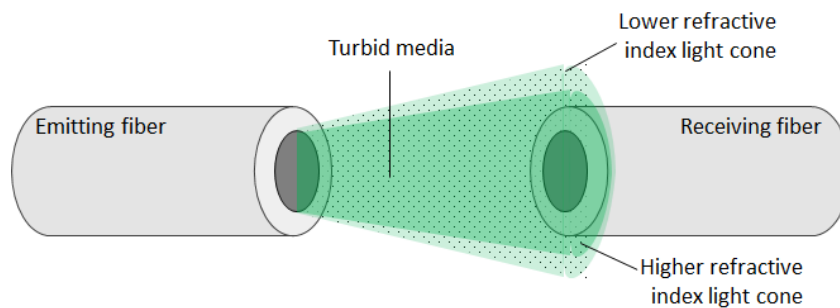
Figure 3.29 – Scheme of the created light cone from an emitting fiber in turbid media with lower and higher refractive indices. The difference of light variation detected in the receiving fiber will be dependent on the concentration of particles.

While with clear solutions there is no effect in the scattering channel since no light reaches the fiber associated to this channel whatever is the aperture magnitude of the emitting light cones, the same cannot be said when in presence of a turbid media. With turbid media, the light will be reaching the scattering fiber and therefore changes in the geometry of the emitting light from the fiber will have effects in the detected scattering light. As discussed, with a low aperture light cone, this is, with a more focused light in the transmission channel, less light will be scattered and detected in the related channel. Again, adding to this effect, there is the lower interaction with the particles to scatter light and so, as a result, even less light is scattered. This will manifest as a decreasing in the detected light with the increase of the refractive index. Experimental values that show this variation can be seen in Figure 3.30 for each light source of the sensor with turbidities of 1250 and 6010 FNU. The respective linear fitting parameters are presented in Table 3.22.
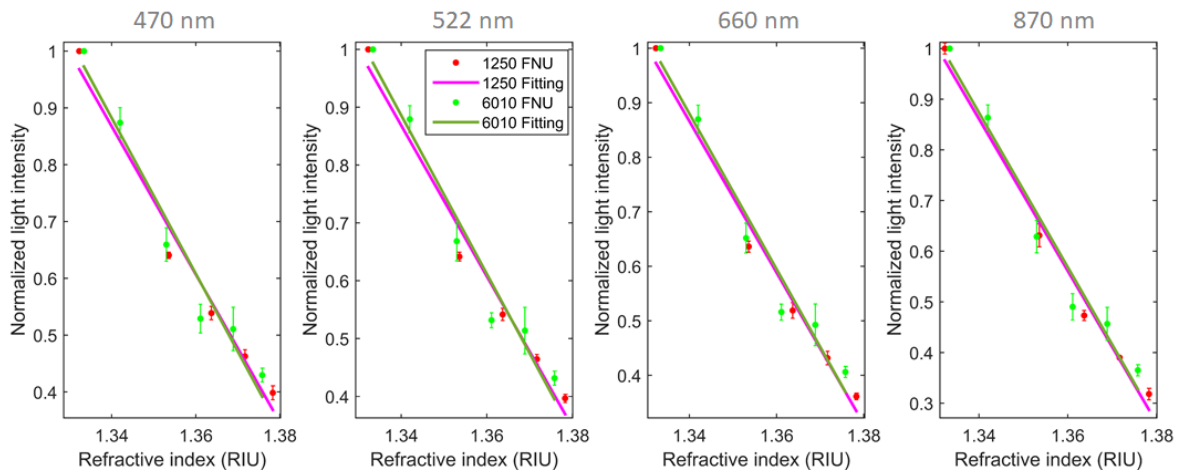
Figure 3.30 – Experimental light intensity variation obtained with the increase of the refractive index and normalized to the refractive index of water for solutions with 1250 and 6010 FNU for all the scattering sensor channels light sources.

Table 3.22 –Linear regression parameters obtained from the experimental values observed in Figure 3.30.

Turbidity value (FNU)		Linear model: $y = ax + b$		R^2	RMSE
		a	b		
470	1250	-13.03	18.33	0.9783	0.0403
	6010	-13.71	19.25	0.9635	0.0481
522	1250	-13.04	18.34	0.9797	0.0390
	6010	-13.72	19.27	0.9657	0.0466
660	1250	-13.88	19.46	0.9861	0.0342
	6010	-14.25	19.98	0.9677	0.0469
870	1250	-14.98	20.94	0.9888	0.0331
	6010	-15.25	21.30	0.9710	0.0475

Since there is light intensity variation dependence with the refractive index of the medium and it plays an important role in the measurement of a solution of interest, this parameter needs to be assessed *a priori* and a compensation carried on in the sensing algorithm to obtain the correct information of the solution Figure 3.31. An in-line and real time low cost refractive index sensor is needed to complement the turbidity/color sensor.
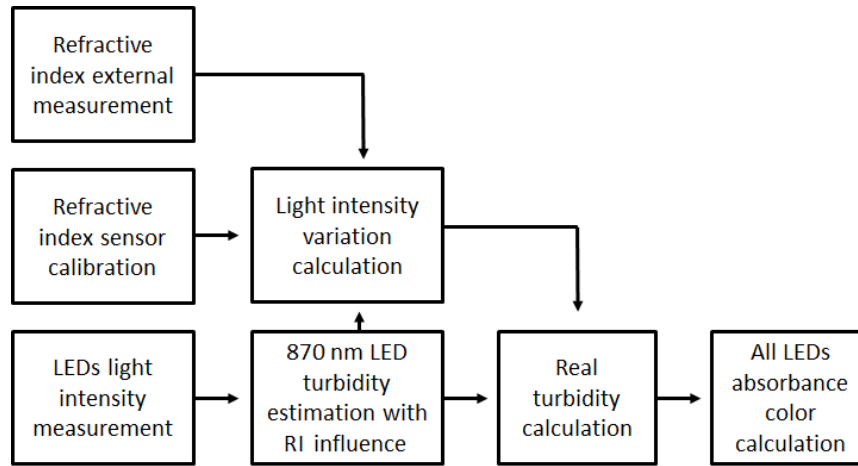
Figure 3.31 – Schematic diagram for possible process to be used for refractive index compensation.

3.3.   Refractive Index Sensor

3.3.1. Working principle

As mentioned before, the technology of which the designed and developed refractive index sensor is based uses the light interaction between two different media, being the mainly one a prism and the other the solution to measure. The sensor was developed with the idea of being used as in-line and real-time monitoring of liquid environments. The working principle of this technology (Figure 3.32) is similar than the one observed for the optical fiber waveguide described in 3.2.1.1 where reflection (R), described by the Fresnel equations (Eq. 3-25), takes the main role.

$$R = \frac{1}{2}\left(\left|\frac{n_1\cos(\theta_i)-n_2\cos(\theta_t)}{n_1\cos(\theta_i)+n_2\cos(\theta_t)}\right|^2 + \left|\frac{n_1\cos(\theta_t)-n_2\cos(\theta_i)}{n_1\cos(\theta_t)+n_2\cos(\theta_i)}\right|^2\right) \qquad \textit{(Eq. 3-25)}$$
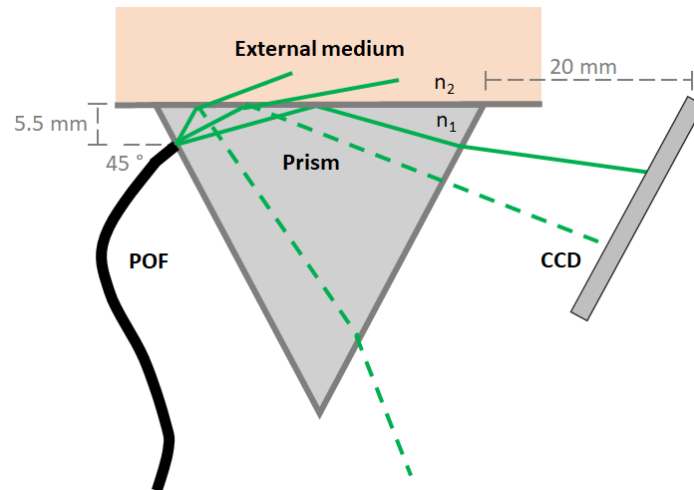
Figure 3.32 – Scheme of the working principle of the refractive index prism base sensor.

A glass prism (30x30x30 mm) with refractive index (n_1) of 1.46 is used as the component that will be in contact with the external medium and where light will propagate. In this component, it is important to guarantee that the material, from which light will interact with the external medium (n_2), has a refractive index higher than the one expected from the latter and which is intended to measure. The fulfillment of this requirement guarantees that total internal reflection can occur.

A light source will irradiate to one surface of the prism, in the direction of the boundary between the prism and the external medium. In the case of the developed sensor, the light will come from a plastic optical fiber (POF GHCP4001 from Mitsubishi) since it has a numeric aperture (NA=0.5) which is ideal since it will not spread much the light in a broad cone as would happen with a direct usage of a LED, but it will be broad enough to have a good range of light ray angles. Using the POF also prevents the usage of a set with an LED and converging lens to focus the light, which simplifies the sensing system in comparison with a traditional refractometer.

When light reaches the interface with an incident angle θ_i below the critical angle of total reflection, this is, when θ_t=90°, it is partly transmitted into the liquid, whereas for higher angles the light is totally reflected, creating a dark and a lit region on the other face of the prism. The angles of which the light will be reflected or transmitted will be totally dependent on the liquid refractive index value. At some distance from the prism, a CCD sensor, model TCD1304DG from Toshiba Electronic with 8 µm pixel resolution, is placed to detect the lit/dark boundary region. Depending on the detected position of this boundary, the refractive index of the liquid can be assessed. Since only the position of the boundary will be the parameter to measure, the light intensity is irrelevant

as far as it is enough to be detected and does not saturate the CCD sensor. The LED light source used was the IF E93 (522 nm) model from Industrial Fiber Optics and presented in chapter 3.2.2.3.

A simulation of the expected CCD detection results from this device was made using all the parameters described for the materials and geometrical measurements presented in Figure 3.32. The simulation starts by calculating the cone of light angles created by the fiber through its numeric aperture. Snell law (Eq. 3-2) is used to calculate the propagation direction of this range of angles, taking in consideration the air-prism boundaries before and after the interaction with the prism-liquid boundary. The results of the simulation performed with a refractive index variation range between 1.333 and 1.350 can be seen in Figure 3.33, where the displacement of the lit/dark boundary is visible.
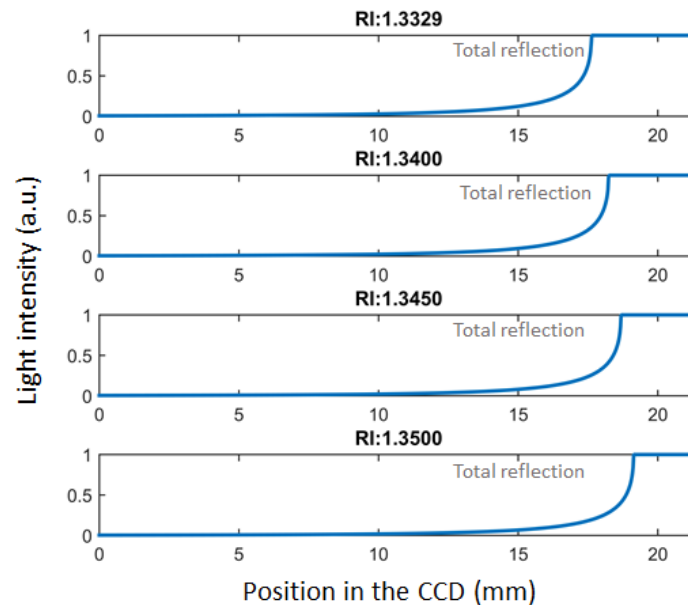
Figure 3.33 – Simulation results obtained for the RI core sensor components used and for values between 1.333 and 1.350.

3.3.2. Sensor design and measurement

For the sensor prototype and to accommodate the presented core components of the sensor in the desired distances and angles, a solid structure is needed. This structure was made by 3D printing which is placed in a waterproof box having a rectangular cut section in the top to place and assure the contact of the prism with the external medium. The box will act as the sensor head when submerged for measurements in the intended liquid. A temperature probe sensor (Maximum Integrated DS18B20) is also present inside the box, in direct contact with the prism. The

measurements will be dependent on its temperature and its compensation is needed after each measurement. The electrical connections to the CCD and to the temperature probe are possible through a lateral cut section of the box, properly sealed, from which the cables will exit to connect to the electronic control board placed in a polymeric protection case. This board ensures the control of the LED, CCD, temperature sensor and wireless communications (Figure 3.34).

Figure 3.34 – Photography of the refractive index stand-alone sensor head.

The LED and CCD are controlled by an electronic board from STMicroelectronics model NUCLEO-F411RE connected to a Raspberry Pi 3 for GSM communication. When activated to measure, it turns on the light during the necessary time to obtain the profile of light obtained by the CCD, this is, the intensity is detected for each one of their effective 3650 pixels. The typical profile of measurement, when in a liquid, can be observed in Figure 3.35 with the line in blue.
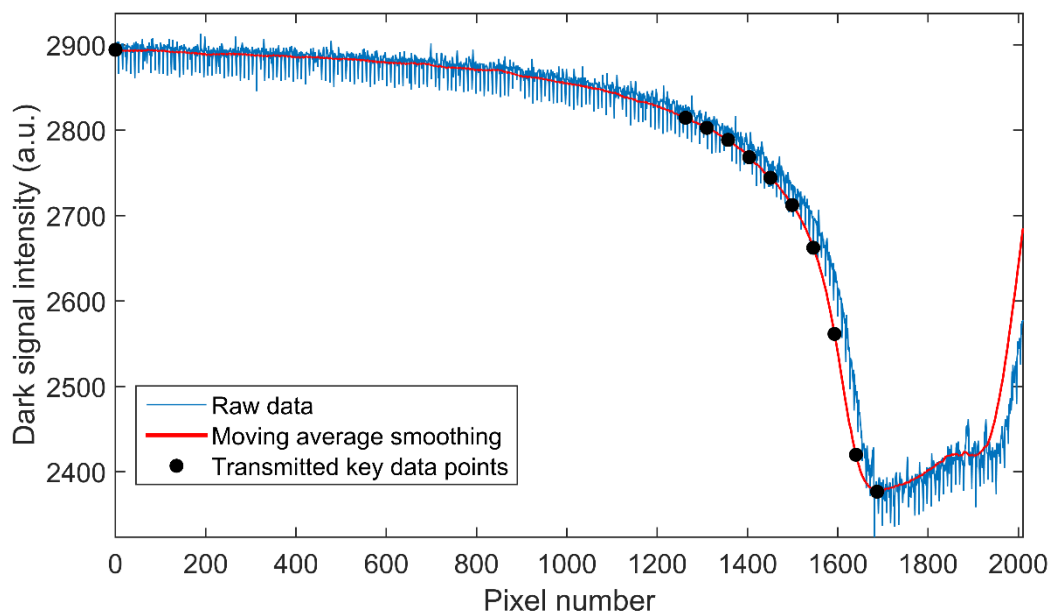
Figure 3.35 – Typical RI profile measurement in a solution obtained from the CCD. A moving average smoothing is performed to obtain key data points from the profile.

As can be seen, the measured signal presents some noise but the profile of the signal is perfectly visible. It is important to notice that the CCD raw measurement is a direct measurement of the dark signal intensity. This means that the maximum values obtained will be related to a darker region, while, when detecting light, the values will be lower. So, by analyzing the typical profile, we can deduce that the first pixels are in a darker region that becomes lesser darker until it reached a spot with total internal reflection. The slope observed reaches the maximum value at the boundary of the dark and light regions. The signal then increases again after the light region since it is a permanent dark region and cannot be used for measurement.

With a data communication system based on wireless transmission, one of the limitations of this approach is the quantity of data possible to send without overwhelming the whole network, mainly if many sensors would be used. This makes impossible to send the whole group information obtained with each pixel and instead simple data manipulation must be done within the hardware of the sensor itself. One step to lower and simplify the data needed for transmission, is the implementation of a smoothing technique based on a moving average (red line in Figure 3.35). This simplify the raw data for a strategic key points gathering, which include: the first value of the curve, which is the maximum; the minimum value of the curve; and 9 representative points of the boundary slope. The latter are important to identify the position of the boundary for an effective refractive index measurement (dark dots in Figure 3.35). These key points are all the necessary information to successfully reconstruct the smoothed curve in a cloud server where the data is sent without major constrains with communication protocols.

When the data is received in the cloud server, the reconstruction of the curve is done using a linear interpolation between the key values with an added extension of the minimal value for more 1000 pixels. This extension will be important for later data analysis using sigmoid fitting. A comparison of the original smoothing curve and the reconstructed curve obtained from the used key points can be seen in Figure 3.36 in blue.
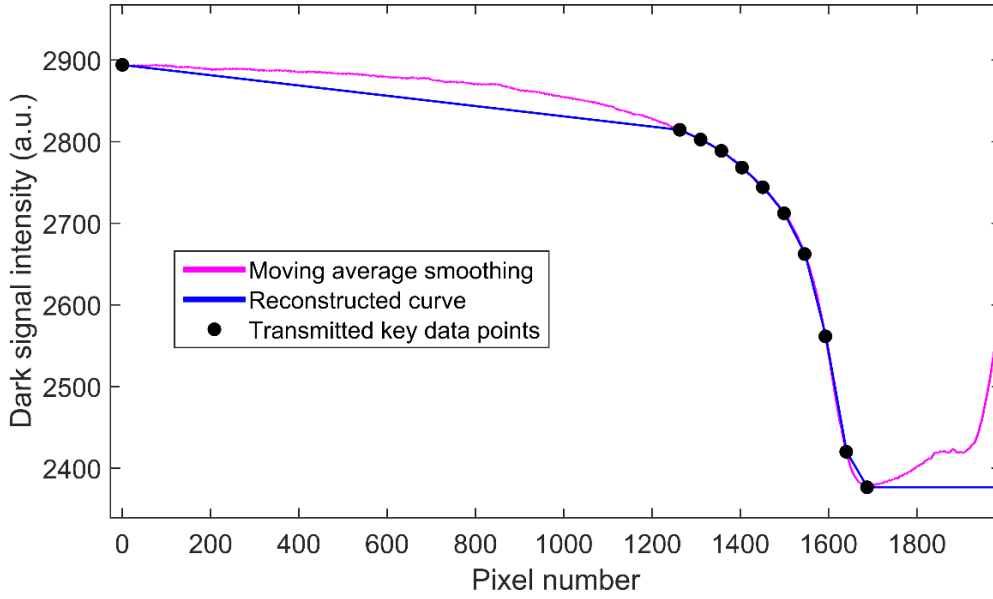
Figure 3.36 – Reconstructed RI profile of a solution performed on the cloud server after receiving the key data points.

The reconstructed curve will then be analyzed to infer the position of the pixel that will be related to the boundary and consequently to the refractive index value measured. This value will be obtained by a sigmoid fitting using the least squares regression with a Levenberg–Marquardt algorithm applied to the reconstructed curve. A sigmoid is described by the following equation:

$$y = \frac{A_1 - A_2}{1 + e^{(x - x_0)/dx}} + A_2 \qquad \text{(Eq. 3-26)}$$

The fitting will calculate 4 parameters that will describe the sigmoid. A_1 and A_2 are the parameters that indicate the value of the sigmoids asymptotes, with A_1 the upper value and A_2 the lower. The slope of the sigmoid is based on the dx value and calculated by $(A_1 + A_2)/4dx$. Finally, the most important parameter is the central point position of the sigmoid which is the one that will be used as the reference point of the boundary. This point has the coordinates of $(x_0, [A1 + A2]/2)$ with x_0 the value that will indicate the measured pixel position (Figure 3.37).
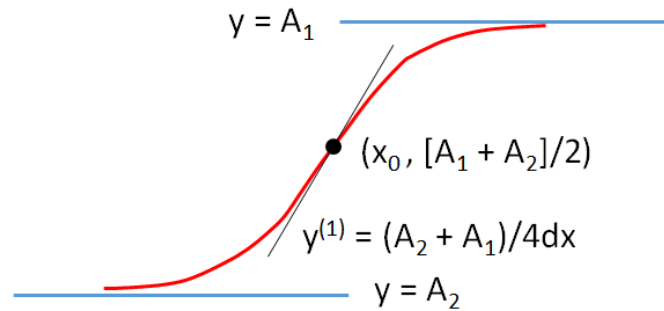
Figure 3.37 – Scheme of the sigmoid fitting with the identification and visual representation of all the parameters necessary to its description.

The central point as the definition of the position of the boundary is more important than other alternatives like the minimum value. This central point of the sigmoid is less affected by light variations and turbid liquids, registering only a variation on the slope value. The same cannot be said for the minimal value which is more sensible and change its position with these variations. An example of a reconstructed curve and its sigmoid fitting is presented in Figure 3.38.
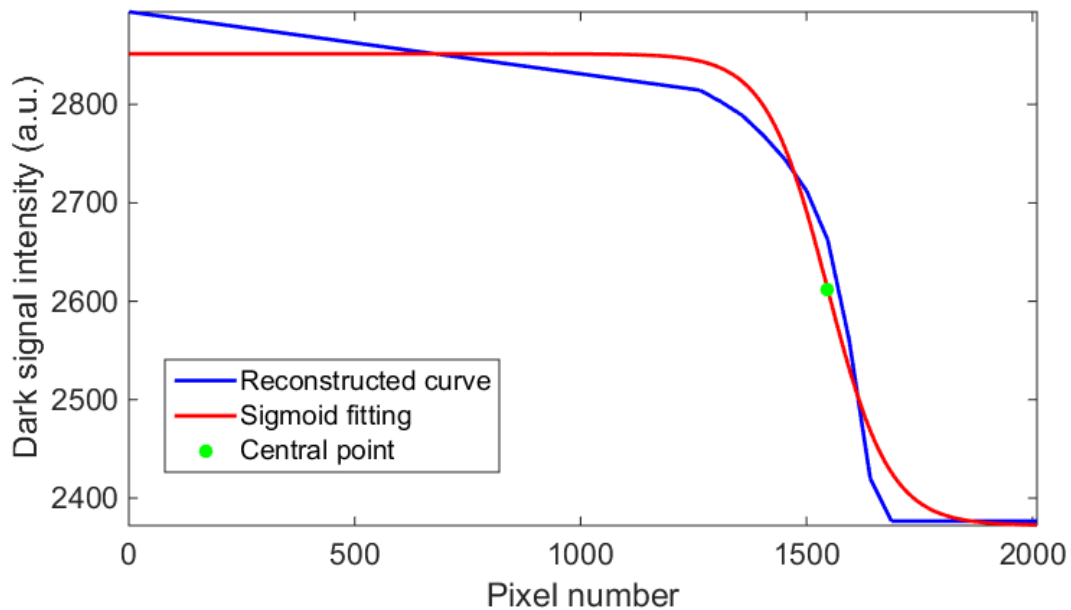
Figure 3.38 – Example of sigmoid fitting applied to a RI reconstructed profile curve with the calculation and visual representation of the central point.

A schematic summary of the whole algorithmic process performed to the raw data obtained described in this chapter is presented in Figure 3.39.
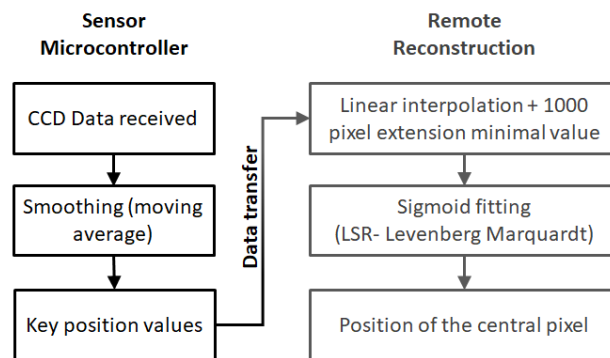
Figure 3.39 – Example of sigmoid fitting applied to a RI reconstructed profile curve with the calculation and visual representation of the central point.

3.3.3. Refractive Index characterization

The refractive index sensor was characterized by increasing the refractive index of a liquid in which the sensor was submerged. Water solutions with sucrose were prepared for this propose and the effective refractive index values were measured in the bench refractometer model Abbemat 200 from Anton Paar GmbH (589 nm), with a temperature of 20 °C. For each sample, three independent measurements were made also at 20 °C. Figure 3.40 presents the evolution of the reconstructed signal with the increase of the refractive index in a range between 1.3329 to 1.3500 RIU. The figure also shows for each curve the related sigmoid fitting with its central point clearly identified in a green dot. The fitting parameters are presented in Table 3.23.
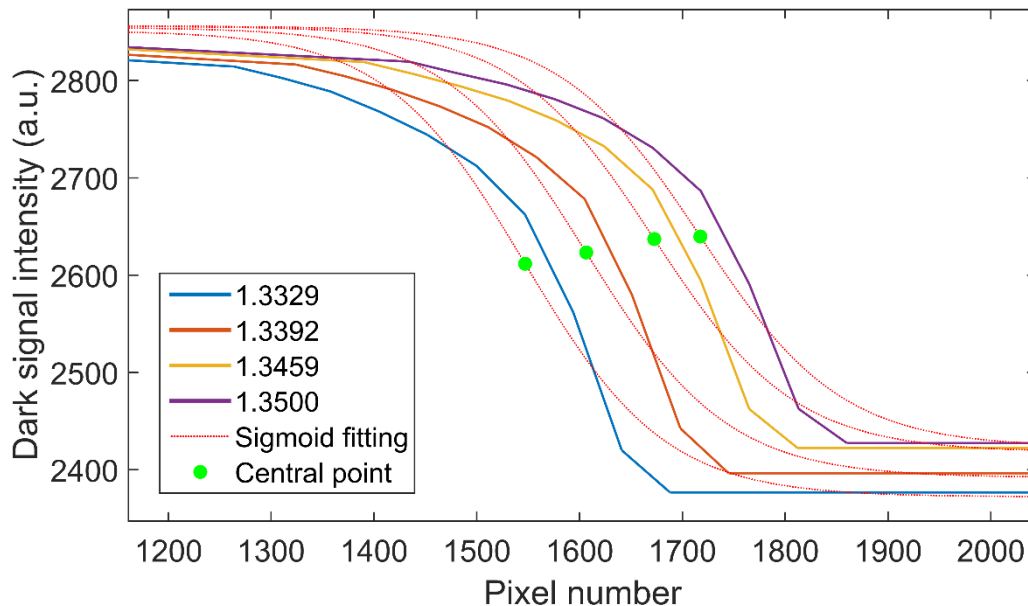
Figure 3.40 – Evolution of the central point obtained from the fittings of the reconstructed signal with the increase of the refractive index in a range between 1.3329 to 1.3500 RIU.

Table 3.23 – Parameters of the sigmoid fittings from Figure 3.40.

Refractive index (RIU)	Curve fit model: $y = \frac{A_1 - A_2}{1+e^{(x-x_0)/dx}} + A_2$				RMSE
	$A1$	$A2$	x_0	dx	
1.3329	2851	2372	1547.0	68.2	1095
1.3392	2855	2392	1606.4	68.6	1128
1.3459	2856	2418	1672.6	68.4	1100
1.3500	2856	2423	1717.5	68.5	1093

As can be seen, there is a clear shift of the central point position with the increase of the refractive index, as expected by the simulations performed in chapter 3.3.1. As the refractive index of the external solution increases, light rays with higher angles will not be totally reflected anymore and so, the dark region detected will be larger. This variation is linear and can be used as a calibration function for the refractive index sensor. A representation of this variation is shown in Figure 3.41 together with the simulated variation (in blue) expected from a sensor with the described dimensions. Both are normalized for the same pixel position attained for water. The parameters obtained for each curve are presented in Table 3.24.
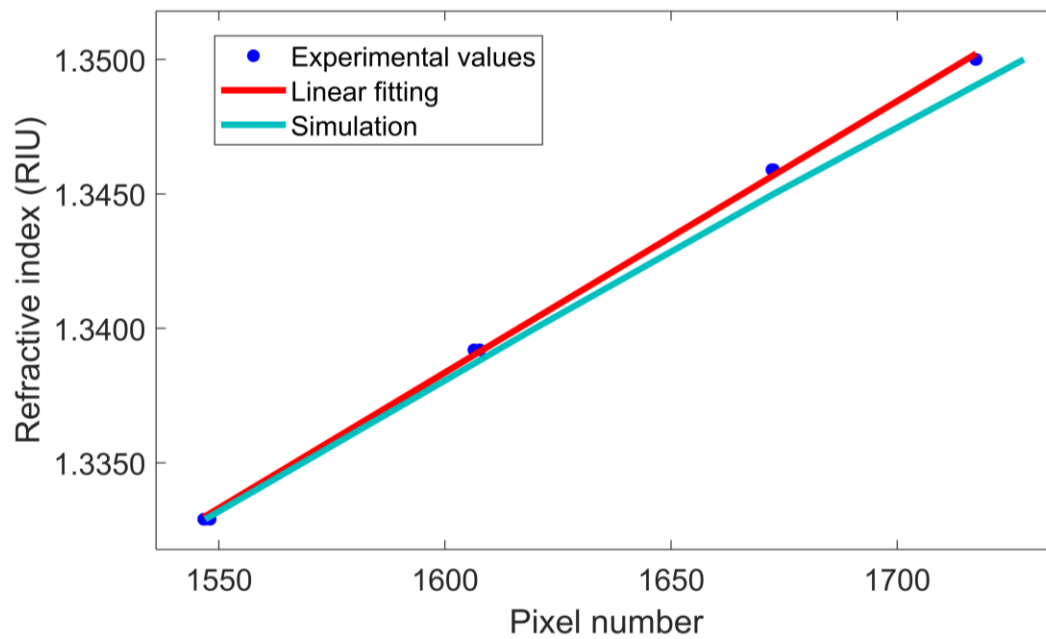
Figure 3.41 – Comparison of the experimental values obtained by the sensor with the increase of refractive index and the expected behavior from the simulation.

Table 3.24 – Parameters of the obtained regression of the experimental values and the simulation from Figure 3.39.

	Linear model: $y = ax + b$		R^2	RMSE
	a (RIU)	b (RIU)		
Linear fitting	1.01 x 10^{-04}	1.177	0.9992	1.98 x 10^{-04}
Simulation	0.95 x 10^{-04}	1.187	-	-

The simulation showed to be close to the regression of the experimental values. The slope, given by the parameter a in the linear model, is a direct value of the resolution obtained from the sensor. The resolution obtained by the experimental values was 1.01 x 10^{-04} RIU while the simulation indicates a theoretical resolution of 0.95 x 10^{-04} RIU, a slightly difference expected by the constrains and non-ideal conditions of a real measurement. Other agreement that both experimental and

simulated values share is the maximum refractive index value of measurement of about 1.3760 RIU for 25 °C. For solutions with a value above that, the boundary slope of measurement will disappear since the light angles necessary to measure values with that magnitude and beyond are not created.

Depending on the chosen geometry and configuration for the sensor, mainly the position and angle of the plastic optical fiber and the distance to the CCD, better resolution could be obtained to the detriment of the measurement range and *vice-versa*. For a specific application, this constrains need to be careful addressed and a proper design made.

It is also important to assess the influence of the measurement of temperature, turbidity and color of the solution since the intended purpose of this sensor is to be used in industrial conditions where liquids and solutions have these different properties. These tests were carried in the following chapters.

3.3.4. Temperature

A temperature characterization of the sensor was made inside a climatic chamber to guarantee a constant temperature of the solution during the measurement process. The positions of the sensor head elements were slightly adjusted to guarantee a better attachment for this characterization. For a better precision of the prism's temperature measuring process, the temperature probe that is in contact with the prism was used. The prism will be the element that will change with higher impact the optical properties of the sensor with temperature. It is then important to be assured that it has the same temperature has the solution to measure. A variation of the refractive index measurement composed by 5 samples between values of 1.3300 to 1.3760 RIU was made for 3 different temperatures: 25, 35 and 45 °C (Figure 3.42). Its real values were measured by the bench refractometer mentioned in the former chapter. For each sample, the measurement was repeated 3 times. A linear fitting was performed to each temperature variation and its parameters presented in Table 3.25.
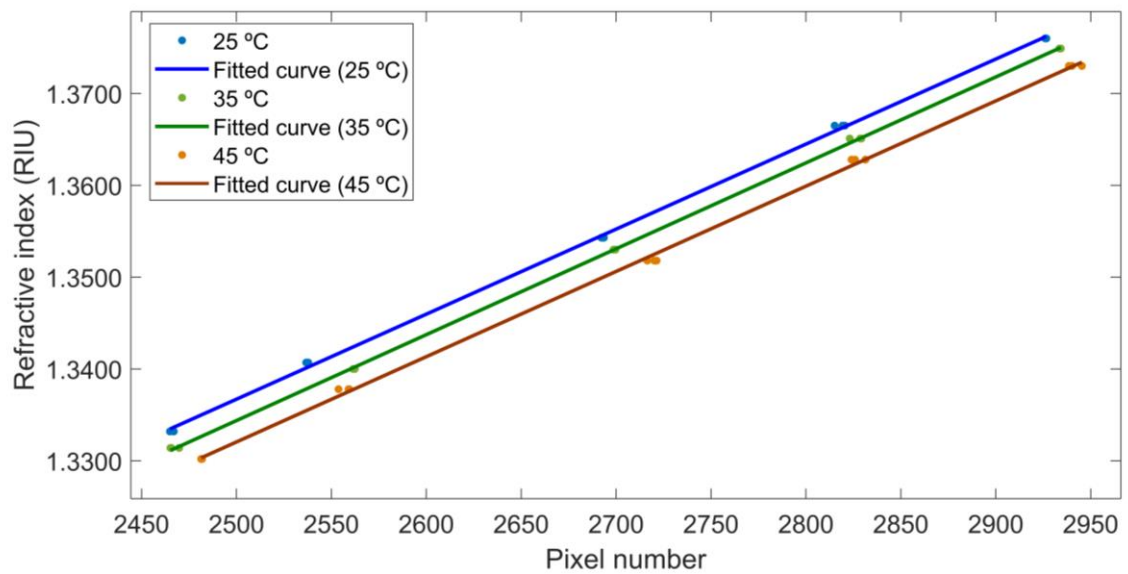
Figure 3.42 – Pixel position increase with the refractive index obtained with the developed sensor when measured in clear solutions with 25, 35, and 45 °C of temperature.

Table 3.25 – Parameters of the obtained regressions from the temperature variation measurements from Figure 3.42.

	Linear model: $y = ax + b$		R^2	RMSE
	a (RIU)	b (RIU)		
Linear fitting (25 °C)	0.93 x 10^{-04}	1.105	0.9994	4.01 x 10^{-04}
Linear fitting (35 °C)	0.94 x 10^{-04}	1.100	0.9999	2.02 x 10^{-04}
Linear fitting (45 °C)	0.93 x 10^{-04}	1.100	0.9992	4.74 x 10^{-04}

The obtained results show that the slopes of the linear regressions are almost the same independently of the temperature of the liquid and the sensor head. With its increase, only a shift of the proportion line that described the relation between the pixel position and the refractive index is obtained. For higher temperatures but with the same refractive index, a higher pixel position number that detects the boundary is necessary. This is expected since the temperature changes both the refractive index of the solution and of the prism. Therefore, the optical properties of the sensor are also changed. In the sensor calibration, this variation of temperature will need to be compensated.

3.3.5. Turbidity

Other important experimental condition that needs to be assessed and that will be found in real measurement conditions is the presence of a turbid solutions. To analyze the behavior of the sensor with these conditions, water solutions containing corn starch particles were prepared with turbidity

values of 3650 and 7370 FNU. For each turbidity value solutions, sucrose was added and measured with the developed sensor. For comparison, a clear solution (0 FNU) was also measured in the same conditions. Every turbidity and refractive index values were measured using the same commercial sensors described previously in this work. Again, for each sample, the measurement was repeated 3 times and a linear regression was made for each turbidity value. The graphical representation of the results and the regression parameters can be observed in Figure 3.43 and Table 3.26 respectively.
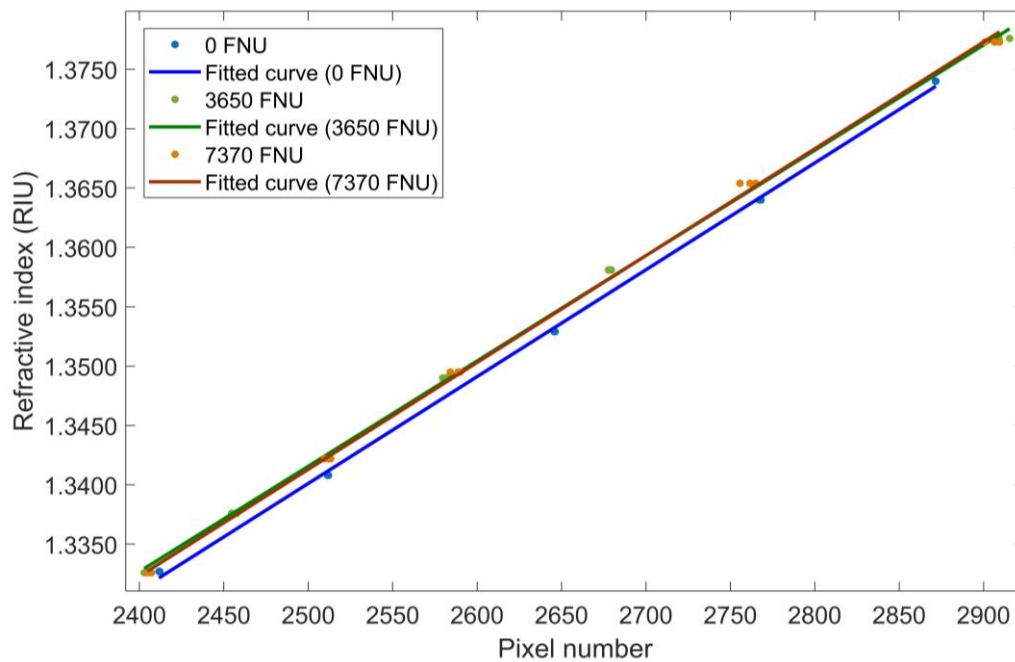
Figure 3.43 – Pixel position variation with the refractive index obtained by the developed sensor when measured in turbid solutions with 3650 and 7370 FNU and its comparison with a clear solution.

Table 3.26 – Parameters of the obtained regressions from the turbidity variation measurements from Figure 3.43.

	Linear model: $y = ax + b$		R^2	RMSE
	a (RIU)	b (RIU)		
Linear fitting (0 FNU)	0.90×10^{-04}	1.115	0.9993	4.22×10^{-04}
Linear fitting (3650 FNU)	0.89×10^{-04}	1.120	0.9993	4.61×10^{-04}
Linear fitting (7370 FNU)	0.90×10^{-04}	1.116	0.9991	5.13×10^{-04}

When the solution to measure presents some degree of turbidity, the sensor will identify the boundary pixel position with a shift from the value expected from a clear solution. The slope of the regression continues to have the same rate, independently of the turbidity value. The shift of the regression due to turbidity is the same from both medium turbidity value of 3650 FNU and high turbidity value of 7370 FNU, which indicates that this is the maximum shift value that can be

observed. The shift will create a displacement of the boundary position of around 15 pixels which will have an equivalency of about 1.5×10^{-3} RIU. The shift is created due to the light that is reflected and scattered by the particles and is redirected to the CCD which do not exist in clear liquids. This light that will not be spatially homogenous will change slightly the profile of the detection signature in the CCD which promotes the shift. The higher the turbidity, the higher will be the shift, until reaching the saturation point observed for both 3650 and 7370 FNU. It is also important to notice that the identification of the boundary pixel position for the same solution has higher dissimilarity when in presence of turbidity, as observed by the increase of the root mean square error with the higher turbid solution values. The shift value of 1.5×10^{-3} RIU will therefore be the new resolution value of this sensor when in presence of turbid solution of unknown value.

3.3.6. Color

The last assessment necessary to be done to the sensor is its behavior when submerged in a solution with high color pigment concentration, mainly the color which the light source used interacts the most. Because of the green 522 nm LED used in the prototype sensor, the color which will interact the most with this wavelength is a red based dye solution, mainly due to absorption when it is diffracted to the exterior of the prism. Other color dyes will interact less or not interact at all. A red dye was heavily added to a water solution (850 ml) with two different concentrations (4 and 14 ml) and increased its refractive index with addition of sucrose. The same procedure was done for a blue dye (10 ml). For comparison, an equivalent solution without dyes (no color) is also made. For each solution, its RI was measured by both the sensor and bench refractometer at 20 °C. The results are presented in Figure 3.44 jointly with the parameters of the linear regressions (Table 3.27) made to the experimental values.
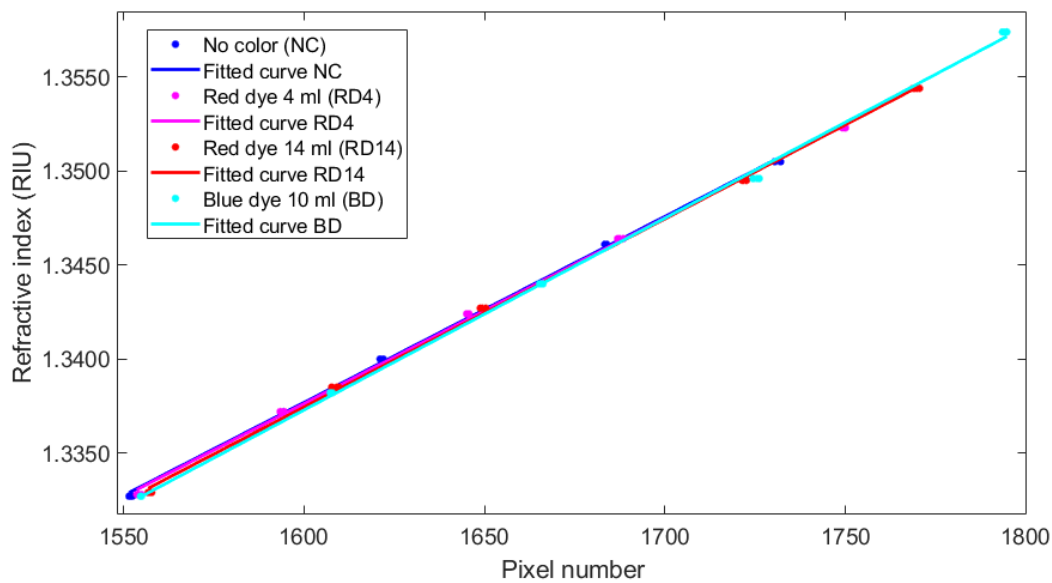
Figure 3.44 – Pixel position variation with the refractive index obtained by the developed sensor when measured in colored solutions with red and blue dyes and comparison with a no colored solution.

Table 3.27 – Parameters of the obtained regressions from the color variation measurements from Figure 3.44.

	Linear model: $y = ax + b$		R^2	RMSE
	a (RIU)	b (RIU)		
Linear fitting (NC)	0.99×10^{-04}	1.179	0.9991	2.00×10^{-04}
Linear fitting (RD4)	0.99×10^{-04}	1.180	0.9992	2.20×10^{-04}
Linear fitting (RD14)	1.00×10^{-04}	1.177	0.9994	2.09×10^{-04}
Linear fitting (BD)	1.02×10^{-04}	1.174	0.9990	2.88×10^{-04}

The results obtained clearly show that the addition of dyes does not affect the measurement of the sensor when in presence of color solutions. This result was expected since in presence of a red dye, the green light that is not totally reflected in the prism and goes to the exterior environment will be absorbed. On the other hand, in water and in the blue dye solutions', the light that is not totally reflected will be transmitted through the solutions. In the end, none of the light of each case will interact with the CCD and therefore the same behavior is obtained. The slightly differences observed in the regressions are due to the intrinsic resolution and error of the sensor and some temporary temperature fluctuations.

4. *Multivariate data analysis*

4.1. Introduction

As seen in chapter 3.2, the developed turbidity and color sensor will have a multivariate data reading that will contain information of both color and turbidity in its 8-dimensional values. Four of them are constituted by the scattering light of the blue, green, red and infrared light sources and the other four from the transmitted light from the same light sources. Instead of using only the infrared information as the primary method, as discussed before, all this dimensional data can be used to extract, in a better and efficient way, the values of turbidity and color of a solution to measure. Two different approaches for the sensor multidimensional data analysis can be performed. One of them is based on the learning experience of previously measured information, named Machine Learning. The other approach, named Data Fusion, tries to combine data in an intelligent way so that the final calculated result can be achieved with a combination of the best characteristics of each data dimension. While the refractive index will play an important role in the final measurement, variations on its value will depend exclusively on the desired application of the sensor. It is also a parameter that is obtained independently of the turbidity and color and will not enter as a factor to the multivariate data analysis of the sensor. Instead, in real measurement conditions and depending on the chosen algorithm, a compensation is made before or afterwards.

4.1.1. Machine learning

Machine learning is the subfield of computer science that gives computers the ability to perform tasks associated with artificial intelligence (AI). These tasks can involve recognition, diagnosis, planning, robot control, prediction, etc. There are several reasons why machine learning is preferred to a direct artificial system design that performs as desired [Nilsson, 1998]:

- Some tasks are only well defined by example. The input and output pairs can be well specified but the relation between them is not known or is difficult to obtain.
- Important relationships and correlations in large piles of data can be hidden if a method of extracting these relationships is not used.

- Some human design may not work well as desired in the field, mainly when the working characteristics of the environment are not well known. On-the-job improvements can solve this issue.

- Certain tasks need a high amount of knowledge that might be impossible for timely human encoding. Gradual learning by a machine might be able to capture more information than humans.

- Changing of the environment over time needs a constant learning adaptation.

Machine Learning uses mathematical models built by statistics theory, with its core task to make inferences from a data sample. It is being used as a multidisciplinary tool for multivariate scientific problems that can be found in engineering [Potdar, 2015], sociology [Habernal, 2013], neuroscience [Hanke, 2010], biology [Swan, 2013], chemistry [Sunny, 2014] or robotics [De Greeff, 2015]. Two phases can be identified as steps to solve a problem using machine learning techniques. In the first phase, training data is used to solve optimization problem with an efficient algorithm. Once the model is learned, a second phase begins with the representation of the solution with efficiency. Efficiency, in this field, is not only associated to its predictive accuracy but also to its space and time complexity [Alpaydın, 2010].

When both input data *X* and output data *Y* are known parameters used in the training phase of the model, it is designated as a supervised learning. In an optimal scenario, it will allow the trained algorithm to correctly determine the class labels for unseen instances. This requires the algorithm to generalize from the training data to unseen situations. The cases which the aim to assign the input data to one of a finite number of discrete categories, are named as classification problems. These can be simple binary classification from two classes, as the example of Bayesian statistics [Shahriari, 2016; Salmerón, 2018] and Support Vector Machines [Orrù, 2012; Zendehboudi, 2018], or multiclass classification as: Decision trees learning and Random forests [Kotsiantis, 2013; Torizuka, 2018]; Artificial neural networks [Behler, 2015; Cao, 2018] and k-nearest neighbors [Dhanabal, 2011; Suominen, 2013]. If the output consists of one or more continuous variables, then the problem is characterized as a regression.

Other pattern recognition problems use a set of input data that has no corresponding target values. This type of problems is named as unsupervised learning which have the aim to find groups of correlated data by clustering it through similarity or by density estimation [Jain, 2010]. Search of "hidden" states in unsupervised temporal data analysis can also be done using the Hidden Markov models [Fink, 2014; Calvo-Zaragoza, 2019]. These types of techniques are not used exclusively for unsupervised data, as they can be used for supervised or semi-supervised data, which are a

category where the training set has some supervised data but most of it are unsupervised. More than one algorithm can also be used to improve the efficiency of learning and inference taking. Sometimes, for high-dimensional space data, a projection of it to a lower dimensional space of two and three dimensions could be advantageous performance-wise, as observed in the examples of the Principal Component Analysis (PCA) [Abdi, 2010] with unsupervised data and Linear Discriminant Analysis (LDA) [Stewart, 2014] with supervised information.

The third type of learning is called reinforcement learning and its main concern is to maximize the reward by, given situation, taking the appropriate action, which is not persistent in space or time. Contrary to the supervised learning, here there aren't any outputs to learn, but instead the algorithm must learn by discovering the outputs through a process of trial and error that can be based in Monte Carlo methods [Robert, 2016]. The learning algorithms that can be used with this methodology are the same based on the supervised information discussed before.

With enough obtained data to create a training set, Machine Learning techniques seem a good approach for the multivariate data analysis necessary for the turbidity and color sensor that can use all the dimensionally gathered information to achieve better accuracy in efficient away.

4.1.2. Data Fusion

Multivariate data fusion, as a subfield of computer science, has the goal of combining information from multiple sources to obtain new or higher quality inferences that a single source cannot achieve by its own. These sources can be from different sensors, human reports or archived data. The subject achieved the emerging technology status during the 1990s with the start of micro and nanoscale sensor proliferation, but today it is considered a mature viable technology that can be applied in real-time due to the emergence of advanced processing techniques, improved hardware and wide-band communication. Multisensory data fusion can be found in several military applications where, historically, it was the field where they were developed first. Examples include: automated target recognition; autonomous vehicles; remote sensing and surveillance; automated threat recognition; health status monitoring; etc. In the last years, nonmilitary applications by technology transference include: monitoring of manufacturing processes; condition-based maintenance of complex machinery, environmental monitoring, robotics, and medical applications [Hall, 2008; Khaleghi, 2016].

In the data fusion techniques, the information obtained from multiple sensors is combined in some sense using an appropriate fusion rule or formula that enhances the observed observation. It

is not, however, a technology that lacks flaws. There are seven challenges or issues in data fusion that one needs to take in consideration when dealing with this topic. These are reported by Hall and Steinberg [Hall, 2000]:

- There is no substitute for a good sensor: a single accurate sensor that measures the phenomena to observe will be better than multiple inaccurate data fusion sensors.

- Downstream processing cannot make up for errors in upstream processing: errors in processing of individual sensors data cannot be corrected by data fusion.

- The fused answer may be worse than the best sensor: failure to accurately model sensor performance will result in corruption of the fused results.

- There are no magic algorithms: there is no perfect algorithm that is optimal under all conditions.

- It is difficult to quantify the value of data fusion: measure of performance can be obtained for sensors or processing algorithms but measures of mission effectiveness are difficult to define.

- Fusion is not a static process: the data fusion is an iterative dynamic process that seeks to continually refine the estimates about an observed situation.

Three physical configurations for sensor data fusion are usually used. In many cases, more than one of these configurations is implemented. They can be categorized as follow [Raol, 2016]:

- **Complementary**: sensors are placed with the goal of giving a more complete picture of the phenomenon that is being observed. It is performed with simultaneously views from different perspectives. An example is the four radars around a geographical area. This arrangement also provides a certain redundancy in measurements (Figure 4.1 a).

- **Competitive**: each sensor provides independent measurements of the same property to measure. This configuration is robust and fault tolerant and each sensor could be identical or non-identical. An example can be observed for three thermometers competing to measure the temperature of a body. Here, each sensor gives a central numerical value with a small uncertainty range around this value. The fusion algorithm returns one value based on the three measurements and their associated uncertainties (Figure 4.1 b).

- **Cooperative**: The data are provided by two independent sensors to derive information that might not be available from only one sensor. The stereoscopic vision system is an example of this category. The two-dimensional (2D) images obtained from two cameras mounted at slightly

different incidence angles can provide a three-dimensional (3D) image of the observed scene (Figure 4.1 c).
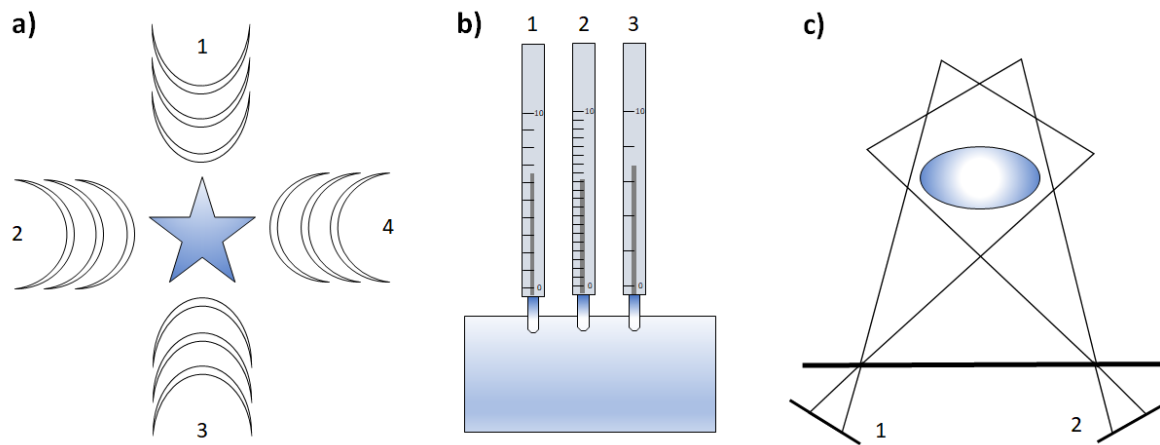
Figure 4.1 – Schematic examples of the a) complementary data fusion using radar; b) competitive data fusing using three different thermometers; c) Cooperative data fusion using stereoscopic vision.

Depending on the problem and on the available information, several data fusion techniques/algorithms can be used. The most simple algorithms relies on weighted averages of prior knowledge of the sensor behavior or in individual quantitative inferences/estimates (voting techniques) [Lin Xiao, 2005]. When in presence of imperfect data, one of the most common used algorithms relies on probability theory, based on the Bayesian theorem. This type of algorithms is capable to expresses the source data uncertainty [Rogova, 2004]. Other techniques are the fuzzy logic and the possibility theory-based fusion, that handles the vagueness of data that is beyond the probability field, trying to give it a meaning to the operator perspective [Saffiotti, 1997]. The Dempster–Shafer (DS) algorithm uses evidential belief theory that can represent uncertain as well as ambiguous data closer to our human perception and reasoning processes [Huadong Wu, 2002]. Hybridizations of these approaches are used, aiming for a more comprehensive treatment of data imperfection. When in presence of highly correlated data for dynamic systems, which are described by difference, or differential equations, the Kalman filter (KF) approach is a very popular algorithm used, requiring either independence or prior knowledge of the cross covariance of data to produce consistent results [Faragher, 2012].

Data fusion techniques are a good approach to a real-time sensor system as the case of the developed turbidity and color sensor, where all the multivariate data information gathered can be processed in a smart way to retrieve the best accuracy possible.

4.1.3. Algorithms used and comparison methodology

Several techniques to correlate the obtained multivariate data were considered and used, but only the more efficient ones applied will be here presented. The methods used can be divided in 4 groups based on its nature: the basic regression analysis, cluster based classification techniques, artificial neural networks and finally a data fusion algorithm.

The first group is based on simple regressions of mean values from known data and will take advantage of the IR light source transmission insensitivity of color to calculate first, the turbidity, and posteriorly, the color.

The second group have two different methods. The first is the implementation of a linear discriminant analysis (LDA) to previously obtained supervised data from the sensor that will merge the data into groups, in a lower dimensional space. These groups can be identified by color first, independently from its turbidity values which is calculated afterwards. The second method will take the advantage of the Gaussian distribution, obtained from the measurements, to create clusters that will be the elements in the regression models that creates a cluster-continuum for each color with the variation of turbidity.

The third group will use artificial neural networks (ANN) that will create a non-linear model that could discriminate the color of the liquid and the turbidity value from supervised data to train.

Finally, the last group will use data fusion techniques, namely a recursive estimator based on a weighted arithmetic mean obtained from the accuracy performance analysis of each dimension. This recursive estimator will also have fuzzy logic rules based on the sensor state of measurement during the recursive cycles, where the minimization of a loss function, related to turbidity, is sought. Knowing the turbidity value to be expected, the color assessment is then performed. Each of these methods of analysis will be discussed in more detail in this chapter.

4.2. Regression analysis

The regression analysis method is the simpler and most direct one. Its prediction is based on the regressions performed to non-color turbid data and therefore does not need the supervised data from colored samples. The regressions discussed in chapter 3.2.3 for the multivariate data variation with turbidity, this is, the 2nd order polynomial for the scattering data and the Beer-Lambert law for the transmitted data, are applied to the non-color solutions to be trained. With the trained regressions, it is possible to assess the color of a new measurement by finding the turbidity

estimative of the sample through the IR light source information, which is color unresponsive (Figure 4.2). This can be done by using both scattered regression, the transmission regression or a combination of both. Because of the lower error presented for higher turbidities from the transmission regression, this is the one chosen to infer the turbidity obtained. Knowing the turbidity value, the color determination is performed by using the regression analysis from the transmitted channel of the other light sources as specified in chapter 3.2.4. It is obtained, in this away, three RGB values of absorbance that characterize the color of that unknown solution. If the refractive index is a changing factor in the measurement, a compensation to all the obtained regressions based in the linear correlations observed in chapter 3.2.5 needs to be made before the application of the classification routine.
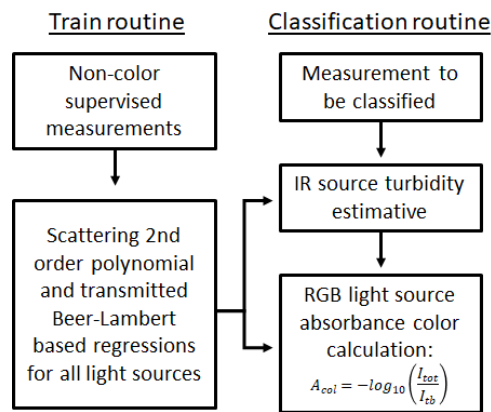
Figure 4.2 – Train and classification routine for the regression analysis. The absorbance values of the RGB light sources are calculated based on the light intensity observed (I_{tot}) and the expected one for the turbid estimation (I_{tb}).

4.3. Cluster classification techniques

4.3.1. Linear Discriminant Analysis

The first method used and based on cluster classification will begin with a linear discriminant analysis (LDA). This technique is most commonly used as dimensionality reduction technique in the pre-processing step for pattern-classification and machine learning applications. The goal of this technique is to project a dataset onto a lower-dimensional space with good class-separability, this is, maintaining the class-discriminatory information avoiding overfitting which can reduce computational costs. In the case of this work, 8-dimensional data from supervised measurements will be reduced to 3 dimensions. LDA computes the directions, this is, the linear discriminants that will represent the axes that maximize the separation between multiple classes.

Considering that one wants to project the d-dimensional data set to a k-dimensional subspace where ($k < d$), the LDA algorithm can be summarized in 5 steps:

1. Compute the d-dimensional mean vectors (m_i) for the different classes from the dataset.

2. Compute the scatter matrices (in-between-class and within-class scatter matrix).

3. Compute the eigenvectors ($e_1, e_2, ..., e_d$) and corresponding eigenvalues ($\lambda_1, \lambda_2, ..., \lambda_d$) for the scatter matrices.

4. Sort the eigenvectors by decreasing eigenvalues and choose k eigenvectors with the largest eigenvalues to form a $d \times k$ dimensional matrix W (where every column represents an eigenvector).

5. Use this $d \times k$ eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the matrix multiplication: $Y = X \times W$ (where X is a $n \times d$-dimensional matrix representing the n samples, and Y are the transformed $n \times k$-dimensional samples in the new subspace).

The within-class scatter matrix (S_W) is calculated by the following equation:

$$S_W = \sum_{i=1}^{c} S_i \qquad \textit{(Eq. 4-1)}$$

Where the scatter matrix (S_i) for every class C is:

$$S_i = \sum_{x \in D_i}^{n} (x - m_i)(x - m_i)^T \qquad \textit{(Eq. 4-2)}$$

With the mean vector (m_i) as:

$$m_i = \frac{1}{n_i} \sum_{x \in D_i}^{n} x_k \qquad \textit{(Eq. 4-3)}$$

The between-class scatter matrix (S_B) is calculated by,

$$S_B = \sum_{i=1}^{c} N_i (m_i - m)(m_i - m)^T \qquad \textit{(Eq. 4-4)}$$

Where m is the overall mean and N_i the sample size of the respective class. To compute the eigenvalues and eigenvectors, one need to solve the generalized eigenvalue problem for the matrix $A_v = S_W^{-1} S_B$. Both eigenvectors and eigenvalues provide information about the distortion of the

linear transformation: the eigenvectors (v) are related to the direction of this distortion, and the eigenvalues (λ_v) are the scaling factor for the eigenvectors that describes the magnitude of the distortion, where,

$$A_v v = \lambda_v v$$

(Eq. 4-5)

The k eigenvectors with the largest eigenvalue are chosen for the dimensional space reduction. In the case of this work, for a 3-dimensional space we have $k = 3$. The supervised samples are then transformed to the new space where the same colored solutions with different turbidities will appear as clusters (Figure 4.3).
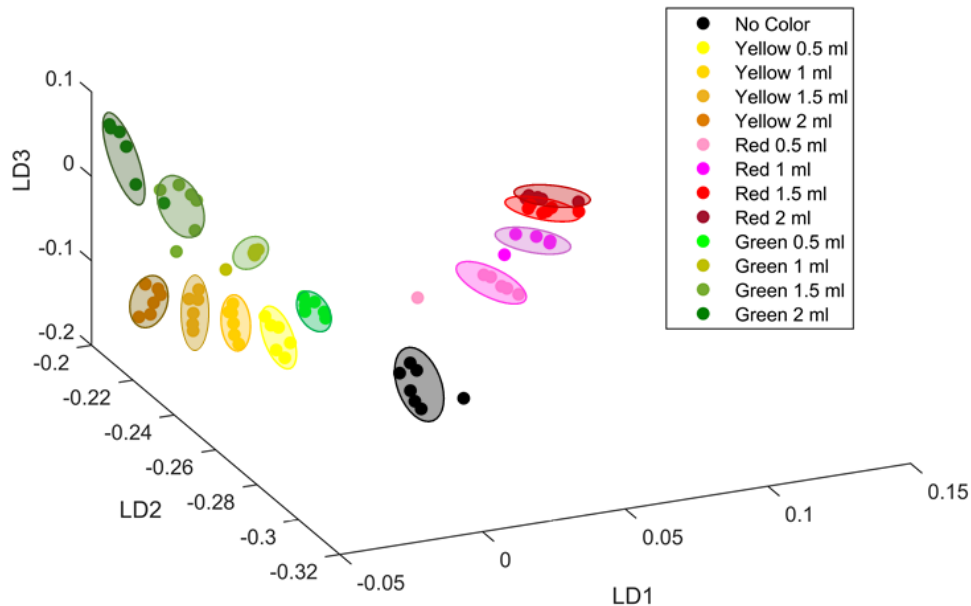
Figure 4.3 – Example of clustering formation in the new LDA 3-dimensional space of colored solutions containing different turbidity values.

In this new space, both mean and covariance values of each color solution class forming a cluster is calculated considering a Gaussian distribution, which finalizes the training routine of this method (Figure 4.4). Both eigenvector information as well as the mean and covariance matrix values will be used in the classification routine.

To estimate the color of a new sample, the first step is to transform its multivariate data to the subspace obtained from the training routine. With the unknown data transformed to this subspace, its mean value will be compared to each one of the clusters representing a color and the weight of each cluster representing this sample is determined from the Gaussians distributions trained. The more weigh a sub-cluster have in relation to the measurement point, the higher the probability of

that point to be of that sub-cluster color. The total color intensity of the sample will be represented by the weight percentage obtained from each cluster. Once the color is estimated, the turbidity is also easily determined by the same equivalent procedure presented for color in chapter 3.2.4. In the case of the refractive index being a parameter to change in the solution, the color calculation will not be affected since effectively the change in the refractive index is interpreted by a change in turbidity. To calculate the correct turbidity value, a compensation is needed based on the linear correlations observed in chapter 3.2.5 before applying the inverse procedure from chapter 3.2.4.
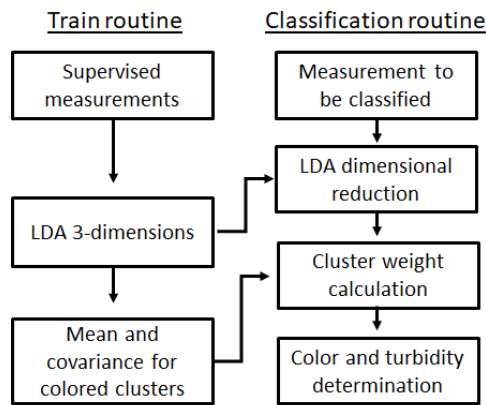
Figure 4.4 – Train and classification routine for the LDA based clustering.

4.3.2. Expectation Maximization Gaussian Mixture

This second method based on cluster classification techniques will take advantage of the intrinsic Gaussian distribution that a measurement, taken over time with several points, will have around a mean value. For this, an Expectation Maximization semi-supervised Gaussian Mixture (EMGM) algorithm is used [Martinez-Uso, 2010; Duarte, 2015, 2019]. The objective of the Expectation-Maximization (EM) algorithm is to find the maximum likelihood solutions for models having latent variables, by an iterative procedure which can be shown to converge to a (local) maximum of the marginal *a posteriori* probability function. For a comprehensive mathematical formulation let one consider the set of all observed data to be defined by X and the set of all latent variables by Z. Let one also consider that the set of all model parameters is denoted by θ. It is now possible to define the log likelihood joint density function by,

$$\ln p(X \mid \theta) = \ln\left\{\sum_{Z} p(X, Z \mid \theta)\right\}. \qquad \text{(Eq. 4-6)}$$

The latent variables values in Z are only given by the posterior distribution $p(Z|X,\theta)$, whose expected values will be useful since the complete-data log likelihood cannot be used. The calculation of this posterior distribution using the actual set of model parameters denoted by θ^{old} will then be used to find the expectation of the complete-data likelihood to a general set of model parameter θ. This expectation generally denoted by Q is given by (Eq. 4-7),

$$Q(\theta, \theta^{old}) = \sum_Z p(Z \mid X, \theta^{old}) \ln p(X, Z \mid \theta). \qquad \textit{(Eq. 4-7)}$$

The goal is to find the new parameters θ^{new} that maximize the expected log likelihood Q of all observations in respect to the latent variables in Z, this is,

$$\theta^{new} = \arg \max_\theta Q(\theta, \theta^{old}). \qquad \textit{(Eq. 4-8)}$$

In summary, the EM algorithm can be described in the following steps:

1. Chose the initial settings for the parameters θ^{old}.
2. Expectation step - Evaluate the posteriors $p(Z|X, \theta^{old})$.
3. Maximization step - Evaluate the new parameters in θ^{old} given by (Eq. 4-8) using the posteriors calculated in step 2.
4. Check for convergence of the log likelihood between two consecutive steps. If convergence is not achieved, set $\theta^{old} \leftarrow \theta^{new}$ and return to step 2.

The aim of a semi-supervised EM algorithm is to improve the final classification or clustering result by providing previously labeled information to the probabilistic model. With this approach a labelled element is not guaranteed to be part of the distribution from which it is labeled but it will have a higher probability to be part of that distribution when the EM algorithm converges. Let us consider the D-dimensional vector x with N elements. Some of these elements are already labeled so we can divide this set of elements in two separate sets. For the labeled elements, we will denote the set by S and for the unlabeled by U where $N = |S| + |U|$.

To identify which set an element belongs, a superscript s or u is used, this is $S = \{x_1^s, \ldots, x_{|S|}^s\}$ and $U = \{x_1^u, \ldots, x_{|U|}^u\}$. By knowing at which k distribution model a variable belongs among the total number of K distributions, it is required that the probabilistic model have a constraint associated to it to ensure that this information is useful. This constraint will be applied to the posterior distribution assuming $p(Z_k|x_n, \theta) = \delta_{nk}$, which returns 1 if x_n belongs to the distribution

k and 0 otherwise. The log-likelihood of all observations of a general mixture of distributions will then have the contribution of the supervised and unsupervised elements,

$$Q(\theta,\theta^{old}) = \sum_{Z^S} \delta(X^S,Z) \ln p(X^S,Z \mid \theta) + \sum_{Z^U} p(Z \mid X^U,\theta^{old}) \ln p(X^U,Z \mid \theta) \qquad \textit{(Eq. 4-9)}$$

For the application of the semi-supervised EM algorithm to the Gaussian mixture model with a D-dimensional vector x with N elements, let us start by considering the multivariate Gaussian distribution in the form of

$$g(x \mid \mu,\Sigma) = (2\pi)^{-\frac{D}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right). \qquad \textit{(Eq. 4-10)}$$

where μ is a D-dimensional mean vector, Σ is a DxD covariance matrix and $|\Sigma|$ denotes the determinant of Σ. The Gaussian mixture distribution can now be written as a linear superposition of K Gaussians in the form

$$p(x) = \sum_{k=1}^{K} \pi_k g(x \mid \mu_k,\Sigma_k). \qquad \textit{(Eq. 4-11)}$$

Here, π_k is the mixing coefficient of the Gaussian number k. Note that $0 \leq \pi_k \leq 1$ and

$$\sum_{k=1}^{K} \pi_k = 1. \qquad \textit{(Eq. 4-12)}$$

As seen in (Eq. 4-9), for the semi-supervised EM algorithm, the expectation of the complete-data log likelihood function for maximization is needed and the one for Gaussian mixture is given by

$$Q = \sum_{n=1}^{S(k)} \sum_{k=1}^{K} \delta_{nk} \ln\left\{ \pi_k g(x_n^s \mid \mu_k,\Sigma_k) \right\} + \sum_{n=1}^{U} \sum_{k=1}^{K} \tau_{nk} \ln\left\{ \pi_k g(x_n^u \mid \mu_k,\Sigma_k) \right\}. \qquad \textit{(Eq. 4-13)}$$

Here, $S(k)$ stands for the subset of supervised instances for a particular distribution k and $\tau_{nk} = E[z_k|x_n,\theta^{old}]$ the expectation of latent variable z_k associated to x_n or the posterior probability of x_n belonging to the k_{th} component of the mixture. Considering $\theta^{old} = \{\pi_k^{old},\mu_k^{old},\Sigma_k^{old}\}$, the posterior probability is calculated by:

$$\tau_{nk} = \frac{\pi_k^{old} g(x_n \mid \mu_k^{old},\Sigma_k^{old})}{\sum_{j=1}^{K} \pi_j^{old} g(x_n \mid \mu_k^{old},\Sigma_k^{old})}. \qquad \textit{(Eq. 4-14)}$$

To find the new parameters, μ^{new} and Σ^{new}, that maximize the expected log likelihood Q, its differentiation with respect of each one of the parameters is needed, this is:

$$\frac{\partial Q}{\partial \mu_k} = 0; \qquad \frac{\partial Q}{\partial \Sigma_k} = 0 . \qquad \textit{(Eq. 4-15)}$$

which lead to the calculation of the new parameters as

$$\mu_k = \frac{\sum_{n=1}^{S(k)} x_n^s + \sum_{n=1}^{U} \tau_{nk}\, x_n^u}{|\, S(k)\,| + \sum_{n=1}^{U} \tau_{nk}} . \qquad \textit{(Eq. 4-16)}$$

$$\Sigma_k = \frac{\sum_{n=1}^{S(k)} M^s + \sum_{n=1}^{U} \tau_{nk}\, M^u}{|\, S(k)\,| + \sum_{n=1}^{U} \tau_{nk}} . \qquad \textit{(Eq. 4-17)}$$

In (Eq. 4-17), to simplify, it was established $M = (x_n - \mu_k)(x_n - \mu_k)^T$. To perform maximization of (Eq. 4-13) with respect to the mixing coefficient π^{new}, one should take in consideration the constraint in (Eq. 4-12). This is achieved by using a Lagrange multiplier and maximize the following,

$$\sum_{n=1}^{S(k)}\sum_{k=1}^{K} \delta_{nk}\, ln\left\{\pi_k\, g(\, x_n^s \,/\, \mu_k, \Sigma_k\,)\right\} + \sum_{n=1}^{U}\sum_{k=1}^{K} \tau_{nk}\, ln\left\{\pi_k\, g(\, x_n^u \,/\, \mu_k, \Sigma_k\,)\right\} + \lambda\left(\sum_{k=1}^{K} \pi_k - 1\right). \qquad \textit{(Eq. 4-18)}$$

which, to conclude with all necessary parameters needed to apply the semi-supervised EM algorithm, we then obtain

$$\pi_k = \frac{|\, S(k)\,| + \sum_{n=1}^{U} \tau_{nk}}{N} . \qquad \textit{(Eq. 4-19)}$$

The EMGM algorithm will pick all the supervised data to be trained and will find the maximum likelihood solution to create a prechosen number of clusters, considering its Gaussian distribution (Figure 4.5). For each cluster a multidimensional mean value, covariance matrix and a mixture coefficient will be calculated as demonstrated above. In this method, a cluster will be calculated for each measurement in the training set. This prevents that an outlier could be considered to its original cluster measurement since it will be integrated to a closer cluster, diminishing the overall error. Having each measurement associated to a cluster, the same

regression models used in chapter 4.2 are applied to the mean values of the clusters, but in this case, it is applied to all of the used colored solutions as well. This process can be seen as a creation of a cluster-continuum for each colored solution that varies with turbidity and that has not a single mean value but instead a "continuous" mean value that is given by the regressions. In this cluster-continuum, a global variance will be considered as the mean variance values of the prior clusters that constitute the cluster-continuum and the mixture coefficient will be the sum of their individual mixture coefficients. With this approach, independently if it was already trained or not, any turbidity value can be associated to a sub-cluster from the cluster-continuum. Finalized the training routine with the determination of each cluster-continuum for each color, the classification routine initializes by estimating the turbidity value of a measurement using the IR regressions just like in the regression analysis method presented in chapter 4.2. Alternatively, an approximate value of turbidity, using only the transmission and scattering dimensional information of the RGB light sources, can be calculated based in a recursive cycle that minimizes and compares the error of the expected turbidity for each individual dimension which makes this approach IR independent but increases the overall error. This turbidity estimation will be very important because it will be the value that will be used to calculate, for each color cluster-continuum, the comparison multidimensional point that will be used later. In other words, the estimated turbidity value will be used to calculate the expected value for each dimension using the regressions obtained by the training routine. Doing this for each color, we will have in the end the measurement point that we want to determinate and, for that expected turbidity, the expected sub-clusters for each color that encompasses the multidimensional mean point, the covariance matrix and the mixture coefficient. It is now possible to determinate the weight or how close is the measurement point to the sub clusters of each color. The more weigh a sub-cluster have in relation to the measurement point, the higher the probability of that point to be of that sub-cluster color. If the calculated weight is mostly distributed between two sub-clusters, then the probably color will be between those two sub-cluster colors. The color is therefore calculated based on a percentage of each sub-cluster color proportional to their weight.

For a changing refractive index parameter in a solution to measure, calibrations using color with a different refractive index value need to be made for a compensation based on the linear correlations observed in chapter 3.2.5 and the position of the clusters continuum recalculated based on the interpolated compensation. This is obviously a more complex process and a disadvantage to this algorithm.
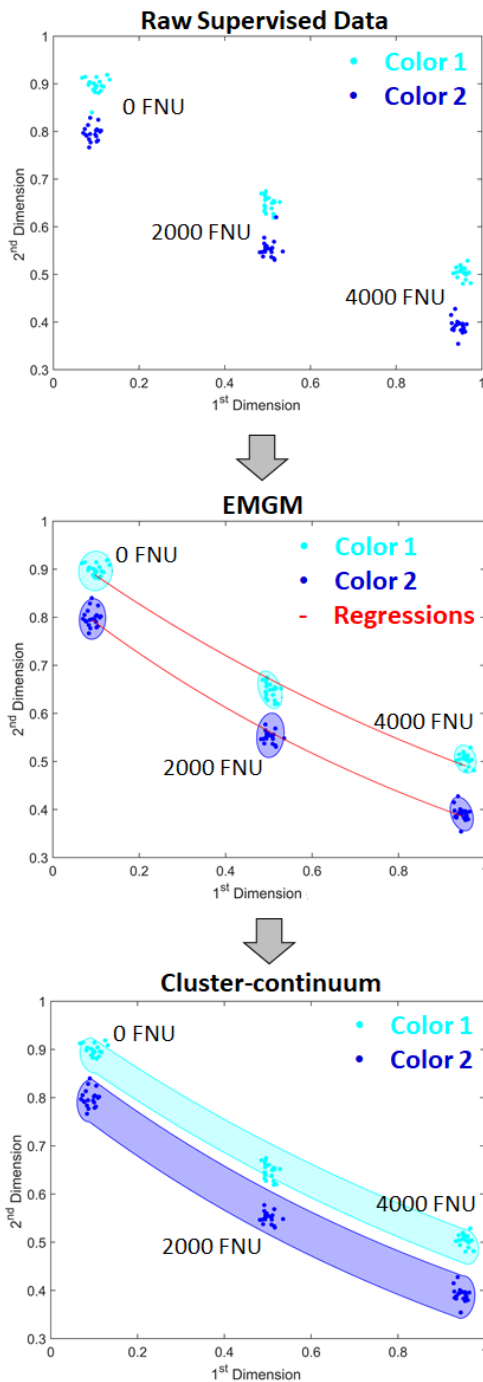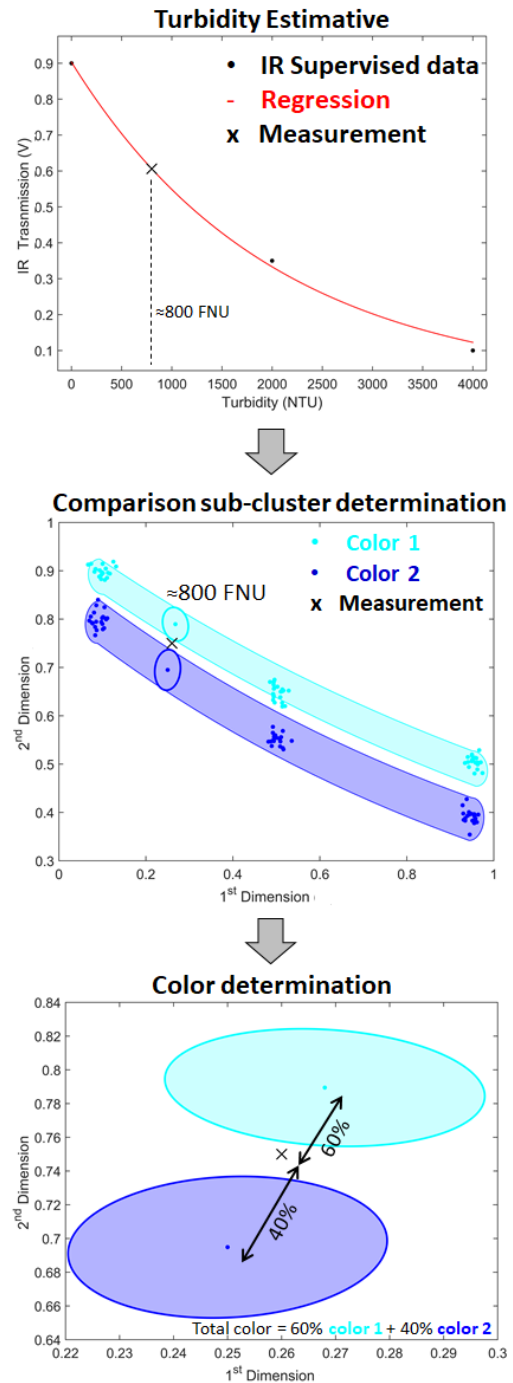
Figure 4.5 – Expectation maximization supervised Gaussian mixture algorithm scheme. Here an example using only the representation of 2 dimensions and 2 colors for data visualization and training. In the classification routine, a random measurement is being classified. In the final step of the routine, the measurement point to be classified have the weight of the 2 sub clusters which makes its color a combination of the other 2.

4.4.   Artificial Neural Networks

One of the most well-established and most used machine learning techniques found in literature is the artificial neural networks (ANN). A more detailed description of its application can be observed in [da Silva, 2017]. In summary, ANN are computational models based on the biological operation and connections of the neurons of living beings. In the same way, a neuron processes and transmits information to other neurons by synapses, the ANN will also have processing units that are connected between each other. Each processing unit will have inputs, represented by the set $\{x_1, x_2, x_3, ..., x_n,\}$. Each input will have an associated weight $\{w_1, w_2, w_3, ..., w_n,\}$ which represent the "synapses" of that neuron. These synapse weights will be applied to the inputs, the external information arriving to the neuron, for an assignment which can have different importance from each input. A bias variable (β) is also added to specify the proper threshold necessary to generate a trigger value toward the neuron output. This weighting input information denoted by u will then be processed with a simple nonlinear activation functions (g) producing, this way, a response as output signal (y) (Figure 4.6). Linear, sigmoid and hyperbolic tangent are example of activation functions typically used. When an output is produced, it will be transmitted to the other neurons by other "synapses". All the "synapses" will have also weights associated to them that are multiplied by the output and deliver to the next neuron.
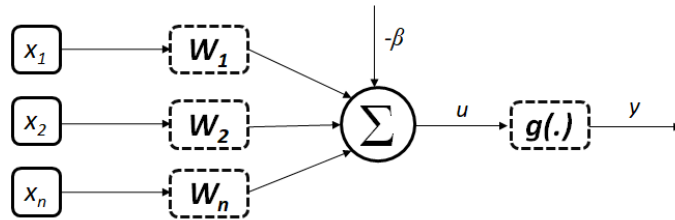
Figure 4.6 – Schematic representation of a single neuron as the elementary processing unit of an artificial neural network.

The two following expressions synthesize the result produced by the artificial neuron:

$$u = \sum_{i=1}^{n} w_i . x_i - \beta \qquad \text{(Eq. 4-20)}$$

$$y = g(u) \qquad \text{(Eq. 4-21)}$$

In the case of this work, the two-activation function that will be used are the linear and sigmoid function described by (Eq. 4-22) and (Eq. 4-23) respectively.

$$y = u \qquad \text{(Eq. 4-22)}$$

$$y = \frac{1}{1+e^{-u}} \qquad \text{(Eq. 4-23)}$$

The general structure of an ANN is made of three parts or layers as can be seen in Figure 4.7. The input layer is responsible for receiving the information to be calculated for a final decision result. In the case of our sensor data, this layer will receive the 8-dimensional measurements. The hidden layers are composed of the neurons responsible to extract the patterns associated with the processes being analyzed. If it is a complex problem to be solve, more hidden layers will be needed. On the other hand, using a large number of hidden layers to solve simple problems could lead to overfitting. In our case, only one hidden layer composed of 8 neurons will be used as it showed to have lower estimation error overall without overfitting. The neurons of this layer will have sigmoids as processing functions. The output layer is also composed of neurons and are responsible for presenting the final solution of the network. The neurons of the output layer used in this work are four and have linear functions since the output will be the three color RGB transmittance and the turbidity values. This method is also infrared independent since it is not mandatory to have its information as an input, providing that the training routine does not use its information also.
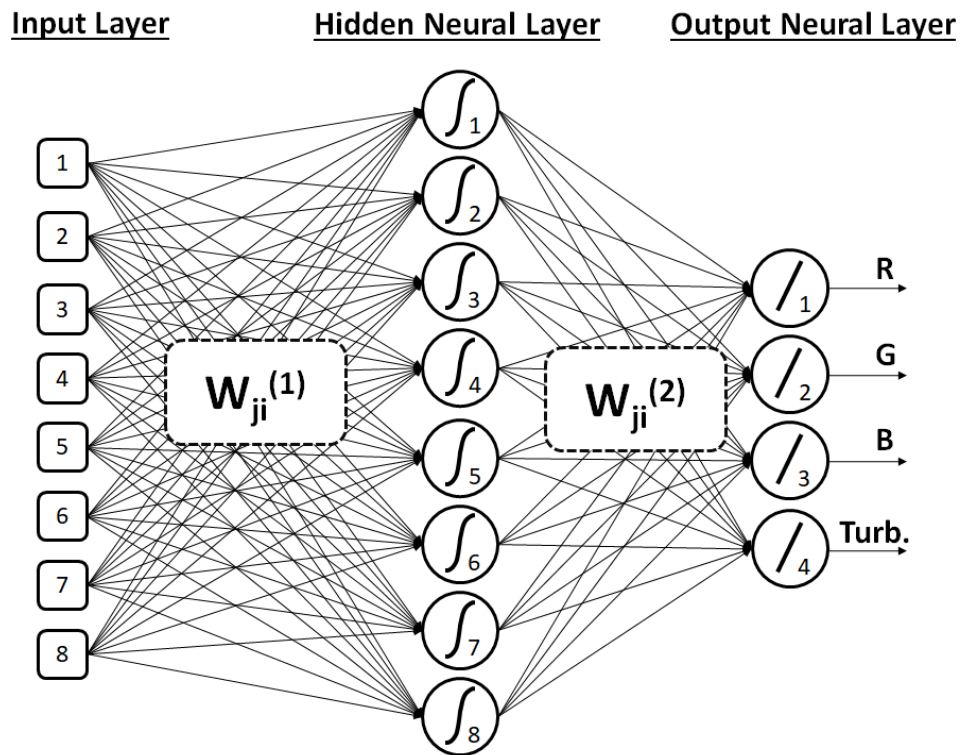
Figure 4.7 – General structure of the artificial neural network used in this work when IR light source information is also used. For the case of not using IR information, the input layer will only have 6 inputs. Between each layer, a group of "synapses" denoted by a subscript number to W is presented, which represent the layer number. The final output will be the RGB color and turbidity values to be determined.

The ANN needs to be trained with the supervised data to have the best final possible results. The most common training process and the one here applied consists in two distinct phases, the

forward propagation and the back propagation. In the forward propagation, we simply apply the set of weights of our network to the input data and calculate the final output. For the first forward propagation, the set of weights are selected randomly. The obtained output values will then be used in the back propagation and a measurement of the margin of error (δ_j) is made in respect to the expected supervised values. The weights, beginning from the last layer backwards, will then be adjusted accordingly, to decrease the final ANN error as,

$$w_{ji} = w_{ji}^{old} + \eta \delta_j y_j \qquad \text{(Eq. 4-24)}$$

Here, η is the rate constant that will control the velocity of adjustment of the weights. Its value was chosen to be 0.5 in this work. The margin of error (δ_j) is calculated based on the output values, this is,

$$\delta_j = y_j(1 - y_j)(D_j - y_j) \qquad \text{(Eq. 4-25)}$$

where D_j is the desired output value. This desired value or values will be directly associated with the expected measurement of the supervised data. In layers prior to the last, this desired quantity will be based on the newly calculated margin of errors and weights, this is,

$$\delta_j = y_j(1 - y_j) \sum_{i=1}^{n} \delta_i w_{ji} \qquad \text{(Eq. 4-26)}$$

When all the new synapse weights are calculated, the algorithm will then again do the forward propagation and the back propagation iteratively. The condition to stop is reached when the total error (E_{ANN}) between two successive epochs is sufficiently small with a defined criterion (ε), this is,

$$\left| E_{ANN}^{current} - E_{ANN}^{previous} \right| \leq \varepsilon \qquad \text{(Eq. 4-27)}$$

After the training is finalized, the classification routine is performed to an unknown sample by simply using the information of the weights that were calculated from the training. Using the weights in a forward propagation will obtain the estimated values of color and turbidity of the unknown sample. The great advantage of the ANN is its easiness to incorporate new parameters in the input layer that will affect the output values. This is the only way to incorporate the refractive index in the algorithm to compensate its influence in the final values of the sensor. Therefore, when performing a calibration for an application that have a changing refractive index, training samples with this variation present would be needed.

4.5.   Data fusion

Data fusion based techniques will be the last method assessed with the objective to analyze and acquire the best results from the multivariate data obtained from the sensor. Basically, two configurations of fusing data will be present and used with the sensor: the complementary and the competitive configurations as discussed in chapter 4.1.2. The complementary is used because not all the light from the light sources will acquire the same properties. As an example, the infrared light source measurement will not detect color, and depending on the color of the solution, other light sources measurements can also be unaffected by it. The same cannot be said from the competitive configuration that will be very present when turbidity solutions are used. Turbidity will affect the measurements from all light sources and though, because each dimensional is not identical, the best characteristics of each can be used to have the best possible result. To do this, one need to know well the behavior of each dimension when in different measurement scenarios, namely, in our case, the variation of the measurement signal with the increase of turbidity and its accuracy. This study was performed with all the used sensors in chapter 3.2.3 and this information will be essential to the data fusion algorithm where a weighed arithmetic mean technique is used, together with fuzzy logic rules.

The data fusion algorithm applied to the turbidity and color sensor data will use both the regressions and accuracy measurements, performed with the variation of turbidity, as the training routine. After this training phase, it will use a recursive method that can be described by the following steps (Figure 4.8):

1. Input data to be classified from the measurement reading.
2. Initial color estimation, that can be 3-random value vector representing the RGB color set.
3. Calculation of turbidity values, taking in consideration the color set calculated in 2, for each of the 8-dimensional data, in an equivalent procedure presented in chapter 3.2.4 for color. It is obtained the turbidity vector τ.
4. Weighed arithmetic mean and fuzzy logic rules to calculate the expected turbidity T_{exp}.
5. Recalculation of the expected color from the expected turbidity obtained in 4.
6. If convergence of values is met, stop the recursive process. If not, return to step 3.
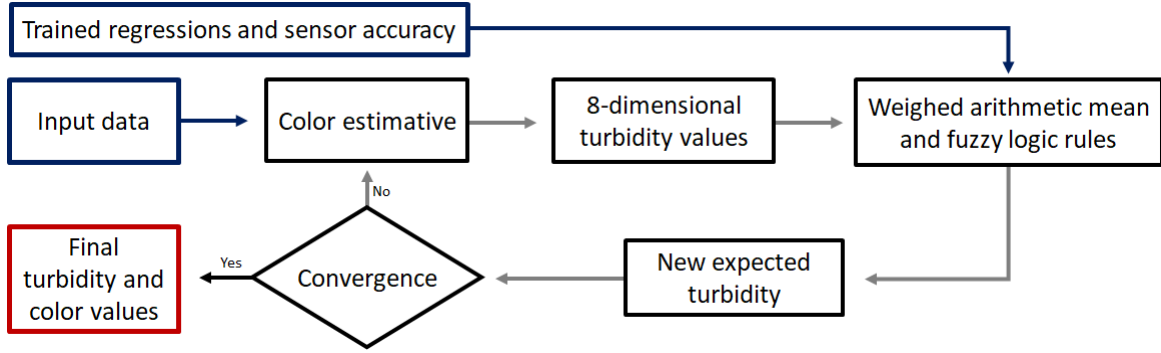
Figure 4.8 – Data fusion recursive algorithm scheme.

The weights associated to the different dimensional data are calculated based on the importance designated to a certain dimension in a direct comparison to other dimensional measurement. This importance is directly obtained from the hypothesis that a certain dimension present lower measurement error than a second dimension of the comparison pair. For example, if the scattering dimension presents lower measurement error than the transmission dimension, for a certain turbidity value, then it will have higher importance and, in consequence, higher weight to the final value of measurement. Considering γ_i the measurement error for a certain turbidity value (τ_i) obtained for dimension i and γ_j the measurement error of dimension j for the same τ_i turbidity value, the calculation of the weighed matrix M, with each element as M_{ij} where $i, j \in D$, being D the number of the measurement dimensional space, can be specified as,

$$M_{ij} = \frac{1-\Gamma_{ij}}{\sum_i \sum_j (1-\Gamma_{ij})}$$

(Eq. 4-28)

where,

$$\Gamma_{ij} = \frac{\gamma_i}{\gamma_i + \gamma_j}$$

(Eq. 4-29)

In the case of the used turbidity and color sensors, a normalized 8 x 8 dimensional M matrix is obtained. Before this weight is applied to the turbidity vector τ, fuzzy logic rules are used to filter some of the weights considering empirical knowledge of the sensor system, mainly to maximize redundancy and fault-proof, but also to take in consideration the increase of uncertainty in the presence of color. These are the type of conditional rules based on the IF, THEN sequence. The first of these rules is the "color rule" where lower weight is given to a certain dimensionality where the presence of color absorption is verified. Here, it is important the identification of each dimension with its correspondent method of measurement and light source to better understand the rules, as seen in Table 4.1.

Table 4.1 – Identification of measurement principle and light sources for each dimension of the data.

Dimension (D)	1	2	3	4	5	6	7	8
Measurement principle	Scattering Blue	Scattering Green	Scattering Red	Scattering IR	Transm. Blue	Transm. Green	Transm. Red	Transm. IR

When in presence of color, a higher uncertainty in the determination of turbidity is expected based on the propagation of uncertainty effect. To lower this uncertainty, lower weight in the weigh matrix M is given to the M_{i*} rows where the i dimension is affected by that specific color, this is,

$$M_{i*}^{new} = M_{i*} \times T_c \qquad \text{(Eq. 4-30)}$$

where T_c is the transmittance value of the respective color in consideration. If the transmittance value is equal to one, this is, there is no effect of color, the M_{i*} row does not change its weight, while if it tends to zero, the weight of this dimension will become also negligible. As an example, if only the green light is affected by the solution to measure, both dimensions using green light (2 and 6) will have its weights lowered, which will increase the preponderance of the other unaffected dimension in the calculation of turbidity.

Other applied rule is a redundancy based mechanism to prevent a non-measurement caused by the obstruction of all the transmission or all the scattering dimensions. That can easily happen if some dirt couples to the tip of one of the fibers related to the two measurement principles. To work around this issue, a rule was created to check if the expected turbidities (τ) for all scattering dimensions, present lower values in relation to the transmission channel and vice versa within its error margin ($\pm\Delta\tau$). In addition, it is guaranteed that the expected turbidity valued for scattering channels needs to be lower than 200 FNU that is the defined minimum safety threshold that the sensor can measure with confidence. When obstruction is detected, all obstructed channels are given a weight of zero, which eliminates its contribution to the new turbidity value calculation. In the case of scattering obstruction, the mathematical formulation can be expressed as,

$$IF \quad \tau_{[1,2,3,4]} + \Delta\tau_{[1,2,3,4]} < \tau_{[5,6,7,8]} - \Delta\tau_{[5,6,7,8]} \quad AND \quad \tau_{[1,2,3,4]} < 200$$
$$THEN \quad M_{i*}^{new} = 0, \quad i \in [1,2,3,4] \qquad \text{(Eq. 4-31)}$$

For the transmission, one has,

$$IF \quad \tau_{[1,2,3,4]} - \Delta\tau_{[1,2,3,4]} > \tau_{[5,6,7,8]} + \Delta\tau_{[5,6,7,8]} \quad AND \quad \tau_{[1,2,3,4]} < 200$$
$$THEN \quad M_{i*}^{new} = 0, \quad i \in [5,6,7,8] \qquad \text{(Eq. 4-32)}$$

To become a more robust and fault-proof measurement system to outlier values that could appear for various reasons, like sudden electrical peaks or momentary dirt blockade to a certain light source during measurement, some rules were created to prevent "absurd" measurements. This is the case where the calculated turbidity for a specific scattering i channel is much higher than the expected from the transmitted $i + 4$ channel of the same color,

$$IF \quad \tau_i - \Delta\tau_i > \tau_{i+4} + \Delta\tau_{i+4}$$
$$THEN \quad M_{i*}^{new} = 0, \quad i \in [1,2,3,4]$$

(Eq. 4-33)

Notice that the inverse, a specific transmitted $i + 4$ channel turbidity being much higher than the expected from the scattered i, is not considered because the influence of color can produce this type of error by uncertainty propagation while the rule applied above is something not expected to occur, since in presence of color, light is absorbed and only turbidity can increase light detection.

Other type of "absurd" errors are detected when the turbidity expected value and its error range is much lower than zero. While negative values of turbidity and their margin error are expected for low turbid solutions, especially when the recursive algorithm is not converged yet, these negative values will be near the zero FNMU turbidity value. If this not happen and the value present a turbidity lower than -200 NTU, then it is certain to affirm that this value, for this i dimension, cannot be considered to the final turbidity value and so the respective dimension must be neglected. So, for the scattering and transmission channels one obtains,

$$IF \quad \tau_i - \Delta\tau_i < -200$$
$$THEN \quad M_{i*}^{new} = 0, \quad i \in D$$

(Eq. 4-34)

By calculating the final M_{ij} coefficients of the weighted matrix, it is necessary to perform a normalization to represent the final weighted matrix M that will be used to estimate the new turbidity value:

$$M = \frac{M_{ij}}{\sum_i \sum_j M_{ij}}$$

(Eq. 4-35)

To calculate the next iteration of new turbidity values, a matrix multiplication is performed to the initial turbidity vector τ which will create the new weighed turbidity vector $T = M\tau$ that is a representation of the contribution of each dimension to the total turbidity value (T_{exp}), which is calculated by the sum of each element of T:

$$T_{exp} = \sum_{k=1}^{D} T_k \qquad\qquad (Eq.\ 4\text{-}36)$$

With the new expected turbidity calculated, the expected color is also easily determined (see chapter 3.2.4), and a convergence value is sought by analyzing the new expected turbidity with the old one, stopping the recursive iteration when the difference of the two values is sufficiently small with a defined criterion *(ε):*

$$\left| T_{exp}^{new} - T_{exp}^{old} \right| \le \varepsilon \qquad\qquad (Eq.\ 4\text{-}37)$$

When in presence of refractive index changes, only a compensation similar to the one performed for the regression analysis based in the linear correlations observed in chapter 3.2.5 needs to be made before the application of the classification routine, this is, new compensated regressions will take the role of the non-compensated ones in the classification routine.

4.6.   Algorithms application and comparative results

To test the proposed algorithms, samples of 500 ml with different colors and turbidities were prepared. Red, yellow and green food dyes were used to add color. The volume of dye added to water was 0.5, 1, 1.5 and 2 ml, respectively, for all colors. For turbidity, corn starch was added to the solution with different concentrations. To create combinations of color-turbidity, for each color, a mass of 0, 0.2, 0.45, 1, 2, and 2.8 g of corn starch was added and mixed. The dyes have no absorption in the IR region. The turbidity value of each sample was measured by the commercial turbidimeter from Libelium S.L. [Libelium Comunicaciones Distribuidas S.L, 2019] which is calibrated with formazine standard and has a 5% accuracy. A total of 78 different samples were obtained (Figure 4.9). Table 4.2 summarizes the characteristics of the prepared samples.

Figure 4.9 – Image of some of the colored prepared samples with different turbidities and dye concentration.

Table 4.2 – Resume table of the characteristics of the prepared samples used to train and classify.

	Dye volume (ml)		Turbidity (FNU)				
No color	0	0	164	449	999	2020	3571
Yellow dye	0.5	0	187	461	1042	2095	3542
	1	0	175	464	1016	2158	3556
	1.5	0	182	468	1033	1989	3456
	2	0	198	462	1052	2080	3575
Green dye	0.5	0	187	415	1021	2065	3721
	1	0	174	431	999	1973	3518
	1.5	0	183	424	1044	2053	3535
	2	0	196	438	1010	2089	3530
Red dye	0.5	0	204	474	996	2176	3688
	1	0	222	450	1023	2143	3529
	1.5	0	203	465	1023	2049	3530
	2	0	205	420	980	2006	3534

With all the samples prepared, a measurement to each solution was performed and repeated 3 times. The short LED sensor (version 1) was used to do these measurements. Each measurement lasted for 3 minutes which encompasses a total of 8 individual points in a Gaussian distribution, intrinsic to the sensor measurement error obtained from the particles Brownian motion.

For each classifying method discussed above, an analysis of the classification error of the color and turbidity will be made with different percentages of trained samples using a random sampling cross validation method [Kotsiantis, 2007] in which the samples that will be chosen for training are randomly selected. The samples that were left of training will be the ones that will have its color and turbidity assessed. The process of training and classification will be performed over 100 cycles and calculated, for each cycle: the error ratio of color and turbidity; the duration of training and the duration of the classification. The total error ratio and duration related to each trained percentage will be calculated by the mean values. Two approaches will be made taking advantages of the proposed algorithms dependency of the IR light source information. The first approach will use the IR light source information to all the classifying methods and then an approach, without using this information, will be done to the algorithms that do not have a mandatory need of IR information, this is, the LDA, EMGM, ANN and Data Fusion. These processes were done in two programming languages, Matlab® R2014b and Python™ 2.7 with Numpy library, to assure that there is an overall comparison and comprehension between the measured durations within the different algorithms and by trying to have independency with the programming language (Figure 4.10). Note that measured durations between algorithms and different programming languages could depend on several factors as better code optimizations and/or libraries or optimization to the hardware. A computer with an Intel® Core™ i5-2430M double core processor at 2.40 GHz and 6 GB of RAM was

used having the 64-bit Windows® 10 as the operative system. The scripts were written for serial processing only and therefore only a single core was used.
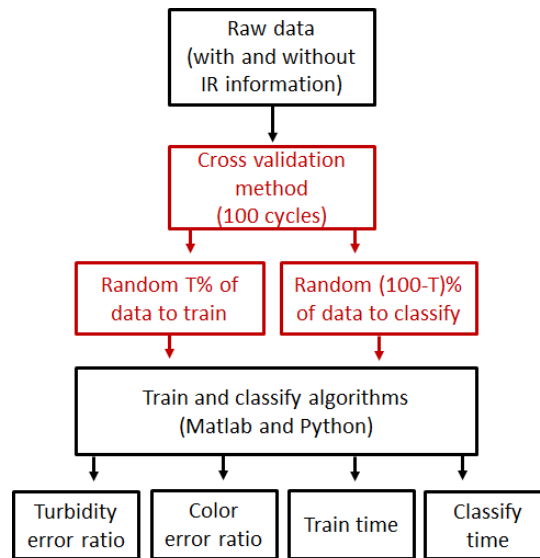
Figure 4.10 – Schematic diagram of the train and classification analysis.

4.6.1. With IR information analysis

4.6.1.1. Cross validation

Although the use of IR light information is not mandatory for the LDA, EMGM, ANN and the Data Fusion algorithms calculation, as stated before, it is essential to the more direct and simple approach of regression models. When assessed with the cross-validation methodology, the regression models' error ratio obtained for the color with trained data of 20% and above, fluctuates practically in the same range values of 0.035 to 0.04. This can be observed in Figure 4.11 represented by the line with dots in both Matlab and Python languages. The error ratio obtained was expected due to the natural intrinsic Gaussian distribution values obtained from turbid media in addition to the own sensor error and the error propagation that, collectively, will have a lower limit value that is achieved with few training values in the regression. Once this minimal error is achieved, it is impossible to go lower with this approach.

Color cross validation
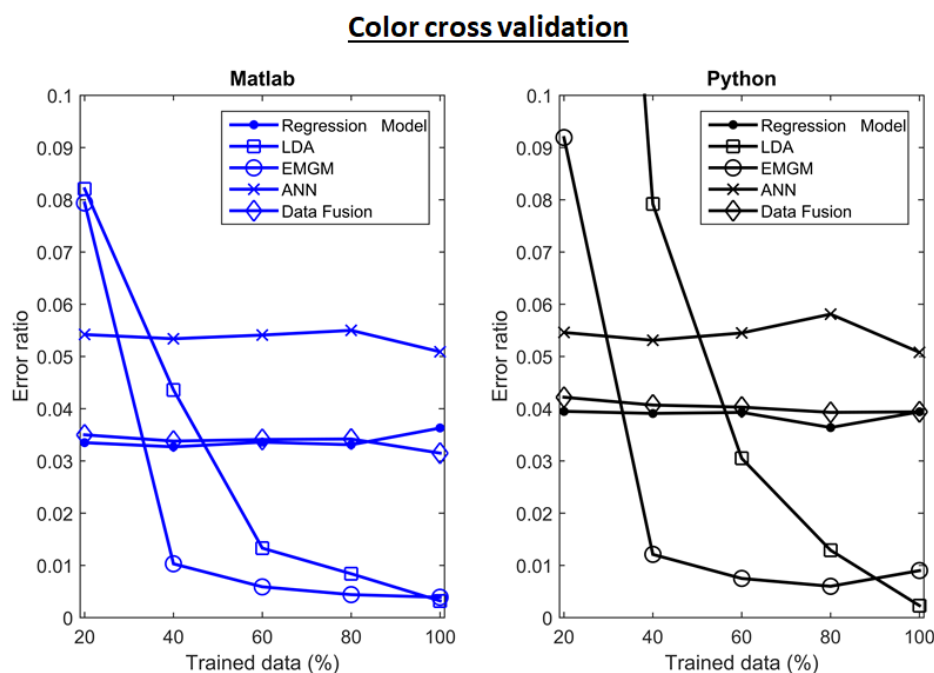
Figure 4.11 – Color cross validation error ratio evolution with the trained data percentage (using IR information) for each classifier methods and for both Matlab and Python languages.

To have a general idea of the performance of each method independently of the trained percentage, a calculation of the overall error ratio was conducted based on the determination of the lower area of the error ratio curves between the 20% to 100% training data, just like an integration. The lower this value is, the higher the overall performance will be considered. For the regression models, the value was 2.68 for Matlab and 3.08 for Python.

An overall lower error ratio value performance was also obtained for the EMGM algorithm, with values of 1.25 for Matlab and 1.52 for Python. The evolution of the error ratio for this algorithm with the trained data percentage can be seen by circles marks in Figure 4.11. The obtained lower values of the overall error ratio, in comparison to the regression models approach, can be attributed to the substantially lower error ratio obtained (roughly an order of magnitude lower) for higher training percentages (40% and above). In contrast, for lower training percentages the error is higher. This proves that this method is excellent for measurements of liquids that have the same cyclic color variation in which good training data set can be attainable and where the traditional method of color determination is outperformed. In contrast if few training samples are available or the liquid to measure does not perform cyclic color variations, then this method cannot be used since its error will have a tendency to infinity when approaching none training values. This same inference is also valid to the results obtained by the LDA method, represented in the figure by square marks. The overall error ratio performance for the LDA was worse than the EMGM with 2.16

for Matlab and 5.50 for Python with the latter having such a high value due to the weight of the 20% error ratio tending to infinity. Note that due to the error intrinsic nature with tendency to infinity for low training data, slightly variations in the random data analyzed can cause high variation in the overall error ratio obtained, as observed in the latter case.

In relation to the ANN approach, represented with crosses in Figure 4.11, it has showed to have also a nearly constant error ratio value, around 0.055 for both Matlab and Python. This reveals that the value could be the lowest error possible to achieve by this approach without overfitting. As expected this higher error values will have influence on the overall performance values, being also the highest from all the methods presented (not considering the LDA python value), with 4.30 and 4.37 for Matlab and Python respectively.

In other hand, the Data Fusion technique, represented by the diamond shape in Figure 4.11, revealed to have an error ratio range very close values to the simple regression method along all the trained percentages (values from 0.035 to 0.04). This evince that this approach is equally effective for color assessment as the regression method. The overall ratio error performance values are, without surprise, close to the regression model, this is, 2.70 for Matlab and 3.22 for Python language.

The same approach of cross validation error ratio assessment was performed for turbidity values and is presented in Figure 4.12.
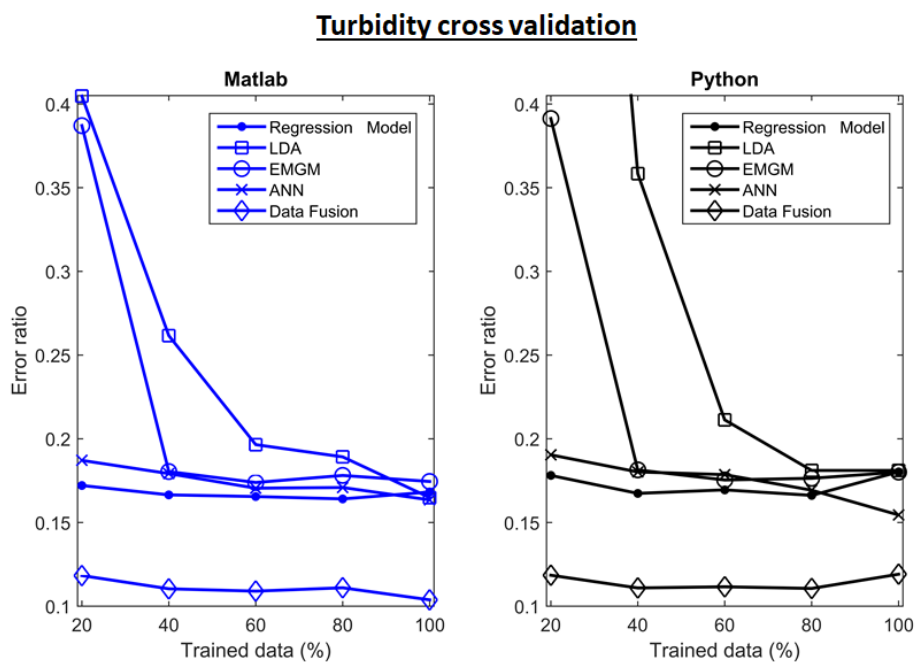
Turbidity cross validation

Figure 4.12 – Turbidity cross validation error ratio evolution with the trained data percentage (using IR information) for each classifier methods and for both Matlab and Python languages.

Again, the results obtained for the simple regression model approach present a nearly constant error ratio value (0.17) along the training data, for both Matlab and Python languages, as can be seen by the line with dots. This 17% high error value is expected for this measuring method since the sensor used to acquire the samples measurement, the short LED sensor - version 1, showed to have equally high measurement error using the transmission IR information, as seen in Figure 3.14, which is used in this method to derive the turbidity value. The overall error ratio performance for regression model is 13.33 and 13.65 for both Matlab and Python respectively.

The EMGM (circle representation) and LDA (square representation) methods present the same expected behavior observed for the color measurement, this is, for higher training data better results will be obtained with lower error ratio. However, in contrast to the color assessment, the error ratio will not be lower than the one obtained from the regression model, but will be very similar, with the same magnitude. For lower training values, the error ratio tends to infinity as observed for the color before. Taking in consideration this information, it is possible to state that the EMGM and LDA approaches will have better performances for color without harming the turbidity values, but again, only when using high training data percentages that can only be possible in cyclic color variations of the liquid of measurement. The EMGM overall error ratio is 16.26 and 16.38 for both Matlab and Python respectively. The obtained LDA algorithm error was 18.64 for Matlab and 26.45 for Python.

The achieved ANN result of error ratio, represented by crosses in Figure 4.12, revealed to be slightly higher than the one seen by the regression model, with values around 0.18 to all the training data range and for both programming languages. Because of the low training data needed to achieve this low error value (at least 20%), which is comparable to the regression models, the ANN approach can be seen as a good candidate for implementation that uses all the dimensional data at once without previous knowledge of the physical behavior of the system. This facilitates the introduction of new and diverse data from other sensors to, globally, obtain also an output result at once, with near results from what is expected. The overall error ratio for the ANN approach was 13.92 and 14.01 for both Matlab and Python, which are, as expected, close to the ones obtained by the regression model.

In last, represented by the diamond shape in Figure 4.12, the Data Fusion algorithm shines in comparison to the other approaches, obtaining a lower error ratio than the regression model of about 0.11, which makes it the best algorithm implemented for turbidity measurement. This is achieved for all the training data range used and it is expected since it is based in the extraction of

the best characteristics of each dimensional data in a competitive mode, which by themselves are based on the turbidity uncertainty. Using lower uncertainties from other dimensions other than the transmission IR, promotes lower error ratios. Considering that the color error ratio for the data fusion approach is also one of the lowest in all training data range, makes this approach the best for implementation in what concerns the error ratio, but other features associated with the algorithms are also important to analyze, namely the processing times of training and classifying which are discussed next. The overall error ratio performance obtained for this algorithm was 8.83 and 9.04 for Matlab and Python respectively.

4.6.1.2. Processing time

During the 100 cycles used to perform the cross validation analysis, the calculated average time that the regression model approach needed to train the data was around 0.4 seconds for Matlab and 1 second for Python, independently of the percentage of trained data used. In Figure 4.13 we see a representation of the training time of EMGM, LDA, ANN and Data Fusion algorithm normalized to the time values obtained for the regression model training routine for each data training percentage and programming language. The normalization will ease the comparison process between the two languages graphically and independently from the language. Due to the great value difference in magnitude of each algorithm, the train time axis is presented in a logarithmic scale.
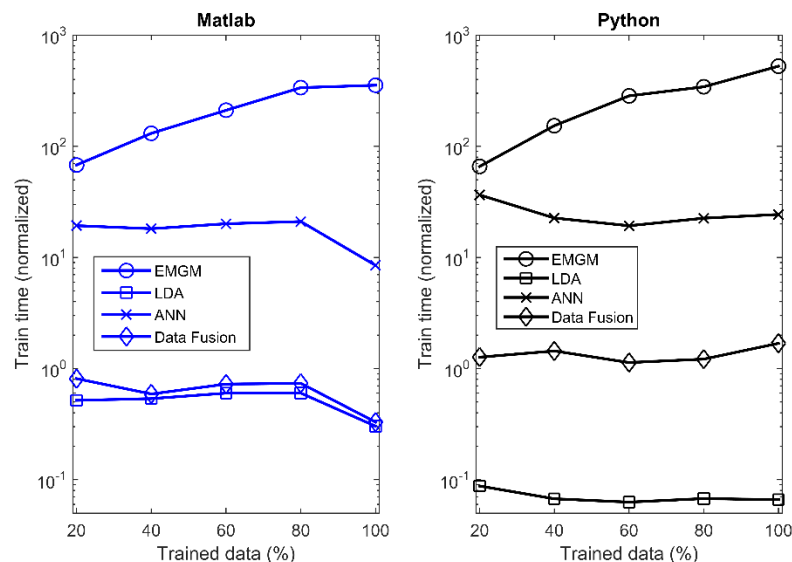
Figure 4.13 – Normalized training time (logarithmic scale), in relation to the regression model approach, with the variation of the training data percentage, for both Matlab and Python languages.

Observing the training times obtained for each algorithm, it is possible to conclude that the EMGM approach is the most computational demanding of them all. The training process of EMGM is, at 20% training data, 65 times slower than the regression model method at the same percentage. When using 99% of the data for training, it becomes 500 times slower at the same percentage. This is almost a linear increase and its easy predicted that the higher of the quantity of data to be trained, the higher the time to train will be necessary. This is clearly a disadvantage compared to the other approaches. The LDA and Data Fusion methods, in turn, does not increases their time to train when more data is available, with their times of training being similar to the regression model. For python, the LDA approach revealed to be even faster than the regression model, achieving 17 times less duration to train, which can be explained by a better adaptation of the script for Python language due to its simpler implementation. The ANN algorithm showed also to have a stable time that is around 20 times slower for both Matlab and Python. The differences observed in processing time between the two languages in some algorithms can be explained, in general, by the lack of optimization done to the python code or libraries used. All the codes were first written and optimized in Matlab and then directly adapted to Python without taking into account its particular differences, as the examples of data structure that in Matlab is column-major while in the Numpy library of Python it is row-major. Matlab also uses Just-in-time (JIT) compilation which is a way of executing computer code that involves compilation during the execution of a program, speeding it up. JIT was not used with python.

While the training routine can be easily performed offline, the classification routine has higher importance due to the demand of high speeds for on-line real-time sensor monitoring, especially in the presence of multiple sensors that retrieve data to the same platform for calculation. If the calculation algorithm is not fast enough, bottleneck processing effects will exist and loss of data is a possibility. The same analysis performed for the obtained training times was also done to the classifying times and can be seen in Figure 4.14.

Figure 4.14 – Normalized classifying time (logarithmic scale), in relation to the regression model approach, with the variation of the training data percentage, for both Matlab and Python languages.

The regression model approach time needed to classify a single measurement for Matlab was 0.006 seconds, while for python was 0.0003 seconds. The EMGM method needed, in Matlab, about tenfold more time to classify the same quantity of training data in relation to the regression model, while in Python it was 60 times over. It can be concluded that the EMGM method is in general a slow process. The LDA approach present a not well-defined behavior since it is 4 times faster to classify than the regression model, using Matlab. It is however 10 times slower using Python. In contrast the ANN method is faster to classify in both Python and Matlab which is expected because of the simple operation process needed. In Matlab it only took 0.02% of the time obtained from the regression models. In Python, the time was 1.25 times faster (0.8% of the regression time), difference that can be explained by the lack of optimization of the data structure as stated before. Nevertheless, it is still faster than the traditional regression models, proving to be highly competitive in applications where low time processing is a priority. For the Data Fusion approach, the classify time seems to be its main disadvantage since it requires about 140 times more time to classify than the simple regression model, which can be observed for both Matlab and Python languages. This higher classification time was expected because of the recursive algorithm present in data fusion that requires more processing time to converge into the final solution.

4.6.2. Without IR information analysis

4.6.2.1. Cross validation

As mentioned before, the infrared light information is not mandatory for the implementation of the EMGM, LDA, ANN and Data Fusion algorithms. While for the LDA and ANN approach the dimensional reduction and the non-linear model are automatically determined by direct implementation to the supervised data, the same does not apply for the EMGM. This is important because the infrared light information was directly used to a first estimate of the turbidity which, in turn, is used to calculate the color in EMGM. Without this information, a recursive cycle similar to the one used in Data Fusion algorithm that minimizes and compares the error of the expected turbidity for each individual dimension is done. This will take more time to classify a measurement and is expected to increase the overall error because of its higher uncertainty due to the reduction of information. In Figure 4.15 it is represented the color error ratio for the EMGM, LDA, ANN and Data Fusion approaches with the different training data percentages and for the both used programing languages.
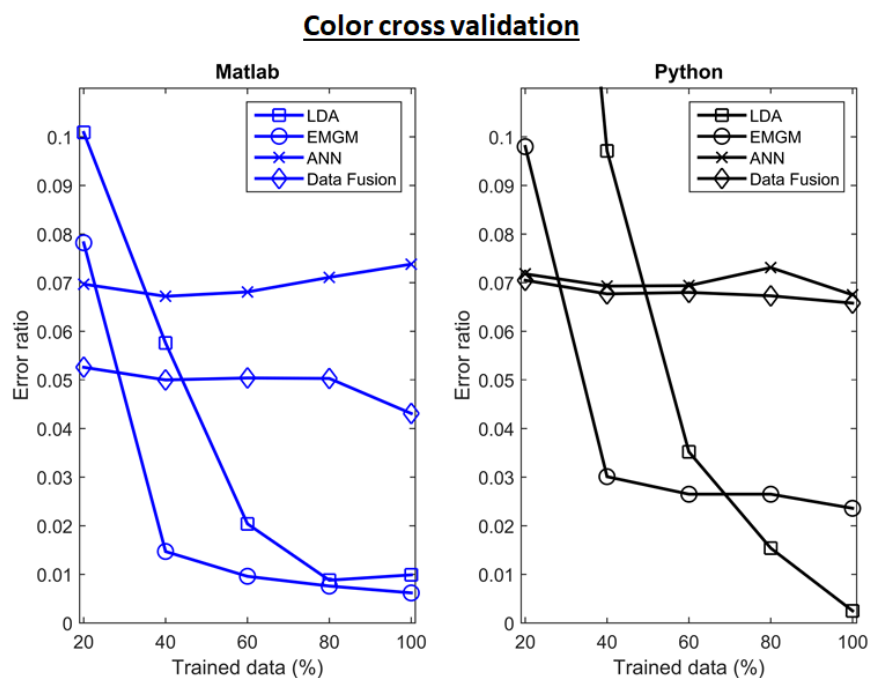
Figure 4.15 – Color cross validation error ratio evolution with the trained data percentage (without IR information) for each classifier methods and for both Matlab and Python languages.

As expected, the overall color error ratio obtained for all the algorithmic approaches without the IR information was higher than the ones obtained with the IR information. Nevertheless, the

increase of error is not that higher that invalidates this approach, which can be very advantageous because it simplifies the necessary hardware of the sensor and lower costs. The same behavior detected for the EMGM and LDA performance with IR is present here, this is, the error ratio is higher with the decrease of trained data but, with sufficient training data, the color error ratio can be less than the ones obtained from the regression model with IR information. The overall EMGM error ratio performance value obtained for Matlab and Python were 1.48 and 2.88 respectively. For LDA the values were 2.84 and 5.70 respectively. The ANN approach has the same constant error behavior as observed before with the IR information but in this case with an error rate around 0.07 instead of the 0.055. The overall performance of the ANN approach was 5.56 and 5.63 for Matlab and Python respectively. The Data Fusion algorithm presents also the same constant behavior as the one obtained with the IR information, but with a higher error ratio of about 0.05 and 0.068 for Matlab and Python respectively. The overall error ratio performance for the Data Fusion algorithm was 3.97 and 5.42. It is proven that, while it is true that the IR light information decreases the overall color error ratio, improving the performance of the sensor, if a simpler and lower cost sensor is needed, the absence of this information will make it still feasible, with a reasonable measurement error.

The turbidity cross validation analysis is presented in Figure 4.16.
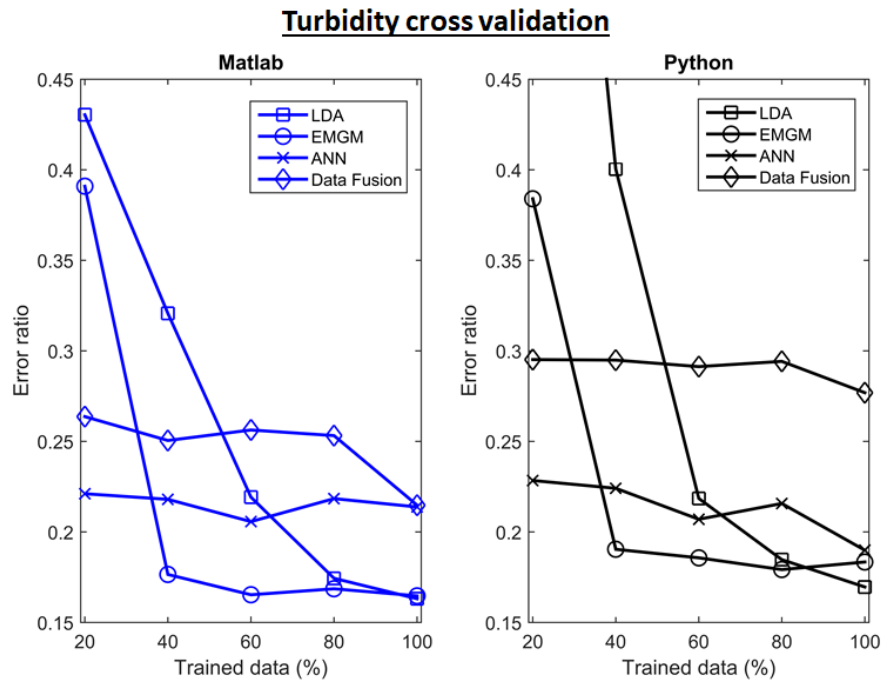
Figure 4.16 – Turbidity cross validation error ratio evolution with the trained data percentage (without IR information) for each classifier methods and for both Matlab and Python languages.

The outcome conclusions obtained from color cross validation error ratio were equally observed for the turbidity cross validation error ratio, that is, the lack of the IR information increases the overall error ratio for all the tested algorithms. The LDA and EMGM present the known decreasing error ratio behavior with the increase of trained data, achieving still equivalent results than the regression model that uses IR information with values of error ratio around 0.17 for high training data percentages. The overall EMGM error ratio performance value obtained for Matlab and Python were 15.76 and 16.78 respectively. For LDA the values were 20.22 and 26.26 respectively.

Contrary to what has been seen previously for the color cross validation, generally the obtained ANN results for turbidity were better (lower error ratio) than the ones obtained from the Data Fusion algorithm, with ANN having error ratio values around 0.21 for both languages and the Data Fusion algorithm with error ratio values around 0.25 and 0.28 for Matlab and Python respectively. The Data Fusion algorithm showed that the IR information is crucial to have high performances with low error ratios while the ANN algorithm was less affected with the lack of it. The overall ANN error ratio performance value obtained for Matlab was 17.19 while for Python was 17.12. In the Data Fusion algorithm, the overall error ratio was 19.99 and 23.33 respectively. With the lack of the IR information, the ANN approach appears to be the best solution for a non-cyclical continuous measurement in a less complex and lower cost sensor solutions.

4.6.2.2.   Processing time

In the processing time analysis, it was performed the same approach used before for the data with IR information, this is, the time values here obtained were all normalized in relation to the previously regression model using IR information, intended to have a fair comparison study with the previous values obtained with the IR information. The achieved results for the training times can be observed in Figure 4.17.
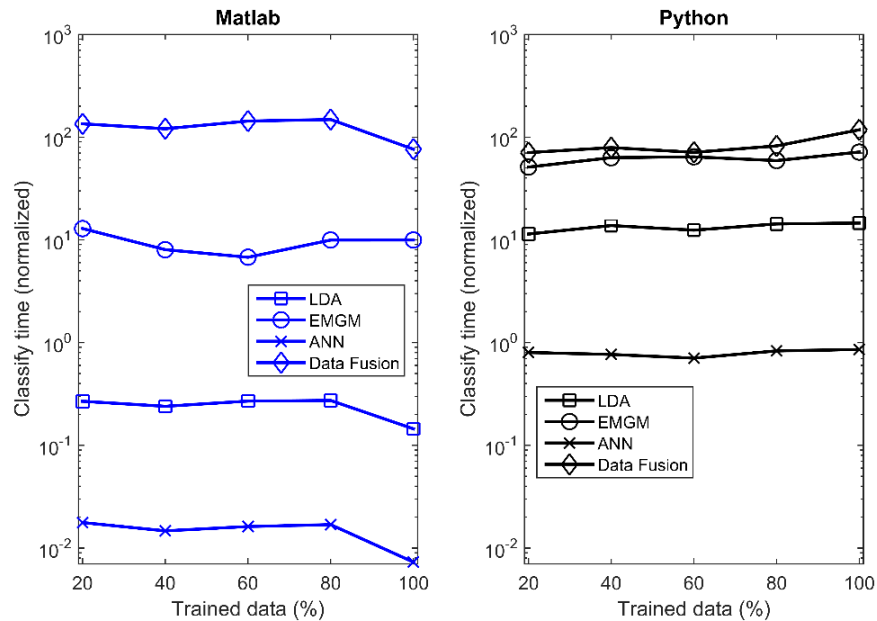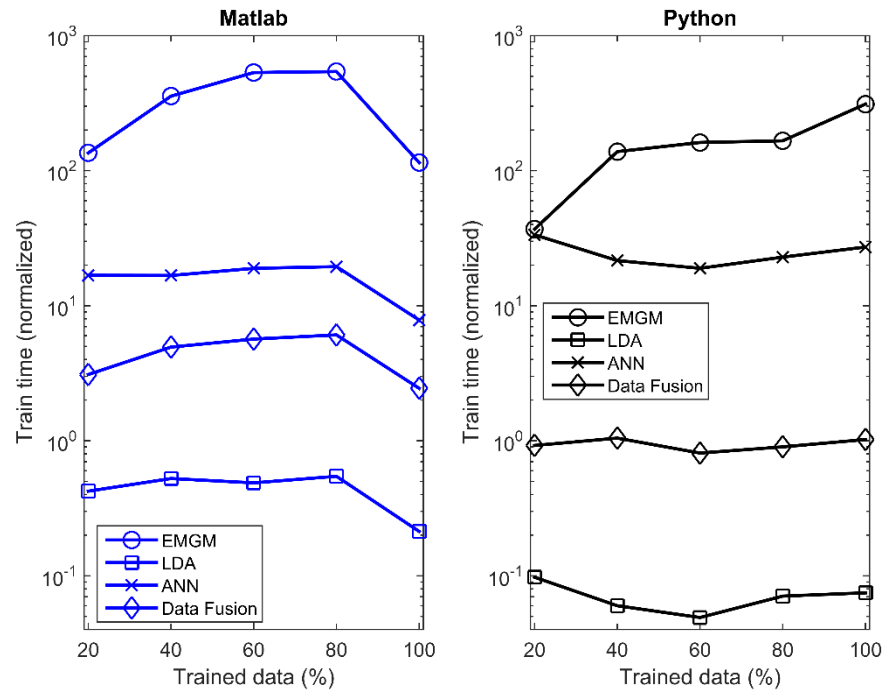
Figure 4.17 – Normalized training time (logarithmic scale), in relation to the regression model approach, with the variation of the training data percentage (without IR information), for both Matlab and Python languages.

Overall there is not much difference between the behavior of the obtained times with IR information and the ones without it. The EMGM approach is the slowest, presenting its increasing behavior with more trained data percentage, achieving 500x more time. The others have a constant behavior with LDA being faster than the regression model and Data Fusion algorithm, having times near the ones observed with the regression model. The ANN was 20 times slower in comparison to the same algorithm using IR information as observed before.

Contrary to the training times that were not affected with less dimensional data by excluding the IR information, slightly differences will be observed for the normalized classification times presented in Figure 4.18.
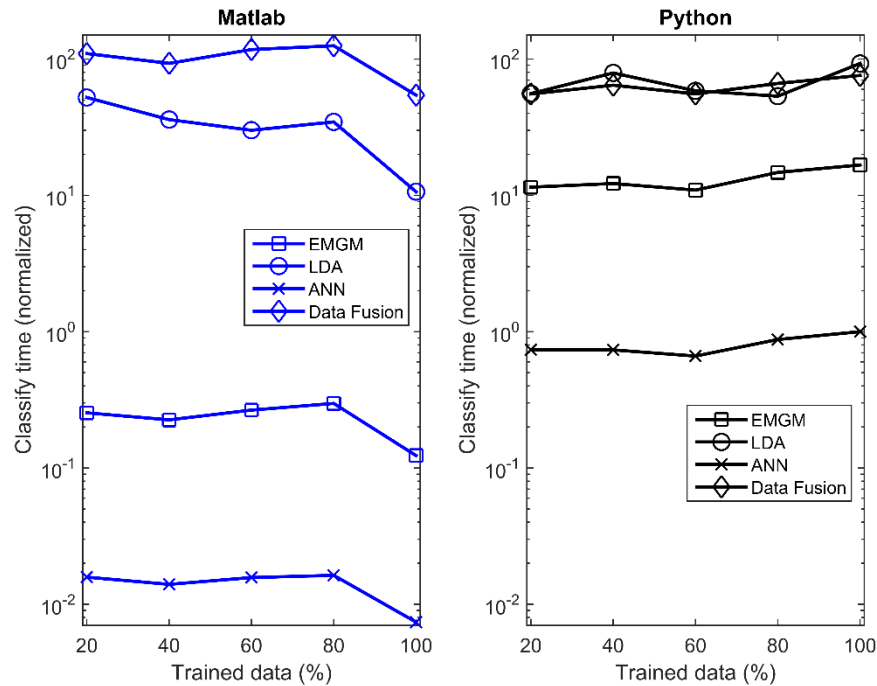
Figure 4.18 – Normalized classifying time (logarithmic scale), in relation to the regression model approach, with the variation of the training data percentage (without IR information), for both Matlab and Python languages.

The main difference observed is the increase of time necessary in Matlab to classify the data for the EMGM process since, as mentioned before, it uses an alternative recursive time costly procedure to initially estimate the turbidity value. This increases its spending time to classify by 3x more in comparison to the normal method. In total, it is 30 times slower than the regression model using IR information. In Python this increase is not observed, with the times being just slightly higher than the ones obtained with the IR information which can be explained by the faster low-level calculation used for the recursive procedure in which Python is efficient.

While the EMGM presented some higher calculation time to classify, the LDA and ANN presented around the same time as seen before for the analysis with IR information, confirming the faster behavior of these algorithms to present a result in comparison to the regression model. The Data Fusion algorithm, in other hand, have slightly lower calculation times in relation to the ones obtained with the IR information, this is, 110x slower than the regression model. This improvement is easily explained by the lower data attributes used (no IR information) that will converge faster to a final result in the recursive algorithm. Nevertheless, the Data Fusion algorithm stands, still, as the slower of the used algorithms.

4.6.3. Vantages and disadvantages summary

All the algorithms here presented have proven to be feasible, containing each one its advantages and disadvantages that can be explored by considering its final application. The traditional approach of regression models offered fast training times while maintaining a low error value. Unfortunately, it needs the presence of the IR information which requires more complex hardware system to be used. In its turn, the LDA can be used without the IR information and is faster than the regression model in training, taking around the same time in classifying. It is however only adequate to be used in measurements where color liquid has a well-defined cyclic pattern and where good data for training is obtainable. In these circumstances, LDA present lower error values than the regression models, but if these criteria are not met, high error values will be acquired. The same conclusion is applied to the EMGM algorithm, but in this case better overall error results are obtainable. The major drawback in comparison to the LDA is its processing times for both classification and training that are the highest between the analyzed algorithms.

Regarding the ANN algorithm, it revealed to be the fastest to classify of all tested methods, having a turbidity error similar to the regression model and a color error slightly higher. It as the great advantage of not needing to know the physical models and can therefore be directly applied to the supervised data with any dimension. It is slow to train but its faster speed to classify makes it perfect for real-time monitoring, mainly avoiding data loss from bottleneck processing when single processing unit is used for multiple sensors. ANN also showed to be the best algorithm to be used when IR information is not an option.

The Data Fusion algorithm achieved its objective to present the lowest error value in turbidity while maintaining the color error value observed in the regression model. The price paid for the higher precision was getting the slowest classification speeds between the algorithms, but with the training time around the same as observed from the regression models. Although without the IR information the Data Fusion algorithm presented worse results than the ANN algorithm in the turbidity error rate, it is still a good algorithm to be used mainly for color determination.

A resume of all the advantages and disadvantages observed in the analyzed algorithm can be seen in Table 4.3.

Table 4.3 – Resume table of the advantages and disadvantages of Regression model, EMGM, LDA, ANN and Data Fusion algorithms for color and turbidity determination.

	Regression Models		EMGM		LDA		ANN		Data Fusion	
	Matlab	Python	Matlab	Python	Matlab	Python	Matlab	Python	Matlab	Python
Physical model	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓
Color overall error ratio performance	2.68	3.08	1.25	4.15	2.16	5.50	4.30	4.37	2.70	3.22
Turbidity overall error ratio performance	13.33	13.65	16.26	16.38	18.64	26.45	13.92	14.01	8.83	9.04
Train speed	Fast		Very slow		Very Fast		Slow		Fast	
Classify speed	Fast		Very slow		Fast		Very fast		Very Slow	
Mandatory IR Information	✓		✗		✗		✗		✗	
Color overall error ratio performance (No IR)	------		1.48	2.88	2.84	5.70	5.56	5.63	3.97	5.42
Turbidity overall error ratio performance (No IR)	------		15.76	16.78	20.22	26.26	17.19	17.12	19.99	23.33
Observations	------		Particularly good for cyclic color variation patterns		Particularly good for cyclic color variation patterns and real-time measurements		Potentially good for multiple dimensional data real-time measurements		Particularly good for higher precision measurements	

5. Case Studies

5.1. Introduction

This chapter will be dedicated to the application of the developed optic sensing platforms in the case studies of grape press, wine fermentation/ageing and cleaning solutions. In grape press and wine fermentation/ageing it will be analyzed the performance of the color and turbidity sensor in real industrial conditions. This sensor was chosen because the parameters used are in the interest of the wine producers and are possible to be monitored with the developed sensor. In the cleaning solution case study, samples obtained from a maintenance and cleaning company of industrial machines was obtained. The cleaning solution used to wash the industrial machines needs to be changed regularly after some usages. Here the objective is to find if the developed refractive index sensor could be used to differentiate the progression of the solution until it becomes unsuitable for cleaning.

To achieve these goals, a collaboration with Watgrid was made, taking advantage of their know-how in the industry and using their internet platform for data communication and visualization. Considering the performance of the sensors developed and tested in chapter 3.2.2 and the conditions of testing in wine, only the laser sensor, the high responsivity LED sensor and the peristaltic sensor (sensor versions 3, 4 and 5 respectively) are used. For data processing, the data fusion algorithm developed in chapter 4.5 is chosen due to its highest turbidity measurement precision, even with its higher processing time, which in this case study it is not problematic. To compensate the turbidity and color sensor for the refractive index variations, the Watgrid density sensor was used and the density values converted in approximation to refractive index values. Since the refractive index sensor developed in chapter 3.3 was not food compliant, it could not be used for compensation. It was however used for the cleaning solutions assessment.

5.1.1. Wine fermentation, ageing and analysis

The production of wine can be traced since pre-historic times, as early as 10 000 years ago based on some evidence obtained in settlements which indicate that wild grapes were gathered to make wine. The evidence of cultivation of the wild European wine grapes suggests that it has a beginning at around 6000 years ago. Ancient Egyptian royalties of 2700 to 2500 BC, already used wine for

religious ceremonies. Its spread through Mediterranean was promoted by the Greeks from 1600 BC to its colonies and later by the romans. The Romans influenced and established the foundation of viticulture and are credited with:

- Classifying grape varieties (varietals) and colors
- Charting observations of ripening characteristics
- Identifying diseases
- Recognizing soil-type preferences
- Improving pruning skills
- Increasing production through irrigation and fertilization
- Using wooden cooperage
- Using glass bottles

Following the fall of the roman empire and the rise of European colonizer powers, production of wine and its consumption was spread through the world becoming today valued as a 28.9 billion euro market with production of around 267 million hectoliters [Jacobson, 2006a; Organisation internationale de la vigne ey du vin, 2017].

The process of winemaking starts with the harvest of the grapes that are taken into a winery. Here, the grapes are pressed or crushed. For white wines only the juice is used which is moved to tanks. For red wines, whole fruit is used (whole must) for fermentation. In this phase, an analysis of the must is already taken to ensure that the must is prepared for fermentation with the conditions pretended by the winemaker. The apparent sugar content is one of the analyzed parameters that will indicate the potential alcoholic content of the must after the fermentation. Other important parameter is the insoluble solids content that consists of small particles (below 2 mm) of cellulose, hemicellulose, pectin, mineral salts, lipids, and insoluble proteins. Figure 5.1 shows the size distribution of a typical red wine must.
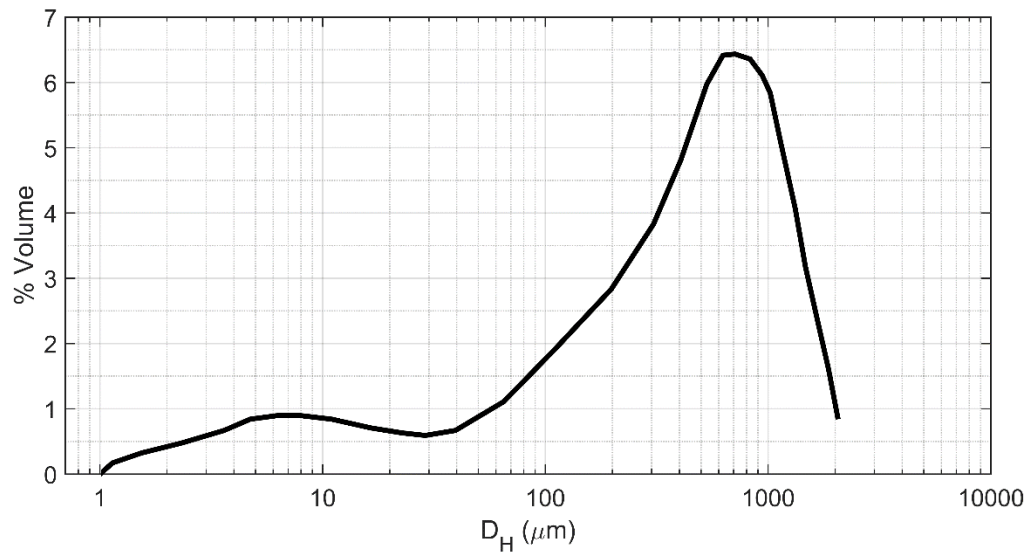
Figure 5.1 – Particle size (Dₕ) distribution in a raw red wine must (adapted from [Vernhet, 2016]*).*

The proper concentration of solids in juice is necessary for the fermentation process proceed to completion, with the added benefit of improving the fruity flavor and quality of white wines [Jacobson, 2006c]. In practice, low turbidity (between 50 and 150 NTU) usually favors the formation of fruity aromas, whereas high-turbidity musts present heavy, herbaceous aromas with sometimes off-odors. High turbidity may also favor browning. On the other hand, excessive clarification may negatively affect fermentation progress [Vernhet, 2016]. With all the finished adjustments made to the must by the winemaker, natural or induced cultured yeasts will start to metabolize, this is, breaking down organic nutrients like sugar ($C_6H_{12}O_6$) to have their source of energy as adenosine triphosphate (ATP). Only a portion of the energy created is used while the rest is dissipated as heat. A by-product of this reaction is the formation of ethanol molecules (CH_3CH_2OH) and carbon dioxide (CO_2). The reaction can be simple expressed as:

$$C_6H_{12}O_6 \rightarrow 2(CH_3CH_2OH) + 2CO_2 + Energy \qquad \text{(Eq. 5-1)}$$
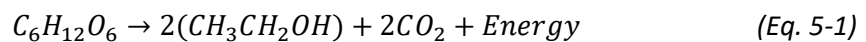
In ideal conditions, most white fermentations can be completed in 2–3 weeks. Red fermentations can usually be completed in a few days, less than the white wines. The fermentation progress of all the juice and must should be closely monitored until it is finalized. Usually, staff from the winery will monitor fermentations once or twice a day, determining the probable sugar content and noting the temperature of the fermentation. These results will depict a trend that is used to analyze if the process is working as expected or if there are potential problems that need to be addressed with anticipation, mainly a stuck fermentation that can have a bacterial spoilage, resulting in the waste of an entire tank.

Usually the control of sugar content during the fermentation is done by floating a hydrometer or a density meter in a column of juice. The hydrometer has been designed to float at a level that will be equal to the correlation with the probable sugar content at various densities. The fermentation starts with density values of around 1.090 g/cm^3, mainly due to the sugar presence, having a matching refractive index value of around 1.3675 RIU. With the transformation of sugar into ethanol, an exponentially decreasing process is observed, *i.e.*, it starts with a higher rate, slowing during the process, reaching a density value around 0.990 g/cm^3 that matches with a solution with a refractive index of about 1.3345 RIU. An example of the evolution of density and refractive index during the fermentation phase can be observed in Figure 5.2.
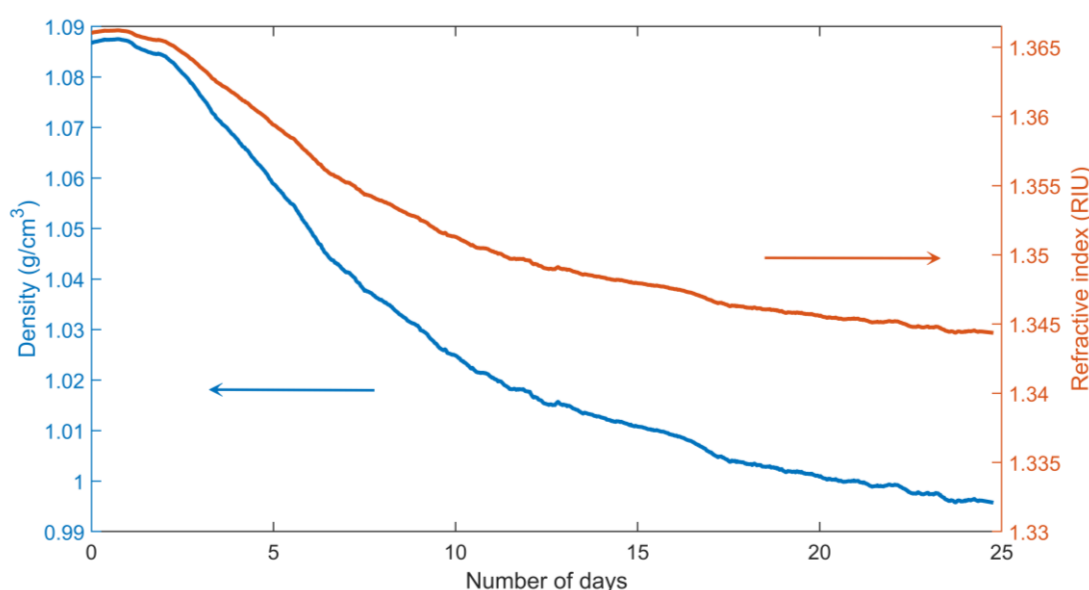
Figure 5.2 – Typical density and refractive index evolution of a white wine fermentation.

Wine color has dependency on the amount of anthocyanins in the grape, the pH, fermentation reduction, and sulfur effects. Anthocyanins are water-soluble pigments that can be found in most of food plants which may appear red, purple, or blue. In the grape, they are mostly found in the skin. Red wines are several days in contact with their skins after harvest for extracting as many anthocyanins and other compounds as possible. This will increase initially the color intensity of the must. The opposite effect is observed by the gradual alcohol content increase in the must with the fermentation kinetics in progress. Alcohol will reduce color intensity because of the hydrogen-ion bonds that are broken duo to its rise [Jacobson, 2006c].

The absorbance spectrum of red wines has a maximum at 520 nm due to the anthocyanins and its combinations with other molecules, and a minimum value in the region of 420 nm. The current and most used approach to color analysis in winemaking, known as Glories method, requires optical

measurements at 420 and 520 nm, with an additional measurement at 620 nm to include the blue component in young red wines [Ribéreau-Gayon, 2006a]. These measurements are usually referred as optical density (*OD*). Color intensity (*CI*) represents the amount of color. It can vary largely from one wine and grape variety to another (usually up to 18 AU) and is calculated as the sum of all the optical density measurements:

$$CI = OD\,420 + OD\,520 + OD\,620 \qquad \text{(Eq. 5-2)}$$

The color hue indicates the development of a color towards orange. In red wines, during fermentation, it has a value on the order of 0.5 - 0.7 which increases under an ageing process, where it can reach an upper limit of around 1.2 - 1.3 [Ribéreau-Gayon, 2006b]. Its values are calculated by the fraction:

$$Hue = \frac{OD\,420}{OD\,520} \qquad \text{(Eq. 5-3)}$$

The color of white wines is more complex to analyze, as the spectrum has no defined maximum in the visible range. Therefore, there is higher difficulty to use this parameter to interpret it with a visual impression. Since the characteristic absorption wavelength of yellow substances is 420 nm, measurements of optical density at this value provide an approximate assessment of color. A typical spectrum of a white and red wine can be seen in Figure 5.3.
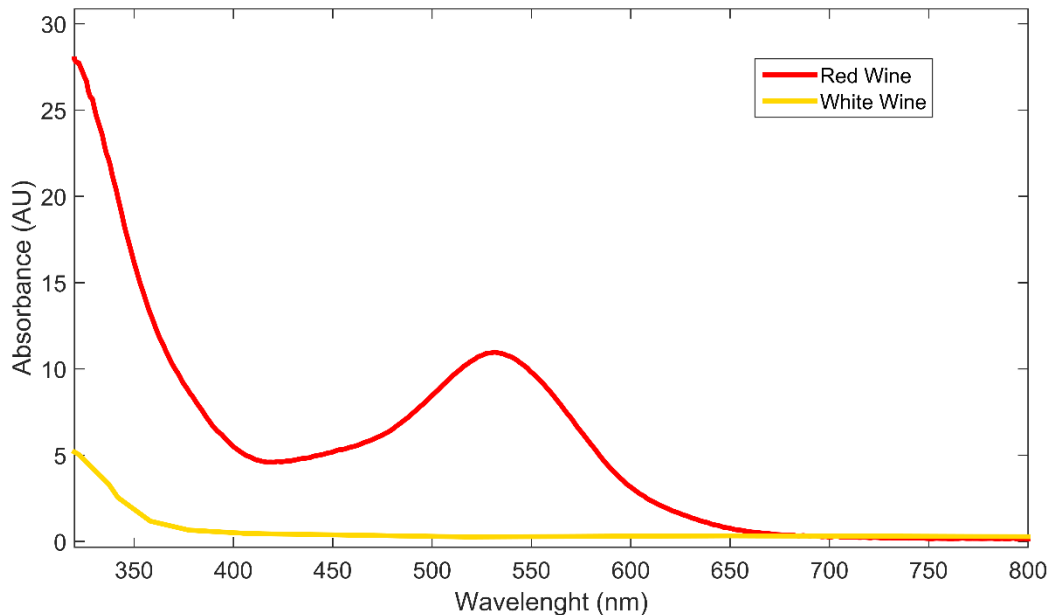
Figure 5.3 – Typical red and white wines optical spectrum in the visible range.

Wine color monitoring can be used as a way to analyze the stability of the colored pigments extracted from the skin of the grapes during fermentation [Yokotsuka, 1997]. There are also strong correlations between the color of the wine and its quality ranking [Chris Somers, 1977; Jackson, 1978]. Adding to this, wine color is also one of the most important characteristics of marketing, which influences the consumers' overall acceptability [Gómez-Míguez, 2007]. Figure 5.4 presents the expected behavior of the color intensity variation during a fermentation with its initial increase due to the red pigments extraction, followed by a decrease due to the higher ethanol content [Yokotsuka, 1997; Ribéreau-Gayon, 2006c].
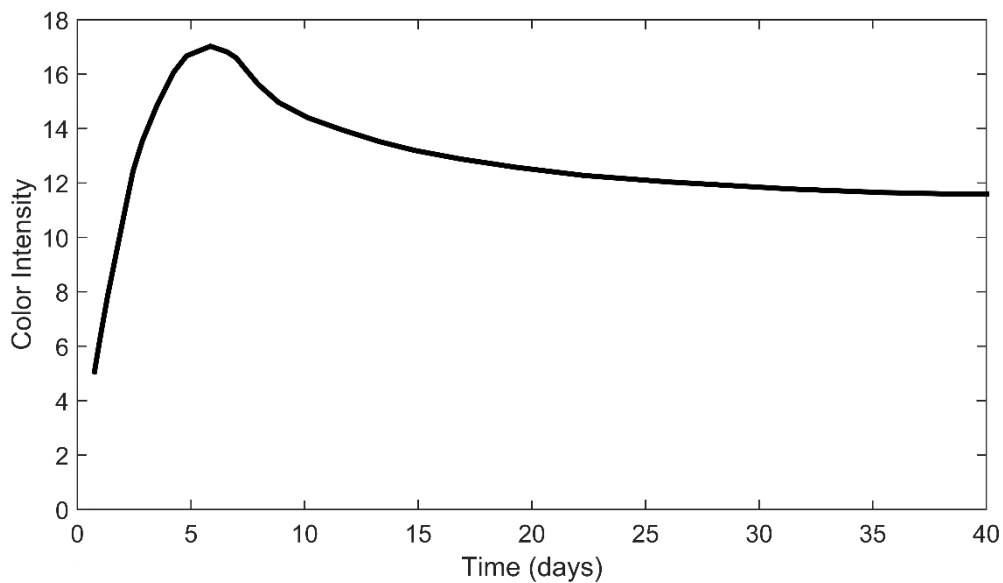
Figure 5.4 – Color intensity evolution of red wines with time of fermentation (adapted from [Ribéreau-Gayon, 2006c]*).*

During fermentation, variations of insoluble solids suspended in the juice can be noticed, mainly in the starting process. Here, an increase of temperature due to the energy output of the metabolic reaction will promote a higher motion of the liquid and by consequence the constituent particles. When it reaches the peak of the metabolic process and the temperature starts to drop, the motion of liquid becomes lower and the quantity of suspended insoluble solids starts also to drop, since it will be slowly deposited in the bottom of the tank. This quantity of insoluble solids can be measured by turbidity units and, as stated, its monitoring can be a useful parameter to evaluate the overall performance of the fermentation process and the quality of the finished product. Turbidity is also related to the fermentation kinetics since it presents similar behavior to the release velocity of CO_2 and to the growth of the yeast number [Garcia, 2008]. Turbidity is usually very low for a finished bottled wine with values below 2 NTUs but, during the fermentation processes, its value can exceed 2000 NTUs [Crespo, 2010] (Figure 5.5).
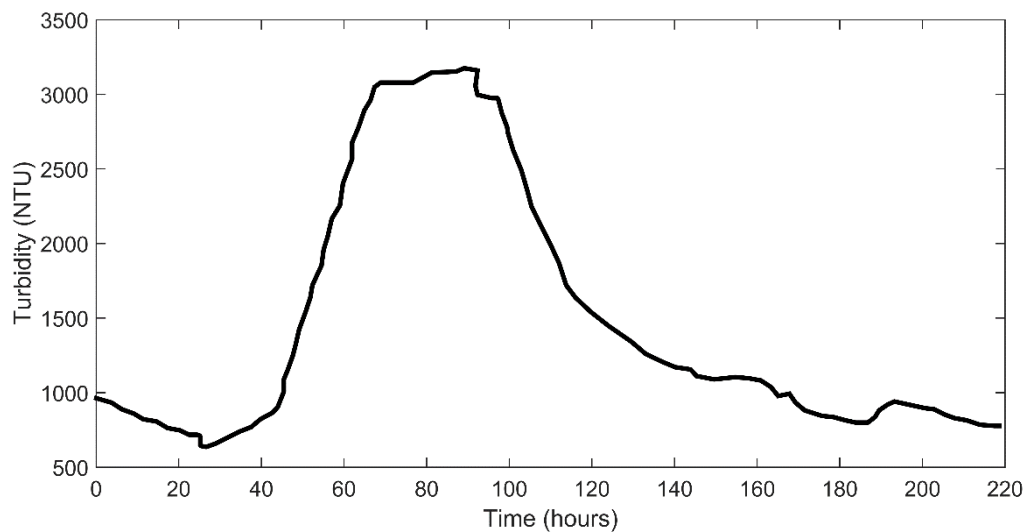
Figure 5.5 – Typical evolution of red wine must turbidity in the fermentation process (adapted from [Garcia, 2008]*).*

With the completion of the fermentation process, the winemaker will then proceed with a plan to address specific aims foreseen for the future of the young wine. This could include the type of protection of the wine, the immediate bottling or ageing in a storage tank or barrel, the time of its bottling, etc. When ageing is the way to follow, some steps should be done: removing unwanted substances; storage the wine in proper containers; protect it from contaminants, oxidation and further fermentation; and finally perform an analysis clarification and stabilization of the wine. Ageing will promote chemical changes and removal of substances in the young wine to improve its quality, with the removal of CO_2 formed from the fermentation as an example. These chemical interactions can, however, have durations from some month to several years. The wine aroma, flavor, taste, and color are parameters that could change with the addition of oxygen (oxidation) and, when stored in barrels, with the extraction of oak compounds. Temperature of storage will also play an important role during this period to control these chemical changes. When the ageing is finalized, the bottling of the wine is initiated [Jacobson, 2006b].

5.1.2. Industrial cleaning service

Maintenance in industrial machines is a requisite that needs to be regularly performed to prevent failures that could jeopardize an entire production line. Industrial and motorized components will need to be carefully cleaned, a task that is usually executed with tanks using manual or automatic cleaning routines Figure 5.6.

Figure 5.6 – Typical tank used for cleaning industrial machinery.

High pressure and high temperature water containing specific detergent is used and a single cleaning solution can be used several times before it becomes not proper to be used again, due to the accumulation of oils or dirt particles. When the cleaning solution is fully dirty, the maintenance company that provided the automatic cleaning machine will need to retrieve the contaminants and change the cleaning liquid, a task that requires trained staff to be summoned to the client company. Usually this procedure is done regularly at predefined period of time, however the client can proceed with cleaning routines with a faster pace than the one originally planned, which will require the unexpected summon of the staff earlier than usually. This provides logistic problems and increased costs to the company due to the unplanned transportation routes needed. On the other hand, the cleaning solution can be overused until the operator realizes, by experience, that it is not proper, which causes a cleaning routine where, in reality, the cleaning is not occurring. At this point, time and energy of the client was also needless spent. The opposite case, this is, if the cleaning solution is not used fully to achieve its cleaning potential due to its early change by the staff, lower profitability will be made. The oils and particles obtained from the machinery washing will have an effect in the refractive index of the cleaning solution that could be detected using a proper sensor. If the sensor has the capability to detect the refractive index and send the data to a cloud server, trigger alarms could be implemented to advise the cleaning company that one of the machines will need to change its cleaning solution in short time. This will allow for better planning routes and avoid cleaning solution that are not used at its potential. It also rises the customer satisfaction and service quality since it is not necessary to contact the cleaning company if the cleaning solution is overused and needs to be changed.

5.1.3. Data collecting and visualization platform

In this work, it was used the data collecting and visualization platform from Watgrid: Winegrid®. The platform allows wireless communication, from both proprietary and third party sensors, to cloud platform which processes, stores and displays the measured data obtained in real time from the sensors [N. Oliveira, 2017] (Figure 5.6). Some of their measured parameters include the temperature, density and liquid level.
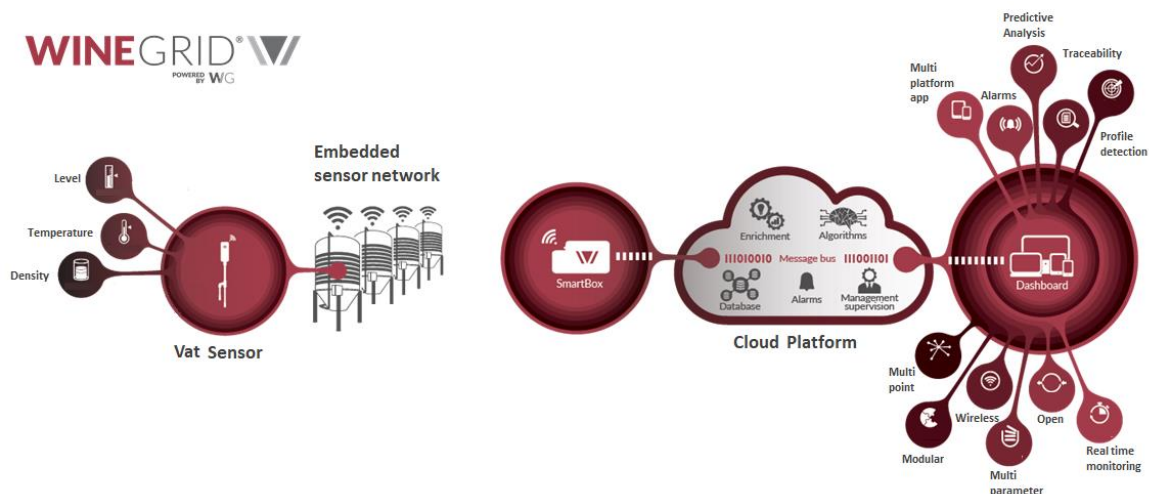
Figure 5.6 – Winegrid full platform schematic.

To display the data, the Winegrid platform has integrated an internet based dashboard for data visualization with potential to integrate sophisticated predictive analysis and alarm functionalities (Figure 5.7). This platform can also provide complex algorithmic calculations for data treatment like the ones developed in chapter 4 and data extraction to local storage. These capabilities and the know-how provided by the company created a synergy between both parties. The access to real winemaking industry and to samples from industrial cleaning service was provided for testing and evaluate the developed sensors and algorithms in pilot case studies. Access to the company's density and temperature sensor data was also used to approximately compensate for refractive index influence in the turbidity and color sensors in the cases where this was necessary, since the developed refractive index sensor was not food grade compliant.
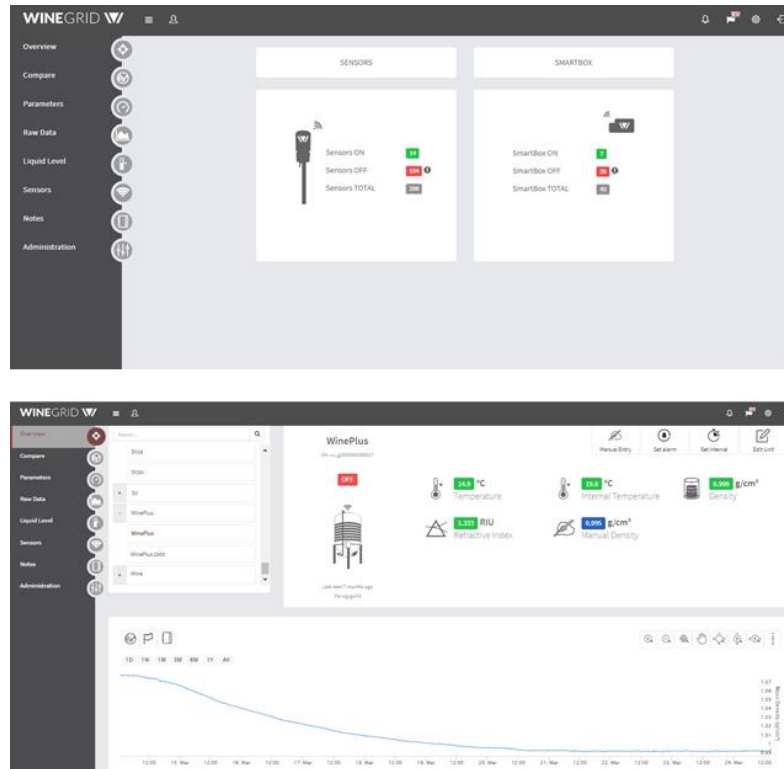
Figure 5.7 – Winegrid platform dashboard for sensor monitoring and data visualization.

5.2. Wine monitoring results

5.2.1. Grape press

The first case study here presented is related to the monitoring of the initial juice obtained by a mechanical pressing machine. This procedure precedes the fermentation which leads to the creation of the wine. Turbidity and color parameters will be measured, with the color having the highest interest for a producer to define the color of the juice before starting the fermentation process. The sensor used was the peristaltic sensor (sensor version 5) which was easier to be placed in-line with a closed circuit. Refractive index measurement for compensation will not be necessary since the sensor is calibrated with the expected refractive index in consideration, due to the constant grape sugar content. The press procedure is taken in steps of increased pressure and release applied to the grapes. In general, the whole procedure lasts for a few hours and so, a continuous and fast data acquisition during the process was performed which takes about 12 second between measurement points. The turbidity measured during the whole press procedure can be observed in Figure 5.8.
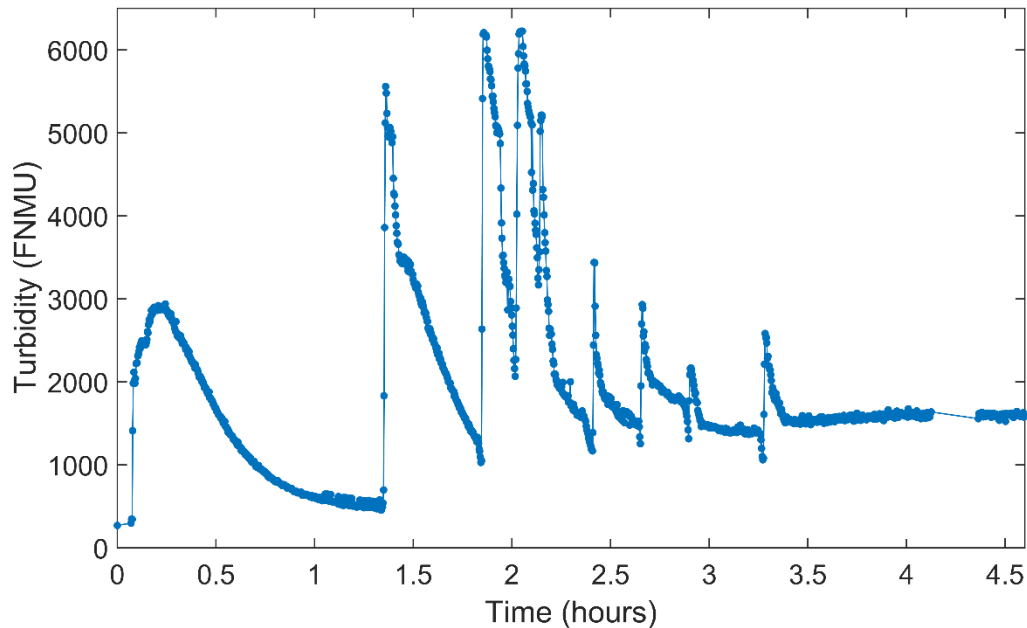
Figure 5.8 – Turbidity measurement during the press procedure where the initiation of pressing and release is clearly seen by the peaks formed.

By analyzing the turbidity behavior, it can be clearly seen the moment when pressure and release is applied to the grapes with instantaneous increase peaks of turbidity, followed by its decrease due to the deposition of particles to the bottom of the tank. This process was repeated 9 times and each one lead to an increase of the turbidity measured when in rest for some time, which is an indicator that more small particles and easier to be in suspension longer times, are being added to the juice from the grapes' skin. The final turbidity obtained was about 1600 FNMU.

The effect of pressing the grapes is also easily visualized by the color intensity obtained with the sensor, where, as expected, clear steps of increasing intensity are observed (Figure 5.9):
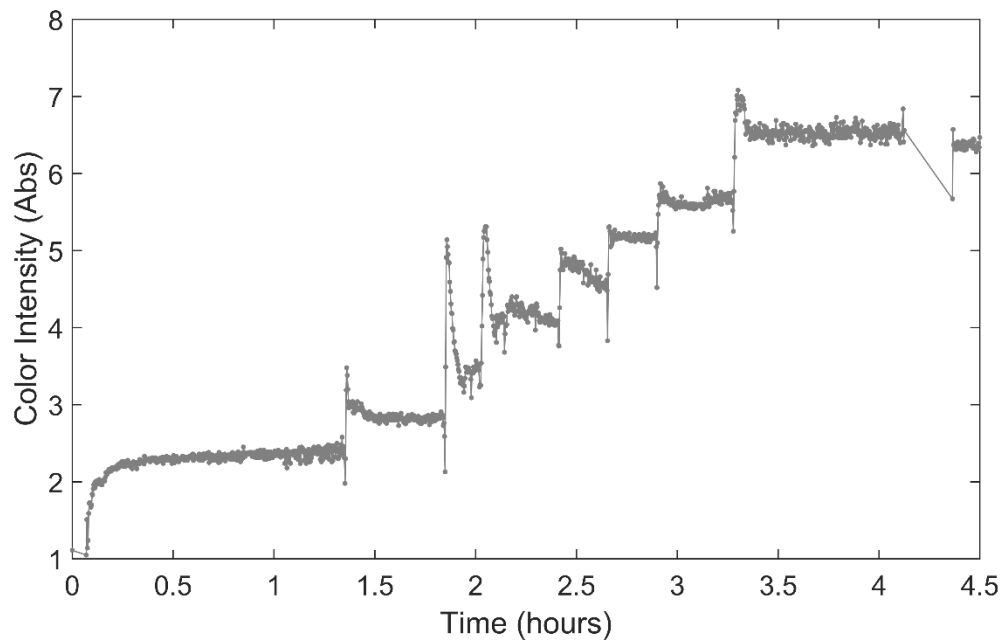
Figure 5.9 – Color intensity measured during the press procedure.

With each pressing, the color increases due to the extraction of the red pigments from the skins of the grapes. It can be seen that, with the release of pressure, the color maintains a very stable measurement until the next pressing, even with the initial decrease of turbidity, showing that the developed data fusion algorithm is performing as intended by correctly differentiating both parameters. The final obtained color intensity was about 6.5 absorbance units, a value expected for a red wine must before its fermentation period.

The extraction of the red pigments with each pressing step is also a conclusion supported by the visualization of the color hue data obtained from the sensor and presented in Figure 5.10. Here, with more red pigments being extracted, a lower color hue values are registered due to the higher absorbance occurring to the 520-nm green light. It decreases from an almost yellowish juice with values near 1, to a heavy reddish juice in its final press, registering a value around 0.65.
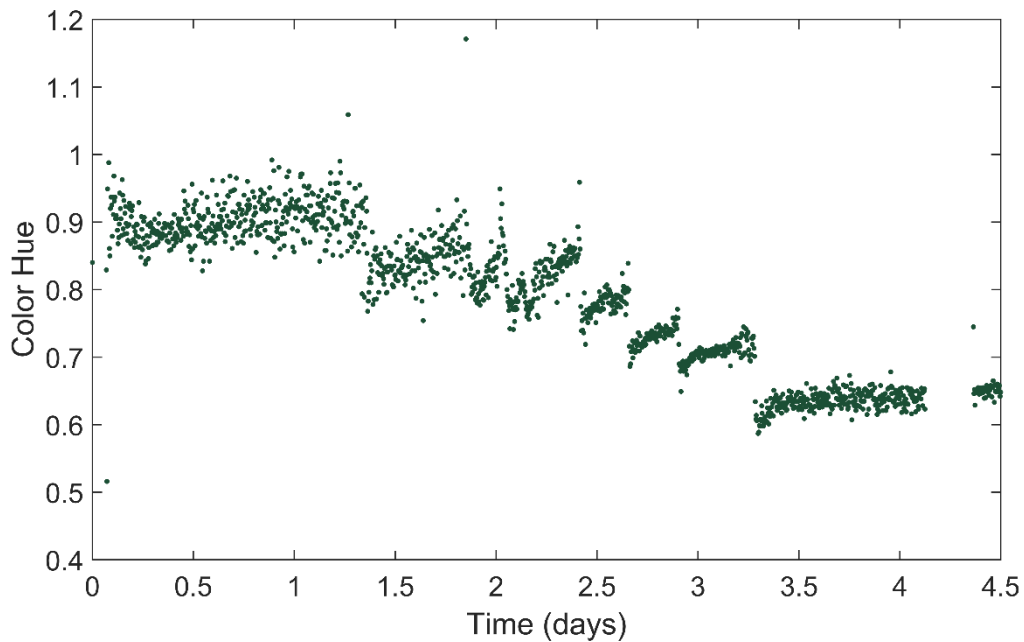
Figure 5.10 – Color hue evolution measured during the press procedure.

5.2.2. White wine fermentation

The opportunity to test the behavior of the turbidity and color sensor with a white wine fermentation in real manufacturing conditions was given by two different winemakers, being one of them in Portugal (white wine A - wwA) and the other located in France (white wine B - wwB). With wwB located abroad, it was also a good assessment in the long-distance behavior of the data communication system used by the sensor and integrated in the Winegrid platform. In both case studies, only the high responsivity LED sensor (sensor version 4) was used for turbidity and color assessment with a modification done to the wwA version where a higher height arm was used to be best adapted for a higher tank. The Watgrid density sensor was used to estimate the refractive index and therefore compensate the influence of it in the turbidity and color measurements. The two fermentations also presented different behaviors, proper of the specific companies' process and consequently requiring different time periods to finalize.

5.2.2.1. White wine A (wwA)

The fermentation period of wwA was of about 5 weeks which is the expected time for a white wine fermentation that takes longer time than red wine that only needs a few days. The density evolution of the fermentation and its equivalent refractive index values can be observed in Figure 5.11.
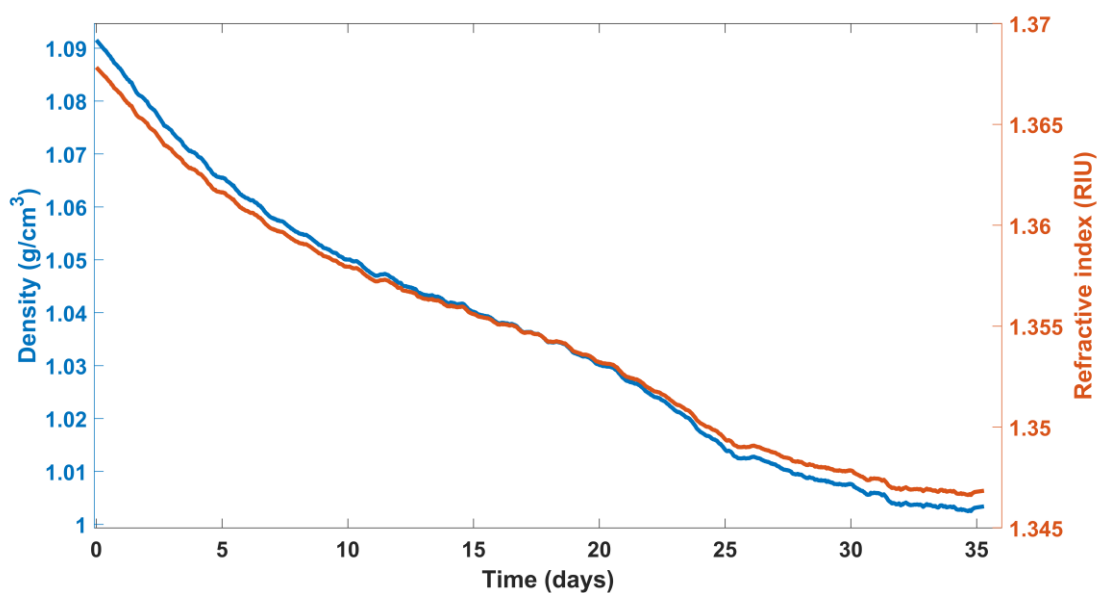
Figure 5.11 – Density and refractive index (calculated) evolution during the fermentation of white wine A.

Because of the longer arm used to measure, less light was reaching the measurement head and the emitting light fluctuation due to changes in the temperature of the sensor electronics and optical system are more pronounced, which can be seen as an oscillating pattern in the measurements. The obtained results for the turbidity evolution during this period can be seen in Figure 5.12, with a smoothing line for a better analysis of its behavior.
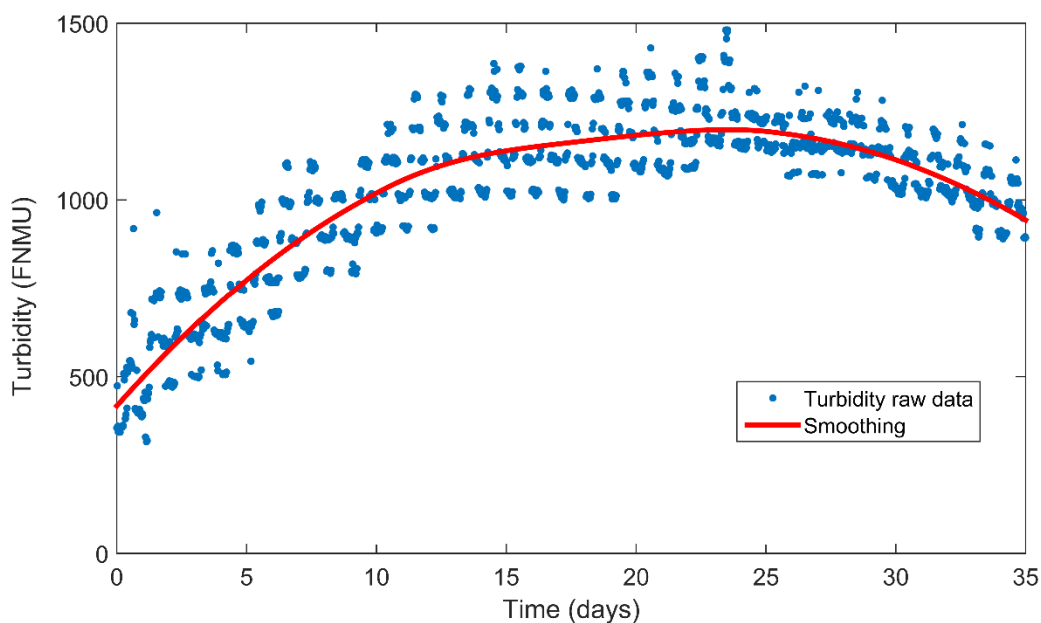
Figure 5.12 – Turbidity evolution during the fermentation of white wine A with oscillating pattern visible due to electronic temperature variation.

As expected, the pattern observed of the initial increase in turbidity is due to the start of the transformation of sugar into ethanol, releasing CO_2. The peak of the fermentation is reached with a value of 1200 FNMU, where the chemical conversion rate slows, promoting lower turbidity measurements. It doesn't appear to have had a high measurement variation when taken in sequence besides the oscillation observed by temperature. This is an indicator that the particles in this juice are very uniform in size.

Being a white wine, the color intensity values expected and measured are not very high. In this type of winemaking the values measured raised from 0.7 Abs units to near 1.0 Abs (Figure 5.13), staying mostly stable during all the process. Due to the interconnection between the turbidity and color measurements, the same temperature based oscillating pattern was observed in the color measurements.
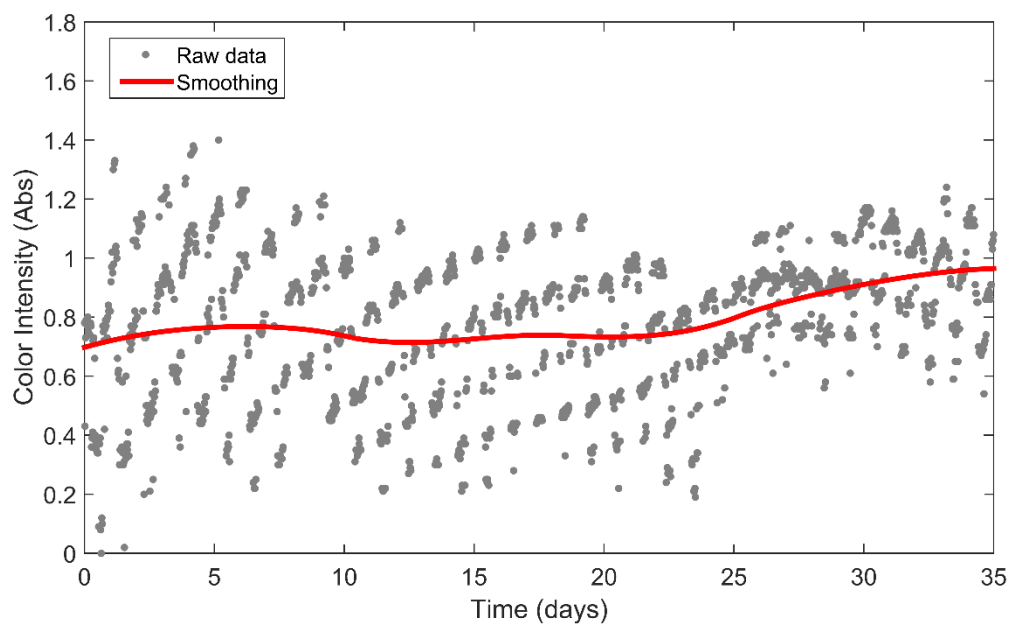
Figure 5.13 – Color intensity evolution during the fermentation of white wine A.

Measuring the color hue in white wines is not as interesting as the other parameters and does not give as much information as when measuring it for rosé and red wines. Nevertheless, its evolution with time during fermentation can be seen in Figure 5.13.
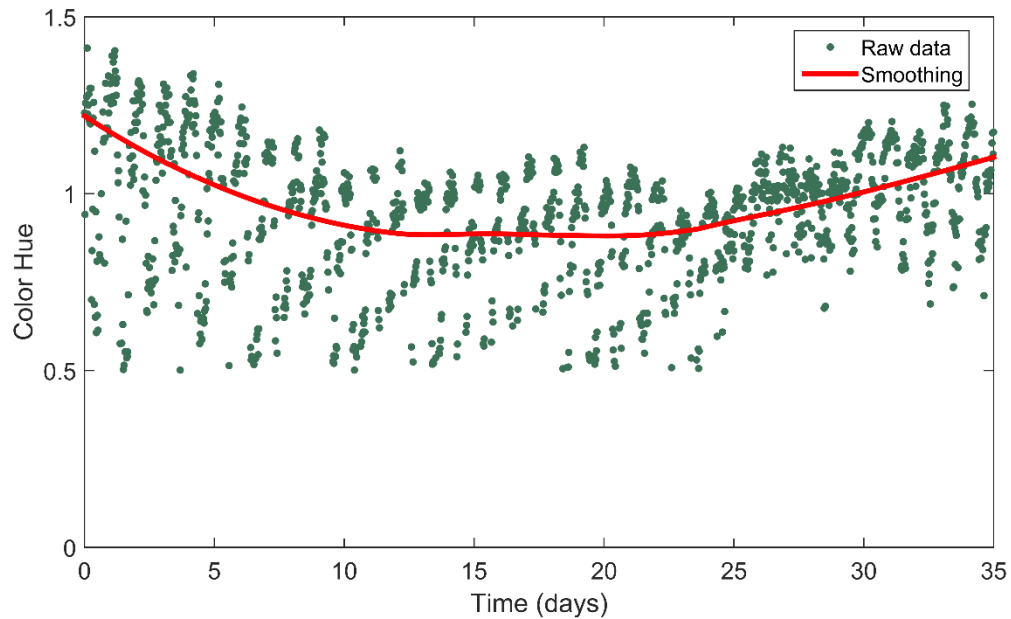
Figure 5.14 – Color hue evolution during the fermentation of white wine A.

With the measured values above 1, there is the indication that blue light was higher absorbed than the green one as is expected from a white wine juice. The decreasing behavior with time suggests that red pigments started to appear, which lowered the absorbance of green light and turned the juice slightly more reddish.

5.2.2.2.   White wine B (wwB)

The fermentation measured from wwB has taken a much shorter period than the observed from wwA. The density evolution obtained by the Watgrid sensor and its refractive index assessment for compensation is presented in Figure 5.15.
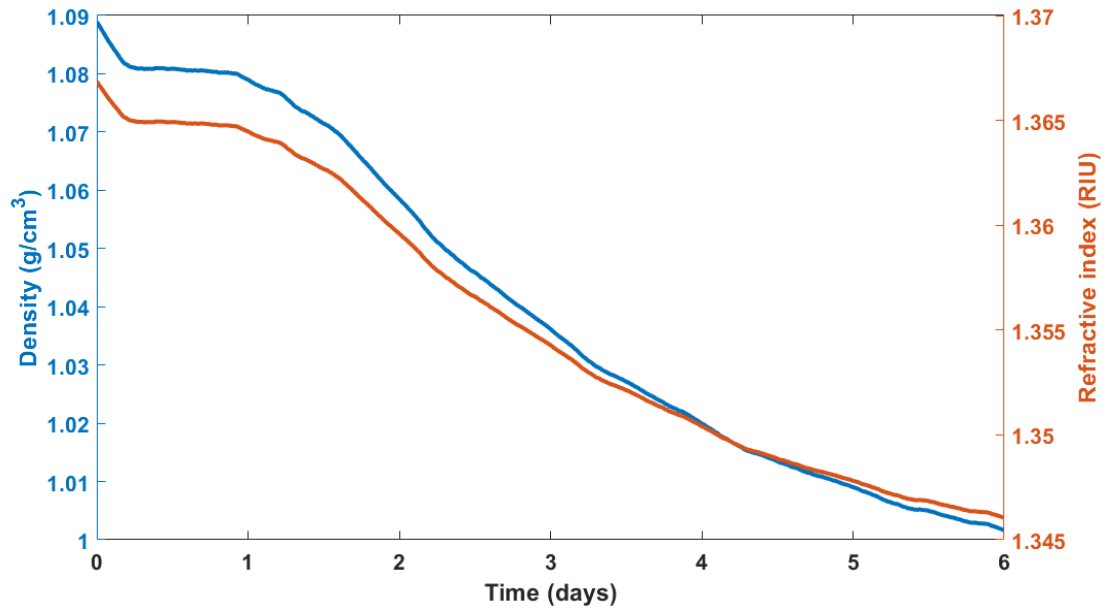
Figure 5.15 – Density and refractive index (calculated) evolution during the fermentation of white wine B.

In 6 days, the fermentation was completed and a turbidity from nearly 200 to a peak of 3000 FNMU was registered (Figure 5.16). Then, its value decreases to 1000 FNMU. Here the expected pattern for a fermentation was better defined than the one observed for the wwA. There is a small-time interval between the second and the third day where no data was collected. Contrary to the wwA monitoring, the wwB was not affected by the temperature oscillation in the electronics and optical system, therefore it is not observed the same periodic behavior as seen in wwA. In contrast, there was a higher variation in consecutive measurements mainly after the peak of fermentation is reached, which is an indication that a more variety of particle sizes was present here than the wwA.
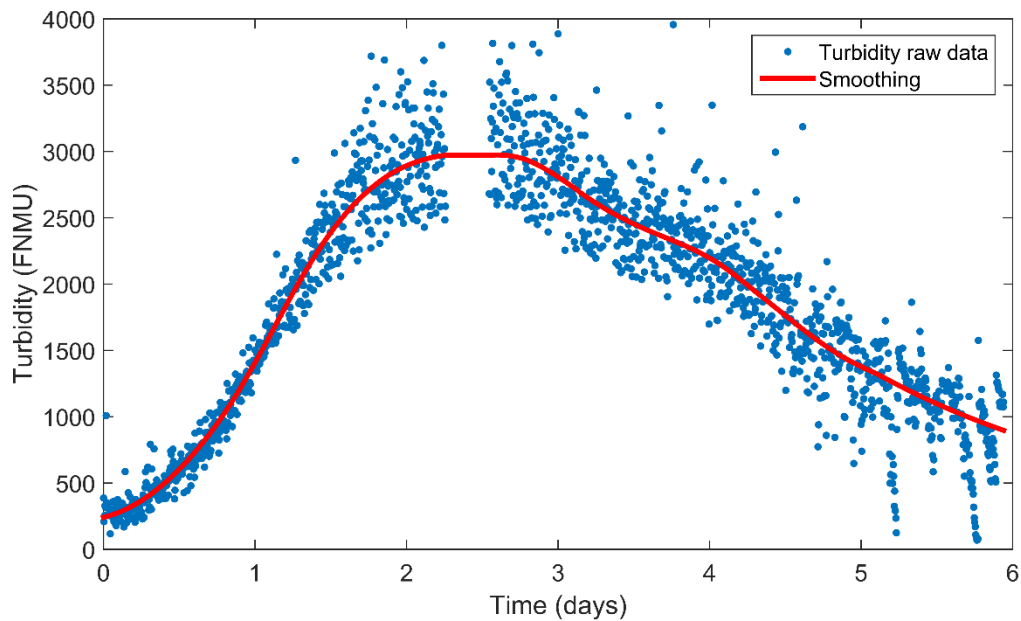
Figure 5.16 – Turbidity evolution during the fermentation of white wine B where a high variation in consecutive measurements is observed.

Other difference that was observed between the two measured white wines was the color. The color of wwB was less intense than the one observed in wwA, being almost a totally transparent liquid reaching an average higher value of 0.05 Abs units which is almost negligible since it's a value lower that the resolution of the sensor (Figure 5.17). Nevertheless, the fermentation behavior of this wine color seems to be closer to the pattern expected with an initial increase of color followed by its decrease in the end. The same high measurement variation observed in the turbidity behavior is also present in the color measurement, which can hinder the interpretation of the results with a behavior that is better understood by the help of the smoothing line.
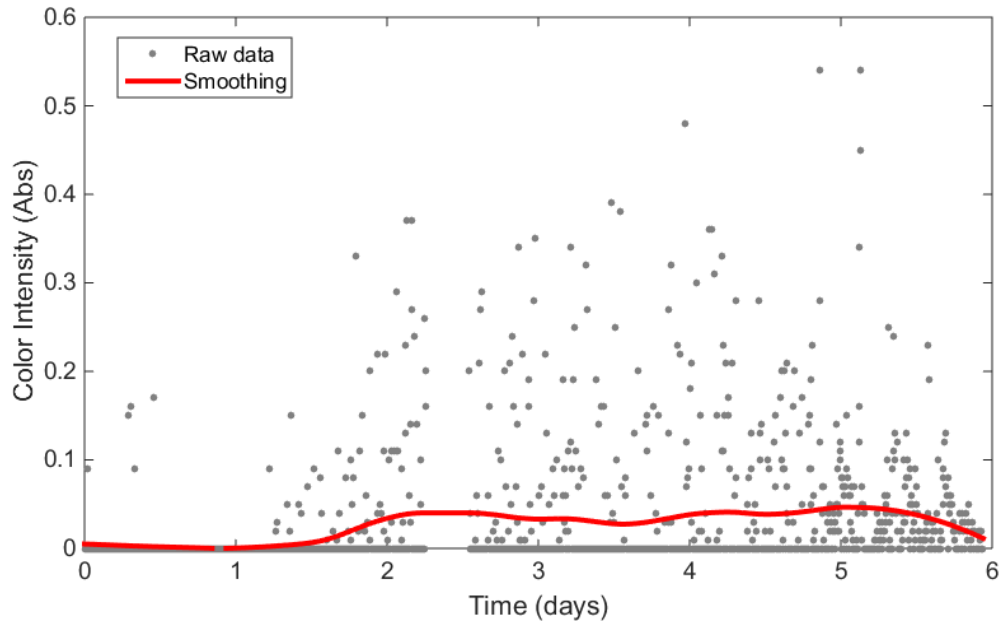
Figure 5.17 – Color intensity evolution during the fermentation of white wine B.

The color hue of this wine was also monitored but because of its low values, mainly zeros, it makes no sense to represent it since is impossible to have any distinct pattern. Color hue will be a more useful parameter when monitoring the rosé and red wines.

Figure 5.18 – Location of the sensor (in the left) for the monitoring of a white wine (in the right).

5.2.3. Rosé wine fermentation

A rosé wine case study was also performed to assess the behavior of the sensor, mainly the color measurement, due to its slightly more intense color than the one observed for the monitored

white wine. A rosé wine is a type of wine that incorporate some of the color of the grape skins by exposing them in contact with the juice, but not enough to classify it as a red wine. For this wine, high responsivity LED sensor (sensor version 4) was also used. Again, Watgrid density sensor will be used to estimate the refractive index value for compensation in the turbidity and color measurements.

The rosé wine monitored has taken a 27-day fermentation period, a value near of what is normally seen with a white wine. The density and refractive index behavior during the fermentation process can be visualized in Figure 5.19.
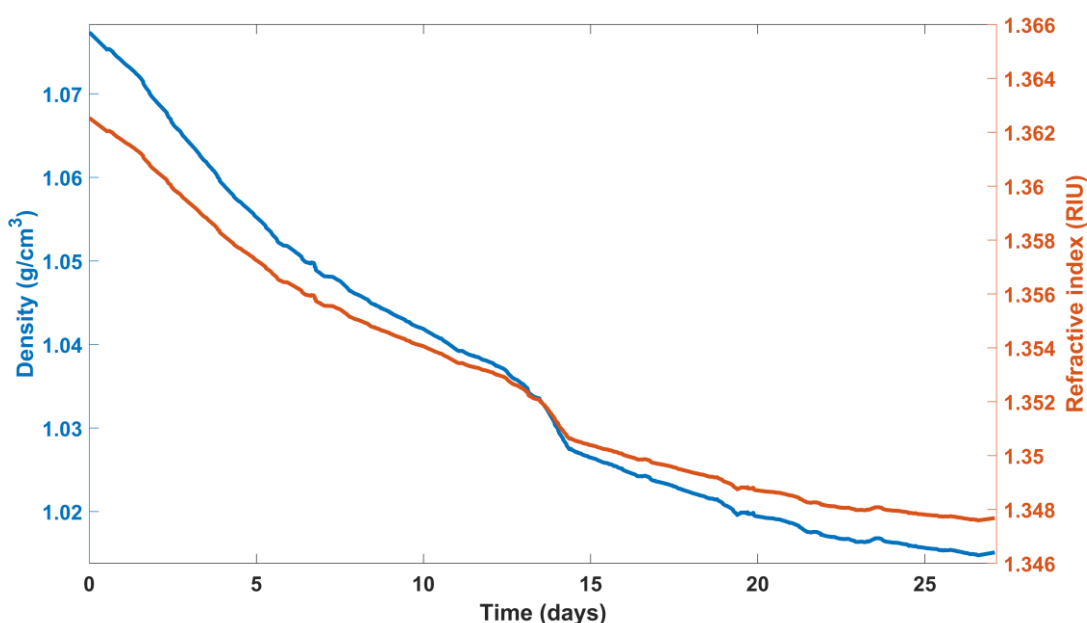
Figure 5.19 – Density and refractive index (calculated) evolution during the fermentation of rosé wine.

The turbidity evolution of this fermentation can be seen in Figure 5.20 together with its smoothing line. It has a halt in measuring by the 14th to the 15th day due to an operation performed by the winemaking staff on the wine which required the sensor removal.
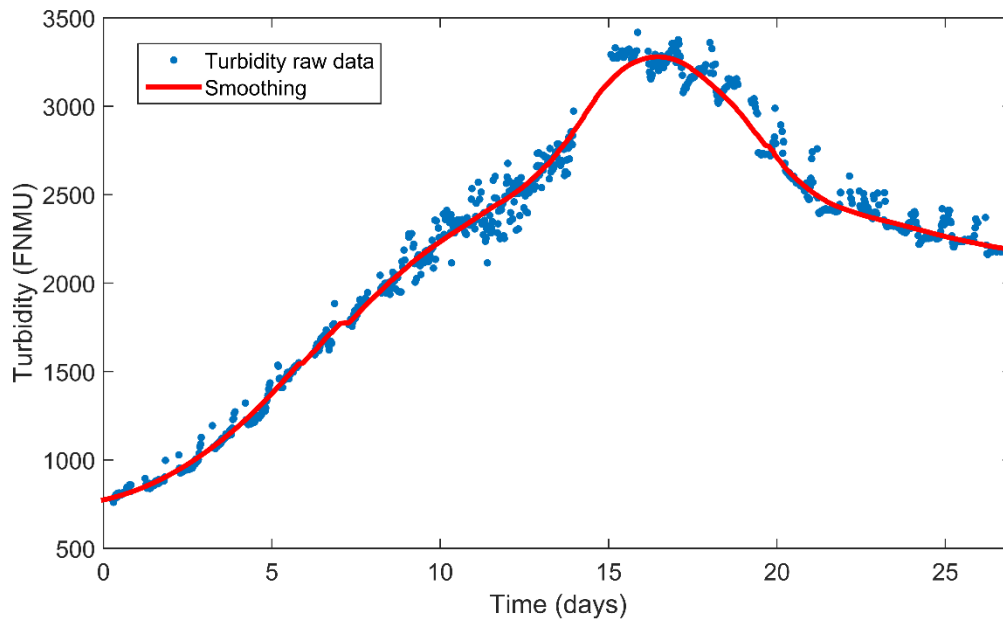
Figure 5.20 – Turbidity evolution during the fermentation of the rosé wine. An intervention was made in the 14th day which required the removal of the sensor.

The juice starts to have a low turbidity value of around 750 FNMU and immediately raises up until the 17th day reaching its peak with 3400 FNMU where the transformation rate of sugar into ethanol is at its maximum. It then decreases to around 2200 FNMU which indicates the decrease in the fermentation rating with less sugar being chemical transformed into ethanol. Again, this is the typical profile expected for a fermentation. The consecutive measurements present some fluctuation, mainly after the 8th day, indicating that at this point there is particles sizes with wider range of values but there is also some influence in temperature in the measurements.

The color evolution during the fermentation was also measured and can be observed in Figure 5.21.
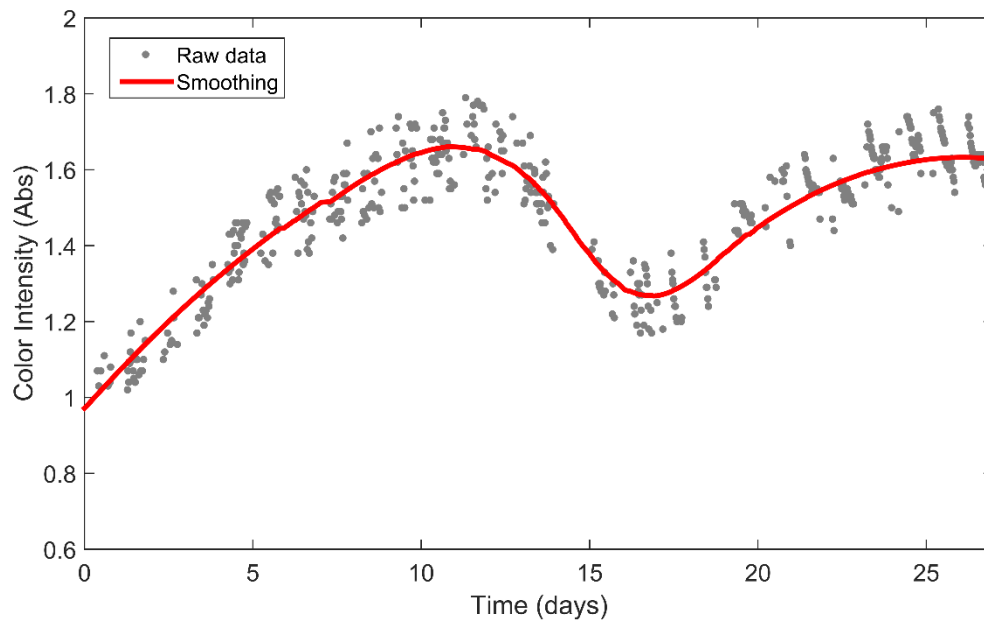
Figure 5.21 – Color intensity evolution during the fermentation of the rosé wine.

The behavior observed for the color during this fermentation is again close to the one expected in the literature with an initial increase due to the red pigments extraction, followed by a decrease due to the increase of alcohol. The perturbation caused in the system by the sensor removal could also be a source of a decrease of color due to the mechanical homogenization of the must and increased particles in suspension. A rise of color intensity is then observed due the further extraction of red pigments in the final process where fermentation almost ends. The stabilization of the juice after the intervention could also contribute to this rise. It is visible that the juice is more intense than the ones measured in the white wine with registered values between 1 to 1.8 Abs units.

The red pigments extraction mechanism can also be seen by analyzing the color hue evolution with the fermentation (Figure 5.22). The color hue value stays stable in the first 5 days. It is then followed by a slightly decrease as a consequence of the rising absorption of the green light which is an indicator of red pigments extraction. After the winemaking staff makes some operation to the wine, the color hue presents a lower value than the one observed before but it and then stabilizes to the initial values as observed with the increase. This could be explained by the increased ethanol which breaks down the red pigmentations, reducing the red absorption.
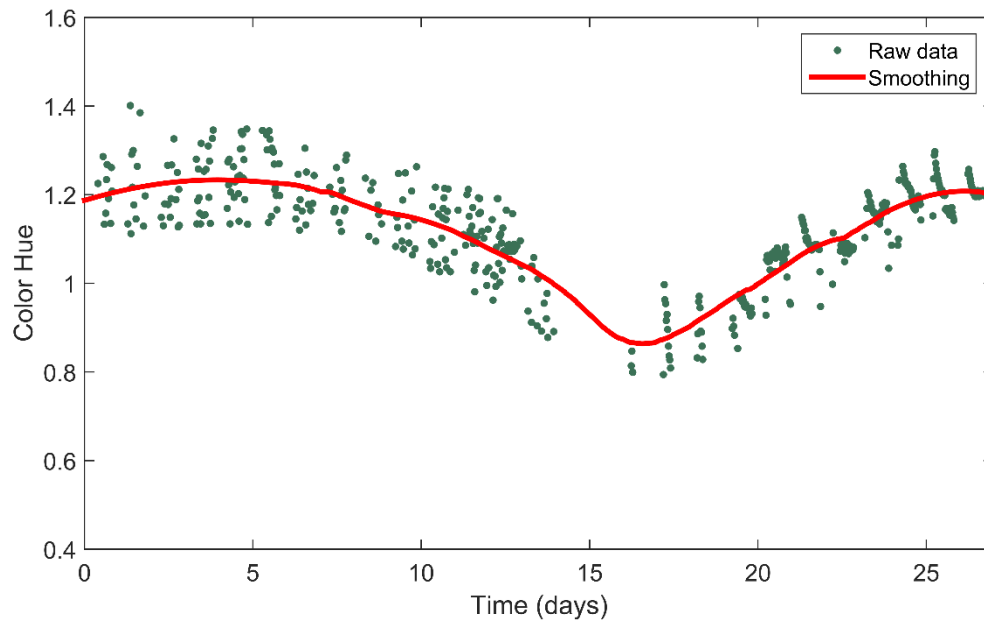
Figure 5.22 – Color hue evolution during the fermentation of the rosé wine.

5.2.4. Red wine fermentation

Two red wine case studies were performed with fermentation analysis performed in the same winemaker. However different sensor versions were used for each case. One of the sensors used was the peristaltic sensor (sensor version 5) to test its viability in an environment with higher expected turbidities than observed in the white and rosé wines. This test will be important to assess the likelihood of deposition of particles in the silicon tubing and in the measurement cell for longer periods than the ones observed in the press machine. It will also be analyzed the behavior of the sensor with high color intensities. For a better assessment, samples of wine were taken and measured in a spectrophotometer to a direct comparison of color intensity with the ones measured by the sensor. To identify this case, it will be named red wine A – rwA. The second case study will use the high responsivity LED sensor (sensor version 4) and therefore a direct comparison with the peristaltic sensor can be made. For identification, the name red wine B – rwB will be used. Both sensors will use the Watgrid density sensor to estimate the refractive index value for compensation in the turbidity and color measurement.

5.2.4.1. Red wine A (rwA)

The red wine A monitored by the peristaltic sensor has taken around a week fermentation period which is expected for a red wine fermentation. The sensor was only ready to be placed when the

fermentation was already occurring as can be seen in Figure 5.23 where the density value starts at a value of 1.065 g/cm^3.
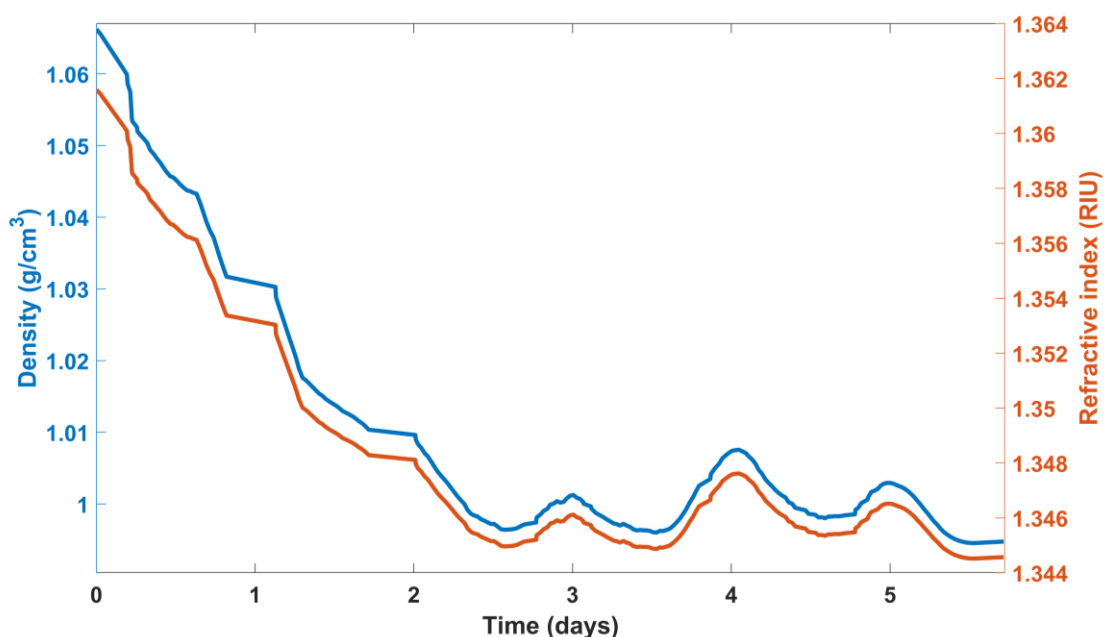
Figure 5.23 – Density and refractive index (calculated) evolution during the fermentation of red wine A.

During the measurement period, the sensor had to be removed several times to allow a rotating mechanism installed in the tank to mix the juice and therefore to promote a homogenous fermentation. The turbidity evolution observed during this period and its smoothing line can be seen in Figure 5.24. The gaps without measured values are related to the period time the sensor was removed.
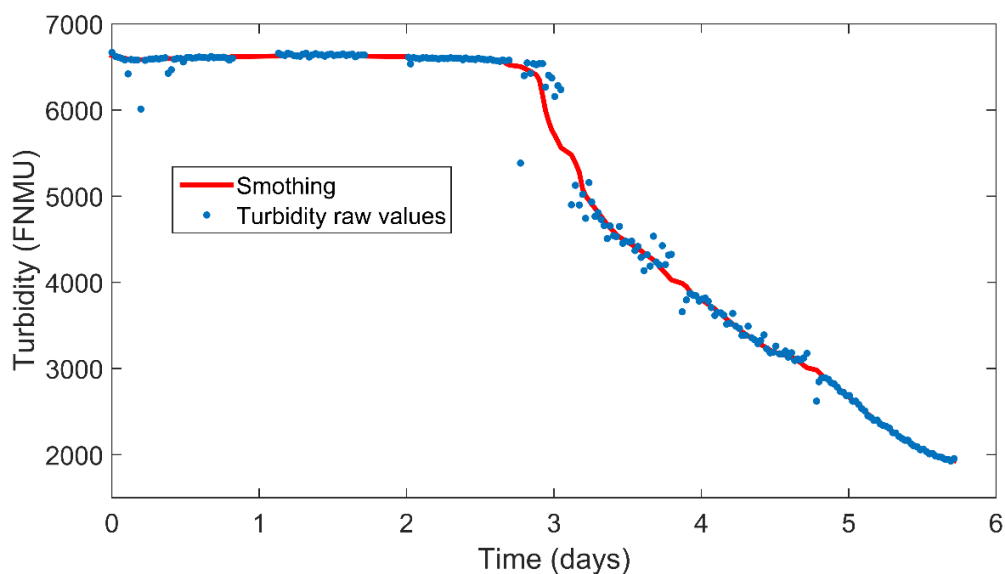
Figure 5.24 – Turbidity evolution during the fermentation of the red wine A.

As can be seen from the figure, due to the fermentation process being already occurring, the initial turbidity of the juice is already very high, with values around 6600 FNMU. The turbidity values obtained from the sensor show a stable behavior during its peak where the chemical conversion and CO_2 release are at its maximum. This stable behavior could be due to saturation on the measurement of turbidity due to the high color intensity present that will have direct influence on the turbidity parameter measurement. In the third day of measurement, the rate of the fermentation decreases and therefore the particles start to slowly set down, as seen by the decrease of the turbidity value. Since the turbidity variation in consecutive measurements is low, it is an indication that the juice particles sizes have a narrow range. When opened to assess the dirty accumulated in the sensor tube system, it was clearly visible the accumulation of particles, not only inside the measurement cell, but also on the silicon tubing. These particles were well stuck and didn't flow with the pumping liquid, demanding regular cleaning or substitution. The presence of these particles could lead to an increase of the parameters error due to obstruction of the optic system when measuring and thus this sensor does not seem so fit to be used regularly in these conditions. A filter located before the reception of the liquid may lower this problem.

Being a red wine, high intensity color is expected to be measured. The evolution of this parameter can be seen in Figure 5.25 together with its smoothing line. It is also presented in the figure measurements taken from centrifuged samples with a spectrophotometer (Shimadzu UV-2100) for comparison. Because of the spectrophotometer measurement range limited to 5 Abs units, dilutions of the juice were necessary to be made and the respective extrapolated value of the original concentration calculated.
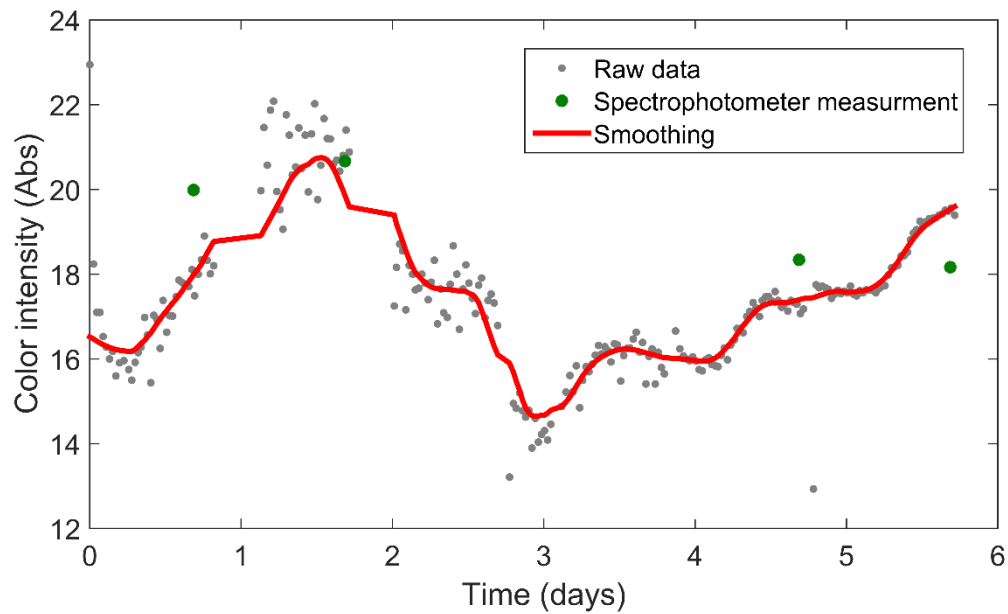
Figure 5.25 – Color intensity evolution during the fermentation of the red wine A with spectrophotometer samples measurement for comparison.

With the fermentation process already occurring when the sensor was placed, the initial measurements still capture the rise of the color intensity from 16 absorbance units until it reaches its peak with around 21 absorbance units in the first and half days of measurement. Then a decrease in the color is registered as expected due to the increase of the ethanol content. When the rate of the fermentation starts to lower with less ethanol being transformed from sugar, the color increases again with the continuation of the red pigments extraction from the skins of the grape, until it reaches the final value of 20 Abs units. The spectrophotometer measurements were not exactly the same obtained with the sensor but are near of what was measured. This discrepancy can be justified by the time of when the spectrophotometer measurements were taken. Due to the lack of availability of the spectrophotometer to be used in the same day as the sample was taken, the samples were saved in low temperatures to prevent further fermentation, and only measured its absorbance the day after the last sampled was acquired. Although in low temperature, some chemical reaction could still have happened and therefore changing the real absorbance value.

The red pigments extraction and their neutralization with the ethanol can also be seen with the interpretation of the color hue evolution with time as observed in Figure 5.26. The decrease in the initial value is a consequence of the increase of absorption of the green light due to the red pigments extraction. Then, the value increases due to the ethanol effect in the pigments, followed by its decrease with more pigments extraction in the final moments of the fermentation. Again, the

spectrophotometer measurements are near the ones measured by the sensor with the differences explained as before.
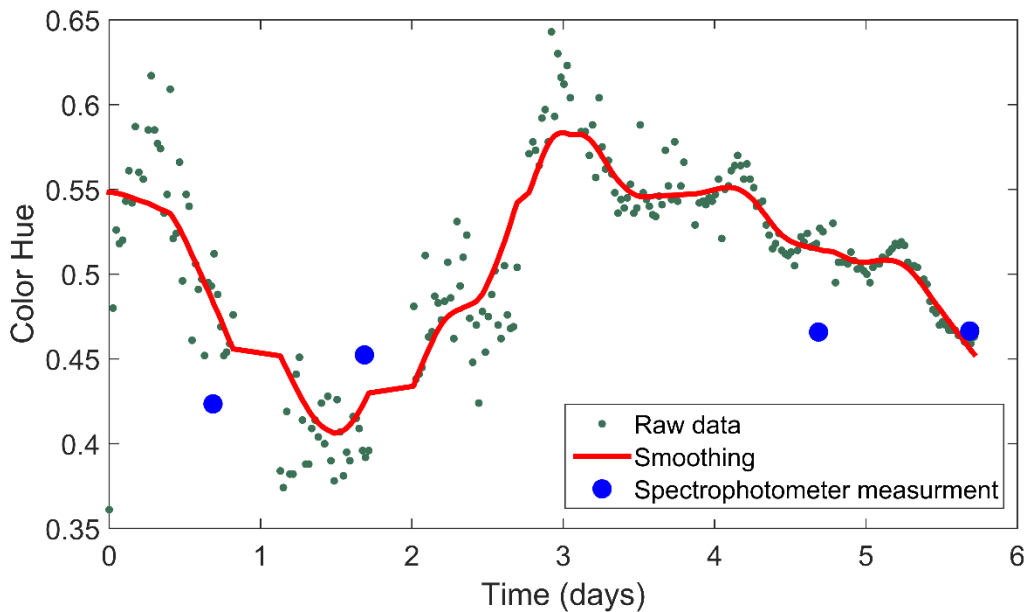
Figure 5.26 – Color hue evolution during the fermentation of the red wine A with spectrophotometer samples measurement for comparison.

5.2.4.2. Red wine B (rwB)

Red wine B was the second fermentation monitored in the same company as rwA but using the high responsivity LED sensor. Like rwA, the monitoring of this wine also started with the fermentation already in process with some days.
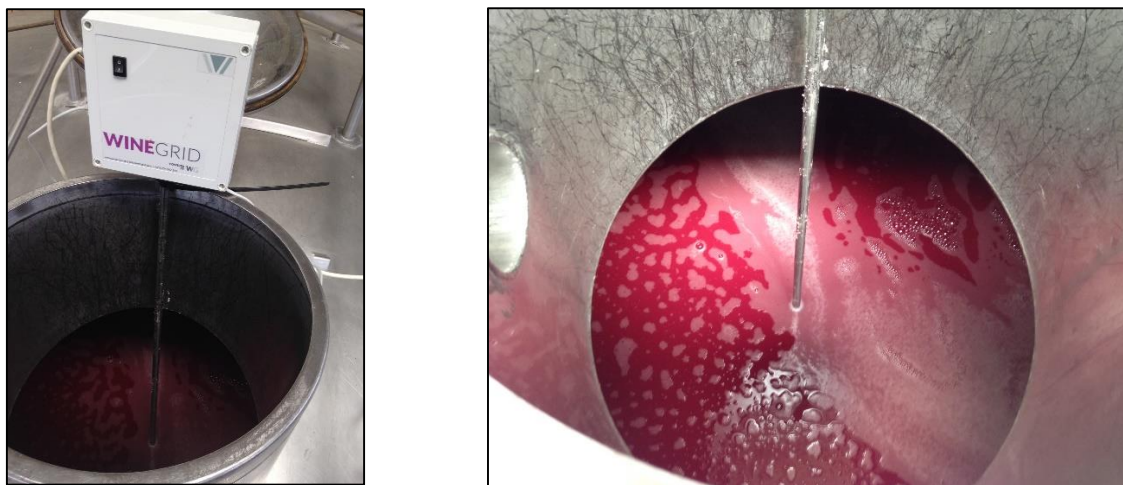
Figure 5.27 – Localization o the sensor (in the left) for the monitoring of a red wine (in the right).

The winemaking process of this fermentation was different to the one used in rwA, mainly due to the different temperature applied which was lower in this case. The first consequence of this is the slower fermentation reaction expected for this wine, which increases the time it needs to complete the full process. This fermentation has taken around two weeks for completion as can be observed in the density and refractive index behavior of Figure 5.28.
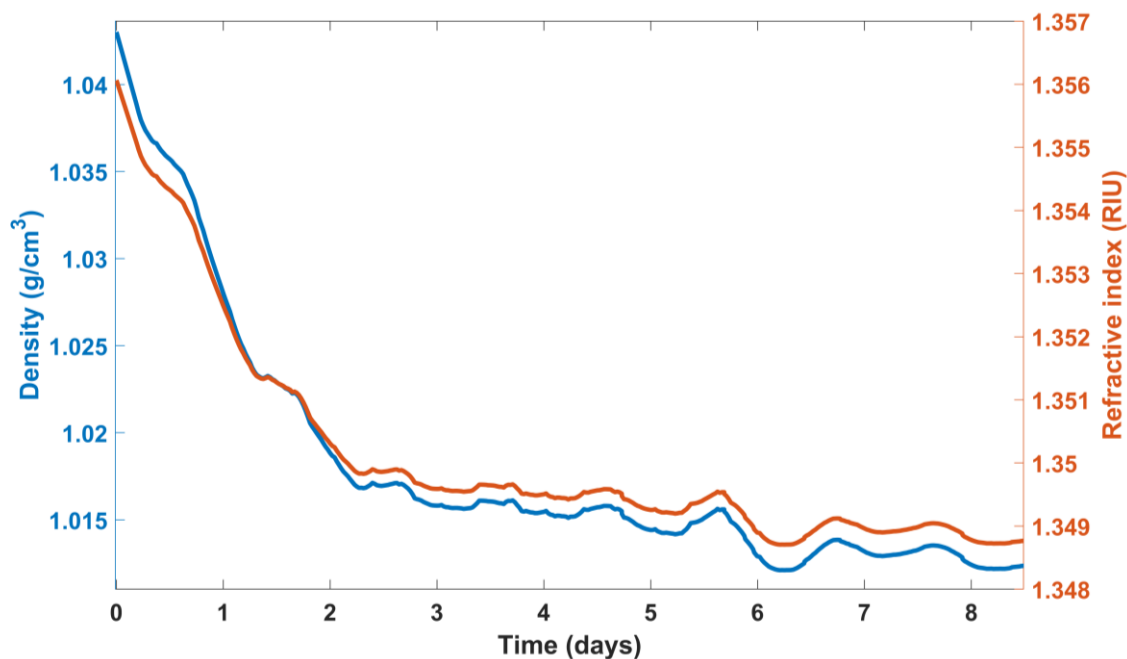
Figure 5.28 – Density and refractive index (calculated) evolution during the fermentation of red wine B.

The evolution of turbidity during this fermentation can be seen in Figure 5.29.
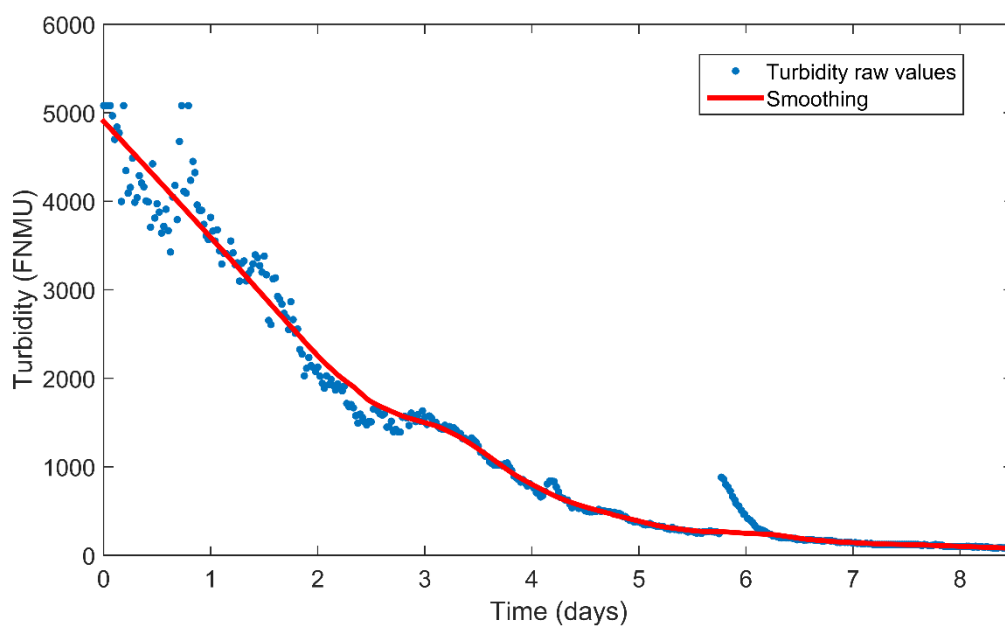
Figure 5.29 – Turbidity evolution during the fermentation of the red wine B.

The turbidity registered by the first measurements from the sensor reveled to be in its maximum values, which is an indication that the fermentation is in its full chemical rate. With a value of 5000 FNMU this process was slightly less turbid than the one registered from rwA. These first turbidity measurements also presented high variance, indicating more agitation in the juice. It begins to decrease gradually, with the fermentation approaching its final reactions. There is a slightly peak observed in the final of day 5 due to an intervention done to the wine, which must have spread some deposited particles, increasing temporarily the turbidity value. It then decreases again reaching a very low turbidity value towards zero.

The color intensity evolution during the fermentation period is presented in Figure 5.30.
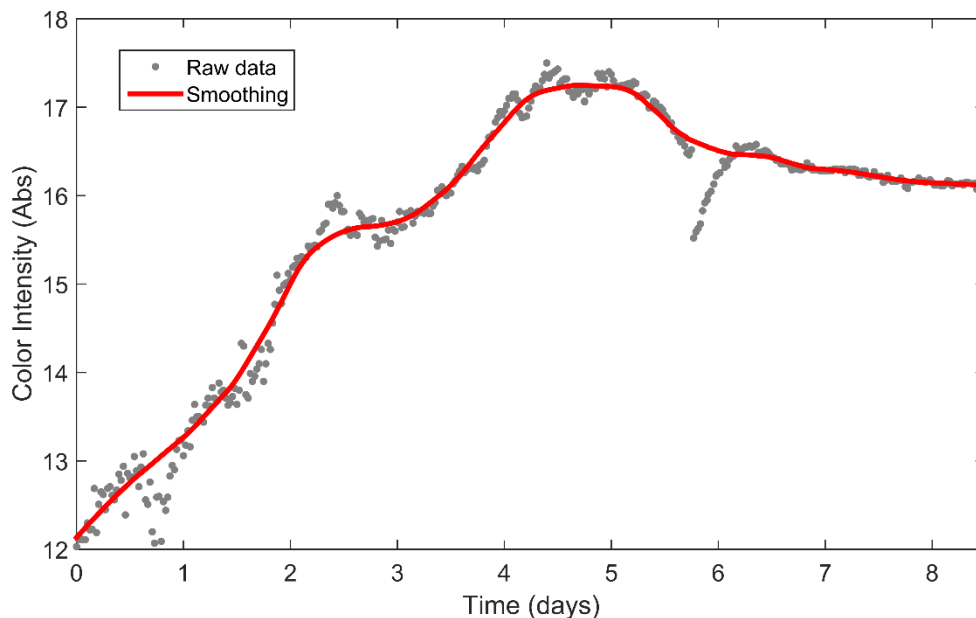
Figure 5.30 – Color intensity evolution during the fermentation of the red wine B.

As observed for rwA, the color intensity of the juice also has an increasing behavior from the start of the measurements by the sensor, reaching a peak in the fourth day. From here, the ethanol produced from the reaction starts to interact with the colored pigments decreasing its color intensity. Contrary to the observed in rwA, there wasn't an increase of color in the final moments of fermentation but instead it stabilized around the 16 Abs units. It is however noticed that the intervention on the wine produced a briefly decrease in the detected color that stabilized again to the behavior measured before. This intervention on the wine could have mixed part of the liquid containing less pigments, bringing it near to the sensor. This can be also verified by analyzing the color hue measurement evolution with its smoothing line in Figure 5.31. It has a decreasing behavior due to the red pigments extraction and then stabilizes around a value of 0.43.
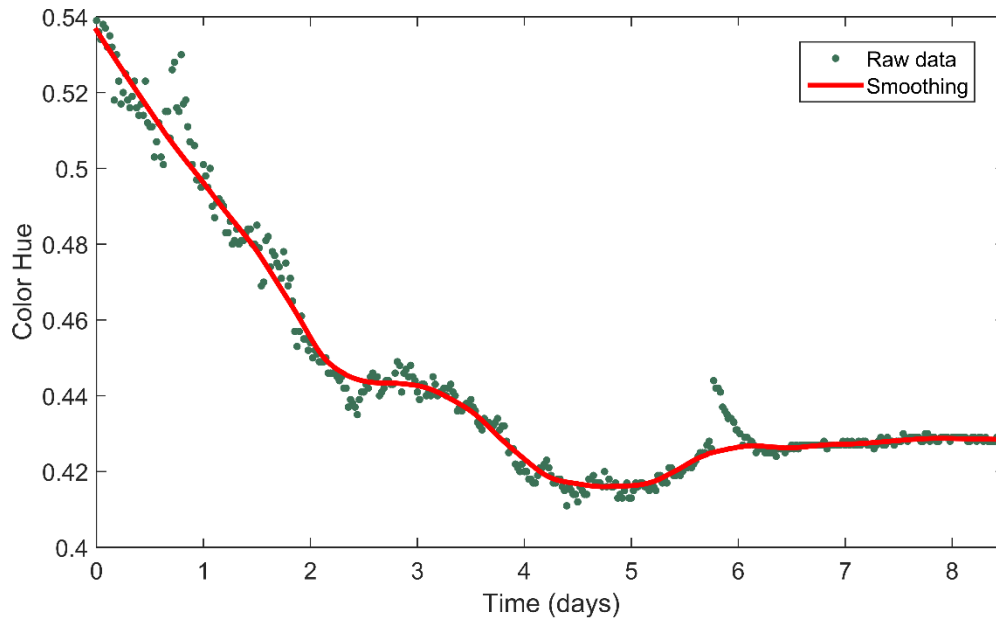
Figure 5.31 – Color hue evolution during the fermentation of the red wine B.

5.2.5. Port wine ageing

The last case study here presented was the monitoring of the ageing of a red Port wine. The sensor used in this case was the laser sensor (sensor version 3) which was the version developed and ready to be used in that moment for this task.

Figure 5.32 – Localization of the sensor for the monitoring of the Port wine ageing.

As mentioned in chapter 3.2.2.4, the laser sensor is developed with laser diodes that have an output light power with high dependence on temperature. This dependence will be visible in the

measurements done during this process with clear daily peaks. These peaks are due to the oscillation in temperature occurring in a typical day. The ageing of this wine has taken over 8 months in which several chemical reactions are occurring to improve its aroma, flavor, taste, and color. No refractive index variation will occur during this process and therefore a dedicated sensor for posterior compensation is not required. The evolution of turbidity measurements and its smoothing line during this process can be observed in Figure 5.33.
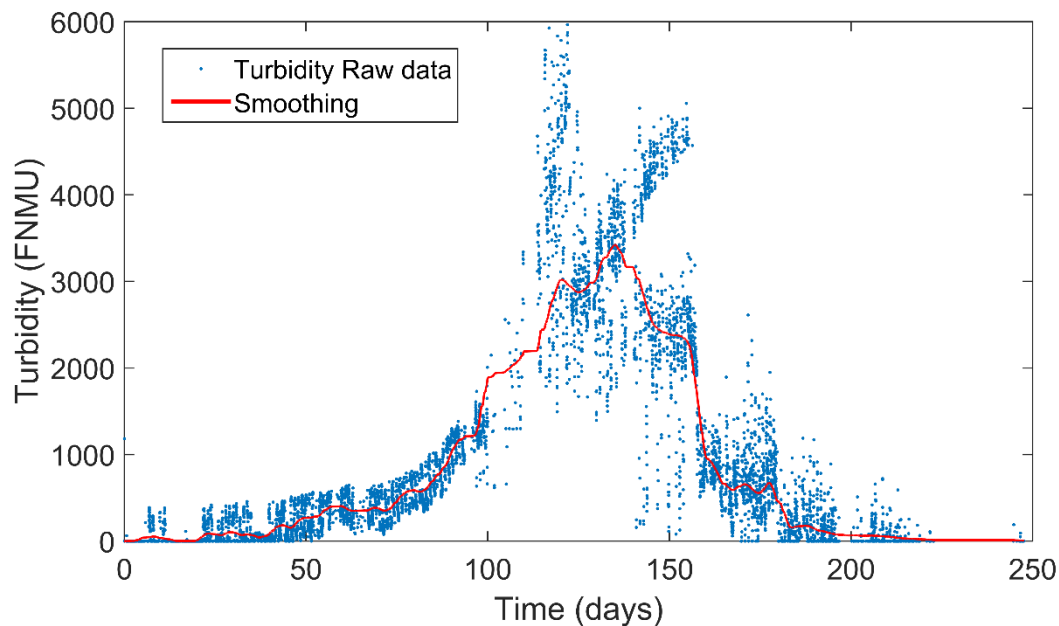
Figure 5.33 – Turbidity evolution during the ageing of a red Port wine.

As expected from a young wine, the turbidity measured from the start of the maturing process is very low with all the remaining particles deposited in the bottom of the tank. It is also important to notice that the temperature of the wine in the start was around 20 °C with the ageing starting in May, middle of spring. When temperature starts to rise, turbidity also starts to be higher, reaching its peak (4000 FNMU) after 4 months of ageing, this is, in the middle of the summer where the maximum of 27 °C was reached. A quick intervention was done to the wine in the 130th day of ageing, which required taking out the sensor from the tank, but it was not noticed any change on turbidity. After reaching this peak, the temperature and consequently the turbidity starts to fall by winter. Both temperature and turbidity returns to the low values observed in the beginning of the ageing that has its process finished in the middle of January.

The other monitored parameter was the color which was expected to be more intense in the end of this process. The evolution of the color intensity during the ageing of this wine is presented in Figure 5.34.
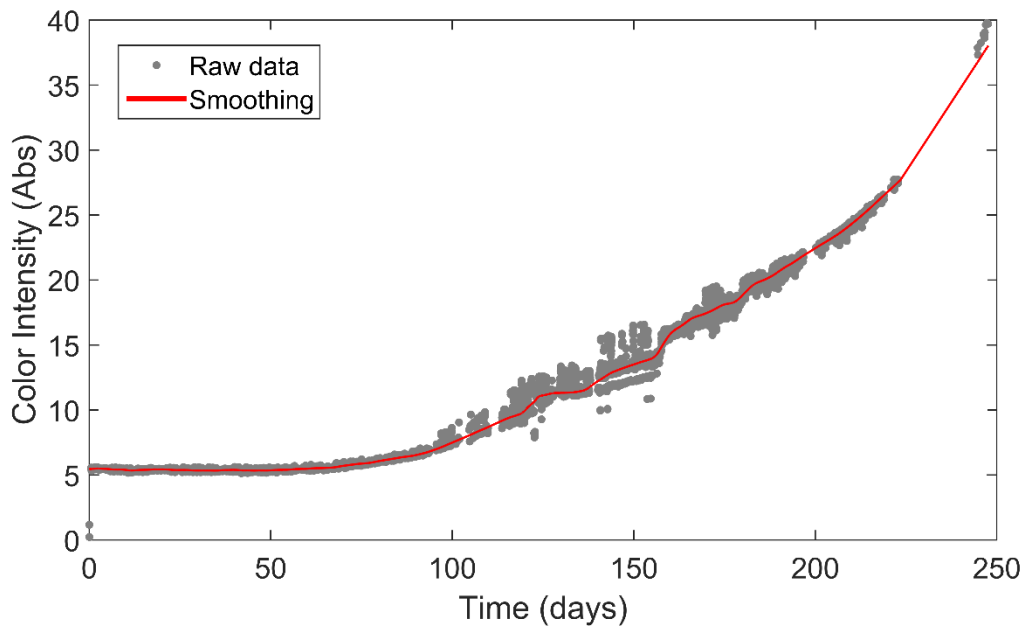
Figure 5.34 – Color intensity evolution during the ageing of a red Port wine.

The color intensity of the wine in the beginning of its ageing process presented a value of 6 absorbance units. This value stays stable in the first 60 days, starting then to increase in a growing rate to the end of the process. The last values measured presented a color intensity of 40 abs units. The intervention done to the wine in the 130th day of ageing have produced a slightly increase in the color intensity which can be explained by the presence of some particles deposited in the sensing head of the sensor during the first 4 months. When the sensor was removed from the wine, the particles were pulled out and the influence of turbidity less considered which increase higher the color intensity. This change is better viewed in the inspection of the color hue presented in Figure 5.35.
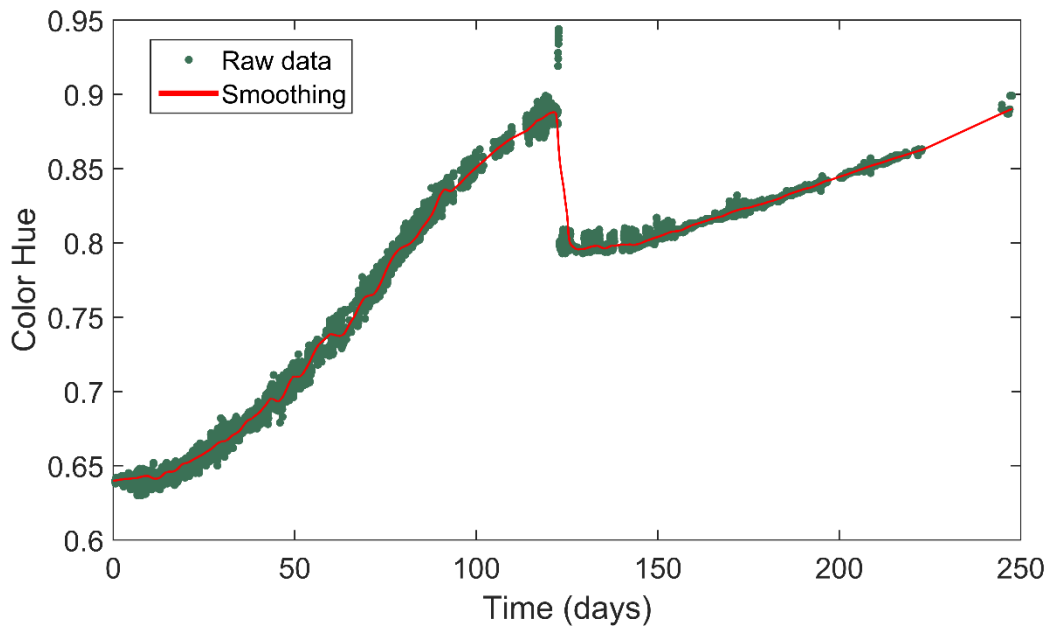
Figure 5.35 – Color hue evolution during the ageing of a red Port wine.

As seen, the color hue measured by the sensor starts with a value of about 0.65, which is a clear indication that the wine has a higher concentration of red pigments in comparison with other pigments color. As the ageing develops, the color hue value rises which is an indication that it is becoming more orange, as expected for an ageing wine. The intervention to the wine and the consequently removal of the sensor is here better visualized with an instantaneous drop from a value of 0.9 to 0.8, as explained before, due the probable pull out of particles from the sensing head. One can infer that the real increase rate for the color hue was lower than the one registered, which have a particle deposition effect added. The color hue then continues to rise reaching the 0.9 value in the end of the process.

5.3.   Cleaning detergent assessment

This case study will have the aim to analyze the possibility of using the refractive index sensor developed in chapter 3.3 to detect variations of refractive index in the cleaning detergent solution used by a specialized cleaning company in one of its machines. If this variation is noticeable using the sensor, a threshold value and a trigger alarm could be used to inform that the liquid has reached its limit. To assess this possibility, 5 samples from a solution used for cleaning were provided by a cleaning company. These samples were taken in a period of 12 weeks with an interval of 3 weeks between each one, where in the first one the cleaning solution is still new, and in the last, the solution is not proper to be used again.

Figure 5.37 – Cleaning detergent samples used for the study.

The developed refractive index sensor was used to measure the refractive index of the samples. For each solution 3 measurement points were taken. A constant temperature of 25 °C was used for the measurements. For comparison, the samples were also measured with the commercial benched refractometer Abbemat 200 from Anton Paar GmbH (589 nm) at the same temperature. Figure 5.38 a) presents the refractive index values measured by the developed sensor for each one of the cleaning solution samples, with a representation of the its error by the error bars. Its correlation with the values measured be the commercial refractometer can be seen in Figure 5.38 b).
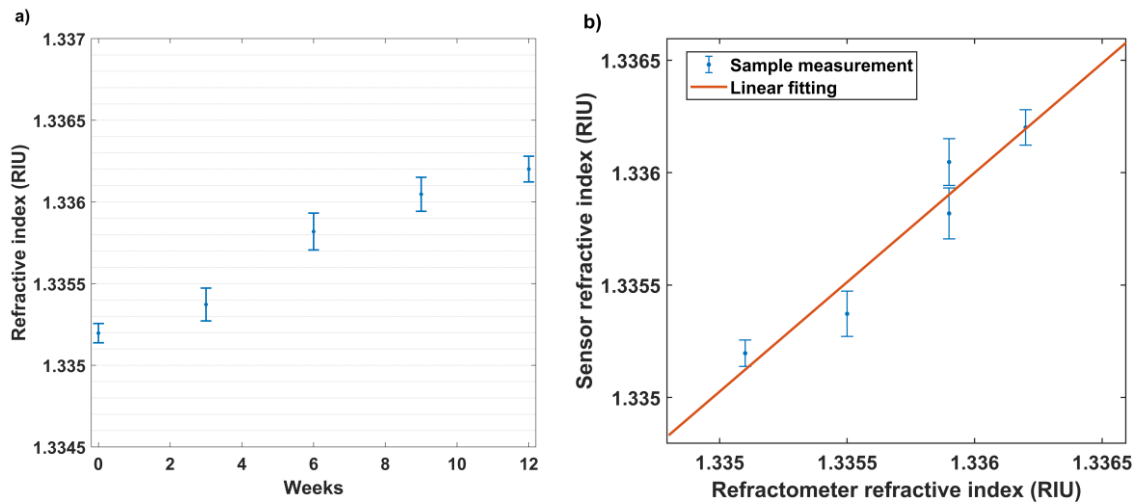
Figure 5.38 – a) Refractive index measurement of the 5 cleaning solutions using the developed sensor and representative of 12 weeks of usage. b) Comparison and correlation of the measurements done with the sensor with measurements done with a commercial refractometer.

It can be observed that with the usage of the cleaning solution with time, the refractive index increases from 1.3352, when its new, to 1.3362 when its improper for use. This demonstrates that

the sensor can be used to differentiate a cleaning solution that is not proper for use anymore. However due to the variation from a cleaning solution to the dirty solution being small, the measurement error is big enough that can cause trouble differentiating if the liquid is improper or not to be used in the last usages for cleaning. To mitigate this problem, high quantity of measurements and the calculation of its mean value will need to be performed.

A correlation analysis with the commercial refractive index was also performed in Figure 5.38 b) where the sensor values are plotted against the refractometer values. A linear fitting was performed to understand how the sensor measurements deviate from the ones measured using the bench refractometer. The fitting parameters can be observed in Table 5.1.

Table 5.1 – Parameters obtained from the linear fitting from Figure 5.38 b).

	Linear model: $y = ax + b$		R^2	RMSE
	a	b (RIU)		
Linear fitting	0.9735	0.03542	0.9284	1.33×10^{-04}

It is observable both graphically and by analyzing the fitting slope that the obtained values for the developed refractive index sensor that the measurements from the sensor are not totally correlated with the ones obtained from the refractometer, with the sensor presenting higher degree of uncertainty. It can be explained by the fact that with the increase of dirt observed in the samples progression, it will have an effect in the measurements as discussed in chapter 3.3.5 where the turbidity influence to the sensor is explored. Higher turbidity will increase the uncertainty of the sensor. It can be concluded however that the uncertainty here presented is small enough that a solution for automatic threshold detection can be implemented if intelligent analysis of the measurements is provided to avoid false positives. Modifications to the cleaning tanks, mainly perforations, will also need to be performed to insert and implement this sensor as an in-line configuration with access to the liquid, since they are usually closed environments.

6. Conclusion and future work

6.1. Conclusion

The importance of real-time measurements of many of today's fields of society is becoming more and more a standard necessity to maintain competitiveness and welfare of a highly, fast moving, interconnected world. The monitoring of liquids, that this work addresses, is not an exception. Industrial sectors that work with liquids and require constant feedback of its quality and characteristics will have highly interest in technology that can provide such response.

In this thesis, optical platforms for the monitoring of liquids were developed with turbidity, color and refractive index as the main parameters of interest in a broad industrial perspective where the developed sensors can be applied. To accomplished this, several conditions and goals were pursued, mainly:

- Its implementation as in-line sensing;
- Presenting low cost in its development and implementation being potential competitive against commercial applications;
- Presenting low complexity and high reproducibility in its development and implementation;
- Having the highest possible resolution by cost rate.

A review of the literature of both commercial and academic optic sensing proposals for liquids was made in chapter 2 with the purpose to identify the best technologies and adaptations possible to implement, trying to fulfill the proposed goals totally, or partially, in a balanced away. Two distinct physical ways of measurement were identified in literature that could be adapted and used in in-line sensing with low cost and low complexity development. The first one used, as its physical phenomena, the transmission and scattering of light between plastic optical fibers in a measurement gap cell to measure turbidity and color of a solution. The second one is based on the total internal reflection of light, with a broad range of incident angles provided by an optical fiber, between two distinct media with different refractive indices. By knowing the refractive index of one medium, then the other can be also measured and deduced. Both strategies were constructed in several prototypes and characterized its behavior in chapter 3, not only with variations of the parameter to measure, but also with variation of external parameters that could affect the

measuring setup. The characterization performed to the refractive index prototype showed resolutions similar to the ones obtained with common bench commercial refractometers, this is, with a value around 1×10^{-4} RIU and ranges up to 1.376 RIU and maintaining the temperature compensation feature, despite being an in-line/portable sensor. Different prototypes for the turbidity and color sensing setup were constructed due to the identification of some problems detected along the development and testing, mainly the low resolution in color observed with higher distances between the electronic board and the measurement head. This was rectified by changing the system components with different light sources and/or with better versions of components that appeared in the market during the testing phase. Therefore, a total of five different prototypes for turbidity and color measurement were presented, showing a diversity of behaviors to the measurement parameters that were characterized. In the end, two main versions were selected by presenting higher measurement quality which were the high responsivity LED sensor (sensor version 4) and the peristaltic sensor (sensor version 5). The first used a sensing head immersed in the liquid and the latter by circulating the liquid throw a measurement cell.

Turbidity and color are two parameters that share similar behavior under the incidence of light, however different wavelengths and physical techniques of detection will present for each parameter dissimilar behaviors detected. This characteristic was taken as an advantage to develop algorithms that could analyze multidimensional measurement data obtained by the usage of the different physical techniques and components with different wavelengths at once, that will evaluate the differentiated characteristics observed with the variation of the parameters to measure. It is then possible to apply algorithms that search and complement the information of each dimension to achieve an answer that provides lower measurement error and higher confidence. In chapter 4 several algorithms were tested and compared by cross validation with the traditional method of measurement, the regression analysis. Machine learning methods included cluster techniques such as LDA and EMGM and artificial neural networks. Data Fusion was also used as an alternative methodology. Each one presented its advantages and disadvantages which ultimately will depend on the application. LDA and EMGM are more appropriate for well-known cyclic behaviors, presenting low measurement error in color when a great amount of training data is available. ANN and data fusion revealed to be the best alternatives to use in a random and more uncontrolled measurement environment. With its lower turbidity error, the data fusion technique was chosen to be used as the algorithm of measurement. Compared to the regression analysis, this methodology decreases the error by a third in turbidity, while maintaining the same error for color.

Measurement accuracy below 5% is guaranteed for the chosen sensor versions used in the case studies having a measurement range up to 8000 NTU.

Case studies were carried on with the application of the turbidity and color sensor in the monitoring of grape pressing, wine fermentation and wine ageing as presented in chapter 5. The behavior of the parameters measured and obtained in each case study were the ones expected theoretically as presented in the introductory section of the chapter 5. In the pressing phase, it was possible to control the color of the finalized juice by proceeding with higher or lower number of pressing cycles. The color of the juice will have an impact in the final wine produced and therefore it is a parameter with high interest. During the fermentation, a very meticulous analysis of the chemical behavior during the transformation of sugar to alcohol can be performed with the proposed sensors, where color and the turbidity of the juice can be monitored as it is fermented. This information could be stored as historical records which can be access by the wine producer for future reference to develop improvements and optimizations, make experimentation with different procedures and identify faults, not only in the process of wine fermentation but also during the pressing and wine ageing. These parameters can also be used to identify fermentation stalling and prevent possible material and monetary losses since the sensor presents real-time capabilities. Measurement of color is also an interesting parameter to analyze during the wine ageing stage since it could be used by the wine producer to identify the ideal moment to proceed for bottling without the necessity of gather samples.

A different case study to the ones stated using the turbidity and color sensor was performed for the evaluation of a detergent solution used in industrial cleaning machines. The goal was to test if it was possible to identify its point of deterioration where a possible trigger alarm could notify the cleaning company to proceed with its substitution. The refractive index sensor developed was used to identify this threshold value. With the repeated usage of the cleaning solution for several times, the dirty agents will change the refractive index of the solution and therefore it is possible to determinate when it is not in conditions to be used again.

Overall, the developed sensors showed that they could be an asset with added value to the manufacturing companies that require constant monitoring of liquids.

6.2.   Future work

The color and turbidity sensor developed in this work have shown to be very promising for several applications and presenting low costs, yet several improvements could still be implemented. Since the sensor is refractive index dependent, the integration of the refractive index sensor also developed with the color and turbidity sensor seems to be the logical step to do. However, to test it with industrial case studies like the one presented for the wine making, it will require that refractive sensor is made with food compliant materials and can sustain high pressures up to 200 Kpa, which will increase the costs due to its high end materials requirement. It will also increase the calibration complexity of the sensor and here the use of the Artificial Neural Network algorithm developed and presented in chapter 4 reveals all its advantages where the increase of parameters of the sensor to measure will maintain the same complexity of calibration since the algorithm will learn from labelled calibration solutions and calibrate for itself without the introductions of complex physical models and correlations.

Other approach that could be taken is the total removal of the refractive index influence to the color and turbidity sensor by the introduction of lens to the optical system that could make the light collimated passing through the measuring cell with parallel rays, instead of the direct conic configuration created by the optical fiber. This will increase the complexity of the optical system since precise position of the lens would be necessary, but it prevents the necessity of refractive index compensation to the system in the applications where its variation is present.

As seen with the usage of Lasers in the color and turbidity version 3 sensor, its use presents advantages in relation to the LEDs since they use single spectral lines instead of a broad wavelength spectrum, which prevents the need of color compensation, if the intended measurement of absorption is a single wavelength value. However, this advantage can only be effective if a temperature control system is implemented. The output of the laser diodes used in the laser based sensor is highly dependent on their own temperature. Therefore, Peltier cooling modules could be implemented to guarantee that the laser measurement is performed at a constant temperature. Even if the LEDs present less dependency with temperature, they could also benefit from a cooling system to prevent light fluctuations, as detected with the long sensor versions when measuring low light intensities.

Other possible improvements could be the introduction of other wavelengths to the color and turbidity sensor such as the 320 nm and 280 nm spectral lines, which have interest mainly in the

wine production industry due to its value being an indication phenols and flavonoids concentration. However, better detectors adapted for this range of values will need to be used as the ones used for the sensors developed have low responsivity to ultraviolet light. A full visible to near infrared spectrum could also be perused using low cost components such as the holographic diffraction gratings and linear CCD detectors this will however increase the complexity of the optical system.

In relation to the refractive index sensor, an optimization to the algorithm used for the determination of the necessary key points of the profile and its reconstruction on the cloud server could be performed which would select better key points representative of the profile and consequently a better reconstruction of it. Ultimately, the best case scenario would be the determination of the pixel position in the sensor microprocessor itself, that would not require the profile reconstruction in the cloud server. However, this determination needs to be computational light and robust to prevent measurement anomalies because of the limitations of the microprocessor.

The refractive index sensor can also be used as a transducer for other sensing mechanisms by serving as a deposition platform for molecular imprinted polymers (MIPs) in the prism surface. These MIPs are designed to be highly selective to certain biological or chemical agents and can its refractive index value is modified when in presence of such agents due to the chemical links performed. Depending on the concentration of these agents, higher shifts in the refractive index value would occur in the MIP. This refractive index shift could then be detected by the configuration here presented for the refractive index sensor.

In conclusion, with the crescent interest and development of low cost IOT sensing technologies, a broad range of challenges and opportunities will surely be available for the massification and implementation of this technology in the most diverse applications of everyday life, which brings a very exciting future perspective.

References

A.KRÜSS Optronic GmbH (2016) 'Refractometers - Professional Solutions for Every Field of Application (V 4.2)'. Available at: https://www.labunlimited.com/CAR1/CAR1-LAB/pdf/kruss-refractometer-4-2.pdf (Accessed: 13 August 2019).

Abdi, H. and Williams, L. J. (2010) 'Principal component analysis', *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), pp. 433–459. doi: 10.1002/wics.101.

Al-Qazwini, Y., Noor, A. S. M., Al-Qazwini, Z., Yaacob, M. H., Harun, S. W. and Mahdi, M. A. (2016) 'Refractive index sensor based on SPR in symmetrically etched plastic optical fibers', *Sensors and Actuators A: Physical*. Elsevier B.V., 246, pp. 163–169. doi: 10.1016/j.sna.2016.04.064.

Alberto, N. J., Marques, C. A., Pinto, J. L. and Nogueira, R. N. (2010) 'Simultaneous strain and refractive index sensor based on a TFBG', in Santos, J. L. et al. (eds) *Proc European Workshop on Optical Fibre Sensors*, p. 765324. doi: 10.1117/12.866344.

Alpaydın, E. (2010) *Introduction to Machine Learning*. Second. Cambridge, Massachusetts: The MIT Press.

Alvarez-Rivera, G., Ballesteros-Vivas, D., Parada-Alfonso, F., Ibañez, E. and Cifuentes, A. (2019) 'Recent applications of high resolution mass spectrometry for the characterization of plant natural products', *TrAC Trends in Analytical Chemistry*, 112, pp. 87–101. doi: 10.1016/j.trac.2019.01.002.

Anderson, C. W. (2005) 'Turbidity (ver. 2.1)', in *U.S. Geological Survey Techniques of Water-Resources Investigations, book 9*, pp. 1–55. Available at: http://pubs.water.usgs.gov/twri9A/.

Anton Paar (2016a) 'Inline Refractometer'. Available at: https://www.anton-paar.com/?eID=documentsDownload&document=57820&L=0 (Accessed: 1 June 2019).

Anton Paar (2016b) 'The Universal Refractometer'. (Abbemat Series). Available at: https://wpstatic.idium.no/www.houm.no/2016/09/Abbemat_Refractometer.pdf (Accessed: 13 August 2019).

Anton Paar (2019) *Sensores de processo*. Available at: https://www.anton-paar.com/?eID=documentsDownload&document=60998&L=19 (Accessed: 13 August 2019).

Arrue, J., Jiménez, F., Aldabaldetreku, G., Durana, G., Zubia, J., Lomer, M. and Mateo, J. (2008) 'Analysis of the use of tapered graded-index polymer optical fibers for refractive-index Sensors.', *Optics express*, 16(21), pp. 16616–16631. doi: 10.1364/OE.16.016616.

Asheim, J., Kvittingen, E. V., Kvittingen, L. and Verley, R. (2014) 'A Simple, Small-Scale Lego Colorimeter with a Light-Emitting Diode (LED) Used as Detector', *Journal of Chemical Education*, 91(7), pp. 1037–1039. doi: 10.1021/ed400838n.

Azimirad, E., Haddadnia, J. and Izadipour, A. (2015) 'A Comprehensive Review of the Multi-Sensor Data Fusion Architectures', *Journal of Theoretical and Applied Information Technology*, 71(1), pp. 33–42.

Behler, J. (2015) 'Constructing high-dimensional neural network potentials: A tutorial review', *International Journal of Quantum Chemistry*, 115(16), pp. 1032–1050. doi: 10.1002/qua.24890.

Beres, C., de Nazaré, F. V. B., de Souza, N. C. C., Miguel, M. A. L. and Werneck, M. M. (2011) 'Tapered plastic optical fiber-based biosensor – Tests and application', *Biosensors and Bioelectronics*. Elsevier B.V., 30(1), pp. 328–332. doi: 10.1016/j.bios.2011.09.024.

Bilro, L., Prats, S. A., Pinto, J. L., Keizer, J. J. and Nogueira, R. N. (2010) 'Design and performance assessment of a plastic optical fibre-based sensor for measuring water turbidity', *Measurement Science and Technology*, 21(10), p. 107001. doi: 10.1088/0957-0233/21/10/107001.

Bilro, L., Alberto, N., Sa, L. M., Pinto, J. de L. and Nogueira, R. (2011) 'Analytical Analysis of Side-Polished Plastic Optical Fiber as Curvature and Refractive Index Sensor', *Journal of Lightwave Technology*, 29(6), pp. 864–870. doi: 10.1109/JLT.2011.2105462.

Bilro, L., Alberto, N., Pinto, J. L. and Nogueira, R. (2012) 'Optical Sensors Based on Plastic Fibers', *Sensors*, 12(12), pp. 12184–12207. doi: 10.3390/s120912184.

Binu, S., Mahadevan Pillai, V. P., Pradeepkumar, V., Padhy, B. B., Joseph, C. S. and Chandrasekaran, N. (2009) 'Fibre optic glucose sensor', *Materials Science and Engineering: C*. Elsevier B.V., 29(1), pp. 183–186. doi: 10.1016/j.msec.2008.06.007.

Bohren, C. F. and Huffman, D. R. (1983) *Absorption and scattering of light by small particles*, *New York*. doi: 10.1038/ncomms1111.

Burns, T. R. and Osborne, J. P. (2015) 'Loss of Pinot noir Wine Color and Polymeric Pigment after Malolactic Fermentation and Potential Causes', *American Journal of Enology and Viticulture*, 66(2), pp. 130–137. doi: 10.5344/ajev.2014.14061.

Cai, Z., Liu, F., Guo, T., Guan, B.-O., Peng, G.-D. and Albert, J. (2015) 'Evanescently coupled optical fiber refractometer based a tilted fiber Bragg grating and a D-shaped fiber', *Optics Express*, 23(16), p. 20971. doi: 10.1364/OE.23.020971.

Cai, Z., Guo, T., Liu, F., Guan, B.-O., Peng, G.-D. and Albert, J. (2015) 'Reflective refractometer based on strong optical coupling between a tilted fiber Bragg grating and a parallel D-shaped fiber', in Kalinowski, H. J., Fabris, J. L., and Bock, W. J. (eds) *Optics Express*, p. 96344I. doi: 10.1117/12.2194702.

Calvo-Zaragoza, J., Toselli, A. H. and Vidal, E. (2019) 'Hybrid hidden Markov models and artificial neural networks for handwritten music recognition in mensural notation', *Pattern Analysis and Applications*, 22(4), pp. 1573–1584. doi: 10.1007/s10044-019-00807-1.

Campbell, C. G., Laycak, D. T., Hoppes, W., Tran, N. T. and Shi, F. G. (2005) 'High concentration suspended sediment measurements using a continuous fiber optic in-stream transmissometer', *Journal of Hydrology*, 311(1–4), pp. 244–253. doi: 10.1016/j.jhydrol.2005.01.026.

Cao, J., Sun, T. and Grattan, K. T. V. (2014) 'Gold nanorod-based localized surface plasmon resonance biosensors: A review', *Sensors and Actuators B: Chemical*. Elsevier B.V., 195, pp. 332–351. doi: 10.1016/j.snb.2014.01.056.

Cao, W., Wang, X., Ming, Z. and Gao, J. (2018) 'A review on neural networks with random weights', *Neurocomputing*, 275, pp. 278–287. doi: 10.1016/j.neucom.2017.08.040.

Caucheteur, C., Shevchenko, Y., Shao, L.-Y., Wuilpart, M. and Albert, J. (2011) 'High resolution interrogation of tilted fiber grating SPR sensors from polarization properties measurement.', *Optics express*, 19(2), pp. 1656–1664. doi: 10.1364/OE.19.001656.

Cennamo, N., D'Agostino, G., Donà, A., Dacarro, G., Pallavicini, P., Pesavento, M. and Zeni, L. (2013) 'Localized Surface Plasmon Resonance with Five-Branched Gold Nanostars in a Plastic Optical Fiber for Bio-Chemical Sensor Implementation', *Sensors*, 13(11), pp. 14676–14686. doi: 10.3390/s131114676.

Cennamo, N., Massarotti, D., Galatus, R., Conte, L. and Zeni, L. (2013) 'Performance Comparison of Two Sensors Based on Surface Plasmon Resonance in a Plastic Optical Fiber', *Sensors*, 13(1), pp. 721–735. doi: 10.3390/s130100721.

Cennamo, N., Zuppella, P., Bacco, D., Corso, A. J., Pelizzo, M. G. and Zeni, L. (2016) 'SPR Sensor Platform Based on a Novel Metal Bilayer Applied on D-Shaped Plastic Optical Fibers for Refractive Index Measurements in the Range 1.38-1.42', *IEEE Sensors Journal*, 16(12), pp. 4822–4827. doi: 10.1109/JSEN.2016.2549271.

Chang, Y.-J., Chen, Y.-C., Kuo, H.-L. and Wei, P.-K. (2006) 'Nanofiber optic sensor based on the excitation of surface plasmon wave near fiber tip', *Journal of Biomedical Optics*, 11(1), p. 014032. doi: 10.1117/1.2165171.

Chen, C.-H. (2010) 'Novel D-type Fiber Optic Localized Plasmon Resonance Sensor Realized by Femtosecond Laser Engraving', *Journal of Laser Micro/Nanoengineering*, 5(1), pp. 1–5. doi: 10.2961/jlmn.2010.01.0001.

Chen, J., Shi, S., Su, R., Qi, W., Huang, R., Wang, M., Wang, L. and He, Z. (2015) 'Optimization and Application of Reflective LSPR Optical Fiber Biosensors Based on Silver Nanoparticles', *Sensors*, 15(6), pp. 12205–12217. doi: 10.3390/s150612205.

Chris Somers, T. and Evans, M. E. (1977) 'Spectral evaluation of young red wines: Anthocyanin equilibria, total phenolics, free and molecular SO2, "chemical age"', *Journal of the Science of Food and Agriculture*, 28(3), pp. 279–287. doi: 10.1002/jsfa.2740280311.

Christopher, C., Subrahmanyam, A. and Sai, V. V. R. (2018) 'Gold Sputtered U-Bent Plastic Optical Fiber Probes as SPR- and LSPR-Based Compact Plasmonic Sensors', *Plasmonics*. Plasmonics, 13(2), pp. 493–502. doi: 10.1007/s11468-017-0535-z.

Clippard, C. M., Hughes, W., Chohan, B. S. and Sykes, D. G. (2016) 'Construction and characterization of a compact, portable, low-cost colorimeter for the chemistry lab', *Journal of Chemical Education*, 93(7), pp. 1241–1248. doi: 10.1021/acs.jchemed.5b00729.

Coelho, L., de Almeida, J. M. M. M., Santos, J. L. . L., Ferreira, R. A. S., André, P. S. and Viegas, D. . (2015) 'Sensing Structure Based on Surface Plasmon Resonance in Chemically Etched Single Mode Optical Fibres', *Plasmonics*, 10(2), pp. 319–327. doi: 10.1007/s11468-014-9811-3.

Crespo, R., Cárcel, L. M., Pérez, M. A., Nevares, I. and del Álamo, M. (2010) 'Suitable at-line turbidity sensor for wine fermentation supervision', in *International Conference on Food Innovation*. Valencia, pp. 1–4.

De-Jun, F., Mao-Sen, Z., Liu, G., Xi-Lu, L. and Dong-Fang, J. (2014) 'D-Shaped Plastic Optical Fiber Sensor for Testing Refractive Index', *IEEE Sensors Journal*, 14(5), pp. 1673–1676. doi: 10.1109/JSEN.2014.2301911.

De-Jun, F., Guan-Xiu, L., Xi-Lu, L., Ming-Shun, J. and Qing-Mei, S. (2014) 'Refractive index sensor based on plastic optical fiber with tapered structure', *Applied Optics*, 53(10), p. 2007. doi: 10.1364/AO.53.002007.

Dhanabal, S. and Chandramathi, S. (2011) 'A Review of various k-Nearest Neighbor Query Processing Techniques', *International Journal of Computer Applications*, 31(7), pp. 14–22.

Dhiman, C. and Vishwakarma, D. K. (2019) 'A review of state-of-the-art techniques for abnormal human activity recognition', *Engineering Applications of Artificial Intelligence*, 77, pp. 21–45. doi: 10.1016/j.engappai.2018.08.014.

Domingues, M. F., Antunes, P., Alberto, N., Frias, R., Ferreira, R. A. S. and André, P. (2016) 'Cost effective refractive index sensor based on optical fiber micro cavities produced by the catastrophic fuse effect', *Measurement*. Elsevier Ltd, 77, pp. 265–268. doi: 10.1016/j.measurement.2015.07.031.

Duan, D.-W., Rao, Y.-J., Xu, L.-C., Zhu, T., Wu, D. and Yao, J. (2011) 'In-fiber Mach–Zehnder interferometer formed by large lateral offset fusion splicing for gases refractive index measurement with high sensitivity', *Sensors and Actuators B: Chemical*, 160(1), pp. 1198–1202. doi: 10.1016/j.snb.2011.09.048.

Duarte, D., Nogueira, R. N. and Bilro, L. B. (2019) 'Semi-supervised gaussian and t-distribution hybrid mixture model for water leak detection', *Measurement Science and Technology*. doi: 10.1088/1361-6501/ab3b48.

Duarte, D. P., Alberto, N. J., Bilro, L. and Nogueira, R. N. (2014) 'Theoretical modeling of an U-shaped SPR fiber sensor in 1550-nm spectral range for sensing applications', *Proceedings of SPIE - The International Society for Optical Engineering*, 9286, pp. 1–5. doi: 10.1117/12.2063724.

Duarte, D. P., Oliveira, N., Georgieva, P., Nogueira, R. N. and Bilro, L. (2015) 'Wine classification and turbidity measurement by clustering and regression models', in *ConfTele 2015*. Instituto de Telecomunicações, pp. 317–320.

Fan, C., Xiao, F., Li, Z. and Wang, J. (2018) 'Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review', *Energy and Buildings*, 159, pp. 296–308. doi: 10.1016/j.enbuild.2017.11.008.

Fang, Z., Zhang, M., Sun, Y. and Sun, J. (2006) 'How To Improve Bayberry (Myrica rubra Sieb. et Zucc.) Juice Color Quality: Effect of Juice Processing on Bayberry Anthocyanins and Polyphenolics', *Journal of Agricultural and Food Chemistry*, 54(1), pp. 99–106. doi: 10.1021/jf051943o.

Faragher, R. (2012) 'Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation [Lecture Notes]', *IEEE Signal Processing Magazine*, 29(5), pp. 128–132. doi: 10.1109/MSP.2012.2203621.

Ferreira, M. F. S., Statkiewicz-Barabach, G., Kowal, D., Mergo, P., Urbanczyk, W. and Frazão, O. (2017) 'Fabry-Perot cavity based on polymer FBG as refractive index sensor', *Optics Communications*. Elsevier B.V., 394(November 2016), pp. 37–40. doi: 10.1016/j.optcom.2017.03.011.

Figura, L. O. and Teixeira, A. a. (2007) 'On-line Sensing', in *Food Physics Physical Properties – Measurement and Applications*. 1st edn. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 445–468. doi: 10.1007/978-3-540-34194-9.

Fink, G. a. (2014) *Markov Models for Pattern Recognition*, *Advances in Computer Vision and Pattern Recognition*. London: Springer London. doi: 10.1007/978-1-4471-6308-4.

Fondriest Environmental Inc. (2014) *Turbidity, Total Suspended Solids and Water Clarity*, *Fundamentals of Environmental Measurements*. Available at: http://www.fondriest.com/environmental-measurements/parameters/water-quality/turbidity-total-suspended-solids-water-clarity/ (Accessed: 1 June 2019).

Gandhi, M., Chu, S., Senthilnathan, K., Babu, P., Nakkeeran, K. and Li, Q. (2019) 'Recent Advances in Plasmonic Sensor-Based Fiber Optic Probes for Biological Applications', *Applied Sciences*, 9(5), p. 949. doi: 10.3390/app9050949.

Gandul-Rojas, B., Gallardo-Guerrero, L., Roca, M. and Aparicio-Ruiz, R. (2013) 'Chromatographic Methodologies: Compounds for Olive Oil Color Issues', in Aparicio, R. and Harwood, J. (eds) *Handbook of Olive Oil Analysis and Properties*. Boston, MA: Springer US, pp. 219–254. doi: 10.1007/978-1-4614-7777-8.

Garcia, M. A. P., Vega, R. M., de la Torre Fernandez, C., de la Fuente, J. A. B. and Carcel, L. M. C. (2008) 'Full-range, true on-line turbidimeter based upon optical fibers for application in the wine industry', in *2008 IEEE Instrumentation and Measurement Technology Conference*. IEEE, pp. 130–134. doi: 10.1109/IMTC.2008.4547017.

Gasior, K., Martynkien, T. and Urbanczyk, W. (2014) 'Effect of constructional parameters on the performance of a surface plasmon resonance sensor based on a multimode polymer optical fiber', *Applied Optics*, 53(35), p. 8167. doi: 10.1364/AO.53.008167.

Georgieva, P., Mihaylova, L. and Jain, L. C. (2013) *Advances in Intelligent Signal Processing and Data Mining*. Edited by P. Georgieva, L. Mihaylova, and L. C. Jain. Berlin, Heidelberg: Springer Berlin Heidelberg (Studies in Computational Intelligence). doi: 10.1007/978-3-642-28696-4.

Ghamisi, P. *et al.* (2019) 'Multisource and Multitemporal Data Fusion in Remote Sensing: A Comprehensive Review of the State of the Art', *IEEE Geoscience and Remote Sensing Magazine*, 7(1), pp. 6–39. doi: 10.1109/MGRS.2018.2890023.

Gippel, C. J. (1995) 'Potential of turbidity monitoring for measuring the transport of suspended solids in streams', *Hydrological Processes*, 9(1), pp. 83–97. doi: 10.1002/hyp.3360090108.

Gli international (1992) 'GLI Method 2: Turbidity', *EPA Guidance Manual*. Great Lakes Instruments, pp. D1–D6. Available at: https://law.resource.org/pub/us/cfr/ibr/004/gli.method.2.2009.pdf.

Gómez-Míguez, M. J., Gómez-Míguez, M., Vicario, I. M. and Heredia, F. J. (2007) 'Assessment of colour and aroma in white wines vinifications: Effects of grape maturity and soil type', *Journal of Food Engineering*, 79(3), pp. 758–764. doi: 10.1016/j.jfoodeng.2006.02.038.

Govindan, G., Raj, S. G. and Sastikumar, D. (2009) 'Measurement of refractive index of liquids using fiber optic displacement sensors', *Journal of American Science*, 5(2), pp. 13–17.

Gowri, A. and Sai, V. V. R. (2016) 'Development of LSPR based U-bent plastic optical fiber sensors', *Sensors and Actuators, B: Chemical*. Elsevier B.V., 230, pp. 536–543. doi: 10.1016/j.snb.2016.02.074.

Gravina, R., Testa, G. and Bernini, R. (2009) 'Perfluorinated Plastic Optical Fiber Tapers for Evanescent Wave Sensing', *Sensors*, 9(12), pp. 10423–10433. doi: 10.3390/s91210423.

De Greeff, J., Belpaeme, T. and Bongard, J. (2015) 'Why robots should be social: Enhancing machine learning through social human-robot interaction', *PLoS ONE*, 10(9). doi: 10.1371/journal.pone.0138061.

Greenberg, A., Clesceri, L. and Eaton, A. (1992) 'APHA Method 2130: Standard Methods for the Examination of Water and Wastewater'. Edited by M. Franson. American Public Health Association, pp. 1–11.

Habernal, I., Ptácek, T. and Steinberger, J. (2013) *Sentiment analysis in czech social media using supervised machine learning*, *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*.

Hach Company (2005) 'Hach Method 10133 - Determination of Turbidty by Laser Nephelometry', pp. 1–8.

Hall, D. L. and Steinberg, A. (2000) 'Dirty secrets in multisensor data fusion', in *Proceedings of the National Symposium on Sensor Data Fusion (NSSDF)*. San Antonio, TX.

Hall, David L. and Llinas, James (2008) 'Multisensor Data Fusion', in Liggins, M. E., Llinas, J, and Hall, D L (eds) *Handbook of Multisensor Data Fusion: Theory and Practice*. 2nd edn. Boca Raton, FL: CRC Press, pp. 1–13.

Hanke, M., Halchenko, Y. O., Haxby, J. V. and Pollmann, S. (2010) 'Statistical learning analysis in neuroscience: Aiming for transparency', *Frontiers in Neuroscience*, 4(MAY), pp. 38–43. doi: 10.3389/neuro.01.007.2010.

Hernández, F., Sancho, J. V. and Pozo, O. J. (2005) 'Critical review of the application of liquid chromatography/mass spectrometry to the determination of pesticide residues in biological samples', *Analytical and Bioanalytical Chemistry*, 382(4), pp. 934–946. doi: 10.1007/s00216-005-3185-5.

Hlubina, P., Kadulova, M., Ciprian, D. and Sobota, J. (2014) 'Reflection-based fibre-optic refractive index sensor using surface plasmon resonance', *Journal of the European Optical Society: Rapid Publications*, 9, p. 14033. doi: 10.2971/jeos.2014.14033.

Homola, J., Yee, S. S. and Gauglitz, G. (1999) 'Surface plasmon resonance sensors: review', *Sensors and Actuators B: Chemical*, 54(1–2), pp. 3–15. doi: 10.1016/S0925-4005(98)00321-9.

Hu, X., Pun, C.-F. J., Tam, H.-Y., Mégret, P. and Caucheteur, C. (2014a) 'Highly reflective Bragg gratings in slightly etched step-index polymer optical fiber', *Optics Express*, 22(15), p. 18807. doi: 10.1364/OE.22.018807.

Hu, X., Pun, C.-F. J., Tam, H.-Y., Mégret, P. and Caucheteur, C. (2014b) 'Tilted Bragg gratings in step-index polymer optical fiber', *Optics Letters*, 39(24), p. 6835. doi: 10.1364/OL.39.006835.

Hu, X., Mégret, P. and Caucheteur, C. (2015) 'Surface plasmon excitation at near-infrared wavelengths in polymer optical fibers', *Optics Letters*, 40(17), p. 3998. doi: 10.1364/OL.40.003998.

Huadong Wu, Siegel, M., Stiefelhagen, R. and Jie Yang (2002) 'Sensor fusion using Dempster-Shafer theory [for context-aware HCI]', in *IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No.00CH37276)*. IEEE, pp. 7–12. doi: 10.1109/IMTC.2002.1006807.

Hussain, I., Ahamad, K. and Nath, P. (2016) 'Water turbidity sensing using a smartphone', *RSC Adv.*, 6(27), pp. 22374–22382. doi: 10.1039/C6RA02483A.

Hydrolab (2018) '4-Beam Turbidity', *OTT Technical White Paper*, pp. 1–2. Available at: http://www.esonetyellowpages.com/datasheets/4beam_1254491289.pdf (Accessed: 1 June 2019).

Iadicicco, A., Campopiano, S., Cutolo, A., Giordano, M. and Cusano, A. (2005) 'Refractive index sensor based on microstructured fiber Bragg grating', *IEEE Photonics Technology Letters*, 17(6), pp. 1250–1252. doi: 10.1109/LPT.2005.846570.

Iadicicco, A., Cusano, A., Campopiano, S., Cutolo, A. and Giordano, M. (2005) 'Thinned fiber Bragg gratings as refractive index sensors', *IEEE Sensors Journal*, 5(6), pp. 1288–1295. doi: 10.1109/JSEN.2005.859288.

International Organization for Standardization (1999) 'Water quality—Determination of turbidity (ISO 7027:1999)', pp. 1–10.

Ishigure, T., Nihei, E. and Koike, Y. (1996) 'Optimum refractive-index profile of the graded-index polymer optical fiber, toward gigabit data links', *Applied Optics*, 35(12), p. 2048. doi: 10.1364/AO.35.002048.

Jackson, M. G., Timberlake, C. F., Bridle, P. and Vallis, L. (1978) 'Red wine quality: Correlations between colour, aroma and flavour and pigment and other parameters of young Beaujolais', *Journal of the Science of Food and Agriculture*, 29(8), pp. 715–727. doi: 10.1002/jsfa.2740290810.

Jacobson, J. L. (2006a) 'Berry to Bottle', in *Introduction to Wine Laboratory Practices and Procedures*. New York, NY: Springer, pp. 83–117. doi: 10.1007/0-387-25120-0.

Jacobson, J. L. (2006b) 'Maturation Matters', in Springer (ed.) *Introduction to wine laboratory practices and procedures*. New York, NY, pp. 200–230. doi: 10.1007/0-387-25120-0.

Jacobson, J. L. (2006c) 'Vineyard to Harvest', in Springer (ed.) *Introduction to wine laboratory practices and procedures*. New York, NY, pp. 118–161. doi: 10.1007/0-387-25120-0.

Jain, A. K. (2010) 'Data clustering: 50 years beyond K-means', *Pattern Recognition Letters*. Elsevier B.V., 31(8), pp. 651–666. doi: 10.1016/j.patrec.2009.09.011.

Jasim, A. A., Hayashi, N., Harun, S. W., Ahmad, H., Penny, R., Mizuno, Y. and Nakamura, K. (2014) 'Refractive index and strain sensing using inline Mach–Zehnder interferometer comprising perfluorinated graded-index plastic optical fiber', *Sensors and Actuators A: Physical*. Elsevier B.V., 219, pp. 94–99. doi: 10.1016/j.sna.2014.07.018.

Ji, W. Bin, Liu, H. H., Tjin, S. C., Chow, K. K. and Lim, A. (2012) 'Ultrahigh Sensitivity Refractive Index Sensor Based on Optical Microfiber', *IEEE Photonics Technology Letters*, 24(20), pp. 1872–1874. doi: 10.1109/LPT.2012.2217738.

Jiang, S. *et al.* (2017) 'A novel U-bent plastic optical fibre local surface plasmon resonance sensor based on a graphene and silver nanoparticle hybrid structure', *Journal of Physics D: Applied Physics*. IOP Publishing, 50(16), p. 165105. doi: 10.1088/1361-6463/aa628c.

Jiménez-Márquez, F., Vázquez, J. and Sánchez-Rojas, J. L. (2015) 'Optoelectronic sensor device for monitoring the maceration of red wine: Design issues and validation', *Measurement*, 63(March 2016), pp. 128–136. doi: 10.1016/j.measurement.2014.12.009.

Jin-Fei Ding, Zhang, A. P., Li-Yang Shao, Jin-Hua Yan and Sailing He (2005) 'Fiber-taper seeded long-period grating pair as a highly sensitive refractive-index sensor', *IEEE Photonics Technology Letters*, 17(6), pp. 1247–1249. doi: 10.1109/LPT.2005.847437.

Jin, Y., Wong, K. H. and Granville, A. M. (2016) 'Developing localized surface plasmon resonance biosensor chips and fiber optics via direct surface modification of PMMA optical waveguides', *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. Elsevier B.V., 492, pp. 100–109. doi: 10.1016/j.colsurfa.2015.11.025.

Jing, N., Zheng, J., Zhao, X. and Teng, C. (2015) 'Refractive Index Sensing Based on a Side-Polished Macrobending Plastic Optical Fiber', *IEEE Sensors Journal*, 15(5), pp. 2898–2901. doi: 10.1109/JSEN.2014.2385308.

Joe, H.-E., Yun, H., Jo, S.-H., Jun, M. B. G. and Min, B.-K. (2018) 'A review on optical fiber sensors for environmental monitoring', *International Journal of Precision Engineering and Manufacturing-Green Technology*, 5(1), pp. 173–191. doi: 10.1007/s40684-018-0017-6.

Jorgenson, R. C. and Yee, S. S. (1993) 'A fiber-optic chemical sensor based on surface plasmon resonance', *Sensors and Actuators B: Chemical*, 12(3), pp. 213–220. doi: 10.1016/0925-4005(93)80021-3.

Kadhum Hisham, H. (2018) 'Optical Fiber Sensing Technology: Basics, Classifications and Applications', *American Journal of Remote Sensing*, 6(1), p. 1. doi: 10.11648/j.ajrs.20180601.11.

Kelley, C., Krolick, A., Brunner, L., Burklund, A., Kahn, D., Ball, W. and Weber-Shirk, M. (2014) 'An Affordable Open-Source Turbidimeter', *Sensors*, 14(4), pp. 7142–7155. doi: 10.3390/s140407142.

Kemtrak (2016) *DCP007 Process Photometer*. Stockholm, Sweden. Available at: www.kemtrak.com.

Khaleghi, B., Khamis, A. and Karray, F. (2016) 'Multisensor Data Fusion A Data-Centric Review of the State of the Art and Overview of Emerging Trends', in Fourati, H. and Iniewski, K. (eds) *Multisensor data fusion : from algorithms and architectural design to applications*. 1st edn. Boca Raton, FL: CRC Press, pp. 15–27.

Kim, C.-B. and Su, C. B. (2004) 'Measurement of the refractive index of liquids at 1.3 and 1.5 micron using a fibre optic Fresnel ratio meter', *Measurement Science and Technology*, 15(9), pp. 1683–1686. doi: 10.1088/0957-0233/15/9/002.

Kotsiantis, S. B. (2007) 'Supervised machine learning: A review of classification techniques', *Informatica*, 31, pp. 249–268.

Kotsiantis, S. B. (2013) 'Decision trees: A recent overview', *Artificial Intelligence Review*, 39(4), pp. 261–283. doi: 10.1007/s10462-011-9272-4.

Kretschmann, E. and Raether, H. (1968) 'Notizen: Radiative Decay of Non Radiative Surface Plasmons Excited by Light', *Zeitschrift für Naturforschung A*, 23(12), pp. 2135–2136. doi: 10.1515/zna-1968-1247.

Kuang, K. S. C., Quek, S. T. and Maalej, M. (2004) 'Assessment of an extrinsic polymer-based optical fibre sensor for structural health monitoring', *Measurement Science and Technology*, 15(10), pp. 2133–2141. doi: 10.1088/0957-0233/15/10/024.

Lakshmanaprabu, S. K., Shankar, K., Gupta, D., Khanna, A., Rodrigues, J. J. P. C., Pinheiro, P. R. and de Albuquerque, V. H. C. (2018) 'Ranking Analysis for Online Customer Reviews of Products Using Opinion Mining with Clustering', *Complexity*, 2018, pp. 1–9. doi: 10.1155/2018/3569351.

Lambrou, T. P., Anastasiou, C. C. and Panayiotou, C. G. (2010) 'A Nephelometric Turbidity System for Monitoring Residential Drinking Water Quality', in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, pp. 43–55. doi: 10.1007/978-3-642-11870-8_4.

Leung, A., Shankar, P. M. and Mutharasan, R. (2007) 'A review of fiber-optic biosensors', *Sensors and Actuators B: Chemical*, 125(2), pp. 688–703. doi: 10.1016/j.snb.2007.03.010.

Liang, W., Huang, Y., Xu, Y., Lee, R. K. and Yariv, A. (2005) 'Highly sensitive fiber Bragg grating refractive index sensors', *Applied Physics Letters*, 86(15), p. 151122. doi: 10.1063/1.1904716.

Liao, C. R., Hu, T. Y. and Wang, D. N. (2012) 'Optical fiber Fabry-Perot interferometer cavity fabricated by femtosecond laser micromachining and fusion splicing for refractive index sensing', *Optics Express*, 20(20), p. 22813. doi: 10.1364/OE.20.022813.

Libelium Comunicaciones Distribuidas S.L (2019) *Smart Water Technical Guide*, *v.7.5*. Available at: http://www.libelium.com/downloads/documentation/smart_water_sensor_board.pdf (Accessed: 5 September 2019).

Lin, H., Huang, C., Cheng, G., Chen, N. and Chui, H. (2012) 'Tapered optical fiber sensor based on localized surface plasmon resonance', *Optics Express*, 20(19), p. 21693. doi: 10.1364/OE.20.021693.

Lin Xiao, Boyd, S. and Lall, S. (2005) 'A scheme for robust distributed sensor fusion based on average consensus', in *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.* IEEE, pp. 63–70. doi: 10.1109/IPSN.2005.1440896.

Lin, Y.-C. (2013) 'Characteristics of optical fiber refractive index sensor based on surface plasmon resonance', *Microwave and Optical Technology Letters*, 55(3), pp. 574–576. doi: 10.1002/mop.27397.

Liu, R.-T., Tao, L.-Q., Liu, B., Tian, X.-G., Mohammad, M., Yang, Y. and Ren, T.-L. (2016) 'A Miniaturized On-Chip Colorimeter for Detecting NPK Elements', *Sensors*, 16(8), p. 1234. doi: 10.3390/s16081234.

Ben Mabrouk, A. and Zagrouba, E. (2018) 'Abnormal behavior recognition for intelligent video surveillance systems: A review', *Expert Systems with Applications*, 91, pp. 480–491. doi: 10.1016/j.eswa.2017.09.029.

Marcus, R. T. (2014) 'Colorometry', in Webster, J. G. and Eren, H. (eds) *Measurement, Instrumentation, and Sensors Handbook, Second Edition: Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement*. Second. CRC Press, pp. 48-1-48–22.

Martinez-Uso, A., Pla, F. and Sotoca, J. M. (2010) 'A Semi-supervised Gaussian Mixture Model for Image Segmentation', in *2010 20th International Conference on Pattern Recognition*. IEEE, pp. 2941–2944. doi: 10.1109/ICPR.2010.721.

Maselli Misure (2017) 'Technical Datasheet n. 90034H3743 - Process Refractometer mod. UR24'. Available at: https://www.merazet.pl/wp-content/uploads/img/cea24b8f.pdf (Accessed: 13 August 2019).

Meeten, G. H. (2014) 'Refractive Index Measurement', in Webster, J. G. and Eren, H. (eds) *Measurement, Instrumentation, and Sensors Handbook, Second Edition: Electromagnetic, Optical, Radiation, Chemical, and Biomedical Measurement*. Second. CRC Press, pp. 50-1-50–11.

Mettler-Toledo GmbH, A. (2014) 'Refracto – reliable refractive index wherever you need it'. Available at: http://www.mt.com/dam/Analytical/Refractometry/RE-PDF/30243620_V12.14_Refracto30PX_30GS_en.pdf (Accessed: 13 August 2019).

Miao, Y., Liu, B. and Zhao, Q. (2009) 'Refractive index sensor based on measuring the transmission power of tilted fiber Bragg grating', *Optical Fiber Technology*. Elsevier Inc., 15(3), pp. 233–236. doi: 10.1016/j.yofte.2008.11.002.

Mileńko, K., Hu, D. J. J., Shum, P. P., Zhang, T., Lim, J. L., Wang, Y., Woliński, T. R., Wei, H. and Tong, W. (2012) 'Photonic crystal fiber tip interferometer for refractive index sensing', *Optics Letters*, 37(8), p. 1373. doi: 10.1364/OL.37.001373.

Mitchell, L. and Mitchell, P. L. (2009a) 'Mitchell Method M5271', *Determination of turbidity by LED nephelometry (Rev. 1.1)*. Grand Junction, CO, pp. 1–8.

Mitchell, L. and Mitchell, P. L. (2009b) 'Mitchell Method M5331', *Determination of turbidity by LED nephelometry (Rev. 1.1)*. Grand Junction, CO, pp. 1–8. Available at: http://www.nemi.gov.

Monzón-Hernández, D., Villatoro, J., Talavera, D. and Luna-Moreno, D. (2004) 'Optical-fiber surface-plasmon resonance sensor with multiple resonance peaks', *Applied Optics*, 43(6), p. 1216. doi: 10.1364/AO.43.001216.

Murphy, K. *et al.* (2015) 'A low-cost autonomous optical sensor for water quality monitoring', *Talanta*. Elsevier, 132(November), pp. 520–527. doi: 10.1016/j.talanta.2014.09.045.

Nath, P., Singh, H. K., Datta, P. and Sarma, K. C. (2008) 'All-fiber optic sensor for measurement of liquid refractive index', *Sensors and Actuators A: Physical*, 148(1), pp. 16–18. doi: 10.1016/j.sna.2008.06.027.

Näykki, T., Koponen, S., Väisänen, T., Pyhälahti, T., Toivanen, T. and Leito, I. (2014) 'Validation of a new measuring system for water turbidity field measurements', *Accreditation and Quality Assurance*, 19(3), pp. 175–183. doi: 10.1007/s00769-014-1052-9.

Nilsson, N. J. (1998) *Introduction to machine learning*. Stanford: An early draft of a proposed textbook. Available at: https://ai.stanford.edu/~nilsson/mlbook.html.

Noiseux, I., Long, W., Cournoyer, A. and Vernon, M. (2004) 'Simple Fiber-Optic-Based Sensors for Process Monitoring: An Application in Wine Quality Control Monitoring', *Applied Spectroscopy*, 58(8), pp. 1010–1019. doi: 10.1366/0003702041655476.

Novo, C., Pinto, J. L., Bilro, L., Antunes, P., Alberto, N., Pereira, V., Marques, J. and Mon-chu Chen (2013) 'Madeira wine online quality control', in *2013 2nd Experiment@ International Conference (exp.at'13)*. IEEE, pp. 186–187. doi: 10.1109/ExpAt.2013.6703065.

Novo, C., Bilro, L., Alberto, N., Antunes, P., Nogueira, R. and Pinto, J. L. L. (2014) 'Plastic optical fibre sensor for Madeira wine monitoring', in Martins Costa, M. F. P. C. and Nogueira, R. N. (eds) *Proceedings of SPIE - The International Society for Optical Engineering*, p. 92862Q. doi: 10.1117/12.2060322.

Nweke, H. F., Teh, Y. W., Mujtaba, G. and Al-garadi, M. A. (2019) 'Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions', *Information Fusion*, 46, pp. 147–170. doi: 10.1016/j.inffus.2018.06.002.

O'Dell, J. W. (1993) *Method 180.1 determination of turbidity by nephelometry (Rev. 2)*. Available at: https://www.epa.gov/sites/production/files/2015-08/documents/method_180-1_1993.pdf.

Oliveira, N. *et al.* (2017) 'Winegrid ® : the remote and real-time wine production process monitoring system', in *40th World Congress of Vine and Wine*. Sofia, Bulgaria, pp. 1–6.

Oliveira, R., Marques, T. H. R., Bilro, L., Nogueira, R. and Cordeiro, C. M. B. (2017) 'Multiparameter POF Sensing Based on Multimode Interference and Fiber Bragg Grating', *Journal of Lightwave Technology*, 35(1), pp. 3–9. doi: 10.1109/JLT.2016.2626793.

Omar, A. F. and MatJafri, M. Z. (2008) 'Water Quality Measurement using Transmittance and 90° Scattering Techniques through Optical Fiber Sensor', in *2008 6th National Conference on Telecommunication Technologies and 2008 2nd Malaysia Conference on Photonics*. IEEE, pp. 17–21. doi: 10.1109/NCTT.2008.4814227.

Organisation internationale de la vigne ey du vin (2017) *State of the Vitiviniculture World Market*.

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G. and Mechelli, A. (2012) 'Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review', *Neuroscience & Biobehavioral Reviews*. Elsevier Ltd, 36(4), pp. 1140–1152. doi: 10.1016/j.neubiorev.2012.01.004.

Osram GmbH. (2015a) *Blue Laser Diode in TO38 ICut Package Version 2.0*. Available at: http://www.osram-os.com/Graphics/XPic5/00193831_0.pdf.

Osram GmbH. (2015b) *Green Laser Diode in TO38 ICut Package Version 1.0*. Available at: http://www.osram-os.com/Graphics/XPic6/00179009_0.pdf.

Otto, A. (1968) 'Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection', *Zeitschrift für Physik A Hadrons and nuclei*, 216(4), pp. 398–410. doi: 10.1007/BF01391532.

Owen, T. (2000) 'Instrumentation', in *Fundamentals of modern UV-visible spectroscopy*. Germany: Agilent Technologies. Available at: https://www.agilent.com/cs/library/primers/Public/59801397_020660.pdf.

Parslow, J., Cressie, N., Campbell, E. P., Jones, E. and Murray, L. (2013) 'Bayesian learning and predictability in a stochastic nonlinear dynamical model', *Ecological Applications*, 23(4), pp. 679–698. doi: 10.1890/12-0312.1.

Pasquini, C. (2018) 'Near infrared spectroscopy: A mature analytical technique with new perspectives – A review', *Analytica Chimica Acta*, 1026, pp. 8–36. doi: 10.1016/j.aca.2018.04.004.

Paz, A. M. C., Acevedo, J. M., Vázquez, A. del R., Peñalver, C. M., Ferreiro, A. L., Meléndez, A. A. N. and Gandoy, J. D. (2010) 'A Multi-Point Sensor Based on Optical Fiber for the Measurement of Electrolyte Density in Lead-Acid Batteries', *Sensors*, 10(4), pp. 2587–2608. doi: 10.3390/s100402587.

Paz, A. M. C. y, Acevedo, J. M., Gandoy, J. D., del Rio Vazquez, A., Freire, C. M.-P. and Soria, M. (2007) 'Plastic Optical Fiber Sensor for Real Time Density Measurements in Wine Fermentation', *2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007*. IEEE, pp. 1–5. doi: 10.1109/IMTC.2007.379330.

Plumb, R. S., Dear, G. J., Mallett, D. N., Higton, D. M., Pleasance, S. and Biddlecombe, R. A. (2001) 'Quantitative analysis of pharmaceuticals in biological fluids using high-performance liquid chromatography coupled to mass spectrometry: a review', *Xenobiotica*, 31(8–9), pp. 599–617. doi: 10.1080/00498250110060941.

Polynkin, P., Polynkin, A., Peyghambarian, N. and Mansuripur, M. (2005) 'Evanescent field-based optical fiber sensing device for measuring the refractive index of liquids in microfluidic channels', *Optics Letters*, 30(11), p. 1273. doi: 10.1364/OL.30.001273.

Portugal, I., Alencar, P. and Cowan, D. (2018) 'The use of machine learning algorithms in recommender systems: A systematic review', *Expert Systems with Applications*, 97, pp. 205–227. doi: 10.1016/j.eswa.2017.12.020.

Potdar, A. A., Longstaff, A. P., Fletcher, S. and Mian, N. S. (2015) 'Application of multi sensor data fusion based on principal component analysis and artificial neural network for machine tool thermal monitoring', *Laser Metrology and Machine Performance XI - 11th International Conference and Exhibition on Laser Metrology, Machine Tool, CMM and Robotic Performance, LAMDAMAP 2015*, pp. 228–237.

Rahman, H. A., Harun, S. W., Yasin, M., Phang, S. W., Damanhuri, S. S. A., Arof, H. and Ahmad, H. (2011) 'Tapered plastic multimode fiber sensor for salinity detection', *Sensors and Actuators A: Physical*. Elsevier B.V., 171(2), pp. 219–222. doi: 10.1016/j.sna.2011.09.024.

Raol, J. . (2016) 'Introduction to Data Fusion Process', in *Data fusion mathematics : theory and practice*. 1st edn. Boca Raton, FL: CRC Press, pp. 1–12.

Reichert Tecnologies Analitical Instruments (2019) 'Refractometer Buyer ' s Guide'. Available at: https://www.reichertai.com/clientuploads/directory/download_pdfs/REFRACTOMETER_BUYER_G UIDE.pdf (Accessed: 13 August 2019).

Ribéreau-Gayon, P., Glories, Y., Maujean, A. and Dubourdieu, D. (2006a) *Handbook of Enology*, *Handbook of Enology, The Chemistry of Wine: Stabilization and Treatments: Second Edition*. Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/0470010398.

Ribéreau-Gayon, P., Glories, Y., Maujean, A. and Dubourdieu, D. (2006b) 'Phenolic Compounds', in *Handbook of Enology Volume 2: The Chemistry of Wine, Stabilization and Treatments*. 2nd edn. John Wiley & Sons, Ltd, pp. 141–203.

Ribéreau-Gayon, P., Glories, Y., Maujean, A. and Dubourdieu, D. (2006c) 'Red Winemaking', in *Handbook of Enology:The Microbiology of Wine and Vinifications*. 2nd edn. John Wiley & Sons, Ltd, pp. 326–395.

Robert, C. P. (2016) 'Monte Carlo Methods', in *Wiley StatsRef: Statistics Reference Online*. Chichester, UK: John Wiley & Sons, Ltd, pp. 1–13. doi: 10.1002/9781118445112.stat03876.pub2.

Rogova, G. L. and Nimier, V. (2004) 'Reliability In Information Fusion : Literature Survey', *Proceedings of the Seventh International Conference on Information Fusion, FUSION 2004*, Volume 2, p. Pages 1158-1165.

Rudin, C. *et al.* (2012) 'Machine Learning for the New York City Power Grid', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(2), pp. 328–345. doi: 10.1109/TPAMI.2011.108.

Saffiotti, A. (1997) 'The uses of fuzzy logic in autonomous robot navigation', *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 1(4), pp. 180–197. doi: 10.1007/s005000050020.

Salmerón, A., Rumí, R., Langseth, H., Nielsen, T. D. and Madsen, A. L. (2018) 'A Review of Inference Algorithms for Hybrid Bayesian Networks', *Journal of Artificial Intelligence Research*, 62, pp. 799–828. doi: 10.1613/jair.1.11228.

Schrider, D. R. and Kern, A. D. (2018) 'Supervised Machine Learning for Population Genetics: A New Paradigm', *Trends in Genetics*, 34(4), pp. 301–312. doi: 10.1016/j.tig.2017.12.005.

Scientific, T. F. (2009) 'Thermo Scientific Orion Method AQ4500', *Determination of turbidity by LED nephelometry (Rev. 1.0)*, pp. 1–10.

Sequeira, F., Duarte, D., Nogueira, R., Rudnitskaya, A., Cennamo, N., Zeni, L. and Bilro, L. (2016) 'Analysis of the roughness in a sensing region on D-shaped POFs', in *Proc. Internacional Conf. on Plastic Optical Fibers - POF*. Birmingham, p. OP36.

Sequeira, Filipa, Duarte, D., Bilro, L., Rudnitskaya, A., Pesavento, M., Zeni, L. and Cennamo, N. (2016) 'Refractive Index Sensing with D-Shaped Plastic Optical Fibers for Chemical and Biochemical Applications', *Sensors*, 16(12), p. 2119. doi: 10.3390/s16122119.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. and de Freitas, N. (2016) 'Taking the Human Out of the Loop: A Review of Bayesian Optimization', *Proceedings of the IEEE*, 104(1), pp. 148–175. doi: 10.1109/JPROC.2015.2494218.

Shao, L.-Y., Zhang, A. P., Liu, W.-S., Fu, H.-Y. and He, S. (2007) 'Optical Refractive-Index Sensor Based on Dual Fiber-Bragg Gratings Interposed With a Multimode-Fiber Taper', *IEEE Photonics Technology Letters*, 19(1), pp. 30–32. doi: 10.1109/LPT.2006.889010.

Shao, M., Qiao, X., Fu, H., Li, H., Jia, Z. and Zhou, H. (2014) 'Refractive Index Sensing of SMS Fiber Structure Based Mach-Zehnder Interferometer', *IEEE Photonics Technology Letters*, 26(5), pp. 437–439. doi: 10.1109/LPT.2013.2295375.

Shao, Y., Xu, S., Zheng, X., Wang, Y. and Xu, W. (2010) 'Optical Fiber LSPR Biosensor Prepared by Gold Nanoparticle Assembly on Polyelectrolyte Multilayer', *Sensors*, 10(4), pp. 3585–3596. doi: 10.3390/s100403585.

Shimada, S. *et al.* (2017) 'Refractive index sensing using ultrasonically crushed plastic optical fibers', *POF 2016 - 25th International Conference on Plastic Optical Fibres, Conference Proceedings*, 012201, pp. 3–5.

Shin, J.-D. and Park, J. (2015) 'High-sensitivity refractive index sensors based on in-line holes in plastic optical fiber', *Microwave and Optical Technology Letters*, 57(4), pp. 918–921. doi: 10.1002/mop.28991.

Shrake, N. L., Amirtharajah, R., Brenneman, C., Boulton, R. and Knoesen, A. (2014) 'In-Line Measurement of Color and Total Phenolics during Red Wine Fermentations Using a Light-Emitting Diode Sensor', *American Journal of Enology and Viticulture*, 65(4), pp. 463–470. doi: 10.5344/ajev.2014.14023.

da Silva, I. N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L. H. B. and dos Reis Alves, S. F. (2017) *Artificial Neural Networks*. Springer International Publishing. doi: 10.1007/978-3-319-43162-8.

Soloman, S. (2010) *Sensors Handbook*. second, *Advanced Manufacturing*. second. McGraw-Hill Education.

Sridhar, M. B., Srinivas, Y. and Krishna Prasad, M. H. M. (2012) 'Software Reuse in Cardiology Related Medical Database Using K-Means Clustering Technique', *Journal of Software Engineering and Applications*, 05(09), pp. 682–686. doi: 10.4236/jsea.2012.59081.

Srivastava, S. K. (2013) 'Fiber Optic Plasmonic Sensors: Past, Present and Future', *The Open Optics Journal*, 7(1), pp. 58–83. doi: 10.2174/1874328501307010058.

Stewart, S., Ivy, M. A. and Anslyn, E. V. (2014) 'The use of principal component analysis and discriminant analysis in differential sensing routines', *Chem. Soc. Rev.*, 43(1), pp. 70–84. doi: 10.1039/C3CS60183H.

Su, H. and Huang, X. G. (2007) 'Fresnel-reflection-based fiber sensor for on-line measurement of solute concentration in solutions', *Sensors and Actuators B: Chemical*, 126(2), pp. 579–582. doi: 10.1016/j.snb.2007.04.008.

Sumriddetchkajorn, S., Chaitavon, K. and Intaravanne, Y. (2014) 'Mobile-platform based colorimeter for monitoring chlorine concentration in water', *Sensors and Actuators B: Chemical*. Elsevier B.V., 191, pp. 561–566. doi: 10.1016/j.snb.2013.10.024.

Sunny, Mishra, V. N., Dwivedi, R. and Das, R. R. (2014) 'Quantification of Individual Gases/Odors Using Dynamic Responses of Gas Sensor Array With ASM Feature Technique', *IEEE Sensors Journal*, 14(4), pp. 1006–1011. doi: 10.1109/JSEN.2013.2292319.

Suominen, L., Ruokolainen, K., Tuomisto, H., Llerena, N. and Higgins, M. A. (2013) 'Predicting soil properties from floristic composition in western Amazonian rain forests: performance of k - nearest neighbour estimation and weighted averaging calibration', *Journal of Applied Ecology*. Edited by P. Brando, 50(6), pp. 1441–1449. doi: 10.1111/1365-2664.12131.

Swan, A. L., Mobasheri, A., Allaway, D., Liddell, S. and Bacardit, J. (2013) 'Application of Machine Learning to Proteomics Data: Classification and Biomarker Identification in Postgenomics Biology', *OMICS: A Journal of Integrative Biology*, 17(12), pp. 595–610. doi: 10.1089/omi.2013.0017.

SWAN Analytische Instrumente AG (2009) 'Continuous Measurement of Turbidity Using a SWAN AMI Turbiwell Turbidimeter', pp. 1–9.

Tarrant, A. W. S. (2010) 'Optical Measurements', in Boyes, W. (ed.) *Instrumentation reference book*. 4th ed. Butterworth-Heinemann/Elsevier, pp. 499–519.

Teng, C., Jing, N., Yu, F. and Zheng, J. (2016) 'Investigation of a Macro-Bending Tapered Plastic Optical Fiber for Refractive Index Sensing', *IEEE Sensors Journal*, 16(20), pp. 7521–7525. doi: 10.1109/JSEN.2016.2601380.

Teng, C., Yu, F., Ding, Y. and Zheng, J. (2017) 'Refractive index sensor based on multi-mode plastic optical fiber with long period grating', in Baldini, F., Homola, J., and Lieberman, R. A. (eds), p. 102311M. doi: 10.1117/12.2265305.

Tian, Z., Yam, S. S.-H. and Loock, H.-P. (2008) 'Single-Mode Fiber Refractive Index Sensor Based on Core-Offset Attenuators', *IEEE Photonics Technology Letters*, 20(16), pp. 1387–1389. doi: 10.1109/LPT.2008.926832.

Tian, Z., Yam, S. S. and Loock, H. (2008) 'Refractive index sensor based on an abrupt taper Michelson interferometer in a single-mode fiber', *Optics Letters*, 33(10), p. 1105. doi: 10.1364/OL.33.001105.

Tiwari, P., Saxena, V. P., Mishra, R. G. and Bhavsar, D. (2015) 'Wireless Sensor Networks: Introduction, Advantages, Applications and Research Challenges', *HCTL Open International Journal of Technology Innovations and Research (IJTIR)*, 14(April), pp. 1–11.

Tkachenko, N. (2006) *Steady State Absorption Spectroscopy*, *Optical Spectroscopy Methods and Instrumentations*.

Torizuka, K., Oi, H., Saitoh, F. and Ishizu, S. (2018) 'Benefit Segmentation of Online Customer Reviews Using Random Forest', in *2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, pp. 487–491. doi: 10.1109/IEEM.2018.8607697.

Tran, N. T., Campbell, C. G. and Shi, F. G. (2006) 'Study of particle size effects on an optical fiber sensor response examined with Monte Carlo simulation', *Applied Optics*, 45(29), p. 7557. doi: 10.1364/AO.45.007557.

Triguero, I., García-Gil, D., Maillo, J., Luengo, J., García, S. and Herrera, F. (2019) 'Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(2), p. e1289. doi: 10.1002/widm.1289.

Vernhet, A., Bes, M., Bouissou, D., Carrillo, S. and Brillouet, J.-M. (2016) 'Characterization of suspended solids in thermo-treated red musts', *OENO One*, 50(1), p. 9. doi: 10.20870/oeno-one.2016.50.1.50.

Wang, Q., Kong, L., Dang, Y., Xia, F., Zhang, Y., Zhao, Y., Hu, H. and Li, J. (2016) 'High sensitivity refractive index sensor based on splicing points tapered SMF-PCF-SMF structure Mach-Zehnder mode interferometer', *Sensors and Actuators B: Chemical*. Elsevier B.V., 225(61371200), pp. 213–220. doi: 10.1016/j.snb.2015.11.047.

Wang, T. and Wang, M. (2012) 'Fabry Pérot Fiber Sensor for Simultaneous Measurement of Refractive Index and Temperature Based on an In-Fiber Ellipsoidal Cavity', *IEEE Photonics Technology Letters*, 24(19), pp. 1733–1736. doi: 10.1109/LPT.2012.2212184.

Wang, X.-D. and Wolfbeis, O. S. (2013) 'Fiber-Optic Chemical Sensors and Biosensors (2008–2012)', *Analytical Chemistry*, 85(2), pp. 487–508. doi: 10.1021/ac303159b.

Webb, D. J. (2015) 'Fibre Bragg grating sensors in polymer optical fibres', *Measurement Science and Technology*. IOP Publishing, 26(9), p. 092004. doi: 10.1088/0957-0233/26/9/092004.

Willets, K. A. and Van Duyne, R. P. (2007) 'Localized Surface Plasmon Resonance Spectroscopy and Sensing', *Annual Review of Physical Chemistry*, 58(1), pp. 267–297. doi: 10.1146/annurev.physchem.58.032806.104607.

Wu, C., Tse, M.-L. V., Liu, Z., Guan, B.-O., Zhang, a P., Lu, C. and Tam, H.-Y. (2014) 'In-line microfluidic integration of photonic crystal fibres as a highly sensitive refractometer', *The Analyst*, 139(21), pp. 5422–5429. doi: 10.1039/C4AN01361A.

Wu, D., Zhu, T., Deng, M., Duan, D., Shi, L., Yao, J. and Rao, Y. (2011) 'Refractive index sensing based on Mach–Zehnder interferometer formed by three cascaded single-mode fiber tapers', *Applied Optics*, 50(11), p. 1548. doi: 10.1364/AO.50.001548.

Wu, D. K. C., Kuhlmey, B. T. and Eggleton, B. J. (2009) 'Ultrasensitive photonic crystal fiber refractive index sensor', *Optics Letters*, 34(3), p. 322. doi: 10.1364/OL.34.000322.

Xue, S., van Eijkelenborg, M. A., Barton, G. W. and Hambley, P. (2007) 'Theoretical, Numerical, and Experimental Analysis of Optical Fiber Tapering', *Journal of Lightwave Technology*, 25(5), pp. 1169–1176. doi: 10.1109/JLT.2007.893028.

Yeoh, S., Matjafri, M. Z., Mutter, K. N. and Oglat, A. A. (2019) 'Plastic fiber evanescent sensor in measurement of turbidity', *Sensors and Actuators, A: Physical*. Elsevier B.V., 285, pp. 1–7. doi: 10.1016/j.sna.2018.10.042.

Yin, M., Gu, B., An, Q.-F., Yang, C., Guan, Y. L. and Yong, K.-T. (2018) 'Recent development of fiber-optic chemical sensors and biosensors: Mechanisms, materials, micro/nano-fabrications and applications', *Coordination Chemistry Reviews*, 376, pp. 348–392. doi: 10.1016/j.ccr.2018.08.001.

Yokotsuka, K., Yajima, M., Seki, T. and Matthews, M. (1997) 'Changes in Color Parameters during Fermentation and Storage of Red Wine Using Thai Roselle under Different pHs and SO2 Concentrations', *Food Science and Technology International, Tokyo*, 3(2), pp. 105–109.

Zendehboudi, A., Baseer, M. A. and Saidur, R. (2018) 'Application of support vector machine models for forecasting solar and wind energy resources: A review', *Journal of Cleaner Production*, 199, pp. 272–285. doi: 10.1016/j.jclepro.2018.07.164.

Zhang, W., Webb, D. and Peng, G. (2012) 'Polymer optical fiber Bragg grating acting as an intrinsic biochemical concentration sensor', *Optics Letters*. Edited by K. Kalli and A. Mendez, 37(8), p. 1370. doi: 10.1364/OL.37.001370.

Zhou, Z., Li, D. X., Qin, Y. and Liu, L. F. (2014) 'Research on a Liquid Color Detecting Device', *Advanced Materials Research*, 981, pp. 602–607. doi: 10.4028/www.scientific.net/AMR.981.602.

Zhu, T., Wu, D., Deng, M., Duan, D.-W., Rao, Y.-J. and Bao, X. (2011) 'Refractive index sensing based on Mach-Zehnder interferometer formed by three cascaded single-mode fiber tapers', in *Applied Optics*, p. 77532P. doi: 10.1117/12.884842.

Zubia, J., Garitaonaindía, G. and Arrúe, J. (2000) 'Passive device based on plastic optical fibers to determine the indices of refraction of liquids', *Applied Optics*, 39(6), p. 941. doi: 10.1364/AO.39.000941.

Žuvela, P., Skoczylas, M., Jay Liu, J., Bączek, T., Kaliszan, R., Wong, M. W. and Buszewski, B. (2019) 'Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography', *Chemical Reviews*, 119(6), pp. 3674–3729. doi: 10.1021/acs.chemrev.8b00246.