

Carlos Augusto Uchôa da Silva

**UM MÉTODO PARA ESTIMAR
OBSERVÁVEIS GPS USANDO
REDES NEURAIS ARTIFICIAIS**

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Doutor em Engenharia Civil: Transportes.

Orientador: Prof. Associado Paulo Cesar Lima Segantine

São Carlos
2003

DEDICATÓRIA

À minha esposa Kátia Cynara e
meus filhos Rayra, Rahym e
Raimi, meus maiores presentes.

AGRADECIMENTOS

A Deus pela esperança, pelos dons, pelo existir.

A Kátia, a mulher que amo, por seu amor, por se tornar pai e mãe de nossos filhos em meus intermináveis devaneios pelo mundo acadêmico. A Maina por sua dedicação e carinho.

A Aristides, meu pai, por seu apoio. Seu caráter, inteligência e criatividade sempre me inspiraram. A Helena, minha mãe, por seu amor. Sua emoção, por vezes inocente, mesmo à distância, sempre me afagaram.

Ao clã dos Uchôas, Conce, Beth e Humberto pela torcida, pela saudade e pelo carinho. A Rita, que com sua amizade, generosidade e desapego, tornou possível mais esse passo.

Aos meus cunhados. Os dois, cada um ao seu tempo, cada um de seu jeito, foram importantes. Antônio no início, na saída de Belém e a César por sua ajuda providencial em nosso regresso.

Ao amigos do forninho. Genival, Tule, Maurício e Rodrigo são os irmãos que escolhi em São Carlos, considerem o “Meio-inteiros” meu agradecimento especial a cada um e a todos os amigos nele citados.

A Rodrigo, em especial por sua essencial ajuda ao desenvolvimento desse projeto, seu talento e doação fizeram dele o autor de diversas das implementações computacionais utilizadas nesta pesquisa.

A meu orientador, Prof. Associado Paulo César Lima Segantine, por suas correções e sugestões para este trabalho.

A Kátia Duarte, do Departamento de geodésia do IBGE pelo fornecimento dos dados da RBMC.

A Profa. Roseli Romero, por sua ajuda na condução da escolha da modelagem neural.

Aos amigos do CEFET-Pa, pela torcida, pela força, um agradecimento especial aos amigos Expedito, Sérgio e Luis Carlos.

Finalmente, a toda minha família, em especial a Rita Barbosa, minha avó, ela se foi em 1994. Acho que herdei dela o seu melhor. Sua perseverança e altivez sempre encontraram eco em mim e definitivamente me impulsionaram a chegar até aqui, sinto saudades.

SUMÁRIO

| | |
|---|-----------|
| LISTA DE FIGURAS | i |
| LISTA DE TABELAS..... | iv |
| LISTA DE SIGLAS | v |
| RESUMO | vii |
| ABSTRACT | viii |
| 1 INTRODUÇÃO | 1 |
| 1.1. Objetivos..... | 4 |
| 1.2. Justificativas | 4 |
| 2 O ESTADO DA ARTE..... | 6 |
| 3 O SISTEMA DE POSICIONAMENTO GLOBAL NAVSTAR/GPS | 9 |
| 3.1. O Sinal GPS | 10 |
| 3.2 As observáveis GPS | 12 |
| 3.2.1. Pseudodistância | 13 |
| 3.2.2 Fase da portadora | 14 |
| 3.3 O Processamento de dados GPS..... | 17 |
| 3.3.1 Posicionamento Pontual Preciso (PPP)..... | 18 |
| 3.3.1.1 O Auto-GIPSY (AG) | 20 |
| 3.4 O formato RINEX..... | 22 |
| 4 REDES NEURAIS ARTIFICIAIS (RNAs)..... | 24 |
| 4.1 Inspiração Biológica..... | 25 |
| 4.2 Neurônio Artificial | 27 |
| 4.3 Topologia ótima..... | 28 |
| 4.4 Características e propriedades úteis das RNAs..... | 29 |

| | |
|--|-----------|
| 4.5 Os Modelos Neurais..... | 30 |
| 4.5.1 O Modelo Perceptron Multi Camadas (Multilayer Perceptrons ou MLP)..... | 31 |
| 4.5.2 O algoritmo backpropagation..... | 33 |
| 4.5.2.1 Algoritmo de treinamento rápido..... | 35 |
| 5 MATERIAIS E MÉTODOS..... | 37 |
| 5.1 Materiais..... | 37 |
| 5.1.1 Critérios adotados para escolha dos dados e desenvolvimento do método proposto..... | 46 |
| 5.2 Metodologia..... | 48 |
| 5.2.1 Etapas realizadas nesta investigação..... | 49 |
| 5.2.2 Procedimentos adotados..... | 50 |
| 5.2.2.1 Formação dos conjuntos de dados para treinamento, validação e testes do modelo neural..... | 51 |
| 5.2.3 Estratégias de Processamento..... | 52 |
| 5.2.3.1 Árvore de testes para definição da topologia adequada à melhor generalização da rede neural..... | 53 |
| 5.2.3.2 Processamentos dos dados GPS..... | 56 |
| 5.3 O Programa MORE-GPS versão 1.0 – Manipulação e Modelagem de Observáveis GPS por Redes Neurais Artificiais..... | 57 |
| 6 APRESENTAÇÃO DOS RESULTADOS..... | 65 |
| 6.1 A topologia escolhida..... | 65 |
| 6.2 O tempo de observação..... | 70 |
| 6.3 Características do agrupamento das estações RBMC mais adequado ao modelamento neural dos dados..... | 74 |
| 6.4 Testes do método para diferentes conjuntos de dados..... | 76 |
| 6.5 Validação do método..... | 79 |
| 6.5.1 Dados para validação..... | 80 |
| 6.5.1.1 Processamento por ponto simples (single point)..... | 80 |
| 6.5.1.2 Processamento relativo estático..... | 85 |
| 6.5.1.3 Processamento Posicionamento Pontual Preciso (PPP)..... | 98 |

| | |
|---|-----|
| 7 CONCLUSÕES E RECOMENDAÇÕES | 101 |
| 7.1 Quanto à modelagem neural | 101 |
| 7.2 Quanto ao sistema MORE-GPS..... | 102 |
| 7.2.1 Limitações do método proposto | 103 |
| 7.3 Quanto ao processamento GPS | 104 |
| 7.3.1 O processamento pontual | 104 |
| 7.3.2 O processamento relativo | 105 |
| 8 REFERÊNCIAS BIBLIOGRÁFICAS | 108 |

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1 – Modulação Bifásica da onda portadora. | 11 |
| Figura 2 – Interpretação geométrica da fase da portadora.. | 17 |
| Figura 3 – Localização das estações que compõe a rede do IGS..... | 19 |
| Figura 4 – Representação resumida do processamento pelo AG..... | 22 |
| Figura 5 – Representação genérica de um neurônio biológico e processo de transmissão sináptica..... | 25 |
| Figura 6 – Modelo não linear de neurônio artificial | 27 |
| Figura 7 – Fac-símile de trecho de um arquivo RINEX de observação da estação BOMJ-RBMC. | 38 |
| Figura 8 – Fac-símile de trecho de um arquivo de efemérides precisas, formato SP3, correspondente ao dia 339 do ano 2001. | 40 |
| Figura 9 – Localização das estações RBMC | 42 |
| Figura 10 – Frequência de observações/estação/ano..... | 44 |
| Figura 11 – N° de estações operando simultaneamente em 2000..... | 45 |
| Figura 12 – N° de estações RBMC operando simultaneamente em 2001. | 45 |
| Figura 13 – N° de estações RBMC operando simultaneamente em 2002. | 46 |
| Figura 14 – Representação gráfica do número Sunspot do Ciclo Solar 23..... | 48 |
| Figura 15 – Tela de inicialização do programa MORE-GPS | 59 |
| Figura 16 – Tela principal do programa MORE-GPS..... | 59 |
| Figura 17 – Fluxograma descritivo do programa MORE-GPS..... | 64 |
| Figura 18 – Desempenho rede 1:6:1 para 5 minutos de observação | 68 |

| | |
|---|----|
| Figura 19 – Desempenho rede 1:10:1 para 5 minutos de observação | 68 |
| Figura 20 – Desempenho alcançado pela estimativas de duas horas de observação..... | 73 |
| Figura 21 – Agrupamento regional das estações RBMC | 75 |
| Figura 22 – Desempenho do treinamento para 5 minutos com agrupamento regional de estações RBMC..... | 76 |
| Figura 23 – Desempenho do treinamento para 5 minutos no dia 17/05/2000 (Dia GPS 138) | 77 |
| Figura 24 – Desempenho do treinamento para 5 minutos no dia 29/01/2002 (Dia GPS 29) | 78 |
| Figura 25 – Desempenho do treinamento para 5 minutos no dia 08/03/2002 (Dia GPS 67) | 78 |
| Figura 26 – Desvios para a estação VICO | 82 |
| Figura 27 – Desvios para a estação UEPP | 83 |
| Figura 28 – Desvios para a estação RIOD..... | 85 |
| Figura 29 – Estações para o processamento I | 86 |
| Figura 30 – Desvios linha base BOMJ-VICO..... | 88 |
| Figura 31 – Desvios linha base BRAZ-VICO | 88 |
| Figura 32 – Desvios linha base UEPP-VICO | 89 |
| Figura 33 – Desvios linha base RIOD-VICO..... | 89 |
| Figura 34 – Estações para o processamento II..... | 90 |
| Figura 35 – Desvios linha base BRAZ-UEPP..... | 92 |
| Figura 36 – Desvios linha base CUIB-UEPP | 93 |
| Figura 37 – Desvios linha base RIOD-UEPP | 93 |
| Figura 38 – Desvios linha base VICO-UEPP | 93 |
| Figura 39 – Desvios linha base PARA-UEPP | 94 |
| Figura 40 – Estações envolvidas no processamento III..... | 95 |
| Figura 41 – Desvios da linha base IMPZ-RIOD..... | 97 |
| Figura 42 – Desvios da linha base BRAZ_RIOD..... | 97 |
| Figura 43 – Desvios da linha base VICO-RIOD..... | 97 |
| Figura 44 – Desvios da linha base UEPP-RIOD..... | 98 |

Figura 45 – Desvios da linha base PARA-RIOD98

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 – Resumo de alguns componentes dos sinais dos satélites..... | 10 |
| Tabela 2- Descrição resumida do Algoritmo de aprendizado backpropagation padrão..... | 34 |
| Tabela 3 – Informações das estações RBMC em maio 2003..... | 43 |
| Tabela 4 – Resumo dos dados escolhidos para modelagem. | 47 |
| Tabela 5 – Comparativo entre diferentes topologias testadas para fase. . | 69 |
| Tabela 6 – Comparativo entre diferentes topologias testadas para código. | 70 |
| Tabela 7 – Desvios para o posicionamento pontual I..... | 82 |
| Tabela 8 – Desvios para o posicionamento pontual II | 83 |
| Tabela 9 – Desvio para o posicionamento pontual III..... | 84 |
| Tabela 10 – Coordenadas oficiais das estações utilizadas no processamento I..... | 87 |
| Tabela 11 – Desvios do processamento I | 87 |
| Tabela 12 - Coordenadas oficiais das estações utilizadas no processamento II | 91 |
| Tabela 13 – Desvios do processamento II..... | 91 |
| Tabela 14 - Coordenadas oficiais das estações utilizadas no processamento III..... | 95 |
| Tabela 15 – Desvios do processamento III | 96 |
| Tabela 16 –Coordenadas geodésicas ITRF2000 das estações VICO, UEPP e RIOD (PPP)..... | 99 |

LISTA SIGLAS

| | |
|----------|--|
| ASCII | American Standard Code for information Interchange |
| AG | Auto Gipsy |
| C/A | Coarse/Acquisition code |
| DEGED | Departamento de Geodésia |
| DOY | Day Of the Year (0-365) |
| ESSC-USP | Escola de Engenharia de São Carlos-Universidade de São Paulo |
| FIR | Filter Inpulse Response |
| FLINN | Fiducial Laboratories for International National Science Network |
| ftp | File Transfer Protocol |
| GIPSY | GPS Inferred Positioning System |
| GLONASS | Global Navigation Satellite System |
| GNSS | Global Navigation Satellite Systems |
| GPS | Global Positioning System |
| IBGE | Instituto Brasileiro de Geografia e Estatística |
| IGS | International GPS Service for Geodynamics |
| ITRF | International Terrestrial Reference Frame |
| JPL | Jet Propulsion Laboratory |
| MLP | Multi-Layer Perceptron |
| MSE | Mean of Squared Error |
| N | Ambiguidade |

| | |
|---------|---|
| NASA | National Aeronautics and Space Administration |
| NGS | National Geodetic Survey, USA |
| NAVSTAR | Navigation System using Time And Ranging |
| NOAA | National Oceanic & Atmospheric Administration |
| OASIS | Orbit Analysis and Simulation Software |
| P | Precise code (Código preciso) |
| PPP | Precise Point Positioning |
| PRN | Pseudo Random Noise |
| RBMC | Rede Brasileira de Monitoramento Contínuo |
| RINEX | Receiver Independent Exchange Format |
| RNAs | Redes Neurais Artificiais |
| RMS | Root Mean Square |
| SA | Selective Availability |
| SAD69 | South American Datum 1969 |
| SGB | Sistema Geodésico Brasileiro |
| SIVAM | Sistema de Vigilância da Amazônia |
| SNNS | Stuttgart Neural Network Simulator |
| SOM | Self Organizing Map |
| SPS | Standard Position Service |
| SP3 | Formato de efemérides usado pelo NGS |
| SVN | Space Vehicle Number |
| TEC | Total Electron Content |
| URL | Uniform Resource Locator |
| WGS84 | World Geodetic System 1984 |
| WWW | World Wide Web |

RESUMO

SILVA, C. A. U. UM MÉTODO PARA ESTIMAR OBSERVÁVEIS GPS USANDO REDES NEURAIAS ARTIFICIAIS. São Carlos, 2003, 112 p. Tese – Escola de Engenharia de São Carlos, Universidade de São Paulo.

O NAVSTAR-GPS, com uma grande variedade de conjuntos receptores e sua aplicabilidade prática em diversas áreas, transformou-se no mais difundido dos sistemas de posicionamento. Porém, necessidades cada vez maiores em termos de precisão trouxeram consigo o ônus de um custo elevado com a aquisição de equipamentos de dupla frequência. Este trabalho consiste no desenvolvimento de um método que possibilite a modelagem das observáveis GPS, através de Redes Neurais Artificiais, bem como a agregação destes dados a um arquivo gerado por um receptor de uma frequência, conferindo-lhe características específicas de arquivos gerados por receptores de dupla frequência e código P. Isto possibilita que dados gerados por receptores de uma frequência, a imensa maioria dos receptores utilizados no Brasil, possam ser processados como vetores de bases longas. Os resultados obtidos indicam que o uso de modelos neurais, treinados por algoritmos de aprendizado supervisionado, são uma alternativa promissora para estimar dados GPS.

Palavras-chave: GPS; Redes Neurais Artificiais, Modelagem, portadora L2, código P.

ABSTRACT

SILVA, C. A. U. A METHOD TO ESTIMATE GPS DATA OBSERVABLES USING ARTIFICIAL NEURAL NETWORKS. São Carlos, 2003. 112 p. Thesis - School of Engineering of São Carlos, University of São Paulo.

The NAVSTAR-GPS, with a great variety of receivers and its practical applicability in several areas, transformed itself in the most known positioning system. But the necessity of improving the results precision brings with it a cost increasing caused by the use of equipments of dual frequency equipments. This work consist on the development of a method that makes possible the GPS data modelling using Neural Networks, as well as the aggregation of these data into a file generated by single frequency receiver, providing to the system specific characteristics of files generated by double frequency an P code receiver. This makes possible that data generated by receivers of single frequency, the majority of receivers in Brazil, can be processed as vectors of long bases. The results obtained indicate that the use of Neural Network models, with algorithms of supervised learning are a promissing alternative to estimate GPS data.

Keywords: GPS; Artificial Neural Networks, Modelling, L2 Carrier, P code.

1

INTRODUÇÃO

O NAVSTAR/GPS, do inglês *Navigation System using Time and Ranging/Global Positioning System*, é um sistema de posicionamento global, fundamentado em rádio-navegação por satélites, que utiliza diferentes observáveis para calcular distâncias entre os centros de fase das antenas dos satélites e de receptores localizados na superfície da terra ou próximo dela, com o objetivo de determinar a posição tridimensional dos pontos de interesse. Neste trabalho, o sistema será tratado apenas por GPS.

Inicialmente desenvolvido e implementado com fins militares, o GPS foi liberado posteriormente pelo congresso americano para uso civil, ainda que com degradações intencionais em sua precisão, PARKINSON (1996).

Após essa liberação, ocorreram profundas transformações no conceito de posicionamento em diversas áreas do conhecimento com diferentes aplicações, tais como: Mensuração e suas áreas afins, Transporte terrestre, aéreo e naval, Telefonia móvel, Monitoramento de segurança, Serviços de emergência, Monitoramento de fauna e flora, Controle de queimadas, Fiscalização de desmatamento e outros.

Resumidamente, pode-se dizer que os primeiros projetos, da chamada “Era espacial”, datam do início dos anos 50. Após duas décadas de pesquisas, os americanos e os russos implantaram o GPS e o GLONASS, respectivamente, que são hoje os únicos sistemas de posicionamento global operacionais. Ambos compõem o chamado GNSS (*Global Navigation Satellite System*).

Existe ainda um projeto liderado pelos países membros da união européia para implantação de um outro sistema, o GALILEO. Este está sendo desenvolvido principalmente para aplicações civis, com previsão de lançamento dos primeiros satélites para 2005.

A operação contínua do sistema GPS, a grande variedade de receptores no mercado, e sua aplicabilidade prática em diversas áreas o transformaram no mais popular dos sistemas de posicionamento. Porém, as necessidades cada vez maiores em termos de precisão no posicionamento trouxeram consigo o ônus de um custo elevado na aquisição de conjuntos receptores e dependendo da estratégia de posicionamento, a necessidade de um aumento considerável nos tempos de coleta e de processamento dos dados.

A possibilidade de se obter uma melhor precisão no posicionamento por ponto simples passou então a ser objeto de estudo de vários pesquisadores, visando eliminar a necessidade de posicionamento relativo, minimizar o tempo da sessão de observação e otimizar o custo dos levantamentos.

Nesse sentido, novas possibilidades têm se apresentado promissoras, tais como, o Posicionamento Pontual Preciso, também chamado de método PPP¹, que proporciona precisão tridimensional da ordem de alguns centímetros e demanda baixo esforço computacional em relação às redes, MONICO (2001). Maiores informações acerca deste método são apresentadas no item 3.3.1.

O termo posicionamento pontual, do inglês *single point*, refere-se tradicionalmente as estimativas das coordenadas de uma determinada estação, através da análise de pseudodistâncias entre as antenas dos satélites e de um receptor, geralmente baseadas no código C/A (com limitações impostas pelos parâmetros contidos na mensagem de navegação, que têm precisão de poucos metros para órbitas e de dezenas de metros para relógios), ZUNBERGE; BERTINGER (1996) apud ZUNBERGE (1997).

O uso de receptores de dupla frequência possibilita ao uso de um modelo de combinação linear dos sinais da portadora L1 e L2, definida como L3, ou

¹ Precise Point Positioning

ainda, *ion free*, onde são eliminados os efeitos da ionosfera na propagação do sinal GPS, geralmente uma fonte de erros importante em se tratando de vetores de linhas-base longas.

Desde o surgimento do GPS, várias foram as questões e dúvidas levantadas sobre ele, quase todas no sentido de melhor compreendê-lo e aprimorá-lo numa incessante tentativa de explorar todas as suas possibilidades. Uma delas é: **seria possível fazer com que dados observados por receptores de uma frequência (L1) possam ser processados de forma a se obter um nível de precisão similar àqueles gerados por receptores de dupla frequência (L1 e L2)?** Uma segunda questão é: **seria possível modelar, prever ou estimar dados GPS?** e ainda uma outra pergunta surge: **Qual ferramenta seria adequada a esta modelagem?**

LUO; UNBEHAUEN (1998) lançaram uma luz sobre estas perguntas quando afirmaram que a utilização de Redes Neurais Artificiais no processamento de sinais estava se ampliando com aplicações em diversas áreas como filtragem, estimativa de parâmetros, detecção de sinais, reconhecimento de padrões, reconstrução de sinais, sistemas de identificação, compreensão e transmissão.

Redes Neurais Artificiais, neste trabalho, tratadas apenas por Redes Neurais ou ainda pela sigla RNAs, têm seu uso indicado nesta investigação em função da excelente capacidade de generalização, classificação, interpolação e extrapolação, tolerância a erros e ruídos e também pelo fato de não haver a necessidade prévia de parametrizações explícitas na modelagem.

A hipótese inicial desta investigação configura-se então na asserção de que é possível estimar dados GPS da portadora L2 e código P2 utilizando modelagem a partir de RNAs, a despeito dos erros inerentes à propagação do sinal GPS, quer intencionais ou não, da época e do local onde tenham sido observados, desde que sejam satisfeitas determinadas restrições.

Estas restrições referem-se à necessidade de existirem dados de um conjunto de estações GPS, equipadas com receptores de dupla frequência e código P, dependendo dos objetivos do usuário, rastreando simultaneamente

àquele que se deseja estimar, dentro de uma determinada região de abrangência. Os limites desta região são objeto de estudos futuros. Existe ainda uma restrição relativa à necessidade de existência de efemérides precisas em formato SP3.

1.1. Objetivos

O objetivo central deste trabalho é desenvolver um método para estimar dados GPS, mais especificamente, a contagem de ciclos da fase da portadora L2 e a pseudodistância a partir do código P2, para serem agregados aos dados de um receptor de única frequência (L1 e C/A), obtendo-se um novo conjunto (L1, C/A, L2Modif, P2Modif). Esses dados serão estimados através da utilização de um modelo de rede neural, usando algoritmos de aprendizado supervisionado, alimentados por dados oriundos de receptores de dupla frequência (L1, C/A, L2, P2) da Rede Brasileira de Monitoramento Contínuo (RBMC), mantida pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Este novo conjunto de dados (L1, C/A, L2Modif, P2Modif) modificado, será testado mediante diferentes condutas de processamentos tais como: ponto simples e processamento relativo de bases longas.

1.2. Justificativas

Esta investigação justifica-se principalmente pela tentativa de elevar a precisão da posição tridimensional de uma dada estação GPS, observada pelo método de posicionamento simples com receptor GPS de uma frequência, a um nível similar às coordenadas de uma estação observada com um receptor de dupla frequência e código P.

Isto proporciona uma expressiva redução de custo no eventual uso de receptores de dupla frequência, função da substituição destes, por receptores de uma frequência. Se utilizado o PPP, ainda se obtém uma segunda vantagem, esta, relativa ao fato de que esse método não necessita do uso simultâneo de dois ou mais receptores, diretamente pelo usuário.

Este trabalho pode também ser legitimado em função do uso que ele proporciona à enorme massa de dados gerados e disponibilizados pelo Departamento de Geodésia do IBGE (DEGED), os dados da RBMC, que geralmente, são usados apenas pelo próprio instituto e também por pesquisadores, com aplicações específicas em seus campos de estudo.

O fato do método proposto utilizar de forma continuada e sistemática, os dados da RBMC confere a esta aplicação, o caráter de democratizar e tornar mais conhecidos e utilizados esses dados. Além de justificar e de tornar imprescindível à implantação, adensamento e manutenção de redes de estações GPS ativas que utilizem conjuntos receptores de dupla frequência.

O Brasil não conta com uma densa rede de estações ativas de receptores de dupla frequência e código P, o que facilitaria diversas aplicações geodésicas, que normalmente requerem um refinamento maior em termos de precisão espacial do posicionamento. Até junho de 2003, a única rede de estações GPS com estas características é a RBMC, composta atualmente de apenas 15 estações e que não são suficientes para cobrir de maneira uniforme toda a extensão territorial brasileira. Isto fica mais evidente ao se observar os imensos vazios na região amazônica, onde atualmente a única estação em operação está localizada em Manaus.

Entretanto, existe um número significativo de estações dotadas de receptores de uma frequência, tal como aquelas que compõem a Rede INCRA de Base Comunitárias do GPS (RIBAC), composta atualmente por 44 estações. A possibilidade de que possam ser utilizadas como se fossem de dupla frequência e código P traria grandes benefícios em termos de precisão, principalmente para o processamento relativo de bases longas.

2

O ESTADO DA ARTE

Esta investigação é um híbrido, fruto da união entre duas áreas de conhecimento, um problema de GPS pela Engenharia - Mensuração e a busca da solução através de Redes Neurais Artificiais pela Ciência da Computação.

Ainda na década de 40 MCCULLOCH & PITTS (1943) desenvolveram o primeiro modelo artificial de um neurônio biológico, onde se concentram muito mais em descrever e apresentar as capacidades computacionais de um neurônio artificial do que apresentar técnicas de aprendizado. Apenas com o trabalho de HEBB (1949) estas técnicas de aprendizado tornaram-se objeto de estudo.

Posteriormente, o trabalho de WIDROW & HOFF (1960) apresentou uma regra de aprendizado baseada no gradiente descendente para minimização do erro da rede, denominada *regra delta*, que ainda hoje é utilizada. ROSENBLATT (1958), criou um novo modelo, o *perceptron*, com sinapses ajustáveis, que se prestaria ao treinamento de certos tipos de padrões.

As Redes Neurais Artificiais ficariam adormecidas durante toda a década de 70, em grande parte por conta do trabalho de MINSKY & PAPER (1969), que chamaram a atenção para algumas tarefas que o perceptron não era capaz de executar, já que só resolve problemas linearmente separáveis. O principal argumento de MINSKY e PAPER era que o problema do crescimento explosivo tanto do espaço ocupado, como do tempo requerido para a resolução de problemas complexos, afetariam as RNAs, inclusive os perceptrons.

HOPFIELD (1982) publicou um artigo que foi o responsável por parte da retomada das pesquisas na área, onde mostrou a relação entre redes recorrentes auto-associativas e sistemas físicos. A descrição do algoritmo de treinamento *backpropagation* por RUMELHART (1986) mostrou que a visão de MINSKY e PAPERT acerca do perceptron era bastante pessimista. Provando que as RNAs de múltiplas camadas são realmente capazes de resolver problemas complexos. O avanço acelerado da microeletrônica também contribuiu para que nas últimas duas décadas, pesquisadores se dedicassem ao estudo e aplicação de RNAs em diversas áreas do conhecimento, BRAGA; CARVALHO; LUDERMIR (2000).

Existem diversos trabalhos que abordam sob os mais variados aspectos tanto a tecnologia GPS quanto RNAs. Porém, na Mensuração a maioria dos trabalhos com aplicação em GPS está voltada à utilização prática que se dá desta tecnologia, principalmente na obtenção de coordenadas com variadas aplicações nos diferentes segmentos tais como: Geodésia, Topografia, Cadastro urbano e rural, Fotogrametria, Sensoriamento Remoto, Sistema de Informações Geográficas, Monitoramento de veículos, Monitoramento de deformações de estruturas e outros.

Há uma outra corrente de pesquisadores que se ocupa com estudo e modelagem dos erros inerentes ao processo de transmissão, propagação ou recepção do sinal GPS, sempre visando obter-se uma maior acuracidade para as coordenadas.

Com relação às redes neurais, sabe-se que o Brasil é um dos países que registra um dos maiores crescimentos mundiais no estudo da inteligência artificial ou inteligência computacional. Segundo levantamento do Conselho Nacional de Redes Neurais, praticamente todas as grandes universidades federais, estaduais e algumas particulares possuem estudos voltados para esta área. São comuns trabalhos que apresentam a utilização de RNAs principalmente em previsão de dados, algoritmos de aprendizado, otimização de sistemas, reconhecimento de padrões e outros.

Dentre os trabalhos, já desenvolvidos, relacionando RNAs aos diferentes segmentos da Mensuração pode-se citar: o processamento digital de imagens de

satélites (Sensoriamento Remoto) PANDEY et al. (2001), a modelagem numérica do terreno (Topografia), BACA (2001) e a predição de anomalia gravitacional (Geofísica), TIERRA; FREITAS (2000).

Freqüentemente, utiliza-se o sistema GPS apenas como instrumento de orientação e navegação. LORA; HEMERLY; LAGES (1998) desenvolveram um método onde a idéia é possibilitar um bom nível de automação em embarcações, que seriam guiadas por este sistema. Nesta aplicação, RNAs são utilizadas no treinamento adaptativo de máquinas e o GPS serve apenas para fornecer parte do padrão de entradas externas, ou seja, as coordenadas de pontos, para a rede neural.

O trabalho de HERNÁNDEZ-PAJARES et al. (1997) propõe e aplica a modelagem do *Total Electrons Content* (TEC) da camada ionosférica e busca quantificar seus efeitos na propagação do sinal GPS utilizando os modelos *Self Organizing Map* (SOM) e Hopfield.

E, mais recentemente, MAIA et al. (2002) apresentaram a possibilidade da predição da contagem de ciclos de uma portadora GPS utilizando o modelo Temporal – FIR MLP. O que indica que se é possível prever o comportamento temporal de uma observável, ainda que por um curto período de tempo, provavelmente também seja possível estimar o comportamento de uma observável em função de outra.

Apesar de uma minuciosa e exaustiva busca, não foi encontrado durante a fase de revisão bibliográfica nenhum trabalho que caminhasse na mesma linha de pesquisa proposta nesta investigação. Este fato impossibilitou a citação expressa de trabalhos anteriores. É desejável que em se tratando de uma tese de doutoramento se exijam como atributos básicos um caráter inédito, uma significativa relevância e a exequibilidade.

3

O SISTEMA DE POSICIONAMENTO GLOBAL NAVSTAR/GPS

Apesar de este ser um assunto cujos fundamentos são amplamente apresentados em literatura específica, existem algumas nuances que podem ser ressaltadas e o serão, como forma de oportunizar um bom entendimento, ainda que superficial devido à complexidade e extensão dos temas, aos profissionais quer sejam eles ligados à Mensuração ou à Ciência da Computação.

Quanto a este capítulo, a abordagem terá como foco principal seu princípio de funcionamento, a caracterização dos sinais e seu processamento. Detalhes pormenorizados tais como histórico e sistemas predecessores, segmentos nos quais está dividido, erros aos quais está sujeito, constelação, configuração orbital e prescrições do tempo foram propositalmente omitidos, em função de serem assuntos amplamente apresentados por diversos autores tais como: SEEBER (1993); LEICK (1995), PARKINSON (1996), HOFFMANN-WELLENHOF (2001), e outros.

Não serão explanadas nenhuma das outras características do sistema que não estejam relacionadas à caracterização, descrição e processamento do sinal recebido pelas antenas receptoras dos equipamentos.

3.1. O Sinal GPS

Sob a ótica da Engenharia, sinais são funções ou seqüências que servem para transportar informação de uma fonte de mensagens a um destinatário. As características específicas dos sinais dependem do canal de comunicação utilizado para este transporte. Este canal é definido pelo tipo de distorção que introduz os sinais. Esta distorção pode ser: determinística linear, determinística não linear ou aleatória, KOVÁCS, 1996.

Os sinais são processados no lado do transmissor com a finalidade de produzi-los e configurá-los, e no lado do receptor para extrair a informação neles contida, se possível com a máxima eficiência. (No caso do sinal GPS, a fonte geradora, são os satélites artificiais GPS, e o canal de comunicação, as portadoras, e o destinatário, o receptor GPS são elementos pré-definidos com características gerais muito bem especificadas).

A Geodésia por Satélites é baseada em dados transmitidos através de ondas eletromagnéticas, oriundas de osciladores a bordo dos satélites que geram uma frequência fundamental f_o . Duas portadoras da banda L são geradas pela multiplicação de constantes por f_o , resultando nas portadoras L1 e L2. Estas são moduladas por códigos para fornecer leituras dos relógios dos satélites ao receptor e também para transmitir informações tais como os parâmetros orbitais, HOFFMANN-WELLENHOF et al. (2001). Os componentes desses sinais encontram-se resumidos na Tabela 1.

Tabela 1 – Resumo de alguns componentes dos sinais dos satélites (Adaptada de HOFFMANN-WELLENHOF et al., 2001).

| Componentes | Frequência (MHz) | |
|------------------------|------------------|------------------------------|
| Frequência fundamental | f_o | = 10,23 |
| Portadora L1 | $154f_o$ | = 1575,42 (\cong 19,0 cm) |
| Portadora L2 | $120f_o$ | = 1227,60 (\cong 24,4 cm) |
| Código P | f_o | = 10,23 |
| Código C/A | $f_o/10$ | = 1,023 |

Os satélites GPS são reconhecidos através de diferentes critérios de nomeação e numeração. Um destes critérios é o número Pseudo-Random-Noise (PRN), ruído falsamente aleatório. Os códigos que formam o PRN são modulados, em fase, sobre as duas portadoras, L1 e L2. Esta modulação possibilita que se executem medições de distâncias a partir do tempo de propagação da modulação (Leick, 1995).

De forma bastante simplificada pode-se dizer que os códigos são uma seqüência com os estados $+1$ e -1 , correspondentes aos valores binários 0 e 1. A chamada modulação bifásica é executada com um efeito de 180° na fase da portadora sempre que ocorrer uma mudança de estado. Esta modulação está ilustrada na Figura 1.

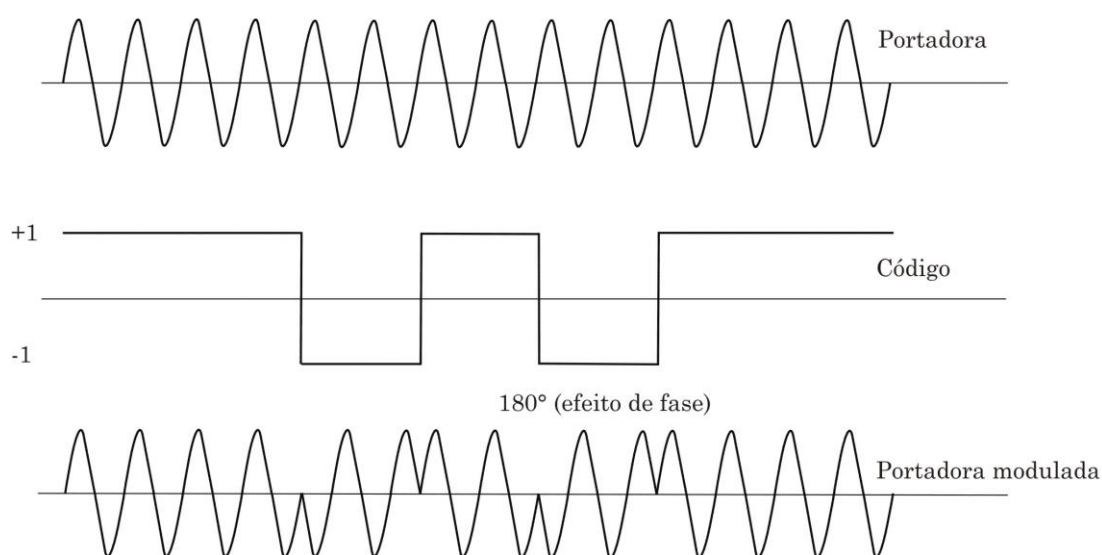


Figura 1 – Modulação Bifásica da onda portadora (Adaptada de HOFFMANN-WELLENHOF et al., 2001).

O principal motivo para o uso de uma segunda frequência nos satélites GPS é minimizar os efeitos da camada ionosférica na propagação do sinal. Assim, o uso de um conjunto receptor/antena de dupla frequência permite que se tenham prontamente disponíveis, as observáveis das duas frequências, cuja composição independe da ionosfera. Usuários que façam uso de receptores de uma frequência normalmente dependem da eliminação dos efeitos ionosféricos através da eliminação de erros que ocorre na dupla diferença e/ou da utilização de modelos ionosféricos.

Diversas razões, para as quais não cabe uma análise mais detalhada neste momento, concorrem para que no Brasil a maioria dos usuários do sistema GPS utilize receptores de uma frequência. Dentre elas pode-se destacar a significativa diferença de custo entre os conjuntos receptores de uma e dupla frequências.

Caso se utilize para observação conjuntos receptores de uma frequência, a estimativa do atraso ionosférico normalmente é realizada a partir da aplicação de diferentes modelos, geralmente alimentados por parâmetros transmitidos na mensagem de navegação. Uma descrição geral desses modelos pode ser encontrada em SILVA (2000). Testes e comparações entre esses modelos podem ser obtidos em LEANDRO (2003).

3.2 As observáveis GPS

O conceito relacionado às observáveis GPS refere-se a distâncias calculadas a partir de medições de fase ou tempo, baseadas na comparação entre sinais gerados internamente pelos receptores e recebidos dos satélites GPS. A forma de medir as distâncias a partir dos sinais recebidos difere dos distanciômetros eletrônicos terrestres pelo uso de dois relógios, um no satélite e outro no receptor. Este fato faz com que as distâncias sejam influenciadas pelos erros de ambos, por isso são chamadas de pseudodistâncias, HOFFMANN-WELLENHOF et al. (2001).

O princípio fundamental da navegação por GPS reside em medições das pseudodistâncias. Ao conhecerem-se as coordenadas dos satélites em um determinado sistema de referência é possível se calcular as coordenadas da antena de um dado receptor, neste mesmo sistema.

Para a determinação das coordenadas de pontos seria necessário observar apenas três satélites, estaria assim solucionado o problema de geometria. Porém devido ao não sincronismo dos relógios do satélite e do receptor, um quarto satélite é necessário, para solucionar este problema.

3.2.1. Pseudodistância

Pseudodistâncias calculadas a partir de leituras de código são função do tempo de viagem do sinal GPS entre as antenas dos satélites e de um dado receptor.

Elas são o resultado da multiplicação da velocidade da luz pelo tempo de propagação do sinal (τ_R^S) resultante do processo de correlação cruzada. Embora os relógios dos satélites possuam osciladores de alta precisão, seu sistema de tempo é diferente dos relógios dos receptores. Estes têm osciladores de menor precisão, mas pode-se relacionar esses dois sistemas de tempo, satélites(t^S) e receptor (t_R), ao sistema de tempo GPS (t_{GPS}), utilizando-se as seguintes expressões:

$$t_{GPS^S} = t^S - dt^S \quad (1)$$

$$t_{GPS_R} = t_R - dt_R \quad (2)$$

onde:

t_{GPS^S} = sistema de tempo GPS do satélite;

t_{GPS_R} = sistema de tempo GPS do receptor;

dt^S = erro do relógio do satélite em relação ao tempo GPS no instante t^S ;

dt_R = erro do relógio do receptor em relação ao tempo GPS no instante t_R .

A partir das equações (1) e (2) pode-se escrever a equação da pseudodistância da seguinte forma:

$$PD_{iR}^S = c\tau_R^S + c(dt_R - dt^S) + \varepsilon_{PDiR}^S \quad (3)$$

onde:

c = velocidade da luz no vácuo;

τ_R^S = tempo de propagação do sinal, contado desde sua geração no satélite até a correlação no receptor;

$\varepsilon_{PDi_R}^S$ = erro da medida de pseudodistância;

i = índice que representativo do tipo de código a partir do qual a pseudodistância foi obtida (C/A, sobre a portadora L1, e ou do código P, sobre as portadoras L1 e ou L2);

Existem diversos fatores que podem ser considerados no cálculo da pseudodistância, e o são por diferentes autores. Entretanto, um detalhe particularmente importante a ser observado refere-se ao tempo de propagação τ_R^S . Seu produto pela velocidade da luz não representa exatamente a distância geométrica ρ_R^S entre as antenas dos satélites e de um dado receptor. Isto ocorre devido aos erros provocados pela atmosfera terrestre (ionosfera = Ii_R^S e troposfera Ti_R^S) e também devido aos efeitos de multicaminhamento ($M_{PDi_R}^S$). MONICO (2000) descreve uma equação possível para a pseudodistância que considere esses fatores e pode ser representada por:

$$PDi_R^S = \rho_R^S + c(dt_R - dt^S) + Ii_R^S + Ti_R^S + M_{PDi_R}^S + \varepsilon_{PDi_R}^S \quad (4)$$

3.2.2 Fase da portadora

Comparativamente à pseudodistância, a fase da portadora ou simplesmente fase, é mais precisa, e por isso é considerada a observável a se utilizar em aplicações geodésicas (MONICO, 2000).

Chamando-se de $\varphi^S(t)$ a fase da portadora reconstruída e recebida com a frequência f^S e de $\varphi_R(t)$ a fase da portadora de referência, gerada internamente no receptor com uma frequência f_R . O parâmetro t é uma época no sistema de tempo GPS, reconhecida a partir de uma época inicial $t_0 = 0$. Assim, pode-se escrever as seguintes equações (HOFFMANN-WELLENHOF et al., 2001):

$$\begin{aligned} \varphi^S(t) &= f^S t - f^S \frac{\ell}{c} - \varphi_o^S \\ \varphi_R(t) &= f_R t - \varphi_{OR} \end{aligned} \quad (5)$$

As fases iniciais, φ_O^S e φ_{OR} , são causadas pelos erros dos relógios e são iguais a:

$$\begin{aligned}\varphi_O^S &= -f^S \delta^S \\ \varphi_{OR} &= -f_R \delta_R\end{aligned}\quad (6)$$

Conseqüentemente a fase $\varphi_R^S(t)$ é dada por:

$$\begin{aligned}\varphi_R^S(t) &= \varphi^S(t) - \varphi_{OR}(t) \\ \varphi_R^S(t) &= -f^S \frac{\ell}{c} + f^S \delta^S + f_R \delta_R + (f^S - f_R)t\end{aligned}\quad (7)$$

O desvio das freqüências f^S , f_R em relação à freqüência nominal f é apenas da ordem de fração de Hz. Isto pode ser constatado ao se considerar que a estabilidade na freqüência em um curto intervalo de tempo é de $\frac{df}{f} = 10^{-12}$, sendo $f \approx 1,5$ GHz, o erro da freqüência $df = 1,5 \cdot 10^{-3}$ Hz.

Assim, o erro da freqüência (aproximadamente $t = 0,07$ segundos) pode ser negligenciado durante a propagação do sinal, o erro máximo do ciclo da fase está abaixo do nível de ruído, por isso os erros dos relógios dos receptores são na ordem de milisegundos e são menos efetivos, podendo-se escrever a equação anterior de forma simplificada:

$$\varphi_R^S(t) = -f \frac{\ell}{c} - f \Delta \delta \quad (8)$$

Inicialmente, o número inteiro N de ciclos da portadora, chamado de ambigüidade entre as antenas dos satélites e de um dado receptor é desconhecido. Porém, se a observação é continuada sem perda de contato, ou seja, sem a chamada perda de ciclo, o número permanece o mesmo e a fase à época t é dada por:

$$\varphi_R^S(t) = \Delta \varphi_R^S \Big|_{t_0}^t + N \quad (9)$$

$\Delta\varphi_R^S|_{t_o}^t$ denota a fração de fase na época t adicionado ao número inteiro de ciclos desde a época inicial t_o .

Uma interpretação geométrica da equação (9) é representada pela Figura 2, onde $\Delta\varphi_i$ é a notação abreviada para $\Delta\varphi_R^S|_{t_o}^{t_i}$ e por simplificação, a fração do batimento de fase inicial, $\Delta\varphi_o$, é assumida como sendo zero. Fazendo a substituição da equação (9) na (8) e nomeando a quantidade observação negativa por $\Phi = -\Delta\varphi_R^S$, a equação da pseudodistância da fase resume-se em:

$$\Phi = \frac{1}{\lambda} \ell + \frac{c}{\lambda} \Delta\delta + N \quad (10)$$

Onde: λ = comprimento de onda.

Multiplicando-se a equação acima por λ , a fase é expressa em ciclos para uma distância que se diferencia da pseudodistância do código apenas por um inteiro multiplicado por λ . A distância ℓ representa a distância entre o satélite na época t da emissão e o receptor na época da recepção $t + \Delta t$. A fase da portadora pode ser medida melhor que 0,01 ciclos, que corresponde a 2 mm de resolução, HOFFMANN-WELLENHOF et al. (2001).

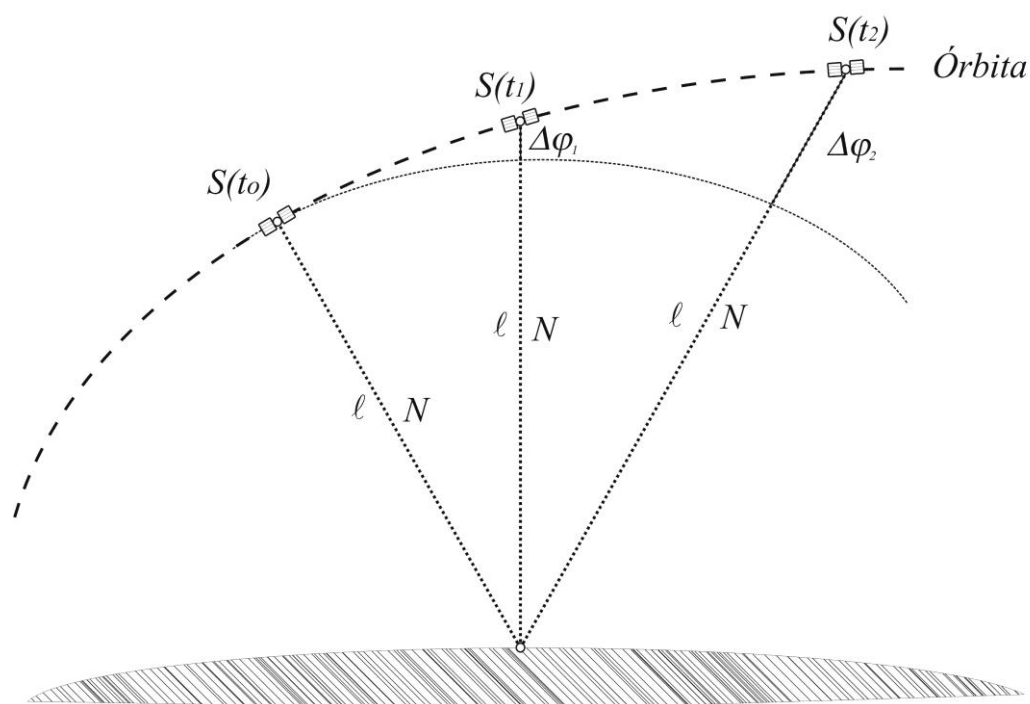


Figura 2 – Interpretação geométrica da fase da portadora. (Adaptada de HOFFMANN-WELLENHOF et al., 2001).

3.3 O Processamento de dados GPS

Não haveria sentido em descrever detalhadamente as formas clássicas de posicionamento e processamento de dados GPS, tais como posicionamento simples e relativo uma vez que estas se encontram bastante difundidas em toda a literatura pertinente. Maiores detalhes podem ser encontrados em SEEBER (1993); LEICK (1995), PARKINSON (1996), HOFFMANN-WELLENHOF (2001), e outros.

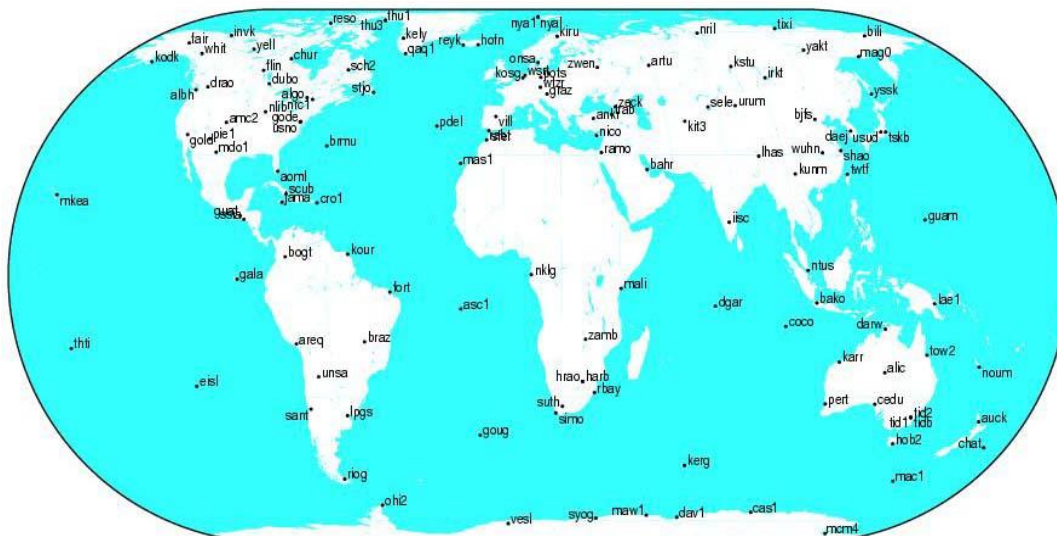
Ainda que esta pesquisa se utilize destas formas de processamento para efeito de análise e comparação de resultados, será descrito apenas, o posicionamento pontual preciso que, apesar de figurar entre as inovações no processamento de dados GPS nos últimos anos ainda é quase que completamente desconhecido tanto por pesquisadores quanto por usuários do sistema GPS no Brasil. Os dados relativos ao uso do Auto-Gipsy por usuários do Brasil, mostram apenas 222 consultas ao sistema desde 1998, <http://milhouse.jpl.nasa.gov/ag/usage.html>, 01/06/2003.

3.3.1 Posicionamento Pontual Preciso (PPP)

O PPP, do inglês *Precise Point Positioning*, é um método que utiliza as informações de fase das portadoras L1 e L2, e pseudodistâncias, a partir dos códigos C/A e P em conjunto com as efemérides precisas e correções de relógio, fornecidas pelo *International GPS Service (IGS)*, para processamento pontual. Este fato provoca uma limitação no seu uso, uma vez que apenas dados oriundos de receptores com estas características podem ser submetidos ao serviço automatizado de processamento. Além disso, os dados só podem ser processados entre uma e duas semanas após sua observação, tempo necessário para cálculo e liberação das efemérides precisas pelo *IGS*, SILVA; SEGANTINE (2001).

Este método possui uma forma de processamento de dados GPS que utiliza uma estratégia relativamente nova. E, como tal, possui vantagens e desvantagens quando comparado às formas tradicionais de processamento. Esta diferença decorre da metodologia utilizada na estimativa das coordenadas para uma determinada estação GPS. O aspecto mais incomum deste tipo de processamento é que o diferencial dos demais é que os erros dos relógios são modelados como coeficientes polinomiais, que são estimados como parâmetros estocásticos e não cancelados como ocorre na dupla diferença de fase, usada pela maioria dos métodos de processamento GPS (WEBB; ZUMBERGE, 1997).

A essência deste método é calcular órbitas e erros dos relógios de todos os satélites disponíveis, através de uma rede global ativa de alta precisão, neste caso a rede IGS, que se encontra ilustrada pela Figura 3, e usar todas estas informações para resolver os parâmetros de uma estação de qualquer local no mundo (posição, relógio e troposfera úmida).



**Figura 3 – Localização das estações que compõe a rede do IGS.
(Fonte: <http://igsceb.jpl.nasa.gov/network/allmaps.html>, 10/03/03).**

É uma ferramenta que pode ser usada para processar desde uma única estação até uma rede com centenas de pontos. A precisão das coordenadas calculadas individualmente no PPP é comparável a soluções de redes, ou seja, variam de poucos milímetros a poucos centímetros. Entretanto, ele é incapaz de identificar correlações entre estações, GREGORIUS (1996).

O *JPL's Flinn Analysis* é o setor, dentro do JPL, responsável pela determinação e fornecimento de órbitas e correções precisas de relógios com uma repetibilidade diária de poucos milímetros na componente horizontal e cerca de um centímetro na vertical das coordenadas dos satélites. A precisão obtida está em um fator entre 100 a 1000 vezes melhor que a possível pela mensagem de navegação, possibilitando pós-processamento de dados de um receptor simples com precisão de poucos centímetros quando usado em conjunto com efemérides precisas (ZUMBERGE, 1998).

O PPP usa para processamento os programas GIPSY e OASIS II, ambos desenvolvidos pelo *Jet propulsion Laboratory (JPL)*. Eles utilizam módulos comuns apesar de serem dois pacotes independentes. Uma das facilidades proporcionada pelo JPL é a análise automática e gratuita de dados GPS, que é um serviço disponível a um usuário qualquer através do uso de e-mail e ftp. Este serviço é designado por *Auto-Gipsy*, descrito a seguir.

3.3.1.1 O Auto-GIPSY (AG)

O JPL possibilita o processamento automatizado de dados GPS com o sistema AG. Qualquer usuário, de qualquer lugar, que utilize e-mail e ftp² e que possua dados, adequados às exigências do PPP pode obter as coordenadas da estação observada, além de diversas outras informações estatísticas, em cerca de 5 a 10 minutos após o envio dos dados. Estas informações podem ser descritas como:

- Tempo de observação \geq uma hora;
- Receptor de dupla frequência e código P;
- Taxa de coleta de dados \leq 30 segundos (taxas menores que 30 devem ser submúltiplos de 30 s);
- Método de posicionamento estático;
- Formato RINEX do arquivo com nomenclatura internacional.

A análise automática funciona regularmente desde 1998 e possibilita a análise de dados GPS coletados por um receptor numa localização arbitrária. Porém, a principal vantagem para o usuário quando se utiliza do serviço automatizado é que não depende tempo de aprendizado, instalação e manutenção de um complexo programa de análise GPS. O procedimento para utilização do AG envolve:

- Disponibilizar o arquivo RINEX de observação (apenas este), nomeado adequadamente, para “*download*” em uma área de ftp anonymous (opcionalmente comprimido);
- Enviar um e-mail para ag@cobra.jpl.nasa.gov. No local destinado ao assunto do e-mail, deve-se escrever “Static”. O corpo do e-mail deve conter a URL³ (endereço de ftp) onde está seu arquivo RINEX a ser processado.

Apenas dados oriundos de satélites com elevação acima de 15° são considerados. O sistema detecta automaticamente a maioria das condições de “anormalidade” nos dados. Caso ocorra problema quanto a informações e ou

² File Transfer Protocol

³ Uniform Resource Locator

formato dos dados, o usuário receberá um e-mail como resposta avisando destes erros. O tempo de resposta, geralmente, é cerca de 10 minutos. Após esse período, o usuário receberá um e-mail com o endereço de ftp onde se encontram oito arquivos para serem copiados com as seguintes informações:

1. Explicações detalhadas de todos os resultados (*readme*);
2. Estimativas de coordenadas (ϕ , λ , h , em WGS 84, com as devidas correções para o ITRF⁴ 2000) do centro de fase da antena receptora, seus desvios e a correlação entre si (*.gd) ;
3. Componentes cartesianas referentes a um determinado ITRF (*.stacvx);
4. Componentes cartesianas sem um ITRF definido (*.stacov);
5. Resíduos posteriores do ajustamento (*.prf);
6. Erros formais em parâmetros estimados tais como: atraso zenital na troposfera úmida, desvio de fase, erro do relógio do receptor e coordenadas da estação (*.tdf);
7. Parâmetros utilizados pelo GIPSY no processamento (*.rgnml);
8. *logfile* (resumo dos dados utilizados).

Uma grande vantagem deste serviço é a análise eficaz de dados observados com receptores GPS de diferentes fabricantes sem a necessidade de formar dados diferenciados com receptores de referência. É alcançada uma precisão diária de alguns poucos milímetros na dimensão horizontal e aproximadamente 1cm na dimensão vertical para receptores estacionários, <http://milhouse.jpl.nasa.gov/ag/agfaq.html>, 15/04/2003.

Uma importante limitação do PPP é o fato de que o receptor utilizado nas seções de observação necessariamente tem que ser capaz de receber os sinais L1, L2, P1 ou C1 e P2. Se o receptor for de dupla frequência, mas não corresponder a este padrão, o AG não processa os dados, SILVA (2000).

É mais uma forma de análise e pós-processamento de dados GPS, que possibilita a um usuário do sistema obter coordenadas com uma precisão compatível com trabalhos topográficos e com trabalhos geodésicos de

⁴ International Terrestrial Reference Frame, referente a correção periódica do deslocamento dos pólos terrestres.

planimetria, de forma simples, rápida e com baixo custo de execução e uso do sistema. A Figura 4 apresenta um resumo do processamento pelo AG.

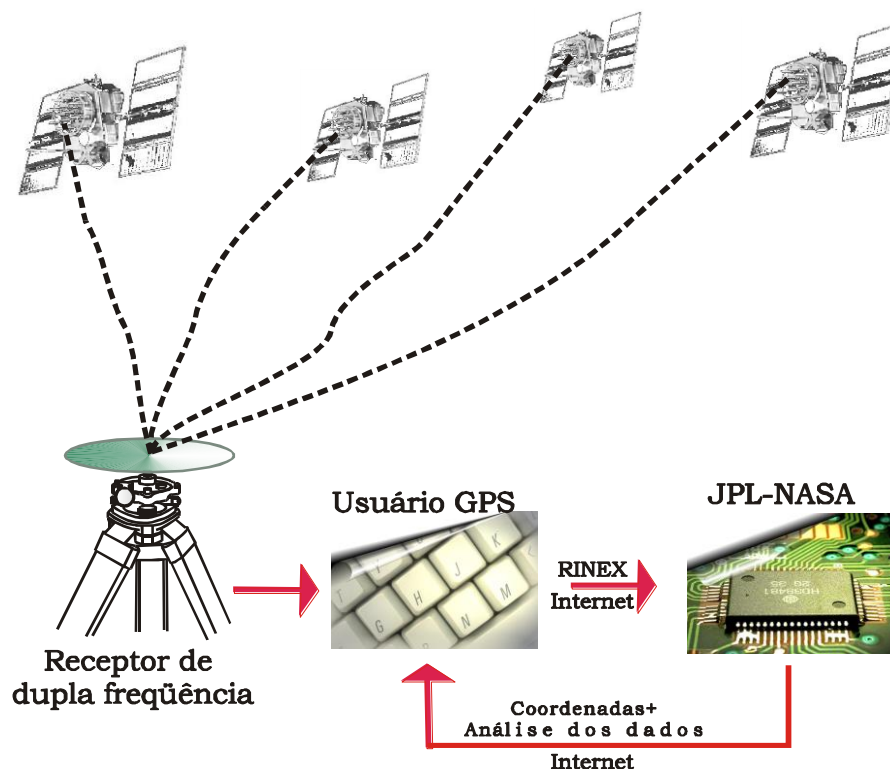


Figura 4 – Representação resumida do processamento pelo AG.

3.4 O formato RINEX

Os dados GPS, frutos da observação e coleta realizados por um receptor, normalmente são armazenados em um formato binário, próprio e específico, que é reconhecido e processado apenas por programas oferecidos pelo fabricante. Entretanto, para resolver o problema de processamento de observações GPS feitas com receptores de diferentes fabricações, definiu-se um formato padrão que independe do tipo e modelo do receptor com o qual ele foi obtido. A este formato deu-se o nome RINEX, do inglês *Receiver Independent Exchange Format*. O RINEX em sua versão atual consiste de seis diferentes arquivos de acordo com o padrão ASCII:

1. Dados de observação;
2. Dados de navegação;
3. Dados meteorológicos;
4. Mensagem de navegação GLONASS;
5. Mensagem de navegação GEO;
6. Dados de relógio de receptor e satélite.

Cada arquivo consiste de uma seção de cabeçalho e uma seção de dados. O cabeçalho contém informações gerais para todo o arquivo e é colocado no início do arquivo. A seção de dados traz informações específicas para cada época observada. Uma descrição detalhada do formato com espaçamentos, tipo de dados, construção e exemplo do formato pode ser obtida em GURTNER (2002).

4

REDES NEURAIS ARTIFICIAIS (RNAs)

As RNAs não são um assunto propriamente novo, pois o primeiro modelo artificial de um neurônio biológico data de 1943, resultado do trabalho do neuro-anatomista e psiquiatra Warren MaCulloch e do matemático Walter Pitts. Entretanto, merecerão ênfase naquilo que possibilite o conhecimento básico de seus fundamentos bem como a descrição dos paradigmas de aprendizagem, evidenciando-se o modelo MLP (*Multi Layer Perceptron*⁵) e o algoritmo *backpropagation*, responsável pelo ressurgimento das RNAs no final da década 80.

Muitos autores conceituam e descrevem Redes Neurais Artificiais, sempre enfatizando as características que mais evidenciam a utilização que se quer dar a elas. Várias dessas definições podem ser encontradas em HERTZ; KROGH; PALMER (1991), HAYKIN (1999); BRAGA; CARVALHO; LUDERMIR (2000) e outros.

Optou-se, neste trabalho, por uma abordagem que faz uma analogia entre RNAs e estatística, numa tentativa de tornar mais claros seus conceitos fundamentais e adequá-los a um jargão mais familiar à engenharia. Porém, deve-se enfatizar que o estudo das RNAs é um importante campo dentro da Inteligência Artificial, área da Ciência da Computação.

SARLE (1994) descreve redes neurais como sendo uma ampla classe de modelos discriminantes e regressão não linear flexível, modelos de redução de dados e sistemas dinâmicos não lineares. Com frequência, compostos por

⁵ Modelo artificial de um neurônio biológico

diversos “neurônios”, isto é, elementos de cálculo lineares ou não, interconectados através de ligações, por vezes complexas e freqüentemente organizados em camadas. RNAs podem ser usadas basicamente de três formas:

1. Como modelos que simulam o sistema nervoso biológico e “inteligência”;
2. Como processador adaptativo de sinais em tempo real ou controladores implementados em hardware para aplicações específicas, tal como em robôs;
3. Como métodos analíticos de dados.

Este trabalho limitar-se-á apenas à estimativa e análise de dados.

4.1 *Inspiração Biológica*

A inteligência potencial, alegada às redes neurais, levou diversos pesquisadores à implementação de redes neurais em hardwares. Contudo, RNAs muito raramente têm mais que centenas ou poucos milhares de neurônios, enquanto que o cérebro humano tem cerca de 100 bilhões de neurônios biológicos, ilustrados pela Figura 5.

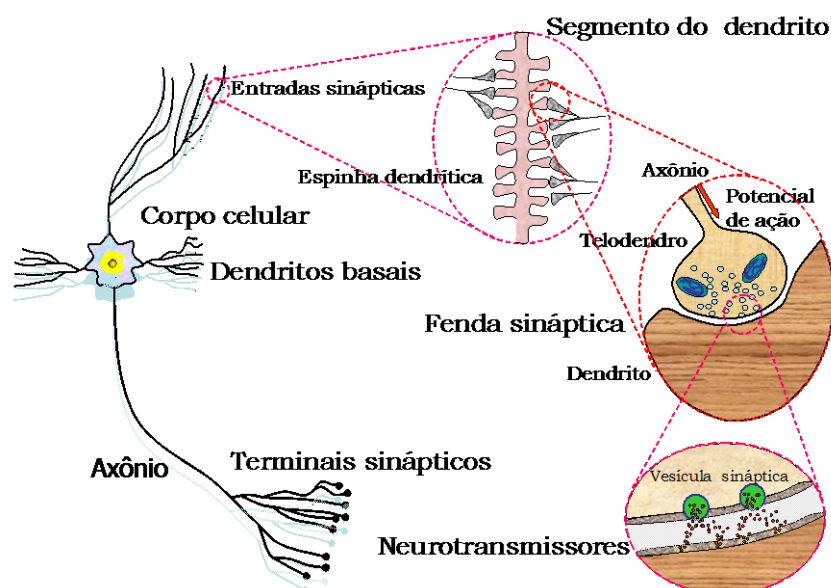


Figura 5 – Representação genérica de um neurônio biológico e processo de transmissão sináptica. (Adaptada de HAYKIN, 1999).

Apesar de existirem computadores com funcionamento paralelo, RNAs comparáveis à complexidade do cérebro humano ainda estão muito além de nossa capacidade tecnológica.

Milhares de fibras nervosas, chamadas dendritos, são conectadas ao corpo celular ou soma do neurônio. Funcionam como conexões de entrada, cada uma com menos de um micron de diâmetro. O corpo celular é capaz de executar operações complexas sobre os sinais de entrada, contudo, estas operações podem ser aproximadas por um somatório simples destes, HAYKIN (1999).

A extensão do soma é uma única e longa fibra nervosa chamada axônio. Esta estrutura recebe e analisa o valor do somatório exterior. Se este valor for superior a um determinado limite mínimo, o axônio produz uma voltagem de saída de 10^{-3} V, também denominada potencial de ação, que é conduzida à sua extremidade.

A extremidade do axônio se ramifica em diversos terminais sinápticos ou telodendros. Estes terminais se conectam aos dendritos ou soma das unidades vizinhas através de complexas estruturas chamadas sinapses. Segundo KOVÁCS (1997), não ocorre um contato direto entre telodendro e dendrito, mas sim uma conexão química temporária (processo de transmissão sináptica), o que provoca a excitação, inibição ou modulação da unidade seguinte.

No telodendro há vesículas com neurotransmissores que são responsáveis pela condução do potencial de ação entre os neurônios. Quando o potencial de ação atinge a membrana do telodendro, abrem-se canais provocados pela fusão entre a membrana do mesmo e as membranas das vesículas. Esta fusão permite a saída dos neurotransmissores, que se difundem pela fenda sináptica e se ligam aos receptores da membrana do dendrito. Esta ligação desestabiliza a membrana, permite a entrada de íons e possibilita a condução do potencial de ação para outro neurônio, HAYKIN (1999).

4.2 Neurônio Artificial

Redes Neurais são sistemas de processamento de informações compostos por um grande número de elementos de processamento simples chamados de neurônios artificiais, ou simplesmente neurônios. Muitos modelos de RNAs têm sido propostos e extensivamente usados para processamento de sinais.

O modelo inicial para um neurônio artificial, foi proposto por McCulloch e Pitts e posteriormente aprimorado por ROSENBLATT (1958), que o nomeou de *perceptron* (Figura 6). Ele representa uma aproximação extremamente simplificada sob a ótica neurofisiológica, porém conserva algumas das características básicas do neurônio biológico e viabiliza a sua implementação computacional.

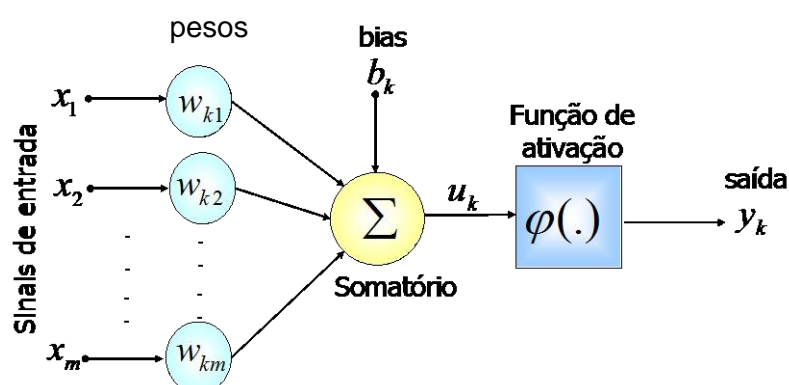


Figura 6 – Modelo não linear de neurônio artificial (Adaptada de HAYKIN, 1999).

Onde x_m são as entradas da rede, w_{km} são os pesos, ou pesos sinápticos, associados a cada entrada, b_k é o termo bias, u_k é a combinação linear dos sinais de entrada, $\varphi(\cdot)$ é a função de ativação e y_k é a saída do neurônio.

Pode-se dizer que é nos pesos que reside todo o conhecimento adquirido pela rede. Os pesos é que são os parâmetros ajustáveis e que mudam e se adaptam à medida que o conjunto de treinamento é apresentado à rede. Assim sendo, o processo de aprendizado supervisionado em uma RNA com pesos, resulta em sucessivos ajustes dos pesos sinápticos de tal forma que a saída da rede seja o mais próximo possível da resposta desejada.

Tipicamente, a ordem de amplitude normalizada da saída do neurônio está no intervalo $[0, 1]$ ou alternativamente $[-1, 1]$. O modelo neural também inclui um termo chamado de bias, aplicado externamente, simbolizado por b_k . O bias b_k tem o efeito do acréscimo ou decréscimo da função de ativação na entrada da rede, dependendo se é positiva ou negativa, respectivamente.

Regiões de decisão são intervalos fixos onde a resposta desejada pode estar. O bias serve para aumentar os graus de liberdade, permitindo uma melhor adaptação, por parte da rede neural, ao conhecimento fornecido a ela.

4.3 Topologia ótima.

A priori, quanto mais camadas de neurônios, melhor seria o desempenho da rede neural pois aumenta a capacidade de aprendizado, melhorando a precisão com que ela delimita regiões de decisão. Mas, na prática, aumentar o número de camadas intermediárias, às vezes, se torna inviável devido ao trabalho computacional.

Existem pesquisadores como HECHT; NIELSEN (1989), que afirmam que com apenas uma camada intermediária na rede neural, já é possível calcular uma função arbitrária qualquer a partir de dados fornecidos. De acordo com estes autores, a camada oculta deve ter por volta de $(2i+1)$ neurônios, onde i é o número de variáveis de entrada.

Outros, como CYBENKO (1988) estudaram o número de camadas intermediárias necessárias à implementação de funções em RNAs. Seus resultados indicam que uma camada intermediária é suficiente para aproximar qualquer função contínua e duas camadas intermediárias aproximam qualquer função matemática.

LIPPMANN (1987) afirma que em havendo uma segunda camada intermediária na rede neural, esta deve ter o dobro de neurônios da camada de saída. No caso de apenas uma camada oculta, ela deverá ter $s(i+1)$ neurônios, onde s é o número de neurônios de saída e i o número de neurônios na entrada.

Outros autores defendem que em redes pequenas o número de neurônios da camada oculta pode ser a média geométrica entre o número de neurônios de entrada pelo número de neurônios de saída.

4.4 Características e propriedades úteis das RNAs

RNAs são capazes de processar uma grande quantidade de dados e fazer previsões, algumas vezes com uma acurácia surpreendente e embora essa característica não as torne “inteligentes” no sentido usual da palavra, redes Neurais Artificiais definitivamente “aprendem”.

O algoritmo de aprendizado é formado por um conjunto de procedimentos bem definidos, capazes de adaptar os parâmetros de uma RNA. Existe um conjunto de ferramentas representadas por diversos algoritmos que diferem entre si através da forma como se dá o ajuste de pesos, BRAGA; CARVALHO; LUDERMIR (2000).

O poder de generalização, atribuído às redes neurais, refere-se à capacidade que elas possuem de produzir saídas razoáveis a entradas que nunca lhe tenham sido apresentadas (aprendizado). Essa característica possibilita a solução de problemas complexos.

A utilização de redes neurais possibilita as seguintes capacidades e propriedades úteis:

- Não linearidade (é uma propriedade importante, particularmente se o sinal de entrada é gerado por mecanismos físicos, inerentemente não lineares);
- Traçado entrada-saída (aprendizado supervisionado);
- Adaptatividade (capacidade de adaptar os pesos sinápticos, modificando-os em função de entradas externas);
- Resposta clara (classificação, interpolação e extrapolação);
- Informação contextual (cada neurônio é potencialmente afetado pela atividade global de todos os outros neurônios da rede);
- Tolerância a erros;

- Analogia neuro-biológica (resolução de problemas complexos).

4.5 Os Modelos Neurais.

Há dezenas de diferentes modelos de RNAs descritos na literatura. À primeira vista, o processo de seleção pode parecer uma tarefa difícil, frente ao número de possibilidades. Porém, é provável que uns poucos modelos forneçam uma solução excelente (MASTERS, 1994).

Muitos modelos neurais são similares ou idênticos a técnicas estatísticas populares tais como: modelos lineares generalizados, regressão polinomial, regressão não paramétrica e análise de discriminantes, análise de grupos e componentes principais, especialmente onde a ênfase é a predição de fenômenos complexos ao invés de uma explicitação das correlações entre as variáveis. Estes modelos neurais podem ser muito úteis em diferentes tipos de aplicação.

Apesar de existirem diversos métodos para o treinamento de redes neurais, eles podem ser agrupados em dois paradigmas principais que são os aprendizados: supervisionado e não supervisionado. Existem ainda os aprendizados por reforço e por competição, mas o método de aprendizado supervisionado é o mais usado em RNAs, quer de neurônios com pesos, quer de neurônio sem pesos, e possui este nome em função de que tanto entradas quanto saídas desejadas são fornecidas à rede durante o processo de treinamento, BRAGA; CARVALHO; LUDERMIR (2000). Esta investigação utilizará apenas o método de aprendizado supervisionado.

SARLE (1994) afirma que embora muitos modelos de redes neurais sejam similares ou idênticos a modelos estatísticos bastante conhecidos, a terminologia na literatura de RNAs é muito diferente que aquela usada por estatísticos. Por exemplo, na literatura de RNAs:

- Variáveis são chamadas de características;
- Variáveis independentes são chamadas de entradas;
- Valores estimados ou preditos são chamados de saídas;
- Variáveis dependentes são chamadas de valores de treinamento;

- Resíduos são chamados de erros;
- Estimar ou prever são chamados de treinamento, aprendizado, adaptação ou auto-organização;
- Critério de estimação é chamado de função erro, função custo, ou função Lyapunov;
- Observações são chamadas de padrões de treinamento;
- Parâmetros estimados são chamados de pesos (sinápticos);
- Transformações são chamadas de links funcionais;
- Regressão ou análise de discriminantes são chamadas de aprendizado supervisionado ou heteroassociação;
- Análise de classes é chamada de aprendizado competitivo ou Adaptive Vector Quantization;
- Interpolação e/ou extrapolação são chamadas de generalização.

Os termos estatísticos, amostra e população não parecem ter equivalentes no jargão de RNAs. Embora os dados sejam sempre divididos em conjunto de treinamento e teste, para que se efetive a validação cruzada.

4.5.1 O Modelo Perceptron Multi Camadas (Multilayer Perceptrons ou MLP).

Um *perceptron*, MCCULLOCH & PITTS (1943), calcula a combinação linear de entradas chamadas de entradas da rede. Então, uma função de ativação possivelmente não linear é aplicada às entradas da rede para produzir uma saída. Um *perceptron* pode ter uma ou mais saídas. Se possuir uma função de ativação linear, é similar a um modelo de regressão linear.

Segundo SARLE (1994) ao introduzir-se na rede uma camada extra de neurônios, chamada de intermediária ou camada oculta, ainda assim, existe apenas uma única camada com pesos a serem estimados. Se o modelo inclui pesos, estimados entre a camada de entrada e a camada oculta, e a camada oculta usa função de ativação não linear, tal como a função logística, o modelo torna-se genuinamente não linear. O modelo resultante é chamado de *Multilayer Perceptron* ou MLP.

As redes de múltiplas camadas são aproximadores universais, RUMELHART et al. (1986) e devem ser usadas quando se tem pouco conhecimento e dados insuficientes acerca do relacionamento entre as variáveis independentes e dependentes. Pode-se variar a complexidade do modelo MLP através da variação do número de camadas intermediárias e do número de neurônios em cada camada intermediária.

Com um pequeno número de neurônios ocultos, uma MLP é um modelo paramétrico que fornece uma alternativa útil à regressão polinomial. Com um número médio de neurônios ocultos, a MLP pode ser considerada como um modelo quase paramétrico. Se o número de neurônios intermediários aumenta com o tamanho da amostra, a MLP torna-se um modelo não paramétrico.

Considerando uma MLP para ajustamento de uma curva de regressão não linear simples, usando uma entrada, uma saída linear e uma camada intermediária com uma função de ativação logística, a curva pode ter muitas ondulações e como existem neurônios ocultos, este tipo de MLP atuaria como uma regressão polinomial ou um ajustamento quadrático.

Diferentes de polinômios, MPLs são fáceis de expandir para múltiplas entradas e múltiplas saídas sem um aumento exponencial do número de parâmetros. MPLs são treinadas normalmente por um algoritmo chamado de regra delta generalizada, sugerida inicialmente por WIDROW; HOFF (1960), que calcula derivadas pela simples aplicação da regra da cadeia chamada *backpropagation*.

Freqüentemente, o termo *backpropagation* é aplicado também ao próprio método de treinamento ou à rede treinada, desta forma. Esta confusão na literatura em redes neurais é geral quando se trata de diferenciar modelos de métodos ou algoritmos de treinamento.

Embora uma MLP com uma camada intermediária seja um aproximador universal, existem diversas aplicações onde o uso de mais de uma camada intermediária pode ser útil. Eventualmente uma função altamente não linear pode ser aproximada com poucos pesos quando se usam múltiplas camadas intermediárias ao invés de apenas uma.

Em estatística necessita-se ter uma informação à priori sobre a estrutura do modelo matemático. Em redes neurais o usuário estima esta estrutura pela escolha do número de camadas e o número de neurônios e as funções de transferência entre os neurônios de uma rede.

4.5.2 O algoritmo *backpropagation*

Para treinar a rede são utilizados vetores de entrada associados a seus respectivos vetores de saída desejada, até que a rede aproxime uma determinada função e possa a partir daí oferecer saídas adequadas a vetores de entrada diferentes daqueles com os quais foi treinada.

O *backpropagation* padrão é um algoritmo gradiente descendente tal como a regra da cadeia. Através dele os pesos das conexões entre os neurônios da rede, são atualizados ao longo de um gradiente descendente de uma determinada função. O termo *backpropagation* refere-se a forma como o gradiente é calculado para redes de múltiplas camadas não lineares.

Existe um número razoável de variações a partir do algoritmo básico, que são baseadas em técnicas de otimização, tais como o gradiente conjugado e métodos de Newton.

O algoritmo *backpropagation* básico tem duas fases, para cada padrão apresentado:

- Feed-forward → as entradas se propagam pela rede, da camada de entrada até a camada de saída;
- Feed-backward → os erros se propagam na direção contrária ao fluxo de dados, indo da camada de saída até a primeira camada escondida.

O *backpropagation* é o algoritmo de Redes Neurais mais utilizado em aplicações práticas de previsão, classificação e reconhecimento de padrões em geral. A Tabela 2 apresenta sucintamente este algoritmo. Maiores detalhes acerca das diferentes funções de ativação, regra de atualização de pesos e uma descrição matemática, podem ser encontrados em HAYKIN (1989). Foi

demonstrado por CYBENKO, (1989) que o *backpropagation* é um aproximador universal capaz de representar qualquer função.

Tabela 2- Descrição resumida do Algoritmo de aprendizado *backpropagation* padrão.

| Inicialização |
|---|
| ↪ pesos iniciados com valores aleatórios e pequenos ($w \leq 0.1$); |
| Treinamento |
| ↪ Loop até que o erro de cada neurônio de saída seja \leq tolerância, para todos os padrões do conjunto de treinamento; |
| 1. Aplica-se um padrão de entrada o de entrada X_i com o respectivo vetor de saída da Y_i desejado; |
| 2. Calculam-se as saídas dos neurônios, começando da primeira camada escondida até a camada de saída; |
| 3. Calcula-se o erro para cada neurônio da camada de saída. Se $\text{erro} \leq \text{tolerância}$, para todos os neurônios, volta ao passo 1; |
| 4. Atualiza-se os pesos de cada neurônio, começando pela camada de saída, até a primeira camada escondida. |
| 5. Volta-se ao passo 1 caso a diferença entre a saída da rede e a resposta desejada seja maior que uma determinada limite, especificado pelo usuário. |

Um dos fatores que influenciam o aprendizado é a forma como ele se dá. O algoritmo básico pode ser implementado tanto para o aprendizado incremental quanto para o aprendizado acumulativo.

No modo de aprendizado incremental os pesos são atualizados a cada apresentação de um novo padrão, ele utiliza o método do gradiente descendente mas tende a aprender melhor o último padrão apresentado.

No modo de aprendizado acumulativo, também chamado de modo *batch*, os pesos são ajustados apenas depois da apresentação de todos os padrões, ou seja, depois de um ciclo completo na apresentação dos padrões. Muitas das variações do algoritmo básico utilizam esta forma de aprendizado como tentativa de acelerar o processo de treinamento da rede.

4.5.2.1 Algoritmo de treinamento rápido

Para um treinamento mais rápido podem-se utilizar diversos artifícios, entre eles pode-se destacar a taxa de aprendizado variável e os algoritmos de gradiente conjugado. Existem ainda como alternativa os chamados de *métodos de Newton* para treinamento acelerado.

A diferença fundamental utilizada nos *métodos de Newton* em comparação ao algoritmo padrão está na matriz Hessiana (segunda derivada) do índice de erro aos valores atuais de pesos e *bias*. Os *métodos de Newton* sempre convergem mais rápido que os métodos de gradiente conjugado, mas, infelizmente, o cálculo da derivada segunda é complexo e demorado para determinar a matriz Hessiana. Assim, utiliza-se uma segunda classe de algoritmos que, apesar de baseados no método de Newton, não necessitam do cálculo da derivada segunda, são chamados de *métodos Quase-Newton* ou ainda de métodos secantes.

Os *métodos Quase-Newton* têm sido publicados com sucesso nos estudos de FLETCHER (1963), GOLDFARB (1970) E SHANNO (1970). Os algoritmos *backpropagation* que utilizam essas técnicas são denominados de “BFGS” (as iniciais dos pesquisadores).

Semelhante aos algoritmos Quase-Newton, o algoritmo Levenberg-Marquardt foi desenvolvido para aproximar velocidade de treinamento de segunda ordem sem ter que calcular a matriz Hessiana. Quando a função erro faz a somatória quadrática, a matriz Hessiana pode ser aproximada como:

$$H = J^T J \quad (11)$$

E o gradiente pode ser calculado como:

$$g = J^T e \quad (12)$$

Onde J é a matriz Jacobiana que contém a primeira derivada dos erros da rede com os respectivos pesos e *bias* e e é o vetor de erros da rede. A matriz Jacobiana pode ser calculada através da técnica *backpropagation* padrão.

O algoritmo Levenberg-Marquardt usa esta aproximação *Quase-Newton* para a matriz Hessiana com a seguinte atualização:

$$x_{k+a} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (13)$$

Quando o módulo de μ é zero, este é o *método de Newton* e usa a aproximação da matriz Hessiana. Quando μ é grande, isto faz com que o gradiente descendente tenha um pequeno incremento a cada passo. Os algoritmos que usam os *métodos de Newton* são mais rápidos e mais precisos quanto menor for o erro mínimo desejado no treinamento.

Assim, μ é diminuído depois de cada passo bem sucedido (redução do erro) e só é aumentado quando o erro aumentar. Deste modo, o erro sempre diminuirá a cada iteração do algoritmo. Este algoritmo aparece na bibliografia como sendo o método mais rápido para o treinamento de redes neurais de tamanho moderado. Além disso, encontra-se implementado de forma bastante eficiente como uma ferramenta do programa MATLAB sob o nome *trainlm* (DEMUTH; BEALE, 1997).

5

MATERIAIS E MÉTODOS

Esta pesquisa, em relação ao método, acatou as idéias de John Dewey, posteriormente difundidas por NORTHROP apud CASTRO (1977) que indicam que se inicie com uma pergunta, a partir da qual se estruturam as explicações, hipóteses e a natureza dos dados que possam respondê-las. Assim sendo, as perguntas que motivaram esta investigação já foram feitas no Capítulo 1 e são repetidas a seguir:

- **Seria possível fazer com que dados observados por receptores de uma frequência (L1) possam ser processados de forma a se obter um nível de precisão similar àqueles gerados por receptores de dupla frequência (L1 e L2)?**
- **Seria possível modelar, estimando o comportamento das observáveis GPS?**
- **Qual ferramenta seria adequada a esta modelagem?**

5.1 Materiais.

Os dados utilizados em todos os procedimentos de investigação desta pesquisa referem-se a:

- Arquivos de dados obtidos nos receptores da RBMC, no formato RINEX, exemplificado na Figura 7, contendo as observáveis L1, C1, L2, P2 e D1;

- ▶ Arquivos de dados GPS no formato RINEX gerados por receptores de frequência simples, que observaram em estações cujas coordenadas já se conhecia de observações anteriores;
- ▶ Arquivos de efemérides precisas disponibilizados pelo IGS⁶;
- ▶ Informações sobre atividade solar disponibilizadas pelo NOAA⁷.

```

BOMJ3391 - Bloco de notas
Arquivo Editar Formatar Exibir Ajuda
2 OBSERVATION DATA RINEX VERSION / TYPE
DAT2RIN 2.35b IBGE/DEGED 06DEC01 2:10:23 GTMPGM / RUN BY / DATE
Agência "HOST" IBGE/BRASIL OBSERVER / AGENCY
16682 TRIMBLE 4000SSI Nav 7.29 sig 3.07 REC # / TYPE / VERS
00001004 DORNE MARGOLIN T ANT # / TYPE
BOMJ - RBMC Bom Jesus da Lapa MARKER NAME
93030 MARKER NUMBER
4510197.1625 -4268322.6928 -1453035.5360 APPROX POSITION XYZ
0.0080 0.0000 0.0000 ANTENNA: DELTA H/E/N
*** Above antenna height is from mark to BOTTOM OF ANTENNA. COMMENT
1 0 WAVELENGTH FACT L1/2
5 L1 C1 L2 P2 D1 # / TYPES OF OBSERV
15 INTERVAL
2001 12 5 0 1 15.000000 TIME OF FIRST OBS
2001 12 6 0 0 0.000000 TIME OF LAST OBS
28 # OF SATELLITES
1 2081 2087 1758 1758 2081 PRN / # OF OBS
2 1631 1634 1076 1076 1631 PRN / # OF OBS
3 1351 1351 1346 1346 1351 PRN / # OF OBS
4 1962 1963 1849 1849 1962 PRN / # OF OBS
5 1649 1651 1613 1613 1649 PRN / # OF OBS
6 1981 1981 1940 1940 1981 PRN / # OF OBS
7 1436 1436 1389 1389 1436 PRN / # OF OBS
8 2087 2091 1836 1836 2087 PRN / # OF OBS
9 1379 1379 1375 1375 1379 PRN / # OF OBS
10 2046 2056 1857 1857 2046 PRN / # OF OBS
11 1382 1385 1286 1286 1382 PRN / # OF OBS
13 1949 1958 1664 1664 1949 PRN / # OF OBS
14 1333 1333 1329 1329 1333 PRN / # OF OBS
15 2073 2073 2047 2047 2073 PRN / # OF OBS
17 2010 2010 1980 1980 2010 PRN / # OF OBS
18 1393 1394 1385 1385 1393 PRN / # OF OBS
20 1533 1534 1350 1350 1533 PRN / # OF OBS
21 1746 1746 1726 1726 1746 PRN / # OF OBS
22 2312 2314 2240 2240 2312 PRN / # OF OBS
23 1313 1313 1282 1282 1313 PRN / # OF OBS
24 1333 1333 1320 1320 1333 PRN / # OF OBS
25 1261 1261 1251 1251 1261 PRN / # OF OBS
26 1803 1813 1622 1622 1803 PRN / # OF OBS
27 2018 2038 1403 1403 2018 PRN / # OF OBS
28 1740 1742 1383 1383 1740 PRN / # OF OBS
29 1495 1495 1490 1490 1495 PRN / # OF OBS
30 1450 1450 1445 1445 1450 PRN / # OF OBS
31 1764 1849 1403 1403 1764 PRN / # OF OBS
END OF HEADER
1 12 5 0 1 15.0000000 0 9 1 2 4 8 13 20 27 28 31
11848599.89810 21980926.07800 3136924.70850 21980941.89140 -1473.35900
-14200220.38610 21170990.81300 4760040.49510 23833393.96900 2015147.16350 23833403.68840 -2115.96900
-2265236.90010 22062074.61700 969928.84050 22062085.18040 -1065.73400
-911.52810 21661844.34400 -113.87500
11216708.46610 22309212.22700 12909489.46150 22309219.69540 -4201.50000
-3566146.90110 21344980.25800 1940020.59750 21344996.44540 -1968.67200
-7080160.82510 22668111.47700 -2053.53100
1680100.64310 23906768.43000 592571.50750 23906783.85540 -735.48400

```

Figura 7 – Fac-símile de trecho de um arquivo RINEX de observação da estação BOMJ-RBMC.

Para o desenvolvimento do método proposto, buscou-se utilizar exclusivamente dados e informações disponíveis gratuitamente na Internet, numa tentativa de que o método possa ser utilizado posteriormente sem a necessidade da compra de dados além da facilidade e rapidez de acesso a eles. Adicionalmente, esta característica dos dados permite que o experimento possa ser reproduzido, a qualquer tempo, a partir de conjuntos de dados diferentes daqueles utilizados neste experimento.

⁶ International GPS Service.

⁷ National Oceanic & Atmospheric Administration.

A aproximação de funções que representem um determinado fenômeno físico, por vezes sob comportamento não linear, através de RNAs nas quais se utilizem modelos com aprendizado supervisionado, tal como proposto nesta investigação, necessitam de um “professor” externo. Ou seja, para que a rede neural aprenda é necessário ensiná-la e isso é feito através da separação do conjunto de dados disponíveis em três conjuntos:

1. Treinamento;
2. Validação;
3. Testes.

Como a própria nomenclatura indica, os conjuntos de treinamento e validação são usados durante o processo de treinamento e o conjunto de testes é usado já para a obtenção das estimativas. Esta é uma técnica estatística chamada de *validação cruzada*, HAYKIN (1999).

A formação de padrões de mapeamento entrada/saída que possibilitou a preparação dos arquivos, para ensinar a rede neural, incluiu além de dados de observação, a inserção de dados de posição precisa dos satélites GPS. Estes dados podem ser obtidos através da Internet como arquivos de efemérides precisas através, por exemplo, do endereço: <ftp://igscb.jpl.nasa.gov/igscb/product>, 02/12/2002. Um exemplo destes arquivos é apresentado na Figura 8.

```

#nP2001 12 5 0 0 0.00000000 96 ORBIT IGS00 HLM IGS
## 1143 259200.00000000 900.00000000 52248 0.00000000000000
+ 28 1 2 3 4 5 6 7 8 9 10 11 13 14 15 17 18 20
+ 21 22 23 24 25 26 27 28 29 30 31 0 0 0 0 0 0
+ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
+ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
++ 4 4 4 4 4 4 4 3 4 4 4 4 4 4 4 4 5
++ 4 4 5 4 4 4 4 4 4 4 4 0 0 0 0 0 0
++ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
++ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
++ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
%c cc cc ccc ccc cccc cccc cccc cccc ccccc ccccc ccccc ccccc
%c cc cc ccc ccc cccc cccc cccc cccc ccccc ccccc ccccc ccccc
%f 0.00000000 0.000000000 0.00000000000 0.0000000000000000
%f 0.00000000 0.000000000 0.00000000000 0.0000000000000000
%i 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
%i 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
/* FINAL ORBIT COMBINATION FROM WEIGHTED AVERAGE OF:
/* cod emr esa gfz jpl ngs sio
/* REFERENCED TO GPS CLOCK AND TO WEIGHTED MEAN POLE:
/* CLK ANT Z-OFFSET (M): II/IIA 1.023; IIR 0.000
* 2001 12 5 0 0 0.00000000
P 1 23965.686604 -1270.807363 -11261.114782 206.242045
P 2 13298.106136 -22178.016214 5505.613407 -116.099129
P 3 20643.406257 12662.855946 -11084.414456 89.123191
P 4 -5324.197618 -25836.006139 1541.795720 393.095600
P 5 -20821.267649 7506.886076 14521.114932 326.343873
P 6 -18705.627781 7945.312031 -17027.714211 -1.521649
n 7 2057 206186 16308 501152 20026 530640 624 400000
Para obter ajuda, pressione F1

```

Figura 8 – Fac-símile de trecho de um arquivo de efemérides precisas, formato SP3, correspondente ao dia 339 do ano 2001.

As efemérides precisas são disponibilizadas como arquivos ASCII, em formatos específicos. O formato adotado nesta investigação é o SP3. De forma bastante resumida, pode-se dizer que o arquivo de efemérides traz consigo informações da posição mais provável dos satélites GPS em intervalos regulares de tempo. São apresentadas as coordenadas cartesianas X^S, Y^S, Z^S e o erro do relógio dos satélites. Maiores detalhes sobre este formato podem ser obtidos no seguinte endereço eletrônico: http://www.ngs.noaa.gov/GPS/SP3_format.html, 02/12/2002.

Inicialmente, os dados retirados ou calculados a partir dos arquivos de observação das estações que compõem a rede GPS para uma determinada época foram: o ângulo de elevação de cada satélite, os valores correspondentes à contagem de ciclos das portadoras L1 e L2, as pseudodistâncias a partir dos códigos C/A e P2 e posição inicial aproximada da estação (X_R, Y_R, Z_R). Do arquivo

de efemérides precisas pode-se retirar as coordenadas (X^S, Y^S, Z^S) para o mesmo satélite, à mesma época.

Na verdade esses dados serão utilizados em função das características dos receptores utilizados pela rede de estações GPS usada para o treinamento da rede neural. Por isso, por exemplo, se utilizará a pseudodistância P2, mas nada impede que se utilize a pseudodistância P1, por exemplo, caso seja esse o padrão dos receptores para treinamento.

A Rede Brasileira de Monitoramento Contínuo (RBMC) é uma rede ativa composta atualmente de 15 estações, sendo quatorze geridas pelo IBGE⁸ e uma pelo INPE⁹. Possui características necessárias e desejáveis à realização do experimento, tais como:

- receptores de dupla frequência e código P
- simultaneidade de operação em diversas estações
- disponibilização gratuita dos dados via rede WWW¹⁰
- Posição da estação bem conhecida.

As estações de Fortaleza e Brasília além de comporem a RBMC, também fazem parte da rede IGS. A localização de todas as estações da RBMC está ilustrada na Figura 9.

⁸ Instituto Brasileiro de Geografia e Estatística

⁹ Instituto Nacional de Pesquisas Espaciais

¹⁰ World Wide Web

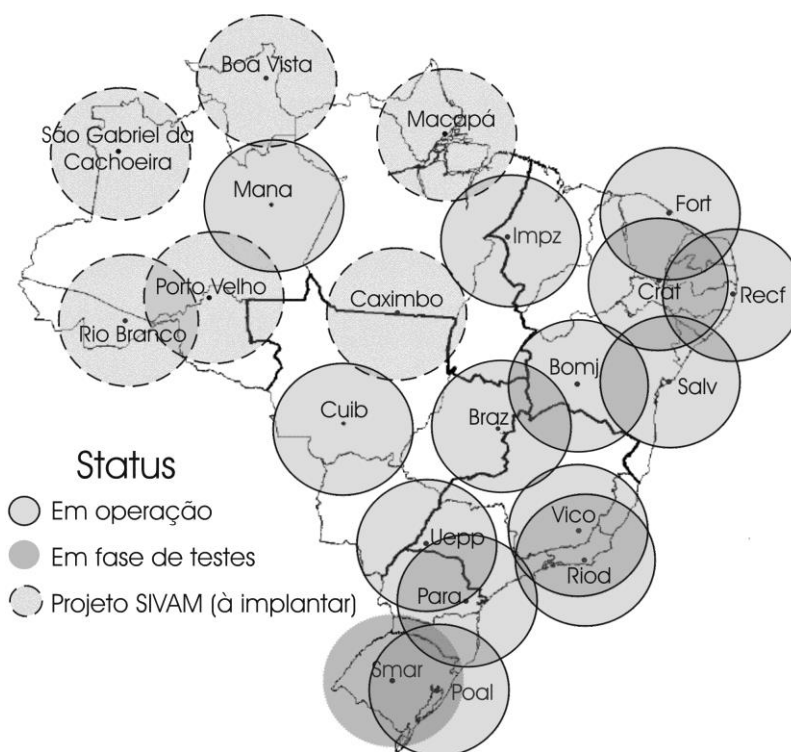


Figura 9 – Localização das estações RBMC. (Adaptada de <http://www.ibge.net/home/geografia/geodesico/rbmc.shtm>, 06/04/2003).

Existe a expectativa da entrada em operação, prevista para 2003, de seis estações adicionais referentes ao Sistema Integrado de Vigilância da Amazônia, designado por projeto SIVAM. Estas recobririam os vazios de observação GPS que hoje ainda persistem na Amazônia Legal e complementariam a RBMC, perfazendo um total de 21 estações.

Há de se notar que, em se tratando de Amazônia, bem como parte do nordeste brasileiro, tem de se ter cuidados especiais na utilização de dados GPS. O equador geomagnético que passa pela Amazônia brasileira, é uma região cuja camada ionosférica freqüentemente está sujeita à ação dos chamados jatos de elétrons. Isso pode gerar uma brusca mudança no TEC e este por sua vez pode ocasionar uma intensa variação em seu índice de refração. Este fato pode provocar diversos efeitos no sinal GPS que vão desde o atraso até a interrupção de sua propagação. Maiores detalhes sobre os efeitos ionosfera no sinal GPS podem ser encontrados em CAMARGO (1999) e FONSECA JUNIOR (2002).

Assim, quando da entrada em operação das estações SIVAM, necessariamente, terão que se realizar testes no sentido de avaliar a funcionalidade do método proposto sob condições adversas de atividade solar geomagnética.

A Tabela 3 contém informações relativas ao estabelecimento da RBMC desde o início de sua implantação e do gradual acréscimo do número de estações entre 1993 e 2002.

Tabela 3 – Informações das estações RBMC em maio 2003.

| Localização/UF | Código IBGE | Início de operação |
|------------------------|--------------------|---------------------------|
| Fortaleza (IGS)/CE | FORT | 13/05/93 |
| Brasília/DF | BRAZ | 03/03/95 |
| Curitiba/PR | PARA | 13/12/96 |
| Presidente Prudente/SP | UEPP | 18/12/96 |
| Bom Jesus da Lapa/BA | BOMJ | 18/02/97 |
| Manaus/AM | MANA | 28/04/97 |
| Viçosa/MG | VICO | 22/05/97 |
| Cuiabá/MT | CUIB | 18/06/97 |
| Imperatriz/MA | IMPZ | 16/02/98 |
| Porto Alegre/RS | POAL | 28/10/98 |
| Salvador/BA | SALV | 20/05/99 |
| Recife/PE | RECF | 06/07/99 |
| Crato/CE | CRAT | 09/04/00 |
| Rio de Janeiro/RJ | RIOD | 21/07/01 |
| Santa Maria/RS | SMAR | 08/11/01 |

As exigências que limitam o uso do método proposto dependem diretamente da disponibilidade de dados simultâneos àquele que se pretende estimar, ou seja, dados de observação dentro de uma determinada região, neste caso, a área de abrangência da RBMC e efemérides precisas para o dia no qual foram feitas as observações. Apesar de vários institutos de pesquisa disponibilizarem as efemérides precisas, aqui se utilizaram aquelas fornecidas pelo IGS, devido ao seu caráter oficial e também à metodologia utilizada em sua determinação.

Embora neste trabalho tenham sido usados dados da RBMC, nada impede que dados de outras redes, implantadas em outros países ou continentes, sejam utilizados para obtenção de estimativas usando o método proposto,

obviamente, desde que possuam características similares àquelas apresentadas pela RBMC e apresentadas anteriormente nesse item.

Como a RBMC foi escolhida como base dos dados, com os quais efetivamente se constrói o modelo, buscou-se fazer um levantamento estatístico dos dados gerados pelo IBGE com a finalidade de ratificar o seu uso para esta aplicação, além de conseguir-se uma visão mais ampla acerca da frequência e qualidade de seus dados.

De posse dos dados referentes a todas as estações da RBMC desde a sua implantação até junho de 2002, pode-se fazer uma seleção no sentido de agrupar os dados e obter um indicativo sobre quantos e quais dados eleger para proceder as estimativas neurais.

Este trabalho gerou informações acerca do período mais adequado a utilizar, quais as estações e dias nos quais havia um maior número de estações operando simultaneamente. Esses resultados são apresentados em forma de gráficos na Figura 10.

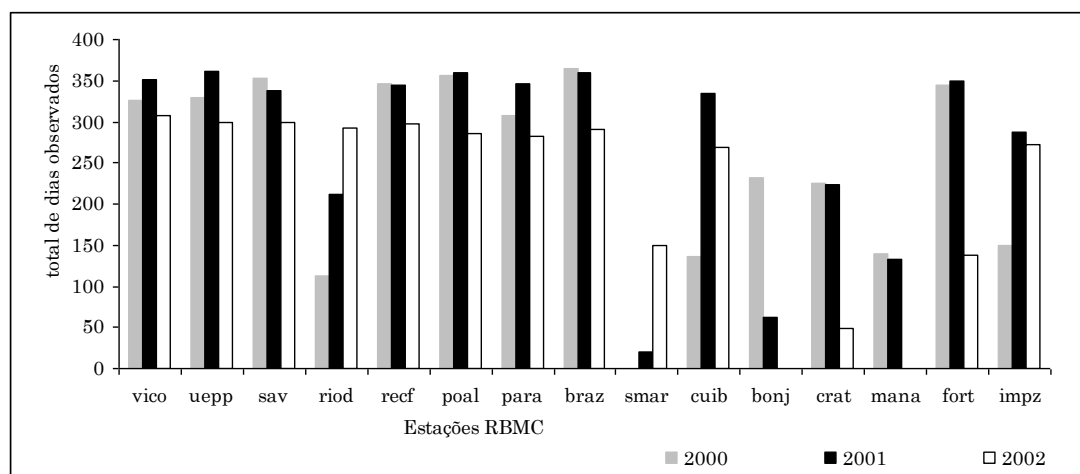


Figura 10 – Frequência de observações/estação/ano.

Uma das necessidades do método proposto é a utilização da maior quantidade disponível de dados, neste caso, dados da RBMC. Este é o motivo pelo qual não serão utilizados dados anteriores ao segundo semestre de 1999, função de só estarem em operação no máximo 11 estações. Assim, utilizou-se para efeito de modelagem e testes, os dados compreendidos entre o ano 2000 e

2002, que foram gentilmente cedidos pelo DEGED¹¹ do IBGE ao Departamento de Transportes da EESC-USP com a estrita finalidade de pesquisa.

Uma visão geral da simultaneidade de observações das estações RBMC para o período escolhido, ou seja, os anos de 2000, 2001 e 2002 são ilustrados nas Figuras 11, 12 e 13 respectivamente. Para o ano de 2002, só se obteve dados até o dia 307 (03/11/2002).

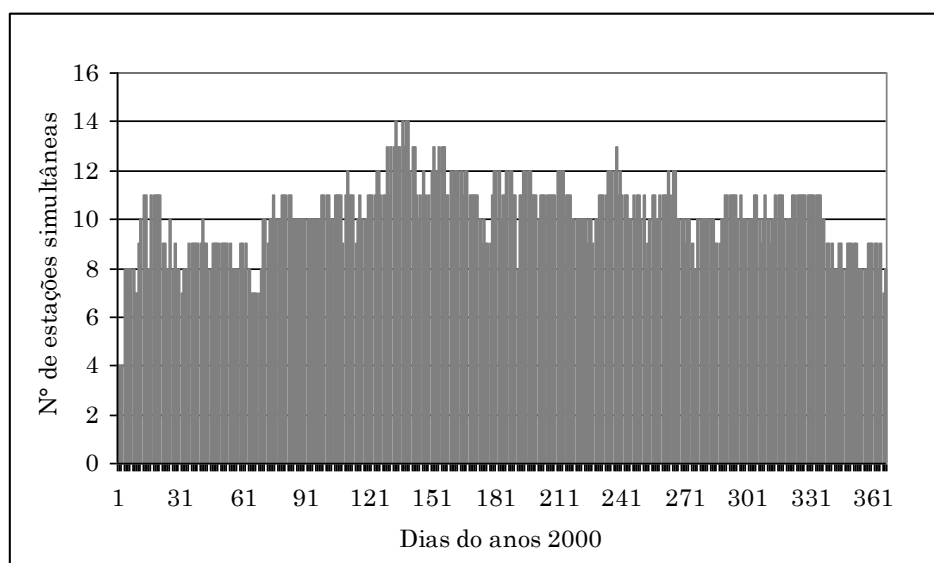


Figura 11 – Nº de estações operando simultaneamente em 2000

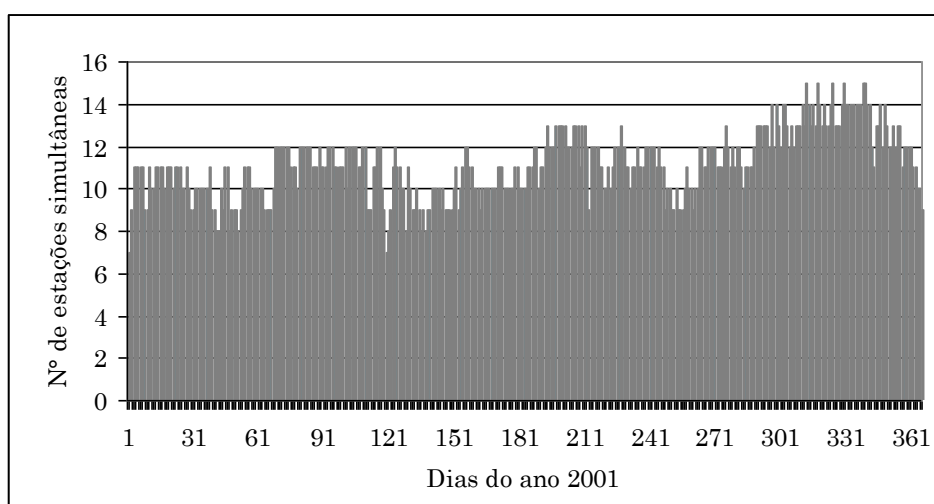


Figura 12 – Nº de estações RBMC operando simultaneamente em 2001.

¹¹ Departamento de Geodésia

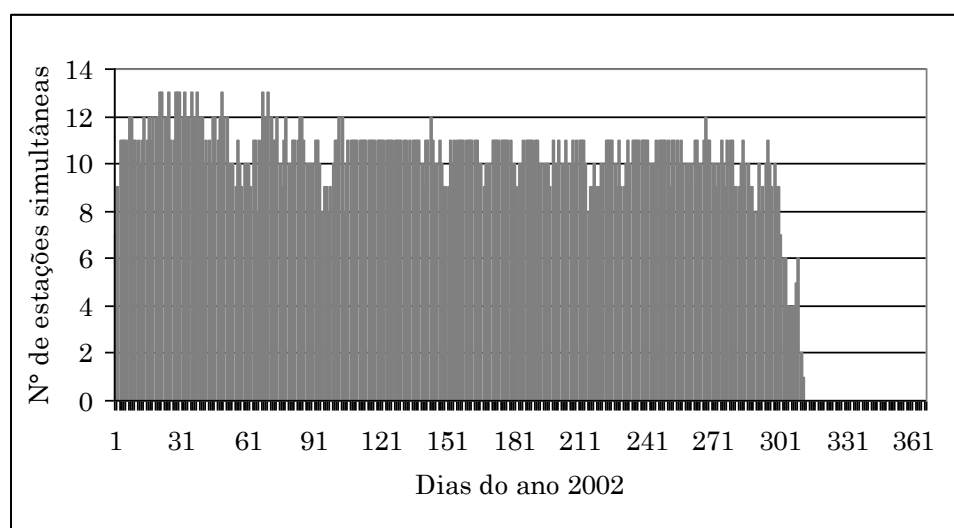


Figura 13 – Nº de estações RBMC operando simultaneamente em 2002.

5.1.1 Critérios adotados para escolha dos dados e desenvolvimento do método proposto

Como critério inicial, buscou-se determinar entre o período de 2000 a 2002, os intervalos com o maior número de estações que coletaram dados simultaneamente. Baseado na estatística de operação simultânea das estações RBMC, evidenciadas pelas Figuras 11, 12 e 13, adotou-se alguns dias como dados para a modelagem.

É importante ressaltar que a estação Fortaleza possui características operacionais que diferem do restante da rede. A principal diferença está na taxa de coleta dos dados. Em todas as estações da RBMC a taxa de coleta é de 15 segundos exceto para Fortaleza que possui taxa de coleta de 30 segundos. A taxa de 15 segundos é de particular interesse para o desenvolvimento do método proposto, pois quanto menor a taxa de coleta, maior será a quantidade de dados (padrões de entrada) disponíveis para treinamento da rede para pequenos intervalos de tempo, o que certamente concorrerá para uma melhor generalização da rede.

FAZAN (2002), submeteu os dados gerados pela RBMC a uma série de testes com o objetivo de verificar a sua qualidade e constatou que a estação de Fortaleza apesar de possuir uma frequência diária dentro do mesmo padrão das

outras estações, possui um aproveitamento das observações relativamente menor que as demais.

Estas duas características da estação Fortaleza fizeram com que as observações geradas por ela fossem ignoradas durante o processo de modelagem, o que não chega a prejudicar a aplicação do método, pois existe uma outra estação que cobre a mesma área, a estação CRAT, localizada em Crato-Ce.

A Tabela 4 resume esses dados, já considerando a exclusão da estação de Fortaleza.

Tabela 4 – Resumo dos dados escolhidos para modelagem.

| Data | Dia GPS | Semana GPS | Nº de estações | Atividade Solar (Emissão de raios-X) |
|-------------|----------------|-------------------|-----------------------|---|
| 12/05/2000 | 133 | 1061 | 13 | Baixa com eventos de alta |
| 17/05/2000 | 138 | 1062 | 13 | Média com eventos de alta |
| 18/05/2000 | 139 | 1062 | 13 | Moderada com eventos de alta |
| 13/11/2001 | 317 | 1140 | 14 | Baixa com eventos de alta |
| 25/11/2001 | 329 | 1142 | 14 | Baixa |
| 05/12/2001 | 339 | 1143 | 14 | Baixa |
| 29/01/2002 | 29 | 1151 | 12 | Moderada |
| 07/02/2002 | 38 | 1152 | 12 | Moderada |
| 08/03/2002 | 67 | 1156 | 12 | Moderada com eventos de alta |

Além disso, buscou-se distribuir os dados em diferentes meses dos anos numa tentativa de validar o método proposto independente do período no qual os dados foram observados, o que comprovará a eficácia da modelagem proposta sob diferentes períodos de observação, conjuntos de estações e condições de atividade solar.

De acordo com os critérios adotados por LEANDRO (2003), o índice de emissão de raios-X pode ser utilizado para estimar a atividade solar classificando-a, por exemplo, em alta, moderada ou baixa. Entretanto, essa é uma classificação relativa, pois a emissão de raios-X é apenas uma dentre várias grandezas que podem ser levadas em consideração, como a emissão de partículas, fluxo de ondas de rádio e número de manchas solares representado pelo número *Sunspot*.

Em janeiro de 2003 o sol passou por um período de alta intensidade que teve seu ápice entre 2000 e 2001. Assim, mesmo os períodos considerados de baixa intensidade são de alta quando comparados ao período próximo a 1997, considerado um período de calma como pode ser visto na Figura 14 para o atual ciclo solar (ciclo 23).

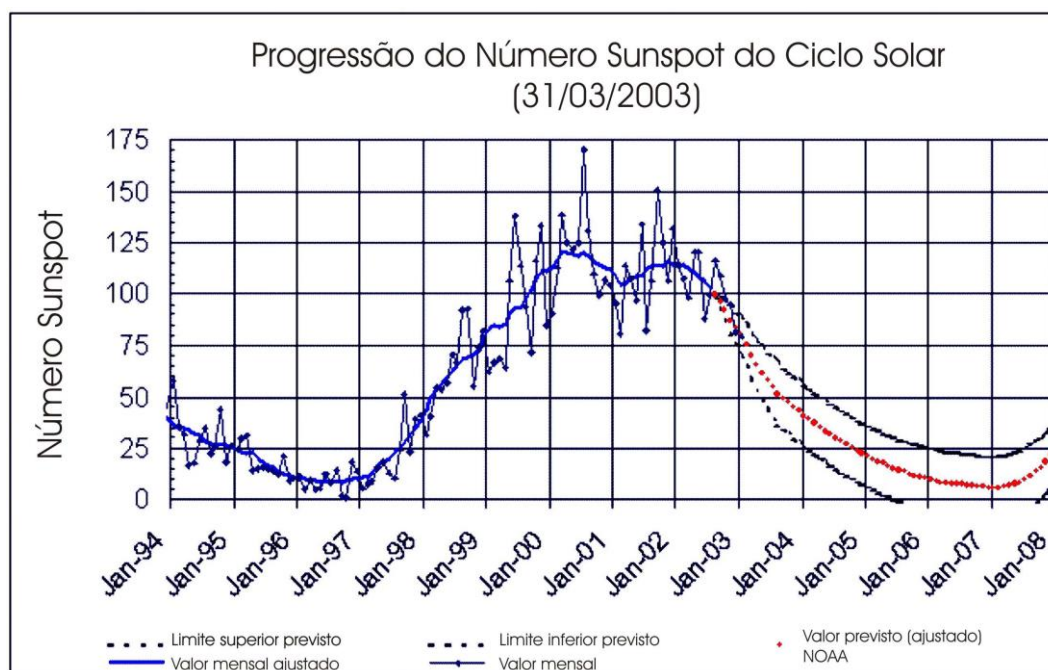


Figura 14 – Representação gráfica do número *Sunspot* do Ciclo Solar 23. (Fonte: <http://www.sec.noaa.gov/SolarCycle/index.html>, 28/01/2003.)

As emissões solares de raios-X podem ser avaliadas a partir de dados e gráficos, obtidos em relatórios de índices geofísicos e solares, medidos através dos satélites GOES-X e disponibilizados na Internet pelo *National Oceanic and Atmospheric Administration* (NOAA) sob o endereço: <http://www.sec.noaa.gov/weekly/>, 28/01/2003.

5.2 Metodologia

De posse de todas as informações e dados descritos no item 5.1, pôde-se delinear a seqüência de etapas, regras de procedimentos e estratégias de processamento adotadas no curso da investigação.

Devido a algumas restrições relativas a observação simultânea das estações RBMC, o conjunto de dados escolhido está exatamente em um período no qual ocorrem os máximos da atividade solar no atual ciclo. Este fato possibilitará, caso a hipótese inicial seja validada, que se comprove que o método proposto funciona, mesmo sob intensa atividade solar, que provoca variações bruscas no TEC e este por sua vez, na propagação do sinal GPS.

5.2.1 Etapas realizadas nesta investigação

1. Revisão Bibliográfica;
2. Busca e definição das estações GPS com receptores L1/L2 e código P que possam ser utilizadas sistematicamente para treinamento da rede neural;
3. Busca e ou implementação do algoritmo neural que melhor se adaptou à extrapolação e interpolação para estimativas de dados GPS;
4. Criação e montagem do padrão de entradas mais representativo com dados obtidos de tal forma a se obter, como respostas, um conjunto de L2 e ou P2 por satélite/época;
5. Desenvolvimento de aplicativos que possibilitem:
 - a. Ler e extrair de um conjunto de arquivos de dados GPS, para um determinado conjunto de épocas específicas (240 épocas por hora a cada 15 segundos);
 - b. Ler e extrair de um arquivo de efemérides precisas através de interpolação os dados de posição para cada a cada 15 segundos;
 - c. A partir das informações obtidas no arquivo GPS e também de efemérides precisas, construir um arquivo com os padrões de entradas e saídas para treinamento da rede (aprendizado supervisionado) adaptado a diferentes formatos de saída;
 - d. Garantir aleatoriedade quando da apresentação dos padrões à rede e se necessário, transformar e normalizar os dados;
6. Implementação e testes com o algoritmo escolhido;

7. Implementação de aplicativo para agregação ao arquivo GPS original, dos dados de L2 e P2 estimados pela RNA, gerando um arquivo RINEX modificado por estimativas neurais;
8. Validação e testes do método proposto:
 - a) Observação com receptor GPS de dupla frequência e código P em pontos que possibilitem o processamento de vetores com diferentes comprimentos de linhas base (curta, média e longa);
 - b) Aplicação do método a esta observação, com a geração e agregação de estimativas neurais;
 - c) Processamento por ponto simples do arquivo original;
 - d) Processamento por ponto simples do arquivo modificado;
 - e) Processamento relativo do arquivo original;
 - f) Processamento relativo do arquivo modificado;
 - g) Processamento PPP do arquivo original;
 - h) Processamento PPP do arquivo modificado;
 - i) Comparação e análise dos resultados obtidos nos diversos processamentos para as estações teste.

5.2.2 Procedimentos adotados

Sabidamente a busca de procedimentos de experimentação exequíveis e adequados que conduzam mais rapidamente à validação ou refutação das hipóteses levantadas é uma constante em todo o método de investigação, porém, nem sempre é possível que isso ocorra.

Tentou-se nesta pesquisa seguir uma seqüência lógica, buscando-se vencer, sem que fosse queimada nenhuma das etapas previstas inicialmente. Na verdade, houve uma série muito maior de tentativas que aquelas explicitadas até aqui. Porém, objetivando deixar a versão final mais resumida, só foram documentadas aquelas que se mostraram relevantes no decorrer da investigação.

5.2.2.1 Formação dos conjuntos de dados para treinamento, validação e testes do modelo neural.

Inicialmente criou-se, a partir dos dados disponíveis e escolhidos, diversos subconjuntos de dados que pudessem representar diferentes combinações de mapeamento entradas/saídas, períodos de observação em horas e agrupamentos espaciais das estações RBMC.

Faz parte das suposições iniciais que os algoritmos de redes neurais que utilizam o conceito de aprendizado supervisionado, conseguissem gerar estimativas adequadas a esta modelagem. Assim, decidiu-se iniciar com o modelo *backpropagation*, o mais difundido entre os modelos neurais conhecidos e bastante utilizado devido a sua excelente capacidade de generalização.

O modelo *backpropagation* aplicado a uma rede de múltiplas camadas conhecida como MPL necessita, além da definição do mapeamento entradas/saídas, da definição do número de camadas intermediárias, do número de neurônios em cada camada e dos parâmetros de treinamento. A esta organização em camadas dá-se o nome de *Topologia da Rede*.

A definição da *Topologia da Rede Neural* é uma das maiores dificuldades na modelagem neural, mas é também um passo importante na sua concepção, pois ela limita o tipo de problema que pode ser abordado pela rede. Assim, para se obter uma boa generalização, deve-se proporcionar à rede a maior quantidade possível de informação. Mas isso implica também em construir uma topologia com uma grande quantidade de neurônios, o que acarreta um outro problema, a complexidade computacional que limita o uso em relação a tempo de processamento e espaço (tamanho de arquivos).

Normalmente, uma topologia adequada é obtida através de sucessivos refinamentos a partir da definição de uma topologia inicial. Esse modo de definir a topologia pode consumir um tempo considerável, fruto do empirismo empregado nesse método, sem que chegue a chamada “topologia ideal”.

Buscando-se diminuir a um mínimo adequado o número de neurônios e, por conseguinte, as conexões entre eles, deve-se fazer uso de algoritmos de otimização sempre que possível. Uma tentativa nesse caminho é a utilização de

algoritmos genéticos combinados a RNAs para otimizar topologia e parâmetros de treinamento da rede. Estes parâmetros podem incluir as funções de ativação, taxa de aprendizado e o termo momento.

Há um projeto de doutoramento em andamento, ligado ao Instituto de Ciências Matemáticas e Computação do Campus da USP em São Carlos - SP, que se propõe ao desenvolvimento de um método de otimização de topologias para RNAs utilizando computação distribuída, algoritmos genéticos e RNAs para varrer parte do espaço possível de topologias e parâmetros na busca da topologia mais adequada aos dados disponíveis. Como este projeto ainda está em desenvolvimento, não foi possível fazer os testes usando esta abordagem.

Foram realizados diversos testes no sentido de orientar a busca da topologia ideal, mais adequada à generalização do problema em questão. Estes testes referem-se a várias topologias diferentes, com variação nos parâmetros durante o treinamento, incluindo a combinação dos elementos abaixo, o que gera um número considerável de diferentes topologias a serem testadas:

- Quantidade de neurônios de entrada: 1, 2, 3, 5 e 6;
- Quantidade de neurônios de saídas: 1 e 2;
- Quantidade de camadas intermediárias: 1, 2, e 3;
- Quantidade de neurônios por camada intermediária: 2,3, 4, 5, 7, 10, 13, 15, 20, 25, 30 e 50.

Realizaram-se testes em todos os subconjuntos de dados, buscando encontrar o melhor conjunto de parâmetros de treinamento com as seguintes variações:

- Taxa de aprendizado (η), variando entre 0; 0,1; 0,2; 0,3; 0,5 e 0,9;
- Termo momento (μ), variando entre 0,0; 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9;

5.2.3 Estratégias de Processamento

Os processamentos tanto neurais quanto GPS foram realizados através de diferentes programas na tentativa da obtenção de respostas adequadas em

um tempo de processamento aceitável, ou seja, que não impedisse a utilização sistemática do método, podendo chegar no máximo a 30 minutos.

5.2.3.1 Árvore de testes para definição da topologia adequada à melhor generalização da rede neural

No caso do processamento neural, realizou-se testes com o EasyNN versão 8.01, um programa comercial desenvolvido e comercializado por Stephen Wolstenholme para treinamento e generalização utilizando o algoritmo *backpropagation* em redes de múltiplas camadas. Este programa não pode ser utilizado em função da sua pouca flexibilidade em relação às saídas da rede.

Tentou-se também a utilização do SNNS¹² versão 4.2, que é um simulador neural implementado e distribuído gratuitamente pelo Institute for Parallel and Distributed High Performance System of *University of Stuttgart* sob o endereço: <http://www-ra.informatik.uni-tuebingen.de/SNNS/>.

O SNNS foi utilizado com bastante eficácia em função das grandes possibilidades que ele oferece em termos de flexibilidade na criação e modificação de topologias e parâmetros. Além disso, possui implementados eficientes modelos neurais. Na maioria dos testes realizados nesta etapa, foi possível usá-lo com sucesso. Apesar dessas características, o SNNS não foi capaz de preencher todas as expectativas e necessidades específicas desta investigação. Assim, buscou-se implementar o modelo algoritmo *backpropagation*, de forma que fossem atendidas as exigências específicas desta modelagem de dados.

Estas exigências referem-se principalmente a sua capacidade de generalização com saídas estimadas com um número mínimo em termos de algarismos significativos que fosse compatível com o nível de normalização exigido e que pudesse após a desnormalização atingir-se uma precisão compatível com as observáveis GPS originais.

Ocorre que o algoritmo *backpropagation* possui algumas características que tornam o aprendizado lento, por vezes estabilizando a solução final em

¹² Stuttgart Neural Network Simulator

mínimos locais. Isto aliado à ineficiência das implementações realizadas, provavelmente em uma linguagem inadequada, tornaram inviável a utilização do algoritmo implementado.

Na literatura foram encontradas diversas variações deste algoritmo com a finalidade de resolver estes e outros problemas de convergência. Algumas destas variações, inclusive o algoritmo *backpropagation Levenberg-Marquardt*. (que foi utilizado para resolver este problema) e que encontram-se descritos no item 4.5.2.1 deste documento.

Escolheu-se então o programa MATLAB versão 6.1, que possui um pacote de ferramentas adequadas a simulação de RNAs, denominado de *Neural Network Toolbox for Use in MATLAB*, cujo algoritmo em questão se encontra implementado.

Adicionalmente, o MATLAB permite que sejam feitas alterações no sentido de se obter um número de significantes desejado na saída da rede neural. A grande flexibilidade na manipulação e formatação de arquivos e diversidade de algoritmos implementados eficientemente o elegeram definitivamente como forma de simulação neural a ser usada nesta pesquisa em conjunto com as implementações já realizadas em Visual Basic versão 5.0.

Para cada conjunto foram gerados os seguintes arquivos de treinamento e teste segundo a seguinte estratégia: data/período/tipo/número de entradas/número de saídas. Para efeito de nomenclatura adotou-se a seguinte:

- L1 é a diferença da contagem de ciclos da portadora L1 entre épocas consecutivas de 15 segundos;
- L2 é a diferença da contagem de ciclos da portadora L2 entre épocas consecutivas de 15 segundos;
- C1 é a pseudodistância a partir do código C/A tomada a cada 15 segundos;
- P2 é a pseudodistância a partir do código P2 a cada 15 segundos;
- AE é o ângulo de elevação de cada satélite a cada época de 15 segundos;

- X^S , Y^S , Z^S são as coordenadas do satélite para épocas múltiplas de 15 segundos, obtidas através da interpolação de Lagrange das efemérides precisas.

Dentro de todos os dados escolhidos, o ano de 2001 apresenta o maior número de estações simultâneas. Assim, o dia 05/12/2001 foi o escolhido.

Avaliação da influência do mapeamento entradas/saídas, nestes testes avaliou-se o comportamento da rede quando se alteravam a quantidade os tipo de entradas e saídas:

- (05/12/2001)/1 h/geral/1 entrada(L1)/ 1 saída(L2);
- (05/12/2001)/1 h/geral/1 entrada(C1)/ 1 saída(P2);
- (05/12/2001)/1 h/geral/2 entradas(L1; AE)/ 1 saída(L2);
- (05/12/2001)/1 h/geral/2 entradas(C1; AE)/ 1 saída(P2);
- (05/12/2001)/1 h/geral/3 entradas (L1; C1; AE)/ 2 saídas(L2; P2);
- (05/12/2001)/1 h/geral/5 entradas(L1; AE; Xs; Ys; Zs)/ 1 saída(L2);
- (05/12/2001)/1 h/geral/5 entradas(C1; AE; Xs; Ys; Zs)/ 1 saída(P2);
- (05/12/2001)/1 h/geral/6 entradas(L1; C1; AE; Xs; Ys; Zs)/ 2 saídas (L2; P2).

Após a determinação da topologia mais adequada, pôde-se avaliar a influência do período de observação, ou seja: Com quantas horas de observação deve ser treinada a RNA para conseguir uma precisão adequada nas estimativas?

- (05/12/2001)/1 hora/geral/melhor ou melhores arquiteturas;
- (05/12/2001)/2 horas/geral/ melhor ou melhores arquiteturas;
- (05/12/2001)/6 horas/geral/ melhor ou melhores arquiteturas;
- (05/12/2001)/12 horas/geral/ melhor ou melhores arquiteturas;
- (05/12/2001)/24 horas/geral/ melhor ou melhores arquiteturas.

Após a determinação da topologia e período de observação, mais adequados, pôde-se avaliar a influência do tipo de agrupamento das estações RBMC, se global ou regional.

- (05/12/2001)/melhor período/global/melhor arquitetura;
- (05/12/2001)/melhor período/norte/ melhor arquitetura;
- (05/12/2001)/melhor período/nordeste/ melhor arquitetura;
- (05/12/2001)/melhor período/centro-oeste/ melhor arquitetura;
- (05/12/2001)/melhor período/sul e sudeste/ melhor arquitetura.

Definidas as melhores generalizações, se geral ou regional, pode-se então comprovar a independência do método proposto, através de repetição para outros conjuntos de dados obtidos em diferentes datas e para diferentes conjuntos de estações:

- (12/05/2000)/melhor período/ melhor tipo/ melhor arquitetura;
- (17 /05/2000)/melhor período/ melhor tipo/ melhor arquitetura;
- (18 /05/2000)/melhor período/ melhor tipo/ melhor arquitetura;
- (13 /11/2001)/melhor período/ melhor tipo/ melhor arquitetura;
- (25 /11/2001)/melhor período/ melhor tipo/ melhor arquitetura;
- (29 /01/2002)/melhor período/ melhor tipo/ melhor arquitetura;
- (07 /02/2002)/melhor período/ melhor tipo/ melhor arquitetura;
- (08 /03/2002)/melhor período/ melhor tipo/ melhor arquitetura.

5.2.3.2 Processamentos dos dados GPS

Neste trabalho o objetivo do processamento GPS tem única e exclusivamente um caráter avaliativo que permita a comparação entre o processamento dos arquivos com dados originais e os arquivos cujos dados são frutos de estimativas neurais.

Assim sendo, não se pretende fazer nenhum tipo de análise ou comparação de estratégias de observação, posicionamento e processamento entre diferentes programas, mas, utilizá-lo apenas como forma de oferecer resultados

claros e objetivos em termos da influência no posicionamento de estações que tenham sido submetidas ao método proposto.

Seguindo essa conduta, adotou-se o programa SKI-Pro versão 2.1 (*GPS Pos-processing Software*) e também o programa SKI versão 2.3, ambos desenvolvidos e comercializados pela Leica Geosystems e disponíveis no Laboratório de Mensuração do Departamento de Transportes-EESC-USP.

5.3 O Programa MORE-GPS versão 1.0 – Manipulação e Modelagem de Observáveis GPS por Redes Neurais Artificiais

Não foi intencional a idéia de se desenvolver e implementar um sistema, mas a premente necessidade de agilizar a pesquisa acabou por obrigar seu desenvolvimento. Ainda que tenha sido fruto da mais estrita necessidade e que permaneça em constante estado de transformação, se adaptando às necessidades que surgiram e certamente ainda surgirão, é uma ferramenta capaz de ajudar bastante aqueles que se lançarem à investigações que necessitem da manipulação de dados GPS.

O programa MORE-GPS é um sistema composto de diversos aplicativos, desenvolvidos e implementados em linguagem *Visual Basic* versão 5.0, com a finalidade de possibilitar a manipulação de um conjunto de arquivos de dados GPS de uma e duas frequências e de efemérides precisas. Atualmente não é possível a utilização das efemérides transmitidas, função da forma como se implementou o programa, mas é uma modificação possível e desejável para trabalhos futuros. Adicionalmente, é capaz de gerar arquivos em formato específico para treinamento e testes de redes neurais e finalmente agregar estas estimativas obtidas a partir de uma rede neural a um determinado arquivo de dados GPS.

A manipulação desses arquivos diz respeito à leitura, extração e transformações dos diversos dados que possibilitem a montagem de um conjunto de padrões de entrada/saída, necessários ao treinamento supervisionado na simulação de uma rede neural.

O programa MORE-GPS permite ainda o uso de uma ferramenta padrão em estatística, que pode ser utilizada para desenvolver o aprendizado *backpropagation*, conhecida como validação cruzada, STONE (1974,1978) apud HAYKIN (1999). Inicialmente o conjunto de dados disponível é randomicamente dividido em um conjunto de treinamento e um conjunto de teste. O conjunto de treinamento é subdividido em dois subconjuntos: o de treinamento utilizado para selecionar e treinar o modelo e o de validação, utilizado para testar ou validar o modelo.

A motivação da *validação cruzada* é ratificar a eficiência do modelo neural com um conjunto de dados diferente daquele que foi utilizado para treinamento. Após a rede ter sido treinada e validada, ela é testada com um conjunto de dados que nunca lhe tenha sido apresentado para testar sua capacidade de generalização.

Uma vez gerados os arquivos de treinamento e testes, pode-se treinar uma ou diversas topologias neurais. Quando a rede estiver adequadamente treinada, ou seja, com o erro menor ou igual aquele necessário para que se atinja uma precisão mínima adequada. Pode-se apresentar a ela os dados de uma estação qualquer dentro da área onde se retiraram os dados para treinamento da rede neural com a finalidade de se obter estimativas adequadas.

Este mapeamento entradas/saídas resultará em um arquivo que pode ser manipulado pelo MORE-GPS para que sejam revertidas todas as transformações a que foram submetidos os dados originais e depois agregados ao arquivo GPS original, dando origem a um arquivo modificado que possuirá o formato RINEX versão 2.1, com a única diferença que a ele foram agregadas estimativas neurais.

Este arquivo modificado poderá então ser utilizado em diferentes condutas de processamento de dados GPS para que se obtenham as coordenadas de uma determinada estação.

O programa MORE-GPS versão 1.0 é um programa compatível com o sistema operacional *Windows* e oferece uma interface amigável com o usuário. Apesar da necessidade, ainda não foi possível construir um guia do usuário, mas certamente o será em trabalhos futuros.

A Figura 15 representa a tela de inicialização do programa e a Figura 16 ilustra a tela principal dos aplicativos de extração, manipulação, geração de arquivos de padrões de treinamento e testes, conversões de formatos e também de agregação de estimativas neurais.



Figura 15 – Tela de inicialização do programa MORE-GPS

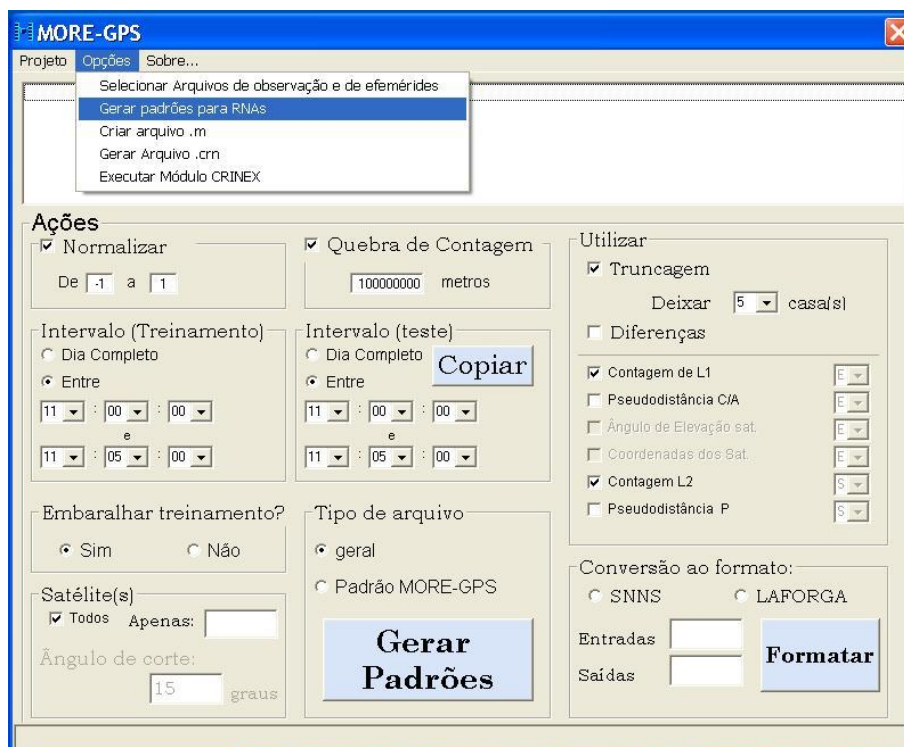


Figura 16 – Tela principal do programa MORE-GPS

A tela principal do MORE-GPS, representada pela Figura 16, além de possibilitar a abertura e manipulação de arquivos de dados GPS e de efemérides precisas, ainda é a tela de chamada dos outros aplicativos que compõem o programa, tais como a geração dos arquivos em formato MATLAB (*.m files*) com extensão (*.m) utilizados para treinamento e estimativas neurais. Possui também a interface para a geração de arquivos fonte e arquivos RINEX.

Os arquivos fonte com extensão (*.crn) são arquivos intermediários, adaptados a um formato específico para posterior geração do arquivo RINEX modificado. Os arquivos fonte possuem as informações de cabeçalho do arquivo original, com as devidas modificações para adaptá-lo à agregação das estimativas, mas permite ao usuário modificar, caso deseje, diversas informações do cabeçalho do arquivo a ser gerado. Possuem também as observáveis originais L1 e C/A e as observáveis estimadas L2 e P2. O módulo CRINEX lê o arquivo fonte e o converte ao formato padrão RINEX versão 2.1.

O menu é composto basicamente por **Projeto**, **Opções** e **Sobre**. O item **Projeto** abre a possibilidade de se criar um projeto novo, abrir um projeto existente ou salvar um projeto. Estas opções foram criadas em função da facilidade de armazenar juntos, arquivos utilizados pelo programa durante sua execução. Isso permite que sejam organizados em projetos diversos trabalhos diferentes e facilita a manipulação de dados diversas vezes, bastando para isso abrir um projeto gerado anteriormente.

O item **Opções** proporciona toda a manipulação de arquivos GPS e efemérides, abertura, inclusão e exclusão, com visualização em tela de nome da estação e data de sua observação, com habilitação ou retirada interativa do processamento. Uma importante observação diz respeito a esse item. Os arquivos de observação e navegação da RBMC referentes ao dia que se deseja estimar, devem estar todos em um mesmo diretório, lá também devem estar o arquivo de observação e navegação da estação que se deseja estimar e o arquivo de efemérides precisas referente a esse dia em formato SP3.

O programa entende que depois de escolhidos todos os arquivos de observação (inclusive o arquivo a ser estimado) e o arquivo de efemérides

precisas ou de carregar um projeto com essa escolha já feita, o arquivo a ser estimado será aquele que estiver sem o sinal de *check* na lista de estações. É fundamental que o usuário retire o sinal de *check* da estação desejada.

O item **Gerar** padrões autoriza ainda a geração e ou conversão de arquivos em formatos específicos adequados ao treinamento neural tanto pelo SNNS quanto pelo próprio aplicativo neural MORE-GPS ou um formato específico para testes que não puderam ser realizados. Além disso, ainda permite a saída em um formato matricial geral, ad-equado ao treinamento neural através do MATLAB.

O programa permite a normalização dos dados dentro de um intervalo especificado pelo usuário. Aceita que seja escolhido um intervalo de tempo dentro daquele em que foram realizadas as observações GPS, e este pode ser diferente para o conjunto de treinamento e teste. Possibilita a aleatoriedade dos dados se necessário e a escolha de quais observáveis ou dados serão utilizados na montagem desses arquivos.

Como se sabe, os valores numéricos no arquivo RINEX relativos à contagem de ciclos das portadoras L1 e L2 não tem nenhuma significância se tomados como valores absolutos, na verdade o que interessa de fato são as diferenças entre esses valores para o mesmo satélite entre duas épocas consecutivas. Entretanto, a contagem inicial destes ciclos pode ser um número qualquer. Este fato dificulta o uso destes dados para o treinamento da rede neural.

Uma rede neural necessita, via de regra, que os dados passem por um pré-processamento no sentido de aplicar transformações visando um bom desempenho no processo de aprendizagem neural e a estabilidade do algoritmo backpropagation durante o processamento. Dentre as transformações, a mais comum é a normalização, que coloca todos os dados comprimidos dentro de um intervalo entre 0 e 1 ou entre -1 a 1.

A operação inversa da normalização é realizada com a aplicação do fator de normalização, fazendo com que eles voltem a ter a mesma amplitude inicial. Porém, se este fator de normalização for muito grande, pode ser que mesmo se

alcançando valores muito bons nas estimativas neurais, ainda assim as estimativas excedam um limite ótimo em termos de precisão. Estes limites são específicos do problema que se deseja tratar.

Especificamente, neste caso onde os dados possuem uma grande amplitude, foi necessário, além de se descartar o valor inicial da contagem de ciclos das portadoras, diminuir ao máximo possível o fator de normalização. Isso foi conseguido com o uso do artifício “*Quebra de contagem*” implementado no programa.

A “*Quebra de contagem*” é a inclusão forçada de falsas perdas de ciclos com o objetivo de “zerar” o contador de ciclos das portadoras. Este recurso permite que a um intervalo escolhido pelo usuário a contagem de ciclos não passe de um determinado limite, limitando assim a amplitude dos valores absolutos das contagens e por conseguinte, reduzindo o fator de normalização.

O segundo recurso também utilizado com o objetivo de diminuir o fator de normalização das observáveis foi denominado “Truncagem” e, como o próprio nome indica, possibilita que o usuário possa truncar os valores das observáveis deixando-as com um número de casas determinadas por ele.

O programa libera também o uso dos valores absolutos da contagem de ciclos das portadoras ou as diferenças de contagem entre épocas consecutivas, esta segunda foi a utilizada neste trabalho, função da melhor capacidade de generalização da rede, diminuindo os erros relativo ao conjunto de treinamento e testes, o que ocasiona estimativas mais próximas dos valores reais.

Implementou-se ainda aplicativos que não fazem parte dessa versão do programa MORE-GPS. São eles os aplicativos de simulação neural implementados em *Visual Basic*, onde se autoriza a definição da topologia da rede, a escolha dos parâmetros de treinamento, treinamento e estimativas, através do algoritmo *backpropagation* padrão.

A não utilização dos referidos aplicativos neurais ocorreu principalmente em função das dificuldades de convergência do algoritmo *backpropagation* padrão, objeto da implementação. Além do fato de que com frequência o

algoritmo estabiliza o processo de aprendizagem, não mais diminuindo os erros do treinamento (são os chamados mínimos locais), existe também o problema do tempo de treinamento, ou seja, é um algoritmo lento, principalmente em se tratando de um grande conjunto de dados, como o que é gerado pelo método proposto.

Certamente, todos estes aplicativos de manipulação de dados GPS podem ser melhorados e o algoritmo *backpropagation* otimizado, mas não é objetivo desta investigação o acesso e domínio de diversas linguagens de programação, mas que se consiga gerar ferramentas adequadas e eficazes à execução da investigação.

É provável que esta limitação, se reflita em um sistema com programas pouco eficientes, geridos por algoritmos que poderiam ser em muito melhorados se otimizados e talvez implementados em linguagens diferentes daquela aqui utilizada. Porém, ainda assim, aquilo que já foi feito pode conduzir ao propósito de comprovar ou refutar as hipóteses. A Figura 17 representa um fluxograma descritivo do programa MORE-GPS.

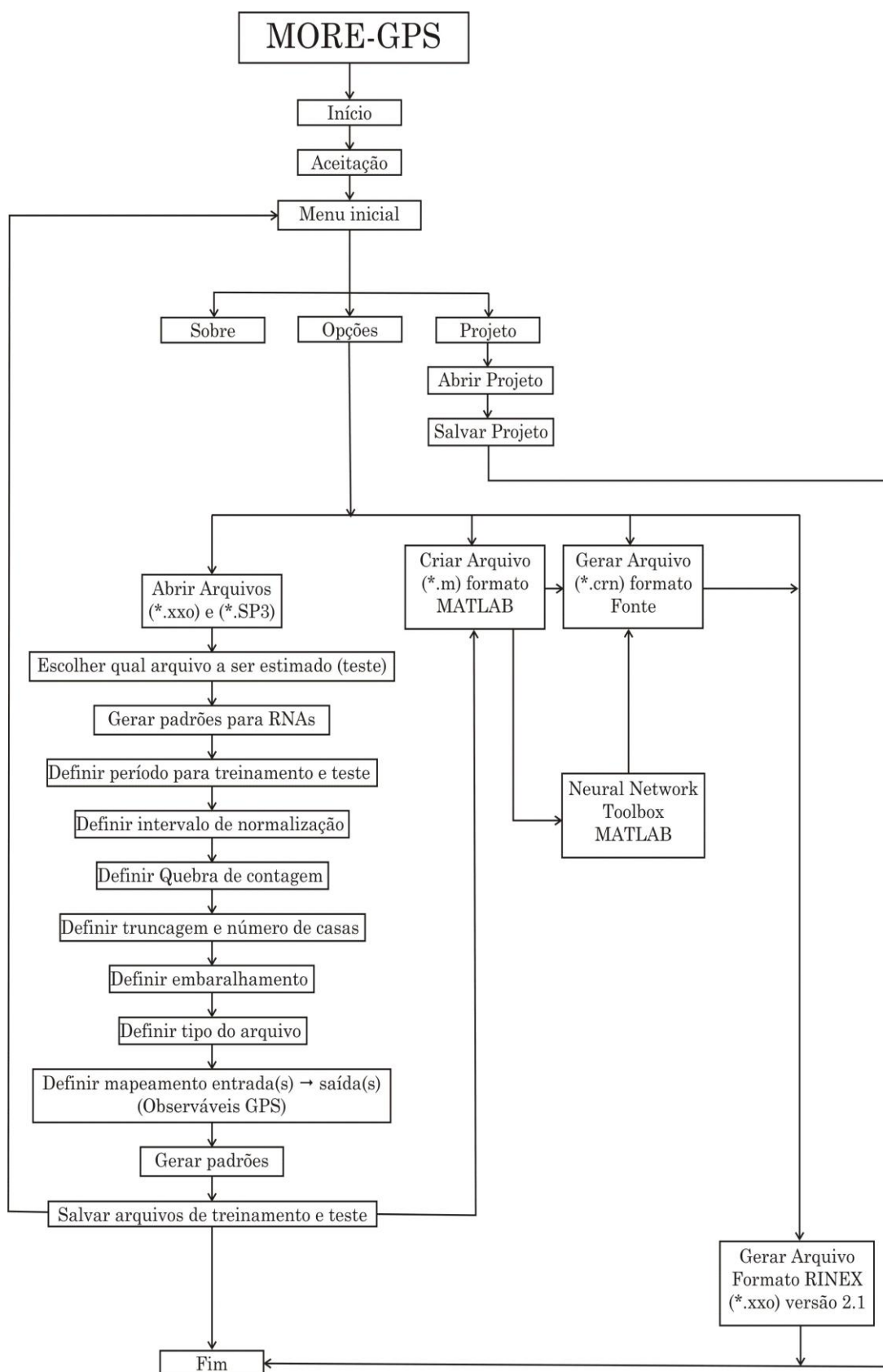


Figura 17 – Fluxograma descritivo do programa MORE-GPS

6

APRESENTAÇÃO DOS RESULTADOS

Confirmando as expectativas iniciais, apesar de se ter testado diversos tipos de redes, tais como a rede *Elman backpropagation*, a rede *backpropagation cascade forward* e a rede de regressão generalizada, o modelo *backpropagation feed-forward* foi mais adequado para que fossem produzidas saídas esperadas, neste caso, contagem de ciclos da portadora L2 e a pseudodistância P2, durante o processo de aprendizagem da rede com o menor erro possível.

Objetivando-se contornar os problemas que com freqüência ocorrem com o treinamento a partir deste algoritmo, principalmente a lentidão na convergência para um resultado satisfatório, utilizou-se uma variação desse algoritmo denominada de *Levenberg-Marquardt*. Este algoritmo encontra-se implementado no *Neural Network Toolbox do MATLAB 6.1*.

6.1 A topologia escolhida

Tornou-se impossível adotar uma conduta mais lógica e racional que aquela adotada aqui na busca e definição da topologia ótima da rede neural. Uma topologia que convergisse para um erro mínimo em um menor tempo de processamento. A topologia adotada nestes resultados e tomada como mais adequada, foi aquela possível de se encontrar dentro das limitações de tempo disponível para a execução dos sucessivos refinamentos necessários.

Durante meses testou-se dezenas, seguramente mais de uma centena de topologias diferentes, fruto das combinações descritas no item 5.2.2.1. Este modo

de definir a topologia depende um tempo considerável e ainda assim é bastante provável que haja uma combinação não testada que responda melhor às expectativas de generalização e tempo de convergência que aquela adotada.

Ficou evidente durante a execução da *árvore de testes*, que não se poderia usar todas as informações retiradas inicialmente do arquivo RINEX original para compor os vetores de padrões de entrada. A *árvore de testes* mostrou claramente que, à medida que se acrescentavam elementos neste vetor, diminuía a capacidade de generalização da rede. Optou-se então por deixar como padrões de entradas, apenas a contagem de ciclos da portadora L1 e a pseudodistância C/A e, como saída, a contagem de ciclos da portadora L2 e a pseudodistância P2. Ainda assim a generalização ficou bastante aquém daquelas observadas quando o vetor de entrada era composto por apenas um elemento.

Decidiu-se então fazer estimativas separadas para contagem de ciclos e pseudodistância. Ocorre que isto implicou em dobrar a quantidade de redes e estimativas. Para um mesmo período, uma rede para estimar L2 a partir de L1 e outra rede para estimar P2 a partir de C/A.

Após a realização de uma outra extensa *árvore de testes* e varrer uma parte significativa do espaço possível, onde deveriam estar as topologias possíveis, obviamente, baseando-se em uma relação custo/benefício. Entenda-se como custo o tempo de convergência e como benefício o erro mínimo no conjunto de treinamento e principalmente no conjunto de teste.

Apesar de se ter feito diversos testes com duas e três camadas intermediárias, adotou-se uma topologia simples, com apenas uma camada intermediária de 6 ou 30 neurônios, dependendo se o treinamento for para fase ou pseudodistância, respectivamente. Esta escolha baseia-se no fato do aumento do número de camadas intermediárias não resultar em acréscimo no tempo de treinamento. Esta decisão foi providencial em função da necessidade de dar rapidez ao método. Já que, à medida que se aumenta o número de camadas e o número de neurônios em cada camada, o tempo despendido no processamento da rede cresce de forma exponencial.

Um exemplo dessa afirmação pode ser visto na comparação entre as Figuras 18 e 19, que ilustram respectivamente o desempenho alcançado por duas redes treinadas com o mesmo conjunto de dados. O conjunto de dados é composto por cinco minutos de observação, de treze estações RBMC, coletados simultaneamente no dia 05/12/2001. Devido ao reduzido intervalo de observação, o tempo de processamento em ambos os casos tornou-se insignificante, variando de poucos segundos até 3 minutos.

Pode-se calcular através da aplicação de estatística diversos índices que demonstram o desempenho do processo de aprendizado durante o treinamento da rede. Adotou-se aqui o *MSE*, do inglês *Mean of Squared Error*, pode ser definido como:

$$MSE = \frac{SSE}{n - p} \quad (14)$$

Onde:

SSE = Somatória dos erros quadráticos do treinamento, que pode ser entendido como as diferenças quadráticas entre os valores de saída da rede neural e os valores esperados como ideais;

n = Número de exemplos do conjunto de treinamento (quantidade de padrões);

P = Número de parâmetros a serem estimados (quantidade de pesos da rede).

Note-se que estes critérios são adequados para modelos lineares. Entretanto, algumas vezes podem ser aplicados diretamente a modelos não lineares caso se esteja trabalhando com um número grande de exemplos de treinamento (JUDGE, 1980).

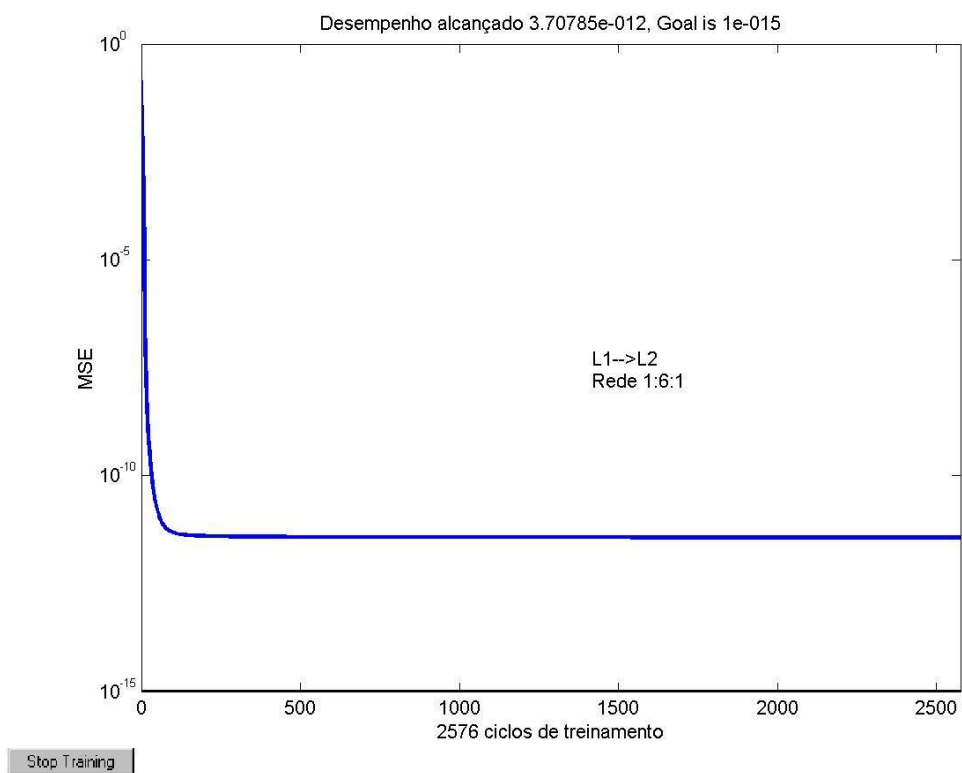


Figura 18 – Desempenho rede 1:6:1 para 5 minutos de observação

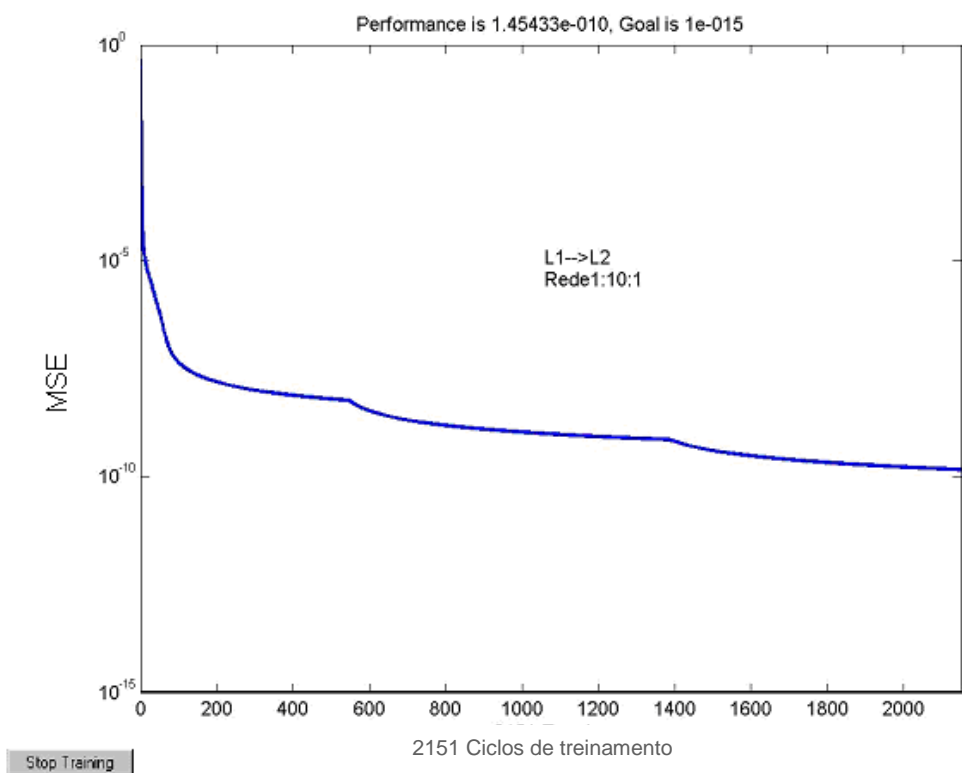


Figura 19 – Desempenho rede 1:10:1 para 5 minutos de observação

A Tabela 5 ratifica a escolha da topologia 1:6:1, a que apresenta o menor MSE, como aquela a ser utilizada para a aproximação de funções que aproximem a observável fase. Além disso, é um exemplo comparativo entre algumas das topologias testadas para dois minutos de treinamento.

Tabela 5 – Comparativo entre diferentes topologias testadas para fase.

| Rede | 1:3:1 | 1:5:1 | 1:6:1 | 1:7:1 | 1:10:1 | 1:15:1 | 1:25:1 | 1:50:1 |
|--------|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Ciclos | 4500 | 2856 | 2576 | 2546 | 2151 | 1551 | 751 | 261 |
| MSE | $2,9.e^{-11}$ | $7,9e^{-12}$ | $3,7e^{-12}$ | $3,9e^{-12}$ | $1,4e^{-10}$ | $2,1e^{-11}$ | $7,8e^{-12}$ | $5,9e^{-12}$ |

Há de se notar que apesar das diferenças de desempenho em relação ao tempo de treinamento e convergência, todas as topologias possuem uma excelente capacidade de generalização, que alcança um erro médio quadrático da ordem de $3,7e^{-12}$ para um curto período de treinamento para a generalização da fase.

Em se tratando da observável pseudodistância a partir dos códigos C/A e P2, o comportamento em relação à topologia difere um pouco da topologia adotada para a generalização da observável fase, pois os resultados e as conclusões em relação a qual topologia utilizar para sua generalização foram diferentes.

A pseudodistância é inerentemente mais ruidosa que a fase, ou seja, a pseudodistância é calculada a partir de informações obtidas no código que se encontra modulado na portadora, mais afetada pela gama de erros inerentes ao processo de propagação do sinal, já a fase é medida pelo receptor através de um processo de sincronização entre os sinais recebidos dos satélites e sinais gerados internamente pelo receptor. Isto deve explicar a diferença de comportamento no treinamento das redes, necessitando de uma quantidade maior de neurônios na camada intermediária. A Tabela 6 apresenta um comparativo de algumas topologias testadas para a pseudodistância.

Tabela 6 – Comparativo entre diferentes topologias testadas para código.

| | | | | | | |
|--------|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Rede | 1:3:1 | 1:5:1 | 1:6:1 | 1:7:1 | 1:10:1 | 1:15:1 |
| Ciclos | 3836 | 2191 | 2147 | 2105 | 1581 | 1156 |
| MSE | 5,2.e ⁻⁹ | 4,9e ⁻⁹ | 4,9e ⁻⁹ | 4,9e ⁻⁹ | 4,8e ⁻⁹ | 4,4e ⁻⁹ |

| | | | | | | |
|--------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Rede | 1:25:1 | 1:30:1 | 1:35:1 | 1:40 | 1:50:1 | 1:100:1 |
| Ciclos | 760 | 450 | 441 | 380 | 306 | 144 |
| MSE | 3,7e ⁻⁹ | 3,4e ⁻⁹ | 3,6e ⁻⁹ | 3,6e ⁻⁹ | 3,3e ⁻⁹ | 2,9e ⁻⁹ |

A topologia adotada como a mais adequada à generalização da pseudodistância foi uma rede com uma camada intermediária de 30 neurônios. A topologia com 100 neurônios tem um ganho muito pequeno em termos de diminuição do erro.

6.2 O tempo de observação

Na seqüência apresentam-se de forma bastante resumida os resultados obtidos na árvore de testes realizada para avaliar a relação entre o tempo de observação GPS e a capacidade de generalização da rede neural. Entenda-se como tempo de observação o período em horas para o qual se deseja gerar estimativas.

O programa MORE-GPS foi desenvolvido com a expectativa de se poder optar pela extração de dados desde um dia completo de observação, ou de vários dias se assim desejar o usuário, até alguns poucos segundos. Possibilita também que esse período seja diferente para os arquivos que serão utilizados para treinamento da rede neural e para o arquivo que se deseja estimar as observáveis.

Existe um inconveniente em relação ao aumento do período de observação GPS. São extraídos dados e informações a cada 15 segundos de todos os arquivos da RBMC que foram observados simultaneamente àquele que se deseja estimar.

Isso provoca um aumento excessivo de dados, com os quais a rede diminui a sua capacidade de generalizar.

Não se deve esquecer que apesar de não aparecerem de forma explícita na modelagem, a composição de todos os erros inerentes à propagação do sinal GPS desde a sua saída do satélite até ser captado pela antena do receptor, deve aparecer como resultado final nas observáveis.

Quanto maior for a taxa de coleta desses dados, provavelmente maiores serão as diferenças entre os erros em épocas consecutivas, fruto da mudança da constelação e da posição relativa de cada um dos satélites em relação à estação em questão.

Pode-se estimar a quantidade de dados da seguinte forma:

$$D = E.S.O.ET \quad (15)$$

Onde:

- D é a quantidade de dados gerados no arquivo de treinamento, com a grande maioria das variáveis calculadas com precisão double e escritas no formato long;
- E é a quantidade de épocas relativa a um dia inteiro de observação a cada 15 segundos;
- S é o número de satélites observados em cada época;
- O é a quantidade de observáveis GPS no arquivo de treinamento;
- ET é a quantidade de estações RBMC operando simultaneamente.

Tomando como exemplo o dia completo 05/12/2001, referente ao dia GPS 339, no qual havia 14 estações RBMC observando simultaneamente, pode-se chegar a uma quantidade de 3.870.720 dados. Isso apenas para um dia de observação, provoca a geração de grandes arquivos, que ocupam mais de 50Mb de memória e de difícil manipulação. O tempo de processamento também aumenta consideravelmente, chegando a algumas horas apenas para a geração dos arquivos de treinamento.

Um exemplo de um arquivo como o descrito acima pode ser:

$$((24h * 240 \text{ épocas}) * 12 \text{ satélites} * 4 \text{ observáveis} * 14 \text{ estações}) = 3870720 \text{ dados}$$

Ainda assim, realizaram-se tentativas com diversos períodos de um dia e chegou-se a conclusão de que não há nenhum sentido em estender-se o período de estimativas pelo simples fato de que a capacidade de generalização da rede é inversamente proporcional ao tamanho da janela de observação. Isto ocorre provavelmente porque à medida que aumenta o tempo de observação, os erros inerentes ao processo de propagação do sinal se modificam, fazendo com que a rede aprenda e desaprenda de forma continuada.

Diante disso, buscou-se diminuir o quanto possível o tamanho da janela de observação e após refinamentos consecutivos determinou-se que a janela de observação máxima para que se obtenham estimativas compatíveis com a precisão das observáveis GPS é da ordem de alguns minutos.

A adoção do intervalo de cinco minutos como o padrão para montagem dos arquivos de treinamento e teste gerou um outro problema. SEGANTINE (1999), afirma que para se obter um resultado que satisfaça critérios previamente estabelecidos, deve-se estar atento ao tempo de observação, pois ele é função de alguns fatores tais como: comprimento da linha base para posicionamento relativo, número de satélites observados, geometria dos satélites (GDOP) e condições atmosféricas, principalmente da ionosfera.

Mesmo que se considerem todos estes fatores, que interferem no processo de medição com o GPS, é muito difícil, senão impossível estabelecer um tempo de observação padrão que garanta o sucesso da operação. Normalmente adota-se como conduta para o método de posicionamento estático um período de observação entre uma e duas horas com taxa de coleta da ordem de 15 segundos.

A árvore de testes possibilitou que fossem testados diferentes agrupamentos temporais dos dados, mas percebeu-se que ao agrupar em um mesmo conjunto de treinamento, dados coletados por um período de cerca de uma hora, a capacidade de generalização da rede neural diminuía bruscamente. O período ótimo para uma boa capacidade de generalização ficou entre 5 e 10 minutos.

Assim sendo, para que se gere um arquivo de uma hora de observação com estimativas neurais, devem-se subdividir os arquivos em períodos de cinco minutos e gerar conjuntos de arquivos (treinamento e teste), tanto para a fase quanto para a pseudodistância. Este procedimento resulta na geração de 48 arquivos diferentes, sendo 24 para treinamento e teste da fase e 24 para treinamento e teste da pseudodistância e o dobro disso caso o período necessário para o processamento GPS seja de duas horas.

Apesar disto, o programa extrai informações e gera automaticamente arquivos para treinamento e teste para quaisquer períodos escolhidos pelo usuário, até mesmo se o usuário escolher diferentes intervalos para treinamento e teste.

As Figuras 18 e 20 indicam a diferença dos erros mínimos alcançado por dois conjuntos de treinamento com tempos de observação diferentes, respectivamente cinco minutos e duas horas. A linha horizontal presente na Figura 20 indica um critério de parada adotado pelo usuário, ou seja, quando deve parar o processo de treinamento da rede.

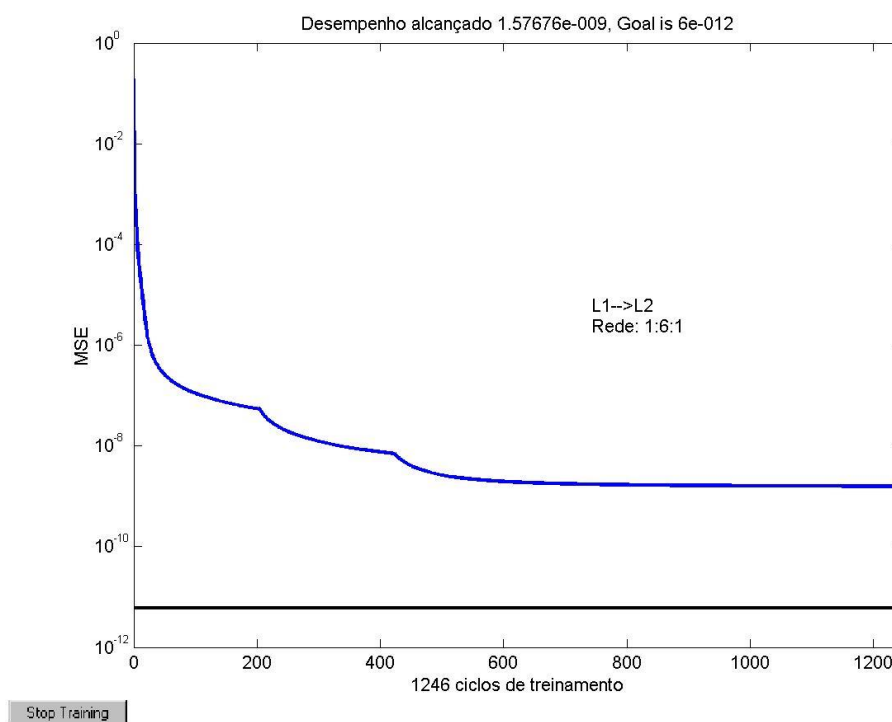


Figura 20 – Desempenho alcançado pela estimativas de duas horas de observação

Adicionalmente, pode-se ressaltar que os tempos consumidos para convergência a estes resultados também guardam uma significativa diferença: 2 minutos para o processamento do arquivo com 5 minutos de observação e 18 minutos para o processamento do arquivo de 2 horas de observação, incluindo todas as estações RBMC disponíveis. Além, é claro, das diferenças nos erros mínimos alcançados que é de $3,7e^{-12}$ para o arquivo de 5 minutos de observação e de $1,57e^{-9}$ para o arquivo com duas horas de observação.

6.3 Características do agrupamento das estações RBMC mais adequado ao modelamento neural dos dados

Depois de definidas topologia e período de observação, pôde-se analisar se a cada estimativa deveriam ser usados dados de todas as estações RBMC ou se ao se agrupar regionalmente parte das estações e usá-las no treinamento se conseguiria reduzir o MSE e com isso aumentar a precisão das estimativas.

Neste sentido houve a necessidade agrupar as estações e adotou-se então, o agrupamento das estações não de acordo com a divisão político geográfica do Brasil, mas denominou-se de região a área de abrangência das quatro estações RBMC mais próximas do ponto a ser estimado.

Através de uma análise rápida da Figura 9, pode-se perceber os imensos vazios na Amazônia, o que atualmente inviabiliza a aplicação do método proposto nesta região. Escolheu-se como área teste a região mais ao centro-sul Brasil, com cinco estações da RBMC (BRAZ, BOMJ, CUIB, UEPP e VICO) como mostrado na Figura 21.

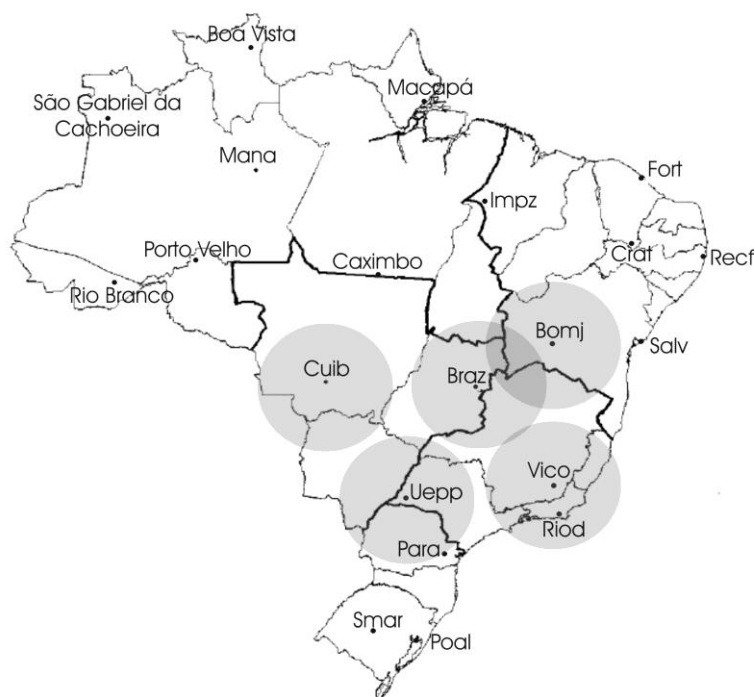


Figura 21 – Agrupamento regional das estações RBMC

As estações BOMJ, CUIB, UEPP e VICO formaram o conjunto de treinamento e a estação BRAZ foi utilizada para geração das estimativas. Na verdade o MSE para o conjunto de treinamento, apresentado na Figura 22, mesmo para uma quantidade de ciclos de treinamento maior, está próximo daquele obtido para as estimativas geradas com todas as estações RBMC. entretanto o MSE para o conjunto de teste é maior, da ordem de 10^{-9} , o que certamente provoca uma diminuição na precisão final das estimativas. Parece que a contribuição das outras estações, ainda que distantes geograficamente da estação em questão, podem contribuir para que a rede consiga uma generalização mais ampla.

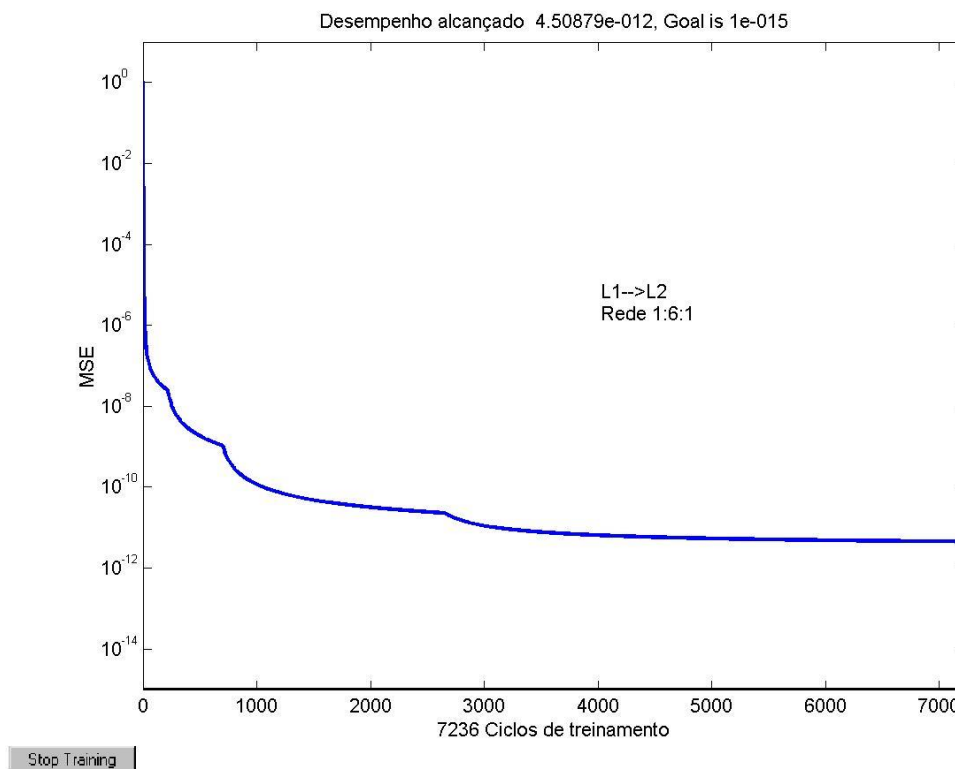


Figura 22 – Desempenho do treinamento para 5 minutos com agrupamento regional de estações RBMC

A partir desta constatação, optou-se pela utilização de todas as estações que coletaram dados simultaneamente àquela que se deseja estimar.

6.4 Testes do método para diferentes conjuntos de dados

Com o objetivo de comprovar a independência do método proposto dos dados com os quais ele foi desenvolvido, lançou-se mão de outros conjuntos de dados observados com outros conjuntos de estações, em outros meses e anos.

Utilizaram-se outros três conjuntos de dados nos dias 17/05/2000, 29/01/2002 e 08/03/2002, cujos resultados para o treinamento se encontram representados pelas Figuras 23, 24 e 25 respectivamente. As estações RBMC em operação em cada um destes dias são:

- ▶ Dia 17/05/2000: 13 estações (BOMJ, BRAZ, CRAT, CUIB, IMPZ, MANA, PARA, POAL, RECF, RIOD, SALV, UEPP e VICO);

- Dia 29/05/2002: 12 estações (BRAZ, CRAT, CUIB, IMPZ, PARA, POAL, RECF, RIOD, SALV, SMAR, UEPP, VICO);
- DIA 08/03/2002: 12 estações (BRAZ, CRAT, CUIB, IMPZ, PARA, POAL, RECF, RIOD, SALV, SMAR, UEPP, VICO).

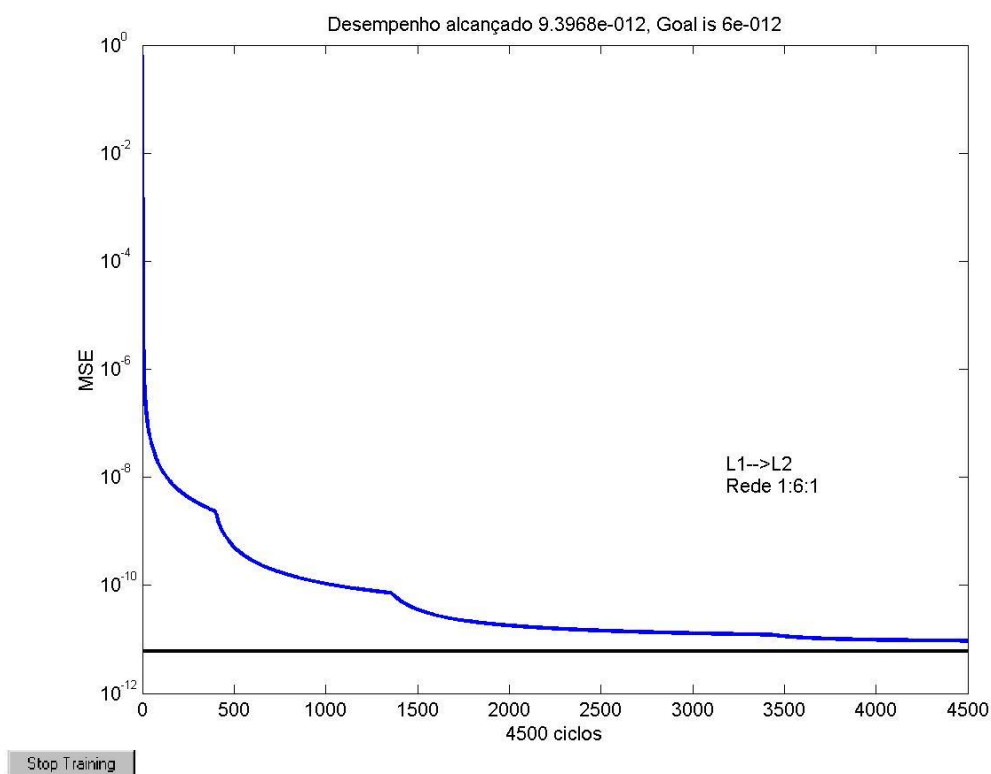


Figura 23 – Desempenho do treinamento para 5 minutos no dia 17/05/2000 (Dia GPS 138)

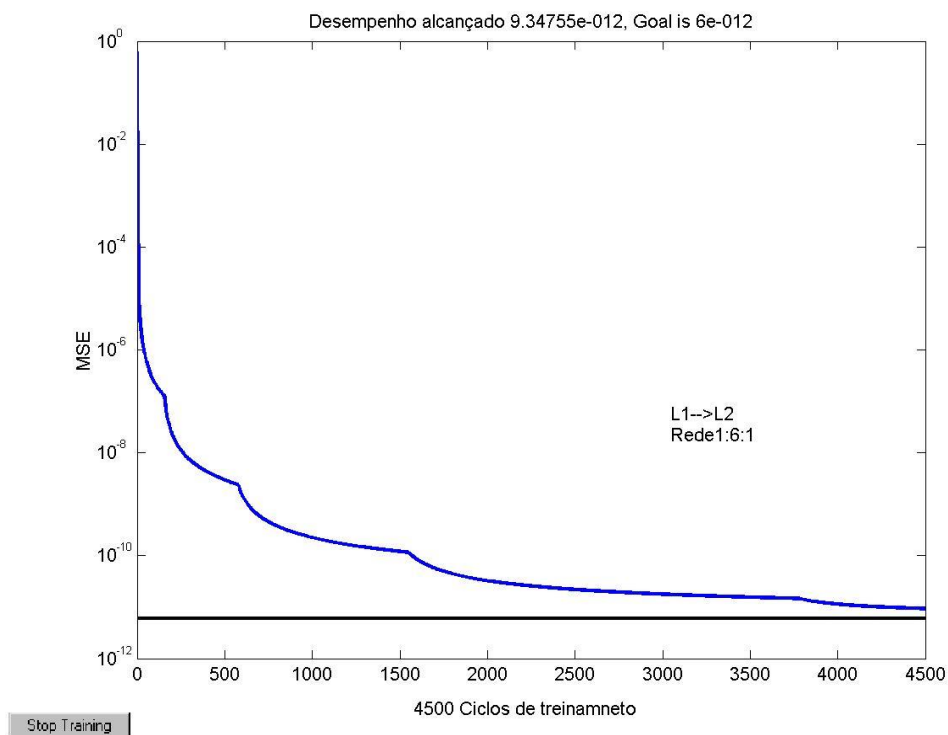


Figura 24 – Desempenho do treinamento para 5 minutos no dia 29/01/2002 (Dia GPS 29)

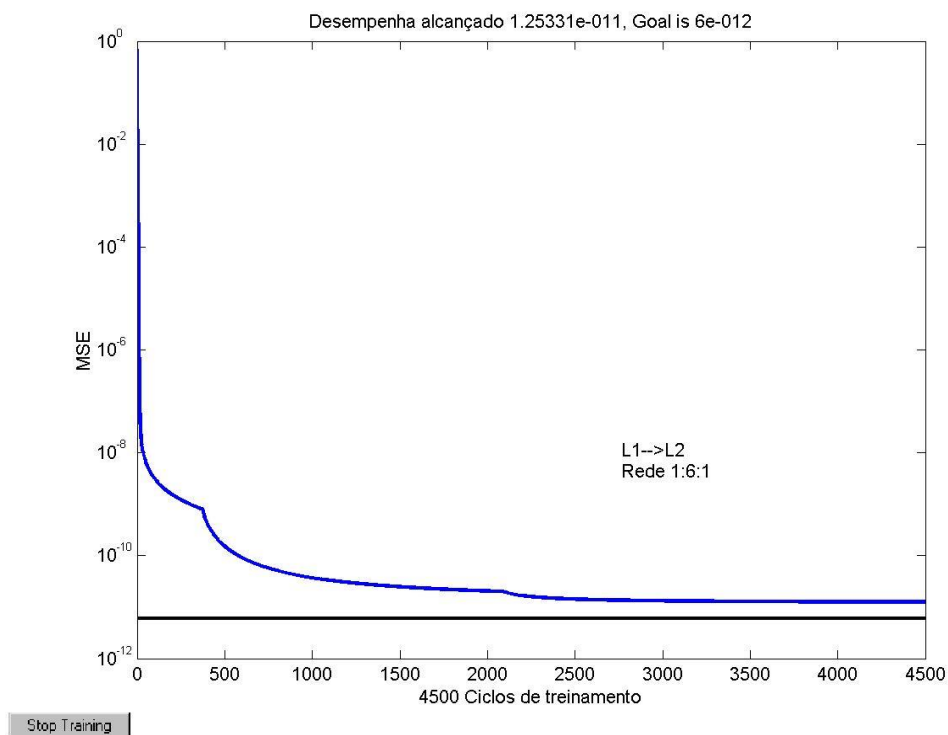


Figura 25 – Desempenho do treinamento para 5 minutos no dia 08/03/2002 (Dia GPS 67)

Em todos os conjuntos de dados testados não ocorreram variações significativas no comportamento da rede neural em relação à magnitude dos erros e estimativas geradas. O que comprova que o método é eficaz, independente do conjunto de treinamento, qualquer que tenha sido o ano, mês ou dia, desde que se tenham dados da RBMC ou alguma outra rede com as mesmas características.

6.5 Validação do método

O desenvolvimento de uma nova técnica, de um novo método, fundamentado em visões e abordagens inovadoras não pode prescindir de um processo de validação que o qualifique como capaz de comprovar ou refutar as hipóteses aventadas inicialmente, além é claro, de permitir que seja reproduzido por outros pesquisadores.

Por não se esperar nenhuma leniência sobre as idéias aqui apresentadas, realizou-se uma série de processamentos com os arquivos originais e com os arquivos modificados pelas estimativas neurais. Apesar disso, sabe-se que ainda existe um número considerável de testes que poderiam ter sido feitos, mas que devido às limitações impostas pelo tempo e pela disponibilidade de dados, não puderam ser realizados ainda.

O desenvolvimento do método e os resultados comparativos apresentados até agora, por si só já comprovam uma das hipóteses iniciais, ou seja, realmente é possível estimar o comportamento de observáveis GPS, ainda que por um curto intervalo de tempo e principalmente, que redes neurais artificiais são uma ferramenta eficaz na obtenção dessas estimativas.

Embora essa seja a tese a ser defendida, não se pode deixar de aplicar o método para atender ao objetivo precípua para o qual ele foi desenvolvido, ou seja, a tentativa de melhorar a precisão posicional de uma determinada estação observada por um receptor de uma freqüência.

Assim sendo, a seqüência de processamentos que se segue tem o estrito objetivo de apresentar resultados em termos de coordenadas e precisão

posicional para as diversas condutas preconizadas no item 5.2.1 e finalmente tentar responder às seguintes perguntas:

- **A aplicação deste método melhora a precisão final das coordenadas de uma determinada estação cujos dados foram coletados por receptor de uma frequência?**
- **Caso melhore, sob que condições essa asserção é verdadeira e em quanto melhora?**

6.5.1 Dados para validação

Apesar do método ter sido desenvolvido visando sua aplicação em arquivos gerados por receptores de uma frequência, não haveria como fazer testes que pudessem comprovar a aplicabilidade e eficácia do método se não fossem feitas observações em pontos cujas coordenadas sejam bem conhecidas. Além disso, havia ainda a necessidade de um comparativo entre os resultados, caso a observação houvesse sido feita por um receptor de dupla frequência e código P.

Buscando-se aplicar o método a estações que possibilitassem o processamento relativo com vetores de diferentes comprimentos de linhas base, escolheu-se dados gerados por receptores de dupla frequência e código P. Destes arquivos, denominados de originais, foram excluídos os valores reais de contagem de ciclos de L2 e pseudodistância P2. Aplicou-se a ele o método descrito. Após gerarem-se as estimativas de L2 e P2, estas foram agregadas ao arquivo, este agora denominado de RINEX modificado.

6.5.1.1 Processamento por ponto simples (single point)

O posicionamento por ponto simples é uma variação do posicionamento absoluto, também chamado de posicionamento instantâneo. Até o dia 1º de maio de 2000, quando o efeito SA, do inglês *Selective Availability*, ainda estava ativo, fazendo com que a posição instantânea de um ponto, usando a pseudodistância a partir do código C/A, possibilitasse uma precisão planimétrica por volta de 100

m. A partir da desativação deste efeito a precisão posicional melhorou cerca de 10 vezes.

O posicionamento pontual refere-se ao posicionamento por ponto simples, do inglês *single point*, também baseado na observável pseudodistância, normalmente a partir do código C/A. Porém, nada impede que se utilize também o código P, caso ele esteja presente no arquivo de observação. A peculiaridade do chamado ponto simples é que o difere do posicionamento instantâneo, é o fato de se fazer um ajustamento do posicionamento instantâneo para diversas épocas.

Processamento Pontual I:

- PERÍODO DE OBSERVAÇÃO: DOY 339/2001 (11:00:00 a 11:04:30);
- ESTAÇÃO VIÇOSA (VICO);
- POSICIONAMENTO PONTUAL (*SINGLE POINT*)-CÓDIGO C/A E P2;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;
- MODELO IONOSFÉRICO STANDARD DO SKI PRO);
- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15°.

A Tabela 7 apresenta as coordenadas obtidas por processamento pontual do código P2 dos arquivos modificado e original para a estação VICO, localizada em Viçosa-MG, que compõe a RBMC e a Figura 26 apresenta os desvios obtidos para o processamento pontual I. As coordenadas oficiais em WGS 84 da estação VICO são:

- Latitude: -20° 45' 41,41107";
- Longitude: -42° 52' 11,93817";
- Altura Geométrica: 666,056m.

Tabela 7 – Coordenadas e Desvios para o posicionamento pontual I (WGS 84)

| Estação | Latitude Longitude Alt. Geométrica | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-----------------|---|---|------------------------|---------------------------|
| VICOmodif P2 | -20 45 41,22649 -42 52 11,72998 687,561 m | -5,5374 -6,2457 21,505 | 8,347 | 23,068 |
| VICOoriginal P2 | -20 45 41,18777 -42 52 11,75400 697,252 m | -6.699 -5.5251 31,196 | 8,684 | 32,382 |

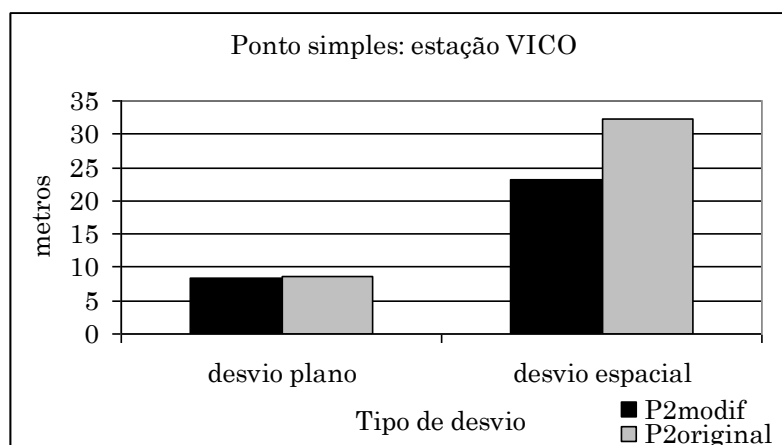


Figura 26 – Desvios para a estação VICO

Processamento Pontual II:

- PERÍODO DE OBSERVAÇÃO: DOY 067/2002 (11:00:00 a 11:04:30);
- ESTAÇÃO PRESIDENTE PRUDENTE (UEPP);
- POSICIONAMENTO PONTUAL (*SINGLE POINT*) - CÓDIGO C/A E P2;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;
- MODELO IONOSFÉRICO STANDARD DO SKI PRO;
- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15°.

A Tabela 8 apresenta as coordenadas obtidas por processamento pontual do código P2 dos arquivos modificado e original para a estação UEPP, localizada em Presidente Prudente-SP, que compõe a RBMC e a Figura 27 apresenta os desvios obtidos para o processamento pontual I. As coordenadas oficiais em WGS 84 da estação UEPP são:

- Latitude: -22° 07' 11,66678”;
- Longitude: -51° 24' 30,69935”;
- Altura Geométrica: 431,039 m.

Tabela 8 – Coordenadas e Desvios para o posicionamento pontual II(WGS 84)

| Estação | Latitude Longitude Alt. Geométrica | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-----------------|--|---|------------------------|---------------------------|
| UEPPmodif P2 | -22 07 11,49903 51 24 30,92609 451,428 m | 5,033 6,802 20,389 | 8,462 | 22,076 |
| UEPPoriginal P2 | -22 07 11,49767 -51 24 30,97202 461,316 | 5,073 8,180 30,277 | 9,626 | 31,770 |

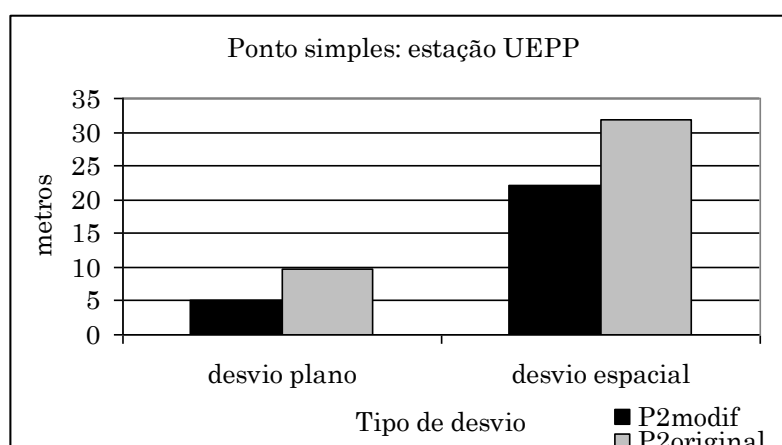


Figura 27 – Desvios para a estação UEPP

Processamento Pontual III:

- PERÍODO DE OBSERVAÇÃO: DOY 038/2002 (11:00:00 a 11:04:30);
- ESTAÇÃO RIO DE JANEIRO (RIOD);
- POSICIONAMENTO PONTUAL (*SINGLE POINT*) - CÓDIGO C/A e P2;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;
- MODELO IONOSFÉRICO STANDARD DO SKI PRO;
- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15°.

A Tabela 9 apresenta as coordenadas obtidas por processamento pontual do código P2 dos arquivos modificado e original para a estação RIOD, localizada em Rio de Janeiro-RJ, que compõe a RBMC e a Figura 28 apresenta os desvios obtidos para o processamento pontual I. As coordenadas oficiais em WGS 84 da estação RIOD são:

- Latitude: -22° 49' 04,24906";
- Longitude: -43° 18' 22,57157";
- Altura Geométrica: 8,745m.

Tabela 9 – Coordenadas e Desvio para o posicionamento pontual III (WGS 84)

| Estação | Latitude Longitude Alt. Geométrica | Desvios em relação às coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-----------------|--|---|------------------------|---------------------------|
| RIODmodif P2 | -22 49 04,17765 -43 18 22,69594 18,083 m | 2,142 3,731 9,338 | 4,302 | 10,281 |
| RIODoriginal P2 | -22 49 04,17446 -43 18 22,69584 28,888 | 2,238 3,728 20,143 | 4,348 | 20,607 |

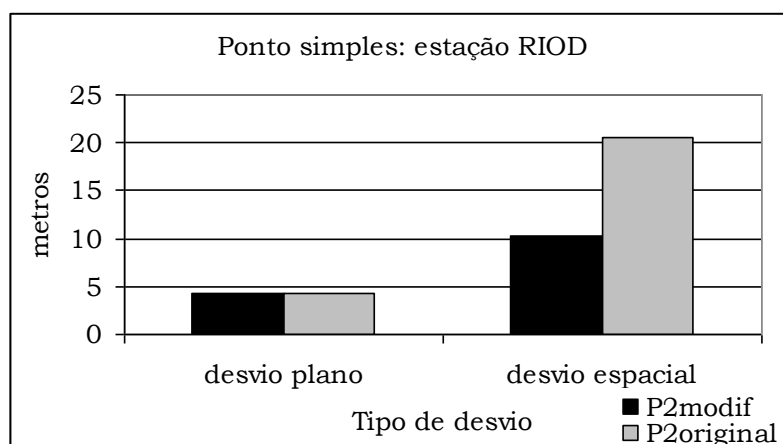


Figura 28 – Desvios para a estação RIOD

6.5.1.2 Processamento relativo estático

“A observável normalmente utilizada para posicionamento relativo estático é a fase da portadora. Neste tipo de posicionamento, dois ou mais receptores, rastreiam simultaneamente, por um período não inferior a 20 minutos, podendo chegar a algumas horas”. (MONICO, 2000).

Segundo SEGANTINE (1995) o objetivo do posicionamento relativo é determinar as coordenadas de um ponto cujas coordenadas são desconhecidas em relação a outro cujas coordenadas se conhece, ou seja, um vetor entre estes dois pontos. Este vetor é chamado de *linha base* ou simplesmente *base*.

Processamento I:

- PERÍODO DE OBSERVAÇÃO: DOY 339/2001 (11:00:00 a 11:04:30);
- PROCESSAMENTO RELATIVO PARA A ESTAÇÃO VICO;
- BOMJ→VICO ($\approx 832km$) BASE LONGA;
- BRAZ→VICO ($\approx 750km$) BASE LONGA;
- UEPP→VICO ($\approx 897km$) BASE LONGA;
- RIOD→VICO ($\approx 232km$) BASE LONGA;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;

- MODELO IONOSFÉRICO STANDARD DO SKI PRO;
- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15°.

A Figura 29 ilustra a posição das estações envolvidas no processamento I (WGS 84).

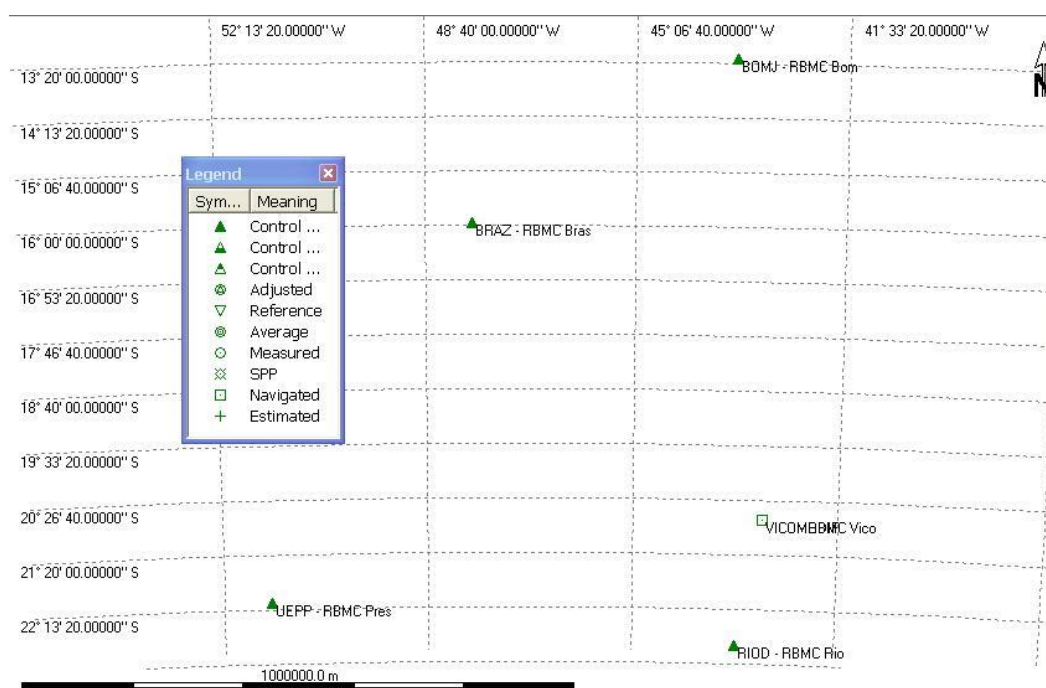


Figura 29 – Estações para o processamento I

A Tabela 10 apresenta as coordenadas geodésicas oficiais em WGS 84 utilizadas como referência no processamento I. A Tabela 11 apresenta os desvios plano e espacial obtidos no processamento I. As Figuras 30, 31, 32 e 33 ilustram uma comparação gráfica entres os resultados obtidos para os processamentos dos arquivos, originais e modificados por estimativas neurais, para as diferentes linhas base do processamento I.

Tabela 10 – Coordenadas oficiais das estações utilizadas no processamento I (WGS 84)

| Estação | Latitude | Longitude | Altura Geométrica (m) |
|---------|-----------------|-----------------|-----------------------|
| BOMJ | -13 15 20,01971 | -43 25 18,22372 | 419,474 |
| BRAZ | -15 56 50,92071 | -47 52 40,30651 | 1105,950 |
| UEPP | -22 07 11,666 | -51 24 30,699 | 431,039 |
| RIOD | -22 49 4,24906 | -43 18 22,57157 | 8,745 |
| VICO | -20 45 41,41107 | -42 52 11,93817 | 666,056 |

Tabela 11 – Coordenadas e Desvios do processamento I (WGS 84)

| Linha base | Latitude Longitude Alt. Geométrica Vetor | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-------------------|---|---|------------------------|---------------------------|
| BOMJ-VICOmodif L3 | -20 45 41,44857 -42 52 11,93861 666,675m 832232,977m | 1,1250 0,0168 0,6192 | 1,125 | 1,284 |
| BOMJ-VICO L1 | -20 45 41,39830 -42 52 11,82686 665,641m 832231,603m | 0,3408 3,2175 0,2275 | 3,235 | 3,243 |
| BRAZ-VICOmodif L3 | -20 45 41,42993 -42 52 11,90530 666,424m 750555,987m | 0,5660 0,9859 0,3677 | 0,986 | 1,052 |
| BRAZ-VICO L1 | -20 45 41,38797 -42 52 11,94492 668,407m 750554,384m | 0,7700 0,5165 2,6529 | 1,086 | 2,867 |
| UEPP-VICOmodif L3 | -20 45 41,39200 -42 52 11,87634 666,492m 897012,819m | 0,5720 1,8550 0,4364 | 1,941 | 1,990 |
| UEPP-VICO L1 | -20 45 41,41962 -42 52 11,99321 667,413m 897009,412m | 0,1420 2,0765 1,8375 | 2,081 | 2,776 |

Tabela 11 – Coordenadas e Desvios do processamento I (WGS 84)

| Linha base | Latitude Longitude Alt. Geométrica Vetor | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-------------------|---|---|------------------------|---------------------------|
| RIOD-VICOmodif L3 | -20 45 41,39799 -42 52 11,94340 666,625m 232120,982m | 0,3923 0,1568 0,5688 | 0,422 | 0,708 |
| RIOD-VICO L1 | -20 45 41,41117 -42 52 11,95969 664,705m 232120,453m | 0,0040 0,7847 1,6074 | 0,785 | 1,789 |

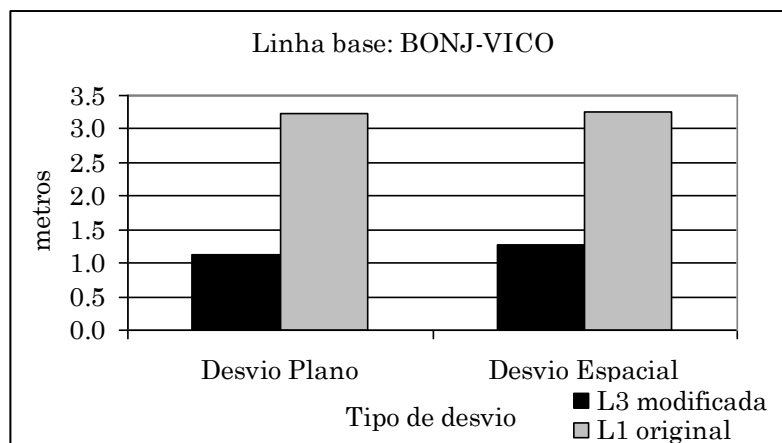


Figura 30 – Desvios linha base BOMJ-VICO

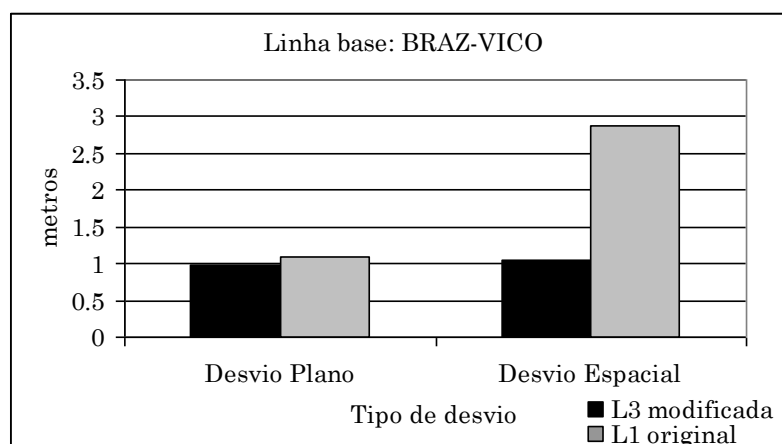


Figura 31 – Desvios linha base BRAZ-VICO

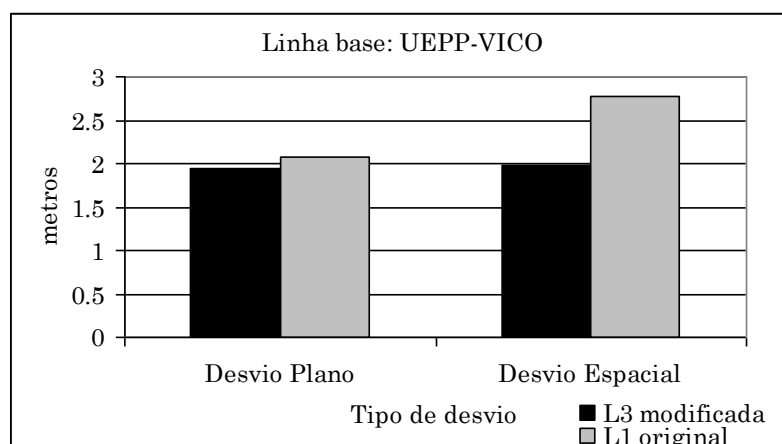


Figura 32 – Desvios linha base UEPP-VICO

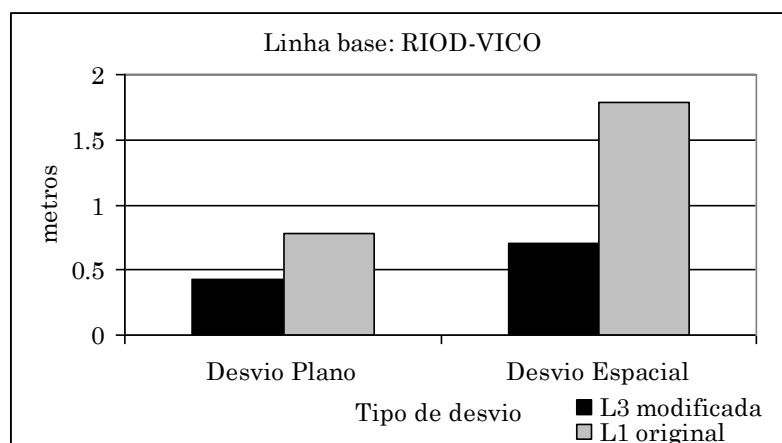


Figura 33 – Desvios linha base RIOD-VICO

Processamento II:

- PERÍODO DE OBSERVAÇÃO: DOY 067/2002 (11:00:00 a 11:04:30);
- PROCESSAMENTO RELATIVO PARA A ESTAÇÃO UEPP;
- BRAZ → UEPP (771km) BASE LONGA;
- CUIB → UEPP (876km) BASE LONGA;
- PARA → UEPP (430km) BASE LONGA;
- VICO → UEPP (897km) BASE LONGA;
- RIOD → UEPP (836km) BASE LONGA;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;
- MODELO IONOSFÉRICO STANDARD DO SKI PRO);

- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15°.

A Figura 34 ilustra a posição das estações envolvidas no processamento II (WGS 84).

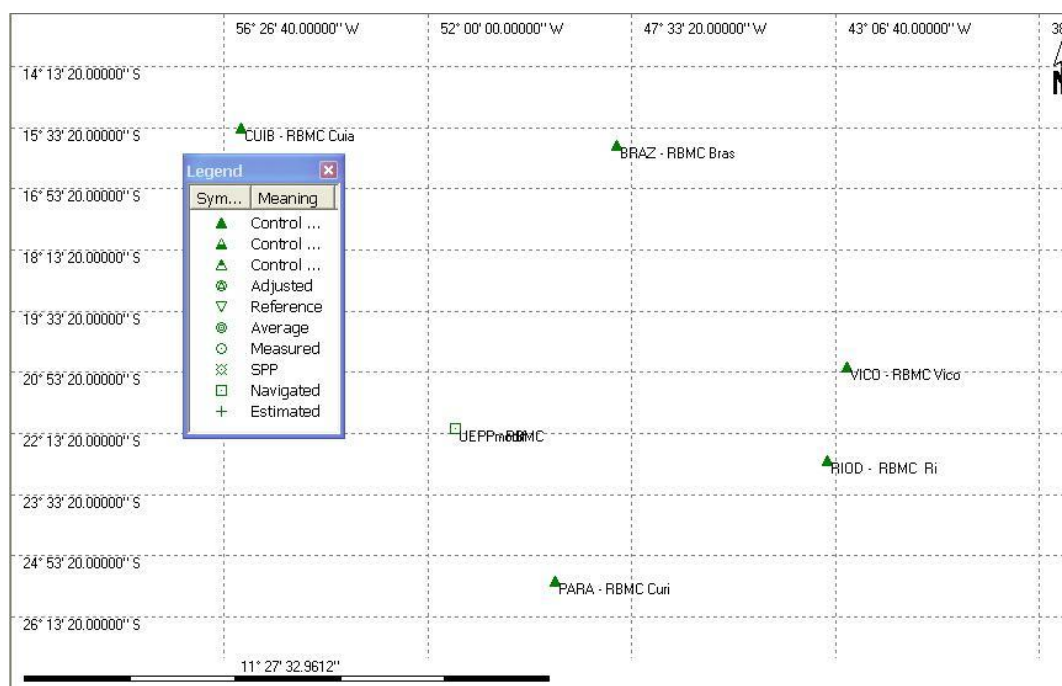


Figura 34 – Estações para o processamento II

A Tabela 12 apresenta as coordenadas geodésicas oficiais em WGS 84 utilizadas como referência no processamento III. A Tabela 13 apresenta os desvios plano e espacial obtidos no processamento II. As Figuras 35, 36, 37, 38 e 39 ilustram uma comparação gráfica entres os resultados obtidos para os processamentos dos arquivos, original e modificado por estimativas neurais, para o processamento II.

Tabela 12 - Coordenadas oficiais das estações utilizadas no processamento II (WGS 84)

| Estação | Latitude | Longitude | Altura Geométrica (m) |
|---------|-----------------|-----------------|-----------------------|
| BRAZ | -15 56 50.92071 | -47 52 40.30651 | 1105.950 |
| CUIB | -15 33 18,95744 | -56 04 11,49752 | 237,358 |
| RIOD | -22 49 4,24906 | -43 18 22,57157 | 8,745 |
| VICO | -20 45 41,41107 | -42 52 11,93817 | 666,056 |
| PARA | -25 26 54,1362 | -49 13 51,4115 | 925,705 |
| UEPP | -22 07 11,66680 | -51 24 30,69935 | 431,039 |

Tabela 13 – Coordenadas e Desvios do processamento II (WGS 84)

| Linha base | Latitude Longitude Alt. Geométrica Vetor | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-------------------|---|---|------------------------|---------------------------|
| BRAZ-UEPPmodif L3 | -22 07 11.69429 -51 24 30.73499 430.164m 777288.044m | 0,8248 1,0692 0,8743 | 1,350 | 2,588 |
| BRAZ-UEPP L1 | -22 07 11.64815 -51 24 30.39423 432.104m 777282.176m | 0,5630 9,1239 1,1623 | 9,141 | 9,215 |
| CUIB-UEPPmodif L3 | -22 07 11.69376 -51 24 30.64403 431.342m 876277.581m | 0,8089 1,6596 0,3034 | 1,846 | 1,871 |
| CUIB-UEPP L1 | -22 07 11.63207 -51 24 30.67417 429.325m 876275.395m | 1,0445 0,8229 1,6266 | 9,141 | 9,215 |
| RIOD-UEPPmodif L3 | -22 07 11.66856 -51 24 30.76582 430.385m 836773.461m | 0,0529 1,9941 0,6533 | 1,995 | 2,099 |
| RIOD-UEPP L1 | -22 07 11.63473 -51 24 30.30827 432.317m 836760.719m | 0,9532 11,7690 1,3640 | 11,807 | 11,886 |

Tabela 13 – Coordenadas e Desvios do processamento II (WGS 84)

| Linha base | Latitude Longitude Alt. Geométrica Vetor | Desvios em relação as coordenadas oficiais (m) | Desvio Plano (m) | Desvio Espacial (m) |
|-------------------|---|---|------------------------|---------------------------|
| VICO-UEPPmodif L3 | -22 07 11.68471 -51 24 30.77165 429.409m 897012.931m | 0,5374 2,1690 1,6291 | 2,235 | 2,765 |
| VICO-UEPP L1 | -22 07 11.64101 -51 24 30.21353 432.042m 896997.128m | 0,7299 14,7579 1,1093 | 14,776 | 14,817 |
| PARA-UEPPmodif L3 | -22 07 11.64424 -51 24 30.70619 431.065m 430236.716m | 0,6767 0,2052 0,027 | 0,722 | 0,723 |
| PARA-UEPP L1 | -22 07 11.65116 -51 24 30.42791 430.959m 430232.469m | 0,4970 8,0802 0,0070 | 8,095 | 8,095 |

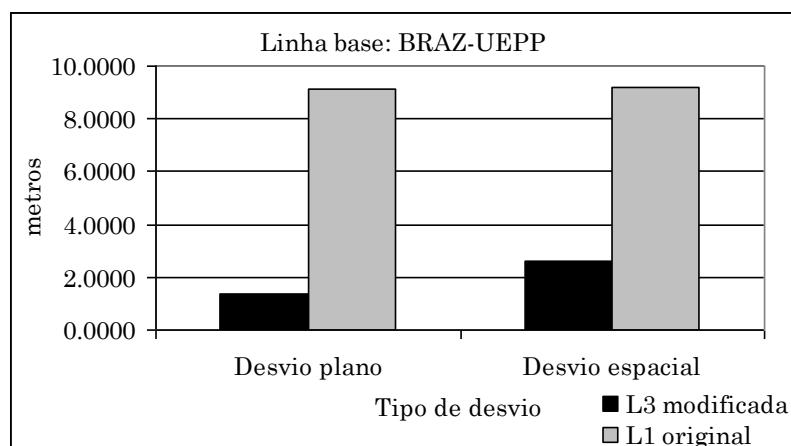


Figura 35 – Desvios linha base BRAZ-UEPP

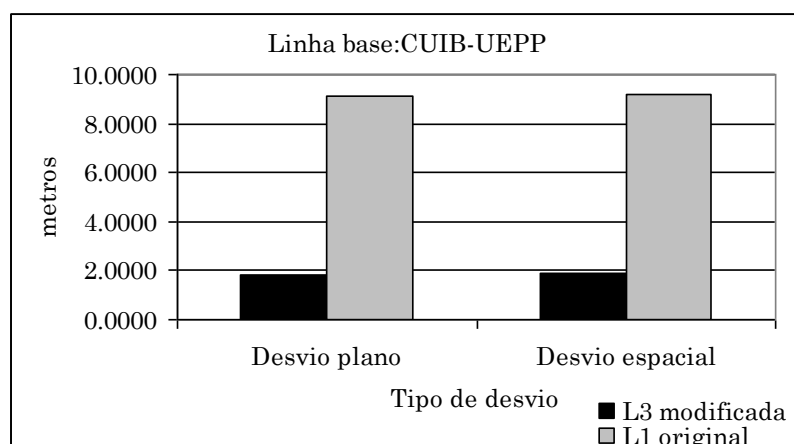


Figura 36 – Desvios linha base CUIB-UEPP

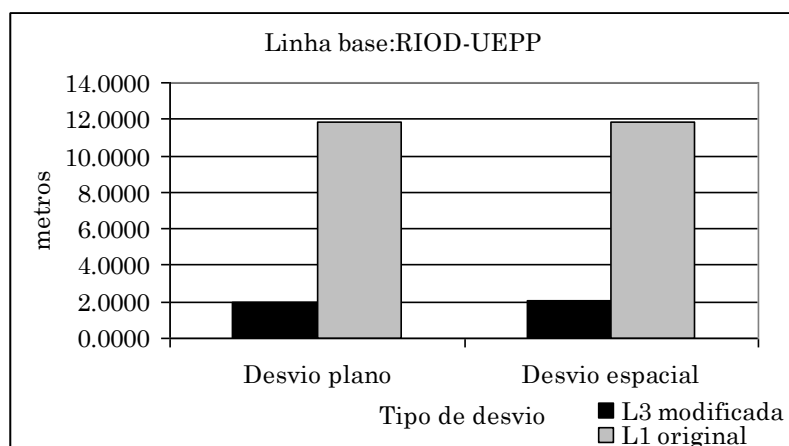


Figura 37 – Desvios linha base RIOD-UEPP

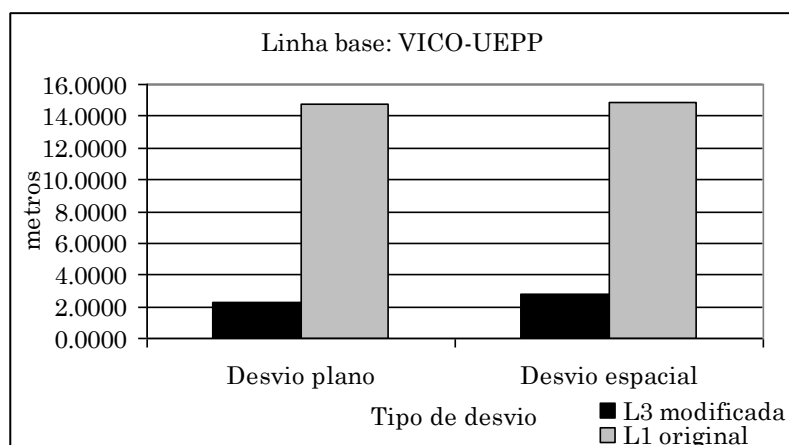


Figura 38 – Desvios linha base VICO-UEPP

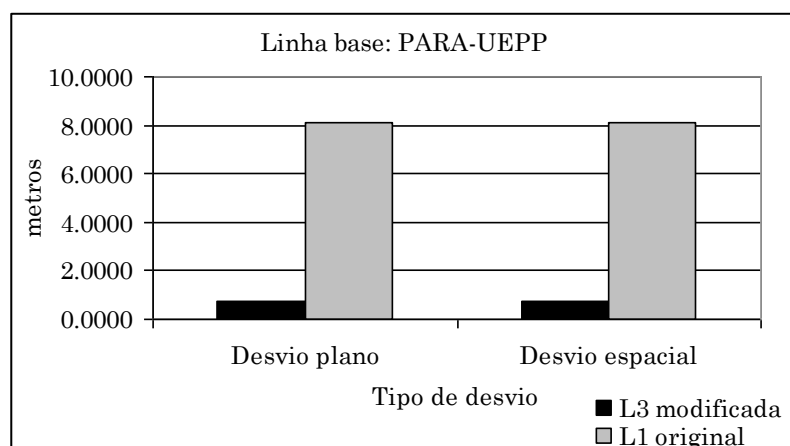


Figura 39 – Desvios linha base PARA-UEPP

Processamento III:

- PERÍODO DE OBSERVAÇÃO: DOY 038/2002 (11:00:00 a 11:04:30);
- PROCESSAMENTO RELATIVO PARA A ESTAÇÃO RIOD;
- UEPP→RIOD ($\approx 836km$) BASE LONGA;
- VICO→RIOD ($\approx 232km$) BASE LONGA;
- PARA→RIOD ($\approx 668km$) BASE LONGA;
- BRAZ→RIOD ($\approx 899km$) BASE LONGA;
- +IMPZ→RIOD ($\approx 1961km$) BASE LONGA;
- MODELO TROPOSFÉRICO DE HOPFIELD SIMPLIFICADO;
- MODELO IONOSFÉRICO STANDARD DO SKI PRO;
- EFEMÉRIDES PRECISAS;
- ÂNGULO DE CORTE: 15° .

A Figura 40 ilustra a posição das estações envolvidas no processamento III.

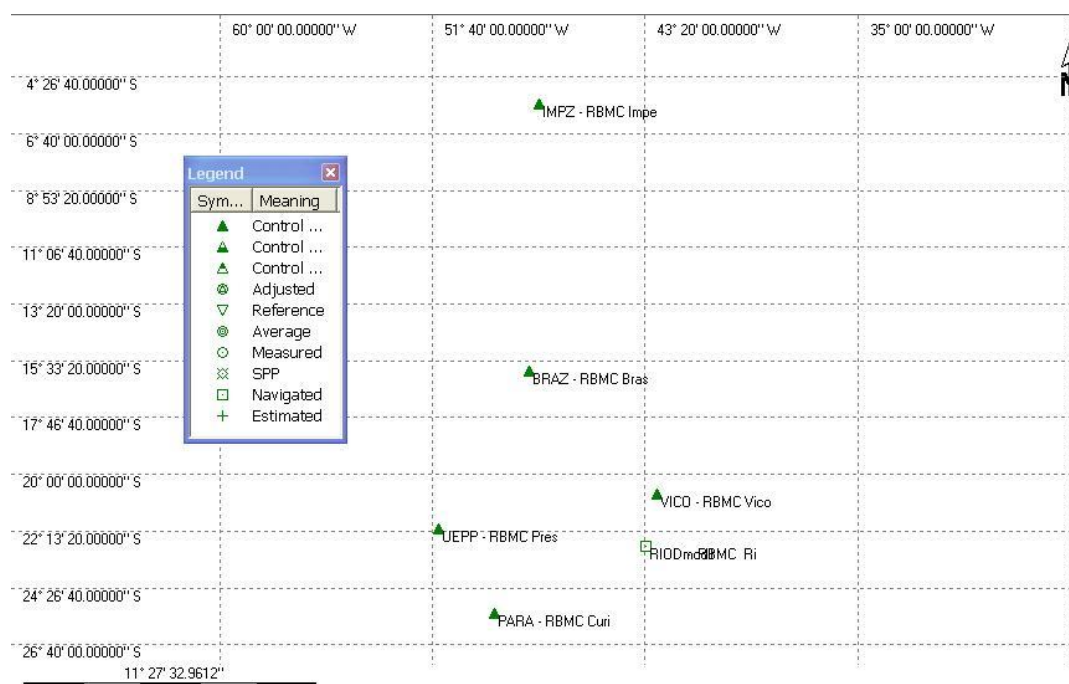


Figura 40 – Estações envolvidas no processamento III (WGS 84)

A Tabela 14 apresenta as coordenadas geodésicas oficiais em WGS 84 utilizadas como referência no processamento III. A Tabela 15 apresenta os desvios plano e espacial obtidos no processamento III. As Figuras 41, 42, 43, 44 e 45 ilustram uma comparação gráfica entre os resultados obtidos para os processamentos dos arquivos original e modificado por estimativas neurais, para o processamento III.

Tabela 14 - Coordenadas oficiais das estações utilizadas no processamento III (WGS 84)

| Estação | Latitude | Longitude | Altura Geométrica (m) |
|---------|-----------------|-----------------|-----------------------|
| IMPZ | -5 29 30,36821 | -47 29 50,02206 | 105,024 |
| BRAZ | -15 56 50,92070 | -47 52 40,30652 | 1105,950 |
| VICO | -20 45 41,41106 | -42 52 11,93817 | 666,056 |
| UEPP | -22 07 11,66680 | -51 24 30,69935 | 431,039 |
| PARA | -25 26 54,13616 | -49 13 51,41145 | 925,705 |
| RIOD | -22 49 4,24905 | -43 18 22,57156 | 8,745 |

Tabela 15 – Coordenadas e Desvios do processamento III (WGS 84)

| Linha base | Latitude Longitude Alt. Geométrica Vetor | Desvios em relação as coordenadas oficiais (m) | Desvio Plano | Desvio Espacial |
|-----------------|---|---|-----------------|--------------------|
| IMPZ-RIODmod L3 | -22 49 04,41497 -43 18 22,44830 8,222m 1961416,184m | 4,9773 3,6981 0,5226 | 6,200 | 6,223 |
| IMPZ-RIOD L1 | -22 49 04,12648 -43 18 22,39136 15,575m 1961409,192m | 3,7293 5,6361 7,2275 | 6,758 | 9,895 |
| BRAZ-RIODmod L3 | -22 49 04,31417 -43 18 22,47967 7,5167m 898583,004m | 1,9533 2,7569 1,2279 | 3,3787 | 3,5949 |
| BRAZ-RIOD L1 | -22 49 04,20820 -43 18 22,32142 16,933m 898583,389m | 1,2258 7,5045 8,1879 | 7,604 | 11,174 |
| VICO-RIODmod L3 | -22 49 04,26456 -43 18 22,57418 8,694m 232121,086m | 0,4650 0,0780 0,0500 | 0,471 | 0,474 |
| VICO-RIOD L1 | -22 49 04,26213 -43 18 22,62223 8,225m 232121,274m | 0,3735 3,2184 0,3650 | 3,240 | 3,260 |
| UEPP-RIODmod L3 | -22 49 04,26503 -43 18 22,44519 8,881m 836775,252m | 0,4791 3,7914 0,1360 | 3,822 | 3,824 |
| UEPP-RIOD L1 | -22 49 04,26601 -43 18 22,52336 14,184m 836773,379m | 0,4812 1,3707 5,5523 | 1,453 | 5,739 |
| PARA-RIODmod L3 | -22 49 04,23825 -43 18 22,47626 10,023m 668631,052m | 0,3243 2,8590 1,2780 | 2,877 | 3,148 |
| PARA-RIOD L1 | -22 49 04,28628 -43 18 22,45644 12,567m 668631,015m | 1,1289 3,1589 3,613 | 3,354 | 4,988 |

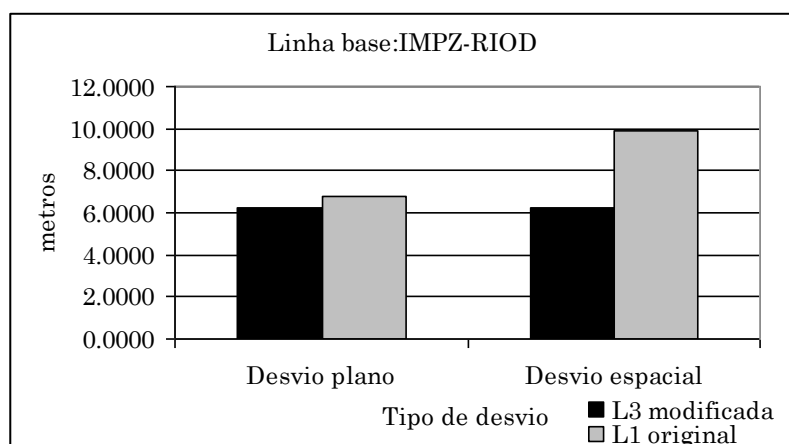


Figura 41 – Desvios da linha base IMPZ-RIOD

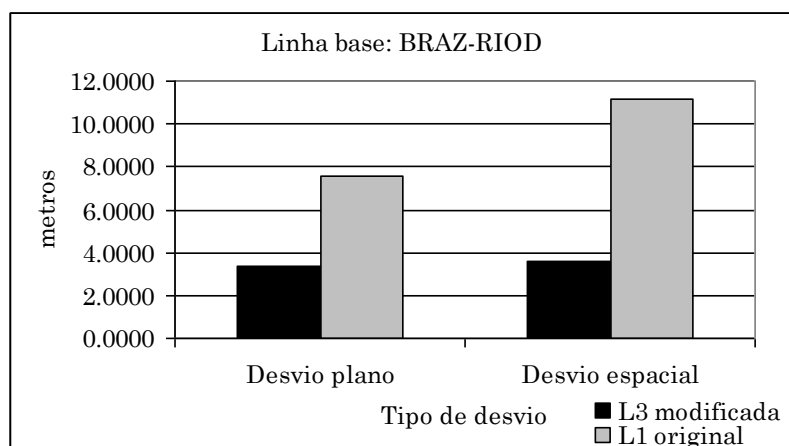


Figura 42 – Desvios da linha base BRAZ_RIOD

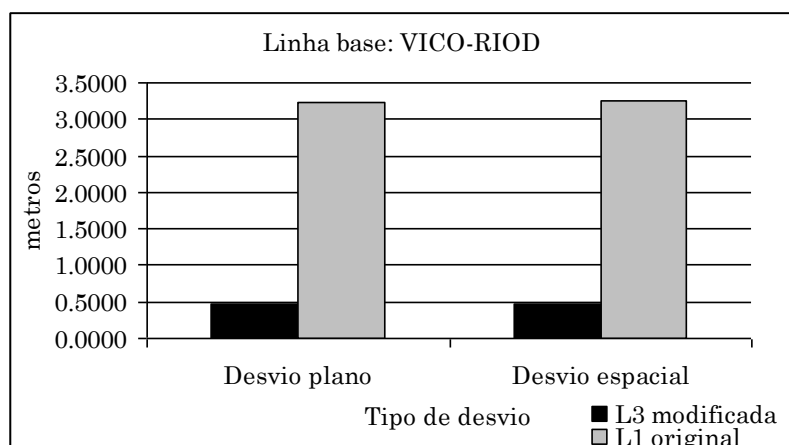


Figura 43 – Desvios da linha base VICO-RIOD

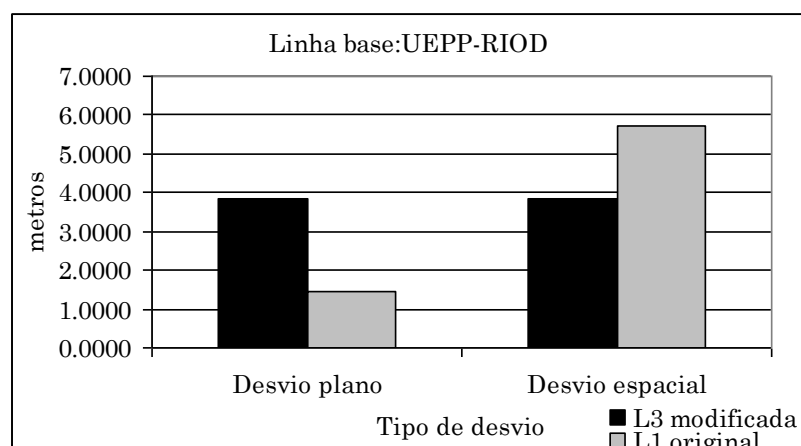


Figura 44 – Desvios da linha base UEPP-RIOD

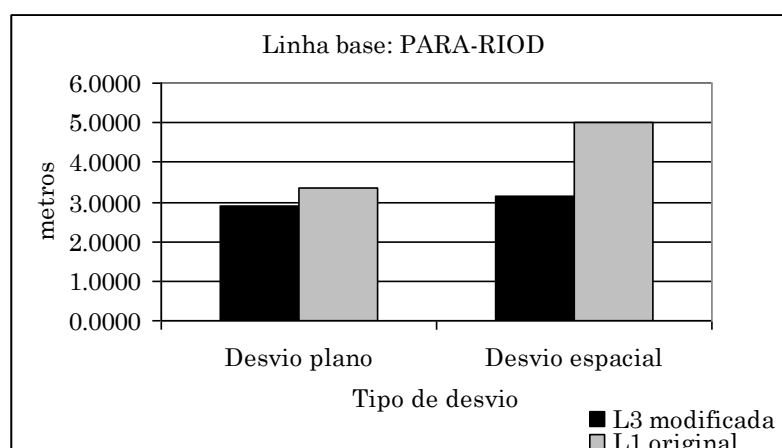


Figura 45 – Desvios da linha base PARA-RIOD

6.5.1.3 Processamento Posicionamento Pontual Preciso (PPP)

Faz parte dos objetivos desta investigação que após o desenvolvimento do método de geração de estimas neurais para dados GPS, os arquivos modificados pudessem ser processados através de diferentes condutas com o objetivo de comparar os resultados para as coordenadas finais do arquivo original e do arquivo modificado. Dentre estas condutas está o método de processamento PPP.

Ocorre que se tornou impossível o processamento PPP em função de algumas das restrições impostas pelo método, o tempo mínimo de observação de uma hora.

O método de geração de estimativas neurais para dados GPS desenvolvido nesta pesquisa obteve sucesso nas estimativas com uma precisão adequada, ou seja, sem que as diferenças entre as estimativas e os valores originais fossem grandes o suficiente para serem consideradas como perdas de ciclos, apenas para curtos períodos de tempo de processamento, neste caso 5 minutos. Isso implica na geração de arquivos para intervalos de tempo muito curtos, neste caso de cinco minutos.

Mesmo o programa MORE-GPS conseguindo gerar arquivos maiores, fruto da união de diversos arquivos gerados com intervalos de tempo menores, ainda permanece a passagem entre esses arquivos menores sendo reconhecida pelos programas de processamento como perdas de ciclos, isso compromete os resultados em termos de precisão de arquivos maiores.

Espera-se em trabalhos futuros conseguir a otimização do método na tentativa de gerar arquivos que preencham as exigências de tempo de observação do PPP, podendo assim, serem processados através desse método.

Assim sendo, apresentam-se na Tabela 16 as coordenadas geodésicas em ITRF2000 dos pontos processados pelo PPP apenas para os arquivos originais.

Tabela 16 –Coordenadas geodésicas ITRF2000 das estações VICO, UEPP e RIOD (PPP-ITRF 2000)

| Estação | Latitude | Longitude | Altura Geométrica (m) |
|---------|------------------|------------------|-----------------------|
| VICO | -20 45 41,401548 | -42 52 11,962092 | 666,091 |
| UEPP | -22 07 11,656200 | -51 24 30,722400 | 431,041 |
| RIOD | -22 49 04,239048 | -43 18 22,596624 | 8,787 |

Mesmo que se consiga no futuro, gerar arquivos de estimativas maiores, superando a restrição do PPP quanto ao tempo de observação, ainda restarão limitações de formato. É importante ressaltar que apesar do RINEX ser um formato cujas especificações são de domínio público e cujas características para a versão atual, o RINEX versão 2.1, serem perfeitamente especificadas por GURTNER (2002), não é uma tarefa muito simples a adequação do formato aos

diversos programas existentes. Essas dificuldades se dão principalmente nos pequenos detalhes que foram incorporados à versão 2, desde o seu lançamento em 1990.

Existem informações no cabeçalho que são consideradas pelo documento que trata das especificações do formato como sendo opcionais, mas que são essenciais para o processamento em alguns programas, tais como o OMNI e o SKI.

Para se ter uma idéia mais geral deste problema, pode-se citar como exemplo o programa SKI-PRO de fabricação Leica. A versão anterior o SKI 2.3, lê e processa com eficiência o arquivo modificado gerado pelo programa MORE-GPS, enquanto o SKI-PRO não o faz, pois apresenta um erro de leitura do arquivo. Para ser processado pelo SKI-PRO o arquivo modificado teve de ser carregado pelo SKI versão 2.3 e gravado com formato RINEX novamente e só então pôde ser lido e processado pelo SKI-PRO.

7

CONCLUSÕES E RECOMENDAÇÕES

7.1 Quanto à modelagem neural

Apesar de se ter feito todos os esforços no sentido de escolher as redes com topologias que oferecessem um erro mínimo em um tempo de processamento razoável, certamente não se esgotou a imensa combinação de possibilidades entre topologia e parâmetros de treinamento. Assim, pode-se utilizar para a modelagem de dados GPS as topologias apresentadas no capítulo 6, mas deve-se considerar que a definição da topologia não é uma questão fechada.

Uma indicação para a realização de trabalhos futuros, pode ser a utilização de um sistema que envolva RNAs, algoritmos genéticos e computação distribuída para a otimização da topologia ótima, de acordo com o proposto por LAFORGA & ROMERO (2002).

Um detalhe que corrobora esta afirmação advém do fato de que mesmo se utilizando redes com topologias diferentes para a generalização das diferentes observáveis L2 e P2, nem sempre se consegue resultados adequados com a topologia indicada neste documento para a generalização da pseudodistância P2, algumas vezes foi necessário dobrar o número de neurônios na camada intermediária para que se obtivesse melhores resultados em termos de generalização da rede. Isso não ocorreu com a observável fase, provavelmente devido suas características de menor ruído e maior precisão.

Um outro fato que se percebeu em toda a árvore de testes que se executou, mesmo entre aqueles que sequer estão documentados neste volume, é a dependência direta entre o erro alcançado no treinamento e o desempenho da

rede na generalização. Porém, não há um ganho significativo em termos de precisão quando se utiliza uma quantidade de ciclos de treinamento muito grande. Nesta investigação sempre se alcançou bons resultados com, no máximo, 5000 ciclos de treinamento.

O algoritmo *backpropagation* realmente atende às necessidades em termos de capacidade de generalização e precisão nas estimativas geradas por Redes Neurais Artificiais onde se utilize como acelerador de convergência, a otimização numérica no ajuste da matriz de pesos da retropropagação. Com frequência o MSE do conjunto de teste e treinamento chega a 10^{-14} . O que é um excelente valor para qualquer treinamento com RNAs.

Apesar de se ter obtido os melhores resultados com o algoritmo *backpropagation* que executa o processo de treinamento da rede neural em modo *batch*, adaptado por Levenberg-Marquard e implementado no MATLAB 6.1, testou-se outras variações que também apresentaram resultados satisfatórios, dentre elas pode-se destacar um outro algoritmo que também opera no modo batch chamado de backpropagation BFG.

Para trabalhos futuros, em relação à modelagem neural de dados GPS, sugere-se que se desenvolva uma árvore de testes ainda maior que aquela executada aqui e que se utilizem outras abordagens, como a de utilizar RNAs e algoritmos genéticos para cobrir um espaço maior das topologias possíveis no sentido de indicar uma topologia ótima para esta generalização.

Ainda vislumbrando trabalhos futuros, sugere-se que se tente a implementação destes algoritmos baseados em otimização numérica dentro do sistema MORE-GPS, isto possibilitará a independência do método à utilização do programa MATLAB.

7.2 Quanto ao sistema MORE-GPS

Ainda que a idéia inicial fosse factível, não havia como testar as hipóteses aqui levantadas em função da absoluta inexistência de programas que conseguissem extrair e manipular dados de arquivos GPS com esta finalidade.

Os programas comerciais e a maioria dos programas ditos científicos, prestam-se simplesmente ao processamento das observáveis com um menor ou maior grau de interação entre o usuário e as diferentes estratégias de processamento.

O programa MORE-GPS foi desenvolvido inicialmente com a finalidade de possibilitar esta extração e manipulação de dados referentes às diferentes observáveis, as épocas nas quais foram observadas, os satélites a partir de onde foram geradas, bem como suas efemérides e seus ângulos de elevação. Contudo, ele foi se transformando em uma ferramenta fundamental sem a qual possivelmente não se chegaria a nenhuma conclusão.

Apesar de ainda carecer de melhorias que consigam otimizá-lo, tornando-o mais eficiente, completo e amigável, o sistema MORE-GPS é a compilação de programas, implementados em Visual Basic 5.0, que possibilitará uma série de diferentes investigações a partir da extração e manipulação de dados GPS.

7.2.1 Limitações do método proposto

Devido à abordagem que se deu para a solução do problema levantado, ou seja, a conduta que se utilizou na modelagem de observáveis GPS, algumas restrições ao uso do método aqui preconizado se fazem necessárias:

- Necessariamente há de se ter dados da maior quantidade possível de estações que observem simultaneamente aquele que se deseja estimar. Obviamente, dados coletados no mesmo dia e período no qual se rastreou;
- Necessidade das efemérides precisas para a data da observação. Sugere-se que em investigações posteriores se possibilite o uso das efemérides transmitidas, do inglês *broadcast*, como mais uma alternativa, no atual formato isto não é possível;
- É um método de pós-processamento, o que inviabiliza aplicações em tempo real;
- Atualmente, para vetores superiores a 230 km, pelo menos para aqueles processados nesta pesquisa, o método, não atende às

exigências de precisão posicional de 0,50 m, normatizado pelo INCRA, que estabelece as exigências e restrições para a obtenção do certificado de cadastro de imóvel rural;

- Não consegue detectar os dados gerados por receptores cujos relógios se encontram muito defasados em relação aos outros, ou seja, é possível que se tenha que realizar um pré-processamento em arquivos com essas características, visando uniformizar a parte inteira dos segundos na contagem de épocas. Essa uniformização acarretará na apresentação na correção das pseudodistâncias, deixando-as com valores mais coerentes que aqueles apresentados em arquivos com a defasagem citada.

7.3 Quanto ao processamento GPS

Não se conseguiu obter resultados satisfatórios para o período de transição diurno/noturno, mais especificamente entre 18:00:00 h e 19:00:00 h, horário local, para algumas estações observadas. Sugere-se que se desenvolva uma investigação mais profunda no sentido de determinar se esse é um fato isolado e se ocorre sistematicamente ou não em toda a área de cobertura da RBMC em diferentes épocas, ou seja, que se façam testes para dados obtido em anos com atividade solar mais baixa que a experimentada no atual ciclo solar. Ainda assim, de um modo geral, o método se mostrou eficaz na geração de observáveis GPS (L2 e P2).

7.3.1 O processamento pontual

O método desenvolvido, além de gerar estimativas de L2, gera também, estimativas para a pseudodistância P2, vislumbra-se ainda a possibilidade de melhora nestas estimativas, função de um refinamento maior na determinação da topologia neural ideal, da aplicação de transformações estatísticas mais adequadas aos dados de forma a se obter um fator de normalização menor, o que certamente degrada menos os resultados das estimativas finais.

Para se ter uma idéia da ordem de grandeza das estimativas atuais, o erro médio das estimativas comparado ao valor original das pseudodistâncias, obtidas no arquivo original para cada satélite/época está em torno de 5 metros.

Os resultados obtidos e apresentados no capítulo 6 para processamento pontual, demonstram claramente que mesmo para curtos períodos de observação, neste caso cinco minutos, se conseguem resultados para as coordenadas, muito próximos aqueles conseguidos pelo processamento do arquivo original.

O processamento pontual realizado com os dados de pseudodistância calculadas a partir do código P2 estimado pelo método proposto, conseguiu, com frequência para o conjunto de dados testado, um desvio plano e espacial menor que aqueles observados pelo processamento pontual do código P2 do arquivo original. O que demonstra a excelente capacidade de generalização da rede neural para estes dados e os qualifica para as mais diferente aplicações onde se necessite da pseudodistância a partir do código P2.

Uma outra característica observada no processamento *single point* dos dados estimados refere-se à constatação que apesar de possuir um desvio espacial muito grande em relação às coordenadas oficiais das estações observadas, o seu desvio plano é menor que aqueles apresentados para o processamento da fase L1 para bases muito longas.

7.3.2 O processamento relativo

Como já foi dito antes, não se deseja fazer nenhuma comparação entre o desempenho e as possibilidades oferecidas por diferentes programas de processamento GPS. Por este motivo, utilizou-se apenas alguns dos programas disponíveis no Departamento de Transportes-EESC/USP.

Para se ter uma idéia da ordem de grandeza das estimativas atuais, o erro médio das estimativas de contagem de ciclos da portadora L2, comparado ao valor original, obtido no arquivo original para cada satélite/época está em torno de 0,082 ciclos da portadora L2, ou seja, 2 cm. Os resultados apresentados no

Capítulo 6 indicam a influência desse erro nas coordenadas finais das estações estimadas para diferentes tipos de processamento.

Os resultados obtidos com os dados disponíveis possibilitam dizer que o método proposto tem uma maior ou menor eficiência dependendo de diversos fatores, dentre os quais pode-se destacar:

- O período de observação (tempo durante o qual se observou);
- Conduta de processamento GPS adotada (no caso da combinação L1+L2 (*ion free*) adoção da solução *fix* ou *float* para ambiguidades);
- Do comprimento da linha-base e;
- Da capacidade de generalização da rede neural para os arquivos de treinamento, ou seja, é possível que a contribuição de arquivos com muitos ruídos, fruto de multicaminhamento ou erros na propagação do sinal, caracterizados por excessiva perda de ciclo possam contribuir para um treinamneto deficiente da rede neural.

Nos testes realizados percebe-se claramente a melhora da precisão nos resultados da coordenadas obtidas para vetores de bases longas quando se usa a combinação *ion free* do arquivo modificado em comparação ao processamento apenas de L1 do arquivo original, essa melhora é maior ou menor dependendo do comprimento da linha base. Para os dados processados pode-se afirmar que vetores cujos comprimentos estão na casa dos 230km, os resultados tanto do desvio plano quanto espacial apresentam excelentes resultados, principalmente se considerando o pequeno intervalo de tempo observado (cinco minutos) e também o fato desses resultados serem baseados em estimativas e não em valores observados.

Sugere-se que se desenvolvam mais testes com dados que possibilitem gerar vetores com diferentes comprimentos de linha base no sentido de aferir a capacidade de gerar estimativas adequadas do método proposto para vetores de diferentes comprimentos e com períodos de observação homogêneos.

Propõe-se ainda que em trabalhos futuros se utilizem diferentes programas de processamento GPS, notadamente aqueles denominados de

científicos, tal como o OMNI, que segundo MADER (1998) é um pacote computacional de processamento de dados GPS, composto por mais de 50 programas, desenvolvidos pelo National Geodetic Survey (NGS). Sugere-se também a aplicação do programa PAGE-NT, disponibilizado pelo NGS. A vantagem com o uso desse programa, dá-se pelo fato dele operar em ambiente Windows 98 ou superior.

Os programas científicos possibilitam uma intervenção maior do usuário no processamento de dados GPS, o que possibilita que sejam feitas comparações mais detalhadas e se obtenham resultados baseados em soluções e parâmetros escolhidos pelo usuário e não adotados como padrão do programa.

Finalmente, conclui-se que é possível modelar o comportamento das observáveis GPS utilizando RNAs como ferramenta de modelagem a partir do algoritmo de treinamento backpropagation, principalmente quando se utiliza o processo de aprendizado supervisionado.

Um efeito positivo, mas não previsto inicialmente, observado quando da aplicação do método proposto, é a “eliminação” de perdas de ciclos para a fase L2. Como se utilizou dados oriundos de arquivos de dupla frequência, pode-se comparar o arquivo original com o arquivo modificado e em alguns casos, havia perdas de ciclos no arquivo original e que não existiam no arquivo modificado, função de terem sido geradas estimativas para todas as épocas.

Como recomendação final, sugere-se que em trabalhos futuros, se utilize o método proposto nesta investigação, repitam-se os experimentos documentados aqui, processem-se os arquivos modificados por estimativas neurais utilizando diferentes programas de processamento de dados GPS, adotando um leque maior de condutas e estratégias de processamento com o intuito de avaliar a eficácia e os limites de aplicação deste método.

Vislumbro ainda uma outra aplicação potencial deste método, que seria a utilização dos arquivos modificados para a geração de modelos regionais dedicados à correção do atraso ionosférico.

8

Referências Bibliográficas

- BACA, J.F.M. (2001). Redes Neurais Artificiais (RNA) Aplicadas na Modelagem Numérica do Terreno. XX Congresso Brasileiro de Cartografia, IX Congresso Nacional de Engenharia de Agrimensura, VIII Conferência Ibero-Americana de SIG, III Seminário Latino-Americano em Educação e Transferência de Tecnologia em Fotogrametria, SR e SIE. Porto Alegre.
- BOON, W.; SEGANTINE, P.C.L. (1999). *Testing JPL's PPP Service. GPS Solutions*, pp 73-76, Vol.3, no. 1.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. (2000), *Redes Neurais Artificiais: Teoria e Aplicações*. Livro Técnico e Científico, Rio de Janeiro.
- CAMARGO, P. O. (1999). *Modelo Regional da Ionosfera para Uso em Posicionamento com receptores GPS de uma Freqüência*. Tese – Universidade Federal do Paraná.
- CASTRO, C.M. (1977). *A prática da pesquisa*. Ed. McGraw-Hill. São Paulo.
- CYBENKO, G. (1988). *Approximation by superpositions of a sigmoidal function*. Urbana, IL: University of Illinois.
- DEMUTH H.; BEALE M. (1997). *Neural Network Toolbox For Use with MATLAB. User's Guide,Version 3.0*. MathWorks.

- DENNIS, J. E.; SCHNABEL, R. B. (1983), Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Englewood Cliffs, NJ: Prentice-Hall.
- FAZAN, J.A. (2002). Controle de Qualidade dos dados GPS da RBMC. Relatório Científico de Bolsa de Pesquisa de Iniciação Científica – FAPESP. Presidente Prudente, julho 2002.
- FLETCHER, R.; POWELL, M.J.D. (1963). *A rapidly convergent descent method for minimization*, Computer Journal, 7:149-154.
- FONSECA JUNIOR, E. S. (2002). O Sistema GPS como Ferramenta a Avaliação da Refração Ionosférica no Brasil. Tese – Escola Politécnica da Universidade de São Paulo.
- FONTANA, R.D.; CHEUNG, W., STANSELL, T. (2001). *The Modernized L2 Civil Signal. GPS World*. September.
- GOLDFARB, D., 1970. *A family of variable metric methods derived by variational means*, Mathematics of Computation, 24:23-26.
- GREGORIUS, T. (1996). Gipsy-Oasis II, *How it works*. Department of Geomatics, University of Newcastle.
- GURTNER W. (2002). RINEX: The Receiver Independent Exchange Format Version 2.10. Astronomical Institute, University of Berne.
- HAGAN, M. T.; MENHAJ, M. (1994). Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks, vol. 5, no. 6, pp.989–993.
- HAYKIN, S. (1999). *Neural networks – A Comprehensive Foundation, second edition*. Prentice Hall. New Jersey.
- HEBB, D. O. (1949). The Organization of Behavior. Wiley.
- HECHT-NIELSEN, ROBERT. (1889). *Neurocomputing*. Addison-Wesley Longman Publishing Co. Inc. Boston.

- HERNÁNDEZ-PAJARES M. et al. (1997). *Neural Network Modeling of Ionospheric Electron Content at Global Scale using GPS data. Radio Science*, v.32, n 3, p.1081-1089, june.
- HERTZ, J.; KROGH, A.; PLAMER, R.G. (1991). *Introduction to the Theory of Neural Computation. Lecture Notes, Vol I, Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley.
- HOFFMANN-WELLENHOF, B. et. al. (2001). *Global Positioning System: Theory and Practice*. Springer-Verlag Wien New York.
- HOPFIELD, J. J. (1982). Neural Networks and physical systems with emergent collective properties. *Proc. Nat. Acad. Sci.*, 79:2554-8.
- JUDGE G. G. et. al. (1980). *The Theory and Practice of Econometrics*. John Wiley & sons.
- KAPLAN, E.D.; WARD, P. (1996). *GPS Satellite Signal Characteristics. Understanding GPS Principles and Applications*. Artech House. Norwood.
- KOVÁCS, Z. L. (1996). *Redes Neurais Artificiais: Fundamentos e Aplicações*, 2ª Ed. Collegium Cognitio e Ed. Acadêmica. São Paulo.
- KOVÁCS, Z. L. (1997). *O Cérebro e a sua Mente: Uma Introdução à Neurociência Computacional*. Ed. Acadêmica. São Paulo.
- LAFORGA,G., ROMERO, R., (2002). A Service-Oriented Architecture applied to Artificial Neural Network. *Proc. ICONIP'2002*, volume 5, pp. 2650-2654, Singapore.
- LEANDRO, R.F. (2003). *A influência da ionosfera no posicionamento GPS por ponto simples*. 79 p. Dissertação – Escola de Engenharia de São Carlos da Universidade de São Paulo.
- LEICK, A.(1995). *GPS Satellite Surveying*. John Wiley & Sons, Inc. A Wiley-Interscience Publication. New York.

- LIPPMANN, R.P. (1987). *An introduction to computing with neural nets. IEEE ASSP Magazine*, vol.4, pp 4-22.
- LORA,F.A.S.; HEMERLY,E.M. e LAGES,W.F. (1998) Sistema para Navegação e Guiagem de Robôs Móveis Autônomos, *Controle & Automação*, Vol. 9, No. 3, pp. 107-118.
- LUO, F.; UNBEHAUEN R. (1998). *Applied Neural Networks for Signal Processing*. Cambridge University Press.
- MACCULLOCH, W.S.; PITTS, W. (1943). *A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics*, vol 5, pp.115-133.
- MADER, G. et. al. (1998). OMNI 4.0 – UPDATE OF USER’S GUIDE- NEW MERGE. Rockville, Maryland.
- MAIA, T.C.;SILVA, C.A.U.;ROMERO, R. A. F.; LEANDRO, R. F.;SEGANTINE, P. C. L. (2002). A Predição da Contagem de Ciclos da Portadora GPS Utilizando Modelagem Conexionista Temporal – FIR ML. VII BRAZILIAN SYMPOSIUM ON NEURAL NETWORKS.
- MASTERS, T. (1994). *Signal and Image Processing with Neural Networks – AC++ Sourcebook*. John Wiley & Sons, Inc, New York.
- MINSKY, M.; PAPERT, S. (1969). *Perceptrons: na introduction to computational geometry*. MIT Press, Massachusetts.
- MOLLER, M. F. (1993), A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, vol. 6, pp. 525-533.
- MONICO, J.F.G. (2000). Posicionamento pelo NAVSTAR-GPS – Descrição, fundamentos e aplicações. Editora UNESP. São Paulo.
- MONICO, J.F.G. (2001). O Posicionamento Pontual Preciso – Descrição, Resultados e Aplicações. Congresso Brasileiro de Cartografia, XX. CT2, Anais. Porto Alegre.

- PANDEY, P. et al. (2001) *Alternate Neural Network Tools for Pattern Classification in Satellite Imageries*. National Environmental Engineering Research Institute, Nehru Marg Nagpur India.
<http://www.gisdevelopment.net/technology/ip/techip0001.htm>. 15/05/2002.
- PARKINSON, B. W. (1996). *Global Positioning System: Theory and Applications*, V. I, V. II. American Institute of Aeronautics and Astronautics, Washington.
- ROSENBLATT, F. (1958). *The perceptron: A probabilistic model for information storage and organization in the brain*. Psychological Review, vol. 65, pp.386-408.
- RUMELHART, D.E. et. al. (1986). *Learning representations by back-propagating errors*. Nature, 323:533-536.
- SARLE, W.S. (1994). *Neural Networks and Statistical Models*. Proceedings of the Nineteenth Annual SAS User Group International Conference. Cary,NC, USA.
- SEEBER, G. (1993). *Satellite Geodesy*. Walter de Gruyter & Co. Berlin.
- SEGANTINE, P.C.L. (1995). Estabelecimento e ajuste de uma rede geodésica no estado de São Paulo, com o Sistema de Posicionamento NAVSTAR/GPS. 222p. Tese - Escola Politécnica da Universidade de São Paulo.
- SEGANTINE, P.C.L. (1999). Sistema de Posicionamento Global. Apostila. São Carlos, SP, publicação EESC-USP.
- SHANNO, D.F., 1970. *Conditioning of quasi-Newton methods for function minimization*, Mathematics of Computation, 24:145-160.
- SILVA, C.A.U. (2000). Análise de Precisão dos Posicionamentos Relativo e por Ponto Simples para Bases Curtas, com o Sistema NAVSTAR/GPS: Estudo de Caso em Belém-Pa. 190 p. Dissertação – Escola de Engenharia de São Carlos da Universidade de São Paulo.

- SILVA, C.A.U.; SEGANTINE, P.C.L. (2001). O Posicionamento Pontual Preciso – PPP. : Congresso Brasileiro de Cartografia, XX. CT2, Anais. Porto Alegre.
- STONE, M. (1974). *Cross-validatory choice and assessment of statistical predictions*. Journal of the Royal Statistical Society, vol. B36, pp.111-113.
- STONE, M. (1978). Cross-validation: A Review, *Mathematische Operationsforschung Statistischen*, serie Statistics, vol. 9, pp.127-139.
- TIERRA, A.R.; FREITAS, S.R.C (2000). Predicting Free-Air Gravity Anomaly Using Artificial Neural Network.
- WAN .E.A. (1994). *Time Series Prediction by Using a Connectionist Network with Internal Delay Lines*. SFI Studies in the Sciences of Complexity. v XVII, Addison-Wesley.
- WEBB, F. H.; ZUMBERGE, J. F. (1997). An Introduction to GIPSY/OASIS-II – Precision Software for the Analysis of data from the Global Posotioning System. Jet Propulsion Laboratory. Pasadena.
- WELLS, D. et al. (1987). *Guide to GPS Positioning*. Canadian GPS Associates. Fredericton.
- WIDROW, B.; HOFF. M.E. (1960). *Adaptative swithing circuits*. Institute of Radio Engineers, Western Electronic Show and Convention.
- ZUMBERGE, J.F. (1998). *Characteristics and Applications of Precise GPS Clock Solutions Every 30 Seconds*. Navigation: Journal of the Institute of Navigation. V 44, No 4, p 449-456.
- ZUMBERGE, J.F. (1999). *Automated GPS Data Analisis Service*. GPS Solutions, v 2, No 3, p 76-78.
- ZUMBERGE, J.F. et al.(1997).*Precise point positioning for the efficient and robust analysis of GPS data from large networks*. Journal Geophysical Research.