

**UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**DISSERTAÇÃO DE  
MESTRADO**

**UMA METODOLOGIA PARA PREDIÇÃO DO CAMPO ELÉTRICO DE  
RADIODIFUSÃO SONORA EM ONDAS MÉDIAS UTILIZANDO INFERÊNCIAS  
BAYESIANAS**

**JULIANA SANTIAGO MONTEIRO COSTA**

**DM 14/2013**

**BELÉM-PA  
2013**

**UNIVERSIDADE FEDERAL DO PARÁ**  
**INSTITUTO DE TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**JULIANA SANTIAGO MONTEIRO COSTA**

**UMA METODOLOGIA PARA PREDIÇÃO DO CAMPO ELÉTRICO DE  
RADIODIFUSÃO SONORA EM ONDAS MÉDIAS UTILIZANDO INFERÊNCIAS  
BAYESIANAS**

Trabalho submetido à Banca Examinadora do Programa de Pós-graduação em Engenharia Elétrica da UFPA para a obtenção do Grau de Mestre em Engenharia Elétrica, elaborada sob a orientação do Prof. Dr. Gervásio Protásio dos Santos Cavalcante e co-orientação da Profa. Dra. Jasmine Priscyla Leite de Araújo.

**UFPA / ITEC / PPGEE**  
**BELÉM/PA2013**

**UMA METODOLOGIA PARA PREDIÇÃO DO CAMPO ELÉTRICO DE  
RADIODIFUSÃO SONORA EM ONDAS MÉDIAS UTILIZANDO INFERÊNCIAS  
BAYESIANAS**

**JULIANA SANTIAGO MONTEIRO  
COSTA**

Esta dissertação foi julgada adequada para o **Exame de Defesa de Mestrado em Engenharia Elétrica**, na área de Telecomunicações, e aprovado na sua forma final pela banca examinadora designada pelo Programa de Pós-Graduação em Engenharia Elétrica do Instituto de Tecnologia da Universidade Federal do Pará em 11 de junho de 2013.

---

**Prof. Tit. Dr. Gervásio Protásio dos Santos Cavalcante – Orientador –PPGEE/UFPA**

---

**Profa. Dra. Jasmine Priscyla Leite de Araújo -Coorientadora.- Faculdade de  
Computação e Telecomunicações do ITEC/UFPA**

---

**Prof. Dr. Ádamo Lima de Santana – membro – PPGEE/UFPA – MEMBRO**

---

**Prof. Dr. Hermínio Simões Gomes – membro externo –ICEN/UFPA**

**VISTO:**

---

**Prof. Dr. Evaldo Gonçalves Pelaes - Coordenador do PPGEE/ITEC/UFPA**

**UFPA / ITEC / PGEE  
BELÉM/PA  
2013**

Não abandoneis, portanto, a vossa confiança; ela tem grande galardão. Com efeito, tendes necessidade de perseverança, para que, havendo feito a vontade de Deus, alcanceis a promessa.

Hebreus, 10:35,36

Aos meus amados pais, João Otávio e  
Glória, com todo carinho e admiração.

## AGRADECIMENTOS

Agradeço a Deus por me iluminar, por permitir meu acesso ao conhecimento e a uma boa educação, e, principalmente, por ter colocado inúmeros anjos em minha vida que me orientam, me guiam e fazem com que eu busque ser uma pessoa melhor.

Meus preciosos pais, João Otávio e Maria da Glória, por cada palavra de incentivo e cada momento de estímulo. Por acreditar em mim, investir e apostar em meus sonhos, além de sempre mostrar o melhor caminho a ser trilhado.

À minha irmã Lorena, pela amizade, por tornar os meus dias mais alegres e me fazer acreditar que tudo é possível.

Ao meu esposo, amigo e companheiro Thiago, pelo carinho e motivação sempre que tudo parecia muito difícil e distante.

Ao Prof. Gervásio Protásio dos Santos Cavalcante, por ter me recebido em seu grupo de estudos, por sua orientação e confiança. Seu bom humor e seu espírito iluminado foram fundamentais para a concretização deste trabalho.

À Prof. Jasmine Araújo, por sua amizade, confiança e dedicação, dando sempre o seu melhor para as inúmeras ocasiões em que sua ajuda por mim foi solicitada.

À Prof. Josiane Rodrigues, por seu interesse em ser minha primeira orientadora ainda na graduação e que me encorajou a seguir com os estudos e iniciar o Mestrado.

Aos colegas do LCT, em especial aos anjos Amaury Salgado, Allan Braga, Bruno Lira e Hermínio Gomes. Sem a ajuda de vocês eu não teria conseguido. Ao colega Bruno Monteiro, por sua valiosíssima contribuição, tendo cedido os dados utilizados neste trabalho e ótimas referências a serem seguidas.

Aos familiares e aos amigos que estiveram sempre na torcida e orações para que eu conseguisse concretizar este trabalho, em especial Adriana Bastos, Danielle Santos e Márcio Rodrigues.

Todos os profissionais do programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará.

Ao Instituto Nacional de Ciência e Tecnologia, por todo o suporte oferecido ao longo deste trabalho.

## SUMÁRIO

<b>LISTA DE FIGURAS</b> .....	08
<b>LISTA DE TABELAS</b> .....	09
<b>LISTA DE SIGLAS</b> .....	10
<b>RESUMO</b> .....	11
<b>ABSTRACT</b> .....	12
<b>CAPÍTULO 1</b>	
<b>INTRODUÇÃO</b> .....	13
<b>CAPÍTULO 2</b>	
<b>MEIOS DE PROPAGAÇÃO E RADIODIFUSÃO SONORA TERRESTRE</b> .....	17
2.1 Introdução .....	17
2.2 Onda Terrestre.....	18
2.3 Modelos de Propagação .....	21
2.3.1 Modelo Espaço Livre .....	21
2.3.2 Modelo Terra Plana .....	23
2.3.3 Modelo Terra Esférica .....	27
2.4 Conclusão do Capítulo .....	28
<b>CAPÍTULO 3</b>	
<b>REDES BAYESIANAS</b> .....	30
3.1 Introdução .....	30
3.2 Estruturas de uma Rede Bayesiana .....	31
3.3 Inferência em Redes Bayesianas .....	33
3.4 Classificadores Bayesianos .....	34
3.5 Classificador Naive Bayes .....	35
3.6 Conclusão do Capítulo .....	38

## **CAPÍTULO 4**

<b>CAMPANHAS DE MEDIÇÃO E METODOLOGIA.....</b>	<b>40</b>
4.1 Introdução.....	40
4.2 Sistemas de Medições.....	41
4.3 Processamento de Dados.....	44
4.4 Considerações Finais.....	48

## **CAPÍTULO 5**

<b>RESULTADOS – UMA METODOLOGIA PARA A PREDIÇÃO DO CAMPO ELÉTRICO DE RADIODIFUSÃO SONORA EM ONDAS MÉDIAS UTILIZANDO INFERÊNCIAS BAYESIANAS.....</b>	<b>49</b>
5.1 Introdução.....	49
5.2 Metodologia.....	50
5.2.1 Inferências Realizadas .....	53
5.3 Considerações Finais.....	61

## **CAPÍTULO 6**

<b>CONCLUSÃO.....</b>	<b>63</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>66</b>



## LISTA DE FIGURAS

Figura 2.1	Ilustração dos campos que formam as ondas terrestres.....	18
Figura 2.2	Representação do modelo terra plana ou modelo de dois raios.....	24
Figura 2.3	Ilustração do método das imagens.....	25
Figura 2.4	Representação do modelo terra esférica.....	28
Figura 3.1	Exemplo de uma Rede Bayesiana (HECKERMAN, 1997).....	32
Figura 3.2	Rede Bayesiana que representa o classificador Naive Bayes .....	35
Figura 4.1	Rotas em que foram realizadas as medidas de intensidade de campo elétrico (MONTEIRO, 2008).....	40
Figura 4.2	Ilustração da disposição dos equipamentos de medição instalados na UMR da ANATEL (LIMA, 2008).....	42
Figura 4.3	Ilustração da disposição dos equipamentos de medição instalados na viatura da Radiobrás (LIMA, 2008).....	43
Figura 4.4	Representação do terreno da Rota 02.....	44
Figura 4.5	Representação do terreno da Rota 03.....	45
Figura 4.6	Representação do terreno da Rota 04.....	45
Figura 4.7	Representação do terreno da Rota 05.....	46
Figura 4.8	Comportamento do sinal na Rota 02 para 10.000 amostras.....	47
Figura 4.9	Comportamento do sinal na Rota 02 para 700 amostras.....	47
Figura 5.1	Rede Bayesiana gerada para a Rota 02.....	50
Figura 5.2	Rede Bayesiana gerada para a Rota 03.....	50
Figura 5.3	Rede Bayesiana gerada para a Rota 04.....	51
Figura 5.4	Rede Bayesiana gerada para a Rota 05.....	51
Figura 5.5	Script do Matlab com a implementação do Teorema de Bayes .....	53
Figura 5.6	Script do Matlab com a implementação das tabelas de Probabilidade Condicional e Variância .....	54
Figura 5.7	Inferência Bayesiana para Emedido entre 67,83 dB $\mu$ V e 83,4 dB $\mu$ V.....	54
Figura 5.8	Inferência Bayesiana para Emedido entre 89,76 dB $\mu$ V e 104,01 dB $\mu$ V.....	55
Figura 5.9	Campo elétrico estimado para a Rota 02.....	56
Figura 5.10	Campo elétrico medido para a Rota 02 .....	56
Figura 5.11	Campo elétrico estimado x Campo elétrico medido para a Rota 03.....	57
Figura 5.12	Campo elétrico estimado x Campo elétrico medido para a Rota 04.....	58
Figura 5.13	Campo elétrico estimado x Campo elétrico medido para a Rota 05.....	58
Figura 5.14	Linhas de transição para a Rota 02.....	59
Figura 5.15	Linhas de transição para a Rota 03.....	60
Figura 5.16	Linhas de transição para a Rota 04.....	60
Figura 5.17	Linhas de transição para a Rota 05.....	61

## LISTA DE TABELAS

Tabela I	Distâncias máximas de transmissão para ondas MF e HF.....	23
Tabela II	Exemplo de treinamento.....	37
Tabela III	Tabela de probabilidade condicional obtida para a Rota 02.....	54
Tabela IV	Tabela de probabilidade condicional obtida para a Rota 03.....	55
Tabela V	Tabela de probabilidade condicional obtida para a Rota 04.....	55
Tabela VI	Tabela de probabilidade condicional obtida para a Rota 05.....	55
Tabela VII	Dados calculados para a Rota 02.....	57
Tabela VIII	Dados calculados para as rotas 03, 04 e 05.....	59

## LISTA DE SIGLAS

AIC	AKAIKE Information Criteria.
AM	Amplitude Modulada.
ANATEL	Agência Nacional de Telecomunicações.
BIC	Bayesian Information Criteria.
CD	Compact Disc.
CPT	Conditional Probability Table.
DAB	Digital Audio Broadcasting.
DRM	Digital Radio Mondiale.
E	Campo Elétrico.
ELF	Extremely Low Frequency.
EMQ	Erro Médio Quadrático.
FI	Frequência Intermediária.
FM	Frequência Modulada.
GPS	Global Positioning System.
HD	High Definition.
HF	High Frequency.
IBOC	In Band On Channel.
ISDB-Tn	Integrated Services Digital Broadcasting – Terrestrial narrowband
ITU-R	International Telecommunications Union Radiocommunications.
LF	Low Frequency.
MD	Mineração de Dados.
MF	Medium Frequency.
RB	Redes Bayesianas.
RF	Radiofrequencia.
SBRD	Sistema Brasileiro de Rádio Digital.
TPC	Tabela de Probabilidade Condicional.
UIT	União Internacional de Telecomunicações.
UMR	Unidade Móvel de Radiomontagem.
VHF	Very High Frequency.
VLF	Very Low Frequency.
WEKA	Waikato Environment for Knowledge Analysis.

## RESUMO

A adoção de sistemas digitais de radiodifusão sonora, que estão em fase de testes no país, permite realizar novos estudos visando um melhor planejamento para a implementação dessas novas emissoras. O que significa reavaliar os principais modelos de radiopropagação existentes ou propor novas alternativas para atender as demandas inerentes dos sistemas digitais. Os modelos atuais, conforme Recomendações ITU-R P. 1546 e ITU-R P. 1812, não condizem fielmente com a realidade de algumas regiões do Brasil, principalmente com as regiões de clima tropical, como a Região Amazônica, seja pelo elevado índice pluviométrico seja pela vasta flora existente. A partir dos modelos adequados ao canal de propagação, torna-se viável desenvolver ferramentas de planejamento de cobertura mais precisas e eficientes. A utilização destas ferramentas é cabível tanto para a ANATEL, para a elaboração dos planos básicos de distribuição de canais quanto para os radiodifusores.

No presente trabalho é apresentada uma metodologia utilizando a inteligência computacional, baseada em Inferências Bayesianas, para predição da intensidade de campo elétrico, a qual pode ser aplicada ao planejamento ou expansão de áreas de cobertura em sistemas de radiodifusão para frequências na faixa de ondas médias (de 300 kHz a 3MHz). Esta metodologia gera valores de campo elétrico estimados a partir dos valores de altitude do terreno (através de análises de tabelas de probabilidade condicional) e estabelece a comparação destes com valores de campo elétrico medidos.

Os dados utilizados neste trabalho foram coletados na região central do Brasil, próximo à cidade de Brasília. O sinal transmitido era um sinal de rádio AM transmitido na frequência de 980 kHz. De posse dos dados coletados durante as campanhas de medição, foram realizadas simulações utilizando tabelas de probabilidade condicional geradas por Inferências Bayesianas.

Assim, é proposto um método para predizer valores de campo elétrico com base na correlação entre o campo elétrico medido e altitude, através da utilização de inteligência computacional. Se comparados a inúmeros trabalhos existentes na literatura que têm o mesmo objetivo, os resultados encontrados neste trabalho validam o uso da metodologia para determinar o campo elétrico de radiodifusão sonora em ondas médias utilizando Inferências Bayesianas.

***Palavras-chave: Campo elétrico; redes Bayesianas; probabilidade condicional, estimado, altitude.***

## ABSTRACT

The adoption of digital sound broadcasting systems, which are under testing in the country, allows new studies aimed a better planning for the implementation of new stations, which means to reassess the major existing radio propagation models or propose new alternatives to meet demands inherent in digital systems. The current models, as Recommendations ITU-R P. 1546 and ITU-R P. 1812, do not match closely with the reality of some regions of Brazil, especially in the tropical regions, such as the Amazon Region, due to the high rainfall and the vast existing flora. Using models suited to the propagation channel, it becomes feasible to develop planning tools covering most accurate and efficient. The use of these tools is applicable both to ANATEL, for the elaboration of the basic plans, as distribution channels for broadcasters.

This paper presents a methodology using a computational intelligence based in Bayesian Networks for prediction of electric field intensity, which can be applied to planning or expanding coverage areas in broadcasting systems for frequencies in the range of medium wave (300 kHz to 3 MHz). This methodology generates electric field values estimated from the values of terrain altitude (through analysis of conditional probability tables) and provides a comparison of these values with the measured electric field.

The data used in this study were collected in Brazil's central region, nearby the city of Brasilia. The transmitted signal was an AM radio signal transmitted at a frequency of 980 kHz. With the data collected during the measurement campaigns, simulations were performed using conditional probability tables generated by Bayesian Networks.

Thus, it's proposed a method for predicting values of electric field based on the correlation between the measured electric field and the altitude through the use of artificial intelligence. Compared to numerous studies in the literature that have the same goal, the results found in this study validate the use of the methodology to determine the electric field in medium wave radio broadcasting using Bayesian Networks.

***Keywords: Electric Field, Bayesian Networks, Conditional Probability, Estimated, Altitude.***

# CAPÍTULO 1

## INTRODUÇÃO

Desde as primeiras transmissões, no início do século XX, a radiodifusão em amplitude modulada (AM) tem sido amplamente utilizada, garantindo o acesso à informação para os mais variados locais do território nacional, chegando esta, inclusive, a ser a única fonte de informação disponível em algumas localidades mais remotas. Com o rápido advento das telecomunicações, onde um novo padrão de tecnologia se torna defasado em períodos cada vez mais curtos, este serviço está perdendo espaço para os novos serviços de comunicações que vem sendo desenvolvidos, tais como a radiodifusão em frequência modulada, televisão e internet. Na tentativa de mudar este panorama e chamar novamente a atenção do público aos serviços de radiodifusão, estão sendo estudados procedimentos para a digitalização do sinal, permitindo manter a abrangência da cobertura nacional associada à qualidade superior dos serviços oferecidos com tecnologia analógica.

São inúmeras as vantagens no que diz respeito à variedade de serviços oferecidos pela digitalização do sinal, como programar o rádio para sintonizar sua emissora favorita e escutar suas músicas prediletas ao som de CD, receber em um display acoplado ao seu aparelho de rádio o boletim meteorológico da sua cidade e, junto a isso, lhe indicar junto a um mapa qual o melhor trajeto, livre de congestionamentos. Tecnicamente, as vantagens da radiodifusão digital estão no aumento na capacidade de transmissão do sinal, o que permite o uso mais eficiente do espectro de radiofrequências e a redução da potência transmitida levando em consideração a mesma área de cobertura do sinal analógico. Esses benefícios representam uma grande melhoria, também, na qualidade do sinal transmitido, bem como um aumento na diversidade das transmissões por parte das emissoras, conforme exemplos citados anteriormente.

São três os modelos do sistema de transmissão terrestre e por satélite conhecidos hoje no mundo: o IBOC (*In-Band On Channel*), sistema norte-americano também

conhecido comercialmente por HD Rádio; os sistemas europeus Eureka 147 DAB (*Digital Audio Broadcasting*), DAB+ e DRM (*Digital Radio Mondiale*); e o ISDB-Tn (*Services Digital Broadcasting – Terrestre narrowband*), sistema japonês, que é o mais recente dentre os três.

No Brasil, o rádio digital começou a ser pensado e testado em meados da década de 2000. Em março de 2010, o então Ministro das Comunicações Hélio Costa publicou a portaria 290 que institui o SBRD: Sistema Brasileiro de Rádio Digital, dando disposições sobre os pré-requisitos que o sistema deveria atender. Porém, não foi escolhido ainda o padrão que será utilizado, esbarrando nos interesses dos pequenos e grandes radiodifusores, bem como das rádios comunitárias, que não teriam recursos suficientes pra realizar essa alteração.

Com base em vários testes realizados, os radiodifusores e a ANATEL (Agência Nacional de Telecomunicações) têm preferência pelos sistemas que permitem que a transmissão do sinal digital seja realizada no mesmo canal do sinal analógico atual, garantindo menor custo de implantação. Com estes requisitos estão os padrões DRM (*Digital Radio Mondiale*) e o IBOC (*In Band On Channel*). Ambos permitem a operação com largura de banda compatível com a utilizada no Brasil e possibilitam a coexistência dos sinais analógicos e digitais para a radiodifusão sonora.

O DRM é um padrão de rádio digital desenvolvido por um consórcio global de nome DRM, com representações em vários países. Trata-se de um padrão aberto, sendo o único reconhecido pela UIT (União Internacional de Telecomunicações) que pode funcionar em todas as bandas de radiodifusão sonora terrestre: Ondas Médias, Ondas Tropicais, Ondas Curtas e o VHF (faixa das rádios FM). Hoje, é um padrão em teste em vários lugares do mundo e encontra-se em fase de implantação na Rússia e na Índia.

O IBOC foi desenvolvido por uma empresa chamada Ibiquity. Esse padrão não tem modo de operação para transmitir na faixa de Ondas Curtas e é conhecido por possuir segredos industriais em sua norma [25]. Além dos Estados Unidos, apenas o México adotou recentemente este padrão.

Nos sistemas de comunicações sem fio a predição das perdas de propagação se torna bastante difícil e complexa, devido às diferentes características observadas no ambiente encontrado entre o transmissor e o receptor, tais como: diferentes estruturas arquitetônicas, arborização variada, diferentes tipos de relevo, propriedades elétricas dos

meios (solo e ar), além dos fenômenos básicos de propagação que as ondas eletromagnéticas estão sujeitas, como difração, reflexão e espalhamento. A fim de otimizar a área de cobertura e a qualidade do sinal, é de fundamental importância o conhecimento prévio da região a ser coberta. Nos casos em que não pode ocorrer linha de visada, isto é, a energia radiada pela antena transmissora atinge a antena receptora com a interceptação de obstáculos, a modelagem matemática detalhada do canal é muito importante para a predição da intensidade do sinal, de forma a mitigar os erros com as aproximações normalmente utilizadas nos cálculos dos efeitos causados pelo canal [1].

O conhecimento dos parâmetros elétricos do solo, condutividade e permissividade, é fundamental para a determinação da zona de cobertura das estações de radiodifusão que operam na faixa de 525kHz a 1705kHz, uma vez que há predominância da onda de superfície, cuja atenuação está relacionada às características do solo[2]. São inúmeros os modelos teóricos de propagação que auxiliam na obtenção dos valores de intensidade do campo elétrico, porém, geralmente, não garantem a precisão necessária para o planejamento ou adaptação dos sistemas de comunicação de rádio, principalmente devido às peculiaridades e desníveis que existem em cada terreno.

Para frequências na faixa de ondas médias (de 300 kHz a 3MHz), que é o caso em análise neste trabalho, predomina a presença, durante o dia, das ondas de superfície, cujas características de propagação são intensamente influenciadas pela terra [3]. Sistemas analógicos e digitais possuem os mesmos procedimentos para avaliação do campo elétrico, tendo neste uma maior dificuldade pois, assim como na TV Digital, se a intensidade do sinal recebido for inferior ao limite imposto para qualidade na recepção, o serviço fica indisponível ao usuário

Sendo assim, para o planejamento e implantação de um sistema de radiodifusão, é importante conhecer o valor real do sinal a ser propagado em condições determinadas, e a intensidade de campo elétrico, no intuito de se calcular com as menores possibilidades de erro a potência transmitida e a se delimitar a área de cobertura do sinal pelos radiodifusores.

Existe hoje a possibilidade de se desenvolver ferramentas que auxiliam, de forma bastante precisa, na obtenção do campo elétrico a partir apenas de características topográficas do ambiente de propagação, sem a necessidade de se conhecer todos os parâmetros e características elétricas do terreno em análise.



Este trabalho propõe uma comparação entre valores de campo (E) medidos durante campanhas de medição com os valores obtidos a partir de simulação por inteligência computacional, considerando uma mesma área de cobertura. Estas simulações foram realizadas utilizando tabelas de probabilidade condicional geradas por Rede Bayesiana. Assim, é proposta uma metodologia para determinar a correlação entre o campo elétrico e altitude através da utilização da técnica de medição e de inteligência computacional.

As Redes Bayesianas são ferramentas bastante adequadas para situações em que se lida com problemas que não são explicáveis logicamente, mas sim probabilisticamente. Na maioria dos casos, as técnicas deste tipo de inteligência artificial estão apoiadas em dados, onde os mesmos são tratados como uma representação compacta de uma tabela de conjunção de probabilidades do universo do problema. Entretanto, em muitos casos, sob o ponto de vista de um especialista, Redes Bayesianas constituem um modelo de grafos que representa de forma simples as relações de causalidade das variáveis de um sistema. Uma das principais vantagens das Redes Bayesianas é permitir a coexistência, na mesma rede, de diversas fontes de conhecimento, independentemente da sua natureza. Outra vantagem bastante relevante é a possibilidade de utilizar dados incompletos, questão frequentemente enfrentada por pesquisadores e analistas.

Este trabalho está dividido da seguinte maneira: no capítulo 2 é apresentada a fundamentação do problema, apontando uma breve revisão sobre propagação das ondas de radiofrequência, explorando a propagação das ondas de superfície e apresentando trabalhos correlatos que serviram como referência para a elaboração deste trabalho. No capítulo 3, explica-se o funcionamento das redes Bayesianas, como estas foram utilizadas para a solução do problema e como foram utilizados os dados experimentais, a fim de ser determinada a metodologia. O capítulo 4 aborda sobre as campanhas de medição em que os dados foram coletados, bem como os tratamentos pelos quais os dados precisaram sofrer para que se adequassem à metodologia utilizada. Os resultados são apresentados no capítulo 5, onde são explicados os parâmetros levados em consideração, bem como são estabelecidas as comparações a fim de se validar o trabalho em questão. Por fim, no capítulo 6, é realizada, de forma consolidada, a conclusão do trabalho e também são indicados os possíveis pontos para continuidade da pesquisa.

## **CAPÍTULO 2**

### **MEIOS DE PROPAGAÇÃO E RADIODIFUSÃO SONORA TERRESTRE.**

#### **2.1 INTRODUÇÃO**

Depois de muitos anos de pesquisa a partir do início da difusão das ondas de rádio, ficou clara a influência que os meios de transmissão e suas propriedades têm no desempenho da propagação do sinal. Pesquisas ao longo do tempo vieram mostrando quais as variáveis que mais afetam o sinal e sua propagação, e os principais fatores que interferem em sua qualidade. Os estudos destas propriedades e variáveis permite a otimização dos projetos de implantação de sistemas de radiopropagação.

Dentre os principais fenômenos responsáveis pela perda de propagação do sinal, os mais conhecidos são aqueles que dependem do tipo de relevo em que o sinal se propaga (se há montanhas, por exemplo) e na intensidade e altura de edificações. Esses fatores são os maiores responsáveis pelas perdas por reflexões, difrações e espalhamentos. Quanto às reflexões no solo, estas dependem, além da sua formação física, das suas características eletromagnéticas e de sua rugosidade em relação ao comprimento de onda.

A onda pode sofrer mudanças em sua trajetória em função de refrações na atmosfera, de difrações em obstáculos ou, ainda, de espalhamentos na superfície do solo. A frequência, a umidade da atmosfera ou mesmo a vibração das moléculas do meio podem contribuir de maneira substancial para a atenuação da amplitude do campo na onda eletromagnética [5]. Além destes fatores devem-se considerar ainda a presença de gases, naturais ou poluentes e partículas de água que se concentram na troposfera, região mais próxima da terra.

A frequência utilizada para transmissão, o direcionamento das antenas envolvidas, a distância entre as antenas transmissora e receptora e a distância das mesmas em relação ao solo afetam diretamente no tipo de influência do ambiente sobre o canal de radiopropagação.

## 2.2 ONDA TERRESTRE

O processo básico de transmissão do sinal consiste na antena transmissora irradiar energia a qual atinge a antena receptora. Como visto anteriormente, para que o sinal chegue à antena receptora ele passa por multipercursos, como se pode observar na Fig.1. A onda terrestre é o conjunto formado pelas ondas com trajetórias próximas ao solo.

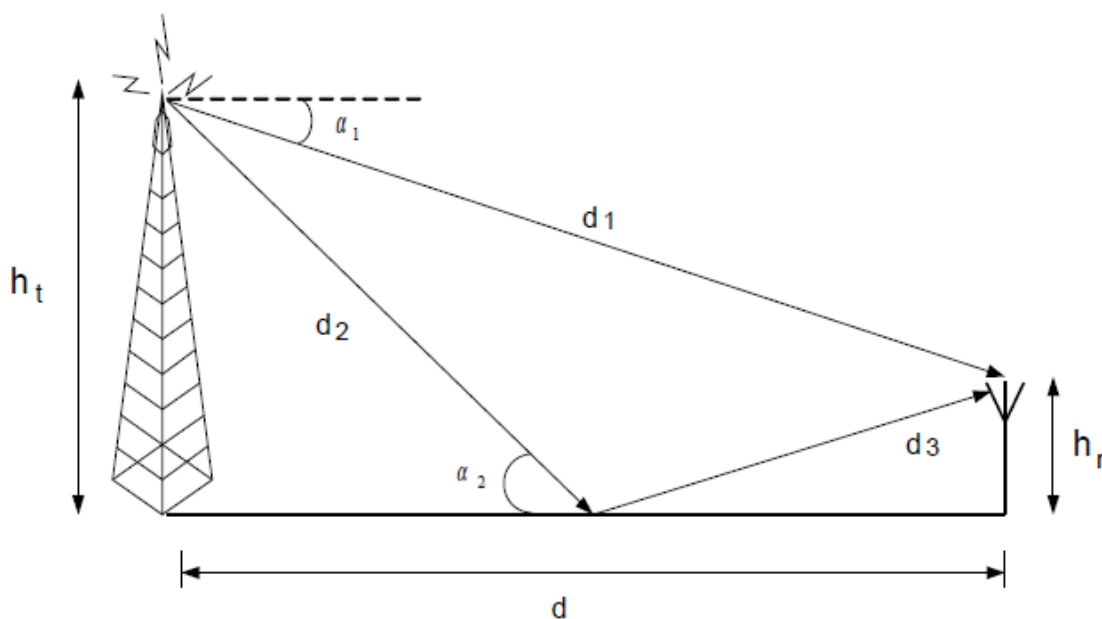


Figura 2.1 – Ilustração dos campos que formam as ondas terrestres

As ondas terrestres aproveitam a condutividade da superfície terrestre ou marítima e são encontradas normalmente em comunicações com alcance médio de 1.000 km, utilizando-se transmissores de alta potência. Sua atenuação é baixa, tornando a transmissão confiável.

Considera-se, em geral, que a onda terrestre é formada por duas componentes, uma onda espacial e outra de superfície. A onda espacial se propaga no espaço imediatamente acima da superfície terrestre por dois caminhos: um que interliga diretamente transmissor e receptor e o outro que ocorre por reflexão na terra antes de alcançar o receptor, sendo que estes dois percursos tem comprimentos diferentes. A onda de superfície é formada pela energia que se propaga guiada pela superfície terrestre de forma semelhante às ondas guiadas em linhas de transmissão, com uma

atenuação dependem das características do terreno e da frequência de transmissão [2]. As ondas de superfície se caracterizam por acompanhar a superfície da terra atingindo longas distâncias. Estas ondas induzem correntes na superfície da terra sobre a qual elas passam, o que induz uma perda por absorção. Essas ondas aparecem em frequências mais baixas, nas faixas de VLF, LF e MF (e algumas aplicações especiais para as bandas ELF e HF). Esse, portanto, é o modo de propagação utilizado na transmissão da faixa de AM comercial que compreende um intervalo de frequência entre 525 a 1705 kHz.

Em frequências baixas, as ondas eletromagnéticas tendem a não se dispersar no espaço, ficando concentradas nas proximidades da superfície que a guiará até o receptor. Estando o solo se comportando como um material condutor, as condições de contorno impõem que o campo elétrico seja praticamente perpendicular à superfície da Terra. Se o solo fosse um condutor perfeito, em suas proximidades só seria possível a existência do campo elétrico na vertical[3]. Em consequência deste fato, as comunicações devem ser feitas, de maneira geral, com ondas polarizadas verticalmente. As ondas polarizadas horizontalmente não são utilizadas em baixas frequências devido à elevada atenuação que estas sofrem mesmo para pequenas distâncias. [3]

A onda de superfície desloca-se ao longo do contorno do solo, excitada por uma antena transmissora localizada imediatamente sobre o solo ou próximo dele, e o campo decresce à medida que a onda se afasta da antena transmissora devido às correntes induzidas provocarem dissipação de potência por Efeito Joule.

O campo elétrico  $E_s$  da onda de superfície pode ser representado pela equação:

$$E_s = (1 - \Gamma)A_s f_s(\theta, \phi) e^{-j|kr|} \quad (2.1)$$

sendo  $f_s$  uma função do diagrama de radiação da antena, com  $\theta$  e  $\phi$  indicando sua direção de radiação.  $\Gamma$  o coeficiente de reflexão no solo,  $A_s$  um fator que afeta a amplitude da onda de superfície, que varia de acordo com características eletromagnéticas do solo e da frequência,  $k$  (s/m) a constante de propagação e  $r$  a distância entre o transmissor e o receptor.

Simplificando a equação do campo elétrico da onda de superfície, pode-se considerar que o seu módulo seja obtido a partir da equação:

$$E \cong \frac{E_0}{r} A \quad (2.2)$$

Com  $E_0$  sendo considerado o campo elétrico no espaço livre e  $A$  representando o fator de atenuação da onda de superfície. Como citado anteriormente, a atenuação da onda é causada pela dissipação de potência no meio, e pode ser obtida pela equação:

$$A \cong \frac{1}{2p - 3,7} \quad (2.3)$$

Considerando-se:

$$p \cong \frac{\pi r \cos^2 b''}{\lambda x \cos b'} , \quad (2.4)$$

onde:

$$x = 60\lambda\sigma$$

$$b' = \text{arctg} \left( \frac{\varepsilon_r - 1}{x} \right)$$

$$b'' = \text{arctg} \left( \frac{\varepsilon_r}{x} \right) \quad (2.5)$$

onde  $\sigma$  é a condutividade elétrica e  $\varepsilon_r$  é a permissividade relativa do meio.

As equações expostas mostram que para se obter o campo elétrico das ondas de superfície faz-se necessário o conhecimento de algumas características eletromagnéticas do meio, que pode variar de acordo com a umidade do terreno, sua salinidade, condutividade, permissividade e sua composição geológica.

## 2.3 MODELOS DE PROPAGAÇÃO

Os modelos de propagação são expressões matemáticas que tentam, por aproximação, simular as características das ondas eletromagnéticas durante a propagação em um determinado meio. Os modelos de propagação são classificados como empíricos (os quais são baseados em medidas e dados estatísticos), determinísticos, que se utilizam de alguma formulação da teoria eletromagnética, ou ainda modelos mistos, que são caracterizados por mesclar teoria e dados estatísticos derivados de medições [6].

Os modelos de propagação podem prever a amplitude média do sinal recebido de acordo com a distância do transmissor e com a variação da amplitude do sinal, como por exemplo, estimar a área de cobertura de um sinal. Quando o objetivo do modelo é este, diz-se que o modelo de propagação é de larga escala. Os modelos de propagação de pequena escala calculam as variações do sinal recebido considerando-se pequenos percursos ou curtos intervalos de tempo.

### 2.3.1 - MODELO ESPAÇO LIVRE

O modelo de propagação no espaço livre é utilizado para prever a intensidade com a qual um sinal transmitido chega ao seu receptor quando as duas antenas apresentam-se em linha de visada direta, isto é, não são obstruídas por nenhum obstáculo. O modelo diz que a potência que chega ao receptor decai à medida que a distância o transmissor e o receptor aumenta. Além disso, fatores como o ganho de ambas as antenas, potência transmitida e perdas sofridas durante a propagação influenciam diretamente no cálculo da intensidade do sinal que é recebido. A potência recebida pela antena receptora pode ser calculada por meio da Fórmula de Friis abaixo:

$$P_R = P_T G_T G_R \left( \frac{\lambda}{4\pi d} \right)^2, \quad (2.6)$$

sendo  $P_T$  a potência transmitida,  $P_R$  a potência recebida,  $G_T$  o ganho da antena transmissora,  $G_R$  o ganho da antena receptora,  $d$  é a distância entre as antenas transmissora e receptora e é dada em metros, e  $\lambda$  é o comprimento de onda da onda

portadora, também dado em metros.

O ganho de uma antena está diretamente relacionado a sua abertura efetiva ( $A_e$ ), e estão relacionados por:

$$G = \frac{4\pi A_e}{\lambda^2} \quad (2.7)$$

A abertura efetiva da antena,  $A_e$ , está relacionada ao comprimento de onda  $\lambda$ , o qual está relacionado à frequência da portadora através de:

$$\lambda = \frac{c}{f} = \frac{2\pi c}{\omega_c} \quad (2.8)$$

Sendo  $f$  a frequência da portadora,  $\omega_c$  a frequência angular da portadora em radianos por segundo e  $c$  a velocidade da luz (em m/s).

Seguindo a equação (2.6) percebemos que a potência recebida decai com o quadrado da distância entre o transmissor e o receptor, o que quer dizer que a potência recebida decai com a distância a uma taxa de 20 dB/década. A perda de propagação para o modelo do espaço livre é dada, então, por

$$PL(\text{dB}) = 10 \log \frac{P_T}{P_R} \quad (2.9)$$

Vale ressaltar que a equação 2.9 só é usada nos casos em que o receptor esteja localizado em uma região de campo distante em relação ao transmissor. A distância mínima,  $d_f$ , para se considerar a situação de campo distante pode ser calculada através da equação a seguir

$$d_f > \frac{2D^2}{\lambda} \quad (2.10)$$

Sendo  $D$  a maior dimensão linear da antena e  $\lambda$  o comprimento de onda. Para estar localizado numa região de campo distante,  $d_f$  deve ser:

$$d_f \gg D \quad (2.11)$$

$$d_f \gg \lambda \quad (2.12)$$

## 2.3.2 – MODELO TERRA PLANA

Também conhecido como Modelo de Dois Raios, o modelo Terra Plana se aplica a situações usuais, onde a onda de superfície se desloca ao longo do contorno do solo. Baseia-se na ótica geométrica e, diferente do Modelo Espaço Livre, considera tanto a onda direta como a onda refletida na superfície do solo. Este modelo foi considerado razoavelmente preciso para predição da amplitude do sinal transmitido em situações com distâncias de vários quilômetros entre o transmissor e o receptor, limitado pela expressão:

$$d \leq \frac{66,94}{\sqrt[3]{f}}$$

Com  $d$  em Km e  $f$  em MHz. Além desta distância, a curvatura da Terra começa a ter efeito significativo sobre a intensidade do campo.

Pode-se relacionar a distância máxima de transmissão de acordo com a frequência pela tabela abaixo:

Tabela I: Distâncias máximas de transmissão para ondas MF e HF

	Frequência	Distância Máxima
MF	300 KHz	100 Km
	3000 KHz	46,4 Km
HF	3 MHz	46,4 Km
	30 MHz	21,5 Km

Além da distância entre transmissor e receptor, este modelo também se adapta a situações em que a antena transmissora está instalada em torres altas (geralmente acima de 50 m do solo).



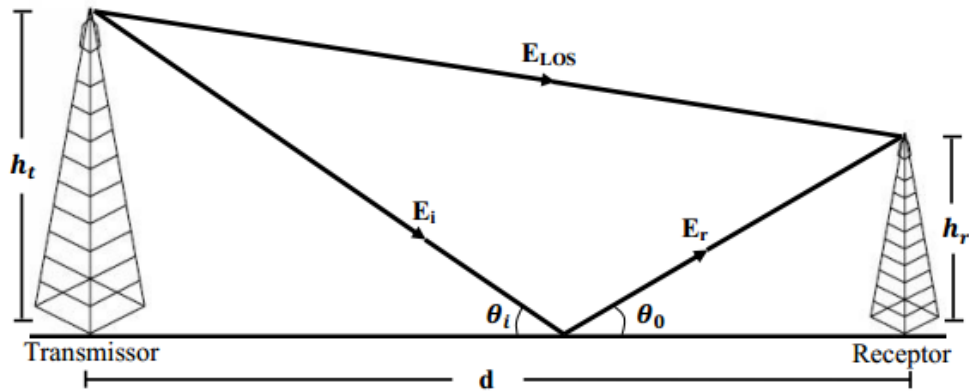


Figura 2.2: Representação do modelo terra plana ou modelo de dois raios

A intensidade de campo elétrico (V/m) com base na Figura 2.2, na condição de espaço livre e a uma distância  $d$  (m) do transmissor é dada por:

$$E(d, t) = \frac{E_0 d_0}{d} \cos\left(\omega_c \left(t - \frac{d}{c}\right)\right) \quad (2.13)$$

Sendo  $E_0$  a intensidade do campo elétrico no espaço livre em V/m,  $d_0$  a distância referencial do transmissor em m, o termo  $\frac{E_0 d_0}{d}$  representa o módulo do campo elétrico a uma distância  $d$ ,  $t$  é o tempo em segundos,  $\omega_c$  é a frequência, em rad/s, da portadora e  $c$  é a velocidade da luz em m/s. Assim, tem-se para intensidade de campo elétrico proveniente do raio direto:

$$E_{LOS}(d', t) = \frac{E_0 d_0}{d'} \cos\left(\omega_c \left(t - \frac{d'}{c}\right)\right) \quad (2.14)$$

e proveniente de raio refletido:

$$E_g(d'', t) = \frac{E_0 d_0}{d''} \cos\left(\omega_c \left(t - \frac{d''}{c}\right)\right) \quad (2.15)$$

Sendo  $d'$  a distância em linha de visada e  $d''$  a distância após reflexões.

Com o coeficiente de reflexão para o solo  $\tau$  assumindo um valor de  $\tau = -1$  para o caso de reflexão perfeita, a intensidade total do campo elétrico  $E_{TOT}$ , no receptor, será a soma vetorial da componente direta com a refletida:

$$|E_{TOT}| = |E_{LOS}| + |E_g| \quad (2.16)$$

que pode ser escrito, portanto, da seguinte forma:

$$E_{total}(d,t) = \frac{E_0 d_0}{R_1} \cos\left(\omega_c\left(t - \frac{R_1}{c}\right)\right) - \frac{E_0 d_0}{R_2} \cos\left(\omega_c\left(t - \frac{R_2}{c}\right)\right) \quad (2.17)$$

Usando o método das imagens, representado na Figura 3, pode-se calcular o valor da diferença de percurso entre a onda direta e a onda refletida.

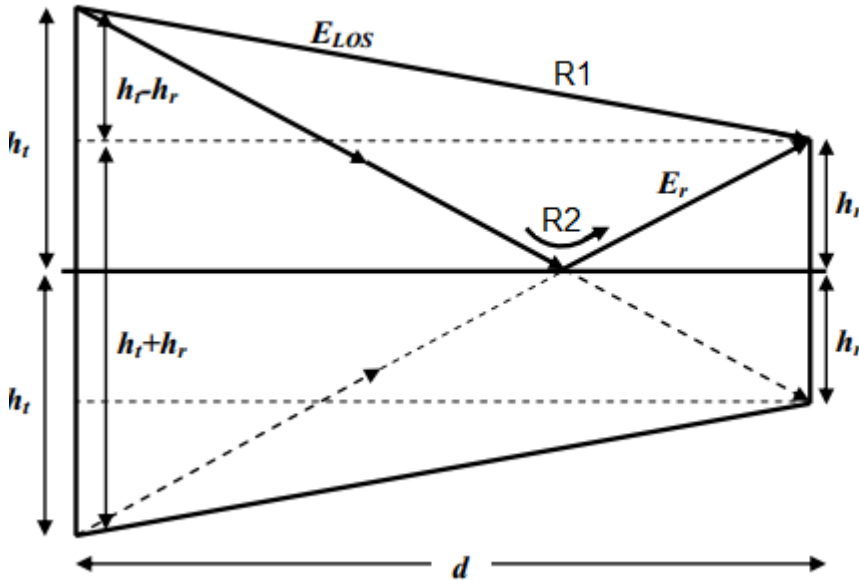


Figura 2.3: Ilustração do método das imagens

Tem-se então que:

$$\Delta = R_2 - R_1 = \sqrt{(h_T + h_R)^2 + d^2} - \sqrt{(h_T - h_R)^2 + d^2} \quad (2.18)$$

sendo  $R_1$  a distância percorrida pela onda direta,  $R_2$  a distância total percorrida pela onda refletida,  $d$  a distância entre as antenas,  $h_T$  a altura da antena transmissora e  $h_R$  a altura da antena receptora.

Se a distância entre o transmissor e o receptor for muito maior que as alturas das antenas, a equação (2.18) pode ser simplificada, usando série de Taylor, da seguinte forma:

$$\Delta = R_2 - R_1 = \frac{2h_T h_R}{d} \quad (2.19)$$

Após o cálculo do valor da diferença de percurso, pode-se calcular a diferença de fase entre as duas componentes do campo elétrico a partir da equação:

$$\theta_\Delta = \frac{2\pi\Delta}{\lambda} = \frac{\Delta\omega_c}{c} \quad (2.20)$$

Fazendo algumas aproximações matemáticas e usando identidades trigonométricas, pode-se comprovar que:

$$|E_{total}(d)| = 2 \frac{E_0 d_0}{d} \text{sen}\left(\frac{\theta_\Delta}{2}\right) \quad (2.21)$$

Para  $\frac{\theta_\Delta}{2}$  menor que 0,3 radianos, pode-se considerar que o valor do seno é o próprio argumento, ou seja,  $\frac{\theta_\Delta}{2}$ . Usando essa simplificação e as equações (2.19) e (2.20), pode-se reescrever a equação (2.19) como:

$$E_{total}(d) = 2 \frac{E_0 d_0}{d} \frac{2\pi h_T h_R}{\lambda d} \left[ \frac{V}{m} \right] \quad (2.22)$$

No espaço livre, a densidade do fluxo de potência pode ser obtida por meio de

$$P_d = \frac{EIRP}{4\pi d^2} = \frac{P_T G_T}{4\pi d^2} = \frac{|E|^2}{\eta} = \frac{|E|^2}{120\pi} \quad [\text{W/m}^2], \quad (2.23)$$

e a potência recebida,  $P_R$ , pode ser obtida a partir da expressão

$$P_R = P_d A_e = \frac{|E|^2}{120\pi} A_e = \frac{P_T G_T G_R \lambda^2}{(4\pi)^2 d^2} = \frac{|E|^2 G_R \lambda^2}{480\pi} \quad [\text{W}] \quad (2.24)$$

Utilizando as equações (2.22), (2.23) e (2.24), é possível expressar a potência recebida por:

$$P_R = P_T G_T G_R \frac{h_T^2 h_R^2}{d^4} \quad (2.25)$$

A equação (2.25) mostra que para grandes distâncias entre transmissor e receptor, tal que  $d \gg \sqrt{h_T h_R}$ , a potência recebida decai com a distância elevada a quarta potência, ou a uma taxa de 40 dB/década, uma atenuação muito maior que aquela apresentada pelo modelo espaço livre (20 dB/década).

### 2.3.3 – MODELO TERRA ESFÉRICA

O Modelo Terra Esférica foi desenvolvido para solucionar uma falha encontrada no Modelo Terra Plana, o qual não responde com precisão em enlaces muito longos por não considerar a curvatura da Terra, diminuindo a intensidade do campo refletido já que tende a acentuar a divergência de seu feixe de raios. As diferenças do Modelo Terra Esférica em relação ao Modelo Terra Plana estão na descrição geométrica das distâncias e na inserção de um fator de divergência ( $D < 1$ ) na expressão do campo refletido, o que resulta na sua atenuação, como segue:

$$E_R = E_D \left( 1 + D\Gamma e^{-jk(R_2 - R_1)} \right) \quad (2.26)$$

sendo  $E_R$  o campo elétrico refletido,  $E_D$  o campo elétrico direto,  $\Gamma$  o coeficiente de reflexão e  $k$  a constante de propagação da onda. Na Figura 4 pode-se perceber o problema do Modelo Terra Esférica.

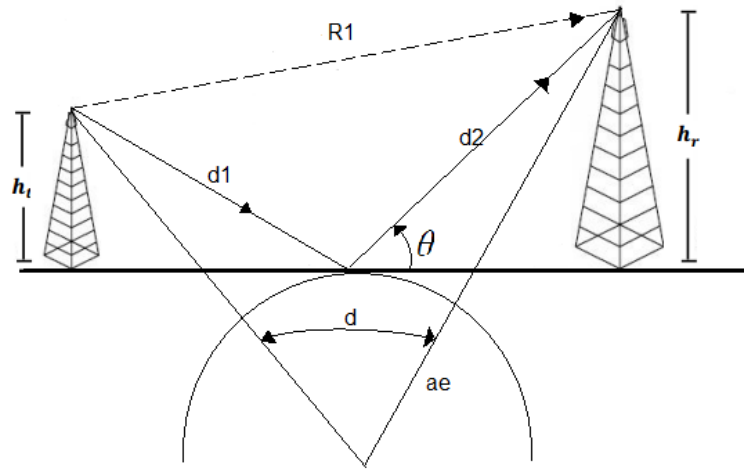


Figura 2.4: Representação do modelo terra esférica

Sendo que  $a_e$  representa o raio da Terra (aproximadamente 6.400 km) e  $d$  refere-se ao comprimento do enlace ao longo da superfície. Para grandes comprimentos de enlace, pode-se considerar que o fator de divergência  $D$  para os campos direto e refletido é o mesmo e pode ser calculado por meio da equação:

$$D = \frac{1}{\sqrt{1 + \frac{2d_1d_2}{a_e \sin \theta (d_1 + d_2)}}} \quad (2.27)$$

Com  $d_1$  e  $d_2$  sendo as distâncias percorridas pela onda refletida, e  $\theta$  o ângulo de reflexão.

## 2.4 CONCLUSÃO DO CAPÍTULO

Neste capítulo ficou clara a difícil análise do comportamento das ondas de superfície em ambientes reais, mesmo havendo muito estudo com diversas técnicas numéricas e aproximações analíticas focados neste problema.

Foram expostos os Modelos de Predição de Propagação da onda terrestre que

consideram a terra plana e a terra esférica para faixas de frequências inferiores a 30MHz.

O próximo capítulo fará uma breve explicação sobre redes bayesianas e seus benefícios em relação a outros tipos de inteligência. Irá mostrar sobre sua estrutura, do que são formadas, organizadas e como são construídas.

Além disso, será demonstrado o que são inferências bayesianas, quais são seus principais métodos de aprendizagem e a definição do classificador Naive Bayes, utilizado neste trabalho.

## CAPÍTULO 3

### CLASSIFICADOR BAYESIANO NAÏVE.

#### 3.1 INTRODUÇÃO

Com o aumento da quantidade de dados a serem manipulados e a limitação humana em relação à capacidade de analisar estes dados, sentiu-se a necessidade de atribuir essas funções aos sistemas de informática, surgindo então as soluções computacionais. Estas soluções são, geralmente, compostas por um conjunto de métodos, técnicas e ferramentas de áreas como bancos de dados, estatística, e inteligência computacional [8]. São exemplos de soluções computacionais as técnicas de Mineração de Dados (MD), que utiliza técnicas de estatística, inteligência artificial, recuperação de informação e reconhecimento de padrões. As técnicas de Mineração de Dados consistem em explorar grandes volumes de dados buscando padrões consistentes, como regras de associação ou sequências temporais, para detectar relacionamentos sistemáticos entre variáveis, detectando novas partições dos conjuntos de dados.

Um dos grandes problemas encontrados no processo de MD é que, em muitos casos reais, não há informações completas sobre o ambiente (seja por falha na coleta dos dados, imprecisão do aparelho de coleta ou até mesmo sendo a informação de impossível obtenção), alterando a qualidade da coleta. Para estes casos, técnicas que trabalham com o raciocínio probabilístico podem ser interessantes.

Uma destas técnicas são as Redes Bayesianas, que se tornou um dos principais métodos para de trabalhar com incerteza no campo da inteligência artificial [9]. Isso se deve às suas propriedades de representar domínios, correlacionar as dependências entre suas variáveis e estudá-las, permitindo mais facilmente visualizar e compreender as relações entre as variáveis do ambiente, considerando a probabilidade como o grau de certeza da ocorrência de um evento.

As Redes Bayesianas também são conhecidas como Redes de Opinião, Redes Causais ou ainda, Gráficos de Dependência Probabilística. São dispostas em grafos que levam a conclusões, baseados na incerteza. Esta técnica vem se tornando

a metodologia padrão para a construção dos sistemas que confiam no conhecimento probabilístico e tem sido aplicada em uma variedade de atividades do mundo real. [10]

Com isso, tem-se que as Redes Bayesianas podem ser consideradas como uma representação visual e informativa da tabela de probabilidade conjunta de todas as variáveis que envolvem o domínio do problema. [11]

## 3.2 ESTRUTURAS DE UMA REDE BAYESIANA

Como citado anteriormente, as Redes Bayesianas são uma representação gráfica de variáveis e como estas se relacionam em um problema específico. A esta representação dá-se o nome de grafo, que é acíclico dirigido e representa as dependências entre suas variáveis, além de avaliar quantitativamente (em termos de probabilidades) essas dependências. O grafo é elemento fundamental da rede.

Os elementos da Rede Bayesiana são chamados de nós e arcos. Os nós são os elementos principais e representam as variáveis aleatórias (podem ser discretas ou contínuas) consideradas no problema e são representados por círculos. São representados por  $X = \{X_1, X_2, \dots, X_n\}$ . Os arcos são setas que representam a conexão direta entre os nós, representando a dependência probabilística direta entre duas variáveis.

Hierarquicamente, definimos os nós da Rede Bayesiana como “pai” e “filho”, referindo-se à relação de dependência direta e causalidade entre dois nós por meio do arco que os conecta (o nó de onde o arco parte é o nó pai e o nó onde o arco chega com sua ponta é designado nó filho).

O nó que não possui pais e, conseqüentemente, inicia a rede, é chamado de “raiz”, e um nó que não possui filhos é chamado de “folha”.

Analogamente a uma árvore genealógica, os nós antecedentes a um determinado nó A, (os pais e seus respectivos pais e assim sucessivamente), são denominados como ancestrais de A. O mesmo ocorre para os nós derivados de determinado nó A (os filhos e seus respectivos filhos e assim por diante), são denominados como descendentes de A.

Outro elemento importante dentro da estrutura de Redes Bayesianas é a tabela de probabilidade condicional (CPT). Trata-se da exibição dos parâmetros de probabilidade



condicional da variável, que quantifica os efeitos que os “nós pais” exercem sobre este nó.

Em suma, pode-se dizer que os arcos significam a existência de uma influência causal direta entre as variáveis conectadas, e a força destas influências é quantificada por probabilidades condicionais.

Um exemplo de Rede Bayesiana é demonstrado na Figura 3.1:

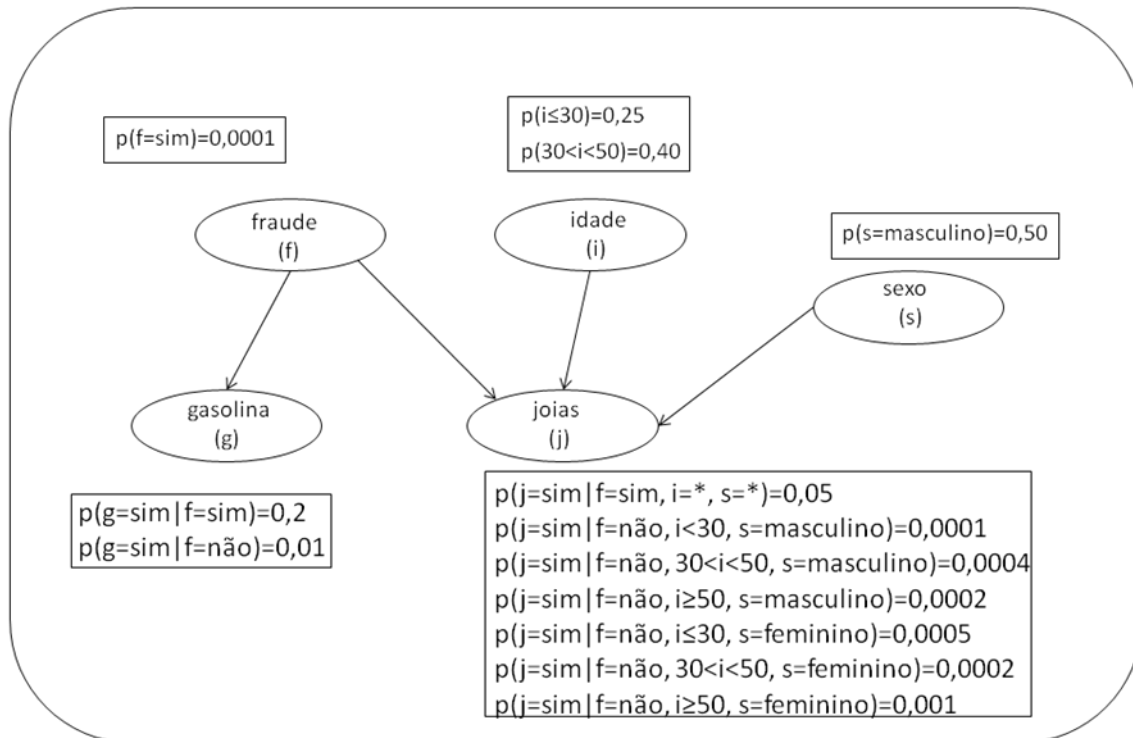


Figura 3.1: Exemplo de uma Rede Bayesiana [18]

Segundo a definição dada anteriormente, temos que os nós fraude, idade, sexo, gasolina e joias são as variáveis do domínio, dadas por  $X=\{\text{fraude, idade, sexo, gasolina, joias}\}$ . As variáveis raiz, fraude, idade e sexo, apresentam apenas probabilidades verdadeiro,  $p(X_i=\text{verdadeiro})$  e falso,  $p(X_i=\text{falso})$ . Os nós filhos, gasolina e joias, apresentam distribuições de probabilidade condicional que dependem quantitativamente dos demais nós da rede,  $p(X_i | \text{Pais}(X_i))$ .

As inferências permitem que qualquer análise que se deseje fazer à rede torne-se possível. Para calcular a inferência são utilizados cálculos de probabilidade condicional e probabilidade condicional conjunta do teorema proposto por Bayes, expresso na equação 3.1:

$$P(A = v_i | B) = \frac{P(A = v_i)P(B | A = v_i)}{\sum_{k=1}^{n_A} P(A = v_k)P(B | A = v_k)} \quad (3.1)$$

O Teorema de Bayes nada mais é do que a junção do teorema da probabilidade condicional e da fórmula de probabilidades totais. Então,  $P(A = v_i)$  pode ser denominada como probabilidade a priori,  $P(B | A = v_i)$  como verossimilhança e  $P(A = v_i | B)$  como probabilidade a posteriori (probabilidade posterior à observação do evento B).

Sendo o denominador a decomposição de  $P(A = v)$ , pode ser considerado como constante normalizadora, permitindo com que a equação 3.1 seja reescrita da seguinte forma:

$$P(A = v_i | B) \propto P(A = v_i)P(B | A = v_i) \quad (3.2)$$

sendo  $\propto$  indicador de proporcionalidade, demonstrando que a probabilidade a posteriori é proporcional à probabilidade a priori multiplicada pela verossimilhança.

Uma das maiores vantagens das Redes Bayesianas é o fato de as relações entre as variáveis  $X_i$  do domínio poderem ser visualizadas graficamente, além de proverem um mecanismo de inferência que permite quantificar, em termos probabilísticos, a magnitude dos efeitos dessas relações [8].

### 3.3 INFERÊNCIA EM REDES BAYESIANAS

Inferência, em Redes Bayesianas, é um mecanismo para o cálculo da distribuição posteriori de probabilidade para um conjunto de variáveis, dado um conjunto de evidências, ou seja, variáveis aleatórias com valores instanciados [12]. Resumindo, a inferência Bayesiana é o processo de obtenção da probabilidade a posteriori a partir da probabilidade a priori. O termo “inferência” também é conhecido, em Redes Bayesianas, como Atualização de Crença.

As inferências são realizadas de diversas formas, levando em consideração os tipos de estruturas e as diferentes quantidades de variáveis. Quanto mais variáveis e mais estruturas complexas, maior a dificuldade no cálculo das probabilidades da rede.

A dificuldade é maior quando há complexidade na modelagem do problema.

Dependendo do caso, uma rede com um número restrito de variáveis pode ter um tempo computacional muito maior do que uma rede com mais variáveis, devido à complexidade de sua estrutura, às vezes inviabilizando a concretização da inferência.

Levando-se em conta as estruturas mais simples, existem dois tipos de algoritmos que podem realizar inferência probabilística, são estes: algoritmos exatos, que produzem resultados mais satisfatórios, porém, com um elevado tempo e esforço computacional na incidência de um elevado número de variáveis na rede; e algoritmos aproximados, que têm maior velocidade de processamento que os algoritmos exatos, porém tem uma precisão inferior.

A especificação das variáveis, bem como seus valores, devem ser precisos o suficiente para satisfazer os requerimentos da situação modelada. Com a situação modelada e as probabilidades calculadas, pode-se inferir qualquer questionamento sobre a situação.

### **3.4 CLASSIFICADORES BAYESIANOS**

Classificadores Bayesianos são classificadores estatísticos que, como seu nome diz, faz a classificação de um objeto em uma determinada classe, considerando a probabilidade deste objeto pertencer à classe em questão. Produz resultados rapidamente, de grande correção quando aplicados a grandes volumes de dados, comparáveis aos resultados produzidos por árvores de decisão e redes neurais.

Para construir um classificador é necessária, primeiramente, a definição de um conjunto de treinamento. Este conjunto possui uma lista de exemplos (que são definidos por atributos), sendo cada um com sua respectiva classe. Então, este conjunto é submetido a um algoritmo de classificação e um modelo é construído. Esse modelo pode ter as mais diversas formas dependendo do algoritmo escolhido. Poder ser representado por uma árvore de decisão, uma fórmula matemática ou até mesmo um cálculo de probabilidades.

Resumindo, pode-se afirmar que, se um agente (responsável pelas informações) conhece a classe, é possível prever os valores das outras características. Em um classificador bayesiano, o agente de aprendizagem constrói um modelo probabilístico dos recursos e usa esse modelo para prever a classificação de um novo exemplo.

A partir do momento em que o modelo foi estabelecido, o classificador pode então atribuir categorias a novos dados nunca vistos antes. Para medir a qualidade do modelo construído, pode-se usar um conjunto de teste que deve ser independente do conjunto de treinamento, cujas classificações também são conhecidas de modo que seja possível averiguar a taxa de acerto do modelo proposto.

Existem vários algoritmos para classificação e o que foi utilizado neste trabalho foi um classificador probabilístico, o Naive Bayes.

### 3.5 CLASSIFICADOR NAIVE BAYES

Do inglês, ingênuo (bayes), o Naive Bayes é um classificador ingênuo considerado um caso especial de uma Rede Bayesiana. O modelo matemático utilizado durante o treinamento do classificador Naive Bayes consiste numa rede Bayesiana na qual todos os atributos são apenas dependentes da classe da imagem, e não possuem relação alguma de dependência entre si, como mostrado na Figura 3.2.

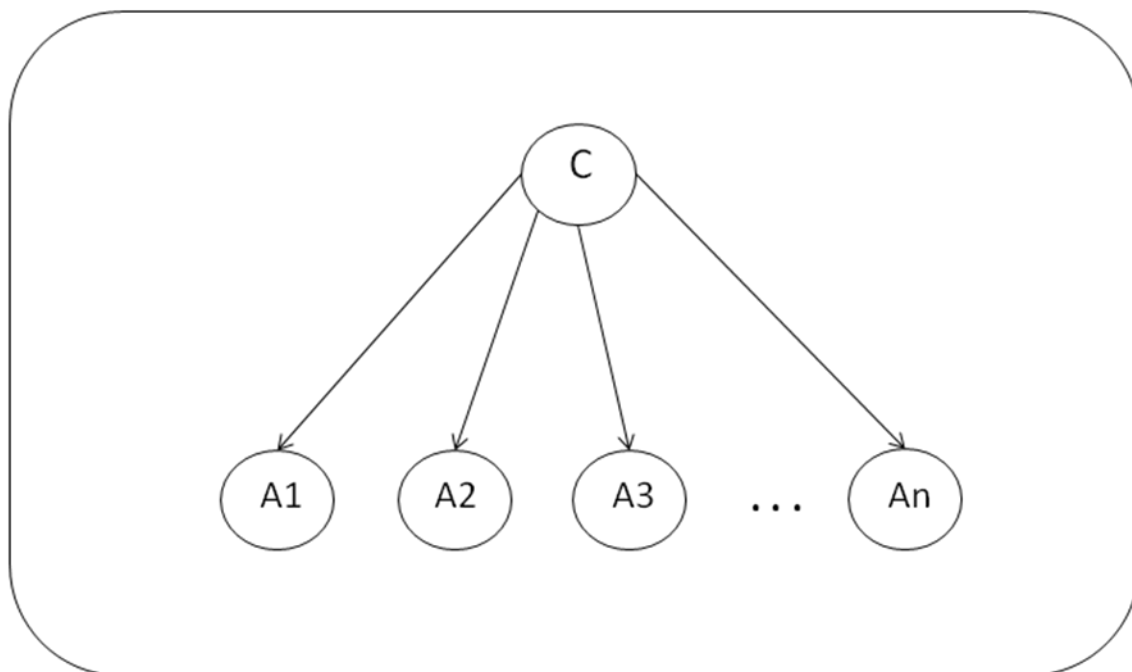


Figura 3.2: Rede Bayesiana que representa o classificador Naive Bayes

Comparativos mostram que os algoritmos Bayesianos, chamados de naive Bayes,

obtiveram resultados compatíveis com os métodos de árvore de decisão e redes neurais [21]. Entre os muitos métodos de classificação existentes ela é uma das mais simples e computacionalmente mais eficiente; sendo também robusta contra ruídos nos dados e atributos irrelevantes, não permitindo a influência nas probabilidades dos outros atributos [22].

Um classificador Naive baseia-se na aplicação do Teorema de Bayes para o cálculo das probabilidades necessárias para a classificação. Para predizer uma classe a partir de uma nova instância  $A=a_1, a_2, \dots, a_n$ , tem-se:

$$P(\text{classe} | A) = \frac{P(A | \text{Classe}) \times P(\text{classe})}{P(A)}$$

Como  $A=a_1, a_2, \dots, a_n$ , tem-se:

$$P(\text{classe} | a_1 \dots a_n) = \frac{P(a_1 \dots a_n | \text{Classe}) \times P(\text{classe})}{P(a_1 \dots a_n)}$$

Para calcular a classe mais provável da nova instância, calcula-se a probabilidade de todas as possíveis classes para, finalmente, escolher a classe com a maior probabilidade como rótulo da nova instância. Isto equivale a maximizar a  $P(\text{classe} | a_1, \dots, a_n)$ . Para isso, deve-se maximizar o valor do numerador  $P(a_1 \dots a_n | \text{Classe}) \times P(\text{classe})$  e minimizar o valor do denominador  $P(a_1 \dots a_n)$ . Como este é uma constante, haja vista que não depende da variável classe que se está procurando, pode-se anulá-lo no Teorema de Bayes, resultando na seguinte fórmula:

$$\arg \max P(\text{classe} | a_1 \dots a_n) = \arg \max P(a_1 \dots a_n | \text{classe}) \times P(\text{classe})$$

Nesta, procura-se a classe que maximize o valor do termo  $P(\text{classe} | a_1 \dots a_n) = P(a_1 \dots a_n | \text{classe}) \times P(\text{classe})$ .

A “ingenuidade” a que o Classificador Naive Bayes se refere é que todos os atributos  $a_1, \dots, a_n$  da instância que se quer classificar são independentes. Assim, o complexo cálculo do valor do termo  $P(a_1, \dots, a_n | \text{classe})$  reduz-se ao simples cálculo  $P(a_1 | \text{classe}) \times \dots \times P(a_n | \text{classe})$ . Então, a fórmula final utilizada pelo classificador é:

$$\arg \max P(\text{classe} | a_1 \dots a_n) = \arg \max \prod_i P(a_i | \text{classe}) \times P(\text{classe})$$

A suposição de independência dos atributos de uma instância muitas vezes é falsa. Ainda assim, o classificador Naive Bayes produz resultados bastante satisfatórios. Quando os atributos são, de fato, independentes, o classificador fornece soluções muito confiáveis.

O cálculo da classe de uma nova instância consiste no cálculo da probabilidade de todas as possíveis classes, escolhendo-se a seguir a classe com maior probabilidade.

Considerando as fórmulas acima, tem-se que  $P(\text{classe})$  é, simplesmente, o número de instâncias com o seu rótulo durante o treinamento do classificador que pertencem à classe em questão sobre o número total de instâncias.  $P(a_i|\text{classe})$  é o número de casos pertencentes à classe em questão com o atributo  $i$  com o valor  $a_i$  sobre o número total de casos.

Como exemplo do uso de um classificador Naive Bayes, pode-se considerar um banco de dados para treinamento, onde o atributo classe é o jogo de tênis, baseado na previsão do tempo [25].

Tabela II: Exemplo de conjunto de treinamento

Caso	Clima	Temperatura	Umidade	Vento	Jogo de Tênis
1	Ensolarado	Quente	Alta	Fraco	Não
2	Ensolarado	Quente	Alta	Forte	Não
3	Nublado	Quente	Alta	Fraco	Sim
4	Chuvoso	Amena	Alta	Fraco	Sim
5	Chuvoso	Frio	Normal	Fraco	Sim
6	Chuvoso	Frio	Normal	Forte	Não
7	Nublado	Frio	Normal	Forte	Sim
8	Ensolarado	Amena	Alta	Fraco	Não
9	Ensolarado	Frio	Normal	Fraco	Sim
10	Chuvoso	Amena	Normal	Fraco	Sim
11	Ensolarado	Amena	Normal	Forte	Sim
12	Nublado	Amena	Alta	Forte	Sim
13	Nublado	Quente	Normal	Fraco	Sim
14	Chuvoso	Amena	Alta	Forte	Não

Para se descobrir a classe mais provável (jogar tênis ou não jogar) para a instância [Clima=ensolarado, Temperatura=frio, Umidade=alta, Vento=forte], com base nos dados de treinamento, deve-se proceder assim:

- Calcular a probabilidade de da ocorrência de cada uma das classes  
 $P(\text{classe}) = (\text{número de casos da classe}) / (\text{número total de casos})$   
 $P(\text{Jogar Tênis}=\text{sim}) = 9/14 = 0,64$   
 $P(\text{Jogar Tênis}=\text{não}) = 5/14 = 0,36$
- Calcular a probabilidade de cada um dos atributos em questão em relação a cada possível classe. Exemplo: Atributo Vento=Forte  
 $P(a_i|\text{Classe}) = (\text{n}^\circ \text{ de casos da classe com atributo } a_i) / (\text{n}^\circ \text{ total de casos da classe})$

$$P(\text{Vento=Forte}|\text{Jogar Tênis=sim})=3/9=0,33$$

$$P(\text{Vento=Forte}|\text{Jogar Tênis=não})=3/5=0,60$$

- Com todas as probabilidades necessárias calculadas, basta calcular a probabilidade da classe ocorrer:

$$P(\text{classe} | a_1 \dots a_n) = \prod_i P(a_i | \text{classe}) \times P(\text{classe})$$

$$\begin{aligned} P(\text{Jogar Tênis=Sim} \mid \text{Clima=Ensolarado, Temperatura=Frio,} \\ \text{Umidade=Alta, Vento=Forte}) = P(\text{Clima=Ensolarado} \mid \text{Jogar Tênis=Sim}) \times \\ P(\text{Temperatura=Frio} \mid \text{Jogar Tênis=Sim}) \times P(\text{Umidade=Alta} \mid \text{Jogar Tênis=Sim}) \\ \times P(\text{Vento=Forte} \mid \text{Jogar Tênis=Sim}) \times P(\text{Jogar Tênis=Sim}) = 0,0053 \end{aligned}$$

$$\begin{aligned} P(\text{Jogar Tênis=Não} \mid \text{Clima=Ensolarado, Temperatura=Frio,} \\ \text{Umidade=Alta, Vento=Forte}) = P(\text{Clima=Ensolarado} \mid \text{Jogar Tênis=Não}) \times \\ P(\text{Temperatura=Frio} \mid \text{Jogar Tênis=Não}) \times P(\text{Umidade=Alta} \mid \text{Jogar Tênis=Não}) \\ \times P(\text{Vento=Forte} \mid \text{Jogar Tênis=Não}) \times P(\text{Jogar Tênis=Não}) = 0,0206 \end{aligned}$$

Assim, a probabilidade de não jogar tênis é maior do que a probabilidade de jogar, com base nos atributos da instância que se pretende classificar. Portanto, a classificação mais provável da nova instância é Jogar Tênis=Não.

Justificada pelos benefícios citados anteriormente e à facilidade em seu entendimento, bem como a importância da independência entre as variáveis, este foi o classificador escolhido para a realização deste trabalho.

### 3.6 CONCLUSÃO DO CAPÍTULO

Neste capítulo foi explicado que as redes Bayesianas utilizam conceitos de mapas causais para modelar domínios. Estes mapas causais descrevem as relações entre as variáveis (nós) e como elas se relacionam, através de arcos (setas) que representam causa e efeito, na forma de um grafo acíclico. A intensidade dessas relações é dada através de tabelas de probabilidade condicional. Cada variável possui uma tabela, que quantifica as probabilidades de ocorrência de um evento.

O Teorema de Bayes é quem calcula as probabilidades utilizando as probabilidades a priori, que são adquiridas por intermédio de um especialista ou através de um banco de

dados.

Desta forma ficou clara a excelente utilização das Redes Bayesianas para representar corretamente um domínio. Foi confirmado por meio de referências bibliográficas que esta pode ser considerada um método bastante atrativo para organização, armazenamento e extração de conhecimento. Por último foi destacada a eficácia do método Naive Bayes, mesmo sendo um método de classificação simples.

Neste trabalho a inferência bayesiana utilizada irá gerar tabelas de probabilidades condicionais e tabelas de variância que serão utilizadas para estimar o campo elétrico e, posteriormente, compará-lo ao campo elétrico medido durante as campanhas de medição.

O próximo capítulo irá abordar como e onde foram feitas as medições, quais equipamentos foram utilizados e como os dados foram tratados para que a rede bayesiana pudesse ser utilizada.



## CAPÍTULO 4

### CAMPANHAS DE MEDIÇÃO E METODOLOGIA.

#### 4.1 INTRODUÇÃO

Em ambientes rurais a onda de superfície se propaga em terrenos com variações no que diz respeito à vegetação e à topografia. O capítulo 2 mostrou os modelos Terra Plana e Terra Esférica, os modelos de predição mais utilizados para determinação da área de cobertura de uma emissora de radiodifusão em onda media, mesmo com sua limitada precisão.

Os dados utilizados para elaboração deste trabalho foram coletados na zona rural ao redor da cidade de Brasília. Esta campanha teve como objetivo determinar o comportamento da onda nesta região e também resultou em diversos trabalhos acadêmicos de grande relevância [2,3,4]. Os dados coletados foram medidas de intensidade de campo elétrico ao longo de seis rotas radiais em torno do Centro de Transmissão da Empresa Brasileira de Comunicação S/A – Radiobrás, situada na SOFS Quadra 1, na cidade de Brasília-DF, que opera em classe A e irradiava os sinais.

A Figura 4.1 mostra a distribuição uniforme das seis rotas onde foram coletados os dados.



Figura 4.1: Rotas em que foram realizadas as medidas de intensidade de campo elétrico

A antena transmissora está situada em  $15^{\circ}49'31,44''S - 47^{\circ}57'49,89''W$ . Opera a uma frequência central de 900KHz, com 10KHz de banda e potência de transmissão de 50kW ERP, tem um campo característico de 321mV/m e está localizada a uma altura de 94m.

Neste trabalho os valores de campo elétrico, medidos dentro da área de cobertura da emissora utilizada, serão comparados aos valores de campo elétrico obtidos através de simulações por inteligência computacional através de Redes Bayesianas, vistas no capítulo 3. [24]

## 4.2 SISTEMAS DE MEDIÇÕES

Os equipamentos utilizados nas medições foram cedidos pela Agência Nacional de Telecomunicações – ANATEL e pela Radiobrás. Os equipamentos da ANATEL foram instalados na Unidade Móvel de Radiomontagem – UMR da própria agência. Os equipamentos foram integrados à UMR pela empresa Rontan em São Paulo, sob a supervisão da empresa Francesa Thales, responsável pela montagem e calibração dos equipamentos. A UMR da ANATEL tem estrutura que permite a realização de missões como radiogoniometria, avaliação de parâmetros técnicos (modulação, frequência e intensidade de campo), análise espectral e taxa de ocupação.

O sinal modulado em amplitude - AM foi captado por uma antena monopolo ativa RN 4203 Rayan, que passava pelo comutador de antenas AEA 196 e era entregue ao receptor REC 108. Este receptor contém dois módulos para a faixa de frequência inferior a 30 MHz: o RF HF 2000, que translada o sinal de Radiofrequência – RF para Frequência Intermediária - FI; e o FI HVU 2000, que amostra o sinal em Frequência Intermediária e o demodula. Após o processamento no receptor, as amostras do sinal em FI são enviadas a unidade central de processamento, KPRIII, que adquire o valor da intensidade de campo de cada amostra em FI utilizando um Processador Digital de Sinais. As coordenadas geográficas, latitude e longitude, de cada amostra são armazenadas através de um receptor GPS Trimble, acoplado na porta RS232 do Computador KPRIII.

Além do GPS Trimble, há um GPS portátil interligado a um Palm Top, com o objetivo de registrar os pontos, através de coordenadas geográficas, onde o sinal apresenta variações bruscas no valor da intensidade do campo elétrico medido. Essas

variações são muitas vezes provocadas pela presença de linhas de transmissão, tuneis e placas de sinalização.

A Figura 4.2 ilustra a organização do sistema de transmissão disposto na UMR da ANATEL.

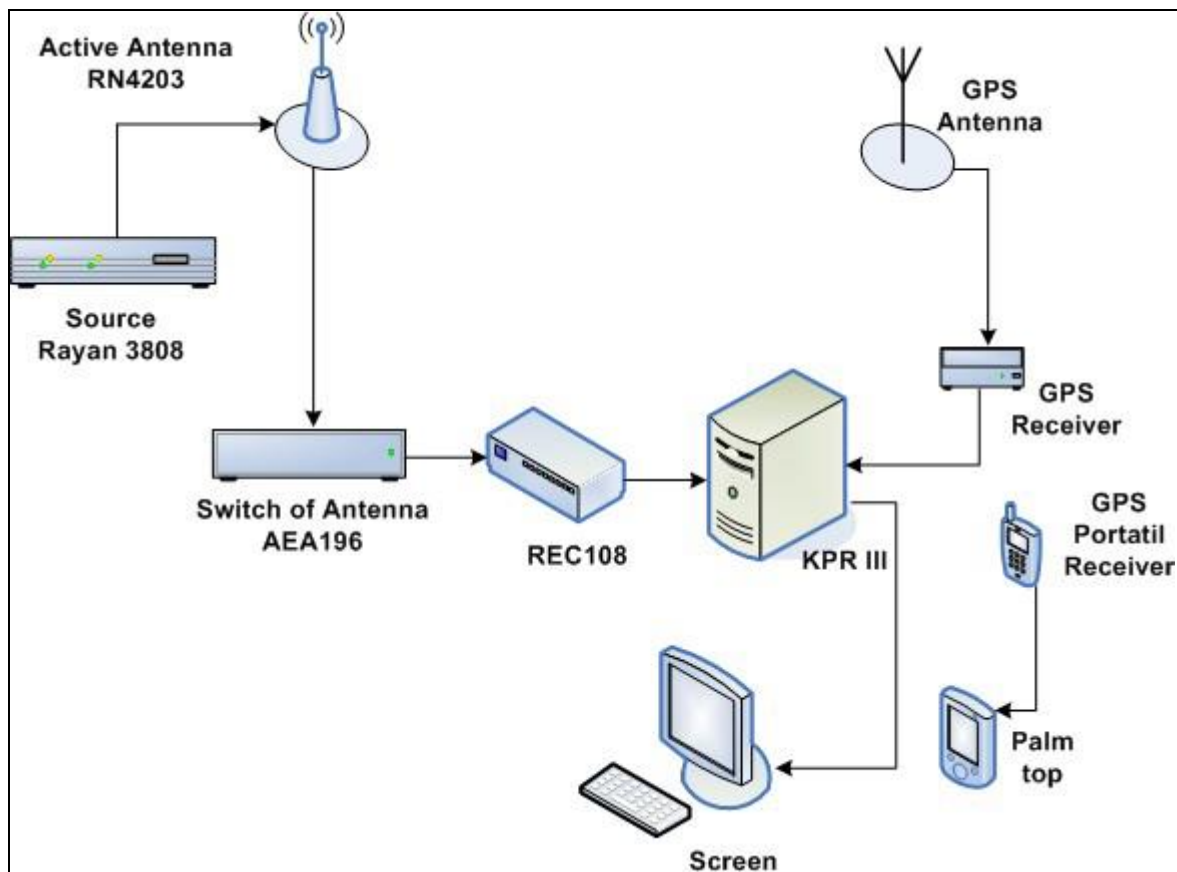


Figura 4.2: Ilustração da disposição dos equipamentos de medição instalados na UMR da ANATEL [4]

A UMR da ANATEL foi responsável pela coleta de dados apenas das rotas 1, 2 e 6. Para as rotas 3, 4 e 5 foi utilizada a viatura da Empresa Brasileira de Comunicação S/A – Radiobrás, sendo que os equipamentos desta viatura foram montados pela equipe do consorcio sistema de radiodifusão digital DRM (*Digital Radio Mondiale*) em Brasília-DF, Brasil. A Rota 01 também foi refeita por esta viatura possibilitando a validação completa ao longo de uma rota.

Para validar as medições realizadas com a viatura da Radiobrás foram feitas comparações dos valores obtidos com os valores coletados anteriormente pela UMR da

ANATEL. Essas comparações foram feitas através de um analisador de espectro instalado no campus da Universidade de Brasília – UnB.

Na viatura da Radiobrás, o sinal AM foi coletado por uma antena monopolo curto ativa, com fator K igual a 15dB, modelo R&S HE010 do fabricante Rohde-Schwarz. O sinal coletado era entregue a um PSU (*Power Suplly Unit* – 10 kHz a 30 GHz) também da Rohde-Schwarz, que interliga a antena de recepção ao receptor.

O receptor profissional EB200 adquire o sinal do PSU em RF e translada para Frequência Intermediária e banda básica, colocando-o no formato IQ, em fase e quadratura. Já no formato IQ, o sinal é entregue a uma placa de som modelo *Creative Sound Blaster* 24 bit diretamente conectada a porta USB de um *LapTop*. Nele, foram instalados os *softwares*, *Dream*, responsável pela demodulação do áudio, e TSR DRM, que faz a captura do valor da intensidade do campo elétrico medido no EB200. Além disso, o *software* TSR DRM também armazena, para cada amostra, o valor do campo elétrico, as coordenadas geográficas obtidas por um GPS modelo 99xx Garmim, a velocidade do veiculo e a altitude do terreno.

Assim, as medições para as rotas 3, 4 e 5 foram realizadas mediante o *layout* da Figura 4.3.

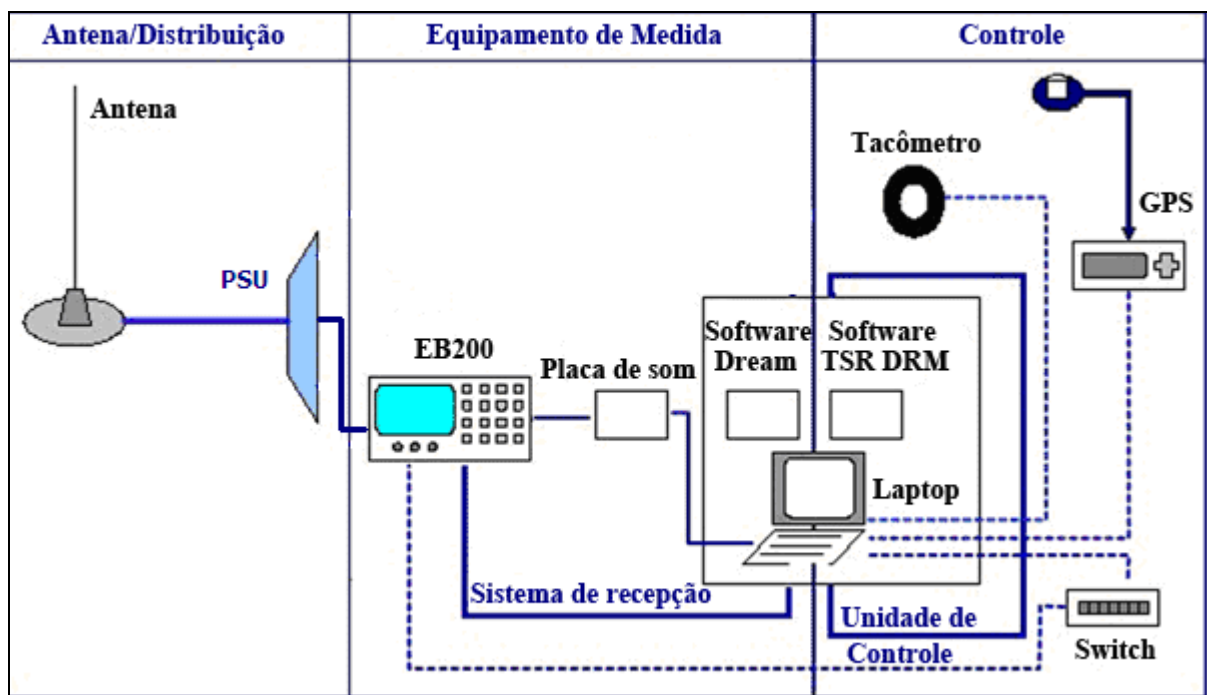


Fig. 4.3: Ilustração da disposição dos equipamentos de medição instalados na viatura da Radiobrás[4]

### 4.3 PROCESSAMENTO DOS DADOS

Os dados utilizados neste trabalho foram gentilmente cedidos pelo Msc. Bruno Monteiro. Foram coletadas aproximadamente 25.000 amostras para cada uma das 6 rotas, cada uma incluindo informações de distância em relação ao transmissor, campo elétrico medido, velocidade das viaturas e localização geográfica dos pontos, com altitude, latitude e longitude.

Os arquivos das rotas 02, 03, 04 e 05 foram recebidos pré-processados, de forma à adequá-los ao estudo em questão. Os arquivos das rotas 01 e 06 não tinham esse processamento, e não foi possível realizá-lo por não haver disponível o software necessário para este fim, portanto, neste trabalho serão consideradas apenas as rotas de 02 à 05.

A topografia do terreno para as rotas 02, 03, 04 e 05 pode ser representada, respectivamente, pela Figura 4.4, Figura 4.5, Figura 4.6 e Figura 4.7.

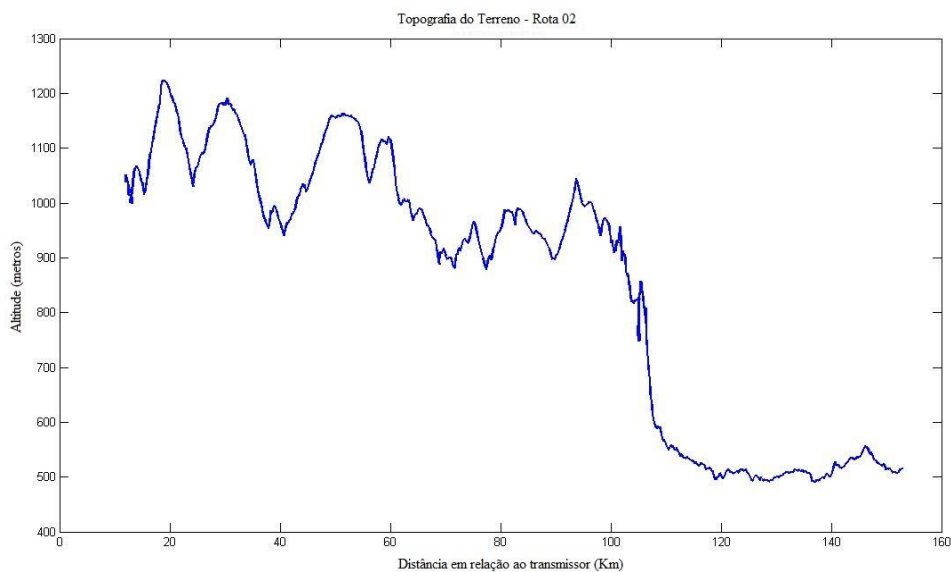


Figura 4.4: Representação do terreno da Rota 02

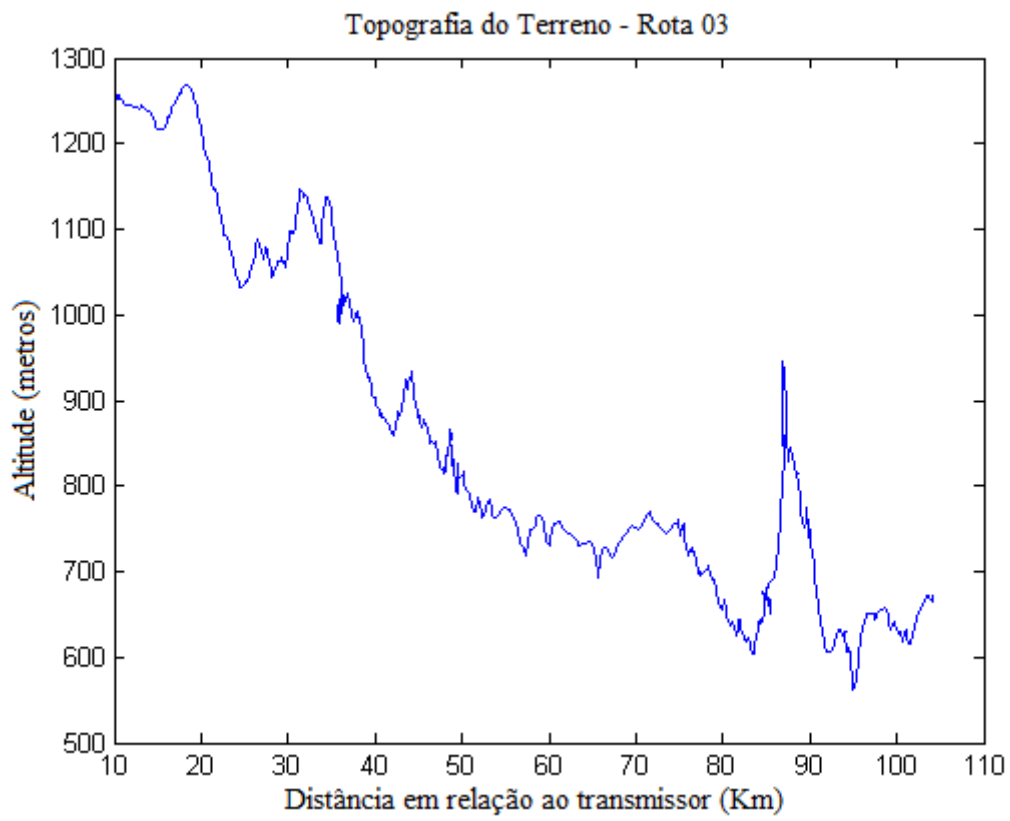


Figura 4.5: Representação do terreno da Rota 03

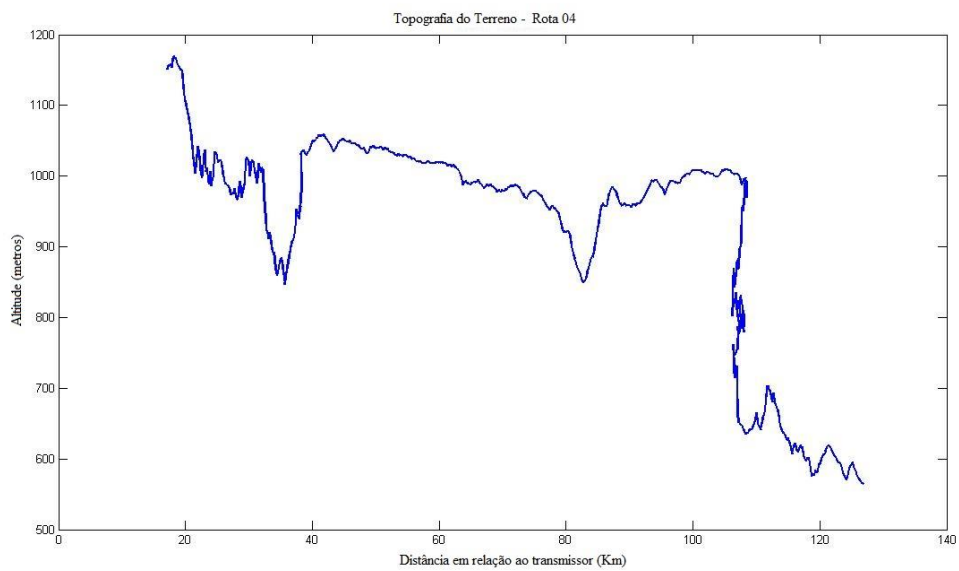


Figura 4.6: Representação do terreno da Rota 04

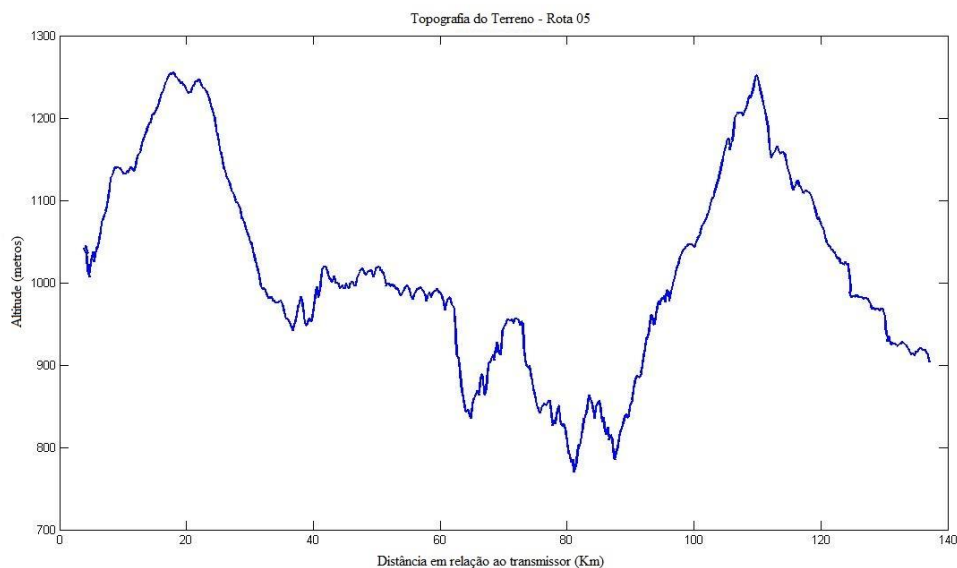


Figura 4.7: Representação do terreno da Rota 05

Após a definição do comportamento do terreno, foi necessário reduzir a quantidade de amostras por rota para um valor de até 700, pois esta é a quantidade máxima suportada pela versão acadêmica do software *Bayseware*, que faz a inferência Bayesiana e permite a exportação das tabelas de probabilidade condicional e variância. O agrupamento de dados, também conhecido por *clusterização*, foi feito através de um programa desenvolvido em *Basic*. O mesmo eliminou redundâncias e reduziu os dados baseados na localização dos pontos próximos.

Primeiramente, com o software *Matlab* foram gerados arquivos que relacionavam “Campo Medido” com a Distância do receptor em relação ao transmissor para todas as rotas em estudo com, aproximadamente, a quantidade de 10.000 amostras em cada rota.

Para avaliar se este agrupamento foi feito de forma adequada, novos gráficos foram gerados no *Matlab*, desta vez com a quantidade pretendida, de 700 amostras. De posse de dois gráficos por rota, os mesmos foram comparados a fim de verificar se houve ou não mudança no comportamento do sinal.

No caso da Rota 02, o comportamento do sinal para a quantidade de 10.000 amostras é representado pela Figura 4.8:

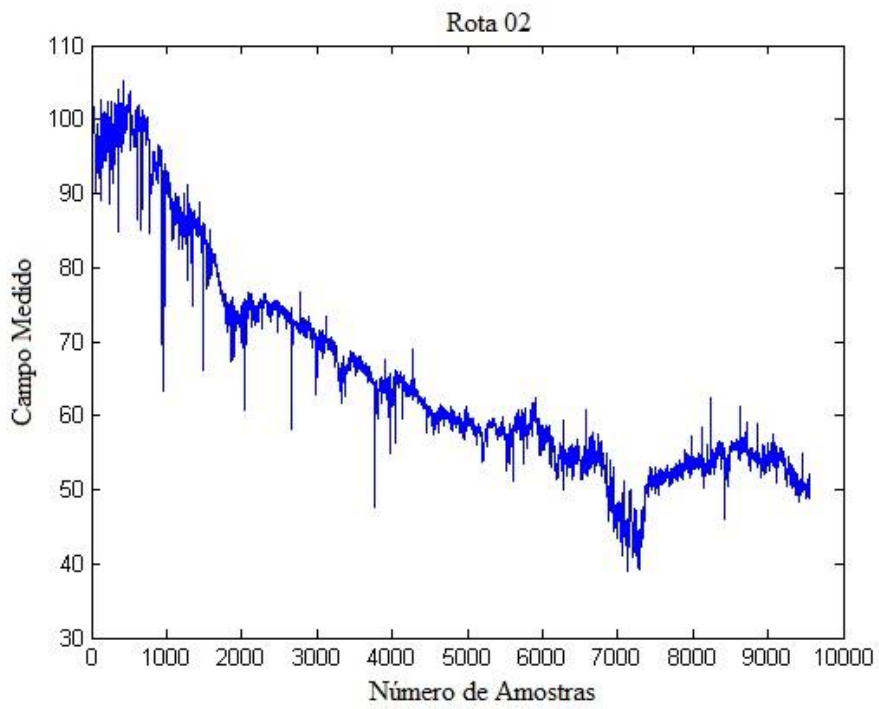


Figura 4.8: Comportamento do Sinal na Rota 02 para 10.000 amostras

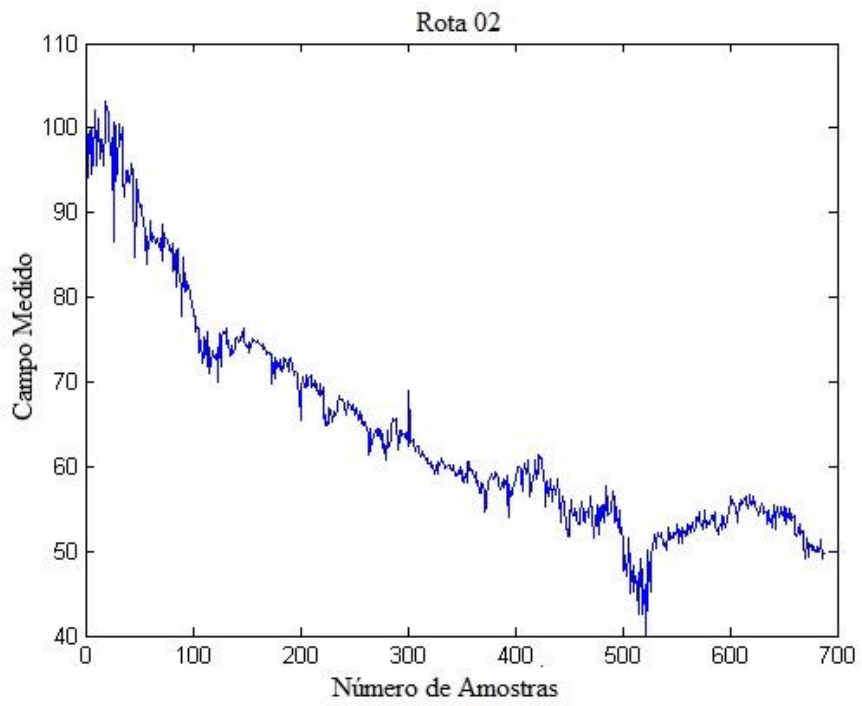


Figura 4.9: Comportamento do Sinal na Rota 02 para 700 amostras



A curva da Figura 4.9 acompanha a curva da Figura 4.8., confirmando que o comportamento do sinal não alterou após o processo de *clusterização*. Assim, este pôde ser validado para utilização nas demais rotas.

Os resultados alcançados com os gráficos comparativos foram satisfatórios para que o tratamento dos dados com até 700 amostras para cada rota pudessem ser utilizados.

Após a *clusterização*, os dados foram usados em uma rede *Bayesiana Naive*, que, como visto no capítulo 3, foi escolhido porque ele se destaca entre os vários métodos de classificação existentes como um dos mais simples e computacionalmente mais eficiente.

#### **4.4 CONSIDERAÇÕES FINAIS**

Este capítulo mostrou como foi dada a campanha de medidas para coleta dos dados utilizados no trabalho, bem como equipamentos utilizados e suas características, incluindo também a ilustração dos percursos seguidos durante a campanha.

Foi exposto como os dados coletados foram tratados de forma que pudessem ser utilizados para execução deste trabalho, e, além disso, foi confirmada a eficácia do método de *clusterização* utilizado.

As informações reduzidas das rotas foram utilizadas em uma rede bayesiana e, com isso, foram geradas tabelas de probabilidades condicionais e tabelas de variância. Esses dados foram utilizados em um script de *Matlab* para traçar os gráficos do campo elétrico estimada contra campo elétrico medido. Estas informações serão apresentadas no próximo capítulo.

## **CAPÍTULO 5**

### **RESULTADOS - UMA METODOLOGIA PARA PREDIÇÃO DO CAMPO ELÉTRICO DE RADIODIFUSÃO SONORA EM ONDAS MÉDIAS UTILIZANDO INFERÊNCIAS BAYESIANAS.**

#### **5.1 INTRODUÇÃO**

No início deste trabalho verificou-se a importância da predição da intensidade de campo elétrico para o planejamento e implantação de novos sistemas de radiodifusão, bem como a reorganização dos sistemas já existentes (minimizando problemas de interferência entre localidades vizinhas, por exemplo). Sua importância se deve ao fato de que estes valores permitem o cálculo da potência transmitida e a área de cobertura das estações de rádio.

Assim, há a necessidade do desenvolvimento de novas ferramentas que possibilitem a predição do campo elétrico de forma confiável, a partir de características topográficas do ambiente, sem a necessidade de se conhecer os demais parâmetros e características elétricas do solo em análise, já que essas informações, muitas vezes, são de difícil acesso.

Como foi citado no Capítulo 3, o interesse pela utilização de Redes Bayesianas como inteligência computacional para estimar valores de campo elétrico em um dado ambiente pode ser explicado por sua eficácia, além de ser vista como uma excelente metodologia de organização, armazenamento de informações e extração de conhecimento.

No capítulo anterior foi explicado como ocorreram as campanhas de medição para coleta dos dados utilizados neste trabalho, o lugar e tipo de relevo em que estas campanhas foram realizadas. O Capítulo 4 apresentou também quais foram os equipamentos utilizados nas campanhas, suas características e como foram dispostos para realizar a coleta.

Neste capítulo será confirmada a eficácia das Inferências Bayesianas para estimar os valores de campo elétrico, mostrando as tabelas de probabilidade condicional encontradas pra cada rota em estudo e, por fim, fazer a comparação entre gráficos do campo elétrico estimado com o campo elétrico medido, bem como avaliar graficamente os erros entre ambos.

## 5.2 METODOLOGIA

A princípio, tentou-se utilizar para a criação e análise da Rede Bayesiana o programa computacional denominado *WEKA (Waikato Environment for Knowledge Analysis)*, por este ser de domínio público e prover um conjunto de algoritmos que implementam diversas técnicas para resolver problemas reais de modelagem de incerteza. A não possibilidade de exportação da tabela de probabilidade condicional, porém, fez com que outra ferramenta fosse utilizada para desenvolver a metodologia em questão.

Então, utilizou-se a versão acadêmica do programa *Bayesware Discoverer*, que proporcionou além da criação e a geração da Rede Bayesiana a partir dos atributos do banco de dados criado, as tabelas de probabilidade condicional de cada atributo.

A figura 5.1, mostra a dependência entre altitude do terreno e campo elétrico para a Rota 02, assim como os valores de discretização obtidos pela rede (mostrados entre colchetes). Em vermelho estão as probabilidades de ocorrência para cada intervalo.

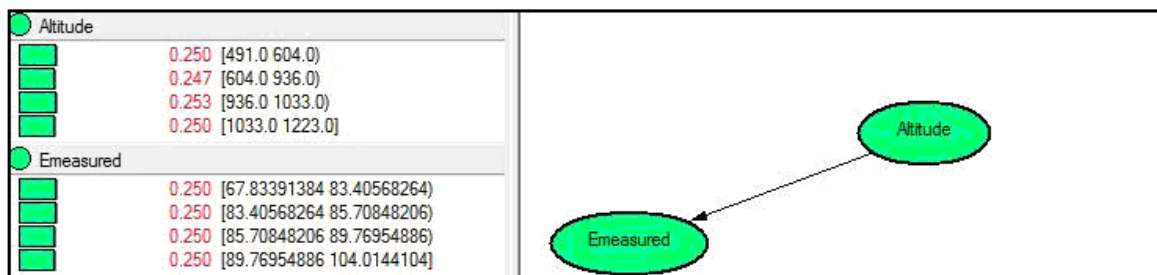


Figura 5.1: Rede Bayesiana gerada para a Rota 02

As Figuras 5.2, 5.3 e 5.4 ilustram, respectivamente, a Rede Bayesiana gerada para as rotas 03, 04 e 05, bem como suas relações de dependência e probabilidades de ocorrência.

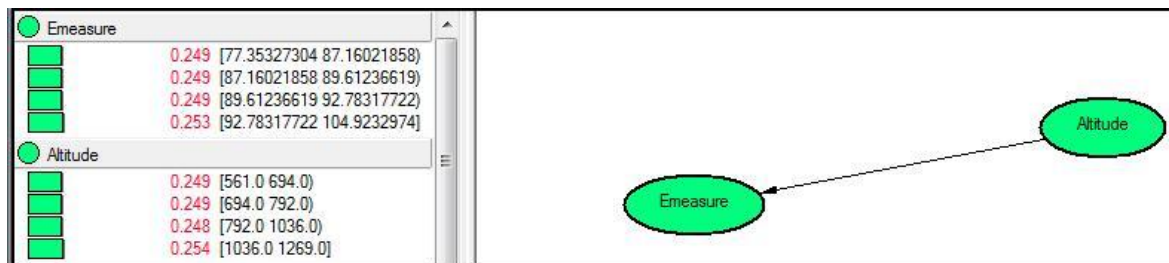


Figura 5.2: Rede Bayesiana gerada para a Rota 03

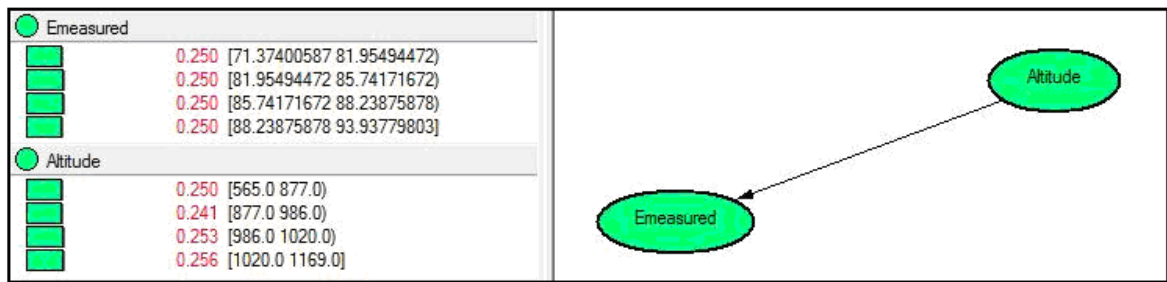


Figura 5.3: Rede Bayesiana gerada para a Rota 04

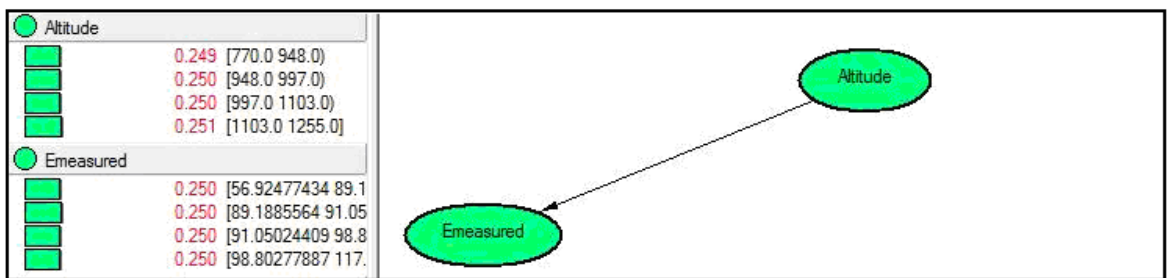


Figura 5.4: Rede Bayesiana gerada para a Rota 05

As *Naive Bayes* são os classificadores Bayesianos mais simples e conhecidos, funcionando com a hipótese de que todos os atributos são independentes. As tabelas de probabilidade condicional foram obtidas e exportadas através da Rede Bayesiana *Naive*, e são mostradas a seguir.

Tabela III: Tabela de Probabilidade Condicional obtida para a Rota 02

Tabela de Probabilidade Condicional	Campo Elétrico			
	[67.8 83.4]	[83.4 85.7]	[85.7 89.7]	[89.7 104.0]
<b>Altitude</b>				
[491 604]	0.534	0.285	0.180	0.000
[604 936]	0.412	0.282	0.300	0.006
[936 1033]	0.046	0.396	0.305	0.253
[1033 1223]	0.012	0.035	0.215	0.738

Tabela IV: Tabela de Probabilidade Condicional obtida para a Rota 03

Tabela de Probabilidade Condicional	Campo Elétrico				
	Altitude	[77.3 87.1]	[87.1 89.6]	[89.6 92.7]	[92.7 104.9]
[561 694]		0.442	0.250	0.273	0.035
[694 792]		0.105	0.343	0.534	0.018
[792 1036]		0.444	0.380	0.146	0.030
[1036 1269]		0.012	0.029	0.046	0.913

Tabela V: Tabela de Probabilidade Condicional obtida para a Rota 04

Tabela de Probabilidade Condicional	Campo Elétrico				
	Altitude	[71.3 81.9]	[81.9 85.7]	[85.7 88.2]	[88.2 93.9]
[565 877]		0.894	0.076	0.018	0.012
[877 986]		0.085	0.491	0.321	0.103
[986 1020]		0.023	0.399	0.376	0.202
[1020 1169]		0.000	0.046	0.286	0.668

Tabela VI: Tabela de Probabilidade Condicional obtida para a Rota 05

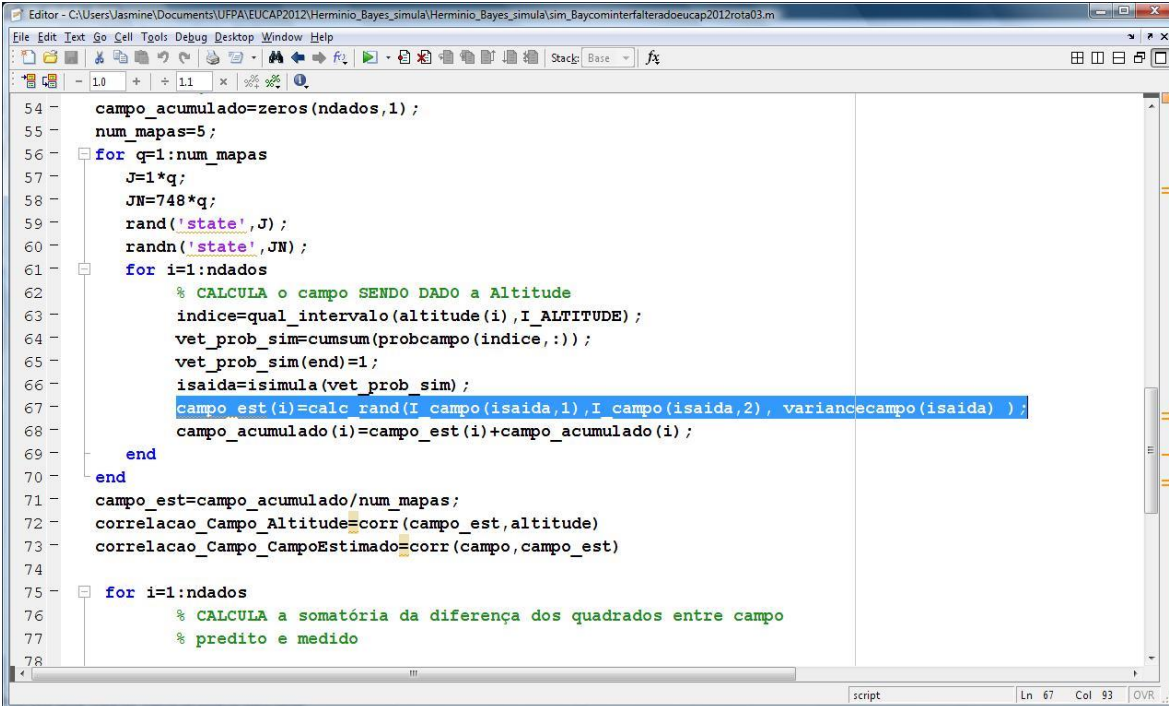
Tabela de Probabilidade Condicional	Campo Elétrico				
	Altitude	[56.9 89.1]	[89.1 91.0]	[91.0 98.8]	[98.8 176.2]
[770 948]		0.249	0.491	0.249	0.012
[948 997]		0.201	0.213	0.465	0.121
[997 1103]		0.190	0.207	0.287	0.316
[1103 1255]		0.360	0.092	0.000	0.548

As tabelas mostram qual a probabilidade de um Evento A ocorrer simultaneamente a um Evento B. Por exemplo, na Tabela III a probabilidade do Campo Elétrico variar entre 67.8 e 83.4 em um nível de altitude entre 491 e 604 é de 0.534.

## 5.2.1 – INFERÊNCIAS REALIZADAS

As tabelas de probabilidades condicionais e tabelas de variância geradas usando a rede bayesiana foram geradas por um script de *Matlab* para traçar os gráficos do campo elétrico estimado, a fim de serem comparados aos gráficos gerados pelo campo elétrico medido. Vale ressaltar que cada rota tem uma tabela específica a ser utilizada.

Primeiramente o Teorema de Bayes foi declarado no Script:



```
54 - campo_acumulado=zeros(ndados,1);
55 - num_mapas=5;
56 - for q=1:num_mapas
57 -     J=1*q;
58 -     JN=748*q;
59 -     rand('state',J);
60 -     randn('state',JN);
61 -     for i=1:ndados
62 -         % CALCULA o campo SENDO DADO a Altitude
63 -         indice=qual_intervalo(altitude(i),I_ALTITUDE);
64 -         vet_prob_sim=cumsum(probcampo(indice,:));
65 -         vet_prob_sim(end)=1;
66 -         isaida=isimula(vet_prob_sim);
67 -         campo_est(i)=calc_rand(I_campo(isaida,1),I_campo(isaida,2), variancecampo(isaida));
68 -         campo_acumulado(i)=campo_est(i)+campo_acumulado(i);
69 -     end
70 - end
71 - campo_est=campo_acumulado/num_mapas;
72 - correlacao_Campo_Altitude=corr(campo_est,altitude)
73 - correlacao_Campo_CampoEstimado=corr(campo,campo_est)
74 -
75 - for i=1:ndados
76 -     % CALCULA a somatória da diferença dos quadrados entre campo
77 -     % predito e medido
78 -
```

Fig. 5.5 – Script do Matlab com a implementação do Teorema de Bayes

Depois, foram implementadas as tabelas de probabilidade condicional e variância:

```

Editor - C:\Users\uzamine\Documents\UFPA\EUCAP2012\Herminio_Bayes_simula\Herminio_Bayes_simula\sim_Baycominterfalteradoeucap2012rota03.m
File Edit Text Go Cell Tools Debug Desktop Window Help
-----
27 nnaoos=max(size(ria));
28 %*****
29
30 I_ALTITUDE= [ 561 694;
31              694 792;
32              792 1036;
33              1036 1269];
34 I_campo= [77.3532 87.1602;
35           87.1602 89.6123;
36           89.6123 92.7831;
37           92.7831 104.9232];
38
39 probaALTITUDE= [0.249 0.249 0.248 0.254];
40 probcampo = [0.442 0.250 0.273 0.035;
41             0.105 0.343 0.534 0.018;
42             0.444 0.380 0.146 0.030;
43             0.012 0.029 0.046 0.913];
44 varianceALTITUDE = [ 0.094 0.094 0.093 0.095];
45 variancecampo= [ 0.058 0.044 0.047 0.008;
46                 0.022 0.053 0.059 0.004;
47                 0.058 0.055 0.029 0.007;
48                 0.003 0.007 0.010 0.019 ];
49
50
51 %-----
script Ln 67 Col 93 OVR

```

Fig. 5.6 - Script do Matlab com a implementação das tabelas de Probabilidade Condicional e Variância

Com isso, as inferências podem ser feitas. Neste trabalho elas foram calculadas a partir do campo elétrico ou ainda a partir da altitude. As figuras abaixo ilustram como foram realizadas as inferências.

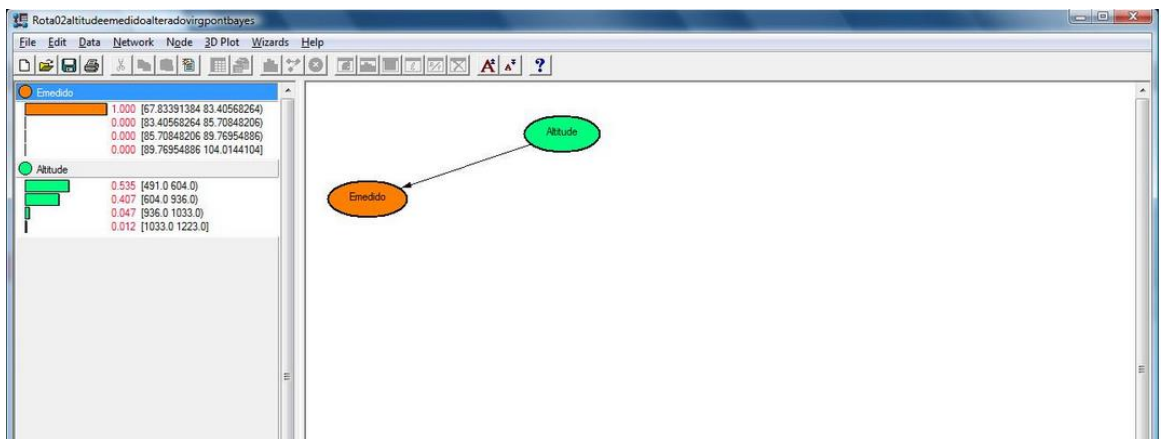


Fig. 5.7 – Inferência Bayesiana para Emedido entre 67,83 dB $\mu$ V e 83,4 dB $\mu$ V.

No caso acima foi feita a inferência considerando-se o campo elétrico variando entre 67.83dB $\mu$ V e 83.4 dB $\mu$ V. Para esta ocasião, percebeu-se que a probabilidade da altitude variar entre 491m e 604m é de 0.535; de variar entre 604m e 936m é de 0.407; entre 936m e 1033m é de 0.047 e, finalmente, entre 1033m e 1223m é de 0.012.

Além destes valores de campo, também foi feita a inferência considerando-se o campo variando entre 89,76 dB $\mu$ V e 104,01 dB $\mu$ V.

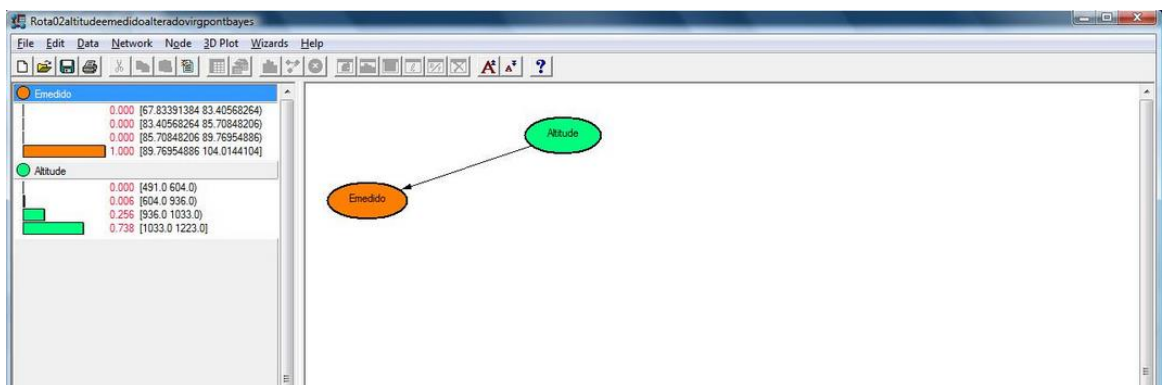


Fig. 5.8 - Inferência Bayesiana para Emedido entre 89,76 dB $\mu$ V e 104,01 dB $\mu$ V

Para tal, verificou-se que não há probabilidade da altitude variar entre 491m e 604m; que a probabilidade da altitude ser de 604m a 936m é de apenas 0.006; de variar entre 936m e 1033m é de 0.256 e, finalmente, é de 0.738 para altitudes entre 1033m e 1223m.

Com a finalidade de confirmar os valores de probabilidade condicional encontrados, as inferências no sentido inverso também foram realizadas, selecionando-se a altitude e verificando as probabilidades condicionais de ocorrência dos valores de campo elétrico. Todos os resultados foram confirmados.

Com as inferências realizadas, foram gerados os gráficos com os campos elétricos estimados. Para a Rota 02, o campo estimado através do uso das probabilidades geradas pela Rede Bayesiana é ilustrado na Figura 5.9 e comparado ao campo elétrico medido, representado pela Figura 5.10.



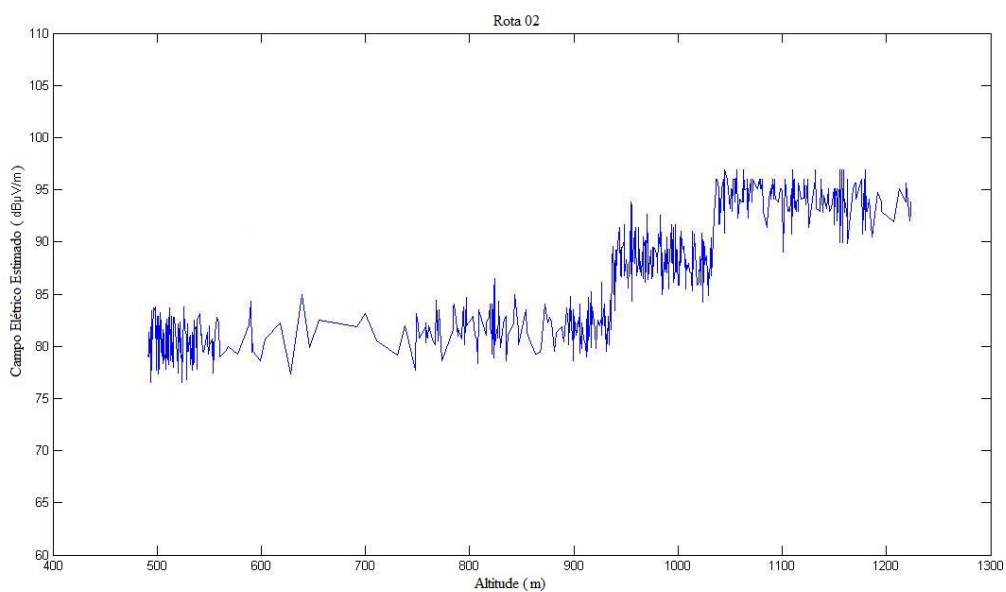


Figura 5.9: Campo Elétrico Estimado para a Rota 02

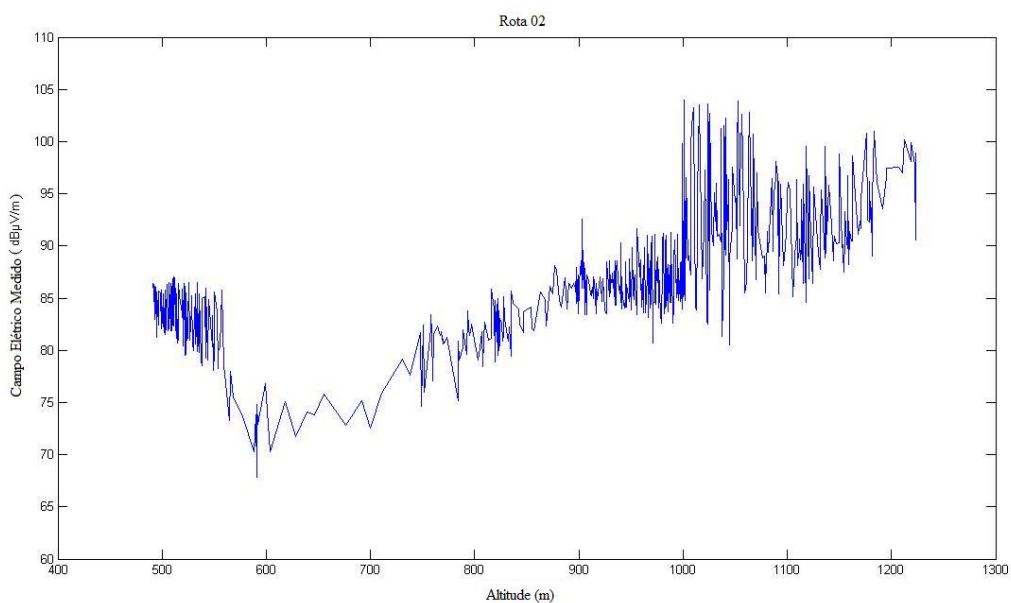


Figura 5.10: Campo Elétrico Medido para a Rota 02

Para validar a metodologia, foram calculados os valores de Erro Médio Quadrático (EMQ) e as correlações entre campo elétrico medido e altitude, e campo elétrico medido e campo elétrico estimado. Estes valores estão dispostos na Tabela VII, e todos foram considerados satisfatórios para este trabalho.

Tabela VII: Dados Calculados para Rota 02

Rota 02	
Erro RMS (dB)	4,74
Correlação Campo Elétrico Medido x Campo Elétrico Estimado	0,6658
Correlação Campo Elétrico Medido x Altitude	0,8319

Para representar as rotas 03, 04 e 05 foram gerados gráficos comparativos entre os valores de Campo Elétrico Medido e Campo Elétrico Estimado, representados, respectivamente, pela Figura 5.11, Figura 5.12 e Figura 5.13. Da mesma forma, os valores Erro Médio Quadrático, Correlação entre Campo Elétrico Medido e Campo Elétrico Estimado, e Campo Elétrico Medido e Altitude também foram calculados e mostrados na Tabela VII. Os gráficos referentes à estas rotas são mostrados abaixo:

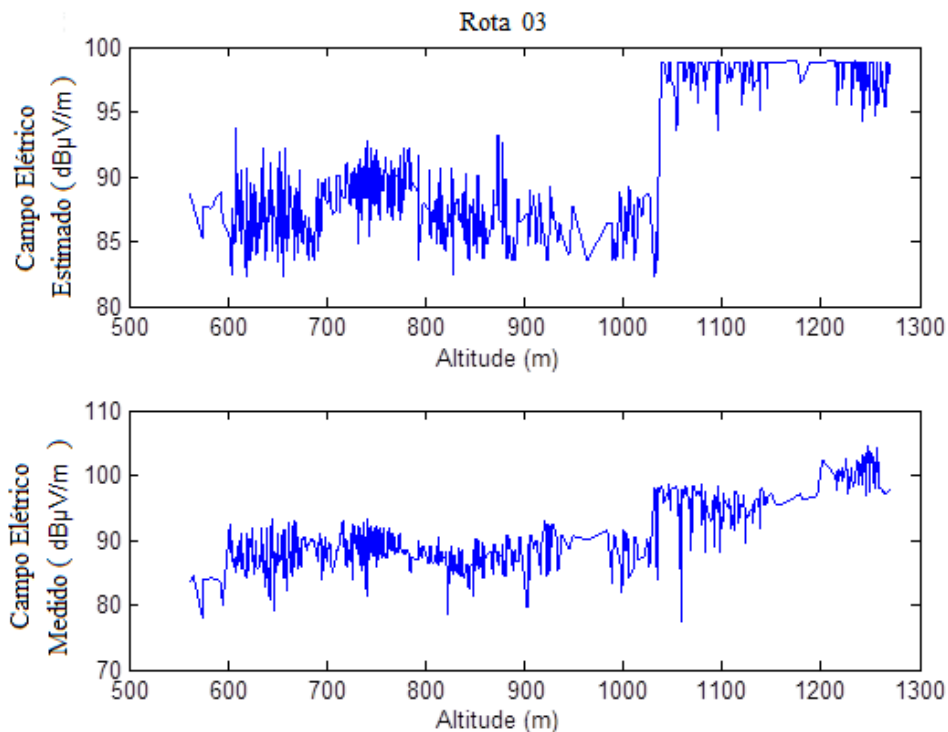


Figura 5.11: Campo Elétrico Estimado x Campo Elétrico Medido para a Rota 03

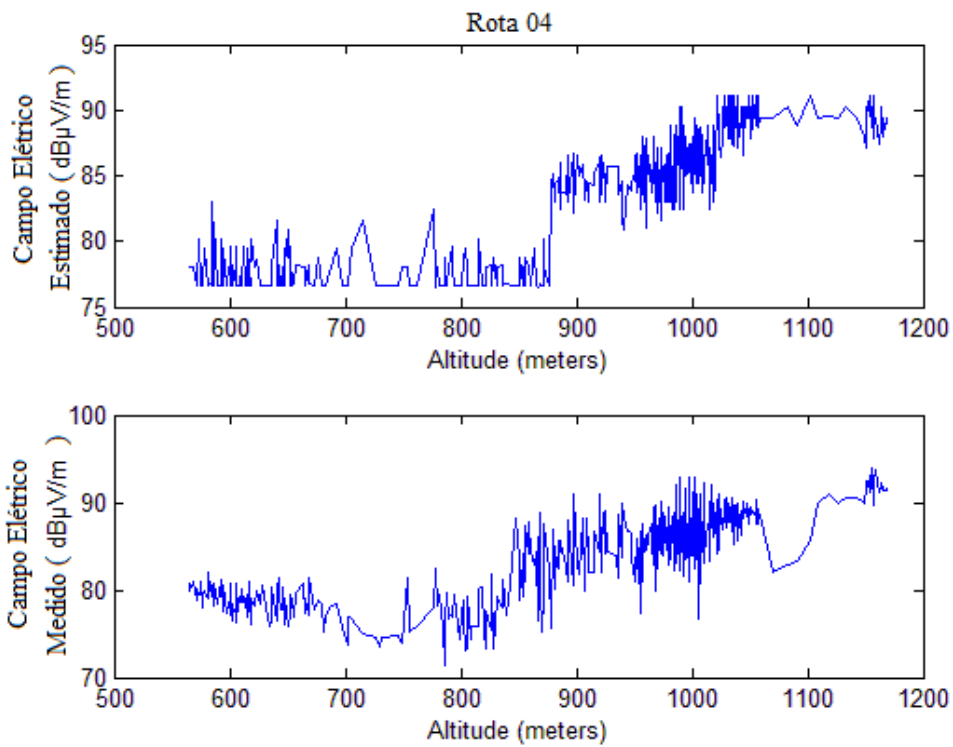


Figura 5.12: Campo Eléctrico Estimado x Campo Eléctrico Medido para a Rota 04

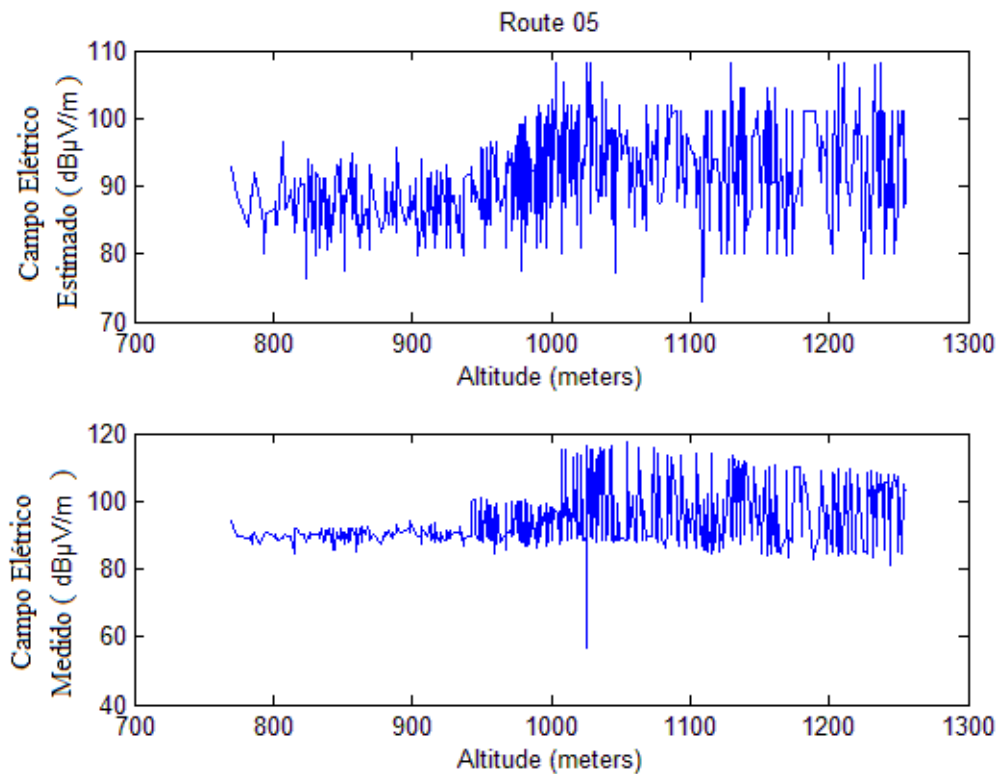


Figura 5.13: Campo Eléctrico Estimado x Campo Eléctrico Medido para a Rota 05

Os valores calculados para as rotas 03, 04 e 05 foram os seguintes:

Tabela VIII: Dados Calculados para as Rotas 03, 04 e 05

	Rota 03	Rota 04	Rota 05
EMQ (dB)	3,56	2,94	9,7
Correlação Campo Elétrico Medido x Campo Elétrico Estimado	0,7413	0,794	0,1563
Correlação Campo Elétrico Medido x Altitude	0,7523	0,8671	0,346

A fim de se solidificar a validação do trabalho, foram utilizadas as linhas de tendência, que são instrumentos muito utilizados em problemas que requerem uma análise técnica. Estas linhas podem ser formadas por mínimos relativos consecutivos (Linha de Tendência de suporte ou Linha de Procura), que foi a utilizada neste trabalho, ou por máximos relativos consecutivos (Linha de Tendência de Resistência ou Linha de Oferta).

Os gráficos a seguir representam uma comparação entre o comportamento das linhas de transição para os campos elétricos medidos e estimados, com a finalidade de demonstrar claramente a correlação entre ambos.

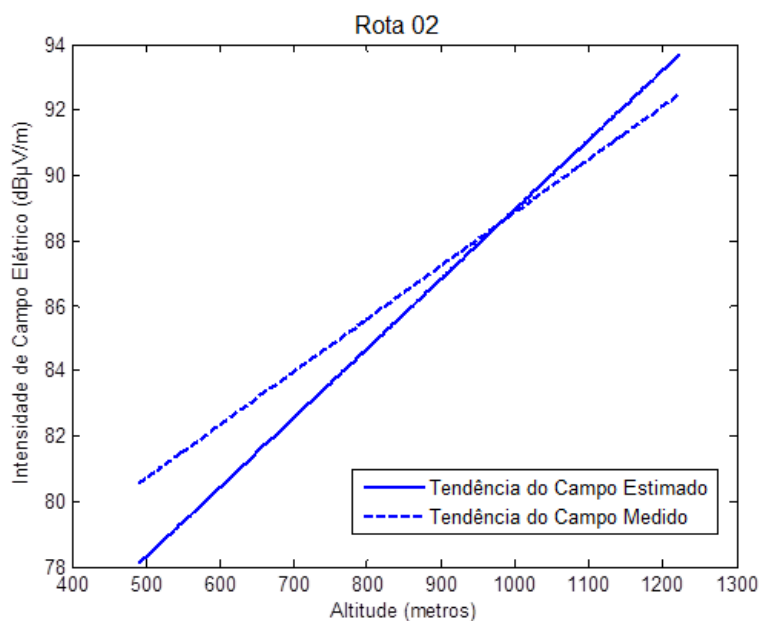


Fig. 5.14: Linhas de transição para a Rota 02

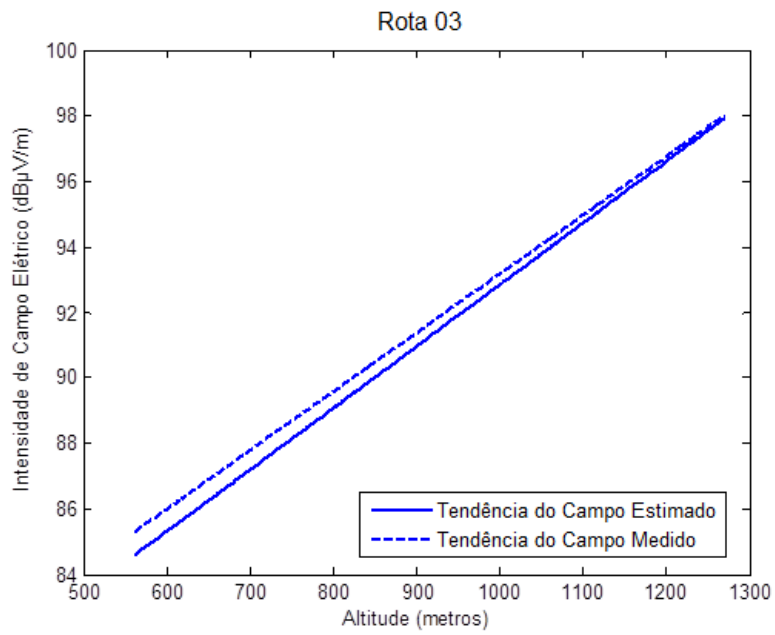


Fig. 5.15: Linhas de transição para a Rota 03

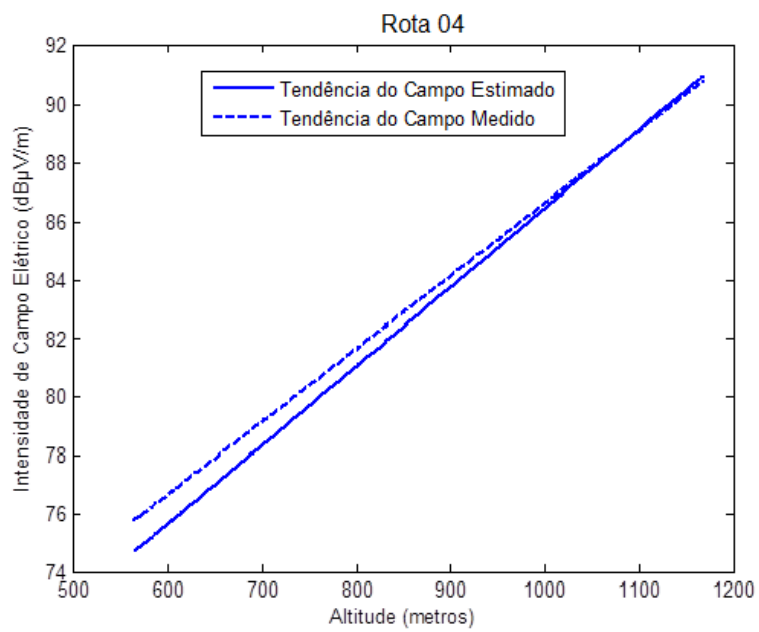


Fig. 5.16: Linhas de transição para a Rota 04

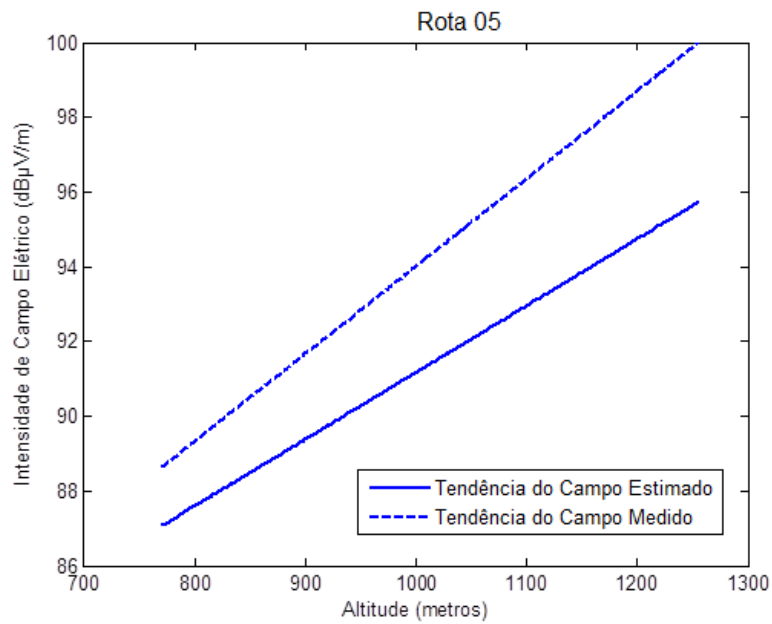


Fig. 5.17: Linhas de transição para a Rota 05

### 5.3 CONSIDERAÇÕES FINAIS

Neste capítulo foi apresentada uma nova metodologia para determinar o campo elétrico de radiodifusão sonora em ondas médias. Utilizando Inferências Bayesianas foi possível desenvolver uma metodologia capaz de estimar os valores do campo elétrico na região do Distrito Federal e áreas vizinhas (que incluíam os estados de Goiás e Minas Gerais).

Os gráficos que ilustram a comparação entre os valores de campo elétrico medido e estimado bem como as linhas de tendência desses sinais e as tabelas geradas pela Rede Bayesiana mostram que a metodologia está funcionando de forma satisfatória para as rotas 02, 03 e 04. Seus valores de Erro Médio Quadrático e correlações entre Campo Elétrico Estimado e Campo Elétrico Medido, e Campo Elétrico Medido e Altitude estão dentro dos padrões considerados aceitáveis.

Os valores encontrados para a Rota 05 não estão de acordo com os níveis de aceitação para validação da metodologia. Essa diferença entre as rotas utilizadas pode ser explicada pela grande oscilação no nível de terreno que existe nesta rota, claramente ilustrada na Figura. 4.7.

Então, pode-se concluir com os resultados encontrados, que a metodologia proposta é viável para utilização em terrenos com poucos aclives e declives, sendo perfeitamente adequada para prever o campo elétrico sem a utilização de campanha de medição.

## CAPÍTULO 6

### CONCLUSÃO

A radiodifusão sonora sempre foi um dos principais meios de comunicação, com seu imprescindível papel social quanto à divulgação de informações e entretenimento. Devido à baixa qualidade oferecida aos usuários desses sistemas, o mesmo vem, gradativamente, perdendo espaço para novas propostas de acesso à informação. Com o objetivo de mudar este cenário, órgãos internacionais sentiram a necessidade de melhorar este serviço. Esta melhora consiste, principalmente, na digitalização dos serviços de radiodifusão sonora, permitindo melhora na qualidade do serviço oferecido aos ouvintes, maior número de emissoras e muito mais opções de entretenimento oferecidas por elas.

Neste processo de modernização do sistema, que pode consistir tanto na expansão de um sistema já existente quanto no planejamento de novos sistemas, uma análise detalhada do canal de rádio é de suma importância. O capítulo 2 abordou os modelos de perda de propagação de sinal que mostram quais são os parâmetros que mais afetam este tipo de propagação. Nos sistemas digitais, a recepção do sinal é muito sensível a pequenas variações da intensidade de campo recebido, sendo que quando os valores de campo forem inferiores ao limiar de recepção, pode ocorrer a perda completa da disponibilidade do serviço. Por isso é imprescindível que os parâmetros utilizados nos modelos de predição sejam mais próximos possíveis dos valores reais.

Essas ondas são muito influenciadas pelo solo (suas características eletromagnéticas, sua rugosidade e condutividade, por exemplo), dificultando a precisão dos modelos de propagação. Além disso, a presença de construções e vegetação, bem como as variações no relevo também afetam os valores de campo elétrico. Este trabalho buscou encontrar a solução para se determinar o campo elétrico em uma onda de superfície, facilitando o projeto de novos sistemas de comunicação de rádio e a expansão de sistemas já existentes. Para isso, utilizou-se como inteligência computacional as Inferências Bayesianas, que foram apresentadas no Capítulo 03.



Foram coletadas amostras ao longo de 06 rotas durante campanhas de medição descritas no Capítulo 04. As informações obtidas pela equipe de medições foram de que as amostras coletadas próximas às linhas de transmissão e placas de sinalização tiveram seus valores alterados, sendo bem mais baixos do que a média.

Após o tratamento dos dados no capítulo 05, teve-se a certeza da eficiência das Redes Bayesianas tanto em sua representação (em forma de grafos) quanto na especificação concisa de qualquer distribuição de probabilidade conjunta, dada pelas tabelas de probabilidade condicional.

Durante a análise dos dados coletados, pôde-se perceber que há uma queda da intensidade do campo elétrico em pontos em que há depressões no terreno, e que algumas amostras têm nível superior à média. Essas diferenças de valores permitiram que fosse feita a *clusterização* seguindo o método adaptativo (eliminando pontos redundantes e reagrupando pontos próximos), que foi inteiramente aceitável por não provocar nenhuma alteração na curva em estudo.

Os valores de campo elétrico estimado pelas redes bayesianas mostraram-se satisfatórios desde os gráficos até os valores de erro médio quadrático, correlação entre campo elétrico medido e altitude e correlação entre campo elétrico medido e campo elétrico estimado encontrados para as rotas 02, 03 e 04. Para a rota 05 os resultados encontrados não foram satisfatórios.

A rota 05 é topograficamente diferente das demais, sendo que nesta há inúmeros aclives e declives. Como relatado anteriormente, características físicas do relevo interferem diretamente no valor de campo elétrico. Pode-se considerar este como o fator responsável pelos valores não satisfatórios para esta rota.

Em resumo, considera-se que o objetivo inicial do trabalho foi alcançado, isto é, uma pesquisa que permite o desenvolvimento de uma nova metodologia para estimar o valor do campo elétrico em ondas superficiais utilizando redes Bayesianas.

Este trabalho possibilita inúmeras opções para trabalhos futuros. Pode-se validar a metodologia utilizando dados coletados em outras regiões, considerando-se as mesmas

condições de medição. Pode-se testar também esta mesma metodologia para transmissões com frequências diferentes da utilizada neste trabalho.

Uma outra proposta é desenvolver um modelo para TV Digital, levando em conta a altitude do terreno. Assim, uma simulação de cobertura será feita para melhor localizar o transmissor, utilizando a mesma metodologia proposta neste trabalho para rádio digital.

Além disso, pode-se aperfeiçoar a rede bayesiana com outra versão do software Bayesware, considerando o número total de amostras, sem a necessidade de utilizar o método de Clusterização.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Parsons, D., *The Mobile Radio Propagation Channel*. John Wiley & Sons, Inglaterra, 1998.
- [2] MONTEIRO, B. *Predição dos parâmetros elétricos do solo para a radiodifusão em onda média utilizando redes neurais*. 2008. 97f. Dissertação (Mestrado em Engenharia Elétrica) – Instituto de Tecnologia, Universidade Federal do Pará, Pará. 2008.
- [3] MELO, K. C. *Predição da intensidade do campo elétrico da onda de superfície utilizando redes neurais artificiais*. 2008. 109f. Dissertação (Mestrado em Engenharia Elétrica) – Departamento de Engenharia Elétrica, Universidade de Brasília, Distrito Federal. 2008.
- [4] LIMA, F. *Estudo da propagação de sinal em ondas médias: Contribuições para a implantação da radiodifusão digital no Brasil*. 2008. 87f. Tese (Doutorado em Engenharia Elétrica) – Faculdade de Tecnologia, Universidade de Brasília, Distrito Federal. 2008.
- [5] Ott, R.H., Vogler, L.E. and Hufford, G.A., *Ground-wave propagation over irregular, inhomogeneous terrain: Comparison of calculations and measurements*. NTIA Report 79-20, 1979. (NTIS Accession No. PB 298668/AS).
- [6] [http://www.wirelessbrasil.org/wirelessbr/colaboradores/marcio\\_rodrigues/propagacao/prop\\_06.html](http://www.wirelessbrasil.org/wirelessbr/colaboradores/marcio_rodrigues/propagacao/prop_06.html), acessado em 17 de janeiro de 2013.
- [7] Rappaport, T. S, *Wireless Communications: Principles and Practice*, 2ndEdition, Prentice Hall, New York, 2002.
- [8] ARAÚJO, J. P. *Estratégia de planejamento e otimização de sistemas sem fio, considerando redes interferentes: abordagem baseada em cross-layer*. 2011. 138f. Tese (Doutorado em Engenharia Elétrica) – Instituto de Tecnologia, Universidade Federal do Pará, Pará. 2011.
- [9] HUANG, H., SONG, H., TIAN, F., LU, Y., WANG, Q. *A comparatively research in incremental learning of Bayesian networks*, *Intelligent Control and Automation*. Fifth World Congress on, Volume 5, 4260-4264 (2004).

- [10] BOBBIO, A.; PORTINALE, L.; MINICHINO, M.; CIANCAMERLA, E.. Improving the Analysis of Dependable Systems by Mapping Fault Trees into Bayesian Networks. *Reliability Engineering & System Safety*, Vol. 71, 2001.
- [11] ARA-SOUZA, Anderson Luiz. **Redes Bayesianas: uma introdução aplicada a credit scoring**. São Paulo: SINAPE, 2010. 98f.
- [12] KORB, K. B.; NICHOLSON, A. E.. *Bayesian artificial intelligence*. London: Chapman & Hall/CRC Press UK, 2004.
- [13] PEARL, J., *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.
- [14] LAM, W. & BACCHUS, F., Learning Bayesian Belief Networks, Na Approach based on the MDL principle. *Computational Intelligence*, 10, 269-293, 1994.
- [15] SUZUKI, J. - A construction of bayesian networks from databases based on MDL scheme. D. Heckerman & A. Mandani Eds., *Uncertainty in Artificial Intelligence: Proceedings of the Ninth Conference*, pp. 266-273, Washington DC, 1993.
- [16] SUZUKI, J - On an efficient mdl learning procedure using branch and bound technique. Technical Report COMP95-27 (1995-06), Institute of Electronics, Information and Communication Engineers, 1995.
- [17] BUNTINE, W. - Operations for learning with graphical models. *Journal of Artificial Intelligence Research*, (2):159-225, 1994
- [18] HECKERMAN, D. Bayesian networks for Data Mining. *Data Mining and Knowledge Discovery*, Kluwer Academic Publishers, (1): 79-119, (1997)
- [19] Heckerman, D., Geiger, D., and Chickering, D., (1995). Learning Bayesian Networks: The combination of knowledge and statistical data. *Machine Learning*, 20:197-243
- [20] DAWID, P. Statistical theory – the frequential approach. *Journal of the Royal Statistical Society*. (147): 178-292 (1984).
- [21] Relatório técnico - Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas. Cássio Oliveira Camilo e João Carlos da Silva
- [22] SANTANA, A. L., FRANCÊS, C.R.L, COSTA, J.C.W. *Hybrid Strategies for improving Bayesian Networks: Applying Mathematical and Computational Intelligence Models to Optimize and Extend the Modelling and Applicability*, Lambert Academic Publishing (2011).

- [23] MONTEIRO, J.; RODRIGUES, J.C. ; FRAIHA, S.C. ; ARAUJO, J. ; GOMES, H.S. ; FERREIRA, A. ; FRANCES, C.R.L. ; CAVALCANTE, G.P.S. **A methodology for determining the electric field of broadcasting medium wave using computational intelligence approach**, 6<sup>th</sup> European Conference on Antennas and Propagation, Prague: 1190-1193, (2012).
- [24] <http://www.drm-brasil.org>, acessado em 16 de julho de 2012.
- [25] Mitchell, T.M. (1997). Machine Learning. McGraw Hill, New York