

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E COMPUTAÇÃO
DEPARTAMENTO DE COMUNICAÇÕES**

**AVALIAÇÃO OBJETIVA DE QUALIDADE DE CODECS DE VOZ NA
FAIXA DE TELEFONIA**

Autor: Jayme Garcia Arnal Barbedo

Orientador: Amauri Lopes

Dissertação submetida à Faculdade de Engenharia Elétrica e Computação, Departamento de Comunicações, como parte dos requisitos para obtenção do título de mestre.

Banca Examinadora

Prof. Dr. Amauri Lopes (presidente)
Prof. Dr. Fábio Violaro
Prof. Dr. João Batista Tadanobu Yabu-uti
Prof. Dr. José Sindi Yamamoto

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

B232a

Barbedo, Jayme Garcia Arnal.

Avaliação objetiva de qualidade de codecs de voz na
faixa de telefonia / Jayme Garcia Arnal Barbedo. --
Campinas, SP: [s.n.], 2001.

Orientador: Amauri Lopes.

Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Codificador de voz - Avaliação. 2. Sistemas
telefônicos. 3. Sistemas de telecomunicação. I. Lopes,
Amauri. II. Universidade Estadual de Campinas.
Faculdade de Engenharia Elétrica e de Computação. III.
Título.

Este projeto foi financiado pela FAPESP

Processo nº 99/01702-0

AGRADECIMENTOS

Agradeço e dedico este trabalho:

A meus pais, que sempre me incentivaram a ir adiante com meus sonhos e que nunca mediram esforços para que eles se realizassem.

À minha esposa, que tornou os desafios muito mais fáceis de serem vencidos, estando sempre a meu lado, nos bons e maus momentos.

A meus colegas, os quais forneceram um apoio inestimável nos momentos de maior necessidade.

A meu orientador, cuja atuação superou a melhor das expectativas, e ao qual devo boa parte dos créditos pelo sucesso deste trabalho.

A todos aqueles que, de uma maneira ou de outra, contribuíram para a realização deste trabalho.

RESUMO

Este trabalho aborda o problema da avaliação objetiva de qualidade de codecs de voz na faixa de telefonia (300-3400 Hz). São apresentados os procedimentos para as avaliações subjetivas e objetivas de qualidade e a modelagem matemática do ouvido humano. O método PSQM, padronizado pela “International Telecommunication Union”, é descrito com detalhes e serviu de base para a implementação de um programa de computador que apresenta uma série de recursos adicionais, dando origem a dois novos métodos, denominados Medida Objetiva de Qualidade de Voz 1 e 2 (MOQV1 e MOQV2). A validação e a caracterização do desempenho e da aplicabilidade destes métodos são feitas através de uma análise estatística detalhada, baseada em um banco de dados da ITU. Por fim, são apresentadas breves discussões sobre a utilização de medidas objetivas de qualidade de voz em telefonia móvel e sobre as principais tendências na avaliação objetiva de qualidade de codecs de áudio.

ABSTRACT

This work boards the problem of the objective quality assessment of telephone-band (300-3400 Hz) speech codecs. It is presented the procedures to the objective and subjective quality assessments and the mathematic modeling of the human ear. The PSQM method, standardized by the International Telecommunication Union, is described in detail and it was the base to the implementation of a computer program that presents some additional resources, and that led to two new methods, named MOQV1 and MOQV2. The validation and characterization of the performance and applicability of these methods were done by a detailed statistical analysis that was based in an ITU's database. At last, briefly discussions about the utilization of objective quality measurement of speech in mobile telephony and about the main tendencies in the objective quality assessment of audio codecs are presented.

ÍNDICE

1. Introdução	1
2. Métodos de Avaliação de Qualidade de Codecs de Voz	3
2.1. Medidas de Avaliação Subjetiva	3
2.1.1. Avaliação Subjetiva de Codecs de Voz	4
2.1.2. Tipos de Teste	4
2.2. Avaliação Objetiva	7
2.2.1. Procedimentos Gerais para a Realização de Medidas Objetivas	7
2.2.2. Medidas Objetivas mais Importantes	8
3. Modelo Psicoacústico para o Ouvido Humano	11
3.1. Fisiologia do Ouvido Humano	11
3.1.1. Modelo Matemático do Ouvido Humano	13
4. Medidas Objetivas Baseadas no Modelo Psicoacústico	21
4.1. Introdução	21
4.2. O Método PSQM	21
4.2.1. Visão Geral	21
4.2.2. Inicialização	22
4.2.3. Cálculo da Medida PSQM	23
4.2.4. Transformação da escala objetiva para escala subjetiva	29
4.2.5. Considerações Finais	30
4.3. O Método PSQM+	31
4.4. Outros Métodos Propostos	32
4.4.1. Método MNB (Measuring Normalizing Blocks)	32
4.4.2. Método PAMS (Perceptual Analyses Measurement System)	34
4.4.3. Método TOSQA (Telecomm. Objective Speech Quality Assessment)	36
4.4.4. Conclusão	37
5. Implementação de Programa de Computador Baseado no Método PSQM	38
5.1. Implementação em Matlab® 5.3	38
5.1.1. Entrada de Dados	38
5.1.2. Sub-rotina “geral”	39
5.1.3. Cálculo da Defasagem	39
5.1.4. Janelamento	40
5.1.5. Cálculo da FFT e da Densidade Espectral de Energia Quadro a Quadro	40
5.1.6. Mapeamento em Bandas	40
5.1.7. Fatores de Escalonamento	41
5.1.8. Filtragem na Faixa de Telefonia e Inserção de Ruído	41
5.1.9. Fator de Ponderação para MOQV2	41
5.1.10. Mapeamento	41
5.1.11. Valores Fornecidos na Saída do Programa	42
5.1.12. Velocidade de Processamento	42
5.1.13. Versão Executável do Programa	42
5.2. Visão Geral do Programa	42

6. Desempenho do Programa MOQV	48
6.1. Descrição do Banco de Dados “Supplement 23”	48
6.1.1. Preparação do Material Fonte	48
6.1.2. Descrição do Primeiro Experimento	49
6.1.3. Descrição do Segundo Experimento	51
6.1.4. Descrição do Terceiro Experimento	52
6.1.5. Visão Geral do Banco de Dados “Supplement 23”	53
6.2. Descrição dos Testes com o Programa MOQV	54
6.2.1. Processo de Mapeamento	54
6.2.2. Cálculo da Correlação entre as Medidas Objetivas após Mapeamento e Medidas Subjetivas	54
6.2.3. Critérios Utilizados para o Mapeamento	55
6.2.4. Obtenção das Curvas de Mapeamento	55
6.3. Resultados de Simulação	56
6.3.1. Resultados para o 1º Critério	57
6.3.2. Resultados para o 2º Critério	87
6.4. Resultados Gerais	92
6.4.1. Comparação entre as correlações obtidas	101
6.5. Conclusões	102
7. Conclusão	104
Bibliografia	106
Apêndice A - Medidas Objetivas de Qualidade em Telefonia Móvel	112
A.1. Introdução	112
A.2. Configuração e Resultados dos Experimentos	112
A.2.1. Primeiro Experimento - Avaliação da Qualidade para Palavras Isoladas.	112
A.2.2. Segundo Experimento - Avaliação Contínua da Qualidade de Voz Variante no Tempo	113
A.3. Avaliação da Qualidade Variável de Voz	115
Apêndice B - Medidas Objetivas de Qualidade para Áudio	118
B.1. Introdução	118
B.2. Propostas Apresentadas	118
B.3. O Método PEAQ	119
B.3.1. Versão Baseada na FFT	121
B.3.2. Versão Baseada em um Banco de Filtros	122
B.3.3. Cálculo das Variáveis de Saída do Modelo (MOV)	124
B.3.4. Estimação da Qualidade Perceptual Básica do Áudio	124
Apêndice C - Definições Complementares	125
C.1. dBA	125
C.2. Pitch	125
C.3. Terças-Oitavas	126
C.4. SPL	127
C.5. Característica de Transmissão IRS Modificada	127
C.6. dBov	127
C.7. MNRU (Unidade de Referência de Ruído Modulado)	128

CAPÍTULO 1

INTRODUÇÃO

O avanço das técnicas de processamento digital de sinais e da tecnologia proporcionou um crescente interesse em métodos e dispositivos de codificação de voz mais eficientes. A avaliação da qualidade de codecs de voz é necessária para o desenvolvimento destes dispositivos e também para o projeto de redes digitais de telecomunicações, as quais fazem uso dos mesmos.

A análise da qualidade da voz recebida envolve fatores que são difíceis de serem ponderados em medidas objetivas, como, por exemplo, o incômodo que uma determinada distorção ou ruído causam e a inteligibilidade de um sinal. Adicionalmente, o conceito de qualidade varia de acordo com a aplicação pretendida e com o público alvo, que pode ser mais ou menos exigente, dependendo de suas características culturais. Portanto, as medidas subjetivas são ainda largamente utilizadas, especialmente em casos em que há a necessidade de resultados realmente confiáveis. Porém, seu custo, complexidade e tempo demandado motivam fortemente a busca de métodos eficazes para a realização de medidas objetivas que estimem a qualidade subjetiva de maneira eficiente.

Assim, há uma forte motivação para a criação de medidas objetivas que possam ser utilizadas para o maior número possível de situações. Até 1980 não havia surgido nenhum método capaz fornecer estimativas precisas da qualidade subjetiva. A fim de preencher esta lacuna, a ITU (*International Telecommunication Union*) incentivou o desenvolvimento de propostas mais eficientes. Após uma série de testes, o método adotado e especificado na Rec. P.861 [1] de 1996 foi o *Perceptual Speech Quality Measure* (PSQM) [2,3,4], desenvolvido no PTT Research da Holanda, o qual serviu de base para o desenvolvimento deste trabalho.

Este trabalho tem como principal objetivo fornecer um melhor entendimento a respeito do funcionamento e aplicabilidade das principais medidas objetivas adotadas atualmente, com particular atenção ao método PSQM, e apontar as principais tendências no desenvolvimento de novos métodos que sejam mais abrangentes e eficientes. Como resultado, o método “Medida Objetiva de Qualidade de Voz” (MOQV), foi desenvolvido. Este, similar, em essência, ao método PSQM original, porém com maiores recursos, foi utilizado em diversos testes que resultaram em uma caracterização estatística detalhada, permitindo assim traçar um roteiro para a sua aplicação prática.

O Capítulo 2 apresenta uma breve descrição dos procedimentos e tipos de testes adotados na realização das medidas subjetivas. São apresentadas também as principais técnicas e abordagens utilizadas no desenvolvimento das medidas objetivas desenvolvidas ao longo do tempo.

No Capítulo 3 é apresentada uma descrição da anatomia do ouvido humano e de sua modelagem matemática, extraíndo daí os principais conceitos utilizados na maioria das medidas objetivas desenvolvidas nos últimos anos.

O Capítulo 4 descreve os procedimentos adotados no desenvolvimento do método PSQM, apresentando detalhadamente o processamento matemático realizado e os respectivos conceitos envolvidos. É feita também a apresentação de uma versão ligeiramente modificada do método PSQM, denominada de PSQM+. Por fim, são descritos, resumidamente, alguns métodos propostos como alternativa ao método PSQM.

No Capítulo 5 é feita a descrição de um programa de computador para a implementação dos métodos MOQV1 e MOQV2, os quais são, basicamente, versões com mais recursos dos métodos PSQM e PSQM+.

O Capítulo 6 apresenta toda a descrição dos testes realizados com os métodos MOQV1 e MOQV2 e os resultados obtidos. Estes permitiram traçar um guia para as aplicações em que estes métodos estimam adequadamente (ou não) os valores subjetivos e forneceram curvas de mapeamento entre os valores objetivos e subjetivos, as quais podem ser utilizadas em situações práticas, sendo, por esse motivo, incluídas no programa desenvolvido.

No Apêndice A, é abordada a problemática do uso de medidas objetivas de qualidade de voz na telefonia móvel.

No Apêndice B são apresentadas as principais características das medidas objetivas de qualidade para áudio. É feita também uma breve descrição do método PEAQ, o qual representa uma compilação dos pontos mais fortes de diversas propostas apresentadas nos últimos anos.

Por fim, o Apêndice C apresenta as definições para vários conceitos que foram utilizados ao longo da dissertação.

CAPÍTULO 2

MÉTODOS DE AVALIAÇÃO DE QUALIDADE DE CODECS DE VOZ

2.1. MEDIDAS DE AVALIAÇÃO SUBJETIVA

Testes subjetivos de qualidade de transmissão são amplamente empregados na avaliação do desempenho de codecs digitais, uma vez que as medidas objetivas simples de desempenho não são suficientes para assegurar um desempenho adequado quando a transmissão é digital e/ou não-linear. Num sistema de comunicação de dados, o objetivo perseguido é minimizar a taxa de erros entre o transmissor e o receptor. Neste caso, a avaliação de desempenho é direta: quanto menor a taxa de erros, melhor o desempenho. Esta é uma medida de desempenho objetiva. Porém, em um sistema com interlocutores, ou quando a mensagem deve ser ouvida ou vista pelo destinatário, há fatores envolvidos no processo de análise da mensagem recebida que são difíceis de serem ponderados em medidas objetivas, como o incômodo que uma determinada distorção ou ruído causam, a inteligibilidade de um sinal ou mesmo suas peculiaridades culturais. Note-se ainda que o conceito de qualidade é relativo, variando muito de acordo com a aplicação pretendida. Citando dois extremos, tem-se as aplicações militares, onde qualidade significa inteligibilidade, enquanto que em áudio profissional a exigência é muito maior, devido à importância dada à fidelidade. O conceito de qualidade varia também de acordo com o público alvo, que pode ser mais ou menos exigente, dependendo de suas características culturais. Então, apesar de as medidas objetivas serem ferramentas muito importantes no processo de desenvolvimento de um sistema com interlocutores, ainda não se pode abrir mão do uso das medidas subjetivas em situações em que há a necessidade de resultados realmente confiáveis.

A Recomendação ITU-T P.800 da International Telecommunication Union [5] especifica os métodos para a determinação subjetiva da qualidade de transmissão em geral, enquanto que a Rec. P.830 [6] trata da medida subjetiva de desempenho de codecs digitais para a faixa de telefonia e faixa larga.

Estas recomendações mostram que os testes subjetivos devem reproduzir em ambiente de laboratório, tanto quanto possível, as condições reais experimentadas pelos usuários do sistema. Isto exige que as condições de conexão e o material falado sejam escolhidos de forma adequada e que os testes sejam administrados de forma conveniente. É importante que as condições simuladas no teste sejam corretamente especificadas, implementadas e continuamente monitoradas, que facilidades auxiliares como discagem e chamada sejam providenciadas e que gravações fiéis dos resultados de cada teste sejam tomadas.

2.1.1. Avaliação Subjetiva de Codecs de Voz

Quando se realiza uma avaliação subjetiva de um codec de voz, deve-se levar em conta a aplicação intencionada para o mesmo. O sistema utilizado para gravar os arquivos de voz, a seleção das condições de teste e o sistema utilizado para apresentar a voz processada aos ouvintes devem ser escolhidos de forma a capturar as características importantes da conexão telefônica. Em particular, os sistemas de gravação e recepção devem refletir as características do transmissor e do receptor que serão utilizados na aplicação do codec [7, 8, 9].

Diversos fatores devem ser avaliados num teste subjetivo de um codec de voz. O grau de importância a ser dado a cada um deles será determinado pelo tipo de aplicação desejada. Dentre os fatores mais importantes, destacam-se: a variação do nível de entrada do sinal a ser codificado; a ocorrência de erros no canal de operação do codec; a possível ocorrência de múltiplas transcódificações; a diversidade de locutores; a presença de ruído ambiente em níveis variáveis; sinais dos quais o usuário requer boa qualidade; e o atraso introduzido pela codificação.

Por fim, deve-se tomar um cuidado especial na escolha quantitativa e qualitativa dos avaliadores a serem utilizados nos testes, bem como na escolha do material-fonte. Este deverá conter uma amostragem significativa dos fonemas existentes, das pronúncias e vocabulários correntes e de tipos de voz, a fim de garantir um rigor suficiente para a análise de desempenho de um codec [7].

2.1.2. Tipos de Teste

Os testes subjetivos de avaliação de codecs de voz podem ser divididos em três grupos: os conversacionais, os de entrevista e os de audição [5].

Os *testes conversacionais* envolvem duas pessoas que tenham sido especificamente treinadas, uma falando e a outra ouvindo (conversação - teste bidirecional), dentro de uma certa metodologia.

Os *testes de entrevista* são realizados através de uma série de questões feitas diretamente aos usuários do sistema sob avaliação [10]. Este é um teste bastante dispendioso, tanto em termos de custos quanto de tempo requerido, já que, para se ter um grau de precisão elevado, é necessário a realização de pelo menos 100 entrevistas. Adicionalmente, existe a desvantagem de não se ter um controle rigoroso das características das conexões telefônicas sob teste. Porém, este método fornece uma apreciação global de como o equipamento se comporta no ambiente real. Este tipo de teste é recomendado apenas para casos em que este tipo de avaliação venha a contribuir de maneira decisiva na análise global do desempenho dos dispositivos sob teste.

Os testes subjetivos mais importantes são os *testes de audição*, os quais são unidirecionais e visam medir a capacidade de um sistema de transmitir adequadamente uma informação. Baseiam-se na avaliação de sua qualidade (de acordo com uma escala apropriadamente escolhida), a partir da audição de sentenças simples, sem relação semântica entre elas, processadas pelo sistema sob avaliação e por sistemas (condições) de referência. Neste tipo de teste, os locutores não devem ser treinados, recebendo apenas instruções quanto à escala de avaliação a ser utilizada.

Existem diversos tipos de testes de audição, sendo que os três mais utilizados são descritos a seguir:

* *Testes de qualidade absoluta (Absolute Category Rating – ACR)*: baseiam-se na avaliação absoluta da qualidade do material processado, sem que o avaliador disponha de material para comparação. Este tipo de teste, por sua vez, utiliza três escalas de opinião:

- *Qualidade de audição (Listening-Quality)*: nesta escala, um sistema de pontuação define a qualidade de pequenos grupos de sentenças descorrelacionadas, cada uma das quais submetida ao processo sob teste. A tabela 2.1 mostra a graduação utilizada nesta escala.

Tabela 2.1 - Escala “Listening-Quality”

Qualidade de voz	Pontos
Excelente	5
Boa	4
Razoável	3
Pobre	2
Ruim	1

- *Esforço de audição (Listening-Effort)*: esta é uma escala muito importante, principalmente nos casos em que há altos níveis de degradação. Neste tipo de situação, seu conhecimento é necessário para evitar erros na interpretação de outras escalas que estejam sendo examinadas, pois existe aqui um maior interesse na inteligibilidade do sinal, ficando sua qualidade em segundo plano. Portanto, uma transmissão pode receber uma pontuação fraca na escala “listening-quality” e, no entanto, ser perfeitamente inteligível (boa pontuação na escala “listening-effort”), o que é aceitável em diversas aplicações, como no caso de comunicações militares. A tabela 2.2 mostra a graduação utilizada nesta escala.

Tabela 2.2 - Escala “Listening-Effort”

Esforço requerido para a compreensão do significado das sentenças	Pontos
Relaxamento completo possível, nenhum esforço requerido	5
Atenção necessária, pouco esforço requerido	4
Esforço moderado requerido	3
Considerável esforço requerido	2
Nenhum significado é reconhecido, qualquer que seja o esforço	1

- *Preferência de sonoridade (Loudness-Preference)*: esta escala define o grau de sonoridade (volume) percebido pelos ouvintes. Sua graduação é mostrada na tabela 2.3.

Tabela 2.3 - Escala “Loudness-Preference”

Preferência de sonoridade	Pontos
Muito mais alto que o ideal	5
Mais alto que o ideal	4
Ideal	3
Mais baixo que o ideal	2
Muito mais baixo que o ideal	1

* *Testes de degradação (Degradation Category Rating – DCR)*: os testes tipo ACR não são apropriados para a distinção entre sistemas de boa qualidade. Por esse motivo, foi criado um novo tipo de teste, o qual avalia a degradação do material processado em relação ao material original, o que o torna mais sensível à distinção de qualidade, em contraste com os testes tipo ACR, onde a avaliação é absoluta. A escala aqui utilizada é mostrada na tabela 2.4.

Tabela 2.4 - Escala de degradação

Degradação	Pontos
A degradação é inaudível	5
A degradação é audível, mas não é incômoda	4
A degradação incomoda um pouco	3
A degradação incomoda	2
A degradação incomoda muito	1

* *Testes de comparação (Comparison Category Rating – CCR)*: este método se distingue do teste tipo DCR apenas pela ordem em que as amostras são apresentadas aos ouvintes. No método DCR, a amostra não-processada sempre é apresentada primeiro, sendo apresentada em seguida a amostra processada. Já neste método, a ordem das amostras é escolhida aleatoriamente, sendo que em metade dos casos a amostra original é apresentada em primeiro lugar. Portanto, neste tipo de teste, os ouvintes têm de responder a duas perguntas: qual dos sinais é melhor e quanto ele é melhor, segundo a escala da tabela 2.5.

Tabela 2.5 - Escala para comparação entre elementos de um par

Comparação entre os sinais de um par	Pontos
Muito melhor	3
Melhor	2
Ligeiramente melhor	1
Aproximadamente igual	0
Ligeiramente pior	-1
Pior	-2
Muito pior	-3

A vantagem do método CCR em relação ao DCR está na possibilidade de se poder avaliar não apenas processamentos de voz em que a qualidade é degradada, como também os casos em que a qualidade é melhorada. A deficiência deste tipo de teste, assim como no DCR, é que apenas desempenhos relativos podem ser obtidos.

O método mais utilizado, para a maioria das aplicações, tem sido o *Absolute Category Rating* (ACR), usando a escala *Listening Quality*. Este método está bem estabelecido e tem sido aplicado a conexões telefônicas digitais e analógicas e dispositivos de telecomunicações. Vários laboratórios em diferentes países realizaram testes subjetivos utilizando este método, nas mesmas condições e com sistemas de transmissão idênticos, conseguindo resultados com alto grau de consistência. A média aritmética dos pontos atribuídos é denominada de “Mean Opinion Score” (MOS). As escalas de degradação e de comparação entre elementos de um par, usando, respectivamente, os métodos de

classificação DCR e CCR, também têm sido largamente utilizadas, e sua média aritmética é denominada “Comparative Mean Opinion Score” (CMOS).

2.2. AVALIAÇÃO OBJETIVA

Em geral, as avaliações realizadas na forma subjetiva são dispendiosas, lentas e exigem infraestrutura adequada e esforços de organização. Assim, é natural a busca por métodos objetivos de avaliação que permitam reduzir a quantidade de testes subjetivos necessários para otimização, homologação ou monitoração de codecs e outros equipamentos e sistemas.

Deseja-se que os métodos objetivos de avaliação apresentem uma complexidade computacional adequada, além de estimar de maneira precisa os valores subjetivos correspondentes. Até o presente momento, ainda não se conseguiu elaborar métodos ou processos que atendam a estes requisitos, porém avanços significativos têm sido alcançados.

A seguir, são apresentados os principais procedimentos padronizados para a realização de medidas objetivas de qualidade de voz.

2.2.1. Procedimentos Gerais para a Realização de Medidas Objetivas

Os procedimentos descritos a seguir, válidos para todas as medidas objetivas de avaliação da qualidade de codecs de voz, foram determinados através de uma série de estudos [1]:

? *Preparação do Material Falado*

Os sinais usados nas medidas objetivas podem ser compostos por vozes reais ou por vozes artificiais, dependendo dos objetivos do experimento [11]. As vozes artificiais podem ser utilizadas na maioria das aplicações, excetuando-se aquelas em que a dependência do locutor é levada em consideração para linguagens particulares.

Vozes Reais: o número mínimo de locutores recomendado em um teste é de duas vozes masculinas e duas vozes femininas. Para testar a dependência dos locutores, o mínimo recomendado é de 8 vozes masculinas, 8 vozes femininas e 8 vozes infantis.

Vozes Artificiais: deve-se utilizar vozes artificiais femininas e masculinas. Estes sinais devem ser submetidos a um filtro com características em frequência apropriadas para simular as características de um sistema telefônico [11].

? *Seleção dos Parâmetros Experimentais*

Para demonstrar o desempenho de um codec, os efeitos de vários fatores de qualidade devem ser examinados [6]:

- níveis de entrada de voz para um codec;
- níveis de audição nos experimentos subjetivos;
- locutores (incluindo locutores múltiplos simultaneamente);
- erros no canal de transmissão entre o codificador e o decodificador;
- taxa de bits em codecs (no caso de haver mais de uma);
- transcodificadores;
- desencontro entre a taxa de bits do codificador e do decodificador;
- ruído de ambiente no lado do transmissor;

- música como a entrada de um codec.

A Rec. P.830 traz maiores informações a respeito destes parâmetros. Como já citado, alguns dos parâmetros acima ainda não podem ser corretamente avaliados através de medidas objetivas, sendo por isso objeto de estudo de diversos grupos de pesquisa.

2.2.2. Medidas Objetivas mais Importantes

As medidas objetivas existentes podem ser classificadas em 3 diferentes classes: medidas temporais, baseadas na análise temporal da forma de onda do sinal [12]; medidas espectrais, baseadas na estimação de parâmetros da representação espectral do sinal [13]; e as medidas psicoacústicas, baseadas em modelos de percepção da fala e que levam em conta o mecanismo de audição do ouvido humano [14]. Existem ainda técnicas mistas, mesclando aspectos de duas ou das três classes [15]. A seguir, são apresentados alguns exemplos de medidas pertencentes a cada uma dessas classes.

- Classe de Medidas Temporais:

Relação Sinal-Ruído (SNR) [16]: é a relação entre a potência de um sinal de referência e a potência de um sinal de erro. Apesar de ser adequada para avaliar a qualidade de sinais que não sejam de voz (dados, por exemplo), sabe-se que esta é uma das piores medidas objetivas de qualidade [17].

Relação Sinal-Ruído Segmentada: neste caso, o cálculo da SNR é feito localmente, segmentando-se consecutivamente o sinal. A SNR é então definida como o valor médio das SNRs dos segmentos analisados, resolvendo o problema das discrepâncias entre a SNR total e a qualidade subjetiva em sinais não-estacionários, para os quais a estimativa da variância (potência) do sinal deve ser local, e não global. Os pesos adotados para os segmentos de alta e baixa potência são semelhantes, a fim de se considerar o efeito do Ruído de Canal Desocupado (ICN), bastante incômodo na avaliação subjetiva, mas mascarado no cálculo da SNR convencional. Esta característica, no entanto, faz com que este método tenda a falhar em trechos de pequena potência, já que a sua contribuição para o valor final da SNR terá um impacto maior que aquele percebido subjetivamente. Algumas modificações, relativamente bem sucedidas, foram propostas, com a finalidade de contornar este efeito [18,19,20].

Este método deu origem às primeiras medidas objetivas bem sucedidas de inteligibilidade para sinais de voz: em 1947, surgiu o “Índice de Articulação” (AI) [21], o qual é baseado no cálculo da média das SNR obtidas em cinco bandas de frequência. Apenas em 1980 surgiria uma nova proposta, conhecida como Índice de Transmissão de Voz (STI) [22]. Este método é, basicamente, uma versão aperfeiçoada do Índice de Articulação, levando em conta diversos fatores adicionais, como por exemplo as distorções provocadas por reverberações e reflexões. A aplicabilidade destes dois métodos, contudo, se mostrou limitada a algumas poucas situações [23].

- Classe de Medidas Espectrais:

Medem a diferença entre as magnitudes dos espectros de um sinal processado e de um sinal de referência. Por serem baseadas na magnitude do espectro do sinal de voz, diferenças de fase não deveriam afetar o cálculo da distorção espectral de longo prazo, o que realmente acontece para sinais estacionários. No entanto, para sinais não-estacionários, faz-se necessário o uso de técnicas de compensação contínua do atraso, como no caso da utilização de pequenos trechos de voz. Nesta classe de medidas, valores próximos de zero

para a diferença espectral indicam sinais semelhantes, enquanto valores distantes indicam fortes distorções.

Medidas de Distância: utiliza-se um conjunto de normas e distâncias, calculadas localmente em segmentos do espectro, calculando-se posteriormente a média sobre todos os segmentos [11,17].

Distância Cepstral: surgida a partir de técnicas de processamento homomórfico [24], onde sinais misturados por convolução podem ser decompostos em excitação e filtro através de processamento adequado. Este processamento calcula o logaritmo do espectro de amplitude, denominado de *cepstrum*. A distância cepstral é obtida a partir da diferença entre os *cepstrums* referentes ao trato vocal para o sinal processado e para o sinal original. Estudos mostram que esta medida apresenta uma alta correlação com as medidas subjetivas de qualidade. Por esse motivo, a NTT do Japão desenvolveu um método de medida objetiva de qualidade utilizando os coeficientes cepstrais dos sinais de entrada e de saída de codecs de voz, a fim de definir uma distância cepstral [25].

Distância Espectral: esta medida distingue-se da distância cepstral em dois aspectos: a distância espectral exige apenas uma limitação de faixa dos sinais, enquanto que a distância cepstral exige um sinal que possa ser decomposto em excitação e filtro através do processamento adequado (o que sempre acontece para voz); além disso, a distância espectral utiliza o espectro total do sinal, enquanto que a distância cepstral se baseia apenas na envoltória do espectro de potência.

Razão de Verossimilhança: é uma medida de distorção para uso em reconhecimento de voz [26]. Ela avalia a diferença entre os modelos LPC para um sinal de teste e para outro de referência, através da comparação do erro de predição para esses sinais. Esta medida, como a distância cepstral, reflete distorções sobre a envoltória, e não sobre o espectro total.

- Técnicas Mistas:

Índice de informação: o índice de informação tenta considerar em sua formulação os erros de transmissão, o ruído ambiente, atenuações e distorções em frequência. Aqui, o sistema auditivo é modelado pela divisão do espectro em 16 bandas críticas, às quais se aplicam ponderações e limiares de audição obtidos previamente, de maneira empírica. O *Centre d'Études des Télécommunications* da França se baseou neste conceito para desenvolver um método de avaliação objetiva [27,28].

Função de Coerência: a função de coerência é uma medida da relação entre sinal e distorção (SDR), levando em conta a sensibilidade de audição, efeitos de limiar de ruído (mascaramento) e sensibilidade do terminal receptor. Neste método, segmentos de voz são classificados em quartis (cada quartil correspondendo a uma faixa de 25% da potência máxima; por exemplo, o primeiro quartil se refere aos segmentos com potência entre 0 e 25%). Calcula-se então os espectros dos sinais de referência e teste, bem como o espectro cruzado, para todos os segmentos de cada quartil. A seguir, calcula-se os espectros médios para cada quartil, os quais são utilizados no cálculo da função de coerência, a qual pode ser interpretada como a correlação entre os sinais para uma dada frequência. Esta função é utilizada no cálculo das potências dos sinais de referência e teste, que por sua vez são utilizadas no cálculo de uma SDR modificada, no qual são levados em conta os efeitos do limiar de audição e a sensibilidade do terminal receptor. Por fim, realiza-se um mapeamento para um valor MOS, onde há uma ponderação para a importância de cada

quartil e também de cada frequência. Este é o método proposto pela *Bell Northern Research* do Canadá [29,30].

- *Método Estatístico (EPR)*: este modelo utiliza vários parâmetros de voz, compilados através da utilização de redes neurais, e um mapeamento não-linear para obter uma medida de qualidade, através de um processo de treinamento com um conjunto de dados, criando um banco de dados de referência para análise das medidas obtidas. Foi desenvolvido pela *National Telecommunication Information Administration* dos EUA [31].

- Classe de Medidas Psicoacústicas:

Os métodos desta classe utilizam modelos matemáticos para o processamento realizado pelo ouvido sobre os sinais acústicos, ou seja, modelam o processo de percepção da fala através de modelos psicoacústicos. Utilizando tais modelos, são geradas versões do sinal original e do sinal codificado, as quais são comparadas através das diferenças entre ambas.

Esta é a classe de medidas mais bem sucedida, sendo por isso a base para a maior parte dos métodos desenvolvidos nos últimos anos. Esta discussão será aprofundada mais adiante, após a apresentação dos modelos psicoacústicos.

CAPÍTULO 3

MODELO PSICOACÚSTICO PARA O OUVIDO HUMANO

O pré-processamento realizado pelo ouvido sobre o sinal acústico é uma atividade objetiva, pois envolve a transformação do sinal acústico que chega ao ouvido externo em impulsos elétricos nos feixes de neurônios distribuídos ao longo da cóclea. O processamento subjetivo será realizado pelas funções superiores do córtex cerebral, baseado neste sinal condensado gerado pelo ouvido [12]. Esses processos ocorrem de maneira bastante homogênea de uma pessoa para outra. Por isso, um bom modelo dos processos envolvidos na percepção da fala pode ser aplicado para um vasto número de pessoas.

3.1. FISIOLOGIA DO OUVIDO HUMANO

Na figura 3.1 é mostrado um corte longitudinal do ouvido. A onda acústica que chega ao pavilhão auditivo é transformada em movimento das estruturas ósseas que compõem o ouvido médio (*martelo*, *bigorna* e *estribo*). Os ossos do ouvido médio estimulam a *cóclea* através da *janela oval*, fazendo com que seu líquido interno se movimente. A cóclea pode ser modelada como um tubo com duas câmaras separadas por uma estrutura chamada *membrana basilar*, como pode ser visto na figura 3.2. Na extremidade oposta à janela oval, existe um orifício sobre a membrana basilar que comunica essas duas câmaras, chamado de *helicotrema*. A membrana basilar apresenta uma resistência (mecânica) que varia ao longo de sua extensão: próximo à janela oval ela é mais fina e tensa, ressoando em frequências mais altas, enquanto no seu final (ápice), ela é espessa e flácida, ressoando então para frequências mais baixas. Sobre a membrana basilar existem ainda duas estruturas: as *fibras basilares* e o *órgão de Corti*. As fibras basilares são cerca de 20.000 espinhas delgadas com comprimentos que variam ao longo da membrana, sendo mais curtas junto à janela oval e mais longas no ápice da cóclea [32]. As ondas geradas pelo estribo, em resposta a um sinal senoidal, viajam ao longo da cóclea, fazendo vibrar a membrana basilar na mesma frequência do sinal de entrada [33]. Com isso, as fibras basilares vibram e estimulam as *células ciliadas*, que compõem o órgão de Corti, que por sua vez transformam o movimento das fibras basilares em impulsos nervosos. Estes são então transmitidos pelo nervo coclear para a região específica do córtex cerebral.

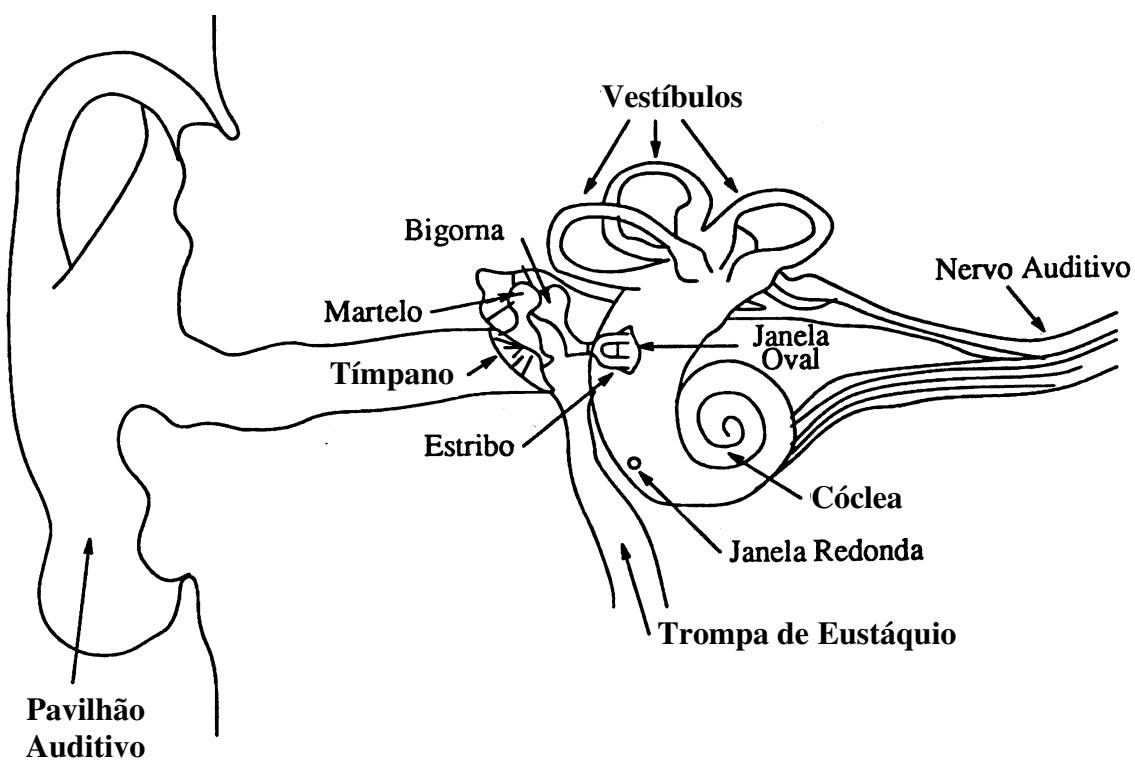


Figura 3.1 - Corte longitudinal do ouvido, com suas principais estruturas.

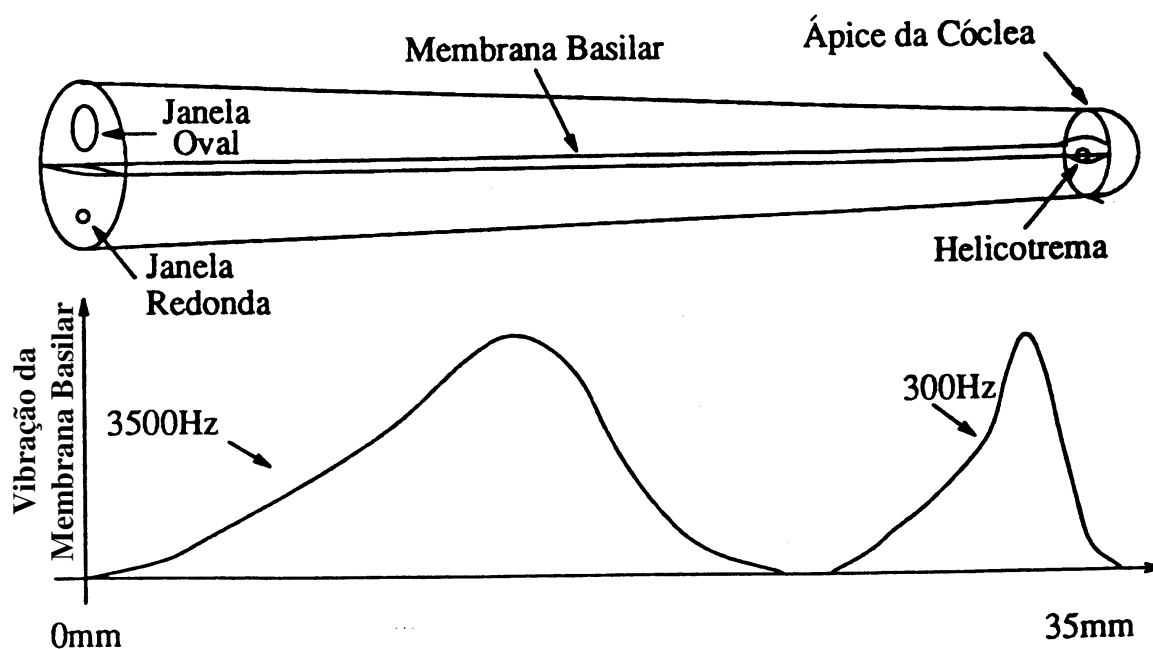


Figura 3.2 - Cóclea, membrana basilar e vibração da membrana em duas diferentes frequências.

Detalhando um pouco mais o que acontece nas duas primeiras porções do ouvido (externo e médio), tem-se o seguinte [34,35,36]:

- *Ouvido externo*: o ouvido externo desempenha um papel passivo no processo da audição. O pavilhão auricular e a cabeça, que possuem propriedades acústicas próprias, podem afetar a propagação do som no espaço e, conseqüentemente, afetar o sinal acústico que chega ao conduto auditivo, agindo como atenuador ou amplificador de sons de determinada frequência (efeito sombra e efeito ilusório). A concha tem uma ressonância de 5 kHz, enquanto que o restante irregular da orelha produz ressonâncias e anti-ressonâncias. A propriedade de ressonância faz com que os sons externos com frequência entre 2 e 5 kHz sofram um ganho de 10 a 15 dB. Este aumento permite a detecção e reconhecimento de sons de pequena energia e alta frequência, como os fricativos (ss, sch).

- *Ouvido médio*: a principal função do ouvido médio é melhorar a transmissão sonora para o ouvido interno, pois ele tem a capacidade de reduzir a reflexão que ocorre quando uma onda sonora incide em uma superfície fluida. Ou seja, pode-se considerá-lo um transformador de impedância que reduz a alta impedância do fluido coclear para um valor semelhante ao do ar. Por outro lado, os ossículos, fazendo a transmissão do som diretamente à janela oval, protegem a janela redonda do som, que chega a ela com menor intensidade. Se isto não ocorresse, o som chegaria ao mesmo tempo às duas janelas com a mesma magnitude e na mesma fase sonora, pois a distância entre elas é muito pequena, e isto não promoveria a movimentação do fluido do ouvido interno.

O processo de transformação do sinal acústico nas ondas do líquido coclear é chamado de função de transferência do ouvido médio. Ele é equivalente a uma filtragem passa-baixas com corte em 5 kHz, com uma sobrelevação na faixa entre 2.000 e 5.000 Hz e um pico em torno de 3.500 Hz [37,38]. Como essa filtragem não altera o espectro de forma significativa, ela é, em geral, desconsiderada para sinais com faixa até 5.000 Hz.

3.1.1. Modelo Matemático do Ouvido Humano

Cada ponto da membrana basilar é mais sensível a uma determinada frequência, chamada de *frequência característica*. Para um ponto específico da membrana basilar, a curva de resposta à frequência de vibração presente na janela oval é equivalente à de um filtro passa-faixa com fator de qualidade aproximadamente constante, resultando numa melhor resolução nas baixas frequências. Assim, as fibras basilares localizadas na região de altas frequências características respondem em uma maior faixa de frequências do que as fibras na região de baixas frequências características.

Um comportamento similar é obtido ao se traçar a curva de resposta ao longo da membrana basilar para um tom numa frequência específica, como se pode ver na figura 3.2. Para cada frequência, há um ponto da membrana basilar em que a vibração é máxima. A posição desse ponto, medida a partir do helicotrema, é aproximadamente proporcional ao logaritmo da frequência do som. Ao redor desse ponto haverá uma faixa, de aproximadamente 1,5 mm, onde a vibração estará presente, atenuando-se conforme se afasta do ponto.

Um aspecto importante da audição é o *mascaramento*, em que um som fica inaudível na presença de outro som. Muitos fenômenos de mascaramento podem ser explicados em termos de faixas de frequências conhecidas como *bandas críticas*, as quais foram determinadas através de experimentos psicoacústicos. Uma banda crítica define uma faixa em torno de uma frequência central. Quando dois sinais estimulam essa banda crítica, o de

maior energia dominará a percepção e mascarará o outro estímulo sonoro. Portanto, dois tons distintos só são distinguidos um do outro quando estão em bandas críticas diferentes. A resolução para a distinção entre uma frequência e outra varia de 100 Hz, nas frequências mais baixas, a mais de 6.000 Hz, nas frequências mais altas.

Para caracterizar melhor estas bandas críticas, deve-se considerar a curva de resposta em frequência do ouvido humano mostrada na figura 3.3. Este gráfico especifica a potência relativa necessária para a audição de um sinal em uma determinada frequência. Pode-se perceber que o ouvido humano é muito mais sensível na faixa de 2 a 4 kHz, sendo que o ponto de mínimo se situa na frequência de 3,3 kHz. A fala humana se situa aproximadamente entre 500 Hz e 2 kHz, sendo que as frequências mais baixas correspondem às vogais e as frequências mais altas às consoantes.

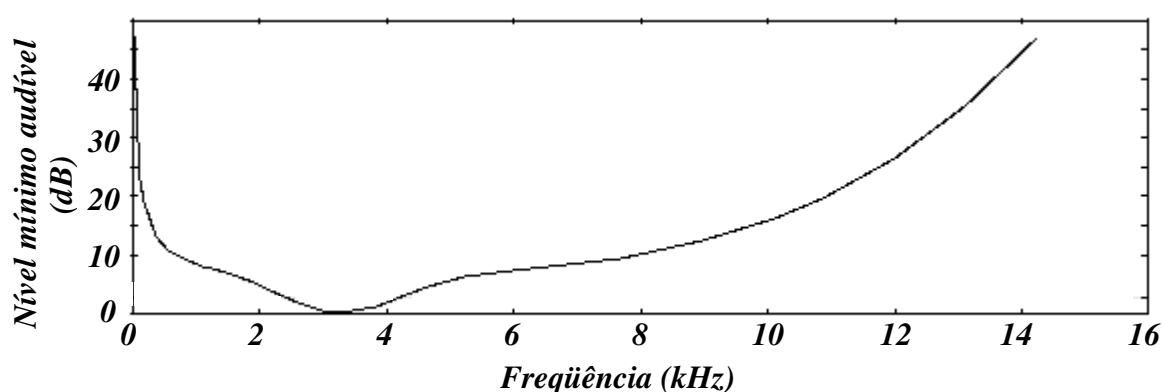


Figura 3.3 - Variação do nível mínimo audível com a frequência

O efeito de mascaramento pode ser sentido se uma experiência simples for feita: estimula-se o ouvido com um sinal de 1 kHz e, simultaneamente, com um outro sinal de frequência variável. Na figura 3.4, a linha tracejada representa a audibilidade do sinal de frequência variável sem a presença do sinal de 1 kHz, enquanto que a linha cheia representa a audibilidade do sinal na presença do tom em 1 kHz.

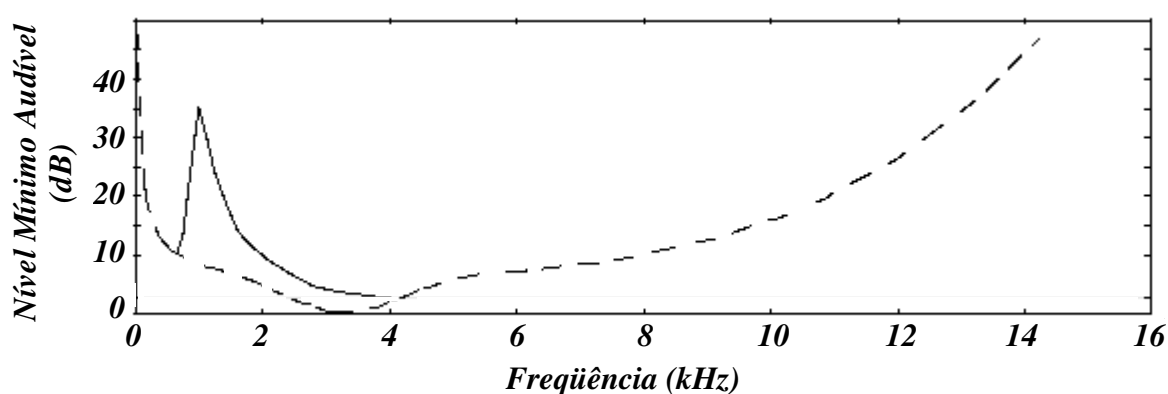


Figura 3.4 - Efeito de mascaramento em 1 kHz

Pode-se perceber que, na proximidade de 1 kHz, o sinal fica completamente mascarado pelo tom em 1 kHz, pois a potência necessária para distingui-lo, que era de aproximadamente 8 dB, passou para quase 40 dB. Pode-se definir uma “região de proximidade”, onde um sinal mascara todos os seus vizinhos, intitulada, como já comentado, de banda crítica. A figura 3.5 mostra o efeito de mascaramento para 4 frequências distintas. É importante notar que o mascaramento abrange uma largura de faixa menor para baixas frequências do que para altas frequências, o que é uma consequência direta da definição das bandas críticas.

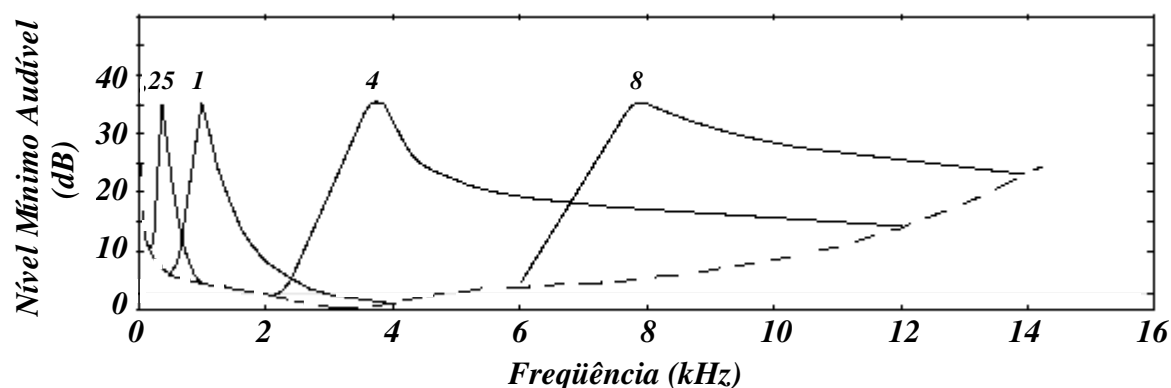


Figura 3.5 - Efeito de mascaramento para 4 frequências distintas

Entretanto, o mascaramento não ocorre apenas no domínio da frequência. Seu efeito pode ser sentido também no domínio do tempo. Após um sinal de grande amplitude, precisa-se de um certo tempo para se recuperar e conseguir ouvir um sinal de pequena amplitude. Pode-se fazer a seguinte experiência: estimula-se o ouvido com um sinal padrão de 1 kHz, a 60 dBA (ver definição de dBA no Apêndice C), e, logo após o término do sinal padrão, estimula-se o ouvido com um outro sinal de teste a uma frequência, por exemplo, de 1,1 kHz. O tempo necessário, em função da potência, para que o sinal de teste se torne audível é mostrado na figura 3.6.

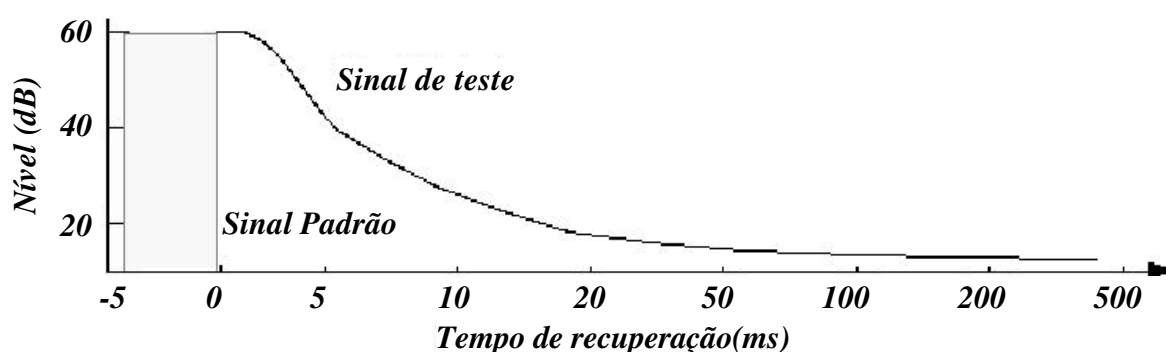


Figura 3.6 - Efeito do mascaramento temporal

De fato, acontece um fenômeno ainda mais curioso no sistema auditivo humano: um sinal forte o suficiente pode mascarar um sinal fraco que ocorra antes do sinal forte ser emitido. A explicação para este fato é que a informação leva um certo tempo para ser processada pelo cérebro. Entretanto, o tempo de pré-mascaramento é relativamente

pequeno (de 2 a 5 ms) em comparação com o tempo de pós-mascaramento, que pode atingir valores da ordem de 100 ms [37].

A figura 3.7 mostra o efeito combinado dos mascaramentos no domínio da frequência e do tempo. Por ocorrer em um período de tempo pequeno, o pré-mascaramento não é representado nesta figura.

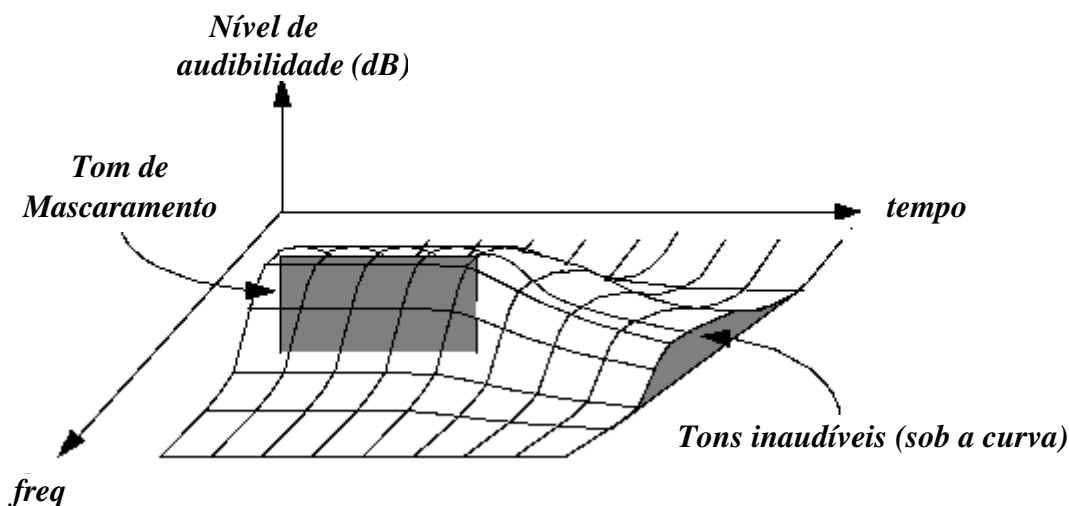


Figura 3.7 - Efeito combinado dos dois tipos de mascaramento

Outra característica interessante da percepção auditiva é a capacidade que o cérebro tem de “decidir”, numa série de harmônicas, qual é a fundamental. Além disso, no caso de remoção desta componente, este tem a capacidade de estimar onde ela deveria estar, a partir das componentes disponíveis, e então atribuir um *pitch* [35] (ver definição de *pitch* no Apêndice C).

É importante ressaltar que as bandas críticas podem ser definidas em torno de qualquer frequência central. A largura de faixa das bandas críticas corresponde a um espaçamento uniforme de 1,5 mm ao longo da membrana basilar, o que corresponde a aproximadamente 100 Hz para frequências abaixo de 500 Hz e de aproximadamente 1/6 da frequência central da banda para frequências acima de 1000 Hz (em direção à janela oval) [39]. Portanto, a resposta de amplitude em frequência, para cada banda crítica, pode ser modelada como a de um filtro passa-faixas com largura de faixa crescente com a frequência. Tais filtros possuem cortes bem acentuados: 65 dB/oitava para as bandas críticas em torno de 500 Hz e 100 dB/oitava em torno de 8 kHz.

Embora exista uma banda crítica ao redor de cada frequência, convencionou-se (com algumas pequenas variações) a adoção dos valores mostrados na tabela 3.1.

Tabela 3.1 - Bandas Críticas

Banda Crítica	Frequência (Hz)		
	Inferior	Superior	Faixa
0	0	100	100
1	100	200	100
2	200	300	100
3	300	400	100
4	400	510	110
5	510	630	120
6	630	770	140
7	770	920	150
8	920	1080	160
9	1080	1270	190
10	1270	1480	210
11	1480	1720	240
12	1720	2000	280

Banda Crítica	Frequência (Hz)		
	Inferior	Superior	Faixa
13	2000	2320	320
14	2320	2700	380
15	2700	3150	450
16	3150	3700	550
17	3700	4400	700
18	4400	5300	900
19	5300	6400	1100
20	6400	7700	1300
21	7700	9500	1800
22	9500	12000	2500
23	12000	15500	3500
24	15500	22050	6550

Zwicker [40] definiu uma escala perceptual, denominada escala *Bark*, que relaciona as frequências acústicas à resolução perceptual de frequência, de modo que 1 Bark cobre uma banda crítica. Foi então definida a relação entre as frequências f em Hz e os valores b na escala Bark, através da equação [38,40]:

$$b = 6 \ln \left(\frac{f}{600} + \sqrt{\frac{f^2}{600^2} + 1} \right) \quad \text{e} \quad f = 600 \sinh(b/6) \quad (3.1)$$

A curva de conversão representada por esta função é mostrada na figura 3.8.

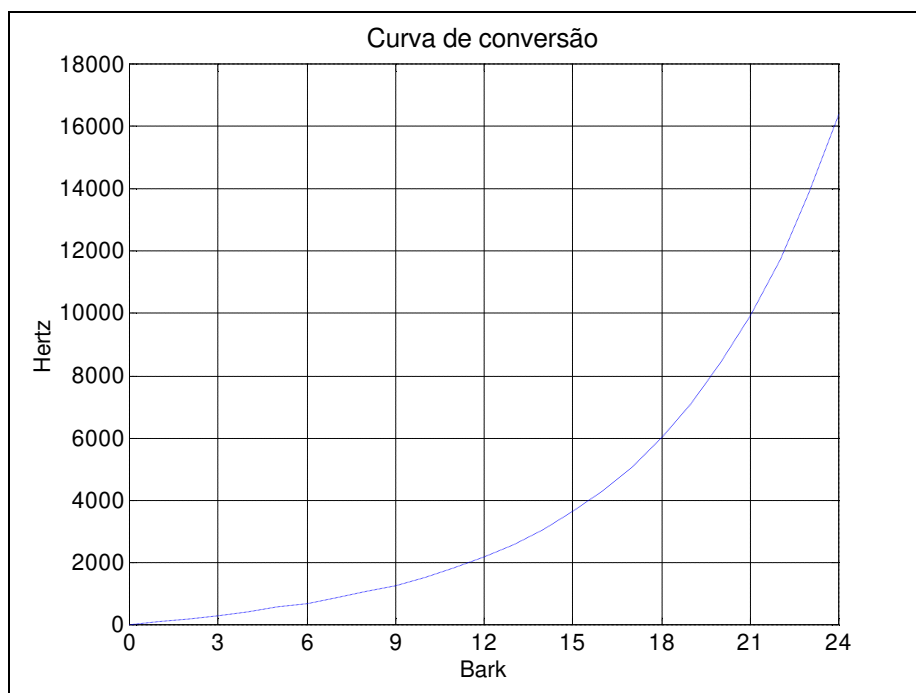


Figura 3.8 - Curva de conversão Hertz/Bark

O formato atribuído às bandas críticas, em que dois picos adjacentes ficam, por definição, espaçados de 1 Bark, é denominado *função espalhamento da membrana basilar*, e sua formulação varia consideravelmente entre os autores, já que alguns aproximam as curvas por segmentos de reta, enquanto outros utilizam funções analíticas. Por exemplo, ter-se-ia para cada uma das B bandas críticas [41]:

$$F(b) = \begin{cases} 0 & b - b_l < -1,3 \\ 10^{2,5(b-b_l+0,5)} & -1,3 < b - b_l < -0,5 \\ 1 & -0,5 < b - b_l < 0,5 \\ 10^{1,0[0,5-(b-b_l)]} & 0,5 < b - b_l < 2,5 \\ 0 & b - b_l > 2,5 \end{cases} \quad (3.2)$$

onde “l” se refere à l-ésima banda crítica e b_l é a frequência central, em Bark, dessa banda. Segundo esta função, duas bandas críticas adjacentes se encontram nas extremidades dos valores de pico, como mostra a figura 3.9. Em algumas formulações, este encontro se dá num ponto 3 dB abaixo do valor de pico [12].

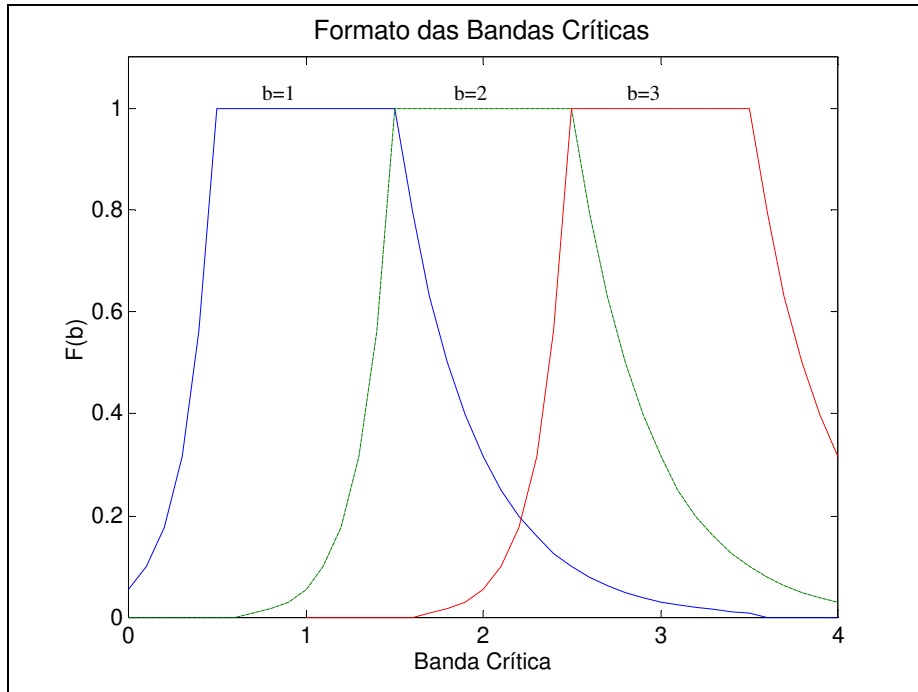


Figura 3.9 - Formato das 3 primeiras bandas críticas

O número total B de bandas críticas depende da faixa dos sinais considerados no modelo. Se um sinal tiver largura de faixa de f_0 Hz, então o número B de bandas críticas será:

$$B = 6 \ln \left(\frac{f_0}{600} + \sqrt{\left(\frac{f_0}{600} \right)^2 + 1} \right) \quad (3.3)$$

Para sinais entre 0 e 5 kHz, a escala Bark varia de 0 a 16,9 Barks. Para a faixa audível como um todo, a faixa considerada situa-se entre 0,5 e 24,5 Barks.

O padrão $D(b)$, gerado ao longo da cóclea por um sinal de faixa estreita, pode ser modelado como a convolução da energia desse sinal no domínio Bark, $Y(b)$, com a função de espalhamento da membrana basilar, $F(b)$:

$$D(b) = F(b) * Y(b) \quad (3.4)$$

O sinal $D(b)$ é chamado *padrão de excitação*, e pode ser entendido como a distribuição de energia ao longo da membrana basilar.

A geração de padrões de excitação para sinais de faixa larga, como a voz, apesar de ser um fenômeno complexo e não-linear, é geralmente aproximado pelo mesmo modelo linear da equação acima.

Com esse sinal em Bark, ainda é necessário considerar que nem todas as bandas críticas têm o mesmo ganho, ou seja, a sensibilidade não é uniforme ao longo da membrana basilar (a resposta auditiva humana começa a cair a partir dos 5 kHz, a uma taxa de 18 dB por oitava). A ponderação do padrão de excitação pela sensibilidade da membrana basilar equivale à conversão de um sinal que está numa escala de intensidade para uma escala de excitação, ou de *sonoridade*, expressa em *fônons*. Por definição, a sonoridade em fônons de um tom com certa frequência e nível é a intensidade em dB de um tom de 1 kHz que soe igualmente a esse tom.

Essa última conversão baseia-se em dados experimentais, e tem sido modelada como um processo de filtragem do padrão de excitação pela curva de correção de sonoridade, $E(b)$ [9,41]:

$$P(b) = E(b) \cdot D(b) \quad (3.5)$$

onde $P(b)$ é o sinal em fônons. Um exemplo de formulação da curva $E(b)$, por conveniência expresso em Hz, é dado por [41]:

$$E(f) = (2\pi)^2 \frac{(f^2 + 1200^2)f^4}{(f^2 + 400^2)(f^2 + 3100^2)} \quad (3.6)$$

No entanto, foi verificado ainda que a escala de sonoridade não é linear em relação à sonoridade percebida pelo ouvido humano. Por exemplo, se um sinal tem sonoridade em torno de 40 fônons, ao se adicionar 10 fônons ao sinal, a sonoridade percebida dobrará; porém, se o sinal estiver próximo do limiar de audição, a sonoridade percebida decuplicará. Fez-se então necessário transformar esta escala não-linear (fônons) para uma escala linear (sônons). Por definição, 1 *sônon* é o aumento de potência que faz dobrar a sonoridade percebida. A função de conversão de fônons para sônons é uma função de compressão não-linear, a qual é aproximada por [41]:

$$L(b) = [P(b)]^{0.33} \quad (3.7)$$

onde $L(b)$ é a sonoridade subjetiva (percebida pelo ouvinte), em contraste com a sonoridade objetiva $P(b)$. A figura 3.10 resume o processamento realizado pela modelagem psicoacústica.

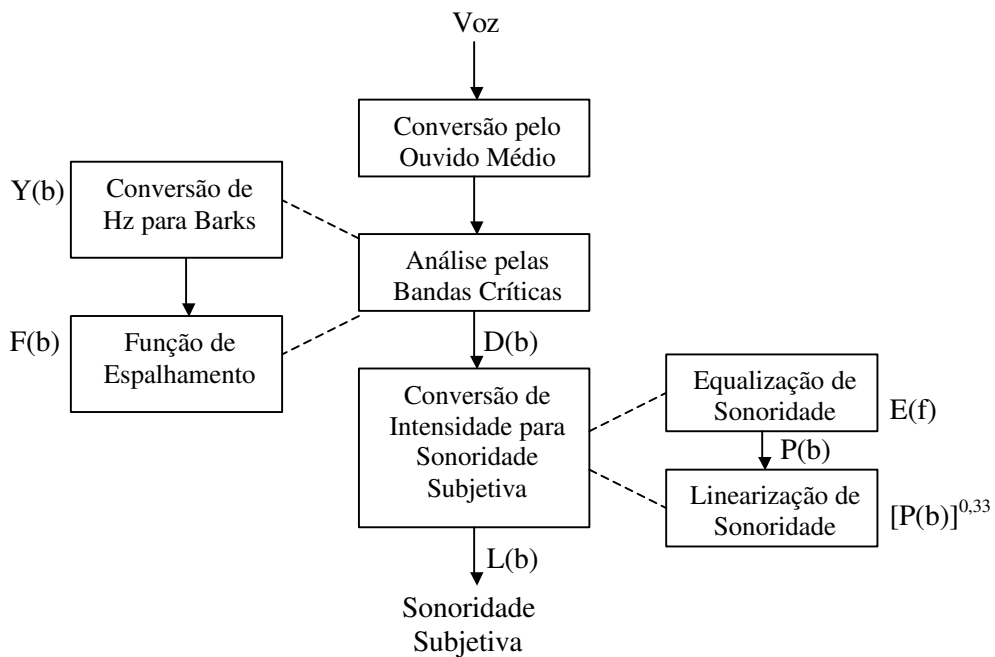


Figura 3.10 - Diagrama de blocos do modelo psicoacústico de percepção da fala.

A sonoridade subjetiva $L(b)$ representa a densidade espectral de energia perceptual enviada para os processos de percepção do córtex cerebral. Portanto, sobre $L(b)$ podem ser empregados vários processamentos, como predição linear, distância espectral perceptual, distância cepstral, etc., visando obter uma medida da distorção de um sinal em relação a outro de referência. Como exemplo, temos:

Distância Espectral Perceptual: considere os sinais de referência x e de teste y , cuja faixa é de F Hz, dos quais calculou-se a densidade espectral de energia perceptual, ou sonoridade subjetiva, $L_x(b)$ e $L_y(b)$. Pode-se então definir a Distância Espectral Perceptual como:

$$PSD = \sqrt{\sum_{b=1}^B [L_x(b) - L_y(b)]^2} \quad (3.8)$$

Esta distância mostra a diferença entre a densidade espectral percebida para os sinais de teste e referência.

Distância Cepstral Perceptual: uma variante de medida perceptual pode ser obtida ao se definir:

$$PCD = 10 \sqrt{\{\log_{10}[L_x(b)] - \log_{10}[L_y(b)]\}^2} \quad (3.9)$$

A modelagem psicoacústica aqui apresentada é utilizada em vários métodos de medida objetiva de qualidade. Os mais importantes serão descritos no próximo capítulo.

CAPÍTULO 4

MEDIDAS OBJETIVAS BASEADAS NO MODELO PSICOACÚSTICO

4.1. INTRODUÇÃO

A maior parte das medidas objetivas de qualidade desenvolvidas nos últimos anos tem se baseado na modelagem psicoacústica do ouvido humano. Dentre estas, serão aqui descritas aquelas que se destacaram por seu bom desempenho, como o MNB, proposto pela NTIA dos EUA [42], o PAMS, desenvolvido pela BT da Grã-Bretanha [43], o TOSQA, apresentado pela DTB da Alemanha [44] e o PSQM, desenvolvido pelo PTT Research da Holanda [2,3]. Este último apresentou os melhores resultados, merecendo por isso uma descrição mais detalhada. A figura 4.1 mostra o esquema básico para a realização das medidas, comum a todos os métodos acima citados.

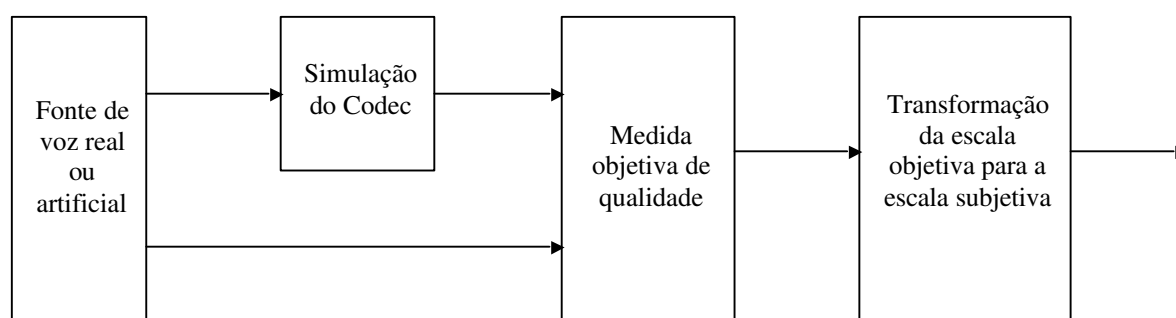


Figura 4.1 - Esquema dos procedimentos para avaliação objetiva

4.2. O MÉTODO PSQM

4.2.1. Visão Geral

O método PSQM utiliza o modelo do sistema auditivo humano descrito no capítulo anterior para gerar uma representação equivalente dos sons no interior do ouvido, denominada *representação interna de sons*. As representações internas para os sinais de entrada e de saída do codec são submetidas a um processamento denominado cognitivo, o qual fornece uma medida de qualidade baseada na diferença entre as duas representações.

A figura 4.2 mostra o esquema da proposta PSQM. Os sinais de entrada e de saída do codec (ou sistema sob teste) são mapeados pelo PSQM em representações psicofísicas similares àsquelas que são realizadas no interior de nossas cabeças. Estas representações internas fazem uso de equivalentes psicofísicos de frequência e intensidade.

A transformação do domínio físico externo para o psicofísico interno é realizada através de três operações:

- mapeamento tempo - frequência via transformada de Fourier;
- transformação de escalas de frequência (Hz ? Bandas Críticas);
- transformação de intensidade.

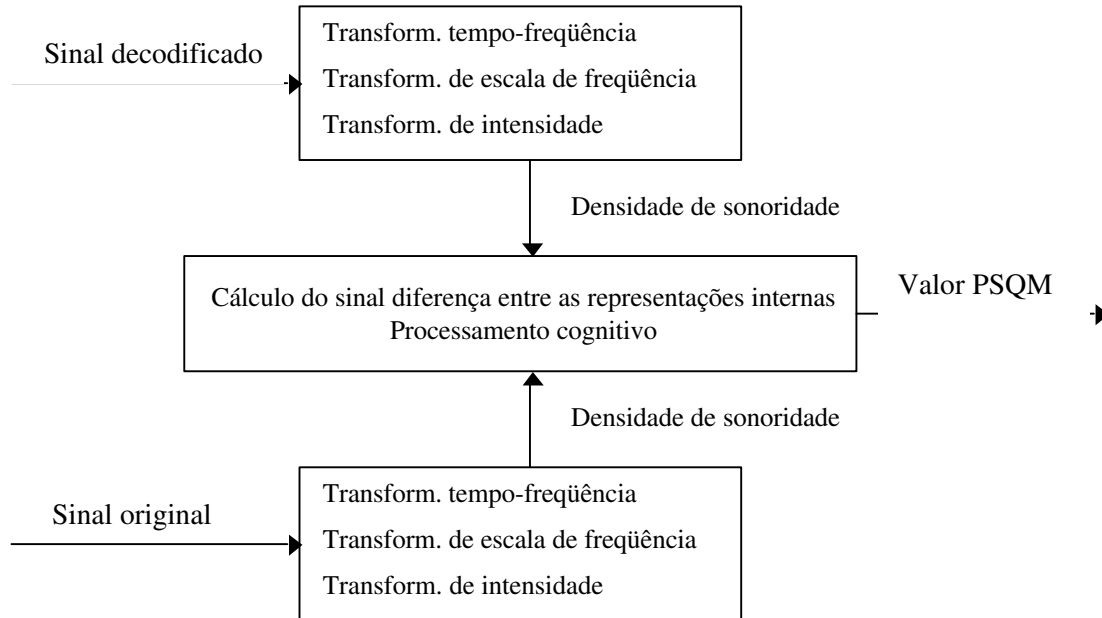


Figura 4.2 - Esquema básico do método PSQM.

Além desta modelagem perceptual, o método PSQM também utiliza um *processamento cognitivo*, o qual consiste de todos os processamentos que são realizados após o cálculo do sinal de diferença entre as representações internas do sinal original e do sinal do codec. A qualidade avaliada pelo PSQM é calculada através deste sinal de diferença, denominado “noise disturbance” ou, como será aqui adotado, *sinal diferença*. O nível médio deste sinal, calculado ao longo do tempo através das suas componentes de frequência, é relacionado diretamente com a qualidade da voz fornecida pelo codec.

Os detalhes dos processamentos envolvidos no método PSQM são descritos a seguir, juntamente com os detalhes de sua utilização, conforme as especificações da Rec. P.861 [1].

4.2.2. Inicialização

Os arquivos de voz utilizados no cálculo de qualquer medida objetiva devem, *a priori*, ser discretizados (com 16 bits por amostra), já que todos os processamentos subsequentes serão realizados no domínio digital. Para a faixa de telefonia, as frequências de amostragem mais utilizadas são as de 8 kHz e 16 kHz.

Antes que se possa realizar o cálculo da medida objetiva PSQM, é necessário realizar as seguintes operações de inicialização:

- ajuste do nível do sinal original para -26 dBov (ver Apêndice C) [6],
- alinhamento temporal,
- escalonamento global: compensação do ganho do codec,
- calibração global: ajuste de audibilidade.

A detecção do início e do final do sinal de voz é realizada via procedimentos padronizados pela Rec. P.861. Considera-se como amostra inicial a primeira cuja magnitude (valor absoluto), somada às magnitudes das quatro amostras anteriores, assume um valor maior ou igual a 200. Da mesma forma, a amostra final é a última cuja magnitude somada às magnitudes das quatro amostras subseqüentes assume valor maior ou igual a 200. A seguir, as amostras do início e do final dos arquivos de voz são descartadas.

O processamento PSQM só pode ser aplicado a sinais alinhados temporalmente. Se o atraso teórico introduzido pelo sistema sob teste é conhecido, este é adotado para o alinhamento. Se esse atraso não é conhecido, é necessário estimá-lo. Nesse caso, a Rec. P.861 recomenda que o alinhamento seja feito através do atraso relativo que proporciona a maior correlação cruzada entre o sinal original e o sinal decodificado. Neste trabalho, este cálculo foi feito utilizando-se um algoritmo de transformada rápida de Fourier (FFT), como será descrito no próximo capítulo (equação 5.2).

O escalonamento global para a compensação do ganho do sistema visa ajustar a energia do sinal decodificado $y[m]$, de modo a igualá-la à do sinal original, $x[m]$. É realizado multiplicando-se $y[m]$ por um fator E_g , definido pela relação entre as energias médias dos dois sinais segundo:

$$E_g = \left[\frac{\sum_{m=mi}^{mf} x^2[m]}{\sum_{m=mi}^{mf} y^2[m]} \right]^{1/2} \quad (4.1)$$

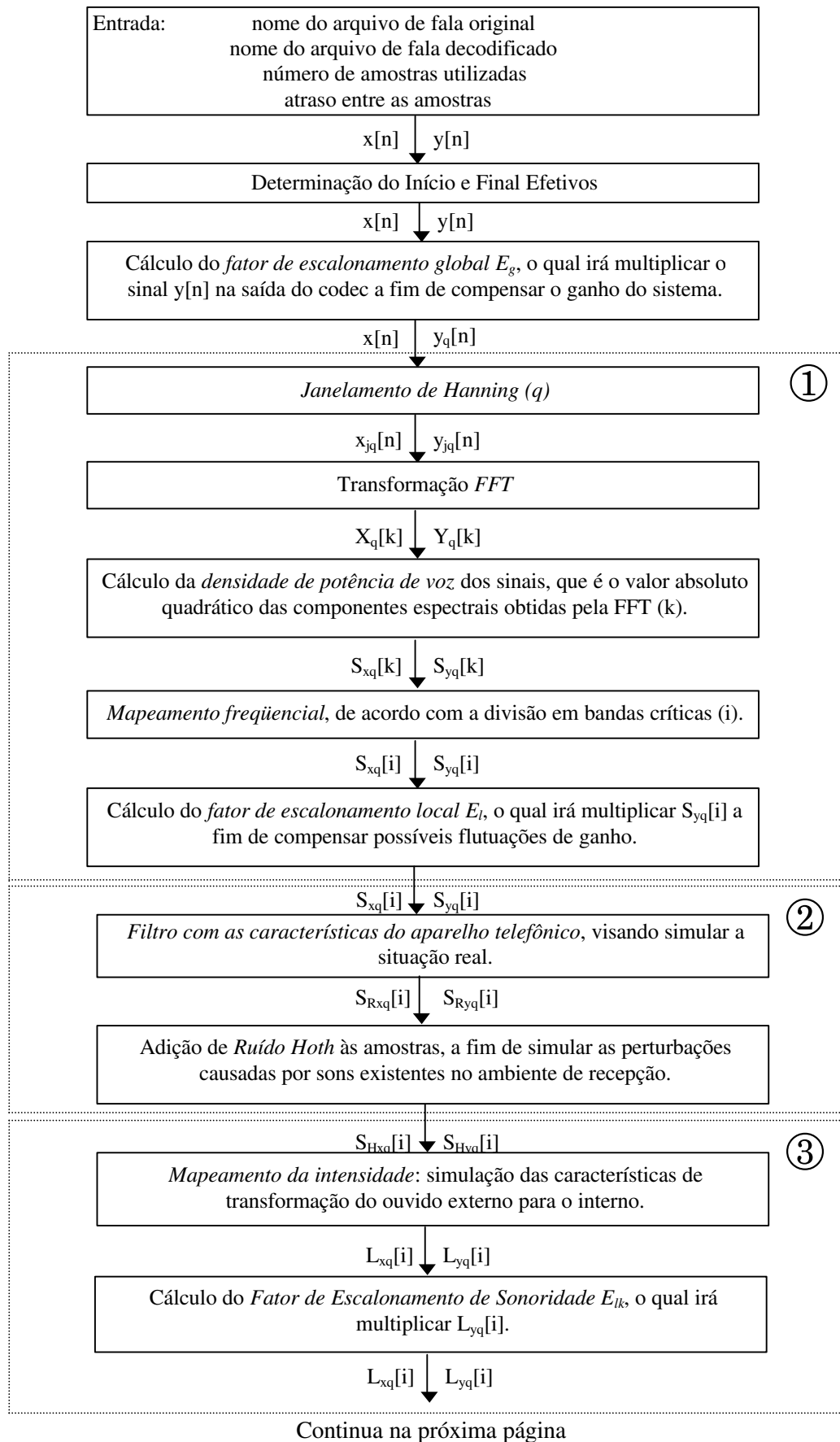
onde mi e mf representam, respectivamente, o início e o final dos sinais.

A calibração consiste em aplicar uma senóide de 1 kHz com amplitude adequada para se obter dois fatores de calibração, os quais são utilizados no ajuste de dois parâmetros do algoritmo PSQM, conforme descrito ao longo desta seção. Este procedimento é realizado apenas uma vez, e os valores obtidos para tais fatores são então usados em todas as medidas a serem realizadas posteriormente.

4.2.3. Cálculo da Medida PSQM

Os demais processamentos realizados no algoritmo PSQM são descritos a seguir, tendo como base o diagrama de blocos mostrado na figura 4.3. Os quadros tracejados indicam as operações básicas, que são:

- 1 - Transformação em frequência e mapeamento em bandas críticas.
- 2 - Simulação das características do aparelho telefônico e ruído ambiente.
- 3 - Transformação de intensidade.
- 4 - Processamento cognitivo.



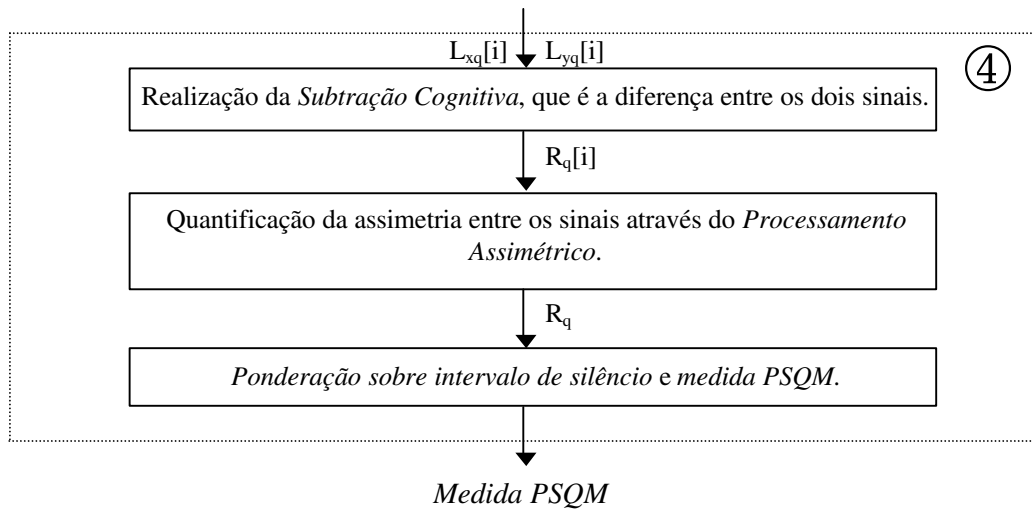


Figura 4.3 - Processamento PSQM

* *Mapeamento tempo-freqüência*: é realizado através da transformada de Fourier, implementada via FFT, aplicada a trechos de voz definidos através de uma janela de Hanning [20]. Os quadros resultantes são compostos por 256 amostras no caso de amostragem a 8.000 amostras/s e por 512 amostras para 16.000 amostras/s, com uma superposição de 50% entre dois quadros adjacentes. Assim, a duração de cada quadro é fixa, com valor de 32 ms. A FFT deve produzir o mesmo número de amostras no domínio da freqüência.

As densidades espectrais de energia de voz (DEE) são denotadas por $S_{xq}[k]$ e $S_{yq}[k]$, onde o subscrito q indica os quadros janelados. Estas amostras são iguais ao valor absoluto quadrático das componentes espectrais fornecidas pela FFT.

* *Transformação em freqüência*: a DEE na escala em Hz é transformada em outra escala mais apropriada para a representação interna, denominada *densidade de energia de pitch*. Tal transformação é realizada quadro a quadro, e a nova escala é baseada na definição das *bandas críticas*. Este mapeamento é ligeiramente diferente daquele descrito no Capítulo 3, já que aqui cada banda crítica é dividida em intervalos correspondendo a $\Delta c = 0,312$ de sua faixa total. Este valor de Δc deriva da definição das “terças-oitavas” (ver Apêndice C). A divisão das bandas críticas origina um total de 56 intervalos ou sub-bandas, denotadas pelo índice i na figura 4.3.

A Rec. ITU P.861 fornece uma tabela completa com os valores dos intervalos de freqüência correspondentes a cada sub-banda, bem como as amostras da FFT correspondentes, a característica de recepção do telefone, o limiar de audibilidade e o valor do Ruído Hoth associados a esses intervalos. A primeira faixa (0 a 15,6 Hz) é descartada no processo, e a freqüência máxima prevista é de 4.193 Hz.

O mapeamento produz uma amostra da densidade espectral de energia a cada sub-banda das bandas críticas. Cada amostra mapeada é denominada $S_{xq}[i]$, para o sinal original, e $S_{yq}[i]$, para o sinal decodificado, as quais são dadas por:

$$S_{xq}[i] = E_p \cdot \frac{\Delta b_i}{\Delta c} \cdot \frac{1}{I_u[i] - I_p[i] + 1} \cdot \sum_{I_p[i]}^{I_u[i]} S_{xq}[k] \quad (4.2)$$

$$S_{yq}[i] = E_p \cdot \frac{\Delta b_i}{\Delta c} \cdot \frac{1}{I_u[i] - I_p[i] + 1} \cdot \sum_{I_p[i]}^{I_u[i]} S_{yq}[k] \quad (4.3)$$

onde $I_p[i]$ e $I_u[i]$ são, respectivamente, os índices da primeira e da última amostra, em Hz, utilizadas para obter a densidade de energia de *pitch* da sub-banda i . Δb_i e Δc são, respectivamente, a largura da sub-banda i , em Hz, e a largura da sub-banda em relação à largura total das bandas críticas (0,312). E_p é o fator de calibração de energia de *pitch*, obtido segundo as recomendações de inicialização ($E_p = 6,4661 \times 10^{-6}$).

Na maioria dos casos, a densidade de energia de *pitch* de cada sub-banda é o resultado do mapeamento de uma única amostra da FFT. Porém, à medida que se avança na escala em Hz, mais amostras são tomadas, alocando um máximo de 7 amostras nas frequências superiores (4.193 Hz), o que era de se esperar, já que, como foi visto no Capítulo 3, a largura de faixa das bandas críticas vai aumentando à medida que se aumenta a frequência.

A última tarefa realizada nesta etapa do processamento é um *escalonamento local*, o qual visa ajustar os níveis de $S_{yq}[i]$ para compensar pequenas flutuações de ganho. Somente as componentes tempo-frequência audíveis são levadas em conta, isto é, apenas aquelas que estão acima do limiar de audibilidade absoluto $S_0[i]$, definido pela Rec. P.861 para cada sub-banda. Inicialmente, são calculadas as energias totais de *pitch* ao longo das sub-bandas dos sinais para o q -ésimo quadro, e um fator de escala E_q é definido da seguinte forma:

$$S_{xq} = \sum_{i=1}^B S_{xq}[i] \quad (4.4)$$

$$S_{yq} = \sum_{i=1}^B S_{yq}[i] \quad (4.5)$$

$$E_q = \frac{S_{xq}}{S_{yq}} \quad (4.6)$$

onde B é o número total de sub-bandas.

Se ambos os valores de S_{xq} e de S_{yq} estiverem acima de 40 dB_{SPL} (ver definição de dB_{SPL} no Apêndice C), os valores de $S_{yq}[i]$ são ajustados segundo:

$$\hat{S}_{yq}[i] = E_q \cdot S_{yq}[i] \quad (4.7)$$

Se os valores de S_{xq} ou de S_{yq} estiverem abaixo de 40 dB_{SPL}, $S_{yq}[i]$ é multiplicado pelo fator de escala E_{av} , definido como a média de todos E_q até então calculados. Para o caso particular em que $S_{xq} = 0$ e $S_{yq} = 0$, o quadro q deve ser ignorado no cálculo de E_{av} .

* *Filtragem na Faixa de Telefonía*: trata-se da filtragem dos valores de $S_{xq}[i]$ e de $S_{yq}[i]$ através das características $T[i]$ de recepção do aparelho telefônico, definidas pela Rec. P. 830 e reproduzidas na Rec. P.861. Esta filtragem é realizada segundo:

$$S_{Rxq}[i] = T[i] \cdot S_{xq}[i] \quad (4.8)$$

$$S_{Ryq}[i] = T[i] \cdot \hat{S}_{yq}[i] \quad (4.9)$$

* *Ruído Hoth*: às amostras $S_{Rxq}[i]$ e $S_{Ryq}[i]$ resultantes da filtragem é adicionado o ruído (Hoth), representado por $H[i]$, o qual simula as perturbações produzidas por sons existentes no ambiente de recepção. Esta operação é realizada segundo:

$$S_{Hxq}[i] = H[i] + S_{Rxq}[i] \quad (4.10)$$

$$S_{Hyq}[i] = H[i] + S_{Ryq}[i] \quad (4.11)$$

A Rec. P.861 apresenta o valor do ruído a ser adicionado à amostra de cada sub-banda.

* *Transformação das intensidades*: tem como objetivo simular as características de transformação do ouvido externo para o interno e geração da sonoridade subjetiva. A escala de intensidade é mapeada em uma escala de *sonoridade subjetiva (loudness)*, dando origem a uma *função densidade de sonoridade comprimida*. As densidades de sonoridade comprimidas, $L_{xq}[i]$ e $L_{yq}[i]$, são calculadas através da função de compressão definida pelas expressões:

$$L_{xq}[i] = E_q \cdot \left(\frac{S_0[i]}{0.5} \right)^\gamma \cdot \left[\left(0.5 + 0.5 \cdot \frac{S_{Hxq}[i]}{S_0[i]} \right)^\gamma - 1 \right] \quad (4.12)$$

$$L_{yq}[i] = E_q \cdot \left(\frac{S_0[i]}{0.5} \right)^\gamma \cdot \left[\left(0.5 + 0.5 \cdot \frac{S_{Hyq}[i]}{S_0[i]} \right)^\gamma - 1 \right] \quad (4.13)$$

onde $S_0[i]$ é o limiar absoluto de audibilidade interno e $\gamma = 0,001$, otimizado em experimentos práticos. E_q é o fator de escala dado pela equação 4.6. Os valores negativos de $L_{xq}[i]$ e $L_{yq}[i]$ são transformados em zero.

O método utilizado para obtenção desta equação foi proposto por Zwicker [9,40] e envolve a obtenção de gráficos dos níveis de pressão de cada banda crítica e cálculo da respectiva área sob o gráfico, a fim de obter a sonoridade. A equação acima é uma aproximação deste método. Informações adicionais podem ser encontradas em [40,45,46].

É importante observar que este não é o único método de cálculo da densidade de sonoridade. Outros métodos são largamente utilizados em outras aplicações, como é o caso do “método A”, desenvolvido por S. S. Stevens [47], bem como aquele descrito pelas equações 3.1 a 3.7. A escolha do método desenvolvido por Zwicker foi feita após testes comparativos, os quais indicaram ser este o mais adequado para a aplicação desejada [2,3].

* *Escala de Sonoridade*: é a última etapa do processamento que é realizado sobre os sinais individualizados. Consiste em um novo escalonamento, agora aplicado às amostras $L_{yq}[i]$ segundo:

$$L'_{yq}[i] = E_{lq} \cdot L_{yq}[i] \quad (4.14)$$

onde o fator de escala E_{lq} é calculado de acordo com:

$$E_{lq} = \frac{L_{xq}}{L_{yq}} \quad (4.15)$$

e

$$L_{xq} = \sum_{i=1}^B L_{xq}[i] \quad (4.16)$$

$$L_{yq} = \sum_{i=1}^B L_{yq}[i] \quad (4.17)$$

Quando L_{xq} ou L_{yq} forem inferiores a 0,02, E_{lq} é feito igual a 1.

* *Modelo Cognitivo*: as operações cognitivas são aquelas realizadas sobre o sinal correspondente à diferença entre as densidades $L_{xq}[i]$ e $L_{yq}[i]$. São compostas pelo cálculo do sinal diferença e subtração do ruído cognitivo interno, pelo cálculo dos fatores de assimetria e pela ponderação dos intervalos de silêncio.

O sinal diferença é definido como a diferença absoluta entre as amostras $L_{xq}[i]$ e $L_{yq}[i]$, representando a diferença entre as sensações provocadas pelos dois sinais. Este sinal será o elemento chave para o cálculo da medida PSQM. Deve-se observar que, se os dois sinais forem idênticos, esta diferença será nula e a qualidade será máxima.

O sinal diferença é calculado como:

$$R_q[i] = |L'_{yq}[i] - L_{xq}[i]| - 0,01 \quad (4.18)$$

onde o fator 0,01 representa o ruído cognitivo interno. Se $R_q[i]$ torna-se negativo, então ele é feito igual a zero.

Para o cálculo da energia do sinal diferença referente a um dado quadro q , são levadas em conta possíveis assimetrias entre os sinais, resultantes, por exemplo, da inserção de componentes estranhas ao sinal ou da supressão de componentes pelo codec.

A assimetria é definida como a diferença de sensação de degradação de qualidade entre os casos em que o codec introduz componentes do sinal (onde um maior impacto é produzido) e aqueles em que este reduz as componentes (menor impacto). Então, a fim de se levar em consideração esta característica, um fator de assimetria $F_q[i]$ é calculado para cada sub-banda, o qual é obtido a partir de $S_{Hxq}[i]$ e $S_{Hyq}[i]$, resultantes do processamento PSQM após a filtragem e a adição de Ruído Hoth, segundo:

$$F_q[i] = \left(\frac{S_{Hyq}[i] + 1}{S_{Hxq}[i] + 1} \right)^{0.2} \quad (4.19)$$

Quando $S_{Hxq}[i]$ e $S_{Hyq}[i]$ são, simultaneamente, menores que $100 \cdot S_0[i]$, $F_q[i]$ é feito igual a 1. Adicionalmente, seu valor máximo é limitado a 2,0.

Assim, os fatores de assimetria são usados para ponderar a energia do sinal diferença de cada quadro q , R_q , que é calculada segundo a expressão:

$$R_q = \sum_{j=1}^B R_q[i] \cdot F_q[i] \cdot \Delta c \quad (4.20)$$

Portanto, cada quadro de voz processado pelo PSQM produz um valor R_q . O valor médio de R_q ao longo dos quadros, levando em conta, de forma explícita, a presença de quadros inativos (silêncio), será a medida PSQM de qualidade.

Quadros de silêncio são definidos como aqueles onde a energia S_{xq} , definida pela equação 4.4, se encontra abaixo de 70 dB_{SPL}. Se a calibração inicial foi realizada, o limiar de silêncio é $S_{xq} = 1,0 \cdot 10^7$. Quadros com S_{xq} menor que este valor são considerados inativos (silêncio).

Após a identificação dos quadros ativos e de silêncio, são calculados dois valores médios, V_{va} e V_{is} , segundo:

$$V_{va} = \frac{1}{N_{va}} \sum_{qva} R_q \quad (4.21)$$

$$V_{is} = \frac{1}{N_{is}} \sum_{q,s} R_q \quad (4.22)$$

onde N_{va} e N_{is} representam, respectivamente, o número de quadros com voz ativa e o número de quadros com silêncio; qva e qs representam, respectivamente, os quadros com voz ativa e os quadros de silêncio.

O valor médio final, V_{psqm} , levando em conta voz ativa e silêncio, depende de N_{va} e N_{is} . Se não houver quadros de silêncio, então $V_{psqm} = V_{va}$. Se o material de voz original contém intervalos de silêncio, a influência destes dependerá da sua duração total. Assim, V_{psqm} é calculado segundo:

$$V_{psqm} = \frac{PV \cdot f_a}{PV \cdot f_a + f_s} \cdot V_{va} + \frac{f_s}{PV \cdot f_a + f_s} \cdot V_{is} \quad (4.23)$$

onde f_s é a fração de quadros de silêncio, f_a é a fração de quadros de voz ($f_s + f_a = 1$) e

$$PV = \frac{1 - PS}{PS} \quad (4.24)$$

onde PS é um fator de ponderação que depende das características do material de voz utilizado no teste. A Rec. P.861 recomenda o valor $PS = 0,2$ para voz com 50% de intervalos de silêncio.

Pode-se observar que se não há quadros de silêncio, então $f_s = 0$, $f_a = 1$ e $V_{psqm} = V_{va}$. Também, se $PS = 1$, então $PV = 0$ e $V_{psqm} = V_{is}$, quaisquer que sejam os valores de f_s e f_a . Por fim, PS não pode ser feito nulo.

A avaliação do desempenho do codec é feita através de V_{psqm} , denominado valor PSQM. Este valor poderá ser utilizado como grau de qualidade, definindo uma escala objetiva com os limites iguais a zero e 6,5 (este limite máximo, sugerido pela Rec. P.861, não é automático, ficando seu controle a cargo do usuário). O valor zero indica a melhor qualidade possível – quando o sinal na saída do codec é igual ao sinal original, exceto pelos limiares e precisões envolvidas. Por outro lado, o valor PSQM também poderá ser utilizado para prever o resultado de testes ACR, DCR ou CCR utilizando testes de audição (LOT), usando-se uma função de mapeamento, como descrito a seguir.

4.2.4. Transformação da escala objetiva para escala subjetiva

O valor PSQM é uma estimativa da medida subjetiva de qualidade numa escala de degradação. Assim, ele não precisa ser transformado quando esta escala é suficiente, como, por exemplo, na comparação ou otimização de codecs. A transformação é necessária apenas quando se deseja expressar a avaliação em escalas subjetivas padronizadas, como o MOS, CMOS ou o Q-Equivalente, definidas pelas Rec. P.800 [5] e P.810 [48].

* *Mean Opinion Score*: os valores da escala MOS dependem da língua e laboratório em que os arquivos de voz foram gravados. Assim, não existe uma única função de transformação entre PSQM e MOS. Ao contrário, curvas devem ser definidas para cada língua, laboratório e, em geral, para o tipo de teste pretendido. E, portanto, dado que os valores MOS estimados via PSQM através de uma dada curva de mapeamento dependem deste contexto, toda vez que os valores MOS obtidos forem apresentados, a curva de transferência utilizada deve ser também apresentada. Tais observações valem também para

a transformação para a escala CMOS. Algumas destas curvas são apresentadas no Capítulo 6.

* *Valores Q-Equivalentes*: a Rec. P.810 define um valor de medida subjetiva, denominada Q-Equivalente. O valor Q é a quantidade de unidades de referência de ruído modulado (MNRU – ver apêndice C) que, ao ser adicionada ao sinal original, produz um valor MOS equivalente àquele produzido pelo codec.

O valor Q pode ser obtido diretamente dos valores PSQM sem a necessidade de transferência para a escala MOS. Neste caso, a curva de transferência PSQM x Q deve ser apresentada.

4.2.5. Considerações Finais

A medida PSQM representa uma estimativa da qualidade média que seria medida via teste subjetivo. Ela não permite que se faça análise da variação de qualidade em função dos ouvintes, uma vez que o valor PSQM representa uma estimativa do valor médio dos votos atribuídos. A estimativa da distribuição dos votos está em estudo [1]. Assim, quando se torna necessária a análise em relação aos ouvintes, será inevitável a realização de experimentos subjetivos.

As diferenças inaudíveis entre o sinal original e aquele resultante do codec, dentro das precisões de representação e julgamento do método PSQM, não influem na pontuação resultante da medida objetiva. Em particular, se os sinais são idênticos, o PSQM estimará uma qualidade perfeita, independente da qualidade dos sinais de entrada.

O *maskamento* só é considerado quando duas componentes coincidem na frequência (mesma sub-banda) e no tempo, e este maskamento é realizado de forma automática, na medida em que o método utiliza a diferença entre os sinais de entrada e de saída.

A questão da aplicabilidade do método PSQM está ligada à existência de diversas condições de teste a serem investigadas na avaliação do desempenho de um codec. Essas condições de teste, descritas na Rec. P.830 [6], englobam os fatores que influenciam a qualidade de um codec, as diferentes tecnologias de codificação existentes e as aplicações desejadas para o codec sob teste. O método PSQM não é capaz de oferecer boas estimativas para todas essas condições. Assim, a Rec. P.861 oferece um guia para a determinação dos fatores de teste, tecnologia e aplicações de codecs que podem ser avaliados seguramente via testes objetivos com o método PSQM. As condições de teste onde pode-se utilizar eficientemente o PSQM são:

- Fatores de teste:
 - Níveis de entrada de um codec de voz;
 - Dependência de locutores;
 - Identificação da taxa de bits (no caso de o codec operar com mais de uma taxa);
 - Transcodificações.

Existem alguns outros fatores de teste para os quais o uso de medidas objetivas requer mais estudos até que se possa garantir uma confiabilidade aceitável, como é o caso dos níveis de audibilidade em experimentos subjetivos, locutores múltiplos simultâneos, erros de canal de transmissão, incompatibilidade de taxa de bits entre o codificador e o decodificador (se o codec opera com mais de uma taxa de bits), música como entrada para um codec e supressão de trechos de voz.

- Tecnologias de codificação:

- Forma de onda;
- CELP e híbridos com taxas acima de 4 kbits/s.

Para o caso de CELPs e híbridos com taxas menores que 4 kbits/s, *vocoders* e outros tipos de codecs, mais estudos são necessários para garantir a confiabilidade das medidas objetivas.

- Aplicações

- Otimização de codecs;
- Avaliação de codecs.

Para o caso de seleção de codecs, mais estudos são necessários. Já para o planejamento de redes telefônicas e teste de rede ao vivo, as medidas objetivas podem, sob certas condições, ser utilizadas [1].

Pode-se observar então que a utilização de medidas objetivas ainda é restrita a poucos casos, e a extensão de seu uso para outras aplicações é objeto de pesquisa em nível internacional.

4.3. O MÉTODO PSQM+

A versão original do PSQM foi, dentre os métodos apresentados durante o período de 1993 a 1996, a que obteve a mais alta correlação entre as medidas subjetivas e objetivas. No entanto, o PSQM não se mostrou eficaz nos casos em que os sinais de voz continham supressão de trechos de voz e/ou distorções provocadas por excesso de volume. Fez-se então necessário um aperfeiçoamento, tendo como resultado o PSQM+ [49].

No PSQM, os efeitos da assimetria são tratados através de um fator $F_q[i]$ (equação 4.19) [50]. Se o sinal codificado contém pequenas distorções, a atuação do fator F_q melhora significativamente a correlação entre os resultados subjetivos e objetivos. Mas, se as distorções são muito grandes, essa correlação diminui. Para distorções por supressão de trechos de voz, as componentes do espectro de energia do sinal degradado $S_{Hyq}[i]$ são muito pequenas, levando a um sinal diferença R_q também muito pequeno, o qual contribui para uma redução do valor médio V_{psqm} na equação 4.24 e, portanto, da medida PSQM.

Por outro lado, quando uma grande distorção por excesso de volume ocorre, $S_{Hyq}[i]$ cresce exageradamente, produzindo o mesmo efeito em R_q e na medida PSQM. Estes problemas podem ser resolvidos de muitos modos, porém um método simples e eficaz consiste em introduzir um fator de ponderação adicional [49]:

$$E_q^+ = \min \left(\frac{S_{xq} + 1}{S_{yq} + 1}, 10000 \right) \quad (4.26)$$

Este fator de ponderação é utilizado no cálculo de R_q , em cada quadro q :

$$R_q = (E_q^+)^{0.25} \sum_{i=1}^B R_q[i] \cdot F_q[i] \cdot \Delta c \quad (4.27)$$

Quando as energias totais do sinal degradado e do sinal de referência são aproximadamente as mesmas, o segundo fator de escala, E_q^+ , é aproximadamente 1,0 e o valor PSQM resultante é semelhante àquele que seria obtido na versão antiga. Porém,

quando uma distorção por excesso de volume ou uma supressão de trechos de voz é introduzida, $F_q[i]$ e E_q^+ têm efeitos opostos no sinal diferença total R_q em cada quadro q , melhorando a correlação com as medidas subjetivas correspondentes.

4.4. OUTROS MÉTODOS PROPOSTOS

Vários métodos foram propostos como alternativa ao PSQM nas situações em que este não mostrou um bom desempenho. Porém, comparando seus desempenhos globais, concluiu-se que nenhum deles apresentava resultados superiores ao PSQM. A seguir, serão descritos, resumidamente, alguns destes métodos.

4.4.1. Método MNB (*Measuring Normalizing Blocks*)

Este método foi proposto como um anexo à Rec. P.861 [42,51].

A figura 4.4 mostra a estrutura geral do método MNB [52].

Os estudos realizados para este método levaram a uma inversão da ênfase tradicional, resultando em um modelo mais simples para o ouvido e um modelo mais sofisticado para o julgamento. Após o estudo de diversos parâmetros, como as funções de transferência dos ouvidos externo e médio, limiares absolutos de audibilidade, curvas de sonoridade e efeitos de mascaramento, os autores da proposta chegaram à conclusão de que estes não eram significativos para o sucesso da estimativa da qualidade percebida para a faixa de telefonia. Por esse motivo, o modelo do ouvido adotado contém apenas um mapeamento da escala em Hertz para uma escala em Bark e uma transformação logarítmica da potência para aproximar a sonoridade percebida [53].

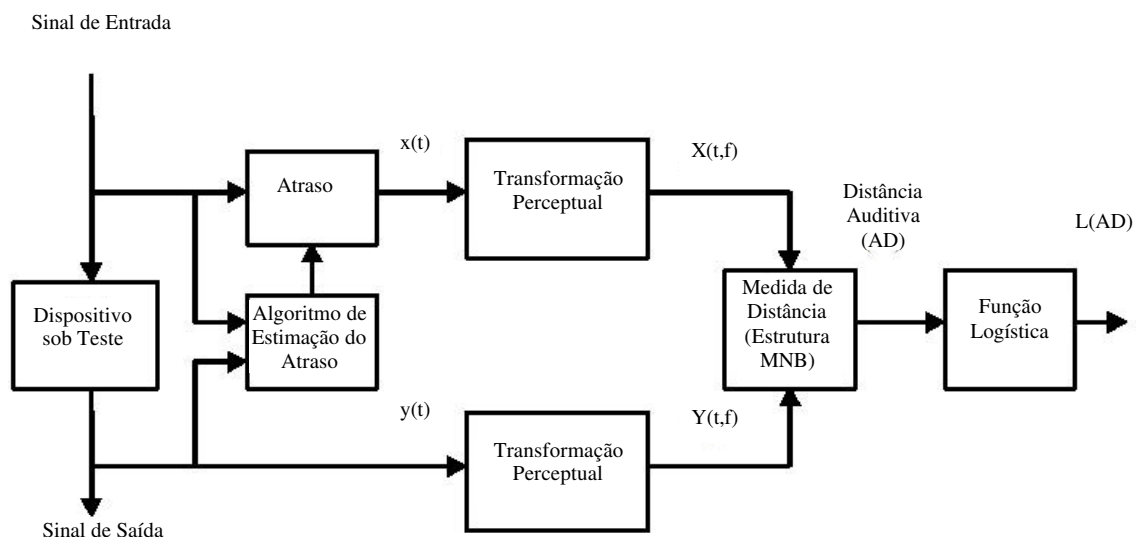


Figura 4.4 - Estrutura geral do método MNB

Medidas de distância procuram simular o julgamento auditivo pela comparação de dois sinais perceptualmente transformados (por exemplo, entrada do codificador e saída do decodificador). Entretanto, muitas medidas de distância convencionais apresentam propriedades que são claramente inconsistentes com o julgamento auditivo. O que se

observa é que os ouvintes se adaptam e reagem diferentemente a desvios espectrais em instantes e frequências distintas, o que sugere a medida de distância em diferentes escalas de tempo e frequência. Então, o problema da estimação da qualidade de voz pode ser abordado a partir de uma família de análises em múltiplas escalas de tempo e frequência, onde os desvios espectrais em uma escala são medidos e removidos para que não sejam novamente levados em consideração como parte dos desvios de outras escalas. Essas escalas podem ser definidas para qualquer faixa (em Bark). Os autores desta medida determinaram, através de uma série de experimentos, que a melhor divisão é aquela mostrada na tabela 4.1.

Tabela 4.1 - Divisão em Diferentes Escalas

Número da Escala	Faixa da Escala (Bark)	Largura de Faixa (Bark)
1	2,5 - 7,5	5
2	7,5 - 12,5	5
3	2,5 - 5	2,5
4	10 - 12,5	2,5
5	12,5 - 15	2,5
6	0,6 - 2,5	1,9

A fim de simular os padrões de adaptação e reação dos ouvintes com mais confiabilidade, trabalha-se das maiores para as menores escalas (ou seja, trabalha-se primeiro com as escalas referentes a uma maior largura de faixa em Bark, como se pode ver na tabela 4.1). À luz dessas observações, escolheu-se formar uma medida de distância na forma hierárquica de Blocos de Medida Normalizados (MNB).

Cada um desses blocos toma os sinais de referência e teste perceptualmente transformados, $X(t,f)$ e $Y(t,f)$, como entradas, e retorna um conjunto de medidas e uma versão normalizada de $Y(t,f)$. Um MNB temporal (TMNB) mede as diferenças entre as integrais de $X(t,f)$ e $Y(t,f)$ conforme a equação 4.28, e modifica $Y(t,f)$ segundo a equação 4.29. As porções positivas e negativas das medidas de desvio são então integradas no tempo, conforme as equações 4.30 e 4.31, gerando as medidas m_{2i-1} e m_{2i} .

$$e(t, fl) = \frac{1}{fu - fl} \int_{fl}^{fu} Y(t, f) df - \int_{fl}^{fu} X(t, f) df \quad (4.28)$$

$$\tilde{Y}(t, f) = Y(t, f) - e(t, fl) \quad (4.29)$$

$$m_{2i-1} = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \max[e(t, fl), 0] dt, \quad i = 1 \text{ a } N \quad (4.30)$$

$$m_{2i} = \frac{-1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \min[e(t, fl), 0] dt, \quad i = 1 \text{ a } N \quad (4.31)$$

onde fl e fu definem o intervalo de frequência considerado e ti define o instante de tempo.

A definição e equações para o MNB frequencial (FMNB) são análogas, com as variáveis de tempo e frequência trocadas.

Os MNB são encadeados em hierarquias, e a medida do desvio para uma dada escala de tempo ou frequência é calculada apenas uma vez. Duas estruturas, MNB-1 e MNB-2, foram escolhidas para ter uma complexidade relativamente pequena, aliada a um bom desempenho na estimação da qualidade de voz percebida [51]. Estas estruturas geram, respectivamente, 13 e 12 medidas. Tais medidas são então linearmente combinadas, a fim de fornecer o valor final para a qualidade de voz (AD), através da equação 4.32:

$$AD = \sum_{i=1}^n m_i \cdot p_i \quad (4.32)$$

onde m_i são os valores obtidos para as medidas, n é o número de medidas e p_i se refere aos pesos adotados para cada medida, os quais podem ser encontrados em [42].

O desempenho do método MNB mostrou-se ligeiramente superior ao PSQM nos casos em que houve presença de erros nos arquivos de voz [54,55]. Já nos casos em que havia ruído de ambiente, o PSQM mostrou o melhor desempenho. Mais importante, notou-se que o método MNB mostrou excelentes resultados quando os arquivos de voz estavam gravados em inglês norte-americano. Porém, o método PSQM mostrou melhor consistência no caso de outros idiomas. Portanto, apesar de ambos os métodos, na média, apresentarem resultados parecidos [43], pode-se afirmar que o método MNB tem aplicação mais restrita que o método PSQM.

4.4.2. Método PAMS (*Perceptual Analyses Measurement System*)

As informações a respeito deste método são restritas, devido ao interesse de seus criadores em proteger seus direitos autorais. As informações a seguir são uma compilação de dados coletados em diversas fontes, que fornecem uma visão geral dos princípios utilizados. A figura 4.10 mostra a estrutura geral do método PAMS.

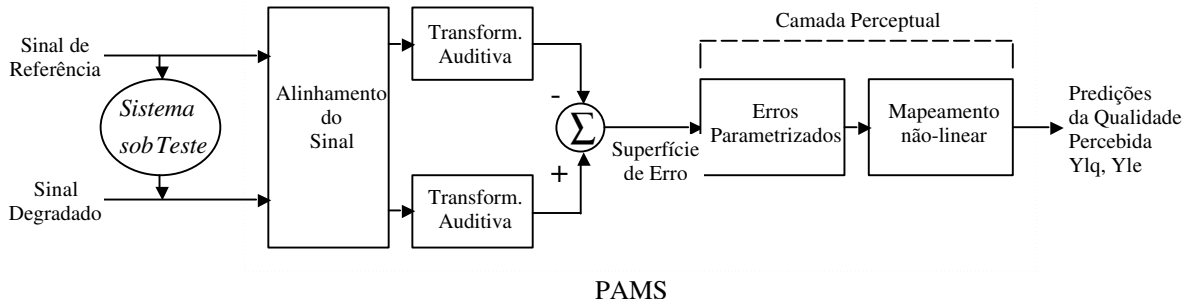


Figura 4.10 - Estrutura do Método PAMS

Neste bem sucedido método, desenvolvido pela *British Telecom* (BT), os sinais de referência e de teste são previamente equalizados e alinhados, e então transformados para levar em conta as propriedades do sistema auditivo humano, como mascaramento e sonoridade. Em seguida, é feita uma representação tridimensional de cada um dos sinais (figura 4.11), e sua diferença dá origem a uma “superfície de erro”, que é uma medida dos erros audíveis introduzidos na codificação.

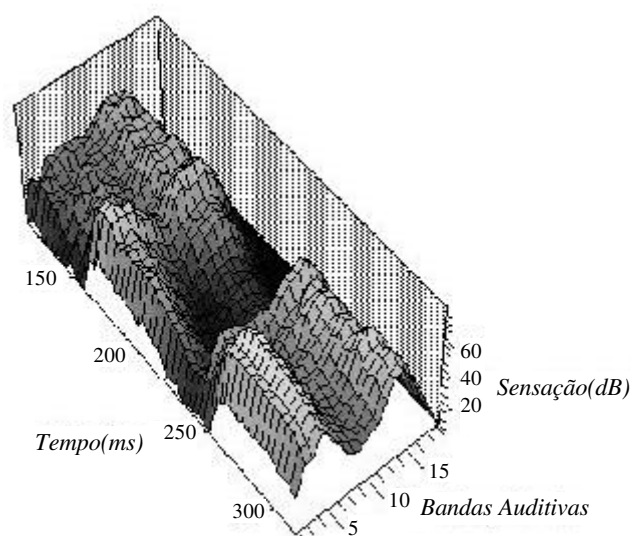


Figura 4.11 - Superfície de Representação do Sinal

Após a aplicação de alguns processamentos, realiza-se um mapeamento em duas escalas de qualidade:

- YLQ: prediz a escala “listening-quality”.
- YLE: prediz a escala “listening-effort”.

Portanto, o método PAMS retorna diretamente uma predição dos valores obtidos em medidas subjetivas, em contraste com o método PSQM, onde o mapeamento para estas escalas é feito posteriormente.

Além desta, são encontradas outras diferenças importantes deste método em relação às demais medidas objetivas:

- *Atraso variável*: este foi o primeiro modelo objetivo capaz de levar em conta atrasos variáveis no tempo. Tal característica é encontrada, em maior ou menor grau, em todos os sinais de voz, e em particular em voz sob IP (*Internet Protocol*) e outras transmissões baseadas em pacotes. A segunda versão do programa, lançada em dezembro de 1998, incorporou a habilidade de identificar o tipo mais comum de variação no atraso – em períodos de silêncio. Por fim, a terceira versão do programa adicionou a capacidade de identificar mudanças de atraso durante os trechos ativos. Um perfil completo do atraso e pontos de mudança de atraso é retornado pelo PAMS.

- *Filtragem em interfaces analógicas*: o método PAMS foi projetado para uso em redes reais. Interfaces analógicas híbridas, entre outras, introduzem uma filtragem que pode fazer com que os modelos objetivos para predição de qualidade se tornem pouco confiáveis. Para lidar com esta situação, este método, em sua terceira versão, apresenta a habilidade de identificar toda a gama de tipos de filtros, e o processo foi totalmente automatizado. A função de transferência do sistema é então apresentada ao usuário para análise e diagnóstico.

É importante destacar que a descrição deste método foi baseada na coleta de informações em uma série de fontes isoladas e não referenciáveis, por se tratarem de textos não formais. Por esse motivo, torna-se impossível fazer uma descrição mais rigorosa, o que explica a presença de alguns trechos apresentando um caráter relativamente vago.

4.4.3. Método TOSQA (*Telecommunication Objective Speech Quality Assessment*)

Este método foi desenvolvido pela *Deutsche Telekom Berkon* [44]. Ele adota a estrutura comum das medidas baseadas em modelos psicoacústicos, mas também considera certos comportamentos dos ouvintes em testes auditivos (reduzindo os efeitos de fatores que têm pequena influência na opinião de qualidade em testes do tipo LOT). Esta proposta modela fortemente os testes LOT para avaliação da qualidade de voz de sistemas de telecomunicações.

Para a avaliação da qualidade de voz de codecs ou sistemas de transmissão similares, geralmente são utilizados testes do tipo LOT em conjunto com um ACR. Neste ambiente de teste, os ouvintes não conhecem o material sem distorção nem o locutor. Eles comparam a qualidade da amostra de voz apresentada com uma referência interna (“gravada” no cérebro) de voz humana. Esta referência é adquirida pela experiência individual de cada um em serviços de telecomunicações.

Baseados nestas informações, os criadores deste método inseriram novos componentes, a saber:

- Uma estimativa da resposta total em frequência do sistema sob teste será usada como uma função de ponderação para o sinal de referência.
- A largura de faixa das bandas críticas não é estática.
- O valor básico para a estimação da qualidade não é a distância entre a referência e o arquivo codificado, mas a similaridade entre as representações.

Vários passos compõem uma primeira etapa de processamento. Primeiro, o atraso total será estimado através de uma função de correlação cruzada. O segundo passo computa os níveis de potência na faixa de telefonia de ambos os sinais e sua envoltória espectral média usando uma estimação da densidade espectral de potência. Uma função auxiliar similar à função coerência é também calculada para obter uma impressão da linearidade do sistema sob teste. Os resultados da primeira parte dos cálculos são usados como parâmetros de controle para a parte principal do programa: a estimação de valores descrevendo a qualidade de voz.

Juntamente com a equalização de nível e atraso para o sinal de entrada, as distorções espectrais são também eliminadas pelo uso da resposta em frequência do primeiro passo. Ambos os sinais, o pré-processado de entrada $x(k)$ e o de saída $y(k)$, são janelados em segmentos de 16 ms e transformados no domínio da frequência usando uma DFT.

Com a ajuda de um limiar de potência, os segmentos são classificados como ativos ou não-ativos. O método TOSQA utiliza o modelo de Zwicker (discutido no Capítulo 3) para cálculo da sonoridade específica. O mascaramento espectral e a compressão são calculados a partir desse modelo. A classificação em bandas críticas é realizada dinamicamente.

Para o cálculo da similaridade entre as características de entrada e saída, utiliza-se os espectros de sonoridade modificados em cada quadro. Neste passo, algumas diferenças entre os padrões de sonoridade que têm pouca influência no teste auditivo serão reduzidas. Utiliza-se também uma ponderação de assimetria. A similaridade computada é o resultado principal da medida de qualidade de voz do TOSQA. O valor médio dos resultados obtidos para cada quadro leva ao valor designado SQ. O valor interno SQ é então transformado em uma escala MOS, predita pelo uso de um mapeamento monotônico não-linear. Este programa é também capaz de calcular valores adicionais para descrever efeitos de alguns fatores de qualidade (por exemplo, ruído de fundo, resposta em frequência ou não-linearidades do sistema sob teste).

4.4.4. Conclusão

O PSQM foi, na média, o método que apresentou o melhor desempenho. Em algumas situações particulares, porém, os outros métodos apresentaram melhores resultados: no caso de sinais contendo distorções por supressão de trechos de voz, isoladamente ou em conjunto com distorções por volume, o método PSQM+ mostrou-se claramente superior ao método PSQM original [45]; o método MNB se mostrou mais eficiente para arquivos de voz gravados na língua inglesa falada nos EUA; por fim, o PAMS obteve resultados melhores para sinais de voz “limpos”, ou seja, livres de ruído e erros [43]. Já o método TOSQA mostrou-se ligeiramente inferior aos métodos PSQM e PAMS. A exceção fica por conta de algumas situações específicas, onde estruturas de distorção bastante complexas são inseridas no sinal a ser analisado [44].

Observou-se ainda que muitos dos métodos desenvolvidos visaram, principalmente, o acúmulo de conhecimento para o desenvolvimento de futuros modelos que corrijam e melhorem aqueles atualmente utilizados, os quais, reconhecidamente, ainda deixam muito a desejar para inúmeras aplicações.

CAPÍTULO 5

IMPLEMENTAÇÃO DE PROGRAMA DE COMPUTADOR BASEADO NO MÉTODO PSQM

O programa de computador original do método PSQM foi implementado na linguagem ANSI C. Este serviu de base para o desenvolvimento de um novo programa, implementado em Matlab[®] (versão 5.3), a fim de se ter maior flexibilidade e versatilidade para a realização de modificações e testes.

5.1. IMPLEMENTAÇÃO EM MATLAB[®] 5.3

Este programa foi implementado através de uma rotina principal, denominada *moqv*, e diversas rotinas auxiliares, as quais serão descritas ao longo deste capítulo. Esta implementação oferece diversas facilidades não encontradas no programa original do método PSQM: alinhamento temporal dos sinais de voz original e decodificado via correlação cruzada; opção de várias taxas de amostragem; detalhamento das características dos sinais de voz, como a evolução ao longo do tempo e audição; ajuste de parâmetros dos sinais, como energia média, valor de pico, etc; diversas opções de mapeamento para os valores subjetivos e processamentos adicionais visando a melhora da correlação entre as medidas objetivas e subjetivas correspondentes. Tais facilidades resultam numa maior interatividade, tornando seu uso adequado tanto para usuários com larga experiência na área, quanto para iniciantes. Este conjunto é aqui denominado de método MOQV, o qual possui duas versões (MOQV1 e 2).

5.1.1. Entrada de Dados

No início do programa, o usuário deverá fornecer alguns parâmetros, que podem ser obrigatórios ou não, como mostra a tabela 5.1.

O padrão é assumido pelo programa quando estes parâmetros não são especificados pelo usuário. Maiores detalhes a respeito dos parâmetros de entrada listados na tabela serão fornecidos ao longo deste capítulo.

Tabela 5.1 - Opções de Entrada do Programa

Parâmetros	Opções	Obrig.	Padrão
Nome Arquivo Original	-	Sim	-
Nome Arquivo Decodificado	-	Sim	-
Tipo de Mapeamento	Medida Subjetiva	Não	MOS
	Língua		Geral
	Mapeamento		3ª Ordem
Detalhamento dos Arquivos	Sim (opção 1)	Não	Não
	Não (opção 2)		
Frequência de Amostragem	8.000 Hz	Não	16.000 Hz
	16.000 Hz		
	32.000 Hz		
	64.000 Hz		
Atraso entre arquivos	1 a 1000	Não	Cálculo via correlação

5.1.2. Sub-rotina “geral”

Se o usuário optar por ter mais detalhes dos sinais de voz, entrará em funcionamento a sub-rotina *geral*. Neste caso, o sinal de voz original é executado, ou seja, faz-se uso dos recursos de áudio do computador para que o usuário possa ouvir o conteúdo do sinal; além disso, são traçadas as curvas (no domínio do tempo) dos dois sinais para efeito de comparação entre ambos. Por fim, é fornecida a energia média dos sinais, através de uma outra sub-rotina, chamada *activlev*, de acordo com as recomendações da ITU para este fim [1,5,6]. Se o usuário desejar mudar o valor da energia média dos sinais de voz, este terá à sua disposição uma terceira sub-rotina (*mudanivel*), a qual não foi incorporada ao programa, uma vez que os valores MOQV não são sensíveis a este parâmetro.

5.1.3. Cálculo do atraso

O alinhamento temporal entre os sinais é necessário, pois a estimativa de qualidade é baseada na diferença entre as representações psicofísicas destes sinais. O procedimento usado neste programa para o cálculo da defasagem é baseado nas sugestões encontradas na Rec. P. 861 [1]. O primeiro método sugerido para o alinhamento é o deslocamento apropriado dos sinais de uma quantidade igual ao atraso nominal produzido pelo codec. Quando não se conhece tal atraso, ou não é possível medi-lo, sugere-se o uso de uma medida de correlação cruzada entre os dois sinais, calculada em função do deslocamento relativo entre ambos. O atraso a ser empregado é aquele correspondente ao ponto de máximo da correlação.

O cálculo da correlação cruzada por sua expressão direta

$$r(n) = \sum_k x(k) \cdot y(n+k) \quad (5.1)$$

não é aconselhável, devido ao esforço computacional requerido. Então, implementou-se uma sub-rotina, chamada *fftp*, onde calcula-se a transformada de Fourier de $r(n)$ através de [57,58]:

$$R(e^{j\omega}) = Y(e^{j\omega})X^*(e^{j\omega}) \quad (5.2)$$

X^* e Y são obtidos via algoritmo de transformada rápida de Fourier (FFT). Por fim, a transformada inversa de $R(e^{j\omega})$ é também calculada via FFT. O índice do valor máximo da correlação representa o deslocamento entre os sinais, e o alinhamento dos mesmos é feito automaticamente.

O cálculo da FFT é feito através da sub-rotina *mfft*, a qual é uma versão modificada daquela encontrada no Matlab®. Esta versão usa apenas as componentes espectrais relevantes em telefonia, descartando, no processo de transformação para o domínio da frequência, todas as componentes acima de 4 kHz, mesmo aquelas com valor igual a 0, a fim de que estas componentes não sejam consideradas nos cálculos, reduzindo desta forma o esforço computacional. Este procedimento reduz consideravelmente o esforço computacional. O tempo requerido pela sub-rotina *mfft* chega a ser 15 vezes menor, dependendo do número de amostras dos arquivos de voz analisados, sem qualquer redução na confiabilidade dos resultados fornecidos. Esta sub-rotina também está incorporada à rotina principal.

5.1.4. Janelamento

Após o cálculo do início e término efetivos dos sinais, realiza-se o escalonamento global (equação 4.1). Em seguida, é feita a divisão dos arquivos de voz em quadros, usando uma janela de *Hanning*, através da sub-rotina *enframe*. Os sinais são divididos em quadros com uma duração fixa de 32 ms, o que equivale a 256 amostras para uma frequência de amostragem de 8 kHz, e a 512 amostras, para uma frequência de 16 kHz. Há uma superposição de 50% entre os quadros.

5.1.5. Cálculo da FFT e da Densidade Espectral de Energia Quadro a Quadro

O próximo passo é o cálculo da FFT para cada quadro, através da sub-rotina *ffts*, criada com a finalidade de melhorar a velocidade de processamento. Como estratégia geral, evitou-se a todo custo a utilização de “laços”, gerados por comandos como *for* ou *while*, ao longo do programa. Como alternativa, utilizou-se operações vetoriais ou matriciais. A única situação em que não foi possível evitar esse tipo de estrutura foi no cálculo da FFT para cada quadro. Verificou-se, no entanto, que fazer o cálculo de todas as FFTs num único “laço” demandava um tempo excessivo. Isto ocorreu devido à necessidade de se armazenar uma enorme quantidade de valores, os quais seriam posteriormente utilizados no cálculo das Densidades Espectrais de Energia (DEE) para cada quadro. É importante salientar que este problema se torna mais grave na medida em que o computador em que se está fazendo as simulações possua uma quantidade de memória RAM inferior a 128 Mbytes. A solução adotada na sub-rotina *ffts* foi realizar o cálculo das FFTs separadamente para pequenos grupos de quadros e utilizar os valores obtidos para o cálculo das respectivas DEEs, descartando-os em seguida. Este procedimento possibilitou um ganho significativo na velocidade de processamento, especialmente se o usuário não dispõe de um equipamento com uma boa capacidade de memória.

5.1.6. Mapeamento em Bandas

Esta etapa consiste no mapeamento da Densidade Espectral de Energia de cada quadro para as bandas críticas, através das equações 4.2 e 4.3, com o auxílio da sub-rotina *auxiliart*. A obtenção dos fatores de escalonamento de *pitch* e de sonoridade, usados nestas expressões, é mostrada a seguir.

5.1.7. Fatores de escalonamento

Os fatores de escalonamento de *pitch* (Ep) e de sonoridade (El) foram calculados de acordo com os procedimentos descritos nas Recs. P.830 [6] e P.861 [1], ou seja, através de uma senóide de 1 kHz. Os valores obtidos foram, respectivamente, $6,4487 \times 10^{-6}$ e 240,1293, ligeiramente diferentes daqueles encontrados para o programa PSQM original ($6,5994 \times 10^{-6}$ e 240,0485), o que pode ser explicado pelas diferenças inerentes ao tipo de processamento executado pelo Matlab® e pelo ANSI C. Esta calibração é feita através da rotina *seno*, a qual não foi incorporada à rotina principal, pois este cálculo é realizado apenas uma vez, já que, se respeitados os níveis estipulados pela Rec. P.861 para os sinais de voz, estes valores não mudam.

5.1.8. Filtragem na Faixa de Telefonia e Inserção de Ruído

Nesta etapa é realizada a filtragem com as características de recepção do aparelho telefônico, segundo as expressões 4.8 e 4.9, e a inserção de ruído do tipo Hoth (expressões 4.10 e 4.11), de acordo com os valores encontrados nas tabelas fornecidas na Rec. P.861 para cada intervalo de banda crítica.

Os processamentos seguintes, como o processamento cognitivo e o cálculo dos valores MOQV finais, são feitos segundo as expressões 4.12 a 4.25, não havendo a necessidade de maiores explicações.

5.1.9. Fator de Ponderação para MOQV2

O programa inclui também o fator de ponderação adicional adotado para o método PSQM+ (ver equações 4.26 e 4.27), o qual é utilizado somente no cálculo do valor MOQV2. É este processamento que faz a distinção entre os métodos MOQV1 e MOQV2.

5.1.10. Mapeamento

O programa fornece 48 diferentes opções de mapeamento, obtidas de maneira empírica através dos testes descritos no Capítulo 6. O usuário pode escolher uma destas opções ou deixar o programa realizar o mapeamento padrão, como mostrado na tabela 5.1. A escolha do tipo de mapeamento obedece a uma estrutura do tipo $map = xyz$, a qual é descrita na tabela 5.2.

Tabela 5.2 - Estrutura para escolha do tipo de mapeamento desejado

Parâmetro	Entrada	Significado
Medida Subjetiva Mapeada	x = 1	MOS
	x = 2	CMOS
Língua dos Arquivos de Voz	y = 1	Francês
	y = 2	Japonês
	y = 3	Inglês
	y = 4	Outra Língua (mapeamento geral)
Ordem do Polinômio de Mapeamento	z = 1	1ª Ordem
	z = 2	2ª Ordem
	z = 3	3ª Ordem
	z = 4	4ª Ordem
	z = 5	5ª Ordem
	z = 6	6ª Ordem

Como exemplo, o mapeamento padrão é dado por $\text{map} = 143$, ou seja, é um mapeamento geral de 3ª ordem para a medida subjetiva do tipo MOS. Esta é a opção sugerida para arquivos de voz em português e outras línguas que não aquelas citadas na tabela.

5.1.11. Valores Fornecidos na Saída do Programa

Após a realização de todo o processamento, o programa retornará para o usuário os seguintes valores:

- “atraso”, correspondendo à defasagem entre o arquivo original e o codificado;
- “moqv1”, correspondendo à medida objetiva moqv1;
- “moqv2”, correspondendo à medida objetiva moqv2;
- “medidasubjetiva1”, correspondendo ao valor obtido através do mapeamento da medida moqv1 para a medida subjetiva desejada;
- “medidasubjetiva2”, correspondendo ao valor obtido através do mapeamento da medida moqv2 para a medida subjetiva desejada;

Estes dois últimos valores, a critério do usuário, podem ou não ser considerados, já que os valores “moqv1” e “moqv2” podem ser suficientes para a avaliação do codec em questão. Este tópico será abordado com mais detalhes na seção 6.5.

5.1.12. Velocidade de Processamento

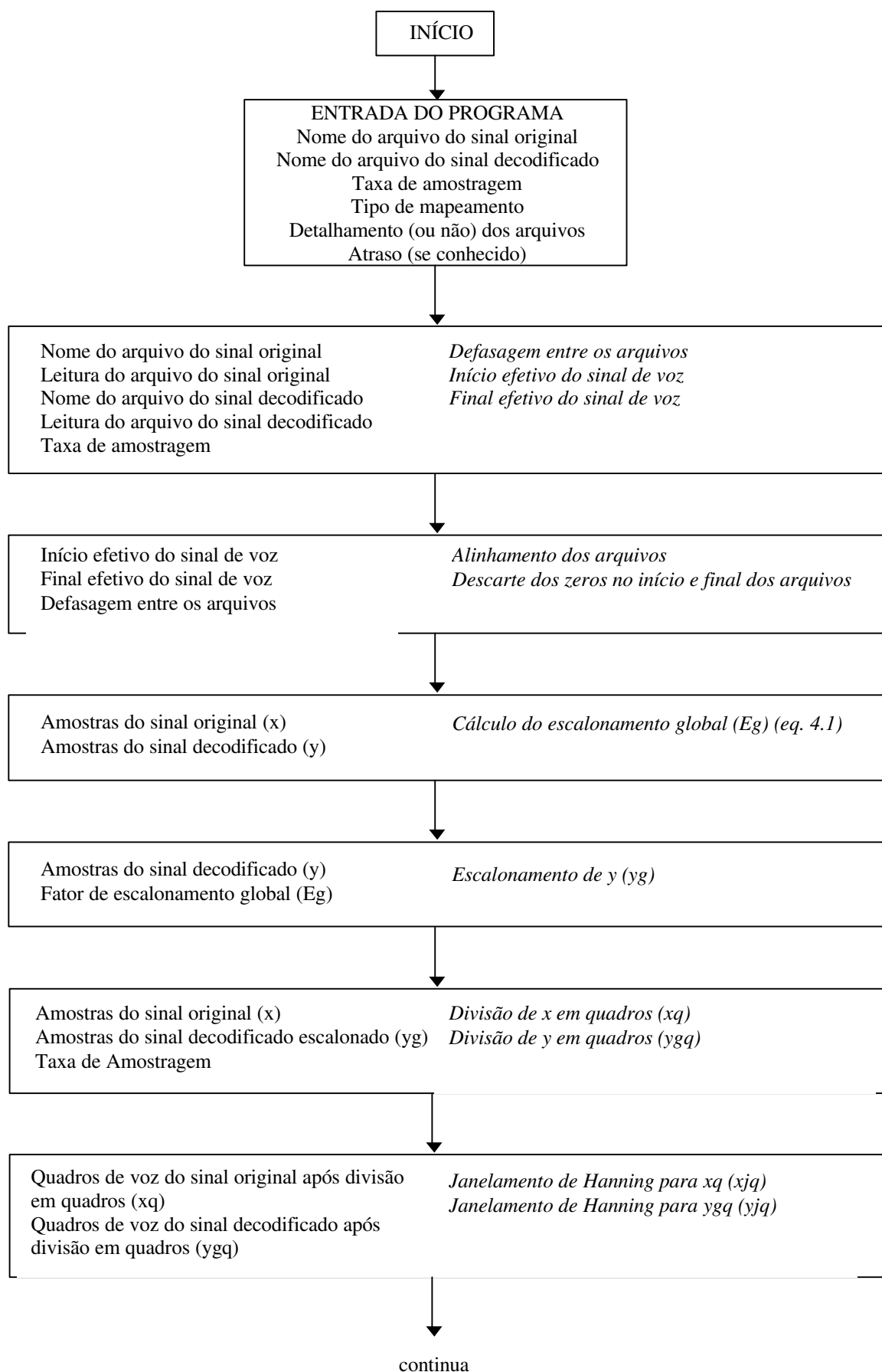
O programa foi testado em um computador PC com processador Intel Pentium III 650 MHz, com 128 Mbytes de memória RAM e sistema operacional Windows ME. Cinco diferentes versões do programa foram desenvolvidas, cada uma contendo modificações com o objetivo de aumentar sua velocidade. As principais medidas tomadas para tornar o programa mais rápido foram a utilização de rotinas mais rápidas para o cálculo das FFTs, redução da necessidade de armazenamento de dados e a substituição da maior parte dos “laços” por operações vetoriais e matriciais. Com isto, o tempo total de processamento do programa, para arquivos contendo 120.000 amostras, é de 4,78 segundos. Este tempo, em diversas situações, chega a ser inferior àquele gasto pelo programa original para o método PSQM, implementado em linguagem C e testado em estações Sun UltraSparc e ambiente Unix. Para arquivos de voz com menos de 30.000 amostras, o tempo gasto é de menos de 1 segundo.

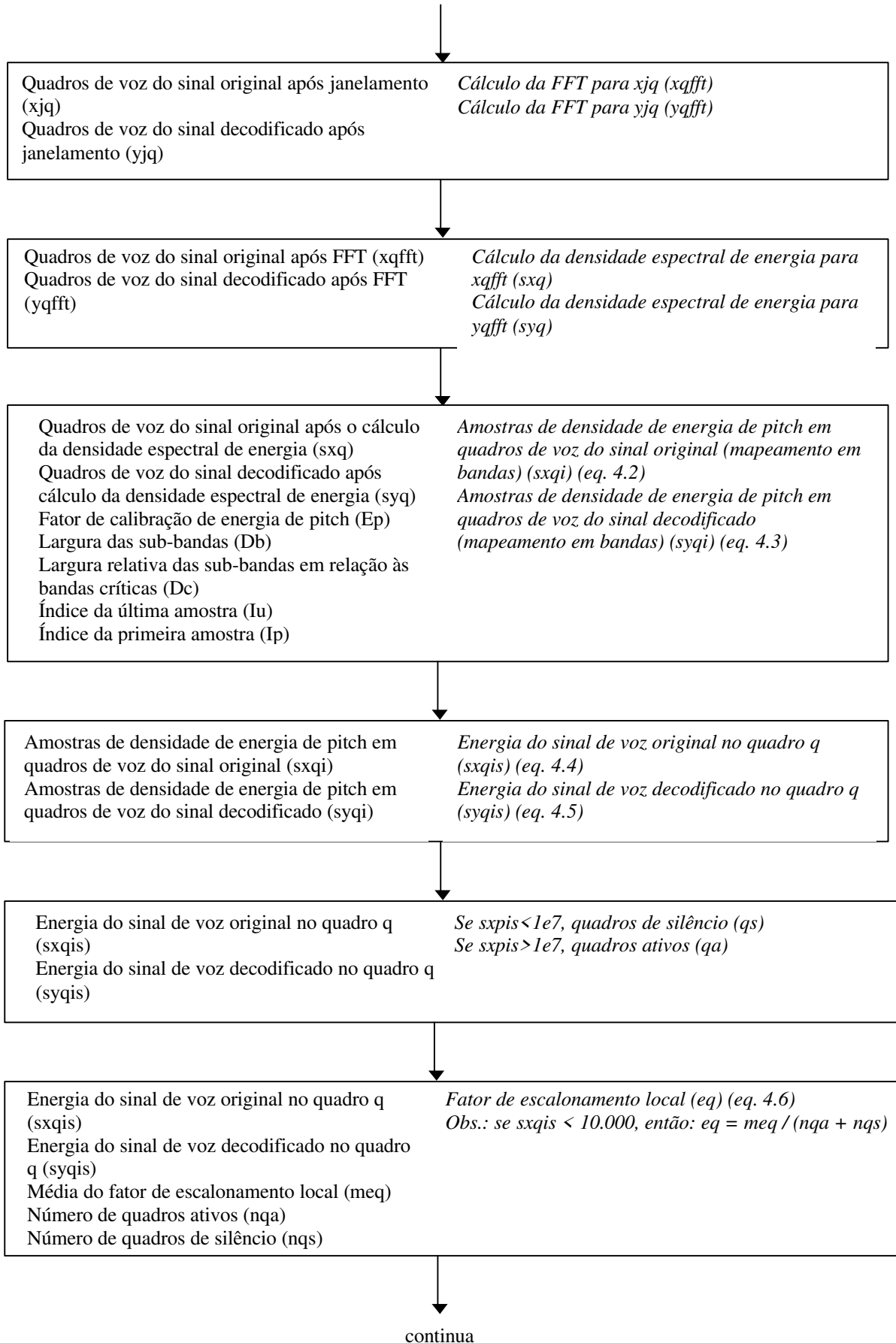
5.1.13. Versão Executável do Programa

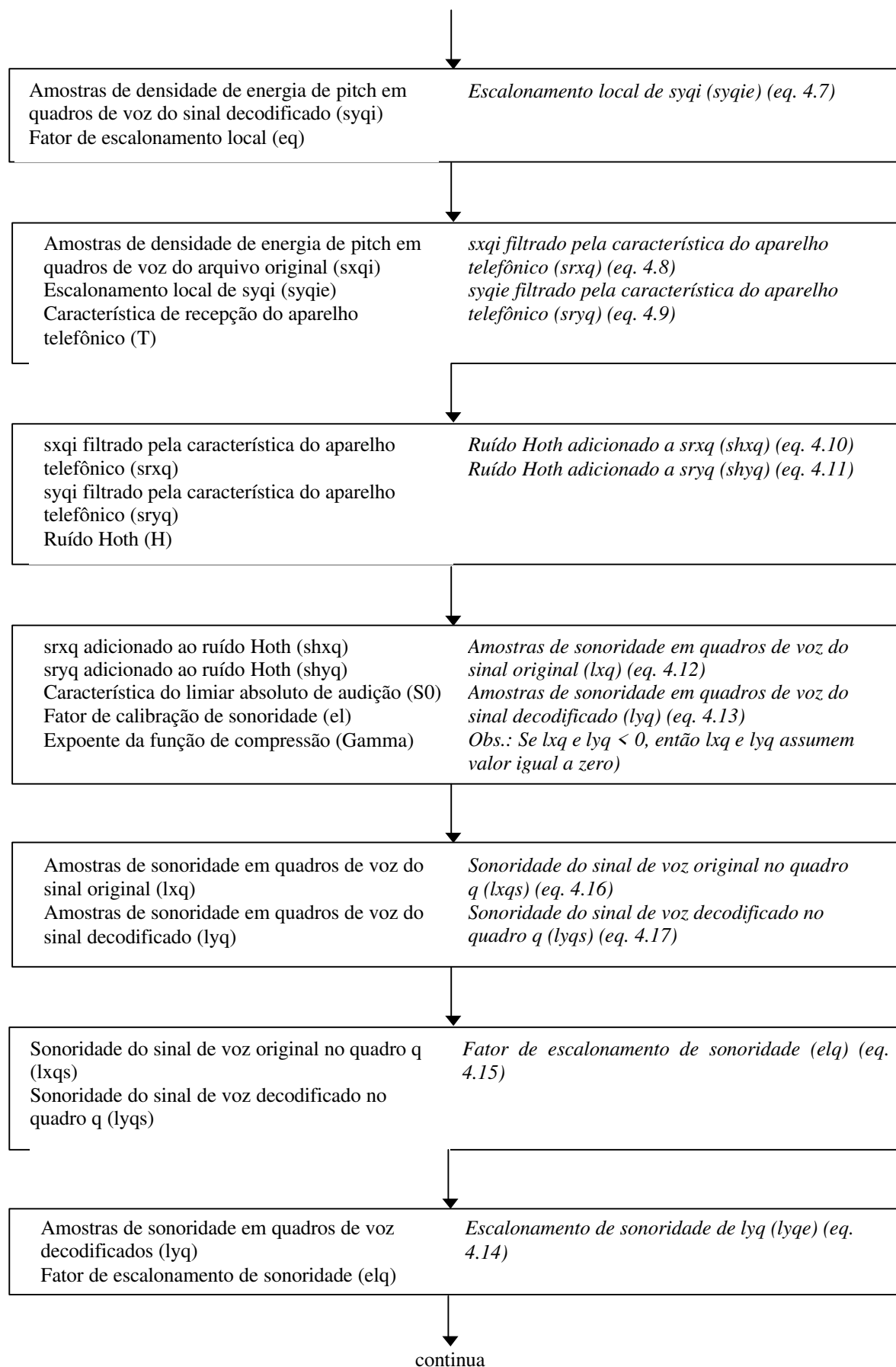
A fim de tornar o programa utilizável em qualquer computador, e não apenas naqueles que possuem instalado o Matlab®, uma versão executável foi gerada através de um compilador da linguagem C. Esta versão não possui as opções para a execução do conteúdo e apresentação das curvas dos sinais no domínio do tempo, porém todas as demais opções estão incluídas. Pequenas diferenças na entrada dos dados podem ser encontradas no arquivo texto que acompanha esta versão.

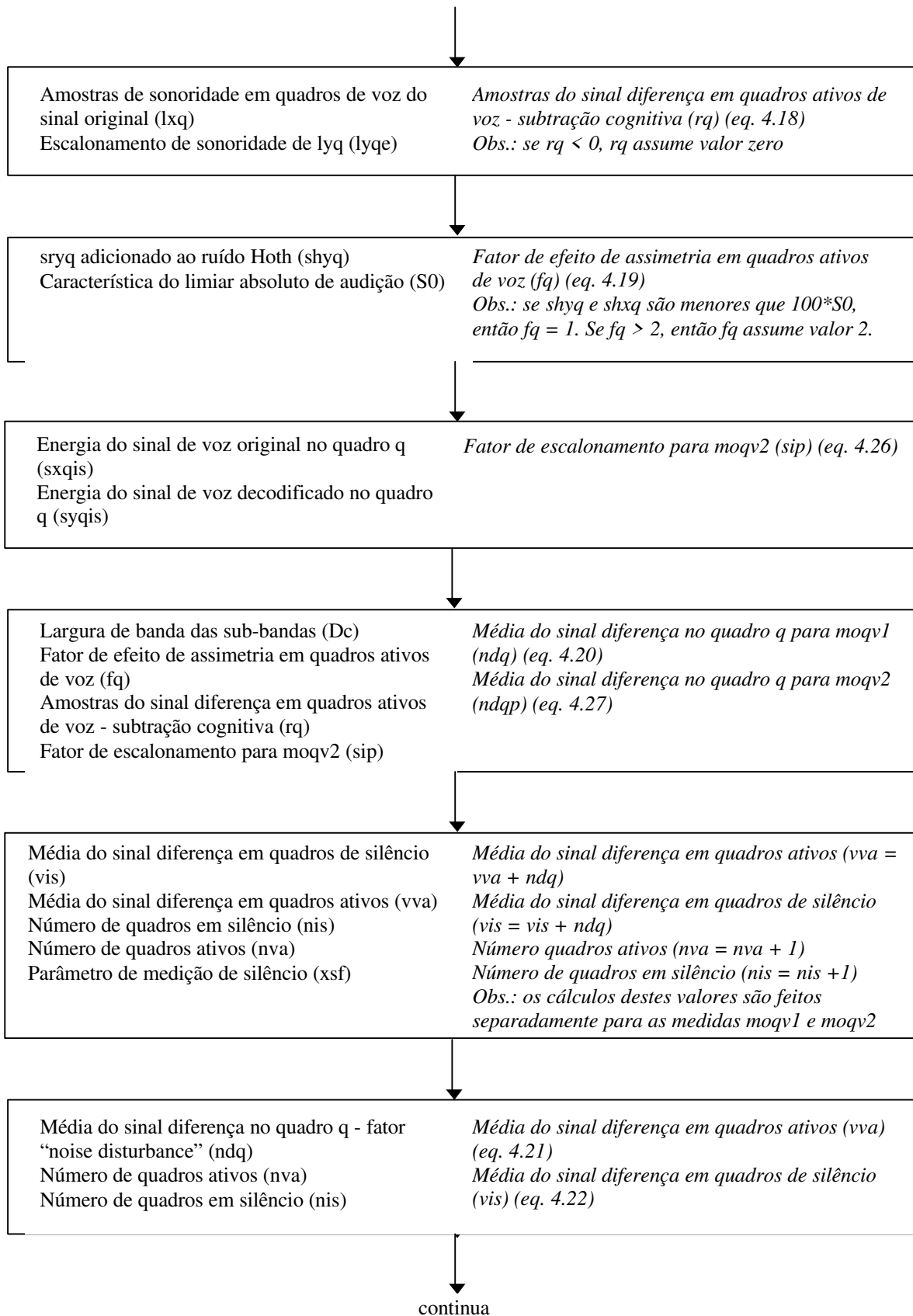
5.2. VISÃO GERAL DO PROGRAMA

O diagrama de blocos das próximas páginas (figura 5.1) representa uma visão detalhada das operações básicas realizadas no cálculo dos valores objetivos MOQV1 e MOQV2. Cada bloco representa uma etapa do programa. Nestes blocos tem-se as entradas (escrita normal) e as saídas (em itálico).









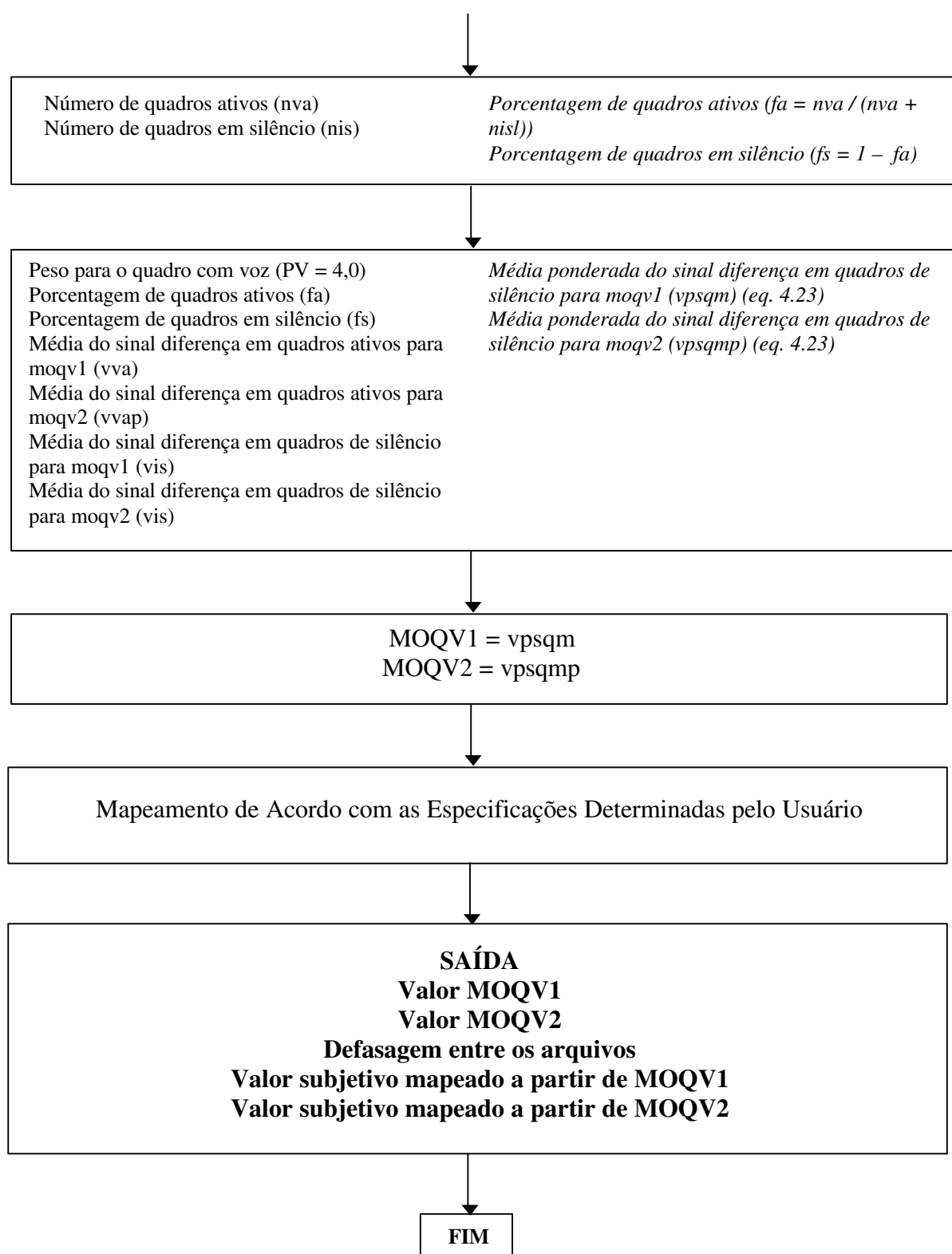


Figura 5.1 - Operações básicas realizadas no programa

CAPÍTULO 6

DESEMPENHO DO PROGRAMA MOQV

Este capítulo apresenta os resultados de testes de desempenho com o programa MOQV. Apresenta também o desempenho proporcionado por diversos tipos de curvas para o mapeamento entre as medidas MOQV e as medidas subjetivas correspondentes.

Todos os testes utilizaram os arquivos de voz contidos no banco de dados “Supplement 23” da ITU [59]. Alguns resultados foram comparados com testes realizados por outros autores, a fim de validar o programa MOQV.

6.1. DESCRIÇÃO DO BANCO DE DADOS “SUPPLEMENT 23”

Este banco de dados é constituído de uma série de arquivos de voz contendo sinais originais, suas versões após o processamento por codecs e as avaliações subjetivas correspondentes, fornecendo um amplo material para a realização de testes de desempenho de métodos de avaliação objetiva. Estes arquivos foram agrupados em três conjuntos, de acordo com as características das condições a que os codecs foram submetidos. A seguir serão descritos os procedimentos empregados na geração destes arquivos, bem como as características básicas de cada um dos três conjuntos, denominados experimentos.

6.1.1. Preparação do material fonte

- *Material de voz*: consiste de sentenças simples, claras e coerentes. O principal critério para a escolha dessas sentenças é a facilidade de entendê-las (sendo retiradas da literatura não-técnica, como jornais, por exemplo). A maior parte delas teve sua duração entre 2 e 3 segundos.

- *Produção do material fonte*: os locutores foram alocados, um de cada vez, em salas com tempo de reverberação inferior a 500 ms e nível de ruído menor que 30 dBA, sem picos dominantes no espectro. Os sistemas de gravação utilizados foram de alta qualidade.

- *Línguas e laboratórios*: o material gravado foi produzido em quatro diferentes línguas, ficando sua geração a cargo dos seguintes laboratórios: AT&T e BNR para a língua inglesa, CNET para a língua francesa, NTT para a língua japonesa e CSELT para a língua italiana.

- *Laboratório anfitrião*: todo o material gerado foi enviado a um laboratório anfitrião (*Deutsche Telekon*, da Alemanha), o qual se incumbiu da tarefa de preparar e pré-processar todo o material, para finalmente aplicá-lo a diversos tipos de codecs e condições de operação destes codecs, de modo a gerar cada um dos três experimentos.

- *Preparação e pré-processamento*: o material original foi amostrado a uma frequência de 16 kHz, com 16 bits por amostra, no sistema complemento de 2. As sentenças produzidas por cada locutor foram então agrupadas de modo a compor módulos de 128.000 amostras, isto é, com cerca de 8 segundos de duração, através do ajuste do intervalo de silêncio entre elas, resultando em arquivos de 250 kbytes. O próximo passo foi submeter os

sinais a um filtro com característica de transmissão IRS modificada (ver Apêndice C), antes de seu processamento através das várias condições de teste. Por fim, o nível médio de energia dos sinais amostrados foi ajustado para -30,15 dBov (ver definição de dBov no Apêndice C).

- *Geração de cada experimento*: o laboratório anfitrião gerou três experimentos distintos com o material obtido. A cada experimento, selecionou-se uma parte do material gravado, submetendo-a então a diversos codecs pré-selecionados. Tais codecs operaram em condições pré-estabelecidas, conforme descrito a seguir para cada experimento.

- *Testes Subjetivos*: os valores subjetivos (MOS e CMOS) fornecidos no banco de dados correspondem à média obtida para os locutores que compõem cada condição.

6.1.2. Descrição do primeiro experimento

Os arquivos de voz presentes neste experimento foram submetidos a diversos tipos de codecs padrão da ITU (G.711, G.726, G.728 e G.729), bem como a codecs utilizados na telefonia celular (GSM francês, VSELP americano e RCR Std 27C japonês). Os codecs foram utilizados individualmente e também em diversas combinações entre eles, como mostrado na tabela 6.1. Alguns arquivos foram submetidos a condições de referência (padrão MNRU com diferentes SNR). No total, o primeiro experimento é composto de 44 condições de teste.

Dois locutores de cada sexo foram utilizados para cada uma das 44 condições, perfazendo um total de 176 arquivos para teste. Estes dados são relativos a cada língua. Como foram utilizadas 3 línguas (francês, japonês e inglês), há um total global de 528 arquivos para o experimento. Cada um desses arquivos foi gravado em três diferentes versões: original, pré-processada (arquivos filtrados e equalizados) e decodificada. Nos testes apresentados neste capítulo, foram usadas as duas últimas.

Tabela 6.1 - Lista das condições - Experimento 1

Condição	1º Codec	2º Codec	3º Codec	4º Codec
1	G.729			
2	G.729	G.729		
3	G.729	G.729	G.729	
4	G.726			
5	G.726	G.726	G.726	G.726
6	G.728			
7	G.711			
8	GSM-FR			
9	IS-54			
10	JDC-HR			
11	G.729	G.726		
12	G.729	G.728		
13	G.729	GSM-FR		
14	G.729	IS-54		
15	G.729	JDC-HR		
16	G.726	G.729		
17	G.728	G.729		
18	GSM-FR	G.729		
19	IS-54	G.729		
20	JDC-HR	G.729		
21	G.729	G.729	GSM-FR	
22	G.729	G.729	IS-54	
23	G.729	G.729	JDC-HR	
24	G.729	G.726	GSM-FR	
25	G.729	G.728	GSM-FR	
26	GSM-FR	G.729	G.729	
27	IS-54	G.729	G.729	
28	JDC-HR	G.729	G.729	
29	GSM-FR	G.726	G.729	
30	GSM-FR	G.728	G.729	
31	GSM-FR	IS-54		
32	IS-54	JDC-HR		
33	JDC-HR	GSM-FR		
34	GSM-FR	G.729	IS-54	
35	IS-54	G.729	JDC-HR	
36	JDC-HR	G.729	GSM-FR	
37	MNRU 5			
38	MNRU 10			
39	MNRU 15			
40	MNRU 20			
41	MNRU 25			
42	MNRU 30			
43	MNRU 35			
44	MNRU 50			

6.1.3. Descrição do segundo experimento

Os arquivos de voz foram submetidos a diversos tipos de ruídos de ambiente - ruído de escritório, ruído de veículo, ruído de rua, ruído branco e música, com uma SNR de 10 ou 20 dB, como pode ser visto na tabela 6.2. As medidas subjetivas utilizadas neste experimento foram as do tipo CMOS, enquanto que para os outros utilizou-se o MOS.

Neste experimento, gerou-se um total de 40 condições. Para as 28 primeiras, utilizou-se dois locutores de cada sexo e, para as últimas 12, apenas um. Há então um total de 136 arquivos de teste para cada uma das 3 línguas utilizadas (francês, inglês e japonês), perfazendo um total global de 408 arquivos. Como no primeiro experimento, cada um destes arquivos foi gravado em três diferentes versões, e os testes se basearam nos arquivos pré-processados e decodificados.

Tabela 6.2 - Lista das condições - Experimento 2

Condição	Codec	Transcod.	Ruído	SNR(dB)	Q(dB)	Caract. de Entrada
1	G.729	1	Limpo	–	–	Lei A, IRS Mod.
2	G.729	1	Escritório	20	–	Lei A, IRS Mod.
3	G.729	1	Veículo	20	–	Lei A, IRS Mod.
4	G.729	1	Veículo	10	–	Lei A, IRS Mod.
5	G.729	1	Rua	20	–	Lei A, IRS Mod.
6	G.729	1	Branco	20	–	Lei A, IRS Mod.
7	G.729	1	Música	20	–	Lei A, IRS Mod.
8	G.729	1	Limpo	–	–	Lei A, IRS Mod.
9	G.729	2	Escritório	20	–	Lei A, IRS Mod.
10	G.729	2	Veículo	20	–	Lei A, IRS Mod.
11	G.729	2	Veículo	10	–	Lei A, IRS Mod.
12	G.729	2	Rua	20	–	Lei A, IRS Mod.
13	G.729	2	Branco	20	–	Lei A, IRS Mod.
14	G.729	2	Música	20	–	Lei A, IRS Mod.
15	G.726	1	Limpo	–	–	Lei A, IRS Mod.
16	G.726	1	Escritório	20	–	Lei A, IRS Mod.
17	G.726	1	Veículo	20	–	Lei A, IRS Mod.
18	G.726	1	Veículo	10	–	Lei A, IRS Mod.
19	G.726	1	Rua	20	–	Lei A, IRS Mod.
20	G.726	1	Branco	20	–	Lei A, IRS Mod.
21	G.726	1	Música	20	–	Lei A, IRS Mod.
22	G.726	2	Limpo	–	–	Lei A, IRS Mod.
23	G.726	2	Escritório	20	–	Lei A, IRS Mod.
24	G.726	2	Veículo	20	–	Lei A, IRS Mod.
25	G.726	2	Veículo	10	–	Lei A, IRS Mod.
26	G.726	2	Rua	20	–	Lei A, IRS Mod.
27	G.726	2	Branco	20	–	Lei A, IRS Mod.
28	G.726	2	Música	20	–	Lei A, IRS Mod.
29	MNRU	–	Limpo	–	6 dB	UPCM, IRS Mod.
30	MNRU	–	Limpo	–	12 dB	UPCM, IRS Mod.
31	MNRU	–	Limpo	–	18 dB	UPCM, IRS Mod.
32	MNRU	–	Limpo	–	24 dB	UPCM, IRS Mod.
33	MNRU	–	Limpo	–	30 dB	UPCM, IRS Mod.
34	Direto	–	Limpo	–	–	UPCM, IRS Mod.
35	Direto	–	Escritório	20	–	UPCM, IRS Mod.
36	Direto	–	Veículo	20	–	UPCM, IRS Mod.
37	Direto	–	Veículo	10	–	UPCM, IRS Mod.
38	Direto	–	Rua	20	–	UPCM, IRS Mod.
39	Direto	–	Branco	20	–	UPCM, IRS Mod.
40	Direto	–	Música	20	–	UPCM, IRS Mod.

Obs.: UPCM = Uniform Pulse Code Modulation.

6.1.4. Descrição do terceiro experimento

As 50 condições deste experimento simulam um canal de comunicação degradado por erros, em particular erros aleatórios e erros de surto (*burst frame erasure*).

Para cada condição e para cada língua foram utilizados dois locutores de cada sexo, perfazendo assim um total parcial de 200 arquivos para teste. Aqui, além das línguas já citadas, há também o italiano, resultando num total de 800 arquivos de teste para o experimento. Novamente, foram utilizadas as versões pré-processadas e decodificadas dos arquivos. A tabela 6.3 descreve cada uma das 50 condições presentes no experimento.

Tabela 6.3 - Alocação de Condições para o Experimento 3

Cond.	Codec	Transc.	Ruído	Tipo de Erro	Taxa de Erro (%)	Característica de Entrada	Valor Q
1	G.729	1	Limpo	–	–	Lei A, IRS Mod.	–
2	G.729	1	Limpo	Quadro Aleatório (Random Frame)	3	Lei A, IRS Mod.	–
3	G.729	1	Limpo	Quadro Aleatório (Random Frame)	5	Lei A, IRS Mod.	–
4	G.729	1	Limpo	Surto (Bursty Frame)	3	Lei A, IRS Mod.	–
5	G.729	1	Limpo	Surto (Bursty Frame)	5	Lei A, IRS Mod.	–
6	G.729	1	Veículo	–	–	Lei A, IRS Mod.	–
7	G.729	1	Veículo	Quadro Aleatório (Random Frame)	3	Lei A, IRS Mod.	–
8	G.729	1	Veículo	Quadro Aleatório (Random Frame)	5	Lei A, IRS Mod.	–
9	G.729	1	Veículo	Surto (Bursty Frame)	3	Lei A, IRS Mod.	–
10	G.729	1	Veículo	Surto (Bursty Frame)	5	Lei A, IRS Mod.	–
11	G.729	1	Rua	–	–	Lei A, IRS Mod.	–
12	G.729	1	Rua	Quadro Aleatório (Random Frame)	3	Lei A, IRS Mod.	–
13	G.729	1	Rua	Quadro Aleatório (Random Frame)	5	Lei A, IRS Mod.	–
14	G.729	1	Rua	Surto (Bursty Frame)	3	Lei A, IRS Mod.	–
15	G.729	1	Rua	Surto (Bursty Frame)	5	Lei A, IRS Mod.	–
16	G.729	1	Hoth	–	–	Lei A, IRS Mod.	–
17	G.729	1	Hoth	Quadro Aleatório (Random Frame)	3	Lei A, IRS Mod.	–
18	G.729	1	Hoth	Quadro Aleatório (Random Frame)	5	Lei A, IRS Mod.	–
19	G.729	1	Hoth	Surto (Bursty Frame)	3	Lei A, IRS Mod.	–
20	G.729	1	Hoth	Surto (Bursty Frame)	5	Lei A, IRS Mod.	–
21	G.729	2	Limpo	–	–	Lei A, IRS Mod.	–
22	G.729	3	Limpo	–	–	Lei A, IRS Mod.	–
23	G.729	2	Limpo	Quadro Aleatório (Random Frame)	3, 3	Lei A, IRS Mod.	–
24	G.729	3	Limpo	Quadro Aleatório (Random Frame)	3, 0, 3	Lei A, IRS Mod.	–
25	G.729	2	Limpo	Surto (Bursty Frame)	3, 3	Lei A, IRS Mod.	–
26	G.729	3	Limpo	Surto (Bursty Frame)	3, 0, 3	Lei A, IRS Mod.	–
27	G.729	2	Veículo	Quadro Aleatório (Random Frame)	3, 3	Lei A, IRS Mod.	–
28	G.729	2	Veículo	Surto (Bursty Frame)	3, 3	Lei A, IRS Mod.	–
29	G.729	1	Limpo	Bit Aleatório (Random Bit)	1	Lei A, IRS Mod.	–
30	G.729	1	Limpo	Bit Aleatório (Random Bit)	3	Lei A, IRS Mod.	–
31	G.729	1	Limpo	Bit Aleatório (Random Bit)	5	Lei A, IRS Mod.	–
32	G.729	1	Limpo	Bit Aleatório (Random Bit)	10	Lei A, IRS Mod.	–
33	G.729	1	Limpo	Surto/Bit Aleatório	3, 1	Lei A, IRS Mod.	–
34	G.729	1	Limpo	Surto/Bit Aleatório	3, 3	Lei A, IRS Mod.	–
35	G.729	1	Limpo	Surto/Bit Aleatório	3, 5	Lei A, IRS Mod.	–
36	G.729	1	Limpo	Surto/Bit Aleatório	3, 10	Lei A, IRS Mod.	–
37	G.726	1	Limpo	–	–	Lei A, IRS Mod.	–
38	G.726	1	Veículo	–	–	Lei A, IRS Mod.	–
39	G.726	1	Rua	–	–	Lei A, IRS Mod.	–
40	G.726	1	Hoth	–	–	Lei A, IRS Mod.	–
41	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	10 dB
42	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	15 dB
43	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	20 dB
44	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	25 dB
45	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	30 dB
46	MNRU	1	Limpo	–	–	UPCM, IRS Mod.	50 dB
47	Direct	–	Limpo	–	–	UPCM, IRS Mod.	–
48	Direct	–	Veículo	–	–	UPCM, IRS Mod.	–
49	Direct	–	Rua	–	–	UPCM, IRS Mod.	–
50	Direct	–	Hoth	–	–	UPCM, IRS Mod.	–

6.1.5. Visão geral do banco de dados “Supplement 23”

A figura 6.1 fornece uma visão geral da estrutura do banco de dados.

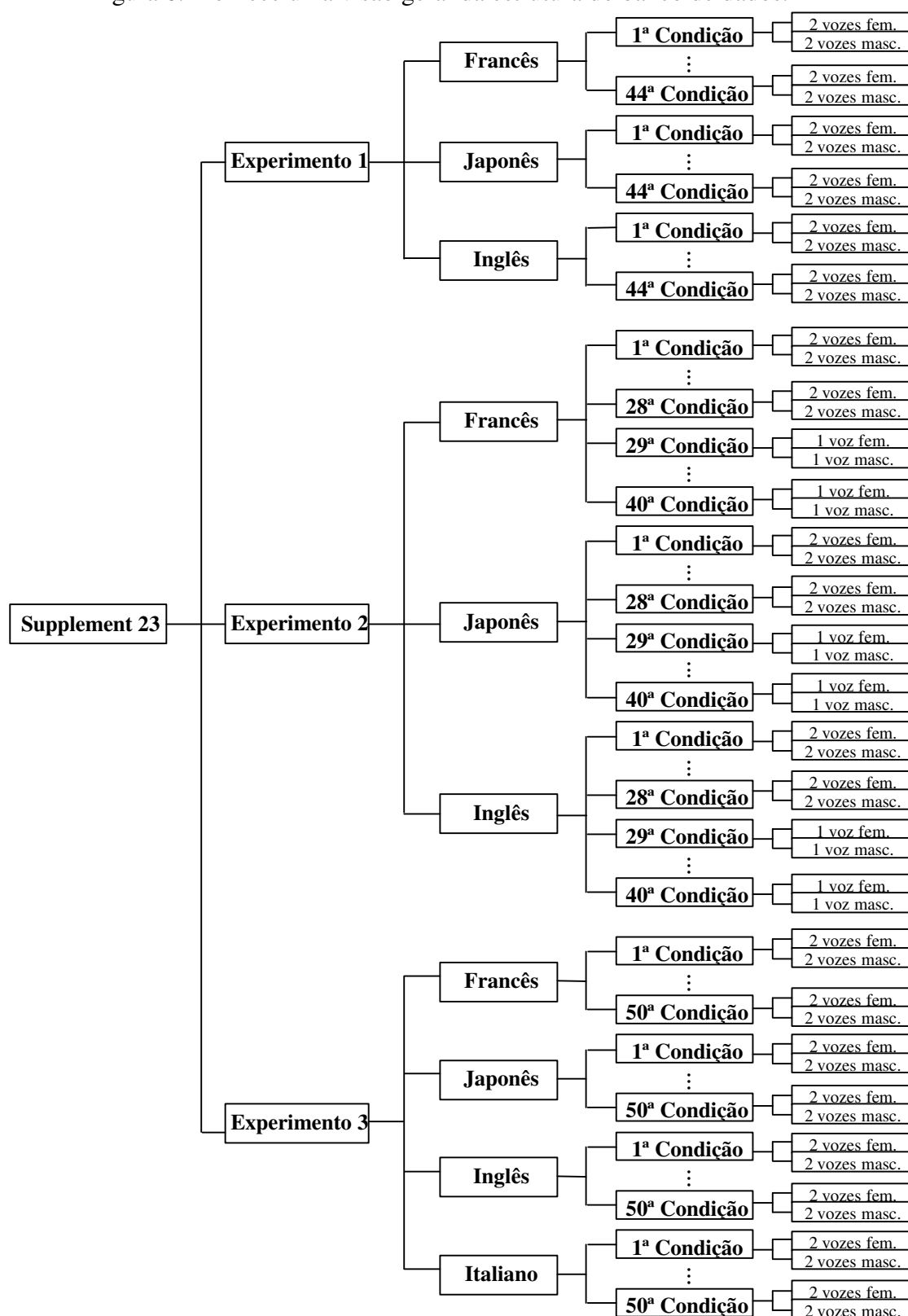


Figura 6.1 - Visão geral do banco de dados “Supplement 23”

6.2. DESCRIÇÃO DOS TESTES COM O PROGRAMA MOQV

O principal parâmetro analisado nos testes foi a correlação entre as medidas MOQV e as medidas subjetivas correspondentes, calculada para curvas de mapeamento do tipo polinomial com várias opções de ordem. Procurou-se abordar o maior número possível de situações, a fim de se traçar um perfil do desempenho do método MOQV. Para isso, considerou-se:

- as características de cada um dos três experimentos submetidos aos testes;
- mapeamento para cada língua e mapeamento geral médio em relação a todas as línguas presentes no banco de dados;
- a influência do valor da ordem polinomial p do mapeamento, permitindo assim determinar as ordens mais adequadas.

Outros parâmetros estatísticos, como desvios e histogramas, foram utilizados, a fim de complementar e reforçar as conclusões obtidas na análise da correlação.

A seguir, serão descritas algumas características dos testes realizados.

6.2.1. Processo de mapeamento

Foi utilizado um mapeamento polinomial do tipo [60]:

$$y_i = \beta_1 + \beta_2 \cdot x_i + \dots + \beta_p \cdot x_i^p \quad (6.1)$$

onde x_i representa as medidas objetivas e y_i representa as medidas subjetivas. Os parâmetros β_i são otimizados, para n medidas resultantes de n condições, pela minimização de

$$Q = \sum_{i=1}^n [y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_i + \dots + \hat{\beta}_p x_i^p)]^2 \quad (6.2)$$

Esta minimização é feita através de três passos:

- 1 - Fazendo as derivadas parciais de Q com respeito a $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$
- 2 - Igualando cada derivada parcial a zero e
- 3 - Resolvendo o sistema resultante de p equações e p incógnitas.

Por exemplo, para o caso de um mapeamento polinomial de 1ª ordem, ter-se-á as seguintes estimativas para os parâmetros:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad e \quad \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} \quad (6.3)$$

onde $\hat{\beta}_2$ é o coeficiente angular da reta e $\bar{x}$ e $\bar{y}$ representam as médias para as medidas objetivas e subjetivas, respectivamente.

6.2.2. Cálculo da correlação entre as medidas objetivas após mapeamento e medidas subjetivas

A correlação entre as medidas objetivas depois de determinado mapeamento e as medidas subjetivas correspondentes é calculada através do coeficiente de correlação de Pearson [61]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(z_i - \bar{z})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (z_i - \bar{z})^2}} \quad (6.4)$$

onde z_i representa os valores objetivos após o mapeamento e $\bar{z}$, sua média. Quanto mais próximo de 1 for o módulo de r , melhor é a qualidade do método e do mapeamento.

Outra maneira de se descrever o desempenho das medidas objetivas de qualidade e do respectivo mapeamento é através do cálculo da raiz quadrada do erro quadrático médio (RMSE) [62], dado por:

$$RMSE = \sqrt{\frac{1}{n} \sum_i (x_i - z_i)^2} \quad (6.5)$$

Quanto mais próximo de zero for este valor, melhor a medida.

6.2.3. Critérios utilizados para o mapeamento

Os testes com o programa MOQV foram feitos segundo dois critérios: o primeiro consiste em se fazer o mapeamento individualmente para cada língua, o que resulta em uma melhor correlação entre as medidas objetivas e os valores subjetivos, já que esta abordagem faz com que os resultados estejam mais bem adaptados às peculiaridades daquela língua e cultura. O segundo critério consiste em se fazer um mapeamento geral, levando-se em consideração todos os arquivos de todas as línguas disponíveis, o que leva a uma fórmula de mapeamento geral. Neste caso, as peculiaridades de cada língua são absorvidas no processo de mediação, fazendo com que a correlação diminua. A Ericsson apresenta alguns resultados de alguns testes utilizando o primeiro critério [63]. Já a KPN apresenta alguns resultados utilizando o segundo critério [62]. Estes resultados parciais foram utilizados para a validação do programa MOQV.

6.2.4. Obtenção das curvas de mapeamento

Todas as curvas de mapeamento testadas são do tipo polinomial. Em todos os casos, variou-se a ordem de mapeamento desde a unidade até a ordem 6, visando estudar a influência da ordem na qualidade da correlação entre as medidas objetivas e subjetivas correspondentes.

Para o primeiro critério, foram obtidos conjuntos de 6 curvas (variando a ordem) para cada língua a cada experimento. Assim, como o primeiro experimento conta com três línguas, foi gerado, neste experimento, um total de 18 curvas para o método MOQV1 e mais 18 para o método MOQV2. Cada uma destas curvas resulta da otimização dos coeficientes do polinômio, considerando todos os arquivos disponíveis para cada língua no experimento. O mesmo procedimento foi adotado nos demais experimentos.

Para o segundo critério, seguiu-se o procedimento do caso anterior, exceto pela particularização por língua. Assim, no primeiro experimento foram geradas 6 curvas (variando a ordem do polinômio) para o método MOQV1 e mais 6 para o MOQV2. Cada curva resultou de uma otimização sobre todos os arquivos empregados no experimento.

A tabela 6.4 resume o conjunto de situações abordadas.

Tabela 6.4 - Curvas Geradas

Critério	Exper.	Língua	Método	n° condições	n° arquivos	Curvas
1°	1°	Japonês	MOQV1	44	176	MOQV1 x MOS
			MOQV2	44	176	MOQV2 x MOS
		Francês	MOQV1	44	176	MOQV1 x MOS
			MOQV2	44	176	MOQV2 x MOS
		Inglês	MOQV1	44	176	MOQV1 x MOS
			MOQV2	44	176	MOQV2 x MOS
	2°	Japonês	MOQV1	40	136	MOQV1 x CMOS
			MOQV2	40	136	MOQV2 x CMOS
		Francês	MOQV1	40	136	MOQV1 x CMOS
			MOQV2	40	136	MOQV2 x CMOS
		Inglês	MOQV1	40	136	MOQV1 x CMOS
			MOQV2	40	136	MOQV2 x CMOS
	3°	Japonês	MOQV1	50	200	MOQV1 x MOS
			MOQV2	50	200	MOQV2 x MOS
		Francês	MOQV1	50	200	MOQV1 x MOS
			MOQV2	50	200	MOQV2 x MOS
		Inglês	MOQV1	50	200	MOQV1 x MOS
			MOQV2	50	200	MOQV2 x MOS
		Italiano	MOQV1	50	200	MOQV1 x MOS
			MOQV2	50	200	MOQV2 x MOS
2°	1°	Todas	MOQV1	132	528	MOQV1 x MOS
			MOQV2	132	528	MOQV2 x MOS
	2°	Todas	MOQV1	120	408	MOQV1 x CMOS
			MOQV2	120	408	MOQV2 x CMOS
	3°	Todas	MOQV1	200	800	MOQV1 x MOS
			MOQV2	200	800	MOQV2 x MOS

A seguir serão mostrados os valores de avaliação objetiva MOQV1 e MOQV2 obtidos para os três experimentos. Todas as medidas foram realizadas com uma ponderação de 0,2 para os intervalos de silêncio (PS na equação 4.24).

6.3. RESULTADOS DE SIMULAÇÃO

A seguir são mostradas, para cada experimento, tabelas com os valores MOQV1 e MOQV2 para cada uma das línguas em que foram gravados os arquivos de voz. Associados a cada tabela, há dois gráficos mostrando as curvas de mapeamento para o primeiro critério e os pontos utilizados em sua determinação, através das equações 6.1 e 6.2. Após, serão apresentados os mapeamentos obtidos a partir do segundo critério. Por fim, será realizada uma análise estatística mais apurada, a fim de se obter subsídios para um exame mais profundo de todo o processo, o qual será detalhado na seção 6.4.

6.3.1. Resultados para o 1º Critério

• Experimento 1

Através da observação das tabelas 6.5 a 6.7 e das figuras 6.2 a 6.7, constata-se um comportamento relativamente homogêneo da avaliação objetiva para todas as línguas analisadas, apresentando correlações elevadas para ambos os métodos, com ligeira vantagem para o MOQV2. A ordem de mapeamento que mostrou ser a mais adequada foi a terceira, já que apresenta resultados significativamente melhores, sem o inconveniente das acentuadas perdas da monotonicidade apresentadas, na maioria dos casos, pelas ordens superiores.

Tabela 6.5 - Valores obtidos através de simulação - Experimento 1 - Francês

<i>Laboratório a (CNET - Francês)</i>											
Condição	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,556	1,436	1,319	1,058	1,342	1,587	1,481	1,359	1,090	1,379	4,16
2	1,822	2,186	1,566	1,765	1,835	1,903	2,270	1,594	1,892	1,915	3,54
3	2,506	2,520	2,555	2,582	2,541	2,727	2,618	2,801	2,852	2,750	3,03
4	0,824	0,753	0,581	0,734	0,723	0,830	0,750	0,581	0,736	0,724	4,25
5	1,642	1,897	1,566	1,401	1,627	1,635	1,896	1,556	1,386	1,618	3,14
6	1,162	1,446	0,839	0,764	1,053	1,164	1,450	0,848	0,768	1,058	3,82
7	0,258	0,247	0,248	0,155	0,227	0,257	0,245	0,248	0,154	0,226	4,22
8	2,000	1,982	1,516	0,916	1,604	2,033	1,996	1,604	0,928	1,640	3,82
9	2,038	2,112	1,546	1,297	1,748	2,078	2,163	1,599	1,342	1,796	3,73
10	2,268	2,436	2,282	1,447	2,108	2,389	2,555	2,422	1,549	2,229	3,15
11	1,508	1,645	1,243	1,081	1,369	1,544	1,672	1,316	1,115	1,412	3,75
12	2,000	1,831	1,633	1,396	1,715	2,042	1,878	1,683	1,448	1,763	3,44
13	2,392	2,306	1,844	1,601	2,036	2,458	2,374	1,967	1,631	2,108	3,33
14	2,580	2,410	2,006	1,801	2,199	2,716	2,460	2,162	1,898	2,309	3,17
15	2,835	2,672	2,648	2,129	2,571	3,036	2,800	2,923	2,215	2,744	2,64
16	1,694	1,709	1,455	1,034	1,473	1,767	1,755	1,537	1,063	1,531	3,78
17	1,680	1,735	1,704	1,338	1,614	1,735	1,790	1,792	1,422	1,685	3,64
18	2,446	2,269	2,150	1,776	2,160	2,561	2,465	2,267	1,840	2,283	3,26
19	2,324	2,195	1,895	1,571	1,996	2,427	2,282	2,075	1,655	2,110	3,35
20	2,837	2,658	2,424	2,350	2,567	3,147	2,843	2,667	2,588	2,811	2,28
21	2,674	2,632	2,075	1,896	2,319	2,942	2,701	2,225	2,003	2,468	2,77
22	2,321	2,393	2,726	1,843	2,321	2,519	2,554	3,009	1,995	2,519	2,61
23	2,896	2,536	3,038	2,779	2,812	3,079	2,716	3,424	3,080	3,075	1,99
24	2,369	2,809	1,856	1,755	2,197	2,444	2,891	1,980	1,858	2,293	3,14
25	2,761	2,840	2,009	1,578	2,297	2,913	2,868	2,122	1,656	2,390	2,90
26	2,624	2,408	2,138	2,063	2,308	2,825	2,463	2,346	2,152	2,447	2,83
27	2,571	2,699	2,789	2,123	2,546	2,685	2,796	2,995	2,255	2,683	2,79
28	2,688	3,101	3,111	2,472	2,843	2,888	3,340	3,531	2,767	3,132	2,02
29	2,356	2,206	1,825	1,548	1,984	2,518	2,308	1,992	1,603	2,105	3,23
30	2,206	2,210	2,283	1,740	2,110	2,373	2,280	2,478	1,789	2,230	3,30
31	2,285	2,311	1,813	1,451	1,965	2,338	2,394	1,960	1,485	2,044	3,36
32	2,919	2,693	2,271	2,001	2,471	3,208	2,910	2,458	2,154	2,683	2,32
33	3,003	2,712	2,178	1,951	2,461	3,173	2,850	2,621	2,145	2,697	2,54
34	2,632	3,104	2,502	1,942	2,545	2,821	3,273	2,776	2,014	2,721	2,79
35	2,883	2,897	3,262	2,580	2,906	3,049	3,185	3,668	2,916	3,205	2,05
36	2,816	3,069	2,359	2,340	2,646	2,928	3,231	2,656	2,435	2,813	2,40
37	5,468	6,487	3,367	3,782	4,776	5,497	6,506	3,401	3,820	4,806	1,18
38	4,899	3,524	2,496	3,075	3,499	4,910	3,540	2,505	3,100	3,514	1,51
39	3,410	2,919	1,905	1,879	2,528	3,416	2,929	1,910	1,884	2,535	2,10
40	1,700	1,893	0,974	1,161	1,432	1,701	1,893	0,982	1,161	1,434	2,99
41	1,289	1,080	0,770	0,560	0,925	1,291	1,085	0,771	0,560	0,927	3,64
42	0,824	0,563	0,375	0,264	0,507	0,827	0,563	0,380	0,265	0,509	4,06
43	0,384	0,334	0,218	0,157	0,273	0,387	0,340	0,218	0,158	0,276	4,35
44	0,122	0,112	0,132	0,053	0,105	0,127	0,113	0,135	0,053	0,107	4,27

Obs.: VF = voz feminina ; VM = voz masculina.

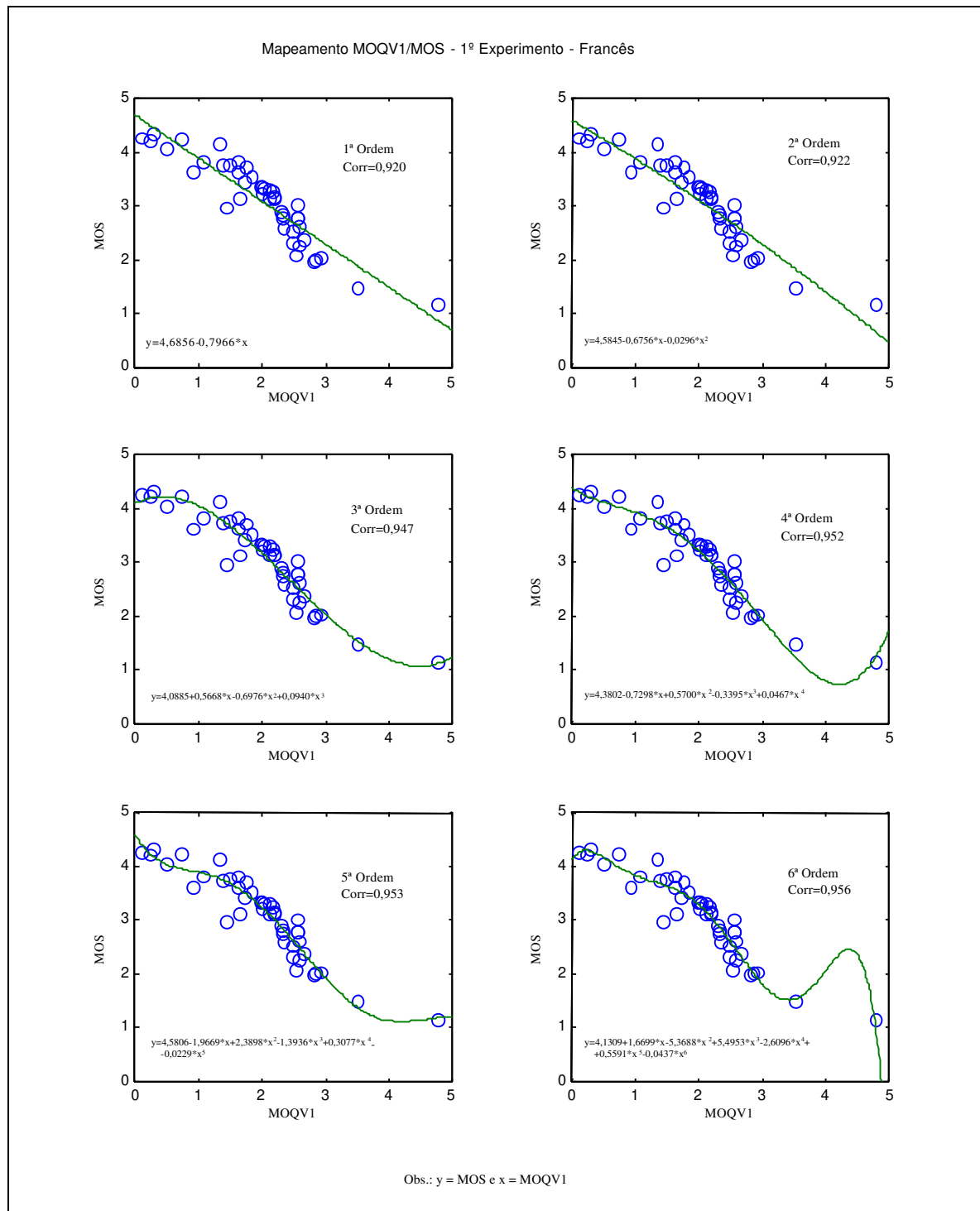


Figura 6.2 - Mapeamento MOQV1/MOS - 1º Experimento - Francês

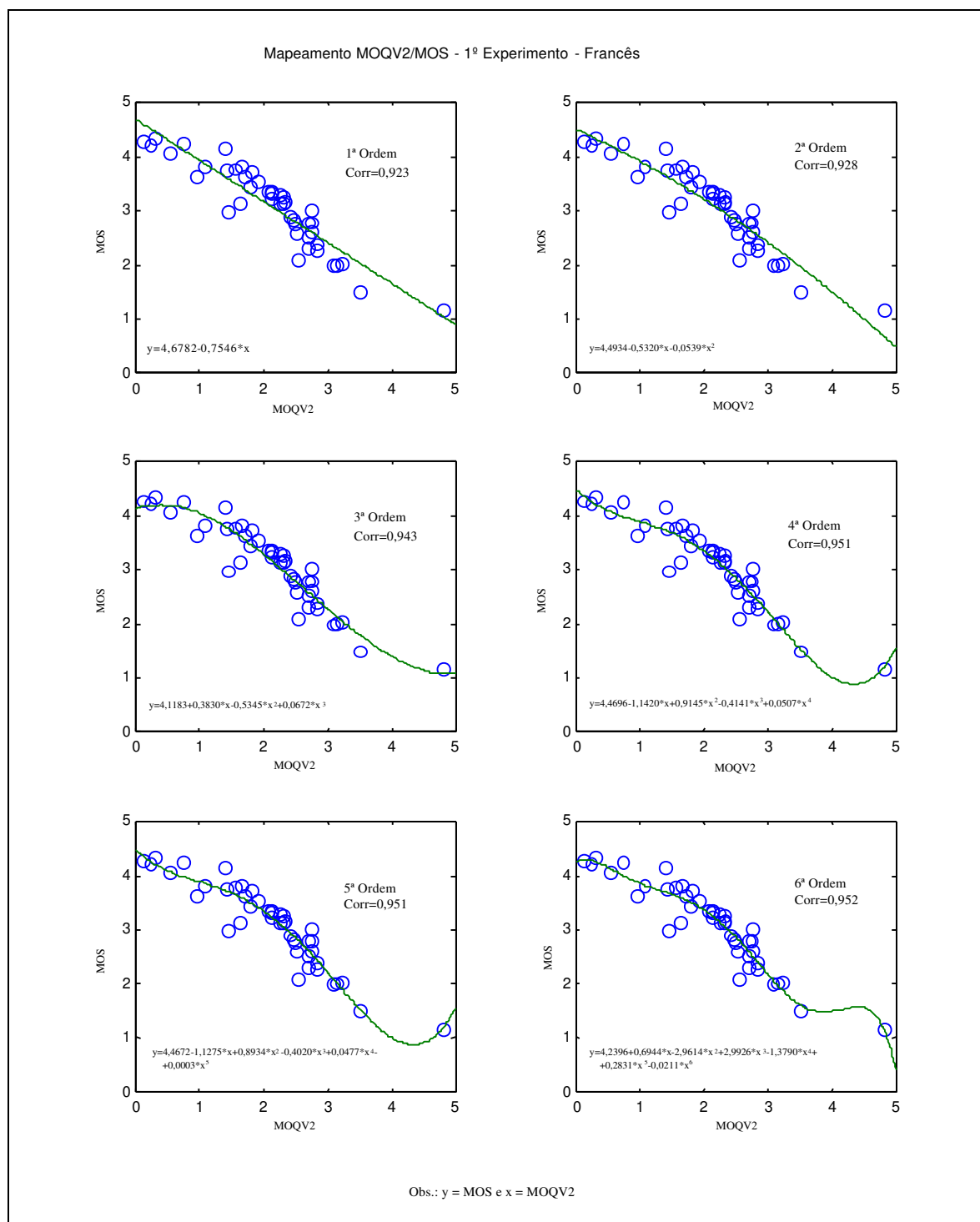


Figura 6.3 - Mapeamento MOQV2/MOS - 1º Experimento - Francês

Tabela 6.6 - Valores obtidos através de simulação - Experimento 1 - Japonês

<i>Laboratório d (NTT - Japonês)</i>											
Condição	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,895	1,813	1,345	1,548	1,650	1,966	1,858	1,399	1,589	1,703	3,52
2	2,452	2,643	2,210	2,188	2,373	2,509	2,824	2,377	2,248	2,490	2,84
3	2,887	2,929	2,495	2,734	2,761	2,988	3,105	2,791	2,862	2,937	2,30
4	0,954	0,947	0,760	1,008	0,917	0,953	0,946	0,765	1,006	0,918	3,39
5	2,254	2,081	1,567	2,027	1,982	2,252	2,074	1,565	2,016	1,977	2,69
6	1,559	1,646	1,170	1,257	1,408	1,565	1,648	1,166	1,254	1,408	3,19
7	0,312	0,239	0,196	0,234	0,245	0,314	0,239	0,198	0,236	0,247	3,56
8	2,002	2,461	1,318	1,340	1,780	2,024	2,502	1,366	1,372	1,816	3,28
9	2,089	2,642	1,876	1,559	2,042	2,178	2,699	1,977	1,592	2,112	3,27
10	2,919	2,816	2,788	2,679	2,801	3,072	3,060	2,891	2,815	2,960	2,69
11	1,810	2,181	1,384	1,608	1,746	1,859	2,233	1,452	1,660	1,801	3,17
12	2,211	2,345	1,692	1,819	2,017	2,253	2,393	1,752	1,864	2,066	2,91
13	2,661	3,177	1,852	2,072	2,441	2,793	3,247	2,042	2,135	2,554	2,64
14	2,798	2,905	2,292	2,617	2,653	2,929	3,066	2,447	2,755	2,799	2,94
15	3,208	3,292	2,812	2,653	2,991	3,453	3,548	3,000	2,818	3,205	2,24
16	1,996	1,886	1,595	1,812	1,822	2,047	1,927	1,647	1,850	1,868	3,29
17	1,961	2,096	1,972	1,902	1,983	1,994	2,174	2,024	1,932	2,031	3,19
18	2,373	2,577	2,333	2,366	2,412	2,454	2,635	2,431	2,486	2,502	2,63
19	2,620	2,896	2,442	2,019	2,494	2,733	3,041	2,609	2,112	2,624	2,78
20	2,870	3,222	2,875	2,833	2,950	3,011	3,448	3,104	3,025	3,147	2,28
21	3,060	3,441	2,575	2,508	2,896	3,291	3,527	2,762	2,610	3,048	2,28
22	2,825	3,457	2,796	3,026	3,026	2,956	3,725	3,036	3,239	3,239	2,38
23	3,535	3,625	3,138	3,197	3,374	3,815	3,904	3,362	3,443	3,631	1,79
24	2,856	2,933	2,003	2,066	2,465	2,971	3,004	2,128	2,152	2,564	2,52
25	2,697	3,532	2,217	2,371	2,704	2,824	3,657	2,335	2,495	2,828	2,32
26	2,834	2,832	2,564	2,977	2,802	2,971	3,073	2,688	3,123	2,964	2,33
27	3,232	3,050	3,109	3,044	3,109	3,414	3,196	3,109	3,199	3,270	2,33
28	3,174	3,508	3,371	3,049	3,276	3,357	3,705	3,675	3,234	3,493	1,75
29	2,382	2,645	2,222	2,008	2,314	2,439	2,713	2,302	2,117	2,393	2,75
30	2,497	2,875	2,390	2,440	2,551	2,576	2,988	2,555	2,592	2,678	2,63
31	2,642	3,197	2,340	1,972	2,538	2,728	3,277	2,481	2,070	2,639	2,53
32	3,364	3,318	3,036	2,955	3,168	3,568	3,556	3,284	3,163	3,393	2,32
33	3,322	3,497	2,734	2,746	3,075	3,513	3,660	3,008	2,841	3,256	2,07
34	2,989	3,407	2,668	3,037	3,025	3,232	3,552	2,846	3,216	3,212	2,60
35	3,504	3,490	3,961	3,191	3,537	3,921	3,871	4,342	3,592	3,932	1,70
36	3,497	3,532	3,180	3,167	3,344	3,650	3,684	3,516	3,269	3,530	1,73
37	6,361	5,783	4,085	3,772	5,000	6,366	5,823	4,119	3,814	5,031	1,10
38	4,376	4,375	2,183	3,383	3,579	4,388	4,397	2,193	3,387	3,591	1,42
39	2,973	3,118	1,391	1,484	2,242	2,976	3,125	1,405	1,495	2,250	2,24
40	2,070	1,852	0,785	1,142	1,462	2,075	1,852	0,792	1,145	1,466	3,02
41	1,102	1,079	0,370	0,536	0,772	1,112	1,085	0,371	0,538	0,777	3,51
42	0,673	0,511	0,251	0,360	0,449	0,674	0,512	0,253	0,363	0,451	3,72
43	0,365	0,300	0,174	0,241	0,270	0,367	0,301	0,176	0,243	0,272	3,68
44	0,114	0,113	0,117	0,226	0,143	0,116	0,114	0,120	0,228	0,145	3,80

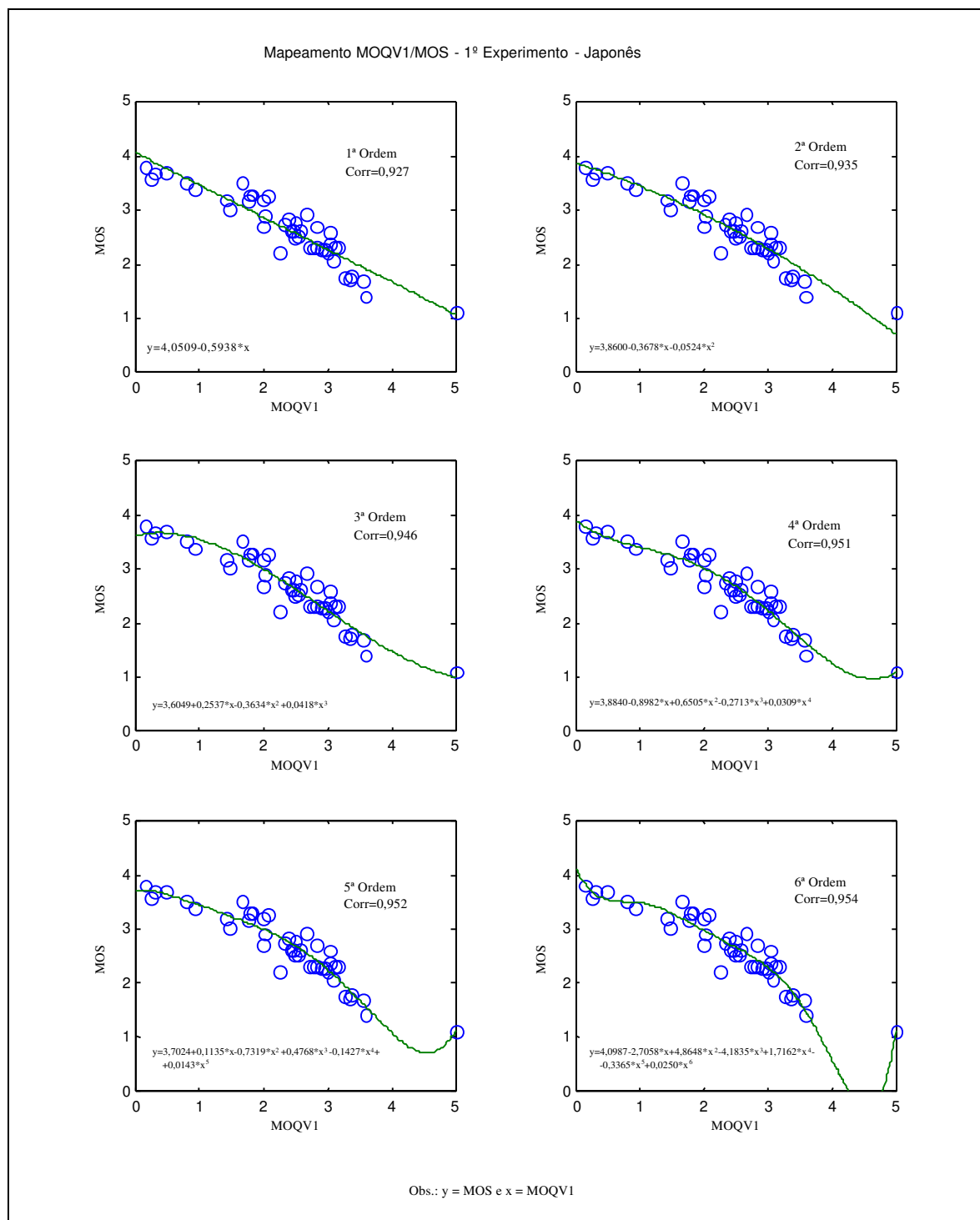


Figura 6.4 - Mapeamento MOQV1/MOS - 1º Experimento - Japonês

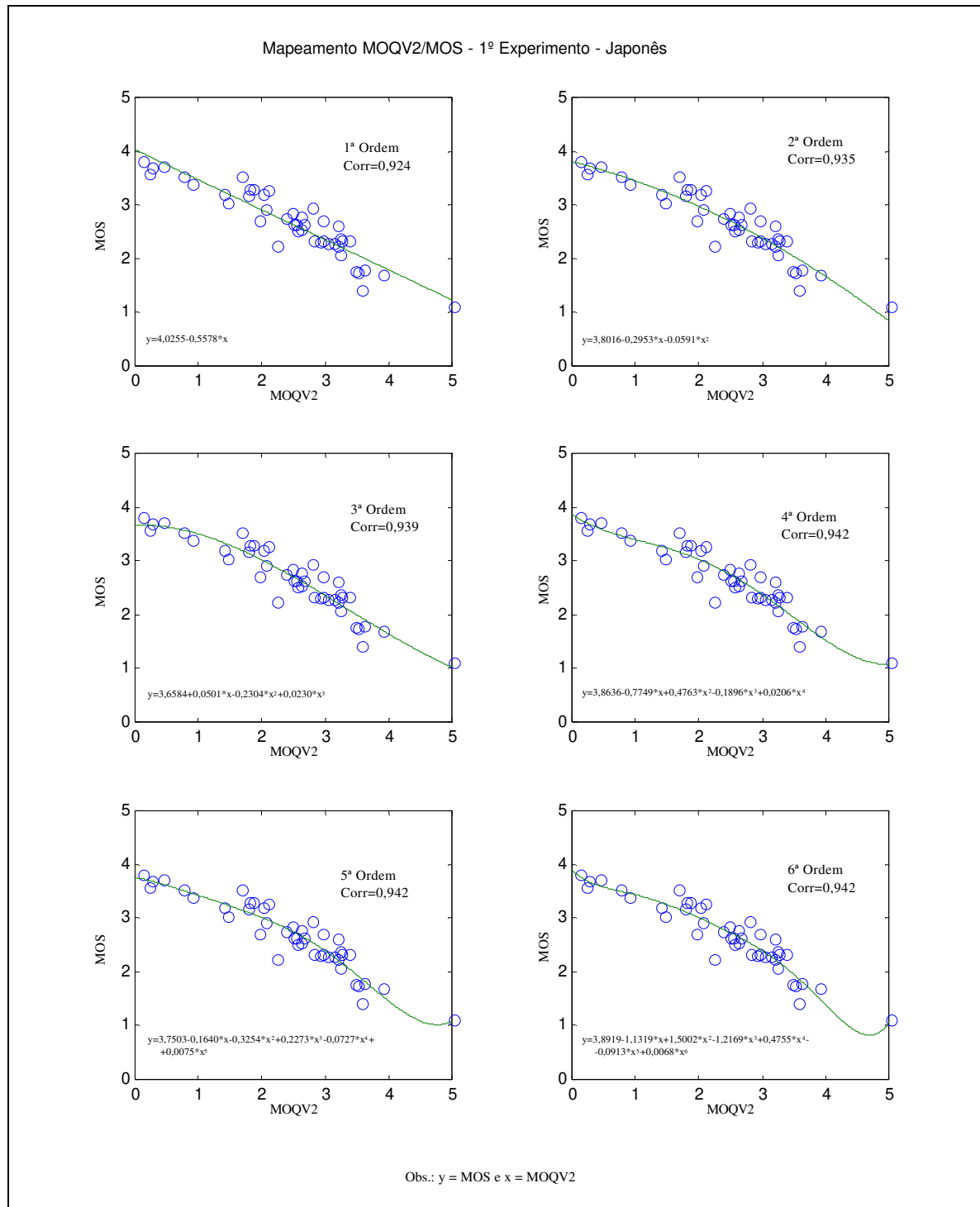


Figura 6.5 - Mapeamento MOQV2/MOS - 1º Experimento - Japonês

Tabela 6.7 - Valores obtidos através de simulação - Experimento 1 - Inglês

<i>Laboratório o (BNR - Inglês)</i>											
Condição	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,833	2,160	2,033	1,874	1,975	1,928	2,328	2,140	2,060	2,114	3,91
2	2,529	2,788	3,050	2,623	2,748	2,690	3,103	3,331	2,740	2,966	3,39
3	2,772	3,364	3,498	3,280	3,229	3,084	3,793	3,922	3,509	3,577	2,65
4	1,188	1,149	0,848	1,014	1,050	1,193	1,161	0,855	1,021	1,058	3,93
5	2,748	1,911	1,757	1,963	2,095	2,770	1,940	1,774	1,962	2,112	3,19
6	1,604	1,557	1,512	1,481	1,539	1,616	1,567	1,521	1,488	1,548	3,71
7	0,276	0,368	0,304	0,275	0,306	0,278	0,385	0,306	0,277	0,312	4,31
8	2,078	1,896	1,828	1,730	1,883	2,219	2,015	1,885	1,835	1,989	3,52
9	2,280	2,230	2,271	2,111	2,223	2,375	2,286	2,358	2,206	2,306	3,67
10	2,965	3,013	2,996	2,886	2,965	3,227	3,422	3,181	3,071	3,225	2,90
11	2,270	2,049	2,220	1,822	2,090	2,347	2,119	2,390	1,899	2,189	3,74
12	2,033	2,261	2,356	1,924	2,144	2,117	2,526	2,578	1,973	2,299	3,53
13	2,561	2,716	2,827	2,491	2,649	2,804	2,908	3,197	2,801	2,928	3,23
14	2,784	2,886	2,747	2,808	2,806	2,968	3,297	3,084	3,113	3,116	3,21
15	3,191	3,584	3,549	3,336	3,415	3,607	4,245	4,018	3,867	3,934	2,34
16	1,873	2,396	2,303	1,861	2,108	1,931	2,511	2,552	1,904	2,225	3,48
17	2,071	2,481	2,386	2,219	2,289	2,165	2,717	2,552	2,375	2,452	3,61
18	2,512	2,911	2,684	2,563	2,668	2,659	3,288	2,844	2,727	2,880	3,42
19	2,691	2,640	2,881	2,719	2,733	2,823	2,781	3,184	2,879	2,917	3,28
20	3,405	3,500	3,617	3,115	3,411	3,873	4,049	4,201	3,397	3,882	2,29
21	2,856	3,278	3,161	2,979	3,069	3,128	3,759	3,491	3,470	3,462	2,68
22	2,964	3,275	3,509	3,329	3,267	3,292	3,744	4,140	3,742	3,725	2,45
23	3,306	3,693	4,012	3,675	3,672	3,820	4,302	4,533	4,143	4,200	2,02
24	3,105	2,700	2,705	2,564	2,769	3,240	2,799	2,995	2,727	2,940	3,17
25	3,180	3,135	3,060	2,755	3,033	3,519	3,282	3,331	3,185	3,329	2,88
26	2,736	3,131	3,262	3,315	3,111	2,989	3,304	3,531	3,701	3,381	2,91
27	3,169	2,986	3,332	3,316	3,201	3,351	3,508	3,643	3,510	3,503	2,50
28	3,912	3,726	4,034	4,012	3,924	4,359	4,340	4,404	4,333	4,359	1,71
29	2,892	2,554	2,565	2,417	2,607	3,169	2,686	2,742	2,617	2,804	3,25
30	2,744	2,587	2,880	2,835	2,762	2,872	2,667	3,046	3,253	2,960	3,14
31	2,887	2,505	2,753	2,543	2,672	3,131	2,614	2,910	2,768	2,856	3,11
32	3,228	3,421	3,739	3,273	3,415	3,545	3,882	4,186	3,673	3,822	2,21
33	3,122	3,401	3,402	3,252	3,294	4,043	3,724	3,751	3,647	3,791	1,91
34	3,195	3,334	3,173	3,328	3,258	3,493	3,667	3,502	3,626	3,572	2,57
35	3,525	4,089	4,122	3,743	3,870	4,241	4,777	4,769	4,337	4,531	1,82
36	3,383	3,905	3,606	3,406	3,575	3,829	4,495	4,570	3,875	4,192	1,78
37	5,935	5,821	5,126	5,370	5,563	5,985	5,947	5,186	5,407	5,631	1,31
38	4,357	3,390	3,000	4,002	3,687	4,392	3,421	3,012	3,996	3,705	2,11
39	3,580	2,445	1,912	2,380	2,579	3,585	2,507	1,931	2,380	2,601	2,95
40	1,557	1,708	1,114	1,770	1,537	1,559	1,738	1,117	1,763	1,544	3,46
41	1,024	0,773	0,733	0,821	0,838	1,027	0,777	0,735	0,826	0,841	4,17
42	0,592	0,587	0,475	0,560	0,554	0,592	0,589	0,484	0,561	0,557	4,25
43	0,340	0,388	0,295	0,331	0,339	0,341	0,392	0,302	0,343	0,345	4,24
44	0,104	0,187	0,226	0,190	0,177	0,105	0,190	0,236	0,193	0,181	4,32

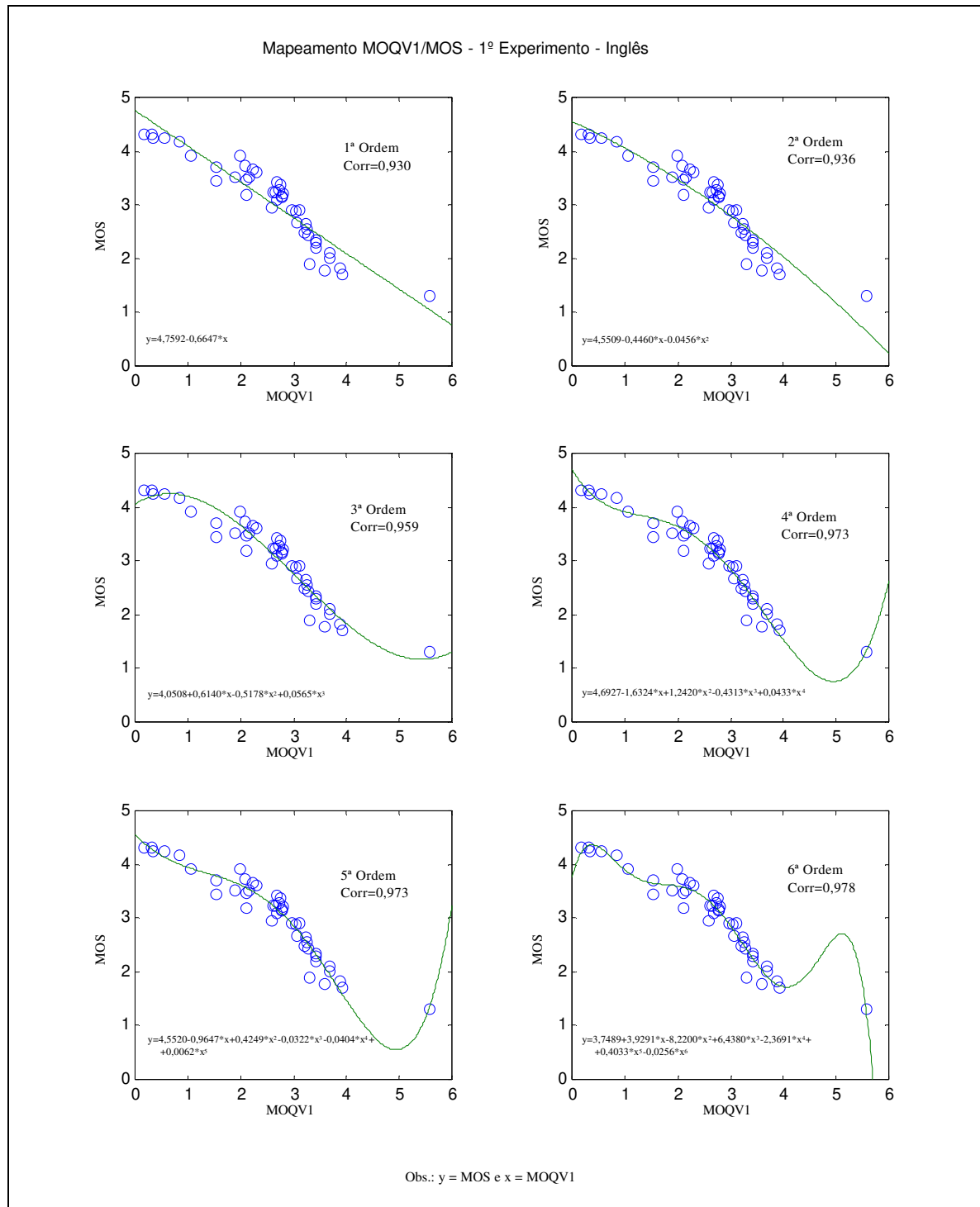


Figura 6.6 - Mapeamento MOQV1/MOS - 1º Experimento - Inglês

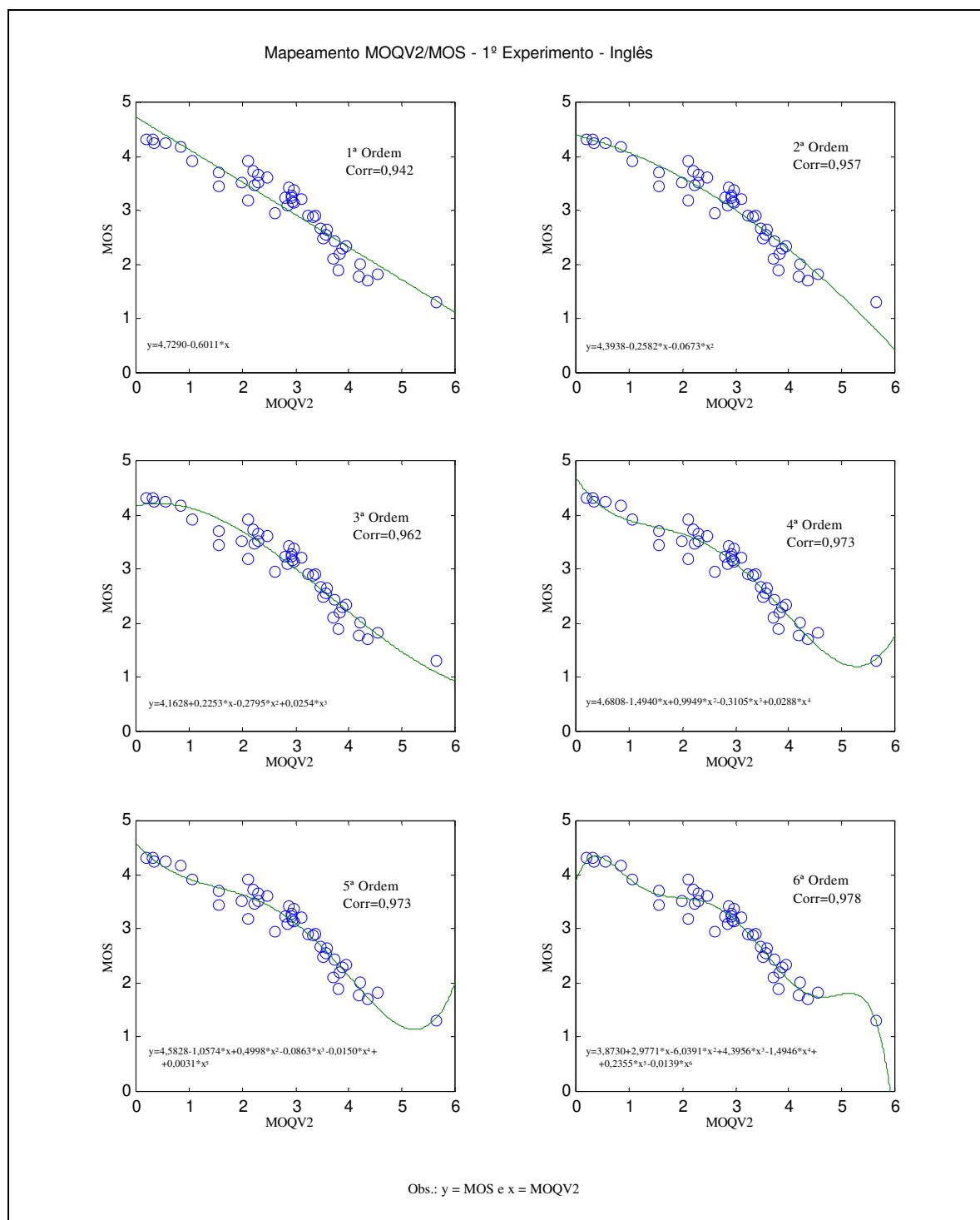


Figura 6.7 - Mapeamento MOQV2/MOS - 1º Experimento - Inglês

Experimento 2

As tabelas 6.8 a 6.10 e as figuras 6.8 a 6.13 mostram um comportamento similar àquele observado no primeiro experimento, porém com valores de correlação mais elevados, revelando que o programa MOQV tem um bom desempenho nas situações em que há a presença de ruído ambiente.

Tabela 6.8 - Valores obtidos através de simulação - Experimento 2 - Francês

<i>Laboratório a (CNET - Francês)</i>											
	MOQV1					MOQV2					CMOS
Condição	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,524	1,418	1,277	1,048	1,317	1,553	1,464	1,324	1,088	1,357	-0,58
2	2,146	2,309	1,838	1,782	2,019	2,271	2,390	1,981	1,901	2,136	-0,64
3	2,178	2,171	1,989	1,902	2,060	2,288	2,263	2,096	2,016	2,166	-0,56
4	2,564	2,557	2,466	2,316	2,476	2,706	2,656	2,598	2,414	2,594	-0,55
5	2,241	2,206	2,180	1,978	2,151	2,353	2,297	2,297	2,091	2,260	-0,60
6	2,407	2,334	1,886	1,989	2,154	2,526	2,461	2,067	2,152	2,302	-0,56
7	1,947	1,890	1,716	1,600	1,788	1,980	1,931	1,777	1,648	1,834	-0,87
8	2,039	1,961	1,774	1,536	1,828	2,102	2,044	1,872	1,619	1,909	-1,35
9	2,693	2,932	2,469	2,418	2,628	3,006	3,131	2,828	2,694	2,915	-1,38
10	2,789	2,776	2,635	2,616	2,704	3,076	3,008	2,900	2,906	2,973	-1,57
11	3,367	3,327	3,284	3,168	3,287	3,714	3,618	3,642	3,424	3,600	-1,41
12	2,846	2,892	2,909	2,669	2,829	3,132	3,116	3,194	2,935	3,094	-1,61
13	2,961	2,945	2,593	2,726	2,806	3,238	3,223	3,006	3,094	3,140	-1,33
14	2,490	2,426	2,325	2,198	2,360	2,566	2,505	2,457	2,295	2,456	-1,55
15	1,027	0,810	0,462	0,540	0,710	1,027	0,810	0,461	0,540	0,710	-0,28
16	0,538	0,585	0,389	0,524	0,509	0,538	0,584	0,389	0,523	0,509	-0,06
17	0,619	0,610	0,546	0,550	0,581	0,619	0,610	0,546	0,551	0,582	-0,04
18	0,445	0,382	0,382	0,491	0,425	0,447	0,382	0,382	0,492	0,426	-0,06
19	0,498	0,597	0,574	0,524	0,548	0,498	0,597	0,574	0,524	0,548	-0,05
20	0,581	0,632	0,449	0,478	0,535	0,581	0,632	0,451	0,478	0,536	-0,01
21	0,808	0,697	0,554	0,524	0,646	0,809	0,697	0,554	0,524	0,646	-0,07
22	1,639	1,272	0,823	0,861	1,149	1,642	1,267	0,819	0,858	1,147	-0,45
23	0,930	0,975	0,633	0,879	0,854	0,929	0,975	0,630	0,875	0,852	-0,06
24	1,035	0,983	0,949	0,893	0,965	1,035	0,981	0,947	0,894	0,964	-0,10
25	0,738	0,681	0,635	0,786	0,710	0,740	0,680	0,634	0,785	0,710	-0,11
26	0,811	1,001	0,887	0,846	0,886	0,811	1,000	0,887	0,844	0,886	-0,06
27	0,953	1,041	0,779	0,759	0,883	0,952	1,039	0,779	0,759	0,882	-0,05
28	1,381	1,162	0,953	0,874	1,093	1,382	1,158	0,951	0,872	1,091	-0,47
29	5,744		3,760		4,752	5,782		3,815		4,799	-2,67
30	7,994		6,368		7,181	36,372		14,753		25,563	-1,96
31	7,591		5,993		6,792	21,453		8,210		14,832	-1,19
32	7,555		6,276		6,916	30,979		11,119		21,049	-0,40
33	7,846		6,771		7,309	37,785		13,072		25,429	-0,10
34	0,024		0,034		0,029	0,024		0,036		0,030	0,01
35	0,011		0,014		0,013	0,011		0,014		0,013	-0,07
36	0,020		0,048		0,034	0,020		0,049		0,035	0,03
37	0,023		0,020		0,022	0,024		0,021		0,023	-0,01
38	0,009		0,038		0,024	0,009		0,038		0,024	-0,13
39	0,030		0,047		0,039	0,030		0,048		0,039	-0,02
40	0,011		0,041		0,026	0,011		0,042		0,027	-0,01

Obs.: as condições 30 a 33 (as quais aparecem hachuradas na tabela) não foram consideradas, uma vez que os arquivos originais e decodificados não contêm a mesma sentença, causando discrepância nos resultados obtidos.

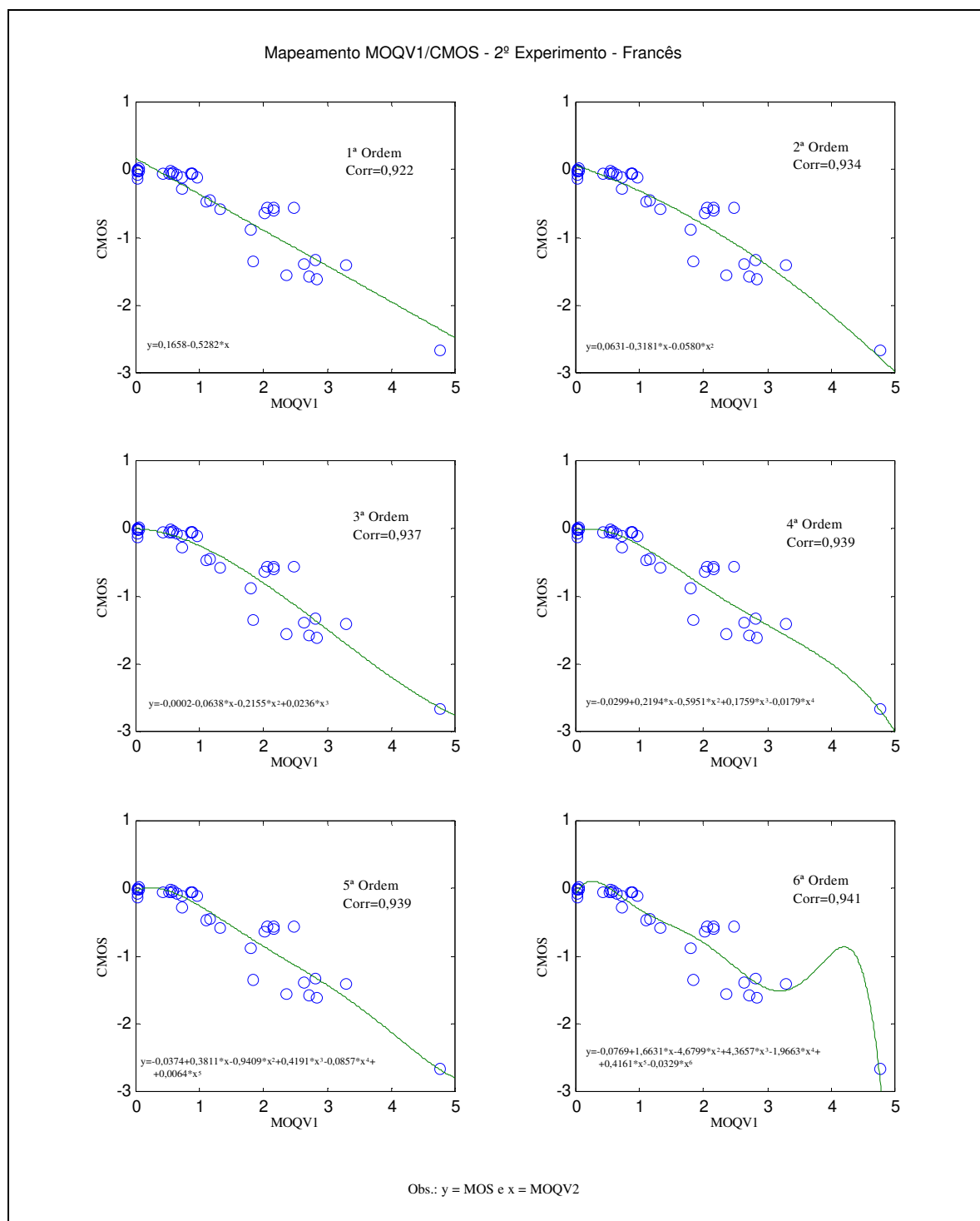


Figura 6.8 - Mapeamento MOQV1/CMOS - 2º Experimento - Francês

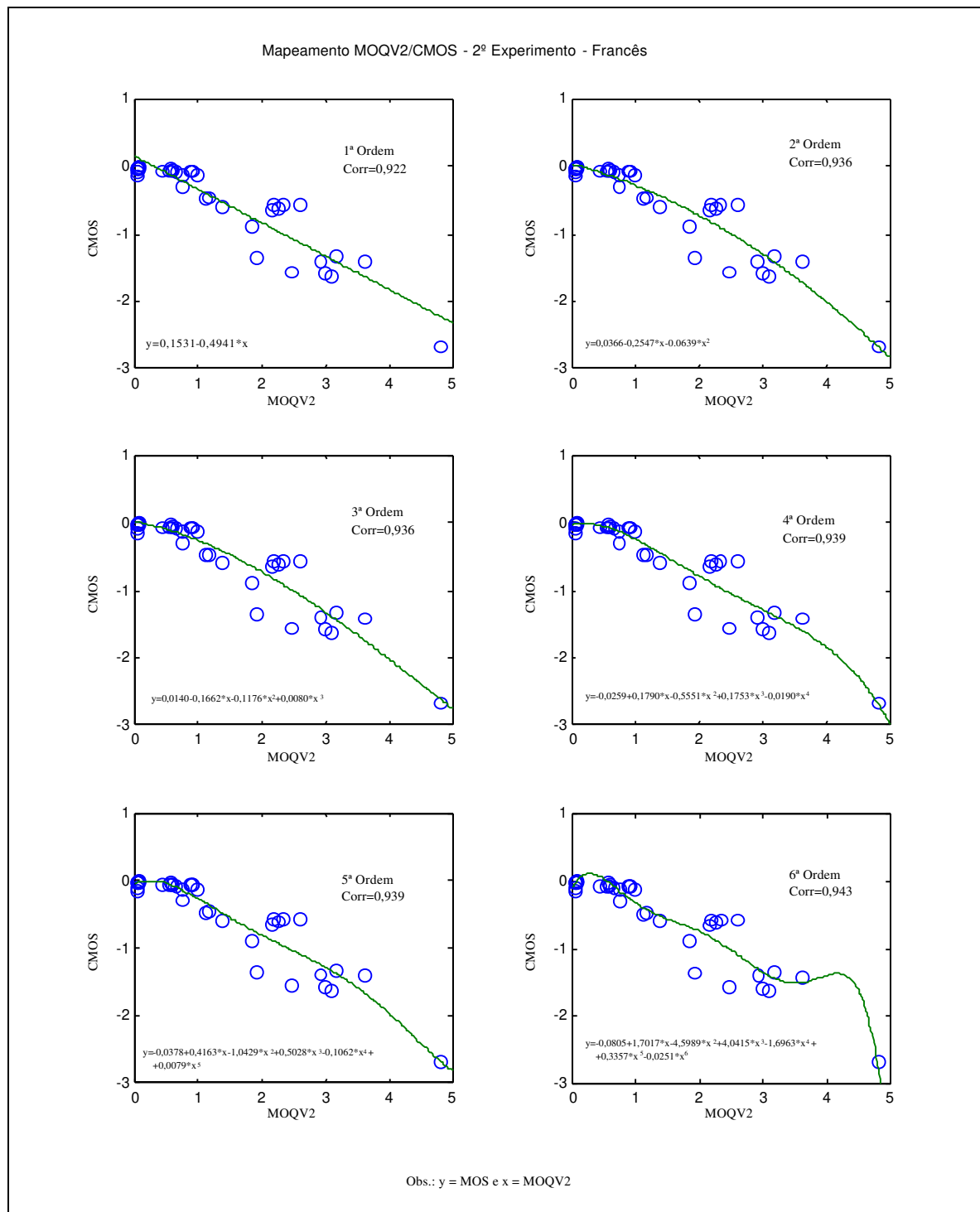


Figura 6.9 - Mapeamento MOQV2/CMOS - 2º Experimento - Francês

Tabela 6.9 - Valores obtidos através de simulação - Experimento 2 - Japonês

<i>Laboratório d (NTT - Japão)</i>											
Condição	MOQV1					MOQV2					CMOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,990	2,002	1,581	1,607	1,795	2,034	2,060	1,671	1,654	1,855	-0,44
2	2,280	2,250	1,790	1,897	2,054	2,380	2,355	1,892	2,000	2,157	-0,65
3	2,387	2,306	1,832	1,917	2,111	2,465	2,382	1,948	2,010	2,201	-0,80
4	2,654	2,653	2,237	2,432	2,494	2,741	2,736	2,354	2,529	2,590	-0,80
5	2,383	2,411	2,021	2,021	2,209	2,460	2,496	2,117	2,125	2,300	-0,60
6	2,411	2,402	1,873	1,959	2,161	2,541	2,501	2,018	2,104	2,291	-0,41
7	2,096	2,123	1,761	1,718	1,925	2,125	2,167	1,815	1,764	1,968	-0,49
8	2,623	2,544	2,283	2,234	2,421	2,734	2,659	2,459	2,332	2,546	-0,84
9	2,927	2,869	2,602	2,634	2,758	3,167	3,129	2,852	2,882	3,008	-1,55
10	3,005	2,914	2,509	2,642	2,768	3,200	3,101	2,793	2,882	2,994	-1,20
11	3,504	3,477	3,012	3,358	3,338	3,761	3,680	3,338	3,624	3,601	-1,23
12	3,007	3,109	2,799	2,743	2,915	3,199	3,305	3,034	3,011	3,137	-1,34
13	3,078	3,060	2,553	2,653	2,836	3,369	3,288	2,897	2,982	3,134	-1,01
14	2,643	2,697	2,485	2,400	2,556	2,707	2,780	2,611	2,499	2,649	-1,66
15	0,925	1,127	0,856	1,186	1,024	0,925	1,127	0,862	1,187	1,025	-0,19
16	0,700	0,703	0,619	0,814	0,709	0,700	0,703	0,619	0,814	0,709	0,11
17	0,579	0,764	0,652	0,818	0,703	0,580	0,763	0,651	0,817	0,703	-0,07
18	0,472	0,505	0,442	0,598	0,504	0,472	0,505	0,442	0,599	0,505	0,11
19	0,600	0,694	0,637	0,743	0,669	0,600	0,694	0,637	0,742	0,668	0,17
20	0,576	0,764	0,577	0,728	0,661	0,576	0,764	0,578	0,728	0,662	-0,10
21	0,794	0,967	0,706	0,942	0,852	0,794	0,968	0,705	0,941	0,852	-0,15
22	1,389	1,717	1,233	1,616	1,489	1,389	1,715	1,237	1,614	1,489	-0,23
23	1,013	1,022	0,924	1,124	1,021	1,011	1,021	0,922	1,122	1,019	-0,16
24	0,932	1,122	0,954	1,137	1,036	0,932	1,121	0,952	1,134	1,035	-0,05
25	0,709	0,800	0,689	0,876	0,769	0,709	0,799	0,688	0,875	0,768	-0,03
26	0,913	1,050	0,941	1,065	0,992	0,913	1,050	0,942	1,064	0,992	0,09
27	0,876	1,118	0,870	1,071	0,984	0,875	1,118	0,871	1,070	0,984	-0,04
28	1,219	1,436	1,054	1,370	1,270	1,218	1,437	1,050	1,367	1,268	-0,21
29	6,388		3,712		5,050	6,398		3,787		5,093	-2,17
30	4,716		2,274		3,495	4,715		2,306		3,511	-1,50
31	2,961		1,211		2,086	2,960		1,227		2,094	-0,84
32	1,553		0,604		1,079	1,552		0,615		1,084	-0,31
33	0,714		0,323		0,519	0,714		0,333		0,524	0,13
34	0,040		0,077		0,059	0,040		0,084		0,062	0,31
35	0,018		0,022		0,020	0,018		0,022		0,020	0,11
36	0,019		0,020		0,020	0,020		0,020		0,020	0,09
37	0,015		0,019		0,017	0,015		0,019		0,017	0,01
38	0,016		0,025		0,021	0,016		0,025		0,021	0,03
39	0,034		0,037		0,036	0,034		0,037		0,036	-0,01
40	0,015		0,026		0,021	0,015		0,026		0,021	-0,03

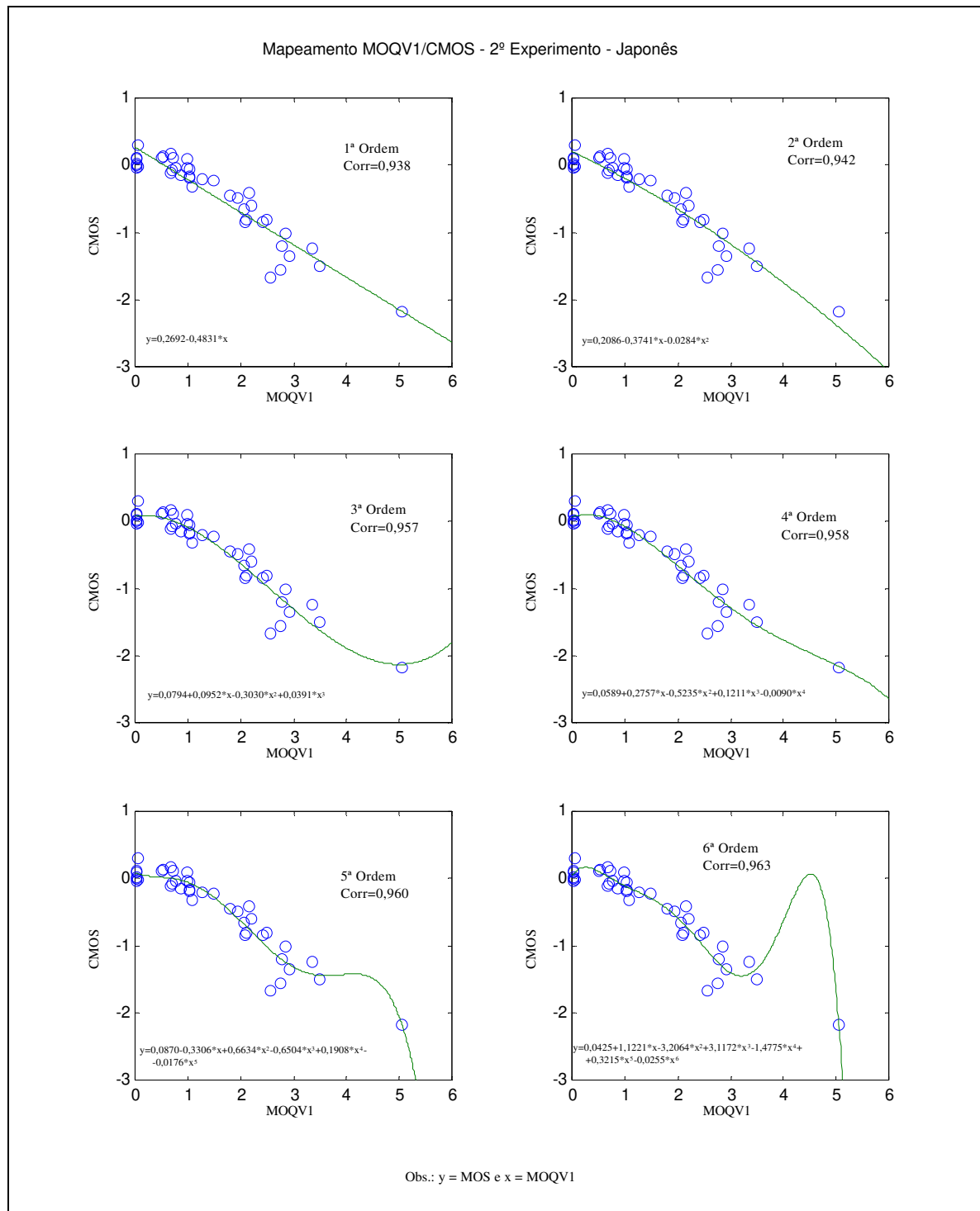


Figura 6.10 - Mapeamento MOQV1/CMOS - 2º Experimento - Japonês

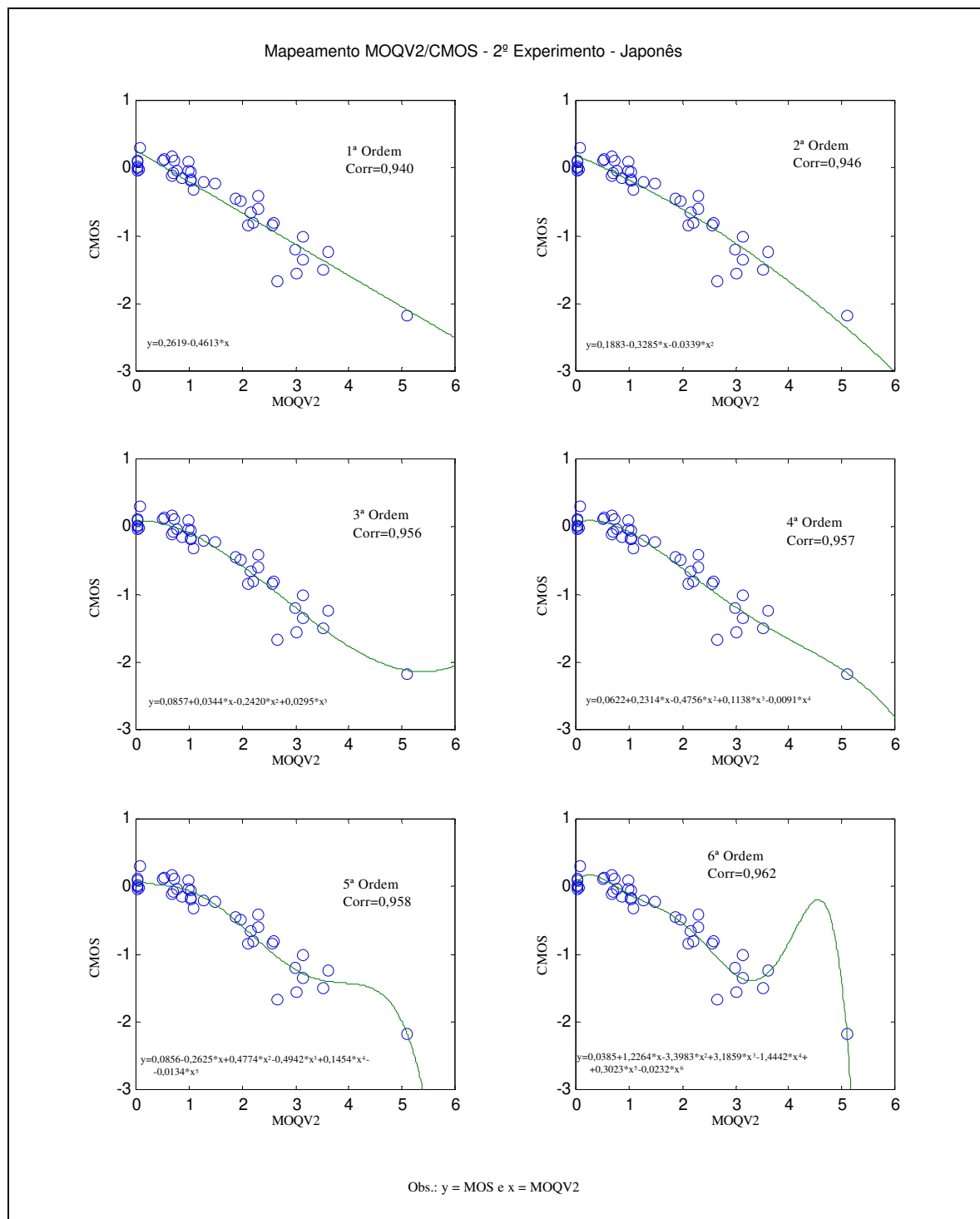


Figura 6.11 - Mapeamento MOQV2/CMOS - 2º Experimento – Japonês

Tabela 6.10 - Valores obtidos através de simulação - Experimento 2 - Inglês

<i>Laboratório e (AT&T - Inglês)</i>											
Condição	MOQV1					MOQV2					CMOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	2,002	1,882	2,089	1,748	1,930	2,114	2,019	2,191	1,805	2,032	-0,40
2	2,198	2,104	2,223	1,875	2,100	2,285	2,201	2,289	1,986	2,190	-0,51
3	2,229	2,140	2,155	2,010	2,134	2,324	2,238	2,249	2,130	2,235	-0,56
4	2,588	2,611	2,528	2,272	2,500	2,710	2,728	2,672	2,406	2,629	-0,44
5	2,370	2,252	2,335	2,157	2,279	2,511	2,419	2,478	2,271	2,420	-0,78
6	2,191	2,200	2,150	1,950	2,123	2,370	2,313	2,266	2,131	2,270	-0,53
7	1,975	1,836	2,219	1,611	1,910	2,025	1,902	2,311	1,673	1,978	-0,71
8	2,598	2,485	2,783	2,448	2,579	2,816	2,746	2,997	2,559	2,780	-1,08
9	2,818	2,714	2,949	2,663	2,786	3,036	2,930	3,130	2,920	3,004	-1,27
10	2,832	2,774	2,982	2,652	2,810	3,060	3,009	3,214	2,964	3,062	-1,33
11	3,455	3,454	3,390	3,080	3,345	3,768	3,734	3,762	3,433	3,674	-1,25
12	3,046	2,904	3,054	2,825	2,957	3,376	3,279	3,369	3,104	3,282	-1,40
13	2,845	2,836	2,855	2,568	2,776	3,254	3,110	3,118	2,980	3,116	-1,24
14	2,633	2,363	2,778	2,302	2,519	2,744	2,505	2,995	2,461	2,676	-1,47
15	1,206	1,051	1,138	1,232	1,157	1,208	1,058	1,149	1,231	1,162	-0,15
16	0,737	0,742	0,839	0,646	0,741	0,737	0,743	0,840	0,647	0,742	-0,02
17	0,759	0,770	0,695	0,600	0,706	0,764	0,773	0,696	0,602	0,709	0,07
18	0,539	0,551	0,495	0,532	0,529	0,541	0,551	0,496	0,532	0,530	-0,03
19	0,672	0,706	0,598	0,706	0,671	0,679	0,714	0,598	0,707	0,675	-0,05
20	0,710	0,748	0,867	0,576	0,725	0,716	0,750	0,868	0,578	0,728	-0,04
21	0,880	0,989	1,112	0,790	0,943	0,881	0,993	1,114	0,792	0,945	-0,13
22	1,708	1,583	1,556	1,721	1,642	1,712	1,592	1,566	1,714	1,646	-0,45
23	1,178	1,111	1,206	1,027	1,131	1,177	1,111	1,205	1,027	1,130	-0,04
24	1,172	1,201	1,009	0,924	1,077	1,180	1,205	1,009	0,923	1,079	-0,06
25	0,856	0,920	0,763	0,840	0,845	0,859	0,920	0,765	0,840	0,846	-0,11
26	1,074	1,157	0,937	1,049	1,054	1,080	1,169	0,936	1,048	1,058	-0,02
27	1,028	1,157	1,207	0,833	1,056	1,035	1,158	1,208	0,833	1,059	-0,04
28	1,373	1,500	1,514	1,157	1,386	1,373	1,507	1,516	1,157	1,388	-0,35
29	5,611		4,595		5,103	5,624		4,646		5,135	-2,36
30	3,860		2,884		3,372	3,861		2,909		3,385	-1,77
31	2,248		1,561		1,905	2,249		1,576		1,913	-0,85
32	1,153		0,802		0,978	1,156		0,816		0,986	-0,37
33	0,586		0,459		0,523	0,591		0,473		0,532	-0,10
34	0,126		0,151		0,139	0,131		0,161		0,146	-0,15
35	0,048		0,086		0,067	0,048		0,087		0,068	0,01
36	0,056		0,068		0,062	0,061		0,069		0,065	-0,13
37	0,058		0,071		0,065	0,059		0,072		0,066	-0,07
38	0,078		0,078		0,078	0,079		0,079		0,079	0,13
39	0,087		0,091		0,089	0,091		0,091		0,091	0,02
40	0,039		0,094		0,067	0,040		0,095		0,068	0,00

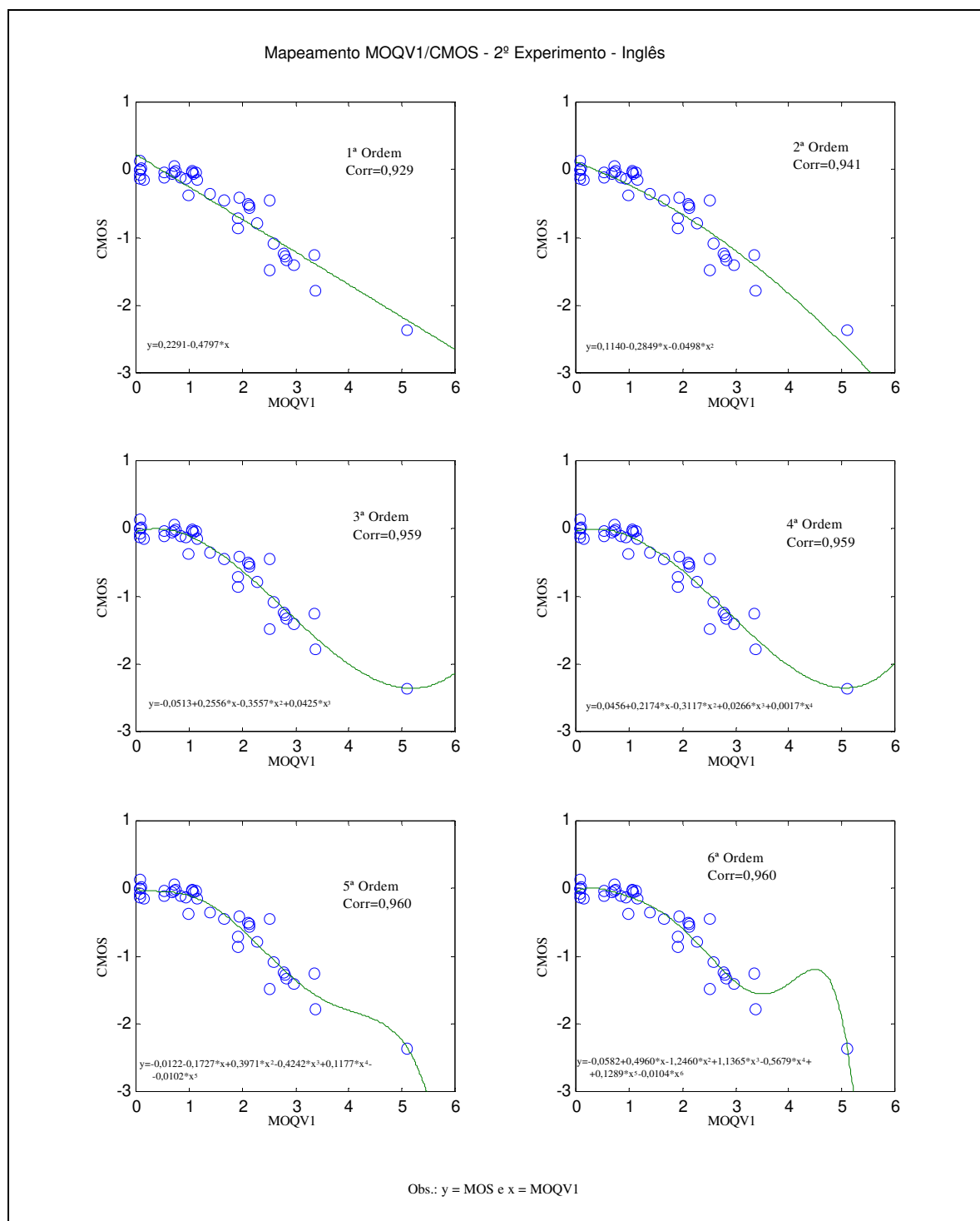


Figura 6.12 - Mapeamento MOQV1/CMOS - 2º Experimento - Inglês

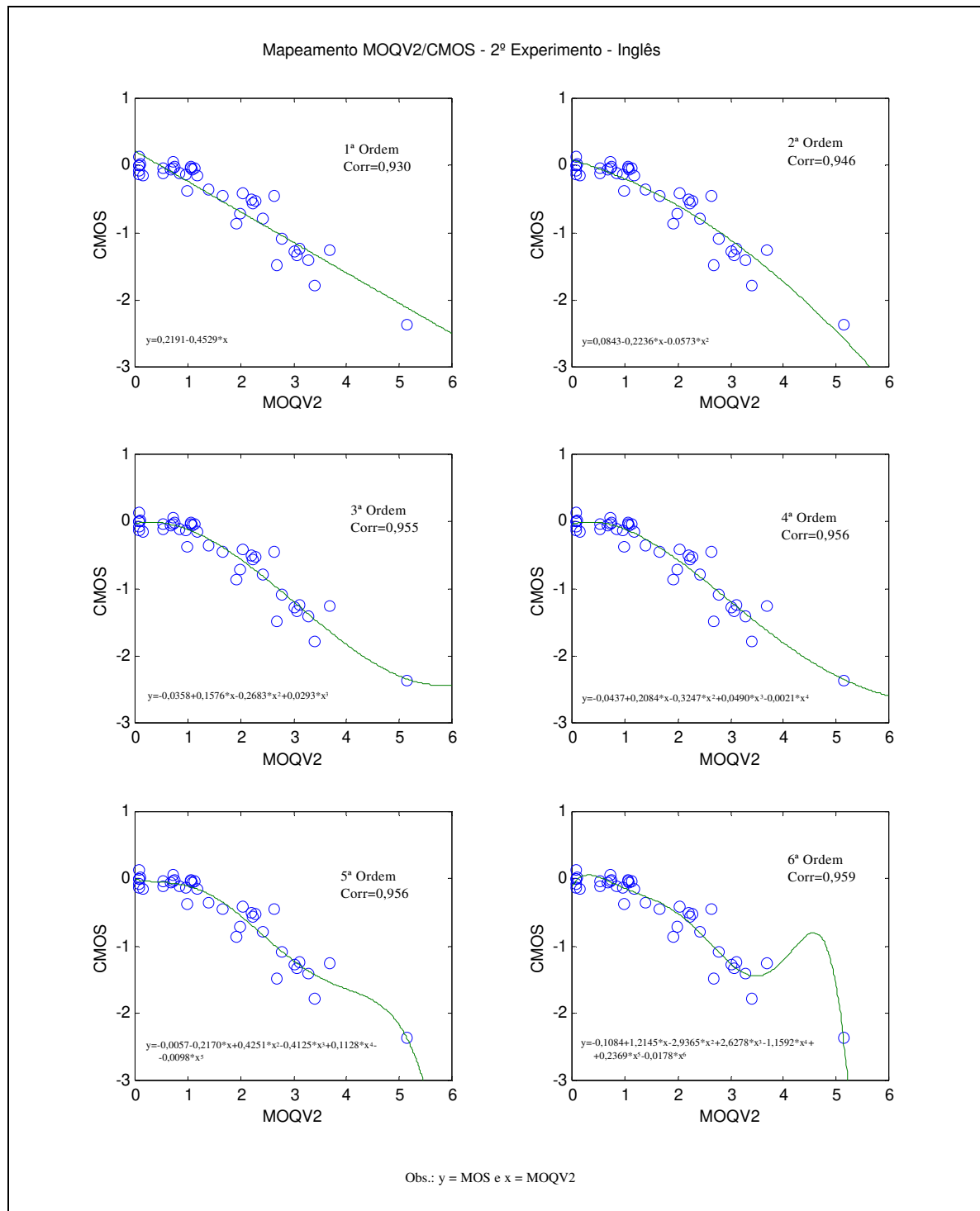


Figura 6.13 - Mapeamento MOQV2/CMOS - 2º Experimento - Inglês

Experimento 3

Os resultados obtidos para este experimento foram muito ruins para todos os métodos, línguas e mapeamentos testados, como se pode observar pelas tabelas 6.11 a 6.14 e figuras 6.14 a 6.21. Isto torna desaconselhável o uso dos métodos aqui testados em situações em que o sinal é corrompido pela presença de erros.

Tabela 6.11 - Valores obtidos através de simulação - Experimento 3 - Francês

EXPERIMENTO 3											
<i>Laboratório a (CNET - Francês)</i>											
Situação	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,538	1,426	1,294	1,068	1,332	1,573	1,474	1,343	1,111	1,375	4,21
2	1,545	1,893	1,205	1,408	1,513	1,710	1,968	1,238	1,602	1,630	3,86
3	1,788	1,950	1,751	1,571	1,765	2,046	2,172	1,907	1,769	1,974	3,58
4	1,658	1,454	1,448	1,678	1,559	1,996	1,738	1,566	2,656	1,989	3,88
5	2,169	1,796	2,202	1,719	1,971	3,752	2,849	3,706	3,066	3,343	3,15
6	2,237	2,202	1,878	1,871	2,047	2,322	2,288	2,000	1,983	2,148	3,51
7	2,415	2,189	2,134	2,036	2,194	2,546	2,332	2,349	2,209	2,359	3,15
8	2,485	2,226	2,262	2,181	2,289	2,727	2,406	2,459	2,570	2,541	2,84
9	2,432	2,295	2,062	2,131	2,230	3,004	2,601	2,174	2,411	2,548	2,93
10	2,335	2,194	2,144	1,993	2,167	3,223	2,380	2,431	2,757	2,698	2,76
11	2,243	2,225	2,084	1,941	2,123	2,347	2,318	2,211	2,058	2,233	3,49
12	2,605	2,471	2,290	2,224	2,397	2,724	2,607	2,504	2,458	2,573	2,91
13	2,762	2,550	2,326	2,114	2,438	3,442	2,813	2,539	2,296	2,772	2,77
14	2,418	2,358	2,353	2,146	2,319	2,534	2,460	3,127	2,541	2,665	3,10
15	2,485	2,510	2,444	2,311	2,437	3,026	3,179	3,536	3,304	3,261	2,23
16	2,407	2,383	2,078	1,898	2,191	2,540	2,526	2,263	2,087	2,354	3,45
17	2,428	2,537	2,242	2,021	2,307	2,695	2,745	2,509	2,277	2,557	2,91
18	2,616	2,476	2,352	2,142	2,397	2,880	2,741	2,579	2,450	2,663	2,74
19	2,647	2,471	2,254	2,025	2,349	3,549	2,966	3,173	2,432	3,030	2,91
20	2,706	2,437	2,288	2,101	2,383	3,629	3,474	3,303	2,315	3,180	2,69
21	2,013	2,057	1,827	1,666	1,891	2,084	2,112	1,941	1,726	1,966	3,75
22	2,387	2,181	2,768	1,799	2,283	2,641	2,324	2,987	1,934	2,471	3,36
23	2,266	1,784	2,244	2,140	2,108	2,509	1,945	2,549	2,547	2,387	3,16
24	2,595	3,366	2,935	2,479	2,844	2,769	3,855	3,340	3,125	3,272	2,30
25	2,659	2,472	2,054	2,122	2,327	3,716	3,308	2,245	2,340	2,902	3,13
26	2,690	2,672	2,285	2,293	2,485	3,313	3,609	2,681	2,743	3,086	2,82
27	2,981	2,990	2,985	2,600	2,889	3,416	3,286	3,391	2,886	3,245	2,64
28	3,029	3,302	3,049	2,894	3,069	3,979	5,407	4,570	4,592	4,637	1,92
29	1,987	1,997	1,564	1,344	1,723	2,098	2,136	1,750	1,455	1,860	3,46
30	2,226	2,295	2,111	1,582	2,054	2,438	2,482	2,295	1,656	2,218	2,88
31	2,702	2,528	2,005	1,622	2,214	3,167	2,728	2,200	1,702	2,449	2,39
32	3,471	3,105	2,276	1,879	2,683	4,624	3,808	2,290	1,919	3,160	1,59
33	1,893	1,809	1,510	1,340	1,638	2,376	1,984	1,902	1,401	1,916	3,74
34	1,922	2,114	2,455	1,381	1,968	2,695	2,239	3,473	1,421	2,457	3,39
35	2,454	2,128	2,209	1,760	2,138	3,166	3,926	2,807	2,085	2,996	2,68
36	2,356	2,473	1,990	2,036	2,214	2,840	2,978	2,262	2,319	2,600	2,64
37	0,831	0,948	0,443	0,702	0,731	0,830	0,947	0,443	0,700	0,730	4,27
38	0,608	0,515	0,468	0,599	0,548	0,608	0,514	0,469	0,599	0,548	3,61
39	0,607	0,481	0,473	0,517	0,519	0,607	0,481	0,472	0,517	0,519	3,79
40	0,610	0,574	0,471	0,471	0,532	0,610	0,574	0,472	0,471	0,532	3,43
41	4,421	3,714	3,340	2,776	3,563	4,432	3,731	3,348	2,780	3,573	2,16
42	3,775	2,858	1,768	1,598	2,500	3,780	2,863	1,776	1,601	2,505	2,77
43	2,204	1,785	1,151	0,871	1,503	2,212	1,790	1,150	0,873	1,506	3,66
44	1,050	1,249	0,615	0,491	0,851	1,054	1,247	0,618	0,492	0,853	4,19
45	0,557	0,651	0,266	0,255	0,432	0,558	0,651	0,267	0,257	0,433	4,47
46	0,077	0,116	0,103	0,031	0,082	0,085	0,116	0,107	0,032	0,085	4,68
47	0,029	0,049	0,074	0,009	0,040	0,036	0,050	0,077	0,010	0,043	4,69
48	0,005	0,019	0,028	0,003	0,014	0,005	0,019	0,029	0,003	0,014	3,71
49	0,010	0,026	0,023	0,009	0,017	0,010	0,026	0,023	0,009	0,017	3,65
50	0,018	0,020	0,038	0,015	0,023	0,018	0,020	0,038	0,016	0,023	3,38

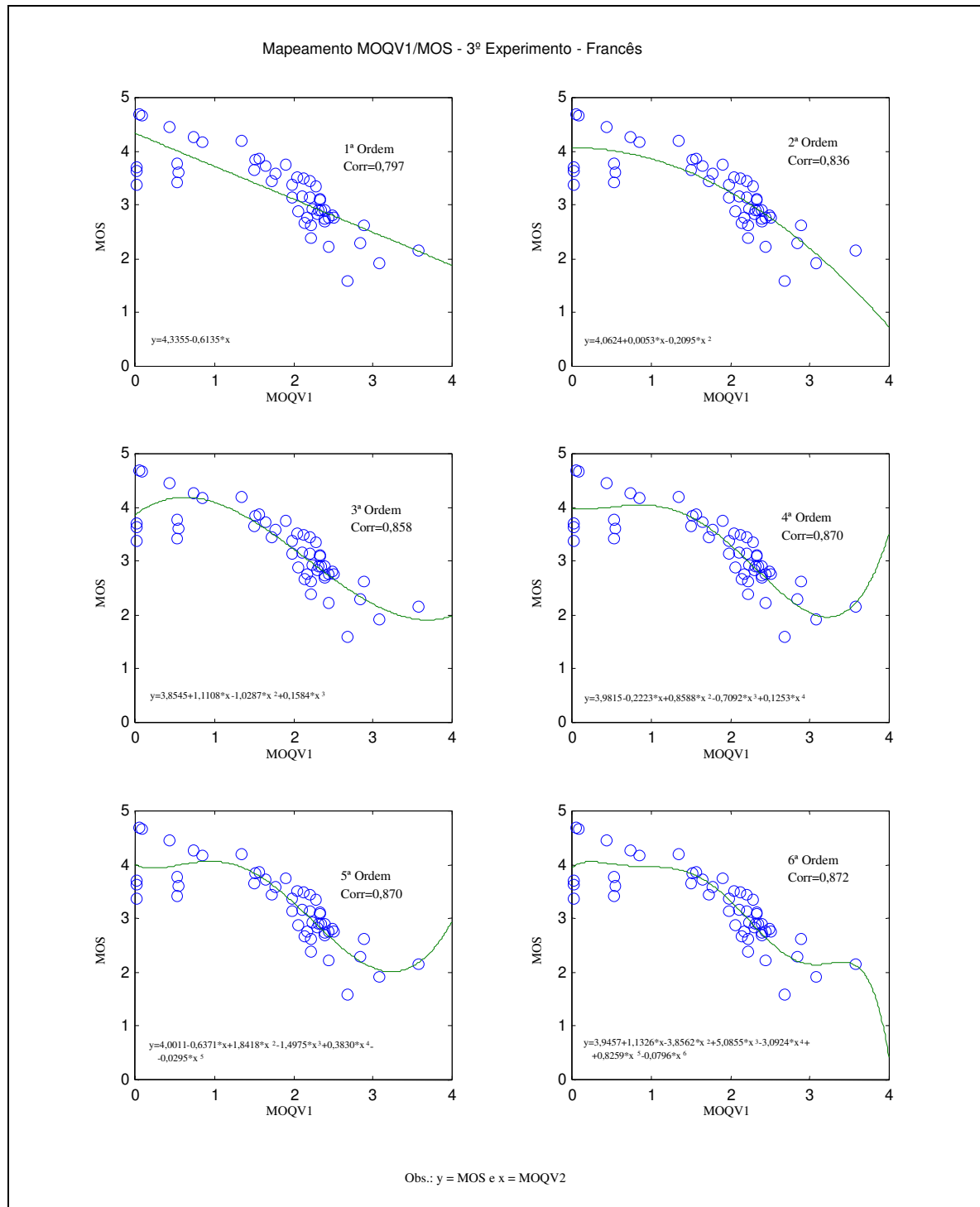


Figura 6.14 - Mapeamento MOQV1/MOS - 3º Experimento - Francês

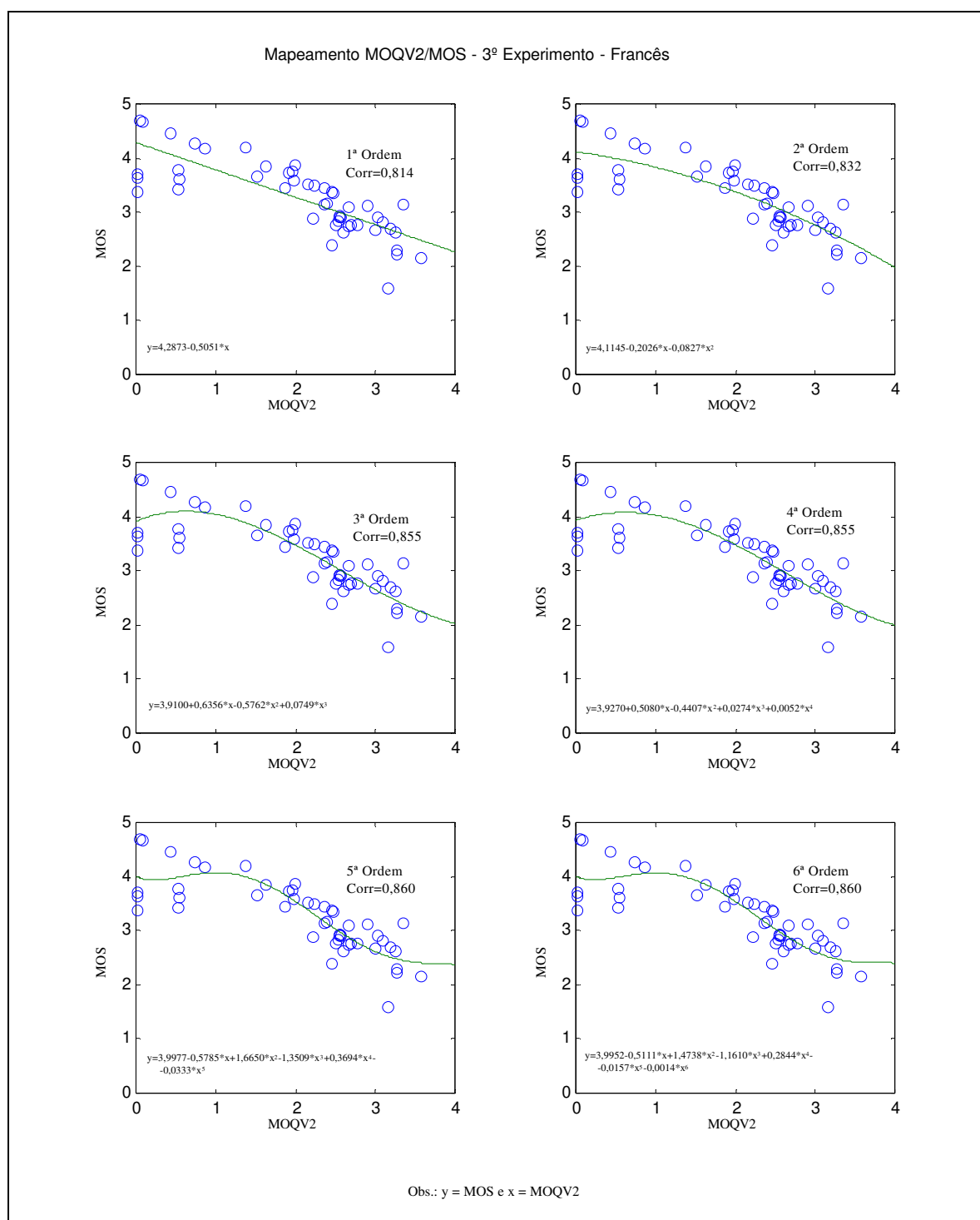


Figura 6.15 - Mapeamento MOQV2/MOS - 3º Experimento - Francês

Tabela 6.12 - Valores obtidos através de simulação - Experimento 3 - Italiano

EXPERIMENTO 3											
<i>Laboratório c (CSELT - Italiano)</i>											
Situação	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,943	1,825	2,006	1,695	1,868	1,970	1,853	2,063	1,741	1,907	4,30
2	1,991	2,114	2,197	1,955	2,064	2,219	2,338	2,322	2,088	2,242	3,79
3	2,023	2,142	2,241	1,995	2,100	2,142	2,428	2,420	2,138	2,282	3,82
4	2,129	2,065	2,156	1,738	2,022	2,778	3,376	2,307	1,861	2,580	3,35
5	2,146	2,087	2,105	2,095	2,108	2,725	2,280	2,442	2,710	2,539	3,15
6	2,256	2,201	2,175	1,994	2,156	2,317	2,271	2,250	2,071	2,227	3,09
7	2,343	2,379	2,274	2,245	2,310	2,452	2,479	2,405	2,448	2,446	2,81
8	2,318	2,477	2,382	2,334	2,378	2,540	2,637	2,573	2,591	2,585	2,47
9	2,454	2,466	2,252	2,240	2,353	3,132	3,186	2,357	2,830	2,876	2,15
10	2,466	2,348	2,528	2,425	2,442	3,296	2,591	3,685	4,007	3,395	1,95
11	2,402	2,443	2,301	2,075	2,305	2,478	2,504	2,356	2,134	2,368	2,80
12	2,396	2,485	2,446	2,280	2,402	2,496	2,583	2,640	2,476	2,549	2,49
13	2,626	2,635	2,484	2,385	2,533	2,933	2,766	2,770	2,523	2,748	2,07
14	2,574	2,465	2,452	2,252	2,436	3,445	2,738	3,225	2,594	3,001	2,17
15	2,460	3,014	2,445	2,361	2,570	2,949	4,327	3,078	3,078	3,358	1,64
16	2,452	2,421	2,327	2,137	2,334	2,533	2,503	2,423	2,230	2,422	2,81
17	2,515	2,601	2,323	2,119	2,389	2,770	2,791	2,442	2,253	2,564	2,68
18	2,490	2,648	2,512	2,382	2,508	2,704	2,973	2,747	2,557	2,745	2,57
19	2,558	2,665	2,261	2,312	2,449	3,078	3,063	2,705	2,783	2,907	2,09
20	2,679	2,617	2,534	2,195	2,506	3,453	3,486	3,927	2,519	3,346	2,08
21	2,610	2,528	2,814	2,585	2,634	2,681	2,599	2,955	2,706	2,735	4,13
22	3,190	3,222	3,433	3,101	3,237	3,329	3,334	3,753	3,298	3,429	3,61
23	2,715	2,979	2,972	2,735	2,850	2,972	3,533	3,550	2,999	3,263	3,49
24	3,321	3,201	3,493	3,343	3,340	3,787	3,818	3,925	3,901	3,858	2,73
25	2,840	2,794	2,859	2,785	2,820	3,694	3,638	3,751	4,406	3,872	2,23
26	3,010	3,453	3,467	3,432	3,340	3,314	4,075	4,122	4,202	3,928	2,55
27	3,087	3,271	3,001	2,978	3,084	3,513	3,794	3,270	3,242	3,455	2,08
28	3,383	3,591	3,274	3,023	3,318	4,674	6,386	4,376	3,877	4,828	1,29
29	2,220	2,328	2,271	2,135	2,238	2,262	2,515	2,406	2,253	2,359	3,86
30	2,591	2,846	2,550	2,336	2,581	2,728	3,111	2,735	2,629	2,801	3,02
31	2,881	3,236	2,906	2,842	2,966	3,155	3,566	3,013	3,081	3,204	2,15
32	3,399	4,320	3,339	3,094	3,538	3,998	5,376	3,765	3,243	4,095	1,45
33	2,033	2,099	2,154	1,966	2,063	2,222	2,218	2,288	2,304	2,258	3,73
34	2,584	2,257	2,768	2,195	2,451	3,025	2,344	3,362	2,305	2,759	3,49
35	2,976	2,623	2,858	2,475	2,733	3,551	3,141	4,100	2,735	3,382	2,71
36	3,153	3,070	3,261	3,188	3,168	4,044	3,297	3,868	4,080	3,822	2,31
37	0,832	1,102	0,799	0,879	0,903	0,833	1,104	0,800	0,880	0,904	4,39
38	0,626	0,721	0,740	0,723	0,703	0,626	0,721	0,740	0,724	0,703	2,99
39	0,594	0,669	0,694	0,664	0,655	0,595	0,669	0,696	0,664	0,656	2,76
40	0,571	0,685	0,739	0,677	0,668	0,571	0,686	0,740	0,677	0,669	2,73
41	3,737	4,460	3,459	3,400	3,764	3,744	4,470	3,481	3,406	3,775	2,22
42	2,588	2,940	2,312	1,923	2,441	2,595	2,945	2,316	1,935	2,448	3,18
43	1,510	2,052	1,220	1,193	1,494	1,509	2,054	1,220	1,196	1,495	3,79
44	0,922	0,954	0,578	0,516	0,743	0,925	0,955	0,578	0,516	0,744	4,20
45	0,439	0,469	0,367	0,319	0,398	0,439	0,470	0,382	0,319	0,403	4,41
46	0,054	0,087	0,115	0,111	0,092	0,055	0,087	0,118	0,111	0,093	4,64
47	0,035	0,041	0,081	0,042	0,049	0,035	0,041	0,089	0,042	0,052	4,60
48	0,010	0,010	0,033	0,030	0,021	0,010	0,010	0,033	0,030	0,021	3,07
49	0,013	0,012	0,046	0,033	0,026	0,013	0,012	0,046	0,033	0,026	2,65
50	0,022	0,028	0,036	0,049	0,034	0,022	0,028	0,036	0,049	0,034	2,71

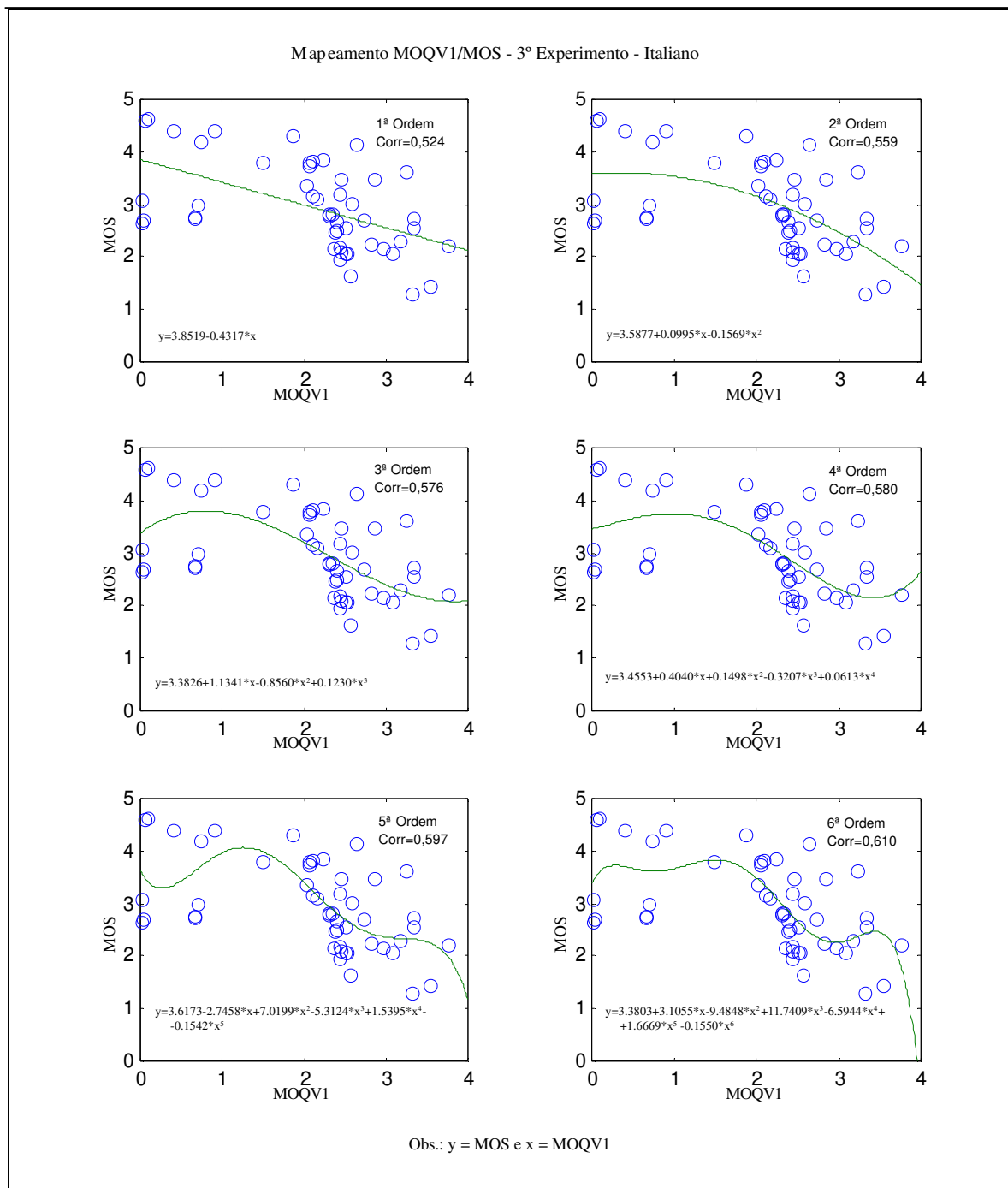


Figura 6.16 - Mapeamento MOQV1/MOS - 3º Experimento - Italiano

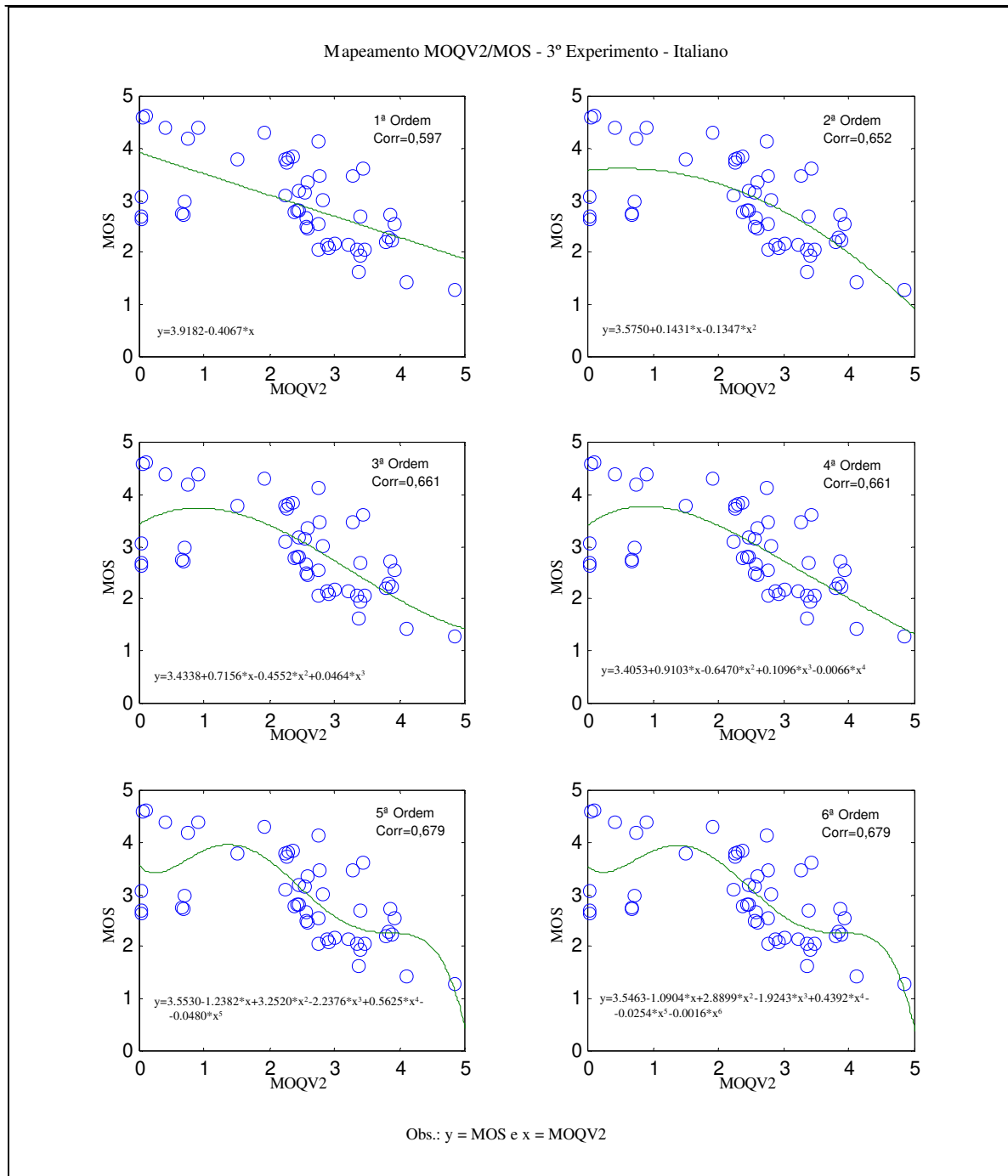


Figura 6.17 - Mapeamento MOQV2/MOS - 3º Experimento - Italiano

Tabela 6.13 - Valores obtidos através de simulação - Experimento 3 - Japonês

EXPERIMENTO 3											
<i>Laboratório d (NTT - Japonês)</i>											
Situação	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,928	1,915	1,517	1,365	1,681	1,983	1,986	1,582	1,408	1,740	3,33
2	2,012	2,031	1,776	1,750	1,892	2,262	2,138	1,972	1,908	2,070	2,68
3	1,956	2,205	1,733	1,766	1,915	2,166	2,549	1,892	2,062	2,167	2,80
4	2,287	2,076	1,651	1,846	1,965	2,983	2,493	1,729	2,470	2,419	2,79
5	2,496	2,581	2,004	1,606	2,172	4,154	3,908	2,728	1,709	3,125	2,40
6	2,305	2,268	1,759	1,897	2,057	2,380	2,353	1,865	1,994	2,148	2,36
7	2,432	2,405	1,968	2,074	2,220	2,573	2,536	2,158	2,300	2,392	1,98
8	2,556	2,504	2,219	2,343	2,405	2,696	2,734	2,502	2,586	2,630	1,88
9	2,597	2,496	1,875	2,181	2,287	3,358	3,020	1,970	2,726	2,768	1,86
10	2,293	2,291	1,980	2,139	2,176	2,777	2,387	2,619	2,647	2,607	1,89
11	2,578	2,396	2,024	2,056	2,264	2,650	2,481	2,108	2,150	2,347	2,24
12	2,632	2,630	2,108	2,199	2,392	2,771	2,843	2,463	2,268	2,586	1,72
13	2,588	2,619	2,218	2,208	2,408	2,863	2,786	2,403	2,349	2,600	1,65
14	2,379	2,444	2,214	2,130	2,292	2,841	2,676	3,134	2,548	2,800	1,79
15	2,604	2,750	2,070	2,442	2,467	3,480	4,144	2,632	3,423	3,420	1,47
16	2,469	2,393	1,854	2,096	2,203	2,592	2,522	2,015	2,203	2,333	2,20
17	2,507	2,586	2,040	2,170	2,326	2,701	2,770	2,214	2,382	2,516	1,93
18	2,531	2,561	2,233	2,346	2,418	2,819	2,835	2,589	2,640	2,721	1,81
19	2,670	2,639	1,980	2,090	2,345	3,147	3,412	2,507	2,589	2,914	1,79
20	2,885	2,628	2,071	2,123	2,427	3,558	2,976	2,379	2,348	2,815	1,90
21	2,533	2,660	2,093	1,958	2,311	2,668	2,766	2,216	2,042	2,423	3,10
22	3,364	2,841	2,707	2,705	2,904	3,518	2,989	3,012	2,897	3,104	2,65
23	2,548	2,829	2,292	2,580	2,562	3,072	3,159	2,610	2,901	2,936	2,27
24	3,460	3,141	2,964	3,098	3,166	3,905	3,567	3,388	3,516	3,594	2,09
25	2,563	2,927	2,413	2,696	2,650	3,950	3,958	2,876	4,997	3,945	2,09
26	3,207	3,087	2,712	2,938	2,986	4,053	4,396	2,994	3,258	3,675	2,41
27	3,181	3,031	2,899	2,849	2,990	3,406	3,395	3,493	3,158	3,363	1,55
28	3,322	3,220	3,070	2,922	3,134	4,756	4,933	4,784	4,797	4,817	1,11
29	2,202	2,132	1,682	1,858	1,969	2,314	2,341	1,726	1,930	2,078	2,80
30	2,841	3,121	1,934	2,259	2,539	3,080	3,541	1,993	2,382	2,749	1,89
31	2,736	3,513	2,357	2,478	2,771	3,306	4,019	2,549	2,571	3,111	1,60
32	3,824	4,042	2,701	2,791	3,340	4,920	4,981	2,933	2,917	3,938	1,15
33	1,967	2,071	1,688	1,795	1,880	2,168	2,301	1,750	1,853	2,018	3,01
34	2,288	2,489	2,145	2,076	2,249	2,388	2,582	2,291	2,751	2,503	2,46
35	2,957	2,730	2,571	2,515	2,693	4,739	3,125	3,870	2,898	3,658	1,98
36	3,098	3,235	2,761	2,704	2,949	3,670	4,235	3,086	3,752	3,686	1,84
37	0,913	1,135	0,656	1,102	0,951	0,919	1,138	0,661	1,105	0,956	3,58
38	0,636	0,701	0,670	0,805	0,703	0,636	0,701	0,668	0,805	0,702	2,45
39	0,708	0,754	0,703	0,724	0,722	0,708	0,754	0,702	0,725	0,722	2,32
40	0,654	0,773	0,634	0,719	0,695	0,654	0,773	0,634	0,719	0,695	2,23
41	4,626	4,461	2,786	2,976	3,712	4,637	4,475	2,793	2,985	3,722	1,79
42	3,307	2,762	1,799	1,537	2,351	3,310	2,768	1,818	1,538	2,359	2,47
43	1,849	1,673	0,974	0,872	1,342	1,848	1,677	0,989	0,874	1,347	3,05
44	1,411	1,168	0,498	0,579	0,914	1,413	1,169	0,503	0,584	0,917	3,58
45	0,619	0,484	0,241	0,324	0,417	0,622	0,485	0,244	0,326	0,419	3,75
46	0,103	0,106	0,101	0,189	0,125	0,107	0,108	0,103	0,191	0,127	4,02
47	0,047	0,037	0,056	0,090	0,057	0,049	0,039	0,056	0,091	0,059	3,85
48	0,016	0,018	0,039	0,046	0,030	0,016	0,018	0,040	0,046	0,030	2,48
49	0,029	0,019	0,022	0,050	0,030	0,029	0,019	0,023	0,050	0,030	2,33
50	0,029	0,025	0,033	0,041	0,032	0,029	0,025	0,033	0,041	0,032	2,18

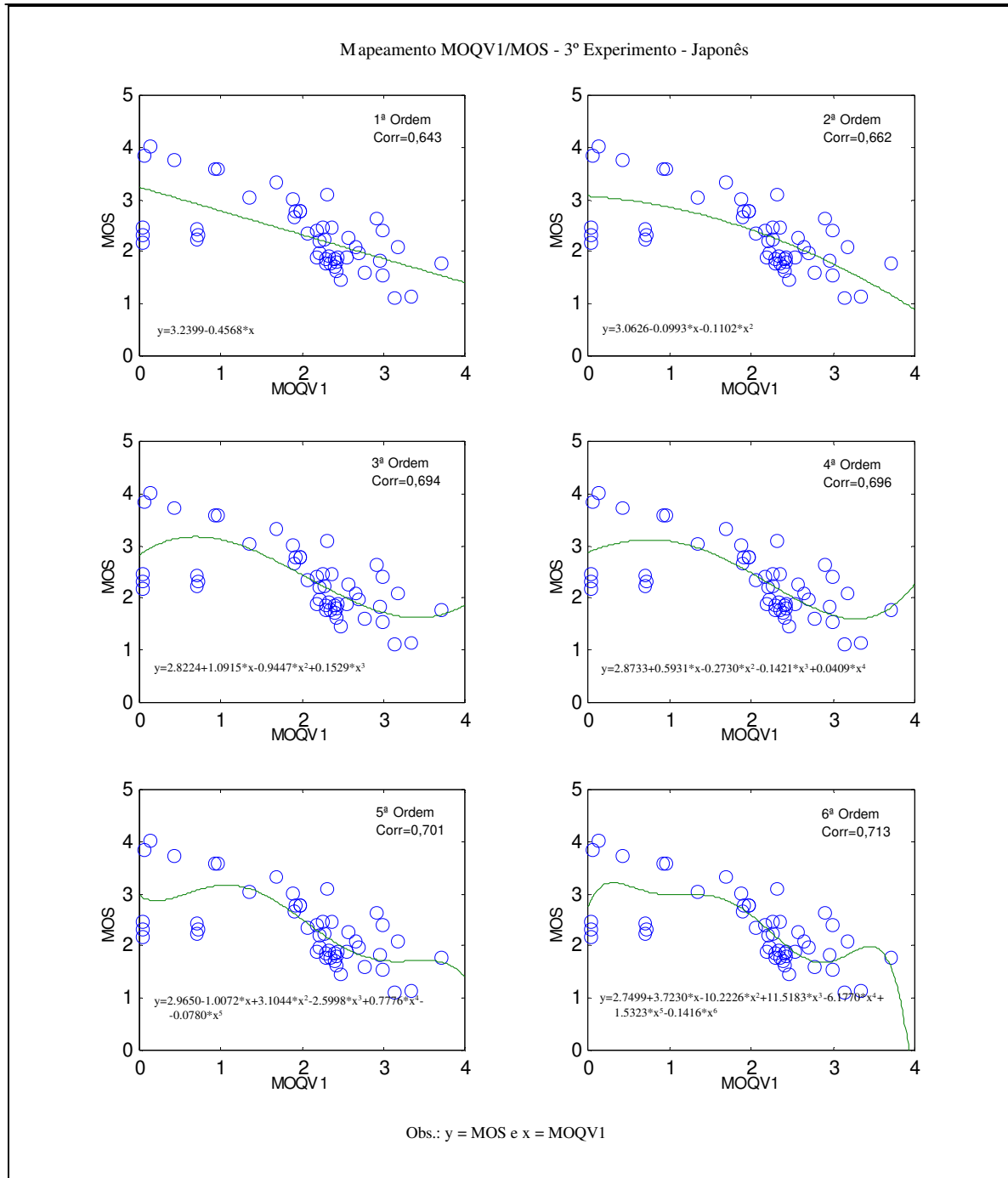


Figura 6.18 - Mapeamento MOQV1/MOS - 3º Experimento - Japonês

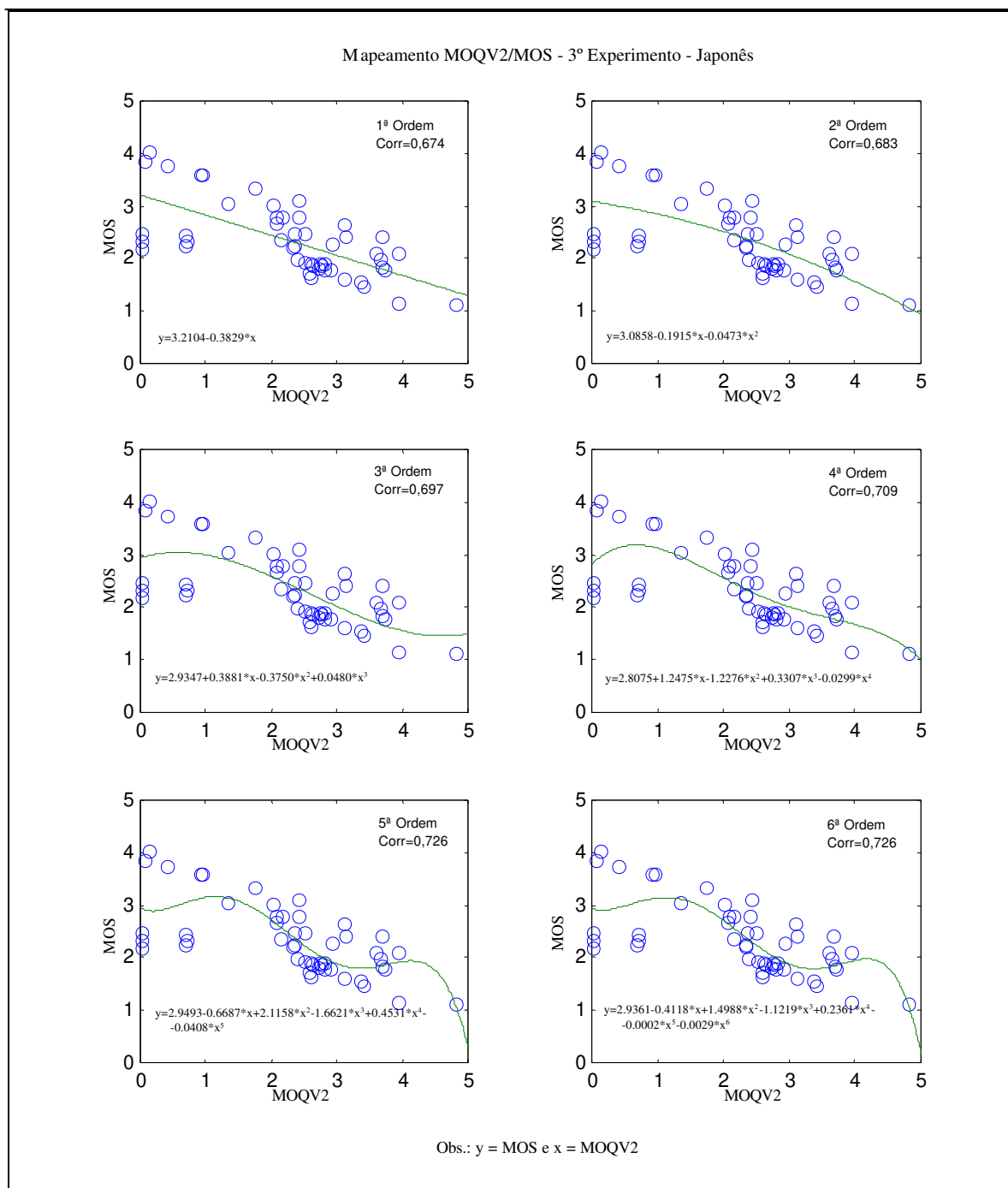


Figura 6.19 - Mapeamento MOQV2/MOS - 3º Experimento - Japonês

Tabela 6.14 - Valores obtidos através de simulação - Experimento 3 - Inglês

EXPERIMENTO 3											
<i>Laboratório o (BNR - Inglês)</i>											
Situação	MOQV1					MOQV2					MOS
	VF1	VF2	VM1	VM2	Média	VF1	VF2	VM1	VM2	Média	Média
1	1,685	2,313	2,014	2,079	2,023	1,732	2,526	2,127	2,153	2,135	3,83
2	2,035	2,452	2,167	2,238	2,223	2,232	2,592	2,361	2,569	2,438	3,15
3	2,026	2,493	2,463	2,167	2,287	2,243	2,746	2,745	2,360	2,523	3,34
4	2,114	2,244	2,249	1,986	2,148	2,985	2,401	2,438	2,358	2,545	3,60
5	2,350	2,286	2,310	2,408	2,339	3,710	2,625	2,758	3,809	3,226	3,19
6	2,034	2,326	2,126	2,149	2,159	2,151	2,429	2,239	2,273	2,273	2,96
7	2,285	2,368	2,175	2,245	2,268	2,413	2,609	2,313	2,373	2,427	2,71
8	2,303	2,400	2,271	2,370	2,336	2,575	2,725	2,454	2,654	2,602	2,38
9	2,255	2,207	2,194	2,577	2,308	2,896	2,366	2,338	3,804	2,851	2,41
10	2,397	2,418	2,511	2,280	2,402	3,933	2,928	3,210	2,539	3,152	2,11
11	2,163	2,454	2,204	2,261	2,270	2,260	2,567	2,329	2,385	2,385	2,74
12	2,372	2,541	2,397	2,290	2,400	2,501	2,693	2,598	2,435	2,557	2,38
13	2,437	2,339	2,474	2,394	2,411	2,610	2,596	2,793	2,573	2,643	2,10
14	2,252	2,317	2,307	2,467	2,336	2,373	2,582	2,604	2,733	2,573	2,28
15	2,297	2,471	2,323	2,360	2,363	2,664	3,048	2,980	2,872	2,891	1,91
16	2,163	2,213	2,215	2,188	2,195	2,318	2,395	2,362	2,348	2,356	2,66
17	2,503	2,490	2,284	2,291	2,392	2,944	2,795	2,433	2,435	2,652	2,51
18	2,378	2,540	2,469	2,400	2,447	2,706	2,949	2,735	2,614	2,751	2,27
19	2,303	2,328	2,184	2,206	2,255	2,627	2,951	2,687	2,731	2,749	2,33
20	2,533	2,406	2,222	2,321	2,370	3,839	2,769	2,416	2,641	2,916	2,33
21	2,392	2,743	2,962	2,561	2,664	2,543	2,859	3,455	2,722	2,895	3,19
22	3,247	2,902	3,566	3,277	3,248	3,552	3,204	4,264	3,760	3,695	2,66
23	3,031	3,033	3,120	2,819	3,000	3,496	3,575	3,513	3,092	3,419	2,63
24	3,421	3,747	3,464	3,576	3,552	3,762	4,853	3,882	4,475	4,243	2,24
25	2,937	2,983	3,267	3,646	3,208	4,560	3,358	4,641	6,977	4,884	2,04
26	2,958	3,971	3,923	3,469	3,580	3,383	5,996	5,801	5,476	5,164	2,14
27	3,082	2,937	3,235	3,002	3,064	3,452	3,478	3,813	3,299	3,510	1,84
28	3,351	3,055	3,139	3,250	3,199	5,392	4,328	4,354	4,562	4,659	1,49
29	2,159	2,129	2,227	2,348	2,216	2,254	2,340	2,453	2,538	2,396	3,23
30	2,778	2,924	2,576	2,630	2,727	2,838	3,338	2,749	2,999	2,981	2,44
31	3,155	3,312	2,956	2,728	3,038	3,657	3,837	3,410	2,935	3,460	1,99
32	3,685	3,673	3,334	3,180	3,468	4,371	4,363	3,704	3,550	3,997	1,43
33	2,141	2,249	2,306	2,152	2,212	2,711	2,390	2,527	2,964	2,648	3,44
34	2,349	2,475	2,955	2,474	2,563	2,458	2,575	4,189	3,033	3,063	2,83
35	2,858	3,036	3,048	2,633	2,894	3,300	3,781	3,759	3,351	3,548	2,38
36	3,088	3,352	3,441	3,074	3,239	3,395	3,562	4,677	3,437	3,768	2,15
37	1,196	1,082	0,806	0,853	0,984	1,197	1,095	0,812	0,861	0,991	3,97
38	0,721	0,692	0,641	0,635	0,672	0,722	0,694	0,642	0,634	0,673	3,03
39	0,714	0,595	0,512	0,561	0,595	0,715	0,595	0,513	0,561	0,596	2,89
40	0,703	0,628	0,480	0,629	0,610	0,706	0,630	0,484	0,632	0,613	2,83
41	4,304	3,864	3,830	3,005	3,751	4,327	3,874	3,879	3,021	3,775	2,02
42	2,707	2,357	1,949	2,611	2,406	2,709	2,353	1,959	2,609	2,408	3,04
43	1,709	1,655	1,344	1,507	1,554	1,712	1,674	1,344	1,513	1,561	3,55
44	1,107	0,815	0,754	0,898	0,894	1,111	0,831	0,767	0,900	0,902	4,19
45	0,395	0,519	0,438	0,452	0,451	0,396	0,521	0,455	0,454	0,457	4,31
46	0,085	0,198	0,160	0,168	0,153	0,087	0,212	0,164	0,179	0,161	4,47
47	0,034	0,098	0,112	0,064	0,077	0,035	0,104	0,130	0,066	0,084	4,28
48	0,022	0,047	0,037	0,037	0,036	0,023	0,048	0,037	0,038	0,036	3,20
49	0,020	0,044	0,041	0,035	0,035	0,020	0,046	0,041	0,036	0,036	2,94
50	0,024	0,036	0,072	0,048	0,045	0,024	0,036	0,077	0,048	0,046	3,08

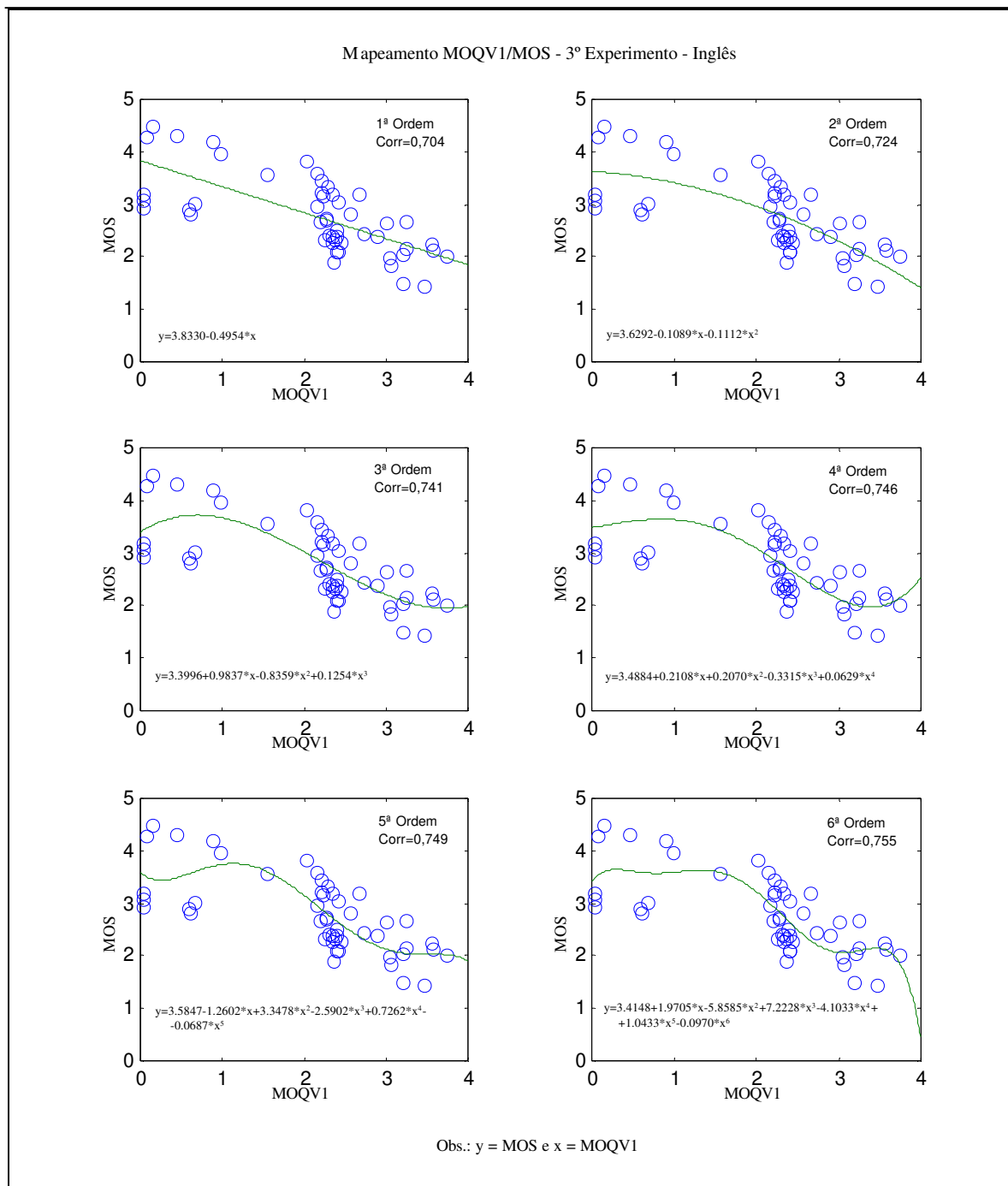


Figura 6.20 - Mapeamento MOQV1/MOS - 3º Experimento - Inglês

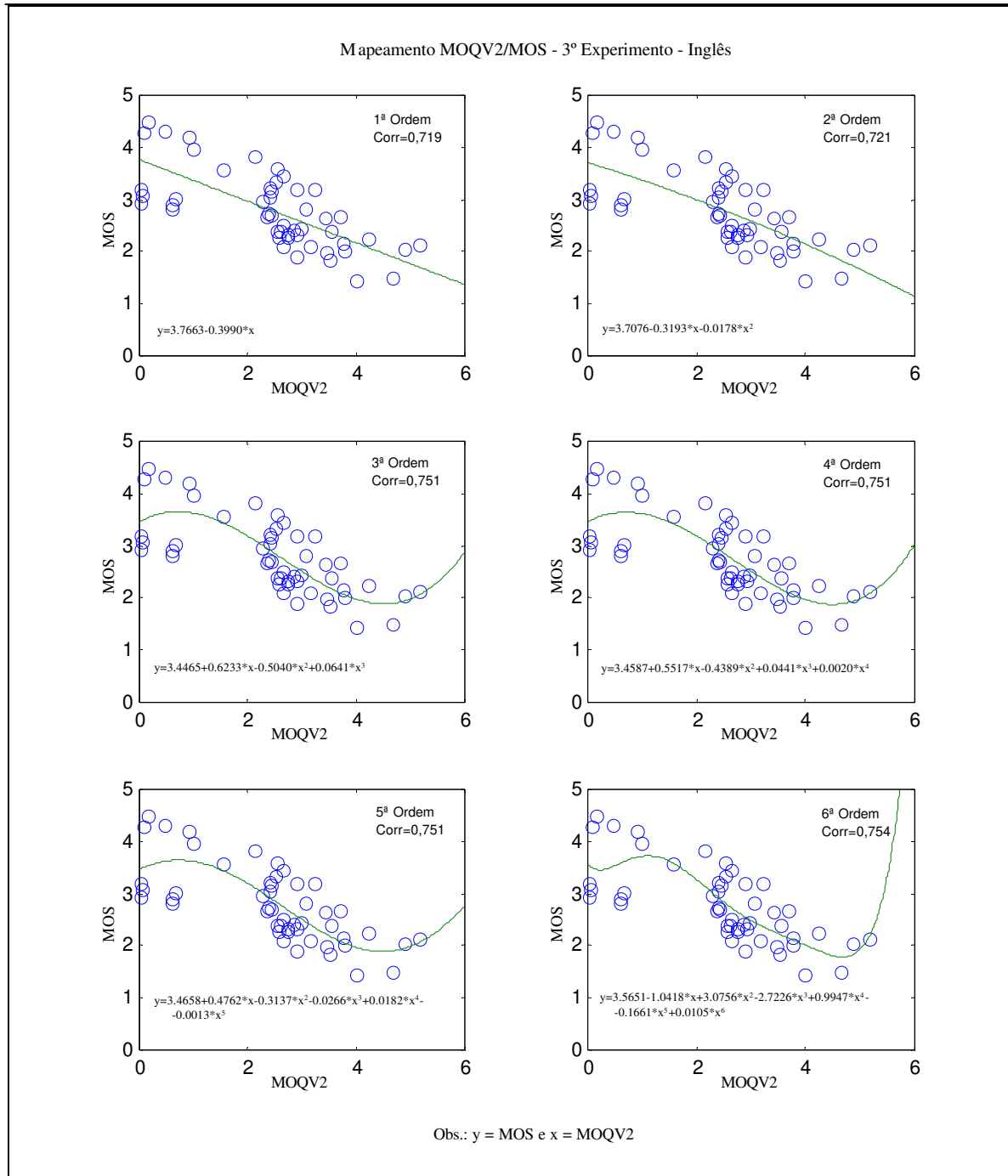


Figura 6.21 - Mapeamento MOQV2/MOS - 3º Experimento - Inglês

6.3.2. Resultados para o 2º Critério

Os resultados obtidos para o segundo critério concordam com aqueles obtidos nos mapeamentos individuais anteriormente descritos, ou seja, bons resultados para os dois primeiros experimentos e ruins para o terceiro, como pode ser visto nas figuras 6.22 a 6.27. Porém, por ser um mapeamento geral, os valores para as correlações são, como era de se esperar, menores. Portanto, seu uso é aconselhável apenas quando não se dispõe de dados para a geração de um mapeamento dedicado à língua em questão.

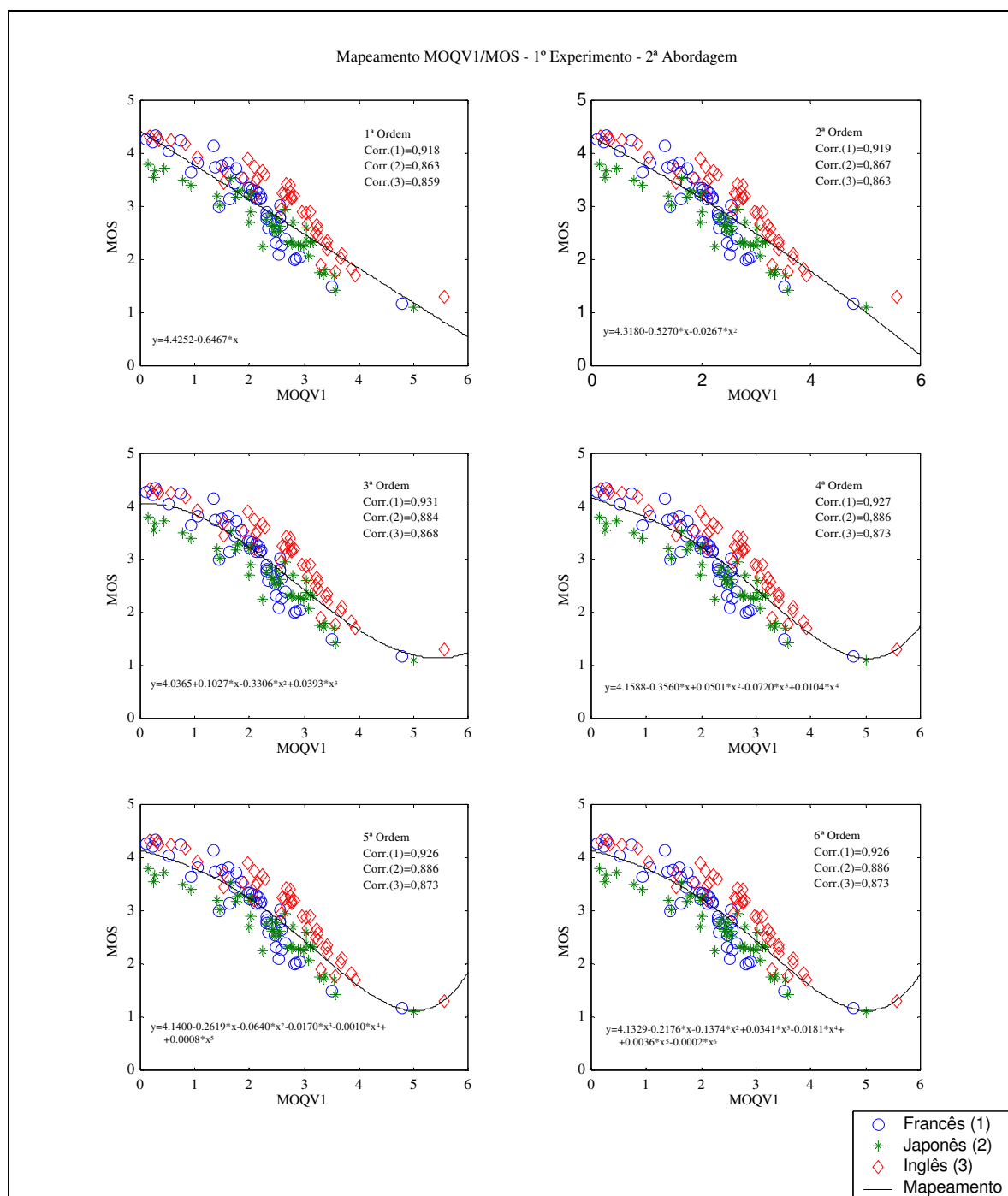


Figura 6.22 - Mapeamento MOQV1/MOS - 1º Experimento - 2º Critério

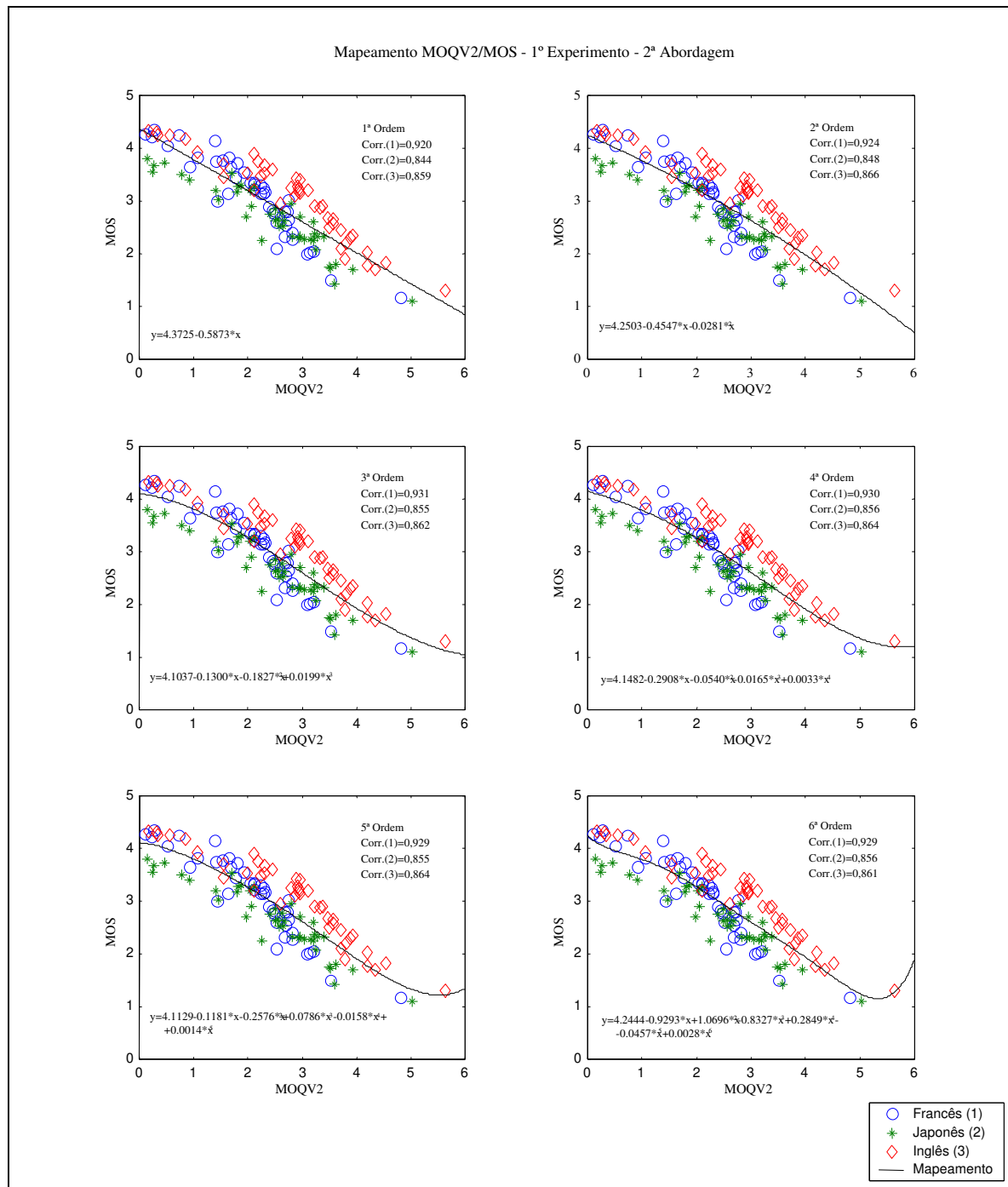


Figura 6.23 - Mapeamento MOQV2/MOS - 1º Experimento - 2º Critério

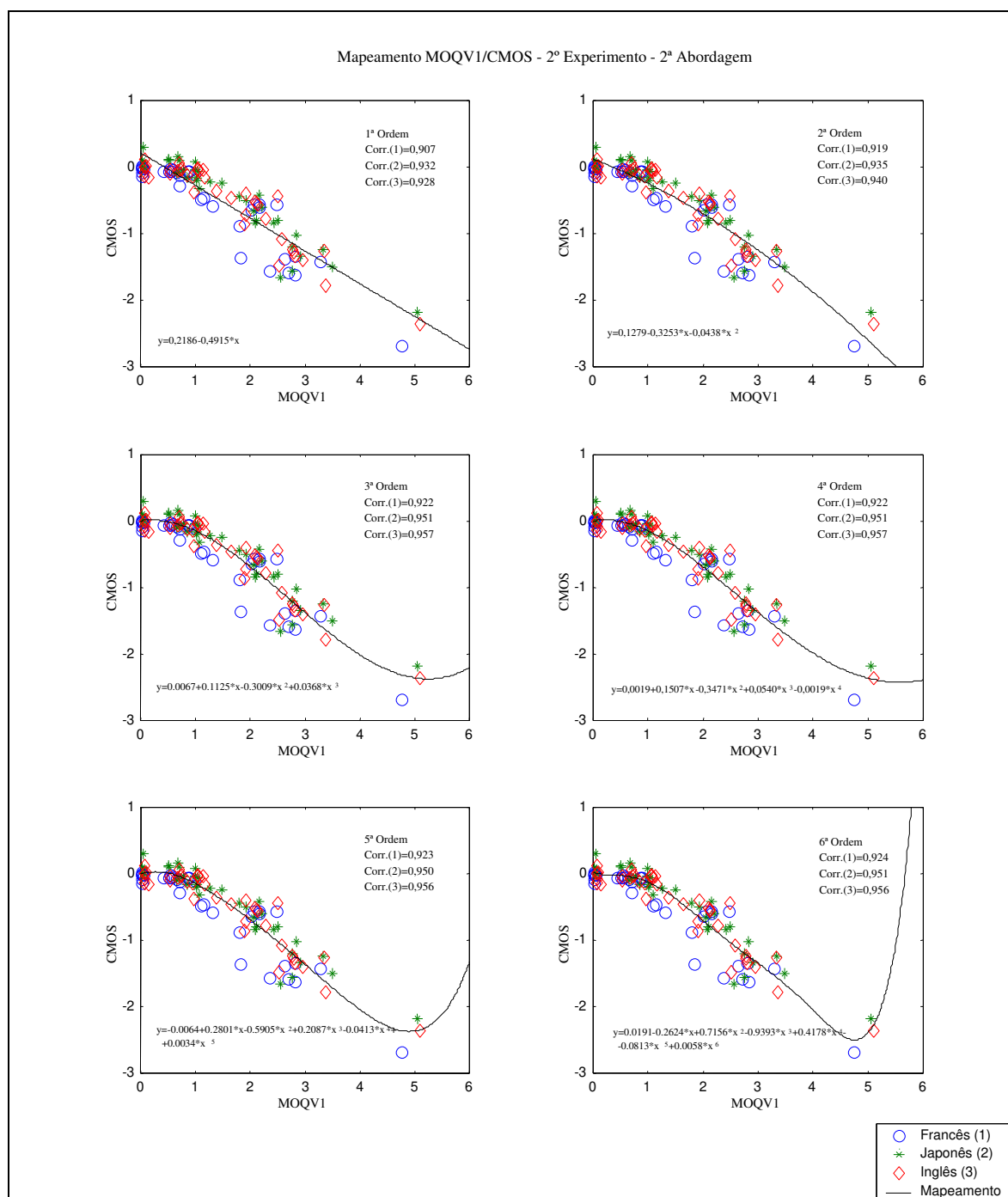


Figura 6.24 - Mapeamento MOQV1/CMOS - 2º Experimento - 2º Critério

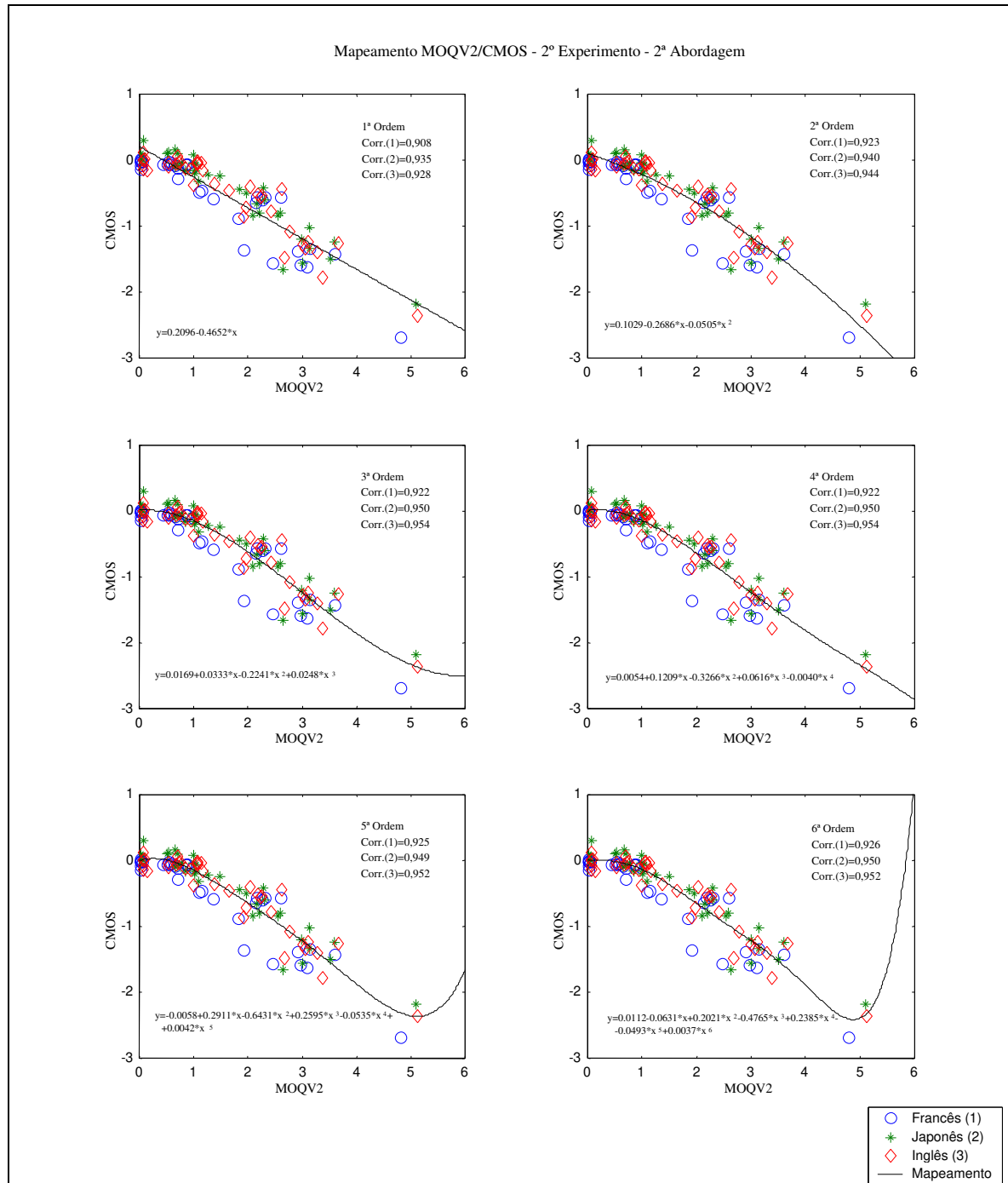


Figura 6.25 - Mapeamento MOQV2/CMOS - 2º Experimento - 2º Critério

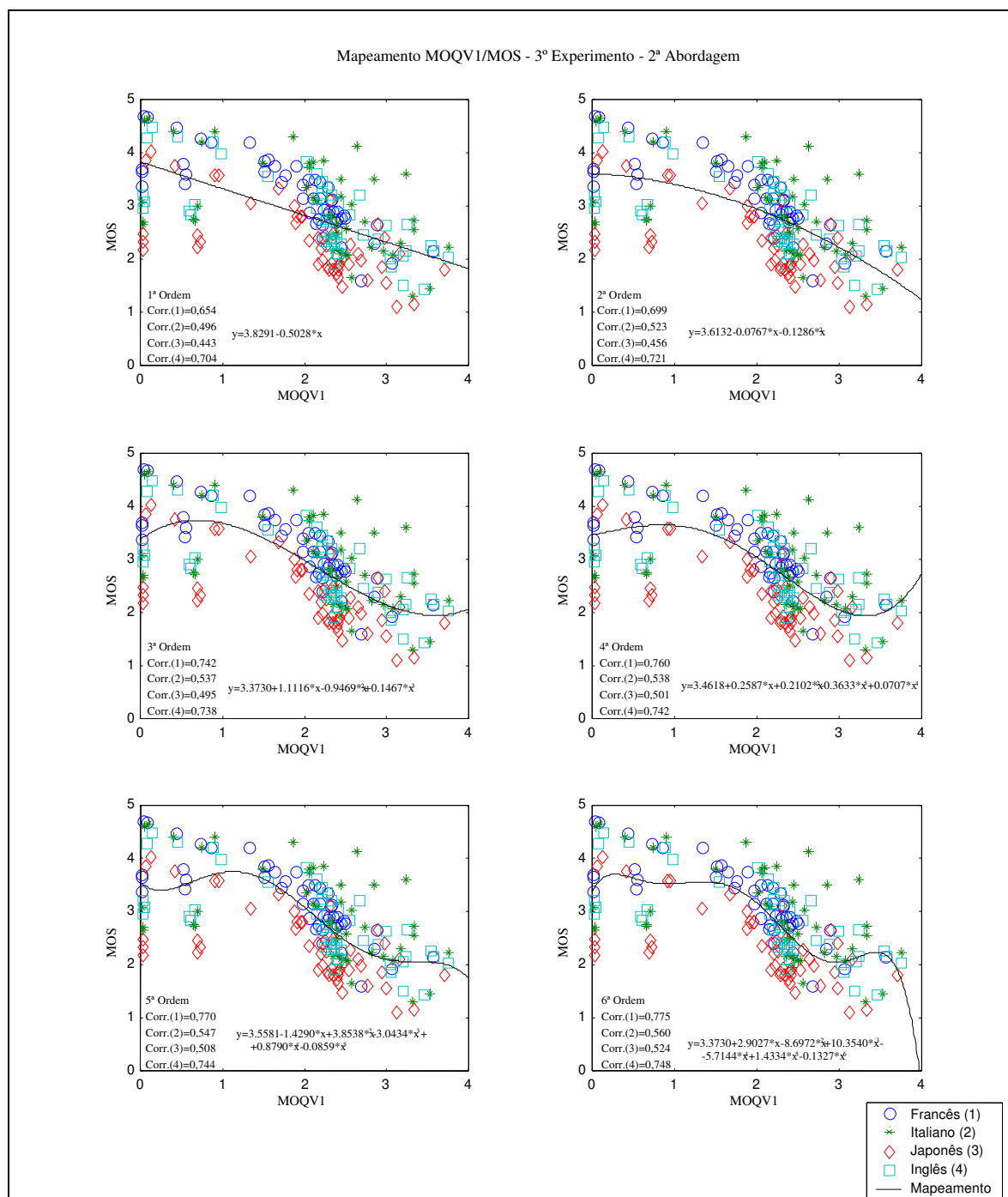


Figura 6.26 - Mapeamento MOQV1/MOS - 3º Experimento - 2º Critério

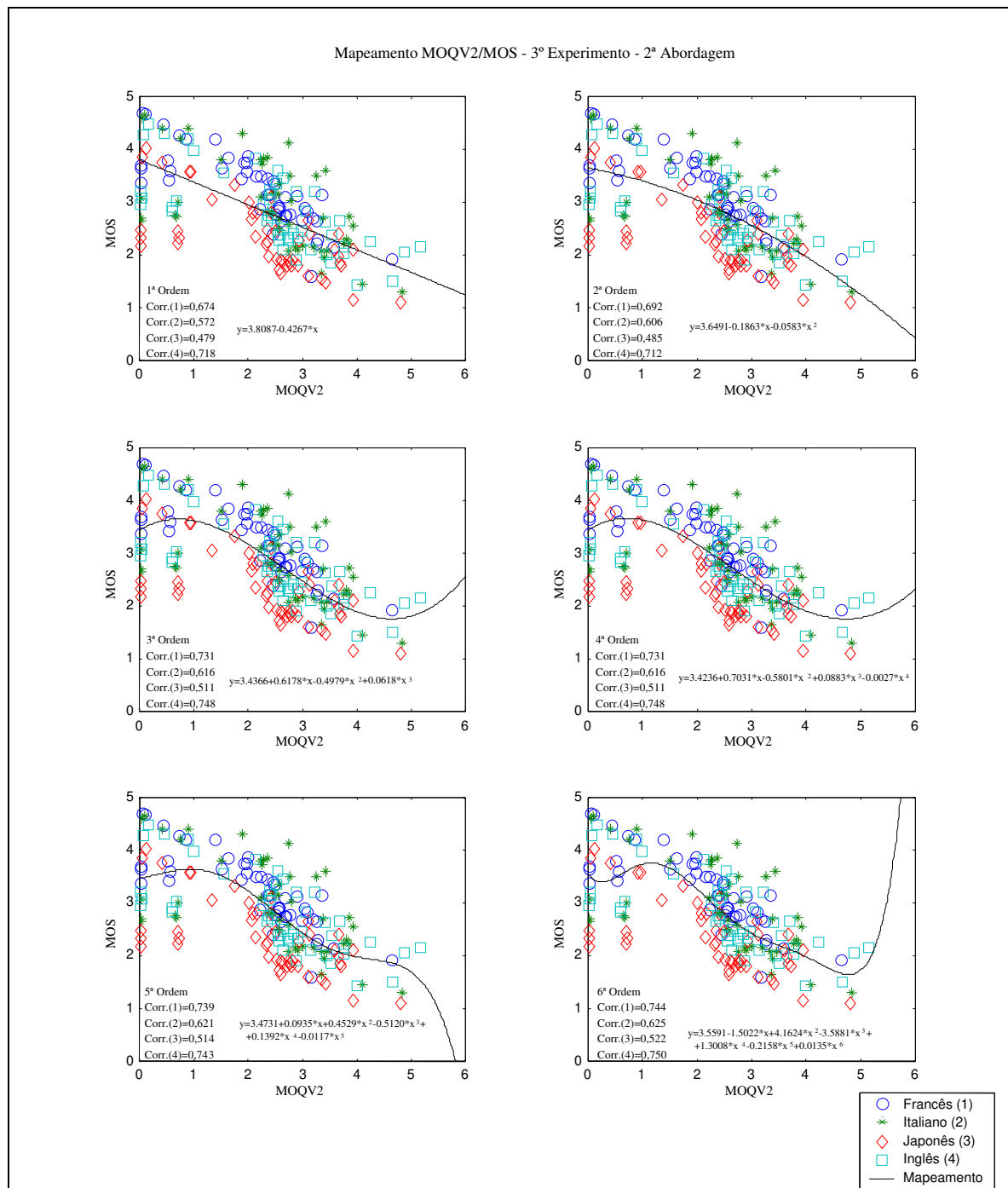


Figura 6.27 - Mapeamento MOQV2/MOS - 3º Experimento - 2º Critério

6.4. RESULTADOS GERAIS

A tabela 6.15 reúne todos os resultados obtidos para a correlação entre as medidas subjetivas e objetivas. Para o segundo critério, houve uma particularização dos resultados para cada língua, a fim de permitir uma comparação destas correlações com aquelas obtidas através do primeiro critério.

As tabelas 6.16 a 6.18 mostram o comportamento dos desvios dos valores objetivos mapeados em relação aos valores subjetivos, contendo as porcentagens por intervalo e o desvio máximo para cada um dos três experimentos. Ao contrário do que foi feito na tabela 6.15, não há aqui uma particularização por língua para o segundo critério, ou seja, os resultados representam o comportamento de todas as línguas agrupadas. Associadas a estas tabelas, as figuras 6.28 a 6.30, na forma de histogramas, descrevem o comportamento dos desvios para os mapeamentos mais importantes (1ª e 3ª ordens). As tabelas 6.19 a 6.21 mostram os valores que determinam os desvios máximos do valor correto correspondentes aos intervalos de confiança escolhidos (90, 95 ou 99%). Como exemplo, tem-se, na tabela 6.19, para a língua francesa, método MOQV1, intervalo de confiança de 90% e 1ª ordem de mapeamento, um desvio máximo de 0,5163. Isto significa que 90% dos valores estarão no intervalo determinado por $MOS \pm 0,5163$ ou, alternativamente, que há uma probabilidade de 90% do valor objetivo obtido estar contido nesse intervalo.

Tabela 6.15 - Resumo dos resultados obtidos para as correlações

INFORMAÇÕES GERAIS				ORDEM DE MAPEAMENTO					
Experim.	Critério	Língua	Versão	1ª	2ª	3ª	4ª	5ª	6ª
1	1	Francês	MOQV1	0,920	0,922	0,947	0,952	0,953	0,956
1	1	Francês	MOQV2	0,923	0,928	0,943	0,951	0,951	0,952
1	1	Japonês	MOQV1	0,927	0,935	0,946	0,951	0,952	0,954
1	1	Japonês	MOQV2	0,924	0,935	0,939	0,942	0,942	0,942
1	1	Inglês	MOQV1	0,930	0,936	0,959	0,973	0,973	0,978
1	1	Inglês	MOQV2	0,942	0,957	0,962	0,973	0,973	0,978
2	1	Francês	MOQV1	0,922	0,934	0,937	0,939	0,939	0,941
2	1	Francês	MOQV2	0,922	0,936	0,936	0,939	0,939	0,943
2	1	Japonês	MOQV1	0,938	0,942	0,957	0,958	0,960	0,963
2	1	Japonês	MOQV2	0,940	0,946	0,956	0,957	0,958	0,962
2	1	Inglês	MOQV1	0,929	0,941	0,959	0,959	0,960	0,960
2	1	Inglês	MOQV2	0,930	0,946	0,955	0,956	0,956	0,959
3	1	Francês	MOQV1	0,797	0,836	0,858	0,870	0,870	0,872
3	1	Francês	MOQV2	0,814	0,832	0,855	0,855	0,860	0,860
3	1	Italiano	MOQV1	0,524	0,559	0,576	0,580	0,597	0,610
3	1	Italiano	MOQV2	0,597	0,652	0,661	0,661	0,679	0,679
3	1	Japonês	MOQV1	0,643	0,662	0,694	0,696	0,701	0,713
3	1	Japonês	MOQV2	0,674	0,683	0,697	0,709	0,726	0,726
3	1	Inglês	MOQV1	0,704	0,724	0,741	0,746	0,749	0,755
3	1	Inglês	MOQV2	0,719	0,721	0,751	0,751	0,751	0,754
1	2	Francês	MOQV1	0,918	0,919	0,931	0,927	0,926	0,926
1	2	Francês	MOQV2	0,920	0,924	0,931	0,930	0,929	0,929
1	2	Japonês	MOQV1	0,863	0,867	0,884	0,886	0,886	0,886
1	2	Japonês	MOQV2	0,844	0,848	0,855	0,856	0,855	0,856
1	2	Inglês	MOQV1	0,859	0,863	0,868	0,873	0,873	0,873
1	2	Inglês	MOQV2	0,859	0,866	0,862	0,864	0,864	0,861
2	2	Francês	MOQV1	0,907	0,919	0,922	0,922	0,923	0,924
2	2	Francês	MOQV2	0,908	0,923	0,922	0,922	0,925	0,926
2	2	Japonês	MOQV1	0,932	0,935	0,951	0,951	0,950	0,951
2	2	Japonês	MOQV2	0,935	0,940	0,950	0,950	0,949	0,950
2	2	Inglês	MOQV1	0,928	0,940	0,957	0,957	0,956	0,956
2	2	Inglês	MOQV2	0,928	0,944	0,954	0,954	0,952	0,952
3	2	Francês	MOQV1	0,778	0,795	0,817	0,820	0,816	0,811
3	2	Francês	MOQV2	0,799	0,807	0,826	0,827	0,829	0,829
3	2	Italiano	MOQV1	0,522	0,552	0,566	0,569	0,580	0,594
3	2	Italiano	MOQV2	0,596	0,633	0,644	0,644	0,649	0,654
3	2	Japonês	MOQV1	0,643	0,661	0,692	0,694	0,699	0,708
3	2	Japonês	MOQV2	0,674	0,683	0,696	0,698	0,703	0,712
3	2	Inglês	MOQV1	0,699	0,715	0,732	0,735	0,736	0,742
3	2	Inglês	MOQV2	0,714	0,707	0,743	0,743	0,737	0,744

Tabela 6.16 - Comportamento dos desvios dos valores mapeados para o 1º experimento

			Intervalo dos Desvios (valores em %)											Desvio Máx.	
Líng.	Tipo	Ord.	<0,1	<0,2	<0,3	<0,4	<0,5	<0,6	<0,7	<0,8	<0,9	<1,0	>1,0		
Francês	MOQV1	1ª	11,36	34,09	65,91	84,09	93,18	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5719
		2ª	13,64	36,36	70,45	81,82	93,18	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5872
		3ª	31,82	56,82	84,09	88,64	97,73	97,73	97,73	100,00	100,00	100,00	100,00	0	0,7556
		4ª	38,64	61,36	84,09	88,64	97,73	97,73	97,73	100,00	100,00	100,00	100,00	0	0,7132
		5ª	38,64	68,18	86,36	90,91	97,73	97,73	97,73	100,00	100,00	100,00	100,00	0	0,7386
		6ª	11,36	68,18	86,36	88,64	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6759
	MOQV2	1ª	13,64	40,91	65,91	86,36	90,91	95,45	100,00	100,00	100,00	100,00	100,00	0	0,6653
		2ª	18,18	54,55	72,73	86,36	93,18	95,45	100,00	100,00	100,00	100,00	100,00	0	0,6988
		3ª	29,55	68,18	86,36	88,64	93,18	95,45	97,73	100,00	100,00	100,00	100,00	0	0,7766
		4ª	45,45	72,73	88,64	88,64	95,45	95,45	97,73	100,00	100,00	100,00	100,00	0	0,7158
		5ª	45,45	72,73	88,64	88,64	95,45	95,45	97,73	100,00	100,00	100,00	100,00	0	0,7155
		6ª	13,64	72,73	88,64	88,64	93,18	95,45	97,73	100,00	100,00	100,00	100,00	0	0,7016
Japonês	MOQV1	1ª	34,09	61,36	72,73	88,64	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5056
		2ª	36,36	61,36	79,55	90,91	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5319
		3ª	38,64	72,73	81,82	95,45	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5776
		4ª	43,18	75,00	88,64	97,73	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6250
		5ª	45,45	75,00	86,36	97,73	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6158
		6ª	34,09	70,45	88,64	97,73	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5755
	MOQV2	1ª	34,09	61,36	75,00	88,64	95,45	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6026
		2ª	40,91	63,64	86,36	93,18	95,45	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5981
		3ª	38,64	72,73	81,82	93,18	95,45	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6270
		4ª	40,91	72,73	86,36	93,18	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6601
		5ª	40,91	70,45	84,09	95,45	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6505
		6ª	34,09	68,18	84,09	95,45	97,73	97,73	100,00	100,00	100,00	100,00	100,00	0	0,6371
Inglês	MOQV1	1ª	15,91	43,18	70,45	86,36	95,45	95,45	100,00	100,00	100,00	100,00	100,00	0	0,6597
		2ª	22,73	56,82	75,00	88,64	93,18	95,45	100,00	100,00	100,00	100,00	100,00	0	0,6771
		3ª	29,55	68,18	86,36	93,18	95,45	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5649
		4ª	50,00	77,27	88,64	95,45	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5698
		5ª	50,00	81,82	88,64	97,73	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5841
		6ª	15,91	84,09	90,91	97,73	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5169
	MOQV2	1ª	25,00	47,73	68,18	88,64	97,73	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5404
		2ª	40,91	61,36	77,27	93,18	95,45	100,00	100,00	100,00	100,00	100,00	100,00	0	0,5379
		3ª	36,36	65,91	84,09	93,18	100,00	100,00	100,00	100,00	100,00	100,00	100,00	0	0,4777
		4ª	43,18	75,00	88,64	93,18	100,00	100,00	100,00	100,00	100,00	100,00	100,00	0	0,4355
		5ª	45,45	79,55	86,36	93,18	100,00	100,00	100,00	100,00	100,00	100,00	100,00	0	0,4435
		6ª	25,00	86,36	90,91	97,73	100,00	100,00	100,00	100,00	100,00	100,00	100,00	0	0,4818
Geral (2º Critério)	MOQV1	1ª	20,45	38,64	56,82	67,42	80,30	89,39	96,21	100,00	100,00	100,00	100,00	0	0,7621
		2ª	21,21	39,39	56,06	70,45	81,82	89,39	96,21	100,00	100,00	100,00	100,00	0	0,7625
		3ª	19,70	39,39	59,09	71,97	82,58	93,18	96,97	99,24	100,00	100,00	100,00	0	0,8080
		4ª	18,94	38,64	56,82	74,24	83,33	93,94	97,73	99,24	100,00	100,00	100,00	0	0,8243
		5ª	18,18	38,64	56,82	74,24	83,33	94,70	97,73	99,24	100,00	100,00	100,00	0	0,8221
		6ª	20,45	38,64	57,58	74,24	83,33	94,70	97,73	99,24	100,00	100,00	100,00	0	0,8222
	MOQV2	1ª	17,42	36,36	56,06	68,94	79,55	90,15	94,70	98,48	100,00	100,00	100,00	0	0,8436
		2ª	16,67	36,36	53,03	71,21	81,06	91,67	94,70	97,73	100,00	100,00	100,00	0	0,8450
		3ª	18,18	35,61	56,06	71,97	80,30	91,67	94,70	98,48	100,00	100,00	100,00	0	0,8732
		4ª	18,94	35,61	53,79	72,73	81,06	92,42	94,70	98,48	100,00	100,00	100,00	0	0,8781
		5ª	19,70	35,61	55,30	72,73	81,06	92,42	94,70	98,48	100,00	100,00	100,00	0	0,8730
		6ª	17,42	37,88	53,79	71,21	80,30	90,91	94,70	97,73	100,00	100,00	100,00	0	0,8711

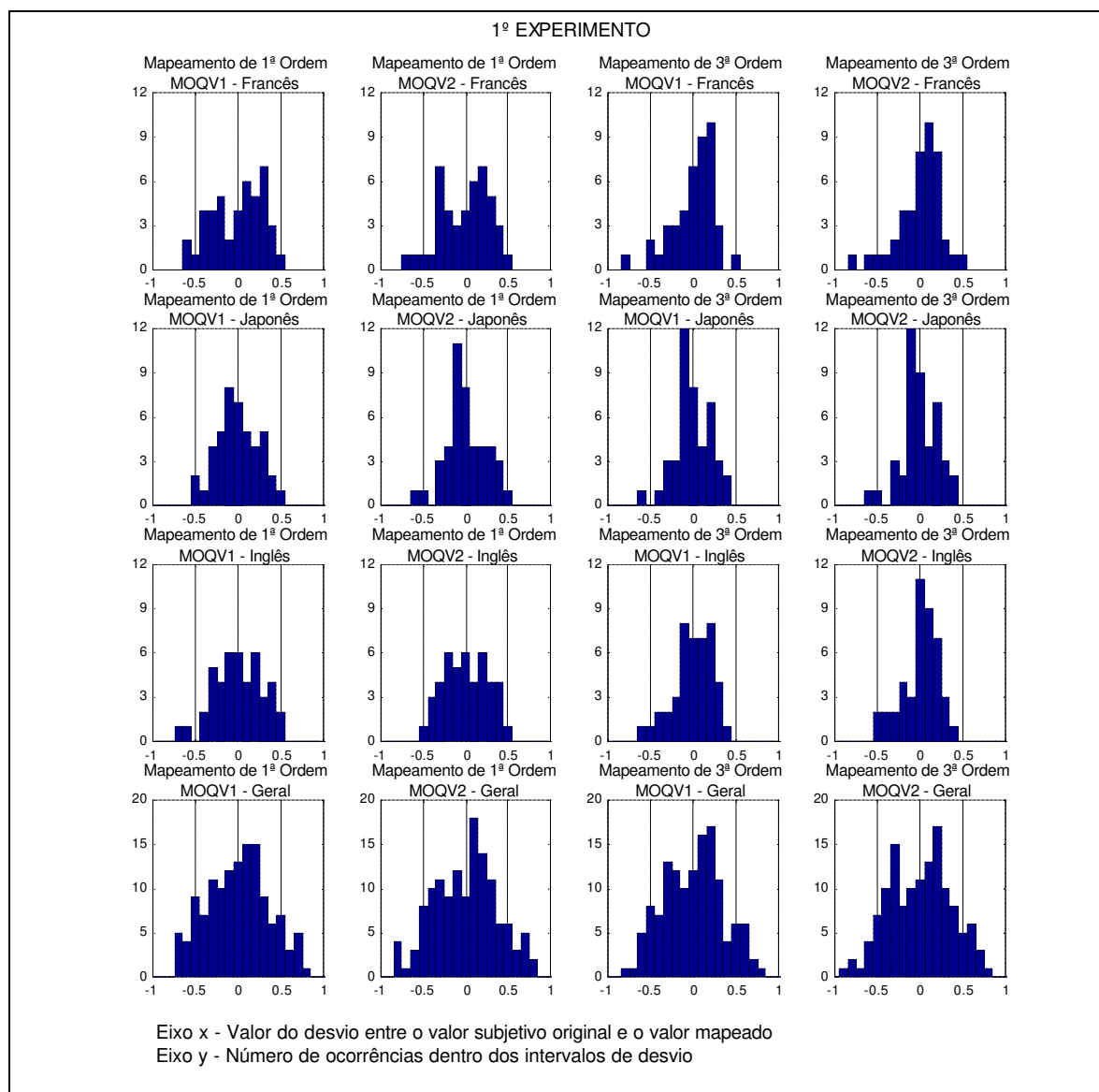


Figura 6.28 - Histogramas caracterizando os desvios entre os valores subjetivos e valores mapeados

Tabela 6.17 - Comportamento dos desvios dos valores mapeados para o 2º experimento

Líng.	Tipo	Ord.	Intervalo dos Desvios (%)											Desvio Máx.
			<0,1	<0,2	<0,3	<0,4	<0,5	<0,6	<0,7	<0,8	<0,9	<1,0	>1,0	
Francês	MOQV1	1ª	27,78	55,56	77,78	88,89	94,44	100,00	100,00	100,00	100,00	100,00	0	0,5920
		2ª	38,89	69,44	83,33	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6378
		3ª	41,67	72,22	88,89	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6573
		4ª	41,67	75,00	80,56	91,67	94,44	97,22	100,00	100,00	100,00	100,00	0	0,6075
		5ª	41,67	72,22	80,56	91,67	94,44	100,00	100,00	100,00	100,00	100,00	0	0,5989
		6ª	27,78	77,78	88,89	91,67	94,44	97,22	100,00	100,00	100,00	100,00	0	0,6518
	MOQV2	1ª	33,33	52,78	80,56	86,11	94,44	100,00	100,00	100,00	100,00	100,00	0	0,5787
		2ª	36,11	72,22	88,89	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6676
		3ª	41,67	72,22	86,11	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6741
		4ª	41,67	75,00	83,33	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6098
		5ª	41,67	72,22	80,56	91,67	91,67	100,00	100,00	100,00	100,00	100,00	0	0,5945
		6ª	33,33	77,78	88,89	91,67	91,67	97,22	100,00	100,00	100,00	100,00	0	0,6434
Japonês	MOQV1	1ª	37,50	72,50	90,00	95,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6944
		2ª	35,00	77,50	92,50	95,00	95,00	97,00	97,00	100,00	100,00	100,00	0	0,7267
		3ª	57,00	87,00	92,50	95,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6562
		4ª	52,50	87,00	92,50	95,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6416
		5ª	52,50	82,00	92,50	97,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5989
		6ª	37,50	82,00	92,50	97,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5874
	MOQV2	1ª	35,00	77,50	92,50	95,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6998
		2ª	37,50	82,00	92,50	95,00	97,00	97,00	97,00	100,00	100,00	100,00	0	0,7403
		3ª	57,00	82,00	90,00	97,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6868
		4ª	60,00	85,00	92,50	97,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6671
		5ª	50,00	85,00	92,50	97,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6269
		6ª	35,00	85,00	92,50	97,00	97,00	97,00	100,00	100,00	100,00	100,00	0	0,6141
Inglês	MOQV1	1ª	27,50	60,00	87,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5302
		2ª	35,00	65,00	92,50	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5500
		3ª	57,00	87,00	92,50	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5307
		4ª	57,00	87,00	92,50	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5274
		5ª	57,00	87,00	90,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5489
		6ª	27,50	85,00	92,50	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5498
	MOQV2	1ª	35,00	60,00	85,00	92,50	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5315
		2ª	37,50	77,50	92,50	92,50	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5459
		3ª	55,00	85,00	87,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5039
		4ª	57,00	85,00	87,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5091
		5ª	57,00	85,00	90,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5362
		6ª	35,00	85,00	90,00	95,00	97,00	100,00	100,00	100,00	100,00	100,00	0	0,5309
Geral (2º Critério)	MOQV1	1ª	25,00	59,48	83,62	90,52	95,69	97,41	100,00	100,00	100,00	100,00	0	0,6702
		2ª	33,62	66,38	87,07	92,24	95,69	97,41	99,14	100,00	100,00	100,00	0	0,7369
		3ª	50,00	71,55	87,93	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7815
		4ª	50,86	72,41	87,07	93,97	96,55	98,28	99,14	100,00	100,00	100,00	0	0,7761
		5ª	48,28	72,41	87,07	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7658
		6ª	25,00	72,41	86,21	93,97	96,55	98,28	99,14	100,00	100,00	100,00	0	0,7670
	MOQV2	1ª	29,31	59,48	84,48	92,24	95,69	96,55	100,00	100,00	100,00	100,00	0	0,6716
		2ª	37,07	69,83	86,21	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7561
		3ª	50,86	73,28	87,07	93,10	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7860
		4ª	50,86	71,55	87,07	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7721
		5ª	49,14	72,41	86,21	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7589
		6ª	29,31	72,41	86,21	93,97	96,55	97,41	99,14	100,00	100,00	100,00	0	0,7571

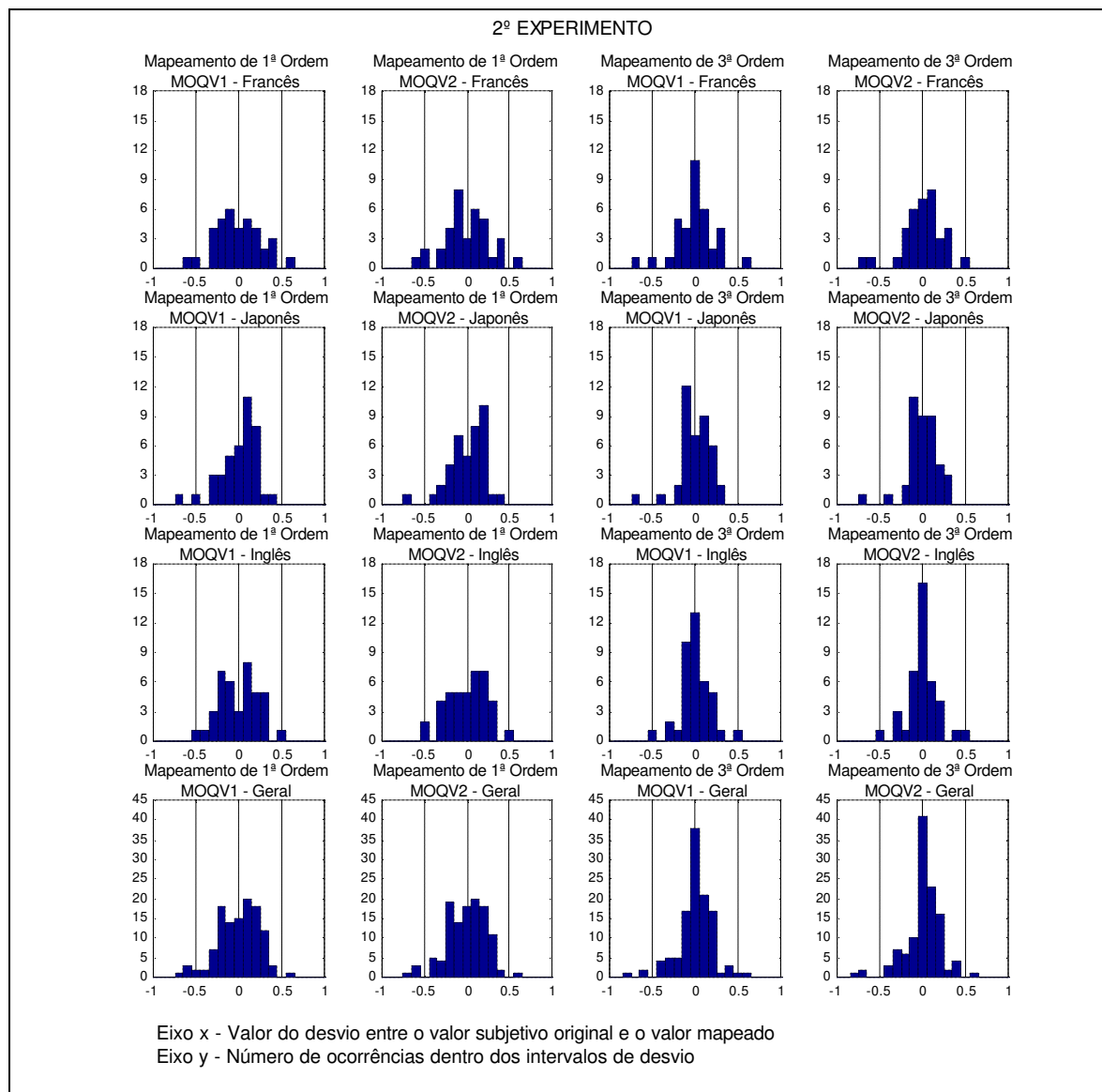


Figura 6.29 - Histogramas caracterizando os desvios entre os valores subjetivos e valores mapeados

Tabela 6.18 - Comportamento dos desvios dos valores mapeados para o 3º experimento

Líng.	Tipo	Ord.	Intervalo dos Desvios (%)											Desvio Máx.
			<0,1	<0,2	<0,3	<0,4	<0,5	<0,6	<0,7	<0,8	<0,9	<1,0	>1,0	
Francês	MOQV1	1ª	22	36	48	66	78	88	96	96	96	98	2	1,0995
		2ª	22	42	50	72	82	86	96	98	98	100	0	0,9784
		3ª	28	48	62	78	86	90	92	98	100	100	0	0,8996
		4ª	28	48	62	74	86	92	94	100	100	100	0	0,7702
		5ª	28	50	62	76	86	90	94	100	100	100	0	0,7836
		6ª	22	36	62	76	88	92	96	100	100	100	0	0,7558
	MOQV2	1ª	20	34	44	68	82	90	96	96	98	98	2	1,1009
		2ª	14	34	54	74	82	92	94	98	98	98	2	1,0583
		3ª	18	46	70	82	86	88	90	98	98	100	0	0,9385
		4ª	14	48	70	80	86	88	90	98	98	100	0	0,9283
		5ª	22	44	70	80	84	88	94	98	98	100	0	0,9344
		6ª	20	44	70	80	84	88	94	98	98	100	0	0,9327
Italiano	MOQV1	1ª	10	20	24	34	40	44	56	66	82	86	14	1,4152
		2ª	8	14	24	32	40	48	56	64	82	88	12	1,3687
		3ª	6	18	30	42	44	50	62	74	80	86	14	1,4512
		4ª	10	24	30	40	44	50	62	74	82	88	12	1,4805
		5ª	8	20	32	36	50	56	64	70	82	86	14	1,5671
		6ª	10	26	30	42	50	56	64	76	84	88	12	1,6547
	MOQV2	1ª	4	18	28	34	44	52	62	72	84	90	10	1,3241
		2ª	6	18	18	34	40	58	68	76	88	92	8	1,1712
		3ª	6	14	20	38	48	62	70	78	86	92	8	1,2036
		4ª	6	18	20	38	50	62	70	80	86	90	10	1,2051
		5ª	12	14	26	44	54	68	74	80	88	90	10	1,2901
		6ª	4	14	26	44	54	68	74	78	88	90	10	1,2883
Japonês	MOQV1	1ª	12	16	32	48	64	70	80	90	96	98	2	1,0455
		2ª	12	20	28	50	62	70	78	92	98	100	0	0,9715
		3ª	14	26	42	54	72	78	86	88	92	98	2	1,0760
		4ª	14	26	44	52	72	80	84	90	92	98	2	1,0774
		5ª	16	30	44	60	72	78	84	88	92	98	2	1,1371
		6ª	12	30	44	60	70	78	84	88	92	100	0	0,9651
	MOQV2	1ª	8	16	30	54	62	72	80	92	98	98	2	1,0181
		2ª	8	20	24	52	66	70	82	96	98	100	0	0,9593
		3ª	2	16	36	56	66	76	84	94	98	98	2	1,0418
		4ª	6	22	36	60	72	78	88	90	94	98	2	1,0729
		5ª	14	30	52	62	72	78	82	90	96	98	2	1,1248
		6ª	8	28	50	60	72	80	82	92	96	98	2	1,1143
Inglês	MOQV1	1ª	16	22	34	46	60	66	78	94	98	100	0	0,9989
		2ª	12	16	22	42	62	72	88	94	100	100	0	0,8758
		3ª	16	26	36	54	66	78	84	88	98	100	0	0,9392
		4ª	8	24	34	52	58	78	84	96	98	100	0	0,9457
		5ª	12	24	36	52	62	76	86	94	98	98	2	1,0085
		6ª	16	22	36	48	58	80	86	96	100	100	0	0,8988
	MOQV2	1ª	6	18	32	48	60	72	80	94	98	100	0	0,9155
		2ª	4	14	30	48	60	72	82	94	100	100	0	0,8850
		3ª	10	20	36	54	64	74	86	94	98	100	0	0,9360
		4ª	10	18	38	54	64	74	86	94	98	100	0	0,9337
		5ª	12	16	38	52	64	74	86	94	98	100	0	0,9359
		6ª	6	20	40	58	64	76	86	96	98	98	2	1,0036
Geral (2º Critério)	MOQV1	1ª	11,0	24,5	34,5	42,0	53,5	60,0	65,5	78,0	88,0	90,5	9,5	1,6332
		2ª	11,0	23,0	32,0	42,0	52,5	63,5	71,5	78,0	85,0	92,0	8,0	1,6113
		3ª	10,0	23,0	37,0	46,5	58,5	66,0	74,0	81,5	86,5	92,0	8,0	1,7175
		4ª	11,0	24,0	35,0	47,0	55,5	66,5	74,0	82,5	88,5	92,5	7,5	1,7624
		5ª	10,5	24,5	38,0	47,5	57,0	66,5	76,5	81,5	88,0	91,0	9,0	1,8006
		6ª	11,0	25,0	38,5	47,0	56,5	66,5	76,0	83,0	89,0	92,0	8,0	1,8954
	MOQV2	1ª	12,5	25,0	34,5	45,0	54,5	60,0	68,5	79,5	88,5	93,0	7,0	1,6150
		2ª	10,5	23,5	34,5	48,0	54,5	63,5	70,5	78,0	87,5	92,0	8,0	1,4631
		3ª	11,5	24,0	36,5	48,5	58,0	66,0	72,0	82,5	89,5	92,5	7,5	1,4636
		4ª	11,5	23,5	36,5	48,5	58,5	65,5	72,5	82,5	90,0	92,5	7,5	1,4680
		5ª	13,5	25,5	38,0	48,0	57,0	67,0	73,0	83,0	89,0	92,5	7,5	1,4897
		6ª	12,5	26,0	38,5	51,0	59,0	68,0	75,0	81,5	89,0	91,5	8,5	1,5231

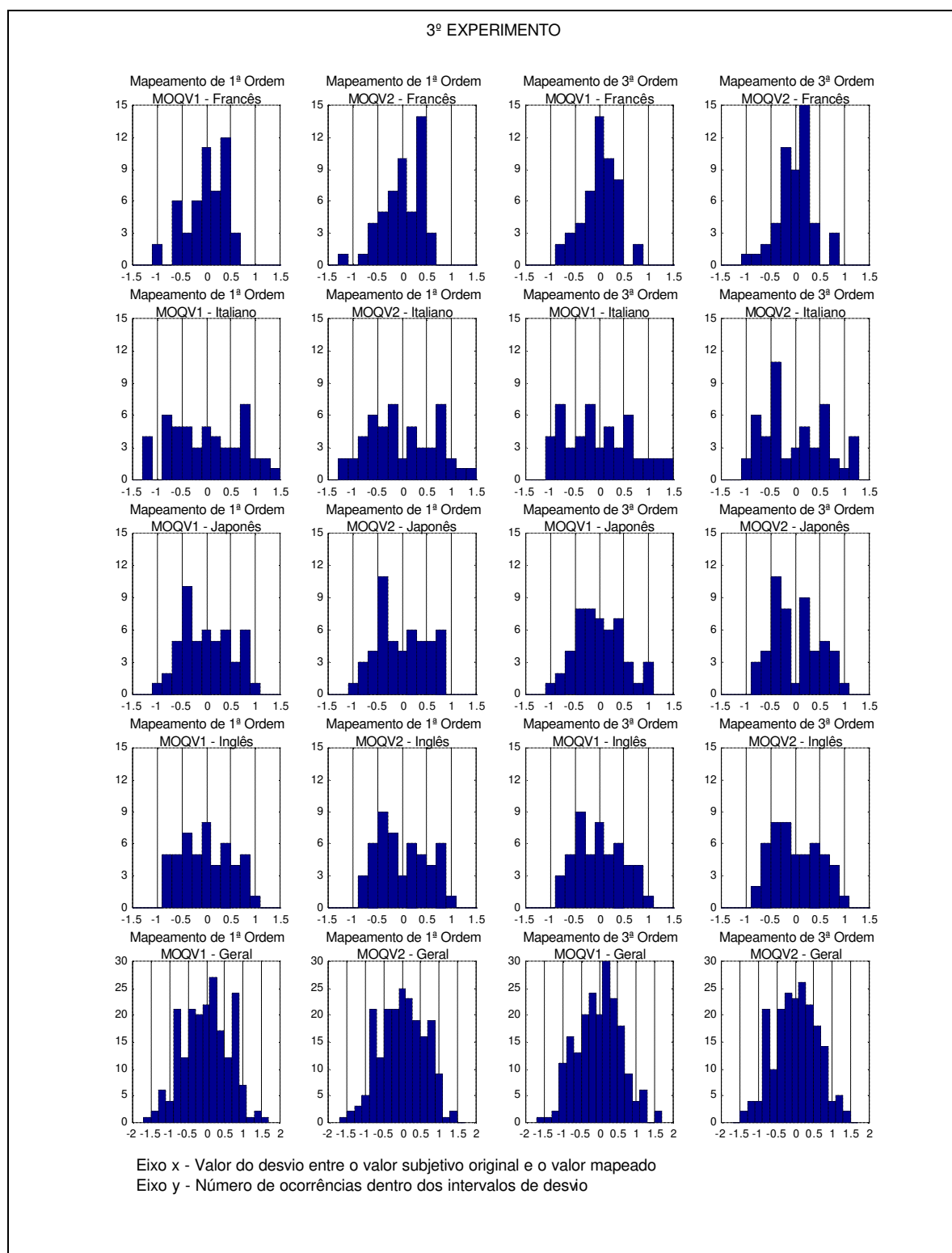


Figura 6.30 - Histogramas caracterizando os desvios entre os valores subjetivos e valores mapeados

Tabela 6.19 - Limites de desvio para determinados intervalos de confiança (1º experimento)

			Ordem de Mapeamento e Desvio Máximo					
Língua	Tipo	Int. Conf.	1ª	2ª	3ª	4ª	5ª	6ª
Francês	MOQV1	90 %	0,5163	0,5227	0,4455	0,4337	0,4382	0,4331
		95 %	0,6194	0,6273	0,5347	0,5207	0,5261	0,5201
		99 %	0,8281	0,8390	0,7155	0,6971	0,7047	0,6970
	MOQV2	90 %	0,5075	0,5012	0,4584	0,4384	0,4487	0,4553
		95 %	0,6089	0,6015	0,5502	0,5262	0,5387	0,5468
		99 %	0,8140	0,8046	0,7362	0,7045	0,7216	0,7328
Japonês	MOQV1	90 %	0,4174	0,4031	0,3791	0,3701	0,3753	0,3759
		95 %	0,5008	0,4838	0,4551	0,4444	0,4506	0,4515
		99 %	0,6696	0,6470	0,6089	0,5949	0,6035	0,6051
	MOQV2	90 %	0,4249	0,4033	0,4019	0,4016	0,4097	0,4182
		95 %	0,5098	0,4840	0,4824	0,4822	0,4919	0,5023
		99 %	0,6816	0,6474	0,6456	0,6455	0,6589	0,6731
Inglês	MOQV1	90 %	0,4978	0,4866	0,4030	0,3359	0,3419	0,3198
		95 %	0,5973	0,5839	0,4837	0,4033	0,4106	0,3841
		99 %	0,7985	0,7810	0,6473	0,5399	0,5499	0,5147
	MOQV2	90 %	0,4532	0,4006	0,3850	0,3344	0,3411	0,3185
		95 %	0,5438	0,4808	0,4621	0,4014	0,4096	0,3825
		99 %	0,7270	0,6431	0,6184	0,5374	0,5487	0,5127
Geral	MOQV1	90 %	0,6060	0,6062	0,5876	0,5894	0,5938	0,5982
		95 %	0,7237	0,7239	0,7017	0,7040	0,7091	0,7145
		99 %	0,9562	0,9566	0,9273	0,9303	0,9372	0,9443
	MOQV2	90 %	0,6227	0,6221	0,6204	0,6247	0,6293	0,6331
		95 %	0,7436	0,7429	0,7409	0,7461	0,7516	0,7562
		99 %	0,9826	0,9817	0,9791	0,9860	0,9933	0,9994

Tabela 6.20 - Limites de desvio para determinados intervalos de confiança (2º experimento)

			Ordem de Mapeamento e Desvio Máximo					
Língua	Tipo	Int. Conf.	1ª	2ª	3ª	4ª	5ª	6ª
Francês	MOQV1	90 %	0,4476	0,4267	0,4279	0,4349	0,4472	0,4516
		95 %	0,5380	0,5130	0,5146	0,5231	0,5381	0,5436
		99 %	0,7222	0,6892	0,6918	0,7038	0,7246	0,7326
	MOQV2	90 %	0,4488	0,4197	0,4303	0,4332	0,4448	0,4459
		95 %	0,5394	0,5046	0,5175	0,5211	0,5352	0,5367
		99 %	0,7242	0,6779	0,6957	0,7011	0,7207	0,7233
Japonês	MOQV1	90 %	0,3687	0,3665	0,3233	0,3285	0,3293	0,3253
		95 %	0,4427	0,4402	0,3884	0,3947	0,3958	0,3910
		99 %	0,5930	0,5899	0,5208	0,5296	0,5313	0,5253
	MOQV2	90 %	0,3625	0,3547	0,3298	0,3344	0,3378	0,3312
		95 %	0,4353	0,4260	0,3961	0,4018	0,4060	0,3982
		99 %	0,5830	0,5709	0,5312	0,5390	0,5450	0,5350
Inglês	MOQV1	90 %	0,3892	0,3645	0,3124	0,3205	0,3267	0,3339
		95 %	0,4674	0,4377	0,3752	0,3851	0,3927	0,4014
		99 %	0,6260	0,5866	0,5032	0,5166	0,5272	0,5392
	MOQV2	90 %	0,3876	0,3508	0,3265	0,3349	0,3407	0,3400
		95 %	0,4654	0,4213	0,3922	0,4024	0,4095	0,4087
		99 %	0,6234	0,5647	0,5260	0,5399	0,5498	0,5491
Geral	MOQV1	90 %	0,4011	0,3830	0,3521	0,3550	0,3576	0,3581
		95 %	0,4791	0,4576	0,4206	0,4241	0,4272	0,4278
		99 %	0,6335	0,6051	0,5563	0,5610	0,5651	0,5659
	MOQV2	90 %	0,3968	0,3708	0,3561	0,3586	0,3610	0,3629
		95 %	0,4741	0,4430	0,4255	0,4285	0,4313	0,4336
		99 %	0,6269	0,5858	0,5627	0,5667	0,5705	0,5736

Tabela 6.21 - Limites de desvio para determinados intervalos de confiança (3º experimento)

Língua	Tipo	Int. Conf.	Ordem de Mapeamento e Desvio Máximo					
			1ª	2ª	3ª	4ª	5ª	6ª
Francês	MOQV1	90%	0,7121	0,6590	0,6300	0,6174	0,6292	0,6394
		95%	0,8536	0,7901	0,7554	0,7404	0,7547	0,7670
		99%	1,1387	1,0543	1,0084	0,9887	1,0082	1,0251
	MOQV2	90%	0,6844	0,6678	0,6355	0,6483	0,6510	0,6645
		95%	0,8204	0,8006	0,7621	0,7775	0,7809	0,7972
		99%	1,0944	1,0684	1,0173	1,0382	1,0432	1,0654
Italiano	MOQV1	90%	1,2407	1,2325	1,2397	1,2602	1,2665	1,2769
		95%	1,4874	1,4777	1,4865	1,5113	1,5191	1,5318
		99%	1,9841	1,9719	1,9844	2,0182	2,0293	2,0471
	MOQV2	90%	1,1680	1,1269	1,1380	1,1606	1,1587	1,1828
		95%	1,4002	1,3510	1,3646	1,3919	1,3899	1,4190
		99%	1,8678	1,8029	1,8216	1,8587	1,8567	1,8963
Japonês	MOQV1	90%	0,8948	0,8942	0,8763	0,8914	0,9036	0,9068
		95%	1,0727	1,0720	1,0508	1,0690	1,0839	1,0878
		99%	1,4309	1,4306	1,4027	1,4275	1,4479	1,4537
	MOQV2	90%	0,8635	0,8708	0,8716	0,8754	0,8708	0,8889
		95%	1,0351	1,0441	1,0452	1,0498	1,0445	1,0663
		99%	1,3809	1,3933	1,3952	1,4019	1,3953	1,4251
Inglês	MOQV1	90%	0,8900	0,8826	0,8762	0,8873	0,9008	0,9106
		95%	1,0669	1,0582	1,0506	1,0641	1,0804	1,0924
		99%	1,4233	1,4120	1,4025	1,4210	1,4433	1,4599
	MOQV2	90%	0,8713	0,8868	0,8620	0,8796	0,8976	0,9106
		95%	1,0446	1,0632	1,0336	1,0549	1,0766	1,0924
		99%	1,3935	1,4188	1,3798	1,4087	1,4383	1,4598
Geral	MOQV1	90%	1,0498	1,0323	1,0117	1,0107	1,0093	1,0029
		95%	1,2528	1,2319	1,2073	1,2061	1,2044	1,1967
		99%	1,6523	1,6247	1,5924	1,5908	1,5886	1,5785
	MOQV2	90%	1,0149	1,0059	0,9855	0,9903	0,9927	0,9920
		95%	1,2111	1,2004	1,1761	1,1818	1,1847	1,1837
		99%	1,5973	1,5832	1,5512	1,5588	1,5626	1,5614

6.4.1. Comparação entre as correlações obtidas

A tabela 6.22 mostra uma comparação entre os valores aqui obtidos para a correlação e RMSE e os valores correspondentes obtidos pela Ericsson [63]. As comparações foram feitas apenas para MOQV1 (este foi o único caso publicado pela Ericsson) e considerando o mapeamento de 1ª ordem.

Tabela 6.22 - Comparação entre os valores obtidos nas simulações e os obtidos pela Ericsson

<i>Experiência</i>	<i>Correlação Obtida</i>	<i>Correlação Ericsson</i>	<i>RMSE Obtida</i>	<i>RMSE Ericsson</i>
Exp. 1 (francês)	0,920	0,91	0,293	0,30
Exp. 1 (japonês)	0,927	0,93	0,237	0,24
Exp. 1 (inglês)	0,928	0,93	0,283	0,29
Exp. 2 (francês)	0,922	-	0,286	-
Exp. 2 (japonês)	0,938	0,93	0,217	0,22
Exp. 2 (inglês)	0,930	0,91	0,218	0,24

Como se pode ver, os valores ficaram bastante próximos, ocorrendo apenas algumas pequenas diferenças. A Ericsson não fez o teste para a língua francesa no segundo experimento devido à presença dos arquivos defeituosos.

A tabela 6.23 mostra a comparação dos resultados obtidos (MOQV1 e MOQV2) com aqueles encontrados pela KPN [62], de acordo com o segundo critério.

Tabela 6.23 - Comparação entre os valores obtidos nas simulações e os obtidos pela KPN

<i>Experiência</i>	<i>Correlação Obtida (MOQV1)</i>	<i>Correlação Obtida (MOQV2)</i>	<i>Correlação KPN (PSQM)</i>	<i>Correlação KPN (PSQM+)</i>
Exp. 1 (francês)	0,918	0,920	0,91	0,92
Exp. 1 (japonês)	0,863	0,844	0,87	0,86
Exp. 1 (inglês)	0,859	0,833	0,86	0,83
Exp. 2 (francês)	0,922	0,922	-	-
Exp. 2 (japonês)	0,937	0,940	0,92	0,96
Exp. 2 (inglês)	0,929	0,929	0,93	0,96

Novamente, os resultados são bastante próximos, especialmente para o primeiro experimento. Há algumas discrepâncias para o experimento 2, não muito acentuadas, o que pode ser explicado pelas diferenças de implementação entre o método MOQV e o PSQM.

6.5. CONCLUSÕES

Apresenta-se aqui as conclusões resultantes da análise de todos os testes realizados, resumindo as informações necessárias para o uso adequado do método MOQV frente às diversas situações encontradas na prática:

1 - O primeiro critério (mapeamento individual para cada língua) é o mais adequado, pois se adapta melhor às peculiaridades de cada língua e cultura. Portanto, cada língua deveria contar com uma curva de mapeamento. Porém, quando não se dispõe de dados adequados para uma determinada língua, pode-se utilizar um mapeamento geral, como aquele aqui calculado utilizando-se os dados para as línguas inglesa, japonesa e francesa. Entretanto, é preciso ter presente que tal mapeamento não oferece os melhores resultados. Por isso, sugere-se que ou sejam feitos experimentos para obter o melhor mapeamento para a língua específica, ou faça-se uso dos valores MOQV brutos, ou seja, analise-se a qualidade diretamente a partir dos valores MOQV obtidos, sem a realização do mapeamento. Esta opção está disponível no programa desenvolvido e é condicionada ao grau de experiência apresentado pelo usuário.

2 - Como regra geral, a correlação entre as medidas objetivas mapeadas e as medidas subjetivas correspondentes, cresce à medida que a ordem do polinômio de mapeamento cresce desde a unidade até 6. Entretanto, a monotonicidade é uma característica desejável neste tipo de mapeamento, e os resultados mostram que ordens de mapeamento acima da terceira introduzem, em certos casos, grandes quebras na monotonicidade. Assim, a ordem de mapeamento mais apropriada para a maioria dos tipos de arquivos é a terceira, fato este observado para todas as línguas presentes nos testes. Isto ocorre devido à sua capacidade de modelar satisfatoriamente o comportamento aproximado dos avaliadores nos testes subjetivos, ou seja, representa bem a tendência de perda de linearidade nos extremos de qualidade (sinal muito bom ou muito ruim), pois os ouvintes tendem a saturar a avaliação nestes pontos. O ganho na correlação representado pela passagem da primeira para a terceira ordem chegou a 0,030 no caso do primeiro e segundo experimentos e 0,065 no caso do terceiro. Os ganhos observados na análise dos intervalos de confiança também se mostraram bastante expressivos, diminuindo os limites de desvio do valor correto em mais de 0,1 em muitos casos. As pequenas quebras na monotonicidade ainda observadas para este caso produziram, no resultado final, um efeito que pode ser considerado desprezível, devido a três fatores: as regiões com quebra de monotonicidade são relativamente

pequenas; os casos testados que apresentaram valores objetivos nestas regiões são em pequeno número; os casos testados que utilizaram estas regiões produziram resultados mapeados próximos dos valores subjetivos correspondentes.

3 - Como já foi comentado, seria desejável um mapeamento individual para cada língua. Uma consideração importante é o fato de que cada país possui características culturais que se refletem na determinação das curvas de mapeamento ótimas. Dentre as línguas analisadas, pode-se perceber claramente que os japoneses são os mais exigentes, seguidos pelos franceses e, com um grau de exigência bem menor, os norte-americanos.

4 - Para a determinação de um mapeamento adequado às peculiaridades da língua portuguesa, em particular aquela falada no Brasil, seria necessário um banco de dados com arquivos de voz e valores de medidas subjetivas. Na ausência de tais dados, a utilização do mapeamento geral pode ser suficiente para grande parte das aplicações desejadas. Alternativamente, pode-se utilizar diretamente o valor MOQV.

5 - Pela análise dos gráficos, pode-se afirmar que o MOQV funcionou adequadamente para os dois primeiros experimentos, mostrando que este método se adapta bem a casos em que o sinal é submetido a um ou mais tipos de codecs, bem como a situações em que há a presença de ruído de fundo. Por outro lado, os resultados para o terceiro experimento foram ruins, denotando uma clara inadequação a situações em que os sinais estão sujeitos a condições de erro severas. Neste caso, as correlações foram sempre ruins, especialmente para as línguas japonesa e italiana. No caso de as condições de erro não serem muito severas (abaixo de 3%), o MOQV pode ser utilizado, desde que observadas as limitações deste procedimento, ou seja, os resultados revelarão apenas uma tendência do desempenho do codec, e não resultados precisos. Esta alternativa, contudo, fica condicionada à aplicação desejada para o codec a ser avaliado.

6 - A escolha entre os valores MOQV1 e MOQV2 depende, basicamente, das condições predominantes a que serão submetidos os sinais a serem processados pelos codecs. Como se pode ver pelas tabelas, o valor MOQV1 é mais apropriado para os casos em que os sinais apresentam pouca distorção ou erros, como ocorre nos dois primeiros experimentos. A adoção do valor MOQV2 está então condicionada a casos em que o sinal está submetido a condições severas de distorção ou erro, como é o caso do terceiro experimento, especialmente para as ordens de mapeamento mais baixas, onde a diferença entre os dois valores pode ser bastante expressiva (até 0,1). Por fim, conforme discutido na seção 4.3, os valores MOQV2 devem ser utilizados quando ocorrem supressão de trechos de voz e/ou distorção por excesso de volume.

7 - De uma maneira geral, as restrições na utilização do MOQV dependem fortemente da aplicação desejada para o codec a ser avaliado, bem como da qualidade da avaliação que se deseja fazer. No caso de codecs a serem utilizados em aplicações menos exigentes, como comunicação militar ou o VoIP (telefonia sob protocolo da Internet), por exemplo, este programa pode ser utilizado sem restrições importantes. Já no caso de aplicações mais exigentes, em que a qualidade é um fator muito importante e decisivo e há a necessidade de uma avaliação mais precisa, pode ser necessário se fazer uso das avaliações subjetivas. Nestes casos, contudo, o método MOQV pode ser utilizado com o objetivo de fornecer uma tendência para o comportamento dos codecs testados, o que pode ser valioso em situações em que não se dispõe de tempo e/ou recursos suficientes para a realização de testes subjetivos exaustivos.

CAPÍTULO 7

CONCLUSÃO

Apresentou-se neste trabalho uma medida objetiva de qualidade de codecs de voz para telefonia baseada no método PSQM, incorporando, ao mesmo tempo, características que melhoram seu desempenho e várias opções de curvas de mapeamento para os valores subjetivos. O programa de computador implementado incorporou recursos de manipulação dos sinais, formando um produto completo que facilita e agiliza o uso do método. Foram descritas suas características básicas e apresentados os resultados alcançados em testes exaustivos, utilizando o banco de dados ITU Supplement 23, o que permitiu traçar um perfil da aplicabilidade do método.

Inicialmente, com o objetivo de situar o leitor em relação aos tipos de testes e procedimentos que vêm sendo adotados na avaliação de codecs de voz, foi feita uma breve descrição das medidas subjetivas existentes. Foram apresentadas também as principais classes de medidas objetivas, as quais definem os tipos de abordagens utilizadas no desenvolvimento dos métodos propostos.

A seguir, foi feita a apresentação dos modelos psicoacústicos, onde descreveu-se a estrutura anatômica do ouvido humano e suas principais características de percepção acústica, permitindo assim extrair uma série de conceitos importantes para a determinação de um modelo matemático eficiente, além de fornecer subsídios para um melhor entendimento dos capítulos subsequentes.

O próximo passo foi a descrição detalhada das diversas características do método PSQM recomendado pela ITU. Foram também descritos os procedimentos e cuidados a serem tomados na realização das medidas e as aplicações para as quais pode-se utilizá-lo eficazmente. Os processamentos que compõem o método PSQM foram explicados e justificados através da teoria apresentada na modelagem matemática do ouvido humano. Apresentou-se também um aperfeiçoamento (PSQM+) que permitiu a este método adaptar-se melhor a situações em que os sinais de voz contenham supressão de trechos de voz e/ou distorções provocadas por excesso de volume.

A seguir, foram apresentados vários métodos propostos como alternativa ao PSQM. Comparando seus desempenhos globais, concluiu-se que nenhum deles apresenta resultados superiores ao PSQM. Contudo, tais métodos apresentam algumas qualidades promissoras, as quais poderão ser incorporadas a futuras propostas, já que existe hoje uma forte tendência em nível internacional nesta direção.

Foi feita a seguir a descrição do programa de computador desenvolvido para implementar o método MOQV aqui proposto. Tal programa, implementado em Matlab®, incorporou diversas facilidades não encontradas naquele sugerido para o método PSQM: alinhamento temporal dos sinais de voz original e decodificado via correlação cruzada; opção de várias taxas de amostragem; detalhamento das características dos sinais de voz, como sua visualização, audição e energia média; ajuste da energia dos sinais; diversas opções de mapeamento para os valores subjetivos e incorporação da proposta PSQM+, visando a melhora de seu desempenho. Tais melhorias resultaram numa maior

interatividade, tornando seu uso adequado tanto para usuários com larga experiência na área quanto para iniciantes.

Por fim, foram realizados diversos testes, utilizando o banco de dados “Supplement 23” da ITU, constituído por arquivos de voz em inglês, francês e japonês, utilizados nos testes subjetivos de vários codecs submetidos a algumas condições de teste. Tal material foi dividido em três grupos de experimentos: no primeiro, os sinais de voz presentes foram submetidos a diversos tipos de codecs padrão da ITU e de telefonia móvel; no segundo, os sinais de voz foram submetidos a diversos tipos de ruídos de ambiente; no terceiro, os sinais codificados foram degradados por erros durante sua transmissão.

Os testes abordaram o maior número de situações possíveis, a fim de se traçar um perfil do desempenho do método MOQV. Para isso, considerou-se:

- as características de cada um dos três experimentos submetidos aos testes;
- mapeamento para cada língua e mapeamento geral médio em relação a todas as línguas presentes no banco de dados;
- a influência da ordem polinomial p da curva de mapeamento.

Os resultados gerais destes testes indicaram que:

- o método MOQV funcionou bem para casos em que o sinal é submetido a um ou mais tipos de codecs, bem como a situações em que há a presença de ruído de fundo, sendo porém inadequado na presença de erros de transmissão;
- deve-se utilizar mapeamentos de 1ª a 3ª ordem, sendo esta última a mais adequada para a maioria das aplicações;
- quando possível, deve-se realizar um mapeamento individual para a língua em questão a fim de se respeitar suas peculiaridades; não sendo possível, pode-se utilizar o mapeamento geral ou o valor MOQV não mapeado;
- devido à falta de dados, não foi possível determinar um mapeamento para a língua portuguesa falada no Brasil; para a maioria das aplicações, porém, o mapeamento geral é suficiente.

Todos os resultados obtidos são também válidos, de forma aproximada, para o método PSQM, e indicam a necessidade da realização de mais pesquisas, a fim de que se possa ampliar a aplicabilidade e melhorar o desempenho de ambos os métodos. Um novo método, conhecido como PESQ (*Perceptual Evaluation of Speech Quality*) [64], promete solucionar vários dos problemas inerentes aos métodos MOQV e PSQM. Este novo modelo, que é, aparentemente, uma compilação dos pontos mais fortes encontrados nos métodos PSQM e PAMS, deverá dar origem a uma nova recomendação da ITU, a Rec. P.862, substituindo a Rec. P.861 a partir do ano de 2001, quando serão então divulgados maiores detalhes.

Por fim, foram acrescentadas, na forma de Apêndices, algumas informações complementares ao trabalho desenvolvido. Foram apresentados diversos conceitos relacionados à avaliação objetiva de sinais de voz com qualidade variável. Houve também a definição de alguns termos cujo uso ao longo do texto demandou um maior grau de detalhamento. Finalmente, foram apresentadas algumas características importantes das avaliações objetivas de qualidade de codecs de áudio, através de uma breve descrição do mais bem sucedido método dos últimos anos, o qual recebeu o nome de PEAQ. Este assunto está intimamente ligado à avaliação objetiva de qualidade de voz. Porém, há neste caso uma maior complexidade e quantidade de conceitos envolvidos. Por esse motivo, o estudo das avaliações objetivas aplicadas a codecs de áudio representa a sequência natural dos estudos realizados neste trabalho, o que deverá dar origem a um projeto de doutorado.

BIBLIOGRAFIA

- [1] ITU-T Recommendation P.861, *Objective quality measurement of telephone-band (300 - 3400) speech codecs*, 1996.
- [2] Royal PTT, Netherlands, *Measuring the quality of audio devices*, CCITT Contribution COM XII-114, Geneva, December 1991.
- [3] Beerends, J.G., Stemerdink, J.A., *A Perceptual Speech-Quality Measure Based on a Psychoacoustic Sound Representation*, J. Audio Eng. Soc., Vol. 42, No. 3, pp. 115-123, March 1994.
- [4] Irii, H., "Comparison of Four Objective Speech Quality Assessment Methods Based on International Subjective Evaluations of Universal Codecs", IEEE International Conference on Communications, June 1991.
- [5] ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, 1996.
- [6] ITU-T Recommendation P.830, *Subjective performance assessment of telephone-band and wideband digital codecs*, 1996.
- [7] Campos Neto, S. F. e Violaro, F., *Metodologias de Avaliação de Algoritmos de Codificação de Voz*, Tese de Mestrado, Unicamp, Julho de 1993.
- [8] South, C.R. et al, *Subjective Test Methodology for a 16 kbit/s speech coder*, Contribution SQ-10.89R, Study Group XII Speech Quality Experts Group, March 1 1990.
- [9] South, C.R., *Phase II Subjective Test Methodology for a 16 kbit/s speech coder*, Contribution SQ-12.91, Study Group XII, April 30 1991.
- [10] CCITT Recommendation P.82, *Method for evaluation of service from the standpoint of transmission speech quality*, 1994.
- [11] ITU-T Recommendation P.50, *Artificial Voices*, 1993.
- [12] Noll, P., *Adaptive Quantizing in Speech Coding Systems*, In Int. Zurich Seminar on Digital Comm., pp. B3.1-B3.6, IEEE, 1974.
- [13] Gray, A.H. and Markel, J.D., *Distance Measures for Speech Processing*, IEEE Trans. On Acoustic, Speech and Signal Processing, ASSP, vol. 24, no. 5, Oct 1976.
- [14] Wang, S.; Sekey, A. and Gersho, A., *An Objective Measure for Predicting Subjective Quality of Speech Coders*, IEEE Journal on Selected Areas in Communications, vol. 10, no. 5, pp. 819-829, June 1992.

- [15] Itoh, K.; Kitawaki, N. and Takehi, K., *Objective Quality Measures for Speech Waveform Coding System*, Review of the Electrical Communication Laboratories, vol. 32, no. 2, pp. 220-228, 1984.
- [16] Crochière, R.E., Rabiner, L.R., Jayant, N.S., Tribolet, J.M., *A Study of Objective Measures for Speech Waveforms Coders*, In Int. Zurich Seminar on Digital Commun., pp. H1.1-H1.7, IEEE, March 1978.
- [17] Quackenbush, S.R.; III Barnwell, T.P. and Clements, M.A., *Objective Measures of Speech Quality*, Signal Processing Series, Prentice-Hall, 1988.
- [18] Crochière, R.E., *An analysis of 16 kb/s sub-band coder performance: dynamic range, tandem connections, and channel errors*, Bell Syst. Tech. J., vol. 57, no. 8, pp. 2927-2952, Oct. 1978.
- [19] Mermelstein, P., *Evaluation of segmental SNR measure as an indicator of the quality of ADPCM coded speech*, J. Acoust. Soc. Am., vol. 66, no. 6, pp. 1664-1667, 1979.
- [20] Goodman, D.J., Scagliola, C., Crochière, R.E., Rabiner, L.R., and Goodman, J., *Objective and subjective performance of tandem connections of waveform coders with an LPC vocoder*, Bell Syst. Tech. J., vol. 58, no. 3, pp. 601-629, Mar 1979.
- [21] French, N.R. and Steinberger, J.C., *Factors Governing the Intelligibility of Speech Sounds*, J. Acoust. Soc. Am., vol. 19, pp. 90-119, 1947.
- [22] Steeneken, H.J.M. and Houtgast, T., *A Physical Method for Measuring Speech-Transmission Quality*, J. Acoust. Soc. Am., vol. 67, pp. 318-326, 1980.
- [23] Hansen, M., *Thesis: Assessment and prediction of speech transmission quality with an auditory processing model*, Oldenburg, Germany, May 1998
- [24] Oppenheim, A.V. e Schafer, R.W., *Discrete Time Signal Processing*, Prentice Hall, New Jersey, 1989.
- [25] NTT, *Transmission performance objective evaluation model for fundamental factors*, CCITT Contribution COM XII-174, November 1983.
- [26] Goodman, D.J., Scagliola, C., Crochière, R.E., Rabiner, L.R., Goodman, J., *Objective and Subjective Performance of Tandem Connections of Waveform Coders with an LPC Vocoder*, Bell Syst. Tech. J., vol. 58, no. 3, pp. 601-629, Mar 1979.
- [27] *Models for Predicting Transmission Quality from Objective Measurements*, series P. Recomm. Suppl. 3, CCITT, 1989.
- [28] Lalou, J., *The Information Index: an Objective Measure of Speech Transmission Performance*, Annales des Télécommunications, vol. 45, no. 1-2, pp. 47-65, CNET/France, 1990.

- [29] *Objective Evaluation of Non-Linear Distortion Effects on Voice Transmission Quality*, Contribution 46, CCITT Study Group XII, March 1986.
- [30] BNR, *Evaluation of non-linear distortion via the coherence function*, CCITT Contribution COM XII-60, April 1982.
- [31] Kubichek, R.F., Quincy, E.A., Kiser, K.L., *Speech Quality Assessment Using Expert Pattern Recognition Techniques*, Proceedings of the IEEE Pacific Rim Conference on Computers, Communication, and Signal Processing, June 1989.
- [32] Guydon, A.C., *Fisiologia Humana*, Ed. Guanabara, 6^a Edição, Rio de Janeiro, 1985.
- [33] Fletcher, H., *Speech and Hearing in Communication*, D. Van Nostrand Co., Toronto, 1953.
- [34] Gulick, W.L., *Hearing - Physiology and Psychophysics*, Oxford Univ. Press, 1971.
- [35] MacKay, I.R.A., *Phonetics: The Science of Speech Production*, College-Hill Publication Little, Brown and Co., 1987.
- [36] Oliveira, J.A.A., *O mecanismo eletrobiomecânico ativo da cóclea*, Rev. Bras. de Otorrinolaringologia, vol. 59, no. 4, 1993.
- [37] Bittencourt, R., *MPEG-Audio*, USP, 1997 (<http://www.lsi.usp.br/~ricardo/mpeg/>).
- [38] Robertson, G., *Critical Bands*, 1996 (<http://tuba.music.gla.ac.uk/~george/audio/psy/psy.html>).
- [39] Fourcin, A.J. et al, *Speech processing by man and machine: Group Report*, pp. 307-351, Bullock, T., 1977.
- [40] Zwicker, E., Zwicker, U.T., *Audio engineering and psychoacoustics: Matching signals to the final receiver, the human auditory system*, J. Audio Engineering Society, vol. 39, no. 3, pp. 115-126, March 1991.
- [41] Hermansky, H., *Perceptual linear predictive (PLP) analysis of speech*, J. Acoust. Soc. Am., vol. 87, no. 4, pp. 88-94, April 1990.
- [42] Atkinson, D.J., *Proposed Annex to Recommendation P.861*, NTIA, ITU Study Group 12 - Contribution COM 12-24-E, December 1997.
- [43] Rix, A., Hollier, M., *Comparison of Speech Quality Assessment Algorithms: BT PAMS, PSQM, PSQM+ and MNB*, BT, ITU Study Group 12, Delayed Contribution D.80, December 1998.
- [44] Berger, J., *TOSQA - Telecommunication Objective Speech Quality Assessment*, ITU Study Group 12, Contribution COM12-34-E, December 1997.

- [45] Quinlan, D., *Subjective Experiments on ISO 532 B*, Proceedings of Inter-Noise 92, p1003, 1992.
- [46] Moore, B., *Introduction to the Psychology of Hearing*, Academic Press, London, 1989.
- [47] Stevens, S.S., *Perceived Level of Noise by Mark VII and Decibels (E)*; JASA 51, no. 2, p575, 1972.
- [48] ITU-T Recommendation P.810, *Modulated Noise Reference Unit (MNRU)*, 1996.
- [49] KPN, *Improvement of the P.861 Perceptual Speech Quality Measure*, The Netherlands, December 1997.
- [50] Beerends, J.G., *Modelling Cognitive Effects that Play a Role in the Perception of Speech Quality*, Speech Quality Assessment, Workshop papers, Bochum, pp. 1-9, November 1994.
- [51] Voran, S., *Objective Estimation of Perceived Speech Quality - Development of the Measuring Normalizing Block Technique*, IEEE Trans. on Speech and Audio Processing, vol. 7, no. 4, pp. 371-390, July 1999.
- [52] Voran, S., *Objective Estimation of Perceived Speech Using Measuring Normalizing Blocks*, NTIA Report 98-347, April 1998.
- [53] Voran, S., *Estimation of Perceived Speech Using Measuring Normalizing Blocks*, Proceedings of the 1997 IEEE Workshop on Speech Coding, September 1997.
- [54] Rosenbluth, J.H., Perkins, M.E., *Comparison of Two Methods for Objective Assessment of the Quality of Coded Speech*, AT&T, ITU Study Group 12, Contribution COM 12-26-E, December 1997.
- [55] Atkinson, D.J., *Summary of Available Performance Results for the MNB Objective Voice Quality Measurement Algorithm*, NTIA, ITU Study Group 12, Contribution COM-25-E, December 1997.
- [56] Press, W.H. et al, *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, Cambridge, Massachusetts, 1988.
- [57] Otnes, R.K., Enochson, L., *Applied time series analyses*, vol. 1, pp. 277-295, John Wiley and Sons, 1978.
- [58] Jenkins, G.M., Watts, D.G., *Spectral Analysis and its Applications*, Holden-Day, San Francisco, 1968.
- [59] *Subjective test plan for characterization of an 8 kbit/s speech codec*, ITU-T Study Group 12, Speech Quality Experts Group, Issue 2.0, 25 September 1995.

- [60] Guthrie, W., *Engineering Statistics Handbook*, Process Modeling - Data Analysis for Process Modeling, 1999.
- [61] Rix, A., Hollier, M., *Performance Metrics for Objective Quality Assessment Systems in Telephony*, ITU Study Group 12 - Delayed Contribution D.79, November 1998.
- [62] Beerends, J.G., Hekstra, A.P., *Comparison of the ITU-T P.861 PSQM, PSQM+ and MNB Objective Speech Quality Measurement Methods*, ITU Study Group 12, Contribution COM 12-58-E, September 1998.
- [63] Folkesson, M., Karlsson, A., *Results of Processing the Supplement 23 Speech Database for Development of the Extended P.861*, Ericsson, ITU Study Group 12, Delayed Contribution D.89, November 1998.
- [64] Rix, A.W., Hekstra, A.P., Beerends, J.G. and Hollier, M.P., *Performance of the Integrated KPN/BT Objective Speech Quality Assessment Model*, ITU Study Group 12, Contribution COM12-D136, May 2000.
- [65] ETSI, TM/TM5/TCH-HS, *Global Analysis of Selection Tests: Basic Data*, Technical Report 91/74, ETSI, 1991.
- [66] ETSI, TM/TM5/TCH-HS, *Selection Test Phase II: Listening Test Results with German Speech Samples*, Technical Report 92/35, FI/DBP- Telekom. Experiment 1, IM 4, 1992.
- [67] Sotscheck, J., Sätze für Sprachg ütemessungen und ihre phonologische Anpassung an die deutsche Sprache, In Fortschritte der Akustik - DAGA '84, pp. 873-876, 1984.
- [68] ITU-R Recommendation BS-1116-1, *Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems*, 1997.
- [69] ITU-R Recommendation BS-1387, *Method for objective measurements of perceived audio quality*, 1998.
- [70] Thiede, T., Kabot, E., *A New Perceptual Quality Measure for Bit Rate Reduced Audio*, Contribution to the 100th AES Convention, preprint 4280, Copenhagen, 1996.
- [71] Brandenburg, K., *Evaluation of quality for audio encoding at low bit rates*, Contribution to the 82nd AES Convention, preprint 2433, London, 1987.
- [72] Sporer, T., *Objective audio signal evaluation - applied psychoacoustics for modeling the perceived quality of digital audio*, 103rd AES-Convention, New York, preprint 4512, October, 1997.
- [73] Beerends, J.G., Stemerdink, J.A., *A perceptual audio quality measure based on a psychoacoustic sound representation*, J. Audio Eng. Soc., vol. 40, pp. 963-978, Dec. 1992.

- [74] Paillard, B., Mabillean, P., Morisette, S., Soumagne, J., *Perceval: Perceptual evaluation of the quality of audio signals*, J. Audio Eng. Soc., vol. 40, pp. 21-31, Jan. 1992.
- [75] Colomes, C., Lever, M., Rault, J.B., Dehery, Y.F., *A perceptual model applied to audio bit-rate reduction*, J. Audio Eng. Soc., vol. 43, pp. 233-240, April 1995.
- [76] Recommendation P.54, *Sound level meters (apparatus for the objective measurement of room noise)*, 1988.
- [77] Terhardt, E. , *Akustische Kommunikation - Grundlagen mit Hörbeispielen*. Springer, Berlin/Heidelberg, 505 S. ISBN 3-540-63408-8, 1998.
- [78] Hartmann, W.M., *Pitch, periodicity, and auditory organization*, J. Acoust. Soc. Am. 100, 3491-3502, 1996.
- [79] Idson, W.L., Massaro, D.W., *A bidimensional model of pitch in the recognition of melodies*, Perceptual Psychophysics, vol. 24, pp. 551-565, 1978.
- [80] *IAC Noise Control Reference Handbook*, 1989.
- [81] Beranek, L.L., *Acústica*, Editorial HASA, Buenos Aires, 1961.
- [82] CCITT Contribution COM XII-79, *Specification for an Intermediate Reference System*, Study Period 1973-1976.
- [83] CCITT Contribution COM XII-104, *Recapitulation and Analysis of the Results of Subjective and Objective Loudness Rating Measurements Carried Out with Eleven Telephone Systems by the CCITT Laboratory*, Study Period 1973-1976.
- [84] ITU-T Recommendation P.48, *Specification for an Intermediate Reference System*, 1988.
- [85] Richard, D.L., *Telecommunication by Speech*; John Wiley, NY and Butterworths, London, 1973.

APÊNDICE A

MEDIDAS OBJETIVAS DE QUALIDADE EM TELEFONIA MÓVEL

Uma das questões recorrentes na utilização de medidas objetivas de qualidade de voz é a sua aplicação na telefonia móvel, onde a qualidade do material varia continuamente com o tempo. Neste apêndice será descrito, de maneira resumida, um método de avaliação objetiva discutido por Hansen [23] no Capítulo 4 de sua tese de doutorado. Para uma melhor caracterização do problema, será discutido também o comportamento dos ouvintes durante uma avaliação subjetiva deste tipo.

A.1. INTRODUÇÃO

A avaliação da qualidade de sinais de voz em sistemas com condições de transmissão variantes com o tempo suscita algumas questões: existe o fenômeno de percepção instantânea de qualidade de trechos de curta duração? Se existe, como é sua classificação pelos ouvintes? Como é a impressão de qualidade global relacionada à variação da qualidade instantânea de curta duração?

Para responder a estas questões, no trabalho de tese mencionado foram realizados dois diferentes experimentos: o primeiro utilizando elementos de voz curtos e isolados, cada um recebendo uma avaliação separada, e o segundo utilizando uma longa sequência de sentenças, havendo aqui uma avaliação contínua.

A.2. CONFIGURAÇÃO E RESULTADOS DOS EXPERIMENTOS

O material utilizado foi selecionado a partir do material disponível na “Deutsche Telekom” [65,66,67]. Foram escolhidas aleatoriamente 5 sentenças para cada uma das duas vozes masculinas, o mesmo ocorrendo para vozes femininas. Este material foi amostrado a 16 kHz e digitalizado com 16 bits de resolução, sendo depois filtrado por um filtro IRS. Todas as sentenças foram calibradas em -30 dBov. A degradação da qualidade foi inserida por meio da utilização do padrão MNRU.

A.2.1. Primeiro Experimento - Avaliação da Qualidade para Palavras Isoladas

Este experimento investigou se os ouvintes são, em princípio, capazes de avaliar uma qualidade instantânea de voz baseados em intervalos de curta duração. Este passo foi necessário para um melhor entendimento dos resultados obtidos no segundo experimento.

As 20 sentenças utilizadas foram divididas em suas palavras constituintes. Apesar da grande variação na duração das palavras, estas não foram divididas em segmentos de

mesmo comprimento (sílabas), a fim de se manter a inteligibilidade, especialmente nos casos de grande degradação.

Os valores utilizados para o parâmetro Q da MNRU, na introdução da degradação, ficaram entre 0 dB (para qualidade muito pobre) e 35 dB (próximo da qualidade de referência).

As palavras foram então individualmente avaliadas pelos ouvintes. A escala utilizada na avaliação foi a ACR (*Absolute Category Rating*).

A figura A.1 mostra os resultados obtidos para as palavras isoladas em relação ao uso de sentenças completas. A linha tracejada corresponde a um mapeamento de segunda ordem otimizado para os dados coletados [23]. Como resultado, verificou-se que a maioria dos ouvintes não conseguiu compreender um pequeno conjunto de palavras curtas, submetidas a condições de degradação bastante severas. Estes dados não foram excluídos, pois a inteligibilidade também faz parte da avaliação. De uma maneira geral, o comportamento dos ouvintes em resposta às palavras isoladas mostrou-se bastante consistente. Nenhuma dependência sistemática relativa aos diferentes locutores foi observada, tornando razoável tomar a média entre os diferentes locutores. Os resultados obtidos suportam a hipótese de que os ouvintes podem avaliar de maneira confiável eventos auditivos breves.

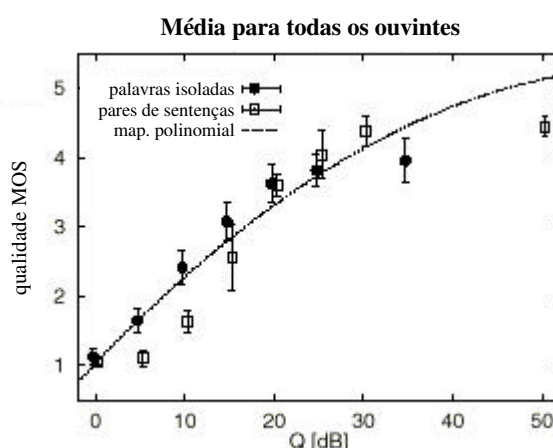


Figura A.1 - Resultados para o primeiro experimento

A.2.2. Segundo Experimento - Avaliação Contínua da Qualidade de Voz Variante no Tempo

Aqui, todas as 20 seqüências foram concatenadas em uma única sentença longa, com pequenas pausas entre elas. O arranjo das sentenças foi feito de duas maneiras: uma iniciando com as vozes femininas e a outra iniciando com as vozes masculinas.

Novamente foi utilizado o padrão MNRU para inserir as degradações, sendo que dois perfis de qualidade foram gerados, como mostra a figura A.2 [23].

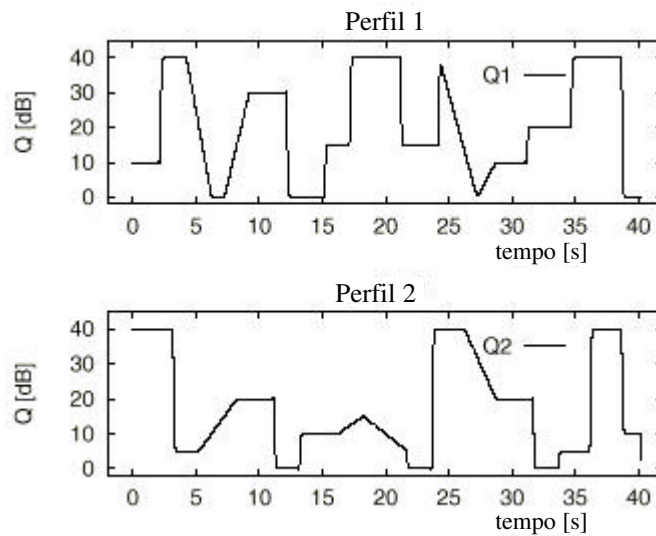
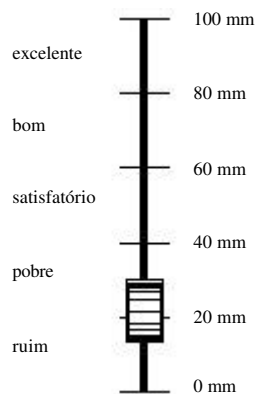


Figura A.2 - Perfis de qualidade utilizados no experimento

Estes perfis de qualidade foram cuidadosamente elaborados de acordo com investigações prévias das reações dos ouvintes. Cada teste foi realizado duas vezes. Como a avaliação deveria ser feita continuamente, uma escala móvel continuamente ajustável foi utilizada, contendo os cinco possíveis níveis de qualidade previstos pela escala ACR, como mostrado na figura A.3:



A.3 - Representação esquemática da escala móvel utilizada

A posição $r(t)$ do botão deslizante na escala de comprimento igual a 100 mm foi amostrada à frequência de 8 kHz e posteriormente dizimada para 8 Hz.

Procedimentos:

O primeiro passo foi a quantificação do atraso entre a percepção auditiva e a reação motora passada para o botão deslizante. Para isso, foi feito o cálculo da correlação cruzada entre os perfis de qualidade mostrados na figura A.2 e os resultados obtidos com os ouvintes, onde se observou um atraso médio de cerca de 1 segundo. Pode-se ver então, através da figura A.4 que, descontado o atraso, os ouvintes estão capacitados a avaliar os diferentes níveis de qualidade dentro dos perfis esperados [23]. A área hachurada em torno da curva de resposta corresponde ao desvio padrão das respostas dos ouvintes. Nota-se claramente o atraso da resposta em relação ao perfil alvo. O eixo vertical da figura A.4

corresponde à escala deslizante. O mapeamento entre esta escala e os valores do parâmetro Q foi feito segundo a Rec. ITU-T P.810 [48].

Os resultados obtidos neste experimento mostram que os ouvintes são capazes de avaliar continuamente a qualidade de voz variante no tempo com um alto grau de confiabilidade e reprodutibilidade. É notável também que todos os ouvintes tendem a fazer uso da escala de resposta de uma maneira reprodutível, o que é indicado pelos baixos desvios encontrados tanto quando se analisa individualmente cada um dos ouvintes, quanto quando se analisa o desvio entre os diversos usuários [23].

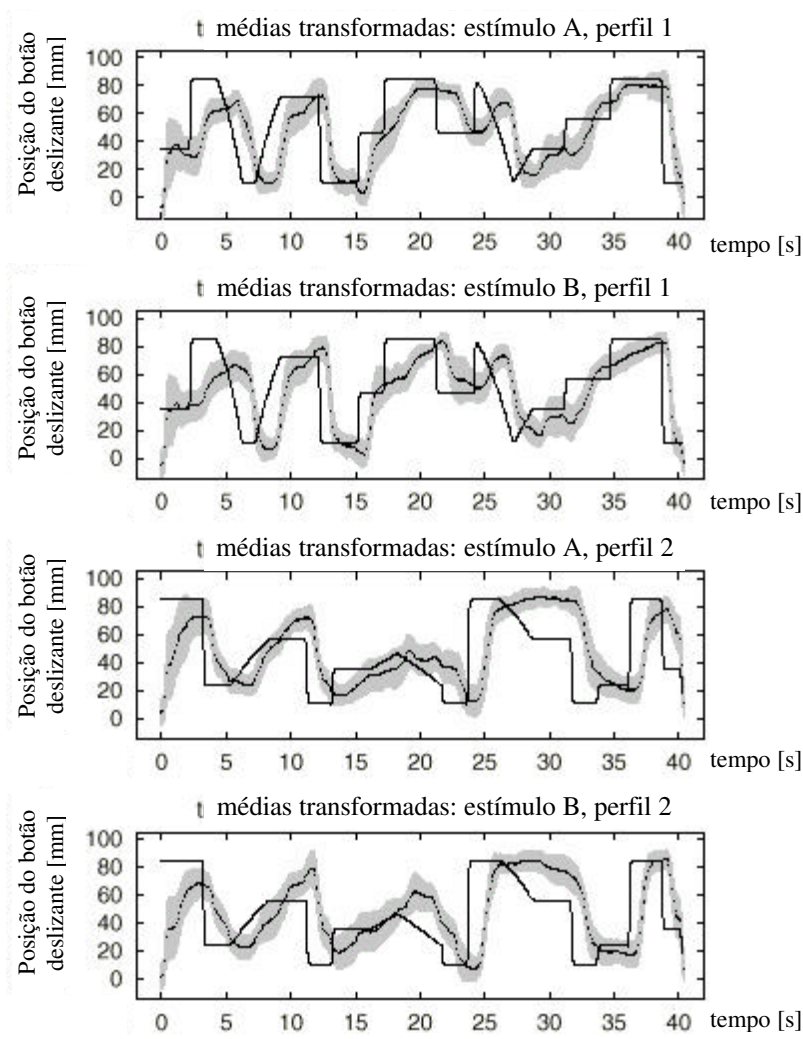


Figura A.4 - Respostas dos ouvintes aos quatro diferentes estímulos em teste

A.3. AVALIAÇÃO DA QUALIDADE VARIÁVEL DE VOZ

Hansen desenvolveu, como parte de seu projeto de doutorado, um método de avaliação objetiva de qualidade de voz, batizado de qC, baseado nos mesmos conceitos utilizados no PSQM, se diferenciando deste em alguns processamentos. Portanto, os

resultados e adaptações produzidas podem, em tese, ser estendidos a outras medidas objetivas que façam uso do modelo psicoacústico.

A medida qC, para condições de transmissão estacionárias, é baseada na correlação cruzada entre as representações internas do sinal de teste e do sinal de referência, ponderados ao longo de toda a duração dos sinais de voz. Essas representações internas são divididas em quadros de duração $\tau = 20$ ms. Os quadros correspondem aos estados de excitação temporal do ouvido em resposta a um sinal de entrada. Essa característica permite que este método seja empregado na estimação contínua da qualidade subjetiva de voz, através do cálculo da medida objetiva individualmente para cada quadro.

A estimação dos valores subjetivos é feita através de um mapeamento exponencial dos valores qC obtidos para valores correspondentes a uma escala móvel ($r_{qC}(t)$). A figura A.5 mostra os resultados obtidos para quadros de 20 ms. Nesta figura são mostradas quatro situações abordadas, onde a linha cheia representa a resposta subjetiva, a linha com traços maiores representa os valores estimados r_{qC} e a linha com traços menores representa o perfil alvo.

A fim de investigar a influência do parâmetro τ , utilizou-se também uma divisão em quadros de 100 ms. Como resultado, notou-se que tal influência é muito pequena (como os gráficos são quase idênticos, não serão mostradas as curvas para $\tau = 100$ ms).

A estrutura ondulada observada nas curvas para os valores estimados não se deve ao valor de τ , e sim à introdução das degradações pelo padrão MNRU, as quais são moduladas com a energia dos sinais de entrada, e esta característica afeta a medida objetiva, ocasionando tais flutuações.

Outra característica observada é que, para intervalos com baixa qualidade de voz ($r(t) < 20$ mm), os valores objetivos mapeados são bastante próximos da resposta subjetiva média. Já para intervalos com boa qualidade de voz, há um pequeno descasamento. Isto ocorre devido à transformação exponencial simples escolhida de qC para a escala em milímetros. O uso de um mapeamento exponencial de dois parâmetros resultaria em uma considerável redução dos desvios [23, Appendix C].

Em suma, observou-se um desempenho satisfatório na utilização do método objetivo para avaliação contínua da qualidade de voz variante no tempo. Isto confirma que o modelo pode prever diferenças perceptíveis variáveis no tempo entre o sinal original e o degradado, baseando-se em suas representações internas.

Com a utilização de métodos objetivos de avaliação contínua, pode-se obter algumas importantes informações a respeito do desempenho de uma combinação individual de esquemas de codificação e transmissão, como, por exemplo, a identificação do momento em que ocorre uma degradação de qualidade instantânea, o que fornece uma pista para a causa da degradação.

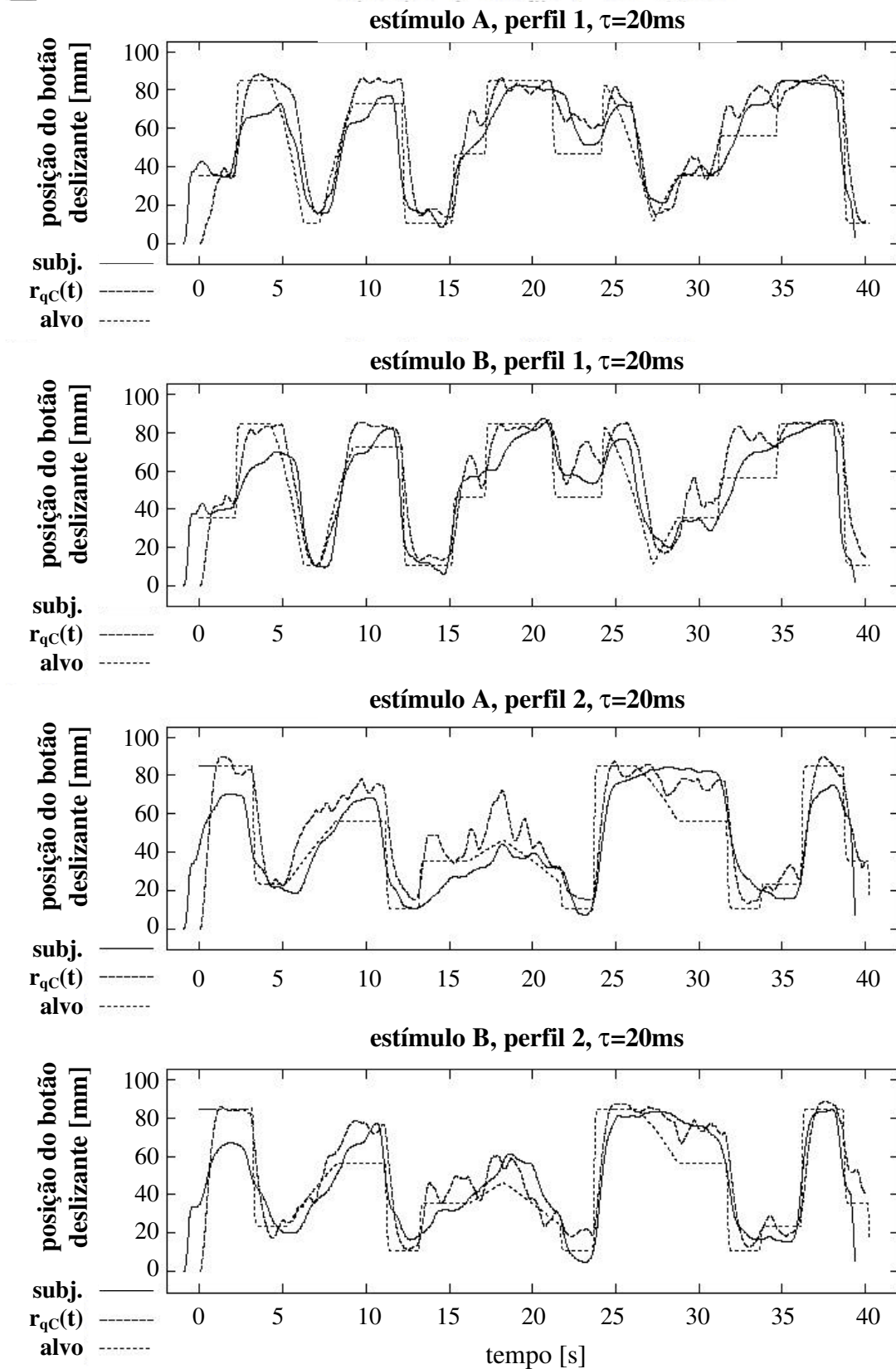


Figura A.5 - Predição da qualidade de voz contínua no tempo pela medida de qualidade $r_{qc}(t)$

APÊNDICE B

MEDIDAS OBJETIVAS DE QUALIDADE PARA ÁUDIO

B.1. INTRODUÇÃO

A questão das medidas objetivas de qualidade para voz está bem definida e delimitada. O problema está limitado à medição da qualidade de um sinal de voz à saída de um determinado codec ou sistema, numa faixa de frequência de 300 a 3.400 Hz ou, excepcionalmente, para sinais de banda larga (até 7 kHz). Por outro lado, o universo das medidas de qualidade para sinais de áudio é bastante amplo. Dentre as situações possíveis, tem-se:

- sistemas de som residenciais;
- casas de espetáculo;
- som automotivo;
- ambiente computacional, etc;
- codecs (MPEG, Dolby, etc.);

Além disso, são vários os equipamentos a serem analisados, como amplificadores, caixas de som, microfones e fones de ouvido; além disso, a acústica do local deve ser considerada.

Dentre os casos citados, apenas o último (avaliação de codecs de áudio) possui regras estabelecidas [68,69]. Para os demais, não existe qualquer tipo de medida objetiva disponível ou em estudo, e nem mesmo medidas subjetivas são padronizadas, ficando a cargo do projetista, através de sua sensibilidade e experiência, avaliar a qualidade do sistema de som em questão.

Os estudos das avaliações objetivas de qualidade de codecs de voz e de áudio estão intimamente ligados e obedecem às mesmas premissas. Este último caso, porém, apresenta uma maior complexidade e quantidade de conceitos envolvidos. Por esse motivo, o estudo das avaliações objetivas aplicadas a codecs de áudio representa a seqüência natural dos estudos realizados neste trabalho. Este apêndice apresenta um estudo preliminar deste assunto, visando o desenvolvimento de futuras pesquisas nesta área, bem como confrontar o tema com aquele desenvolvido neste trabalho.

B.2. PROPOSTAS APRESENTADAS

A necessidade de se desenvolver um método objetivo de avaliação de áudio, capaz de preencher todos os pré-requisitos preconizados na Rec. ITU-R BS.1116 [68], motivou a ITU-R a fazer, em 1994, uma chamada aberta de propostas. Sete métodos foram apresentados:

- DIX (*Disturbance Index*): é baseado em um banco de filtros simulando as características auditivas (bandas críticas), através do qual se avalia a estrutura das envoltórias temporais. O índice de qualidade é dado pela combinação das distorções lineares e não-lineares obtidas [70].
- NMR (*Noise-to-Masked-Ratio*): avalia a diferença de nível entre o limiar de mascaramento e o sinal de ruído. Seus principais valores de saída são a porcentagem de quadros com distorções audíveis e a distância média entre a energia de erro e o limiar de mascaramento [71].
- OASE (*Objective Audio Signal Evaluation*): usa um banco de 241 filtros para analisar os sinais de entrada. O índice de qualidade é dado pela probabilidade média de detecção de distorções [72].
- PAQM (*Perceptual Audio Quality Measure*): avalia a diferença, mapeada por um processamento cognitivo, entre os sinais de referência e degradado [73].
- Perceval (*Perceptual Evaluation*): modela as características de transferência do ouvido médio para o interno, a fim de formar uma representação interna do sinal [74]. Este método calcula a diferença entre as representações dos sinais de referência e teste.
- POM (*Perceptual Objective Measurement*): quantifica a degradação através da comparação entre as representações psicoacústicas de ambos os sinais [75].
- *Toolbox approach*: descreve a representação perceptual das diferenças entre os dois sinais de áudio [69].

Após uma análise minuciosa, chegou-se à conclusão de que nenhuma proposta atendia a todas as exigências requeridas. Foi então sugerido que todos os esforços fossem direcionados para um único método que incorporasse os pontos mais fortes de cada uma das propostas, o que deu origem ao método PEAQ e à Recomendação ITU-R BS.1387 [69].

B.3. O MÉTODO PEAQ

A figura B.1 fornece uma visão global da estrutura utilizada no método PEAQ [69].

Este método consiste de um modelo de simulação do ouvido, vários processamentos intermediários (aqui referidos como “pré-processamento dos padrões de excitação”) e o cálculo das variáveis de saída do modelo (MOV), a partir das quais obtém-se um único valor representando a qualidade de áudio básica do sinal sob teste.

Duas diferentes versões foram implementadas: a primeira, chamada “versão básica”, é baseada na FFT, e tem como principal característica a baixa complexidade computacional, permitindo a implementação em tempo real; a segunda, chamada “versão avançada”, é baseada em um banco de filtros, e tem como principal característica um melhor desempenho, a fim de se adequar às aplicações mais exigentes, porém com uma complexidade computacional mais elevada. A estrutura geral, à exceção de alguns poucos processamentos, é semelhante para as duas versões.

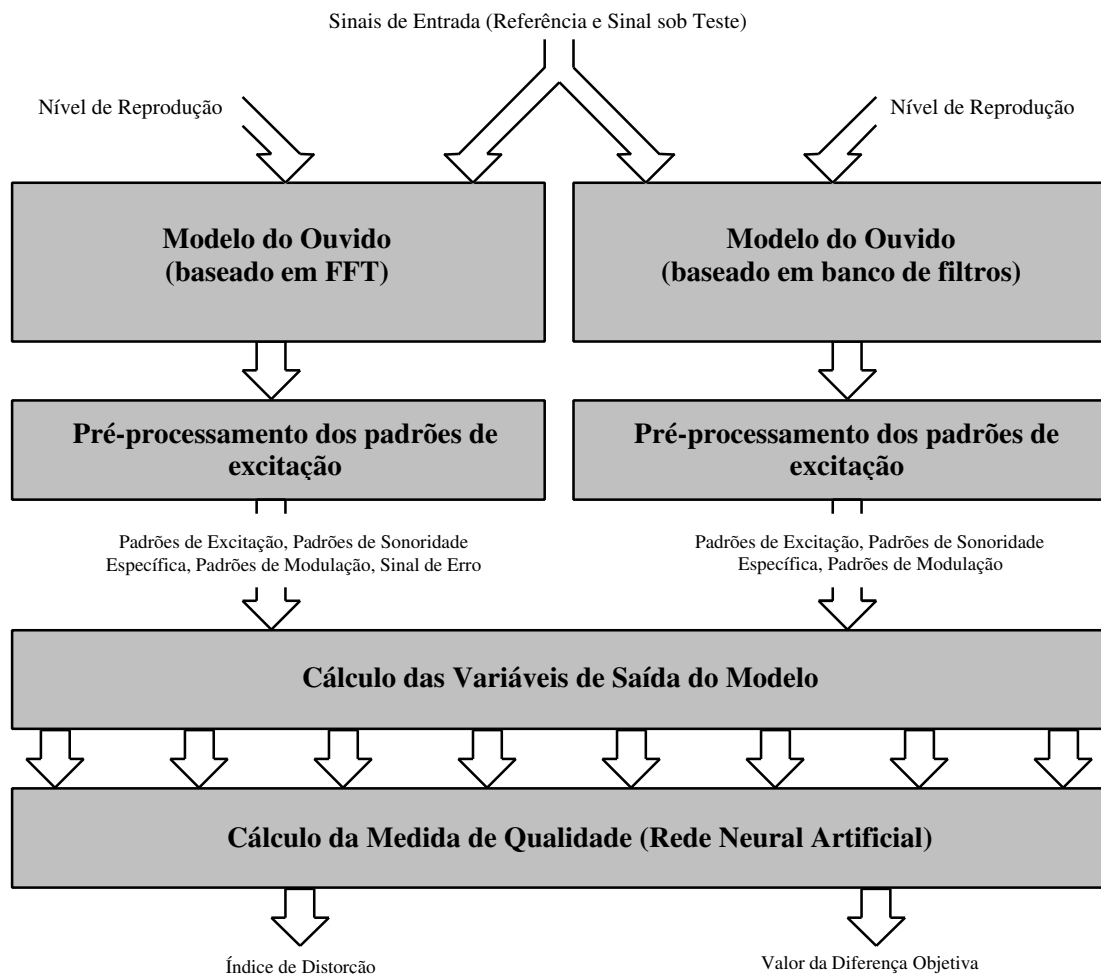


Figura B.1 - Diagrama de blocos genérico do método PEAQ

As entradas para o cálculo das MOV, obtidas tanto para o sinal de referência quanto aquele sob teste, são:

- Os *padrões de excitação*, obtidos pela distribuição da energia dos sinais em diferentes regiões de *pitch*.
- Os *padrões de excitação espectralmente adaptados*, os quais são computados apenas no modelo baseado no banco de filtros, permitindo a obtenção de resultados mais precisos.
- Os *padrões de sonoridade específica*, simulando a sonoridade percebida pelo ouvinte.
- Os *padrões de modulação*, calculados a partir dos padrões de excitação.
- O *sinal de erro*, calculado como a diferença espectral entre os sinais (somente para o modelo do ouvido baseado na FFT).

Todas as computações para sinais em estéreo são feitas independentemente e da mesma maneira para os canais direito e esquerdo.

A seguir, será feita uma breve descrição das principais etapas do método PEAQ.

B.3.1. Versão Baseada na FFT

A estrutura adotada na versão baseada na FFT é mostrada na figura B.2.

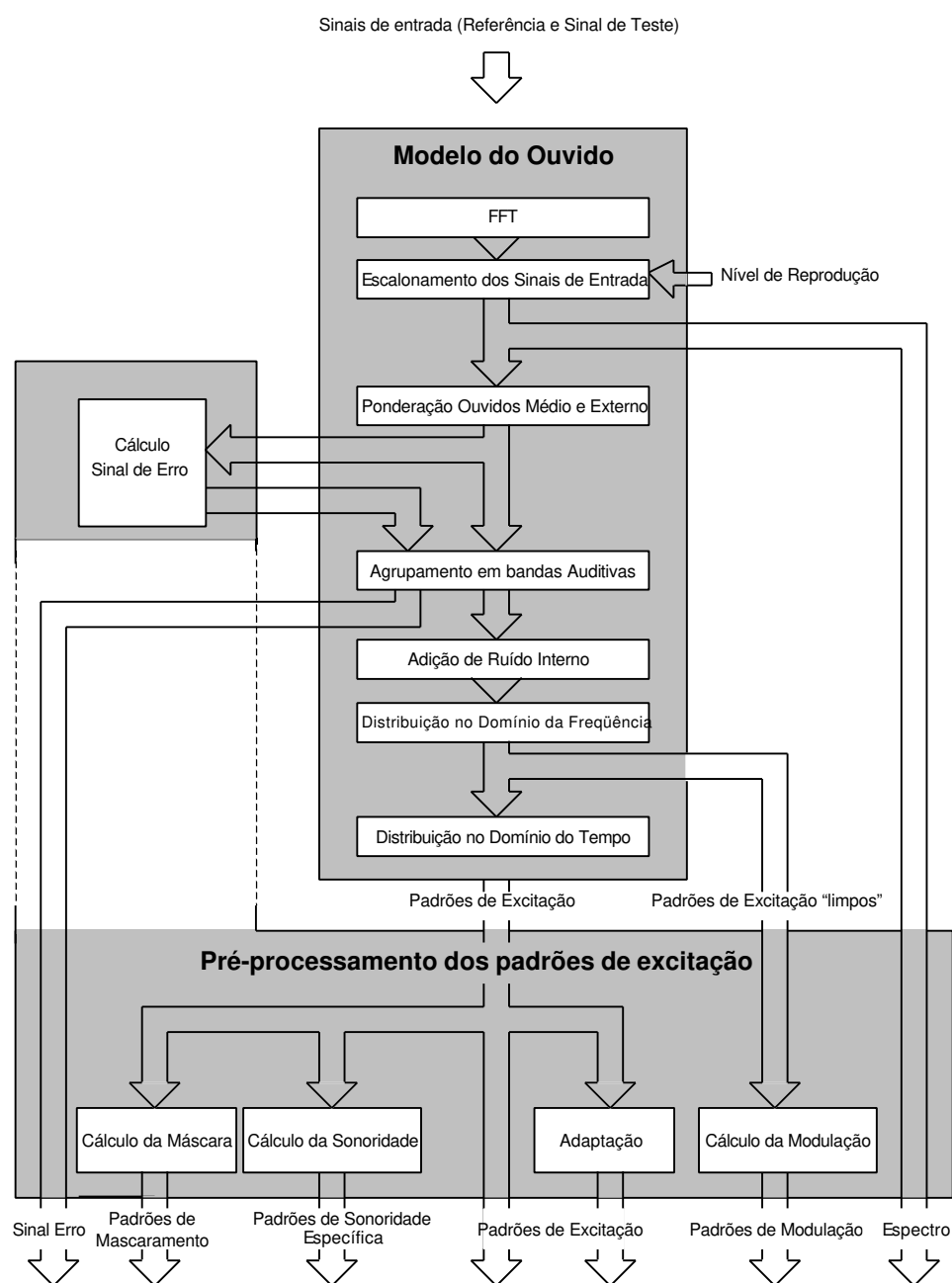


Figura B.2 - Esquema da Versão Baseada na FFT

As entradas deste modelo, que são os sinais de referência e de teste alinhados temporalmente e amostrados a uma frequência de 48 kHz, são divididas em blocos de aproximadamente 0,042 segundo, com uma superposição de 50%. Cada bloco é transformado para o domínio da frequência usando uma janela de Hanning e uma FFT aplicada a cada trecho, e escalonada para o nível de reprodução (se tal nível é

desconhecido, recomenda-se a adoção do valor de 92 dB_{SPL}). Uma função de ponderação é aplicada sobre os coeficientes espectrais, os quais, após a ponderação, modelam a resposta em frequência dos ouvidos externo e médio. A transformação para a representação de *pitch* é feita agrupando-se os coeficientes espectrais ponderados em bandas críticas. Uma compensação dependente da frequência é adicionada para simular o ruído interno no sistema auditivo. Uma função de espalhamento é usada para modelar o espectro dos filtros auditivos (determinados a partir das bandas críticas [40]) no domínio da frequência. A seguir, faz-se o mesmo no domínio do tempo, a fim de se levar em conta os efeitos de mascaramento.

Os padrões de excitação obtidos, que são a maneira como a energia dos sinais se distribui entre as diferentes regiões de *pitch*, são usados para computar os *padrões de sonoridade específica* e os *padrões de mascaramento*. Os padrões anteriores ao espalhamento final no domínio do tempo (padrões de excitação “limpos”) são usados para calcular os *padrões de modulação*.

O sinal de erro é obtido através da diferença entre os espectros de potência para os sinais de referência e de teste, presentes na saída do filtro, representando as características em frequência dos ouvidos externo e médio. É então mapeado para uma escala de *pitch*, pelo agrupamento em bandas críticas. Esse sinal, juntamente com os padrões de excitação, sonoridade específica, modulação e mascaramento, é usado no cálculo dos valores das Variáveis de Saída do Modelo (MOV). Os espectros dos sinais, indicados como uma das saídas do modelo, não são usados neste cálculo, possuindo apenas um caráter de informação complementar.

B.3.2. Versão Baseada em um Banco de Filtros

A estrutura adotada na versão baseada em um banco de filtros é mostrada na figura B.3.

Na entrada deste modelo, os sinais de referência e teste são ajustados para assumir o nível de reprodução e são submetidos a um filtro passa-altas, a fim de remover as componentes DC e subsônicas dos sinais. Os sinais são então decompostos em sinais passa-banda por filtros de fase linear, os quais são igualmente distribuídos sobre uma escala de *pitch* perceptual. Uma ponderação dependente da frequência é aplicada a esses sinais, a fim de modelar as características espectrais dos ouvidos externo e médio. A resposta em frequência dos filtros auditivos é modelada fazendo-se a convolução, no domínio da frequência, destes sinais com uma função de espalhamento dependente do nível.

As envoltórias dos sinais são calculadas tomando-se a transformada de Hilbert dos sinais passa-banda (retificação), e uma convolução temporal com uma função janela é aplicada a fim de modelar o mascaramento regressivo (espalhamento no domínio do tempo 1). A seguir, uma compensação dependente da frequência é adicionada, levando em conta o ruído interno no sistema auditivo e modelando o limiar de silêncio. Finalmente, uma segunda convolução temporal é realizada, usando uma função de espalhamento exponencial que leva em conta o mascaramento progressivo (espalhamento no domínio do tempo 2) [69].

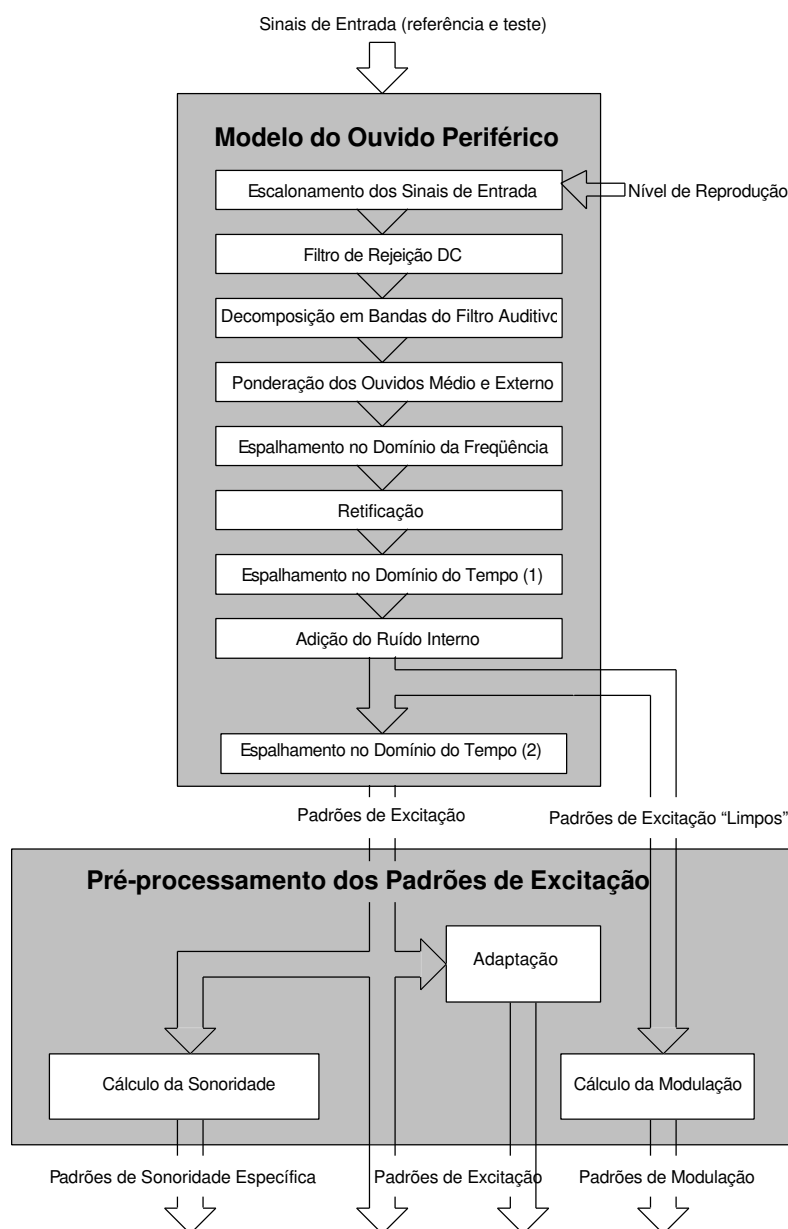


Figura B.3 - Esquema da Versão Baseada em um Banco de Filtros

Os padrões de excitação, obtidos pela divisão da energia dos sinais entre as bandas críticas, são usados para computar os *padrões de sonoridade específica*, enquanto que os padrões antes do espalhamento final no domínio do tempo (padrões de excitação “limpos”) são usados para calcular os *padrões de modulação*. Estes, junto com os padrões de excitação, formam a base para o cálculo das Variáveis de Saída do Modelo (MOV). A fim de separar a influência da resposta em frequência do equipamento sob teste de outras distorções, os padrões de excitação dos sinais de referência e teste são também espectralmente adaptados um ao outro (adaptação). Os padrões de modulação e de sonoridade específica são calculados tanto a partir dos padrões de excitação adaptados quanto dos não-adaptados.

B.3.3. Cálculo das Variáveis de Saída do Modelo (MOV)

As variáveis de saída do modelo, referentes à figura B.1, são:

- *Diferença de Modulação*: diferenças na modulação das envoltórias temporais dos sinais de teste e de referência são medidas pelo cômputo da diferença da modulação local para cada canal em cada filtro auditivo. A seguir, a diferença de modulação momentânea é calculada pela soma das diferenças de modulação locais. Neste cálculo, o limiar de silêncio é levado em conta através de um fator de ponderação dependente do nível, calculado a partir dos padrões de excitação modificados e da função de ruído interno.
- *Sonoridade do Ruído*: estas variáveis de saída do modelo estimam a sonoridade parcial de distorções aditivas na presença do sinal de referência.
- *Largura de Faixa*: estima a largura de faixa média dos sinais de referência e teste.
- *Taxa de Ruído/Mascaramento*: estes valores são calculados a partir dos valores de ruído e mascaramento obtidos.
- *Valor Relativo de Quadros com Distúrbios*: mede a quantidade de quadros com distúrbios em relação ao número total. Quadros com pequena energia são desconsiderados.
- *Probabilidade de Detecção*: é a probabilidade de detecção de uma diferença de nível (ganho ou atenuação) para determinado canal do sinal de teste. Esta probabilidade é tanto maior quanto maior for a diferença de nível. Este valor é usado na estimação do número de quadros em que distorções serão detectadas pelos ouvintes.
- *Estrutura Harmônica de Erro*: um sinal de referência contendo harmônicas fortes tem um espectro caracterizado por um número de picos regularmente espaçados, separados por vales profundos. Sob determinadas condições, o sinal de erro pode herdar tal estrutura, porém ocupando o lugar dos vales, resultando em uma distorção com atributos tonais que poderiam aumentar a importância do erro. Por isso, um parâmetro de qualidade melhor que a magnitude da estrutura de erro harmônica é sua variabilidade sobre a sequência completa, a qual é dada pelo desvio padrão da magnitude da estrutura harmônica sobre os quadros (Variação Harmônica de Erro - EHV).

B.3.4. Estimação da Qualidade Perceptual Básica do Áudio

Esta estimação é realizada pelo mapeamento das Variáveis de Saída do Modelo (MOV) para um único valor, através do uso de uma estrutura de rede neural artificial com uma camada oculta. Esta rede neural usa como função de ativação uma sigmóide assimétrica. A versão básica do método utiliza três nós na camada oculta da rede neural, enquanto que a versão avançada utiliza quatro. As entradas são então mapeadas para um *índice de distorção*, o qual é diretamente relacionado à qualidade perceptual básica do áudio.

APÊNDICE C

DEFINIÇÕES COMPLEMENTARES

C.1. dBA

A escala dBA consiste da aplicação da ponderação “A” de frequência à escala dB_{SPL} , realizada através do filtro cuja resposta é mostrada na figura C.1 [76]. Esta ponderação visa reproduzir a característica do ouvido humano de ser mais sensível à faixa de frequências de 2 a 4 kHz. Por se adaptar bem às características auditivas, este é o decibel mais popular para a medição do nível de pressão sonora. Entretanto, é importante observar que este filtro não simula as características do ouvido com a mesma precisão da modelagem do Capítulo 3 e, portanto, não pode substituir tal modelo. Desta forma, os sinais por ele ponderados não podem ser diretamente relacionados com a sonoridade expressa em sôns.

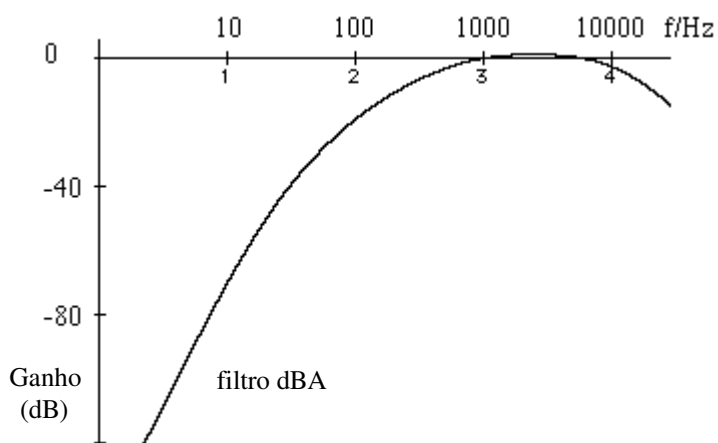


Figura C.1 - Resposta em frequência do filtro dBA

C.2. PITCH

O termo *pitch* tem sido usado com dois sentidos diferentes:

- na área de processamento de voz, o termo é freqüentemente utilizado para designar a frequência de oscilação da glote (vibração das cordas vocais);
- em psicoacústica, é usado como um atributo da sensação auditiva, segundo a definição encontrada na ANSI (*American National Standards Institute*), a qual estabelece que *pitch* é o atributo auditivo do som, de acordo com o qual os sons podem ser ordenados, em uma escala de frequência, de baixo a alto. Este é o sentido adotado neste trabalho.

Os estudos da percepção humana do *pitch* são complexos. Maiores informações podem ser encontradas em [77,78,79].

C.3. TERÇAS-OITAVAS

A divisão do espectro de frequências audível em oitavas se faz de maneira que a frequência central de uma faixa corresponda ao dobro da frequência central da faixa anterior [80]. A expressão geral é dada então por:

$$fc(n) = c \cdot 2^n, \quad n = 0, 1, 2, 3 \dots$$

onde $fc(n)$ representa a frequência central n e c é a primeira frequência central desejada.

Então, como se pode observar, se a frequência central de uma banda tem como valor $c \cdot 2^n$, as frequências centrais anterior e posterior terão os valores $c \cdot 2^{n-1}$ e $c \cdot 2^{n+1}$, respectivamente. Da mesma forma, os limites inferior e superior para a faixa definida por essa frequência central serão, respectivamente:

$$lib = c \cdot 2^{n-\frac{1}{2}} = 0,7071 \cdot fc$$

$$lsb = c \cdot 2^{n+\frac{1}{2}} = 1,4142 \cdot fc,$$

onde lib é o limite inferior e lsb é o limite superior.

Então, por exemplo, para 1 kHz tem-se 500 Hz para uma oitava abaixo e 2 kHz para uma oitava acima, e tem-se 707 Hz para o limite inferior e 1.414 Hz para o limite superior.

A divisão em terços de oitavas segue o mesmo princípio, sendo que a expressão geral, para este caso, é dada por:

$$fc(n) = c \cdot 10^{\frac{n}{10}}$$

Da mesma maneira, se a frequência central de uma banda tem como valor $c \cdot 10^{n/10}$, as frequências centrais anterior e posterior terão os valores $c \cdot 10^{(n-1)/10}$ e $c \cdot 10^{(n+1)/10}$, respectivamente. Então, os limites inferior e superior para a faixa definida por essa frequência central serão, respectivamente:

$$lib = c \cdot 10^{\frac{\left[n - \left(\frac{1}{2} \right) \right]}{10}} = 0,8913 \cdot fc$$

$$lsb = c \cdot 10^{\frac{\left[n + \left(\frac{1}{2} \right) \right]}{10}} = 1,1220 \cdot fc$$

É interessante observar ainda que $10^{\frac{n}{10}} \cong 2^{\frac{n}{3}}$. Consequentemente, a expressão para as terças-oitavas pode ser reescrita como:

$$fc(n) \cong c \cdot 2^{\frac{n}{3}}$$

ou seja, é a mesma expressão encontrada para as oitavas, porém com o expoente dividido por três, daí a adoção do nome terças-oitavas.

O valor $\Delta c = 0,312$ utilizado no algoritmo PSQM é a relação aproximada entre as larguras das faixas determinadas pelas oitavas e terças-oitavas.

C.4. SPL

Definição de SPL – Sound Pressure Level (Nível de Pressão Sonora) [81]: tem como unidade o decibel SPL (dB_{SPL}), e é dado pela expressão $20\log_{10}(P/P_{\text{ref}})$, onde P é a pressão sonora do sinal que se está medindo e P_{ref} é a pressão sonora de referência, a qual pode assumir dois valores:

$$\text{a) } P_{\text{ref}} = 0,0002\mu\text{B} \quad (2 \times 10^{-5} \text{ N/m}^2)$$

$$\text{b) } P_{\text{ref}} = 0,1\mu\text{B} \quad (0,1 \text{ N/m}^2)$$

onde μB é a pressão em microbars.

É importante observar que a pressão sonora de referência dada no item (a) é mais utilizada nas medições relacionadas com a audição e nas medições de nível sonoro no ar e nos líquidos, enquanto que aquela dada no item (b) tem maior aplicação na calibração de transdutores e certos tipos de medição de nível sonoro em líquidos. Os dois níveis de medição diferem um do outro em aproximadamente 74 dB. Por essa razão, é necessário indicar explicitamente o nível de referência adotado (neste trabalho, usou-se o primeiro).

C.5. CARACTERÍSTICA DE TRANSMISSÃO IRS MODIFICADA

As características de frequência de transmissão e recepção IRS consistem das características médias de resposta em frequência de transmissão e recepção encontradas em telefones analógicos, a partir de uma extensa série de medidas feitas no início dos anos 70 [82,83]. Porém, para se adequar às características de sonoridade encontradas nos telefones analógicos, foi também necessário incluir um filtro passa-bandas, na faixa de 300-3400 Hz, conhecido como filtro SRAEN [84]. Este procedimento funcionou bem para as transmissões analógicas de longa distância. Com o advento das conexões digitais, tal abordagem perdeu sua efetividade. Para as propostas de avaliação de codecs operando a baixas taxas de bits, uma melhor característica pode ser obtida pela modificação da resposta IRS, através da remoção do filtro SRAEN, originando assim a característica de transmissão IRS modificada [6].

C.6. dBov

dBov é o nível relativo ao limiar de saturação (*overload*) de um sinal em um computador ou codec digital [6]. Por exemplo, para um sistema de 16 bits, o nível máximo 0 dBov corresponde a um nível DC igual a 32767. Este decibel é largamente utilizado em implementações digitais.

C.7. MNRU (UNIDADE DE REFERÊNCIA DE RUÍDO MODULADO)

Historicamente, o MNRU foi idealizado para gerar uma distorção não-linear que simulasse a distorção produzida por sistemas de codificação PCM logarítmicos [85], os quais produzem uma relação sinal-ruído de quantização independente do nível do sinal de entrada. O resultado corresponde subjetivamente à adição, ao sinal original, de um ruído gaussiano (branco na faixa desse sinal) cuja amplitude seja proporcional à amplitude desse sinal. Isto equivale à modulação do ruído pelo sinal de voz, advindo daí o nome “ruído modulado”. A relação entre o sinal e tal ruído, em dB, é referida como Q [7].