



DANIELE BARROCA MARRA ALVES

**MÉTODO DOS MÍNIMOS QUADRADOS
COM PENALIDADES: APLICAÇÃO NO
POSICIONAMENTO RELATIVO GPS**

DANIELE BARROCA MARRA ALVES

**MÉTODO DOS MÍNIMOS QUADRADOS COM
PENALIDADES: APLICAÇÃO NO
POSICIONAMENTO RELATIVO GPS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Cartográficas da Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista, para obtenção do título de Mestre em Ciências Cartográficas.

Orientador: João Francisco Galera Monico

Co-orientador: Messias Meneguette Jr.

A478m

Alves, Daniele Barroca Marra

Método dos mínimos quadrados com penalidades: aplicação no posicionamento relativo GPS / Daniele Barroca Marra Alves. – Presidente Prudente : [s.n.], 2004.

131 f. : il. ; 29 cm.

Dissertação (Mestrado). - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia.

Orientador: João Francisco Galera Monico

Co-orientador: Messias Meneguette Junior

1. Posicionamento Geodésico. 2. GPS. 3. Erros Sistemáticos. I. Alves, Daniele Barroca Marra. II. Título.

CDD (18.ed.) 623.71

DADOS CURRICULARES

Daniele Barroca Marra Alves

Nascimento 29/12/1980 - Presidente Prudente - SP

Filiação Iderval Rojas Marra
Neuza Barroca Marra

1998-2001 Curso de Graduação
Licenciatura em Matemática
Faculdade de Ciências e Tecnologia - UNESP

2002-2004 Curso de Pós-Graduação
Mestrado em Ciências Cartográficas
Faculdade de Ciências e Tecnologia - UNESP

Ao meu amado marido, Alessandro da Rocha Alves, pelo grande apoio e força nos momentos difíceis.

Aos meus maravilhosos pais, Iderval Rojas Marra e Neuza Barroca Marra, que sempre foram o alicerce da minha vida, me ajudando e incentivando em todas as situações.

Às minhas queridas irmãs, Soellyn, Aline e Luana, pela amizade e carinho que sempre tiveram por mim.

AGRADECIMENTOS

Desejo externar os meus sinceros agradecimentos a todos que colaboraram com o desenvolvimento desse trabalho, em especial:

A Deus, que sempre me guiou pelos melhores caminhos, pois sem sua ajuda não conseguiria seguir em frente para alcançar meus objetivos.

Ao professor Dr. João Francisco Galera Monico, meu orientador, pela confiança e contribuição no desenvolvimento dessa pesquisa. Ao professor Dr. Messias Meneguette Jr., meu co-orientador, pelo auxílio e apoio.

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP, pelo auxílio financeiro nessa pesquisa, sob forma de bolsa de demanda social.

A todos os professores do PPGCC, do Departamento de Cartografia e do Departamento de Matemática, que tiveram grande contribuição na minha formação acadêmica. Em especial, ao prof. Dr. José Roberto Nogueira que me ajudou no estudo inicial das splines.

Ao professor Dr. Mike Stewart, da Curtin University, por ter fornecido textos que auxiliaram no embasamento teórico deste trabalho.

Ao Dr. Luiz Paulo Souto Fortes, pelas sugestões dadas no exame de qualificação.

Aos funcionários da universidade que conviveram comigo, em especial, ao Sr. Geraldo e Ítalo, aos funcionários Edmilson e Pedro do Departamento de Matemática, à secretária Graça do Departamento de Cartografia, aos funcionários Márcia, Erinate, Ivonete e Washington da seção de Pós-Graduação, à bibliotecária Fátima, ao motorista Pedro e à Dona Cida da limpeza, que sempre foram muito atenciosos e contribuíram de forma indireta com o desenvolvimento dessa pesquisa.

À minha grande amiga Eniuce, que compartilhou todos os momentos bons e ruins dos últimos nove anos ao meu lado, pela força, garra e grande amizade.

A todos os amigos do PPGCC, em especial a Edinéia, Chris, Giovane, Marcelo, Tadashi e Sapucci, pelo companheirismo, incentivo e ajuda.

Aos familiares e amigos que sempre me incentivaram e acreditaram em mim.

RESUMO

O *Global Navigation Satellite System* (GNSS), que congrega os vários sistemas de posicionamento por satélite existentes, tem como principal objetivo viabilizar o posicionamento de baixa, média e alta precisão. Dentre os sistemas de posicionamento que integram o GNSS, o *Global Positioning System* (GPS) tem grande destaque. Mas as observáveis GPS, tal como todas as outras observáveis envolvidas nos processos de medidas, estão sujeitas a erros aleatórios, sistemáticos e grosseiros. Os erros aleatórios são inevitáveis, sendo, portanto, considerados uma propriedade inerente das observações. Erros grosseiros (outliers) devem ser eliminados através do processo de controle de qualidade. Erros sistemáticos podem ser parametrizados ou eliminados por técnicas apropriadas de observação. Eles degradam a acurácia do posicionamento realizado com o GPS. Esses erros incluem erros da órbita dos satélites GPS, multicaminho, erros de refração atmosférica, dentre outros. Dessa forma, alguns trabalhos recentes têm utilizado o modelo semiparamétrico e o método dos mínimos quadrados com penalidades (MMQ com penalidades) para atenuar os efeitos desses erros residuais, utilizando dados de receptores de monofrequência. No modelo semiparamétrico as variáveis estimadas são divididas em uma parte paramétrica (coordenadas da estação e ambigüidades), que é de interesse do usuário, e uma parte não-paramétrica (funções de erros que variam suavemente com o tempo). Assim, devido ao número de incógnitas ser maior que o usual, é utilizado o MMQ com penalidades. Essa técnica utiliza uma spline cúbica natural, cuja suavidade é determinada pelo parâmetro suavizador, calculado pela validação cruzada generalizada. Nesse método, os erros são modelados como funções que variam suavemente com o tempo. E mais, as funções de erros sistemáticos, ambigüidades e coordenadas de interesse são estimadas simultaneamente. Como resultado, as ambigüidades e as coordenadas de interesse são estimadas com melhor confiança e acurácia do que com o MMQ convencional. Além disso, a solução requer um menor período de coleta de dados, minimizando custos. Nesse trabalho foi implementado o MMQ com Penalidades juntamente com o modelo semiparamétrico, com o objetivo de atenuar os erros sistemáticos no posicionamento relativo GPS. A revisão teórica, resultados e análises são apresentados nessa dissertação. Para analisar a performance do método, foram realizados dois experimentos. O primeiro foi realizado com uma linha de base curta, no qual o principal erro envolvido era o multicaminho. Já no segundo experimento, foram utilizadas quatro linhas de base, com comprimento variando de 15 a 120 km, onde os erros predominantes eram a refração ionosférica, troposférica e o erro das órbitas dos satélites GPS. No primeiro experimento, utilizando 5 min de dados, as discrepâncias máximas das coordenadas em relação aos valores verdadeiros chegaram a 1,6 cm e 3,3 cm em h para o MMQ com Penalidades e o MMQ convencional recursivo, respectivamente. Já no segundo experimento, com 5 min de dados, as discrepâncias máximas foram de 56,6 cm em N para o MMQ com Penalidades e 3,45 m em h para o MMQ convencional, para a linha de base mais longa. Em todos os testes realizados, foi possível verificar uma considerável melhora na acurácia das coordenadas e na solução das ambigüidades utilizando o MMQ com Penalidades em relação ao MMQ convencional, com um reduzido intervalo de tempo de coleta de dados.

Palavras-chave: Modelo Semiparamétrico; MMQ com Penalidades; Spline Cúbica Natural; Parâmetro Suavizador; Validação Cruzada Generalizada; Erros sistemáticos.

ABSTRACT

The Global Navigation Satellite System (GNSS), that encompasses several satellite positioning systems, has as main goal to make available the low, medium and high precision positioning. Among the positioning systems that integrate GNSS, the Global Positioning System (GPS) has a great importance. But the GPS observables, like all other observables involved in a measurement process, are subject to random, systematic and outliers errors. The random errors are inevitable, being, therefore, considered an inherent property of the observations. Outliers should be eliminated through the quality control process. Systematic errors can be modeled or eliminated by appropriate observation techniques. The systematic errors degrade the accuracy of the positioning accomplished by GPS. These errors are those related to GPS satellites orbits, multipath, atmospheric refraction among others. Thus, some authors have been using the semiparametric model and the penalised least squares technique to mitigate these residual errors, using single frequency receiver data. In a semiparametric model the estimated variables are divided into a parametric part (station coordinates and ambiguities), which is of interest to the users, and a nonparametric one (composed by error functions that vary smoothly with time). However, due to the unknowns number being larger than the usual, the penalised least squares is used. This technique uses a natural cubic spline, whose smoothness is determined by a smoothing parameter, computed by using the generalized cross validation. In this method, the errors are modeled as functions which vary smoothly in time. And more, the systematic errors functions, ambiguities and station coordinates are estimated simultaneously. As a result, the ambiguities and the station coordinates are estimated with better reliability and accuracy than the conventional least square method. Therefore, the solution requests a shorter data collection interval, minimizing costs. In this work, the penalised least squares was implemented with the semiparametric model, with the objective of mitigating systematic errors involved in GPS relative positioning. The theoretical revision, results and analyses are presented in this dissertation. So, to analyze the method performance, two experiments were carried out. The first one was accomplished with a short baseline, where the main error was the multipath. In the second experiment, four baselines from 15 to 120 km were used. In this case, the predominant errors were due to the ionosphere and troposphere refraction and GPS satellites orbits. In the first experiment, using 5 minutes of data collection, the largest coordinates discrepancies in relation to the true values reached 1.6 cm and 3.3 cm in the h coordinate for Penalised Least Squares and the conventional recursive Least Squares, respectively. In the second one, also using 5 minutes of data, the discrepancies were 56.6 cm in N for the Penalised Least Squares and 3.45 m in h for the conventional Least Squares, for the longer baseline. In all accomplished tests, it was possible to verify a considerable improvement in the coordinates accuracy and in the ambiguities resolution using the Penalised Least Squares in relation to the conventional Least Squares, with a reduced data collection time interval.

Keywords: Semiparametric Model; Penalised Least Squares; Natural Cubic Spline; Smoothing Parameter; Generalized Cross Validation; Systematic Errors.

LISTA DE FIGURAS

FIGURA 01 - Simples diferença.....	19
FIGURA 02 - Dupla diferença.....	21
FIGURA 03 - Camadas da Ionosfera.....	31
FIGURA 04 - Variação de longo período da Ionosfera para os últimos cinco ciclos.....	35
FIGURA 05 - Comportamento da refratividade R em função da altura; para a troposfera (R_t) e ionosfera (R_i)	36
FIGURA 06 - Efeito do multicaminho.....	39
FIGURA 07 - Conjunto de dados unidos por linhas retas	44
FIGURA 08 - Dados interpolados por uma curva com segunda derivada contínua.....	45
FIGURA 09 - Curva que minimiza $S(g)$ para α dado pela validação cruzada generalizada	47
FIGURA 10 - Curva que minimiza $S(g)$ para um valor grande de α	47
FIGURA 11 - Curva que minimiza $S(g)$ para um valor pequeno de α	48
FIGURA 12 - Spline Mecânica moldada por presilhas	49
FIGURA 13 - Spline suavizadora para a DD^{2-15} do código	74
FIGURA 14 - Spline suavizadora para a DD^{2-31} do código	75
FIGURA 15 - Etapas utilizadas na implementação do algoritmo do MMQ com Penalidades... 87	
FIGURA 16 - Receptor Trimble 4600 LS utilizado no levantamento	89
FIGURA 17- Antena e receptor Ashtech ZXII utilizados no levantamento	89
FIGURA 18 - Estação UEPP que foi adotada como estação base	90
FIGURA 19 - Coleta de dados com superfície refletora realizada na empresa Takigawa no Município de Presidente Prudente.....	90
FIGURA 20 - Municípios onde foram realizados os experimentos.....	91
FIGURA 21 - Coleta de dados realizada no sítio (Santo Antônio) em Regente.....	92
FIGURA 22 - Coleta de dados realizada em Venceslau	92
FIGURA 23 - Coleta de dados realizada na fazenda São Pedro - Quintana	93
FIGURA 24 - Coleta de dados realizada no Campus da UNESP na cidade de Assis	93
FIGURA 25 - Ângulo de elevação dos satélites da coleta realizada na empresa Takigawa.....	95
FIGURA 26 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Takigawa com intervalo de 5 min de dados	96
FIGURA 27 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Takigawa com intervalo de 10 min de dados	96
FIGURA 28 - Teste <i>ratio</i> para o experimento UEPP-Takigawa com intervalo de 5 min e 10 min de dados	97
FIGURA 29 - Ângulo de elevação dos satélites da coleta realizada em Regente.....	101
FIGURA 30 - Ângulo de elevação dos satélites da coleta realizada em Venceslau	101
FIGURA 31 - Ângulo de elevação dos satélites da coleta realizada em Quintana	101
FIGURA 32 - Ângulo de elevação dos satélites da coleta realizada em Assis	102
FIGURA 33 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Regente (≈ 18 km) com intervalo de 5 min de dados.....	103
FIGURA 34 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Venceslau (≈ 52 km) com intervalo de 5 min de dados	103
FIGURA 35 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Quintana (≈ 102 km) com intervalo de 5 min de dados	104
FIGURA 36 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Assis (≈ 116 km) com intervalo de 5 min de dados coletados	104

FIGURA 37 - Teste <i>ratio</i> para os experimentos com intervalo de 5 min de dados coletados por volta das 14 h local	105
FIGURA 38 - Teste <i>ratio</i> para os experimentos com intervalo de 5 min de dados coletados durante a noite	106
FIGURA 39 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Regente (≈ 18 km) com intervalo de 10 min de dados....	108
FIGURA 40 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Venceslau (≈ 52 km) com intervalo de 10 min de dados ...	108
FIGURA 41 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Quintana (≈ 102 km) com intervalo de 10 min de dados ..	109
FIGURA 42 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Assis (≈ 116 km) com intervalo de 10 min de dados	109
FIGURA 43 - Teste <i>ratio</i> para os experimentos com intervalo de 10 min de dados coletados por volta das 14 h local	110

LISTA DE TABELAS

TABELA 01 - Fontes e efeitos dos erros envolvidos no GPS	29
TABELA 02 - Efeito do atraso de propagação, devido à ionosfera sobre as distâncias medidas com observações de uma frequência, e erros residuais para observações de dupla frequência	36
TABELA 03 - Produtos IGS	41
TABELA 04 - Arquivo de entrada do software GPSeq	82
TABELA 05 - Solução das ambigüidades GPS para o experimento UEPP-Takigawa com intervalo de 5 min de dados.....	97
TABELA 06 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Takigawa com intervalo de 5 min de dados	98
TABELA 07 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Takigawa com intervalo de 10 min de dados	98
TABELA 08 - Intervalo de tempo de processamento nas estações dos municípios de Regente, Venceslau, Quintana e Assis (hora local).....	100
TABELA 09 - Média absoluta, média aritmética e desvio padrão dos resíduos para os experimentos com intervalo de 5 min de dados	106
TABELA 10 - Intervalo de tempo de processamento nas estações dos municípios de Regente, Venceslau, Quintana e Assis (hora local).....	107
TABELA 11 - Média absoluta, média aritmética e desvio padrão dos resíduos para os experimentos com intervalo de 10 min de dados	111
TABELA 12 - Demanda computacional para 5 min de dados com o MMQ com Penalidades	112
TABELA 13 - Demanda computacional para 10 min de dados com o MMQ com Penalidades	112
TABELA 14 - Demanda computacional para 5 min de dados com o MMQ convencional.....	113
TABELA 15 - Demanda computacional para 10 min de dados com o MMQ convencional....	113

SUMÁRIO

1 INTRODUÇÃO	13
1.1 Objetivos	14
1.2 Justificativa	15
1.3 Conteúdo do relatório	15
2 POSICIONAMENTO RELATIVO COM GPS	17
2.1 Pseudodistância	17
2.2 Fase de batimento da onda portadora	18
2.3 Simples diferença	19
2.4 Dupla diferença	21
2.4.1 Matriz variância covariância	22
2.4.2 Variância a posteriori	24
2.5 Posicionamento relativo	25
2.5.1 Posicionamento relativo estático	26
2.5.2 Posicionamento relativo estático rápido	27
3 ERROS ENVOLVIDOS NAS OBSERVÁVEIS GPS	29
3.1 Ionosfera	30
3.1.1 Características da ionosfera	31
3.1.2 Efeitos da ionosfera nos sinais GPS	33
3.2 Troposfera	38
3.3 Multicaminho	39
3.4 Erros na órbita dos satélites GPS	40
4 MÍNIMOS QUADRADOS COM PENALIDADES E MODELO SEMIPARAMÉTRICO	42
4.1 Aproximações por regressão	42
4.1.1 Regressão linear	42
4.1.2 Regressão polinomial	43
4.2 Penalidade de aspereza	44
4.2.1 O ajuste da curva	44
4.2.2 Quantificando a aspereza de uma curva	45
4.2.3 Regressão por Mínimos Quadrados Penalizados	46
4.3 Splines	48
4.3.1 Origem	49
4.3.2 Definição de spline	50
4.3.3 Splines cúbicas naturais	51
4.3.3.1 Representação do valor da segunda derivada	52
4.3.4 Interpolando com splines	54
4.3.4.1 Construindo a spline cúbica natural interpolante	55
4.3.4.2 Propriedades de uma spline cúbica natural interpolante	56
4.3.5 Suavizando com splines	57
4.3.5.1 O algoritmo de Reinsch	58
4.3.6 Suavização com splines de forma ponderada	59
4.3.6.1 Propriedades básicas da formulação ponderada	60
4.3.6.2 O algoritmo de Reinsch para suavização ponderada	61
4.4 Modelo semiparamétrico	62
4.4.1 Matriz de incidência	63
4.5 MMQ com Penalidades para o modelo semiparamétrico	64
5 PARÂMETRO SUAVIZADOR	67
5.1 Introdução	67
5.2 Validação cruzada	68
5.2.1 Cálculo da função de validação cruzada	69
5.2.2 Encontrando os elementos da diagonal de $H(\alpha)$	70
5.2.2.1 As diagonais centrais da inversa de uma matriz em banda	71
5.2.2.2 Matriz $H(\alpha)$ otimizada	72
5.3 Validação Cruzada Generalizada	73

5.3.1 Exemplo da utilização da GCV	74
5.4 Validação cruzada e validação cruzada generalizada para suavização ponderada	75
5.5 Validação cruzada generalizada no modelo semiparamétrico	76
6 SOFTWARES DISPONÍVEIS E IMPLEMENTAÇÃO COMPUTACIONAL	78
6.1 O software GAS	78
6.2 O Software GPSeq	80
6.2.1 Dados de entrada	82
6.2.2 Modelos matemáticos e estratégias adotadas	82
6.2.2.1 Modelo funcional e estocástico	83
6.2.2.2 Cálculo das coordenadas dos satélites	83
6.2.2.3 Estimativa recursiva e controle de qualidade	84
6.2.2.4 Solução das ambigüidades e validação	85
6.3 Implementação do MMQ com Penalidades e do modelo semiparamétrico no GPSeq	86
7 COLETA DE DADOS, RESULTADOS E ANÁLISES	89
7.1 Coleta de dados	89
7.2 Resultados e análises dos experimentos	94
7.2.1 Resultados e análises de experimento com presença de multicaminho	94
7.2.2 Resultados e análises de experimentos realizados com linhas de base variando de 18 a 120 km	99
7.2.2.1 Resultados e análises de experimentos com intervalos de 5 min de dados	99
7.2.2.2 Resultados e análises de experimentos com intervalos de 10 min de dados	107
7.3 Eficiência computacional	111
8 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES	114
9 REFERÊNCIAS BIBLIOGRÁFICAS	117
10 BIBLIOGRAFIA	122
APÊNDICE A – Teorema que determina o mínimo de uma spline cúbica natural	124
APÊNDICE B – Teorema fundamental da GCV	126
APÊNDICE C – Tabelas dos resíduos para cada DD	128

1 INTRODUÇÃO

As observáveis básicas do *Global Navigation Satellite System* (GNSS), em especial o GPS, que permitem determinar posição, velocidade e tempo, tal como todas as outras observáveis envolvidas nos processos de medidas, estão sujeitas a erros aleatórios, sistemáticos e grosseiros. Para obter resultados confiáveis, o modelo matemático estabelecido deve ser válido para a realidade física que se tenta descrever e ser capaz de detectar problemas nas observáveis utilizadas. Dessa forma, todos os tipos de erros envolvidos no processo de medida devem ser considerados. Dentre eles, os erros aleatórios são inevitáveis, sendo, portanto, considerados uma propriedade inerente das observações (GEMAEL, 1994).

Erros grosseiros (outliers) devem ser eliminados através do processo de controle de qualidade. Um procedimento extensivamente utilizado no campo da Geodésia é denominado *Detection, Identification and Adaptation* (DIA) (TEUNISSEN, 1998b).

Erros sistemáticos podem ser parametrizados, isto é, modelados como parâmetros adicionais, ou eliminados por técnicas apropriadas de observação e/ou processamento. Quando se utiliza o GPS, as principais fontes de erros sistemáticos no posicionamento relativo de alta precisão, envolvendo linhas de base de média e longa distância, são a refração troposférica e a refração ionosférica (MONICO, 2000; HOFMAN-WELLENHOF, 1997). Esses erros podem não somente impedir uma confiável resolução de ambigüidades, como também degradar a acurácia dos resultados.

Para atenuar os efeitos desses erros e para melhorar a confiabilidade da resolução das ambigüidades e da estimativa das coordenadas de interesse, alguns estudos têm sido iniciados no sentido de se aplicar o método dos Mínimos Quadrados com Penalidades (MMQ com penalidades) (JIA, TSAKIRI e STEWART, 2001), dentro do contexto do modelo semiparamétrico, usando uma spline cúbica natural (GREEN e SILVERMAN, 1994).

No modelo semiparamétrico os erros são modelados como funções que variam suavemente com o tempo. E mais, as funções de erros sistemáticos, ambigüidades e coordenadas de interesse são estimadas simultaneamente. Como resultado, a ambigüidade e as coordenadas de interesse são estimadas com melhor confiança do que com o MMQ convencional. Além disso, a solução requer um menor intervalo de tempo de coleta de dados.

No entanto, estudos dessa natureza não têm sido realizados no território brasileiro, onde a influência da ionosfera é significativa. Além disso, tal funcionalidade não está disponível em software de uso comum, o que requer implementação para a investigação desejada.

1.1 Objetivos

O objetivo principal deste trabalho é investigar, adaptar e implementar um algoritmo que reduza o intervalo de tempo necessário para a solução das ambigüidades GPS no posicionamento relativo com presença de erros sistemáticos, sem perda significativa da qualidade dos resultados. Para atingir esse objetivo pretende-se:

- Apresentar o ajustamento com penalidades visando atenuar erros sistemáticos no posicionamento relativo GPS, utilizando dados de receptores de simples frequência;
- Verificar se no MMQ com penalidades os efeitos sistemáticos dos resíduos são atenuados e a acurácia das coordenadas ajustadas é melhorada em relação ao MMQ convencional;
- Contribuir com o desenvolvimento científico e tecnológico nacional.

1.2 Justificativa

O projeto de pesquisa realizado trata-se de uma inovação, o que vai de encontro com os objetivos do Programa de Pós-Graduação em Ciências Cartográficas, onde foi desenvolvido como uma dissertação de Mestrado, e foge à simples utilização de um software do tipo *Black Box*.

O algoritmo foi implementado no software GPSeq que está em desenvolvimento na FCT/UNESP. Como os resultados foram promissores, o produto final poderá ser repassado para a indústria, contemplando a atual política do Ministério de Ciências e Tecnologia, bem como a da própria FAPESP.

Além disso, cabe acrescentar que, com a implementação do algoritmo, o intervalo de tempo de coleta de dados pode ser reduzido, minimizando custos de trabalho de campo.

Adicionalmente pode-se acrescentar que essa metodologia, embora já tenha mostrado resultados promissores, não foi testada considerando a realidade brasileira, onde os efeitos causados pela refração ionosférica afetam consideravelmente o posicionamento com o GPS.

1.3 Conteúdo do relatório

A organização deste trabalho é descrita a seguir, com o conteúdo de cada capítulo.

No capítulo 2 é descrito o posicionamento relativo, em especial o posicionamento relativo estático, que é utilizado no desenvolvimento dessa pesquisa. Para tal,

foi introduzido o conceito de simples e dupla diferença e suas respectivas matrizes variância covariância, visto que as duplas diferenças são as observáveis fundamentais do posicionamento relativo GPS.

O capítulo 3 é dedicado às fontes de erros envolvidas nas observáveis GPS, em especial a refração ionosférica, troposférica, multicaminho e o erro nas órbitas dos satélites GPS, pois esse trabalho tem como objetivo atenuar o efeito dos erros sistemáticos no posicionamento relativo GPS.

Como nessa pesquisa é aplicado o MMQ com penalidades juntamente com o modelo semiparamétrico, no capítulo 4 os conceitos teóricos desse método são descritos. Além disso, este capítulo trata das splines, bem como da interpolação e suavização da spline cúbica natural, pois o MMQ com penalidades utiliza tal função.

O capítulo 5 descreve um método que otimiza a escolha do parâmetro suavizador, bem como suas propriedades.

O capítulo 6 apresenta alguns conceitos fundamentais dos softwares utilizados e da implementação. Já o capítulo 7 descreve a coleta de dados, os experimentos e a análise dos resultados.

Finalmente, o capítulo 8 é dedicado as considerações finais e as recomendações dessa pesquisa.

2 POSICIONAMENTO RELATIVO COM GPS

Para se realizar o posicionamento relativo com o GPS utilizam-se, em geral, as duplas diferenças (DD), as quais são formadas a partir das simples diferenças (SD) (MONICO, 2000, p.205). Para obter as SD e as DD pode-se utilizar as observáveis pseudodistância e fase da onda portadora.

2.1 Pseudodistância

A medida de pseudodistância é obtida a partir da correlação entre o código gerado pelo satélite no instante de transmissão (t^s) e sua réplica gerada no receptor no instante de recepção (t_r). A equação da pseudodistância entre o satélite s e o receptor r pode ser escrita como (SEEBER, 2003, p.255):

$$PD_r^s = \rho_r^s + c(dt_r - dt^s) + I_r^s + T_r^s + dm_r^s + d\rho_r^s + \varepsilon_{PD_r^s}, \quad (2.1)$$

onde:

- ρ_r^s é a distância geométrica entre o satélite s no instante de transmissão do sinal e o receptor r no instante de recepção, em metros;
- c é a velocidade da luz no vácuo, em m/s;
- dt_r é o erro do relógio do receptor em relação ao tempo GPS no instante de recepção, em segundos;
- dt^s é o erro do relógio do satélite em relação ao tempo GPS no instante de transmissão, em segundos;
- I_r^s é o erro causado pela refração ionosférica, em metros;

- T_r^s é o erro causado pela refração troposférica, em metros;
- dm_r^s é o erro causado pelo multicaminho, em metros;
- $d\rho_r^s$ é o erro causado pela órbita do satélite, em metros;
- $\varepsilon_{PD_r^s}$ é o erro da pseudodistância devido aos efeitos não modelados e aleatórios, em metros.

2.2 Fase de batimento da onda portadora

A medida da fase de batimento da onda portadora é obtida a partir da diferença entre a fase gerada pelo satélite, no instante de transmissão do sinal, e sua réplica gerada pelo receptor, no instante de recepção do sinal. Apenas uma medida fracionária é obtida, restando um número inteiro de ciclos desconhecido, denominado ambigüidade (N). A equação da fase de batimento da onda portadora pode ser escrita como (SEEBER, 2003, p.255):

$$\phi_r^s(t) = f \left(\frac{\rho_r^s - I_r^s + T_r^s + dm_r^s + d\rho_r^s}{c} \right) + f(dt_r - dt^s) + (\phi_t^s(t_0) - \phi_r(t_0)) + N_r^s + \varepsilon_{\phi_r^s}, \quad (2.2)$$

onde:

- f a frequência nominal da fase, em Hz;
- $\phi_t^s(t_0)$ é a fase inicial no satélite, correspondente à época de referência t_0 , em ciclos;
- $\phi_r(t_0)$ é a fase inicial no receptor, correspondente à época de referência t_0 , em ciclos;
- N_r^s é a ambigüidade da fase no instante inicial de rastreo, em ciclos;
- $\varepsilon_{\phi_r^s}$ é o erro da fase da portadora devido aos efeitos não modelados e aleatórios, em metros.

Os demais termos são os mesmos da equação 2.1.

2.3 Simples diferença

Simple differences of GPS observables can be formed between two receivers, two satellites, or two epochs. Usual combinations involve differences between satellites and stations (MONICO, 2000, p.172). The SD between two receivers is illustrated in figure 01. The fundamental idea is that the two receivers (r_1 and r_2) track simultaneously the same satellite.

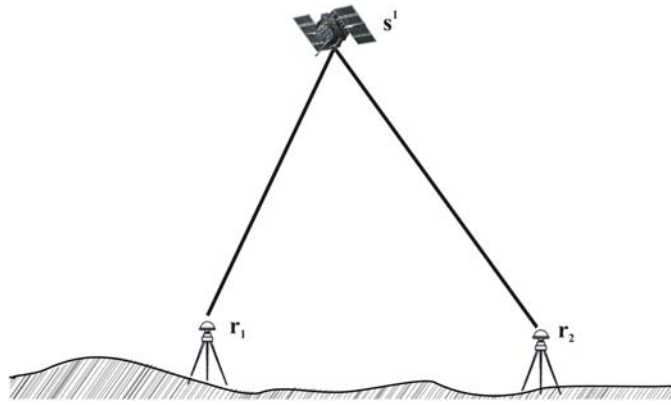


FIGURA 01 - Simples diferença

The difference between the pseudoranges observed simultaneously in two stations is the SD of the pseudorange. The observation equation is given by (MONICO, 2000, p.173):

$$PD_{1,2}^1 + v_{PD_{SD}} = \rho_{1,2}^1 + c(dt_1 - dt_2) + I_{1,2}^1 + T_{1,2}^1 + dm_{1,2}^1, \quad (2.3)$$

onde:

- $v_{PD_{SD}}$ é o resíduo da SD da pseudodistância, em metros;
- $\rho_{1,2}^1 = \rho_1^1 - \rho_2^1$, sendo ρ a distância geométrica entre o satélite no instante de transmissão do sinal e o receptor no instante de recepção, em metros;
- dt_1 e dt_2 são os erros dos relógios dos receptores, em segundos;
- $I_{1,2}^1 = I_1^1 - I_2^1$, sendo I o erro causado pela refração ionosférica, em metros;

- $T_{1,2}^1 = T_1^1 - T_2^1$, sendo T o erro causado pela refração troposférica, em metros;
- $dm_{1,2}^1 = dm_1^1 - dm_2^1$, sendo dm o erro causado pelo multicaminho, em metros;

Com as mesmas considerações expostas, a SD da fase da onda portadora é expressa por (MONICO, 2000, p.173):

$$\phi_{1,2}^1 + \nu_{\phi_{SD}} = f \left(\frac{\rho_{1,2}^1 - I_{1,2}^1 + T_{1,2}^1 + dm_{1,2}^1}{c} \right) + f(dt_1 - dt_2) + \phi_{1,2}(t_0) + N_{1,2}^1, \quad (2.4)$$

sendo:

- $\phi_{1,2}(t_0) = \phi_1(t_0) - \phi_2(t_0)$, em ciclos;
- $N_{1,2}^1 = N_1^1 - N_2^1$, em ciclos.

A principal vantagem da SD é que a maioria dos erros comuns do satélite é cancelada, como, por exemplo, o erro do relógio do satélite (dt^s), a fase inicial no satélite correspondente à época t_0 e o atraso causado no hardware do satélite (LEICK, 1995, p.260).

Os erros causados pelo satélite existem porque, embora o tempo nominal de recepção do sinal seja o mesmo, o tempo de emissão difere um pouco, por causa das diferentes distâncias entre o satélite e as estações r_1 e r_2 . No entanto, as SD podem ser consideradas livres do efeito do erro do relógio do satélite. Mas são sensíveis ao erro do relógio do receptor (LEICK, 1995, p.261).

Além disso, erros devidos às posições dos satélites e à refração atmosférica são minimizados nas SD, especialmente em bases curtas, onde os efeitos da ionosfera e da troposfera são similares em cada estação. Para bases longas, a refração troposférica pode ser modelada e a ionosférica pode ser reduzida pelo uso da combinação linear *Ion-Free*, caso em que se necessita de um receptor de dupla frequência. Uma opção seria simplesmente ignorar tais efeitos, o que deterioraria os resultados. Os erros não modelados ou não totalmente eliminados são assumidos como de natureza aleatória, fazendo parte do resíduo da observação em questão (MONICO, 2000, p.173).

2.4 Dupla diferença

A dupla diferença é a diferença entre duas SD. Envolve, portanto, dois receptores e dois satélites (MONICO, 2000, p.173). Uma ilustração é mostrada na figura 02.

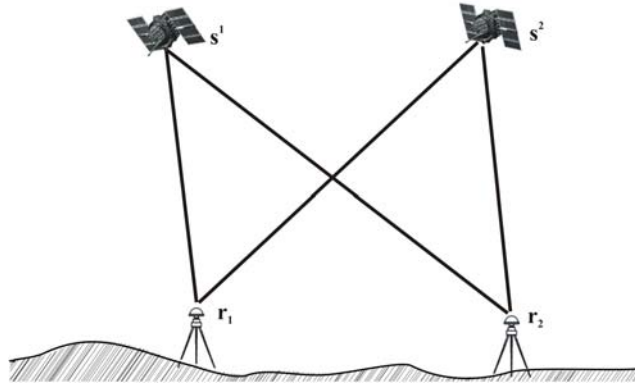


FIGURA 02 - Dupla diferença

A equação da DD para a pseudodistância é dada por (MONICO, 2000, p.174):

$$PD_{1,2}^{1,2} + v_{PD_{DD}} = \rho_{1,2}^{1,2} + I_{1,2}^{1,2} + T_{1,2}^{1,2} + dm_{1,2}^{1,2}, \quad (2.5)$$

onde $\rho_{1,2}^{1,2} = \rho_{1,2}^1 - \rho_{1,2}^2$, $I_{1,2}^{1,2} = I_{1,2}^1 - I_{1,2}^2$, $T_{1,2}^{1,2} = T_{1,2}^1 - T_{1,2}^2$, e $dm_{1,2}^{1,2} = dm_{1,2}^1 - dm_{1,2}^2$.

A equação correspondente para a fase da onda portadora é da forma (MONICO, 2000, p.174):

$$\phi_{1,2}^{1,2} + v_{\phi_{DD}} = f \left(\frac{\rho_{1,2}^{1,2} - I_{1,2}^{1,2} + T_{1,2}^{1,2} + dm_{1,2}^{1,2}}{c} \right) + N_{1,2}^{1,2}, \quad (2.6)$$

com $N_{1,2}^{1,2} = N_1^1 - N_2^1 - N_1^2 + N_2^2$.

A característica mais importante das DD é a eliminação dos erros dos relógios dos receptores (dt_{r_1} e dt_{r_2}) em adição à anulação do erro do relógio do satélite que foi realizada pela SD. Os erros do relógio dos receptores se cancelariam completamente se as observações dos satélites s^1 e s^2 fossem realizadas simultaneamente. Além disso, a fase inicial

do receptor também é cancelada. O multicaminho não é eliminado nas DD, pois depende da geometria entre receptor, satélite e refletor, ou seja, das condições de reflexão do sinal na região onde está localizada a antena do receptor (LEICK, 1995, p.261). Além disso, para as linhas de base médias e longas os erros causados pela órbita dos satélites GPS, refração ionosférica e troposférica também não são eliminados.

A equação de DD é normalmente a observável preferida nos processamentos de dados GPS envolvendo a fase da onda portadora. Ela parece proporcionar a melhor relação entre o ruído resultante e a eliminação de erros sistemáticos envolvidos nas observáveis originais (MONICO, 2000, p.174).

2.4.1 Matriz variância covariância

As observações da fase ou pseudodistância originais são supostas não correlacionadas no tempo e no espaço. A partir do momento em que novas observáveis (SD ou DD) são produzidas com base na combinação de várias observáveis originais, elas tornam-se correlacionadas, devendo tal correlação ser considerada no ajustamento (MONICO, 2000, p.175).

Um vetor ϕ_i , contendo as observações coletadas nas estações 1 e 2, a partir de n satélites, durante uma época t_i e arranjadas da seguinte forma:

$$\phi_i^T = [\phi_1^1, \phi_1^2, \dots, \phi_1^n, \phi_2^1, \phi_2^2, \dots, \phi_2^n], \quad (2.7)$$

tem sua Matriz Variância Covariância (MVC) dada por (MONICO, 2000, p.175):

$$\sum_{\phi_i} = \sigma^2 I_{2n}, \quad (2.8)$$

onde:

- I_{2n} é a matriz identidade de ordem igual ao número de observações, isto é, $2n$;

- σ^2 é a variância da observação.

As observações de SD podem ser escritas como (MONICO, 2000, p.176):

$$\phi_{SD_i} = [I_n; -I_n] \phi_i, \quad (2.9)$$

onde ϕ_{SD_i} é um vetor $(n \times 1)$ contendo as SD. Aplicando a lei de propagação de covariâncias, obtém-se a MVC do vetor das SD (MONICO, 2000, p.176):

$$\sum_{\phi_{SD_i}} = \phi_{SD_i} \sum_{\phi_i} \phi_{SD_i}^T = 2\sigma^2 I_n. \quad (2.10)$$

As $((n-1) \times n)$ observáveis de DD independentes contidas no vetor ϕ_{DD_i} são obtidas a partir das SD, e podem ser escritas como (MONICO, 2000, p.176):

$$\phi_{DD_i} = C \phi_{SD_i}. \quad (2.11)$$

A matriz C , de ordem $((n-1) \times n)$, contendo as informações para a obtenção das DD, pode ser definida de várias formas. Na prática, somente duas formas são extensivamente usadas, as quais são denominadas diferença seqüencial e satélite de referência ou base. No método da diferença seqüencial, a matriz C é dada por (MONICO, 2000, p.176):

$$C = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 1 & -1 \end{pmatrix}, \quad (2.12)$$

e no método do satélite base, com o satélite 1 definido como base (MONICO, 2000, p.176):

$$C = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \dots & 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.13)$$

Aplicando a lei de propagação de covariâncias, e considerando a matriz C dada por (2.12), obtém-se:

$$\begin{aligned}
\sum_{DD_i} &= C \sum_{\phi_{SD_i}} C^T = 2\sigma^2 \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & \vdots \\ 0 & -1 & \dots & 0 \\ \vdots & 0 & \dots & 0 \\ 0 & \vdots & \dots & 1 \\ 0 & 0 & \dots & -1 \end{pmatrix} \Rightarrow \\
&\Rightarrow \sum_{DD_i} = 2\sigma^2 \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & 0 & \dots & 0 & 0 & -1 & 2 \end{pmatrix}. \quad (2.14)
\end{aligned}$$

No caso em que se considera um satélite base (equação (2.13)), independentemente do satélite escolhido, tem-se:

$$\Rightarrow \sum_{DD_i} = 2\sigma^2 \begin{pmatrix} 2 & 1 & \dots & 1 & 1 \\ 1 & 2 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 2 \end{pmatrix}. \quad (2.15)$$

É importante frisar que a escolha do método para formar as DD não afeta os resultados do processamento e que apenas observáveis independentes são utilizadas. As DD não são correlacionadas entre épocas. Portanto, a MVC de, por exemplo, k épocas é composta por k blocos diagonais, similares aos da equação (2.14) ou (2.15).

2.4.2 Variância a posteriori

A variância a posteriori dá uma indicação sobre a validade do ajustamento como um todo, está relacionada com a confiabilidade do processo adotado (MONICO, 2000, p.166).

O cálculo da variância a posteriori é dado por (GEMAEL, 1994, p.122):

$$\hat{\sigma}_0^2 = \frac{V^T P V}{S}, \quad (2.16)$$

onde:

- V é o vetor dos resíduos;
- P é a matriz peso;
- S é o grau de liberdade.

2.5 Posicionamento relativo

Para realizar o posicionamento relativo o usuário deve dispor de dois ou mais receptores. No entanto, com o advento dos chamados Sistemas de Controle Ativos (SCA)¹, o usuário que dispõe de apenas um receptor pode efetuar o posicionamento relativo. Deve, para tal, acessar os dados de uma ou mais estações pertencentes ao SCA; no caso do Brasil, a Rede Brasileira de Monitoramento Contínuo (RBMC) (FORTES, 1997, p.7). Nesse caso, a realização do sistema de referência do SCA será introduzido na solução do usuário via as coordenadas das estações utilizadas como referência.

No posicionamento relativo utilizam-se em geral as DD como observáveis fundamentais (MONICO, 2000, p.205), pois os erros comuns às estações são reduzidos durante o processo de dupla diferenciação. Por isso, o posicionamento relativo tem se tornado muito popular e útil em levantamentos (LEICK, 1995, p.279).

Os métodos de posicionamento relativo são (MONICO, 2000, p.205):

- Estático;

¹ Num SCA, receptores rastreiam continuamente os satélites visíveis e os dados podem ser acessados via um sistema de comunicação.

- Estático rápido;
- Semicinemático;
- Cinemático.

Esses métodos de posicionamento podem ser realizados utilizando as seguintes observáveis (MONICO, 2000, p.206):

- Pseudodistância;
- Fase da onda portadora;
- Fase da onda portadora e Pseudodistância.

O princípio fundamental do posicionamento relativo é que os dois ou mais receptores envolvidos rastreiem, simultaneamente, um grupo de pelo menos dois satélites (MONICO, 2000, p.206).

Como nessa pesquisa será utilizado o posicionamento relativo estático, na próxima seção é realizada uma breve descrição desse método de posicionamento, e na seção seguinte são apresentadas as características relevantes do posicionamento estático rápido.

2.5.1 Posicionamento relativo estático

A observável adotada no posicionamento relativo estático é a DD da fase de batimento da onda portadora, muito embora possa também ser utilizada a DD da pseudodistância, ou ambas. Os casos em que se têm as duas observáveis proporcionam melhores resultados em termos de acurácia (MONICO, 2000, p.207).

Nesse tipo de posicionamento, dois ou mais receptores rastreiam, simultaneamente, os satélites visíveis por um período de tempo que pode variar de dezenas de minutos (em geral no mínimo 20 min), até algumas horas. Normalmente, os casos envolvendo

períodos curtos de ocupação, até 20 minutos, serão tratados como método relativo estático rápido.

Como no posicionamento relativo estático o período de ocupação é relativamente longo, somente as DD da fase da onda portadora são normalmente incluídas como observáveis. Como a precisão da fase da onda portadora é muito superior à da pseudodistância, a utilização dessa última não melhora os resultados de forma significativa. Mesmo assim, as pseudodistâncias devem estar disponíveis, pois elas são utilizadas no pré-processamento para estimar o erro do relógio do receptor, ou calcular o instante aproximado de transmissão do sinal pelo satélite (MONICO, 2000, p.208).

Trata-se da técnica mais utilizada em posicionamento geodésico, particularmente em softwares comerciais (MONICO, 2000, p.208).

Além disso, como a duração da coleta de dados é relativamente longa, as ambigüidades, exceto em alguns casos com problemas não esperados, são facilmente solucionadas no processo de ajustamento. Isso se deve à alteração da geometria dos satélites durante a sessão.

2.5.2 Posicionamento relativo estático rápido

O posicionamento relativo estático rápido segue, em linhas gerais, o mesmo princípio que o do posicionamento estático. A diferença fundamental diz respeito ao período de ocupação da estação de interesse. Neste caso, as ocupações, em geral, não excedem 20 minutos, ao passo que no posicionamento relativo estático as ocupações podem durar várias horas. A utilização do método estático rápido é propícia para levantamentos em que se deseja alta produtividade, mas há muitas obstruções entre as estações a serem levantadas. Além

disso, esse tipo de posicionamento é realizado em linhas de base curtas. Pode-se utilizar neste caso receptores de simples (L1) ou dupla frequência (L1 e L2) (MONICO, 2000, p.212).

Um receptor serve como base, permanecendo fixo sobre uma estação de referência, coletando dados, enquanto um outro receptor percorre as estações de interesse (receptor móvel), onde permanece parado cerca de 5 a 20 minutos, para coletar dados. Não há necessidade de continuar rastreando durante o deslocamento entre as estações, o que permite desligar o receptor móvel (MONICO, 2000, p.212).

Os dados coletados simultaneamente na estação de referência e nas estações a determinar, formando várias linhas bases, são processados. Para que os resultados apresentem razoável nível de precisão, o vetor de ambigüidade envolvido em cada linha deve ser solucionado, isto é, fixado como inteiro (MONICO, 2000, p.212). Deve-se, portanto, utilizar um algoritmo adequado de solução da ambigüidade, ou mesmo aqueles envolvidos nas técnicas *On-The-Fly* (OTF), como o método *Least Squares Ambiguity Decorrelation Adjustment* (LAMBDA) (TEUNISSEN, 1998a).

3 ERROS ENVOLVIDOS NAS OBSERVÁVEIS GPS

As observáveis GPS, tal como outras observáveis envolvidas nos processos de medidas, estão sujeitas a erros aleatórios, sistemáticos e grosseiros. Para obter resultados confiáveis, o modelo matemático (funcional e estocástico) estabelecido deve ser válido para a realidade física que se tenta descrever, e capaz de detectar problemas. Dessa forma, as fontes de erro envolvidas nos processos de medidas devem ser consideradas (MONICO, 2000, p.120).

Os diversos erros, agrupados pelas possíveis fontes, envolvidos no posicionamento GPS são apresentados na tabela 01. No que se refere à estação, é bom frisar que marés terrestres, carga dos oceanos e da atmosfera e movimento do pólo não se tratam especificamente de erros, mas de variações que devem ser consideradas para os casos de posicionamento de alta precisão, especialmente em linhas de bases longas.

TABELA 01 - Fontes e efeitos dos erros envolvidos no GPS

FONTES DE ERRO	ERROS
SATÉLITE	Erro da órbita Erro do relógio Relatividade Atraso entre as duas portadoras no hardware do satélite
PROPAGAÇÃO DO SINAL	Refração troposférica Refração ionosférica Perdas de ciclo Multicaminhamento ou sinais refletidos Rotação da Terra
RECEPTOR/ANTENA	Erro do relógio Erro entre os canais Centro de fase da antena
ESTAÇÃO	Erro nas coordenadas Multicaminhamento Marés terrestres Movimento do Pólo Carga dos oceanos Pressão da atmosfera

Fonte: MONICO, 2001, p. 121.

Nesse trabalho, pretende-se atenuar os erros sistemáticos que degradam a acurácia das coordenadas de interesse no posicionamento relativo GPS. Dentre esses erros pode-se citar, por exemplo, a refração ionosférica e troposférica, o multicaminho e os erros nas órbitas dos satélites GPS. Portanto, nas próximas seções são descritas as principais características desses erros. Será dada maior ênfase à refração ionosférica, visto que para linhas de base de comprimento médio e longo a ionosfera é a maior fonte de erro sistemático (ALVES, MONICO e MENEGUETTE, 2003).

3.1 Ionosfera

A ionosfera é uma importante fonte de erro para usuários do GPS que requerem medidas com boa acurácia. Às vezes, os erros causados pela troposfera e ionosfera podem ser comparados, mas a variabilidade dos erros da ionosfera é muito maior que da troposfera, e também é mais difícil para se modelar. Os erros da ionosfera podem variar de poucos metros a dezenas de metros, enquanto que na troposfera os erros no zênite estão geralmente entre dois e três metros (KLOBUCHAR, 1996, p.485).

A ionosfera é um meio dispersivo para a faixa de frequência do GPS, no qual o índice de refatividade é função da frequência utilizada. Assim, os usuários de GPS podem utilizar receptores de dupla frequência para tirar vantagem dessa propriedade, eliminando com isso os erros de primeira ordem (KLOBUCHAR, 1996, p.485).

3.1.1 Características da ionosfera

A radiação solar causa a fotoionização da atmosfera terrestre nas altas altitudes, criando, na atmosfera superior, regiões parcialmente ionizadas, conhecidas como ionosfera, que variam de, aproximadamente, 50 a 1000 km (CAMARGO, 1999, p.12).

A ionosfera é um plasma fracamente ionizado, ou gás, que pode afetar a propagação de ondas de rádio. Diferentes regiões da ionosfera são produzidas por diferentes substâncias químicas. A ionosfera é formada pelas regiões D, E, F₁ e F₂, nomeadas em ordem crescente de altitude (figura 03). Essas regiões da ionosfera são produzidas por diferentes comprimentos de ondas da radiação solar (KLOBUCHAR, 1996, p.486).

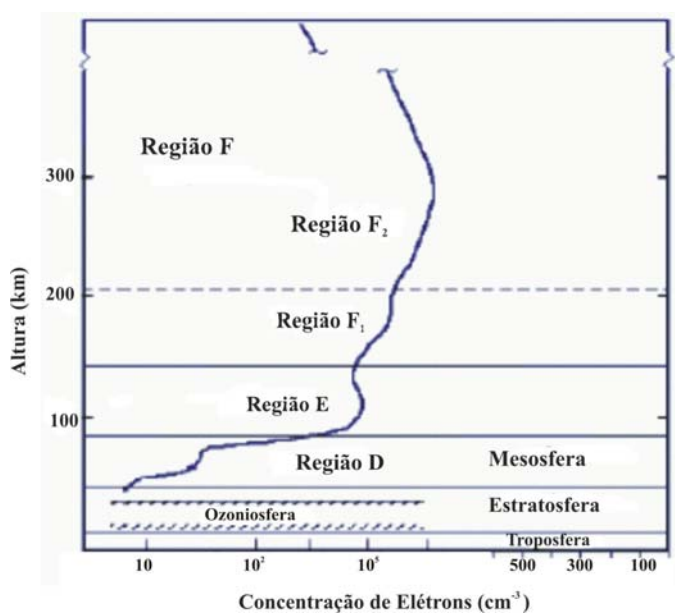


FIGURA 03 - Camadas da Ionosfera

Fonte: Adaptado de <http://ion.le.ac.uk/ionosphere/profile.html> (University of Leicester, 2003).

As principais características dessas regiões são (KLOBUCHAR, 1996, p.487; CAMARGO, 1999, p.14):

- **Região D** (50 – 90 km) - Na região D, região mais baixa da ionosfera, a concentração máxima de elétrons ocorre próxima à altitude de 80 km e é da ordem de 10^3 elétrons/cm³

(el/cm^3). Essa região desaparece durante a noite, devido à falta de ionização e recombinação dos elétrons. Ela é importante na propagação de ondas e atua como uma fonte absorvedora de energia eletromagnética das ondas de frequência média (MF), de frequência alta (HF) e de frequência muito alta (VHF), refletora dos sinais de frequência baixa (LF) e de frequência muito baixa (VLF). Devido à alta densidade de gás, faz com que as colisões dos elétrons sejam altas. À noite, quando a concentração de elétrons é menos acentuada, a propagação de ondas não é muito perturbada. Dessa forma, essa região tem efeitos não mensuráveis nas frequências GPS;

- **Região E (90 – 140 km)** - A região E apresenta uma concentração de elétrons da ordem de $10^5 \text{ el}/\text{cm}^3$. A densidade de elétrons é maior próxima do meio dia local, com comportamento quase simétrico ao longo do dia. Variações na densidade de elétrons ocorrem próximo das altitudes de 90-120 km ou maior, gerando uma fina camada com poucos quilômetros de espessura, designada de esporádica E, representada por Es. Nas regiões próximas ao equador magnético, ocorrências diurnas da esporádica E são bem comportadas e apresentam feições regulares com pouca variação sazonal. A Es ocorre ocasionalmente à noite, para altas latitudes, e durante o dia próximo ao equador magnético. Nas médias latitudes, a ocorrência é maior durante o verão do que no inverno. Com respeito à propagação de ondas, a esporádica Es reflete ondas de rádio, com frequência acima de aproximadamente 100 MHz. A região normal E tem efeito mínimo no GPS. A região esporádica E também tem efeitos desprezíveis no GPS;
- **Região F₁ (140 – 210 km)** - A região normal F₁, combinada com a região E, pode causar até 10% do atraso de tempo ionosférico encontrado no GPS. A região F1 apresenta uma concentração de elétrons variando de $2,5 \times 10^5 \text{ el}/\text{cm}^3$ a $4 \times 10^5 \text{ el}/\text{cm}^3$, para ocorrências de manchas solares mínimas e máximas, respectivamente. Porém, à noite, essa região desaparece;

- **Região F₂** (210 – 1000 km) - A região F₂ é a mais densa e tem a maior variação, causando a maioria dos efeitos nos sistemas de recepção GPS. A altura do pico da densidade de elétrons geralmente varia de 250 a 400 km. A região F₂ é produzida principalmente pela ionização de átomos de oxigênio, o qual é o principal componente da atmosfera neutra nesta altitude. A região F₂, e até certo ponto a região F₁, causam a maior parte dos problemas da propagação de ondas de rádio nas frequências GPS.

3.1.2 Efeitos da ionosfera nos sinais GPS

Os sinais GPS, no seu caminho entre o satélite e a antena da estação de rastreamento, propagam-se, através de uma atmosfera dinâmica, atravessando camadas de diferentes naturezas e estados variáveis. Assim sendo, sofrem diferentes tipos de influência, que podem provocar variações na direção de propagação, na velocidade de propagação, na polarização e na potência do sinal (SEEBER, 2003, p.309).

Esse meio de propagação compreende a troposfera e a ionosfera, com características bem diferentes. A ionosfera, como um meio dispersivo para a faixa de frequência GPS, afeta a modulação e a fase da portadora, fazendo com que sofram, respectivamente, um retardo e um avanço (CAMARGO, 1999, p.21).

O efeito da refração ionosférica depende da frequência usada e, conseqüentemente, do índice de refração. O efeito da refração é proporcional ao Conteúdo Total de Elétrons (*Total Electron Contents* - TEC), ou seja, o número de elétrons presentes ao longo do caminho percorrido pelo sinal entre o satélite e o receptor, e inversamente proporcional ao quadrado da frequência (MONICO, 2000, p.135). Além disso, outros parâmetros que influenciam a refração ionosférica são principalmente a atividade solar e o

campo geomagnético. A refração ionosférica também depende da frequência, da localização geográfica e do tempo (SEEBER, 2003, p. 309).

Dessa forma, o TEC apresenta os seguintes tipos de variação:

- **Diária** - Máxima por volta das 14 h local, com possibilidade de um segundo máximo por volta das 22 h na região equatorial (FORTES, 2002, p.36; LIU, 2001, p.37);
- **Sazonal** – No Hemisfério Norte o TEC é mínimo no verão e máximo próximo aos equinócios (Março e Setembro) e no inverno. O TEC é de 2 a 3 vezes maior no inverno que no verão. Já no Hemisfério Sul as condições são opostas, isto é, o TEC é menor no inverno (igual ao verão do Hemisfério Norte) e máximo no verão (FORTES, 2002, p.36);
- **Geográfica** – Os valores de pico do TEC são geralmente encontrados na região equatorial. Há ainda alta concentração de elétrons nas baixas latitudes (até $\pm 15^\circ$ a 20°) situados em ambos os lados do equador magnético. Esse fenômeno é denominado anomalia equatorial (FONSECA, 2002, p.26). Já as regiões de latitudes médias são consideradas relativamente livres das anomalias ionosféricas, enquanto as regiões polares não são muito previsíveis (CAMARGO, 1999, p. 31);
- **Longo período** - Com ciclo de aproximadamente 11 anos, também designada de variação do ciclo solar, é associada à ocorrência de manchas solares, que provocam um correspondente aumento na ionização (CAMARGO, 1999, p.29; GIZAWY, 2003, p.24). A figura 04 ilustra esta variação para os últimos cinco ciclos.

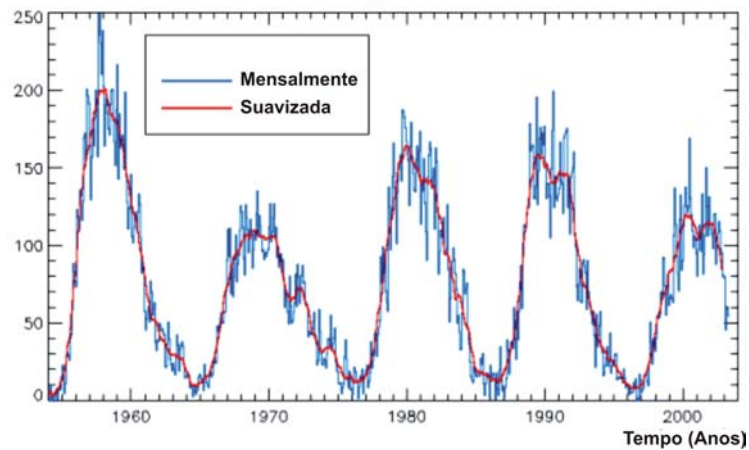


FIGURA 04 - Variação de longo período da Ionosfera para os últimos cinco ciclos
 Fonte: <http://sidc.oma.be/html/wolfmms.html> (SIDC, 2003)

Em Seeber (2003, p.309) são desenvolvidas as equações do índice de refração ionosférico para a fase da onda portadora e para o grupo (código), respectivamente dadas por:

$$n_f = 1 - \frac{40,3n_e}{f^2} \quad e \quad n_g = 1 + \frac{40,3n_e}{f^2}, \quad (3.1)$$

onde:

- n_e é a densidade de elétrons (n° de elétrons/ m^3);
- f é a frequência (Hz).

Como $n_f = \frac{c}{V_f}$ e $n_g = \frac{c}{V_g}$, sendo c a velocidade da luz, nota-se que ocorre

atraso no grupo e avanço na fase, pois a velocidade do grupo V_g é menor que a velocidade da fase V_f . Dessa forma, resulta em um aumento nas distâncias obtidas a partir dos códigos modulados sobre a portadora, e uma diminuição nas obtidas a partir da fase, de uma mesma quantidade (MONICO, 2000, p.138).

A figura 05 mostra o comportamento da refratividade R em função da altitude. Para a troposfera R é positivo e independente da frequência usada. Para a ionosfera, a refratividade é negativa, e depende da frequência (SEEBER, 2003, p. 53).

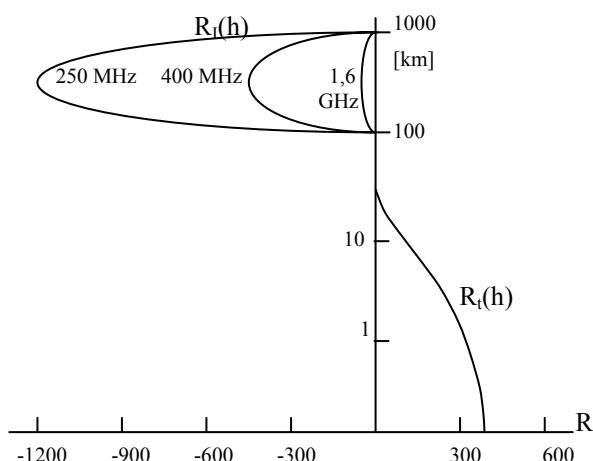


FIGURA 05 - Comportamento da refratividade R em função da altura; para a troposfera (R_t) e ionosfera (R_i)
 Fonte: Seeber, 2003, p.53.

De acordo com as equações (3.1) a refratividade da ionosfera torna-se menor quando a frequência aumenta (ver tabela 02). Mas altas frequências são tecnicamente exigentes. Frequências acima de 10 GHz não podem ser usadas facilmente com a tecnologia existente (SEEBER, 2003, p. 52).

TABELA 02 - Efeito do atraso de propagação, devido à ionosfera sobre as distâncias medidas com observações de uma frequência, e erros residuais para observações de dupla frequência

Simples frequência	400 MHz	1600 MHz	2000 MHz	8000 MHz
Efeito médio	50 m	3 m	2 m	0,12 m
Para 90% menor que	250 m	15 m	10 m	0,6 m
Efeito máximo	500 m	30 m	20 m	1,2 m
Dupla frequência	150/400 MHz	400/2000 MHz	1227/1572 MHz	2000/8000 MHz
Efeito médio	0,6 m	0,9 cm	0,3 cm	0,04 cm
Para 90% menor que	10 m	6,6 cm	1,7 cm	0,21 cm
Efeito máximo	36 m	22 cm	4,5 cm	0,43 cm

Fonte: SEEBER, 2003, p. 53.

Embora sejam muito importantes, o avanço da fase e o atraso do grupo não são as únicas manifestações da ionosfera na propagação do sinal; pode ocorrer também a cintilação, a qual pode, em algumas latitudes, causar uma rápida oscilação na amplitude do sinal e da fase recebidas (SPILKER Jr. e PARKINSON, 1996, p.50; OLYNIC, 2002, p.12).

Este termo cintilação é análogo à variação em intensidade de estrelas luminosas quando vistas através de uma atmosfera turbulenta (LEICK, 1995, p.297).

Para estimar a densidade de elétrons, vários modelos têm sido desenvolvidos. Para a correção de medidas GPS o modelo de Klobuchar geralmente é aplicado. Esse modelo corrige cerca de 50% do efeito total da ionosfera. Uma melhor alternativa é obtida quando o coeficiente c_2 (utilizado para estimar o índice de refração) é determinado a partir de observações simultâneas dos sinais transmitidos pelos sinais GPS em duas frequências diferentes (SEEBER, 2003, p. 54; CAMARGO, 1999, p.24).

O erro causado pela ionosfera no posicionamento relativo com GPS está geralmente entre 1 e 2 partes por milhão (ppm). Entretanto, já foram encontrados valores acima de 15 ppm na região auroral e acima de 40 ppm na região equatorial (FORTES, 2002, p.39).

Os efeitos da refração ionosférica podem ser praticamente eliminados quando dados oriundos de receptores de dupla frequência estiverem disponíveis (realizando combinações lineares entre as portadoras L_1 e L_2 é possível eliminar erros de primeira ordem). Por outro lado, os usuários de receptores de monofrequência têm de negligenciar os efeitos ou, quando for o caso, corrigi-los a partir de modelos existentes. Uma outra possibilidade é realizar apenas o posicionamento relativo envolvendo linhas de base curtas (MONICO, 2000, p.144). Mas, com o algoritmo desenvolvido nesse projeto de mestrado, usuários de simples frequência poderão reduzir de forma significativa os efeitos causados pela ionosfera no posicionamento relativo GPS.

3.2 Troposfera

A troposfera é a camada mais superficial da atmosfera, com espessura média de aproximadamente 50 km (MONICO, 2000, p.126). Para a frequência dos sinais GPS, o atraso troposférico não depende da frequência, ou seja, a troposfera é um meio não dispersivo. Portanto ela não pode ser corrigida utilizando medidas de dupla frequência (SEEBER, 2003, p.314).

Usualmente, as componentes úmida e hidrostática (seca) expressam a influência da troposfera nas medidas GPS (SAPUCCI, 2001, p.1). A componente úmida depende da quantidade de vapor d'água na atmosfera e é difícil de modelar. Mas ela é responsável por apenas 10% da refração troposférica total. Já a componente seca é precisamente descrita (com acurácia de $\pm 1\%$) por modelos. Os erros causados pelas componentes úmida e seca no zênite podem alcançar 0.80 m e 2.3 m, respectivamente (SPILKER Jr., 1996 p.524), aumentando aproximadamente 10 vezes próximo ao horizonte (10° de elevação) (SEEBER, 2003, p.315).

No posicionamento diferencial, o erro residual troposférico desaparece quase completamente para linhas de base curtas (SEEBER, 2003, p.315). Mas, quando a distância entre as estações é grande, ou quando a diferença de altitude é grande (em regiões montanhosas), isso não ocorre pois as condições atmosféricas locais não são suficientemente correlacionadas. Assim, segundo Fortes (2002, p.25) o erro causado pela troposfera varia de 0.2 a 0.4 ppm, depois da aplicação de um modelo. Além disso, antes de aplicar um modelo, pode variar de 1 a 4 ppm, dependendo do ângulo de elevação do satélite.

3.3 Multicaminho

O multicaminho é o fenômeno pelo qual um sinal chega à antena do receptor por caminhos múltiplos, devido à reflexão (Figura 06) (BRAASCH, 1996, p.547).

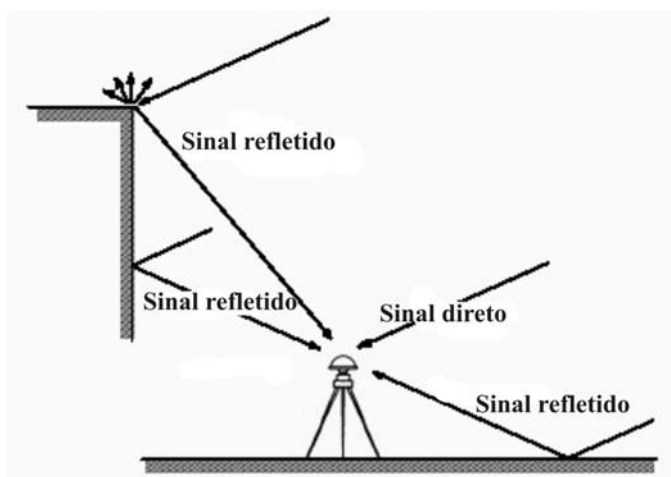


FIGURA 06 - Efeito do multicaminho
Fonte: Hannah, Walker e Kubik, 1998.

O multicaminho é causado principalmente por reflexões do sinal em superfícies próximas ao receptor, tais como construções, carros, árvores, colinas, etc (MONICO, 2000, p.145). Efeitos secundários são causados por reflexões no próprio satélite e durante a propagação do sinal (HOFMANN-WELLENHOF, 1997, p.126).

A aplicação do método relativo (DD) reduz erros, como, por exemplo, do relógio do satélite, da órbita e da atmosfera, menos o de multicaminho. Isto provém do fato de que o multicaminho é um fenômeno altamente localizado. As fontes de multicaminho que afetam o receptor de uma estação de referência não causam, necessariamente, erros em um receptor móvel. Da mesma forma, fontes de multicaminho que afetam o receptor móvel podem não afetar a estação de referência (BRAASCH, 1996, p.547).

O multicaminho causa erros nas medidas de pseudodistância e de fase da onda portadora, dependendo da geometria do cenário envolvendo as antenas, os satélites e os

objetos refletores e da natureza do material reflexivo (FARRET, 2000, p.1). Conseqüentemente, a posição do objeto de interesse terá sua precisão e acurácia deterioradas. O erro causado pelo multicaminho na fase da onda portadora pode chegar a 1/4 do comprimento de onda. Já para a pseudodistância esse erro pode alcançar ± 150 m (RAY, 2000, p.85).

Nesse sentido, várias técnicas têm sido desenvolvidas para atenuar o multicaminho. Estas técnicas incluem o uso de antenas especiais, arranjo de várias antenas, estratégia de localização da antena, técnicas de software, etc. Nesse trabalho, o MMQ com penalidades e o modelo semiparamétrico são utilizados para atenuar os efeitos desses erros.

3.4 Erros na órbita dos satélites GPS

Informações sobre as órbitas dos satélites GPS podem ser obtidas através das efemérides transmitidas pelos satélites ou das efemérides precisas fornecidas pelo *International GPS Service* (IGS). A partir das efemérides são calculadas as posições dos satélites GPS, normalmente, injuncionadas como fixas durante o processo de ajustamento dos dados GPS. Assim, erros nas coordenadas dos satélites se propagam para a posição do usuário.

No posicionamento relativo, os erros orbitais são praticamente eliminados pelo processo de diferenciação. Mas erros remanescentes degradam a acurácia da linha de base na medida em que essa se torna mais longa. Uma regra que expressa esse erro é dada por (MONICO, 2000, p. 123):

$$\Delta b = b \frac{\Delta r}{r}, \quad (3.2)$$

onde:

- Δb é o erro resultante na linha de base;
- b é o comprimento da linha de base;
- Δr é o erro na posição do satélite;
- r é a distância do satélite ao receptor.

A acurácia das efemérides transmitidas, de acordo com IGS (2004), é de aproximadamente 2 m. Elas estão disponíveis em tempo real, pois são transmitidas com os arquivos de observação. A tabela 03 apresenta a acurácia das efemérides precisas:

TABELA 03 - Produtos IGS

ÓRBITAS IGS	ACURÁCIA	LATÊNCIA	ATUALIZAÇÕES
Ultra-rápida (predita)	≈ 10 cm	Tempo real	Duas vezes por dia
Ultra-rápida (observada)	< 5 cm	3 horas	Duas vezes por dia
Rápida	< 5 cm	17 horas	Diariamente
Final	< 5 cm	≈ 13 dias	Semanalmente

Fonte: IGS (2004)

4 MÍNIMOS QUADRADOS COM PENALIDADES E MODELO SEMIPARAMÉTRICO

O Método dos Mínimos Quadrados (MMQ) com penalidades e o modelo semiparamétrico tem sido utilizado por alguns autores na determinação das ambigüidades e das coordenadas de interesse (JIA, STEWART e TSAKIRI, 2001).

Nesse capítulo, um estudo dessa técnica será realizado, já que nessa dissertação de mestrado utiliza-se o MMQ com penalidades e o modelo semiparamétrico para realizar o ajustamento das observáveis GPS.

4.1 Aproximações por regressão

Antes de se introduzir a aproximação por penalidade de aspereza (utilizada no MMQ com penalidades), será realizada uma breve discussão sobre as regressões linear e polinomial.

4.1.1 Regressão linear

A regressão linear é uma das mais antigas e utilizadas técnicas estatísticas. Dados (t_i, Y_i) , $i = 1, \dots, n$, a regressão linear é dada pelo modelo:

$$Y = a + bt + \text{erro}. \quad (4.1)$$

Existem duas finalidades principais para a regressão linear. A primeira é proporcionar um resumo ou redução dos dados observados através de uma aproximação linear, e o segundo propósito é utilizar o modelo (4.1) para uma predição dos dados (GREEN e SILVERMAN, 1994, p.1).

4.1.2 Regressão polinomial

Existem muitos conjuntos de dados onde é inapropriado ajustar um modelo linear como (4.1). Assim o seguinte modelo pode ser utilizado (GREEN e SILVERMAN, 1994, p.2):

$$Y = g(t) + \text{erro}, \quad (4.2)$$

onde g é um polinômio.

A aproximação clássica utiliza g com o menor grau polinomial possível, e os seus coeficientes são estimados pelo MMQ convencional.

A regressão polinomial é uma técnica popular, mas apresenta desvantagens. Uma delas é que observações individuais podem exercer uma influência, de modo inesperado, na curva. Uma outra dificuldade ocorre em relação ao aumento do grau do polinômio. Seria desejável que o aumento do grau da função interpolante fosse suave, e não discreto como ocorre com os polinômios, pois esse aumento causa mudanças significativas na curva final (GREEN e SILVERMAN, 1994, p.2).

4.2 Penalidade de aspereza

A penalidade de aspereza é utilizada no MMQ com penalidades, razão pelo qual será realizada uma breve descrição desse assunto.

4.2.1 O ajuste da curva

Em sua forma mais simples, a aproximação por penalidade de aspereza é um método para relaxar a rigidez do modelo usado na regressão linear clássica, com abordagem um pouco diferente da regressão polinomial (GREEN e SILVERMAN, 1994, p.2).

Considere, primeiro, o que aconteceria se um modelo da forma (4.2) fosse ajustado pelo MMQ, sem colocar nenhuma restrição na curva g . Neste caso, a soma dos quadrados dos resíduos poderia ser reduzida a zero se g fosse escolhida de forma que interpolasse o conjunto de dados (GREEN e SILVERMAN, 1994, p.2). Uma ilustração disso é mostrada na figura 07.

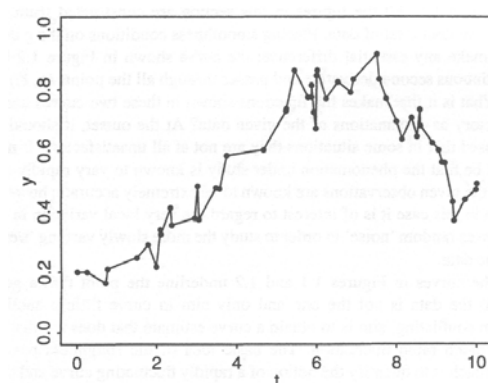


FIGURA 07 - Conjunto de dados unidos por linhas retas
Fonte: GREEN e SILVERMAN, 1994, p.3.

Colocando condições de suavidade em g , não ocorreriam diferenças essenciais. A curva mostrada na figura 08 (para o mesmo conjunto de dados) tem a derivada segunda contínua e passa por todos os pontos (t_i, Y_i) (GREEN e SILVERMAN, 1994, p.4).

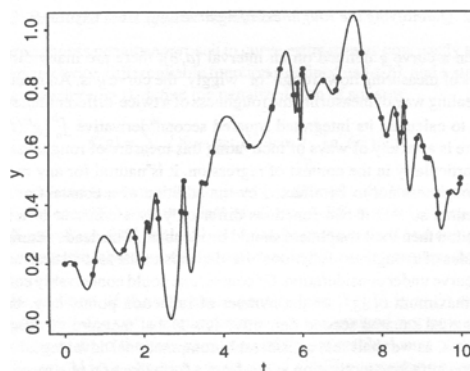


FIGURA 08 - Dados interpolados por uma curva com segunda derivada contínua
Fonte: GREEN e SILVERMAN, 1994, p.3.

Em algumas situações pode-se dizer que essas curvas são satisfatórias, pois pode ser que o fenômeno em estudo tenha grande variação e que as observações sejam extremamente acuradas. Porém, até mesmo neste caso, é de interesse considerar a variação local na curva como um ruído aleatório para se estudar a variação na tendência dos dados (GREEN e SILVERMAN, 1994, p.4).

4.2.2 Quantificando a aspereza de uma curva

Seja g uma curva duas vezes diferenciável definida em um intervalo $[a, b]$.

Um modo de medir sua aspereza é através do cálculo da integral de sua segunda derivada ao

quadrado $\int_a^b \{g''(t)\}^2 dt$.

Existem muitas formas para se medir a aspereza de uma curva, dentre elas pode-se considerar o número de pontos de inflexão em g , mas a $\int_a^b \{g''(t)\}^2 dt$ é uma medida global de aspereza que tem vantagens computacionais (GREEN e SILVERMAN, 1994, p.4).

Uma motivação para se usar esse método que determina a aspereza de uma curva surgiu de um dispositivo mecânico que foi muito usado (antes do surgimento da computação gráfica) para desenhar curvas suaves. Esse dispositivo consistia em um pedaço fino e flexível de madeira, chamado de spline, que era curvado para moldar o gráfico de g (GREEN e SILVERMAN, 1994, p.4). Mais detalhes sobre splines serão vistos na seção 4.3.

4.2.3 Regressão por Mínimos Quadrados Penalizados

Dada uma curva g duas vezes diferenciável definida em um intervalo $[a, b]$, e um parâmetro suavizador $\alpha > 0$, a soma dos quadrados penalizada é dada por (GREEN e SILVERMAN, 1994, p.5):

$$S(g) = \sum_{i=1}^n \{Y_i - g(t_i)\}^2 + \alpha \int_a^b \{g''(t)\}^2 dt. \quad (4.3)$$

A estimativa $\hat{g}$ do MMQ com penalidades é a função que minimiza $S(g)$ na classe das funções g duas vezes diferenciáveis.

A adição do termo da penalidade de aspereza $\alpha \int g''^2$ em (4.3) assegura que $S(g)$ de uma curva particular é determinada não somente pela sua aderência aos dados, quantificada pela soma dos quadrados dos resíduos $\sum \{Y_i - g(t_i)\}^2$, mas também pela sua aspereza $\int g''^2$ (GREEN e SILVERMAN, 1994, p.5). Um exemplo dessa aproximação usando

o MMQ com penalidades é dado na figura 09. Nesta figura, α é determinado pela validação cruzada generalizada (capítulo 5).

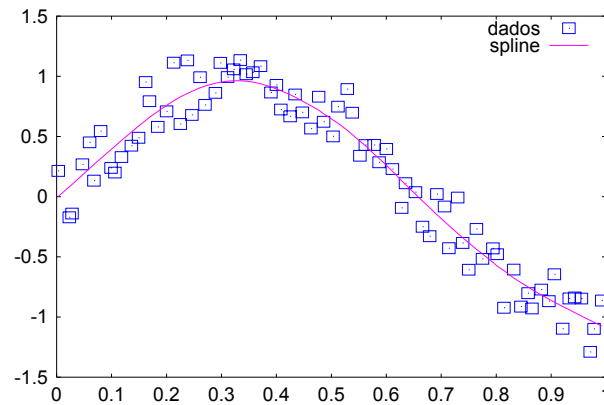


FIGURA 09 - Curva que minimiza $S(g)$ para α dado pela validação cruzada generalizada

Se α é grande, o principal componente em $S(g)$ será o termo de penalidade de aspereza, e, portanto, a curva que minimiza $S(g)$ exibirá uma curvatura pequena. Um exemplo é dado na figura 10. No caso limite onde α tende para o infinito o termo $\int \hat{g}''^2$ será forçado a zero e a curva $\hat{g}$ se aproximará de uma regressão linear (GREEN e SILVERMAN, 1994, p.6).

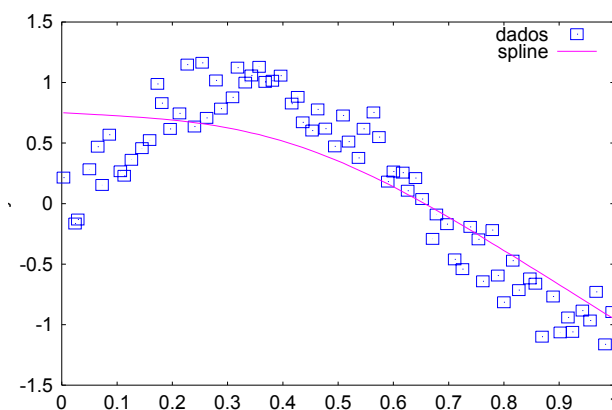


FIGURA 10 - Curva que minimiza $S(g)$ para um valor grande de α

Por outro lado, se α é relativamente pequeno, então a principal contribuição para $S(g)$ será a soma dos quadrados dos resíduos, como por exemplo na figura 11. No caso

limite onde α tende a zero, $\hat{g}$ se aproximará da curva mostrada na figura 08 (GREEN e SILVERMAN, 1994, p.6).

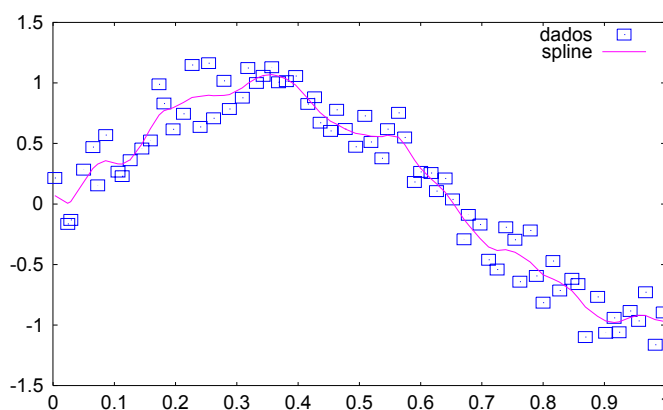


FIGURA 11 - Curva que minimiza $S(g)$ para um valor pequeno de α

A questão sobre como escolher o valor de α que melhor se adapte a um determinado problema será discutida no capítulo 5.

4.3 Splines

Para se trabalhar com o ajustamento e interpolação de funções, a aproximação por polinômios é muito conveniente, uma vez que os polinômios têm várias propriedades interessantes, dentre estas a analiticidade, que torna possível calcular as derivadas, de qualquer ordem, dos polinômios. Entretanto, a necessidade de muitas derivadas da função que está sendo aproximada por polinômios pode ser muito restritiva (CUNHA, 2000, p.125).

Uma solução para isso é a utilização de polinômios por partes, pois dessa forma pode-se escapar da analiticidade no intervalo inteiro, permitindo descontinuidades das derivadas de ordem mais elevadas em alguns pontos. Essa característica de

“pseudo analiticidade” confere às funções polinomiais por partes, chamadas splines, boas propriedades de aproximação, convergência e estabilidade (CUNHA, 2000, p.125).

Embora existam indicações de que funções suaves por partes tenham sido usadas por Euler (1707-1783) e outros autores, no início do século XX, deve-se a R. Courant e I. Shoemberg, respectivamente em 1943 e 1946, a formulação de splines que se usa hoje. A teoria das splines foi desenvolvida a partir de necessidades práticas de aproximação (CUNHA, 2000, p.125).

4.3.1 Origem

A origem do nome spline vem de uma régua elástica, usada em desenhos de engenharia, que pode ser curvada de forma a passar por um dado conjunto de pontos (x_i, y_i) .

Essa régua flexível que constrói a spline é uma longa e estreita tira de madeira ou plástico moldada por “presilhas” (figura 12). Pela variação da posição e do número de presilhas, a spline passa pelos pontos especificados resultando em uma curva suave (ROGERS e ADAMS, 1990, p.251).

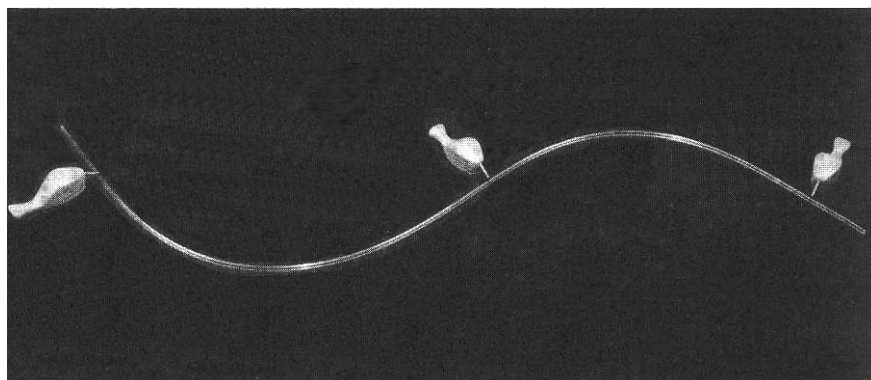


FIGURA 12 - Spline Mecânica moldada por presilhas
Fonte: ROGERS e ADAMS, 1990, p.252.

A curva definida pela régua pode ser descrita aproximadamente como uma função por partes, cada qual um polinômio cúbico, de tal forma que ela e suas duas primeiras derivadas são sempre contínuas. A terceira derivada, entretanto, pode ter descontinuidade nos pontos x_i (RUGGIERO, 1996, p.245).

Essas splines mecânicas não são mais utilizadas hoje em dia, mas elas foram muito utilizadas na construção de navios e estradas de ferro (GREEN e SILVERMAN, 1994, p.14).

4.3.2 Definição de spline

As funções splines estão associadas a partição de um intervalo $[a, b]$ onde se pretende trabalhar. Uma partição I será definida pelos pontos $x_0, x_1, \dots, x_m$, tal que:

$$I: a = x_0 < x_1 < \dots < x_{m-1} < x_m = b \quad (4.4)$$

Em cada subintervalo (x_{i-1}, x_i) , $i = 1, 2, \dots, m$, as splines são polinômios de um determinado grau n . Estes “pedaços” de polinômios são colocados convenientemente para que algumas derivadas, da ordem ditada pelo problema, existam em todo o intervalo $[a, b]$. Existe uma relação entre o grau dos “pedaços” dos polinômios e a ordem das derivadas exigidas nos pontos da partição. Assim, algumas restrições devem ser impostas na definição geral das splines (CUNHA, 2000, p.126).

Definição: Uma função $s(x)$ é chamada de spline de grau n , associada a uma partição (4.4) de $[a, b]$, se (CUNHA, 2000, p.126):

- $s(x)$ é um polinômio de grau n em cada subintervalo (x_{i-1}, x_i) ;
- $s(x)$ tem $n - 1$ derivadas contínuas em cada x_i , e portanto em $[a, b]$.

Para mais detalhes sobre a definição das splines, e também sobre bases de splines, consulte Alves (2003).

4.3.3 Splines cúbicas naturais

O MMQ com penalidades utiliza uma spline cúbica natural, portanto, nesta seção serão apresentados detalhes de tais funções.

Dados os números reais $t_1, \dots, t_n$ em algum intervalo $[a, b]$, satisfazendo $a < t_1 < t_2 < \dots < t_n < b$, uma função g definida sobre $[a, b]$ é uma spline cúbica se valerem as seguintes condições:

- Em cada intervalo $(a, t_1), (t_1, t_2), (t_2, t_3), \dots, (t_n, b)$, g é um polinômio cúbico;
- A primeira e segunda derivadas são contínuas em cada t_i e, portanto, contínuas no intervalo $[a, b]$.

Uma forma óbvia para escrever os quatro coeficientes de cada parte cúbica de uma spline pode ser, por exemplo (GREEN e SILVERMAN, 1994, p.11):

$$g(t) = d_i(t - t_i)^3 + c_i(t - t_i)^2 + b_i(t - t_i) + a_i \text{ para } t_i \leq t \leq t_{i+1}, \quad (4.5)$$

dadas as constantes $a_i, b_i, c_i, d_i, i = 0, \dots, n$, com $t_0 = a$ e $t_{n+1} = b$.

Uma spline cúbica sobre um intervalo $[a, b]$ será dita uma spline cúbica natural se as segunda e terceira derivadas são nulas em a e b . Essa condição implica que $d_0 = c_0 = d_n = c_n = 0$ e assim g é linear em $[a, t_1]$ e $[t_n, b]$ (GREEN e SILVERMAN, 1994, p.12).

4.3.3.1 Representação do valor da segunda derivada

A forma (4.5) não é conveniente para representar, computacionalmente, uma spline cúbica natural, ou mesmo para uma discussão matemática sobre a mesma. Deve-se especificar uma spline cúbica natural dando o seu valor e o de sua segunda derivada em cada nó t_i . Essa especificação será chamada de representação do valor da segunda derivada (GREEN e SILVERMAN, 1994, p.12).

Suponha que g é uma spline cúbica natural com nós $t_1 < \dots < t_n$, então define-se:

$$g_i = g(t_i) \text{ e } \gamma_i = g''(t_i) \text{ para } i = 1, \dots, n.$$

Pela definição de spline cúbica natural $\gamma_1 = \gamma_n = 0$. Assim, seja $\mathbf{g}$ o vetor $(g_1, \dots, g_n)^T$ e γ o vetor $(\gamma_2, \dots, \gamma_{n-1})^T$. Note que o vetor γ tem $(n-2)$ entradas γ_i . Isto ajudará em considerações posteriores (GREEN e SILVERMAN, 1994, p.12).

Nesta seção serão discutidas as condições necessárias e suficientes para que os vetores $\mathbf{g}$ e γ representem uma spline cúbica natural para uma dada sequência de nós. Essas condições dependem de duas matrizes em banda, Q e R , que serão definidas a seguir.

Seja $h_i = t_{i+1} - t_i$ para $i = 1, \dots, n-1$. Seja Q uma matriz $n \times (n-2)$ com entradas q_{ij} , para $i = 1, \dots, n$, dada por (GREEN e SILVERMAN, 1994; FESSLER, 1991):

$$q_{j-1,j} = h_{j-1}^{-1}, \quad q_{jj} = -h_{j-1}^{-1} - h_j^{-1} \text{ e } q_{j+1,j} = h_j^{-1},$$

para $j = 2, \dots, n-1$ e $q_{ij} = 0$ para $|i - j| \geq 2$. Assim, tem-se:

$$Q = \begin{pmatrix} h_1^{-1} & 0 & 0 & \dots & 0 \\ -h_1^{-1} - h_2^{-1} & h_2^{-1} & 0 & \dots & 0 \\ h_2^{-1} & -h_2^{-1} - h_3^{-1} & h_3^{-1} & \dots & 0 \\ 0 & h_3^{-1} & -h_3^{-1} - h_4^{-1} & \ddots & 0 \\ 0 & 0 & h_4^{-1} & \ddots & h_{n-2}^{-1} \\ \vdots & \vdots & \vdots & \ddots & -h_{n-2}^{-1} - h_{n-1}^{-1} \\ 0 & 0 & 0 & \dots & h_{n-1}^{-1} \end{pmatrix}. \quad (4.6)$$

As colunas de Q são enumeradas do mesmo modo que as entradas de γ , começando em $j = 2$, de forma que o elemento do topo esquerdo de Q é q_{12} (GREEN e SILVERMAN, 1994, p.12).

A matriz simétrica R é $(n - 2) \times (n - 2)$ com elementos r_{ij} dada por (GREEN e SILVERMAN, 1994; FESSLER, 1991):

$$r_{ii} = \frac{1}{3}(h_{i-1} + h_i) \quad i = 2, \dots, n-1.$$

$$r_{i,i+1} = r_{i+1,i} = \frac{1}{6}h_i \quad i = 2, \dots, n-2,$$

e $r_{ij} = 0$ para $|i - j| \geq 2$. Assim:

$$R = \begin{pmatrix} \frac{1}{3}(h_1 + h_2) & \frac{1}{6}h_2 & 0 & \dots & 0 \\ \frac{1}{6}h_2 & \frac{1}{3}(h_2 + h_3) & \frac{1}{6}h_3 & \dots & 0 \\ 0 & \frac{1}{6}h_3 & \frac{1}{3}(h_3 + h_4) & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \frac{1}{6}h_{n-2} \\ 0 & 0 & 0 & \frac{1}{6}h_{n-2} & \frac{1}{3}(h_{n-2} + h_{n-1}) \end{pmatrix}. \quad (4.7)$$

A matriz R é estritamente diagonal dominante, pois $|r_{ii}| > \sum_{j \neq i} |r_{ij}|$ para cada i (GREEN e SILVERMAN, 1994, p.13). Assim, usando o teorema de Gershgorin (GOLUB e LOAN, 1983, p.200) e Todd (1962, p.287), todos os autovalores de R são positivos. Portanto, R é estritamente positiva definida, e conseqüentemente, é inversível. Pode-se então definir uma matriz K dada por:

$$K = QR^{-1}Q^T. \quad (4.8)$$

Assim, a propriedade fundamental de uma spline cúbica natural pode agora ser apresentada.

Teorema 4.1 Os vetores $\mathbf{g}$ e γ especificam uma spline cúbica natural g se, e somente se, a condição

$$Q^T \mathbf{g} = R\gamma, \quad (4.9)$$

é satisfeita. Se (4.9) é satisfeita então na penalidade de aspereza tem-se:

$$\int_a^b g''(t)^2 dt = \gamma^T R\gamma = \mathbf{g}^T K \mathbf{g}. \quad (4.10)$$

Esse teorema está demonstrado em Alves (2003) e apresentado em Green e Silverman (1994).

4.3.4 Interpolando com splines

Suponha que são dados os valores $z_1, \dots, z_n$ aos pontos $t_1, \dots, t_n$. Deseja-se encontrar uma curva suave g que interpole os pontos (t_i, z_i) , isto é, $g(t_i) = z_i$ para $i = 1, \dots, n$.

Existem muitas maneiras de construir uma função interpolante g . A mais simples, e provavelmente a mais usada, é a interpolação linear, onde os pontos (t_i, z_i) são unidos por linhas retas. Mesmo sendo utilizada em muitas aplicações, esse tipo de interpolação não fornece uma curva suave, pois a função resultante g tem descontinuidades nas derivadas em cada ponto t_i . Além disso, pode ser mostrado matematicamente que uma curva suave interpolante, adequadamente escolhida, pode fornecer uma melhor aproximação da curva “verdadeira” do que a interpolação linear (GREEN e SILVERMAN, 1994, p.14).

Um método de interpolação que pode ser desenvolvido vem da definição da penalidade de aspereza, discutida na seção 4.2. Seja $S[a,b]$ o espaço de todas as funções g definidas em $[a,b]$ que têm duas derivadas contínuas. As funções pertencentes a $S[a,b]$ serão

chamadas de funções suaves. Se a curva interpolante deve ser a mais suave possível, então deve-se procurar entre todas as curvas que interpolam os dados, aquela que possua o valor de $\int g''^2$ mínimo (GREEN e SILVERMAN, 1994, p.14).

Entre todas as curvas g em $S[a,b]$ que interpolam os pontos (t_i, z_i) , a curva que minimiza $\int g''^2$ é uma spline cúbica natural com nós t_i . Além disso, se $n \geq 2$, existe uma única spline cúbica natural que interpola os dados. Essas afirmações serão discutidas nas próximas seções (GREEN e SILVERMAN, 1994, p.14).

4.3.4.1 Construindo a spline cúbica natural interpolante

Teorema 4.2 *Suponha $n \geq 2$ e $t_1 < \dots < t_n$. Dados os valores $z_1, \dots, z_n$, existe uma única spline cúbica natural g com nós nos pontos t_i satisfazendo:*

$$g(t_i) = z_i \text{ para } i = 1, \dots, n.$$

A demonstração deste teorema está em Green e Silverman (1994, p.15).

Para construir uma spline cúbica natural interpolante é necessário utilizar também o teorema 4.1 que especifica uma spline cúbica natural. De acordo com a seção 4.3.3.1, R é uma matriz tridiagonal e portanto $R\gamma = x$ pode ser resolvido para γ sem achar R^{-1} , em um número de operações de ordem linear. A natureza tridiagonal de Q (seção 4.3.3.1) assegura que $Q^T g$ também pode ser encontrado em um número linear de operações (GREEN e SILVERMAN, 1994, p.15). Dessa forma para $i = 2, \dots, n-1$, tem-se:

$$(Q^T g)_i = \frac{g_{i+1} - g_i}{h_i} - \frac{g_i - g_{i-1}}{h_{i-1}}. \quad (4.11)$$

Pode-se concluir que o algoritmo que segue encontrará uma spline cúbica natural interpolante, para n pontos (t_i, z_i) (GREEN e SILVERMAN, 1994, p.15).

Algoritmo para a interpolação de uma spline cúbica natural

Passo 1: Faça $g_i = z_i$ para $i = 1, \dots, n$.

Passo 2: Faça $x = Q^T g$ (usando a fórmula 4.11) e resolva $R\gamma = x$ para γ .

Para calcular os valores de g para outros pontos t_i , consulte Alves (2003) ou Green e Silverman (1994, p.22).

4.3.4.2 Propriedades de uma spline cúbica natural interpolante

Foi referido na seção 4.3.4 que uma spline cúbica natural interpolante possui um valor mínimo de $\int g''^2$ entre todas as curvas suaves que interpolam os dados. O teorema seguinte prova essa afirmação, sua demonstração encontra-se no apêndice A.

Seja $S_2[a, b]$ o espaço das funções que são diferenciáveis sobre $[a, b]$ e tenham a primeira derivada absolutamente contínua. Isto significa que uma função g pertencente a $S_2[a, b]$ é contínua e diferenciável sobre $[a, b]$ com derivada g' , e que existe uma função integrável g'' tal que $\int_a^x g''(t)dt = g'(x) - g'(a)$ para todo x em $[a, b]$. Assim

$S_2[a, b]$ contém todas as funções em $S[a, b]$ (GREEN e SILVERMAN, 1994, p.16).

Teorema 4.3 Suponha $n \geq 2$, e que g é uma spline cúbica natural interpolante para os valores $z_1, \dots, z_n$ nos pontos $t_1, \dots, t_n$ satisfazendo $a < t_1 < \dots < t_n < b$. Seja $\bar{g}$ uma função em $S_2[a, b]$ no qual $\bar{g}(t_i) = z_i$ para $i = 1, \dots, n$. Então $\int \bar{g}''^2 \geq \int g''^2$, com igualdade somente se $\bar{g}$ e g são idênticas.

4.3.5 Suavizando com splines

Como na seção 4.3.4, suponha que $t_1, \dots, t_n$ são pontos em $[a, b]$ satisfazendo $a < t_1 < \dots < t_n < b$ e $Y_1, \dots, Y_n$ são as observações. Será assumido que $n \geq 3$ para que as condições do teorema 4.1 sejam satisfeitas.

Seja $S(g)$ a soma dos quadrados penalizada (GREEN e SILVERMAN, 1994, p.17):

$$S(g) = \sum_{i=1}^n \{Y_i - g(t_i)\}^2 + \alpha \int_a^b \{g''(t)\}^2 dt,$$

como definido na seção 4.2.3, sendo g uma função em $S_2[a, b]$.

O mínimo de $S(g)$ será denotado por $\hat{g}$ na classe $S_2[a, b]$ de todas as curvas suaves em $[a, b]$. O próximo teorema sintetiza as propriedades de $\hat{g}$.

Teorema 4.4 Suponha $n \geq 3$ e que $t_1, \dots, t_n$ são pontos que satisfazem $a < t_1 < \dots < t_n < b$. Dados os pontos $Y_1, \dots, Y_n$, e o parâmetro suavizador estritamente positivo α , seja $\hat{g}$ uma spline cúbica natural com nós nos pontos $t_1, \dots, t_n$ para o qual:

$$\mathbf{g} = (\mathbf{I} + \alpha \mathbf{K})^{-1} \mathbf{Y}. \quad (4.12)$$

Então, para qualquer g em $S_2[a, b]$, $S(\hat{g}) \leq S(g)$, cuja igualdade só ocorre se $\hat{g}$ e g são idênticas.

Toda a teoria desenvolvida para demonstrar esse teorema está descrita em Alves (2003) e apresentada em Green e Silverman (1994).

Na prática é ineficiente usar diretamente $\mathbf{g} = (I + \alpha K)^{-1} \mathbf{Y}$ para calcular o vetor $\mathbf{g}$ e portanto a curva $\hat{g}$ (GREEN e SILVERMAN, 1994, p.19). Na próxima seção será desenvolvido um algoritmo a partir dessa definição.

4.3.5.1 O algoritmo de Reinsch

A idéia básica do algoritmo de Reinsch (1967-1971) é construir um sistema não singular de equações lineares para a segunda derivada γ_i de $\hat{g}$ nos nós t_i (GREEN e SILVERMAN, 1994, p.19).

Seja γ o vetor $(\gamma_2, \dots, \gamma_{n-1})^T$ como definido na seção 4.3.3.1. Substituindo (4.8) em (4.12) tem-se:

$$\mathbf{g} = (I + \alpha QR^{-1}Q^T)^{-1} \mathbf{Y} \Rightarrow \mathbf{Y} = (I + \alpha QR^{-1}Q^T) \mathbf{g}. \quad (4.13)$$

Rearranjando (4.13) tem-se:

$$\mathbf{g} = \mathbf{Y} - \alpha QR^{-1}Q^T \mathbf{g}.$$

Pelo teorema 4.1, $Q^T \mathbf{g} = R\gamma$, assim:

$$\mathbf{g} = \mathbf{Y} - \alpha QR^{-1}R\gamma = \mathbf{Y} - \alpha Q\gamma. \quad (4.14)$$

Aplicando Q^T em todos os membros e utilizando novamente o teorema 4.1 tem-se:

$$\begin{aligned} Q^T \mathbf{g} &= Q^T \mathbf{Y} - Q^T \alpha Q\gamma \Rightarrow R\gamma = Q^T \mathbf{Y} - Q^T \alpha Q\gamma \Rightarrow R\gamma + Q^T \alpha Q\gamma = Q^T \mathbf{Y} \Rightarrow \\ (R + \alpha Q^T Q)\gamma &= Q^T \mathbf{Y}. \end{aligned} \quad (4.15)$$

Esta equação é a essência do algoritmo de Reinsch (GREEN e SILVERMAN, 1994, p.20).

A matriz $(R + \alpha Q^T Q)$ é simétrica e positiva definida, visto que R e $Q^T Q$ satisfazem essas duas propriedades. Além disso, $(R + \alpha Q^T Q)$ tem tamanho de banda 5, pois $Q^T Q$ tem tamanho de banda 5. Assim, pode ser aplicada a decomposição de Cholesky em $(R + \alpha Q^T Q)$ da seguinte forma:

$$R + \alpha Q^T Q = LDL^T,$$

onde D é uma matriz diagonal estritamente positiva e L é uma matriz em banda triangular inferior com $L_{ij} = 0$ para $j < i - 2$ e $j > i$, e $L_{ii} = 1$ para todo i . Portanto, o algoritmo para determinar a spline suavizadora é descrito a seguir.

Algoritmo para a spline suavizadora

Passo 1: Determine o vetor $Q^T Y$, usando a fórmula (4.11).

Passo 2: Encontre as diagonais não nulas de $R + \alpha Q^T Q$, e portanto os fatores L e D da decomposição Cholesky.

Passo 3: Escreva (4.15) como $LDL^T \gamma = Q^T Y$ e resolva esta equação para γ .

Passo 4: Use (4.14) para encontrar g .

O **passo 1** precisa ser executado uma única vez para cada conjunto de dados. Ele não precisa ser repetido para um novo valor do parâmetro suavizador α . Além disso, se novos dados Y são usados, mas o *design* dos pontos permanece inalterado, o **passo 2** pode ser omitido (GREEN e SILVERMAN, 1994, p.21).

4.3.6 Suavização com splines de forma ponderada

Até agora, o termo de penalidade de aspereza, $\int g''^2$, tem sido adicionado à soma dos quadrados dos resíduos, $\sum \{Y_i - g(t_i)\}^2$, no MMQ com penalidades.

Nesta seção, será considerada uma forma mais geral, no qual os resíduos são ponderados.

Supondo que $w_1, \dots, w_n$ são pesos estritamente positivos, a soma dos quadrados dos resíduos ponderados é definida por (GREEN e SILVERMAN, 1994, p.40):

$$\sum_{i=1}^n w_i \{Y_i - g(t_i)\}^2 \quad (4.16)$$

Existem algumas aplicações em que é apropriado avaliar o ajuste da curva g nos pontos (t_i, Y_i) pela soma dos quadrados dos resíduos ponderados. Uma aplicação desse procedimento pode ser realizada em dados onde os Y_i são distribuídos com média $g(t_i)$, mas, as variâncias de Y_i não são iguais. Neste caso, é natural que os pesos sejam inversamente proporcionais a variância das observações.

4.3.6.1 Propriedades básicas da formulação ponderada

Seja W uma matriz diagonal com elementos w_i . Dada uma função g em $S_2[a, b]$, a soma dos quadrados penalizada ponderada $S_W(g)$ é definida por (GREEN e SILVERMAN, 1994, p.41):

$$S_W(g) = \sum_{i=1}^n w_i \{Y_i - g(t_i)\}^2 + \alpha \int g''^2, \quad (4.17)$$

onde, como usual, α é o parâmetro suavizador estritamente positivo, e $\hat{g}$ é o mínimo de $S_W(g)$.

Os resultados da seção 4.3.5 podem ser facilmente estendidos para a formulação ponderada. O resultado correspondente ao teorema 4.4 está descrito a seguir, cuja demonstração se encontra em Alves (2003) ou Green e Silverman (1994).

Teorema 4.5 - Seja $n \geq 3$ e $a < t_1 < \dots < t_n < b$. Suponha que o parâmetro suavizador α e os pesos w_i , $i = 1, \dots, n$ são todos estritamente positivos. Dados os valores $Y_1, \dots, Y_n$, a soma dos quadrados penalizada ponderada $S_W(g)$ é minimizada unicamente pela spline cúbica natural g em $S_2[a, b]$, sendo:

$$\mathbf{g} = (W + \alpha K)^{-1} WY \quad (4.18)$$

4.3.6.2 O algoritmo de Reinsch para suavização ponderada

É fácil modificar o algoritmo de Reinsch da seção 4.3.5.1 para incorporar a matriz de pesos W . De (4.18) tem-se (GREEN e SILVERMAN, 1994, p.41):

$$\begin{aligned} \mathbf{g} &= (W + \alpha K)^{-1} WY = (W + \alpha QR^{-1}Q^T)^{-1} WY \Rightarrow WY = (W + \alpha QR^{-1}Q^T) \mathbf{g} \\ &\Rightarrow W\mathbf{g} = WY - \alpha QR^{-1}Q^T \mathbf{g} \Rightarrow \mathbf{g} = Y - \alpha W^{-1}QR^{-1}Q^T \mathbf{g} \end{aligned}$$

Assim, como $\gamma = R^{-1}Q^T \mathbf{g}$, tem-se:

$$\mathbf{g} = Y - \alpha W^{-1}Q\gamma. \quad (4.19)$$

Multiplicando todos os elementos por Q^T e usando $Q^T \mathbf{g} = R\gamma$, tem-se:

$$Q^T \mathbf{g} = Q^T Y - \alpha Q^T W^{-1}Q\gamma \Rightarrow Q^T Y = R\gamma + \alpha Q^T W^{-1}Q\gamma \Rightarrow Q^T Y = (R + \alpha Q^T W^{-1}Q)\gamma. \quad (4.20)$$

Como W é uma matriz diagonal estritamente positiva definida, a matriz $(R + \alpha Q^T W^{-1}Q)$ é uma matriz em banda, com largura de banda 5 e positiva definida. Assim, pode ser aplicada a decomposição de Cholesky em $(R + \alpha Q^T W^{-1}Q)$ da seguinte forma:

$$(R + \alpha Q^T W^{-1}Q) = LDL^T$$

onde D é uma matriz diagonal estritamente positiva e L é uma matriz em banda triangular inferior com diagonal unitária. O algoritmo resultante, pode agora ser declarado (GREEN e SILVERMAN, 1994, p.42).

Algoritmo para spline suavizadora ponderada

Passo 1: Determine o vetor $Q^T W Y$, usando a fórmula (4.11).

Passo 2: Encontre as diagonais não nulas de $(R + \alpha Q^T W^{-1} Q)$, e portanto os fatores L e D da decomposição Cholesky.

Passo 3: Escreva (4.20) como $LDL^T \gamma = Q^T W Y$ e resolva esta equação para γ .

Passo 4: Use (4.19) para encontrar g .

É necessário ressaltar que essa seção 4.3.6 trata da suavização ponderada, mas nesse caso o peso é apenas diagonal. Nas aplicações com o GPS isso não acontece, ou seja, a matriz peso é cheia. Mas, apenas pequenas alterações devem ser feitas nessa seção para se adaptar ao caso do GPS. Na próxima seção a matriz peso já será tratada de acordo com a aplicação desse trabalho.

4.4 Modelo semiparamétrico

No modelo semiparamétrico, o cálculo das variáveis é dividido em duas partes: a parte paramétrica e a não-paramétrica. Normalmente, a parte paramétrica é a de interesse do usuário. No caso do GPS, ela pode ser as DD das ambigüidades e as coordenadas da estação. A parte não-paramétrica representa uma combinação de algumas funções de erros que variam suavemente com o tempo (JIA, STEWART e TSAKIRI, 2001).

O vetor semiparamétrico pode ser expresso por (JIA, STEWART e TSAKIRI, 2001):

$$y_i = A_i x + M_i g(t_i) + \varepsilon_i \quad i = 1, 2, \dots, n, \quad (4.21)$$

onde:

- $y_i \in \mathbb{R}^m$ são as observações na i -ésima época (DD da fase da onda portadora e/ou da pseudodistância);
- $A_i \in \mathbb{R}^{m \times p}$ é a matriz dos coeficientes da equação (matriz *design*);
- $x \in \mathbb{R}^p$ é o vetor dos parâmetros a ser estimado envolvendo as ambigüidades da fase da onda portadora e as coordenadas de interesse;
- $M_i \in \mathbb{R}^{m \times q}$ é a matriz de incidência (seção 4.4.1);
- $g(t_i) \in \mathbb{R}^q$ são as funções de erros;
- t_i é o índice de tempo;
- $\varepsilon_i \in \mathbb{R}^m$ é o vetor dos erros aleatórios na i -ésima época;
- n é o número de épocas;
- m é o número de observações por época;
- q é o número de funções de erros;
- p é o número de parâmetros de interesse a estimar.

4.4.1 Matriz de incidência

No modelo semiparamétrico é utilizada a matriz de incidência (denotada por M). Essa matriz é usada em casos onde os t_i não são distintos ou ordenados. Assim, a matriz M é construída da seguinte forma (GREEN e SILVERMAN, 1994, p.65):

- Denote os valores $t_1, t_2, \dots, t_n$ ordenados e distintos por $s_1, s_2, \dots, s_q$, com $q \leq n$;
- $m_{ij} = 1$ se $t_i = s_j$ e caso contrário $m_{ij} = 0$.

Dessa forma, M é uma matriz $n \times q$ que faz a conexão entre $t_1, t_2, \dots, t_n$ e $s_1, s_2, \dots, s_q$. Além disso, se os t_i são distintos e ordenados, a matriz de incidência é igual a identidade.

4.5 MMQ com Penalidades para o modelo semiparamétrico

A equação (4.21) contém $m \times n$ observações e $q \times n + p$ incógnitas. Dois casos podem ser considerados para tal equação. Primeiro, se o número de incógnitas é maior que o número de observações. Neste caso, a equação (4.21) não pode ser resolvida usando o MMQ convencional. Já no outro caso, mesmo com o número de incógnitas menor que o número de observações, a equação (4.21) pode não fornecer uma solução estável quando o MMQ convencional é usado, devido ao fato do número de incógnitas ser maior que o usual, pois as funções de erros ($g(t_i)$) também devem ser calculadas.

Para que se obtenha uma solução confiável, injunções adicionais, que correspondem à penalidade do ajustamento, devem ser adicionadas. Trata-se do MMQ com penalidades. Dessa forma, a função a ser minimizada é dada por (JIA, STEWART e TSAKIRI, 2001):

$$\sum_{i=1}^n \left\{ (y_i - A_i x - M_i g(t_i))^T \sum_i^{-1} (y_i - A_i x - M_i g(t_i)) \right\} + \sum_{j=1}^q \alpha_j \int (g_j''(t))^2 dt = \min, \quad (4.22)$$

onde $\sum_i^{-1}$ é a matriz peso das observações, α_j é o parâmetro suavizador e $g_j''(t)$ é a segunda derivada da j -ésima função em relação ao tempo.

A equação (4.22) define a forma quadrática penalizada. A primeira parte da equação (4.22) refere-se a forma quadrática dos resíduos do MMQ, e a segunda parte ao termo de penalidade de aspereza.

Mas, para desenvolver os cálculos computacionalmente, a penalidade de aspereza pode ser expressa por (FESSLER, 1991):

$$\sum_{j=1}^q \alpha_j \int (g_j(t))^2 dt = ((R^{-1} \otimes I_q)(Q^T \otimes I_q)g)^T (R \otimes D(\alpha))((R^{-1} \otimes I_q)(Q^T \otimes I_q))g, \quad (4.23)$$

onde:

- Q e R são matrizes relacionadas com o índice de tempo t_i (seção 4.3.3.1);
- I_q é a matriz identidade $q \times q$;
- $D(\alpha) = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_q)$;
- $\otimes$ denota o produto de Kronecker.

Dessa forma, substituindo (4.23) em (4.22), tem-se:

$$\begin{aligned} & \sum_{i=1}^n (y_i - A_i x - M_i g(t_i))^T \sum_i^{-1} (y_i - A_i x - M_i g(t_i)) + \\ & + ((R^{-1} \otimes I_q)(Q^T \otimes I_q)g)^T (R \otimes D(\alpha))((R^{-1} \otimes I_q)(Q^T \otimes I_q))g = \min. \end{aligned} \quad (4.24)$$

Minimizando (4.24) em relação a x e g , obtém-se respectivamente:

$$A^T (I_n \otimes \Sigma^{-1})Ax + A^T (I_n \otimes \Sigma^{-1})(I_n \otimes M)g = A^T (I_n \otimes \Sigma^{-1})y \quad (4.25)$$

e

$$\begin{aligned} & (I_n \otimes M^T)(I_n \otimes \Sigma^{-1})Ax + ((I_n \otimes M^T)(I_n \otimes \Sigma^{-1})(I_n \otimes M) + \\ & + (Q \otimes D(\alpha))(R^{-1} \otimes I_q)(Q^T \otimes I_q))g = (I_n \otimes M^T)(I_n \otimes \Sigma^{-1})y, \end{aligned} \quad (4.26)$$

sendo:

- $A = (A_1, A_2, \dots, A_n)^T$;
- I_n a matriz identidade $n \times n$;
- $\Sigma^{-1} = \Sigma_1^{-1} = \Sigma_2^{-1} = \dots = \Sigma_n^{-1}$;
- $y = (y_1, y_2, \dots, y_n)^T$;
- $g = (g_1^1, g_1^2, \dots, g_1^q, g_2^1, g_2^2, \dots, g_2^q, \dots, g_n^1, g_n^2, \dots, g_n^q)$.

As equações (4.25) e (4.26) podem ser resolvidas pelo método direto (GREEN e SILVERMAN, 1994, p.69), ou seja, substituindo a equação (4.26) em (4.25). Assim, utilizando o método direto em tais equações e realizando algumas manipulações matemáticas, tem-se:

$$\hat{x} = \left[A^T (I_n \otimes \Sigma^{-1}) A - A^T (I_n \otimes \Sigma^{-1}) A_S \right]^{-1} \left[A^T (I_n \otimes \Sigma^{-1}) y - A^T (I_n \otimes \Sigma^{-1}) y_S \right], \quad (4.27)$$

onde:

- $A_S = SA$ é a matriz *design* suavizada;
- $y_S = Sy$ é o vetor de observações suavizadas;
- S é a matriz de suavização dada por:

$$S = \frac{(I_n \otimes M) \left((I_n \otimes M^T) (I_n \otimes \Sigma^{-1}) (I_n \otimes M) + (Q \otimes D(\alpha)) (R^{-1} \otimes I_q) (Q^T \otimes I_q) \right)^{-1}}{(I_n \otimes M^T) (I_n \otimes \Sigma^{-1})} \quad (4.28)$$

Além disso, para estimar $\hat{g}$, usa-se:

$$\hat{g} = \frac{\left((I_n \otimes M^T) (I_n \otimes \Sigma^{-1}) (I_n \otimes M) + (Q \otimes D(\alpha)) (R^{-1} \otimes I_q) (Q^T \otimes I_q) \right)^{-1}}{(I_n \otimes M^T) (I_n \otimes \Sigma^{-1})} (y - A\hat{x}). \quad (4.29)$$

Convém salientar que, na teoria apresentada, está sendo considerado que o número de observações em cada época é o mesmo. Isto significa, no caso do GPS, que os mesmos satélites estão sendo rastreados. Se ocorrer mudanças, algumas alterações devem ser realizadas nessa seção.

5 PARÂMETRO SUAVIZADOR

Como foi apresentado na seção 4.2.3 (p.46), a escolha do parâmetro suavizador é muito importante no MMQ com penalidades. Portanto, esse capítulo tem como principal objetivo apresentar um método que encontre tal parâmetro suavizador, de forma que otimize a solução do problema.

5.1 Introdução

Existem duas filosofias diferentes para a escolha do parâmetro suavizador. No primeiro procedimento, considera-se a escolha livre do parâmetro suavizador um vantajoso critério. Variando o parâmetro suavizador, algumas características dos dados podem ser exploradas. Mas, se uma única estimativa é desejada, esta pode ser obtida através de uma escolha subjetiva. Já o outro procedimento se opõe ao primeiro pois, nesse caso, a escolha do parâmetro suavizador deve ser automática, isto é, calculado a partir dos dados (GREEN e SILVERMAN, 1994, p.29).

O procedimento automático é essencial em aplicações onde o método é aplicado para um grande número de conjuntos de dados, ou quando esse método é um componente de um processo mais complicado.

Como nesse trabalho será utilizado um método automático para escolher o parâmetro suavizador, as próximas seções descreverão os dois métodos mais conhecidos e utilizados para tal fim: a validação cruzada e a validação cruzada generalizada.

5.2 Validação cruzada

A validação cruzada (*Cross Validation* - CV) é uma técnica usada para estimar o erro de predição, para um modelo ajustado aos dados. O erro de predição mede a capacidade de um modelo prever a resposta de uma observação futura (MORETTIN, 1997, p.31).

A validação cruzada usa parte dos dados para estimar o modelo e o restante para avaliar se ele é adequado ou não. Usualmente, o algoritmo de validação cruzada é o seguinte (MORETTIN, 1997, p.31):

- Dadas n observações $y_1, \dots, y_n$, ajusta-se o modelo para cada conjunto de observações, deixando a observação y_i de fora, e calcula-se o valor previsto para esta i -ésima observação, denotado por $\hat{y}_i^{-i}$;
- Calcula-se $CV = n^{-1} \sum_{i=1}^n \{y_i - \hat{y}_i^{-i}\}^2$.

Assim, a técnica de validação cruzada cria uma situação de “nova observação” para conjuntos de dados onde a mesma não está disponível. Um exemplo disso ocorre quando o método de suavização é utilizado para um único conjunto de dados (GREEN e SILVERMAN, 1994, p.30). O desenvolvimento matemático da validação cruzada necessário neste trabalho é descrito a seguir.

Seja α o parâmetro suavizador. Considere a observação Y_i de t_i como sendo uma nova observação, e a omita do conjunto de dados utilizado para estimar a curva. Denote por $\hat{g}^{(-i)}(t; \alpha)$ a curva estimada dos dados restantes, usando α para o parâmetro suavizador, de modo que $\hat{g}^{(-i)}(t; \alpha)$ seja o mínimo de:

$$\sum_{j \neq i} \{Y_j - g(t_j)\}^2 + \alpha \int g''^2. \quad (5.1)$$

Como a escolha da observação que será omitida é arbitrária, a eficácia global do procedimento com o parâmetro suavizador α pode ser quantificada pela função de validação cruzada dada por (GREEN e SILVERMAN, 1994, p.30):

$$CV(\alpha) = n^{-1} \sum_{i=1}^n \{Y_i - \hat{g}^{(-i)}(t_i; \alpha)\}^2. \quad (5.2)$$

A idéia básica da validação cruzada é escolher o valor de α que minimize $CV(\alpha)$. Como não há garantia de que a função CV tenha um único mínimo, cuidados devem ser tomados com sua minimização. Uma rede de procura é, provavelmente, o melhor caminho a ser seguido. Além disso, qualquer método de minimização que for utilizado envolverá o cálculo de $CV(\alpha)$ para um número de valores de α e, por isso, é importante que se utilize um método eficiente para o cálculo de CV (GREEN e SILVERMAN, 1994, p.30).

5.2.1 Cálculo da função de validação cruzada

Num primeiro momento, observando (5.2), parece que para calcular a função $CV(\alpha)$ é necessário resolver n problemas de suavização separadamente, para encontrar n curvas $\hat{g}^{(-i)}$. Entretanto, será visto nessa seção que isto não é necessário.

O primeiro passo na simplificação é recordar do teorema 4.4 (p.57) que os valores da spline suavizadora $\hat{g}$ dependem linearmente dos dados Y_i , pela equação:

$$\mathbf{g} = H(\alpha)\mathbf{Y}, \quad (5.3)$$

onde a matriz $H(\alpha)$ é definida por:

$$H(\alpha) = (I + \alpha QR^{-1}Q^T)^{-1}. \quad (5.4)$$

Um resultado muito importante no desenvolvimento do cálculo da função de validação cruzada é descrito no próximo teorema (GREEN e SILVERMAN, 1994, p.31).

Teorema 5.1 - Os valores da validação cruzada satisfazem

$$CV(\alpha) = n^{-1} \sum_{i=1}^n \left(\frac{Y_i - \hat{g}(t_i)}{1 - H_{ii}(\alpha)} \right)^2, \quad (5.5)$$

onde $\hat{g}$ é a spline suavizadora calculada através do conjunto de dados $\{(t_i, Y_i)\}$ com parâmetro suavizador α .

A prova deste teorema será realizada através de um lema e de um corolário que seguem no apêndice B. Esse teorema também é demonstrado em Alves (2003) e Green e Silverman (1994).

Este teorema mostra que se os valores $H_{ii}(\alpha)$ da diagonal de H são conhecidos, a validação cruzada pode ser calculada dos resíduos $Y_i - \hat{g}(t_i)$. Mas, de (5.4), pode-se perceber que o cálculo de $H_{ii}(\alpha)$ não é uma tarefa muito fácil, pois existem várias inversões envolvidas em sua definição. Assim, a próxima seção apresentará um algoritmo que calcula $H_{ii}(\alpha)$ minimizando esse esforço computacional.

5.2.2 Encontrando os elementos da diagonal de $H(\alpha)$

Existem dois componentes importantes desse método: o primeiro encontra as diagonais centrais da inversa de uma matriz em banda, e o segundo reescreve a matriz $H(\alpha)$ de forma que ela possa ser calculada sem inverter matrizes completas. Ambos os componentes podem ser aplicados para matrizes de larguras de banda arbitrárias. No entanto, somente o caso especial em que a largura de banda é 5, o qual coincide com as splines cúbicas naturais, será descrito.

5.2.2.1 As diagonais centrais da inversa de uma matriz em banda

Suponha que B é uma matriz em banda simétrica positiva definida, com largura de banda 5, de forma que os elementos (i, j) de B são nulos se $|i - j| > 2$. Para encontrar B^{-1} são necessárias $O(n^2)$ operações. Mas, existe um algoritmo para calcular apenas as cinco diagonais centrais de B^{-1} , que será descrito a seguir (GREEN e SILVERMAN, 1994, p.33).

Decomponha B em $B = LDL^T$ onde L é uma matriz em banda triangular inferior com diagonal unitária, e D é uma matriz diagonal com elementos d_i . Suponha que B^{-1} tem elementos $\bar{b}_{ij}$. Então, pela definição:

$$B^{-1} = L^{-T} D^{-1} L^{-1} \Rightarrow L^T B^{-1} = D^{-1} L^{-1}. \quad (5.6)$$

Assim, somando B^{-1} dos dois lados em (5.6) tem-se:

$$B^{-1} = D^{-1} L^{-1} + B^{-1} - L^T B^{-1} = D^{-1} L^{-1} + (I - L^T) B^{-1}. \quad (5.7)$$

A matriz L^{-1} é triangular inferior com diagonal unitária, então, $D^{-1} L^{-1}$ é triangular inferior com elementos d_i^{-1} na diagonal. Além disso, $(I - L^T)$ é triangular superior com diagonal nula. Considerando os elementos da diagonal principal e das duas diagonais superiores, para $i = 1, \dots, n-2$, tem-se:

$$\bar{b}_{i,i} = d_i^{-1} - l_{i+1,i} \bar{b}_{i+1,i} - l_{i+2,i} \bar{b}_{i+2,i}$$

$$\bar{b}_{i,i+1} = -l_{i+1,i} \bar{b}_{i+1,i+1} - l_{i+2,i} \bar{b}_{i+2,i+1}$$

$$\bar{b}_{i,i+2} = -l_{i+1,i} \bar{b}_{i+1,i+2} - l_{i+2,i} \bar{b}_{i+2,i+2}.$$

Além disso:

$$\begin{cases} \bar{b}_{n,n} = d_n^{-1} \\ \bar{b}_{n-1,n} = -l_{n,n-1} \bar{b}_{n,n} \\ \bar{b}_{n-1,n-1} = d_{n-1}^{-1} - l_{n,n-1} \bar{b}_{n-1,n} \end{cases}$$

Estas fórmulas podem ser usadas para calcular as cinco diagonais centrais de B^{-1} . A ordem na qual os elementos podem ser encontrados é: (n, n) , $(n-1, n)$, $(n-2, n)$; $(n-1, n-1)$, $(n-2, n-1)$, $(n-3, n-1)$; ...; $(2, 2)$, $(1, 2)$; $(1, 1)$. Pode ser ressaltado que os elementos de B^{-1} fora das 5 diagonais, em geral, não são nulos, mas esses elementos não são relevantes na realização deste trabalho (GREEN e SILVERMAN, 1994, p.34).

5.2.2.2 Matriz $H(\alpha)$ otimizada

Das equações (4.14) e (4.15) do algoritmo de Reinsch na seção 4.3.5.1 (p.58), a spline suavizadora é definida pelos vetores $\mathbf{g}$ e γ , que satisfazem:

$$\gamma = (R + \alpha Q^T Q)^{-1} Q^T Y.$$

$$\mathbf{g} = Y - \alpha Q \gamma = Y - \alpha Q (R + \alpha Q^T Q)^{-1} Q^T Y = \{I - \alpha Q (R + \alpha Q^T Q)^{-1} Q^T\} Y. \quad (5.8)$$

Pode ser visto de (5.8) que $H(\alpha)$ tem uma expressão alternativa:

$$H(\alpha) = \{I - \alpha Q (R + \alpha Q^T Q)^{-1} Q^T\}.$$

Assim (GREEN e SILVERMAN, 1994, p.34):

$$I - H(\alpha) = \alpha Q (R + \alpha Q^T Q)^{-1} Q^T. \quad (5.9)$$

Seja B uma matriz em banda simétrica dada por $(R + \alpha Q^T Q)$, que tem tamanho de banda 5. Como Q é uma matriz tridiagonal, para $i = 1, \dots, n$ segue que:

$$\begin{aligned} (QB^{-1}Q^T)_{i,i} &= q_{i,i-1}^2 \bar{b}_{i-1,i-1} + q_{i,i}^2 \bar{b}_{i,i} + q_{i,i+1}^2 \bar{b}_{i+1,i+1} + 2q_{i,i-1}q_{i,i} \bar{b}_{i-1,i} + 2q_{i,i-1}q_{i,i+1} \bar{b}_{i-1,i+1} + \\ &+ 2q_{i,i}q_{i,i+1} \bar{b}_{i,i+1}. \end{aligned} \quad (5.10)$$

É claro, de (5.10), que somente os elementos $\bar{b}_{ij}$ com $|i - j| \leq 2$ precisam ser conhecidos para calcular os elementos da diagonal de $QB^{-1}Q^T$, e portanto, os elementos $I - H_{ii}(\alpha)$ necessários para calcular os valores da validação cruzada.

Para mais detalhes em relação ao desenvolvimento das equações das seções 5.2.2.1 e 5.2.2.2, consulte Alves (2003).

5.3 Validação Cruzada Generalizada

A validação cruzada generalizada (*Generalized Cross Validation* - GCV), uma forma modificada da validação cruzada, é um método popular para encontrar o parâmetro suavizador (GREEN e SILVERMAN, 1994, p.35).

A idéia básica da GCV é substituir o fator $1 - H_{ii}(\alpha)$ pela sua média, $1 - n^{-1}trH(\alpha)$. Assim, a GCV é construída por analogia com a validação cruzada (GREEN e SILVERMAN, 1994, p.35). Dessa forma:

$$GCV(\alpha) = n^{-1} \frac{\sum_{i=1}^n \{Y_i - \hat{g}(t_i)\}^2}{\{1 - n^{-1}trH(\alpha)\}^2}. \quad (5.11)$$

Da mesma maneira que na CV, a escolha do parâmetro suavizador na GCV é realizada pela minimização da função $GCV(\alpha)$ em α .

Se todos os $H_{ii}(\alpha)$ são iguais, por exemplo, se todos os t_i são igualmente espaçados, a GCV seria idêntico a $CV(\alpha)$. Mas, geralmente, existirão algumas diferenças entre as duas aproximações.

5.3.1 Exemplo da utilização da GCV

Como foi apresentado na seção 4.2.3 (p.46), o valor de α interfere de forma significativa no ajuste da curva ao conjunto de dados. Dessa forma, um algoritmo que determine o valor de α de forma automática, como a GCV, é indispensável neste trabalho.

Para ilustrar a potencialidade desse método, um algoritmo para a GCV foi implementado em linguagem FORTRAN 77/90. Para tanto foi utilizado um algoritmo desenvolvido por Hutchinson (1986).

Para o desenvolvimento desse exemplo foram utilizados dados coletados por um receptor Trimble 4000 SSI na estação UEPP (estação base) da RBMC e por um receptor Trimble 4600 LS na estação EP02 do IBGE. Ambas localizadas na FCT/UNESP. A taxa de coleta foi de 15 segundos e o ângulo de elevação de 15°.

Nas figuras 13 e 14 são plotadas as splines cúbicas naturais obtidas pelo algoritmo de Reinsch (seção 4.3.5.1, p.58) utilizando α que minimiza a GCV (seção 5.3). Como dados de entrada foram utilizadas as DD da pseudodistância, para os satélites 2 e 15 (DD^{2-15}) e 2 e 31 (DD^{2-31}).

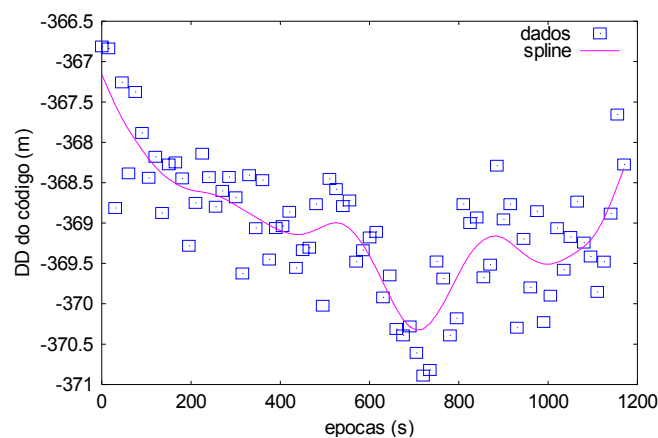


FIGURA 13 - Spline suavizadora para a DD^{2-15} do código

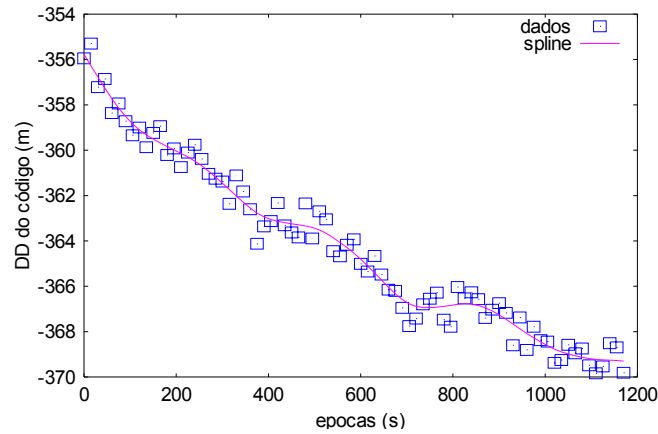


FIGURA 14 - Spline suavizadora para a DD^{2-31} do código

Note que como esperado a spline cúbica natural ajusta o conjunto de dados de forma suave, acompanhando a tendência dos dados.

5.4 Validação cruzada e validação cruzada generalizada para suavização ponderada

O método de validação cruzada é facilmente estendido ao caso com ponderação. Como usual, seja $\hat{g}^{(-i)}(t_i)$ a curva estimada omitindo os pontos (t_i, Y_i) . Assim, a função de validação cruzada (5.2) é modificada para (GREEN e SILVERMAN, 1994, p.42):

$$CV(\alpha) = n^{-1} \sum_{i=1}^n w_i \{Y_i - \hat{g}^{(-i)}(t_i; \alpha)\}^2, \quad (5.12)$$

que será minimizada pela escolha de α .

Da mesma forma que no caso não ponderado, não é necessário resolver n problemas de suavização para encontrar $CV(\alpha)$. Seja $H_W(\alpha)$ uma matriz para o caso ponderado. Usando os mesmos argumentos do teorema 5.1, a validação cruzada pode ser escrita por:

$$CV(\alpha) = n^{-1} \sum_{i=1}^n w_i \left(\frac{Y_i - \hat{g}(t_i)}{\{I - H_w(\alpha)\}_{ii}} \right)^2, \quad (5.13)$$

onde $\hat{g}$ é o mínimo da soma dos quadrados penalizada ponderada.

Para a GCV, a equação (5.11) é modificada para:

$$GCV(\alpha) = n^{-1} \frac{\sum_{i=1}^n w_i \{Y_i - \hat{g}(t_i)\}^2}{\{1 - n^{-1} \text{tr} H(\alpha)\}^2}. \quad (5.14)$$

Note que nessa seção, novamente, a matriz peso é apenas diagonal. Na próxima seção a teoria será estendida para o caso onde a matriz peso é cheia.

5.5 Validação cruzada generalizada no modelo semiparamétrico

Como nessa pesquisa é utilizada a GCV, será realizada uma discussão detalhada da utilização dessa função no modelo semiparamétrico.

Reescrevendo (5.14) para o modelo semiparamétrico, tem-se (JIA, STEWART e TSAKIRI, 2001; FESSLER, 1991):

$$GCV(\alpha) = \frac{1}{n} \frac{V^T (I_n \otimes \Sigma^{-1}) V}{\left\{ \frac{1}{n} \text{tr}(I_{nm} - H) \right\}^2}, \quad (5.15)$$

onde:

- n é o número de épocas;
- m é o número de observações por época;
- V é o vetor de resíduos;
- Σ^{-1} é a matriz peso das observações;

- $H(\alpha) = S + (I_{nm} - S)A(A^T(I_n \otimes \Sigma^{-1})(I_{nm} - S)A)^{-1}A^T(I_n \otimes \Sigma^{-1})(I_{nm} - S)$, sendo S a matriz dada pela equação (4.28) na seção 4.5 (p.66).

De acordo com a equação (4.22) da seção 4.5 (p.64), no modelo semiparamétrico o resíduo é dado por:

$$V_i = y_i - A_i x - M_i g(t_i).$$

Assim:

$$GCV(\alpha) = \frac{1}{n} \frac{(y - A\hat{x} - (I_n \otimes M)\hat{g})^T (I_n \otimes \Sigma^{-1})(y - A\hat{x} - (I_n \otimes M)\hat{g})}{\left\{ \frac{1}{n} tr(I_{nm} - H) \right\}^2}$$

$$\Rightarrow GCV(\alpha) = \frac{1}{n} \frac{(y - A\hat{x} - (I_n \otimes M)\hat{g})^T (I_n \otimes \Sigma^{-1})(y - A\hat{x} - (I_n \otimes M)\hat{g})}{(m - trH / n)^2}, \quad (5.16)$$

No modelo semiparamétrico o número de graus de liberdade é dado por (GREEN e SILVERMAN, 1994, p. 71):

$$graus\ de\ liberdade = tr\{I_{nm} - H\}. \quad (5.17)$$

Além disso, com analogia ao modelo paramétrico, a variância a posteriori do modelo semiparamétrico é dada por:

$$\hat{\sigma}^2 = \frac{V^T (I_n \otimes \Sigma^{-1}) V}{tr(I_{nm} - H)}. \quad (5.18)$$

Assim, utilizando (5.16), (5.17) e (5.18) a função GCV é dada por:

$$GCV(\alpha) = \frac{n}{grau\ de\ liberdade} \times \hat{\sigma}^2. \quad (5.19)$$

Da mesma forma, a escolha do parâmetro suavizador na GCV é realizada pela minimização da função GCV em α . Por isso, um método de minimização deve ser utilizado. Uma alternativa é utilizar o método *Gold Search* (PRESS et al, 1986). Mais informações sobre esse método de minimização serão apresentadas no capítulo 6, seção 6.3.

6 SOFTWARES DISPONÍVEIS E IMPLEMENTAÇÃO COMPUTACIONAL

Neste capítulo será realizada uma breve descrição do software GAS, que foi utilizado para estimar as coordenadas “verdadeiras” dos dados GPS coletados. Além disso, serão abordados os conceitos fundamentais do software GPSeq que foi utilizado para o processamento dos dados GPS, bem como alguns aspectos da implementação do MMQ com penalidades e do modelo semiparamétrico efetuada no GPSeq.

6.1 O software GAS

O *GPS Analysis Software* (GAS) é um software desenvolvido pela Universidade de Nottingham (UK). Esse software foi projetado principalmente para o processamento de redes GPS usando DD com dados de dupla frequência. Para realizar tal processamento, o GAS utiliza como dados de entrada os arquivos de efemérides transmitidas no formato *Receiver Independent Exchange Format* (RINEX) (IGS, 2003), ou efemérides precisas fornecidas pelo *International GPS Service* (IGS) no formato SP3, e os arquivos de observação no formato RINEX.

No software GAS são utilizados os seguintes módulos (STEWART, 1997, p. 4-1):

- **FILTER** – A função primária do FILTER é converter dados GPS RINEX para dados no formato NOT, pois um outro módulo do GAS aceitará dados apenas no formato NOT. Durante esta conversão o FILTER também detectará e corrigirá as grandes perdas de ciclo nos dados da fase da onda portadora. O FILTER também tem a capacidade de realizar o

posicionamento por ponto através das pseudodistâncias se o usuário desejar checar as coordenadas aproximadas dadas no cabeçalho dos arquivos RINEX;

- **PANIC** – O *Program for the Analysis of Networks using Interferometric Carrier phase or Code* (PANIC) é o principal módulo de processamento de dados GPS utilizando o software GAS. O formato NOT dos arquivos de dados e o formato SP3 das efemérides são lincados pelo programa, via um arquivo de controle ASCII. O programa produz várias soluções e arquivos de saída no formato padrão ASCII. O PANIC pode operar em vários módulos, incluindo a correção de perdas de ciclo, resolução da ambigüidade e posicionamento cinemático. Adicionalmente, muitas outras opções estão disponíveis, como, por exemplo, a modelagem atmosférica;
- **SLIPCOR** – O *Slip Corrections* (SLIPCOR) é utilizado juntamente com o PANIC no módulo da detecção de perdas de ciclo. O módulo do PANIC de detecção de perdas de ciclo produz um arquivo de perdas de ciclo de uma determinada estação; este arquivo é denominado arquivo “slip”. Depois que todas as perdas de ciclo foram detectadas no arquivo slip, os arquivos de dados podem ser corrigidos utilizando o SLIPCOR. O SLIPCOR corrige os dados GPS no formato NOT das perdas de ciclo listadas no arquivo slip, pois todos os arquivos devem ser corrigidos antes do ajuste de rede realizado pelo PANIC.

O processamento dos dados no software GAS pode ser dividido nas seguintes fases (STEWART, 1997, p. 3-1):

- **Pré-processamento (Filter)** – Os arquivos de efemérides e de dados GPS no formato padrão RINEX são convertidos para dados no formato NOT. Neste estágio, é possível descartar épocas que contenham menos que um certo número de satélites, definir um novo intervalo de dados, por exemplo, dados coletados a cada 15 segundos para dados coletados a cada 30 segundos, remover dados de satélites que apresentem algum problema, detectar e

corrigir grandes saltos de ciclos na fase de batimento da onda portadora, especificar a janela de processamento para os arquivos de observação e realizar o posicionamento por ponto através das pseudodistâncias;

- **Deteção de perdas de ciclo (Panic, Slipcor)** – Este é o estágio que utiliza o principal módulo de processamento GPS, o **PANIC**, em seu módulo de detecção de perda de ciclo. A saída do PANIC é um arquivo de perdas de ciclos (o arquivo slip) que o usuário então tem a opção de editar manualmente. Através desse arquivo é possível saber quantas iterações são necessárias no PANIC. O programa corrige internamente as perdas de ciclo contidas no arquivo slip e procura novas perdas de ciclos, atualizando o arquivo slip quando necessário. Este procedimento é realizado iterativamente até que perdas de ciclos não sejam mais encontradas. O arquivo slip final é então utilizado para corrigir os arquivos de dados GPS através do módulo SLIPCOR;
- **Solução da rede (PANIC)** – Todos os arquivos de dados corrigidos são arquivos de entrada no PANIC, o qual é executado para realizar a solução da rede. Neste estágio, muitas opções adicionais, como modelagem atmosférica e resolução da ambigüidade, estão disponíveis para produzir as melhores coordenadas finais da estação.

6.2 O Software GPSeq

O GPSeq é um software para processamento de dados oriundos de posicionamento relativo de bases curtas com o NAVSTAR-GPS. Está sendo desenvolvido na FCT/UNESP em linguagem FORTRAN 77 (Compilador Lahey Fujitsu Fortran 95) e realiza um ajustamento recursivo utilizando como observáveis as DD da fase de batimento da onda

portadora e da pseudodistância a partir do código C/A (MACHADO e MONICO, 1999, 2002).

Inicialmente, o software utilizava o método sequencial para solucionar as ambigüidades e coordenadas, atendendo apenas aplicações estáticas. Tal software foi posteriormente modificado com a implementação do método *Least Squares Ambiguity Decorrelation Adjustment* (LAMBDA), desenvolvido na Universidade de Tecnologia de Delft, Holanda, para a solução das ambigüidades (MACHADO, 2001, p.76). Este método apresenta a possibilidade da estimação rápida dos números inteiros de ciclos, tendo em vista que faz uma transformação no conjunto das ambigüidades. Isto possibilita a reformulação do problema original em um novo problema, que por sua vez é muito mais fácil de resolver (MACHADO e MONICO, 1999, 2002).

Além disso, foi implementada por Machado (2001) uma estratégia de processamento que permite a utilização do software tanto para processamentos de dados coletados através de posicionamento relativo cinemático, quanto para relativo estático (MACHADO, 2001, p.76). Também implementou-se uma rotina de controle de qualidade, a qual está baseada no processo *Detection, Identification and Adaptation* (DIA) (TEUNISSEN, 1998b). A parte de adaptação encontra-se, ainda, em fase de implementação.

No presente momento, o software permite processar apenas linhas de base rastreadas com receptores de simples frequência. Uma versão para receptores de dupla frequência está em fase de testes.

6.2.1 Dados de entrada

Os dados de entrada devem estar no formato RINEX, permitindo que dados de receptores de fabricantes distintos possam ser processados (MACHADO e MONICO, 1999).

Além dos arquivos contendo as observações e os elementos necessários para calcular a posição dos satélites, o software requer um arquivo de entrada contendo o endereço de onde se encontram os arquivos RINEX, o valor limite para assumir que as observações coletadas pelos receptores base e móvel sejam simultâneas, bem como a precisão das observações de fase de batimento da onda portadora (L1) e pseudodistância (C/A) e o número *Pseudo Random Noise* (PRN) do satélite base (MACHADO, 2001, p.77). Um exemplo de arquivo pode ser visto na tabela 04.

TABELA 04 - Arquivo de entrada do software GPSeq

c:\meusdocumentos\dados\2003\uepp0360.03o	rinex obs file station 1 (base)
c:\meusdocumentos\dados\2003\05790360.03o	rinex obs file station 2 (rover)
c:\meusdocumentos\dados\2003\uepp0360.03n	rinex navigation file (base)
0.25	synchronization
0.003	standard deviation L1
0.3	standard deviation C/A
31	base satellite

6.2.2 Modelos matemáticos e estratégias adotadas

Nessa seção serão apresentados os modelos matemáticos e as estratégias de implementação adotados no desenvolvimento do software GPSeq.

6.2.2.1 Modelo funcional e estocástico

O modelo funcional implementado no software é o das DD da fase de batimento da onda portadora e da pseudodistância. A escolha desse modelo matemático se deve à presença dos efeitos sistemáticos que degradam a precisão das observáveis GPS, os quais, se não forem adequadamente tratados, tornam tais observáveis com nível de precisão inferior ao comprimento de onda, impossibilitando que as ambigüidades sejam estimadas como número inteiro (MACHADO, 2001, p.78). Utilizando as DD, a maior parte desses erros é significativamente reduzida para linhas de base curtas.

Conforme visto na seção 2.4.1 (p.22), as DD podem ser calculadas através de duas técnicas: satélite base e seqüencial. Na implementação do software GPSeq, as DD são calculadas a partir da técnica do satélite base, onde combina-se as observações oriundas de um dos satélites rastreados, o base, com as observações dos demais satélites. Porém, se ocorrer do sinal do satélite escolhido como base não ser, por algum motivo, observado em alguma das estações envolvidas no levantamento, não será possível calcular as DD, mostrando uma deficiência da técnica adotada.

No que se refere ao modelo estocástico, adotou-se o modelo (2.15) (p.24). Esse modelo tem sido muito utilizado no decorrer dos anos, apresentando resultados satisfatórios na maioria das aplicações (MACHADO, 2001, p.79).

6.2.2.2 Cálculo das coordenadas dos satélites

As coordenadas dos satélites GPS no software GPSeq são calculadas através das efemérides transmitidas.

O software calcula as coordenadas dos satélites e suas respectivas velocidades apenas para uma das estações envolvidas nas DD, como, por exemplo, para a estação base. Conhecendo-se as coordenadas e velocidades dos satélites para a estação base, bem como a diferença dos instantes de recepção entre as duas estações, pode-se calcular as coordenadas vinculadas à outra estação de interesse.

Para determinar as coordenadas do satélite, tendo em vista a dependência existente entre a posição dos satélites e o tempo, o instante de recepção do sinal é corrigido dos erros do relógio dos satélites e do receptor (MACHADO, 2001, p.81).

Outro efeito que é considerado no cálculo das coordenadas do satélite é a rotação da Terra durante o intervalo de tempo de propagação do sinal, pois as coordenadas dos satélites são calculadas num sistema fixo à Terra (WGS-84 realização G-1150).

6.2.2.3 Estimação recursiva e controle de qualidade

Devido ao número de observações, DD de pseudodistância e da fase de batimento da onda portadora, ser maior que o número de incógnitas ao final do levantamento, necessita-se ajustar as observações para se estimar o vetor dos parâmetros e calcular sua MVC, a partir da coerência entre as observações, o modelo matemático e os parâmetros a serem estimados (MACHADO, 2001, p.84).

Entretanto, a grande quantidade de observações que geralmente é coletada no posicionamento com NAVSTAR-GPS faz com que o Filtro de Kalman torne-se uma ferramenta bastante adequada ao processo de estimação dos parâmetros para a solução *float*, sendo, portanto, adotado no desenvolvimento do software.

Na implementação do software GPSeq, o controle de qualidade do Filtro de Kalman é realizado através do processo DIA (TEUNISSEN, 1998b).

6.2.2.4 Solução das ambigüidades e validação

Na estimativa das DD das ambigüidades, utiliza-se o método LAMBDA, o qual requer como dados de entrada as DD de ambigüidades estimadas como números reais (solução *float*) e a respectiva MVC, ambos provenientes do Filtro de Kalman (MACHADO, 2001, p.85).

No processo de estimação pelo Filtro de Kalman, os parâmetros estimados na última época representam a melhor solução, pois esse estimador considera as informações de todas as épocas anteriores. Portanto, adotou-se como estratégia para solução das DD de ambigüidades utilizar a solução *float* da última época como dados de entrada para o LAMBDA.

Antes de aceitar a solução de números inteiros, recomenda-se verificar a qualidade desses parâmetros através do processo de validação da solução da ambigüidade. Portanto, pode ser realizado um teste de discriminação. Esse teste consiste em verificar quanto o vetor que produz a melhor solução das ambigüidades ($\tilde{n}_1$) é diferente do vetor que produz a segunda melhor solução ($\tilde{n}_2$).

Um dos testes de discriminação mais conhecidos e utilizados é o teste *ratio*, o qual consiste na razão entre a variância a posteriori de $\tilde{n}_2$ e $\tilde{n}_1$, isto é (TEUNISSEN, 1998a, p. 330):

$$\tilde{\sigma}_2^2 / \tilde{\sigma}_1^2 = \Omega_2 / \Omega_1 > e_r, \quad (6.1)$$

onde:

- $\tilde{\sigma}_2^2$ e $\tilde{\sigma}_1^2$ representam a variância a posteriori de $\tilde{n}_2$ e $\tilde{n}_1$, respectivamente;
- Ω_1 e Ω_2 representam a forma quadrática para $\tilde{n}_2$ e $\tilde{n}_1$, respectivamente;
- e_r é um valor crítico que pode ser definido empiricamente.

Em Jia, Stewart e Tsakiri (2001) é considerado um valor crítico de 1,5 para a validação da ambigüidade.

Embora alguns autores assumam que o teste da equação (6.1) tenha distribuição F, isso não representa a realidade, pois Ω_1 e Ω_2 não são estatisticamente independentes.

6.3 Implementação do MMQ com Penalidades e do modelo semiparamétrico no GPSeq

O MMQ com Penalidades juntamente com o modelo semiparamétrico foi implementado no software GPSeq disponível na FCT/UNESP. Para tanto, o algoritmo foi desenvolvido em linguagem FORTRAN 77 (Compilador Lahey Fujitsu Fortran 95) e realiza o ajustamento utilizando como observáveis as DD da fase de batimento da onda portadora e da pseudodistância a partir do código C/A (MACHADO e MONICO, 1999, 2002).

Os dados utilizados são de receptores de simples frequência. Para a solução da ambigüidade é utilizado o método LAMBDA, já implementado no GPSeq.

A figura 15 sumariza as principais etapas realizadas na implementação do método.



FIGURA 15 - Etapas utilizadas na implementação do algoritmo do MMQ com Penalidades

A etapa que diz respeito à utilização do método direto para estimar os valores ajustados de x (coordenadas e DD das ambigüidades) e g (spline suavizadora), utiliza as equações desenvolvidas na seção 4.5. A equação (4.27) (p.66) é utilizada para estimar os valores de x , e a equação (4.29) (p.66) é utilizada para estimar os valores de g .

Em relação ao cálculo da função GCV, foi utilizada a equação (5.16) da seção 5.5 (p.77). Mas, como relatado anteriormente, o parâmetro suavizador (α) deve ser aquele que minimize a função GCV. Portanto, um método de minimização que calcule novos valores para α foi utilizado. Nesta pesquisa o método de otimização utilizado foi o *Gold Search* (PRESS et al, 1986).

O método *Gold Search* encontra o mínimo de uma função unidimensional. Para utilizar esse método foram implementadas duas rotinas principais:

- **Inicialização** - Dada a função GCV e 2 valores iniciais para cada parâmetro suavizador (definidos pelo usuário), esta rotina pesquisa a direção que minimiza a função e retorna 3 novos valores de α que limitam o mínimo da função. Também retorna o valor da GCV nesses 3 pontos;
- **Minimização** - Dada a função GCV e os 3 pontos calculados pela rotina anterior, essa rotina determina o valor de α que minimiza a função GCV.

Além disso, na implementação, foi desenvolvida uma rotina que retorna os parâmetros ajustados (x e g) e o valor da função GCV. Como ilustrado na figura 15, essa rotina é utilizada várias vezes durante o ajustamento, até que o α seja escolhido de forma que minimize a função GCV pelo método *Gold Search*. Essa rotina é a que dispende maior tempo no processamento dos dados, visto que a maior parte dos cálculos é realizada nela. Na seção 7.3 (p.111) é descrito o intervalo de tempo utilizado no processamento dos dados.

7 COLETA DE DADOS, RESULTADOS E ANÁLISES

Uma vez que o método proposto nessa dissertação teve os fundamentos teóricos descritos e sua implementação efetivada, a próxima etapa do trabalho envolveu a realização de alguns experimentos, objetivando analisar a performance do mesmo.

7.1 Coleta de dados

Os dados foram coletados com receptores de simples (Trimble 4600 LS) e de dupla (Ashtech ZXII) frequência (figuras 16 e 17), todos disponíveis na FCT/UNESP.



FIGURA 16 - Receptor Trimble 4600 LS utilizado no levantamento



FIGURA 17- Antena e receptor Ashtech ZXII utilizados no levantamento

Como estação base, foi adotada a estação UEPP (localizada no Campus da Unesp de Presidente Prudente) da RBMC, que é mantida pelo IBGE (FORTES, 1997, p.7). Essa estação dispõe de um receptor de dupla frequência Trimble 4000 SSI, cuja antena é ilustrada na figura 18.



FIGURA 18 - Estação UEPP que foi adotada como estação base

O primeiro experimento teve como objetivo verificar a eficácia do método proposto na redução do multicaminho. Portanto, utilizou-se parte dos dados coletados por Souza (2004), cujo objetivo principal do trabalho era a redução do multicaminho. Esses dados foram coletados em uma empresa de Presidente Prudente (Takigawa) com uma linha de base de aproximadamente 2 km. Inicialmente, os dados foram coletados nas proximidades de uma superfície refletora (uma carreta), que distava 6 m dos receptores, estava a 1,3 m do chão e tinha 13 m e 2,5 m de comprimento e altura respectivamente (figura 19). Posteriormente, a carreta foi retirada e uma outra coleta de dados foi realizada no mesmo local e horário. Em ambos os casos, foram coletados aproximadamente 3 h de dados, nos dias 13 e 20 de setembro de 2003, respectivamente. Esse experimento será denominado UEPP-Takigawa.



FIGURA 19 - Coleta de dados com superfície refletora realizada na empresa Takigawa no Município de Presidente Prudente

Nos outros experimentos, o objetivo foi o de verificar a eficácia do método para linhas de base médias, onde os efeitos da ionosfera e troposfera são mais relevantes. Assim, optou-se por uma coleta de dados de modo que o comprimento das linhas de base variasse de 18 a aproximadamente 120 km. Para tanto, a coleta de dados foi realizada em 4 locais distintos, cuja distribuição pode ser vista na figura 20. A localização da estação base (UEPP – Presidente Prudente) também é ilustrada nesta figura.

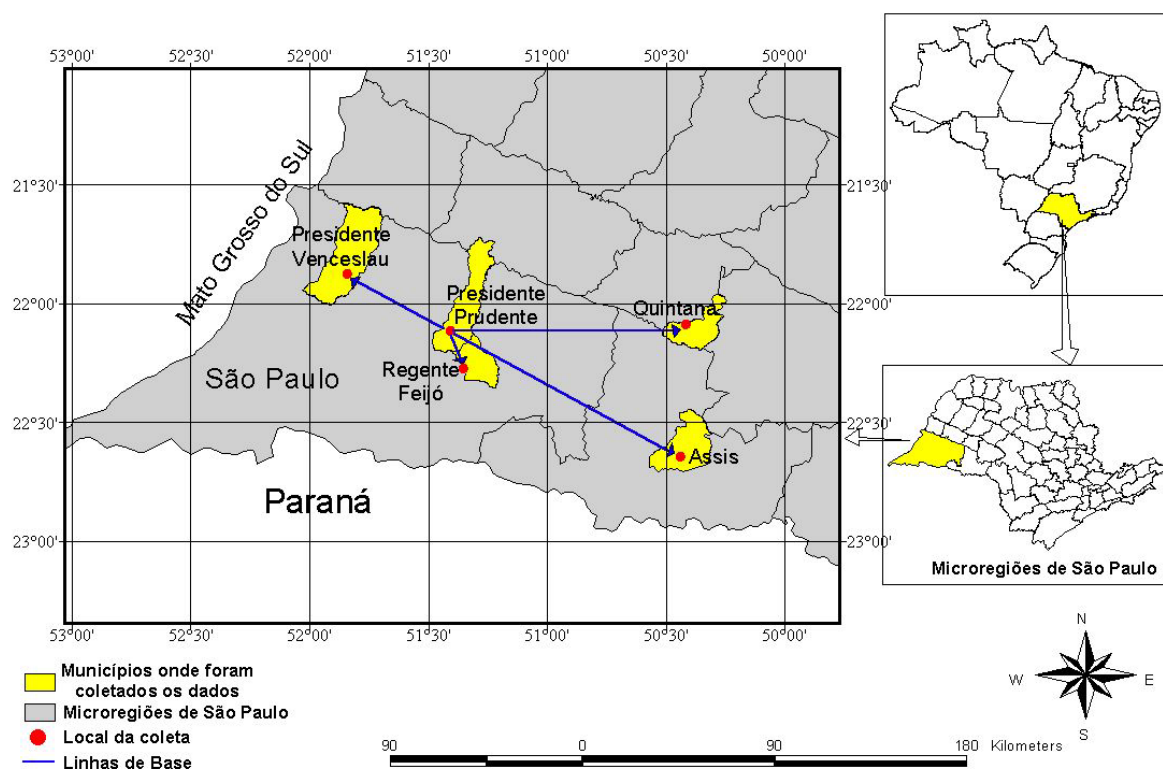


FIGURA 20 - Municípios onde foram realizados os experimentos

No Município de Regente Feijó (Regente), os dados foram coletados num sítio (figura 21), que dispunha de todas as condições de segurança. O comprimento da linha de base é de aproximadamente 18,5 km.

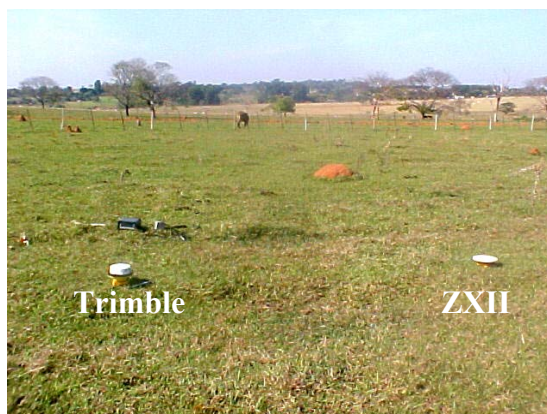


FIGURA 21 - Coleta de dados realizada no sítio (Santo Antônio) em Regente

É necessário ressaltar que, por precaução, mais um receptor Trimble 4600 LS foi instalado nos outros pontos de coleta de dados (Presidente Venceslau, Quintana e Assis), visando garantir a disponibilidade dos dados.

Para coletar os dados no Município de Presidente Venceslau (Venceslau), montou-se um suporte no telhado de uma casa para instalar os receptores (1 receptor ZXII e 2 receptores 4600 LS). A figura 22 ilustra a situação. Nesse caso, o comprimento da linha de base é de aproximadamente 52 km.



FIGURA 22 - Coleta de dados realizada em Venceslau

Posteriormente, os dados foram coletados na fazenda São Pedro, localizada no Município de Quintana (Quintana), com comprimento de linha de base de aproximadamente 102 Km. Nesta fazenda, os mesmos receptores também foram colocados

sobre um suporte no telhado de uma casa. A figura 23 ilustra a disposição dos receptores nessa coleta de dados.

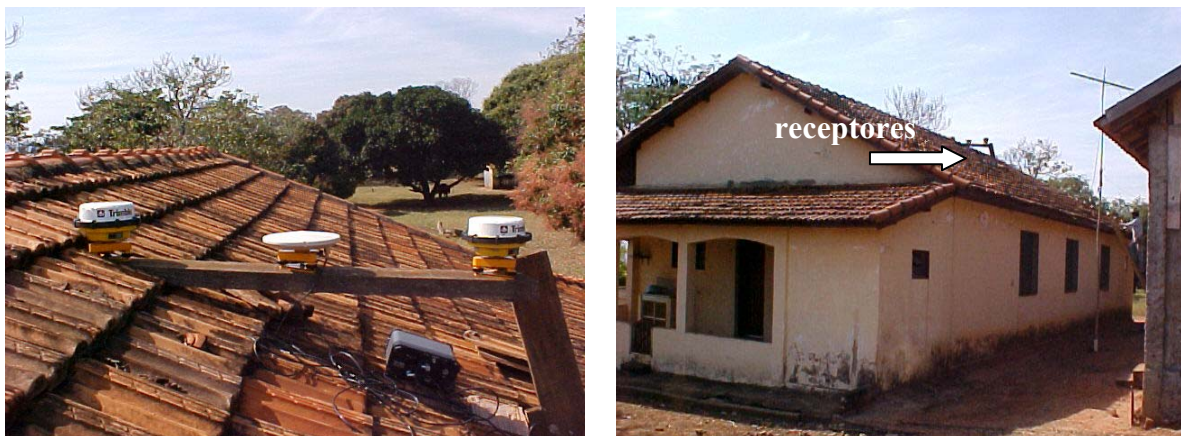


FIGURA 23 - Coleta de dados realizada na fazenda São Pedro - Quintana

No município de Assis os dados foram coletados no campus da UNESP. Os receptores foram instalados sobre os tripés, em cima da caixa de água, como ilustrado na figura 24.



FIGURA 24 - Coleta de dados realizada no Campus da UNESP na cidade de Assis

Para a estação em Assis, o comprimento da linha de base é de aproximadamente 116 km.

Em todas as estações os dados foram coletados apenas para ângulos de elevação superiores a 15° (máscara de elevação de 15°) e taxa de coleta de 15 s. Para as

estações ilustradas na figura 20, a duração da coleta em cada local foi de aproximadamente 24 h. Além disso, os dados foram coletados em julho de 2003. Nos dias 1 e 2 em Regente, 15 e 16 em Assis, 16 e 17 em Quintana, e 17 e 18 em Venceslau.

7.2 Resultados e análises dos experimentos

Essa seção descreve os resultados e as análises dos experimentos realizados para validar o método proposto nessa pesquisa. Primeiramente serão analisados os resultados do experimento que mostra a eficiência do método na redução, principalmente, do multicaminho. Posteriormente, serão apresentados os resultados e análises dos experimentos em que o método é aplicado para linhas de base médias, onde os efeitos atmosféricos são mais significantes.

7.2.1 Resultados e análises de experimento com presença de multicaminho

Para realizar esse experimento foram utilizados os dados coletados na estação base (UEPP), ilustrada na figura 18, e na empresa Takigawa, ilustrada na figura 19, com o receptor Trimble 4600 LS (figura 16). Como não se tratava de uma estação com coordenadas conhecidas, para determinar as coordenadas consideradas “verdadeiras” foram processados no software GAS (seção 6.1) os dados coletados com o receptor Trimble 4600 LS sem a presença do objeto refletor. Para tanto, foram processadas 3 horas de dados, das 15 às 18 h (horário local), com máscara de elevação de 15°. Com isso, foi obtido o valor de

3,98 mm para o sigma a posteriori das DD (raiz quadrada da variância a posteriori, seção 2.4.2, p.25).

Para avaliar a eficiência do método proposto, os dados coletados pelo receptor Trimble 4600 LS com a presença do objeto refletor foram processados no software GAS, utilizando o MMQ convencional (MMQ1), e no software GPSeq (seção 6.2), utilizando o MMQ convencional recursivo (MMQ2) e o MMQ com Penalidades (MMQP) (seção 6.3). Para realizar o processamento foram utilizados intervalos de 5 e 10 min de dados, com a taxa de coleta original, ou seja, 15 s. O processamento de 5 min foi realizado com dados das 15:20 às 15:25 (sessão A1) e das 16:40 h às 16:45 h (sessão B1) (horário local). Já para o processamento de 10 min foram utilizados dados das 15:20 às 15:30 (sessão A2) e das 16:40 h às 16:50 h (sessão B2). Em relação ao satélite base, foi escolhido o satélite 23 para a sessão A e o 14 para a sessão B, pois eles apresentavam o maior ângulo de elevação nos respectivos intervalos de tempo de processamento (figura 25).

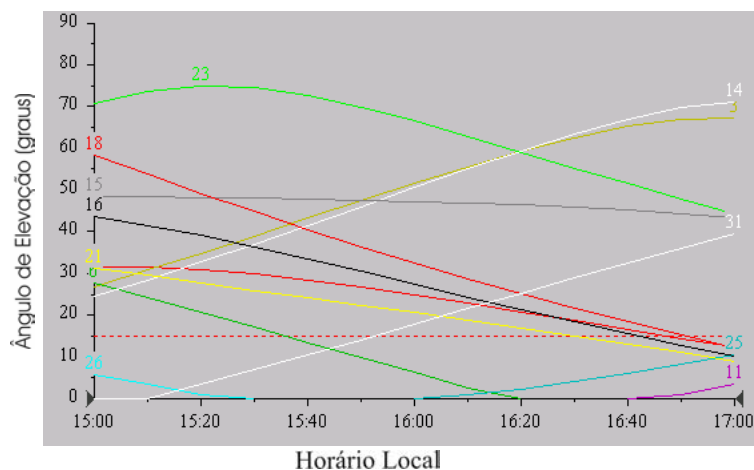


FIGURA 25 - Ângulo de elevação dos satélites da coleta realizada na empresa Takigawa

A figura 26 mostra as discrepâncias entre as coordenadas E, N (UTM) e h (altura geométrica) consideradas verdadeiras e as obtidas pelo MMQ1, MMQ2 e o MMQP, para 5 min de dados. Já na figura 27 as discrepâncias são descritas para 10 min de dados. Em

todos os casos, 8 satélites foram utilizados no processamento. É necessário ressaltar que todas as coordenadas foram calculadas utilizando a ambigüidade fixa.

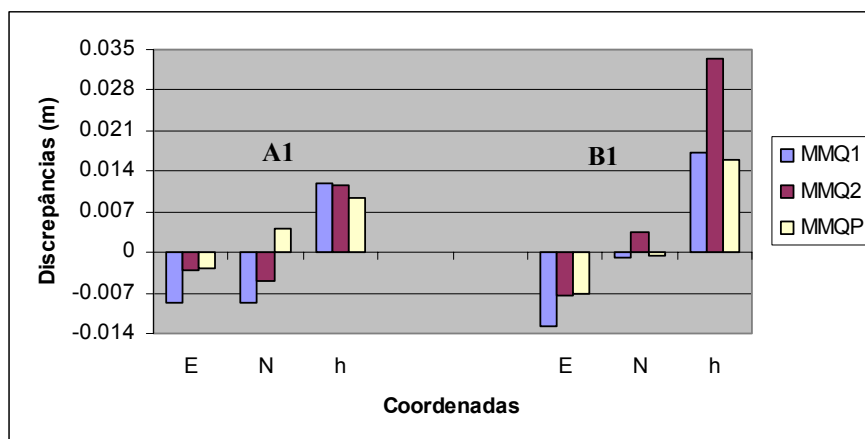


FIGURA 26 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Takigawa com intervalo de 5 min de dados

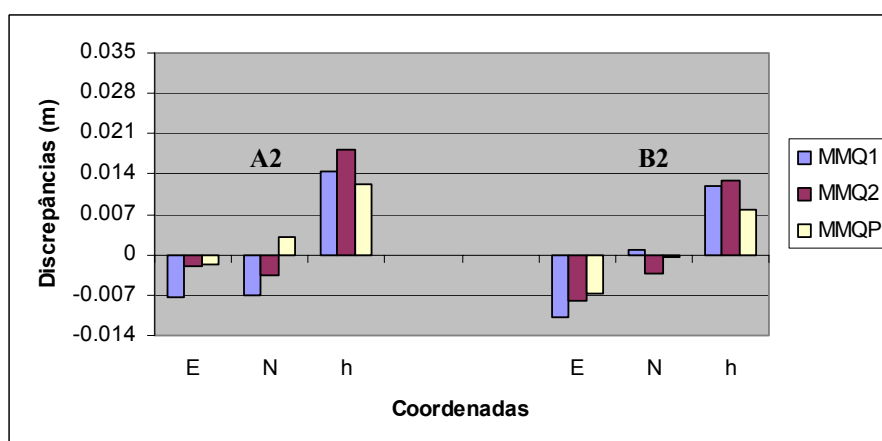


FIGURA 27 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Takigawa com intervalo de 10 min de dados

Note que, para os dois conjuntos de dados, as discrepâncias obtidas pelo modelo semiparamétrico utilizando o MMQP em relação às coordenadas verdadeiras são muito menores do que no MMQ. Isto se deve, provavelmente, a eficiência do método proposto na mitigação do multicaminho sistemático. Para o MMQP, no primeiro caso (figura 26), a discrepância máxima foi de 1,6 cm em h para a sessão B1 e no segundo (figura 27) 1,2 cm em h para a sessão A2. Já para o MMQ convencional as maiores discrepâncias foram de

3,3 cm em h para o MMQ2 na sessão B1, e 1,8 cm em h para o MMQ2 na sessão A2, no primeiro e segundo casos, respectivamente.

Para analisar a qualidade da solução das ambigüidades GPS, foi realizado o teste *ratio* (seção 6.2.2.4, p.85). Apenas valores para o MMQ2 e MMQP serão apresentados, pois o software GAS (MMQ1) não disponibiliza esse tipo de informação. A figura 28 mostra os resultados.

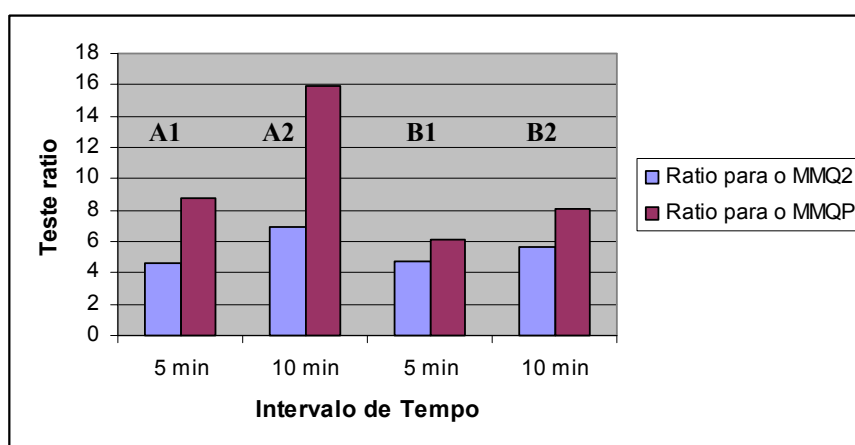


FIGURA 28 - Teste *ratio* para o experimento UEPP-Takigawa com intervalo de 5 min e 10 min de dados

Observando a figura 28 pode-se notar que o MMQP apresenta os maiores valores para o teste *ratio*. Isso indica que a solução possui melhor confiabilidade que no MMQ2. A tabela 05 mostra os valores reais e inteiros do vetor de ambigüidades ajustado pelos dois métodos com 5 min de dados para a sessão A1. Para ambos os casos o método LAMBDA foi utilizado para solucionar a ambigüidade.

TABELA 05 - Solução das ambigüidades GPS para o experimento UEPP-Takigawa com intervalo de 5 min de dados

MMQ2		MMQP		Ambigüidade inteira
Ambigüidade real	Desvio Padrão	Ambigüidade real	Desvio Padrão	
5233121,748	0,231	5233120,835	0,174	5233121,0
-4128753,028	0,267	-4128754,161	0,121	-4128754,0
-6423002,350	0,288	-6423000,820	0,159	-6423001,0
24471958,216	0,232	24471957,949	0,085	24471958,0
10386373,119	0,181	10386372,886	0,117	10386373,0
17797442,788	0,256	17797443,020	0,139	17797443,0
50625330,105	0,387	50625331,099	0,253	50625331,0

Note na tabela 05 que, utilizando o MMQP, a solução da ambigüidade poderia ser obtida a partir do arredondamento ao inteiro mais próximo. Já, para o MMQ2, isso não ocorre. Em relação aos outros conjuntos de dados, com 5 e 10 min, a ambigüidade também poderia ser obtida a partir do inteiro mais próximo no MMQP. Além disso, para 10 min de dados, os desvios padrão são menores para os dois métodos.

Para analisar os resíduos, foram realizadas comparações entre os valores obtidos pelo MMQP implementado nesse trabalho e o MMQ1 processado com o software GAS. Não se comparou com o MMQ2 pelo fato de se tratar de um método recursivo, que calcula resíduos preditos. As tabelas 06 e 07 sumarizam a média absoluta, a média aritmética e o desvio padrão dos resíduos das DD da fase de batimento da onda portadora para os conjuntos de dados de 5 e 10 min da sessão B.

TABELA 06 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Takigawa com intervalo de 5 min de dados

Métodos	Resíduos (ciclos)	DD¹⁴⁻⁰²	DD¹⁴⁻⁰³	DD¹⁴⁻¹⁵	DD¹⁴⁻¹⁶	DD¹⁴⁻¹⁸	DD¹⁴⁻²³	DD¹⁴⁻³¹
MMQ1	Média absoluta	0,032	0,017	0,027	0,022	0,028	0,015	0,034
	Média aritmética	-0,027	-0,002	-0,027	0,012	0,002	0,001	-0,003
	Desvio padrão	0,034	0,022	0,029	0,027	0,033	0,018	0,037
MMQP	Média absoluta	0,018	0,011	0,010	0,016	0,018	0,006	0,019
	Média aritmética	2,2E-12	-7,3E-12	-3,5E-13	8,0E-12	-2,8E-12	-2,2E-12	-9,5E-12
	Desvio padrão	0,022	0,012	0,012	0,020	0,022	0,008	0,021

TABELA 07 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Takigawa com intervalo de 10 min de dados

Métodos	Resíduos (ciclos)	DD¹⁴⁻⁰²	DD¹⁴⁻⁰³	DD¹⁴⁻¹⁵	DD¹⁴⁻¹⁶	DD¹⁴⁻¹⁸	DD¹⁴⁻²³	DD¹⁴⁻³¹
MMQ1	Média absoluta	0,031	0,019	0,021	0,021	0,036	0,011	0,037
	Média aritmética	-0,017	0,004	-0,021	0,005	0,009	-0,008	-0,003
	Desvio padrão	0,035	0,023	0,025	0,026	0,042	0,013	0,044
MMQP	Média absoluta	0,015	0,014	0,012	0,019	0,021	0,007	0,020
	Média aritmética	1,4E-11	2,5E-12	-1,2E-12	-2,2E-12	1,1E-12	2,7E-12	-2,9E-13
	Desvio padrão	0,017	0,018	0,015	0,023	0,028	0,010	0,022

Observando as tabelas 06 e 07 pode-se concluir que a média absoluta, a média aritmética e o desvio padrão dos resíduos obtidos pelo MMQP são sempre melhores que o MMQ1. Para a sessão A, ocorre situação similar.

7.2.2 Resultados e análises de experimentos realizados com linhas de base variando de 18 a 120 km

Para realizar esses experimentos foram utilizados os dados da estação base (UEPP), ilustrada na figura 18, e dos Municípios de Regente Feijó, Presidente Venceslau, Quintana e Assis, figuras 21, 22, 23 e 24, respectivamente. Para obter as coordenadas consideradas “verdadeiras” também foi utilizado o software GAS (seção 6.1).

Primeiramente transportou-se as coordenadas para o centro de fase do receptor de dupla frequência (figura 17). Foram obtidos os valores de 5,28 mm, 10,23 mm, 8,10 mm e 13,8 mm para o sigma a posteriori das DD (seção 2.4.2, p.25) das estações nos municípios de Regente Feijó, Presidente Venceslau, Quintana e Assis, respectivamente. Posteriormente, o receptor de dupla frequência dessas estações foi considerado estação base e realizou-se o posicionamento relativo para determinar as coordenadas das estações ocupadas com os receptores de simples frequência. Nesse processamento, os valores do sigma a posteriori das DD obtidos no GAS foram: 2,26 mm, 3,62 mm, 3,69 mm e 3,74 mm. Nos dois processamentos foram utilizados dados coletados entre 1 e 6 h da madrugada (horário local), com máscara de elevação de 15°.

7.2.2.1 Resultados e análises de experimentos com intervalos de 5 min de dados

Para avaliar a eficiência do método proposto, para linhas de base médias, os dados coletados com o receptor Trimble 4600 LS de simples frequência foram processados no software GAS, utilizando o MMQ convencional (MMQ1), e no software GPSeq, utilizando o MMQ convencional recursivo (MMQ2) e o MMQ com Penalidades (MMQP). Para realizar o

processamento, foram utilizados intervalos de 5 min de dados, com a taxa de coleta original, ou seja, 15 s.

Foram processados, em todas as estações, dados coletados por volta das 14 h local (sessões T1 e T2), pois neste horário os efeitos da ionosfera são maiores, e por volta das 3 h local (sessões N1 e N2), durante a noite. A tabela 08 descreve o intervalo de tempo de processamento para todas as estações.

TABELA 08 - Intervalo de tempo de processamento nas estações dos municípios de Regente, Venceslau, Quintana e Assis (hora local)

ESTAÇÕES	PERÍODOS	INTERVALOS	SESSÕES
Regente	Tarde	13:50 h às 13:55 h	T1
		14:00 h às 14:05 h	T2
	Noite	02:10 h às 02:15 h	N1
		02:15 h às 02:20 h	N2
Venceslau	Tarde	14:00 h às 14:05 h	T1
		14:10 h às 14:15 h	T2
	Noite	03:20 h às 03:25 h	N1
		03:25 h às 03:30 h	N2
Quintana	Tarde	14:35 h às 14:40 h	T1
		15:05 h às 15:10 h	T2
	Noite	03:40 h às 03:45 h	N1
		03:45 h às 03:50 h	N2
Assis	Tarde	14:20 h às 14:25 h	T1
		14:30 h às 14:35 h	T2
	Noite	03:40 h às 03:45 h	N1
		03:45 h às 03:50 h	N2

Em Regente, no período da tarde, apenas 6 satélites foram observados. Para os demais municípios e horários, 7 satélites foram utilizados no processamento.

Em relação ao satélite base, foi escolhido aquele que apresentava o maior ângulo de elevação no intervalo de tempo de processamento. As figuras 29 e 30 ilustram o ângulo de elevação dos satélites para os dados coletados em Regente e Venceslau, durante a tarde.

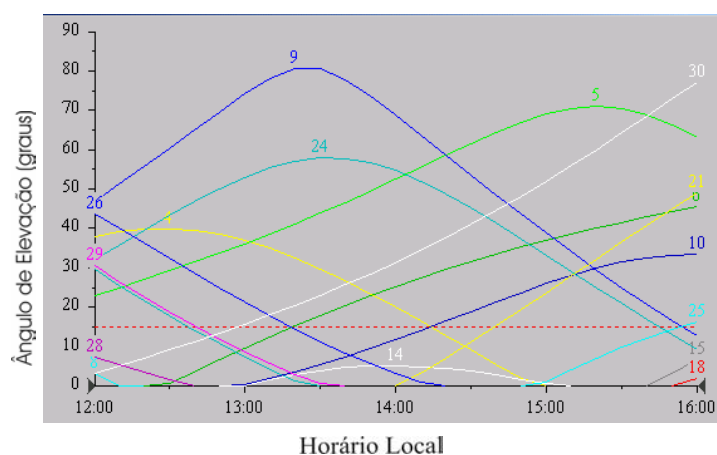


FIGURA 29 - Ângulo de elevação dos satélites da coleta realizada em Regente

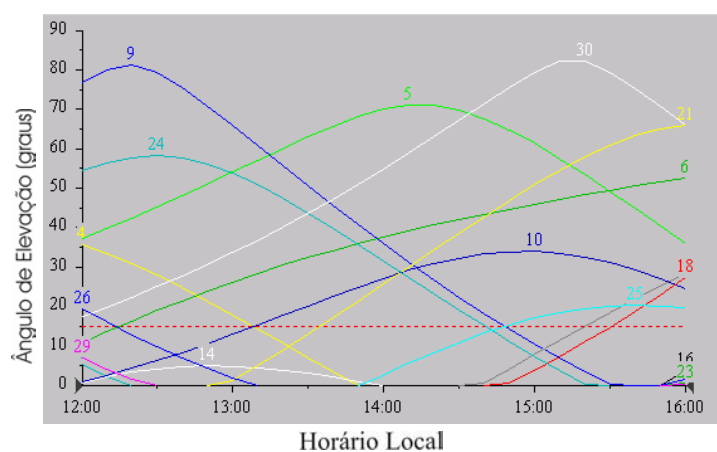


FIGURA 30 - Ângulo de elevação dos satélites da coleta realizada em Venceslau

Note que para Regente o satélite base deve ser o 9. Já para Venceslau, o satélite 5 deve ser escolhido. Para os dados coletados em Quintana e Assis, novamente durante a tarde, as figuras 31 e 32 ilustram o ângulo de elevação dos satélites.

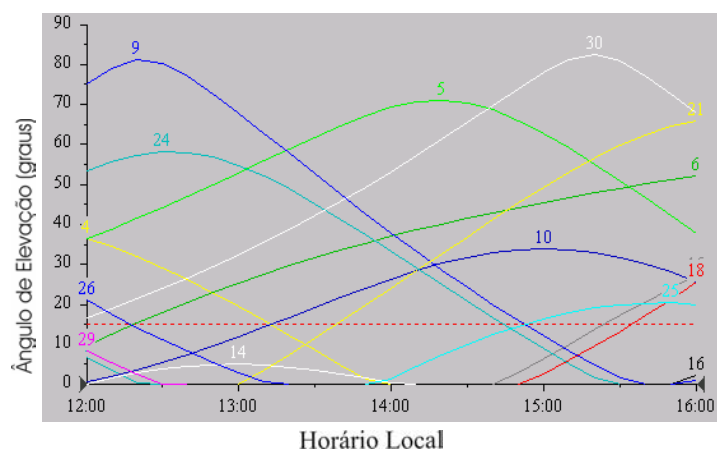


FIGURA 31 - Ângulo de elevação dos satélites da coleta realizada em Quintana

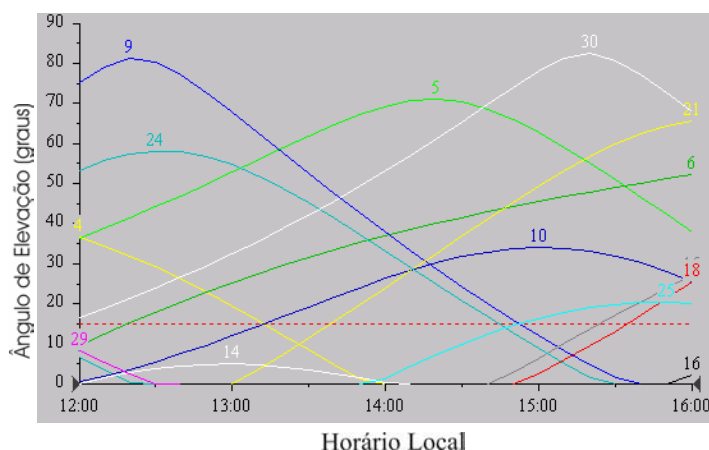


FIGURA 32 - Ângulo de elevação dos satélites da coleta realizada em Assis

Observe que para os dados coletados em Quintana deve ser utilizado o satélite 30 como base. No caso de Assis, o satélite base deve ser o 5. Para os dados coletados durante a noite, o mesmo procedimento foi adotado para a escolha do satélite base.

As figuras 33 e 34 ilustram as discrepâncias entre as coordenadas E, N (UTM) e h (altura geométrica) consideradas verdadeiras e as obtidas pelo MMQ1, MMQ2 e o MMQP nos experimentos UEPP-Regente e UEPP-Venceslau, durante a tarde (sessões T1 e T2) e durante a noite (sessões N1 e N2). É necessário ressaltar que apenas as coordenadas da estação de Regente foram calculadas utilizando a ambigüidade fixa. Para as demais estações foram utilizados os valores *float* da ambigüidade, devido à pequena confiabilidade na solução (figuras 37 e 38). Além disso, no MMQ1, apenas as coordenadas da sessão N2 foram calculadas utilizando valores fixos da ambigüidade.

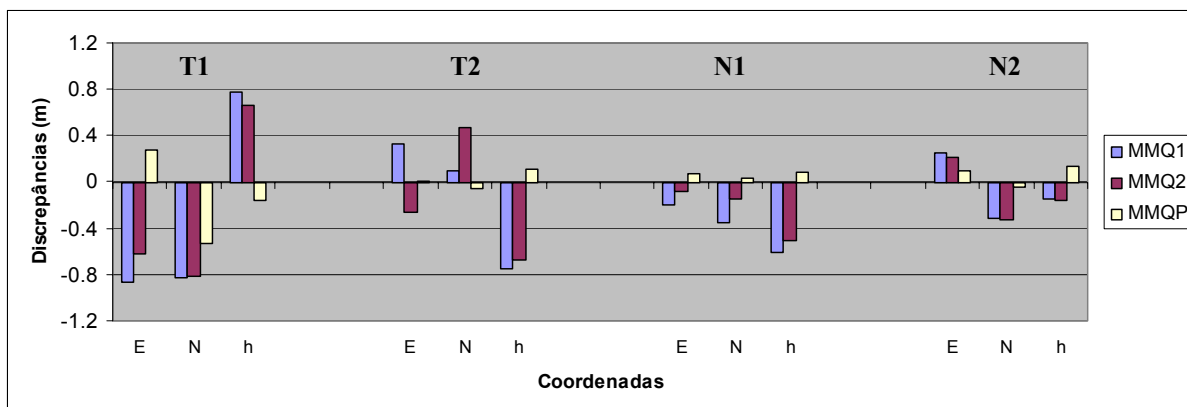


FIGURA 33 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Regente (≈ 18 km) com intervalo de 5 min de dados

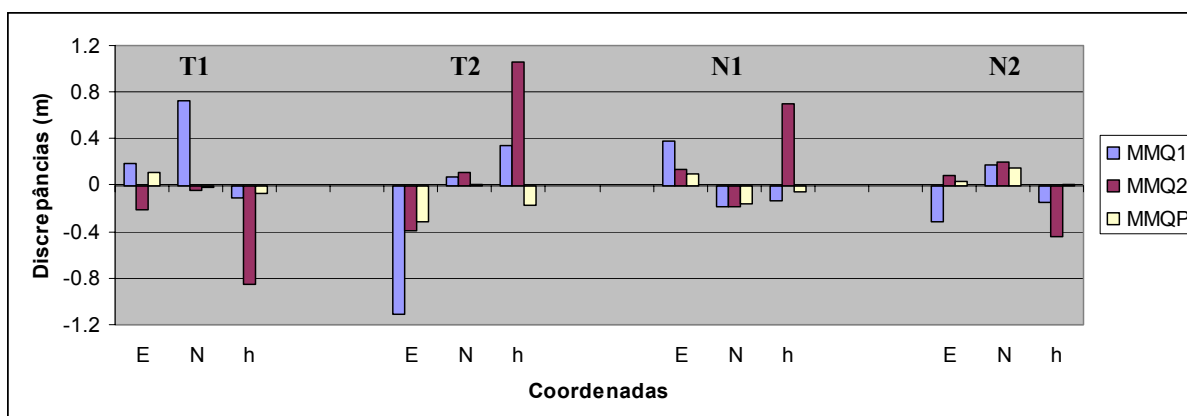


FIGURA 34 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Venceslau (≈ 52 km) com intervalo de 5 min de dados

Durante o período da tarde, no experimento UEPP-Regente, a maior discrepância no MMQ convencional foi de 86,5 cm em E para o MMQ1 na sessão T1. Já para o MMQP foi obtido o valor máximo de 52,8 cm em N para a mesma sessão. No experimento da noite, as discrepâncias máximas foram de 61,0 cm e 13,8 cm em h para o MMQ convencional (MMQ1) e MMQP, nas sessões N1 e N2 respectivamente.

No experimento UEPP-Venceslau, durante a tarde, foram obtidos os valores máximos de 1,10 m e 0,31 cm em E na sessão T2 para o MMQ convencional (MMQ1) e o MMQP, respectivamente. Durante a noite, os maiores valores foram de 70,5 cm em h para o MMQ convencional (MMQ2) e 15,8 cm em N para o MMQP, ambos na sessão N1.

As figuras 35 e 36 ilustram as discrepâncias das coordenadas para os experimentos UEPP-Quintana e UEPP-Assis.

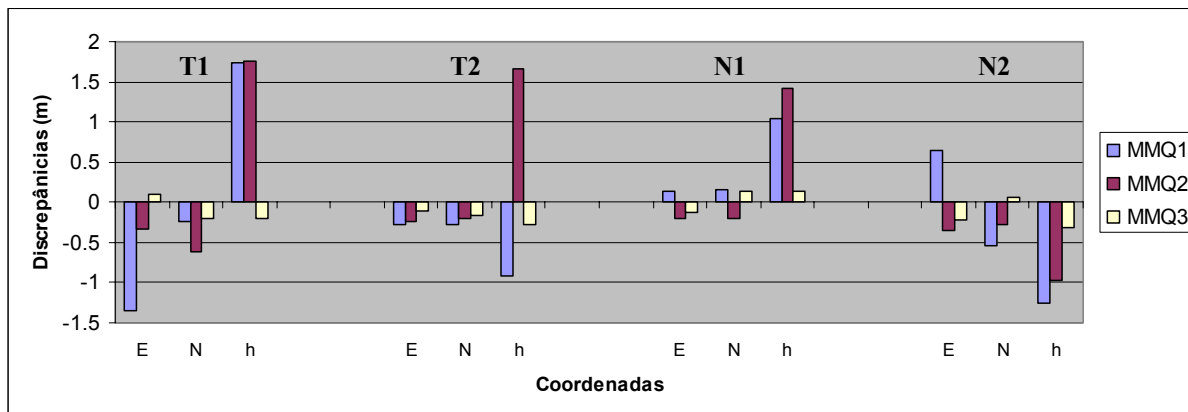


FIGURA 35 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Quintana (≈ 102 km) com intervalo de 5 min de dados

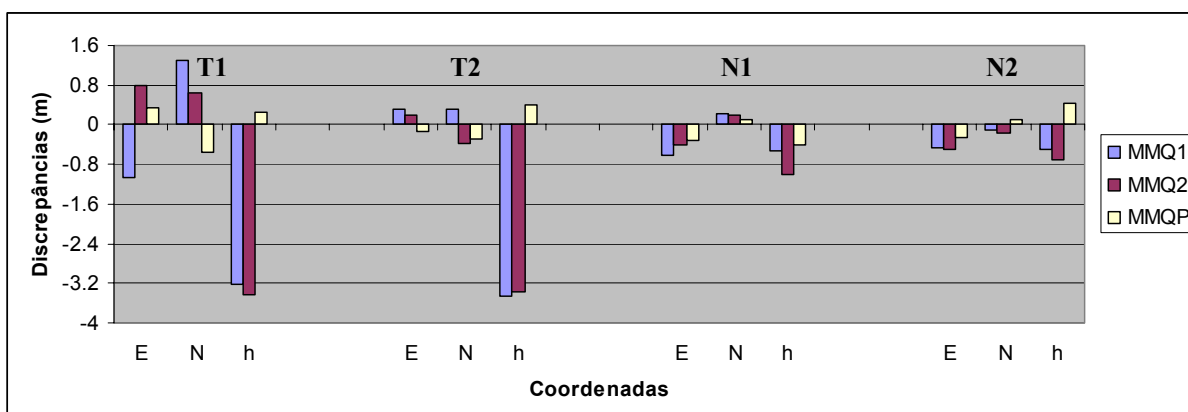


FIGURA 36 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Assis (≈ 116 km) com intervalo de 5 min de dados coletados

No experimento UEPP-Quintana, durante a tarde, foram obtidas as maiores discrepâncias para a coordenada h, sendo de 1,75 m para o MMQ convencional (MMQ2) e 27,2 cm para o MMQP, nas sessões T1 e T2 respectivamente. No período da noite, as discrepâncias máximas foram de 1,41 m em h com o MMQ convencional (MMQ2) para a sessão N1 e 31,4 cm em h com o MMQP para a sessão N2.

Em relação ao experimento UEPP-Assis, durante a tarde, as maiores discrepâncias foram de 3,45 m em h e 56,6 cm em N para o MMQ convencional (MMQ1) e o MMQP nas sessões T2 e T1 respectivamente. Já no experimento realizado a noite, as maiores

discrepâncias foram em h com valores de 1,0 m e 42,8 cm para o MMQ convencional (MMQ2) e o MMQP nas sessões N1 e N2 respectivamente.

Note que para todos os conjuntos de dados, as discrepâncias obtidas pelo modelo semiparamétrico utilizando o MMQP em relação às coordenadas verdadeiras são muito menores do que no MMQ, mostrando a eficiência do método proposto para linhas de base de comprimento médio e curto intervalo de tempo. Além disso, é necessário ressaltar que, como esperado, para as linhas de base mais longas as discrepâncias também aumentam no MMQP. Portanto, nesses casos, dependendo da acurácia requerida, é indicado utilizar dados de receptores de dupla frequência. Jia (2000) indica a utilização de dados de dupla frequência para linhas de base maiores que 30 km.

Além disso, as discrepâncias foram maiores, em geral, para os dados coletados durante a tarde. Isso ocorre provavelmente porque nesse período os efeitos atmosféricos nos sinais GPS são mais relevantes.

Para analisar a qualidade da solução do vetor de ambigüidades GPS, foi realizado o teste *ratio* (seção 6.2.2.4, p.85). As figuras 37 e 38 ilustram os resultados para os dados coletados durante a tarde e a noite, respectivamente.

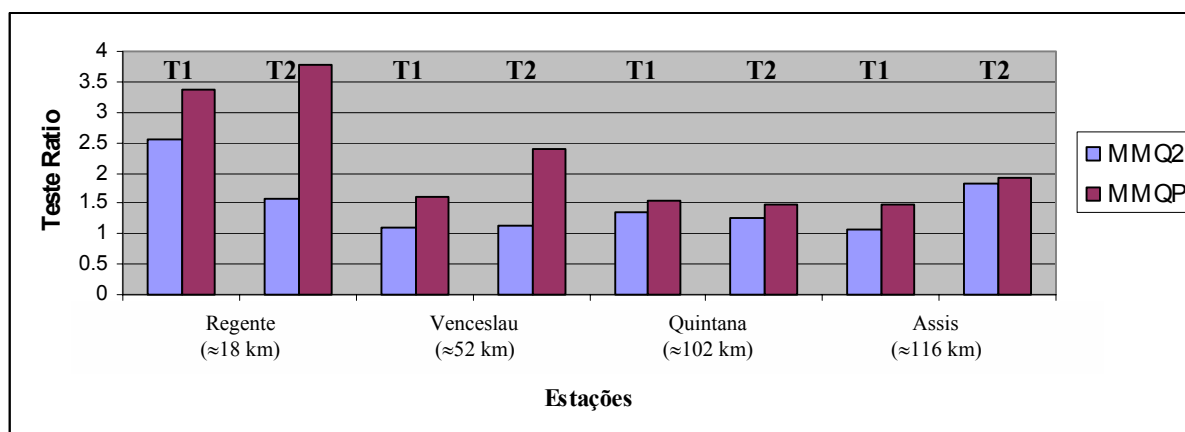


FIGURA 37 - Teste *ratio* para os experimentos com intervalo de 5 min de dados coletados por volta das 14 h local

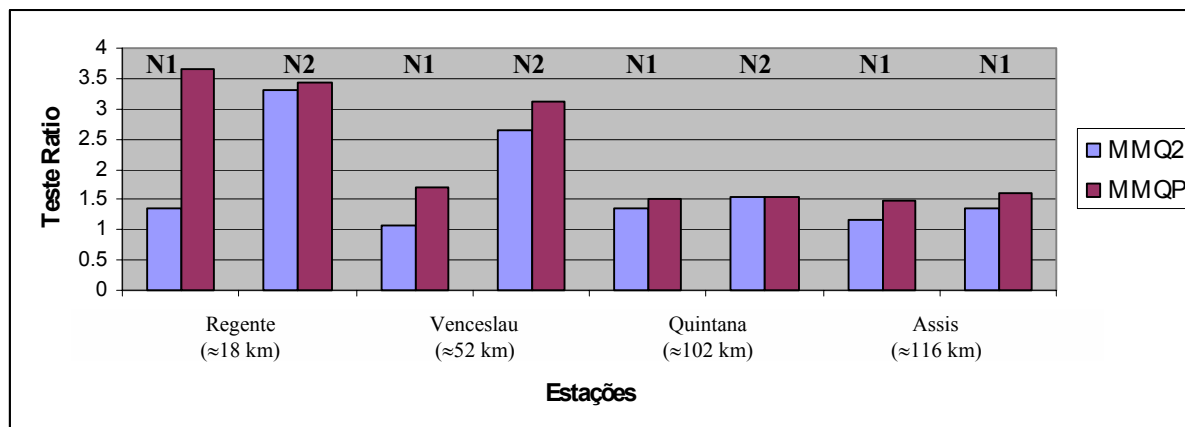


FIGURA 38 - Teste *ratio* para os experimentos com intervalo de 5 min de dados coletados durante a noite

Note que, como esperado, para as linhas de base mais longas o valor do teste *ratio* diminui em ambos os casos. No entanto, para o método proposto, o valor do teste *ratio* foi sempre maior. Além disso, pode-se ressaltar que, para o experimento UEPP-Regente, no MMQP as ambigüidades aproximadas para o inteiro mais próximo já proporcionava a solução correta.

Em relação aos resíduos, foi calculada a média absoluta (MAB), a média aritmética (MA) e o desvio padrão (DP) dos resíduos das DD da fase de batimento da onda portadora para todos os experimentos. Os menores (min) e maiores (max) valores para as sessões T1 e N1 estão sumarizados na tabela 09.

TABELA 09 - Média absoluta, média aritmética e desvio padrão dos resíduos para os experimentos com intervalo de 5 min de dados

Período	Método	Resíduos (ciclo)	Regente		Venceslau		Quintana		Assis	
			min	max	min	max	min	max	min	max
Tarde	MMQ1	MAB	0,013	0,134	0,007	0,082	0,016	0,139	0,034	0,258
		MA	-1,6E-04	1,0E-04	-1,6E-04	1,0E-04	-1,6E-04	1,0E-04	-1,1E-04	1,1E-04
		DP	0,016	0,158	0,009	0,102	0,021	0,165	0,039	0,301
	MMQP	MAB	0,011	0,025	0,007	0,050	0,009	0,049	0,013	0,041
		MA	1,5E-16	5,8E-16	-9,2E-16	-6,1E-16	-8,4E-17	4,7E-18	7,6E-17	2,2E-16
		DP	0,014	0,031	0,009	0,057	0,012	0,059	0,016	0,049
Noite	MMQ1	MAB	0,009	0,032	0,005	0,013	0,010	0,080	0,011	0,048
		MA	-5,2E-05	1,1E-04	-1,6E-04	3,6E-19	-5,2E-05	1,0E-04	-1,1E-04	1,1E-04
		DP	0,012	0,036	0,007	0,016	0,012	0,094	0,015	0,054
	MMQP	MAB	0,009	0,017	0,005	0,011	0,007	0,024	0,009	0,020
		MA	-5,5E-17	-1,5E-17	2,8E-17	6,0E-17	-2,2E-17	1,3E-17	-1,0E-16	-2,0E-17
		DP	0,010	0,021	0,007	0,014	0,008	0,028	0,010	0,024

Observando a tabela 09 pode-se notar que os resultados obtidos pelo MMQP são sempre melhores do que no MMQ1. Além disso, para as linhas de base mais longas, a magnitude dos resíduos e o desvio padrão aumentam em ambos os métodos, tal como esperado. Mas a redução dos valores dos resíduos no MMQP é ainda mais significativa. Em relação aos dados coletados durante a noite, os valores são menores se comparados com os da tarde. Para verificar os valores obtidos para cada DD nas quatro linhas de base, consulte o apêndice C, tabelas C1 a C8.

7.2.2.2 Resultados e análises de experimentos com intervalos de 10 min de dados

Também foram realizados experimentos com intervalo de 10 min de dados por volta das 14 h local (sessões T1 e T2). Da mesma forma que na seção anterior, foram utilizados dados de simples frequência com taxa de coleta de 15 s. A tabela 10 ilustra o intervalo de tempo de processamento para todas as estações.

TABELA 10 - Intervalo de tempo de processamento nas estações dos municípios de Regente, Venceslau, Quintana e Assis (hora local)

ESTAÇÕES	INTERVALOS	SESSÕES
Regente	13:50 h as 14:00 h	T1
	14:00 h as 14:10 h	T2
Venceslau	14:00 h as 14:10 h	T1
	14:10 h as 14:20 h	T2
Quintana	14:35 h as 14:45 h	T1
	15:05 h as 15:15 h	T2
Assis	14:15 h as 14:25 h	T1
	14:30 h as 14:40 h	T2

No experimento UEPP-Regente foram utilizados 6 satélites no processamento. Já nos outros experimentos foram utilizados 7 satélites.

Em relação ao satélite base, foi escolhido aquele que possuía o maior ângulo de elevação no intervalo de tempo de processamento. Os satélites base desse processamento coincidem com aqueles selecionados na seção anterior.

As figuras 39 e 40 ilustram as discrepâncias entre as coordenadas E, N e h consideradas verdadeiras e as obtidas pelo MMQ1, MMQ2 e o MMQP nos experimentos UEPP-Regente e UEPP-Venceslau respectivamente. Além disso, como na seção anterior, apenas as coordenadas da estação de Regente foram calculadas utilizando a ambigüidade fixa. Mas para o MMQ1, as coordenadas são obtidas a partir da solução real das ambigüidades.

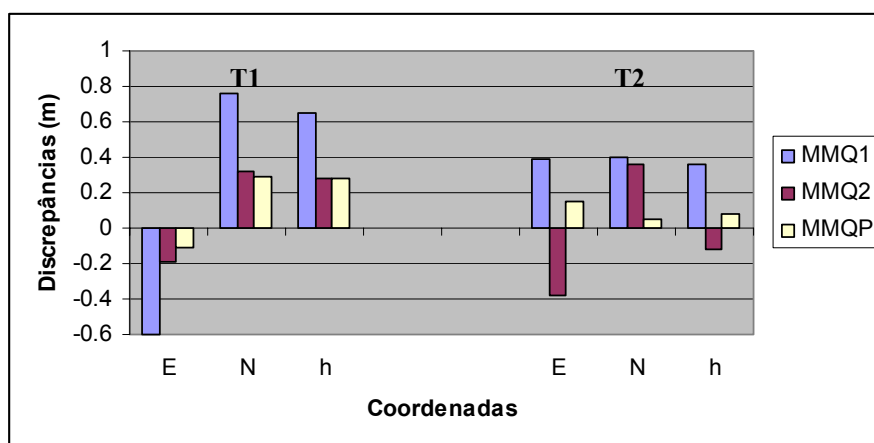


FIGURA 39 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Regente (≈ 18 km) com intervalo de 10 min de dados

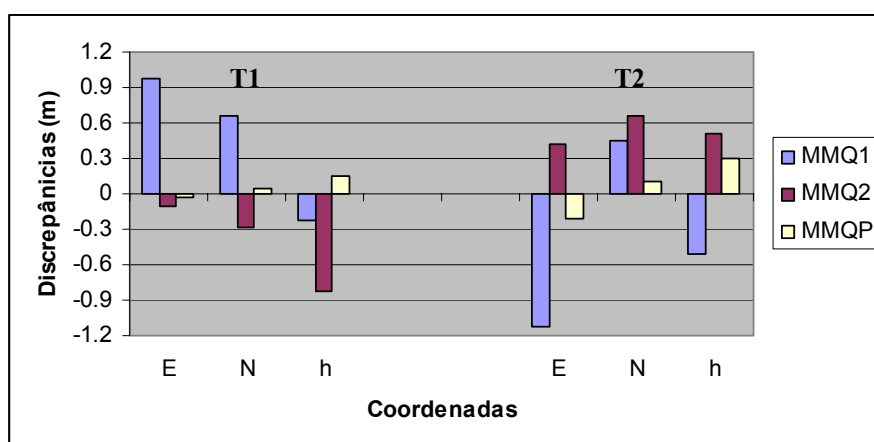


FIGURA 40 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Venceslau (≈ 52 km) com intervalo de 10 min de dados

No experimento UEPP-Regente as maiores discrepâncias foram de 76,4 cm e 28,5 cm em N para o MMQ convencional (MMQ1) e o MMQP respectivamente, ambos na sessão T1. Já no experimento UEPP-Venceslau as discrepâncias máximas foram 1,13 m em E para o MMQ convencional (MMQ1) e 30,3 cm em h para o MMQP, ambos na sessão T2.

As figuras 41 e 42 ilustram as discrepâncias das coordenadas nos experimentos UEPP-Quintana e UEPP-Assis.

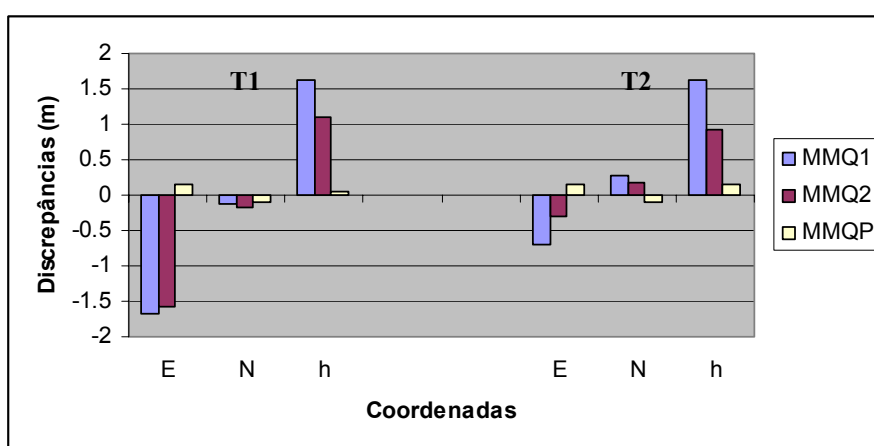


FIGURA 41 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Quintana (≈ 102 km) com intervalo de 10 min de dados

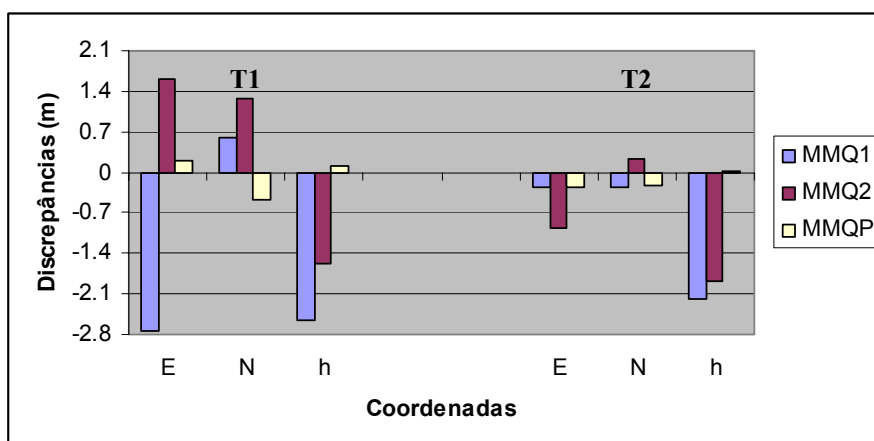


FIGURA 42 - Discrepâncias em relação às coordenadas consideradas verdadeiras para o experimento UEPP-Assis (≈ 116 km) com intervalo de 10 min de dados

No experimento UEPP-Quintana as maiores discrepâncias foram na coordenada E com valores de 1,68 m e 15,0 cm para o MMQ convencional (MMQ1) e o MMQP respectivamente, ambos na sessão T1. Em se tratando do experimento UEPP-Assis,

os maiores valores foram de 2,7 m em E para o MMQ convencional (MMQ1) e 48,1 cm em N para o MMQP, ambos na sessão T1.

Da mesma forma que anteriormente, para todos os conjuntos de dados, as discrepâncias obtidas pelo modelo semiparamétrico utilizando o MMQP em relação às coordenadas verdadeiras são muito menores do que no MMQ, mostrando a eficiência do método proposto. Pode-se notar que, tanto para o MMQ, como para o MMQP, principalmente para as linhas de base mais longas, as discrepâncias diminuíram em relação ao processamento realizado com 5 min de dados. Isso é esperado, pois, como foi dito anteriormente, para essas linhas de bases existe a necessidade de uma maior quantidade de dados, além de ser indicado utilizar dados de receptores de dupla frequência.

Em relação à solução da ambigüidade, a figura 43 ilustra os valores do teste *ratio* obtidos nesses experimentos.

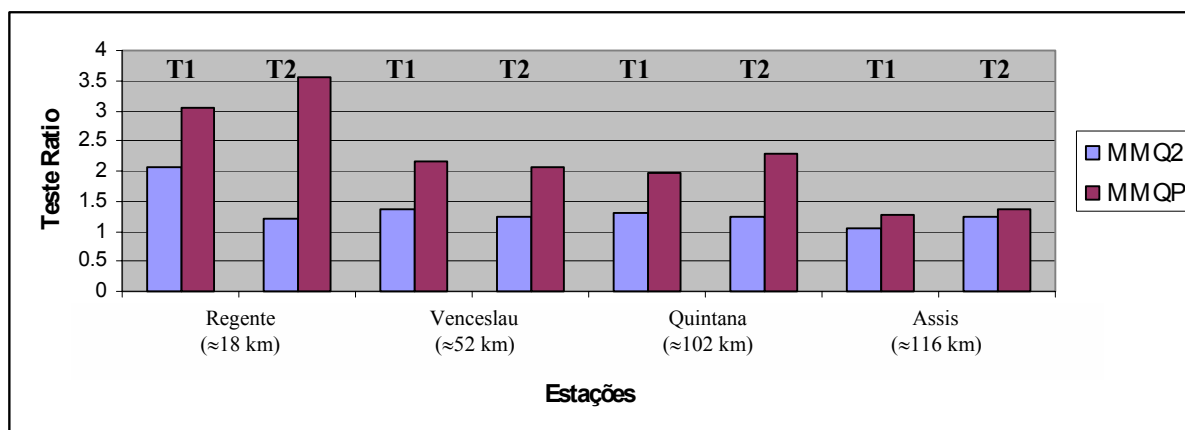


FIGURA 43 - Teste *ratio* para os experimentos com intervalo de 10 min de dados coletados por volta das 14 h local

Note que, novamente, o valor do teste *ratio* diminui para as linhas de base mais longas. Fazendo uma comparação com o experimento realizado com 5 min de dados, pode-se perceber que com 10 min os valores do teste *ratio* são, em geral, maiores. Além disso, como anteriormente, pode-se ressaltar que, para o experimento UEPP-Regente, no

MMQP as ambigüidades aproximadas para o inteiro mais próximo já proporcionava a solução correta.

Para analisar os resíduos, também foi calculada a média absoluta (MAB), a média aritmética (MA) e o desvio padrão (DP) dos resíduos das DD da fase de batimento da onda portadora para todos os experimentos. Os menores (min) e maiores (max) valores para as sessões T1 estão sumarizados na tabela 11.

TABELA 11 - Média absoluta, média aritmética e desvio padrão dos resíduos para os experimentos com intervalo de 10 min de dados

Período	Método	Resíduos (ciclo)	Regente		Venceslau		Quintana		Assis	
			min	max	min	max	min	max	min	max
Tarde	MMQ1	MAB	0,037	0,175	0,048	0,109	0,026	0,157	0,037	0,294
		MA	-7,6E-05	6,0E-18	-5,1E-05	2,5E-05	-3,1E-05	8,6E-18	-0,023	2,6E-05
		DP	0,049	0,200	0,055	0,131	0,032	0,172	0,043	0,347
	MMQP	MAB	0,022	0,041	0,046	0,088	0,021	0,063	0,028	0,138
		MA	-3,6E-17	2,9E-17	-9,2E-17	1,5E-16	-1,5E-16	-4,4E-17	-1,6E-16	5,5E-16
		DP	0,026	0,044	0,055	0,098	0,026	0,075	0,036	0,160

Note que os resultados apresentados pelo MMQP são melhores em relação ao MMQ1. Para maiores detalhes em relação aos resíduos para cada DD consulte o apêndice C, tabelas C9 a C12.

7.3 Eficiência computacional

Considerando os resultados apresentados, pode-se concluir que em termos de qualidade, o MMQP é superior ao MMQ. No entanto, uma desvantagem do MMQP em relação ao MMQ diz respeito à eficiência computacional.

Como foi descrito na seção 6.3 (p.86), para implementar o MMQP foi desenvolvida uma rotina principal que determina os valores ajustados de x e g , bem como os valores da função GCV. Essa rotina é utilizada várias vezes no processamento, devido ao

método de minimização. Portanto, para quantificar a demanda computacional requerida pelo MMQP, o intervalo de tempo necessário para executar essa rotina uma única vez foi avaliado. Para tanto, foram utilizados dois computadores Pentium III com processadores diferentes. Os resultados para 5 e 10 min de dados com taxa de coleta de 15 s seguem, respectivamente, nas tabelas 12 e 13.

TABELA 12 - Demanda computacional para 5 min de dados com o MMQ com Penalidades

COMPUTADORES	Demanda computacional		
	6 satélites	7 satélites	8 satélites
Processador 750 MHz Memória 128 MB	44,35 s	1 min e 15,8 s	1 min e 58,83 s
Processador 1 GHz Memória 256 MB	34,10 s	57,78 s	1 min e 30,59 s

TABELA 13 - Demanda computacional para 10 min de dados com o MMQ com Penalidades

COMPUTADORES	Demanda computacional		
	6 satélites	7 satélites	8 satélites
Processador 750 MHz Memória 128 MB	5 min 9,24 s	8 min e 44,75 s	13 min e 45 s
Processador 1 GHz Memória 256 MB	3 min 49,79 s	6 min 38,71 s	11 min e 3,88 s

Observando as tabelas 12 e 13 pode-se concluir que para o MMQP o aumento da quantidade de dados aumenta em muito o intervalo de tempo necessário para o processamento dos dados. Além disso, é possível notar uma diminuição no intervalo de tempo de processamento para um computador de melhor desempenho.

É necessário ressaltar que, conforme descrito anteriormente, as tabelas 12 e 13 mostram a demanda computacional da rotina principal do programa. Essa rotina é utilizada de 0 a 15 vezes para cada parâmetro suavizador. Portanto o tempo total de processamento varia bastante.

Em relação ao MMQ, a demanda computacional está descrita nas tabelas 14 e 15.

TABELA 14 - Demanda computacional para 5 min de dados com o MMQ convencional

COMPUTADORES	Demanda computacional (s)		
	6 satélites	7 satélites	8 satélites
Processador 750 MHz Memória 128 MB	1,27 s	1,29 s	1,32 s
Processador 1 GHz Memória 256 MB	1,03 s	1,13 s	1,18 s

TABELA 15 - Demanda computacional para 10 min de dados com o MMQ convencional

COMPUTADORES	Demanda computacional (s)		
	6 satélites	7 satélites	8 satélites
Processador 750 MHz Memória 128 MB	1,65 s	2,54 s	3,66 s
Processador 1 GHz Memória 256 MB	1,21 s	1,79 s	1,97 s

Observando as tabelas 14 e 15 pode-se notar que no MMQ o processamento dos dados é bastante rápido.

Face a demanda computacional exigida no MMQP, não se pôde, até o momento, explorar vários períodos do dia, bem como sessões de dados mais longas, para uma avaliação mais detalhada do método. Assim, para melhorar a demanda computacional do MMQP, deve ser realizada uma otimização do algoritmo, visto que grande parte das matrizes utilizadas tem estrutura em banda. Além disso, computadores com maior eficiência também podem auxiliar na diminuição do intervalo de tempo de processamento dos dados.

8 CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

O GPS tem como principal objetivo viabilizar a navegação de baixa, média e alta precisão. Mas suas observáveis, bem como todas as outras observáveis envolvidas nos processos de medidas, estão sujeitas a erros. Essa pesquisa teve como objetivo atenuar o efeito dos erros sistemáticos no posicionamento relativo GPS, em princípio, para usuários que dispõem apenas de receptores de simples frequência.

No posicionamento relativo com o GPS, para linhas de base curtas, a principal fonte de erro sistemático é o multicaminho. Já, para linhas de base de média e longa distância, as principais fontes de erros sistemáticos são a refração ionosférica e troposférica e o erro causado pela órbita dos satélites GPS. Esses erros sistemáticos podem não somente impedir uma confiável resolução de ambigüidades, como também degradar a acurácia dos resultados.

Para atenuar os efeitos desses erros e para melhorar a confiabilidade da resolução das ambigüidades e da estimativa das coordenadas de interesse para usuários de receptores de simples frequência, foi avaliada a utilização do MMQ com penalidades, dentro do contexto do modelo semiparamétrico, usando uma spline cúbica natural. Além disso, como tal funcionalidade não está disponível em software de uso comum, teve-se que realizar sua implementação para a investigação desejada.

Assim, para desenvolver essa pesquisa foi realizada uma extensa revisão bibliográfica sobre as principais características do GPS, bem como suas observáveis, tipos de posicionamento e erros envolvidos nos processos de medida. Dentre esses erros, foi dada especial atenção aos erros sistemáticos, como aqueles causados pela refração ionosférica e troposférica, o multicaminho e os erros das órbitas dos satélites GPS. Além disso, foi

realizado um extenso e rigoroso estudo sobre as splines, o MMQ com Penalidades e o modelo semiparamétrico. Todos esses conceitos foram essenciais no desenvolvimento do trabalho.

Cabe ainda acrescentar que o método proposto foi implementado no software GPSeq que está em desenvolvimento na FCT/UNESP. E, com isso, vários experimentos foram realizados com o software para validar o método proposto.

No primeiro experimento, o objetivo principal era verificar se o método proposto seria eficaz na redução de erros causados pelo multicaminho estático. Para tanto, foi introduzido um objeto refletor próximo ao receptor. Com esse experimento foi possível observar uma significativa melhora nas coordenadas e na solução das ambigüidades GPS com o MMQ com Penalidades em relação ao MMQ convencional. Utilizando 5 min de dados, as discrepâncias máximas das coordenadas em relação aos valores verdadeiros chegaram a 1,6 cm e 3,3 cm em h para o MMQ com Penalidades e o MMQ convencional recursivo, respectivamente. Assim, pode-se concluir que o MMQ com Penalidades é eficiente na atenuação do multicaminho, utilizando dados de simples frequência (base curta), com um pequeno intervalo de tempo de coleta de dados.

Em se tratando do segundo experimento, foram coletados dados em diferentes comprimentos de linhas de base. Em todos os testes realizados com o método proposto, as discrepâncias em relação às coordenadas consideradas verdadeiras foram menores se comparados com o MMQ convencional. Utilizando 5 min de dados, as discrepâncias máximas foram de 56,6 cm em N para o MMQ com Penalidades e 3,45 m em h para o MMQ convencional, para a linha de base mais longa.

Além disso, no segundo experimento foi possível notar que, com o aumento do comprimento das linhas de base, as discrepâncias em relação às coordenadas consideradas verdadeiras aumentaram. Nesse caso, o intervalo de tempo de processamento deveria ser aumentado, o que não foi realizado devido à alta demanda computacional exigida no MMQ

com Penalidades até o momento. Uma outra possibilidade seria a utilização de dados de dupla frequência para as linhas de base mais longas (JIA, STEWART e TSAKIRI, 2001). Porém, como um dos objetivos desse trabalho era verificar até que ponto dados de receptores de simples frequência poderiam ser utilizados no posicionamento relativo GPS, não foram utilizados dados de dupla frequência.

Mas, apesar dos resultados satisfatórios apresentados pelo MMQ com Penalidades, esse método apresentou uma deficiência: uma alta demanda computacional. O intervalo de tempo gasto para o MMQ com Penalidades é muito maior se comparado com o MMQ convencional. Para atenuar esse problema, uma recomendação é realizar a otimização do algoritmo implementado, visto que muitas das matrizes utilizadas no modelo semiparamétrico e no MMQ com Penalidades são em banda. Além disso, computadores de maior eficiência podem ser utilizados para realizar o processamento. Com isso, com certeza o intervalo de tempo utilizado será reduzido significativamente.

Outra recomendação diz respeito ao processamento dos dados. Poderão ser realizados experimentos modificando a taxa de coleta de dados e a duração das sessões processadas. Também poderá ser testada a utilização de dados de dupla frequência para verificar as melhorias para as linhas de base mais longas. Tais alternativas serão viáveis a partir da otimização do programa.

9 REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, D. B. M. **MMQ com Penalidades: Atenuação de erros sistemáticos no posicionamento relativo GPS**. Presidente Prudente: Departamento de Matemática, 2003. Relatório Técnico. 2/2003, 70p.

ALVES, D. B. M.; MONICO, J. F. G.; MENEGUETTE, M. Atenuação dos efeitos da ionosfera no posicionamento relativo GPS de alta precisão utilizando a técnica dos mínimos quadrados penalizados. In: III COLÓQUIO BRASILEIRO DE CIÊNCIAS GEODÉSICAS, 2003, Curitiba. **Anais...**, 2003.

BRAASCH, M. S. A Signal Model for GPS, **Navigation**. v. 37, n. 4, p. 363-377, 1991.

CAMARGO, P. O. **Modelo regional da ionosfera para uso em posicionamento com receptores GPS de uma frequência**. 1999. 191f. Tese (Doutorado em Ciências Geodésicas) – Universidade Federal do Paraná, Curitiba.

CUNHA, M. C. **Métodos Numéricos**. 2.ed. Unicamp, 2000. 263p.

FARRET, J. C. **O efeito do multicaminho estático nas medidas da fase das portadoras GPS**. 2000. 132 f. Tese (Doutorado em Ciências Geodésias) - Universidade Federal do Paraná, Curitiba.

FESSLER, J. A. Nonparametric Fixed-Interval Smoothing With Vector Splines. In: IEEE TRANSACTIONS ON SIGNAL PROCESSING, 1991. **Proceedings...** v.39, p.852-859.

FORTES, L. P. S. **Operacionalização da Rede Brasileira de Monitoramento Contínuo do Sistema GPS (RBMC)**. 1997. 152f. Dissertação (Mestrado em Ciências em Sistemas e Computação) – Instituto Militar de Engenharia (IME), Rio de Janeiro.

_____. **Optimising the Use of GPS Multi-Reference Stations for Kinematic Positioning**. 2002. 355f. Tese (PhD) – University of Calgary, Calgary.

FONSECA, E. S. **O sistema GPS como ferramenta para a avaliação da refração ionosférica no Brasil**. 2002. 200f. Tese (Doutorado em Engenharia de Transportes) – Escola Politécnica, Universidade de São Paulo, São Paulo.

GEMAEL, C. **Introdução ao ajustamento de observações: aplicações Geodésicas**. Curitiba: UFPR, 1994. 319p.

GIZAWY, M. L. **Development of an Ionosphere Monitoring Technique Using GPS Measurements for High Latitude GPS Users**. 2003. 174p. Dissertação (MSc) – University of Calgary, Calgary.

GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations**. United States of America: The Johns Hopkins University Press, 1983. 476p.

GREEN, P. J.; SILVERMAN, B. W. **Nonparametric Regression and Generalized Linear Models: a roughness penalty approach**. 1.ed. London: Chapman & Hall, 1994. 182p.

HANNAH, C. D.; WALKER, R. A. e KUBIK, K. Towards a Complete Virtual Multipath Analysis Tool. In: INTERNATIONAL TECHNICAL MEETING, 11, 1998, Nashville. **Proceedings...** Nashville: The Satellite Division of the Institute of Navigation, 1998. p. 1055-1063.

HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; COLLINS, J. **GPS Theory and Practice**. 4 ed. Wien: Springer-Verlag, 1997. 326p.

HUTCHINSON, M. Algorithm 642: a fast procedure for calculating minimum cross validation cub smoothing splines. **ACM Transactions on Mathematical Software**, v.12, n.2, p.150-153, 1986. Disponível em: <<http://cm.bell-labs.com/netlib/toms/642.gz>>. Acesso em: 04 mar. 2003.

IGS - International GPS Service . **IGS PRODUCTS**. Disponível em: <<http://igsb.jpl.nasa.gov/components/prods.html>>. Acesso em: 13 jan. 2004.

IGS - International GPS Service . **GPS/GLONASS RINEX 2.1.0 data**. Disponível em: <<http://igsb.jpl.nasa.gov/igsb/data/format/rinex210.txt>>. Acesso em: 10 mar. 2003.

JIA, M. Mitigation of Systematic Errors of GPS Positioning Using Vector Semiparametric Models. In: ION GPS 2000, Salt Lake City, 2000. **Proceedings...**, 2000. p.1938-1947.

_____; TSAKIRI, M. STEWART, M. Mitigation Multipath Errors Using Semi-parametric Models for High Precision Static Positioning. In: IAG SYMPOSIA, 2000. **Geodesy Beyond 2000 – The Challenges of the First Decade**, v.121, p.393-398.

JIA, M.; STEWART, M.; TSAKIRI, M. Mitigation of Ionospheric Errors by Penalised Least Squares Technique for High Precision Medium Distance GPS Positioning. In: KIS 2001, Banff, Canada. **Proceedings...**, 2001. 1 CDROM.

KLOBUCHAR, J. A. Ionospheric Effects on GPS. In: PARKINSON, B. W.; SPILKER Jr., J. **J. Global Positioning System: Theory and Applications**. Cambridge: American Institute of Aeronautics and Astronautics, 1996. v.1, p.485-515.

LEICK, A. **GPS – Satellite Surveying**. 2.ed. John Wiley & Sons, 1995. 560p.

LIU, G. C. **Ionosphere weighted global positioning system Carrier phase ambiguity resolution**. 2001. 173p. Dissertação (MSc) – University of Calgary, Calgary.

MACHADO, W. C. **Solução Rápida das Ambigüidades GPS para Aplicação no Posicionamento Relativo DE Bases Curtas**. 2001. 121f. Dissertação (Mestrado em Ciências Cartográficas) – Faculdade de Ciências de Tecnologia, Universidade Estadual Paulista, Presidente Prudente.

_____; MONICO, J. F. G. Solução rápida das ambigüidades GPS utilizando o software GPSeq. In: XIX Congresso Brasileiro de Cartografia, Recife, 1999. **Proceedings...**, 1999.

_____. Utilização do software GPSeq na solução rápida das ambigüidades GPS no posicionamento relativo cinemático de bases curtas. **Pesquisa em Geociências**, Porto Alegre, p.89-99, 2002.

MONICO, J. F. G. **Posicionamento pelo NAVSTAR-GPS: Descrição, Fundamentos e Aplicações**. 1.ed. São Paulo: Unesp, 2000. 287p.

MORETTIN, P. A. **Ondaletas e seus usos na Estatística**. In: 7ª Escola de Séries Temporais e Econometria, 1997, Canela, Rio Grande do Sul. 85p.

OLYNIC, M. C. **Temporal characteristics of GPS error sources and their impact on relative positioning**. 2002. 141f. Dissertação (MSc) – University of Calgary, Calgary.

PRESS, W. H. et al. **Numerical Recipes: The Art of Scientific Computing**. Cambridge University Press, 1986.

RAY, J. K. **Mitigation of GPS Code and Carrier Phase Multipath Effects Using a Multi-Antenna System**. 2000. 260p. Dissertação (MSc) – University of Calgary, Calgary.

REINSCH, C. H. Smoothing by Spline Functions. **Numerische Mathematik**, v.10, p.177-183, fev. 1967.

_____. Smoothing by Spline Functions II. **Numerische Mathematik**, v.16, p.451-454, mar. 1971.

ROGERS, D. F.; ADAMS, J. A. **Mathematical Elements for Computer Graphics**. 2. ed. McGraw-Hill, 1990. 611p.

RUGGIERO, M. A. G.; LOPES, V. L .R. **Cálculo Numérico: Aspectos Teóricos e Computacionais**. 2.ed. São Paulo: MAKRON Books, 1996. 406p.

SAPUCCI, L. F. **Estimativa do vapor d'água atmosférico e a avaliação do atraso zenital troposférico utilizando GPS**. 2001. 167 f. Dissertação (Mestrado em Ciências Cartográficas) - Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente.

SEEBER, G. **Satellite Geodesy: Foundations, Methods, and Applications**. Berlin, New York: Walter de Gruyter, 2003. 589p.

SIDC – Solar Influence Data Analysis Center. **Sunspot Index Graphics**. Disponível em: <<http://sidc.oma.be/html/wolfmms.html>>. Acesso em: 23 jun. 2003.

SPIKER Jr., J. J.; PARKINSON, B. W. Overview of GPS Operation and Design. In: PARKINSON, B. W.; SPIKER Jr., J. J. **Global Positioning System: Theory and Applications**. Cambridge: American Institute of Aeronautics and Astronautics, 1996. v.1, p.29-55.

SPIKER Jr., J. J. Tropospheric Effects on GPS. In: PARKINSON, B. W.; SPIKER Jr., J. J. **Global Positioning System: Theory and Applications**. Cambridge: American Institute of Aeronautics and Astronautics, 1996. v.1, p.517-546.

STEWART, M.P.; FOULKES-JONES, G.H.; CHEN, W.; OCHIENG, W.Y.; SHARDLOW, P.J.; PENNA, N.T. **GPS Analysis Software (GAS)**, Version 2.4 User Manual, IESSG Publication, UK, 1997.

TEUNISSEN, P. J. G. GPS Carrier Phase Ambiguity Fixing Concepts. In: TEUNISSEN, P. J. G.; KLEUSBERG, A. **GPS for Geodesy**. 2ed. Berlin: Springer Verlage, 1998a, p.271-318.

TEUNISSEN, P. J. G. Quality Control and GPS. In: TEUNISSEN, P. J. G.; KLEUSBERG, A. **GPS for Geodesy**. 2ed. Berlin: Springer Verlage, 1998b, p.271-318.

UNIVERSITY OF LEICESTER. **Ionospheric Physics**: The electron concentration profile. Disponível em: <<http://ion.le.ac.uk/ionosphere/profile.html>>. Acesso em: 23 jun. 2003.

TODD, J. **Survey of Numerical Analysis**. McGraw-Hill Book Company, 1962. 589p.

10 BIBLIOGRAFIA

BAZARAA, M. S.; JARVIS, J. J.; SHERALI, H. D. **Linear Programming and Network Flows**. 2.ed. 1990. 684p.

BOOR, C. A **Practical Guide to Splines**. New York: Springer-Verlag, 1978. 392p.

BRODLIE, K. W. **Mathematical Methods in Computer Graphics and Design**. Academic Press, 1980. 147p.

EUBANK, R. L. **Spline smoothing and nonparametric regression**. New York: Marcel Dekker, 1988. 435p.

HOOG, F. R.; HUTCHINSON, M. F. An efficient method for calculating smoothing splines using orthogonal transformations. **Numerische Mathematik**. v.50, p.311-319, 1987.

HUTCHINSON, M. F.; HOOG, F. R. Smoothing Noisy Data with Spline Functions. **Numerische Mathematik**. v.47, p.99-106, 1985.

MORITZ, H. **Advanced Physical Geodesy**. 2.ed. New York: Wiley & Son, 1989. 560p.

PRENTER, P. M. **Splines and variational methods**. John Wiley & Sons, 1975. 323p.

SATIRAPOD, C.; WANG, J.; RIZOS, C. Modelling residual systematic errors in GPS positioning: Methodologies and comparative studies. In: IAG Scientific Meeting, Budapest, Hungary, 2001a. **Proceedings...** Disponível em: <www.gmat.unsw.edu.au/snap/publications/satirapod_etal2001f.pdf>. Acesso em 10 nov. 2002.

_____ et al. GPS analysis with the aid of wavelets. In: 5th Int. Symp. on Satellite Navigation Technology & Applications, Canberra, Australia, 2001b. **Proceedings...**

SHULTZ, M. H. **Spline Analysis**. Prentice Hall, 1973. 165p.

STEWART, M.P.; FOULKES-JONES, G.H.; CHEN, W.; OCHIENG, W.Y.; SHARDLOW, P.J.; PENNA, N.T. **GPS Analysis Software (GAS)**, Version 2.4 User Manual, IESSG Publication, UK, 1997.

TIBERIUS, C. C. J. M. **Recursive data processing for Cinematic GPS surveying**. Delft University of Technology, 1998. 248 p.

WANG, J.; SATIRAPOD, C.; RIZOS, C. Stochastic assessment of GPS carrier phase measurements for precise static relative positioning. **Journal of Geodesy**, v.76, n.2, fev.2002.

APÊNDICE A – Teorema que determina o mínimo de uma spline cúbica natural

Teorema - Suponha $n \geq 2$, e que g é uma spline cúbica natural interpolante para os valores $z_1, \dots, z_n$ nos pontos $t_1, \dots, t_n$ satisfazendo $a < t_1 < \dots < t_n < b$. Seja $\bar{g}$ uma função em $S_2[a, b]$ no qual $\bar{g}(t_i) = z_i$ para $i = 1, \dots, n$. Então $\int_a^b \bar{g}''^2 \geq \int_a^b g''^2$, com igualdade somente se $\bar{g}$ e g são idênticas.

Prova: Seja h uma função em $S_2[a, b]$ dada por $h = \bar{g} - g$. Ambas funções g e $\bar{g}$ interpolam os valores z_i , assim h é nula em todos os pontos t_i para $i = 1, \dots, n$. Utilizando integração por partes tem-se (GREEN e SILVERMAN, 1994, p.16):

$$\begin{aligned} \int_a^b g''(t) h''(t) dt &= (uv)_a^b - \int_a^b du v = (g''(t)h'(t))_a^b - \int_a^b g'''(t)h'(t) dt \\ \text{com, } u &= g''(t) \quad du = g'''(t)dt \\ dv &= h''(t) \quad v = h'(t)dt \end{aligned}$$

Como pela definição de spline cúbica natural g''' é zero em a e b , tem-se:

$$\int_a^b g''(t) h''(t) dt = - \int_a^b g'''(t)h'(t) dt.$$

Também pela definição de spline cúbica natural, g''' é nula em (a, t_1) e (t_n, b) , além disso, g''' é constante em cada intervalo (t_j, t_{j+1}) com valor $g'''(t_j^+)$, assim:

$$\begin{aligned} \int_a^b g''(t) h''(t) dt &= - \int_a^b g'''(t)h'(t) dt = - \sum_{j=1}^{n-1} g'''(t_j^+) \int_{t_j}^{t_{j+1}} h'(t) dt = \\ &= - \sum_{j=1}^{n-1} g'''(t_j^+) \{h(t_{j+1}) - h(t_j)\} = 0 \end{aligned} \tag{A1}$$

Portanto:

$$\int_a^b \bar{g}''^2 = \int_a^b (g'' + h'')^2 = \int_a^b g''^2 + 2 \int_a^b g''h'' + \int_a^b h''^2 \stackrel{(A1)}{=} \int_a^b g''^2 + \int_a^b h''^2 \geq \int_a^b g''^2 \tag{A2}$$

A igualdade ocorrerá em (A2), se e somente se, $\int h''^2$ é nula. Dessa forma, h é linear em $[a, b]$. Mas como h é nula nos pontos $t_1, \dots, t_n$ com $n \geq 2$, isto só pode acontecer se h é identicamente zero, em outras palavras, se $\bar{g}$ e g são a mesma função (GREEN e SILVERMAN, 1994, p.17). ■

APÊNDICE B – Teorema fundamental da GCV

Teorema - Os valores da validação cruzada satisfazem

$$CV(\alpha) = n^{-1} \sum_{i=1}^n \left(\frac{Y_i - \hat{g}(t_i)}{1 - A_{ii}(\alpha)} \right)^2 \quad (B1)$$

onde $\hat{g}$ é a spline suavizadora calculada através do conjunto de dados $\{(t_i, Y_i)\}$ com parâmetro suavizador α .

Lema B1 – Para valores fixos de α e i , seja $\mathbf{g}^{(-i)}$ o vetor com componentes $g_j^{(-i)} = \hat{g}^{(-i)}(t_j; \alpha)$, e $\mathbf{Y}^*$ o vetor definido por:

$$\begin{cases} Y_j^* = Y_j \text{ para } j \neq i \\ Y_i^* = \hat{g}^{(-i)}(t_i) \end{cases}$$

Então:

$$\mathbf{g}^{(-i)} = A(\alpha) \mathbf{Y}^* \quad (B2)$$

Prova: Para uma curva suave g ,

$$\begin{aligned} \sum_{j=1}^n \{Y_j^* - g(t_j)\}^2 + \alpha \int g^{n2} &\geq \sum_{j \neq i} \{Y_j^* - g(t_j)\}^2 + \alpha \int g^{n2} \stackrel{\text{def. de } \hat{g}^{(-i)}}{\geq} \sum_{j \neq i} \{Y_j^* - \hat{g}^{(-i)}(t_j)\}^2 + \alpha \int \hat{g}^{(-i) n2} \\ &= \sum_{j=1}^n \{Y_j^* - \hat{g}^{(-i)}(t_j)\}^2 + \alpha \int \hat{g}^{(-i) n2} \end{aligned}$$

Segue que $\hat{g}^{(-i)}$ é o mínimo de $\sum_{j=1}^n \{Y_j^* - g(t_j)\}^2 + \alpha \int g^{n2}$, e dessa forma,

$\mathbf{g}^{(-i)} = A(\alpha) \mathbf{Y}^*$, como requerido pelo lema.

Como corolário, pode-se obter uma expressão para o resíduo não considerado $Y_i - \hat{g}^{(-i)}(t_i)$. Assim, escrevendo A em lugar de $A(\alpha)$, tem-se (GREEN e SILVERMAN, 1994, p.32):

$$\begin{aligned}
\hat{g}^{(-i)}(t_i) - Y_i &\stackrel{\text{lema B1}}{=} \sum_{j=1}^n A_{ij} Y_j^* - Y_i = \sum_{j \neq i} A_{ij} Y_j + A_{ii} \hat{g}^{(-i)}(t_i) - Y_i = \sum_{j=1}^n A_{ij} Y_j - Y_i + A_{ii} \{\hat{g}^{(-i)}(t_i) - Y_i\} \\
&= \hat{g}(t_i) - Y_i + A_{ii} \{\hat{g}^{(-i)}(t_i) - Y_i\}.
\end{aligned} \tag{B3}$$

Segue de (B3) que:

$$\begin{aligned}
\hat{g}^{(-i)}(t_i) - Y_i &= \hat{g}(t_i) - Y_i + A_{ii} \{\hat{g}^{(-i)}(t_i) - Y_i\} \Rightarrow 1 = \frac{\hat{g}(t_i) - Y_i}{\hat{g}^{(-i)}(t_i) - Y_i} + A_{ii} \Rightarrow \frac{\hat{g}(t_i) - Y_i}{\hat{g}^{(-i)}(t_i) - Y_i} = 1 - A_{ii} \\
\Rightarrow \hat{g}^{(-i)}(t_i) - Y_i &= \frac{\hat{g}(t_i) - Y_i}{1 - A_{ii}} \Rightarrow Y_i - \hat{g}^{(-i)}(t_i) = \frac{Y_i - \hat{g}(t_i)}{1 - A_{ii}(\alpha)}.
\end{aligned} \tag{B4}$$

Elevando (B4) ao quadrado, somando todos os n elementos e multiplicando

$\frac{1}{n}$ nos dois lados, tem-se:

$$CV(\alpha) = n^{-1} \sum_{i=1}^n \left(\frac{Y_i - \hat{g}(t_i)}{1 - A_{ii}(\alpha)} \right)^2$$

completando a prova do teorema. ■

APÊNDICE C – Tabelas dos resíduos para cada DD

Nesse apêndice são apresentadas a média absoluta, a média aritmética e o desvio padrão dos resíduos para os experimentos UEPP-Regente, UEPP-Venceslau, UEPP-Quintana e UEPP-Assis para 5 min e 10min de dados com o MMQ1 e o MMQP.

TABELA C1 – Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Regente com intervalo de 5 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ⁹⁻⁴	DD ⁹⁻⁵	DD ⁹⁻⁶	DD ⁹⁻²⁴	DD ⁹⁻³⁰
MMQ1	Média absoluta	0,034	0,015	0,134	0,015	0,013
	Média aritmética	1,6E-18	1,0E-04	-1,6E-04	5,2E-05	-1,1E-04
	Desvio padrão	0,042	0,020	0,158	0,020	0,016
MMQP	Média absoluta	0,025	0,013	0,016	0,014	0,011
	Média aritmética	5,8E-16	6,8E-16	1,5E-16	4,7E-16	3,9E-16
	Desvio padrão	0,031	0,016	0,019	0,016	0,014

TABELA C2 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Venceslau com intervalo de 5 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ⁵⁻⁶	DD ⁵⁻⁸	DD ⁵⁻¹⁰	DD ⁵⁻²¹	DD ⁵⁻²⁴	DD ⁵⁻³⁰
MMQ1	Média absoluta	0,007	0,046	0,082	0,035	0,025	0,071
	Média aritmética	5,2E-05	-1,1E-04	1,0E-04	2,1E-08	5,2E-05	-1,6E-04
	Desvio padrão	0,009	0,052	0,102	0,042	0,035	0,081
MMQP	Média absoluta	0,007	0,020	0,050	0,044	0,025	0,011
	Média aritmética	-9,2E-16	-6,1E-16	-7,9E-16	-6,8E-16	-6,3E-16	-6,6E-16
	Desvio padrão	0,009	0,025	0,057	0,052	0,031	0,015

TABELA C3 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Quintana com intervalo de 5 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ³⁰⁻⁵	DD ³⁰⁻⁶	DD ³⁰⁻⁹	DD ³⁰⁻¹⁰	DD ³⁰⁻²¹	DD ³⁰⁻²⁴
MMQ1	Média absoluta	0,016	0,022	0,049	0,053	0,090	0,139
	Média aritmética	-1,1E-04	1,0E-04	-1,0E-18	5,2E-05	5,2E-05	-1,6E-04
	Desvio padrão	0,021	0,026	0,058	0,064	0,103	0,165
MMQP	Média absoluta	0,015	0,022	0,018	0,009	0,049	0,039
	Média aritmética	-1,2E-17	-5,7E-18	9,9E-19	4,7E-18	-8,4E-17	-6,3E-17
	Desvio padrão	0,021	0,027	0,023	0,012	0,059	0,052

TABELA C4 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Assis com intervalo de 5 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD⁵⁻⁶	DD⁵⁻⁹	DD⁵⁻¹⁰	DD⁵⁻²¹	DD⁵⁻²⁴	DD⁵⁻³⁰
MMQ1	Média absoluta	0,118	0,258	0,183	0,110	0,034	0,203
	Média aritmética	1,1E-04	1,2E-17	1,1E-04	-1,1E-04	-1,5E-18	-5,5E-05
	Desvio padrão	0,135	0,301	0,206	0,131	0,039	0,239
MMQP	Média absoluta	0,041	0,032	0,038	0,015	0,013	0,017
	Média aritmética	1,9E-16	1,5E-16	2,2E-16	7,6E-17	1,9E-16	2,1E-16
	Desvio padrão	0,049	0,037	0,043	0,018	0,016	0,021

TABELA C5 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Regente com intervalo de 5 min de dados coletados durante a noite

Métodos	Resíduos (ciclos)	DD¹¹⁻¹	DD¹¹⁻²	DD¹¹⁻¹³	DD¹¹⁻¹⁶	DD¹¹⁻²⁰	DD¹¹⁻²⁵
MMQ1	Média absoluta	0,014	0,011	0,032	0,021	0,010	0,009
	Média aritmética	5,2E-05	-5,2E-05	1,1E-04	5,2E-05	1,8E-19	9,1E-20
	Desvio padrão	0,017	0,013	0,036	0,025	0,013	0,012
MMQP	Média absoluta	0,017	0,011	0,007	0,012	0,009	0,009
	Média aritmética	-1,5E-17	-4,3E-17	-2,7E-17	-5,5 E-17	-3,3 E-17	-3,8E-17
	Desvio padrão	0,021	0,013	0,010	0,016	0,011	0,012

TABELA C6 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Venceslau com intervalo de 5 min de dados coletados durante a noite

Métodos	Resíduos (ciclos)	DD²⁰⁻¹	DD²⁰⁻²	DD²⁰⁻⁴	DD²⁰⁻⁸	DD²⁰⁻¹³	DD²⁰⁻²⁷
MMQ1	Média absoluta	0,005	0,013	0,012	0,011	0,012	0,006
	Média aritmética	-3,6E-19	-5,2E-05	-1,6E-04	-5,2E-05	-5,2E-05	3,6E-19
	Desvio padrão	0,007	0,016	0,015	0,014	0,015	0,008
MMQP	Média absoluta	0,005	0,010	0,008	0,011	0,008	0,005
	Média aritmética	3,3E-17	4,3E-17	3,1E-17	6,0E-17	3,8E-17	2,8E-17
	Desvio padrão	0,007	0,014	0,010	0,014	0,009	0,007

TABELA C7 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Quintana com intervalo de 5 min de dados coletados durante a noite

Métodos	Resíduos (ciclos)	DD²⁷⁻¹	DD²⁷⁻²	DD²⁷⁻⁴	DD²⁷⁻⁸	DD²⁷⁻¹³	DD²⁷⁻²⁰
MMQ1	Média absoluta	0,010	0,057	0,032	0,047	0,080	0,036
	Média aritmética	2,6E-19	-5,2E-05	1,2E-05	4,0E-18	1,0E-04	5,2E-05
	Desvio padrão	0,012	0,064	0,040	0,055	0,094	0,041
MMQP	Média absoluta	0,010	0,016	0,024	0,018	0,007	0,007
	Média aritmética	-1,5E-17	-1,1E-17	-2,2E-17	7,6E-18	1,3E-17	-1,6E-17
	Desvio padrão	0,012	0,021	0,028	0,022	0,008	0,009

TABELA C8 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Assis com intervalo de 5 min de dados coletados durante a noite

Métodos	Resíduos (ciclos)	DD ²⁷⁻¹	DD ²⁷⁻²	DD ²⁷⁻⁴	DD ²⁷⁻⁸	DD ²⁷⁻¹³	DD ²⁷⁻²⁰
MMQ1	Média absoluta	0,030	0,048	0,040	0,026	0,038	0,011
	Média aritmética	-5,5E-05	5,5E-05	-1,1E-04	7,7E-19	3,0E-18	1,1E-04
	Desvio padrão	0,036	0,054	0,046	0,032	0,043	0,015
MMQP	Média absoluta	0,016	0,020	0,009	0,009	0,009	0,011
	Média aritmética	-2,8E-17	-1,0E-16	-2,0E-17	-3,9E-17	-4,2E-17	-3,5E-17
	Desvio padrão	0,020	0,024	0,011	0,012	0,010	0,015

TABELA C9 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Regente com 10 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ⁹⁻⁴	DD ⁹⁻⁵	DD ⁹⁻⁶	DD ⁹⁻²⁴	DD ⁹⁻³⁰
MMQ1	Média absoluta	0,093	0,048	0,175	0,037	0,040
	Média aritmética	6,0E-18	-2,5E-05	-2,5E-05	-7,6E-05	-5,1E-05
	Desvio padrão	0,102	0,055	0,200	0,050	0,049
MMQP	Média absoluta	0,041	0,022	0,034	0,029	0,041
	Média aritmética	2,9E-17	7,6E-18	1,2E-17	-3,6E-17	-1,4E-17
	Desvio padrão	0,047	0,026	0,038	0,034	0,050

TABELA C10 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Venceslau com intervalo de 10 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ⁵⁻⁶	DD ⁵⁻⁸	DD ⁵⁻¹⁰	DD ⁵⁻²¹	DD ⁵⁻²⁴	DD ⁵⁻³⁰
MMQ1	Média absoluta	0,048	0,058	0,079	0,105	0,099	0,109
	Média aritmética	-5,1E-05	-2,4E-18	2,5E-05	-5,1E-05	3,2E-18	2,5E-05
	Desvio padrão	0,055	0,068	0,103	0,124	0,109	0,131
MMQP	Média absoluta	0,046	0,056	0,088	0,067	0,063	0,063
	Média aritmética	-5,7E-17	-9,2E-17	8,1E-18	1,5E-16	2,5E-17	2,3E-17
	Desvio padrão	0,054	0,065	0,098	0,074	0,080	0,078

TABELA C11 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Quintana com intervalo de 10 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD ³⁰⁻⁵	DD ³⁰⁻⁶	DD ³⁰⁻⁹	DD ³⁰⁻¹⁰	DD ³⁰⁻²¹	DD ³⁰⁻²⁴
MMQ1	Média absoluta	0,026	0,043	0,080	0,078	0,157	0,089
	Média aritmética	4,3E-19	8,6E-19	8,6E-18	-3,1E-05	-1,9E-17	-1,8E-05
	Desvio padrão	0,032	0,051	0,099	0,100	0,172	0,103
MMQP	Média absoluta	0,025	0,034	0,021	0,038	0,063	0,048
	Média aritmética	-7,4E-17	-1,5E-16	-1,1E-16	-9,8E-17	-1,0E-16	-1,1E-16
	Desvio padrão	0,030	0,041	0,026	0,044	0,075	0,058

TABELA C12 - Média absoluta, média aritmética e desvio padrão dos resíduos para o experimento UEPP-Assis com intervalo de 10 min de dados coletados por volta das 14 h local

Métodos	Resíduos (ciclos)	DD⁵⁻⁶	DD⁵⁻⁹	DD⁵⁻¹⁰	DD⁵⁻²¹	DD⁵⁻²⁴	DD⁵⁻³⁰
MMQ1	Média absoluta	0,175	0,401	0,294	0,190	0,037	0,288
	Média aritmética	2,6E-05	-3,7E-17	-2,6E-05	2,6E-05	-0,023	-7,8E-05
	Desvio padrão	0,208	0,471	0,347	0,230	0,043	0,338
MMQP	Média absoluta	0,063	0,082	0,121	0,138	0,028	0,074
	Média aritmética	5,5E-16	-1,6E-16	1,5E-16	-1,6E-17	-6,3E-17	3,2E-17
	Desvio padrão	0,077	0,098	0,153	0,160	0,036	0,087