UNIVERSITY OF PELOPONNESE

DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS

DOCTORAL THESIS

Experimental Assessment of the Atmospheric Effects on Laser Communications On Maritime Environment

Antonios Lionis

Tripoli, December 2023

Περίληψη

Στην παρούσα διδακτορική διατριβή ερευνώνται οι ασύρματες οπτικές επικοινωνίες, γνωστές και ως επικοινωνίες λέιζερ. Η εν λόγω τεχνολογία κάνει χρήση του οπτικού και υπερύθρου φάσματος για την μετάδοση της πληροφορίας και προσφέρει σημαντικά πλεονεκτήματα σε σχέση με την υφιστάμενη κατηγορία των επικοινωνιών με ραδιοκύματα, όπως αυξημένες δυνατότητες ρυθμού μετάδοσης δεδομένων, ασφάλεια έναντι υποκλοπών και παρεμβολών, χαμηλό κόστος και ευκολία εγκατάστασης και τέλος κανέναν περιορισμό στις χρησιμοποιούμενες συχνότητες εκπομπής. Η μεγαλύτερη πρόκληση που αντιμετωπίζουν τα συστήματα επικοινωνίας λέιζερ, είναι οι εγγενείς περιορισμοί τους από τις περιβαλλοντικές συνθήκες και ιδιαίτερα από το φαινόμενο της ατμοσφαιρικής τυρβώδους ροής. Η τεχνολογία αυτή έχει τόσο επίγειες όσο και υποβρύχιες και διαστημικές εφαρμογές. Κάθε μια από τις ανωτέρω εφαρμογές, καλείται να αντιμετωπίσει διαφορετικές δυσκολίες και περιορισμούς, λόγω των διαφορετικών ατμοσφαιρικών συνθηκών που η διαδιδόμενη δέσμη λέιζερ συναντά στην πορεία της.

Στην παρούσα διατριβή εξετάζεται, πειραματικά, η απόδοση μιας ασύρματης οπτικής ζεύξης που λειτουργεί σε θάλασσιο περιβάλλον. Το φαινόμενο της ατμοσφαιρικής τυρβώδους ροής σε ένα τέτοιο περιβάλλον εμφανίζει αρκετές διαφορές με το αντίστοιχο άνωθεν επίγειου περιβάλλοντος και είναι αυτό που ευθύνεται για το φαινόμενο του σπινθηρισμού, δηλαδή της χωρικής και χρονικής μεταβολής της λαμβανόμενης οπτικής ισχύος από τον δέκτη. Η ταχύτητα των μεταβολών αυτών απαιτεί την χρήση στατιστικών κατανομών για την μελέτη τους. Διάφορες τέτοιες κατανομές έχουν μελετηθεί και αποδειχθεί ότι μπορούν να περιγράψουν με αρκετή

ακρίβεια τις διακυμάνσεις της λαμβανόμενης ισχύος σε άμεση συνάρτηση με την ένταση του φαινομένου της ατμοσφαιρικής τυρβώδους ροής, μετρούμενης με την παράμετρο κατανανομής του δείκτη διάθάσης στην ατμόσφαιρα, C_n^2.

Βασικός στόχος της παρούσας μελέτης είναι η, σε πειραματικό επίπεδο, συσχέτιση του φαινομένου αυτού με την απόδοση μιας ασύρματης οπτικής ζεύξης άνωθεν θαλασσίου περιβάλλοντος. Συγκεκριμένα, με χρήση κατάλληλων πειραματικών διατάξεων, τόσο στο κτίριο εργαστηρίων της Σχολής Ναυτικών Δοκίμων (ΣΝΔ) όσο και στις εγκαταστάσεις του Naval Postgraduate School (NPS) στο Μοντερέυ της Καλιφόρνια, γίνεται συλλογή βασικών ατμοσφαιρικών δεδομένων (ταχύτητα ανέμου, σχετική υγρασία, θερμοκρασία αέρα, ηλιακή ακτινοβολία, σημείο δρόσου, ατμοσφαιρική πίεση και ρυθμός βροχόπτωσης), τιμών της παραμέτρου κατανομής δείκτη διάθλασης της ατμόσφαιρας (C_n^2) με διάφορες μεθόδους και του λαμβανόμενου σήματος στον δέκτη του συστήματος. Εν συνεχεία, γίνεται μοντελοποίηση παλινδρόμησης και ταξινόμησης του λαμβανόμενου σήματος και της παραμέτρου κατανομής δείκτη διάθλασης της ατμόσφαιρας βάσει των ατμοσφαιρικών δεδομένων τόσο με χρήση βασικών πολυωνυμικών μοντέλων, όσο και με χρήση διαφόρων αλφορίθμων μηχανικής μάθησης, όπως Random Forest, Gradient Boosting, Decision Trees, k-Nearest Neighbors καθώς και νευρωνικών δικτύων μονού και πολλαπλών επιπέδων. Εξετάζεται η αξιοπιστία διάφορων βασικών κατανομών πυκνότητας πιθανότητας του λαμβανομένου σήματος του συστήματος με χρήση μεθόδων θεωρίας της πληροφορίας (Kullback-Leibler και Jensen-Shannon divergence). Εξάγεται με χρήση θεωρητικού μοντέλου η πιθανότητα αποκοπής της ζεύξης και εν συνεχεία μοντελοποιείται η ταξινόμηση της σε λειτουργική ή μη, με χρήση βαθέων νευρωνικών δικτύων. Τέλος γίνεται μοντελοποίηση πειραματικών δεδομένων της παραμέτρου κατανομής δείκτη διάθλασης με χρήση αλγορίθμων

μηχανικής μάθησης (Random Forest, Neural Network και Gradient Boosting) καθώς και ταξινόμηση αυτών αναλόγως της έντασης, με χρήση βαθέων νευρωνικών δικτύων.

Abstract

This dissertation investigates the free space optical communications (FSOCs) technology, the so-called laser communications (LaserComm). The FSO technology operates in the optical and infrared spectrum for data transmission and offers significant advantages comparing to their radio frequency (RF) counterparts. These include increased data rate transmission, greater security and immunity, lower cost of installation and, finally, no license restrictions. The major challenge that FSO systems face, is the degradation effects of the atmosphere and in especially the optical turbulence. FSO technology applications span from terrestrial and underwater to space. Each FSO application face different kind of challenges due to the variety of the atmospheric conditions that a propagating beam encounters.

An FSO link over a maritime environment is investigated. The optical turbulence phenomenon in a maritime environment exhibits significant differences with a terrestrial one and causes the effect of scintillation, that is the spatial and temporal variation of the irradiance in the receiver. To study these variations, statistical distributions are required. Already, many different probability density functions (PDFs) have been studied and related the received irradiance with the optical turbulence effect, quantified by the refractive index structure parameter, C_n^2.

The major goal of this dissertation is the experimental assessment of the optical turbulence effect towards an FSO link over maritime environment. We utilized the experimental site of the Hellenic Naval Academy (HNA), Piraeus, Greece as well as the site of the Naval Postgraduate School (NPS) in Monterey, California to collect macroscopic meteorological parameters (wind speed, relative humidity, air

temperature, solar radiation, dew point, air pressure and rainfall rate), C_n^2 values with different methods and the received signal strength (RSSI) on the FSO terminal. Based on this data, initially we developed second-order polynomial models for the RSSI parameter, using regression modeling, to quantify its relation with the macroscopic environmental parameters and demonstrated the prediction accuracy of various machine learning based algorithms (Random Forest, Gradient Boosting, Decision Trees, k-Nearest Neighbors, Neural Networks) for an FSO link performance. The similarity between different probability distributions is investigated, by employing an information theoretic method, namely the so-called Jensen–Shannon divergence, a symmetrization of the Kullback–Leibler divergence. The outage probability of the link is estimated by utilizing a theoretical model and the operational status of the link, either functional or not, is modeled by a classification deep neural network (DNN). Finally, the refractive index structure parameter, C_n^2, is modelled using ML-based regression (Random Forest, ANN και Gradient Boosting) and classification algorithms (DNN).

Preface

The work presented in this thesis was carried out as a part of my Ph.D. project in the period from January 20th, 2020 to October 20th, 2023. This project was conducted as a student of the Information and Telecommunications Department of the University of Peloponnese, at Tripoli, Greece. The primary experimental setup utilized to obtain the analyzed data sets, was located on the premises on the Hellenic Naval Academy (HNA) under the Division of Combat Systems, Naval Operations, Sea Sciences, Navigation, Electronics & Telecommunications Sector. However, the facilities of the Naval Postgraduate School (NPS), Monterey, California, have also been exploited in terms of experimental data collection. Specifically, as will be discussed in the following chapters in detail, two different experimental setups have been utilized, the first across the Monterey Bay and the second at the roof of the Spanagel Hall building of the university.

This Ph.D. project was supervised by:

- Konstantinos Peppas (supervisor), Assistant Professor, Department of Information and Telecommunications, University of Peloponnese, Tripoli, Greece
- Andreas Tsigopoulos (Ph.D. committee), Professor, Division of Combat Systems, Naval Operations, Sea Sciences, Navigation, Electronics & Telecommunications Sector of the Hellenic Naval Academy, Piraeus, Greece
- Hector Nistazakis (Ph.D. committee), Section of Electronic Physics and Systems, Department of Physics, National and Kapodistrian University of Athens, Panepistimiopolis Zografou, Athens, Greece

On the context of my Ph.D. project and apart from the research work presented in this thesis, several other related tasks have also been undertaken, which briefly include the following:

- Collaboration with the Directed Energy Group of the NPS in topics such as data exchange, model validation, expertise exchange and collaborative publications.
- Collaboration with the Naval Information Warfare Center (NIWC), Pacific in topics such as experimental data and expertise exchange as well as collaborative publications in international conferences and peer reviewed publications.
- Research project proposal for the 5th/20 call for funding of the Hellenic Foundation for Research and Innovation (HFRI).
- Research project proposal for the Defense Innovation Challenge announced by the Hellenic Ministry of Defence in November 2021. Our proposal was rank in the 7^{th} place out of two hundred twelve proposals totally.
- Collaboration with the EMTECH SPACE consortium on the context of the "Greek CubeSats In-Orbit Validation" project as a subject matter expert in optical communications atmospheric modeling.

Acknowledgements

It would be ungrateful not to mention that such an accomplishment requires significant support in many different aspects of it. That being said, I need to express my deep gratitude to the following:

- My supervisor, Assistant Professor Konstantinos Peppas, because from the very beginning of our cooperation till the end of it, he provided me with trust and gave me the absolutely required flexibility and space to move forward towards the research goals we had agreed.
- Professor Andreas Tsigkopoulos for his valuable assistance with data collection and experimental setup superintendence. Without his support and given my professional responsibilities, my progress would have been impossible.
- Professor Hector Nistazakis, for his constant support and his recognition of my work, which provided me with the necessary confidence in moving forward.
- Professor Keith Cohn, for his kind offer of sharing data collected at NPS and during the experimental campaign of his research group. This fact, gave a significant boost to my research goals.

Contents

List of Figures

List of Tables

Ph.D. Publications

The following publications have resulted from this Ph.D. project.

Articles in peer-reviewed journals

[J1] Lionis, A.; Peppas, K.; Nistazakis, H.E.; Tsigopoulos, A.D.; Cohn, K. Experimental Performance Analysis of an Optical Communication Channel over Maritime Environment. Electronics 2020, 9, 1109.

[J2] Lionis, A.; Peppas, K.; Nistazakis, H.E.; Tsigopoulos, A.D.; Cohn, K. Statistical Modeling of Received Signal for an FSO Channel Over Maritime Environment. Oct 2020. Optics Communications 2020. Volume 489, 2021,126858.

[J3] Lionis, A.; Peppas, K.; Nistazakis, H.E.; Tsigopoulos, A.D.; Cohn, K. and Zagouras A. Using Machine Learning Algorithms for Accurate Received Optical Power Prediction of an FSO Link over Maritime Environment. Photonics 2021, 8(6), 212

[J4] Lionis, A.; Sklavounos, A.; Stassinakis, A.; Cohn, K.; Tsigopoulos, A.; Peppas, K.; Aidinis, K.; Nistazakis, H. Experimental Machine Learning Approach for Optical Turbulence and FSO Outage Performance Modeling. Electronics 2023, 12, 506.

[J5] Lionis, A.; Peppas, K.; Tsigopoulos, A.; Nistazakis, H.; Cohn, K.; Drexler, K. Supervised Machine Learning for Refractive Index Structure Parameter Modeling. Quantum Beam Science 2023, 7, 18.

[J6] Lionis, A.; Peppas, K.; Nistazakis, H.E.; Tsigopoulos, A.D. RSSI probability density functions comparison using Jensen-Shannon divergence and Pearson distribution. Technologies 2021, 9(2), 26.

[J7] Lionis, A.; Peppas, K.; Nistazakis, H.E.; Tsigopoulos, A.D.; Cohn, K., Chaskakis, G. and Blau, J. Optical Turbulence Measurements and Modeling over Monterey Bay and Potential Effects on Free Space Optical Communication Performance. Optical Communications, Vol. 520, 1 Oct 22.

[J8] Lionis, A., Tsigopoulos A., 2022. High Energy Laser Weapon Integration Issues for the Future Hellenic Frigate. NAUSIVIOS CHORA 2022 ed.

[J9] Lionis, A.; Cohn, K.; Tsigopoulos, A.D. An Application of Artificial Neural Networks to Estimate the Performance of High-Energy Laser Weapons in Maritime Environments. Technologies 2022, 10(3), 71.

[J10] Lionis, A.; Peppas, K.; Tsigkopoulos, A.; Nistazakis, H.E.; Cohn, K. and Drexler, K.R. "Supervised Machine Learning for Refractive Index Structure Parameter Modeling," *2023 12th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, Athens, Greece, 2023, pp. 1-4, doi: 10.1109/MOCAST57943.2023.10176724.

Contributions to international peer-reviewed conferences

[C1] Lionis, A.; Peppas, K.; Tsigopoulos, A.; Nistazakis, H.; Cohn, K.; Drexler, K. Supervised Machine Learning for Refractive Index Structure Parameter Modeling. 12th International Conference on Modern Circuits and Systems Technologies (MOCAST) on Electronics and Communications. 28 – 30 Jun 2023. Athens, Greece.

CHAPTER 1

Introduction

Over the last few decades, a new application of a rather older technology arose, the so-called free space optical (FSO) communication or else, the laser communication (LaserComm). As appears in its definition, the FSOC technology makes use of a laser beam as its medium carrier for the information transfer through the free space. The capacity requirements for the information channels in the 21st century calls for an effective alternative of the traditional radio frequency (RF) technology and the FSO technology aims to fill this gap.

1.1 Background

Initially driven by the requirements of the defense sector, the first attempts for FSO systems developments were focused on military solutions [1]. Later on, commercial applications were also developed, spanning in any kind of environment, from space to underneath the sea surface. Perhaps, the most evident asset of FSOC is their potential for increased data rate (several 10s of Gb), as compared to their RF counterpart, due to their inherently very high carrier frequency (on THz level) and therefore bandwidth [1]. Despite being applicable for operation in almost any kind of environment, currently the terrestrial applications are the most widely implemented. Links up to several kilometers can be established, given certain limitations that can affect this range. As a result, prior to the installation of an FSO link, the typical local atmospheric conditions must have been measured in order to be able to predict the performance of the link. Other advantages of the FSO systems include high security

and immunity against electromagnetic interference (LPD/LPI), very low installation expenditures and power requirements as compared to RF links and no license requirements for operation [2]. The broad spectrum of the potential FSO applications can be classified into three main categories: (i) Optical Wireless Satellite Networks (OWSNs), (ii) Optical Wireless Terrestrial Networks (OWTNs) and (iii) Optical Wireless Home Networks (OWHNs) (Figure 1.1) [2]. OWSNs are supposed to exploit the satellite infrastructure in order to provide a global backbone wireless network and cover almost any terrestrial point. OWSNs can constitute an alternative to the existing wired internet, especially for the maritime communications that are actualized through optical fibers laid on the seabed. The OWTNs refer to wireless networks between certain points on any terrain, but limited from the line-of-sight (LOS) requirement. However, they can provide a variety of solutions such as “last mile” applications, complementing the existing optical fiber cables, bridge geographical separated locations, such as ship to ship or building to building and finally provide an alternative operational mode to RF networks in order to improve their availability. OWHNs are wireless networks confined in the interior of a building or a house. They can be either LOS or non-LOS links, also known as diffused links, exploiting the backbone infrastructure of the building [2].

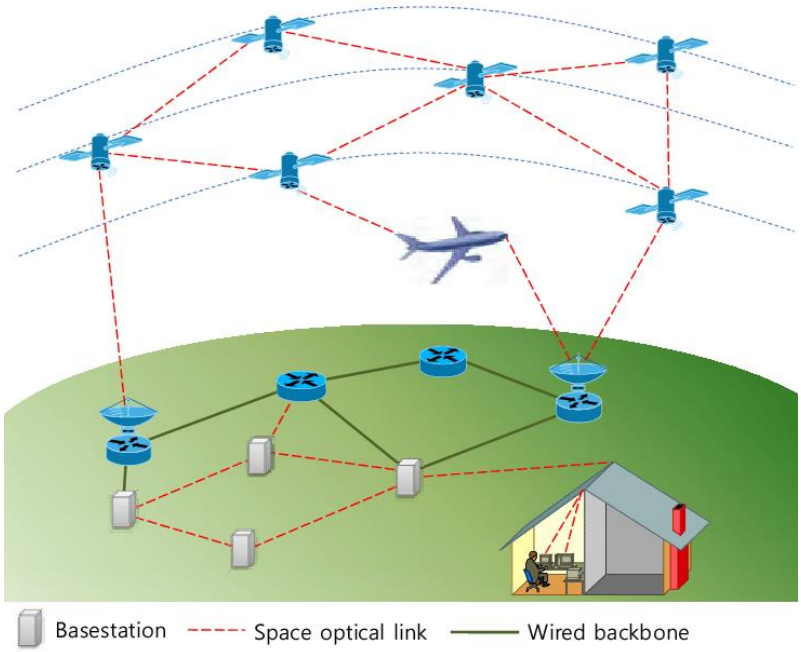


Figure 1.1. The conceptual topology of integrated optical wireless networks [3].

The transmitting light can be modulated by frequency, amplitude, phase or polarization. Most frequent are either amplitude modulation combined with direct detection or phase modulation combined with homodyne or heterodyne receiver [3]. The technically simplest scheme is the on-off keying (OOK), where “on” refers to transmission of light and “off” for a pause and correspond to a logic “1” and “0”, respectively. This modulation comes with a non-return-to-zero (NRZ) or return-to-zero (RZ) coding. Both of them are sensitive to long strings of ones or zeros transmission and this could lead to loss of synchronization. The addition of 8B10B coding along with the NRZ assures that no inter-symbol interference will occur because the average signal remains constant [3]. Coherent modulation systems are also used, such as binary phase shift keying (BPSK) or differential phase shift keying (DPSK) which are much more sensitive than the OOK systems but also more complex and vulnerable to phase distortions of the received beam [3]. Generally, OOK systems are considered much more robust in terms of atmospheric effects and with higher supportability from market components, consequently they are preferred for free space links.

A significant drawback of the FSO systems that could potentially lead to a total blackout of the link, is the atmospheric effects. Since the atmosphere consists of many different particles like aerosols, dust, smoke etc. or large precipitation due to rain, haze or snow, the laser beam propagating through this may face significant power loss because of the attenuation or scattering effects [4]. The main factors that cause power loss to the propagating beam are the following:

a. Absorption and Scattering Losses: Constitute the main attenuation factor. At the specific spectrum part of interest for FSO systems (IR and visible) the molecular absorption has the strongest effect. The optical depth term, τ, correlates the power transmitted P_T and received P_R, whose ratio defines the transmittance, as follows [4]:

$$P_R = P_T \exp(-\tau) \tag{1.1}$$

The total attenuation coefficient of the propagating signal is comprised of four factors, that are the absorption and scattering coefficients of the molecules and aerosols [4]:

$$\gamma(\lambda) = a_m(\lambda) + a_a(\lambda) + \beta_m(\lambda) + \beta_m(\lambda) \tag{1.2}$$

The absorption coefficients are strongly dependent on the wavelength; thus the known atmospheric windows have to be exploited in order to avoid high absorption losses. The scattering coefficients depend on the radius of the particles that the beam encounters and classified in the Rayleigh scattering if the wavelength is larger and Mie scattering if their size is comparable. For smaller wavelengths, the scattering can be described using the diffraction theory.

b. Free-Space Losses

The propagation of the light in free space causes, usually the highest losses to an FSO system. The free space losses are given by [4]:

$$L_s = \left(\frac{\lambda}{4\pi R}\right)^2 \tag{1.3}$$

where R is the range of the link and is significantly higher than the RF equivalent free-space losses.

c. Beam Divergence Loss

The diffraction phenomenon causes the propagating beam to spread out as it travels away from the transmitter. This fact along with the narrow field of view (FOV) of the FSO receiver, allows only a small amount of the transmitted energy to be collected, which can be approximated as follows [4],

$$P_R \approx P_T \left(\frac{D_T D_R}{\lambda R}\right)^2 \tag{1.4}$$

where D_T and D_R are the antenna diameter of the transmitter and the receiver, respectively. Despite preferring a narrow beam divergence, a careful trade-off with the pointing and tracking requirements has to be made, because these two features are inversely proportional.

d. Weather Conditions Losses

As mentioned previously, the propagating beam faces power reduction due to existing particles along its path. The most prominent factor because its size is comparable to the beam's wavelength, is fog, which can ultimately lead to link outage. Visibility is a measure that is used to estimate the optical attenuation of the beam, characterized by the specific attenuation, defined as the attenuation per unit length expressed in dB/km, given by [4],

$$\beta(\lambda) = \frac{1}{R} 10 \log\left(\frac{P_T}{P_R}\right) \tag{1.5}$$

The specific attenuation (Eq. 1.5) due to fog, snow and rain takes a different form, mainly based on the visibility and has been extensively modeled.

e. Pointing Losses

FSO systems, either stationary or mobile are very sensitive in terms of alignment. Any imperfection can lead to link outage, therefore either a perfect alignment with frequent corrections for stationary systems or an advanced acquisition, tracking and pointing (ATP) subsystem for mobile systems, are absolutely necessary.

The unique features of the FSO technology have triggered the FSO community to put a lot of effort towards their further development. However, plenty of issues have yet to be solved, in order to fully benefit from them. The most prominent research challenges that still need to be overcome are the following [2]:

a. *Channel modeling*: There is a significant need to fully describe an optical channel in terms of optical turbulence effects on the laser propagation, so as we could come up with ways to improve it. The numerous potential applications make this issue even more complicate. Therefore, new efficient ways to model the channel are required, such as machine learning algorithms.

b. *Pointing, Acquisition and Tracking (PAT)*: The extremely narrow transmitted beam (few mrads) and the narrow FOV of the receiver, result in very challenging PAT requirements, especially for applications with extended propagation distance. Several hardware design approaches exist, which propose either a coarse-pointing or a coarse-sensing beacon signal, for easier receiver detection [5, 6]. Advanced hardware architecture and improved synchronization methods are still two major issues that need to be solved [2]. A way to increase the bandwidth of a link -of FSO systems that has been adopted recently, is the use of space division multiplexing, where in order to multiplex many encoded data channels within a communication system orthogonal or spatially separable optical modes are used [7]. This type of FSO systems require very accurate PAT, a fact that limits the use of this technology due to the increased cost of a high performance PAT subsystem. Abadi et. al. have proposed

a very low cost prototype FSO system with auto locate (via GPS) and align capability based on commercial of the shelf (COTS) components [7]. This prototype includes a telescope mount controlled by a computer and a Small Form Pluggable (SFP) transceiver operating at 1550 nm which uses an algorithm to self-align the transceiver with nearby transceivers that have been located by the GPS.

c. *Advanced hardware design*: FSO systems are quite vulnerable to atmospheric effects that directly affect the reliability of the link. However, several solutions have been adopted, in terms of hardware design, to overcome this issue, such as the hybrid RF/FSO and the Multiple Input Multiple Output (MIMO) architecture. It has been shown that this spatial diversity technique can decrease the outage probability and increase the power gain [8]. Examples of advanced techniques on hardware design include the control of transmitting divergence and receiving FOV as well as the incorporation of the Wavelength-Division Multiplexing (WDM) [2]. An architecture has been proposed by Born et. al. to allow multidirectional communications in an asymmetric link using a passive uplink with retro-modulation without requiring the increase of power and the number of elements to steer the beam. Such a link uses an active link which is being modulated by the retromodulator and reflected as a passive link. This architecture reduces significantly the PAT requirements and allow a number of transceivers, within the FOV of the retromodulator, to be linked [9]. It is ideal for implementation between a stationary and a mobile platform, where the former does not have size and power restrictions, whereas the latter is able to compensate with the inevitable platform jitter disturbances. However, these come at the cost of the data rate capacity. Recently, multiple-quantum-well (MQW) modulators in conjunction with corner cubes and cat's eye retroreflectors have been demonstrated and allow data rates at the level of several tens of Mbps with the potential for Gbps speeds [9]. Quintana et.

al. presented the design and operation of a MQW-based modulator embedded in an asymmetric Fabry-Perot cavity. This modulator was tested in an outdoor environment, operating at 150 Mbps with a bit error rate (BER) of 1.22×10^{-6} in a distance of 200 m [10].

d. *FSO networking*: This FSO related research area include topics such as integrated topology design, topology design with QoS guarantees, distributed topology design and dynamic rate allocation [2].

e. *Transport layer issues*: Up to now, there have not been proposed successful transport protocols [11]. The extended propagation distance and high data rate transmission, cause poor efficiency in typical transport protocols (TCP). Currently, existing TCPs possess a throughput of less than 10%, despite using advanced hardware techniques, resulting in lack of broadband services exploitation [2].

An eminent FSO project that leads the way to the future of laser communications technology, is the NASA's Lunar Laser Communication Demonstration (LLCD), consisted of one space (onboard Lunar Atmospheric Dust and Environment Explorer spacecraft) and three alternative ground terminals (the primary in White Sands, NM and two alternatives, one in JPL's Optical Communications Telescope Laboratory, California and one in Tenerife, Canary Islands) [12]. The operating wavelength of the system was 1.5 μm, supporting 4 uplinks and 16 downlinks at up to 20 and 622 Mbps, respectively. The LLST was comprised of one optical and two electronic modules (modem and control electronics). The optical module produced the ~15 μm divergence laser beam and was coupled through optical fibers with the modem to generate the downlink (at a power of 0.5 W) and process the uplink signal [12]. The main features of the White Sands station are four 15 cm telescopes for the uplink signal (at a power of 10 W) and four 40 cm telescopes for receiving the downlink signal all of which are

carried by a single gimbal. The overall LLCD project characterized as absolutely successful [12].

In contrast to traditional FSO links, i.e. symmetric, a more sophisticated approach is the asymmetric links or else indirect links, enabled through modulating retro-reflectors (MRR). In this case, instead of two active trancseivers the link is accomplished through a single laser beam which is passively retro-reflected back to the interrogator [13]. In such a link, the power of the laser, the beam divergence and PAT requirements are determined by the interrogator, whereas the MRR characteristics that affect the performance of the link are the gain of its optical antenna and the efficiency and bandwidth of the modulation [13]. It must be noted that an indirect link possesses twice as much free space losses, therefore the antenna optical gain plays a crucial role in overcoming this pitfall, since the MRR acts both as a receiver and a transmitter. Typically, an MRR is a corner cube prism, where the modulator can either cover the face of the corner cube or act as a reflector in one of the sides of the prism [13]. The major challenge for an MRR is the tradeoff between the size and the data rate capability, two parameters that are inherently inversely proportional, yet this technology would allow an FSO system to meet the size, weight and power (SWaP) requirements of small mobile platforms [13].

A key design issue for an FSO system is the wavelength of operation. The most common used spectra are the 780-850 nm (near-infrared) and the 1520-1600 nm (short-wave-infrared) [14]. In terms of eye safety, the latter is much safer since above 1400 nm the radiation is absorbed before reaching the eye retina. This fact, allows FSO systems operating in the 1520-1600 nm range to have higher power with the same safety standards. On the other hand, in the Near-Infrared (NIR) spectrum, a variety of low-cost and high efficiency optoelectronic devices exist, whereas in the Short-Wave

Infrared (SWIR) spectrum, PIN or APD based InGaAs detectors are optimized to operate in 1550 nm [14]. However, this FSO design parameter is also highly dependent on the atmospheric transmission and has to be taken into consideration.

Apparently, the anticipated environment of operation for an FSO system will dictate the performance requirements of the link. As mentioned above, the two main issues for the availability of an optical link are the atmospheric attenuation and turbulence. A different approach that has been proven to be necessary for an all-weather and long range link, is the combination of a lasercomm with a high bandwidth directional RF system [15]. The Applied Physics Laboratory of the John Hopkins University, has been involved in numerous developments and demonstrations of lasercomm systems, from a laboratory environment to a complex system of systems level. Special focus has been given to hybrid RF/FSO systems which ultimately lead to the successful demonstration of the complex system of systems DARPA FOENEX program, whose goal was the development and demonstration of a multinode airborne hybrid RF/FSO communications network [15].

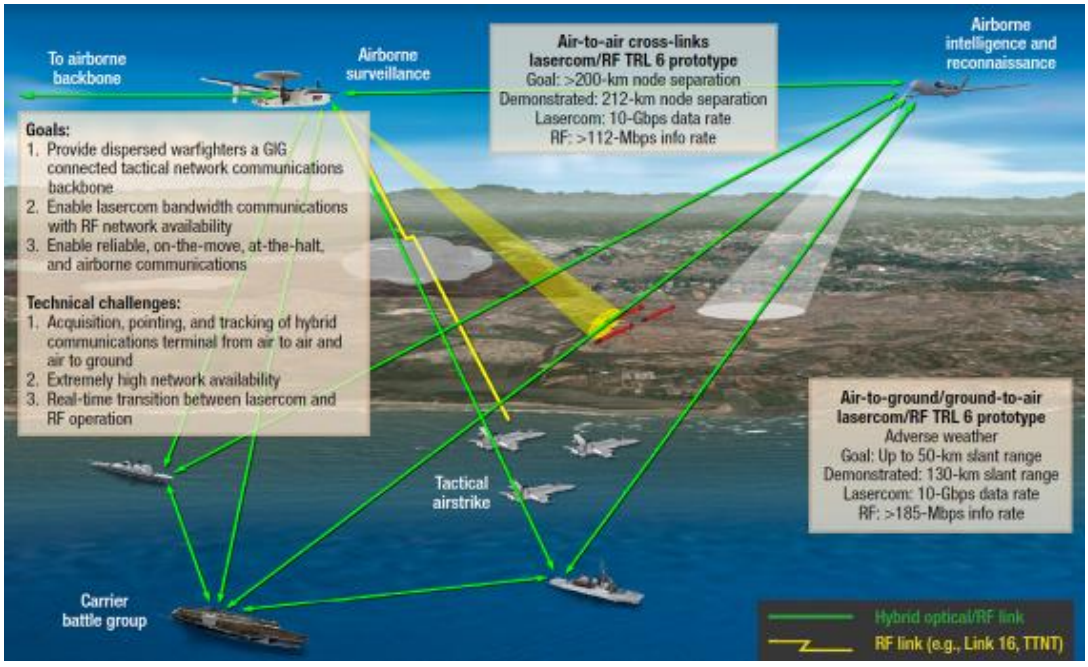


Figure 1.2. Overview chart of the DARPA FOENEX program [15].

The sum of the degradation effects, i.e turbulence and attenuation, can result in received signal fading for an FSO system, which are fluctuating randomly, since the propagating wave irradiance is also a randomly fluctuating value [16]. The reliability of the optical link, meaning the likelihood that the received signal is above a certain level, can be determined by using adequate statistical methods. The most common measure for the irradiance fluctuations is their probability density function (PDF) [16]. A joint PDF of the irradiance fluctuations and its derivative, can give us other fade issues, such as expected number of fades below threshold and their mean time [16]. Developing a closed form mathematical expression for the PDF, whose parameters will directly depend on the existing atmospheric conditions, that will be valid under any conditions, is highly desirable. Up to now, an extended number of PDFs exist, that can describe the irradiance fluctuations, in either strength regime of turbulence conditions, or, a few of them in any regime [16]. In any case, the PDF of the irradiance fluctuations is nonstationary and depends upon the turbulence parameters, the characteristics of the beam and the parameters of the receiver [17]. Since a reliable PDF is absolutely necessary for a robust FSO link, a thorough background of the existing PDFs will be given in a later chapter.

1.1.1 Motivation

Superiority in the modern warfare dictates the importance of information sharing, a concept that leads to the future network-centric operations, for example the access and distribution of sensor-based data is an eminent principle in such operations [18]. Equivalently, modern societies depend heavily on the available capacity of information exchange links in order to support the various “tele” requirements, either for working,

studying, buying, or even health purposes. The capabilities of the existing RF links seem to have reached their limits. On the other hand, the inherent capabilities of the FSOC technology, promise to fill this gap and act as a valuable complement to the RF networks [18]. The much higher operating frequencies of the FSO systems result in higher available bandwidth as compared to RF, the collimated laser beam has increased LPI and LPD properties as well as high protection against interference [19]. Therefore, it is crucial for these FSO benefits to be exploited, so as to get the money spent for high quality communications worth. However, several constraints also exist that can degrade the performance of an FSO link, with the atmospheric effects to be the most unpredictable and perhaps severe. For that reason, a major research area that has been identified is channel modeling, in order to quantify those effects on the performance of the link [2]. Special attention, should be given at the optical turbulence properties and their impact to the atmospheric laser propagation in the free space [20]. Therefore, it is indicated that further research in the link performance improvement through optical turbulence modeling is required, so as more efficient FSOC systems will be employed.

1.1.2 Historical Background

The historical roots of the FSO technology can be found at the end of the 19th century, when Alexander Graham Bell invented the so called Photophone, an instrument able to send optical signals to another, demonstrating the principle of optical communications [1]. An FSO system is a mean of communication which uses the free space as the transmission medium of optical signals in high data rates. The research on actual FSO systems started back in the 1960s' [2]. Three are the main temporal eras of the FSO experimental development, the first, spans from 60s to 90s where the laser invention resulted in the idea of optical communications, the second during 1990s,

where the first ground to space and vice versa experiments took place and finally after 2000, where numerous other applications of optical communications came into play and were demonstrated experimentally [1]. In particular, the invention of LASER in 1960 by Maiman, offered a great chance to enhance the communication capabilities by using beams propagating at the free space over long distances. The first FSO experiment, took place in 1968 in Rome, Italy, with a red laser between two buildings 4 km apart [1]. Another critical milestone in FSO development, was the invention of semiconductor light source in 1970 by Alferov. Between 1994 and 1996, the first symmetric ground to space link experiment was demonstrated in Tokyo, Japan, between the ETS-VI satellite and the Communications Research Laboratory (CRL) [1]. Another important date that changed the path of FSO development, was the 11th of September 2001, where the laser communication links, immediately substitute the demolished communication networks in New York, USA. As stated at the Deep Space Optical Communications Project of the National Aeronautics and Space Administration (NASA) its goal was "to develop key technologies for the implementation of a deep space optical transceiver and ground receiver that will enable greater than 10X the data rate of a state-of-the-art deep space RF system (Ka-band) for similar spacecraft mass and power" [20]. On the other hand, the European Space Agency (ESA) has started investing in laser communications in summer 1977 with an ultimately similar goal [2]. Notable projects that can be mentioned include: a) the Semiconductor Intersatellite Link Experiment (SILEX) in 2001, between GEO-LEO satellites and GEO – ground, b) the Geosynchronous Lightwave Technology Experiment (GeoLite) experiment between GEO satellites, ground and aircrafts and c) the Mars Laser Communication Demonstration (MLCD) for a deep space link from Earth to Mars [1]. Despite the fact that FSO technology indeed promised a new communications era, after several decades

of developmental efforts, the result did not meet the standards. However, the recent advances in the FSO optical and communication components, gave a new boost for FSO implementation in the wireless access networks [2].

1.1.3 Previous Work

Since the main goal of this dissertation is the investigation, characterization and modeling of the atmosphere to the laser beam propagation in the free space and consequently the effects towards an FSO system performance, a summary brief of previous experimental campaigns with similar objectives is presented.

The most comprehensive work has been done by the US Naval Research Laboratory (NRL), which included both electro-optical components development and atmospheric characterization, for more than fifteen years [21]. NRL has focused on the quantification of atmospheric effects on an FSO link in a maritime environment and for that reason a long range (16 km), maritime lasercomm test facility (LCTF) has been developed across Chesapeake Bay, Maryland, where transmission, scintillation index and angle-of-arrival measurements can be taken [22, 23]. With regard to link availability, a 100 Mbps FSO link with packet error rate measurements, assessed the long term availability of the link, in a reduced to 100 mW output power mode due to unattendance restrictions, resulted in 76% availability. This probability would have exceeded 90% if full output power and error correction protocols have been used [24]. With regard to atmospheric transmission measurements, a portable scatter-based transmission monitor, operating at the 1550 nm has been developed, to determine the local visibility, defined as the distance where the optical signal has been reduced by 5% of its initial value. A reasonable question is whether such a point measurement is representative of a long range path. For light to moderate rain this could be true,

whereas for heavy rain it depends on the size of the rain drop [25]. Numerous measurements under different weather conditions, time of year and location, have been executed to measure optical turbulence [26]. The Transportable Atmospheric Test System (TATS) was developed to measure scintillation through a variety of instruments and through analysis of the scintillation index, σ_I^2 and the power spectral density (PSD) of the scintillation [27]. An interesting result of those measurements, was that the diurnal scintillation variation of over water measurements, was less than over land but surprisingly very similar to a desert environment. A comparison of a 95% availability versus range for a 1550 nm direct FSO link in a maritime and a desert environment, showed that for the latter is much further, as shown in Figure 1.3. Statistical analysis of TATS measurements, showed a very good fit of the log-normal distribution for any environment.

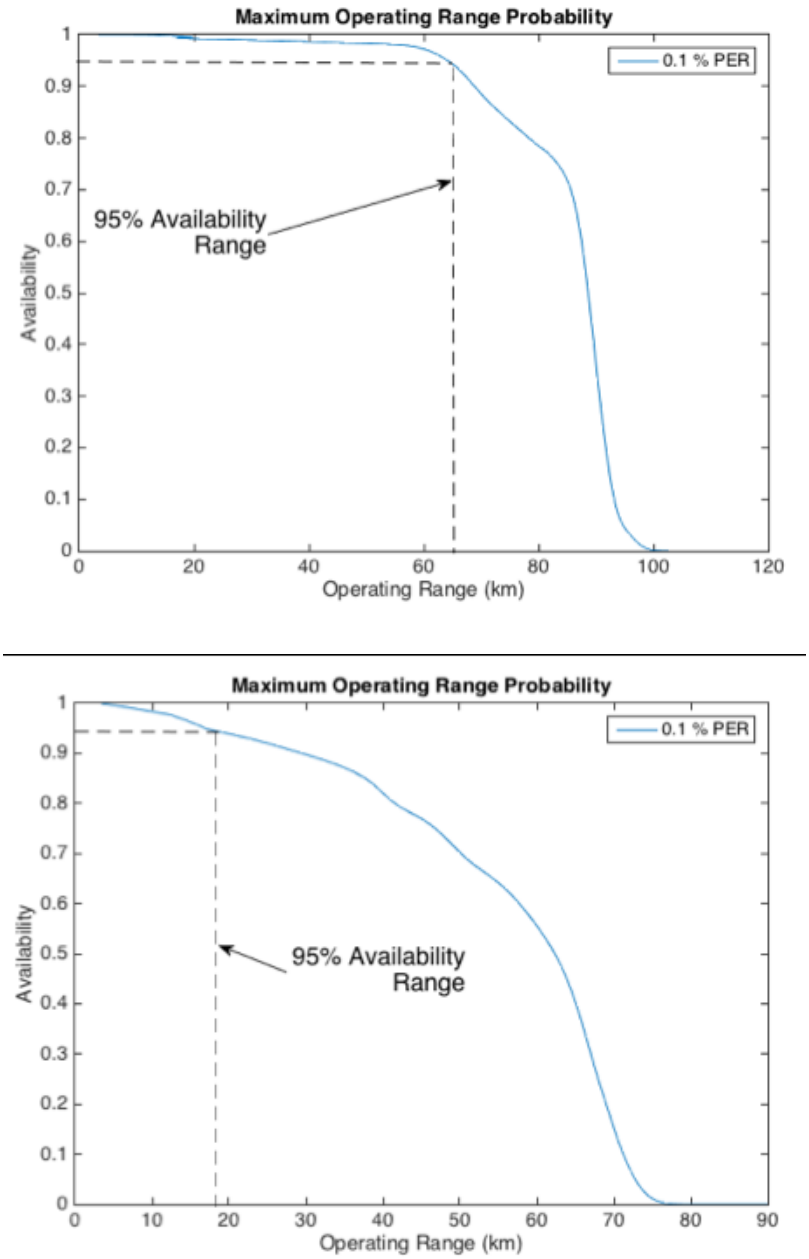


Figure 1.3. FSO link availability versus operating range for (upper) desert and (lower) maritime environment [24].

The fiber optics communications industry has already developed and improve a number of components that can also be adopted from the FSO community, such as laser and amplifiers. However, other components exist, that are either not useful or not appropriate for fiber optics. NRL have made a significant progress in filing that gap by developing new components or use COTS when they are available. The component with the biggest number of different possible options, is the photodetector, where many

alternatives can result in equivalent performance but with an impact to the other required components. For example, optical preamplifiers are used in the fiber optics communications, but in FSO either adaptive optics or adequate receiver aperture diameter is required. NRL, on the other hand, selected a larger area high performance InGaAs detector approach, based on market available components and eye safe. From mid-2000s, NRL has been collaborating with Optogration Inc. and University of Virginia to develop more sophisticated Avalanche Photodetectors (APD) with better noise performance and also APD arrays, that combine a larger focal plane than a single one, while keeping high bandwidth and sensitivity. Another area of great research advancements in NRL is MRRs, which allow SWaP restricted platforms to be able to transmit a lot of data [28, 29]. The core of an asymmetric program in NRL included two basic MQW MRR architectures. The first is a corner cube based retroreflector and the second a cat's eye approach [30]. Since the traditional squared quantum wells required a large applied voltage that limit the available devices for MRRs, NRL developed a new design based on strain balanced coupled quantum wells, that decreased the voltage requirements, from 20 V to 5 V, with an additional power consumption by a factor of 10.

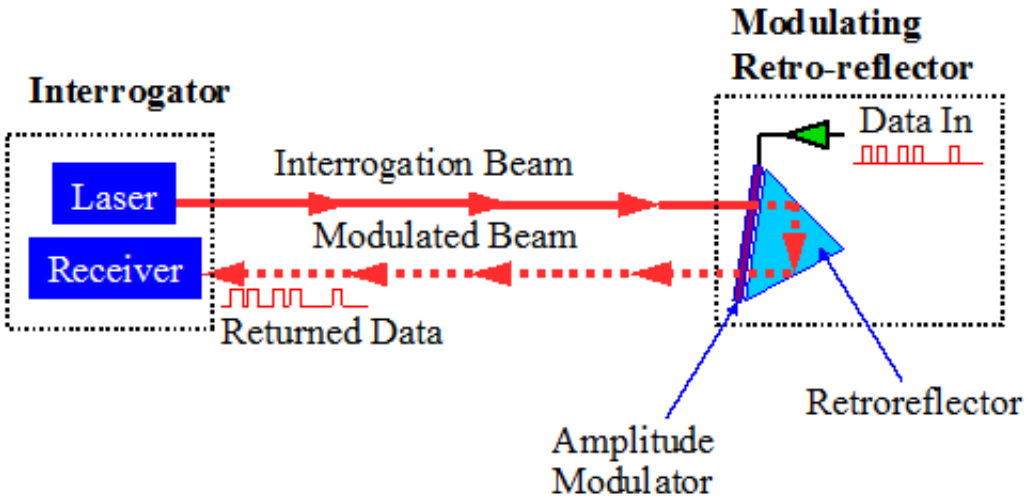


Figure 1.4. A diagram of a Modulating Retro-reflector [24].

Another representative long period experimental campaign that took place in False Bay, South Africa, called FATMOSE and included a variety of weather conditions [31]. The experimental setup was between the Institute of Marine Technology (IMT) and the National Sea Rescue Institute (NSRI) in Strandfontein, with total length of 15.7 km. Various sensors were mounted in both locations and in different heights above sea level. The campaign lasted for ten months and several atmospheric parameters were measured, such as wind speed, absolute humidity, visibility, air-sea temperature difference and air temperature and their effect on the performance of electro-optical systems were assessed. Of great practical importance is the result of the performance improvement that longer wavelengths provide in a laser beam propagation [31]. The geographical peculiarities presented a high variability in atmospheric parameters, such as the air – sea temperature difference which cause variations of C_n^2 along the optical path and visibility due to local fog presence. All these environmental parameters were used for several model proposals, such as an aerosol model based on Junge particle size distribution and the TARMOS model, for temperature and C_n^2 profile prediction, which was then validated by real time scintillation measurements [32]. Due to the long range of the path, the strong turbulence theory was used and showed a good agreement between measured and predicted SI values, which mainly appeared due to refraction effects large turbulence eddies [32].

Another maritime laser communications trial took place at Port Wakefield, Australia, executed by the Command, Control, Communications and Intelligence Division of the Australian Defense Science and Technology Organization during 2011 and demonstrated a novel analogue frequency modulated (FM) ship – to – shore laser communications system [33]. In close collaboration with NRL, which provided an MRR for trial needs, an optical communications link between a maritime platform and

the coast was established, as the precursor of a ship – to – ship optical link. A bidirectional optical link, transferred audio and video data at a distance of 3 km, using modulating retro-reflectors, for the first time [33]. The success of these trials, indicated the high potential of laser communications technology to be adopted for ship – to – ship communication purposes, offering additionally communication capabilities during emission controlled situations, alternative means of communication and adequate bandwidth capacity for network operations [33].

In [34] an experimental series of measurements, verified the classical theory of optical turbulence, by obtaining intensity data for a multi-wavelength laser beam propagating in a horizontal distance of 7 km near the ground. The refractive index structure parameter was evaluated for two wavelengths (λ_1=532 nm and λ_2=1064 nm) and then compared to theoretical predictions. The experimental setup included four fast-framing cameras synchronized by a software developed for that reason and the results showed a good agreement between measurements and predictions [34].

In [35, 36] a radio over FSO (RoFSO) system is presented and assessed under turbulent conditions. The experimental demonstration of this system included transmission for an extended period of time of different kind of services, such as 3GPP cellular, WLAN and terrestrial digital broadcasting TV (ISDB-T) signals at a distance of 1 km. In such a system, the RoFSO signal is directly optically amplified before emitted in the free space and the received optical signal is focused directly into a single mode fiber [36]. The operating wavelength of the RoFSO system is 1550 nm, with an output power of 100 Mw and a beacon beam for rough tracking operating at 850 nm.

A coastal-oriented experiment was held at the King Abdullah University of Science and Technology (KAUST) in Saudi Arabia, were the attenuation coefficient of an optical channel was modelled, based on atmospheric parameters such as temperature,

relative humidity and dew point, utilizing an outdoor FSO system operating at 1310 nm and 1550 nm at a distance of 70 m [37, 38]. This model was further validated against indoor laser link of the same operating wavelengths. The contribution of this work has to do with the illustration of the effects of temperature and humidity variations, which are prominent in a maritime environment, on the quality of an FSO channel. The key outcome of that analysis was that the 1550 nm wavelength is the least sensitive to humidity and temperature effects [37].

In [39], a successful demonstration of an error-free communication link between two FSO terminals 5.4 km apart developed by the MIT Lincoln Laboratory is presented, to simulate a low angle air-to-ground link with turbulence-induced effects mitigation capability, based on the presence of multiple small aperture receivers and a new encoding and interleaver hardware. The effectiveness of spatial diversity was demonstrated by the fading-induced received optical power probability density function width reduction by a factor of 2 [39].

1.2 Structure of this thesis

Regardless of the FSO communication application, the turbulent atmosphere is a major drawback that must be overcome in order to achieve a reliable optical link. Significant impact on this have the environment, the local atmospheric conditions, the geometry and the time of the day. In any case, the importance of quantifying those effects, has been identified as a paramount goal.

Many different experimental turbulence measurement methods exist, including predictive modelling of C_n^2 [39]. Using micrometeorology is one way to measure optical turbulence [40]. However, there are certain models based on local macro-meteorological parameters that can be obtained from point measurements. This method

offers high flexibility in terms of predictive ability, however it's accuracy might be limited because upper and lower bounds of those values exist, that constrain the validity area of the models.

An emerging topic for optical communications that will certainly gain more interest in the near future is Machine learning, which includes a broad spectrum of tools that allows interpretation and understanding of data through a trained algorithm. Thus it is expected to have much higher prediction accuracy, regardless of the terrain, location and time, with only prerequisite adequate training of a ML-based algorithm.

The reported work on this thesis includes both extensive regression modeling and ML-based training algorithms for FSO performance modeling on a maritime environment based on meteorological experimental data.

Section 2, provides the background theory of the optical turbulence phenomenon, including the fundamental physics beyond the impact of turbulence in the laser atmospheric propagation, the classical theory of optical turbulence and the structure functions. The differences between weak and strong fluctuation theory as well as the power spectrum models are also analyzed and a comprehensive report of various optical turbulence modeling experimental efforts is described.

Section 3, gives an introduction to the FSO performance metrics, assesses the atmospheric effects on the reliability and availability of an optical link and gives an overview of the channel modeling probability density functions. An extensive number of experimental campaigns for FSO performance modeling in different terrains and applications are presented.

Section 4, gives an introductory background of machine learning, machine learning algorithms and applications for optical communications modeling and fading mitigation.

In Section 5, two new mathematical models are proposed, to predict the received signal strength of an FSO optical link. The models have the form of a second-order polynomial with seven macroscopic meteorological parameters as the independent variables.

In Section 6, a machine learning based scheme was introduced to estimate the RSSI parameter of an FSO link over a maritime environment based on macroscopic meteorological measurements. To test the proposed approach, a large experimentally derived data set was used, which included seven parameters, i.e. air temperature, wind speed, solar flux, dew point, relative humidity, air pressure and air-sea temperature difference, obtained over a twelve-month period from a commercial FSO system and a weather station, respectively. Five popular ML algorithms were trained in order to construct a robust model to accurately predict the link's performance in terms of received signal strength.

Section 7 is comprised of two main parts, which present a thorough analysis of C_n^2 and FSO outage probability modeling, by leveraging machine learning algorithms. The first part of the chapter is two-fold. Initially the regression analysis results for C_n^2 is presented. We utilized six common ML algorithms and trained them on four different data sets. Then, a thorough analysis for the outage probability of an FSO link is executed. Initially, the corresponding P_{out} for the measured meteorological conditions was derived, based on an existing in the literature mathematical formula. These P_{out} estimations were used to classify the link status as functional or non-functional depending on a required availability of 99%, which corresponds to a 1% outage probability. We then trained a DNN classifier to model the status of the link based on the six measured meteorological parameters. Finally, an empirical mathematical model for outage probability estimation was developed based upon those meteorological

parameters and refractive index. The second part of this chapter, utilized four common ML algorithms and trained them on a preliminary data set consisting of six experimentally obtained macroscopic meteorological parameters.

In Section 8, we employ an information theoretical method, namely the so-called Jensen-Shannon divergence, a symmetrization of the Kullback-Leibler divergence, to measure the similarity between different probability distributions. In doing so, a large experimental dataset of received signal strength measurements from the HNA FSO link is utilized. Additionally, the Pearson family of continuous probability distributions is also employed to determine the best fit according to the mean, standard deviation, skewness and kurtosis of the modeled data.

Section 9, summarizes the key findings and concludes the thesis.

1.3 FSO Applications

Optical communications can be applied in a big variety of different scenarios, covering the airspace, terrestrial and underwater environment. In this section we provide a representative amount of such FSO applications for real world scenarios of high military and commercial interest.

1.3.1 Space Applications

The use of lasers in space is no longer something new. They are already being used for crosslinks between satellites offering order of magnitudes higher capacity with order of magnitudes less power. On the other hand, linking terrestrial or airspace moving platforms with geosynchronous satellites constitutes the future of FSO technology [41]. The space optical information network layers are composed of, 1)

satellite optical communication network, 2) ground optical communication relay stations, 3) ground microwave communication terminals and 4) point-to-point laser communications ground terminals [42].

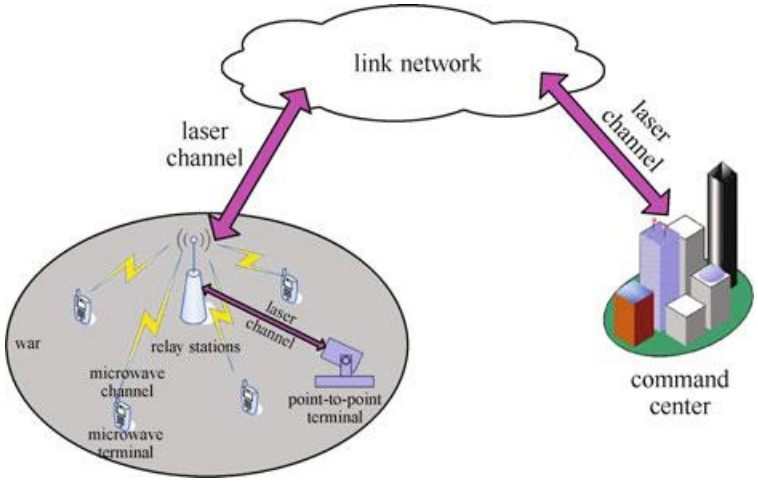


Figure 1.5. Notional architecture of an earth-space information network [42].

Unlike any uplink or downlink, the inter-satellite links are not limited by the weather conditions or the visibility, because they are far beyond the atmosphere. In this case though, the acquisition and tracking requirements are very high due to their relative movement [34]. Phase coherent techniques, i.e. homodyne or heterodyne are more suitable for this scenario in order to achieve the power requirements of the propagation distance. The highest published rate between two LEO satellites is a homodyne BPSK with 5.6 Gbps [34]. The main challenges of inter-satellite FSO network are the following [34]:

a. Point-ahead-angle: Refers to the required offset of the returned signal due to the relative motion of the terminals, called point-ahead-angle (PAA). PAA is in the order of hundreds of micro-radians for deep space links and tens of micro-radians for inter-satellite or ground to space links. If PAA is larger than the isoplanatic angle from the tracking direction, given by [34],

$$\theta_0^{-5/3} = 2.91k^2 \int_0^L h^{5/3} C_n^2(h) dh \quad (1.6)$$

we observe an effect called point ahead angular anisoplanatism.

b. Doppler shift: The frequency changes in the received signal due to the relative motion leads to Doppler effect. The amount of shift that has to be compensated in data relay systems is around 7.5 GHz which can be even more for LEO to LEO links, which can result in a 140 kHz frequency shift in a 2 GHz clock used for signaling. An optical phase lock loop or optical injection locking technique can reduce the Doppler effect.

c. Satellite vibration and tracking: A very challenging task is the acquisition and tracking of the received signal due to disturbances caused either by the satellite or laser communication assembly. External disturbances caused by the satellite include, solar panels, momentum wheels, gimbal packages and thrusters and different noise sources of the lasercomm assembly include, relative intensity noise, thermal noise, dark current shot noise, signal shot noise and background shot noise. The combination of all the aforementioned disturbances can result in misalignment between the terminals.

d. Background noise sources: These types of noises depend on the detection technique and whether optical pre-amplification is used or not. In case of direct detection, this noise is caused by the detector, the receiver amplifier and shot noise from the signal itself. For coherent detection, the major contributor is the local oscillator.

The overall comparison among RF and optical systems in space, has indicated that each system performs better under certain conditions provided its design is suitable for the system's mission. As an example, let's consider a Ka-band, where for frequencies from 32 to 38 GHz the typical bandwidth is around 500 MHz, whereas for

a 1.5 μm laser beam the bandwidth is more than 1000 times bigger [43]. At the same time however, the small divergence of the laser beam increases significantly the pointing and tracking requirements for a satellite crosslink, the EM energy on the laser spectrum is more sensitive to atmospheric effects and suffers from higher free space loss as a function of the base-ten logarithm $\log_{10}(f)$. Considering the beam divergence comparison between RF and laser systems is given by the following [41],

$$\theta \approx \frac{wavelength}{aperture\ diameter} \approx \frac{\lambda}{D_r} \tag{1.7}$$

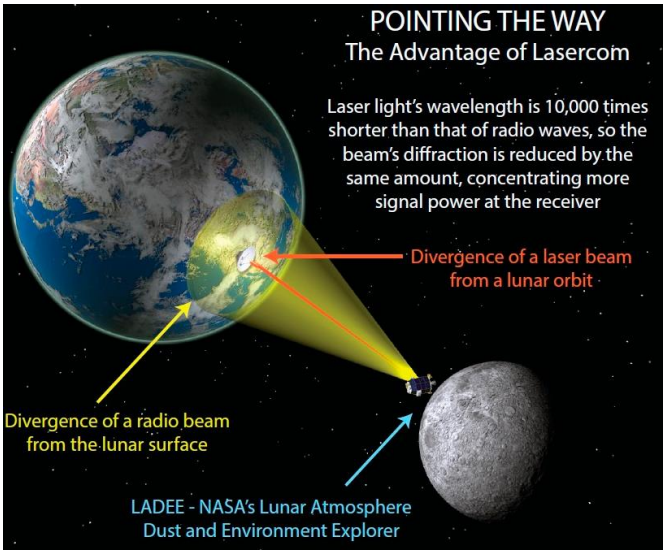


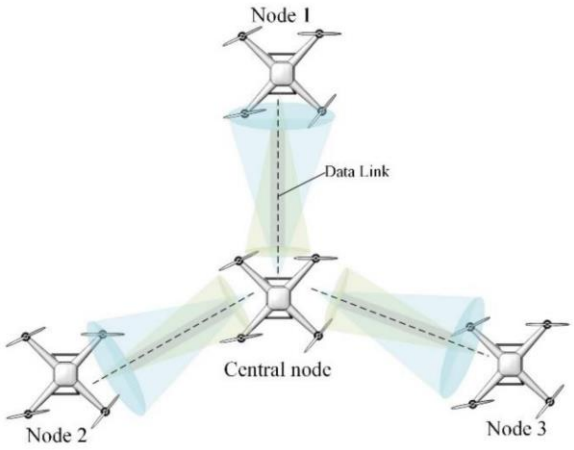
Figure 1.6. A comparison between RF and optical systems divergence [41].

The small divergence of a laser beam, apart from reducing the signal loss, provides a higher degree of security in the data transfer and immunity against interference and jamming. Additional security level can be provided by using quantum key distribution, which uses individual light quanta in quantum superposition states in order to assure security of communication between two terminals. Despite being achievable to a few hundreds of kilometers in terrestrial applications, in space environment this is not the case [41]. Already, several research projects have demonstrated satellite links utilizing QKD [44 – 46].

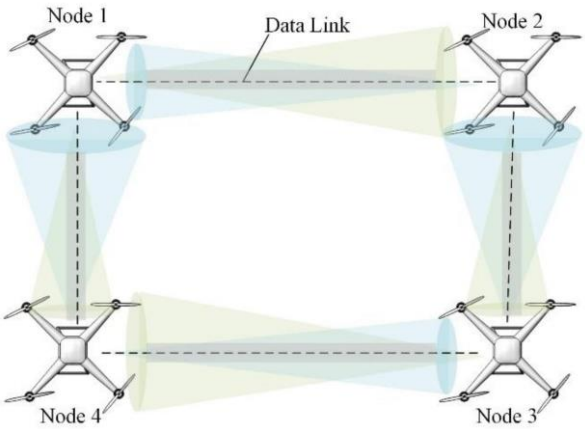
1.3.2 Unmanned Air Systems Applications

The future of information networks calls for the integration of space and terrestrial networks. The development of such a network in a cost-effective manner and with high enough capacity, makes the involvement of UAVs inevitable. A UAV often needs to interact with other assets in speeds that traditional RF links cannot support, such as emergency cases. Given that the RF links are subject to interference, capacity, security and spectrum availability issues and the space to ground links are subject to high cost, time delay and interference issues, the air to ground UAV links can fill this gap [47]. The integration of laser communications and UAV platforms is focused to UAV to satellite terminals, UAVs, UAV to ground terminals, and UAV to sea surface terminals.

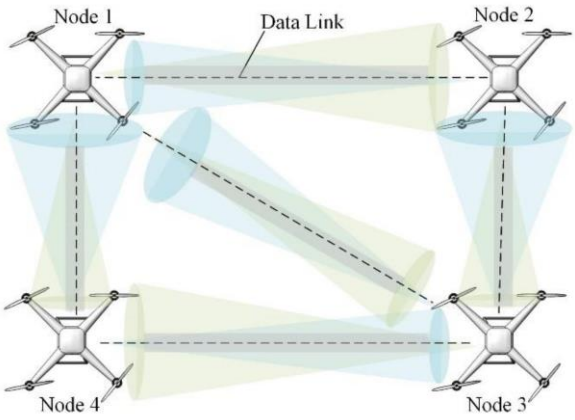
The terrestrial 5G networks are usually constrained in terms of coverage, especially in cases such as rural areas, post disaster situations, and maritime communications. Satellites and drones can provide the solution in the next network generation [47]. Lately, there is a lot of ongoing research on how to develop reliable optical communication links between a high-altitude platform and a satellite, by solving the various atmospheric effects and bridging the gap between air, space, ground and sea platforms [47]. Another issue that was identified, was the inability of a single UAV to perform several tasks independently, thus the improvement of the cooperation among multiple UAVs is of paramount importance. A technology that strongly supports the laser communication link between UAVs at the hardware level, is the three dimensional (3D) printing [47]. For higher reliability, a star and ring combination networks (Figure 1.7) provides the best solution.



(a)



(b)



(c)

Figure 1.7. UAV networks: (a) star architecture; (b) ring architecture; and (c) meshed architecture [47].

The UAVs can also provide a very effective alternative to ground communications infrastructure in case of a link outage or lack of performance (Figure 1.8) and with limited pointing and SWaP requirements due to the MRR technology [48]. Finally, a crucial environment that could be covered better in terms of broadband communications, is the ocean.

In [49] a drone-based communication network is considered, consisted of many drones loitering above a certain area to provide mobile remote radios. The geometric loss caused by the position and orientation fluctuations of the drones is quantified. Upper and lower bounds that correspond to approximate expressions, and a closed-form statistical model for the geometric loss, are derived and validated through simulations [49]. Finally, a very interesting application of UAVs that is of great importance for FSO technology, is turbulence measurements in the atmospheric boundary layer [50].

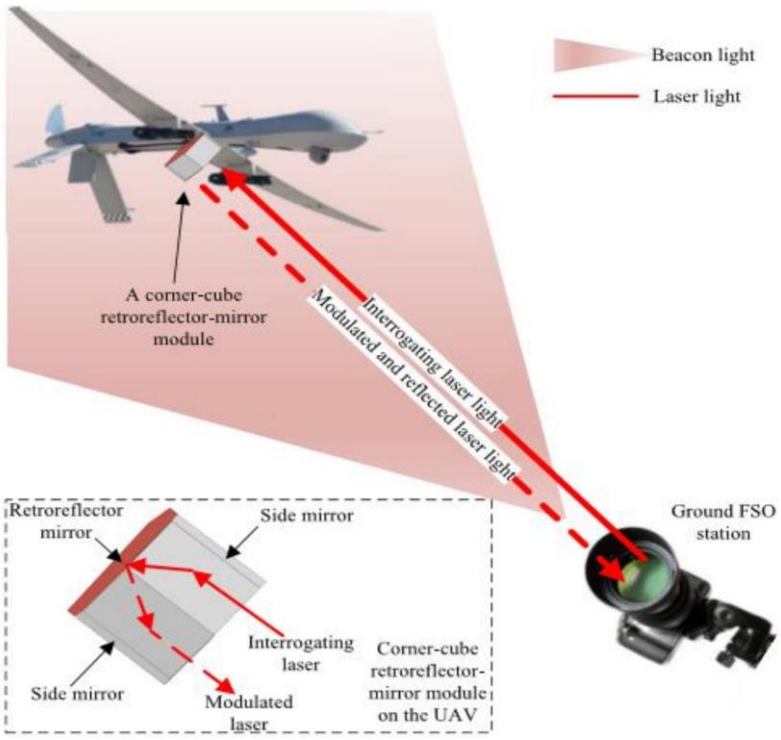


Figure 1.8. A schematic of a hybrid ground ATP mechanism and an MRR mounted in a UAV [47].

1.3.3 Underwater Systems Applications

Apart from the more noted space, air and terrestrial applications of optical communications, the underwater communications is also worth to mention, despite being the more challenging because of the numerous different physical processes that exist in the maritime environment [51]. The current technology for underwater

communications uses acoustic waves that suffer from limited bandwidth, high transmission losses, Doppler spread and high latency [51]. Ultimately, these factors limit the available data rates in a few tens of kbps for long ranges and a few hundreds of kbps for short ranges. For mobile underwater platforms that require higher bandwidth, the underwater optical communications offer an alternative with high potential. On the other hand, they are significantly affected by the water temperature fluctuations, dispersion, scattering and beam steering [51]. A transparency window exists though, that allow much better transmission conditions and lies in the blue-green wavelengths (450 – 550 nm). In general, the absorption of an optical beam underwater is higher for higher wavelengths. The possible alternative architectures for an underwater optical wireless communication network (UOWC) can be categorized with respect to their spatial coverage, the mobility of the sensor nodes and the channel [52]. Classification of underwater optical wireless applications are presented in Figure 1.10.

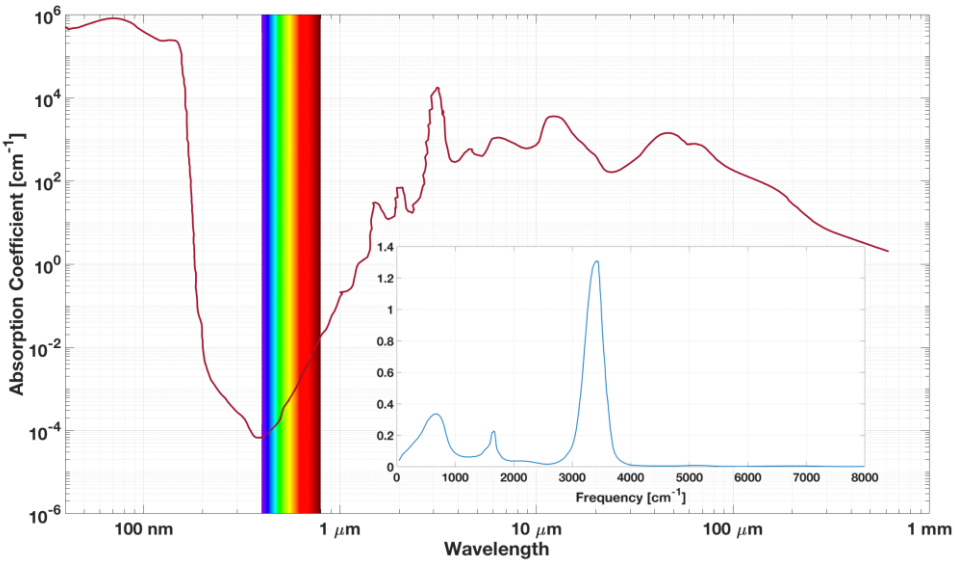


Figure 1.9. Absorption coefficient for optical waves in aquatic medium [52].

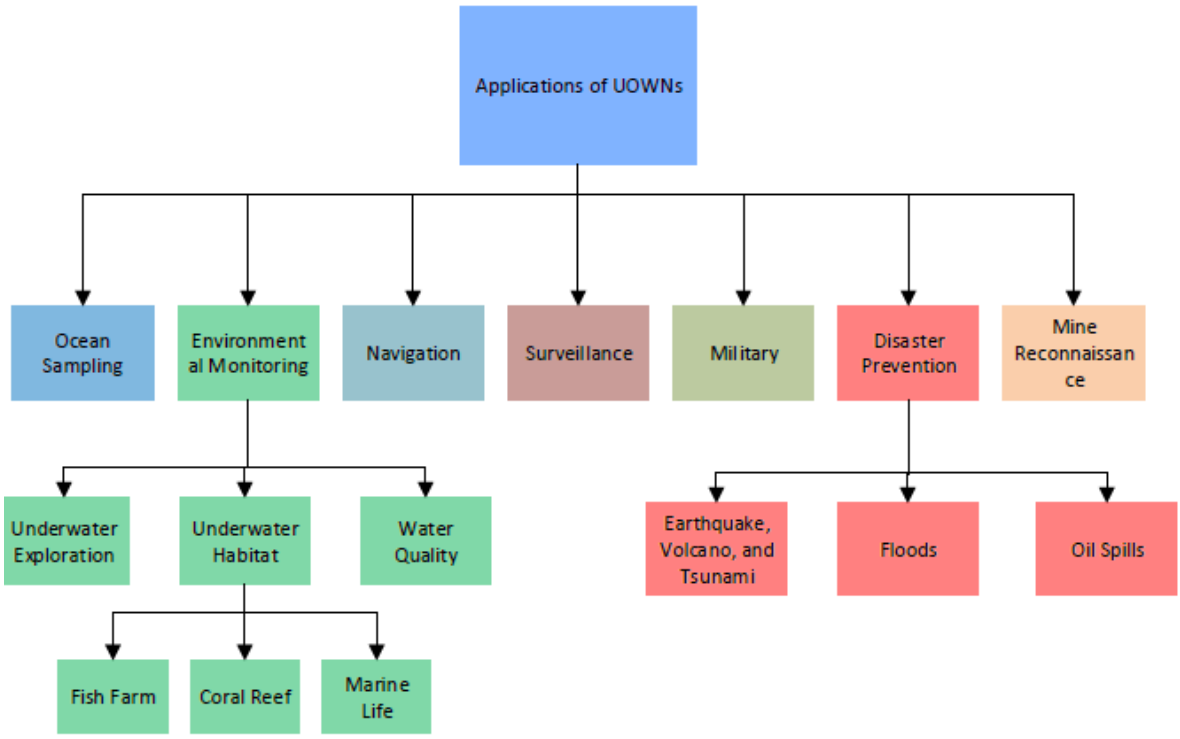


Figure 1.10. Underwater Optical Wireless Network applications [52].

Towards the building of the future smart cities, underwater optical wireless networks can play a critical role in the support of the new concept of Internet of Underwater Things (IoUT), defined as "the network of smart interconnected underwater objects" [53]. The most probable applications of IoUT can be classified in either of the following five types: (1) environmental monitoring, (2) disaster prevention, (3) military, (4) underwater exploration and (5) others. The main challenges that these applications face are related to the following: (1) the propagation speed, (2) the transmission range, media and rate, (3) the difficulty to recharge a sensor, (4) the mobility and the (5) reliability [53].

A way to improve the performance of UOWC is the use of optical pre-amplification in single-input single-output (SISO) systems that decreases the sensitivity of the receiver. In spite of the existing mismatch between the operating wavelength of the current semiconductor optical amplifiers (SOA) and the optimum spectrum for underwater transmission, the recent development of Gallium Nitride (GaN) – based optoelectronic components, promises to fix it, since they are able to operate in a wide

spectral range. Combining optical pre-amplification with spatial diversity can improve even further the performance of UOWC [54].

CHAPTER 2

Optical Turbulence Theory

The scope of this chapter is to understand and model the atmospheric effects on the performance of a free space optical communication link over a maritime environment. Special attention has been given in the phenomenon of optical turbulence. Thus, this chapter is devoted to conceptually and mathematically describe turbulence in fluids and in particular in the atmosphere. Both laminar and turbulent flow states for any fluid have been extensively researched since 1950s, where the impact of their physical properties to the way they flow was verified. The main attributes of a fluid are the flow velocity, the fluid viscosity and the scale length of the flowing field [39].

The turbulence in the atmosphere is generated by temperature variations that cause various effects on optical waves. In the case where the Earth's surface is hotter than the air, it also causes the lower air masses to be hotter than the higher ones, resulting in a negative temperature gradient that causes a light beam parallel to the surface to bend upwards, forming the so called phenomenon of *mirage* [55]. The opposite occurs when the Earth surface is colder than the air and results in the phenomenon of *looming* which can allow to see objects, like stars in the sky, even beyond the horizon. The temperature gradient leads to changes in the refractive index in the atmosphere known as *optical turbulence*. The propagation of an optical or IR wave through the turbulent atmosphere can be described using statistics.

2.1 Introduction

The disturbances of the air refractive index caused by the temperature gradients result in the formation of cells, called optical turbules. The random fluctuations of the refractive index cause several effects on a travelling optical wave, such as temporal, spatial and phase fluctuations of the irradiance [55]. Two very important concepts that are assumed to mathematically describe a turbulent medium is the statistical *homogeneity* and *isotropy*. *Homogeneity* (uniformity along the entire optical path), refers to a constant mean value of a random field and that any random fluctuations are irrelevant with the observed points and depend only on their vector separation. *Isotropy* (independence of direction), means that correlations between different points depend only on the magnitude of the vector separation between them [55].

As the air is also considered a fluid, it possesses the phenomenon of turbulence and follows the laws of fluid dynamics, where two distinct states of motion exist, a laminar and a turbulent. The difference between them is that the velocity flow characteristics in the former are uniform or change in a regular way, whereas in the latter this is not the case, acquiring random subflows, called *turbulent eddies*. A non-dimensional quantity exists, namely the *Reynolds number*, defined as following [55],

$$Re = Vl/v \tag{2.1}$$

where V the characteristic velocity, l the flow dimension and v the kinematic viscosity. In general, near the ground the flow is expected to be highly turbulent. A certain value of this number defines the transition from one state to the other. The mathematical background that governs the highly nonlinear phenomenon of turbulence, is the Navier-Stokes equations. However, because the solution of these equations is very hard, Kolmogorov developed his own theory for turbulence, based on statistics and dimensional analysis and provided a much more simplified approach for turbulence analysis with few approximations [55].

To better understand the formation of atmospheric turbulence we shall adopt the energy cascade theory of turbulence, developed by Kolmogorov. According to this theory, the increase of wind speed leads to the overcome of the critical value for the Reynolds number and this results in the formation of the *turbulent eddies*. Due to inertial forces, these initial larger turbulent eddies, with size equal to the outer scale size, L_0, break up into smaller ones, until their size reaches the inner scale value, l_0 [4]. The range between the outer and the inner scale is called the *inertial subrange*. The outer scale is about a few tens of meters and increases proportionally with height, whereas the inner scale is of the order of a few millimeters near the ground. Eddies with size smaller than the inner scale dissipate in the form of heat. Figure 11, depicts the energy cascade theory, developed by Richardson, for the outer and inner scale of turbulent eddies [4].

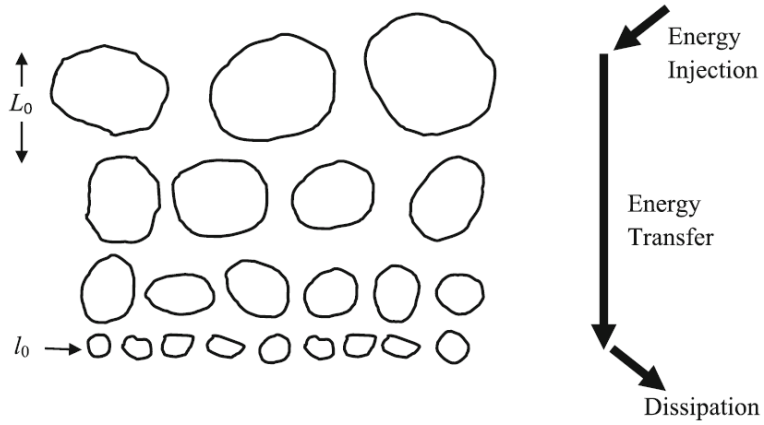


Figure 2.1. Energy Cascade model by Richardson, from the outer, L_0, to the inner, l_0, turbulent eddies size [4].

The size of the inner scale, l_0 and the velocity of these cells, u_0, determine the rate of heat dissipation into the environment, ϵ [55],

$$\epsilon \approx \frac{v u_0^2}{l_0^2} \tag{2.2}$$

Equation (2.2) can be modified in order to apply in any scale size and the transferred energy to the immediately smaller scale size [55],

$$\epsilon \approx \frac{u_l^3}{l} = \frac{u_0^3}{l_0} \tag{2.3}$$

The energy dissipation rate, ϵ, and the inner scale size, l, are the only contributors to the velocity and temperature fluctuations within the turbulent field. Combining Equations (2.2) and (2.3) and eliminating u_0, we can extract a theoretical value for the inner scale [55],

$$l_0 \approx \eta \approx \sqrt[4]{\frac{\nu^3}{\epsilon}} \tag{2.4}$$

The inverse relation between the inner scale size and the energy dissipation rate designates that the stronger the turbulence, the smaller the inner scales. The most critical parameter that affects the optical wave propagation is the refractive index of the air defined as [55],

$$n(\vec{r}) = n_0 + n_1(\vec{r}) \tag{2.5}$$

where $\vec{r}$ a random point in space, n_0 the mean value of refractive index, equal to one, and $n_1(\vec{r})$ the random deviation of $n(\vec{r})$. Therefore $\langle n_1(\vec{r}) = 0 \rangle$ and Eq. (2.5) can be rewritten as [55],

$$n(\vec{r}) = 1 + n_1(\vec{r}) \tag{2.6}$$

The index of refraction is directly related to the pressure and the temperature of the atmosphere [4],

$$n(r) = 1 + 7.66 \times 10^{-6}(1 + 7.52 \times 10^{-3}\lambda^{-2})\frac{P'(r)}{T'(r)}$$

$$\cong 1 + 79 \times 10^{-6}\left(\frac{P'(r)}{T'(r)}\right) \tag{2.7}$$

where λ the wavelength in μm, *P* the atmospheric pressure in mbar and *T* the air temperature in Kelvin. The humidity fluctuations contribute only in the far-IR region, whereas the pressure fluctuations are usually negligible [56].

2.2 Kolmogorov Theory of Turbulence

Kolmogorov was the first who realized in 1941 that he could explain the optical turbules from dimensional and reliable heuristic arguments. He managed to notice that this analysis resulted in universal laws in the statistical sense. The turbulent velocity field should be considered as a stochastic variable in the overall sense of statistical mechanics. It is independent of how the turbulence is initialized and the way the energy is injected is out of scope. Therefore, the chosen force statistics has no effect on the statistics of the turbulence [57].

The law that governs the dynamics of the fluids is the Navier-Stokes equation that is similar to the second law of Newton and the law for the conservation of the momentum. This equation is given by Eq. (2.8) and it can be modified depending on the type of the fluid and the flow conditions [57],

$$\frac{\partial}{\partial t}\mathbf{v} + (\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} - v\Delta\mathbf{v} = \frac{1}{\rho}(\boldsymbol{F} - \nabla p) \tag{2.8}$$

where **v** (**r**, t) is the velocity field, *v* the kinematic viscosity (with dimensions [*v*] = L^2/T), ρ the density, p the pressure and $\boldsymbol{F}$ the external force. It is important to note that this equation possesses the feature of scale invariance. So Eq. (2.8) can be converted into the following dimensionless [57],

$$\frac{\partial}{\partial \tilde{t}}\tilde{\boldsymbol{v}} + \left(\tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{\nabla}}\right)\tilde{\boldsymbol{v}} - \left(\frac{v\tau}{l^2}\right)\Delta\tilde{\boldsymbol{v}} = \frac{1}{\tilde{\rho}}(\tilde{\boldsymbol{F}} - \tilde{\nabla}\tilde{p}) \tag{2.9}$$

where l and τ the characteristic system length and time, respectively. The Reynolds number, which was introduced in Eq. (2.9), is a scale dependent parameter

and its value defines the flow type, either as laminar, for $R_e(l) << 1$, where a regular flow occurs, while for $R_e(l) > 10^4$ an obvious spatial disorder exists in the flow. For Reynolds number values that tend to infinity, we have a fully developed turbulent flow. Close to the ground the characteristic scale size is $l \sim 2$ m, the characteristic wind speed is from 1 to 5 m/s, and $v \sim 0.15 \times 10^{-4}$ m^2/s, leading to large Reynolds numbers on the order of $R_e \sim 10^5$. In such cases the motion is considered highly turbulent. Apparently, a turbulent flow designates that the inertial forces have outweighed the viscous forces.

2.3 Structure Functions

The classical theory of turbulence, as developed by Kolmogorov, predicts that the magnitude and direction of the velocity field of a fluid, fluctuates randomly. His theory assumed that for large Reynolds numbers, the small scale structure of turbulence, is homogeneous and isotropic, as well as independent of the large-scale structure. Additionally, the motion of the small-scale is determined by the kinematic viscosity, v, and the energy rate dissipation, ϵ, of the turbulent energy per unit mass of the fluid.

Using dimensional analysis, Kolmogorov proved that the longitudinal structure function of the wind velocity within the inertial range (parallel to **r** vector that connects two points in space), satisfies the universal two thirds power law [57]:

$$D_{RR}(R) = \langle (V_1 - V_2)^2 \rangle = C_V^2 R^{2/3}, \; l_0 \ll R \ll L_0 \tag{2.10}$$

where V_1 and V_2 the velocity components of two points in a distance of R and C_V^2 the *velocity structure function*, which gives the total amount of energy in the turbulence. The velocity structure function and the energy dissipation rate are related by [57],

$$C_V^2 = 2\epsilon^{2/3} \tag{2.11}$$

In case of $R \ll l_0$, there is a quadratic relation of the distance *R* with the structure function, ant the proportionality constant is such that the two-thirds power law exists for $R = l_0$ [57],

$$D_{RR}(R) = \begin{cases} C_V^2 l_0^{-4/3} R^2, \ R \ll l_0 \\ C_V^2 R^{2/3}, \ l_0 \ll R \ll L_0 \end{cases} \tag{2.12}$$

The intersection of the R^2 and $R^{2/3}$ curves define the inner scale, l_0. No similar description for $R \gg L_0$ exists, because the assumption of isotropy is not valid in that region. The 2/3 power law behavior of the structure function in the inertial range is equivalent to the power spectrum in three dimensions given by [57],

$$\Phi_{RR}(\kappa) = 0.0066\varepsilon^{2/3}\kappa^{-11/3} =$$

$$= 0.0033 C_V^2 \kappa^{-11/3}, 1/L_0 \ll \kappa \ll 1/l_0 \tag{2.13}$$

where κ is the scalar spatial frequency (rad/m). We can observe that the power spectrum exhibits a -11/3 power law, corresponding to a one-dimensional spectrum with a -5/3 power law. The power laws validity for the structure function and the power spectrum have been demonstrated over a wide range of experiments. However, a number of works suggest a modest change in the power law behavior of the structure function, which may have little effect on second order statistical quantities, but may be important in higher-order statistics.

Traditionally, the basic ideas about turbulence have been developed concerning velocity fluctuations. However, Kolmogorov's fundamental ideas have been also applied to passive scalar physical quantities, such as potential temperature, which does not exchange energy with the velocity temperature. An equivalent inner and outer size scale form the boundaries of the inertial range, which in case of temperature fluctuations, is the *inertial-convective range*. Additionally, the analogue to the viscosity as the dissipation mechanism for velocity fluctuations, is the molecular diffusion. By

applying the Kolmogorov theory towards temperature fluctuations, yet assuming homogeneity and isotropy, we take the same two-third power law [57],

$$D_T(R) = \langle (T_1 - T_2)^2 \rangle = \begin{cases} C_T^2 l_0^{-4/3} R^2, & R \ll l_0 \\ C_T^2 R^{2/3}, & l_0 \ll R \ll L_0 \end{cases} \quad (2.14)$$

where T_1 and T_2 the temperature at two points separated by R and C_T^2 the temperature structure function (in $deg^2/m^{2/3}$). The inner scale for temperature fluctuations is given by [57],

$$l_0 = 5.8(D^3/\varepsilon)^{1/4} \quad (2.15)$$

where *D*, the diffusivity of heat in air (in m^2/s). The 3-D spectrum of temperature fluctuations takes the -11/3 power law form [57],

$$\Phi_T(\kappa) = \frac{1}{4\pi}\beta\chi\varepsilon^{-1/3}\kappa^{-11/3} =$$

$$= 0.0033 C_T^2 \kappa^{-11/3}, 1/L_0 \ll \kappa \ll 1/l_0 \quad (2.16)$$

where β the Obukhov – Corrsin constant and χ the rate of dissipation of the mean-squared temperature fluctuations. Hill and Clifford have noted that there is a small "bump" in the temperature spectrum, close to $1/l_0$, which can play a significant role in various quantities related to optical wave propagation.

The statistical description of the random field of the atmospheric refractive index caused by turbulence fluctuations, is similar to those caused by the velocity fluctuations. Again, an equivalent inertial subrange exists, bounded by an upper and a lower limit, where in the statistical homogeneity and isotropy are also valid assumptions. Because $\langle n_1(\boldsymbol{R}) \rangle = 0$, the covariance function of $n(\boldsymbol{R})$ can be expressed as [57],

$$B_n(\boldsymbol{R_1}, \boldsymbol{R_2}) \equiv B_n(\boldsymbol{R_1}, \boldsymbol{R_1} + \boldsymbol{R}) = \langle n_1(\boldsymbol{R_1}) n_1(\boldsymbol{R_1} + \boldsymbol{R}) \rangle$$

which is a function of $\mathbf{R} = \mathbf{R_1} - \mathbf{R_2}$, for homogeneous field. In case isotropy also exists, the covariance function reduces to $R = |\boldsymbol{R_1} - \boldsymbol{R_2}|^2$. The refractive index structure function exhibits the asymptotic behavior [57],

$$D_n(R) = 2[B_n(0) - B_n(R)] = \begin{cases} C_n^2 l_0^{-4/3} R^2, & R \ll l_0 \\ C_n^2 R^{2/3}, & l_0 \ll R \ll L_0 \end{cases} \tag{2.17}$$

where C_n^2 the *refractive index structure parameter* (in $m^{-2/3}$), whereas the inner scale [57],

$$l_0 = 7.4\eta = 7.4(v^3/\varepsilon)^{1/4} \tag{2.18}$$

From a physical point of view, C_n^2 is a measure of the strength of the refractive index fluctuations and is directly related to the temperature structure constant, C_T^2[58],

$$C_n^2 = \left(79 \times 10^{-6} \frac{P}{T^2}\right)^2 C_T^2 \tag{2.19}$$

2.4 Power Spectrum Models for Refractive Index Fluctuations

Once again we assume a homogeneous and isotropic turbulent field, where the energy is constant. Therefore, when we consider the turbulent energy transportation, this will be in wavenumber, $\boldsymbol{\kappa}$, rather than in the coordinate space. So the spatial power spectral density of refractive index fluctuations, $\Phi_n(\kappa)$, is related with the covariance function by the three dimensional Fourier transform [57],

$$\Phi_n(\kappa) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} B_n(R) exp[-i\boldsymbol{\kappa} \cdot \boldsymbol{R}] d^3R =$$

$$= \frac{1}{2\pi^2 \kappa} \int_0^{\infty} B_n(R) \sin(\kappa R)\, R dR \tag{2.20}$$

The last integral has been obtained by using the spherical symmetry and $\kappa = |\kappa|$ is the scalar wave number. Based on inverse Fourier transform properties, it follows that [57],

$$B_n(R) = \frac{4\pi}{R} \int_0^\infty \kappa \Phi_n(\kappa)\, sin(\kappa R)\, d\kappa \tag{2.21}$$

Therefore, the spectrum can be expressed in terms of the structure function as follows [57],

$$D_n(R) = 2[B_n(0) - B_n(R)] =$$

$$= 8\pi \int_0^\infty \kappa^2 \Phi_n(\kappa) \left(1 - \frac{sin(\kappa R)}{\kappa R}\right) d\kappa \tag{2.22}$$

The temperature variations in the atmosphere are the main cause for refractive index fluctuations. In other words, humidity and pressure variations can be ignored. It is then generally accepted, that the functional form of the spatial power spectrum for the refractive index is the same as is for the temperature and additionally that temperature fluctuations follow the spectral law that velocity fluctuations also do. Based on the two third power law for the structure functions, it can be inferred that the associated power spectral density for the refractive index fluctuations within the inertial subrange is defined by the *Kolmogorov power-law spectrum* [57],

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3}, 1/L_0 \ll \kappa \ll 1/l_0 \tag{2.23}$$

which is widely used in theoretical calculations due to its simplicity, even though it is valid only within the inertial subrange. The extension of its use outside the inertial subrange may lead to divergent integrals.

Whenever scale effects cannot be neglected, other proposed models exist to make the necessary calculations. In order to extend the power law beyond the inertial subrange, the introduction of a function that truncates the spectrum at high wave number is required. To do so, Tatarski introduced a Gaussian function leading to the spectrum model [57],

$$\Phi_n(\kappa) = 0.033C_n^2\kappa^{-11/3}exp\left(-\frac{\kappa^2}{\kappa_m^2}\right), \kappa \gg \frac{1}{L_0};\ \kappa_m = 5.92/l_0 \qquad (2.24)$$

Equation (2.24) is better known as *Tatarski spectrum*, and before having been adopted for refractive index fluctuations, it has been developed by Novikov for velocity fluctuations. For κ = 0 there is a singularity for the limiting case $1/L_0 = 0$ which means that while the structure function can be calculated, the covariance function, cannot.

A third model also exists, known as *von Karman*, allowing wave numbers in the input range $\kappa < \frac{1}{L_0}$ [57],

$$\Phi_n(\kappa) = \frac{0.033C_n^2}{\left(\kappa^2+\kappa_0^2\right)^{-11/6}}, 0 \le \kappa \ll \frac{1}{l_0}, \kappa_0 = 2\pi/L_0 \qquad (2.25)$$

If we adequately modify the Tatarski model, it can be valid for any wave number and be called the *modified von Karman spectrum* [57],

$$\Phi_n(\kappa) = \frac{0.033C_n^2}{\left(\kappa^2+\kappa_0^2\right)^{-11/6}} exp\left(-\frac{\kappa^2}{\kappa_m^2}\right), 0 \le \kappa < \infty \qquad (2.26)$$

Within the inertial subrange, the Tatarksi and the von Karman spectra are reduced to the Kolmogorov spectrum and due to their mathematical convenience they are widely used.

Finally, another model valid outside the inertial subrange, is the *exponential spectrum* [57],

$$\Phi_n(\kappa) = 0.033C_n^2\kappa^{-11/3}\left[1 - exp\left(-\frac{\kappa^2}{\kappa_0^2}\right)\right], 0 \le \kappa \ll \frac{1}{l_0} \qquad (2.27)$$

where $\kappa_0 = C_0/L_0$, where the scaling constant C_0 depends on the application.

2.5 Free Space Optical Wave Propagation

The turbulent eddies, formed by the refractive-index variations, cause various distortions to an optical wave propagating through the atmosphere. There, energy is

injected to form the outer scale size eddies and thereinafter is transferred to smaller scale size until it is dissipated.

The random amplitude and phase variations of the travelling wave, can be theoretically treated by solving the wave equation for the electric field and its respective statistical moments, derived from the *stochastic Helmholtz equation* [58],

$$\nabla^2\vec{E} + k^2n^2(\vec{r})\vec{E} = 0 \tag{2.28}$$

where $k = 2\pi / \lambda$ is the wavenumber, $\vec{r}$ a random point in space and $n(\vec{r})$ as given in Eq. (2.6). The actual equation that traditionally needs to be solved is the *scalar* stochastic Helmholtz equation [58],

$$\nabla^2U + k^2n^2(\vec{r})U = 0 \tag{2.29}$$

which corresponds to either components of electric field. The Born and Rytov approximations have traditionally been used to solve Eq. (2.29). These solutions assume that backscattering and depolarization effects are neglected, the refractive-index is assumed uncorrelated in the propagating direction and we can use the paraxial approximation.

2.5.1 Weak and strong fluctuation conditions

Theoretical studies of optical wave propagation have been traditionally classified to belong to either weak or strong fluctuation theory. For using the Kolmogorov spectrum, we use the *Rytov variance* to distinguish these two cases [58],

$$\sigma_R^2 = 1.23C_n^2k^{7/6}L^{11/6} \tag{2.30}$$

where C_n^2 the refractive index structure parameter. For $\sigma_R^2 < 1$, fluctuations are considered as *weak* and the Rytov variance represents the irradiance fluctuations of an unbounded plane wave. For Rytov variance values more than unity, fluctuations are considered *strong* and for very high values $\sigma_R^2 \to \infty$, *saturation regime* is reached. For

a Gaussian-beam wave, this classification based solely on the Rytov variance is not adequate, but the scintillation index throughout the beam profile has to be less than unity. Therefore, for a Gaussian-beam wave, weak fluctuations require both conditions [55],

$$\sigma_R^2 < 1 \text{ and } \sigma_R^2 \Lambda^{5/6} < 1 \tag{2.31}$$

where $\Lambda = 2L/kW^2$ and W is the free space beam radius at the receiver. If either conditions do not hold, fluctuations are classified as moderate to strong.

2.5.2 Born Approximation

The most popular approaches towards the solution of Eq. (2.29) are the Born and Rytov perturbation methods. The difference between those two methods, is that the former *adds* whereas the latter *multiplies* the perturbation terms.

The square of the refractive index in Eq. (2.29) can be written [58],

$$n^2(\boldsymbol{r}) = [n_0 + n_1(\boldsymbol{r})]^2 \cong 1 + 2n_1(\boldsymbol{r}), \ \ |n_1(\boldsymbol{r})| \ll 1 \tag{2.32}$$

therefore, we can neglect the term $n_1(\boldsymbol{r})$ as compared to $n_1^2(\boldsymbol{r})$. For an optical wave travelling on the positive *z*-axis, we can express the optical field at *z* = L as [58],

$$U(r) = U_0(r) + U_1(r) + U_2(r) + \cdots, \tag{2.33}$$

where $U_0(r)$ is the unperturbed part of the field while the rest of the terms the first-order, second-order etc. scattering, caused by random inhomogeneities. Substituting Eq. (2.32) and (2.33) into Eq. (2.29) and equating terms of the same order we take [58],

$$\nabla^2 U_0 + k^2 U_0 = 0 \tag{2.34}$$

$$\nabla^2 U_1 + k^2 U_1 = -2k^2 n_1(r) U_0(r) \tag{2.35}$$

$$\nabla^2 U_2 + k^2 U_2 = -2k^2 n_1(r) U_1(r) \tag{2.36}$$

and so forth for higher order perturbations. The most significant advantage of this approach, is the transformation of Eq. (2.29) which contains random and space-dependent coefficients to a homogeneous equation and a system of non-homogeneous equations, all with constant coefficients [55].

2.5.3 Rytov Approximation

The Rytov approximation for the solution of Eq. (2.29) was initially used by Obukhov in a wave propagation problem in random media. Bounded only to the weak fluctuation conditions, this approach expresses the field of the electromagnetic wave as following [58],

$$U(\boldsymbol{r}) \equiv U(\boldsymbol{r}, L) = U_0(\boldsymbol{r}, L) exp[\psi(\boldsymbol{r}, L)] \tag{2.37}$$

where $\psi(\boldsymbol{r}, L)$ is a *complex phase perturbation* due to turbulence and is expressed as [58],

$$\psi(\boldsymbol{r}, L) = \psi_1(\boldsymbol{r}, L) + \psi_2(\boldsymbol{r}, L) + \cdots \tag{2.38}$$

where $\psi_i(\boldsymbol{r}, L)$ the ith-order phase perturbation. As known, the addition of these phase perturbations results in a multiplication of the exponential functions of Eq. (2.37).

These perturbation terms can be directly related to the previously calculated Born perturbations. We introduce the *normalized Born perturbations* defined by [57],

$$\Phi_m(\boldsymbol{r}, L) = \frac{U_m(\boldsymbol{r}, L)}{U_0(\boldsymbol{r}, L)}, \; m = 1, 2, 3, \ldots \tag{2.39}$$

We then equate the first-order perturbation terms of both approximations [57],

$$\begin{aligned} U_0(\boldsymbol{r}, L) exp[\psi_1(\boldsymbol{r}, L)] &= U_0(\boldsymbol{r}, L) + U_1(\boldsymbol{r}, L) \\ &= U_0(\boldsymbol{r}, L)[1 + \Phi_1(\boldsymbol{r}, L)] \end{aligned} \tag{2.40}$$

We then divide both terms by $U_0(\boldsymbol{r}, L)$ and take their natural logarithm [57],

$$\psi_1(\boldsymbol{r}, L) = ln[1 + \Phi_1(\boldsymbol{r}, L)] \cong \Phi_1(\boldsymbol{r}, L), \; |\Phi_1(\boldsymbol{r}, L)| \ll 1 \tag{2.41}$$

where

$$\Phi_1(\boldsymbol{r}, L) = \frac{U_1(\boldsymbol{r}, L)}{U_0(\boldsymbol{r}, L)} =$$

$$= \frac{k^2}{2\pi} \int_0^L dz \iint_{-\infty}^{\infty} d^2 s exp\left[ik(L-z) + \frac{ik|\boldsymbol{s}-\boldsymbol{r}|^2}{2(L-z)}\right] \frac{U_0(\boldsymbol{s},z)}{U_0(\boldsymbol{r},L)} \frac{n_1(\boldsymbol{s},z)}{(L-z)} \quad (2.42)$$

where $U_0(\boldsymbol{r}, L)$ represents the optical field in the receiver plane ($z = L$), whereas $U_0(\boldsymbol{s}, z)$ represent the optical field at an arbitrary point along the propagation path. By following the exact same procedure for up to second-order perturbation terms we take [57],

$$\psi_2(\boldsymbol{r}, L) = \Phi_2(\boldsymbol{r}, L) - \frac{1}{2}\Phi_1^2(\boldsymbol{r}, L) \quad (2.43)$$

where,

$$\Phi_2(\boldsymbol{r}, L) = \frac{U_2(\boldsymbol{r}, L)}{U_0(\boldsymbol{r}, L)} =$$

$$= \frac{k^2}{2\pi} \int_0^L dz \iint_{-\infty}^{\infty} d^2 s exp\left[ik(L-z) + \frac{ik|\boldsymbol{s}-\boldsymbol{r}|^2}{2(L-z)}\right] \frac{U_0(\boldsymbol{s},z)}{U_0(\boldsymbol{r},L)} \frac{\Phi_1(\boldsymbol{r},L) n_1(\boldsymbol{s},z)}{(L-z)} \quad (2.44)$$

The first-order perturbation of the Rytov approximation which is directly proportional to the first-order perturbation of the Born approximation, is called the *single scattering* approximation. By using this approach, we can calculate many statistical quantities of interest (log-amplitude variance, phase variance etc.), but not any statistical moments (i.e. mean value etc.), where the second-order perturbation is required.

2.6 Physical Effects

2.6.1 Angle-of-arrival fluctuations

In the presence of turbulence, the laser beam partially arrives on the receiver off-axis, due to deflection. These variations in the angle with respect to the optical axis of the receiver, represent the concept of angle-of-arrival fluctuations. These fluctuations,

which depend directly to the turbulence strength and optical path length, are given by [58],

$$\langle \beta_a^2 \rangle = 2.91 C_n^2 L (2W_G)^{-1/3} \quad (2.45)$$

Where W_G is the soft aperture radius and is related to the receiver diameter D [58],

$$D^2 = 8W_G^2 \quad (2.46)$$

In order to deal with this effect, a combination of fast steering mirrors and adaptive optics algorithms is required.

2.6.2 Beam Wander

In case of existence of turbulent eddies in the path of the optical beam whose size is bigger than the transmitter's beam size, then we have random deflection of the beam's original path, as a whole. This phenomenon is called beam wander and causes serious pointing errors towards the receiver [4]. It is mostly related to the instantaneous field of the received optical signal. On the other hand, the short and long term fields, present a closer to a perfect Gaussian field profile. A simulated comparison of these three cases, namely the instantaneous, the short and the long term, are presented in Figure 2.2, assuming a path length of L = 2 km, $C_n^2 = 0.6 \cdot 10^{-14} m^{-2/3}$, λ = 1064 nm and W_0 = 2 cm. The exposition time of the long term field is 34 times the short time. Fante, related the beam's centroid with the short-term, W_{ST}, and the long-term W_{LT}, spot size as following [59],

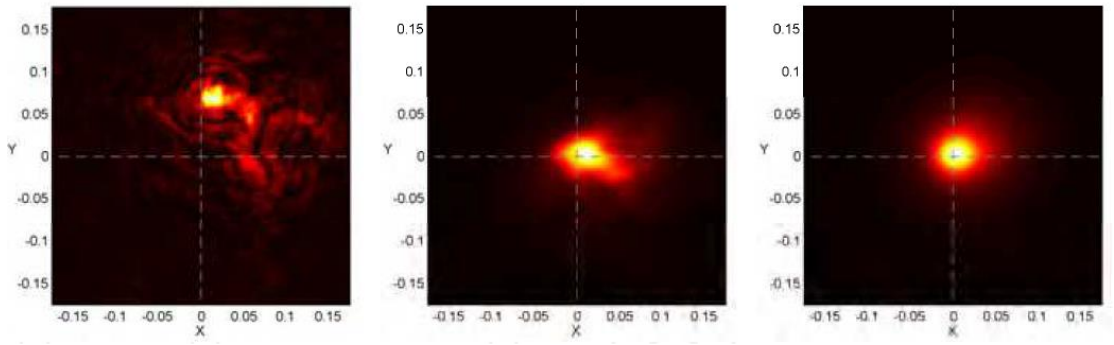
$$\langle r_c^2 \rangle = W_{LT}^2 - W_{ST}^2 \quad (2.47)$$

where

$$W_{LT}^2 = W^2 \left[1 + 1.33 \sigma_R^2 \Lambda^{5/6}\right], \ \Lambda = \frac{2L}{kW^2} \quad (2.48)$$

where W, the pure diffraction beam radius at the receiver plane. The short-term profile is given by,

$$W_{ST}^2 = W^2\left\{1 + 1.33\sigma_R^2\Lambda^{5/6}\left[1 - 0.66\left(\frac{\Lambda_0^2}{1+\Lambda_0^2}\right)^2\right]\right\} \quad (2.49)$$



(a) Instantaneous beam profile (b) Short time beam profile (c) Long time beam profile

Fig. 2.2. Different profiles of a propagating Gaussian beam [58].

2.6.3 Scintillation

The random fluctuations of the refractive index in the free space where an optical beam is propagating, cause temporal and spatial irradiance fluctuations at the receiver and this phenomenon is known as scintillation, a very serious factor that degrades the performance of an FSO system and many times its availability and reliability. In order to quantify the scintillation effect on the performance of an FSO system, we define several statistical quantities, the most common of which is the scintillation index [60],

$$\sigma_I^2 = \frac{\langle I^2\rangle - \langle I\rangle^2}{\langle I\rangle^2} = \frac{\langle I^2\rangle}{\langle I\rangle^2} - 1 \quad (2.50)$$

where I is the irradiance of the optical signal at the receiver and $\langle\cdot\rangle$ denotes the ensemble average. Classical studies on optical wave propagation, are classified into the strong and weak fluctuation theory. Commonly this is done using the value of the Rytov variance, as described in Section 2.5.1. Andrews et. al have developed a new approach, based on the extended Rytov theory combined with the solution of the Helmholtz equation, that treats the turbulence effects separately, namely those caused by the small-scale eddies, assumed to be diffractive inhomogeneities, and those caused

by the large-scale eddies, assumed as refractive inhomogeneities. Mathematically, this is expressed as [58],

$$\hat{I} = XY \tag{2.51}$$

where X and Y the independently mean processes caused by the large and small scale turbulence, respectively. The irradiance can also be written as [58],

$$I = Aexp(2\chi) \tag{2.52}$$

where χ the log-amplitude of the optical wave. When this amplitude is normally distributed, it can be expressed in terms of scintillation index [58],

$$\sigma_I^2 = \exp\left(4\sigma_\chi^2\right) - 1 = \exp(\sigma_{lnI}^2) - 1 \tag{2.53}$$

where σ_{lnI}^2 the variance of the log-irradiance, which depends on the large and small scale variances as follows [58],

$$\sigma_{lnI}^2 = 4\sigma_\chi^2 = \sigma_{lnX}^2 + \sigma_{lnY}^2 \tag{2.54}$$

2.7 Optical Turbulence Modeling

Besides the theoretical background of the optical turbulence, a thorough experimental analysis has been executed over the last 20 years from many academic institutes and research organizations, in order to better understand the effects of optical turbulence in laser beam propagation in the free space and be able to predict it.

One of the most prominent research campaigns towards the aforementioned goals has been executed by the Naval Information Warfare Center (NIWC), San Diego, ex. Space and Naval Warfare Systems Center (SPAWAR). NIWC has focused on the estimation of the refractive index structure parameter in a maritime environment. Specifically, several experimental campaigns have been executed along a 7-km path over San Diego Bay concurrently with meteorological measurements obtained from a sensor located on the sea surface at the midpoint of the path.

A significant difference of a typical diurnal profile of atmospheric turbulence between a terrestrial and an over-water propagation path is that the latter does not exhibit reduced values around sunrise and sunset. That means that the C_n^2 strength does not follow the characteristic bell-shaped diurnal profile but a random one during the day. Additionally, the C_n^2 strength over water is generally an order of magnitude lower than over land [61]. The FSO link availability as a function of range was measured for specific set of FSO parameters for a desert (China Lake, CA) and a maritime (Chesapeake Beach, MD) environments, and the former, as it would be expected, was much larger [61]. Over water, the atmospheric structure constant was found to have a strong dependence on the air-sea temperature difference (ASTD). Additionally, different beam propagation characteristics were observed for each temperature gradient sign. For colder air temperatures the beam transport is improved. Apart from the expression of Eq. 2.17 for the refractive index structure parameter (valid for the inertial-subrange), C_n^2 can be expressed in terms of the temperature structure parameter, C_T^2 , the specific humidity structure parameter, C_q^2, and the temperature-specific humidity cross-structure parameter, C_{Tq}, as follows,

$$C_n^2 = A^2 C_T^2 + 2ABC_{Tq} + B^2 C_q^2 \tag{2.55}$$

where A and B are the partial derivatives of refractive index with respect to temperature and specific humidity, respectively [26]. The Meteorology Department of the NPS has developed a bulk C_n^2 prediction model based upon mean atmospheric layer properties, with an emphasis on the ASTD. The model's basis is the Monin-Obukhov similarity (MOS) theory, which assumes conditions to be horizontally homogeneous and stationary and turbulent fluxes of momentum, sensible and latent heat, to be constant with height. These conditions are most likely to be valid in the open ocean rather than in a coastal location; therefore, the bulk model is believed to perform better

in the former environment [62]. Scaling parameters for wind speed, temperature and specific humidity are defined [62],

$$u_* \equiv \langle -w'u' \rangle^{1/2} \tag{2.56a}$$

$$T_* \equiv -\frac{\langle w'T' \rangle}{u_*} \tag{2.56b}$$

$$q_* \equiv -\frac{\langle w'q' \rangle}{u_*} \tag{2.56c}$$

where u the streamwise wind component and w the vertical wind component. Based on the MOS theory, any dimensionless dynamic surface-layer property by the proper scaling parameters can be expressed as a universal parameter ξ [62],

$$\xi = \frac{zkg(T_* + 0.61Tq_*)}{\theta_u u_*^2} \tag{2.57}$$

where z is the height above the surface, k is the von Karman constant (≈ 0.4), g the gravitational acceleration and θ_u is the vertical potential temperature. The mean vertical profiles of wind speed, temperature and specific humidity are also defined according to MOS theory [62],

$$U(z) = U_0 + \frac{u_*}{k}\left[ln\left(\frac{z}{z_{oU}}\right) - \Psi_U(\xi)\right], \tag{2.58a}$$

$$T(z) = T_0 + \frac{T_*}{k}\left[ln\left(\frac{z}{z_{oT}}\right) - \Psi_T(\xi)\right], \tag{2.58b}$$

$$q(z) = q_0 + \frac{q_*}{k}\left[ln\left(\frac{z}{z_{oq}}\right) - \Psi_q(\xi)\right], \tag{2.58c}$$

where Ψ functions are the integrated forms of the respective dimensionless profile functions and z_{oU}, z_{oT} and z_{oq} the heights where the log-z profiles of U, T and q, respectively, reach their surface values. Solving those profile expressions for the scaling parameters and combining the expressions of the structure parameters in terms of the scaling parameters with Eq. 2.17, results in the NAVSLaM model [63],

$$C_n^2 = \frac{f(\xi)k^2[A^2\Delta T^2 + 2ABr_{Tq}\Delta T\Delta q + B^2\Delta q^2]}{z^{2/3}[\ln\left(\frac{z}{z_{oT}}\right) - \Psi_T(\xi)]^2} \tag{2.59}$$

and

$$\xi = \frac{zg(\Delta T + 0.61T\Delta q)[\ln\left(\frac{z}{z_{oU}}\right) - \Psi_U(\xi)]^2}{\theta_u \Delta U^2 [\ln\left(\frac{z}{z_{oT}}\right) - \Psi_T(\xi)]} \quad (2.60)$$

where $f(\xi)$ is an empirical determined dimensionless function. By iteratively solving Eqs. (2.59) and (2.60), C_n^2 can be estimated after parameterizing z_{oT} and z_{oU} in terms of known quantity [62]. After extended experimental analysis of the model, C_n^2 is found to have strong dependence on the absolute value of the ASTD, increasing for higher wind speeds and negative ASTD and decreasing for positive ASTD values [62]. Compared to relative humidity, C_n^2 is directly proportional for positive ASTD, except for very small positive values [62]. The transition between positive and negative ASTD values was found to have a significant effect on the laser beam pointing. Finally, its dependence on height was found to scale as $z^{-4/3}$ for unstable ($\xi < 0$), $z^{-2/3}$ for neutral ($\xi = 0$) and constant for stable ($\xi > 0$) conditions [62].

Sadot and Kopeika have derived experimentally two models, one for practical use and the other for scientific understanding [40]. The model predictions presented a 90% correlation with actual measurements over a wide range of environmental conditions. An interesting finding was the effect of aerosols on scintillation, which have been taken into account into their modeling. The models are simple mathematical formulas for C_n^2 strength prediction, based on macroscopic meteorological parameters which can be easily obtained from a local weather station. The first model can be mathematically expressed as [40],

$$C_n^2 = 3.8 \times 10^{-14} W(t) + 2 \times 10^{-15} T - 2.8 \times 10^{-15} RH + 2.9 \times 10^{-17} RH^2 - 1.1 \times 10^{-19} RH^3 - 2.5 \times 10^{-15} WS + 1.2 \times 10^{-15} WS^2 - 8.5 \times 10^{-17} WS^3 - 5.3 \times 10^{-13} \quad (2.61)$$

where *W(t)* is a weight function, *T* is the air temperature in Kelvin, *RH* the relative humidity in hPa and *WS* the wind speed in m/s. The second model, apart from wind speed and relative humidity, takes into account the solar flux in Cal/(cm^2 * min) and the total cross-sectional area of particles in cm^2/m^3, namely [40],

$$C_n^2 = 5.9 \times 10^{-15} W(t) + 1.6 \times 10^{-15} T - 3.7 \times 10^{-15} RH + 6.7 \times 10^{-17} RH^2 - 3.9 \times 10^{-19} RH^3 - 3.7 \times 10^{-15} WS + 1.3 \times 10^{-15} WS^2 - 8.2 \times 10^{-17} WS^3 + 2.8 \times 10^{-14} SF - 1.8 \times 10^{-14} TCSA + 1.4 \times 10^{-14} TCSA^2 - 3.9 \times 10^{-13} \quad (2.62)$$

The C_n^2 strength is highly height dependent. The highest values are observed at almost zero altitude, whereas at higher altitude decrease rapidly [64]. The above models have used a height of 15 meters, therefore all subsequent users need to scale them in the desired height. A typical diurnal profile of C_n^2 is characterized by higher values during the day, with a peak around midday and lower ones during night. The lowest values appear around sunrise and sunset. In order to emphasize this profile, both models include a weight function, calculated on the basis of the temporal hour that relates the actual time to the times of sunrise and sunset [39],

$$H_T = 12 \frac{H_{actual} - H_{sunrise}}{H_{sunset} - H_{sunrise}} \quad (2.63)$$

where H_T is the temporal hour, H_{actual} is the actual time, $H_{sunrise}$ is the sunrise time and H_{sunset} the sunset time. Then the weight factor can be assigned based upon Table 2.1.

Table 2.1. Weight Function.

Temporal Hour Interval	Weight Factor
until −4	0.11
−4 to −3	0.11
−3 to −2	0.07
−2 to −1	0.08
−1 to 0	0.06
0 to 1	0.05
1 to 2	0.10
2 to 3	0.51
3 to 4	0.75
4 to 5	0.95

5 to 6	1.00
6 to 7	0.90
7 to 8	0.80
8 to 9	0.59
9 to 10	0.32
10 to 11	0.22
11 to 12	0.10
12 to 13	0.08
over 13	0.13

An extended number of other independent experimental campaigns toward optical turbulence modeling exist. The US Army Research Laboratory (ARL) has been working on its Atmospheric Laser Optics Testbed (A_LOT), which is a 2.33 km near horizontal optical path. A representative amount of work includes scintillation measurements and focal spot displacement to obtain optical turbulence information [65], point measurements of temperature variance and its relationship to refractive index fluctuations [66] and topographic and meteorological influences upon infrared imaging and laser optics communications [67].

The United States Naval Research Laboratory and Naval Academy can also present a significant research work on optical turbulence [68-70]. Specifically, they have developed and evolved ever since, an optical turbulence model known as PAMELA [71,72]. The estimated solar irradiance R, determines the radiation class $c_r = R/300$ and for wind speed u_0, we define the wind speed class c_w. The Pasquill stability category P, is then defined [26],

$$P = \frac{-(4-c_w+c_r)}{2} \tag{2.64}$$

The Obukhov buoyancy length scale L can be calculated, for a roughness length for the shrub-covered dry lakebed, $z_r = 0.16$m, as follows [26],

$$L = \left[(a_1 P + a_2 P^3) z_r{}^{-(a_3 - a_4|P| + a_5 P^2)}\right]^{-1} \tag{2.65}$$

where,

$$a_1 = 0.004349, a_2 = 0.003724, a_3 = 0.5034$$

$$a_4 = 0.231, a_5 = 0.0325$$

The wind shear $\varphi_m(\zeta)$ and the potential temperature gradient $\varphi_h(\zeta)$ are dimensionless and can be expressed as a function of the scaled buoyancy parameter $\zeta = z / L$. The friction velocity u_* is then given by [26],

$$u_* = \frac{\kappa u_0}{\ln(\frac{z}{z_r})} \tag{2.66}$$

and the characteristic temperature T_*,

$$T_* = \frac{-H}{c_p \rho u_*} \tag{2.67}$$

where H the heat flux, c_p the specific heat and ρ the mass density. The eddy dissipation rate, ϵ is [26],

$$\epsilon = \frac{u_*^3(\varphi_m - \zeta)}{0.4z} \tag{2.68}$$

and for a constant $b \approx 2.8$ we can evaluate C_n^2 [26],

$$C_n^2 = \frac{bK_h}{\epsilon^{1/3}}\left(\frac{dn}{dz}\right)^2 \tag{2.69}$$

where K_h the turbulent exchange coefficient for heat and $\frac{dn}{dz}$ given by [26],

$$\frac{dn}{dz} = -\frac{77.6\times10^{-6} P_a T_* \varphi_h}{0.4zT^2} \tag{2.70}$$

In [73], a new approach for optical turbulence estimation was developed, the method of available power. This theory does not take into account the optical intensity distribution of the laser beam. It is based on an analysis of the optical intensity profile and works with the redistribution of the optical intensity within the laser beam. The method's name comes from the fact the intensity's integration result in the available optical power [74],

$$V_{AP} = \iint f(x,y)dxdy \tag{2.71}$$

where *f(x,y)* the three-dimensional available power function obtained by the local minimum in each of the scanned pixels of the laser beam intensity profile. The available

power function *f(x,y)* needs to be positive in order for the method to be valid. The integration limits are defined by the size of the receiving optical lens. The relative volume of the available power, $V_{AP,r}$, is given as the ratio of the volume of available power to the volume of available power in non-turbulent atmosphere, $V_{AP,o}$. Consequently, the turbulence attenuation is given by [74],

$$\alpha_{AP} = 10logV_{AP,r} \tag{2.72}$$

Another bulk method approach for optical turbulence modeling is the Thierman-Kohnle concept according to which the vertical profile of C_T^2 is [75],

$$C_T^2 = 4\beta \frac{T_\emptyset^2}{(kz)^{2/3}} \left[1 + 7\frac{z}{L_\emptyset} + 20\left(\frac{z}{L_\emptyset}\right)^2\right]^{1/3} \tag{2.73}$$

where β is an empirical constant set to 35 W/m^2, $T_\emptyset^2$ the turbulent temperature scale, z the elevation, k the von Karman constant equal to 0.35 and $L_\emptyset$ the Monin-Obukhov length, in case of stability ($\frac{z}{L_\emptyset} > 0$) where air is warmer than the ground, whereas [75],

$$C_T^2 = 4\beta \frac{T_\emptyset^2}{(kz)^{2/3}} \left[1 - 7\frac{z}{L_\emptyset} + 75\left(\frac{z}{L_\emptyset}\right)^2\right]^{-1/3} \tag{2.74}$$

in case of instability ($\frac{z}{L_\emptyset} < 0$), where air is colder than the ground. Furthermore, the turbulent temperature scale and the Monin-Obukhov length can be expressed in terms of the wind velocity, $u_\emptyset$, measured at height z_u [75],

$$u_\emptyset = uk\left[ln\left(\frac{z_u}{z_0}\right) - \psi(L_\emptyset)\right]^{-1} \tag{2.75}$$

$$T_\emptyset = -\frac{Q_0}{u_\emptyset} \tag{2.76}$$

$$L_\emptyset = \frac{u_\emptyset^2 T}{kgT_\emptyset} \tag{2.77}$$

where z_0 the roughness length of the ground surface in meters, Q_0 the vertical turbulent kinematic heat flux (in K times meters per second) and g the acceleration due

to gravity. Depending on the stability of the conditions, the parameter $\psi(L_\emptyset)$ is expressed as [75],

$$\psi(L_\emptyset) = \begin{cases} 2ln\left(\frac{1+y}{2}\right) + ln\left(\frac{1+y^2}{2}\right) - 2tan^{-1}y + \frac{\pi}{2}, & \frac{z}{L_\emptyset} < 0 \\ -\frac{4.7z_u}{L_\emptyset}, & \frac{z}{L_\emptyset} > 0 \end{cases} \quad (2.78)$$

where y is,

$$y = \left(1 - \frac{15z_u}{L_\emptyset}\right)^{1/4} \quad (2.79)$$

2.7.1 Numerical Weather Prediction Modeling

A different way to predict the optical turbulence in the atmosphere and exploit it for FSO systems performance assessment and improvement, is the numerical weather prediction (NWP) models. For experiments that use optical devices which are sensitive to scintillation effects caused by fluctuations in the atmospheric refractive index, it is highly desirable to have a method for forecasting the scintillation levels. This requires a forecast model of atmospheric conditions as a starting point. However, since most NWP models cover a large area, the grid spacing of even high-resolution models is necessarily coarse relative to the scales of scintillation [76]. Most NWP models suffer from smoothing and filtering effects so that the smallest scales produced by the model are in fact underresolved.

Tatarski proposed a theoretical model in 1961 to compute C_n^2 defined as [77],

$$C_n^2(z) = \gamma M^2 L_0(z)^{4/3} \quad (2.73)$$

where $\gamma \approx 2.8$ a constant, L_0 the outer scale of the turbulence, and *M* the vertical gradient of the generalized potential refractive index of the air [77],

$$M = -80 \cdot 10^{-6} \frac{P}{gT} N^2 \quad (2.74)$$

where *P* the atmospheric pressure in hPa, T the air temperature in K, g the gravity acceleration and N the buoyancy frequency given by [77],

$$N^2 = \frac{g}{\theta}\frac{\partial \theta}{\partial z} \tag{2.75}$$

where θ the potential temperature in K, given by [77],

$$\theta = T\left(\frac{1000}{P}\right)^{R/c_P} \tag{2.76}$$

where R = 287 $JK^{-1}kg^{-1}$ the gas constant of the air and c_P = 1004 $JK^{-1}kg^{-1}$ the specific heat capacity at a constant pressure [77]. This model, coupled with a Weather Research Forecast (WRF) model have been used to predict the refractive index in the vertical profile [77-80]. The exact process is forecasting routine meteorological parameters by a WRF model and then predict C_n^2 by a turbulence model (i.e. Monin-Obukhov Similarity Theory based model) based on these forecasts.

WRF model is a mesoscale NWP model both for professional forecasts and atmospheric research. WRF has been developed by the National Center of Environment Prediction (NCEP) and the National Center of Atmospheric Research (NCAR) of the United States. WRF is initialized by the Global Forecast System (GFS) data, with a horizontal resolution of 1^o X 1^o (longitude and latitude). WRF model exports a big variety of routine meteorological parameters which depend upon the physical schemes that have been chosen for the simulation. The available schemes are the following [77]:

(1) The micro-physics process uses the WRF Single-Moment 3 class (WSM-3) scheme which is suitable for medium-scale grid dimension, and contains three kinds of water materials: water vapor, cloud water or cloud ice, rainwater or snow.

(2) The Rapid Radiative Transfer Model (RRTM) scheme is used for the longwave radiation. The longwave process is caused by water vapor, ozone, carbon dioxide and other gases, as well as by the optical depth of cloud.

(3) The shortwave radiation uses the Goddard scheme which is suitable for cloud resolution models.

(4) The planetary boundary layer uses the Yonsei University (YSU) scheme which is suitable for ocean environment simulation and adds the process of dealing with entrainment at the top of planetary boundary layer.

(5) The surface layer uses the Monin-Obukhov scheme which is based on the Monin-Obukhov similarity theory (MOST).

(6) The cumulus parameterization uses the Kain-Fritsch scheme which consists of a cloud model concerning the water vapor lifting and subsidence, with the phenomena, such as, entrainment, detrainment, air-current ascension and subsidence covered.

2.8 Optical Turbulence Modeling in Monterey Bay

An experimental campaign took place during September and October of 2020 over the Monterey Bay in California. The main goal of this campaign was to measure atmospheric turbulence over the water and compare the results with a theoretical model called the Navy Surface Layer Model (NAVSLaM), developed by the Meteorology Department at the Naval Postgraduate School (NPS), as well as conduct a regression analysis for turbulence predictive modeling based on environmental parameters. The results showed very good agreement between theory and experiment.

The experimental measurements took place over the southern part of the Monterey Bay, California. The experiment utilized an MZA DELTA turbulence monitor, which consisted of a telescope located on the Coast Guard pier and a target board located on a coastal bluff (Figure 2.3). The optical path distance was 1563 meters, at a height that varied from ~2.8 meters above sea level at the telescope to ~10 meters

above sea level at the target board; these heights fluctuated due to tidal variation. More than 95% of the path was over the water. The DELTA uses a Celestron f/10, 6-inch diameter aperture Schmidt-Cassegrain telescope equipped with a Point Grey 3.2 mega-pixel Grasshopper 3 USB Camera in order to obtain images of the target board. A laptop computer was connected to the camera through a USB cable, in order to store and analyze the images. The MZA software calculates C_n^2 along the image path by tracking target board features as they jitter due to turbulence. Specifically, it measures differential jitter of feature pairs as a function of angular separation. The C_n^2 turbulence profiles are estimated by observing the jitter over many length scales, as illustrated in Figure 2.4. Turbulent cells close to the telescope (the red box in Figure 2.4) are associated with jitter correlations over larger scales (the red line in the target board image on the right in that figure), while cells close to the target board (the green box) cause jitter correlations over smaller scales (the green line). Thus, by applying a path weighting function, it is possible to obtain a C_n^2 profile along the beam path. DELTA measurements are separated into ten bins. Path-weighting functions in each bin are heavily weighted at the receiver, extending to the midpoint of the path, but effectively zero at the target. Path weighting functions model how turbulence along the path contributes to the expected value of the tilt-variance is observed. The system was set up to take one measurement every minute and the output file contained the turbulence profile. Figure 2.5 shows the locations of the DELTA telescope, the target board, and the weather stations.

Figure 2.3 The MZA DELTA telescope and the Ambient WS-2000 weather station located on the Coast Guard Pier (left) and the target board (right) [J7].

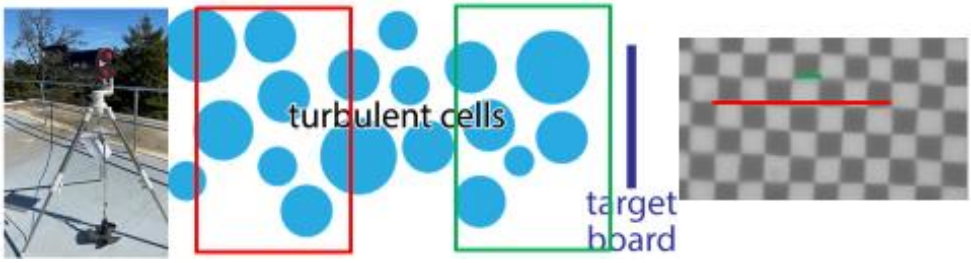


Figure 2.4. Differential jitter measurements vs angular separation using MZA DELTA system [J7].

Figure 2.5. The maritime path over Monterey Bay, CA. The MZA DELTA telescope was located on the Coast Guard Pier and the target board on a coastal bluff [J7].

The experimental setup also included weather stations located on the Coast Guard Pier next to the DELTA (Ambient WS-2000) and on the Municipal Pier (NOAA weather station). Real time data from these weather stations were used as inputs to the NAVSLaM model for C_n^2 predictions. The WS-2000 weather station was installed at a height of ~4.8 meters above sea level, providing meteorological data such as air temperature, relative humidity, air pressure, wind speed, solar radiation, and rainfall rate; the NOAA weather station provided additional measurements, including tidal height and sea temperature.

CHAPTER 3

FSO Performance Modeling

Two are the main factors in order to characterize the performance of an FSO link, that is the channel modeling and system performance metrics. This section provides the background of channel modeling, including the atmospheric effects that affect a laser beam propagating in the atmosphere and the most well-known probability density functions (pdf) that are being utilized towards irradiance fluctuations modeling. Additionally, the performance metrics on the system-level are provided, to include bit error rate (BER), outage probability, average capacity and outage capacity. Finally, representative experimental FSO performance modeling efforts, are also presented.

3.1 Introduction

In the FSO optics, the modulated optical signal of the transmitter is propagating through the dynamically changing atmosphere. This fact, has been proved, both theoretically and experimentally, that causes link performance degradation and even total outage. The main root causes are two, the free space propagation effects and the system itself.

The optical signal before being transmitted in the free space is being modulated, with multiple possible techniques which influence parameters, such as bandwidth and energy efficiency. Perhaps, the most commonly used techniques are the On-Off Keying (OOK), the phase shift keying (PSK), the differential phase shift keying (DPSK) and orthogonal modulation formats [81]. The modulation formats exhibit different

sensitivity on the turbulence flow in the atmosphere, which results in received optical power decrease, something that affects the bit-error-rate (BER). The advantage of OOK modulation is its simplicity, however it is more sensitive to atmospheric turbulences and other disturbing influences for received level fluctuations. In this case, the BER is defined as follows [81],

$$BER_{OOK-NRZ} = \frac{1}{2} erfc\left(\frac{1}{2\sqrt{2}}\sqrt{SNR}\right) \quad (3.1)$$

In [82], it is shown that an Optical Wireless Communication (OWC) system which uses an Optical Code Division Multiple Access (OCDMA) based on Carbon Nanotubes (CNTs), exhibits improved data rate, BER and signal-to-noise ratio (SNR). Another approach for FSO performance enhancement in terms of modulation scheme, is the optical spatial modulation (OSM) as a mean to achieve spatial diversity in coherent FSO systems [83]. By using a generic analytical framework for obtaining the average bit error rate probability (ABEP) of uncoded and coded OSM with coherent detection in the presence of turbulence, it was shown that OSM can offer comparable performance with conventional coherent FSO schemes [83]. For the case of multiple-input multiple-output (MIMO) FSO communication systems, it is shown that the performance is worse for a spatial multiplexing scheme when OOK modulation is employed, even with the optimal maximum likelihood detection at the receiver, as compared to the repetition coding (RC) case for equal transmission rate [84]. In such a transmitter architecture (MIMO), by controlling the amplitude and phase of the optical field at each transmitter, based on turbulence state information fed back from the receiver, it has been shown that the system's performance was significantly increased by exploiting the instantaneous structure of the turbulence [85].

Another factor that seriously affects the performance of an FSO communication link is the pointing errors, especially at high range links. Simulations have shown that

the maximum pointing error allowed to achieve a 10^{-9} BER for clear weather, can reach 13.53 μrad. In case of heavy fog, which is the worst case scenario for an FSO link, it is shown that even very small pointing errors can lead to total link failure [86].

Another essential element of supporting the engineering development and operational employment of aerospace electro-optical sensor systems, is the extinction coefficient measurements, which allow the calculation of their range performance, either for current or possible future applications, such as ranging, weapons, remote sensing and possible planetary exploration missions. Novel methods were proposed in [87], based on laser energy measurement incident on surfaces with known geometric and reflection characteristics by using IR detectors. They also proposed algorithms that allow a direct determination of the atmospheric transmittance and spatial characteristics of the laser spot, as [87, Tables 7 - 9]. Assuming an extended target, the power at the detector is expressed as follows [87],

$$P = K_{SYS} \cdot \rho \cdot \frac{1}{d_0^2} \cdot e^{-2\gamma d_0} \tag{3.2}$$

where ρ the target reflectivity, d_0 the distance between the transmitter and the target, γ the extinction coefficient and K_{SYS} a constant including all relevant transmitter system parameters.

The first step in studying the mean fade time in an FSOC system, is the temporal analysis of the irradiance at the detector plane. A novel technique to execute this analysis for a laser beam (i.e Gaussian) propagating through the turbulent atmosphere is by means of computer simulation, and specifically by a known numerical method to generate long phase screens [88-89]. In order to simulate the wind effect, these screens are displaced in a transverse direction as the wave propagates. The temporal power spectrum S_I can be derived from the temporal covariance B_I with a Fourier transform [88],

$$S_I(\omega) = 2\int_0^{\infty} B_I(\tau)\exp(-j\omega\tau)\,d\tau \tag{3.3}$$

The most relevant parameter needed to assess the link's availability is the mean fade time, defined as the continuous average period of time in which the received power is below a certain level. This time period is expressed as follows [89],

$$\langle T(F_0)\rangle \equiv \frac{P\{F<F_0\}}{\langle n(F_0)\rangle} \tag{3.4}$$

where F the fading level below the average power, F_0 a certain fading value, $\langle n(F_0)\rangle$ the expected number of fades per second and $P\{F < F_0\}$ the fading probability.

When assessing the performance of hybrid radio-frequency/FSOC system, we need to investigate the noise propagation effect, which can be very serious when battery-charged sensor nodes have very limited transmit power. In [90], an exact expression for the cumulative distribution function (CDF) is presented as follows,

$$F_{\gamma FSO}(x) = \sum_{j=0}^{\infty} \frac{b_j(\alpha,\beta,\delta)}{(\eta^w \mu_w)^{\frac{j+\beta}{w}}} x^{\frac{j+\beta}{w}} + \frac{b_j(\alpha,\beta,\delta)}{(\eta^w \mu_w)^{\frac{j+\beta}{w}}} x^{\frac{j+\beta}{w}} \tag{3.5}$$

where μ_w the average electrical SNR, α and β atmospheric parameters, δ a ratio between the equivalent beam radius and the pointing error displacement standard deviation at the receiver and [90],

$$b_j(\alpha,\beta,\delta) = \frac{a_j(\alpha,\beta)\delta^2}{\{(j+\beta)(\delta^2-j-\beta)\Phi^{j+\beta}\}} \tag{3.6}$$

Outdoor experiments for FSO performance validation and assessment, including different applications, terrains, geometry and environments are perhaps the safest way to extract reliable insights about their feasibility and potential for future use [91]. However, outdoor experiments are usually expensive and difficult to reproduce, therefore it is often helpful to develop realistic numerical and experimental simulations. Spatial light modulators (SLMs), have been demonstrated as a well-suited alternative for simulating different turbulent conditions in a laboratory [92].

3.2 Environmental Effects

The performance of FSO links is subject to various environmental factors like fog, haze, rain, etc. that lead in received signal power degradation. From these factors, the atmospheric attenuation is usually dominated by fog, since its particle size is comparable with the wavelength of interest in FSO system. It can alter the optical signal characteristics or even completely block the light passage because of absorption, scattering, and reflection. The atmospheric visibility is a useful measure of atmospheric environmental conditions prediction and is defined as the distance that a parallel luminous beam travels through the atmosphere until its intensity drops 2% of its original value. In order to estimate the optical attenuation based on the visibility statistics to predict the availability of an FSO system, the relationship between visibility and attenuation has to be known. In order to characterize the attenuation of an optical signal propagating through a medium, we use a term called "specific attenuation" which is defined as the attenuation per unit length expressed in dB/km and is given as [93],

$$\beta(\lambda) = \frac{1}{R} \cdot 10log\left(\frac{P_0}{P_R}\right) = \frac{1}{R} \cdot 10log\left(e^{\gamma(\lambda)R}\right) \tag{3.7}$$

where R the link length, P_0 the emitted optical power, P_R the optical power at a distance R and $\gamma(R)$ the atmospheric attenuation coefficient.

(i) Fog effect: The fog particle is composed of very fine water droplets or ice, or combination of them near the earth's surface, which scatter the light and hence reduce the visibility. Fog is correlated with a visibility of less than 1 km and relative humidity that reaches the saturation level (100%) [93]. The most important parameter which describes fog, is the particle size distribution, usually modeled by the modified gamma distribution, as follows [93],

$$n(r) = N_0 r^m exp(\Lambda r^\sigma) \tag{3.8}$$

where $n(r)$ denotes the number of particles per unit volume per unit increment of the particle having radius r and N_0, m, Λ and σ are the four adjustable parameters that characterize the particle size distribution.

In FSOC systems, the operating wavelength is chosen to fall within the low absorption bands. Therefore, absorption contribution to the total attenuation coefficient becomes very small as compared to the scattering effect [94]. Therefore, studying fog particles scattering is important in order to predict the attenuation for wireless network planning and installation. Determining the size, and water content in fog particles is important to predict the attenuation. The fog particle radii differ in different climatic regions and therefore, we can observe different attenuation for optical wave propagating through fog conditions even at the same wavelength. Assuming spherical shape, fog particles can be categorized into following three classes based on their radii [95]:

a) Aitken particles and ultra-fine particles: Fine particles with an average size ranging between 0.001 and 0.1 μm.

b) Fine particles: The size range of these particles lie within 0.1 and 1 μm.

c) The larger particles: The size of these particles lie within 1 and 100 μm.

Apparently, this information is difficult to achieve and not always available at the FSO link installation site. Therefore, researchers have proposed empirical models that depend on visibility data, which are widely available from meteorological stations in cities. Such models have been developed from Kruse, Kim and Al Nabulsi and use this approach and predict specific attenuation using visibility. The specific attenuation in dB/km for both Kim and Kruse model is given by [96],

$$a_{spec} = \frac{10logV\%}{V(km)}\left(\frac{\lambda}{\lambda_0}\right)^{-q} \tag{3.9}$$

where V(km) is the visibility, $V\%$ is the transmission of air drops to percentage of clear sky, λ the operating wavelength and λ_0 the wavelength of reference (550 nm).

For Kruse model [96],

$$q = \begin{cases} 1.6, & V > 50km \\ 1.3, & 6km < V < 50km \\ 0.585V^{1/3}, & V < 6km \end{cases} \tag{3.10}$$

For Kim model [96],

$$q = \begin{cases} 1.6, & V>50km \\ 1.3, & 6km<V<50km \\ 0.16V + 0.34, & 1km < V < 6km \\ V-0.5, & 0.5km<V<1km \\ 0, & V<0.5km \end{cases} \tag{3.11}$$

The specific attenuation in dB/km, both for advection and radiation fog is given by Al Naboulsi as follows [96],

$$a_{spec}\left(\frac{dB}{km}\right) = \frac{10}{ln10}\gamma(\lambda) \tag{3.12}$$

Finally, Ijaz model is given, for wavelengths from 0.6 to 1.6 μm, by [93],

$$a_{spec}\left(\frac{dB}{km}\right) = \frac{17}{V}\left(\frac{\lambda}{0.55}\right)^{-q(\lambda)} \tag{3.13}$$

where $q(\lambda) = 0.1428\lambda - 0.0947$. This model is valid only for visibilities between 15 m and 1 km.

It has been shown, that the fog attenuation effect does not have linear relationship with wavelengths in the region from 0.6 to 1.6 μm, for $V < 0.5$ km, whereas the effect of smoke attenuation is found to be linearly decreasing from 0.83 to 1.55 μm from very dense smoke ($V < 0.07$ km) to moderate smoke ($V < 0.5$ km) [97]. In [98] a fog sensor is presented, which obtains environmental data, process them mathematically and can exploit them for FSO link evaluation as well as in process of switching FSO link with some form of back up radio link (RF/FSO).

The Wavelength Division Multiplexing (WDM) technique, allows many signals to be multiplexed together and transmitted as one signal. When applied in FSOC systems, multiple modulating signals modulate different optical carriers, which are then multiplexed and sent through a single laser beam. An improved version of this technique, to better cope with the atmospheric effects, is the multibeam WDM, where more than one beam of the multiplexed signal traverse through the free space and reach the receiver [99]. Since each beam follows a different path, faces different attenuation too. In [99] it was shown that a multibeam WDM-FSO system exhibits a significantly improved performance under heavy haze conditions.

(ii) Rain effect: Rain attenuation, although heavily dependent on the precipitation microphysics, is of the same order of magnitude as at mm-waves. That being said, when fog is non-negligible, FSOCs are outperformed by microwave links in high-availability applications. On the other hand, in locations where fog is rare or not present at all, rain is the main atmospheric phenomenon that limits the performance of an FSO link. Several models have been developed to predict the attenuation caused by rain at millimeter frequency level, such as the International Telecommunication Union – Radio communication sector (ITU-R), the Brazilian, the Lin and the Moupfouma [100]. Although several models for FSO applications exist, they are usually location dependent and based on limited local data.

Specific attenuation due to rain, in the optical transmission windows that are usually adopted by commercial FSO systems, is wavelength independent, and is often calculated from the rain rate R (in mm/h) as follows [100],

$$\gamma = kR^{a} \tag{3.14}$$

where k and α are coefficients that are very sensitive to the distribution of the raindrop size. A typical example is the experimental calculation of these coefficients

for a typical tropical region, as k = 2.03 and α = 0.74 by using least square mean equation (LSME) method with Levenberg–Marquardt optimization based on one year collected heavy rain data [101]. Typical values for rain attenuation prediction model coefficients are shown in Table 3.1 [102].

Table 3.1. Rain attenuation prediction model for FSO

Attenuation	**Relation**
Light Rain (R<3.8 mm/hr)	$0.509R^{0.63}$
Mean Rain (3.8<R<7.6 mm/hr)	$0.319R^{0.63}$
Strong Rain (R>7.6 mm/hr)	$0.163R^{0.63}$
Rain	$0.365R^{0.63}$

Specific rain attenuation in dB/km can be calculated by integrating all drop sizes as follows [103],

$$\gamma = 4.343 \int Q(D, \lambda, m) N(D) dD \tag{3.15}$$

In order to be able to build a global model, a methodology able to synthesize realistic rain fields globally is presented in [100]. Afterwards, the relationship between specific attenuation and rain rate is derived from scattering theory and finally the rain attenuation impairing a link of an arbitrary length is obtained by integrating specific rain attenuation along that path. The results are presented in the form of the complementary cumulative distribution function (CCDF) of rain attenuation, a key element supporting the design of wireless communication systems.

(iii) Snow effect: Snow attenuation depends on the size of the snowflakes and the snowfall rate. Since snowflakes are larger than the raindrops, they cause deeper fades in the signal. Snowflakes can be as large as 20mm in diameter, therefore, they can completely block the path of the optical signal, based on the width of the beam. The

snow attenuation can be classified into wet and dry. The specific attenuation (dB/km) for snow rate S in mm/hr is given as follows [4],

$$\beta_{snow} = aS^b \tag{3.16}$$

where the values of parameters a and b in dry and wet snow are [4],

$$\text{Dry snow: } a = 5.42 \times 10^{-5} + 5.4958776, \quad b = 1.38$$

$$\text{Wet snow: } a = 1.023 \times 10^{-4} + 3.7855466, \quad b = 0.72 \tag{3.17}$$

The snow attenuation can also be approximated based on the visibility as follows [4],

$$a_{snow} = \frac{58}{V} \tag{3.18}$$

3.3 Channel Statistical Modeling

In this section, the various performance metrics of the wireless channels environment in modern communications schemes, whose availability and reliability is closely related to the signal fading conditions, are presented. The energy propagation in a mobile radio environment is characterized by several effects, like multi-path fading and shadowing, which result in channel strength variations over time and frequency. These variations can be classified into two categories, the large scale fading, due to path loss and shadowing and the small scale fading, due to reflection, scattering, diffraction and absorption of the propagating waves with physical obstacles [104]. The mathematical modeling of wireless channels is also facilitated by another classification of the fading phenomena, the slow and the fast, based on the rate at which the magnitude and phase of the received signal varies with respect to channel changes. This classification is related to the channel's coherence time T_c, a measure of the minimum time required for the magnitude change of the channel to become uncorrelated from its

previous value. The coherence time is related to the channel Doppler spread f_d as follows [104],

$$T_c \cong 1/f_d \tag{3.19}$$

A wireless channel is classified as slow fading, when T_c, is large relative to the delay requirement of a specific application, and fast fading when it is small. It is important to notice that the characterization of a fading channel as slow or fast depends strongly on the bit rate of the link, which means that as data rates increase, a wireless channel is better described as slow fading and vice versa.

Initially, statistics of irradiance referred only on single family distributions, such as the modified Rician and lognormal (LN) distribution. Later, the nonstationary nature of atmospheric turbulence, showed that these distributions could not describe turbulence accurately [105]. There has been a significant interest, in the development of an accurate PDF for received irradiance description of an optical wave after propagating through a turbulent path. The goal is to derive a distribution valid throughout the strength region of turbulence. An ideal PDF of irradiance, would be valid in all turbulence regimes, for any receiving aperture size, has parameters related to physical atmospheric quantities, and is described by a closed mathematical form [105].

Malaga Distribution

In this distribution the small-scale fading characteristic of the atmospheric channel is primarily due to diffraction and is modelled by three different signal components, the line-of-sight (LOS) field component, U_L, and two scattered components due to small-scale fluctuations [106]. The first, U_S^C, is the quasi-forward signal that is scattered by eddies within the propagation axis, assumed to be coupled with U_L. The second, U_C^G, is the scattered energy by off-axis eddies and considered

statistically independent from the other two. Therefore, the optical field is described as follows [106],

$$U = (U_L + U_C^G + U_S^C)\exp(\chi + jS) \tag{3.20}$$

where χ and S are real random variables (RVs) which represent the log-amplitude and phase fluctuations of the optical field, respectively. Both variables model the large-scale fading characteristic of the channel which is due to refractive effects [107]. The LOS components are defined as follows [107],

$$U_L = \sqrt{G}\sqrt{\Omega}\exp(j\varphi_L)$$

$$U_S^C = \sqrt{G}\sqrt{\xi_c}\exp(j\varphi_C) \tag{3.21}$$

where G is a random variable that follows the Gamma distribution and E[G]=1 and φ_L and φ_C deterministic phases of the LOS and the coupled-to-LOS components. Ω is the LOS average optical power represented by $\Omega = E[|U_L|^2]$ and ξ the average optical power of the total scatter components, represented by $\xi = E\left[\left|U_S^C\right|^2 + \left|U_C^G\right|^2\right] = \xi_c + \xi_g$. The amount of scattering power is represented by the parameter ρ, which relates the two scattering components and ranges from 0 to 1 [108]. Thus we take,

$$\xi_c = \rho\xi \; and \; \xi_g = (1-\rho)\xi \tag{3.22}$$

and the total average optical power [108],

$$E[I] = \Omega + \xi \tag{3.23}$$

From Eq. (3.20) the irradiance in the receiver is expressed as follows [108],

$$I = \left|U_L + U_S^C + U_C^G\right|^2 \exp(2\chi) = YX \tag{3.24}$$

where $X = exp(2\chi)$ and $Y = \left|U_L + U_S^C + U_C^G\right|^2$ are independent variables for the large and small-scale fluctuations, respectively [108]. The large-scale parameter follows a log-normal distribution but for better mathematical tractability it is modeled with a Gamma pdf, whereas the small-scale parameter, is obtained from a combination

of a Nakagami-m distribution and a Rayleigh random phasor. Thus, the small-scale fluctuation pdf is given by [108],

$$f_Y(y) = \frac{1}{\xi_g}\left[\frac{\beta\xi_g}{\Omega'+\beta\xi_g}\right]^{\beta} exp\left[-\frac{y}{\xi_g}\right] \ {}_1F_1\left(\beta; 1; \frac{\Omega'}{\Omega'+\beta\xi_g}\frac{y}{\xi_g}\right) \tag{3.25}$$

where $\Omega' = \Omega + \xi_c + 2\sqrt{\Omega\xi_c}\cos(\varphi_L - \varphi_C)$ is the average power from the coherent contributions, β the shape parameter of the Nakagami distribution, whereas ${}_1F_1(\cdot)$ the Kummer confluent hypergeometric function of the first kind [109]. Finally, we take the pdf of the received irradiance *I* as follows [110],

$$f_I(I) = A\sum_{k=1}^{\beta} a_k I^{\frac{a+k}{2}-1} K_{a-k}\left(2\sqrt{\frac{\alpha\beta I}{\xi_g\beta+\Omega'}}\right) \tag{3.26}$$

where,

$$A = \frac{2a^{\frac{a}{2}}}{\xi_g^{1+\frac{a}{2}}\Gamma(a)}\left(\frac{\xi_g\beta}{\xi_g\beta+\Omega'}\right)^{\beta+\frac{a}{2}} \tag{3.27}$$

and

$$a_k = \binom{\beta-1}{k-1}\frac{(\xi_g\beta+\Omega')^{1-\frac{k}{2}}}{\Gamma(k)}\left(\frac{\Omega'}{(\xi_g}\right)^{k-1}\left(\frac{\alpha}{\beta}\right)^{\frac{k}{2}} \tag{3.28}$$

Gamma-Gamma Distribution

The Gamma-Gamma distribution is a two parameter model based on the doubly stochastic scintillation theory which assumes that the large scale irradiance fluctuations modulate the small scale irradiance fluctuations of the propagating beam and both of them are governed by two independent Gamma distributions [111]. A significant advantage of this distribution, is that it's parameters are directly related to the local atmospheric turbulence conditions, as well as with the parameters of the link, such as the link's length, the aperture diameter of the receiver and the operational wavelength.

We can express the second moment of Eq. (3.24) as follows [111],

$$\langle I^2\rangle = \langle X^2\rangle\langle Y^2\rangle = (1+\sigma_x^2)\left(1+\sigma_y^2\right) \tag{3.29}$$

where σ_x^2 and σ_y^2 are the normalized variances of *X* and *Y* respectively. To develop an irradiance pdf model for this distribution, it is assumed that both large-scale and small-scale irradiance fluctuations are governed by the Gamma distribution [112],

$$f(x) = \frac{a(ax)^{a-1}}{\Gamma(a)} \exp(-ax)\,, x > 0, a > 0 \tag{3.30}$$

By fixing *x* and setting *y = I/x*, we can take the conditional pdf [111],

$$p_y(I|x) = \frac{\beta(\beta I/x)^{\beta-1}}{x\Gamma(\beta)} \exp(-\frac{\beta I}{x}) \tag{3.31}$$

where *x* is the conditional mean value of *I*. Forming the average of Eq. (3.31) over the gamma distribution, leads to the Gamma-Gamma distribution [111],

$$p(I) = \int_0^\infty p_y\left(\frac{I}{x}\right) p_x(x)dx = \frac{2(a\beta)^{(a+\beta)/2}}{\Gamma(a)\Gamma(\beta)} I^{(a+\beta)/2-1} K_{a-\beta}\left[2(a\beta I)^{1/2}\right], I > 0 \tag{3.32}$$

The parameter α represents the effective number of large-scale cells of the scattering process and β the effective number of small-scale cells. The gamma-gamma pdf results in,

$$\langle I^2 \rangle = \left(1 + \frac{1}{a}\right)(1 + \frac{1}{\beta}) \tag{3.33}$$

The gamma-gamma distribution has been extensively used for turbulence channel modeling in FSO link performance studies [113 – 118].

A new and unifying statistical model for irradiance fluctuations and valid under all range of turbulence conditions was proposed in [119], named Double Generalized Gamma. Actually, it demonstrated an excellent match to simulation data and proved to be superior as compared to Gamma-Gamma model. By integrating (3.32) we take the Double Generalized Gamma distribution [120],

$$f_I(I) = \frac{\gamma_2 p p^{m_2-1/2} q^{m_1-1/2} (2\pi)^{1-(p+q)/2} I^{-1}}{\Gamma(m_1)\Gamma(m_2)} \times$$

$$G_{1,p+q+1}^{p+q,1}\left[\left(\frac{I^{\gamma_2}}{\Omega_2}\right)^p \frac{m_1^q m_2^p}{p^p q^q \Omega_1^q} \Big|_{\Delta(q: m_1), \Delta(p: m_2), 0}^{1}\right] \tag{3.34}$$

where G [.] the Meijer G-function, m_i correspond to parameters alpha and beta of Eq. (3.32).

Double-Weibull Distribution

If we consider a random variable, x, with a Weibull distribution we take [121],

$$f_x(x) = \frac{\beta x^{\beta-1}}{\Omega} \exp\left(-\frac{x^\beta}{\Omega}\right) \tag{3.35}$$

where $\beta > 0$ irradiance fluctuation strength related parameter of the distribution and $\Omega > 0$ is related to the average power. The n-order moment of x is given by [121],

$$E\langle x^n \rangle = (\Omega)^{n/\beta} \Gamma\left(1 + \frac{n}{\beta}\right) \tag{3.36}$$

where $E\langle . \rangle$ denotes the expectation and Γ (.) the Gamma function. The second moment of the irradiance is given by Eq. (3.29). Then the scintillation index is given by,

$$\sigma_I^2 = \frac{E\langle I^2 \rangle}{E\langle I \rangle^2} - 1 = (1 + \sigma_x^2)\left(1 + \sigma_y^2\right) - 1 \tag{3.37}$$

If we assume that both irradiance scales are governed by the Weibull distribution, then using the generalized statistical model proposed in [122] and after some simplifications we take the Double-Weibull pdf of I,

$$f_I(I) = \frac{\beta_2 k (kl)^{\frac{1}{2}}}{(2\pi)^{\frac{l+k}{2}-1}} I^{-1} \times G_{k+l,0}^{0,k+l}\left[\left(\frac{\Omega_2}{I^{\beta_2}}\right)^k k^k l^l \Omega_1^l \middle|_{-}^{\Delta(l;0),\,\Delta(k;0)}\right] \tag{3.38}$$

I-K Distribution

The *I-K* distribution has been found to be applicable for the received irradiance fluctuations modeling under all range of turbulence conditions. It includes two parameters, namely α and ρ_y, which represent the effective number of scatters and a coherence parameter, respectively. The pdf of the I-K distribution is expressed as follows [123],

$$f_{IK,I_r}(I_r) = \begin{cases} 2ay\left(\frac{yI_r}{\rho}\right)^{\frac{a-1}{2}} K_{a-1}\left(2\sqrt{a\rho}\right) I_{a-1}\left(2\sqrt{ayI_r}\right), & for\ I_r < \frac{\rho}{y} \\ 2ay\left(\frac{yI_r}{\rho}\right)^{\frac{a-1}{2}} I_{a-1}\left(2\sqrt{a\rho}\right) K_{a-1}\left(2\sqrt{ayI_r}\right), & for\ I_r > \frac{\rho}{y} \end{cases} \quad (3.39)$$

where $I_r = I/I_n$ the normalized irradiance, I and I_n the instantaneous and average received irradiance, respectively. Additionally, $y = 1+ \rho$ and $I(.)$ and $K(.)$ the modified Bessel function of the first and the second kind of order v, respectively [124].

Log-Normal (LN) Distribution

The log-normal distribution models pretty accurately the behavior of the channel under weak turbulence conditions. One of the main advantages of this model, as stated in Gamma-Gamma distribution too, is that the link's parameters are directly related with the ones of the distribution. The log-normal distribution model of the normalized irradiance I is expressed as follows [60],

$$f_I(I) = \frac{1}{I\sigma\sqrt{2\pi}} exp\left(-\frac{\left(\ln(I)+\sigma^2/2\right)^2}{2\sigma^2}\right) \quad (3.40)$$

where σ^2 the log-irradiance variance, depending on the channel characteristics and expressed for a plane wave propagation as follows [60],

$$\sigma^2 = exp\left[\frac{0.49\sigma_R^2}{\left(1+0.65d^2+1.11\sigma_R^{12/5}\right)^{7/6}} + \frac{0.51\sigma_R^2\left(1+0.69\sigma_R^{12/5}\right)^{-5/6}}{1+0.9d^2+0.62d^2\sigma_R^{12/5}}\right] - 1 \quad (3.41)$$

while for a spherical wave propagation the expression is as follows [60],

$$\sigma^2 = exp\left[\frac{0.49\sigma_R^2}{\left(1+0.18d^2+0.56\sigma_R^{12/5}\right)^{7/6}} + \frac{0.51\sigma_R^2\left(1+0.69\sigma_R^{12/5}\right)^{-5/6}}{1+0.9d^2+0.62d^2\sigma_R^{12/5}}\right] - 1 \quad (3.42)$$

where $d = \sqrt{kD^2/4L}$ the receiver diameter.

Fisher-Snedecor F Distribution

The F-Distribution is also a two-parameter distribution based on the doubly stochastic turbulence induced fading, which assumes that the small-scale irradiance variations are modeled by a gamma distribution and are subject to the large-scale

variations which are modeled by an inverse gamma distribution [125]. The irradiance can be expressed in the form of Eq. (3.24) that corresponds to the multiplication of the small (I_S) and large (I_L) scale irradiance fluctuations. This model assumes that the small-scale parameter is modeled by a gamma distribution given by Eq. (3.30). Many previous works, suggested that the large scale parameter is better expressed by the log-normal distribution. More recently, the inverse gamma distribution has also been suggested as an accurate approximation to the log-normal distribution. Additionally, the inverse gamma is related to the gamma distribution and exhibits semi heavy-tailed characteristics that allows it to model the large-scale fading [125]. The pdf for the large-scale is given by [126],

$$f_{I_L}(y) = \frac{(b-1)^b y^{-b-1}}{\Gamma(b)} exp\left(-\frac{b-1}{y}\right), b > 1 \tag{3.43}$$

I_S and I_L are independent, therefore the pdf expression for I is given by [126],

$$f_I(I) = \int_0^\infty x^{-1} f_{I_s}(x) f_{I_L}(Ix^{-1}) dx \tag{3.44}$$

Thus, if we substitute the two previous mentioned expressions for I_S and I_L in Eq. (3.44) we take the F-distribution for I in closed form [126],

$$f_I(I) = \frac{a^a (b-1)^b I^{a-1}}{B(a,b)(aI+b-1)^{a+b}} \tag{3.45}$$

where the scintillation index can be computed as follows [126],

$$\sigma_I^2 = \left(1+\frac{1}{a}\right)\left(1+\frac{1}{b-2}\right) - 1 \tag{3.46}$$

and $\sigma_S^2 = \frac{1}{a}$ and $\sigma_L^2 = \frac{1}{b-2}$ the normalized variances for I_S and I_L, respectively.

Exponentiated Weibull Distribution

In [127], the authors also proposed a different model, with performance comparable to the Log-normal and the Gamma-Gamma models, the so-called Exponentiated Weibull model. This model is developed on the idea that for a non-stationary process, where the signal statistics vary significantly over the time of

observation, a mixture of models, with weighted summation of different distributions, is more suitable. The main concept accounts for one on-axis term and an unknown number of independent off-axis terms. To provide the necessary degrees of freedom to the mathematical model to account for all (with unknown number) uncorrelated terms, a generic average is used as follows [128],

$$I^p = \sum_{j=1}^{m} w_j I_j^p \tag{3.47}$$

where I_j are the Weibull random variables and w_j their weighting factors. Then, the PDF of a random variable I described by the exponentiated Weibull (EW) distribution is given as follows [129],

$$f_I(I;\alpha,\beta,\eta) = \frac{\alpha\beta}{\eta}\left(\frac{I}{\eta}\right)^{\beta-1} exp\left[-\left(\frac{I}{\eta}\right)^{\beta}\right]\left\{1 - exp\left[-\left(\frac{I}{\eta}\right)^{\beta}\right]\right\}^{a-1} \tag{3.48}$$

3.4 FSO Link Analysis

It is absolutely essential to be able to predict an FSO link performance in order to design and operate a practical, and cost-effective system. The system must be able to establish a laser communication link between transmitting and receiving terminals with a certain availability (outage probability) and a probability of error lower than a specified bit-error rate (BER).

A laser beam propagating through the atmosphere is attenuated by absorption and scattering due to the presence of aerosols, dust, smoke, fog, clouds, rain, snow, and atmospheric molecules. The induced photocurrent from the received optical signal is given as follows [130],

$$i = \frac{GP_{REC}\lambda q}{hc} \tag{3.49}$$

where P_{REC} the received power, q the electronic charge and G the avalanche photon detector (APD) gain. The mean-square fluctuations in the signal current i_{sig} is given as follows [130],

$$\langle i^2_{sig_{noise}} \rangle = 2qMFi_{sig}BW_{det} \tag{3.50}$$

where F the excess-noise factor for the APD. Equivalent expressions can be deduced for the currents due to background and the detector noise. If we assume the mean-square fluctuations in the photo detector current due to signal, background, and detector noise are all independent and uncorrelated, the total root-mean-square (rms) noise for a binary 1 or 0 can be calculated as follows [130],

$$\sigma_1 = \sqrt{\langle i^2_{sig_{noise}} \rangle + \langle i^2_{bkg_{noise}} \rangle + \langle i^2_{det_{noise}} \rangle}, \text{ for transmission of binary 1} \tag{3.51}$$

$$\sigma_0 = \sqrt{\langle i^2_{sig_{noise}}(\eta) \rangle + \langle i^2_{bkg_{noise}} \rangle + \langle i^2_{det_{noise}} \rangle}, \text{ for transmission of binary 0} \tag{3.52}$$

where $i^2_{sig_{noise}}(\eta)$ is the extinction ratio of the laser signal, therefore, accounts for the incomplete modulation of the laser from an ON state to an OFF state. The probability of detecting either a 1 or 0 in error is given as follows [130],

$$error1 = \int_{-\infty}^{0.5I} \frac{exp[-(i-I)^2/2\sigma_1^2]}{\sqrt{2\pi\sigma_1^2}} di \tag{3.53}$$

$$error0 = \int_{0.5I}^{\infty} \frac{exp[-i^2/2\sigma_0^2]}{\sqrt{2\pi\sigma_0^2}} di \tag{3.54}$$

Then, the overall BER is the mean of these two probabilities,

$$\text{BER} = (\text{error1}+\text{error0})/2 \tag{3.55}$$

The goal for a lasercomm system is to transmit the maximum number of bits per second over the maximum possible range with the fewest errors. Typically, a “1” denotes a pulse of a transmitted light whereas “0” denotes no optical light transmission. The bit rate of the link is defined as the number of either “1s” or “0s” per second.

The BER of the system depends on the modulation format, and the signal-to noise ratio (SNR), where the noise contributions come from all possible sources which include signal shot noise, dark current noise, thermal/Johnson noise in the electronics following the photo detector, and the background noise. In a turbulent channel, the SNR is a fluctuating term, therefore it is appropriate to extract the average (mean) value which can be expressed as follows [130],

$$\langle SNR \rangle = \frac{SNR_0}{\sqrt{\frac{P_{S_0}}{\langle P_S \rangle} + \sigma_I^2(D) SNR_0^2}} \tag{3.56}$$

where SNR_0 the signal-to-noise ratio in the absence of turbulence, P_{S0} the signal power in the absence of atmospheric effects, $\langle P_S \rangle$ the mean input signal power and $\sigma_I^2(D)$ the aperture-averaged scintillation index. In presence of turbulence, the probability of error for a wireless IM/DD system using OOK signaling technique, is related to the SNR as follows [130],

$$BER = \frac{1}{2}\int_0^\infty p_I(s) erfc\left(\frac{\langle SNR \rangle s}{2\sqrt{2}\langle i_S \rangle}\right) ds \tag{3.57}$$

In the case of a channel modeled by the Malaga pdf the average BER for a system with IM/DD (OOK), can be written in closed-form expression as follows [131],

$$P_b(e) = \frac{2^{a-1}A}{8\pi\sqrt{\pi}} B^{\frac{a}{2}} \sum_{k=1}^{\beta} 2^k B^{\frac{k}{2}} a_k \times G_{5,2}^{2,4}\left(\frac{8R^2P_t^2}{\sigma_N^2} B^2 \middle| \begin{matrix} \frac{1-\alpha}{2}, \frac{2-\alpha}{2}, \frac{1-k}{2}, \frac{2-k}{2}, 1 \\ 0, \frac{1}{2} \end{matrix}\right) \tag{3.58}$$

with $B = \frac{\gamma\beta + \Omega'}{\alpha\beta}$ and the rest of parameters as previously defined in par. 3.3.

The BER of IM/DD with OOK is given by $P_b(e) = p(1)p(e|1) + p(0)p(e|0)$, where $p(1)$ and $p(0)$ the probabilities of sending 1 and 0 bits, respectively and $p(e|1)$ and $p(e|0)$ are the conditional bit error probabilities when the transmitted bit is 1 and 0, respectively. If we also consider that $p(1) = p(2) = 1/2$ and $p(e|1) = p(e|0)$, then the conditioned on h is given as follows [132],

$$P_b(e|h) = p(e|1,h) = p(e|0,h) = Q\left(\frac{\sqrt{2}P_t h}{\sigma_N}\right) \tag{3.59}$$

Where $Q(.)$ the Gaussian Q function. Then, the average BER can be obtained by averaging over the pdf of h [132],

$$P_b(e) = \int_0^\infty f_h(h) P_b(e|h) dh \tag{3.60}$$

Using a K turbulence model for the irradiance and expressing the $K_v(.)$ and erf(.) integrands as Meijer's G functions, a closed-form solution of BER for a FSO system under misalignment-induced fading is given as follows [132],

$$P_b(e) = \frac{2^{\alpha-3}\gamma^2}{\sqrt{\pi^3}\Gamma(\alpha)} G_{6,3}^{2,5}\left(\frac{16P_t^2A_0^2}{\sigma_N^2 a^2}\middle| \begin{matrix} \frac{2-\gamma^2}{2}, \frac{1-\alpha}{2}, \frac{2-\alpha}{2}, 0, \frac{1}{2}, 1 \\ 0, \frac{1}{2}, \frac{-\gamma^2}{2} \end{matrix}\right) \tag{3.61}$$

Assuming a maritime Visible Light Communication (VLC) multi-hop (maritime transceiver) IM/DD (OOK) communication link system using DF relays, which includes k hops and (N-1) DF relays, the equivalent end-to-end SNR at the receiver is given as follows [133],

$$\mu = \left(\sum_{k=1}^{N} \frac{1}{\mu_k}\right)^{-1} \tag{3.62}$$

where $\mu_k = (\eta I_k)^2/N_0$ the instantaneous SNR at the kth hop. Assuming that all hops have the same statistical behavior, an approximated bit error rate (BER) of a DF-based multi-hop FSO is given as follows [133],

$$BER \approx \frac{1}{2}(1 - (1 - 2BER_k)^N) \tag{3.63}$$

where BER_k the BER at the kth hop.

A variety of experimental estimation research works for BER of FSO links under all turbulence strength regions and for differently modeled irradiance fluctuations exist in the literature [134-137].

As described in the previous paragraphs, the atmospheric turbulence causes fading on the propagating laser beam, which within the context of laser communications

affects the reliability of an FSO link. The optical turbulence effects on the beam include the intensity decrease of the received signal and below a threshold value it can even cause outage of the link. The frequency of the signal intensity fluctuations, as compared to the bit rate of the channel, characterizes the fading statistics that describe this channel, as either fast or slow. Fast fading statistics refer to the case where the fluctuations are much more rapid than the bit rate, whereas slow statistics refer to the case where these fluctuations are slow as compared to the bit rate.

Another performance metric, apart from the bit-error-rate (BER), that is useful for both fast and slow fading statistics is the outage probability (P_{out}) of the link. P_{out} is defined as the probability that the instantaneous SNR in the receiver falls below a critical threshold that is determined by the sensitivity of the receiver. Apparently, the lower this probability the more reliable the FSO link. The relationship between the irradiance and the SNR on the receiver is given by the expression [60],

$$I = \sqrt{\frac{\gamma}{\mu}} \tag{3.64}$$

where γ and μ are the instantaneous and the average electrical SNR, defined as $\gamma = (\eta I)^2/N_0$, and $\mu = (\eta\langle I\rangle)^2/N_0$, respectively. Substituting Eq. (3.64) in the PDF for a Gamma-Gamma modeled channel [Eq. (3.32)] and with a power transformation of I, we obtain the PDF of the irradiance in terms of the instantaneous SNR on the receiver [60],

$$f_\gamma(\gamma) = \frac{(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} \frac{\gamma^{((\alpha+\beta)/4)-1}}{\mu^{((\alpha+\beta)/4)}} K_{\alpha-\beta}\left(2\sqrt{\alpha\beta\sqrt{\frac{\gamma}{\mu}}}\right), \quad I > 0 \tag{3.65}$$

Therefore, as stated before, if we assume γ_{th} to be the threshold value of the instantaneous SNR in the receiver, we define the outage probability for a channel modeled with the Gamma Gamma distribution as follows [60],

$$P_{out} = Pr(\gamma \le \gamma_{th}) = F_{\gamma}(\gamma_{th}) =$$

$$\frac{(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{\gamma_{th}}{\mu}\right)^{\frac{a+\beta}{4}} G_{1,3}^{2,1}\left(\alpha\beta\sqrt{\frac{\gamma_{th}}{\mu}}\middle|\begin{matrix}1-\frac{a+\beta}{2}\\ \frac{a-\beta}{2},\frac{\beta-a}{2},-\frac{a+\beta}{2}\end{matrix}\right) \tag{3.66}$$

Based on the Shannon theorem, the capacity of a communication channel is given as follows [60],

$$C = Blog_2(1+\Gamma) \tag{3.67}$$

where B the bandwidth and Γ the SNR at the receiver. However, since the SNR fluctuates randomly, it is considered a random variable and Eq. (3.67) does not have a significant meaning, thus we have to define the average capacity for a fast fading channel and the outage capacity for a slow fading channel. The expression for the evaluation of the average capacity of an optical link modeled by a Gamma Gamma PDF is given as follows [60],

$$\langle C\rangle = \frac{B(a\beta/\sqrt{\mu})^{\frac{a+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)\ln(2)}\int_0^{\infty}\ln(1+\gamma)\gamma^{\frac{\alpha+\beta}{4}-1}K_{\alpha-\beta}\left(2\sqrt{\alpha\beta\sqrt{\frac{\gamma}{\mu}}}\right)d\gamma \tag{3.68}$$

Other expressions for the average (ergodic) capacity of dual-hop FSO communication system employing amplify-and-forward (AF) relaying [138] and of a MIMO FSO system [139] exist.

As previously mentioned, the appropriate metric for the estimation of the performance of the FSO channel for the cases of slow fading statistics is the outage capacity. Assuming again a Gamma Gamma modeled channel, the expression for the outage (normalized) capacity is given as follows [60],

$$f_{\tilde{C}} = \frac{2^{\tilde{C}}\ln(2)(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)}\frac{\left(2^{\tilde{C}}-1\right)^{\frac{\alpha+\beta}{4}-1}}{\mu^{\frac{\alpha+\beta}{4}}}K_{\alpha-\beta}\left(2\sqrt{\alpha\beta\sqrt{\frac{2^{\tilde{C}}-1}{\mu}}}\right) \tag{3.69}$$

The outage probability estimation of an FSO link is extremely significant in the case of a hybrid RF/FSO link. In that case, the FSO sub-system of the hybrid system

initiates the transmission with the selected link as long as its instantaneous signal-to-noise ratio (SNR) at the optical receiver is above a certain threshold. When the SNR falls below this threshold, the system switches to the RF sub-system, while putting the FSO sub-system on a standby mode. The outage probability and average symbol error rate (SER) of a proposed hybrid FSO/RF system was extensively analyzed in [140] and the following expression was given,

$$P_{out} = F_{\gamma}^{FSO}(\gamma_{th}) F_{\gamma}^{RF}(\gamma_{th}) \tag{3.70}$$

where,

$$F_{\gamma}^{FSO}(\gamma_{th}) = \prod_{j=1}^{N_f} F_j^{FSO}(\gamma_{th}) \tag{3.71}$$

the cumulative distribution of the instantaneous SNR at the FSO receiver.

In [141] the outage probability of an FSO link was estimated based on actual visibility data collected from different sites of India.

Another outage probability expression was proposed in [142], for an FSO system using multi-pulse pulse position modulation (MP-PPM) over Gamma-Gamma turbulence channel,

$$P_e(h) \leq \frac{\binom{M}{k}-1}{2\sqrt{\pi}} G_{1,2}^{2,0}\left(\mu^2 h^2 \middle| \begin{matrix} 1 \\ 0, \frac{1}{2} \end{matrix}\right) \tag{3.72}$$

where G is the Meijer G-function, h the optical intensity fluctuations resulting from the atmospheric attenuation h_a, atmospheric turbulence h_t, and pointing error effects h_p, μ the mean value of the beam displacement due to pointing errors and

$$\binom{M}{k} = \frac{M!}{k!(M-k)!} \tag{3.73}$$

the number of unique symbols generated by k pulses among M slots.

In the case of a relay-assisted FSO channel (A, B and C nodes), the outage probability can be derived for a bit-detect-and-forward (BDF) protocol using time diversity, assuming a statistical channel model as follows [143],

$$Y_{BDF} = \frac{1}{2} X I_{AC} + Z_{AC} + \frac{1}{2} X^* I_{BC} + Z_{BC} \quad (3.74)$$

where $Z_{AC}, Z_{BC} \sim N(0, N_0/2)$ and X^* a random variable corresponding to the information detected at node B, therefore, equal to X when the bit has been detected correctly at B and $X^*=d_E - X$ when it has been detected incorrectly and d_E an Euclidean distance, $d_E = 2P_{opt}\sqrt{T_b \xi}$, with P_{opt} the average optical power transmitted, T_b the bit period and ξ the square of the increment in the Euclidean distance due to the use of a pulse shape of high peak-to-average optical power ratio. The outage probability which corresponds to the BDF cooperative protocol is given as follows [143],

$$P_{out}^{BDF} = P_{out}^0 \cdot (1 - P_b^{AB}) + P_{out}^1 \cdot P_b^{AB} \quad (3.75)$$

where P_{out}^0 and P_{out}^1 the outage performance when the bit is correctly and incorrectly detected at B and P_b^{AB} the BER corresponding to the A-B link.

In [144], the performance of different MIMO FSO communication schemes has been analyzed. In particular, different FSO diversity methods have been compared using the outage probability in a gamma-gamma modeled turbulence channel. The diversity gains of MIMO repetition coding (RC), MIMO transmit laser selection (TLS), MIMO all-active relaying (AR) and MIMO selective relaying (SR) was derived and was proved that MIMO-RC and MIMO-TLS, on one hand, and MIMO-AR and MIMO-SR, on the second hand, achieve the same diversity gain. By comparing the MIMO and The main result of the study was that it is always better, from a diversity gain point of view, to add more apertures to the source and destination rather than adding more relays in their vicinity despite the fact that the fading variance along FSO links decreases with the distance.

3.5 Fading Mitigation Techniques

As previously mentioned, the atmospheric turbulence causes irradiance fluctuations and/or beam wander effect of the received signal, something that can lead to an increased BER in the system. These effects, result in deep signal fades lasting from 1 to 100 μs. Assuming a link operating at a nominal data rate of 1 Gbps, this fading duration would result in a loss of 10^5 consecutive bits. This loss would definitely degrade the performance and availability of the link, therefore effective mitigation techniques must be employed in order to compensate this issue.

3.5.1 Aperture Averaging

The received beam will intensively fluctuate due to turbulence, if the size of the receiver aperture is much smaller than the beam diameter. The atmosphere will provide an acceptance angle for the receiver assuming detector at the receiver to be omnidirectional and only the scattered optical signal from turbulent cells within this acceptance cone will contribute to the received signal power. The largest acceptance cone will be for smallest eddy size (inner scale), l_0, and is given as follows [4],

$$\theta_{max} \cong \frac{\lambda}{l_0} \tag{3.76}$$

In case that the maximum width of the cone, given by $R\theta_{max}$, is bigger than l_0, the acceptance cone may contain smaller cells and the received power will fluctuate much more as long as receiver aperture is less than beam diameter. Increasing the receiver aperture in order to overcome irradiance fluctuations is called aperture averaging. A parameter used to quantify the power fluctuations reduction by aperture averaging is called aperture averaging factor, A_f, defined as the ratio of normalized variance of the irradiance fluctuations from a receiver with aperture diameter D_R to that from a point receiver [4],

$$A_f = \frac{\sigma_I^2(D_R)}{\sigma_I^2(0)} \tag{3.77}$$

where $\sigma_I^2(D_R)$ and $\sigma_I^2(0)$ the scintillation indices for a receiver with aperture D_R and a point receiver ($D_R \sim 0$), respectively.

3.5.2 Spatial Diversity

Increasing the receiver aperture size is not always the best solution because, doing so beyond a certain point will lead increased background noise. Therefore, the best alternative is the replacement of the bigger size aperture diameter with an array of small apertures, in either end of the link, in sufficient distance apart, but certainly greater than the coherence length of the atmosphere, r_0, so that multiple beams are independent and at least uncorrelated. This technique, is known as spatial diversity and the systems that apply it are called either single-input-multiple-output (SIMO), multiple-input-single-output (MISO) or multiple-input-multiple-output (MIMO). A single beam propagating through the atmosphere will end up splitting into various small beam segments. These segments will then independently reach the receiver either in or out of phase with respect to each other, causing signal fading. If instead multiple independent and uncorrelated beams are used, then any overlapping at the receiver will result in addition of power from different beams. Consequently, the deep fading probability will be reduced significantly. Another advantage of this technique, is that it allows for lower power transmission leading to a safer laser beam for a human eye. The improvement in the performance of an FSO that applies spatial diversity technique, expressed as the summed output power is given as follows [4],

$$I_r = \eta \sum_{j=1}^{N} \left(I_{s,j} + I_{n,j} \right) \tag{3.78}$$

where N the number of the statistically independent detectors, η the optical to electrical conversion efficiency and $I_{s,j}$ and $I_{n,j}$ are the signal and noise currents corresponding to the j^{th} receiver, respectively. Additionally, the mean rms *SNR* is given as follows [4],

$$\langle SNR_N \rangle = \sqrt{N} \langle SNR_1 \rangle \tag{3.79}$$

where $\langle SNR_1 \rangle$ the mean *SNR* of a single detector receiver. Eq. (3.79) designates that the output *SNR* from the N independent detectors can improve the system performance by a factor of $\sqrt{N}$. The number of the detectors or the number of transmitted beams required to achieve a given BER depends upon the strength of the atmospheric turbulence but in principle, the received irradiance statistics are improved with the increase of both numbers. In [145], an FSO system with one output and two inputs has been used, with two tunable iris diagrams for effective receiver aperture diameter selection, to measure the aperture averaging and the spatial diversity effect in a controlled laboratory environment, for aperture diameters of 10mm, 15mm, 20mm, 30mm and 35mm and for weak turbulence conditions. The experiment results, indeed, verified the theoretical beneficial effect of both aforementioned techniques.

3.5.3 Adaptive Optics

Another fading mitigation technique that helps to deliver an undistorted beam through the atmosphere is the employment of adaptive optics, a closed-loop control which pre-corrects the beam by putting the conjugate of the atmospheric turbulence before transmitting it into the atmosphere, reducing that way the spatial and time fluctuations. An adaptive optics system consists of wavefront sensor to measure the closed-loop phase front, corrector to compensate for the phase front fluctuations, and a deformable mirror that is driven by a suitable controller. In [146], an adaptive optical

transceiver telescope was used to investigate the possibility of correcting wavefront aberrations under strong atmospheric turbulence conditions over a distance of several kilometers. A fiber laser was connected to a fiber positioner within the telescope, which acted as the transmitter by sending a laser beam at 1550 nm through the turbulent atmosphere to a retro reflector mounted on the top of a water tower at a distance of 2.33 km. The reflected laser light was received and focused onto the fiber tip, guided to a photo detector by a splitter and recorded by a PC in order to be used as the feedback signal for the adaptive optics controller, which controlled the fiber-tip positioner and a six-channel adaptive mirror using a stochastic parallel gradient descent optimization algorithm. The experiments showed that for different turbulence conditions, the low-order adaptive optics system in the transceiver telescope, controlled using an SPGD algorithm, increased the average returned intensity of the back-reflected light from the retro reflectors on the water tower.

3.5.4 Semiconductor Optical Amplifier

Finally, another novel fade mitigation technique, applicable on outdoor optical wireless systems. The key idea is to leverage the nonlinear power-dependent gain properties of a semiconductor optical amplifier (SOA), in order to provide unbalanced amplification between faded and non-faded instances of the optical wireless signal. In [147], this power equalization process to smooth out fade-induced power fluctuations and drastically reduce the probability of the system being in a fade state, was demonstrated. The results predicted that the fade probability, under medium to strong turbulence conditions governed by gamma-gamma statistics, can be reduced by over 80% when the SOA is introduced at the optical wireless receiver. It was also shown that the duration of remaining fades was reduced by a sizeable percentage, and a percentile

reduction of the average fade duration of over 85% could be achieved at the SOA output. In [148], it was also shown that with respect to first order statistics, a SOA-assisted system exhibits significantly lower average BERs, higher link availabilities. In the same way, the sensitivity improvement can be utilized to partially or fully compensate for the fade margin in medium-to-strong turbulence for link lengths that do not exceed several hundreds of meters. With respect to the second order statistics, the presented analytical results on the amplified system demonstrated a drastic improvement of the average fade duration (AFD) irrespective of the link length.

CHAPTER 4

ML for Laser Communications

4.1 Introduction

The fields of Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL), have been evolved and proved tremendously successful over the last few decades. Traditionally, DL is considered a sub-field of ML which is considered a sub-field of AI. A very general definition for ML is that it is a set of methods and algorithms whose goal is the relationship prediction of an input and an output by means of training certain parameters [149]. In principle, these algorithms constitute a black box with regard to the specific task they are intended to complete, rather they are trainable based on data. This fact, does not mean that every algorithm is appropriate for every kind of problem or equally effective. This decision, remains up to the user to make it. ML approach has several advantages; ML algorithms can be utilized where no sufficient domain knowledge is available to build expert systems, they can be used for many different tasks that are similar in structure, only after being trained for the new problem.

Machine learning belongs to the computer science field and basically trains a computer from data or from the interaction with a real or virtual environment via statistical methods. It can be divided into three sub-categories, based on the training

nature they provide to the computer, the supervised, the unsupervised and the reinforcement learning.

4.1.1 Supervised Learning

Supervised learning utilizes a given dataset, such as an input where the output is known, in order to train a ML algorithm, which, thenceforth, is able to label unknown data. Input variables -or predictors- are denoted with an *X* whereas the output variable or response is denoted with a *Y*. Usually, the predictors are more than one and it is assumed that a general relationship among them exists [150],

$$Y = f(X) + \varepsilon \tag{4.1}$$

where *f* is an unknown function and ε a zero mean random error, statistically independent to *X*. Supervised learning is focused on estimating this function *f* for prediction and inference purposes. Having estimated *f*, one can predict the output for a specific set of input variables. The accuracy of this prediction depends upon the so-called reducible and irreducible errors. The reducible error refers to the inherent inaccuracy due to any inappropriateness of the function *f*, whereas the irreducible error is included otherwise since *Y* is a function of ε. On the other hand, inference refers to the understanding of the relationship between inputs and outputs and the impact of certain changes in inputs on their corresponding output [150]. Problems with a quantitative response are called regression problems whereas problems with a qualitative response are called classification problems.

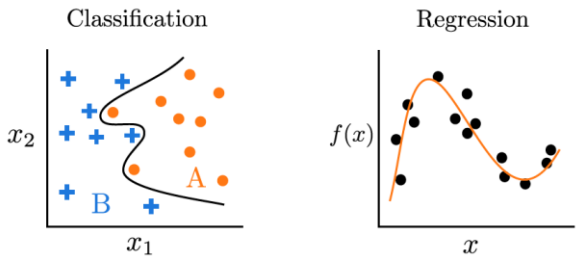


Figure 4.1. The two main paradigms for the supervised learning, classification and regression [149].

A classic example is image classification. For a given set of images depicting cats and dogs, a trained ML algorithm should be able to distinguish a new picture whether it is a cat or a dog. The training process follows a division of the entire dataset into three parts, the larger is for training and the other two for validating and testing the algorithm. The ultimate goal is to minimize the error of the predicted values, something that rates the performance of the trained algorithm. In case of having better performance for the test set than for the training set, we then face overfitting of our data. Overfitting is a negative sign for our training process and indicates a highly complex model that follows the errors too closely, thus is biased and will perform poorly with new data. The performance on the validation set, is used after the training is completed, for comparing different algorithms, machine learning models architectures and hyperparameters. Apart from a classification problem that distinguishes among two or more discrete values or features, there are regression problems, where the solution has a continuous value. Since the algorithm is not explicitly programmed in supervised learning, insufficient data or falsely labeled data will lead to an incorrect algorithm, therefore the dataset quality is very important.

Most ML algorithms are categorized either as parametric or non-parametric, that is a known or not form of *f*. Parametric methods involve a two-step approach for model development, initially the assumption of the form of function *f* and the training procedure of this function. Non-parametric methods, on the other hand, do not make explicit assumptions about the functional form of *f*. Instead they seek an estimate of *f* that gets as close to the data points as possible without being too rough or wiggly.

4.1.2 Unsupervised Learning

In contrast with the supervised learning, unsupervised learning uses unlabeled data and its goal is to find patterns that can adequately describe the data, called features, and keep only them to develop a new dataset with reduced redundancy. In the previously mentioned example of the difference between pictures with cats and dogs, an unsupervised learning process would lead to the recognition that pictures of dogs have common features and there is contrast with the pictures with cats. This process is also called clustering, where cluster refers to the collection of pictures of either dogs or cats. This procedure would also be applicable for outlier or anomaly detection on a dataset. For example, if we add a picture depicting a fox, then a third cluster would be recognized, detecting that way the anomaly. Another common application for unsupervised learning is the face recognition, where pictures are translated into a domain in which they can be compared mathematically. A face recognition system, uses a dataset of human faces, to learn how to extract features that preserve the most distinguishable patterns of a human face. That way, the system translates every human face into a vector representation of these most important features. Therefore, by comparing these vectors, we can distinguish different faces.

4.1.3 Reinforcement Learning

Reinforcement learning, is a different method where an algorithm learns by rewarding itself for actions that favor a desired outcome and is based on a Markov decision process [151]. A famous example, is the AlphaGoZero algorithm, trained by Google Deepmind team, to learn the game of Go, by taughting itself through self-play and reinforcing/rewarding moves which led to victory.

4.1.4 Assessing Model Accuracy

In order to estimate the function *f* that best fits a data set we have to be aware of the special attributes that possesses in terms of accuracy and interpretability. For example, a linear regression algorithm may be very clear to interpret but also very restrictive, especially when a non-linear relationship between input and output actually holds. On the other hand, algorithms such as support vector machines or boosting methods, could lead to such complicated models that may be very difficult to interpret [150]. However, if prediction accuracy is the desired outcome, models with high flexibility seem to be the best choice. This is not always the case though, since overfitting of the data will, despite decreasing the training error, increase the testing error. Therefore, a careful bias – variance tradeoff has to take place before selecting the most appropriate model. The term variance refers to how much *f* will change if the same model is used on a different data set, whereas bias refers to the error introduced by modeling a real-life problem. In general, more flexible models tend to increase the variance and decrease the bias [150].

4.2 Machine Learning Algorithms

In this section, we will shortly introduce some of the most common ML methods, either for supervised or unsupervised learning, such as random forests, support vector machines, artificial neural networks and k-means.

4.2.1 Random Forests

The first algorithm, namely Random Forest (RF), belongs to the tree-based methods which are suitable both for regression and classification problems. What these

types of algorithms do, is segmenting the predictor space of the data into small regions and use the mean or mode value of these regions to make the final prediction of the response value. Since the splitting rules of these regions can be summed up in a tree, these algorithms are known as decision-tree methods [150]. The simple decision-tree methods have proved not so effective in terms of accuracy; therefore, more advanced methods exist to deal with more complex data, i.e. random forests and boosting. The advantages of decision-tree methods include: 1) easy to explain, 2) they mimic better the human decision-making process, 3) they can be easily depicted in a graph and 4) they can handle qualitative predictors without requiring dummy variables [150]. The training of a decision tree is translated into finding the best set of decision rules for training set division, which will be used towards new data. In general, decision trees are overfitting and are highly vulnerable in training dataset changes, since it is difficult to generalize these decision rules in other problems. The two basic steps of the DT process include the following [29]:

a. Predictor space division into *J* distinct and non-overlapping regions, $R_1, R_2, \ldots R_J$. The criterion to determine the optimal split point is to minimize the RSS given by,

$$\sum_{j=1}^{J} \sum_{i \in R_j} \left(y_i - y_{R_j} \right)^2 \tag{4.2}$$

where y_{Rj} is the mean response for the training observations in the *j-th* box.

b. For every observation falling into a certain region, the prediction emerges from the response mean value based on the training observations that belong to the same region.

A general process followed in decision tree based statistical learning methods in order to achieve a lower variance, is bagging, in other terms, to take several training subsets from the entire dataset, create a model for each one of these subsets and average

the predictions of each model. An improvement of bagging is random forest, that decorrelates the trees (subsets). Instead of considering the whole set of *p* predictors for each split criterion, a random sample of *m* predictors is only considered. Typically, we choose $m \approx \sqrt{p}$, that is, the number of predictors considered at each split is approximately equal to the square root of the total number of predictors. This random selection is repeated in each and every split step, meaning that a new sample is chosen. Therefore, a strong predictor will not be considered in each tree and consequently we will avoid the correlation among them, hence less variable.

The method of random forests was originally proposed by Ho [152], who extended the notion of the single decision tree model to an ensemble of individual decision trees, to improve the generalization properties of the method. The main idea is that each individual decision tree is trained only on a random subset of the training data, or receives only a randomly selected subset of the data features as input data. The overall label of a class in the data is obtained by selecting the most prevalent within the individual predictions, something that improves the generalization capability of the decision tree method.

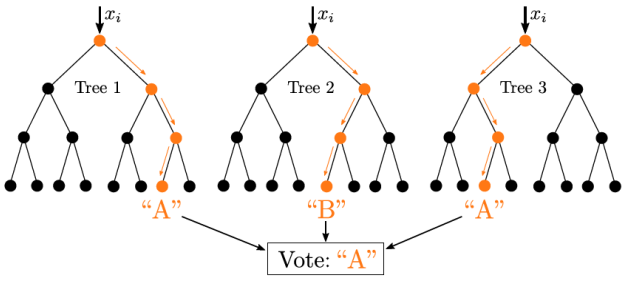


Figure 4.2. A schematic view of the random forest approach [149].

Another general approach for prediction accuracy improvement is boosting. In contrast with bagging, this approach follows a sequential procedure to train the model, by using information from the previously trained tree [150]. Specifically, the first

decision tree is fitted to the training data set, the second is fitted using the residuals of the previous one, then this new tree is added to the fitted function in order to update the residuals. This procedure is followed by fitting rather small trees -few number of terminal nodes- to the residuals and that way we improve the model in areas where it does not perform as well. An additional shrinkage parameter λ exists, that allows more and different trees to be fitted in certain areas where higher residuals appear.

4.2.2 Support Vector Machines

The support vector machine (SVM) algorithm is based on the selection of a suitable hyperplane that will split the input features of every class in the right side of the hyperplane, a flat affine subspace of dimension *p*-1, for a *p*-dimensional space. The mathematical definition of a hyperplane in a *p*-dimensional space is given as follows,

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p = 0 \qquad (4.2)$$

where any $X = (X_1, X_2, \ldots, X_p)^T$ for which (4.2) holds is a point on the hyperplane.

The simplest case of a SVM, is the maximal margin classifier, where the selection of the hyperplane is selected based on the maximum perpendicular distance of each training point. SVM, developed in the 1990s, were primarily applied for supervised classification tasks, but can also be used in regression as well as unsupervised tasks, i.e. clustering. In case of classification, the new data can then be classified based on the side of the hyperplane they reside to, i.e a point *X* that does not verify Eq. (4.2). Because of the apparent high nonlinearity of most of the training datasets, high-dimensional feature spaces were proposed, in which the linear separation of the data will be always achievable, as shown in Figure 4.3. A certain trick is used in order to keep the computational cost in such high dimension, low, the so-called *kernel trick*. By selecting and applying a suitable kernel function, the coordinates of the data points never have to

be computed explicitly in the high-dimensional feature space, while computations can be carried out implicitly. In other words, the enlarged feature space is in fact linear, but in the original feature space, the decision boundary has the form of a quadratic polynomial with non-linear solutions.

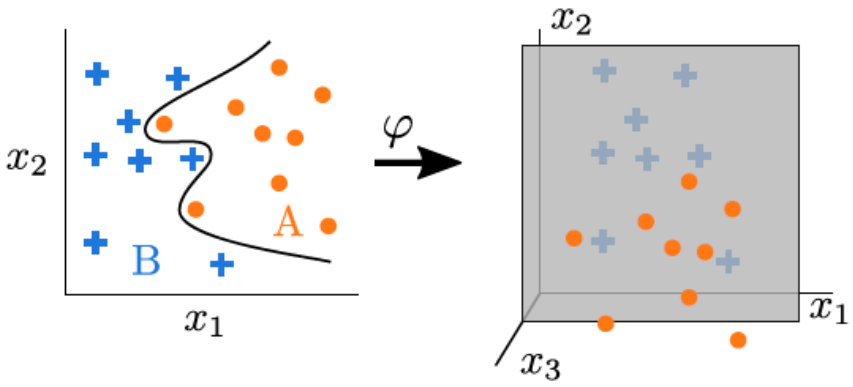


Figure 4.3. The idea of the higher-dimension feature space used in SVM algorithm [149].

4.2.3 K-nearest Neighbors

The K-nearest Neighbors (KNN) algorithm is a simple model to implement for regression and classification problems. To evaluate the effectiveness and adequacy of any algorithm we need to account for output interpretation ease, calculation time, and power of prediction. The main advantages of KNN algorithm are the easiness of its output interpretation and the significantly low time for calculation. KNN is a simple algorithm that stores all available cases and predicts the numerical target based on a similarity measure, e.g., distance functions. A simple implementation of KNN regression is to calculate the average of the numerical target of the *K* nearest neighbors. Another approach uses an inverse distance weighted average of the *K* nearest neighbors. KNN regression uses the same distance functions as KNN classification. Choosing the optimal value for *K* is best done by first inspecting the data. In general, a large *K* value is more precise as it reduces the overall noise; however, the compromise is that the

distinct boundaries within the feature space are blurred. Cross-validation is another way to retrospectively determine a good *K* value by using an independent data set to validate the *K* value. Mathematically, the KNN regression algorithm can be expressed as the average of all training points in N_0,

$$f(x_0) = \frac{1}{K}\sum_{x_i \in N_0} y_i \tag{4.3}$$

where x_i is the input and y_i is the output of the model, x_0 a prediction point and N_0 the K points that are closer to x_0. A similar process is followed for a clustering problem. Initially, we need to choose the *k* data points at random, as centroids, then all data points are assigned to the respectively nearest cluster centroid, and thus each data point is assigned to one of the *k* clusters, and finally an update of the cluster centroids is taking place by computing the center of all points assigned to the respective cluster. The second and third step are repeated till every cluster is finalized.

4.2.4 Neural Networks

A Neural Network (NN) mimics the functionality of the human brain, where the human neurons are represented by the nodes and the various connections among them by adjusted weights among the nodes. A node (perceptron) consists the fundamental component of a neural network which receives several x_i input signals with a corresponding w_i weight and an *b* applied bias, a factor associated with the storage of information. The sum of these signals enters the node where a certain activation function takes place and process the information according to the type of the function. All these complex combinations of input signals, weights, activation functions and biases for each layer, allows for a sufficient modeling of highly non-linear relationships. A NN structure is defined in two dimensions, namely the number of layers and the number of nodes in each layer. A NN with only one layer is called a single perceptron

model whereas a network with multiple layers is called multiple layer perceptron model or a deep neural network. The best number for either dimensions is a matter of research and different in each model. The perceptron transfer function can then be expressed as [153],

$$y = f\left(\sum_{j=1}^{n} w_{ij} x_i + b\right) \tag{4.4}$$

The internal layer perceptrons, those in between the input and output layers, receive the summed weighed values of the input parameters and provide the "input" signal for the output layer. All layers are fully connected to each other. The architecture of a multi-layer perceptron model, i.e. shallow neural network, and the input-output mapping can be represented as follows [153],

$$y = g\left(\sum_{j=1}^{m_1} w_j f\left(\sum_{i=1}^{n} w_{ij} x_i + b\right) + c\right) \tag{4.5}$$

where g is the transfer function and c the bias for the output layer. Other ANN architectures could include more than a single hidden layer. The transfer function to characterize the hidden layer, can be one of the following functions, [153],

a. Tangent Hyperbolic Function: $f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$

b. Sigmoid Function: $f(x) = \frac{1}{1+e^{-x}}$

c. Rectified Linear Unit (ReLU) Function: $f(x) = \max(0, x)$

d. Gaussian Function: $f(x) = e^{-x^2}$

e. Linear Function: $f(x) = x$

A NN can represent a much wider range of functions than a single neuron, in cost of having more free parameters. It can represent arbitrarily complex functions by increasing the number of hidden nodes, which is equivalent to increasing the number of free parameters. The amount of free parameters determines its computational

complexity and with an increasing number of parameters, a NN becomes prone to overfitting.

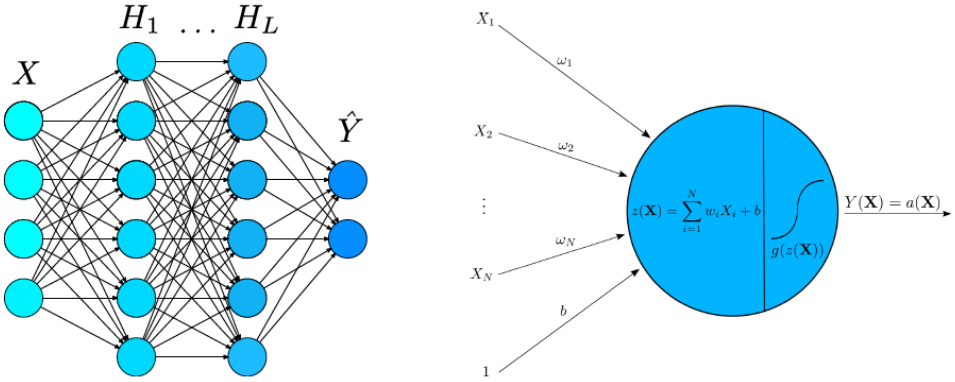


Figure 4.4. A general view of a feed-forward NN with two hidden layers (left) and the design of a perceptron neuron (right) [149].

During the training of a NN through a supervised learning process, the samples of the training set are used to adjust the weights of the neural network, such that the output vector becomes similar to the target vector for every sample of the training set,

$$f_\theta\left(x^{(n)}\right) = y^{(n)} \approx t^{(n)} \tag{4.6}$$

where $x^{(n)}$ a set of input vectors and $t^{(n)}$ their corresponding outcome. The first step to do so is the loss function definition, which is application specific. For classification problems, we typically choose the cross-entropy loss, whereas for regression problems is the mean squared error (MSE) [153],

$$L_{cross-entropy}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\left[-\sum_{d=1}^{D} t_d^{(n)} log\left(y_d^{(n)}\right)\right]_\theta \tag{4.7a}$$

$$L_{MSE}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\left[\frac{1}{2D}\sum_{d=1}^{D}\left|t_d^{(n)} - y_d^{(n)}\right|^2\right]_\theta \tag{4.7b}$$

where N is the number of samples in the training set and D is the number of output nodes. The loss function minimization typically is achieved by the gradient descent optimization, where the gradient of the loss function w.r.t. the weights of the neural

network is computed. A learning rate is also set up to determine the speed of algorithm learning. The training process of a single weight update, is called batch gradient descent, because only a single batch was involved. The term batch refers to a small subset of the whole training set and is used for a more rapid and flexible training process. Another common term is epoch, which is the number of times a NN has been trained with all batches. The whole training process goal is to find the local minima of the loss function, however, whether the trained neural network performs sufficient is determined with the validation and test phase. If the loss of test set increases while the loss of the training set decreases, then our NN is overfitting on the training set.

4.3 Machine Learning for FSO Modeling

Machine Learning use for Optical Communications research is a recently emerged topic that will definitely continue to increase rapidly. The value of this approach is derived mainly from the unique characteristics of the FSO research field rather than the applied ML algorithms. By reviewing the literature on this topic, we can notice that a number of ML algorithms or a number of specific FSO-related applications have not been extensively leveraged. After all, a balanced mix of both fields must be achieved, in order to get the best results from this synergy [154].

An important challenge in using machine learning techniques for scientific applications is to be able to use domain knowledge and data-driven algorithms in synergy. Machine learning algorithms have already been exploited by the optical communications community and in particular the free space optical communications community. For example, in [155], a data-driven fiber channel deep learning (DL) modeling method is introduced in an optical communication system. Specifically, a bidirectional long short-term memory is selected to perform fiber channel modeling for

on-off keying (OOK) and pulse amplitude modulation 4 signals. A deep learning based -convolutional neural network- atmospheric turbulence compensation method to correct the distorted vortex beam and improve the performance of orbit angular momentum (OAM) multiplexing communication is presented in [156]. FSO related channel modeling, is studied using DL techniques in [157] for all turbulence strength regimes and results published therein indicate that DL can provide performance that is reasonably close to the perfect channel estimation scheme. The ability to mitigate the negative effects of atmospheric turbulence for FSO systems performance is studied in [158-160] using ANN, Generative Machine Learning (GML) and convolutional neural networks (CNN) through simulated and experimentally obtained data. ML techniques using DT, RF and ANN, for C_n^2 estimation, is presented in [161,162] and compared against macroscopic meteorological parameters. Other ML algorithms have also been used to predict RSSI for a hybrid RF/FSO detector by developing both regression and classification models [163, 164]. Finally, an overview of the, as expected, next generation of the ANN, the so-called optical neural networks (ONNs) and previous studies on the field are reviewed in [165]. The novelty of the current study lies in the comparison of various classical ML algorithms that are presented for the first time in predicting RSSI measurements, especially for the particular domain of interest, Piraeus, Greece. The contribution of this paper is that it a) provides an extension of previous related works in terms of the utilized ML algorithms i.e. KNN and ANN which both performed sufficiently enough and b) provides a unique experimental data analysis in terms of the terrain of the link i.e. maritime which exhibits different characteristics from a terrestrial one.

In [166], several evolutionary algorithms (EAs) in conjunction with the Levenberg - Marquardt (LM) back-propagation algorithm are used, in order to train

different ANNs and in particular the L-SHADE algorithm, which self-adapts control parameters and dynamically adjusts population size. Five new hybrid training methods were designed by combining LM with self-adaptive Differential Evolution (DE) strategies. These new training methods obtained better performance to ANN weight optimization than the original LM method. The concept of atmospheric turbulence characterization by using laser light backscattered off a moving unresolved target or a moving target with a glint is considered and analyzed through wave-optics numerical simulations in [167]. This technique is based on analysis of the autocorrelation function and variance of the power signal measured by the target-in-the-loop atmospheric sensing (TILAS) system composed of a single-mode-fiber-based optical transceiver and the moving target. It is shown that the TILAS received power signal autocorrelation function strongly depends on the turbulence distribution and is weakly sensitive to the turbulence strength.

Optical turbulence severely affects different types of optoelectronic and adaptive optics systems. Direct measurements of the refractive index structure parameter are incomplete, for many types of environments, particularly in the maritime, because it is difficult and expensive to deploy the appropriate instrumentation [168]. To fill this gap, a backpropagation neural network (BPNN) approach has been proposed for the forecasting and verification of optical turbulence profiles in the offshore atmospheric boundary layer. In [169], the performance of this BPNN approach is further evaluated against the Holloman Spring 1999 thermosonde campaigns (HMNSP99) model for outer scale, and the Hufnagel/Andrew/Phillips (HAP) model for a single parameter. The results have shown that the agreement between the BPNN approach and the measurements is very close. On the other hand, in [170], the atmospheric turbulence

refractive index structure parameter is predicted using deep neural network (DNN)-based processing of short-exposure laser beam intensity scintillation patterns.

4.4 Machine Learning Applications for Fading Mitigation

A vortex beam (VB) is a structured light which carries orbital angular momentum (OAM) and presents a “doughnut” intensity distribution, due to the phase singularity in the beam cross-section. The VBs possess many significant optical properties and show promising applications in optical trapping, optical microscopy and imaging, quantum information, optical storage, and optical communications [171]. In particular, by multiplexing OAM modes of VBs, the capacity of an optical communication link can reach the Tbps-class. However, the spiral phase of VB is very sensitive to the transmission environment and can easily be distorted. The demodulation and demultiplexing technologies for OAM communications have been studied, by utilizing the information capture and automatic classification capabilities of deep learning [172]. The turbulence correction in conjugate superposition OAM modes has been realized through deep learning based compensation method and proved to be more accurate and faster in terms of correction ability than AO system. In [171, 173], a novel deep learning based turbulence compensation method is introduced, for correcting the distorted VB and improving the performance of OAM multiplexing communication. Gaussian probe beam (GPB) is introduced, as information extraction object, in order to simultaneously compensate the turbulence distortion for multiple VBs. A convolutional neural network (CNN) model, through supervised learning, can learn to produce compensation phase screen through the intensity distribution of GPB in randomly changed turbulent environments and result in a significantly improved mode purity of the distorted VB. By achieving a communication link through OAM multiplexing, the BER of each OAM

channel is reduced, after it is counterbalanced by the trained CNN, by almost two orders of magnitude in moderate to strong turbulence conditions. The results demonstrated that the proposed CNN model predicts accurately the turbulent phase and showed good generalization ability in compensating quickly and accurately the distorted VBs. In [158], an optical feedback network is designed, utilizing an ANN scheme which only relies on measuring the intensity profile of the distorted modes and demonstrated its turbulence effects correction capabilities via simulations. Initially, a laser beam directed onto a spatial-light modulator (SLM) with a given phase profile is simulated, used to convert each incident optical beam from a Gaussian mode profile into a Laguerre-Gauss mode. The resulting mode profile, has then a "petal pattern" of bright and dark spots, in a circular configuration. In order to simulate the beam, the Fresnel Transfer Function (TF) propagator and the Fresnel Impulse Response (IR) propagator was used. Additionally, the addition of this unsupervised learning scheme can be extended to demodulate more complex optical profiles which are difficult to be labeled and classified with current supervised techniques [174].

CHAPTER 5

Experimental Statistical Modeling of FSO Performance in Maritime Environment

This chapter includes the results that have been published in the peer-reviewed journals [J1] and [J2].

5.1 Introduction

Free space optical (FSO) communications is a significant application of the laser technology initially developed in 1960. Since then, much research has been conducted into FSO communications and different applications have been demonstrated, including terrestrial, maritime, space and deep space applications. Despite the initial uncertainties about its potential, the ongoing development of optoelectronic devices and its proven success in military applications provided the required boost to continue investments in the field [175]. The ever-increasing demands on reliable and high-speed data transfer lead the way to the current research efforts to advance the optical wireless communications (OWC) technology so it can be exploited by both small and light platforms.

FSO technology offers significant advantages over its RF counterpart and benefits applications in platforms with increased weight and space limitations. Therefore, platforms with size, weight and power constraints can benefit from more compact communication systems while achieving higher performance. FSO technology

applications span from fixed point to point networks, e.g. LAN, last mile access, fiber backup, to high speed moving platforms, e.g. UAVs, aircrafts. The applied laser technology level allows not only for terrestrial but also for underwater and space applications. The principal advantages of FSO communications include increased bandwidth, greater security and immunity, lower cost of installation and, finally, no license restrictions. However, FSO communications are susceptible to various atmospheric effects and phenomena, including molecular and aerosol absorption and scattering as well as atmospheric turbulence. The performance of a laser communication link is highly affected by these phenomena that can ultimately cause temporary link interruption. These deleterious effects exhibit a rather different behavior in a maritime environment as compared to the terrestrial one [176, 177]. Therefore, exhaustive research in such an environment is required in order to characterize the local weather effects on the FSO performance. Atmospheric turbulence can be a major degradation factor for an FSO link and, therefore, extensive theoretical and experimental research work has been devoted to quantify its effects on atmospheric laser propagation.

This chapter is divided into two parts, both of which model the performance of an FSO link in a maritime environment. In the first part of this chapter, the performance of a commercial FSO link in a maritime environment is explored. Since measuring directly an optical link over sea is rather difficult, it is very helpful to construct simple models for optical link performance quantification based upon routinely single point-measured environmental parameters. To this end, a second-order polynomial model is proposed to predict the RSSI of the system based upon local macroscopic parameter measurements. The collected data spanned over a period of approximately 40 days, within which the fluctuations of these parameters were quite intense. The model was

validated twice against observed data in later periods and proved to be very accurate, i.e., a correlation > 0.8. The model includes basic meteorological parameters, including wind speed, air temperature, humidity, air pressure, solar radiation, dew point, and rainfall rate. Finally, the probability density function of the RSSI data has been compared against standard channel models, i.e., Gamma, Lognormal, Weibull, and the best fit is estimated using the Kullback-Leibler (KL) divergence.

The main goal of the second part of this chapter, is to use a large data set obtained from the same FSO link, located at the entrance of Piraeus port, during the winter of 2020, in order to gain a better understanding of the effects of atmospheric conditions to the performance of the link. Apparently, experimenting with a laser link in the open sea for extended period of time is not trivial. Therefore, we utilized an established link between two fixed points on the land that crosses a maritime environment and allows for adequate experimental data to be obtained. A twenty two-day data collection period has been used to construct our proposed model for RSSI as a function of local environmental conditions (air temperature, wind speed, dew point, humidity, pressure, solar radiation and air-sea temperature difference) with a very decent approximation (R^2 = 71%). The model has been validated and exhibited better accuracy as compared to our, elsewhere [J1], proposed RSSI model. Finally, a C_n^2 parameter estimation model in maritime environment (NAVSLaM), proposed by the meteorology department of the Naval Postgraduate School, has been employed to correlate the measured RSSI parameter with the estimated atmospheric turbulence strength.

5.2 Experimental Setup

The experimental instrumentation was located on the roof of the Hellenic Naval Academy (HNA), i.e., primary terminal, and the lighthouse of Psitalia island, i.e.,

remote terminal. The horizontal optical link is located 35 meters above the sea and crosses the entrance of the Piraeus port; nearly the entire path is over the water and, thus, clearly a maritime environment. Figure 5.1 shows the exact spots of both terminals in the map, as well as the ambient environment that the 2958-m-long optical link operates.

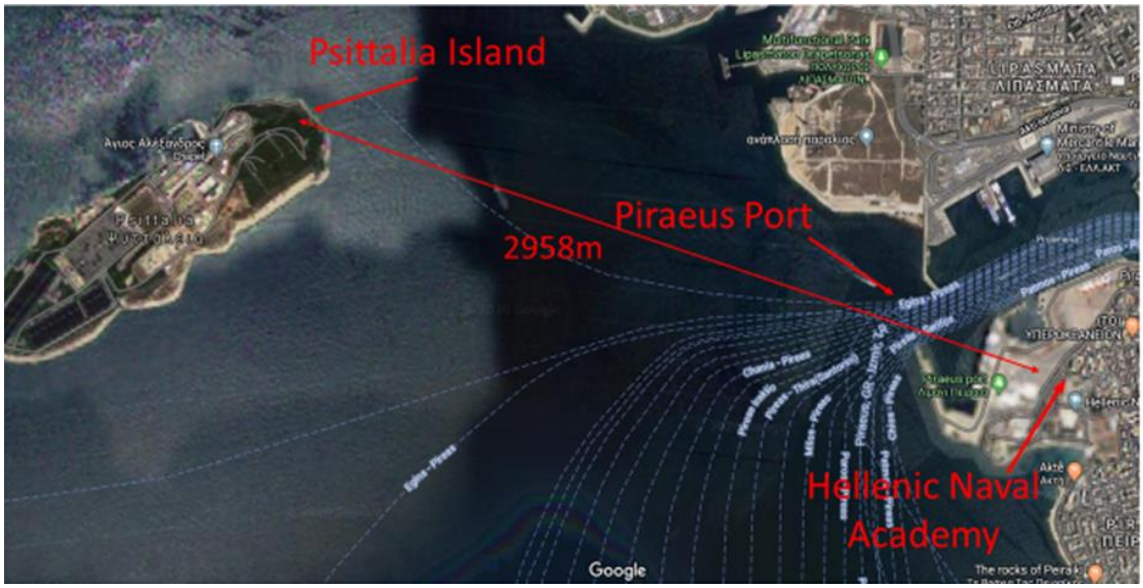


Figure 5.1. The laser communications link located across Piraeus harbor entrance [J1].

This link will be disrupted whenever a vessel taller than 35 m crosses the path; therefore, to minimize these disruptions, the experiment was carried out during the winter when fewer cruise ships visit the port. The FSO system used in the experiment was an MRV TS5000/155 model. The setup consisted of two terminals with operational characteristics available in Table 5.1. The system's scheme used is intensity modulation/direct detection (IM/DD) and it operates in a data rate of 155 Mbps. It uses an open protocol to automatically identify and lock on the current data rate and clock.

Table 5.1. FSO System Parameters.

Parameter	Value
Operating Wavelength	850 nm
Light Source	3 Lasers
Total Output Power	150 mW
Beam Divergence	2 mrad
Detector Type	APD
FOV	2 mrad
Sensitivity	−46 dBm

Both terminals utilized stand-alone PCs in order to send and receive/store data. The interface between them is achieved through an SFP multimode fiber cable, operating at 1310 nm, which drives the optical signal from the detector through an O-E converter directly to the PC. The RSSI data is then stored and is available to export for further analysis. The terminal over Psitalia island (Fig. 5.2) can be remotely operated from the HNA through the optical link.

Figure 5.2. The FSO link from the Psitalia Island (left) and the HNA' building (right) point of view [J1].

Additionally, an Ambient Weather (WS-2000) weather station is co-located with the HNA FSO terminal (Fig. 5.3) to provide real time measurements of macroscopic meteorological parameters that include wind speed, wind direction, air temperature, relative humidity, air pressure, dew point, solar radiation and rainfall rate. These measurements are then stored and readily available to export, analyze and study.

Figure 5.3. The MRV TS5000/155 FSO system on the Hellenic Naval Academy and the co-located ambient weather WS-2000, weather station [J1].

5.3 Results and Analysis (Base Model)

The experimental measurements for the first part of this chapter, spanned over a period of two months, from 30 November 2019 until 31 January 2020. During the first part (30 November 2019–10 January 2020), the observed data were utilized to build the model. Due to technical reasons, the data collection during that period was not perpetual. During the second part, which consisted of two sub-periods (10–15 January 2020 and 24–31 January 2020), the model was validated against real data. The obtained data were stored and exported every few days for further analysis using spreadsheets and MATLAB. The location where the experiment took place, along with the diverse meteorological conditions, provided a very challenging environment for a laser communications link. During that period, the FSO link operated successfully in warm and sunny, rainy, cold and windy days. Therefore, the model was trained on a wide variety of conditions. Figures 5.4 and 5.5 show the fluctuations of the observed meteorological parameters over the data collection period, including air temperature, dew point, relative humidity, air pressure, wind speed, solar radiation and rainfall rate.

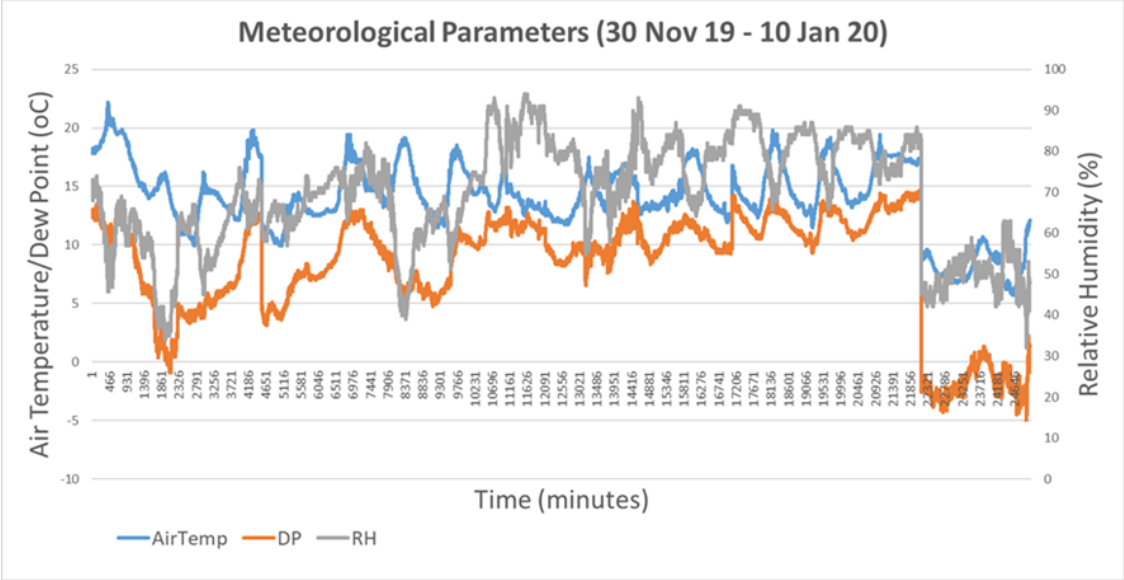


Figure 5.4. Air Temperature, Dew Point and Relative Humidity fluctuations over the data collection period.

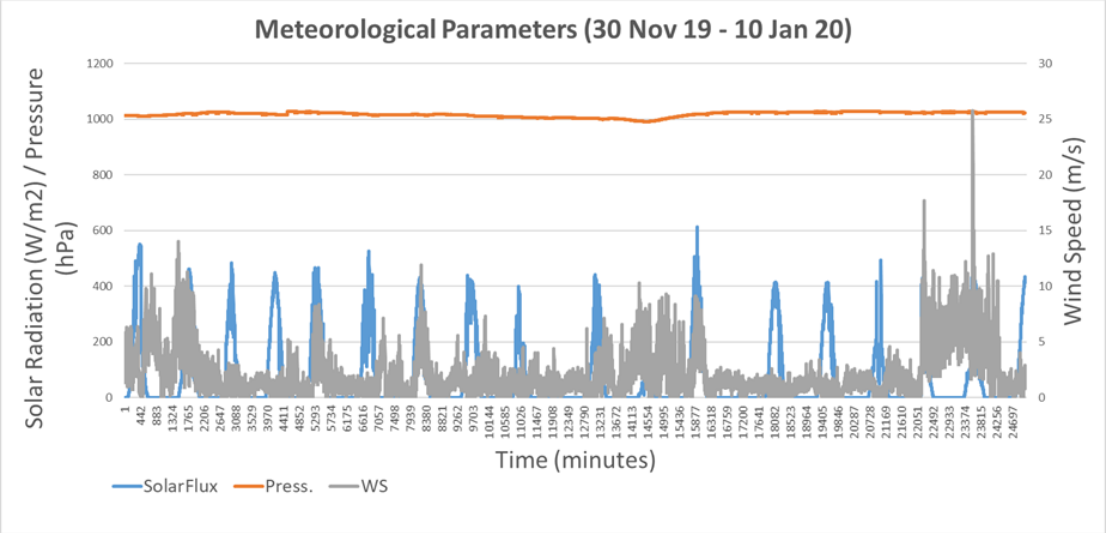


Figure 5.5. Air Pressure, Wind Speed and Solar Radiation fluctuations over the data collection period.

Table 5.2 summarizes the mean, minimum and maximum values of the meteorological parameters observed during the data collection period.

Table 5.2. Collection of Mean, Minimum and Maximum observed values of meteorological parameters during the data collection period.

	Air Temperature (°C)	**Dew Point (°C)**	**Relative Humidity (%)**	**Air Pressure (hPa)**	**Wind Speed (m/s)**	**Solar Radiation (W/m^2)**
Mean Value	14.07	8.3	69.34	1017.77	2.09	70.5
Min Value	5.70	−4.9	32.00	990.70	0.00	0.0
Max Value	22.20	14.7	94.00	1028.70	25.80	613.3

5.3.1 Regression Model

A set of 25,056 data points (one measurement/minute) within a period of 42 days was utilized to deduce the empirical model for RSSI prediction. The following second-order polynomial has been selected to provide a good fit among seven independent parameters (wind speed, relative humidity, air temperature, air pressure, solar radiation, dew point and hourly rainfall rate) and the dependent RSSI,

$$RSSI = -61236.2 - 4.7678 * P + 0.002386 * P^2 + 461.42 * T - 0.8294 * T^2 - 0.6145 * RH - 0.0236 * RH^2 + 8.2251 * DP + 0.2627 * DP^2 - 0.1626 * WS - 0.011 * WS^2 + 0.04889 * SF - 3.8313 * 10^{-5} * SF^2 - 3.75634 * HR \quad (5.1)$$

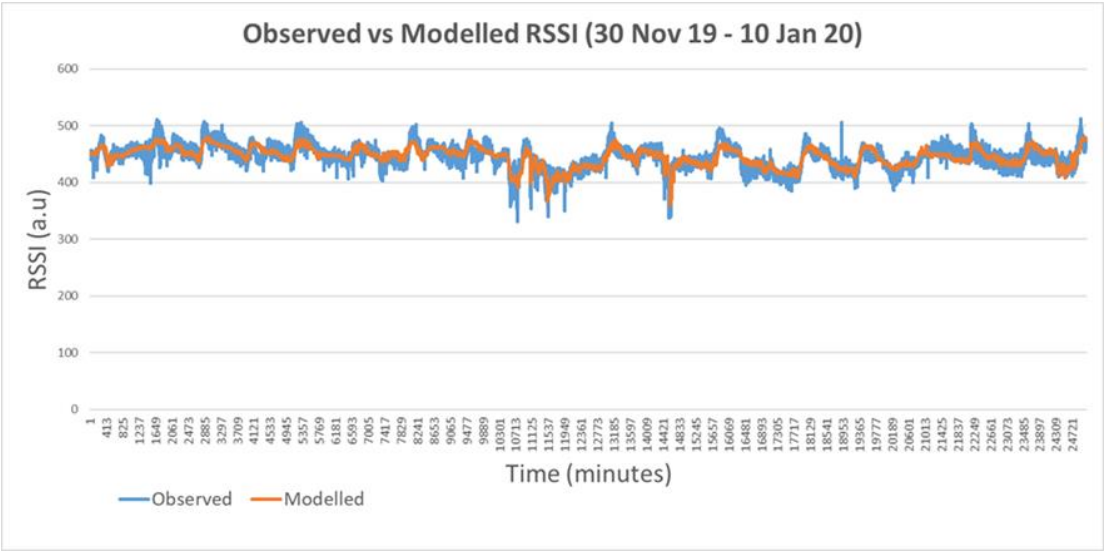


Figure 5.6. Comparison between observed and modeled RSSI for the data collection period.

where T stands for the air temperature in Kelvin, P is the air pressure in hPa, RH is the percentage of relative humidity, DP is the dew point in Celsius, WS represents the wind speed in meters per second, SF being the solar flux in Watts per square meter and HR is the rain rate in mm/hour. The derived model demonstrated a very decent accuracy with an R^2 of 68.2%. Figure 5.6 shows the predicted values of RSSI based on the results of the regression analysis versus the observed one.

By using linear correlation coefficients, we further investigated the correlation coefficients of the considered parameters with RSSI. In the case under consideration, relative humidity appeared to have the most significant correlation with RSSI, with a negative value of −0.56557, indicative of its adverse effect to the link's performance. Table 5.3 summarizes the correlation coefficients of all seven parameters with RSSI.

Table 5.3. Matrix of linear correlation coefficients for measured meteorological parameters and received signal strength from 30 November 2019 to 10 January 2020.

	Air Pressure	**Air Temperature**	**Relative Humidity**	**Dew Point**	**Wind Speed**	**Solar Flux**	**Hourly Rain Rate**	**RSSI**
Air Pressure	1							
Air Temperature	−0.26383	1						
Relative Humidity	−0.24084	0.234145	1					
Dew Point	−0.31862	0.77857	0.788697	1				
Wind Speed	−0.02946	−0.17976	−0.42763	−0.39007	1			
Solar Flux	0.058201	0.279619	−0.33786	−0.04767	0.140114	1		
Hourly Rain Rate	−0.26519	−0.02916	0.236718	0.117156	−0.01767	−0.04762	1	
RSSI	0.13399	0.219768	−0.56557	−0.20549	0.131537	0.44691	−0.35769	1

Emphasis should be given to the impact of rain on the performance of the optical link. During the data collection period, i.e., totally 22.5 hours, exhibited a non-zero precipitation rate, thus allowing further investigation on these effects. As shown in Table 4, the rain has a moderate anti-correlated relation (−0.35769) with RSSI.

Additional analysis on MATLAB showed that a 32% of the total RSSI variance is explained by the variance of the hourly rain rate, i.e., R^2 in Figure 5.7.

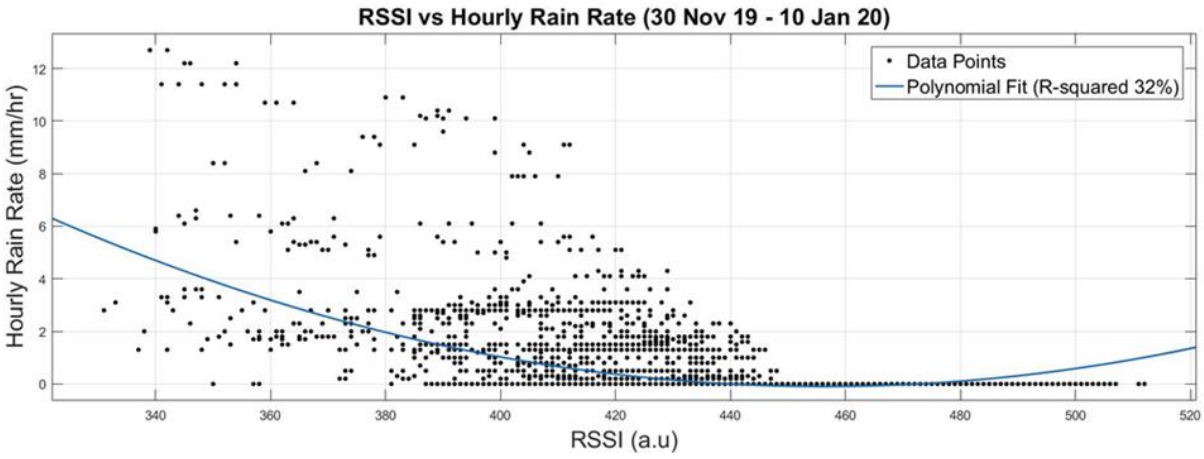


Figure 5.7. versus hourly rainfall rate measured data.

5.3.2 Model Validation

The validation period of the first model was from 10 – 15 January 2020. During this period, in general, the range of the observed parameters is smaller than the data collection period, the mean air temperature is 3 degrees lower, and the mean air pressure was more than 7 hPa higher. Qualitatively, the modeled RSSI estimations, as seen in Figure 5.8, demonstrated a very good fit with the observed values with a few exceptions where the observed values exhibited strong fluctuations. In this part, the model predicted a smoother form with less fluctuations. Quantitatively, the observed RSSI values as compared to the predicted values for this period had an R^2 of 76%. The percentage of the predicted values variance is explained very well by the variance of the observed values: that is the R^2 reached a value of 69%. The model also correlated very well with a linear correlation coefficient of 0.8327. Apart from the modeled and observed values graph, Figure 5.8 shows the RSSI values graph predicted by the model proposed in [Eq. 6, 178] by J. Latal et al. A relatively significant constant offset of

approximately 80 – 100 RSSI units is observed during the entire period. It should be pointed out that the model proposed by J. Latal et al. [Eq. 6, 178] referred to a terrestrial terrain in contrast with our model, which focuses on a maritime environment.

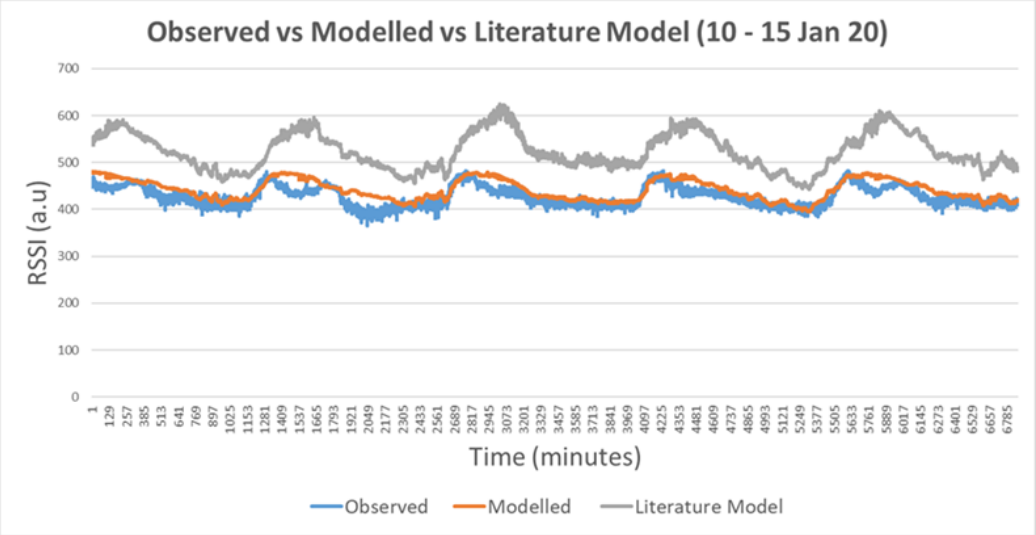


Figure 5.8. Model evaluation for measured RSSI data for the period 10 – 15 January 2020. The grey line shows the literature modeled RSSI proposed by J. Latal et al. [178].

The second validation period was from 24 to 31 January 2020. During this period, the mean values as well as the range of the observed parameters were closer to those of the data collection period. Qualitatively, the modeled RSSI estimations (Eq. 5.1), as seen in Figure 5.9, again yielded a very good match with the observed values. It lacked accuracy on the parts where observed RSSI values exhibited abrupt "peaks". In these parts, the model seemed to follow better the peak values of the observations. Quantitatively, the observed RSSI values, as compared to the meteorological parameters of this period, had an even better R^2 (84%). The percentage of the predicted values variance is also explained better by the variance of the observed values, that is the R^2, reached a value of 74%. The model also correlated very well with the observed values with a linear correlation coefficient of 0.8645. Finally, the model proposed in

[Eq. 6, 178] had an even larger offset which at some points reached values of around 100–120 RSSI units.

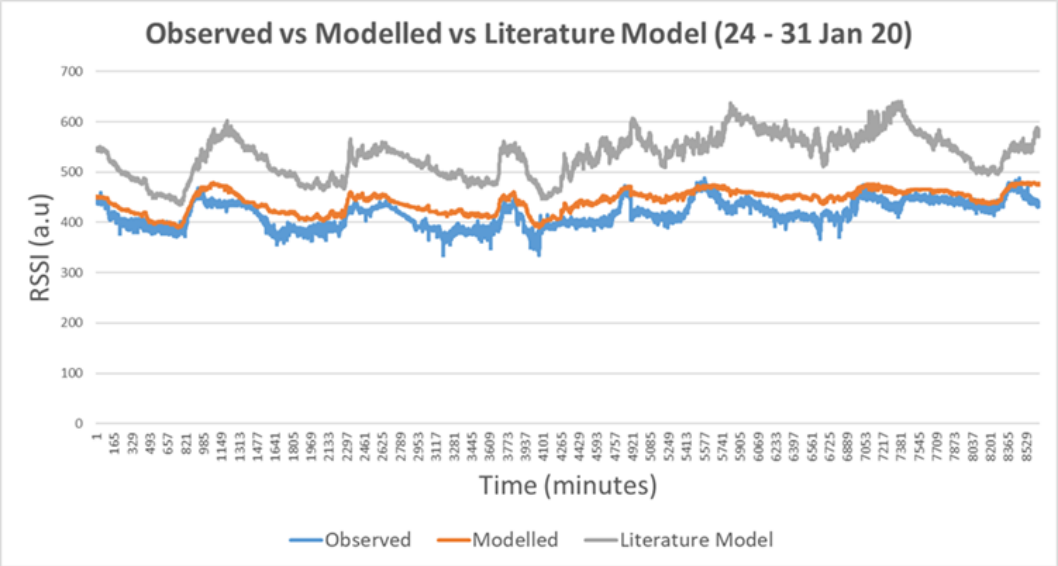


Figure 5.9. Model evaluation for measured RSSI data for the period 24–31 January 2020. The grey line shows the literature modeled RSSI proposed by J. Latal et al. [Eq. 6, 178].

5.3.3 RSSI Distribution Fitting Analysis

The Kullback-Leibler (KL) divergence provides a very useful mathematical tool to measure the difference of two probability distributions [179]. KL divergence is a non-symmetric measure between two probability distributions *p(x)* and *q(x)*. The KL divergence, denoted DKL (*p(x)* || *q(x)*), represents the information lost: the smaller the KL divergence, the more the two distributions are similar. In case of a discrete random variable *x*, assume two probability distributions *p(x)* and *q(x)*, both non-negative and both summing up to unity for any *x* in X. The definition of KL divergence is then [179],

$$D_{KL}(p(x)||q(x)) = \sum_{x \in X} p(x) ln \frac{p(x)}{q(x)} \quad (5.2)$$

Typically, *p(x)* represents the "true" or a theoretical calculated distribution of the observed data, whereas *q(x)* represents a model or an approximation of *p(x)*. In case of a continuous variable *x*, then

$$D_{KL}(p(x)||q(x)) = \int_{-\infty}^{+\infty}(x) ln\frac{p(x)}{q(x)}dx \quad (5.3)$$

Based on the KL divergence theory, the theoretical distribution that best fits the observed RSSI data during the collection period (30 November 2019–10 January 2020) has been deduced. Utilizing the distribution fitting application of MATLAB, initially we estimated the parameters of the probability density function of three theoretical distributions, namely lognormal, Weibull and gamma, for the considered RSSI values range, i.e., min. 331–max. 512. The empirical PDF of the RSSI data has also been evaluated and the corresponding results are available both graphically in Figure 5.10 and numerically in Table 5.4, from where it becomes evident that among the three considered probability distributions, the gamma distribution yields the best fit, a fact that is difficult to be ascertained from an inspection of Figure 5.10.

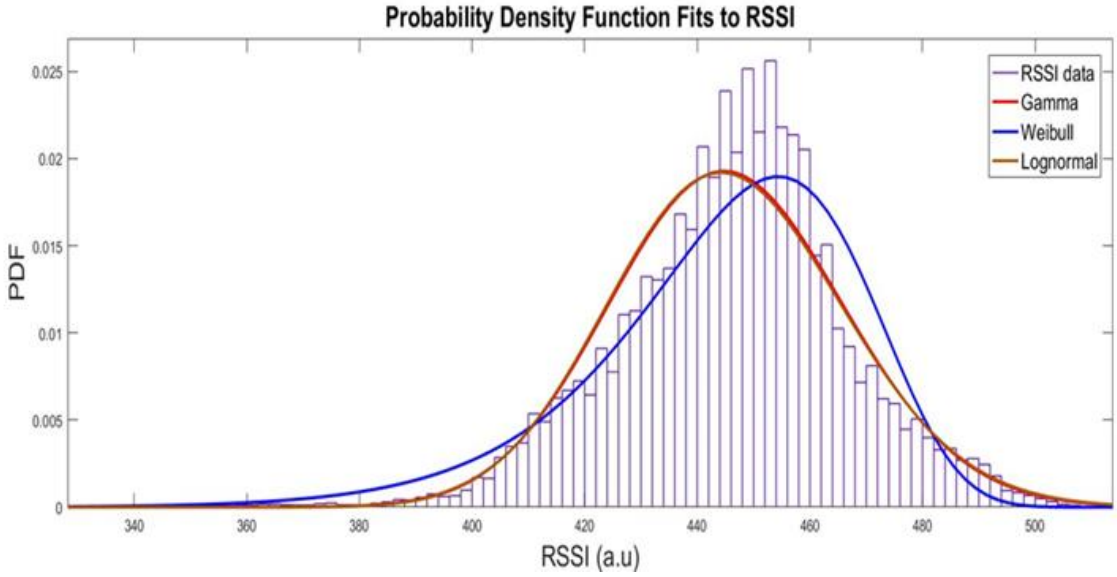


Figure 5.10. PDF fits of gamma, lognormal and Weibull distributions for observed RSSI data for the time period from 30 November 2019 to 10 January 2020.

Table 5.4. divergence for Gamma, Lognormal and Weibull probability distributions.

	$DKL\ (p(x)\ ||\ q(x))$
Gamma	4.17×10^{-2}
Lognormal	4.38×10^{-2}
Weibull	7.95×10^{-2}

5.4 Results and Analysis (Improved Model)

The experimental measurements for the improved model, took place during the winter of 2020. An initial period of twenty-two days (24 Jan to 14 Feb) was devoted for data collection, analysis and model construction. Once per minute, the RSSI parameter was collected and logged from the MRV FSO system. That frequency was selected in order to easily correlate those measurements with the weather station's data collection, which include wind speed (WS), pressure (P), air temperature (T), dew point (DP), solar radiation (SF) and relative humidity (RH). Additionally, the improved model included as an independent variable the air-sea temperature difference (ASTD). An online weather statistics database [180] was exploited for the sea temperature measurements. The range of the environmental parameters values over this period is presented in Table 5.5.

Table 5.5. The value range of the environmental parameters for the period 21 Jan to 14 Feb 20.

Parameter	Min Value	Mean Value	Max Value
RSSI	335	422.39	517
P (hPa)	987.7	1017.6	1035.3
T (°C)	273.8	282.8	290.7
RH(%)	41	69,61	93
DP (°C)	-5.5	7.29	13.4
WS (m/s)	0	2.89	20
SF (W/m^2)	0	103.53	735.7
ASTD (°C)	-11.1	-2.23	5.7

As known, the relative humidity exhibits a significant anti-correlated behavior with temperature. This is the case between ASTD and RH too, as shown in Figure 5.11.

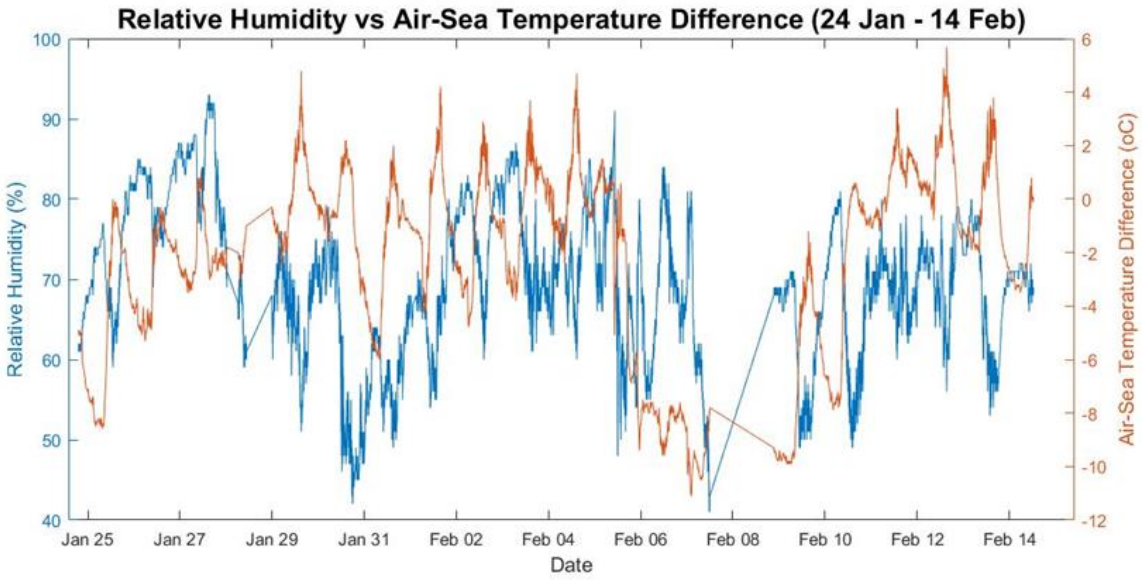


Figure 5.11. The highly anti-correlated relation between relative humidity and air-sea temperature difference during the period 24 Jan to 14 Feb 20.

On the other hand, the observed RSSI shown a good agreement with the ASTD in terms of their fluctuation trends. The measurements "gap" during the 8th of February is due to a temporary technical issue on the FSO system.

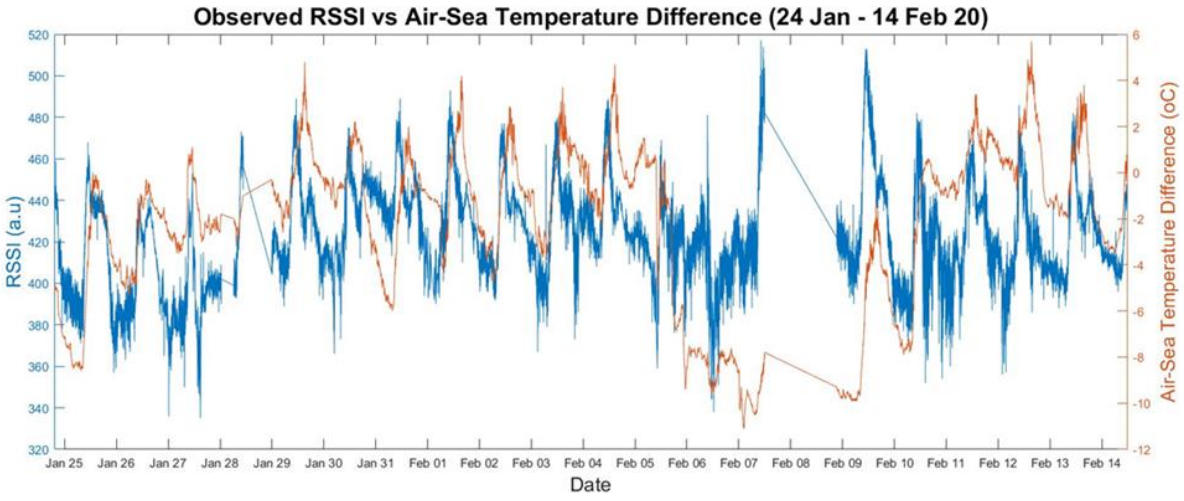


Figure 5.12. Observed RSSI and ASTD measurements comparison during the period 24 Jan to 14 Feb 20.

5.4.1 RSSI Statistical Modeling

The data collected during that period (24 Jan to 14 Feb 20) gave 26.238 data points that allowed an accurate model construction to relate the RSSI with seven macroscopic environmental parameters. We used a linear regression analysis to construct a second-order model that would allow RSSI predictions based on environmental parameters point measurements. The regression analysis summary output showed a significant similarity between measured and predicted RSSI values, that is certified by an R^2 parameter value of 71,1%. A linear correlation coefficient was also used to numerically evaluate their analogy (0,843). The resulting regression model is given as,

$$RSSI = 112100{,}8 + 21{,}4 * P - 0{,}01 * P^2 - 878{,}1 * T + 1{,}57 * T^2 + 10{,}47 * RH - \ 0{,}066 * RH^2 - 19{,}83 * DP + 0{,}52 * DP^2 - 0{,}62 * WS - 0{,}09 * WS^2 + 0{,}12 * SF - 0{,}0001 * SF^2 - 7{,}8 * ASTD - 2{,}3 * ASTD^2 \quad (5.4)$$

Fig. 5.13 presents a comparison between the RSSI values directly measured from the MRV FSO system (blue line) and those predicted by our model (red line).

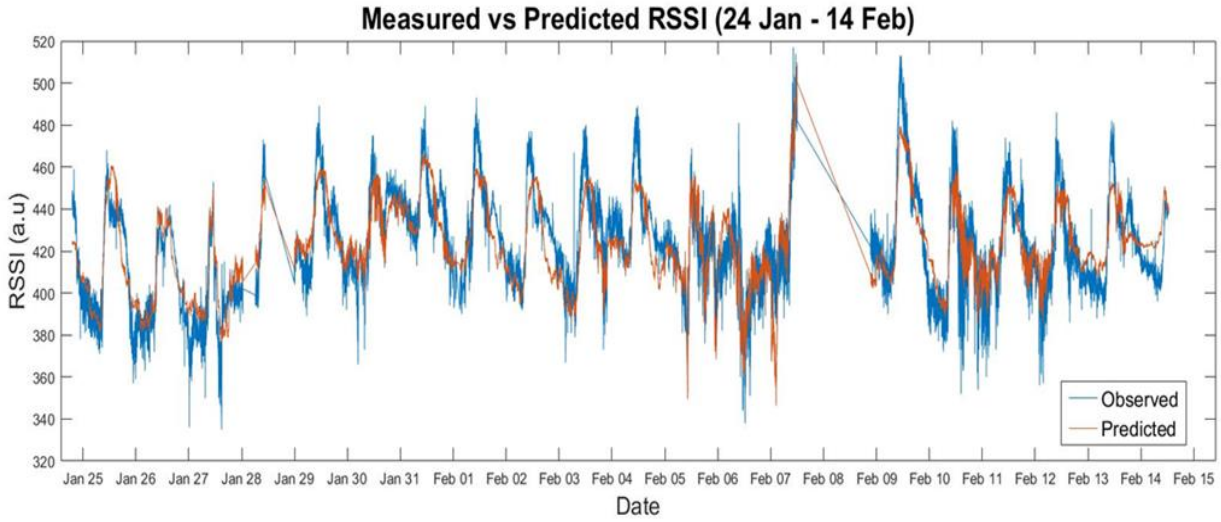


Figure 5.13. Comparison between measured from FSO system and predicted from the model RSSI values for the period 24 Jan to 14 Feb 20.

5.4.2 Improved Model Validation

In order to validate our model and its ability to predict the RSSI parameter from macroscopic environmental point-measurements, we selected two distinct periods, the first from 20 to 26 Feb 20 and the second from 7 to 11 Mar 20. The same environmental measurements were taken, and the observed RSSI value was compared with the predicted one. The environmental conditions during this period was pretty much the same with the collection data period. This fact is assumed to favor our model in terms of its predictability. As a comparison to the improved model, the RSSI parameter is also computed by our base model presented elsewhere [J1]. It is observed that the inclusion of the ASTD as an independent parameter plays a key role to its improvement comparing to the base model. Fig. 5.14 shows the comparison between the observed (blue line), base (red line) and improved (yellow) model. Our improved model achieves an R^2 parameter of 70,5%, which quantitatively supports its statistical significance. Qualitatively, it can be shown from Fig. 5.14 that the predicted parameters have a very good fit with the observed ones, even in harsh RSSI value differences as during the morning of the 23 February. Two minor failures are observed during the midday of the 24^{th} and 25^{th} of February where the model seems to underestimate the observed RSSI values. Overall, the improved model achieves a significant linear correlation coefficient of 0,78 with the observed values as compared to 0,73 of the base model.

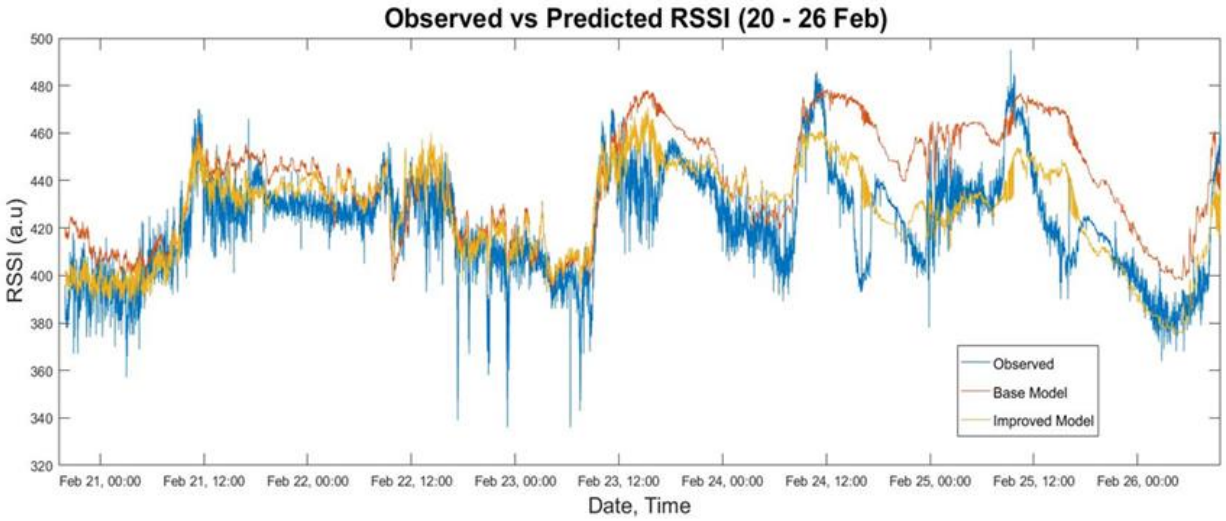


Figure 5.14. Observed RSSI parameter as compared to both the base and the improved model for the period from 20 to 26 February 2020.

The exact same validation procedure was followed for the period from 7 to 11 of March 20, where again our model proved its ability to make legitimate predictions for the RSSI parameter. Figure 5.15 shows the comparison between the observed, base and improved models. Qualitatively, the improvement of the predictability of our model is apparent as compared to the base model which overestimated the RSSI parameter throughout the whole period, except two peak values observed during midday of 10^{th} and 11^{th} of March. The linear correlation coefficient comparison, is slightly better for the improved model (0,81 compared to 0,79) however this validity check has to do mainly with how good the model follows the general trend of the real measurements. As already stated, the base model has indeed an adequate correlation with the observed values, however it fails to accurately predict the real values. Finally, the improved model during this period exhibits less accuracy in terms of its R^2 parameter (66%), however, it is still reliable for bulk estimations.

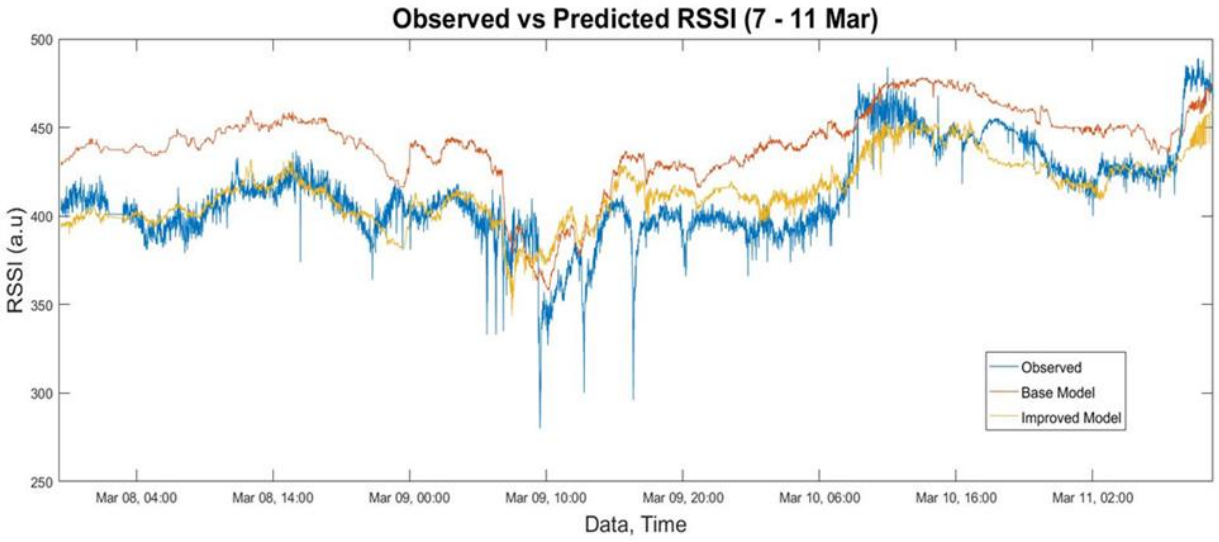


Figure 5.15. Observed RSSI parameter as compared to both the base and the improved model for the period from 7 to 11 March 2020.

5.4.3 NAVSLaM C_n^2 Predictions

A maritime environment exhibits different atmospheric phenomena and specialized models are required for atmospheric turbulence predictions in such an environment, as compared to predictions for a terrestrial one. This paper [J2] utilized such a model (NAVSLaM) to predict the C_n^2 strength along the propagation path of the link based on point measurements of macroscopic environmental parameters. Special attention is given to the effect of the ASTD, since this parameter is found to have a significant impact to the turbulence behavior [181]. By solving Eqs. (2.59) and (2.60) in an iterative process, we estimated the C_n^2 during both validation periods and as a comparison we also estimated the C_n^2 utilizing other empirical models that are not focused on a maritime environment [Eqs. 12,13, 182]. A significant disagreement between model predictions is observed, a fact that justifies the different mechanisms of atmospheric turbulence over a maritime environment, so that maritime atmospheric turbulence cannot be predicted by a model focused on laser propagation over land.

Figures 5.16 and 5.17 show the turbulence fluctuation over two period of times, 20 to 26 Feb and 7 to 11 Mar 20.

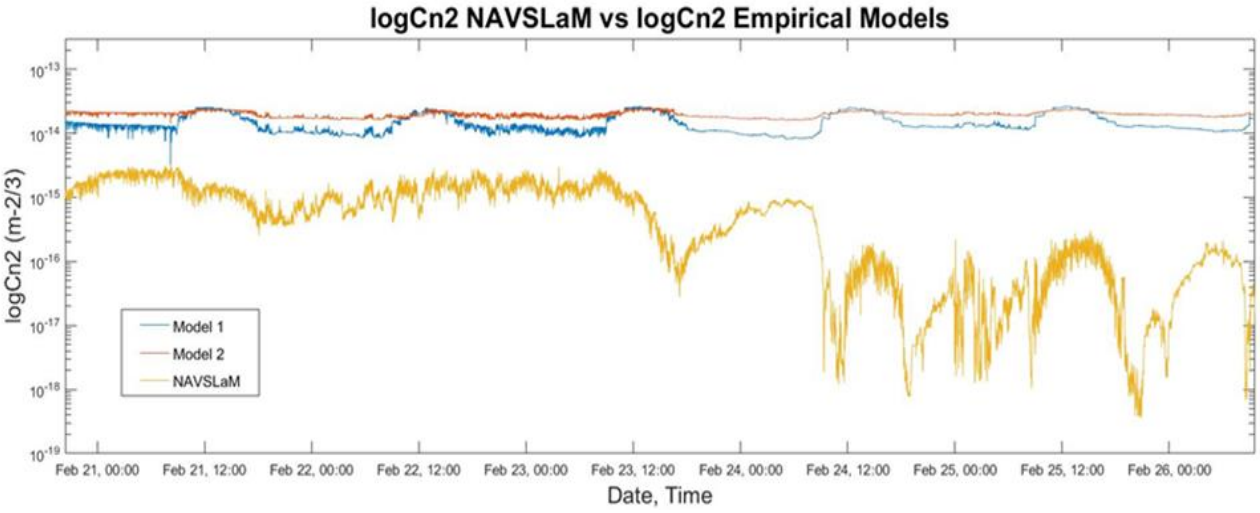


Figure 5.16. Prediction of C_n^2 over the period from 20 to 26 Feb 20. NAVSLaM model (yellow line) as compared to empirical models.

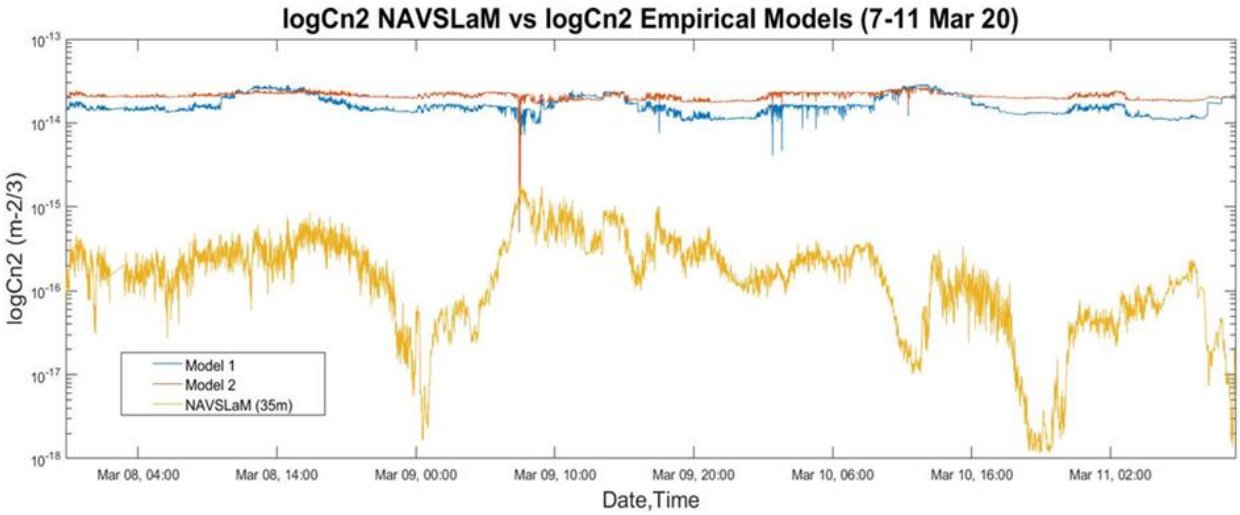


Figure 5.17. Prediction of C_n^2 over the period from 7 to 11 Mar 20. NAVSLaM model (yellow line) as compared to empirical models.

The C_n^2 mean value for the first period is approximately 7.5×10^{-16}, with a minimum of 3.5×0^{-19} and a maximum of 3.2×10^{-15}. The equivalent values for the second period are, 2.2×10^{-16}, 1.2×10^{-18} and 1.2×10^{-15}. The values plotted in both figures are in logarithmic scale and a significant difference between the empirical models and

NAVSLaM is apparent. Specifically, the prediction based on the empirical models has a mean value of approximately two order of magnitudes higher and as known they exhibit a diurnal bell-shaped profile, with a maximum value around midday and minimum around the sunrise and sunset. This diurnal profile is not the case for turbulence over a maritime environment as shown in Figures 5.16 and 5.17.

The C_n^2 predictions over both periods allowed an analysis of the effect of the turbulence strength to the received signal. Therefore, in Figures 5.18 and 5.19 we plotted the RSSI parameter and the C_n^2 in logarithmic scale. In both periods, a strong anti-correlated relation was observed, demonstrating the deleterious effect of turbulence on the laser propagation.

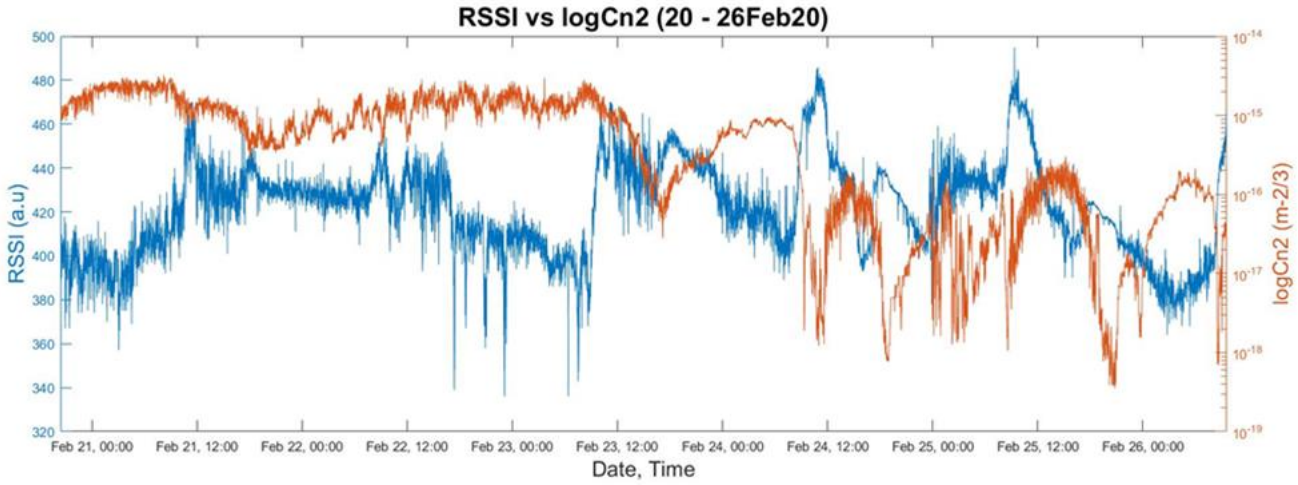


Figure 5.18. RSSI vs C_n^2 for the period from 20 to 26 Feb 20. A strong anti-correlated relation is observed. The C_n^2 is plotted in a logarithmic scale.

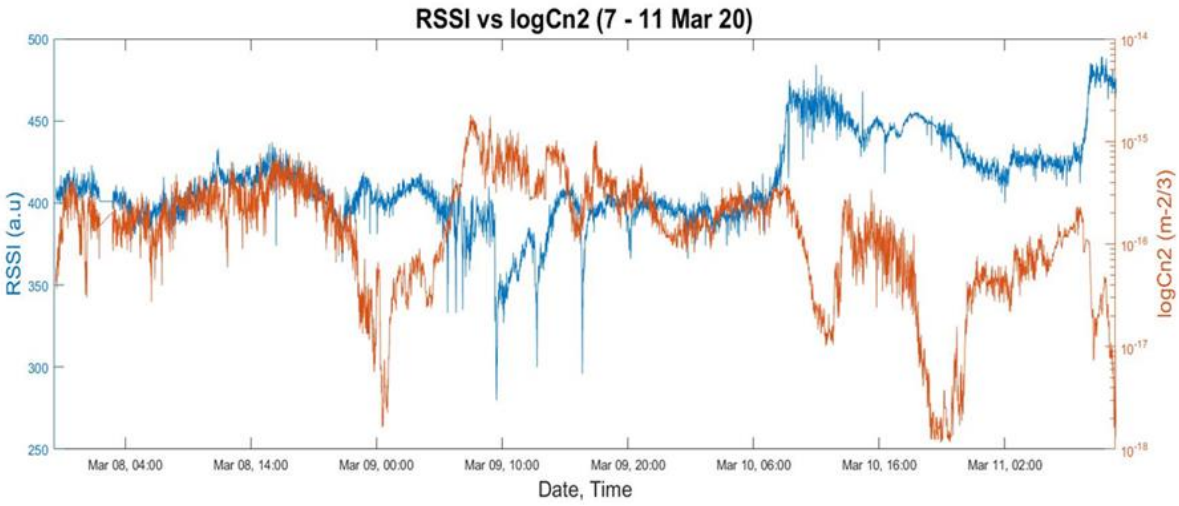


Figure 5.19. RSSI versus C_n^2 for the period from 7 to 11 Mar 20. Again a highly anti-correlated relation is observed. The C_n^2 value is plotted in a logarithmic scale.

As previously stated the ASTD parameter value has a significant impact on the turbulence strength. During the two validation periods, both stable (ASTD > 0) and unstable (ASTD < 0) conditions occurred. During the first period (20 - 26 Feb) the ASTD had a mean value of -3.33 ºC whereas during the second period the mean value was -0.57 ºC. As shown in Figures 5.20 and 5.21, as the ASTD parameter value approached to zero the atmospheric turbulence strength decreases, whereas for greater absolute values the turbulence increases.

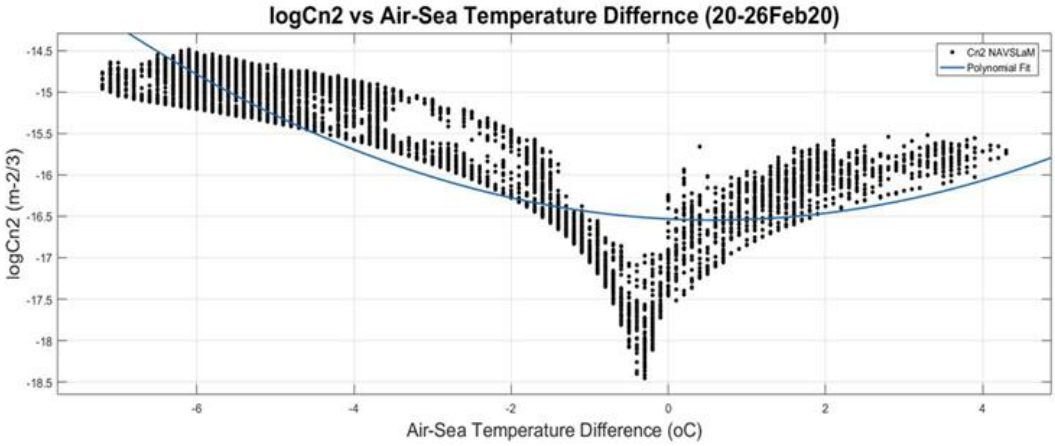


Figure 5.20. C_n^2 measurements versus air-sea temperature difference for the period from 20 to 26 Feb 20.

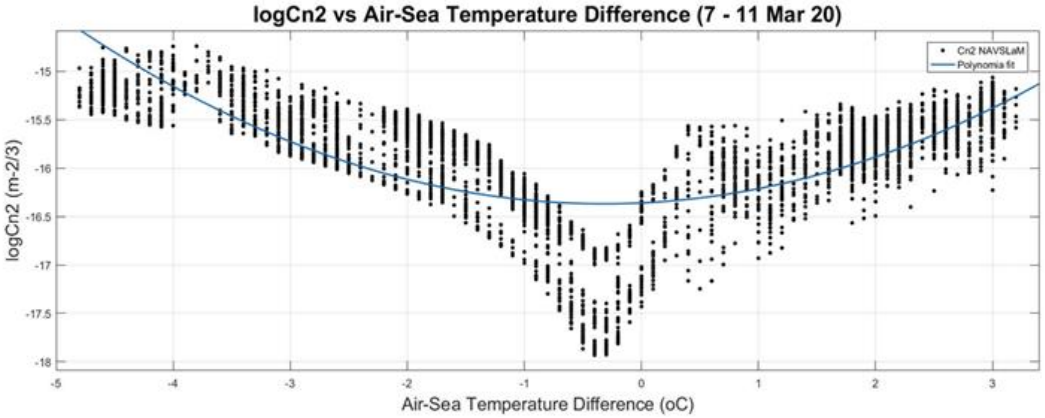


Figure 5.21. C_n^2 measurements versus air-sea temperature difference for the period from 7 to 11 Mar 20.

The atmospheric turbulence is highly height-dependent. The highest values are observed at zero altitude, whereas significantly lower at higher altitudes. For positive ASTD values the C_n^2 profile initially has larger gradient with height near zero height and then becomes less dependent on height than for negative ASTD values. Figure 5.22 shows the height dependence of C_n^2 for a typical midday and sunset time. It is obvious that the turbulence strength differs by an order of magnitude between the value at the surface and a height of 35m, where our experimental setup is located. This fact would favor FSO systems in maritime platforms that are set in the highest possible location.

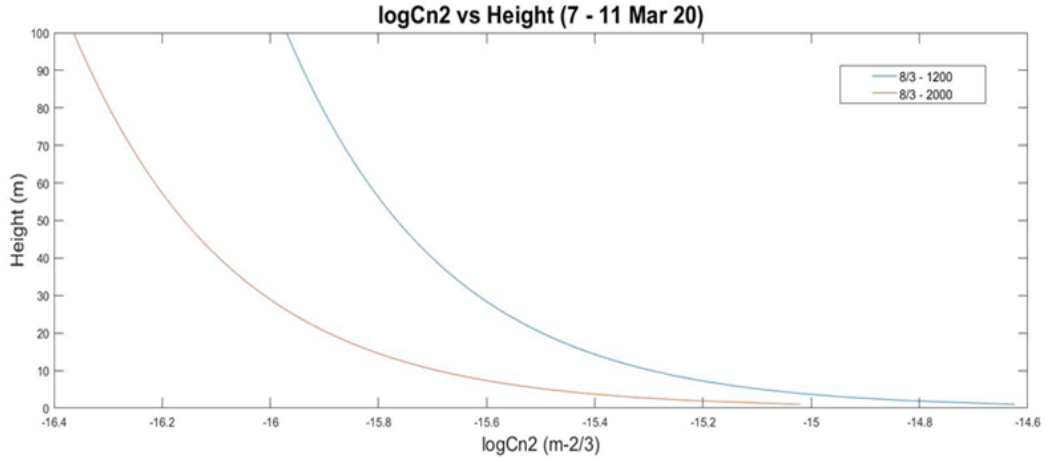


Figure 5.22. Height dependence of C_n^2 for a typical midday (blue line) and sunset (red line) time.

5.5 Summary

In this chapter, two new mathematical models are proposed, to predict the received signal strength of an FSO optical link. The models have the form of a second-order polynomial with seven macroscopic meteorological parameters as the independent variables.

The predicted RSSI values of the base model, fitted the observed values quite well, yielding an R^2 value of 68.2%. The correlation of all seven parameters to the RSSI has been calculated to deduce the weight of each one's effect. Emphasis has been given to the rain effect, where 32% of the RSSI variance was explained by the rainfall rate variance (R^2 = 0.32). The proposed model has been validated against real data in two separate periods and the R^2 and correlation coefficient between the observed and modeled RSSI values has been computed to check how good the fit was. Both periods exhibited high R^2 and correlation coefficient, namely 69% and 0.8327, respectively.

The second proposed model is an improved edition of the base model. A closed form expression is constructed that predicts RSSI parameter based upon point measurements of local environmental parameters, including wind speed, air temperature, humidity, pressure, dew point, solar flux and air–sea temperature difference. The latter has been found to have a significant impact on the laser beam propagation over sea. The regression analysis output showed a significant fit between measured and predicted RSSI values, that is certified by an R^2 parameter value of 71.1%. A linear correlation coefficient was also used to numerically evaluate their analogy, i.e. 0.843. We then validated our model and its ability to predict the RSSI parameter from macroscopic environmental point-measurements, during two distinct periods, the first from 20 to 26 February 2020 and the second from 7 to 11 March 2020. Overall, the improved model achieved a significant linear correlation coefficient of 0.78

with the observed values as compared to 0.73 of the base model for the first period and 0.81 as compared to 0.79 for the second period. The goodness of fit parameter R^2 between the improved model prediction and the observed values was 70.5% and 66% for the two periods, respectively.

CHAPTER 6

Machine Learning Algorithms for Received Optical Power Prediction of an FSO Link Over a Maritime Environment

This chapter includes the results that have been published in the peer-reviewed journal [J3].

6.1 Introduction

Across the past 20 years, there has been extensive research on how to estimate how well an optical communications link would work when it is used across water. Numerous atmospheric phenomena and turbulence effects have been well studied, and lengthy experimental data have made it possible to build straightforward models. This chapter's goal is to provide a variety of machine learning (ML) methods that can accurately anticipate how well a free-space optical communication (FSO) link would function given current atmospheric conditions. An extensive data set of the received signal strength indicator (RSSI) parameter for an FSO link was collected and analyzed against seven local atmospheric parameters, i.e. wind speed, pressure, temperature, humidity, dew point, solar flux and air-sea temperature difference. The k-nearest-neighbors (KNN), decision trees (DT), random forest (RF), gradient boosting regressor (GBR) and artificial neural networks (ANN) algorithms have been employed and compared among each other, using the root mean square error (RMSE) and the

coefficient of determination (R^2) of each model as the primary performance indices. The analysis showed an excellent fit for all ML models, proving their ability to offer a significant improvement in FSO performance modeling as compared to empirical regression models. The ANN model performed best and achieved an R^2 equal to 0.94867, whereas RF model achieved the optimum RMSE result (7.37).

Based on the knowledge that the seven measured parameters can have a significant impact on FSO performance, which was learned from the findings reported in the previous chapter, the variance selection of the seven measured parameters was made. With the exception of the ANN, the creation of ML algorithms followed the following process: 80% of the input data set was used to train the model, and the remaining 20% was utilized to verify the model's accuracy using the RMSE and R^2 parameters as performance measures. In case of the ANN model development, the dataset was split in three subsets, i.e. 70%, 15% and 15%, allocated to a training, validation and testing phase, respectively. The role of the validation phase is the fine tuning of the model's hyperparameters. The analysis resulted in an excellent fit for all models (R^2>0.85) and also to the fact that the ambient temperature proved to be statistically the most important meteorological parameter for RSSI prediction.

A recently emerging topic, machine learning for optical communications, will undoubtedly get more attention in the near future. It comes with a wide range of tools that let the trained algorithm evaluate and comprehend the data. The methods of supervised learning for regression models are the main topic of this chapter. In order to estimate the function that best fits a data set, we have to be aware of the special attributes that possesses in terms of accuracy and interpretability. For example, a linear regression algorithm may be very clear to interpret but also very restrictive, especially when a non-linear relationship between input and output actually holds. On the other

hand, algorithms such as support vector machines or boosting methods, could lead to such complicated models that may be very difficult to interpret. However, if prediction accuracy is the desired outcome, models with high flexibility seem to be the best choice. This is not always the case though, since overfitting of the data will, despite decreasing the training error, increase the testing error. Therefore, a careful bias – variance tradeoff has to take place before selecting the most appropriate model. The term variance refers to how much will the function be changed, if the same model is used on a different data set, whereas bias refers to the error introduced by modeling a real-life problem. In general, more flexible models tend to increase the variance and decrease the bias.

6.2 Experimental Setup

In order to measure the received signal strength for a real FSO system, a 2958-meter propagation path was established over Piraeus port entrance at the Hellenic Naval Academy (HNA), Piraeus, Greece. A commercial FSO system was used to take RSSI measurements, to compare with the model predictions. Specifically, the FSO system used in the experiments is a MRV TS500/155 transceiver, operating at 850nm with a maximum output power of 150-mW and data rate of 155-Mbps. It utilizes three laser sources with a beam divergence of 2mrad each and a single receiver with a diameter of 20-cm, sensitivity of -46 dBm, and an avalanche photodetector (APD). It also uses an open protocol to automatically identify and lock on the current data rate and clock. The first terminal was located on the roof of a building at HNA, whereas the second was located on the Psytalia Island lighthouse. The horizontal propagation path was approximately 35 m above the surface of the sea and over 95% of it over water. Tidal changes at that area are minimal, therefore altitude changes of the propagation path were ignored. In addition to collecting measured readings from the FSO system, a WS-

2000 weather station near the MRV was deployed to measure macroscopic meteorological parameters, including ambient temperature, air pressure, relative and absolute humidity, dew point, solar flux, ultra-violet index, rainfall rate, wind speed and direction. An online weather statistics database, [180] was also exploited to take sea temperature measurements from which the air-sea temperature difference was estimated. Measurements from both sensors were taken every minute, twenty-four hours per day.

6.3 Results and Analysis

The results of this paper were extracted based on a large data set of measurements recorded between the 30th November 2019 and the 27th October 2020. Several technical issues, including system resets and line-of-sight blockages due to maritime traffic, did not allow continuous measurements. Additional data cleaning before analysis, reduced the data points to a total amount of 144803. Both sensors internal clocks were synchronized and the measurements were taken at one-minute intervals. The input data obtained from the WS-2000 station, was stored into a .xlsx file format, screened and redundant data was excluded to result in a clean dataset, consisted of eight columns including wind speed in m/s, air temperature in K degrees, dew point in °C, relative humidity (%), air-sea temperature difference in °C, solar flux in $W \cdot m^{-2}$ and relative pressure in mbar. An additional column was added to include the water temperature from the online database. Each data row was then compounded with the respective output value of RSSI for the same date/time. Any measurements with missing or non-physical values were excluded. For example, heavy maritime vessel traffic entering or leaving the Piraeus port caused many of the RSSI values inaccurate because they blocked the LOS.

The first step of the ML development process, was to take the set of all observations and divide it randomly into two subsets. The first subset was used to train the model. The second subset was used to test the model once it had been trained. This second dataset is referred to as the test subset. The objective was to estimate the performance of the machine learning model on new data, i.e., data not used to train the model. There is not an optimum percentage to split the data set for training and testing, rather the need to account for the computational cost to train and test the model as well as an adequate representativeness on both subsets. We chose to follow a split of 80% for training and 20% for the testing subset of our models, except for ANN, where a 70%, 15%, 15% scheme was selected, accounting for the validation subset too. Specific performance measures should be used in order to measure how did the models perform on predicting the RSSI value based on unseen meteorological parameters. Among others, we selected to apply the RMSE which represents the square root of the variance and the coefficient of determination, R^2. The RMSE is defined as,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \tag{6.1}$$

where N is the total number of observations, x_i is the i-th actual RSSI observation, while $\hat{x}_i$ stands for the i-th predicted one. The R^2 metric has the form,

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{6.2}$$

where SS_{res} defined as $\sum_{i=1}^{N}(x_i - \hat{x}_i)^2$, and SS_{tot} as $\sum_{i=1}^{N}(x_i - \overline{x_i})^2$, where $\overline{x_i}$ represents the mean value of the RSSI observations.

The flowchart diagram of the followed ML approach to model the RSSI, is shown in Figure 6.1.

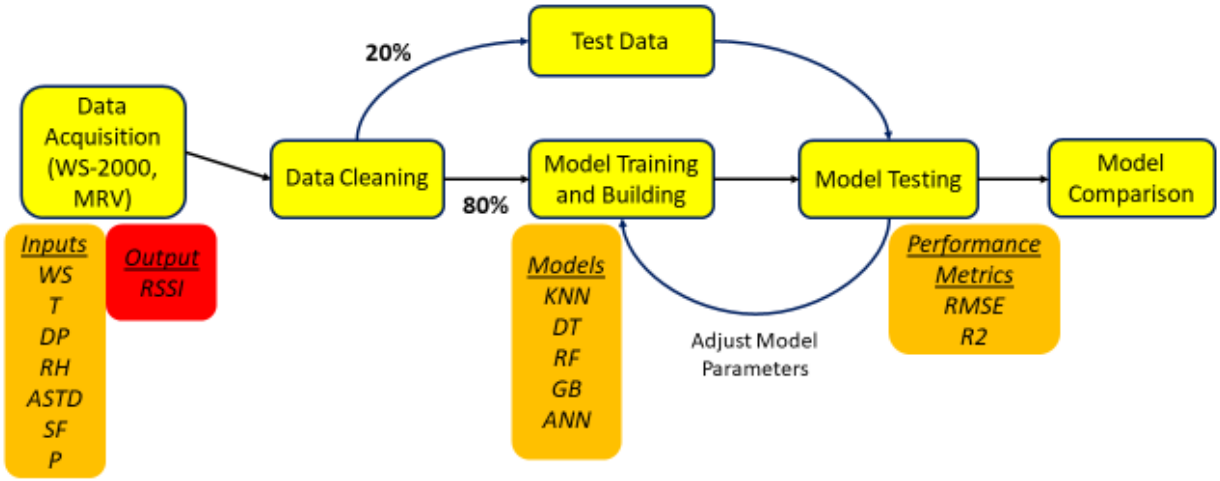


Figure 6.1. Flow chart diagram for the ML-based methodology for FSO system RSSI modeling.

The data analysis was performed using the programming language Python (version 3.8). Specifically, Jupyter Notebook, a web-based interactive computing notebook, which allows the implementation of various libraries, e.g. Pandas, Numpy, Matplotlib, Seaborn, for advanced data analytics and visualization. Matlab (version 2016a) was also used to run the ANN model.

The meteorological data spanned throughout all seasons of a year, therefore the observed values covered a large range, of environmental conditions. Apart from the descriptive statistics of each meteorological parameter that are presented in Table 6.1, it is worth observing the correlation between them and the independent value of RSSI. These correlations are shown in a matrix representation in Figure 6.2, where darker colors indicate higher anti-correlation, whereas lighter colors indicate higher correlation.

Table 6.1. The value range of the meteorological parameters for the period 30 Nov 2019 to 27 Oct 2020.

Parameter	Min Value	Mean Value	Max Value
RSSI	187	420.4	517

P (hPa)	987.7	1040.6	1015
T (ºC)	273.8	306	288.7
RH(%)	22	63.5	94
DP (ºC)	-5.5	11.3	24.6
WS (m/s)	0	2.95	25.8
SF (W/m^2)	0	140.1	1149.5
ASTD (ºC)	-11.1	-1.39	10

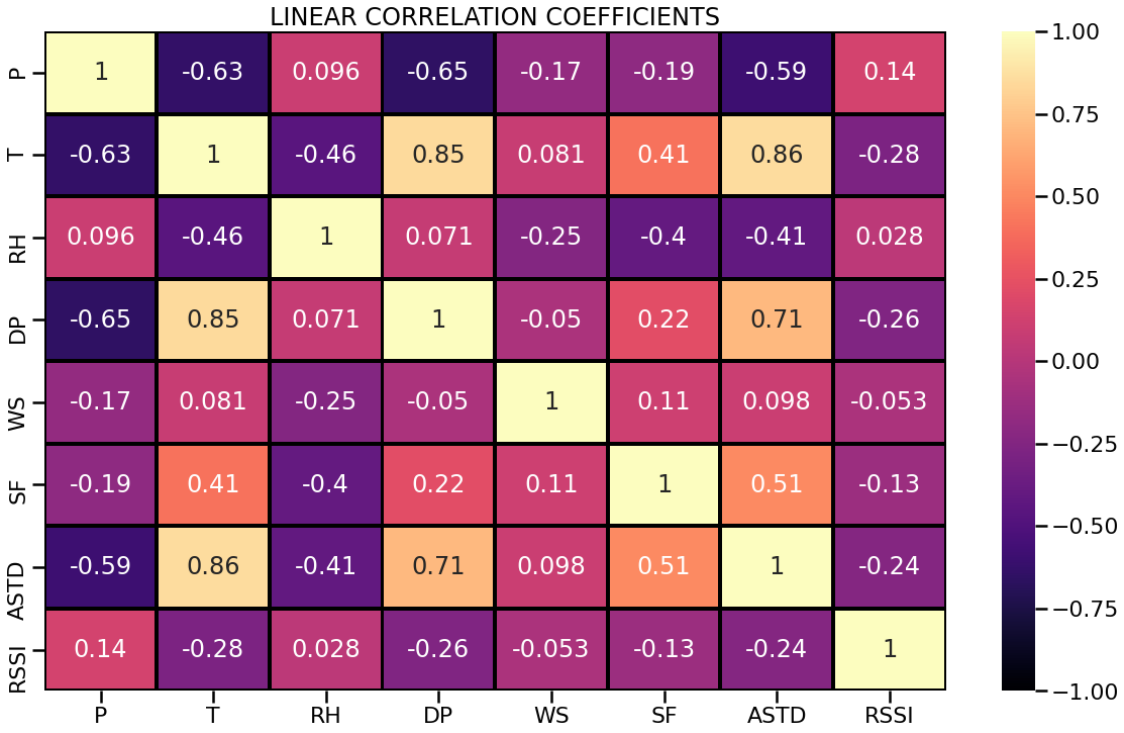


Figure 6.2. Matrix of linear correlation coefficients for measured meteorological parameters and received signal strength.

6.3.1. K-Nearest Neighbors Algorithm

The KNN algorithm is a simple model to implement for regression and classification problems. The main advantages of KNN algorithm are the easiness of its output interpretation and the significantly low time for calculation. KNN is a simple algorithm that stores all available cases and predicts the numerical target based on a similarity measure, e.g., distance functions. A simple implementation of KNN regression is to calculate the average of the numerical target of the K nearest neighbors.

Another approach uses an inverse distance weighted average of the K nearest neighbors. KNN regression uses the same distance functions as KNN classification. Choosing the optimal value for K is best done by first inspecting the data. In general, a large K value is more precise as it reduces the overall noise. The coefficient of determination, R^2, and RMSE for our experimental data set are plotted against various number of neighbors (K) in Figure 6.3, starting from 2 up to 50. As shown in Figure 6.3, the coefficient of determination for the training subset, initiates from a value of 0.95 and then gradually decreases to 0.75 when K = 50. The equivalent values for the testing subset had a peak value of 0.85 for K = 3, meaning that this is the optimal value for the number of neighbors that best fits KNN model to the data set, therefore is the value we selected for our model A structured process of scanning the training subset, i.e. grid search, to determine the optimal value for a parameter is employed to determine the optimal K value (K = 3).

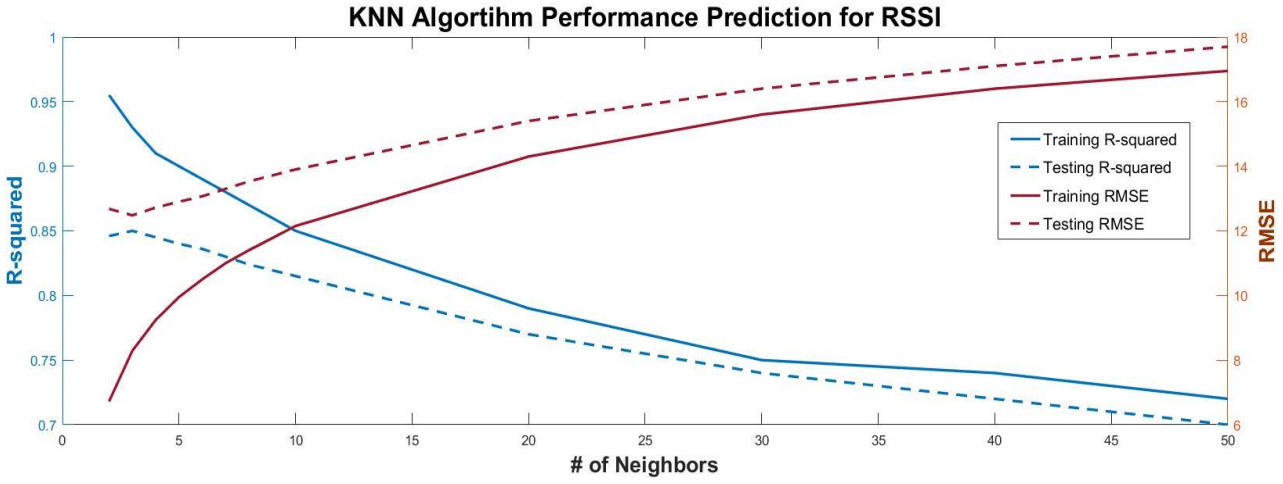


Figure 6.3. Performance metrics plot for KNN algorithm against different number of Neighbors.

6.3.2 Decision Tree Algorithm

A second simple and easy to interpret algorithm is the DT. Despite not so accurate compared to other decision tree methods, e.g. random forests, gradient boosting, is still a good alternative for regression and classification problems. This algorithm splits the predictor space into several regions and make predictions based on the mean of the training observations of the corresponding region. Starting from the root node, a decision tree is constructed by recursive partitioning. Each node is then split into two child nodes, based on the observation parameter that results in the maximum information gain. This process, known as recursive binary splitting, continues until the leaves are pure, i.e. samples at each node belong to the same class. The two basic steps of the DT process include the following:

a. Predictor space division into *J* distinct and non-overlapping regions, $R_1, R_2, \ldots R_J$. The criterion to determine the optimal split point is to minimize the RSS given by,

$$\sum_{j=1}^{J} \sum_{i \in R_j} \left(y_i - y_{R_j} \right)^2 \tag{6.3}$$

where y_{R_j} is the mean response for the training observations in the *j-th* box.

b. For every observation falling into a certain region, the prediction emerges from the response mean value based on the training observations that belong to the same region.

Figure 6.4 shows the first three levels of the decision tree to predict RSSI. The root node assigns observations where T<=299.65 to the left branch with a corresponding mean RSSI value of 423.452. The right branch, i.e. false outcome, leads to a child node that is further split based on the solar flux value and a corresponding mean RSSI value of 375.092, as shown in more details in Figure 6.5. That is, for any

training observation that includes a solar flux parameter value less than 4.9, the left branch will be selected and the corresponding values for the dew point parameter will then be examined.

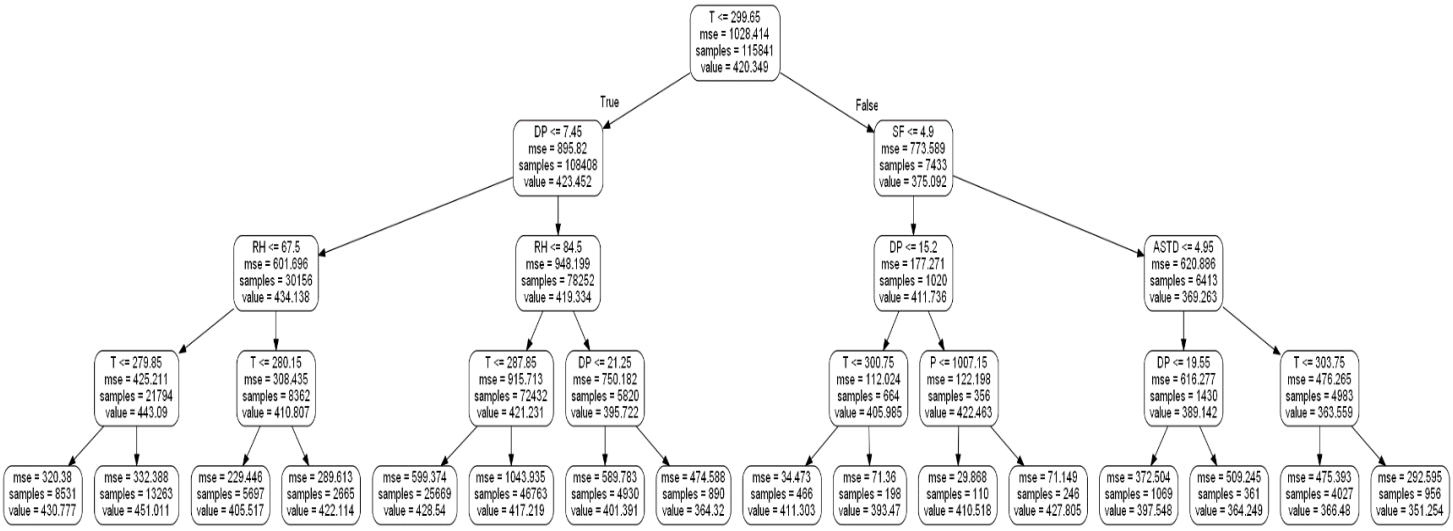


Figure 6.4. Decision Tree model for RSSI prediction. Each node depicts the split selection parameter, the resulted mean squared error (MSE) of the split, the number of samples included to the specific region and the mean RSSI value based on the samples of the region.

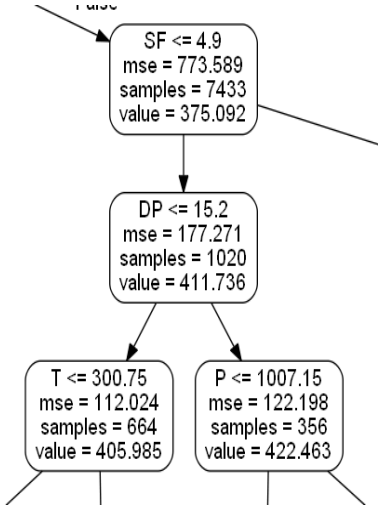


Figure 6.5. Resulting child node for RSSI.

The process described above can be followed until different depth levels are reached. However, there is a critical number of levels that improve the fit to the training data but

decreases the fit to the test data; in other words, an overfitting effect comes in to play. This issue is shown in Figure 6.6, where up to a tree depth of 9, the training and testing R^2 values coincide. For larger tree depths, the testing R^2 values tend to converge in a value of 0.9 and no more improvement can occur. For this model, a tree depth of twenty is selected as it provides the optimum performance metrics values for the testing subset. Larger tree depth values can result in a decrease of model performance.

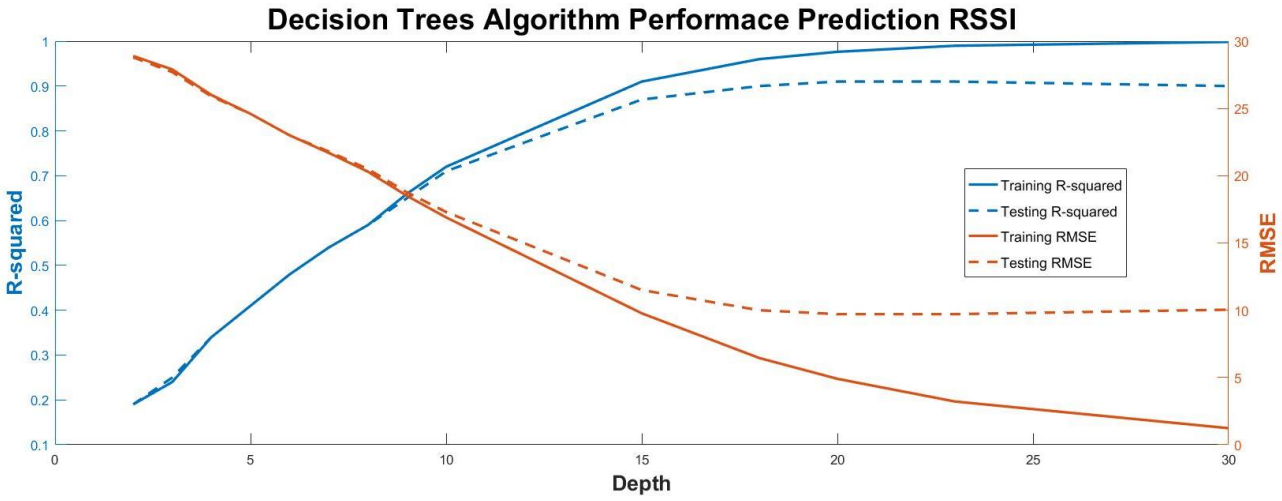


Figure 6.6. Performance metrics plot for Decision Tree algorithm against different number of Neighbors.

6.3.3 Random Forest

The decision tree algorithm exhibits high variance that can correspond to less prediction accuracy. That means that if we apply a certain model built from the DT algorithm to different data sets, we will probably observe significantly different results. Instead, the prediction accuracy of a model with low variance would be consistent in any data set. To improve this issue, the method of bagging is applied, to improve DT algorithm's performance. This method makes use of the general principle that averaging different sets of data will result in lower variance. Therefore, if we train our

DT model for different samples from a single training data set and average their predictions, we will achieve a model with higher prediction accuracy.

The RF algorithm elaborates on this principle and by decorrelating the trees to achieve an even higher performance. The way to do so, is described previously, in Chapter 4.

The results of running a RF algorithm in our data set, improved compared to the single decision tree method. A coefficient of determination, R^2, for the model evaluation of 0.95 is achieved and is indicative that this occurs with not so many different trees, as shown in Figure 6.7. Approximately, 20 different trees are sufficient to construct a model with the highest possible accuracy. Therefore, considering the lower complexity of a RF model with less trees and since more trees do not provide better results, we selected a value of twenty for our model. It is also important to notice that training and testing R^2 do not coincide for any number of trees; instead there is a constant difference of 0.05. On the other hand, the RMSE begins with a value of 4.5 for two trees and rapidly reduces to less than 3 for the rest of the number of trees. A significant advantage of the RF algorithm is that by adding more trees we do not risk overfitting.

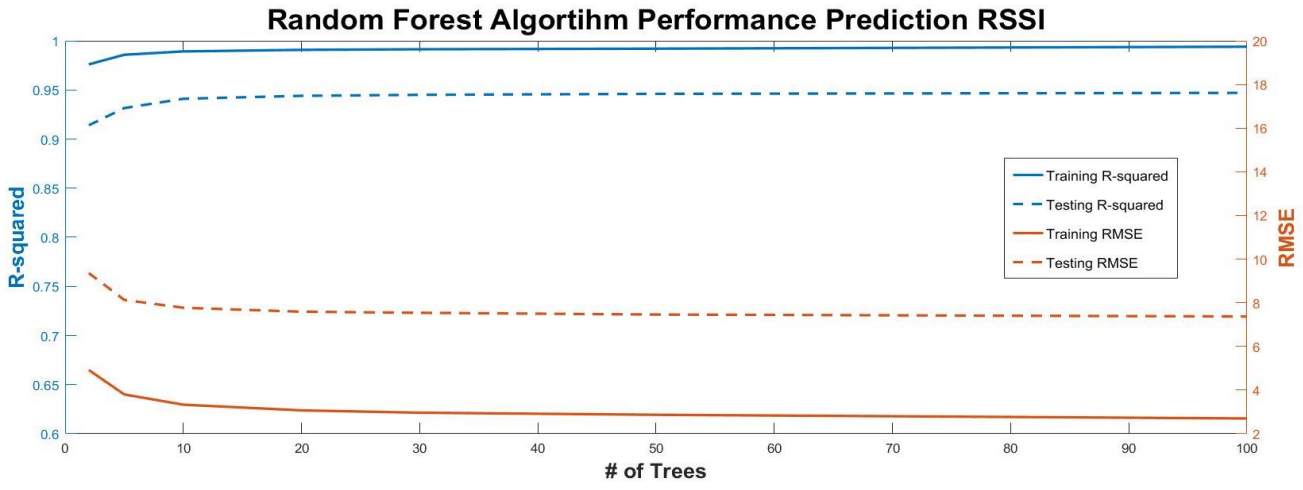


Figure 6.7. Performance metrics plot for Random Forest algorithm against different number of Trees.

6.3.4 Gradient Boosting Algorithm

Another approach to improve prediction performance is gradient boosting, which can be applied to many machine learning methods. The major difference of this approach is that instead of training the model concurrently, as is done in decision tree and random forest methods, it trains it sequentially. Specifically, the first decision tree is fitted to the training data set, the second is fitted using the residuals of the previous one, then this new tree is added to the fitted function in order to update the residuals. This procedure is followed by fitting rather small trees -few number of terminal nodes- to the residuals and that way we improve the model in areas where it does not perform as well. An additional shrinkage parameter λ exists, that allows more and different trees to be fitted in certain areas where higher residuals appear.

We selected to fit our Gradient Boosting Regressor model to our data set by using a tree number -iterations- of 5000 and a learning rate of 0.05. The model performed sufficiently and was comparable to the RF by achieving the value R^2=0.941 and an RMSE=7.71. In Figure 6.8 the model improvement is illustrated in terms of predictive accuracy as compared to the boosting iterations, which in the case of regression problems equals the number of trees, since each iteration uses a single tree.

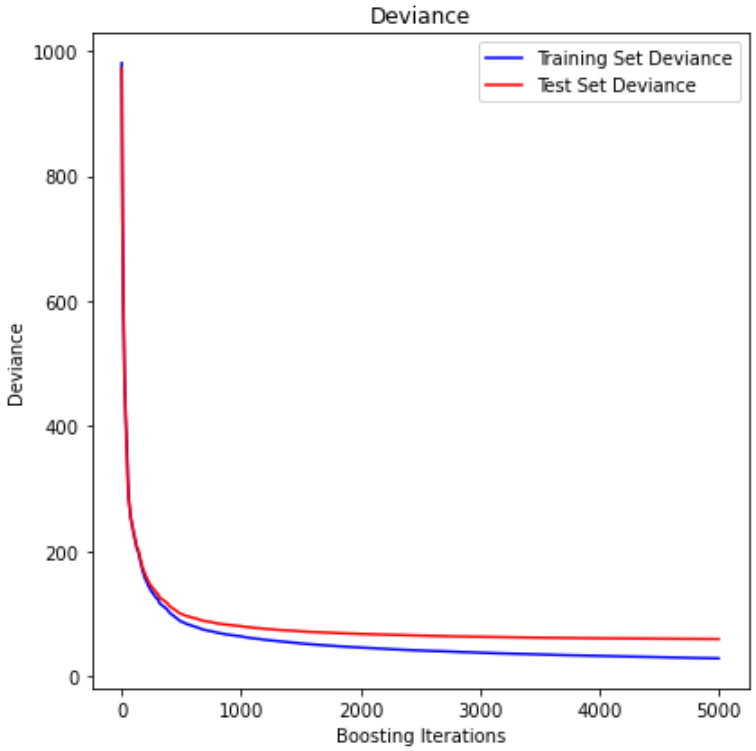


Figure 6.8. Training and testing MSE as n trees are added to the GBR algorithm.

Impurity-based feature importance can be misleading for features with many unique values. As an alternative, we can compute the permutation importance. Permutation feature importance is a model inspection technique that can be used for any fitted estimator when the data is tabular, something very useful for non-linear estimators. The permutation feature importance is defined as the decrease in a model score when we randomly shuffle the value of a single feature. This procedure breaks the relationship between the feature and the target, thus the drop in the model score is indicative of how much the model depends on the feature. This technique benefits from being model agnostic and can be calculated many times with different permutations of the feature. For the current test subset, both the impurity-based and permutation methods -as a box-plot figure- identify the air temperature as the most significant predictive feature, whereas the rest of the ranking differs as shown in Figure 6.9.

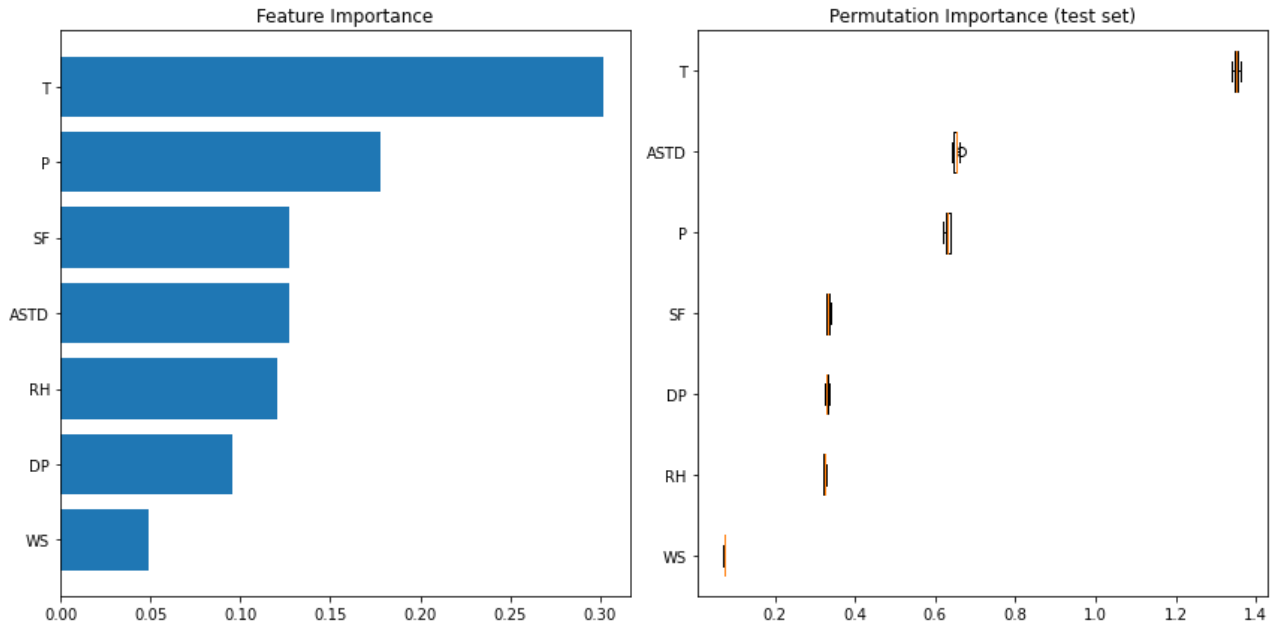


Figure 6.9. Feature Importance (left) and Permutation Importance (right)for the GBR model estimators.

6.3.5 Artificial Neural Network Algorithm

Another powerful form for ML model's implementation is the ANN. An ANN resembles the function principles of a human's brain. The analogy is that in place of the brain neurons, an ANN model uses the nodes that deliver the signal's information within the network. The basis of an ANN is the perceptron, which receives an input signal x_i with a weight of w_i, a bias b, added through a summation junction and then passed through an activation function f to provide the final output, as shown in Figure 6.10. All these complex combinations of input signals, weights, activation functions and biases for each layer, allows for a sufficient modeling of highly non-linear relationships.

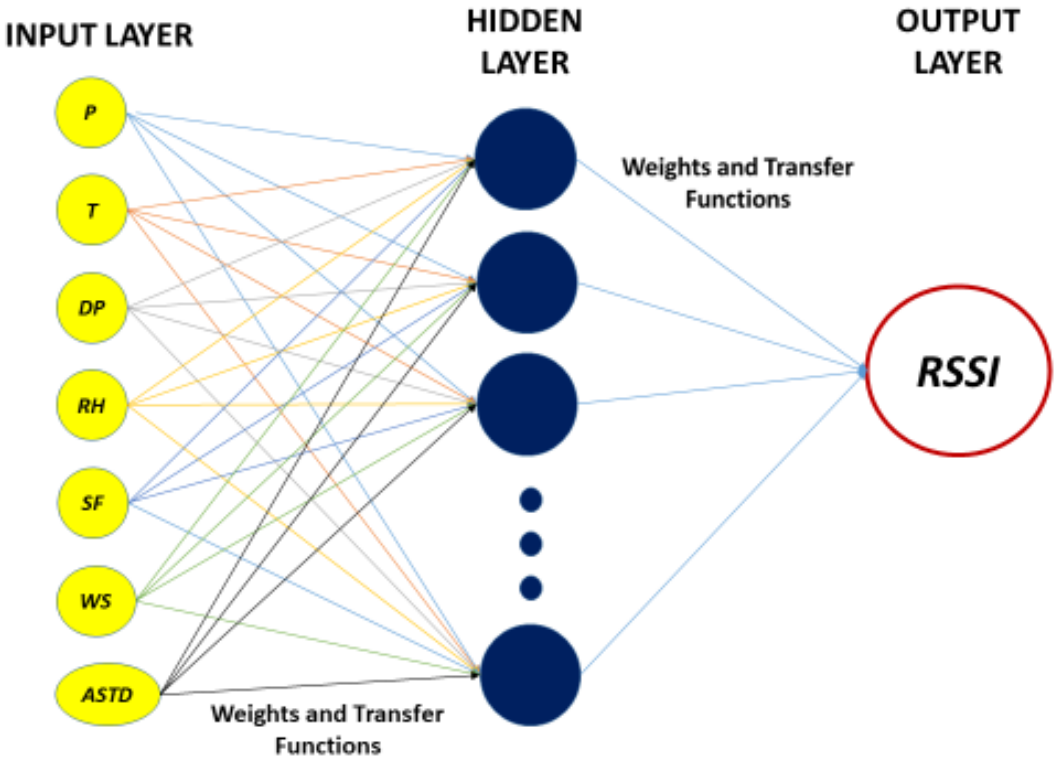


Figure 6.10. The multilayer perceptron network architecture for RSSI prediction.

The internal layer perceptrons, those in between the input and output layers, receive the summed weighed values of the input parameters, i.e. temperature, pressure, dew point, solar flux, wind speed, relative humidity and air-sea temperature difference, and provide the "input" signal for the output layer. All three layers are fully connected to each other. The proposed ANN architecture contains seven nodes in the input layer and a single node in the output layer RSSI, which receives the weighted input signals from the hidden layers along with a bias and represents the values of the predicted variable. Other ANN architectures could include more than a single hidden layer.

We used a standard Levenberg-Marquardt learning method to train a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons, for an extended number of nodes, i.e. 2-100, in order to map the macroscopic meteorological parameters obtained from the weather station to the RSSI parameter obtained from the

FSO system. The Neural Net Fitting application of MATLAB was used to perform all the ANN computations.

The ANN algorithm applies a three-fold division of the entire data set, i.e. the training, validation and testing subsets. The training data subset is presented to the network during training and the network is adjusted according to its error. The validation subset is used to measure the generalization of the network and stop the training when generalization stops improving. The testing subset is used to independently measure the performance of the network after training. Seventy percent of the entire data set was assigned to the training of the network, 15% to the training validation, and 15% for the network's performance testing. Figure 6.11 shows the performance of the MLP ANN algorithm for the RSSI prediction based on the seven meteorological parameters. We trained, validated and tested the algorithm for different number of nodes of the hidden layer, which resulted in significant decrease of the model prediction error approximately up to the 20-th node. Further increase of hidden layer nodes, provides slightly better results but also a substantially bigger computational cost. Figure 6.12 shows the error distribution for the training, validation and testing subsets, which appears to follow a normal distribution indicative of a very good performance of the prediction regression model, also shown in Figure 6.13, for a node number of 50 and a coefficient of determination $R^2 = 0.921$, both for the training and the testing phase.

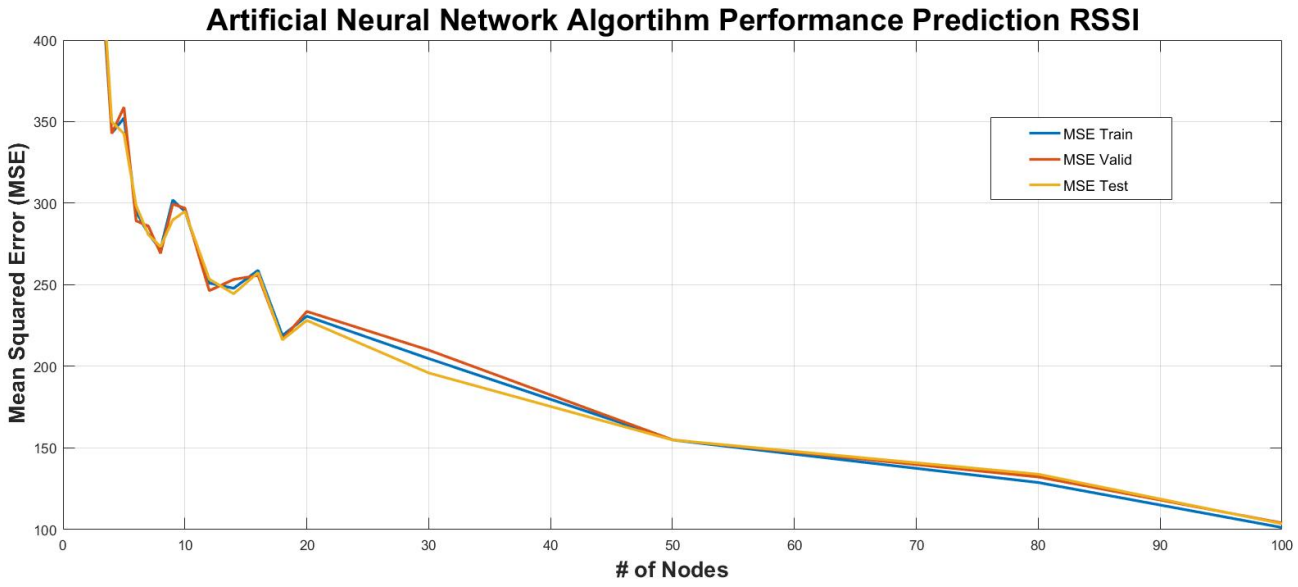


Figure 6.11. Performance metrics plot for MLP ANN algorithm against different number of hidden layer's nodes.

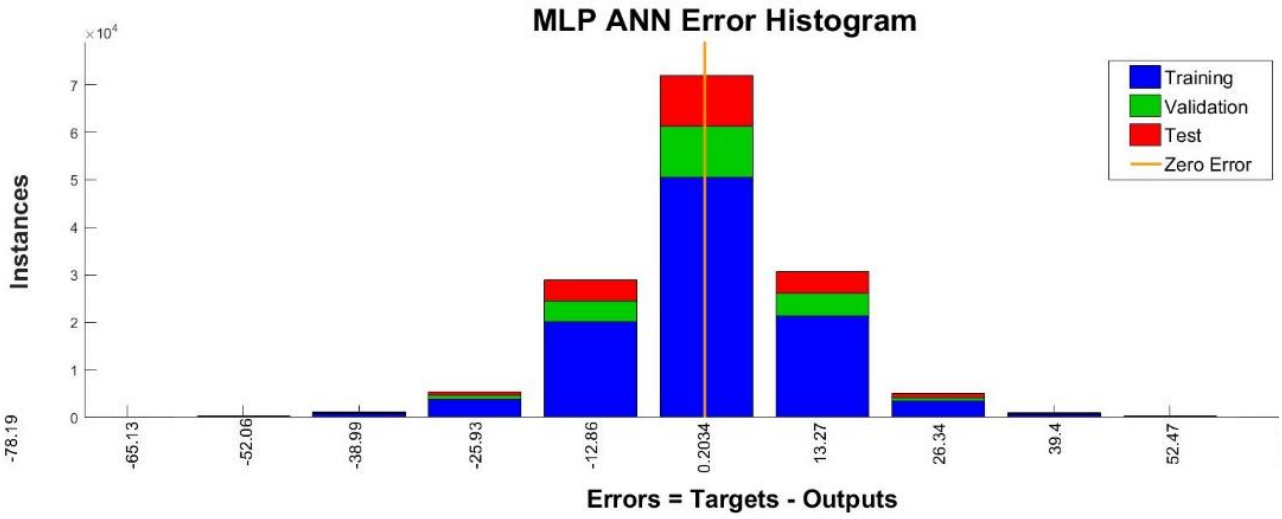


Figure 6.12. Error Histogram of the proposed ANN architecture for RSSI prediction for the training, validation and testing phases.

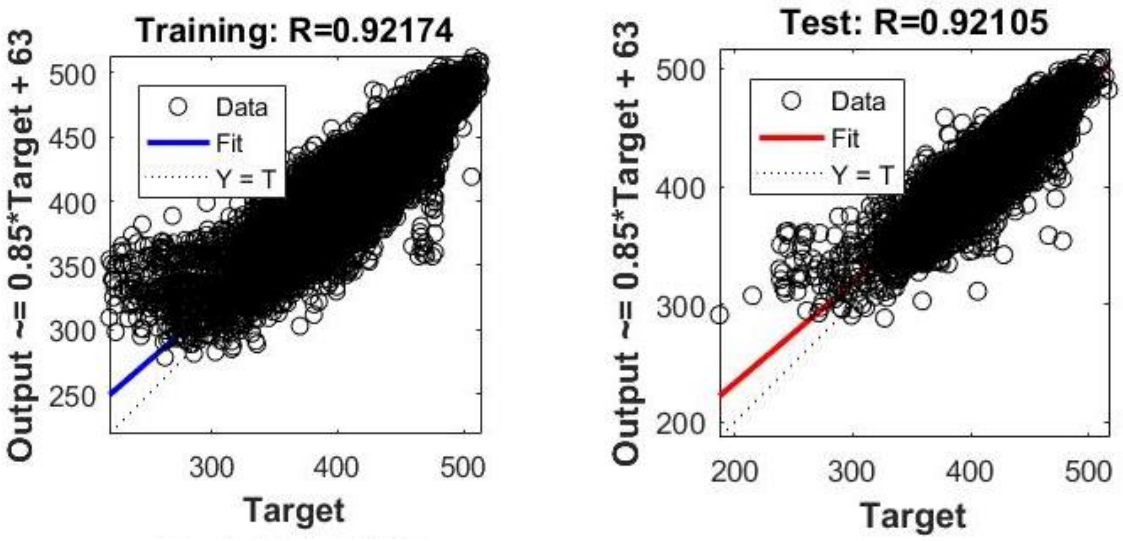


Figure 6.13. Regression Fits for the proposed ANN architecture for RSSI prediction.

6.3.6 Model Comparison and Discussion

To create a forecasting model for the RSSI parameter of an FSO connection, five ML algorithms were used. The challenge of fitting a conventional linear or non-linear model, i.e., polynomial, was extremely difficult due to the size of the data set and the significant variation of the independent variables. A second-order polynomial in particular performed quite poorly when it was first fitted to the data set, with R^2 values of 0.36 for the training subset and 0.05 for the testing subset. Therefore, more complex models were needed to get appropriate RSSI parameter prediction accuracy. Five distinct machine learning (ML) algorithms were trained and evaluated, and their performance was assessed using the coefficient R^2 and RMSE as performance indicators. Table 6.2 summarizes the results for each model, for the training, validation (if applicable) and testing phase. The ANN proved to be the most accurate model with an $R^2 = 0.94867$ but also very high in computational time cost (approximately three hours to train, validate and test the model). On the other hand, all three tree-methods

provided comparable results within a relatively short training period of time. Finally, the KNN algorithm resulted in a slightly less accurate model, yet much better than the polynomial and still statistically significant for RSSI prediction.

Despite using a much larger data set that covered an entire year and included measurements from various seasons, this study's ML algorithm proved to be significantly more reliable and accurate than the polynomial model created in the previous chapter, according to comparisons with earlier studies. The value range of the observed meteorological parameters is substantially wider than those in [J1, J2], as indicated in Table 6.1. However, the prediction accuracy performance of the models used in this study, outperformed the previous ones since they achieved an R^2>0.9, which is significantly higher comparing to an R^2~0.7 of the previous models. Regression models for RSSI prediction developed in [183, 184], also did not achieve an accuracy higher than 70%. Definitely, the ML algorithms provide a more efficient method to predict the RSSI parameter of an FSO link. However, the ML algorithms provide in a less intuitive prediction model and require a deeper comprehension in order to interpret the results.

Table 6.2. Performance comparison of the Baseline and five ML algorithm models for RSSI prediction in terms of R-squared (R^2) and root mean square error (RMSE).

Approach	R^2			RMSE		
	Training	Validation	Test	Training	Validation	Test
Baseline	0.36	-	0.05	-	-	-
KNN	0.93	-	0.85	8.29	-	12.48
DT	0.9764	-	0.91	4.9	-	9.71
RF	0.994	-	0.947	2.7	-	7.37

GBR	-	-	0.9417	-	-	7.71
ANN	0.9496	0.9468	0.94867	10.06	10.19	10.17

6.4 Summary

A machine learning-based method for estimating the RSSI parameter of an FSO connection over a marine environment based on macroscopic meteorological observations was described in this chapter. Seven parameters, including air temperature, wind speed, solar flux, dew point, relative humidity, air pressure, and air-sea temperature difference, were obtained over the course of a year from a commercial FSO system and a weather station, respectively, and were used to test the proposed approach. Five popular ML algorithms were trained in order to construct a robust model to accurately predict the link's performance in terms of received signal strength. The results showed a significant improvement compared to traditional regression modeling techniques, i.e polynomials, and their prediction accuracy performance measured by the coefficient of correlation, R^2 and the RMSE was extremely promising for even more complex predictive modeling. The significance of each of the independent variables was also studied with the gradient boosting regression algorithm by using two different approaches. Both of them showed that air temperature influences the output parameter the most. While all five ML methods showed a high degree of RSSI prediction accuracy, the ANN approach resulted in the most accurate model in terms of R^2, i.e. 0.94867, while the RF in terms of RMSE values, i.e. 7.37. ANN and GBR did require a significant computational time, while the three other methods gave their results in a much shorter training time. In conclusion, this work offered a full understanding of RSSI prediction accuracy utilizing several machine learning techniques that demonstrated to be remarkably accurate to model such a connection of complicated

systems. The superior performance of ML approaches compared to the widely used regression method shows that ML is the best modelling option when overall prediction is the goal and the volume of data is high, that ML approaches allow for the ability to tune hyperparameters to enable optimal performance, and that trained models can either be used for continuous streamflow RSSI predictions or improved by adding more data.

CHAPTER 7

Machine Learning Algorithms for Optical Turbulence Prediction over a Maritime Environment

This chapter includes the results that have been published in the peer-reviewed journals [J4 and J5]. It consists of two parts, the first, that presents the ML-based turbulence modeling approach for the data acquired at the NPS premises and the second that presents the equivalent approach for the data acquired at the HNA premises.

7.1 Introduction

The purpose of the first part of this chapter is two-fold. First, to introduce the application of machine learning algorithms in modeling the refractive index structure parameter (C_n^2) and estimate its value through regression analysis of macroscopic meteorological parameters, obtained from the NPS site. Secondly, it applies well known mathematical expressions to estimate the outage probability of a notional FSO link, based on the strength of the optical turbulence and model the link status (On-Off) based on macroscopic meteorological parameters, by utilizing a DNN classification algorithm.

The second part, presents the ML-based modeling approach of the refractive index structure parameter (C_n^2), based on macroscopic meteorological parameters and C_n^2 values, obtained from the HNA experimental site. To do so, four machine learning

algorithms were employed, including a Random Forest (RF), a Gradient Boosting Regressor (GBR) and two Neural Networks (a single layer and a deep network).

The fluctuations of the irradiance intensity on the receiver of a terrestrial FSO system, caused by the propagation through the atmospheric turbulence, result in a phenomenon called "scintillation". It can lead to significant power losses and eventually communication disruption. Turbulence strength can be classified in the weak, moderate and strong regime. It can be derived from theoretical formulas that shorter wavelengths experience smaller irradiance fluctuations, which is significantly important to be modeled. For that reason, many studies have been devoted to develop the probability density function (PDF) of the irradiance fluctuations to the receiver for every turbulence strength regime. This chapter utilizes experimentally derived C_n^2 data, with values spanning over every regime, therefore we apply the gamma gamma (GG) distribution which is the most suitable and can also be expressed in closed form. The PDF of this distribution is given by Eq. (3.32).

As described in the previous chapters, the atmospheric turbulence causes fading on the propagating laser beam, which within the context of laser communications affects the reliability of an FSO link. The optical turbulence effects on the beam include the intensity decrease of the received signal and below a threshold value it can even cause outage of the link. The frequency of the signal intensity fluctuations, as compared to the bit rate of the channel, characterizes the fading statistics that describe this channel, as either fast or slow. A performance metric that is useful for both fast and slow fading statistics is the outage probability (P_{out}) of the link, which is given by Eq. (3.66).

7.2 NPS Experimental Setup

The whole experimental setup has been installed on the roof of Spanagel Hall at the Naval Postgraduate School, at an elevation of about 20 meters above sea level, and was a follow up experimental campaign to that conducted in a near maritime environment [J3]. Figure 1 depicts the experimental apparatus, consisted of a tripod with four instruments: an IRGASON (Integrated CO2/H2O Open-Path Gas Analyzer and 3D Sonic Anemometer), an infrared radiometer, a net radiometer, and a global positioning system (GPS) receiver. The entire system was oriented to the northwest (facing Monterey Bay). This allowed the bay's onshore winds to flow across the device, while the building prevented southeasterly winds. From late January till late November 2021, continuous data collection was conducted. The obtained data included air temperature, ground (roof) temperature, solar flux, wind speed, sonic temperature and water vapor concentration readings.

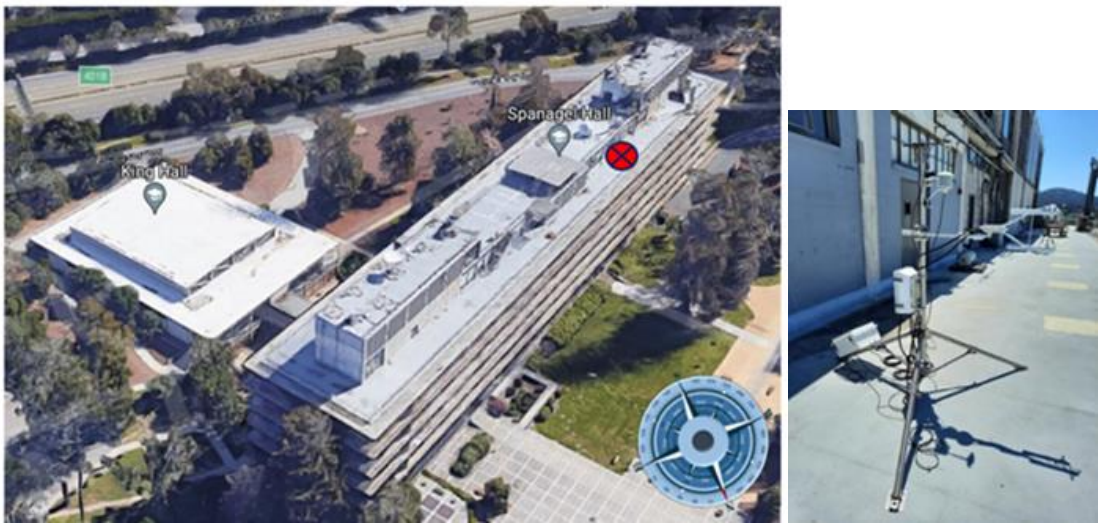


Figure. 7.1. The experimental setup location on the roof of Spanagel Hall (left) and the instruments utilized to obtain the macroscopic meteorological parameters.

The IRGASON is a three-dimensional sonic anemometer with an integrated optical gas analyzer. By synchronizing gas, wind, pressure, and temperature measurements, the optical turbulence values were calculated. The infrared radiometer is a sensor used to detect surface temperature by monitoring the surface's blackbody radiation. Part of the shortwave light (between 280 and 4000 nm) emitted by the sun is reflected by the Earth's surface. Atmospheric and terrestrial molecules absorb solar radiation and emit it as longwave thermal radiation (between 4 and 100 micrometers). Net radiation flux is the difference between inbound (downwelling) and outgoing (upwelling) radiation. Differences between incoming and outgoing shortwave radiation provide information regarding the amount of solar radiation absorbed by the ground. Consequently, the SN500SS net radiometer was utilized to compute the solar flux. Finally, the GPS receiver GPS16X-HVS served as a clock for this experiment, since the datalogger's clock was set to synchronize with the GPS data automatically.

The C_n^2 values were extracted from the sonic anemometer based upon the Kolmogorov's mathematical expression that relates, within the inertial subrange, the amplitude and frequency of the sonic temperature fluctuations as follows,

$$PSD(f) \propto C_T^2 f^{-5/3} \tag{7.1}$$

which results in a -5/3 slope on a log-log plot of the PSD against frequency. Therefore, fitting a line on PSD gives us the C_T^2 and from Eq. (2.19) we get C_n^2. The ML algorithms were then trained on the logarithm (base 10) of those estimated C_n^2 values.

7.3 Results and Analysis (NPS Site)

The data collected from the instruments presented on the previous section included air temperature, wind speed, water vapor concentration, solar flux, sonic

temperature and ground temperature as independent macroscopic meteorological parameters. As also described previously, we utilized Kolmogorov theory to extract the refractive index structure parameter from the sonic temperature and wind speed obtained from the IRGASON, which used as the dependent variable for our modeling purposes. An almost 40-second interval power spectral density (PSD) of the sonic temperature (i.e 2048 observations at a 50Hz frequency) provided the C_n^2. Average values for all other measurements were calculated over the same time interval. That gave us totally almost 525000 observations (point measurements) of turbulence strength, i.e. C_n^2, and routine meteorological parameters which were stored for further analysis. The complete dataset can be found in [185].

7.3.1 Data Analysis

The data obtained spanned over a period of almost eleven months. That means that the meteorological parameters, primarily the temperatures, had a significant variance due to their seasonality dependence. To include this seasonality dependence on our analysis, as well as to have a more effective manipulation of the datasets, we divided the whole dataset in four distinct sets of almost equal size. These different sets represent, but not precisely coincide, with the four seasons. However, since they are almost match to the four seasons, we treated each one of them as a season representative (i.e. autumn (1/9 – 20/10), winter (25/1 – 31/3), spring (1/4 – 30/6) and summer (1/7 – 31/8)), as shown in Datasets 1 through 4 [191]. Our focus was primarily to the seasonal effect on the C_n^2. Figure 7.2 presents the distribution plot of C_n^2, in a logarithmic scale, for every season. Apparently, the refractive index parameter is a very complex metric to extract definite conclusions, since multiple non-linear phenomena affect its value,

however an approximate analysis can be done. To that end, Figure 7.3 shows the daily moving average for the entire period of every season. We notice that during winter, C_n^2 exhibits, in average, the lower values in contrary with summer where we observe the higher values. This is partially supported by the fact that the mean observed value of C_n^2 in winter is the lowest, whereas in summer is the highest. It is interesting though to mention that during winter we observed a higher maximum value for C_n^2 than in summer. During spring and autumn, the corresponding mean values reside within the rest two. In general, the time series of all seasons, follow the same pattern. Another characteristic of C_n^2 values, is that winter and autumn exhibit much lower variance as compared to summer and spring (Figure 7.2).

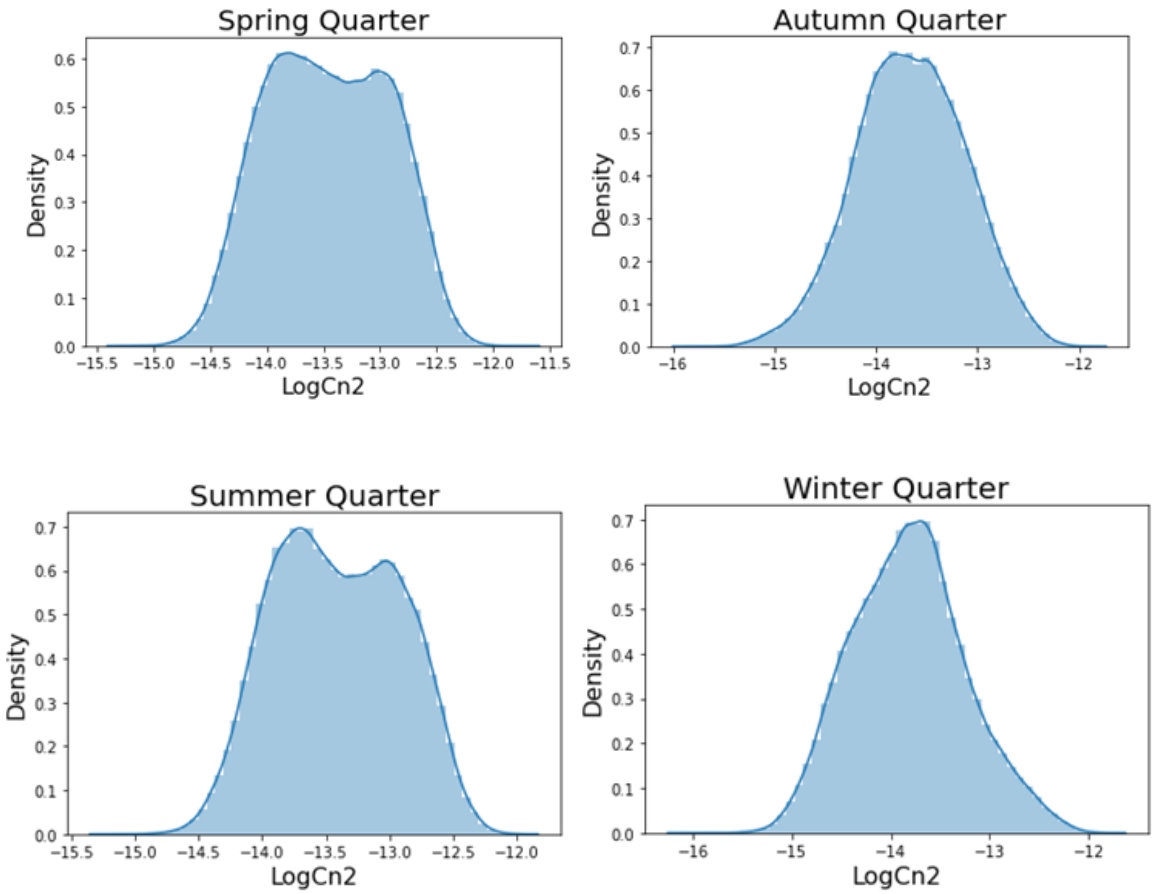


Figure. 7.2. Distribution plots for $LogC_n^2$ per season.

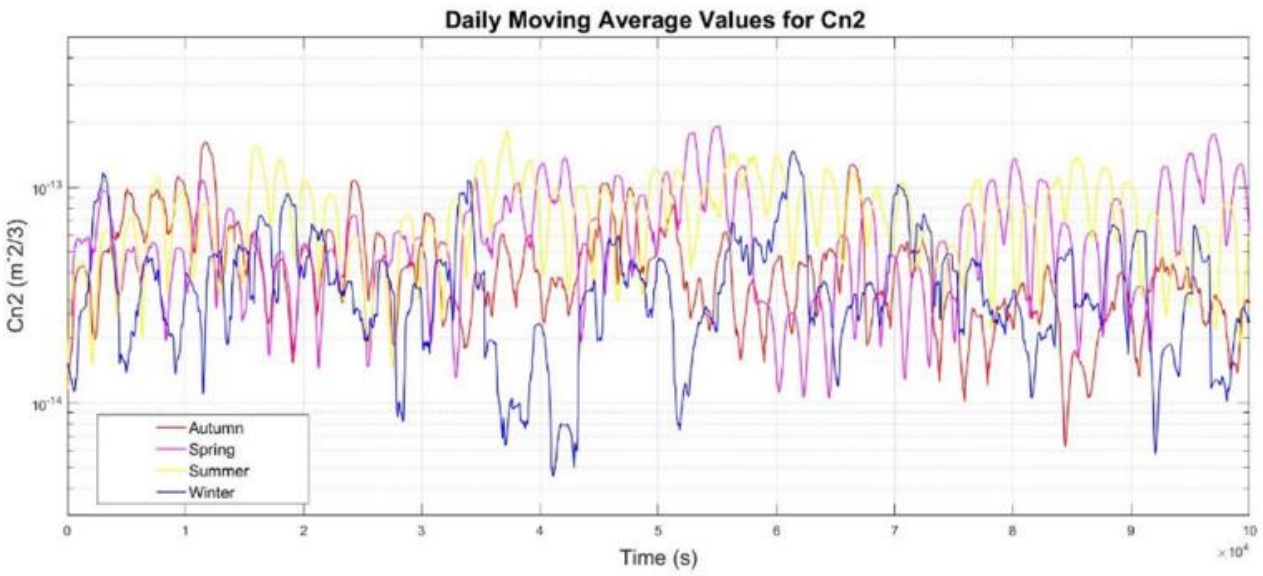


Figure. 7.3. The daily moving average of the C_n^2 for every season.

Finally, the correlation between the measured atmospheric parameters and C_n^2 for the entire period, is depicted in the heatmap of Figure 7.4. In this figure, we have included an additional parameter D_{temp}, defined as the difference between the ground and the air temperature, which has a significant negative effect on C_n^2. The darker a cell the stronger the anti-correlation between the corresponding two parameters. On the other hand, a lighter cell corresponds to a stronger correlation.

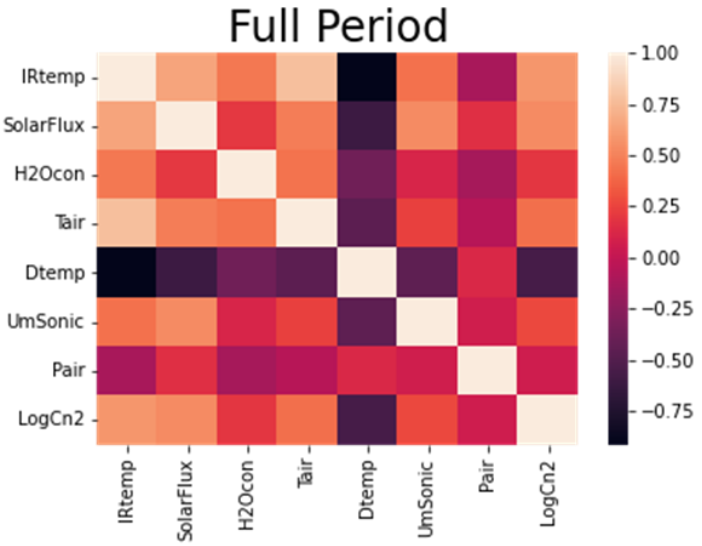


Figure. 7.4. Correlation matrix for the macroscopic meteorological parameters and C_n^2.

7.3.2 Regression Analysis for C_n^2 Modeling

The analysis of this chapter is based on the extended number of observations taken from January to November 2021. During this period, we didn't face any technical issues, therefore our dataset can be considered continuous. Additional data cleaning before analysis, reduced the data points to a total amount of 524786, on a ~40 second interval. The measured data obtained from all sensors, was logged into an .xlsx file format, i.e. six columns. Each data row was then compounded with the respective output value of C_n^2 for the same date/time. We excluded measurements with missing or non-physical values. That is, C_n^2 values above $2x10^{-12}$ $m^{-2/3}$ that are not considered realistic. After collecting, compiling and cleaning the data set, the ML process was initiated. The first step was to take the set of all observations and divide it randomly into two subsets. The first subset was used to train the model. The second subset was used to test the model once it had been trained. This second dataset is referred to as the test subset. The objective was to estimate the performance of the ML model on unknown data, i.e. data not used to train the model. There is not an optimal percentage in which to split the data set, rather the need to account for the computational cost to train and test the model as well as an adequate representativeness in the two subsets. We chose to split at 80% for training and 20% for testing. Certain performance measures should be used in order to measure how well the models performed comparing to each other, on predicting C_n^2 values based on unseen meteorological parameters. We applied the RMSE which represents the square root of the variance and the coefficient of determination, R^2, a number between 0 and 1 that measures how well a statistical model predicts an outcome. All data during the preparation phase has been scaled to a 0 to 1 scale so as to account for the different variance of each parameter. The data analysis was performed using the programming language Python (version 3.8) in a

Jupyter Notebook environment, a web-based interactive computing notebook, which allows the implementation of various libraries, e.g. Pandas, Numpy, Matplotlib, Seaborn, for advanced data analytics and visualization.

To execute the regression modeling analysis, we utilized six well-known ML algorithms, Artificial Neural Networks (ANN), Random Forest (RF), Gradient Boosting Regressor (GBR), k-Nearest Neighbor (KNN), Decision Trees (DT) and Deep Neural Networks (DNN). As mentioned above, the same analysis was repeated for all seasons and each algorithm was fine tuned to achieve the best accuracy. The following sections present these results and comment on the selected hyperparameters for each algorithm.

The ANN that best fitted the observed data during autumn was a single layer perceptron model that included 100 neurons in its hidden layer. The optimum batch size was 32 and the model trained over 400 epochs. A standard Levenberg-Marquardt learning method was used to train the feed-forward network with sigmoid hidden neurons and linear output neurons. The results of this algorithm showed a moderate to low accuracy, with $R^2 = 0.55$ and RMSE= 0.0916.

The results for the RF algorithm in our data set gave a significantly improved coefficient of determination, R^2, for the model evaluation of 0.78 which was the best value among all algorithms for autumn, as shown in Figure 7.6. Approximately, 300 different trees are sufficient to construct the model with the highest possible accuracy. Additionally, an RMSE = 0.064 showed a great improvement in the error of the predicted values.

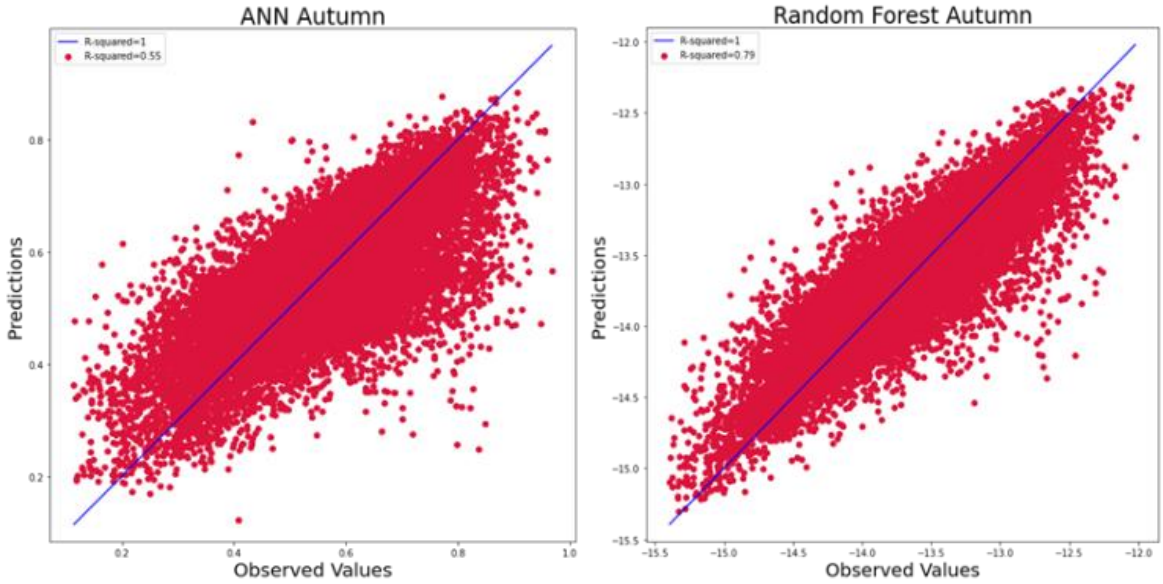
A Gradient Boosting Regressor model was used to fit the autumn data set by using a tree number -iterations- of 2000, maximum depth = 6, minimum sample split = 12

and a learning rate of 0.05. The model performed sufficiently and was comparable to the RF by achieving a value of $R^2 = 0.7$ and an RMSE = 0.075.

The KNN algorithm achieved its best performance for a value of k = 15 which resulted in $R^2 = 0.71$ and an RMSE = 0.073. That is, it slightly outperformed the GBR algorithm.

The DT algorithm, which is a simplified version of the RF model, as expected had poorer performance than RF. The optimum depth of trees was found to be 15 for this occasion with a corresponding $R^2 = 0.637$ and an RMSE = 0.083.

Finally, perhaps the more complex and demanding to fine tune, DNN algorithm, comprised of three layers of neurons (1st hidden layer=50, 2nd hidden layer = 30 and 3rd hidden layer = 10), ran over batches of 32 for a total number of 400 epochs. The results for this model were $R^2 = 0.61$ and an RMSE = 0.085. Figure 7.5 collectively presents the scattering plot for each algorithm.



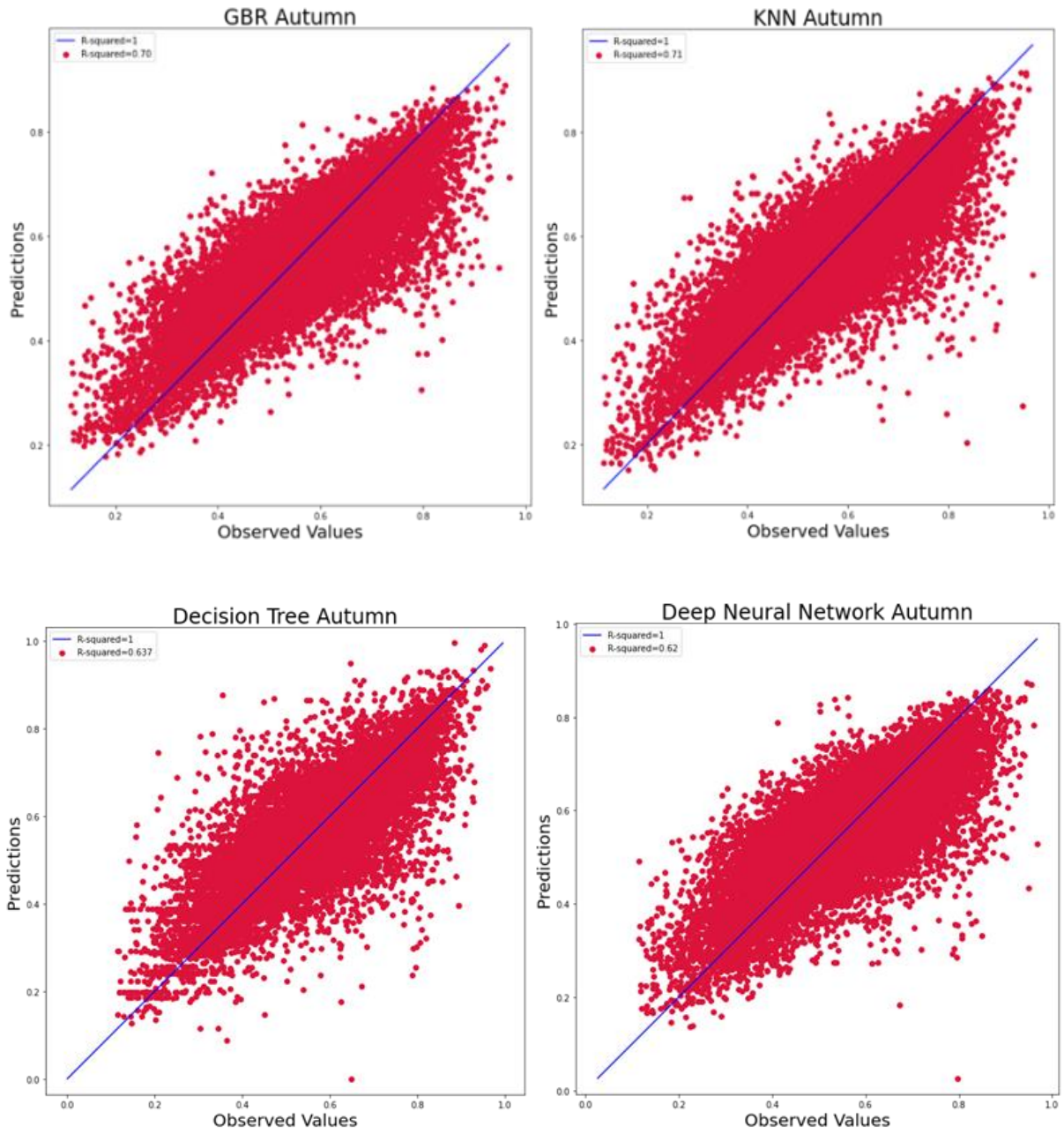


Figure. 7.5. Scattering plots of the six ML algorithms for the autumn data set.

The results of the performance for all ML algorithms are presented collectively in Figure 7.6, where the performance ranking is clearly depicted and shows that the RF exhibited the best fit.

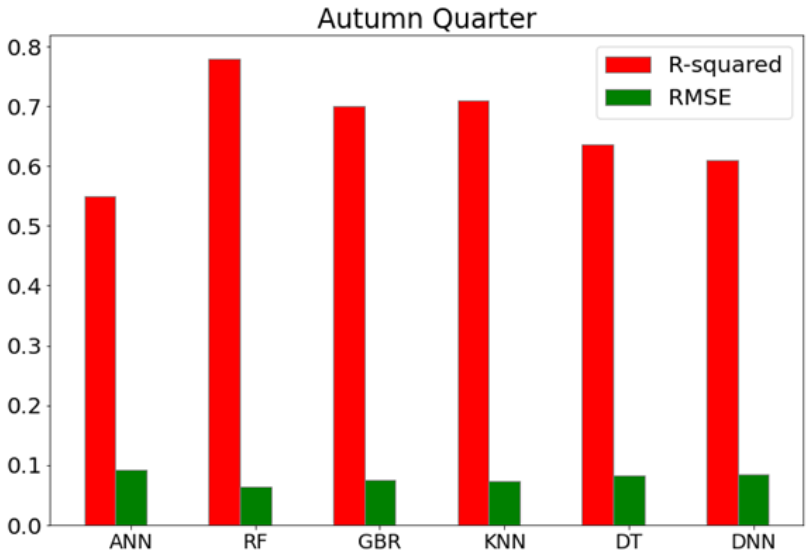


Figure. 7.6. R^2 and RMSE metrics for the ML algorithms autumn data set.

We presented detailed results for the performance of each algorithm with regard to the autumn dataset, but the same procedure was followed repeatedly for every seasonal dataset. Therefore, we directly plot in Figures 7.7 through 7.9 the corresponding algorithm performance metrics comparison.

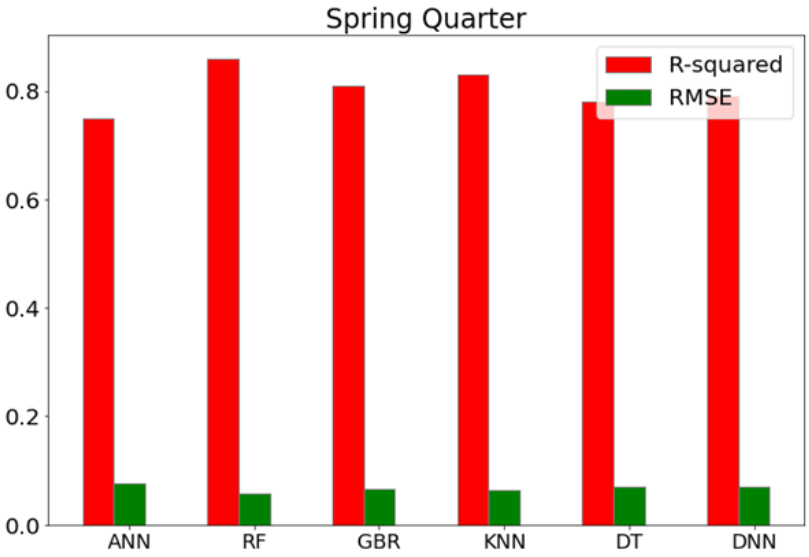


Figure 7.7. R^2 and RMSE metrics for the ML algorithms spring data set.

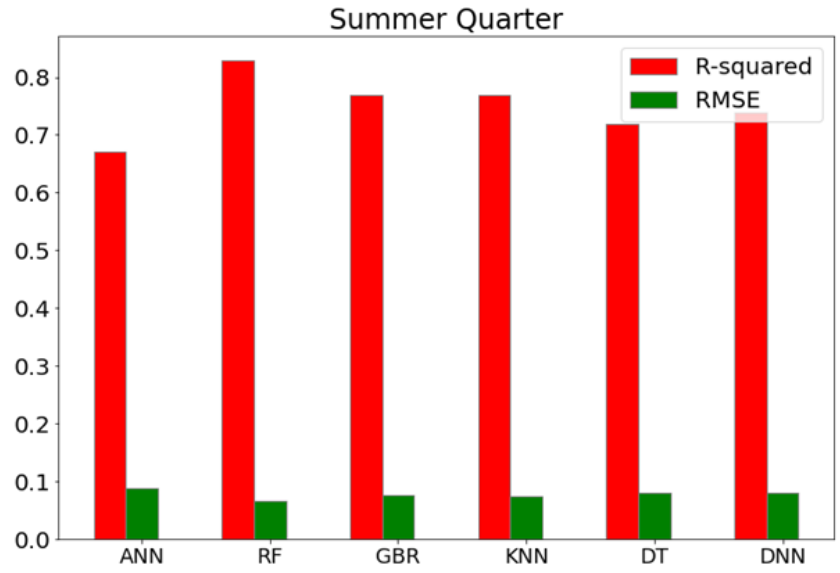


Figure. 7.8. R^2 and RMSE metrics for the ML algorithms summer data set.

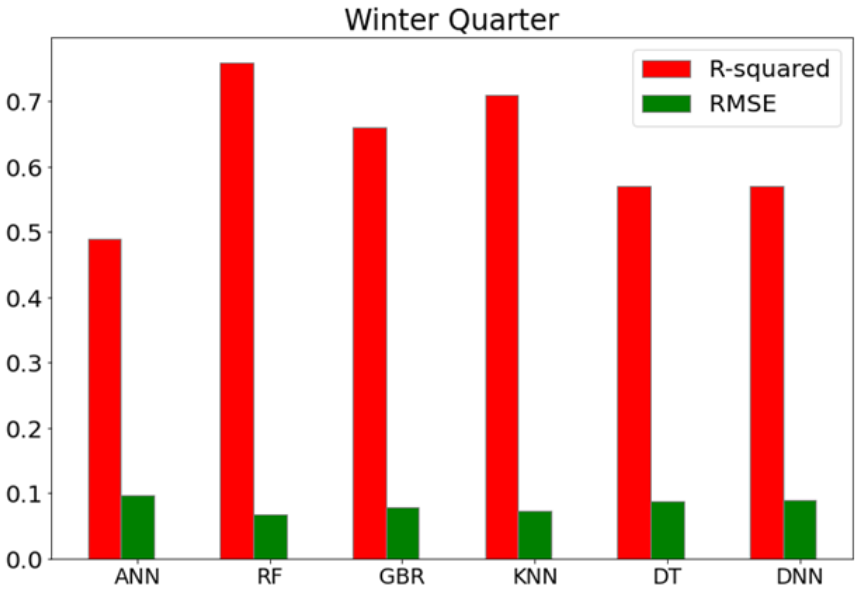


Figure. 7.9. R^2 and RMSE metrics for the ML algorithms winter data set.

To conclude this section, it is important to comment on the overall performance of each algorithm. RF appeared to be the best, since it achieved the highest metrics values on every season. Interestingly, the one followed was KNN, pretty close to GBR. DT and DNN also demonstrated similar performance to each other, but poorer than GBR. Finally, ANN had the worst fit over all datasets.

7.3.3 Outage Probability Calculation

The next step in our analysis was the estimation of the outage probability for a notional FSO link based on the experimental meteorological and optical turbulence data. Initially, the experimental C_n^2 data was utilized to compute the Rytov variance, Eq. (2.30) for a plane wave, σ_R^2, along with a typical wavelength λ = 850 nm and a link range, R = 3000 meters. The same parameters plus $d = \sqrt{KD^2/4L}$, with D = 20 cm for the diameter of the FSO receiver, allowed to calculate the environmental conditions dependent parameters α and β from the following equations,

$$\alpha = \left[exp\left[\frac{0.49\sigma_R^2}{\left(1+0.18d^2+0.5\sigma_R^{12/5}\right)^{7/6}} \right] - 1 \right]^{-1} \tag{7.2}$$

$$\beta = \left[exp\left[\frac{0.51\sigma_R^2\left(1+0.69\sigma_R^{12/5}\right)^{-5/6}}{\left(1+0.9d^2+0.62d^2\sigma_R^{12/5}\right)^{7/6}} \right] - 1 \right]^{-1} \tag{7.3}$$

We assumed a typical SNR level $I = \sqrt{\frac{\gamma}{\mu}} = 10\ dB$. By applying Eq. (3.66) we were then able to compute the corresponding outage probability for every C_n^2 measurement instance, thus a total of almost 525000 data points for all four seasons. The last step described in section 7.3.4, was to model the outage probability based on the six measured meteorological parameters and build a mathematical expression. A flowchart of the overall process is depicted in Figure 7.10.

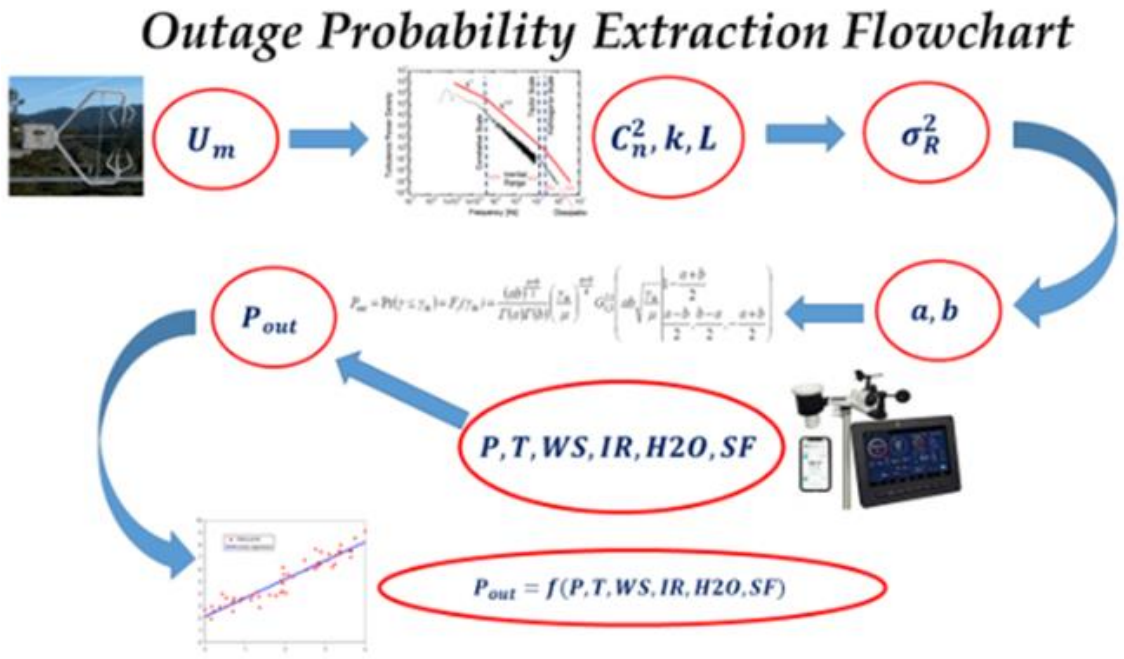


Figure. 7.10. The outage probability of a notional FSO link extraction flowchart.

7.3.4 DNN Classification

A received signal below a certain strength level, i.e. threshold, in an FSO link, would eventually result in fading and perhaps a total link interruption. To cope with this inherent disadvantage of laser communications technology, a different approach is followed, which brings into play a hybrid scheme, to include both laser and radio frequency, although the corresponding data rate in the radio frequency mode will be less than the laser mode. That way, whenever the received signal strength is below that threshold value of the receiver sensitivity, the operation mode of the link automatically switches to RF mode, in order to compensate the atmospheric, e.g. fog, haze, or other physical, e.g. LOS blockage, effects. Therefore, it is of great interest to be aware of the certain availability of the link.

This section presents a DNN approach to model the status of a notional FSO link, that is either functional (On) or non-functional (Off) in the laser mode. That way, we can be able to predict the status of the link, whether it will operate in the laser or radio frequency mode. To do so, we utilized the environmental data set per season as described in Section 7.3.1 and built a deep neural network (Fig. 7.11) to classify the link's operational status as functional (On) or Non-functional (Off). We assumed a required 99% availability for our link, that is any outage probability more than P_{out} = 0.01. Therefore, a "0" was attached to every row in our dataset with $P_{out} > 0.01$ and a "1" for $P_{out} < 0.01$. The resulted split of our experimental data for all seasons was quite unbalanced and showed that the vast majority of the observations imply a non-functional on its laser mode FSO link, under the aforementioned assumptions of outage probability estimation. However, it gives interesting insights on a realistic operational performance of an FSO link. Figure 7.12 presents the results on the "0" sand "1" s of the link. Similar allocation among "0" s and "1" s exhibited by the rest three seasons too.

The network has three hidden layers with 50, 30 and 10 neurons, respectively. A feed-forward back propagation algorithm was used, with a dropout rate of 0.5 per layer. The activation function for the three hidden layers was a rectifier (ReLU) whereas for the output layer a sigmoid function. In order for the algorithm to monitor the progress of the algorithm fitting, a binary cross entropy loss function was used and the Adam optimizer to adapt the gradient descent of the loss function. The algorithm was trained against the 80% of the dataset and tested over the rest 20%. An early stopping criterion was also introduced to avoid overfitting the model, which lead to a total of 189 training epochs for a batch size of 32.

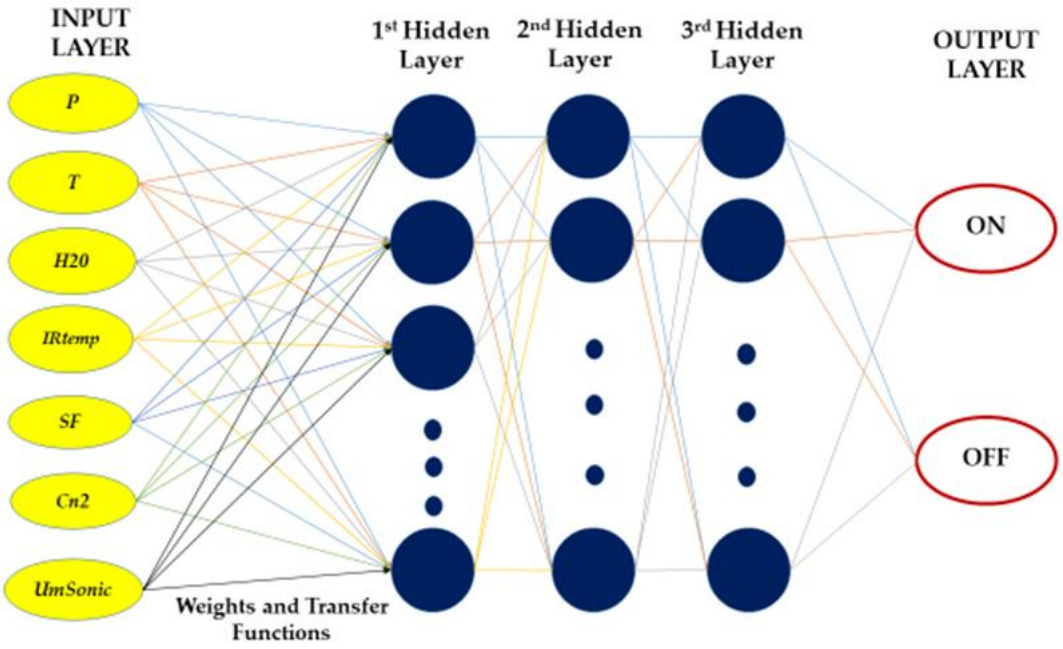


Figure. 7.11. The Deep Neural Network approach to classify the FSO link laser mode status.

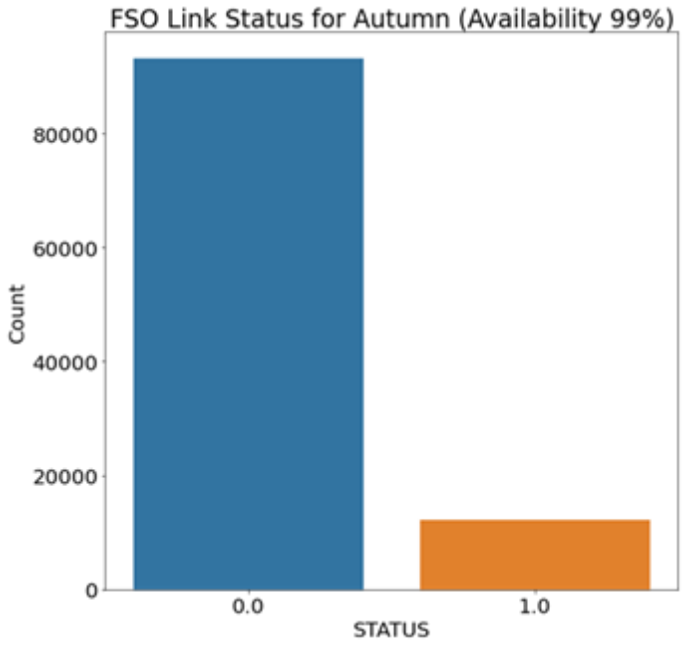


Figure. 12. The cumulative results of the FSO operational status for the Autumn season. “0” for non-functional and “1” for functional laser operation.

Figure 7.13 shows the progressive performance of the model throughout training, measured by the accuracy and loss both for the training and the validation set. The model exhibits a significant accuracy early on the epochs iteration, which after slight

variations seems to get stabilized after around 150 epochs. The early stopping criterion interrupted training at the 189th epoch. Figure 7.14, presents the confusion matrix of the DNN classification model. By definition a confusion matrix C is such that $C_{i,j}$ is equal to the number of observations known to be in group i and predicted to be in group j. Thus, in a binary classification the count of true negatives is $C_{0,0}$, false negatives is $C_{1,0}$, true positives is $C_{1,1}$ and false positives is $C_{0,1}$. As shown in Figure 7.13, we confirmed the excellent classification performance of the DNN, since we observe that false negatives $C_{1,0}$ = 0 and false positives is $C_{0,1}$= 39, which means that out of the almost 2400 "1"s only 39 false predictions made instead.

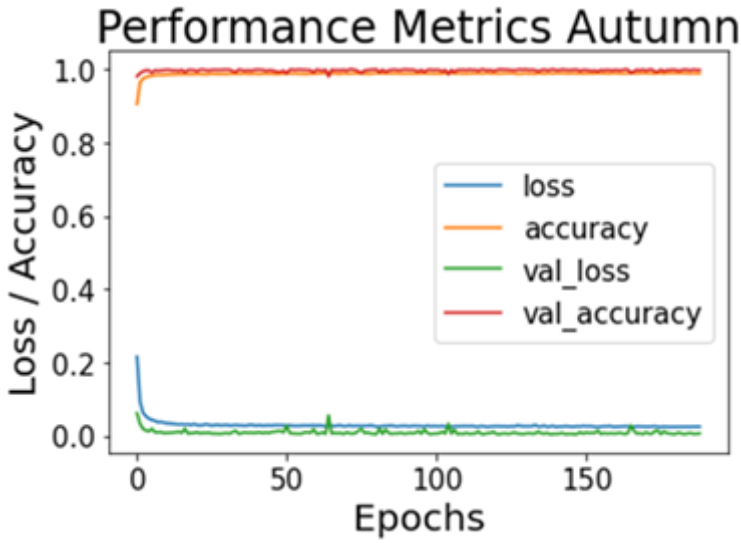


Figure 7.13. The loss/accuracy performance of the deep neural classifier for the Autumn season. An early stopping criterion interrupted the training after 189 epochs.

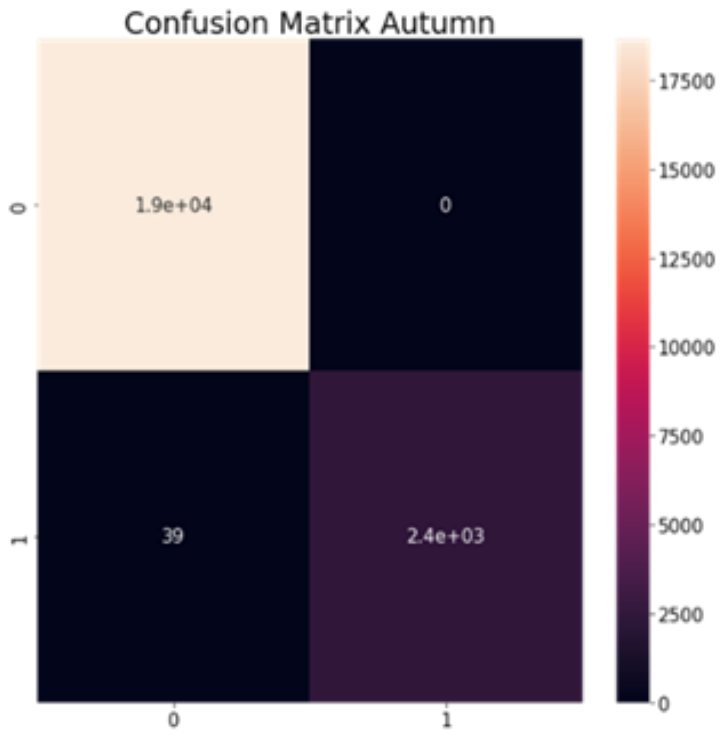


Figure 7.14. The confusion matrix for the DNN classifier during Autumn.

7.3.5 Modeling Outage Probability

The last section of the first part of this chapter, aimed to develop an easily interpretable mathematical model for outage probability estimation, based upon routine meteorological parameters and refractive index. A first order polynomial has been selected to provide a decent fit among seven independent parameters, i.e. IR temperature, solar flux, atmospheric water concentration, air temperature, wind speed, atmospheric pressure and logarithmic refractive index structure parameter, and the dependent P_{out} to the autumn dataset with a coefficient, R^2= 0.55. The mathematical expression that gives the predicted outage probability is given,

$$P_{out} = 0.349 - 2.423 * 10^{-4} IRtemp + 4.911 * 10^{-6} Solar\ Flux - 3.391 * 10^{-4} H2Ocon + 4.249 * 10^{-4} Tair + 6.489 * 10^{-4} UmSonic - 7.467 * 10^{-4} Pair + 1.849 * 10^{-2} LogC_n^2 \quad (7.4)$$

Figure 7.15 plots the predicted values of P_{out} obtained from Eq. (7.4) against the measured values obtained from Eq. (3.66) for the autumn dataset. The daily mean value – instead of raw data - for both values has been used for a clearer depiction. We notice that the model's predictions follow the general trend of the experimental P_{out}, thus it can be used to give very good estimates for the performance of an FSO link.

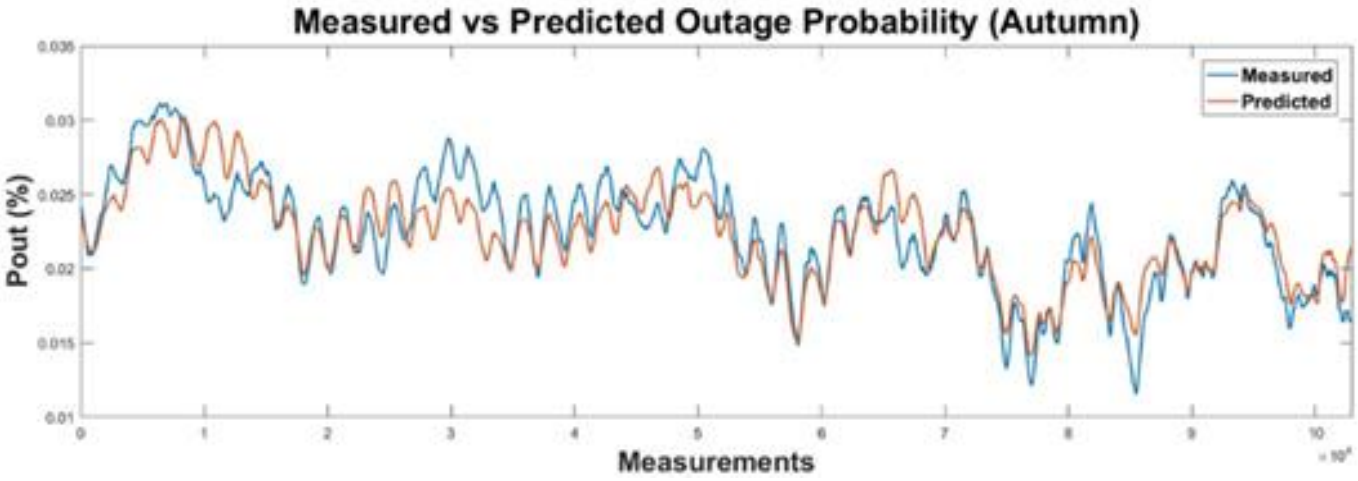


Figure 7.15. Measured vs Predicted outage probability time series for autumn.

7.4 HNA Experimental Site

The HNA experimental site is as presented in Chapters 5 and 6, with the addition of a BLS450 scintillometer, located in the vicinity of the MRV and WS-2000 to measure the atmospheric turbulence and heat flux over the path length. A scintillometer measures turbulence along the path between an optical transmitter and a receiver. Resulting in a path integrated C_n^2 measurement. Its operation principle is based on the modulation of light by atmospheric refractive index fluctuations in the air, the phenomenon called scintillation. Compared to conventional turbulence measurements with point sensors, a scintillometer gathers spatially representative results with lower statistical scatter and shorter averaging times. As a double-ended remote sensing system, the BLS450 also allows access to such a terrain (i.e. over water) without need

to install in-situ sensors. Figure 7.16 shows the MRV TS5000/155 FSO system, co-located with the BLS450 scintillometer and the ambient weather WS-2000, weather station on the roof of the laboratory building of the Hellenic Naval Academy.

Figure 7.16. The MRV TS5000/155 FSO system, co-located with the BLS450 scintillometer and the ambient weather WS-2000, weather station on the roof of the laboratory building of the Hellenic Naval Academy.

7.5 Results and Analysis (HNA Site)

The experimental measurements that took place during the last week of May 2022 and comprised the first compete dataset from the upgraded instrumentation setup of the HNA experimental site. An initial period of a week (24 to 31 May) was devoted for the data collection, analysis and ML-based models construction and validation. Furthermore, the observed meteorological data was used for C_n^2 predictions based on two theoretical models. The main goals of this research analysis were i) the regression modeling of the refractive index structure parameter (C_n^2) using ML algorithms and the assessment of their prediction accuracy, and ii) the application of ML algorithms for the classification modeling of the strength level of the C_n^2 parameter (i.e. low, high).

7.5.1 Data Set

During the aforementioned period, the observed C_n^2 parameter values from the BLS450 scintillometer were logged once per minute. The same time interval was used for the atmospheric data collection and storage from the WS-2000 weather station in order to accurately match them with the C_n^2 measurements and compile them in an .xlsx file. A few technical issues, such as system resets and line-of-sight link blockages due to maritime traffic, resulted in a few missed measurements. The meteorological conditions during the experiment were pretty stable, with air temperature values ranging from 20 to 29 ^{o}C, relative humidity within 45 and 85 percent and very low average wind speed. A key parameter of the meteorological data was the air-sea temperature difference (ASTD). To extract this parameter, we used the online weather statistics database [186]. The entire dataset was screened and redundant recorded data excluded to end up with a clean dataset including 8055 rows and eight columns with the meteorological parameters and the respective output value of C_n^2 for the same date/time. Therefore, a single user-friendly file was compiled for further process and analysis.

7.5.2 Regression Modeling Results

The first part of the analysis is devoted to the modeling of the refractive index structure parameter by using four machine learning based regression algorithms, namely a single layer neural network applied in the *Neural Fitting* application of Matlab, a deep neural network, a gradient boosting regressor and a random forest applied in a *Jupyter* notebook of *Anaconda* environment using Python language. Two different software application approaches were followed to compare the prediction

accuracy of a built-in model with a user defined model which allows for much more flexibility.

The *Neural Fitting* application allows for data selection, creation and training a network and performance evaluation according to the mean square error and regression analysis. A single hidden layer feed-forward network with sigmoid hidden neurons and linear output neurons was created in order to fit the seven meteorological parameters (inputs) to the $logC_n^2$ (output). The network was trained either with a Levenberg – Marquardt backpropagation algorithm or with a Bayesian Regularization algorithm. The first one requires more memory but less time to train the model. Training stops automatically when generalization stops improving, as indicated in the mean square error of the validation samples. The second algorithm requires more time but can result in good generalization for difficult, small or noisy datasets. Training stops according to adaptive weight minimization. The network was trained several times using different training algorithms (Levenberg – Marquardt and Bayesian Regularization) and number of nodes. The best outcome came from a network with 70 nodes, trained with a Levenberg – Marquardt algorithm which result in an R^2 of 0.896 and a mean square error (MSE) of 0.0834. R^2 measures the correlation between outputs and target values. The closer its value to 1, the closer their relationship. The MSE is the average of the summation of the squared difference between the actual output value and the predicted output value. The split of the data followed an 80/10/10 mode for training, validation and testing. The results of the network fitting are shown in Figure 7.17.

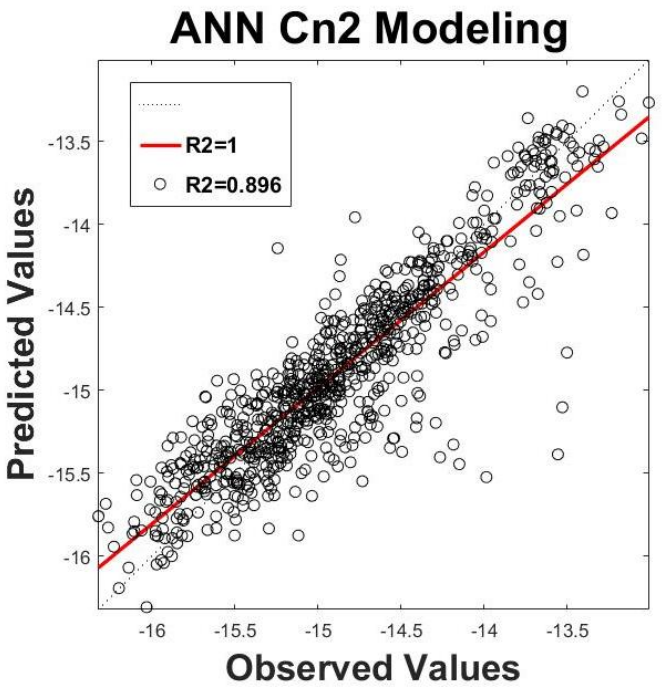


Figure 7.17. Regression plot for the Single Hidden Layer Network.

To develop the Random Forest model, we used the sklearn module and specifically the RandomForestRegressor function. There are several different parameters we can select for an RF model. In our case, after executing a grid search and cross validation, we found the optimal set of the following parameters, the number of decision trees that will be running in the model (n_estimators = 80), the criterion (loss function) used to determine the model outcome (criterion = MSE), the maximum possible depth of each tree (default value allows for leaves expansion until they are all pure) and the maximum number of features under consideration in each split (equal to the number of estimators). The results showed a very good agreement between model predictions and observed values, i.e. an R^2 of 0.865 and a RMSE of 0.241 and are plotted in Figure 7.18.

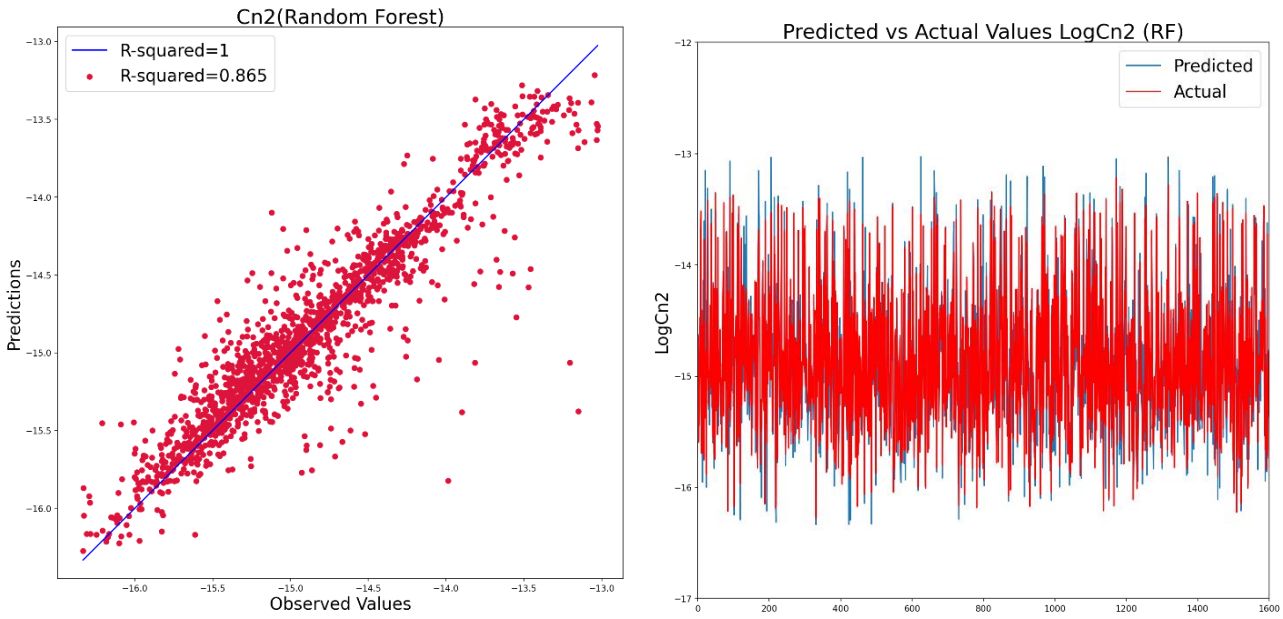


Figure 7.18. Scatter plot (left) and line plot (right) for observed vs predicted C_n^2 values for the Random Forest regression model.

Gradient boosting (GB) is one of the variants of ensemble methods where weak learners are created in series in order to produce a string ensemble model. GB makes use of the residual error for learning. The main training steps for a GB model are the following: i) an initial tree estimates the label value, ii) afterwards, the residual error is calculated, iii) then, another model is created to predict the error based on the previous model, not the label and iv) the label prediction is updated based on the error prediction. Again, the GB model includes several hyperparameters that can be initially selected and tuned adequately. The hyperparameters selected for this model are, i) the number of boosting stages to perform (n_estimators = 1000), ii) the learning rate of the model (learning_rate = 0.05), iii) the maximum depth of the individual regression estimators (max_depth = 6) and iv) the minimum number of samples required to split an internal node (min_samples_split = 12). The results again showed good agreement between model predictions and observed values, i.e. an R^2 of 0.851 and a RMSE of 0.252 and are plotted in Figure 7.19.

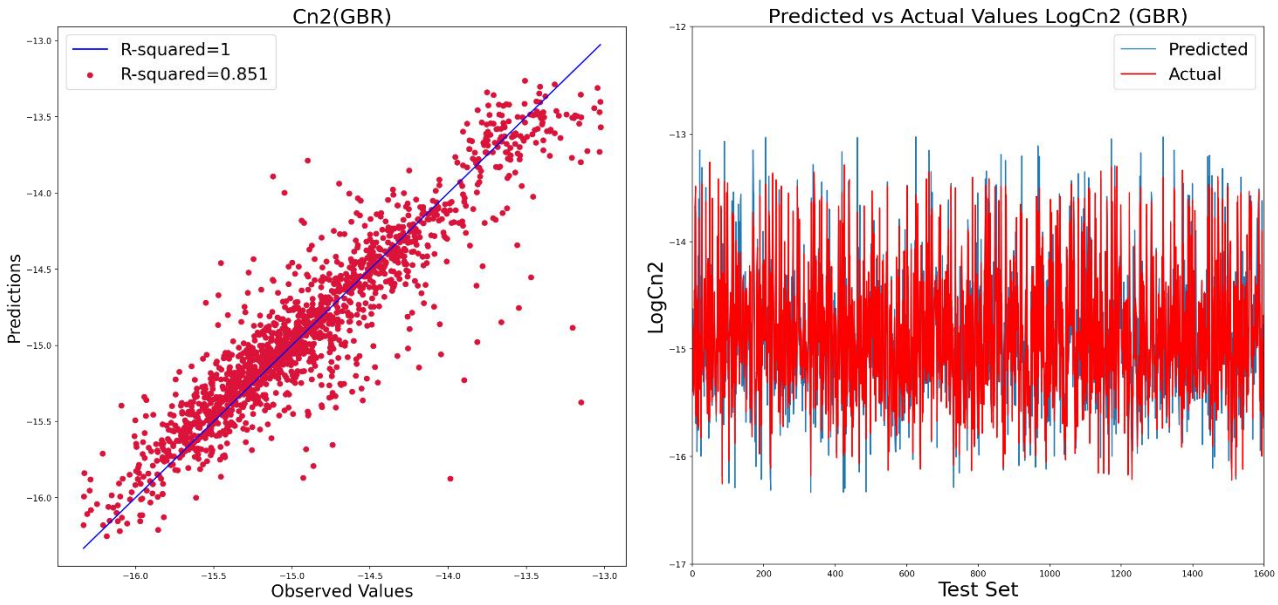


Figure 7.19. Scatter plot (left) and line plot (right) for observed vs predicted C_n^2 values for the Gradient Boosting regression model.

The last algorithm explored was a deep neural network, the evolution of a single layer network in order to overcome inherent limitations. Practically, a deep neural network is a single neural network with added hidden layers. The number of the hidden layers and the number of nodes in each layer control the neural network model capacity and depends on the specific problem we want to solve. As the dataset was not too big, we limited the number of layers in the deep learning model in order to save time and avoid overfitting. For this model, a three hidden layer architecture with a sequentially decreasing number of nodes in each layer (30/20/10), was selected and ran over batches of 16 for a total number of 350 epochs. A ReLu activation function was used to connect the hidden layer nodes and a linear for the output node because it was a regression model. The ReLu activation outputs the input directly if it is greater than 0; otherwise it returns zero. The loss function (Figure 7.20) was based on the mean squared error with an Adam optimizer. The line plot shows the expected behavior. Namely, that the model rapidly learns the problem, decreasing the loss function down to about 0.01 in

about 75 epochs and remains pretty stable thereafter. The line plot also shows that train and test performance remain comparable during training, whereas the training line is a bit bumpy. Figure 7.21 presents the scattering and line plots for the DNN algorithm. The results for this model were $R^2 = 0.79$ and an RMSE = 0.088.

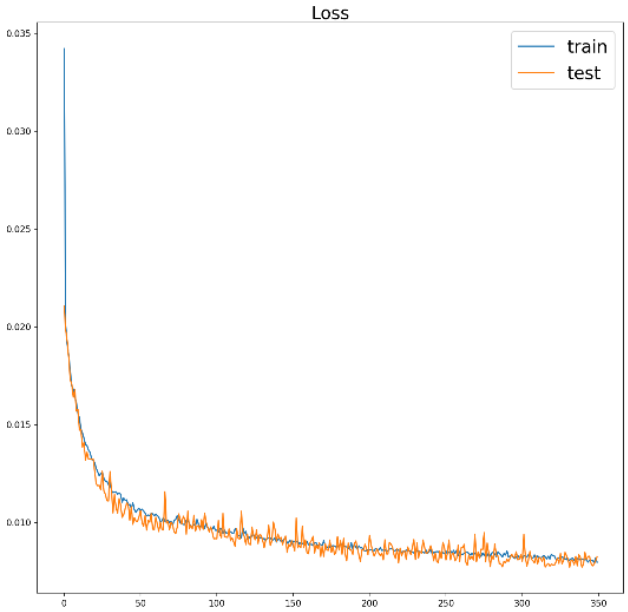


Figure 7.20. The line plot of the loss function during DNN model training.

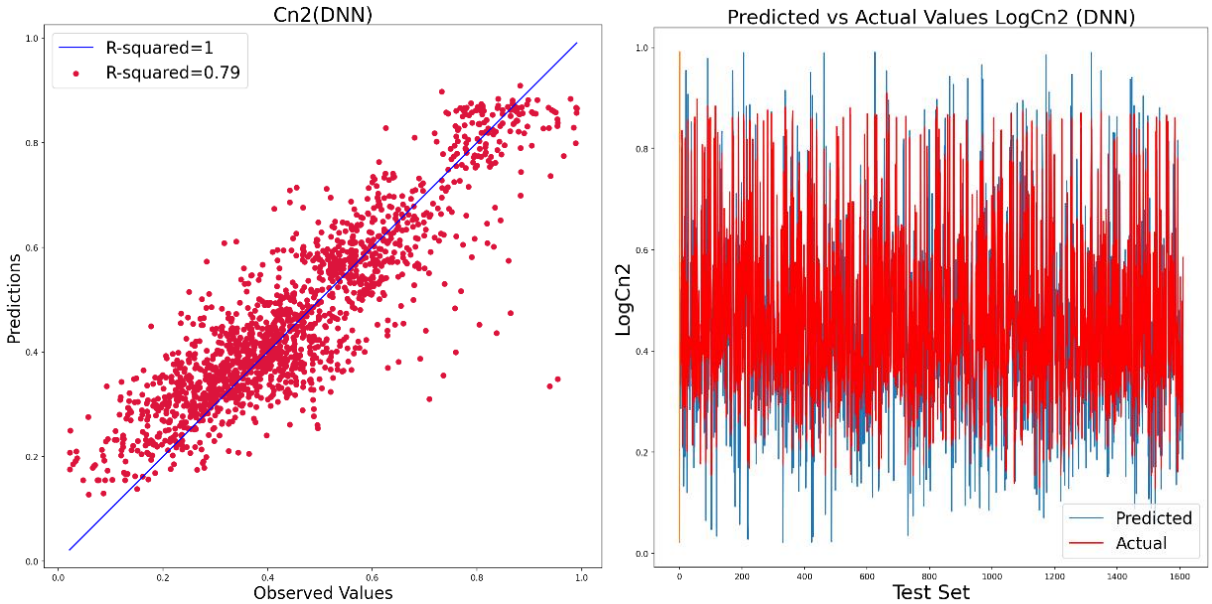


Figure 7.21. Scatter plot (left) and line plot (right) for observed vs predicted C_n^2 values for the DNN regression model.

7.5.3 Turbulence Classification Modeling

Many studies have been performed to create a mathematical model for the probability density function (pdf) of the received irradiance. The result of these studies, is the development of various statistical models for the scintillation induced by the atmospheric turbulence for a range of atmospheric conditions. The turbulence strength has been divided into two levels, the weak and the strong, defined by the value of the Rytov variance, σ_R^2. For σ_R^2 values less than unity, the statistics of irradiance can be adequately described by the lognormal model. In case of higher turbulence strength, the lognormal pdf is not so accurate and therefore not appropriate for strong turbulence level irradiance modeling. For σ_R^2 higher than unity, the statistics for received irradiance can be well described by the negative exponential or the gamma-gamma pdf. Apart from these two models, numerous others exist that can sufficiently describe the irradiance statistics in either turbulence level or some for both.

This section aims to describe a DNN approach to model the turbulence strength level, that is either strong or weak. That way, we can be able to use the applicable statistical model to describe the channel based on its current status. To do so, we utilized the environmental data set as described in Section 7.5.1 and built a deep neural network as shown in Fig. 7.22 to classify the turbulence strength level.

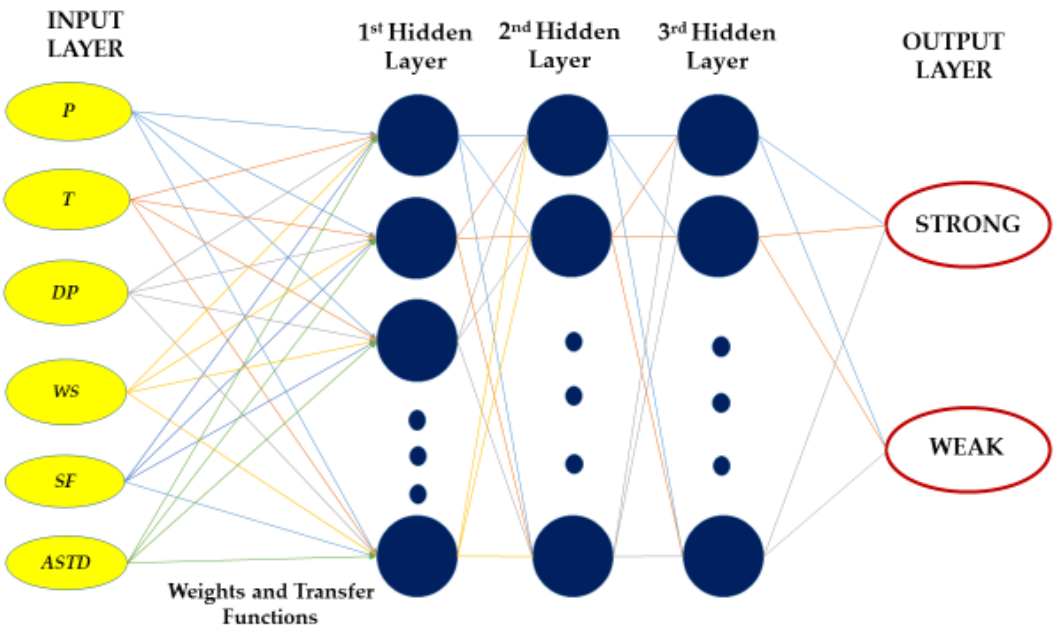


Figure 7.22. The three-layer deep neural network for turbulence strength classification.

In the raw dataset described in Section 7.5.1, we assumed a notional value of refractive index structure parameter, $C_n^2 > 5 \cdot 10^{-15}$, to characterize it as strong and for $C_n^2 < 5 \cdot 10^{-15}$ weak. Given Eq. (2.30), for $C_n^2 = 5 \cdot 10^{-15}$, λ = 850 nm and L = 3000 meters, $\sigma_R^2 \approx 1$. Therefore, a "0" was attached to every row in our dataset with $C_n^2 < 5 \cdot 10^{-15}$ and a "1" for $C_n^2 > 5 \cdot 10^{-15}$. The resulted split of our experimental data was quite balanced and showed that strong values slightly outnumbered the weak, as shown in Figure 7.23.

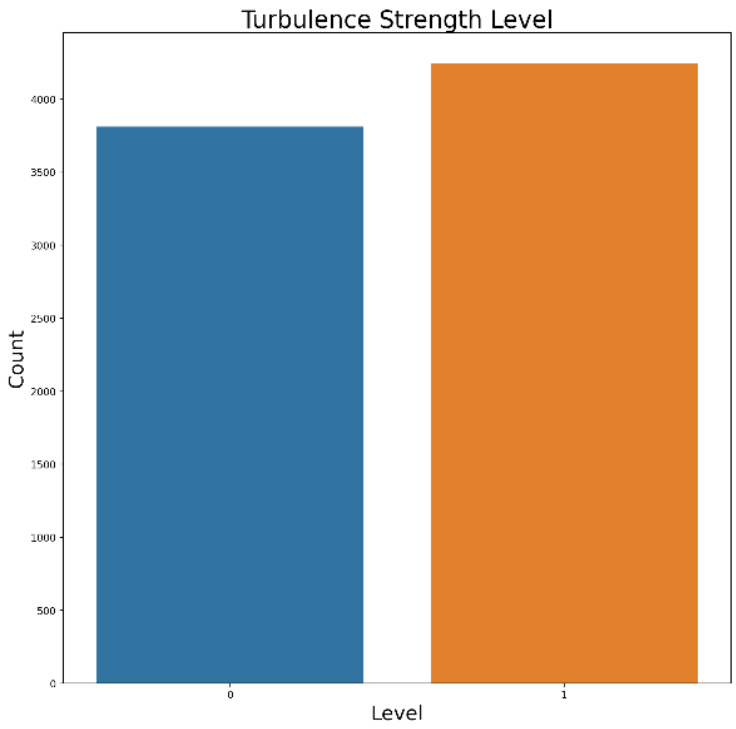


Figure 7.23. The cumulative results of the turbulence strength level. “0” denotes weak and “1” strong turbulence conditions.

The network has three hidden layers with 30, 20 and 10 neurons, respectively. A feed-forward back propagation algorithm was used, with a dropout rate of 0.5 per layer. The activation function for the three hidden layers was a rectifier (ReLU) whereas for the output layer a sigmoid function. In order for the algorithm to monitor the progress of the algorithm fitting, a binary cross entropy loss function was used and the Adam optimizer to adapt the gradient descent of the loss function. The algorithm was trained against the 80% of the dataset and tested over the remaining 20%. We used a total number of 500 training epochs for a batch size of 8. Figure 7.24 shows the progressive performance of the model throughout training, measured by the accuracy and loss function both for training and validation sets. The model exhibited a significant accuracy early on the epochs iteration. After approximately 200 epochs we observe a

slight divergence between the training and validating measurements, which remain pretty constant throughout the entire epochs number.

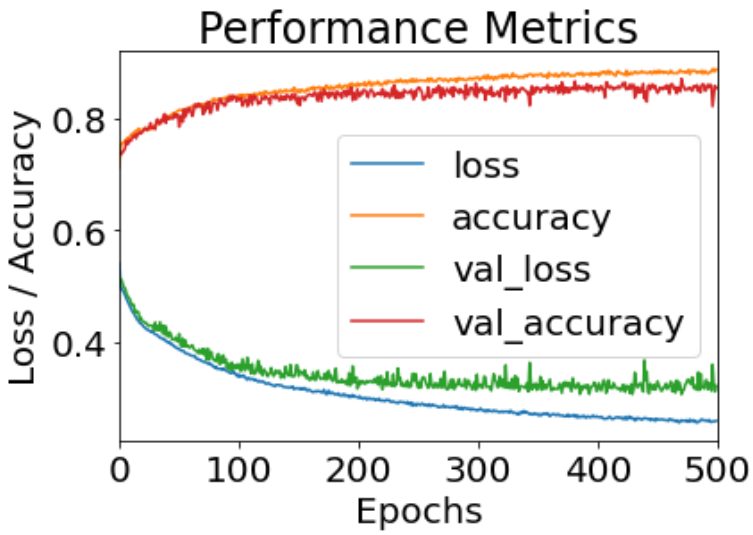


Figure 7.24. The loss/accuracy performance of the deep neural classifier of the turbulence strength for a total of 500 epochs.

Figure 7.25, presents the confusion matrix of the DNN classification model. By definition a confusion matrix C is such that $C_{i,j}$ is equal to the number of observations known to be in group i and predicted to be in group j. Thus, in a binary classification the count of true negatives is $C_{0,0}$, false negatives is $C_{1,0}$, true positives is $C_{1,1}$ and false positives is $C_{0,1}$. As shown in Figure 7.25, our model exhibited a very acceptable classification performance, given the high variability of our target value (C_n^2), since we observe that false negatives are only $C_{1,0}$ = 87 and false positives are $C_{0,1}$ = 132, which is translated in an accuracy value of 0.87. In other words, 87% of the model's prediction will be correct.

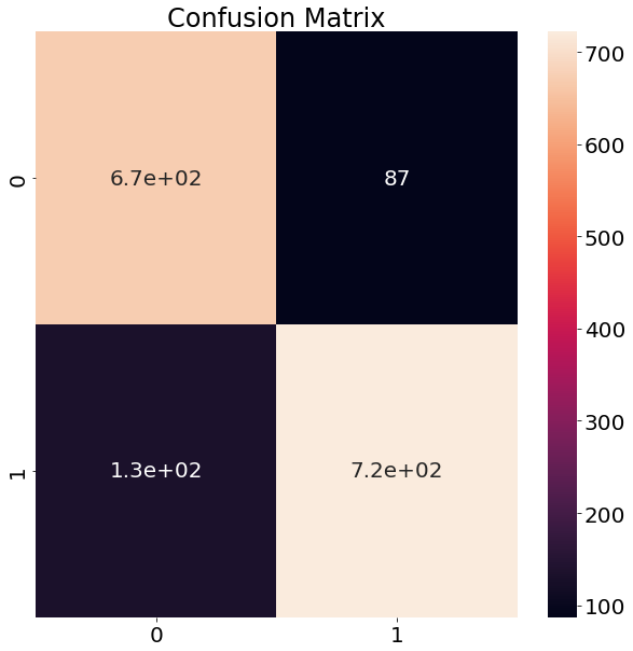


Figure 7.25. The confusion matrix for the DNN classifier.

7.6 Summary

This chapter is comprised of two main parts, which present a thorough analysis of C_n^2 and FSO outage probability modeling, by leveraging machine learning algorithms.

The first part of the chapter is two-fold. Initially the regression analysis results for C_n^2 is presented. We utilized six common ML algorithms and trained them on four different data sets. The results showed a great variance on the prediction accuracy of each model. Specifically, the ANN algorithm, a single layer perceptron model, that included 100 neurons in its hidden layer, with a training batch size of 32, showed a moderate to low accuracy, with R^2 = 0.55 and RMSE = 0.0916. The RF algorithm gave a significantly improved coefficient of determination for the model evaluation, R^2= 0.78, which was the best value among all algorithms for every season. Additionally, an RMSE = 0.064 showed a great improvement in the error of the predicted values. The

Gradient Boosting Regressor model, used a tree number -iterations- of 2000 and also performed very sufficiently. It was comparable to the RF by achieving a value of R^2 = 0.7 and an RMSE = 0.075. The KNN algorithm achieved its best performance for a value of k = 15, which resulted in R^2 = 0.71 and an RMSE = 0.073. That is, it slightly outperformed the GBR algorithm. The DT algorithm, with an optimum depth of 15 trees resulted in R^2 = 0.637 and RMSE = 0.083. Finally, the DNN algorithm, comprised of three hidden layers of neurons (1^{st} hidden layer = 50, 2^{nd} hidden layer = 30 and 3^{rd} hidden layer = 10), resulted in an R^2 = 0.61 and an RMSE = 0.085. The second part of the paper was devoted on a thorough analysis for the outage probability of an FSO link. Initially, the corresponding P_{out} for the measured meteorological conditions was derived, based on an existing in the literature mathematical formula. These P_{out} estimations were used to classify the link status as functional or non-functional depending on a required availability of 99%, which corresponds to a 1% outage probability. We then trained a DNN classifier to model the status of the link based on the six measured meteorological parameters. Finally, an empirical mathematical model for outage probability estimation was developed based upon those meteorological parameters and refractive index. Both the DNN classifier and the regression formula showed a very good performance. On the author's knowledge, an experimentally obtained outage probability analysis for an FSO link does not exist in the literature, therefore the presented results of this paper is of significant importance for the FSO technology community.

The second part of this chapter, utilized four common ML algorithms and trained them on a preliminary data set consisting of six experimentally obtained macroscopic meteorological parameters. The results showed very good prediction accuracy for every model. Specifically, the ANN algorithm, a single layer perceptron model that included

70 neurons in its hidden layer, with a training batch size of 32, trained with a Levenberg – Marquardt algorithm, resulted in an R^2 of 0.896 and a mean square error (MSE) of 0.0834. The RF algorithm, comprised from 80 estimators, also gave a highly acceptable coefficient of determination, an R^2 of 0.865 and a RMSE of 0.241. The Gradient Boosting Regressor model, with 1000 boosting stages (n_estimators), a learning rate of 0.05, maximum depth of the individual regression estimators equal to six and minimum number of samples required to split an internal node equal to twelve, resulted in an R^2 of 0.851 and a RMSE of 0.252. Finally, the DNN algorithm, comprised of three hidden layers of neurons (1st hidden layer = 30, 2nd hidden layer = 20 and 3rd hidden layer = 10), ran over batches of 16 for a total number of 350 epochs, resulted in an $R^2 = 0.79$ and a RMSE = 0.088. Additionally, we developed a DNN approach to classify the turbulence strength level as either strong or weak utilizing the same data set. A notional value of refractive index structure parameter, $C_n^2 = 5 \cdot 10^{-15}\, m^{-2/3}$ was set to distinguish between strong and weak region and the resulted split of the experimental data was quite balanced. The network has three hidden layers with 30, 20 and 10 neurons, respectively and a dropout rate of 0.5 per layer. The algorithm was trained against the 80% of the dataset and tested over the remaining 20%, for a total number of 500 training epochs and a batch size of 8. The model exhibited a very acceptable classification performance, given the highly variability of our target value (C_n^2), since we observed an accuracy of 87% on the model's prediction.

CHAPTER 8

RSSI probability density functions comparison using Jensen-Shannon divergence and Pearson distribution

This chapter includes the results that have been published in the peer-reviewed journal [J6].

8.1 Introduction

Traditionally, a large number of probability density functions (PDFs) have been utilized to model the received optical power of an FSO link. The most widely used and accepted PDF models include the Log-Normal (LN), Gamma-Gamma (GG), Gamma, Weibull, K, I-K, Malaga, the Fisher-Snedecor F and the negative exponential. In order to evaluate the performance of an optical communication system, various link parameters can be employed, such as the probability of detection, the outage probability, the average capacity and the outage capacity. In any case, knowledge of the PDF of the received optical power is required; however, it is rather difficult to determine which PDF better fits the statistics of the received power data.

In the open technical literature, several PDF models are available for weak, moderate and strong turbulence. For example, the Log-Normal and the Gamma-Gamma distributions describe accurately the weak and moderate regime whereas the gamma-gamma, the negative exponential and the K-distribution yield a better fit at the strong regime [4]. Apart from selecting a distribution corresponding to the correct turbulence

strength regime, it is risky to assume that only a certain one fits accurately experimental data. Instead, it is critical to investigate the best fit among different distributions.

Motivated by the above, the main contribution of this chapter is to propose a methodology of comparing different theoretical distribution models, for any turbulence strength regime. To achieve this, several theoretical PDF models are fitted to the dataset's histogram using a sizable empirically measured dataset of the received signal intensity of an FSO connection in the marine environment. Initial comparisons of five theoretical models are made using the Kullback-Leibler (KL) and Jensen-Shannon (JS) divergences, and the best match is decided by the absolute value of the associated divergence. The type that best matches experimental data is also determined using the Pearson family of probability distribution functions, which consists of seven different continuous probability distributions.

8.2 Experimental Setup

The commercial FSO communications system located on the Hellenic Naval Academy's building roof and on Psitalia's island lighthouse, which has been already describe in the previous chapters, was utilized for the data acquisition of the received signal strength indicator (RSSI) of the optical link. The system has been set to obtain a measurement of the received signal strength indicator (RSSI) every minute, during each day, and then store these values in a connected PC for further analysis.

The data collection period of the RSSI parameter spanned over a year, from the 30th November of 2019 to the 27th October 2020, to include totally 144802 data points. At the same time, several macroscopic meteorological parameters were also measured from a standalone weather station located close to the FSO system so that their fluctuations can be monitored and estimate their influence to RSSI. This dataset of the

RSSI measurements is used to apply the PDF fits comparison. The descriptive statistics of the dataset is shown in Table 8.1.

Table 8.1. Descriptive statistics for the RSSI measurements from 30th November 2019 to 27th October 2020.

Statistic	Value
Mean	420.385927
Standard Error	0.084275342
Median	425
Mode	445
Standard Deviation	32.06917633
Kurtosis	1.481024233
Skewness	-0.798160493
Maximum	187
Minimum	517

8.3 Results and Analysis

In this section, we examine the fitting comparison techniques of different PDFs to an experimentally gathered dataset. This process will benefit the performance analysis of an FSO link and allow for credible reliability and availability estimations.

8.3.1 Kullback-Leibler Divergence

A fundamental concept of the information theory is the Kullback-Leibler (KL) divergence, a method to measure the conventionally defined “distance” among two distributions [186]. The entropy of a random variable is the information required to describe it [187]. We assume two distinct distributions defined on the same probability space, the first, *p*, which refers to a theoretical probability distribution and is measured

against an experimental or modeled distribution, q. The KL divergence, also called the relative entropy, measures the inefficiency of assuming a distribution to be q when in reality is p. The relative entropy or KL divergence between two distributions is then given [187],

$$D_{KL}(p||q) = \sum_i p_i log_2 \frac{p_i}{q_i} \tag{8.1}$$

The KL divergence is non-negative and asymmetric in p and q, meaning that,

$$D_{KL}(p||q) \neq D_{KL}(q||p) \tag{8.2}$$

The lower the value of the D_{KL}, the less the distance between p and q, which ultimately can become zero if the two distributions are identical. Additionally, in case an i exists where $p_i > 0$ and $q = 0$, then the KL divergence goes to infinity. Cover and Thomas, give a technical interpretation of the KL divergence, as the "coding penalty" associated with a distribution q selected to compare with a distribution p [187]. A relevant important measure is the mutual information, which gives the amount of a random variable explained from another. Let two random variables X, and Y, then their mutual information is given by the relative entropy of their joint distribution $p(x, y)$ and the product distribution $p(x)p(y)$ [187],

$$I(X,Y) = D(p(x,y)||p(x)p(y)) \tag{8.3}$$

Five theoretical distribution models where utilized to apply the Kullback-Leibler divergence Eq. (8.1) in order to evaluate and compare their fit performance to the real-valued RSSI data. The theoretical PDFs used where:

$$\text{Weibull: } f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \tag{8.4}$$

$$\text{Gamma: } f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, for\ x > 0\ and\ \alpha, \beta > 0 \tag{8.5}$$

$$\text{Log} - \text{Normal (LN)}: f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left(-\frac{\left(\ln(x)+\sigma^2/2\right)^2}{2\sigma^2}\right) \quad (8.6)$$

$$\text{Burr}: f(x) = ck\frac{x^{c-1}}{(1+x^c)^{k+1}} \quad (8.7)$$

$$\text{Extreme} - \text{Value (EV)}: f(x) = \frac{1}{\beta}e^{-(x+e^{-x})} \quad (8.8)$$

The Kullback-Leibler divergence was calculated for these five PDFs and the best fit to the observed RSSI data (30th Nov 19 – 27th Oct 20) was deduced. Initially the frequency occurrence of each RSSI observation is calculated. The probability density of the RSSI data was then developed by dividing the frequency of each observation by the total number of observations. We then utilized the *distfit* MATLAB application to obtain the parameters of the theoretical distributions within the observed RSSI values range (187-517). By applying Eq. (8.1) for each model, the KL divergence was calculated (Figure 8.1). The Burr distribution is proved to best fit our data and thus exhibited the lowest KL divergence value, 2.77×10^{-2}, followed by Weibull distribution with a value of 3.07×10^{-2}.

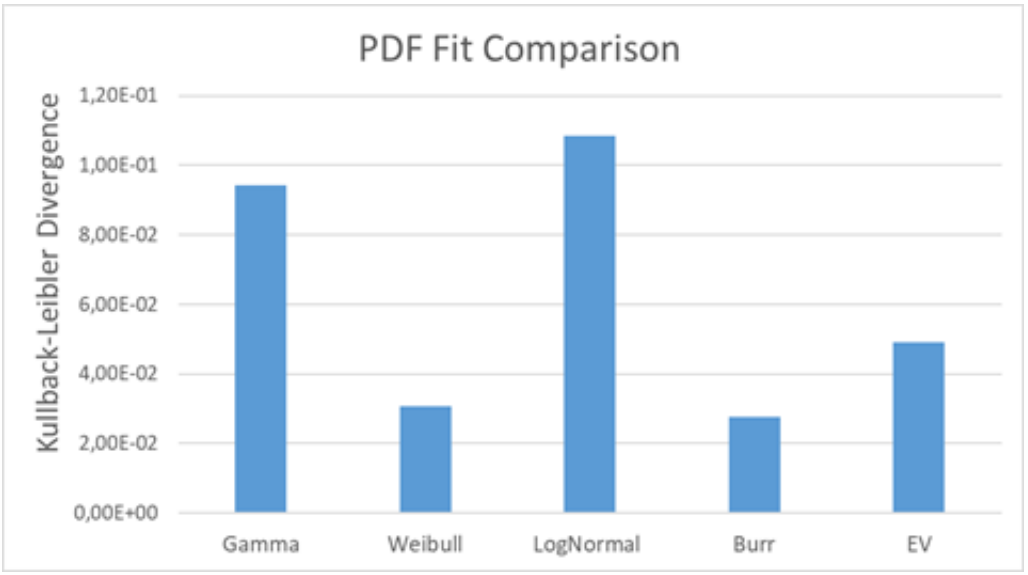
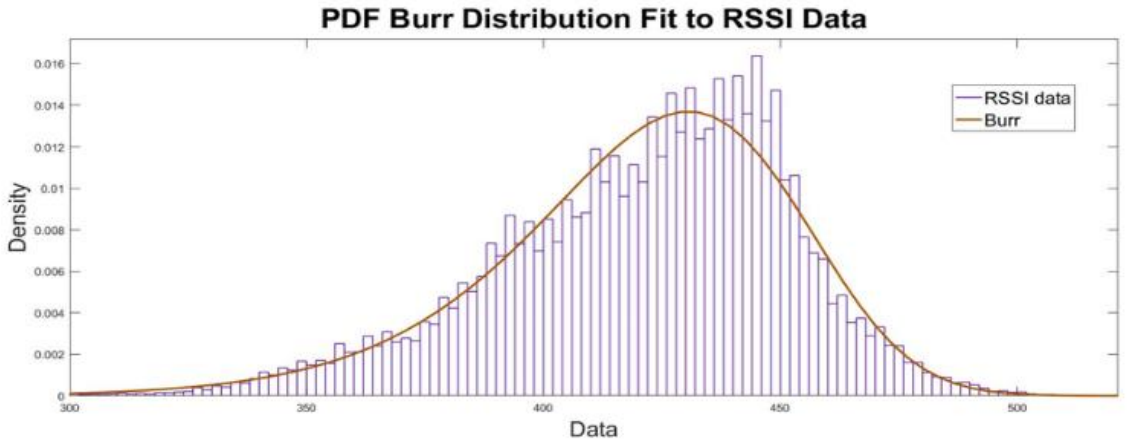
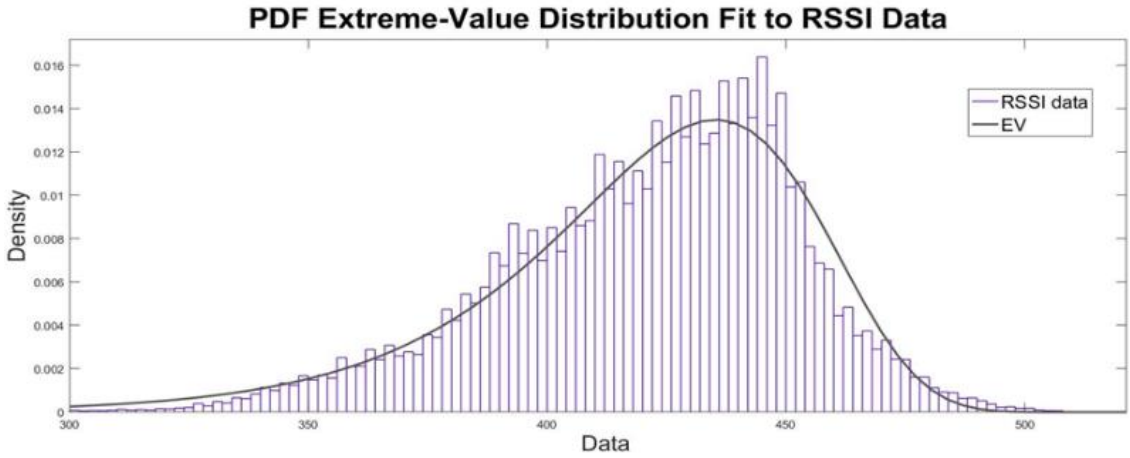
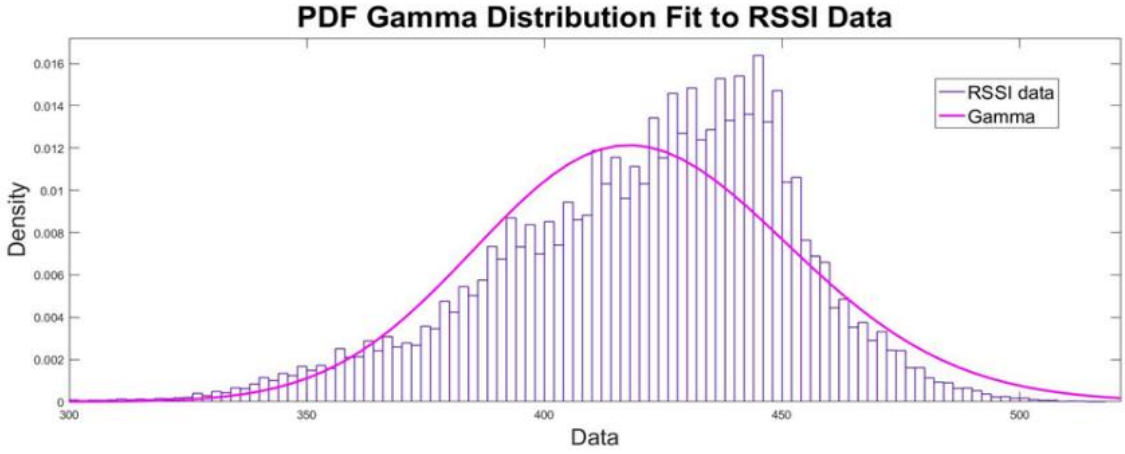


Figure 8.1. The Kullback-Leibler divergence values for Gamma, Weibull, Lognormal, Burr and Extreme-Value distribution.

The probability density functions for each theoretical model against the observed RSSI data are depicted in Figure 8.2, and their corresponding probability plots in Figures 8.3-8.5.







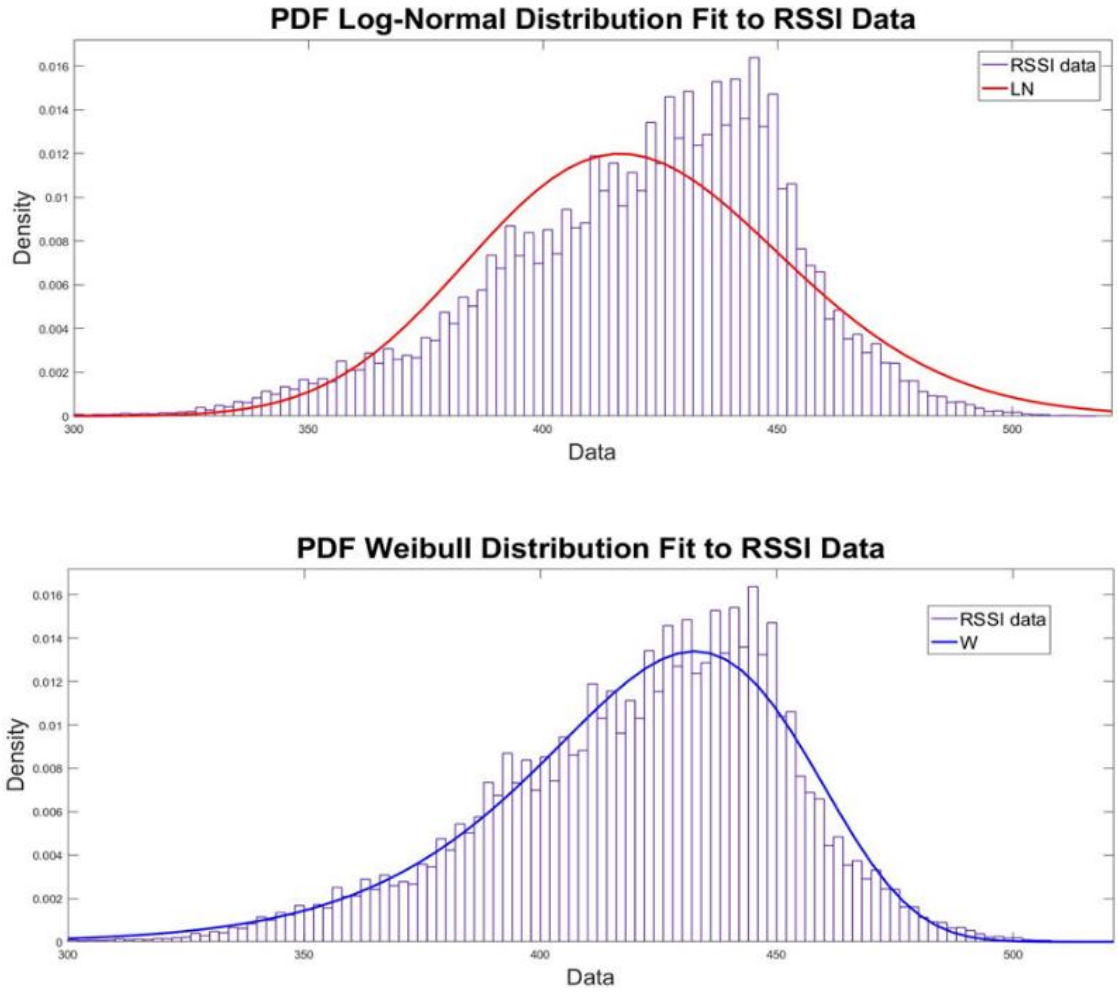


Figure 8.2. Probability density functions fits to RSSI data.

It is obvious graphically from the probability plots that the Burr distribution best fits the RSSI data. In particular, it presents an excellent agreement with the experimental data with an exception in the far upper tail of the distribution. In contrast, the rest four of the tested distributions exhibit a noteworthy disagreement in both tails.

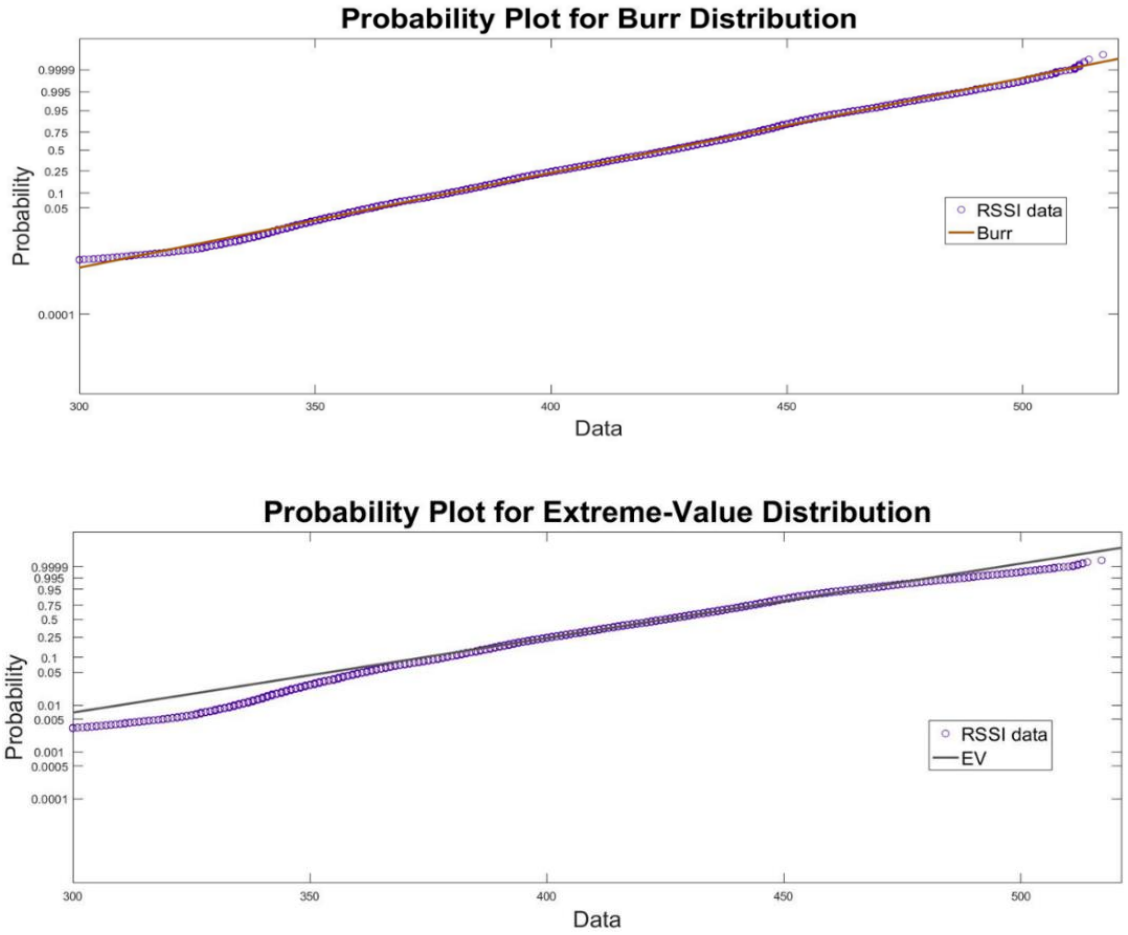
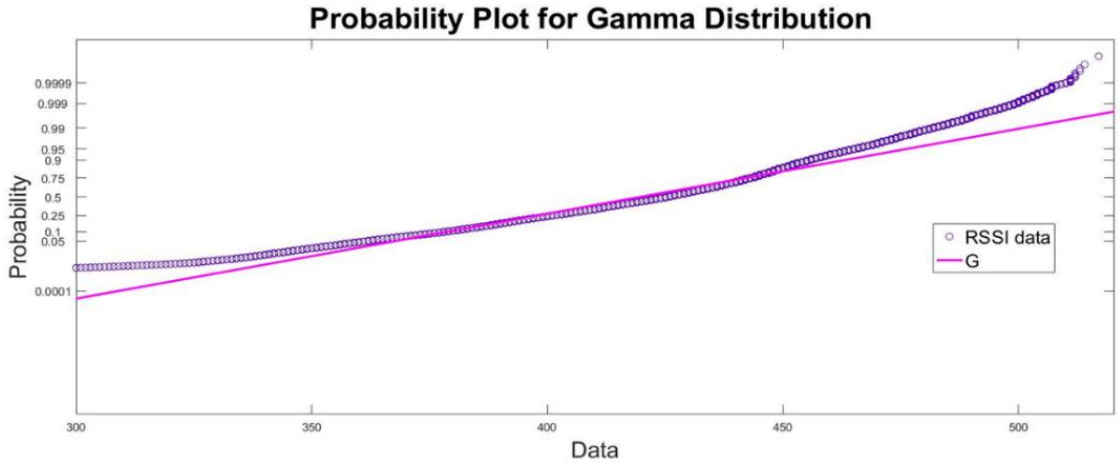


Figure 8.3. Probability plots for Burr and Extreme-Value distribution against RSSI data.



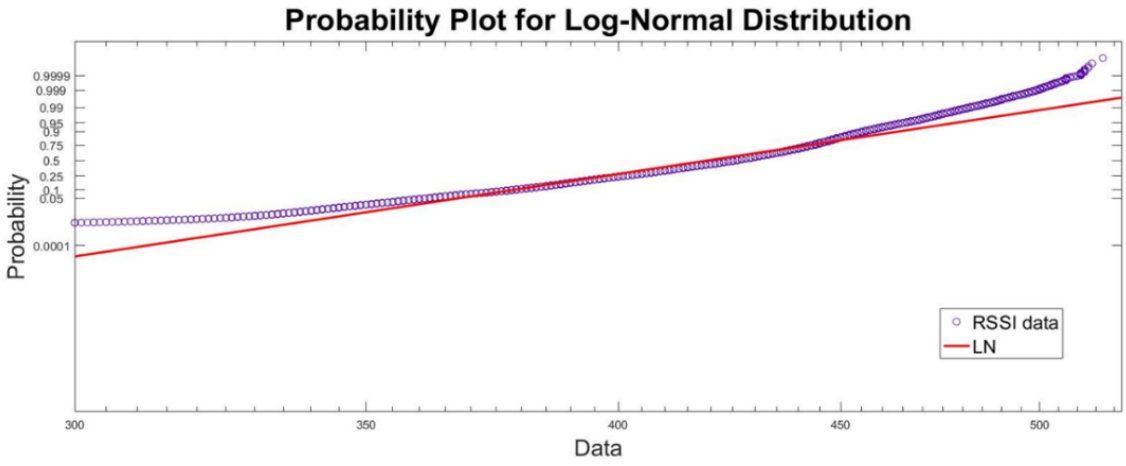


Figure 8.4. Probability plots for Gamma and Log-Normal distribution against RSSI data.

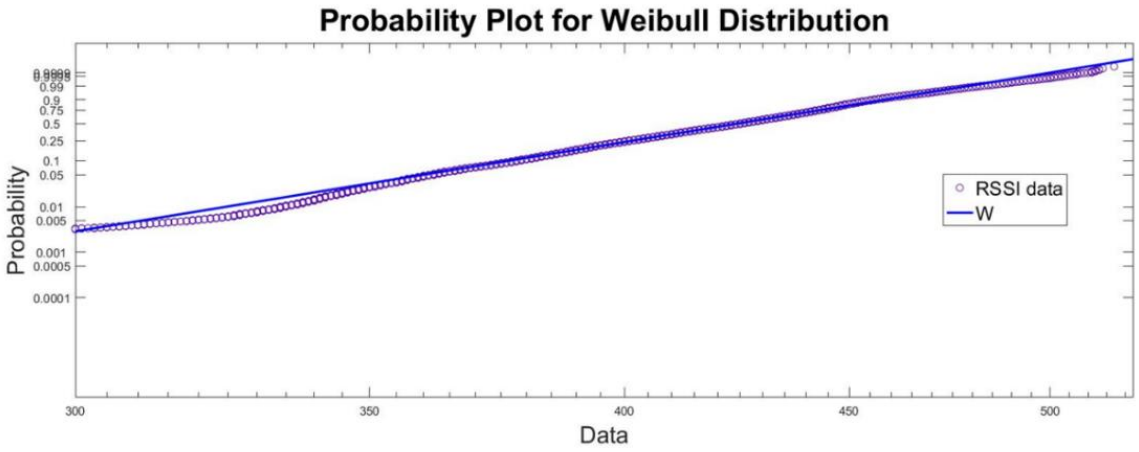


Figure 8.5. Probability plots for Weibull distribution against RSSI data.

8.3.2 Jensen-Shannon Divergence

Another method of measuring the similarity of two probability distributions is the Jensen-Shannon divergence (JSD). The JSD is based on the KL divergence, however it exhibits a few but important differences with it, that it is symmetric and it always has a finite value. The square root of JSD is also defined as JS distance [188] and it is given by,

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M) \quad (8.9)$$

where, *M=(P+Q)/2*, and *D(P||M)* and *D(Q||M)* are the KL divergences as defined by Eq. (8.1). Provided that at least one of the two measured distributions uses base 2 logarithm, it is upper bounded by 1,

$$0 \leq JSD(P||Q) \leq 1 \quad (8.10)$$

The JSD, also called *capacitory discrimination*, can be understood as the total KL divergence to the average distribution, *M* and can be also mathematically defined as a Jensen divergence of the Shannon information h [22],

$$JS(p;q) = h\left(\frac{p+q}{2}\right) - \frac{h(p)+h(q)}{2} \quad (8.11)$$

where the notation ';' is used instead, to emphasize its symmetry. Other symmetrizations of the KL divergence can be found in the literature, including among other the extrinsic Jensen-Shannon divergence, the Jeffreys divergence and the extended Kullback-Leibler divergence [189]. In terms of the JSD, we can also define it as the mutual information between a random variable X and the binary indicator variable Z, used to switch between *P* and *Q*, whose mixture distribution is associated with *X*.

In order to calculate the JSD of each theoretical distribution against the observed RSSI data we utilized Eq. (8.9). The *M* parameter was deduced by taking the average between each model and the RSSI data, a total of five. Following that, we calculated the KL divergence between each distribution, including the RSSI data distribution, and the corresponding *M* parameter, a total of ten distinct KL values. Finally, the JSD is calculated and the results are shown in Figure 8.6.

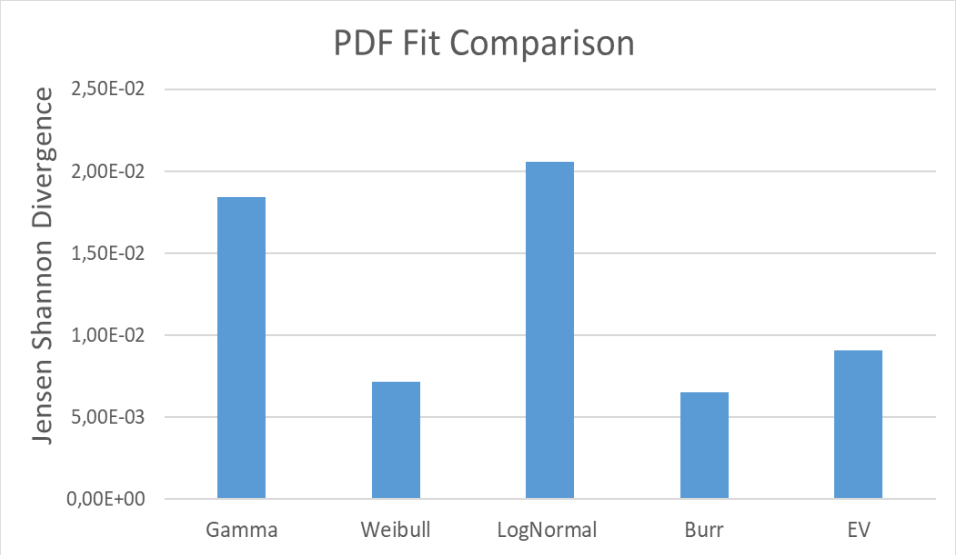


Figure 8.6. The Jensen-Shannon divergence values for Gamma, Weibull, Lognormal, Burr and Extreme-Value distribution.

8.3.3 Pearson Distribution Family

The Pearson family of continuous probability distributions, has become popular within the research community due to its ability to take into account the skewness and kurtosis of a probability distribution. This fact is really valuable when theoretical probability distributions need to be fit to experimental data that exhibits skewness, that is the asymmetry of a real-valued data around its mean. Skewness can take negative, positive or zero values, where negative indicates the distribution tail's on the left, whereas positive on the right. Pearson, originally identified four types of distributions, characterized by two quantities, β_1 and β_2. Any valid solution of Eq. (8.12) defines a Pearson type PDF [196].

$$\frac{df(x)}{dx} + \frac{a+(x-\lambda)}{d(x-\lambda)^2+c(x-\lambda)+b} = 0 \qquad (8.12)$$

where,

$$\begin{cases} b = \frac{4\beta_2 - 3\beta_1}{10\beta_2 - 12\beta_1 - 18}\mu_2 \\ a = c = \sqrt{\mu_2}\sqrt{\beta_1}\frac{\beta_2 + 3}{10\beta_2 - 12\beta_1 - 18} \\ d = \frac{2\beta_2 - 3\beta_1 - 6}{10\beta_2 - 12\beta_1 - 18} \end{cases} \quad (8.13)$$

where λ the location parameter of the distribution.

The solution for each type of Pearson distribution differs to the values of parameters *a, b, c* and *d*. Many types included to Pearson family distributions, are commonly used today in various applications, such as the beta, gamma and t-distribution [191].

The value of the discriminant of the quadratic function,

$$f(x) = dx^2 + ax + b \quad (8.14)$$

distinguishes the two main cases which group the distribution types.

We applied a standard *pearsrnd* function in MATLAB in order to determine the type of Pearson distributions that best fits the observed RSSI data. The mean value, standard deviation, skewness and kurtosis of the RSSI data were required to calculate the *d, a*, and *b* parameters of Eq. (8.14), where *d* = 0,0548, *a* = -0,3116 and *b* = 0,08355. The results showed that a Type IV Pearson PDF would best fits the data given by [191],

$$f(x) = \frac{A}{[A_0 + d(x + A_1)^2]^{1/(2d)}} exp\left[-\frac{b - A_1}{\sqrt{dA_0}} arctan\left(\frac{x + A_1}{\sqrt{A_0/d}}\right)\right] \quad (8.15)$$

where $A_0 = b - c^2(4d)^{-1}$ and $A_1 = c(2d)^{-1}$. The Pearson Type IV cumulative distribution function (CDF) against the RSSI data is plotted in Figure 8.7 and shows a very good agreement.

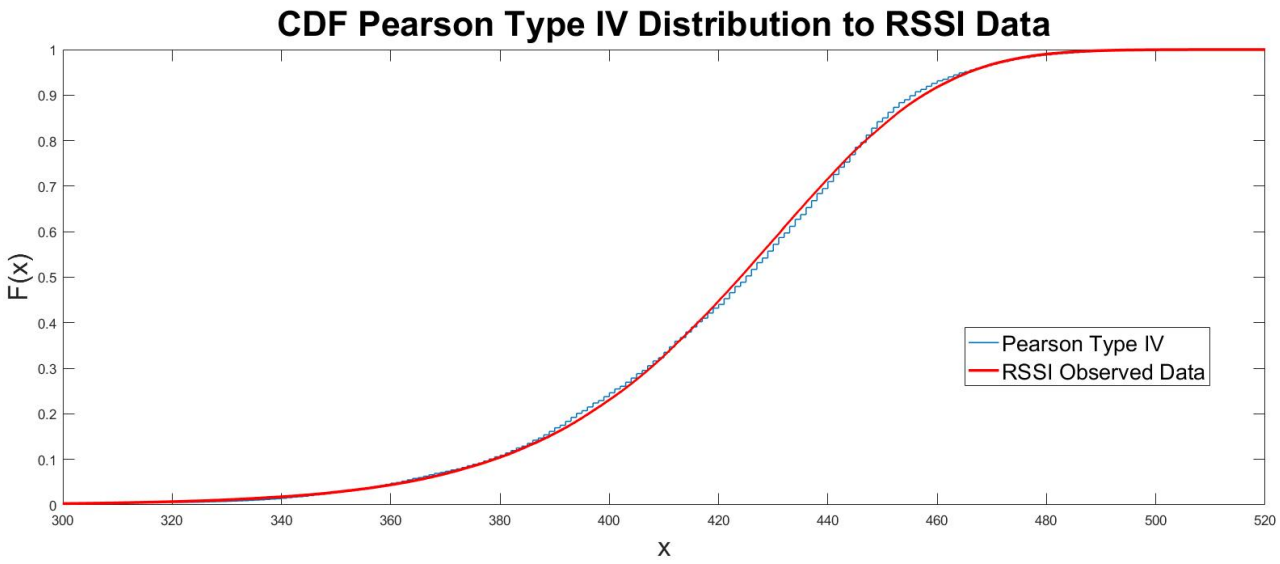


Figure 8.7. The cumulative distribution function for Pearson Type IV against RSSI data.

8.4 Summary

The modeling accuracy degree of five probability density functions to a large experimental dataset of the received signal strength for an FSO link has been studied. The focus of this paper is to identify among those PDFs the one that best fits to the observed data based on the available comparison methods. To that end, the Gamma, Lognormal, Extreme-Value, Burr and Weibull distributions have been utilized to calculate initially their Kullback-Leibler divergence and a symmetrization of it, the Jensen-Shannon divergence, with the RSSI data. The Burr distribution was found to best fir the experimental data, with a KL and JS divergence values of 2.77×10^{-2} and 6.53×10^{-3}, respectively. The Weibull and EV distributions also exhibited a comparable accuracy in contrast with the Gamma and Lognormal, whose "distance" from real data was of an order of magnitude higher. The Pearson distribution family of continuous probability functions was also used and the results showed that a type IV function yield the best fit. Both methods offer a straightforward process of comparing among different PDFs for fitting accuracy on real data. The two aforementioned methods follow a different approach to select the most appropriate PDF to fit to a dataset. KL and JD, allow for the

calculation of the performance metric of each candidate PDF, thus designate the best one as the one with the lowest divergence. On the other hand, the Pearson approach, compares among predefined PDFs and the algorithm determines which is the best. From a practical point of view, the KL and consequently the JD metrics are more useful to be applied in a real-world dataset, since they directly provide a certain output that allows any PDF to be examined, whereas the Pearson method selects within a more limited number of options.

CHAPTER 9

Conclusions

Over the last few decades, a new application of a rather older technology arose, the so-called free space optical (FSO) communication or else, the laser communication (LaserComm). The capacity requirements for the information channels in the 21st century calls for an effective alternative of the traditional radio frequency (RF) technology and the FSO technology aims to fill this gap. Perhaps, the most evident asset of FSOC is their potential for increased data rate (several 10s of Gb), as compared to their RF counterpart, due their inherently very high carrier frequency (on THz level) and therefore bandwidth. A significant drawback of the FSO systems that could potentially lead to a total blackout of the link, is the atmospheric effects. Since the atmosphere consists of many different particles like aerosols, dust, smoke etc. or large precipitation due to rain, haze or snow, the laser beam propagating through this may face significant power loss because of the attenuation or scattering effects. Special attention has been given in the phenomenon of optical turbulence. Two are the main factors in order to characterize the performance of an FSO link, that is the channel modeling and system performance metrics.

The fields of Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL), have been evolved and proved tremendously successful over the last few decades. Machine Learning use for Optical Communications research is a recently emerged topic that will definitely continue to increase rapidly. The value of this approach is derived mainly from the unique characteristics of the FSO research field

rather than the applied ML algorithms. By reviewing the literature on this topic, we can notice that a number of ML algorithms or a number of specific FSO-related applications have not been extensively leveraged.

This thesis presented a review of both traditional and machine learning methods, and applied them in novel contexts to model and predict the performance of optical communication systems and optical turbulence over maritime environment, thus increasing their availability. In this section, the results of this work are summarized and an outlook on potential future research directions is presented.

The reviewed traditional and machine learning methods were applied in FSO performance and optical turbulence modeling as follows:

Statistical Modeling of FSO Performance in Maritime Environment

In this chapter, two new mathematical models were proposed, to predict the received signal strength (RSSI) of an FSO optical link over maritime environment. The models have the form of a second-order polynomial with seven macroscopic meteorological parameters as the independent variables. The first proposed model (base model) was validated against real data in two separate periods and the R^2 and correlation coefficient between the observed and modeled RSSI values were computed to check how good the fit was. Both periods exhibited high R^2 and correlation coefficient, namely 69% and 0.8327, respectively. The second proposed model was an improved edition of the base model by including the air-sea temperature difference, as an additional independent variable. Overall, the improved model achieved a significant

linear correlation coefficient of 0.78 with the observed values as compared to 0.73 of the base model for the first period and 0.81 as compared to 0.79 for the second period. The goodness of fit parameter R^2 between the improved model prediction and the observed values was 70.5% and 66% for the two periods, respectively.

Machine Learning Algorithms for Received Optical Power Prediction

In this chapter, a machine learning based scheme was introduced to estimate the RSSI parameter of an FSO link over a maritime environment based on macroscopic meteorological measurements. To test the proposed approach, a large experimentally derived data set was used. Five ML algorithms were trained in order to construct a robust model to accurately predict the link's performance in terms of received signal strength. The results showed a significant improvement, as compared to traditional regression modeling techniques, and their prediction accuracy performance measured by the coefficient of correlation, R^2 and the RMSE was extremely promising for even more complex predictive modeling. The superlative performance of the ML approaches comparing to the common-used regression method indicates that, first, ML is the appropriate modelling choice when overall prediction is the goal and the volume of data is high, second, allows the ability to tune hyper parameters per ML approach to enable optimal performance, and, finally, the trained models can be either improved by adding more data or be used for continuous streamflow RSSI predictions.

Machine Learning Algorithms for Optical Turbulence Prediction

This chapter introduces the application of machine learning algorithms in modeling the refractive index structure parameter (C_n^2) and estimate its value through regression analysis of macroscopic meteorological parameters, obtained from the NPS site. Secondly, it applies well known mathematical expressions to estimate the outage probability of a notional FSO link, based on the strength of the optical turbulence and model the link status (On-Off) based on macroscopic meteorological parameters, by utilizing a DNN classification algorithm.

The second part, presented the ML-based modeling approach of the refractive index structure parameter (C_n^2), based on macroscopic meteorological parameters and C_n^2 values, obtained from the HNA experimental site. To do so, four machine learning algorithms were employed, including a Random Forest (RF), a Gradient Boosting Regressor (GBR) and two Neural Networks (a single layer and a deep network).

RSSI probability density functions comparison using Jensen-Shannon divergence and Pearson distribution

This chapter studied the modeling accuracy degree of five probability density functions to a large experimental dataset of the received signal strength for an FSO link, in order to identify among those PDFs the one that best fits to the observed data based on the available comparison methods. To that end, the Gamma, Lognormal, Extreme-Value, Burr and Weibull distributions have been utilized to calculate initially their

Kullback-Leibler divergence and a symmetrization of it, the Jensen-Shannon divergence, with the RSSI data. The Pearson distribution family of continuous probability functions was also used and the results showed that a type IV function yield the best fit. The two aforementioned methods follow a different approach to select the most appropriate PDF to fit to a dataset. KL and JD, allow for the calculation of the performance metric of each candidate PDF, thus designate the best one as the one with the lowest divergence. On the other hand, the Pearson approach, compares among predefined PDFs and the algorithm determines which is the best. From a practical point of view, the KL and consequently the JD metrics are more useful to be applied in a real-world dataset, since they directly provide a certain output that allows any PDF to be examined, whereas the Pearson method selects within a more limited number of options.

References

1. D. M. et al., “Free Space Optical Technologies”, Trends in Telecommunications Technologies. InTech, Mar. 01, 2010. doi: 10.5772/8488.

2. In Keun Son, Shiwen Mao, A survey of free space optical networks, Digital Communications and Networks, Volume 3, Issue 2, 2017, Pages 67-77.

3. Henniger, Hennes and Otakar Wilfert. “An Introduction to Free-space Optical Communications.” (2010).

4. Kaushal, Hemani& Jain, Vk&Kar, Subrat. (2017). Free-Space Optical Channel Models. 10.1007/978-81-322-3691-7_2.

5. B. Epple, H. Henniger, Discussion on design aspects for free-space optical communication terminals, IEEE Commun. Mag. 45 (October (10)) (2007) 62- 69.

6. G. Baister, P. Gatenby, Pointing, acquisition and tracking for optical space communications, Electron. Commun. Eng. J. 6 (December (6)) (1994) 271–280.

7. Abadi, M.M., Cox, M.A., Alsaigh, R.E. *et al.* A space division multiplexed free-space-optical communication system that can auto-locate and fully self align with a remote transceiver. *Sci Rep* **9**, 19687 (2019). https://doi.org/10.1038/s41598-019-55670-1.

8. E.J. Lee, V.W. Chan, optical communication over the clear turbulent atmospheric channel using diversity, IEEE J. Sel. Areas Commun. 22 (November (9)) (2004) 1896–1906.

9. Brandon Born, Ilija R. Hristovski, Simon Geoffroy-Gagnon, and Jonathan F. Holzman, "All-optical retro-modulation for free-space optical communication," Opt. Express 26, 5031-5042 (2018).

10. C. Quintana et al., "High Speed Electro-Absorption Modulator for Long Range Retroreflective Free Space Optics," in IEEE Photonics Technology Letters, vol. 29, no. 9, pp. 707-710, 1 May1, 2017, doi: 10.1109/LPT.2017.2680842.

11. V.W. Chan, Free-space optical communications, IEEE/OSA J. Light. Technol. 24 (December (12)) (2006) 4750–4762.

12. B. S. Robinson, D. M. Boroson, D. A. Burianek, D. V. Murphy, "Overview of the lunar laser communications demonstration," Proc. SPIE 7923, Free-Space Laser Communication Technologies XXIII, 792302 (21 February 2011); https://doi.org/10.1117/12.878313.

13. S. Das et al., "Requirements and challenges for tactical free-space Lasercomm," MILCOM 2008 - 2008 IEEE Military Communications Conference, San Diego, CA, USA, 2008, pp. 1-10, doi: 10.1109/MILCOM.2008.4753049.

14. Mikołajczyk, Janusz; Bielecki, Zbigniew; Bugajski, Maciej; Piotrowski, Józef; Wojtas, Jacek; et al. Metrology and Measurement Systems; Warsaw Vol. 24, Iss. 4, (2017): 653-674. DOI:10.1515/mms-2017-0060

15. Young, D.W. & Hurt, H.H. & Sluz, J.E. & Juarez, J.C. (2015). Development and demonstration of laser communications systems. Johns Hopkins APL Technical Digest (Applied Physics Laboratory). 33. 122-138.

16. Stromqvist Vetelino, Frida, "Fade Statistics For A Lasercom System And The Joint Pdf Of A Gamma-gamma Distributed Irradiance And Its Time Derivative" (2006). Electronic Theses and Dissertations. 824.

17. Wayne, David, "The Pdf Of Irradiance For A Free-space Optical Communications Channel: A Physics Based Model" (2010). Electronic Theses and Dissertations. 4300.

18. J. C. Juarez, A. Dwivedi, A. R. Hammons, S. D. Jones, V. Weerackody and R. A. Nichols, "Free-Space Optical Communications for Next-generation Military Networks," in IEEE Communications Magazine, vol. 44, no. 11, pp. 46-51, November 2006, doi: 10.1109/MCOM.2006.248164.

19. Casey, Charles J., John Gibson, Gurminder Singh, Charles Prince and Peter Ateshian. “Suitability of free space optical communication in military environments.” (2015).

20. NASA, Deep Space Optical Communications (DSOC), 2016, [online]. Available: ⟨https://gameon.nasa.gov/projects/deep-space-optical-communications-dsoc/⟩.

21. Moore, Christopher & Burris, Harris & Stell, Mena & Wasiczko, Linda & Suite, Michele & Mahon, Rita & Rabinovich, William & Gilbreath, Gcharmaine & Scharpf, William. (2005). Atmospheric Turbulence Studies of a 16 km Maritime Path. Proceedings of SPIE - The International Society for Optical Engineering. 5793. 12. 10.1117/12.606019.

22. M. F. Stell; C.I. Moore; H.R. Burris; M.R. Suite; M.J. Vilchek; M.A. Davis; R. Mahon; E. Oh; W.S. Rabinovich; W.J. Scharpf; G. C. Gilbreath and A.E. Reed. Passive Optical Monitor for Atmospheric Turbulence and Windspeed. Proc. SPIE 5160, Free-Space Laser Communications and Active Laser Illumination III, Bellingham, WA, 2004; doi: 10.1117/12.508051.

23. Wasiczko, L.M.; Moore, C.I.; Burris, H.R.; Suite, M.; Stell, M.; Murphy, J.; Gilbreath, G.C.; Rabinovich, W.; Scharpf, W. Characterization of the Marine Atmosphere for Free-Space Optical Communication. In Proceedings of the SPIE 6215, Atmospheric Propagation III, Orlando (Kissimmee), FL, USA, 17 May 2006.

24. Rabinovich, W.S.; Moore, C.I.; Mahon, R.; Goetz, P.G.; Burris, H.R.; Ferraro, M.S.; Murphy, J.L.; Thomas, L.M.; Gilbreath, G.C.; Vilcheck, M.; Suite, M.R.; Free-

space optical communications research and demonstrations at the U.S. Naval Research Laboratory,Appl. Opt. 2015,Vol. 54, F189-F200.

25. C. I. Moore; H. R. Burris; W. S. Rabinovich; L. Wasiczko; M. R. Suite; L.A. Swingen; R. Mahon; M. F. Stell; G. C. Gilbreath; W. J. Scharpf; Overview of NRL's maritime laser communication test facility. Proc. SPIE 5892. Free-Space Laser Communications V, Bellingham, WA, 2005. doi: 10.1117/12.622252.

26. Stephen Doss-Hammel, Eun Oh, Jennifer C. Ricklin, Frank D. Eaton, G. Charmaine Gilbreath, Dimitri Tsintikidis, "A comparison of optical turbulence models," Proc. SPIE 5550, Free-Space Laser Communications IV, (20 October 2004); https://doi.org/10.1117/12.563746.

27. Rita Mahon, Christopher I. Moore, Harris R. Burris, William S. Rabinovich, Michele R. Suite, Linda M. Thomas, "Power spectra of a free space optical link in a maritime environment," Proc. SPIE 7464, Free-Space Laser Communications IX, 746407 (21 August 2009); doi: 10.1117/12.828860.

28. H. R. Burris; C. I. Moore; L. A. Swingen; M. J. Vilcheck; D. A. Tulchinsky; R. Mahond; L. M. Wasiczko; M. F. Stell; M. R. Suite; M. A. Davis; S. W. Moore; W. S. Rabinovich; J. L. Murphy; G. C. Gilbreath; J. Scharpf; Latest Results from the 32km Maritime Lasercom Link at the Naval Research Laboratory, Chesapeake Bay Lasercom Test Facility. Proc. SPIE 5793. Atmospheric Propagation II, Bellingham, WA, 2005; doi: 10.1117/12.606030.

29. Rabinovich, William & Mahon, Rita & Burris, Harris & Spie, Member & Gilbreath, Gcharmaine & Goetz, Peter & Moore, Christopher & Stell, Spie & Vilcheck, Michael & Witkowsky, Jennifer & Swingen, Lee & Suite, Michele & Oh, Eun & Koplow, Jeffrey. (2004). Free-space optical communication link at 1550 nm using

multiple-quantum-well modulating retroreflectors in a marine environment. Optical Engineering. 23. 10.1117/12.507737.

30. G. C. Gilbreath; W. S. Rabinovich; C. I. Moore; H. R. Burris; R. Mahon; K. J. Grant; P. G. Goetz; J. L. Murphy; M. R. Suite; M. F. Stell; M.L. Swingen; L. M. Wasiczko; S. R. Restaino; C. Wilcox; J. R. Andrews; W. J. Scharpf; Progress in Laser Propagation in a Maritime Environment at the Naval Research Laboratory. Proc. SPIE 5892. Free-Space Laser Communications V, Bellingham, WA, 2005; doi: 10.1117/12.633390.

31. De Jong, A.N.; Schwering, P.B.; Van Eijk, A.M.; Gunter, W.H.; Validation of Atmospheric propagation models in littoral waters. Optical Engineering 2013, Vol. 52(4), 046002.

32. de Jong, A. N., Schwering, P. B. W., Benoist, K. W., Gunter, W. H., Vrahimis, G., and October, F. J., “Long-term measurements of atmospheric point-spread functions over littoral waters as determined by atmospheric turbulence”, in Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII, 2012, vol. 8355. doi:10.1117/12.917718.

33. Grant, K. J.; Murry, K.A.; Clare, B.A.; Perejma, A.S.; Martinsen, W.S. Maritime Laser Communications Trial. General Document, Defense Science and Technology Organization, Edimburg South Australia, June 2012

34. H. Kaushal and G. Kaddoum, "Optical Communication in Space: Challenges and Mitigation Techniques," in IEEE Communications Surveys & Tutorials, vol. 19, no. 1, pp. 57-96, First quarter 2017, doi: 10.1109/COMST.2016.2603518.

35. Ni, Wei & Miyamoto, Yuichi & Wakamori, Kazuhiko & Kazaura, Kamugisha & Matsumoto, Mitsuji & Higashino, Takeshi & Tsukamoto, Katsutoshi & Komaki,

Shozo. (2009). Experimental Study of Atmospheric Turbulence Effects on RoFSO Communication Systems. Piers Online. 5. 65-70. 10.2529/PIERS080907042345.

36. Kazaura, Kamugisha & Suzuki, Toshiji & Wakamori, Kazuhiko & Matsumoto, Mitsuji & Higashino, Takeshi & Tsukamoto, Katsutoshi & Komaki, Shozo. (2009). Experimental Demonstration of a Radio on Free Space Optics System for Ubiquitous Wireless. Piers Online. 5. 235-240. 10.2529/PIERS080907022238.

37. Alheadary, W.G.; Park, K.-H.; Alfaraj, N.; Guo, Y.; Stegenburgs, E.; Ng, T.K.; Ooi, B.S.; Alouini, M.-S. Free-space optical channel characterization and experimental validation in a coastal environment. Opt. Express 2018, 26, 6614–6628.

38. Alheadary, W.G.; Park, K.-H.; Ooi, B.S.; Alouini, M.-S. Free-space optical channel characterization in a coastal environment. J. Commun. Inf. Networks 2017, 2, 100–106.

39. Oermann, R. J. Novel Methods for the Quantification of Atmospheric Turbulence Strength in the Atmospheric Surface Layer. PhD Thesis, School of Chemistry and Physics, University of Adelaide, Adelaide SA, Australia, 2014.

40. Sadot, D.; Kopeika, N. S. Forecasting optical turbulence strength on the basis of macroscale meteorology and aerosols: models and validation. Opt. Eng. 1992, 31 doi:10.1117/12.56059.

41. Dmytryszyn, M.; Crook, M.; Sands, T. Lasers for Satellite Uplinks and Downlinks. Sci 2021, 3, 4. https://doi.org/10.3390/sci3010004

42. Jing, M., Liying, T. & Siyuan, Y. Technologies and applications of free-space optical communication and space optical information network. J. Commun. Inf. Netw. 1, 61–71 (2016). https://doi.org/10.1007/BF03391546

43. Munemasa, Yasushi & Saito, Yoshihiko & Carrasco-Casado, Alberto & Trinh, Phuc & Takenaka, Hideki & Kubo-oka, Toshihiro & Shiratama, Koichi & Toyoshima,

Morio. (2019). Feasibility study of a scalable laser communication terminal in NICT for next-generation space networks. 211. 10.1117/12.2536131.

44. Kevin Günthner, Imran Khan, Dominique Elser, Birgit Stiller, Ömer Bayraktar, Christian R. Müller, Karen Saucke, Daniel Tröndle, Frank Heine, Stefan Seel, Peter Greulich, Herwig Zech, Björn Gütlich, Sabine Philipp-May, Christoph Marquardt, and Gerd Leuchs, "Quantum-limited measurements of optical signals from a geostationary satellite," Optica 4, 611-616 (2017).

45. Sharma, V., Banerjee, S. Analysis of atmospheric effects on satellite-based quantum communication: a comparative study. Quantum Inf Process 18, 67 (2019). https://doi.org/10.1007/s11128-019-2182-0

46. Takenaka, H., Carrasco-Casado, A., Fujiwara, M. et al. Satellite-to-ground quantum-limited communication using a 50-kg-class microsatellite. Nature Photon 11, 502–508 (2017). https://doi.org/10.1038/nphoton. 2017.107

47. Ding, J.; Mei, H.; I, C.-L.; Zhang, H.; Liu, W. Frontier Progress of Unmanned Aerial Vehicles Optical Wireless Technologies. Sensors 2020, 20, 5476. https://doi.org/10.3390/s20195476.

48. P. V. Trinh et al., "Experimental Channel Statistics of Drone-to-Ground Retro-Reflected FSO Links With Fine-Tracking Systems," in IEEE Access, vol. 9, pp. 137148-137164, 2021, doi: 10.1109/ACCESS.2021.3117266.

49. M. Najafi, H. Ajam, V. Jamali, P. D. Diamantoulakis, G. K. Karagiannidis and R. Schober, "Statistical Modeling of FSO Fronthaul Channel for Drone-Based Networks," 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 2018, pp. 1-7, doi: 10.1109/ICC.2018.8422552.

50. Witte, Brandon M., "Development of an Unmanned Aerial Vehicle for Atmospheric Turbulence Measurement" (2016). Theses and Dissertations--Mechanical Engineering. 82. https://uknowledge.uky.edu/me_etds/82

51. H. Kaushal and G. Kaddoum, "Underwater Optical Wireless Communication," in IEEE Access, vol. 4, pp. 1518-1547, 2016, doi: 10.1109/ACCESS.2016.2552538.

52. Nasir Saeed, Abdulkadir Celik, Tareq Y. Al-Naffouri, Mohamed-Slim Alouini, Underwater optical wireless communications, networking, and localization: A survey, Ad Hoc Networks, Volume 94, 2019, 101935, ISSN 1570-8705, https://doi.org/10.1016/j.adhoc.2019.101935.

53. Kao, C.-C.; Lin, Y.-S.; Wu, G.-D.; Huang, C.-J. A Comprehensive Study on the Internet of Underwater Things: Applications, Challenges, and Channel Models. *Sensors* **2017**, *17*, 1477. https://doi.org/10.3390/s17071477.

54. A. C. Boucouvalas, K. P. Peppas, K. Yiannopoulos and Z. Ghassemlooy, "Underwater Optical Wireless Communications With Optical Amplification and Spatial Diversity," in IEEE Photonics Technology Letters, vol. 28, no. 22, pp. 2613-2616, 15 Nov.15, 2016, doi: 10.1109/LPT.2016.2607278.

55. Andrews L.C.; Phillips R.L. Laser Beam Propagation through Random Media, 2nd ed.; SPIE Optical Engineering Press: Bellingham, WA, USA, 2005.

56. Heba Yuksel. Studies of the Effects of Atmospheric Turbulence on Free Space Optical Communications. PhD Thesis. Electrical Engineering Departmen. University of Maryland, College Park, 2005.

57. Angel Fernandez. Experiments for Laser Beam Propagation through Optical Turbulence: Development, Analysis and Applications. Mathematical Physics. Université d'Angers; Universidad técnica Federico Santa María (Valparaiso, Chili), 2016.

58. R. Barrios and F. Dios, 'Wireless Optical Communications Through the Turbulent Atmosphere: A Review', Optical Communications Systems. InTech, Mar. 07, 2012. doi: 10.5772/34740.

59. R.L. Fante, Electromagnetic beam propagation in turbulent media: An update, Proceedings of the IEEE, 10.1109/PROC.1980.11882, 68, 11, (1424-1443), (1980).

60. Nistazakis, H.E.; Katsis A.; Tombras, G.S. On the reliability and performance of FSO and hybrid FSO communication systems over turbulent channels; Nova Science Publishers: Hauppauge, NY, USA, 2011; pp. 69–112.64.

61. Rabinovich, W.S.; Moore, C.I.; Mahon, R.; Goetz, P.G.; Burris, H.R.; Ferraro, M.S.; Murphy, J.L.; Thomas, L.M.; Gilbreath, G.C.; Vilcheck, M.; Suite, M.R.; Free-space optical communications research and demonstrations at the U.S. Naval Research Laboratory, *Appl. Opt.* **2015**,*Vol. 54*, F189-F200.

62. Frederickson, P.A.; Davidson, K.L.; Zeisse, C.R.; Bendall, C.S. Estimating the refractive index structure parameter (Cn2) over the ocean using bulk methods. Journal of Applied Meteorology2000, 39, 1770-1783.

63. Frederickson, P.; Hammel, S.; Tsintikidis, D. Measurements and modeling of optical turbulence in a maritime environment. Proc SPIE 2006. 10.1117/12.683017.

64. Lionis A.; Cohn K.; Pogue C. Experimental Design of a UCAV-based High Energy Laser Weapon. Master's Thesis, Naval Postgraduate School, Monterey, CA, USA, December 2016.

65. Tunick, A. Optical turbulence parameters characterized via optical measurements over a 2.33-km free-space laser path. Opt. Express 2008, 16, 14645–14654.

66. Tunick, A. Statistical analysis of optical turbulence intensity over a 2.33 km propagation path. Opt. Expres 2007, 15, 3619–3628.

67. Tunick, A.; Nikolay, T.; Mikhail, V.; Gary, C. Characterization of optical turbulence (Cn2) data measured at the ARL A_LOT facility; Technical Report for U.S. Army Research Laboratory: Adelphi, MD, USA, September 2005.

68. Chang, Mark; Gilbreath, Gcharmaine; Oh, Eun; Distefano, Emi; Restaino, Sergio; Wilcox, Christopher and Santiago, Freddie. (2007). Comparing horizontal path C_n^2 measurements over 0.6 km in the tropical littoral environment and in the desert - art. No. 65510I. Proc SPIE. 10.1117/12.718257.

69. Jellen, Christopher, Charles Nelson, Cody J. Brownell, John Burkhardt and Miles Oakley. "Measurement and analysis of atmospheric optical turbulence in a near-maritime environment." IOP SciNotes 1 (2020).

70. Olga Korotkova, Svetlana Avramov-Zamurovic, Reza Malek-Madani, and Charles Nelson, "Probability density function of the intensity of a laser beam propagating in the maritime environment," Opt. Express 19, 20322-20331 (2011).

71. Oh, Eun & Ricklin, Jennifer & Gilbreath, Gcharmaine & Vallestero, Neil & Eaton, Frank. (2004). Optical turbulence model for laser propagation and Imaging applications. Proceedings of SPIE - The International Society for Optical Engineering. 25-32. 10.1117/12.504556.

72. Oh, Bun & Ricklin, Jennifer & Eaton, Frank & Gilbreath, CHarmaine & Doss-Hammel, Steve & Moore, Chris & Murphy, James & Oh, Yeonju & Stell, Mena. (2004). Estimating optical turbulence using the PAMELA model. Proceedings of SPIE - The International Society for Optical Engineering. 12. 10.1117/12.561481.

73. Wilfert, Otakar and Lucie Dordová. "Calculation and Comparison of Turbulence Attenuation by Different Methods." (2010).

74. Wilfert, Otakar and Lucie Dordová. Laser beam attenuation determined by the method of available optical power in turbulent atmosphere. Journal of Telecommunications and Information Technology, 2009, no. 2, p. 53-57.

75. Hongxing Wang, Bifeng Li, Xiaojun Wu, Chuanhui Liu, Zhihui Hu & Pengfei Xu (2015): Prediction model of atmospheric refractive index structure parameter in coastal area, Journal of Modern Optics, DOI: 10.1080/09500340.2015.1037801.

76. Frehlich, Rod & Sharman, Robert & Vandenberghe, Francois & Yu, Wei & Liu, Yubao & Knievel, Jason & Jumper, George. (2010). Estimates ofCn2from Numerical Weather Prediction Model Output and Comparison with Thermosonde Data. Journal of Applied Meteorology and Climatology. 49. 1742-1755. 10.1175/2010JAMC2350.1.

77. Alohotsy Rafalimanana, Christophe Giordano, Aziz Ziad, Eric Aristidi, "Prediction of atmospheric turbulence by means of WRF model for optical communications," Proc. SPIE 11852, International Conference on Space Optics — ICSO 2020, 118524G (11 June 2021); doi: 10.1117/12.2599659.

78. Chun Qing, Xiaoqing Wu, Xuebin Li, Wenyue Zhu, Chunhong Qiao, Ruizhong Rao, and Haipin Mei, "Use of weather research and forecasting model outputs to obtain near-surface refractive index structure constant over the ocean," Opt. Express 24, 13303-13315 (2016).

79. Cheinet, S., Beljaars, A., Weiss-Wrana, K. et al. The Use of Weather Forecasts to Characterise Near-Surface Optical Turbulence. Boundary-Layer Meteorol 138, 453–473 (2011). https://doi.org/10.1007/s10546-010-9567-z.

80. Chun Qing, Xiaoqing Wu, Honghua Huang, Qiguo Tian, Wenyue Zhu, Ruizhong Rao, and Xuebin Li, "Estimating the surface layer refractive index structure constant over snow and sea ice using Monin-Obukhov similarity theory with a mesoscale atmospheric model," Opt. Express 24, 20424-20436 (2016).

81. A. Vanderka; L. Hajek; J. Latal; J. Vitasek and P. Koudelka. Design, Simulation and Testing of the OOK NRZ Modulation Format for Free Space Optic Communication

in a Simulation Box. Advances in Electrical and Electronic Engineering Journal. Vol. 12, No 6. 2014.

82. Fahad A.Rida, Jafaar & Rida, A & Bhardwaj, A K & Jaiswal, A. (2014). Design Optimization of Optical Wireless Communication (OWC) Focusing On Light Fidelity (LI-FI) Using Optical Code Division Multiple Access (OCDMA) Based on Carbon Nanotubes (CNTS) International Journal of Advanced Research in Engineering and Technology (IJARET). Vol. 5. Issue 10, October 2014, pp. 69-103.

83. K. P. Peppas and P. T. Mathiopoulos, "Free-Space Optical Communication with Spatial Modulation and Coherent Detection Over H-K Atmospheric Turbulence Channels," in Journal of Lightwave Technology, vol. 33, no. 20, pp. 4221-4232, 15 Oct.15, 2015, doi: 10.1109/JLT.2015.2465385.

84. G. Yang, M. -A. Khalighi, T. Virieux, S. Bourennane and Z. Ghassemlooy, "Contrasting space-time schemes for MIMO FSO systems with non-coherent modulation," 2012 International Workshop on Optical Wireless Communications (IWOW), Pisa, Italy, 2012, pp. 1-3, doi: 10.1109/IWOW.2012.6349694.

85. A. L. Puryear. 2011. Optical Communication Through the Turbulent Atmosphere with Transmitter and Receiver Diversity, Wavefront Control, and Coherent Detection. PhD Thesis. Massachusetts Institute of Technology. Department of Electrical Engineering and Computer Science.

86. Aly, Moustafa. (2012). Pointing Error in FSO Link under Different Weather Conditions. International Journal of Video & Image Processing and Network Security. 12. 6-9.

87. Sabatini, Roberto and Richardson, Mark. "New techniques for laser beam atmospheric extinction measurements from manned and unmanned aerospace vehicles"

Open Engineering, vol. 3, no. 1, 2013, pp. 11-35. https://doi.org/10.2478/s13531-012-0033-1.

88. Federico Dios, Jaume Recolons, Alejandro Rodríguez, and Oscar Batet, "Temporal analysis of laser beam propagation in the atmosphere using computer-generated long phase screens," Opt. Express 16, 2206-2220 (2008).

89. Lars Sjöqvist, Markus Henriksson, Ove Steinvall, "Simulation of laser beam propagation over land and sea using phase screens: a comparison with experimental data," Proc. SPIE 5989, Technologies for Optical Countermeasures II; Femtosecond Phenomena II; and Passive Millimetre-Wave and Terahertz Imaging II, 59890D (3 November 2005); https://doi.org/10.1117/12.630655.

90. Jeong, D.-K.; Park, C.-S.; Kim, D. Statistical Analysis of Noise Propagation Effect for Mixed RF/FSO AF Relaying Application in Wireless Sensor Networks. Sensors 2020, 20, 979. https://doi.org/10.3390/s20040979.

91. He, D.; Wang, Q.; Liu, X.; Song, Z.; Zhou, J.; Wang, Z.; Gao, C.; Zhang, T.; Qi, X.; Tan, Y.; Ren, G.; Qi, B.; Ren, J.; Cao, Y.; Huang, Y. Shipborne Acquisition, Tracking, and Pointing Experimental Verifications towards Satellite-to-Sea Laser Communication. Appl. Sci. 2019, 9, 3940. https://doi.org/10.3390/app9183940.

92. Carolina Rickenstorff, José A. Rodrigo, and Tatiana Alieva, "Programmable simulator for beam propagation in turbulent atmosphere," Opt. Express 24, 10000-10012 (2016).

93. M. A. Esmail, H. Fathallah and M. -S. Alouini, "Outdoor FSO Communications Under Fog: Attenuation Modeling and Performance Evaluation," in IEEE Photonics Journal, vol. 8, no. 4, pp. 1-22, Aug. 2016, Art no. 7905622, doi: 10.1109/JPHOT.2016.2592705.

94. I. I. Kim, B. McArthur, and E. J. Korevaar, “Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications,” in Proc. SPIE Opt. Wireless Commun. III, Feb. 2001, vol. 4214, pp. 26–37.

95. Khan, Muhammad Saeed & Grabner, M & Muhammad, Sajid & Awan, M.s & Leitgeb, Erich & Kvicera, V. & Nebuloni, Roberto. (2012). Empirical Relations for Optical Attenuation Prediction from Liquid Water Content of Fog. Radioengineering. 21. 911-916.

96. Nadeem, F., & Leitgeb, E. (2010). Dense Maritime Fog Attenuation Prediction from Measured Visibility Data. Radioengineering, 19(2), 223-227.

97. M. Ijaz, Z. Ghassemlooy, H. Le Minh, S. Rajbhandari and J. Perez, "Analysis of fog and smoke attenuation in a free space optical communication link under controlled laboratory conditions," 2012 International Workshop on Optical Wireless Communications (IWOW), Pisa, Italy, 2012, pp. 1-3, doi: 10.1109/IWOW.2012.6349680.

98. J.Toth, M.Tatarko, L.Ovsenik, et.al. Free Space Optics Availability and Reliability. Carpathian Journal of Electronic and Computer Engineering 7/2 (2014) 19-23.

99. Grover, Marvi; Singh, Preeti and Kaur, Pardeep. (2017). Performance Analysis of Multibeam WDM-FSO System in Clear and Hazy Weather Conditions. 10.1007/978-981-10-1708-7_21.

100. Korai, U.A.; Luini, L.; Nebuloni, R. Model for the Prediction of Rain Attenuation Affecting Free Space Optical Links. Electronics 2018, 7, 407. https://doi.org/10.3390/electronics7120407.

101. Samir A. Al-Gailani, Abu Bakar Mohammad, Usman U. Sheikh, Redhwan Q. Shaddad, Determination of rain attenuation parameters for free space optical link in tropical rain, Optik, Volume 125, Issue 4, 2014, Pages 1575-1578, ISSN 0030-4026, https://doi.org/10.1016/j.ijleo.2013.10.018.

102. Zabidi, S. A., Islam, M. R., Al-Khateeb, W., & Naji, A. W. (2012). Analysis of Rain Effects on Terrestrial Free Space Optics based on Data Measured in Tropical Climate. IIUM Engineering Journal, 12(5). https://doi.org/10.31436/iiumej.v12i5.232.

103. Basahel, Ahmed A., Md. Rafiqul Islam, Mohamed Hadi Habaebi and Suriza Ahmad Zabidi. “A proposed rain attenuation prediction method for free space optical link based on rain rate statistics.” (2015).

104. K. P. Peppas, H. E. Nistazakis, and G. S. Tombras, ‘An Overview of the Physical Insight and the Various Performance Metrics of Fading Channels in Wireless Communication Systems’, Advanced Trends in Wireless Communications. InTech, Feb. 17, 2011. doi: 10.5772/15028.

105. Wayne, David, "The Pdf of Irradiance for A Free-space Optical Communications Channel: A Physics Based Model" (2010). PhD Thesis. School of Electrical Engineering and Computer Science. University of Central Florida. Orlando, FL.

106. José M. Garrido-Balsells, F. Javier Lopez-Martinez, Miguel Castillo-Vázquez, Antonio Jurado-Navas, and Antonio Puerta-Notario, "Performance analysis of FSO communications under LOS blockage," Opt. Express 25, 25278-25294 (2017).

107. López-González, F. J., Jurado-Navas, A., Garrido-Balsells, J. M., Castillo-Vázquez, M., & Puerta-Notario, A. (2017). Characterization of sub-channel based Málaga atmospheric optical links with real β parameter. Optica Applicata, 47(4), 545-556. https://doi.org/10.5277/oa170405.

108. José María Garrido-Balsells, Antonio Jurado-Navas, José Francisco Paris, Miguel Castillo-Vazquez, and Antonio Puerta-Notario, "Novel formulation of the $\mathscr{M}$ model through the Generalized-K distribution for atmospheric optical channels," Opt. Express 23, 6345-6358 (2015).

109. Dror, Itai; Atar, S.; Grossman, Shlomit; Kopeika, Norman S. Accurate method for prediction of atmospheric transmission according to weather. Optical Engineering, vol. 35, pp. 2548–2555, 1996. doi:10.1117/1.600863.

110. L. Kong, W. Xu, L. Hanzo, H. Zhang and C. Zhao, "Performance of a Free-Space-Optical Relay-Assisted Hybrid RF/FSO System in Generalized M-Distributed Channels," in IEEE Photonics Journal, vol. 7, no. 5, pp. 1-19, Oct. 2015, Art no. 7903319, doi: 10.1109/JPHOT.2015.2470106.

111. Ammar Al-Habash, Larry C. Andrews, and R. L. Phillips "Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media," Optical Engineering 40(8), (1 August 2001). https://doi.org/10.1117/1.1386641.

112. Bourazani, D.; Stasinakis, A.N.; Nistazakis, H.E.; Varotsos, G.K.; Tsigopoulos, A.D.; Tombras, G.S.; Experimental Accuracy Investigation for Irradiance Fluctuations of FSO Links Modeled by Gamma Distribution. In Proceedings of the 8th International Conference from Scientific Computing to Computational Engineering, Athens, Greece, 4–7 July 2018.

113. H. E. Nistazakis, M. P. Ninos, A. D. Tsigopoulos, D. A. Zervos and G. S. Tombras (2016) Performance study of terrestrial multi-hop OFDM FSO communication systems with pointing errors over turbulence channels, Journal of Modern Optics, 63:14, 1403-1413, DOI:10.1080/09500340.2016.1149626.

114. H. E. Nistazakis, G. S. Tombras, A. D. Tsigopoulos, E. A. Karagianni and M. E. Fafalios, "Capacity estimation of optical wireless communication systems over moderate to strong turbulence channels," in Journal of Communications and Networks, vol. 11, no. 4, pp. 384-389, Aug. 2009, doi: 10.1109/JCN.2009.6391352.

115. Antonio García-Zambrana, Carmen Castillo-Vázquez, Beatriz Castillo-Vázquez, and Rubén Boluda-Ruiz, "Bit detect and forward relaying for FSO links using equal gain combining over gamma-gamma atmospheric turbulence channels with pointing errors," Opt. Express 20, 16394-16409 (2012).

116. N. D. Chatzidiamantis, D. S. Michalopoulos, E. E. Kriezis, G. K. Karagiannidis and R. Schober, "Relay selection protocols for relay-assisted free-space optical systems," in Journal of Optical Communications and Networking, vol. 5, no. 1, pp. 92-103, Jan. 2013, doi: 10.1364/JOCN.5.000092.

117. C. K. Datsikas, K. P. Peppas, N. C. Sagias and G. S. Tombras, "Serial Free-Space Optical Relaying Communications Over Gamma-Gamma Atmospheric Turbulence Channels," in Journal of Optical Communications and Networking, vol. 2, no. 8, pp. 576-586, August 2010, doi: 10.1364/JOCN.2.000576.

118. M. Uysal, Jing Li and Meng Yu, "Error rate performance analysis of coded free-space optical links over gamma-gamma atmospheric turbulence channels," in IEEE Transactions on Wireless Communications, vol. 5, no. 6, pp. 1229-1233, June 2006, doi: 10.1109/TWC.2006.1638639.

119. M. A. Kashani, M. Uysal and M. Kavehrad, "A novel statistical model for turbulence-induced fading in free-space optical systems," 2013 15th International Conference on Transparent Optical Networks (ICTON), Cartagena, Spain, 2013, pp. 1-5, doi: 10.1109/ICTON.2013.6602795.

120. M. A. Kashani, M. Uysal and M. Kavehrad, "On the performance of MIMO FSO communications over Double Generalized Gamma fading channels," 2015 IEEE International Conference on Communications (ICC), London, UK, 2015, pp. 5144-5149, doi: 10.1109/ICC.2015.7249140.

121. N. D. Chatzidiamantis, H. G. Sandalidis, G. K. Karagiannidis, S. A. Kotsopoulos and M. Matthaiou, "New results on turbulence modeling for free-space optical systems," 2010 17th International Conference on Telecommunications, Doha, Qatar, 2010, pp. 487-492, doi: 10.1109/ICTEL.2010.5478872.

122. N. C. Sagias, G. K. Karagiannidis, P. T. Mathiopoulos, and T. A. Tsiftsis, “On the performance analysis of equal-gain diversity receivers over generalised gamma fading channels,” IEEE Trans. Wireless Commun., vol. 5, no. 10, pp. 2967–2975, Oct. 2006.

123. G.K. Varotsos, A.N. Stassinakis, H.E. Nistazakis, A.D. Tsigopoulos, K.P. Peppas, C.J. Aidinis, G.S. Tombras, Probability of fade estimation for FSO links with time dispersion and turbulence modeled with the gamma–gamma or the I-K distribution, Optik, Volume 125, Issue 24, 2014, Pages 7191-7197, ISSN 0030-4026, https://doi.org/10.1016/j.ijleo.2014.08.047.

124. H. Nistazakis; A. Tsigopoulos; M. Hanias; C. Psychogios; D. Marinos; C. Aidinis and G. Tombras. (2011). Estimation of Outage Capacity for Free Space Optical Links over IK and K Turbulent Channels. Radioengineering. 20.

125. K. P. Peppas, G. C. Alexandropoulos, E. D. Xenos and A. Maras, "The Fischer–Snedecor F -Distribution Model for Turbulence-Induced Fading in Free-Space Optical Systems," in Journal of Lightwave Technology, vol. 38, no. 6, pp. 1286-1295, 15 March15, 2020, doi: 10.1109/JLT.2019.2957327.

126. S. K. Yoo, S. L. Cotton, P. C. Sofotasios, M. Matthaiou, M. Valkama and G. K. Karagiannidis, "The Fisher–Snedecor F Distribution: A Simple and Accurate Composite Fading Model," in IEEE Communications Letters, vol. 21, no. 7, pp. 1661-1664, July 2017, doi: 10.1109/LCOMM.2017.2687438.

127. Maged A. Esmail, Experimental performance evaluation of weak turbulence channel models for FSO links, Optics Communications, Volume 486, 2021, 126776, ISSN 0030-4018, https://doi.org/10.1016/j.optcom.2021.126776.

128. R. Barrios. 2013. Exponentiated Weibull fading channel model in free-space optical communications under atmospheric turbulence. PhD Thesis. Department of Signal Theory and Communications. Universitat Polit`ecnica de Catalunya. Barcelona, Spain.

129. R. Barrios and F. Dios. (2012). Exponentiated Weibull model for the irradiance probability density function of a laser beam propagating through atmospheric turbulence. Optics & Laser Technology. 45. 10.1016/j.optlastec.2012.08.004.

130. A.K. Majumdar. Free-space laser communication performance in the atmospheric channel. J. Opt. Fiber Commun. 2005, 2, 345–396.

131. A. Jurado-Navas, J. M. Garrido Balsells, J. Francisco Paris, M. Castillo-Vázquez, and A. Puerta-Notario, "General analytical expressions for the bit error rate of atmospheric optical communication systems," Opt. Lett. 36, 4095-4097 (2011).

132. H. G. Sandalidis, T. A. Tsiftsis, G. K. Karagiannidis and M. Uysal, "BER Performance of FSO Links over Strong Atmospheric Turbulence Channels with Pointing Errors," in IEEE Communications Letters, vol. 12, no. 1, pp. 44-46, January 2008, doi: 10.1109/LCOMM.2008.071408.

133. K. Hyeong-Ji, T. Samrat Vikramaditya and C. Yeon-Ho, "Multi-hop relay-based maritime visible light communication," Chin. Opt. Lett. 14, 050607- (2016).

134. J. Libich, M. Komanec, S. Zvanovec, P. Pesek, W. O. Popoola, and Z. Ghassemlooy, "Experimental verification of an all-optical dual-hop 10 Gbit/s free-space optics link under turbulence regimes," Opt. Lett. 40, 391-394 (2015).

135. L. Li, R. Zhang, Z. Zhao et al. High-Capacity Free-Space Optical Communications Between a Ground Transmitter and a Ground Receiver via a UAV Using Multiplexing of Multiple Orbital-Angular-Momentum Beams. Sci Rep 7, 17427 (2017). https://doi.org/10.1038/s41598-017-17580-y.

136. Le Minh, H., Ghassemlooy, Z., Ijaz, M., Rajbhandari, S., Adebanjo, O., Ansari, S., Leitgeb, E. (2010) 'Experimental study of bit error rate of free space optics communications in laboratory controlled turbulence', Workshop on Optical Wireless Communications in conjunction with the IEEE Globecom 2010, Miami, Florida, USA, 6-10 December. Institute of Electrical and Electronics Engineers Globecom Workshops, pp. 1072-1076.

137. Garlinska, M.; Pregowska, A.; Gutowska, I.; Osial, M.; Szczepanski, J. Experimental Study of the Free Space Optics Communication System Operating in the 8–12 μm Spectral Range. Electronics 2021, 10, 875. https://doi.org/10.3390/electronics10080875.

138. K. Peppas, A. Stassinakis, H. Nistazakis and G. Tombras. (2013). Capacity Analysis of Dual Amplify-and-Forward Relayed Free-Space Optical Communication Systems Over Turbulence Channels with Pointing Errors. Optical Communications and Networking, IEEE/OSA Journal of. 5. 1032-1042. 10.1364/JOCN.5.001032.

139. Trung, Ha & Ai, Duong & Pham, Anh. (2015). Average Channel Capacity of Free-Space Optical MIMO Systems Over Atmospheric Turbulence Channels. ASEAN Engineering Journal Part A, Vol 5 No 2, ISSN 2229-127X p.57. 5. 57-66.

140. S. Sharma, J. Tan, A. S. Madhukumar and R. Swaminathan, "Switching-Based Transmit Antenna/Aperture Selection in a MISO Hybrid FSO/RF System," 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647284.

141. A.J. Kshatriya, Y.B. Acharya, A.K. Aggarwal. et al. Estimation of FSO link availability using climatic data. J Opt 45, 324–330 (2016). https://doi.org/10.1007/s12596-016-0327-4.

142. T. Ismail, E. Leitgeb, and M. Al-Nahhal. (2018). Performance Improvement of FSO System using Multi-Pulse PPM and SIMO under Atmospheric Turbulence Conditions and with Pointing Errors. IET Networks. 7. 10.1049/iet-net.2017.0203.

143. C. Castillo-Vázquez, R. Boluda-Ruiz, B. Castillo-Vázquez and A. García-Zambrana, "Outage performance of DF relay-assisted FSO communications using time diversity," 2015 IEEE Photonics Conference (IPC), Reston, VA, USA, 2015, pp. 423-426, doi: 10.1109/IPCon.2015.7323709.

144. C. Abou-Rjeily, "Performance Analysis of FSO Communications With Diversity Methods: Add More Relays or More Apertures?" in IEEE Journal on Selected Areas in Communications, vol. 33, no. 9, pp. 1890-1902, Sept. 2015, doi: 10.1109/JSAC.2015.2432526.

145. Y. Guowei; S. Rajbhandari; M.A. Khalighi and S. Bourennane. (2012). Experimental Works on Free-Space Optical Communications with Aperture Averaging and Receive Diversity in a Controlled Laboratory Environment. Conference: Actes des Journées d'études Algéro-Françaises de Doctorants en Signal, Image & Applications. Alger, Algeria.

146. Polnau, Ernst & Vorontsov, Mikhail & Carhart, Gary & Weyrauch, Thomas & Beresnev, Leonid. (2007). Adaptive compensation over a 2.33 km propagation path

with retro reflectors under strong scintillation conditions. Atmospheric Optics: Models, Measurements, and Target-in-the-Loop Propagation. Proc. of SPIE Vol. 6708 67080C-1.

147. K. Yiannopoulos, N. C. Sagias and A. C. Boucouvalas, "Fade Mitigation Based on Semiconductor Optical Amplifiers," in Journal of Lightwave Technology, vol. 31, no. 23, pp. 3621-3630, Dec.1, 2013, doi: 10.1109/JLT.2013.2285260.

148. K. Yiannopoulos, N. C. Sagias and A. C. Boucouvalas, "On the Performance of Semiconductor Optical Amplifier-Assisted Outdoor Optical Wireless Links," in IEEE Journal on Selected Areas in Communications, vol. 33, no. 9, pp. 1869-1876, Sept. 2015, doi: 10.1109/JSAC.2015.2433052.

149. A. Beck and M. Kurz. A Perspective on Machine Learning Methods in Turbulence Modelling. GAMM-Mitteilungen 44. https://doi.org/10.1002/gamm.202100002.

150. G. James, D. Witten, T. Hastie and R. Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. Springer New York Heidelberg Dordrecht London.

151. R. S. Sutton and A. G. Barto, *Introduction to reinforcement learning*. MIT press Cambridge, 1998, vol. 135.

152. T. K. HO, Random decision forests, in Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1, IEEE, 1995, pp. 278–282.

153. Cifuentes, Jenny; Marulanda, Geovanny; Bello, Antonio; Reneses, Javier. 2020. Air Temperature Forecasting Using Machine Learning Techniques: A Review. Energies 13, no. 16: 4215.

154. M.A. Amirabadi. A survey on machine learning for optical communication [machine learning view]. arXiv preprint arXiv:1909.05148 (2019).

155. D. Wang, Y. Song, J. Li, J. Qin, T. Yang, M. Zhang, X. Chen and A. Boucouvalas. Data-driven Optical Fiber Channel Modeling: A Deep Learning Approach. Journal of Lightwave Technology. vol. 38, no. 17, pp. 4730-4743, 1 Sept.1, 2020, doi: 10.1109/JLT.2020.2993271.

156. J. Liu, P. Wang, X. Zhang, Y. He, X. Zhou, H. Ye, Y. Li, S. Xu, S. Chen, and D. Fan, "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," Opt. Express 27, 16671-16688 (2019).

157. M. Amirabadi, M.Kahaei, S. A. Nezamalhosseini and V. T. Vakili. Deep Learning for channel estimation in FSO communication system. Optics Communications, Vol. 459, 2020, 124989, ISSN 0030-4018, https://doi.org/10.1016/j.optcom.2019.124989.

158. S. Lohani and R. Glasser. Turbulence correction with artificial neural networks. Opt. Lett. 43, 2611-2614 (2018).

159. S. Lohani, E.M. Knutson and R.T. Glasser. Generative machine learning for robust free-space communication. Communations Physics 3, 177 (2020). https://doi.org/10.1038/s42005-020-00444-9.

160. P. Mishra, Sonali, A. Dixit and V. K. Jain, "Machine Learning Techniques for Channel Estimation in Free Space Optical Communication Systems," 2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), GOA, India, 2019, pp. 1-6, doi: 10.1109/ANTS47819.2019.9117976.

161. C. Jellen, J. Burkhardt, C. Brownell, and C. Nelson, "Machine learning informed predictor importance measures of environmental parameters in maritime optical turbulence," Appl. Opt. 59, 6379-6389 (2020).

162. Y. Wang and S. Basu, "Using an artificial neural network approach to estimate surface-layer optical turbulence at Mauna Loa, Hawaii," Opt. Lett. 41, 2334-2337 (2016).

163. Haluška, R.; Šuľaj, P.; Ovseník, Ľ.; Marchevský, S.; Papaj, J.; Doboš, Ľ. Prediction of Received Optical Power for Switch-ing Hybrid FSO/RF System. Electronics 2020, 9, 1261.

164. J. Tóth, L. Ovseník, J. Turán, L. Michaeli, M. Márton, Classification prediction analysisof RSSI parameter in hard switch-ing process for FSO/RF systems, Measurement (2017), doi: https://doi.org/10.1016/j.measurement.2017.11.044

165. Runqin Xu, Pin Lv, Fanjiang Xu, Yishi Shi. A survey of approaches for implementing optical neural networks. Optics & Laser Technology, Volume 136, 2021, 106787, ISSN 0030-3992, https://doi.org/10.1016/j.optlastec.2020.106787.

166. S. K. Goudos, G. V. Tsoulos, G. Athanasiadou, M. C. Batistatos, D. Zarbouti and K. E. Psannis, "Artificial Neural Network Optimal Modeling and Optimization of UAV

Measurements for Mobile Communications Using the L-SHADE Algorithm," in IEEE Transactions on Antennas and Propagation, vol. 67, no. 6, pp. 4022-4031, June 2019, doi: 10.1109/TAP.2019.2905665.

167. V.A. Kulikov; S.L. Lachinova; M.A. Vorontsov; V.S.R. Gudimetla. Characterization of Localized Atmospheric Turbulence Layer Using Laser Light Backscattered off Moving Target. Appl. Sci. 2020, 10, 6887. https://doi.org/10.3390/app10196887.

168. C. Bi; C. Qing; P. Wu; X. Jin; Q. Liu; X. Qian; W. Zhu; N. Weng. Optical Turbulence Profile in Marine Environment with Artificial Neural Network Model. Remote Sens. 2022, 14, 2267. https://doi.org/10.3390/rs14092267.

169. M. Xu; S. Shao; Q. Liu; G. Sun; Y. Han; N. Weng. Optical Turbulence Profile Forecasting and Verification in the Offshore Atmospheric Boundary Layer. Appl. Sci. 2021, 11, 8523. https://doi.org/10.3390/app11188523.

170. A.M. Vorontsov; M.A. Vorontsov; G.A. Filimonov; E. Polnau. Atmospheric Turbulence Study with Deep Machine Learning of Intensity Scintillation Patterns. Appl. Sci. 2020, 10, 8136. https://doi.org/10.3390/app10228136.

171. J. Liu, P. Wang, X. Zhang, Y. He, X. Zhou, H. Ye, Y. Li, S. Xu, S. Chen, and D. Fan, "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," Opt. Express 27, 16671-16688 (2019).

172. J. Li, M. Zhang, and D. Wang, “Adaptive demodulator using machine learning for orbital angular momentum shift keying,” IEEE Photonics Technol. Lett. **29**(17), 1455–1458 (2017).

173. W. Xiong, P. Wang, M. Cheng, J. Liu, Y. He, X. Zhou, J. Xiao, Y. Li, S. Chen, and D. Fan, "Convolutional Neural Network Based Atmospheric Turbulence Compensation for Optical Orbital Angular Momentum Multiplexing," J. Lightwave Technol. 38, 1712-1721 (2020).

174. Lohani, S., Knutson, E.M. & Glasser, R.T. Generative machine learning for robust free-space communication. Commun Phys 3, 177 (2020). https://doi.org/10.1038/s42005-020-00444-9.

175. Ghassemlooy, Z.; Popoola, W.O. Terrestrial Free-Space Optical Communications. In *Mobile and Wireless Communications Network Layer and Circuit*

Level Design; Fares, S.A., Adachi, F., Eds.; Books on Demand: Metro Manila, Philippines, 2010; pp. 355–392.

176. S. Doss-Hammel, D. Tsindikidis, D. Merritt, J. Fontana, Atmospheric characterization for high energy laser beam propagation in the maritime environment, in: Atmospheric Tracking, Imaging and Compensation, Proceedings of the SPIE 5552, Bellingham, WA, 2004; Michael T. Valley, Mikhail Vorontsov.

177. A.N. De Jong, P.B. Schwering, A.M. Van Eijk, W.H. Gunter, Validation of atmospheric propagation models in littoral waters, Opt. Eng. 52 (4) (2013) 046002.

178. Latal, J.; Vitasek, J.; Hajek, L.; Vanderka, A.; Koudelka, P.; Kepak, S.; Vasinek, V. Regression Models Utilization for RSSI Prediction of Professional FSO Link with Regards to Atmosphere PhenomenaIn Proceedings of the 2016. International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom), Graz, Austria, 14–16 September 2016.

179. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* 1951, *22*, 79–86.

180. https://weather-stats.com/greece/athenes/sea_temperature#details.

181. R. Mahon, C. Moore, H. Burris, W. Rabinovich, M. Suite, L. Thomas, Power spectra of a free space optical link in a maritime environment, in: Proc. SPIE 7464, Free Space Laser Communications IX, San Diego, CA, 2009, http://dx.doi. org/10.1117/12.828860.

182. D. Sadot, N.S. Kopeika, Forecasting optical turbulence strength on the basis of macroscale meteorology and aerosols: models and validation, Opt. Eng. 31 (1992) http://dx.doi.org/10.1117/12.56059.

183. Latal, J.; Vitasek, J.; Hajek, L.; Vanderka, A.; Koudelka, P.; Kepak, S.; Vasinek, V. Regression Models Utilization for RSSI Prediction of Professional FSO Link with

Regards to Atmosphere Phenomena In Proceedings of the 2016. International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom), Graz, Austria, 14–16 September 2016.

184. Hajek, L.; Vitasek, J.; Vanderka, A.; Latal, J.; Perecar, F.; Vasinek, V. Statistical prediction of the atmospheric behavior for free space optical link. In Proceedings of the SPIE 9614, Laser Communication and Propagation through the Atmosphere and Oceans IV, San Diego, CA, USA, 4 September 2015.

185. A. Lionis, K. Peppas, A. Tsigkopoulos, A. Sklavounos, A. Stasinakis, H. Nistazakis, K. Kohn, K. Aidinis, September 21, 2022, "Experimental Machine Learning Approach for Optical Turbulence and FSO Outage Performance Modeling", IEEE Dataport, doi: https://dx.doi.org/10.21227/8bqw-gy72.

186. Kullback, S., Leibler, R.: On Information and Sufficiency. Annals of Mathematical Statistics 22(1) (1951) 79-86.

187. Cover, T.M.; Thomas, J.A. Elements of Information Theory; John Wiley & Sons: Hoboken, NJ, USA, 2012.

188. Fuglede, B.; Topsoe, F. (2004). "Jensen-Shannon divergence and Hilbert space embedding". Proceedings of the International Symposium on Information Theory, 2004. IEEE. p. 30. doi:10.1109/ISIT.2004.1365067. ISBN 978-0-7803-8280-0.

189. Nielsen F. On a Generalization of the Jensen-Shannon Divergence and the Jensen-Shannon Centroid. Entropy (Basel). 2020 22(2):221. doi: 10.3390/e22020221.

190. Lahcene, Bachioua. (2013). On Pearson families of distributions and its applications. African Journal of Mathematics and Computer Science Research. Vol. 6(5), pp. 108-117. DOI: 10.5897/AJMCSR2013.0465.

191. Wei-Liem, Loh. (2004). On the characteristic function of Pearson type IV distributions. Institute of Mathematical Statistics. (Lecture Notes-Monograph Series). Vol. 45. 171-179.