1

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services Branch

Direction des acquisitions et
des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file / Votre référence

Our file / Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Statistical Range Prediction Algorithms For Indoor Wireless Systems

by

**Angsuman Rudra, B.Tech. (Hons.)
Indian Institute of Technology, Kharagpur.**

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of

Master of Engineering

**Ottawa-Carleton Institute for Electrical Engineering,
Faculty of Engineering,
Department of Systems and Computer Engineering,
Carleton University,
Ottawa, Canada.**

September 1995

© copyright
1995, Angsuman Rudra

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-08984-3

Canada

Name ANGSUMAN RUTRA

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

STATISTICAL RANGE PREDICTION ALGORITHMS FOR INDOOR

SUBJECT TERM

WIRELESS SYSTEMS

0544

U·M·I

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language
General 0679
Ancient 0289
Linguistics 0290
Modern 0291
Literature
General 0401
Classical 0294
Comparative 0295
Medieval 0297
Modern 0298
African 0316
American 0591
Asian 0305
Canadian (English) 0352
Canadian (French) 0355
English 0593
Germanic 0311
Latin American 0312
Middle Eastern 0315
Romance 0313
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion
General 0318
Biblical Studies 0321
Clergy 0319
History of 0320
Philosophy of 0322
Theology 0469

SOCIAL SCIENCES

American Studies 0323
Anthropology
Archaeology 0324
Cultural 0326
Physical 0327
Business Administration
General 0310
Accounting 0272
Banking 0770
Management 0454
Marketing 0338
Canadian Studies 0385
Economics
General 0501
Agricultural 0503
Commerce-Business 0505
Finance 0508
History 0509
Labor 0510
Theory 0511
Folklore 0358
Geography 0366
Gerontology 0351
History
General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0398
Political Science
General 0615
International Law and Relations 0616
Public Administration 0617
Recreation 0314
Social Work 0452
Sociology
General 0626
Criminology and Penology 0627
Demography 0938
Ethnic and Racial Studies 0631
Individual and Family Studies 0628
Industrial and Labor Relations 0629
Public and Social Welfare 0630
Social Structure and Development 0700
Theory and Methods 0344
Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture
General 0473
Agronomy 0285
Animal Culture and Nutrition 0475
Animal Pathology 0476
Food Science and Technology 0359
Forestry and Wildlife 0478
Plant Culture 0479
Plant Pathology 0480
Plant Physiology 0817
Range Management 0777
Wood Technology 0746
Biology
General 0306
Anatomy 0287
Biostatistics 0308
Botany 0309
Cell 0379
Ecology 0329
Entomology 0353
Genetics 0369
Limnology 0793
Microbiology 0410
Molecular 0307
Neuroscience 0317
Oceanography 0416
Physiology 0433
Radiation 0821
Veterinary Science 0778
Zoology 0472
Biophysics
General 0786
Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodetic 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
Health Sciences
General 0566
Audiology 0300
Chemotherapy 0992
Dentistry 0567
Education 0350
Hospital Management 0769
Human Development 0758
Immunology 0982
Medicine and Surgery 0564
Mental Health 0347
Nursing 0569
Nutrition 0570
Obstetrics and Gynecology 0380
Occupational Health and Therapy 0354
Ophthalmology 0381
Pathology 0571
Pharmacology 0419
Pharmacy 0572
Physical Therapy 0382
Public Health 0573
Radiology 0574
Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
Chemistry
General 0485
Agricultural 0749
Analytical 0486
Biochemistry 0487
Inorganic 0488
Nuclear 0738
Organic 0490
Pharmaceutical 0491
Physical 0494
Polymer 0495
Radiation 0754
Mathematics 0405
Physics
General 0605
Acoustics 0986
Astronomy and Astrophysics 0606
Atmospheric Science 0608
Atomic 0748
Electronics and Electricity 0607
Elementary Particles and High Energy 0798
Fluid and Plasma 0759
Molecular 0609
Nuclear 0610
Optics 0752
Radiation 0756
Solid State 0611
Statistics 0463
Applied Sciences
Applied Mechanics 0346
Computer Science 0984

Engineering
General 0537
Aerospace 0538
Agricultural 0539
Automotive 0540
Biomedical 0541
Chemical 0542
Civil 0543
Electronics and Electrical 0544
Heat and Thermodynamics 0348
Hydraulic 0545
Industrial 0546
Marine 0547
Materials Science 0794
Mechanical 0548
Metallurgy 0743
Mining 0551
Nuclear 0552
Packaging 0549
Petroleum 0765
Sanitary and Municipal 0554
System Science 0790
Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

The undersigned recommend to the Faculty of Graduate Studies
and Research acceptance of the thesis

**“Statistical Range Prediction Algorithms For Indoor Wireless
Systems”**

submitted by Angsuman Rudra, B.Tech.(Hons.)
in partial fulfilment of the requirements for
the degree of Master of Engineering.

Prof. S.A.Mahmoud, Thesis Supervisor.

Dr. G.Boudreau, Thesis Supervisor.

Chair, Department of Systems and Computer Engineering.

Carleton University
December 1995

Abstract

There has been a phenomenal growth in wireless communications during the past decade and the trend is expected to continue. The next century would see the wide spread availability of PCN services and thus accurate range prediction is an important issue for researchers. This thesis examines accurate and easy to use methods for predicting ranges for indoor wireless systems. Two approaches for range prediction has been proposed: Direct method and the Regression Analysis method. The Direct method involves directly including the slow fading component (shadow) in the prediction technique. The second approach (Regression Analysis approach) examines methods to enhance existing range prediction models. New variables and site specific functional classification have been added to the existing model to improve its accuracy and to enable a non-technical person (salesperson) to identify these variables for a new site. Linear and logarithmic regression analysis has been carried out to predict the dependant variable range as a function of the independent variables identified. Predictions have been done for known and unknown sites to verify the model selected. The prediction results are in agreement with the actual results obtained by experimentation.

Acknowledgement

I express my sincere gratitude to my thesis supervisors, Prof. Samy Mahmoud and Dr. Gary Boudreau for their guidance and encouragement throughout the programme. I thank Dr. Tim Speight for his help in conducting the experiments at Carleton University and providing insight to the data processing. Help from the departmental staff: Ms. Janet Scharf, Ms. Christy Harris, Ms. Darlene Hebert, Ms. Vivienne Gilchrist, Mr. Don Kerr, Mr. Naren Mehta, Mr. Danny Lemay, Mr. John Pope is also appreciated. Special thanks are due to all my friends and fellow graduate students for their help and the illuminating discussions that made my stay here enjoyable. In this short space it is not possible to name all of them, and I will not try to do the impossible. I would also like to thank Mr. Brian Egan, Mr. Derek Yu, Mr. Ralph Carsten, Mr. William Bird, Mr. Koon Teo, Mr. Padraig McTiernan of Bell Northern Research for welcoming me in their group and taking time out of their busy schedule to help me. Thanks are also due to Bell Northern Research for supplying me with the data needed for my work.

In addition I am grateful to Carleton University and Bell Northern Research for providing me with the financial support needed to complete my studies.

I should express my sincere thanks to CGI, Ottawa and RES International, Ottawa for allowing me to conduct experiments at their premises.

Finally I would like to appreciate the love, encouragement and support I received from my parents, brother and other family members.

Table of Contents

List of Tables	viii
List of Figures.....	ix
List of Abbreviations and Acronyms	xii
List of Symbols	xiv
List of Appendices.....	xvii
1 INTRODUCTION.....	1
1.1 Thesis Objective.....	4
1.2 Thesis Organization	5
2 INDOOR RADIO CHANNEL MODEL AND PREDICTION TECHNIQUES.....	6
2.1 Introduction.....	6
2.2 Impulse Response Profile of the Indoor Channel	6
2.2.1 Characterization of the Impulse Response.....	9
2.2.1.1 Distribution of the path amplitudes.....	9
2.2.1.2 Distribution of the arrival times.....	12
2.2.1.3 Other Channel Parameters	14
2.2.2 Correlation and Power Spectral Density Functions of the Impulse Response	15
2.3 Prediction Techniques.....	17
2.3.1 Direct Prediction Method.....	17

2.3.2	Prediction using Regression Analysis.....	19
2.4	Summary	20
3	VARIATION OF σ WITH DISTANCE	21
3.1	Introduction.....	21
3.2	Experimental Setup.....	21
3.2.1	2.0 GHz.....	21
3.2.2	The 900 MHz Band.....	22
3.3	Data Post Processing.....	24
3.3.1	2.0 GHz.....	24
3.3.2	The 900 MHz Band.....	26
3.4	Indoor Environment.....	26
3.5	Statistical Testing of Experimental Distribution.....	26
3.5.1	The Kolmogorov-Smirnov goodness of fit test.....	27
3.5.1.1	Illustration of the K-S test with an example	27
3.6	Results.....	30
3.6.1	University Environment.....	30
3.6.2	Office Environment.....	41
3.7	Summary	41
4	REGRESSION ANALYSES FOR RANGE PREDICTION	45
4.1	Introduction.....	45
4.2	Experimental Set Up.....	45
4.3	Variables Affecting Ranges.....	46
4.3.1	Existing Variables	46
4.3.2	New Variables	49
4.3.2.1	Building Type.....	49
4.3.2.2	Variables based on Functionality	52
4.4	Regression Analysis.....	55
4.4.1	Combining Log and Linear Prediction	57

4.5	Summary	60
5	RESULTS AND VERIFICATION OF REGRESSION ANALYSES	61
5.1	Introduction	61
5.2	Regression Results	61
5.3	Verification of the Model	82
5.3.1	Known Sites	82
5.3.2	Unknown Sites	84
5.4.	Summary	87
6	CONCLUSIONS	88
6.1	Thesis Summary	88
6.2	Future Work	89
Appendix A	DERIVATION OF THE CUMULATIVE DISTRIBUTION FUNCTION FOR A NORMALIZED RAYLEIGH VARIABLE	91
Appendix B	SOURCE CODE FOR THE COMPUTATION OF THE CUMULATIVE DENSITY FUNCTION.....	93
Appendix C	SOURCE CODE FOR THE KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST	94
Appendix D	SOURCE CODE FOR SIMULATING THE NORMALIZED POWER IN PRESENCE OF DOPPLER AND RAYLEIGH FADING.....	96
References		98

List of Tables

TABLE 3.1	Results of the KS test for the CDFs for various distance separation.....	31
TABLE 4.1	Definition of Area Types for Indoor Channels	47
TABLE 4.2	Definition of Building Types for Indoor Channels.....	50
TABLE 4.3	Definition of Functional Area Types	53
TABLE 4.4	Statistics of the Functional Area Types	54
TABLE 5.1	Results for Regression Analysis	61
TABLE 5.2	Mapping of Building Type into Average Values for Area Type	64
TABLE 5.3	Coefficients from the analysis VIIIa (using “Range (m)”).....	78
TABLE 5.4	Coefficients from the analysis VIIIb (using “log(range)”).....	80
TABLE 5.5.	Cell Count from Predicted Range using analysis VIII a & b	82
TABLE 5.6	Summary of the geometric variables for the unknown sites.....	84
TABLE 5.7	Summary of the functional variable for the unknown sites.....	85
TABLE 5.8	Cell Count from the predicted Range using analysis VIII a and b for unknown sites	86
TABLE 5.8.	Summary of the Statistics of the Residuals from the experimental data for the unknown sites	86

List of Figures

FIGURE 2.1.	Schematic representation of the channel. (a) represents the Mathematical Model of the Channel and (b) represents the response of the time-variant multipath channel to narrow pulses.	8
FIGURE 2.2	Correlation and Power Spectral Density Functions of the Channel Impulse Response. (a) shows the delay power spectrum, the spaced-frequency correlation function, the coherence bandwidth and the multipath spread of the channel. (b) shows the spaced-time correlation function, the Doppler power spectrum, the coherence time and the Doppler spread of the channel.	18
FIGURE 3.1	The experimental setup for the continuous wave Measurements at 2 GHz.	23
FIGURE 3.2	Schematic of the experimental procedure for 2.0 GHz.	23
FIGURE 3.3	Schematic of the experimental set up for the 900 MHz band	25
FIGURE 3.4	Theoretical (broken line) and Simulated (solid line) Cumulative Distribution Function.	29
FIGURE 3.5	Variation of σ with distance for the Minto Case Building in Carleton University.	32
FIGURE 3.6	Variation of σ with distance for the Mackenzie Building in Carleton University.	33
FIGURE 3.7	Variation of σ with distance for the Library in Carleton University.	34
FIGURE 3.8	Variation of σ with distance for the Dunton Tower in Carleton University.	35
FIGURE 3.9	Experimental vs. Theoretical Cumulative Distribution Function for the Library.	37
FIGURE 3.10	Experimental vs. Theoretical Cumulative Distribution Function for the Minto Case Building. The plots for the three receive antennae are shown separately.	38
FIGURE 3.11	Experimental vs. Theoretical Cumulative Distribution Function for the Mackenzie Building. The plot for the four receive antennae are shown separately.	39

FIGURE 3.12	Experimental vs. Theoretical Cumulative Distribution Function for the Dunton Tower. The plots for the four receive antennae are shown separately.	40
FIGURE 3.13	Variation of σ with distance in BNR (Carling) Lab 3. Each point on the graph represents several location with the distance being the same.	42
FIGURE 3.14	Variation of σ with distance for BNR (Carling), Administration Building. The two plots are for two different positions of the base station.....	43
FIGURE 3.15	Variation of σ with distance for BNR (Carling), Administration Building. This is the composite plot i.e. combining the two graphs in Figure 3.14.....	44
FIGURE 4.1	The cell boundary (at any pre-determined RSS) is determined by walking away from the base station till the requisite RSS is reached.....	48
FIGURE 4.2	Example of a mixed kind of environment (clearly demarcated areas).....	51
FIGURE 4.3	Definition for Rb. (a): Log Analysis, error in log; (b): Transformed to the linear Domain; (c): Range Analysis, error in metres and (d): Superimposing (b) and (c).	59
FIGURE 5.1	Residuals vs. Fitted "log(range)" (case VIII b). Note that the residual error is approximately constant.....	65
FIGURE 5.2	Residuals vs. Fitted Range for "log (range)" analysis (case VIII b), transformed to the linear domain (in metres).	65
FIGURE 5.3	Residuals Histogram for "log (range)" analysis (case VIII b). Note that it is very close to the Gaussian pdf.	66
FIGURE 5.4	Residuals Probability plot for "log (range)" analysis (case VIII b). The straight line represents the plot for Gaussian distribution.	66
FIGURE 5.5	Residuals vs. Fitted Range for "Range (m)" analysis (case VIII a). Note that the fan shaped nature is much more pronounced than that for the "log (range)" analysis (case VIII b) (Figure 5.1, Figure 5.2)	67
FIGURE 5.6	Residuals Histogram for "Range (m)" analysis (case VIII a). Note that there is a deviation from the Gaussian pdf.	67
FIGURE 5.7	Residuals Probability Plot for the "Range (m)" analysis (case VIII a). Note that there is a deviation from the Gaussian case.	68

FIGURE 5.8	Residuals vs. Fitted Range for the "Range (m)" analysis (case I a). Note that the fan shaped nature is pronounced.....	68
FIGURE 5.9	Residuals Histogram for "Range (m)" analysis (case I a). Note there is a deviation from the Gaussian pdf.....	69
FIGURE 5.10	Residuals Probability Plot for "Range (m)" analysis (case I a).....	69
FIGURE 5.11	Residuals vs. Fitted Range for case IV a.....	70
FIGURE 5.12	Residuals Histogram for case IV a.....	70
FIGURE 5.13	Residuals vs. Fitted log(range) for case IV b.....	71
FIGURE 5.14	Residuals vs. Fitted Range (transformed to Linear Domain) for case IV b.....	71
FIGURE 5.15	Residuals Histogram for Case IV b.....	72
FIGURE 5.16	Residuals vs. Fitted Range for case VI a.....	72
FIGURE 5.17	Residuals Histogram for case VI a.....	73
FIGURE 5.18	Residuals vs. Fitted Range for case VI b.....	73
FIGURE 5.19	Residuals transformed to linear domain (case VI b).....	74
FIGURE 5.20	Residuals Histogram for case VI b.....	74
FIGURE 5.21	Residuals vs. Fitted Range for case VII a.....	75
FIGURE 5.22	Residuals Histogram for case VII a.....	75
FIGURE 5.23	Residuals vs. Fitted log(range) for case VII b.....	76
FIGURE 5.24	Residuals vs. Range transformed to the linear domain for case VII b.....	76
FIGURE 5.25	Residuals Histogram for case VII b.....	77
FIGURE 5.26	Residuals Probability plot for "log (range)" analysis (case VII b).....	77

List of Abbreviations and Acronyms

AMPS	Advanced Mobile Phone Systems
Adj	Adjusted
BNR	Bell Northern Research
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CW	Continuous Wave
dB	Decibel
dBm	dB milli-watts
FDMA	Frequency Division Multiple Access
ft	Feet
GHz	Giga Hertz = 10^9 Hz
GSM	Group Speciale Mobile
HCMT	High Capacity Mobile Telephone System
Hz	Hertz
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
KHz	Kilo Hertz = 10^3 Hz
K-S	Kolmogorov-Smirnov
LOS	Line of Sight
MHz	Mega Hertz = 10^6 Hz
m	Metre
ms	Milli-second = 10^{-3} sec
mw	Milli-Watt = 10^{-3} Watt
NMT	Nordic Mobile Telephone
ns	Nano-second = 10^{-9} sec
OBS	Obstructed
PBX	Private Branch Exchange

PCN	Personal Communication Network
pdf	Probability density function
RF	Radio Frequency
RMS	Root Mean Square
RSS	Received Signal Strength
Rx	Receiver
sec	Second
TDMA	Time Division Multiple Access
TACS	Total Access Communication Systems
Tx	Transmitter

List Of Symbols

B_d	Doppler spread of the channel
D_N	Maximum distance between the Theoretical and the Experimental CDFs: used in the K-S test
d	Distance between the transmitter and the receiver
$E[]$	Expectation Operator
$F_0(x)$	Theoretical Cumulative Distribution Function
$S_N(x)$	Experimental Cumulative Distribution Function
f_c	Carrier Frequency
$(\Delta f)_c$	Coherence bandwidth
g_r	Receive Antenna gain
g_t	Transmit Antenna gain
$h(r;\tau)$	Channel Impulse Response
$Loss$	Path Loss due to distance only, defined as $P_{Tx} - P_{mean}$
$\ln()$	Natural Logarithm
$\log()$	Log base 10
$N(\tau)$	Number of Multipath Components
N	number of points
n	Distance Loss exponent; also used as an integer in the simulation of Normalized power in presence of Doppler and Rayleigh fading
$n(t)$	Complex valued Additive Gaussian Noise
$Prob$	Probability
P_0	Free space received power at unity distance
$P(d)$	Received power at distance d
P_{mean}	Mean (local) power
P_{Rx}	Received power
P_{Shadow}	Shadow component, defined as: $P_{Rx} - P_{mean}$
P_{Tx}	Transmitted Power

$p(r)$	Probability Density Function (pdf) of r
R	Range
R_b	Break-even Range
r	Random Variable (mainly Amplitude)
$s(t)$	Transmitted Signal
T_m	Multipath spread of the channel
t	Observation Time
$(\Delta t)_c$	Coherence time
v	Velocity of the mobile
X	Matrix of the regression variables (predictors)
$x_i(t)$	In-phase component used in the simulation of Normalized power in presence of Doppler and Rayleigh fading
$x_q(t)$	Quadrature-phase component used in the simulation of Normalized power in presence of Doppler and Rayleigh fading
$x(t)$	The envelope of the simulated signal in presence of Doppler and Rayleigh fading
Y	Vector of observations (response), in Regression Analysis
$y(t)$	Received signal
α	Level of significance (K-S test)
$a_k(t)$	Randomly varying Amplitude
B	Vector of the regression coefficients
$\hat{B}$	Least Squares estimate of B
β	regression coefficient
Γ	Parameter of the pdf
$\Gamma(m)$	Gamma Function
E	Vector of the random errors associated with each observation (Regression Analysis)
$\epsilon_{N,\alpha}$	Critical value for the K-S test
$\theta_k(t)$	Randomly varying Phase
Λ	Cluster Arrival Rate
λ	Wavelength

λ_p	Ray Arrival Rate
$\lambda_0(t)$	Mean Arrival Rate of paths
π	Constant 3.141592.....
σ	Standard Deviation of P_{Shadow}
τ	Application time of the impulse
$\tau_k(t)$	Randomly varying Arrival time
χ^2	Chi-squared (test)
ω_m	Doppler Frequency shift

List of Appendices

APPENDIX A	DERIVATION OF THE CUMULATIVE DISTRIBUTION FUNCTION FOR A NORMALIZED RAYLEIGH VARIABLE	91
APPENDIX B	SOURCE CODE FOR COMPUTATION OF THE CUMULATIVE DENSITY FUNCTION	93
APPENDIX C	SOURCE CODE FOR THE KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST	94
APPENDIX D	SOURCE CODE FOR SIMULATING THE NORMALIZED POWER IN PRESENCE OF DOPPLER AND RAYLEIGH FADING	96

Chapter 1

INTRODUCTION

There has been a phenomenal growth in the demand for wireless communications services during the past decade and the trend is expected to continue. Cordless telephones, alpha-numeric pagers, cellular systems are becoming common and the demand for such services are increasing rapidly. It is expected that the number of cellular users would reach 22 million by 1995 in the US and 50 million by the year 2000 in Europe [1].

This growth in demand for wireless services has motivated system designers to improve on the first generation technology of the 1980s which was characterized mainly by analogue voice transmission. The cordless telephones and most of the current cellular systems use a single channel for speech and network control messages (e.g. handoff, power control), thereby limiting the capacity of the total system [1]. Examples include AMPS (Advanced Mobile Phone Systems) in North America, TACS (Total Access Communication Systems) in the UK, NMT (Nordic Mobile Telephone) in the Scandinavian countries and HCMTS (High Capacity Mobile Telephone Systems) in Japan.

Recently, new wireless systems were introduced commercially in order to meet the requirement for higher capacity, improved quality and increased ranges [2]. The second generation wireless systems carry digital speech and have enhanced capability of exchanging signalling and control information within the network. Second generation vehicular mobile radio systems, such as the fully digital pan-European GSM (Group Special Mobile) and the dual mode North American IS-54 standard, have dedicated channels for exchange of control information between mobile terminals and the remainder of the network.

The continuing demand for wireless communication services has prompted manufacturers and service providers to invest in the development of third generation wireless systems. Digital cordless telephony has already enabled more sophisticated applications

like wireless Private Branch Exchanges (PBXs) and public-access Telepoint systems [3], [4]. This opens up the vision of Personal Communication Network (PCN) which has now become the focus of enormous effort on the part of the telecommunications industry [5]. PCN is expected to integrate several communication systems into one interconnected network, with the capability of supporting several wireless devices and services [5]. Such a large and diverse service cannot be supported by a single vendor (service provider) and thus all the interfaces of the sub-networks have to be standardized to allow for the interconnection and inter-working of the diverse sub-networks. PCN is expected to enable subscribers to initiate and receive calls (voice or any other information like data, fax, video) anywhere in the world by maintaining an unique identity throughout the network. The video transmission requires large bandwidth for transmission and thus would probably be routed via the wireline network to a video terminal. Thus instead of identifying places (i.e. individual fixed telephones), the network is expected to identify people (i.e. portable sets) [6]. The third most important feature that PCN is expected to provide will be wireline voice quality and data communications through a small pocket size, lightweight, low power, personal communicator.

The small size of the envisioned portable units means much lower transmitted power as compared to the present day vehicular cellular technology. The lower transmit power is also necessary in order to minimize the health hazards¹ posed by high power RF radiations within buildings. Moreover, the low power units have the added advantage of having longer battery lives. These portable radio systems are expected to operate at power levels around 10 milliwatts [5]. Operating at lower powers in turn means a reduction of the cell size (i.e. the coverage area of a single base station). Thus for indoor communications micro-cellular and pico-cellular architectures are being proposed as opposed to the large cells of the vehicular cellular systems. For example, in indoor environments with the transmitted power around 10 dBm, cells with radii 30 to 100 metres are practical, to provide acceptable quality of service. This compares with cell radii of a few tens of kilometres for vehicular cellular radio systems.

1. The health hazards of Electro-Magnetic waves particularly at Radio frequencies have been studied by many researchers. However, there is no consensus among the researchers. There has been a detailed coverage of the various issues regarding this in [7].

The evolution of current technologies that provide wireless access to people on the move is detailed in [3], [8]. As discussed in [5], [6] and [8] a pocket size low power radio system design is widely different from the present days high power vehicular cellular systems. These differences arise due to the differences in transmit power levels, target coverage areas, speed of the mobile, cost of the fixed radio sites etc. [5], [8]. This means that the futuristic hand held personal communicator would co-exist with the vehicular cellular systems available at present. Cox suggests in [5] that a reasonable approach would be to allow the expansion of cellular mobile systems optimized for vehicular communications and to deploy new tetherless access systems optimized to support the low power pocket sized personal communicator. The user identity (much like a credit card number) could be programmed in the personal communicator or in a "smart card" inserted in the set. When entering an automobile this card or the communicator could be inserted into the vehicular system and the user identification transferred to the cellular set. The network will have the intelligence to correctly route calls to the user. This will also require co-ordination of system registration between the vehicular and the low-power tetherless access systems.

For successful operation of an indoor wireless PCN system, determining the cells that fully cover a given area is an important issue. As the cell sizes are small, so the cell count can be large for large spaces (e.g. Shopping Malls, large retail stores etc.). An important parameter to be determined is the cell size and the locations of the base stations. In theory, experiments can be conducted to determine the cell contours for a particular received signal power. However, with the expansion of this service and its widespread use, it would be costly and impractical to perform experimental deployment for every site. Finding an accurate prediction method greatly simplifies the deployment process and allows the portable systems to be cost-competitive with the wired PBX systems.

It is obvious that cell sizes are closely related to the average ranges for a given environment. One solution for the problem at hand would be to predict the average range for the given system in a particular environment (i.e. type of building: malls, retail stores etc.). The work presented in this thesis investigates ways to predict ranges and cell count for an indoor wireless system.

1.1 Thesis Objective

This thesis proposes methods for predicting the ranges for indoor wireless PCN systems. The methods proposed are general methods and as such could be adapted for any channel access scheme. The reason for this is that each receiver has a minimum RF power level for signal detection. Our purpose here is to predict the maximum range of the fixed base station. Thus the methods outlined can be used for TDMA, FDMA as well as CDMA access scheme. Currently these schemes are widely used for personal communication systems.

Although range prediction has always been an important issue, accurate prediction of ranges for indoor systems has been a difficult problem due to the large number of parameters that affect the signal propagation in the indoor environment.

This work is aimed at developing accurate and easy to use methods for range prediction. The motivation behind the work is to develop a model that can be easily used by a non-technical person (probably a sales person), to predict ranges with reasonable degree of accuracy. Thus the inputs to the models should be clear but should describe the site information, as ranges will change with various environments (i.e. different sites).

With this objective in mind, we have investigated the ease of statistical methods to solve the problem. Experiments have also been performed which will be described in subsequent chapters.

For predicting the ranges, two approaches have been proposed. These can be identified as:

- Direct Method
- Regression Analysis

The direct method can be characterized by directly including the slow fading¹ component (Shadow) in the prediction technique. Shadow (in dB) is typically modelled as

1. These relevant terms will be explained in Chapter 2

a Gaussian variable and the standard deviation, σ of this process is estimated.

The second approach that has been considered in this thesis is based on performing Regression Analysis. The independent variables have been chosen so that it best describes the propagation phenomenon. The approach has been made feasible because of the large amount of data that had been collected.

1.2 Thesis Organization

The thesis is organized as follows:

Chapter 2 describes the indoor channel models that have been proposed in the existing literature. The two prediction approaches are described briefly, therein.

Chapter 3 investigates the standard deviation σ obtained in various environments as well as the variation with distance.

Chapter 4 describes the various variables identified, which have a significant impact on the ranges. This chapter presents the mathematical basis of the Regression analysis. Expressions for goodness of fit parameters such as the RMS error, R^2 and adjusted R^2 values are presented. This chapter also describes the various Regression analyses performed (with different combinations of the variables identified). Comparison of the different Regression analysis performed has been carried out based on the goodness of fit (characterized by the RMS error and the Adjusted R^2 value) for each analysis.

Chapter 5 presents the results for eleven Regression analysis models tested. Based on the goodness of fit parameters one of these models is chosen for the prediction process. The chosen model has been tested for several test cases containing sixteen known sites and four unknown sites.

Finally, Chapter 6 summarizes the work presented in the thesis. Conclusions are presented and possible directions for future research work are suggested.

Chapter 2

INDOOR RADIO CHANNEL MODEL AND PREDICTION TECHNIQUES

2.1 Introduction

The performance of any communication system is greatly affected by the transmission media. This is especially true for the wireless systems, where multipath fading, doppler spread, delay spread and interference are some of the major impairments. The inherent nature of the channel cannot be changed to suit our needs. However, we can design our system to match the conditions of a given channel, such that the errors associated with these impairments are minimized. In achieving this objective, it is necessary that the channel be known and characterized so that the transmitter and the receiver be accordingly designed. The above discussion holds true for both the indoor and the outdoor wireless channels. However, as this work is aimed at indoor systems, the indoor channel is reviewed in detail.

In this chapter, a brief overview of the indoor wireless channel is presented. A summary of the prediction techniques studied is also presented. In subsequent chapters we will describe these methods in greater details and present results based on measurements and analysis.

2.2 Impulse Response Profile of the Indoor Channel

The indoor channel is difficult to characterize deterministically as the number of potential reflectors, scatterers and absorbers is large. Moreover, the relative position and orientation of the reflectors, scatterers and absorbers vary considerably as the portables move. Thus in general, the impulse response of the channel can only be modelled statistically. If extremely short radar like pulses, ideally impulses, are transmitted over a

time-varying multipath channel, the received sequence will be a train of pulses, whose number, amplitude and relative delay will be random.

Traditionally the channel is modelled as a time varying filter with a discrete low pass complex response given by: [9], [10], [11]:

$$h(t;\tau) = \sum_{k=1}^{N(\tau)} a_k(t) \delta(\tau - \tau_k(t)) e^{j\theta_k(t)} \quad (2.1)$$

where t is the observation time and τ is the application time of the impulse. $N(\tau)$ is the number of multipaths, $a_k(t)$, $\tau_k(t)$ and $\theta_k(t)$ are randomly time varying amplitude, arrival time and phase respectively. The phase $\theta_k(t)$ is given by:

$$\theta_k(t) = 2\pi \cdot f_c \cdot \tau_k(t) \quad (2.2)$$

where f_c is the carrier frequency. In the continuous domain, the low pass channel is:

$$h(t;\tau) = \alpha(t;\tau) e^{-j2\pi f_c \tau} \quad (2.3)$$

This model is a wide band model which can be used to obtain the received signal $y(t)$:

$$y(t) = \int_{-\infty}^{\infty} s(\alpha) h(t - \alpha) d\alpha + n(t) \quad (2.4)$$

where $s(t)$ is the transmitted signal and $n(t)$ is the low pass complex valued additive Gaussian noise (Figure 2.1).

As the propagation channel is continuously changing, the impulse response exhibits temporal and spatial variations. As the user moves from place to place, a different impulse response is experienced. These variations depend on the environment and the separation between the transmitter and the receiver. There are three types of variations as described in the following ([10], [12]):

- 1) Small-Scale Variations: Impulse response profiles collected at points separated by

Transmitted Signal

Received Signal

(b)

FIGURE 2.1 Schematic representation of the channel. (a) represents the Mathematical Model of the Channel and (b) represents the response of the time-variant multipath channel to narrow pulses.

a few wavelengths are similar as the channel characteristics are not significantly different. This is referred to as Small-Scale variation.

2) **Midscale Variations**: Receivers with the same antenna separation (from the transmitter) but at different locations exhibit this kind of variation. Thus a receiver placed in a room has a widely different impulse response profile than one in the corridor. This is referred to as Midscale variation.

3) **Large-Scale Variations**: As the transmitter receiver distance changes, the impulse response profile undergoes drastic variations. The primary reason is that the number and characteristic of intervening objects changes significantly. Increasing the antenna separation between transmitter and receiver results in an increase in path loss. This path loss is typically characterized by d^n loss, where d is the separation between transmitter and receiver and n is usually a constant for a particular channel.

2.2.1 Characterization of the Impulse Response

Recently, considerable attention has been given by many researchers to the study of the distribution for the path amplitude. The study of the arrival times have also received some attention. Other parameters of interest are path losses, delay spread, Doppler spread, etc. In the following sub-sections, we will review these parameters.

2.2.1.1 Distribution of the path amplitudes

There has been a lot of work to try to characterize the path amplitude and to find out statistics that best describe the distribution of path amplitude. In a multipath environment, amplitude fading may follow different distributions depending upon the area under observation and the presence or absence of a dominant component. We describe below some of the commonly used distributions.

1) **Rayleigh Distribution**: This is a widely used distribution for small-scale rapid amplitude fluctuations in a multi-path environment where there is no single strong received component. The probability density function (pdf), $p(r)$, of the Rayleigh fading is

given by [9], [10], [11], [13]:

$$p(r) = \begin{cases} \frac{r}{\Gamma^2} \exp\left\{-\frac{r^2}{2\Gamma^2}\right\} & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (2.5)$$

where, Γ is the Rayleigh parameter, r is the Rayleigh random variable (amplitude in this case). The mean and variance of this distribution is $\sqrt{\pi/2}\Gamma$ and $(2 - \pi/2)\Gamma^2$ respectively [14], [10]. Theoretically the model can be derived by assuming that the transmitted signal reaches the receiver from N directions such that all these components have equal amplitude and uniformly distributed phase, and adding all these incident waves vectorially. Moreover the Inphase (I) and Quadrature (Q) components are independent and by the central limit theorem are Gaussian. The pdf of $r = \sqrt{I^2 + Q^2}$ can be derived ([14]) and is given in Eq. (2.5). Although the assumption that all the received paths have equal attenuation is unrealistic, it has been shown that the Rayleigh distribution can explain the amplitude fluctuation even if the individual path attenuations are unequal provided that there is no dominant path. Measurements, at various frequencies and environments, show that the small scale non Line of Sight (LOS) fading is predominantly Rayleigh [15], [16], [17].

2) **Rician Distribution**: When there is a strong LOS path in addition to a large number of low power scattered path, the amplitude fading follows a Rician distribution whose pdf, $p(r)$, is given by [9], [10], [11]:

$$p(r) = \begin{cases} \frac{r}{\Gamma^2} \exp\left\{-\frac{r^2 + v^2}{2\Gamma^2}\right\} I_0\left(\frac{rv}{\Gamma^2}\right) & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (2.6)$$

where, I_0 is the zeroth-order modified Bessel function of the first kind, v is the magnitude (envelope) of the strong component and Γ^2 is proportional to the power of the scatter component. As expected the Rician distribution (Eq. (2.6)) becomes Rayleigh (Eq. (2.5))

when the strong path is eliminated (i.e. $v = 0$ or $\frac{v}{2\Gamma^2} \ll \frac{r^2}{2\Gamma^2}$). On the other hand when the power in the dominant path is considerably larger than the combined power in the scattered path, the Rician distribution is well approximated by the gaussian distribution.

Measurements seem to validate the proposition that in presence of a Line of Sight component the amplitude follows a Rician distribution [16], [18].

3) **Lognormal Distribution**: This distribution is often used to explain the variation over the mean value. It has been reported that when the fast fading is averaged out the variations over the local mean followed a lognormal distribution. The pdf, $p(r)$, of a lognormal variable r is given by [10], [11]:

$$p(r) = \begin{cases} \frac{1}{\sqrt{2\pi}\Gamma r} \exp\left\{-\frac{(\ln(r) - \mu)^2}{2\Gamma^2}\right\} & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (2.7)$$

With this distribution $\ln(r)$ has a Gaussian distribution. There have been a lot of measurements which confirm that the mean of the local data is lognormal [10], [19], [20].

4) **Nakagami Distribution**: This distribution (also known as the m -distribution) has been generally neglected in radio propagation literature [10]. In describing the Rayleigh distribution it was assumed that the received multipath signals were of equal amplitude. However if we let the received signals to be of random magnitude, the Nakagami distribution is obtained with the pdf, $p(r)$, given as [1], [10]:

$$p(r) = \begin{cases} \frac{2m^m r^{2m-1}}{\Gamma(m) \Omega^m} \exp\left\{-\frac{mr^2}{\Omega}\right\} & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (2.8)$$

where, $\Gamma(m)$ is the Gamma function, $\Omega = E[r^2]$ and $m = \frac{(E[r^2])^2}{Var[r^2]}$, with the

constraint $m \geq 1/2$. When $m = 1$, the distribution reduces to Rayleigh distribution and

with $m = 1/2$ it becomes an one-sided Gaussian distribution.

There has been little effort to apply this distribution to experimental data. It has been reported in [21] that the Nakagami distribution provided a better fit for the probability distribution function for the measured data, than the Rayleigh distribution for populated and unpopulated environments. It has also been reported in [22] that analytical ray tracing technique showed that the fading component followed the Nakagami distribution.

5) **Suzuki Distribution**: This distribution, first proposed by H.Suzuki, tries to explain the transition from a local to global statistics [12]. The Suzuki distribution is a mixture of the Rayleigh and lognormal statistics, whose pdf, $p(r)$, is given by [12], [10]:

$$p(r) = \int_0^{\infty} \frac{r}{\sigma^2} \exp\left\{-\frac{r^2}{2\sigma^2}\right\} \frac{1}{\sqrt{2\pi}\sigma\lambda} \exp\left\{-\frac{(\ln(r) - \mu)^2}{2\lambda^2}\right\} d\sigma \quad (2.9)$$

which effectively is the integral of the Rayleigh distribution over all values of σ , weighted by the lognormal pdf of σ . This distribution has also been neglected in mobile radio literature probably because of its complexity.

2.2.1.2 Distribution of the arrival times

The arrival times (τ_k in Eq. (2.1)) have not received much attention in literature. The arrival times form a discrete random variable and several distributions that describe it are reviewed here.

1) **Standard Poisson Model**: This is a preliminary model which proposes that the excess delay arrivals (i.e. the delay of the rays after the arrival of the first ray) follow a Poisson distribution [23]. The Poisson distribution is often used to model completely random events like initiation of phone calls, occurrences of automobile accidents etc. [10]. In the indoor channel this hypothesis presumes that the reflectors are arranged in a random fashion. Poisson fit to experimental data has been observed when the receiver sensitivity is low [10]. Several researchers have reported the inadequacy of this hypothesis in different

environments [10], [17], [23], [24].

2) **Modified Poisson—the Δ -K Model**: This model was first suggested by Turin *et al.* and later developed by Suzuki to explain the tendencies of rays to arrive in clusters [23], [10]. There are two states: S_1 and S_2 , with mean arrival rate of paths as $\lambda_0(t)$ and $K\lambda_0(t)$ respectively. If a path arrives at time t , a transition is made to S_2 for the interval $[t, t+\Delta]$ and a failure of a path arrival during this interval resets it back to S_1 . K and Δ are constant parameters of the model. When $K=1$ or $\Delta=0$, the process reverts back to a standard Poisson process. A good fit to experimental data has been observed in [24] with optimum values of K less than 1, indicating that the paths are more evenly distributed than in the standard Poisson model. However, application of this model to data collected in a factory environment has not been successful [19].

3) **The Neyman-Scott Clustering Model**: This model has the physical base in the observation that rays in an indoor environment tend to arrive in clusters. The arrival times are modelled as a two step process. The arrival times of the clusters (i.e. the first ray of each cluster) are modelled as a Poisson process with a fixed rate Λ and within each cluster the rays are modelled to arrive according to another Poisson process with a fixed rate of λ_p [17]. Typically it is expected that $\lambda_p \gg \Lambda$. This model had been originally utilized in cosmology to study the distribution of galaxies [10]. Data collected from a large office building shows a good fit to this double Poisson process [17]. The cluster arrival rate Λ has been estimated as $1/300 \text{ ns}^{-1}$, while the ray arrival rate λ_p has been estimated as $1/5 \text{ ns}^{-1}$ [17].

4) **Piecewise Functions**: Measurements done in a factory environment show that the data exhibit a poor fit to the Poisson models presented earlier [19]. Instead the researchers [19] computed the empirical probability of path arrival by counting the total number of detected multipath components at a particular excess delay time and dividing by the total number of possible multipath components for each excess delay interval. The authors have modelled these probabilities for multipath arrivals as three piecewise functions of excess delay for the LOS topography and as two functions for the Obstructed (OBS) situation.

2.2.1.3 Other Channel Parameters

1) Distribution of the Phases: The performance of any digital indoor communication system would be strongly dependent on the phase sequence ($\theta_k(t)$ in Eq. (2.1)). Although the issue is important, there has not been any model, backed up by experiments, for the phase sequence. The probable reason for this is the difficulty of measuring the phases of each of the received paths [10]. One important aspect about the phase is that it changes by 2π as the path length changes by a wavelength. For the frequency of interest (e.g. 1 GHz) this distance is 30 cm. At higher frequencies this distance is even shorter. The phase is thus assumed to be Uniformly distributed [10]. Researchers have considered this distribution without experimental verification [17].

2) Number of Multipath Components: One of the important parameters for analysing the performance of various types of diversity, modulation and equalization techniques (e.g. RAKE receiver [9]) is the number of multipath components received. Rappaport *et. al.* studied a statistical model for the number of rays arriving at the receiver in a factory and open plan building [19]. For this environment it has been reported that the number of multipath component is normally distributed with a random mean [19]. The mean is found to be uniformly distributed between 9 and 35 for LOS topographies and between 11 and 36 for OBS topographies. The standard deviation is linearly related to the mean.

3) Large Scale Path Loss: The impulse response approach characterizes the channel at a microscopic level, while path loss models describe it at a macroscopic level. Knowledge about the path loss in indoor environment is important for determining the cell sizes and thus the coverage area. Extensive work has been carried out in this area by many researchers [16], [17], [18], [25], [26], [27]. A much more comprehensive discussion and reference list is given in [10]. The most widely used is the inverse exponent model, where the received power $P(d)$, at a distance of d metres from the transmitting antenna is given as:

$$P(d) = P_0 \cdot d^{-n} \quad (2.10)$$

where P_0 is the free space received power at unity distance ($d=1$) and n is dependent on

the environment. P_0 is given by [13]:

$$P_0 = P_t \cdot \left(\frac{\lambda}{4\pi} \right)^2 \cdot g_t \cdot g_r \quad (2.11)$$

where, P_t is the transmitted power, λ is the wavelength, g_t and g_r are the transmit and receive antenna gains respectively. As P_0 is constant for a given system, the propagation is proportional to d^n . On a logarithmic scale this translates to a straight line, whose slope is proportional to n . The measured values for the exponent differs widely. For the LOS case values of 1.55 and 2.15 have been reported for 1.75 GHz and 37.2 GHz respectively, while for the Non-LOS path these values are 2.7 and 3.5 respectively [27]. The environment considered for the above work was university. A refinement of this is made by including the information about the wall loss [18], [27]. Wall loss values have been reported to vary between 1.9 to 5.0 dB for 1.75 GHz while for 37.2 GHz, it has been to vary between 5.2 to 16.2 dB [27]. Measurements at 60 GHz show that the loss for a wooden door is 7 dB, while that for a concrete wall is 27 dB [18].

2.2.2 Correlation and Power Spectral Density Functions of the Impulse Response

In this sub-section a number of correlation and power spectral density functions that define the characteristics of the fading multipath channel are described. The continuous domain lowpass equivalent impulse response function (Eq. (2.3)) is considered to be complex-valued zero mean Gaussian random process in the t variable [9]. Assuming that the impulse response, $h(t;\tau)$ is wide sense stationary, the auto-correlation function of $h(t;\tau)$, $\phi_h(\Delta t;\tau_1, \tau_2)$ is defined as [9], [14]:

$$\phi_h(\Delta t;\tau_1, \tau_2) = \frac{1}{2} E [h^*(t;\tau_1) \cdot h(t + \Delta t;\tau_2)] \quad (2.12)$$

where $E[\cdot]$ is the Expectation operator and the factor $\frac{1}{2}$ appears because $h()$ is complex-valued. For uncorrelated path delays, the Eq. (2.12) is modified to:

$$\frac{1}{2} E [h^*(t; \tau_1) \cdot h(t + \Delta t; \tau_2)] = \phi_h(\Delta t; \tau_1) \cdot \delta(\tau_1 - \tau_2) \quad (2.13)$$

Putting $\Delta t = 0$, the resulting auto-correlation function $\phi_h(\tau)$, represents the average output power as a function of the time delay τ and is called the multipath intensity profile or the delay power spectrum of the channel [9]. The range of values of τ , for which $\phi_h(\tau)$ is non-zero is called the multipath spread of the channel and is denoted by T_m . The autocorrelation function of $H(t; f)$ (which is the Fourier Transform of $h(t; \tau)$), denoted by $\phi_H(\Delta t; \Delta f)$, is called the spaced-time spaced-frequency correlation function of the channel [1], [9]. Putting $\Delta t = 0$ the auto-correlation function in the frequency variable $\phi_H(\Delta f)$ can be shown to be the Fourier Transform of the multipath intensity profile $\phi_h(\tau)$. As a result of this Fourier transform relationship, the width, $(\Delta f)_c$, of $\phi_H(\Delta f)$ is reciprocal of the multipath spread of the channel (T_m) and is called the coherence bandwidth of the channel. Thus two signals, whose frequency separation is greater than $(\Delta f)_c$, would be affected differently by the channel.

Based on the relationship between the Coherence Bandwidth and the signal bandwidth, the channel can be classified into two categories. If the signal bandwidth is large compared to the Coherence Bandwidth, the channel is said to be frequency-selective. In this case, the signal experiences severe distortions. If the signal bandwidth is small compared to the Coherence Bandwidth, the channel is called frequency-nonselective and the signal is undistorted, although the amplitude may fade, due to the channel.

The temporal variations of the channel are manifested as the Doppler effects. These effects are studied by considering the Fourier Transform of $\phi_H(\Delta t; \Delta f)$ (the spaced-frequency spaced-time correlation function.), with respect to the variable Δt and denoting this by $S_H(\lambda; \Delta f)$. Putting $\Delta f = 0$, the Doppler power spectrum of the channel $S_H(\lambda)$ is obtained. The range of values of λ for which the Doppler power spectrum is non-zero is called the Doppler Spread of the channel and is denoted by B_d [9]. The reciprocal of the Doppler Spread, $(\Delta t)_c$, is known as the coherence time of the channel. The Doppler spread of a channel is directly proportional to the speed of the mobile [13]. For indoor channels, with the carrier frequency of 1 to 2 GHz, Doppler spread is not a serious issue as the

mobile users are not fast moving. For outdoor systems, especially vehicular systems, Doppler is an important issue. However, at higher frequencies (e.g. 60 GHz) the Doppler effects can become significant even for indoor systems. The relationships between all these functions discussed in this sub-section are illustrated in Figure 2.2.

2.3 Prediction Techniques

This thesis studies techniques to carry out range prediction for indoor wireless systems. In this section the two techniques studied are briefly introduced. The two methods investigated are the Direct Method and the Regression Analysis method.

2.3.1 Direct Prediction Method

In this approach, the system design can be carried out by directly including the shadow component in the prediction technique. The shadow component, P_{shadow} is defined as:

$$P_{shadow} = P_{Rx} - P_{mean} \quad (2.14)$$

where P_{Rx} is the received power and P_{mean} is the mean (local) power. The underlying assumption for Eq. (2.14) is that the fast fading has been averaged out. The term P_{mean} is assumed to have been caused only by path loss (i.e. due to distance) and can be written as:

$$P_{mean} = P_{Tx} - Loss \quad (2.15)$$

where, P_{Tx} is the transmitted power and $Loss$ is the path loss (caused solely by the distance between the transmit and receive antennas).

FIGURE 2.2 Correlation and Power Spectral Density Functions of the Channel Impulse Response. (a) shows the delay power spectrum, the spaced-frequency correlation function, the coherence bandwidth and the multipath spread of the channel. (b) shows the spaced-time correlation function, the Doppler power spectrum, the coherence time and the Doppler spread of the channel.

As discussed in § 2.2.1.3 the term *Loss* can be determined from Eq. (2.10) and Eq. (2.11) and thus mean (local) power P_{mean} can be determined from Eq. (2.15). The shadow component P_{shadow} is log-normally distributed, i.e. expressed in dB s, the shadow is Gaussian. The outage rate for any designed system can be determined if the standard deviation σ for the distribution is known or conversely a system can be designed for any specified outage rate. It would be interesting to investigate if σ varies with the distance between the transmitter and receiver. In Chapter 3 we have investigated this.

For designing a system with a certain outage probability, p , we have to ensure that:

$$Prob\{P_{RA} \leq Threshold\} \leq p \quad (2.16)$$

which reduces to (using Eq. (2.14)):

$$Prob\{P_{shadow} \leq Threshold - P_{mean}\} \leq p \quad (2.17)$$

The maximum distance between the transmitter and receiver that satisfies Eq. (2.17) would be the range for the wireless system. Solving this equation would be an iterative process as the σ (and thus P_{shadow}) as well as the mean (local) power, P_{mean} , depend on distance. The important requirement in this kind of design would be to know the σ accurately. Of course σ would vary with different kinds of environment (e.g. office, laboratory, cafeteria etc.).

2.3.2 Prediction using Regression Analysis

The other method for range prediction would be to perform a Regression analysis based on experimental data collected. This appears to be more pragmatic and easier to perform. For a large amount of experimental data, Regression analysis would yield consistent and accurate results. In Chapter 4, a brief discussion on Regression analysis is presented. For utilizing this method, variables (both discrete and continuous), that affect radio propagation, have to be identified. This technique treats the range prediction issue as a statistical problem without going into the physics behind the propagation phenomenon. The success of the method depends on the data collected and the variables used in the regression analysis. The process of identifying the variables accounts for the propagation

phenomenon and the effects that differing environments have on radio transmission, albeit in an indirect way.

In Chapter 4, the Regression analysis method and the various variables identified have been described. The data collected and the various Regression analysis performed are also described.

2.4 Summary

In this chapter, the efforts of various researchers in characterizing the indoor channel have been reviewed. A mathematical model has been presented and the various parameters that characterize the channel have been discussed. As reported by Hashemi [10], different values have been suggested in the literature for these channel parameters. The variations in the parameter values reported may be attributed to insufficient data samples or inaccurate measurements. Finally two methods of prediction techniques have been introduced and these are studied in greater details in subsequent chapters.

Chapter 3

VARIATION OF σ WITH DISTANCE

3.1 Introduction

In this chapter we will look into the variation of the standard deviation of the log-normal shadow, with distance. As discussed in the last chapter, σ can be utilized to predict the coverage range along the lines explained in Chapter 2. Experiments have been conducted to determine σ at various separation between the transmitter and receiver. In subsequent sections we will describe the experimental setup utilized, data processing technique and results obtained. We have considered two frequencies in the work presented in this chapter: 2 GHz and the 900 MHz band. We will describe the experiment and the data collection method for both these frequencies. Each of the sections in this chapter examines the two different frequency bands and discuss issues particular to the respective frequency band. This chapter also contains a section on the Kolmogorov-Smirnov goodness of fit test. We will also apply this test to test the hypothesis that the shadow fading is log-normal in nature.

3.2 Experimental Setup

This section has two subsections that describe the experimental set up for the two frequencies considered: 2.0 GHz and the 900 MHz band.

3.2.1 2.0 GHz

The experiments were carried out by transmitting a continuous wave at a frequency of 2 GHz. The block diagram of the setup is shown in Figure 3.1. The 2 GHz source shown was implemented in two different ways. In the first method, a Qualcomm 1 GHz RF source was connected in series to a frequency doubler, while in the second method the 2 GHz signal had been generated from the Hewlett-Packard Network Analyser

(HP 8753 C). The transmit antenna is omnidirectional. The receiver is a four channel coherent receiver designed by Todd [28]. The receiver uses four omnidirectional receive antennae spaced throughout the coverage area. The power input to the transmit antenna is -8 dBm. The absolute value of the receive power (and hence the transmit power) is not important as we are concerned with the attenuation and not the actual power.

Keeping the receive antennas fixed the transmitter is pushed along a straight path at a constant speed (Figure 3.2) The data is logged by a computer connected to the receiver. The data post processing technique is described in § 3.3.1.

3.2.2 The 900 MHz Band

The experiments in this frequency band have been carried out using the Companion Deployment Tool (CDT)¹, from Nortel. The CDT is a modification of Nortel's wireless PBX Companion system basestation so that the power received by the base station from the portable transmitter can be measured.

The CDT consists of a modified base station and a portable handset so that the measured value of the received signal power (at the base station) is transmitted back to the portable and can be read off the portable display panel. For the measurements in this study a one second time average of the instantaneous received power was employed so as to remove the fast fading. The specification of the system for CT2 Plus protocol is given below:

- i) Transmit Power for Base Station and Portable — 7 to 10 dBm.
- ii) Individual Channel bandwidth — 100 KHz.
- iii) Frequency band used — 944 MHz to 948.5 MHz

Measurements are done at various antenna separation (Figure 3.3). The small scale spatial variations for each of the locations are averaged out by moving the portable along the arc shown in the Figure 3.3. Moving along the arc ensures that the distance between

1. TM of Nortel (Nortel Telecom)

FIGURE 3.1 The experimental setup for the continuous wave Measurements at 2 GHz.

FIGURE 3.2 Schematic of the experimental procedure for 2.0 GHz.

the base station and the portable remains constant for a given location. The length of the arc is small as compared to the distance between the base station and the portable. At each location a recording is made for about 3 minutes, which leads to about 200 values for the received power. A computer is connected to the CDT and the received power measured by the basestation can be logged. The data post processing procedure is explained in § 3.3.2.

3.3 Data Post Processing

The basic objective for both frequency bands is to determine the shadow component as discussed in Chapter 2 (Eq. (2.14)). As the experimental procedure is different for these cases, the data post processing also is different for these two frequency bands. In the two subsections we discuss the data processing technique for the two frequency bands under consideration.

3.3.1 2.0 GHz

The data sampling rate in this case is 1000 samples per second. As the transmitter moves along the designated path, instantaneous values of the received power are recorded. The post processing is a two step process. In the first step, the data is averaged locally to remove the fast fading, which typically manifests itself as a Rayleigh fading. The averaging is done by sliding a window of width 2000 samples (i.e. 2 sec) over the collected data. Assuming that the transmitter moves at a speed of 2 m/s, a 2 second window translates to a window of about 27 times the wavelength (at 2 GHz). In the second step the path loss associated with the data is accounted for by fitting a straight line through the data. As the received power is expressed as a dB value, this translates to a d^n loss estimate. The difference of the averaged data (obtained after the first step) from this fitted line is the shadow component (Eq. (2.14), Eq. (2.15)).

The shadow component corresponding to a predetermined separation between the transmit and the receive antenna is collected for the different paths traversed and the standard deviation of the ensemble is estimated. The process is repeated for various distance between the transmit and receive antennas. This is used to investigate the variation

FIGURE 3.3 Schematic of the experimental set up for the 900 MHz band

of the standard deviation with distance. The σ obtained, is plotted against the respective distance separation and a best fit curve is obtained.

3.3.2 The 900 MHz Band

The underlying principle of the data processing is the same. However, there are slight differences in the data post processing to account for the different equipment used in the two cases. In this case the experiments have been conducted using the CDT as described earlier (§ 3.2.2). As the CDT removes the fast fading by providing the averaged value, the first step as outlined in the last section is not required. The mean for a location is calculated. The mean is computed by transforming the individual dB values into linear units and then converting back the mean into dB. The deviation of the individual measured values from this computed mean gives the variation about the mean. The standard deviation of the variation for each location is computed and thus the distance dependence can be investigated.

3.4 Indoor Environment

We have studied two major kinds of indoor environments: university and office. However, these two different environment have been studied with the different frequency bands. Carleton University served as the university environment, while Bell Northern Research (Carling , Ottawa), served as the office environment. These two environments are different and probably represent two of the typical environments that would be encountered when PCN becomes widely available.

3.5 Statistical Testing of Experimental Distribution

One of the important question that arises when comparing a derived (experimental) distribution with a theoretical distribution, is the goodness of fit. Extensive work has been done to find non-parametric goodness of fit tests. The most well known of these is the χ^2 test, which has been widely used. However, in this work we consider another simple test called the Kolmogorov-Smirnov (KS) goodness of fit test, which is described in

the following subsections.

3.5.1 The Kolmogorov-Smirnov goodness of fit test

The KS test is a better all round test than the χ^2 test [29]. This test is applicable for cases when the cumulative distribution function is continuous. The test computes the maximum “distance” between a theoretical distribution $F_0(x)$ and the experimental distribution $S_N(x)$ (i.e. $S_N(x) = k/N$, where k is the number of data points less than or equal to x and N is the total number of data points in the sample) and then compares this value with a critical value to determine whether the experimental distribution passes the test. The maximum distance, D_N , is defined as [20], [29]:

$$D_N = \max_x (|F_0(x) - S_N(x)|) \quad (3.1)$$

where $F_0(x)$ and $S_N(x)$ are as explained. The critical value, $\epsilon_{N,\alpha}$ is defined such that:

$$Prob[D_N \leq \epsilon_{N,\alpha}] = 1 - \alpha \quad (3.2)$$

where α is the level of significance. For an experimental cumulative distribution to pass the test, it must satisfy $D_N \leq \epsilon_{N,\alpha}$. For large N , the critical value $\epsilon_{N,\alpha}$ is given by:

$$\epsilon_{N,\alpha} = \sqrt{-\frac{1}{2N} \ln\left(\frac{\alpha}{2}\right)} \quad (3.3)$$

3.5.1.1 Illustration of the K-S test with an example

We will illustrate the test considering an example. Let us consider a way to simulate a Rayleigh fading characteristics. The technique uses N_0+1 oscillators to generate the in-phase and quadrature-phase signals, which are then combined to get the simulated Rayleigh envelope. The in-phase component, $x_i(t)$ and the quadrature-phase component, $x_q(t)$ are written as [13]:

$$x_i(t) = 2 \sum_{n=1}^{N_0} \cos \beta_n \cos(\omega_n t) + \sqrt{2} \cos \alpha \cos(\omega_m t) \quad (3.4)$$

$$r_q(t) = 2 \sum_{n=1}^{N_0} \sin \beta_n \cos(\omega_n t) + \sqrt{2} \sin \alpha \cos(\omega_m t) \quad (3.5)$$

where, ω_m is the Doppler frequency shift ($=2\pi v/\lambda$, v is the speed of the mobile and λ is the wavelength of the carrier), $\omega_n = \omega_m \cos(2\pi n/N)$ and $N = 2(2N_0 + 1)$. The envelope, $x(t)$ of the received signal is given by:

$$x(t) = \sqrt{x_i^2(t) + x_q^2(t)} \quad (3.6)$$

Putting $\alpha = 0$ and $\beta_n = \pi n/(N_0 + 1)$ in Eq. (3.4), Eq. (3.5) we get the Rayleigh distributed envelope.

We will illustrate the K-S test by comparing the Cumulative Distribution Function (CDF) obtained through this simulation process with the theoretical CDF of a normalized Rayleigh variable. The CDF, $F_\rho(x) = \text{Prob}[\rho \leq x]$, of a normalized Rayleigh variable, ρ , expressed as dB is given in Eq. (3.7). The derivation can be found in the Appendix A.

$$F_\rho(x) = 1 - \exp\left[-10^{\frac{x}{10}}\right] \quad (3.7)$$

To apply the test we generate our simulated data using Eq. (3.4), Eq. (3.5), Eq. (3.6) and compute the simulated CDF. The code for simulating the received normalized power in presence of Doppler and Rayleigh fading can be found in Appendix D. The code for computing the CDF is included in Appendix B. Appendix C lists the source code for the Kolmogorov-Smirnov goodness of fit test.

Figure 3.4 shows the two cumulative distributions, where the parameters for the simulated CDF is shown in the figure. The range of values for the simulated normalized power was -30 dB to 10 dB. The entire range is not shown to magnify the effects, as in the range displayed the deviation is maximum. The dotted lines in the figure represents a band of $\epsilon_{N,\alpha}$ drawn around the theoretical CDF. After carrying out the K-S test for these two CDFs we conclude that the simulated CDF passes the test with a level of significance of 0.05, as it lies fully within the band of $\epsilon_{N,\alpha}$.

FIGURE 3.4 Theoretical (broken line) and Simulated (solid line) Cumulative Distribution Function. The Simulated CDF passes the Kolmogorov-Smirnov test as it lies within a band of $\epsilon_{N,\alpha}$ of the Theoretical CDF.

We will apply this test to test the hypothesis that the shadow component at various distances between the transmit and receive antennas, is log-normal.

3.6 Results

In this section we will discuss some of the results obtained, for the distance dependance of the standard deviation of the shadow, σ . We will present the results for the two frequency bands studied, in two sub-sections.

3.6.1 University Environment

As explained in § 3.2.1 the experiments for this environment was carried out by transmitting a continuous wave at 2 GHz. Several distinct locations in Carleton University were studied. They included the Minto Case Building, the Mackenzie Building, the Dunton Tower and the Library. The experiments in the Dunton Tower considered inter-floor propagation, where receive antennae were placed in two floors.

It has been observed that the standard deviation σ is constant for university environment. The case of the Minto Case building is interesting to note (Figure 3.5). We observe that the σ remains constant up to about 19 metres and then starts to decrease. However, the decrease is minimal as compared to the rate of increase for office environment (§ 3.6.2). This may be explained by the fact that in this case the larger distances correspond to paths which include intervening structural walls. Although the distance dependance part has been removed (as explained in § 3.3.1) the effect of walls might mean small variations over the mean, which in turn means smaller σ . For the Mackenzie Building (Figure 3.6), σ is almost constant and is approximately equal to 2.5 dB. In the figure, we have also shown the plot excluding the point A. Due to the less number of locations considered, we are unable to comment whether the point A is an outlier. For the library (Figure 3.7) also, σ is almost constant and approximately equals 3 dB. Measurements in the Dunton Tower enabled us to investigate the effects on σ due to inter-floor propagation in a modern high rise building. σ for this case is approximately constant (with a value of 2 dB), when the two points A and B (in Figure 3.8) are removed. In this case also we are not

in a position to comment whether these two points are statistically outliers.

One common observation for all these sites is that the σ is almost constant. The variation of σ , if any, is observed to be minimal.

We have also investigated the Cumulative Distribution Function of the shadow component for these distances in the buildings considered earlier. The experimental CDFs have been compared with the theoretical Gaussian CDF using the KS test, as discussed in § 3.5. As the shadow is expressed in dB, so a Gaussian distribution translates to a log-normal distribution for the power in absolute terms. The results are summarized in Table 3.1

TABLE 3.1. Results of the KS test for the CDFs for various distance separation.

Location	Level of Significance α	No. of different distance separations considered	No. of CDFs that pass the KS test
Mackenzie	0.05	5	5
Minto Case	0.05	11	11
Library	0.05	7	6
Dunton Tower	0.05	8	8
All Locations	0.05	31	30

Out of the four locations studied at Carleton, we observe that the σ is highest for the library. The library is characterized by open spaces with bookshelves, exhibiting high losses. It was found in [27] that the bookshelves exhibited the highest losses for the 1.75 GHz band. Moreover, at this frequency the propagation exponent was also highest for the library and the bookshelves [27].

FIGURE 3.5 Variation of σ with distance for the Minto Case Building in Carleton University.

FIGURE 3.6 Variation of σ with distance for the Mackenzie Building in Carleton University.

FIGURE 3.7 Variation of σ with distance for the Library in Carleton University.

FIGURE 3.8 Variation of σ with distance for the Dunton Tower in Carleton University.

As the standard deviation of the log-normal shadow, σ , is observed not to vary with distance, we can examine the behaviour of the shadow over the entire path (§ 3.2.1 and Figure 3.2). This approach provides us with a subjective picture about the whole path as a whole and not the effects at a particular distance. Figure 3.9 compares the experimental CDF with the theoretical CDF for the library for each of the four channels. Ideally the plots should be a straight line with slope -45° . From the figure, we observe that except for some of the Antenna 2 measurements, the other measurements yield shadows that are approximately log-normal. Five separate paths have been considered and each path has been repeated a number of times.

Figure 3.10 shows the corresponding plots for the Minto Case Building. In this case, three receive antennae were used. Six separate paths were considered and each path was repeated seven times.

Figure 3.11 shows the corresponding plots for the Mackenzie Building. In this case four receive antennae were used. Seven separate paths were considered and each path was repeated six times.

Figure 3.12 shows the corresponding plots for the Dunton Tower. In this case four receive antennae were used and were placed in two floors. Six separate paths were considered and each path was repeated a number of times.

The theoretical CDFs in all these plots have been calculated using σ that has been estimated using the experimental lognormal fading values for the entire path. The results presented in Table 3.1 are computed using the estimated σ for a fixed separation between the transmitter and the receiver (§ 3.3.1). Figure 3.9 through Figure 3.12 show that the experimental CDF is close to the theoretical CDF, which is a subjective confirmation of our hypothesis that σ remains constant with distance for the University environment.

FIGURE 3.9 Experimental vs. Theoretical Cumulative Distribution Function for the Library. The plots for the four receive antennae are shown separately.

FIGURE 3.10 Experimental vs. Theoretical Cumulative Distribution Function for the Minto Case Building. The plots for the three receive antennae are shown separately.

FIGURE 3.11 Experimental vs. Theoretical Cumulative Distribution Function for the Mackenzie Building. The plot for the four receive antennae are shown separately.

FIGURE 3.12 Experimental vs. Theoretical Cumulative Distribution Function for the Dunton Tower. The plots for the four receive antennae are shown separately.

3.6.2 Office Environment

As explained earlier (§ 3.2.2) the experiments for this environment was conducted using a CT2 Plus system. Two different buildings were considered: BNR (Carling), Lab 3 and BNR (Carling), Administration building. The former is an open plan office while the latter is a closed office. The data post processing technique has already been explained in § 3.3.2. The standard deviation, σ , is estimated for each of the distance separation between the base station and the mobile. It is found that a straight line describes the variation of σ quite accurately

Figure 3.13 shows the results for the BNR (Carling), Lab3. The site is an open plan office with partitions. It is found that σ increases rapidly with distance. The σ at each distance separation has been obtained by combining the data collected for that distance at several locations.

Figure 3.14 shows the variations in BNR (Carling) Administration building. The site is a closed office. The experiments were conducted for two different positions of the base station. The results for these two locations are shown in the two plots in Figure 3.14. For this building also, the σ increases linearly with distance. Figure 3.15 is the plot obtained by combining the results of these two locations in the BNR (Carling) Administration Building. we observe that the rate of increase of σ is smaller in this case than in BNR (Carling), Lab 3. (Figure 3.13). This may be due to the fact that an open plan office has more reflectors, scattered in a random manner than a closed office environment.

3.7 Summary

In this chapter we have discussed the experimental procedure and the data processing technique for the two bands investigated: the 900 MHz band and the 2.0 GHz band. We have also discussed the Kolmogorov-Smirnov goodness of fit test and provided an illustration to explain the test. We observe that for university environment σ is almost constant, while for the office environment, σ increases linearly with distance. We also note that the rate of increase for closed office is smaller than that for open plan office.

FIGURE 3.13 Variation of σ with distance in BNR (Carling) Lab 3. Each point on the graph represents several location with the distance being the same.

FIGURE 3.14 Variation of σ with distance for BNR (Carling), Administration Building. The two plots are for two different positions of the base station.

FIGURE 3.15 Variation of σ with distance for BNR (Carling), Administration Building. This is the composite plot i.e. combining the two graphs in Figure 3.14.

Chapter 4

REGRESSION ANALYSES FOR RANGE PREDICTION

4.1 Introduction

As already noted the range of a radio system depends on a wide variety of factors. Qualitatively we are able to define these variables, but finding the exact analytical dependence with the ranges is difficult. If we have a large number of (experimental) data available at our disposal, then a pragmatic solution to the problem would be performing Regression Analysis. The coefficients obtained could then be used to predict ranges for other sites. This is the second of the two techniques introduced in Chapter 2 (§ 2.3).

For carrying out a Regression Analysis, we have to identify the independent variables, also called the predictor variables. We can predict the range or any function of the range in terms of these predictor variables.

In this chapter we will describe the equipment and the method used to collect the data. We will briefly introduce the mathematical fundamentals of Regression Analysis and will briefly discuss the variables that have a pronounced effect on the radio propagation. We will discuss two broad ways of identifying these variables and thus to classify the sites. In the next chapter we will utilize the analysis to predict the ranges and hence cell count for some known and unknown sites. The utility of the regression analysis method (for range and cell count prediction) would then be realized.

4.2 Experimental Set Up

The experiments were done using the Companion Deployment Tool (CDT)¹. The

1. TM of Nortel (Northern Telecom)

majority of the data was collected previously as part of the Companion deployment process. The CDT is a modification of the wireless PBX Companion that enables quick measurement of the ranges. The CDT and the Companion system use CT2 Plus protocol, the specification of which is briefly given in § 3.2.2.

The deployment procedure consisted of finding the cell centres and the cell boundary (for a predetermined signal level) so that the system coverage extends over the entire floor. The required signal strength is decided based on the receiver sensitivity and Quality of Service (QOS). Spot readings at some other points were also recorded. In this procedure the handset is the mobile transmitter and the base station is the fixed receiver. The signal value displayed on the handset is the received signal power measured by the base station. The receiver measured a one second average of the received signal strength. This is expected to average out the short term fading which typically exhibits itself as a Rayleigh fading [15], [16], [17]. The exercise is illustrated in Figure 4.1. The database consisted of 1016 measurements in various sites.

4.3 Variables Affecting Ranges

In [30] some variables were identified which have an important effect on the propagation (and thus the range). Previous efforts [30] have been directed in identifying variables and utilizing them for regression analysis to provide an accurate way of predicting ranges. In the following subsections we will describe the old variables that were used and the new variables that were identified to improve the performance.

4.3.1 Existing Variables

Over the years efforts had been made to come up with variables that accurately predict the radio propagation in various environments. The database consisted of records for each "vector"¹ measured with the CDT. The important independent variables are:

1. "Vectors" means the independent and the response variables as defined later. Thus one vector would be identified by a site name, cell number, vector number, RSS, Area, longest Length Range (m) and all the other relevant variables.

- i) Site;
- ii) Cell Identification;
- iii) Vector Number;
- iv) CDT Angle of Orientation;
- v) Received Signal Strength (RSS);
- vi) Ceiling Height (in feet);
- vii) Area (in sq. m);
- viii) Longest Length (in m);
- ix) $\gamma = \text{Area}/(\text{Longest Length})^2$;
- x) Four Kinds of Area Types:
 - 1. %Sparse;
 - 2. %Open;
 - 3. %Closed;
 - 4. %Dense;

The major indoor environments that constitute these four kinds of Area Types are discussed in [30]. We summarize the results in Table 4.1 (taken from [30]). Each vector is assigned values for these variables depending on the path traversed. The $\{A_i\}$'s in Figure 4.1 denote this.

TABLE 4.1. Definition of Area Types for Indoor Channels

Area Type	Definition	Examples of Area Types
Sparse	Large open areas with few or no obstructions and a ceiling above 4 metres.	Atrium, Recreational Facilities, Swimming Pool and Low Density Labs
Open	An open area with no intervening walls with obstruction up to 2 metres in height	Open office, hallways and warehouses
Closed	An area with obstructing walls	enclosed offices, apartments, medium and high density labs
Dense	An area surrounded by structural walls	Washrooms, stairwells, elevator, electrical/mechanical rooms, shielded rooms

FIGURE 4.1 The cell boundary (at any pre-determined RSS) is determined by walking away from the base station till the requisite RSS is reached.

A practical problem that arises in this method of classification is the subjectivity of the definitions. Thus what might be perceived as a sparse area by one person is often defined as an open area by another. To overcome this difficulty and to bring about greater objectivity in the classification process, new variables were identified which are described in the following subsections.

4.3.2 New Variables

The new way of classifying and partitioning a site was carried out so that greater objectivity was obtained. Two separate ways of achieving this objective is described. The first one involves defining a non-numeric variable called "Building Type" which would be assigned to the vectors. The second method defines a number of building types based on the functionality and a percentage value is assigned to each of these variables depending on the site. The details of both these approaches are described in the following subsections. As noted earlier the purpose of the exercise was to define new variables that would serve primarily two important functions. Firstly, the variables so defined have to be simple and easy to define by a non-technical sales person and secondly they should give fairly accurate prediction results.

4.3.2.1 Building Type

In this method a new variable (non-numeric) called "Building Type" was defined. The various buildings identified were as follows:

- i) Office;
- ii) Lab/Computer Rm;
- iii) Cafeteria/Atrium;
- iv) Recreation/Fitness Centre;
- v) Factory/WorkShop;
- vi) Condominium.

The vectors were assigned one of these types depending on their nature. The assignment was done such that the larger portion of the vector was confined to the build-

ing type being assigned. An example will make it clear. If a vector was mainly passing through a lab then it would be assigned the type "Lab/Computer Rm". Of course a part of this vector might pass through an office type environment but that part would be smaller. Table 4.2 describes the various areas that the above building types would include.

TABLE 4.2. Definition of Building Types for Indoor Channels

Building Type	Areas
Office	Open and Closed Offices (including the hallways)
Lab/Computer Rm	All Labs and Computer Rooms, Kitchens
Cafeteria/Atrium	Cafeteria, Dining Rooms, Atrium, Lobby
Recreation/Fitness Centre	Fitness Centres, Gyms, Recreation Centres (Indoor Pools)
Factory/WorkShop	Factory, Workshops, Auto Repair
Condominium	Condominium Residences

A difficulty that might arise during range prediction is defining a site which consists of different kinds of building types. A solution is to define separate building types for each of the clearly partitioned areas in a site (Figure 4.2) and carrying out separate predictions for these clearly demarcated areas. However this not an easy task as clearly demarcating the area is not possible in all cases. Also arriving at cell counts with differing predicted ranges (obviously the different regions would give different ranges) is not simple.

A variation of this method is to map a building type (as described before) into average values for the Area Type (i.e. "%Sparse", "%Open", "%Closed", "%Dense") and proceeding with the prediction based on these average values. For a mixed kind of environment a weighted average of the Area Types ("%Sparse", "%Open", "%Closed" and "%Dense") can be used. Thus if 25% of the area is lab and the rest is office (as in Figure 4.2) then the effective "%Sparse" (or "%Open", "%Closed", "%Dense") value would be a weighted average (in the ratio of 1:3) of the average value of "%Sparse" (or the other 3 variables) obtained by the mapping process.

FIGURE 4.2 Example of a mixed kind of environment (clearly demarcated areas).

This variation is refined to produce the second method of classification. Before ending this discussion we would like to remind the reader that the six "Building Types" as listed at the beginning of this sub-section are those used in the classification of the 1016 data that are present in the measured database. As and when new sites are encountered, these definitions can be modified or expanded. Thus some of the new "types" that could be included are: i) Retail/Grocery; ii) Warehouse; iii) Hospital; iv) Hotel etc.

4.3.2.2 Variables based on Functionality

The second method of identifying variables is similar in approach to the method of classification based on area types (§ 4.3.1). Here we define functional area types variables (numeric) which denote the percentage of the total area that is included in that variable. The various variables are as follows:

- i) %Office;
- ii) %Lab/Cmp/Kitc;
- iii) %Wash/Stair_etc;
- iv) %Caf/Atr/Lob;
- v) %Fact/WShop;
- vi) %Ship&Rec/Pack;
- vii) %Rec/Fitness;
- viii) %Condo;
- ix) %Storage;

Values were assigned to these variables for each floor in a site. Thus for a given floor in a site all the vectors would have the same values for these variables. (Note that in the previous method (§ 4.3.2.1) each vector was assigned a "Building Type"). Table 4.3 describes the various area types the above variables would include. This kind of classification is much more objective than the method utilizing area types (§ 4.3.1) as well as the method discussed in § 4.3.2.1 as the classification of the coverage area is done by their functionality rather than their density or contents. Moreover, it is much easier for a non-technical person to identify these variables than the ones outlined in the previous sub-sec-

tions. Table 4.4 summarizes the statistics of the functional area type variables identified. Q1 and Q2 are the first and third quartile respectively. For a set of n ordered samples $x_1, x_2, \dots, x_n$ ($x_1 < x_2 < \dots < x_n$), where $n = 4i + j$, i and j being integers Q1 is [35]:

$$Q1 = \begin{cases} (x_i + x_{i+1}) / 2 & j = 0 \\ (x_i + 3x_{i+1}) / 4 & j = 1 \\ x_{i+1} & j = 2 \\ (3x_{i+1} + x_{i+2}) / 4 & j = 3 \end{cases} \quad (4.1)$$

Similarly Q3 is given by [35]:

$$Q3 = \begin{cases} (x_{n-i} + x_{n-i+1}) / 2 & j = 0 \\ (3x_{n-i} + x_{n-i+1}) / 4 & j = 1 \\ x_{n-i} & j = 2 \\ (x_{n-i-1} + 3x_{n-i}) / 4 & j = 3 \end{cases} \quad (4.2)$$

TABLE 4.3. Definition of Functional Area Types

Variable	Areas
% Office	Open and Closed Office (including corridors).
% Lab/Cmp/Kitc	Labs, Computer Rooms, Kitchens.
% Wash/Stair_etc	Washrooms, Elevators, Stairwells, Electrical, Mechanical and Telephone Rooms etc.
% Caf/Atr/Lob	Cafeteria, Dining Rooms, Atrium, Lobby (especially large) and large open spaces.
% Fact/WShop	Factory, Work Shops, Auto Repairs etc.
% Ship&Rec/Pack	Shipping and Receiving, Packaging Areas.
% Rec/Fitness	Recreation (including pools) and Fitness Centres, Gyms
% Condo	Condominium Residences, Apartment Blocks.
% Storage	Storage Area (especially in a factory environment)

TABLE 4.4. Statistics of the Functional Area Types

Functional Area Type	Min	Q1	Median	Q3	Max
%Office	0	48.4	66.5	75	97.5
%Lab/Cmp/Kitc	0	3.4	5.6	12.2	91.4
%Wash/Stair_etc	0	2.2	6.1	8.1	15
%Caf/Atr/Lob	0	0.7	5.8	18.1	59.7
%Fact/WShop	0	0	0	4.1	72.1
%Ship&Rec/Pack	0	0	0	0	7.1
%Rec/Fitness	0	0	0	0	97.7
%Condo	0	0	0	0	83.9
%Storage	0	0	0	0	11.7

Table 4.4 gives an indication about the variations of the values of these variables. The statistics presented in Table 4.4 was obtained from the database under study. Thus we find that for 50% of the cases, the “%Office” variable has a value less than 66.5. Q1 and Q2 are the first and third quartile respectively indicating the 25% and 75% mark. Knowing the range of the value for the variables allows us to apply the prediction models judiciously. Thus we should be cautious when applying the model to predict for sites where the variables assume values that lie outside the range.

The variables defined above are by no means comprehensive. We have tried to keep the number of variables as low as possible and so the reader would observe that sometimes two or three functional areas are grouped together. This grouping is done where the radio propagation characteristics are expected to be of similar nature. This grouping also takes into account the variations in nomenclature that might arise from site to site. We again note that these definitions were used for the existing database. There has to be an expansion or redefinition of these variables when installation is carried out in new

sites. As noted in § 4.3.2.1 we might have to include new variables to describe sites like: i) Hospital; ii) Retail/Grocery iii) Hotel etc. One point we want to note is that the variable “%Storage” might include warehouses. However we are not in a position to confirm this as there is no data for installations in warehouse. Also the definition of “%Fact/WShop” includes high and very high density labs (e.g. labs with big machinery, upright metal tanks etc.: typically found in IC fabrication facilities). Similarly hotels may be treated as condominiums (ball rooms may be treated as cafeteria) or a new variable may have to be included. This can only be tested when data for these kinds of sites are available. This type of grouping (i.e. extending the definitions) requires a judgement call by the user (loss of objectivity) and should only be done when the RF propagation characteristics of the new kind of area (functional) closely matches the characteristics of an existing definition.

4.4 Regression Analysis

Regression analysis can be defined as the analysis of relationship among variables [31], [32]. There is an enormous amount of literature on Regression analysis. [31], [32], [33], [34] are a small sample. The multiple regression analysis model can be written in matrix notation as [31], [32], [33], [34]:

$$Y = XB + E \quad (4.3)$$

$$\text{where } Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, B = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_k \end{bmatrix} \text{ and } E = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix}. Y \text{ is the vector of}$$

the observations (response), X is the matrix of the regression variables (predictors), B is the vector of the regression coefficients and E is the vector of the random errors associated with each observation. The aim is to find the least squares estimator $\hat{B}$, which minimizes the squared error. The solution is given by:

$$\hat{B} = (X^T X)^{-1} X^T Y \quad (4.4)$$

where $\hat{B}$ is the estimate of B .

For the purpose of this work we have also looked into the effects of adding the cross product between the predictors (independent variables). A logarithmic transformation is carried out on the response variable and regression is carried out to test the suitability of the transformation. Multiple regression analysis was performed to predict the dependent variable Range (R) as function of the independent variables (or predictors) identified earlier. For this case, Eq. (4.3) can be modified so that any function of Range, $g(R)$ can be written as:

$$g(R) = \beta_0 + \sum_i \beta_i x_i + \sum_{i>j} \sum_{i \neq j} \beta_{ij} x_i x_j + \epsilon \quad (4.5)$$

where x_i s are the predictor variables β s are constants and ϵ is the error. The response functions considered in this thesis are the (measured) Range (m) and the $\log_{10}(\text{range})$. The predictor variables are chosen from the independent variables described earlier.

Regression analysis has been carried out using various combinations of the independent variables, described before. The underlying assumption for all these analyses is that the geometrically shortest path between the transmitter and receiver is the dominant mode of propagation. This assumption is generally true, except for some special kind of geometry. Inter floor propagation is not explained in this analysis.

The goodness of fit of the analysis is determined by the parameters: RMS error and the Adj R^2 value. R^2 value is the proportion of the variation of the data that is explained by the regression model. Adj R^2 is a version of R^2 that has been adjusted for the number of terms used in fitting the model. The adjustment is performed in order to ensure that any increase in R^2 is not due to random chance. The Root Mean Square (RMS) Error, s is defined as [35]:

$$s = \sqrt{\left(\sum_i (y_i - \hat{y}_i)^2 \right) / (n - p)} \quad (4.6)$$

where, y_i is the response value of the i^{th} observation (Eq. (4.3)), $\hat{y}_i$ is the estimated (fitted)

value, n is the number of observations and p ($=k+1$) is the number of terms in the model. The term $n-p$ represent the residual degrees of freedom. The R^2 is defined as [35]:

$$R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2} \quad (4.7)$$

where, $\bar{y}$ is the sample mean of the response values and the other symbols are as defined before. The Adjusted R^2 is defined as:

$$AdjR^2 = 1 - \left(\left(1 - R^2 \right) \cdot \frac{n-1}{n-p} \right) \quad (4.8)$$

where the symbols are as defined earlier. The $n-1$ term is due to the constant term β_0 in Eq. (4.5).

4.4.1 Combining Log and Linear Prediction

As already noted, the response functions, $g(R)$, (Eq. (4.5)) considered in this thesis are the (measured) Range and the $\log_{10}(\text{Range})$. It is expected that for certain situations, combining these two response functions would yield a better prediction technique. The technique is basically utilizing these two prediction methods (i.e. the Log and the Linear) over two different regions. The boundary between these two regions is defined by the break-even range R_b .

The break-even range, R_b , is defined as the range at which the average error (in the linear domain i.e. in metres) due to the log analysis (i.e. taking the “log(range)” as the response variable) equals the RMS error due to the normal analysis (i.e. using “Range (m)” as the response variable). Figure 4.3 explains the definition of the break-even range R_b . Let us develop this notion more formally. We note that if the RMS error in the log analysis (domain) is Δ , then the corresponding errors in the linear domain (i.e. in metres) ϵ_1 and ϵ_2 are given by:

$$\epsilon_1 = 10^{\log R + \Delta} - R = R \left(10^{\Delta} - 1 \right) \quad (4.9)$$

$$\epsilon_2 = R - 10^{\log R - \Delta} = R(1 - 10^{-\Delta}) \quad (4.10)$$

where R is the actual range. Now it can be easily shown that:

$$\epsilon_1 - \epsilon_2 = R \cdot 10^\Delta \cdot (1 - 10^{-\Delta})^2 \quad (4.11)$$

The Right Hand Side of Eq. (4.11) is a positive quantity for all values of Δ . Thus $\epsilon_1 > \epsilon_2$. So we can deal with the greater value (i.e. ϵ_1), and for simplicity we will drop the subscript and call it ϵ . The break-even range, R_b , is thus the solution of:

$$\epsilon = \epsilon_1 = R \cdot (10^\Delta - 1) = \epsilon_N \quad (4.12)$$

where ϵ_N is the RMS error in the linear analysis (i.e. using Range). Solving for R in Eq. (4.12) we get R_b :

$$R_b = \frac{\epsilon_N}{10^\Delta - 1} \quad (4.13)$$

Thus we can use the linear prediction technique for ranges greater than the break even range R_b . The method can be summarized as follows:

- Use log method (case **b** in § 5.2) and predict ranges: R_p .
- If R_p is greater than R_b , then use the normal method (case **a** in § 5.2) to get R_p' .
- Else keep R_p .
- Use R_p or R_p' as per as the above rule.

The basic assumption in this process is that the errors are constant over the whole range of interest and equal to the RMS error. This may not be strictly true as is evident from Figure 5.5. However, for cases when the assumption holds, combining log and linear prediction technique will give better results.

FIGURE 4.3 Definition for R_b . (a): Log Analysis, error in log; (b): Transformed to the linear Domain; (c): Range Analysis, error in metres and (d): Superimposing (b) and (c).

4.5 Summary

In this chapter, we have introduced the mathematical foundation of Regression analysis. We have also discussed the variables that characterize the indoor radio propagation environment and which can be subsequently utilized to build up a regression prediction model. The various area types that these variables would include is also tabulated. We have also summarized the statistics of the variables for the database used in the analysis. In the next chapter we will present the results of the Regression analysis carried out. We will perform the analysis with various combinations of these variables and the comparisons would be presented. The most accurate one would be utilized for predicting the range (and hence the cell count) for some known and unknown sites.

Chapter 5

RESULTS AND VERIFICATION OF REGRESSION ANALYSES

5.1 Introduction

In the previous chapter we discussed the theoretical concepts behind the Regression analysis technique and the variables that affect indoor radio propagation had been identified. In this chapter we will utilize these variables to determine a model that can be used for range prediction. We will compare some of the models that can be used and use the most accurate one to predict ranges for some of the known and unknown sites.

5.2 Regression Results

We have already defined the variables and discussed the basis of Regression Analysis in Chapter 4. In this section we will present results of the Regression Analysis of some of the models studied. Table 5.1 provides a summary of these results. The mathematical expressions for RMS error and Adj R^2 values are given in Eq. (4.6), Eq. (4.7), Eq. (4.8).

TABLE 5.1. Results for Regression Analysis

SI #	Independent Variables	SI #	Response	RMS	Adj R^2	R_b (m)
I	RSS, Ceiling, Area, gamma, % Sparse	a	Range (m)	12.180	0.497	33.12
		b	log(range)	0.136	0.505	
II	RSS, Ceiling, Area, gamma	a	Range (m)	13.138	0.415	34.25
		b	log(range)	0.141	0.464	

TABLE 5.1. Results for Regression Analysis

SI #	Independent Variables	SI #	Response	RMS	Adj R ²	R _b (m)
III	RSS, Ceiling, Area, Longest Length, gamma, % Sparse	a	Range (m)	10.775	0.607	33.86
		b	log(range)	0.120	0.614	
IV	RSS, Ceiling, Area, Longest Length, gamma	a	Range (m)	12.105	0.504	35.32
		b	log(range)	0.128	0.559	
V	RSS, Ceiling, Area, gamma, Building Type	a	Range (m)	12.774	0.447	35.96
		b	log(range)	0.132	0.530	
VI	RSS, Ceiling, Area, Longest Length, gamma, Building Type	a	Range (m)	11.720	0.535	36.13
		b	log(range)	0.122	0.603	
VII	RSS, Ceiling, Area, Longest Length, gamma, % Sparse, % Open, % Closed, % Dense	a	Range (m)	9.485	0.695	36.23
		b	log(range)	0.101	0.725	
VIII	RSS, Ceiling, Area, Longest Length, gamma, % Office, % Lab/Cmp/Kitc, % Wash/Stair_etc, % Caf/Atr/Lob, % Fact/WShop, % Rec/Fitness, % Ship&Rec/Pack, % Condo, % Storage	a	Range (m)	10.986	0.591	38.91
		b	log(range)	0.108	0.687	
IX	RSS, Ceiling, Area, Longest Length, gamma, % Op&%Sp, % Closed, % Dense	a	Range (m)	10.336	0.638	37.00
		b	log(range)	0.107	0.691	
X	RSS, Ceiling, Area, gamma, % Sparse, % Dense	a	Range (m)	11.405	0.559	34.20
		b	log(range)	0.125	0.584	
XI	RSS, Ceiling, Area, Longest Length, gamma, % Sparse, % Dense	a	Range (m)	10.229	0.646	34.77
		b	log(range)	0.112	0.662	

The best possible prediction would be using the analysis identified by VII (pro-

vided that we can accurately define the "Area Type" variables: "% Sparse", "% Open", "% Closed", "% Dense"). The worst case scenario is utilizing the analysis **IV** (if we consider the models with the "Longest Length" variable included), which uses only the geometry of the site. In this context we note that adding the variable "Longest Length" has a marked effect on improving the analysis. Comparison of **I & III**, **II & IV**, **V & VI** and **X & XI** makes this obvious. In all the cases tabulated, the log model (i.e. using the variable "log(range)": identified as **b**) gives a better performance (as far as the Adj R^2 value is concerned) than the linear model (i.e. using the "Range (m)" variable: case **a**). Focussing our attention on case **VIII**, we see this that the analysis **a** is comparable to case **III a** (which is a modified version of **Ia**: with the "Longest Length" term added); while the analysis **b** is much better than **III b**. Also it is worth noting that the case **VIII b** is quite close to the best case **VII** (which uses the Area Type variables: "% Sparse", "% Open", "% Closed", "% Dense"). We have also investigated the effects of combining the "% Sparse" and "% Open" into a single variable the "% Op&% Sp" (case **IX**). As discussed in § 4.3.2.2, this method of classification is much more objective, which would be of great help in defining the parameters objectively for predicting ranges in a new site.

In § 4.3.2.1 we proposed that another alternate approach would be to map "Building Type" into average values for "Area Type" ("% Sparse", "% Open", "% Closed", "% Dense") and utilize these average values for predicting ranges. In Table 5.2 we give these mean values, grouped by the various Building Types. The mapping method is not expected to be as accurate as the method using functional area types (case **VIII**), because by using the average values for the Area Type variables ("% Sparse", "% Open", "% Closed", "% Dense") the site specific information is lost. Moreover to get the weighted average values, we have to identify the functional area types. Thus the method using the functional area types (case **VIII**) is a better method and it has been subsequently utilized to predict the ranges.

Figure 5.1 through Figure 5.26 show the results of some of the analyses summarized in Table 5.1. The residuals versus the fitted (or predicted) variable plot is a qualitative test of the goodness of fit. If the residuals increase, then the model fails to explain some of the variation and other terms or higher orders need to be considered. Figure 5.1

shows the residuals for the log analysis with the functional area type variables (case **VIII b**). It is observed that the residuals are almost constant and so this model is suitable. Figure 5.2 shows these residuals transformed to the linear domain (i.e. residuals have been expressed in metres). As explained in § 4.4.1 and Figure 4.3, there is a linear relationship between the residuals, expressed as metres and the residuals expressed as log. The other figures show similar kinds of behaviour for some of the other cases studied. The residual histograms also give a qualitative idea about the model. As is evident from Figure 5.3, the residual histogram for the case **VIII b** is close to a Gaussian density function, which satisfies the assumptions for the validity of the Regression analysis.

TABLE 5.2. Mapping of Building Type into Average Values for Area Type

Building Type	%Sparse	%Open	%Closed	%Dense	%Op&%Sp
Cafeteria/Atrium	44.57	25.27	25.71	4.61	69.84
Condominium	0	12.67	58.64	28.76	12.67
Factory/WorkShop	63.75	3.60	29.38	3.36	67.35
Lab/Computer Rm	23.78	40.56	29.71	5.95	64.35
Office	13.12	50.25	29.63	7.01	63.37
Recreation/Fitness Centre	67.92	5.00	3.85	23.31	72.92

From the regression results summarized in Table 5.1, we observe that the case **VII** gives the minimum RMS error. However the variables used in this case are difficult to specify and are subjective in nature. We thus conclude that the log analysis using the functional area type variables (case **VIII b**) is the best model to use for range prediction. In this case (**VIII**), the variables used can be objectively defined leading to better prediction results. Although the goodness of fit parameters give an idea about how well the model fits the data, the ultimate test of the model is in predicting the ranges for unknown sites. The prediction can be done using Eq. (4.5), where β 's are the coefficients obtained from the models. The prediction results are given in the next section. The coefficients obtained from the models **VIII a** and **b** are given in Table 5.3 and Table 5.4 respectively.

FIGURE 5.1 Residuals vs. Fitted “ $\log(\text{range})$ ” (case VIII b). Note that the residual error is approximately constant.

FIGURE 5.2 Residuals vs. Fitted Range for “ $\log(\text{range})$ ” analysis (case VIII b), transformed to the linear domain (in metres).

FIGURE 5.3 Residuals Histogram for “log (range)” analysis (case VIII b). Note that it is very close to the Gaussian pdf.

FIGURE 5.4 Residuals Probability plot for “log (range)” analysis (case VIII b). The straight line represents the plot for Gaussian distribution.

FIGURE 5.5 Residuals vs. Fitted Range for “Range (m)” analysis (case VIII a). Note that the fan shaped nature is much more pronounced than that for the “log (range)” analysis (case VIII b) (Figure 5.1, Figure 5.2)

FIGURE 5.6 Residuals Histogram for “Range (m)” analysis (case VIII a). Note that there is a deviation from the Gaussian pdf.

FIGURE 5.7 Residuals Probability Plot for the “Range (m)” analysis (case VIII a). Note that there is a deviation from the Gaussian case.

FIGURE 5.8 Residuals vs. Fitted Range for the “Range (m)” analysis (case I a). Note that the fan shaped nature is pronounced.

FIGURE 5.9 Residuals Histogram for “Range (m)” analysis (case I a). Note there is a deviation from the Gaussian pdf.

FIGURE 5.10 Residuals Probability Plot for “Range (m)” analysis (case I a).

FIGURE 5.11 Residuals vs. Fitted Range for case IV a.

FIGURE 5.12 Residuals Histogram for case IV a.

FIGURE 5.13 Residuals vs. Fitted log(range) for case IV b.

FIGURE 5.14 Residuals vs. Fitted Range (transformed to Linear Domain) for case IV b.

FIGURE 5.15 Residuals Histogram for Case IV b

FIGURE 5.16 Residuals vs. Fitted Range for case VI a.

FIGURE 5.17 Residuals Histogram for case VI a

FIGURE 5.18 Residuals vs. Fitted Range for case VI b

FIGURE 5.19 Residuals transformed to linear domain (case VI b)

FIGURE 5.20 Residuals Histogram for case VI b.

FIGURE 5.21 Residuals vs. Fitted Range for case VII a.

FIGURE 5.22 Residuals Histogram for case VII a

FIGURE 5.23 Residuals vs. Fitted log(range) for case VII b.

FIGURE 5.24 Residuals vs. Range transformed to the linear domain for case VII b.

FIGURE 5.25 Residuals Histogram for case VII b.

FIGURE 5.26 Residuals Probability plot for “log (range)” analysis (case VII b).

2

of/de

2

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

TABLE 5.3. Coefficients from the analysis Villa (using "Range (m)")

Terms (Variables)	Coefficient	Std Error	Significance
Constant	13492.876	4736.793	0.004
RSS	-0.921	0.126	0
Area	0.048	0.017	0.005
Longest Length	-4.024	1.778	0.024
Ceiling	-20.367	10.612	0.055
gamma	-6298.62	3117.928	0.044
% Office	-132.699	46.507	0.004
% Lab/Cmp/Kite	-139.264	46.993	0.003
% Wash/Stair_etc	-140.773	48.155	0.004
% Caf/Atr/Lob	-132.926	46.475	0.004
% Fact/WShop	-124.254	38.557	0.001
% Ship&Rec/Pack	-27.809	34.693	0.423
% Rec/Fitness	6.239	64.767	0.923
% Condo	-114.918	39.081	0.003
% Storage	-109.317	36.15	0.003
Area * RSS	5.309 e-05	1.535 e-05	0.001
% Storage * RSS	-0.099	0.035	0.005
Area * Ceiling	-0.003	0.002	0.040
% Wash/Stair_etc * Area	-7.158 e-04	2.529 e-04	0.005
% Caf/Atr/Lob * Area	-3.660 e-04	1.504 e-04	0.015
Ceiling * Longest Length	0.362	0.180	0.045
Longest Length * gamma	-0.751	0.436	0.085
% Caf/Atr/Lob * Longest Length	0.025	0.015	0.085
% Lab/Cmp/Kite * Ceiling	0.774	0.229	0.001

TABLE 5.3. Coefficients from the analysis VIIIa (using "Range (m)")

Terms (Variables)	Coefficient	Std Error	Significance
%Cat/Atr/Lob * Ceiling	-0.68	0.334	0.042
%Fact/WShop * Ceiling	0.786	0.255	0.002
%Ship&Rec/Pack * Ceiling	-8.286	1.939	0
%Rec/Fitness * Ceiling	-11.006	5.44	0.043
%Storage * Ceiling	1.283	0.741	0.084
%Office * gamma	62.162	31.019	0.045
%Lab/Cmp/Kite * gamma	61.565	31.087	0.048
%Wash/Stair_etc * gamma	64.310	33.047	0.052
%Cat/Atr/Lob * gamma	65.322	32.208	0.043
%Lab/Cmp/Kite * Office	-0.007	0.003	0.023
%Office * %Wash/Stair_etc	0.094	0.032	0.004
%Cat/Atr/Lob * Office	0.030	0.010	0.003
%Cat/Atr/Lob * %Lab/Cmp/Kite	0.071	0.026	0.006
%Cat/Atr/Lob * %Wash/Stair_etc	0.588	0.208	0.005
%Condo * %Wash/Stair_etc	0.124	0.041	0.003
%Cat/Atr/Lob * %Fact/WShop	-3.684	1.41	0.009
%Cat/Atr/Lob * %Storage	-3.355	1.251	0.007
%Fact/WShop * %Ship&Rec/Pack	4.994	1.717	0.004
%Fact/WShop * %Storage	-7.037	2.763	0.011
%Ship&Rec/Pack * %Storage	33.507	14.076	0.017
%Rec/Fitness * %Storage	286.475	112.050	0.011

TABLE 5.4. Coefficients from the analysis VIIIb (using "log(range)")

Terms (Variables)	Coefficient	Std Error	Significance
Constant	153.154	26.73	0
RSSI	-0.051	0.011	0
Area	3.404 e-04	9.589 e-05	0
Longest Length	-0.038	0.011	0.001
Ceiling	-0.100	0.068	0.141
gamma	3.542	1.678	0.035
% Office	-1.530	0.265	0
% Lab/Cmp/Kitc	-1.610	0.266	0
% Wash/Stair_etc	-1.589	0.266	0
% Caf/Atr/Lob	-1.499	0.262	0
% Fact/WShop	-1.737	0.285	0
% Ship&Rec/Pack	-0.629	0.293	0.032
% Rec/Fitness	-0.393	0.586	0.502
% Condo	-1.537	0.265	0
% Storage	-1.418	0.270	0
Ceiling * RSSI	9.308 e-04	1.488 e-04	0
RSSI * gamma	0.066	0.023	0.005
% Storage * RSSI	-6.706 e-04	2.704 e-04	0.013
Area * Longest Length	1.449 e-07	5.337 e-08	0.007
Area * Ceiling	-2.252 e-05	9.128 e-06	0.014
% Wash/Stair_etc * Area	-9.135 e-06	2.256 e-06	0
% Caf/Atr/Lob * Area	-4.269 e-06	9.263 e-07	0
Ceiling * Longest Length	0.003	0.001	0.005
% Caf/Atr/Lob * Longest Length	2.229 e-04	6.826 e-05	0.001

TABLE 5.4. Coefficients from the analysis VIIIb (using "log(range)")

Terms (Variables)	Coefficient	Std Error	Significance
%Lab/Cmp/Kite * Ceiling	0.009	0.002	0
%Cal/Atr/Lob * Ceiling	-0.01	0.002	0
%Fact/WShop * Ceiling	0.007	0.002	0.001
%Ship&Rec/Pack * Ceiling	-0.093	0.014	0
%Rec/Fitness * Ceiling	-0.104	0.048	0.031
%Lab/Cmp/Kite * %Office	-1.523 e-04	3.061 e-05	0
%Office * %Wash/Stair_etc	7.902 e-04	2.815 e-04	0.005
%Cal/Atr/Lob * %Office	3.728 e-04	7.757 e-05	0
%Cal/Atr/Lob * %Lab/Cmp/Kite	0.001	0	0
%Cal/Atr/Lob * %Wash/Stair_etc	0.007	0.002	0
%Condo * %Wash/Stair_etc	0.001	0	0
%Cal/Atr/Lob * %Fact/WShop	-0.078	0.024	0.001
%Cal/Atr/Lob * %Storage	-0.036	0.011	0.001
%Fact/WShop * %Ship&Rec/Pack	0.102	0.029	0.001
%Fact/WShop * %Storage	-0.108	0.032	0.001
%Rec/Fitness * %Ship&Rec/Pack	-0.641	0.234	0.006
%Ship&Rec/Pack * %Storage	0.538	0.166	0.001
%Rec/Fitness * %Storage	4.161	1.197	0.001

The "Std Error" in Table 5.3 and Table 5.4 is a measure of the uncertainty of the estimated coefficient. The "Significance" provides an idea about the relevance of the particular term in the model. Terms with high values of Significance can be eliminated from the model.

5.3 Verification of the Model

As already pointed out, the acid test of any model selected is its ability to correctly predict the ranges for any unknown sites. Towards this end, we will investigate how well our models predict ranges. We will present the results for the model using functional area type variables (cases **VIII a** and **b**). The predicted cell count is obtained based on a geometric consideration of optimally packing minimum number of circles within a rectangle. We have carried out the testing on known and unknown sites. Known sites refer to those sites which are included in the database and thus have been used in the Regression analysis. Unknown sites are new sites which are not included in the data base.

5.3.1 Known Sites

Testing the model on known sites can be considered as a first step in the verification process. Failure to predict the cell count for known sites should make the model invalid. Table 5.5 gives a summary of the prediction results from some of the typical sites from the database. The RSS value used for these sites is -70 dBm.

TABLE 5.5. Cell Count from Predicted Range using analysis VIII a & b

Site	Predicted Range		Predicted Cell Count		Actual Cell Count
	VIII a	VIII b	VIII a	VIII b	
Site 1 (1 st Floor)	36.93	36.46	4	4	4
Site 2 (2 nd Floor)	34.47	34.55	5	5	5
Site 3 (1 st Floor)	18.07	17.91	4	4	4
Site 4	20.09	19.81	3	3	4
Site 5 ^a	22.70	21.70	5	5	5
Site 6 (14 th Floor)	27.82	27.19	2	2	1
Site 6 (3 rd Floor)	26.72	26.21	2	2	2

TABLE 5.5. Cell Count from Predicted Range using analysis VIII a & b

Site	Predicted Range		Predicted Cell Count		Actual Cell Count
	VIII a	VIII b	VIII a	VIII b	
Site 7 (5 th Floor)	29.10	27.08	1	1	3
Site 7 (6 th Floor)	25.90	26.13	1	1	2
Site 8	25.64	24.52	1	1	2
Site 9	27.80	27.33	1	1	1
Site 10	29.95	29.00	1	1	1
Site 11	54.68	51.43	3	3	3
Site 12 ^b	53.77	51.18	2	2	3
Site 13 (1 st Floor) ^c	43.13	41.86	6	7	7
Site 14 (1 st Floor)	33.73	33.09	7	7	7

a. with weighted average of Ceiling Height=11.95 ft.

b. with weighted average of Ceiling Height=20.16 ft.

c. with weighted average of Ceiling Height=10.55 ft.

We note that this model is effective in predicting the cell count. The error obtained for Site 6 (14th Floor) can be explained because of the unusual shape of the building. In this case the rectangular approximation of the floor may not be justified. Also the whole floor was not covered in the deployment process. It can be seen for the same site: 3rd Floor, that 2 cells were used. For the Site 12, we note that this is partly factory and partly office. The factory part has high ceiling height. This wide difference of the ceiling heights may have caused the slight error. However the errors are small and acceptable. The interesting point to note is that in all the erroneous cases the prediction model tends to overestimate the ranges, leading to a smaller cell count.

5.3.2 Unknown Sites

Four unknown sites were considered. Of these two were in Carleton University: the Mackenzie building and the Dunton tower, while the other two were offices in high rise. A brief description of these sites is provided below.

The two sites in Carleton University were the fourth floor of Mackenzie Building and the third floor of Dunton Tower. The Mackenzie Building has concrete block interior walls. It has a central laboratory area ringed by classrooms and offices. The Dunton Tower is a modern multi-storey building with centrally spaced elevators. The third floor consists of classrooms. The third site investigated was an open plan office of CGI in a multi-storey building in downtown Ottawa. The last site investigated was also an open plan office of RES. International in downtown Ottawa.

Values for the relevant functional variables are assigned and the prediction is carried out for the linear and the log cases (cases **VIII a** and **b**). Table 5.6 and Table 5.7 give a summary of the values for the variables used in the prediction process.

TABLE 5.6. Summary of the geometric variables for the unknown sites

Site	Area (sq m)	Longest Length (m)	RSS	Ceiling Height (ft)
Mackenzie Building: 4 th Floor	1106	48	-62	11.19 ^a
Dunton Tower: 3 rd Floor	692	39	-70	8
CGI, Ottawa	1197	57	-70	8
RES, Ottawa	597	36	-70	8

a. Weighted average of 8.33 ft for washrooms, corridors and 12.33 ft for the rest.

As already noted that the Mackenzie Building has concrete internal walls, which has high attenuation. This has also been reported in [27]. As the Regression analysis does not consider the effects of wall attenuation explicitly, we include a 8 dB loss and predict using a higher cell boundary value (i.e. a higher RSS value — -62 dBm instead of -70 dBm). This 8dB wall loss value used is obtained from [27]. This is done for the Mackenzie

Building because this building is quite different from modern buildings. Table 5.8 summarizes the cell count results for these three unknown sites and Table 5.8 presents a summary of the statistics of the residuals (or the errors) in prediction using the cases **VIII a & b**.

TABLE 5.7. Summary of the functional variable for the unknown sites^a

Site	%Office	%Lab/ Cmp/Kite	%Wash/ Stair_etc	%Storage
Mackenzie Building: 4 th Floor	42.1	45.3	12.6	0
Dunton Tower: 3 rd Floor	77.5	0	22.5	0
CGI, Ottawa	70.8	17.5	10.4	1.3
RES, Ottawa	76.3	8.1	15.6	0

a. Functional variables not shown in the table are zero.

We observe that the best prediction is obtained for the two office sites: CGI, Ottawa and RES, Ottawa. This is because that the database utilized for the Regression Analysis consisted mainly of office areas. The CGI and the RES site were also office areas and thus the prediction yielded good results. The other two sites were University locations and as this environment was not represented in the database so the prediction yielded poor results. The decision about good or poor results are made based on the mean and standard deviations of the residuals (Table 5.9) and based not on the cell count values (Table 5.8). This is because that for the two university cases, although we get large residuals, the cell count error is not large, because the cell count for these sites are small. Thus, even if we predict very large ranges (i.e. large errors) we will still arrive at a cell count of 1.

We can also get a subjective idea about the utility of the combined log and linear prediction which we discussed in § 4.4.1. For the Mackenzie Building we see (in Table 5.9) that the log analysis (case **VIII b**) produces larger error (in terms of metres) than the linear analysis (case **VIII a**). We also note that the predicted range is 36.78 m using the log analysis (case **VIII b**) which is close to the break even range (R_b) value of 38.91 m (Table 5.1). For the other sites we note that the log analysis (case **VIII b**) pro-

duces smaller mean residuals. In all these cases the predicted range is smaller than the break even range (R_b).

TABLE 5.8. Cell Count from the predicted Range using analysis VIII a and b for unknown sites

Site	Predicted Range (m)		Predicted Cell Count		Actual Cell Count
	VIII a	VIII b	VIII a	VIII b	
Mackenzie Building: 4 th Floor	26.85	37.54	1	1	2
Dunton Tower: 3 rd Floor	37.00	28.45	1	1	1
CGI, Ottawa	26.25	18.35	1	2	2
RES, Ottawa	30.84	22.47	1	1	1

TABLE 5.9. Summary of the Statistics of the Residuals from the experimental data for the unknown sites

Site	Predicted Range (m)		Mean Residuals ^a (m)		Residual Standard Deviation (m)	
	VIII a	VIII b	VIII a	VIII b	VIII a	VIII b
Mackenzie Building: 4 th Floor	26.85	37.54	-9.19	-20.29	4.47	4.52
Dunton Tower: 3 rd Floor	37.00	28.45	-18.31	-10.29	2.97	2.47
CGI, Ottawa	26.25	18.35	-6.10	0.96	3.61	2.85
RES, Ottawa	30.84	22.47	-7.06	0.58	3.53	3.60

a. Residuals is defined as: (Actual Range - Predicted Range) for all the experimental data.

5.4 Summary

This chapter compares the various Regression analysis performed. We find that the analysis using the functional variables (case **VIII**) is simple and accurate. Moreover, these variables are easy to identify and can be objectively defined. This model (with the two subcases **a** & **b**) have been verified on test cases consisting of known and unknown sites. The model was found to be most suitable for office areas. This was expected to be the case as the database consisted of mainly office environment. We have found that the log model (case **VIII b**) is more accurate than the linear model (case **VIII a**), provided that the predicted range is less than the break even Range, R_b . We have also observed that combining the log and linear prediction technique gives a smaller error.

Chapter 6

CONCLUSIONS

6.1 Thesis Summary

The work presented in this thesis has the objective of providing a method for range prediction, which is an important issue for the deployment of both indoor and outdoor radio systems. With the rapid expansion of universal portable and personal communication systems, range prediction is increasingly becoming important. Accurate range prediction would enable service providers to deploy systems without going through a lengthy and costly experimental process to determine the base station requirements. This thesis has examined two different methods for range prediction for indoor wireless systems: the direct method and the Regression analysis method.

The thesis began with a study on the indoor wireless channel. Literature available for the indoor channel was studied and various channel parameters were reviewed. Brief discussions on the major impairments for the indoor channel like shadowing, frequency selective fading etc. have been presented in Chapter 2. A brief review of the two prediction methods is also presented.

The direct method of prediction has been studied in details in Chapter 3. An important aspect of this method is the study of the variation of the standard deviation of the log-normal shadowing (σ) with distance. Two different classes of indoor environments have been studied: the University environment and a typical Office environment. It was found that σ was almost constant for the University, while for the office σ increased linearly with distance. Actual prediction was not performed using the direct method, although a general method was suggested to design systems (and thus to predict ranges) based on certain outage probability. A discussion on the Kolmogorov-Smirnov goodness of fit test has also been included in this chapter and an illustration of this test has been shown by comparing the Cumulative Distribution Function of a simulated Rayleigh enve-

lope with the theoretical one. The test has been performed to check whether the CDF of the shadow component, obtained from experiments, are log-normal. It has been found that almost all the CDFs pass the K-S test with a 0.95 confidence level for log-normal distribution.

The mathematical principles behind Regression Analysis has been presented in Chapter 4. The different goodness of fit parameters like RMS error, R^2 and Adjusted R^2 values have been explained. The variables that affect radio propagation inside buildings have been identified and these variables would be utilized for the analysis. This exercise is based on the work done in [30] and explores ways to enhance that work.

The results for the Regression analysis have been presented in Chapter 5. We have considered eleven models each with two sub-cases: the log and the linear. It is observed that the log model using the functional area type variables, which have been identified in Chapter 5, (case **VIII b**), is an accurate prediction technique. We have also found that the variable “Longest Length” improves the prediction, by reducing the RMS error. It has also been observed that the log model (cases **b**) produces a better Adjusted R^2 value than the corresponding linear models (cases **a**).

6.2 Future Work

It has been shown that the Regression model developed is a pragmatic and accurate way to tackle the issue of range prediction. It can, however, be improved. The stability of the model should be studied rigorously. One way is to investigate the effects of using the Principal Components instead of the original variables in the Regression models tested. Moreover some new variables can be added to the existing model to improve the accuracy. The effect of including the variable “Length of Internal Walls” should also be examined. It is expected that the presence or absence of internal structural walls has an important effect on radio propagation. This new variable has the potential of improving the model. The database can also be expanded to include some important classes of indoor environments such as shopping malls, super-stores, hospitals, hotels etc. This would require the definition of new variables for these sites would have to be defined. Another

area of future work is to investigate the behaviour of including higher order terms like squared or cubed terms, in the regression analysis equation, and the stability of models incorporating such higher order terms.

The direct method of range prediction can also be studied in greater details. An important aspect would be to utilize such method to predict ranges. This process would utilize the knowledge about the variation of the standard deviation, σ , with distance, which has been studied in this work.

Of course the ultimate dream for any researcher investigating the range prediction issues is to develop an analytical model that completely describes the physical phenomenon of radio wave propagation in indoor environments. However, the enormity of the task can be very easily realized and statistical methods to deal with the problem appears to be simpler. We thus feel that statistical range prediction techniques seem to be the practical alternative and should be investigated thoroughly.

Appendix A

DERIVATION OF THE CUMULATIVE DISTRIBUTION FUNCTION FOR A NORMALIZED RAYLEIGH VARIABLE

Let us start with the probability density function (pdf), $p(r)$ of a Rayleigh distributed envelope:

$$p(r) = \begin{cases} \frac{r}{\Gamma} \exp\left[-\frac{r^2}{2\Gamma}\right] & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (\text{A } 1)$$

where r is the received envelope and Γ is the mean received power. Let us define the normalized received power, ρ as:

$$\rho = 10 \log\left(\frac{1}{2}r^2\right) - 10 \log \Gamma \quad (\text{A } 2)$$

Solving for r in Eq (A2) we get:

$$r = 10^{(\rho/20)} \sqrt{2\Gamma} \quad (\text{A } 3)$$

Our next step would be to determine the probability density function (pdf) of the variable ρ . This is a simple exercise by utilizing the principle of statistics. If $Y = g(X)$ and if the pdf of the random variable X is $f_X(x)$, then the pdf, $f_Y(y)$, of the random variable Y is given by:

$$f_Y(y) = \sum \frac{f_X(x)}{|g'(x)|} \bigg|_{x=g^{-1}(y)} \quad (\text{A } 4)$$

For this case $g(r) = 10 \log\left(\frac{1}{2}r^2\right) - 10 \log \Gamma$ which implies $g'(r) = \frac{20}{\ln 10} \cdot \frac{1}{r}$.

Substituting in Eq (A4) and simplifying, we get:

$$p(\rho) = \frac{\ln 10}{10} \cdot 10^{\frac{\rho}{10}} \cdot \exp \left[-10^{\frac{\rho}{10}} \right] \quad (\text{A } 5)$$

Integration of Eq (A5) yields the Cumulative Density Function, $F_p(x) = \text{Prob}[\rho \leq x]$ of ρ :

$$F_p(x) = \int_{-\infty}^x p(\rho) d\rho = 1 - \exp \left[-10^{\frac{x}{10}} \right] \quad (\text{A } 6)$$

Appendix B

SOURCE CODE FOR COMPUTATION OF THE CUMULATIVE DENSITY FUNCTION

```
function cdf = cdf_general(x,t)
% cdf_general calculates the CDF of any data.
%
% cdf = cdf_general(...t)
%
% x : vector of the data.
% t : vector of the abscissa points ie cdf=prob(x<=t)
%
% if any of the vectors is a matrix the first row is considered
%
[m,length_x]=size(x);
[m,length_t]=size(t);
cdf = zeros(1,length_t);
for index_t=1:length_t,
    counter =0;
    for index_x=1:length_x,
        if x(1,index_x)<= t(1,index_t),
            counter=counter+1;
        end
    end
    cdf(1,index_t)=counter/length_x;
end
```

Appendix C

SOURCE CODE FOR THE KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST

```
function [pass,D_max,counter] = KS_dist(cdf,y,alpha,data_points)
% KS_dist tests whether the experimental CDF passes the KS test with respect
% to a theoretical distribution.
%
% [pass,D_max,counter] = KS_dist(expt,theory,alpha,points)
%
% expt   : vector containing the experimental CDF
% theory : vector containing the theoretical CDF
% alpha  : level of confidence
% points : number of points considered to calculate the experimental CDF
% pass   : pass=1 means test passes, pass=0 means fails;
% D_max  : the max distance between the experimental & theoretical CDFs
% counter : denotes the position where the D_max occurs
%
% If any of the vectors is a matrix, the first row is considered.
% If the length of the vectors expt and theory do not match then the least
% length is considered and the other data are not considered.

[m,n]=size(cdf);
[p,m]=size(y);
clear p;
counter=1;
if m>n,
    m=n;
end
```

```
epsilon=sqrt(-1/(2*(data_points)*log(alpha/2)));  
D_max = 0.0;  
for index = 1:m,  
    D= cdf(1:index)-y(1:index);  
    D = abs(D);  
    if D>D_max,  
        D_max = D;  
        counter = index;  
    end  
end  
if D_max<=epsilon,  
    pass = 1;  
else  
    pass = 0;  
end
```

Appendix D

SOURCE CODE FOR SIMULATING THE NORMALIZED POWER IN PRESENCE OF DOPPLER AND RAYLEIGH FADING

```
function [ray_power,time] = ray_simu(freq,vel,N0,points)
%
% [ray_power,time] = ray_simu(freq,vel,N0,points)
%
% simulates the received power(normalized) in dB in Doppler
%
% freq      : carrier frequency in Hz
% vel       : velocity of the mobile in m/sec
% N0        : number of oscillators in the simulation model
% points    : number of points in the simulated profile (suggested >750)
%
% ray_power : output vector containing the power profile (normalized)
% time      : the range of the times

ray_power = zeros(1,points);
time = zeros(1,points);
lamda = 3e8/freq;
wm = 2*pi* vel/lamda; % vel in m/sec
tmax = 50*lamda/vel;
t_inc = tmax/(points-1);
pow_total = 0.0;
N = 2*(2*N0+1); % N0 is the # of simulators
t = 0.0;
```

```

for index = 1 : points,
    xc = 0.0;
    xs = 0.0;
    for n = 1 : N0,
        bn = pi*n/(N0+1);
        wn = wm*cos(2*pi*n/N);
        xc = xc+2*cos(bn)*cos(wn*t);
        xs = xs+2*sin(bn)*cos(wn*t);
    end
    xc = xc+sqrt(2)*cos(wm*t);
    pow_total = pow_total+xc*xc+xs*xs;
    ray_power(index) = 10*log10(xc*xc+xs*xs);
    time(index) = t;
    t = t+t_inc;
end
avg_power = 10*log10(pow_total/points);
ray_power = ray_power - avg_power;

```

References

- [1] S.Venkataraman, "Filter Structure for Equalization and Diversity Combining", *M.Eng. thesis*, Dept. of Systems and Computer Engineering, Carleton University 1994.
- [2] D.J.Goodman, "Trends in Cellular and Cordless Communications", *IEEE Communications Magazine*, June 1991, pp. 31-40.
- [3] J.E.Padgett, C.G.Gunther and T.Hattori, "Overview of Wireless Personal Communications", *IEEE Communications Magazine*, Jan 1995, pp. 28-41.
- [4] R.Steele, "Deploying Personal Communications Networks", *IEEE Communications Magazine*, September 1990, pp. 12-15.
- [5] D.Cox, "Personal Communications- A Viewpoint", *IEEE Communications Magazine*, Nov. 1990, pp. 8-20.
- [6] D.Cox, H.Arnold and P.Porter, "Universal Digital portable Communications: A Systems Perspective", *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No.5, June 1987, pp. 764-773.
- [7] T.Perry, "Today's View of Magnetic Fields", *IEEE Spectrum Magazine*, December 1994, pp. 14-23.
- [8] D.Cox, "Portable Digital Radio Communications- An Approach to Tetherless Access", *IEEE Communications Magazine*, July 1989, pp. 30-40.
- [9] J.G.Proakis, *Digital Communications*, McGraw Hill, New York, 2nd edition, 1989.
- [10] H.Hashemi, "The Indoor Radio Propagation Channel", *Proceedings of the IEEE*, vol 81, No. 7 July 1993, pp. 943-967.

- [11] H.Hashemi, "Principles of Digital Indoor Radio Propagation", *Proceedings of IASTED International Symposium on Computers, ELectionics, Communication and Control, CECC '91*, Calgary, April 1991.
- [12] J.D.Parsons and J.G.Gardiner, *Mobile Communications Systems*, Blackie, Glasgow, Scotland, 1989.
- [13] W.Jakes (Jr.) ed., *Microwave Mobile Communications*, Wiley, New York, 1974.
- [14] A.Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw Hill, New York, 1984.
- [15] T.Rappaport and C.McGillem, "UHF Fading in Factories", *IEEE Journal on Selected Areas in Communications*, vol.7 No.1, January 1989, pp. 40-48.
- [16] G.Kalivas, M.El-Tanany and S.Mahmoud, "Millimetre-wave Channel Measurements for Indoor Wireless Communications", *Proceedings of IEEE Vehicular technology Conference, VTC 1992*, Denver, Colorado, May 1992, pp. 609-612.
- [17] A.Saleh and R.Valenzuela, "A Statistical model for Indoor Multipath Propagation", *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, No. 2, Feb. 1987.
- [18] A.Tharek and J.McGeehan, "Indoor propagation and Bit Error rate Measurements at 60 GHz Using Phase Locked Oscillators", *Proceedings IEEE Vehicular Technology Conference*, '88, Philadelphia, June 1988, pp. 127-133.
- [19] T.Rappaport, S.Seidel and K.Takamizawa, "Statistical Channel Impulse Response Models for factory and Open Plan Building Radio Communication System Design", *IEEE Transaction on Communications*, vol. 39, No.5, May 1991, pp. 794-807.
- [20] H.Nikookar and H.Hashemi, "Statistical Modelling of Signal Amplitude Fading of Indoor Radio Propagation Channels", *IUPAC '93*, pp. 84-88.

- [21] M.Handforth, A.Sheikh and M.Abdi, "Statistical Characterization of the Indoor Mobile Radio Channel at 946 MHz", *Proceedings of the International Conference on Universal Personal Communications, ICUPC '93* pp. 955-959.
- [22] D.Laurenson, A.Sheikh and S.McLaughlin, "Characterization of the Indoor Mobile Channel Using a Ray Tracing Technique", *Proceedings IEEE International Conference on Selected Topics in wireless Communications '92*, Vancouver, June 1992, pp. 65-68.
- [23] G.Turin, F.Clapp, T.Johnston, S.Fine and D.Lavry, "A Statistical Model of urban Multipath Propagation", *IEEE Transaction on Vehicular Technology*, vol. VT-21, No. 1, February 1972, pp. 1-9.
- [24] H.Hashemi, "Impulse Response Modelling of Indoor Radio Propagation Channels", *IEEE Journal on Selected Areas in Communications*, vol 11, no 7, Sept. 1993, pp. 967-978.
- [25] S.Todd, M.El-Tanany and S.Mahmoud, "Space and Frequency Diversity Measurements of the 1.75 GHz Indoor Radio Channel for Portable Data terminals and Telephones", *Proceedings IEEE Vehicular Technology Conference, VTC '92*, Denver, May 1992, pp. 613-616.
- [26] R.Bultitude, "Measurement, Characterization and Modelling of Indoor 800/900 MHz Radio Channels for Digital Communications" *IEEE Communications Magazine*, vol. 7, no. 1, pp. 20-30, Jan 1989.
- [27] S.Todd, M.El-Tanany, G.Kalivas and S.Mahmoud, "Indoor Radio Path Loss Comparison between 1.7 GHz and 37 GHz", *ICUPC 1993*, October 1993.
- [28] S.Todd, "L-Band Diversity Receiver with High Dynamic Range I-Q Demodulator", *Internal Document Dept. of Systems and Computer Engineering, Carleton University*, No. SCE-93-03, October 17, 1993.
- [29] F.Massey Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit", *American Sta-*

Statistical Association Journal, March 1951, pp. 68-78.

- [30] G.Boudreau, W.Bird, M.Caskey and R.Hafez, "CT2/CT2Plus Indoor Wireless Propagation Modelling and Range Prediction", *Proceedings of Wireless 94*, Calgary, Alberta, pp. 124-132.
- [31] S.Chatterjee and B.Price, *Regression Analysis by Example*, 2nd ed., John Wiley & Sons, Inc., New York, 1991.
- [32] D.C.Montgomery and E.A.Peck, *Introduction to Linear Regression Analysis*, John Wiley & Sons, Inc., New York, 1982.
- [33] G.W.Snedecor and W.G.Cochran, *Statistical Methods*, The Iowa State University Press, Ames, 5th ed. 1956.
- [34] C.R.Rao, *Linear Statistical Inference and Its Applications*, John Wiley & Sons, Inc., New York, 1965.
- [35] *BBN/Cornerstone Regression Manual: Appendix A*, BBN/Cornerstone Release 1.0, Regression 2529003, May 1993, Appendix A, pp. 76.

END

0|5-06-9|6

FIN