Universitat de les Illes Balears

Shallow Water Remote Sensing Using Sonar Improved with Geostatistics and Stochastic Resonance Data Processing

Doctoral Thesis

Universitat de les Illes Balears

Autor
André Luis Sousa Sena

2018

Universitat
de les Illes Balears

DOCTORAL THESIS

2018

Doctoral Programme of *Electronic Engineer*

Shallow Water Remote Sensing Using Sonar Improved with Geostatistics and Stochastic Resonance Data Processing

André Luis Sousa Sena
(doctorate)

Thesis Supervisors

Gerardo Gabriel Acosta
Wojciech Korneta
Alejandro Rozenfeld

Thesis tutor

Rodrigo Picos Gaya

Doctor by the Universitat de les Illes Balears

Shallow Water Remote Sensing Using Sonar Improved with Geostatistics and Stochastic Resonance Data Processing

By

André Luis Sousa Sena

B.S. (Faculdade de Ciência e Tecnologia – Área 1) 2005
M.S. (Universitat de Les Illes Balears – UIB) 2009

THESIS

Presented to the Physics Department of
Universitat de les Illes Balears
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR
in
ELECTRONIC ENGINEER

February
2018

Universitat de les Illes Balears

Universitat de les Illes Balears

Prof. Dr. Gerardo Gabriel Lazo Acosta, Prof. Dr. Hab. Wojciech Korneta e Prof. Dr. Alejandro Fabian Rozenfeld, as the thesis advisor, and Prof Rodrigo Picos Gaya acting as presenter from the UIB Physics Department, certify that this thesis with title **"Shallow Water Remote Sensing Using Sonar Improved with Geostatistics and Stochastic Resonance Data Processing"**, has been carried out by Mr. André Luis Sousa Sena, in order to obtain the degree of Doctor in Electronic Engineer at Universitat de les Illes Balears. Because that they sign the present document.

Palma, Mallorca, Illes Balears, 10 of January of 2018

Prof. Dr. Gerardo Gabriel Lazo Acosta
(Director)

Prof. Dr. Hab. Wojciech Korneta
(Director)

Prof Dr. Alejandro Fabian Rozenfeld
(Director)

Prof Dr. Rodrigo Picos Gaya
(Ponente)

Abstract

The main objective proposed in this doctoral thesis was focused on the study and development of a solution for the remote sensing of the submarine topographic relief, using for this, inexpensive equipment. Here we focus in three works that altogether will improve the remote sensing process for underwater medium using sonars as the main relief data acquisition system. The problem was systematically addressed since the previous work in the master thesis, constituting three stages: 1) building a prototype data extraction platform, 2) data acquisition and 3) data processing. In the first stage was developed a prototype taking as a reference the modular structure and the software design applied in AUVI project (Acosta 2008), and besides it was used the model of autonomous navigation system developed to AutoTracker project (Acosta et al., 2005), this stage and part of the following one were developed in the master thesis. In the second stage, it was done the survey planning and the acoustic data extraction including navigational data in three different places: i) On the cove "Estancia", located in Palma, Mallorca island/Spain, where we used the prototype developed in the first stage of this work as a platform to data extraction; ii) "Alfeite Arsenal" Port, located on the Tejo River, in Almada/Portugal into the context of robotics exercises promoted by the Navy of Portugal (REX2014). In this location, it was used the robotic vehicle ZARCO from the Oceansys Lab, through cooperative work with the Oceansys Lab. research group at the University of Porto (UP)/INESC in Portugal and finally, iii) in the "Bay of All Saints", near the city of Salvador/Brazil, where were performed some missions to collect data using hydrographic survey boats in cooperation with the Federal University of Bahia (UFBA) and Belov Engineering - Port Engineering and Hydrographic Services Company, both located in Salvador/Bahia. Finally, the third stage, that is the main body of this thesis, was characterized by the data analysis and comparison between several datasets. In this stage, studies had been conducted to verify the feasibility of the use of spatial statistical algorithms in the process of bathymetric data interpolation without any ancillary information to support the prediction. We determined an optimized procedure for estimating the unsampled points, hence it was validated using a regular cross-validation method and a comparative validation method to compare the estimated data with a second dataset acquired in the same region and acting as a Merit Figure. The average discrepancy between the estimated data and Merit Figure data value was 25 cm, it is below the acceptable error for bathymetric data at depths below 30m (IHO 2012). In addition, an algorithm based on the Stochastic Resonance (SR) theory was developed. It consists in applying white noise in an optimal intensity level to improve the contrasts of acoustic images generated by a side Sonar Scan (SSS). The SR theory also, was used as a basis for development of a weak signals detection algorithm in sensing applications. Regarding the sensors application for measure remote sensing physical variables, we could cite the magnetic field meters (magnetometers), and inertial sensors (accelerometers and gyroscopes), in this study was performed a simulation of Chua's circuit operating in a chaotic regime as a sensor, where we could determine successfully the region of system solution into the strange attractor, using, for this, the technique of residence time, which will be defined along this thesis. The entire set of modules, techniques and processes described in this work proposed one solution to the remote sensing problem applied to the underwater environment, and give the opportunity to develop in more deep way future works in sensor integration, algorithms and data acquisition platform.

Resum

L'objectiu principal proposat en aquesta tesi doctoral es va centrar en l'estudi i desenvolupament d'una solució per a la detecció remota del relleu topogràfic submarí, utilitzant per a això un equip econòmic. Aquí ens centrem en tres treballs que en total milloraran el procés de teledetecció per al medi submarí utilitzant sonars com el principal sistema d'adquisició de dades de relleu. El problema es va abordar sistemàticament des del treball anterior a la tesi de màster, que constava de tres etapes: 1) construcció d'una plataforma d'extracció de dades prototip, 2) adquisició de dades i 3) tractament de dades. En la primera etapa es va desenvolupar un prototip prenent com a referència l'estructura modular i el disseny de programari aplicat en el projecte AUVI (Acosta 2008), a més d'utilitzar el model de sistema de navegació autònom desenvolupat per al projecte AutoTracker (Acosta et al., 2005), aquesta etapa i una part de la següent es van desenvolupar en la tesi de màster. En la segona etapa, es va realitzar la planificació de l'enquesta i l'extracció de dades acústiques, incloses les dades de navegació en tres llocs diferents: i) A la cala "Estancia", situada a Palma, illa de Mallorca / Espanya, on utilitzem el prototip desenvolupat a la primera etapa d'aquest treball com a plataforma d'extracció de dades; ii) Port "Alfeite Arsenal", ubicat al riu Tajo, a Almada / Portugal, en el context dels exercicis de robòtica promoguts per l'Armada de Portugal (REX2014). En aquesta ubicació, es va utilitzar el vehicle robotitzat ZARCO del laboratori Oceansys, a través del treball cooperatiu amb l'Oceansys Lab. grup d'investigació de la Universitat de Porto (UP) / INESC a Portugal i, finalment, iii) a la "Badia de Tots Sants", a prop de la ciutat de Salvador / Brasil, on es van realitzar algunes missions per recollir dades utilitzant vaixells d'enquestes hidrogràfiques en cooperació amb la Universitat Federal de Bahia (UFBA) i Belov Engineering - Port Engineering and Hydrographic Services Company, ambdós ubicats a Salvador / Bahia. Finalment, la tercera etapa, que és el cos principal d'aquesta tesi, es va caracteritzar per l'anàlisi de dades i la comparació entre diversos conjunts de dades. En aquesta etapa, s'han realitzat estudis per verificar la viabilitat de l'ús d'algoritmes estadístics espacials en el procés d'interpolació de dades batimètriques sense cap tipus d'informació auxiliar per a la predicció. Es va determinar un procediment optimitzat per estimar els punts sense mostrejar, per tant, es va validar mitjançant un mètode de validació de mètodes regular i un mètode de validació comparatiu per comparar les dades estimades amb un segon conjunt de dades adquirit a la mateixa regió i actuant com a Figura de mèrit. La discrepància mitjana entre les dades estimades i el valor de dades de Merit Figure va ser de 25 cm, per sota de l'error acceptable per a dades batimètriques a profunditats inferiors a 30 m (IHO 2012). A més, es va desenvolupar un algorisme basat en la teoria de la ressonància estocàstica (SR). Consisteix en aplicar soroll blanc en un nivell d'intensitat òptima per millorar els contrastos d'imatges acústiques generades per un costat Sonar Scan (SSS). La teoria del SR també es va utilitzar com a base per al desenvolupament d'un algoritme de detecció de senyals feble en la detecció d'aplicacions. Pel que fa a l'aplicació de sensors per a la mesura de les variables físiques de control remot, podríem citar els mesuradors de camp magnètic (magnetòmetres) i els sensors inercials (acceleròmetres i giroscopis), en aquest estudi es va realitzar una simulació del circuit de Chua que funciona en un règim caòtic com a sensor, on podríem determinar amb èxit la regió de la solució del sistema en l'estrany atractor, utilitzant, per a això, la tècnica del temps de residència, que es definirà al llarg d'aquesta tesi. Tot el conjunt de mòduls, tècniques i processos descrits en aquest treball proposen una solució al problema de teledetecció aplicat a l'entorn submarí i permeten desenvolupar de manera més profunda futurs treballs en integració de sensors, algorismes i plataforma d'adquisició de dades..

Resumen

El principal objetivo propuesto en esta tesis doctoral se centró en el estudio y desarrollo de una solución para la detección remota del relieve topográfico submarino, utilizando para esto, equipos de bajo costo. Aquí nos enfocamos en tres trabajos que en conjunto mejorarán el proceso de detección remota para medios subacuáticos usando sonares como el principal sistema de adquisición de datos de alivio. El problema fue abordado sistemáticamente desde el trabajo anterior en la tesis de maestría, constituyendo tres etapas: 1) construcción de una plataforma prototipo de extracción de datos, 2) adquisición de datos y 3) procesamiento de datos. En la primera etapa se desarrolló un prototipo tomando como referencia la estructura modular y el diseño de software aplicado en el proyecto AUVI (Acosta 2008), y además se utilizó el modelo de sistema de navegación autónomo desarrollado para el proyecto AutoTracker (Acosta et al., 2005), esta etapa y parte de la siguiente se desarrollaron en la tesis de maestría. En la segunda etapa, se realizó la planificación de la encuesta y la extracción de datos acústicos, incluyendo datos de navegación en tres lugares diferentes: i) En la cala "Estancia", ubicada en Palma, isla de Mallorca / España, donde utilizamos el prototipo desarrollado en el primera etapa de este trabajo como una plataforma para la extracción de datos; ii) Puerto "Alfeite Arsenal", ubicado en el río Tajo, en Almada / Portugal en el contexto de ejercicios de robótica promovidos por la Armada de Portugal (REX2014). En esta ubicación, se utilizó el vehículo robótico ZARCO del Laboratorio Oceansys, a través del trabajo cooperativo con el Laboratorio Oceansys. grupo de investigación en la Universidad de Oporto (UP) / INESC en Portugal y finalmente, iii) en la "Bahía de Todos los Santos", cerca de la ciudad de Salvador / Brasil, donde se realizaron algunas misiones para recopilar datos utilizando barcos hidrográficos en cooperación con la Universidad Federal de Bahía (UFBA) y Belov Engineering - Compañía de Ingeniería Portuaria y Servicios Hidrográficos, ambas ubicadas en Salvador / Bahía. Finalmente, la tercera etapa, que es el cuerpo principal de esta tesis, se caracterizó por el análisis de datos y la comparación entre varios conjuntos de datos. En esta etapa, se realizaron estudios para verificar la viabilidad del uso de algoritmos estadísticos espaciales en el proceso de interpolación de datos batimétricos sin ninguna información auxiliar para respaldar la predicción. Determinamos un procedimiento optimizado para estimar los puntos no muestreados, por lo que se validó utilizando un método de validación cruzada regular y un método de validación comparativa para comparar los datos estimados con un segundo conjunto de datos adquiridos en la misma región y actuando como una figura de mérito. La discrepancia promedio entre los datos estimados y el valor de los datos de Merit Figure fue de 25 cm, está por debajo del error aceptable para los datos batimétricos a profundidades por debajo de 30 m (OHI 2012). Además, se desarrolló un algoritmo basado en la teoría de la Resonancia Estocástica (SR). Consiste en aplicar ruido blanco en un nivel de intensidad óptimo para mejorar los contrastes de las imágenes acústicas generadas por un Sonar Scan lateral (SSS). La teoría SR también se usó como base para el desarrollo de un algoritmo de detección de señales débiles en aplicaciones de detección. En cuanto a la aplicación de sensores para medir variables físicas de teledetección, podríamos citar los medidores de campo magnético (magnetómetros) y sensores inerciales (acelerómetros y giroscopios), en este estudio se realizó una simulación del circuito de Chua operando en un régimen caótico como sensor, donde pudimos determinar con éxito la región de la solución del sistema en el atractor extraño, utilizando, para ello, la técnica del tiempo de residencia, que se definirá a lo largo de esta tesis. El conjunto completo de módulos, técnicas y procesos descritos en este trabajo propuso una solución al problema de teledetección aplicado al entorno subacuático, y brinda la oportunidad de desarrollar de manera más profunda futuros trabajos de integración de sensores, algoritmos y plataforma de adquisición de datos.

Resumo

O objetivo principal proposto nesta tese de doutorado foi focado no estudo e desenvolvimento de uma solução para o sensoriamento remoto do alívio topográfico submarino, usando para isso, equipamentos baratos. Aqui nos concentramos em três trabalhos que, em conjunto, melhorarão o processo de sensoriamento remoto para o meio subaquático, utilizando os sonares como o principal sistema de aquisição de dados de alívio. O problema foi sistematicamente abordado desde o trabalho anterior na tese de mestrado, constituindo três etapas: 1) construção de um protótipo de plataforma de extração de dados, 2) aquisição de dados e 3) processamento de dados. Na primeira etapa foi desenvolvido um protótipo tomando como referência a estrutura modular e o projeto de software aplicado no projeto AUVI (Acosta 2008), e além disso, utilizou-se o modelo de sistema de navegação autônomo desenvolvido para o projeto AutoTracker (Acosta et al., 2005), este estágio e parte do seguinte foram desenvolvidos na tese de mestrado. Na segunda etapa, foi feito o planejamento da pesquisa e a extração de dados acústicos, incluindo dados de navegação em três lugares diferentes: i) Na enseada "Estancia", localizada em Palma, Ilha de Maiorca / Espanha, onde usamos o protótipo desenvolvido no primeira etapa deste trabalho como plataforma para a extração de dados; ii) Porto do "Alfeite Arsenal", localizado no rio Tejo, em Almada / Portugal no contexto de exercícios de robótica promovidos pela Marinha de Portugal (REX2014). Nessa localização, utilizou-se o veículo robotizado ZARCO do Laboratório Oceansys, através do trabalho cooperativo com o Oceansys Lab. grupo de pesquisa na Universidade do Porto (UP) / INESC em Portugal e, finalmente, iii) na "Baía de Todos os Santos", perto da cidade de Salvador / Brasil, onde foram realizadas algumas missões para coletar dados usando embarcações hidrográficas em cooperação com a Universidade Federal da Bahia (UFBA) e a Engenharia Belov - Empresa de Engenharia de Portos e Hidrográficos, ambos localizados em Salvador / Bahia. Finalmente, o terceiro estágio, que é o corpo principal desta tese, foi caracterizado pela análise de dados e comparação entre vários conjuntos de dados. Nesta fase, foram realizados estudos para verificar a viabilidade do uso de algoritmos estatísticos espaciais no processo de interpolação de dados batimétricos sem qualquer informação auxiliar para sustentar a predição. Determinamos um procedimento otimizado para estimar os pontos não amostrados, portanto, foi validado usando um método de validação cruzada regular e um método de validação comparativa para comparar os dados estimados com um segundo conjunto de dados adquirido na mesma região e atuando como uma Figura de mérito. A discrepância média entre os dados estimados eo valor de dados da Figura de Mérito foi de 25 cm, está abaixo do erro aceitável para dados batimétricos a profundidades abaixo de 30 m (IHO 2012). Além disso, um algoritmo baseado na teoria da Ressonância Estocástica (SR) foi desenvolvido. Consiste na aplicação de ruído branco em um nível de intensidade ótimo para melhorar os contrastes de imagens acústicas geradas por um lado de Sonar Scan (SSS). A teoria SR também foi utilizada como base para o desenvolvimento de um algoritmo de detecção de sinais fracos em aplicações de detecção. Em relação ao aplicativo de sensores para medir as variáveis físicas de sensoriamento remoto, podemos citar os medidores de campo magnético (magnetômetros) e sensores inerciais (acelerômetros e giroscópios), neste estudo realizou-se uma simulação do circuito de Chua operando em regime caótico como sensor, onde podemos determinar com sucesso a região da solução do sistema no atrativo estranho, usando, para isso, a técnica de tempo de residência, que será definida ao longo desta tese. Todo o conjunto de módulos, técnicas e processos descritos neste trabalho propôs uma solução para o problema de sensoriamento remoto aplicado ao ambiente subaquático e oferece a oportunidade de desenvolver de forma mais profunda os futuros trabalhos em integração de sensores, algoritmos e plataforma de aquisição de dados.

Acknowledgment

There are a lot of people that I would like to thank for the indispensable support to carry out this work.

I would like to thank God and my guardian angels for protection and health. A very special thanks to my family, who along the way of this thesis has grown, beautiful, gentle and sublime. Thank you for your support, understanding, love and at the same time I apologize for the physical absence at many moments, but everyone was always in my heart.

I would like to thank to develop and ending the thesis, Eugenio Garcia Moreno that was one of the best professors that I have in my life; Gerardo Gabriel L. Acosta, my reference in engineer, research project management and research team guidance; Wojciech Korneta, my reference in Physics applied to Engineering problems solutions that teach me to take advantages from stochastic resonance systems in sensing and signal processing and Alejandro Rozenfeld, my reference to Physics and computing systems applied to Spatial Statistical, we have worked hard over simulations and data processing to extract interesting correlations in geostatistics.

I would like to thank a lot the UIB professors Rodrigo Picos Gayá, Jaume Segura, Sebastian Bota and Victor Homar for all help and support given to me during this time.

Inspiration is the motor that moves forward life and world, this will conduct actions that promote development. Inspiration will create objects and solutions in two ways: solid form, that reflect objects that surround us, or abstract form, it is something that we feel as a soul expressions. In my life, I have had many sources of inspiration, each one very important to me, involving a large range into my professional and emotional development. I will not be able to comment the name of all the people, but I will thank directly to the people most involved in all my work and my life.

Em primeiro lugar, eu tive uma inspiração muito especial para aceitar e envolver-me no desafio da realização deste trabalho. Minha inspiração especial é a minha esposa Flora Souza Bacelar que já conheço, admiro e amo à quase metade da minha vida, desculpe pela ausência em muitos momentos e muito obrigado por ser o meu exemplo e minha musa inspiradora.

Te amo minha gata, estamos tocando e bailando juntos o palo da nossa vida.

Minha outra inspiração é a da vida, garra, esforço, gratidão, caridade e carinho. Ainda vejo os olhos de preocupação, de amor, e de compreensão entre as subidas e descidas da vida. Quero ainda pedir perdão pela ausência e por ainda não ter podido retribuir na mesma altura tudo que recebi e recebo. Muito obrigado a estrelinha guia da minha vida, quero dizer, Estelita guia da minha vida, minha Mãe. Muito obrigado Mainha, muito obrigado por me colocar e guiar-me neste nosso mundo, te amo.

Agradeço também a minha mãe-avó, que ao lado da minha mãe foi extremante importante para mim, muito obrigado minha vó Loka, a Luz (ia) da minha vida. Muito obrigado!!!

Tenho muita sorte em ter um irmão, que é um amigo, me apoia sempre nos meus sonhos, projetos e planos, me sinto muito protegido e amado. Além disso, me trouxe uma irmã e um lindo presentinho. Muito obrigado Sandro, Gisa e Mariana, vocês também iluminam a minha vida. O curioso é que eu e meu irmão, em muitos casos, temos pensamentos e opiniões divergentes, mas sempre respeitamos um a opinião do outro, é uma relação muito bonita. Sandro meu irmão, muito obrigado, te amo muito e estou dos eu lado para qualquer coisa.

Depois de conhecer a minha esposa, a minha família se expandiu, ganhei de presente muito carinho, amor e compreensão da sua família, que é a minha família também, ganhei outro Papi, outra Mami e outro irmão. Muito obrigado Seu Ademir, Dona Sandra, Rudá, Debora e Davi. Vamos seguindo juntos em frente.

Tenho que agradecer muito pela oportunidade de conhecer dois lugares muito especiais e que considero como minha casa fora da minha cidade, lugares onde me sinto em casa, respeito a cultura e amo as pessoas e suas coisas do jeito que elas são a minha ilha do coração Mallorca e a minha cidade em Portugal, o Porto. Ambos os lugares representam minha referência de cultura, educação e cidadania. Em ambas as cidades encontrei amigos e mestres que influenciaram e influenciam a minha vida, posso até afirmar que saí do Brasil com uma cabeça e volto ao Brasil com outro entendimento do mundo, das pessoas e suas relações. Em relação à universidade, minha conexão com a Universidad de Las Islas Baleares (UIB) ficou tão sólida que apesar de não haver me graduado nela, se me

perguntam: Em qual universidade estudei? Minha resposta será sempre UIB, porque se tornou a minha referência e a minha casa acadêmica.

M'agradaria donar les gràcies a tots els amics guanyat i es va trobar que eren part de la meva vida a Mallorca: Iacyel, Isabel, Xisca, Javier, Dana, Diana Calvo, Matías Calvo, Sebastian Calvo, Patri, Charles, Ísis, Marina, Leozinho, Natália, Joana, Sigrid, Joab, Jhonna, Antônio, Eva, Matias.

Moltes gràcies a a totes les persones del meu grup de recerca i amics del departament de Fisica de la UIB, He tingut l'oportunitat i el plaer de conèixer, socialitzar i aprendre molt. Por los bons moments aquesta bonica isla, salut a tots: Eugenio Garcia Moreno, Rodrigo Picos, Sebastian Bota, Ivan de Paul, Carol de Benito, Jaume Segura, Eugeni Isern, Vicenzs Canals, Pep Lluis, Miguel Roca, Xavi, Joan, Gabriel Torrens, Salvá, Miguel, Julian, Gemma, Mohamad Moner, Miquel Alomar, Kay Suenaga, Maria, Rúben, Andreu, Aquilina, Borja, Carlo, Diego i Fábio.

Moltes gràcies a todos los funcionarios de la universidad por la paciencia y apoio sempre que los necesité, especialmente mi gustaría agradecer a Francesc Bonet, Rafael Pizà, Melcior Rotger, Antònia Tous y Xisco Litteras. Además de los amigos Lianca, Tomeu, Nuria, Cati, …..

I would like to give special thanks in memoriam to my first friend, director and his family that without him I cannot arrived here in Mallorca, Oscar Calvo Ibañez, we work together for a very brief period of time, but I could see the beautiful person, very dedicated to his family and his work.

Eu gostaria de agradecer a minha segunda casa de formação acadêmica, onde conheci e pude interagir com professores excelentes e inovadores, com uma estrutura acadêmica forte e competente. Além disso, alí conheci pesquisadores, técnicos e estudantes muito competentes e dedicados, alguns que posso chamar orgulhosamente de amigos apesar da distância. Foi e sempre será uma experiência maravilhosa compartilhar o tempo e os conhecimento com vocês, obrigado ao Ocean System Laboratory da Faculdade de Engenharia do Porto (FEUP) pela oportunidade e muito obrigado ao professor Nuno Cruz. Além do professor Nuno Cruz, gostaria de nominar as pessoas mais fantásticas que conheci ali, professor Anibal Matos e Antônio Moreira, Rui Almeida, Bruno Ferreira, José Melo, Nuno Abreu, Miguel Pinto.

Thanks to all Electronic Engineering Group (GEE) and System Electronic Group (GSE) from UIB physics department and All UIB departments and Educations Ministry of Spain, involved in this research trajectory, who proportioned a good background knowledge from all disciplines, research collaborations experiences laboratory structures and administrations support.

Este trabajo de investigación y formación professional fue suportado por el Ministério de Educación de España, dentro del program de Formacón de Personal Investigador (FPI) Program Scholarship. Los fundos que soportaran la beca son vinculados al proyecto IOGECS AUVICOPS.

I would like to give special thanks to Intelymec Research Group from Universidad del Centro de La provincia de Buenos Aires (UNICEN-BA), located in Olavarria, Argentina. From the UNICEN-BA I need to give a special thanks to Sebastian Villar, that help me a lot in many scientific works that we made together.

When I arrived in Bahia in 2011, I found an important support related to data collect and knowledge about some hydrographic surveis principles and given by a specialized brazilian company Engineering and maritime survey services called Belov Engenharia LTDA. I would like to give special thanks to some peoples from this company as Petrônio Achiles, Juracy Costa e Sérgio Correia.

Finally, I would like to give thanks to the research team and company that I have my first contact to field survey and sonar image processing in Scotland. Thanks, Seebyte and their researches Jonathan Evans, David Lane, Scott Reed and Ioseba Tena.

Index

Preface

At beginning of this work the focus are pointed at the development of low-cost underwater remote sensing solutions applying to robotic platforms. The common purpose to this solution was addressed to perform underwater topographic relief surveys and object recognition (natural or man-made) using sonar technology. The research developed and documented in this thesis was mainly supported by the Ministry of Education and Science of Spain into the framework of the IOGECS project (Calvo Ibáñez 2006). Remote sensing in underwater medium has several requirements and constraints, most of these relates to (i) mechanical paradigms to make electronical and electrical devices operating in high pressure medium and in waterproof mode; (ii) communications and navigation, since in underwater medium we cannot apply the common carrier communications based in radio frequency devices as radio-links and GPS navigational system; and finally, (iii) sensing task, we are limited by the use of acoustical devices to sense the topographic relief since the acoustic waves has low attenuations than electromagnetics sensing devices in underwater environments. In conclusion, to address all requirements and constraints, the underwater remote sensing solution, proposed in this thesis, was divided in two parts: hardware development (Underwater Vehicle Platform) and Algorithm development (Stochastic Resonance and Geostatistics applied to Acoustic Data).

The main research question was addressed to reach the convergence of hardware and software into the focus called Underwater Remote Sensing. The research starts with the development of a low-cost underwater, an unmanned vehicle prototype, capable to support acoustic remote sensing devices as a Single Beam Echo Sounder (SBES) and a Side Scan Sonar (SSS). The second stage of work was composed by two tasks, the first was the acoustic data acquisition in a specific site survey to extract real data using the SBES and SSS acoustic devices, it was used a conventional hydrographic boat to carry the devices as a prelaminar step to integrate the sensors to the underwater vehicle prototype. The second task was related to algorithm study, development of data enhancement and data processing based in geostatistics to increase the resolution for point sampling data from SBES. In addition it was used stochastic resonance to enhance the quality of acquired data by SSS, in order to increase the success of target perception into acoustic image data.

During the work to integrate navigational and sonar data one question arouse, how to increase the accuracy of sensors that sense weak signals from some physical variables, for example in our navigational data set, given by inertial sensors one application can be done by magnetometers and thus, this study can be applied to increase the accuracy of inertial sensors and biosensors. It was an addendum to what was initially proposed in this thesis, but it added an interesting research line for my academic formation. Therefore, using the expertise of thesis supervisors, a new technique to using stochastic resonance to detect changes in weak dc signal was studied.

Regarding the real data used in this research that was obtained by hydrographic survey, two distinct types of acoustical sensors were used SBES and SSS as the main research dataset. And the data from the Multi Beam Echo Sounder (MBES) was used as a data reference to estimate the error from the geostatistical regression method applied to SBES device samples. Then, the SBES was used to extract topographic underwater relief data to perform the research study in geostatistical regression techniques, thereby, increase data resolution using Kriging interpolator. The SSS data processing was uncorrelated with the two mentioned sensors. The main objective related to SSS data processing was the acoustical image enhancement. This was due the fact that the SSS has an intrinsic noise embedded in his acoustical data, which may be due from instabilities of remote sensing platform during the survey process or some interferences from electrical or environmental conditions. Improvements in SSS acoustic data before the processing stage to object detection and recognition can increase the rate of successful for automatic and human based processing task.

A set of underwater remote sensing real data (sonar samples) were acquired using both the Low-Cost Underwater Remote Sensing, developed in the first stage of this research work, and a hydrographic boat used as the main source of data to get results reported here. Relating to the topographic relief data (bathymetric data) acquired and processing the results reached will attend the requirements of errors of vertical uncertainty according the standard from IHO-Order 1 specifications (IHO 2008).

Contribution of This Thesis

In this thesis, the work enfolds the research and development of techniques and algorithms that allow the improvement of signal detection tasks, which involves the development of detectors that use the phenomenon of stochastic resonance as a support for the detection of signals of low intensity, or low frequency such as sensors of magnetic field, acoustic sensors and biosensors. Simulation-based studies have been successful demonstrating that

the stochastic resonance process can be used to increase the level of signal detection in sensor applications. On the other hand, the study of stochastic resonance was advancing in the direction of the analysis and processing of signals involving acoustic images acquired with a lateral scanning sonar, which are widely used in submarine inspection systems, remotely detecting morphological and structural characteristics of the submerged relief or even in the inspection of structures to support navigation or mineral extraction used in the oil industry.

Another scientific contribution was the development of a study involving a spatial statistical tool called geostatistics. The objective was to determine if the predictions made with its linear regression technique would leverage a sufficient level of confidence to be accepted within the international standard for hydrographic surveys. Thus, we would confirm the efficiency of this tool in the prediction of bathymetric sample values, which are useful in the elaboration of hydrographic maps to support navigation, underwater infrastructure management and geomorphology. The results that will be demonstrated throughout the work confirm the acceptance of the method in this sense.

The three objects of the scientific work described in this thesis converge to a single application focus area: the development of devices and signal processing techniques geared to the remote sensing area. The theories and techniques of detectors and processing presented can be used in the development of devices for both marine and submarine sensing and remote sensing by satellites, where magnetic and inertial field sensors are used as support for image data acquisition. In turn, require a processing system based on spatial statistics, which confirm or estimate values in certain regions, where the data could not be identified or acquired correctly, corrupted by some effect unrelated to the regular operation of the equipment. In this way, this work serves as a contribution of discussion to the area of remote sensing.

Thesis Layout

This thesis is divided in two parts. The firsts part is related to basic knowledge about mathematical paradigms, tools, devices and procedures used to acquire, develop and processing data from this research. The main purpose of the first part is to introduce mathematical, physical and methodological concepts using equations, expressions and illustrations to support all knowledge needed to interpret and understand the work developed and described in the second part. The second part, described as original work, will revels the core mathematical and physical assumptions, accompanied by data processing and analysis to extract the relevant information’s from the worked signals and data to the remote sensing area.

Related Projects and Colaborations Efforts

The work presented in this thesis was made with support of some cientific projects, and their main reseraches and colaborators. The two projects involved was IOGECS and AUVICOPS, both with support from Spain Ministry of Eduation and their main researches Oscar Calvo, Gerardo Gabriel Acosta and Alejandro Rozenfeld. Besides these project, a big and helpfull colaboration and support, witout this I cannot ending the thesis work was done by Electronic Engineering Group from UIB university and their main researches and colaborators involved Eugenio Garcia Moreno, Wojciech Korneta and Rodrigo Picos Gayà.

Publications Lists

During Doctoral Thesis Development (2014-2016)

Article in Scientific Journals

2015

[1] W. Korneta, E. Garcia-Moreno, A.L. Sena - Noise activated dc signal sensor based on chaotic Chua circuit Communications in Nonlinear Science and Numerical Simulation 24(1-3) · July 2015
Impact Factor: 2.87 · DOI: 10.1016/j.cnsns.2014.12.010

2014

[2] Villar Sebastián A., Sousa André L., Rozenfeld Alejandro y Acosta Gerardo G. (2013) "Evaluation of an efficient approach for target tracking from acoustic imagery for the perception system of an Autonomous Underwater Vehicle," INTERNATIONAL JOURNAL OF ADVANCED ROBOTIC SYSTEMS (INTECH), vol. 11, no. 24, pp. 1-13. ISSN/ISBN: 1729-8806. DOI: http://dx.doi.org/10.5772/56954. (on-line http://cdn.intechopen.com/pdfs-wm/46265.pdf).

Complete Works Publish in Proceedings of Conferences

2016

[1] André L. Sousa, Sebastián A. Villar, Wojciech Korneta, Gerardo G. Acosta and Alejandro Rozenfeld - Resonancia Estocástica para el Mejoramiento del Contraste y Calidad en Imágenes Acústicas de Sonar de Barrido Lateral - IEEE ARGENCON 2016 – Congresso Bienal de IEEE Argentina 2016.

2014

[2] André L. Sena, Eduardo Simas Filho, Pedro Nery - Segmentação e Extração da Informação de Profundidade Local nas Imagens de um Sidescan Sonar Utilizando Redes Neurais de Pulso Acoplado, XI Encontro de Tecnologia em Acústica Submarina. Organizado pelo Instituto de Pesquisas da Marinha/COOPE-UFRJ, 03/11/2014 à 05/11/2014.

Pre Doctoral Phase (2007-2013)

Article in Scientific Journals

2010

[3] "A guidance and control system proposal for autonomous pipeline inspections", A. Rozenfeld, **G. G. Acosta**, A. Sousa, H. Curti and O. Calvo, Transactions on Systems, Signals & Devices, Vol. 5, Nº. 1, pp. 5-27, Shaker-Verlag, Germany, November 2010, ISSN: 1861-5252.

2009

[4] "Low-Cost Autonomous Underwater Vehicle for Underwater Acoustic Inspections", O. Calvo, A. Sousa, J. Bibiloni, G. Acosta, A. Rozenfeld, Journal of Maritime Research (JMR), ISSN 1697-9133, Vol. 6, Nº. 2, pp. 37-52, SEECMAR, 2009. (on-line: http://www.jmr.unican.es/index.php/jmr/article/view/114/111)

Complete Works Publish in Proceedings of Conferences

2013

[3] Villar, Sebastián A., Sousa, Andre, Rozenfeld, Alejandro, Acosta, Gerardo G. (2013) "Pipeline detection system from acoustic images utilizing CA-CFAR," OCEANS 2013 MTS/IEEE San Diego- Estados Unidos. IEEE Marine Technology Society. (On-line: http://ieeexplore.ieee.org/)

2012

[4] Villar, Sebastián A. André Sousa Sena, Alejandro Rozenfeld, and Gerardo G. Acosta (2012) "Sistema automático de detección de tuberías utilizando un novedoso método de promediado de celdas con tasa constante de falsas alarmas acumulado (CA-CFAR acumulado)," VII Jornada Argentinas de Robótica JAR2012. Facultad de Ingeniería-Universidad Nacional del Centro de la Provincia de Buenos Aires.

[5] Villar, Sebastián A., Sousa, Andre, Rozenfeld, Alejandro, Acosta, Gerardo G. (2012) "Sistema de detección de tuberías en imágenes acústicas utilizando CA-CFAR," 23º Congreso Argentino de Control Automático – AADECA'12 Buenos Aires Argentina.

2010

[6] Villar, Sebastián A., Martinez Mariano A., Sousa, Andre, Rozenfeld, Alejandro and Acosta, Gerardo G. (2010) "Detección de tuberías submarinas desde imágenes de sonar de barrido lateral," 22º Congreso Argentino de Control Automático - AADECA '10 Buenos Aires Argentina. Publicado en CD, trabajo AF-001945.

[7] "Sensor module interconnection alternatives for low-cost AUV prototype", Rossi S.R., de la Vega R.J., Cogliatti J.I., Sousa Sena A., G. G. Acosta, publicado en CD, trabajo 6, Proceedings del IEEE/OES South America International Symposium, 12 al 14 de Abril 2010, Buenos Aires, Argentina.

2009

[8] "Smooth path planning for autonomous pipeline inspections", O. Calvo, A. Rozenfeld, A. Sousa, G. G. Acosta Conferencia Invitada al 6th IEEE International Multi-Conference on Systems, Signals and Devices, Túnez, África, 23-26 Marzo 2009, (Accepted to International Journal of Modelling, Identification & Control).

[9] "Inspección Submarina de tuberías mediante Robots Autónomos utilizando sonares", O. Calvo, A. Sousa y G.G. Acosta, Anales de las XXX Jornadas de Automática, 2-4 Septiembre de 2009, Valladolid, España.

2008

[10] "Experimental tests on underwater pipeline inspections with an AUV prototype", O. A. Calvo, A. Sousa, A. F. Rozenfeld, J. Bibiloni, **G. G. Acosta,** H. J. Curti, J. Ghommam, Proc. del MTS/IEEE Oceans'08, Quebec, Canadá, September 15-18, 2008.

[11] "Experimental results on smooth path tracking with application to pipe surveying on inexpensive AUV", O. A. Calvo, A. Rozenfeld, A. Sousa, F. Valenciaga, P.F Puleston, **G. G. Acosta**, Proc. del IEEE/RSJ Int. Conf. on Intelligent Robots and Systems – IROS 2008, Niza, France, September 22-26, 2008.

Part I: General Theory Introduction

1.0 Introduction

A remote sensing problem consists mainly in search and investigate a phenomena or parameter under study through the solution of an inverse problem (Legleiter & Roberts 2009). The inverse problem is the basic process behind sensing systems and promotes the relationship between system input and output, mapping parameters and relates it to measurements from sensing elements. Into the context of remote sensing, parameters are measured without direct physical contact with sensing element, and the process to acquire remote sensing data is called survey or technical mission survey. The term survey in the research scientific area is commonly used to the process of gathering information or acquire a set of data to model any characteristics about a space, system or object. Remote sensing applications is defined as the techniques used to measure something or extract some information about one system or object without direct contact (Schowengerdt 2007). Here in this work, into the context of underwater remote sensing, the term survey will be used to define any process of data collection, in continuous or discrete form to acquire a landscape information using underwater acoustics. It will generate some information to infer physical characteristics capable to predict some topographic relief parameters and trends to help the solution of inverse problem related to underwater mapping for underwater navigation, features extraction and data fusion.

2.0 Dynamic Systems and Sensing

The general inverse problem can be summarized, according many authors, as the problem to identify the Cause-Effect (CE) explicit factor or relationship from a set of data sampling behaviour. For example, Calvettti and Somersalo, in their work, defined the inverse sampling problem as "The problem of retrieving information of unknown quantities by indirect observations" (Calvetti & Somersalo 2007), in another work from Kubo, the definition for the inverse problem was "determination of input or source from output or response" (Kubo 1993).

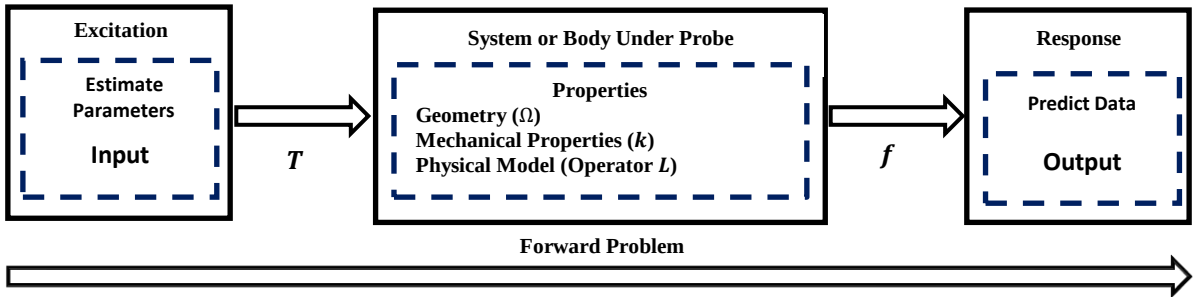


Figure 2.1: *Diagram of direct problem relationship*

Different tools can be applied to model an inverse problem (Tarantola 2005), for the sake of simplicity and coherence, the approach used by Guillermo Rus and Rafael Gallego in this work (Rus & Gallego 2002) and similarly in (Rus & Gallego 2007), was chosen as a base for our definition in this work. The result of a comparison between inverse problem and forward estimation problem is very interesting. As described in following works (Rus & Gallego 2002) and (Rus & Gallego 2007), regarding direct problem, all the mathematical laws governing the studied physical model are known, as well as the values of the parameters and their responses when subjected to define values, resulting in the property of reproducibility. Respect to our problem research, related to discrete data the work (Bertero et al. 1985) has the mathematical bases to apply the inverse modelling in discrete dataset. In a short and direct definition, the forward problem, relates to observe a physical phenomenon, and in order to describe how it works on a system (internally, externally or both), we will discover physical laws, and thus, we can extract some model parameters useful to make predictions/estimations on the results of measurements of observed variable. In other hand, think in a reverse case. Inverse problem solution will address a values or results based in some measures or both, measures and model parameters of observed variable (Tarantola 2005). As an example of forward problem, we depict in the *Figure 2.1* and inverse problem in the *Figure 2.2* as an illustration of the process of the relationship among excitation signal, system under study and response according the forward problem defined by (Rus & Gallego 2002).

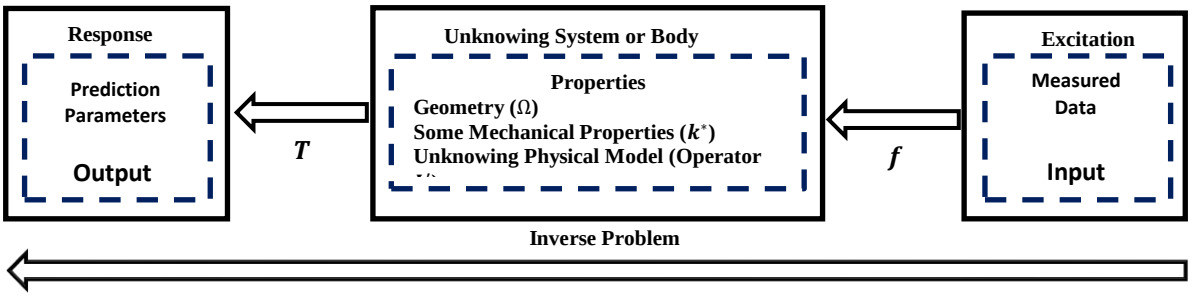


Figure 2.2: *Diagram of inverse problem relationship*

Considering one hypothetical system under study, we need to have a map between objects of interest (parameters), extracted from some acquired information about these objects (data or measurements). Assuming we have a system, response defined by the displacements u and the tractions q of the object under a known boundary conditions (some previous known values about u and q). We can extract the direct mapping expression or forward problem expression for a given force (T), applied to body as a function defined by the hypothetical, *equation 2.1*. That relates the correspondence between the parameter k and the data $f(u, q)$.

In this case, considering typically the Hilbert space Hin the space of data G, and parameter $L \in H$, given a function $f(u,q) \in G$, that relates the displacements u and traction q.

$$f(u,q) = L(k) \cdot T \qquad \textit{eq. 2.1}$$

The correspondence relationship to solve the inverse problem should invert the operator L, and finding parameters values $k \in H$ [1]from a set of knowledge data $(u,q) \in G$, then the inverse problem is expressed by *equation 2.2* that. Observing that L is a linear operator.

$$T = L(k)^{-1} \cdot f(u,q) \qquad \textit{eq. 2.2}$$

In remote sensing applications, a common way to specify an empirical relationship to predict some unknown properties, parameters or features related to a set of measured data taken from a target involves methods to solve an inverse problem. Here, the term "target" is used to define any object, region or surface under study. Environmental and biological information, spatially or temporally distributed, can be predicted as an attribute of some physical measurements over a surveyed objective.

To estimate directly the relationship between measures and physical properties as an attribute of interest in underwater acoustical data, such as bathymetry or acoustic images, some pre-emptive knowledge about physical properties of medium and surface target, is required to apply any inverse problem algorithm. The underwater acoustic data cannot follow a regular pattern, it will depend on geomorphologic topographic relief, distance from target surface and interaction characteristics among signal apply to sense, medium and target, these will change according to the magnitude and bottom type (Lesser & Mobley 2007).

Underwater acoustical remote sensing uses active devices called sonars to sense the interior of any water body (lakes, ocean, rivers, etc.). Sonar devices are applied to extract characteristics or features such as: (i) target distance from source, (ii) target shapes and, in some cases (iii) the intensity of backscatter energy reflected from the target. Some of these features or properties, emerge from the relationship between sensing signal and target surface, however, in our main application, only the first feature, (i), was used to represent the topographic relief sampling technically called bathymetry.

2.1 Sensing Linear and Non-Linear Systems

The use of the technology to interact with an environment, things and persons, gives the science an opportunity to understand complex processes. These interactions are commonly carried out by devices named sensors. Regarding the robotic application, and most specifically, the underwater robotics application and sonar mapping, we need one or more sensors to get the information about the location and operation of the robot. Any sensor must be capable to receive an excitation signal as a stimulus or input signal and translate it into electrical signal called response or output signal (Fraden 2010), this is shown in *Figure 2.3*.

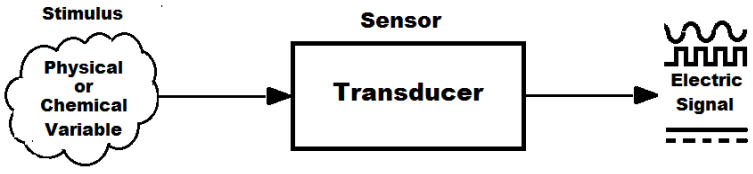


***Figure 2.3**:Block Scheme of the Information Transfer*

The sensor is often used as a part of a control system as it is shown in *Figure 2.4*. There are thus two applications of sensors: passive and active. Their measurement of physical variable can be absolute, i.e. independent of the measurement conditions, or relative, i.e. related to a referential system.

There is a set of constraints which should be considered in sensor development and in designing signal processing systems. Some of these constraints are related to a sensitivity and the excitation input range known as a "span". The most important in our study was the accuracy, the stability, the decrease of hysteresis effects and the linearity

[1] Hilbert Space generalizes the Euclidian Space extending the vector algebra and calculus to a finite or infinite number of dimensions.

of the response. Relates to automatic control system, a simple scheme that involves a sensor as an element to measure the efficiency and accuracy of control acting over a generic system is show in *Figure 2.4*

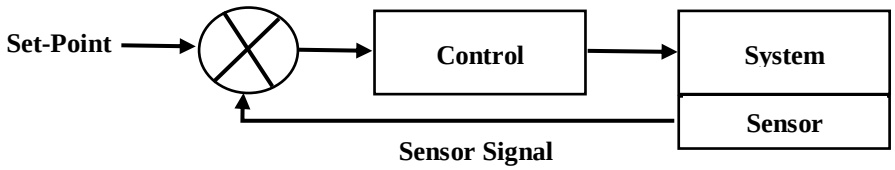


Figure 2.4: *Scheme of the Control System Using Sensor*

Considering an automatic control system, a sensor device uses sensing elements to read some physical or chemical variable to generate inputs and send this information and help to solve some control problem, mapping or simply store or share a log of data. The underwater robotics and mapping surveys are basically sensor-based systems in a mission survey. The robotic vehicle uses data information from sensors to improve navigation and log data from the environment. In a task of autonomous underwater systems, one should consider energy restrictions, reliability, immunity to noise and exposure to extreme environmental conditions. Furthermore, deep in the water, the online monitoring system is excluded. Besides, it is not possible to verify the system state, task execution and system diagnostics by remote monitoring system communication due to restrictions in sending and receiving high data transmission through the water by radio frequency spectrum. In some cases, one can use an acoustics communication system, but the slow data rate and energy consumption requirements limits its use in autonomous systems because of power consumption constraints.

To abstract the concept of sensor applied to measure a system behaviour we introduce our model system related to environmental survey, specifically to underwater topographic landscape, and in this abstraction, our sensor system consists in acoustic transducers called SOund Navigation And Ranging (SONAR) and our system is the underwater landscape and sometimes any structure lying on its. The effort of investigation developed in this work was made to solve problems and paradigms involving data processing from two basic sonars, Single Beam Echo Sounder (SBES) and Side Scan Sonar (SSS).

Regarding to the effort of sensor works, in this thesis we consider a sensor simulation model based on electronic Chua's circuit and the stochastic resonance phenomenon which can be used, e.g., acoustic transducer response, inertial magnetometer. Magnetometer sensors are used massively to improve an autonomous navigation in both indoor and outdoor environments. In underwater robotics magnetometers give a control system the attitude and heading information without GPS signal, due to its sensing element is based on the measure of the magnetic earth's field.

The use of dynamic systems in the description of natural phenomena involves mathematical modelling by differential equations and computer simulation. Dynamical systems are represented by a set of variables interconnected by causal relationships. In a dynamical system, some characteristics of the interacting elements change over time. Continuous dynamical systems can be modelled by the following set of ordinary differential equations (L. Gammaitoni et al. 1998), as represented in the *equation 2.3*.

$$\dot{x}_i(t) = F_i[x_1(t), x_2(t), \dots, x_n(t)] \qquad \textit{eq. 2.3}$$

For $i = 1, \dots, n$. By $x_1(t), x_2(t), \dots, x_n(t)$ are denoted system state variables and $\dot{x}_i(t)$ is the time derivative. $F_i(\dots)$ is a function which describes the variation rate of state variables. The system time evolution in n-dimensional phase space determines the trajectory called orbit. For discrete dynamical systems, the time changes discretely. One thus has a default time increment or spacing Δt. The time evolution of each system state variable is given by the map in the following *equation 2*.4.

$$X_{i+1} = F[X_i, X_{i-1}, X_{i-2}, \dots, X_{i-k}] \qquad \text{eq. 2.4}$$

Where n-dimensional vector X_i represents the system state variables in the *i*th time step and *F* is a function describing evolution of the system. The system state variable is thus a function of past values of state variables. In many cases, it is convenient to use a discrete set of times and describe continuous system evolution as discrete jumps from one configuration to another. Another paradigm to model a system behaviour is based in a linear dynamical system, and it is completely defined if the system under study has two properties: superposition and homogeneity.

The superposition property: any two equations or solutions can be combined by simple addition to generate a third solution or equation. The state obtained by the superposition property for two different inputs x and y, in the context of the function F(...) will describe a dynamical system given by the *equation 2.5*.

$$F(x+y) = F(x) + F(y) \qquad \textit{eq. 2.5}$$

The homogeneity property: for any real number k and for the function $F(\dots)$, we can describe a dynamical system according the *equation 2.6*.

$$F(kx) = kF(x) \qquad \textit{eq. 2.6}$$

Any function that does not satisfy superposition and homogeneity is nonlinear. The linear approximation of equations describing dynamical system can be used as good approximations to nonlinear ones in situations where the trajectory stays very close to a stable fixed point or limit cycle. Nonlinear system equations describe its time evolution including at least one term that contains the square, power different from one, a product of system variables, one term in a logarithm form, i.e., any nonlinear function of the state variable. Thus, the addition of two solutions does not yield a valid new solution. In all real systems deviations of large enough amplitude require nonlinear terms in the relevant model. The consequences of nonlinearity are profound. Most importantly, nonlinear systems may contain multiple attractors, each with its own basin of attraction. The attractor can be a single point, limit cycle, a finite set of points, a curve, a manifold, or even a complicated set with fractal structure known as a strange attractor. Let F be a map from the n-dimensional space to itself. We say that a compact set $A \in \mathbb{R}^n$ is an attractor for the map F if it satisfies the following properties:

- $F(A) = A$, i.e. A is invariant;
- A contains an initial point whose trajectory travels throughout A.

The portion of a trajectory corresponding to the progress toward the attractor is called a transient. The set of points in state space that lie on transients associated with a attractor is called the basin of attraction of the attractor. In a stable linear system, all points in state space lie in the same basin of attraction. Bifurcation is an abrupt change in the properties of a system when system parameters change. There are different types of bifurcations (Gammaitoni et al. 1998), (Benzi et al. 1982) and (Kapitaniak & Bishop 1999). We shortly describe some of them, such as the saddle-node bifurcation, pitchfork bifurcation and Hopf bifurcation. Bifurcations are usually represented by the bifurcation diagram where the system state variable and control parameter are represented. The stable and unstable solutions are depicted by continuous and dotted lines respectively. In the phase space diagrams stable fixed points are denoted by black circles and unstable by white circles.

The saddle-node bifurcation is a bifurcation in which two fixed points one stable (node) and one unstable (saddle) coalesce into a single point and annihilate each other (Gammaitoni et al. 1998) and (Bacelar 2010). In the *Figure 2.5*A we show an example of phase space diagram of 1D system with two fixed points: one stable (left) and one unstable (right). In *Figure 2.5* B we show the corresponding bifurcation diagram with two solutions x^* for this system, one stable $x^* < 0$ for $r < 0$ and the other unstable $x^* > 0$ for $r < 0$, where r is a system control parameter.

Pitchfork bifurcation occurs in systems with inversion or reflection symmetry (Bacelar 2010). This bifurcation has two types: supercritical (appear two stable equilibria) or subcritical (appear two unstable equilibria). In *Figure 2.6* A, we show supercritical pitchfork bifurcation with one stable fixed point $x^* = 0$. In *Figure 2.6* B we show the bifurcation diagram with one stable solution for $x^* = 0$ for $r < 0$, one unstable solution for $x^* = 0$ with parameter $r > 0$, and finally two stable solutions for $x \neq 0$, when $r > 0$, considering in all cases, r is a system parameter.

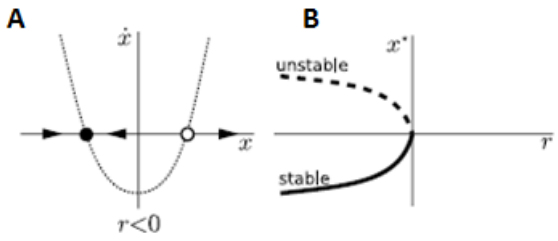


Figure 2.5: *Saddle node bifurcation: the phase space and the bifurcation diagrams* (Gammaitoni et al. 1998)

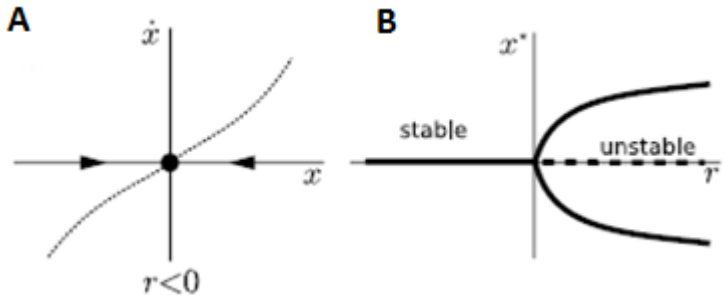


Figure 2.6: *Phase space and the bifurcation diagrams of supercritical Pitchfork Bifurcation*(Gammaitoni et al. 1998)

The Hopf bifurcation is a bifurcation in which a limit cycle arises from an equilibrium in dynamical system when the equilibrium changes stability via a pair of purely imaginary eigenvalues. In the phase portrait of nonlinear systems limit cycle is an isolated closed trajectory. An isolated trajectory means the absence of other closed trajectories infinitely close. Therefore, the neighbouring trajectories must approach or move away from the limit cycle which is a periodic attractor or repeller.

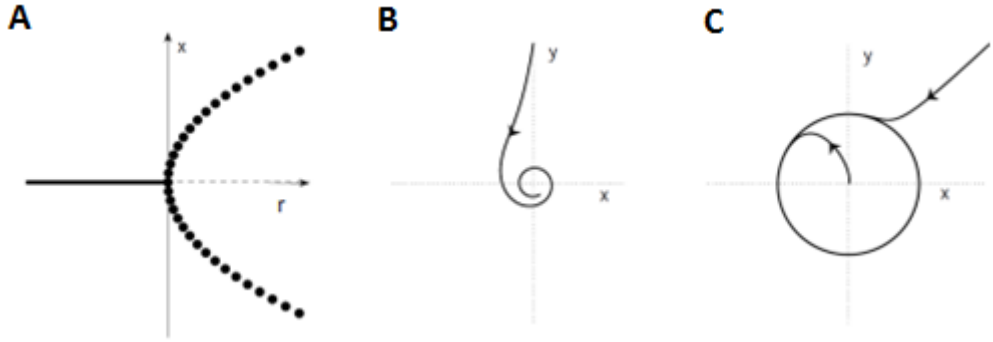


Figure 2.7*:Bifurcation and the phase space diagrams of supercritical Hopf Bifurcation*(Bacelar 2010)

Like Pitchfork bifurcation, the Hopf bifurcation has two types: supercritical (stable limit cycle) or subcritical (unstable limit cycle). To obtain this sort of bifurcation minimally a two-dimensional system is necessary (Bacelar 2010). In *Figure 2.7*A example of the bifurcation diagram for the supercritical Hopf bifurcation is presented, whereas in *Figure 2.7* B and *Figure 2.7* C examples of phase space diagrams are shown before and after bifurcation, r is a system parameter, whereas x and y are system state variables.

In a short definition, a bifurcation can be understood as change in the equilibrium, in some special systems the solution will be periodic orbits, and changing in this orbit will represent a bifurcation point too. Then bifurcations mean a change in system stability properties, according a parameters variation. This can be visualized as changes in the topology of the trajectories in the phase space.

2.2 Stochastic Resonance and Noise

The theory of stochastic resonance (SR) comes from the study of non-linear dynamical systems and the optimal level of noise to enhance weak signal. The stochastic resonance occurs in a bistable threshold-based system driven by a periodic external force in the presence of noise. The phenomenon involves a rapid increase in the signal-to-noise ratio with an optimal noise variance value (Andò & Graziani 2000). We present below the basic concepts of the stochastic resonance phenomenon and introduce its characteristics.

In the (Gammaitoni et al. 1998) the phenomenon of stochastic resonance is explained analysing the motion of heavily damped particle with mass m in symmetric double well potential $V(x)$ represented in the *Figure 2.8*. We can identify three fixed points: two stable x_a, and x_c and one unstable x_b. There are thus two regions of stable state with the minimum energy and one unstable region with the maximum energy. If a particle receives enough energy, it can jump between the two potential minima. Applied external periodic excitation source causes

transitions between potential wells with a rate given by the Kramers formula (Gammaitoni et al. 1998) (Brinkman 1956)(Hanggi et al. 1990), the *equation 2.7* express this relationship:

$$r_k = \frac{w_0 w_b}{2\pi\gamma} \exp(-\frac{\Delta V}{D}) \qquad \textit{eq. 2.7}$$

Where, γ denotes viscous friction, $w_0^2 = V'' \frac{(x_m)}{m}$ is the squared angular frequency of the potential in the potential minima x_m located at x_a and x_c ; $w_b^2 = \left|V'' \frac{(x_b)}{m}\right|$ is the squared angular frequency at the top of the potential barrier located at x_b and ΔV is the height of the potential barrier separating the two minima and D is the noise intensity.

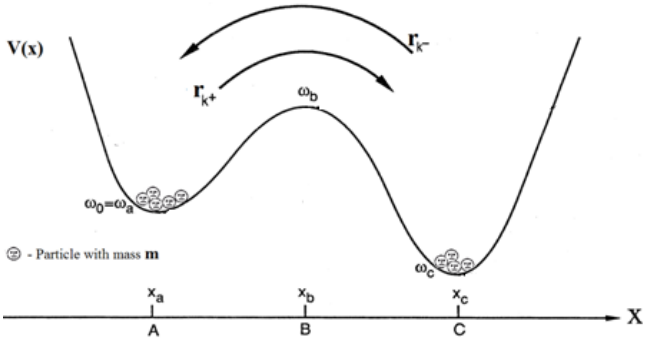


Figure 2.8: *Double well potential with mass* (Gammaitoni et al. 1998), (Hanggi et al. 1990)

The noise intensity is related to the temperature (T) by the formula $D = K_B T$, where K_B is the Boltzmann constant. External periodic excitation tilts double-well potential asymmetrically up and down, as it is shown in *Figure 2.9* (subplots 2 and 4), periodically raising and lowering the potential barrier. Although the periodic forcing is too weak to let the particle roll periodically from one potential well into the other, the noise induces hops between potential wells.

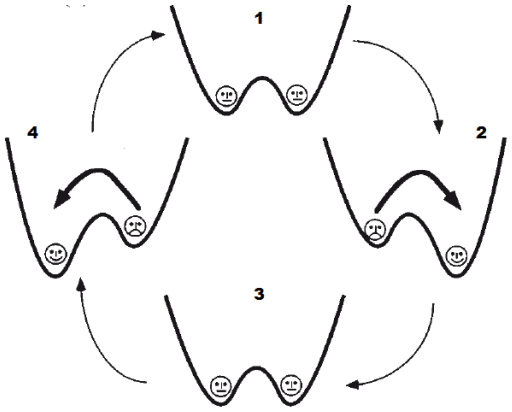


Figure 2.9: *Double-Well potential with external excitation* (Gammaitoni et al. 1998)

Statistical synchronization of particle jumps with external periodic forcing takes place when the average waiting time $T_k(D) = 1/r_k$ between two noise-induced inter-well transitions is comparable with half of the period of periodic forcing. This yields the time-scale matching condition for stochastic resonance.

The noise is a random signal that cannot be described by a specific function ahead of its occurrence (Andò & Graziani 2000). Engineers and scientists for many years tried to avoid any type of noise in a signal under study. In recent years, it was shown that noises can promote an enhanced desired information signal and often, need to be added to make a most realist model of physical system behaviour. Today a relevant number of natural phenomena can be explained by the presence of some type of noise (Andò & Graziani 2000). In electronics, noises are well understood and their study are fundamental to reduce undesirable effect over electronic components and circuits that can degrade the desirable level of signal. For example: a resistor has a noise of thermal origin and if

it is in an audio amplifier circuit the output signal can be degraded. The thermal noises are close to the white noise, because their spectral density is a flat spectrum. The white noise is highly random signal and its spectral statistics demonstrates that its intensity is the same over all frequencies.

There are some types of the noise that differ basically by the frequency distribution or random fashion. For example, biological or geological phenomena produce low frequency noise and this is known as flicker or pink noise. This is a type of called coloured noise which is different from a white noise in its spectral distribution.

2.3 Chaotic Systems

Chaos theory is basically applied to the study of dynamical behaviour of deterministic systems modelled by a set of differential equations. The long-term prediction of dynamical behaviour of deterministic systems highly sensitive to initial conditions is impossible (Ghys 2012) and (Lorenz 1963). Their dynamical behaviour can be predictable for a while and then 'appear' to become random. Systems or signal that present this behaviour are called chaotic. Chaos is characterized by stretching and folding mechanism, nearby trajectories of a dynamical system are pulled apart exponentially and folded back together, repeatedly.

A chaotic signal has a broad Fourier spectrum in intermediate form between a periodic signal and a stochastic signal (noise) (Silva 2006). Primarily in systems with chaotic dynamics, scientists only identified chaotic regions to avoid them. Later, more complete studies of chaos control and chaos synchronization initiated chaos theory applications in biology, physics, engineering and in computer sciences (Silva 2006). Nowadays scientists, engineers and technicians are faced with a non- linear paradigm. They have a large amount of data with nonlinear behaviour to analyse and understand.

Chaotic dynamical behaviour can be observed and studied in simple dynamical systems. One of the simplest and most widely studied real nonlinear dynamical system is called Chua's Circuit (Silva 2006), (Kennedy 1993b), (Kennedy 1993a), and (Boccaletti et al. 2000).

2.4 Chua's Circuit

Chua's electronic circuit is shown in *Figure 2.10*. It consists of linear inductor L, a linear resistor R, two linear capacitors C1 and C2, and a single voltage controlled nonlinear resistor N_R called a Chua diode (Silva 2006) and (Kennedy 1993b).

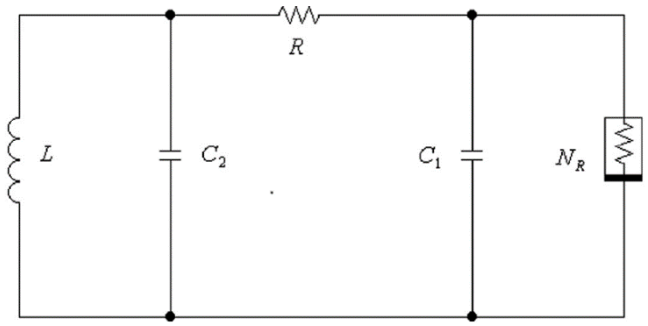


***Figure 2.10**: Chua's Electronic Circuit*

Differential equations describing Chua's circuit are the following in the *system equation 2.8*.

$$\begin{cases} \dfrac{di}{dt} = -\dfrac{1}{L}V_2 \\ \dfrac{dV_2}{dt} = \dfrac{1}{C_2}iL - \dfrac{1}{RC_2}(V_2 - V_1) \\ \dfrac{dV_1}{dt} = \dfrac{1}{RC_1}(V_2 - V_1) - \dfrac{1}{C_1}\ g(V_1) \end{cases} \qquad \textit{eq. 2.8}$$

where $\mathbf{V_1}$ and $\mathbf{V_2}$ are voltages across capacitors C1, C2, i is the current of the inductor L and $g(V_R)$ is a piecewise-linear function shown in *Figure 2.11* representing the change in resistance vs. current across the Chua Diode.

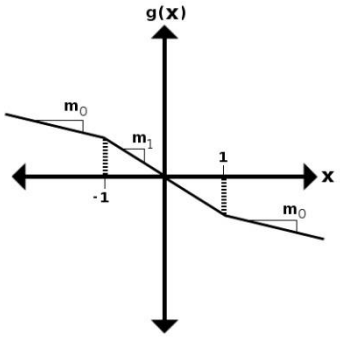


Figure 2.11:Generic response curve of Chua diode: $\boldsymbol{m_1}$ is the slope of the middle segment and $\boldsymbol{m_0}$ is the slope of the two outer segments.

Chua's circuit allows one to change parameters and observe many different phenomena related to chaotic dynamical behaviour (Silva 2006). It exhibits a large variety of bifurcations and different approaches to chaotic dynamical behaviour (Silva 2006), (Kennedy 1993b) and (Kennedy 1993a). This circuit is easy construct at low cost. Some applications involving Chua's circuits provide a controlling of chaos and selection of desired periodic orbit from orbits embedded in the attractor. This control applies small perturbations like an optimal noise to stabilize selected orbit. The Chua's circuit as a nonlinear dynamic system with chaotic behaviour is sensitive to a noise and to initial conditions. Small perturbation can give rise to a very large response in the course of time. The optimal choice of a perturbation can direct the trajectory to wherever one wants in the attractor and to produce a series of desired dynamical states. This is exactly the idea of targeting (Kennedy 1993b), (Kennedy 1993a), (Boccaletti et al. 2000) and (Grebogi & Lai 1997).

2.4.1 Chua's Diode

Chua's diode is the main component to Chua's electronic circuit. It is a type of nonlinear active resistor built using resistors, capacitors and operational amplifiers. In new designs of Chua circuit, Chua's diode is replaced by a component called memristor. This was proposed by Chua in order to describe a new non-linear passive two-terminal electrical component relating flux linkage $\varphi(t)$ and the amount of electric charge that has flowed $q(t)$ (Matsumoto 1984), (Chua 1992) and (Kennedy 1992). The memristor is characterized by its memristance function $M(q)$ describing the charge-dependent rate of change of flux with charge and defined as *equation 2.9*.

$$M(q(t)) = \frac{d\varphi}{dq} = \frac{d\varphi/dt}{dq/dt} = \frac{V(t)}{I(t)} \qquad \textit{eq. 2.9}$$

In the equation 9, we can see that the memristance is thus charge-dependent resistance. The memristor that replacing Chua's diode is a flux controlled memristor that is characterized by its memductance $W(\varphi)$ and modelled in the *equation 2.10*.

$$W(\varphi(t)) = \frac{dq}{d\varphi} = \frac{dq/dt}{d\varphi/dt} = \frac{I(t)}{V(t)} \qquad \textit{eq. 2.10}$$

The relation $I(t) = g(V_1)$ will be given by *equation 2.11*.

$$I(t) = W\big(\varphi(t)\big)V_1(t) = \frac{dq}{d\varphi}V_1(t) \qquad \textit{eq. 2.11}$$

In which nonlinear dependence of q on φ is introduced. The first Nano-scale TiO2 memristor was constructed in May 2008 and the result was published in Nature Letters (Strukov et al. 2008). The memristor is currently under development by various research groups including in the University of Balearic Island, that works in their modelling (Picos et al. 2015), (Garcia-Moreno et al. 2015) and (Tetzlaff et al. 2016).

2.4.2 Residence Time

Considering a system described in the subsection 2.2, and depicted in the Figure 8, this model is characterized by a bistable potential energy function, in which two stable states, are separated by an unstable point represented by the maximum of the potential. Considering a sensor application using the model of double well potential from Figure 8, to move particle between the two potential minima, one external periodic excitation signal is required. The system subjected to a periodic deterministic excitation $A(t)$ and some random force $\xi(t)$, promotes the particle jumps between minima according Kramers rate, modelled according equation 7 (L. Gammaitoni et al. 1998) and (Brinkman 1956).

Hence, considering the deterministic signal, $A(t)$, the target signal to be sensed, and it is monitored through the particle dynamics or the particle state between the two minima, this will be done using a statistical approach based on monitoring the difference of the mean residence time. The Residence Time (RT) is the time that particle spends in each of the two minima stable states (Dari et al. 2010).

3.0 Underwater Acoustics: Physical Interactions and Data Extraction

Despite the underwater acoustic waves studies comes from the 15th century with Leonardo da Vinci experiments (Urick 1983), studies on acoustic propagation remained without developing their mathematical basis to model their behaviour until the 17th century. This started with Newton's study to model a sound wave propagation, published in "Mathematical Principles of Natural Philosophy", and posterior application to underwater sound propagation that was carried out by Charles Sturm and Jean-Daniel Colladon on the 19th century, in order to measure the velocity of the sound in water (Urick1983) and (Plumian1837). Relevant researches on the underwater sound propagation applied to sonar devices were only perform after 19th century and specifically after the sinking of the Titanic in 1912, to civil navigation security and after the Second World War to military application (Urick 1983).and (Medwin & Clay 1998). Studies to Submarine Warfare applications was conducted by Ewing and Worzel in 1947 when they discovered the deep sound channel (Urick 1983), (Medwin & Clay 1998), (Blondel 2001) and (Lurton 2010).

The device developed to use sound waves energy as well as the transport vehicle to acquire underwater data is a sonar. The first active sonar was patented after the sinking of the Titanic in 1912. Today there are some types of sonars to cover a diversity of application such as detection of objects, navigation and mapping (Brouwer 2008) and (Bartberger 1966). Nowadays, ceramic transducers are the main component of the sonar system; they are made mainly by Piezoelectric sensor element.

Piezoelectric ultrasonic transducer is the most common underwater acoustic sensing component used in an underwater sonar system. The sensor response is produced using the properties of piezoelectricity (Rongxing Li 1992), (Ting 1992) and (Ensminger & Bond 2011), despite the existence of many technologies to acoustic transducer like Lasers, electromagnetic and mechanical coupling (Ensminger & Bond 2011). Some researches to develop other type of sensors elements as MEMS underwater acoustic sensors used a silicon wafer or Silicon Insulator (SOI) as the base material (Li et al. 2007). A piezoelectric transducer can work as a simple ultrasonic or acoustic sensor, which acquire a charge and generate an electrical pulse when compressed by a mechanical excitation or perform a mechanical excitation when submitted to an electrical charge variation (Li, Zhiqun Daniel Deng, et al. 2012) and (Ensminger & Bond 2011).

The field of ultrasonic transducers application increase in classical and new areas of their use due to discovers and development of more sensitive and compact piezoelectric sensors (Li, Zhiqun Daniel Deng, et al. 2012), (Ting 1992), and (Ensminger & Bond 2011). Their application can be found in different areas such as industry, metrology, chemistry, biology and material science. One of the fundamental difference between acoustic and electromagnetic energy propagation is the necessity of mass particle to perform a mechanical wave propagation in the medium in the case of acoustics waves. Electromagnetic waves don't need mass particles to propagate, can propagate in the vacuum medium (Ensminger & Bond 2011).

Actually, a piezoelectric acoustic transducer is a core of the underwater sonar system, it can work in a passive or active mode to acquire an acoustic information about the environment in which is immersed. Developments in Lead Zirconate Titanate (PZT) (Li, Z. Daniel Deng, et al. 2012) and (Martins et al. 2012) leads this composite to be the most used piezoelectric ceramics for sonar underwater applications. In our work, we use an active underwater sonar system. On the other hand, in a passive mode, the underwater sonar system is common called hydrophone, and sense an acoustic mechanical wave energy propagated thrown the medium from other sources of sound like animals, boats, geological activities, noise background active sonars.

An active sonar system generates an acoustic pulse (mechanical energy) and it is transmitting a mechanical energy in an acoustic beam spread to a mass of particles. The mass of particle refers to a volume element that contains so much elements, that it may be consider as continuum fluid. Despite this, yet small enough so that the acoustics properties of pressure, density and velocity can be considered as constant throughout the volume element of the considered fluid medium. The acoustic energy travels thrown the medium by energy transfer from particle to particle. This is exemplified in the *Figure 3.1*A, which represents a portion of continuum fluid that have not disturbance or pressure applied, the distribution of particles mass is uniform. If we consider a disturbance signal, like a sinusoidal acoustic pressure illustrating in the middle of the *Figure 3.1* applied into the medium, this will propagate in it, and regions of high pressure (compressed particles) and low pressures (uncompressed particles) will appear, *Figure 3.1* B.

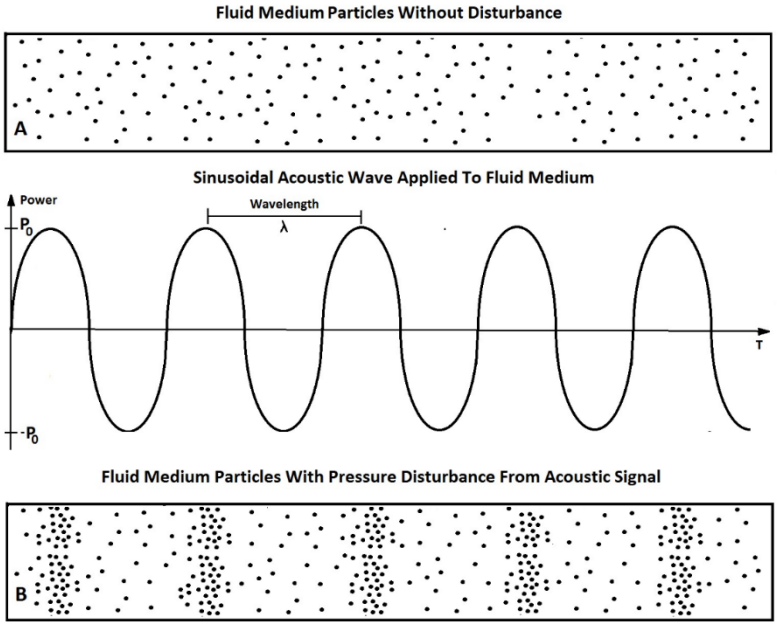


Figure 3.1:*Underwater medium distortions over acoustic signal disturbance*

When the wave find an obstacle (target), it scatters a portion of acoustic emitted energy back to the sonar transducer (Ensminger & Bond 2011). The acoustic backscatter signal has attenuations as a result of beam spreading, absorption, and other time dependent effects inherent to the acoustical propagating wave into the fluid medium (Lurton 2000), (Blondel 2001), (Blondel 2009), (Clifford 1980) and (Urick 1983).

There are some types of sonars devices and their uses depend on the surveys constraints, namely, needed information, effective cost and time response. The pulse beam spreading, wavelength and duration are very important to survey resolution from a sonar system data acquisition, it will directly influence in the object detect and mapping application.

Applications and uses of a sonar system depend on the physical construction and sensor arrangement, including their electronic systems. On the simplest way a single active sonar transducer works emitting an acoustic pulse in the medium (firing), normally to one specific direction. The sonar system waits to receive a backscatter portion of the emitted pulse from a target reflection, during a time ΔT, according the maximum range of the pulse propagation setting. It can be used to measure a single distance between a sonar system to a target object simply using the wave velocity propagation (c) in the medium and the wave flying time (t) using the *equation 3.1*.

$$dist_{from\ sonar} = c.t \qquad \textit{eq. 3.1}$$

Some sonar application needs to take the reflectivity information from a specific target or surface, not the distance. Despite this, it can be inferred with indirect methods, using the wave spreading and their relationship with geometric characteristics of the data extracted to form an acoustic image. In this case the sonar system converts a sequence of received acoustic pulse, based on successive sampling at each discrete space time (Δt_i). According the strength of the echo signal backscatter to sonar, a specific level of gray-scale is determined to this point, or element sample of the backscatter information vector, along the entire propagation time (ΔT). The entire backscatter vector represents a 1-D acoustic information of the surveyed area or target and can be called an acoustic image line. A set of acoustic backscatter vector or 1-D information also can be extract if we have a single sensor scanning one region or a row of sensors passing throw one region. When the sonar sensor acquiring a set of vectors and maintaining these together, side by side, it can represent a covered scanning region, and then perform a 2-D or a 3-D representation.

In conclusion, the quantity of information acquired, surface representation or dimension, 1-D to 3-D, will depend on sensors arrangement to capture a single discrete information or the spatial distribution of the acoustic backscatter when scanning a region and finally this information is processed.

3.1 Interactions with Environment and Losses

The acoustic waves transmitted by the sonar transducers interact with the underwater medium, fluid, bottom surface and subsurface bottom, depending on some factors like sonar frequency, the composition of underwater floor and their physical arrangement. All the previous exposure, refers to the relation about medium impedance and their relationship. The impedance for each medium is dependent on wave velocity propagation (v) and medium density (ρ). At each medium interface the amount of energy reflected is proportional to the acoustic impedance contrast at each frontier interface. Thus, in the propagation trough fluid medium, the acoustic energy can interact in different ways. Acoustic waves can backscatter to transducer, reflect, be absorbed and refract. These can be seen in a simple representation in the *Figure 3.2*.

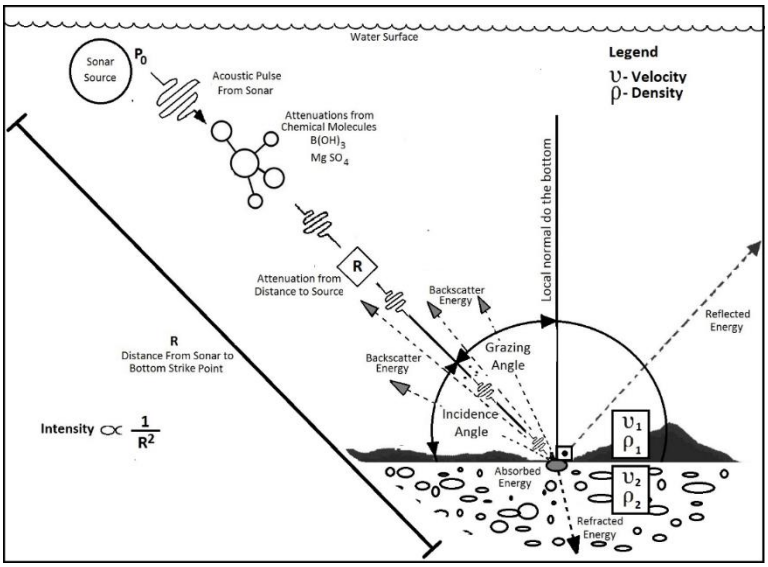


Figure 3.2: *Acoustic pulse interaction with seabed*

In the geometry of sidescan data acquisition, we consider some sensor inner characteristics, settings and environmental correlation. Regarding image formation geometry, and data acquire, we need to distinguish some relationships about physical and geometrical approach. The propagation of acoustical pulse in the medium will imply in some attenuations and distortions (Lurton2002), (Urick1983) and (Klein1985), and it directly depend on the type and characteristics of fluid (Chemical components, density and temperature) and the maximum range distance spread.

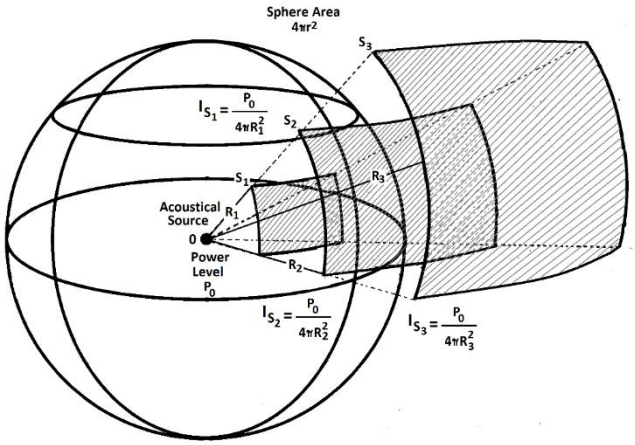


Figure 3.3: *Spherical spreading of an acoustical pulse over consecutives surface layers and the energy intensity distribution over an element surface of each wave front.*

The wave amplitude and polarity reflection is proportional to the acoustic impedance. The acoustic pulse propagates away from transducer, in a spherical spreading, the energy will spread over one increased surface and,

consequently, the emitted initial energy intensity will be distributed over this surface. The result of the energy distribution over a continuous increased area is the intensity loss by unit of area covered, thus, some notion of energy spreading modelling and attenuations are fundamental to understand the acoustical image physical generation.

The acoustic energy in the wave front per unit of area or intensity, decrease in a square of the distance. In quantitative terms, the acoustic energy will loss of 6 dB for each doubling of the distance travelled to 1/4 (25 %) of the sound intensity initial value. (Klein1985) and (Urick1983). This relation is summarized by *equation 3.2*.

$$\text{Intensity} \propto \frac{1}{R^2} \qquad \textit{eq. 3.2}$$

The *Figure 3.3* shows the spherical wave spreading and the intensity propagation in consecutives spherical layers of wave front. The energy intensity, according the distance from power source can be calculated if we have the source power level:

$$I_{Si} = \frac{P_0}{4\pi R_i^2} \qquad \textit{eq. 3.3}$$

Where:

P_0 - is the power leaves the transducer;
R_i - is the distance from sonar.

If we consider a homogenous medium, in a far distance from transducer, the wave front tends to be planar. In this work, all considerations about energy propagation and interaction with the medium will consider this approximation (Urick1983) and (Randall2010). For the case of plane wave propagation, the concept of sound or acoustic pressure P (amplitude) is related to the particle velocity v as shown in the *equation 3.4* and represents a called Ohm's law for acoustics, where acoustic pressure is analogue to voltage, the particle velocity is analogue to the current, and the specific acoustic resistance is considered as ρc (impedance), this is analogue to the electrical resistance. Sound pressure decrease inversely proportional to the distance from the sound source and the acoustic sound pressure level decreases by $-6\,dB$ for doubling of the distance from the source to $\frac{1}{2}$ (50%) of the sound pressure initial value (Lurton2002), (Urick1983) and (Klein1985).

$$P = \rho v c \qquad \textit{eq. 3.4}$$

Where: P - is the acoustic pressure;
ρ - is the density;
v - is the particle velocity;
c - is the velocity of plane wave propagating.

For seawater, P has the value $1.5 \times 10^5\ g/cm\,s^2$ and for air is $42\ g/cm\,s^2$. In theoretical investigations, the sound pressures are often expressed in $Newton/m^2$ and sound intensities in $Watt/m^2$. However, historically because of the large range of audible intensities, covering several orders of magnitude, the common way to describe these quantities is using a logarithmic scale. The intensity level unit is *bel*, but in practice it is convenient to use the *decibel* = 0.1 *bel*. The *decibel* scale is used to compress a wide range of absolute values into a manageable range, this is a relative measure, and thus we need a reference value to infer the relationship to the quantity measured value. To define the decibel measure we use the *equation 3.5*.

$$dB = 10 log_{10} \frac{Quantity\ measured}{Reference\ value} \qquad \textit{eq. 3.5}$$

Then the intensity level (IL) of sound intensity I in decibel is:

$$IL = 10 \log \frac{I}{I_0} \qquad \textit{eq. 3.6}$$

Where the intensity reference, I_0, or threshold level and the logarithmic scaling makes the values manageable from 0 to $140dB$, if the reference value I_0 ranges from $10^{-12}\,W/m^2$ (threshold of audibility) to $100\ W/m^2$. The

Standard Reference Sound Intensity $I_0 = 1\,pW/m^2 = 10^{-12}\,W/m^2 = 0\,dB$ (Lurton2002, Urick1983 and Klein1985). For the pressure level, the similar relationship can be expressed in relation to a pressure reference (P_0) using the *equation 3.7*.

$$PL = 10\log\frac{P}{P_0} \qquad \textit{eq. 3.7}$$

The reference pressure level in underwater acoustics is one micropascal ($1\mu Pa$) and it corresponds to an intensity of $0.65\ x\ 10^{-18}$ W/m^2. Differences between air and water reference pressure level differ by approximately 62 dB. (Urick1983, Klein1985 and IAGC_OGP2008). The source level (SL), can be defined as the relationship between the intensity of the energy source at standard range (S_0) and the reference intensity (I_0). This can be seen in the *equation 3.8*.

$$SL = 10\log\frac{S_0}{I_0} \qquad \textit{eq. 3.8}$$

The medium impedance depends from both, fluid density and speed wave propagation, also known as the specific acoustic resistance. The medium impedance is the opposition that the medium offers to a longitudinal wave motion and this property is analogous to resistance or impedance in electrical circuit theory. The unit for acoustic impedance is rayl, equal to $p.s/m$ (Pascal second per metre) or $kg/m^2\,s$ (kilogramme per square meter second). Lurton2002, Urick1983, Blondel2009). To measure the impedance value (Z), we can use the *equation 3.9*. Where the ρ and v express the fluid density and the particle mass velocity introduced in this section.

$$Z = \rho v \qquad \textit{eq. 3.9}$$

When two dissimilar mediums, such water and bottom or a specific object including vegetation are adjacent to each other, the boundary between the two is called a discontinuity. When sound energy is traveling through one medium and encounters a discontinuity, part of the energy will be transferred across the boundary, part will be absorbed and another part will be reflected into the original medium. It can be viewed in the Figure 15, and the relation between incident angles (grazing angle), reflected and transmitted angle can be expressed by de Snell law. Considering the sonar acoustic image system, we must consider some acoustic reflectivity characteristics for acoustic image formation. Concerning the acoustic reflectivity, it depends on the bottom material and their impedance, grazing angle and distribution of reflective bottom surface. Respecting the impedances, the greater the difference between the characteristic impedances, the greater will be the percentage of energy reflected (Lurton2002), (Urick1983) and (Blondel2009)

At the interfaces, interactions between propagation energy or acoustic pulse level and surface targets will be lead to help the characterization of the surveyed environment. To describe relationship between the angles and the velocities of waves that have a strong dependence with impedances in two different mediums we have used the Snell's Law.

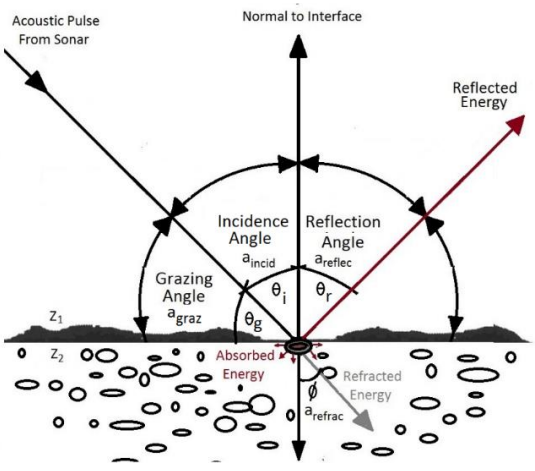


Figure 3.4: *Path ways of acoustic energy after environmental medium interaction*

Consider an acoustic wave propagation through two distinct mediums with impedances Z_1 and Z_2. If the angle of incident acoustic wave is θ_i, then, the reflected angle will be θ_r , and so $\theta_i = \theta_r$. We can see this figurative relationship in the *Figure 3.4*.

The Snell law is represented in the *equation 3.10*, and the ratio between propagate medium velocities c_1 and c_2 equates to the grazing angles θ and refracted angle $\emptyset$ ratio.

$$\frac{\sin(a_{incid})}{\sin(a_{refrac})} = \frac{\sin(\theta_i)}{\sin(\theta_r)} = \frac{c_1}{c_2} \qquad \textit{eq. 3.10}$$

Then, the amplitude of the reflected wave is a function of the reflected coefficient R and is a function of grazing angle, density and velocity of the medium. The numeric value of the reflected amplitude of the reflected energy can be seen in *equation 3.11*.

$$R(\theta) = \frac{\rho_2 c_2 \sin(\theta_i) - \rho_1 c_1 \sin(\theta_r)}{\rho_2 c_2 \sin(\theta_i) + \rho_1 c_1 \sin(\theta_r)} \qquad \textit{eq. 3.11}$$

In traveling through a fluid, an underwater sound signal becomes delayed, distorted, and weakened. It is known as transmissions loss, spreading loss and attenuations loss. The interaction between fluid medium, irradiating energy including mechanical disturbances, surfaces and objects together with its boundaries, forms a complex medium for the propagation of sound. It can create many diverse effects upon the sound emitted from an underwater source. An important factor that influences the acoustic signal response and attenuations in the underwater medium is the wavelength, which correspond to the spatial correspondence to the time periodicity of acoustic signal.

Absorption is a frequency dependent factor to lose some portion of the propagated energy through the process of molecular relaxation. The physical-chemical main process to absorb energy consists in dissociation of ionic compounds caused by the Magnesium Sulphate ($MgSO_4$) and Boric Acid ($B(OH)_3$) in seawater. Some works show and relate the absorption frequency dependency to energy propagation in water using empirical or simulation methods, and to summarize, the higher is the frequency, greater will be the absorption (Sawant 2010), (Pinkerton 1947) and (Li & Liu N.D.). On the other hand, if we have in mind the physical constraints, frequency implies in the absorptions, and it is related to the period of the local pressure variation. If it is greater than the molecule relaxation time the process is reproduced at every cycle and dissipate energy permanently, so the attenuation due to this process appears at frequency lower than the characteristic relaxation frequency of the relevant compound. Many models have been proposed to describe this process and the model most used is the Francois-Garrison Model (Lurton 2010) and (Blondel 2009).

To specify the Transmission Loss, we need to compare the amount of intensity of the signal at a specific range from the source. Normally in the literature it is used the distance at one yard as reference:

$$TL = 10\log\frac{Intensity\ reference}{Intesity\ in\ target\ distance} = 10\log\frac{I_{(at\ 1\ yard)}}{I_{(target\ from\ the\ source)}} \qquad \textit{eq. 3.12}$$

Where: $I_{(at\ 1\ yard)}$ - Is the intensity reference at 1 yard from the source;
$I_{(target\ from\ the\ source)}$ – Is the intensity at the target point from the source.

The received source level (L_s) can be extracted from the relationship between the intensity source level (SL) from the equation (19) and the transmission loss (TL) from the *equation 3.12*. These relations are expressed by the equation 3.13.

$$\begin{aligned} TL_S &= LS - TL = 10\log\frac{I_{(at\ 1\ yard)}}{I_0} - 10\log\frac{I_{(at\ 1\ yard)}}{I_{(target\ from\ the\ source)}} \\ L_S &= LS - TL = 10\log\frac{I_{(target\ from\ the\ source)}}{I_0} \\ L_S &= 10\log\frac{I_{(target\ from\ the\ source)}}{I_0} \end{aligned} \qquad \textit{eq. 3.13}$$

The quantity of transmission losses will consider spreading loses and absorption loses and besides it will depend on the type of geometric spreading, medium characteristics (density and temperature), and finally the wave frequency.

3.2 Acoustic Data Extraction

The acoustic backscatter can be used to perform an imaging systems and covers thus all processes related to image processing techniques. Since we perform an image from objects and environment considering the physical characteristics about sensors that convert acoustics pressure to electric signals, these signals will be sampled and quantized into digital signals and then they can be processed by an electronic processing system based on microcontrollers or microprocessors (SeaBeam Instruments 2000), (Blondel 2009) and (Lurton 2000).

The goal is attaining a spatial representation (1-D to 2-D) of an acoustic sampled area and extract a signal backscatter from an object or surface, and represent it conveniently with a referenced space where the sonar is localized. Supposing that we know besides the physics characteristics of the sensor, a geometry of received signal in the sensor and the environmental properties of the medium. A 2-D spatial acoustic representation can be called as a simple acoustic image, and is an approximation of objects and surface of the sampled area like a photo. Using techniques and algorithmic of image processing for optics images, with some considerations and corrections inherent to acoustic image formation, we can extract physicals characteristics about the surveyed area and objects into it (Lurton 2000), (Lurton 2010), (Urick 1983) and (Medwin & Clay 1998). In this work, we used acoustic images of the bottom surface acquired from a sidescan sonar, generating a digital gray scale image, at two-dimensional (2-D) surface. We worked with a non-bathymetric sidescan, that is a type of sensor that we have no information about bathymetry directly because of his characteristics in the geometric acquisition and physical arrangement of the sensor array. Despite this, we can get a high-resolution representation from a bottom survey like an acoustical photo. The image resolution acquired at least, depends on system settings as a maximum range, acoustical pulse frequency and distance to bottom from sensor. A special sidescan sonar device that incorporate a distinctive physical arrangement of sensor arrays and special algorithms, can infer the bathymetry from bottom represented in the image. It is reached using a backscatter signal processing considering the phase difference between two separated sonar array sensors, this process is also called interferometry (Lurton 2010). A sidescan sonar represent a useful tool to surveys underwater bottom, applications come from civil, scientific and military interests. For civilian and some researches application the hobbyists and professional divers normally searches shipwreck (Diercks et al. 2010), (Blondel 2009). Industry uses the acoustic representation of underwater environment to survey navigation channel, underwater infrastructures, petroleum products pipelines, communication cables and etc. (Blondel 2009). Researches in geology sciences can study geological facies and sediment characterization, with a space-temporal sampling method can study additional characteristics about sedimentology as sediment displacement and redistribution (Blondel 2001), (Lurton 2000) and (Collier & Brown 2005). Archeologic discovery can be carried by a sidescan image analysis too (Quinn et al. 2005), (Atallah et al. 2005). In Biology, scientists use acoustic representation in studies about benthic habitat including a reef formation, structure and conservation(Brown & Collier 2008), (Cochrane & Lafferty 2002). Military applications is about port and coastal security, underwater target detection as mines and some structures that implies in navigation security and strategic underwater subsurface mineral resources (Reed et al. 2003), (Rao et al. 2009), (E. Dura et al. 2005) and (Dura 2011).

3.3 Sonar Device System and Process: A Basic Overview

Sonar devices are largely used in hydrographic area to mapping surveys application, in this case we have to take into account the footprint coverage of the sonar beam, some transducers will project over a surface a circular footprint, and others will project an ellipsoid footprint. There are some types of sonars, each of them can be called by a short name that describe your using and some of your characteristics.

A Single-Beam Echo Sounder (SBES) is a type of single element sonar, the SBES devices use the most basic function of an echo sounder, they measure the range to the target underwater, and with this information it is possible to extract the estimated distance from an object or the ocean floor by emitting a single acoustic pulse in a specific direction and sensing your backscatter from it. Range to target can be infered by the total pulse travel time, for that we need the wave velocity in the propagation medium. Despite new technologies developed to acoustic surveys, single beam echo sounders (SBES) still remain, and is a traditional equipment used on hydrographic surveys (IHO 2008) and (El-Hattab 2014), this type of sonar can extract the bathymetric information from one specific location with each ping, measuring the depth to a specific location on the underwater bottom, directly below the vessel or underwater vehicle doing the survey. This task is knowing in the Hydrography area as a Bathymetric survey. Nowadays this device is used combined with any global system reference as GPS and

inertial sensors to correct motions of the sensor or platform. With this set, we can extract an information about depth at specific point on the underwater bottom (Umbach 1976)and (IHO 2008). Thus, according the survey vessel or the vehicle movement it takes a sample point over bottom along the line track survey, the bathymetric accuracy and resolution will depend on the sensor applied, surveys settings and environmental conditions (Blondel 2009).

One type of multi-element sonar is based on a set of transducers in a line array, this physical arrangement is used in a type of sonar called Side Looking Sonar or more commonly Side Scan Sonar (SSS), the SSS and their survey characteristic methods are designed to provide "acoustic images" of the seafloor (Lurton 2010) and (Lurton 2000). The sidescan sonar system ensonifies the underwater bottom using two array of ultrasonic sensors, one in each side of the body device. The resultant acoustic beam is large in the plane perpendicular to the sensor trajectory and narrow in the plane of the trajectory (Lurton 2000) and (Blondel 2009). Each pulse generated by the sonar transducer will back to the sonar, the amplitude of the backscattered echoes will be sampled and quantized to form a row, the vector that each column is composed by gray level pixels. The value of each pixel is dependent of the backscatter intensity corrected by a Time Varying Gain (TVG) amplifier, which normally has a gain proportional to the square of the elapsed time measured from the sidescan firing pulse to the backscatter intensity reached to sensor for each specific pixel (Blondel 2009) and (clifford1980). A disadvantage regarding the sidescan TVG corrections, is about the system that has a constant gain characteristics, in this case the possibility of input range-dependent variations, causing distortions on the image are increased (Anstee 2001).

The sidescan device will acquire the acoustic backscatter line by line, according the sensor travelling on the trajectory to compose an acoustical image. The acoustic beam will ensonifier the bottom with an incident grazing angle and will result in an acoustic image representing the echo strength from the bottom, and represent it numerically into a gray level range from 0 to 255. For example, a more or less medium gray level of reflective surfaces, can represent a flat bottom; high gray level values is from, as rigid objects or rigid surfaces that has orthogonality with the incident acoustic pulse propagation, low and very low gray level from reflective surfaces, as shadow can indicate a region that no acoustic pulse ensonifier (Lurton 2000), (Lurton 2000), (Urick 1983). An especial type of sidescan sonar that can extract the bathymetric information and at the same time measure the bottom reflectance is called bathymetric sidescan sonar and generally it uses the interferometry principle, it means, the phase difference between the signals received on two separate receivers of the sonar arrays.

A Single Beam Mechanical Scanning sonar (SBMS) is a special type of Single-Beam Echo Sounder that the sensor body enclosure a mechanical scanning device to rotate the transducer over 360° in relation to the sensor axis. As the sidescan sonar a SBMS is an image sonar, but different to the SBES their acoustic beam has an elliptical cone resulting in an elliptical footprint over the bottom surface. The acoustic image around the sensor is forming by the backscatter from the scanning ensonifier.

Since the sidescan is a type of multi-element sonar, and it consists of a transmitter and receiver array, the common arrangement is on array perpendicular to each other (SeaBeam Instruments 2000). Each orthogonal array consists on several piezoelectric elements and each of these can receive a portion of the backscatter at specific angle and correlate it, combining and summing in phase to extract a bathymetric information from each surface bottom portion. Then, the multibeam sonar is capable of resolving the arrival angle from several directions simultaneously(SeaBeam Instruments 2000) and (Mussai 2010).

3.3.1 Remote Sensing Using a SBES

Bathymetric data survey, in the first approach is the process to acquire a set of measures representing topographic bed relief relating to a standard reference. As a standard reference of Earth's geometry, we have two distinct models, ellipsoid and a geode approximation. The Ellipsoid is a mathematical approach that uses an approximation of Earth's geometry to represent the Earth's shape. In this representation, the Earths shape are characterized by an ellipsoid of revolution. The Geometrically description can be seen in the *Figure 3.5*, where we have an Earth's ellipsoid representation. In this ellipsoid, the two axes in horizontal plane (x, y), represented by a and b symbols have equal values and in vertical plane (z), represented by c, has distinct and shortest value than a and b. The Ellipsoid model was calculated based on the hypothetical equipotential gravitational surface, this is defined by either the semi-major axis, represented in the figure 16 by a and b $(a = b)$ and by the semi-minor axis c or can by switched with the representation using a and the flattening $f = (a - c)/a$. This model has a smooth surface and can be described mathematically, and because of this, the ellipsoid model can be used as basis for map projections.

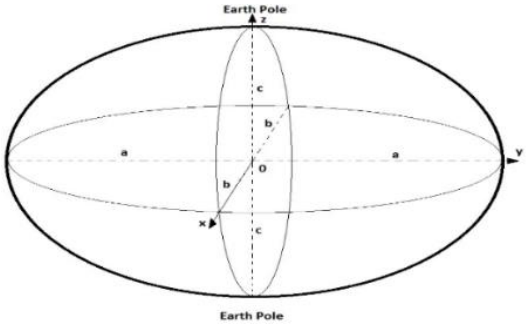


Figure 3.5: *Illustration of Ellipsoid Earth's Surface Approximation*

The Earth's geometry isn't a perfect ellipsoid shape and in some locations the ellipsoid approximation don't describe some irregularities, hence, there isn't one single ellipsoidal representation that can be used to every part of Earths (IHO 2012), (Bowditch 2002) and (Schowengerdt 2007). For this reason, we will refer the Earth's shape as a geoide structure and different models of ellipsoids to model the Earth's shape are used to cover different regions of the Earth. For this work, we used the latest revision of the World Geodetic System, WGS-84, as the reference system. The geoide shape uses the Earth's gravity field to model the specificity and regional irregularities of Earths, and generate the best shape approximation for a specific region, therefore as a reference for elevation data calculous, the surface of ocean would align with the geoide surface, and this is the best approximation of the mean sea level. In some places the geoide and the mean sea level are not coincident for some environmental factors and one variable is used to separate the geoide and the mean sea level, this variable is called Sea Surface Topography (SST) (Kovalevsky et al. 2011), (Intellingence Agency of EUA 1984) and (IHO 2012).

The standard reference to depth measure is the Mean Sea Level (MSL) acceding a specific geoide approximation, but this cannot fit the real shape of Earth's surface in some regions. According this, the depth value of any point laying in the Earth's surface must be related to other local environmental variable as an instantaneous tide value. Using an instantaneous tide measurement and correlate it with historical local variations of tide, we can estimate the best local value for correlate the elevation data to MSL and the best fit to store a bathymetric topographic measure, it means, the distance from bed to Local MST. Concluding, to perform the bathymetric survey, there are some considerations related to the sonar transducer used, survey vehicle, geodetic reference (datum) and dynamic of local water level (tide). In the *Figure 3.6*, we can see an illustration of a conventional boat survey and a sonar transducer installed in his hull, observing the difference between distance from transducer to bottom and the water surface to bottom. Regarding to an accurate mapping process, it is necessary some corrections in vertical and horizontal directions because the off-set distance between sensors, reference systems and periodical changes in environmental dynamics. If we consider a concrete instant of measure, the offset value between the sonar distance measure D_{SB} and the distance to water surface is represented in the *Figure 3.6* as the value D_{SW}. To obtain the correct value of the topographic relief of bottom D_{BW} , the offset D_{SW} need to be added to the value measured by the sonar, this relationship is expressed by *equation 3.14*.

$$D_{BW} = D_{SB} + D_{SW} = BP \qquad \textit{eq. 3.14}$$

In the *equation 3.14*, the variable BP is used to name a past point measured, adding a sequential number to each new point acquired and, store the value surveyed into the structure of data survey log. A set of measured points BP_1 to BP_n into the hole of surveyed region is called bathymetric survey or bathymetry. To store bathymetric information, the spatial positioning information to each bathymetric point are relevant to identify each surveyed point in any georeferenced map representation and it is stored together in the same log file. The spatial localization can be stored in any acceptable georeferenced coordinates system and datum, describing the position of a geographic location on the earth's surface representing the survey track for horizontal measures and elevation level for vertical.

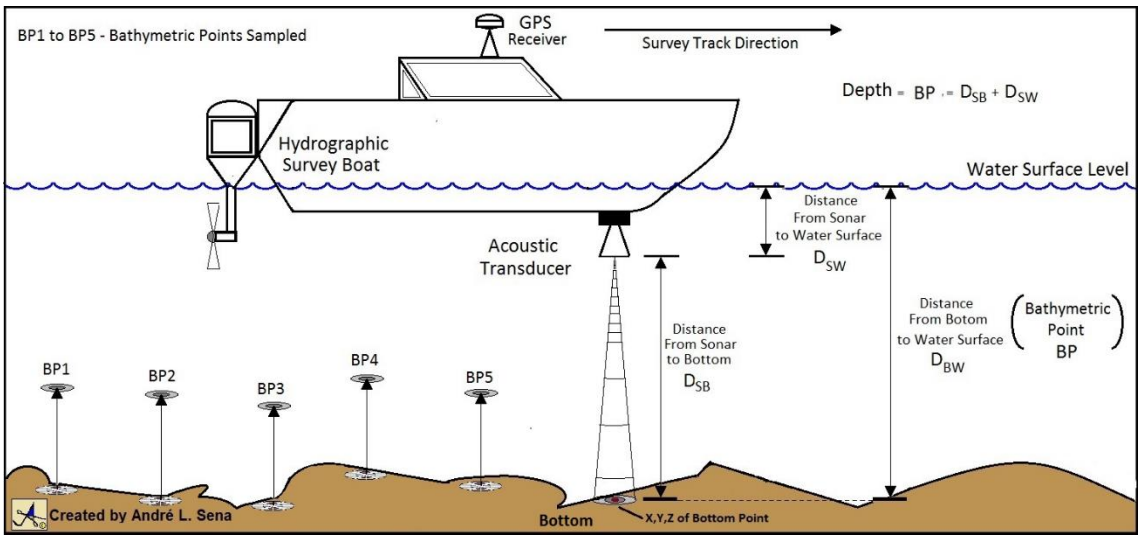


Figure 3.6: *Illustration of an example of boat survey using an acoustic transducer to get samples of bottom depth.*

The survey track follows a determined horizontal navigation trajectory commonly pass by a set of predefined way-points and the sequence of point sampled is usually called transect. Generally, some physical considerations related to data, acquire in space and time domain, are needed to ensure a data conditioning and corrections.

3.3.2 Depth Reference, Tide and Vertical Datum

There are two main applications to use acquire underwater elevation data, that could come from acoustical sensors using marine vehicle as manned survey boat or unmanned robotic vehicle. The first approach is related to vehicle navigation aid, i.e., when the vehicle navigates through a trajectory and uses the distance information from his submerged hull to underwater bed, to avoid collision with respect to underwater bed structures. The second application relates the use of acoustic measures to bathymetric surveys, when the relative distance measure from acoustic source (acoustic transducer or sonar) to target (point of underwater landscape) will be used in topographic relief representation and mapping. In this case, this distance isn't the correct topographic elevation according to a local geodetic reference, and represent only the distance between acoustic source device to the surface target or underwater bed (Bowditch 2002)(Medwin & Clay 1998)(Randall 1997). The standard reference to topographic survey is related to the distance from underwater bottom surface to local water mean surface level also called Mean Sea Level (MSL), this represents the average sea level over time and, in some cases, over space (Kovalevsky et al. 2011). One approximation is the representation of a bathymetric point symbolized by the variable *BP* in the *Figure 3.6* and in the *expression 3.14* reveals the instantaneous sea level. In modern topographic survey, elevation points that are measured above the water surface are called traditionally as a topographic elevation data. Topographic elevation data are stored with respect to any geodetic reference model. Otherwise, bathymetric data, is a special case of topographic elevation data, that was acquired from underwater Earth's surfaces, and is related to a tidal datum. A set of reference points can be installed in some places spatially distributed over land as a geodetic reference to positioning and elevation, mainly for applications in local measurements. In case of bathymetric elevation data, the use of a local tide measure gauge, calibrated with respect to geodetic reference point, helps bathymetric surveys to perform tide corrections to store a more accurate elevation data than use an estimated geodetic model to predict the tide movement.

Tide is a natural and periodic movement of sea surface level, and it is characterized by a long-period waves that roll around the planet (NOAA Coastal Services Center 2009) and (NOAA Coastal Services Center 2005). The local tide information is an import factor in the process of bathymetric surveys. The movement of sea surface will affect the accuracy of data extracted according the time expended to complete the survey, the region of survey, period of the day and lunar cycle. The tidal movement is mainly characterized by changes in tidal range, i.e., the rise and fall along the time, this movement will characterize the predictions and establishment of a local tidal datum. This range will change in distinct modes in some regions of the Earth's surface depending mainly by changings in attractive forces caused by lunar and solar cycles and the interaction with Earth's rotation. The tidal movements can affect some aspects of coastal oceanographic and geomorphologic dynamics (Bowditch 2002), (IHO 2012) and (Medwin & Clay 1998). The previous exposure factors, will affect the survey process to acquire

a bathymetric data in coastal zones. Measures of underwater topographic surface using sonar extract the distance from transducer to underwater bed, and can be used in two distinct ways at the same time: the instantaneous real distance from sensor to underwater bed and to calculate the topographic elevation with reference to a reference MSL or a local tidal datum. Both are important and are used in different ways, the first one is applied in navigation tasks and the second one to bathymetric surveys (mapping applications).

The tide movement has three main vertical reference levels; they are influenced by the interaction among Sun, Moon and Earths movement, e.g., the Mean Low Water (MLW), the Mean High Water (MHW) and Mean Sea Level (MSL). But the extreme values of tide are very important in port operations and are named as highest high waters (MHHW) and the lowest low waters (MLLW) of a tide cycle. The MHHW and MLLW occur according to the Moon, Sun and Earth, they are nearly in alignment, representing the full or new moon, the periodicity of the tide movement is approximated 2 times per month and this is also called Spring Tide (NOAA Coastal Services Center 2005) and (IHO 2012). An illustration of tidal datum can be seen in the *Figure 3.7*. To survey an elevation point over the Earth's surface (depth or height), we need a reference to a specific vertical datum (sets of data) from which all elevations can be determined. A vertical data represents the reference for elevation heights measures in a consistent system, all elevation measure data will be referred to specific local, regional or global vertical datum. Despite the existence of many different types of vertical datum to bathymetric use, the predominance are based in tidal and geodetic datum (Willumsen et al. 2007) and (IHO 2012).

A distinctive type of vertical datum is the WGS-84, as mentioned before, this is due the large use in the GPS system and the recommendation of the International Hydrographic Organization (IHO 2012) as the standard geocentric reference system. In this work, this datum was used to acquire the bathymetric data because the WGS-84 was the best geodetic model to fit the surface of surveyed area.

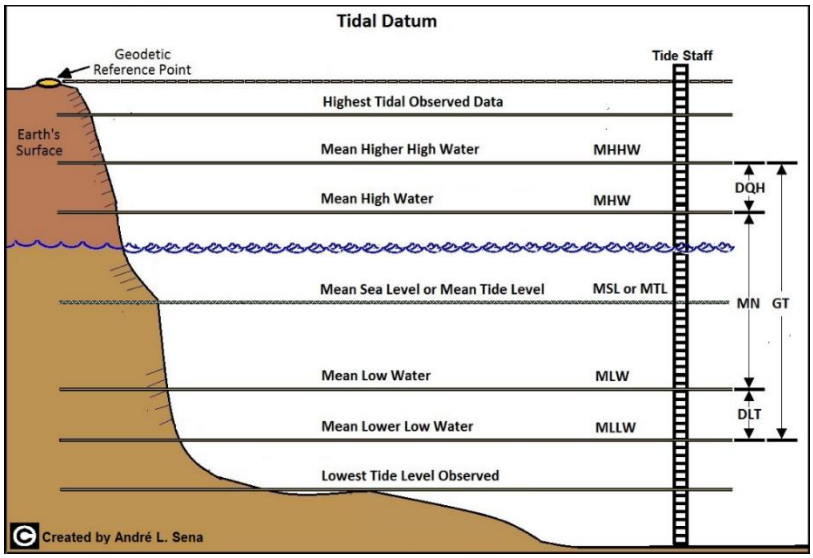


Figure 3.7: *Tidal Datum Illustration. where GT is the Great Diurnal Range, MN is the Mean Range Tide, DHQ is the Mean Diurnal High Water Inequality and DLQ is the Mean Diurnal Low Water Inequality*

The MSL is a main parameter into the vertical datum representation, the MSL is not an equipotential surface, and the measure of tide variations in specific site is performed using a tide gauge station, the station measures the rise and fall of sea level continuously. The MSL is not a constant and can be affected by variations in currents, wind, hydrologic factors (Kovalevsky et al. 2011). In operations of bathymetric surveys, a common sense is install a MSL station gauge in the surveyed region, with a system calibrated to any close georeferenced point or station to measure the water sea level variations while the bathymetric survey is executed and the data login file of the gauge need to be stored with a timestamp and the data log of the bathymetric survey too. It will give the opportunity to compare the data stored from the bathymetric survey at the same time of sea level are taken to perform data corrections according tide variations during the survey.

3.3.3 Tide Correction in Bathymetric Surveys

Bathymetric surveys can be performed in a variety of area extension and it will imply in the time to acquisition the complete hole of data to conclude the survey process. Therefore, some corrections need to be done in the

bathymetric data to behold changes in the local tide dynamic resulting in changes in the instantaneous level of local tide regarding the MSL during the data acquisition. The tidal dynamic is a result of some factors including earth's dynamics and environmental interactions. The seasonal cycle of MSL at the coast is affected by gravitational interactions and meteorological/hydrological contributions that acts in the sea level time series (Tsimplis & Woodworth 1994), the tidal range will vary cyclically according these interaction and mainly by the combined effects of gravitational interaction forces among Earth's, Moon, Sun and the Earth's rotation (Medwin & Clay 1998). During the bathymetric survey, the time to complete the data extraction of survey can affect the accuracy of measures. The effect of changes in tide level affects the relative vertical positioning of the boat survey or surface robotic vehicle and will lead an incorrect measure of underwater bed relating the MSL or geodetic reference system. As the surface boat or robotic survey platform rise and fall together with the tide changes, the relative position between the transducer and underwater bed will change. To store and represent the correct value of topographic relief, the application of a correction factor is necessary to perform the correction of tide change during the bathymetric survey, adjust heights based on the measurements of a local and calibrated tide gauges, adjust the MSL measures data based in others stations far from the survey area. The *Figure 3.8* is a hypothetical representation of a tide measure gauge insert in the marine environment and they theoretical interactions with mains tidal references and survey boat measurements. Instantaneous measure of sea level usually is referenced with tidal datum using a local tide measure gauges, and may be considering three components to do some correction and represent the measure to a geodetic consistent model and extract an observed measure (OM) as follows: actual tide (AT) also called instantaneous tide measured, mean sea water level (MSL), and meteorological residuals (MR)(IOC-UNESCO 1985). The relationship among these components are described by the *equation 3.15*.

$$OM \approx AT + MSL + MR \qquad \textit{eq. 3.15}$$

The component MR or meteorological irregular residuals results of variations from local weather and are related to a non-tidal component which remain after removing the tides by analysis. The MR components will have a relative importance according the latitude, period of year, proximity and extensions of shallow water areas. Storms and the inverted barometer effect, the last caused by changes in atmospheric pressure, acts vertically on the sea and can affect the tide level (IOC-UNESCO 1985). The instrument to measure the tide level along time is called tide gauge, and records water levels at fixed intervals of time. The right measure of bathymetric data need to be corrected before it is stored, the correction can be done using the GNSS system or a local chart datum using a tide gauge measure stations with reference

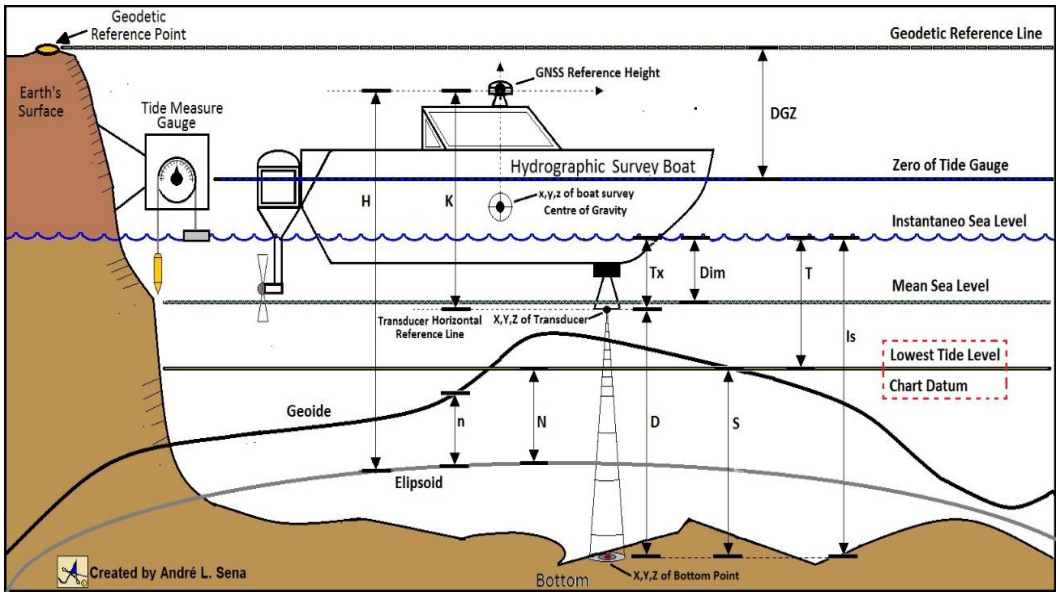


Figure 3.8: *Tide References and Gauge Measurements Relationship: Bathymetric hydrographic surveys, executed by a survey boat using a single beam sonar, extract the depth just below you line track or survey trajectory. The depth extract represented by D is not the right topographic underwater relief measure.*

The GNSS based system can be used in the boat survey to acquire the positioning information from GPS constellation satellite (X, Y, Z position information) the bathymetric elevation of underwater relief sounding can be obtained by subtracting the height of the vessel's GPS antenna above the WGS84 ellipsoid. The value **N**, represented in the Figure 19, is the distance between the ellipsoid chart and the datum reference, needed to calculate

the depth value. In this case we don't need the tide measure data from a tide gauge measure to estimate the bathymetric depth, this will be derived from GPS and sounding data(Surveyors, International Federation Greenland & Higgins 2006). The *equation 3.16* is a simple mathematical expression to estimate the depth or topographic elevation relief value from a chart datum using the GPS signal system.

$$S = K + D - H - N \qquad \textit{eq. 3.16}$$

Where according the figure 19 we have the variables definition as:

S - Is the corrected depth value according official chart datum;
K - Is the distance from GPS receiver a type of Global Navigation Satellite System (GNSS);
D - Is the distance measure from sonar transducer (Ex. A Single Beam Echo Sounder) to a point in the underwater bed;
H - Is the distance from GPS receiver to an ellipsoid reference level; and
N - The distance from the official local chart datum to the ellipsoid reference level.

The bathymetric survey uses only the single GPS system information without instantaneous tide measure cannot separate the model between chart datum (CD) and WGS-84, this results in depth measure inaccuracy because the tide measure is one fundamental information to do corrections and without is we cannot know the CD relationship and instantaneous water level measurement. In the Appendix we show the flowchart published by (Surveyors, International Federation Greenland & Higgins 2006), in this appendix it is related the flow process to calculate the separation values between CD and WGS-84 at a discrete point. Some studies were performed to compare a new implementation of GPS system information to obtain an accurate three axis positioning, Arroyo et al have been presented a study of Global Differential GPS System (Arroyo-Suarez et al. 2005), the GDGPS structure is composed by a Global GPS Network, the Real-Time GIPSY developed by the Jet Propulsion Laboratory (JPL) of the National Aeronautics Space Administration (NASA) in Spring 2001 and a signal-in-space provided via Inmarsat. This paper demonstrates level of accuracy for positioning in altitude information about 95%, which is acceptable according Table 1 in (IHO 2008). On the other hand, a unique tide gauge measure calibrated in respect to any or some better proximal geodetic reference points can be used to measure the instantaneous local tide, value represented by **T** in the *Figure 3.8*. The surveyor can obtain their depth with respect to WGS84 accurately. In this case, and adding the instantaneous information of local sea level we will infer the depth using the equation 3.17.

$$S = D + Tx - T \qquad \textit{eq. 3.17}$$

Where Tx is the vertical distance from instantaneous sea water level to sonar transducer. All data used in this work was extracted using a manned survey boat and a single beam echo sounder, to correct the depth information we work with a DGPS real time system information.

3.3.4 Remote Sensing Using a Sidescan Sonar

As mentioned before, the portion of sound energy scatter back to the sensor after hit the target (object or underwater bottom) is called backscatter. Generally, it is a function of the pulse incidence angle to target and material properties. The dependence of material properties has a relationship with the impedance between medium boundaries. A backscatter can contain geometric information about the target or surface being imaged, the survey detail level depends on acoustic transducer type, transducer arrangement and the method used to acquire data. Despite some limitations, the reconstruction of targets or seafloor can be performed using acoustics backscatter information through inverse acoustic models (Lurton 2010) and (Urick 1983). Consider a flat bottom, for an image sonar, the backscatter strength dependence with the frequency and grazing angle variation, these can be used to reduce radiometric and geometric distortions over the data set (Blondel 2009) and (Lurton 2000). The signal frequency and wavelength are one of the main operational parameters used to characterize the acoustic system and their interaction in the medium and characterize the distance over which the wave's shape repeats. The *equation 3.18* represent the relationship between frequency and wavelength.

$$\lambda = \frac{c}{f} \qquad \textit{eq. 3.18}$$

Where: λ - is the wavelength of acoustic wave;
c - is the mean velocity of acoustic waves;
f - is the frequency of acoustic waves.

For a sound velocity of 1,500 m/s, underwater acoustic wavelength will be 150 m at 10Hz, 1.5m at 10kHz and 0.0015m at 1MHz. The main constraints on the frequency usable for an application. The sound wave attenuation in water, limiting the maximum usable range whose effect increases very rapidly with frequency.

The sonar system used to acquire the data in this work uses the Compressed High Intensity Radar Pulse (CHIRP). This technique uses two different frequencies to emitting the acoustic pulse, the chirp sweeps across a band of frequencies. The difference between start frequency (f_s) and end frequency (f_e) will lead the Bandwidth. Thus, it will generate a resultant signal with a resultant frequency equal to the centre of this frequency difference. For example, considering a sonar system that use two frequencies to generate the acoustic pulse with values: $f_s = 430\ kHz$ and $f_e = 470\ kHz$, the bandwidth (BW) will be extracted using the *equation 3.19*.

$$BW = f_e - f_s \qquad \textit{eq. 3.19}$$

The resultant frequency (f_r) can be expressed using the half part of BW as described in *equation 3.20*.

$$f_r = f_s + \frac{BW}{2} \qquad \textit{eq. 3.20}$$

In practice, the frequency change will give a unique and short acoustic signature than the low frequency used (f_s) and the high frequency (f_e) used to produce the chirping pulse. Using a short acoustic signature carried by chirp techniques, results in an increment of the resolution for target detection. We will have a better compromise between power needed, ping duration, resolution to provide even greater detail and information when processed.

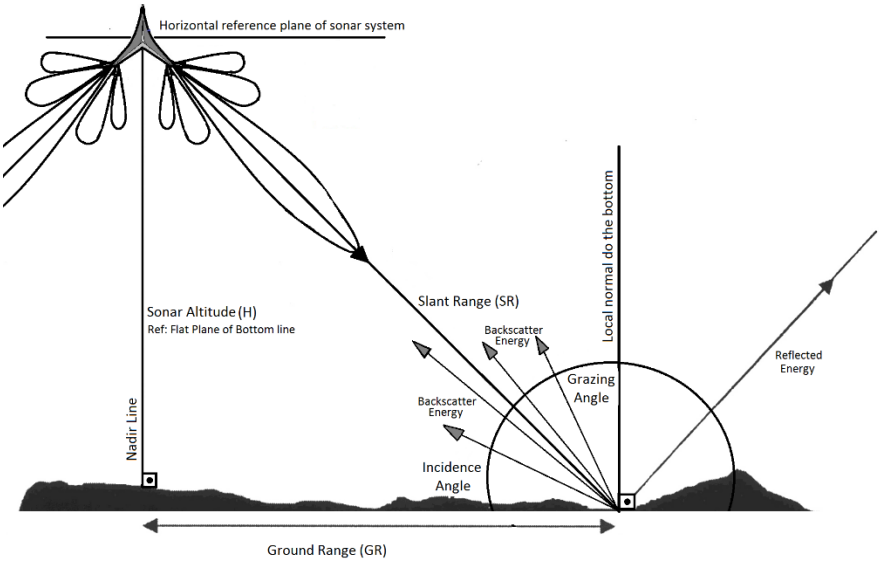


Figure 3.9: *Illustration of sidescan data acquire geometry to image formation (adapted from Blondel 2009).*

The distance travelled by the acoustical wave front, from the transducer to the target (bottom or object), is called the slant-range (SR), see *Figure 3.9*. The distance between the vertical imaginary line below the transducer and the target is called ground-range (GR), this is also understood as a projection of the SR line over the bottom plane. The angle between the incoming wave and the reference plane of the bottom is called angle of incidence and the angle formed by the incoming wave and the vertical plane (normal to the bottom), is called grazing angle, and represent, 90° - angle of incidence (Blondel2009).

Sidescan sonars work emitting an acoustic pulse at specific time rates (firing time) and pulse lengths, T. A short pulse emission will produce a thinner spatial pulse length resulting in a higher resolution. Short acoustics pulses need more power to reach the same range of large pulse to compensate the transmission losses and abortions (Lurton 2000), (Urick 1983) and (Mazel 1985). In this work, we can reach a good resolution using a chirp sidescan sonar system. The resolution decrease as far as the acoustic pulse travels (Lurton 2000), (Urick 1983) and (Blondel 2009).

The resolution concept is one characteristic to measure the sonar system performance and can be defined as the capability to distinguish separated entities into the environment or surveyed area and represent it into the acoustic image. The sidescan system spreads their bearing to target, object or surface in a solid angle, and when hit one

generic surface will ensonifier it as we can see in the *Figure 3.9*, forming a general two-dimensional footprint area with two perpendicular resolution components: along-track and across-track resolution.

3.3.4.1 Along-Track Resolution

The resolution in the along-track direction, ($jResol$) is dependent of sonar system velocity, distance to target from sonar system (Range) and horizontal aperture beam angle (Azimuth), also knowing as beam with (θ_{AZ}). Concerning the last the beam width will be determined by array characteristic, mainly the length of the sonar array (Lurton 2000), (Lurton 2010), (Waite 2002) and (Blondel 2009). Conventional sidescan sonar system devices need to travel at constant velocity and between levels defined by each manufacturer. The along track resolution will be also dependent with the range (GR)(Blondel 2009).

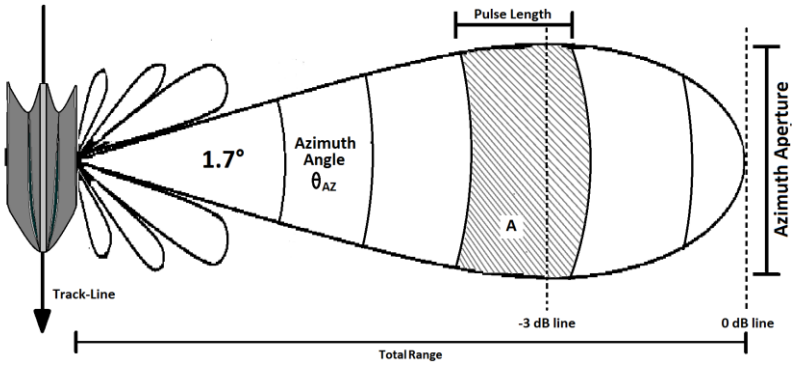


Figure 3.10: *Illustration of SSS Along-track beam*

A short and good definition about along-track resolution is written by (Blondel 2009): " The along-track resolutions is the smallest of the distance travelled over ground during the reception interval, and width of the beam on the ground". In the *Figure 3.10* we can see an illustration of along-track SSS beam.

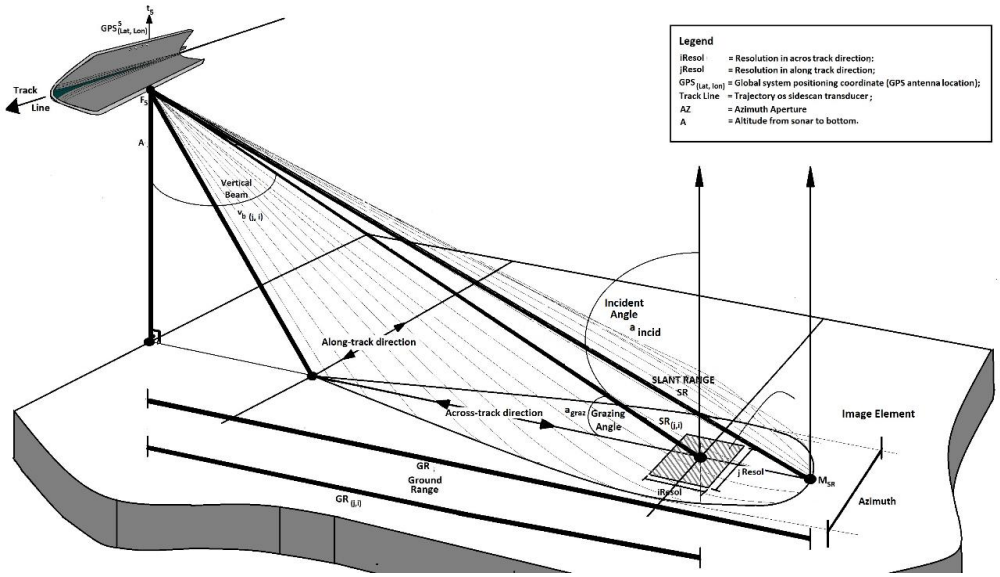


Figure 3.11: *Sidescan footprint and element of image (pixel), with geometrics correlations*

The along-track resolution ($iResol$) can be defined by the mathematical *equation 3.21*, considering the sonar system travelling in a constant velocity (Blondel 2009):

$$iResol = Range \times \mathrm{Sin}(\theta_{AZ}) \qquad \textit{eq. 3.21}$$

If we have the sonar system altitude from bottom (A), length of the sidescan sonar array (L), and the pulse wavelength, we can have one expression to describe the azimuth resolution in function to the incidence angle (a_i),

if we maintain the constant altitude. We can use the *equation 3.22* to describe the numerical along-track resolution too.

$$iResol = \frac{A \times \lambda}{\mathrm{L} \times \cos(a_i)} \qquad \textit{eq. 3.22}$$

The along-track resolution is a main factor to take into account when measuring the capability to distinguish into the sidescan acoustic mage two objects related to the horizontal distance (along-track direction). If the distance between them is shorter than along-track resolution they will be represented as only one object into the sidescan image. We can see it, in the *Figure 3.12*, objects (C, D, H and I), "C and D" are represented as only one object in the Sidescan Segmented Starboard Image. The same effect is represented for the targets "H and I" in the Sidescan Segmented Image representation for portboard.

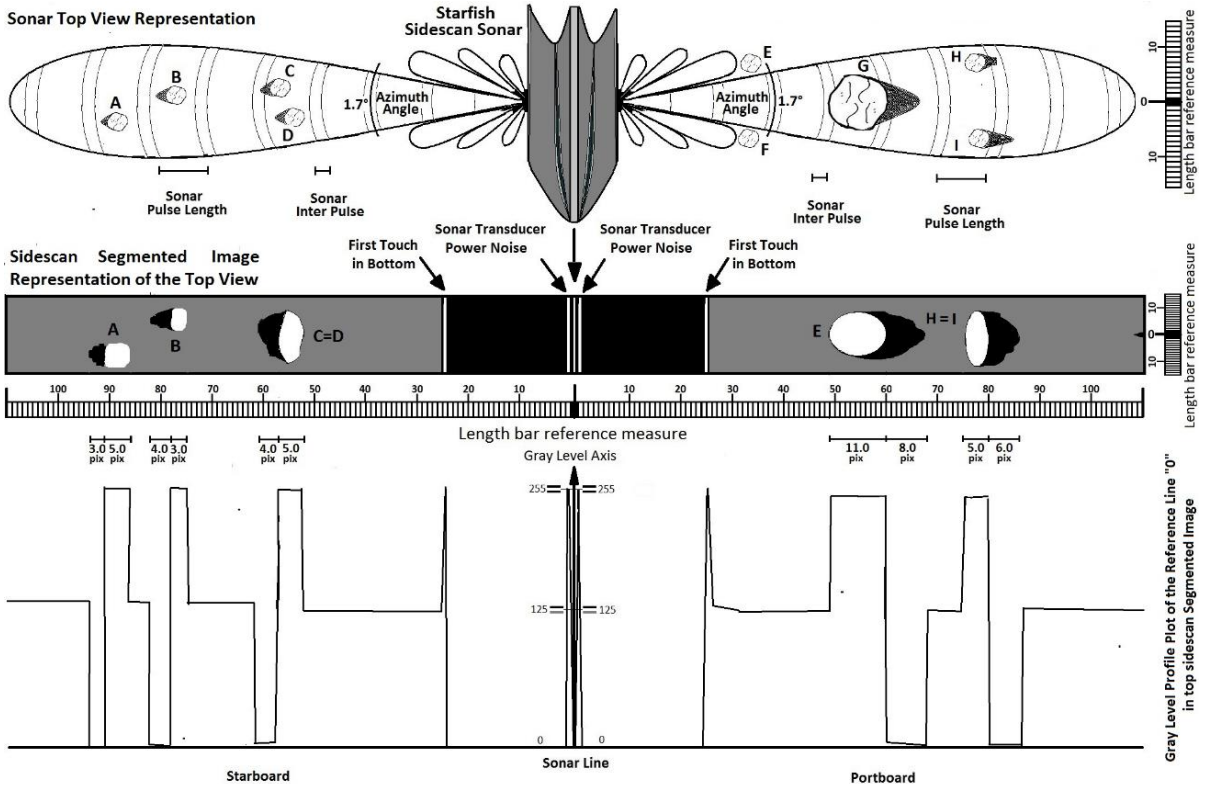


Figure 3.12: *Top view of SSS beam (up illustration), digital acoustic image representation (meddle illustration), and line profile of backscatter signal (bottom illustration)*

In the *Figure 3.12* we can observe in the centre of figure, an illustration of the three gray levels class representation of the sidescan acoustic image. In this representation, a background or bottom have a mean of gray level values about 125 in gray scale, targets are represented with a high reflectance area about 255 in gray level and the area that we have not an acoustic backscatter is a shadow region, and numerically is 0 in gray scale. The bottom plot in the Figure 23 describes the changes profile levels between three classes in the line "0" of the segmented image.

3.3.4.2 Across-Track Resolution

In a sidescan sonar system, across-track resolution can be related to the minimum distance that two targets can be distinguish in the orthogonal direction to the sidescan sonar system travels (Sonar Track). The across-rack resolution is dependent of the pulse length, each pulse will occupy an equivalent "distance" related to its pulse duration (τ), the sidescan acoustic wave propagation velocity (c) and the grazing angle (a_{graz}). This is referred as "range resolution", and it can be approximated by the *expression 3.23* when the $a_{graz} = 0$, in case we don´t have the grazing angle information.

$$R_{resolution} = \frac{(Pulse\ Length\ \ x\ Velocity\ of\ Sound)}{2} = \frac{\tau \times c}{2} \qquad \textit{eq. 3.23}$$

Another way to calculate the across-track resolution is through the equation 35 if we have a grazing angle information.

$$R_{resolution} = \frac{(Pulse\ Length\ \ x\ Velocity\ of\ Sound)}{2} \times \frac{1}{\cos(a_{graz})} = \qquad \textit{eq. 3.24}$$

$$= \frac{\tau \times c}{2} \times \frac{1}{\cos(a_{graz})}$$

The velocity of acoustic propagation (c) depends on some factors, such as water temperature, salinity(density) and pressure (Lurton 2010), (Urick 1983) and (Mazel 1985). The value of 1500 m/s is usually taken as the average sound velocity in seawater, *Figure 3.13* shows the sidescan sonar image.

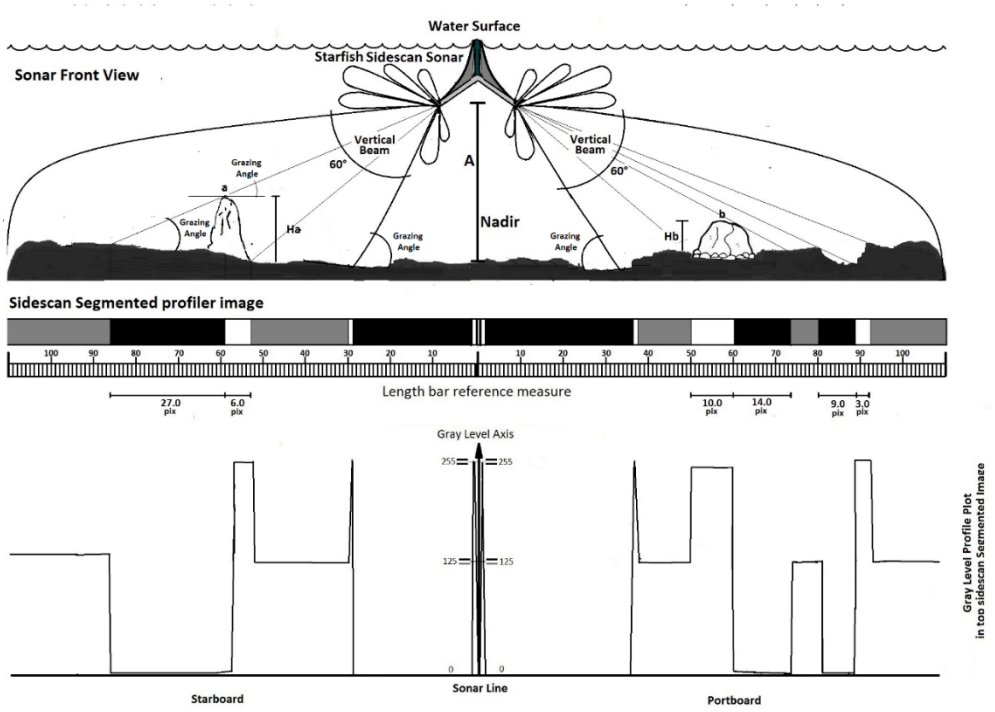


Figure 3.13: *Front view of SSS beam (up illustration), digital acoustic image representation of one data line (meddle illustration), and line profile of backscatter signal (bottom illustration)*

In the middle of the *Figure 3.13*, we can see the grayscale strip image (Sidescan Segmented Profiler Image) related to the geometry, some detected targets characteristics and bottom interaction. The bottom illustration, in the Figure 24, is a profile plot into three gray scale level classes, from that we can detect targets, shadows and background. When we have the sonar altitude information and a grazing angle (a_{graz}), we can infer the height of target object from bottom simply measuring the distance from sonar (GR) and the length of shadows.

For a sidescan sonar systems the information from Table 1 classify SSS qualitative resolution according to the operating frequency used, one summarized relationship between sidescan survey frequency, wavelength and maximum range was published by the Woods Hole Coastal and Marine Science Center (C. and M. C. Woods Hole 2016).

3.3.4.3 Slant Range Correction

The entire raw Image formation from a sidescan system, extracted from the surveyed area, is reached using an acquired acoustic row concatenation. However, this row image is not a correct real space representation, this is a Slant Range (SR) representation or time per time representation ($\Delta t \; x \; \Delta t$) and so called Slant Range Image (SRI). In the Figure 3.14, we can see the strip sample of SRI image, it can be understood as the bottom representation from the sonar point of view, approximates the real space representation and have some distortions.

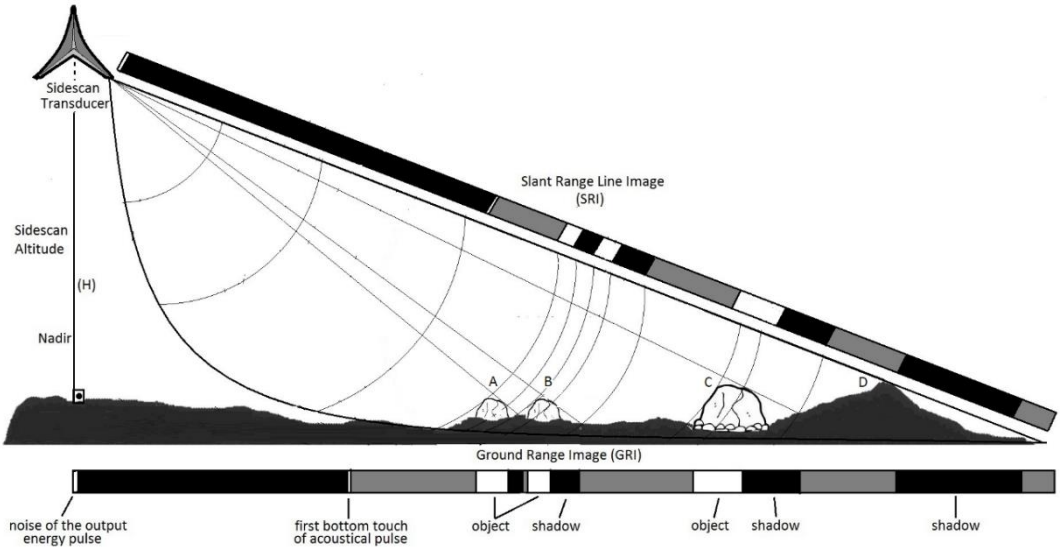


Figure 3.14: *Representation of the geometric differences between a Slant Range and Ground Range Images and how targets are represented in a sidescan acoustic image.*

If we need the ground referenced view to associate with any geographical information system, we need to perform a slant range correction and thus generate a Ground Range Image (GRI). The slant range correction will give the space per space representation ($\Delta s \; x \; \Delta s$), also called GRI and can be georeferenced to be more precise and to have a standard positioning.

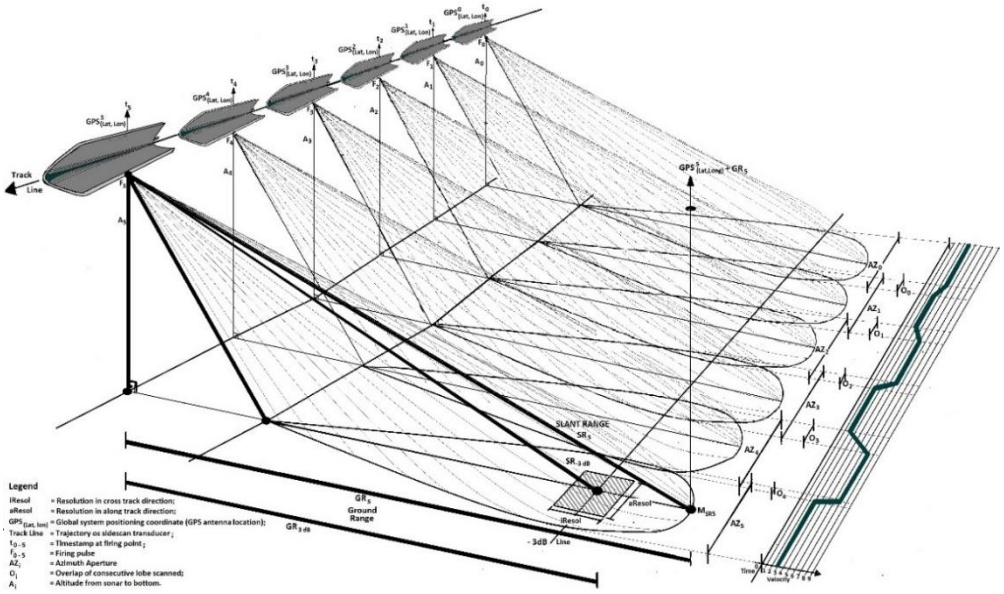


Figure 3.15: *Illustrative sidescan survey track-line, bottom coverage and ping overlapping according survey velocity*

The total distance composed by two sides range is also called swath width. Altitude information is very important on the processing of SR corrections; we use this in the process of conversion from SRI to GRI. In the *Figure 3.15* we can see a simplified illustration of a linear trajectory of a SSS and the representation of beam interaction and

the relationship between GPS, platform velocity and beam slant, ground range and resolutions.

4.0 Spatial Statistics and Regionalized Variable

Statistics will help any research to understand the meaning behind the numerical data universe encapsulated in a cohesive dataset. The significance and interpretation of a numerical data can be correlated with the spatial or temporal distribution, the main objective is to understand the significance behind data and their trends. Several authors have been divided statistics analyses and applications in three main stages: Data Description; Interpretation and Estimation (Cressie 1993) and (Lucieer 2007). A data description will give to us the opportunity to know our dataset universe through exploratory data analysis and hence, we can describe characteristics data from some statistical descriptors as moments, average, deviance, correlation and etc. Using some information from data descriptors we can perform a data interpretation and thus estimations can be done based on mathematical models driven by data knowledge. Several works related to statistical researches were used to develop tools, methodologies and procedures to solve statistical problems in data mining, interpretation, estimation and predictions related to various sample data, however, we gave a special attention to work with a branch of statistics called Spatial Statistics and into this, methods and techniques related to Geostatistics (Cressie 1993), (Hengl et al. 2008) and (Diggle & Ribeiro Jr. 2007).

4.1 Spatial Statistics

Spatial Statistics techniques allow us the opportunity to analyse a spatial dataset, such as, in our case, bathymetry dataset or acoustic images. We consider the spatial distribution and their neighbourhood interaction as the main influences to explain some behaviour driven by a dataset and, also, it is possible to predict their trends based on data correlations. Hence, a study of spatial correlation among data points can be done considering the distance between them. It will be tested the level of correlation to a set of samples into a specific dominium, along a specific direction or around several directions from an certain point or spatial entity (Cressie 1993), (Diggle & Ribeiro Jr. 2007) and (Wackernagel 1995). In a simple way, differences between classical statistical analysis and spatial statistical techniques, have only the insertion of a spatial component as the inflexion point to decision making process and weighting. Thus, spatial statistics have some advantages of georeferenced data samples available for each spatial entity to improve the prediction task to uncover relationships between the spatial objects and phenomena. Some models have been developed to cover a variety of applications, related to this thesis, the main guidance and inspiration are described in the following papers: (Cressie 1993), (Wackernagel 1995), (Diggle & Ribeiro Jr. 2007), (Brenning & Boogaart n.d.) and (Hengl 2007). Geostatistical based approach is used here to classify, quantify, and then, increase the survey resolution to find relevant information occulted into the original dataset, to make a conceptual information map based on concepts of raster map. This map allows the construction of navigations map as a background to perform an autonomous robotic navigation based on track topographical relief objects. Another application is a resolution increase of bathymetric surveys to costal environment using SBES in application, where a multibeam sonar cannot be applied because surveys limitation such as survey vehicle capacity, environmental limitation or costs.

In the geostatistical concepts like a branch of classical statistics, measures quantities defined by moments will play an important role supported by the concept of variance/covariance to analyse the spatial dependence for a regional variable (Cressie 1993). The first and second moments, respectively mean and variance are the base to develop a reasoning path to reach a covariance concept, resulting in the possibility to use a Multiple Linear Regression (MLR), a covariance function approach is used to model a set of samples which represent a domain of a sampled surface. The MLR is a core of a Geostatistical predictor called simple Kriging (Diggle & Ribeiro Jr. 2007) and (Wackernagel 1995). This work has a core fundaments based on Kriging based interpolation to solve the problem of weight average to make a prediction of a value for an unsampled point (unknowing), using information from sampled points spreads in a spatial distribution around the unknowing point. A simple version of Kriging interpolation cannot be applied unrestrictedly to any type and distribution of spatial sampling, some requirements need to be verified before their implementation variable regionalization and trends. Therefore, values of statistical moments and covariance among sampled points, as a method to estimating statistical parameters need to be checked and consider the development of the prediction process with Kriging. The method of statistical moments was developed over a set of e independent random variables.

4.2 Regionalized Variables

According the Matheron thesis (Matheron 1970), “Geostatistics is the application of the theory of regionalized variable of mineral deposit (with all that this implies)”. But, we need to answer a question: What is a Regionalized Variable? In a simple way, Matheron describe that the Regionalized Variables (RV) are described by a phenomenon spreader in a space with the exhibition of a spatial structure (Matheron 1970). A simple function over a spatial distribution, can be considered a RV, but when the variable represents a sample from any natural or environmental domain, it will represent a very irregular function (Matheron 1970) and a simple linear model can fail if applied unrestrictedly over all dataset. It is because the contradictory aspect from theory of RV, represented by the random nature of study variable and the structural aspect imposed by RV theory. Thus, the original theory from (Matheron 1970) aims to solve the problem of estimate a value of RV at the unsampled point.

4.2.1 Regionalized Variable (RV) into the Context of Topographic Relief Elevation Dataset

Considering a real surface represented by "s" extracted from a portion of earth’s surface "S", both represent a continuous surface that the area differs from zero (Wackernagel 1995) and (Diggle & Ribeiro Jr. 2007). Samples points from "s" will generate a set of elevation points samples $z(n)$, where n is the number of samples over all surveyed area represented by (s). Considering "x" as georeferenced coordinate point into "s", a concise and mathematical representation of a generic data location can be described in a 2-dimensional Euclidian space according (Cressie 1993), and is expressed in the *expression 4.1*.

$$s(x) \in \mathbb{R}^2 \qquad \textit{exp. 4.1}$$

Values sampled from region (s) has the survey realization independent from one point to another into this region. Into our practical application to extract a real data, each one sampled point was taken at distinct time in a sequential transect sampling method but in the same survey uninterrupted process. Each sample will generate a potential datum $z(n)$ as an output value, this is a regionalized value as the outcome of some random natural mechanism and this random mechanism could produce different property at each point of a region (Wackernagel 1995). Now, into any georeferenced coordinate system, which coordinates are given by "x" we consider the value $z(x)$ (Elevation level value “Bathymetric Point”), one realization (Measure) in the spatial location $s(x)$ (Location), then, it will vary over an index set Domain $D \subset \mathbb{R}^2$, and the process will generate the multivariate random field so called random process (Cressie 1993), represented by mathematical *expression 4.2*.

$$\{Z(x): s(x) \in D\} \qquad \textit{exp. 4.2}$$

According (Diggle & Ribeiro Jr. 2007), the mathematical *expression 3.26* can be consider a partial realization of a stochastic process into the 2-dimensional Euclidian ($\mathbb{R}^2$) surveyed area. However, the sampling design assumed for locations $s(x)$ has a deterministic fashion, because the extraction data has quasi-equal distance from one sampled to another. Besides that, the survey trajectory is in linear transect method, and if we delimitate a closed subarea or portion into the entire surveyed surface, it seems like a regular grid. Additionally, in the core of sample process we have stochastically intrinsic components into the survey realization, where each sampled realization is uncorrelated to another one, and small variations can be result from dynamics of ocean, survey boat track and noises. In this way, each sample $z(x)$ will compose a set of measures that define a random variable of a stochastic process $Z(x)$. With this, D can be consider as a random set domain, according (Cressie 1993) it means D “is a measurable mapping from a probability space onto a measure space of subsets of $\mathbb{R}^d$”, in our case $d = 2$.

The bathymetric dataset from a SBES device that was used in this work, is a type of survey realization to extract discrete elevation data samples, it is able to meet all requirements of a regionalized variable according definitions given by (Matheron 1971b), (Wackernagel 1995), (Diggle & Ribeiro Jr. 2007) and (Cressie 1993). We apply a variogram geostatistical analyse to probe the assumption of our dataset in a priori context, complies all requirements and constraints to be a regionalized variable field.

A proposed geostatistical model needs to be met in the first approach two main constraints (**I**) and (**II**), but internally to definition of sampling survey we include a third requirement (**III**):

I. **The stationarity of RV –** This is relating to a trend in data, and will explicit the need that some characteristics of RV remain the same when shifting a given set of n sampled points from one region $s(x_b)$ to another $s(x_a)$ into the same survey domain $\{\boldsymbol{Z}(\boldsymbol{x}): \boldsymbol{s}(\boldsymbol{x}) \in \boldsymbol{D}\}$**;**

II. **Gaussian Process –** The samples values should have a Gaussian behaviour, then, statistical algorithms and methods can model their characteristics;

III. **Mutually Independence of Sampling Realization –** Data extracted from survey realization need to be uncorrelated from one generic point $z(x_a)$ to another $z(x_b)$, each sample are an independent event. The mathematical model expression to obey this assumption is given by $\boldsymbol{p}(z(x_a) \cap z(x_b) = \boldsymbol{p}\big(z(x_a)\big).\boldsymbol{p}\big(z(x_b)\big)$, where $\boldsymbol{p}(z(x_n))$ represent the probability of event $z(x_n)$ and represent a generic event or sampling process.

Overall, if the dataset meets the all requirements from I to III, in first approximation, we can apply a geostatistical approach to analyse sampled data and predict values at unsampled points.

4.2.2 Basic Statistical Support

In a concise way, we describe briefly some characteristics about statistical parameters emerging from a set of sampled values that demonstrate both a sufficiently strong central tendency to cluster it in a domain of regionalized variable and the weighted covariance value that correlates values between them. Regardless of, relationships among subsets of regionalized variables, those statistical parameters are proposed and used as decision make prediction factor.

Considering a set of discrete random variables $Z(x)$, this type of variable is the most common type of stochastic model, then we will refer here to the process represented by our model as a stochastic process. Into the context of our dataset the process will be defined by a Cumulative Distribution Function (CDF) as a Gaussian Random Variable (GRV) behaviour, with this we can map elements from the sample using the space probability law (Nisbet et al. 2009) and (Ross 2009).

Let $Z(x)$, defined here as a random variable, each element is mapped by $Z(x)$ into "n" values. Then, we can partition the sample surface "s" into "n" disjoint subsets $(s_1, s_2, \dots, s_n)$, having the probability mass function $p(x) = P(Z(x) = s_n)$ for any value of $s_n \in s$ is represented here by its CDF in the equation 4.3.

$$F\big(x_j\big) = \sum_{n \le j} p(s_n) \qquad \textit{eq. 4.1}$$

Another definition from our dataset comes from the assumption of stationary Gaussian model.

The random variable $\{Z(x): x \in \mathbb{R}^2\}$, represent an instance of one Regionalized Variable (RV), and it is a Gaussian process with mean μ, variance $\sigma^2 = Var\{Z(x)\}$ with correlation function $\rho(u) = Corr\{Z(x), Z(x')\}, where\ u = \|x - x'\|$ and $\|\cdot\|$ denotes distance.

Observing that the random variable is conditional on $\{Z(x): x \in \mathbb{R}^2\}$, where each individual measure $z(i)$ represent a mutually independent random variable, normally distributed with conditional means $E[Z_i|Z(\cdot)] = S(n)$ and conditional variance τ^2.

4.2.2.1 Variance

Let $Z(x)$, in the context of the RV, we denote the variance $Var(Z(x))$, the quantity expressed by:

$$\sigma^2 = Var\big(Z(x)\big) = E([Z(x) - E(Z(x)]^2) \qquad \textit{eq. 4.2}$$

The calculous of variance will give us the simple measures of the spread, or variability, of the distribution.

Obs.: The variance defined here don't obey the scale properties, we only work with SI unit, where meter is the standard unit to represent length and distance measures.

4.2.2.2 Standard Deviation

The Standard Deviation (SD) can be used to remedy the scaling problem from variance approach and mathematically is described by *equation 4.3*.

$$\sigma\big(Z(x)\big) = \sqrt{Var\big(Z(x)\big)} \qquad \textit{eq. 4.3}$$

4.2.2.3 Covariance

The covariance between two samples into the concept of the random variable will compute the dependence between them, and will be described by the *equation 4.9*.

$$Cov(z(x_a), z(x_b)) = E[(z(x_a) - E(z(x_a)) \cdot (z(x_b) - E(z(x_b))] \qquad \textit{eq. 4.4}$$

Obs.: In the same way of the variance definition, the covariance defined here do not obey the scale properties, to overcome this problem in application, we need to apply the concept of correlation.

5.0 Basic Geostatistical Regression Support

Geostatistical is a branch of spatial statistics data processing and has, as fundamentals aims, the numerical study to determine the correlation among measured samples, it will give us the notion of similarity coupled to a distance in one specific direction or in omnidirectional form. Then, the way to perform geostatistical analysis and estimations starts when are applied the theory of Regionalized Variables to the problem of predicting spatial phenomena (Matheron 1962) and (Matheron 1970). The study starts with a graphical analysis of spatial distribution and values of data samples using the so called Variogram or Semivariogram function ($\gamma(h)$), in this thesis, I will use the semivariogram. The semivariogram function is the graphical tool to describe the spatial correlation of a phenomenon, and will give us the correlation function among samples measured into the measured area according the distance lag (h). The *Figure 5.1* represent some samples from the surface S.

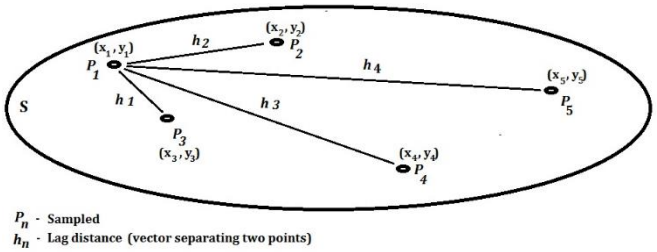


Figure 5.1: *Illustration of Lag distance among some points*

To construct the semi-variogram graphical representation, we compare the sample values against distance in space using the *equation 5.1*, and each value found will be plot in the semi-variogram graphic represented in the *Figure 5.2*. After the first analysis using a geostatistical process to plot the empirical semivariogram, we can predict values at unsumpled locations using a predictor, or a geostatistical interpolation method called Kriging. Thus, inserting information (parameters) from semivariogram analysis we can extract parameters values obtained using the graphical representation after fit the semivariogram cloud using any adequate theoretical semivariogram model, in order to perform estimations (Chiies & Chauvet 1974) and (Matheron et al. 1987).

$$\gamma(h) = \frac{1}{N}\sum_{\alpha=1}^{N}(z(x+h) - z(x))^2 \qquad \textit{eq. 5.1}$$

Where

$\gamma(h)$ – Is the semivariogram value;
$z(x)$ and $z(x+h)$ – Represent the measure of RV at origin and at distance h from origin;
N – Represent the number of measures into the survey sampling.

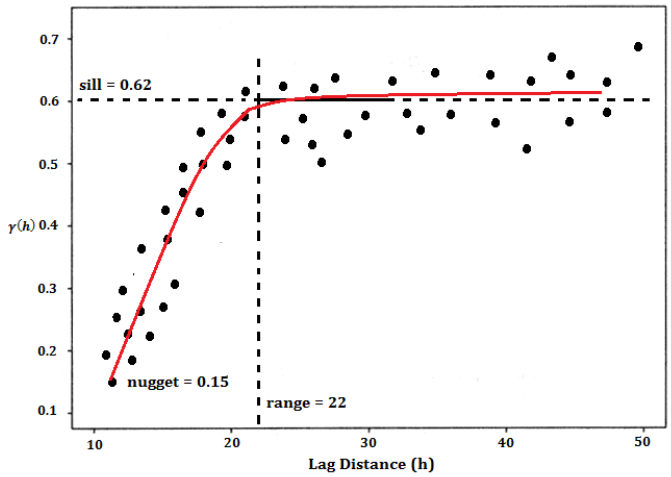


Figure 5.2: *Semi-variance cloud fitted*

The experimental variogram is usually computed using $\boldsymbol{h}$ vectors wit length inferior to half of interest region diameter (Diggle & Ribeiro Jr. 2007) and (Wackernagel 1995), it will prevent the use of samples near the survey border region in estimations at the centre of region that will be not representative of the hole dataset. To calculate the experimental semivariogram, ones ideally need to have any knowledge about the RV phenomena under survey, and it will give some definition about the phenomena to define an isotropic or anisotropic behaviour. Hence, we can calculate the semivariogram function adequately for each specific direction or considering all direction. Usually we choose four directions to analyse: E-W, N-S, NE-SW and NW-SE. Considering now, the simplified graphical representation of experimental semivariogram in the *Figure 5.3*, to extract parameters of interest we need to fit by one of the formals theoretical semivariogram models. The use of theoretical semivariogram model allow us the opportunity to apply any adequate regression method based on this model.

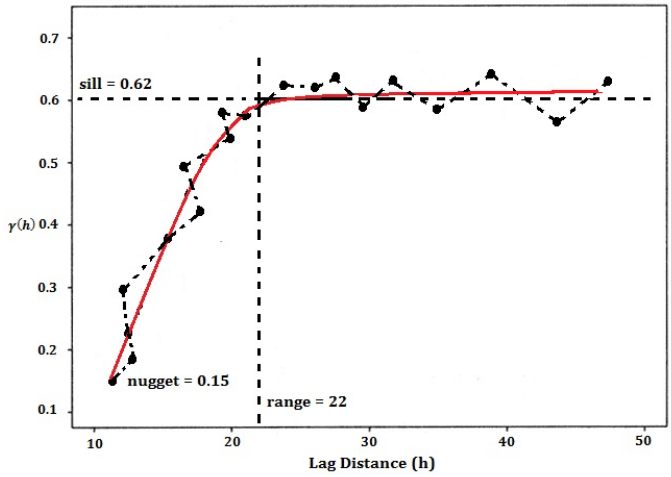


Figure 5.3: *Simplified Experimental semivariogram fitted*

From the theoretic model we can extract the three parameters called, the nugget effect value, range and sill, all before are described in context by (Diggle & Ribeiro Jr. 2007), (Clark 1979), (Matheron, G. Kleingeld 1987) and (Wackernagel 1995) as:

- **Nugget effect:** Represent the semivariogram value at lag distance equal to zero "0", this value represent variability at small scales, with this, the discontinuity at the origin can be identified. The correct model definition will give to the kriging process the possibility to avoid numerical predictions instability. The nugget effect characterizes the variogram shape near the origin;

- **Range:** The lag distance at the variogram model fit becomes constant, define the existence of autocorrelation value among sample measures, and this parameter has the lag value equal to the first

flattens out. In the geostatistical model theory, next the lag distance equal to range, the autocorrelation is 0;

- **Sill:** The semivariogram value that attains at the range value, can be equal to the data variance.

The theoretical semivariogram model is defined under two assumption related to the random functions, but an intrinsically stationary random function does not need to have a constant mean or a constant variance:

- The mean is invariant for any sampling location into the surveyed region (ideally must not have any drift);

- The variance of increments in $\boldsymbol{h}$ need to has a finite value and equal to $2\boldsymbol{\gamma}(\boldsymbol{h})$, and will depend only on the length and the orientation of $\boldsymbol{h}$ vector.

Mathematically we can define these two assumption using *equations 5.2* and *5.3* that describe respectively the expectation and variance (Wackernagel 1995).

$$[z(x+h) - z(x)] = m(h) = 0 \qquad \textit{eq. 5.2}$$

$$var[z(x+h) - z(x)] = 2\gamma(h) \qquad \textit{eq. 5.3}$$

According the semivariogram behaviour we can estimate the smoothness of the RV phenomena from the theoretical variogram model fitting over the sampling realization. In a concise way, it can be described by three types smoothness indicators (Wackernagel 1995), described by the *Table 5.1*.

Table 5.1: *Smoothness indicator and their relationship with regionalized variable aspect*

Properties of Semivariogram Cloud	Behaviour of $\boldsymbol{\gamma}(\boldsymbol{h})$ at origin	Regionalized Variable Aspect	Observation
$\boldsymbol{\gamma}(\mathbf{0}) = \mathbf{0}$	Differentiable	smooth	
$\boldsymbol{\gamma}(\boldsymbol{h}) \geq \mathbf{0}$	Continuous but not differentiable	rough	
$\boldsymbol{\gamma}(\boldsymbol{h}) = \boldsymbol{\gamma}(-\boldsymbol{h})$	Discontinuous	speckled	Nugget effect – RV changes abruptly at small scale

The geostatistical method for data analysis and prediction was chosen here due the increased information from a bathymetric survey, mainly because their characteristics of recognition the spatial variability applied to complex surface as topographic relief. The geomorphologic change in some way has a random profile and there are often complexities to be model by a simply linear regression method. In some mathematical approaches specifics geomorphologic attributes can be smoothed by some regression algorithms. The geostatistics analysis developed by George Matheron and improved by Kriging estimator become one powerful regression tool to estimate analysis values using the RV theory. Originally developed to ore mineral deposit estimation and then it was used as a general-purpose estimation if the sampled studied that obeys the properties of RV. Remember, regionalized variables try to characterize physical phenomenon spread over a space as a random process defined by the spatial structure (Matheron, 1965).

This will imply in a contradictory or complementary relationship, described by (Matheron, 1971), in his work Matheron has two main purposes to apply a geostatistical analysis and estimation:

- Express structural properties in an adequate form; and
- Solve the problem of regionalized variable estimation from fragmentary sampling data.

The previous purposes are defined for samples that falls into the same region and the error estimation will depend on the structural characteristics (Matheron, 1971). Geostatistical approaches assume that the samples measured are independent, has a spatial correlation and has no trend.

If we have a system composed by two variables Z_1 and Z_2, we can represent their relationship in a scatter diagram, we will have a cloud of points in a plane represented by the Z_1 and Z_2 axis according *Figure 5.4*. According (Wackernagel 1995), considering the centre of mass from data points cloud defined by the two means points m_1 and m_2.

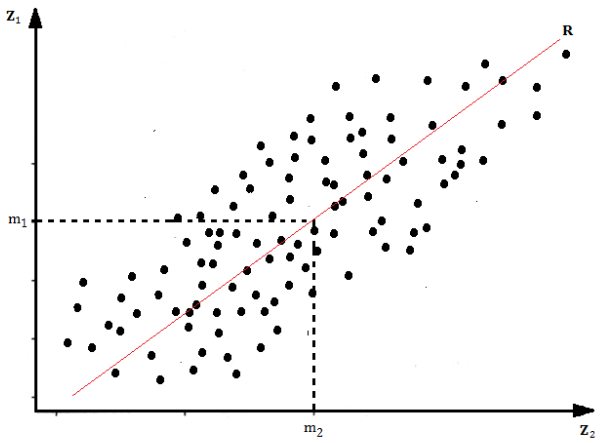


Figure 5.4: *Scatter Plot of sample Cloud and their Relationship between variables measured*

The data cloud dispersion around this centre of mass is obtained multiplying the difference between values of one variable and its means (residual). Then, the covariance is defined as the average of this products, and modelled by *equation 5.4* (Diggle & Ribeiro Jr. 2007) and (Wackernagel 1995).

$$cov(Z_1, Z_2) = \frac{1}{n}\sum_{\alpha=1}^{n} (z_{1\,Residual})(z_{2\,Residual}) = \qquad \textit{eq. 5.4}$$

$$= \frac{1}{n}\sum_{\alpha=1}^{n} (z_1^{\alpha} - m_1)(z_2^{\alpha} - m_2)$$

Covariance values will give the information about two variables relationship into a cloud of measures. This will be done extracting the level of correlations or linking strength and their behaviour, negative or positive correlation. Because this definition, often, the covariance will be performed using pairs of variables. Concluding, covariance function will describe the relationship between variances from pairs of measures related with their distance h. One relationship can be done between semivariogram and covariance function using the *equation 5.5*.

$$\gamma(h) = cov(0) - cov(h) \qquad \textit{eq. 5.5}$$

Where

$\gamma(h)$ - Is the semivariogram values between pairs of measures located at 0 and h distance;
$cov(0)$ - Represent the covariance of data sample;
$cov(h)$ - Represent the covariance at distance h.

Ideally, covariance function requires the stationarity assumption, as in case of some variogram function applications, but in the case of semivariogram the assumption of stationarity does not need to have the assumption of stationarity in the data, represented by a constant mean and variance (Wackernagel 1995)(Matheron 1962).

There are a set of theoretical variogram used to fit an experimental variogram realization, here we can see four representative models of these from expression 5.6 to 5.9.

Linear Model:

$$\gamma(h) = \begin{cases} 0, & if\ |h| = 0 \\ c_0 + ph, & if\ |h| \neq 0 \end{cases} \qquad \textit{eq. 5.6}$$

Exponential Model:

$$\gamma(h) = \begin{cases} 0, & if\ |h| = 0 \\ c_0 + c[1 - exp(-3|h|/a)], & if\ |h| \neq 0 \end{cases} \quad \textit{eq. 5.7}$$

Spherical Model:

$$\gamma(h) = \begin{cases} 0, & if\ |h| = 0 \\ c[1.5(|h|/a - 0.5(|h|/a)^3], & if\ 0 < |h| \leq a \\ c_0 + c, & if\ |h| \geq a \end{cases} \quad \textit{eq. 5.8}$$

Gaussian Model:

$$\gamma(h) = \begin{cases} 0, & if\ |h| = 0 \\ c_0 + c[1 - exp(-3(h/a)^2)], & if\ |h| \neq 0 \end{cases} \quad \textit{eq. 5.9}$$

Where:

- p - Represent the slop of linear model;
- c_0 - Is the nugget value;
- c - Is the sill;
- a - Represent the range.

The previous models are all isotropic, it means, the spatial correlations have the same behaviours in all directions, but in many cases of spatial data we might expect spatial correlation in one direction but not another, anisotropic case, and thus we need to apply another approach. One alternative is to apply the geometric anisotropy technique followed by the linear transformation (Cressie 1993).

5.1 Modelling Spatial Continuity Using Geostatistics

The Spatial continuity, from some surveyed surface can be inferred using an adequate set of data sampling submitted to a geostatistical process. Considering this geostatistical inference system, empirical semivariogram cloud acts as a support to estimates values at unsampled locations using kriging predictor algorithm. In this work, I use a set of discrete elevation samples from a surface S extracted using SBES device. The S region represent a portion of a random domain D. To validate predictions, another set of sample from a MBES device were extracted in the same region to be considered as the merit figure to compare and determine the quality of geostatistical method to estimate unsampled points in bathymetric dataset. In this way, the domain D can be considered as a random set domain, according (Cressie 1993) it means D "is a measurable mapping from a probability space onto a measure space of subsets of $\mathbb{R}^d$". Each sample $z(x)$ will compose a set of measures that define a random variable of a stochastic process $Z(x)$ to extract discrete elevation data samples. It can meet all requirements of a RV. Thus, we consider each discrete sampling as x and measures, $(x + h)$, another realization separated by the lag distance h as a RV, and applying the function modelled by the mathematical expression 46 an empirical semivariogram (ES) cloud will be generated. The ES will need to be fitted using any Theoretical Semivariogram (TS) assuming a specific mathematical model behaviour. Aspect and behaviour of semivariogram slop will indicate the spatial structure of RV into the data sampled, including their changes. It can be summarized in the *Table 5.2*.

Table 5.2: *Sampling Dissimilarity and Spatial Structure Relationship*

Semivariogram Average Dissimilarity Between Sampling Data Separated by lag Distance h	**Spatial Structure Characteristics in the Data**
Constant	There is no spatial structure
Non-zero slop near the origin	There are some spatial structures
Abrupt change in slop	Passage to a different structuration

To perform the prediction process model, and estimate spatial continuity geostatistical analysis, we start with an empirical variogram fitted by one theoretical semivariogram model, and in some cases, supported by one type of ancillary information. These assumptions are often assumed in geostatistical literature, and ancillary information

is extracted from the physical knowledge of the study phenomena including any robust method to evaluate anisotropy directions, ratio and range.

The key question is: how is the best set of ancillary information or just the unique auxiliary information to perform a prediction task?

The modelling process involving geostatistics, can be done using two procedures described in (Goovaerts 1997), the first is called Black-Box (BB), it involves a computational automatic choice and fitting the experimental semivariogram, and the second is Semi-Automatic (SA) process, that sill and range can be chosen by user (Goovaerts 1997). According (Goovaerts 1997), The BB estimation must be avoided because it cannot consider ancillary information. However, if we have an experimental semivariogram well fitted, the user decision can be supported by any efficient computation tool to get better parameters, and additionally backed by experimental data or ancillary information (Goovaerts 1997). Another basic structural information to support the prediction task is related to the correct definition of better semivariogram model, we have two choices, that is, isotropic or anisotropic. It will be done comparing the experimental variogram in some directions. Usually, at least as reference directions, we consider four basics cardinal directions as South-North (SN), West-East (WE), North-East (NE), and North-West (NW). Often, when computation direction is considered, we determine an angle of tolerance to input in the semivariogram computation.

The spatial structure and values of parameters extracted from variogram analysis will guide the geostatistical estimation process using one interpolation method. The official and most used geostatistical method to predict unsampled values is based on Kriging algorithm (Chiies & Chauvet 1974) and (Hengl 2007). The Kriging algorithm will consider the correlation from sampled values extracted around an unsampled points, according a model based mainly in the spatial distribution of samples analysed by geostatistical Variogram analysis. Kriging is the direct application of regression analysis proposed by D.G. Krige to solve the problem of estimate of ore reserve using sparse sampling block of ore. It is correlated to some others elements spatially distributed around them (Matheron, G. Kleingeld 1987), (Matheron 1971b), (Gumiaux et al. 2003) and (Diggle et al. 2002). Despite the existence of others interpolation methods to estimate unsampled values, Kriging estimators has the advantage to prevent the problem related to optimal parameter estimation, because it uses in their interpolation process, parameters based on preview spatial structure analysis. Then, we can use definitions from (Cressie 1993), (Diggle & Ribeiro Jr. 2007) and (Wackernagel 1995) as the fundamental bases to investigate the process of spatial continuity modelling and their confidence level to construct an underwater mapping representation based on discrete samples spread in the region of interest.

5.2 Regression Using Kriging

The problem of estimating values in a continuous surface, that uses information from discrete sample, is the main matter of this part of the thesis. In this research work we explore some structural spatial dependence in discrete samples realization to infer structural characteristics of the surface and thus, predict values at unsampled regions. The work was improved using a type of linear regression called Kriging supported by geostatistical spatial statistical analysis. The focus here was the study of two simple and common Kriging algorithms, their advantages in solve the prediction problem related to continuous surface estimation using discrete bathymetric sampled from a SBES sonar without any ancillary information contrasting it with the common assumption of many geostatistical bibliographies that relates the importance of the use ancillary information to reach a satisfactory result in geostatistical prediction process as is the case of Kriging.

Considering a Kriging based application over a set of random variables, described here as a regionalize variable $z(u_\alpha)$, where $i = (1, \dots, i)$ is the number of sample points, and if this random variable is a subset of a second-order stationary random function Z. The regionalized variable $z(u_\alpha)$ will be the reference base sampling to perform a regression analysis, and because this, it will be called "regressors" on a "regressand" $z(x_u)$ where $u = (1, \dots, n)$ is the number of sampling points that are need a prediction (Wackernagel 1995). With this assumption, we gave two main relationships between sample distribution into the specific domain according equations 46, 49, and 50 (Wackernagel 1995). The basic problem is to find a type of weight average to make estimations at unsampled point. Geostatistical and kriging methods take a knowledge of covariance among random variables and their structure distribution.

All Kriging application has a base algorithm designed from a linear regression application defined as *equation 5.10*, following the notation described in (Goovaerts 1997).

$$Z^*(u) - \mu(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)[Z(u_\alpha) - \mu(u_\alpha)] \quad \textit{eq. 5.10}$$

Where:

$Z(u)$ – Predicted value;
$\mu(u)$ –Expected values of RV represented by $Z(u)$;
$\lambda_\alpha(u)$ – Weight assigned to the sample;
$z(u_\alpha)$ – Sample;
$\mu(u_\alpha)$ – Expected values of RV represented by $Z(u_\alpha)$.

Estimations error and predictions can be done using the relationship $Z(u) - Z(u_\alpha)$, and the common objective of kriging algorithm variations has the same constraints, minimize the estimation error variance $\sigma_E^2(u)$, considering the unbiasedness characteristic of kriging estimator, represented by *equation 5.11*.

$$\sigma_E^2(u) = Var\{Z^*(u) - Z(u)\} \quad \textit{eq. 5.11}$$

And their minimization assumption given by the *equation 5.12*.

$$E\{Z^*(u) - Z(u)\} = 0 \quad \textit{eq. 5.12}$$

The Kriging model based estimation is sensible to the adopted model for the random function $Z(u)$ (Goovaerts 1997). The random function $Z(u)$ can be decomposed in two independent terms.

$$Z(u) = R(u) + \mu(u) \quad \textit{eq. 5.13}$$

The residual component is modelled as a stationary random function with zero mean and covariance $C_R(u)$ according *expression 5.1* and *equation 5.14*.

$$E\{R(u)\} = 0 \quad \textit{exp. 5.1}$$

$$Cov\{R(u), R(u+h)\} = E\{R(u) \cdot R(u+h)\} = C_R(u) \quad \textit{eq. 5.14}$$

The expected values of RV at location u, *expression 5.2*, is thus the value of the trend component at the location.

$$E\{Z(u)\} = \mu(u) \quad \textit{exp. 5.2}$$

Just using the bases of Kriging estimator theory, we can explicit three main types or variants based on distinctive approach for trend $m(u)$, as follows in *equations 5.15*, *5.16*, and *5.17*.

5.2.1 Simple Kriging (SK)

Simple Kriging (SK) algorithm considers the knowledge about mean $m(u)$, and it is constant throughout the study area S:

$$\mu(u) = \mu, \text{ known } \forall\, u \in S \quad \textit{eq. 5.15}$$

5.2.2 Ordinary Kriging (OK)

Ordinary Kriging (OK) accounts for local fluctuations of the mean by limiting the domain of stationarity of the mean to the local neighbourhood $W(u)$:

$$m(u) = constant\ but\ unknown\ \ \forall\, u' \in W(u) \qquad \textit{eq. 5.16}$$

Unlike SK, here the mean is deemed unknown.

5.2.3 Kriging with Trend (KT)

Kriging with trend (KT), like OK, it considers the unknown local mean $m(u')$. The mean can vary smoothly for each part of region or neighbourhood $W(u)$, hence over the entire study area S. The trend component is modelled as a linear combination of functions $f_k(u)$ of the coordinates:

$$m(u') = \sum_{k=0}^{K} a_k(u') f_k(u') \qquad \textit{eq. 5.17}$$

With $a_k(u') \approx a_k$ constant but unknown $\forall\, u' \in W(u)$

The coefficients $a_k(u')$ are unknown and deemed constant within each local neighbourhood $W(u)$. By convention, $f_0(u') = 1$, hence the case $K = 0$ is equivalent to OK (constant but unknown mean a_0).

6.0 Bathymetric Surveys: Data Uncertainty and Their Requirements

The knowledge about the uncertainty in bathymetric elevation measures is necessary in order to monitor the performance and system reliability, despite the criteria used, here is specified by hydrographic surveys manuals (IHO 2008), (IHO 2012), (Bowditch 2002) and (Surveyors, International Federation Greenland & Higgins 2006). Some factors can influence the bathymetric data accuracy and as cited in the Standard (IHO 2008), for shallow water surveys and considering a datum error irrelevant , due the fact that the distance to geodetic station reference is too small, the set of main factors related to vertical and horizontal uncertainty is described below:

I. Vertical:

- Vertical positioning system errors;
- Tidal measurement errors;
- Instrument errors;
- Vessel motion errors;
- Seabed slope; and
- Time synchronisation and latency.

I. Horizontal:

- Positioning system errors;
- Range and beam errors;
- The error in vessel heading;
- System pointing errors resulting from transducer misalignment;
- Sensor location;
- Vessel motion sensor errors;
- Sensor position offset errors;
- Time synchronisation and latency.

Survey errors comes from a variety of factors, these contributing elements can be combined by applying the Law of Propagation of Random Variances, resulting in estimations of Total Propagated Uncertainty (TPU) for both the vertical (depth) and horizontal position (Positioning). It will include two of main uncertainty elements, one is related to systematic errors and other is the random component due to noise in the survey measure process (IHO 2008). In a basic definition, the uncertainty relates the range of values or interval that will contain the true value measured at the required confidence level. According to the Fifth Edition of the IHO Standards (IHO 2008) , the total error in depth measurements should not exceed the confidence level of 95%, this means 1.96 x Standard Deviation (STD). If surveyed data has a normal distribution of errors with a probability of 90%, 0.3 meters for depths less than 30 meters or 1% of depths greater than 30 meters. To hydrographic surveys operations the IHO standard (IHO 2008), classify a surveyed area according the needed or requirements for security navigations and maximum depth to sampling in four types, named order as follow:

- **Special Order:** Areas where under-keel clearance is critical, generally less than 40 metres;
- **Order 1A:** Areas shallower than 100 metres where under-keel clearance is less critical but features of concern to surface;
- **Order 1B:** Areas shallower than 100 metres where under-keel clearance is not considered to be an issue for the type of surface shipping transit the area;
- **Order 2:** Areas generally deeper than 100 metres where a general description of the sea floor is considered adequate. shipping may exist.

In this thesis, we will consider as the minimum standard for quality control to the proposed process of geostatistical estimation, the standard related to the Special Order (Ideal) to 1A (Acceptable). The Special order generally relates to shallow water depth less than 40 meters where under-keel clearance (distance from boat hull o bottom) is critical. For 1A Order, about shallow water areas that depth is below 100 meters, under-keel clearance (distance from boat

hull o bottom) is critical (Special Order) or less critical but features of concern to surface shipping may exist (Order 1A).

6.1 Vertical Uncertainty

A simple standard calibration technique was performed to a Single Beam Echo-Sounder device (SBES) to eliminate systematic errors occurrence from survey instruments before acquisition of data. A simple "Bar Check" method was applied for calibrating both Multi Beam devices (MBES) and Single Beam devices (SBES) (CHS 2008). The procedure involves a piece of metal with conical format or plate lowered to a maximum depth reference. The Figure 6.1 illustrates the physical disposition of devices and support elements to the process of bar check calibration using a port structure. The process involves the installation of a fixed ruler bar and the reference plate next to a pier of port structure, installing and positioning the transducer in a straight line down to the plate, and after that starts the bar check calibration. In the calibration process, the acoustical pulse will propagate directly to the plate surface and after receive the backscatter energy from it, we can measure the distance between them. Recording this measure, called Water Depth from Transducer to Plate (WDT) and knowing the bar ruler measure by manual reading, called here Measure in Ruler (MR), we can compare these measures to extract an Index of Error (IE) for measures by the transducer, and we can correct it in the data extraction or store process, or later when the data will be conditioned or interpreted.

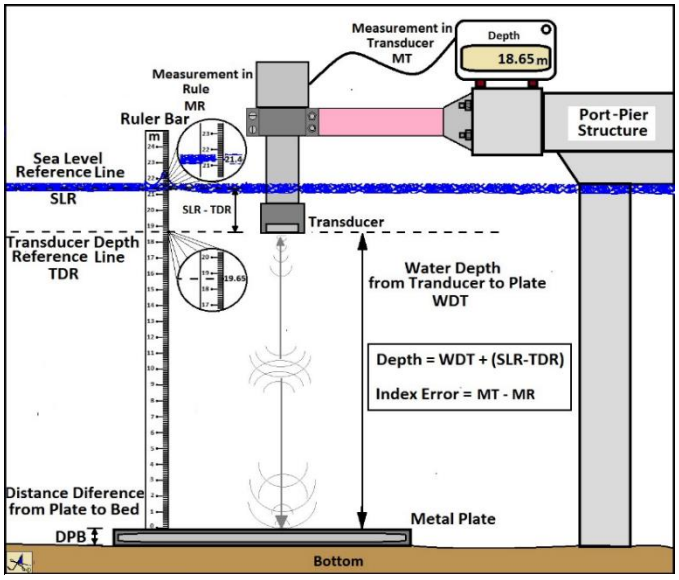


Figure 6.1: *Illustration of an underwater echo sounder bar check calibration using a port structure support.*

If environmental conditions and support structures enable the bar check process with the transducer installed in the boat survey, this will be the best choice because we can extract the level references and off-set displacements among sensors at the same time. To compute the confidence level we use the standard definition published in (IHO 2008) to reach the maximum allowable Total Vertical Uncertainty (TVU) we can use the *equation 6.1* published in the same report.

$$\pm\sqrt[2]{a^2 + (b \times d)^2} \qquad \textit{eq. 6.1}$$

Where parameters $\boldsymbol{a}$ is the uncertainty that does not vary with depth, also called, depth independent error (the sum of all constant error) and $\boldsymbol{b}$ is the uncertainty that varies with depth, called the factor dependent error, while $\boldsymbol{d}$ is the depth. The term resulting from $\boldsymbol{b} \times \boldsymbol{d}$ is called the depth dependent error and represent the sum of all depth dependent errors (IHO 2008). The *Table 6.1* specifies the value of parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ applied to each survey order. We can use the *equation 6.1*, to find the value of TVU, considering the survey depth of data acquire. The equation will compute the 95% of confidence level to depth measurements.

Table 6.1: *Parameters Values According Survey Order*

	Special Order	**Order 1A**	**Order 1B**	**Order 2**
a	0.25	0.5	1	1
b	0.0075	0.0013	0.023	0.023

In the *Table 6.2*, we demonstrate results of calculous to compute the confidence level using the equation 65 to depth values from 5 to 30 meters. Our work dataset depth reached between 20 and 30 meters.

Table 6.2: *Confidence Level for Depths Between 5 and 30 meters for all orders of hydrographic Surveys*

Depth (m)	**Special Order**	**Order 1A**	**Order 1B**	**Order 2**
5 m	0.25 m	0.50 m	1.01 m	1.01 m
10m	0.26 m	0.50 m	1.03 m	1.03 m
15m	0.34 m	0.50 m	1.06 m	1.06 m
20 m	0.27 m	0.50 m	1.10 m	1.10 m
25 m	0.29 m	0.50 m	1.00 m	1.10 m
30 m	0.31 m	0.50 m	1.21 m	1.21 m

As we can see in the section 3.3.3, the depth measure of bottom also called bathymetry can be carry out using the *equation 3.16*, considering that a local tide gauge measures are available. All components of the equation 66 have not relationship among them, resulting in independent terms, then, knowing the uncertainty of each individual variable, we can apply the method of propagation of uncertainty based in the *equation 6.2* (Lockhart et al. 2001) and (Minami 2013).

$$\sigma_S^2 = \sigma_D^2 + \sigma_{Tx}^2 - \sigma_T^2 \qquad \textit{eq. 6.2}$$

Where each individual uncertainty is represented by sigma (σ) subscript the equation 27 term for each component, it is represented in the *equation 6.2*. This error model is related to prior estimations, posteriori error estimations models differ from it and the most usual method is related to both Root Mean Square Error (RMSE) for a set of measures or population and the standard deviation applied to a single data sample (Fisher & Tate 2006). The RMSE is a general method applied to compare a sample population with another reference dataset, usually used to compare a set of predicted values with a reference base. To apply the RMSE analysis we need to define the accuracy (random and systematic components) and precision (random component) (Willumsen et al. 2007), (Byrnes et al. 2002) and (Umbach 1976). In general, a degree of randomness in samples is much more acceptable than in systematic fashion. Thus, we can resume the relationship between systematics (biases) and non-systematics uncertainties (random), to calculate the statistical error associated to depth observations to calculous the TVU related to vertical measurements in the following *equation 6.3* (Byrnes et al. 2002) and (Wolberg 2006).

$$\sigma_{Total} = \sqrt{\sum \sigma^2_{Normal\ Distribuited}} + Systematic\ Errors \qquad \textit{eq. 6.3}$$

To confidence level required for (IHO 2008), we can apply the *equation 6.4* relationship.

$$CL = 1.96 * \sigma_{Total} \qquad \textit{eq. 6.4}$$

Where CL represent the confidence level.

6.2 Horizontal Uncertainty

Horizontal uncertainties are related to the uncertainty at the position of the sounding device (transducer) and hence, the horizontal position of the sampled point within the geodetic reference frame. According the standard for hydrographic published surveys (IHO 2008), the Total Horizontal Uncertainty (THU) for survey dataset will have confidence level applied to total vertical, that is, 95 percent. At the same way applied to vertical uncertainty case, there are four (4) groups for hydrographic survey to classify a horizontal positioning orientation used to navigation and topography aids in signalization purpose. The *Table 6.3* is a partial transcription of Table 1 from (IHO 2008).

Table 6.3: *Minimum Standards for Hydrographic Surveys Related to Positioning*

Description of Objective	**Special Order**	**Order 1A**	**Order 1B**	**Order 2**
Positioning of fixed aids to navigation and topography significant to navigation. (95% Confidence level)	2	2	2	5
Positioning of the Coastline and topography less significant to navigation. (95% Confidence level)	10	20	20	20

In hydrographic surveys uses SBES, the device is generally installed in the survey boat hull, some cases it is set in the steam boat or, in other cases installed in the boat hull edge using a mechanical support to adapt the transducer body to boat hull fixed point. In this case, we can consider the system composed by the SBES and Boat a rigid body, the position of sounding device or SBES is the same position of the survey boat considering plus/minus an off-set. The off-set results in the distance separation from the antenna of the boat geo-positioning system in our case a Differentia Global Positioning System (DGPS) device and the SBES. The *Figure 6.2* illustrates the relationship between positioning system receptor (GPS Receptor) and the transducer device SBES.

The uncertainty in horizontal positioning for extracted data from SBES can be considered the same of vessel positioning system, and approximated by the THU approach. Some standard of bathymetric survey recommends maintaining the horizontal positioning precision in deliverable surveys reports into decimetres, at least, but in some cases it can reach few meters (NOS 2013) and (IHO 2008). Nowadays with advances in the GPS systems, a mainly in the DGPS approach to use Real Time Kinematic (RTK) techniques, we can give a horizontal positioning with order of centimetres accuracy. The reference of horizontal and vertical positioning of the GPS system is based on MSL of chart datum.

The GPS system has two types of services, the standard Positioning Service (SPS) and the Precise Positioning Service (PPS). The PPS service provide the GPS standard codes and signal in a broadcast mode by all GPS constellation of Navstar satellites in L1 frequency (USA 2008). This service is available for peaceful civil, commercial, and scientific use. The PPS service provides by way of authorized access in broadcast at the GPS L1 and L2 frequencies. In the case of PPS service there are one precision (P) code ranging that is reserved for authorized use. This P code is encrypted and altered to become the Y-code. The Y-code will be available only for usuries that has a valid cryptographic keys(USA 2007).

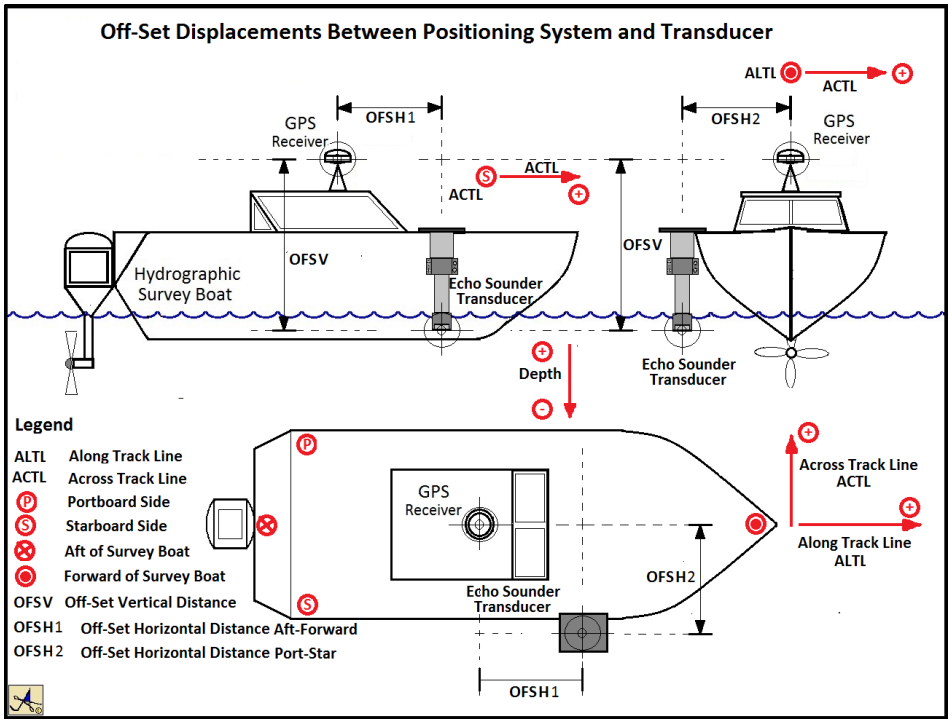


Figure 6.2: *Hydrographic survey boat scheme to SBES data extractions a correlation among off-sets*

In this work, we use a special service of Navistar to have a DGPS adjust to precise positioning system from the GPS PPS service, we reach a positioning mark with 4 cm of precision.

Part II: Original Research

7.0 Sensing Based in Stochastic Resonance and Chaos Approach

The process to measure physical variables from some environmental signal or effect often starts from sensing elements based on dynamical sensors. Many sensors have nonlinear input-output characteristics and conventional techniques cannot be used to infer a linear and proportional input-output relationship commonly used to a linear sensor. To solve this problem, and to answer the research question related to non-linear sensing based devices a study to solve the problem of nonlinear response sensor behaviour was proposed and simulated using a sensing based on a Chua's circuit chaotic regime.

In the underwater remote sensing context using a manned or unmanned marine platform some nonlinear data are acquired using inertial/navigational and sonar devices. Today, techniques to increase the quality and the confidence level of remote sensing data will impact directly in the survey efficiency and results, one of the most common techniques uses standards sensors, and in some cases with hardware replication, to performing a data fusion among sensing devices. For example, regarding to remote sensing path planning and navigation process, the use of data fusion between positional and attitude data from inertial, navigational and image sensors is most common in order to achieve its self-localization and mapping tasks in underwater inspection surveys. Sensors data fusion will reduce the estimative error in attitude and positioning but can increase a time delay in sensing, data processing or communication stages (Robertsson et al. 2011) and (Kim 1997). This can influence controller time response and timestamp synchronization between all sensors using acquired data. Advances in research involving dynamic nonlinear systems and concepts to treat nonlinearity and chaos in Chua's circuit have inspired this work to improve stochastic resonance in chaotic regime in sensing applications.

In this work, we focus on a new paradigm to process a signal by applying a noise activated nonlinear bistable dynamical system. It can have many applications in detecting weak dc or low-frequency response signals (Kapitaniak & Bishop 1999) and (Bacelar 2010). As an example of this type of sensors, I cite the Fluxgate Magnetometer, that is a core sensor component of an electronic compass system. An electronic compass system is one of the main sensors embedded in mobile robotic platform and it is one of components embedded in any autonomous navigation system. Its task is to give the vehicle attitude and heading information. A common way in navigation and self-localization task is to perform data fusion between inertial sensors and GPS receivers and to improve the navigation data passed to system controller. In an underwater survey application performed with underwater data acquisition platform there is no GPS receiver information. Data from compass and other inertial sensors are the only way to infer the attitude, position and heading data. Estimative in attitude, heading and position is better when the sensor and signal processing are better and faster.

The aim of this work is to simplify and speed up data interpretation from one of the navigation sensor element, a compass based fluxgate magnetometer, and to get higher accuracy in data acquisition and fast data processing in self-localization and mapping application.

7.1 Stochastic Resonance Chua's Circuit and Residence Time Differences

In this study, nonlinear bistable systems are used to detect and quantify small DC or low frequency signals. The proposed sensor is based on the stochastic resonance phenomenon (SR) and the residence time difference (RTD) technique. RTD technique is free from the second type of error (ii) mentioned above and reduces complexity of the electronic board. There is no feedback loop in RTD technique. The existence of chaotic systems with two attractors inspired us to use them as sensors of small DC target signals. We considered a sensor based on Chua's electronic circuit. It is a very simple circuit with a variety of different dynamic behaviours. Its chaotic dynamical behaviour is characterized by two single scroll attractors. The period of switching between attractors is called as residence time and this switching has a random component (Bulsara et al. 2003). The residence time distribution has been studied e.g. in neurophysiological experiments based on spiking stimulus of neuron transmissions. The external target signal produces an asymmetry in the nonlinear states potential what leads to concomitant difference in the mean residence times at two attractors. Performing numerical simulations, we obtained that the residence time difference is proportional to potential asymmetry caused by DC target signal (Bulsara et al. 2003).

Here in this work, we propose how to detect DC or low frequency voltage signals based on the stochastic resonance phenomenon. A simulation data and processing was verified in a practical implementation of Chua's circuit operating in a chaotic state in which two single scroll attractors coexist and the noise activates bistable dynamic sensor (Gammaitoni & Bulsara 2002) and (Dari & Gammaitoni 2009). The Chua's circuit is an easy and accessible circuit at low cost and now can be implemented in a single component based on bulk-silicon as a monolithic chip (Zhong & Ayrom 1985) and (Cruz & Chua 1992). The capability to make Chua's circuit as miniaturized chip is important in its sensor applications. In our applications, the external periodic bias signal is not necessary (suprathreshold scenery) and the noise drives the sensor between attractors and use this observable as a quantifier of the target signal. The work of the sensor operation was conducted using numerical simulations. The best results were archived by applications of an optimal noise. In the sensor design, there are two main difficulties. One is the relationship between input and output signals known as sensor response which should be linear and disrupting in presence of a noise (Gammaitoni & Bulsara 2003). The second difficulty is the observation time (Kennedy 1992). The first difficulty is overcome in our sensor because the noise is inherent in the proposed circuit, and for the optimal noise level there is practically linear dependence between the output time proportion and the input dc target signal voltage. For the second difficulty, we show fast decrease of the error with the observation time for our sensor driven by the optimal noise.

7.2 Sensor Description, Detection Strategy and Operating Range

The proposed design of a sensor based on Chua's circuit that we have studied by numerical simulations and tested in practical experiments based on classical circuit proposed in ref.(Matsumoto 1984) is made up of one resistor, one inductor, two capacitors and one nonlinear element called Chua's diode characterized by a three-segment piecewise-linear current–voltage curve (Dari & Gammaitoni 2009) and (Chua 1992). The Chua's diode can also be made using multiple discrete components like resistors, operational amplifiers, diodes and transistors (Matsumoto 1984) and (Chua 1992), or integrated circuits (Chua 1992). Nowadays in many works Chua's diode is replaced by a new type of a component called Memristor (Chua 1971), (Strukov et al. 2008) and (Wang et al. 2011). The sensor based on Chua's circuit proposed in this work has a standard configuration diagram with the innovative difference, a source of a noise and the signal from sensing element are inputted in the inductor circuit loop. The schematics of the proposed circuit is presented in the *Figure 7.1*.

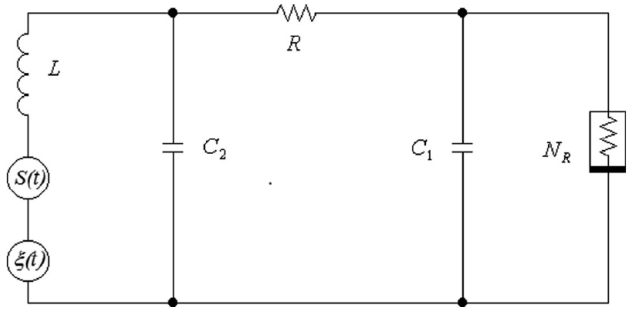


Figure 7.1: *Noise activated voltage sensor based on Chua's circuit*

The current versus voltage dependence of Chua's diode is defined analytically as follows:

$$I = f(V) = m_1 V + \frac{1}{2}(m_0 - m_1)(|V + V_0| - |V - V_0|) \qquad \textit{eq. 7.1}$$

where:

- $f(V)$ – represents the function Voltage-Current;
- V – is the potential voltage over the Chua's diode;
- V_0 – is the potential breakpoint of the Chua's diode response curve;
- m_0 – is the slop curve into the linear region between $-V_0$ and V_0;
- m_1 – is the slop curve out of limits $(-V_0 , V_0)$, also linear from $-V_0\ to -\infty$ and $V_0\ to +\infty$.

In *Figure 7.2*, we show graphical representation of the voltage-current curve and in *Table 7.1* we list values of parameters used.

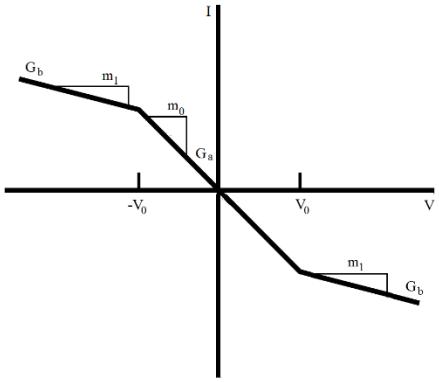


Figure 7.2: *Graphical representation voltage-current*

Table 7.1: *Voltage-Current parameter values*

Parameter	Value
V_0	1.08 Volts
m_0	-0.758 mA
m_1	-0.409 mA

The sensor based on Chua's circuit has dynamical behaviour described by three state variables, modelled by differential equations. These state equations are the following:

$$L\frac{dI_L}{dt} = -V_2 + S(t) + \xi(t) \qquad \textit{eq. 7.2}$$

$$C_2\frac{dV_2}{dt} = I_L - G(V_2 - V_1) \qquad \textit{eq. 7.3}$$

$$C_1\frac{dV_1}{dt} = G(V_2 - V_1) - f(V_1) \qquad \textit{eq. 7.4}$$

where:

$G = \frac{1}{R}$ - represent the inverse of the resistance R which influences dynamical speed;
$S(t)$ - is the target signal to be detected;
$\xi(t)$ - is the noise intensity.

Values of circuit components relates to parameters given in *Table 7.1* was defined in the *Table 7.2*.

Table 7.2: *Components Values to make an Autonomous Chua's Circuit in the Figure 7.2*

Component	Value	Observation
L (Inductor)	0.018 H	
C_1 (Capacitor)	10 nF	
C_2 (Capacitor)	100 nF	
R (Resistor)	1.9~2kΩ	R is a linear variable potentiometer

The autonomous Chua's circuit with these components exhibits rich variety of chaotic attractors, bifurcation phenomena and routes to chaos (Matsumoto 1984). The sensor based on Chua's circuits has different dynamical behaviour based on the value of the resistance R. This is described in *Table 7.3*.

Table 7.3: *Relationship between value of the resistance R and Chua's circuit dynamical behaviour*

Input		Range Value selected (R)	System Behaviour
$S(t)$	$\xi(t)$		
0	0	R > 2.1 kΩ	The sensor is in the stable equilibrium point, figure 36A
0	0	~2030 Ω < R < 2.1 kΩ	The sensor loses the stability of the equilibrium point
0	0	≈2030 Ω	Stable limit circle emerges (Hopf bifurcation), figure 36B
0	0	~1994 Ω< R <~2030 Ω	Gives rise to period-2, period-4 etc. limit circles (Bifurcations),
0	0	~1969 Ω<R <~1994 Ω	Two single scroll attractors coexist, figure 36C
0	0	R< 1969 Ω	The single scroll attractor abruptly enlarges itself and the double-scroll attractor is created, figure 36D
0	0	=1970 Ω	Sensor operates near the birth of the double scroll attractor

For R values (~1969Ω <R < ~1994Ω), the system trajectory will converge to a particular attractor if the initial conditions are chosen in its basin of attraction *Figure 7.3*. The two coexisting attractors we observe in experiments depends on the initial state of the sensor. The single scroll attractor has a structure of Rössler's spiral-type attractor (Zhong & Ayrom 1985). The dynamics of a sensor is not strictly periodic, but its frequency varies in time around some preferred value which in our case is 2900 Hz. This dynamic behaviour mimics the oscillatory processes occurring in natural biological systems(Moss 2004).

The noise activated dc signal sensor proposed in this work operates in the resistance R range in which two single scroll attractors coexist *Figure 7.3 D*. By increasing R in this range one increases the speed of target signal detection because noise activated crossings between attractors become faster.

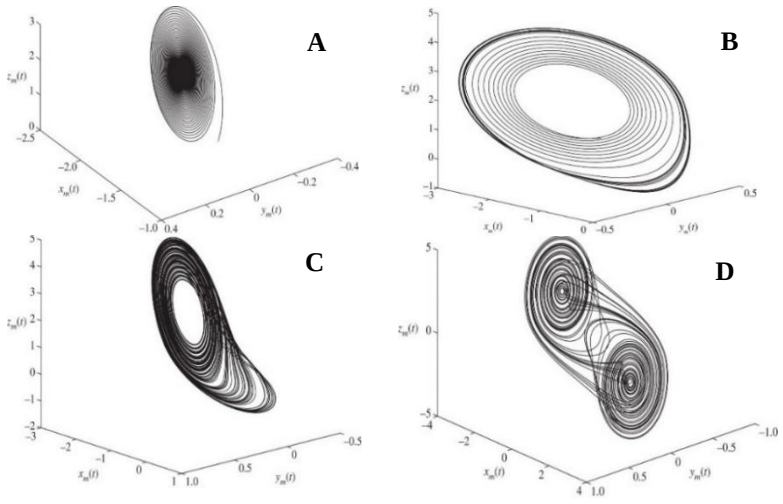


Figure 7.3: *Different dynamical behaviour of Chua's circuit in the Phase space.*

We selected R=1970Ω, i.e. the sensor operates near the birth of the double scroll attractor. In recent years Chua's diode with three-segment piecewise-linear current–voltage characteristic has been replaced by Chua's diode with smooth cubic current–voltage characteristic (Rössler 1979). Chua's circuit with cubic nonlinearity has a period-1 oscillation frequency in the MHz frequency range, so it is more than 300 times faster than Chua's circuit used in this paper.

7.4 Numeric Simulations

We have performed computer simulations using the set of three differential *equations 7.4* to *7.6*. The time step was $\Delta t = 10^{-6}$ s. Each simulation run we started from different initial conditions. We assumed the noise $\xi(t)$ to be the Gaussian white noise with zero mean. The noise intensity, i.e. the noise standard deviation we denote by σ. The Box–Muller algorithm (Box & Muller 1958) was used to generate Gaussian random variables from uniformly distributed random variables in the interval [0, 1]. When the target and noise signals are absent, the dynamic trajectory of the sensor is in one single scroll chaotic attractor. When only dc target signal is present, dynamical

trajectory remains on one attractor. Switching on the noise source causes crossings of dynamic trajectory between coexisting single scroll attractors. We followed the value of the voltage $V_1(t)$ and used the step function $\theta(t) = \pm 4$ V to identify the attractor in which the trajectory is located. In this procedure, additional crossing levels ± 4 V were used to eliminate anomalous switching events where the trajectory crosses zero but remains on the same attractor (Gammaitoni et al. 1998).

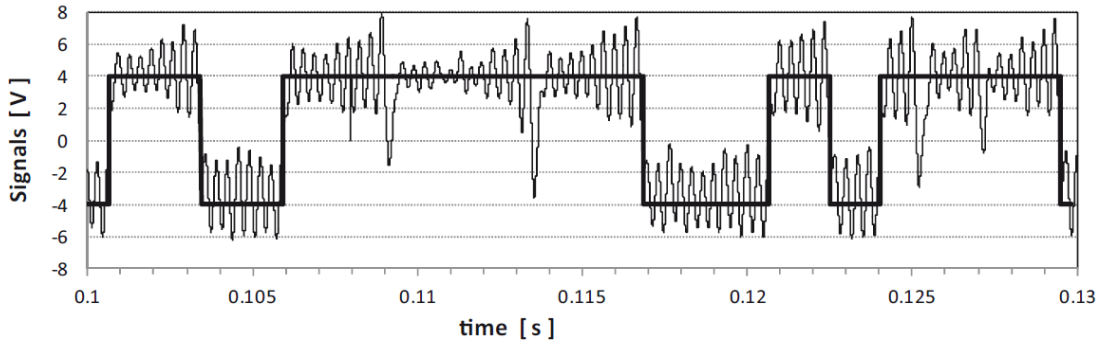


Figure 7.4: *The voltage* $\boldsymbol{V_1(t)}$ *(thin line) and the step function* $\boldsymbol{\theta(t)}$ *(thick line) for dc target signal* $\boldsymbol{S(t)}$ *= 60* $\boldsymbol{mV}$ *and the noise intensity* $\boldsymbol{\sigma = 0.5}$ *V.*

The example of the voltage $V_1(t)$ and the step function $\theta(t)$ temporal evolution for the weak dc target signal 60 mV and for the noise intensity $\sigma = 0.5$ V is shown in *Figure 7.*4. The information about the target signal is contained in unequal residence times of the sensor dynamic trajectory in two stable single scroll attractors. The sensor converts dc target signal voltage into square waveform with stochastic switching between two states. The sequence of crossings between attractors can be linked to the spike train recorded from real sensory neurons (Wellens et al. 2004), (Lindner 2004), (Bulsara & Gammaitoni 1996) and (Moss 2004).

We propose to quantify dc target signal by the proportion T_u/T of the time T_u which dynamic trajectory stays in the single scroll attractor with the center at the positive voltage +4 V to the total observation time T. We choose the time proportion and not the mean residence time difference in coexisting attractors, like in RTD technique, to simplify the readout scheme and avoid averaging. The time proportion T_u/T depends both on dc target signal voltage and on the noise intensity as it is shown in *Figure 7.5*.

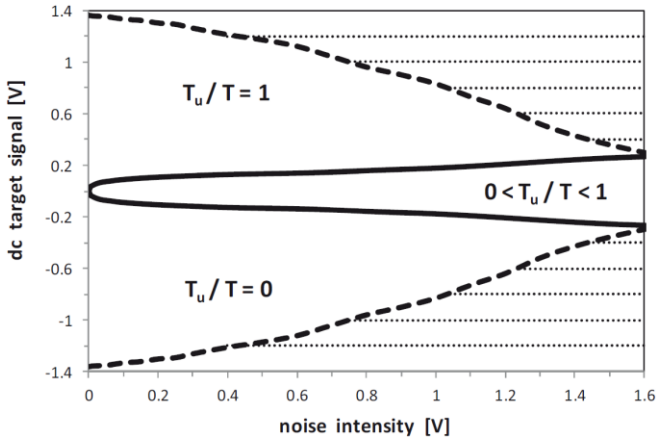


Figure 7.5: *The diagram showing the time proportion* $\boldsymbol{T_u/T}$ *for different values of dc target signal and the noise intensity. In regions denoted by dotted lines the sensor is unstable.*

The dependence of the time proportion T_u/T on the noise intensity for different positive dc target signals is presented in *Figure 7.6*. For negative dc target signals this dependence is symmetric to the dependence shown in *Figure 7.6*. with respect to the line $T_u/T = 0.5$.

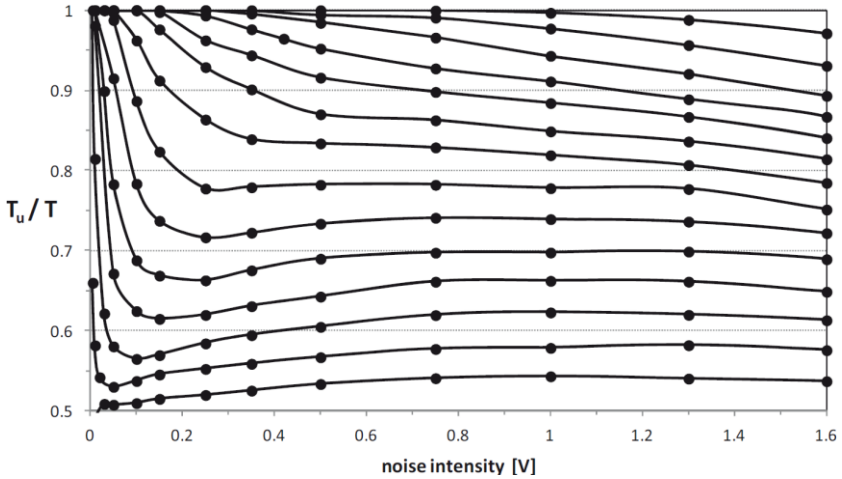


Figure 7.6: *The dependence of the time proportion Tu/T on the noise intensity for different positive dc target signals. The curves from bottom to top are for S(t) equal: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 125, 150 and 200 [mV].*

7.5 The Optimal Noise Level for Signal Detection

To perform the study of signal detection we assume that dc target signal to be detected is in the range (-127.5 mV,127.5mV). In *Figure 7.7*, we present enlarged part of *Figure 7.5*. Thick dashed lines show the selected range of dc target signals for three noise intensities: 0.005, 0.5 and 1.6 V. One can notice that this range for the very low noise intensity 0.005 V is mainly in the region where Tu/T = 1, so most of dc target signals are not detected. For the noise intensity 1.6 V the selected range spans the range of time proportions (-0.9, 0.9), so the sensor resolution for time proportions outside this range is not used. Higher noise intensities deform attractors too much and assigning dynamic trajectory to a given attractor becomes impossible. The sensor thus cannot operate for σ> 1.6 V.

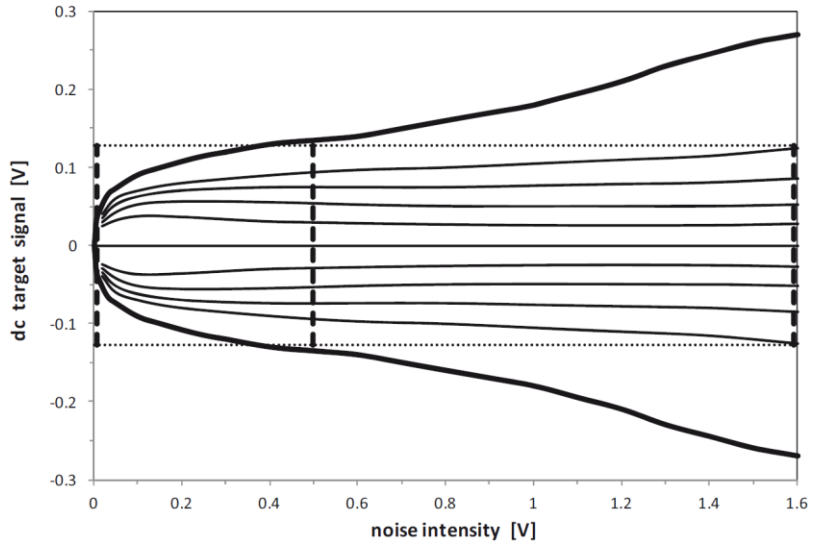


Figure 7.7*: The enlarged part of the diagram from Figure 16. The dotted lines are for ± 125.5 mV. The thick dashed lines are for noise intensities 0.005,0.5 and 1.6 V. The thin lines correspond to time proportions T_u/T: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9 (from bottom to top).*

The optimal noise intensity for the range (-127.5 mV, 127.5 mV) of dc target signals is near 0.5 V. In *Figure 7.8*, we present the dependence between the output time proportion T_u/T and dc target signal voltage obtained in this case for two observation times. This dependence is close to the linear dependence denoted by grey line in the

figure. One can notice that the error in measurement of dc target signal decreases with the observation time. To quantify this effect, we divided the input dc signal range (-127.5 mV, 127.5 mV) into 256 uniformly distributed voltages and we performed computer simulation runs for all these voltages over given observation time.

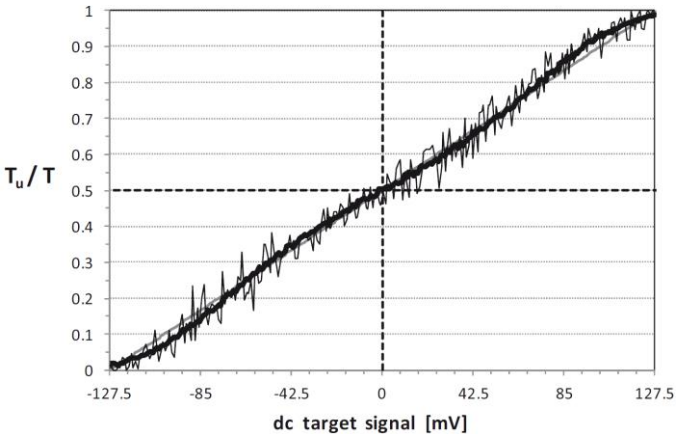


Figure 7.8: *The dependence of the output time proportion on the input dc signal voltage for the noise intensity 0.5 V. The observation time was 0.25 s (thin line) and 15 s (thick line).*

Next we rescaled obtained output time proportions into output voltages using the formula $S_{out} = 255(T_u/T - 0.5)$. The dc target signal measurement error $E(t)$ we then calculated using the following formula from *equation 7.7* as follow.

$$E(t) = \langle \sqrt{\sum_{i=0}^{255} \frac{[S_{out}(i,t) - S_{out}(i,t=15)]^2}{256}} \rangle \qquad \textit{eq. 7.5}$$

Where the sum runs over 256 uniformly distributed voltages and brackets < . . .> denote the average over several runs starting with random initial conditions. In this formula $S_{out}(i, t = 15)$ is the output voltage after 15 s obtained from time proportions shown in *Figure 7.8*. In *Figure 7.9* we show how the measurement error decreases within the first 1.5 s of the observation time.

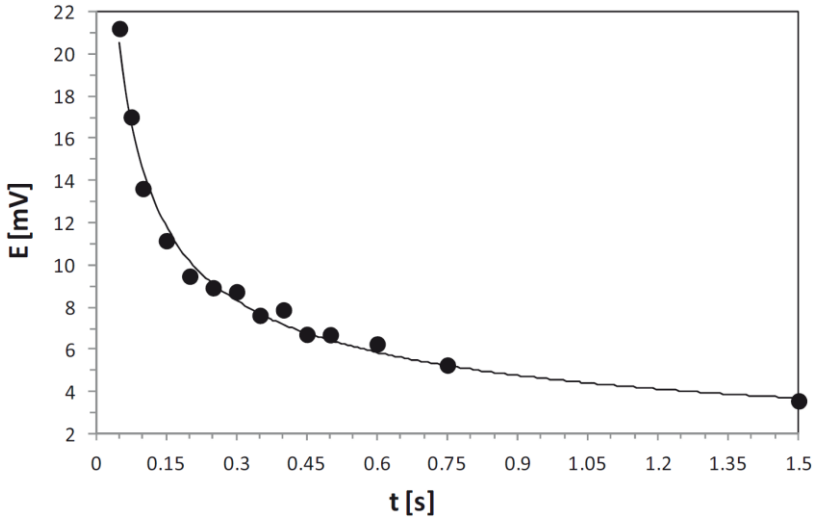


Figure 7.9: *The dependence of the measurement error on the observation time for the noise intensity 0.5 V The continuous line is defined by the power law with the exponent -0.5.*

This decrease can be well fitted by the power law $E(t) = 4.53/\sqrt{t}$, which is shown by continuous line in this figure. This dependence was expected and it was also obtained for the measurement error in sensors based on RTD measurement technique(Nikitin et al. 2003).

7.6 Conclusion

In this section, we proposed the new stochastic resonance noise activated sensing detector applied to measure a weak dc signal or also a low frequency voltage target signal in sensing applications. The sensor is based on Chua's electronic circuit operating in chaotic range where two single scroll Rössler's spiral-type attractors coexist. We selected standard parameters of electronic components, so this sensor can practically be implemented at very low cost. Moreover, Chua's circuits can now be fabricated in a single chip, so this sensor can be made very small. Small-size sensors are especially needed and used in many applications. The detection strategy of the proposed sensor is based on the monitoring of crossings between attractors activated only by the noise added to a target signal. The proportion of time the sensor stays in one attractor is used to quantify its output. The operational scenario described in this paper has many advantages. The sensor performs well in the presence of noise, because the noise is inherent in its functioning.

The readout scheme is very simple to implement practically without complicated electronics. The sensor operates in the absence of a time-periodic bias signal commonly used in conventional bistable dynamic sensors to drive them between their stable states. Performing numerical simulations, we obtained the diagram showing the output time proportion for different values of the input dc target signal voltage and the noise intensity. This diagram allows one to determine the range of noise intensities optimal for detection of a given target signal. We found the optimal noise intensity 0.5 V for dc target signals in the range (-127.5mV;127.5mV). We obtained nearly linear output–input relation in this case, what is one of the most desirable features of a good sensor. The short observation times often limit the applicability of noise activated sensors operating without the bias signal. For the optimal noise level, we obtained that the measurement error decreases with the observation time according to the power law with the exponent 0.5 and it is smaller than 4 mV after 1.5 s. Finally, we would like to mention very close relation between the sensor described in this paper and natural biological systems. Its internal dynamics within the attractor mimics the natural oscillatory processes, whereas crossings between attractors can be linked to the spike train recorded from real sensory neurons.

8.0 Contrast Modification of Acoustic Images Based on Stochastic Resonance

In this section we present the model of contrast enhancement of acoustic images developed in (Sousa et al. 2016) that was inspired by environmental noise applications in biological systems. Visual detectability of biological systems attracted considerable attention over the last decades. It was shown that visual perception characteristics of the input subthreshold signal are enhanced at some optimal level of added noise (Simonotto et al. 1997) and (Siegel & Read n.d.). This process comes from a phenomenon called the stochastic resonance (L. Gammaitoni et al. 1998) and (Lindner et al. 2004). Recently, new form of stochastic resonance phenomenon has been introduced under the name of suprathreshold stochastic resonance (SSR) to illustrate the fact that it is not restricted to subthreshold signals (Stocks 2000) and (McDonnell et al. 2012). This phenomenon occurs in an array of comparators (threshold devices) subject to the same input signal and independent identically distributed noise. The output from each device is summed to give an overall output. It has been shown that above a certain noise intensity the optimal threshold settings of comparators occur when all thresholds are equal to the signal mean (McDonnell et al. 2006) and (Mcdonnell et al. 2005). The array of comparators subject to signal and noise is a useful model for a range of systems including flash parallel analog-to-digital converters (McDonnell et al. 2005), networks of sensory neurons (Ashida & Kubo 2010) and digital sonar arrays e.g. DIMUS (Digital Multibeam Steering) sonar array which was one of the first digital beamforming sonar systems (Anderson 1960)(Rudnick 1960). It has been shown that pre-processor based on the phenomenon of SSR can be used to improve the performance of bearing estimation especially in an environments contaminated with leptokurtic (heavy-tailed) noise (Hari et al. 2009a)(Hari et al. 2009b). This non-Gaussian noise is typical for ocean acoustic environments (Machell et al. 1989a). The performance of the pre-processor was optimized by appropriate selection of the standard deviation of the noise added to comparators and their number.

Sonar systems represent an efficient tool for obtaining acoustic images from underwater relief (Lurton 2010). These systems provide a portrait of underwater sounded area in the form of high-resolution images even in waters where the visibility is poor for optical systems or human vision (Celik & Tjahjadi 2011). The processing of these images is the basis for solving different problems in automated target identification, underwater pipeline and cable inspections and in the navigation (Chapple 2008) (Acosta & Villar 2015). The lateral side scan sonar (SSS) is one of the most widely used type of lateral scanning sonars, constituting the main visualization tool for obtaining acoustic images of the sea floor (Celik & Tjahjadi 2011) (Esther Dura et al. 2005) (Reed et al. 2006) (Petillot et al. 2002). The SSS images present a low dynamic range of pixel intensities and it is difficult or often impossible to distinguish targets or objects details in these images. Thus, they must be improved to perform tasks like the automated target detection and recognizing. Additionally, SSS images present an inherent noise known as speckle pattern. This noise is a common disturbance in remote sensors that use coherent radiation as a source. When the energy pulse in the form of acoustic waves is emitted, after the impact or interaction with a target or underwater bed these waves are no longer in phase. There are thus positive or destructive interferences, which generate abnormally high or low returns. Speckle noise is difficult to distinguish automatically, and its elimination is complex. At present, there is no universal algorithm for the improvement of SSS acoustic images, although there are several strategies used which depend on the sonar device used, the wavelength, acquisition depth, processing time and other factors (Lurton 2010) (Chapple 2008) (Blondel 2009).

In this section the technique to modify digital sonar grayscale image contrast using noise based on the threshold elements is investigated. Many of our sonar images have very low dynamic range of the intensity values and therefore needed to be enhanced before being displayed. Moreover, the performance of processors used for underwater source localization degrades rapidly as the quality of the sonar image reduces. It was shown that pre-processing of sonar images extends their usability (Hari et al. 2009a) and (Hari et al. 2009b). We apply the algorithm for contrast modification like the algorithm proposed by Marks et.al. (Marks et al. 2002). In this algorithm, the noise is added repeatedly to an image and is successively thresholder followed by overall averaging. Our aim is to demonstrate how the contrast and the quality of digital sonar image depends on the number of binary representations used to get the output composite image and on the noise mean, intensity and probability distribution. We consider the generalized Gaussian noise parameterized by positive exponent p (Nadarajah 2005). This noise family is widely used in underwater acoustics and sonar application (Machell et al. 1989b). We selected the Uniform noise described by the rectangular probability density function (RPDF with $p=\infty$), the white noise described by the Gaussian probability density function (GPDF with $p=2$) and the Laplacian noise described by the Laplace probability density function (LPDF with $p=1$). The Gaussian white noise was successfully used for contrast enhancement of low-contrast and low-quality images (Chouhan 2012) and (Jha et al. 2012), the Uniform

noise was used by Marks et.al. (Marks et al. 2002) to present the suprathreshold stochastic resonance for the image signal, whereas the Laplacian noise is used to model underwater acoustic noise (Machell et al. 1989a) and it can effectively be used in pre-processing of sonar images (Hari et al. 2009a)(Hari et al. 2009b). In experimental situations, it is often difficult to distinguish between stochastic noise and deterministic chaos. We have also replaced the stochastic noise by the chaotic time-series generated by the logistic map. The logistic map is a very simple quadratic difference equation parameterized by positive parameter R (May 1976). For the R=4 case, the iterate sequence of numbers in the interval (0.1) generated by this map is chaotic for almost all initial conditions. The probability density function of these numbers increases fast and reaches maxima near 0 and 1, so it is very different from generalized Gaussian probability density functions. It was observed that the amount of improvement offered by the stochastic resonance is dependent on the signal-to-threshold distance (Korneta et al. 2006) and (Stocks 2000). This distance can be changed by moving the noise mean or the threshold. In order to show how the contrast and quality of the output composite image depend on the number of its binary representations, we have obtained output images for the number of binary representations from 1 to 10000. It has proven difficult to determine a metric for the contrast of natural images. Bex et.al. (Bex & Makous 2002) showed that rms contrast was the best metric for natural images among the commonly used Michelson contrast, root-mean-square (rms) contrast, and band-limited contrast metrics (Peli 1990). The rms contrast is a single value assigned to the whole image simple to calculate. We use rms metric to quantify image contrast in this work. (Collins et al. 2002) proposed cross-correlation measure as appropriate measure to quantify aperiodic stochastic resonance. The correlation coefficient was used e.g. to measure the performance of preprocessors enhancing sonar images (Hari et al. 2009b). We measure the quality of the processed image in terms of its correlation coefficient with the original image. We also compare results of image contrast modification considered in this work with results of image histogram equalization, histogram stretching and contrast and brightness modification procedures in GIMP image processing software.

8.1 Method Description

In this section, the method of contrast and quality enhancement applied to SSS images is described. Let's denote the original input image of size $m \times k$ by I . $I(i,j)$ represents the intensity level of a pixel in $i-th$ row and $j-th$ column with $i = 1, \dots, m$, and $j = 1, \dots, k$. We denote by b the possible number of intensities and for 8 bits $b = 256$. The proposed method consists of generating the composite image I_c by producing and combining N binary images. Each binary image $g(i,j)$ is generated according to the following formula:

$$g(i,j) = \begin{cases} e_1 \; if \; I(i,j) + (b-1) - \xi(i,j) \leq T \\ e_2 \; if \; I(i,j) + (b-1) - \xi(i,j) > T \end{cases} \qquad \textit{eq. 8.1}$$

Where $\xi(i,j)$ denotes random variable extracted from the probability distribution of the noise at the position (i,j). The random variable $\xi(i,j)$ is generated and added to each pixel intensity of the image $I(i,j)$. The result of the operation $I(i,j) + (b-1) - \xi(i,j)$ is thresholded using a single threshold T. If it is less than or equal to the threshold T, it establishes a label e_1 (for example, the lowest possible pixel intensity value $e_1 = 0$). Otherwise, a label e_2 is established (for example the highest possible pixel intensity value $e_2 = b-1$). Generally, a constant threshold T is set equal to $(b-1)$. The resulting image will contain only two possible tag values e_1 and e_2, generating a binary image. This procedure is repeated up to N times generating a set of binary images $g_1(i,j), \dots, g_N(i,j)$. These images are combined to form a single composite image I_c by averaging its pixel values. Each composite image I_c will contain $N+1$ pixel intensity values. The composite image I_c is finally approximated to intensity of pixel values between 0 and $b-1$.

In order to evaluate the composite image quality and contrast we have different metrics that emerge from the theory of digital image processing (González 2010). To analyse the contrast of a digital image I we have:

- Michelson's contrast (Albert A Michelson 1927) useful for periodic patterns defined as the quotient between maximum intensity $(I_{\max})$ and minimum intensity (I_{min}) difference and the sum of these image boundary intensities I:

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad \textit{eq. 8.2}$$

- Webber contrast (Peli 1990), useful for well-defined objects and background defined as the difference between the intensity at a point $I(i,j)$ and the average intensity $< I >$of the image I divided by the average intensity $< I >$.

$$C = \sum_{i=1}^{m}\sum_{j=1}^{k}\frac{I(i,j) - \langle I \rangle}{\langle I \rangle} \qquad \textit{eq. 8.3}$$

- Contrast expressed by the root mean square value (RMS) (Peli 1990), which uses the standard deviation σ_I of the pixel intensity levels in the image I defined as.

$$C = \sigma_I = \sqrt{\frac{1}{mk-1}\sum_{i=1}^{m}\sum_{j=1}^{k}[I_n(i,j) - \langle I_n(i,j)\rangle]^2} \qquad \textit{eq. 8.4}$$

Where I_n represents the normalized intensity level such that $0 \leq I_n \leq 1$.

On the other hand, in order to analyses the quality of a digital image, the correlation coefficient ρ between the input and output image (Collins et al. 1995) (Mcdonnell et al. 2002) can be used. This coefficient between the composite output image I_c and the original image I is defined as:

$$\rho = \frac{\sum_{i=1}^{m}\sum_{j=1}^{k}(I_C(i,j) - \langle I_C \rangle)(I(i,j) - \langle I \rangle)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{k}(I_C(i,j) - \langle I_C \rangle)^2 \sum_{i=1}^{m}\sum_{j=1}^{k}(I(i,j) - \langle I \rangle)^2}} \qquad \textit{eq. 8.5}$$

Therefore, if the correlation coefficient ρ presents a high value, close to 1, it means that the images I_c and I are similar. It thus represents an appropriate quantitative measure of similarity.

8.2 Sidescan Device and Image Acquisition

The acoustic images were acquired in *Todos los Santos* bay, near the Salvador city in, state of Bahia in Brazil. For this, the side scan sonar (SSS) Starfish 450F of the company Blueprint was used (S. Blueprint 2016). The SSS was mounted on the front of a research vessel's hull using a steel bracket Stainless as shown in *Figure 8.1*. The SSS device comprises a dual acoustic echo channel employing two frequencies 430 kHz and 470 kHz to produce a chirping type signal. The maximum sonar range was set to remain three to four times greater than the depth of the water. The SSS provide an acoustic pulse for each side and each of these called beam. These beams are acquired perpendicular to the trajectory that the boat makes. Transducers array are tilted 30° to the horizontal reference or water surface. The main acoustic beam had 60° of vertical aperture (perpendicular to the sonar track line) and 1.7° of horizontal aperture or azimuth (parallel to the sonar track line)(S. Blueprint 2016). Figures and 2 show some SSS device beam characteristics. If the acoustic pulses move through the column of water without refraction, the pulse strikes the frontal surface of an object that returns to sonar being much stronger than the marine surface around the object. To survey vessel navigation followed the best practices of the National Oceanic and Atmospheric Administration (NOAA), (Umbach 1976) and (Bowditch 2002).

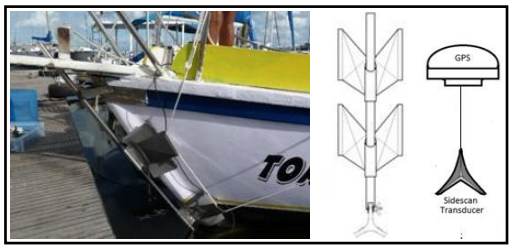


Figure 8.1: *Survey boat, sonar from the left with sonar bracket and GPS-Sonar head relative position.*

The vessel speed was less than 4.5 m/s for depths of between 5 m to 20 m and was reduced between 1.5 m/s and 2 m/s in deeper water (> 20 m). These speeds are set by the manufacturer of the acoustic device for correct

coverage. In this way, by joining successive contiguous beams, it is possible to construct acoustic SSS images. These images are saved in bitmap format with 8 bits per pixel (intensity levels between 0 and 255). To acquire and insert a spatial data from standard georeferenced system we used a SIF III GPS device connected to the survey by Universal Serial Bus (USB)(S. Blueprint 2016).

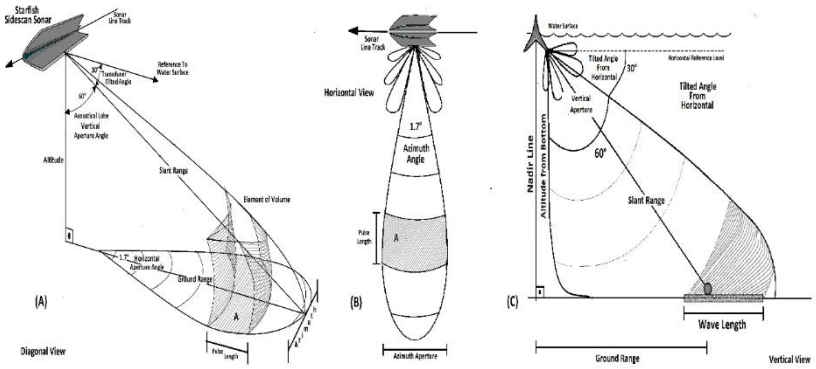


Figure 8.2: *Sidescan beam characteristics and coverage*

Data extraction are performed in shallow waters then we assume that the sound pulse moves in the water column without refraction, the returned reflected pulse from the front surface of the object made of denser material (pipeline, rock a etc.) is much stronger than from the background (flat seabed). Our images were saved in the bmp file format with 8 bits per pixel, i.e. pixel grayscale intensities were between 0 and 255 (S. Blueprint 2016).

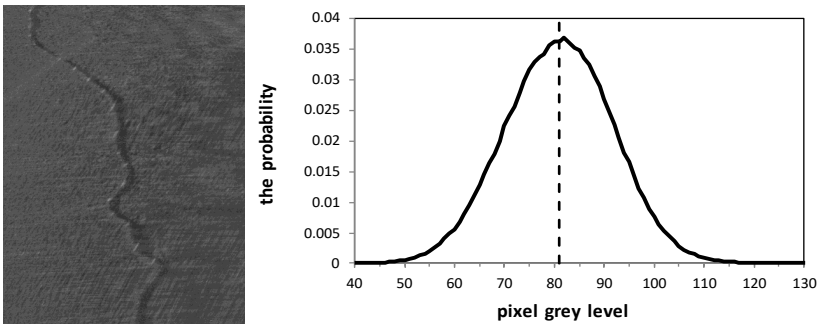


Figure 8.3: *Sidescan image seeing a real pipeline extracted in a hydrographic mission survey (left) and their pixel grayscale level distribution (right)*

Trial and surveys were performed in two stages. The first stage was initialized at a latitude and longitude of -12 ° 51'19.5 " S, -38 ° 32'23.03"W respectively and finished at -12 ° 52'23.28 "S, -38 ° 33 ' 48.48 W. The system has collected 50500 beams of acoustical data which provided 51 images of size 1000x1000. The second stage began at a latitude and longitude of -12 ° 53 '33.04 ", -38 ° 33' 48.14" W respectively and ended at -12 ° 52 '16.1 "S and -38 ° 31' 37.14" W. The system collected 47000 beams of acoustic data providing a total of 47 images of the same size as in the first stage. One of the obtained images was cut out for its best presentation (size 604x800 pixels) and it is shown in Figure 8.3. In this figure, we present the real acoustic image of a pipeline lying at the bottom at 6 m depth from water surface acquired from SSS device at left and its histogram defined as the probability distribution of pixel grey levels at the right graph.

8.3 Nonlinear Image Processing and Analysis

Vision is the most powerful way of environmental perception to human and animals. The process of the vision can be a single static image, like a photo, being spatial representation of environment, an animal or an object. Besides this, a movement of creatures and objects can be detected like a movie simple called scene. Today image processing is a tool used in several areas and applications as industrial, medical, security, entertainment etc. The aim of image processing is often to extract and classify information and objects at images. Often it is required to extract subparts of an image and focus in a specific area or object in a region of interest (ROI) according to its certain properties. It permits to get some relevant information in ROI, make an analysis and specific decisions. Here we divided these procedures in two hierarchical groups: Pre-processing and Processing. The pre-processing

operation is composed by a set of procedures to prepare or conditioning images as normalization, selective focus, binarization, filtering, resampling etc. It can significantly increase the reliability of acoustic image surveys. We can perform image pre-processing by applying some filtering algorithm in spatial or frequency domain. In pre-processing image processing stage, computational algorithm extracts the information of interest to make decisions faster and exactly as required. Here, we focus on one type of filtering i.e. acoustic image enhancement, a pre-processing task to prepare adequately the raw image to the next processing stage.

8.3.1 Nonlinear Approach applied to Acoustic Images

Active sonars are commonly monostatic signal acquiring systems what means that the source and the receiver are in the same location. With this, the acquiring process is the foillowing:

I. A pulse is generated by the acoustic source and transmitted to the medium;
II. The acoustic pulse travels through the medium suffering some losses imposed by physical properties of the propagation medium. Signal absorption signal is determined mainly by the pulse frequency[2] and the distance (Stojanovic & Freitag 2006);
III. When the pulse reaches a target, a fraction of its original intensity is reflected and/or spread through the medium, depending on physical characteristics of the target and the medium;
IV. The portion of a pulse that is backscatter to a sonar platform will suffer losses too and it reaches the sonar receptor.

The acoustics pulse energy is propagated and absorbed to depending on the frequency, the range, the density and elasticity modulus of the medium. These processes are modelled using a square root and logarithmic dependence. The acquisition signal has thus non-linear characteristic. In addition, the target reflection can have random reflectivity and absorption depending on their properties.

The acoustic image has itself non-linear properties and because of this our procedure of sonar image processing uses non-linear approach like the stochastic resonance. We apply this technique to enhance acoustic images and prepare them to a next stage of image processing i.e. selection of ROI and extraction of the information.

8.3.2 Image Enhancement Using Stochastic Resonance

Image enhancement is one of the pre-processing techniques which facilitates the next stage of image processing i.e. feature extraction. It performs balance between image pixel values by expanding a dynamic range of pixel level representation. Researches in this area develop solutions using global or local approach. In the first case this technique is applied to all pixels of an image or scene. The local approaches are applied into a specific area, ROI of image or scheme. A care in the image enhancement application must be taken to maintain a coherence between pixel levels. In acoustic image processing perception of some features and information is limited by small differences in their grayscale.

The stochastic resonance effect usually requires the following three components: subthreshold input signal, a noise with small correlation time and a nonlinear system. The stochastic resonance phenomenon was studied mainly in bistable and excitable nonlinear systems. The periodic stimulus is assumed as the input signal in most of the literature on the stochastic resonance (L. Gammaitoni et al. 1998) and (Korneta et al. 2006). In the case of aperiodic signals which are common in real world the term aperiodic stochastic resonance was introduced (Gang 1992) and (Harmer & Abbott 2000). The aperiodic stochastic resonance is characterized by the maximum of cross-correlation between the stimulus signal and the system response. Typically, in theoretical and experimental studies of the stochastic resonance and aperiodic stochastic resonance phenomena the white noise with the Gaussian probability distribution and zero mean value is added to the input signal (L. Gammaitoni et al. 1998), (Wellens et al. 2004) and (Korneta et al. 2006). In some studies, the stochastic noise was replaced by deterministic chaos which models dynamical input from an environment.

We performed image enhancement using contrast modification based on the stochastic resonance. The technique we applied to digital acoustic image being digital representation of an acoustic underwater sonar system. The inspiration comes from processing of visual signals observed in biological systems. The physical visual effect is like separation of the distance between the brightest and the darkest pixels in grayscale image. With this we get an emphasis in grayscale range i.e. shadows are darker and highlights are brighter. The purpose is to have better

[2] The signal frequency determines the absorption occurs because of the transfer of acoustic energy into heat.

visible details in the image what facilitate the next stage of image processing to get relevant information for feature extraction. Acoustic underwater image acquisition has random nature, because some parameters of the acquisition process and of the underwater medium have nonlinear aspect. Our proposition was to apply a nonlinear approach to acoustic image enhancement based on the stochastic resonance. By adding an optimal noise to the image we increased its contrast.

In this method, we search for an optimal noise level that applied to digital sonar grayscale image enhances its contrast. The output image is formed by averaging binary images obtained by repeatedly adding a noise to the input image and thresholding. The output image contrast and quality are quantified by rms contrast and the output-input correlation function. The dependence of these quantities on the number of binary representations forming the output composite image and on the noise mean, intensity and distribution is discussed.

8.4 Algorithm for Image Contrast Modification by Noise and Quantitative Performance Metrics

The method of image contrast modification applied in this work to SSS images is based on the method used by Marks et.al. (Marks et al. 2002) to present the stochastic resonance effect in the image visualization. The method was presented in the ARGENCON 2016 Conference (Sousa et al. 2016). Let's denote by Image(i,j) the grey level of a pixel in the ith raw and jth column of the image, where i=1,…,m and j=1,…,k . We denote by ξ(i,j) random variable drawn from a given noise probability distribution at the position (i,j) . We generate for each image pixel random variable 255-ξ(i,j) and add it to grey level of the pixel. The obtained sum Image (i, j) + (255-ξ(i,j)) is then subjected to a threshold and it equals 0 for values smaller than 255 or 255 otherwise. This procedure forms binary image. Repeating this procedure N times for different realizations of the noise ξ(i, j) it is obtained a set of N binary images. The average of these binary images forms a composite image with (N+1) grey levels between 0 and 255. This composite image is finally approximated by the output image with 256 gray levels between 0 and 255. In this work, we selected the root mean square (rms) contrast to measure the contrast in the image. This contrast characterizes the image globally. The rms contrast is a common way to compare contrast of images. We found it the most suitable contrast metric to quantify the performance of the technique of contrast modification using the noise for our sonar images. In this paper, we use normalized rms contrast defined as the standard deviation of pixel grey levels σ_I divided by 127.5 which is the maximal possible value of σ_I. The formula for normalized rms image contrast C is thus the following:

$$C = \frac{\sigma_I}{127.5} = \frac{\sqrt{\langle (Image - \langle Image \rangle)^2 \rangle}}{127.5} \qquad \textit{eq. 8.4}$$

Where $\langle . \rangle$ denotes an average. We obtained for the image shown in *Figure.8.3* $\sigma_I = 10.87$, so its normalized rms contrast $C = 0.085$ is very small. The normalized rms contrast $C \sim 0.1$ was typical for our sonar images. The average of pixel grey levels for the image is $\mu_I = 81$ and it equals to their median value. The measure often used to characterize aperiodic stochastic resonance and suprathreshold resonance is the correlation coefficient between the output and input signals (Collins et al. 2002), (Harmer & Abbott 2000), (Mcdonnell et al. 2002), (Hari et al. 2009b) and (Chapeau-Blondeau 2000). The correlation coefficient between the output composite image $Image_C$ and the original image $Image_O$ is defined as:

$$\rho = \frac{\langle (Image_C - \langle Image_C \rangle)(Image_O - \langle Image_O \rangle) \rangle}{\sqrt{\langle (Image_C - \langle Image_C \rangle)^2 \rangle \langle (Image_O - \langle Image_O \rangle)^2 \rangle}} \qquad \textit{eq. 8.5}$$

The correlation coefficient is large when images $Image_C$ and $Image_O$ are similar and it is an appropriate quantitative measure of their similarity. We use the correlation coefficient ρ to quantify the quality of the output composite image.

8.5 Application of the Method to the Sonar Image and Quantitative Representations of Results

In this section, we apply the technique of image contrast modification using noise to the sonar image shown in *Figure 8.3*. By μ and σ we denote the noise mean and standard deviation (i.e. the noise intensity) respectively. The random variable ξ(i,j) is computed from the Gaussian distribution with the mean value equal to the average pixels grey levels of the image i.e. μ=μ_I. We first selected the Gaussian white noise, because it is the most often used to present the stochastic resonance and suprathreshold resonance phenomena in many different systems (L. Gammaitoni et al. 1998) and (Korneta et al. 2006) and in image processing methods based on these phenomena (Simonotto et al. 1997), (Harmer & Abbott 2000), (Chapeau-Blondeau 2000), (Ryu et al. 2011), (Ye et al. 2004), (Chouhan 2012) and (Chouhan 2012). We have used the Box–Muller algorithm (Box & Muller 1958) to generate Gaussian random variables from uniformly distributed random variables in the interval [0, 1]. We have obtained a series of output composite images for different values of the noise intensity and different number of binary representations. Generally, in studies of aperiodic stochastic resonance or suprathreshold resonance phenomenon one usually plots the dependence of the measure of the output-input fidelity, defined e.g. by the correlation coefficient, on the noise intensity, i.e. on the noise standard deviation (Collins & Kuperman 1994), (Harmer & Abbott 2000) and (Mcdonnell et al. 2002). In image visualization, the image contrast is another factor important for its perception. The normalized rms contrast we use in this paper depends on the noise intensity. In order to optimize the output composite image one should thus observe how both its rms contrast and its correlation coefficient with the original image depend on the noise intensity. Moreover, one should consider how these two images metrics depend on the number of binary representation forming the output composite image. We propose to represent the efficiency of the technique of image contrast modification using noise in the plot which takes into account all these factors.

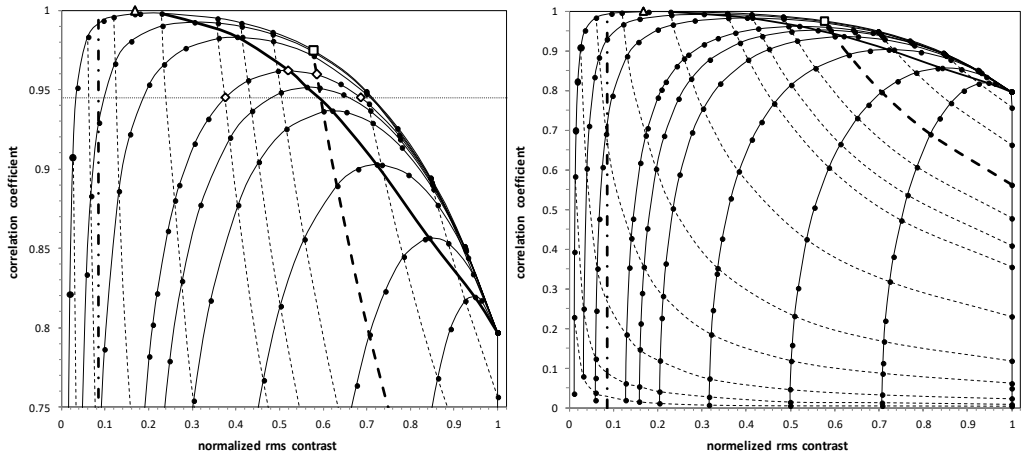


Figure 8.4: *The CvC plot for the Gaussian noise represents the correlation coefficient ρ versus contrast C. Thin lines from the right to the left are for 1, 2, 4, 10, 24, 40, 63, 256, 1000 and 10000 binary representations. The vertical dot-dash line marks the original image contrast. Thin dashed lines from the right to the left are for the noise intensities 1, 2, 4, 5, 6, 10, 20, 40, 100, 300 and 1000 times σ_I/3. The thick line connects points corresponding to the maximal. Four points marked by diamonds correspond to composite images shown in fig.4. Points marked by the square and the triangle correspond to images obtained in the GIMP software after histogram equalization and histogram stretching of the image Track1 respectively. The thick dashed line corresponds to* ⟦σ=σ⟧ *_I and images with equalized histograms. The right-side plot is the magnification of the lower plot.*

It allows one to select properties of the optimal output composite image with respect to specific applications and to obtain the noise intensity and the number of binary representations (the number of threshold elements) to get this image. We call this plot as CvC plot i.e. the correlation versus contrast plot. This plot for output composite images obtained for the image in *Figure 8.3* is shown in *Figure 8.4*. The noise intensity is treated in this plot as a parameter which connects the normalized rms contrast of the output composite image to its correlation coefficient with the original input image. In CvC plot each thin line is for different number of binary representation forming the output composite image. The noise intensity varies along these lines. The noise intensity along dashed lines is the same. All thin lines start at the same point corresponding to zero noise intensity σ=0 and ξ(i,j)=μ_I. In this case, grey level of pixels satisfying the condition $Image_o(i,j) < \mu_I$ is set to 0 and grey level of other pixels is set to 255. As an example of underwater background acoustic image, *Figure 8.5* shows three composite images I_c

composed of $N = 64$ binary representations for noise intensities equal to equal to $20.24 \approx 2\sigma_I$, $13.49 \approx \frac{4\sigma_I}{3}$ and $6.75 \approx \frac{2\sigma_I}{3}$ respectively.

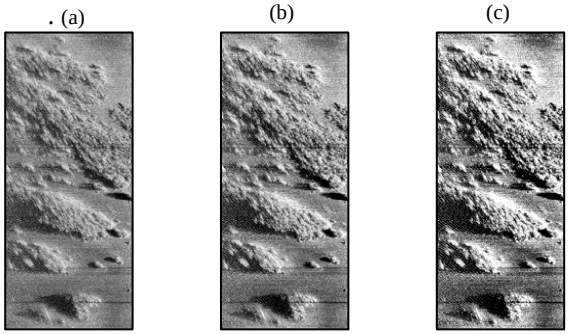


Figure 8.5: Composite output images with N = 64 binary representations for noise intensities σ equal (a) $20.24 \approx 2\sigma_I$, (b) $13.49 \approx \frac{4\sigma_I}{3}$ and (c) $6.75 \approx \frac{2\sigma_I}{3}$.

Images in *Figure 8 5* (a) (with ρ = 0.945 and C≈0.247) and *Figure 8 5* (b) (with $\rho \approx 0.964$ and C≈0477) have higher similarity with the original image. On the other hand, in the image in *Figure 8 5* (c) (ρ = 0.949 and C≈ 0.345) light and dark intensities are more noticeable. These three images show the importance of optimizing both the correlation coefficient ρ and the contrast C to select the output image that satisfies the task to be solved (Sousa et al. 2016).

As our main research line is related to pipeline track survey, we choose the pipeline Track1 image shown in *Figure 8.3* to present the method. The composite image has the maximal normalized rms contrast C=1, because the average and the median of pixel grey levels for the image Track1 are the same. The correlation coefficient between the black and white image for C=1 and the original image Track1 is ρ=0.8, so some features of the original image has been lost. All thin lines for more than 1 binary representation have a maximum for a non-zero level of noise.

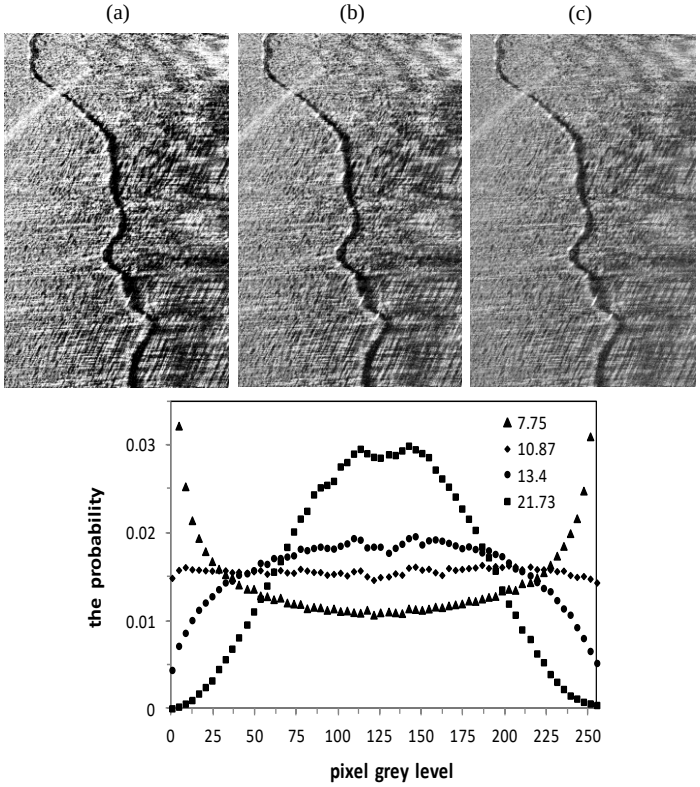


Figure 8.6 *Images composed with 63 binary representations of the image Track1 for noise intensities 7.75, 13.4 and 21.73 (from the left to the right) and their histograms below the images track.*

This maximum represents the output composite image for which the correlation coefficient with the original input image is maximal. The thick line connects these maxima. One can notice that in fact there is a range of noise intensities near each maximum in which the correlation coefficient changes slightly, whereas the rms composite image contrast changes a lot. Let's consider in *Figure 8.6*, the range of noise intensities $\sigma = (\frac{2\sigma_I}{3}, 2\sigma_I)$, for the output image composed of 63 binary representations, i.e. 64 possible pixel grey levels. This range corresponds to the upper part of the thin line above dotted horizontal line drawn for ρ=0.945. The correlation coefficient at the maximum of this line is at ρ=0.962, so ρ changes in this range less than 2%. The correlation coefficient in the considered range is high, so the output composite image has almost all features of the original image in this case. The normalized rms contrast at the maximum is at 0.52, so it is around 7 times higher than the normalized rms contrast of the original image. In the considered range, C decreases from 0.686 to 0.376 with increasing noise intensity, i.e. about 45%. This affect the visual perception of the output composite images with approximately the same ρ. In *Figure 8.6*, we present three composite images and their histograms corresponding to points denoted by diamonds in *Figure 8.4*. The left and the right image (a) and (c), have ρ=0.945 and rms contrast C=0.376 and C=0.686 respectively. The image in the middle (b), has the maximum correlation coefficient ρ=0.962 and the contrast C=0.52. The form of the image histogram depends on the noise intensity. Histograms of composite images obtained for the noise intensity below $\sigma = \sigma_I$ increase near pixel grey levels 0 and 255. The low noise intensity thus moves pixel grey levels of the image to their extreme values. Histograms of composite images obtained for $\sigma > \sigma_I$ have the maximum in the middle of the grey scale. The higher is the noise intensity the faster is the decrease of the histogram from its maximum. The range of grey levels used in the output composite image decreases with increasing the noise intensity. The histogram of the composite image obtained for $\sigma = \sigma_I$ is flat. Histograms of the left image and the middle image in *Figure 8.6*, expand over the whole grey scale and many details are better visible. On the other hand, very bright points on the pipeline of the right image are better visible. Images in *Figure 8.6*, show the importance of both the correlation coefficient and rms contrast in the output image optimization and selection. Histograms of output composite images corresponding to points along the thick dashed line in *Figure 8.4* are flat up to 256 of their binary representations. When the number of binary representation is large and increases, these histograms tend to the histogram of the image obtained from the image Track1 after its histogram equalization in the GIMP software. The operation of histogram equalization is performed by remapping graylevel of the image Track1, based on its own graylevel probability distribution shown in *Figure 8.3*. The image Track1 graylevel probability distribution can well be fitted by the Gaussian distribution with μ=μ_I an $\sigma = \sigma_I$ is used as the noise distribution for points along the thick dashed line in *Figure 8.4*. The described variation of the output composite image histogram for images corresponding to points along this line is presented in *Figure 8.8*. The asymptotic output image composed of very large number of binary representations is the same as the image obtained from the image Track1 after remapping its grey levels by the noise cumulative distribution function multiplied by 255. Histogram of the image obtained after such remapping are plotted in *Figure 8.8* for comparison. The asymptotic behaviour of the output composite image was briefly discussed in ref. (Ye et al. 2007).

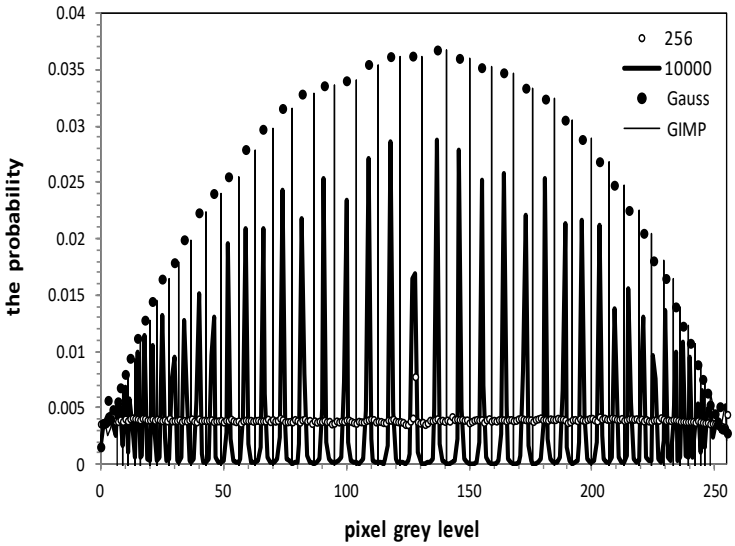


Figure 8.7: *Histograms of output images composed of 256 and 10000 binary representations obtained using the Gaussian noise with μ=μ_I and σ=σ_I. Histograms of images obtained from the image Track1 after its histogram equalization in GIMP software and after remapping of its grayscale by the noise cumulative distribution are shown for comparison.*

The normalized rms contrast C and the correlation coefficient ρ for output images composed of 1000 and 10000 binary representations and images obtained from the image Track1 after its histogram equalization in GIMP software and after remapping of its gray levels by the noise Gaussian distribution are practically the same. Points corresponding to these images in *Figure 8.4* coincide with the point denoted by the square. In *Figure 8.4*, we denoted by the triangle the point corresponding to the image obtained from the image Track1 after its histogram stretching in GIMP. One can notice that the output image with nearly the same C and ρ was obtained for 10000 binary representations and the Gaussian noise. The number of pixel grey levels increases with the number of binary representations and the correlation coefficient of the output composite image with the input image increases. For small number of binary representations, the maximum of the correlation coefficient increases fast with their number. For the number of binary representations above 256 this increase becomes very slow. Moreover, in *Figure 8.4*, lines for different number of binary representations overlap in a certain range of noise intensities. This range allows one the selection of the minimum number of threshold elements used in the algorithm to get the output image with nearly the same C and ρ. This justifies different selection of the minimal number of comparators in different applications (Hari et al. 2009b), (Yang 1998), (Yang 1998), (Ryu et al. 2011), (Jha et al. 2012) and (Jha et al. 2012).

In the *Figure 8.4*, points on thin lines have been obtained for noise intensities from zero up to very high noise values $\sigma \approx 100\sigma_I$. For very high noise intensities all pixel grey levels concentrate very close to 127.5 and the rms contrast of the output image drops below the rms contrast of the original image. In this case, there are output images with high correlation coefficient and very small contrast, like e.g. the output image obtained for 10000 binary representations and the noise intensity $\sigma = 73\sigma_I$. Its correlation coefficient ρ=0.7 and the contrast C=0.016, i.e. it is more than 5 times lower than the normalized rms contrast of the input image Track1. Details in this image are almost invisible what shows that the correlation coefficient cannot be the only metrics quantifying the output image properties.

8.5.1 The Effect of Moving the Noise Mean on The Output Composite Image

The CvC plot in *Figure 8.4* was obtained assuming the mean of the Gaussian noise used in the algorithm equal to the average of the image Track1 pixels grey levels. In this section, we present how this plot changes when one moves the noise mean value. By moving the noise mean one changes the position of the threshold with respect to processed image pixel grey levels.

The number of black and white pixels in the binary output image corresponding to the common starting point for thin lines in *Figure 8.4* is the same. This is because the noise mean value coincides with the median value of the image Track1 pixels grey levels. The rms contrast at this point is thus $C = 1$. For different noise mean values, the number of black and white pixels in the black and white image corresponding to the starting point of thin lines in CvC plot is different and the rms contrast of this image is $C < 1$. In CvC left plot in *Figure 8.8* the thick dashed line connects positions of starting points for different noise mean values. Both C and ρ for black and white images corresponding to these points decrease when the distance of the noise mean value to the median of pixel grey levels of the input image increases. The starting point for thin lines obtained by applying the noise with μ above the highest pixel grey level in the image Track 1 has coordinates (0,0). In *Figure 8.8* we present CvC plots obtained for the Gaussian noise mean $\mu = \mu_I + 1.5\sigma_I$ (the left plot) and $\mu = 255$ (the right plot). By plotting some dotted lines obtained for $\mu = \mu_I$ one can notice that at high noise intensities any thin line obtained for the noise with any mean value overlaps with the dotted line obtained for the same number of binary representations. For very large noise intensities the output image with the same C and ρ can thus be obtained using noises with any mean value. This property and the position of the starting point divide thin lines in two groups: lines turning to the right and to the left of the starting point. This division of lines helps one e.g. to select the number of binary representations above which thin lines start to overlap in a certain range of noise intensities, or to predict the contrast range of the output image for a given number of binary representations. Dotted lines with points having the contrast C for high noise intensities near the contrast C of the thin lines starting point determine the number of binary representation for which the output composite images have very small rms contrast range. In the left plot of *Figure 8.8* thin line obtained for 6 binary representations first turns to the right and then to the left. The contrast modification of the output image composed of 6 binary representations using noise is in this case very limited.

Assuming $\mu = \mu_I$ one assigns pixels with grey level μ_I in the input image in the middle of the grey scale in the output image, regardless of the value of μ_I. For images with high and low μ_I values this means a significant change in the mean brightness of the processed image.

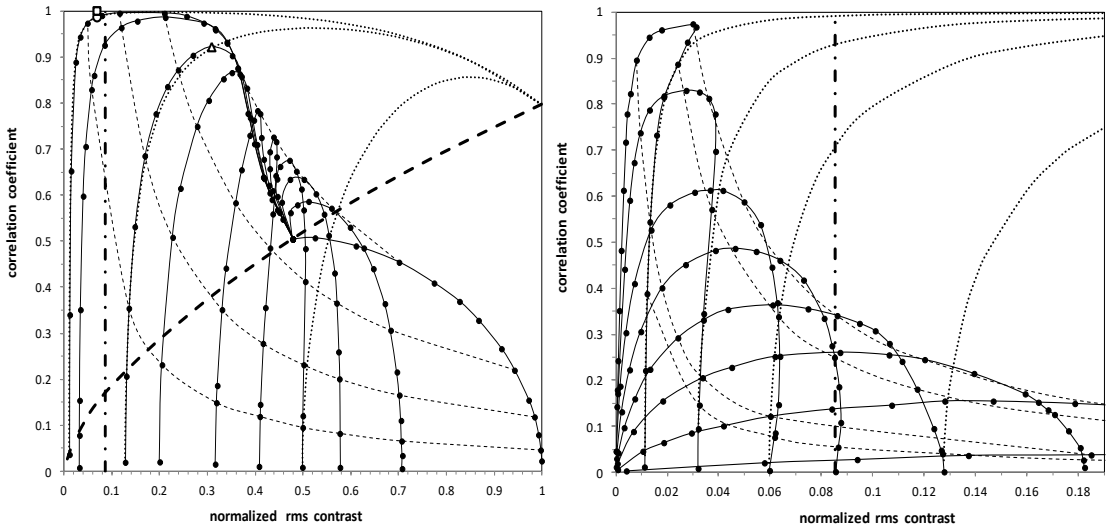


Figure 8.8: *The CvC plot for the Gaussian noise with* $\langle\!\langle \mu=\mu \rangle\!\rangle\_I+1.5\sigma\_I$ *(upper plot–UP) and* $\mu=255$ *(lower plot – LP). Thin lines from the right to the left are for 1, 2, 3, 4, 6, 8, 25, 63, 1000, 10000 (UP) and 1, 10, 30, 63, 128, 256, 1000 and 10000 (LP) binary representations. The vertical dot-dash line marks the original image contrast. Thin dashed lines from the right to the left are for the noise intensities 3, 10, 20, 50 (UP) and 40, 100, 20, 500 (LP) times in/3. Dotted lines are for the Gaussian noise with* $\langle\!\langle \mu=\mu \rangle\!\rangle\_I$ *and 4, 63, 10000 (UP) and 63, 256, 1000 and 10000 (LP) binary representations. The thick dashed line denotes positions of thin lines starting points for different noise mean values. Points denoted by unfilled square and circle correspond to the sonar image after modification of its brightness (+50) in the GIMP software and to the composite image with approximately the same properties. Composite images corresponding to the point denoted by the triangle are shown in Figure.8.4.*

For fixed C and ρ values of the output image one can adjust its brightness by moving the noise mean value. In *Figure 8-9* we present two output images with $C \approx 0.31$ and $\rho \approx 0.92$ obtained using two different noise mean values: $\mu = \mu_I = 81$ and $\mu = \mu_I + 1.5\sigma_I = 97.3$. The point denoted by triangle in *Figure 8.8* in which dotted and thin lines overlap corresponds to these images. The average intensity value is 127.5 for the left image and 64 for the right image, so the left image is much brighter. This shows the importance of the noise mean value used in the algorithm on the outlook of the output image.

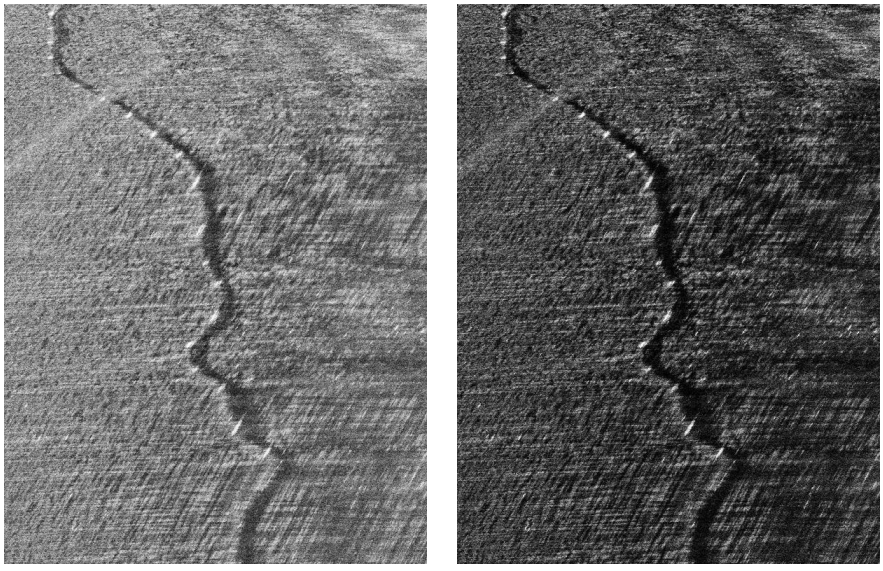


Figure 8.9: *Images composed with 63 binary representations obtained for the Gaussian noise with* $\boldsymbol{\mu = \mu_I}$ *and* $\boldsymbol{\sigma = 8\sigma_I/3}$ *(left plot) and* $\boldsymbol{\mu = \mu_I + 1.5\sigma_I}$ *and* $\boldsymbol{\sigma = 2\sigma_I}$ *(right plot) corresponding to the point denoted by the triangle in fig.23.*

To show that using the Gaussian noise one can modify the image brightness similar as in the image processing software, we have modified the Track1 image brightness in the GIMP image processing software by +50. The obtained image and its histogram are shown in *Figure 8-10* and the point corresponding to this image is denoted by the square in *Figure 8.8*. Using the Gaussian noise with $\mu = \mu_I + 1.5\sigma_I$ and $\sigma = 12\sigma_I$, we have obtained the

output image composed of 10000 binary representations with the same outlook and the histogram shown in *Figure 8.10*. The point corresponding to this image is denoted by the circle in *Figure 8.10*. One can notice that C and ρ values for both images obtained by two different methods are approximately the same. The average value and the standard deviation of histograms in *Figure 8.10* are the same, but the histogram obtained in the GIMP software have several peaks. The histogram of the composite image is smooth because of finite number of binary representations. The GIMP procedure does not redistribute pixel intensities to produce smooth histogram. It is useful for highlighting borders and edges between different objects, but it reduces intensity variations within these objects. The image brightness modification using the noise in some applications can thus give better results.

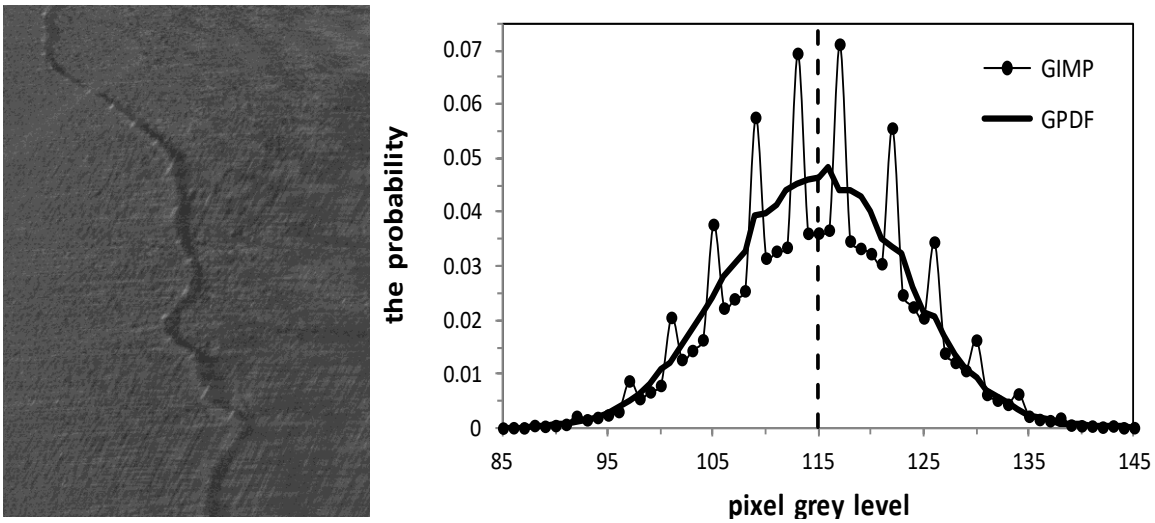


Figure 8.10: *The sonar image after modification of its brightness (+50) in the GIMP software and its histogram. The thick line represents the histogram of the output image composed of 10000 binary representations using the Gaussian noise with* $\boldsymbol{\mu = \mu_I + 1.5\sigma_I}$ *and* $= \mathbf{12}\boldsymbol{\sigma_I}$ *The dashed line denotes the position of the average grey level of both histograms.*

The right CvC plot in *Figure 8.8* was obtained for the Gaussian noise with the mean value μ=255. This value is the highest grey scale level and the threshold value in the technique of image contrast modification applied in this paper. Assuming μ=255 in the algorithm is equivalent of adding the Gaussian noise with the zero mean to input image pixel grey level. The white Gaussian noise with zero mean is commonly added noise to the input signal in many experimental and theoretical investigations of the stochastic resonance and suprathreshold resonance phenomenon. The CvC plot shows that in this case it is impossible to obtain the output composite image with high correlation coefficient and with the rms contrast higher than the rms contrast of the original image Track1. This is because the original image pixel grey levels are too distant to the threshold. All curves start at the point (0,0) and the output image one can obtain using the noise with the intensity $\sigma = 3\sigma_I$. The output–input correlation coefficient becomes bigger than 0.5 for the output image composed of the minimum 125 binary representations and it tends to 0.97 for 10000 binary representations. The rms contrast of output images with high ρ values is however very small.

8.5.2 The Effect of Noise Distribution on the Output Composite Image

In previous sections the output composite image was obtained using the Gaussian white noise (GPDF). In this section, we present results of using non-Gaussian noise. The non-Gaussian noise the most frequently used in the image enhancement algorithms based on the stochastic resonance is the Uniform noise with the rectangular probability density function (RPDF) (Yang 1998), (Ye et al. 2007) and (Jha et al. 2012). We selected this noise and the Laplacian noise (LPDF) which is widely used in sonar applications (Hari et al. 2009b). The RPDF, GPDF and LPDF noises are members of noise family described by the generalized Gaussian probability density function (GGD) defined as expressed in *equation 8.6*.

$$f(\xi) = \frac{A}{\sigma} exp\left(-B\left|\frac{\xi-\mu}{\sigma}\right|^{p}\right) \qquad \textit{eq. 8.6}$$

where $A = (p/2)\Gamma[3/p]^{1/2}/\Gamma[1/p]^{3/2}$, $B = [\Gamma(3/p)/\Gamma(1/p)]^{p/2}$ and Γ(.) is the gamma function. One has for p=2 the Gaussian distribution, for p=1 the Laplace distribution, for p→+∞ the Uniform distribution and for $p \to 0^{+}$ one has the impulse probability function at ξ=μ. The function $f(\xi)$ has tails that are heavier than normal for p<2 (leptokurtic distribution) or lighter than normal for p>2 (platykurtic distribution). At present, GGD noise is widely used in engineering in the area of signal processing (Yu et al. 2012) and to model the impulsive heavy-tailed noises like e.g. the environmental noise in the ocean (Machell et al. 1989a).

The random variable ξ which has Laplacian distribution with the mean μ and the intensity σ we generated from the following eqution:

$$\xi = \mu - \sigma\, sgn(U-0.5) ln(1-2|U-0.5|)/\sqrt{2} \qquad \textit{eq. 8.7}$$

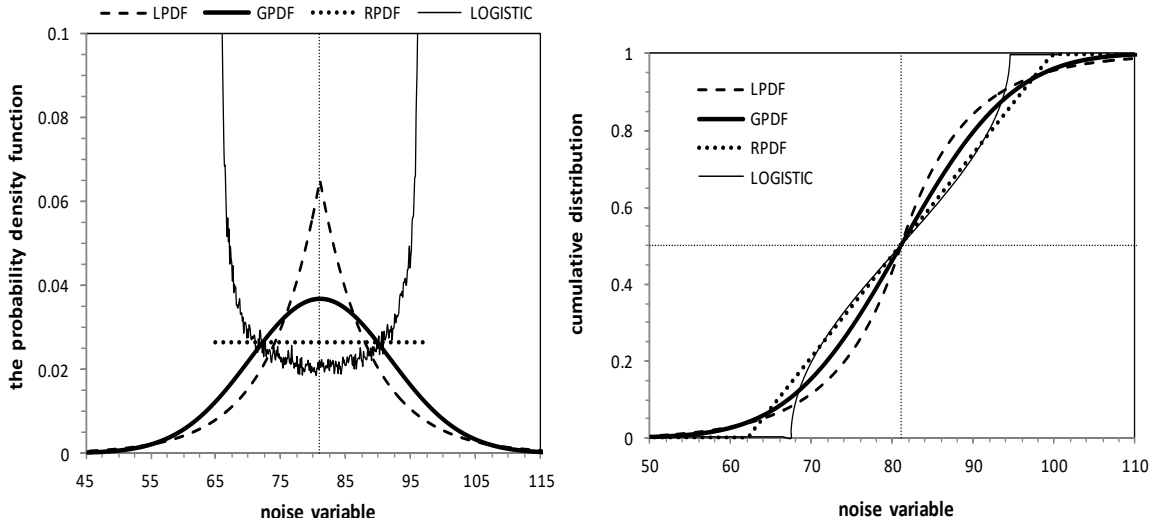


Figure 8.11: *The probability density functions and cumulative distributions for Laplace, Gaussian and Uniform noises and for the chaotic noise generated by logistic time-series (tails are up to around 0.37). The mean value is 81 and the standard deviation is 10.87*

In the *equation 8.6*, *U* is a random variable drawn from the uniform distribution in the interval (0, 1). This formula follows from the inverse cumulative distribution function. The probability density functions for RPDF, GPDF and LPDF noises and their cumulative distributions are shown in *Figure 8.11*. The GPDF and LPDF functions are maximal for ξ=μ, whereas RPDF function is flat.

In dynamical environments, whose complex behaviour is neither purely random nor perfectly predictable the stochastic noise is often replaced by the chaotic time-series. The simplest nonlinear equation generating the chaotic time-series is the logistic map (May 1976)(Sprott 2003). It was used to present the chaotic resonance in bistable system (Ippen et al. 1993)(Castro & Sauer 1997). We selected as the fourth noise added to the original image the chaotic noise defined by the logistic time series (LOGISTIC).

$$\eta_l = 4\eta_{l-1}(1-\eta_{l-1}) \qquad \textit{eq. 8.8}$$

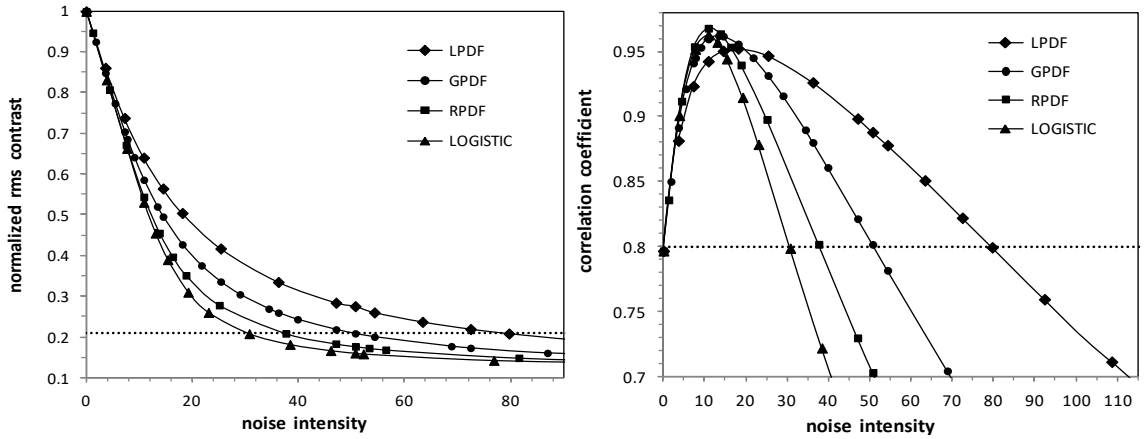


Figure 8.12: *The dependence of the normalized rms contrast and the correlation coefficient of the output image composed with 63 binary representations on the noise intensity for different noise distributions and the noise mean* $\boldsymbol{\mu = \mu_I}$. *Dotted lines denote parameters of the composite image corresponding to unfilled circle point in Figure 8.13.*

We obtained the successive numbers η_l of the logistic time-series by iterating the following map assuming any random number in the range (0, 1) with 0.25, 05 and 0.75 excluded as the initial number η_0. Numbers η_1. for l=1,2, … form a chaotic set of numbers between 0 and 1. The probability density function of these numbers is symmetric and minimal at 0.5 and it fast increases close to 0 and 1. It is thus very different from the generalized Gaussian probability density functions. The probability density function and the cumulative distribution for the chaotic variable ξ defined by logistic time-series with the mean $\mu = \mu_I$ and the standard deviation $\sigma = \sigma_I$ is shown in *Figure 8.11*.

In *Figure 8.12* we present the dependence of the output-input correlation coefficient on the noise intensity obtained for 63 binary representations using noises with different distributions and the mean $\mu = \mu_I$. This dependence is typically plotted to present the suprathreshold stochastic resonance (Mcdonnell et al. 2002). The correlation coefficient has the maximum for nonzero noise intensity for all noises. We also present in *Figure 8.13* the dependence of the normalized rms contrast of the output composite image on the noise intensity. This contrast is the biggest for Laplacian noise and the smallest for chaotic noise for all noise intensities. The CvC plot obtained in this case is shown in *Figure 8.13*. Points corresponding to RPDF, LPDF and LOGISTIC noises are on the line for GPDF noise. For a given number of binary representations, the output image with a given rms contrast and the correlation coefficient can thus be obtained using any considered noise. In *Figure 8.12* we denoted by dotted lines values of C and ρ corresponding to the unfilled circle point in *Figure 8.13*. The intensity of noises used to obtain the output image corresponding to this point are the following: 79.69 (LPDF), 50.71 (GPDF), 37.64 (RPDF) and 30.74 (LOGISTIC). In *Figure 8.13* we marked by unfilled symbols positions of points corresponding to the output image obtained using different noises having the same intensity 50.71. The difference between noise intensities necessary to get a given composite image becomes very pronounced for high noise intensities.

In *Figure 8.14* (left) we present CvC plot obtained for different noises with the mean $\mu = 128$. We assumed this mean value to show the relation between image contrast modification using the noise and applying the standard contrast modification procedure in the GIMP image processing software. By thick dashed line we denoted C and ρ values of images obtained by modifying the contrast of the image Track1 in GIMP software. One can observe that points corresponding to output images obtained using RPDF noise for 1000 and 10000 binary representations are on this thick dashed line in a large range of noise intensities.

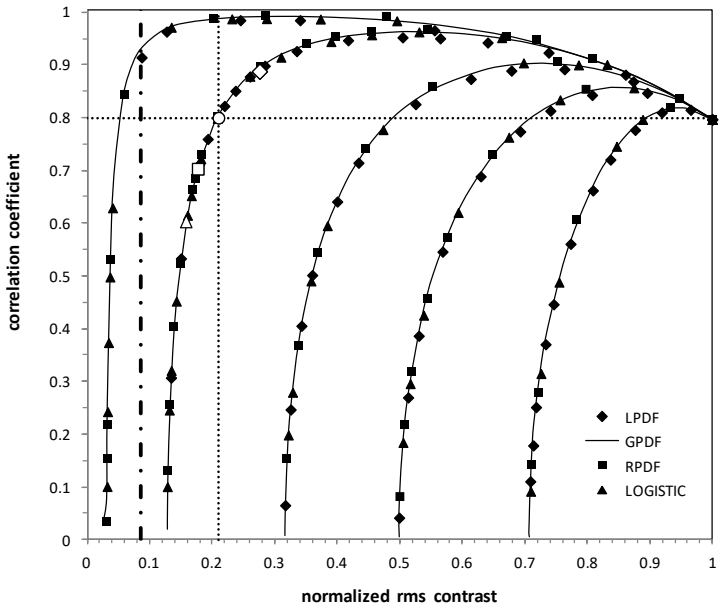


Figure 8.13: *The CvC plot for the noise with* $\boldsymbol{\mu = \mu_I}$ *and different probability distributions. Lines from the right to the left are for 2, 4, 10, 63 and 1000 binary representations. The vertical dot-dash line marks the original image contrast. Points marked by unfilled symbols correspond to the noise intensity* $\boldsymbol{\sigma = 14\sigma_I/3}$ *(the circle represents GPDF noise).*

This range increases with the number of binary representations. In *Figure 8.14* (rigth) we also present CvC plot obtained for different noises with the mean $\mu = 255$ i.e. the value of the threshold assumed in this paper. This case is equivalent to typical observations of the stochastic resonance phenomenon in which the noise with zero mean is added to subthreshold image. One can notice that using only the LOGISTIC noise one can get the composite image with the contrast higher than the contrast of the original image.

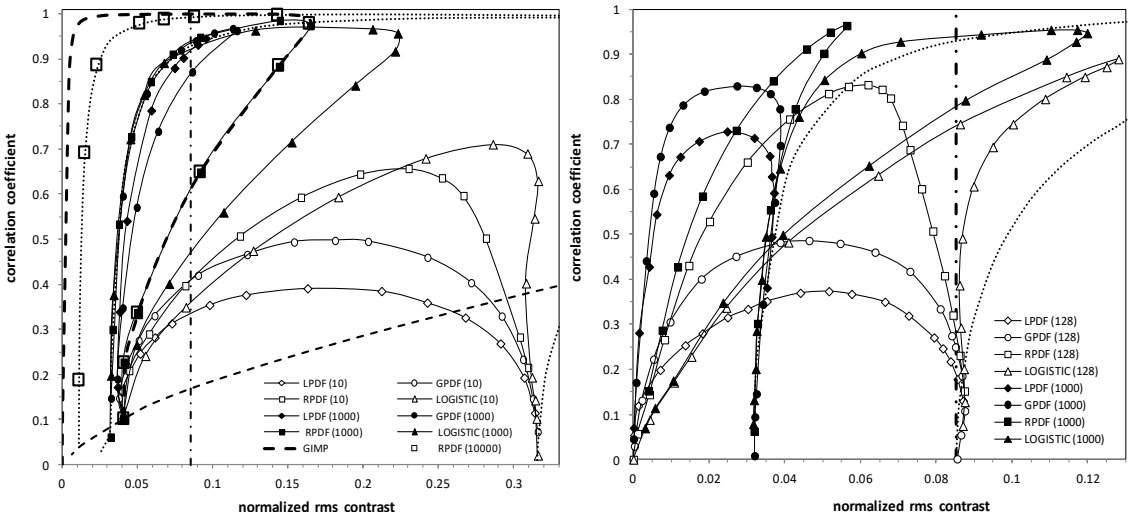


Figure 8.14: *The CvC plot for the noise with μ=128 (left plot - LF) and μ=255 (rigth plot – RG) and different probability distributions. The number of binary representations is given in brackets. The thick dashed line corresponds to the sonar image after modification of its contrast (from +127 to -127) in the GIMP software. The thin dashed line denotes positions of lines starting points for different noise mean values. The vertical dot-dash line marks the original image contrast. Dotted lines are for the Gaussian noise with* $\boldsymbol{\mu = \mu_I}$ *and 10, 1000, 10000 (LF) and 128, 1000 (RG) binary representations*

Lines in *Figure 8.14* obtained for different noises and the same number of binary representations start at the same point corresponding to zero noise intensity and they do not overlap in a wide range of noise intensities. However, for high noise intensities these lines are close to and finally overlap with lines presented in *Figure 8.4* obtained for the GPDF noise with the mean $\mu = \mu_I$. The noise mean and distribution thus become unimportant for very high noise intensities. The correlation coefficient of the output composite image obtained in this case with the input image is however small. Using different noises to modify the image Track1 contrast we have the same (*Figure 8.13*) or different (*Figure 8.14*) output composite images.

8.6 Conclusion

We applied the technique of image contrast modification using non-dynamical system of threshold elements and the noise to digital grayscale sonar image. The threshold was set at the maximum of the grey scale. We presented results obtained for the chaotic noise generated by the logistic map and for Gaussian, Uniform and Laplace noises. The contrast and quality of the output image was quantified by the normalized rms contrast and the output-input correlation coefficient. These quantities are the most important for visual image perception. We proposed CvC plot (correlation vs. contrast plot) to show the effect of the noise properties and the number of binary representations on the output composite image. Lines in CvC plot determine the correlation coefficient and rms contrast of the output composite image for a given number of its binary representations and varying noise intensity. They allow to select the number of binary representations and the noise properties to get the output image with required contrast and quality. This is important e.g. to minimize the time of image processing. For very high noise intensities lines obtained for any noise mean value and distribution but the same number of binary representations overlap. The point in CvC plot which belongs to lines obtained for the same number of binary representations using the noise with different mean values allows one to select the output image brightness having the same its contrast and quality. The dependence of the output-input correlation coefficient on the noise intensity has the maximum for any considered noise, any position of its mean value (i.e. both for subthreshold and suprathreshold images) and any number of binary representations bigger than 1. Varying the noise intensity, one modifies also the output image contrast. The range of contrast modification depends on the number of binary representations and on the noise, mean value and distribution. We showed that for high number of binary representations image contrast modification using noise can replace image contrast and brightness modification procedures of the GIMP image processing software. Images obtained using both techniques have the same contrast and quality. Their histograms have the same mean and standard deviation, but the histogram of the composite image is smoother because of finite binary representation error. For increasing number of binary representations, the output composite image tends to the image obtained by remapping pixel grey levels of the input image by the noise cumulative distribution function.

9.0 Study of Reliability of Unsampled Bathymetric Points Estimations Using Spatial Statistics, a Geostatistics Approach

The technical and scientific community makes efforts to discover or increase knowledge about hidden properties in some materials, structures or environmental behaviour. To this end, some techniques aim to explore and understand the environmental objects or properties around us, based on discrete samples for the representation of a continuous surface. It is considered the specific case of spatial sampling involving measurements of the elevation level of an underwater surface or submerged topography, this process is called bathymetry, see chapter 3. The resolution of the data acquired in the sampling, especially in geospatial data, will depend on several factors such as sensor resolution, the coverage area of the inspection, the limitation of the operating time and the memory of the equipment to support data collection.

A variety of devices and technologies can be used to acquire discrete data through a sampling process. The researches in the area of remote sensing seek to develop a model to explain or predict some phenomenon, or explain characteristics of the surface under study, in the context of the inverse problem, from a limited set of samples. Many techniques are applied to extract characteristics of a topographic relief considering samples taken by sensors that are above the surface of the water. Among them, we can highlight an optical acquisition system, technically known as LIDAR, operating on a variety of light frequency spectra (Höfle et al. 2013) and Radio Wave Detection and Ranging (Wandinger 2005).

In recent years, the use of laser scanning systems installed in airborne and terrestrial systems and vehicles have become the most widely used remote sensing devices for acquiring topographic data to cover large regions above sea level (Shan & Toth 2008) and (Hell 2011). Nowadays, the research focuses on the sensor fusion paradigm, using more than one type of technology to acquire the data. The objective is to reduce the uncertainty of estimates related to the studied characteristics by comparing information from the same target or acquired by two distinct devices or by adding complementary auxiliary information from two or more different sensors (Hermosilla et al. 2005), (Lockhart et al. 2008), (Coiras et al. 2003) and (James & David 2008).

Despite the availability of all the technologies and techniques mentioned above, it still configures as a problem when the object of study is a submerged landscape. The problem comes from the physical restrictions at high frequencies, which results in the dissipation of the energy flow when it is transmitted through the water, resulting from the interaction with a non-zero conductivity medium (Lurton 2010), where the dielectric properties can affect electromagnetic waves, avoiding that the transport of information is adequate, especially in sea water. These effects were discussed in the work of (Jiang et al. 2008) and (Karagianni 2015), which discussed the problems of wireless networks in underwater environments, addressing some physical constraints involving the information transmission process and the influences of general physical properties in this medium on electromagnetic waves of 2.4 GHz.

In the work of (Karagianni 2015), it was explained that, in underwater environment, the electromagnetic waves used in optical systems and by devices based on lasers undergo a strong attenuation/absorption. The main causes are due to the intrinsic impedance, conductivity and molecular relaxation of the medium, which have influence on the delay of the particles of the medium to respond to the change of wave field, resulting in energy lost to the medium. Despite this, some radar-based technologies and LIDAR systems can be used by aircraft and satellites to estimate underwater topographic relief, but with certain known limitations (Lockhart et al. 2008), (Gao 2009) and (Penrose et al. 2005).

In order to solve the inverse problem for the construction of the topographic maps, objective after the acquisition of data in underwater environment, acoustic devices based on sonar technology are currently applied. These devices were originally developed to assist ships and boats in navigation, to collision avoidance, and are also adapted for other civil and military applications (Medwin & Clay 1998) and (Peyvandi et al. 2011). In fact, underwater acoustic propagation is the most widely used physical principle to perform underwater communications (Nelson 2009), (Stojanovic 2008) and (Stojanovic & Preisig 2009), as well as applications in remote sensing (Tacconi & B 1981). Some techniques use specific characteristics of acoustic devices and characteristics of elastic (Medwin & Clay 1998), (Daniel R. Raichel 2006) and (Taroudakis 2013), to perform the

underwater signal processing and solve the inverse problem involving acoustic signals (Stojanovic 2008)(Vaccaro 1998).
The sonar devices used in bathymetric applications employ sensing energy principles based on backscattering of acoustic waves when impacted on a target. The propagation of acoustic waves in a medium can undergo refraction and absorption according to the physical characteristics of the fluid or medium in question (Medwin & Clay 1998), (Lurton 2010) and (Daniel R. Raichel 2006). The temperature, density and salinity in an aquatic environment are important factors that influence the propagation of the acoustic signal (Lurton 2010) and (Medwin & Clay 1998). The interaction between the medium and the propagated energy allows us to apply a digital processing paradigm of acoustic signals that, in some cases, presents analogies when compared to conventional optical systems. In the case of sonar images, for example, techniques that involve the relationship between shading and highlights of objects and regions can be used to assist in the processing of underwater relief reconstruction tasks and in the detection of submerged objects submersos (Coiras et al. 2007), (Rongxing Li 1991), (Bikonis et al. 2005) and (Rongxing Li 1992).

For the survey of bathymetric data, sonar devices may be transported by a conventional research vessel, surface robotic vehicle (SRV) or underwater robotic vehicle (ROV or AUV). Each of them introduces specific uncertainties in data acquisition. Technicians and researchers generally need to consider this fact in the stages of conditioning or pre-processing of data (Baker & Li 2002), (Johnston 2002) and (Elmore et al. 2009). The type of described trajectory for the sampling of an underwater area depends on the sensor, the vehicle, the environmental characteristics and, in some cases, the final applications of the data collected (Lurton et al. 2015) and (Penrose et al. 2005).

In relation to the bathymetric survey, it is common to use two types of sonars called Single Beam Echo Sounder (SBES) and Multi Beam Echo Sounder (MBES) (Haris & Chakraborty 2013) and (Gao 2009). The first device, SBES, consists of a single active transducer element, which emits only one acoustic pulse for the target at any given time, recording the echo received after the backscatter process after impact on the target. Thus, knowing the velocity of propagation of the acoustic wave to the medium, we can infer the distance between the sensor and the target (Hell 2011), (Lurton et al., 2015) and (Brouwer 2008). Some sonar systems can record the echo intensity received, revealing some additional physical characteristics of the target (Lurton et al., 2015). The systems that use MBES are composed of a set of active transducer elements, in a specific physical arrangement. The device will emit an acoustic pulse from each transducer element simultaneously and will receive several echo signals, each signal referring to a transducer element, covering a large survey area, with a better resolution than SBES.
The use of SBES or MBES in a bathymetric survey process is conditioned by some questions such as: need for data accuracy, application, environmental characteristics, level of knowledge of the technicians involved and survey vehicle (IRO, 2005), (Penrose et al., 2004) and (Lurton et al., 2015)). In some cases, SBES is chosen due to economic issues (available budget) and simplicity of work, but the level of detail and precision required may result in an inefficient survey for certain applications. In the case of MBES, although the evolution of the technology of these devices produces better resolution results, this device presents a relatively high cost (Lurton et al., 2015) and (Penrose et al., 2005).

The use of MBES in surveying with underwater autonomous robotic vehicles is restricted to industrial and commercial applications mainly in offshore industries and in some cases in the area of port infrastructures and coastal military security. Bathymetric surveying applications for simpler systems such as small port infrastructures, coastal civil management, rivers and dams are generally performed using SBES. In this case, the sampling will have lower resolution compared to the same operation performed by an MBES. The resolution of the sampling process can be increased through analyses and adequate statistical inferences, aimed at estimating values at non-sampled points.

Statistical analyses are composed of three specific stages: description, indication and estimation (Cressie 1993). In the first stage, the work is directed to the summarization of the data, performed by the exploratory analysis of the data. This stage will be completely defined, when the best theoretical model that allows an adjustment of the data is defined. The second phase is related to the extraction of parameters (indicators), extracted from the theoretical model that adjusts the experimental distribution of data. Finally, in the third phase, we are able to estimate the study variable value in some places or points on the study surface, where we do not have measurements made. The conventional sequence of operations in statistical analysis described in this paragraph will be used in the application of geostatistics to bathymetric data.

Geostatistics is a methodological process to predict values of variables in points not previously sampled. The predictions are based on the structural analysis with the use of spatial statistics, from discrete samples collected on the region of interest. In many cases, the prediction is supported by another auxiliary variable, which has a

correlation with the variable of interest. The geostatistical prediction technique consists mainly of the application of specific spatial statistics algorithms developed by Matheron in his work (Matheron 1970) and (Matheron, G. Kleingeld 1987). The geostatistical method of interpolation or prediction, is based on the algorithm of Kriging, that initially was used to solve the problem of the prediction of the gold reserves. In our research, we applied a geostatistical approach using spatial correlation on the submerged topographic relief, using as basis discrete depth samples.

9.1 Exploratory Data Analysis

In this work, we used a data set composed by 424 sample points extracted from the researched area, shown in the scatter plot of *Figure 9.1*. This scatter plot represents the values of the samples confronted with the geographic coordinates of the survey obtained in the Bay of All Saints (a more detailed description can be seen in Appendix B, item 2.4). The colour of the point expresses the value of the regionalized variable (Depth) of the underwater topographic relief. The deeper topographic values, i.e., higher depression is indicated in light yellow, on the other hand, lower topographic values, represented by shallower regions, are represented in dark blue colour. In the graph, we observed a trend in the transverse direction, that is, specifically in the direction Northeast (NE).

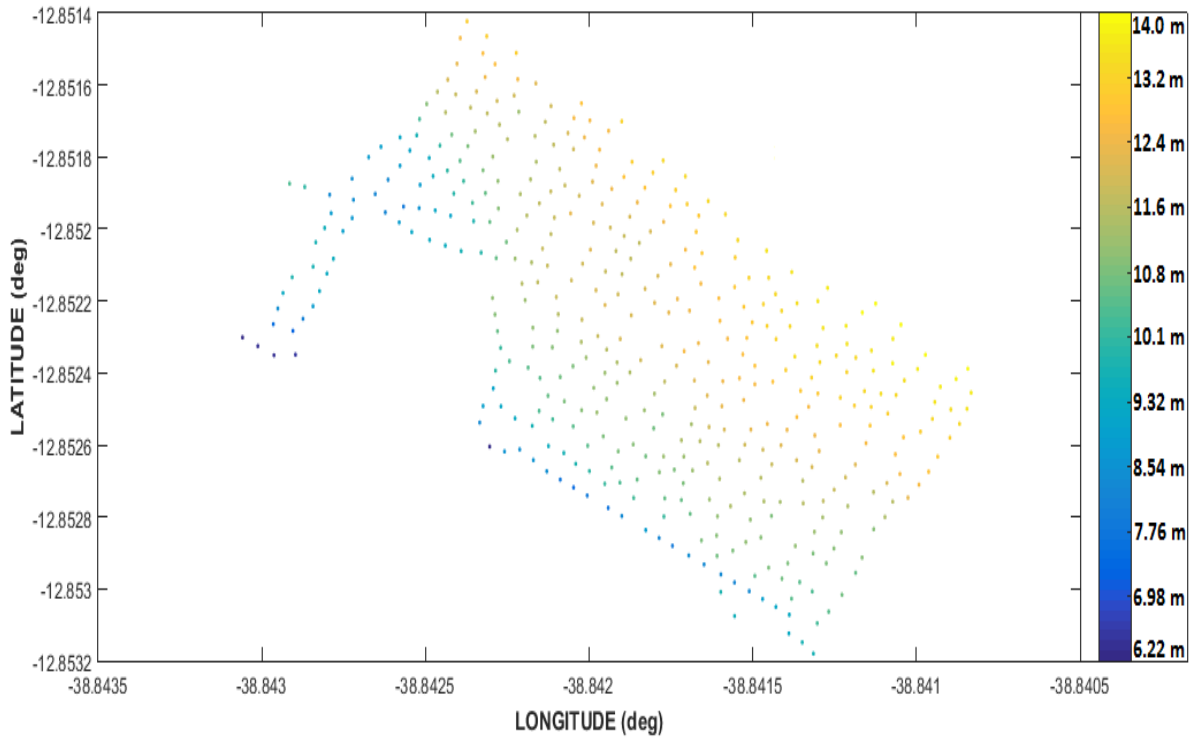


***Figure 9.1**: Bathymetric Dataset Scatter Plot from a SBES*

The hypothesis of considering a trend surface in the direction (NE) was investigated throughout this work. In contrast to a constant average depth assumption, we applied a spatial trend function to vary the mean depth throughout the work, which showed little efficiency over the approach used. Referring to *Figure 9.1*, the sample set does not have a perfectly regular grid, and may vary slightly, the distance from one point to another along the trajectory in the direction of the survey is approximately 7 meters, while the value of the distance between trajectories is approximately 5 meters.

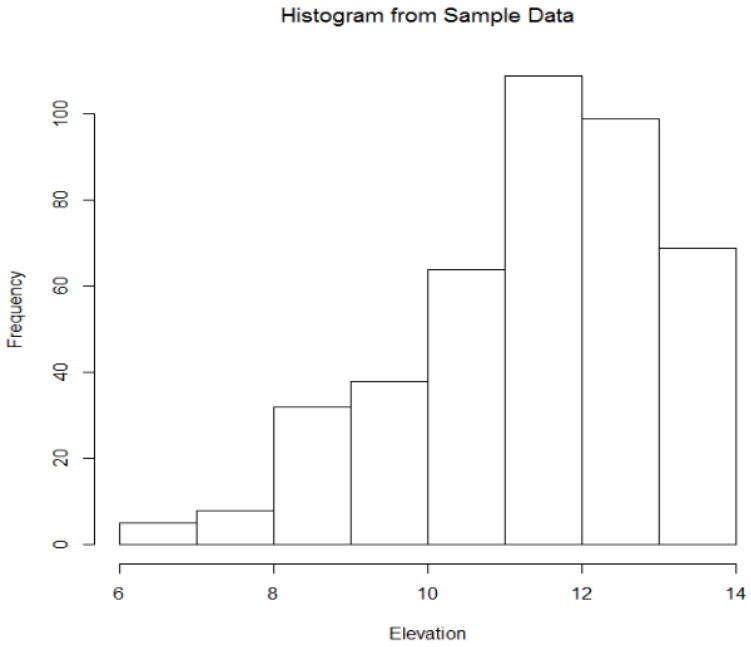


***Figure 9.2**: Frequency distribution of elevation (in meters) sampling.*

The sampled data were represented by a histogram, it is intended to make a first characterization, *Figure 9.2*. In this histogram, we observe that the data do not have symmetrical distribution, i.e., presents a non-Gaussian distribution. Thus, it is not possible to apply conventional and simple statistical algorithms. The data presented a distorted behaviour, with predominance of values for greater depths. In addition, no outliers appear.

However, the histogram analysis is a simple statistical analysis, as shown in *Table 9.1*, by itself does not contain details about spatial structure. In the representation of data through the graph of *Figure 9.2*, it will not be possible to determine changes in spatial variability such as roughness. Thus, the statistical summary through the histogram is a very simple representation, in this work the exploratory data analysis step, which involves the histogram, was performed with the geoR (Diggle & Ribeiro Jr. 2007) and (Ribeiro Jr. & Diggle 2001). The characteristics on sample distribution were discussed in (Cressie 1993)

Consider a set of data with sequentially collected samples through a continuous survey path. The sequence is plotted in *Figure 9.3* **A**, which shows the relationship between the navigation trajectory and the geographic coordinates. In the *Figure 9.3* **B** is shown the relationship between the sample sequence and the bathymetry values.

***Table 9.1:** Dataset Summary*

Dataset Summary					
Geographical Coordinates and Distance					
	Latitude		**Longitude**		**Data Samples Distance (m)**
Minimum	-12.85318		-38.84306		5.04
Maximum	-12.85142		-38.84083		248.41
Data Samples Non-Spatial Statistics (Depth)					
Minimum (m)	**1st. Qu. (m)**	**Median (m)**	**Mean (m)**	**3st. Qu. (m)**	**Maximum (m)**
6.22	10.53	11.62	11.36	12.58	14.00

Consider a set of data with sequentially collected samples through a continuous survey path. The sequence is plotted in *Figure 9.3* **A**, which shows the relationship between the navigation trajectory and the geographic coordinates. In the *Figure 9.3* **B** is shown the relationship between the sample sequence and the bathymetry values.

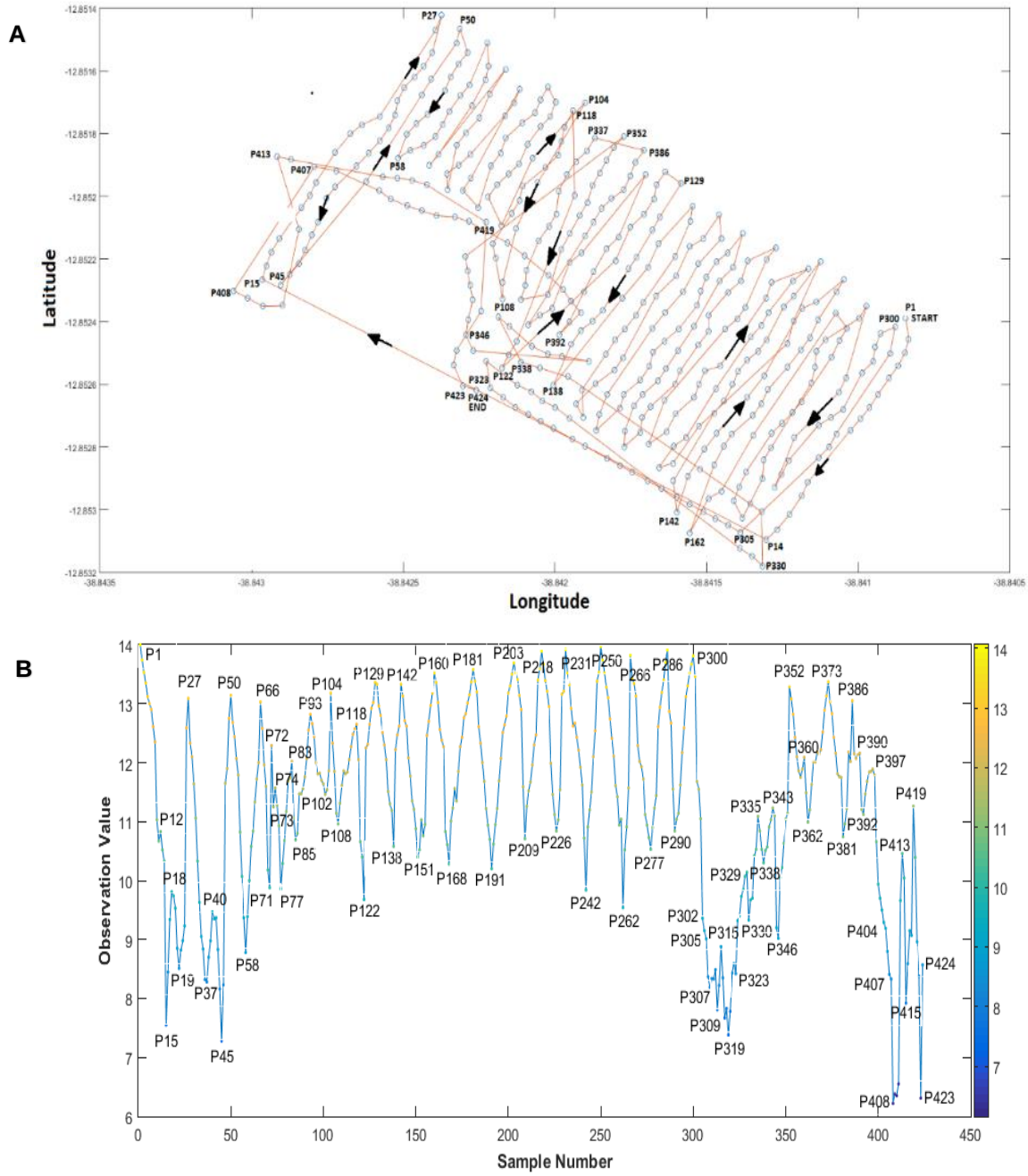


Figure 9.3: *Dataset path line track survey (**A**) and sequence sample (**B**).*

Conventional statistical analysis was initially applied in order to discover any hidden spatial tendency or correlation in the data. It was observed that there were different averages associated with different sites within the sampling region, proving the non-Gaussian distribution. This feature is an evidence of a correlation structure that depends on spatial location. This problem is adequately described in (Cressie 1993), (Diggle & Ribeiro Jr. 2007) and (Goovaerts 1999). In a conventional approach, the path could be to apply a mathematical transformation to solve the problem of a non-Gaussian distribution, allowing the use of common statistical tools in solving the prediction problem. Another way to approach the system is to apply a spatial statistical approach, in our case we use the tools of geostatistics. We performed the exploratory data analysis using a scatter plot in order estimate the

spatial trend, but presenting the dispersion with respect to each geographic coordinate separately. In *Figure 9.4*, we can see the measured depth in latitude (left) and longitude (right). The graphs confirm a clear trend in the Data vs. Longitude, where the trend is confirmed in the direction (NE). Thus, we have one more indication that the assumption of constant mean cannot be applied in the forecast of points not sampled. For best results, we can apply any Kriging methodology based on non-constant mean $\mu(u)$, such as the OK process, consider a model as a trend as KT, or address the problem according to segmentation of the dataset and apply different parameters to Different regions.

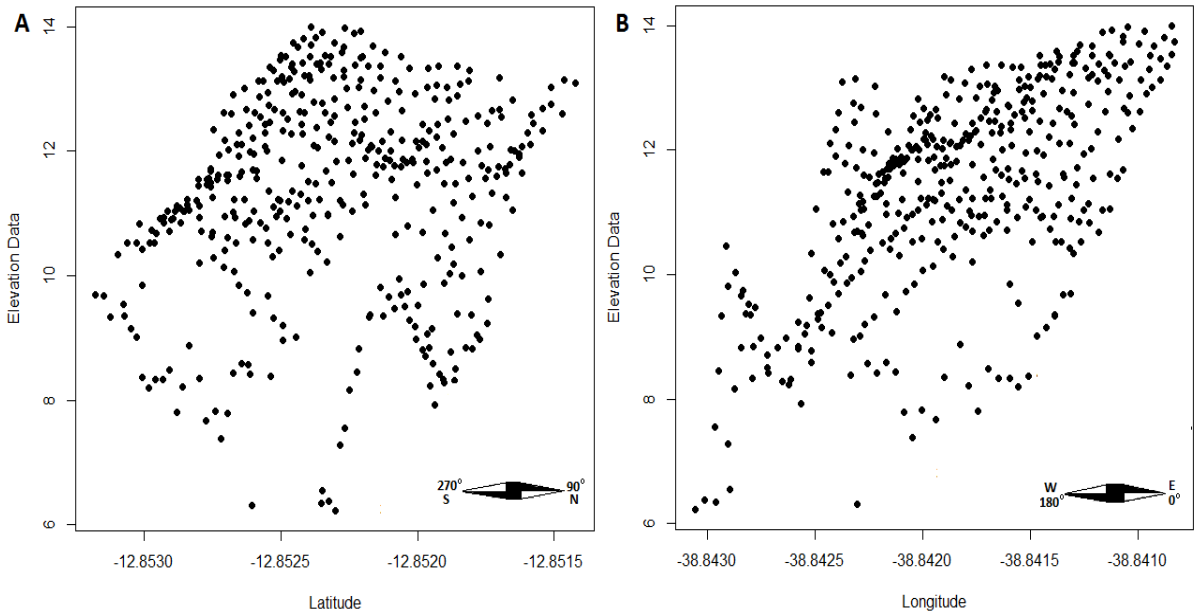


Figure 9.4: *Dataset scatter plot relative to each survey geographical coordinate*

The study presented here was developed using the correlation structure of the regionalized variable considering some patterns and spatial variations, and with a non-constant $\mu(u)$.

The most suitable model for the prediction of data was determined after the analysis of the three Kriging processes, considering the exploratory data analysis under the simple and spatial statistical paradigm. We consider the structure of the mean determined by equation 9.1, according to (Diggle & Ribeiro Jr. 2007).

$$\mu(u) = \beta_0 + \sum_{j=1}^{p} \beta_j(u) d_j(u) \qquad \textit{eq. 9.1}$$

Where β_0 represents the initial average parameter, $\beta_j(u)$ is the average iterative parameter for the real region and $d_j(u)$ is the spatial explanatory variable. The average parameter $\beta_j(u)$ has initial value determined by an ordinary least squares criterion and p is the total number of samples available (Diggle & Ribeiro Jr. 2007).

9.2 Empirical Semivariogram Analysis

The calculations of the function related to the semivariogram can be performed under the assumption of isotropy or anisotropy, so we must determine what type of behaviour can be framed the data of the surveyed surface by checking the spatial structure. In this work, the calculation and representation on the experimental semivariogram graph uses the classical method (Cressie 1993) and (Diggle & Ribeiro Jr. 2007). In *Figure 9.5*, we present the graph of the experimental semivariogram, calculated from the samples collected.

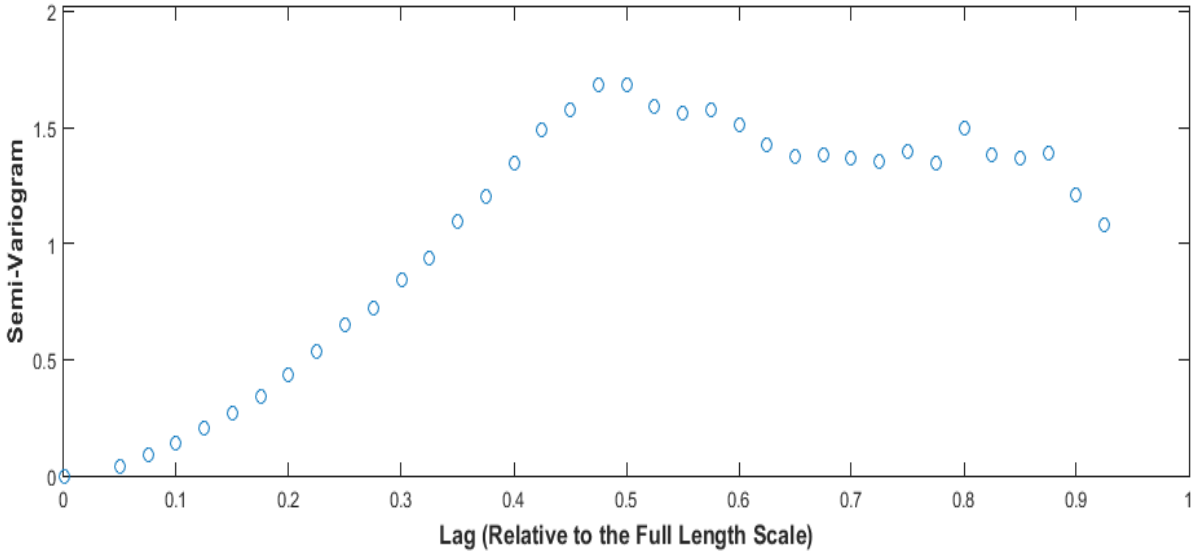


Figure 9.5: *Semivariogram cloud*

The final analysis of the data set expressed in the representation of the experimental semivariogram relates the correlation between the values of the samples and the directional effect of the spatial structure of interest. In *Figure 9.5*, the visual analysis can be performed taking into account the three parameters of analysis presented in chapter 5, that is, nugget, range and sill. We observed that the value represented by the nugget is extremely low, this means that the sampling process was performed correctly, in relation to the number of data collected, spacing and regularity in the sampling. The maximum correlation distance between the sample values reaches almost 50% of the sample space, that is, we obtain correlation in samples ranging up to 200 meters. However, we should note that this correlation may not be confirmed in all directions, as we will see more clearly in *Figure 9.6*.

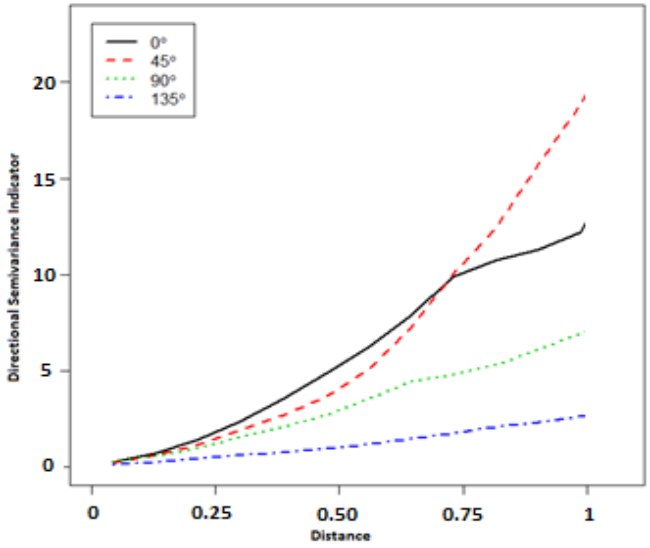


Figure 9.6: *Directional variogram*

Still with respect to the analysis of *Figure 9.5*, the graph does not present a constant sill value until the end of the graph, which represents the calculation of the correlation between points at a great distance. The lack of a constant plateau may represent predictability instability for some regions. The directional analysis of the spatial correlation can be performed using the anisotropic semivariogram, where we can observe the correlation for each direction chosen in a segmented way. A set of four direction indicators is usually used for the calculation of the directional semivariogram (0 °, 45 °, 90 ° and 135 °).

The result of the directional semivariogram analysis is shown in *Figure 9.6*. We observed that there are three directions in which the values of the samples have a higher correlation between them (0°, 90° and 135°), although they have differences of scale in the behaviour, in addition a certain level of anisotropy is seen in the direction of 90°. The behaviour shown for the direction (45°) presents the greatest dissimilarity, resulting in a level of anisotropy in the structure. At this point, we observe that a simple and unique prediction model should not satisfy the requirements for a good prediction result.

In the presence of an anisotropic behaviour in some directions, the application of the omnidirectional experimental semivariogram is valid to observe more clearly the level of anisotropy. In our analysis, shown in *Figure 9.7*, we observed the strong anisotropy level for the 45 ° direction, which represents a very representative trend in this direction. In the direction of 0°, the trend is lower, but it should still be treated. The best result regarding non-existence of tendencies lies in the direction of 135 °, in this direction, simple predictions can be carried out with good results.

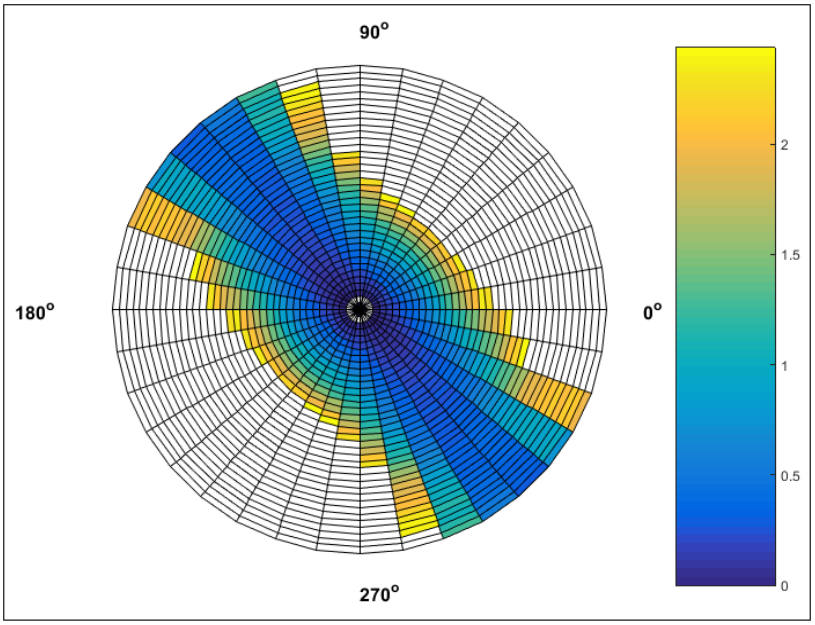


Figure 9.7: Omni-directional semivariogram map

The anisotropic semivariogram map is used to confirm that the direction of the maximum anisotropy corresponds to 135 °, corresponds to the results of the directional semivariogram present in *Figure 9.7*. Confirms the trend assumption depicted in *Figure 9.6*.

9.2.1 Analysis and Results - General Concerns

After performing the initial analyses involving the exploratory data analysis and the construction of the experimental semivariogram, we need to adjust the distribution of points in the experimental semivariogram through one of the theoretical models described in chapter 5, and thus extract some parameters to be used in the predictor of kriging. In particular, we applied a set of five prediction validation processes, of which four are cross-validation tools available in the EasyKrig3.0 Matlab Toolbox (Chu 2004), until the prediction error presented is as minimal as possible. The fifth and last process was implemented from a second set of data acquired with an MBES, performed in the same region and at the same time we collected the data with the SBES.

The difference between the two sets of data is related to the resolution. The reference data have a minimum resolution five times higher than the data used in the estimation, which guarantees an excellent reference source for the validation of the prediction process. This process was called the Inter-Samples (IS) validation, in which each estimated point is compared to a nearest reference data. The final result will be accepted if a minimum reference requirement is reached based on an international standard for bathymetric data (IHO 2008). The process of predicting non-sampled points was performed using a conventional geostatistical approach using the classical Kriging predictor methods: simple, universal and ordinary. We use the algorithms in the geostatistical package EasyKrig3.0 without any adaptation or modification. We evaluated the best predictor, which in our study, based on the analysis from the validation processes was the Ordinary Kriging in "Block to Block" mode.

The result of applying techniques to remove the linear trend found in the original data for certain directions was not satisfactory to achieve the desired results according to (IHO 2008). So, the problem was faced with a focus on the segregation in closed regions of samples that present spatial correlation. The set of regions was evaluated by the application of a partitioning technique called "K-mean clustering". We then apply geostatistical analysis to each cluster using a set of different kriging parameters. However, the same OK kriging model was applied to all of them.

In the search for the best solution for the application of gestational techniques in prediction applications without the use of an auxiliary variable, the work proposal consisted of dividing the study into two paradigms. In the first case, the predictions were made considering the entire set of data. In the second case, we considered the segmentation of the data, classified the same in groups of samples with higher correlation in the same prediction group. Thus, we predicted by groups, considering the prediction for the total area as the junction of the predictions of each group. We compared both prediction paradigms to prove the efficiency of each method to mark our conclusion.

9.3 Geostatistical Semivariogram Analysis, Kriging Estimations and Results without Clustering

Here we will describe in more detail the characteristics of our data set, as well as the geostatistical analysis about them. The whole process was oriented according to the studies of reference works (Goovaerts 1999), (Houlding 2000), (Wackernagel 1995), (Matheron 1962) and (Diggle & Ribeiro Jr. 2007). The set of n samples was acquired in the same hydrographic surveying operation, considering the universe of measurements in the context of regionalized variables $Z(x)$, as described in Chapter 5.

9.3.1 Theoretical Semivariogram Analysis and Model Parameters Estimation

All analyses were performed using tools on generic computing support platforms such as R and Matlab. The experimental variogram was extracted with the 424 samples acquired in the bathymetric survey carried out in the *Bay of All Saints (Bahia de Todos os Santos)*, in the state of Bahia, Brazil. The processing extractions and parameters were supported by the EasyKrig 3.0 Matlab Toolbox (Chu 2004) on the Matlab platform. The result is shown in Figure 9.8.

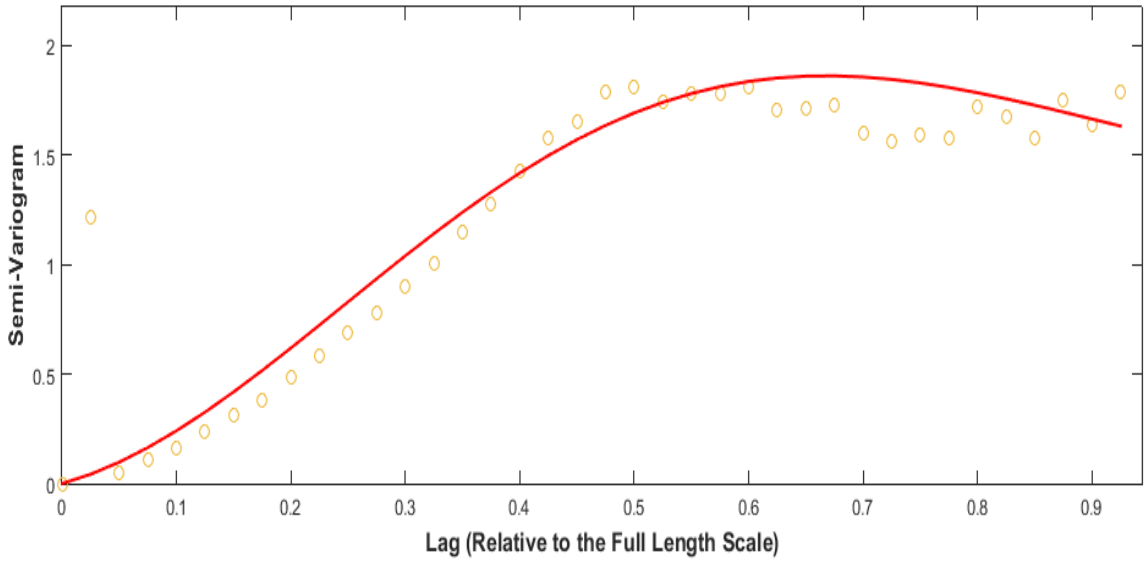


Figure 9.8: *Experimental Semivariogram Fitted using a Theoretical Semivariogram Model for Both Fashion Dataset*

The value of the Nugget, identified by the first semivariogram point, is located on the origin of the axes. This behavior means that the data collection operation was executed correctly, in relation to the number of points and spacing between them. The adjustment was performed with the function of the theoretical semivariogram of the Exponential-Bessel type.

Table 9.2: *Parameters used to adjust the experimental semivariogram*

Survey Dataset Theoretical Model Parameters	
Model	Exponential Bessel
Nugget	3.2947×10^{-13}
Sill	1.5337
Length	1.0659
Power	
Hole Scl	5.4956
Range	0.93586

Considering the initial experimental semivariogram, we need to adjust it using one of the accepted theoretical semivariogram models, so we can extract the value of the parameters to use in the kriging predictor. The best theoretical model and the parameters used to adjust the experimental semivariogram are described in *Table 9.2*. After completing the process of extracting parameters from the model, we started the prediction process using Kriging.

9.3.2 Kriging Prediction

The best Kriging predictor model was determined following a test sequence between the three basic estimation models, simple (SK), ordinary (OK) and universal Kriging with trend (KT). The predictor based on the OK offered the best performance, mainly for two reasons: (i) the bathymetric values acquired in the surveyed surface do not have a constant average throughout its length, varying the trend behavior depending on the direction, in this case SK will fail. In addition, the use of KT was affected by variation in the type of trend, which was linked to the direction of the study. (ii) The trend presented in the distribution of data values is extremely smooth in a specific direction, around 135 °, the trend in the 90 ° direction has an intermediate value and in the orthogonal direction at 135 °, that is, 45 °, the trend is extremely pronounced.

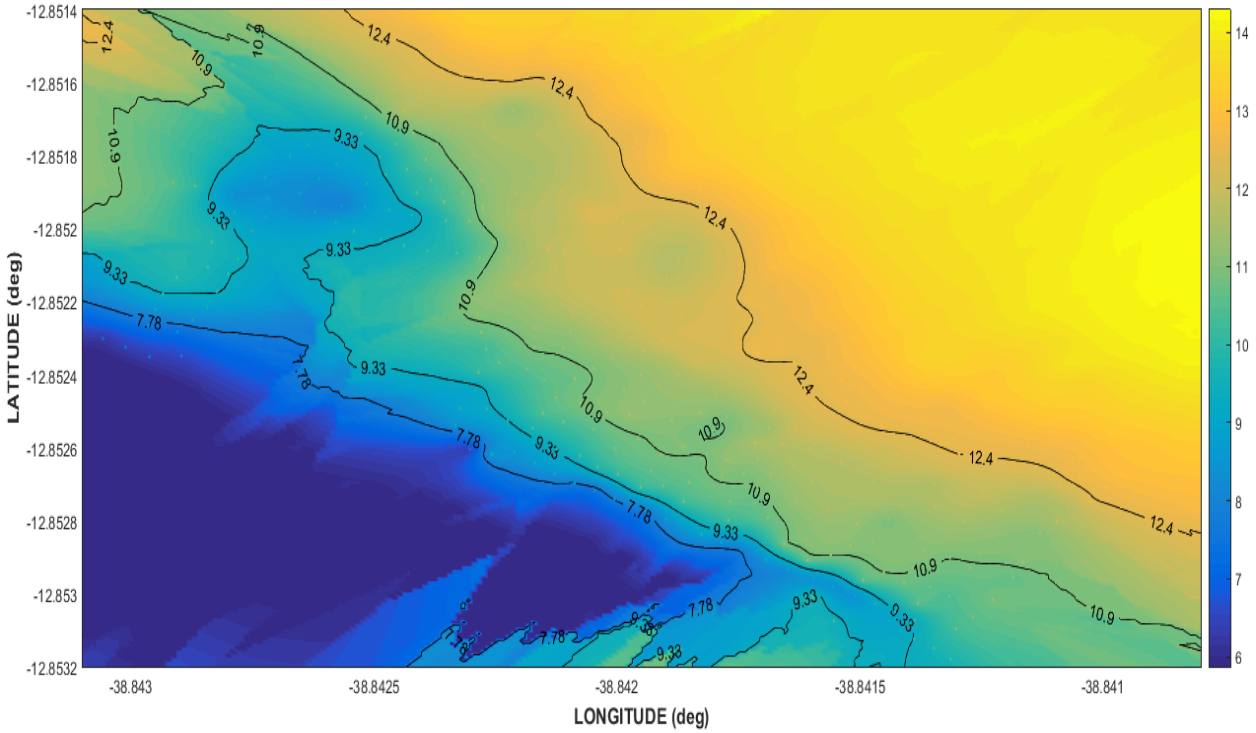


Figure 9.9: *Bathymetric map representation after Kriging Prediction process*

The comparative analysis of the prediction was performed using the variance values, comparing estimated values with the actual data sampled. Thus, the variance map shown in *Figure 9.10* was used. *Figure 9.9* shows the results of the map predicted by OK. In the semivariogram shown in *Figure 9.8*, it is possible to observe a behaviour in the region of the sill, represents a behaviour of cyclicity and in principle, could provoke an instability in the prediction, known as prediction hole effect (Wackernagel 1995), (Matheron 1962) and (Diggle & Ribeiro Jr. 2007).

The hole effect results in the inability to automatically extract the parameter from the theoretical model related to the Sill. However, the parameters were found semi-automatically, automatic extraction was used to obtain Sill reference values, and then manual adjustment was performed between values ranging from -10% to 10% of the reference. The cyclicity presented in the graphical representation of the experimental semivariogram may be a result of the anisotropy in the spatial structure as observed in the directional variogram of *Figure 9.6* and in the anisotropic variogram of *Figure 9.7*. The problem related to cyclicity may affect kriging estimates.

In *Figure 9.10* we present the map of variance relative to the prediction OK, this map shows a small value of variance in the predicted values within the surveyed region. In this figure, the region surveyed is represented by small white dots, one point for each sample acquired with the SBES sensor. The small error level, that is, low variance in prediction, is found within the surveyed area, the map in *Figure 9.10* is represented by the dark blue colour, which is dominant in the region.

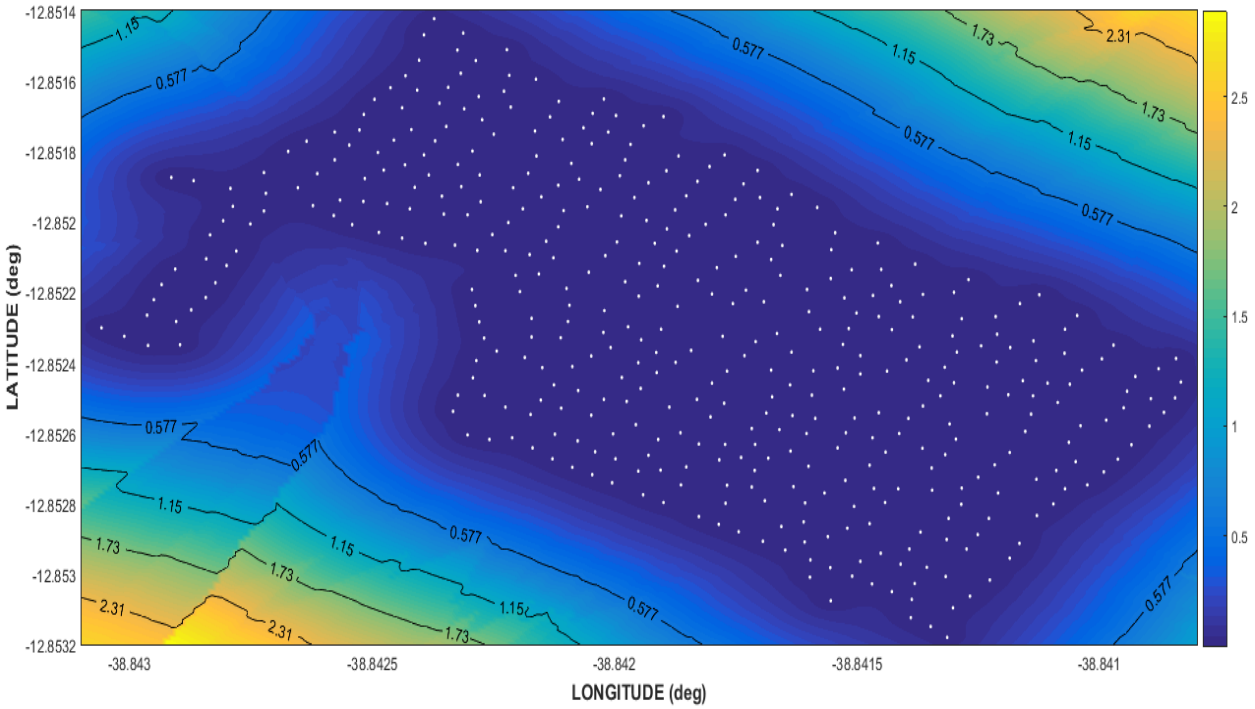


Figure 9.10: *Variance map from Kriging prediction model used in the first phase of this work*

In order to obtain the most accurate variance map, we change the scale of representation of the map of variance, we represent the result in *Figure 9.11*, which shows the small-scale variations within the sampled area. In *Figure 9.11*, we note that the error does not reach 0.10 meters. Based on the results expressed by *Figure 9.1*, the Kriging estimates lead us to believe that the method is appropriate and sufficient for the estimation of bathymetric data.

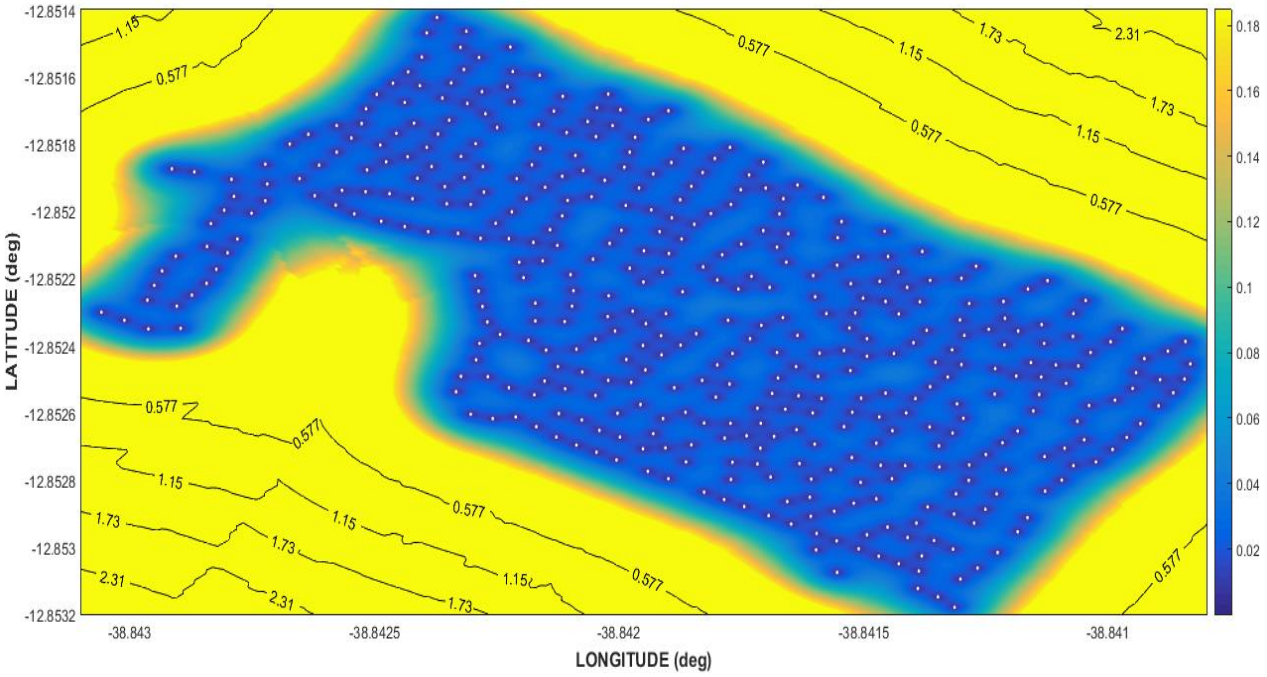


Figure 9.11: *Small scales variance map*

Despite the good results obtained for the processes related to interpolation, this model is certainly not capable of extrapolating. In this case, for the outer area or limit the area sampled shows the much greater variance error, reaching values over two meters, as we can see in *Figures 9.10* and *9.11*.

9.3.3 Cross-Validation Process

The cross-validation process is a simple numerical technique capable of estimating the efficiency and robustness of the prediction model based on a comparative study of observed (sample) and predicted data. Some cross-validation methods are based on statistical tests intrinsically dependent on residues. In this work, we apply two methods based on the calculation of residues, both provided by EasyKrig 3.0 (Chu 2004), they are identified as methods Q_1 and Q_2, both described in the work (Kitanidis 1997). Both methods are used to validate predictions of kriging. Both methods take the values sampled and the values predicted for the same point, leaving these values to estimate the deviation.

The validation process based on methods Q_1 and Q_2 begins with the construction of the normalized residual matrix (ε_k) (Kitanidis 1997). The value of the indicator Q_1 represents the mean of the residuals (ε_k) being calculated using equation 9.7:

$$Q_1 = \frac{1}{n-1}\sum_{k=2}^{n} \varepsilon_k \qquad \textit{eq. 9.2}$$

We note that Q_1 will be calculated on a normally distributed set of values with mean 0 and variance $\frac{1}{n-1}$ (Samui & Sitharam 2011). The probability density function (pdf) of Q_1 will be expressed by equation 9.8:

$$f(Q_1) = \frac{1}{\sqrt{\frac{2\pi}{n-1}}}\, exp\left(-\frac{{Q_1}^2}{\frac{2}{n-1}}\right) \qquad \textit{eq. 9.3}$$

Where n represents the number of data points. If the value of the indicator Q_1 is superimposed on the graph of the Gaussian distribution of the residues under analysis, it should remain within the range considered acceptable, ideally on or near zero of the distribution. This will indicate that the test demonstrates the validity of the model used in the predictions (Samui & Sitharam 2011) and (Kitanidis 1997). In the case of the indicator Q_2 is related to the variance of the matrix values (ε_k), the mathematical expression for its calculation is presented in equation 9.9.

$$Q_2 = \frac{1}{n-1}\sum_{k=2}^{n} {\varepsilon_k}^2 \qquad \textit{eq. 9.4}$$

The calculation of the indicator Q_2 follows the chi-square distribution with $(n-1)$ degrees of freedom. Where n is the number of data points (Samui & Sitharam 2011). The mean and variance of Q_2 are respectively 1 and $\frac{2}{n-1}$ (Samui & Sitharam 2011). The expression that models the computation of the probability density function (pdf) is given by equation 9.10.

$$f(Q_2) = \frac{(n-1)^{\frac{(n-1)}{2}}\, {Q_2}^{\frac{n-3}{2}} exp\left(-\frac{(n-1)Q_2}{2}\right)}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)} \qquad \textit{eq. 9.5}$$

Where Γ represents the gamma function, which is an extension of the factorial function for a set of real and complex numbers, since the argument is subtracted from 1, the pdf value of the Q_2 indicator should be close to 1 to indicate the validity of the model prediction.

The results of cross-validation for the predictions made with our bathymetric data can be seen in Figure 9.12. In this figure, we confirm that the best result for the prediction performance is provided by the Q_1 indicator. In the case of the Q_2 indicator, it does not show a good validity for the predictions made. Referring to the reference (Chu 2004), we confirm that it is difficult to obtain the indication of validation of the model by both indicators. The numerical results of the cross-validation for methods Q_1 and Q_2 are provided in *Table 9.3*.

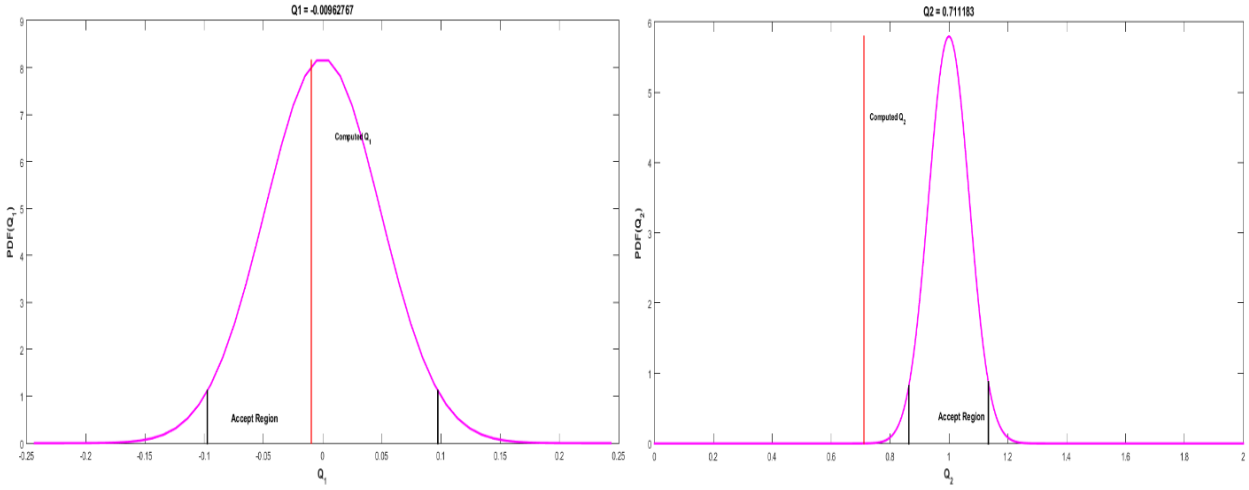


Figure 9.12: *Cross-validation* $\boldsymbol{Q_1}$ *and* $\boldsymbol{Q_2}$ *results.*

Table 9.3: *Cross-Validation Numerical Results for* $\boldsymbol{Q_1}$ *and* $\boldsymbol{Q_2}$ *methods*

Cross-Validation Result for Q_1 and Q_2 Method			
Q_1		Q_2	
Value	Deviance Error	Value	Deviance Error
-0.00962767	0	0.711183	0.288817

After analysing the results for the indicators Q_1 and Q_2, it was decided to use two more methods of cross-validation provided by EasyKrig 3.0, the Double-Kriging (DK) and Leave One Out (LOO) methods. The simplest application of cross-validation by DK cross-validation is determined by a two-step process described in (Chu 2004) and (Piazza et al. 2009). In the first step, the sampled original values are reserved as input for a comparison process. In the second step, a Kriging process is performed considering the average value of the actual samples and the parameters extracted from the theoretical model of the semivariogram extracted automatically. Then the result of the second

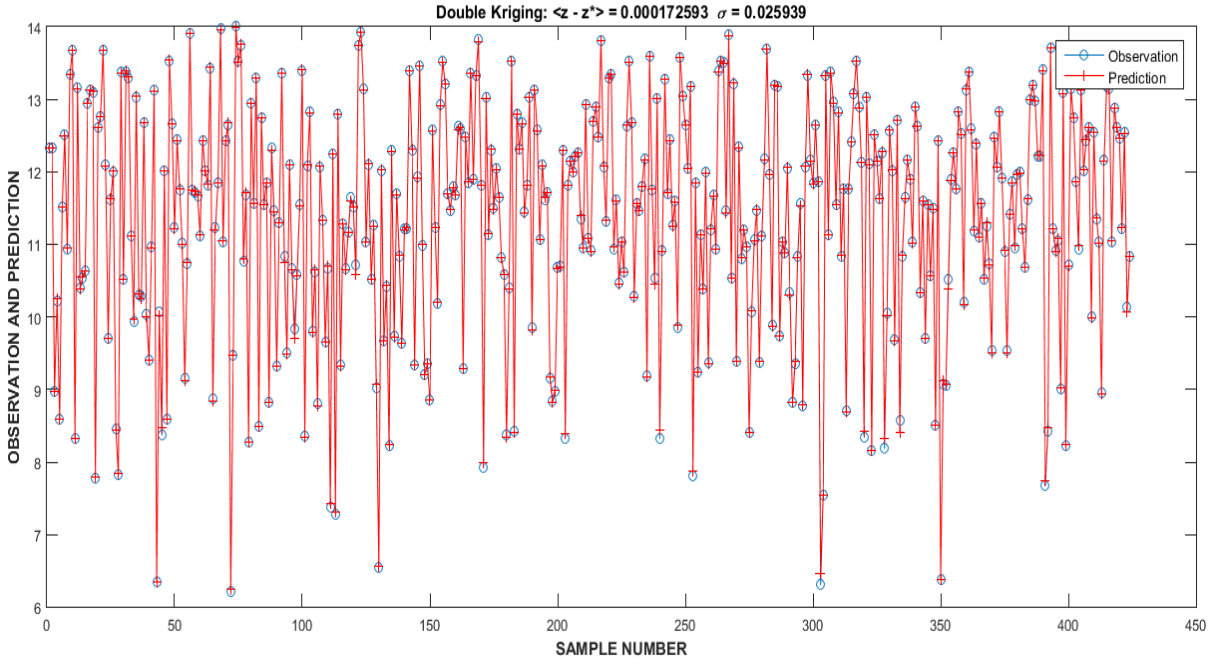


Figure 9.13: *Double-kriging cross validation result*

kriging phase is compared with the original sampling data. Presenting as output a graph where the predicted and sampled values are represented simultaneously.

In *Figure 9.13* illustrate the results of the cross-validation process using DK, the numerical results reveal a good prediction level with the difference between observed and predicted values in the order of 0.000172 meters, while the variance was 0.026. This numerical data is presented in *Table 9.4*.

***Table 9.4**: Numerical results from Double-Kriging Cross-Validation process*

Double-kriging Cross-Validation Numerical Results	
Observed-Predicted Difference	**Variance**
0.000172593	0.025939

The LOO method works similarly to the DK, it considers the sampled data z_n as one of the input data of the comparison step. The process consists in the consecutive withdrawal of sampled points, and subsequent estimation of this point using the neighbouring points that were estimated by the Kriging prediction process. In this way, the mean error is calculated and used to evaluate the model. In terms of accuracy, LOO can often result in a large variation of results for validation of a predictor (Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel et al. 2011).

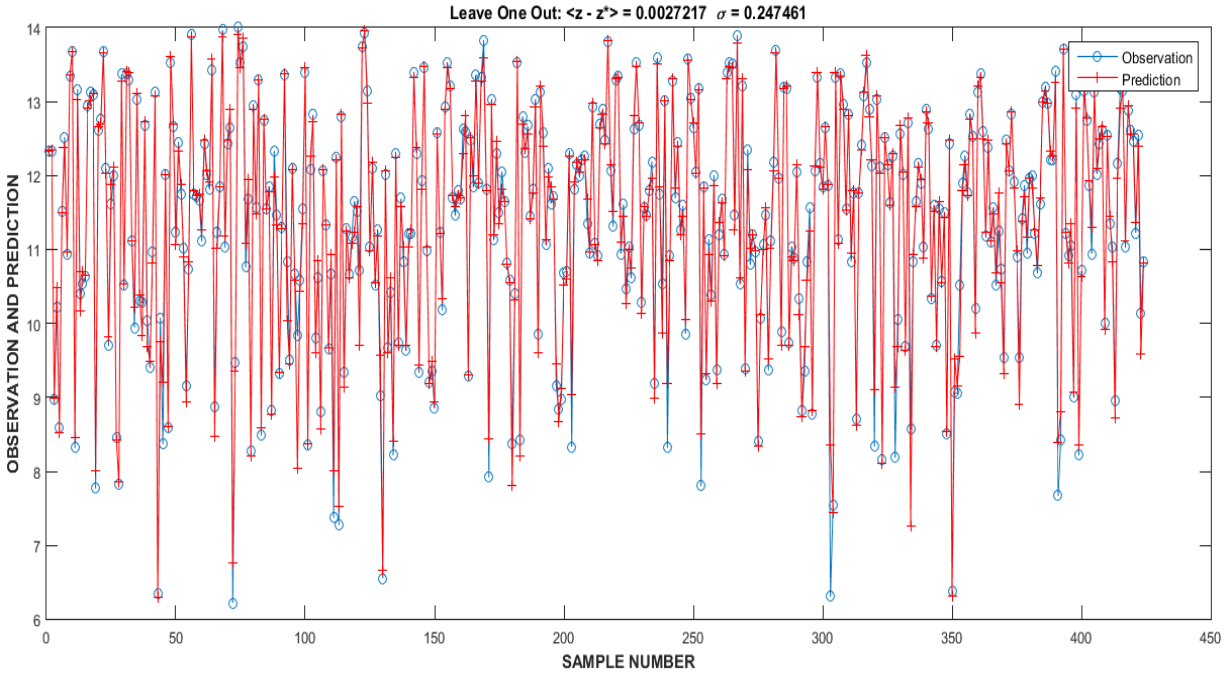


***Figure 9.14:** Leave one out cross validation result*

In the cross-validation by the LOO method, shown graphically in *Figure 9.14* and numerically in *Table 9.5*, the interpolation process performed by the Kriging technique was shown to have a good estimation quality. This conclusion is based on the differences between observed and predicted values, but in this case the result was not classified as better than the one indicated by the DK cross validation technique. Looking at *Figures 9.13* and *9.14* we have identified that both DK and LOO methods have some points that were underestimated.

***Table 9.5:** Numerical results from Leave-One-Out*

Leave-One-Out Cross-Validation Numerical Results	
Observed-Predicted Difference	**Variance**
0.0027217	0.247461

The results obtained by the study of the various methods of cross-validation demonstrate ambiguities between some of these methods, thus, we developed the proposal of applying a second set of data obtained with another type of sensor, an MBES. The set of data collected with MBES was collected in the same SBES data collection used in the study of this thesis. Both sample points were collected independently and with different variances. In addition, data collection using MBES results in a data set with sample densities greater than one hundred times the number of points collected with the SBES within the region of interest.

The region of interest was defined as the region of perfect overlap between SBES and MBES data. The proposed method that confronts the predicted values with the actual values acquired with the MBES will be called here as the Inter-Sensor (IS) method, being considered here as our main method of validation. The set of reference data having a much higher density of sample points allows more precisely quantifying the prediction error, or deviations obtained using the method of interpolation by the Kriging method. *Figure 9.15* shows the two sets of SBES and MBES data and the region in which the collected data match, that is, we have measurements of the same area performed by both sensors

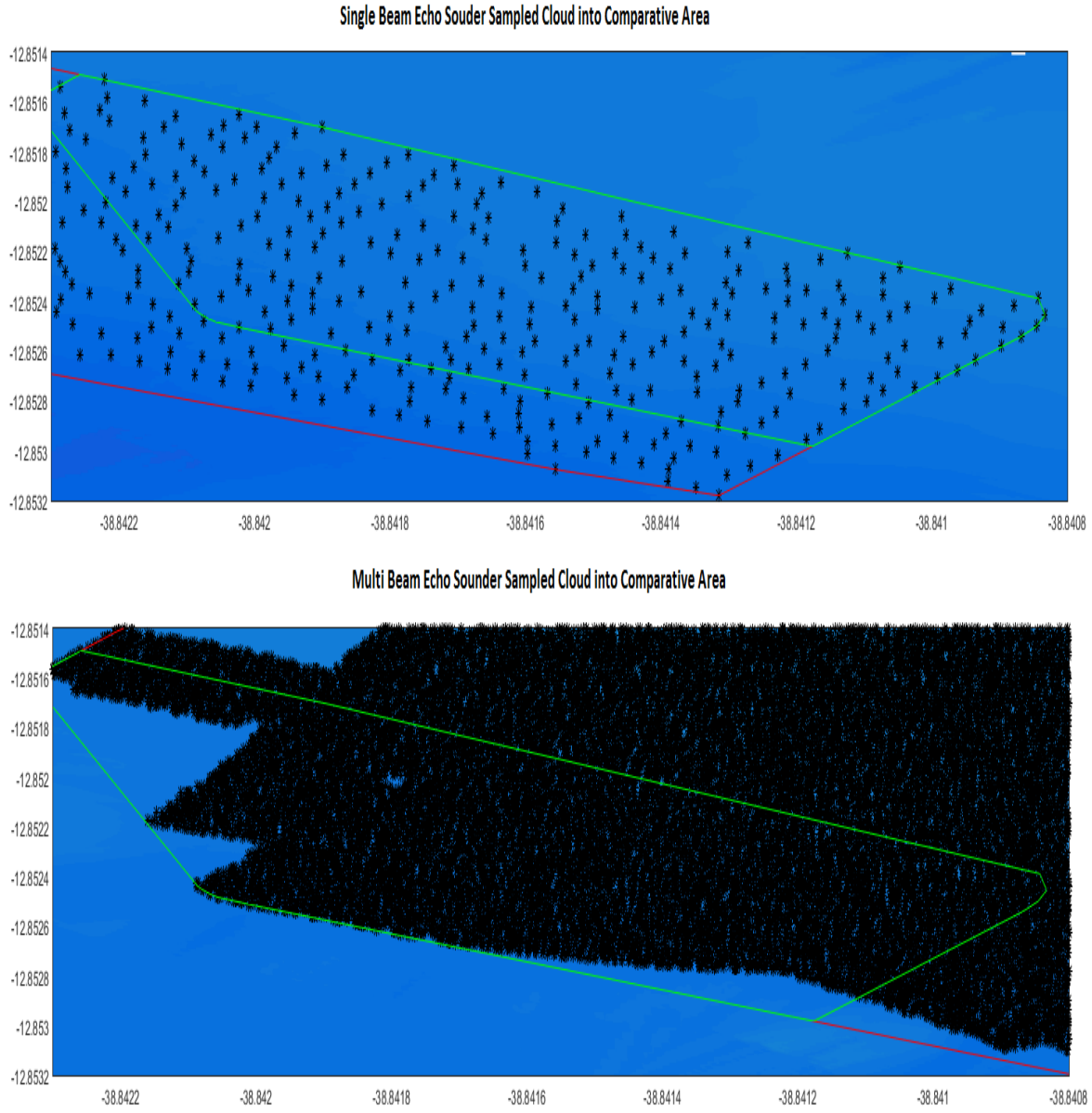


Figure 9.15: *Interest region to use in IS validation method from, worked dataset (SBES) and reference dataset (MBES). The regions bounded by the yellow line represent the data that coincides in the same geographic region*

In the upper illustration, you will find the representation of the cloud of sample points of the SBES and in the illustration below the data referring to the MBES. To obtain a more consistent view of all the sample space in which the bathymetric survey was performed, we present in *Figure 9.16* an overlap of the data collection region for both sensors

After the kriging interpolation process, the original data set, composed of 424 samples and represented in *Figure 9.16* as black circles, was incremented to 50,000 estimated points. The final result of the estimated points is a regular network, with estimated points equally spaced in the directions of the reference axes. This spacing was determined by one (1) meter, determined based on the average spacing of our reference data. The network of predicted points is represented in *Figure 9.16* as red squares, which due to the density of points we observe as a continuous surface. The reference data set, i.e. the MBES, is represented in this same figure by a total of 51,131 real points, represented by the blue area which are blue squares on a small scale. Further details on the number of points for each sensor and for the predicted data are given in *Table 9.6*.

After the kriging interpolation process, the original data set, composed of 424 samples and represented in *Figure 9.16* as black circles, was incremented to 50,000 estimated points. The final result of the estimated points is a regular network, with estimated points equally spaced in the directions of the reference axes. This spacing was determined by one (1) meter, determined based on the average spacing of our reference data. The network of predicted points is represented in *Figure 9.16* as red squares, which due to the density of points we observe as a continuous surface.

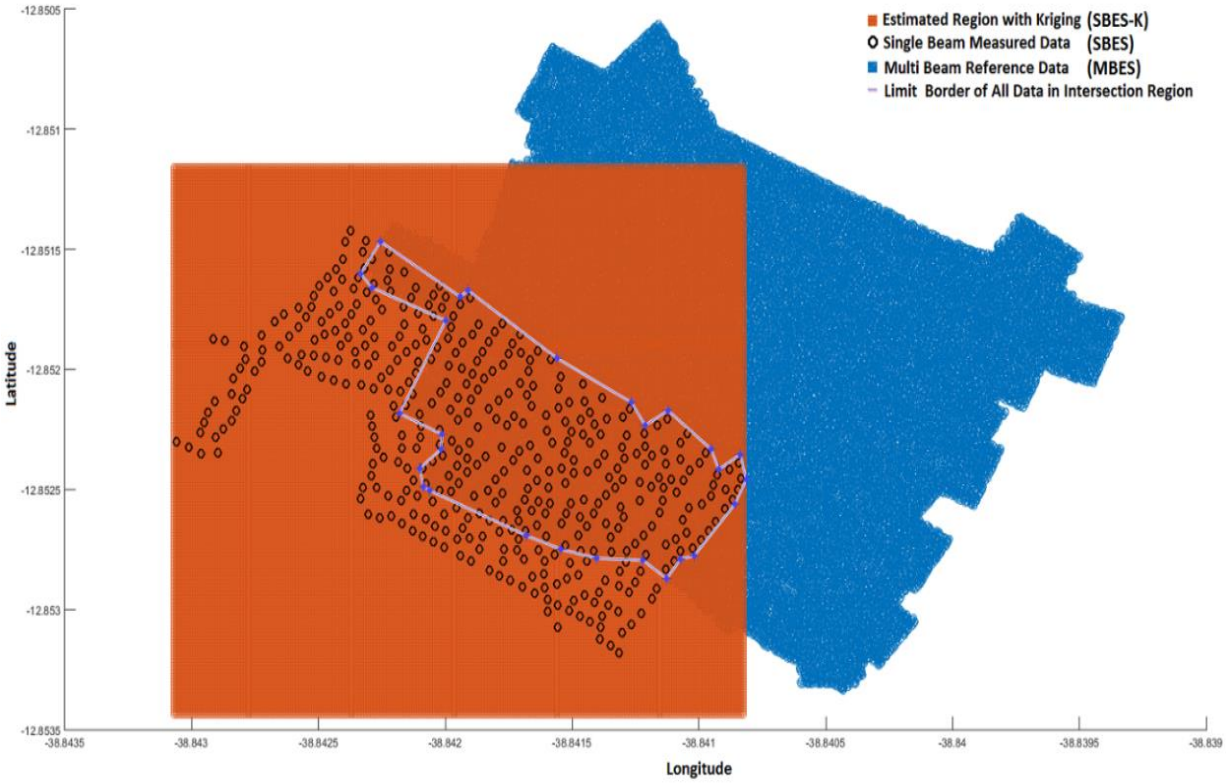


Figure 9.16: *Graphical representation of dataset crossing*

The reference data set, i.e. the MBES, is represented in this same figure by a total of 51,131 real points, represented by the blue area which are blue squares on a small scale. Further details on the number of points for each sensor and for the predicted data are given in *Table 9.6*.

Table 9.6: *Number of points and aspect for each dataset.*

Data Type	Number of Total Data Points	Number of Data Points into the Intersect Region	Distribution
SBES Sonar	424	222	Irregular
SBES Kriging	50000	19430	Regular
MBES Sonar	51131	23792	Partially Irregular

The IS method will consider the three SBES datasets, estimated based on kriging (SBES-K) and obtained by the MBES survey, to perform the comparative analysis within the intersection area, bounded by the blue line in Figure 9.16. The comparative analysis will be implemented following the following algorithm:

- 1° Consider a single point (P_{SBES-K}) within the SBES-K data set;
- 2º. Search the sampled point within the MBES data set closest to P_{SBES-K};
- 3° Use equation 9.11 to evaluate the difference between the sampled and predicted values represented by the chosen points.

The comparison is performed by performing the operation represented by equation 9.11, in which the residual value is extracted for the chosen points within the two previously defined data sets SBES_K and MBES.

$$\delta_{IS} = z_p - z_r \qquad \textit{eq. 9.6}$$

Where z_p represents the estimated value, while z_r represents the reference value of the nearest neighbor of the estimated point. The value of the residual, represented by δ_{IS}, will be stored in a stock accumulation matrix to be graphically plotted as a histogram. *Figure 9.17* shows a small intercession region of the two SBES-K and MBES data sets, showing that there are a large number of reference points, represented by blue circles with "x" neighbouring an estimated point, represented by a red square.

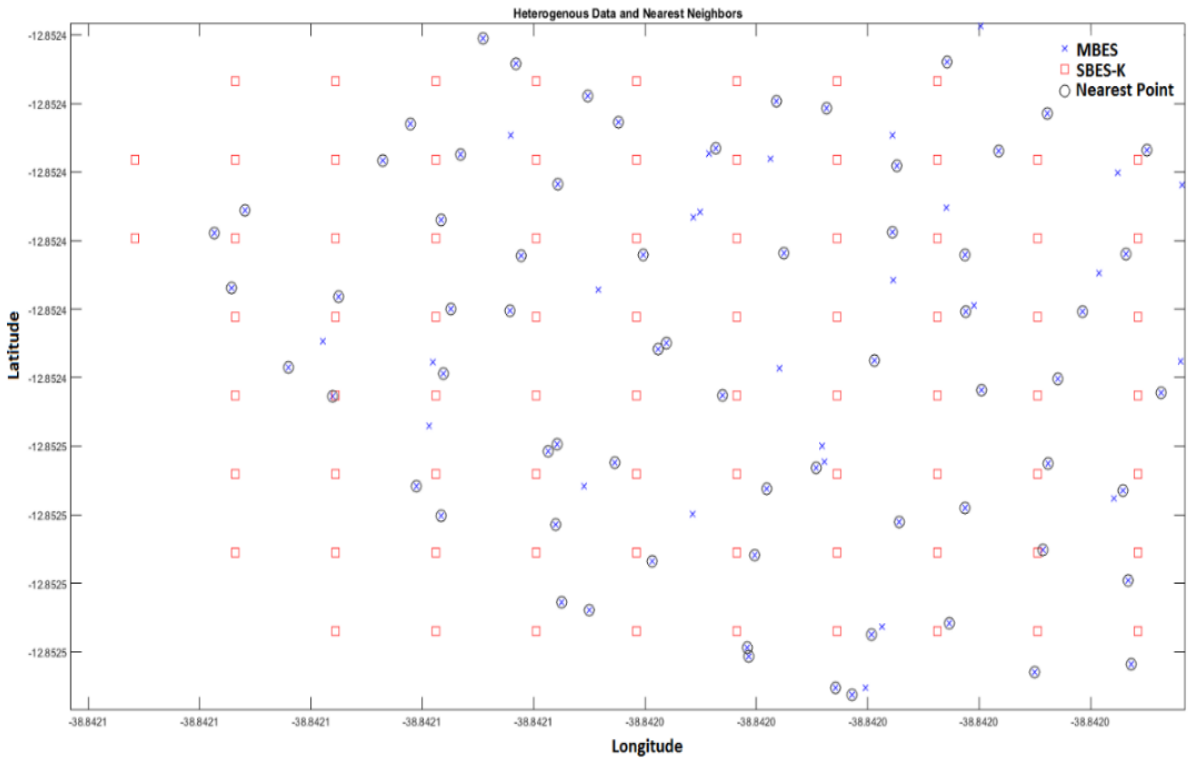


Figure 9.17: *Distribution of predicted and reference points with the nearest neighbour selected*

Looking at *Figure 9.17*, we conclude that we cannot guarantee that we will always have a reference point correlated to an estimated point over the same geographic coordinate to compare its values. However, there are many points in a nearby neighbourhood that ranges from 2 to just over 0.8 meters which can ensure the formation of a pair, estimated point - reference point. To determine the nearest neighbour in relation to each estimated point, a simple non-parametric test type was performed, based on the *k-nearest* method. For the calculation of the metric distance used in this work, we consider the Euclidean distance algorithm presented in equation 9.12.

$$d_{st}^2 = (x_s - y_t)(x_s - y_t)' \qquad \textit{eq. 9.7}$$

The nearest neighbour computation applying the k-nearest classification algorithm was performed within the intersection region shown in *Figure 9.18* for both sets of data using the Matlab program.

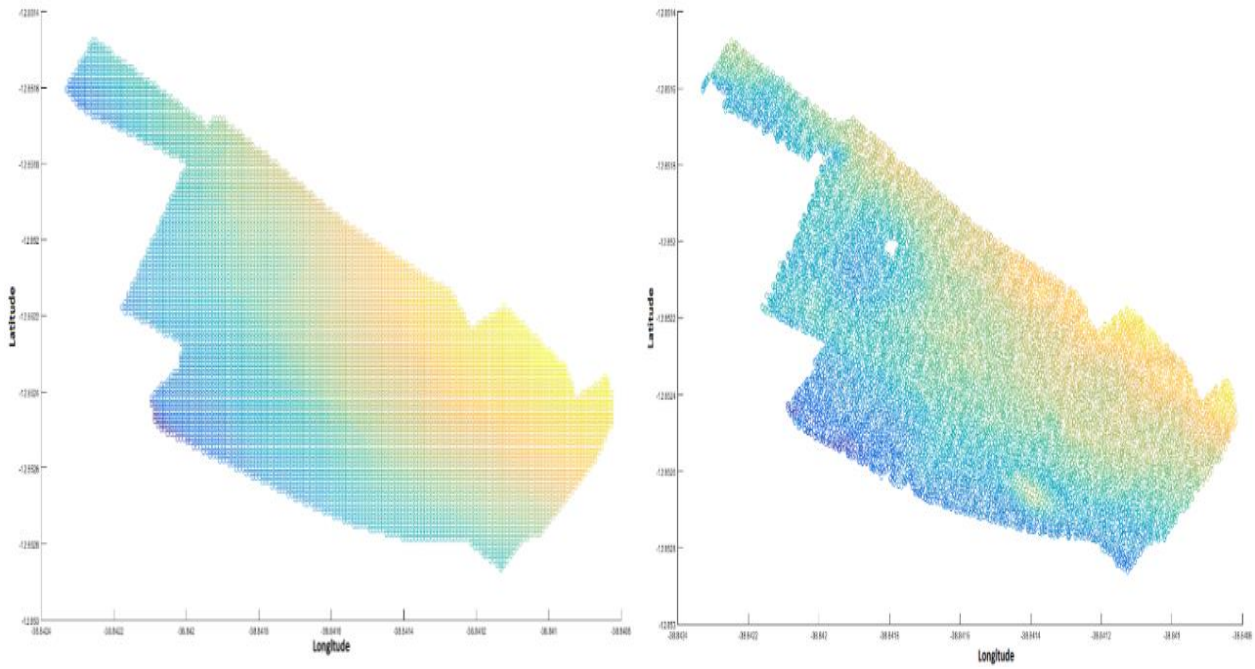


Figure 9.18: *Intersect regions extracted from predicted dataset (left) and reference dataset (right)*

As a result of the IS validation process, a map was generated that reflects the residue value for each estimated point by the kriging process. In *Figure 9.19*, we can see the residual map considering the nearest reference points with distance ≤ 1.0 meter.

In *Figure 9.19* the differences between the predicted and reference values are plotted in a colour code, with each colour representing a range of error, it would be a form of classification in clusters of errors. The smallest prediction errors are represented as red, numerically representing an error value between 0.30 and -0.30 meters. If we compare these values in term of percentage error, this represents an approximate value of 2.42% in relation to the maximum value recorded in the dataset and 4.82% in relation to the minimum value. According to the International Standard for Hydrographic Surveys (IHO 2012), this error value is acceptable for classifying bathymetric surveys and estimates within the IHO-S44 *Special Order* (IHO 2008). On the other hand, the region of green dots represents a small prediction error that is between 0,31 and 0,50 meters. In this case, the estimated points can be accepted as being valid for classification under the IHO-S44 *Order 1A* and *Order 1B* standard, depending on the total depth in the second sound area (IHO 2008). For the other areas, the predictive value did not reach the minimum necessary for classification within the standards (IHO 2008), these regions are represented by the light blue and dark blue points.

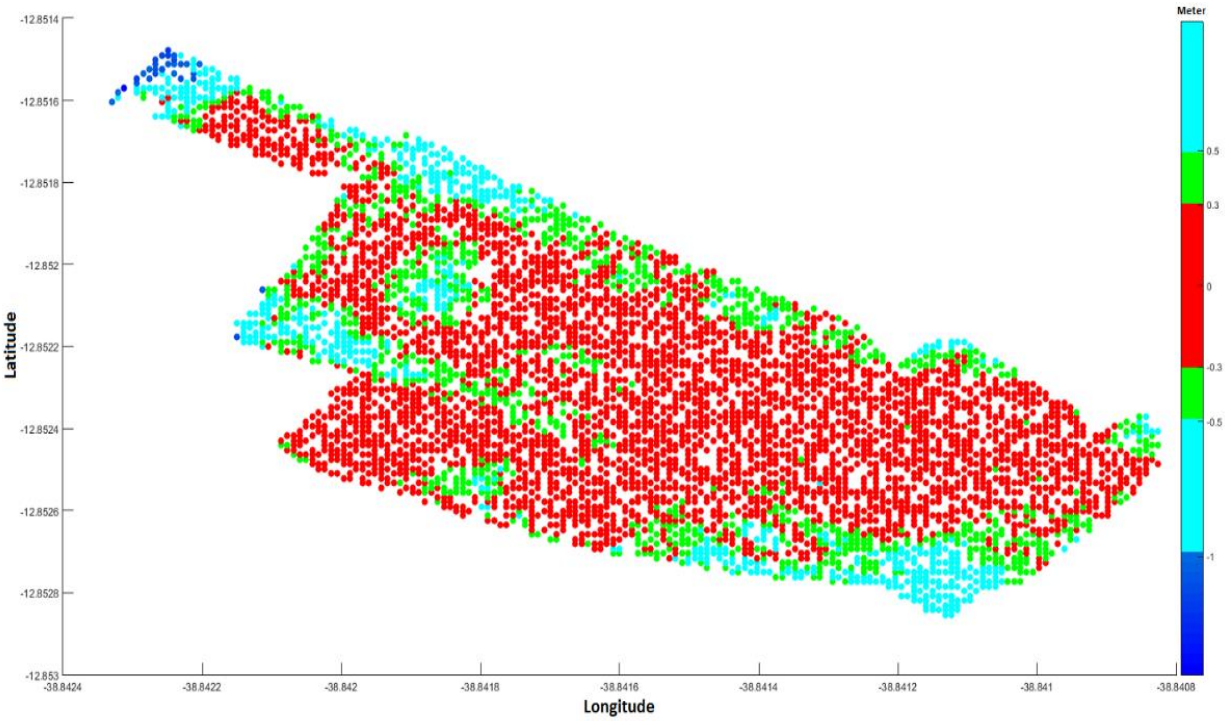


Figure 9.19: *Residual map generated after the IS validation process*

Observing the result of the errors associated with the estimated values presented in *Figure 9.19*, some regions experienced a high level of error, in this case, we have investigated the three overlapping data sets (SBES, SBES-K and MBES). We observed that there is a region of intercession involving not only the two previously studied (SBES-K and MBES). *Figure 9.20* shows this region of intersection, the largest errors were found outside the zone of intersection of the three data sets, that is, we were considering not only the interpolated data, but also the extrapolated data. As we noted in section 9.2 and *Figures 9.10* and *9.11*, the prediction using Kringing is not valid for extrapolations, the results obtained by the residue analysis strengthened the previous statement in section 9.2.

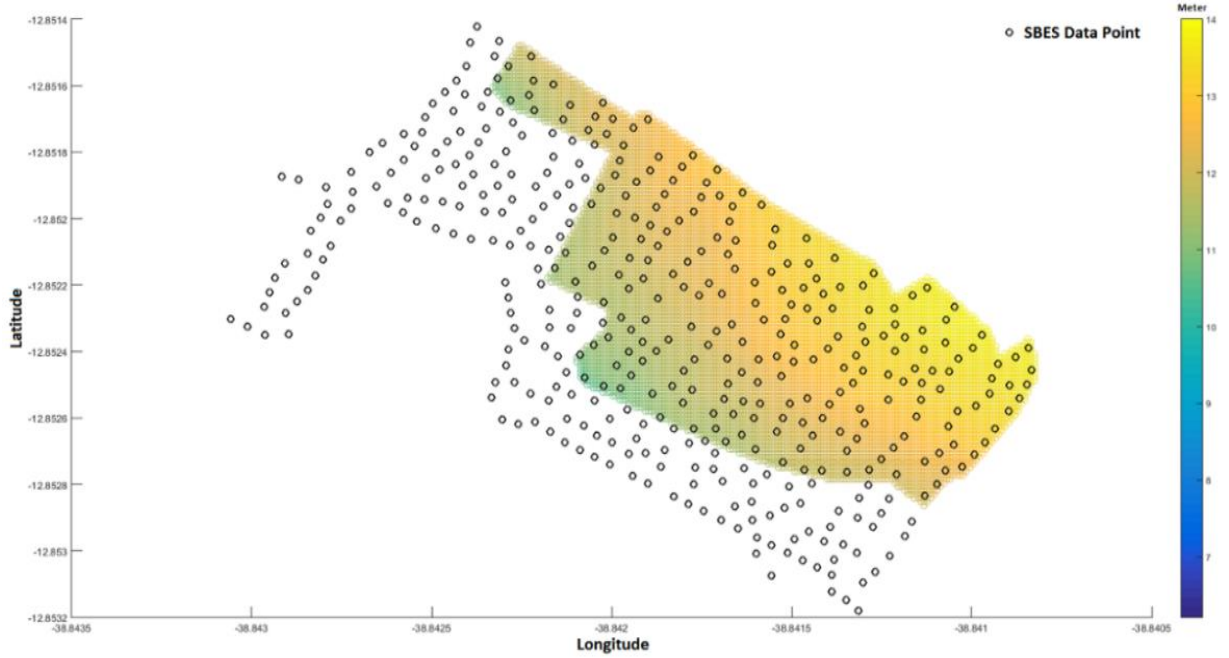


Figure 9.20*: Intercept region between real point dataset and predicted points*

In the intersection region involving the three data sets we have 5,117 points, of which 4,517 represent valid estimated points, representing a total of 87.82% of the total predicted points and with classification capacity within the standard IHO-S44 *Special, 1A* and *1B Order* (IHO 2008), but only 3,234 points can be classified as *Special Order*, with an estimated prediction error below 0.30 meters, representing 63.20% of the total estimated points.

Table 9.7: *Kriging results data summary*

Kriging Results				
Description	**Std. Dev**	**Bins**	**%**	**IHO-S44 Order Validity**
Number of Points into Region	-	5,117	100%	
Number of Points with Difference Value (DV) <=0.3	1 σ	3,234	63,20%	Special Order
Number of Points with Difference Value (DV) <=0.5	1.61 σ	4,517	87,82%	Special Order and Order 1A and 1B
0.50<Number of Points with Difference Value (DV) <=0.60	2 σ	4,917	8.27%	Acceptable only to depth > 100 m
Number of Points with Difference Value (DV) >0.6	-	200	3,91%	Not Acceptable in IHO Standard

In *Table 9.7* we present the summary of the IS validation process with the percentage of valid points for classification according to IHO-S44 *Special, 1A* and *1B Order* (IHO 2008). In this table, the number of invalid points for classification is also displayed.

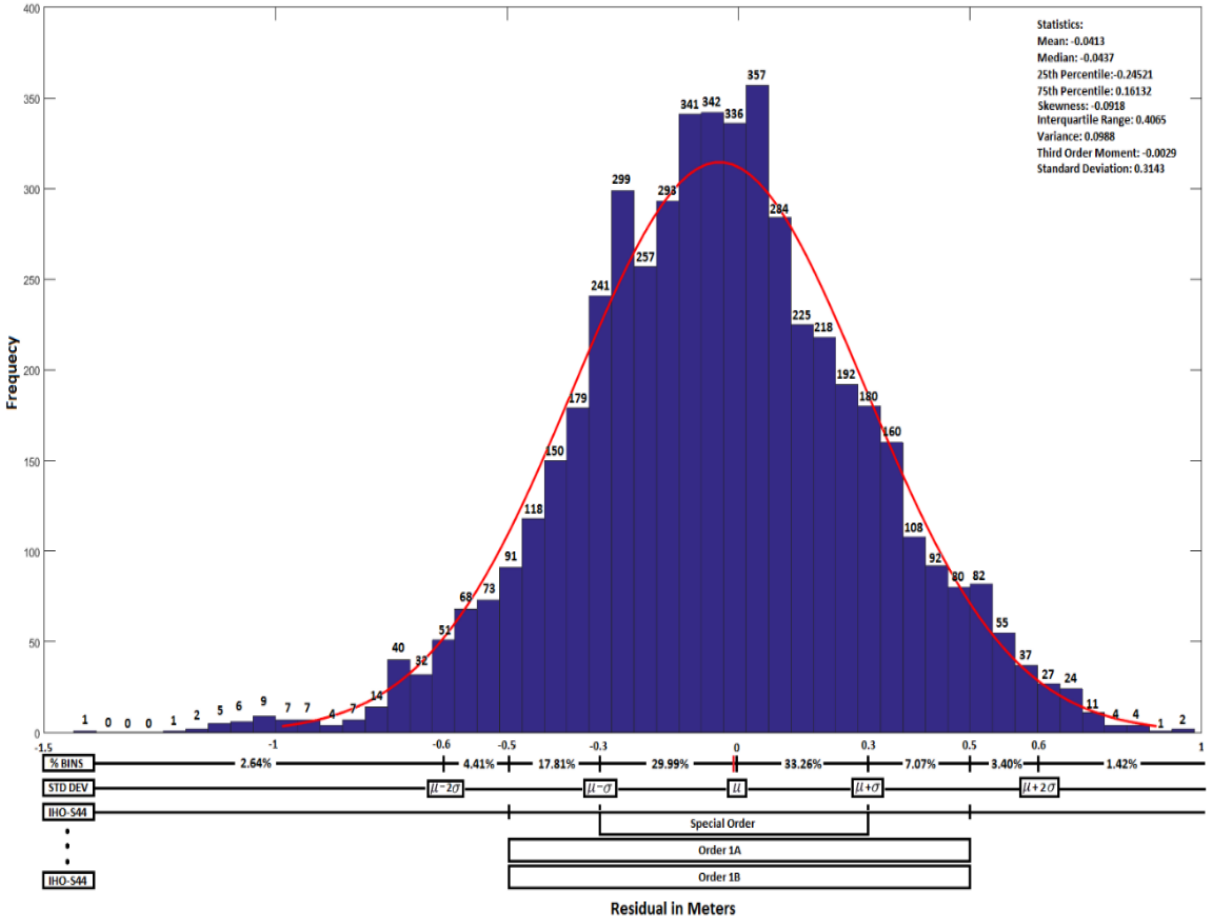


Figure 9.21: *Estimated error distribution against Gaussian distribution*

In this case, we consider as invalid points all points at which the residual error value is above 0.5 meters. Considering this case, we will have a total of 607 points, which represents 11.87% of the total estimated points. Considering the need to analyse the content of the residues more graphically, in *Figure 9.21*, we represent these values in the form of a histogram of the residual Gaussian distribution. This representation allows an analysis of the mean magnitudes of the residues alongside some descriptive statistics and indicators may help to better understand the quality of the estimates generated by kriging.

9.3.4 Average Magnitude of Errors Analysis

The next stage of study and literature review was established in order to study the methods of the mean magnitude of errors evaluation. Thus, we used some techniques to measure and compare the mean error amplitude using different interpolation methods that were applied to the data. In this context, for example, the reference (Robinson & Metternicht 2006) compare the accuracy of some interpolation methods such as kriging, Inverse Distance (IDW) and Splines. In this work, the authors used a set of reference data and another set of estimated data to extract some statistical information about the mean magnitude of errors, evaluating the efficiency of the prediction involving spatially distributed solid properties.

In the analysis of (Robinson & Metternicht 2006) the main techniques used in the evaluation of the mean magnitude of errors are implemented through the Mean Error (ME) and Mean Squared Root Error (RMSE) indicators, which led to similar results involving cross validation and RMSE. In another study, (Curtarelli et al. 2015) carried out a work involving a comparison of interpolation methods for bathymetric data of a water reservoir located in the Amazon Basin. The results were evaluated according to the application of the analysis based on the RMSE and Mean Absolute Error (MAE) indicators, in order to choose the best interpolation method. The results confirmed that the geostatiscal-based interpolation method can map important features obtained from the bathymetric survey. The numerical values found for the RMSE and the MAE were respectively 0.92m and 0.45m.

Table 9.8: *Numerical results for the global average magnitude of the error metrics*

Global Average Magnitude of the Errors Analysis					
RMSE	**RMSE (%)**	**MAE**	**MAE (%)**	**Correlation Coefficient**	**Standard Deviation**
0.320	2.21	0.25	1.72	0.9098	0.3143

Another interesting work was done to estimate the evolution of the bathymetry of a river and discharge calculation (Yoon et al. 2012). In this application, the achieved results were confirmed by the RMSE value of 0.52 m. Thus, the use of statistical indicators corresponding to the mean magnitude of the error constitutes a frequent evaluation of the predictions in some scientific works for the estimation of environmental variables. We then identified three widely used indicators based on RMSE, MAE and MSE. Among these, we chose RMSE as our main indicator and MAE with a secondary indicator. They were used here as indicators of the estimation performance. Our numerical results for measuring the magnitude of the error were provided in *Table 9.8* for the numerical results for our estimate.

Observing our results, reflected in the numerical results of *Table 9.8*, we have the notion of the accuracy level of the predictions performed with the Kriging method in the context of bathymetric surveys without the use of complementary information. The values found for RMSE and MAE were 0.32 m and 0.25 m, respectively.

Table 9.9 shows the results of the study involving the descriptive statistics on the data set, comparing the values and correlations between the observed (SBES), estimated (SBES-K) and reference (MBES) data. This statistical study indicates some differences between extreme values related to the different datasets, resulting in a high value for the discrepancy between observed and estimated values.

Table 9.9: *Descriptive statistics for all bathymetric data*

Descriptive Statistics Measures with All Bathymetric Data									
Descriptors	**Original Data (A)**	**Estimated (B)**	**Reference (C)**	**Discrepancy (A-B)**		**Discrepancy (A-C)**		**Discrepancy (B-C)**	
	meters	meters	meters	meters	%	meters	%	meters	%
Minimum Value	6.22	10.54	10.49	4.27	38.03	4.27	34.13	0.05	0.40
Maximum Value	14.00	13.82	14.46	0.18	1.58	0.46	3.68	0.64	5.10
Average	11.36	12.51	12.55	1.14	10.04	1.19	9.51	0.04	0.32
Median	11.62	12.50	12.54	0.88 (	7.57	0.92	7.36	0.04	0.32
Standard Deviation	1.61	0.68	0.75	0.42	26.09	0.86	126.47	0.07	9.33

The *Table 9.9* shows the results of the study involving the descriptive statistics on the data set, comparing the values and correlations between the observed (SBES), estimated (SBES-K) and reference (MBES) data. This

statistical study indicates some differences between extreme values related to the different datasets, resulting in a high value for the discrepancy between observed and estimated values.

Note that in this case we are not considering data in all regions, coincident or not, so we find high values of discrepancy. In *Table 9.10*, in the used data sets is considered only points within an intersection area, where there must be samples of the three data sets considered. In *Table 9.10*, it is observed that, when the area is limited to the coincident region, the discrepancy is reduced drastically. The result was expected, but the underestimation for some points still exists.

9.3.5 Discussion and Preliminary Conclusions

In this section, we will discuss the results presented in Table 9.7, comparing with the requirements of the international standard for bathymetric surveys (IHO 2008) and finally compare with the conclusions published by Cressie 1993 that is our main reference in the study of spatial statistics. In *Table 9.7*, line 3, only 87.82% of the estimated data can be accepted within the standard (IHO 2008) for classification between Special Order, Order 1A and Order 1B. In chapter 1, of the reference (Cressie 1993), the author developed a comparative study on two situations. The first situation considers that the sample set with mean μ and variance σ_0^2/n. In the second case, the data set of the samples has a positive correlation between the values of the samples, in this case, the correlation between the samples decreases as the distance between the samples grows.

In the first case, for both sides of the Gaussian distribution, the confidence interval is defined as 95% in relation to the mean according to expression 9.13.

$$(\bar{Z} - (1.96)\,\sigma_0/\sqrt{n}\,, \bar{Z} + (1.96)\,\sigma_0/\sqrt{n}) \qquad \textit{exp. 9.13}$$

Table 9.10: *Descriptive statistics for data into the survey coincident region*

Descriptive Statistics Measures with Only Bathymetric Data into Intersected Region									
Descriptors	**Original Data Inter (A)**	**Estimated Data (B)**	**Ref. Data (C)**	**Discrepancy (A-B)**		**Discrepancy (A-C)**		**Discrepancy (B-C)**	
	meters	**meters**	**meters**	**meters**	**%**	**meters**	**%**	**meters**	**%**
Minimum Value	10.74	10.54	10.49	0.20	1.60	0.25	2.00	0.05	0.40
Maximum Values	14.00	13.82	14.46	0.18	1.44	0.46	3.68	0.64	5.10
Average	12.48	12.51	12.55	0.03	0.24	0.07	0.56	0.04	0.32
Median	12.48	12.50	12.54	0.02	0.16	0.06	0.48	0.04	0.32
Standard Deviation	0.77	0.68	0.75	0.09	0.69	0.02	2.94	0.07	9.33

In the second case, where there is a positive correlation between the sample values, the covariance is defined by equation 9.13.

$$cov(z_i, z_j) = \sigma_0^2 \cdot \rho^{|i-j|} \qquad \textit{eq. 9.13}$$

Where i, $j = 1, \ldots, n$, and $0 < \rho < 1$. As a result, we will arrive at a first-order autoregressive process (Cressie 1993). In this case, for a spatial data set in $\mathbb{R}^1$, where $z(0)$ or $z(2/3)$ predictions represent similarities for any $z(n+1)$. Then, considering $\bar{Z}$ as an estimator for this process, having mean μ and variance σ_0^2/n, the process variance will be given by:

$$var(\bar{z}) = \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} cov(z(i), z(j)) \right\} =$$

eq. 9.14

$$= \left\{ \frac{\sigma_0^2}{n} \right\} [1 + 2\{\frac{1}{(1-\rho)}\}\{\frac{1}{(1-n)}\} - \frac{2\left\{\frac{\rho}{(1-\rho)}\right\}^2 (1-\rho^{n-1})}{n}]$$

Considering the value $n = 10$, $\rho = 0.26$ and e $var(\bar{Z}) = \{\sigma_0^2/10\}[1.608]$, the 95% confidence interval for both sides of the Gaussian distribution for the mean μ will be defined as:

$$(\bar{Z} - (2.485)\,\sigma_0/\sqrt{10}\,, \bar{Z} + (2.485)\,\sigma_0/\sqrt{10}) \qquad \textit{eq. 9.15}$$

In this case, the hypothesis of assuming the application of non-spatial statistics in the estimates leads to inconsistent results. If we consider the previously studied confidence intervals, in the second case, the confidence interval of the estimates reaches 87.8% instead of the 95% reached in the first case. Comparing these results with the results obtained in our study, expressed in *Table 9.7*, line 3, the result is exactly equal to the result found by (Cressie 1993). Therefore, we have concluded that the use of geostatistics for the estimation of bathymetric data at non-sampled points, without the use of auxiliary information to support the process, will give similar results to the application of simple statistics.

In contrast, we perceive three different regions for the behaviour of the estimated values compared to those observed from the results and using a graphical representation that relates the estimated and observed data in a plane, represented by *Figure 9.22*.

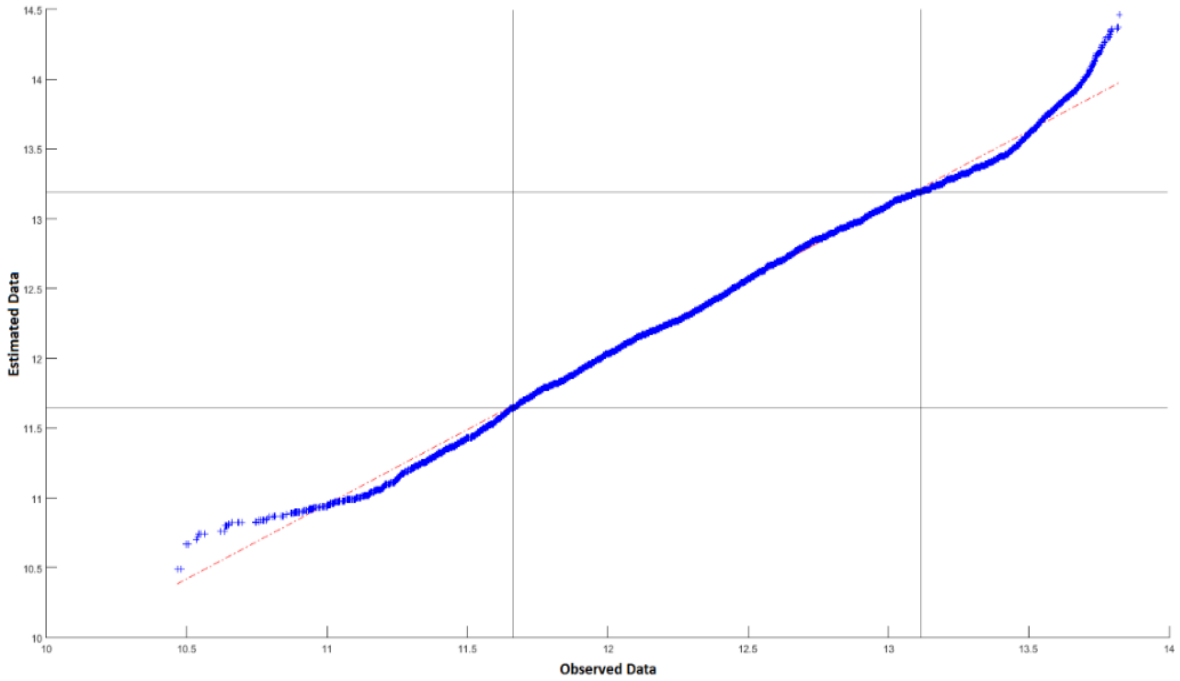


Figure 9.22: *Observed data against estimated data relationship*

In *Figure 9.22* we observe the bad estimates related to the values of bathymetry above 13.1 meters and below 11.7 meters. The region between 11.7 and 12.1 has a good prediction result. Thus, although the data tends to follow a trend of normal distribution in intermediate values, we confirm the existence of two regions with distinct behaviour, as shown in *Figure 9.23*. These two regions exert influence over the regression algorithm based on kriging, and therefore estimates using the entire dataset result in failures for some locations.

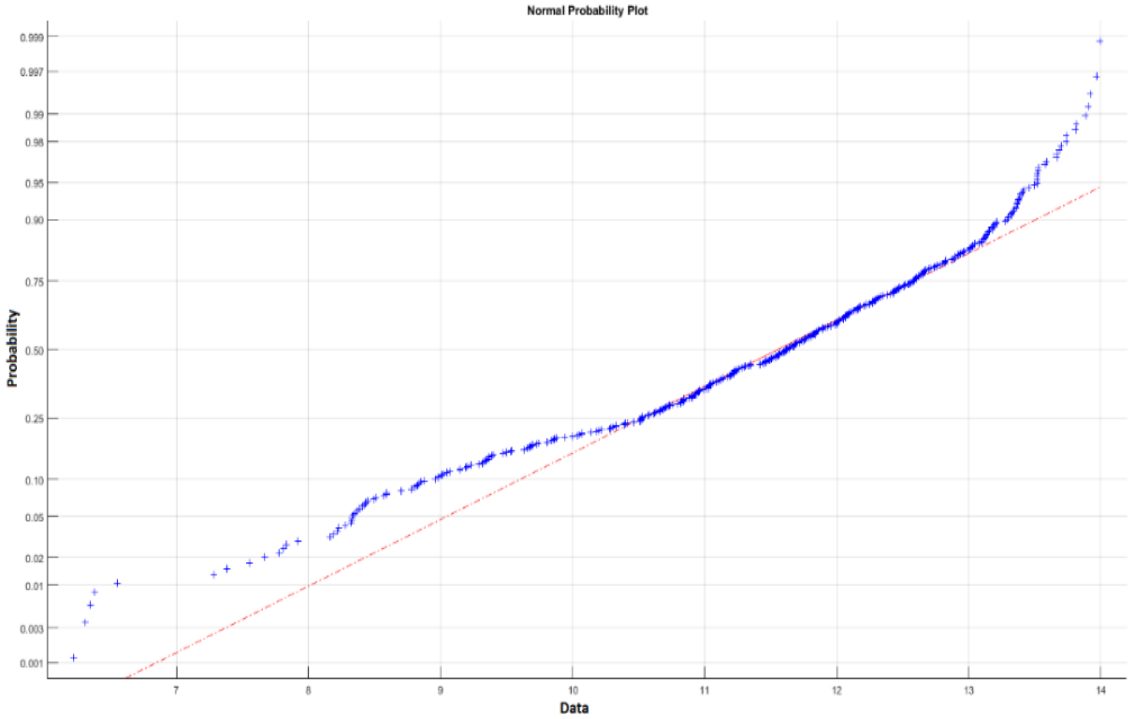


Figure 9.23: *Normal probability plot from original data*

We note that there is a visual similarity between *Figures 9.22* and *9.23*, which means that we can evaluate the adequacy of the original distribution of the data values to verify the feasibility of making geostatistical estimates. If the data distribution represents a deviation from the normal distribution, there will probably be a failure in the prediction process using Kriging. In this case, we intuit that according to the data distribution, we can divide the dataset into distinct regions, extracting the main prediction parameters for each region. Therefore, we could perform the kriging process for each region separately, joining the results in a single set later. In this case, it is necessary to apply or develop an analytical method to divide the regions before the geostatistical prediction process.

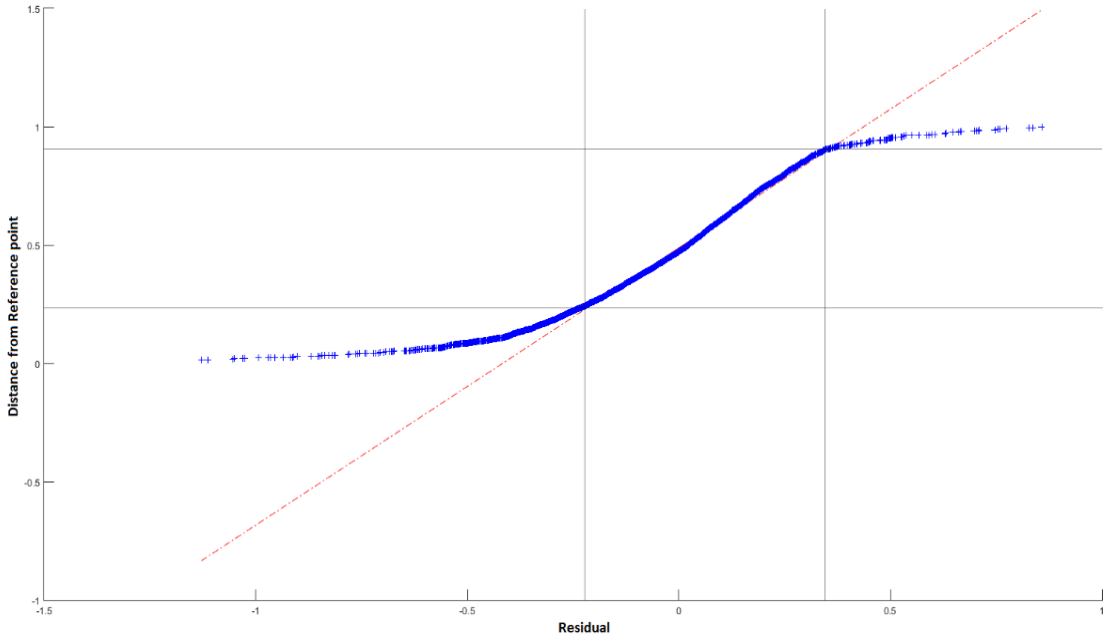


Figure 9.24: *Plot of relationship between residual and reference point distance from estimated point*

As an additional consideration to guarantee a minimum level of confidence, we will choose the reference point using as a parameter of choice the half of the maximum distance for surveying bathymetric points according to the standard (IHO 2008), representing a distance of 1 meter, based on the comparative analysis carried out by the IS validation process. Thus, we will have some assurance that the estimates will be accepted by the international standard reference.

Figure 9.24 shows the illustration of the residual value related to the distance between the observed points and the reference points. We observed that the maximum distance between points was 1 meter, so the method used to choose the reference points is within the limits established by our previous hypothesis.

9.4 Geostatistical Semivariogram Analysis, Kriging Estimations and Results Appling Clustering

After the application of the geostatistical estimation and analysis presented in the previous sections, it was observed that only 87.82% of the estimated data can be accepted within the IHO (IHO 2008) standard for classification between Special Order, Order 1A and Order 1B. This would make it impossible to apply the method for the construction of bathymetric maps to support navigation in certain coastal regions or rivers that demand a more accurate confidence level. In this case, we propose a second way of approaching the estimation problem, which, like the previous approach, will not use any auxiliary information to support the geostatistical prediction. It was observed that the desired confidence level around 95% was not reached due to the trend in some directions of the sampled surface. Thus, the extraction of parameters for the geostatistical estimation using only one adjustment on the experimental semivariogram is not enough to capture all the nuances of variability and behaviour of the sample values.

The spatial data analysis applied in this paper actually relates to the quest for spatial dependence between observations, the basic and intuitive assumption will be: "if we have two samples close together, their values are more correlated than distant ones. ". Therefore, the idea of seeking a spatial correlation between acquired samples using a transect method in a distribution of measurements on a plane was mathematically formalized by Whittle in his work (Whittle 1954). After that, Besag (1972) concludes that self-covariance related to some spatial process cannot be mathematically expressed using a standard analytic technique. Considering a discrete sample scheme, Besag suggests the use of conditional expectation models against the autoregressive model in bilateral spatial situations involved in normal processes. On the other hand, Haining in his work (Haining 1991) assumes that pairs of spatial observations are independent for parametric and non-parametric association test applications. In his paper, Haining noted that correlation measurements may reflect the properties of covariance that are dependent on the location and variables means involved in a bivariate system. Thus, this dependence can be used as an indicator of the spatial correlation existing in the distribution of the sample data, but with the exception of non-use as the only indicator.
A modelling proposal aimed at solving the problem of identifying the different characteristics and property dependent on location was proposed in (Ripley 1977) to approach spatially distributed processes. In this work, one of the solutions was based on a grouping process. This solution divided the sample space into subsets because the total set of samples would apparently not be covered by the assumption of the Poisson distribution, mainly due to trends in certain directions Distribution of samples on the surface (Nelson et al. 2003), (Ross 2009) and (Ripley 1977). The division of the process into subsets G through a grouping technique such as *k-means* allows regions with similar characteristics to be grouped, avoiding the effect of trends, approaching the characteristics of each subset closer to a Poisson distribution in R^2. Summarizing, we propose a technique of geostatistical analysis from the segmentation of the initial set of samples, resulting in the creation of subsets with similar trends using the grouping technique called *k-means*.

The K-means clustering technique uses an unsupervised learning algorithm (Hastie et al. 2009). The K-means algorithm finds clusters based on a centre and a radius from this centre, considering a set of unmarked data (Wang 2006), (Fred & Jain 2006) and (Hastie et al. 2009). Thus, we must define the number of desired centres, which will consequently result in the number of clusters that will be generated as a processing result, having as preliminary information only the data of the bathymetry samples. The approximate determination of this number of centres, or clusters based on existing data, became the first challenge of solving the problem. The division of the sample dataset into subsets will permit the adjustment of the experimental semivariogram more precisely, resulting in more reliable predictions through the use of geostatistics. In order to guide the process of determining the number of subsets in this work, an empirical expression called the Grouping Factor (GF) was developed. The GF correlates the values of the correlation coefficients of the data with respect to the coordinate axes, which allows to determine the probable number of regions that we must segment the total of samples in disjoint regions for application of the clustering algorithm.

The determination of GF is initiated by calculating the correlation coefficient, derived from descriptive statistics, the coefficient and correlation measures the degree of correlation between two variables (Richardson & Hemon 1981), (Hauke & Kossowski 2011) and (Haining 1991). The correlation coefficient can be extracted according to different methods, receiving different names for each type of method. In this work, three different types of methods were used to calculate the correlation coefficient corresponding to each coordinate axis, these are called Pearson's, Spearman's and Kendall's correlation coefficients. The process of extracting the GF consists of two steps. In the first step two of the methods of calculating the correlation coefficient (ρ) were used, on the bivariate system composed by an axis of the geographical coordinates described in Latitude and Longitude (Lat, lon) and by the

value of the sample (z). In this case, we will find the correlation coefficients for each axis as ρ_{Lat} for the bivariate system (Lat, z) and (Lon, z). In the second step, we will extract numerical values, which are indicators of the correlation coefficient for each axis, that is, a coefficient indicator on the vertical axis and coefficient indicator on the horizontal axis.

In this work, we follow the open-set abstraction with inherent properties of the topological space, more concretely, the total sample space set will be defined as a union of subsets, in which the mathematical formalist is described in (Benedikt et al. 2006). In the representation of topological spaces, the open sets can be represented by a subspace in which each one includes different characteristics, since they represent different spaces.

9.4.1 Dataset Correlation Coefficient

The original sampling data of the bathymetric survey consists of a digital file in which each registered sample is stored with its corresponding geographical coordinate. The generated digital file format represents the sample value on the Latitude and Longitude plane in the format (x, y, z), where z is the sample value. The digital sample file is decomposed into a bivariate arrangement: latitude versus sample (x, z) and Longitude versus sample (y, z). Thus, considering the bivariate correlation coefficient analysis (ρ) for each bivariate system, the correlation, or dependence between the variables of the new arrangement will be expressed through a numerical value ranging from -1 (strong negative correlation), through 0 (indicating lack of correlation or total randomness) to value 1 (strong positive correlation). In this case, a question arises within the context of our study: How can we extract useful information about spatial dependence within the bathymetric data using the bivariate correlation coefficient analysis in the two directions of interest? A preliminary study stage was started seeking the understanding and use of studies on the correlation of the variables.

9.4.1.1 Pearson's Correlation Coeficient

From Pearson's paper (Pearson 1900), we will have one formal definition for describing the linear correlation coefficient of two random variables within a data set. This work was a consequence of published work Pearson & Filon 1898). The Pearson correlation coefficient or product-moment correlation coefficient is the linear type, it is currently the most used form to calculate the correlation coefficient for a bivariate system (x_i, y_i),, where i varies from 1 to n (Press et al. 1992). The Person correlation coefficient was formally defined as the covariance between two variables divided by the product of the standard deviations of each variable (Pearson 1900). The mathematical formulation is expressed in equation 9.16.

$$\rho_{x_i,y_j} = correlation(x, y_j) =$$

eq. 9.16

$$= \frac{covariance(x_i, y_j)}{\sqrt{variance(x_i)}\sqrt{variance(y_i)}} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})^2}}$$

Where $\bar{x}$ is the mean of the measurement values x_i and $\bar{y}$ represents the mean of the measurement values for y_j. Using the Matlab function "corrcoef" and the algorithm of the "coor" function, the values of the correlation coefficients ρ_{x_i,y_j} are generated in the form of a $2x2$ matrix. Note that the value of ρ_{x_i,y_j} calculated using the algorithm of Person ignores individual characteristics of the point distributions. Therefore, we can not probe the null hypothesis in its distribution for all bands considering a single bivariate system, but using the composition of systems such as two bivariate systems, we increase this possibility.

9.4.1.2 Spearman Correlation Coeficient

Spearman's correlation coefficient, or classification order correlation (Press et al. 1992), can evaluate the monotonic relationship between two continuous or discrete variables, that is, the variables tend to change together, but not necessarily to a constant rate. Spearman's correlation has no assumption that the relationship between variables is linear. Thus, Spearman's correlation is a nonparametric correlation measure, this means, unlike Pearson's correlation, that there is no requirement for normality nor the need to measure the strength of association between two variables. The expression for calculating the Spearman correlation coefficient is given in equation 9.17.

$$\rho_{x_i,y_j} = 1 - \frac{6\sum d_i^2}{n^3 - n} \qquad \textit{eq. 9.17}$$

Where the value d_i is the classification distance between variables, considering the bivariate system defined as (lat, y_j) for the latitude axis e (lon, x_i) for the length axis. The value of n represents the number of pairs that measure the correlation (Press et al. 1992). The existence of a zero value (0) as a result of the Spearman correlation calculation, different from the case where we calculated the correlation coefficient using Pearson's, does not imply that there is no relation between the variables. This is a motivation not to use just one kind of correlation coefficient in a decision-making process within our method.

9.4.1.3 Kendall τ Correlation Coeficient

This coefficient was introduced by (Kendall 1938), and can be used as an alternative to the Spearman correlation coefficient for organized row data (Hauke & Kossowski 2011). The Kendall coefficient τ performs a mean operation between pairs of samples. The main advantage of the Kendall τ coefficient according to the work (Hauke & Kossowski 2011) is summarized, with the expression: "The main advantage of using Kendall τ lies in the fact that its distribution has slightly better statistical properties in terms of probabilities, according to concordant and discordant pairs are observed. " As an example, we can consider N data points (x_i, y_i), and $\frac{1}{2}N(N-1)$ pairs of points, where a point of the data cannot be paired with itself, in addition, points in any order can be considered as a pair (Press et al., 1992). We can define the technique using the formalism of (Press et al. 1992) and (Christensen et al. 2005), in which it describes that the Kendall correlation τ between two sets of random variables $X = (x_i)$ and $Y = (y_i)$ with $i = (1,2,3,\ldots,N)$, and it can be determined using the assumptions published in (Press et al. 1992):

i. We call a pair of concordants if the relative ordering of the rows of the two elements x_i is the same relative to the relative ordering of the rows of the two corresponding elements y_i;

ii. We call a pair of discordants if the relative ordering of the rows of the two elements x_i is opposite to the relative ordering of the rows of the two corresponding elements y_i;

iii. If there is a tie in the classifications of the two elements x_i or in the rows of the two elements y_i, then we do not classify the pair as concordant or discordant.

- If the tie is in the x, we will call the pair of "extra pair y".
- If the tie is on y, we will call the pair an "extra pair x".
- If the tie is in the x's and y's, we do not name it.

The Kendal coefficient τ can be described as a simple combination of the previously mentioned terms defined as:

$$\boldsymbol{\rho_\tau = \frac{concordant - discordant}{\sqrt{concordante + discordant + extra - y}\sqrt{concordante + dicordante + extra - x}}} \qquad \textit{eq. 9.18}$$

As the correlation coefficients of Pearson and Spearman, Kendal τ will range from -1 to 1.

In our experimental implementation, we identified a very similar result in the values of the correlation coefficient calculated with the Spearman and Kendall τ algorithms, on the other hand, the most discordant value was found with the Pearson algorithm. The best way to determine a numerical value for the cluster factor is to use at least two methods, if the values are divergent, we will apply the third method as a tie-breaking criterion. This method of determining a value that represents the ideal value of clusters within a dataset was inspired by the reference (Wang 2006), this work was applied to gene expression data. Thus, we could find a relation between the coefficients of correlation between the coordinated axes and the techniques of groupings applied in spatial data.

9.4.2 Determine of the Cluster Factor (CF) by the Correlation Coefficient

The determination of the Cluster Factor (CF), as a function of the correlation coefficient, intends to provide the cluster algorithm with a notion of the ideal number of clusters within the data set. Thus, we can increase the confidence level of the predictions involving the study and application of the geostatistical method insofar as it helps to divide the data set into groups whose elements have statistical characteristics that respect the assumption of the Poisson distribution. Therefore, each cluster encompasses samples with structural characteristics and similar trends. After the data characterization using the experimental (empirical) representation, that relates the values of the correlation coefficients of the bivariate system to the generation of the CF is developed considering the following steps:

1. Choice of two distinct methods to calculate the correlation coefficient (ρ) of the bivariate system, in our case:

   - ρ_{lat} - For the bivariate system (latitude, sample); and
   - ρ_{lon} - For the bivariate system (Longitude, sample).

2. Extraction of the correlation coefficient indicators for each axis of the bivariate system:

   - Vertical Axis Coefficient Indicator (VA_i); and
   - Horizontal axis coefficient indicator (HA_i).

The process of calculation and determination of the indicators will be carried out according to the chosen method for the calculation of the correlation coefficient.

1) For Pearson's Correlation Coefficient:

1. We extract the determinant of the matrix ρ, which represents the correlation coefficient generated by the Pearson algorithm, resulting in the value $\rho_{\det}$ for each bivariate system;

2. We compare the values of $\rho_{\det}$ for both bivariate systems and consider the highest value of $\rho_{\det}$ as the Indicator VA_i. Thus, the lower value will be the Indicator HA_i.

2) For Spearman e Kendall τ Coeficiente de Correlação:

1. Comparamos os valores do coeficiente de correlação ρ para ambos os sistemas bivariados, determinamos o maior dos valores do coeficiente de correlação como VA_i e o menor como HA_i.

2. We compare the values of the correlation coefficient ρ for both bivariate systems, we determine the highest values of the correlation coefficient as VA_i and the smaller one as HA_i

After the comparisons and adjustments of the correlation coefficients relative to each bivariate system generated, we can calculate the CF, thus extracting the best indication of the ideal number of clusters that we will enter as parameter in the K-means aggregation method, whose operation is represented in the expression 9.14

$$CF = rounded\left(\frac{VA_i}{HA_i}\right) \qquad exp.\ 9.14$$

Note 1: If the values of the CF, after rounding are different, a third method for extraction of the correlation coefficient should be used as a tie-breaking criterion between the previously calculated CF values. Note that we consider only absolute values to represent the value of ρ.

To illustrate graphically the definition operation of CF, we developed the graphical idea of the correlation plane ($\rho_{lat,z_i}, \rho_{lon,z_j}$). In this plan, we represent the distribution or spreading form of the data samples in relation to the two axes, representing the value of the CF in relation to this distribution. The graphical representation is valid only for the calculation of the correlation coefficient using the methods of Pearson and Spearman.

The correlation plane (CP), represented in *Figure 9.25*, will graphically represent the reliability regions for use or not of the correlation plane method to determine how many clusters the set of samples can be divided to maximize the confidence level of the Predictions using the geostatistical method.

The correlation plane can be simplified due to the symmetry between the occurrences of correlation coefficient values and the regions described in *Figure 9.25*. Thus, we can present the CP using only the absolute values of the correlation coefficients. In this way, the simplified representation of the CP represented by *Figure 9.26*, includes the ideal values of clusters related to the relationship of the correlation coefficients with respect to the two bivariate systems. Recalling that this plan was only validated for the use of correlation coefficients using the algorithms of Pearson and Spearman.

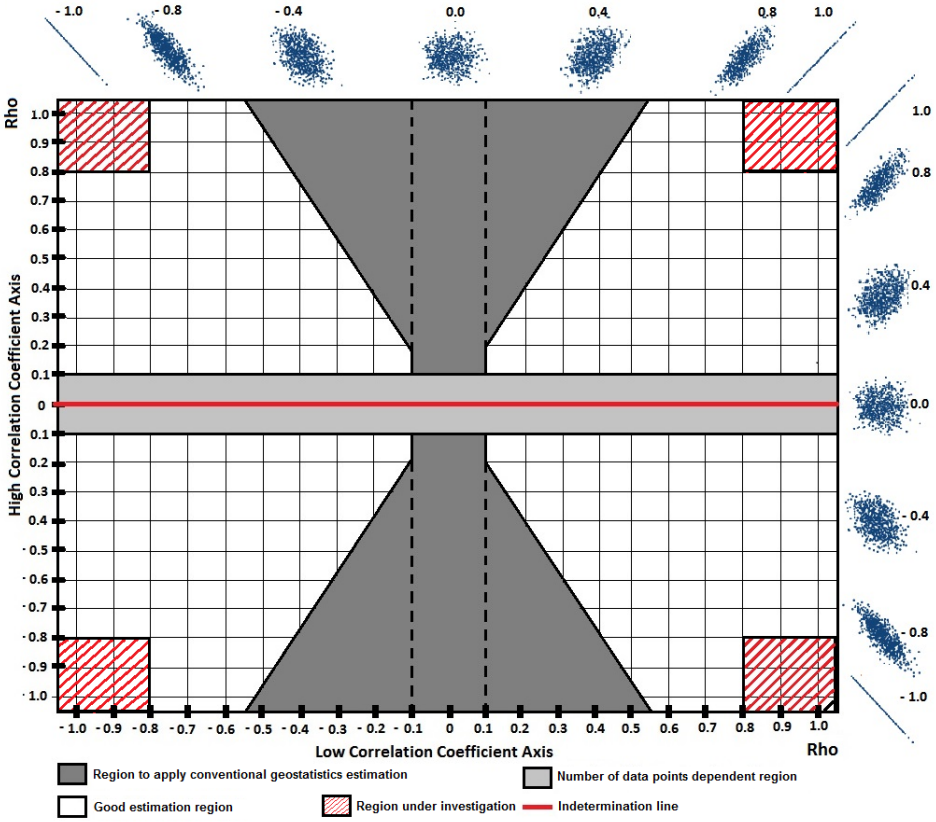


Figure 9.25: *Correlation coefficient to cluster number map grid, scatter plot representation*

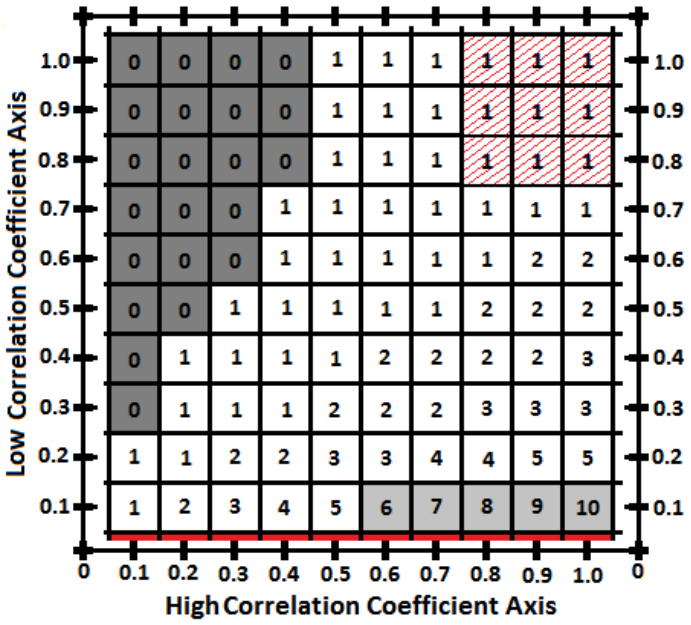


***Figure 9.26**: Correlation coefficient to cluster simplified map diagram*

In the correlation plan model, we represent only 100 relationship sub-regions, because in this case, we only need to consider only one decimal digit. Depending on the number of existing data points, the different dimensions and trends present in the data collection region, we can consider increasing the number of decimal digits, so we will increase the cluster number. The information extracted from the correlation plane will be similar to the information of the calculation of the cluster factor (CF), this information will be used as a parameter for the k-means cluster algorithm. The clustering process, as well as geostatistical data processing, was performed on a computer with 64, Windows 2010 and Matlab software architectures.

Considering our data, we will use the expression (9.14), to extract the ideal cluster number or CF, the analytically calculated result was 3. The methodology used to extract this information was done through the three-step procedure described below:

Step 1: We chose as the first two algorithms for the extraction of the correlation coefficient and subsequent calculation of the correlation indicators the algorithms of Pearson and Spearman.

Step 2: We then determine the vertical and horizontal correlation indicators VA_i and HA_i:

Pearson's Case:

$$\rho_{lat,z} = \begin{pmatrix} 1 & 0.1741 \\ 0.1741 & 1 \end{pmatrix}; \text{ then, } \rho_{\det\_lat} = det\begin{pmatrix} 1 & 0.1741 \\ 0.1741 & 1 \end{pmatrix} = 0.9697 \qquad \textit{eq. 9.19}$$

$$\rho_{lon,z} = \begin{pmatrix} 1 & 0.6298 \\ 0.6298 & 1 \end{pmatrix}; \text{ then, } \rho_{\det\_lon} = det\begin{pmatrix} 1 & 0.6298 \\ 0.6298 & 1 \end{pmatrix} = 0.6034 \qquad \textit{eq. 9.20}$$

Comparing both $\rho_{\det\_lat}$ and $\rho_{\det\_lon}$ values, we will have $\rho_{\det\_lat} > \rho_{\det\_lon}$, then:

$$VA_i = \rho_{\det\_lat} = 0.9697 \qquad \textit{exp. 9.15}$$

$$HA_i = \rho_{\det\_lon} = 0.6034 \qquad \textit{exp. 9.16}$$

Spearman Case:

$$\rho_{lat} = 0.1951 \qquad \textit{exp. 9.17}$$

$$\rho_{lon} = 0.5958 \qquad \textit{exp. 9.18}$$

Comparing both ρ_{lat} and ρ_{lon} values, we will have $\rho_{lon} > \rho_{lat}$, then:

$$VA_i = \rho_{\text{lon}} = 0.5958 \qquad \textit{exp. 9.19}$$

$$HA_i = \rho_{\text{lat}} = 0.1951 \qquad \textit{exp. 9.20}$$

Step 3: Determining values for:

Pearson's Case:

$$CF_P = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.9697}{0.6034} = round(1.60) = 2; \qquad \textit{eq. 9.21}$$

Spearman Case:

$$CF_S = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.5958}{0.1951} = round(3.05) = 3; \qquad \textit{eq. 9.22}$$

As the values of CF_P and CF_S are divergent, we need to calculate the CF using a third method, Kendall τ.

Kendall τ Case:

$$\rho_{lat} = 0.1368 \qquad \textit{exp. 9.21}$$

$$\rho_{lon} = 0.4375 \qquad \textit{exp. 9.22}$$

Comparing both ρ_{lat} and ρ_{lon} values, we will have $\rho_{lon} > \rho_{lat}$, then:

$$VA_i = \rho_{\text{lon}} = 0.4375 \qquad \textit{exp. 9.23}$$

$$HA_i = \rho_{\text{lat}} = 0.1368 \qquad \textit{exp. 9.24}$$

$$CF_S = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.4375}{0.1368} = round(3.19) = 3; \qquad \textit{eq. 9.23}$$

In this case, since most of the results point to a CF = 3, we use this value as the input parameter of the k-means algorithm to partition the entire data set into three disjoint regions. Thus, after division, we can apply the geostatistical method, making predictions by the kriging algorithm with parameters suitable for each of the grouping regions.

9.4.3 Clustering Bathymetric Dataset Based in K-means Algorithm

The technique of dividing the dataset into disjoint groupings will be applied to our data set. The objective will be to increase the reliability in the predictions made through the geostatistical analysis, as a result of the elimination or decrease of the trend level present in the data of each grouping. Therefore, from the initial consideration, it consists of a set of measurement data (z_i) extracted from a single bathymetric surveying operation, which forms a single P region within the context of regionalized variables (Matheron 1971b). The region P will be expressed as a set of sub regions p_i, mathematically represented by the expression 9.25.

$$P = \bigcup_{i \leq n} p_i \qquad \textit{exp. 9.25}$$

Where $p_i \cap p_j = \emptyset$, if $i \neq j$. Thus, the clustering operation performed on the region P will form disjoint groupings $p_1, \ldots, p_n$, where n represents the total number of portions or grouping. The size of each cluster may be the same

or different, but must have a minimum number of data, approximately 50 samples, ideally the value should be above 100 samples. The cluster algorithm k-means uses an iterative relocation technique based on similarity, in order to improve the position of the centroids within each group, this algorithm represents a generalization of the classic clustering criterion of the sum-of-squares method SSQ clustering criterion (Hans-Hermann 2008). The application of the FC, developed in this work, intends to satisfy the classical request of the k-means algorithm, which consists in determining the desired number of classes in which the data set must be divided into subsets that do not overlap. The algorithm will interactively minimize the Euclidean distance between the reference centroid and each sample, to include points in each partitioned region. The result is a set of compact clusters that are as separate as possible.

The data set was plotted in *Figure 9.1*. The total set of available samples has 424 points, but only 220 points can be used because they are the points that are in the same region of the reference dataset, as observed in section 9.3.3, and represented in *Figure 9.20* by the intersection between Black spots (SBES samples) and coloured (MBES samples). In this way, we can proceed to the study of prediction reliability by comparing predictions and reference data. So, we proceeded with the initial analysis to find the ideal number of clusters to divide the data.

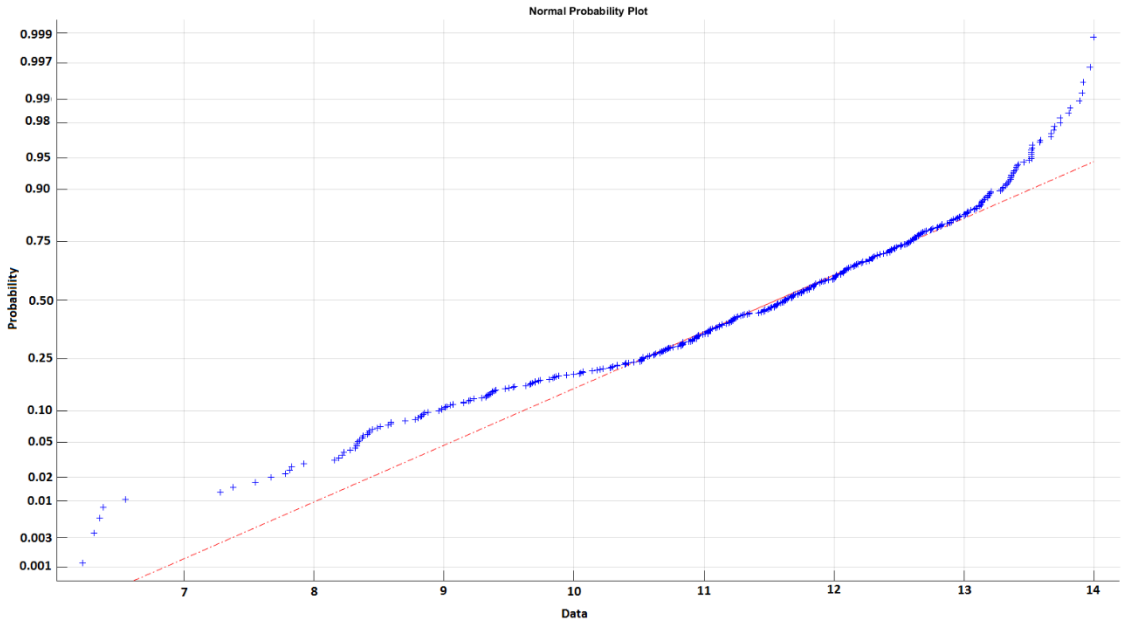


Figure 9.27: *Probability plot against normal distribution of 220 data points selected*

Observing the distribution of the data samples values in front of the line of a normal distribution, through the graph of probabilities distribution, we can observe deviations or outliers in relation to the normal distribution. Deviations or outliers may represent the existence of a trend in the data. This graphical representation can be visualized in *Figure 9.27*.

Table 9.11: *Simple statistical analysis*

Parameters	**Values**
Minimum	10.74
1st Qu.	11.86
Mean	12.45
Median	12.45
3rd Qu.	13.13
Maximum	14.00
Standard Deviation	0.79
Skewness	-0.11

We determine the outliers with the denomination of atypical points, that is, all points above and below two standard deviations (2σ). In *Table 9.11*, we summarize the statistical characteristics for the set of considered samples.

1) Determination the number of clusters using CF

According to the algorithm discussed in this chapter, in order to guarantee a more accurate prediction of values for non-sampled points, we will perform the process of determining the ideal cluster number for dividing the data samples in a few steps as shown below.

Step 1: Determination of the correlation coefficients ($\rho_{lat,z}$ and $\rho_{lon,z}$), in addition to calculating the correlation indicators ($VA_i$and HA_i) for the bivariate system represented by the relation. The coordinate axes and the value of the samples. We first chose the Person and Spearman methods.

Pearson's Case:

$$\rho_{lat,z} = \begin{pmatrix} 1 & 0.2289 \\ 0.2289 & 1 \end{pmatrix}; \text{ then, } \rho_{\det\_lat} = det\begin{pmatrix} 1 & 0.2289 \\ 0.2289 & 1 \end{pmatrix} = 0.9476 \quad \textit{eq. 9.24}$$

$$\rho_{lon,z} = \begin{pmatrix} 1 & 0.5661 \\ 0.5661 & 1 \end{pmatrix}; \text{ then, } \rho_{\det\_lon} = det\begin{pmatrix} 1 & 0.5661 \\ 0.5661 & 1 \end{pmatrix} = 0.6795 \quad \textit{eq. 9.25}$$

Comparing both $\rho_{\det\_lat}$ and $\rho_{\det\_lon}$ values, we will have $\rho_{\det\_lat} > \rho_{\det\_lon}$, then:

$$VA_i = \rho_{\det\_lat} = 0.9476 \quad \textit{exp. 9.26}$$

$$HA_i = \rho_{\det\_lon} = 0.6795 \quad \textit{exp. 9.27}$$

Spearman Case:

$$\rho_{lat} = 0.2772 \quad \textit{exp. 9.28}$$

$$\rho_{lon} = 0.5671 \quad \textit{exp. 9.29}$$

Comparing both ρ_{lat} and ρ_{lon} values, we will have $\rho_{lon} > \rho_{lat}$, then:

$$VA_i = \rho_{\text{lon}} = 0.5671 \quad \textit{exp. 9.30}$$

$$HA_i = \rho_{\text{lat}} = 0.2772 \quad \textit{exp. 9.31}$$

Passo 2: Determine values to CF.

Pearson's Case:

$$CF_P = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.9476}{0.6795} = round(1.39) = 1; \quad \textit{eq. 9.26}$$

Spearman Case:

$$CF_S = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.5671}{0.2772} = round(2.04) = 2; \quad \textit{eq. 9.27}$$

As the values of CF_P and CF_S are divergent, we need to calculate the CF using a third method, Kendall τ again.

Kendall τ Case:

$$\rho_{lat} = 0.1731 \quad \textit{exp. 9.32}$$

$$\rho_{lon} = 0.4012 \quad \textit{exp. 9.33}$$

Comparing both ρ_{lat} and ρ_{lon} values, we will have $\rho_{lon} > \rho_{lat}$, then:

$$VA_i = \rho_{\text{lon}} = 0.4012 \quad \textit{exp. 9.34}$$

$$HA_i = \rho_{\text{lat}} = 0.1731 \quad \textit{exp. 9.35}$$

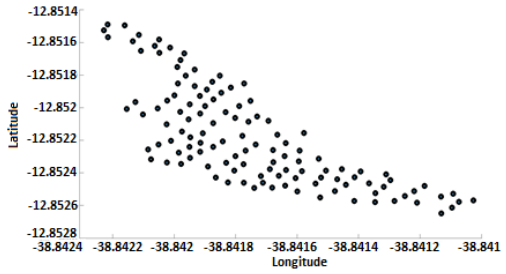
$$CF_S = rounded\left(\frac{VA_i}{HA_i}\right) = \frac{0.4012}{0.1731} = round(2.31) = 2; \quad \textit{eq. 9.28}$$

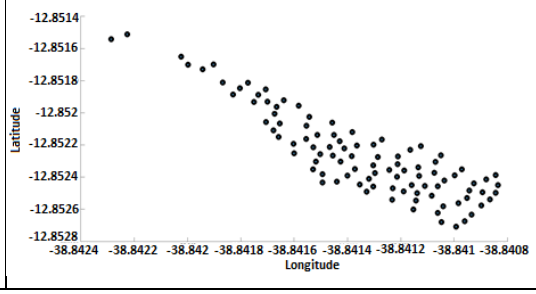
In this case, as the majority results points to $CF = 2$, we will use this in the K-means cluster algorithm to partitioning our dataset in two disjoints regions.

2) Appling K-Means Clustering using the Cluster Factor (CF)

The K-Means algorithm that was used in this work is based on the algorithm inspired by (Helmuth Spath 1985) and (Seber 2004). The CF value, previously calculated, will be used to divide the dataset into two regions, both regions are indicated by Region 1 and Region 2. Therefore, after dividing the region, we must start the geostatistical analysis and processing of each region separately. *Table 9.12* shows the previously separated regions and a statistical summary of the samples.

Table 9.12: *Scatter data from both disjoints regions created by K-Means cluster*





Exploratory Data Analysis Region 1	
Parameters	**Values**
Minimum	10.74
1st Qu.	11.55
Mean	11.96
Median	11.86
3rd Qu.	12.22
Maximum	12.58

Exploratory Data Analysis Region 2	
Parameters	**Values**
Minimum	12.59
1st Qu.	12.92
Mean	13.18
Median	13.20
3rd Qu.	13.42
Maximum	14.00

Table 9.12 shows the differences between the extreme values, that is, the maximum and minimum values of the sample set in relation to the mean for each region. The values can be seen in *Table 9.13.*

Table 9.13*: Differences from mean to extreme sample values for each region and entire dataset.*

Parameters	Region R1	Region R2	Both Regions Together
Dif. Mean-Maximum (m)	0.62	0.82	1.55
Dif. Mean-Minimum (m)	1.22	0.59	1.71

Reducing the differences between the extreme values of the samples in relation to the mean after dividing the regions will result in a greater precision in the kriging estimates, regardless of the method used. *Figure 9.28* shows the histogram of both separated regions. The R1 region is represented by the left side illustration and the right side is the region 2 representation. Despite the non-Gaussian distribution model in both regions, the decrease of the differences between the extreme values of the samples and the mean, allows the predictions to have a smaller error, when compared to the values obtained if we used the whole set of samples in the same prediction operation.

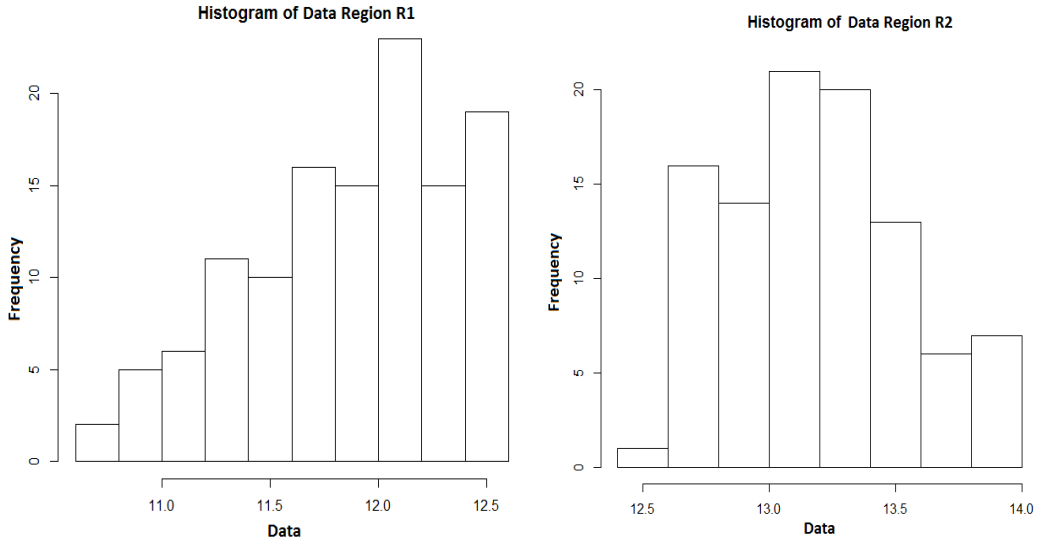


***Figure 9.28**: Box-plot from both clustered regions*

The parameters values extracted after adjustment of the experimental half-variogram, using a theoretical semi-variogram model for both regions, are shown in *Figure 9.29* graphically and numerically in *Table 9.14*. These values will be used in the estimation process using the kriging algorithm, which will result in value propositions for not sampled points within the regions of interest.

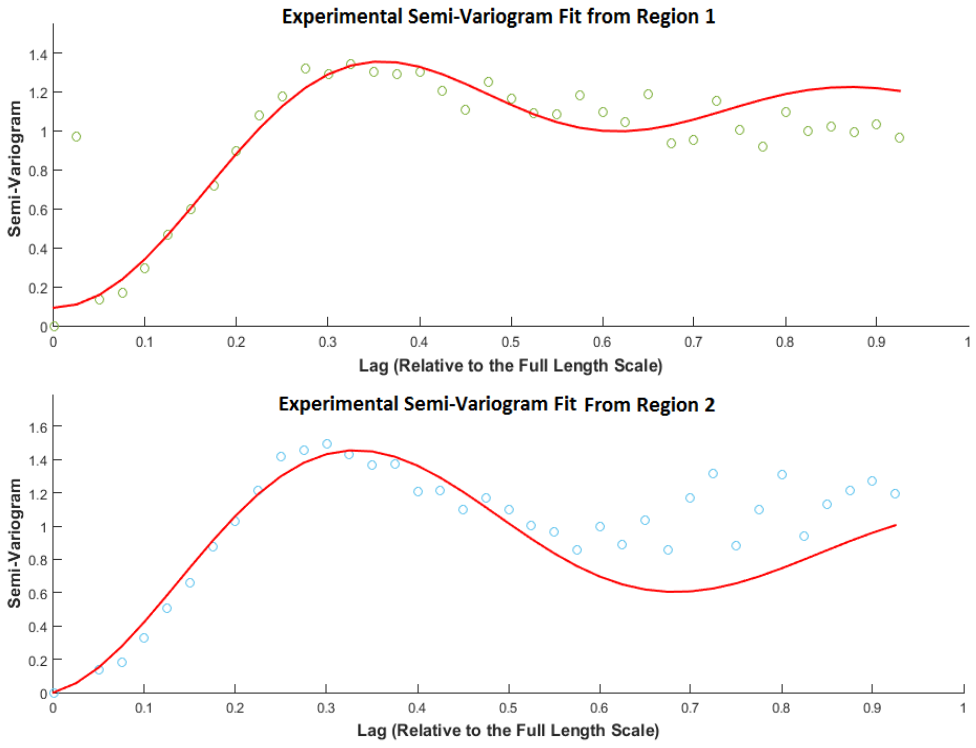


***Figure 9.29:** Experimental semi-variogram fit for both regions*

***Table 9.14**: Variogram parameters from both regions R1 and R2*

Variogram Parameter from Region R1		**Variogram Parameter from Region R2**	
Theoretical Model Fit	Gaussian-Bessel	**Theoretical Model Fit**	Exponential-Cosine I
Nugget	0.0478	**Nugget**	2.1842 x 10^{-13}
Sill	1.1307	**Sill**	0.91632
Length	---------	**Length**	0.64398
Hole Scl.	12.5271	**Hole Scl.**	8.9425
Range	0.95	**Range**	0.95
Resolution	0.025	**Resolution**	0.025

After an iterative process to choose the best model of theoretical semi-variogram, we arrived at the adjustment using the Gaussian-Bessel and Exponential-Cosine 1 models for regions 1 and 2, respectively. The parameters used in the kriging estimator are expressed in *Table 9.15* in addition to the definition of some other parameters.

***Table 9.15:** Additional kriging parameters to estimate data at unsampled points for both regions*

Additional Kriging Parameter to Estimates Over Region R1		**Additional Kriging Parameter to Estimates Over Region R2**	
Model	Ordinary Kriging	**Model**	Ordinary Kriging
Scheme	Point to Point	**Scheme**	Point to Point
Search Radius	0.3	**Search Radius**	0.3
Min. Kriging Points	10	**Min. Kriging Points**	10
Max. Kriging Points	30	**Max. Kriging Points**	30
Relative Variance	2.5	**Relative Variance**	2.5
Estimation Resolution	1 meter	**Estimation Resolution**	1 meter

As previously described, the process of prediction or estimation of values in non-sampled points began in the division of the sampling region into disjoint sub-regions (R1 and R2).

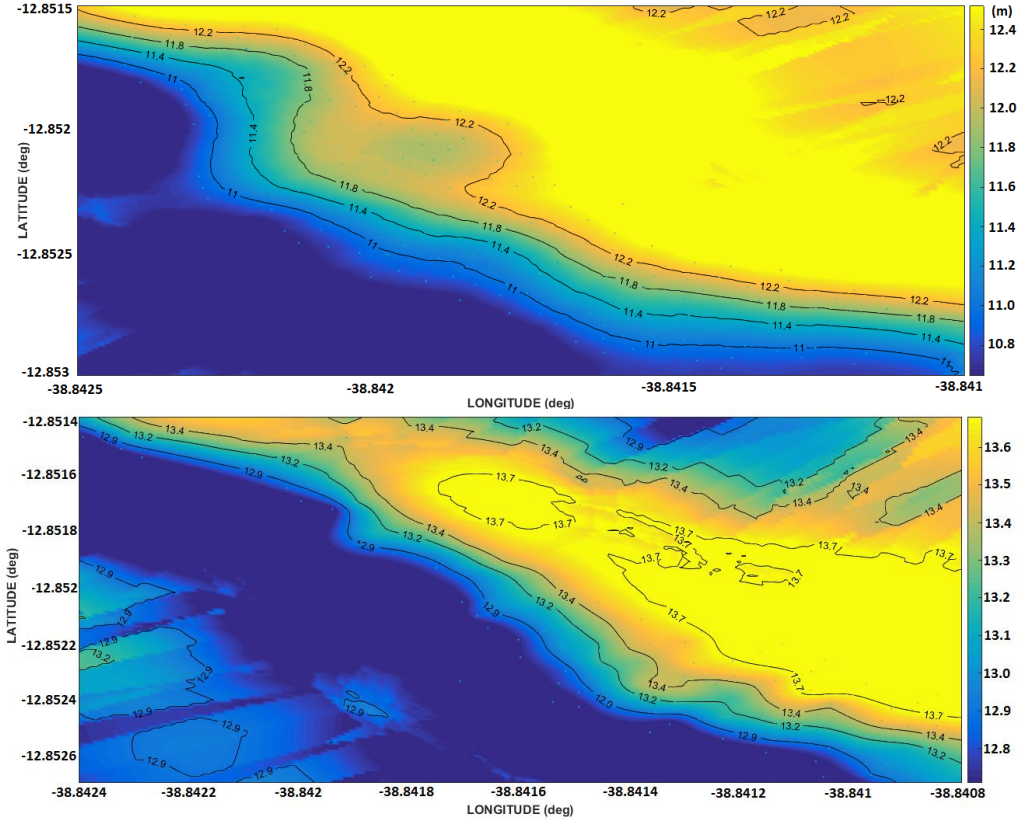


***Figure 9.30:** Contour map from both regions estimated separately*

Under each region, the operation of prediction value is applied, with different parameters for each region, according to the specific characteristics of the regions, keeping only the parameter related to the desired resolution at the end of the individual prediction process. The final result will be composed of the union of the two predicted regions, grouping them and forming a single estimated region. *Figure 9.30* shows the estimate for Regions R1 (illustration above) and R2 (illustration below).

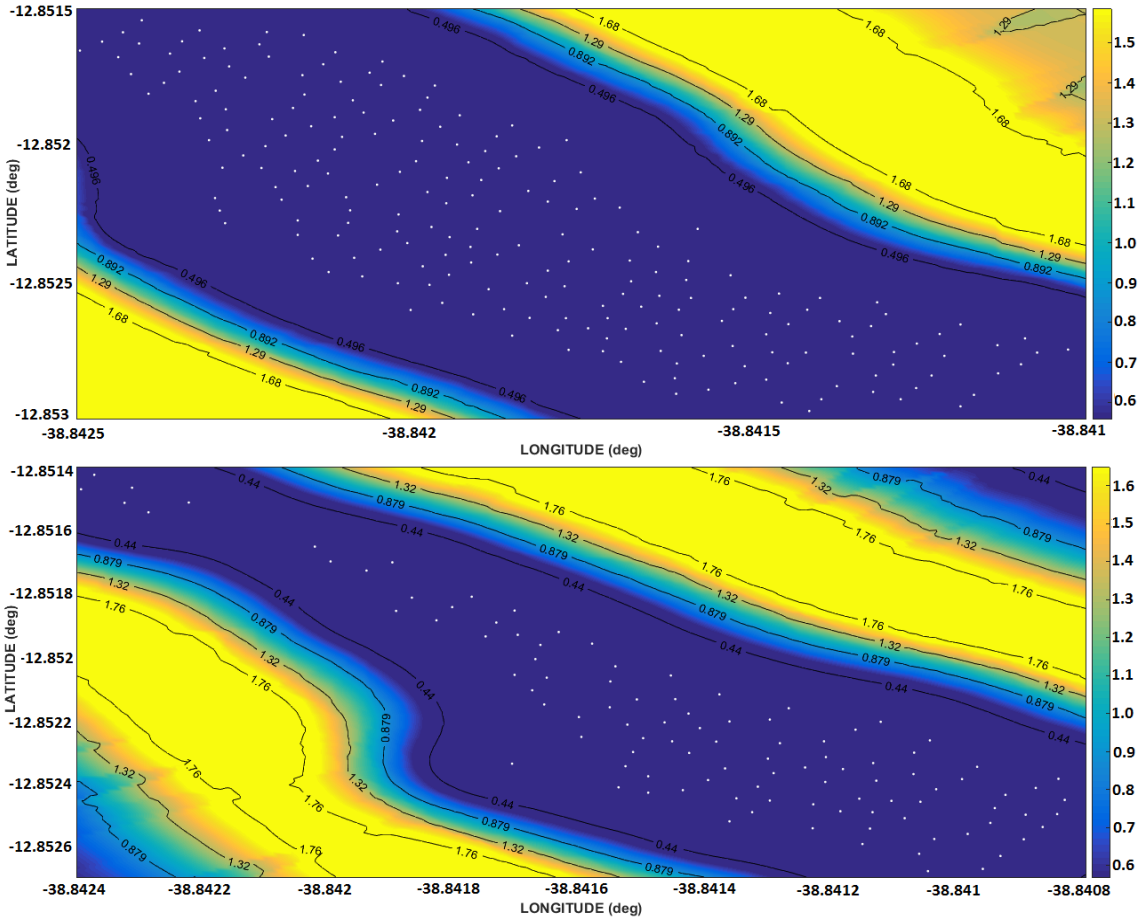


Figure 9.31: *Variance map from kriging estimations of both regions.*

The EasyKrig 3.0 Matlab Toolbox module has its own algorithm to calculate the error of the estimates, generating a variance map that allows to evaluate the accuracy in the prediction process. The maps of *Figure 9.31* were generated from this tool, in which we can see both regions, R1 (in the upper illustration) and R2 (in the lower illustration).

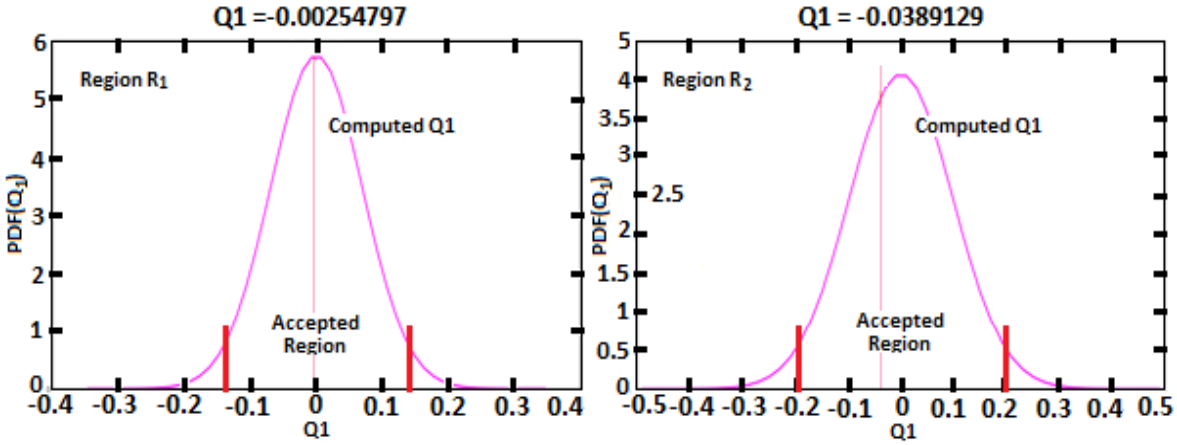


Figure 9.32: *Cross-validation results using* $\boldsymbol{Q_1}$ *method for both regions, R1(left) and R2 (right)*

In *Figure 9.31*, we can see that in the R1 region the prediction error is about 0.5 meters and for the R2 region, about 0.44 meters. If we compare this value with the prediction result considering the entire data set, shown in Figure 9.10, the error presented was less than 0.5 meters, however, we still must cross-validate and analyse the residuals for the second approach. In Figure 9.32, we present the cross-validation result for both regions of *Figure 9.30*.

Figure 9.32 shows the cross-validation result Q_1 for the two regions. This validation method confirms the good results of the estimation. The magenta line representing the calculated value of the Q_1indicator is within the accepted region as valid (among the red bars), indicating the suitability of the regression process for the dataset. The numeric result for each region is specified on each graph, being -0.00254797 for the R1 region and -0.0389129 for the R2 region.

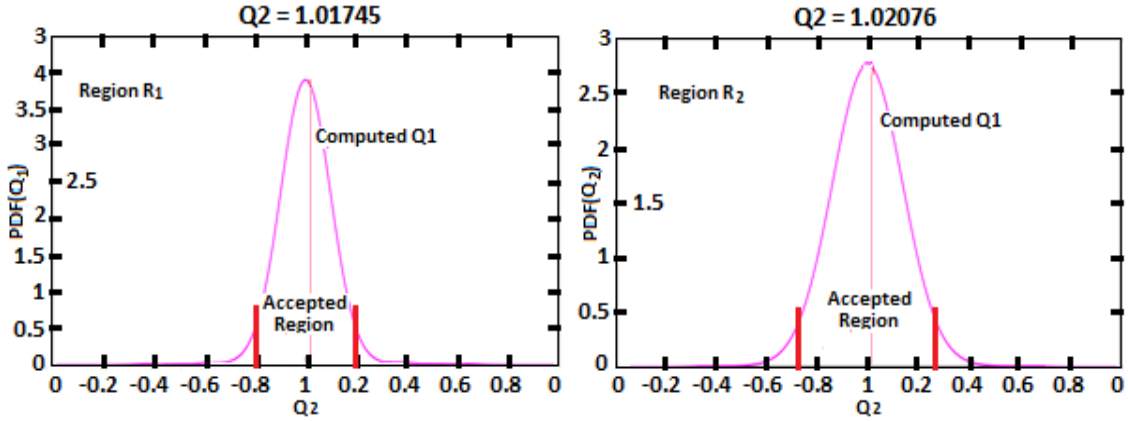


Figure 9.33: *Cross-validation results using* Q_2 *method for both regions, R1(left) and R2 (right).*

The cross-validation result using the Q_2 method is shown in *Figure 9.33*, both regions, R1 and R2, have similar values, the calculated value is approximately 1.0, this is an interesting result because according to the reference (Chu 2004) it is difficult to obtain good cross-validation results for both methods for the same estimation process. Despite this, it was decided to continue the cross-validation process using two other methods provided by EasyKrig 3, such as, the Leave One Out (LOO) and Double Kriging (DK) methods.

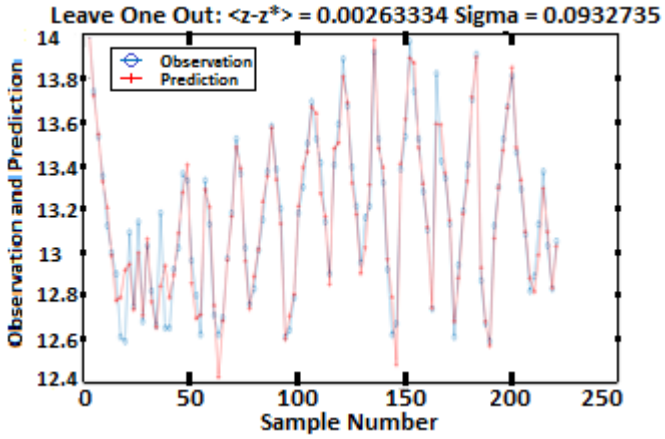


Figure 9.34:*Cross-validation results using LOO method for both regions.*

Considering the cross-validation method using LOO for both regions, the results are presented in *Figure 9.34*, we can see some underestimates occurring in situation of sudden changes values of regionalized variables, where they occur in local maxima and minima. Numerically, the value of the residue between the observed and predicted values was about 0.0026 with a variance of 0.09. The last cross-validation method applied was the DK and the graphical results are presented in *Figure 9.35*. The cross-validation DK presents the numerical result for the values of residue and variance equal to 0.0014 and 0.13 respectively.

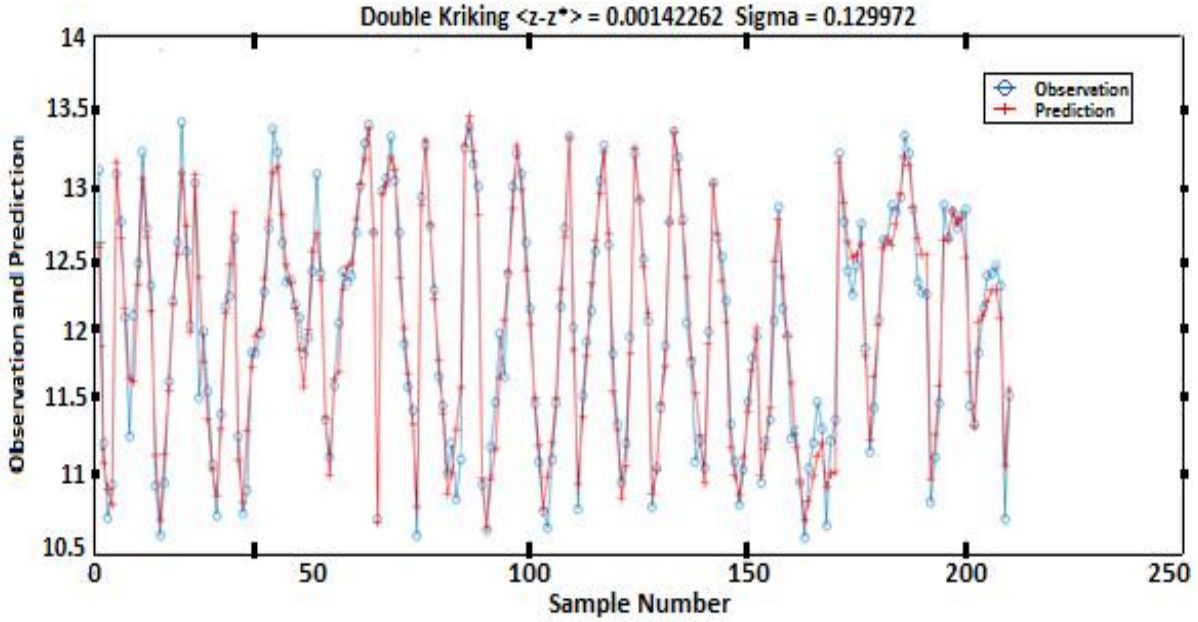


***Figure 9.35**: Cross-validation results using DK method for both regions.*

All cross-validation methods validate positively the predictions results; however, we proceed with the application of the evaluation method based on the comparison of the predicted values with real samples of an IS reference dataset. The advantage of using the IS method is the independence of the data between the estimated and observed samples, both from independent sources, but are taken from the same region. It guarantees independence in the evaluation of our estimates, calculating the residue between the observed data and the estimated data. *Figure 9.36* shows the variance map generated after the application of the IS method. We can see the clear difference in the variance values when compared to the application of the geostatistical method for the whole dataset without the clustering process, shown in *Figure 9.19*.

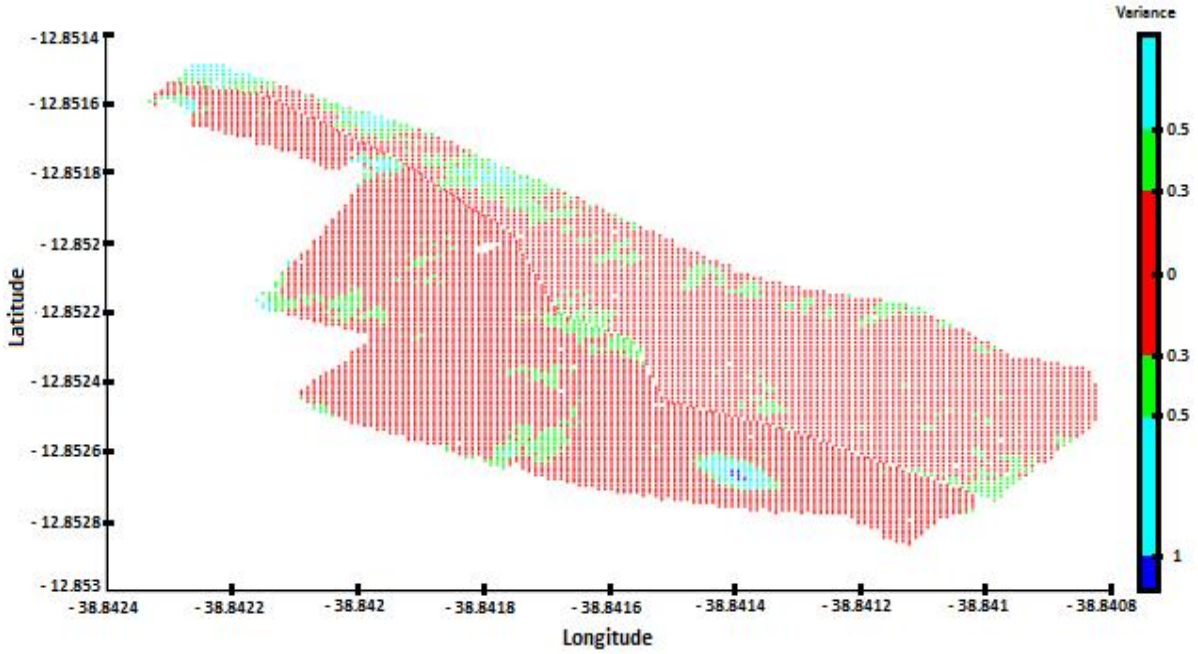


***Figure 9.36**: We observe the decrease of estimation error, and there are more regions with error bellow 0.3 meters ideal to consider into the specification to SO of IHO-S44 Standard.*

In *Figure 9.36*, we observed visually, through the colour indication, that the number of points whose estimation error is below 0.30 meters increased considerably in relation to the geostatistical prediction operation without the initial clustering process. The percent error was numerically described in *Table 9.16*. In *Figure 9.37* we graphically present the residual analysis on the residue distribution histogram graph, adding some information about the confidence level study.

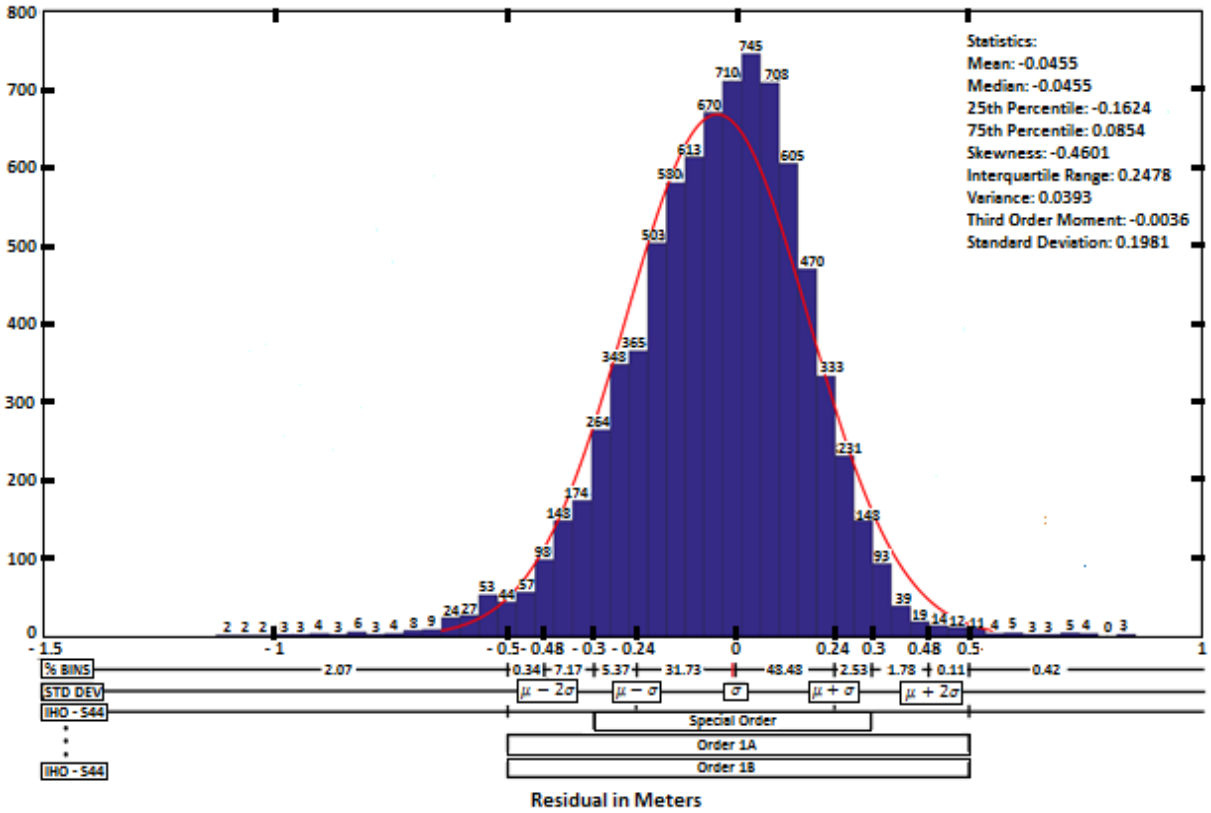


***Figure 9.3**: We observe the decrease of estimation error, and there are more regions with error bellow 0.3 meters ideal to consider into the specification to SO of IHO-S44 Standard.*

The processing approaches considered in this work are divided into: (i) Processing of the entire data set using the same model and parameter for the entire region; and (ii) to process the data set divided into clusters according to similarity assessments among samples and to apply different models and parameters, both processing was approached in order to meet the standard of the (IHO 2008).

The standard (IHO 2008) requires a minimum confidence level for measurements in bathymetric survey operations. Thus, depending on the level of confidence we can build bathymetric maps, or maps highlighting characteristics of interest for a sampled region.

These maps can serve as support to the navigation, being classified according to the criticality of the operation linked to the level of confidence achieved in the representation. In a brief review, we consider the reference values of parameters a and b of Table 1 of the reference (IHO 2008) as input of the expression of the calculation to determine the level of vertical uncertainty. Therefore, two equations were considered, equation 9.29 related to the confidence level for the Special Order (SO) and equation 9.30 for Order 1A and 1B.

$$CL = \pm\sqrt{(0.25)^2 + (0.0075 \times d)^2} \qquad \textit{eq. 9.29}$$

$$CL = \pm\sqrt{(0.5)^2 + (0.013 \times d)^2} \qquad \textit{eq. 9.30}$$

Within our sample set the maximum depth reached was $z_i = 14$ m, this will be our value d, represented in equations 9.29 and 9.30. Applying equation 9.29, for SO, our 95% confidence level is numerically about $\pm$ 0.27 m, this value was approximated to $\pm$ 0.3 m in our calculations and representations. Therefore, this value (CF_{SO}) was found to be:

$$CF_{SO} = \pm\sqrt{(0.06)^2 + (0.0075 \times 14)^2} = \pm\, 0.27\ m = \cong \pm\, 0.3\ m \qquad \textit{eq. 9.31}$$

Considering now the level of confidence to reach the orders of classification Order 1A and 1B, when we apply the equation 9.30, we will have as the value of CL approximately equal to ± 0.53 m, approximated here to ± 0.5 m, found of the following expression:

$$CF_{1A\ and\ 1B} = \pm\sqrt{(0.5)^2 + (0.013 \times 14)^2} = \pm\, 0.53\ m \cong \pm\, 0.5\ m \qquad \textit{eq. 9.31}$$

In the analysis of the comparative study between both paradigms (i) and (ii), we compared the global results using the same parameters for both, and thus, we can have a notion of precision involving both cases. Considering the first case, the kriging algorithm generates 8810 estimated sample values. From the total set of estimated points, we look for the points within the reference data set with a maximum distance of 0.5 m from an estimated point using the k-nearest neighbour algorithm, considering the Euclidean distance. This process resulted in 5117 reference points, so in the first case, we performed the analysis using these points. In the second case, the prediction process using the data grouped in clusters generated 8490 estimated values, similarly to the first case, points with a maximum distance of 0.5m within the set of reference samples were searched. In this case 8182 reference points were found with distance equal to or less than 0.5m.

The result of the kriging application is summarized in *Table 9.14*, we note that in the first case approach, the confidence level for the classification of Special Order estimates of 63.20% of the total, that is, the estimation error for 63.20% of the Estimates reached a maximum of 0.3 meters. In the situation where we analysed and processed the data after a previous cluster operation, the use of the estimates within the confidence level for the Special Order was 88.11%. Considering the estimated values for Orders 1A and 1B, for the first processing paradigm, the estimation reached a valid confidence level of 87.82% of the total, while for the data of the second case the confidence level reached was 97.51% of the total estimated points.

Table 9.2: *Results and comparisons in kriging estimations for both cases*

Kriging Results							
Description	**Data without Cluster (DWC)**			**Clustered Data (DC)**			**IHO - S44**
	Std. D.	**Samples**	**%**	**Std. D.**	**Samples**	**%**	
Number of Points into the Region	xxx	5,117	100%	xxx	8,182	100%	
Number of Points into one Standard Deviation	1 σ	3,234	63,20%	1 σ	6,563	80,21%	SO for Both
Number of Points with Difference Value (DV) <=0.3	xxx	3,234	63,20%	xxx	7,209	88.11%	SO for Both
Number of Points with Difference Value (DV) <=0.5	xxx	4,517	87,82%	xxx	7,978	97.51%	SO/1A/ 1B
Number of Points into two Standard Deviation	2 σ	4,917	96.09%	2 σ	4,917	97.06%	
Number of Points with Difference Value (DV) >0.5	xxx	607	11.87%	-	200	2,49%	Invalid

If we consider the standard confidence level for statistical estimates in which the 95% confidence level must be reached for occurrence within any Gaussian distribution a value of two standard deviations, that is, 2σ, the two approaches to the problem reach this requirement. However, according to the international standard (IHO 2008), only the second case of application of the kriging estimator will meet the level of confidence required for three valid Cases, namely Special Order, Order 1A and 1B.

9.4.4 Classical Metrics to Evaluate Results from Residual Analysis

A classical estimation method was applied using the Root Mean Square (RMSE) and Mean Absolute Error (MAE) indicators, which are based on residuals. This allowed a second evaluation of the solution used, the calculated values are described in Table 9.15.

Table 9.3: *Numerical results for the global average magnitude of the error metrics*

Global Average Magnitude of the Errors Analysis				
	RMSE	**RMSE (%)**	**MAE**	**MAE (%)**
Estimation without cluster	0.3200	2.21	0.2500	1.72
Estimation with cluster	0.2033	1.45	0.1541	1.10
Accuracy increment	0.117	65.00	0.0959	61,64

In Table 9.15 it can be observed that the best estimate was reached in the methodology when the clustering was applied. In this sense, the evaluation should take into account that the best estimation must present the lowest values for RMSE and MAE. In this case, we obtained a great improvement in the precision in the estimations for the regions divided by the clustering process, fact that confirms the comparative method of measurement of the precision used in this work.

9.5 Conclusion

In this section, we proposed the main objective related to the use of the geostatistical linear regression method, called Kriging (Cressie 1993)(Matheron 1971a), on a dataset from bathymetric survey that extracted discrete samples of the submerged surface depth value within a delimited region using an SBES device. The bathymetry samples were extracted in a sampling operation along a transect path, collecting samples at points with almost regular spacing of approximately 5 (five) meters. The main objective is verifying the viability of a Kriging predictor in the task of increasing the resolution of the bathymetric survey, estimating depth values in non-sampled geographic points, thus considering the discrete depth samples as a regionalized variable under the theory of estimation and geostatistical prediction.

The resolution of the trend problem in the data, which results in instability that decreases the confidence level of the predictions, has been solved by applying a data-grouping process that aggregates a set of samples with similar statistical characteristics into disjoint groupings prior to the application of the Analysis and geostatistical prediction. The evaluation of the quality and confidence level of the prediction process was performed using cross-validation methods provided in the Toolbox used in the work, EasyKrig 3.0 (Chu 2004), in addition to a method that compares the predicted values with a second data set with a much higher resolution extracted from the same working region. In the international standard for hydrographic surveys published in the document (IHO 2008), the value of the confidence level of bathymetric information in the construction of aid maps to coastal navigation operations should be at least 95%, in our study the application of the predictor Kriging together with the data grouping method, a prediction confidence level of approximately 97% was achieved.

10.0 Overall Conclusion

All sections described in this thesis has one focus, remote sensing applications. Although remote sensing covers a very wide field, remote sensing techniques in underwater medium often is required to acquire data in some places that cannot be directly collected by humans. This will improve advances in data extractions to mapping the underwater landscape and any structures immersed in this medium.

In this thesis, we pay attention in the main definition of remote sensing as the measure some information about an object or phenomenon under study without making physical contact. Into this context, we focus in underwater remote sensing applying sonars and their sensing elements. We probe the improvement in data information collected and the process gain using spatial statistics to predict a dense bathymetric data from few numbers of bathymetric measured points and stochastic resonance approach to increase the sensitivity of sensing elements and to decrease the noise level present in acoustic images.

The key aspect related to research in sensing in this thesis is the monitoring applied to crossings between attractors activated only by an optimal level of noise added to a target signal. Thus, we improve the sensibility and sensing velocity to detect weak dc signal using only a simple approach of the proportion of time the sensor stays in one or another state.

The focus in acoustic image processing was pointed to image contrast modification using stochastic resonance, it means chaotic noise generated by the logistic map and for Gaussian to conduct an image enhancement in their digital grayscale levels non-dynamical system of threshold elements and the noise to digital grayscale sonar image. This will result in a better acoustical image to use a visual or automatic feature extraction computer algorithm.

Spatial statistics improves feature extraction and underwater landscape to map generation using data points generated from the geostatistics regression tool called Kriging from some samples points extracted with a Single Beam Echo-sounder. It means, we increase the resolution of an underwater survey with high confidence level.

The set of study and solutions developed and present in this thesis will help the underwater acoustical survey area. It improves the data extraction from underwater landscape and can be used to mineral resources as petroleum and gas; telecommunications and energy routing and distribution in underwater medium and environmental management.

Appendices

A - Sonars and Devices to Support Data Acquired

A1 Sidescan and Single Beam Sonar

The work presented here starts from a bathymetric survey using an acoustical transducer to acquire a underwater depth in many area from Todos os Santos Bay (TSB), and their estuaries. The TSB is localized near the Salvador city, in Bahia, Brazil. Each data consists in a discrete depth point information taken just below the survey vehicle or boat. We use three types of sonar devices, a Starfish Sidescan Sonar, Teledyne Echotrack CV100 Single Beam Sonar and Multibeam R2 Sonic 2024.

B - Surveys and Dataset Acquired

The data format to data extracted using SBES was stored in a XYZ file format, where three columns of data represent in our case, longitude, latitude and a depth value. This stored data format is very closely related to geo-referenced raster and grids representation, but some considerations need to be done to represent a XYZ data file into a raster or grid data space, but in our application, we work in the first approach with a vector format of map representation, the so-called contour map to do a cluster analysis. We can consider that each sampled point value will represent a unit portion of the entire area, each portion will be called grid cell, or simply consider each sampled point as a discrete measure spatially distributed. A point measurements can be allocated at the vertices of a grid cell or in the interior of it (centre of observation points). Here we consider each sample will represent a value centred on observation points cell value according guidance from (McBratney & Webster 1981).

B1 Survey Mission Region 1: Palma Bay – Mallorca – Spain

The surveys works was realized in the Palma Bay and was performed using two small cove beach, Cala Estància and Cala Gamba, the first survey was used to extract navigational data to probe the control and navigations modules. The second mission were performing to extract acoustic image data extract using a Sidescan Sonar and a type of kayak boat survey.

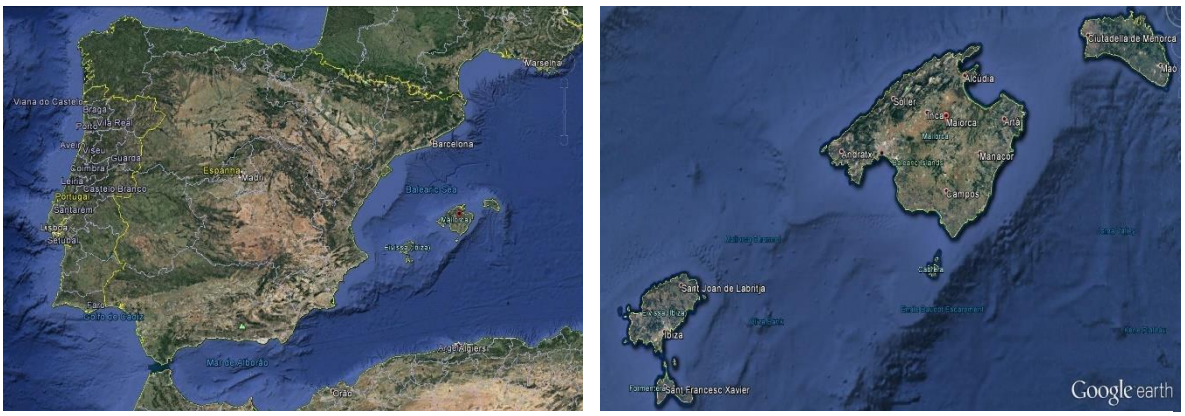


Figure B.1: *localization/situation map of Iberian Peninsula (left) and the Balearic Island archipelago (right), both extracted from Google Earths.*

The Palma Bay, is a biggest bay from Mallorca island, and it is show in the *Figure B.1*, in this figure Mallorca is the biggest island about the centre of the Balearic Archipelago, and this archipelago is one of the autonomous community of Spain. The environmental and geographical conditions allies to academic knowledge of marine data extraction/processing and technological development is an attractive to joint to UIB research groups to develop works in the marine technologic and environmental data processing area, and the first and last stage of this work development reflect these assumptions.

B1.1 Cala Estància – Palma

This mission survey was performed in three days to extract navigational data to probe the control system, communications and instrumentation modules and mechanical design of the underwater robotic vehicle. The control and navigation model used was designed with an inspiration in the work (Acosta et al. 2005) and (Calvo et al. 2005). While the mechanical design was inspired by the work developed in the Marie Curie - European Project - MIF2-CT-2004-003027 from Prof. Dr. Gerardo G. Acosta (Acosta 2005). In the *Figure B.2* (left), we can see the localization/situation map of Cala Estància beach, a small cove beach into the Palma bay that we extract navigational data and probe navigation and control algorithm using the Low-cost underwater vehicle developed. The *Figure B.2* (right), is the picture of the underwater vehicle developed taken at this beach.

The coloured area selected into the four-vertex region was the selected as a best testbed according the shallow mean depth and the calm water characteristics, details about the vertex points are expressed in the *Table B.1*, and the total worked area covered and limited by the vertex points are about 6000 m^2.

Table B.1: *Geo-located Vertex points of Cala Estància survey*

Point	Latitude	Longitude
Vertex 1	39°32'4.39"N	2°42'44.31"L
Vertex 2	39°32'3.48"N	2°42'48.73"L
Vertex 3	39°32'2.13"N	2°42'47.23"L
Vertex 4	39°32'2.08"N	2°42'44.11"L

Figure B.2: *localization/situation map of Cala Estància beach (left) extracted from Google Earths and a view of an Autonomous Underwater Vehicle developed(right).*

Two published work was generated using the navigational data extracted in this survey mission and here I can cite (O. Calvo et al. 2009) and (Calvo et al. 2008). These works relate experiments and results of control of two methods for governing an AUV to perform lawnmower searches of pipelines and cables based on Lyapunov and PI controller to achieve better performance, passing through a series of waypoints with guaranteed convergence.

B1.2 Cala Gamba – Palma

The Cala Gamba mission survey was performed using a small boat like a kayak with small adaptations including an electric propulsion and manned control. The main objective in this mission was try to extract acoustic images using a Starfish Sidescan Sonar in a shallow water region that we can travel in a long straight line trajectory, some images were taken with some distortions and we cannot extract a good set of GPS data. In the *Figure B.3*, we can see the situation map of Cala Gamba beach, like the Cala Estància this is a small cove beach into the Palma bay too.

Figure B.3: *localization/situation map of Cala Gamba beach extracted from Google Earths.*

The problem related to the Sidescan image distortions was mainly provoked by the inefficient survey vehicle velocity from 0.3 to 0.5 m/s, two unites of sealed mercury marine trolling motor from the set of Torpedo2000 Diver DPV devices (Torpedo 2016) were used attached in the kayak boat hull as propeller, one in each side.

Despite the motor trolling propulsion technical specification published at (Torpedo 2016) relates a speed about 0.88 m/s for each trolling motor, the sea conditions, drag from the kayak boat hull and the payload composing by a sonar device and operator decrease the propulsion efficiency and the velocity using two motors was about 0.4 m/s with variations about 1 m/s. According the Starfish Sidescan user guide manual (Blueprint 2010), the velocity to acquire acoustic imagine need to be constant and between 0.51 m/s and 2 m/s. In other way, the GPS receptor based on sirf-star iii chipset module from (D. E. L. Blueprint 2016), only receive the signal from 5 to 7 satellites, and it can result in an inaccurate navigational data. In the *Figure B.4* we can see a set of 6 acoustics images from this site survey, without any special features or object detected but a lot of distortions can be viewed mainly because the inefficient survey boat velocity and perturbation from environmental conditions heave, roll and yaw movements.

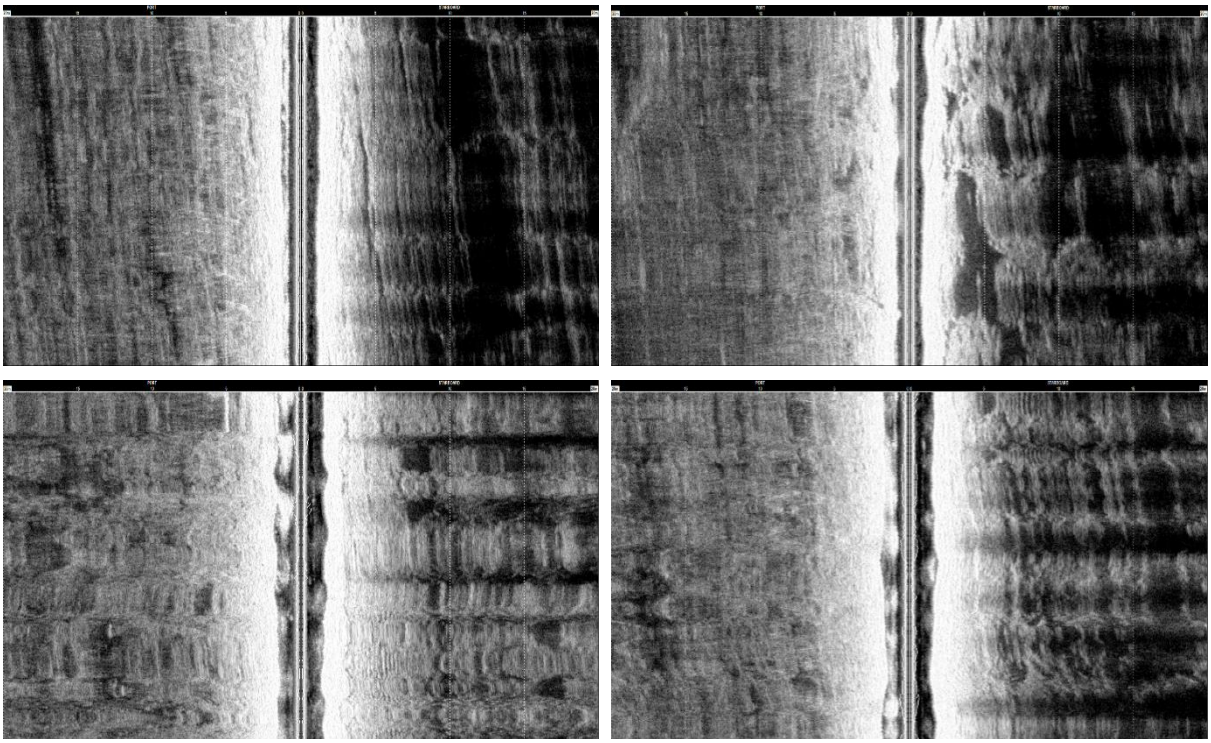


Figure B.4: *Small set of acoustic images extracted with a Starfish Sidescan Sonar in the Cala Gamba survey mission, this acostic dataset was not processed because some acquisition distortions in the data.*

In the *Table B.2*, is described the vertex points that limits the polygonal survey area, shown in the *Figure B.3*, this area is delimited by a polygon with 4 vertexes. The total worked area covered and limited by the vertex points are about 32000 m^2.

Table B.2: *Geo-located Vertex points of Cala Gamba survey*

Point	Latitude	Longitude
Vertex 1	39°32'47.79"N	2°42'44.31"L
Vertex 2	39°32'44.55"N	2°41'52.31"L
Vertex 3	39°32'42.94"N	2°41'40.42"L
Vertex 4	39°32'39.76"N	2°41'43.82"L

The dataset generated by this survey work cannot be used to do studies mainly because distortions in the acoustic images, but given the knowledge base to planning and develop a survey mission to acquire sonar and navigational data using a boat.

B2 Survey Mission Region 2: Todos os Santos Bay – Salvador – Brazil

The survey process was performed in four stages using a boat survey, three in collaboration with "Belov Engenharia" (www.belov.com.br) (Belov 2015), a local private company specialized in port engineering and

hydrographic data acquire and processing with focus in private and public companies from brazil and around the world. The boat navigation tracking method used follow a zig-zag path to cover an entire region in topographic underwater relief data extraction and in linear fashion to survey pipelines or pier structures. Todos os Santos Bay is the second biggest by in Brazil and is located in Bahia. This bay has a fundamental importance to the concave coastline regions around to Port operations, related to commercial and people transport, connecting some cities and regions through the bay. Furthermore, the military interest is evident because it has some military units of Brazilian army, one of the biggest Brazilian petroleum refinery and a complex habitat interaction with some others internal bays, small islands and delta rivers generating mangroves habitats.

In the *Figure B.5* we can see the localization/situation map of Bahia state, localized in the southwest of Brasil (left), and the localization/situation map of Todos os Santos Bay localized in Bahia state (right), we can see a diversity of environmental interactions among small and big islands and distinct water bodies from delta rivers, interior bay.

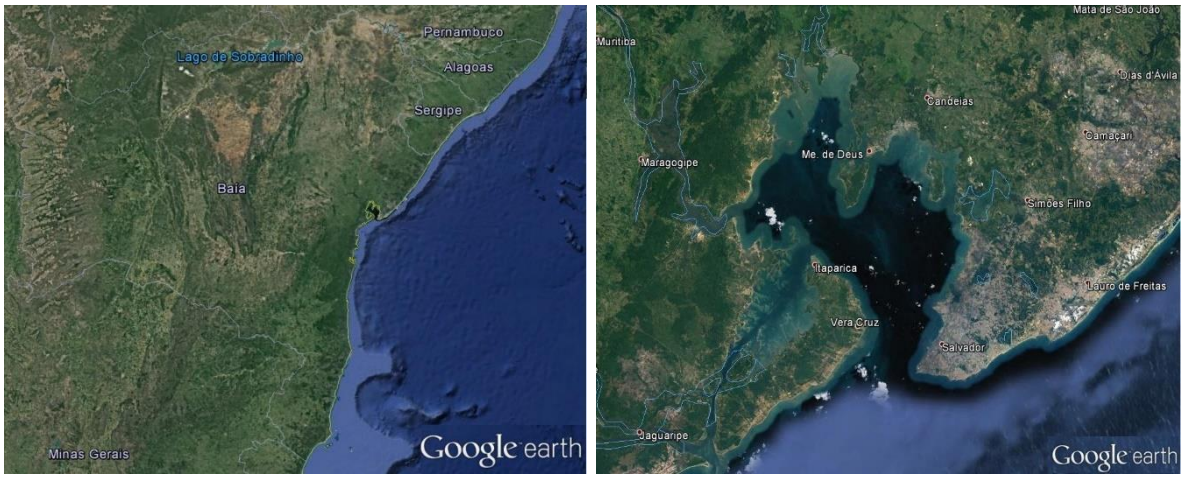


Figure B.5: *localization/situation map of Bahia in the Southwest of Brasil (left) and the Todos os Santos Bay (right), both extracted from Google Earths.*

The mission survey to acquire data was designed to cover basics task into the concept of hydrographic and underwater structure search survey, both very important to navigation and port security. The survey objective was extract a gas pipeline acoustic images using a sidescan, the others three was to extract the navigation channel data using a sidescan and single beam echo sounder in the last case from three navigations channel: Madre de Deus, Aratú Bay and Paraguaçu River.

B2.1 – Pipeline Track – Gas Pipeline Itaparica-Aratú

The gas pipeline Itaparica-Aratú is a ridged carbon steel pipeline with 33 km of length and 8 inches' diameter, it connects the continental area (Aratú Port) and the Itaparica island, the pipeline has capability to transport gas with a rate about 100 m^3/h and is laying in the sea bottom around 20 meters' depth. The main objective in this survey mission is collect sidescan data using a boat survey with sonar adapter designed to fix the sonar sidescan device to a boat hull, and with this acquire the acoustic data to probe the image quality to track a target far away from 10 meters' depth and prove the limitations to get data to track small pipeline targets at this condition. The *Figure B.6* show the context map of the pipeline track survey region into the Todos os Santos Bay.

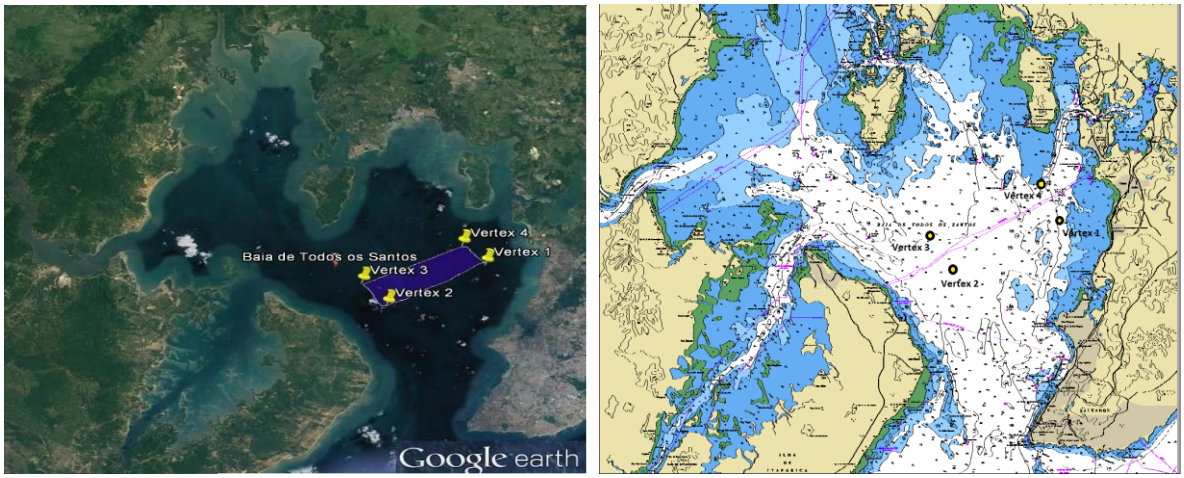


Figure B.6: *Survey area selected into the Todos os Santos bay limited by a set of vertex points (left), source: Google Earths. Nautical chart from the same region to illustrate the gas pipeline (blue line between vertex point) into the bay, connecting Aratú port (Continental region) and Itaparica island.*

In this survey, the main difficult were track the pipeline with a manned boat survey without an automatic navigation close loop control, in some time we lost the pipeline according the boat had changing their heading because tidal currents and wave interactions in the hull boat. The total area choosing to do a pipeline search and survey is limited to four vertex points, and represent a polygonal area with 8.9 km x 2.6 km in its side, representing 32 km^2, with vertex coordinates presented in *Table B.3*.

Table B.3: *Geo-located Vertex points of Todos os Santos bay to pipeline track survey*

Point	Latitude	Longitude
Vertex 1	12°51'9.02"S	38°31'14.73"O
Vertex 2	12°53'29.90"S	38°35'40.24"O
Vertex 3	12°52'11.71"S	38°36'47.28"O
Vertex 4	12°50'3.26"S	38°32'17.31"O

To get the sidescan data from pipeline track we use a regular boat presented in the figure 36 with a special structure installed in his hull steam line to support the sidescan device. The boat named Tonhozé is 16 feet long with a single central motor, and the sidescan adapter support was made using a stainless steel, the position at the hull steam line is to guarantee the most solid fixation at the boat hull when the boat navigates forward and to maintaining the sonar device completely into the water because the pipeline track was made orthogonally to the predominant sea current, causing a constant boat roll movement.

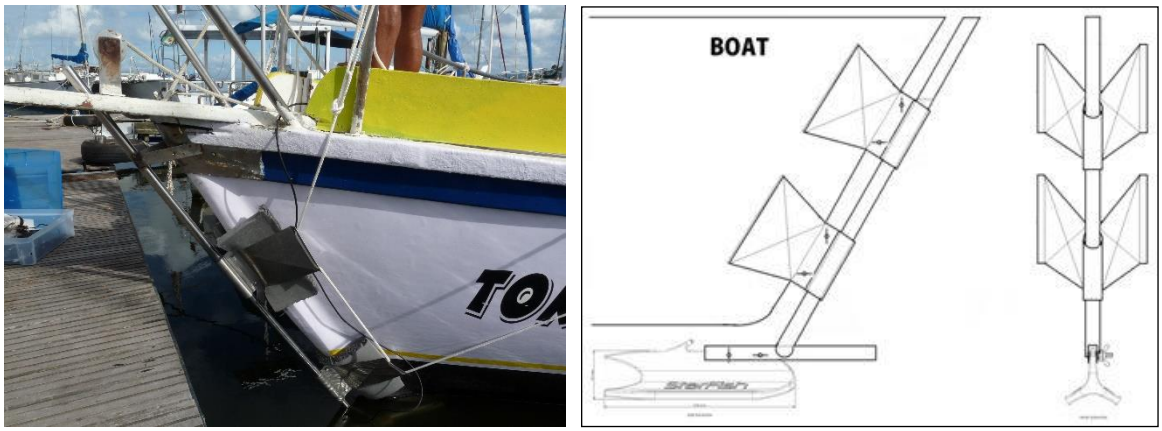


Figure B.7: *Survey area selected into the Todos os Santos bay limited by a set of vertex points (left), source: Google Earths. Nautical chart from the same region to illustrate the gas pipeline (blue line between vertex point) into the bay, connecting Aratú port (Continental region) and Itaparica island.*

In the *Figure B.8* we can see some acoustic image samples acquired in this pipeline track survey some of these were used in the study to develop these three works (Sousa et al. 2016) , (Villar et al. 2014) and (Villar et al. 2013).

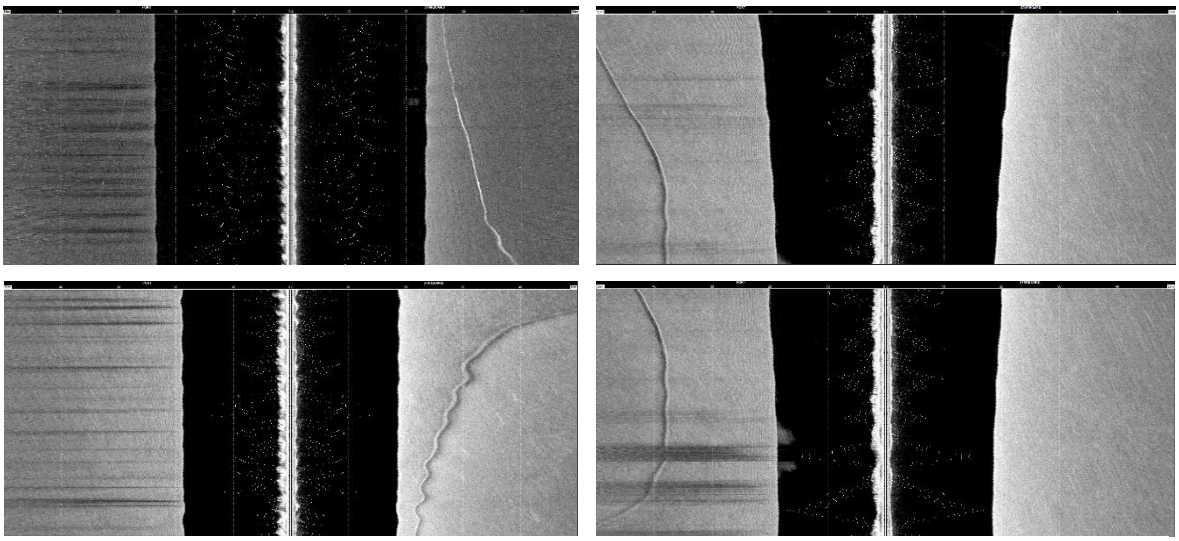


***Figure B.8**: Pipeline track sample images*

In *Figure B.8*, we can see some features into the acoustic images and two of this are shading and pipeline distortion, in the first case, shading is a result of sea stirring and tide current from navigation channel, transverse to the pipeline track trajectory shocking in the boat hull side, causing a boat roll movement, take out from water one of the sensor side, despite the care in positioning the sensor in a boat hull position to minimize this disturbance. The second feature, pipeline distortion, is caused by the same problem of sea stirring and tide currents, but in this case, affect the navigation of boat causing oscillations in heading direction, yaw movement. The effect of yaw movement into images on the *Figure B.7*, can be viewed as a sinuous pipeline shape. In several moments in the track survey operations we lost the pipeline track because the yaw disturbance and to continue the survey, we perform the zig-zag search to find and track the pipeline again.

B2.2 – Navigation Channel Track – Salvador – Madre de Deus

The navigation channel Salvador-Madre de Deus is main way to oil products transport and logistic of Bahia and in on of this extreme has the second biggest petroleum maritime terminal.

***Figure B.9:** Survey area selected into the Todos os Santos bay limited by a set of vertex points (left), source: Google Earths.*

This is an artificial dragged channel and starts at the latitude and longitude position (12º 49’ 2” S, 038º 34” 0’ W), nearest to Salvador cost and ending at the Madre de Deus Terminal (12º 45’12” S, 038º 37’ 27” W). The objective in this survey is acquire some data using a sidescan sonar to identify some physics features from the navigation

channel as dimensions and structural components, as we pass beside the Madre de Deus Maritime Terminal we try to acquire some data from pier structure. The *Figure B.9* show the context map of the navigation channel into the Todos os Santos Bay.

The total area choosing to do a navigation channel structure search and survey is limited to four vertex points, and represent a polygonal area, representing a total area around 2.6 km^2, with vertex coordinates presented in Table B.4.

Table B.4: *Geo-located Vertex points of Todos os Santos bay to Salvador-Madre de Deus, navigation channel survey*

Point	Latitude	Longitude
Vertex 1	12°46'10.04"S	38°36'2.74"O
Vertex 2	12°46'0.90"S	38°35'38.95"O
Vertex 3	12°45'12.66"S	38°37'42.17"O
Vertex 4	12°45'3.27"S	38°37'42.44"O

Some difficult were found to perform this survey task, first of all is the navigation channel dimensions and a not sharp change from the low to higher depth, in this survey only sidescan data was extracted. During the survey, I cannot found any feature or relevant structures from the navigation channel formation, probably because the underwater seabed is very smoothed formed majority by loamy and sand soil type. Because the before explained, we decide to start a port structure survey and acquire some data from navigation buoy and pier pillars around the Madre de Deus Terminal. In the *Figure B.9* we can see some acoustic image extracted in this mission survey.

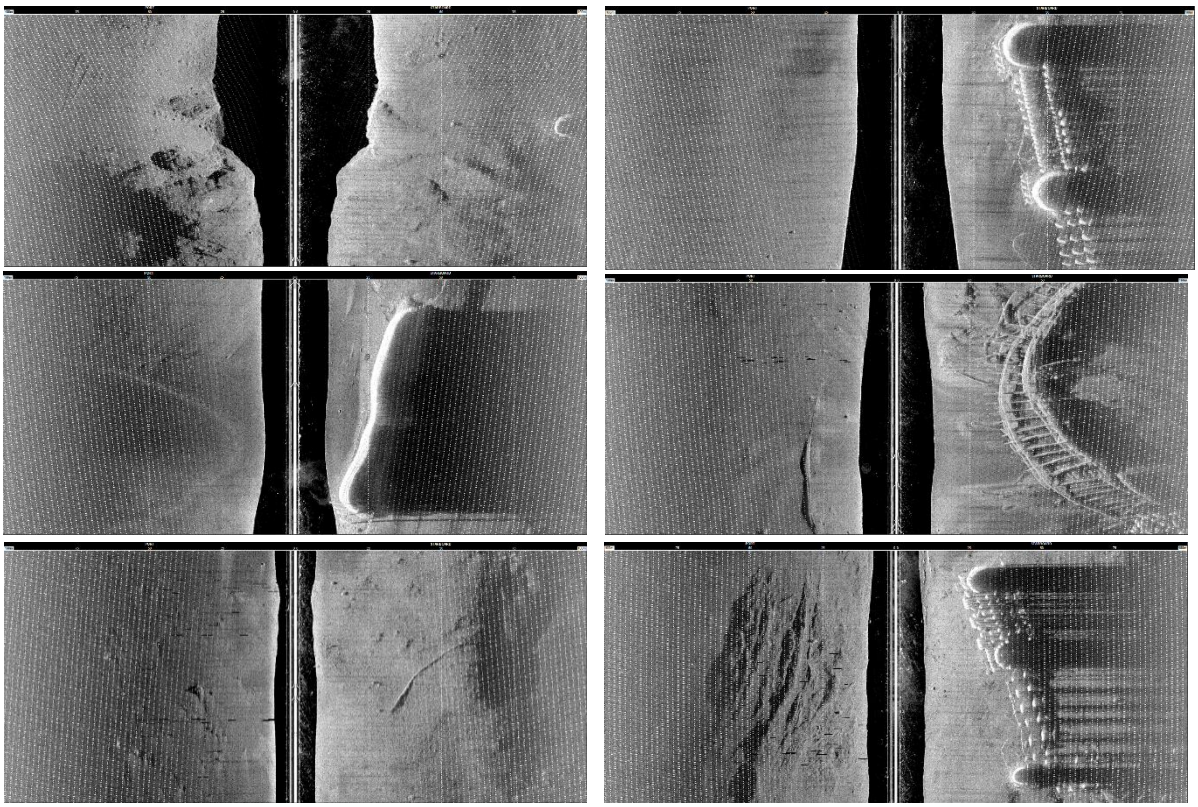


Figure B.10*: Sample images from navigation channel Salvador-Madre de Deus, in these images are show actual pillars pier structure, some debris and an old pier structure which was flooded after an accident with an oil cargo tanker vessel few decades before.*

In the *Figure B.10*, we can observe some structures and debris found lying in the navigation channel seabed, as port pillars, a sunken small boat and some debris, this data were user to probe the sidescan sonar system and GPS signal integration, but we cannot apply it in any published work.

B2.3 – Navigation Channel and Port structure of Aratú – Aratú Bay

The second navigation channel survey after the first experience was performed in the Aratú Bay, an interior bay with access through the Todos os Santos Bay, that start the Aratú Port Navigation Channel. This channel conduct mainly chemical and solid/dry freight using cargo ship. The channel starts at the latitude and longitude position (12º 50' 2" S, 038º 31' 4" W), and ending at the Aratú Terminal (12° 47' 71" S, 38° 28' 48"O). The objective in this survey is try to acquire some data using a sidescan sonar to identify some physics features from the navigation channel as dimensions and structural components and if possible, extract some data from pier/port structure or some debris laying in the seabed. The *Figure B.11* show the context map of the navigation channel into the interior Aratú Bay.

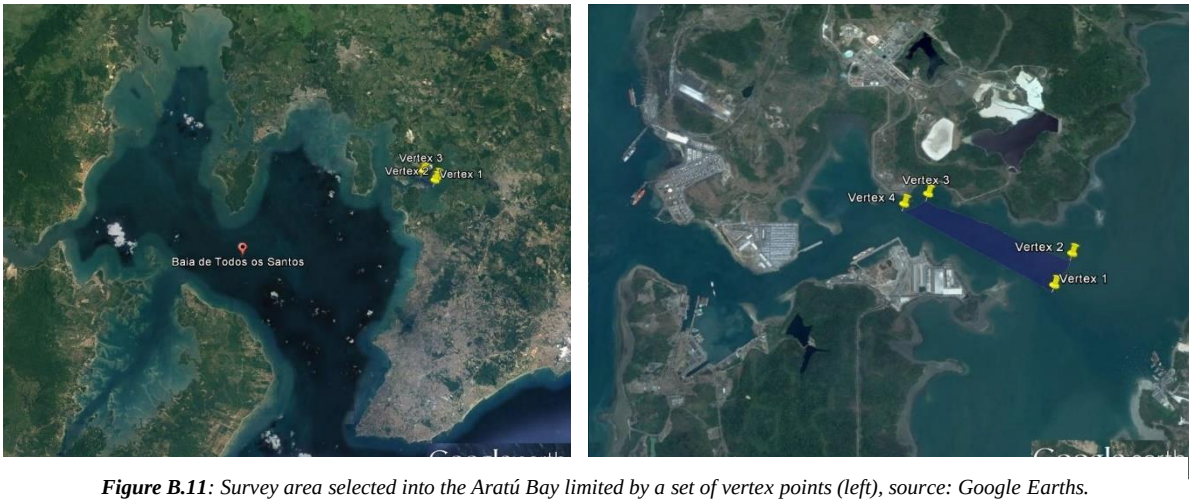


Figure B.11*: Survey area selected into the Aratú Bay limited by a set of vertex points (left), source: Google Earths.*

The total area choosing to do a navigation channel and structure search and survey is limited to four vertex points, and represent a polygonal area, representing a total area around 0.6 km^2, with vertex coordinates presented in *Table B.5*.

Table B.5: *Geo-located Vertex points of Todos os Santos bay to Salvador-Madre de Deus, navigation channel survey*

Point	Latitude	Longitude
Vertex 1	12°47'19.53"S	38°27'55.34"O
Vertex 2	12°47'19.53"S	38°27'55.34"O
Vertex 3	12°47'19.53"S	38°27'55.34"O
Vertex 4	12°47'19.53"S	38°27'55.34"O

This survey was performed with support of Hydrography Department of Belov Engineering LTDA, and data was acquired with a Starfish Sidescan Sonar and the Starfish GPS. The most important feature detected was a pier

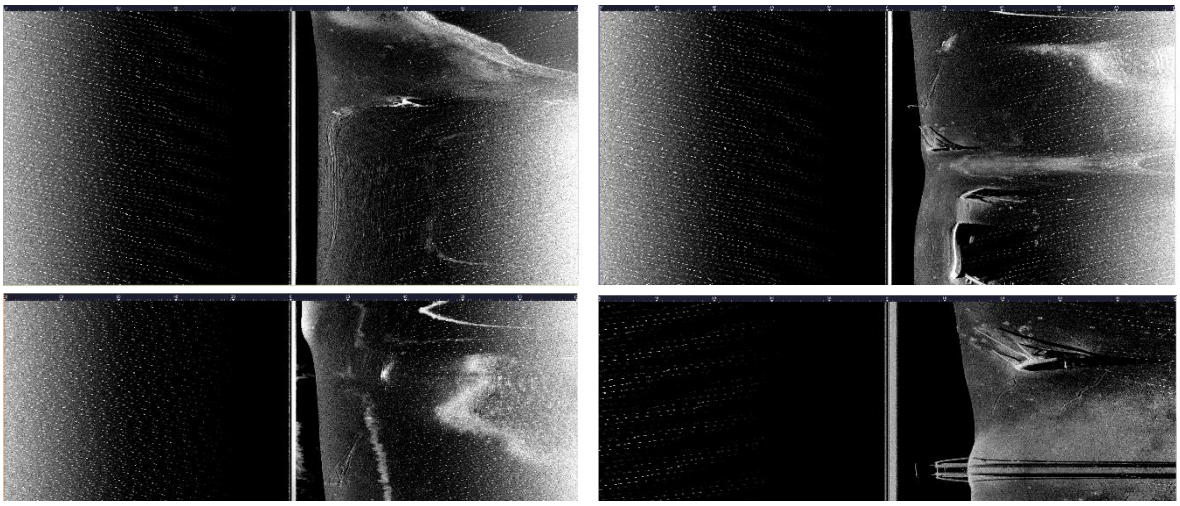


Figure B.12: *Sample images from Aratú navigation channel, in these images are show some pillars pier structure, and some old pier dolphin structure which was flooded and lying on the seafloor.*

structure flooded into the water, laying in seabed and some pillars from the pier structure. In the *Figure B.12* we can see some acoustic image extracted in this mission survey.

In the bottom image at left side, we can identify a pier structure called mooring dolphin flooded lying in the seafloor and some grooves in the seafloor sediments. Moring Dolphin is a man-made structure construct beside a pier and out of shore to provide a fixed point facility to keep a ship stable and fixed at the berth. The image context can be an indication that it could have been dragged by a boat and their pillar made the furrows in the seafloor before it sank completely.

B2.4 – Paraguaçu River – São Roque do Paraguaçu

The third survey was performed in the Paraguaçu River, near to the Paraguaçu Delta into the Todos os Santos Bay, this mission was made by Belov complete team and devices and is the best set of datasets because has the complete information to do some research study. The *Figure B.13* show the context map from Paraguaçu river region and the region that the data was acquired.

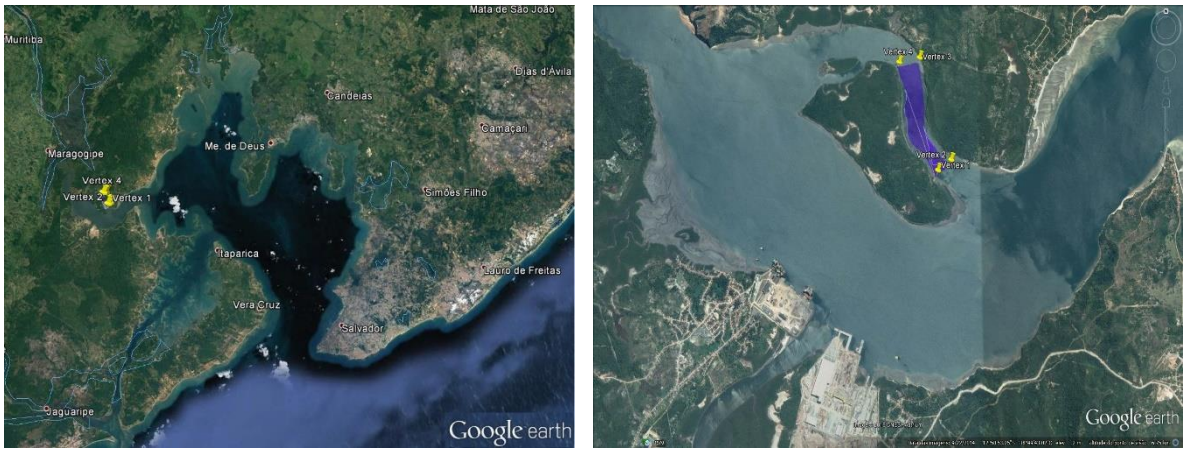


Figure B.13: *Context map from Paraguaçu river, in the left, the location of Paraguaçu river into the Todos os Santos Bay and at right the survey region into Paraguaçu river delimited by fou Vertex points.*

The vertex points that limit the survey region are expressed in the *Table B.6*. The total area surveyed is around 0.22 km^2, and the area into the vertex points is around 0.32 km^2.

Table B.6: *Geo-located Vertex points of Paraguaçu river that the data was acquired*

Point	Latitude	Longitude
Vertex 1	12°50'34.17"S	38°49'21.48"O
Vertex 2	12°50'30.13"S	38°49'15.52"O
Vertex 3	12°49'50.11"S	38°49'29.32"O
Vertex 4	12°49'51.89"S	38°49'38.63"O

This survey is being acquired the best set of sonar data and was the main dataset used to perform the work based in geostatistics approaches. Additionally, in future works these datasets will be used to develop studies in data fusion to increase the capability of feature extraction and digital elevation model construction (DEM) using simultaneously Single Beam Echo Sounder (SBES), Multi Beam Echo Sounder (MBES) and Sidescan Sonar.

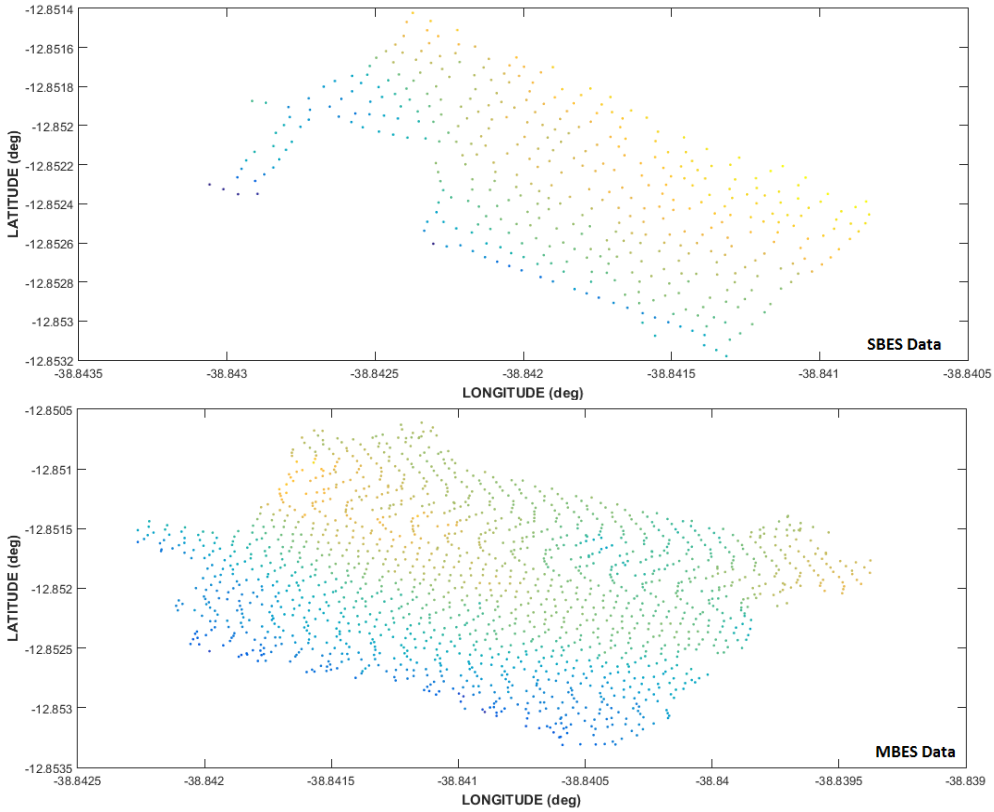


Figure B.14: *SBES and MBES datasets acquired from the Paraguaçu river survey, source from Belov Engenharia Hydrographic Department.*

The *Figure B.14* show a scatter plot from the SBES with 442 data points, and MBES with 51,131 data points. This two datasets were used in the geostatistics work developed in this thesis, and in the *Figure B.15* we have the Sidescan data from the same region. The complete dataset from three sensors give the opportunity to compare the data from each sensor in order to evaluate the information among these sensors and how we can increase the region information using data fusion in future works.

Figure B.15: *Sample of Sidescan image acquired in the Paraguaçu river survey.*

The complete dataset from three sensors give the opportunity to compare the data from each sensor in order to evaluate the information among these sensors and how we can increase the region information using data fusion in future works.

C IOGECS Project – Sketch of Preliminary research work an project that support this thesis

A Low-Cost Experimental Platform to Perform Underwater Remote Sensing

In this part of the thesis I described efforts to develop a Low-Cost Underwater Remote Sensing Platform to acquire some data from underwater relief or structures to support underwater operations as mineral resources exploitations. The human history has been driven by the need to overcome the man limits to adapt in the most adverse situations. Almost always this results in a technologic knowledge evolution. Often, the men felt the need to discovery new frontiers, beyond the limits imposed by the physical, chemical and biological. The human limitation to interact with aggressive environments characteristics is the start-up point to develop solutions to ovecome it, based in learning with earlier works to apply in new developments, this is the normal way to lead the science evolution to increase the human knowledge about their body and mind, technology, nature and interacions into the earth's enviroment. The global aim of this work is involving the study and development of hardware (robotic platform) and software (algorithm) to effort a knowledge based on survey the underwater environment based on data extraction of underwater topographic relief using acoustics sensor (sonars). To perform the work of data conditioning and signal processing new approaches using stochastics resonance and geostatistics has been used. On this part of the thesis I will describe some aspects about vehicle development, the hardware and software architecture concept are described,

2.1 Generic Architecture Guideline

The prototype design was based in a generic architecture proposed in some research works with some simplifications. A modular philosophy was used in the design and development of the small AUV, re-using many modules previously developed and tested in other applications. The dynamic Mission Planner (DMP) (Acosta et al. 2005) and the Obstacle Avoidance Software (OAS) are in cascade. Thus, if the OAS does not detect any object through the forward-looking sonar, its output will be simply the desired trajectory from the DMP. On the contrary, if an obstacle is detected, the OAS changes the necessary waypoints in the trajectory provided by the DMP.

The software architecture, running on a Debian Linux CPU, is based on communications messages between modules implemented using the UDP layer. The system is divided in modules specialized in a particular task (O Calvo et al. 2009). We have four main structures to do fundamentals vehicle tasks: Navigation Module, Dynamic Mission Planner (DMP), Sensor Fusion and Guidance & Control Systems. According to figure 1 we can see that the Navigation Module gets the vehicle position information from GPS and Inertial Sensors to navigate in surface. When the vehicle is submerged, the underwater navigation is need to be performed, another type of sensor is necessary. Then, a navigation Doppler system called Doppler Velocity Logs (DVL), is used. The DVL is a sonar based device that measures the vehicle velocity with respect to the sea bottom by taking advantage of the Doppler effect (Acosta et al. 2009). In our case, we have not a DVL sensor to support the navigation data enhancement so we apply a GPS receiver.

The Sensor Fusion takes signals from the sensors responsible to recognize and track underwater structures. Among these sensors we can find sonar, magnetic trackers and optical sensors like cameras or scanners lasers. This module is very important to obtain a high quality mission, due to the acquired navigation information must be sent to the DMP, to confirm the correct track path, any deviation or inconsistency between the expected environmental or target characteristic. The DMP module takes information from the environment (self-position and target position), the exclusion zones in the mission data bank, and the static mission plan, provided by the user. The DMP decides the trajectory to follow, this trajectory is verified against the exclusion zones and probable obstacles detected by the Obstacle Avoidance Sonar (OAS) and modified accordingly before is sent to the path planner.

The Guidance & Control Systems module provide the settings for propellers and rudders to reach the desired trajectory or a specific waypoint. This module yields its output taking into account the Dynamical Mission Planning module output and the broadcast information from the Navigation module. The path planner also receives information from a Static Mission planner, that could take control at any time modifying the final trajectory (i.e. safety reasons, beginning of mission, abort, etc)(O Calvo et al. 2009)(Acosta et al. 2003)(Acosta et al. 2009).

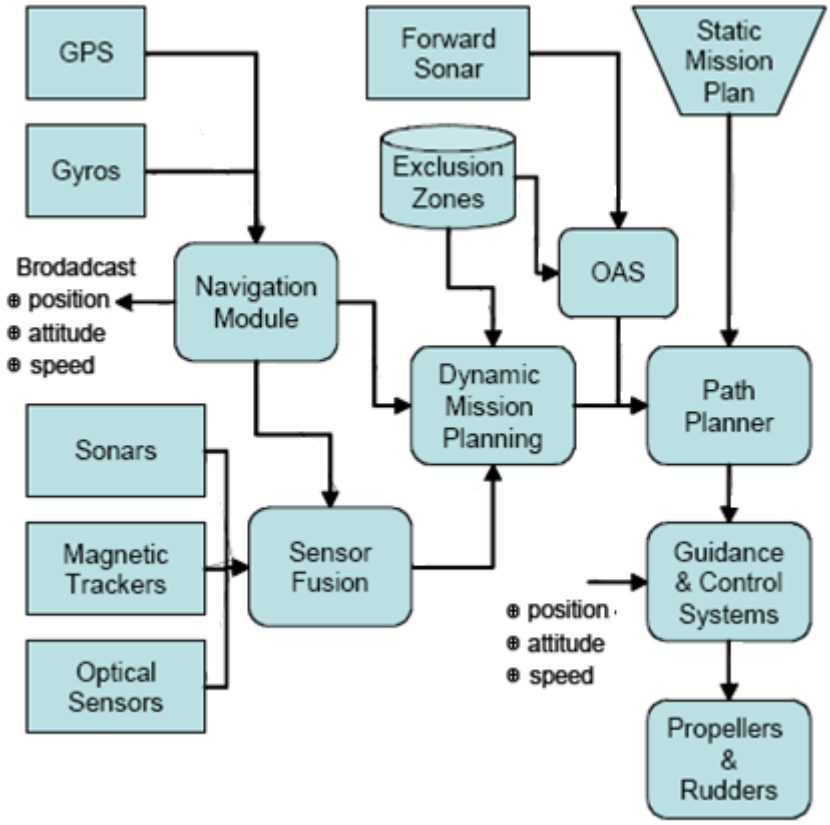


Figure C.1:*Hardware Modules Used to the Vehicle Development*

Finally, the actual command is sent to the Autopilot that controls thrusters, rudders and buoyancy. Then, this DMP is capable to redefine its original mission plan in an adaptive way, and hence it was called dynamic.

Relative to the remote sensing tasks devices, the main requirement is to integrate mainly two types of sonar sensors a Side Scan Sonar (SSS) and a Mechanical Scanning Echo Sounder (MSES) used as Single Beam Echo Sounder (SBES). The data extracted from sonars are used in two distinct ways, as an input to the sensor fusion module to improve the Navigation Module tasks in on-line operations and signal processing and in off-line processing using the sensor data stored from the survey operation performed.

The mechanical design constraint was inspired mainly by the before work developed by the UIB researcher member Dr. Gerardo Gabriel Acosta in his research project called Autonomous Underwater Vehicle for Inspections (AUVI), supported by the FP6 Marine Science and Technology Project (MIF1- CT-2004-003027). The AUVI prototype developed in this project uses the structure of two parallel hulls propeller (horizontal plane), using a binary actuation to navigate in the horizontal plane (yaw movement). The hull based components used is based in two singles units of Torpedo Dive Propulsion (TDP) (Torpedo 2016). The two hulls shape models were used in other projects along the world as the Deep Blue AUV and the Benthic Explorer AUV (O. I. Woods Hole 2016) and as the model of dynamics behavior to perform a navigation the catamaran feature was selected to avoid problems with rudders.

The control system was inspired in the Cormoran AUV project developed by Instituto Mediterráneo de Estudios Avanzados (IMEDEA) in collaboration with the University of the Balearic Islands (G. G. Acosta et al. 2007)(Valenciaga et al. 2007). The Cormoran AUV is a torpedo marine vehicle to support at shallower depths; your propulsion system is composed by one main electrical thruster with a guided mechanism based in a rudder. The immersion control was based on a density change by use of piston mechanism to environmental water admittance. The mathematical model for a physics interaction between vehicle and environment has been proved and publish in some papers(G. G. Acosta et al. 2007)(Valenciaga et al. 2007).

5.0 Hardware Architecture

To perform the AUV control system, the vehicle hardware was composed by two main processor boards: a low-level control and a high level control. The low-level control was based on an Ingenia microcontroller developer board: the iCm4011 development board, which uses a dsPIC 30F4011 as a main processor.

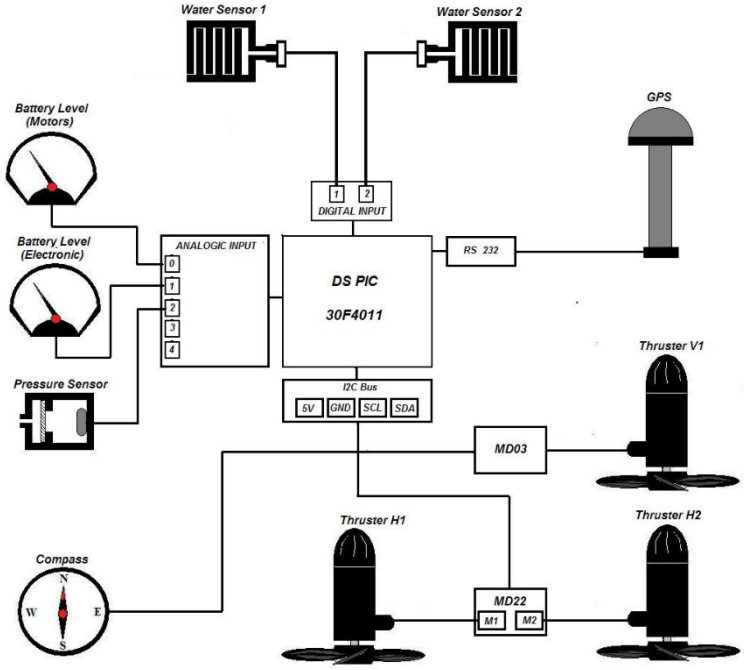


Figure C.2: *Sketch of hardware architecture from AUV prototype developed in IOGECS project.*

This board provides I2C and RS232 interfaces to communicate with sensors, actuators and high level CPU. The *Figure C.2* shows the low-level control system. At this level, we have some basics functions to read some sensors and to coverage an actuator controller to move the vehicle as proposed by the guidance system.

Literature

Acosta, G.G. et al., 2003. An Expert Navigator for autonomous underwater vehicle. In *The Argentine Symposium on Artificial Intelligence (ASAI'03)*. Buenos Aires.

Acosta, G.G. et al., 2009. Some Issues on the Design of a Low-Cost Autonomous Underwater Vehicle with an Intelligent Dynamic Mission Planner for Pipeline and Cable Tracking. *Underwater Vehicles, I-Tech Online Books, Robotics Series, I-Tech Education and Publishing KG, Viena, Austria*, (December), pp.1–18. Available at: http://www.intechopen.com/source/pdfs/6203/InTech-Some_issues_on_the_design_of_a_low_cost_autonomous_underwater_vehicle_with_an_intelligent_dynamic_mission_planner_for_pipeline_and_cable_tracking.pdf.

Acosta, G.G., Curti, H.J. & Calvo, O.A., 2005. Autonomous underwater pipeline inspection in AUTOTRACKER PROJECT: the navigation module. In *Europe Oceans*. Brest, France: IEEE, pp. 389–394.

Acosta, G.G. & Villar, S.A., 2015. Accumulated CA–CFAR Process in 2-D for Online Object Detection From Sidescan Sonar Data. *IEEE Journal of Oceanic Engineering*, 40(3), pp.558–569.

Acosta, G.G.L., 2005. Autonomous Underwater Vehicle for Inspections - AUVI.

Acosta, G.G.L., 2008. Autonomous Underwater Vehicle for Inspections (AUVI).

Albert A Michelson, 1927. *Studies in optics* T. U. of C. Press, ed., The University of Chicago Press.

Anderson, V.C., 1960. Digital Array Phasing. *The Journal of the Acoustical Society of America*, 32(7), pp.14–17.

Andò, B. & Graziani, S., 2000. *Stochastic Resonance - Theory and Application.pdf*, Boston/Dordretch/London: Kluwer Academic.

Anstee, S., 2001. *Removal of Range-dependent Imagery*, Fishermans Bend Vic, Australia. Available at: http://dspace.dsto.defence.gov.au/dspace/handle/1947/3648.

Arroyo-Suarez, E.N. et al., 2005. Evaluating a global differential GPS system for hydrographic surveying. In *Proceedings of MTS/IEEE OCEANS, 2005*. IEEE, pp. 1–7.

Ashida, G. & Kubo, M., 2010. Suprathreshold stochastic resonance induced by ion channel fluctuation. *Physica D: Nonlinear Phenomena*, 239(6), pp.327–334.

Atallah, L., Shang, C. & Bates, R., 2005. Object detection at diferent resolution in archaeological side-scan sonar images. In *Europe Oceans*. Brest, France: IEEE, pp. 287–292.

Bacelar, F.S., 2010. *PhD Thesis Nonlinear Dynamics and Regime Shifts in Ecosystems*. Universidad de Las Islas Baleares.

Baker, J.L. & Li, F., 2002. Quantifying Potential Measurement Errors and Uncertainties Associated with Bathymetric Change Analysis. , (September), pp.1–17.

Bartberger, C.L., 1966. *Lecture Notes on Underwater Acoustics*, Johnsville, Pennsylvania. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/468869.pdf.

Belov, E., 2015. BELOV Engenharia Ltda. Available at: www.belov.com.br.

Benedikt, M. et al., 2006. A Characterization of First-Order Topological Properties of planar spatial data. *Acm*, 53(2), pp.273–305.

Benzi, R. et al., 1982. Stochastic resonance in climatic change. *Tellus*, 34, pp.10–16. Available at: http://www.tellusa.net/index.php/tellusa/article/viewFile/10782/12407.

Bertero, M., Mol, C. De & Pike, E.R., 1985. Linear inverse problems with discrete data . I . General formulation and singular system analysis. *Inverse Problems l - The Institute of Physics*, p.301–330. Available at: http://iopscience.iop.org/0266-5611/1/4/004.

Besag, Y.J.E., 1972. On the correlation structure of some two-dimensional stationary processesr. *Oxford University Press on behalf of Biometrika Trust*, 68(2), pp.373–379.

Bex, P.J. & Makous, W., 2002. Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A*, 19(6), pp.1096–1106.

Bikonis, K. et al., 2005. THREE-DIMENSIONAL IMAGING OF SUBMERGED OBJECTS BY SIDE-SCAN SONAR DATA PROCESSING. In *Proceedings of the International Conference "Underwater Acoustic Measurements: Technologies &Results*. Crete, Greece.

Blondel, P., 2001. Seabed classification at ocean margins. *Ocean Margin Systems*, pp.125–141.

Blondel, P., 2009. *The Handbook of Sidescan Sonar* 1st ed., Springer-Verlag Berlin Heidelberg. Available at: http://www.springer.com/us/book/9783540426417.

Blueprint, D.E.L., 2016. StarFish GPS Receiver. , pp.4–7. Available at: http://www.blueprintsubsea.com/pages/product.php?PN=BP00070.

Blueprint, D.E.L., 2010. Starfish Towed Sonar System User Guide. , p.28. Available at: http://blueprintsubsea.com/starfish/support.php.

Blueprint, S., 2016. StarFish Towed Sonar User Guide. *Blueprnt Subsea Inc*. Available at: http://blueprintsubsea.com/starfish/index.php.

Boccaletti, S. et al., 2000. The Control of Chaos : Theory and Application. *Physics Reports Elsevier*, 329, pp.103–197.

Bowditch, N., 2002. *The American Practical Navigator*, Bethesda, Maryland. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:THE+AMERICAN+PRACTICAL+NAVIGATOR#4.

Box, G.E.P. & Muller, M.E., 1958. A Note on the Generation of Random Normal Deviates. *Ann. Math. Statist*, 29(2), pp.610–611. Available at: http://projecteuclid.org/euclid.aoms/1177706645.

Brenning, A. & Boogaart, K.G. Van Den, Geostatistics without Stationarity Assumptions within GIS. , (i), pp.1–14.

Brinkman, H.C., 1956. Brownian motion in a field of force and the diffusion theory of chemical reactions. II. *Physica*, 22(1–5), pp.149–155.

Brouwer, P.A.., 2008. *Seafloor classification using a single beam echosounder*. Delft University of Technology, Netherlands.

Brown, C.J. & Collier, J.S., 2008. Mapping benthic habitat in regions of gradational substrata: An automated approach utilising geophysical, geological, and biological relationships. *Estuarine, Coastal and Shelf Science*, 78(1), pp.203–214. Available at: http://www.sciencedirect.com/science/article/pii/S0272771407005288 [Accessed April 11, 2016].

Bulsara, A.R. et al., 2003. Signal detection via residence-time asymmetry in noisy bistable devices. *Physical Reviews E*, 67(1). Available at: http://journals.aps.org/pre/abstract/10.1103/PhysRevE.67.016120.

Bulsara, A.R. & Gammaitoni, L., 1996. Tuning in to Noise. *Physics Today*, 49(3), p.39. Available at: http://scitation.aip.org/content/aip/magazine/physicstoday/article/49/3/10.1063/1.881491 [Accessed April 10, 2016].

Byrnes, M.R., Baker, J.L. & Li, F., 2002. *Quantifying Potential Measurement Errors and Uncertainties Associated with Bathymetric Change Analysis*,

Calvetti, D. & Somersalo, E., 2007. *Introduction to Bayesian scientific computing: ten lectures on subjective ...* L. S. S.S. Antman, J.E. Marsden, ed., Available at: http://books.google.com/books?id=HIG0CRf6msoC&pgis=1.

Calvo, O. et al., 2008. Experimental results on smooth path tracking with application to pipe surveying on inexpensive AUV. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*.

Calvo, O. et al., 2009. Low-Cost Autonomous Underwater Vehicle for Underwater Acoustic. *Journal of Maritime Research*, VI(Ii), pp.37–52. Available at: http://www.jmr.unican.es/index.php/jmr/article/view/114/111.

Calvo, O. et al., 2009. Smooth path planning for autonomous pipeline inspections. In *2009 6th International Multi-Conference on Systems, Signals and Devices, SSD 2009*.

Calvo, O., Acosta, G. & Curti, H., 2005. ATMODULE- Modulo planificador de trayectorias de AUTOTRACKER.

Calvo Ibáñez, O.A., 2006. Inspección autónoma de oleoductos, gasoductos, emisarios y cables submarinos - IOGECS.

Castro, R. & Sauer, T., 1997. Chaotic Stochastic Resonance: Noise-Enhanced Reconstruction of Attractors. *Physical Review Letters*, 79(6), pp.1030–1033. Available at: http://link.aps.org/doi/10.1103/PhysRevLett.79.1030.

Celik, T. & Tjahjadi, T., 2011. A Novel Method for Sidescan Sonar Image Segmentation. *IEEE JOURNAL OF OCEANIC ENGINEERING*, 36(2), pp.186–194.

Chapeau-Blondeau, F., 2000. Noise, Oscillators and Algebraic Randomness: From Noise in Communication Systems of Number Theory Lectures of a School Held in Chapelle des Bois, France, April 5--10, 1999. In M. Planat, ed. *Noise, Oscillators and Algebraic Randomness*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 137–155. Available at: http://dx.doi.org/10.1007/3-540-45463-2_7.

Chapple, P., 2008. *Automated Detection and Classification in High-resolution Sonar Imagery for Autonomous Underwater Vehicle Operations*, Edinburgh South Australia. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/a501900.pdf.

Chiies, by J. & Chauvet, P., 1974. Kriging : a Method for Cartography of the Sea Floor. In *OCEANEXPO, 1974*. Bord eaux , October 1974.: International Hydrographic Review. Available at: https://journals.lib.unb.ca/index.php/ihr/issue/view/1798.

Chouhan, R., 2012. Image Denoising using Dynamic Stochastic Resonance in Wavelet domain. In *2012 12th International Conference on Intelligent Systems Design and Applications (ISDA)*. Kochi: IEEE, pp. 58–63.

Christensen, D. et al., 2005. Fast algorithms for the calculation of Kendall â€™ s. *Computational Statistics*, pp.51–62.

CHS, C.H.S., 2008. *Canadian Hydrographic Service STANDARDS FOR HYDROGRAPHIC SURVEYS*,

Available at: http://www.charts.gc.ca/data-gestion/hydrographic/standards-normes-eng.pdf.
Chu, D., 2004. EasyKrig. Available at: http://globec.whoi.edu/software/kriging/easy_krig/EasyKrigDescription.htm.
Chua, L.O., 1971. Memristor-The missing circuit element. *IEEE Transactions on Circuit Theory*, 18(5), pp.507–519.
Chua, L.O., 1992. *The genesis of Chua's circuit*, Berkeley. Available at: http://www.eecs.berkeley.edu/Pubs/TechRpts/1992/1924.html.
Clark, I., 1979. Practical Geostatistics. Available at: http://www.kriging.com/PG1979/#Chapter_1 [Accessed March 8, 2016].
Clifford, P.J., 1980. Time varying gain amplifier for side scan sonar applications. *The Journal of the Acoustical Society of America*, 68.
Cochrane, G.R. & Lafferty, K.D., 2002. Use of acoustic classification of sidescan sonar data for mapping benthic habitat in the Northern Channel Islands, California. *Continental Shelf Research*, 22(5), pp.683–690. Available at: http://www.sciencedirect.com/science/article/pii/S0278434301000899 [Accessed April 11, 2016].
Coiras, E., Petillot, Y. & Lane, D.M., 2003. Bathymetric side-scan backscatter map restoration based on data fusion. *Oceans 2003. Celebrating the Past ... Teaming Toward the Future (IEEE Cat. No.03CH37492)*, 1, pp.287–291.
Coiras, E., Petillot, Y. & Lane, D.M., 2007. Multiresolution 3-D reconstruction from side-scan sonar images. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 16(2), pp.382–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17269632.
Collier, J.S. & Brown, C.J., 2005. Correlation of sidescan backscatter with grain size distribution of surficial seabed sediments. , 214, pp.431–449.
Collins, J.J., Chow, C.C. & Imhoff, T.T., 1995. Aperiodic stochastic resonance in excitable systems. *Physical Review E*, 52(4). Available at: http://link.aps.org/doi/10.1103/PhysRevE.52.R3321.
Collins, J.J., Chow, C.C. & Imhoff, T.T., 2002. Stochastic resonance without tuning. *Nature*, (376), pp.236–238.
Collins, M.D. & Kuperman, W.A., 1994. Inverse problems in ocean acoustics. *Inverse Problems*, 10(5), p.1023. Available at: http://stacks.iop.org/0266-5611/10/i=5/a=003.
Cressie, N.A.C., 1993. *Statistics for spatial data / Noel A.C. Cressie.*, New York : J. Wiley,. Available at: https://encore.uib.es/iii/encore/record/C__Rb1016356;jsessionid=983B3968F5F312C13063E7B550EF37F1?lang=cat.
Cruz, J.M. & Chua, L.O., 1992. A CMOS IC Nonlinear Resistor for Chua's Circuit. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 39(12), pp.985–995.
Curtarelli, M. et al., 2015. Assessment of Spatial Interpolation Methods to Map the Bathymetry of an Amazonian Hydroelectric Reservoir to Aid in Decision Making for Water Management. *ISPRS International Journal of Geo-Information*, 4(October), pp.220–235. Available at: http://www.mdpi.com/2220-9964/4/1/220/.
Daniel R. Raichel, 2006. *THE SCIENCE AND APPLICATIONS OF ACOUSTICS THE SCIENCE AND APPLICATIONS OF ACOUSTICS SECOND EDITION*,
Dari, A., Bosi, L. & Gammaitoni, L., 2010. Nonlinear sensors : An approach to the residence time detection strategy. *PHYSICAL REVIEW E 81*, pp.1–10.
Dari, A. & Gammaitoni, L., 2009. Applications of Nonlinear Dynamics: Model and Design of Complex Systems. In V. In, P. Longhini, & A. Palacios, eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 225–235. Available at: http://dx.doi.org/10.1007/978-3-540-85632-0_18.
Diercks, A. et al., 2010. NIUST AUV ' s study Shipwrecks in the Northern Gulf of Mexico. In *2010 IEEE/OES Autonomous Underwater Vehicles*. Monterey, CA: IEEE, pp. 1–5. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5779679.
Diggle, P.J. & Ribeiro Jr., P.J., 2007. *Model-based Geostatistics* First. Springer, ed., New York, EUA: Springer.
Diggle, P.J., Ribeiro & Paulo, J., 2002. Bayesian inference in Gaussian model-based geostatistics. Available at: http://www.informaworld.com/smpp/content~content=a713668617~db=all~order=page.
Dura, E. et al., 2005. Active learning for detection of mine-like objects in side-scan sonar imagery. *IEEE Journal of Oceanic Engineering*, 30(2), pp.360–371.
Dura, E. et al., 2005. Active learning for detection of mine-like objects in side-scan sonar imagery. *IEEE Journal of Oceanic Engineering*, 30(2), pp.360–371.
Dura, E., 2011. Image Processing Techniques for the Detection and Classification of Man MAde Objects in Side-Scan Sonar Images. *IntechOpen Sonar Systems*. Available at: http://www.intechopen.com/articles/show/title/image-processing-techniques-for-the-detection-and-classification-of-man-made-objects-in-side-scan-so.
El-Hattab, A.I., 2014. Single beam bathymetric data modelling techniques for accurate maintenance dredging. *Egyptian Journal of Remote Sensing and Space Science*, 17(2), pp.189–195.

Elmore, P. a. et al., 2009. Uncertainty estimation of historical bathymetric data from Bayesian networks. *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*, pp.1–25.

Ensminger, D. & Bond, L.J., 2011. *Ultrasonics: Fundamentals, Technologies, and Applications, Third Edition* Third., Boca Raton, FL: CRC Press. Available at: https://www.crcpress.com/Ultrasonics-Fundamentals-Technologies-and-Applications-Third-Edition/Ensminger-Bond/9780824758899.

Fisher, P.E.F. & Tate, N.J.J., 2006. Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30(4), pp.467–489.

Fraden, J., 2010. *Handbook of Modern Sensors* Fourth., New York, EUA: Springer Science.

Fred, A.L.N. & Jain, A.K., 2006. Learning Pairwise Similarity for Data Clustering. In IEEE, ed. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*. Hong Kong.

G. G. Acosta, F. et al., 2007. Guiado y Control del Vehículo Autónomo Submarino CORMORÁN combinando acción por adelanto y lógica difusa. In *Anales XII RPIC*. Río Gallegos, Argentina.

Gammaitoni, L. et al., 1998. Stochastic resonance. *Rev Mod Phys*, 70(1), pp.223–287.

Gammaitoni, L. & Bulsara, A.R., 2002. Noise Activated Nonlinear Dynamic Sensors. *Phys. Rev. Lett.*, 88(23).

Gammaitoni, L. & Bulsara, A.R., 2003. Nonlinear sensors activated by noise. *Physical Reviews A*, pp.152 – 164.

Gang, H., 1992. Stochastic resonance in a nonlinear system driven by an aperiodic force. *Phys. Rev. Lett. A*, 46(6), pp.3250–3254.

Gao, J., 2009. Bathymetric mapping by means of remote sensing: methods, accuracy and limitations. *Progress in Physical Geography*, 33(1), pp.103–116.

Garcia-Moreno, E., Picos, R. & Al-Chawa, M.M., 2015. SPICE model for unipolar RRAM based on a flux-controlled memristor. In IEEE, ed. *2015 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*. Ixtapa: IEEE, pp. 1–4.

Ghys, É., 2012. The Butterfly Effect. In S. J. Cho, ed. *12th International Congress on Mathematical Education*. COEX, Seoul, Korea: Springer Open, pp. 19–39.

González, R.C., 2010. Digital Image Processing / Rafael C. González, Richard E. Woods. Available at: https://encore.uib.es/iii/encore/record/C__Rb1997595;jsessionid=A893C7E5FBE99097EF4A32508BC67C73?lang=cat [Accessed February 1, 2017].

Goovaerts, P., 1997. Geostatistics for natural resources evaluation / Pierre Goovaerts.

Goovaerts, P., 1999. Geostatistics in soil science: state-of-art and perspectives. *Geoderma*, 89, pp.1–45.

Grebogi, C. & Lai, Y.-C., 1997. Controlling chaotic dynamical systems. *System & Control Letters - Elsevier Science*, 31, pp.307–312.

Gumiaux, C., Gapais, D. & Brun, J.P., 2003. Geostatistics applied to best-fit interpolation of orientation data. , 376, pp.241–259.

Haining, R., 1991. Bivariate Correlation with Spatial Data. *Geographical Analysis*, 23(3), pp.210–227.

Hanggi, P., Talkner, P. & Borkovec, M., 1990. Reaction-rate theory - fifty years after kramers.pdf. *Reviews of Modern Physics - The American Physical Society*, 62(2), pp.251–341. Available at: http://www.physik.uni-augsburg.de/theo1/hanggi/Papers/112.pdf.

Hans-Hermann, B., 2008. Origins and extensions of the k-means algorithm in cluster analysis. *Journal Electronique d'Histoire des Probabilités et de la Statistique*, 4(December), pp.1–18.

Hari, V., Anand, G. & Premkumar, A., 2009a. Optimal suprathreshold stochastic resonance based nonlinear detector. In *Signal Processing Conference, 2009 17th European*. Glasgow: IEEE, pp. 2062–2066.

Hari, V., Anand, G. & Premkumar, A., 2009b. Preprocessor based on suprathreshold stochastic resonance for improved bearing estimation in shallow ocean. *Oceans 2009, Mts*, (65), pp.1–8. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5422099.

Harmer, G.. & Abbott, D., 2000. Simulation of circuits demonstrating stochastic resonance. *Microelectronics Journal*, 31(7), pp.553–559. Available at: http://linkinghub.elsevier.com/retrieve/pii/S002626920000029X.

Hastie, T., Tibshirani, R. & Friedman, J., 2009. *The Elements of Statistical Learning - Data Mining, Infrence and Predictions* Second. Springer, ed., Springer. Available at: http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf.

Hauke, J. & Kossowski, T., 2011. Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *Quaestiones Geographicae*, 30(2), pp.87–93.

Hell, B., 2011. *Mapping bathymetry*.

Helmuth Spath, 1985. *Cluster dissection and analysis: theory, FORTRAN programs, examples.* Chichester, ed., Ellis Horwood.

Hengl, T., 2007. *A Practical Guide to Geostatistical Mapping og Environmental Variables*,

Hengl, T. et al., 2008. Geostatistical modeling of topography using auxiliary maps. *Computers and Geosciences*, 34(12), pp.1886–1899.

Hermosilla, T. et al., 2005. Combining Features Extracted From Imagery and Lidar Data for Object-Oriented Classification of Forest Areas. *The International Archives of the Photogrammetry, Remote Sensing and*

Spatial Information Sciences, XXXVIII, pp.3–8.
Höfle, B., Griesbaum, L. & Forbriger, M., 2013. GIS-Based detection of gullies in terrestrial lidar data of the cerro llamoca peatland (peru). *Remote Sensing*, 5(11), pp.5851–5870.
Houlding, S., 2000. Practical geostatistics : modeling and spatial analysis / Simon W. Houlding. Available at: https://encore.uib.es/iii/encore/record/C__Rb1234782;jsessionid=FBA71FF30C77D2FB88C885DD84996E75?lang=cat [Accessed March 8, 2016].
IHO, I.H.O., 2012. *International Hydrographic Organization (IHO) GEOSPATIAL STANDARD - Bathymetric Surface Product Specification S-102*, Monaco. Available at: https://www.iho.int/mtg_docs/com_wg/HSSC/HSSC3/Draft_S-102_Ed1.0.0.pdf.
IHO, I.H.O., 2008. *STANDARDS FOR HYDROGRAPHIC SURVEYS - Special Publication No. 44*, Monaco. Available at: https://www.iho.int/iho_pubs/standard/S-44_5E.pdf.
Intellingence Agency of EUA, N.-N.G., 1984. World Geodetic System 1984. *WGS-84*, p.3.
IOC-UNESCO, I.O.C., 1985. *Manual on Sea Level*, United Kingdom. Available at: http://unesdoc.unesco.org/images/0014/001477/147773e.pdf.
Ippen, E., Lindner, J. & Ditto, W.L., 1993. Chaotic resonance: A simulation. *Journal of Statistical Physics*, 70(1), pp.437–450. Available at: http://dx.doi.org/10.1007/BF01053980.
James, L. & David, H., 2008. *Handbook of Multisensor Data Fusion*, Available at: http://dx.doi.org/10.1201/9781420053098.ch1.
Jha, R.K., Chouhan, R. & Biswas, P.K., 2012. Noise-induced contrast enhancement of dark images using non-dynamic stochastic resonance. In *Communications (NCC), 2012 National Conference on*. Kharagpur, India: IEEE, pp. 1–5.
Jiang, X. et al., 2008. Study and implementation of the methods of the side-scan sonar image processing. *Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008*, 6, pp.109–112.
Johnston, B.S., 2002. Uncertainty in Bathymetric Surveys. *Advances*, (September), pp.1–25.
Kapitaniak, T. & Bishop, S.R., 1999. The Illustrated Dictionary of Nonlinear Dynamics and Chaos. , p.277.
Karagianni, E.A., 2015. Electromagnetic Waves under Sea : Bow-Tie Antenna Design for Wi-Fi Underwater Communications. *Progress In Electromagnetics Research M*, 41(January), pp.189–198. Available at: http://www.jpier.org/PIERM/pierm41/19.15012106.pdf.
Kendall, A.M.G., 1938. Biometrika Trust A New Measure of Rank Correlation. , 30(1), pp.81–93. Available at: https://www.jstor.org/stable/2332226?seq=1#page_scan_tab_contents.
Kennedy, M.P., 1992. Robust op amp realization of Chua's circuit. *Frequenz*, 46, pp.66–80.
Kennedy, M.P., 1993a. Thee Steps to Chaos - Part II – A Chua's Circuit Primer. *Evolution, IEEE Transactions on Circuits and Systems*, 40(10), pp.657–674.
Kennedy, M.P., 1993b. Three Steps to Chaos. I. Evolution,.pdf. *Evolution, IEEE Transactions on Circuits and Systems*, 40(10), pp.640–656.
Kim, J.H.I., 1997. *Remote Monitoring and Control of Autonomous Underwater Vehicles*. Massachusetts Institute of Technology.
Kitanidis, P.K., 1997. *Introduction to Geostatistics: Applications in Hydrogeology*, California: Stanford University.
Korneta, W. et al., 2006. Experimental study of stochastic resonance in a Chua's circuit operating in a chaotic regime. *Physica D: Nonlinear Phenomena*, 219(1), pp.93–100.
Kovalevsky, J., Mueller, I.I. & Kolaczek, B., 2011. *Reference Frames: In Astronomy and Geophysics* 2nd ed. Springer, ed.,
Kubo, S., 1993. Classification of Inverse Problems Arising in Field Problems and Their Treatments. In M. Tanaka & H. D. Bui, eds. *Inverse Problems in Engineering Mechanics International Union of Theoretical and Applied Mechanics*. Tokyo: Springer Berlin Heidelberg Copyright Holder, pp. 51–60.
L. Gammaitoni et al., 1998. Stochastic resonance. *Reviews of Modern Physics*, 70(1), pp.223–287.
Legleiter, C.J. & Roberts, D.A., 2009. Remote Sensing of Environment A forward image model for passive optical remote sensing of river bathymetry. *Remote Sensing of Environment*, 113(5), pp.1025–1045. Available at: http://dx.doi.org/10.1016/j.rse.2009.01.018.
Lesser, M.P. & Mobley, C.D., 2007. Bathymetry, water optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery. *Coral Reefs*, 26(4), pp.819–829.
Li, H., Deng, Z.D., et al., 2012. Design parameters of a miniaturized piezoelectric underwater acoustic transmitter. *Sensors (Switzerland)*, 12(7), pp.9098–9109.
Li, H., Deng, Z.D. & Carlson, T.J., 2012. Piezoelectric Materials Used in Underwater Acoustic Transducers. *Sensor Letters*, 10(3). Available at: http://jsats.pnnl.gov/publications/peer/2012/2012_li_etal_pzt_review_paper_sensor_letters.pdf.
Li, Q. & Liu, Y., The Research on Measuring the Coefficient of Sound Absorption in Turbid Seawater. , pp.6557–6562.

Li, S.Y. et al., 2007. A novel design of piezo-resistive type underwater acoustic sensor using SOI wafer. *OCEANS 2006 - Asia Pacific*, pp.9–12.

Lindner, B., 2004. Effects of noise in excitable systems. *Physics Reports*, 392(6), pp.321–424. Available at: http://www.sciencedirect.com/science/article/pii/S0370157303004228 [Accessed April 10, 2016].

Lindner, B. et al., 2004. Effects of noise in excitable systems. *Physics Reports*, 392(6), pp.321–424.

Lockhart, C., Lockhart, D. & Martinez, J., 2001. *Comparing LiDAR and Acoustic Bathymetry Using Total Propagated Uncertainty (TPU) and the Combined Uncertainty and Bathymetry Estimator (CUBE) Algorithm Authors: Carol Lockhart, Doug Lockhart, José Martinez (Fugro Pelagos, Inc.)*, Seattle. Available at: http://www.fugro-pelagos.com/papers/ComparingLiDARandAcousticBathymetryUsingTPUandCUBE.pdf.

Lockhart, C., Lockhart, D. & Martinez, J., 2008. Total Propagated Uncertainty (TPU). *International Hydrographic Review*, 9(2), pp.19–27. Available at: https://journals.lib.unb.ca/index.php/ihr/article/view/20821.

Lorenz, E.N., 1963. Deterministic Nonperiodic Flow. *Journal of Athmospheric Science*, 20, pp.130–141.

Lucieer, V.L., 2007. Spatial uncertainty estimation techniques for shallow coastal seabed mapping. Available at: http://eprints.utas.edu.au/1919.

Lurton, X., 2010. *An Introduction to Underwater Acoustics*, Springer-Verlag Berlin Heidelberg.

Lurton, X., 2000. Swath bathymetry using phase difference: theoretical analysis of acoustical measurement precision. *IEEE Journal of Oceanic Engineering*, 25(3), pp.351–363.

Lurton, X., Lamarche, G. & at all, 2015. *Backscatter measurements by seafloor - mapping sonars Guidelines and Recommendations*, Available at: http://geohab.org/wp-content/uploads/2013/02/BWSG-REPORT-MAY2015.pdf.

Machell, F.W., Penrod, C.S. & Ellis, G.E., 1989a. Statistical Characteristics of Ocean Acoustic Noise Processes. In E. J. Wegman, S. C. Schwartz, & J. B. Thomas, eds. *Topics in Non-Gaussian Signal Processing*. New York, NY: Springer New York, pp. 29–57. Available at: http://dx.doi.org/10.1007/978-1-4613-8859-3_3.

Machell, F.W., Penrod, C.S. & Ellis, G.E., 1989b. Topics in Non-Gaussian Signal Processing. In E. J. Wegman, S. C. Schwartz, & J. B. Thomas, eds. *Topics in Non-Gaussian Signal Processing*. New York, NY: Springer New York, pp. 29–57. Available at: http://dx.doi.org/10.1007/978-1-4613-8859-3_3.

Marks, R.J.. I. et al., 2002. Stochastic resonance of a threshold detector: image visualization and explanation. *2002 IEEE International Symposium on Circuits and Systems. Proceedings (Cat. No.02CH37353)*, 4, pp.1–3.

Martins, M. et al., 2012. Optimization of piezoelectric ultrasound emitter transducers for underwater communications. *Sensors and Actuators A: Physical*, 184, pp.141–148. Available at: http://www.sciencedirect.com/science/article/pii/S0924424712003792 [Accessed April 11, 2016].

Matheron, G. Kleingeld, W.J., 1987. The Evolution of Geostatistics. In S. A. I. of M. and Metallurgy, ed. *APCOM 87 - Twentieth International Symposium on the Application of Computers and Mathematics in the Mineral Industries*. Johannesburg: South African Institute of Mining and Metallurgy, pp. 9–12. Available at: https://inis.iaea.org/search/search.aspx?orig_q=RN:19095057.

Matheron, G., 1970. La Théorie des Variables Régionalisées et ses Applications. , 5.

Matheron, G., 1971a. The theory of regionalized variables and its applications. *Les Cahiers du Centre de Morphlogie Mathematique*, 5, p.218.

Matheron, G., 1971b. *The Theory of Regionalized Variables and Its Applications*. École Nationale Supérieure des Mines de Paris. Available at: http://cg.ensmp.fr/bibliotheque/public/MATHERON_Ouvrage_00167.pdf.

Matheron, G., 1962. *Traité de géostatistique appliquée*, Paris: Editions Technip.

Matheron, G., Kleingeld, W. & Africa, S., 1987. The Evolution of Geostatistics. In Johannesburg, pp. 9–12.

Matsumoto, T., 1984. A Chaotic Attractor from Chua's Circuit. *IEEE Transactions on Circuits and Systems*, 31(12), pp.1055–1058.

May, R.M., 1976. Simple mathematical models with very complicated dynamics. *Nature*, (261), pp.459–467.

Mazel, C., 1985. *Side Scan Sonar Training Manual*, Salem, NH: Klein Associates, Inc.

McBratney, A.B. & Webster, R., 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables-II. Program and examples. *Computers and Geosciences*, 7(4), pp.335–365.

Mcdonnell, M.D. et al., 2005. Analog to digital conversion using suprathreshold stochastic resonance of Mathematical Sciences. *Biomedical Engineering*, 5649(1).

McDonnell, M.D. et al., 2006. Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 352(3), pp.183–189.

McDonnell, M.D. et al., 2005. Quantization in the presence of large amplitude threshold noise. *Fluctuation and Noise Letters*, 5(3), pp.L457–L468. Available at:

http://www.worldscientific.com/doi/abs/10.1142/S0219477505002884.
McDonnell, M.D. et al., 2012. *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization* 1st ed., Cambridge University Press.
Mcdonnell, M.D., Abbott, D. & Pearce, C.E.M., 2002. A characterization of suprathreshold stochastic resonance in an array of comparators by correlation coefficient. *Fluctuation and Noise Letters*, 2(3), pp.213–228.
Medwin, H. & Clay, C.S., 1998. *Fundamentals of acoustical oceanography* Academic Press, ed., Academic Press.
Minami, H., 2013. Estimation of total vertical uncertainty for the bathymetry acquired by autonomous underwater vehicle in deep water. *Report of hydrographic and oceanographic researches*, (50), pp.58–64.
Moss, F., 2004. Stochastic resonance and sensory information processing: a tutorial and review of application. *Clinical Neurophysiology*, 115(2), pp.267–281. Available at: http://www.sciencedirect.com/science/article/pii/S1388245703003304 [Accessed January 15, 2016].
Mussai, K., 2010. *Imaging and counting of targets with a high resolution multibeam sonar*. University of Cape Town.
Nadarajah, S., 2005. A generalized normal distribution. *Journal of Applied Statistics*, 32(7), pp.685–694.
Nelson, J.K., 2009. Signal Processing for Underwater Acoustic Communications. , (January), pp.90–96.
Nelson, P.R., Coffin, M. & Copeland, K.A.F., 2003. Introductory Statistics for Engineering Experimentation. Available at: https://encore.uib.es/iii/encore/record/C__Rb1953408;jsessionid=FF39E08C8ABA8DEC0FB2409A4FB0922C?lang=cat.
Nikitin, A., Stocks, N.G. & Bulsara, A.R., 2003. Asymmetric bistable systems subject to periodic and stochastic forcing in the strongly nonlinear regime: switching time distributions. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 68(1).
Nisbet, R., Elder, J. & Miner, G., 2009. Handbook of statistical analysis and data mining applications. Available at: https://encore.uib.es/iii/encore/record/C__Rb1910439;jsessionid=713BED61575D2765EFCBD41272EEF2E5?lang=cat.
NOAA Coastal Services Center, 2009. Coastal Inundation Mapping Guidebook. , (August), p.26.
NOAA Coastal Services Center, 2005. Elements of Geodesy - Datums. *NOAA Coastal Services Center*, p.46. Available at: http://www.ngs.noaa.gov/datums/vertical/ [Accessed January 28, 2016].
NOS, N.N.O. and A.A. of E., 2013. *Nos - Hydrographic Specifications and Deliverablse*,
Pearson, K., 1900. On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling. *Philosophical Magazine Serie S5*, 50(302), pp.157–172.
Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel, V. et al., 2011. Scikit-learn - Machine Learning in Python. *Journal of Machine Learning Research*, 12, pp.2825--2830. Available at: http://scikit-learn.org/stable/.
Peli, E., 1990. Contrast in complex images. *Journal of the Optical Society of America A*, 7(10), pp.2032–2040.
Penrose, J.D. et al., 2005. *Acoustic Techniques for Seabed Classification*,
Petillot, Y.R., Reed, S.R. & Bell, J.M., 2002. Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder. *Oceans '02 Mts/Ieee*, 1, pp.1234–1239.
Peyvandi, H. et al., 2011. Sonar Systems. In N. Z. Kolev, ed. *SONAR Systems and Underwater Signal Processing : Classic and Modern Approaches*. Intech Open, pp. 173–206. Available at: http://cdn.intechopen.com/pdfs/18876/InTech-Sonar_systems_and_underwater_signal_processing_classic_and_modern_approaches.pdf.
Piazza, A. Di, Piazza, M. Di & Vitale, G., 2009. A Kriging-Based Partial Shading Analysis in a Large Photovoltaic Field for Energy. *International Conference on Renewable Energies and Power Quality (ICREPQ'09)*. Available at: http://www.icrepq.com/ICREPQ'09/475_di_piazza.pdf.
Picos, R. et al., 2015. Semiempirical Modeling of Reset Transitions in Unipolar Resistive-Switching Based Memristors. *RADIOENGINEERING*, 24(2), pp.420–424. Available at: http://www.radioeng.cz/fulltexts/2015/15_02_0420_0424.pdf.
Pinkerton, J.M.M., 1947. A Pulse Method for the Measurement of Ultrasonic Absorption in Liquids: Results for Water. *Nature*, (160), pp.128–129.
Press, W.H. et al., 1992. *Numerical recipes in C++ : the art of scientific computing*, New York: Cambridge University Press.
Quinn, R. et al., 2005. Backscatter responses and resolution considerations in archaeological side-scan sonar surveys: a control experiment. *Journal of Archaeological Science*, 32(8), pp.1252–1264. Available at: http://www.sciencedirect.com/science/article/pii/S0305440305000701 [Accessed April 11, 2016].
Randall, R.E., 1997. Elements of Ocean Engineering. , pp.215–237.
Rao, C. et al., 2009. Underwater mine detection using symbolic pattern analysis of sidescan sonar images. In

American Control Conference. St. Louis, MO: IEEE, pp. 5416–5421.
Reed, S. et al., 2006. The fusion of large scale classified side-scan sonar image mosaics. *IEEE Transactions on Image Processing*, 15(7), pp.2049–2060.
Reed, S., Petilot, Y. & Bell, J., 2003. A model based approach to mine detection and classification in sidescan sonar. In *IEEE Oceans 2003. Celebrating the Past ... Teaming Toward the Future*. San Diego, CA, USA: IEEE, pp. 1402–1407.
Ribeiro Jr., P.J. & Diggle, P.J., 2001. {geoR}: a package for geostatistical analysis. *R-NEWS*, 1(2), pp.15–18. Available at: http://cran.r-project.org/doc/Rnews.
Richardson, S. & Hemon, D., 1981. On the Variance of the Sample Correlation between Two Independent Lattice Processes. *Journal of Applied Probability*, 18(4), pp.943–948. Available at: https://www.jstor.org/stable/3213069?seq=1#page_scan_tab_contents.
Ripley, B.D., 1977. Modelling Spatial Patterns. *Journal of the Royal Statistical Society*, 39(2), pp.172–212.
Robertsson, A., Berntorp, K. & Arz, K., 2011. Sensor Fusion for Motion Estimation of Mobile Robots with Compensation for Out-of-Sequence Measurements. *Control, Automation and Systems (ICCAS), 2011 11th International Conference on*, pp.211–216.
Robinson, T.P. & Metternicht, G., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. *Elsevier Computer and Electronics in Agriculture*, 50, pp.97–108.
Rongxing Li, 1991. Improvement of Bathymetric Data Bases By Shape From Shading. In IEEE, ed. *OCEANS "91. Ocean Technologies and Opportunities in the Pacific for the 90"s. Proceedings*. Honolulu, USA: IEEE, pp. 0–4.
Rongxing Li, 1992. RECONSTRUCTION OF SEAFLOOR SURFACE MODELS BY SHAPE FROM SHADING. In *isprs*. Washington, D.C., USA, pp. 780–784.
Ross, S.M., 2009. Introduction to probability and statistics for engineers and scientists. Available at: https://encore.uib.es/iii/encore/record/C__Rb1891387;jsessionid=672E00A0B9433D134E6DE5D660C7BDF6?lang=cat.
Rössler, O.E., 1979. CONTINUOUS CHAOS—FOUR PROTOTYPE EQUATIONS. *Annals of the New York Academy of Sciences*, 316(1), pp.376–392. Available at: http://dx.doi.org/10.1111/j.1749-6632.1979.tb29482.x.
Rudnick, P., 1960. Small Signal Detection in the DIMUS Array. *The Journal of the Acoustical Society of America*, 32(7).
Rus, G. & Gallego, R., 2002. Optimization algorithms for identification inverse problems with the boundary element method. *Engineering Analysis with Boundary Elements*, 26(4), pp.315–327. Available at: http://www.ugr.es/~grus/publications/conferences/rus_gallego_eabe.pdf.
Rus, G. & Gallego, R., 2007. Solution of identification inverse problems in elastodynamics using semi-analytical sensitivity computation. *Engineering Analysis with Boundary Elements*, 31(4), pp.343–360.
Ryu, C., Kong, S.G. & Kim, H., 2011. Enhancement of feature extraction for low-quality fingerprint images using stochastic resonance. *Pattern Recognition Letters*, 32, pp.107–113.
Samui, P. & Sitharam, T.G., 2011. Application of Geostatistical Models for Estimating Spatial Variability of Rock Depth. *Engineering*, 3(9), pp.886–894. Available at: http://www.scirp.org/journal/PaperDownload.aspx?DOI=10.4236/eng.2011.39108.
Sawant, S.R., 2010. Modeling of Acoustic Wave Absorption in Ocean. , 9(12), pp.19–24.
Schowengerdt, R.A., 2007. *Remote Sensing* Third. R. A. Schowengerdt, ed., Burlington: Academic Press. Available at: http://dx.doi.org/10.1016/B978-012369407-2/50000-0.
SeaBeam Instruments, L.-3 C., 2000. Multibeam Sonar Theory of Operation. , p.107.
Seber, G.A.F., 2004. *Multivariate Observations*, Wiley-Interscience.
Shan, J. & Toth, K.C., 2008. *Topographic Laser Ranging and Scanning: Principles and Processing* J. Shan & K. C. Toth, eds., CRC Press. Available at: https://www.crcpress.com/Topographic-Laser-Ranging-and-Scanning-Principles-and-Processing/Shan-Shan-Toth-Toth/9781420051438.
Siegel, R.M. & Read, H.L., Models of the temporal dynamics of visual processing. *Journal of Statistical Physics*, 70(1), pp.297–308. Available at: http://dx.doi.org/10.1007/BF01053969.
Silva, I.G. da, 2006. *Aspectos de sincronización en un sistema caótico Claudio Rubén Mirasso Santos y*. Universidad de Las Islas Baleares.
Simonotto, E. et al., 1997. Visual Perception of Stochastic Resonance. *Physical Review Letters*, 78, pp.1186–1189.
Sousa, A.L. et al., 2016. Resonancia Estocástica para el Mejoramiento del Contraste y Calidad en Imágenes Acústicas de Sonar de Barrido Lateral. In *IEEE ARGENCON 2016 – Congresso Bienal de IEEE Argentina 2016*. Buenos Aires: IEEE.
Sprott, J.C., 2003. *Chaos and Time-Series Analysis*, Oxford Press.
Stocks, N., 2000. Suprathreshold stochastic resonance in multilevel threshold systems. *Physical review letters*, 84(11), pp.2310–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11018872.

Stojanovic, M., 2008. Underwater acoustic communications: Design considerations on the physical layer. *Wireless on Demand Network Systems and …*, 1(2), pp.1–10. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4459349%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4459349.

Stojanovic, M. & Freitag, L., 2006. Multichannel detection for wideband underwater acoustic CDMA communications. *IEEE Journal of Oceanic Engineering*, 31(3), pp.685–695.

Stojanovic, M. & Preisig, J., 2009. Underwater acoustic communication channels: Propagation models and statistical characterization. *IEEE Commun. Mag.*, 47(1), pp.84–89. Available at: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4752682.

Strukov, D.B. et al., 2008. The missing memristor found. *Nature*, 453(7191), pp.80–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18451858.

Surveyors, International Federation Greenland, A. & Higgins, M., 2006. *FIG Guide on the Development of a Vertical Reference Surface for Hydrography*, Copenhagen, Denmark. Available at: http://www.fig.net/pub/figpub/pub37/pub37.pdf.

Tacconi, G. & B, W., 1981. The Remote Sensing of Factors Influencing Underwater Acoustics. In L. Bjørnø, ed. *Underwater Acoustics and Signal Processing*. Copenhagen: Springer Netherlands, pp. 93–98.

Tarantola, A., 2005. *Inverse problem theory and methods for model parameter estimation*, Available at: http://www.ncbi.nlm.nih.gov/pubmed/17069280.

Taroudakis, M., 2013. Monitoring The Sea Enviroment Using Acoustics, The Role of The Acoustical Observatories. In *44º CONGRESO ESPAÑOL DE ACÚSTICA ENCUENTRO IBÉRICO DE ACÚSTICA EAA EUROPEAN SYMPOSIUM ON ENVIRONMENTAL ACOUSTICS AND NOISE MAPPING*. Valladolid, Spain, pp. 1108–1123.

Tetzlaff, R. et al., 2016. Memristors-Devices, Models, Circuits, Systems and Applications. *International Journal of Circuit Theory and Applications*.

Ting, R.Y., 1992. A review on the development of piezoelectric composites for underwater acoustic transducer applications. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT,*, 41(1), pp.2181–2204.

Torpedo, D.P., 2016. Torpedo Dive Propulsion. Available at: http://www.torpedodpv.com/models.php.

Tsimplis, M.N. & Woodworth, P.L., 1994. The global distribution of the seasonal sea level cycle calculated from coastal tide gauge data. *Journal of Geophysical Research*, 99, p.16031.

Umbach, B.M.J., 1976. *Hydrographic Manual Fourth Edition*, Rockville, Md.

Urick, R.J., 1983. Principles of underwater / Robert J. Urick. Available at: https://encore.uib.es/iii/encore/record/C__Rb1197145;jsessionid=DA6677B136E1131EC76E6636DEDBDFC9?lang=cat.

USA, D.O.D., 2007. *Global Positioning System Precise Positioning Service (PPS) Performace Standard*, Available at: http://www.gps.gov/technical/ps/#spsps.

USA, D.O.D., 2008. *Global Positioning System Standard Positioning Service - SPS Service*, Available at: http://www.gps.gov/technical/ps/2008-SPS-performance-standard.pdf.

Vaccaro, R.., 1998. The Past, Present and Future of Underwater Signal Processing.pdf. *Signal Processing Magazine, IEEE*, 15(4), pp.21–51.

Valenciaga, F. et al., 2007. Trajectory Tracking of the Cormoran AUV Based on a PI-MIMO Approach. In *OCEANS - Europe*. Aberdeen, Schotland: IEEE, pp. 1–6.

Villar, S.A. et al., 2014. Evaluation of an efficient approach for target tracking from acoustic imagery for the perception system of an autonomous underwater vehicle. *International Journal of Advanced Robotic Systems*.

Villar, S.A. et al., 2013. Pipeline detection system from acoustic images utilizing CA-CFAR. In IEEE, ed. *OCEANS*. San Diego, CA, USA: IEEE, pp. 1–8.

Wackernagel, H., 1995. *Multivariate Geostatistics* First. Springer, ed., New York, EUA: Springer.

Waite, A.D., 2002. *Sonar for Practising Engineers* 3rd ed., Wiley.

Wandinger, U., 2005. Introduction to Lidar. In C. Weitkamp, ed. *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*. Leipzing, German: Springer Series in Optical Sciences, Volume 102, 2005, pp. 1–18. Available at: http://www.bookmetrix.com/detail/book/775bb1c2-b95b-40df-91e5-a32647f0d663#citations.

Wang, Q., 2006. Identifying Coexpressed Genes. In W. Hardle, Y. Mori, & P. Vieu, eds. *Statistical Methods for Biostatistics and Related Fields*. Berlim: Springer Berlin Heidelberg, p. 373.

Wang, W., Wang, G. & Tan, D., 2011. A New Memristor Based Chaotic Circuit. In IEEE, ed. *Fourth International Workshop on Chaos-Fractals Theories and Applications*. Hangzhou: IEEE, pp. 57–60.

Wellens, T., Shatokhin, V. & Buchleitner, A., 2004. Stochastic resonance. *Reports on Progress in Physics*, 67(1), p.45. Available at: http://stacks.iop.org/0034-4885/67/i=1/a=R02.

Whittle, P., 1954. On Stationary Processes in the Plane. *Biometrika*, 41(3), pp.434–449.

Willumsen, A.B., Hagen, O.K. & Boge, P.N., 2007. Filtering depth measurements in underwater vehicles for

improved seabed imaging. *2007 IEEE OCEANS-Europe*, pp.1–6. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4302249.

Wolberg, J., 2006. Data Analysis Using the Method of Least Squares [Recurs electrònic] : Extracting the Most Information from Experiments / by John Wolberg. Available at: https://encore.uib.es/iii/encore/record/C__Rb1903647;jsessionid=D0A9272ADEC05B1961A4C49ABE751050?lang=cat.

Woods Hole, C. and M.C., 2016. Woods Hole Coastal and Marine Center. *Woods Hole Coastal and Marine Center*. Available at: http://woodshole.er.usgs.gov/operations/sfmapping/sonar.htm [Accessed April 12, 2016].

Woods Hole, O.I., 2016. Autonomous Benthic Explorer - AUV. Available at: http://www.whoi.edu/page.do?pid=10779.

Yang, T., 1998. Adaptively optimizing stochastic resonance in visual system. *Physics Letters A*, 245(August), pp.79–86.

Ye, Q. et al., 2007. STOCHASTIC RESONANCE AD CONVERSION AND ITS EFFECT ON IMAGE. In *IEEE International Conference on Multimedia and Expo*. Beijing, China: IEEE, pp. 1675–1678.

Ye, Q., I, H.H. & I, C.Z., 2004. Image enhancement using stochastic resonance. *2004 International Conference on Image Processing (GIP)*, 100080(1), pp.263–266.

Yoon, Y. et al., 2012. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, 464–465, pp.363–375.

Yu, S., Zhang, A. & Li, H., 2012. A Review of Estimating the Shape Parameter of Generalized Gaussian Distribution. *Journal of Computational Information Systems*, 21(8), pp.9055–9064.

Zhong, G.-Q. & Ayrom, F., 1985. Experimental confirmation of chaos from Chua's circuit. *International Journal of Circuit Theory and Applications*, 13(1), pp.93–98. Available at: http://dx.doi.org/10.1002/cta.4490130109.