UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Physics

Final Dissertation

Water detection in multiphase mixtures

Thesis supervisor
Prof. Cinzia Sada
Thesis co-supervisor
Dott. Diego Tescaro

Candidate
Marco Simionato

Academic Year 2019/2020

Abstract

The measurement of the water component in multiphase mixtures is key in the Oil and Gas industry. My thesis work was developed in Pietro Fiorentini Spa which is a company who provides services to the oil and gas industries from the extraction to the transport at the final user. These services can includes gas meters for the houses, remote control of a gas pipeline and extraction analysis. Actually Pietro Fiorentini Spa is developing a microwave sensor that measures the complex permittivity of the mixture in order to compute the water fraction in the extraction process of an oil wells. The final measurement accuracy also depends on the modelling of the multiphase mixture (liquid and gas). After the study of the application, the Thesis foresees the improvement of the analytical model in use through the analysis of available data. Finally, a novel data-driven analysis approach with "machine learning" techniques will be developed and compared to the previous

Contents

Chapter 1

Introduction

1.1 Introduction to multiphase flow measurement

A multiphase mixture is a solution in which we can encounter two or more phases, this can be seen in many simple examples of our lives such as falling rain (liquid) through the air (gas) or bubbles (gas) on a glass of water (liquid). Multiphase flow can be encountered in many industrial applications such as food, pharmaceutical and paper factory and in Gas & Oil industries. In most of these applications there is no need to measure the fraction of the different phases composing a particular flow but this becomes important in Gas & Oil industries where we need to monitor the quantity of gas and oil flowing inside a pipe to understand the efficiency of a given oil well, but also to see if they can respect the quality assurance protocol and trustworthy monitoring for safety reasons.
Before entering into the details of flow regimes inside a pipe and the measuring technique we give a brief resume on the terminology and aspects characterising a multiphase flow:

- **Phase mass fraction:** is the ratio between the volumetric fraction of a phase and the total, in particular the most used is the GVF, gas volume fraction which is *GVF=gas volumetric flow rate/total volumetric flow rate*;
- **Water cut:** or also called water liquid ratio (**WLR**) which is the ratio between the water volumetric flow rate and the total liquid volumetric flow rate, *WLR = water volumetric flow rate/total liquid volumetric flow rate*;
- **Phase slip:** is the difference between the velocity of the gas and the liquid inside the pipe, $\nu_R = \nu_{gas} - \nu_{liq}$, so it is a measure of the relative velocity between the phases;
- **Superficial phase velocity:** is the velocity a single phase would have if flowing alone in the pipe;
- **Wet gas:** is a flow in which the gas component is very high respect to the liquid. There are multiple definition for a wet gas regime, the simpler consider the GVF and defines the wet gas regime when the GVF is above 90 or 95%;
- **Water continuous:** is a flow in which the liquid part is made of dispersed particles of oil in a medium of water;
- **Oil continuous:** is a flow in which the liquid part is made of dispersed particles of water in a medium of oil;
- **Inversion region:** is the region in which there is a switch between oil/water continuous regime. This region is not fixed and depends on the type of flow and velocity regime, for

example at the mixer we can have a switch at 50 % WLR, but exist some oil wells with a switch at WLR 30 % or also at 70 %;

These are the most used definition in Oil industry and they will help us understand what will follow in the next sections.
We will now introduce the flow conditions which we can encounter in industrial applications and in the next chapter we will see the measuring techniques used to analyze the flow.

1.2 Type of flows

Inside a pipe we can have very different situations based on the velocities and the ratio between the phases, let's consider for the moment a two phases flow in which there is a liquid and a gas component respectively.
For the moment we start considering a vertical pipe in which the gravity force has the same direction of the pipe and so the flow remains cylindrical symmetrical respect to the axis of the pipe.

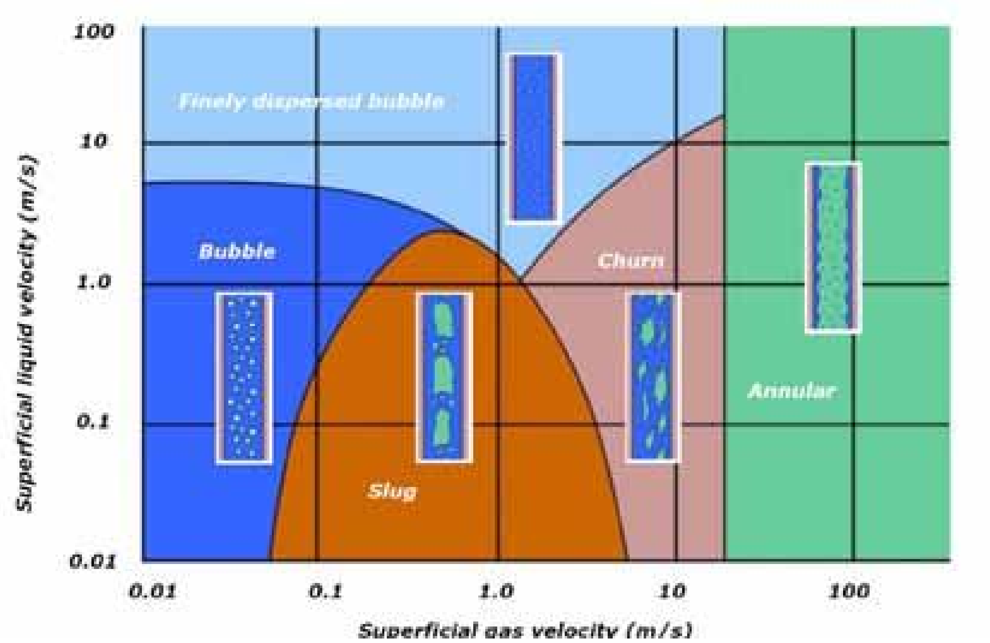

Figure 1.1: Type of flows in a vertical pipe [1]

In Fig. 1.1 we can notice five different type of flow depending on the superficial liquid and gas velocities

- **Bubble:** is a continuous phase of liquid in which there are dispersed bubbles of gas;
- **Slug:** it is characterized by an alternating region of high and low concentration of gas in a continuous host of liquid;
- **Churn:** At high velocities of gas there is an irregular flow of gas slugs;
- **Finely dispersed bubbles:** at higher liquid velocities the bubbles of gas becomes smaller and the liquid fraction increases;
- **Annular:** at very high gas velocities the liquid will create a ring on the walls of the pipe and the gas will flow inside the remaining empty region.

For a horizontal pipe the behavior is altered by the presence of gravity which causes the liquid to lie on the bottom of the pipe, for this reason we have more flow regime conditions.

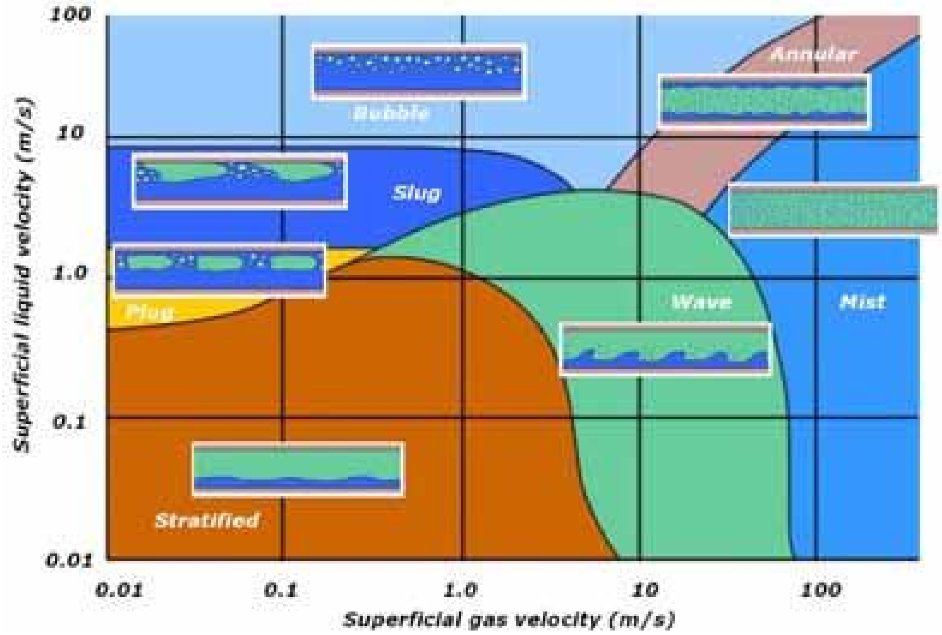

Figure 1.2: Type of flows in a horizontal pipe [1]

- **Stratified:** in this regime the liquid and gas velocity are low and so the two phases do not mix. The liquid component lies on the bottom of the pipe and the gas on top;
- **Wave:** is similar to the stratified regime but the higher velocity of the gas creates a wavy pattern on the interphase between the two phases
- **Plug:** in this regime we have an alternating region of gas in a continuous medium of liquid. For the gravity influence the gas will lie on the upper part of the pipe;
- **Slug:** is similar to the plug regime but we have an higher flow rate of gas resulting in a greater region of gas;
- **Bubble:** is similar to the vertical situation;
- **Annular:** is similar to the vertical situation;
- **Mist:** when the gas velocity becomes higher the gas starts to pick up some of the liquid on the bottom of the pipe which becomes suspended in the gas in the form of small droplets.

As it can be seen in a horizontal pipe the behavior of the flow is much more complicated and has an higher dependence on the velocities of the two phases and has also a strong dependence on the angle. The angle dependence can be seen in Fig. 1.3 where with different colors we are representing the flow for the different phases, we can notice that even a deviation on the angle of one degree causes a very different flow regime. In the lower pictures we can see the boundaries regions of different types of flow regimes and we can notice that they can change widely respect to the angle. Let's consider for example the case of the stratified flow (SS). In this case when the angle is 89° the pipe has a small increase in height respect to the flow direction and we can see that this regime occurs only at specific superficial velocities. On the other hand, when the angle is 90° or 91° this regime can occur in a wider range of velocities since the flow is not hindered by the gravity force. In particular on the last case the gravity force favors the stratification of the two phases resulting in a an increased superficial velocities area at which this regime can occur.

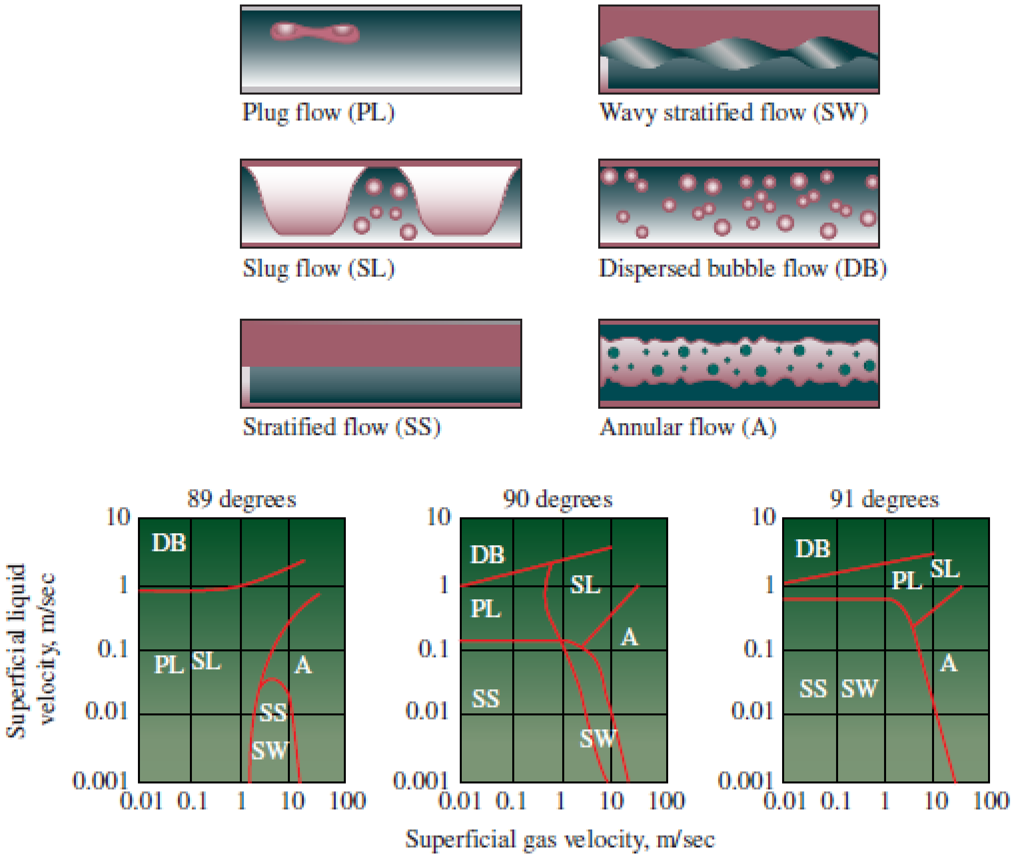

Figure 1.3: Flow rate dependence on the angle [5]

For these reasons we prefer to measure the WLR using a vertical pipe. With this configuration we limit the number of flow regime we can have and we are independent of any error of alignment of the pipe.
In particular the vertical pipe is usually preceded by a blind tee configuration, Fig. 1.4, which has the advantage of mixing the flow and creates homogenization in the flow upstream the blind tee.
The blind tee helps us to reduce the complexity of the analysis since the mixing effect for the liquid part makes the flow homogeneous and so the slip velocity between the water and oil phases becomes almost zero.
In figures 1.5 and 1.6 we can see a simulation for the flow rate inside the pipe and notice the mixing effect. These simulations are performed using a three dimensional computational fluid dynamics (CFD). In the first simulation the MultiPhaseEulerFoam solver from the OpenFoam CFD software [4] was used to simulate both dispersed and separated flow within the same geometry. In the first image there is a gas liquid flow with low velocity and with the red color we refer to pure gas. On the horizontal pipe the two phases are divided and we have gas on top of the pipe while on the vertical pipe the flow becomes mixed and we have a portion of gas inside the pipe mixed together with liquid. In the second picture with the different colors we refer to the velocity of different streamlines of a laminar flow of a single phase.We can notice that at the end of the pipe the streamlines have a circular motion resulting in a mixing effect. In this case what happens is that the fluid inside the end section of the blind tee mixes with the incoming flow and a portion of the fluid will go upward in the pipe. In this simulation the

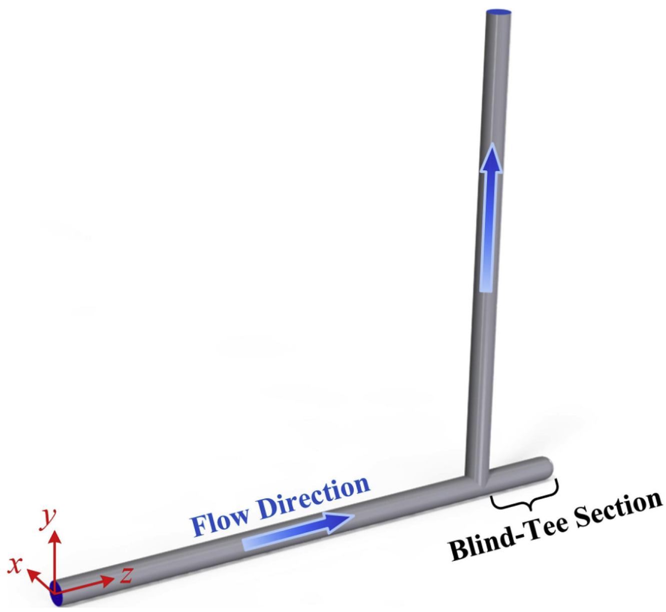

Figure 1.4: Blind tee configuration [2]

boundaries conditions use a cylindrical symmetry to simplify the computations, the blind tee section length is set at 3 diameters and the Reynolds number was set at 1000.

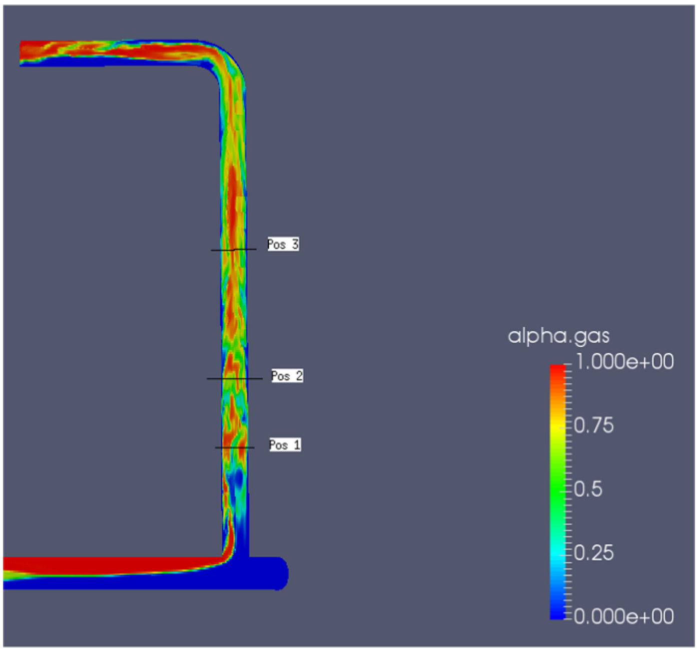

Figure 1.5: CFD simulation with gas liquid distribution [3]

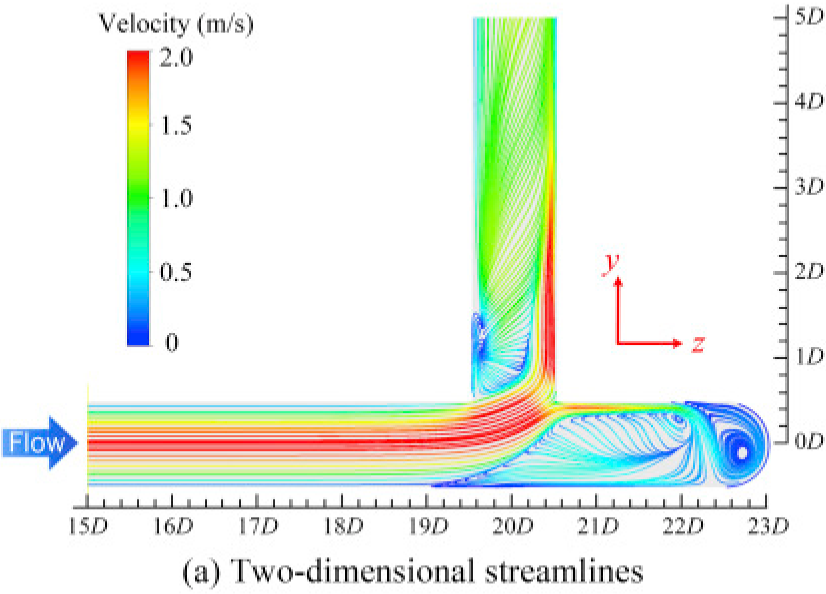

Figure 1.6: Streamlines on the symmetrical plane [2]

1.3 Importance to measure the WLR

Since the beginning of oil wells, the oil industries have worked with a multiphase flow. This condition occurs because even if the reservoir is made of one phase, because of the high pressure, when we extract these substances the different pressure at ground level will cause the oil to evaporate and lead to a two or three phase flow. The importance of measuring the WLR started to gain importance in offshore wells, in this facilities the small area available and the weight are a big concern so we need to occupy the least area and save more weight as possible. For this reason a separator, Fig. 1.7, to measure the water cut cannot be used. The principle of a separator is to collect the fluids from the well and to divide the phases thanks to the deposition effect coming from different densities.

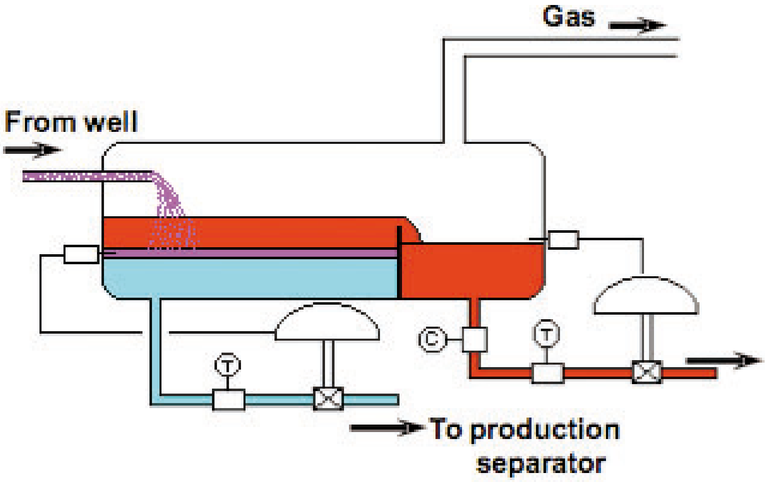

Figure 1.7: Design of a separator [1]

This configuration has many disadvantages for the offshore use:

- **Time:** since the separation occurs for densities difference and oil and water have a small difference we need to wait a lot of time for the deposition to occur;
- **Weight:** To increase the production we need to have a very large separator which will increase the weight on the platform;
- **Mechanics:** a separator needs to have three pumps: one for the well fluids and two for the separated oil and liquid phase.

Furthermore, due to this mechanism the separator is not able to provide a real time analysis of the flow but we gain information on it only at the end of the process. Some other problems can occur like carryover and carryunder: the first is when some liquid escapes with some gas and the second when some gas escapes with the liquid.
So there was the need to provide a sensor able to detect the water cut in the best efficient way. To overcome these problems a multiphase flow meter is the best configuration since it gives a real time analysis on the flow properties, it is connected to the principal pipe and there is the need of only one pump to extract the oil from the well.

Chapter 2

Measuring technique

2.1 Multiphase flow meter

There are many types of multiphase flow meter available on the market and they have many different setup of sensors chosen on the type of flow to analyze. For example a meter designed for a wet gas will have different component respect to a meter designed for measuring a flow with small GVF only in oil continuous regime. We will return later on the type of meter we can use and which types of sensors characterize them, first we summarize the type of analysis that can be done to analyze a multiphase flow.
There are 4 routes (Fig. 2.1) in which we can divide all the types of meters and for each of this route we have a different type of measurement and sensors we can use. Two steps are used to divide the approach in our analysis:

- **1^{st} step:** based on the flow regime we are analysing and divides the approach in a homogeneous flow or in a non homogeneous flow;
- **2^{nd} step:** based on the use of a small sample of the flow to help the analysis or using a non sampling approach.

An important feature on the routes division is the use of a homogeniser which is the trigger for the 1^{st} step as seen on figure 2.1. As we saw on chapter 1 the homogeniser can be a blind tee configuration but can also be a mechanical device installed inside the pipe. In principle it is better to avoid the second type of configuration since we perturb the flow and we can have a pressure drop inside the pipe which has to be avoided for maximizing the flow rate. In the case of a non homogeneous flow there are different phase velocities and so we need to measure the velocity of each of these phases in order to analyze each stream individually. On the other hand if we have a homogeniser we have to use a set of sensors to analyze the flow properties such as density, velocity, momentum and mass flow and we need to perform a measurement on the mixture properties. To measure the mixture properties we can extract a sampling from the flow to analyze the different phases ratio or we can use a set of different sensors to analyze the flow without interfering. These sensors have to be coupled with the flow properties measurement in order to measure the individual phases of oil or water and they use a wide range of fluid properties to accomplish that.
In our case we choose to use a homogeniser in order to reduce the complexity of the problem and we preferred not to use a sampling extraction since this type of analysis do not give a real time information on the flow and requires more maintenance work and cost. So for our analysis we will discuss in detail the route number 2 by analyzing the flow properties and measure the individual phases using permittivity measurement. The other routes will not be discussed in

this thesis, further information on these approaches are reported on [5].
Now we focus on the measuring technique for our case and the type of data we can measure to characterize the flow. The data at which we can have access in a homogenised flow are [5]:

- **Density:** for measuring the total density of the flow we can use the absorption of neutron or gamma ray or instead measure the weight of a section of the pipe
- **Velocity:** can be measured using a neutron source, an electromagnetic field or using the cross correlation for many type of quantities such as gamma ray and neutron absorption, acoustic waves, conductivity or impedance
- **Mass Flow:** is measured using a vibrating tube and using the Coriolis effect
- **Momentum flux:** the best and most robust device is a Venturi which uses a differential pressure to measure the momentum

We will now look in detail each type of measurement and available sensor.

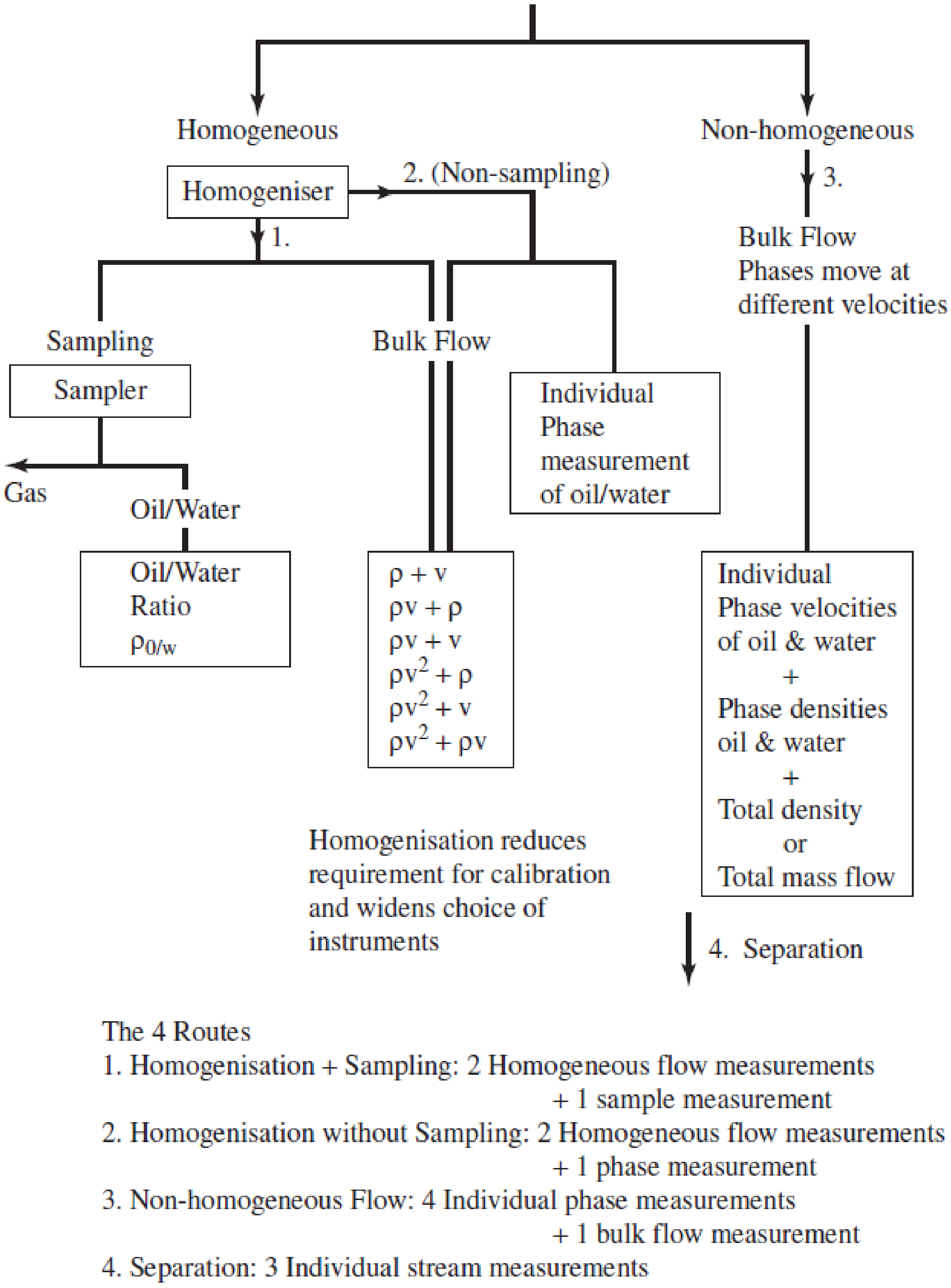

Figure 2.1: 4 routes for multiphase flow metering [5]

2.1.1 Density measurement

There are three methods to measure the density of the flow: weighting of the tube, gamma ray and neutron absorption. The first method weights the quantity of fluid inside a portion of the pipe, this can be done by measuring the deflection of a horizontal pipe and relate it to the weight of the pipe assuming a spring behavior of the tube.
The other two methods use the absorption properties of the fluid inside the pipe to measure the density, these techniques can be used with a collimated beam, Fig. 2.2, or a broad beam. For a gamma ray there are three effects for which a photon interact with the matter: photoelectric effect, pair production and scattering Compton. During photoelectric effect a photon is absorbed by an electron and the electron escapes the atom with an energy equal to the difference between the energy of the incoming photon and the bound energy $E = E_{ph} - E_{bound}$. In a pair production the incoming photon creates a pair of electron and positron, the positron then decays in two photon of energy 0.51MeV. Since the energy of the two outgoing photons is much smaller than the energy needed for the pair production we can assume a complete absorption for these two photons in the medium. The scattering Compton occurs at higher incoming energies of the photon, in this case the photon interacts with an atomic electron and loses energy and is scattered so it changes direction.

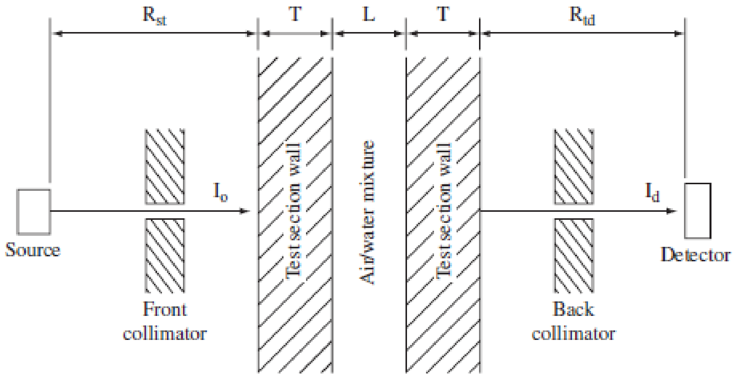

Figure 2.2: Typical arrangement of a single beam gamma ray densitometer [5]

The absorption of a beam with an initial intensity I_0 is determined by the law $I = I_0 e^{-\mu z}$ where μ is the linear absorption coefficient and z is the distance travelled inside the absorbing medium. In our case we have absorption due to the pipe wall and fluid inside the pipe. The absorption due to pipe can be considered constant over time and must be taken into account when calibrating the gamma ray densitometer while for the fluid we can write the contribution of each phase inside the pipe

$$I = I_0 e^{-(\mu_1 z_1 + \mu_2 z_2 + \mu_3 z_3 + \mu_a z_a)} \tag{2.1}$$

where with $\mu_i, \; i = 1, 2, 3$ we define the linear absorption coefficient of the different phases, for example water, oil and gas and z_i are the relative distances. μ_a is the air linear absorption coefficient for the distance source-wall and can be considered zero so after a calibration needed to consider the wall pipe contribution we can measure the intensity and this value will depend only on the fluid flowing inside the pipe.
The linear absorption coefficient is related to the density by $\mu = \mu_m \cdot \rho$ where μ_m is the mass attenuation coefficient. Now we have a formula to relate the measured intensity to the density but we have no information on the distances z_i so we cannot use the formula directly. To get

information on the density we need to calibrate the setup in the case where there is only oil or only water inside the pipe. In this two cases we get

$$\begin{aligned} I_1 &= I_0 exp(-\mu_w z_w) exp(-\mu_a z_a) exp(-\mu_1 D) \\ I_2 &= I_0 exp(-\mu_w z_w) exp(-\mu_a z_a) exp(-\mu_2 D) \end{aligned} \tag{2.2}$$

where D is the diameter of the pipe which is the distance travelled by the gamma ray considering a one phase flow and the subscript w indicates the pipe wall. By doing some algebra on these value we get

$$\begin{aligned} ln\left(\frac{I}{I_1}\right) &= (\mu_1 - \mu_2) z_2 \\ ln\left(\frac{I}{I_2}\right) &= (\mu_2 - \mu_1) z_1 \\ ln\left(\frac{I_1}{I_2}\right) &= (\mu_2 - \mu_1) D \end{aligned} \tag{2.3}$$

and if we consider a gas-liquid flow we can find that the gas fraction can be computed as

$$g_f = \frac{ln(I) - ln(I_l)}{ln(I_g) - ln(I_l)} \tag{2.4}$$

where I_i is the intensity measured when only the i phase is present inside the pipe. The density then can be computed as

$$\rho = g_f \rho_g + (1 - g_f)\rho_l \tag{2.5}$$

This is the computation for a single beam, if we use a broad beam gamma densitometer we can use this type of computation for each single channel and we obtain the final density by averaging the densities computed.

2.1.2 Velocity measurement

There are many methods to measure the velocity of a fluid in motion, we will explain a method which can be used with many measured quantities and its principles. The cross correlation technique consists in measuring the fluctuations of a quantity in two different positions and correlate the fluctuations of the measure between the two sensors. This can be seen in Fig. 2.3 where sensor number 1 measures a fluctuations and after a time τ which depends on the flow velocity and distance of the sensors the same fluctuations is measured on sensor 2. The velocity is then found as the distance between the two sensors divided by the time, $V = L/\tau$.

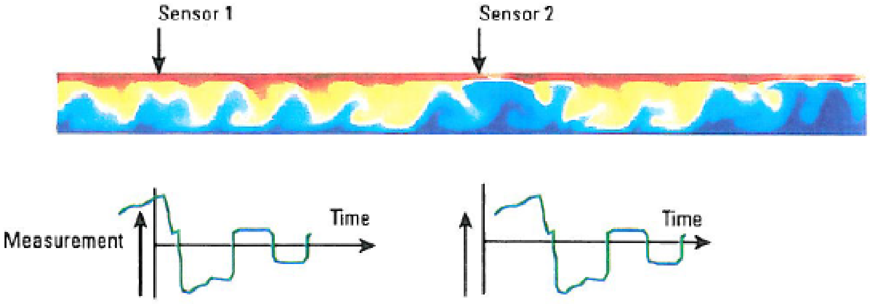

Figure 2.3: Typical design of a cross correlation sensor [6]

This method can be used with many quantities such as gamma ray absorption, acoustic waves, conductivity and impendance measurement. When doing this type of measurement we must pay attention to the acquisition rate of the measure and the sensors distance. In fact if we have a not fast enough acquisition rate we loose information on all the fluctuations which occur with a faster rate inside the pipe. The same can happen if the distance between the sensors is greater than the decay length of the fluctuations which we define as the distance a given fluctuation travels inside the pipe before changing shape.

2.1.3 Mass flow measurement

A setup used to measure the mass flow rate is the Coriolis , Fig. 2.4, this setup consists of a c-shaped tube and a T leaf spring which oscillates. The curve oscillates thanks to an electromagnetic forcer and the T oscillates in a plane normal to the curve. This type of oscillation forces each particle inside the pipe to have a Coriolis type acceleration, this acceleration hence creates a deflection on the c-shaped pipe which is inversely proportional to the stiffness of the pipe and proportional to the mass flow rate.
This deflection is measured using two optical pick off which creates a pulsed signal with width related to the mass flow.

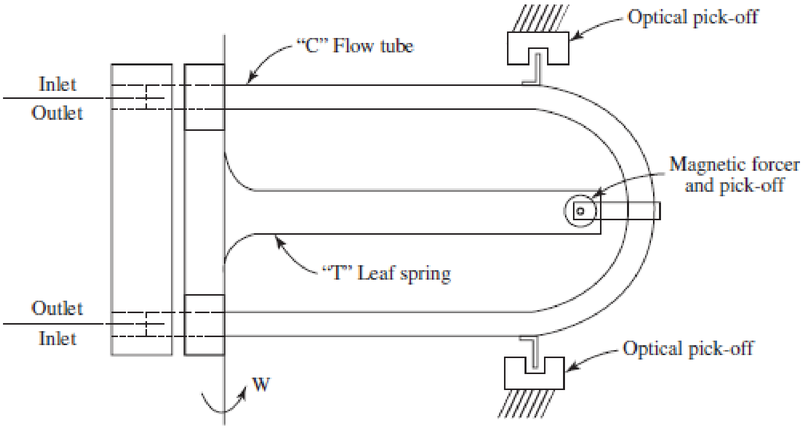

Figure 2.4: Typical design of a Coriolis [5]

2.1.4 Momentum flux measurement

The most robust and used configuration for measuring the momentum flux is a Venturi which uses the differential pressure in a shrinked portion of the throat of the pipe. Inside the throat the area of the pipe will decrease and so the velocity of the fluid must increase and the pressure decrease. These phenomenon is governed by the Bernoulli's equation:

$$p + \frac{\rho V^2}{2} = const \tag{2.6}$$

If we deal with a multiphase fluid the Bernoulli's equation may not work because of the slip velocity between the phases, in this case we need to consider another equation to balance the momentum difference of the different phases. This problem is overcome if a mixing device before the Venturi throat is used, in this case the velocities between the phases can be considered equal and since the Bernouilli's equation uses the density of the fluid we only need to measure it. As seen above in our configuration the mixing effect is made by the blind tee configuration.

2.1.5 Permittivity measurement

To compute the WLR the most useful information to use is the permittivity of the material, in fact this value is very different between oil and water phases. In the case of oil we have

$\epsilon = 2.1$ Fm^{-1} while for water it can vary around a value of $\epsilon = 80$ Fm^{-1}. The permittivity for a material has a real part related to the energy storage and an imaginary part related to the dissipation of the electromagnetic field.

$$\epsilon_m = \epsilon' - j\frac{\sigma}{\omega\epsilon_0} = \epsilon' - j\epsilon'' \tag{2.7}$$

where ϵ_0 is the permittivity of air.
The real part of the permittivity is also called loss factor and is associated to the capacitance behavior of the material while the imaginary part is associated to a conductive behavior. The sensors for measuring the conductivity are divided using the range of frequency at which they work. We have low frequency sensors which work in a range $30 - 300$ kHz, medium frequency 300 kHz $-$ 3 MHz, high frequency $3 - 30$ MHz and microwave sensors 30 MHz $-$ 30 GHz. We will present here briefly the different setup and the measurement principles, a further explanation on the microwave sensors will be done later on chapter 4.

Conductance and capacitance measurement

This type of measurement is associated to all the regimes outside the microwave region.
The fluid is composed of two phases made by water and oil, the single properties of the two fluids for the electromagnetic field are very different: water is a polar fluid and is conductive while oil behaves as an insulator. These two different behaviors help us to understand why we need two sensors to compute the WLR. In the water continuous regime we have a fluid which has a conductive part as medium so we measure the conductance of the fluid to get the permittivity while in oil continuous the medium behaves as an insulator so we need to measure the capacitance of the fluid to achieve information on the permittivity. So the setup must be able to switch between these two configurations. These configurations can have three major problems: the presence of gas inside the pipe as shown in Fig. 2.5, the inversion region and the possibility of abrasion and contamination of the electrodes. In the first case the gas presence alters the capacitance and conductivity of the fluid creating an excluded region for this measurement. To avoid this effect we can correct the measure of the permittivity by doing a posterior analysis to include the gas presence. But in order to have an exact correction we need to know the exact fraction of gas inside the flow. The second problem is difficult to manage since the continuous phase in the inversion region is not well defined and so these methods do not work. For this reason these sensors cannot cover the entire range of water cut measurement but works only in a defined continuous regime. The third problem can be a major issue because in oil wells we have presence of hydrates inside the flow which can attach to the electrodes and alter significantly the measure.
For the water continuous case exists another method to measure the conductivity which consists of two toroidal solenoids around the pipe. The first solenoid is used to create a current which then is measured using Lenz law by the second solenoid. Also in this case we must take into account the presence of gas inside the pipe since it alters the measure.

Microwave resonators

When we work with microwave sensors we have a very high frequency rate and in Eq. 2.7 we have that $\sigma \approx \omega(\epsilon_0\epsilon')$ and so $\epsilon' \approx \epsilon''$. The frequency shift is mainly due to the permittivity of the mixture while the absorption of the signal is related to the conductivity. Using the frequency shift we can measure the permittivity but we need to build a model for the water fraction. Many models are used and the most famous are Bruggemann and Rao and Ramu but these models are not invertible so we cannot compute directly the WLR.

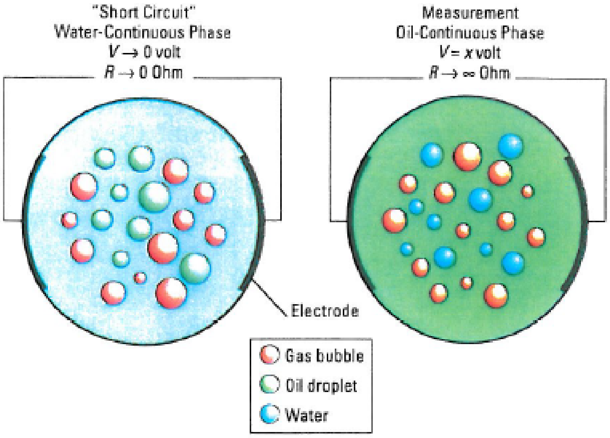

Figure 2.5: Section of a pipe with the electrodes and the two regime [6]

At these frequencies we can measure the signal shift in frequency and the attenuation of the transmitted or reflected signal. This type of sensors can work either in water continuous and in oil continuous and has no constraints on the flow regime.
Besides, this configuration presents also some limits: it undergoes erosion, the measure is altered by dirt presence and we can measure only a fraction of the liquid phase so we need another setup for measuring the gas fraction. Although these problems this configuration is used widely in many applications and many type of microwave resonator exist.

We now analyze another type of setup which is used for describing the multiphase flow and which uses some properties of the techniques described above.

2.1.6 Tomographic measurement

The tomography technique is used to visualize the flow inside a pipe using an array of sensors and can give information on the density and the fraction of each phase present. There are different sensors used in tomography to analyze the flux and we can divide them in 3 categories

- nuclear based non ionising;
- nuclear base ionising;
- non nuclear.

In the first class we can find elements such as gamma ray and X ray tomography while for the second class we have the nuclear magnetic resonance and the last class includes the electrical capacitance, the optical, the ultrasonic and microwave tomography respectively. The principle at the basis of tomography is to measure a given properties of the fluid and to correlate all the signals coming from the different sensors. The signal correlation can be done with sensors at the same height along the pipe, at different height and even with multiple adjacent sensors. In the earlier days of this technique its application for industrial use was limited due to the high cost for the array sensors and the high computational cost. Nowadays the problem for the high computational costs do not exist anymore but still these setups have a high cost and maintenance problems since an higher number of sensors leads to a major probability of breaking.

Chapter 3

Mixture permittivity

Before introducing the models for the mixtures permittivity we draw some background informations on what is the permittivity of a material and how its properties changes with the frequency.
The permittivity is a value which quantifies how much a material will polarize itself when an external electric field is applied. The response of the material is indicated using a complex numbers

$$\epsilon = \epsilon' + i \cdot \epsilon'' \tag{3.1}$$

The real part of the permittivity is associated to the energy that the material store when it polarizes, in particular if the real part is positive the material will store energy against the external field while if the real part is negative the polarization will increase the external field.
The imaginary part is related to the dissipation of energy by the material and is proportional to the conductivity so for an insulator the imaginary part of the permittivity is 0.
Since the permittivity depends on the capability of the material to polarize, it will depends also on the frequency of the field applied since at different frequencies we have different type of response. We can divide the response of the material in four different regions according to the scale of the mechanism involved (Fig. 3.1):

- electronic polarisation;
- atomic polarisation;
- orientational polarisation;
- interfacial polarisation.

The electronic polarisation occurs at very high frequencies ($\approx 10^{16}$Hz), at this frequency the material is not able to polarize since the time scale is very low. The only response that occurs is for the electronic scale, in this case the response can be seen as the displacement of the electronic clouds around a nucleus. For example this effect can be observed, undisturbed by other effects, in noble gases.
The atomic polarisation ($\approx 10^{13}$Hz) involves the scale of the molecules of the material. When a bond is formed the shared electrons are, usually, not equally distributed between two different atoms, in this case the external field will modify these distributions and the atoms in the molecules acquire charges. This is the example of the carbon dioxide molecule.
The orientational polarisation ($\approx 10^{10}$Hz) occurs when a molecule already posses a dipole moments in the absence of an electric field, such as the molecule of water. In this case the molecules tend to align its dipole moment along the electric field and so all the molecule is interested in this movement. Since the molecule has to orient its dipole moment the time scale is much slower

respect to the previous cases. At the end we have a medium with dipole moments aligned with the electric field and some oscillations of the dipoles due to the thermal noise.
The last case is the less important for homogeneous medium but can have some useful application in biology. Even if a medium is non conductive it can often have some mesoscale structures which can be conductive. In this case when a low frequency field is applied these structures can behave as polarisable macromolecules. When the field is applied the charge is accumulated at the interfaces of the conducting and non conducting region and so affect the permittivity measurement.

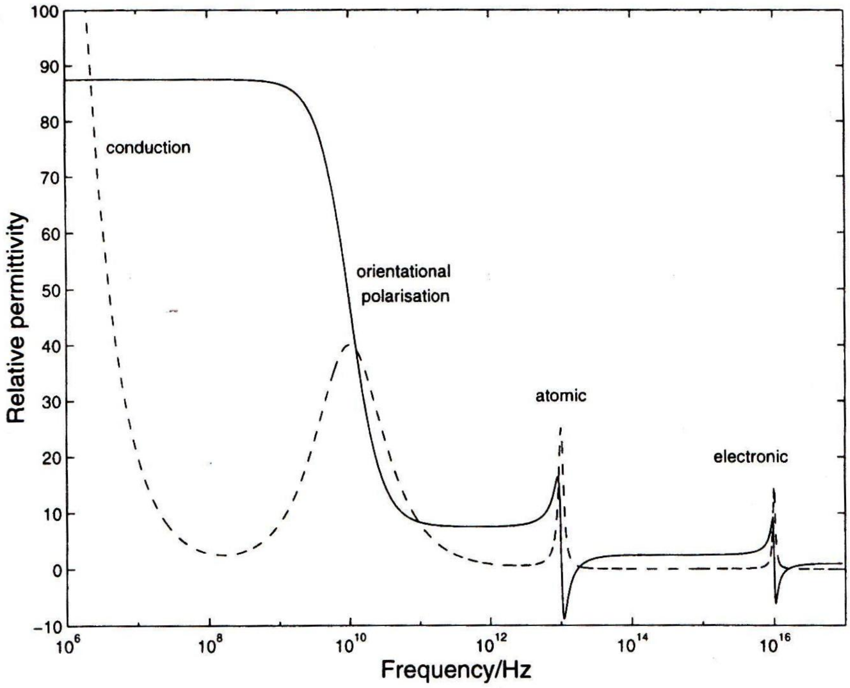

Figure 3.1: Permittivity dependence on frequency [8]

3.1 Classical mixing

When we consider a heterogeneous medium we would like to derive its properties starting from the properties of the single component inside the medium. This can be easy in some cases such as the density where, if we consider f the volumetric fraction of a single component, can be computed simply using $\rho = \sum \rho_i f_i$. For some other properties this approach does not work and that is the case of the permittivity of a mixture.
To derive the permittivity of a mixture we can use many different models and assumptions such as the shape of the inclusions or even use a non linear response of the polarisability. For our purposes we will not need some advanced mixing models but we will focus on simpler models assuming a spherical shape of the inclusions.
The term **Inclusions** refers to the droplets of a phase which are surrounded by a **Medium**. The medium is also called host, matrix or environment and represents the dominant phase in our mixtures. The volume fraction is the volume occupied by the inclusions, if the inclusions occupy a volume f, the host occupy a volume $1 - f$.
We will now introduce the most known mixing model and then compare the results we can obtain using them.

3.1.1 Maxwell Garnett rule

The Maxwell Garnett model derives the permittivity of the mixtures using an effective medium approach which describes the macroscopic effects of the material by averaging the microscale effects of the components. In this approach we relate the effective permittivity to the average effect on the field and the flux density so we use.

$$< D >= \epsilon_{eff} < E > \tag{3.2}$$

By weighting on the volume fraction and using the permittivity of the inclusions, ϵ_i and environment ϵ_e we get the relations.

$$\begin{aligned} < D >&= f\epsilon_i E_i + (1-f)\epsilon_e E_e \\ < E >&= fE_i + (1-f)E_e \end{aligned} \tag{3.3}$$

Using these equations we get that the effective permittivity is

$$\epsilon_{eff} = \frac{f\epsilon_i A + \epsilon_e(1-f)}{fA + (1-f)} \tag{3.4}$$

where A is the ratio between the internal and external field, $A = E_i/E_E$.
To compute A we must consider the response of a single sphere to the external electric field but let's suppose we know the result $A = 3\epsilon_e/(\epsilon_i + 2\epsilon_e)$, if we substitute in the equation we get

$$\epsilon_{eff} = \epsilon_e + 3f\epsilon_e \cdot \frac{\epsilon_i - \epsilon_e}{\epsilon_i + 2\epsilon_e - f(\epsilon_i - \epsilon_e)} \tag{3.5}$$

This formula is known as the Maxwell Garnett mixing formula and we can notice that in the limit of no inclusions: $f \to 0$ we get $\epsilon_{eff} = \epsilon_e$ and in the limit $f \to 1$ we get $\epsilon_{eff} = \epsilon_i$ as we would be expect.
The Maxwell Garnett formula can be generalized for multiphase mixtures in which there are more than one type of inclusions but we still recognize a host.

$$\epsilon_{eff} = \epsilon_e + 3f\epsilon_e \cdot \frac{\displaystyle\sum_{k=1}^{N} f_k \frac{\epsilon_k - \epsilon_e}{\epsilon_k + 2\epsilon_e}}{1 - \displaystyle\sum_{k=1}^{N} f_k \frac{\epsilon_k - \epsilon_e}{\epsilon_k + 2\epsilon_e}} \tag{3.6}$$

Although (3.6) gives the effective permittivity directly another method can be more convenient and consists on using the mixing formula for two phases repetitively. In the last case we treat each inclusions alone by considering all the other as hosts and then we obtain the same result for a multi inclusions model.
This model is the basis for many others models and can be modified to include the ellipsoidal shape of the inclusions and their orientation. We are not interested in this type of analysis but we will discuss the case in which the inclusions are made of lossy material as is the case of water.
Until now we have treated the material as pure dielectrics in which there is no charge flow but this is not the case for conductive materials. In this case we have already seen that the permittivity is written as a complex variable $\epsilon = \epsilon' - j\sigma/\omega\epsilon_0$ in which the imaginary part depends on the conductivity of the material. This case is easy to implement on equation (3.5) since we simply have to write the permittivity as a complex parameter. The effective permittivity can have a different response according to a small imaginary part or a large imaginary part as we can see in figures 3.2 and 3.3.

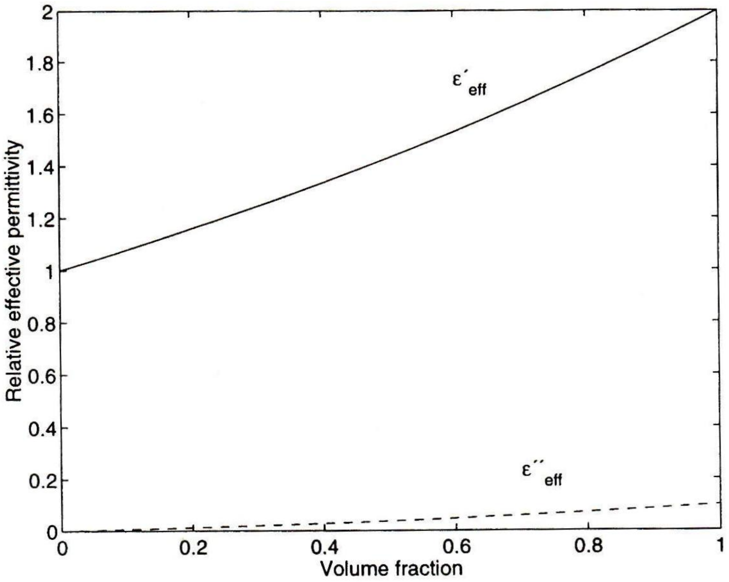

Figure 3.2: Effective permittivity with lossy inclusion, ϵ'' small [8]

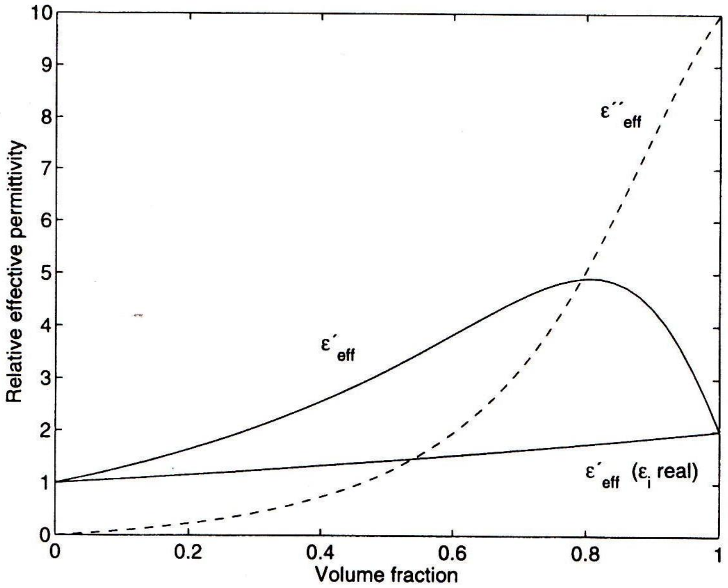

Figure 3.3: Effective permittivity with lossy inclusion, ϵ'' big [8]

In both pictures we are using a pure dielectric as environment (only real part of ϵ_e) but in the first picture we are using a ratio $\epsilon_i/\epsilon_e = 2 - 0.1j$ and in the second we are using $\epsilon_i/\epsilon_e = 2 - 10j$. In the first picture the presence of ϵ_i'' does not alter the result for the real part of ϵ_{eff} while in the other case we can notice two different behaviors. The real part of the permittivity increases with a maximum before decreasing to the permittivity of the inclusion while the imaginary part increase until it reaches the imaginary part of the inclusion.
Since in our case we have a mixture of oil, which has only the real part of the permittivity, and water, which for the frequency range we are using has an imaginary part of the permittivity of the order of the hundreds, we will have a behavior like Fig. 3.3.

It is important to specify an approssimation used. Up to now we have used a static external electric field to derive the Maxwell Garnett formula, but this may not be the case in many applications where we use it to find the permittivity. The formula for the permittivity is still valid in a time dependent field description because a high frequency field can be considered constant in a small region at a given instant. The question is then under which condition a time dependent field breaks our model for the permittivity and how large the spatial variation of the field in our medium is. A simple answer to these questions is to consider a field as quasi-static if it satisfies the equation $\lambda/2\pi > \delta$ where δ is the size of the inclusions in our medium and λ is the wavelength of the field. Using this rule we approximate the field around an inclusion as constant and we extend the range for the effective permittivity formula.

The Maxwell Garnett formula presents some limitations in its construction, for example the permittivity is found using the polarizability of a single scatterer by considering the surroundings as a homogeneous medium with a uniform polarisation density. Of course this is an approximation which does not take into account the heterogeneous distribution of the scatterers and their influence is not simply an average polarisation density. Another limitation we have already seen is the quasi static approximation. We saw that in order to have meaning the formula must be applied only in cases where the size of the scatterers is smaller than the wavelength of the field. Although this is known, it is not clear how to define the size of the scatterers: in a polydisperse medium the scatterers will have different sizes and we can have also situation in which the permittivity varies in a continuous way and so we should abandon the concept of inclusions.
These limitations led to new models to compute the effective permittivity of a mixture and we will discuss the most used in the sections below.

3.1.2 Bruggeman formula

The Bruggeman formula is also known in material science as effective-medium model and is different from the Maxwell-Garnett formula since it does not consider a hierarchy between host and inclusions but treats the phase in the mixture in an equal way. In this vision the homogenised medium is considered as the background in which polarisations are measured. For each phases f_j with sums equal to 1 and considering spherical inclusions we have

$$\sum_{j=1}^{N} f_j \frac{\epsilon_j - \epsilon_{eff}}{\epsilon_j + 2\epsilon_{eff}} = 0 \tag{3.7}$$

and we can rewrite it for a two phases mixture as

$$(1-f) \cdot \frac{\epsilon_e - \epsilon_{eff}}{\epsilon_e + 2\epsilon_{eff}} + f \cdot \frac{\epsilon_i - \epsilon_{eff}}{\epsilon_i + 2\epsilon_{eff}} = 0 \tag{3.8}$$

From this equation we can see that the two phases are symmetrical, in fact if we make the substitutions $f \to (1-f)$, $\epsilon_i \to \epsilon_e$ and $\epsilon_e \to \epsilon_i$ we get the same equation as before. This symmetry properties is the distinction between the Maxwell Garnett formula and the Bruggeman formula.
This formula is not explicit for ϵ_{eff} but we can find the result for it using some computation work by varying a given ϵ_{eff} and computing (3.8) until we find the value for which it is equal to zero.

3.1.3 Coherent Potential formula

This formula is similar in the derivation to the Maxwell Garnett formula but uses the coherent potential approximation to compute the permittivity. In Maxwell Garnett formula the dipole

moment and local field were found using a single scatterer isolated surrounded by the background. In this technique the polarisation density used to compute the permittivity is not the background but it has to be the one of the effective medium.
Using this procedure we get the equation

$$\epsilon_{eff} = \epsilon_e + f(\epsilon_i - \epsilon_e) \cdot \frac{3\epsilon_{eff}}{3\epsilon_{eff} + (1-f)(\epsilon_i - \epsilon_e)} \tag{3.9}$$

This equation is implicit in terms of the effective permittivity and we can notice it is similar to the Maxwell-Garnett.

3.1.4 Unified mixing rule

We have seen three models to compute the effective permittivity, each with its different properties and derivation but we wish to unify them under one equation. This idea is supported by the fact that the result for $f = 0$ and $f = 1$ using the different models must be equal and since we notice that in the limit $f \ll 1$ we get the same first order approximation for all the models, $\epsilon_{eff} \approx \epsilon_e + 3f\epsilon_e(\epsilon_i - \epsilon_e)/(\epsilon_i + 2\epsilon_e)$.
This unification can be done by introducing a new variable ν which can assume the values $\nu = 0, 2, 3$ and we can collect all the models in one single equation.

$$\frac{\epsilon_{eff} - \epsilon_e}{\epsilon_{eff} + 2\epsilon_e + \nu(\epsilon_{eff} - \epsilon_e)} = f \cdot \frac{\epsilon_i - \epsilon_e}{\epsilon_i + 2\epsilon_e + \nu(\epsilon_{eff} - \epsilon_e)} \tag{3.10}$$

And we can retrieve the models using ν

- $\boldsymbol{\nu = 0}$: gives the Maxwell Garnett formula
- $\boldsymbol{\nu = 2}$: gives the Bruggeman formula
- $\boldsymbol{\nu = 3}$: gives the Coherent potential formula

In Fig. 3.4 we can see the effective permittivity computed with the three different models and using a ratio for the permittivity of $\epsilon_i/\epsilon_e = 3$ and $\epsilon_i/\epsilon_e = 20$. We can notice that when the ratio between the inclusion and the environment is small the three models are similar but when the ratio is high the Maxwelll Garnett formula behaves in a different way from the others two models. In our case the ratio between water permittivity and oil is around $\approx 30 - 40$ so we are in the second case and in our analysis we should not use the Maxwell Garnett formula.

3.1.5 Differential mixing models

Another approach to compute the effective permittivity is based on differential analysis. When we use this type of derivation we get a power 1/3 in the formula which helps us to recognize what type of analysis is done. The simplest mixing rule is the Looyenga formula which is simply an average permittivity of the inclusion and environment using the permittivities with a power 1/3.

$$\epsilon_{eff}^{1/3} = f\epsilon_i^{1/3} + (1-f)\epsilon_e^{1/3} \tag{3.11}$$

This equation has the same symmetry properties of inversion between environment and inclusion as the Bruggemann.
Another type of equation which is derived by differential analysis are the Bruggeman non symmetric formulas

$$\frac{\epsilon_i - \epsilon_{eff}}{\epsilon_i - \epsilon_e} = (1-f) \cdot \left(\frac{\epsilon_{eff}}{\epsilon_e}\right)^{1/3} \tag{3.12}$$

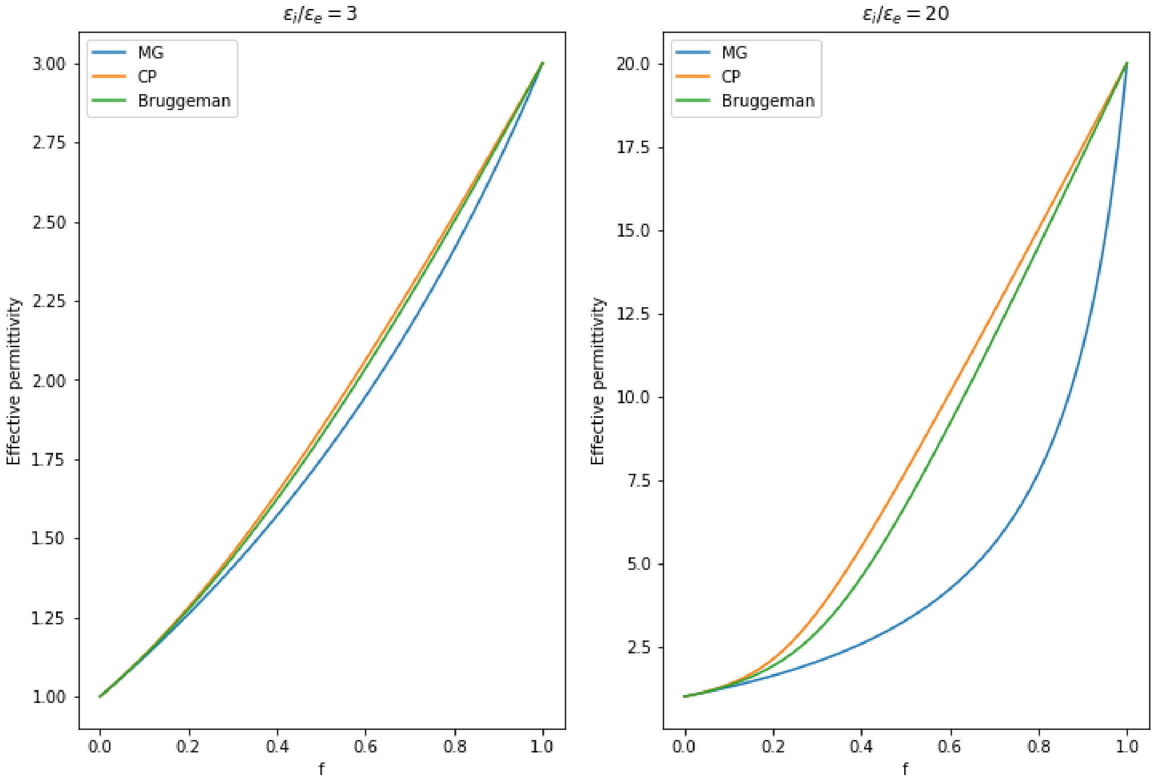

Figure 3.4: Comparison of the three models with different ratio permittivity

$$\frac{\epsilon_{eff}-\epsilon_e}{\epsilon_i-\epsilon_e} = f\cdot\left(\frac{\epsilon_{eff}}{\epsilon_i}\right)^{1/3} \tag{3.13}$$

The first formula was derived by Bruggeman [9] while the second was derived by Sen, Scala and Cohen [10].
They are "complementary" since the the first equation deals with inclusion ϵ_i in an environment ϵ_e while the second deals with the case in which we have the substitution $f \to (1-f)$, $\epsilon_i \to \epsilon_e$ and $\epsilon_e \to \epsilon_i$. Also this model can be unified to the unified mixing model. To see how we can include this model we expand (3.12) centred in $f=0$

$$\epsilon_{eff} = \epsilon_e + 3f\epsilon_e\cdot\frac{\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e} + 3f^2\epsilon_e\left(\frac{\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e}\right)^2\frac{2\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e} + \ldots \tag{3.14}$$

and if we expand (3.10) around $f=0$ and keeping the variable ν we get

$$\epsilon_{eff} = \epsilon_e + 3f\epsilon_e\cdot\frac{\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e} + 3f^2\epsilon_e\left(\frac{\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e}\right)^2\left(1+\nu\cdot\frac{\epsilon_i-\epsilon_e}{\epsilon_i+2\epsilon_e}\right) + \ldots \tag{3.15}$$

So we can notice that the two equations are the same up to the second order if we set ν equal to 1.
In the next figure we can see the previous Fig. 3.4 where we add the result for the permittivity using the Bruggeman non symmetric model. We can see that this model stands between the MG models and the other two.
It is worth noticing that the red line is computed using only the Bruggeman non symmetric formula so it describes a medium in which the inclusions are spherical and grows in size until they cover all the volume of the medium. In our case where we have oil inclusions in a medium of water this phenomenon cannot happen, in fact due to the mixing we have a range of volume fraction in which from the water inclusion case we pass to a system in which we have water as medium and oil as inclusion. When this happens we must use the Sen formula to derive the

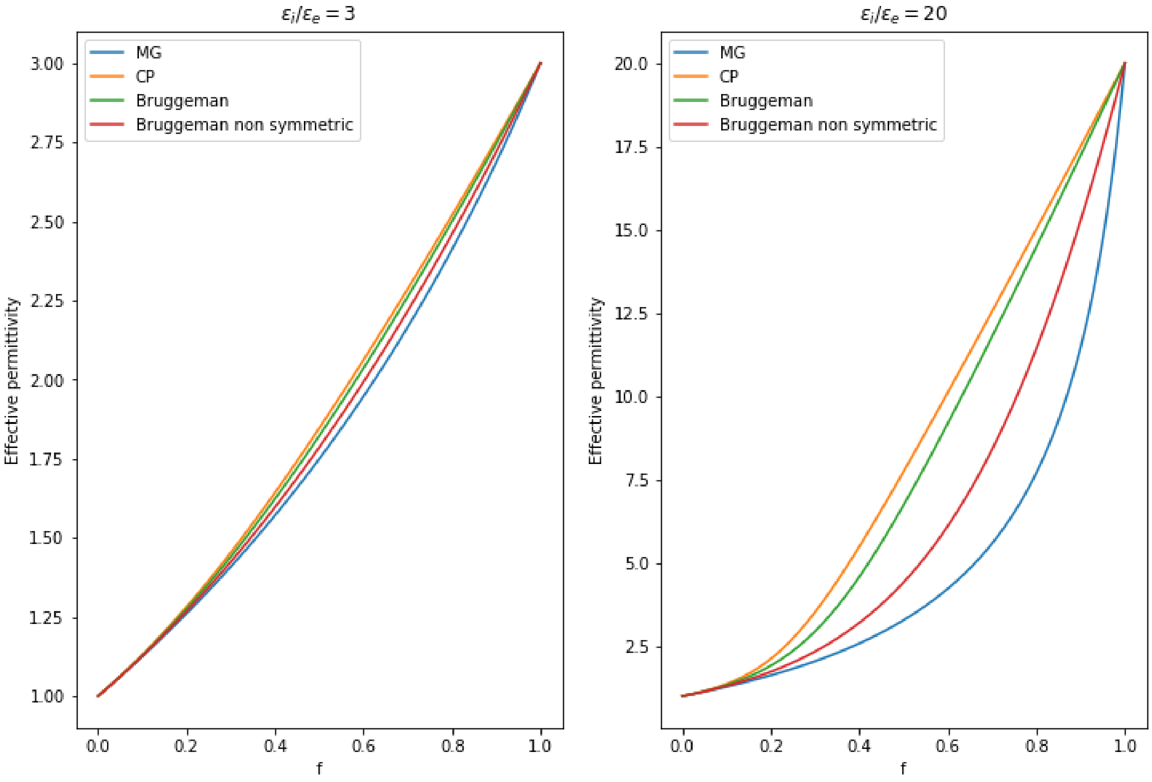

Figure 3.5: Comparison of the four models with different ratio permittivity

effective permittivity. In Fig 3.6 we can see the behavior of the two differential mixing models, in our case when we have the inversion region and we need to pass from a description to the other there is a jump in the permittivity the mixture can have. These forbidden values of the permittivity are greater when the ratio between the inclusions and the environment is higher, Fig. 3.6. This effect was also seen in the calibration and loop measurement and we will return and explain them in chapter 5.

In the book by Sihvola [8] the author performs simulations of a two dimensional mixtures and compute the permittivity using a Finite Difference Time Domain (FDTD) in order to compare the different models with a reference. In the simulation he uses a ratio $\epsilon_i/\epsilon_e = 16$ so we have a situation similar to the pictures on the right shown before. The models displayed are derived from (3.10) using $\nu = 0, 1, 2$ so we have the Maxwell Garnett model ($\nu = 0$), the Bruggemann non symmetric ($\nu = 1$) and the coherent potential model ($\nu = 2$). From the results we can notice that the best model which interpolates the data is the Bruggeman model even if overextimate a little the simulation results. For this reason and for the explanation of the forbidden region for the effective permittivity of the mixtures we will use the differential mixing models in our analysis.

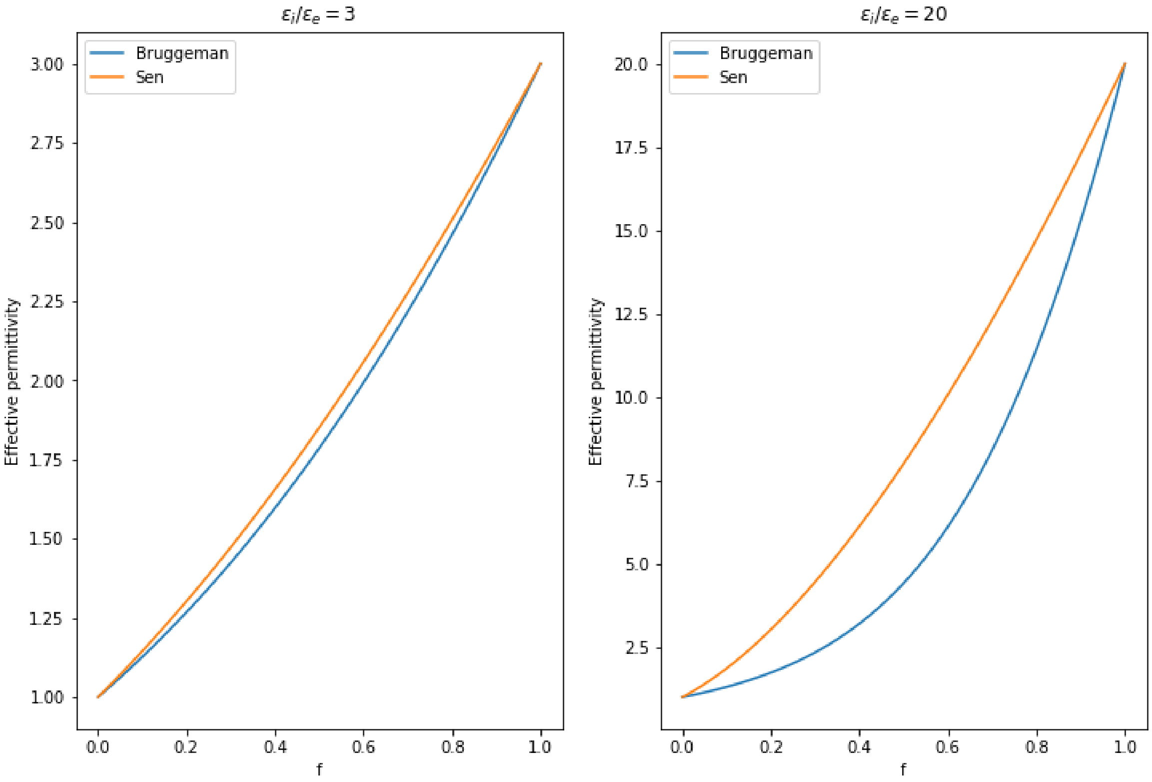

Figure 3.6: Comparison of the differential mixing models with different ratio permittivity

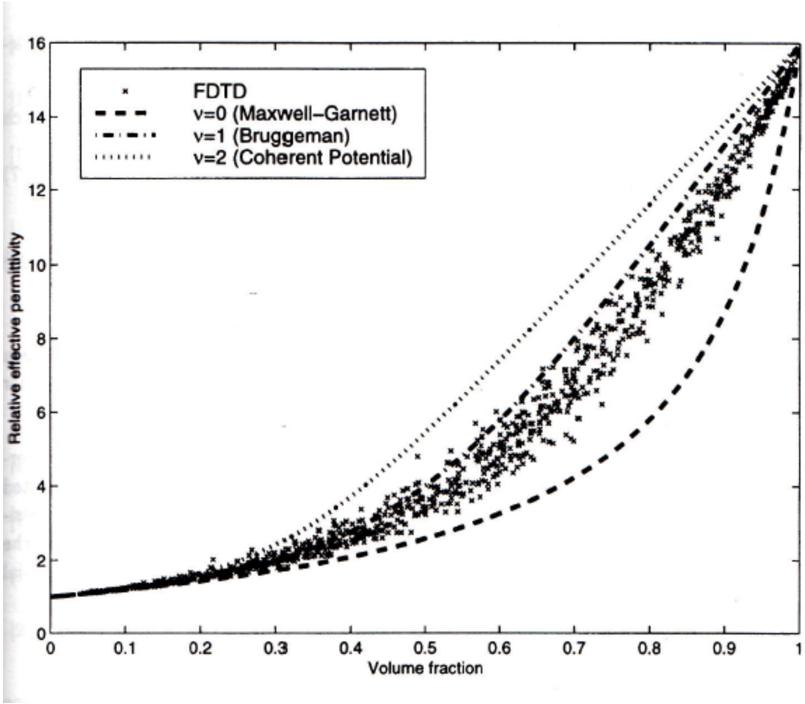

Figure 3.7: Simulations of a two dimensional mixtures using FDTD

Chapter 4

Microwave resonators

In this chapter we will briefly discuss the different type of resonators and how they works and then we will define in a better way the type of sensors we had. These sensors present some advantages respect to other configurations:

- they do not need mechanical contact with the medium and so they can perform measurement from a given distance;
- microwaves penetrates all the volume of the medium and so we measure all the material and not the surface;
- the resonant frequency depends only on the length of the cavity so they are very stable;
- at low frequencies the dc conductivity dominates the electrical properties of the material but in microwave frequencies we lose this influence;
- they have a fast acquisition rate;
- they measure water with an high contrast respect to the other materials.

In particular the last properties is the best suited for our type of experiment. Resonators sensors cover a wide range of configurations but their working principle is to measure the shift of the resonance peak when the sensors is in contact with a medium. So most of the resonators have a resonant cavity in which the electric field is stored and a dielectric surface which is usually in contact with the medium to analyze. Each sensor has a defined resonance and the presence of the medium is able to modify this peak and we are able to measure the shift in frequency, amplitude and phase of the curve.

4.1 Working principle

As said before each resonator sensor has a resonant cavity which can be filled with air or even with the medium to analyze. Inside this cavity an antenna transmits an alternating electric field at different frequencies and only the frequency which matches the intrinsic resonance frequency of the sensors is able to survive while the others are suppressed.
Since the electric field is a wave the only components of the field which survives inside the cavity are the ones for which half of the wavelength is a multiple of the length of the cavity.

$$l = n \cdot \frac{\lambda}{2} \qquad\qquad n \in \mathbb{N} \tag{4.1}$$

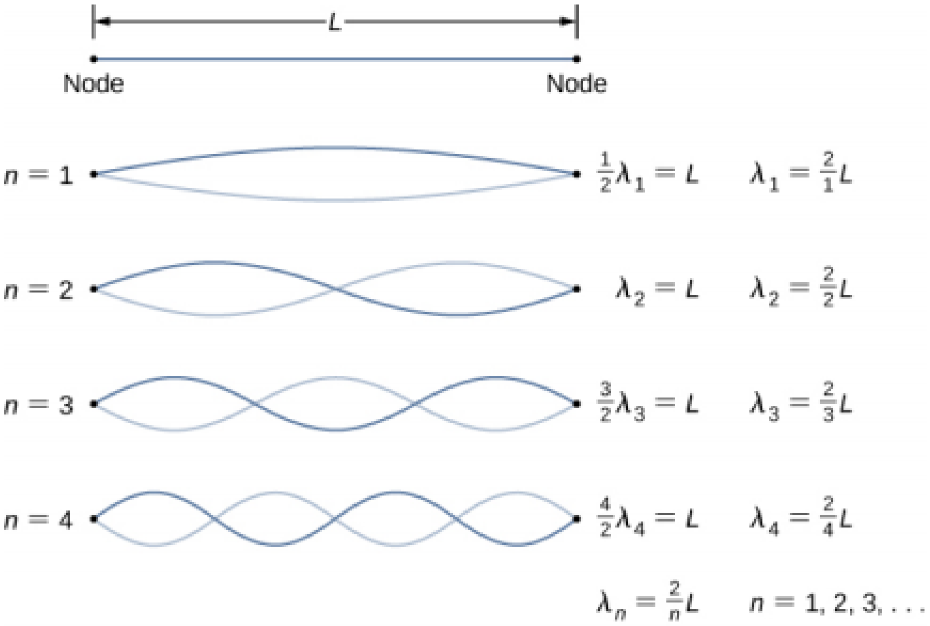

Figure 4.1: Standing wave solutions inside a cavity

So we can see that the first wave that satisfies this equation is the wave which has its first zero at the edge of the cavity. All the others frequencies not satisfying this equation corresponds to an electric field which loses strength and after few reflection are completely dumped inside the cavity.

The dependence of the amplitude respect to frequency is described by a Lorentzian curve:

$$y = y_0 + \frac{H}{(1 + ((fr - fr_0)/\mathrm{FWHM})^2)} \tag{4.2}$$

where FWHM is the Full Width at Half Maximum and fr_0 is the intrinsic resonance frequency of the cavity.

So with a resonator we can sweep the amplitude at various frequency and get the parameter of the curve. In Fig. 4.2 we can see some resonance curves from our mixer with different WLR.

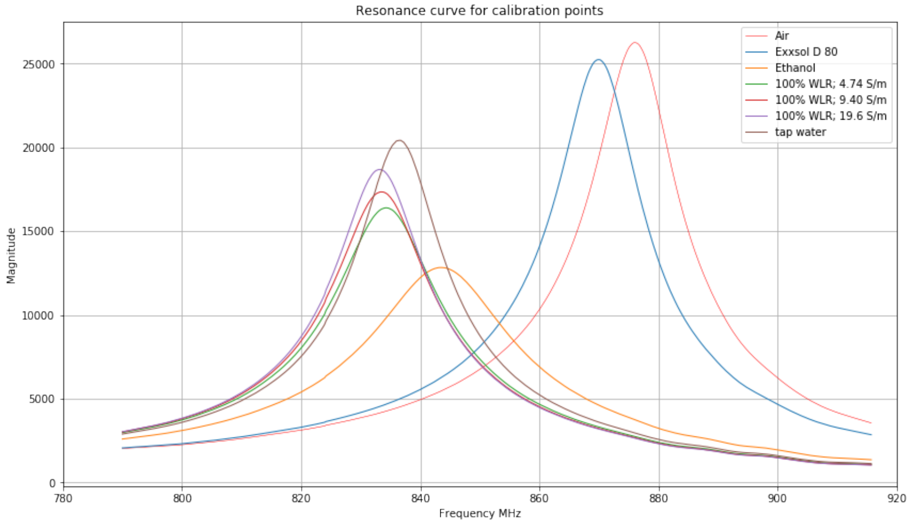

Figure 4.2: Resonance curve for different mixtures and conductivities

After this brief introduction we describe the different type of microwave resonators and their properties and applications.

4.2 Type of Microwave resonators

4.2.1 Free space Transmission Sensors

This configuration has two antennas, one transmitting the signal and one receiving the signal. The two antennas are placed on opposite walls of the pipe so the electric field passes through the flow of material and we can measure the shift in amplitude and phase of the signal. This configuration has the advantage of being very simple but it has also some other issues such as the reflection of the signal due to the pipe walls and the dielectric window or some other reflections in the interfaces of the flow under measurement. These reflections cause ripples on the frequency response and the amplitude is much more affected than the phase. So if the sensors need to measure only a parameter of the signal the best result is obtained using the phase instead of the amplitude. The signal reflections depends also on the type of material inside the pipe. If the material is lossy, the contribution of the reflection from the pipe is lower and so we can neglect it, instead if the material has small losses the pipe will affect strongly the measurement. The best solution to correct these reflections is to use a technique called frequency modulated continuous wave (FMCW) [7]. With this technique we are able to discriminate the signal in time and so we can exclude the reflections which arrives at the sensors sligthly later than the principal signal.

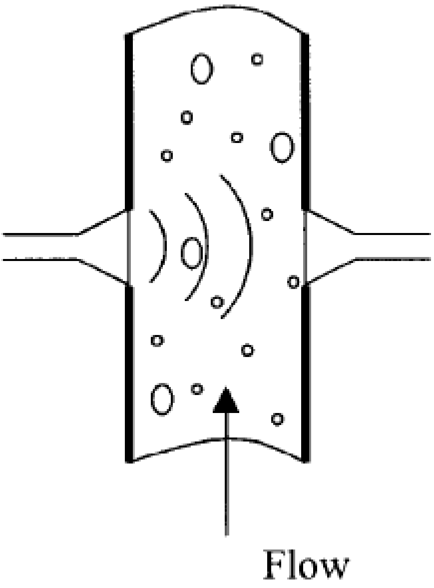

Figure 4.3: Free space transmission sensors configuration [7]

An application of this configuration is the microwave consistency analyzer (MCA) by Valmet Automation which is used to measure the pulp consistency in paper industry. This particular sensor is composed of two horns antennas inside the pipe and has a frequency sweep of $2-3$GHz.

4.2.2 Special Transmission sensors

A transmission sensor needs a a reference channel to make the phase measurement and usually the reference is external to the sensor. In this setup the phase reference is directed inside the material but in a different channel respect to the main channel. This setup has two receiving antennas, using this setup we can achieve more features respect than a single channel. For example if the distance between the receiving antennas is known the signals can be correlated to measure the velocity of the flow.

4.2.3 Guided Wave Transmission sensors

In this configuration the microwaves do not propagate freely inside the material but instead they are guided in a ring inside the pipe using a dielectric or a coaxial cable as seen in Fig. (4.4). This configuration has the advantage of having a better matching for impedance measurement but has some limits which make this setup not suited for all the type of analysis. The electric field is affected by the presence of the flow but since the microwaves moves inside the ring the sensible region detected is only a fraction of the total volume of the pipe. For this reason this setup gives good result only if the flow inside the pipe is homogeneous while for the case of the pulp measurement it is not suited since the pulp has a higher density and so for the gradient velocity it flows near the pipe walls.
Another type of problems this configuration can have is the sensibility to contamination. Since the field is located around the ring, if the ring presents some dirt like sand or other types of unwanted particles in the flow the measure can be compromised.

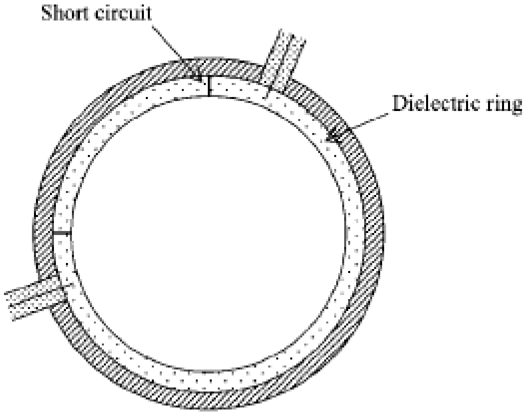

Figure 4.4: Guided wave sensors configuration [7]

4.2.4 Open ended resonators

This class of sensors is based on the reflection of the signal from the end of the transmission line which is in contact with the material. An example is the open ended sensor which is a coaxial cable in which the end is open and in contact with the material and so the resonance condition is affected by the external environment. This setup is widely used in laboratories and to analyze biological tissue and has the advantage of having a wide range of frequency at which it can work. The use of this setup for the oil industry is instead not recommended since the presence of bubble of gas in the sample affects the measurement. Instead for biological tissues we have a continuum of material so the open ended sensors works well. This configuration has the problem of contamination since the fringing field of the antenna does not cover the entire volume of the material and the presence of dirt on the end line of the sensor alters the measure.

Our microwave is an open ended type but instead of having the end of the coaxial cable in contact with the environment we have a dielectric cap. In the paper By Prafull Sharma, Liyun

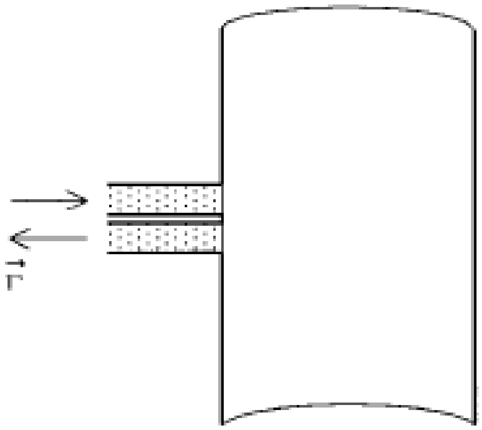

Figure 4.5: Open ended configuration [7]

Lao and Gioia Falcone [11] they present a similar type of microwave resonator from the one we had in our lab. In Fig. 4.6 we can see the setup for the sensor and its sensible region which is lower than the volume of the pipe. The flow of material affects the resonance of the cavity and so we can retrieve the permittivity from the shift of the resonance properties. On Fig. 4.7 we can see a simulation of the electric field inside the resonator, we can notice that the field is stronger near the antennas and is quite constant inside the cavity and then decreases in the medium. For this simulation they used as medium only oil so this is not a mixture and we cannot appreciate the effect that droplets with different permittivity have on the electric field.

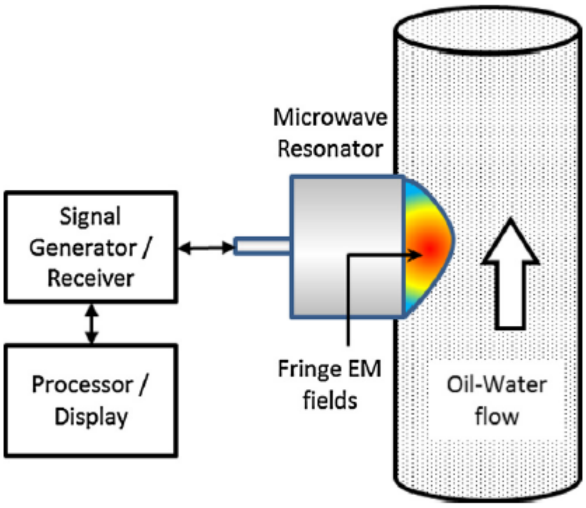

Figure 4.6: Configuration for a microwave resonator [11]

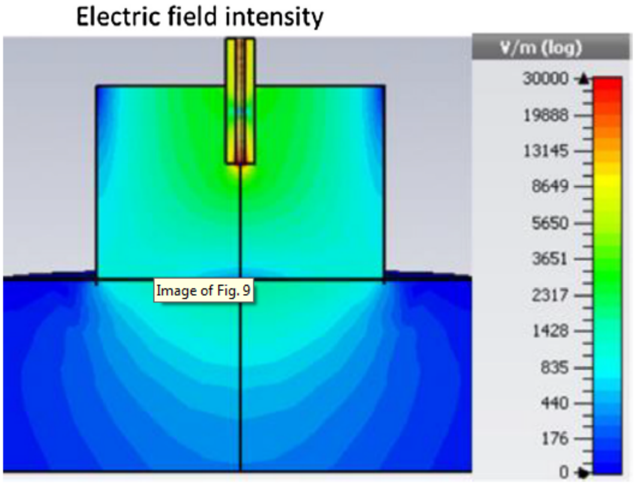

Figure 4.7: Simulation of the electric field inside the resonator [11]

Chapter 5

Analysis and data

5.1 Introduction

In this chapter we will analyze how the informations we gathered in the previous chapters can be used to compute the WLR of a multiphase mixture. We know that using a microwave sensors we have information on the permittivity of the fluid flowing inside the pipe and we need a method to relate the mixture permittivity with the WLR. From the Bruggemann mixing model we know how to compute the WLR by knowing ϵ_{mix}, the permittivities of the single phases and their fraction inside the mixture. The measure of the single permittivity is easy for the oil since it is an insulator and is a constant but it is more difficult for water since it is a conductor and depends on the temperature, conductivity and frequency of the electric field applied.
We will discuss the setups used, the measure and computation needed, the calibration of the probe and all the problems we encountered in order to get the correct result for the WLR.

5.2 Setups

In Fig. 5.1 we can see a photo of one of the probes used in this thesis which has a dielectric cap window of 10 mm and a cavity with no medium inside. This probe has two antennas which are connected to an electronic which has a frequency range of ≈ 200 MHz in the region of 0.8 GHz. This electronics is able to perform a frequency sweep each $2s$ and each measure last 3 minutes so for each datapoint we have around 80 resonance curves to analyze. These curves were used to create a representative curve for the specific point. In particular we analyzed two different curves: the first is the mean curve which is taken as the mean of the single curves, the second is what we call the max curve which for each single frequency value it takes the max amplitude. These two curves are used to find the curve parameters and compute the permittivity, we will discuss later the use of each curve and why one is better than the other in some situations.

Figure 5.1: Picture of a probe

We had two different setups for the data acquisition: a mixer and the loop. At the mixer we had a chamber which can contain a volume of 250 ml and a mixer used for the mixing with a velocity from 0 to 6000round/min. Using this setup we performed the calibration for each probe and we did a first analysis to understand which model could be used. The loop is the circuit shown in Fig. 5.2, we can notice the blind tee corner, as seen in Fig. 1.4, for the mixing and a Venturi throat in the vertical pipe. In order to control the flow inside the pipe we had two different tanks with Exxsol D80 and water with a volume of 1000 l. For each tank we had two pumps which were able to control the liquid flow rate on a range from 0.03 to 2 m^3/hr. The GVF of the flow was controlled using a valve. In order to simulate a real working condition we had another mixing device placed at the entrance of the circuit in order to achieve a fluid already mixed in the first horizontal pipe. In the loop setup we had different positions in which we could place the probes and we were able to acquire data from two different positions at the same time.
Using this loop configuration we had informations on these quantities:

- temperature;
- pressure;
- velocity;
- salinity of water.

The information on the salinity is related to the conductivity of the water and affects the permittivity: higher conductivities gives a lower permittivity. This value can be considered as an input in our analysis in order to correct the water permittivity in Bruggeman mixing model. Another useful input we will use on the analysis is the GVF to correct for the gas presence but we will see later in which situations this value is useful. Starting from this values and from the resonance curve we would like to achieve a resolution on the WLR measurement of 95 % confidence level inside the 5% of error.

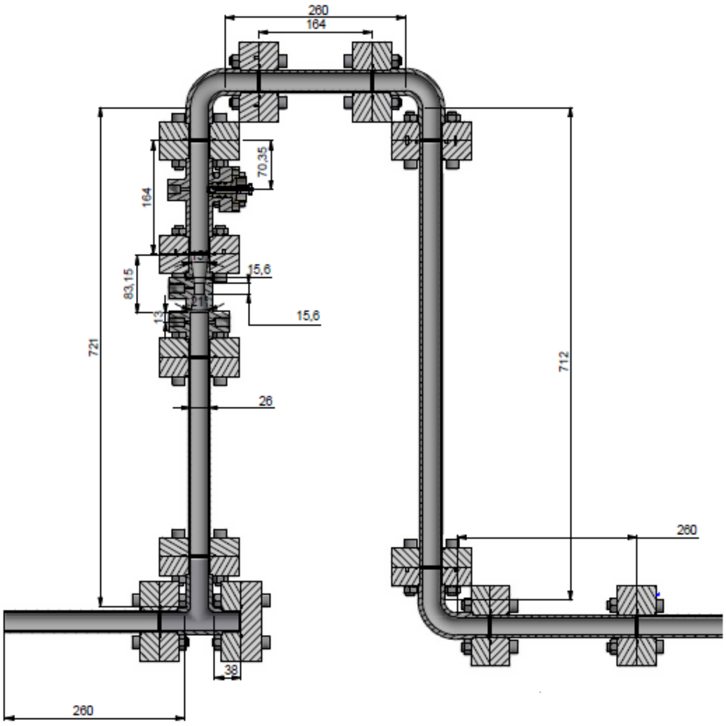

Figure 5.2: Loop circuit

5.3 Loop Data

In order to simulate the working condition of most of the oil wells we decided to divide the points acquisition in the loop according to GVF, WLR, gas superficial velocity and conductivity.
In Fig. 5.3 we can see a plot in which the points are categorized using GVF and WLR while the title of the plot represents the gas superficial velocity inside the pipe. These points were recorded at different conductivities from a range of tap water to 16 S/m. These points were selected in order to cover all the possible values of WLR from 0 to 100 and to have the GVF from 95 to 99.5 which corresponds to the wet gas regime. The points at 99.5 GVF will not be used later in the analysis since we found that the signal in this condition was very noisy and was impossible to extract. For this region of GVF from 99.5 to 99.9 exist another type of meters specific for this region which is called high GVF. On the lower GVF values we had some problems in the flow control since we reached the upper limit for the pump power so for example the measure at WLR 100 or 0 is missing and we were able to acquire points only in the range 30-70 WLR. As said before each point is taken for a time of 3 minutes and 1 minutes was needed to change flow but some problems may arise when we needed to record some points with lowest liquid single component flow rate. In this case the liquid flow rate can be very low and results in having a superficial velocity of 1 or 2 m/s so to avoid any type of data with acquisition while the flow is changing we waited 2 or 3 minutes before the acquisition.

After the introduction of the setups we used we now discuss how we can compute the water permittivity in order to use it in the Bruggemann mixing formula to compute the WLR.

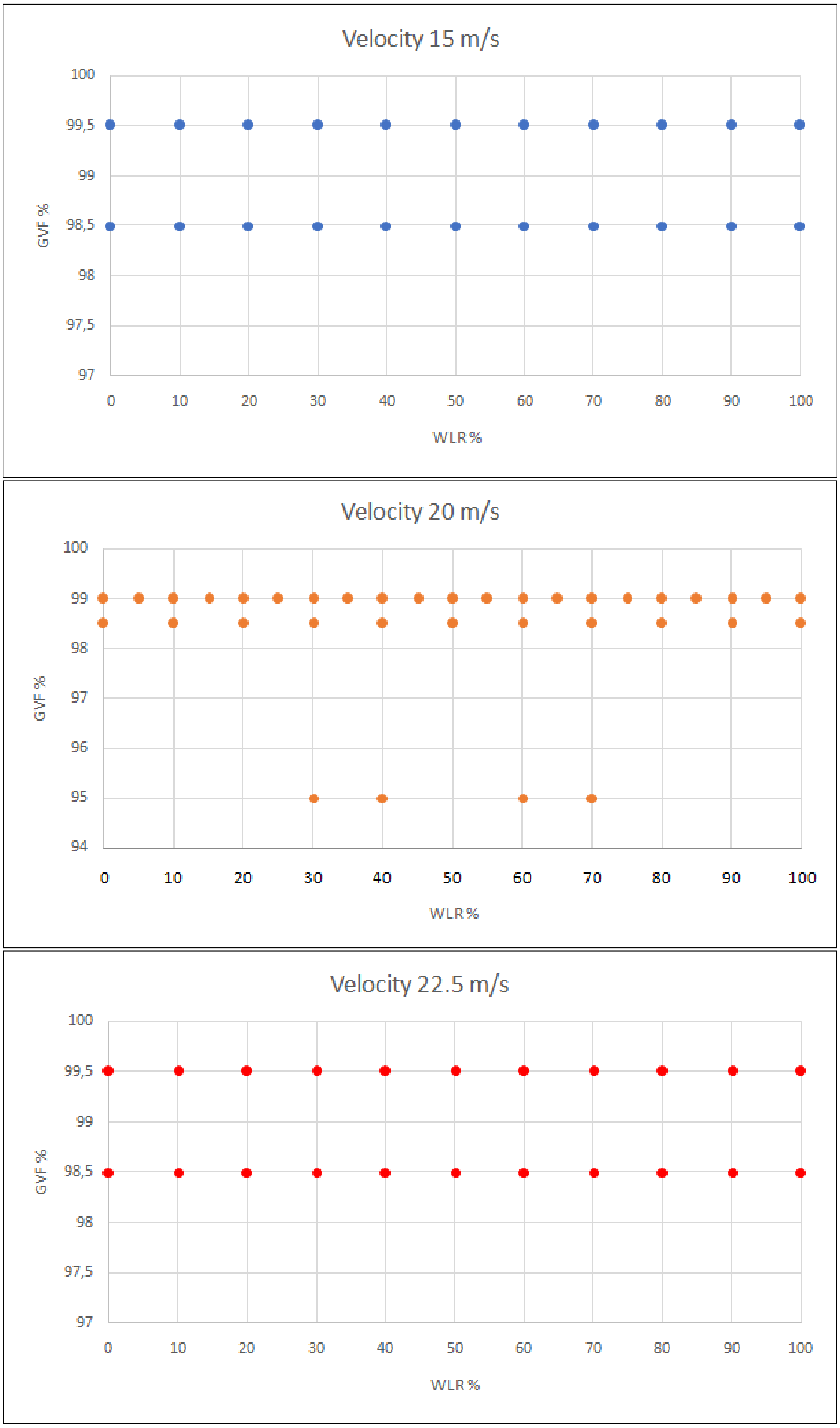

Figure 5.3: Datapoints for the loop flow

5.4 Water permittivity

In order to compute the WLR we start from the measure of the mixture permittivity and then we use the Bruggemann mixing formula to compute the WLR. In order to achieve that we need the informations on the single phases permittivities. The Exxsol is an insulator and so it has only the real part of the permittivity and is independent on the frequency, $\epsilon_o = 2.1$. For water permittivity it is more complicated since it is a conductor and its permittivity depends on the frequency, on its conductivity and on temperature. To compute the permittivity of water we use the Stogryn model [12]. This model uses the water conductivity and the frequency of the applied field to compute the water permittivity. The equation for the permittivity is:

$$\epsilon = \epsilon_\infty + \frac{\epsilon_w - \epsilon_\infty}{1 - i2\pi\tau\nu} + \frac{i\sigma}{2\pi\epsilon_0\nu} \tag{5.1}$$

where

- $\boldsymbol{\epsilon_\infty}$ is the permittivity of water at infinite frequency and is equal to 5.6;
- $\boldsymbol{\epsilon_w}$ is the permittivity of water at zero frequency;
- $\boldsymbol{\tau}$ is the relaxation time;
- $\boldsymbol{\nu}$ is the frequency of the applied field;
- $\boldsymbol{\sigma}$ is the conductivity;
- $\boldsymbol{\epsilon_0}$ is the permittivity of free space;

In this equation we have two terms which have a dependence on the temperature: $\epsilon_W(T, N)$ and $2\pi\tau(T)$. For these two values the dependency on the temperature is

$$\epsilon_W(T, 0) = \epsilon_W - 0.44T + 9.39 \cdot 10^{-4}T^2 - 1.41 \cdot 10^{-6}T^3 \tag{5.2}$$

$$2\pi\tau(T) = 1.11 \cdot 10^{-}2 - 3.82T + 6.93 \cdot 10^{-2}T^2 - 5.096 \cdot 10^{-4}T^3 \tag{5.3}$$

$\epsilon_W(T, N)$ depends also on the normality of the solution of salt and water. The normality of a solution is related to the concentration of the solute in the solvent. In our case the normality of the solution can be derived from its conductivity and we can use this value to correct the permittivity

$$2\pi\tau(T, N) = 2\pi\tau(T, 0) \cdot (0.146 \cdot 10^{-2}NT + 1 - 0.048N - 0.029N^2 + 5.64 \cdot 10^{-}3N^3) \tag{5.4}$$

In Fig. 5.4 we can see the behavior of water permittivity with different conductivities at a temperature of 25° in a frequency range from 0 to 10^{10} Hz. We can notice that for the imaginary part of the permittivity at our frequency regime the values seems to be zero for all the curves but this is only a problem of scaling. In Fig. 5.5 we present the same curve plotted for a lower range of frequency in the range of the GHz which are our working frequencies. We notice that the imaginary of ϵ is of the order of some hundreds. For the real part of the permittivity we have an high dependence on the conductivity which can decrease ϵ of 50% of the initial value such as in the case of $\sigma = 19S/m$ respect to the tap water case.

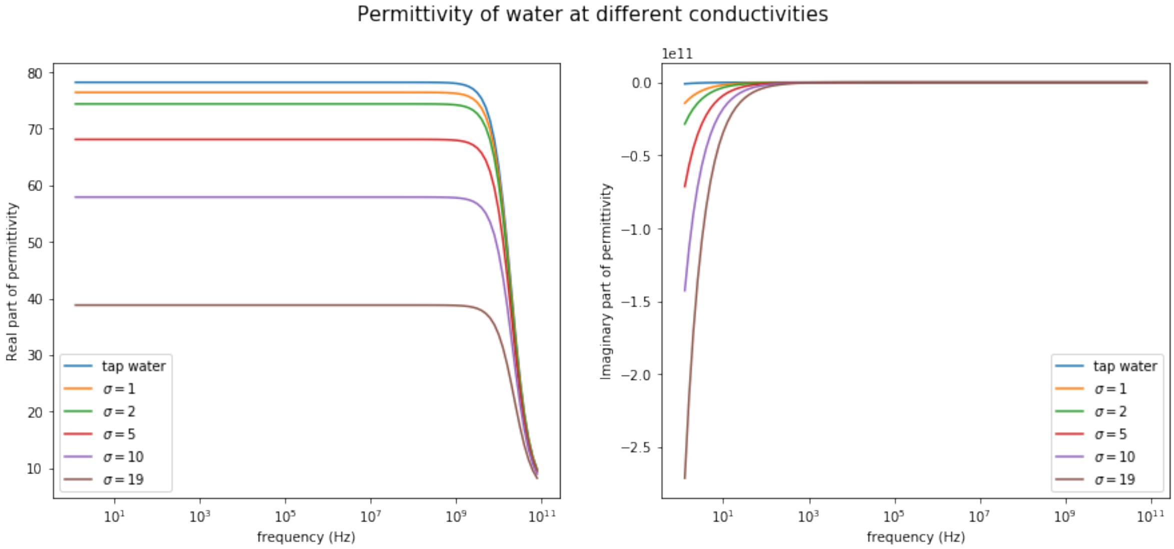

Figure 5.4: Permittivity of water with different conductivities at T = 25°

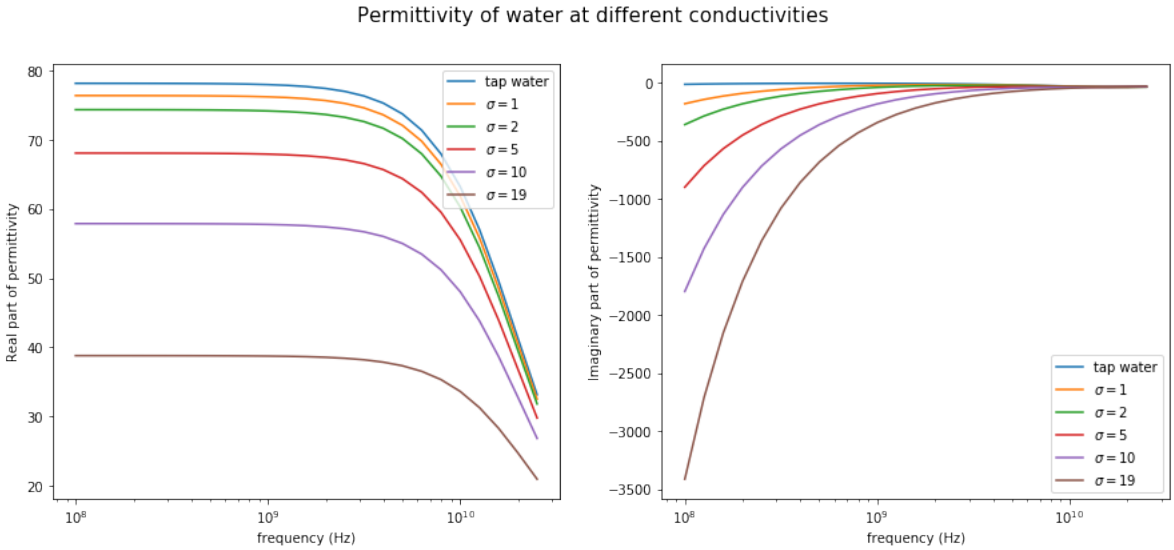

Figure 5.5: Permittivity of water with different conductivities at T = 25°

Now that we have all the informations we need in order to compute the WLR we will see the calibration process for the probes and how the results we found will apply to the loop data

5.5 Calibration

To calibrate the models we used the mixer setup in order to avoid any influence of the gas presence and measure only the mixture permittivity given by oil and water fractions. Then we used the Bruggemann mixing formula using the two ϵ_o and ϵ_w to compute ϵ_{mix} and relate this value to the resonance curve parameters measured. ϵ_w was computed using the Stogryn formula, (5.1), and since the WLR was known we could compute ϵ_{mix} and we used this value as a reference for the calibration.
For this type of analysis we used the mean curve of the three minute of data because the measure was stable and so we didn't have much fluctuations on the resonance frequency or amplitude. We took data at different salinities and with a WLR from 0 to 100 at steps 10. We

found out that the behavior of ϵ_{mix} could be described using three different models.
In Fig. 5.6 we can see the dependency of the permittivity respect to the curve parameters for the different models. We can notice that using these different models we were able to describe three different regime of permittivity behavior. The different colors on the plot represent data with different salinities and we can notice in all models that the points lies on the same curves. We can see the dependency on salinity on model 3, in this figure we have points from two different values for the salinity and the darker points are the ones with lower values. In this model data with higher salinity at the same WLR will have a lower permittivity value and a higher curve parameter. For example if we consider the points at 4, 10 and 19 S/m at 100 WLR and we use the 4 S/m as comparison the curve parameter of 10 S/m increases of 7 % and for the 19 S/m increases of 13 % while for the permittivity parameter the value decreases of 7 % in the 10 S/m and of 23 % in the 19 S/m.
We found that the minimum calibration point needed to build these models are

- air;
- exxsol D80 which is a refined oil;
- ethanol;
- tap water;
- water at different conductivities: 0.5, 1, 2, 5, 10, 19 S/m.

All these points are taken with static fluid condition while we need some additional mixing data using tap water and exxsol with WLR from 60 to 100 to obtain a full calibration. In Fig. 5.7 we can see the different curves for the calibration of a probe and we can immediately notice some dependencies: the resonance frequency increases when the permittivity decreases and the amplitude of the curve decreases when the conductivity increases. In particular we can notice that there is a gap in the resonance frequency between the oil continuous and water continuous regime due to the permittivity gap introduced in section 3.1.5. This fact becomes useful when we need to understand in which regime we are by using the resonance frequency. In fact we found that the resonance frequency of ethanol lies in this gap region and so we can use this value as discriminant to understand which is the host medium.

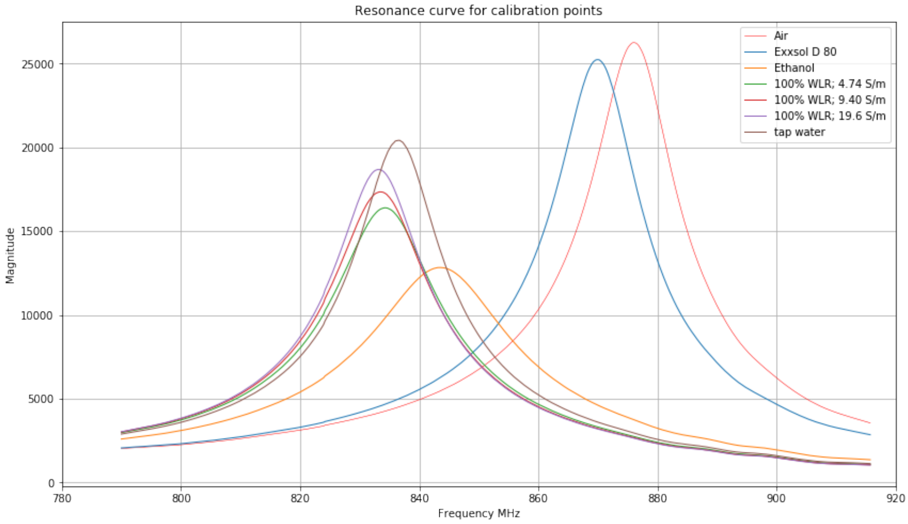

Figure 5.7: Resonance curve for different mixtures and conductivities

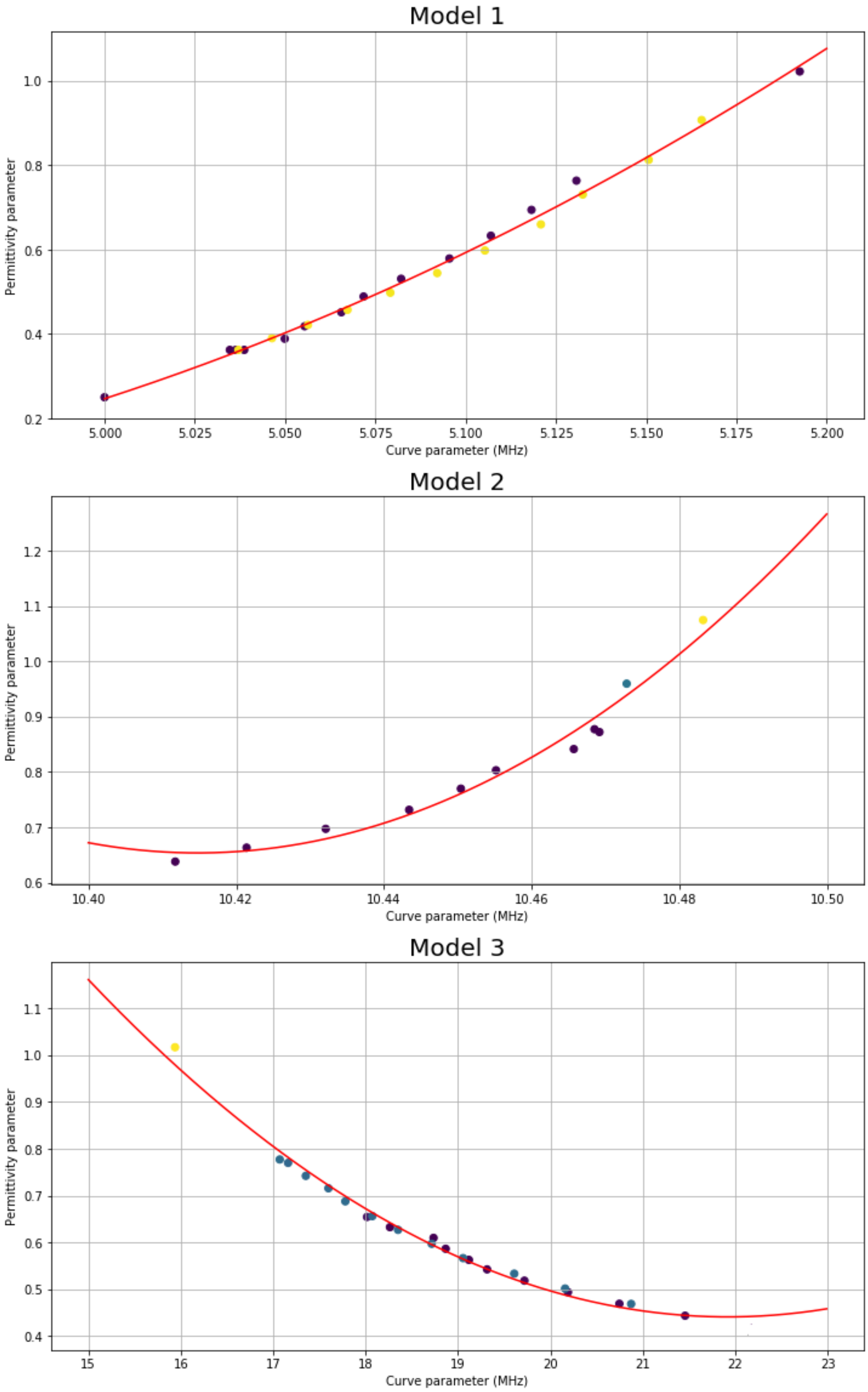

Figure 5.6: Three different models found for different conditions

5.5.1 Curve fitting

In this section we will analyze all the problems to overcome when doing an average for the curve signal. Since in the loop we have a gas flow which is of the order of 95 % of the total volume fraction we have a very small volume occupied by the liquid part. This leads to a very noisy signal because of the fluctuations of gas inside the sensible region of the probe. In particular the major problems are in the region of inversion where the points have a not defined behavior in terms of the host and inclusions and we end up having in the same point curve with different resonances. By taking the mean curve of the curves this effect is lost but we have a bias of the mean curve or we can have some spikes in the tail of the Lorentzian which leads to errors in the fit procedure.

We will now analyze some of the problems listed above using some curves from the data acquisition. In the next pictures we will show all the curves for a single point in red, in green the mean curve and in blue the max curve which is made by the max value of amplitude for each frequency.

- **Points with high noise**
  In some positions on the loop we have more noise due to an high presence of gas as we can see in Fig. 5.8. In this particular case we can see that the mean curve does not represent the set of curves with accuracy since we have a large spread on amplitude near the peak. For this type of curves we found that the best representative curve is given by the max curve. To explain why this is the best curve to represent the data we can think of the curve as given by the liquid part while all the fluctuations seen during the measure are given by the gas presence. These fluctuations have the only effect of lowering the amplitude of the resonance curve, in fact in the frequency range near this peak the air resonance curve seen in Fig. 5.7 has a very low amplitude. So if we have a gas slugs near the probe its effect is to shift the resonance curve towards higher resonance frequencies and so the amplitude in the range 800 − 840MHz lowers. However the resonance shift does not occurs since the gas flow rate is much higher than the acquisition rate of the

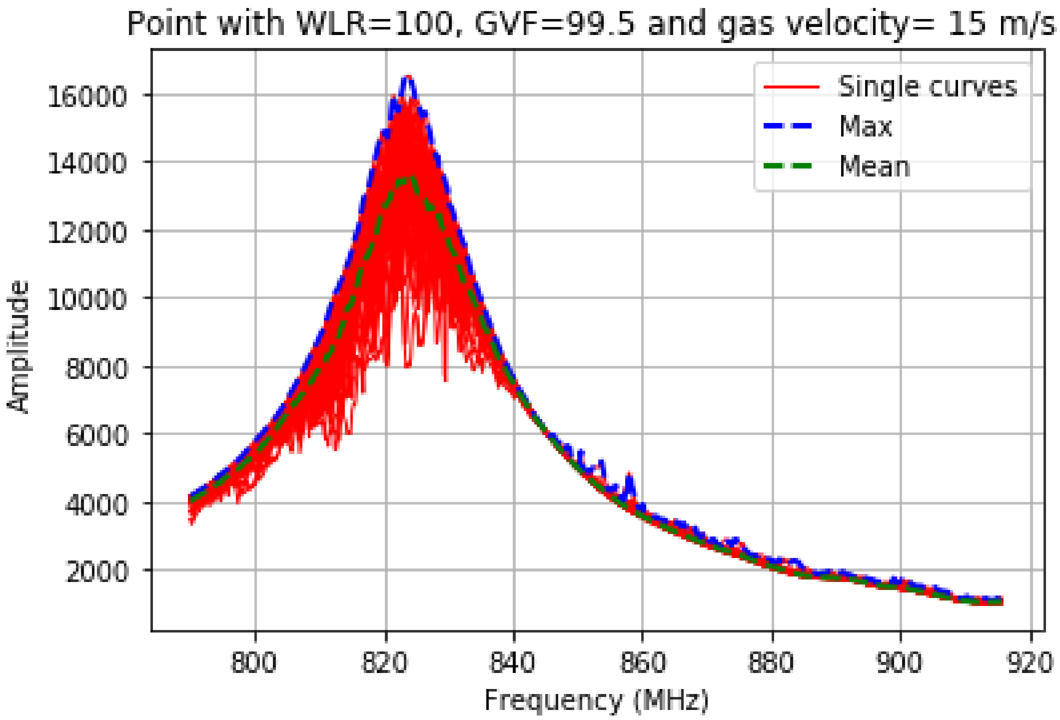

Figure 5.8: Point with high noise curves

electronics and so we are not able to detect a single gas slugs passing near the probe but instantaneously we can see the lowering in amplitude at a given frequency value.

- **Signal with spikes**
  In this case the signal is very noisy but the max curve do not represent the curves as the example above. This spiky curves can be explained by the slug presence of gas and oil in a water medium. The mean curve can represent the set of curves since the spikes are divided by the number of curves and the spread in amplitude is lower. However this procedure is an approximation of what the real resonance frequency can be; it allows to estimate that the frequency difference can be of the order of $2 - 3$MHz

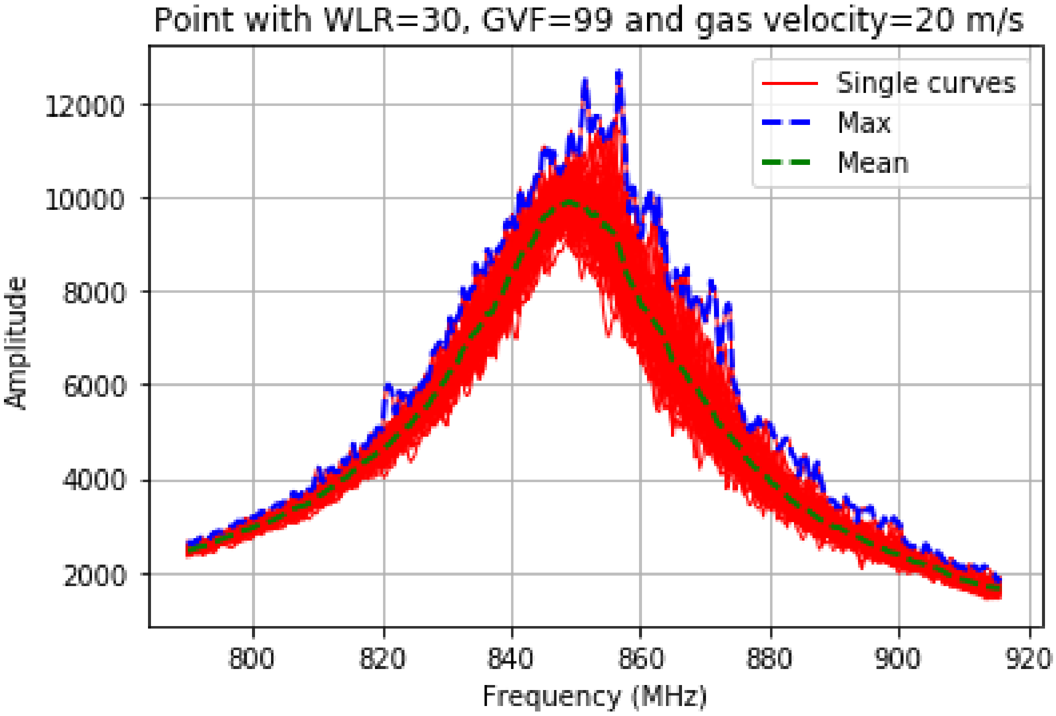

Figure 5.9: Point with high noise curves

- **Double peak signal**
  In this case the signal presents a double peak, this phenomenon can be explained by a slugs of liquid which is detected by the probe. The second peak arises from the other curves because we are near an inversion point and so even a small change in WLR can cause a big change in permittivity. When this happens, the mean curve is biased in the direction of the second peak and the max curve cannot be used. These points are the most difficult to analyze and we were not able to define a single curve able to represent this points with a double peak.

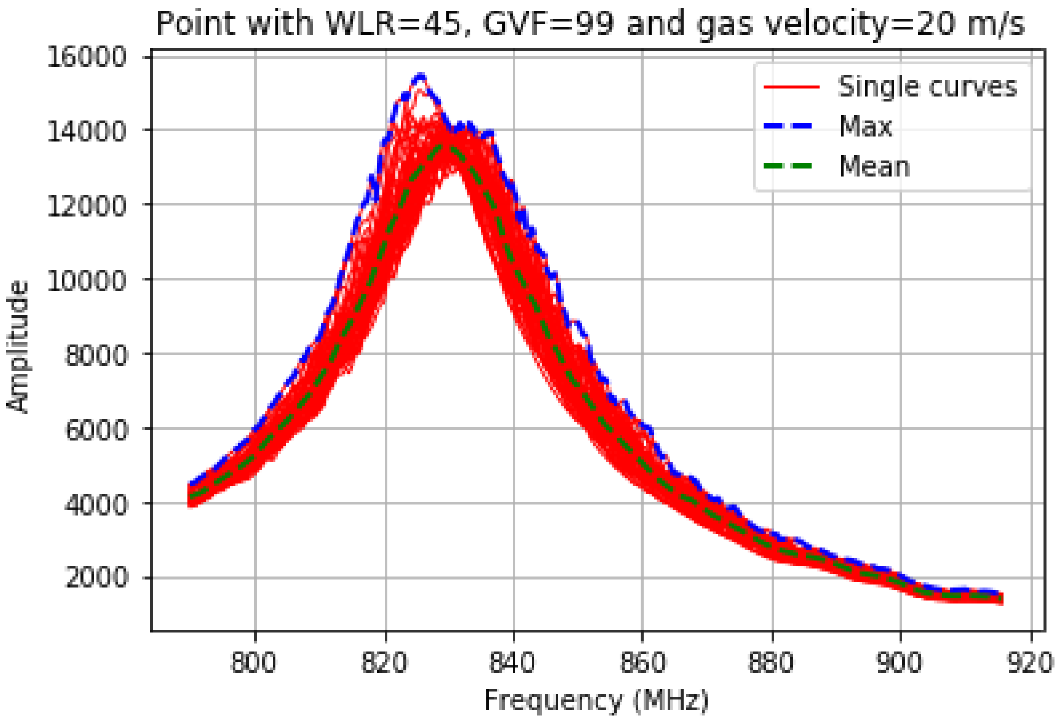

Figure 5.10: Point with high noise curves

These are the type of analysis used to overcome some of the difficulties in the signal extraction that we may encounter.

There is another type of signal extraction which we use to correct the curve parameters. Since we use an electronics to extract the signal we have a non zero noise level also in absence of signal. This noise do not alter the peak of the curve but modifies the tails outside the peak. This effect leads to an error in the fitting procedure and so if we fit the entire frequency range we introduce a systematic error on the measurement as seen in Fig. 5.11.

In Fig. 5.12 we can notice the fitting procedure done on a window range of width 40MHz and we can see that the fit is more overlapped with the initial curve. Using this frequency cut the resonance frequency can vary of 1-2 MHz so we introduce a small correction but for the FWHM we can have a change of the order of $5 - 7$MHz. Since the width for the curves is of the order of 13MHz this correction can change substantially the result of the model.

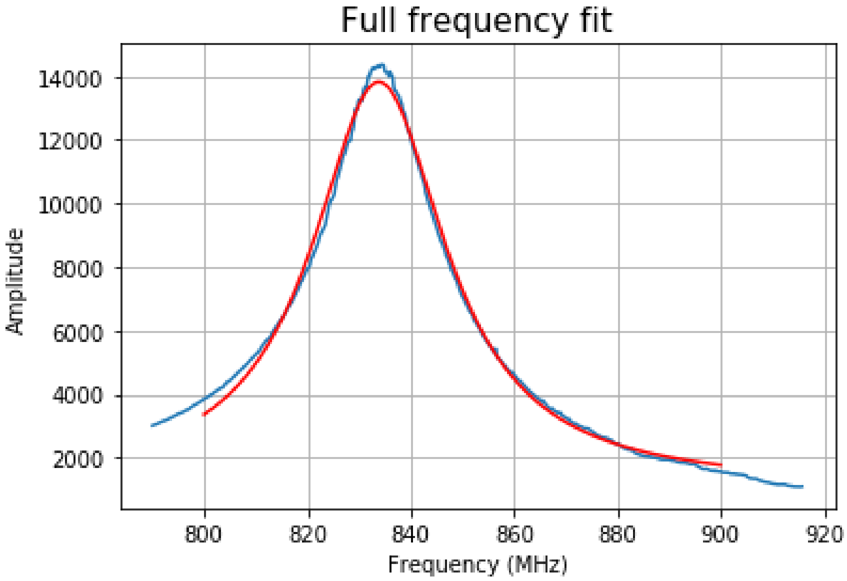

Figure 5.11: Lorentzian fitting on full frequency range

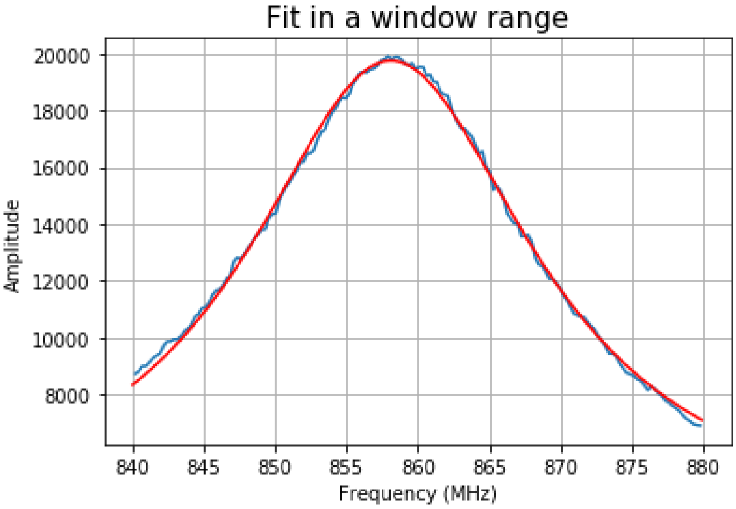

Figure 5.12: Lorentzian fitting on a window frequency range

5.6 Results

On the loop we had various different positions where we placed the probe and here we present the best results obtained. We will analyze three different positions and see how we can correct the results according to the problems each positions had.

- **Position A** This is the position for which we had more points in different flow regime and points from all the conductivities used in the loop. In Fig. 5.13 we can see the plot of the actual WLR vs the result from the models and the residual plot of the data respect to GVF and conductivity. The color code represents the conductivity of the points where with the darker color we refer to lower conductivities.

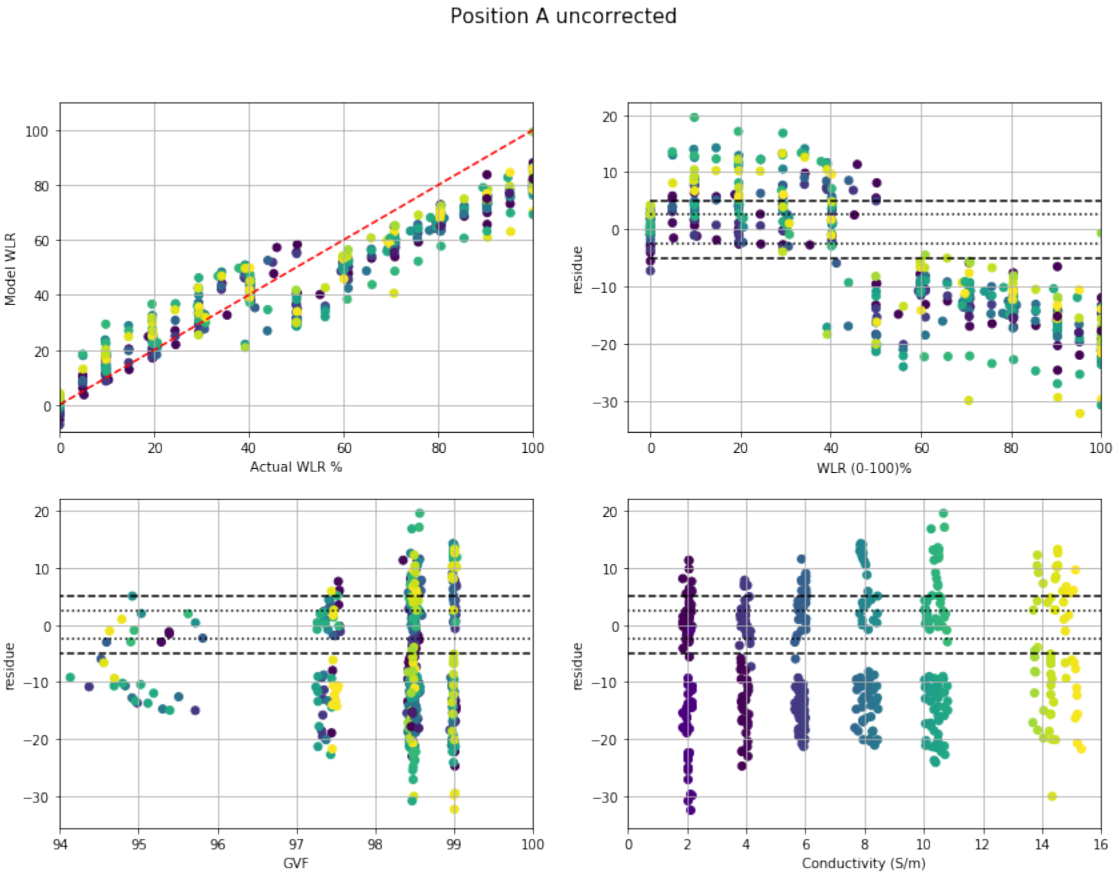

Figure 5.13: Position A result

In the first plot we can see the different flow conditions in which we had a water continuous medium or an oil continuous and the inversion region which is around 50 % WLR in which we have a step in the models result. In particular we can notice that in oil continuous the results overestimate the real values while in water continuous the results underestimate them. These two phenomenons can be explained by making some assumptions based on the position of the probe. In the water continuous case the permittivity is lowered by gas bubbles inside the sensible region of the probe. For the oil continuous case the results can be explained by some water presence which is not washed away from the liquid flow. Of course we have gas bubbles also in the oil continuous case but since oil and air has a small difference in permittivity the gas presence is hard to detect while even a small fraction of water can provide a change in permittivity due to its very high permittivity compared to oil. The first problem can be solved by introducing the gas presence in the

Bruggeman mixing model while for the second it is harder to find a solution since it is difficult to know the exact water fraction.
From the residue plots we can see the tendency of the points respect GVF and conductivity. If we look at the GVF residue plot we can notice that the residue has a mean different from zero while it should be zero, in particular at GVF near 95 % we can divide the points in two regions: the oil continuous which lies inside the dotted lines and the water for which we have a mean around -10. So from this plot we can see that we have to introduce a correction for the water part respect to the gas fraction. On the conductivity plot we can see that the points spread remains quite constant between the different values so the model is able to correct the conductivity dependence of the resonance curve.

- **Position B** For this position we encountered some problems due to gravity. In fact this probe was placed in a horizontal part of the pipe and so even if the fluid was well mixed we could have a water layer on the bottom which was detected by the probe. This fact can be seen in Fig. 5.15 where we can notice that the effect on the probe was to measure a mixture in water continuous regime even at WLR lower than 40 %.

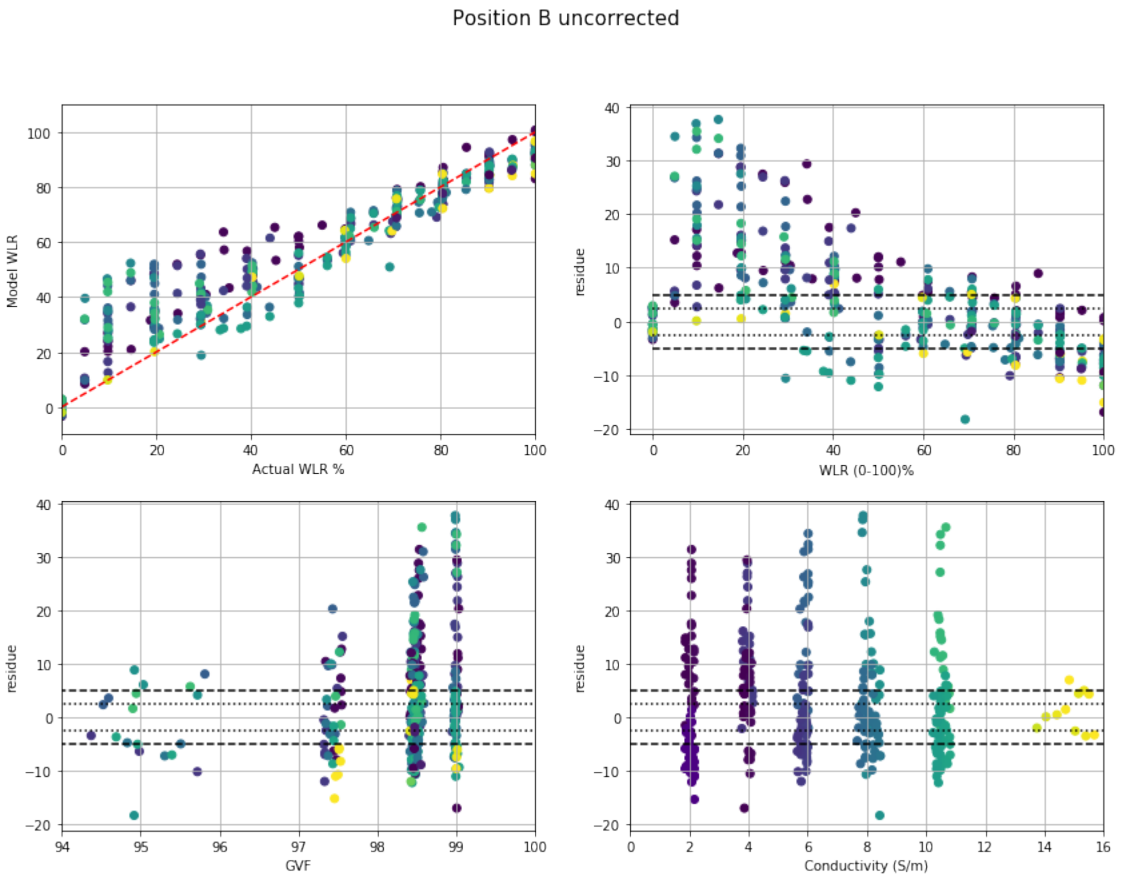

Figure 5.14: Position B result

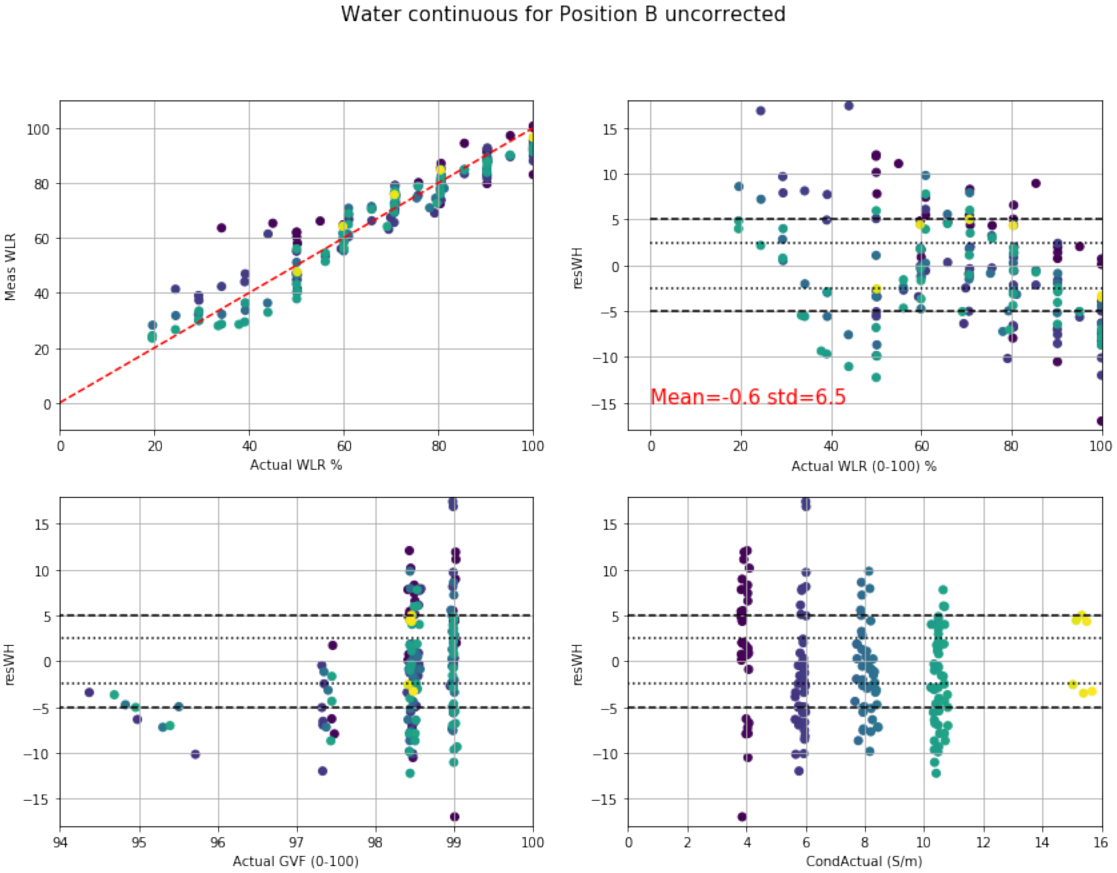

Figure 5.15: Position B result - water continuous regime

Although this problem in this position we do not need to correct for the gas presence since on the loop we were able to see a liquid layer on this sensors which was clearly greater than the sensible region of the probe. The GVF residue for the water continuous case is in contrast with the previous statement since the residues are not centered around zero. We will discuss later in the gas correction this pattern in the GVF residue.

- **Position C** For this configuration we had the minor number of points since we changed a piece of the loop to insert the probe. For this reason we have points on this setup only with the conductivities of 10 and 16S/m.

  In the plot we can see that the model in this position underestimates the real value of WLR, this is because this position is in a part of the loop where the gas presence is very high and so it alters the permittivity measure. We can also notice that there is a tendency on the residue plot respect to WLR so there is also another correction to include but we will see later that we can use the gas correction to correct both this tendencies.

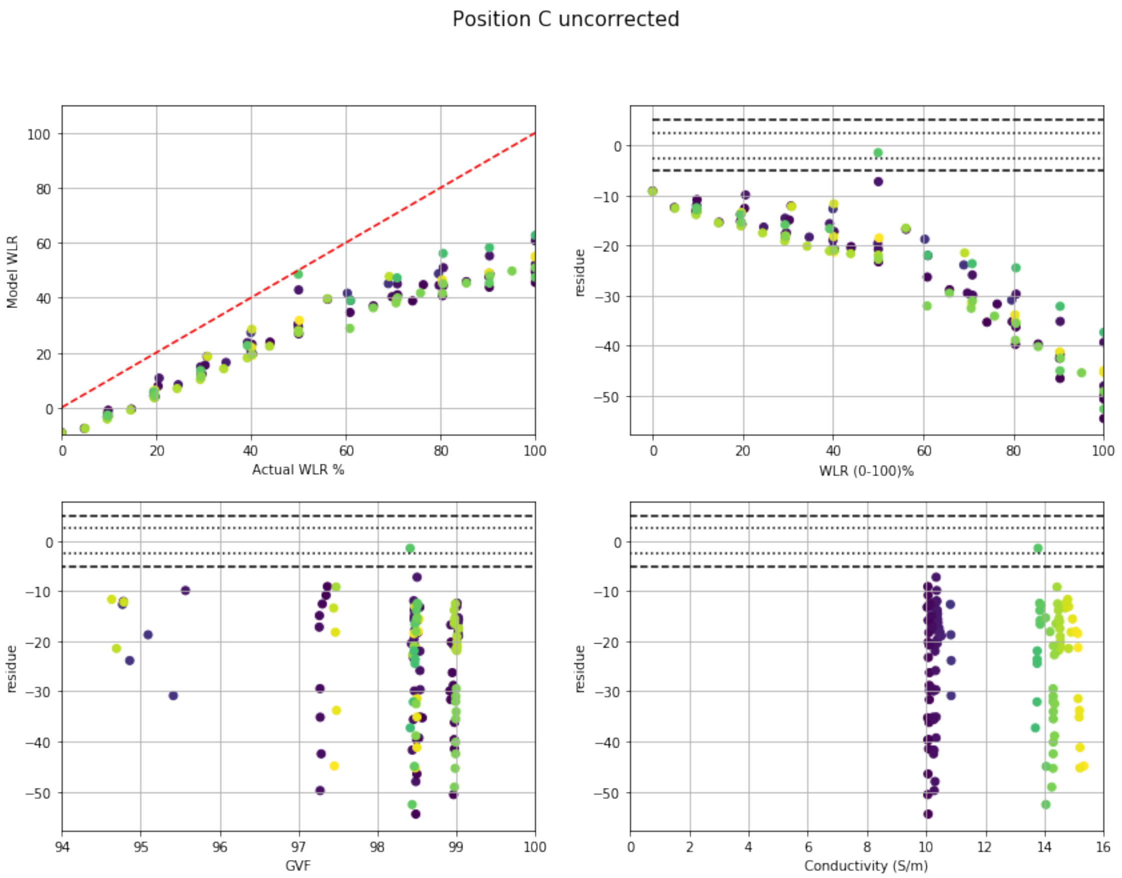

Figure 5.16: Position C result uncorrected

5.6.1 Gas correction

In order to correct the gas presence we can include the gas fraction in the Bruggeman model, but the gas presence depends on the position of the probe and depends also on the GVF of the flow. To find the gas correction we selected a subset of the dataset and use it to compute the gas presence. So we had the reference result for the permittivity using the Bruggemann model and we wanted to add a fraction of gas to our model computation in order to equalize the two results.

- **Position A** For this position the gas presence is low and it has a small dependence on the GVF. For this reason we can correct the underestimation of the result using a constant gas fraction inside the sensible region of the probe.

- **Position B** As said before in this position we should not correct for the gas presence since in this region there is a liquid layer greater than the sensible region of the probe. Besides we saw that the residues respect the GVF have not a zero mean so we must try to correct this offset. By applying a 3 % of gas presence we obtain Fig. 5.17. By comparing this image with Fig. 5.15 we can notice the influence of the gas correction. In this case we may notice that the residues at lower GVF are now centered but from the residue plot vs WLR we can notice that there are some tendency which are uncorrected in both of the two results. So the gas correction is not able to correct the behavior of the points but only shifts the mean value of the residues. As a consequence the gas correction do not work in this case and the behavior of the residues have to be considered influenced by the water layer present in this position.

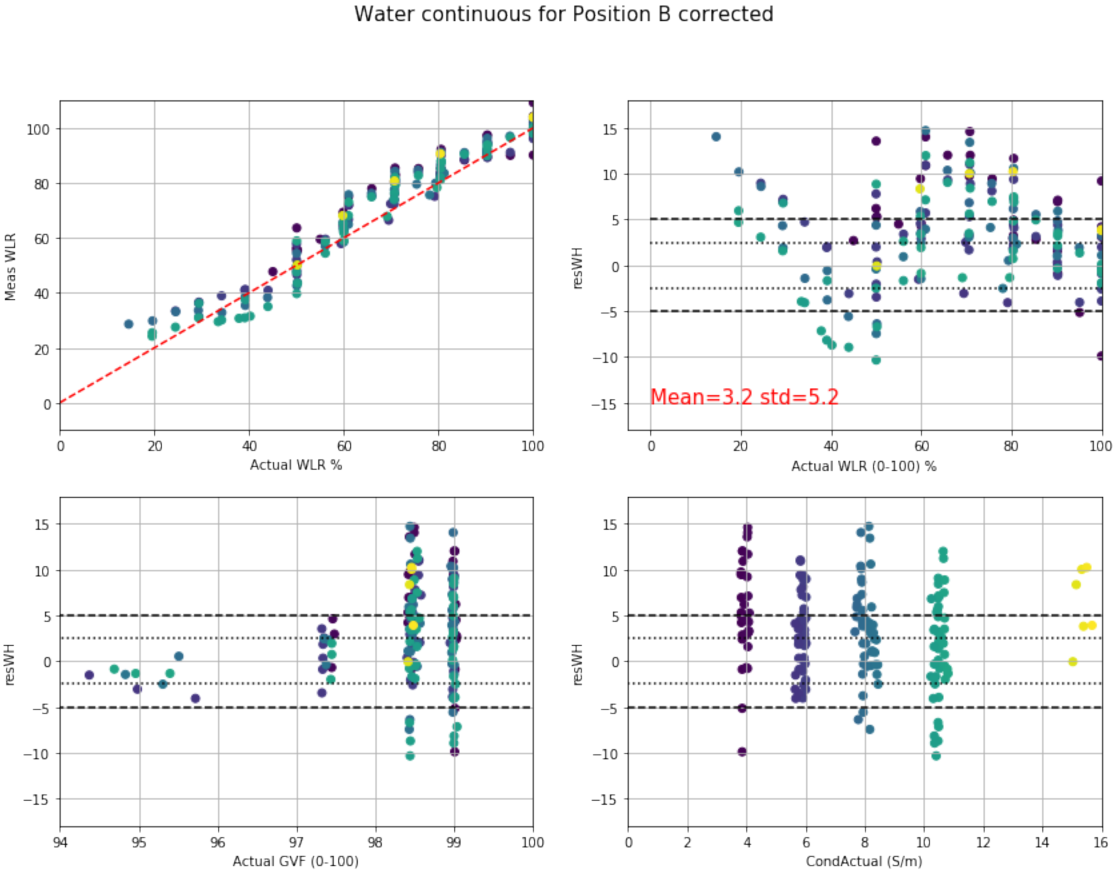

Figure 5.17: Position B result - gas correction

- **Position C** In this position the gas fraction inside the sensible region of the probe can reach values near the 45% and depends on the GVF of the flow as we can notice from the dependence on Fig. 5.18 and 5.19. At higher GVF the gas fraction in the sensible region becomes higher while at lower GVF it lowers as we should expect. For this probe we found that the right amount of gas to correct depends also on WLR, in fact the water presence influences the field near the probe and since water has a high permittivity it absorbs the field and reduces the sensible region. This fact can be seen in the gas correction for oil where in the first plot we refer to the WLR with the different points color listed in the legend. We notice that the brigther colors at a defined GVF lies at a lower percentage of gas which means that when the WLR increases the gas percentage in the sensible region decrease. In the water continuous case this fact is less visible since in this case we are in a medium of water which has a stronger contribution to the absorption respect to an oil medium. There are some outlier in the plot but their influence in the red line which interpolates the points is very small and so we do not need to filter this data.
  For this reason we have to correct the gas presence for this position using two different models according to the host medium and the GVF. The dependence of a given regime on the WLR is less significant and consequently provides a smaller contribute that was neglected.

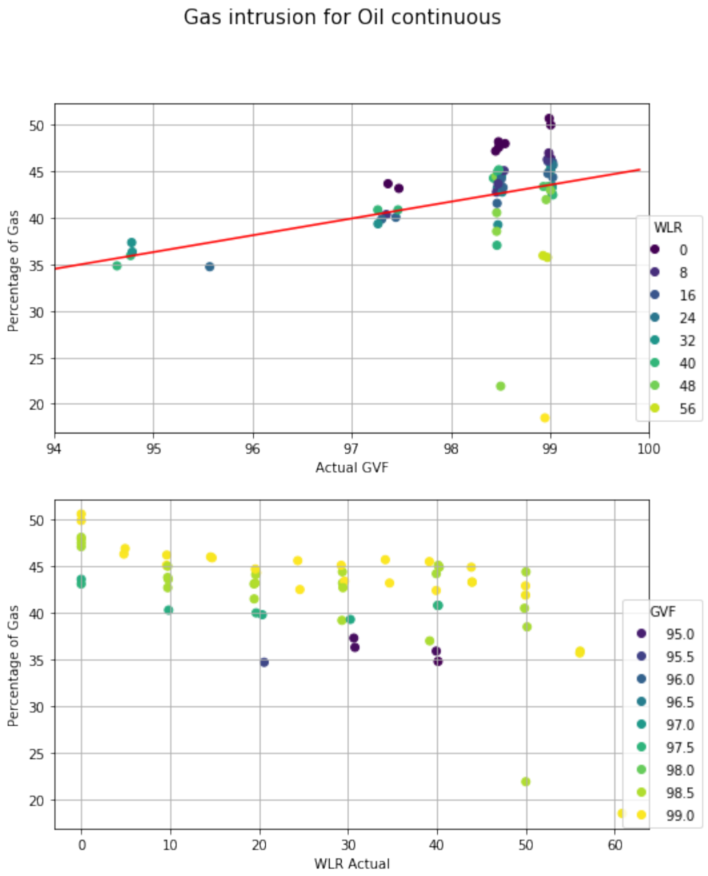

Figure 5.18: Position C gas correction for water continuous regime

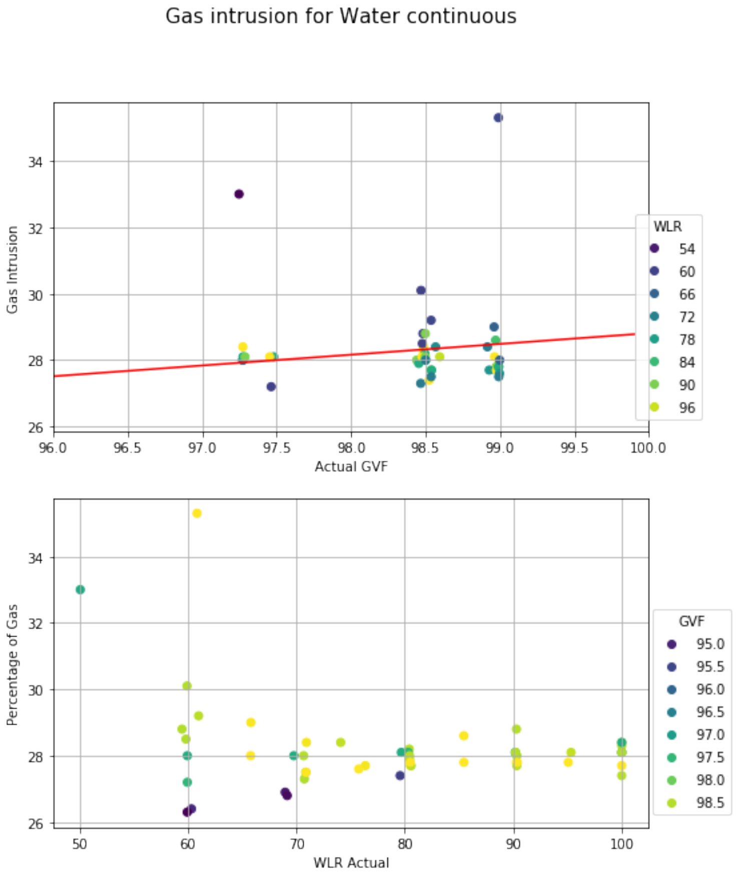

Figure 5.19: Position C gas correction for water continuous regime

5.7 Final Results

- **Position A** In this position only the water continuous region was corrected using the gas correction. In Fig. 5.20 we can see the result obtained and we can notice that in the region between 40-50 WLR there is a not well defined behavior which differs from the water and oil continuous branches. This happens because, has explained in the curve fitting section, in this region it is hard to find a representative curve for the point and so there is a bias in the result which we were not able to corret. In Fig. 5.21 it is displayed the result for the water continuous region where we filtered the data to have solution only over 60 % WLR. In this region we have a mean for the residue which is almost 0 and a standard deviation of 3.4. The fact that the mean is almost 0 is a hint that the applied constant gas correction was made with the correct value and we can see it also from the residue plot vs GVF in which the residual do not have a tendency but lies around 0. We can notice that there are some outlier at lower conductivities, for these points we were not able to understand their deviation but the hypothesis is that they belong to a different model.

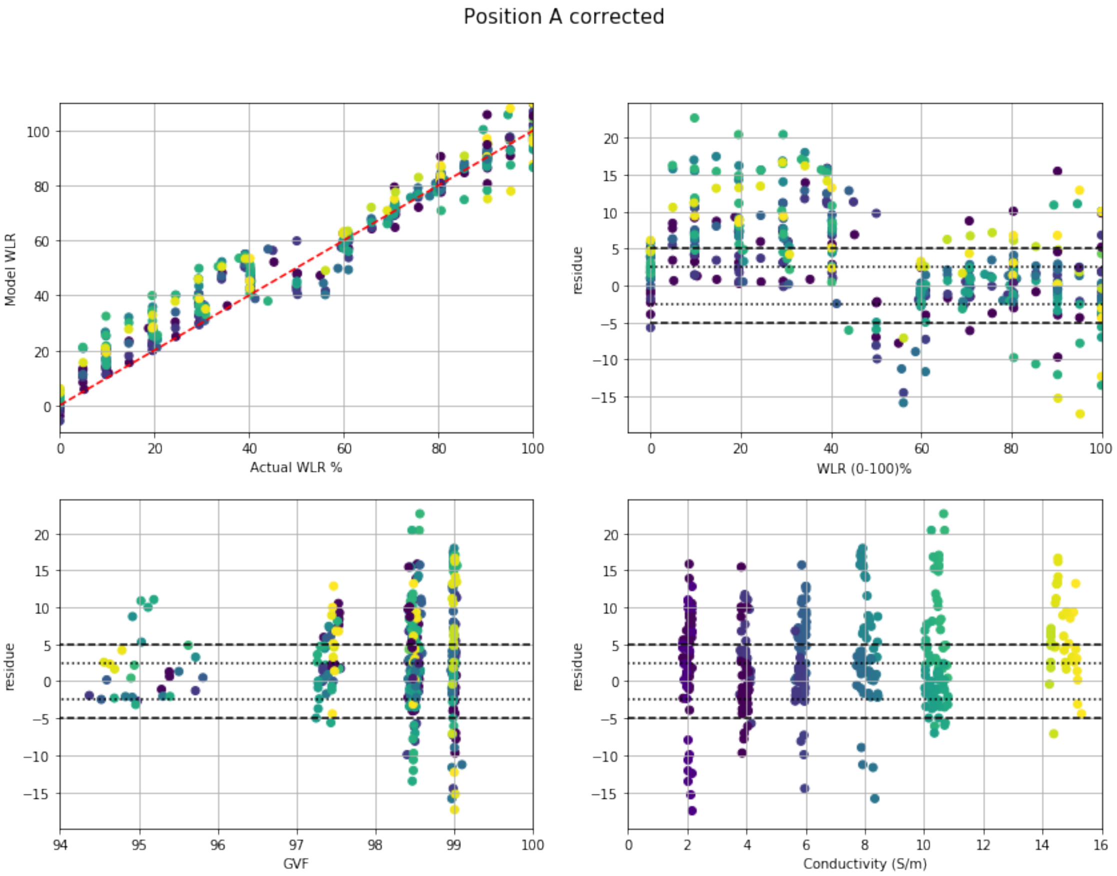

Figure 5.20: Position A result corrected

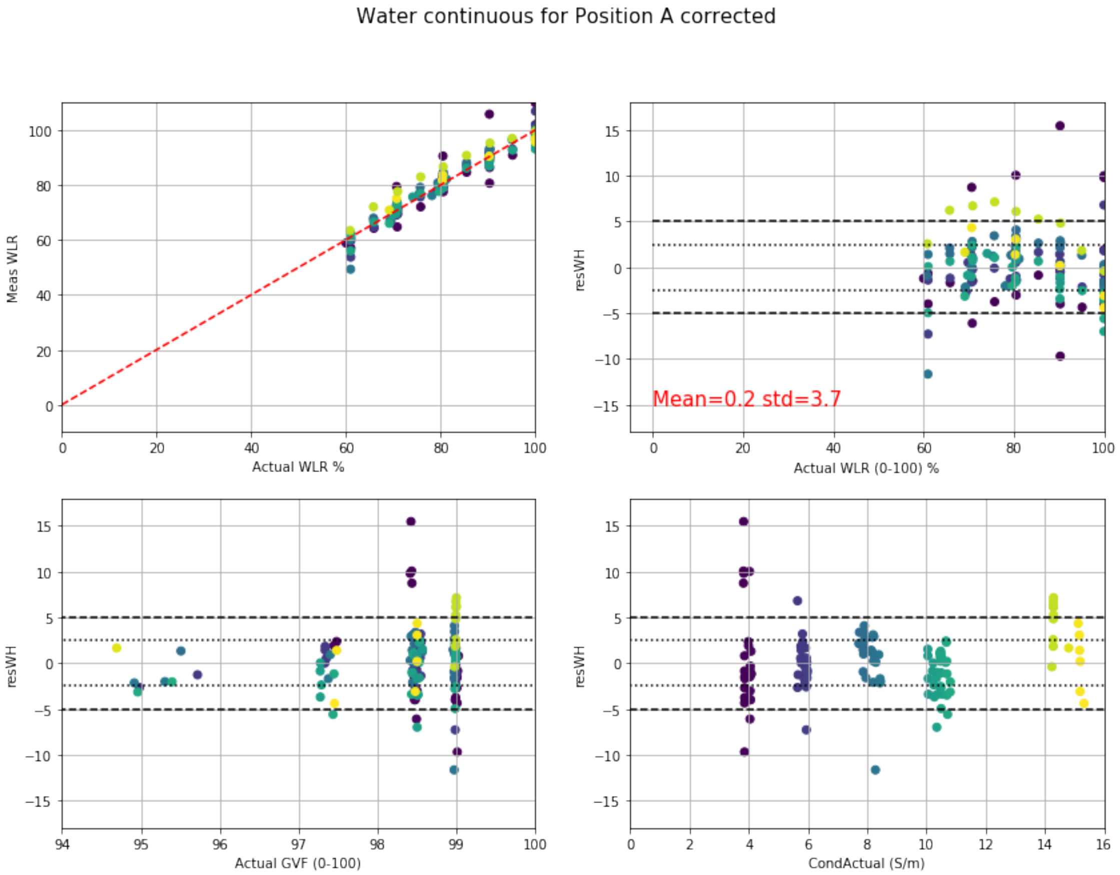

Figure 5.21: Position A result corrected - water continuous

- **Position B** In this position we have not performed any further correction in the analysis so the result are shown in Fig. 5.14 but we focus now on the water continuous regime applying a filter for the result at an actual WLR over 50 %. In this case we can notice that the mean of the residue is not zero and also that we have a tendency of the residue respect GVF but as said before this must be considered affected by the water layer which lies on the bottom of the pipe.

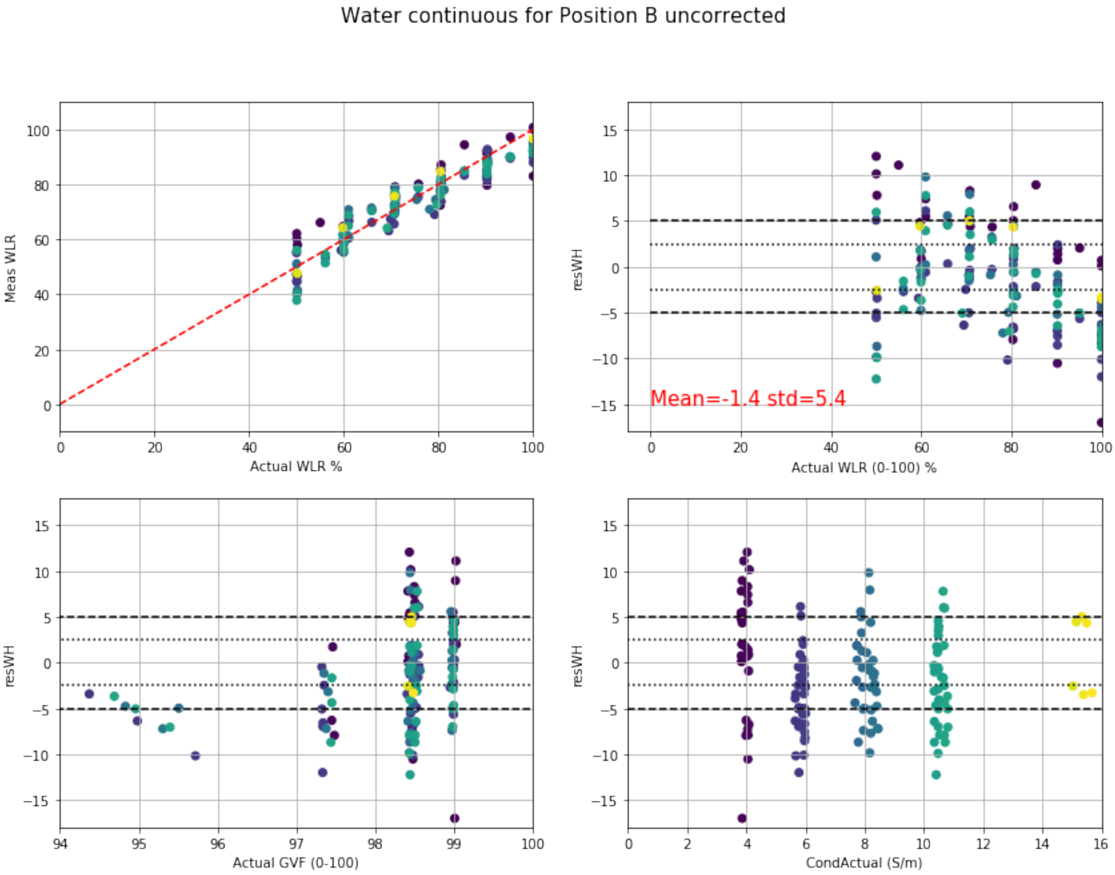

Figure 5.22: Position B result filtered - water continuous

- **Position C** For position C we applied a gas correction for the permittivity computation and by comparing Fig. 5.23 with Fig. 5.16 we can notice that when we apply the gas correction we are able to correct the underestimation of the points. In particular we can notice that the results for the oil continuous part have a small variance while for the water continuous the spread gets higher.
  In Fig. 5.24 we can see the plot result only for the oil continuous case and we can notice that the mean value of the residue is not 0 and have a standard deviation of 1.3. Although the variance of the points is very low but we can notice that there can be some underestimation on the gas fraction used to correct the data and this can be seen also on the residue vs GVF plot. In fact we can notice that residual increase with the GVF so there is still a tendency which we should correct to have a correct result. Despite this further correction the results obtained for the oil continuous part are satisfying for this position.

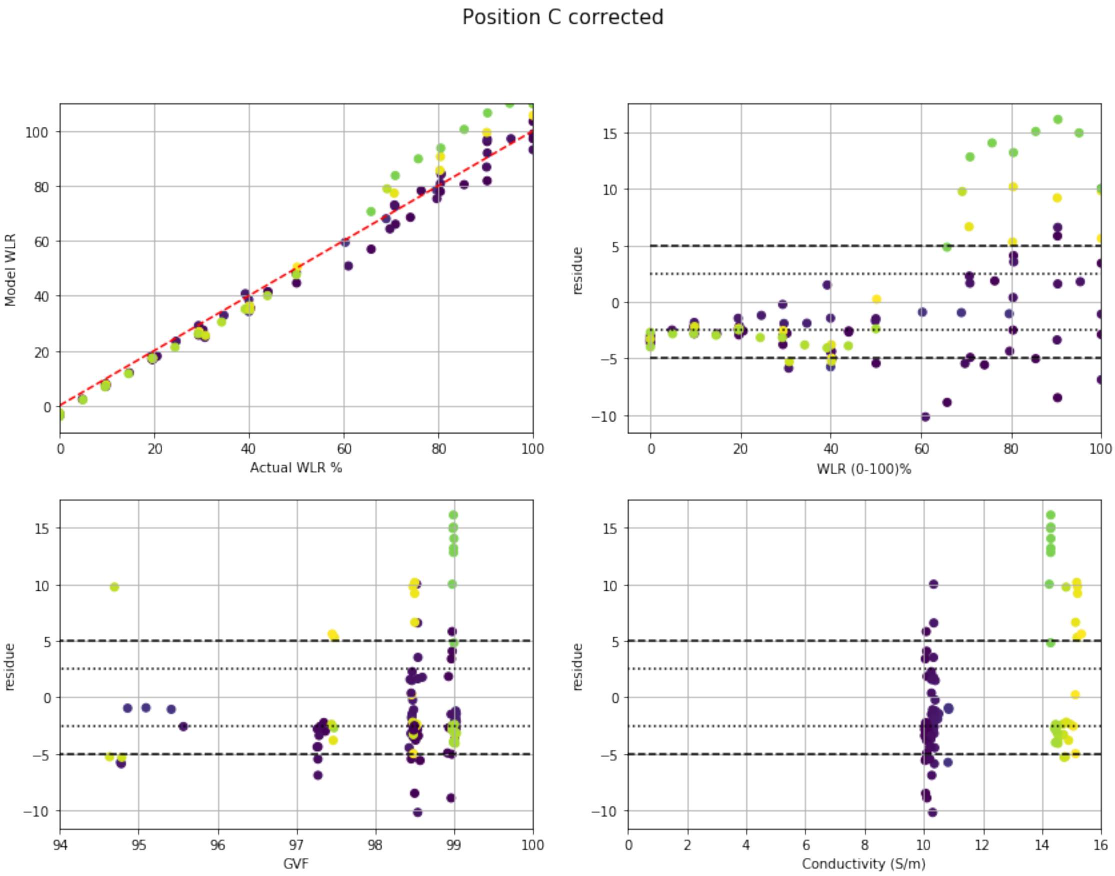

Figure 5.23: Position C result corrected

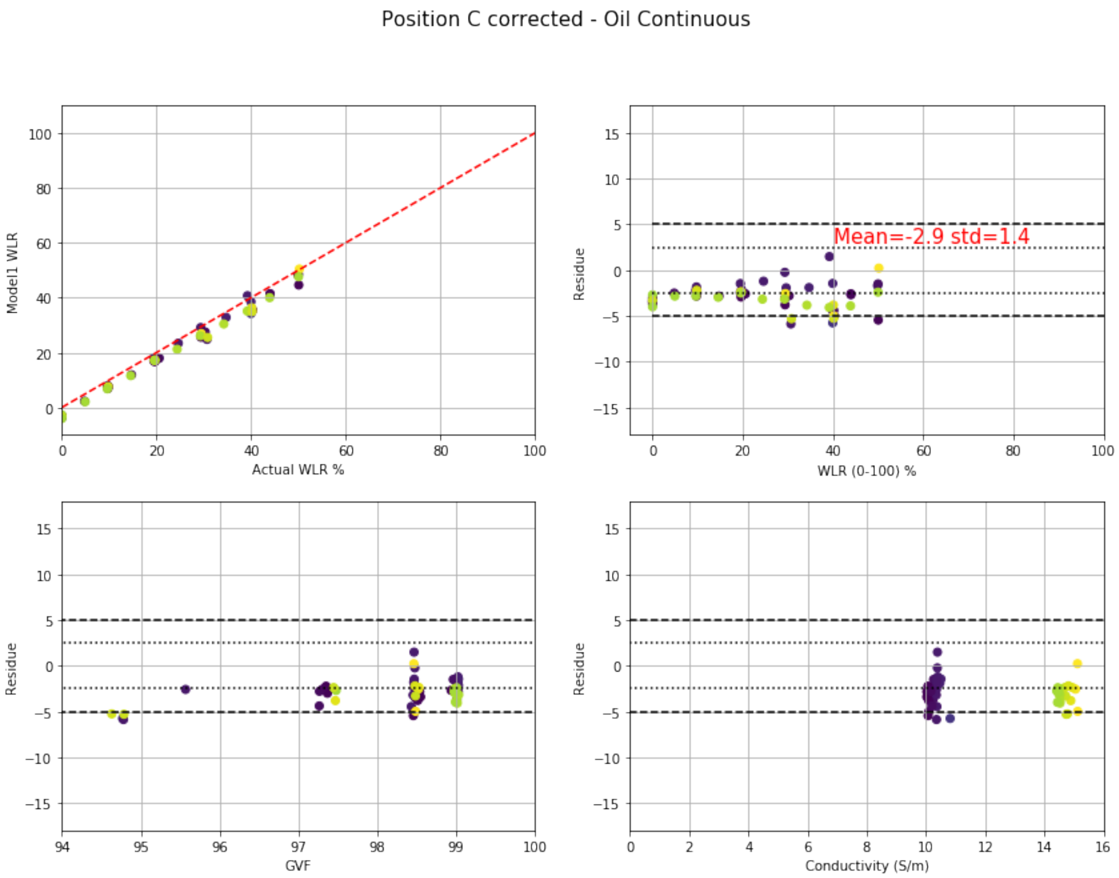

Figure 5.24: Position C result corrected - Oil continuous

The single position we analyzed did not give a result as expected but we can see that combining two of them we can achieve a good result. This is the example of position A for water continuous and position C for oil continuous. Using a combination of these two we can have a system of probes with a standard deviation of 1.1 in oil continuous and 3.4 in water continuous. So in the oil continuous case the 95 % of the points lies in the region -2.9 ± 1.1 so we have achieved the required accuracy. For the water continuous regime we have the 95 % of the points who lies in the region 0.2 ± 6.4 so we are not inside the ± 5 region accuracy. If we take the 90 percentile we obtain a region 0.2 ± 3.3 which now is compatible with the accuracy wanted. In the end the best result we can obtain using a combination of position A and C are:

Best result obtained		
Regime	Mean ± percentile	Score
Water continuous	0.2 ± 3.3	90 percentile
Oil continuous	-2.9 ± 1.1	95 percentile

Chapter 6

Machine Learning

In this chapter we will introduce the machine learning process and we will see some models which can be applied to our analysis and which result we can obtain.

6.1 Introduction

The definition of machine learning is given by Tom Mitchell and recites: "A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."
This simple definition can help us understand how machine learning algorithms work and why they are useful nowadays. In simple words the machine learning algorithms learn by itself how to treat and analyze the data given as input and then is able to predict the result for a new value using what it learned. These class of algorithms are largely used in the big data era, we can find its use in spam classification, hand writing recognition, image recognition and in many other examples.
The Machine learning algorithms can be divided into two big classes:

- **Supervised:** in this class the algorithm is trained using the input values and the correct output;
- **Unsupervised:** in this class of algorithm only the input values are given.

For our purposes we will deal only with the first type of class but also the second class is very important and is used for example in the spam identification and the recommended section on famous search engine or marketing website.
In the first class we divide the dataset in a train set and in a test set. The train set is used to train the algorithm in order to find a model which describes the data and then the test set is used to see how well the model is able to predict the results. This class of algorithm is also divided into two subclasses which are:

- **Regression:** the algorithm finds a pattern in the data and use this pattern to predict the new values;
- **Classification:** the algorithm divides the inputs in different classes and use this classification to predict the class of a new value.

The most used example for the regression algorithm is the prediction of the price of an house. In this example we can have as input values the dimension of the house, the construction year, the zone of the city and the number of bedrooms. Using these informations and a set of train

houses for which we know the price we can predict the new house's price.
For the classification problem an example can be the classification of the species of iris flower. In this case we can use as input the length of the petal, their color, the separation in length between the petals and the petal width. In Fig. 6.1 we can see the result obtained for three specie of Iris: setosa, versicolor and virginica. The different colors represent the different species and we can notice how a single flower is classified into a species using its characteristics. We now focus our attention in the regression algorithms and we will begin from the simpler model of linear regression in order to explain some features of the regression problems in the easier example and then we will discuss the models used to analyze our problem.

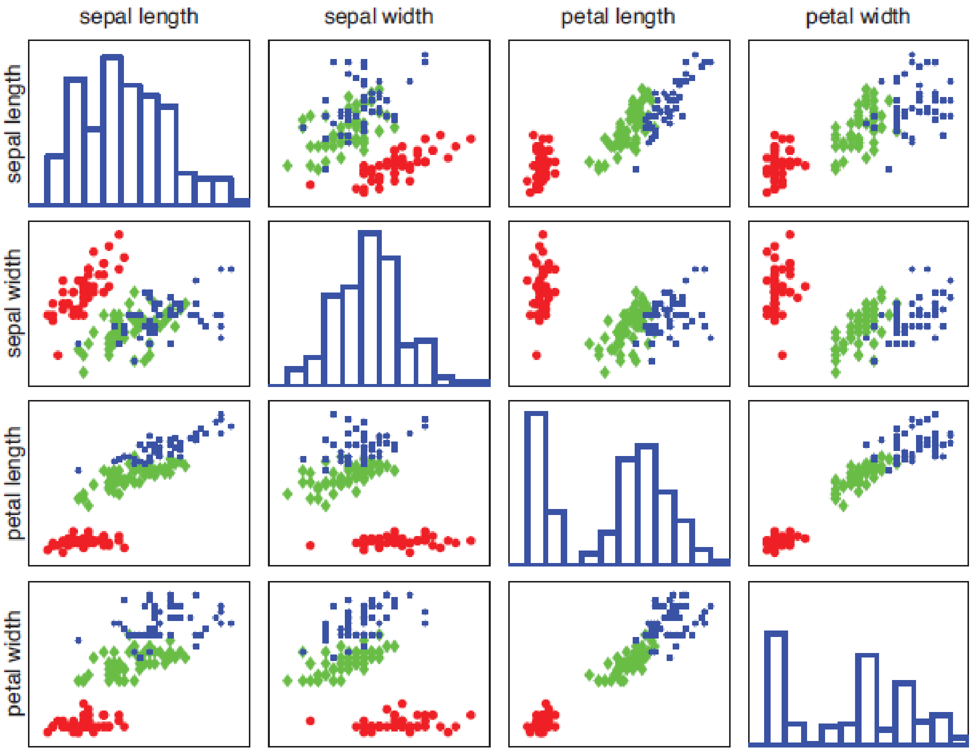

Figure 6.1: Visualization of the data from Iris flowers [13]

6.2 Linear regression

The simplest model for a linear regression is a linear combination of the input variables

$$y(\boldsymbol{x},\boldsymbol{w}) = w_o + w_1x_1 + w_wx_2 + ... + w_Dx_D \tag{6.1}$$

where D is the dimension of the input and w are the hypothesis parameters. To find the solution of the Machine learning problem we must introduce the cost function which is

$$J(\boldsymbol{w}) = \frac{1}{2}\sum_{i=1}^{N}(y_n - w_0 - \sum_{j=1}^{D} w_jx_{ij})^2 = \frac{1}{2}\sum_{i=1}^{N}(y_i - \mathbf{w^T x_i}) \tag{6.2}$$

where $\mathbf{w}$ and $\mathbf{x}$ are the vector of the hypothesis and inputs where we added the value $x_0 = 1$ in order to have the same dimension. There are mathematical demonstrations which shows that the cost function has a global minimum as function of the hypothesis. For example if the linear regression has only one input parameter x_1 the cost function will be a parabola which has a minimum in w_1 which is the hypothesis parameter which minimizes J. The machine learning algorithm finds the parameters for which the cost function has its global minimum and these parameter are the ones which has the best fit for the data. This model is simple because is a linear function of the parameter w_i but of course has some limitations since it is also linear in the input parameters and so will not be able to predict a complex model. To compute different models we can introduce the higher order power for the input values or we can use some functions called **basis function**, ϕ_j. Using the basis function we write the equation as

$$y(\boldsymbol{x}, \boldsymbol{w}) = w_0 + \sum_{i=1}^{D} w_j \phi_j(X) = \mathbf{w}^{\mathbf{T}} \boldsymbol{\phi}(\mathbf{x}) \tag{6.3}$$

in this manner the equation is still linear in the parameter but if $\boldsymbol{\phi(x)}$ is non linear the equation will be non linear in the input values. In this way the problem remains simple but we can model data which has a non linear pattern.
An example of basis function is the radial basis function also called Gaussian basis function [14]

$$\phi_j(x) = exp\left\{ - \frac{(x - \mu_j)^2}{2s^2} \right\} \tag{6.4}$$

where μ_j is the location of the basis function in the input space and s is their spatial scale. Another basis function is the sigmoidal basis function

$$\phi_j(x) = \sigma\left(- \frac{(x - \mu_j)^2}{s} \right) \tag{6.5}$$

where σ is the sigmoidal function

$$\sigma(a) = \frac{1}{1 + exp(-a)} \tag{6.6}$$

We can choose the best basis function or polynomial degree to fit our model and minimize the cost function J but if we use a high complexity for the model we can introduce an error in the process and have an overfitting problem. In Fig. 6.2 we have an example of underfitting and overfitting problems in a fitting procedure. In the first case we have a model which is too simple and so it is not able to describe the pattern of the data, in the second case we are using a complex model and so by minimizing the cost function it introduces some error on the prediction. The last statement will be explained later in a more complete way but now we for greater clarity we go back to the first example of the houses's price. In this case we know that the price of the houses will increase if their size increase, so if we need to choose a model to fit these data we cannot choose an even power law as max degree for the polynomial function because by doing so we can have a good result for the training data but we end up having a paradox in which at a certain point the prize will decrease when increasing the house's size. So before choosing the model it is always recommended to have in mind the type of function we want to model and the pattern expected from the data.

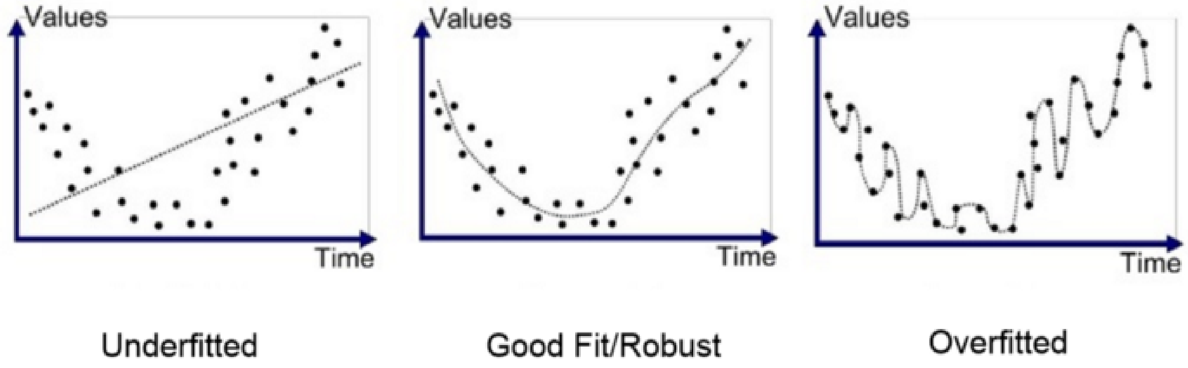

Figure 6.2: Underfitting and overfitting examples

We now return on the statement about the error on the prediction done by a complex model. In Fig. 6.3 we can see the prediction error done on the training and on the test set. We can notice that the train error decreases with the model complexity, this happens because we try to fit in the best way the train data and so the result obtained in this set will be very close to the real result. For the test prediction error instead the model is not trained on this data and so the error will decrease until having a minimum and then it will increase because we are introducing an high variance in the points error.

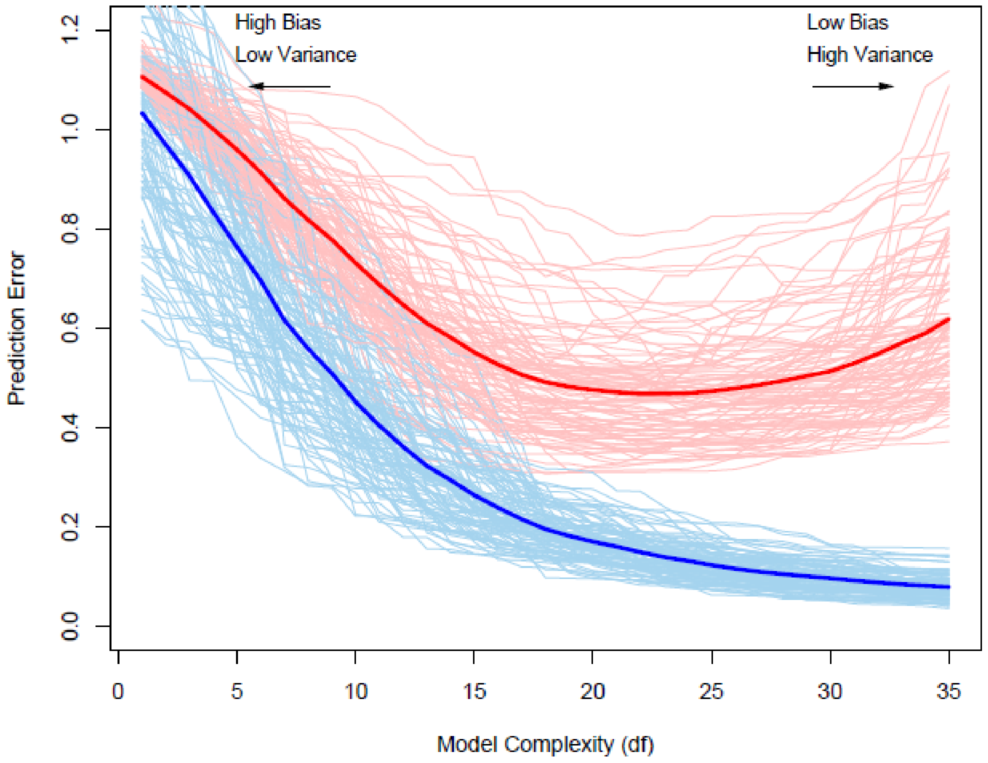

Figure 6.3: Train error and test prediction error at different model complexity

This problem can be solved introducing a penalization on the hypothesis parameter dimension. Two examples which use this penalization are the Ridge and Lasso model.

6.2.1 Ridge Linear model

In this model the cost function has an additional element which is the square of the hypothesis parameters.

$$J(\boldsymbol{w}) = \frac{1}{2}\sum_{i=1}^{N}(y_n - w_0 - \sum_{j=1}^{D} w_j x_{ij})^2 + \lambda \sum_{i=1}^{D} w_i^2 \tag{6.7}$$

The last term depends on the penalization term λ which has the constraint $\lambda > 0$ and on the square of the parameters w_i. Using this constraint is as we we want to find the solution of Eq.(6.2) and imposing the condition that the parameters must be inside the sphere with radius t in the parameter space

$$\lambda \sum_{i=1}^{D} w_i^2 < t^2 \tag{6.8}$$

6.2.2 Lasso linear model

This model is similar to the Kernel but now the penalization term on the cost function is the absolute value of the hypothesis parameters

$$J(w_i) = \frac{1}{2}\sum_{i=1}^{N}(y_n - w_0 - \sum_{j=1}^{D} w_j x_{ij})^2 + \lambda \sum_{i=1}^{D} |w_i| \tag{6.9}$$

using this penalization the constraint on the parameters are

$$\lambda \sum_{i=1}^{D} |w_i| < t \tag{6.10}$$

which is the equation of a square with diagonal $2t$ in the parameter space.

To understand the different constraints for the two models we can look at Fig. 6.4 in which we can see the contour plot of the cost function and the constraints of the two models. Since we have a two parameters dependence in this case the constraint are: for the kernel model $w_1^2 + w_2^2 < t^2$ and for the Lasso $|w_1| + |w_2| < t$. We can see that the solution to the machine learning algorithm is given by the point for which the contour plot and the constraint of the models are tangent. In this example the advantage of using the Lasso model is that the tangent is quite always in one of the edges of the square so the final solution will set some of the parameter to zero. For example in figure we can see that the tangent is in a point of the square for which the parameter $\beta_1 = w_1$ is zero so in the final solution we will have reduced the number of parameters on which the model depends. In this way the input values associated with the parameter w_1 are not used and this allows us to find the input values which are not essential for the model.

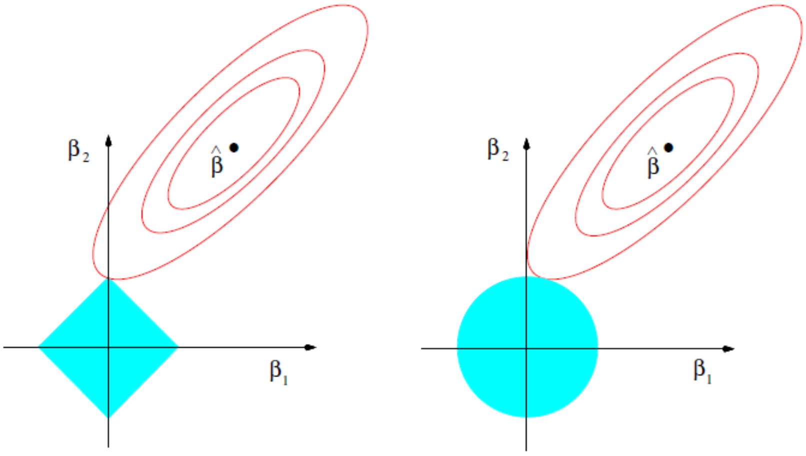

Figure 6.4: Kernel ridge constraint (right) and Lasso Constraint (left), in red the contour plot of the cost function

We will now introduce the models used in our analysis and we will focus our attention on their properties and on the set parameters they need as input and their meaning.

6.3 Kernel Models

This class of models is based on the use of a kernel space instead of using the parameter space to predict the result. Using this technique we are able to compute the cost function using the kernels which are defined as

$$k(x, x') = \boldsymbol{\phi}(x)^{\boldsymbol{T}} \boldsymbol{\phi}(x') = \sum_{i=1}^{D} \phi_i(x)\phi_i(x') \tag{6.11}$$

and so we can avoid the use of radial basis function. Now we will briefly see how we can replace the parameters using the kernel in the cost function and then we will discuss the advantages using this substitution.
The cost function written using vector notation in the parameter space is

$$J(\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\phi}(\boldsymbol{x_n}) - y_n)^2 + \frac{\lambda}{2}\boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{w} \tag{6.12}$$

and we can write the solution for **w** using a linear combination of the basis functions

$$\boldsymbol{w} = \sum_{n=1}^{N} a_n \boldsymbol{\phi}(\boldsymbol{x_n}) = \boldsymbol{\Phi}^{\boldsymbol{T}}\boldsymbol{a} \tag{6.13}$$

Where $\boldsymbol{\Phi}$ is the matrix having as n^{th} row $\boldsymbol{\phi}(\boldsymbol{x_n})^T$. If we insert this solution in the cost function and define the matrix $\boldsymbol{K} = \boldsymbol{\phi}\boldsymbol{\phi}^{\boldsymbol{T}}$ we have

$$J(\boldsymbol{a}) = \frac{1}{2}\boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{K}\boldsymbol{K}\boldsymbol{a} - \boldsymbol{a}^{\boldsymbol{t}}\boldsymbol{K}\boldsymbol{y} + \frac{1}{2}\boldsymbol{y}^{\boldsymbol{t}}\boldsymbol{y} + \frac{\lambda}{2}\boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{K}\boldsymbol{a} \tag{6.14}$$

from this cost function we can get the solution of the vector **a** which minimize $J(\boldsymbol{a})$ and then we can write the solution for $y(\boldsymbol{x})$ as

$$y(\boldsymbol{x}) = \boldsymbol{k}(\boldsymbol{x})^{\boldsymbol{T}}(\boldsymbol{K} + \lambda\boldsymbol{I_N})^{-1}\boldsymbol{y} \tag{6.15}$$

where $\boldsymbol{k}$ is the vector whose elements are $k_n(x) = k(x_n, x)$. So using this substitutions we have shown that we can write the solution for the model by using only the kernel functions. So we can modify the problem by inverting the kernel matrix instead of inverting the basis function matrix. In the first case we have a **NxN** matrix where **N** is the dimension of the train set while in the parameter space we have an **MxM** matrix where **M** is the dimension of our parameter. In general **N** is much larger than **M** so it needs more computational power in order to invert the matrix but the use of the kernel notations allows us to avoid the introduction of the basis function and use a feature space with an higher dimension.
In Fig. 6.5 we have an example of construction of a kernel function using a polynomial, a Gaussian and a sigmoidal basis function where we vary x and keep fixed the point $x' = 0$.

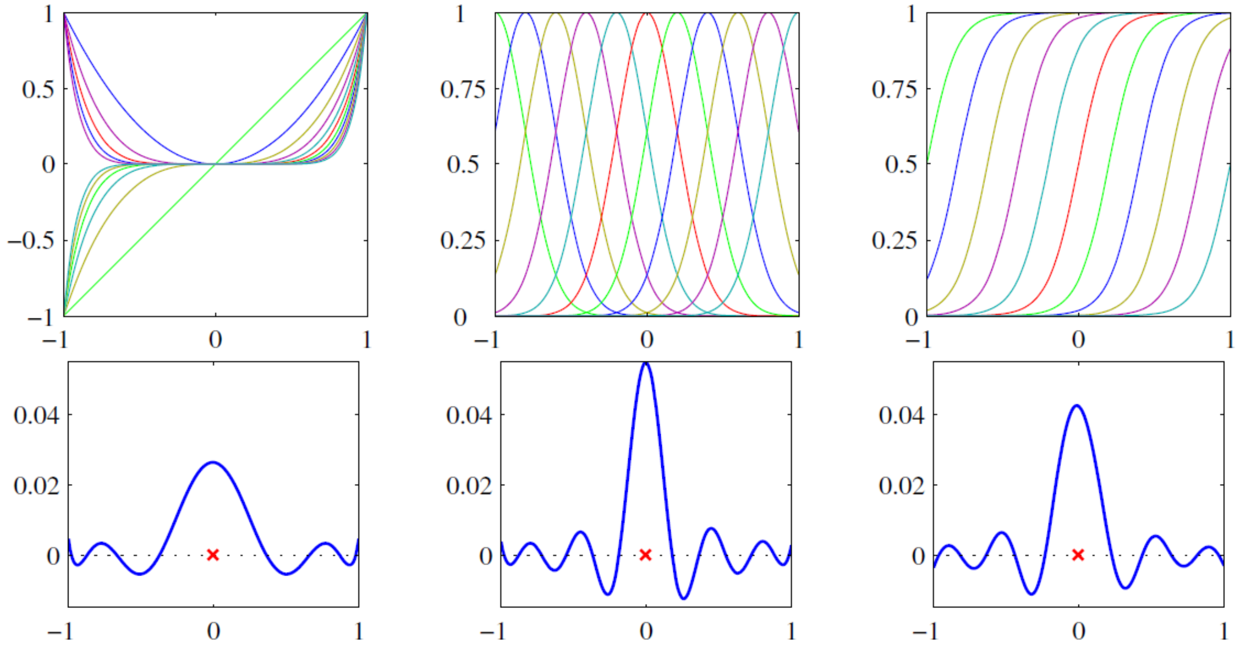

Figure 6.5: Examples of train construction using $x' = 0$, on the left we have a polynomial basis function, in the middle a Gaussian and on the right a sigmoidal one

Now that we have the information on what is a kernel functions and how they are constructed we will explain how they works and how they are able to predict the result for a given input vector. The kernel methods are called memory based since the prediction of a new result will be done using the entire training dataset or a subset of it. So when a new feature vector is given as input the algorithm searches the similar input values of the training dataset and uses them to compute the result of the new value. This approach is completely different respect to the parameter optimization since in the parameter case we find a set of optimized parameters which models the data and after that the training dataset becomes useless and is not used in the prediction.
The kernel methods require less time to train the model but will require more time to predict new values since the algorithm needs to use the kernel function to compute the new result.
We now see some of the kernel functions we will use in our analysis: the first example is the polynomial kernel which is:

$$k(\boldsymbol{x}, \boldsymbol{x'}) = (\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{x'} + c)^M \tag{6.16}$$

and the second is the radial basis function which we already saw in Eq. (6.4) but now it has at exponential the euclidean distance of the two vector $\boldsymbol{x}$ and $\boldsymbol{x'}$

$$k(x, x') = exp\Big(- \frac{||\boldsymbol{x} - \boldsymbol{x'}||^2}{2\sigma^2}\Big) \tag{6.17}$$

6.3.1 Gaussian process regressor

In this method we use the concept of bayesan probability to find the probability distribution of the result, $\boldsymbol{y}$, given an input vector $\boldsymbol{x}$, $p(\boldsymbol{y}|\boldsymbol{x})$. The key idea on this process is that the result $\boldsymbol{y}$ must be a collection of random variables and have a joint Gaussian distribution. In order to do so we introduce a noise in the target result and obtain a new function $t_n = y_n + \epsilon_n$ where ϵ_n follows a Gaussian distribution. In order to be a gaussian process the mean of ϵ must be zero and the covariance can be chosen on the basis of the data variance. So after we introduced a prior distribution in the function space which results in a prior distribution on y we use the training set data to find a posterior distribution and compute the final result and its distribution. In Fig. 6.6 we can see on the left a prior gaussian process in which we have a high noise and an high length scale and so the points distribution is explained by the noise while on the right we have a posterior distribution and we can see that the shaded area which is the standard deviation is limited by the noise obtained from the train set (crossed points). The colored curves are different random functions which satisfy the boundary conditions given by the points. We can notice that on the right side of the plot where there are no train points the standard deviation of the model increases and we can have a lot of functions which can lies inside this area.

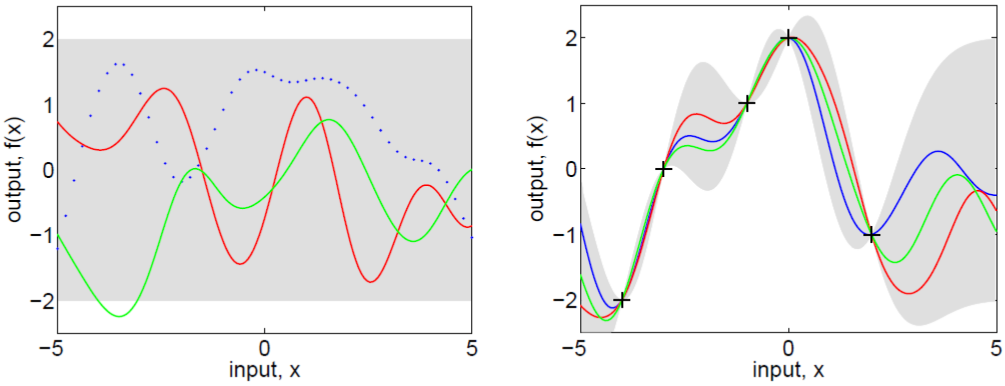

Figure 6.6: Examples of prior distribution on the left with two random functions and in dotted the y values, on the right we have three random functions computed using the posterior distribution of the noise obtained by training points [15]

In this model we have to fix the kernel to use, α which is an indicator of the strength of the noise and the length scale of the kernel which is introduced inside the kernel function in Eq. (6.17) as σ. When choosing the kernel we can also introduce some white noise which can be used as identical and equally distributed on all the training set.

6.3.2 Kernel Ridge

This method combines the penalization term seen in the Ridge method with the kernel trick. In this method the algorithm learns a linear function in the space which is induced by the kernel and the input data. So if we use a non linear kernel the function in the original space will be non linear. In this model we can set the kernel, a parameter α which is called regularization strength and a parameter gamma which corresponds to $\gamma = 1/2l^2$ in Eq.(6.17).

6.3.3 Support vector machine

The support vector machine has a similar process to the kernel ridge but in this case the penalization factor on the cost function is different. In this case the cost function is linear and has a region of insensitive error as can be seen in Fig. 6.7. In this region there is no penalization on the cost function and so all the error which are less than ϵ has zero contributions. The cost function for support vector machine is

$$C\sum_{n=1}^{N} E_\epsilon(y(x_n) - y_n) + \frac{1}{2}||\boldsymbol{w}||^2 \tag{6.18}$$

where

$$E_\epsilon(y(x) - t) = \begin{cases} 0 & if|y(x) - y| < \epsilon \\ |y(x) - t| - \epsilon & otherwise \end{cases} \tag{6.19}$$

The advantage of using this model is that using the parameter ϵ we can save memory and reduce the number of steps the algorithm will do. In fact the model doesn't use all the training dataset but if two points are close to each other respect to the parameter ϵ one of them will be discarded and so we reduce the number of training points.

In this model we have to set the parameter ϵ, the parameter C which is related to the model complexity and the kernel to use.

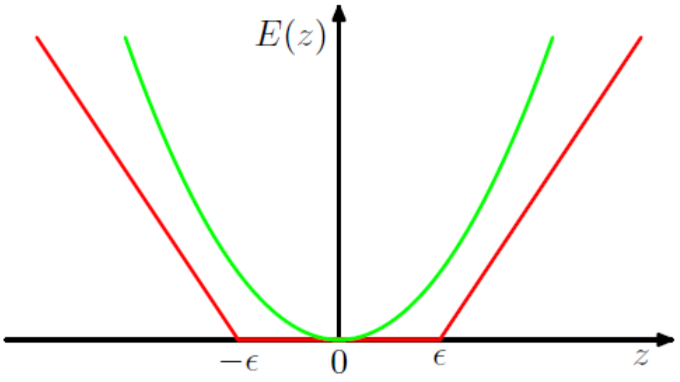

Figure 6.7: Support vector machine cost function [14]

6.4 Random forest regressor

This class of algorithm uses a random process and a set of decision trees to build a model. In Fig. 6.8 and 6.9 we can see an example of decision tree in which the space is divided using some random values and then the points are classified using the conditions on the tree leaves. When we use more than one decision tree on the same input space we use a method called Random forest. Using more decision trees allows us to gain higher precision on the prediction since we are able to divide the input space in very narrow region. This class of algorithm is most used in classification problem since starting from a new point we can easily compute to which region it belongs using the decision trees but we can use it also for regression problems. In the last case the result of the decision trees are averaged in order to improve the predictive accuracy and avoid overfitting problems. In fact if we have two points which are very close to each other in the input space using only the decision trees we will predict the same result values for both of them. This problem leads to an overfitting procedure on the training set which we should avoid and so the average between near region is used to avoid overfitting.

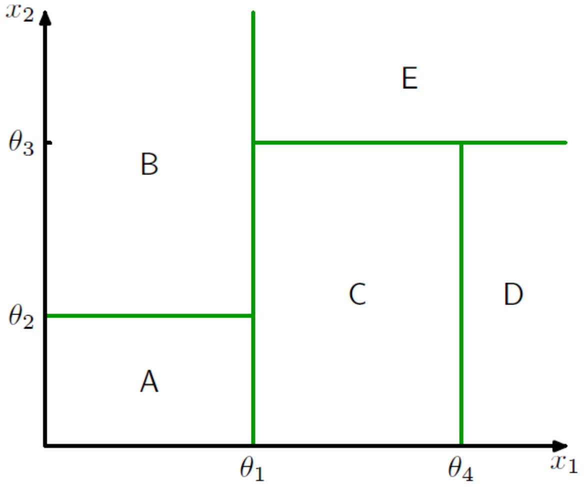

Figure 6.8: The input space is divided into five regions using the parameter θ_i [14]

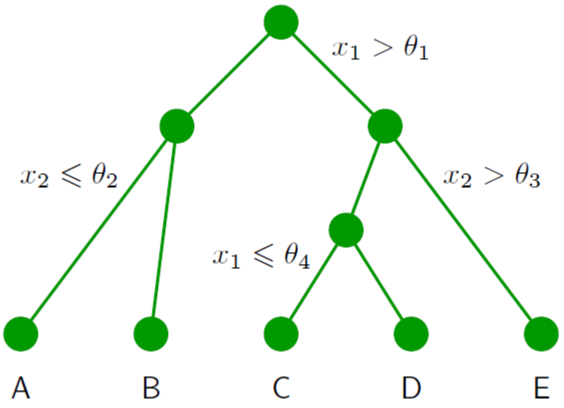

Figure 6.9: Decision tree used to classify the points [14]

6.5 Analysis

In our analysis we chose the input parameters using the Lasso model and varying the α parameter. In fact if we plot the coefficient of a given feature respect the α parameter we can see the weight each feature has and which of them can be neglected in the analysis. We can see in Fig. 6.10 the result we obtained and notice that the amplitude of the curve has zero weight and so we can neglect it in our analysis. We should have expected this result since the amplitude of the curve depends on the electronics used to acquire the data and has some dependence also with the temperature.

As input values for the algorithm we use the frequency and the phase of the resonance curve, the conductivity and the phase shift of the signal.

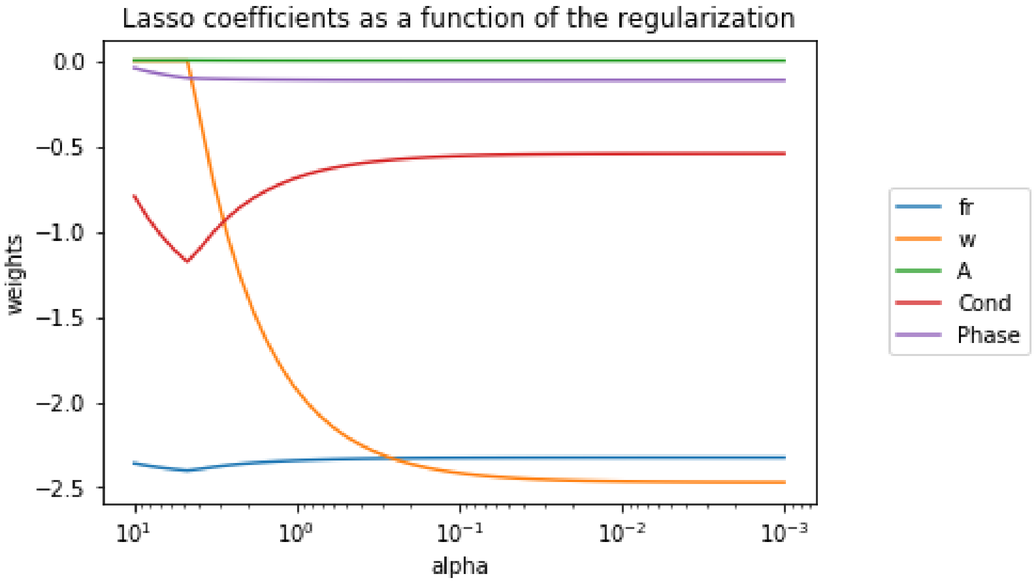

Figure 6.10: Lasso model respect the regularization parameter α

Since the algorithm requires a huge number of data points to perform the training we achieved results with the best accuracy on positions A dataset and we will display only them in this analysis. The training subset for this analysis was fixed at 75 % and so the test subset for which we plot the results contains 25 % of the total number of data.

During the analysis we performed an iteration on the different algorithm parameters in order to found the best parameter for each method and we did some other type of analysis to validate the results. In the next sections we will discuss briefly the type of analysis used to validate the model.

6.5.1 Cross validation

The cross validation is a method used in training model but can also be used to verify if the result are biased using a given subset. The idea of cross validation is to perform the training of the model using a given fraction of the dataset and keep the other subset to validate the model. Then this procedure is repeated using another subset as validation set and then again until all the subset are used as validation. For example if we keep the validation subset fixed at 25% of the dataset this method will perform 4 run on the subset to validate the model as can be seen in Fig. 6.11 in which all the dataset is divided in four subset and each run uses a different subset as validation.

This procedure can be used also in the result prediction since given the test subset we can end up having a prediction biased on the training set or having some outlier which are not predicted by the model. So in order to avoid having results coming from a lucky or unlucky case we perform a cross validation changing the train and test dataset and see the result on each subset. If the model doesn't have overfitting problems the result of all the different subsets should be very similar.

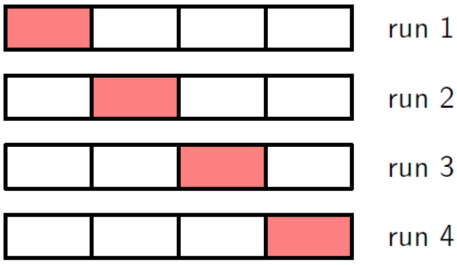

Figure 6.11: Cross validation example [14]

6.5.2 Learning curves

In this type of analysis we can plot the standard deviation of a given model respect the training set size. So we see the result of the prediction using a given fraction of the dataset as train and then we vary this fraction and record the results. In this way we can see how many points in the training set we must have in order to achieve a given standard deviation. In particular in our analysis the points on the learning curve are obtained as the average of the standard deviation of the residue on a cross validation of the set and the shaded area represents the 2 standard deviation of this mean.

6.6 Results

6.6.1 Gaussian process

In this case the best kernel for the analysis was a polynomial with power degree equal to 2. But we can see that this kernel was not able to correct some of the pattern in the data, in fact we can notice that in the water continuous part we have an increasing tendency for the residue respect the WLR. We can also notice that the histogram of the residue is not a Gaussian as one should expect since it is altered from this residue behavior. In the box plot on the upper right we can see the cross validation for this model using different training and data set, each box represents a different train set but the number of points inside the subset doesn't changes. We can notice that the boxes are similar and has a little number of outlier so the model can be considered coherent. With this model we obtain the results in Tab. 6.1.

Result obtained from Gaussian process			
	Mean	std	95 percentile
Gaussian process	1.0	6.4	11.8

Table 6.1: Gaussian process.

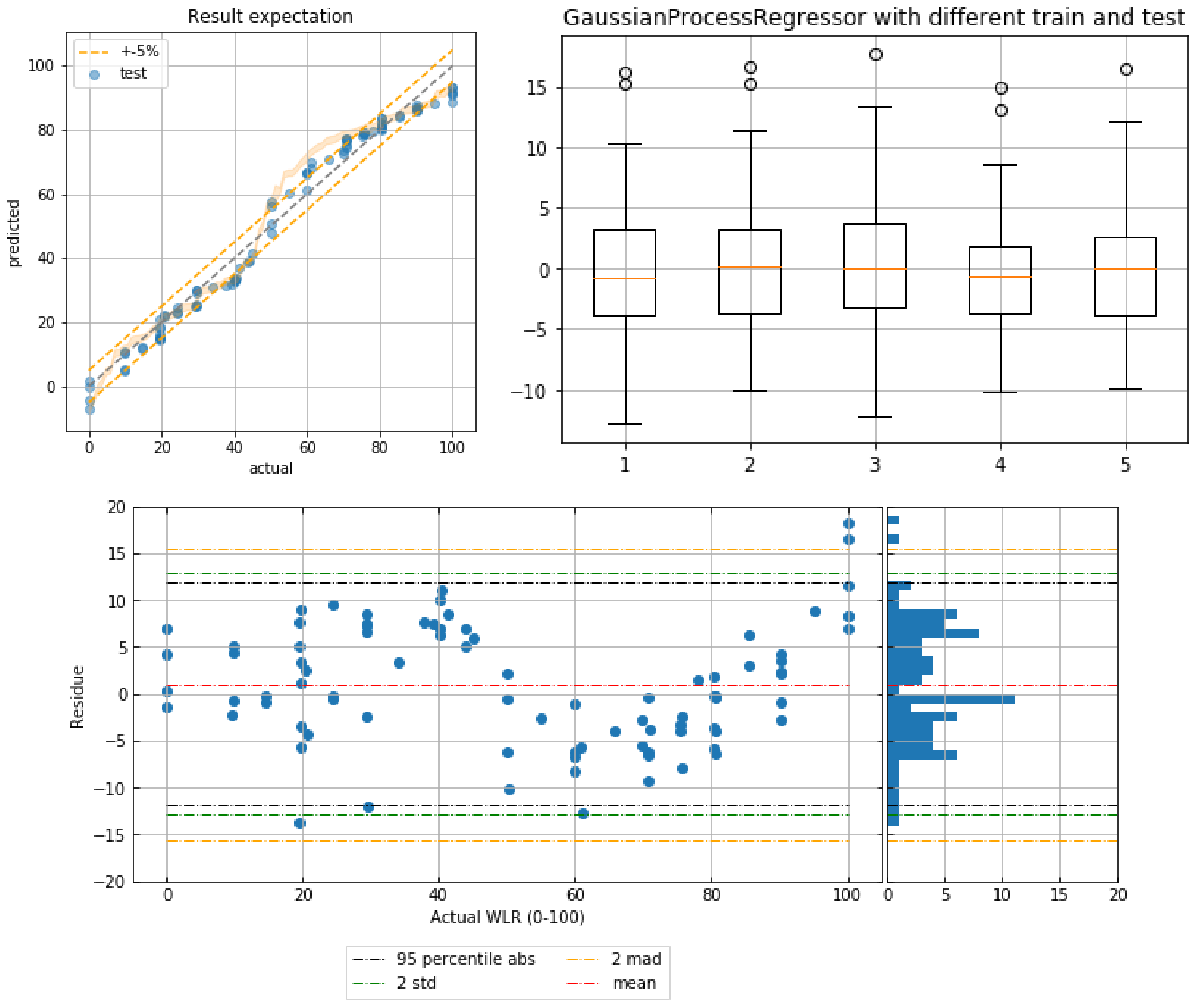

Figure 6.12: Result from the Gaussian process model

In Fig. 6.13 we can see the learning curve for the gaussian process. In particular we can notice that the model reaches a value for 2 std around 10 at 250 training points and then remain constant. This means that in order to achieve a 2 std of 10 we need at least 250 points to train our model. The fact that when we increase the training set we do not increase the accuracy can be seen as a positive or negative fact. It is positive since we can use the least number of training points and still reach the best accuracy for the model but it can also be seen in a negative way since the model do not increase its prediction score when we add more training data. So in this case we reached the limit on the best accuracy we can obtain for this model.

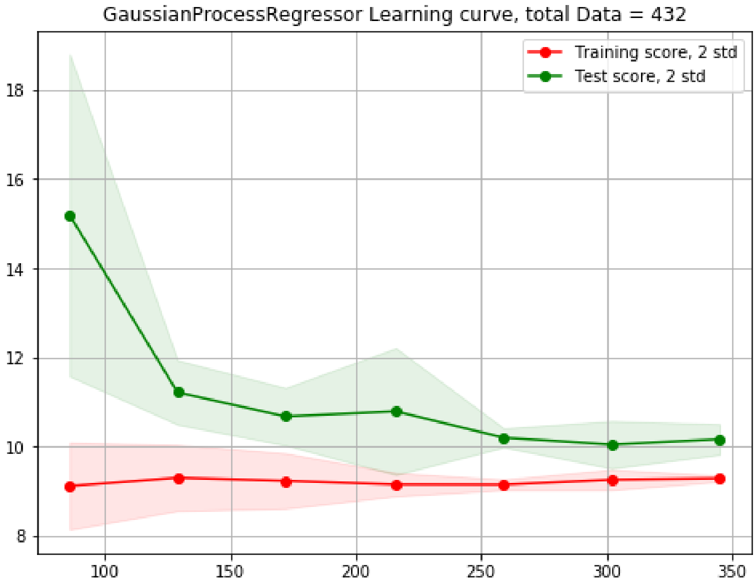

Figure 6.13: Learning curve for a gaussian process

6.6.2 Kernel Ridge

In this model the best kernel was a radial basis function instead of a polynomial function. We can notice from the residue plot and the histogram that this kernel is able to predict the points pattern. In fact the residue do not follows any tendency and the histogram is a Gaussian with mean equal to zero.
In this case we can see from the box plot that the first and third quartile of the residue are smaller respect the previous case but we can also notice that for a test subset we can have few outlier with a large residue. These outlier we have on the box 5 can be explained by a combinations of points which lies in a region where there were fewer or no data points in the training set and so their result is biased. This particular combination happens only on the last test subset so the model is still valid in predicting the result and we can also see that the quartile lines are similar and so also the median values.

Result obtained from Kernel Ridge

	Mean	std	95 percentile
Kernel Ridge	0.0	2.6	4.9

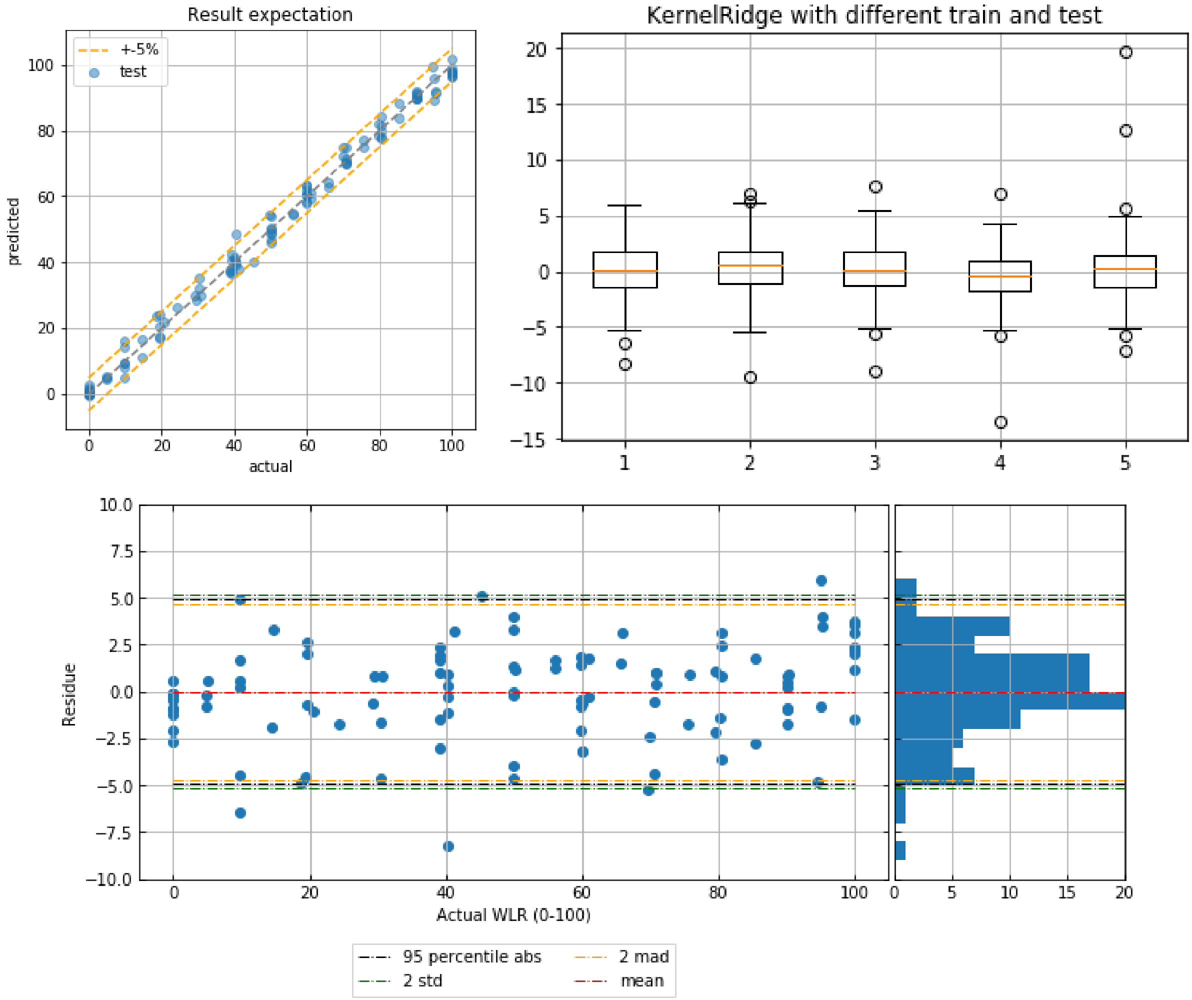

Figure 6.14: Result from the Kernel Ridge model

In this learning curve we can see that we have a small decrease even at higher size for the training subset so in principle we can hypothetically increase the prediction result accuracy by adding new training points. This increase in the prediction accuracy although come at a very high cost in terms of data acquisition time since we need a big number of data to gain a little increase in accuracy. Since the accuracy achieved with the size used as train in the analysis is already in the wanted accuracy there is no need to increase the number of data points.

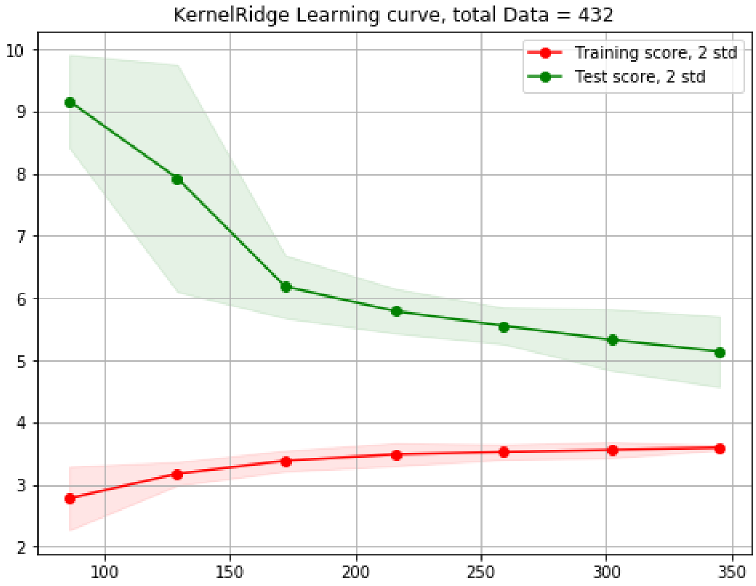

Figure 6.15: Learning curve for Kernel Ridge

6.6.3 Support Vector Machine

For this model the result are similar with the ones obtained with the kernel ridge and also in this case the best kernel functions is the radial basis function. Besides this model has some problems respect to the kernel ridge. In fact in the box plot we can see that we have more than one test subset with large residue which means that using SVM we are not able to predict some of the data and we has as result an outlier.

Overall the result displayed in table are similar but the presence of this residue lead us to prefer the use of the previous method. Also in this case we have a gaussian histogram for the residue but we can notice some patterns which are not corrected by the model. The same pattern can also be seen in the Kernel ridge but in the SVM method they are clearly visible respect to the previous case.

Result obtained from Support Vector machine			
	Mean	std	95 percentile
SVM	-0.2	2.7	5.5

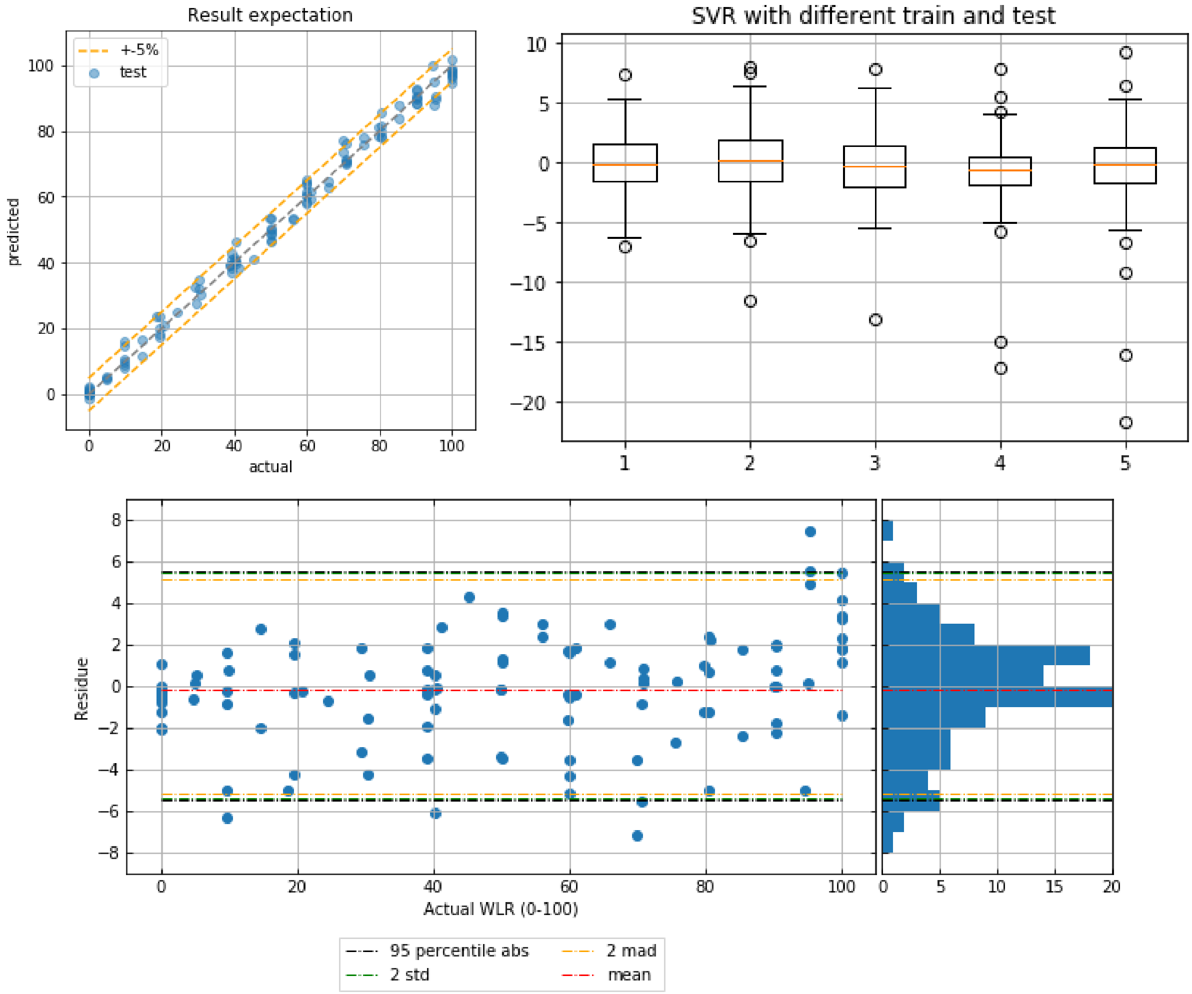

Figure 6.16: Result from the Support vector machine model

Also in the learning curve we can notice the presence of outliers since the shaded are which is the 2 times the standard deviation of the mean is larger. This represents that the mean value of the residue has a larger variance in this case respect to the kernel ridge where all the mean value of the residue were similar.

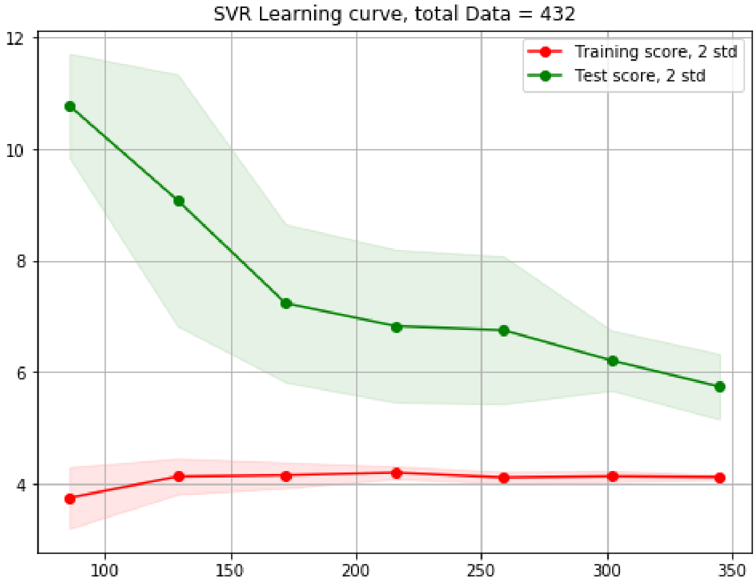

Figure 6.17: Learning curve for Support vector machine

6.6.4 Random Forest

This model has not need to set any parameter so there was no need to select a kernel. However because of the random process of the training we expect the prediction to vary respect to the training set in use.

From the residue plot we can notice that also this model is not able to correct some pattern of the data since we see an increasing tendency of the residue in water continuous regime.

In the box plot we can also notice that the boxes are different one from the other and this represents the random process of this model as we stated before. In principle if the model had a Gaussian histogram for the residue we would have not seen this difference in the box plot but we would expect the result of being independent of the training subset.

Result obtained from Random Forest			
	Mean	std	95 percentile
Random Forest	-0.4	3.5	7.6

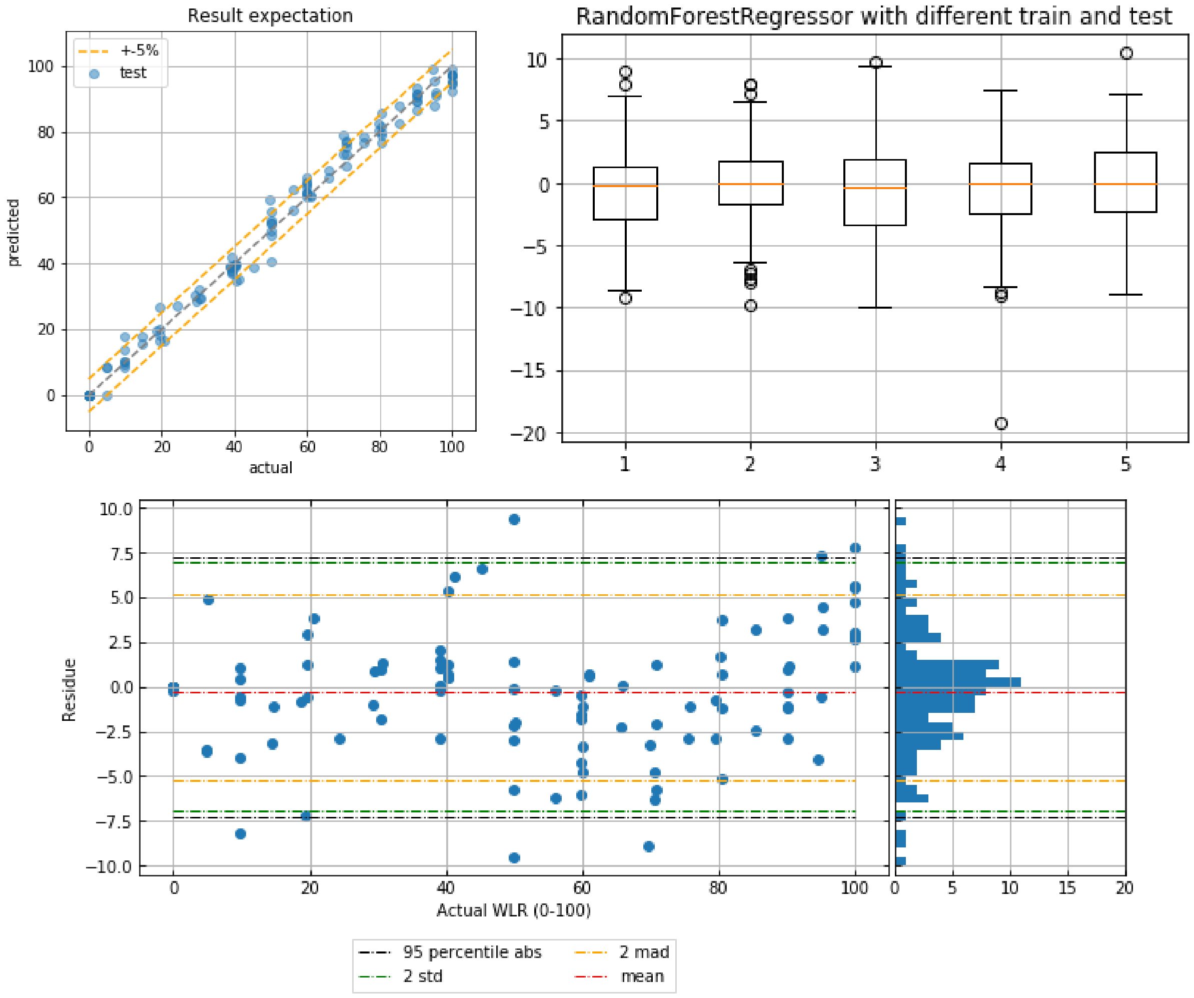

Figure 6.18: Result from the Random Forest model

From the learning curve we can see that the decrease of the curve respect to the training size is larger in this method respect to the previous. This is because for this model if we increase the train set we will have more leaves in the decision trees and so we are able to increase the classification of the points. Beside this increase in accuracy the overall result obtained with this method has a lower accuracy respect to the kernel ridge and SMV.

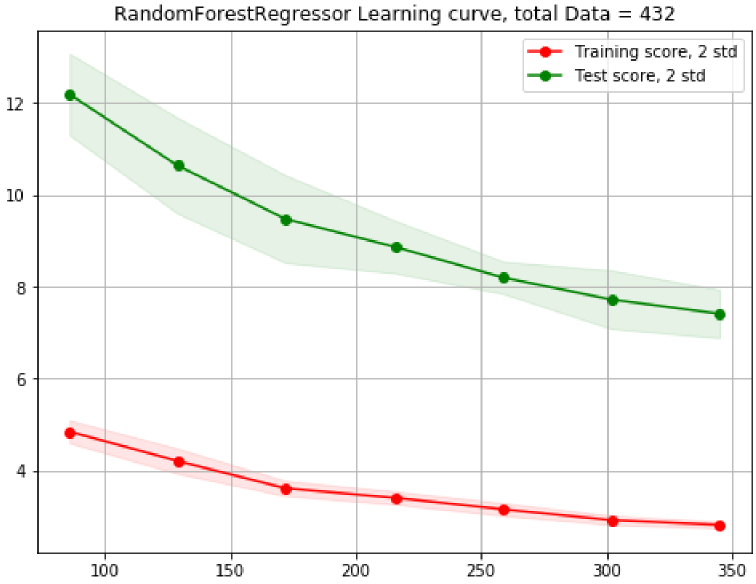

Figure 6.19: Learning curve for Random Forest

6.6.5 Result Comparison

In the next table we resume the result obtained with the different methods used in the analysis. As explained above the Gaussian process and the random forest methods were not able to correct all the patterns in the dataset and so we have a larger 95 percentile and standard deviation. In particular we can also see that for these models the mean is not close to zero which means that we have a deviation on the residue as can be seen also from the residue histogram.
For the other two models we have similar results but from the box plot and the learning curve we can notice that the Support vector machine method has more outlier respect to the Kernel Ridge model. For this reason we prefer to use the Kernel Ridge model to analyze our data.
For the Kernel ridge model we obtain a 95 percentile of 4.9 which is inside the wanted accuracy for the prediction but we reach this level of accuracy only using the 75 % of the dataset as training. If we reduce the training set we reduce the accuracy on the prediction and we will not have the 95 percentile inside the region ±5 as can be seen in the learning curve plot, Fig. 6.15.

Result obtained from different models

	Mean	std	95 percentile
Gaussian process	1.0	6.43	11.8
Kernel Ridge	0.0	2.6	4.9
SVM	-0.2	2.7	5.5
Random Forest	-0.4	3.5	7.6

Chapter 7

Conclusions

We used two different approaches to compute the WLR starting from the resonance curve parameters and the conductivity of water. In the first case we used the Bruggeman mixing model to compute the permittivity of the mixture and we related this value to the resonance curve properties. In the second we used a machine learning process to learn a model and predict the result for new values.
In the first case we found as best result:

Best result obtained

Regime	Mean ± percentile	Score
Water continuous	0.2 ± 3.3	90 percentile
Oil continuous	-2.9 ± 1.1	95 percentile

These results are promising since we can achieve a good accuracy in the oil continuous regime while in the water continuous the accuracy decreases. Although these results are obtained using a combination of two different probes and some problems may arise. The major problem is the switch process between the use of a probe and the other which is not defined yet. Another problem in this configuration is the inversion region which is critical for the signal extraction and causes a loss of accuracy on the measure. In particular it is worth to notice that the results in oil continuous are obtained using only the points at 10 and 16 S/m since we had to change a piece in the circuit to place the probe. For this reason the results do not cover the entire range of conductivities which can occur in a working regime for an oil well and we need to do further investigation on this position.
Using the machine learning algorithms the best result obtained was given by the Kernel ridge model:

Result obtained from different models

	Mean	std	95 percentile
Kernel Ridge	0.0	2.6	4.9

We can notice that the machine learning algorithm offers a better accuracy on the prediction over all the regimes of data but some statement must be done since these two methods are different and must be applied in different cases.
In the first case to obtain the result we must have only a small dataset of calibration points which are taken in a static fluid condition. On the other hand, to achieve the accuracy shown in table for the machine learning algorithm we need to have a training set of data with at least 300 points, furthermore this process has some other limitations. In fact the machine learning

algorithm trains its model on the data taken in certain conditions. If for example we change the dimension of the pipe, the model used to predict the data for position A will not work anymore since the conditions are changed. Instead the first method can in principle be applied to different system conditions and give the same result. This type of analysis is not done yet but will be completed when we will finish the point acquisition at DNV facilities in Holland where a working condition and a flow similar to the industrial ones is used.

So the key question is on the type of result we want to achieve and if the production of the microwave sensors is to be in series or if any meter can be calibrated under its own flow conditions. In the first case we have seen we can achieve the accuracy wanted in oil continuous while in water we must use the 90 percentile to be inside the ±5 region for the residue. In the second case we reached the accuracy wanted for the entire spectrum of WLR but each meter must be calibrated and have its own training set built using the same working conditions of the meter. The first type of analysis requires less work on calibration and will hypothetically work under any flow conditions but do not have the accuracy wanted while the machine learning will need a lot of calibration work and time in order to achieve the given accuracy for the meter.

Further investigation can be done in order to increase the accuracy. The major problem we had during the analysis was the signal extraction, especially in the inversion region where the signal was very noisy. Further studies can be done on this process of signal extraction in order to achieve a better accuracy. Also in the machine learning algorithms we can increase the analysis process by introducing an end-to-end dual convolutional neural network. This type of analysis uses as input the entire resonance curve in order to build a model to predict the new results. In this way the algorithm is able to use some other informations as input values such as the noise of the curves and the resonance frequency variability respect to the only curve parameters.

These further adjustment on the analysis process are going to be introduced after the preliminary results coming from the DNV facilities in order to include them using the data coming from an industrial regime working situation.

The results obtained nowadays are promising since there are not existing meters suited for the wet gas regime which can work in all the WLR range using only two microwave probes or a simpler setup. This allows to introduce a new meter on the market which can be competitive in terms of cost and accuracy. Although its effectiveness on an industrial working regime must be verified and we need to perform more measurement using the probe on position C in order to verify its accuracy.

Bibliography

[1] NEL, *An introduction to multiphase flow measurements*, National Measurement system

[2] Fenghui Han, Muk Chen Ong, Yihan Xing, Wenhua Li, *Three-dimensional numerical investigation of laminar flow in blind-tee pipes*, Ocean Engineering, Vol. 217, 2020

[3] B.T. Hjertaker, S.-A. Tjugum, A. Hallanger, R. Maad,*Characterization of multiphase flow blind-T mixing using high speed gamma-ray tomometry*, Flow Measurement and Instrumentation, Volume 62, 2018, Pages 205-212

[4] OpenFoam, Web page [online]. Available: `http://www.openfoam.com/`

[5] Gioia Falcone, G.F. Hewitt, Claudio Alimonti, *Multiphase Flow Metering Principles & Applications*, Developments in Petroleum Science, Elsevier, Volume 54, 2009

[6] Schlumberger, *Fundamentals of multiphase metering*, Schlumberger, Houston, USA

[7] E.G. Nyfors, *Cilindical microwave resonators sensors for measuring materials under flow*, Helsinki University of Technology, Department of Electrical and Communications Engineering, Helsinki, 2000

[8] Sihvola A., 1999,*Electromagnetic Mixing Formulas and Applications*, IET Electromagnetic Waves Series 47, The Institution of Engineering and Technology, London, UK

[9] Bruggeman, D. A. G. *Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen*, Annalen der Physik, 1935, Series 5, pp.639-679

[10] P. N. Sen, C. Scala, and M. H. Cohen *A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads* GEOPHYSICS 46: 781-795, 1981.

[11] Prafull Sharma, Liyun Lao, Gioia Falcone, *A microwave cavity resonator sensore for water-in-oil measurement*, Oil and gas engineering centre, Cranfield University, United Kingdom

[12] D H Gadani, *Effect of salinity on the dielcetric properties of water*, Indian Journal of Pure & applied Physics, vol. 50, 2012, pp.405-410

[13] Kevin P. Murphy, *Machine Learning, a probabilistic perspective*, MIT press, 2012

[14] Christopher M. Bishop, *Pattern recognition and machine learning*, Springer science, Cambridge, 2006

[15] C. E. Rasmussen & C. K. I. Williams, *Gaussian Processes for Machine Learning*, the MIT Press, 2006