



Ministério da  
**Ciência, Tecnologia  
e Inovação**



sid.inpe.br/mtc-m19/2011/06.06.12.20-TDI

# **USO DE REDES NEURAIS OTIMIZADAS PARA RECUPERAÇÃO DO PERFIL DE CONCENTRAÇÃO DE GASES TRAÇO ATMOSFÉRICOS A PARTIR DE DADOS DE SATÉLITES**

Adenilson Roberto Carvalho

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada,  
orientada pelos Drs. Fernando Manuel Ramos, e João Carlos Carvalho, aprovada  
em 27 de junho de 2011.

URL do documento original:

<<http://urlib.net/8JMKD3MGP7W/39R454E>>

INPE  
São José dos Campos  
2011

**PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@sid.inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE (RE/DIR-204):****Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

**Membros:**

Dr<sup>a</sup> Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr<sup>a</sup> Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr<sup>a</sup> Regina Célia dos Santos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

Dr. Horácio Hideki Yanasse - Centro de Tecnologias Especiais (CTE)

**BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Deicy Farabello - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

**REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

**EDITORAÇÃO ELETRÔNICA:**

Vivéca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da  
**Ciência, Tecnologia  
e Inovação**



sid.inpe.br/mtc-m19/2011/06.06.12.20-TDI

## **USO DE REDES NEURAIAS OTIMIZADAS PARA RECUPERAÇÃO DO PERFIL DE CONCENTRAÇÃO DE GASES TRAÇO ATMOSFÉRICOS A PARTIR DE DADOS DE SATÉLITES**

Adenilson Roberto Carvalho

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada,  
orientada pelos Drs. Fernando Manuel Ramos, e João Carlos Carvalho, aprovada  
em 27 de junho de 2011.

URL do documento original:

<<http://urlib.net/8JMKD3MGP7W/39R454E>>

INPE  
São José dos Campos  
2011

Dados Internacionais de Catalogação na Publicação (CIP)

---

Carvalho, Adenilson Roberto.

C253u      Uso de redes neurais otimizadas para recuperação do perfil de concentração de gases traço atmosféricos a partir de dados de satélites / Adenilson Roberto Carvalho. – São José dos Campos : INPE, 2011.

xxxii+176 p. ; (sid.inpe.br/mtc-m19/2011/06.06.12.20-TDI)

Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2011.

Orientadores : Drs. Fernando Manuel Ramos, e João Carlos Carvalho.

1. problemas inversos. 2. redes neurais. 3. efeito estufa. 4. sensoriamento remoto. 5. computação evolutiva. I.Título.

CDU 004:528.835

---

Copyright © 2011 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, reprográfico, de microfilmagem ou outros, sem a permissão escrita do INPE, com exceção de qualquer material fornecido especificamente com o propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2011 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, microfilming, or otherwise, without written permission from INPE, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

Aprovado (a) pela Banca Examinadora  
em cumprimento ao requisito exigido para  
obtenção do Título de Doutor(a) em  
Computação Aplicada

Dr. Haroldo Fraga de Campos Velho



---

Presidente / INPE / São José dos Campos - SP

Dr. Ezzat Selim Chalhoub



---

Vice-Presidente / INPE / SJCampos - SP

Dr. Fernando Manuel Ramos



---

Orientador(a) / INPE / SJCampos - SP

Dr. João Carlos Carvalho



---

Orientador(a) / ANATEL / Rio de Janeiro - RJ

Dr. José Carlos Becceneri



---

Membro da Banca / INPE / SJCampos - SP

Dr. Rodrigo Augusto Ferreira de Souza



---

Convidado(a) / JEA / Manaus - AM

Dr. Ivan Bergier Tavares de Lima



---

Convidado(a) / EMBRAPA / Corumbá - MS

Este trabalho foi aprovado por:

( ) maioria simples

unanimidade

Aluno (a): Adenilson Roberto Carvalho

São José dos Campos, 27 de junho de 2011



*“Eu não sei o que o mundo pensa sobre mim, mas, quanto à mim, eu me sinto apenas como um garoto brincando na areia da praia, e me divertindo de vez em quando em encontrar uma pedra mais lisa ou uma concha mais bonita do que o normal, enquanto o grande oceano da verdade estende todo o desconhecido diante de mim.”.*

*“I don't know what I may seem to the world, but, as to myself, I seem to have been only like a boy playing on the sea shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.”.*

SIR ISAAC NEWTON (1642-1727)





*A minha mãe Nairce Aparecida e ao meu pai José Carlos  
por me fornecerem o caráter e equilíbrio necessários para me  
manter de pé sempre que possível, a coragem para se levantar  
após a queda inevitável e sobretudo as condições para que pudesse  
transformar meus sonhos em realidade.*



## AGRADECIMENTOS

A Deus pela existência.

A meus familiares pelo incentivo irrestrito, apoio e carinho nas horas difíceis.

À minha esposa Flávia minha fiel companheira na hora da tribulação por suportar momentos de ausência e ainda assim ficar ao meu lado.

A meu irmão, amigo e co-orientador João Carlos pela inspiração, pelas críticas construtivas e acima de tudo por me recolocar nos trilhos novamente.

Ao meu Orientador e amigo Fernando Ramos por acreditar neste trabalho, pelo acompanhamento e diretrizes em todas as suas etapas e por fazer renascer em mim o interesse em fazer ciência.

Aos meus amigos do INPE por me ajudarem muitas vezes sem perceber que estavam me ajudando.

A todos aqueles que ao longo desses anos contribuíram de uma forma ou de outra para a realização deste trabalho.

Ao professor Dr. José Demísio Simões da Silva (in memoriam).

E por último ao CNPq e à CAPES pelo apoio e suporte financeiro.



## RESUMO

Atualmente, o monitoramento da concentração de gases traço na atmosfera é feito através de medições realizadas *in situ*, em pontos esparsos sobre diferentes regiões do planeta. Contudo, as fontes e sumidouros de tais gases ( $CO_2$  e  $CH_4$ , por exemplo) variam muito em escala temporal e espacial; e o uso de medidas por sensoriamento remoto a partir de satélites pode fornecer um conjunto mais denso de medidas aumentando a resolução espacial e temporal das observações. O sensoriamento remoto da atmosfera permite, a partir de um conjunto de radiâncias espectrais realizadas em diferentes comprimentos de onda, determinar a estrutura vertical de parâmetros atmosféricos. Se os canais a bordo do satélite estão localizados em uma região espectral que é sensível à concentração do gás traço, a intensidade desta radiância pode nos fornecer informações acerca do perfil do gás na atmosfera num determinado tempo e local. Neste trabalho desenvolvemos um modelo baseado em Redes Neurais capaz de estimar a concentração de  $CO_2$  e  $CH_4$  na atmosfera a partir de medidas de radiância obtidas pelo sensor SCIAMACHY a bordo do satélite ENVISAT. Através da análise do conteúdo de informação no sentido de Shannon das linhas espectrais do sensor foi possível selecionar conjuntos reduzidos e auto-suficientes de canais para recuperação de perfis verticais de GEE. Outras duas abordagens para seleção de canais foram testadas com sucesso: a seleção de canais através da extração dos graus de liberdade do sinal (método GLS) e a escolha de canais através de análise da sensibilidade das linhas com relação à variação do gás (método OSP). Adicionalmente propomos uma metodologia baseada em metaheurísticas de busca e computação evolucionária para otimização dos parâmetros do modelo de inversão. Construímos uma nova abordagem para estimação da complexidade computacional da arquitetura de uma Rede Neural Artificial (RNA) baseada no número de neurônios e no número de épocas necessários para o treinamento dos padrões apresentados à rede. Aplicou-se o esquema proposto ao problema da recuperação de perfis atmosféricos de concentração de gases de efeito estufa o qual permitiu a obtenção de uma arquitetura de rede otimizada. Os resultados mostraram que a abordagem computacional proposta apresenta melhor performance em relação a um especialista humano, oferecendo ao mesmo tempo muitas vantagens se comparada a abordagens similares encontradas na literatura. Os resultados foram validados utilizando-se observações *in-situ* independentes obtidas pela divisão de monitoramento global da NOAA. Essas medições foram realizadas no observatório de Mauna Loa, no Havaí (latitude  $19.54^\circ N$ , longitude  $155.58^\circ W$ ) à altitude de 3397 m e no Pólo Sul (latitude  $-89.98^\circ S$ , longitude  $-24.80^\circ W$ ) à altitude de 2800 m acima do nível do mar. As comparações entre os  $CO_2$  e  $CH_4$  prevista pelo modelo e aqueles observados pela NOAA ao longo do ano de 2005 e 2007 nas regiões de interesse mostraram, em geral, boa concordância. O resíduo detectado é menor que 0.9% para ambos os gases.



# USE OF OPTIMIZED NEURAL NETWORKS FOR RETRIEVAL OF ATMOSPHERIC TRACE GAS CONCENTRATION PROFILE BY USING SATELLITE DATA

## ABSTRACT

Currently, the  $CO_2$  and  $CH_4$  concentration monitoring is accomplished through *in situ* measurements, only present in small amounts and numbers and often spread over large areas along different regions of the planet. However, the sources and sinks of these trace gases vary too much across spatial and temporal scale; and the use of satellite remote sensing measurements may provide us a more compact measurements set enhancing, this way, the spatial and temporal resolution observations. The remote sensing of the atmosphere allows, from a spectral radiance set achieved in different wavelengths, to determine the vertical structure of atmospheric parameters. If the sensors aboard satellite are placed at a spectral region sensitive to changes in the greenhouse effect concentration gases such as  $CO_2$  and  $CH_4$ , the obtained radiance intensity can supply us with information concerning the vertical profile of the gases in the atmosphere at a certain time and place. This work aims at developing a neural network model to estimate the  $CO_2$  and  $CH_4$  atmospheric concentration from radiance measurements acquired by SCIAMACHY sensor aboard ENVISAT satellite. By analyzing the information content in the Shannon sense of the sensor spectral lines it was able to select a reduced and self-sufficient channel set for retrieving vertical profiles of trace gases. Two other approaches have been successfully tested: the selection of dedicated channels through determination of the number of degrees of freedom (DFS) for signal and the choice of channels through sensitivity analysis for the gas spectral lines (OSP method). Additionally we propose a methodology based on metaheuristics and evolutionary computing to optimize the parameters of the inverse model. We built a new approach to estimate the computational complexity of the ANN architecture based on the number of neurons and the number of epochs required for the network training. We applied the proposed scheme to the ill-posed problem of retrieving concentration vertical profiles of atmospheric greenhouse gases which has lead to optimized network architectures. The results showed that the computational approach proposed has better performance compared to a human expert, while offering many advantages compared to similar approaches in the literature. The results were validated using independent in-situ observations obtained by the Global Monitoring Division of NOAA. These measurements were performed at the Mauna Loa observatory, Hawaii (latitude  $19.54^\circ N$ , longitude  $155.58^\circ W$ ) at an altitude of 3397 m and at the South Pole (latitude  $-89.98^\circ S$ , longitude  $-24.80^\circ W$ ) at an altitude of 2800 m above sea level. Comparisons between  $CO_2$  and  $CH_4$  predicted by the model and those observed by NOAA over the years 2005 and 2007 in regions of interest showed good agreement. The residue is detected to be less than 0.9% for both gases.





## LISTA DE FIGURAS

	<u>Pág.</u>
1.1 representação esquemática do Efeito Estufa. . . . .	1
1.2 Concentrações atmosféricas mundiais de dióxido de carbono ( $CO_2$ ), metano ( $CH_4$ ), e óxido nitroso ( $N_2O$ ). Apresenta as mudanças nas concentrações atmosféricas nos últimos 2000 anos. . . . .	2
1.3 Variações temporais e latitudinais de $CO_2$ na superfície observadas em MAUNA LOA, US pelas estações de medida da NOAA. . . . .	3
1.4 Variações temporais e latitudinais de $CH_4$ na superfície observadas em MAUNA LOA, US pelas estações de medida da NOAA. . . . .	4
1.5 Chip intel ETANN. . . . .	9
3.1 a) Foto do SCIAMACHY a bordo do ENVISAT; b) Esquema da faixa espectral utilizada pelo instrumento SCIAMACHY. . . . .	18
3.2 Visão do satélite OCO-2. . . . .	22
3.3 Visão detalhada do satélite Ibuki (GOSAT). . . . .	24
4.1 Medida simulada de radiância (normalizada pelo espectro solar) do SCIAMACHY em modo nadir, associada a uma atmosfera tropical, calculado com o SCIATRAN (modo k-correlacionado (c-k); depois de efetuada convolução com função fenda Gaussiana de 0.3 nm (FWHM). Cenário: atmosfera U.S. Standard, espalhamento múltiplo, ângulo solar zenital $60^\circ$ , albedo 0.1, ausência de nuvens. . . . .	33
4.2 Medida simulada de radiância (normalizada pelo espectro solar) do SCIAMACHY em modo nadir, associada a uma atmosfera subártica, calculada com o SCIATRAN (modo k-correlacionado (c-k); depois de efetuada convolução com função fenda Gaussiana de 0.3 nm (FWHM). Cenário: atmosfera U.S. Standard, espalhamento múltiplo, ângulo solar zenital $60^\circ$ , albedo 0.1, ausência de nuvens. . . . .	34
4.3 Representação esquemática de problemas direto e inverso. . . . .	35
5.1 Um Neurônio Biológico. . . . .	49
5.2 Elementos básicos de um neurônio artificial. . . . .	50
5.3 Exemplos de Função de ativação:(a) Função Degrau; (b) Função Sig-móide; (c) Função linear; (d) Função gaussiana . . . . .	53
5.4 Função sigmóide tanh (azul) e sua derivada (vermelho) . . . . .	54

5.5	Ilustração de uma rede multicamadas. . . . .	56
5.6	Rede Neural recorrente. . . . .	63
6.1	Diagrama com a metodologia empregada. As cores servem para identificar e agrupar um conjunto de tarefas com características comuns. . . . .	66
6.2	Jacobiano do $CO_2$ (figura inferior) para o canal 6 do SCIAMACHY e ampliação da região espectral com maior variação observada (topo) . . .	70
6.3	Jacobiano do $CH_4$ (figura inferior) para o canal 6 do SCIAMACHY e ampliação da região espectral com maior variação observada (topo) . . .	70
6.4	Exemplos de codificação binária de duas soluções candidatas à arquitetura de rede e sua representação decimal correspondente. . . . .	97
6.5	Crescimento exponencial do espaço de busca. Gráfico que exhibe o número de possíveis arquiteturas de RNA versus número de bits na solução codificada. . . . .	98
6.6	Critério de parada baseado na perda de generalização da RNA . . . . .	100
6.7	Síntese da metodologia empregada para o problema da otimização de uma arquitetura de RNA do tipo <i>feedforward</i> . . . . .	101
6.8	Visualização tridimensional da função de penalidade aplicada ao problema.	105
7.1	Canais SCIAMACHY utilizados pelo modelo de inversão proposto para inferir perfis verticais de concentração de $CO_2$ (superior) e $CH_4$ (inferior) selecionados pelos métodos OSP (azul), conteúdo de informação (vermelho) e pelo método dos graus de liberdade (preto). . . . .	107
7.2	(a) Sensibilidade para uma perturbação de 8 ppmv no perfil de $CO_2$ vs. comprimento de onda na janela espectral de 1000-1750 nm do SCIAMACHY (canal 6). (b) Sensibilidade para uma perturbação de 0.05 ppmv no perfil de $CH_4$ vs. comprimento de onda na janela espectral de 1000-1750 nm do SCIAMACHY (canal 6). (c) RSI (ver seção 6.1.1 do capítulo anterior) para o $CO_2$ . (d) RSI para o $CH_4$ . . . . .	108
7.3	Relação Sinal-Ruído (RSR) obtida para os canais do sensor SCIAMACHY durante testes de calibração. As relações foram determinadas para os sinais de radiância máximo e mínimo. (gráficos: DLR-FMI). . . .	110
7.4	Relação Sinal/Ruído (RSR) para o canal 6 do sensor SCIAMACHY. . . .	110
7.5	Sensibilidade dos canais do SCIAMACHY para (a) $H_2O$ , (b) $CO$ , (c) $N_2O$ e (d) T para a faixa de comprimento de onda de 1000-1750nm (banda do canal 6) e uma situação representativa tropical. . . . .	111
7.6	Classificação de canais do sensor SCIAMACHY na seleção do conteúdo de informação para o gás $CO_2$ em uma atmosfera tropical. . . . .	115

7.7	Classificação de canais do sensor SCIAMACHY na seleção do conteúdo de informação para o gás $CH_4$ em uma atmosfera tropical. . . . .	115
7.8	Ganho em informação para uma situação representativa de uma atmosfera tropical após a seleção de 1,2,3,4,5,10 e 25 canais respectivamente, usando o método do conteúdo de informação para o gás $CO_2$ . . . . .	119
7.9	Ganho em informação para uma representativa situação atmosférica tropical após a seleção de 1,2,3,4,5,10 e 25 canais respectivamente, usando o método do conteúdo de informação para o gás $CH_4$ . . . . .	119
7.10	Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude tropical para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	127
7.11	Erro (rms) obtido na recuperação de um perfil (tropical) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	127
7.12	Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude média no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	128
7.13	Erro (rms) obtido na recuperação de um perfil (latitude média no inverno) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	128
7.14	Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude média no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	129
7.15	Erro (rms) obtido na recuperação de um perfil (latitude média no verão) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	129

7.16	Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude subártica no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	130
7.17	Erro (rms) obtido na recuperação de um perfil (latitude subártica no inverno) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	130
7.18	Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude subártica no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	131
7.19	Erro (rms) obtido na recuperação de um perfil (latitude subártica no verão) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	131
7.20	Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do $CO_2$ para o cenário 1 (inversão global ou generalizada). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica. . . . .	132
7.21	Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do $CH_4$ para o cenário 1 (inversão global ou generalizada). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica. . . . .	132

7.22	Resultados para a recuperação da taxa de mistura (em ppmv) de $CO_2$ e $CH_4$ em região de latitude tropical para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	135
7.23	Erro (rms) obtido na recuperação de um perfil (tropical) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	135
7.24	Resultados para a recuperação da taxa de mistura (em ppmv) de $CO_2$ e $CH_4$ em região de latitude média no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	136
7.25	Erro (rms) obtido na recuperação de um perfil (latitude média no inverno) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	136
7.26	Resultados para a recuperação da taxa de mistura (em ppmv) de $CO_2$ e $CH_4$ em região de latitude média no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	137
7.27	Erro (rms) obtido na recuperação de um perfil (latitude média no verão) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	137

7.28	Resultados para a recuperação da taxa de mistura (em ppmv) de $CO_2$ e $CH_4$ em região de latitude subártica no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	138
7.29	Erro (rms) obtido na recuperação de um perfil (latitude subártica no inverno) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	138
7.30	Resultados para a recuperação da taxa de mistura (em ppmv) de $CO_2$ e $CH_4$ em região de latitude subártica no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI). . . . .	139
7.31	Erro (rms) obtido na recuperação de um perfil (latitude subártica no verão) vertical de concentração de $CO_2$ e $CH_4$ ordenado por estratégia de seleção de canais e camadas atmosféricas. . . . .	139
7.32	Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do $CO_2$ para o cenário 2 (inversão Local ou distribuída). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica. . . . .	140
7.33	Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do $CH_4$ para o cenário 2 (inversão Local ou distribuída). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica. . . . .	140

7.34	Ganho comparativo (em termos de erro RMS) do cenário 2 (Inversão local distribuída) em relação ao cenário 1 (Inversão generalizada ou global) na recuperação do $CO_2$ . O quociente $r$ corresponde à divisão entre o erro RMS obtido no cenário 2 e o erro rms obtido no cenário 1 agrupado e totalizado por camada atmosférica (Inferior direito), conjunto de canais selecionados (Superior direito) e faixa latitudinal (Esquerdo). . . . .	142
7.35	Ganho comparativo (em termos de erro RMS) do cenário 2 (Inversão local distribuída) em relação ao cenário 1 (Inversão generalizada ou global) na recuperação do $CH_4$ . O quociente $r$ corresponde à divisão entre o erro RMS obtido no cenário 2 e o erro rms obtido no cenário 1 agrupado e totalizado por camada atmosférica (Inferior direito), conjunto de canais selecionados (Superior direito) e faixa latitudinal (Esquerdo). . . . .	143
7.36	Perfis verticais de $CO_2$ (ppbv) para o dia 11 de outubro de 2005 sobre Mauna Loa (Havaí), usando três subconjuntos de canais SCIAMACHY no modelo inverso, juntamente com dados observados e o desvio padrão (erro de medida). . . . .	145
7.37	Ampliação da figura 7.36 mostrando detalhes dos perfis recuperados pelos 3 conjuntos de canais do sensor SCIAMACHY juntamente com dados observados e o desvio padrão (erro de medida). . . . .	145
7.38	Perfis verticais de $CH_4$ (ppbv) para o dia 11 de outubro de 2005 sobre Mauna Loa (Havaí), usando três subconjuntos de canais SCIAMACHY no modelo inverso, juntamente com dados observados e o desvio padrão (erro de medida). . . . .	146
7.39	Ampliação da figura 7.38 mostrando detalhes dos perfis recuperados pelos 3 conjuntos de canais do sensor SCIAMACHY juntamente com dados observados e o desvio padrão (erro de medida). . . . .	147
7.40	(Superior) Recuperação de Fração Molar de Dióxido de Carbono Atmosférico ( $CO_2$ ) (em ppmv) obtido a partir de medições quasi-contínuas em Mauna Loa, no Havaí para o ano de 2005. Para a inversão utilizou-se uma RNA PMC com duas camadas ocultas, uma com 3 neurônios em cada camada e função de ativação Logarítmica. (Inferior) Erro residual (em porcentagem). . . . .	147

7.41 (Superior) Recuperação de Fração Molar de Metano Atmosférico ( $CH_4$ ) (em ppmv) obtido a partir de medições quasi-contínuas em Mauna Loa, no Havaí para o ano de 2005. Para a inversão utilizou-se uma RNA PMC com função de ativação do tipo Tangente Hiperbólica duas camadas ocultas, uma com 21 neurônios e outra com 1. (Inferior) Erro residual (em porcentagem). . . . .	148
7.42 (Superior) Recuperação de Fração Molar de Dióxido de Carbono Atmosférico ( $CO_2$ ) (em ppmv) obtido a partir de medições quasi-contínuas no Pólo Sul para o ano de 2007. A tabela 7.11 Para a inversão utilizou-se uma RNA PMC com uma única camada oculta de 30 neurônios e função de ativação sigmoidal. (Inferior) Erro residual (em porcentagem). . . . .	148
7.43 (Superior) Recuperação de Fração Molar de Metano Atmosférico ( $CH_4$ ) (em ppmv) obtido a partir de medições quasi-contínuas no Pólo Sul para o ano de 2007. Para a inversão utilizou-se uma RNA PMC com função de ativação Sigmoidal com duas camadas, uma com 4 neurônios e outra com 6. (Inferior) Erro residual (em porcentagem). . . . .	149
7.44 Passagem do satélite sobre a região de Mauna Loa no dia 11 de outubro de 2005. . . . .	149
7.45 Passagem do satélite sobre a região do Pólo Sul dia 13 de setembro de 2007. . . . .	150



## LISTA DE TABELAS

	<u>Pág.</u>
3.1 Características dos canais espectrais do SCIAMACHY. . . . .	19
4.1 Símbolos usados na ETR . . . . .	31
6.1 Intervalo de valores assumidos pelos parâmetros que definem uma arquitetura de rede. . . . .	96
7.1 Canais do SCIAMACHY (em nm) utilizados para recuperar o perfil vertical de concentração de $CO_2$ escolhidos pelo método OSP. . . . .	112
7.2 Canais do SCIAMACHY (em nm) utilizados para recuperar o perfil vertical de concentração de $CH_4$ . . . . .	112
7.3 Os primeiros 29 canais selecionados para recuperação de $CO_2$ com máxima contribuição para o total do conteúdo de informação. . . . .	117
7.4 Os primeiros 25 canais selecionados para recuperação de $CH_4$ com máxima contribuição para o total do conteúdo de informação. . . . .	118
7.5 Os primeiros 23 canais selecionados para recuperação de $CO_2$ com maior número de graus de liberdade. . . . .	120
7.6 Os primeiros 21 canais selecionados para recuperação de $CH_4$ com maior número de graus de liberdade. . . . .	121
7.7 Distribuição dos níveis de altitude/concentração para cada camada atmosférica e região climática do perfil atmosférico. . . . .	123
7.8 Melhores arquiteturas de RNA para recuperação de $CO_2$ atmosférico encontradas pelos algoritmos VNS, SA, GEO e AG para 5000 avaliações da função objetivo. Nesses experimentos os fatores de ponderação atribuídos ao erros de treinamento ( $\rho_1$ ) e generalização ( $\rho_2$ ) foram de 0.1 e 1.0, respectivamente. . . . .	124
7.9 Melhores arquiteturas de RNA para recuperação de $CH_4$ atmosférico encontradas pelos algoritmos VNS, SA, GEO e AG para 5000 avaliações da função objetivo. Nesses experimentos os fatores de ponderação atribuídos ao erros de treinamento ( $\rho_1$ ) e generalização ( $\rho_2$ ) foram de 0.1 e 1.0, respectivamente. . . . .	124
7.10 Arquitetura de rede, obtida pela aplicação da estratégia otimizante, para a recuperação de perfil vertical de concentração de $CO_2$ e $CH_4$ agrupada por método de seleção de canais. . . . .	126

7.11	Arquitetura de rede obtida para a recuperação de perfil vertical de concentração de $CO_2$ no cenário 2 (Inversão localizada distribuída) agrupada por método de seleção de canais e camada atmosférica. . . . .	134
7.12	Arquitetura de rede obtida para a recuperação de perfil vertical de concentração de $CH_4$ no cenário 2 (Inversão localizada distribuída) agrupada por método de seleção de canais e camada atmosférica. . . . .	134

## LISTA DE ABREVIATURAS E SIGLAS

ADEOS	– Advanced Earth Observing Satellite
AG	– Algoritmo Genético
AGS	– Algoritmo Genético Simples
AIRS	– Advanced Infrared Radiation Sounder
AIRS-TCA	– AIRS Team Core Algorithm
ANN	– Artificial Neural Network
AMF	– Air Mass Factor
CDIAC	– Carbon Dioxide Information Analysis Center
CE	– Computação Evolutiva
CI	– Conteúdo de Informação
CMDL	– Climate Monitoring and Diagnostics Laboratory
CNES/EUMETSAT	– Centre National d’Etudes Spaciales/ European Organisation for the Exploitation of Meteorological Satellites
ENVISAT	– Environmental Satellite
EOS	– Earth Observing System
ESA	– European Space Agency
ETANN	– Electrically Trainable Analog Neural Network
ETR	– Equação de Transferência Radiativa
FBR	– Função de Base Radial
FWHM	– Full Width at Half Minimum
FRESCO	– Fast Retrieval Scheme for Cloud Observables
GDP	– GOME Data Processor
GEE	– Gás de Efeito Estufa
GEO	– Generalized Extremal Optimization
GLOBALVIEW-CO2	– Cooperative Atmospheric Data Integration Project
GLS	– Graus de Liberdade do Sinal
GMD	– Global Monitoring Division
GOME	– Global Ozone Monitoring Instrument
GOSAT	– Greenhouse gases Observing Satellite
HIRS	– High Resolution Infrared Radiation Sounder
IASI	– Infrared Atmospheric Sounder Interferometer
IUP/IFE	– Institut für UmweltPhysik / Institut für FernErkundung
IMG	– Interferometric Monitor for Greenhouse gases)
IMLM	– Iterative Maximum Likelihood Method
INPE	– Instituto Nacional de Pesquisas Especiais
IPCC	– Intergovernmental Panel on Climate Change
LR	– Learning Rate
MPI	– Message Passing Interface
NASA	– National Aeronautics And Space Administration

NIR	–	Near Infrared
NOAA	–	National Oceanic and Atmospheric Administration
OCO	–	Orbit Carbon Observatory
OMI	–	Ozone Monitoring Instrument
OSP	–	Optimal Sensitivity Profile
PAD	–	Processamento de Alto Desempenho
PMC	–	Perceptron de Múltiplas Camadas
PMD	–	Polarization Mode Dispersion
PVM	–	Parallel Virtual Machine
RMS	–	Root Mean Square
RNA	–	Rede Neural Artificial
RSI	–	Relação Sinal Interferência
RSR	–	Relação Sinal Ruído
SA	–	Simulated Annealing
SCIAMACHY	–	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SCIATRAN	–	
SGA	–	Simple Genetic Algorithm
STI	–	Signal To Interference
SWI	–	Short Wave Infrared
TANSO-CAI	–	Thermal And Near infrared Sensor for carbon Observation - Cloud and Aerosol Imager
TANSO-FTS	–	Thermal And Near infrared Sensor for carbon Observation - Fourier Transform Spectrometer
TCCON	–	Total Carbon Column Observing Network
TDNN	-	Time Delay Neural Network
TIGR	–	Thermodynamic Initial Guess Retrieval
TIR	-	Thermal Infrared
TIROS	–	Television Infrared Observation Satellite
TOA	–	Top Of Atmosphere
TOSOMI	–	Total Ozone retrieval scheme for SCIAMACHY based on the OMI DOAS algorithm
TOVS	–	Tiros Operational Vertical Sounder
UTC	–	Universal Time Coordinated
VNS	–	Variable Neighborhood Search
WDCGG	–	World Data Center for Greenhouse Gases
WFM-DOAS	–	Weighting Function Modified Differential Optical Absorption Spectroscopy
WMO	–	World Meteorological Organization

## LISTA DE SÍMBOLOS

$A$	– operador diferencial
$B_\lambda$	– função de Planck
$d$	– Vetor alvo (desejado) na saída das RNA
$E_{trein}$	– erro de treinamento
$E_{gen}$	– erro de generalização
$F$	– espaço de soluções
$g$	– aceleração da gravidade
$H$	– função peso ou Jacobiano
$I_\lambda$	– intensidade de radiação monocromática (radiância monocromática)
$K$	– matriz Jacobiana normalizada pela matriz de covariância do erro
$p$	– pressão atmosférica
$p_0$	– pressão no topo da atmosfera
$p_s$	– pressão na superfície
$o$	– vetor obtido na saída das RNA
$q$	– razão de mistura
$r$	– coeficiente de correlação de Pearson
$S_a$	– matriz de covariância da estimativa a priori
$S_M$	– matriz de covariância do erro da medida
$S_F$	– matriz de covariância do erro do modelo direto
$S_e$	– covariância do erro global
$T$	– temperatura
$U$	– espaço de observações
$x$	– vetor de estado atmosférico
$x_a$	– estimativa a priori
$z$	– medida de distância na direção normal ao plano de estratificação (altura)
$y_j$	– Sinal na saída do neurônio
$w$	– vetor de pesos de uma RNA
$\varepsilon_\lambda$	– emissividade da superfície
$\theta$	– ângulo zenital
$\kappa_\lambda$	– coeficiente de absorção
$\lambda$	– Comprimento de Onda
$\mu$	– coseno do ângulo zenital
$\tau$	– profundidade ótica
$\phi$	– função resposta do sensor
$\varphi$	– função de ativação
$\mathfrak{S}_\lambda$	– transmitância
$\Delta\lambda$	– largura da banda espectral
$\alpha$	– taxa de aprendizagem
$\zeta$	– razão de resfriamento

- $\rho_1$  – fator de ponderação atribuído ao erro de treinamento
- $\rho_2$  – fator de ponderação atribuído ao erro de generalização
- $\eta$  – constante de momentum

## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
<b>2 PRINCIPAIS GASES DE EFEITO ESTUFA (GEE)</b> . . . . .	<b>11</b>
2.1 Dióxido de Carbono . . . . .	11
2.2 Metano . . . . .	13
<b>3 OS SENSORES EMBARCADOS</b> . . . . .	<b>17</b>
3.0.1 O Sensor Sciamachy . . . . .	18
3.0.2 O Sensor AIRS . . . . .	20
3.0.3 O Observatório Orbital de Carbono (OCO) . . . . .	22
3.0.4 GOSAT . . . . .	23
<b>4 TEORIA DA SONDAÇÃO REMOTA DE GASES DE EFEITO ESTUFA</b> . . . . .	<b>27</b>
4.1 A Equação de Transferência Radiativa (ETR) . . . . .	27
4.2 Modelo Direto . . . . .	30
4.2.1 SCIATRAN . . . . .	31
4.3 Modelo Inverso . . . . .	32
4.3.1 Conceito de Problema Inverso . . . . .	34
4.3.2 O conceito de problemas Mal-Postos . . . . .	36
4.3.3 Os principais algoritmos de inversão baseados no sensor SCIAMACHY . . . . .	37
4.3.3.1 WFM-DOAS . . . . .	37
4.3.3.2 AIRS-Team Retrieval For Core Products and Geophysical Parameters . . . . .	39
4.3.3.3 O Método Iterativo de Máxima Verossimilhança (IMLM) . . . . .	40
4.3.3.4 O esquema 4D-VAR . . . . .	41
4.3.3.5 Esquema de Inversão baseada em Estimação Ótima . . . . .	43
4.3.3.6 O algoritmo TOSOMI . . . . .	44
4.3.3.7 Considerações gerais sobre os métodos de inversão . . . . .	45
<b>5 REDES NEURAIS ARTIFICIAIS</b> . . . . .	<b>47</b>
5.1 Breve histórico acerca das Redes Neurais Artificiais . . . . .	47
5.1.1 O Neurônio biológico . . . . .	49

5.1.2	O Neurônio Artificial . . . . .	50
5.2	Arquitetura de Redes Neurais . . . . .	54
5.2.1	Redes Neurais Multicamadas . . . . .	55
5.2.2	Perceptrons Multi-Camadas . . . . .	56
5.2.2.1	Capacidade de representação do PMC . . . . .	57
5.2.2.2	Aprendizagem da Rede . . . . .	57
5.2.3	Outros tipos de Redes Neurais . . . . .	60
5.2.3.1	Rede Função de Base Radial . . . . .	60
5.2.3.2	Redes Neurais Recorrentes . . . . .	61
5.2.3.2.1	- Redes de Elman e Jordan . . . . .	62
<b>6</b>	<b>METODOLOGIA PROPOSTA . . . . .</b>	<b>65</b>
6.1	Seleção de Canais . . . . .	66
6.1.1	Método OSP . . . . .	68
6.1.1.1	Cálculo dos Jacobianos . . . . .	69
6.1.1.2	Análise de Variância . . . . .	71
6.1.2	Seleção de canais baseada no Conteúdo de Informação (CI) e nos Graus de Liberdade do Sinal (GLS) . . . . .	73
6.1.2.1	Conteúdo de Informação . . . . .	73
6.1.3	Graus de Liberdade do Sinal (GLS) . . . . .	75
6.1.3.1	Descrição Detalhada do Método . . . . .	77
6.2	Geração da Base de Dados de Treinamento da Rede . . . . .	78
6.3	Otimização dos Parâmetros da Rede . . . . .	82
6.3.1	Computação Evolucionária . . . . .	84
6.3.2	Método Proposto . . . . .	86
6.3.3	Variable Neighbourhood Search (VNS) . . . . .	87
6.3.4	Otimização extrema generalizada (GEO) . . . . .	88
6.3.5	Recozimento Simulado . . . . .	91
6.3.6	Algoritmos Genéticos . . . . .	94
6.3.7	Representação de uma solução candidata . . . . .	96
6.3.8	Aprendizado e teste da rede . . . . .	99
6.3.9	Função objetivo . . . . .	100
6.3.9.1	Detalhando o critério usado para compor a função objetivo . . . . .	102
6.3.9.1.1	- O critério do Erro de Treinamento . . . . .	102
6.3.9.1.2	- Critério do erro de generalização . . . . .	102
6.3.9.1.3	- Penalidade devida à complexidade da arquitetura de rede . . . . .	103



6.4	Recuperação dos Perfis de Concentração de Metano e Dióxido de Carbono	105
6.4.0.1.4	- Cenário 1: Inversão global	106
6.4.0.1.5	- Cenário 2: Inversão localizada distribuída por camadas atmosféricas	106
<b>7</b>	<b>RESULTADOS</b>	<b>107</b>
7.1	Resultados para Seleção de Canais	107
7.1.1	Método OSP ( <i>Optimal Sensibility Profile</i> )	108
7.1.1.1	Resultados dos Estudos de sensibilidade para os canais do SCIA-MACHY	108
7.1.2	Conteúdo de Informação (CI) de Shanon para o sinal de $CO_2$ e $CH_4$	113
7.1.3	Graus de liberdade para o Sinal de $CO_2$ e $CH_4$	117
7.2	Resultados para a Recuperação de Perfis de Concentração de $CO_2$ e $CH_4$	121
7.2.0.1	Cenário 1: Inversão global	122
7.2.1	Cenário 2: Inversão localizada distribuída por camadas atmosféricas	133
7.3	Comparação entre os dois cenários de inversão propostos	142
7.4	Validação do Modelo Proposto	144
<b>8</b>	<b>CONCLUSÃO</b>	<b>153</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>157</b>



# 1 INTRODUÇÃO

Uma das grandes preocupações da comunidade científica na última década refere-se às mudanças climáticas e suas consequências para a humanidade. Os três anos mais quentes dos últimos mil anos da história da Terra aconteceram na Última década. Segundo o “Intergovernmental Panel on Climate Change” (IPCC), a temperatura média global aumentou entre 0.3 e 0.6°C desde o final do século passado, e pela escala dos diferentes cenários desenvolvidos pelo IPCC em seu relatório “Special Report on Emission Scenarios” (SRES) a média global da temperatura do ar, projetada por modelos numéricos, sofrerá uma elevação de 1.4° a 5.8° em 2100. É provável que a maior parte desse aquecimento seja decorrente da emissão, por atividades humanas, de gases que retêm radiação térmica. As moléculas de alguns gases presentes na atmosfera como o dióxido de carbono ( $CO_2$ ), o monóxido de carbono (CO), o metano ( $CH_4$ ), o ozônio ( $O_3$ ), entre outros, interagem com a radiação eletromagnética na faixa do espectro conhecida como infravermelho termal e com isso dificultam a perda para o espaço da radiação térmica, produzindo o aquecimento da superfície da Terra. Esse fenômeno é denominado “efeito-estufa”. A Figura 1.1 ilustra este processo.



Figura 1.1 - representação esquemática do Efeito Estufa.  
Fonte: IPCC (2007)

Esses gases são fundamentais para manter o equilíbrio climático e condições ambientais adequadas para a vida na Terra, isto é, temperaturas que permitam a existência de água nas formas líquida e gasosa essencial à vida e ao ciclo hidrológico, respectivamente. Atualmente, as análises sistemáticas do IPCC ([INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE \(IPCC\), 2007](#)), que sintetizam o conhecimento científico sobre o sistema climático, levam a um razoável consenso de que o aquecimento global observado nos últimos 50 anos é explicado principalmente pelas emissões humanas de gases de efeito estufa (GEE) e de aerossóis e não por eventual variabilidade natural do clima.

Metano ( $CH_4$ ) e dióxido de carbono ( $CO_2$ ) são os mais importantes gases de efeito estufa. Apesar da ratificação do protocolo de Kyoto, a previsão é que a liberação de  $CO_2$  e  $CH_4$  na atmosfera continuará a aumentar ano a ano ([INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE \(IPCC\), 2007](#)). Estudos recentes mostram que as concentrações atmosféricas desses dois gases aumentaram muito desde os tempos pré-industriais (ver Figura 1.2). ([LELIEVELD J.; CRUTZEN; DENTENER, 1998](#)) e ([MCCARTHY et al., 2001](#)) demonstraram que mais da metade das emissões atuais desses gases são de origem antropogênica ([SCIENCES; CLIMATE, 2008](#)).

Nas Figuras 1.3 e 1.4 exibimos, a título de ilustração, uma representação tridimensional da distribuição global de  $CO_2$  e  $CH_4$  atmosféricos para a camada limite planetária. Foram utilizados dados da "Cooperative Air Sampling Network". O espaçamento da grade regular de 7 dias no tempo e  $10^\circ$  na latitude é obtido através de suavização e interpolação dos dados. Os dados reais utilizados para construir a superfície estão disponíveis gratuitamente a partir de GMD, CDIAC e WDCGG WMO (ver lista de siglas).

Estas duas espécies de gases absorvem radiação emergente do sistema terra-atmosfera na faixa espectral do infravermelho e desempenham um papel importante no efeito estufa e nas mudanças climáticas sobre o globo terrestre. Além disso, ambos são gases de longa vida média na atmosfera, de modo que as quantidades atuais emitidas naturalmente ou por ações antropogênicas podem permanecer na atmosfera e influenciar o clima no planeta durante décadas, caso do metano, ou centenas de anos, caso do dióxido de carbono, de acordo com Hartman ([HARTMANN, 1994](#)).

Entre os principais efeitos causados pelo aumento dos GEE sobre diversas regiões do globo terrestre, incluindo o Brasil, já foram observados os seguintes: intensi-

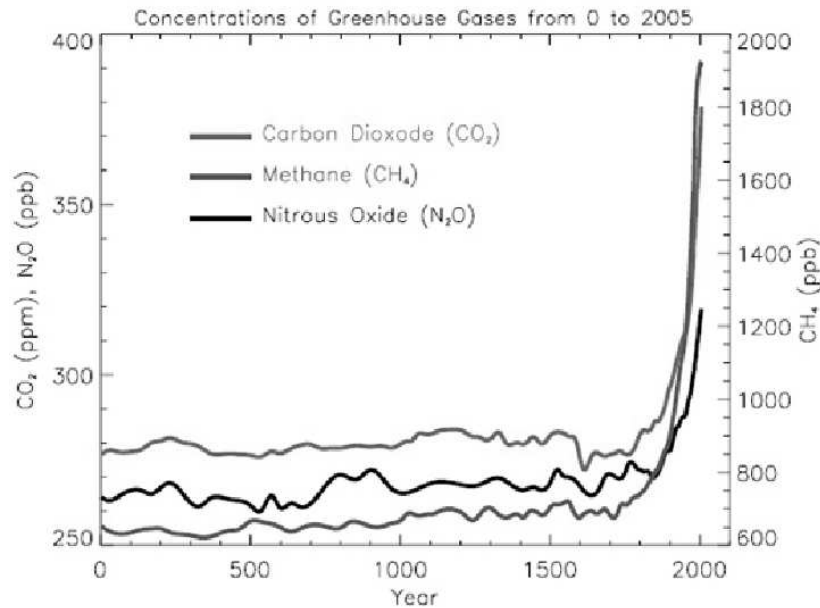


Figura 1.2 - Concentrações atmosféricas mundiais de dióxido de carbono ( $CO_2$ ), metano ( $CH_4$ ), e óxido nitroso ( $N_2O$ ). Apresenta as mudanças nas concentrações atmosféricas nos últimos 2000 anos.  
Fonte: IPCC (2007)

ficção da atividade convectiva, déficits de precipitação sobre regiões semi-áridas, elevação das temperaturas da superfície durante o verão de cada hemisfério, elevação do nível do mar, entre outros (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2007). Como conseqüências, são previstas mais secas, enchentes e queimadas nos trópicos e nos subtropicos, redução geral no potencial de produção agrícola em vários países, prejuízos à fauna e à flora, maior demanda por refrigeração durante o verão, avanço do mar sobre algumas cidades litorâneas, e outros (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2007). Simulações numéricas realizadas com mais de 20 modelos climáticos diferentes mostraram ainda que, para uma duplicação na concentração do dióxido de carbono na atmosfera a partir do nível de 250 ppmv (partes por milhão de volume), ter-se-ia um aumento na temperatura média da superfície entre 2-5 K e, conseqüentemente, sérias variações climáticas no globo terrestre [ver (HARTMANN, 1994),(INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2001) e (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2007)].

O conhecimento da distribuição das fontes e dos sumidouros do dióxido de carbono,

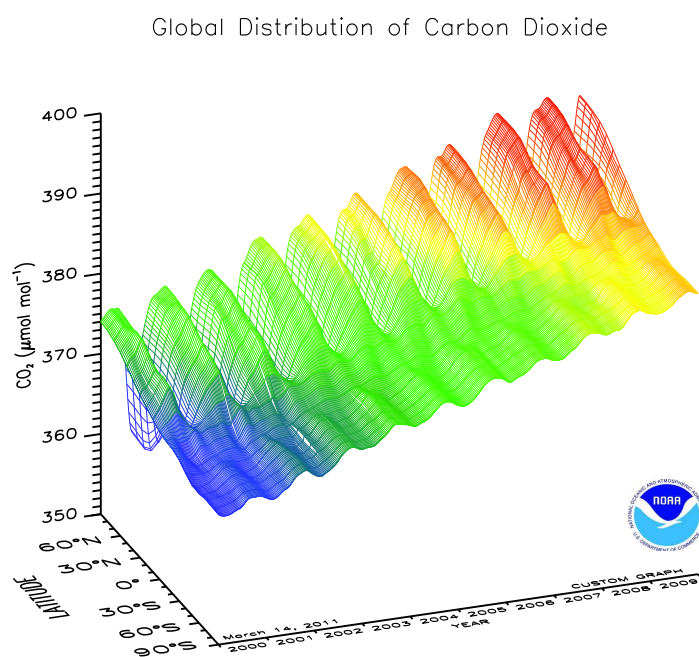


Figura 1.3 - Variações temporais e latitudinais de  $CO_2$  na superfície observadas em MAUNA LOA, US pelas estações de medida da NOAA.  
 Fonte: <http://www.esrl.noaa.gov/gmd/ccgg/iadv/> (2010)

do metano e do óxido nitroso sobre o globo terrestre é crucial para quantificar, entender e prever a evolução do ciclo dos gases constituintes do efeito estufa no sistema terra-atmosfera, e para o desenvolvimento e aplicação de novas políticas de redução da emissão destes gases no futuro. Este aspecto é relevante não apenas para os países mais desenvolvidos, que detêm os maiores parques industriais do planeta, mas também para o Brasil, que possui um parque industrial razoável a ponto de contribuir para a emissão do dióxido de carbono e do óxido nitroso na atmosfera, sobretudo na região Sudeste (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

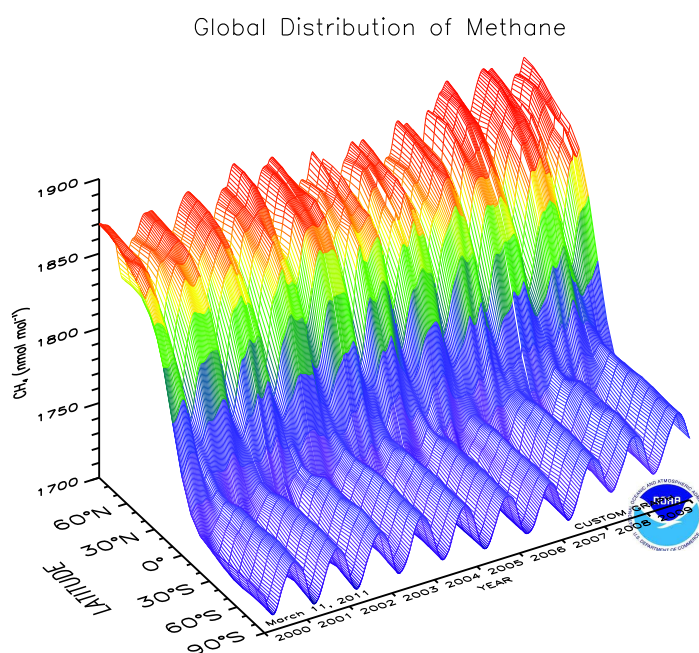


Figura 1.4 - Variações temporais e latitudinais de  $CH_4$  na superfície observadas em MAUNA LOA, US pelas estações de medida da NOAA.  
 Fonte: <http://www.esrl.noaa.gov/gmd/ccgg/iadv/> (2010)

(IPCC, 2007). O crescente desflorestamento na bacia Amazônica e a ampla utilização de pastagens agrícolas também fazem do Brasil um dos países de maior emissão de dióxido de carbono e metano na atmosfera ((NOBRE C. A.; SELLERS; SHUKLA, 1991); (CHU P. S.; YU; HASTENRATH, 1994); (GASH et al., 1996); (FISCH et al., 1997); (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2007)). Deste modo, é fundamental para o Brasil a realização de estudos profundos a respeito do ciclo dos gases do efeito estufa sobre o seu território e do papel de cada uma das suas regiões na geração, na manutenção e na dissipação destes gases na atmosfera.

O monitoramento de gases minoritário da atmosfera é uma tarefa relevante e que deve ser realizada continuamente. Mesmo que a produção de energia mundial deixe de ser baseada em carbono, precisamos monitorar aerossóis e gases minoritários para previsão de tempo e monitoramento climático.

O modelamento inverso de emissões de gases traço a partir de medidas de satélite foi realizado com sucesso para espécies de tempo de vida média curta tais como o  $NO_2$  (MARTIN et al., 2003),  $O_3$  (LIU et al., 2005) e CO ((PÉTRON et al., 2004) e (BUCHWITZ et al., 2007)), e a combinação de ambos (MULLER J. F.; STAVRAKOU, 2005). Considerando os GEE de longa vida média, têm sido relatados estudos para a recuperação de metano atmosférico ((SCHNEISING et al., 2008b);(FRANKENBERG et al., 2008);(XIONG et al., 2008)) e  $CO_2$  ((BUCHWITZ et al., 2006);(BUCHWITZ et al., 2007) e (CREVOISIER, 2004)).

Rayner et al. (2002) investigaram a utilidade de observações de  $CO_2$  a partir de satélites para a identificação de fontes e/ou sumidouros e calcularam a influência da órbita do satélite, geometria de varredura e nuvens (RAYNER et al., 2002). Adicionalmente, Houweling et al. compararam os potenciais benefícios da utilização de três diferentes (existentes e planejados) instrumentos de medida de  $CO_2$  (HOUWELING et al., 2004).

O objetivo desta tese é fazer a recuperação de perfis verticais de  $CO_2$  e  $CH_4$  atmosférico a partir de dados de satélite, utilizando um modelo de inversão baseado em redes neurais artificiais. Existem muitos algoritmos atualmente operacionais que realizam a recuperação da coluna integrada ou da taxa de mistura como mostrado no capítulo 4 na seção 4.3.3. Porém existem poucos registros na literatura que comprovem a realização de inversão para recuperação de perfis verticais de concentração de gases de efeito estufa tais como o  $CO_2$  e o  $CH_4$  (FOUCHER et al., 2010). Mesmo quando consideramos a inversão da coluna integrada desses gases poucas foram as tentativas de realização dessa tarefa utilizando-se modelos inversos baseados em redes neurais artificiais como é proposto nesse projeto de pesquisa. Além disso, a maioria delas, senão todas, limitou-se à inferência em apenas uma camada da atmosfera, a camada troposférica. Neste trabalho propomos estender a inferência até as camadas superiores da mesosfera e termosfera.

O conhecimento da estrutura termodinâmica da atmosfera (temperatura, vapor d'água, ozônio) e da concentração de seus principais componentes atmosféricos per-



mite, por intermédio da equação de transferência radiativa e de modelos diretos, calcular a radiância e a temperatura de brilho medida pelos diferentes canais dos instrumentos embarcados a bordo de satélites. Nesse caso, o problema inverso associado consiste em estimar o estado da atmosfera a partir das radiâncias medidas por estes instrumentos. Particularmente, estaremos interessados em resolver o problema inverso que consiste em estimar as concentrações de  $CO_2$  e  $CH_4$  atmosférico a partir de radiâncias medidas pelos canais do sensor SCIAMACHY. Este problema inverso apresenta inúmeras dificuldades. É preciso salientar que a equação de transferência radiativa (ver equação 4.3), que conecta o "mundo termodinâmico" ao "mundo radiativo", é uma equação integral de Fredholm de primeira espécie; a inversão de tal equação tem sido, por um longo tempo, reconhecido como um exemplo de problema mal-posto (ver seção 4.3.2 do capítulo 4). Problemas inversos são frequentemente problemas mal-postos: os dados de que dispomos são em geral brutos e, se uma solução existe, parâmetros diferentes podem conduzir a observações idênticas de onde surge a não-unicidade da solução (ver seção 4.3.2 do capítulo 4). Além disso, dentro da equação de transferência radiativa, a concentração intervém através da função de transmissão e o problema inverso é, portanto, não-linear. E como se não bastasse, os sinais de  $CO_2$  e  $CH_4$  são relativamente fracos, mesmo para os canais mais sensíveis às variações de concentração do gás, e são da ordem do ruído do instrumento (CREVOISIER C.; CHEDIN; SCOTT, 2003).

Estes diferentes aspectos do problema inverso representado acima fazem com que as técnicas clássicas de estimação se tornem pouco apropriadas. Neste trabalho propomos a utilização de Redes Neurais como uma possível solução para o problema da recuperação do perfil atmosférico de concentração de gases de efeito estufa. A utilização de redes neurais dentro de um quadro de inferência permite trazer uma resposta satisfatória a estes diferentes aspectos de estimação de gases atmosféricos. Por serem não-lineares, o que evita a necessidade de linearização do problema inverso, as redes neurais se revelam bastante eficazes para filtrar os ruídos e não necessitam da escrita de um funcional. Na técnica de inversão direta, uma rede neural é usada para estimar diretamente o mapeamento entre as observações SCIAMACHY e as variáveis geofísicas requeridas. Na realidade, a rede treinada é um modelo estatístico da ETR inversa, fornecendo de uma só vez uma inversão global. Na maioria dos casos o algoritmo de aprendizagem, que é a parte computacional mais intensa, é executado de modo offline e apenas uma vez. Então, a aplicação do modelo de redes neurais para a inversão de observações SCIAMACHY é quase imediata no estágio

operacional: nenhuma regressão e nenhuma computação de jacobiano são requeridos.

A estrutura das redes neurais é essencialmente paralela, favorecendo a velocidade de processamento e a sua implementação em um ambiente de processamento paralelo de alto desempenho (PAD) utilizando bibliotecas de programação paralela tais como MPI (Message Passing Interface) ou PVM (Parallel Virtual Machine). Outra vantagem sobre técnicas físico-estatísticas clássicas é que não é exigido uma boa condição inicial para a inversão. Além disso, o armazenamento de memória requerido é muito pequeno. Não existe necessidade de um modelo direto rápido (necessário em algoritmo de inversão iterativos) no qual a velocidade é geralmente obtida às custas de uma linearização da ETR e assumindo erros Gaussianos descorrelacionados.

Em um sistema de rede neural, a informação pode parecer ter representação redundante, porém, o fato de que ela se encontre distribuída por todos os elementos da rede significa que mesmo que parte da rede seja destruída, a informação contida nesta parte ainda estará presente na rede, e poderá ser recuperada. Portanto, a redundância na representação de informações em uma rede neural, diferente de outros sistemas, transforma-se em uma vantagem, que torna o sistema tolerante a falhas. Os atributos de uma rede neural, tais como aprender através de exemplos, generalizações redundantes, e tolerância a falhas, proporcionam fortes incentivos para a escolha de redes neurais como uma ferramenta apropriada para aproximação e modelagem de sistemas biológicos.

Com o avanço em tecnologias de hardware, existem componentes com funções dedicadas à sistemas com implementações voltadas para redes neurais, o que traz uma velocidade adicional à computação neural. Já se encontra disponível no mercado circuitos integrados (dedicados) que implementam a rede, permitindo a compactação do hardware, ganho na velocidade de processamento, e maior nível de distribuição no controle. A Figura 1.5 abaixo mostra um microchip ETANN (acrônimo para rede neural analógica eletronicamente treinável) produzido pela intel nos anos 90 (SERBEDZIJA, 1996).

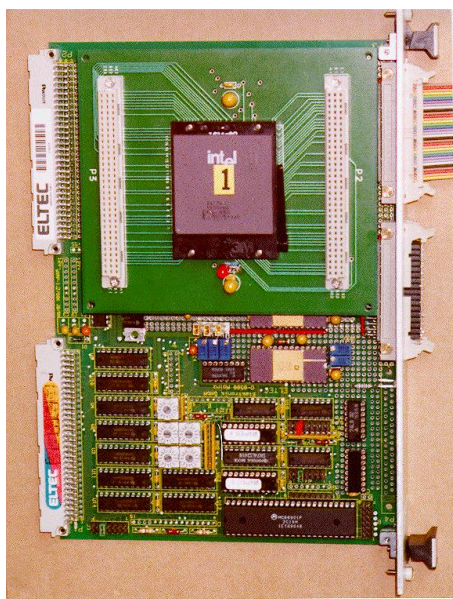


Figura 1.5 - Chip intel ETANN.  
Fonte: Holler (1989)

Outra importante contribuição nessa área de pesquisa foi dada pelo trabalho de (VELHO H. F.; SHIGUEMORI; SILVA, 2007) que desenvolveu um sensor para estimar o perfil vertical de temperatura atmosférica. Neste sensor uma RNA é implementada em um dispositivo lógico programável chamado FPGA (Field Program Gateway Away).

As redes neurais já foram utilizadas com sucesso dentro do quadro da modelagem direta de parâmetros radiativos ((RIEU et al., 1996); (CHEVALLIER et al., 1998)) e dentro do quadro da estimação de parâmetros atmosféricos: variáveis termodinâmicas ((ESCOBAR-MUNOZ et al., 1993); (SHIGUEMORI et al., 2006)), perfis de umidade e características da superfície ((AIRES et al., 2001); (SHIGUEMORI et al., 2006)), taxa de mistura de CO ((CLERBAUX et al., 1999);(CLERBAUX et al., 2005)) e taxa de mistura de Ozônio (TURQUETY S.; HADJI-LAZARO; CLERBAUX, 2002). E, por último, elas se revelam particularmente adequadas à estimação da taxa de mistura de  $CO_2$  a partir de observações TOVS (CHEDIN et al., 2003) e do sensor AIRS (CREVOISIER C.; CHEDIN; SCOTT, 2003).

Em uma primeira etapa propõe-se a utilização do modelo direto SCIATRAN (ver seção 4.2.1 do capítulo 4) para simularmos os dados de radiância na faixa espectral

do infravermelho próximo a partir de perfis perturbados do gás. Através da aplicação do método da perturbação sobre um conjunto de perfis atmosféricos e climatológicos presentes na base de dados do modelo iremos proceder à seleção de canais mais adequados à tarefa de recuperação do GEE. Essa tarefa é explicada em maiores detalhes no capítulo 6.1. Uma vez escolhidos os canais mais apropriados para a recuperação do perfil partimos para a inversão propriamente dita. Para tanto utilizamos o modelo de inversão próprio baseado em redes neurais. Uma breve explanação a respeito da teoria das redes neurais e de sua utilização como instrumento de auxílio nessa área do sensoriamento remoto é apresentada no capítulo 5. No capítulo 6.3 apresentamos uma metodologia desenvolvida para otimizarmos os parâmetros de configuração do modelo inverso utilizado. No caso específico das redes neurais artificiais foram implementados algoritmos baseados em metaheurísticas híbridas de busca. Por fim é realizada a validação dos resultados obtidos pelo modelo.

## 2 PRINCIPAIS GASES DE EFEITO ESTUFA (GEE)

### 2.1 Dióxido de Carbono

O uso de combustíveis fósseis e a prática do desflorestamento visando suprir a demanda mundial de energia e alimentos têm conduzido a um aumento das concentrações desse gás. Além do impacto ambiental, a questão de mudanças globais tem grande efeito no mercado econômico. Um exemplo disto é o surgimento do mercado de créditos de carbono. Neste mercado, os países industrializados, que são responsáveis por grande parte da emissão de  $CO_2$  na atmosfera, são obrigados a adquirir cotas de emissão de carbono de países que emitem pouco  $CO_2$ . Países como o Brasil podem investir na geração de energia limpa, além disso, existe bastante espaço para o plantio de florestas, que serviriam como absorvedores de  $CO_2$ . Portanto, é importante para o país aprimorar o conhecimento científico acerca do papel que nossas florestas, lagos, mares e agroecossistemas desempenham dentro do balanço global de carbono. Para isto é necessário obter dados e gerar modelos capazes de quantificar as fontes e sumidouros de carbono.

As fontes e sumidouros de carbono variam muito em escala temporal e espacial. Modelos de transporte atmosférico podem ser usados para avaliar as medidas atmosféricas em termos de fontes regionais e sumidouros usando procedimento de inversão (RAYNER et al., 2002). O conhecimento a respeito do ciclo do  $CO_2$ , tal qual a identificação das regiões que agem como fonte ou sumidouros de  $CO_2$ , está diretamente associado com a acurácia e a resolução espacial e temporal das medidas de concentração destes gases. Atualmente o monitoramento da concentração de  $CO_2$  na atmosfera é feito através de medições realizadas *in situ*, em pontos esparsos sobre diferentes regiões do planeta. A mais completa rede de observação de  $CO_2$  é mantida pelo Carbon Cycle Greenhouse Gases Group CMDL/NOAA, que coordena um esforço internacional para a realização contínua de medidas discretas a partir de observatórios, bases localizadas em torres altas, bases fixas de medidas espalhadas pela superfície da terra e oceano, aeronaves e navios comerciais, etc (GLOBALVIEW-CO2, 2001). Estas medições registram a distribuição temporal e espacial do ciclo de diversos gases atmosféricos, sendo praticamente a única ferramenta disponível atualmente para a compreensão do ciclo global do carbono.

Apesar da boa acurácia, o número limitado de medidas disponíveis continua a ser o maior obstáculo para o avanço do conhecimento do ciclo global do  $CO_2$ . Neste sen-

tido, o uso de medidas por sensoriamento remoto a partir de satélites pode fornecer um conjunto mais denso de medidas aumentando a resolução espacial e temporal das observações, pois as medições podem ser feitas em intervalos de tempo menores e em mais regiões do globo. As medições convencionais podem ser usadas para calibrar os modelos de ciclo de carbono e validar as medidas feitas do espaço. Os sondadores Advanced Infrared Radiation Sounder (AIRS) da NASA, Infrared Atmospheric Sounding Interferometer (IASI) da CNES/EUMETSAT, SCIAMACHY a bordo do ENVISAT e GOSAT são alguns exemplos de instrumentos que podem ser usados com este objetivo. Além disso, o conjunto de dados já existente também poderia servir para este objetivo. Como exemplo, podemos citar os satélites de órbita polar da série TIROS/NOAA (SMITH W. L.; WOOLF; HAYDEN, 1979), que desde 1979 vêm fornecendo medidas da atmosfera e superfície terrestre, totalizando até o presente momento uma série contínua de 23 anos de dados.

Rayner et al. buscaram estabelecer a acurácia mínima necessária na observação da concentração de  $CO_2$  integrado na coluna atmosférica, para que estes dados pudessem ser úteis na identificação dos fluxos de  $CO_2$  na superfície, usando modelos de inversão de fontes e sumidouros de carbono (RAYNER et al., 2002). Os resultados simulados pelo modelo de estimação apontaram que é necessária uma acurácia de 2.5 ppmv ( $< 1\%$ ) na média mensal de  $CO_2$  integrado, com uma resolução horizontal de  $8^\circ \times 10^\circ$ , para ultrapassar a performance dos dados convencionais de superfície [(MILLER et al., 2007); (CHEVALLIER F.; BREON; RAYNER, 2007)].

Diferentes regiões do espectro eletromagnético são sensíveis a variações na concentração de  $CO_2$  na coluna atmosférica. Contudo, algumas bandas espectrais na faixa do infravermelho e infravermelho próximo possuem grande potencial para inferência da coluna integrada de carbono na atmosfera. Engelen et al. (2001) usou um algoritmo de inversão que se baseia na teoria da estimação ótima, para investigar a capacidade da nova geração de sensores de alta resolução espectral que operam na faixa espectral do infravermelho em detectar a coluna integrada de  $CO_2$  (ENGELEN et al., 2001). As simulações mostraram ser possível obter a concentração de  $CO_2$  integrado na coluna troposférica com uma acurácia melhor que 1 ppmv. Contudo, fatores como nuvens e erros nos modelos de transferência radiativa podem degradar os resultados se nenhum cuidado for tomado no sentido de filtrá-los e corrigí-los. Strow et al. comenta que variações climatológicas (hemisféricas e sazonais) na concentração de carbono são grandes o bastante para serem observadas com o sensor

AIRS, que se encontra a bordo do satélite Aqua, principalmente com os canais espectrais localizados na banda espectral de absorção do  $CO_2$  em  $4.26 \mu m$  e  $15 \mu m$  (STROW et al., 2002).

Rayner et al. discutem sobre a viabilidade de se estimar a concentração de  $CO_2$  atmosférico na faixa do infravermelho próximo, usando para isso as informações contidas na luz solar refletida em direção ao espaço na banda de absorção do  $CO_2$  em  $1.61 \mu m$  (RAYNER et al., 2002). O princípio físico que permite esta inferência é baseado no fato que diferentes frequências possuem respostas distintas com relação à absorção por  $CO_2$ , e, por consequência, apresentam diferentes espessuras óticas. Assim, medidas de radiância (em condições de céu claro) em duas frequências diferentes, onde na primeira o  $CO_2$  absorve mais fortemente a radiação e na outra mais fracamente, podem ser usadas para estimar a diferença entre as espessuras óticas de ambas as frequências, que por sua vez está relacionada à concentração de  $CO_2$ . Para reduzir os erros devido à presença de nuvens e aerossóis, as estimativas são calculadas relativamente às medidas de radiância na banda de absorção do  $O_2$  em  $1.27 \mu m$ . As simulações sugerem que o erro na estimativa da coluna de  $CO_2$  é de aproximadamente 0.5%, quando o perfil de temperatura é fornecido com uma precisão de  $\pm 1K$ .

(BOSCH et al., 2006) e (SCHNEISING et al., 2008a) mostraram que  $CO_2$  pode ser recuperado a partir do SCIAMACHY com uma precisão relativa de medida de 1 – 2% assumindo condições de céu claro para médias mensais em uma resolução espacial de cerca de  $7^\circ \times 7^\circ$ . Recentemente, Buchwitz et. al. utilizou o algoritmo WFM-DOAS para recuperar a taxa de mistura desse gás traço a partir de dados do SCIAMACHY [(BUCHWITZ et al., 2006);(BUCHWITZ et al., 2007)]. Também foram desenvolvidos algoritmos para recuperação de  $CO_2$  atmosférico utilizando a estratégia de recuperação baseada no método da estimação ótima de Rodgers (RODGERS, 1997) a partir do sensor AIRS (CREVOISIER C.; CHEDIN; SCOTT, 2003) e GOSAT.

## 2.2 Metano

Concentrações do metano atmosférico aumentaram cerca de 150% desde o período pré-industrial. Os níveis atuais de metano estão em torno de 1.7-1.8 ppmv e são os mais altos de que se tem registro, incluindo os registros de amostras do núcleo das geleiras que representam um período anterior a 1850 a.C. Isto corresponde a um forçamento radiativo de  $0.48 W m^{-2}$ , que corresponde a 20% do forçamento radiativo



total devido aos gases de efeito estufa [(INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2007);(LELIEVELD J.; CRUTZEN; DENTENER, 1998)].

Mais da metade das emissões atuais de  $CH_4$  é de origem antropogênica, sendo que as fontes mais importantes são a queima de combustível fóssil, ruminantes domésticos, cultivo de arroz e manuseio do lixo orgânico (LELIEVELD J.; CRUTZEN; DENTENER, 1998). Emissões naturais se originam principalmente de pântanos e regiões alagadas do planeta. As incertezas no fluxo na maioria dessas categorias são da ordem de 50% em uma escala global e podem ser maiores em escalas regionais. A taxa de crescimento de metano tem grande variabilidade, e as causas desta variabilidade estão sujeitas a consideráveis debates. Portanto, um conhecimento mais aprofundado da distribuição das fontes e sumidouros de metano se faz necessário para compreender quantitativamente as mudanças na sua concentração atmosférica.

Estimativas existentes de fontes de metano podem ser melhoradas via modelagem inversa utilizando um modelo de transporte para inferir as emissões a partir de medidas de concentrações atmosféricas. Esta técnica tem sido aplicada extensivamente com o auxílio de observações de superfície [(HEIN R.; CRUTZEN; HEIMANN, 1997); (HOUWELING et al., 1999); (CHEN Y. H.; PRINN, 2006); (DENTENER et al., 2003); (BUTLER et al., 2005); (FLETCHER et al., 2004); (BERGAMASCHI et al., 2005)]. Contudo, as medidas de superfície estão disponíveis somente a partir de um número limitado de estações de monitoramento, principalmente em regiões remotas. Estas medidas contêm informações estatísticas significativas sobre as emissões somente em escalas (sub-) continentais porque a distância entre as fontes de metano e a localização das observações é grande (HOUWELING et al., 2004). Observações de superfície também estão disponíveis em locais com emissões bem delimitadas. Estas medidas têm uma relação mais direta com as fontes, e, portanto, são úteis para estimação de fontes de escalas regionais (BERGAMASCHI et al., 2005), embora uma representação correta dessas medidas por modelos de transporte permaneça um desafio. Especificamente, se as medidas são influenciadas por fontes locais a sua interpretação é difícil. Observações por satélite representam uma promessa futura para superar esses problemas por causa da sua cobertura global e homogênea e de seu grande volume de amostras, levando a menos problemas de representatividade. Observações de  $CH_4$  a partir do espaço têm sido feitas tanto em modo limb (tangente à atmosfera) como em nadir (transversal à atmosfera). Observações limb estão restritas principalmente à estratosfera, e, portanto, dificilmente contêm alguma informação sobre a distribuição de



fontes na superfície (RANDEL et al., 1998). Observações nadir de metano atmosférico foram realizadas pelo instrumento IMG (Interferometric Monitor for Greenhouse gases) a bordo do ADEOS (ADvanced Earth Observing Satellite) (CLERBAUX et al., 2003). Contudo, sendo um espectrômetro infravermelho, IMG é relativamente insensível à baixa troposfera, e sua relação entre anomalias nas concentrações medidas e suas fontes é indireta.

Recentemente, observações de metano a partir de satélite na faixa espectral do infravermelho próximo têm sido disponibilizadas pelo sensor SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) ((BOVENSMANN et al., 1999);(GOEDE et al., 2001)) a bordo do satélite ambiental ENVISAT da ESA (European Space Agency), lançado em março de 2002 (FRANKENBERG et al., 2005). No infravermelho próximo a maior parte da radiação que alcança o satélite foi espalhada pela superfície da Terra, resultando em medidas com uma sensibilidade quase constante da estratosfera até a camada limite planetária. Metano é um gás de efeito estufa, com uma vida média de 10 anos. Portanto, a variabilidade na sua concentração é pequena. Medidas da coluna atmosférica total devem ter uma alta precisão (da ordem de 1%) para que sejam úteis. É um desafio atingir tais níveis de precisão. Além disso, existem muitos fatores que podem levar a erros sistemáticos nas observações, tais como erros nos perfis verticais de pressão, temperatura e vapor d'água (FRANKENBERG et al., 2005), incertezas na presença e caracterização de nuvens e aerossóis, erros na calibração das medidas de radiância, e incertezas em dados de referência espectroscópicos ou parâmetros instrumentais, tais como a função fenda(ou função resposta) do detector (KLEIPOOL et al., 2005).



### 3 OS SENSORES EMBARCADOS

Nesta seção faremos uma breve descrição dos principais instrumentos de sondagem atmosférica em operação na atualidade e que estão envolvidos na medição e coleta de dados para recuperação de perfil de concentração de gases traço.

Atualmente, existem apenas alguns instrumentos a bordo de satélites em órbita que são capazes de medir o  $CO_2$  e o  $CH_4$  atmosférico. O HIRS (High Resolution Infrared Radiation Sounder) ((CHEDIN et al., 2002) e (CHEDIN et al., 2003)), o AIRS (Atmospheric InfraRed Sounder ((ENGELEN et al., 2004); (AUMANN H. H.; GREGORICH; GAISER, 2005) apud (AUMANN et al., 2003);(STROW et al., 2006);(MADDY et al., 2008))), e o IASI (The Infrared Atmospheric Sounding Interferometer)(CREVOISIER et al., 2009a) realizam medidas de  $CO_2$  na região do infravermelho termal (TIR) do espectro, ou seja, estes instrumentos não detectam radiação solar refletida, mas a radiação térmica emitida pela superfície e atmosfera.

Isso traz a vantagem de que as medições são possíveis não só durante o dia, mas também durante a noite. No entanto, a desvantagem de tais medidas é a sua falta de sensibilidade na baixa troposfera, onde o sinal mais forte devido a várias fontes e sumidouros podem ser esperados. Em contraste com isso, a sensibilidade de instrumentos que medem radiação solar refletida no infravermelho próximo (NIR) / região espectral do infravermelho de ondas curtas (SWIR) é muito mais constante (com altura) e tipicamente exibe os valores máximos próximo à superfície. Atualmente, SCIAMACHY a bordo do ENVISAT lançado em 2002 (BOVENSMANN et al., 1999), AIRS e TANSO (Thermal And Near infrared Sensor for carbon Observation) a bordo do GOSAT (Greenhouse gases Observing SATellite) lançado em 2009 (YOKOTA et al., 2004) são os únicos instrumentos em órbita medindo radiação no Infravermelho próximo (NIR) em bandas de absorção apropriadas ao redor de 0.76, 1.6, e 2.0  $\mu\text{m}$  com resolução espectral mínima requerida para recuperação de perfis de concentração de  $CO_2$  e  $CH_4$ . Outro satélite de observação de dióxido de carbono foi o OCO (Orbiting Carbon Observatory) (CRISP et al., 2004). OCO foi projetado para medir dentro da mesma região espectral. Infelizmente, o satélite foi perdido logo após a decolagem, em 24 de Fevereiro de 2009 (PALMER P. I.; RAYNER, 2009).

Contrariamente ao TANSO, SCIAMACHY não foi especialmente concebido para a recuperação de  $XCO_2$  (fração molar de  $CO_2$ ) com a precisão e acurácia necessárias para aumentar nosso conhecimento sobre as fontes e os sumidouros através de mod-

elagem inversa. Devido à menor resolução espacial e espectral do SCIAMACHY, a precisão e acurácia atingidas devem ser menores em relação a instrumentos como o TANSO. No entanto, dentro do período de tempo compreendido entre 2002 e 2009 SCIAMACHY foi o único instrumento capaz de medir  $CO_2$  e outros GEE do espaço com sensibilidade significativa também para a baixa troposfera. Portanto, o desenvolvimento de algoritmos derivando gases traço a partir do SCIAMACHY de forma tão precisa quanto possível, com estimativas de erro realista é fundamental para iniciar uma consistente série temporal de longo prazo de observações por satélite.

### 3.0.1 O Sensor Sciamachy

O sensor SCanning Imaging Absorption spectroMeter for Atmospheric CHartograPhY (SCIAMACHY) ((BURROWS, 1995);(BOVENSMANN et al., 1999);(SKUPIN et al., 2005)) é parte do equipamento de química atmosférica a bordo do satélite de monitoramento ambiental ENVISAT da Agência Espacial Européia (ESA), lançado em março de 2002. O ENVISAT voa em uma órbita quase-polar solar síncrona cruzando o equador às 10:00 a.m. (hora local). O SCIAMACHY (ver Figura 3.1(a)) é um espectrômetro multicanal que mede o espectro de radiação solar espalhada, refletida e transmitida e que cobre a faixa espectral de 240-2385 nm (ver Figura 3.1(b)) nos modos nadir, limb, ocultação solar e lunar com uma resolução que varia de 0.2-1.6 nm. O sensor consiste de oito canais espectrais principais (cada um equipado com um arranjo linear de detectores com 1024 pixels de detecção) e sete aparelhos de medida de Polarização espectral de banda larga (PMDs); maiores detalhes são fornecidos em (BOVENSMANN et al., 1999). A Tabela 3.1 lista as características dos 8 canais espectrais do SCIAMACHY.

Tabela 3.1 - Características dos canais espectrais do SCIAMACHY.

Canal $n^\circ$	Faixa de Comprimento de Onda (nm)	Resolução do pixel do detetor (nm)	Resolução espectral (nm)
1	240-314	0.12	0.24
2	309-405	0.13	0.26
3	394-620	0.22	0.44
4	604-805	0.24	0.48
5	785-1050	0.27	0.54
6	1000-1750	0.74	1.48
7	1940-2040	0.11	0.22
8	2265-2380	0.13	0.26

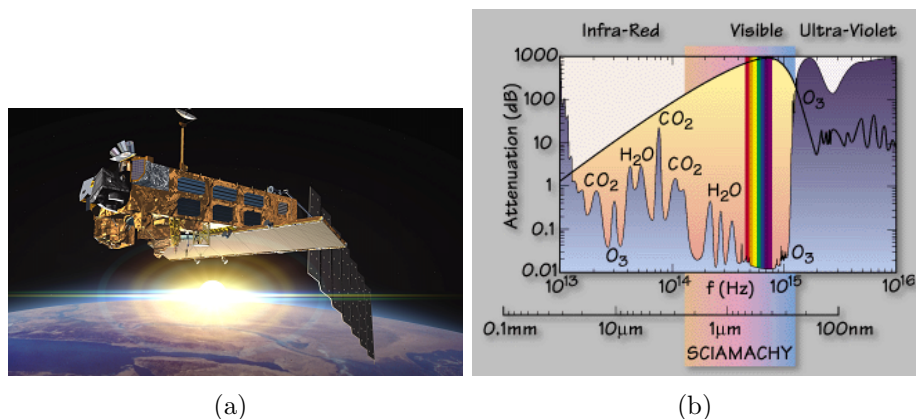


Figura 3.1 - a) Foto do SCIAMACHY abordo do ENVISAT; b) Esquema da faixa espectral utilizada pelo instrumento SCIAMACHY.  
 Fonte: <http://www.esa.int> (2002)

Para o SCIAMACHY a resolução espacial depende do intervalo espectral e posição orbital. Para o canal 8, a resolução espacial, ou seja, o tamanho da cobertura espacial de uma única medida em modo nadir, é de  $30 \times 120 km^2$  correspondendo a um tempo de integração de 0.5s. Para o canal 6 a resolução espacial é de  $30 \times 60 km^2$ , correspondendo a um tempo de integração de 0.25s. Para grandes ângulos polares zenitais o tamanho do pixel aumenta, podendo chegar a  $(30 \times 240 km^2)$ . Além disso, é um dos primeiros instrumentos que realizam observações nadir dentro da região espectral do infravermelho (NIR) (i.e., ao redor de  $2 \mu m$ ). Em contraste às regiões espectrais do Ultra-Violeta (UV) e visível, onde detectores de alta performance têm sido fabricados por um longo tempo, nenhum detector apropriado atuando na faixa do infravermelho-próximo estava disponível quando o SCIAMACHY foi concebido. Os detectores para o infravermelho próximo em InGaAs (material composto de arseneto de índio-gálio) foram um desenvolvimento especial para o SCIAMACHY. A performance óptica global do instrumento pode ser descrita como muito boa se comparado com outros instrumentos similares. Medidas de calibração em órbita, junto com informações de solo, resultam em radiâncias e irradiâncias precisamente calibradas, permitindo a recuperação de parâmetros geofísicos com a alta acuidade (LICHTENBERG et al., 2006).

Como resultado, este sensor tem um bom potencial para estimar o dióxido de carbono e outros GEE, usando o princípio da radiação refletida no infravermelho próximo.

Resultados preliminares obtidos por (BUCHWITZ et al., 2004) mostraram que o canal 7 (1940-2040 nm) e o canal 8 (2265-2385 nm) podem ser usados para a inferência da coluna de  $CO_2$  e  $CH_4$  na atmosfera, respectivamente. Em outros experimentos similares, Buchwitz mostra que também o canal 6 (nm 1000-1750) pode ser usado para inferência de perfis atmosféricos tanto de  $CO_2$  quanto de  $CH_4$  (BUCHWITZ et al., 2004). Resultados obtidos com dados reais, mostraram um erro de  $\pm 10$  ppmv ou 2.7% na concentração de  $CO_2$  na coluna atmosférica sobre o continente. Embora estes resultados ainda sejam pouco acurados, espera-se uma gradual redução dos erros, principalmente devido ao aprimoramento dos algoritmos de inversão e também com uma diminuição dos erros de calibração do satélite.

### 3.0.2 O Sensor AIRS

O AIRS foi o primeiro sondador infravermelho de alta resolução espectral desenvolvido pela NASA em conjunto com a NOAA. O sondador encontra-se em operação a bordo do satélite EOS-AQUA desde março de 2002, e foi desenvolvido com o objetivo de monitorar os gases do efeito estufa na atmosfera, detectar mudanças climáticas no globo terrestre, evidenciar o papel das nuvens e do vapor d'água atmosférico na manutenção da temperatura da superfície terrestre, estimar o balanço hidrológico e o balanço energético e aprimorar as previsões de tempo ((AUMANN et al., 2003); (ENGELEN R. J.; STEPHENS, 2004)). O satélite EOS-AQUA descreve uma órbita polar em torno do globo a uma altitude de quase 705 km e com periodicidade de cerca de 98 minutos. Este satélite realiza o imageamento de uma mesma região do globo a cada 12 horas.

O AIRS realiza uma varredura do globo terrestre com um ângulo de inclinação de  $\pm 49.5^\circ$  (com o nadir) e com 90 posições, resultando numa faixa imageada com largura de 2000 km na superfície terrestre e uma resolução horizontal variando de aproximadamente 2.3 km (para os canais do visível e do infravermelho próximo) a 13.5 km (para os canais do infravermelho termal). O espectrômetro do AIRS possui uma cobertura espectral em aproximadamente 2400 bandas nas faixas espectrais do infravermelho termal (3.7-15  $\mu m$ ), infravermelho próximo e do visível (0.4-1.0  $\mu m$ ), resultando numa resolução espectral de 1200 ( $\lambda/\Delta\lambda$ , onde  $\lambda$  é o comprimento de onda e  $\Delta\lambda$  é a largura da banda espectral). Estas faixas foram definidas especificamente para estimar perfis de temperatura, umidade e conteúdo integrado de diferentes gases na atmosfera. Gases atmosféricos tais como o dióxido de carbono, o vapor d'água, o ozônio, o metano e o óxido nitroso absorvem fortemente a radiação nos comprimentos

de onda situados na faixa espectral do infravermelho (LIU, 1980). Devido a essas regiões do espectro serem sensíveis às variações na concentração atmosférica destes gases, além da alta resolução espectral dos seus canais, as medidas de radiâncias emergentes do sistema terra-atmosfera feitas pelo AIRS possibilitam inferir a coluna integrada destes gases na atmosfera.

Uma variedade de algoritmos de inversão da equação da transferência radiativa (ETR) tem sido recentemente desenvolvida pela comunidade científica para quantificar os gases do efeito estufa na atmosfera a partir das radiâncias do AIRS. (ENGELN et al., 2001) elaboraram um algoritmo de inversão da ETR baseado na teoria da estimação ótima de Rodgers (RODGERS, 1976) para investigar a capacidade do AIRS em detectar o dióxido de carbono na coluna atmosférica. Neste estudo foram geradas medidas sintéticas do conteúdo integrado do dióxido de carbono na coluna atmosférica utilizando um modelo de circulação geral, simulando as radiâncias do AIRS por meio de um modelo de transferência radiativa na atmosfera (modelo direto). As simulações realizadas mediante o processo de inversão da ETR mostraram ser possível quantificar o dióxido de carbono com uma acurácia de 1 ppmv, desde que certos cuidados sejam tomados com respeito à existência de nuvens cirrus e de possíveis erros de precisão nos cálculos de transferência radiativa. Um outro algoritmo importante é o AIRSTeam Core Algorithm (AIRS-TCA), que foi elaborado por (CHAHINE et al., 2001) para inferir perfis verticais de temperatura, umidade e o conteúdo dos gases atmosféricos do efeito estufa, e baseia-se no nível de ruído das radiâncias de cada canal do AIRS e em técnicas iterativas de inversão.

(SUSSKIND J.; BARNET; BLAISDELL, 2003) aperfeiçoaram o AIRS-TCA de modo a eliminar o efeito da nebulosidade no processo de inversão da ETR. Mediante o uso de radiâncias sintéticas do AIRS simuladas por um modelo direto com a presença de nuvens, os autores observaram uma acurácia do algoritmo equivalente ao ruído do canal do AIRS para a inferência de perfis de temperatura e umidade atmosférica.

Devido ao grande número de canais do AIRS (alta resolução espectral), importantes estudos também foram realizados com o intuito de selecionar os canais mais adequados do sensor nas faixas espectrais do infravermelho próximo e do infravermelho termal para a inferência dos gases do efeito estufa na atmosfera ((RABIER et al., 2002);(CREVOISIER C.; CHEDIN; SCOTT, 2003)). (RABIER et al., 2002) observaram que a aplicação de métodos iterativos selecionando sequencialmente os canais do AIRS com maior conteúdo de informação é uma das mais convenientes para este

fim, e permite a utilização de uma mesma série de canais selecionados para inferir os gases do efeito estufa em diferentes condições atmosféricas. Com o mesmo enfoque, (CREVOISIER C.; CHEDIN; SCOTT, 2003) aplicaram o método Optimal Sensitivity Profile (OSP) a diferentes perfis atmosféricos para selecionar os principais canais do AIRS associados ao dióxido de carbono, metano e óxido nitroso atmosférico. Este método, que é baseado na sensibilidade dos canais do AIRS às variações da distribuição vertical destes gases, permitiu selecionar um total de 42 canais cobrindo toda a coluna atmosférica do dióxido de carbono.

### 3.0.3 O Observatório Orbital de Carbono (OCO)

A missão de lançamento do instrumento Observatório Orbital de Carbono (OCO) da NASA em 2009 falhou. Porém, está planejado o lançamento de um outro sensor, o OCO-2, cópia do primeiro, com lançamento previsto para começo de 2013. Esta missão que será exploratória, terá uma duração nominal de 2 anos. O OCO-2 fará medições espaciais de  $CO_2$  atmosférico com a precisão, resolução e cobertura necessária para caracterizar as fontes e os sumidouros de  $CO_2$  em escalas regional e quantificar sua variabilidade ao longo do ciclo sazonal. Este sensor irá voar a uma altitude de 705 km com uma órbita polar sol-síncrona, que irá proporcionar cobertura quase global do hemisfério iluminado com um ciclo de repetição de 16 dias (“ground track repeat cycle”). A Figura 3.2 ilustra o satélite e sua provável órbita operacional.

OCO leva um único instrumento que incorpora três espectômetros de grade de alta resolução que irão fazer as medições da luz solar refletida nas bandas de absorção na faixa do infravermelho próximo do oxigênio molecular ( $O_2$ ) e de  $CO_2$ . Estas medidas serão combinadas para fornecer estimativas espacialmente resolvido da fração molar de  $CO_2$ ,  $XCO_2$ . O instrumento coleta de 12 a 24 sondagens/segundo de  $XCO_2$  sobre a porção iluminada da órbita, gerando de 200 a 400 sondagens por grau de latitude, ou de 7 a 14 milhões de sondagens a cada 16 dias. Grossas nuvens e aerossóis reduzirão o número de sondagens disponíveis para recuperação de  $XCO_2$  em 80 – 90%, mas para os dados remanescentes é esperado a geração de estimativas de  $XCO_2$  com acurácia de aproximadamente 0.3% a 0.5 (1 a 2 ppm) em escala regional a cada mês. Para verificar a exatidão dos dados  $XCO_2$  obtidos por sensoriamento, o programa de validação OCO-2 vai utilizar a Rede de Observação de Coluna de Carbono (TCCON) para atrelar os dados de  $XCO_2$  espaciais com o padrão da Organização Meteorológica Mundial (WMO) para o  $CO_2$  atmosférico, que se baseia em observações *in situ* de  $CO_2$  a partir de medições por balão, torres e aeronaves.





Figura 3.2 - Visão do satélite OCO-2.

Fonte: <http://oco.jpl.nasa.gov/> (2010)

Os projetistas da missão OCO selecionaram três bandas específicas de comprimento de onda no infravermelho para ajudá-los a medir o  $CO_2$  atmosférico. O instrumento OCO mede intensidade ao longo destas três bandas na mesma posição sobre a superfície da Terra simultaneamente: uma banda fraca de  $CO_2$  centrada em torno de  $1.61\mu m$ , uma banda de oxigênio ( $O_2$ ) em  $0.76\mu m$  e uma banda de  $CO_2$  forte com centro em torno de  $2.06\mu m$ . Cada uma das três bandas de comprimento de onda selecionadas fornece uma contribuição específica para a precisão da medição. O objetivo científico principal da missão OCO é coletar dados globais de  $CO_2$  que ajudem a distinguir as fontes e os sumidouros. Contudo, a missão OCO não medirá diretamente fontes e os sumidouros de  $CO_2$ . Modelos computacionais de assimilação de dados que utilizam o  $XCO_2$  realizam a inferência da localização destas fontes e sumidouros (ver (PAES F. F., 2009) e (PAES et al., 2008)).

### 3.0.4 GOSAT

O satélite observador de gases de efeito estufa (GOSAT) é um satélite construído para monitorar o dióxido de carbono ( $CO_2$ ) globalmente, e objetiva contribuir para

os esforços internacionais para prevenir o aquecimento global, tais como o protocolo de Kioto. Foi desenvolvido pela agência espacial japonesa com o objetivo de observar as distribuições globais de  $CO_2$  e  $CH_4$  e suas mudanças do espaço com uma acurácia esperada da densidade da coluna do  $CO_2$  de 1-4 ppmv e para o  $CH_4$  de 6-18 ppbv. A Figura 3.3 ilustra o satélite e sua órbita operacional.

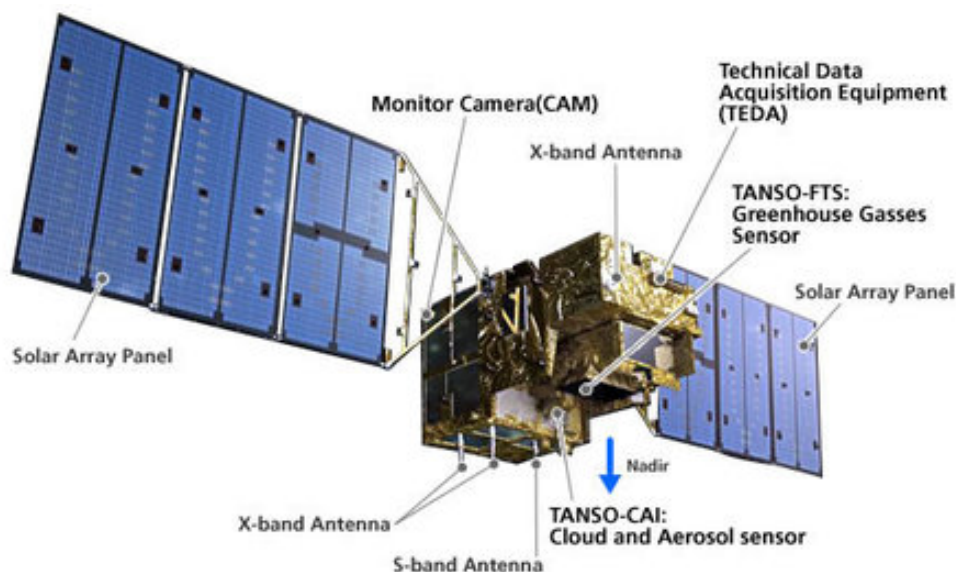


Figura 3.3 - Visão detalhada do satélite Ibuki (GOSAT).  
Fonte: <http://www.esa.int> (2010)

A estimaco da taxa de emisso/absoro em escala sub-continental foi tambm planejada. Espera-se melhorar a acurcia em relao aos sensores atuais de um fator de pelo menos 2. GOSAT inclui um sensor termal e de infravermelho-prximo (TANSO-FTS) para observao de carbono e um imageador de nuvens e de aerossis (TANSO-CAI) (SAKUMA *et al.*, 2009). O TANSO-FTS  um espectmetro baseado na transformada de Fourier e cobre uma ampla faixa do espectro atmosfrico de  $0.75$  a  $14.3\mu m$  a uma resoluo espectral de  $0.2cm^{-1}$ . J o sensor TANSO-CAI  usado para compensar os erros que so causados pela presena de nuvens e aerossis. Um mecanismo de dois eixos permite ao GOSAT observar qualquer ponto escolhido na superfcie da Terra. Com uma sequncia de observao em ziguezague, observaoes globais tornam-se possveis em uma malha de 180 km (WATANABE *et al.*, 2008).

Os alvos da misso so a observao da coluna de  $CO_2$  atmosfrico com 1% (4 ppmv)

de acurácia relativa em uma resolução espacial subcontinental e a redução dos erros na identificação de fontes e sumidouros de GEE em escala subcontinental de 50% aliando os dados obtidos pelo satélite GOSAT e os dados coletados por instrumentos *in situ*. A banda de  $0.76\mu m$  é usada para derivar a densidade de coluna de  $O_2$  e determinar o comprimento exato do caminho óptico. As bandas de  $1.6\mu m$  e  $2.0\mu m$  são utilizadas para observar a densidade de coluna de  $CO_2$ . A banda de  $1.6\mu m$  é usada para observar a densidade de coluna de  $CH_4$ . A banda de  $5.5 - 14\mu m$ , por sua vez, é usada para recuperar o  $CO_2$ ,  $CH_4$ , vapor de água, e temperatura atmosférica. Perfis verticais de  $CO_2$  e  $CH_4$  também estão disponíveis com esta banda.



## 4 TEORIA DA SONDAGEM REMOTA DE GASES DE EFEITO ESTUFA

Neste capítulo introduzimos o conceito de transferência radiativa na atmosfera. Apresentamos também, de forma simplificada, uma aproximação matemática para o conteúdo de radiação que chega ao topo da atmosfera e sensibiliza os sensores a bordo do satélite. Em seguida é apresentada uma breve descrição do modelo direto empregado para simulação dos dados de radiância do sensor SCIAMACHY. Ele serviu de base para a geração de radiâncias sintéticas e estudos de sensibilidade das bandas espectrais correspondentes aos canais do sensor envolvidos na recuperação dos perfis de concentração de dióxido de carbono e metano. Por último, explicamos o conceito de problema direto (PD) e problema inverso (PI). A explanação desses dois assuntos é relevante, visto que, na metodologia são abordados os modelos numéricos (modelo direto e inverso) utilizados para solução desses problemas.

### 4.1 A Equação de Transferência Radiativa (ETR)

A radiação eletromagnética provém de uma variação temporal dos campos elétricos e magnéticos descrita pelas equações de Maxwell. Este processo, não necessita de um meio material para se efetuar, e transporta energia e momento. A velocidade de propagação das ondas eletromagnéticas no vácuo é uma constante universal ( $3.0 \times 10^8 m/s$ ). Devido à sua natureza ondulatória, a radiação é caracterizada por uma amplitude e por ter uma frequência característica cuja composição dos diferentes comprimentos de onda forma o espectro eletromagnético. A base física que possibilita a inferência de parâmetros atmosféricos a partir de medidas de radiação efetuadas remotamente por satélites está associada às interações entre radiação e matéria, e pode ser descrita pela equação geral da transferência radiativa. Esta equação descreve os processos físicos (absorção, emissão e espalhamento) que um feixe de radiação sofre ao atravessar um meio material. Contudo, quando trabalhamos no domínio dos processos atmosféricos em sensoriamento remoto, normalmente se faz uso de algumas considerações físicas que simplificam o seu tratamento matemático. A equação de Schwartzchild é um caso particular da ETR, e descreve as interações entre radiação e a matéria em um meio não espalhante e em equilíbrio termodinâmico local, comportando-se como um corpo negro. Para uma atmosfera plano-paralela, onde se deseja determinar a radiação normal ao plano de estratificação da atmosfera, e supondo que a radiação atmosférica não seja função do ângulo azimutal, a equação de Schwartzchild pode ser escrita como:

$$\mu \frac{dI_\lambda(\tau, \mu)}{d\tau} = I_\lambda(z, \mu) - B_\lambda(T) \quad (4.1)$$

onde  $I_\lambda$  é a intensidade de radiação monocromática,  $B_\lambda(T)$  é a função de Planck (ver Equação 4.2),  $\lambda$  é o comprimento de onda (nm),  $z$  é a medida de distância na direção normal ao plano de estratificação,  $\mu = \cos\theta$  (onde  $\theta$  é o ângulo zenital) e  $\tau$  é a profundidade ótica normal ao plano de estratificação. Podemos definir profundidade ótica como o parâmetro que mede o quão opaco um meio se apresenta em relação à radiação que o atravessa. É medido ao longo de uma vertical conhecida como caminho ótico, que encontra-se na direção normal à superfície a ser medida (LIOU, 1980).

A partir da Equação 4.1 podemos obter a radiação que chega ao topo da atmosfera e sensibiliza os sensores a bordo do satélite. Se considerarmos o satélite posicionado próximo ao ângulo zenital ( $\mu \approx 1$ ), temos que:

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5 [\exp(hc/\lambda kT) - 1]} [Wm^{-2}sr^{-1}\mu m^{-1}] \quad (4.2)$$

$$I(p_0) = I_\lambda(p_s)\tau_\lambda(p_s) + \int_{p_s}^{p_0} B_\lambda(T) \mathfrak{S}_\lambda(p) \frac{\partial \mathfrak{S}_\lambda(p)}{\partial p} dp \quad (4.3)$$

A equação 4.3 está em coordenadas de pressão (usando aproximação hidrostática), onde  $p$  é a pressão atmosférica;  $p_0$  a pressão no topo da atmosfera;  $p_s$  a pressão na superfície;  $T$  a temperatura e  $\mathfrak{S}_\lambda$  a transmitância entre o ponto referenciado e o topo da atmosfera  $p_0$ . O primeiro termo do lado direito da equação representa a contribuição da radiação provinda da superfície terrestre. Geralmente a radiância emitida pela superfície pode ser aproximada por  $\varepsilon_\lambda B_\lambda(T)$ , onde  $\varepsilon_\lambda$  é a emissividade da superfície. O segundo termo representa a contribuição atmosférica para a radiância medida pelo sensor para cada nível diferencial de pressão  $dp$ , onde o termo  $\frac{\partial \mathfrak{S}_\lambda(p)}{\partial p}$  é conhecido como função peso. A transmitância monocromática é dada em coordenadas de pressão por

$$\mathfrak{S}_\lambda(p) = \exp \left[ -\frac{1}{g} \int_{p_0}^p \kappa_\lambda(p') q(p') dp' \right] \quad (4.4)$$

onde  $q$  é a razão de mistura do gás em consideração;  $g$  a aceleração da gravidade local e  $p_0$  a pressão no topo da atmosfera. O sensor a bordo do satélite tem uma resolução finita e pode distinguir somente uma banda espectral finita  $\phi(\bar{\lambda}, \lambda)$  onde  $\phi$  e  $\bar{\lambda}$  denotam a função resposta do sensor e o comprimento de onda médio da banda, respectivamente. Na prática, quando nos referimos ao canal do satélite com certo comprimento de onda  $\lambda$ , na verdade estamos nos referindo ao comprimento de onda médio deste canal e, a radiância medida neste canal é o resultado da contribuição das radiâncias monocromática para as frequências incluídas dentro do intervalo de resolução do sensor ponderadas pela função resposta deste sensor. A equação da transferência radiativa é utilizada para estimar os perfis de concentração de gases traço a partir de bandas de absorção do gás. Neste caso o perfil vertical de temperatura (e conseqüentemente  $B_\lambda(T)$ ) é suposto ser conhecido (determinado por métodos de inversão), e se deseja determinar  $q(p)$  (razão de mistura do gás). Convém ressaltar que  $q(p)$  se encontra dentro de uma integral, que faz parte de uma função exponencial que determina a transmitância, dificultando, assim, o processo de inversão. Além disso, existem os erros associados ao perfil de temperatura, também inferido pela inversão da ETR, e a influência de outros gases na faixa de absorção medida pelo sensor. Tudo isto diminui consideravelmente a qualidade dos perfis verticais de umidade inferidos remotamente via satélite.

Ressaltamos que o modelo matemático para a ETR utilizado neste estudo deve levar em consideração os efeitos do espalhamento anisotrópico de Rayleigh (molecular) e Mie (por partículas) e, ao mesmo tempo, ignorar o termo correspondente à emissão de ondas longas no infravermelho termal ( $\sim 4\mu\text{m}$ ) pela superfície terrestre pois nossos experimentos irão se concentrar na faixa espectral do infravermelho próximo de 1000-1750 nm (canal 6 do sensor SCIAMACHY). Se transformamos as coordenadas de pressão em coordenadas de altitude e levamos em consideração o termo de espalhamento a expressão utilizada para a ETR, neste caso específico, teria formulação idêntica à expressa em (ROZANOV et al., 1997) e (LENOBLE, 1985) e poderia

ser escrita da seguinte forma:

$$\mu \frac{dI(z, \mu, \varphi)}{dz} = -c(z)I(z, \mu, \varphi) + \frac{b(z)}{4\pi} \int_0^{2\pi} \int_{-1}^1 p(z, \mu, \mu', \varphi, \varphi') I(z, \mu', \varphi') d\mu' d\varphi',$$

com

$$\begin{aligned} a(z) &= a_M(z) + \sum_{i=1}^k \alpha_i \eta_i(z) \\ b(z) &= b_M(z) + b_R(z) \\ c(z) &= a(z) + b(z) \end{aligned}$$

onde  $\mu$  é o coseno do ângulo zenital  $\theta$ . As condições de contorno no topo da atmosfera e na superfície terrestre levam a

$$\begin{aligned} I^-(z_0, \mu, \varphi) &= \delta(\mu - \mu_0, \varphi - \varphi_0) \pi F \\ I^+(0, \mu, \varphi) &= \frac{A}{\pi} \int_0^{2\pi} \int_0^1 I^-(0, \mu', \varphi') \mu' d\mu' d\varphi' \end{aligned}$$

onde  $I^-(z_0, \mu, \varphi)$  é a radiância em  $z_0$  cujo fluxo aponta de fora para dentro da terra e  $I^+(0, \mu, \varphi)$  é a radiância cujo fluxo parte da superfície terrestre. Por questão de simplicidade a dependência do comprimento de onda  $\lambda$  é omitida e um albedo de superfície Lambertiano é assumido. Os outros símbolos usados na equação junto com seus significados são exibidos na Tabela 4.1 abaixo:

Aqui, utilizamos um modelo direto que implementa a ETR e simula observações do sensor SCIAMACHY para estimar os perfis de concentração de gases traço a partir de bandas de absorção apropriadas dentro da janela de transmissão no infravermelho próximo. Neste caso, o perfil vertical de temperatura é determinado por um procedimento de inversão independente e precisamos apenas recuperar a concentração do gás representada pela sua taxa de mistura que, na formulação apresentada em (ROZANOV et al., 1997), é parte da equação. O modelo direto utilizado nesse trabalho



Tabela 4.1 - Símbolos usados na ETR

Símbolo	Significado
$I$	radiância total (i.e., campo de radiação)
$z$	altitude;
$z_0$	altitude no topo da atmosfera
$\mu, \mu'$	cosenos dos ângulos zenitais $\theta, \theta'$ ;
$\mu_0$	coseno do ângulo solar zenital $\theta_0$ ;
$\varphi, \varphi'$	ângulos azimutais em relação à linha de visão projeção sobre a superfície da terra
$\varphi_0$	ângulo azimutal do Sol em relação à linha de visão projeção sobre a superfície da terra
$\kappa$	número de gases traço
$c$	coeficiente de extinção total
$a$	coeficiente de absorção total (soma de gases traço e partículas coeficientes de absorção)
$b_R$	coeficiente de espalhamento de Rayleigh
$b_M$	coeficiente de espalhamento por partículas
$b$	coeficiente de espalhamento total (soma do espalhamento de Rayleigh e espalhamento por partículas)
$p_M$	função de fase de espalhamento de Rayleigh
$p$	função de fase de espalhamento total = $(b_R p_R + b_M p_M)/(b_R + b_M)$
$\alpha_i$	seção cruzada de absorção de gases traço para todos gases relevantes
$n_i$	concentração de gases traço
$a_M$	coeficiente de absorção por partículas
$\pi F \mu_0$	fluxo solar no topo da atmosfera
$A$	reflectância espectral (Lambertiana) da superfície da terra

é detalhado na seção 4.2 seguinte onde descrevemos as suas principais características e as adaptações efetuadas para este trabalho.

## 4.2 Modelo Direto

Para a maioria dos algoritmos de recuperação ser bem sucedida deve haver um modelo direto que calcula o espectro de saída e, conseqüentemente, integra a equação de transferência radiativa para um determinado estado atmosférico. Muito trabalho tem sido feito na tentativa de construir modelos diretos precisos e computacionalmente eficientes que são capazes de descrever uma ampla faixa de estados atmosféricos que o satélite irá observar. O modelo direto integra a equação de transferência radiativa, desde a superfície até a altura do satélite para uma série de comprimentos de onda que compõem o vetor de medição. Esta é uma tarefa computacionalmente intensiva e intrincada para todos os cenários, exceto os mais triviais, e foi detalhada de forma

exaustiva e pioneira por Chandrasekhar, (1950) (CHANDRASEKHAR, 1950).

#### 4.2.1 SCIATRAN

O SCIATRAN, escrito em FORTRAN 95, foi desenvolvido para realizar a modelagem de transferência radiativa em qualquer geometria de observação apropriada para medidas da radiação solar espalhada na atmosfera da terra, e tem sido usado como modelo direto na recuperação de constituintes atmosféricos. Ele resolve a equação de transferência radiativa utilizando o método de diferenças finitas para uma atmosfera verticalmente não-homogênea plano-paralela levando em consideração o espalhamento múltiplo. A partir da versão 2.0 permite adicionalmente cálculos de transferência radiativa em uma atmosfera esférica. A faixa de comprimento de onda coberta pelo modelo de transferência radiativa é 175-2380 nm, incluindo as bandas de absorção de oxigênio de Schuman-Runge e Herzberg (ROZANOV et al., 2002). Por “default”, é empregado no SCIATRAN um banco de dados climatológicos obtido utilizando um modelo dinâmico-químico desenvolvido por Brühl e Crutzen (BRUHL C.; CRUTZEN, 1993). Este banco de dados contém distribuição vertical mensal e latitudinal da taxa de mistura do volume de gases traço, pressão e temperatura entre 0 e 60 km.

A Figura 4.1 mostra um espectro de radiância normalizado (sun-normalized) associado a uma atmosfera tropical. Ele cobre toda a faixa espectral do instrumento SCIAMACHY (modo nadir) e foi calculado utilizando o modo k-correlacionado do software SCIATRAN. No painel superior, observa-se um forte declínio da intensidade abaixo de 300 nm devido à absorção por ozônio. As características remanescentes são devido à absorção por vapor d’água e oxigênio. No painel intermediário, observa-se uma forte banda de absorção de  $CO_2$  com sobreposição da banda de  $H_2O$ . No painel inferior, as estruturas claramente visíveis são devido à absorção por  $H_2O$  e  $CH_4$  com fraca sobreposição de  $N_2O$  e absorção por CO na primeira e segunda parte do canal 8, respectivamente. A título de comparação é exibido na Figura 4.2 seguinte o espectro de radiância associado a um perfil subártico no inverno. Embora a forma do envelope de sinal calculado pelo SCIATRAN seja o mesmo pode-se observar diferenças principalmente com relação à amplitude da intensidade do sinal. Outra vantagem do SCIATRAN é que ele possibilita também determinar as derivadas das radiâncias com respeito a inúmeros parâmetros atmosféricos, que são necessárias no processo de inversão da ETR (BUCHWITZ M.; ROZANOV; BURROWS, 2000) e na análise de sensibilidade dos canais do SCIAMACHY com relação à variação dos parâmetros

atmosféricos para seleção de canais.

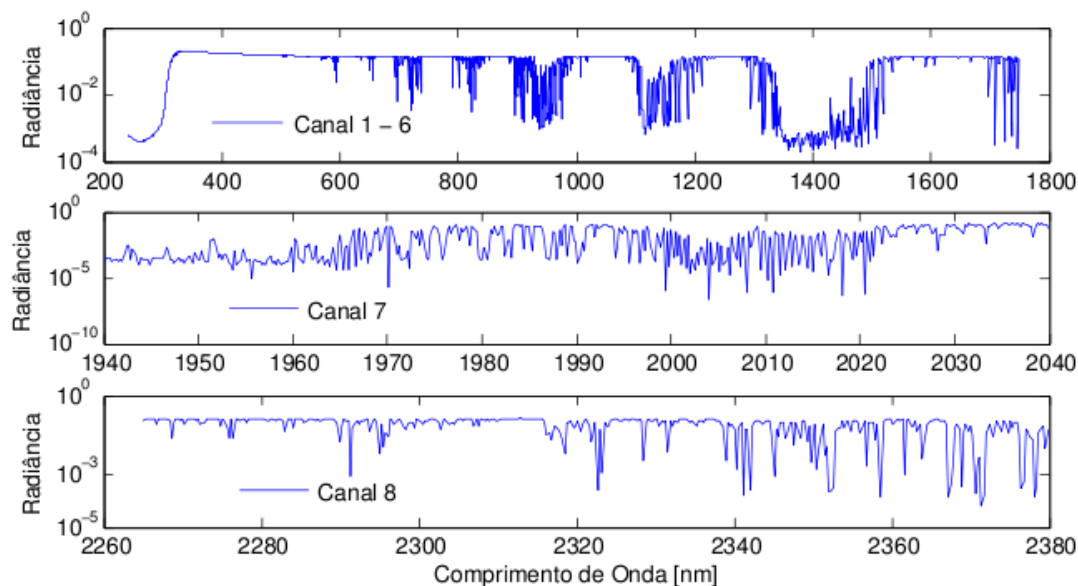


Figura 4.1 - Medida simulada de radiância (normalizada pelo espectro solar) do SCIAMACHY em modo nadir, associada a uma atmosfera tropical, calculado com o SCIATRAN (modo k-correlacionado (c-k); depois de efetuada convolução com função fenda Gaussiana de 0.3 nm (FWHM). Cenário: atmosfera U.S. Standard, espalhamento múltiplo, ângulo solar zenital  $60^\circ$ , albedo 0.1, ausência de nuvens.

### 4.3 Modelo Inverso

Para recuperar variáveis de perfis atmosféricos a partir de medidas radiativas no topo da atmosfera precisamos inverter a equação de transferência radiativa (ETR). A inversão analítica desta equação não é possível; somente uma abordagem de inferência pode ser utilizada nesses casos (TWOMEY, 1977). Diferentemente do problema direto, que pode ser estimado com alta precisão por um algoritmo físico, o problema inverso necessita de um método de resolução baseado na representação estatística da equação inversa (desconhecida). Existem duas abordagens gerais: utilizando um esquema de inversão para cada observação (abordagem de inversão local) ou modelando a inversa da ETR de uma só vez (abordagem de inversão global). A inversão local geralmente requer uma boa estimativa inicial para restringir a solução e um modelo direto rápido e preciso (RODGERS, 1976). Os modelos de inversão global não exigem uma estimativa inicial (AIRES et al., 2001), e nenhum modelo direto é

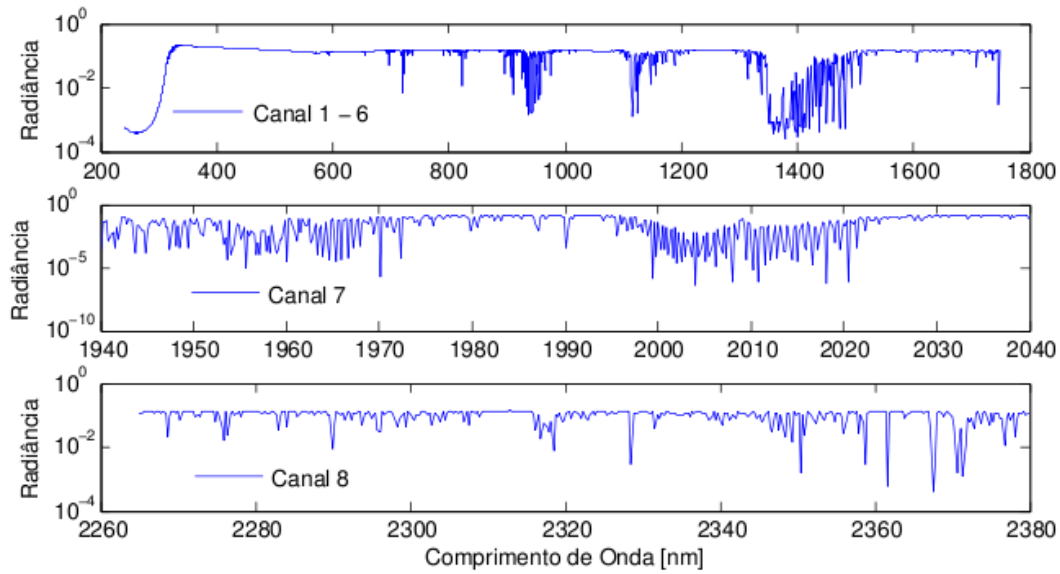


Figura 4.2 - Medida simulada de radiância (normalizada pelo espectro solar) do SCIAMACHY em modo nadir, associada a uma atmosfera subártica, calculada com o SCIATRAN (modo k-correlacionado (c-k); depois de efetuada convolução com função fenda Gaussiana de 0.3 nm (FWHM). Cenário: atmosfera U.S. Standard, espalhamento múltiplo, ângulo solar zenital 60°, albedo 0.1, ausência de nuvens.

necessário durante o uso operacional. Embora a inversão global não tenha essas duas limitações, ela constitui um problema muito mais ambicioso do ponto de vista matemático-computacional.

Neste trabalho utilizamos uma abordagem de inversão global, a qual permitiu a introdução de informação *a priori* no método de inversão. O conhecimento *a priori* é qualquer informação sobre a solução do problema que complementa a informação contida nas observações. Em técnicas estatísticas usuais, tais como regressão, superar o conceito de modelagem do tipo “caixa-preta” (que não leva em consideração prerrogativas sobre a física do problema) tende a melhorar os resultados finais. Neste trabalho em particular, aproximamos a relação inversa dada pela ETR com um modelo baseado em Redes Neurais Artificiais (RNA) para uma inversão direta dos perfis atmosféricos de gases traço.

#### 4.3.1 Conceito de Problema Inverso

O conceito de um problema direto pode ser entendido como um processo onde a partir de determinadas condições iniciais um sistema sofre transformações e desta

forma evolui para novas configurações. Os problemas diretos em geral pertencem à classe de problemas bem-postos, podendo ser descritos matematicamente, para um caso linear pela relação:

$$u = Az, \quad (4.5)$$

sendo  $z \in F$  e  $u \in U$  onde  $(F, U)$  define um par de espaços métricos,  $u$  corresponde ao vetor de dados observados e  $A$  representa um operador diferencial, algébrico ou integral, com um domínio definido não nulo  $D_A$  e operando em  $U$ . Uma formulação inversa para este problema implica em proceder a uma estimativa das causas do problema (condições de contorno e parâmetros iniciais) a partir dos seus efeitos, ou seja, resolver a Equação 4.5 para  $z$ , expressa por:

$$z = A^{-1}u, \quad (4.6)$$

onde  $A^{-1}$  representa o operador inverso.

A Figura 4.3 abaixo mostra de maneira pictórica a relação entre problema direto e inverso. Causas, num modelo matemático, são as condições iniciais e de contorno, termo de fontes/sumidouro e propriedades do sistema (material). Efeitos são as propriedades calculadas a partir de um modelo direto, como o campo de temperatura, radiância, concentração, corrente elétrica, etc.

Os problemas inversos frequentemente incluem-se na classe de problemas do tipo mal-postos, caracterizados por serem matematicamente difíceis de resolver.

### 4.3.2 O conceito de problemas Mal-Postos

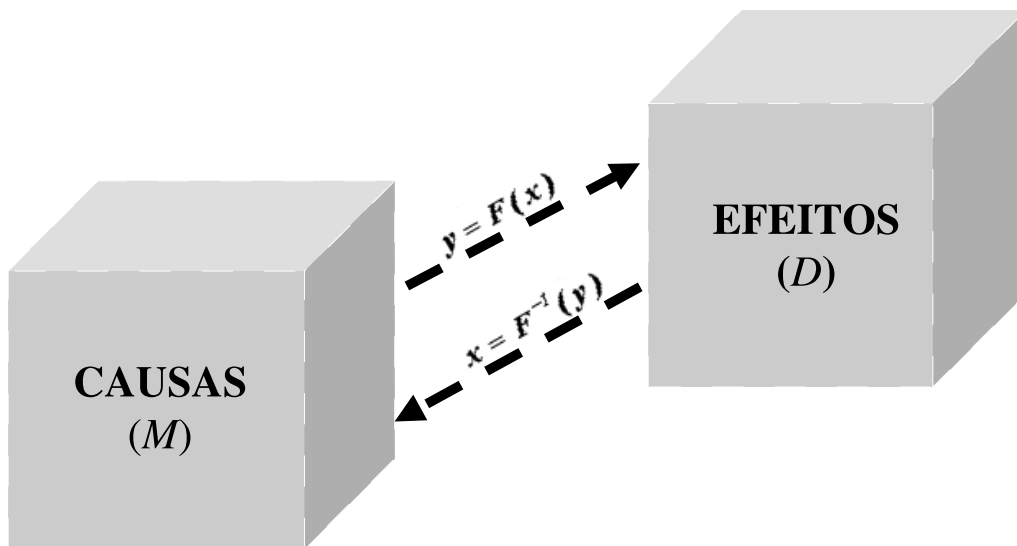
Dada a formulação descrita na Equação 4.6, onde se deseja determinar uma solução  $z$  pertencente ao espaço  $F$ , a partir dos dados observados  $u$  contido no espaço  $U$ , um problema é considerado bem posto em relação ao par de espaços métricos  $(F, U)$  se as seguintes condições forem satisfeitas:

- a) Existência:  $\forall u \in U \Rightarrow \exists z \in D_A$  tal que  $Az = u$ ,
- b) Unicidade :  $\forall z_1, z_2 \in D_A$  tal que  $Az_1 = Az_2 \Rightarrow z_1 = z_2$ ,
- c) Estabilidade  $\forall z_1, z_2 \in D_A$  tal que  $Az_1 = u_1, Az_2 = u_2 \Rightarrow z_1 \rightarrow z_2$ , em  $F$ ,

com  $u_1 \rightarrow u_2$  em  $U$ , ou seja: variações infinitesimais em  $u$  correspondem a variações infinitesimais em  $z$ .

Problemas que não satisfazem estas condições, são considerados problemas do tipo mal-postos. Lembrando que esta afirmação só é válida para um determinado espaço métrico  $(F, U)$ , desde que o mesmo problema pode se bem posto com relação a outros espaços métricos (TIKHONOV A. N.; ARSENIN, 1977). As duas primeiras condições caracterizam a determinação matemática do problema, já a terceira condição está relacionada a sua determinação física e com a possibilidade ou não de aplicação de métodos numéricos para resolvê-lo a partir de um certo número de dados iniciais aproximados.

Na prática, é comum contar-se com poucos dados, fornecidos através de medições realizadas por aparelhos resultando em uma malha discreta de pontos. Isto torna o sistema impossível de se resolver para um espaço contínuo, tornando-se necessário a utilização de uma formulação discreta para o problema. Além disso geralmente busca-se estimar um grande número de parâmetros a partir de poucos resultados, o que gera sistemas indeterminados. Assim informação complementar deve ser de



<b><math>M</math></b>	espaço de parâmetros ou modelos	<b><math>x</math></b>	modelo inverso
<b><math>D</math></b>	espaço de dados ou observações	<b><math>y</math></b>	modelo direto

Figura 4.3 - Representação esquemática de problemas direto e inverso.

alguma forma fornecida ao algoritmo de inversão a fim de que este se torne viável e o sistema possa ser resolvido. Apesar da dificuldade inerente ao problema, existe um esforço muito grande em se buscar um maior conhecimento acerca desta classe de problemas tendo em vista a relevância dos mesmos para diversas áreas do conhecimento. Twomey (TWOMEY, 1977) apresenta uma lista de aplicações de problemas inversos a situações reais, como nas áreas de transferência de calor, física médica, engenharia, medicina, reconstrução de imagens, meteorologia, geofísica, entre outras. Existem diferentes métodos para a solução de Problemas Inversos: Inversão Direta; Decomposição em Valores Singulares; Mínimos Quadrados; Métodos de Regularização; Métodos Variacionais; Redes Neurais; Algoritmos Genéticos; Colônias de Formigas, entre outros ((MUNIZ W. B.; CAMPOS VELHO; RAMOS, 1999); (MUNIZ W. B.; RAMOS; VELHO, 2000);(RAMOS et al., 1999); (VELHO, 2001)).

Para o problema caracterizado pela determinação da concentração de gases de efeito estufa a partir de radiâncias de satélite, o modelo direto consiste em determinar o conjunto finito de intensidades de radiância  $\{I_i\}_{i=1}^N$  que atingem o satélite a partir dos parâmetros atmosféricos  $\{C_j\}_{j=1}^M$  que representam a concentração do gás traço nos níveis de pressão  $j$ . E o modelo inverso (problema inverso) em questão envolve a determinação dos parâmetros  $\{C_j\}_{j=1}^M$  a partir das radiâncias  $\{I_i\}_{i=1}^N$ .

Erros sistemáticos de recuperação podem ser causados pela variabilidade não contabilizada de aerossóis e nuvens residuais que levam à erros relacionados ao espalhamento do sinal radiante. Nuvens podem atuar como um escudo para as espécies absorventes localizadas abaixo da cobertura de nuvens. Neste caso, é necessário introduzir algoritmos para identificar os pixels que são nuvens e eliminá-los do cálculo. Além de ruído do instrumento um grande número de outras fontes de erro contribuem para o erro global tais como erros decorrentes da variabilidade de perfis de temperatura e vapor d'água, aerossóis, nuvens cirrus sub-visual não detectadas, e os efeitos do albedo. Os erros resultantes são parcialmente sistemáticos e parcialmente aleatórios, ou seja, eles influenciam tanto a precisão (bias) quanto a acurácia (dispersão) das colunas recuperadas. Neste trabalho, a recuperação de abundância de  $CO_2$  atmosférico é feita a partir de observações SCIAMACHY sob condições de céu claro e ar limpo. Os pixels produzidos pelo sensor mostrando nuvens são identificados e descartados antes de nosso estudo.

### 4.3.3 Os principais algoritmos de inversão baseados no sensor SCIAMACHY

#### 4.3.3.1 WFM-DOAS

O WFM-DOAS (em fase operacional) é o principal algoritmo de inversão da ETR que vem sendo aplicado para inferir o conteúdo integrado de diferentes gases atmosféricos a partir das medidas de radiância feitas pelo SCIAMACHY na faixa espectral do infravermelho próximo [(BUCHWITZ M.; ROZANOV; BURROWS, 2000);(BUCHWITZ et al., 2004);(BUCHWITZ et al., 2006);(SCHNEISING et al., 2008a);(SCHNEISING et al., 2008b)]. Foi desenvolvido principalmente para a recuperação da coluna total de CO, CH<sub>4</sub>, CO<sub>2</sub>, H<sub>2</sub>O e N<sub>2</sub>O, contudo, não é limitado a esta aplicação e também tem sido empregado com sucesso na recuperação da coluna total de Ozônio utilizando dados do sensor GOME (COLDEWEY-EGBERS et al., 2004) e na recuperação da coluna total de vapor d'água utilizando o espectro dos sensores GOME e SCIAMACHY ao redor dos 700 nm no modo nadir (NOEL S.; BUCHWITZ; BURROWS, 2004). O WFM-DOAS é baseado na minimização de uma função objetivo expressa pela norma quadrática do logaritmo da saída do modelo de transferência radiativa linearizado  $I_{mod}$  mais um polinômio de baixa ordem  $P$ , menos o logaritmo da radiância nadir medida, isto é, radiância  $I_{obs}$  (normalizada pelo espectro solar). A equação de mínimos quadrados WFM-DOAS pode ser escrita como segue (parâmetros de ajuste estão sublinhados):

$$\| \ln I_i^{obs} - \ln I_i^{mod}(\hat{\underline{V}}) \|^2 \equiv \| RES_i \|^2 \rightarrow min. \quad (4.7)$$

onde o modelo de transferência radiativa linearizado é dado por

$$\ln I_i^{i mod}(\hat{\underline{V}}) = \ln I_i^{i mod}(\bar{\underline{V}}) + \sum_{j=1}^J \frac{\partial \ln I_i^{i mod}}{\partial V_j} \times (\hat{V}_j - \bar{V}_j) + P_i(\underline{a}_m) \quad (4.8)$$

O índice  $i$  refere-se ao comprimento de onda central  $\lambda_i$  do píxel número  $i$  do detector. Os componentes do vetor  $V$ , denotado  $V_j$ , são as colunas verticais de todos os gases traço que tem linhas de absorção dentro da janela espectral de ajuste selecionada. Os parâmetros de ajuste são as colunas verticais dos gases traço desejados  $\hat{V}_j$  e os coeficientes polinomiais  $a_m$ . Um parâmetro de ajuste adicional também utilizado (mas omitido nas equações 4.7 e 4.8) é o deslocamento (em  $K$ ) de um perfil de tem-



peratura pré-selecionado. Este parâmetro de ajuste foi adicionado afim de levar em consideração a dependência das seções cruzadas de absorção do gás traço em relação à temperatura. Os valores desses parâmetros de ajuste são determinados através da minimização (no sentido de mínimos quadrados) da diferença entre observação ( $\ln I_i^{obs}$ ) e modelo WFM-DOAS ( $\ln I_i^{mod}$ ), isto é, ajuste residual  $RES_i$ , para todos pontos espectrais  $\lambda_i$  simultaneamente. A derivada, ou função peso, com relação à coluna vertical refere-se à uma mudança na radiância no topo da atmosfera causada pela variação de um perfil vertical de um gás traço pré-selecionado. O espectro de referência WFM-DOAS é o logaritmo das radiâncias (normalizadas pelo espectro solar) e suas derivadas. Eles são computados com um modelo de transferência radiativa (BUCHWITZ M.; ROZANOV; BURROWS, 2000) para colunas (climatológicas) médias  $\bar{V}$ . O espalhamento múltiplo é totalmente levado em consideração. O problema de mínimos quadrados (eqs. 4.7 e 4.8) pode também ser expresso na seguinte notação matricial/vetorial: Minimize  $\|y - Ax\|^2$  com relação à  $x$ . A solução é  $\hat{x} = C_x A^t y$  onde  $C_x \equiv (A^t A)^{-1}$  é a matriz de covariância da solução  $\hat{x}$ . Os erros das colunas estimadas pelo WFM-DOAS são estimados como segue (PRESS et al., 1992):

$$\sigma_{\hat{V}_j} = \sqrt{(C_x)_{jj} \times \sum_i RES_i^2 / (m - n)} \quad (4.9)$$

onde  $(C_x)$  é o  $j$ -ésimo elemento da diagonal da matriz de covariância,  $m$  é o número de pontos espectrais na janela de ajuste e  $n$  é número de parâmetros de ajuste linear ( $RES_i$  é o resíduo de ajuste espectral, ver eq. 4.7).

Qualquer algoritmo que objetive a recuperação de informações atmosféricas a partir de dados de satélite não somente tem de ser preciso, mas também tem de ser suficientemente rápido a fim de processar enormes quantidades de dados (o espectro do SCIAMACHY compreende cerca de 8000 pontos de dados), que são registrados várias vezes por segundo a uma taxa de transferência de dados de 400 kbit por segundo).

Acurácia e velocidade são requisitos conflitantes e um compromisso apropriado tem de ser encontrado. O algoritmo WFM-DOAS em sua implementação corrente emprega uma abordagem do tipo *look-up table* e é muito rápido. O processamento de uma órbita inteira de um espectro SCIAMACHY pré-processado requer cerca de 2 minutos por janela espectral de ajuste em um computador PC padrão, o que é mais

do que uma ordem de magnitude mais rápida do que em tempo real (uma órbita demora cerca de 100 min).

#### **4.3.3.2 AIRS-Team Retrieval For Core Products and Geophysical Parameters**

Um outro algoritmo importante é o AIRS-Team Core Algorithm (AIRS-TCA), que foi elaborado por (CHAHINE et al., 2001) para inferir perfis verticais de temperatura, umidade e o conteúdo dos gases atmosféricos do efeito estufa a partir das radiâncias do sensor AIRS. Este algoritmo é baseado numa matriz de covariância que leva em consideração o nível de ruído das radiâncias de cada canal do AIRS e em técnicas iterativas que visam a estimação de diferentes parâmetros atmosféricos a partir da minimização das diferenças entre a radiância medida pelo AIRS e a radiância simulada por um modelo direto. O AIRS-TCA utiliza como modelo direto o AIRS-RTA. O AIRS-RTA é um modelo híbrido que parametriza as transmitâncias atmosféricas de uma grande variedade de gases e se utiliza de uma espectroscopia específica para efetuar rapidamente os cálculos de transferência radiativa na faixa espectral do infravermelho sob condições de céu claro (STROW et al., 2002). (SUSSKIND J.; BARNET; BLAISDELL, 2003) aperfeiçoaram o AIRS-TCA de modo a eliminar o efeito da nebulosidade no processo de inversão da ETR. Mediante o uso de radiâncias sintéticas do AIRS simuladas por um modelo direto com a presença de nuvens, os autores observaram uma acurácia do algoritmo equivalente ao ruído do canal do AIRS para a inferência de perfis de temperatura e umidade atmosférica.

#### **4.3.3.3 O Método Iterativo de Máxima Verossimilhança (IMLM)**

Este método é baseado no ajuste de um modelo do sinal esperado do detetor às medidas variando-se as quantidades totais dos gases traço que são relevantes dentro da janela espectral selecionada. Em outras palavras, o modelo direto inclui não somente a absorção atmosférica mas também as características do instrumento. Em particular a convolução com a função fenda do instrumento (“slit function”) é realizada de uma maneira lógica, evitando a necessidade da denominada correção  $I_0$  que é requerida quando o sinal observado no detetor é convertido para refletividade como é feito na maioria dos métodos de recuperação de perfis do tipo DOAS. Como os resultados do modelo e as contagens medidas pelo detetor podem ser diretamente comparadas, a influência das peculiaridades do instrumento tais como sinal escuro (“dark signal”), efeitos pixel-a-pixel, etc. são tratados mais facilmente. Por outro lado o uso de cal-

ibrações solares torna-se ligeiramente mais complexo. Um espectro de radiância é computado utilizando um espectro solar modelo  $I_0$  em uma unidade astronômica, com valores estimados correntemente das densidades de coluna inclinada e um valor de albedo da superfície fixo  $A$  assumindo reflexão lambertiana.

$$I(\lambda) = I_0(\lambda)e^{-\tau \cos(\theta)}A\pi^{-1}R_\theta^{-2} \quad (4.10)$$

onde  $R_\theta$  é a distância sol-terra. Este espectro é então alimentado para dentro do modelo do instrumento produzindo as medidas do detetor modelado. O resultado é multiplicado por um termo de fechamento (um polinômio de ordem  $n$ , onde normalmente  $n = 1$  ou  $2$ ) afim de combinar o nível do sinal medido e levar em conta os erros de calibração de baixa-frequência. As medidas computadas pelo detetor são corrigidas para diferenças entre o espectro solar do modelo e o espectro solar medido pelo SCIAMACHY utilizando calibração solar. O modelo de instrumento também calcula o ruído experimental esperado pela avaliação das várias fontes de ruído. Além disso, o modelo direto calcula as derivadas das medidas esperadas com relação aos parâmetros de ajuste, isto é, as colunas totais e os termos do polinômio de fechamento. O processo de ajuste objetiva descobrir os valores para as colunas totais dos gases (assumindo um perfil vertical fixo) e os coeficientes do fator de escala polinomial que correspondem à verossimilhança estatística máxima para as medidas do detetor. Por causa das características não-lineares do modelo, isto requer um procedimento iterativo. Em cada iteração o modelo e suas derivadas devem ser recalculados. Nós denotamos o resultado do modelo direto como o vetor  $I(p) = I_1(p_1 \dots p_m) \dots I_k(p_1 \dots p_m)$ , consistindo, portanto, de  $K$  elementos  $I_k$  cada um dependendo de  $M$  parâmetros  $p_m$ . As medidas são representadas pelo vetor de  $K$  elementos  $N$  e uma matriz de covariância de medida  $\Sigma$ . Então dentro da  $i$ -ésima iteração um valor atualizado de  $p$  é calculado como

$$p_{i+1} = p_i + (D^T \Sigma^{-1} D)^{-1} \times D^T \Sigma^{-1} \times (N - I(p_i)) \quad (4.11)$$

onde  $D$  é uma matriz  $K \times M$  com elementos  $D_{km} = \frac{\partial I_k}{\partial p_m}$ . A covariância dos parâmetros estimados (isto é, o erro devido ao ruído instrumental) é igual a

$$(D^T \Sigma^{-1} D)^{-1} \quad (4.12)$$

Estas expressões podem ser avaliadas por procedimentos numéricos padrão, baseadas na decomposição de Cholesky. Os valores iniciais para as colunas totais de gás traço podem ser convenientemente assumidos como sendo iguais a zero. A convergência é alcançada depois de 4 iterações. As colunas verticais totais são computadas assumindo-se um fator ar-massa (AMF) geométrico calculado a partir de um ângulo solar zenital e um ângulo de visão em uma geometria esférica. O algoritmo tem sido testado extensivamente sobre inúmeros espectros sintéticos. Ele tem se mostrado robusto e os resultados independem dos valores iniciais dentro de uma iteração (KLEIPOOL et al., 2005).

#### 4.3.3.4 O esquema 4D-VAR

Um sistema de assimilação de dados 4D-Var é uma formulação prática da teoria da estimação Bayesiana para o caso particular de um problema (quase) linear com erros Gaussianos não tendenciosos (LORENC, 1986). Ele procura por uma trajetória modelo que é estatisticamente consistente com a informação fornecida pelas observações  $y^o$  para a janela de análise de tempo  $[t_0, t_n]$  e pela informação fornecida por um modelo de estado *a priori*  $x^b$  denominado “estado de fundo”. Este estado de fundo é geralmente encarado como uma previsão de curto alcance. A trajetória do modelo em si é completamente definida pelo estado inicial  $x_0$  no tempo  $t_0$  através do uso de um modelo de previsão físico-dinâmico.

A análise de correção ( $\delta x(t_0)$ ) para o estado inicial do modelo é visto como uma combinação das informações das observações e do estado de fundo utilizando uma função custo objetivo com dois termos (COURTIER P.; THEPAUT; HOLLINGSWORTH, 1994):

$$J(\delta x(t_0)) = \frac{1}{2} \delta x(t_0)^T + \frac{1}{2} \sum_{i=0}^n [H_i \delta x(t_i) - d_i]^T R^{-1} [H_i \delta x(t_i) - d_i] \quad (4.13)$$

o termo de fundo e o termo observação. As observações de saída ( $d_i$ ) são as diferenças entre as radiâncias observadas e as radiâncias simuladas pelo modelo

$$d_i = y_i^o - H_i [x^b(t_i)] \quad (4.14)$$

onde  $H_i$  é o operador não-linear de observação na forma de transferência radiativa para o modelo de transferência radiativa do instrumento. Para experimentos que envolvam assimilação de gases traço inclue-se o gás como um perfil variável no modelo de transferência radiativa.  $H_i$ , que aparece na equação 4.13, é o operador de observação tangente linear que é parte do modelo direto. Os valores de fundo no tempo  $t_i$ , necessários para o cálculo das observações de saída  $d_i$ , são evoluídos de acordo com o modelo de previsão não linear  $\Phi$ :

$$x^b(t_i) = \Phi [x^b(t_0)] \quad (4.15)$$

Os incrementos evoluem com o tempo de acordo com o modelo tangente linear  $\Phi$ :

$$\delta x(t_i) = M \delta x(t_0) \quad (4.16)$$

Finalmente,  $B$  e  $R$  são a matriz de covariância dos erros de fundo e a matriz de covariância dos erros de observação, respectivamente. A função custo é então minimizada com relação aos incrementos do estado inicial  $\delta x(t_0)$ . Estes incrementos são adicionados ao estado de fundo para se obter a análise  $x(t_0)$ :

$$x(t_0) = x^b + \delta x(t_0) \quad (4.17)$$

A vantagem de um sistema completo de assimilação de dados é que ele procura combinar todas as observações disponíveis de uma maneira (quase) ótima. Alguns esforços consideráveis têm sido feitos para a implementação do 4DVar durante as últimas décadas ((ZUPANSKI D.; MESINGER, 1993);(ZUPANSKI D.; MESINGER, 1995);(COURTIER P.; THEPAUT; HOLLINGSWORTH, 1994);(ENGELEN et al., 2004) ;(FRANKENBERG et al., 2008);(ERRERA et al., 2008)). Contudo, devido ao grande custo computacional, implementar o esquema 4DVar completo em um ambiente operacional é praticamente impossível sem substanciais modificações (RABIER et al., 2000).

### 4.3.3.5 Esquema de Inversão baseada em Estimação Ótima

Este algoritmo de inversão se baseia na técnica de estimativa ótima para encontrar o estado atmosférico mais provável dada uma medida SCIAMACHY e algum conhecimento prévio. Quase todas as expressões matemáticas utilizadas por essa abordagem, assim como suas derivações e a notação podem ser encontradas no livro texto de Rodgers (RODGERS, 1997).

Considere o modelo direto  $F$  uma função vetorial que calcula para um dado estado atmosférico medidas simuladas correspondentes, ou seja, espectros SCIAMACHY simulados. As entradas para o modelo direto são o vetor de estado  $x$  e o vetor de parâmetros  $b$ . O vetor de estado consiste de todas as variáveis desconhecidas que deveriam ser recuperadas a partir da medição ( $CO_2$ , por exemplo). Os parâmetros que são considerados conhecidos com exatidão, mas que afetam a transferência radiativa (por exemplo, geometria de visada) são os elementos do vetor de parâmetros. O vetor de medidas  $y$  consiste de radiâncias SCIAMACHY, normalizadas pelo espectro solar, obtidas a partir da fusão de duas janelas espectrais concatenando as medições na janela de ajuste de  $CO_2$  e  $O_2$ . A diferença entre medida e correspondente simulação pelo modelo direto é dada pelo vetor de erro  $\varepsilon$  compreendendo imprecisões do instrumento e do modelo:

$$y = f(x, b) + \varepsilon \quad (4.18)$$

De acordo com a equação. 5.3 de (RODGERS, 1997), o objetivo é encontrar o vetor de estado  $x$  que minimiza a função custo  $\chi^2$ :

$$\chi^2 = [y - F(x, b)]^T S_\varepsilon^{-1} [y - F(x, b)] + (x - x_a)^T S_a^{-1} (x - x_a) \quad (4.19)$$

nesta equação,  $S_\varepsilon$  é a matriz de covariância de erro correspondente ao vetor de medição,  $x_a$  é o vetor de estado *a priori* que detém o conhecimento prévio sobre os elementos do vetor de estado e  $S_a$  é a matriz de covariância de erro *a priori* correspondente, que especifica as incertezas dos elementos do vetor de estado *a priori*, bem como suas correlações cruzadas. (RODGERS, 1997) sugere a utilização do método de Gauss-Newton para encontrar iterativamente o estado vetor  $x$  que minimiza a função custo:

$$x_{i+1} = x_i + \widehat{S}[K_i^T S_\varepsilon^{-1}(y - F(x_i, b)) - S_a^{-1}(x - x_a)] \quad (4.20)$$

Nesta equação,  $K$  é a matrix Jacobiana ou de função peso qua consiste da derivada do modelo direto em relação aos elementos do vetor estado  $K = \partial F(x, b)/\partial x$ . Em caso de convergência,  $x_{i+1}$  é a solução mais provável, dada a medida e o conhecimento *a priori* e é então denotada como a máxima solução *a posteriori*  $x$  do problema inverso.  $S$  é a matriz de covariância correspondente composta das variâncias dos elementos do vetor estado recuperados e suas correlações.

#### 4.3.3.6 O algoritmo TOSOMI

O TOSOMI (“SCIAMACHY total ozone retrieval algorithm“) é uma aplicação do algoritmo TOGOMI para SCIAMACHY. O algoritmo TOGOMI (VALKS P.; VAN OSS, 2003) é baseado no algoritmo DOAS [(PLATT U.; PERNER, 1983);(PLATT, 1994)] desenvolvido para o instrumento OMI (VEEFKIND J. P.; DE HAAN, 2001). Os instrumentos OMI, SCIAMACHY e GOME são espectrômetros similares que medem a radiação solar espalhada na faixa do UV e visível.

O algoritmo TOSOMI consiste de 4 passos. Primeiro, aplica-se um conjunto de ferramentas (software) aos arquivos de nível-1 do Sciamachy para melhorar a precisão da radiância terrestre medida e do espectro de irradiância solar. No segundo passo, o método DOAS é utilizado para ajustar o espectro de absorção diferencial de referência do ozônio ao espectro de radiância terrestre e ao espectro de irradiância solar, para obter a densidade de coluna inclinada (slant column density) do gás. A densidade de coluna inclinada do gás é a quantidade total de ozônio ao longo de um caminho médio que os fótons percorrem a partir do sol, através da atmosfera em direção ao sensor do satélite. No terceiro passo a densidade da coluna do gás é traduzida em densidade de coluna vertical utilizando o assim chamado fator ar-massa (AMF). O fator ar-massa  $M$  é definido como a razão entre a densidade de coluna inclinada,  $N_s$ , e a densidade de coluna vertical,  $N_v$ , i.e.,  $M \equiv \frac{N_s}{N_v}$ . O quarto passo consiste em uma correção para efeitos de nuvens através da utilização do algoritmo FRESCO (KOELEMEIJER et al., 2001).

#### 4.3.3.7 Considerações gerais sobre os métodos de inversão

Ainda que existam muitos algoritmos de inversão de gases traço em operação na atualidade, cada um com as suas características específicas, a maioria deles apresenta desvantagens em relação aos algoritmos baseados em RNAs.

Algoritmos como o 4D-Var precisam resolver uma função custo e dependem de uma boa estimativa inicial, além de necessitarem de um bom modelo para cálculo dos valores de fundo. Já o esquema de inversão baseada em Estimação ótima além de exigir uma boa estimativa inicial, requer o cálculo de uma matriz Jacobiana, a execução repetida de um modelo direto e a resolução de uma função custo. Tudo isso agrega alto custo computacional ao processo de inversão. Tanto o algoritmo IMLM quanto o WFM-DOAS necessitam, em cada iteração, computar uma função peso através de um modelo de transferência radiativa e resolver uma função custo em tempo de execução. Embora o WFM-DOAS seja um algoritmo rápido e muito utilizado atualmente a sua dependência com o modelo direto nunca poderá ser ignorada.

As RNAs permitem inversão sem a necessidade de resolver o modelo direto a cada iteração já que, após treinadas, exigem apenas um ciclo para a inversão. Modelos baseados em RNAs não precisam resolver uma função custo nem exigem uma boa estimativa inicial. Além disso o código resultante é facilmente paralelizável e podem ser implementados e disponibilizados em hardware.



## 5 REDES NEURAIS ARTIFICIAIS

Neste capítulo apresentamos uma breve introdução às Redes Neurais artificiais, o qual servirá como alicerce para o modelo de inversão de constituintes atmosféricos ( $CO_2$  e  $CH_4$ ) proposto em nosso trabalho. Apresentamos também uma breve revisão sobre o desenvolvimento deste ramo da computação científica ao longo das últimas décadas. Na sequência introduzimos o modelo do neurônio biológico e artificial. Discutimos as arquiteturas das RNA exploradas neste texto. Descrevemos o funcionamento das RNAs e, em especial, as fases de treinamento e ativação, bem como os paradigmas e regras de aprendizagem que aparecem na literatura com maior frequência. Reservamos uma seção para a descrição dos diversos tipo de Redes Neurais existentes entre elas as redes Perceptron Multi-Camadas (PMC), Rede de Função de Base Radial (FBR), Rede Neural de Elman (RNA-E) e Redes Neurais de Jordan (RNA-J). Por fim discutimos algumas razões que motivam o emprego de RNA e algumas dificuldades encontradas na aplicação desta ferramenta.

### 5.1 Breve histórico acerca das Redes Neurais Artificiais

Por muitas décadas, tem sido um objetivo da ciência e da engenharia desenvolver máquinas inteligentes com um grande número de elementos simples. Referências a este assunto podem ser encontradas na literatura científica do século 19. Durante a década de 1940, pesquisadores querendo replicar a função do cérebro humano, desenvolveram modelos de hardware simples (e mais tarde software) de neurônios biológicos e seus sistemas de interação. McCulloch e Pitts ([MCCULLOCH W. W.; PITTS, 1943](#)) publicaram o primeiro estudo sistemático de redes neurais artificiais. Quatro anos depois, os mesmos autores exploraram paradigmas de rede para o reconhecimento de padrões usando um perceptron de camada simples ([MCCULLOCH W. W.; PITTS, 1947](#)). Na década de 1950 e 1960, um grupo de pesquisadores combinaram essas percepções biológicas e psicológicas para produzir a primeira rede neural artificial (RNA) ([MCCLELLAND J. L.; RUMELHART, 1986](#)) e ([ROSENBLATT, 1961](#)). Inicialmente implementados como circuitos eletrônicos, foram posteriormente convertidos em um meio mais flexível de simulação por computador. No entanto, alguns anos depois, pesquisadores como Minsky e Papert ([MINSKY M.; PAPERT, 1969](#)) questionaram estes trabalhos. Eles acreditavam firmemente que sistemas inteligentes eram essencialmente processamento simbólico do tipo facilmente modelado em um computador de Von Neumann.

Por uma série de razões, a abordagem de processamento simbólico tornou-se o método dominante. Além disso, o perceptron como proposto por Rosenblatt acabou por ser mais limitado do que se previa inicialmente. Embora as pesquisas com RNA continuaram durante a década de 1970 por vários pesquisadores pioneiros como Grossberg, Kohonen, Widrow e outros, seus trabalhos receberam relativamente pouca atenção. Foi somente no início da década de oitenta, após a obtenção de alguns importantes resultados teóricos destacando-se, dentre eles, a descoberta da retro-propagação do erro e o aumento da capacidade de processamento devido ao desenvolvimento de avanços em hardware que o interesse pelas redes neurais ressurgiu.

Hopfield (HOPFIELD, 1982) formulou um modelo de rede recorrente com conexões sinápticas simétricas, supondo que os sistemas devam convergir para um estado de mínima energia. O modelo de Hopfield é usualmente referenciado na literatura como sinônimo de rede recorrente. Foi neste trabalho que pela primeira vez o princípio do armazenamento de informação em redes dinamicamente estáveis foi explícito. Assim, o trabalho de Hopfield dá um novo impulso ao estudo de RNA. Além disso a fundação matemática realizada por Hopfield possibilitou a compreensão da dinâmica de uma importante classe de redes até então pouco exploradas e compreendidas. Em 1983 foi desenvolvido um novo procedimento denominado recozimento simulado para resolver problemas de otimização combinatória. Esta idéia levou ao desenvolvimento da máquina de Boltzman (ACKLEY D.; HINTON; SEJNOWSKI, 1985), que foi a primeira realização bem sucedida de uma rede neural de múltiplas camadas. Como destacado por Haykin (HAYKIN, 2001), apesar de o algoritmo de aprendizagem da máquina de Boltzman não ter se mostrado tão eficiente do ponto de vista computacional como o algoritmo de retropropagação, ele superou o impasse psicológico, mostrando que a especulação de Minski e Papert não estava corretamente embasada. Rumelhart et al. (RUMELHART D. E.; HINTON; WILLIAMS, 1986) propuseram a utilização do algoritmo de retropropagação em aprendizagem de máquina, que tornou-se o algoritmo de aprendizagem mais popular para o treinamento do Perceptron Multi-Camadas (PMC).

O interesse em redes neurais vem de sua capacidade de aprender e responder. Como resultado, as redes neurais têm sido utilizadas em um grande número de aplicações complexas e provaram ser eficazes em uma variedade de campos. Estes incluem o reconhecimento de padrões, classificação, visão, sistemas de controle e predição

((WIDROW B.; HOFF, 1960) e (FAUSETT, 1994)). Este interesse renovado é refletido no número de cientistas, quantidade de fundos, número de grandes conferências realizadas, e no número de revistas científicas associados com redes neurais. A adaptação ou aprendizagem é o principal foco de investigação de RNA atualmente e que fornece um grau de robustez ao modelo. Na modelagem preditiva, o objetivo é mapear um conjunto de padrões de entrada para um conjunto de padrões de saída. Uma RNA realiza essa tarefa através da aprendizagem de uma série de conjunto de dados de entrada/saída apresentados à rede. A rede treinada é então utilizada para aplicar o que foi assimilado durante a aprendizagem para aproximar ou prever a saída correspondente (HAYKIN, 2001).

### 5.1.1 O Neurônio biológico

A estrutura de um neurônio artificial em rede neural é inspirada no conceito de neurônio biológico, mostrado na Figura 5.1.

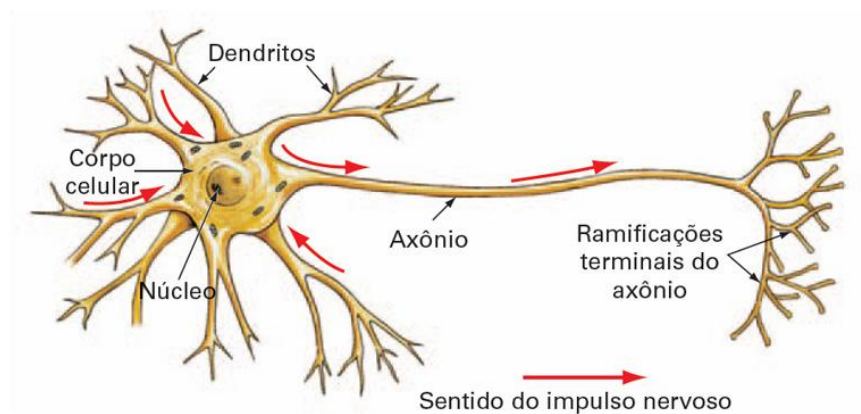


Figura 5.1 - Um Neurônio Biológico.  
Fonte: Wasserman (1989)

O cérebro é o elemento central do sistema nervoso humano, consistindo de quase  $10^{10}$  neurônios biológicos que são conectados uns aos outros através de sub-redes. O neurônio é delimitado por uma fina membrana celular que possui determinadas propriedades, essenciais ao funcionamento da célula. A partir do corpo celular projetam-se extensões filamentosas, os dendritos, e o axônio. Os neurônios são definidos como células polarizadas capazes de receber sinais em seus dendritos e transmitir informações por seus axônios. Ao ser excitado, um neurônio transmite informações,

através de impulsos, chamados potenciais de ação, para outros neurônios. Estes sinais são propagados como ondas pelo axônio da célula e convertidos para sinais químicos nas sinapses. O neurônio biológico pode ser visto como o dispositivo computacional elementar do sistema nervoso, composto de muitas entradas e uma saída. As entradas são formadas através das conexões sinápticas que conectam os dendritos aos axônios de outras células nervosas. Os sinais que chegam por estes axônios são pulsos elétricos conhecidos como impulsos nervosos ou potenciais de ação e constituem a informação que o neurônio processa para produzir como saída um impulso nervoso no seu axônio (KOVACS, 1996). Dependendo dos sinais enviados pelos axônios as sinapses podem ser excitatórias ou inibitórias. Uma conexão sináptica excitatória contribui para a formação de um impulso nervoso no axônio de saída, enquanto uma sinapse inibitória age no sentido contrário. A partir do conhecimento da estrutura e do comportamento dos neurônios naturais foram extraídas suas características fundamentais, utilizadas na criação de modelos de neurônios artificiais que simulam os reais. Estes neurônios artificiais são utilizados na formação das Redes Neurais Artificiais (OSORIO, 1991).

### 5.1.2 O Neurônio Artificial

O elemento básico que forma uma rede neural artificial é o neurônio artificial, conhecido também por nó ou elemento processador (ver Figura 5.2). Seu projeto foi baseado no funcionamento de um neurônio natural.

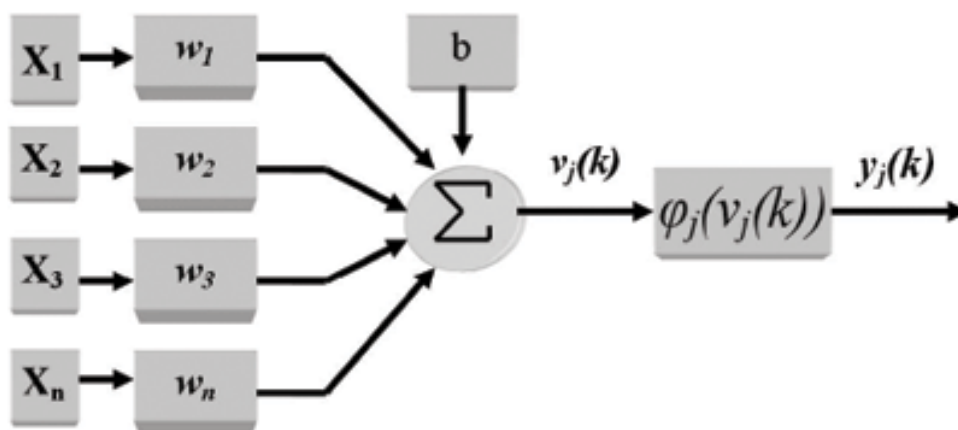


Figura 5.2 - Elementos básicos de um neurônio artificial.

O modelo do neurônio artificial ilustrado na Figura 5.2 acima pode ser equacionado da seguinte forma:

$$v_j(k) = \sum_{i=0}^n X_i w_i + b$$

$$y_j(k) = \varphi_j(v_j(k))$$

onde  $n$  é o número de sinais de entrada do neurônio,  $X_i$  é o  $i$ -ésimo sinal de entrada do neurônio,  $w_i$  é o peso associado com o  $i$ -ésimo sinal de entrada,  $b$  é o limiar de cada neurônio,  $v_j(k)$  é a resposta ponderada do  $j$ -ésimo neurônio em relação ao instante  $k$ ,  $\varphi_j(\cdot)$  é a função de ativação para o  $j$ -ésimo neurônio e  $y_j(k)$  é o sinal de saída do  $j$ -ésimo neurônio em relação ao instante  $k$ . Cada neurônio artificial é capaz de computar os sinais de entrada e a respectiva saída. A função de ativação usada para calcular o sinal de saída é tipicamente não-linear. As RNAs que processam dados analógicos, que estão envolvidas nesta aplicação, têm como função de ativação a sigmóide, a tangente hiperbólica, a Logarítmica e a Gaussiana (ver Figura 5.3). O processo de ajuste dos pesos da rede  $w_j$  associados ao  $j$ -ésimo neurônio de saída é feito pelo cálculo do sinal de erro em relação à  $k$ -ésima iteração ou ao  $k$ -ésimo vetor de entrada. Este sinal de erro é calculado pela seguinte equação:

$$e_j(k) = d_j(k) - y_j(k) \quad (5.1)$$

onde  $d_j(k)$  é a resposta desejada do  $j$ -ésimo neurônio de saída. Somando todos os erros quadráticos produzidos pelos neurônios de saída da rede em relação à  $k$ -ésima iteração, tem-se:

$$E(k) = \frac{1}{2} \sum_{i=0}^p e_j^2(k) \quad (5.2)$$

onde  $p$  é o número de neurônios da saída. Para uma configuração de pesos ótima,  $E(k)$  é minimizado pelo ajuste dos pesos sinápticos  $w_{ji}$ . Os pesos associados à camada de saída da rede são recalculados usando a seguinte expressão:

$$w_{ji}(k) \leftarrow w_{ji}(k) - \eta \frac{\partial E(k)}{\partial w_{ji}(k)} \quad (5.3)$$

onde  $w_{ji}$  é o peso conectando o  $j$ -ésimo neurônio da camada de saída ao  $i$ -ésimo neurônio da camada anterior e  $\eta$  é a constante que determina a taxa de aprendizado do algoritmo de retro-propagação. O ajuste dos pesos pertencentes às camadas escondidas é feito de forma análoga. Os passos para o ajuste destes pesos são detalhados em (HAYKIN, 2001).

O modelo do neurônio artificial proposto por McCulloch e Pitts (MCCULLOCH W. W.; PITTS, 1943) é uma simplificação do que se sabia na época a respeito do neurônio biológico. A sua descrição matemática se traduziu em um modelo com  $n$  portas de entrada (dendritos) com valores  $x_1, x_2, \dots, x_n$  e apenas um terminal de saída  $y$  representando o axônio. Para representar o comportamento das sinapses, os terminais de entrada possuem pesos associado  $w_1, w_2, \dots, w_n$ , alguns com sinais excitatórios (+) e outros com sinais inibitórios (-). Os valores de entrada e ativação das células podem ser discretos, nos conjuntos 0, 1 ou -1, 0, 1 ou contínuos nos intervalos [0,1] ou [-1,1]. O efeito de uma sinapse particular  $i$  no neurônio pós-sináptico é dada por  $x_i w_i$ . Pode-se dizer que os pesos determinam quantitativamente a importância dos sinais de disparo que ocorrem naquela conexão.

Um neurônio biológico dispara quando a soma dos impulsos que ele recebe ultrapassa um certo limiar de excitação pré-determinado. Esse comportamento do neurônio biológico, por sua vez, é representado no modelo artificial por um mecanismo que faz a soma ponderada dos valores  $x_i w_i$  recebidos pelo neurônio e decide se ele deve ser ativado (nesse caso a saída assume o valor 1) ou não (saída igual a 0), comparando a soma obtida com o limiar de excitação do neurônio. No modelo PMC (ver seção 5.2.2 abaixo) a ativação do neurônio é conseguida através da aplicação de uma função de ativação, que ativa ou não a saída de acordo com a soma ponderada das entradas. A função de ativação restringe a amplitude do sinal na saída de um neurônio. Geralmente se restringe o sinal ao intervalo [0,1] ou ao intervalo [-1,1]. A função utilizada para o cálculo de ativação geralmente é algum tipo de função não-linear, a qual garante a plena funcionalidade das redes neurais com múltiplas camadas de neurônios. Redes neurais que processam dados analógicos devem utilizar funções com formato sigmoidal como tangente hiperbólica, seno, etc. Já em redes utilizadas para processar valores discretos a função deve ser do tipo degrau. A

Figura 5.3 abaixo exibe algumas das funções de ativação usualmente utilizadas em RNAs:

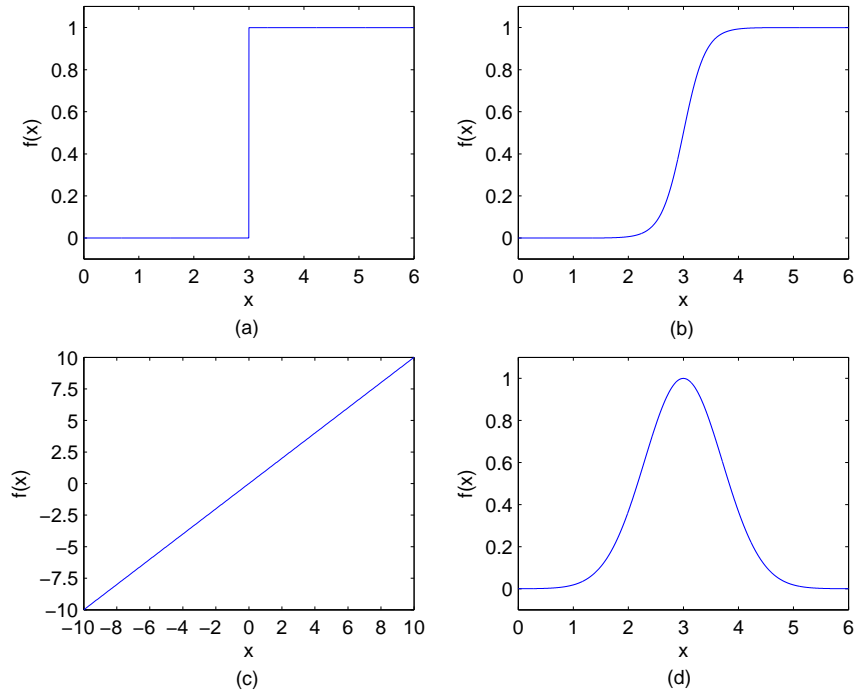


Figura 5.3 - Exemplos de Função de ativação:(a) Função Degrau; (b) Função Sigmóide; (c) Função linear; (d) Função gaussiana

Uma das funções de ativação utilizadas em nosso estudo foi a função tangente hiperbólica. Ela é contínua e é computacionalmente simples para qualquer computador. Esta função tem intervalo de existência entre -1 e 1, sendo assintótica nos dois ramos. Isto é, embora ela tenda a -1 e a 1, tais valores não são alcançados nunca. Ela é definida por:

$$\sigma(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5.4)$$

e a derivada é dada por

$$\sigma'(x) = 1 - \tanh(x)^2 = 1 - \sigma^2(x) \quad (5.5)$$

A função e sua derivada são traçadas na Figura 5.4 abaixo. A escolha dessa função de ativação implica que as saídas dos neurônios estão todas entre -1 e 1. Se a entrada é superior a 1 em valor absoluto, a saída do neurônio é bloqueada à  $\pm 1$  e dizemos que o neurônio está saturado.

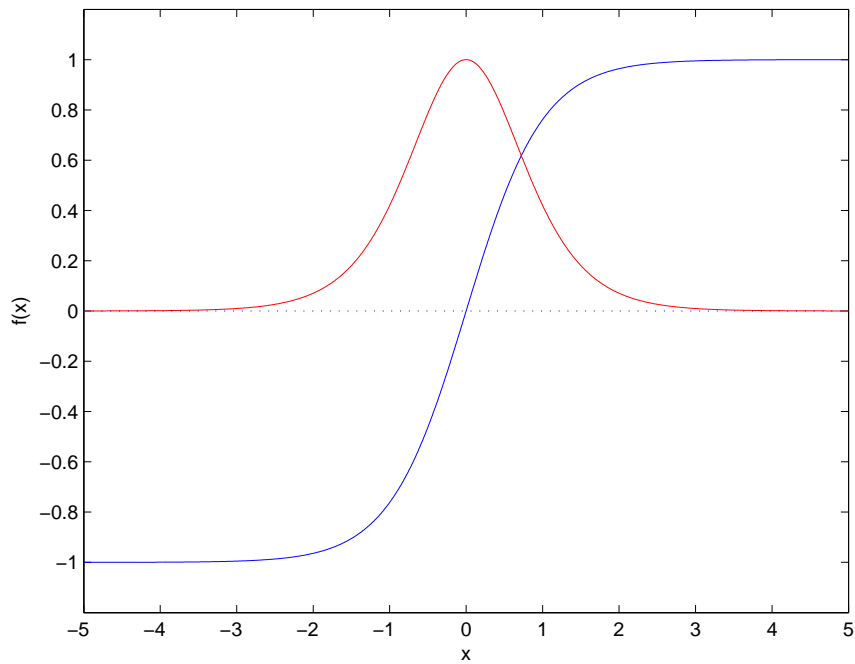


Figura 5.4 - Função sigmóide tanh (azul) e sua derivada (vermelho)

De acordo com uma ponderação dos sinais de entrada, realizada pela função de ativação, o neurônio pode ser ativado, enviando um sinal de saída. Este sinal de saída será propagado de acordo com a topologia de interconexão da rede de neurônios.

No início das pesquisas envolvendo redes neurais, estas eram emuladas através de circuitos analógicos. Atualmente o mais comum é a utilização de software para a geração de redes neurais, ou seja, o funcionamento dos neurônios é simulado por programas. Uma tendência recente é a implementação das redes neurais em circuitos digitais, visando a maximização da velocidade de processamento.



## 5.2 Arquitetura de Redes Neurais

As diferentes arquiteturas de RNAs são formadas pela combinação de neurônios artificiais e são definidas pelo tipo de conexão entre as redes. O mais importante caso especial é a arquitetura com alimentação para frente (*feedforward*) em que os neurônios são divididos em camadas. Cada neurônio transfere o seu sinal apenas para os neurônios que se encontram em uma das camadas subseqüentes, ou seja, não há retro acoplamento entre os neurônios. Três tipo de camadas são identificadas:

- a) Camada de entrada: é a interface de entrada, onde os sinais de entrada externos são alimentados para dentro da rede.
- b) Camada de saída: os sinais de saída da rede são capturados a partir dos nós de saída.
- c) Camadas Ocultas: Definem a representação interna do mapeamento e não tem ligação direta com o usuário externo.

Os neurônios podem ser dispostos de maneira a formar uma ou mais camadas. Em parte da literatura as entradas são consideradas a primeira camada da rede. Neste texto as entradas serão denominadas de camada de entrada, a primeira camada sem contato com o meio externo de primeira camada oculta e assim sucessivamente para as demais camadas ocultas. Os valores na saída da rede são denominados de camada de saída.

Destaca-se ainda na arquitetura das RNAs, os tipos de conexão entre os neurônios, ou seja, diz-se que a rede é totalmente conectada quando todas as saídas de uma camada estão conectadas em todos os neurônios da próxima camada. Na falta de alguma dessas conexões se diz que a rede é parcialmente conectada. A definição da arquitetura de uma RNA é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser tratado pela rede. RNA com uma única camada de nodos, por exemplo, só consegue resolver problemas linearmente separáveis. Redes recorrentes, por sua vez, são mais apropriadas para resolver problemas que envolvam processamento temporal (HAYKIN, 2001).

### 5.2.1 Redes Neurais Multicamadas

Redes neurais multicamadas são arranjos ou arquiteturas em que neurônios são organizados em duas ou mais camadas. A Figura 5.5 representa um esquema típico de uma rede neural artificial com múltiplas camadas. As entradas são conectadas aos elementos processadores básicos, que são por sua vez interconectados com elementos de outras camadas e/ou a saída da rede.

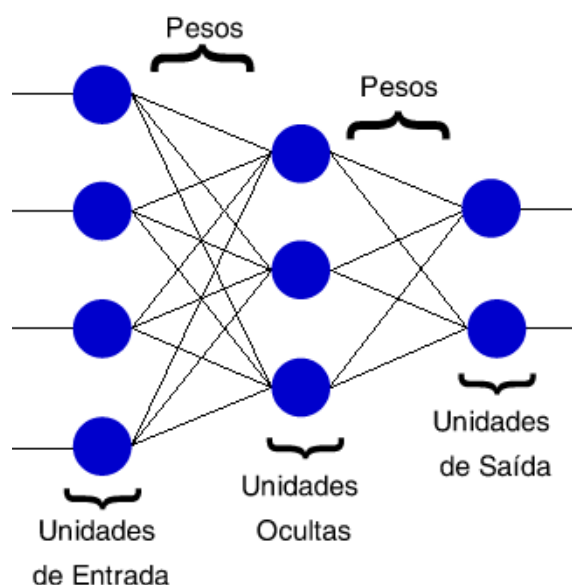


Figura 5.5 - Ilustração de uma rede multicamadas.

A camada de entrada é especial pois não realiza qualquer processamento. Ela apenas distribui os valores de entradas para todos os neurônios da 1ª camada de processamento.

### 5.2.2 Perceptrons Multi-Camadas

Esta é talvez a mais popular arquitetura de rede em uso na atualidade e sua criação é creditada à Rumelhart e McClelland (RUMELHART D. E.; HINTON; WILLIAMS, 1986) e discutida extensamente na maioria dos livros textos (BISHOP, 1995)). Em uma rede Perceptron Multi-Camadas (PMC) cada unidade executa uma soma ponderada de suas entradas e transmitem esse nível de ativação através de uma função de transferência para produzir uma saída, e as unidades são arranjadas em uma topologia de camadas. A rede, portanto, tem uma interpretação simples como uma forma de

modelo entrada-saída, com os pesos e vieses como parâmetros livres do modelo. Tais redes podem modelar funções de complexidade arbitrária com o número de camadas e o número de unidades em cada camada determinando a complexidade da função. Importantes temas relacionados ao design de Perceptrons Multi-Camadas (PMC) incluem a especificação do número de unidades ocultas e o número de unidades dentro dessas camadas ((HAYKIN, 2001);(BISHOP, 1995)). A definição do número de neurônios das camadas de entrada e saída é realizada de acordo com o problema ao qual a rede solucionará. O número de neurônios intermediários, ou mesmo o número de camadas ocultas é definido de forma intuitiva, não havendo, portanto, uma regra que defina o seu número. Se o número de neurônios for muito alto, isto pode fazer com que apenas alguns neurônios especializem-se em características úteis, porém, se o contrário ocorrer e o número de neurônios for insuficiente, pode acontecer da rede não conseguir aprender os padrões desejados. Neste trabalho optamos pela adoção desta arquitetura por apresentar melhor relação de custo computacional versus desempenho e também por ser relativamente fácil de se implementar em software e hardware. Porém, desenvolvemos uma heurística de otimização para a definição dos parâmetros da rede, inclusive do número de neurônios ocultos (Para saber mais detalhes ver seção 6.3 do capítulo 6).

### 5.2.2.1 Capacidade de representação do PMC

Numerosos trabalhos teóricos foram publicados sobre a capacidade de representação das Perceptron Multi-Camadas ((MINSKY M.; PAPERT, 1969); (HORNIK, 1993)). O teorema principal que resultou deles pode ser enunciado da seguinte forma: Toda função contínua sobre um compacto (conjunto fechado e limitado) pode ser aproximada uniformemente e também precisamente por um PMC com uma camada oculta. De onde vêm o qualificativo usual das redes PMC de aproximador universal. Este teorema é de fato um resultado importante, mas ele nos fornece somente uma informação de existência. Em particular, nada é dito sobre a arquitetura do PMC e de sua complexidade. Na prática, o PMC pode ser visto como um sistema parametrizado que associa um espaço  $I \subset R^n$  a um espaço  $J \subset R^p$  por intermédio de uma função  $f : I \longrightarrow J$  que a uma entrada  $\bar{x} \in I$  associa uma saída  $\bar{y} = f(\bar{\omega}, \bar{x} \in I) \in J$ , onde  $\bar{\omega}$  contém os parâmetros da rede (os pesos sinápticos) que são determinados ao longo da fase de aprendizagem.

No caso de uma aprendizagem supervisionada do PMC, a fase de aprendizagem consiste em estabelecer as conexões entre as entradas e saídas de uma base de  $N$  ex-

emplos, de maneira a prever, para os casos não-observados dentro da base, os valores das saídas conhecendo-se os valores das entradas. Configura-se, nesse caso, o problema da representatividade da base de aprendizagem: Até que ponto o PMC é capaz de se aproximar da função a ser estimada a partir de um número finito de exemplos de realização dessa função ? Em particular, um caso delicado a tratar, embora atual, ocorre quando as variáveis a estimar não podem ser totalmente explicadas pelas variáveis de entrada. Podemos então mostrar que a rede tem tendência a estimar a esperança das variáveis de saída conhecendo as variáveis de entrada (BISHOP, 1995).

### 5.2.2.2 Aprendizagem da Rede

Os pesos sinápticos são determinados por um algoritmo conhecido como algoritmo de aprendizagem que consiste em modificar os pesos em função de um conjunto de dados apresentados à entrada, e algumas vezes à saída, da rede. O objetivo dessa aprendizagem é, como foi indicado, de permitir à rede aprender, em função dos exemplos que lhes são apresentados, as operações que nós desejamos que sejam executadas. Uma vez a aprendizagem (corretamente) realizada, os pesos são fixos e a rede possui uma certa capacidade de generalizar, isto é, é capaz de fornecer um sinal de saída correspondente a um sinal de entrada que não tenha sido aprendido. A aprendizagem é dita supervisionada quando, da apresentação de uma certa entrada, obriga-se a rede a convergir na direção de um estado final preciso. Nesse caso, é necessário possuir uma base de aprendizagem que, a um conjunto de dados de entrada, associe um conjunto de dados de saída. É o tipo de aprendizagem utilizada quando queremos treinar a rede para modelar uma certa função. Dentro de uma aprendizagem não-supervisionada o conjunto de saída não é mais necessário, ou seja, não existe saída *a priori* e a rede é deixada livre para convergir na direção de qualquer estado. É a técnica de aprendizagem utilizada quando queremos fazer a classificação ou a compressão de dados, das quais os mapas topológicos de Kohonen são um exemplo clássico (KOHONEN, 1995). Finalmente, temos a aprendizagem semi-supervisionada, ou por reforço. Neste tipo de aprendizagem, ao invés de fornecer as saídas corretas para a rede, relativo a cada treinamento individual, a única informação de realimentação fornecida à rede é se uma determinada saída está correta ou não. Isto é, não é fornecida à rede a resposta correta para o padrão de entrada ((HECHT-NIELSEN, 1989); (FREEMAN J.; SKAPURA, 1991); (BEALE R.; JACKSON, 1990)). O aprendizado por reforço pode ser visto como um caso particular de aprendizado supervisionado. A principal diferença entre o aprendizado supervisionado clássico e o aprendizado por

reforço, é a medida de desempenho usada em cada um dos sistemas. No aprendizado supervisionado, a medida de desempenho é baseada no conjunto de respostas desejadas usando um critério de erro conhecido, enquanto que no aprendizado por reforço o desempenho é baseado em qualquer medida que possa ser fornecida ao sistema. Em termos mais práticos, redes neurais são ferramentas estatísticas de modelagem de dados não-lineares. Elas podem ser usadas para modelar complexos relacionamentos entre entradas e saídas ou para descobrir padrões nos dados.

As regras comumente aplicadas com o paradigma supervisionado são correção de erros e a regra delta, que foi generalizada para o treinamento do Perceptron Multi-Camadas, conhecido como algoritmo de retropropagação do erro. O algoritmo de retropropagação é a mais poderosa técnica para o aprendizado do Perceptron Multi-Camadas e é o algoritmo de aprendizagem escolhido neste trabalho. A seguir resume-se o algoritmo de retropropagação, de acordo com Haykin (HAYKIN, 2001), utilizado no treinamento das redes Perceptron Multi-Camadas, rede de Funções de Base Radial e redes recorrentes (Jordan e Elman).

- a) Inicializar os pesos: assumindo nenhuma informação prévia disponível, os pesos sinápticos e limiares são iniciados com números aleatórios que seguem uma distribuição uniforme.
- b) Apresentação dos exemplos de treinamento: apresenta-se uma época de exemplos de treinamento à rede. Para cada exemplo do conjunto de treinamento executa-se os passos c e d apresentados a seguir.
- c) Cálculo para Frente (Propagação): sendo um exemplo do conjunto de treinamento representado por  $(x(n), d(n))$ , com o vetor de entrada  $x(n)$  aplicado a camada de entrada de nós sensoriais e o vetor resposta desejada  $d(n)$  apresentado à camada de saída de nós computacionais, calcula-se os campos locais induzidos e os sinais funcionais camada por camada da rede. O campo local induzido  $v_j^{(l)}(n)$  para o neurônio  $j$  na camada  $l$  é

$$v_j^{(l)}(n) = \sum_{i=0}^{m_0} w_{ji}^{(l)}(n) y_i^{l-1}(n), \quad (5.6)$$

onde  $y_i^{l-1}(n)$  é o sinal de saída do neurônio  $i$  na camada interior  $l - 1$ , na iteração  $n$ , e  $w_{ji}^{(l)}(n)$  é o peso sináptico do neurônio  $j$  da camada  $l$ , que é

alimentado pelo neurônio  $i$  da camada  $l-1$ . Para  $i = 0$ , temos  $y_i^{l-1}(n) = +1$  e  $w_{ji}^{(l)}(n) = b_j^l(n)$  é o viés aplicado ao neurônio  $j$  na camada  $l$ . O sinal de saída do neurônio  $j$  na camada  $l$  é

$$y_j^{(l)} = \varphi_j(v_j(n)), \quad (5.7)$$

Se o neurônio  $j$  está na primeira camada oculta (i.e.,  $l = 1$ ), faz-se  $y_j^0(n) = x_j(n)$ ,  $x_j(n)$  é o  $j$ -ésimo elemento do vetor de entrada  $x(n)$ . Se o neurônio  $j$  está na camada de saída (i.e.,  $l = L$  onde  $L$  é denominado a profundidade da rede), se faz  $y_j^{(L)} = o_j(n)$ . Calcula-se o sinal de erro

$$e_j^{(n)} = d_j^{(n)} - o_j^{(n)}, \quad (5.8)$$

onde  $d_j^{(n)}$  é o  $j$ -ésimo elemento do vetor resposta desejada  $d(n)$ .

- d) Cálculo regressivo (Retropropagação): Calcula-se os gradientes locais da rede definidos por

$$\delta_j^{(l)}(n) = \begin{cases} e_j^{(L)}(n)\varphi_j'(v_j^L(n)) & \text{para o neurônio } j \text{ na camada de saída } L, \\ \varphi_j'(v_j^l(n)) \sum_k \delta_k^{(l+1)}(n)w_{kj}^{(l+1)}(n) & \text{para o neurônio } j \text{ na camada oculta } l \end{cases}$$

onde o apóstrofo em  $\varphi_j'(\cdot)$  representa a diferenciação em relação ao argumento. Assim, ajusta-se os pesos sinápticos da rede na camada  $l$  de acordo com seguinte regra, conhecida como regra delta generalizada:

$$w_j^{(n+1)} = w_j^{(n)} + \alpha[w_j^{(n)} - w_j^{(n-1)}] + \eta\delta_j^{(l)}(n)y_i^{(l-1)}(n), \quad (5.9)$$

onde  $\eta$  é a taxa de aprendizagem e  $\alpha$  é a constante de momento.

- e) Iteração: repete-se os itens c e d, apresentando à rede novos exemplos de treinamento até que o critério de parada seja satisfeito. Neste trabalho, a ordem de apresentação dos exemplos de treinamento seguem sempre a mesma seqüência, ao invés de serem apresentados de maneira aleatória, de época para época, como sugerido em Haykin (HAYKIN, 2001). O parâmetro de momento é constante, mas difere de experimento para experimento enquanto a taxa de aprendizagem é variável ao longo do treinamento.

O paradigma não-supervisionado é comumente aplicado com os algoritmos de aprendizagem Hebbiana (HEBB, 1949) e competitivo (KOHONEN, 1977) apud (KOHONEN, 1995). Durante o desenvolvimento deste trabalho, explorou-se a aprendizagem competitiva na implementação de Rede Função de Base Radial (RBF), ao se calcular o centro das funções gaussianas. Como os resultados para a rede RBF foram insatisfatórios ou em última instância equivalente aos obtidos pela rede PMC, resolvemos não exibí-los neste texto.

### 5.2.3 Outros tipos de Redes Neurais

#### 5.2.3.1 Rede Função de Base Radial

Redes com uma única camada oculta são capazes de aproximar com uma precisão pré-definida qualquer função contínua, pois uma função contínua e limitada sobre um certo intervalo, pode ser considerada como uma superposição linear de gaussianas. Estas gaussianas por sua vez, podem ser implementadas na camada oculta da RNA. A precisão da aproximação dependerá apenas do número de funções gaussianas, ou seja, do número de neurônios da camada oculta (KOVACS, 1996). Redes Neurais representando funções reais cujo valor real dependa apenas da distância de um ponto ao centro  $c$ , tal que  $\phi(x, c) = \varphi(\|x - c\|)$ , implementadas nos neurônios da camada oculta, denominam-se Funções de Base Radial (FBR). Podemos citar como exemplos de Funções de Base Radial as funções gaussianas, multi-quádricas, Spline poli-harmônica e Thin-plate Spline. Em um trabalho recente Shiguemori, (2008) introduziu uma nova forma de RBF conhecida como RBF não-extensiva com aplicações interessantes em várias áreas do conhecimento (SHIGUEMORI E. H.; CAMPOS VELHO; SILVA, 2008).

Do ponto de vista de uma FBR, aprender é equivalente a encontrar uma superfície, em um espaço multidimensional que forneça o melhor ajuste para os dados de treinamento, com o critério de melhor ajuste sendo medido num sentido estatístico (BENNETT, 1992). Neste trabalho, tanto a FBR como a PMC foram implementadas com aprendizagem por correção de erro (também conhecido como algoritmo de retropropagação do erro ou Regra Delta de Widrow) e paradigma de aprendizagem supervisionada. As redes são totalmente conectadas, apresentam uma camada de entrada, apenas uma camada escondida (no caso da rede RBF) e uma camada de saída. A rede FBR explorada neste trabalho difere da rede PMC apenas pelo

fato da função de ativação ser gaussiana (com média zero e variância unitária) e não tangente hiperbólica, Sigmoidal ou Logarítmica.

### 5.2.3.2 Redes Neurais Recorrentes

Muitos algoritmos de treinamento das RNAs não são capazes de implementar mapeamentos dinâmicos, como por exemplo o algoritmo de retropropagação simples, que pode apenas aprender mapeamentos estáticos. Um artifício utilizado para processamento temporal utilizando estas redes envolve o uso de janelas de tempo, onde a entrada da rede utiliza trechos dos dados temporais como se eles formassem um padrão estático. Entretanto, esta solução não é a mais indicada para o processamento temporal. A principal questão, portanto, é como estender a estrutura das redes Perceptron Multi-Camadas para que assumam um comportamento que varie com o tempo, sendo assim capaz de tratar sinais temporais. O tempo pode ser representado pelo efeito que acarreta no processamento de sinais, que significa oferecer características dinâmicas ao mapeamento realizado pela rede, tornando-a sensível a sinais que variem com o tempo. Para uma RNA ser considerada dinâmica é preciso que possua memória (ELMAN, 1990). Existem basicamente duas maneiras de prover memória a uma RNA. A primeira seria introduzindo atraso no tempo, como as técnicas TDNN (Time Delay Neural Network), e FIR (Finite Impulse Response) Multilayer Perceptron. A segunda maneira, seria a utilização de redes neurais recorrentes, tais como *Backpropagation Through Time*, *Real-Time Recurrent Learning*, *Cascade Correlation* recorrente, redes de Elman e redes de Jordan (CARVALHO A. C. P. L. F.; BRAGA; LUDERMIR, 1997). Pode-se definir redes recorrentes como aquelas que possuem conexões de realimentação que proporcionam comportamento dinâmico. Há dois tipos de redes recorrentes: aquelas em que o padrão de entrada é fixo e a saída caminha, dinamicamente, para um estado estável e aquelas em que ambas, entrada e saída, variam com o tempo, sendo estas últimas mais gerais e utilizadas com maior frequência. Há muitas variações de arquiteturas de redes recorrentes, sendo que algumas delas permitem o uso de algoritmos de treinamento mais simples ou adaptados a uma tarefa particular.

#### 5.2.3.2.1 Redes de Elman e Jordan

Elman introduziu memória às RNAs (ELMAN, 1990). Nas redes de Elman, além das unidades de entrada, intermediárias e de saída, há também unidades de contexto, como nas redes parcialmente recorrentes em geral. As unidades de entrada e saída



interagem com o ambiente externo, enquanto as unidades intermediárias e de contexto não o fazem. As unidades de entrada são apenas unidades de armazenamento (buffer) que passam os sinais sem modificá-los. As unidades de saída são unidades lineares que somam os sinais que recebem. As unidades intermediárias podem ter funções de ativação lineares ou não lineares, e as unidades de contexto são usadas apenas para memorizar as ativações anteriores das unidades intermediárias e podem ser consideradas como atraso no tempo em um passo. As conexões *feedforward* são modificáveis e as conexões recorrentes são fixas, motivo pelo qual a rede de Elman é apenas parcialmente recorrente. A Figura 5.6 ilustra o modo de operação de uma rede de Elman. Em um instante de tempo específico  $k$ , as ativações das unidades intermediárias (em  $k-1$ ) e as entradas correntes (em  $k$ ) são utilizadas como entradas da rede. Em um primeiro estágio, *feedforward*, estas entradas são propagadas para frente a fim de produzir as saídas. Posteriormente, a rede é treinada com o algoritmo de aprendizagem de retropropagação padrão. Após este passo de treinamento, as ativações das unidades intermediárias no tempo  $k$  são reintroduzidas através das ligações recorrentes nas unidades de contexto, sendo salvas nestas unidades para o próximo passo do treinamento ( $k+1$ ). No início do treinamento, as ativações das unidades intermediárias são desconhecidas e, geralmente, são inicializadas para a metade do valor máximo que as unidades intermediárias podem ter.

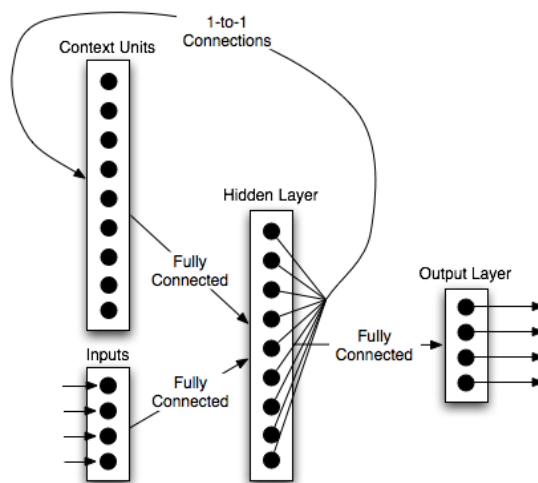


Figura 5.6 - Rede Neural recorrente.  
Fonte: Jordan (1986)

Na rede de Jordan (JORDAN, 1986), a saída da rede é copiada para a unidade de contexto. Adicionalmente, as unidades de contexto são localmente recorrentes. A grande diferença em termos de topologia entre as duas redes é que a recorrência na rede de Elman é feita da camada oculta para as entradas, enquanto que na rede de Jordan a recorrência é feita das saídas para as entradas. As redes neurais recorrentes foram utilizadas em trabalhos que abordam seu emprego no auxílio à tarefa de proteção de sistemas elétricos, como em (SEGATTO E. C.; COURY, 2005); (SEGATTO E. C.; COURY, 2006). Em assimilação de dados, um tipo importante e especial de Problema Inverso, as redes de Elman e Jordan foram utilizadas com sucesso e obtiveram um bom desempenho. A ideia básica foi introduzir um comportamento dinâmico aproximando desta forma as redes neurais do comportamento típico de um filtro de Kalman (ver (HARTER, 2004) e (HARTER F. P.; CAMPOS VELHO, 2005)).

## 6 METODOLOGIA PROPOSTA

Neste capítulo expomos a metodologia empregada neste projeto de pesquisa para resolver o problema da recuperação de perfis verticais de concentração de metano e dióxido de carbono atmosférico. Em uma primeira etapa propõe-se a utilização do modelo direto SCIATRAN (ver seção 4.2 do capítulo 4) para simularmos os dados de radiância na faixa espectral do infravermelho próximo a partir de perfis perturbados do gás.

Inicialmente procedemos à uma seleção dos canais que serão utilizados no processo de recuperação da concentração dos GEE. Foram implementados 4 métodos diferentes para a tarefa de escolher um conjunto restrito de canais a partir de um amplo espectro de comprimentos de onda. Cada método produziu um conjunto de canais independentes e auto-suficientes. O cálculo dos jacobianos dos canais para o  $CO_2$  e  $CH_4$ , necessário à cada uma das abordagens, foi realizado com o auxílio de um modelo direto. Esses e outros detalhes concernentes ao processo de seleção são explicados na seção 6.1 deste capítulo.

Uma vez escolhidos os canais montamos um conjunto de dados de treinamento, validação e teste para o modelo de inversão proposto. Esse conjunto de dados irá consistir basicamente de um conjunto de tuplas  $R, S$  onde  $R$  e  $S$  correspondem respectivamente a um vetor observação e ao vetor de estado atmosférico do qual o perfil de concentração é parte integrante. Propusemos a utilização do modelo direto para simular dados de radiância na faixa espectral do infravermelho próximo a partir de perfis perturbados do gás. Essa tarefa de geração da base de treinamento é detalhada na seção 6.2.

Na seção 6.3 apresentamos uma metodologia desenvolvida para otimizarmos os parâmetros de configuração do modelo inverso utilizado. No caso específico das redes neurais artificiais foram implementados algoritmos baseados em metaheurísticas híbridas de busca. A próxima etapa consiste da inversão propriamente dita. São realizados experimentos de recuperação de concentração com cada um dos conjuntos independentes de canais. Para tanto utilizamos o modelo de inversão próprio baseado em redes neurais. Uma breve explanação a respeito da teoria das redes neurais e de sua utilização como instrumento de auxílio nessa área do sensoriamento remoto é apresentada no capítulo 5. No capítulo 7 é conduzida a análise e validação dos resultados obtidos pelo modelo. A figura 6.1 abaixo ilustra os principais passos

da metodologia científica aplicada à trabalho de pesquisa.

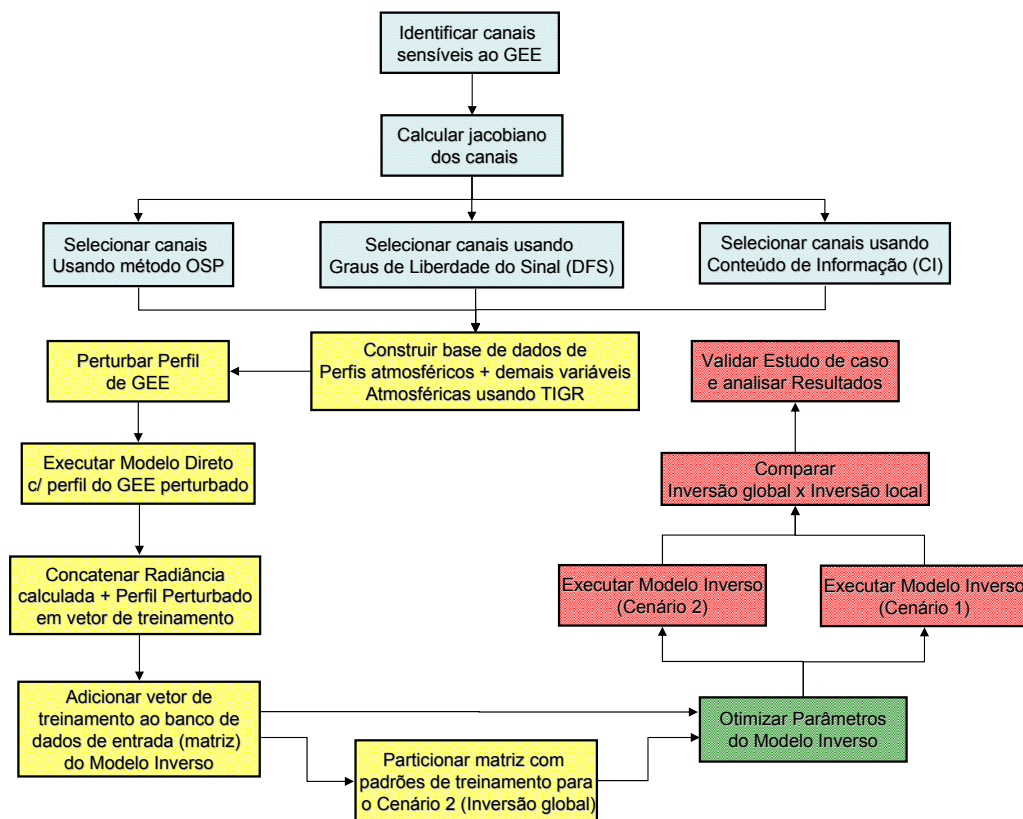


Figura 6.1 - Diagrama com a metodologia empregada. As cores servem para identificar e agrupar um conjunto de tarefas com características comuns.

## 6.1 Seleção de Canais

Nesta seção, descrevemos a metodologia utilizada para a seleção dos canais do sensor SCIAMACHY aproveitados no processo de recuperação do perfil de concentração dos gases  $CO_2$  e  $CH_4$ . É possível utilizar uma abordagem via RNA, em princípio, para mapear qualquer espaço vetorial de entrada para qualquer espaço vetorial de saída. No entanto, na prática, a representação de dados afeta significativamente a qualidade dos resultados. Podemos usar técnicas de redução de dimensão para apresentar não somente uma representação mais compacta, mas também informações mais relevantes para a entrada da rede (ver, por exemplo, (HAYKIN, 2001) e (ROJAS, 1996)). A “maldição da dimensionalidade” estipula que é difícil aplicar uma técnica

estatística para dados com alta dimensão espacial. O número de parâmetros (pesos na RNA) aumenta com o número de entradas. Isto pode permitir excessivos graus de liberdade no interpolador neural e introduzir dados pouco informativos (ou seja, o ruído ou informação espectral descorrelacionada com as quantidades recuperadas), o que pode distorcer o processo de aprendizagem (BELLMAN, 1961); (BISHOP, 1995).

Assim, o objetivo da redução da dimensão é apresentar à RNA as informações mais relevantes a partir de dados brutos iniciais, ou seja, medidas físicas ruidosas. Podemos reduzir a dimensionalidade dos dados de entrada através da extração de características (uma transformação linear, ou não, dos dados brutos) ou usando seleção de características através da seleção de canais específicos nos dados de entrada ((BISHOP, 1995); (JAIN A.; ZONGKER, 1997)). Em nosso trabalho optamos pela segunda abordagem (seleção de características)(RABIER et al., 2002). Para a recuperação de uma variável geofísica, selecionamos canais que são, na medida do possível, sensíveis somente a este parâmetro atmosférico. Ao estudar as derivadas das radiâncias em relação a cada um dos parâmetros geofísicos (Jacobianos da ETR), é possível inspecionar as informações comuns entre as radiâncias medidas e as variáveis geofísicas. No entanto, precisamos fazer um balanço entre reduzir a dimensionalidade dos dados e preservar o conteúdo de informação dos dados a fim de compensar os efeitos do ruído.

O sensor SCIAMACHY conta com 8 canais ou bandas espectrais principais cobrindo grande parte do espectro do infravermelho próximo que podem ser utilizados para inferência de variáveis atmosféricas (por ex: concentração de gases traço). Cada um desses canais é subdividido em centenas de outros subcanais totalizando alguns milhares de canais ou comprimentos de ondas candidatos. Utilizar todos os canais para recuperar  $CO_2$ ,  $CH_4$  ou qualquer outro gás traço seria proibitivo em termos de seu custo operacional e ao mesmo tempo desnecessário, haja vista que somente uma pequena fração desses canais é sensível à variação da concentração desses gases. Desse modo, precisamos extrair ou selecionar de alguma forma um subconjunto restrito de canais para realizarmos a inferência desses gases. Na literatura são descritos vários métodos para a concretização de tal tarefa.

Primeiro, podemos utilizar o método OSP (*Optimal Sensitivity Profile*) para estudar a sensibilidade dos canais em relação aos gases traço e outros componentes da atmosfera e da termodinâmica da superfície (CREVOISIER C.; CHEDIN; SCOTT, 2003). Outra forma de selecionar os canais é fazer um estudo do conteúdo de informação

dos canais através da teoria de Shannon sobre conteúdo de Informação. Associado a este método existe ainda na literatura um método baseado no estudo dos graus de liberdade para o sinal (GLS) do sensor, sendo esta uma medida do número de quantidades estatisticamente independentes de qualquer medição. Por último, sugerimos como método alternativo, que ainda necessita ser testado e validado, calcular a variância dos canais jacobianos calculados e elencar os canais com maior variabilidade intrínseca para realizar a inferência.

A maioria dos artigos publicados sobre a seleção de canais individuais são baseados em critérios objetivos ((RODGERS, 1996);(LERNER J. A.; WEISZ; KIRCHENGAST, 2002); (RABIER et al., 2002)). No entanto, todos eles lidam com temperatura e umidade. Neste trabalho ampliamos o escopo de aplicação desses métodos para o estudo das emissões de  $CO_2$  e outros gases. Particularmente, concentramo-nos em canais no infravermelho próximo que apresentam certas características desejáveis, tais como um baixo nível de ruído no sinal; alta sensibilidade a mudanças no perfil de concentração; mínima contaminação por linhas de absorção de outros gases como  $H_2O$ ,  $O_3$ ,  $N_2O$ ,  $CO$ ; e baixa sensibilidade à variações no perfil de temperatura (RABIER et al., 2002).

O canal 7 do SCIAMACHY possui 455 linhas espectrais e é sensível a variações na concentração de  $CO_2$ , sendo, portanto, adequado para a reconstrução de perfis verticais deste gás. Por outro lado, o canal 8, com 443 linhas espectrais, é mais apropriado para a recuperação de  $CH_4$ . No entanto, ambos os canais 7 e 8 estão contaminados pelo vapor d'água, que gradualmente se condensa sobre as lentes frias do detector destes canais, depositando, assim, uma camada de gelo cuja espessura cresce lentamente com o tempo. Como o gelo absorve e espalha estes comprimentos de onda isso afeta negativamente a recuperação de concentração de gases traço atmosférico, reduzindo a relação sinal/ruído e modificando a função fenda (ou função resposta) do instrumento ((KLEIPOOL et al., 2005);(BUCHWITZ et al., 2005); (LICHTENBERG et al., 2006)).

Desta forma, embora as linhas espectrais de absorção para os canais 7 e 8 apresentem um sinal mais intenso e estejam em maior número, os resultados apresentados neste trabalho foram derivados a partir de características de absorção do  $CO_2$  e do  $CH_4$  no canal 6 (1000-1750 nm), com 507 linhas espectrais, que não são afetadas pela contaminação por deposição de gelo, pois seus detectores podem operar a temperaturas mais elevadas.

### 6.1.1 Método OSP

Para ser selecionado, uma canal deve apresentar, como foi dito anteriormente, o melhor potencial para detectar mudanças na concentração do gás traço a partir de medidas de radiância realizadas no topo da atmosfera (o "sinal") e tem de apresentar a menor sensibilidade possível a outros gases ou variáveis termodinâmicas da atmosfera (o "ruído"). Desta forma, uma maneira de selecionar canais é estudar suas sensibilidades a diferentes componentes e então selecionar os canais apresentando a maior relação sinal-ruído. Assim, o primeiro critério para a seleção de canais é estudar a relação entre o sinal do gás traço a ser inferido com relação à soma dos sinais de interferência dos outros sinais, de agora em diante referido como a relação RSI (relação sinal-interferência) (CREVOISIER C.; CHEDIN; SCOTT, 2003).

#### 6.1.1.1 Cálculo dos Jacobianos

A resposta de um dado canal à perturbação de um componente atmosférico relevante (gás, características da superfície,...) pode ser obtida pelo produto, camada de pressão por camada de pressão, do componente Jacobiano e do perfil de perturbação. Tanto no método OSP quanto para o método da variância utilizou-se para a identificação dos canais SCIAMACHY sensíveis aos componentes  $CO_2$  e  $CH_4$  uma adaptação do método da perturbação, proposto por Schaerer e Wilheit (SCHAERER G.; WILHEIT, 1979), também conhecido na literatura como Método do Jacobiano (AIRES et al., 2001). Basicamente, este método consiste em perturbar o perfil do constituinte atmosférico que se deseja estudar e em seguida simular o espectro de radiâncias com o modelo direto. Neste caso, procedemos ao cálculo das derivadas das radiâncias (R) com relação à concentração do gás (C) em um determinado nível de altitude z dado por  $\Delta R(\lambda)/\Delta C(z)$ . Posteriormente, faz-se uma comparação do espectro de radiância do perfil inicial com o do perfil perturbado, para diferentes níveis de pressão/altitude atmosférica, a fim de verificar quais são os canais mais sensíveis às alterações introduzidas. No nosso caso as perturbações foram feitas para os 50 níveis de pressão atmosférica presentes no modelo climatológico do SCIATRAN. Nas figuras 6.2 e 6.3 abaixo, são exibidos os gráficos bidimensionais dos jacobianos calculados para o gás metano e dióxido de carbono.

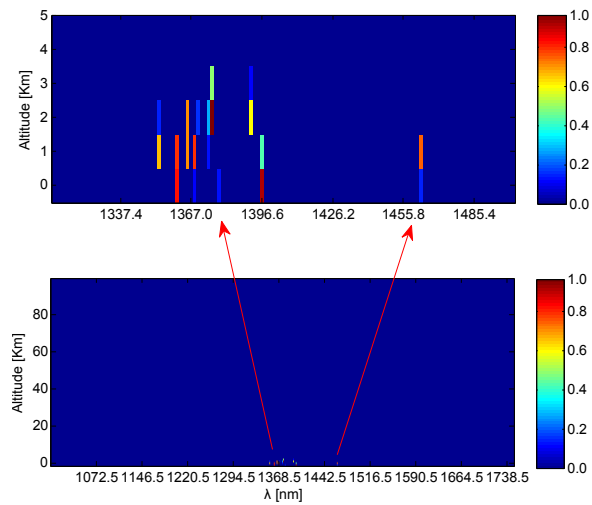


Figura 6.2 - Jacobiano do  $CO_2$  (figura inferior) para o canal 6 do SCIAMACHY e ampliação da região espectral com maior variação observada (topo)

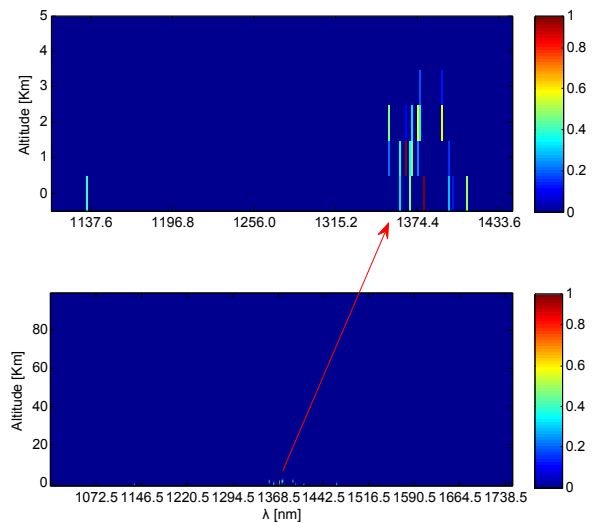


Figura 6.3 - Jacobiano do  $CH_4$  (figura inferior) para o canal 6 do SCIAMACHY e ampliação da região espectral com maior variação observada (topo)

Resumindo, três critérios são usados para selecionar canais: a razão RSI, o sinal devido a emissões de gás traço e a altitude máxima dos Jacobianos. Para cada camada de uma determinada massa de ar, os canais para os quais o Jacobiano do gás ( $CO_2$  ou  $CH_4$ ) é máximo são selecionados primeiro. Então, a razão RSI é calculada, e os canais para os quais tanto o sinal devido ao gás quanto a razão RSI são demasiado baixos são eliminados e, em seguida, os canais restantes são ordenados,



de tal modo que o primeiro apresente o maior RSI. Finalmente, para cada camada, canais com uma RSI dez vezes menor que o RSI do primeiro canal selecionado na camada são rejeitados. Aplicamos o método para cada uma das massas de ar consideradas (tropical, temperada e subártica) sendo que para cada uma delas, a ordem de seleção dos canais é naturalmente diferente. Para se chegar a um único conjunto de canais, um peso é dado a todos os canais para cada massa de ar. O peso máximo sendo igual a  $n$ , onde  $n$  é o número total de canais, é atribuído ao primeiro canal selecionado em cada camada. Para o segundo canal selecionado, o peso é igual a  $n - 1$ , e assim por diante. Então, para cada camada os canais estarão ordenados pela média dos seus pesos.

### 6.1.1.2 Análise de Variância

O método OSP, na forma como foi exposto na seção anterior, poderia ser complementado à luz de uma técnica simples, porém inovadora, que idealizamos para a seleção dos canais. Propomos, como uma abordagem alternativa, uma análise ao longo das colunas da matriz Jacobiana  $J_{ref}$ . Entretanto, nesse caso, estamos interessados em quantificar a dispersão de medidas de um determinado canal classificando-os de acordo com a sua variância. Os canais mais heterogêneos serão elencados e escolhidos mediante a aplicação de um limiar pré-determinado. A fórmula utilizada para o cálculo da variância dos canais é a da variância amostral e pode ser expressa por

$$\sigma_i^2 = \frac{1}{n-1} \sum_{i=1}^n (J_{ij}^{ref} - \overline{J_i^{ref}})^2 \quad (6.1)$$

onde  $n$  representa o número de canais amostrados e  $J_{ij}^{ref}$  representa o vetor com os jacobianos para o  $i$ -ésimo canal do  $j$ -ésimo nível de altitude/pressão do gás. Após selecionarmos os canais de maior variância temos de aplicar um raciocínio similar ao aplicado no método OSP com relação à eliminação dos canais interferentes. O procedimento é descrito pelo algoritmo seguinte:

Agora é só eliminarmos os canais com  $\Delta_i \geq r$  (Em nosso estudo adotamos  $r = 0.01$  ou 1%). Dessa forma ficaremos com um conjunto reduzido de canais cujo processo de escolha foi todo construído a partir do conceito de variância intercanais e do pressuposto de que canais cujo jacobiano com relação à determinado gás traço apresentam maior variância serão também os canais mais sensíveis à variações na concentração da coluna de gás.

- 1.1 Seja  $M_{n \times 1}^{ref}$  o vetor com as variâncias  $\sigma_i^2$  ( $i = 1, \dots, n$ ) calculadas para cada canal do gás traço a ser recuperado e  $M_{n \times 1}^{int}$  um vetor que irá armazenar a somatória das variâncias dos gases interferentes;
- 1.2 **para cada canal  $i$  do gás faça**
- 1.3     **para cada gás traço  $G$  interferente faça**
- 1.4         Calcular a matriz de Jacobianos  $J_G$  para  $G$ ;
- 1.5         Aplicar a equação 6.1 e calcular a variância  $[\sigma_i^G]^2$  para os Jacobianos de cada canal de  $G$ ;
- 1.6          $M_i^{int} \leftarrow M_i^{int} + [\sigma_i^G]^2$ ;
- 1.7     **fim**
- 1.8     Gerar a matriz de Jacobianos  $J_T$  para a Temperatura  $T$  (alterando-se o perfil de temperatura em 10K por ex);
- 1.9     Aplicar a equação 6.1 e calcular a variância  $[\sigma_i^T]^2$  para os Jacobianos de cada canal de  $J_T$ ;
- 1.10      $M_i^{int} \leftarrow M_i^{int} + [\sigma_i^T]^2$ ;
- 1.11     Calcular a matriz de Jacobianos  $J_W$  para o vapor d'água  $W$  (alterando-se o perfil de vapor d'água em 10% por ex);
- 1.12     Aplicar a equação 6.1 e calcular a variância  $[\sigma_i^W]^2$  para os Jacobianos de cada canal de  $W$ ;
- 1.13      $M_i^{int} \leftarrow M_i^{int} + [\sigma_i^W]^2$ ;
- 1.14      $\Delta_i \leftarrow \frac{M_i^{ref}}{M_i^{int}}$ ;
- 1.15 **fim**
- 1.16 **Retorne  $\Delta$**

**Algoritmo 1:** Algoritmo gerado para a eliminação de canais contaminados por algum tipo de interferência ( Ex: sobreposição de linhas de absorção, alterações no perfil de Temperatura e vapor d'água).

Ressaltamos que esta estratégia de seleção não foi aplicada a este trabalho e, pelo que consta, ainda não existem pesquisas que envolvam a sua utilização como método de seleção de canais. Sendo assim, a sua eficácia e eficiência necessitam ser comprovadas em trabalhos futuros.

### **6.1.2 Seleção de canais baseada no Conteúdo de Informação (CI) e nos Graus de Liberdade do Sinal (GLS)**

Outra forma de selecionarmos canais com o objetivo de recuperar a concentração de gases traço a partir de dados de satélite é baseada no estudo da informação contida nas observações. Duas quantidades são de interesse: O conteúdo de informação (CI) dos canais (no sentido de Shannon) e o número de graus de liberdade para o sinal (GLS) do gás.

O conteúdo de informação de uma medida, um conceito desenvolvido no contexto da teoria da informação por Shannon (SHANNON, 1948), pode ser definido pela quantidade de bits com que o conhecimento de uma determinada quantidade ou estado é melhorado realizando-se uma medição. É uma generalização do conceito de relação sinal-ruído. Já o grau de liberdade para um sinal consiste em uma medida do número de quantidades independentes úteis que existem na medida. Ambos os métodos são descritos a seguir. Uma descrição mais detalhada pode ser encontrada em (RODGERS, 1996).

Primeiro, empregamos a análise baseada no conteúdo de informação (CI) para cada canal afim de determinarmos os graus de liberdade (GLS) e conteúdo de informação (CI) para o metano e para o dióxido de carbono. Em seguida, os canais (em cada banda) são classificados em ordem decrescente de CI. Verificamos que os canais com alto CI de  $CO_2$  e  $CH_4$  são aqueles com absorção intermediária. Isto ocorre porque, para canais muito fracos, há pouca absorção do gás traço que consiga nos fornecer um sinal útil, enquanto que para os canais saturados, a absorção é demasiado alta para ter qualquer sensibilidade à concentração do gás. Nós aplicamos o mesmo procedimento para outros parâmetros, como temperatura e vapor de água. Em seguida, classificamos os canais de forma semelhante. Na banda espectral de absorção do canal 6 os canais com alto CI para os gases traço considerados são na sua maioria diferentes daqueles para a temperatura e vapor de água, o que facilita o processo de seleção de canais.

### 6.1.2.1 Conteúdo de Informação

Aplicamos a análise de conteúdo de informação para escolher os canais que contenham a maior parte do conteúdo de informação para o  $CO_2$  e  $CH_4$  e que sejam, ao mesmo tempo, insensíveis a outros parâmetros, tais como temperatura, vapor de água e pressão na superfície. Uma descrição mais rigorosa do ponto de vista matemático-formal do método empregado para o cálculo do CI é transcrita nos parágrafos seguintes.

Vamos considerar os problemas em que o vetor de medidas  $y$ , de tamanho  $m$ , está relacionado ao vetor de estado  $x$ , de tamanho  $n$ , por uma matriz  $m \times n$  denominada "Função Peso" (ou Jacobiano)  $H$  e um vetor de erro randômico  $\epsilon = \epsilon_m + \epsilon_F$ :

$$y - y_0 = H(x - x_0) + \epsilon_m + \epsilon_F \quad (6.2)$$

onde  $\epsilon_m$  e  $\epsilon_F$  são respectivamente, o erro da medida e o erro do modelo direto e  $x_0$  e  $y_0$  definem um ponto de linearização arbitrária. Assume-se que esses erros sejam Gaussianos e não-tendenciosos com matrizes de covariância  $S_M$  e  $S_F$ . A matriz de covariância do erro global é denotada  $S_\epsilon = S_M + S_F$ . A teoria da informação será utilizada no contexto de uma abordagem de estimação ótima ou Bayesiana para a recuperação, na qual o conteúdo de informação é preservado. Antes que uma medição seja feita, o conhecimento do vetor de estado pode ser descrito por uma função densidade de probabilidade (f.d.p.)  $P(x)$ , e após a medição pode ser descrito por uma f.d.p. *a posteriori* do estado dada a medida,  $P(x|y)$ . A f.d.p. *a posteriori* está relacionada com a f.d.p. *a priori* pelo teorema de Bayes:

$$P(x|y) = P(y|x)P(x)/P(y) \quad (6.3)$$

onde  $P(y|x)$  é a f.d.p. da medida, se o estado fosse fornecido, e  $P(y)$  é a f.d.p. *a priori* da medida, descrevendo a distribuição da medida conhecida apenas a f.d.p. *a priori* do Estado. Para efeitos de ilustração, vamos considerar todas as f.d.p. como sendo gaussianas, por exemplo:

$$P(x) = \frac{1}{(2\pi)^{n/2}|S_a|^{1/2}} \exp\left\{-\frac{1}{2}(x - x_a)^T S_a^{-1}(x - x_a)\right\} \quad (6.4)$$

onde  $x_a$  é a estimativa *a priori* de  $x$  e  $S_a$  é a sua matriz de covariância *a priori*. Com este modelo, a estimativa *a posteriori* de  $x$  (isto é, "a recuperação"),  $x$ , tem uma f.d.p. gaussiana com covariância dada por (por exemplo ver (RODGERS, 1976)):

$$\hat{S} = S_a - S_a H^T (H S_a K^T + S_\epsilon)^{-1} H S_a \quad (6.5)$$

ou de maneira equivalente

$$\hat{S}^{-1} = H^T S_\epsilon^{-1} H + S_a^{-1} \quad (6.6)$$

onde  $S_\epsilon$  é a covariância do erro de medida.

A informação obtida ao fazer uma medição pode ser descrita por comparar  $P(x|y)$  com  $P(x)$  de alguma forma, ou equivalentemente  $P(y|x)$  com  $P(y)$ . A quantidade adequada para comparar foi definida por Shannon (SHANNON, 1948) como a entropia de uma função densidade de probabilidade:

$$H(P) = - \int P(x) \log[P(x)/M(x)] dx \quad (6.7)$$

onde a integral é sobre todo o espaço de estado, e  $M(x)$  é uma função de medida que podemos considerar como constante. Qualquer base pode ser usada para o logaritmo; Shannon usou a base 2, expressando a entropia em "bits". A entropia pode ser interpretada como o logaritmo do volume de espaço de estado dentro de um contorno particular da f.d.p.  $P(x)$ . O conteúdo de informação de uma medida é a mudança na entropia consequente de se realizar a medida:

$$S = H(P(x)) - (P(x|y)) \quad (6.8)$$

e pode ser interpretado como o logaritmo de uma generalização da relação sinal/ruído, onde o "sinal" é o volume *a priori* do espaço de estado, e o "ruído" é o volume *a posteriori* do espaço de estados no qual o estado se situa. Para uma distribuição Gaussiana com covariância  $S$  pode ser mostrado que  $H(P) = \frac{1}{2} \log|S|$ .

### 6.1.3 Graus de Liberdade do Sinal (GLS)

O conteúdo da informação é claramente uma abordagem útil e que pode ser utilizada para otimizar o desempenho de um sistema de sensoriamento remoto. Porém, esta abordagem apresenta um pequeno problema conceitual: um mesmo valor do

conteúdo da informação pode ser obtido tanto se medirmos um componente escalar do vetor de estado com alta acurácia quanto se realizamos a medida de vários componentes com menor acurácia. Uma acurácia muito alta pode não ser útil para além de algum nível, ao passo que um maior número de graus de liberdade, medida com menor acurácia pode ser preferível. Abaixo demonstramos como podemos encontrar, à luz do formalismo matemático do trabalho de (RODGERS, 1976), os graus de liberdade do sinal para o canal 6 do sensor SCIAMACHY, utilizado em nosso trabalho de pesquisa.

O maior número de graus de liberdade possível é claramente determinado pelo número de elementos no vetor de estado (ou o vetor de medição, se for menor). Formalmente, o número de graus de liberdade medido é o posto (rank) da matriz de função peso  $K$ , mas devido à presença de ruído e a possibilidade de dependência linear entre as linhas de  $K$ , nem todos estes graus de liberdade podem ser medidos sem risco de redundância. Para obtermos uma abordagem conceitual para identificar o número de graus de liberdade úteis, considere transformações lineares tanto do espaço de estado quanto do espaço de medidas, de tal forma que as medições sejam ortogonais e estatisticamente independentes. Em primeiro lugar, transformamos o espaço de estado usando  $x' = S_a^{-\frac{1}{2}}x$ , de tal modo que a matriz de covariância *a priori*  $S'_a$  se torna uma matriz unitária, em seguida, transformamos o espaço de medidas de forma similiar:  $y' = S_\epsilon^{-\frac{1}{2}}y$ , de modo que a matriz de erro de medição também se torne uma matriz unitária. A matriz de covariância *a posteriori*  $\hat{S}'$  neste sistema de coordenadas não será necessariamente diagonal, mas isto pode ser obtido rotacionando a base do espaço de estado para coordenadas determinadas pelos seus autovetores (eixos principais). Isso deixa a matriz de covariância inalterada, porque é uma matriz unitária.

Sem entrar em maiores detalhes da álgebra envolvida, pode ser mostrado que os componentes do estado correspondente aos autovetores de  $\hat{S}'$  são medidos de forma independente e, o fator pelo qual a variação de cada um desses autovetores é reduzida é dada pelo autovalor correspondente,  $\lambda_i$ . Cada um dos autovetores corresponde a um grau de liberdade medido independentemente, com uma relação sinal-ruído  $[(1 - \lambda_i)/\lambda_i]^{\frac{1}{2}}$ . Se o autovalor correspondente é próximo de zero, significa que aquele grau de liberdade foi medido com exatidão, e pode ser designado como um "Grau de liberdade para o Sinal" mas se o autovalor é próximo da unidade significa que o grau de liberdade é medido com imprecisão, e é designado como um "Grau de liberdade

para o ruído”. Os valores intermediários são atribuídos de maneira fracionada. Assim, o número total de graus de liberdade para o ruído é a soma dos autovalores,  $d_n = \sum \lambda_i$ , e o número de graus de liberdade para o sinal,  $d_s$  é  $n - d_n$  onde  $n$  é o tamanho do vetor de estado. Podemos associar uma contribuição de  $-\frac{1}{2}\log\lambda_i$  ao conteúdo de informação, de forma independente para cada um dos vetores. Assim

$$d_s = n - \sum \lambda_i = \text{tr}(I - \hat{S}') \quad (6.9)$$

e

$$H = -\frac{1}{2}\log\lambda_i = -\frac{1}{2}\log\hat{S}' \quad (6.10)$$

Não é necessário realmente calcular os autovalores, desde que precisamos apenas de sua soma e seu produto, que são iguais ao traço e ao determinante respectivamente de  $\hat{S}'$ .

### 6.1.3.1 Descrição Detalhada do Método

Seguindo a abordagem proposta por (RODGERS, 1996), uma abordagem sequencial é utilizada para selecionar os canais. Inicialmente, o canal que maximiza o ganho de informação a partir de um estado de fundo inicial (*a priori*) é selecionado. Um novo Estado, representado por uma nova matriz de covariância obtida a partir desse estado *a priori* e modificado pelo canal escolhido, é então obtido. O próximo passo é selecionar o canal que maximiza o ganho de informação para este novo estado e, em seguida fazer uma iteração do método até que o número desejado de canais seja atingido. Vamos primeiro normalizar a matriz Jacobiana pela matriz de covariância de erro. Isto produzirá a nova matriz

$$K = S_\epsilon^{-\frac{1}{2}}H \quad (6.11)$$

Vamos supor que  $i - 1$  canais já tenham sido selecionados. A matriz de covariância para este passo é denotada  $\hat{S}_{i-1}$ . Seguindo a 6.6, a nova matriz de covariância obtida para o passo  $i$  através da seleção do canal  $j$  será igual

$$\hat{S}_i^{-1} = \hat{S}_{i-1}^{-1} + k_j k_j^T \quad (6.12)$$

onde  $k_j$  é o vetor associado ao  $j$ -ésimo canal, isto é, a  $j$ -ésima linha de  $K$ . Antes de qualquer canal ter sido selecionado,  $\hat{S}_0 = S_a$ . A outra abordagem, através da 6.5

atualiza sequencialmente  $\hat{S}_i$  ao invés de  $\hat{S}_i^{-1}$ :

$$\begin{aligned}\hat{S}_i &= \hat{S}_{i-1} - (\hat{S}_{i-1}k_i)(\hat{S}_{i-1}k_i)^T/[1 + (\hat{S}_{i-1}k_i)^T k_i] \\ &= \hat{S}_{i-1}\{I_n - k_i\hat{S}_{i-1}k_i^T/[1 + (\hat{S}_{i-1}k_i)^T k_i]\}\end{aligned}$$

que parece mais complicada mas é mais conveniente para o cálculo da mudança no traço e no determinante. O ganho no conteúdo de informação ao introduzir o canal  $i$  é dado por

$$\delta S_i = \frac{1}{2} \ln \|\hat{S}_{i-1}\| - \frac{1}{2} \ln \|\hat{S}_i\| \quad (6.13)$$

onde  $\|\cdot\|$  representa o determinante, e os graus de liberdade para o sinal por

$$d_j = \text{tr}(\hat{S}_{i-1}) - \text{tr}(\hat{S}_i) \quad (6.14)$$

Esta equação pode ser simplificada para

$$\delta S_i^j = \frac{1}{2} \ln(1 + k_j^T \hat{S}_{i-1} k_j) \quad (6.15)$$

e

$$d_j = \frac{(\hat{S}_{i-1}k_i)^T (\hat{S}_{i-1}k_i)}{\frac{1}{2} \ln(1 + k_j^T \hat{S}_{i-1} k_j)} \quad (6.16)$$

O canal selecionado neste passo é o canal  $j$  que maximiza  $\delta S_i^j(CI)$  ou  $d_j(GLS)$ ,  $j$  variando sobre todos os canais ainda não selecionados.

## 6.2 Geração da Base de Dados de Treinamento da Rede

Para treinarmos uma rede neural do tipo Perceptron Multi-Camadas (nosso modelo inverso) necessitamos de um conjunto de padrões de entrada e saída para o seu treinamento (ver seção 5.2.2.2 do capítulo 5). Para realizarmos esta tarefa empregamos o modelo direto SCIATRAN (ver seção 4.2). Para melhorarmos a caracterização do perfil do gás recuperado pelo modelo inverso incluímos em sua base de dados climatológica perfis de umidade e temperatura extraídos do banco de dados de perfis conhecido como TIGR (Thermodynamic Initial Guess Retrieval). Este banco agrupa 2311 situações atmosféricas selecionadas por métodos estatísticos sobre mais de 150000 radiosondagens ((CHEDIN et al., 1985);(CHEVALLIER et al., 1998)) e é ar-



mazenado em um único arquivo ASCII, perfil por perfil. Para cada perfil, o arquivo fornece:

- a) número TIGR,
- b) latitude e longitude,
- c) Data da medida do perfil, formato é YYMMDD,
- d) Temperatura nos 40 níveis de pressão (K),
- e) Temperature (K) e pressão (mb) ao nível da superfície,
- f) Concentração de vapor d'água nos 40 níveis de pressão,
- g) Concentração de Ozônio nos 40 níveis de pressão.

Para os experimentos realizados consideramos como variáveis os perfis de umidade, temperatura e do gás traço de interesse, o qual estamos interessados em recuperar. Todos os demais constituintes atmosféricos foram assumidos como conhecidos e constantes sendo que os perfis de temperatura e vapor d'água foram interpolados e, se necessário, extrapolados, sobre 40 níveis de pressão, do solo à 0.05hPa (cerca de 70km).

Parte do trabalho consistiu em reduzir o conjunto de 2311 situações ou perfis atmosféricos a um sub-conjunto contendo 459 perfis ( $\sim 20\%$  do total de perfis) sem perda de conteúdo de informação significativa. Estes perfis foram agrupados em 5 classes: tropical, temperado (inverno), temperado (verão), polar (inverno) e polar (verão), os quais foram suficientes para representar a dinâmica das variáveis atmosféricas consideradas aqui a um menor custo computacional. A distribuição dos perfis selecionados em cada uma dessas classes se deu da seguinte forma: tropical (174 situações), temperado (147) e polar (138). Cada situação é descrita por sua temperatura, vapor d'água e perfis de ozônio, a atmosfera sendo dividida em 50 níveis de pressão, regularmente espaçados. Para reduzirmos a base de perfis ou situações atmosféricas do TIGR realizamos o seguinte procedimento:

- a) Geramos uma matriz  $P_{igr}$  com os perfis de temperatura e umidade presentes na base de dados original onde cada coluna representa um perfil de temperatura concatenado a um perfil de umidade correspondente.

- b) Em seguida calculamos a correlação entre cada um dos perfis (colunas da matriz) através de

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6.17)$$

onde  $x_1, x_2, \dots, x_n$  e  $y_1, y_2, \dots, y_n$  são os valores medidos de ambas as variáveis com médias aritméticas  $\bar{x}$  e  $\bar{y}$ . A análise correlacional indica a relação entre 2 variáveis lineares de modo que os valores para  $r$  sempre estarão no intervalo entre +1 e -1. O valor 1 (-1) indica dois vetores perfeitamente (anti-) correlacionados.

- c) Geramos uma matriz de correlação  $C$ ,  $n \times n$ , com  $n$  igual ao número total de perfis, a partir da aplicação sucessiva da equação 6.17. Calculamos adicionalmente uma matriz de p-valores  $H_{tigr}$  para testar a hipótese de nenhuma correlação. O p-valor fornece uma medida de quanta evidência nós temos contra a hipótese nula. A hipótese nula, tradicionalmente representada pelo símbolo  $H_0$ , representa a hipótese de nenhuma mudança ou nenhum efeito. Cada p-valor corresponde à probabilidade de obter uma correlação tão grande quanto a observada por um evento ao acaso, quando a verdadeira correlação é zero. Se  $H(i, j)$  é pequena, digamos menor que 0.05, então a correlação  $C(i, j)$  é significativa.
- d) A partir da análise dos valores dos coeficientes de correlação de  $C$  e da matriz de p-valores ordenamos os perfis em ordem decrescente de grau de correlação. No passo seguinte agrupamos estes perfis em outra matriz  $P_{tigr}^{ord}$  de tal forma que os primeiros perfis correspondem a perfis fortemente correlacionados. Nas últimas fileiras ficaram os perfis que apresentam grau de correlação mais fraco.
- e) Considerando-se as duas primeiras fileiras  $P_1$  e  $P_2$  de  $P_{tigr}^{ord}$ , ou seja, os dois perfis com maior valor de correlação, escolhemos ao acaso um deles digamos  $P_1$ . Em seguida retiramos  $P_1$  de  $P_{tigr}$  gerando, desta forma, uma nova matriz  $P'_{tigr}$ , subconjunto de  $P_{tigr}$  com uma fileira a menos.
- f) Repetimos os passos anteriores (a-e) para a matriz  $P'_{tigr}$  até obtermos uma matriz  $P_{tigr}^{red}$  de tamanho reduzido  $n < N$  pré-definido contendo apenas os perfis com menor índice de correlação.

O próximo passo, após a seleção dos canais mais promissores e do subconjunto de perfis climatológicos do TIGR, foi a geração de um banco de dados de radiâncias para o sistema SCIAMACHY que serviu como entrada para o nosso modelo inverso. Para cada situação atmosférica e para cada canal SCIAMACHY, computamos a radiância que seria vista pelo sensor, assumindo um cenário livre de nuvens, utilizando um modelo direto de ETR com a opção de permitir que os Jacobianos (derivada parcial da radiância do canal com relação a uma variável física da camada tal como a taxa de mistura do gás, temperatura ou emissividade) sejam calculados analiticamente para todas as 507 linhas espectrais do canal 6 do sensor.

O procedimento para criar o banco de dados de treinamento pode ser descrito pelos seguintes passos: Para um determinado gás, selecionamos a partir do conjunto de dados climatológicos do modelo direto (SCIATRAN), um perfil de concentração de referência  $P_{gás}^{ref} = [C_1, C_2, \dots, C_m]$ , onde  $m$  é o número de níveis verticais. Modificamos o perfil de referência para gerar um conjunto de perfis  $P_{base}^{gás}$  compreendendo novos perfis  $P_j$ , através da relação

$$P_j = P_{gás}^{ref} \times \left[1 + \frac{j}{100}\right] \text{ com } j = -k, -(k-1), \dots, 0, \dots, k-1, k,$$

onde  $k$  é numericamente equivalente ao número de perfis de concentração desejados com concentração inferior ou superior ao do perfil de referência. Para este projeto de pesquisa adotamos  $k = 6$  para o  $CO_2$  e  $k = 11$  para o  $CH_4$ . Dessa maneira, a partir de cada perfil de referência conseguimos construir 13 novos perfis de concentração para o metano e 23 para o dióxido de carbono, respectivamente.

Em seguida, para todo perfil de umidade e temperatura  $P_i^{tigr} \in P_{tigr}^{red}$  e para todo elemento  $P_j \in P_{base}^{gás}$  realizamos o seguinte procedimento: Inserimos  $P_i^{tigr}$  e  $P_j$  no modelo direto  $F$ . Executamos o modelo direto para  $P_j$  produzindo um vetor de radiância  $R_{ij} = [R_1, R_2, \dots, R_n]$ , onde  $n$  é o número de canais selecionados ou comprimentos de onda escolhidos previamente. Acrescentamos esse vetor  $R_{ij}$  à matriz de radiâncias  $R$ .

Na etapa seguinte derivamos  $R_{perturb}$  a partir de  $R$  perturbando estocasticamente  $R$  através de

$$R_{\sigma}^{ij} = R_{ij} \times (1 + \sigma \times \mu) \quad (6.18)$$

onde  $\sigma$  é o desvio padrão do ruído e  $\mu$  é uma variável aleatória retirada de uma distribuição Gaussiana com média zero e variância unitária. Em nossos experimentos utilizamos  $\sigma = 0.01$  (ruído de 1%) o que corresponde à uma boa estimativa do erro de medida obtido pelo sensor. Esta estimativa de erro em nosso modelo corresponde ao erro do modelo direto (SCIATRAN) na estimativa do sinal. Por sua vez o ruído de 1% corresponde ao pior cenário possível para a relação sinal-ruído do sensor SCIAMACHY para o canal 6 no modo nadir (ver Figura 7.3 e 7.4 da seção 7.1.1.1 do capítulo 7). Portanto, armazenamos em  $R_{perturb}$  o conjunto de radiâncias associada à cada perfil de umidade, temperatura e concentração do gás que irá compor a base de dados de padrões de treinamento de nosso modelo inverso. Procedemos desta forma para obtermos um conjunto de dados mais representativo do estado real da atmosfera. Neste estudo, foram considerados diferentes perfis de referência  $P_{gas}^{ref}$  representando o clima em regiões temperadas (latitudes médias), regiões polares (ártico e sub-ártico) e regiões equatoriais do globo para as duas estações bem definidas de inverno e verão. O produto final desse processo é uma base de dados contendo 5967 padrões de treinamento para o  $CO_2$  e 10557 para o  $CH_4$ . Finalmente, esse conjunto de dados é dividido em três subgrupos seguindo o que foi sugerido em Haykin (HAYKIN, 2001):

- a) Conjunto de treinamento, com 4177 amostras para o  $CO_2$  e 7390 amostras para o  $CH_4$ , incluindo os padrões de treinamento utilizados durante a fase de treinamento da RNA ( $\sim 70\%$  do total de amostras).
- b) Conjunto de validação, com 1491 amostras para o  $CO_2$  e 2639 amostras para o  $CH_4$ , utilizado na validação (PERRONE, 1993) para avaliar o desempenho da RNA ( $\sim 25\%$  do total de amostras).
- c) Conjunto de Generalização, com 299 amostras para o  $CO_2$  e 528 amostras para o  $CH_4$ , empregado para o teste efetivo da RNA, conhecido como teste de generalização ( $\sim 5\%$  do total de amostras).

### 6.3 Otimização dos Parâmetros da Rede

Neste seção apresentamos uma maneira sistemática e automatizada de estabelecer uma arquitetura de RNA otimizada para o problema da estimação de perfil vertical de concentração de gases traço. Para tanto adaptamos quatro metaheurísticas de busca para encontrarmos uma arquitetura de RNA do tipo *feedforward* que satisfaça

algum critério de otimalidade. A inovação ficou por conta da introdução de um novo critério para quantificar a performance de uma RNA baseado na combinação de erro de treinamento e generalização, e o emprego de heurísticas que nunca foram utilizadas neste tipo de aplicação, tais como o Generalized Extremal Optimization (GEO) e o *Variable Neighbourhood Search* (VNS) (ver seções 6.3.4 e 6.3.3).

Adicionalmente, propusemos uma nova metodologia para estimação da complexidade computacional da arquitetura de uma RNA baseada no número de neurônios e no número de épocas necessários para o treinamento dos padrões apresentados à rede. Aplicou-se o esquema proposto ao problema da recuperação de perfis atmosféricos de concentração de gases de efeito estufa o qual permitiu a obtenção de uma arquitetura de rede otimizada. Os resultados mostraram que a abordagem computacional proposta apresenta melhor performance em relação a um especialista humano, oferecendo ao mesmo tempo muitas vantagens se comparada a abordagens similares encontradas na literatura.

Redes neurais artificiais tem sido um tópico de interesse nos institutos de pesquisa do mundo todo. Embora muito tenha sido proposto e estudado sobre elas ainda existem muitas questões que dizem respeito ao modelo de RNAs que necessitam ser melhor compreendidas. Não existe uma garantia prévia de que modelo adotado será ou não bem sucedido na tarefa para o qual ele foi designado. Muito tem sido estudado e pesquisado na tentativa de se encontrar uma maneira sistemática de ajustar um modelo de rede para um problema em particular objetivando otimizar a sua performance global. Estas pesquisas são baseadas em quase todos os aspectos do modelamento de RNAs tais como diferentes tipos de função de ativação, a inicialização de pesos, a coleção de dados de treinamento, o pré e pós processamento, os algoritmos de treinamento e as funções de erro. Contudo, baseado em evidências empíricas e teóricas encontradas na literatura, enfatizamos a busca e definição de uma arquitetura ótima de RNA como o fator mais importante quando priorizamos a performance e a eficiência computacional do modelo (BERNARDOS P. G.; VOSNIAKOS, 2004).

Podemos então formular o problema correspondente a encontrar uma arquitetura de rede ótima ou quase-ótima como uma busca em um espaço de arquiteturas onde cada ponto representa uma possível arquitetura. Se associarmos a cada ponto ou solução (arquitetura) um nível de performance baseado em algum critério de otimalidade tal como complexidade, mínimo erro quadrático ou velocidade de treinamento,

podemos construir e exibir uma superfície em um espaço tri-dimensional. Encontrar uma arquitetura ótima nesse caso seria equivalente encontrar o ponto mais alto (ou mais baixo) desta superfície, isto é, seus pontos extremos globais dependendo do critério utilizado (YAO X.; LIU, 1997). Diante desse cenário, técnicas das mais variadas têm sido usadas para otimizar RNAs com o intuito de tirar do projetista essa responsabilidade, eliminando assim os erros decorrentes da má escolha de uma arquitetura de rede.

Neste trabalho, aplicamos um grupo de metaheurísticas para resolver este problema. Não há nenhum registro na literatura de qualquer tentativa de usar metaheurísticas de busca que implementam o GEO e VNS como um método de otimização para o problema da melhor arquitetura de rede. Além disso, definimos a complexidade computacional de uma arquitetura de RNA do tipo *feedforward* como sendo uma função do número total de pesos e bias presentes em suas estruturas e do tempo necessário para a aprendizagem da rede. A partir disto derivamos um termo de penalidade usado para avaliar a função objetivo de modo a evitar arquiteturas de rede muito complexas. O uso combinado destes elementos é também uma contribuição deste trabalho.

Outra contribuição foi a inclusão de dois fatores de ponderação ( $\rho_1$  e  $\rho_2$ , ver a seção 6.3.9), que medem a influência dos erros de generalização e treinamento no processo de busca algorítmica, permitindo-nos ajustar adequadamente o andamento da otimização.

### 6.3.1 Computação Evolucionária

O uso de RNAs na solução de determinados problemas não se mostra uma tarefa fácil. O projeto de uma RNA é um trabalho empírico em que vários dos parâmetros ajustáveis que a constituem devem ser estimados pelo projetista da rede neural. Não existe uma indicação clara e inequívoca de como podemos encontrar a melhor arquitetura entre as várias opções apresentadas. Além disso, ainda não sabemos de antemão a correta topologia de rede a ser aplicada. Na prática, este problema geralmente é resolvido (em parte), utilizando técnicas empíricas baseadas em métodos repetitivos de tentativa e erro. Neste caso, poderíamos dizer que o sucesso deste método depende quase única e exclusivamente de experiências anteriores e da intuição do especialista humano por trás da cena, envolvendo, portanto, um elevado grau de incerteza e subjetividade. Nestes métodos, utiliza-se muito tempo da CPU em sim-

ulações desnecessárias e que não abrangem a totalidade das opções de arquiteturas disponíveis para um determinado problema. Além disso, quase sempre encontramos uma solução subótima que corresponde a uma solução inferior se comparada com a solução que poderia ser obtida através de um método de busca autônoma.

É, portanto, uma tarefa desgastante projetar e construir um RNA para um problema específico. Segundo Jeffrey et al. (JEFFREY J. L.; VITTER, 1991), a busca por essa arquitetura "ótima" é um problema NP-difícil, ou seja, possui ordem de complexidade exponencial. Em outras palavras, o esforço computacional para a sua resolução cresce exponencialmente com o tamanho do problema. Existem alguns poucos princípios de projetos rigorosos que estão disponíveis para definir as RNAs e, no entanto, muitos parâmetros para se ajustar. Os métodos construtivos são uma alternativa para a concepção de redes neurais artificiais, mas o custo computacional necessário para certas aplicações pode ser proibitivo, além da dificuldade apresentada no tratamento de espaços de entrada de dimensão elevada. Portanto, o objetivo aqui é considerar técnicas avançadas de computação evolutiva, que também promovam a obtenção de arquiteturas de RNA dedicadas, específicas para cada problema.

A computação evolutiva (CE) é um ramo da ciência da computação que propõe um paradigma alternativo ao processamento de dados convencional. Este novo paradigma, diferentemente do convencional, não exige, para resolver um problema, o conhecimento prévio de uma maneira de encontrar uma solução. A CE é baseada em mecanismos evolutivos encontrados na natureza, tais como a auto-organização e o comportamento adaptativo (FARMER J. D.; TOFFOLI; WOLFRAM, 1983), (GOLDBERG D.E.; HOLLAND, 1988). Estes mecanismos foram descobertos e formalizados por Darwin e Wallace e condensados posteriormente no que é chamado de *teoria da evolução natural de Darwin-Wallace*. Esta teoria preconiza que o desenvolvimento e (des-)aparecimento das espécies na terra seria o resultado de um processo de seleção, pelo meio ambiente, dos mais aptos e adaptados, e por isto mesmo com mais chances de gerar maior número de descendentes (DARWIN C. R.; WALLACE, 1858). A diversidade da vida, associada ao fato de que todos os seres vivos compartilham uma bagagem genética comum, pelo menos em termos de seus componentes básicos, é um exemplo eloqüente das possibilidades do mecanismo de evolução natural.

Podemos dividir a computação evolutiva e as propostas de combinação com redes neurais em dois grupos principais: computação evolutiva para treinar RNA ou para ajustar os seus pesos e computação evolutiva para definição de arquiteturas de rede

(IYODA, 2000). De um modo semelhante (MILLER G. F.; TODD; HEDGE, 1991) identificou duas abordagens para codificar a arquitetura de RNA em uma seqüência binária ou em uma representação decimal. Uma delas é o esquema de especificação forte (ou esquema de codificação direta), onde a arquitetura de uma rede é explicitamente codificada. O outro é um esquema de especificação fraco (ou esquema de codificação indireto), onde o padrão de conectividade não é explicitamente representado. Podemos incluir como exemplos da aplicação do esquema de especificação forte (WHITLEY D.; STARKWEATHER; BOGART, 1989), (SCHAFER J. D.; CARUANA; ESHELMAN, 1990), (MILLER G. F.; TODD; HEDGE, 1991), (MENCZER F.; PARISI, 1992) e (ARIFOVIC J.; GENCAJ, 2001). Por outro lado, as aplicações do esquema de especificação fraca incluem (HARP S.; SAMAD; GUHA, 1989), (YAO X.; LIU, 1997), (CHEN Z.; XIAO; CHENG, 1997), (KITANO, 1994), (KITANO, 1990) e (BERNARDOS P. G.; VOSNIAKOS, 2004).

Uma extensa comparação entre os diversos métodos de codificação podem ser encontrados em (KOEHN, 1994). A nossa abordagem pertence a esta última categoria, uma vez que utiliza um esquema de codificação indireta para encontrar a melhor arquitetura capaz de resolver um problema em particular. Porém, ao contrário da maioria das outras abordagens, além de empregarmos uma estratégia puramente evolutiva tal qual algoritmos genéticos (AG) acrescentamos alguns outros métodos combinatoriais alternativos tais como Simular Annealing (SA) ou Busca de Vizinhança Variável (VNS) que foram muito bem sucedidos na resolução de problemas com muitos parâmetros e que apresentam uma clara tendência exponencial no número de soluções viáveis geradas. É importante ressaltar que outras alternativas igualmente interessantes poderiam ser utilizadas para o problema em questão. Modelos de otimização baseados no método das colônias de formiga foram utilizados com sucesso para reconstruir o perfil vertical de concentração de clorofila em águas oceânicas ((SOUTO et al., 2008) e (CARVALHO et al., 2008)). Luz et al. se baseou num modelo de otimização por enxame de partículas para estimar fontes de poluição atmosférica (LUZ et al., 2007). O algoritmo baseado na colisão de partículas é empregado em um ambiente de processamento de alto desempenho por Luz et al. também resultando em boas soluções para problemas de otimização. Cuco et al. aplica uma versão modificado do algoritmo genético combinada com o GEO para resolver um problema inverso de adsorção molecular (CUCO et al., 2009). Esses são exemplos de aplicação de alguns dos muitos algoritmos com bom potencial para serem utilizados nesse tipo de problema.



### 6.3.2 Método Proposto

Foi desenvolvida uma abordagem para otimizar arquiteturas de RNA do tipo Perceptron Multi-Camadas com alimentação para frente (*feedforward*) baseadas na utilização de metaheurísticas de busca global. Em nosso trabalho adaptamos quatro metaheurísticas para a resolução do problema representado pela busca ótima de uma arquitetura de RNA. Abaixo descrevemos sucintamente cada uma dessas metaheurísticas.

### 6.3.3 Variable Neighbourhood Search (VNS)

O primeiro algoritmo utilizado foi o *Variable Neighbourhood Search* (VNS). O algoritmo VNS é uma meta-heurística recente usada para resolver problema combinatoriais complexos. Ele foi proposto e descrito em trabalhos de (HANSEN P.; MLADENOVIC, 2003), (HANSEN P.; MLADENOVIC, 1995) e (MLADENOVIC N.; HANSEN, 1997). VNS é uma técnica que tenta escapar de um ótimo local modificando a estrutura de vizinhança de uma maneira sistemática. O VNS básico obtém uma solução vizinha à solução corrente, executa uma busca local até que um ótimo local seja alcançado e se move para ele se existir uma melhora em modificar a estrutura de vizinhança. Caso contrário a vizinhança é modificada sistematicamente.

Contrariamente a outras metaheurísticas, o método VNS não segue uma trajetória, mas sim explora vizinhanças gradativamente mais "distantes" da solução corrente e focaliza a busca em torno de uma nova solução, se e somente se um movimento de melhora é realizado.

Mais especificamente, essa heurística parte de uma solução inicial qualquer e, a cada iteração, seleciona aleatoriamente um vizinho dentro da vizinhança  $N_k$  da solução corrente. Esse vizinho é então submetido a um procedimento de busca local. Se a solução ótima local for melhor que a solução corrente, a busca continua a partir desta, recomeçando da primeira estrutura de vizinhança. Caso contrário, a busca prossegue a partir da próxima vizinhança,  $N_{k+1}$ . Esta heurística é encerrada quando uma condição de parada for atingida, tal como o tempo máximo de processamento ou o número máximo de iterações consecutivas sem melhoramento. Os vizinhos da solução corrente são gerados aleatoriamente de forma a evitar ciclagem, situação que pode ocorrer se alguma regra determinística for usada. A seguir apresentamos um pseudo-algoritmo com as etapas correspondentes à versão do algoritmo VNS

aplicado ao problema.

```
2.1 Seja  $s_0$  uma solução inicial e  $r_0$  o número de estruturas de vizinhança
2.2  $s \leftarrow s_0$ ; {Solução corrente}
2.3 enquanto critério de parada não satisfeito faça
2.4    $k \leftarrow 1$ ; {Tipo de estrutura de vizinhança}
2.5   enquanto  $k \leq r$  faça
2.6     Gere um vizinho qualquer  $s' \in N^{(k)}(s)$ 
2.7      $s'' \leftarrow BuscaLocal(s')$ 
2.8     se  $(f(s'') < f(s))$  então
2.9        $s \leftarrow s''$ 
2.10       $k \leftarrow 1$ 
2.11     senão
2.12        $k \leftarrow k + 1$ 
2.13     fim
2.14   fim
2.15 fim
2.16 Retorne  $s$ 
```

**Algoritmo 2:** *Variable Neighbourhood Search (VNS)*

### 6.3.4 Otimização extrema generalizada (GEO)

Uma outra metaheurística implementada nesse trabalho foi o método da Otimização Extrema Generalizada que é um uma meta-heurística de busca global ((SOUSA F. L.; RAMOS, 2002),(SOUSA et al., 2003),(SOUSA et al., 2004)), baseada em um modelo de evolução natural (BAK P.; SNEPPEN, 1993), e especialmente criado para ser usado em problemas de otimização complexa (SOUSA F. L.; VLASSOV; RAMOS, 2003). Ele tem seus fundamentos na teoria da criticalidade auto-organizada, que vem sendo utilizada para explicar as assinaturas de leis de potência que emergem de muitos sistemas complexos (BAK, 1996). O algoritmo é um método estocástico, assim como o SA e o AG, não faz uso de derivadas e pode ser aplicado a problemas não-convexos ou disjuntos. Ele pode também lidar com qualquer tipo de variável, seja contínua, discreta ou inteira. O algoritmo possui apenas um parâmetro livre ( $\tau$ ) que nos permite ajustar o grau de determinismo da busca. A busca pode partir de uma marcha aleatória conhecido como *caminho aleatório* ( $\tau = 1$ ) até uma busca totalmente determinística ( $\tau \rightarrow \infty$ ). Na prática o que tem sido observado é que existe um valor de  $\tau$  ótimo para cada problema mas que, em geral, fica restrito ao intervalo [1, 5]. As variáveis do modelo são codificadas em strings binárias, à semelhança do que ocorre em algoritmos genéticos simples (AGS). Contudo, GEO associa um número que caracteriza um índice de adaptabilidade à cada bit da string, também chamado de espécie, ao invés de associá-lo à string toda, como ocorre no AGS. O algoritmo modifica os bits na tentativa de encontrar a configuração que fornece o menor valor para a função objetivo. Apresentamos abaixo um pseudo-algoritmo no qual podemos observar em detalhes todas as etapas de sua implementação.

```

3.1 inicialize aleatoriamente uma sequência binária de comprimento  $L$  que
    codifica  $N$  variáveis de projeto em sub-sequências de comprimento
     $l_j (j = 1, \dots, N)$ 
3.2 repita
3.3   para uma configuração inicial de bits  $C$  faça
3.4     calcule o valor da função objetivo  $V$ 
3.5      $C_{melhor} \leftarrow C$ 
3.6      $V_{melhor} \leftarrow V$ 
3.7   fim
3.8   para cada cada bit  $i$  da sequência faça
3.9     mude o valor do bit (de 0 para 1 ou 1 para 0)
3.10    calcule o valor da função objetivo  $V_i$ , da configuração de bits  $C_i$ 
3.11     $\Delta V \leftarrow V_i - V_{melhor}$ 
3.12    {Atribui ao bit um índice de adaptabilidade que indica o ganho (ou perda) que se
    têm ao mudar o valor do bit, comparado com o melhor valor encontrado para a
    função objetivo até o momento}
3.13    retorne o bit ao seu valor original
3.14  fim
3.15  ordene os bits de cada variável de  $k = 1$ , para o menos adaptado desta
    variável, até  $k = l_j$ , para o mais adaptado, onde  $l_j$  é o número de bits de
    cada variável  $j$ 
3.16  {Em um problema de minimização valores altos de  $\Delta V_i$  terão maior “rank”, enquanto
    que em problemas de maximização ocorre o oposto. Se ocorrer de dois ou mais bits
    apresentarem o mesmo valor para  $\Delta V_i$  eles são ordenados aleatoriamente com
    distribuição uniforme}
3.17  para cada variável  $j$  da sequência  $L$  faça em paralelo
3.18    escolha com igual probabilidade um bit candidato  $i$  para sofrer
    mutação (mudar de 0 para 1 ou de 1 para 0)
3.19    gere um número aleatório  $RAN$ , com distribuição uniforme, no
    intervalo  $[0, 1]$ 
3.20     $mudou \leftarrow false$ 
3.21    enquanto  $mudou \neq true$  faça
3.22      se  $(P_i(k) \leftarrow k^{-\tau}) \geq RAN$  então
3.23        Inverte bit
3.24         $mudou \leftarrow true$ 
3.25      fim
3.26    fim
3.27  fim
3.28  para cada bit  $i$  escolhido para sofrer mutação faça
3.29     $C \leftarrow C_i$ 
3.30     $V \leftarrow V_i$ 
3.31  fim
3.32  se  $V < V_{melhor}$  então
3.33    {para um problema de maximização utiliza-se  $V > V_{melhor}$  }
3.34     $V_{melhor} \leftarrow V$ 
3.35     $C_{melhor} \leftarrow C$ 
3.36  fim
3.37 até Critério de parada satisfeito

```

### 6.3.5 Recozimento Simulado

O algoritmo de recozimento simulado (ou simulated annealing) foi a terceira meta-heurística implementada e adaptada neste trabalho para abordar o problema da descoberta da arquitetura de RNA ótima. O algoritmo de recozimento simulado (SA) é um método de busca local motivado por uma analogia com recozimento em sólidos. A idéia de SA vem de um artigo publicado por Metropolis et al. em 1953 (METROPOLIS et al., 1953). O algoritmo neste artigo simulou o resfriamento de material aquecido através da imersão em um líquido. Este é um processo conhecido como recozimento. Se você aquecer um sólido além do seu ponto de fusão e, em seguida, resfriá-lo, as propriedades estruturais dos sólidos dependerão da taxa de resfriamento. Se o líquido é resfriado de forma lenta o suficiente, grandes cristais se formarão. No entanto, se o líquido é resfriado rapidamente (por exemplo introduzindo o metal quente em água) os cristais irão conter imperfeições. O algoritmo de Metropolis, também conhecido por Algoritmo de Metropolis-Hastings, simula os materiais como um sistema de partículas. O algoritmo simula o processo de resfriamento diminuindo gradualmente a temperatura do sistema até que ela convirja para um estado estável e congelado. Em 1982, Kirkpatrick et al. (KIRKPATRICK S.; GELATT JR.; VECCHI, 1983) usou a ideia do algoritmo de Metropolis e aplicou-a a problemas de otimização. A ideia é usar o recozimento simulado para buscar soluções viáveis e convergir para uma solução ótima.

O algoritmo começa sua busca a partir de uma solução inicial qualquer. O procedimento principal consiste em um laço de repetição que gera aleatoriamente, em cada iteração, um único vizinho  $s'$  da solução corrente  $s$ . A cada geração de um vizinho  $s'$  de  $s$ , é testada a variação  $\Delta$  ( $\Delta = f(s') - f(s)$ ) do valor da função objetivo. Se  $\Delta < 0$ , o método aceita a solução e  $s'$  passa a ser a nova solução corrente. Caso  $\Delta \geq 0$  a solução vizinha candidata também poderia ser aceita, mas neste caso, com uma probabilidade  $e^{-\frac{\Delta}{T}}$ , onde  $T$  é um parâmetro do método, chamado de temperatura e que regula a probabilidade de aceitação de soluções com custo pior.

A temperatura  $T$  assume, inicialmente, um valor elevado. Após um número fixo de iterações (o qual representa o número de iterações necessárias para o sistema atingir o equilíbrio térmico em uma dada temperatura), a temperatura é gradativamente diminuída por uma razão de resfriamento  $\zeta$ , tal que  $T_n = \zeta \times T_{n-1}$ , sendo  $0 < \zeta < 1$ . Com esse procedimento, dá-se, no início uma chance maior para escapar de ótimos locais e, à medida que  $T$  se aproxima de zero, o algoritmo comporta-se como o *Hill*

*Climbing*, uma vez que diminui a probabilidade de se aceitar movimentos de piora ( $T \rightarrow 0 \Rightarrow e^{\frac{-\Delta}{T}} \rightarrow 0$ ).

O procedimento pára quando a temperatura chega a um valor próximo de zero e nenhuma solução que piore o valor da melhor solução é mais aceita, isto é, quando o sistema está estável. A solução obtida quando o sistema encontra-se nesta situação evidencia o encontro de um ótimo local. Os parâmetros de controle do procedimento são a razão de resfriamento  $\zeta$ , o número de iterações para cada temperatura ( $SA_{max}$ ) e a temperatura inicial  $T_0$ . Algoritmos baseados em recozimento simulado geralmente incluem reaquecimento seguido de um novo processo de resfriamento, utilizado quando a quantidade de movimentos consecutivamente rejeitados é alta. É comum também trabalhar nas temperaturas mais altas com taxa de resfriamento menor e aumentá-la quando a temperatura reduzir. Abaixo apresentamos um esboço correspondente ao algoritmo de recozimento simulado básico utilizado nesse trabalho.

```

3.1 procedimento SA ( $f(\cdot), N(\cdot), \zeta, SA_{max}, T_{min}, T_0, s$ )
4.2  $s^* \leftarrow s$ ; {Melhor solução obtida até então}
4.3  $iterT \leftarrow 0$ ; {número de iterações na temperatura  $T$ }
4.4  $T \leftarrow T_0$ ; {Temperatura corrente}
4.5 enquanto  $T > T_{min}$  faça
4.6     enquanto  $iterT < SA_{max}$  faça
4.7          $iterT \leftarrow iterT + 1$ 
4.8         Gere um vizinho ( $s'$ ) aleatoriamente na vizinhança  $N^k(s)$ 
4.9          $\Delta \leftarrow f(s') - f(s)$ 
4.10        se  $\Delta < 0$  então
4.11             $s \leftarrow s'$ 
4.12            se  $f(s') < f(s^*)$  então
4.13                 $s^* \leftarrow s'$ 
4.14            fim
4.15        senão
4.16            Tome  $x \in [0, 1]$ 
4.17            se  $x < e^{\frac{-\Delta}{T}}$  então
4.18                 $s \leftarrow s'$ 
4.19            fim
4.20        fim
4.21    fim
4.22     $T \leftarrow T \times \zeta$ 
4.23     $iterT \leftarrow 0$ 
4.24 fim
4.25 Retorne  $s^*$ 

```

Algoritmo 4: Recozimento Simulado

### 6.3.6 Algoritmos Genéticos

Por último, construímos e adaptamos uma versão simplificada do algoritmo genético em sua forma canônica sem grandes modificações ou mudanças em sua ideia original para resolver o problema de recuperação proposto. A Computação Evolucionária é uma abordagem que se inspira em mecanismos evolucionários naturais para desenvolver algoritmos computacionais. Os algoritmos desenvolvidos segundo essa abordagem utilizam um reservatório de informações codificado em um formato que imita a informação genética (cromossomos) e opera segundo uma lógica que permite que tais unidades troquem informação entre si (YAO X.; LIU, 1997). Os Algoritmos Genéticos (HOLLAND, 1975) constituem uma das mais conhecidas classes de algoritmos evolucionários. Eles são baseados em uma população de cromossomos, desenvolvendo sua fonte de diversificação através de operadores de cruzamento e mutações aleatórias. No contexto computacional "evolução" pode ser entendida como um método de busca dentro de um espaço de soluções viáveis. Os operadores de cruzamento são uma forma de re-combinar informações contidas nos cromossomos dos pais, produzindo filhos que representam novos pontos no espaço de busca. Apresentamos abaixo um pseudo-código do algoritmo genético (AG) utilizado em nossos experimentos computacionais:

No algoritmo descrito no Quadro acima o procedimento  $cruzamento(S_1, S_2)$  é o responsável pela recombinação dos cromossomos pais.  $S_1$  e  $S_2$  são cromossomos escolhidos em uma sub-população  $P' \subset P$ . O procedimento leva à formação de um novo cromossomo, filho. A variável  $f(S_i)$  guarda o valor da função objetivo da solução representada pelo cromossomo  $S_i$ . Uma vez que o problema tratado nesse trabalho é de minimização, quanto menor o valor dessa variável, melhor será a adequação de um dado indivíduo. O valor da função objetivo da solução representada pelo cromossomo filho é comparado com o valor da função objetivo das soluções representadas pelos cromossomos pais, através da variável  $S_{aux}$  que guarda o cromossomo pai com o pior valor da função objetivo. Dentre os três cromossomos (um filho e dois pais), retornam à população os dois que representam as melhores soluções (com menor valor da função objetivo). O procedimento  $mutacao(S_j)$  realiza mutações nos cromossomos escolhidos em  $P$ , introduzindo variabilidade no processo.



```

5.1 Inicializar população  $P$ 
5.2 Selecione uma subpopulação  $P'$  //Pais da próxima geração
5.3 repita
5.4   para  $i \leftarrow 1$  até  $nr\_cruzamento$  faça
5.5     escolha  $S_1, S_2 \in P'$ , aleatoriamente
5.6      $filho \leftarrow cruzamento(S_1, S_2)$ 
5.7     se  $f(S_1) \geq f(S_2)$  então
5.8        $S_{aux} \leftarrow S_1$ 
5.9     senão
5.10       $S_{aux} \leftarrow S_2$ 
5.11     fim
5.12     se  $f(S_{aux}) \geq f(filho)$  então
5.13       filho substitui  $S_{aux}$  em  $P$ 
5.14     fim
5.15   fim
5.16   para  $i \leftarrow 1$  até  $nr\_mutações$  faça
5.17     Selecione um cromossomo  $S_j$  em  $P$ 
5.18      $S_j \leftarrow mutacao(S_j)$ 
5.19   fim
5.20 até Critério de parada satisfeito

```

**Algoritmo 5:** Exemplo de um Algoritmo Genético (AG) simples

A grande vantagem do algoritmo genético em relação aos algoritmos de otimização tradicionais, é o fato de ele ser menos propenso em sua busca por uma solução sub-ótima a ficar preso em um mínimo local. Isso ocorre porque ele não faz uso de regras de transição aplicadas a um ponto único ao se mover de uma instância para outra no espaço de soluções. Ao invés disso, o algoritmo genético tira proveito de todo um conjunto de soluções espalhadas através do espaço de soluções, às quais estão experimentando muitos ótimos potenciais.

No entanto, para que os algoritmos genéticos possam trabalhar eficazmente, alguns critérios devem ser atendidos:

- a) Deve ser relativamente fácil avaliar o quão “boa” uma solução potencial é em relação à outras soluções possíveis.
- b) Deve ser possível quebrar uma possível solução em partes distintas, que podem variar de forma independente. Estas peças se tornam os “genes” no algoritmo genético.

- c) Finalmente, algoritmos genéticos são mais adequados para situações em que uma resposta “boa” será suficiente, mesmo se não é a melhor resposta absoluta.

### 6.3.7 Representação de uma solução candidata

Em nossa abordagem, utilizamos os algoritmos mencionados acima para evoluir: (i) o número de camadas ocultas; (ii) o número de elementos de processamento (neurônios) presentes em cada camada oculta; (iii) a taxa de aprendizagem; (iv) o termo de momentum e (v) a função de ativação.

A taxa de aprendizagem ( $\alpha$ ) pode assumir qualquer um dos valores gerados pela soma abaixo:

$$\alpha = \sum_{i=1}^k a_k \times 2^{-k} \quad (6.19)$$

onde  $a_k \in \{0, 1\}$ .

Uma expressão similar é adotada para o termo de Momentum ( $\eta$ ). Neste trabalho adotamos  $k = 6$  para ( $\alpha$ ) e  $k = 4$  para ( $\eta$ ). Os valores possíveis para os parâmetros remanescentes são mostrados na Tabela 6.1 abaixo:

Tabela 6.1 - Intervalo de valores assumidos pelos parâmetros que definem uma arquitetura de rede.

Parâmetro	Valores
Função de Ativação	[ Tanh <sup>1</sup>   Sigmóide   Logarítmica   Gaussiana ]
Nº de camadas ocultas	[ 1   2   3 ]
Nº de neurônios em cada camada	[ 1   ...   32 ]

A RNA construída a partir de informações codificadas em estruturas binárias é treinada com dados gerados pelo modelo direto SCIATRAN, que é um modelo bem conhecido e amplamente utilizado em estudos de recuperação de constituintes atmosféricos, e com dados obtidos a partir de um banco de dados climatológicos bem conhecido (ver seção 4.2).

<sup>1</sup>Tangente Hiperbólica

Para a simulação realizada neste trabalho, codificamos os parâmetros do problema em uma seqüência binária de 19 bits. O primeiro e segundo grupos de 6 e de 4 bits corresponde à taxa de aprendizagem e momentum, respectivamente. Nós representamos o tipo de função de ativação empregado no treinamento da rede e o número de camadas ocultas permitida na solução de rede pelo terceiro e quarto grupos consistindo de 2 bits cada um que se segue. O último grupo de 5 bits cada, está relacionado ao número de neurônios em cada camada. Devido ao esquema de codificação binária e o número de bits utilizado, o número de neurônios em cada camada oculta é restrito à  $[0, 31]$  já que  $1 + 2 + 4 + 8 + 16 = 31$ . Para esclarecer ainda mais o esquema de codificação aplicado, duas soluções candidatas diferentes, juntamente com as suas formas decodificadas, são fornecidas na Figura 6.4.

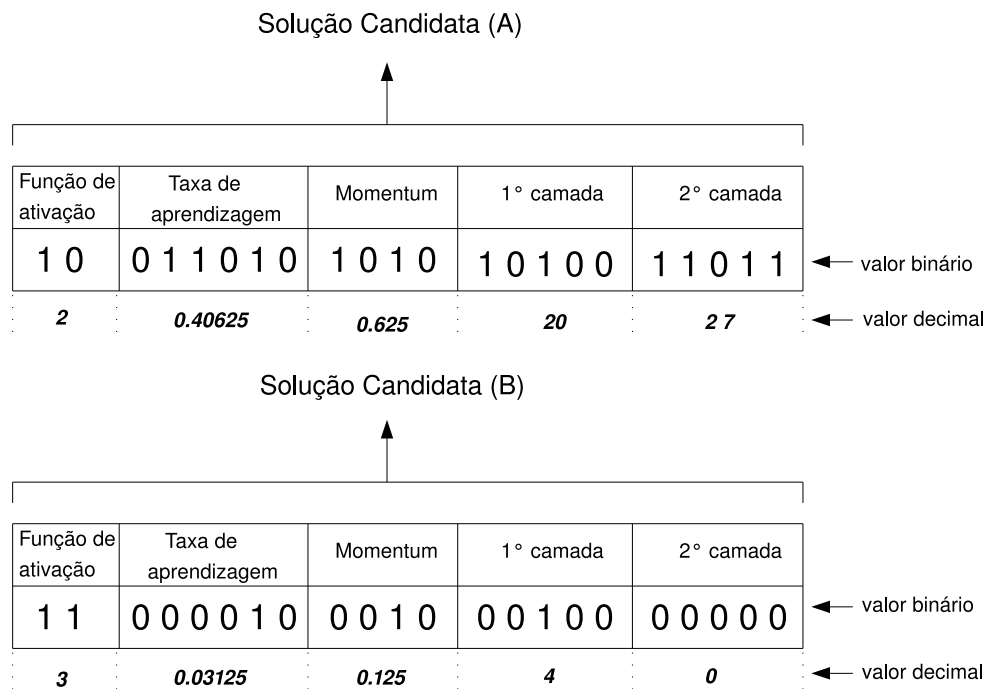


Figura 6.4 - Exemplos de codificação binária de duas soluções candidatas à arquitetura de rede e sua representação decimal correspondente.

Quando diminuimos ou aumentamos o tamanho da seqüência de bits estamos am-

pliando ou reduzindo o número de possíveis arquiteturas (a dimensão do espaço de soluções), assegurando assim a generalidade da metodologia. Podemos observar na Figura 6.5 abaixo a evolução no número de possíveis arquiteturas de RNA versus o número de bits na cadeia binária da solução.

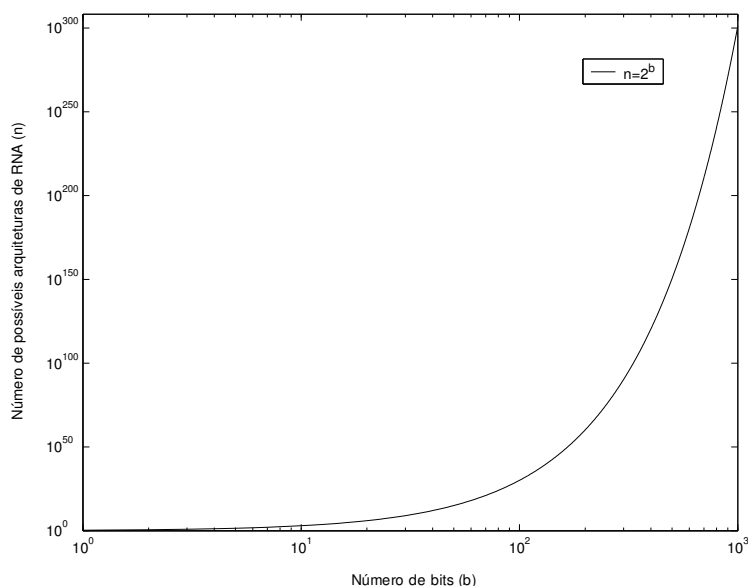


Figura 6.5 - Crescimento exponencial do espaço de busca. Gráfico que exhibe o número de possíveis arquiteturas de RNA versus número de bits na solução codificada.

É fácil demonstrar que o número de soluções possíveis representados por strings binárias segue uma simples lei de potência dada pela equação  $n = 2^{(b)}$  onde  $b$  é o número de bits e  $n$  o número de soluções. Portanto, neste trabalho o nosso espaço de busca foi formado por 524.288 soluções representando arquiteturas de rede possíveis o que corresponde a  $2^{(19)}$  combinações binárias.

O limite máximo de duas camadas escondidas e 32 neurônios em cada camada foi adotada neste trabalho apenas por razões práticas (tempo computacional). Estudos têm mostrado que, para RNAs do tipo *feedforward* com função de ativação contínua não-linear, uma camada oculta com um número de neurônios arbitrariamente grande seja suficiente para aproximar qualquer mapeamento contínuo devido à sua propriedade de aproximação universal ((BISHOP, 1995), (RIPLEY, 1996) e (HORNIK, 1993)). Este teorema é um teorema de existência, pois ele fornece uma justificativa para a aproximação de funções (ou mapeamentos) contínuas. Entretanto, o teorema não afirma que uma única camada é um número ótimo de camadas ocultas. De

acordo com Haykin, o problema com Perceptron Multi-Camadas que apresentam apenas uma camada intermediária é que os neurônios internos tendem a interagir uns com os outros de maneira global e dessa forma, em situações complexas esta interação dificulta a aproximação em um ponto sem que ocorra uma degradação em algum outro ponto (HAYKIN, 2001). Em redes com duas camadas ocultas esse processo é mais maleável com características locais sendo extraídas na primeira camada oculta e características globais na segunda. Portanto, este limite estabelecido pela aplicação de 2 camadas ocultas é mais do suficiente para a maioria, senão a totalidade das aplicações.

### 6.3.8 Aprendizado e teste da rede

Após codificarmos cada solução candidata em uma arquitetura de rede passamos para a etapa seguinte do algoritmo que corresponde ao treinamento da rede. Ao término do treinamento a rede é testada com o conjunto de dados de teste, que corresponde ao teste de generalização do modelo. Neste estágio, a RNA "aprendeu" a identificar os perfis de concentração do gás e é capaz de prever padrões de saída desconhecidos. Posteriormente, atribuímos um valor para a solução gerada de acordo com uma função objetivo (ver seção 6.3.9) previamente definido. Assim, construímos uma nova solução candidata utilizando o mesmo método descrito na seção 6.3.7 e o processo se repete.

Em nossos experimentos, utilizamos uma função de inicialização de pesos e bias estocástica com valores extraídos de uma distribuição uniforme entre  $[-1, 1]$ . Um termo momentum foi especificado para filtrar os distúrbios de alta frequência da superfície de erro. A estratégia de treinamento com parada antecipada (BISHOP, 1995) foi utilizada para evitar o "overfitting" dos dados. Durante a fase de treinamento, a capacidade de generalização da rede é avaliada através de um conjunto de dados chamado conjunto de validação. Quando o erro do conjunto de validação tende a aumentar, caracteriza-se perda de generalização, e o treinamento é encerrado conforme podemos verificar na Figura 6.6.

A taxa de aprendizagem (LR) variou da seguinte forma:

- a) Se  $E(t) > E(t - 1)$  então  $LR = LR \times (1 - step/100\%)$ ,
- b) Se  $E(t) \leq E(t - 1)$  então  $LR$  permanece inalterado,

onde "step" é um parâmetro ajustável entre 0 e 1 (*tipicamente*  $\in [0, 0.2]$ ), que controla a velocidade da diminuição da taxa de aprendizagem. Em seguida, a taxa de aprendizagem é reduzida para refinar a busca do mínimo global.

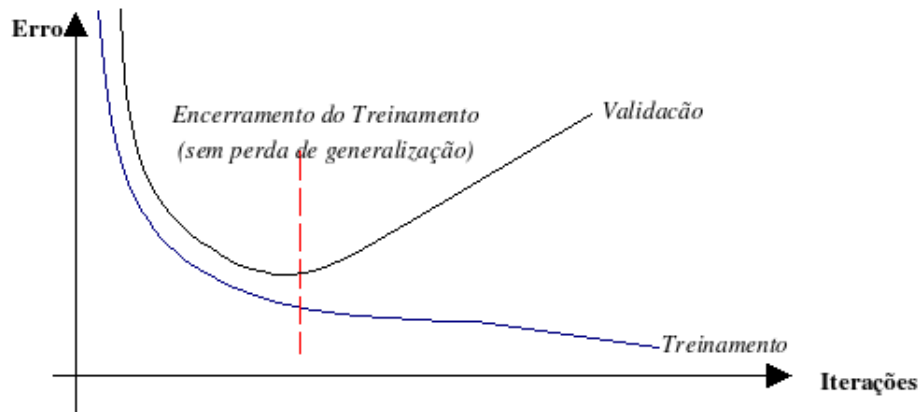


Figura 6.6 - Critério de parada baseado na perda de generalização da RNA

O processo de otimização de um modelo de rede que visa obter uma arquitetura ótima ou sub-ótima segue etapas similares não importando qual seja a meta-heurística utilizada. Essas etapas comuns aos algoritmos de busca são melhor compreendidas e descritas pelo fluxograma exibido na Figura 6.7.

### 6.3.9 Função objetivo

Neste trabalho lidamos com um problema de otimização combinatória relativamente complexo onde estamos interessados em encontrar uma arquitetura de custo mínimo, que apresente melhor desempenho com menor sobrecarga computacional possível. Por sua vez, otimização é o processo de encontrar a melhor solução em um espaço de busca formado por diversas soluções possíveis. Um problema de otimização é formado por uma função objetivo, que se deseja maximizar ou minimizar dependendo do problema, e um conjunto de restrições que precisam ser satisfeitas (PAPALAMBROS P. Y.; WILDE, 1988).

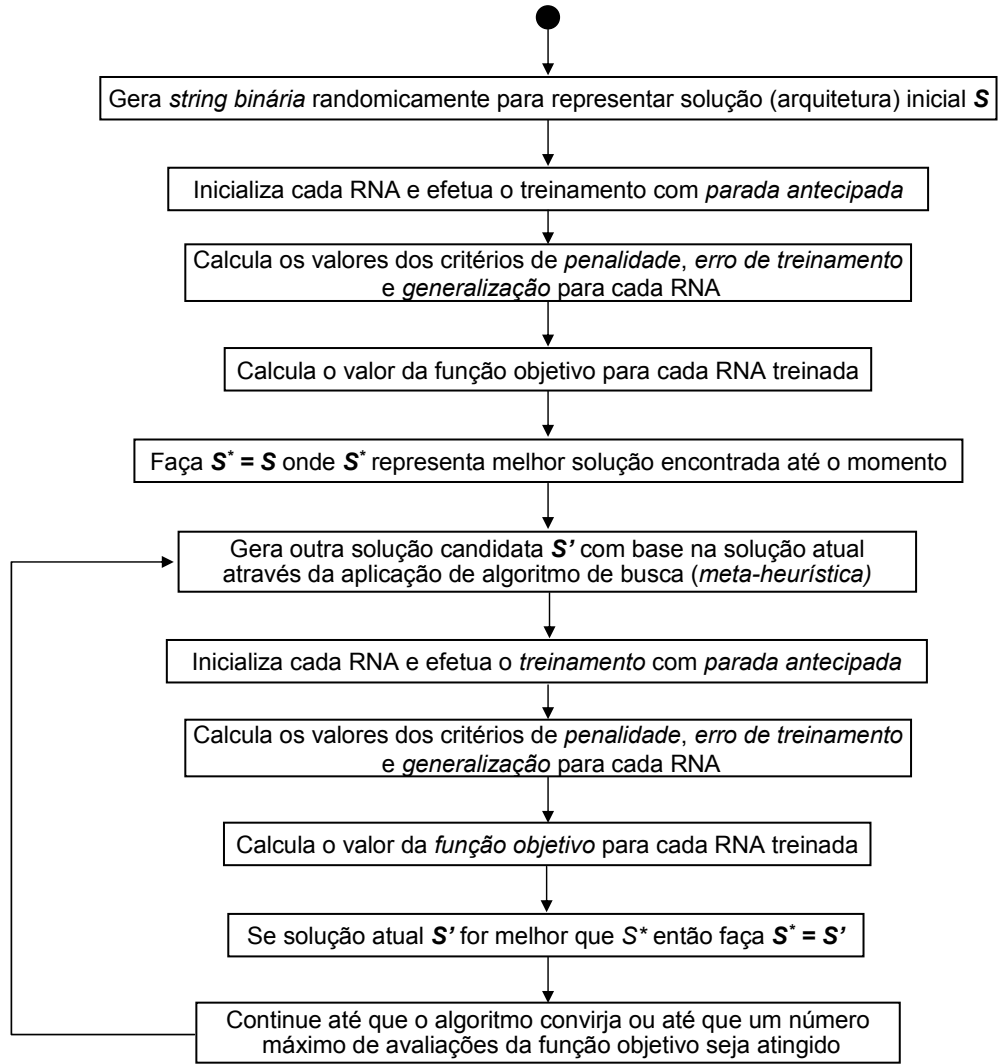


Figura 6.7 - Síntese da metodologia empregada para o problema da otimização de uma arquitetura de RNA do tipo *feedforward*

Para tanto, construímos uma função objetivo que consiste em uma combinação de 2 critérios de erro e um fator de penalização e que pode ser expressa por:

$$f_{obj} = penalidade \times \left( \frac{\rho_1 \times E_{trein} + \rho_2 \times E_{gen}}{\rho_1 + \rho_2} \right) \quad (6.20)$$

onde  $\rho_1$  e  $\rho_2$  ( $\rho_1, \rho_2 > 0; \rho_1, \rho_2 \in \mathbb{R}$ ) são fatores de ajuste utilizados para ponderar o grau de importância atribuído ao erros de treinamento e generalização respectivamente. A presença desses parâmetros de ajuste permite uma maior flexibilidade

na hora de avaliarmos a função objetivo. Na maioria dos casos teremos  $\rho_1 < \rho_2$  pois estamos interessados em modelos que dêem boas respostas quando apresentamos à rede um conjunto de valores de entrada que não pertença ao conjunto de treinamento, ou seja, que saibam generalizar. Em particular, para este trabalho, adotamos  $\rho_1 = 1$  e  $\rho_2 = 0.1$ . A função  $f_{obj}$  é a função cujo valor deve ser minimizado pelos algoritmos implementados. Desta forma, fica claro que  $f_{obj}$  é constituída pela soma dos erros de treinamento e generalização multiplicado pela penalidade devida à complexidade da arquitetura de rede em questão. O mínimo valor para  $f_{obj}$  corresponde a uma arquitetura simples e que exibe comportamento consistente no espaço de soluções aliado a baixos erros de treinamento e generalização. Aqui entendemos por "simples" uma solução cuja arquitetura seja reduzida em termos de números de pesos em tempo de aprendizagem.

### 6.3.9.1 Detalhando o critério usado para compor a função objetivo

#### 6.3.9.1.1 O critério do Erro de Treinamento

Este termo é importante porque nos fornece uma indicação quantitativa do nível de treinamento do modelo, e está diretamente relacionada à capacidade de memorização da rede. A fórmula do erro quadrático médio usada para calcular o erro de treinamento foi a seguinte:

$$E_{train} = \sqrt{\frac{1}{N \times M} \times \sum_{i=1}^N \sum_{j=1}^M [Y_{m_{ij}} - Y_{d_{ij}}]^2} \quad (6.21)$$

onde  $E_{train}$  é o erro de treinamento,  $Y_{d_{ij}}$  é o valor de saída desejado (alvo) do  $i$ -ésimo padrão de treinamento para o  $j$ -ésimo neurônio da camada de saída da rede,  $Y_{m_{ij}}$  é a resposta da RNA (modelo) no  $i$ -ésimo padrão de treinamento para o  $j$ -ésimo neurônio da camada de saída da rede,  $N$  é o número de padrões de treinamento ou exemplos apresentados à rede e  $M$  é o número de neurônios na camada de saída.

#### 6.3.9.1.2 Critério do erro de generalização

Refere-se à capacidade da RNA para identificar e responder aos padrões que são semelhantes mas não idênticos aos padrões com os quais a rede foi treinada. Afigura-se mais importante do que o erro de treinamento como um parâmetro sinalizador



do desempenho do modelo de RNA na maioria das aplicações. A fórmula utilizada para definir o erro de generalização é a seguinte:

$$E_{gen} = \sqrt{\frac{\sum_{i=1}^M [Y_{m_i} - Y_{d_i}]^2}{M}} \quad (6.22)$$

onde  $E_{gen}$  é o erro de generalização  $Y_{d_i}$  é o valor alvo para o ( $i$ -ésimo) neurônio de saída,  $Y_{m_i}$  é o valor previsto pelo ( $i$ -ésimo) neurônio de saída da RNA,  $M$  é o número de neurônios na camada de saída.

### 6.3.9.1.3 Penalidade devida à complexidade da arquitetura de rede

O terceiro critério determina a influência de uma arquitetura de RNA sobre os valores obtidos pela função objetivo. Podemos definir a complexidade computacional de uma arquitetura de RNA do tipo *feedforward* como o número total de pesos e bias presentes em sua estrutura. Quando falamos de otimização de um arquitetura de RNA estamos interessados na obtenção de modelos ou arquiteturas que apresentem um desempenho tão próximo quanto possível de um ótimo global para o problema em questão ou, pelo menos, que produzam respostas sub-ótimas melhores do que as respostas fornecidas por um modelo não-otimizado. Em outras palavras, procuramos arquiteturas de rede com baixos erros de treinamento e generalização. Para este fim, desenvolvemos um critério que favorece arquiteturas leves, aplicando um termo de penalidade à função objetivo. Ao utilizar arquiteturas mais simples, evitamos o sobre-ajuste de dados de treinamento da rede, conhecido como “overfitting” (HAYKIN, 2001), e aceleramos o processo de treinamento, porque haverá menos elementos de processamento e, por conseguinte, menos pesos para serem calculados além de aumentar a sua capacidade de generalização na medida certa. Em um problema de otimização tal termo de penalidade é incorporado à função objetivo de tal forma a limitar o universo de possíveis arquiteturas de rede. Ele penaliza arquiteturas maiores e complexas com muitos neurônios nas camadas ocultas ou que precisem de muito tempo de CPU (tempo de aprendizagem). Idealmente, a penalidade deveria ser uma função exponencial da complexidade de RNA e não uma função linear como se poderia esperar. Isto ocorre porque a soma do número de pesos e bias também aumenta de maneira não-linear com o número de neurônios e camadas ocultas. Assim, a penalidade seria gradualmente crescente enquanto o número total de pesos

e bias é baixo e cresceria mais rapidamente quando a arquitetura fosse complexa. Portanto, a forma geral do critério para a penalidade de uma arquitetura de RNA seria dada por

$$P_1 = a \times e^{b \times f(x)} \quad (6.23)$$

onde um  $a$  e  $b$  são constantes e  $f(x)$  uma função do total de pesos e bias, normalmente indicado por uma função exponencial. Também incluímos restrições sobre o tempo de aprendizagem dos padrões presentes na entrada da rede. Para esse fim, introduzimos um segundo termo de penalidade que leva em consideração o número de ciclos ou épocas necessárias para realizar o treinamento. Neste trabalho definimos este fator de penalidade como uma função linear do número de épocas, ou seja,  $g(y)$ , que tem a forma

$$P_2 = c \times y + d,$$

onde  $c$  é uma constante que mede a inclinação da reta e  $y$  representa o número de épocas ou ciclos de treinamento. Então, agrupando os termos definidos acima, temos a seguinte expressão global para a penalidade aplicada ao modelo:

$$Penalidade = P_1 + P_2 \quad (6.24)$$

Neste trabalho, especificamente, assumimos a seguinte expressão geral para a penalidade devida à complexidade do modelo:

$$Penalidade = \underbrace{e^{5 \times 10^{-8} \times x^2}}_{P_1} + \underbrace{5 \times 10^{-5} \times y + 1}_{P_2} \quad (6.25)$$

onde  $x$  é o número de pesos nas conexões e  $y$  o número de épocas necessárias no treinamento da rede. É importante notar que tanto  $P_1$  como  $P_2$  podem ser representados por qualquer função monótona contínua e estritamente crescente no intervalo  $[1, \infty]$ . A Figura 6.8 abaixo exhibe um modelo tridimensional para a função de penalidade aplicada neste trabalho. Como foi mencionado previamente, ao aplicarmos tais funções ao nosso problema damos preferência à soluções com menos neurônios

em cada camada oculta e que, ao mesmo tempo, não necessitem de muitas iterações para convergir para uma solução adequada.

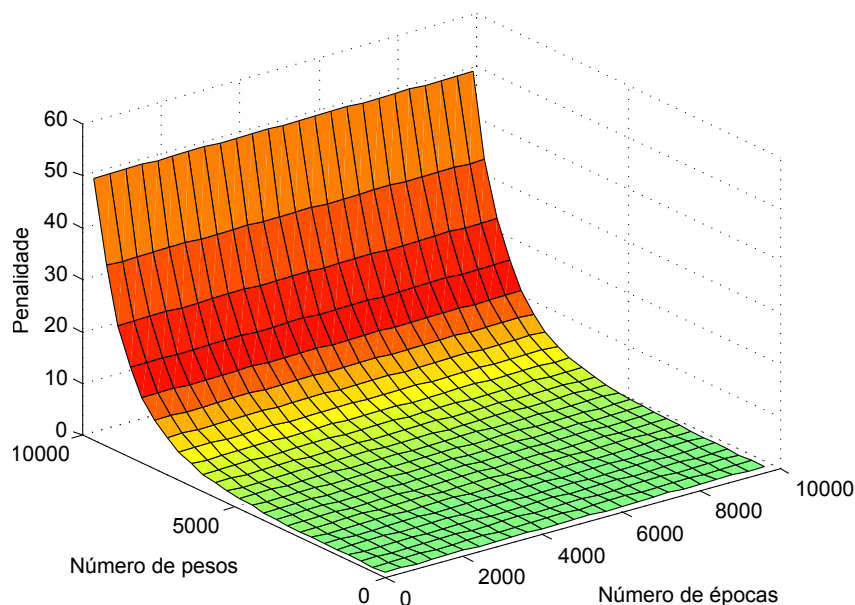


Figura 6.8 - Visualização tridimensional da função de penalidade aplicada ao problema.

Os resultados computacionais obtidos pelas execuções sucessivas de cada metaheurística são exibidos na seção 7.2 do próximo capítulo.

#### 6.4 Recuperação dos Perfis de Concentração de Metano e Dióxido de Carbono

As RNAs apresentam algumas peculiaridades e características que foram exploradas em nosso trabalho. Uma delas é a sua plasticidade na resolução de um problema e o seu paralelismo intrínseco. Procuramos por meio da técnica da divisão e conquista aferir o desempenho do modelo e verificar como a rede se comporta quando treinada com conjuntos de padrões de entrada-saída diferentes.

Em nosso trabalho testamos dois cenários alternativos de inversão para o problema da recuperação da concentração de gases traço atmosféricos. Nos parágrafos abaixo, fornecemos mais detalhes dos cenários testados e da metodologia de teste. Procuramos verificar através dessas experimentações se existe algum ganho significativo na adoção de um cenário em particular. Os resultados dos testes e a análise de cada alternativa proposta são apresentados no capítulo 7.

#### 6.4.0.1.4 Cenário 1: Inversão global

Em um primeiro experimento utilizamos os subgrupos de treinamento, validação e teste gerados anteriormente, em sua totalidade, para treinar uma rede neural única. Essa rede neural foi treinada com perfis de massas de ar tropicais, temperadas e subárticas e cada perfil era constituído por 50 pontos de concentração ao longo de 4 camadas atmosféricas que vão da superfície à altitude de 120 km. Treinamos uma rede neural para cada conjunto de canais do sensor. Ao final do treinamento a rede foi apresentada à cinco padrões de entrada desconhecidos e os perfis produzidos foram comparados com os perfis reais conhecidos à priori.

#### 6.4.0.1.5 Cenário 2: Inversão localizada distribuída por camadas atmosféricas

Por último, criamos um cenário misto, no qual treinamos uma rede neural para cada camada atmosférica e massa de ar isoladamente. Para tanto segmentamos o conjunto de dados de treinamento gerado (ver seção 6.2) em grupamentos de treinamento de tal forma que em cada cluster armazenamos somente radiâncias e perfis pertencentes a uma determinada massa de ar (pertencente a uma região climática) e a uma determinada camada atmosférica (troposfera, estratosfera, mesosfera ou termosfera). Desta forma ao final desse procedimento foram criados 20 subconjuntos de dados de treinamento. Cada rede neural, com topologia (configuração dos parâmetros ajustáveis da rede) otimizada pelo algoritmo de otimização detalhado na seção 6.3, foi treinada com um desses subconjuntos. Para esse cenário existem diversas combinações possíveis de redes. Podemos ter, por exemplo, uma rede neural específica para perfis tropicais estratosféricos, outra para perfis subárticos de inverno troposféricos e assim por diante.

Os parâmetros da RNA (modelo de inversão) foram encontrados mediante a aplicação de uma estratégia de otimização paramétrica global explicada em maiores detalhes na seção 6.3 deste capítulo. Os valores iniciais para a taxa de aprendizado, termo de momento, função de ativação, número de camadas e de neurônios em cada camada para cada uma das redes utilizadas nos cenários 1 e 2 foram encontrados após sucessivas aplicações dos algoritmos de busca e otimização. Os valores desses parâmetros foram exibidos nas tabelas 7.8, 7.9, 7.11 e 7.12 do capítulo 7 (ver seção 7.2).

## 7 RESULTADOS

Inicialmente mostraremos os resultados para a seleção de canais do sensor e em seguida exibimos os parâmetros de configuração e topologia de rede que foram encontrados para o modelo inverso proposto. Esses resultados foram resultantes da aplicação da metodologia de otimização exposta na seção 6.3 do capítulo anterior. Finalmente mostramos os produtos (perfis) gerados pela aplicação do modelo para a recuperação dos perfis de concentração dos gases traço e traçamos uma análise qualitativa e quantitativa *a posteriori*.

### 7.1 Resultados para Seleção de Canais

As 3 abordagens para seleção de canais utilizadas neste estudo e apresentadas no capítulo anterior resultaram num conjunto mínimo de canais para recuperação do perfil de concentração de gases traço. Na figura 7.1 abaixo podemos ver o resultado final, com as linhas espectrais identificadas pela sua frequência para cada abordagem de seleção.

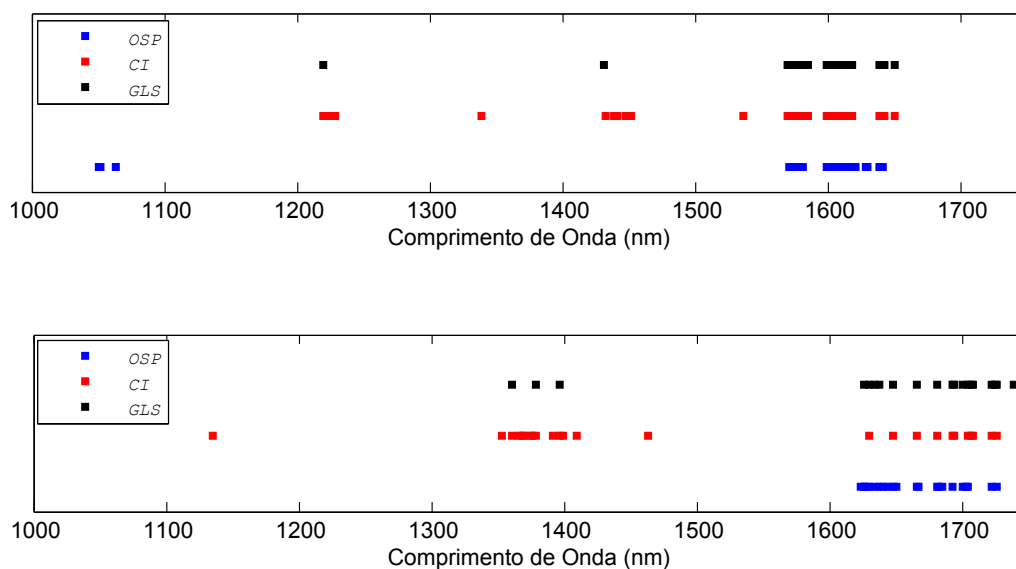


Figura 7.1 - Canais SCIAMACHY utilizados pelo modelo de inversão proposto para inferir perfis verticais de concentração de  $CO_2$  (superior) e  $CH_4$  (inferior) selecionados pelos métodos OSP (azul), conteúdo de informação (vermelho) e pelo método dos graus de liberdade (preto).

Nas seções seguintes analisamos mais à fundo a figura acima fornecendo detalhes

sobre a escolha dos canais feita por cada abordagem e sua localização no espectro.

### 7.1.1 Método OSP (*Optimal Sensibility Profile*)

#### 7.1.1.1 Resultados dos Estudos de sensibilidade para os canais do SCIAMACHY

Para  $CO_2$  optou-se por perturbar o perfil de metade do valor da amplitude média de pico-a-pico do ciclo sazonal do gás, o que equivale a 4 ppmv para situações tropicais e 9 ppmv para situações temperadas e polares (CONWAY et al., 1994). O sinal correspondente para o  $CO_2$  obtidos nas faixas do canal 6 do sensor é mostrado na Figura 7.2 (a), para um perfil representativo tropical. Para o  $CH_4$  utilizamos 0.05 ppmv de perturbação média para a mesma faixa espectral como podemos observar na Figura 7.2(b).

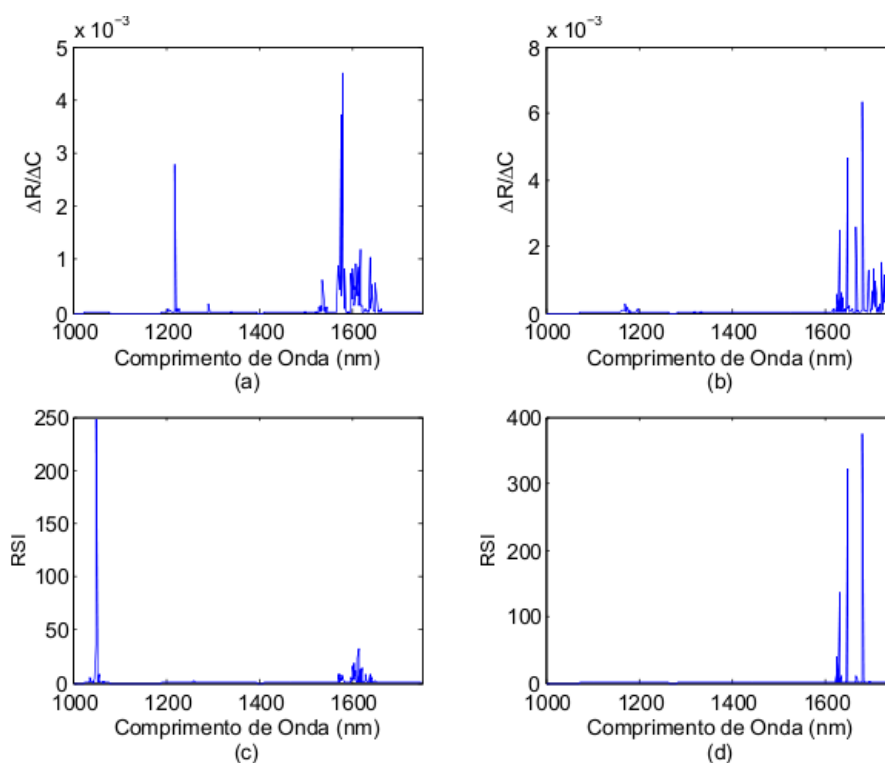


Figura 7.2 - (a) Sensibilidade para uma perturbação de 8 ppmv no perfil de  $CO_2$  vs. comprimento de onda na janela espectral de 1000-1750 nm do SCIAMACHY (canal 6). (b) Sensibilidade para uma perturbação de 0.05 ppmv no perfil de  $CH_4$  vs. comprimento de onda na janela espectral de 1000-1750 nm do SCIAMACHY (canal 6). (c) RSI (ver seção 6.1.1 do capítulo anterior) para o  $CO_2$ . (d) RSI para o  $CH_4$

Constatamos que as linhas de  $CO_2$  muito sensíveis do canal 6 do sensor SCIAMACHY localizam-se principalmente dentro de duas janelas espectrais, uma centrada em 1200 nm e outra na faixa de 1580-1650 nm. Para o  $CH_4$  essas linhas localizam-se uma janela única que vai de 1650 nm a 1700 nm às quais correspondem às regiões de absorção desses gases. Naturalmente, algumas linhas espectrais podem também ser sensíveis à outros constituintes atmosféricos. O que ocorre nesses casos é a existência de uma sobreposição de bandas de absorção em determinadas faixas espectrais. Este efeito também deve ser considerado no processo de seleção desde que ele interfere nas medidas de  $CO_2$  e  $CH_4$ . Nesse trabalho, estudamos o efeito da seção cruzada das linhas de  $N_2O$  e  $CH_4$  para a recuperação de  $CO_2$  atmosférico e das linhas de  $N_2O$  e  $CO_2$  para a recuperação de  $CH_4$ .

Altos valores de RSI, até 300 por ex., podem surgir a partir de valores muito baixos de interferência, que é o caso para regiões “estratosférica“ enquanto que, em regiões ”troposféricas“, pode ser tão baixo quanto 0.01. Isto é devido ao fato de que tanto vapor d’água, características da superfície, ou outros gases como o  $N_2O$  e CO, afetam essencialmente os canais troposféricos. Nos experimentos realizados, o critério de poda utilizado para eliminação de linhas espectrais foi  $RSI \geq 2$  para o  $CO_2$  e  $RSI \geq 1$  para o  $CH_4$ . No entanto, valores baixos de interferência também podem resultar em grandes valores RSI mesmo que o sinal do gás traço a ser recuperado seja baixo. Como conseqüência, um segundo critério tem de ser introduzido: o valor do numerador, o sinal do gás, deve ser superior a um limiar fixo, aqui tomado igual a 0.004 (ou 0.4% do sinal) para o  $CO_2$  e 0.002 (0.2% do sinal) para o  $CH_4$ . Este valor foi escolhido para evitar a seleção de muitos canais, considerando a perturbação pelos valores utilizados. Nas figuras 7.2(c) e 7.2(d), podemos observar os valores de RSI para o  $CH_4$  e  $CO_2$  na faixa espectral do canal 6 do sensor SCIAMACHY.

Outra característica relevante em modelos de retrieval que utilizam dados de satélite é a relação sinal/ruído do instrumento. Evidentemente, faixas nas quais o ruído no sinal é excessivo interferem negativamente no processo de retrieval e devem ser evitadas. Um gráfico demonstrando o ruído observado nos 8 canais do sensor SCIAMACHY é exibido na Figura 7.3. Na gráfico da Figura 7.4 é apresentado o ruído intra-canal especificamente para o canal 6.

Globalmente, o desempenho do detector é caracterizado por baixo nível de ruído e altas taxas de transmissão e recepção de dados. Isto permite medir a luz incidente com a relação sinal-ruído muito alta requerida, um pré-requisito para a obtenção

dos parâmetros geofísicos considerados nesse trabalho.

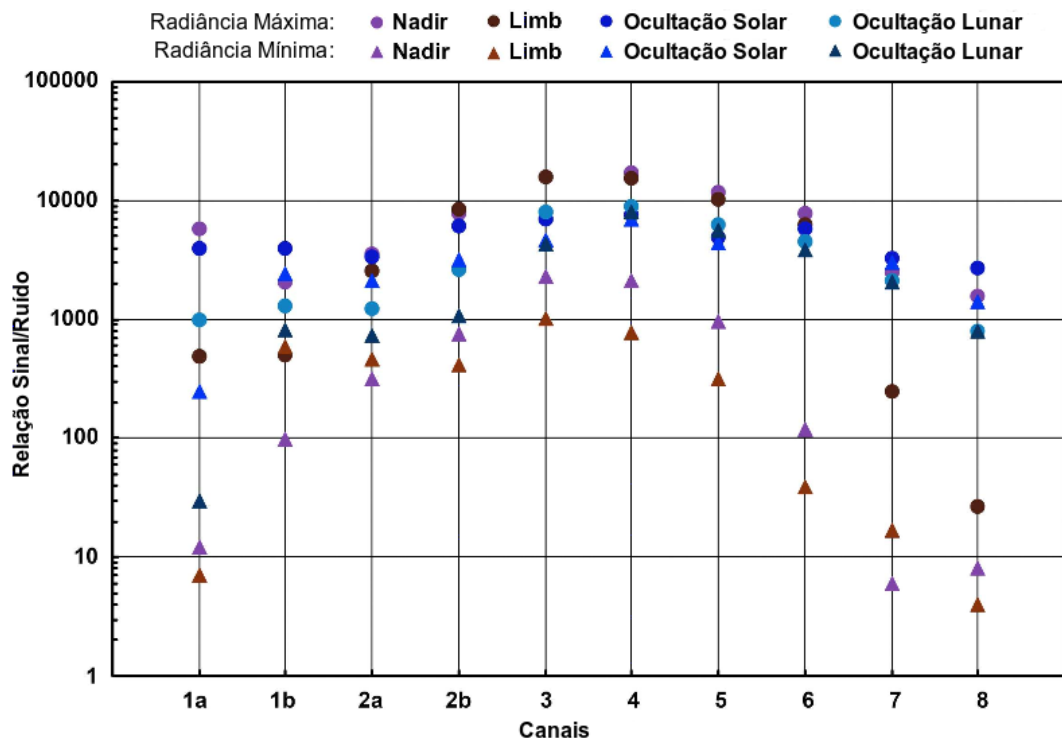


Figura 7.3 - Relação Sinal-Ruído (RSR) obtida para os canais do sensor SCIAMACHY durante testes de calibração. As relações foram determinadas para os sinais de radiância máximo e mínimo. (gráficos: DLR-FMI).

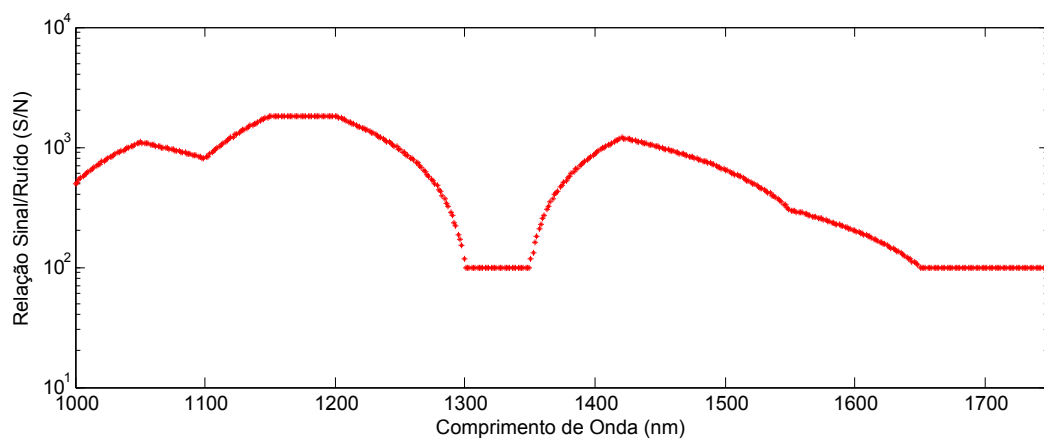


Figura 7.4 - Relação Sinal/Ruído (RSR) para o canal 6 do sensor SCIAMACHY.



A temperatura da superfície e emissividade são supostamente conhecidas dentro da faixa 1K e 0.01, respectivamente. Esses valores são os erros de recuperação esperados (ver (CHEDIN et al., 2002)). Os sinais são mostrados nas Figuras 7.5(d). Para os três gases restantes, da mesma forma que as emissões de  $CO_2$ , assume-se perturbações constantes e iguais a 4% para o  $N_2O$  e 40% para o  $CO$  (INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC), 2001). Os sinais são representados nas Figuras 7.5(b) e 7.5(c). Para o vapor d'água assumimos um erro dado por um limite superior de 20% de perturbação em um perfil médio e o seu sinal é mostrado na Figura 7.5(a).

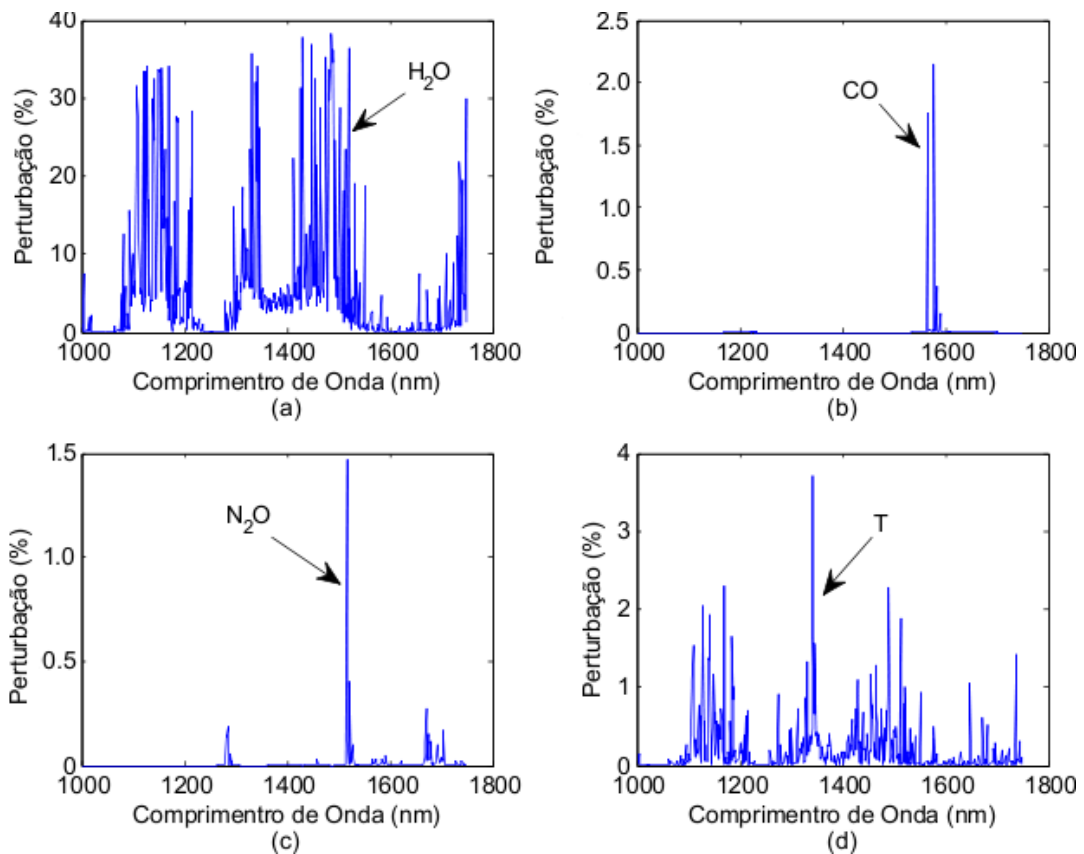


Figura 7.5 - Sensibilidade dos canais do SCIAMACHY para (a)  $H_2O$ , (b)  $CO$ , (c)  $N_2O$  e (d) T para a faixa de comprimento de onda de 1000-1750nm (banda do canal 6) e uma situação representativa tropical.

A variação na intensidade do sinal captado pelo sensor no canal 6, quando perturbamos os perfis de vapor d'água e os perfis de temperatura segue um padrão similar.

Ambas apresentam forte sensibilidade às mudanças no perfil em quase toda a banda de comprimento de onda considerado. Notamos que algumas linhas no canal 6 são fortemente influenciadas pela umidade e deveriam ser evitadas. No caso específico do vapor d'água a variação capturada pelo sinal do sensor é causada principalmente por um distúrbio ao nível da troposfera, que concentra mais de 80% de todo vapor d'água presente na atmosfera. Contudo, é possível observar a existência de algumas janelas atmosféricas insensíveis a tais perturbações. É justamente nessas janelas que estão localizados os canais selecionados para a recuperação de  $CO_2$  e  $CH_4$ .

Embora os canais sensíveis ao  $CO_2$  e ao  $CH_4$  sejam, como qualquer outro canal, antes de tudo sensíveis à variações no perfil de temperatura, o efeito das variações do gás traço segue uma tendência (bias) bem definida, enquanto que o sinal gerado a partir do perfil de temperatura perturbado apresenta, em geral, uma distribuição aleatória ao longo de um período de tempo razoável (cerca de duas semanas) e deve, portanto, ser diluído na média, quando uma amostra grande de radiâncias é considerada (CHEDIN et al., 2002). Por esse motivo, não levamos em consideração neste momento, a sensibilidade dos canais com relação à temperatura. Contudo, existem condições onde os erros de temperatura no modelo e os erros de  $CO_2$  e  $CH_4$  podem estar correlacionados e precisarão ser estudados em mais detalhes com dados *in situ*.

O resultado final, com as linhas espectrais identificadas pela sua frequência está representado nas Tabelas 7.1 e 7.2. No geral, 26 linhas foram escolhidas para recuperação de  $CO_2$  e 21 linhas para o  $CH_4$ .

Tabela 7.1 - Canais do SCIAMACHY (em nm) utilizados para recuperar o perfil vertical de concentração de  $CO_2$  escolhidos pelo método OSP.

1050.32	1051.80	1063.64	1569.78	1572.74	1574.22
1575.72	1578.68	1580.14	1597.90	1599.38	1600.86
1602.34	1603.82	1605.32	1606.80	1608.28	1611.24
1612.72	1615.68	1617.16	1620.12	1627.52	1629.00
1637.88	1640.84				

Observou-se através da aplicação do método OSP para seleção de canais que as linhas espectrais ficaram mais concentradas, tanto para  $CH_4$  quanto para  $CO_2$ , em torno de janelas espectrais mais restritas (para  $CH_4$  ao redor de 1620-1700 nm e para

Tabela 7.2 - Canais do SCIAMACHY (em nm) utilizados para recuperar o perfil vertical de concentração de  $CH_4$ .

1623.08	1626.04	1627.52	1630.48	1633.44	1637.88
1640.84	1643.80	1648.24	1651.20	1666.00	1667.48
1680.80	1682.28	1685.24	1692.64	1700.04	1703.00
1704.48	1722.24	1726.68			

$CO_2$  em torno de 1600 nm) . Como veremos nas próximas seções as linhas espectrais selecionadas pelas outras duas abordagens embora contemplem as linhas escolhidas por este método abrangem também outras micro-janelas espectrais distribuídas ao longo do espectro do canal 6 do sensor.

### 7.1.2 Conteúdo de Informação (CI) de Shanon para o sinal de $CO_2$ e $CH_4$

Em um esforço para aumentar significativamente a performance de algoritmos de recuperação de gases traço atmosféricos para o SCIAMACHY, realizamos uma análise do conteúdo de informação do canal 6 deste sensor. Ao término desse procedimento foi possível identificar 29 sub-canais (linhas de absorção espectral) para o  $CO_2$  . Recuperações efetuadas com estes canais foram capazes de fornecer tanto quanto 80% do conteúdo total de informação de  $CO_2$  em comparação com recuperações usando todos os mais de 500 canais disponíveis nessa região espectral. Para o  $CH_4$  identificamos 25 canais ao todo, que são responsáveis por mais de 83% de todo o CI para o canal 6 do sensor. As recuperações de  $CO_2$  e  $CH_4$  usando os canais selecionados têm uma precisão melhor que 1% (conforme exibido na seção 7.2 abaixo). Esta técnica pode ser aplicada para a recuperação de outras variáveis geofísicas (por exemplo, temperatura, umidade, etc), ou modificada para outros instrumentos, como AIRS, OCO ou GOSAT.

Em nossos experimentos computacionais, adotamos  $x$  como sendo o perfil atmosférico do gás traço a ser recuperado ( $CO_2$  ou  $CH_4$ ). Assume-se que vapor d'água, demais gases traço e características da superfície são conhecidos dentro de alguma indeterminação residual e são levados em consideração no cálculo do erro da medida. O método é aplicado para uma situação atmosférica tropical. Assumimos que a matriz de covariância do erro global  $S_\epsilon$  seja diagonal. A matriz de covariância do erro da medida  $S_M$  leva em conta o ruído do instrumento e as interferências. Para cada canal, o ruído do instrumento é calculado para a situação atmosférica estudada. O

“ruído” devido à indeterminação residual do vapor d’água, gases traço e características de superfície, é definida como a raiz quadrada da média das perturbações na Radiância. A variância do erro do modelo direto em  $S_F$  é tida constante e igual a 0.01% de  $\mu_y$  (média das medidas de radiância do canal 6) para todos os canais. Esta representação do ruído do modelo direto não é realista e é um ponto a ser explorado em trabalhos futuros afim de melhorarmos nosso conhecimento acerca do erro do modelo direto. No entanto, este valor foi usado na maioria das publicações encontradas na literatura ((RODGERS, 1996); (PRUNET P.; THÉPAUT; CASSÉ, 1998)).

Infelizmente, o conhecimento prévio de  $CO_2$  e  $CH_4$  atmosférico é bastante pobre devido à escassez de medidas que, no entanto, indicam valores relativamente constantes para a concentração de  $CO_2$  na troposfera e na estratosfera. O metano se distribui de maneira uniforme e constante na troposfera mas nas camadas superiores, onde esse comportamento não ocorre, necessitamos ajustar os valores de maneira mais cuidadosa. Portanto, assumimos que a matriz de covariância *a priori*  $S_a$  do  $CO_2$  tem uma variância constante e igual a  $10^2$  ppmv para a estratosfera e a camada limite, e igual a  $8^2$  ppmv para a troposfera enquanto que para o  $CH_4$  ela equivale a  $0.09^2$  ppmv para a troposfera e camada limite e  $(0.09)^2 \times r$  onde o fator  $r$  corresponde à razão  $\left(\frac{p_i}{p_{sup}}\right)$  entre a pressão no nível  $i$  e a pressão à nível do mar e, portanto, segue uma aproximação proporcional ao longo do perfil atmosférico. Os maiores valores para o  $CO_2$  foram tomados para a estratosfera e a camada limite, pois espera-se que seja mais difícil recuperar a concentração de  $CO_2$  nessas duas partes da atmosfera do que na troposfera. Essas estimativas para variância poderão ser melhoradas no futuro com os primeiros resultados obtidos a partir de estudos em curso (campanhas de aeronaves, a análise da NOAA e dados ENVISAT/SCIAMACHY, etc.). Verificamos que os valores das variâncias em si não afetam a ordem de seleção dos canais. Todos as matrizes aqui consideradas são diagonais e para os elementos que não pertencerem à diagonal principal assume-se que serão iguais a zero.

Abaixo, as Figuras 7.6 e 7.7 mostram a localização espectral dos 507 canais e sua posição ou classificação na seleção do CI, para uma situação representativa de uma atmosfera tropical para os dois gases traço considerados nesse estudo. Entre eles, o canal rotulado como 1 na ordenada e localizado na posição 390 do espectro na Figura 7.6 e 257 na Figura 7.7 foi o primeiro a ser escolhido: é o que apresenta o maior ganho em informação relativo a uma configuração de estado inicial. Após a seleção deste canal, a matriz de covariância é atualizada e, em seguida é calculado

o CI dos canais remanescente em relação a este novo estado.

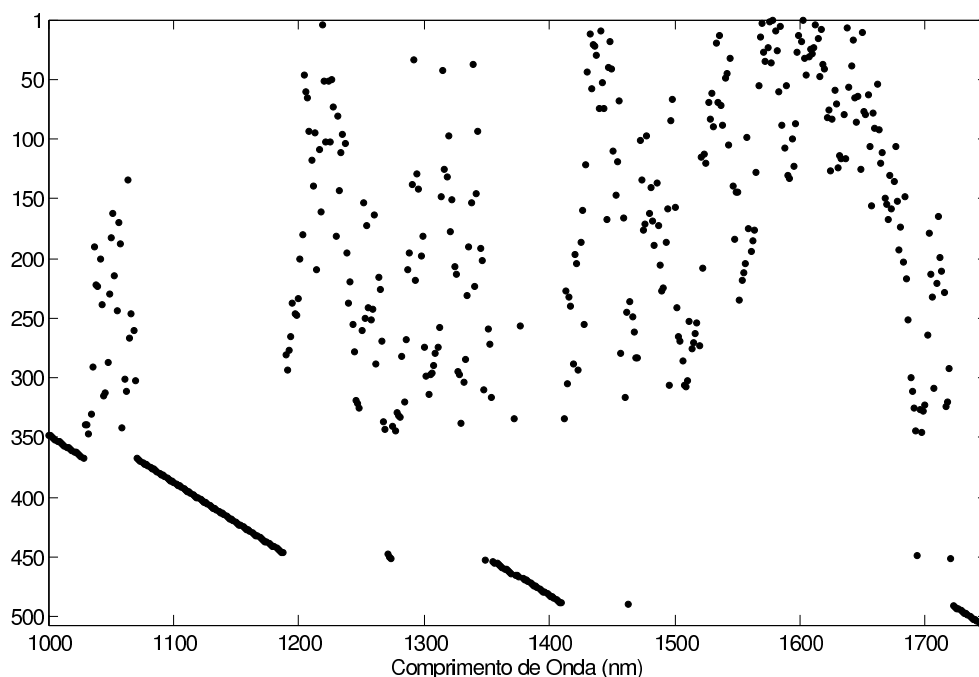


Figura 7.6 - Classificação de canais do sensor SCIAMACHY na seleção do conteúdo de informação para o gás  $CO_2$  em uma atmosfera tropical.

Selecionar, por exemplo, dez canais consiste em manter os dez canais que apresentaram o maior ganho em informação relativo a um estado à priori. Isto é diferente do que foi realizado para o método OSP. Agora, ordenamos os canais pelo seu CI e canais estratosféricos e troposféricos estão misturados. Sendo assim, a idéia de ordenar os canais, nível por nível, como realizado no método OSP, é perdida.

Na Figura 7.6, é possível observar que para o  $CO_2$  algumas regiões espectrais, ao redor de 1000, 1100 (uma grande faixa que vai de 1080 a 1200 nm), 1400 e 1740 nm parecem não ter muita relevância por não acrescentarem informação alguma ao processo. O mesmo vale para o metano (Figura 7.7) nas bandas que se encontram nos intervalos espectrais de 1000 – 1100, 1300 e 1575 nm. Nas tabelas abaixo, listamos os canais com maior CI selecionados pelo processo descrito acima para os dois gases

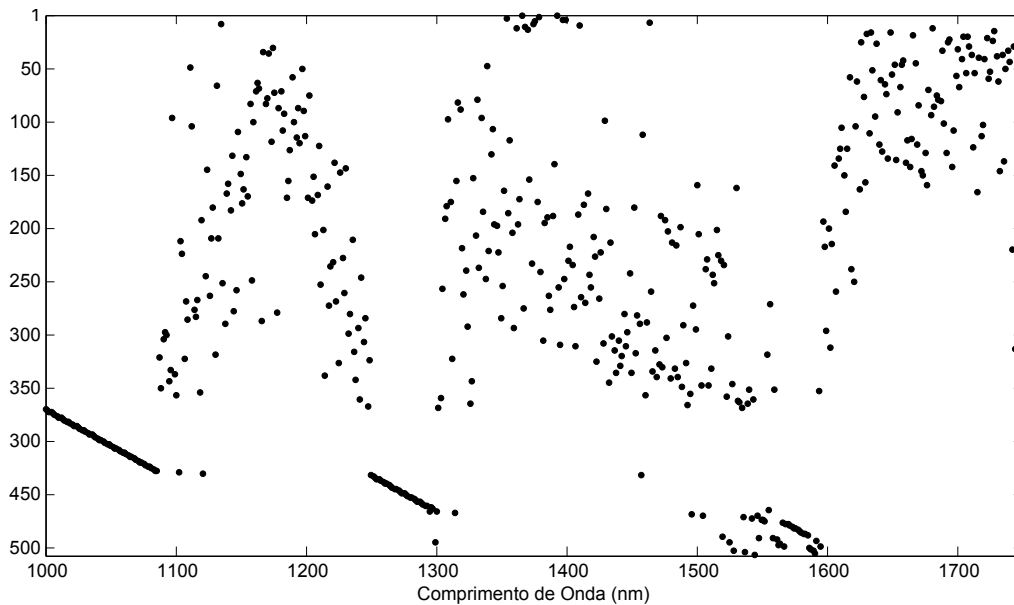


Figura 7.7 - Classificação de canais do sensor SCIAMACHY na seleção do conteúdo de informação para o gás  $CH_4$  em uma atmosfera tropical.

traço analisados e o conteúdo de informação correspondente para cada canal.

O conteúdo de informação inicial de cada canal pode ser graficado contra o número do canal ou comprimento de onda como um “espectro de informação”. Após a seleção do primeiro canal, um novo espectro de informação pode ser calculado. Para ver o ganho em informação, as Figuras 7.8 e 7.9 exibem a diferença entre o espectro de informações inicial e o espectro obtido após a seleção de 1,2,3,4,5,10 e 25 canais. O ganho em CI obtido com os primeiros 5 canais selecionados é importante. Entretanto, o aumento de informação quando se passa de 20 canais (linha amarela) para 25 canais (linha verde claro) é menor e de 25 para qualquer outro número superior a esses é praticamente nulo. Desta forma, a seleção de um conjunto reduzido de canais é o suficiente para obtermos o máximo do conteúdo da informação de todo o conjunto de canais.

Para o  $CO_2$  os primeiros canais selecionados estão localizados em 2 bandas do gás, mais precisamente, em 3 domínios restritivos: um centrado em 1575 nm conjugado com um segundo em 1615 nm e por último um terceiro em 1225 nm. Já para o  $CH_4$  os canais com maior ganho encontram-se na faixa que vai de 1350 – 1410 nm com mais 2 canais isolados em torno de 1130 nm e 1465 nm.

Tabela 7.3 - Os primeiros 29 canais selecionados para recuperação de  $CO_2$  com máxima contribuição para o total do conteúdo de informação.

	Número do Canal	Localização (nm)	Conteúdo de Informação
1	390	1575.72	1.30e+0
2	392	1578.68	4.34e-1
3	414	1611.24	1.06e-1
4	386	1569.80	6.94e-2
5	149	1219.04	5.51e-2
6	405	1597.92	2.57e-2
7	389	1574.24	1.22e-2
8	155	1227.92	1.06e-2
9	412	1608.28	1.00e-2
10	230	1338.92	9.42e-3
11	432	1637.88	8.57e-3
12	385	1568.32	8.12e-3
13	407	1600.88	6.59e-3
14	440	1649.72	6.41e-3
15	393	1580.16	5.58e-3
16	418	1617.16	4.89e-3
17	297	1438.08	4.83e-3
18	363	1535.76	3.97e-3
19	408	1602.36	3.84e-3
20	396	1584.60	3.61e-3
21	388	1572.76	3.49e-3
22	299	1441.04	3.25e-3
23	409	1603.84	3.13e-3
24	152	1223.48	2.69e-3
25	303	1446.96	2.67e-3
26	435	1642.32	2.54e-3
27	306	1451.40	2.29e-3
28	293	1432.16	2.13e-3
29	415	1612.72	1.92e-3

### 7.1.3 Graus de liberdade para o Sinal de $CO_2$ e $CH_4$

Uma vez mais, os primeiros canais a serem selecionados estão localizados nas duas bandas mais sensíveis de  $CO_2$  e  $CH_4$ , em domínios restritos próximos aos obtidos com o método do CI. Nas tabelas abaixo listamos os canais selecionados pelo método GLS para recuperação da concentração dos dois gases traço analisados e os graus de liberdade para cada canal considerado.

Tabela 7.4 - Os primeiros 25 canais selecionados para recuperação de  $CH_4$  com máxima contribuição para o total do conteúdo de informação.

	Número do Canal	Localização (nm)	Conteúdo de Informação
1	257	1378.88	3.98e+0
2	248	1365.56	3.49e+0
3	255	1375.92	3.15e+0
4	240	1353.72	2.03e+0
5	439	1648.24	3.53e-1
6	266	1392.20	1.91e-1
7	250	1368.52	6.40e-2
8	254	1374.44	5.28e-2
9	269	1396.64	1.70e-2
10	314	1463.24	1.55e-2
11	271	1399.60	8.86e-3
12	461	1680.80	8.03e-3
13	245	1361.12	7.79e-3
14	427	1630.48	2.99e-3
15	251	1370.00	2.60e-3
16	278	1409.76	1.10e-3
17	451	1666.00	3.58e-4
18	92	1134.68	3.02e-4
19	489	1722.24	9.59e-5
20	477	1704.48	5.99e-5
21	469	1692.64	5.73e-5
22	479	1707.44	4.10e-5
23	480	1708.92	3.62e-5
24	470	1694.12	2.35e-5
25	492	1726.68	2.27e-5

O critério para a escolha de canais pôde ser derivado da razão  $q$  entre o valor do grau de liberdade para o sinal calculado e a somatória dos graus de liberdade de todos os outros canais selecionados, em termos percentuais, e pode ser expresso por

$$q = \left( \frac{d_j}{\sum_{i=1}^{j-1} d_i} \right) \times 100 \quad (7.1)$$

Assumimos  $q < 0.1\%$  para o  $CO_2$  e  $q < 0.01\%$  para o metano pois apresenta menos canais sensíveis na faixa espectral considerada e um número muito reduzido de canais seria selecionado caso fosse adotado o mesmo limiar utilizado para o  $CO_2$ .



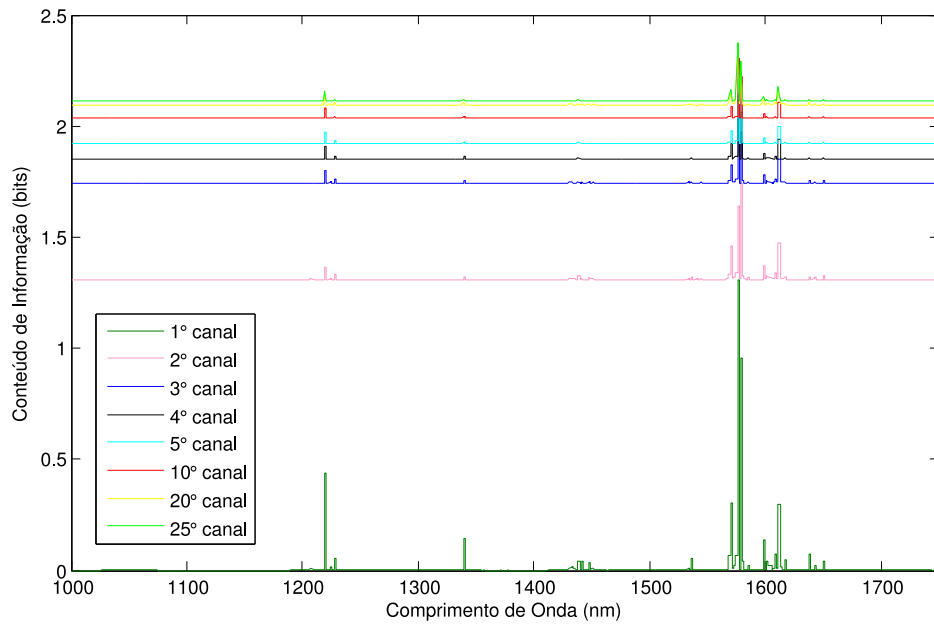


Figura 7.8 - Ganho em informação para uma situação representativa de uma atmosfera tropical após a seleção de 1,2,3,4,5,10 e 25 canais respectivamente, usando o método do conteúdo de informação para o gás  $CO_2$ .

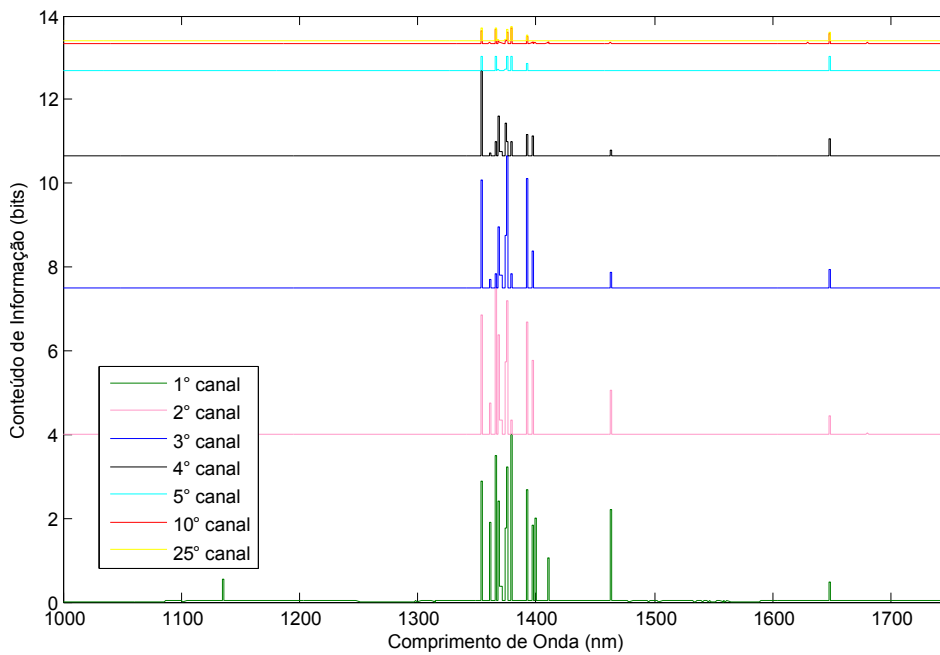


Figura 7.9 - Ganho em informação para uma representativa situação atmosférica tropical após a seleção de 1,2,3,4,5,10 e 25 canais respectivamente, usando o método do conteúdo de informação para o gás  $CH_4$ .

Tabela 7.5 - Os primeiros 23 canais selecionados para recuperação de  $CO_2$  com maior número de graus de liberdade.

	Número do Canal	Localização (nm)	Graus de Liberdade
1	392	1578.68	7.15e+0
2	385	1568.32	2.23e+0
3	414	1611.24	2.23e+0
4	386	1569.80	9.52e-1
5	149	1219.04	5.27e-1
6	405	1597.92	3.53e-1
7	389	1574.24	1.95e-1
8	432	1637.88	1.63e-1
9	390	1575.72	1.48e+1
10	412	1608.28	1.10e-1
11	440	1649.72	9.53e-2
12	418	1617.16	9.01e-2
13	407	1600.88	7.77e-2
14	393	1580.16	7.16e-2
15	408	1602.36	5.95e-2
16	396	1584.60	5.23e-2
17	409	1603.84	4.68e-2
18	388	1572.76	4.58e-2
19	415	1612.72	3.56e-2
20	435	1642.32	3.02e-2
21	411	1606.80	2.88e-2
22	413	1609.76	2.74e-2
23	292	1430.68	2.72e-2

Dos 23 canais selecionados para a recuperação de  $CO_2$  de acordo com a sua contribuição para os graus de liberdade total para o sinal, 20 deles são idênticos aos selecionados de acordo com seu conteúdo de informação, como mostram as tabelas 7.5 e 7.3. Eles são quase indistinguíveis dos selecionados pelo conteúdo de informação, o que indica que o critério de seleção não é crítico. Também é mostrado na tabela os graus de liberdade para o sinal do gás. Fato semelhante ocorre quando analisamos os dados referentes ao  $CH_4$ . Dos 25 primeiros canais elencados de acordo com o seu conteúdo de informação e graus de liberdade 14 deles são comuns aos dois métodos como mostram as tabelas 7.6 e 7.4.

Tabela 7.6 - Os primeiros 21 canais selecionados para recuperação de  $CH_4$  com maior número de graus de liberdade.

	Número do Canal	Localização (nm)	Graus de Liberdade
1	257	1378.88	8.50e-2
2	439	1648.24	3.98e-2
3	461	1680.80	6.21e-4
4	427	1630.48	2.93e-4
5	451	1666.00	4.11e-5
6	269	1396.64	1.34e-5
7	489	1722.24	7.62e-6
8	477	1704.48	4.53e-6
9	469	1692.64	4.40e-6
10	479	1707.44	3.46e-6
11	480	1708.92	3.28e-6
12	470	1694.12	2.03e-6
13	492	1726.68	1.76e-6
14	491	1725.20	1.65e-6
15	245	1361.12	1.59e-6
16	500	1738.52	7.44e-7
17	432	1637.88	7.09e-7
18	505	1745.92	6.95e-7
19	424	1626.04	5.66e-7
20	474	1700.04	5.30e-7
21	429	1633.44	4.20e-7

## 7.2 Resultados para a Recuperação de Perfis de Concentração de $CO_2$ e $CH_4$

Recuperamos os valores de concentração de metano e dióxido de carbono, graficamos com os perfis correspondentes e tabelamos os resultados para cada um dos cenários expostos na metodologia (ver seção 6.4, 6.4.0.1.5 e 6.4.0.1.4 do capítulo 6). Apresentamos gráficos que confrontaram os perfis de concentração do gás, que foram obtidos através da utilização de conjunto de canais selecionados pelas técnicas OSP, CI e GLS, com o perfil exato conhecido à priori. Os erros RMS dos resultados de simulação para cada gás traço foram calculados através da relação:

$$Erro = \sqrt{\sum_{i=0}^N (C_i^{exato} - C_i^{modelo})^2 / N} \quad (7.2)$$

onde  $N$  é o número de pontos de amostragem do perfil, que varia de  $p_s$  (pressão

na superfície) a  $p_t$  (pressão no topo da atmosfera (TOA)). Todas as situações atmosféricas utilizadas durante as fases de aprendizagem e teste são descritas por 50 pontos de medidas de concentração atmosférica (níveis LOWTRAN até 0.000036 hPa ou 120 km de altura) e as correspondentes radiâncias selecionadas calculadas pela SCIATRAN.

Para resolver o problema da recuperação da concentração de  $CO_2$  e de  $CH_4$  foi desenvolvido um algoritmo, escrito em linguagem Java, que implementa uma RNA Perceptron Multi-Camadas ((HAYKIN, 2001) e (BISHOP, 1995)). Cada uma dessas redes foi treinada utilizando-se o algoritmo *backpropagation* para busca dos pesos da rede. Ao final dessa etapa teremos como resultado as redes treinadas e prontas para serem usadas no processo de recuperação do perfil original desejado. A entrada da rede será composta por um vetor contendo as radiâncias captadas pelo sensor em cada canal (comprimento de onda) selecionado e a saída corresponderá a um vetor contendo as concentrações do gás em cada nível considerado (perfil). Todos os experimentos foram conduzidos sob o sistema operacional Linux, em um microcomputador com processador AMD Athlon (tm) 64 Processor 3200 +, 1,53 GHz e 1 GB MB de RAM. Os resultados dos testes e a análise de cada alternativa proposta são apresentados abaixo.

### 7.2.0.1 Cenário 1: Inversão global

Nesse cenário utilizamos os subgrupos de treinamento, validação e teste contendo dados de radiância e perfis atmosféricos do gás coletados pela metodologia exposta na seção 6.2 do capítulo 6. Esses dados englobam perfis completos, com 50 pontos de medição de concentração em todas as latitudes do globo terrestre, sejam elas tropicais, temperadas ou subárticas e em 4 camadas atmosféricas: troposfera, estratosfera, mesosfera e termosfera. A distribuição dos pontos para cada camada atmosférica é diferente para cada faixa latitudinal do globo e foi resumida na tabela 7.7 abaixo.

Inicialmente aplicamos a metodologia de otimização dos parâmetros de rede exposta na seção 6.3 do capítulo anterior para encontrarmos as melhores opções de rede para esse cenário em particular. Realizamos 3 baterias de experimentos para a recuperação de cada gás traço com a possibilidade de escolha entre quatro funções de ativação distintas: Tanh, Sigmóide, Logarítmica e Gaussiana. Em um primeiro experimento atribuímos um peso maior à capacidade da RNA de aprender os padrões que lhe são apresentados em detrimento da sua capacidade de generalização.

Tabela 7.7 - Distribuição dos níveis de altitude/concentração para cada camada atmosférica e região climática do perfil atmosférico.

Tipo de Perfil	Camada	Nível de Altitude (km)
Tropical	Troposfera	<i>0,1,2,...,18</i>
	Estratosfera	<i>19,20,...,25,27.5,30,32.5,35,37.5,40,42.5,45,47.5,50</i>
	Mesosfera	<i>55,60,65,70,75,80,85,90</i>
	Termosfera	<i>95,100,105,110,115,120</i>
Lat Média Inverno	Troposfera	<i>0,1,2,...,9</i>
	Estratosfera	<i>10,11,...,25,27.5,30,32.5,35,37.5,40,42.5,45,47.5,50</i>
	Mesosfera	<i>55,60,65,70,75,80,85,90</i>
	Termosfera	<i>95,100,105,110,115,120</i>
Lat Média Verão	Troposfera	<i>0,1,2,...,9</i>
	Estratosfera	<i>10,11,...,25,27.5,30,32.5,35,37.5,40,42.5,45,47.5,50</i>
	Mesosfera	<i>55,60,65,70,75,80,85,90</i>
	Termosfera	<i>95,100,105,110,115,120</i>
Subártico Inverno	Troposfera	<i>0,1,2,...,7</i>
	Estratosfera	<i>8,9,...,25,27.5,30,32.5,35,37.5,40,42.5,45,47.5,50</i>
	Mesosfera	<i>55,60,65,70,75,80,85,90</i>
	Termosfera	<i>95,100,105,110,115,120</i>
Subártico Verão	Troposfera	<i>0,1,2,...,7</i>
	Estratosfera	<i>8,9,...,25,27.5,30,32.5,35,37.5,40,42.5,45,47.5,50</i>
	Mesosfera	<i>55,60,65,70,75,80,85,90</i>
	Termosfera	<i>95,100,105,110,115,120</i>

No segundo experimento o erro de aprendizagem e o erro de generalização da RNA tiveram pesos iguais na composição da função objetivo. Por último, emulamos um cenário mais realístico, no qual a capacidade de generalizar, ou seja, encontrar um valor de saída correto para um padrão desconhecido para a RNA, possui maior peso do que a capacidade de aprendizagem do modelo. O terceiro experimento resultou em soluções mais acuradas e precisas (erros rms cerca de 20-30% menor) com arquitetura de rede menos complexas em relação aos outros experimentos. As Tabelas 7.8 e 7.9 exibem os parâmetros correspondentes à melhor solução encontrada para um determinado número de avaliações da função objetivo.

A análise dos resultados obtidos neste trabalho (ver seção 7.2) evidenciaram as virtudes de se utilizar uma abordagem evolutiva automatizada na busca ótima de parâmetros de uma RNA. Foram geradas arquiteturas com grande capacidade de generalização para os dados de entrada da rede e com baixo custo computacional, ou seja, com número reduzido de conexões ou pesos para ajuste, como por exemplo nos experimentos 1 e 11 da tabela 7.8 e 1, 9 e 12 da tabela 7.9.

Tabela 7.8 - Melhores arquiteturas de RNA para recuperação de  $CO_2$  atmosférico encontradas pelos algoritmos VNS, SA, GEO e AG para 5000 avaliações da função objetivo. Nesses experimentos os fatores de ponderação atribuídos ao erros de treinamento ( $\rho_1$ ) e generalização ( $\rho_2$ ) foram de 0.1 e 1.0, respectivamente.

Exp	Conjunto de Canais	Metaheurística	Arquitetura	$\alpha$	$\eta$	$\phi(\cdot)$	$F_{obj}$
#1	OSP	VNS	$26 \times 3 \times 3 \times 50$	0.9375	0.5	Gauss	4.946
#2	OSP	SA	$26 \times 11 \times 50$	0.140625	0.6875	Tanh	7.1927
#3	OSP	AG	$26 \times 6 \times 14 \times 50$	0.25	0.1875	Tanh	6.3429
#4	OSP	GEO	$26 \times 4 \times 4 \times 50$	0.296875	0.125	Logarítmica	5.7898
#5	CI	VNS	$29 \times 3 \times 3 \times 50$	0.671875	0.8125	Gauss	7.5530
#6	CI	SA	$29 \times 4 \times 6 \times 50$	0.34375	0.125	Sigmóide	6.8903
#7	CI	AG	$29 \times 16 \times 50$	0.25	0.375	Tanh	5.9211
#8	CI	GEO	$29 \times 6 \times 11 \times 50$	0.265625	0.125	Logarítmica	8.2450
#9	GLS	VNS	$23 \times 12 \times 30 \times 50$	0.34375	0.1875	Tanh	6.6640
#10	GLS	SA	$23 \times 3 \times 5 \times 50$	0.25	0.375	Logarítmica	6.1132
#11	GLS	AG	$23 \times 7 \times 7 \times 50$	0.015625	0.875	Tanh	5.4770
#12	GLS	GEO	$23 \times 6 \times 10 \times 50$	0.109375	0.625	Sigmóide	6.0378

Tabela 7.9 - Melhores arquiteturas de RNA para recuperação de  $CH_4$  atmosférico encontradas pelos algoritmos VNS, SA, GEO e AG para 5000 avaliações da função objetivo. Nesses experimentos os fatores de ponderação atribuídos ao erros de treinamento ( $\rho_1$ ) e generalização ( $\rho_2$ ) foram de 0.1 e 1.0, respectivamente.

Exp	Conjunto de Canais	Metaheurística	Arquitetura	$\alpha$	$\eta$	$\phi(\cdot)$	$F_{obj}$
#1	OSP	VNS	$21 \times 15 \times 8 \times 50$	0.796875	0.9375	Sigmóide	0.0324
#2	OSP	SA	$21 \times 21 \times 1 \times 50$	0.484375	0.5625	Sigmóide	0.0450
#3	OSP	AG	$21 \times 14 \times 10 \times 50$	0.046875	0.75	Gauss	0.0355
#4	OSP	GEO	$21 \times 13 \times 50$	0.0625	0.5	Logarítmica	0.0534
#5	CI	VNS	$25 \times 7 \times 19 \times 50$	0.03125	0.375	Sigmóide	0.047
#6	CI	SA	$25 \times 17 \times 12 \times 50$	0.640625	0.625	Sigmóide	0.0392
#7	CI	AG	$25 \times 3 \times 5 \times 50$	0.0156250	0.75	Tanh	0.0369
#8	CI	GEO	$25 \times 6 \times 2 \times 50$	0.140625	0.1875	Tanh	0.0571
#9	GLS	VNS	$21 \times 5 \times 50$	0.1875	0.0625	Sigmóide	0.0311
#10	GLS	SA	$21 \times 3 \times 2 \times 50$	0.390625	0.75	Gauss	0.0396
#11	GLS	AG	$21 \times 23 \times 6 \times 50$	0.890625	0.0	Tanh	0.0393
#12	GLS	GEO	$21 \times 13 \times 50$	0.765625	0.625	Logarítmica	0.0295

A melhor arquitetura de rede encontrada para recuperar o dióxido de carbono foi obtida após 5000 iterações do algoritmo VNS e corresponde aos parâmetros de rede exibidos no primeiro experimento da tabela 7.8. A solução aponta para uma arquitetura de rede com duas camadas ocultas e poucos neurônios em cada uma delas (3 e 3, respectivamente) e sugere a aplicação de uma função de ativação do tipo Gaussiana nas camadas intermediárias. Curiosamente uma solução similar foi obtida atribuindo-se pesos equivalente tanto para treinamento quanto para generalização da rede. Para o metano a melhor solução (experimento 12 da tabela 7.9) foi obtida através da aplicação do algoritmo GEO e consistiu de uma RNA com apenas uma camada intermediária com 13 neurônios e função de ativação do tipo logarítmica.

Outras soluções igualmente interessantes do ponto de vista do custo computacional requerido e desempenho oferecido foram obtidas nos experimento 4 da tabela 7.8 e nos experimentos 3 e 7 da tabela 7.9. Ademais constatamos que os fatores de ponderação atribuídos ao erro de treinamento e generalização têm influência na solução de arquitetura final obtida pelo algoritmo de busca. Embora tenhamos omitido os resultados obtidos para as outras configurações dos fatores de ponderação  $\alpha$  e  $\beta$ , eles eram, em média, piores do que os que foram exibidos nas tabelas acima. Isso se explica devido ao fato de atribuirmos um peso maior ao erro de treinamento em detrimento do erro de generalização na composição da função objetivo. Na prática, o que realmente importa para uma rede neural é a sua capacidade de generalizar sobre padrões desconhecidos e não sobre aqueles padrões para os quais a rede foi treinada.

Comparativamente, podemos inferir dos resultados que as arquiteturas produzidas pelas metaheurísticas VNS e GA foram ligeiramente superiores às outras embora a melhor solução obtida para a recuperação do metano tenha sido obtida através do algoritmo *GEO*. A razão entre a melhor e a pior solução obtida nesses experimentos, levando-se em consideração o mesmo número de avaliações da função objetivo em cada caso, foi para o metano de 0.94, 0.75 e 0.52 para 100, 1000 e 5000 avaliações e para o  $CO_2$  de 1.04, 0.82 e 0.6 para 100, 1000 e 5000 avaliações, respectivamente. Esses valores revelam outro fato interessante a respeito da dinâmica evolutiva do modelo proposto: com o aumento no número de avaliações de função objetivo alguns algoritmos ou metaheurísticas tendem a se sobressair em relação aos demais.

A partir da análise dos resultados apresentados nas tabelas acima foi possível selecionar dentre as soluções geradas por cada metaheurística as que produziram perfis suaves e com menor erro residual. Desta forma conseguimos restringir as opções de soluções encontradas para cada conjunto de canais do sensor e que foram efetivamente utilizadas para reconstruir os perfis definitivos desse cenário. A Tabela 7.10 apresenta as combinações precisas dos parâmetros utilizados em cada simulação para o conjunto de dados de  $CO_2$  e  $CH_4$ . Para cada conjunto de canais espectrais pré-definidos (OSP, GLS e CI) é realizada a recuperação do perfil vertical do gás correspondente. As Figuras 7.11(a)- 7.19(b), exibem os resultados dos testes de generalização do modelo para o cenário de inversão global em comparação com os dados de radiância sintético (aqui chamado de perfil real) para  $CO_2$  e  $CH_4$ . Nas figuras 7.20 e 7.21, exibimos um quadro resumo com os erros RMS obtidos por cada RNA nesse

cenário agrupados por camada atmosférica, faixa latitudinal e conjunto de canais selecionados. Nas figuras os índices referem-se a porcentagens relativas em relação ao erro total do agrupamento.

Tabela 7.10 - Arquitetura de rede, obtida pela aplicação da estratégia otimizante, para a recuperação de perfil vertical de concentração de  $CO_2$  e  $CH_4$  agrupada por método de seleção de canais.

Conjunto de canais	Gás	Número de camadas	Neurônios em cada camadas oculta <sup>1</sup>	$\alpha$	$\eta$	$\phi(\cdot)$
OSP	$CO_2$	2	3×3	0.984375	0.5	Gauss
OSP	$CH_4$	1	18	0.484375	0.59375	Sigmóide
CI	$CO_2$	1	30	0.375	0.5	Sigmóide
CI	$CH_4$	1	21	0.25	0.125	Tanh
GLS	$CO_2$	1	25	0.015625	0.875	Sigmóide
GLS	$CH_4$	1	5	0.1875	0.03125	Logarítmica

<sup>1</sup>Notação =  $[x_1 \times x_2 \times \dots \times x_n]$  onde  $x_i$  é igual ao número de neurônios na  $i$ -ésima camada oculta



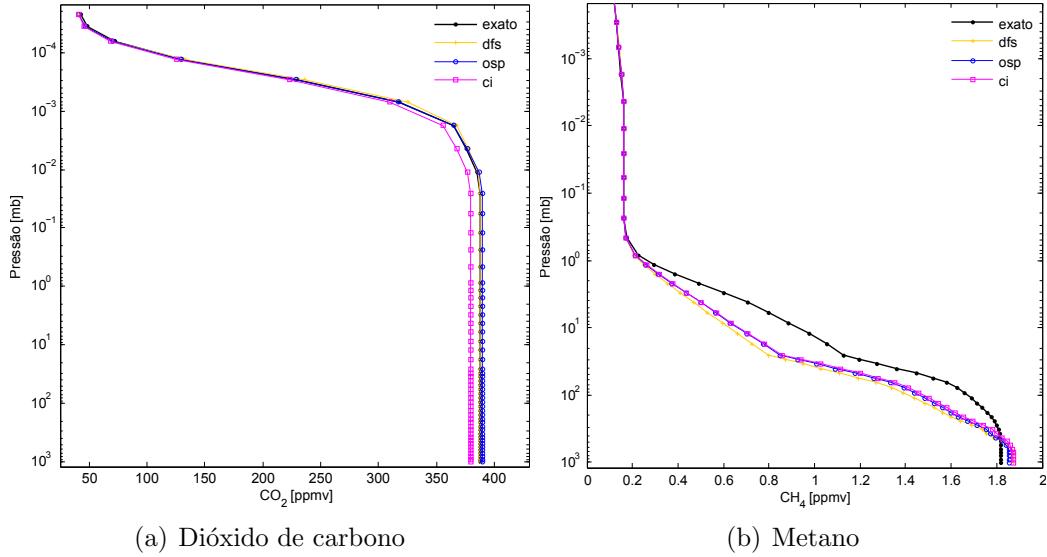


Figura 7.10 - Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude tropical para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	1.8639	mesosfera	OSP	1.7942
troposfera	GLS	0.3722	mesosfera	GLS	0.3726
troposfera	CI	8.2575	mesosfera	CI	8.3437
estratosfera	OSP	1.8635	termosfera	OSP	1.182
estratosfera	GLS	0.3810	termosfera	GLS	4.043
estratosfera	CI	8.2709	termosfera	CI	6.3091

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.11513	mesosfera	OSP	0.003503
troposfera	GLS	0.1421	mesosfera	GLS	0.003637
troposfera	CI	0.10474	mesosfera	CI	0.003693
estratosfera	OSP	0.22498	termosfera	OSP	0.001766
estratosfera	GLS	0.27073	termosfera	GLS	0.003318
estratosfera	CI	0.21487	termosfera	CI	0.002044

(b) Metano

Figura 7.11 - Erro (rms) obtido na recuperação de um perfil (tropical) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

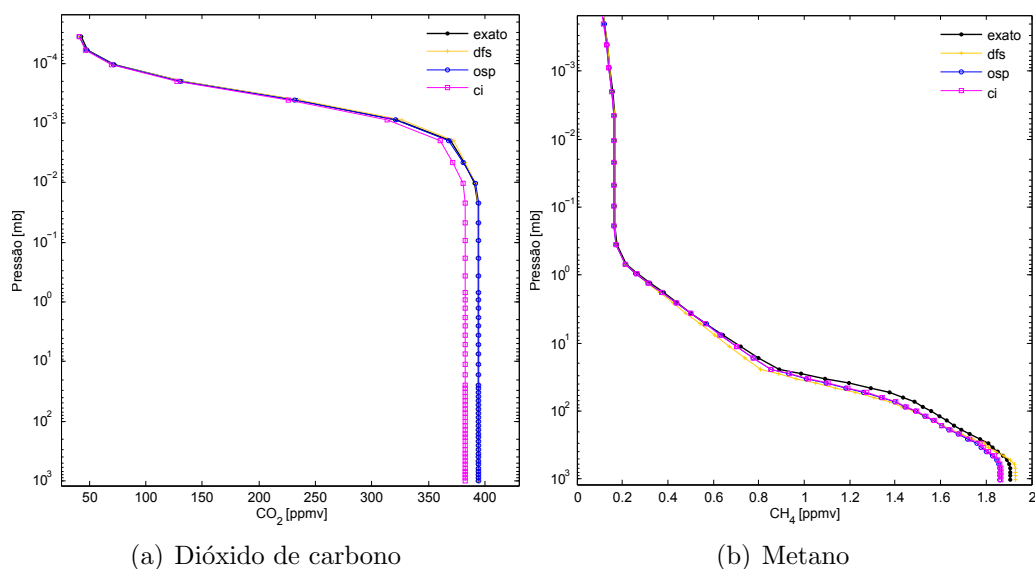


Figura 7.12 - Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude média no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	1.2181	mesosfera	OSP	1.1437
troposfera	GLS	0.3137	mesosfera	GLS	0.3335
troposfera	CI	10.962	mesosfera	CI	10.831
estratosfera	OSP	1.2176	termosfera	OSP	1.4751
estratosfera	GLS	0.3120	termosfera	GLS	2.1347
estratosfera	CI	10.949	termosfera	CI	6.7

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.05396	mesosfera	OSP	0.004605
troposfera	GLS	0.042371	mesosfera	GLS	0.003013
troposfera	CI	0.04872	mesosfera	CI	0.005130
estratosfera	OSP	0.06007	termosfera	OSP	0.004060
estratosfera	GLS	0.088407	termosfera	GLS	0.00167
estratosfera	CI	0.054567	termosfera	CI	0.004510

(b) Metano

Figura 7.13 - Erro (rms) obtido na recuperação de um perfil (latitude média no inverno) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

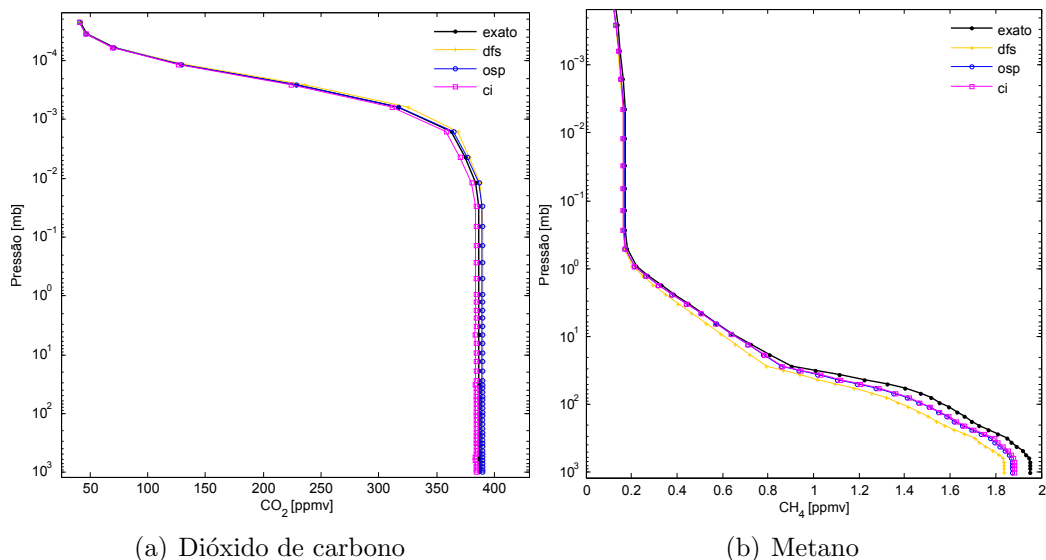


Figura 7.14 - Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude média no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	2.8828	mesosfera	OSP	2.8042
troposfera	GLS	3.3409	mesosfera	GLS	3.3461
troposfera	CI	3.057	mesosfera	CI	3.1963
estratosfera	OSP	2.8823	termosfera	OSP	0.8225
estratosfera	GLS	3.3575	termosfera	GLS	4.8473
estratosfera	CI	3.0607	termosfera	CI	3.8563

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.078407	mesosfera	OSP	0.007303
troposfera	GLS	0.14335	mesosfera	GLS	0.011228
troposfera	CI	0.067486	mesosfera	CI	0.007417
estratosfera	OSP	0.071368	termosfera	OSP	0.006058
estratosfera	GLS	0.13273	termosfera	GLS	0.007159
estratosfera	CI	0.063079	termosfera	CI	0.006209

(b) Metano

Figura 7.15 - Erro (rms) obtido na recuperação de um perfil (latitude média no verão) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

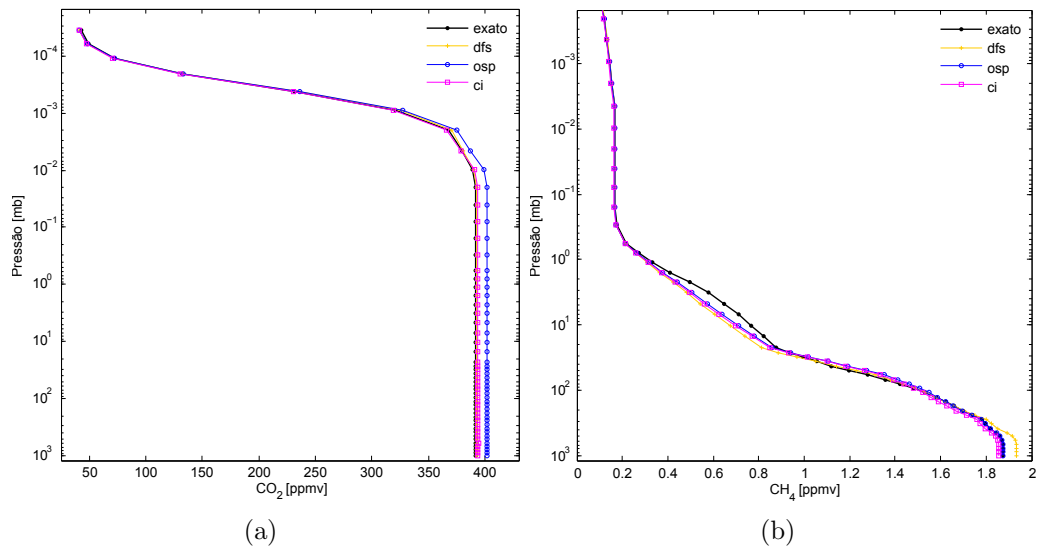


Figura 7.16 - Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude subártica no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	10.21	mesosfera	OSP	10.152
troposfera	GLS	1.012	mesosfera	GLS	1.017
troposfera	CI	2.4139	mesosfera	CI	2.328
estratosfera	OSP	10.21	termosfera	OSP	4.7587
estratosfera	GLS	1.0262	termosfera	GLS	2.8915
estratosfera	CI	2.389	termosfera	CI	1.5913

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.0123	mesosfera	OSP	0.001
troposfera	GLS	0.038294	mesosfera	GLS	0.000885
troposfera	CI	0.021991	mesosfera	CI	0.0034537
estratosfera	OSP	0.0521	termosfera	OSP	0.001192
estratosfera	GLS	0.06	termosfera	GLS	0.002068
estratosfera	CI	0.053135	termosfera	CI	0.003018

(b) Metano

Figura 7.17 - Erro (rms) obtido na recuperação de um perfil (latitude subártica no inverno) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

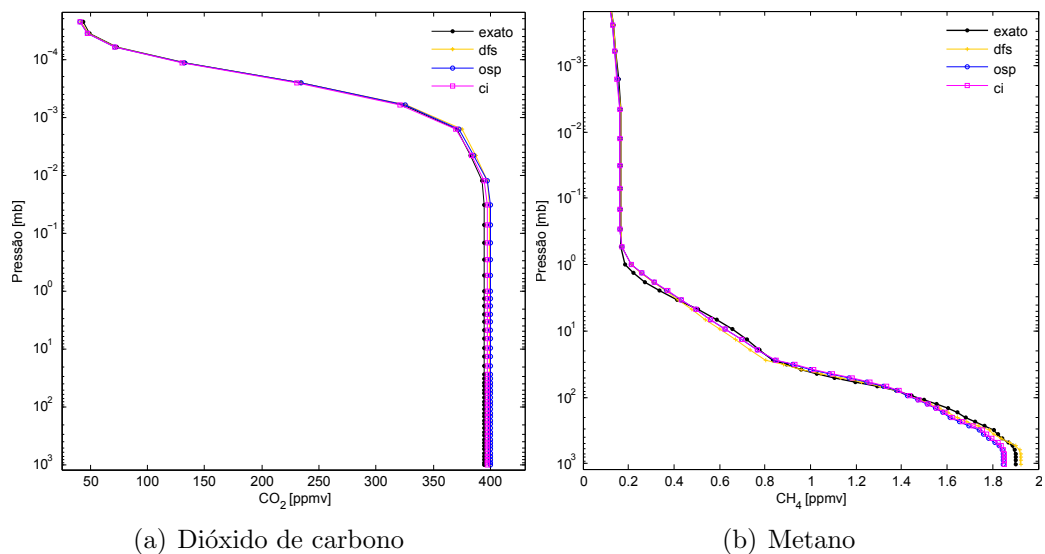


Figura 7.18 - Resultados para a recuperação da taxa de mistura (em ppmv) em região de latitude subártica no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	5.2945	mesosfera	OSP	5.2151
troposfera	GLS	4.6291	mesosfera	GLS	4.6604
troposfera	CI	2.6567	mesosfera	CI	2.5379
estratosfera	OSP	5.2941	termosfera	OSP	1.8178
estratosfera	GLS	4.6223	termosfera	GLS	2.6497
estratosfera	CI	2.6564	termosfera	CI	2.0404

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.05719	mesosfera	OSP	0.004946
troposfera	GLS	0.025611	mesosfera	GLS	0.002960
troposfera	CI	0.047942	mesosfera	CI	0.005232
estratosfera	OSP	0.033474	termosfera	OSP	0.004588
estratosfera	GLS	0.030698	termosfera	GLS	0.001446
estratosfera	CI	0.039696	termosfera	CI	0.004864

(b) Metano

Figura 7.19 - Erro (rms) obtido na recuperação de um perfil (latitude subártica no verão) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

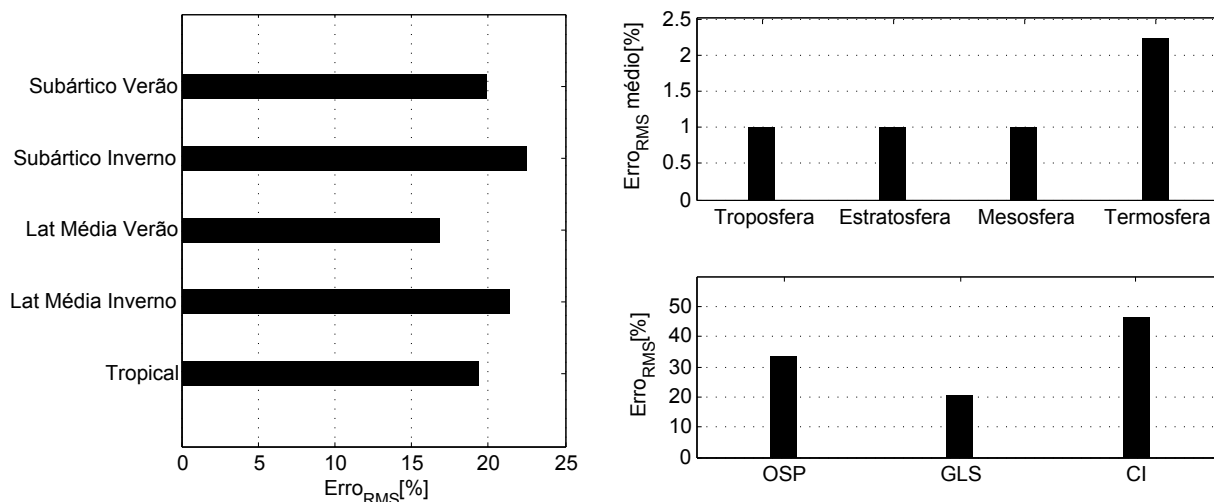


Figura 7.20 - Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do  $CO_2$  para o cenário 1 (inversão global ou generalizada). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica.

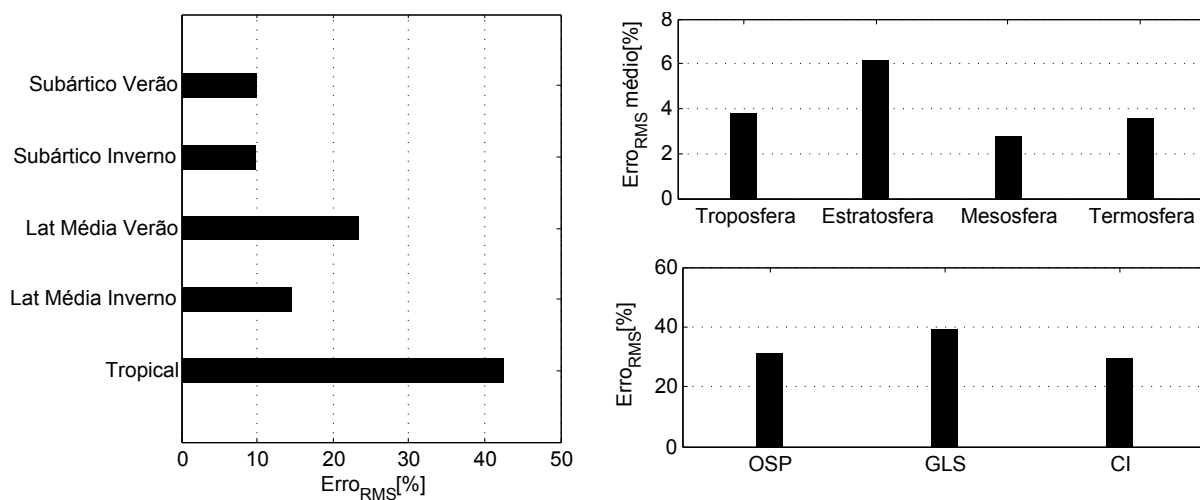


Figura 7.21 - Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do  $CH_4$  para o cenário 1 (inversão global ou generalizada). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica.

As Figuras 7.20 e 7.21 nos forneceu informações a respeito da distribuição do erro quadrático médio para a recuperação de  $CO_2$  e  $CH_4$  pelo modelo baseado em RNA. Os conjuntos de canais OSP e CI sobressairam quando recuperamos  $CH_4$  e o GLS quando recuperamos o  $CO_2$ . Para perfis tropicais de metano o erro residual foi maior e respondeu sozinho por quase 40% do erro total acumulado. Entretanto para perfis de dióxido de carbono os erros tiveram boa distribuição entre perfis de diferentes latitudes. Nesse cenário no qual utilizamos uma única rede neural otimizada para cada conjunto de canais, recuperamos perfis de  $CO_2$  com um erro quadrático médio que oscilou entre  $\approx 2.2\%$  (acima de 90 km de altitude) e  $\sim 1\%$  ( $0 \mapsto 90$  km). Porém, na média, o resíduo se manteve abaixo do limiar mínimo aceitável de 1% ou  $\approx 4$  ppmv para a troposfera, estratosfera e parte da mesosfera. Acima de 80 km esse erro aumenta consideravelmente atingindo um bias médio acima de 2% ou  $\approx 9$  ppmv. Para o metano o erro quadrático médio aumentou e ficou bem próximo dos 4% na troposfera, atingiu elevados 6% na estratosfera e permaneceu em  $\sim 4\%$  nas demais camadas. Importante ressaltar que esses valores residuais tornariam inviável a utilização desse modelo, na forma em que se encontra, para delimitarmos e quantificarmos as fontes e sumidouros de metano em regiões continentais (ver seção 2.1 do capítulo 2). Para o dióxido de carbono seria possível, com alguns ajustes de escala pontuais, a recuperação de concentração próximo à superfície terrestre com relativa acurácia e precisão.

### 7.2.1 Cenário 2: Inversão localizada distribuída por camadas atmosféricas

Por último, criamos um cenário híbrido, no qual treinamos uma rede neural para cada camada atmosférica e faixa de latitude terrestre. Para esse cenário existem diversas combinações possíveis de redes. Podemos ter por exemplo, uma rede neural específica para perfis tropicais estratosféricos, outra para perfis subárticos de inverno troposféricos, etc. Novamente procuramos encontrar a melhor opção arquitetural possível para cada subgrupo de treinamento através da utilização da estratégia de otimização implementada e discutida nas seções anteriores. As Tabelas 7.11 e 7.12 apresentam as combinações precisas dos parâmetros utilizados em cada simulação para  $CO_2$  e  $CH_4$ . Para esse cenário, por questão de clareza e síntese e pelo fato de existirem muitas combinações possíveis de topologias de redes, exibimos apenas as topologias de RNAs obtidas para um perfil tropical. Nas figuras 7.32 e 7.33 exibimos um quadro resumo com os erros RMS obtidos por cada RNA nesse cenário agrupados

por camada atmosférica, massa de ar e conjunto de canais selecionados. Nas figuras os índices referem-se a porcentagens relativas em relação ao erro total do agrupamento.

Tabela 7.11 - Arquitetura de rede obtida para a recuperação de perfil vertical de concentração de  $CO_2$  no cenário 2 (Inversão localizada distribuída) agrupada por método de seleção de canais e camada atmosférica.

Conjunto de canais	Camada Atmosférica	Número de camadas	Neurônios em cada camadas oculta <sup>1</sup>	$\alpha$	$\eta$	$\phi(\cdot)$
OSP	Troposfera	2	4-4	0.296875	0.125	Logarítmica
OSP	Estratosfera	2	15-8	0.796875	0.9375	Tanh
OSP	Mesosfera	2	5-7	0.109375	0.4375	Tanh
OSP	Termosfera	1	25	0.015625	0.9375	Sigmóide
CI	Troposfera	1	16	0.25	0.375	Tanh
CI	Estratosfera	2	17-12	0.75	0.5	Logarítmica
CI	Mesosfera	1	13	0.765625	0.625	Tanh
CI	Termosfera	2	3-1	0.703125	0.6875	Sigmóide
GLS	Troposfera	2	6-10	0.109375	0.625	Gauss
GLS	Estratosfera	2	7-3	0.984375	0.5	Logarítmica
GLS	Mesosfera	2	5-5	0.890625	0.0625	Tanh
GLS	Termosfera	1	22	0.453125	0.0	Tanh

Tabela 7.12 - Arquitetura de rede obtida para a recuperação de perfil vertical de concentração de  $CH_4$  no cenário 2 (Inversão localizada distribuída) agrupada por método de seleção de canais e camada atmosférica.

Conjunto de canais	Camada Atmosférica	Número de camadas	Neurônios em camadas ocultas	$\alpha$	$\eta$	$\phi(\cdot)$
OSP	Troposfera	2	4-4	0.296875	0.125	Tanh
OSP	Estratosfera	2	6	0.1875	0.03125	Tanh
OSP	Mesosfera	2	17	0.046875	0.8125	Sigmóide
OSP	Termosfera	2	2-6	0.796875	0.9375	Tanh
CI	Troposfera	1	3-3	0.25	0.375	Tanh
CI	Estratosfera	1	5	0.1875	0.03125	Sigmóide
CI	Mesosfera	1	6-11	0.0625	0.5	Gauss
CI	Termosfera	1	7-12	0.078125	0.375	Gauss
GLS	Troposfera	2	6-9	0.109375	0.625	Sigmóide
GLS	Estratosfera	1	15	0.765625	0.5625	Sigmóide
GLS	Mesosfera	1	7-5	0.1875	0.875	Tanh
GLS	Termosfera	1	3-5	0.203125	0.8125	Logarítmica

As Figuras 7.23(a) - 7.31(b) abaixo mostram os resultados dos testes de generalização do modelo para o cenário de inversão local em comparação com os dados de radiância sintético (aqui chamado de perfil real).

<sup>1</sup>Notação =  $[x_1 \times x_2 \times \dots \times x_n]$  onde  $x_i$  é igual ao número de neurônios na  $i$ -ésima camada oculta



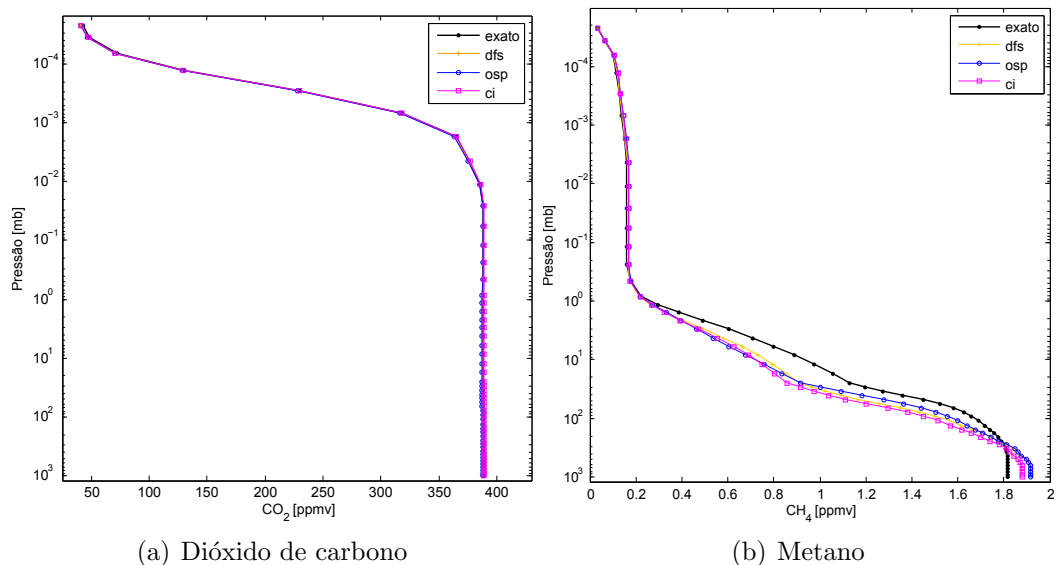


Figura 7.22 - Resultados para a recuperação da taxa de mistura (em ppmv) de  $CO_2$  e  $CH_4$  em região de latitude tropical para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.2231	mesosfera	OSP	0.4510
troposfera	GLS	0.379	mesosfera	GLS	0.0848
troposfera	CI	1.3525	mesosfera	CI	0.6659
estratosfera	OSP	0.7656	termosfera	OSP	1.8675
estratosfera	GLS	0.3090	termosfera	GLS	1.4735
estratosfera	CI	1.1394	termosfera	CI	1.0437

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.075968	mesosfera	OSP	0.009118
troposfera	GLS	0.088212	mesosfera	GLS	0.002894
troposfera	CI	0.093499	mesosfera	CI	0.007037
estratosfera	OSP	0.16196	termosfera	OSP	0.006733
estratosfera	GLS	0.18738	termosfera	GLS	0.001782
estratosfera	CI	0.22815	termosfera	CI	0.005488

(b) Metano

Figura 7.23 - Erro (rms) obtido na recuperação de um perfil (tropical) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

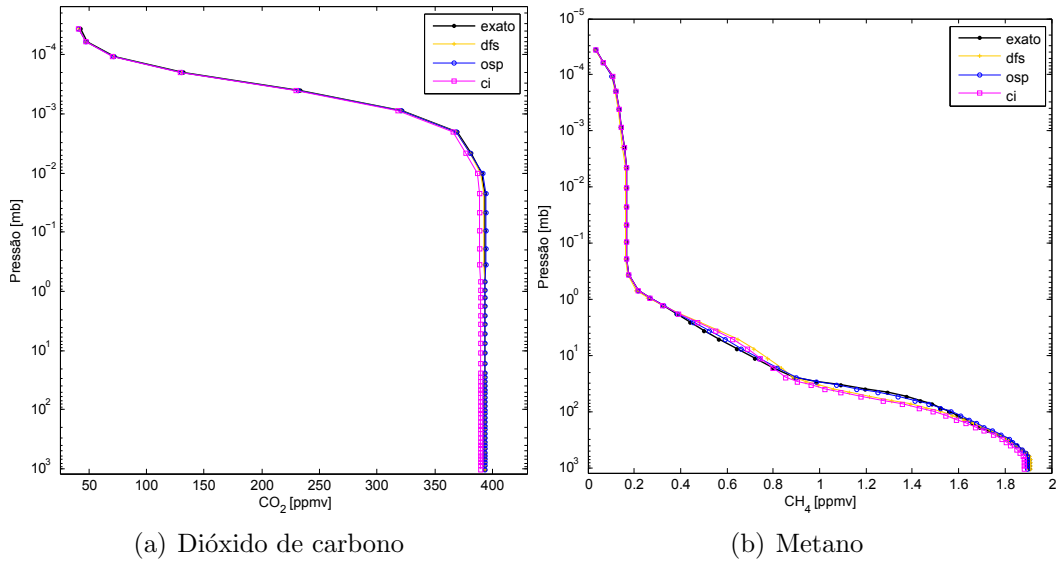


Figura 7.24 - Resultados para a recuperação da taxa de mistura (em ppmv) de  $CO_2$  e  $CH_4$  em região de latitude média no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.1973	mesosfera	OSP	0.8862
troposfera	GLS	0.1469	mesosfera	GLS	0.9128
troposfera	CI	4.0484	mesosfera	CI	4.3679
estratosfera	OSP	0.2201	termosfera	OSP	1.2883
estratosfera	GLS	0.2066	termosfera	GLS	1.3013
estratosfera	CI	4.1965	termosfera	CI	3.1497

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.010776	mesosfera	OSP	0.000748
troposfera	GLS	0.018002	mesosfera	GLS	0.007506
troposfera	CI	0.037689	mesosfera	CI	0.0005
estratosfera	OSP	0.020889	termosfera	OSP	0.001250
estratosfera	GLS	0.084627	termosfera	GLS	0.005503
estratosfera	CI	0.10086	termosfera	CI	0.000576

(b) Metano

Figura 7.25 - Erro (rms) obtido na recuperação de um perfil (latitude média no inverno) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

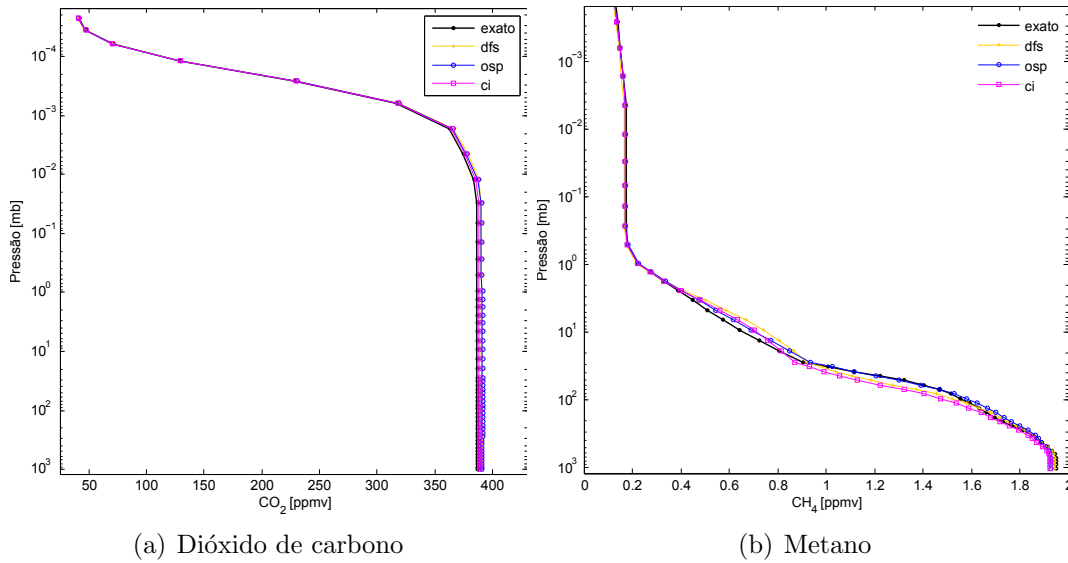


Figura 7.26 - Resultados para a recuperação da taxa de mistura (em ppmv) de  $CO_2$  e  $CH_4$  em região de latitude média no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	4.2359	mesosfera	OSP	4.1321
troposfera	GLS	4.0878	mesosfera	GLS	4.2171
troposfera	CI	2.2029	mesosfera	CI	1.8659
estratosfera	OSP	4.5482	termosfera	OSP	1.5138
estratosfera	GLS	4.2169	termosfera	GLS	2.2644
estratosfera	CI	2.0211	termosfera	CI	0.98197

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.025173	mesosfera	OSP	0.000313
troposfera	GLS	0.013248	mesosfera	GLS	0.007292
troposfera	CI	0.031881	mesosfera	CI	0.00149
estratosfera	OSP	0.026534	termosfera	OSP	0.001133
estratosfera	GLS	0.07695	termosfera	GLS	0.005543
estratosfera	CI	0.096926	termosfera	CI	0.001412

(b) Metano

Figura 7.27 - Erro (rms) obtido na recuperação de um perfil (latitude média no verão) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

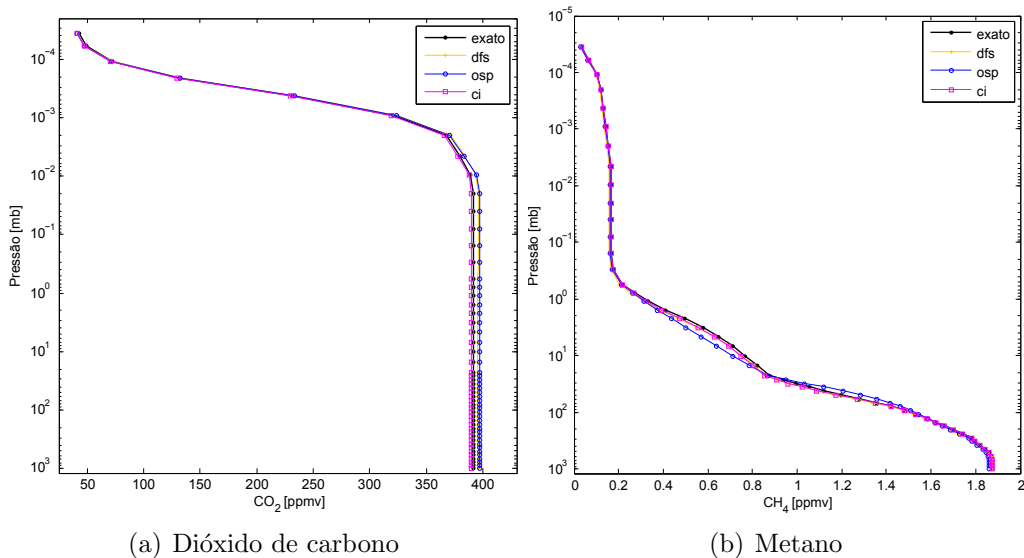


Figura 7.28 - Resultados para a recuperação da taxa de mistura (em ppmv) de  $CO_2$  e  $CH_4$  em região de latitude subártica no inverno para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	5.8644	mesosfera	OSP	5.8115
troposfera	GLS	4.6195	mesosfera	GLS	4.9086
troposfera	CI	1.7275	mesosfera	CI	1.2768
estratosfera	OSP	5.8349	termosfera	OSP	2.3682
estratosfera	GLS	4.4228	termosfera	GLS	2.8868
estratosfera	CI	1.7397	termosfera	CI	1.8246

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.016763	mesosfera	OSP	0.002004
troposfera	GLS	0.01055	mesosfera	GLS	0.008365
troposfera	CI	0.000975	mesosfera	CI	0.000749
estratosfera	OSP	0.056275	termosfera	OSP	0.001973
estratosfera	GLS	0.023221	termosfera	GLS	0.005775
estratosfera	CI	0.020676	termosfera	CI	0.000366

(b) Metano

Figura 7.29 - Erro (rms) obtido na recuperação de um perfil (latitude subártica no inverno) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

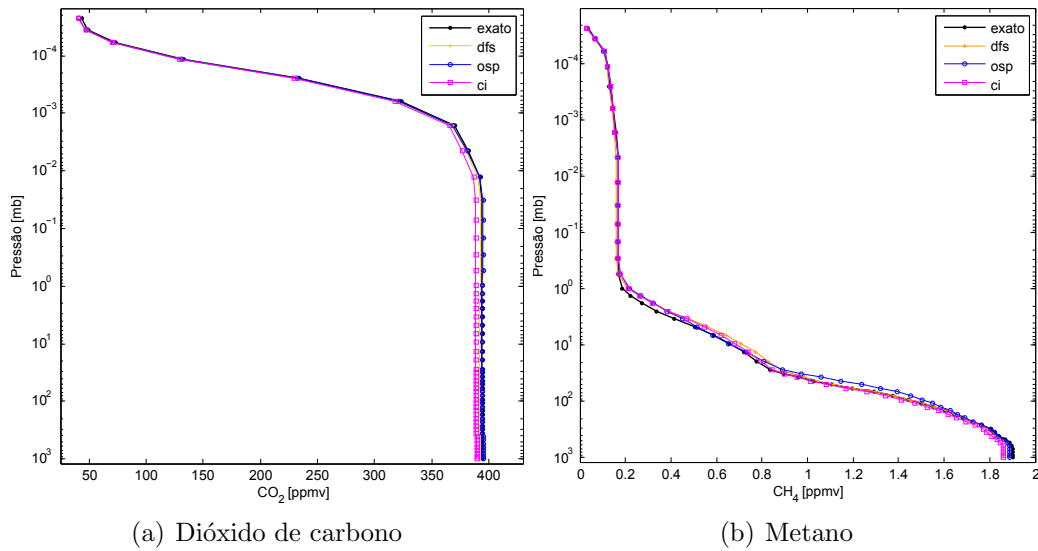


Figura 7.30 - Resultados para a recuperação da taxa de mistura (em ppmv) de  $CO_2$  e  $CH_4$  em região de latitude subártica no verão para dados corrompidos com 1% de ruído gaussiano branco. Utilizamos três subconjuntos de canais SCIAMACHY. (azul) canais obtidos pelo método da sensibilidade ótima (OSP); (vermelho) canais obtidos pelo método dos graus de liberdade (GLS); (amarelo) canais obtidos pelo método do Conteúdo de Informação (CI).

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.8016	mesosfera	OSP	0.4219
troposfera	GLS	1.6561	mesosfera	GLS	1.4495
troposfera	CI	5.4219	mesosfera	CI	5.3654
estratosfera	OSP	0.2882	termosfera	OSP	1.9213
estratosfera	GLS	1.5635	termosfera	GLS	2.282
estratosfera	CI	5.6399	termosfera	CI	4.0199

(a) Dióxido de carbono

Camada atmosférica	Conjunto de canais	Rms (ppmv)	Camada atmosférica	Conjunto de canais	Rms (ppmv)
troposfera	OSP	0.021775	mesosfera	OSP	0.003147
troposfera	GLS	0.010527	mesosfera	GLS	0.007843
troposfera	CI	0.033771	mesosfera	CI	0.002822
estratosfera	OSP	0.07079	termosfera	OSP	0.002033
estratosfera	GLS	0.04016	termosfera	GLS	0.006403
estratosfera	CI	0.031462	termosfera	CI	0.001183

(b) Metano

Figura 7.31 - Erro (rms) obtido na recuperação de um perfil (latitude subártica no verão) vertical de concentração de  $CO_2$  e  $CH_4$  ordenado por estratégia de seleção de canais e camadas atmosféricas.

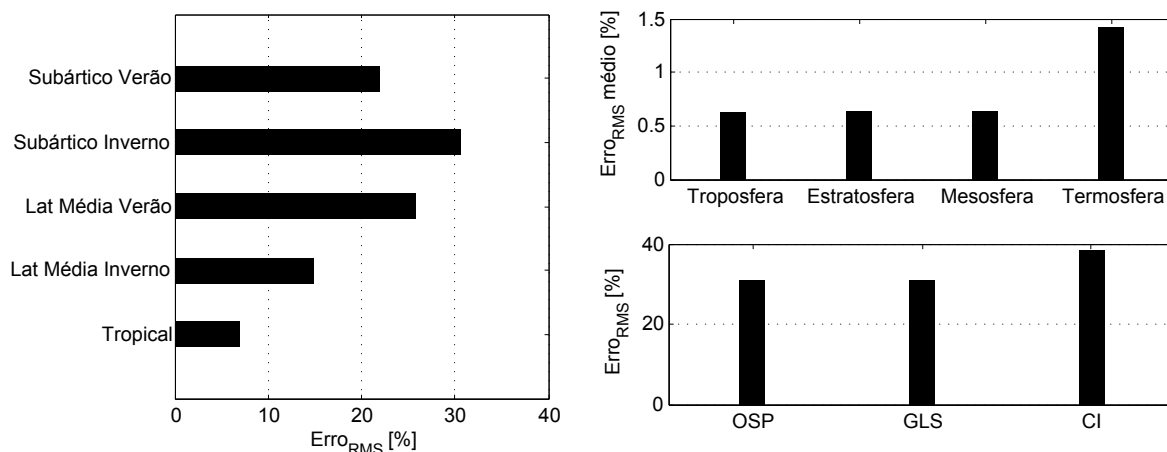


Figura 7.32 - Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do  $CO_2$  para o cenário 2 (inversão Local ou distribuída). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica.

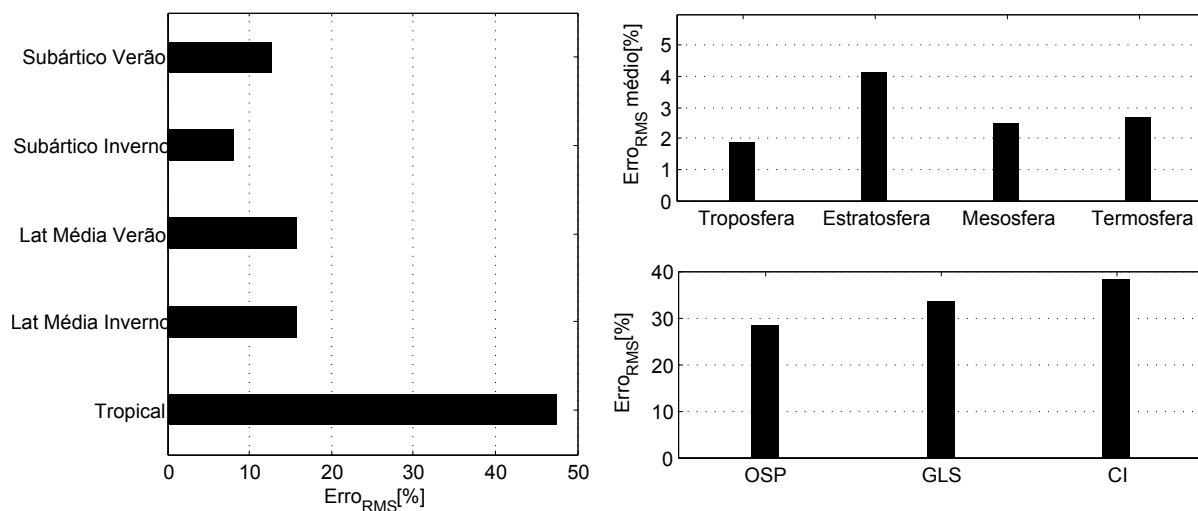


Figura 7.33 - Distribuição (em porcentagem) do erro quadrático médio (RMS) obtido pelo modelo na recuperação do  $CH_4$  para o cenário 2 (inversão Local ou distribuída). Os valores estão expressos em porcentagem relativa agrupados e totalizados por conjunto de canais selecionados (Inferior direito) e faixa latitudinal (Esquerdo). Na parte Superior direita observamos o erro RMS médio experimentado por cada camada atmosférica.

O dióxido de carbono é um gás bem misturado que mantém uma proporção relativa do número de moléculas em relação ao ar seco praticamente constante até a termosfera. A análise dos resultados levando-se em conta essa informação mostrou que, nesse cenário, foi possível recuperar a concentração (taxa de mistura) desse gás com uma acurácia inferior a  $\pm 1\%$  para a troposfera, a estratosfera e para a mesosfera em todas as latitudes testadas. Na termosfera é que obtivemos os piores resultados, com acurácia próxima a 2%. Mesmo assim este resultado não pode ser considerado de todo ruim considerando-se que nenhum dos três conjuntos de canais possui elevado conteúdo de informação quando nos afastamos da estratosfera e atingimos as partes mais altas da atmosfera acima dos 50km de altitude.

Já o metano tem um perfil mais heterogêneo com distribuição mais variável ao longo das camadas. Ainda assim foi possível obter perfis para o metano com erro rms inferior a 1.5% na maioria dos casos. Os perfis de  $CO_2$  e  $CH_4$  obtidos com a rede Perceptron Multi-Camadas (PMC) estão em excelente concordância com o modelo verdadeiro. Os perfis reconstruídos são suaves e apresentam na baixa troposfera erros em relação a um perfil de referência (retirado da base de dados de perfis climatológicos do SCIATRAN) da ordem de  $\pm 1\%$  ( $\pm 3.8$  ppmv em 380 ppmv) para o  $CO_2$  e  $\pm 1.5\%$  ( $\pm 27$  ppbv em 1800 ppbv) para o metano. Podemos considerar este resultado como o erro devido ao próprio modelo da rede neural somado aos erros da observação (ruído do sinal e erros sistemáticos de medida) e ao erro da própria linguagem de programação em que foi implementado o modelo de RNA. Mesmo com dados corrompidos pela adição de 1% de ruído gaussiano os perfis apresentaram uma forma suave e homogênea. Em particular, o desempenho global do modelo para ambos os gases na região da camada limite também é satisfatório apresentando um erro RMS entre  $\pm 0.5\%$  e  $\pm 1\%$ . Um bom nível de acerto com a curva de perfil experimental também é obtido nas camadas superiores (mesosfera e parte da termosfera), mas decresce substancialmente em altitudes mais elevadas (acima de 100 km), onde os erros alcançam um máximo de 3%. Na recuperação de  $CO_2$  utilizando o método de inversão local os conjuntos de canais GLS e OSP obtiveram os melhores resultados com um desempenho similar. Para a recuperação de metano o conjunto OSP foi consideravelmente melhor sobretudo para perfis de latitude média. Para ambos os gases, o conjunto de canais CI obteve um desempenho global um pouco inferior aos demais, embora particularmente para perfis subárticos de  $CH_4$  reconstruídos com esse conjunto de canais os valores rms tenham sido menores.

A distribuição dos erros com relação às camadas atmosféricas ocorreu de tal forma que os pontos com maior erro rms para a recuperação de metano concentraram-se na camada estratosférica e na termosfera para o dióxido de carbono. Especificamente no caso do metano obtivemos bons resultados para a troposfera, que é a camada com maior relevância na maioria dos estudos realizados com gases de efeito estufa. Para essa camada conseguimos recuperar a taxa de mistura do gás com um erro inferior a  $\pm 18$  ppbv ( $\sim 1\%$ ) na grande maioria dos casos. Na literatura existem poucos relatos [ex: (CREVOISIER et al., 2009b) e (XIONG et al., 2008)] de casos de sucesso envolvendo recuperação de metano que apresentassem tais níveis de qualidade condizentes com o requerido para se aferir fontes e sumidouros (ver seção introdutória do capítulo 1).

### 7.3 Comparação entre os dois cenários de inversão propostos

A simples observação e análise gráfica dos resultados alcançados pelos dois cenários já indicava uma redução significativa do erro quadrático médio dos perfis recuperados pelo modelo ao se utilizar o cenário 2 em vez do cenário 1. Para uma comparação quantitativa dos cenários calculamos a razão  $r$  entre o erro quadrático médio (agrupado e totalizado por camada atmosférica, conjunto de canais selecionados e região climática) do cenário 1 e o erro quadrático médio conseguido no segundo cenário. Nas figuras 7.34 e 7.35 abaixo mostramos os valores desse quociente para o dióxido de carbono e para o metano.

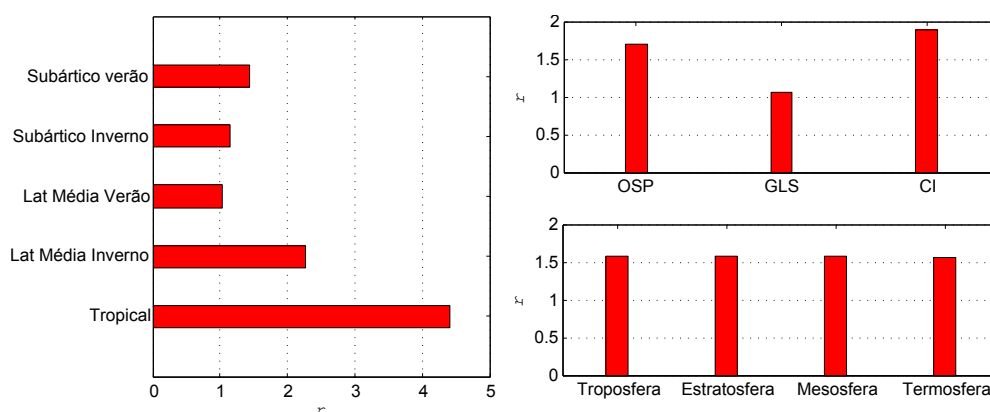


Figura 7.34 - Ganho comparativo (em termos de erro RMS) do cenário 2 (Inversão local distribuída) em relação ao cenário 1 (Inversão generalizada ou global) na recuperação do  $CO_2$ . O quociente  $r$  corresponde à divisão entre o erro RMS obtido no cenário 2 e o erro rms obtido no cenário 1 agrupado e totalizado por camada atmosférica (Inferior direito), conjunto de canais selecionados (Superior direito) e faixa latitudinal (Esquerdo).



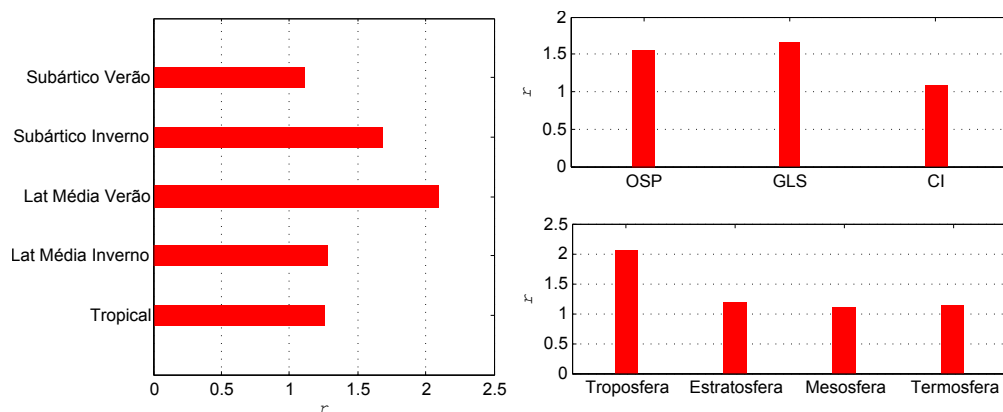


Figura 7.35 - Ganho comparativo (em termos de erro RMS) do cenário 2 (Inversão local distribuída) em relação ao cenário 1 (Inversão generalizada ou global) na recuperação do  $CH_4$ . O quociente  $r$  corresponde à divisão entre o erro RMS obtido no cenário 2 e o erro rms obtido no cenário 1 agrupado e totalizado por camada atmosférica (Inferior direito), conjunto de canais selecionados (Superior direito) e faixa latitudinal (Esquerdo).

Para ambos os gases houve uma melhora significativa na recuperação do perfil. Para o  $CO_2$  percebemos um ganho (diminuição do erro RMS) relativo de 50% em todas as camadas atmosféricas. Para o metano esse ganho foi de 100% na troposfera e oscilou entre 10 a 20% nas demais camadas. Além disso quando consideramos a totalização dos erros acumulados para os 3 conjuntos de canais selecionados do sensor tivemos redução em cerca de 50% para a recuperação de  $CO_2$  usando os conjuntos OSP e CI e de 50% com os conjuntos OSP e GLS para o  $CH_4$ . Por último apontamos também diminuição do resíduo para perfis em determinadas faixas climáticas chegando em alguns casos a cair pela metade, como para os perfis de Latitude Média no Inverno de  $CO_2$  e Latitude Média de Verão para o Metano. O maior ganho ocorreu para perfis tropicais de  $CO_2$  diminuindo em até 400% o resíduo total. Com base nos resultados de inversão para os cenários estudados concluímos que a abordagem da “divisão e conquista” aplicada ao cenário de inversão local se mostrou eficiente nesse problema em particular produzindo perfis mais homogêneos, acurados e precisos. Embora seja necessário um trabalho adicional para otimizar e treinar um número maior de redes, esse esforço é compensado pela redução significativa nos erros residuais para os perfis recuperados. Devemos ressaltar que cada ppmv de concentração é importante quando logramos recuperar concentrações de gases com reduzida variabilidade sazonal e geográfica como é o caso do  $CO_2$  e do  $CH_4$ .

## 7.4 Validação do Modelo Proposto

Afim de avaliarmos a eficiência do procedimento desenvolvido e validarmos as concentrações recuperadas pelo modelo contra observações independentes de  $CO_2$  e  $CH_4$ , realizamos uma série de experimentos para uma amostra de mais de 24 grânulos de medidas reais do SCIAMACHY. Cada uma dessas medidas é realizada em uma faixa de passagem do satélite ENVISAT sobre as áreas de estudo e cada pixel de interesse é livre de contaminação por nuvens ou aerossóis. Inicialmente investigamos a disponibilidade de observações *in situ* independentes para a área de interesse durante o período de funcionamento do SCIAMACHY/ENVISAT, ou seja, a partir do ano de 2002. Verificou-se ser apropriado utilizar os resultados das medições de amostras de ar coletadas em frascos de vidro que foram realizadas pela divisão de monitoramento global da NOAA. Essas medições foram realizadas no observatório de Mauna Loa, no Havaí (latitude  $19.54^\circ N$ , longitude  $155.58^\circ W$ ) à altitude de 3397 m e no Pólo Sul (latitude  $-89.9800^\circ S$ , longitude  $-24.80^\circ W$ ) à altitude de 2800 m acima do nível do mar. A acurácia absoluta das medidas é estimada em 0.2% para o  $CO_2$  e 0.4% para o  $CH_4$ . Maiores detalhes sobre as medições *in situ* e sobre as duas áreas de estudo podem ser encontrados em (DLUGOKENCKY E. J.; LANG, 2010) e (CONWAY T. J.; LANG, 2010).

No primeiro estudo realizado recuperamos os perfis de metano e dióxido de carbono, de acordo com a estratégia de inversão local e distribuída por camadas do cenário 2, com o auxílio dos 3 conjuntos de canais pré-selecionados (OSP, GLS e CI) e comparamos com um dado observacional *in situ*. O resultado é apresentado nas Figuras 7.36-7.38.

Em um segundo experimento realizamos a recuperação para um ponto da superfície apenas e não para todos os 50 pontos (perfil atmosférico) como no primeiro experimento. Os dados são distribuídos uniformemente por todo o ano de 2005 e 2007, respectivamente e correspondem a um conjunto de 12 observações para cada local. Nas figuras 7.40, 7.41, 7.42 e 7.43, é feita uma comparação entre os dados obtidos pelo modelo de inversão proposto e os dados *in situ* coletados no observatório. Pelo fato das estimativas realizadas com os 29 canais mais sensíveis ao  $CO_2$  e com os 25 canais mais sensíveis ao  $CH_4$  que foram selecionados pelo método do conteúdo de informação (CI) terem apresentado melhor desempenho, comparado com os dados observados, os resultados que seguem serão apresentados tomando como referência esse subconjunto de canais.

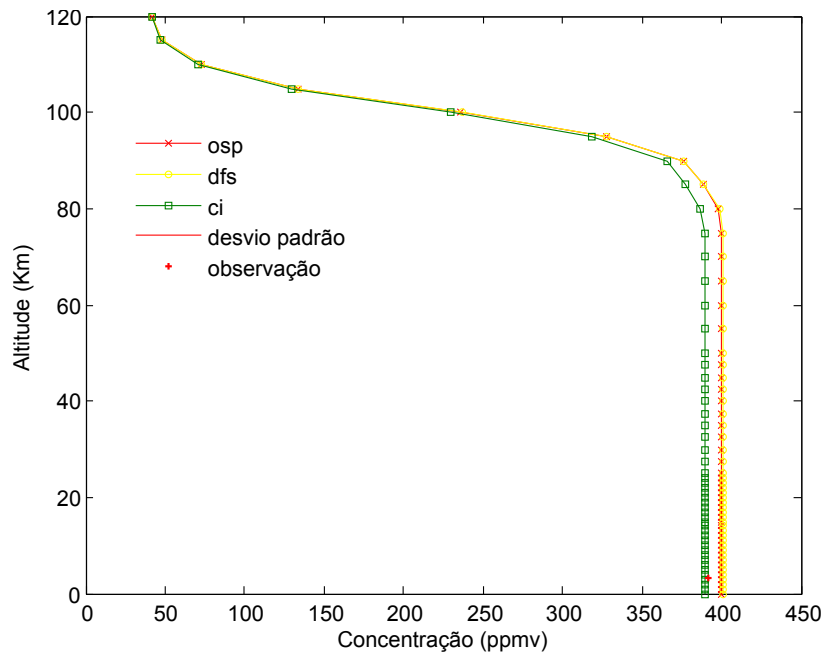


Figura 7.36 - Perfis verticais de  $CO_2$  (ppbv) para o dia 11 de outubro de 2005 sobre Mauna Loa (Havaí), usando três subconjuntos de canais SCIAMACHY no modelo inverso, juntamente com dados observados e o desvio padrão (erro de medida).

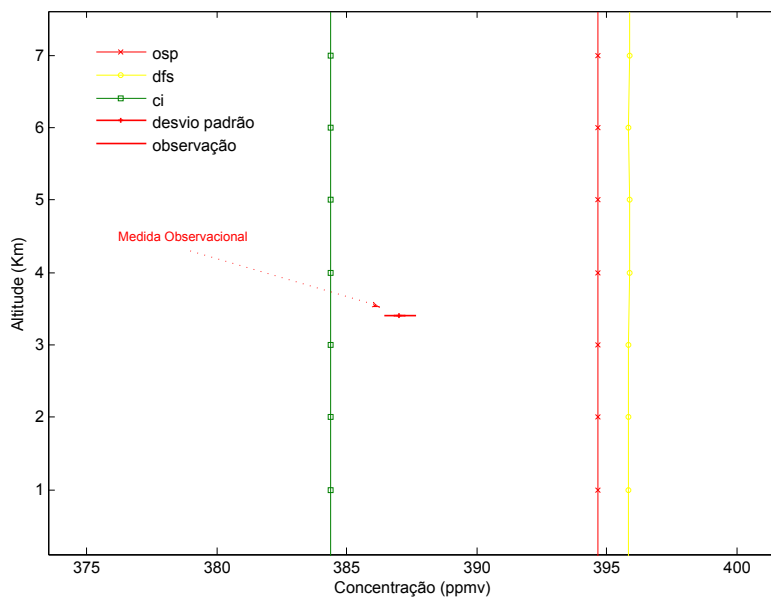


Figura 7.37 - Ampliação da figura 7.36 mostrando detalhes dos perfis recuperados pelos 3 conjuntos de canais do sensor SCIAMACHY juntamente com dados observados e o desvio padrão (erro de medida).

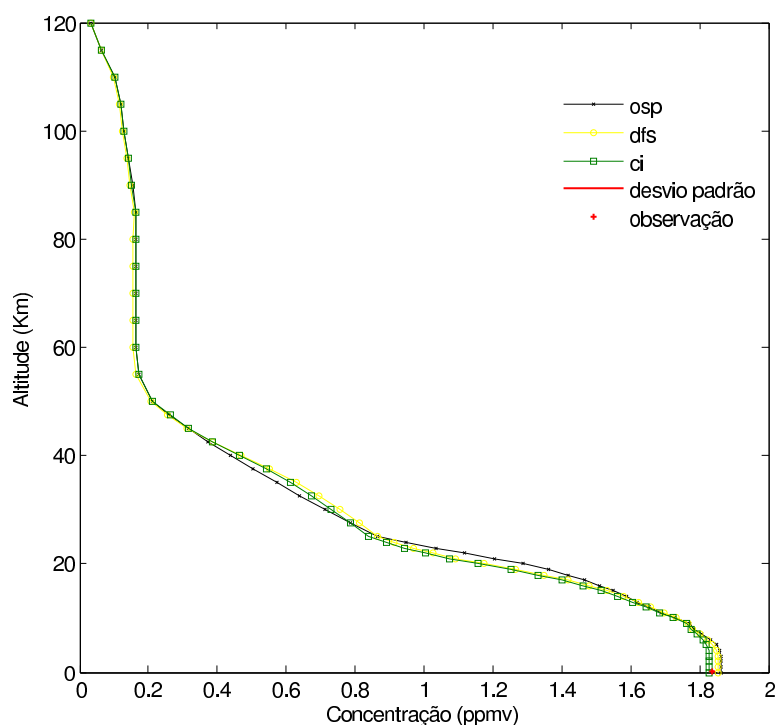


Figura 7.38 - Perfis verticais de  $CH_4$  (ppbv) para o dia 11 de outubro de 2005 sobre Mauna Loa (Havaí), usando três subconjuntos de canais SCIAMACHY no modelo inverso, juntamente com dados observados e o desvio padrão (erro de medida).

As Figuras 7.44 e 7.45 exibem a faixa de passagem do satélite ENVISAT em dois instantes distintos no tempo, no modo nadir, sobre e a região de Mauna Loa, Havaí e sobre o Pólo Sul. Em geral, os erros obtidos pelo modelo estão próximos do limite estabelecido por (RAYNER et al., 2002) para o erro máximo permitido para estimar coluna integrada de  $CO_2$  (ver capítulo 1). A precisão (erro aleatório) da taxa de mistura de  $CO_2$  recuperado devido ao ruído instrumental é de  $\pm 1\%$  para a janela espectral usada para este estudo (para um albedo de 0.3 e um ângulo solar zenital  $53^\circ$ ) (BUCHWITZ M.; BURROWS, 2004).

As comparações entre os  $CO_2$  e  $CH_4$  previstas pelo modelo e aqueles observados pela NOAA ao longo do ano de 2005 e 2007 nas regiões selecionadas mostram, em geral, boa concordância. O resíduo detectado é menor que 0.9% para os dois gases. A correlação entre as séries temporais dos dados *in situ* e as médias mensais do modelo para área de estudo de Mauna Loa foi de  $\sim 0.69$  e  $\sim 0.7943$  para  $CO_2$  e  $CH_4$ , respectivamente.

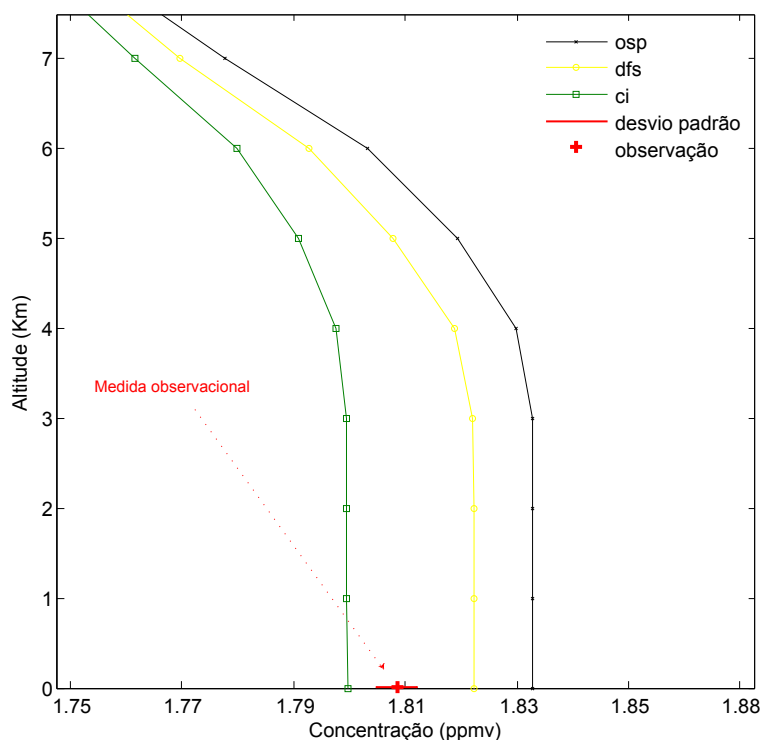


Figura 7.39 - Ampliação da figura 7.38 mostrando detalhes dos perfis recuperados pelos 3 conjuntos de canais do sensor SCIAMACHY juntamente com dados observados e o desvio padrão (erro de medida).

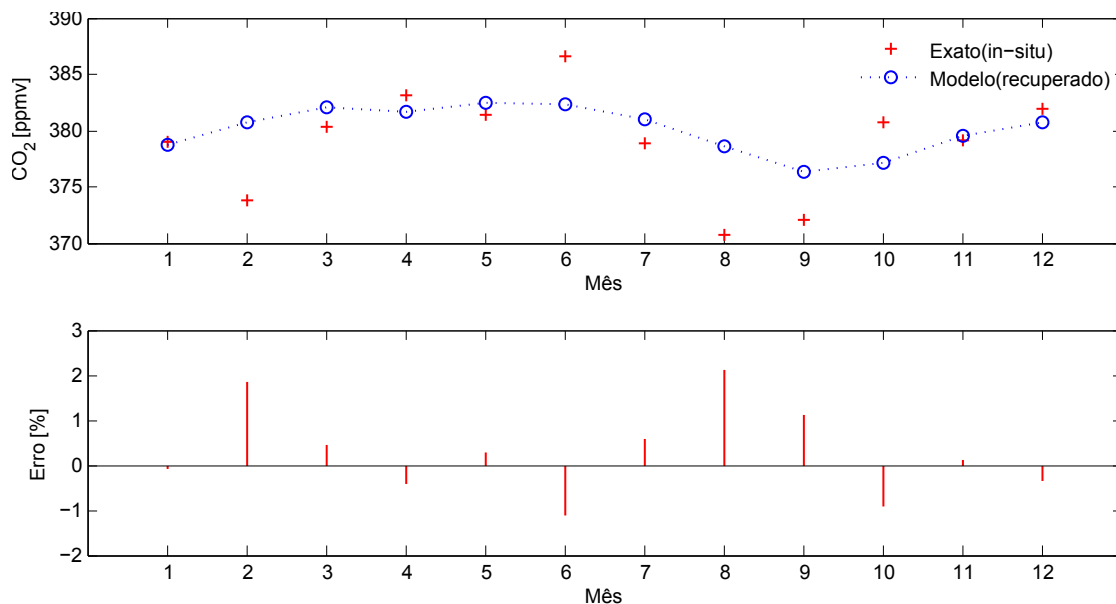


Figura 7.40 - (Superior) Recuperação de Fração Molar de Dióxido de Carbono Atmosférico ( $CO_2$ ) (em ppmv) obtido a partir de medições quasi-contínuas em Mauna Loa, no Havaí para o ano de 2005. Para a inversão utilizou-se uma RNA PMC com duas camadas ocultas, uma com 3 neurônios em cada camada e função de ativação Logarítmica. (Inferior) Erro residual (em porcentagem).

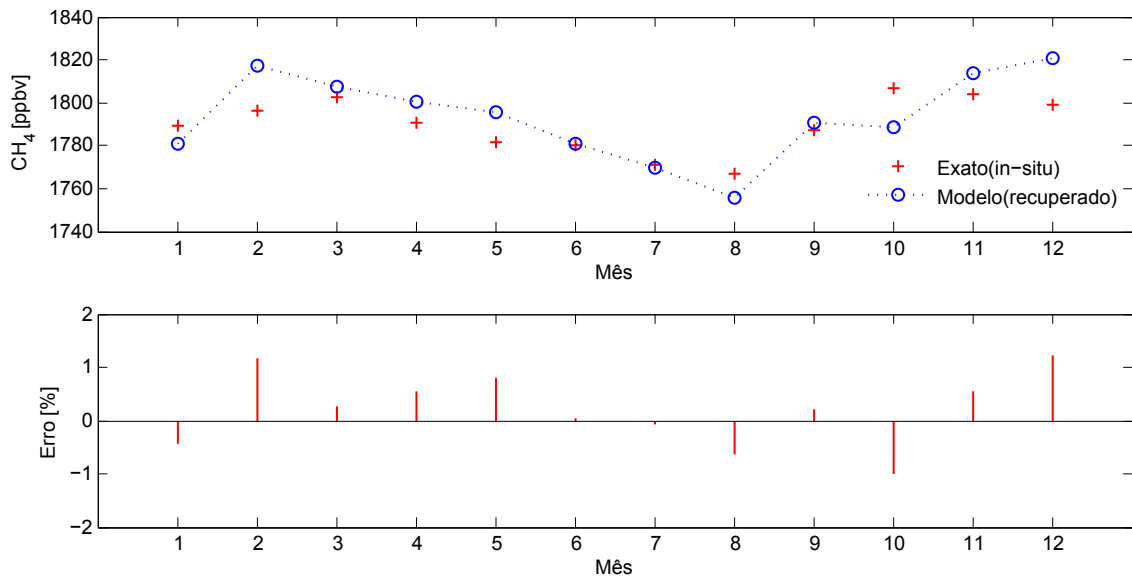


Figura 7.41 - (Superior) Recuperação de Fração Molar de Metano Atmosférico ( $CH_4$ ) (em ppmv) obtido a partir de medições quasi-contínuas em Mauna Loa, no Havaí para o ano de 2005. Para a inversão utilizou-se uma RNA PMC com função de ativação do tipo Tangente Hiperbólica duas camadas ocultas, uma com 21 neurônios e outra com 1. (Inferior) Erro residual (em porcentagem).

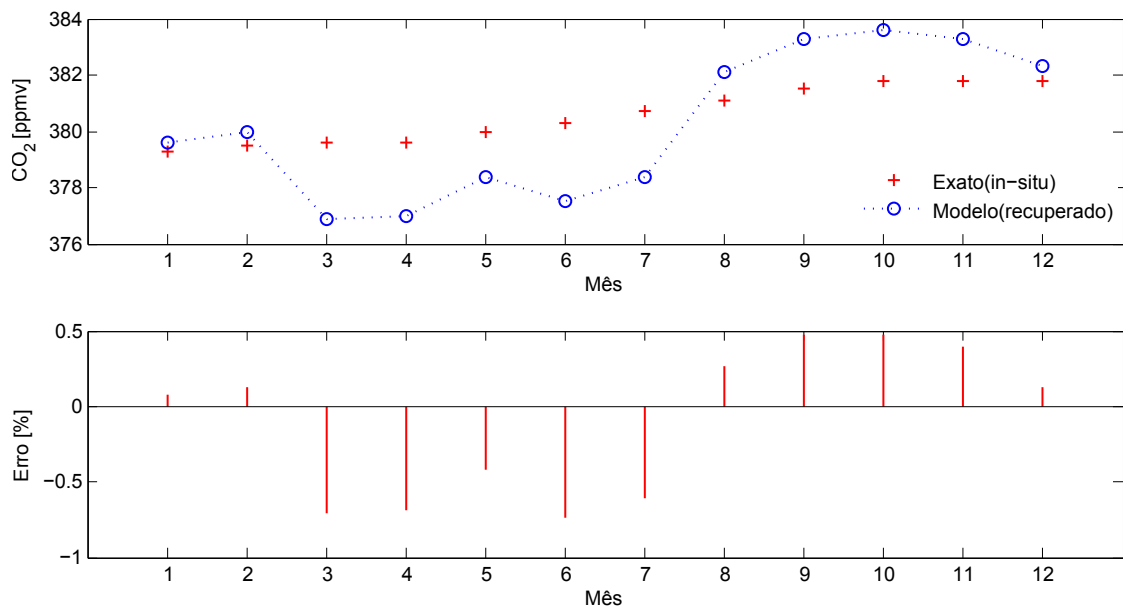


Figura 7.42 - (Superior) Recuperação de Fração Molar de Dióxido de Carbono Atmosférico ( $CO_2$ ) (em ppmv) obtido a partir de medições quasi-contínuas no Pólo Sul para o ano de 2007. A tabela 7.11 Para a inversão utilizou-se uma RNA PMC com uma única camada oculta de 30 neurônios e função de ativação sigmoidal. (Inferior) Erro residual (em porcentagem).

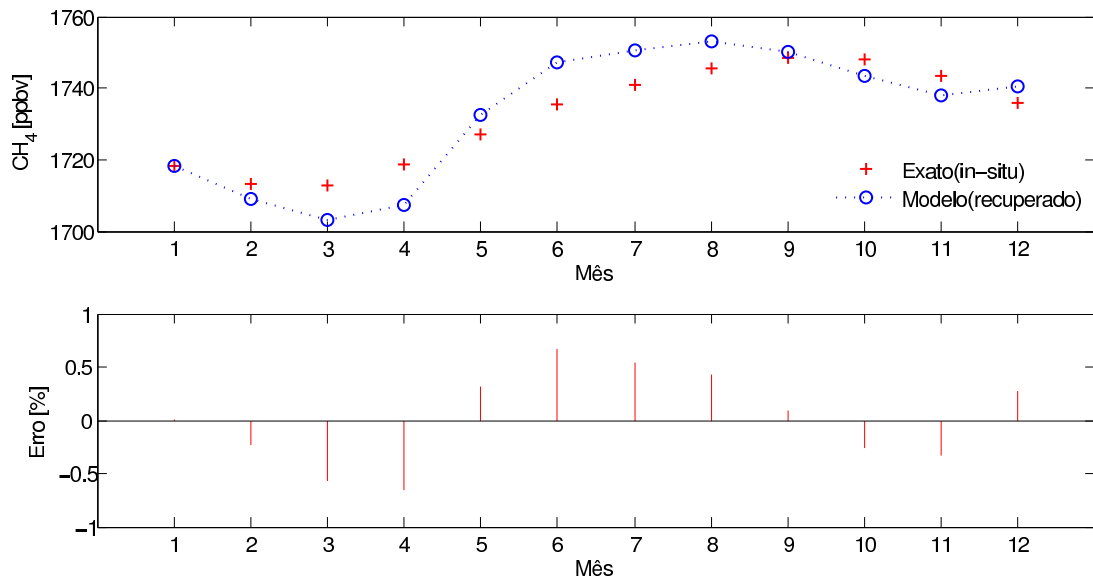


Figura 7.43 - (Superior) Recuperação de Fração Molar de Metano Atmosférico ( $CH_4$ ) (em ppmv) obtido a partir de medições quasi-contínuas no Pólo Sul para o ano de 2007. Para a inversão utilizou-se uma RNA PMC com função de ativação Sigmoidal com duas camadas, uma com 4 neurônios e outra com 6. (Inferior) Erro residual (em porcentagem).

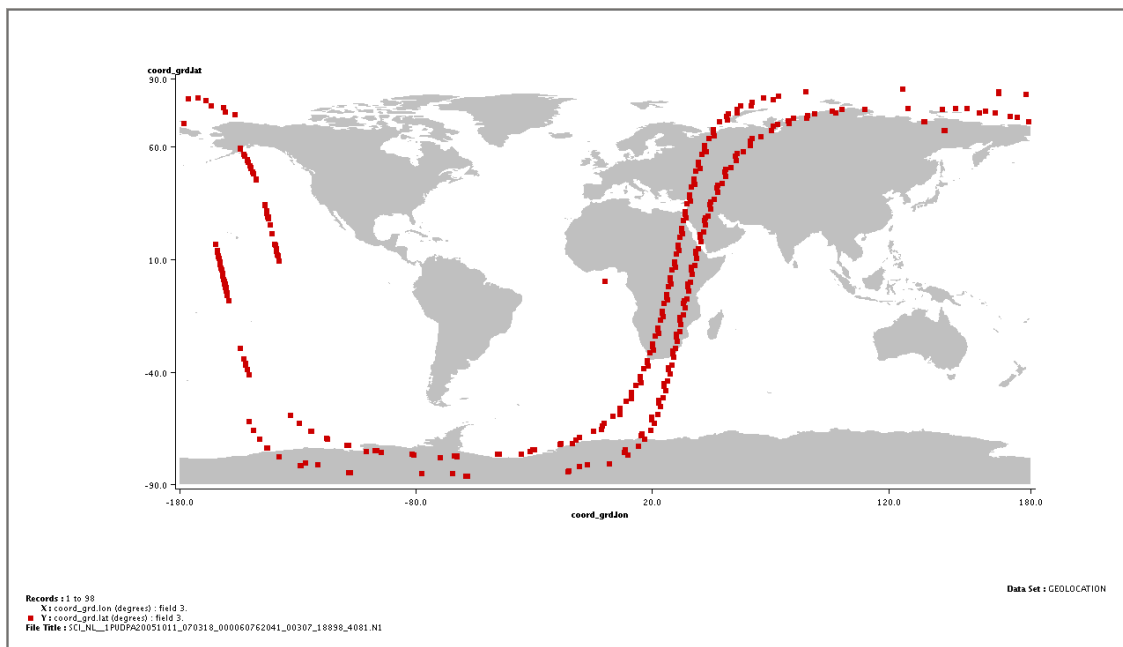


Figura 7.44 - Passagem do satélite sobre a região de Mauna Loa no dia 11 de outubro de 2005.

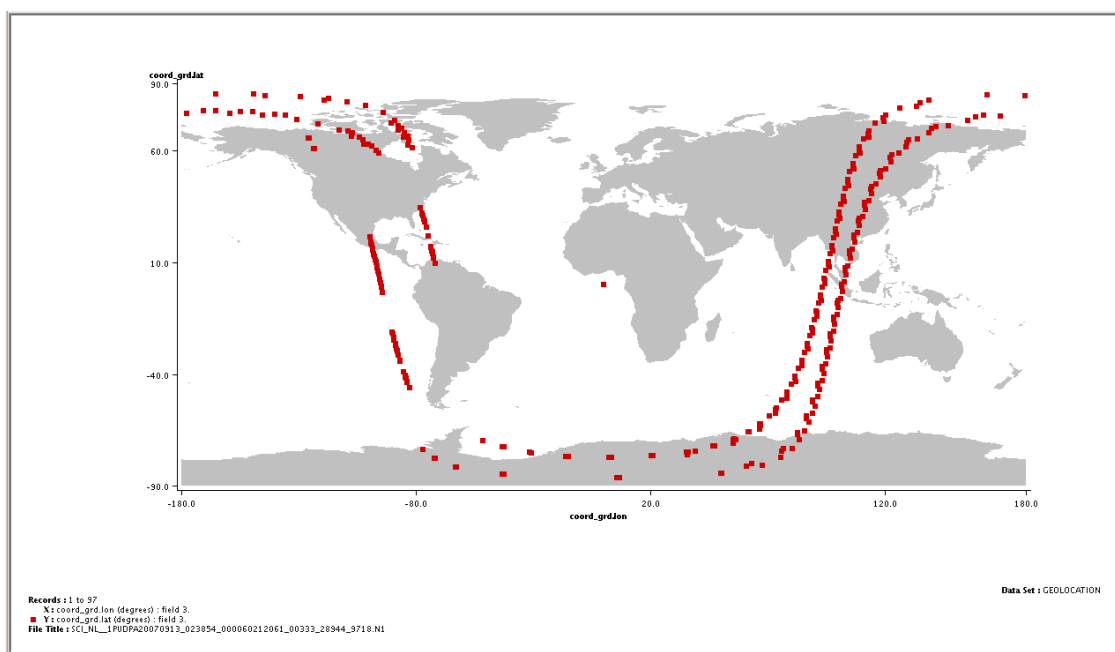


Figura 7.45 - Passagem do satélite sobre a região do Pólo Sul dia 13 de setembro de 2007.

Para o segundo cenário, o Pólo Sul, as correlações foram de  $\sim 0.61$  para  $CO_2$  e  $\sim 0.9274$  para o  $CH_4$ . Esses valores de correlação indicam uma habilidade moderada a alta do algoritmo de inversão recuperar as mudanças sazonais nas concentrações de  $CO_2$  e  $CH_4$ .

No entanto, o modelo RNA mostra uma tendência a subestimar os valores de concentração do gás nos 6 primeiros meses e superestimar os valores nos meses seguintes, como observado nos dois gráficos ao longo dos anos 2005 e 2007, o que não pode ser explicado ainda. Além disso, o modelo foi capaz de detectar tendências e uma amplitude do ciclo sazonal para as previsões feitas em ambos os cenários. À luz deste estudo, a precisão global (desvio padrão) e a acurácia (bias) das razões de mistura recuperadas (em relação à observação) para o metano no primeiro cenário foram estimadas em  $\sim 0.69\%$  e  $0.22\%$ , respectivamente. No segundo cenário (Pólo Sul) a precisão ficou em  $\sim 0.44\%$  e o bias foi de  $0.02\%$ . Em ambos os cenários nesse caso houve uma tendência de superestimação do bias. Para o dióxido de carbono essas mesmas variáveis, em idênticas condições de temperatura e pressão, foram estimadas em  $\sim 1\%$  (precisão) e  $0.3\%$  (bias) no primeiro cenário e  $\sim 0.49\%$  (precisão) e  $\sim -0.10\%$  (bias) no segundo cenário. Também devemos salientar que não aplicamos fatores de escala em nenhuma fase dos experimentos para a recuperação das taxas



de mistura dos gases, ao contrário do que observamos em outros estudos similares. No geral, as diferenças entre RMS medidos e calculados ficarem abaixo dos 1% para ambos os gases o que parece razoável se levarmos em conta os vários tipos de erros que podem contribuir para as diferenças entre o modelo e os dados observacionais. Esses erros incluem:

- a) O ruído instrumental (erros aleatórios);
- b) Os erros sistemáticos do instrumento;
- c) Os erros do modelo direto (podem incluir ambos os componentes sistemático e aleatório, ver (STROW *et al.*, 2002));
- d) Erros de amostragem e do Modelo atmosférico. Para determinar o vetor de estado completo utilizado pelo modelo direto em seus cálculos, os dados de análise foram complementados através de extrapolação dos perfis de temperatura (T) e umidade (U) retirados do TIGR, para além de 0.0026hPa.

Embora estes resultados sejam encorajadores ainda não estão com a qualidade desejada para a modelagem inversa. Espera-se que novas melhorias no algoritmo de recuperação superem esse problema no futuro. Podemos citar como exemplo de melhorias: (i) uma melhor calibração dos dados SCIAMACHY; (ii) um incremento na qualidade dos dados de entrada utilizados pelo modelo de RNA em seu treinamento, incluindo perfis reais obtidos por campanhas meteorológicas e (iii) ajustes no algoritmo de otimização para produzir soluções de topologia de rede cada vez melhores.



## 8 CONCLUSÃO

Perfis verticais de concentração de  $CO_2$  e  $CH_4$  atmosférico foram recuperados nos trópicos em (latitude  $19.54^\circ N$ ,  $155.58^\circ W$ ) e no Pólo Sul (latitude  $-89.98^\circ S$ , longitude  $-24.80^\circ W$ ) usando medições no canal 6 do sensor SCIAMACHY (1000-1700 nm) no modo nadir do infravermelho próximo. O método de inversão se baseia em um esquema de inferência por regressão não-linear utilizando redes neurais. Para isso, aplicamos uma arquitetura do tipo Perceptron Multi-Camadas (PMC) com aprendizado supervisionado e algoritmo de retro-propagação para o treinamento da rede.

O uso de redes neurais foi capaz de resolver este difícil problema inverso, mesmo quando os dados foram contaminados com ruído. A precisão global e a acurácia dos perfis verticais são estimadas em aproximadamente 1.0% e no pior caso inferiores a 3.0% respectivamente. Uma estimativa aproximada do erro médio relativo é de cerca de 0.5% para os dados sintéticos e 1% para o experimento com dados de satélite. Por fim, as recuperações por satélite foram comparados com medições *in situ* a partir de amostras de ar coletadas em flascos de vidro realizados no pela divisão de monitoramento global da NOAA ESRL. A partir deste trabalho, duas vantagens do uso de redes neurais na recuperação de gases-traço ficaram evidentes. Em primeiro lugar, após a fase de treinamento, o algoritmo de reconstrução é muito mais rápido (na ordem de milissegundos) do que os métodos de inversão clássicos. Em segundo lugar, ele pode ser facilmente implementado em um ambiente paralelo. Um trabalho futuro poderia ser comparar o desempenho do modelo com outras abordagens operacionais e utilizar esta fonte de resultados para produzir mapas de distribuição de fontes e sumidouros de  $CO_2$  e  $CH_4$ . Mais estudos, uma calibração melhorada e refinamento do algoritmo são necessárias para permitir recuperações mais precisas de colunas de  $CO_2$ .

Realizamos estudos de sensibilidade e criamos uma metodologia para selecionar os sub-canais mais adequados para a recuperação dos perfis. Mostramos ser possível selecionar conjuntos reduzidos de canais para recuperação de perfis verticais de GEE mediante análise prévia do conteúdo de informação no sentido de Shannon das linhas espectrais do sensor (método CI). Outras duas abordagens para seleção de canais foram testadas com sucesso: a seleção de canais através da extração dos graus de liberdade do sinal (método GLS) e a escolha de canais através de análise da sensibilidade das linhas com relação à variação do gás (método OSP). As três

abordagem obtiveram bons resultados para a tarefa de recuperação dos perfis, no entanto o método baseado no conteúdo de informação resultou em perfis mais acurados e suaves. Sugerimos um novo método que tem como fundamento a variância ao longo das colunas da matriz jacobiana da radiância em relação à concentração do gás. Este método, apesar de bem simples e contar com algoritmo bem definido, não foi implementado nesse trabalho de pesquisa e fica, portanto, como uma sugestão para trabalhos futuros.

A partir dos estudos de sensibilidade e da análise dos Jacobianos dos canais constatamos que o SCIAMACHY apresenta baixa sensibilidade e, portanto, apresenta pouca informação nas camadas da termosfera e estratosfera. Entretanto, o modelo inverso baseado em RNA foi capaz de extrapolar o perfil de concentração para tais níveis atmosféricos sem grandes dificuldades. Ainda não foi possível obter uma explicação do ponto de vista físico ou meteorológico para tal comportamento da rede. Uma provável explicação para tal comportamento seria o fato de que a rede consegue realizar um ajuste do que ocorre na estratosfera e termosfera com base nos dados da troposfera, que contribuem com informação necessária e suficiente para as radiâncias.

Adicionalmente propomos uma nova metodologia para a escolha de arquitetura de RNA feedforward com complexidade mínima e desempenho otimizado. Nós adaptamos e comparamos quatro meta-heurísticas de busca global para treinar e descobrir arquiteturas de rede melhoradas. A seleção da arquitetura de rede considerou como critério de poda não apenas o erro de treinamento da rede, mas também o erro de generalização. Além disso, o erro de generalização não é simplesmente medido mas diretamente calculado durante o processo evolutivo. Depois disso, a arquitetura é treinada usando um subconjunto que não pertence ao conjunto de teste. Um novo critério para medir a complexidade do modelo de inversão com base no número de pesos presentes no arranjo da rede e no número de épocas necessárias para o treinamento é criado e incluído na metodologia. Aplicou-se o esquema proposto ao problema da recuperação de perfis atmosféricos de concentração de gases de efeito estufa o qual permitiu a obtenção de uma arquitetura de rede otimizada. O primeiro estudo de caso consistiu em estimar corretamente o perfil vertical de concentração utilizando um cenário clássico de inversão onde utilizamos apenas uma rede neural para recuperar o perfil todo, incluindo níveis em todas as camadas atmosféricas e uma RNA por massa de ar e conjunto de canais selecionados.

Em um segundo cenário utilizamos para o problema uma abordagem de “divisão e conquista” através da criação de um cenário híbrido, no qual treinamos uma rede neural para cada camada atmosférica e faixa de latitude terrestre. Para esse cenário geramos diversas combinações possíveis de redes. Os resultados mostraram que a abordagem computacional proposta apresenta melhor performance em relação a um especialista humano, oferecendo ao mesmo tempo muitas vantagens se comparada a abordagens similares encontradas na literatura. Além disso, definimos a complexidade computacional de uma arquitetura de RNA do tipo feedforward como sendo uma função do número total de pesos e bias presentes em suas estruturas e do tempo necessário para a aprendizagem da rede. A partir disto derivamos um termo de penalidade usado para avaliar a função objetivo de modo a evitar arquiteturas de rede muito complexas.

Com base nos resultados obtidos e naqueles encontrados na literatura, devemos ressaltar que é plenamente justificável evoluir a arquitetura de RNA se o nosso interesse é obter uma solução dedicada e adaptada ao contexto do problema. O tempo necessário para evoluir uma arquitetura e encontrar um ajuste apropriado para o problema é plenamente justificado, se comparado com o tempo gasto em um procedimento empírico de “tentativa-e-erro”. Contudo, o custo computacional traduzido em tempo de processamento para uma máquina monoprocessada executar dezenas de milhares de avaliações da função objetivo, ainda é alto e pode se tornar impraticável, se adotamos uma maior granularidade do espaço de busca. Uma alternativa promissora seria paralelizar o código do algoritmo de otimização e executá-lo em um ambiente multiprocessado, usando o fato que RNAs são um algoritmo intrinsecamente paralelo. Outra alternativa seria a utilização de tecnologia de processamento paralelo distribuído pela computação em grade onde seria possível atingir uma alta taxa de processamento dividindo as tarefas de processamento entre várias máquinas. Estes processos seriam executados quando as máquinas não estivessem sendo utilizadas pelo usuário evitando o desperdício de processamento da máquina. Em ambos os casos, com pequenos ajustes de código e melhorias no hardware, poderíamos reduzir o tempo de busca em mais de uma ordem de grandeza.

É importante ressaltar que o foco desse estudo em Otimização Combinatória não diz respeito à utilização e posterior comparação das 4 metaheurísticas aqui empregadas, mas sim a destacar e propor alternativas à estratégia de busca de “às cegas” de uma arquitetura de rede, realizada mediante o processo repetitivo de tentativa-e-

erro. Obviamente, os resultados obtidos e exibidos nas tabelas acima poderiam também ser obtidos com a utilização de outras metaheurísticas. O que podemos inferir desses e de outros resultados encontrados na literatura ((YAO X.; LIU, 1997);(MILLER G. F.; TODD; HEDGE, 1991);(ARIFOVIC J.; GENCAJ, 2001);(BERNARDOS P. G.; VOS-  
NIAKOS, 2004) e (CHEN Z.; XIAO; CHENG, 1997)) é que o simples emprego de algum método sistemático e automatizado de busca como o proposto nesse trabalho é preferível à sua ausência. Por fim, podemos dizer que, embora a abordagem tenha sido aplicada para a resolução de um problema específico, o da estimação da concentração de gases traço atmosférico, ela é bastante genérica e mostra-se robusta o suficiente para ser adaptada e aplicada a qualquer problema que possa ser resolvido através de RNAs. Uma vantagem do algoritmo proposto, baseado em RNAs, em relação à outras abordagens é a possibilidade da implementação em hardware, por exemplo em um chip utilizando FPGA (VELHO H. F.; SHIGUEMORI; SILVA, 2007). Trata-se de uma nova era de sensores "inteligentes", com inversão quase em tempo real.

A metodologia implantada neste trabalho foi direcionada para a inferência de perfis de concentração a partir de dados do sensor SCIAMACHY porém, ela é extensível a outros instrumentos de sondagem atmosférica, mesmo que estes atuem fora da faixa do Infravermelho Próximo. Neste contexto, sugerimos para trabalhos futuros a utilização de sensores com maior resolução espacial e espectral tais como o GOSAT ou o OCO-2 que, certamente, irão produzir resultados com maior precisão e acurácia.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ACKLEY D.; HINTON, G. E.; SEJNOWSKI, T. J. A learning algorithm for boltzmann machines. **Cognitive Science**, v. 9, p. 147–169, 1985. 48
- AIRES, F.; PRIGENT, C.; ROSSOW, W.; ROTHSTEIN, M. A new neural network approach including first-guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature and emissivities over land from satellite microwave observations. **J. Geophys. Res.**, v. 106, p. 14887–14907, 2001. D14, doi:10.1029/2001JD900085. 9, 33, 69
- ARIFOVIC J.; GENCAY, R. Using genetic algorithms to select architecture of a feedforward artificial neural network. **Physica A**, v. 289, p. 574–594, 2001. 86, 156
- AUMANN, H. H.; CHAHINE, M. T.; GAUTIER, C.; GOLDBERG, M. D.; KALNAY, E.; MCMILLIN, L. M.; REVERCOMB, H.; ROSENKRANZ, P. W.; SMITH, W. L.; STAELIN, D. H.; STROW, L. L.; SUSSKIND, J. Airs/amsu/hsb on the aqua mission: design, science objectives, data products, and processing systems. **IEEE Transactions on Geoscience and Remote Sensing**, v. 41, p. 253–264, 2003. 17, 20
- AUMANN H. H.; GREGORICH, D.; GAISER, S. Airs hyper-spectral measurements for climate research: Carbon dioxide and nitrous oxide effects. **Geophys. Res. Lett.**, v. 32, 2005. L05806, doi:10.1029/2004GL021784. 17
- BAK, P. **How nature works**. New York: Copernicus, Springer-Verlag, 1996. 88
- BAK P.; SNEPPEN, K. Punctuated equilibrium and criticality in a simple model of evolution. **Physical Review Letters**, n. 24, v. 71, p. 4083–4086, 1993. 88
- BEALE R.; JACKSON, T. **Neural computing: an introduction**. New York, NY: Adam Higler Bristol, 1990. 58
- BELLMAN, R. E. **Adaptive control processes**. Princeton, NJ: Princeton University Press, 1961. 67
- BENNETT, K. P. Decision tree construction via linear programming. In: MIDWEST ARTIFICIAL INTELLIGENCE AND COGNITIVE SCIENCE SOCIETY CONFERENCE, 4., 1992, Utica, Illinois. **Proceedings...** Utica: M. Evans, ed., 1992. p. 82–90. 61

BERGAMASCHI, P.; KROL, M.; DENTENER, F.; VERMEULEN, A.; MEINHARDT, F.; GRAUL, R.; RAMONET, M.; PETERS, W.; DLUGOKENCKY, E. J. Inverse modelling of national and european CH<sub>4</sub> emissions using the 5 atmospheric zoom model tm5. **Atmos. Chem. Phys.**, v. 5, p. 2431–2460, 2005. 14

BERNARDOS P. G.; VOSNIAKOS, G. C. Optimizing feedforward artificial neural network architecture. **Engineering Applications of Artificial Intelligence**, n.3, v. 20, p. 365–382, 2004. 83, 86, 156

BISHOP, C. **Neural networks for pattern recognition**. 1. ed. Oxford, UK: Oxford University Press, 1995. 56, 57, 67, 98, 99, 122

BOSCH, H.; TOON, G. C.; SEN, B.; WASHENFELDER, R. A.; WENNBERG, P. O.; BUCHWITZ, M.; BEEK, R. de; BURROWS, J. P.; CRISP, D.; CHRISTI, M.; CONNOR, B. J.; NATRAJ, V.; YUNG, Y. L. Space-based nearinfrared CO<sub>2</sub> measurements: Testing the orbiting carbon observatory retrieval algorithm and validation concept using sciamachy observations over park falls, wisconsin. **J. Geophys. Res.**, v. 111, p. 127–150, 2006. D23302, doi:10.1029/2006JD007080. 13

BOVENSMANN, H.; BURROWS, J. P.; BUCHWITZ, M.; FRERICK, J.; NOEL, S.; ROZANOV, V. V.; CHANCE, K. V.; GOEDE, A. Sciamachy: Mission objectives and measurement modes. **J. Atmos. Sci.**, v. 56, p. 127–150, 1999. 15, 17, 18

BRUHL C.; CRUTZEN, P. J. Mpic two-dimensional model. **In The atmospheric effects of stratospheric aircraft, NASA Ref. Publ.**, v. 1292, 1993. 32

BUCHWITZ, M.; BEEK, R. de; NOEL, S.; BURROWS, J. P.; BOVENSMANN, H.; BREMER, H.; BERGAMASCHI, P.; KORNER, S.; HEIMANN, M. Carbon monoxide, methane and carbon dioxide columns retrieved from sciamachy by wfm-doas: year 2003 initial data set. **Atmos. Chem. Phys. Discuss.**, **SRef-ID:1680-7375/acpd/2005-5-1943**, v. 5, p. 1943–1971, 2005. 68

BUCHWITZ, M.; BEEK, R. de; NOEL, S.; BURROWS, J. P.; BOVENSMANN, H.; SCHNEISING, O.; KHLYSTOVA, I.; BRUNS, M.; BREMER, H.; BERGAMASCHI, P.; KORNER, S.; HEIMANN, M. Atmospheric carbon gases retrieved from sciamachy by wfm-doas: version 0.5 CO and CH<sub>4</sub> and impact of calibration improvements on CO<sub>2</sub> retrieval. **Atmos. Chem. Phys.**, v. 6, p. 2727–2751, 2006. 6, 13, 37



BUCHWITZ, M.; KHLYSTOVA, I.; BOVENSMANN, H.; BURROWS, J. P. Three years of global carbon monoxide from sciamachy: Comparison with mopitt and first results related to the detection of enhanced CO over cities. **Atmos. Chem. Phys.**, v. 7, p. 2399–2411, 2007. 6

BUCHWITZ, M.; NOEL, S.; BRAMSTEDT, K.; ROZANOV, V. V.; EISINGER, M.; BOVENSMANN, H.; TSVETKOVA, S.; BURROWS, J. P. Retrieval of trace gas vertical columns from sciamachy/envisat near-infrared nadir spectra: first preliminary results. **Advances in Space Research**, v. 34, p. 809–814, 2004. 19, 20, 37

BUCHWITZ, M.; SCHNEISING, O.; BURROWS, J. P.; BOVENSMANN, H.; NOTHOLT, J. First direct observation of the atmospheric CO<sub>2</sub> year-to-year increase from space. **Atmos. Chem. Phys.**, v. 7, p. 4249–4256, 2007. 6, 13

BUCHWITZ M.; BURROWS, J. P. Retrieval of CH<sub>4</sub>, CO, and CO<sub>2</sub> total column amounts from sciamachy near-infrared nadir spectra: Retrieval algorithm and first results. In: SPIE 5235, REMOTE SENSING OF CLOUDS AND THE ATMOSPHERE, 8., 2004, Barcelona, Spain. **Proceedings...** Barcelona: K. P. Schäfer and A. Comeron and M. R. Carleer and R. H. Picard (Editors), 2004. p. 375–388. 146

BUCHWITZ M.; ROZANOV, V. V.; BURROWS, J. P. A near infrared optimized doas method for the fast global retrieval of atmospheric CH<sub>4</sub>, CO, CO<sub>2</sub>, H<sub>2</sub>O, and N<sub>2</sub>O total column amounts from sciamachy/envisat nadir radiances. **Journal Of Geophysical Research**, v. 105, p. 15231–15246, 2000. D12, 2000, doi:10.1029/2000JD900191. 32, 37, 38

BURROWS, J. P. Sciamachy - scanning imaging absorption spectrometer for atmospheric cartography. **Actu Astr**, v. 35, p. 445–451, 1995. 18

BUTLER, T. M.; RAYNER, P. J.; SIMMONDS, I.; LAWRENCE, M. G. Simultaneous mass balance inverse modeling of methane and carbon monoxide. **J. Geophys. Res.**, v. 110, 2005. D21310, doi:10.1029/2005JD006071. 14

CARVALHO A. C. P. L. F.; BRAGA, A. P.; LUDERMIR, T. B. Fundamentos de redes neurais artificiais. In: ESCOLA BRASILEIRA DE COMPUTAÇÃO, 11., 1997, Rio de Janeiro: Sba. **Proceedings...** Rio de Janeiro, 1997. 62

CARVALHO, A. R.; VELHO, H. C.; STEPHANY, S.; SOUTO, R. P.; BECCENERI, J. C.; SANDRI, S. Fuzzy ant colony optimization for estimating chlorophyll concentration profile in offshore sea water. **Inverse Problems in Science & Engineering**, v. 16, p. 705–715, 2008. 86

CHAHINE, M. T.; AUMANN, H.; GOLDBERG, M.; MCMILLIN, L.; ROSENKRANZ, P.; STAELIN, D.; STROW, L.; SUSSKIND, J. **AIRS level2 algorithm theoretical basic document (ATBD), version 2.2**. JPL, 2001. 198p. JPL D-17006 Relatório técnico. 21, 39

CHANDRASEKHAR, S. **Radiative transfer**. [S.l.]: Dover, New York, 393 pp. (2nd Edit., 1960), 1950. 31

CHEDIN, A.; SCOTT, N.; WAHICHE, C.; MOULINIER, P. The improved initialisation inversion method: A high resolution physical method for temperature retrievals from satellite of the tiros-n series. **J. Clim. Appl. Meteorol.**, v. 24, p. 128–143, 1985. 78

CHEDIN, A.; SERRAR, S.; HOLLINGSWORTH, A.; ARMANTE, R.; SCOTT, N. A. Annual and seasonal variations of atmospheric CO<sub>2</sub>, N<sub>2</sub>O, and CO concentrations retrieved from noaa/tovs satellite observations. **Geophys. Res. Lett.**, v. 29, p. 110–1–110–4, 2002. Doi:10.1029/2001GL014082. 17, 111, 112

CHEDIN, A.; SERRAR, S.; SCOTT, N.; CREVOISIER, C.; ARMANTE, R. First global measurement of midtropospheric CO<sub>2</sub> from NOAA polar satellites: tropical zone. **J. Geophys. Res.**, v. 108 (D18), p. 13 pp., 2003. 4581, doi:10.1029/2003JD003439. 9, 17

CHEN Y. H.; PRINN, R. G. Estimation of atmospheric methane emissions between 1996 and 2001 using a threedimensional global chemical transport model. **J. Geophys. Res.**, v. 111, 2006. Doi:10.1029/2005JD006058. 14

CHEN Z.; XIAO, J.; CHENG, J. A program for automatic structure search. In: INTERNATIONAL CONFERENCE ON NEURAL NETWORK, 1997, Houston, TX. **Proceedings...** Houston, TX: IEEE, 1997. p. 308–311. 86, 156

CHEVALLIER, F.; CHERUY, F.; SCOTT, N. A.; CHEDIN, A. A neural network approach for a fast and accurate computation of a longwave radiative budget. **J. Appl. Meteorol.**, v. 37, p. 1385–1397, 1998. 8, 78

CHEVALLIER F.; BREON, F. M.; RAYNER, P. J. Contribution of the orbiting carbon observatory to the estimation of CO<sub>2</sub> sources and sinks: Theoretical study in a variational data assimilation framework. **J. Geophys. Res.**, v. 112, 2007. D09307, doi:10.1029/2006JD007375. 12

CHU P. S.; YU, P.; HASTENRATH, S. Detecting climate change concurrent with deforestation in the amazon basin: Which way has it gone? **Bulletin Of American Meteorological Society**, v. 75, p. 579–583, 1994. 5

CLERBAUX, C.; COHEUR, P. F.; HURTMANS, D.; BARRET, B.; CARLEER, M.; COLIN, R.; SEMENIUK, K.; MCCONNELL, J. C.; BOONE, C.; BERNATH, P. Carbon monoxide distribution from the ace-fs solar occultation measurements. **J. Geophys. Res.**, v. 32, p. 4 PP., 2005. 9

CLERBAUX, C.; HADJI-LAZARO, J.; PAYAN, S.; CAMY-PEYRET, C.; MEGIE, G. Retrieval of co columns from img/adeos spectra. **IEEE Trans. Geosci. Remote Sens.**, v. 37, p. 1657–1662, 1999. 9

CLERBAUX, C.; HADJI-LAZARO, J.; TURQUETY, S.; MEGIE, G.; COHEUR, P. F. Trace gas measurements from infrared satellite for chemistry and climate applications. **Atmos. Chem. Phys.**, v. 3, p. 2027–2058, 2003. 14

COLDEWEY-EGBERS, M.; WEBER, M.; LAMSAL, L. N.; BEEK, R. de; BUCHWITZ, M.; BURROWS, J. P. Total ozone retrieval from gome uv spectral data using the weighting function doas approach. **Atmos. Chem. Phys. Discuss.**, v. 4, p. 4915–4944, 2004. 37

CONWAY, T. J.; TANS, P. P.; WATERMAN, L. S.; THONING, K. W.; KITZIS, D.; MASARIE, K. A.; ZHANG, N. Evidence for interannual variability of the carbon cycle from the national oceanic and atmospheric administration climate monitoring and diagnostics laboratory global air sampling network. **J. Geophys. Res.**, **22**, v. 99, p. 831–855, 1994. 108

CONWAY T. J.; LANG, P. M. M. K. A. **Atmospheric carbon dioxide dry air mole fractions from the NOAA ESRL carbon cycle cooperative global air sampling network, 1968-2009**. 2010. Disponível em: <<ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/>>. 144

COURTIER P.; THEPAUT, J. N.; HOLLINGSWORTH, A. A strategy for operational implementation of 4d-var, using an incremental approach. **Q. J. R. Meteorol. Soc.**, v. 120, p. 1367–1388, 1994. 42, 43

CREVOISIER, C. **Etude de la distribution du CO<sub>2</sub> atmosphérique à partir des observations infrarouges à haute résolution spectrale de l'instrument Aqua/AIRS**. 238 p. PhD Thesis, Université Denis Diderot-Paris VII — Université Denis Diderot-Paris, 2004. 6

CREVOISIER, C.; CHEDIN A. AN MATSUEDA, H.; MACHIDA, T.; ARMANTE, R.; SCOTT, N. A. First year of upper tropospheric integrated content of CO<sub>2</sub> from iasi hyperspectral infrared observations. **Atmos. Chem. Phys.**, v. 9, p. 4797–4810, 2009. 17

CREVOISIER, C.; NOBILEAU, D.; FIORE, A. M.; ARMANTE, R.; CHEDIN, A.; SCOTT, N. A new insight on tropospheric methane in the tropics - first year from iasi hyperspectral infrared observations. **Atmos. Chem. Phys. Discuss.**, v. 9, p. 6855–6887, 2009. 142

CREVOISIER C.; CHEDIN, A.; SCOTT, N. A. Airs channel selection for CO<sub>2</sub> and other trace-gas retrievals. **Quarterly Journal of the Royal Meteorological Society**, v. 129, p. 2719–2740, 2003. 7, 9, 13, 21, 67, 69

CRISP, D.; ATLAS, R.; BREON, F.; BROWN, L.; BURROWS, J.; CIAIS, P.; CONNOR, B.; DONEY, S.; FUNG, I.; JACOB, D. The orbiting carbon observatory (oco) mission. **Adv. Space Res.**, v. 34, p. 700–709, 2004. 17

CUCO, A. P. C.; NETO, A. J. S.; VELHO, H. F. C.; SOUZA, F. Solution of an inverse adsorption problem with an epidemic genetic algorithm and the generalized extremal optimization algorithms. **Inverse Problems in Science & Engineering**, v. 17, p. 289–302, 2009. 86

DARWIN C. R.; WALLACE, A. R. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. **Journal of the Proceedings of the Linnean Society of London**, v. 3, p. 46–50, 1858. 85

DENTENER, F.; PETERS, W.; KROL, M.; WEELE, M. V.; BERGAMACHI, P.; LELIEVELD, J. Interannual variability and trend of CH<sub>4</sub> lifetime as a measure for

OH changes in the 1979-1993 time period. **J. Geophys. Res.**, v. 108, p. 6855–6887, 2003. 14

DLUGOKENCKY E. J.; LANG, P. M. M. K. A. **Atmospheric methane dry air mole fractions from the NOAA ESRL carbon cycle cooperative global air sampling network, 1983-2009**. 2010. Disponível em: [<ftp://ftp.cmdl.noaa.gov/ccg/ch4/flask/event/>](ftp://ftp.cmdl.noaa.gov/ccg/ch4/flask/event/). 144

ELMAN, J. L. Finding structure in time. **Cognitive Science**, v. 14, p. 179–211, 1990. 62

ENGELEN, R. J.; ANDERSSON, E.; CHEVALLIER, F.; HOLLINGSWORTH, A.; MATRICARDI, M.; MCNALLY, A. P.; THEPAUT, J. N.; WATTS, P. D. Estimating atmospheric CO<sub>2</sub> from advanced infrared satellite radiances within an operational 4d-var data assimilation system: Methodology and first results. **Journal of Geophysical Research**, v. 109, 2004. 17, 43

ENGELEN, R. J.; DENNING, A. S.; GURNEY, K. R.; STEPHENS, G. L. Global observations of the carbon budget, 1. expected satellite capabilities in the EOS and NPOESS eras. **Geophys. Res.**, 20, v. 106, p. 20055–20068, 2001. 12, 21

ENGELEN R. J.; STEPHENS, G. L. Information content of infrared satellite sounding measurements with respect to CO<sub>2</sub>. **Journal of Applied Meteorology**, v. 43, p. 373–378, 2004. 20

ERRERA, Q.; DAERDEN, F.; CHABRILLAT, S.; LAMBERT, J. C.; LAHOZ, W.; VISCARDY, S.; BONJEAN, S.; FONTEYN, D. 4d-var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses. **Atmos. Chem. Phys. Discuss.**, v. 8, p. 8009–8057, 2008. 43

ESCOBAR-MUNOZ, J.; CHEDIN, A.; CHERUY, F.; SCOTT, N. A. Réseaux de neurones multi-couches pour la restitution de variables thermodynamiques atmosphériques à l'aide de sondeurs verticaux satellitaires. **Comptes Rendus de l'Académie des Sciences de Paris, in French.**, v. 317, p. 911–918, 1993. 9

FARMER J. D.; TOFFOLI, T.; WOLFRAM, S. Cellular automata. In: INTERDISCIPLINARY WORKSHOP, 1983, Los Alamos, New Mexico. **Proceedings...** Los Alamos, New Mexico, 1983. 85

FAUSETT, L. **Fundamentals of neural networks**. New York, NY: Englewood Cliffs: Prentice-Hall, 1994. 48

FISCH, G.; LEAN, J.; WRIGHT, I. R.; NOBRE, C. A. Simulações climáticas do efeito do desmatamento na região amazônica: estudo de um caso em Rondônia.

**Revista Brasileira De Meteorologia**, n.1, v. 12, p. 33–48, 1997. 5

FLETCHER, S. E. M.; TANS, P. P.; BRUHWILER, L. M.; MILLER, J. B.; HEIMANN, M. CH<sub>4</sub> sources estimated from atmospheric observations of CH<sub>4</sub> and its 13c/12c isotopic ratios: 2. inverse modeling of CH<sub>4</sub> fluxes from geographical regions. **Global Biogeochemical Cycles**, GB4005, v. 18, p. 15 pp, 2004. 14

FOUCHER, P. Y.; CHEDIN, A.; ARMANTE, R.; BOONE, C.; CREVOISIER, C.; BERNATH, P. First carbon dioxide atmospheric vertical profiles retrieved from space observation using ACE-FTS solar occultation instrument. **Atmos. Chem. Phys. Discuss.**, v. 10, p. 26473–26512, 2010. 6

FRANKENBERG, C.; BERGAMASCHI, P.; BUTZ, A.; HOUWELING, S.; MEIRINK, J. F.; NOTHOLT, J.; PETERSEN, A. K.; SCHRIJVER, H.; WARNEKE, T.; ABEN, I. Tropical methane emissions: A revised view from sciamachy onboard Envisat. **Geophys. Res. Lett.**, v. 35, 2008. 6, 43

FRANKENBERG, C.; MEIRINK, J. F.; WEELE, M. V.; PLATT, U.; WAGNER, T. Assessing methane emissions from global space-borne observations. **Science**, v. 308, p. 1010–1014, 2005. 15

FREEMAN J.; SKAPURA, D. **Neural networks**. New York, NY: Reading MA: Addison-Wesley, 1991. 58

GASH, J. H. C.; NOBRE, C. A.; ROBERTS, J. M.; VICTORIA, R. L. **Amazonian deforestation and climate**. New York, NY: John Wiley and Sons Ltd, 1996. 5

GLOBALVIEW-CO<sub>2</sub>. **Cooperative atmospheric data integration project-carbon dioxide**. Boulder, Colorado, 2001. CD-ROM. Disponível em: <<ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/>>. 11

GOEDE, A. P. H.; AARTS, H. J. M.; BAREN, C. van; BURROWS, J. P.; CHANGE, K. V.; HOEKSTRA, R.; HOLZLE, E.; PITZ, W.; SCHNEIDER, W.; SMORENBURG, C. Sciamachy instrument design. **Adv. Space Res.**, v. 11, p. 243–246, 2001. 15

GOLDBERG D.E.; HOLLAND, J. Genetic algorithms and machine learning: Introduction to the special issue on genetic algorithms. **Machine Learning**, v. 3, 1988. 85

HANSEN P.; MLADENOVIC, N. A variable neighborhood algorithm: a new metaheuristic for combinatorial optimization. In: OPTIMIZATION DAYS N. 112, 1995, Montreal, Canada. **Proceedings...** Montreal, Canada, 1995. 87

\_\_\_\_\_. Variable neighborhood search. **International Series in Operations Research & Management Science**, v. 57, p. 145–184, 2003. 87

HARP S.; SAMAD, A.; GUHA, A. Toward the genetic synthesis of neural networks. In: **Proceedings...** San Mateo, CA: [s.n.], 1989. p. 762–767. 86

HARTER, F. P. **Redes neurais recorrentes aplicadas a assimilação de dados em dinâmica não-linear**. Tese (Doutorado em Computação Aplicada)- Instituto Nacional de Pesquisas Espaciais — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2004. 63

HARTER F. P.; CAMPOS VELHO, H. F. Recurrent and feedforward neural networks trained with cross validation scheme applied to the data assimilation in chaotic dynamics. **Revista Brasileira de Meteorologia**, v. 20, p. 411/3–420, 2005. 63

HARTMANN, D. L. **Global physical climatology**. New York, NY: Academic Press, San Diego, 1994. 4, 5

HAYKIN, S. **Neural networks: a comprehensive foundation**. New York, NY: Prentice Hall, 2001. 48, 49, 52, 55, 56, 58, 60, 66, 82, 99, 103, 122

HEBB, D. O. **The organization of behavior**. New York, NY: John Wiley & Sons, New York, 1949. 60

HECHT-NIELSEN, R. **Neurocomputing**. San Diego: Addison-Wesley, 1989. 58

HEIN R.; CRUTZEN, P. J.; HEIMANN, M. An inverse modeling approach to investigate the global atmospheric methane cycle. **Glob. Biogeochem. Cycles**, v. 11, p. 43–76, 1997. 14

HOLLAND, J. H. **Adaptation in natural and artificial systems**. Ann Arbor:: University of Michigan Press, 1975. 94

HOPFIELD, J. Neural networks and physical systems with emergent collective computational abilities. In: NATIONAL ACADEMY OF SCIENCES OF THE USA, 9., 1982, Washington, DC. **Proceedings...** Washington, DC, 1982. 48

HORNIK, K. Some new results on neural network approximation. **Neural Networks**, v. 6, p. 1069–1072, 1993. 57, 98

HOUWELING, S.; BREON, F. M.; ABEN, I.; RODENBECK, C.; GLOOR, M.; HEIMANN, M.; CIAIS, P. CO<sub>2</sub> sources and sinks using satellite data: A synthetic inter-comparison of measurement techniques and their performance as a function of space and time. **Atmos. Chem. Phys.**, v. 4, p. 523–538, 2004. 6, 14

HOUWELING, S.; KAMINSKI, T.; DENTENER, F.; LELIEVELD, J.; HEIMANN, M. Inverse modeling of methane sources and sinks using the adjoint of a global transport model. **J. Geophys. Res.**, d21, 26,137-26,160, v. 104, p. 137–26,160, 1999. 14

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). **Climate change 2001: the physical science basis**. contribution of working group i to the third assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom, 2001. 881p. Relatório técnico. Disponível em: <[http://www.grida.no/publications/other/ipcc\\_tar/](http://www.grida.no/publications/other/ipcc_tar/)>. 5, 111

\_\_\_\_\_. **Climate change 2007: the physical science basis**. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom and New York, NY, USA, 2007. 996p. Relatório técnico. Disponível em: <[http://www.ipcc.ch/publications\\_and\\_data/publications\\_ipcc\\_fourth\\_assessment\\_report\\_wg1\\_report\\_the\\_physical\\_science\\_basis.htm](http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm)>. 2, 5, 13

IYODA, E. M. **Inteligência computacional no projeto automático de redes neurais híbridas e redes neurofuzzy heterogêneas**. Tese de Doutorado - Universidade Estadual de Campinas (UNICAMP) - Campinas, SP: [s.n.] — Universidade Estadual de Campinas (UNICAMP), Campinas, SP, 2000. 85

JAIN A.; ZONGKER, D. Feature selection: Evaluation, application, and small sample performance. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, p. 153–158, 1997. 67



JEFFREY J. L.; VITTER, J. S. Complexity results on learning by neural nets. **machine learning**, p. 211–230, 1991. 85

JORDAN, M. I. Attractor dynamics and parallelism in a connectionist sequential machine. In: ANNUAL CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY, 8., 1986, Englewood Cliffs, NJ. **Proceedings...** Englewood Cliffs, NJ, 1986. p. 531–546. 63

KIRKPATRICK S.; GELATT JR., C. D.; VECCHI, M. P. Optimization by simulated annealing. **Science**, N. 4598, v. 220, p. 671–680, 1983. 91

KITANO, H. Designing neural networks using genetic algorithms with graph generation system. **Complex Systems**, v. 4, p. 461–476, 1990. 86

\_\_\_\_\_. Evolution, complexity, entropy and artificial reality. **Physica D**, v. 75, p. 239–263, 1994. 86

KLEIPOOL, Q. L.; JONGMA, R. T.; GLOUDEMANS, A. M. S.; SCHRIJVERA, H.; LICHTENBERG, G. F.; HEESA, R. M. van; MAURELLISA, A. N.; HOOGEVEEN, R. W. M. In-flight proton-induced radiation damage to sciamachy's extended-wavelength ingaas near-infrared detectors. **Infrared Physics & Technology**, 2005. 15, 41, 68

KOEHN, P. **Combining genetic algorithms and neural networks: the encoding problem**. Master Thesis, University of Tennessee, Knoxville, EUA — University of Tennessee, Knoxville, EUA, 1994. 86

KOELEMEIJER, R. B. A.; STAMMES, P.; HOVENIER, J. W.; HAAN, J. F. de. A fast method for retrieval of cloud parameters using oxygen a-band measurements from global ozone monitoring experiment. **J.Geophys.Res.**, v. 106, p. 3475–3490, 2001. 45

KOHONEN, T. **Associative memory: a system theoretic approach**. New York. Berlin-Heidelberg-New York: Berlin: Springer-Verlag, 1977. 60

\_\_\_\_\_. **Self-organizing maps**. New York: Berlin: Springer-Verlag, 1995. 58, 60

KOVACS, K. L. **Redes neurais artificiais: fundamentos e aplicacoes**. 2. ed. São Paulo: Editora Acadêmica, São Paulo, 1996. 50, 61

LELIEVELD J.; CRUTZEN, P.; DENTENER, F. Changing concentration, lifetime and climate forcing of atmospheric methane. **Tellus** **50B**, p. 128–150, 1998. [2](#), [13](#), [14](#)

LENOBLE, J. **Radiative transfer in scattering and absorbing atmospheres**: standard computational procedures. Virginia: A. Deepak Publishing, Hampton, Virginia, 1985. [29](#)

LERNER J. A.; WEISZ, E.; KIRCHENGAST, G. Temperature and humidity retrieval from simulated infrared atmospheric sounding interferometer (iasi) measurements. **J. Geophys. Res.**, **ACH 4-1-4-12**, v. 107, 2002. D14, doi:10.1029/2001JD900254. [68](#)

LICHTENBERG, G.; KLEIPOOL, Q.; KRIJGER, J. M.; SOEST, G. van; HEES, R. van; TILSTRA, L. G. Sciamachy level 1 data: Calibration concept and in-flight calibration. **Atmos. Chem. Phys. Discuss.**, v. 6, p. 5347–5367, 2006. [19](#), [68](#)

LIU, K. N. **An introduction to atmospheric radiation**. New York, NY: Academic Press, Orlando, 1980. [20](#), [28](#)

LIU, X.; CHANCE, K.; SIORIS, C. E.; SPURR, R. J. D.; KUROSU, T. P.; MARTIN, R. V.; NEWCHURCH, M. J. Ozone profile and tropospheric ozone retrievals from the global ozone monitoring experiment: Algorithm description and validation. **Journal of Geophysical Research**, v. 110, p. 1177–1194, 2005. D20307, doi:10.1029/2005JD006240. [6](#)

LORENC, A. C. Analysis methods for numerical weather prediction. **Quart. J. Roy. Meteor. Soc.**, **112**, v. 112, p. 1177–1194, 1986. [41](#)

LUZ, E. F. P.; VELHO, H. F. C.; BECCENERI, J. C.; ROBERTI, D. R. Estimating atmospheric area source strength through particle swarm optimization. In: INVERSE PROBLEMS, DESIGN AND OPTIMIZATION SYMPOSIUM, 2007, Miami, FL. **Proceedings...** Miami, FL, 2007. p. 354–359. [86](#)

MADDY, E. S.; BARNET, C.; GOLDBERG, M.; SWEENEY, C.; LIU, X. CO<sub>2</sub> retrievals from the atmospheric infrared sounder: Methodology and validation. **J. Geophys. Res.**, v. 113, 2008. D11301, doi:10.1029/2007JD009402. [17](#)

MARTIN, R. V.; JACOB, D. J.; THOMAS, P. K.; PAUL, I. P.; MATHEW, J. E. Global inventory of nitrogen oxide emissions constrained by space-based

observations of NO<sub>2</sub> columns. **J. Geophys. Res.**, v. 108, 2003. D17, 4537, doi: 10.1029/2003JD003453. 6

MCCARTHY, J. J.; CANZIANI, O. F.; LEARY, N. A.; DOKKEN, D. J.; WHITE, K. S. **Climate change 2001: impacts, adaptation, and vulnerability**. New York: Cambridge University Press, Cambridge, 1033. 11, 2001. 2

MCCLELLAND J. L.; RUMELHART, D. E. Parallel distributed processing-explorations in the microstructure of cognition. **Psychological and Biological Models**, MIT Press, Cambridge, MA, v. 2, 1986. 47

MCCULLOCH W. W.; PITTS, W. A logical calculus of ideas imminent in nervous activity. **Bull. Math. Biophys.**, v. 5, p. 115–133, 1943. 47, 52

\_\_\_\_\_. How we know universals. **Bull. Math. Biophys.**, v. 9, p. 127–147, 1947. 47

MENCZER F.; PARISI, D. Evidence of hyperplanes in the genetic learning of neural networks. **Biol. Cybernet.**, v. 66, p. 283–289, 1992. 86

METROPOLIS, N.; ROSENBLUTH, A.; ROSENBLUTH, M.; TELLER, A.; TELLER, E. Equation of state calculations by fast computing machines. **Journal of Chemical Physics**, v. 21, p. 1087–1092, 1953. 91

MILLER, C. E.; CRISP, D.; DECOLA, P. L.; OLSEN, S. C.; RANDERSON, J. T. Precision requirements for space-based XCO<sub>2</sub> data. **J. Geophys. Res.**, v. 112, 2007. D10314, doi:10.1029/2006JD007659. 12

MILLER G. F.; TODD, P. M.; HEDGE, S. U. Designing neural networks using genetic algorithms. **Neural Networks**, v. 4, p. 53–60, 1991. 85, 86, 156

MINSKY M.; PAPERT, S. **Perceptron: an introduction to computational geometry**. Cambridge, MA: MIT Press, 1969. 47, 57

MLADENOVIC N.; HANSEN, P. Variable neighbourhood search. **Computers and Operations Research**, v. 24, p. 1097–1100, 1997. 87

MULLER J. F.; STAVRAKOU, T. Inversion of CO and NO<sub>x</sub> emissions using the adjoint of the images model. **Atmos. Chem. Phys.**, v. 5, p. 1157–1186, 2005. 6

MUNIZ W. B.; CAMPOS VELHO, H. F.; RAMOS, F. M. Entropy- and tikhonov-based regularization techniques applied to the backwards heat equation. **Journal of Computational and Applied Mathematics**, v. 103, p. 145–163, 1999. 37

MUNIZ W. B.; RAMOS, F. M.; VELHO, H. F. C. Entropy- and tikhonov-based regularization techniques applied to the backwards heat equation. **Computers & Mathematics with Applications**, v. 40(8/9), p. 1071–1084, 2000. 37

NOBRE C. A.; SELLERS, P. J.; SHUKLA, J. Amazonian deforestation and regional climate change. **Journal of Climate**, v. 4, p. 957–988, 1991. 5

NOEL S.; BUCHWITZ, M.; BURROWS, J. P. First retrieval of global water vapour column amounts from sciamachy measurements. **Atmos. Chem. Phys.**, v. 4, p. 111–125, 2004. 38

OSORIO, F. S. **Um estudo sobre reconhecimento visual de caracteres através de redes neurais**. Dissertação de Mestrado., Porto Alegre: CPGCC., 1991. 50

PAES, F. F.; CAMPOS, H. F.; LUZ, E. P.; CARVALHO, A. R. Identifying atmospheric pollutant sources using artificial neural networks. In: AGU 2008 JOINT ASSEMBLY, 2008, Fort Lauderdale. **Proceedings...** Fort Lauderdale (FL), USA: SBMAC, 2008. p. 131–134. 23

PAES F. F., C. V. H. F. Neuro-estimador do ciclo diurno de CO<sub>2</sub>. **Ciência e Natura**, vol. esp, p. 293–296, 2009. 23

PALMER P. I.; RAYNER, P. Failure to launch. **Nature Geosci.**, v. 2, p. 247, 2009. 17

PAPALAMBROS P. Y.; WILDE, D. J. **Principles of optimal design: modeling and computation**. Second edition. New York, NY: Cambridge: Univ. Press, England, 1988. 100

PERRONE, M. **Improving regression estimation: averaging methods for variance reduction with extensions to general convex measure optimization**. 1993. 194 p. PhD thesis, Brown University — Brown University, 1993. 82

PÉTRON, G.; GRANIER, C.; KHATTATOV, B.; YUDIN, V.; LAMARQUE, J. F. Monthly CO surface sources inventory based on the 2000-2001 mopitt satellite data. **Geophys. Res. Lett.**, v. 31, 2004. L21107, doi: 10.1029/2004GL020560. 6

PLATT, U. **Differential optical absorption spectroscopy (DOAS), air monitoring by spectroscopic techniques**. Second edition. [S.l.]: M. Sigrist, ed., John Wiley & Sons, Inc., 1994. 44

PLATT U.; PERNER, D. Measurements of atmospheric trace gases by long path differential uv/visible absorption spectroscopy. **In Springer Ser. Opt. Sci., Optical and Laser Remote Sensing (D. K. Killinger A. Mooradian)**. Heidelberg: Springer., v. 39, p. 95–105, 1983. 44

PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. **Numerical recipes in C: the art of scientific computing**. Second edition. Cambridge: Cambridge University Press, 1992. 39

PRUNET P.; THÉPAUT, J. N.; CASSÉ, V. The information content of clear sky iasi radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 124, p. 211–241, 1998. 114

RABIER, F.; FOURRIER, N.; CHAFAI, D.; PRUNET., P. Channel selection methods for infrared atmospheric sounding interferometer radiances. **Quart J. Roy. Meteor. Soc.**, v. 111, p. 974–975, 2002. 21, 67, 68

RABIER, F.; JARVINEN, H.; KLINKER, E.; MAHFOUF, J. F.; SIMMONS, A. The ecmwf operational implementation of four-dimensional variational assimilation, i: Experimental results with simplified physics. **Quart J. Roy. Meteor. Soc.**, v. 126(A), p. 1143–1170, 2000. 43

RAMOS, F. M.; VELHO, H. F. C.; CARVALHO, J. C.; FERREIRA, N. J. Novel approaches on entropic regularization. **Inverse Problems**, v. 15, p. 1139–1148, 1999. 37

RANDEL, W. J.; WU, F.; III, J. M. R.; ROCHE, A.; WATERS, J. W. Seasonal cycles and qbo variations in stratospheric CH<sub>4</sub> and h<sub>2</sub>o observed in uars haloe data. **J. Atmos. Sci.**, v. 55, p. 163–185, 1998. 14

RAYNER, P. J.; LAW, R. M.; O'BRIEN, D. M.; BUTLER, T. M. Global observations of the carbon budget. 3. initial assessment of the impact of satellite

- orbit, scan geometry, and cloud on measuring CO<sub>2</sub> from space. **Journal of Geophysical Research**, v. 107 (D21), 2002. 6, 11, 12, 13, 146
- RIEU, H.; ESCOBAR, J.; SCOTT, N. A.; CHEDIN, A. Ssm/t forward modelling using neural networks. **J. Quant. Spectros. Radiat. Transfer**, issue 6, v. 56, p. 821–833, 1996. 8
- RIPLEY, B. D. **Pattern recognition and neural networks**. Cambridge: Cambridge University Press, 1996. 98
- RODGERS, C. D. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. **Reviews of Geophysics and Space Physics**, n.2, v. 14, p. 609–624, 1976. 21, 33, 75, 76
- \_\_\_\_\_. Information content and optimization of high spectral resolution measurements. In: OPTICAL SPECTROSCOPIC TECHNIQUES AND INSTRUMENTATION FOR ATMOSPHERIC AND SPACE RESEARCH, SPIE, 2830, 2., 1996, San Jose, California. **Proceedings...** San Jose, California: SPIE, 1996. p. 136–147. 68, 73, 77, 114
- \_\_\_\_\_. **Inverse methods for atmospheric sounding: theory and practise**. [S.l.]: World Scientific Publishing Co. Ltd., 1997. 13, 43, 44
- ROJAS, R. **Neural networks: a systematic introduction**. Berlin: Springer-Verlag, 1996. 66
- ROSENBLATT, F. **Principles of neurodynamics**. New York, NY: Spartan Press, Washington, DC, 1961. 47
- ROZANOV, V. V.; BUCHWITZ, M.; EICHMANN, K. U.; BEEK, R. de; BURROWS, J. P. Sciatran - a new radiative transfer model for geophysical applications in the 240- 2400 nm spectral region: The pseudo-spherical version. **Advances in Space Research**, 11, v. 29, p. 1831–1835, 2002. 32
- ROZANOV, V. V.; DIEBEL, D.; SPURR, R. J.; BURROWS, J. P. Gometran: a radiative transfer model for the satellite project gome - the plane-parallel version. **J. Geophys. Res.**, v. 102 (D14), p. 16683–16695, 1997. 29, 30
- RUMELHART D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations of back-propagating error. **Nature**, v. 323, p. 533–536, 1986. 48, 56

- SAKUMA, F.; BRUEGGE, C. J.; RIDER, D.; BROWN, D.; GEIER, S.; KAWAKAMI, S.; KUZE, A. Oco/gosat preflight cross-calibration experiment. **Geoscience and Remote Sensing, IEEE Transactions, issue 1, part 2**, v. 48, p. 585–599, 2009. 24
- SCHAERER G.; WILHEIT, T. T. A passive microwave technique for profiling of atmospheric water vapor. **Radio Sci.**, v. 14, p. 371–375, 1979. 69
- SCHAFER J. D.; CARUANA, R. A.; ESHELMAN, L. J. Using genetic search to exploit the emergent behavior of neural networks. **Physica D**, v. 42, p. 244–248, 1990. 86
- SCHNEISING, O.; BUCHWITZ, M.; BURROWS, J. P.; BOVENSMANN, H.; REUTER, M.; NOTHOLT, J.; MACATANGAY, R.; WARNEKE, T. Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite - part 1: Carbon dioxide. **Atmos. Chem. Phys.**, v.8, v. 8, p. 3827–3853, 2008. 13, 37
- SCHNEISING, O.; BUCHWITZ, M.; BURROWS, J. P.; BOVENSMANN, H.; BERGAMASCHI, P.; PETERS, W. Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite - part 2: Methane. **Atmos. Chem. Phys. Discuss.**, v. 8, p. 443–465, 2008. 6, 37
- SCIENCES, B. on A.; CLIMATE. **Understanding and responding to climate change, highlights of national academies reports**. Washington, DC, 2008. Disponível em: <[http://americasclimatechoices.org/climate\\_change\\_2008\\_final.pdf](http://americasclimatechoices.org/climate_change_2008_final.pdf)>. 2
- SEGATTO E. C.; COURRY, D. V. A power transformer protection with recurrent ann saturation correction. In: POWER ENGINEERING SOCIETY GENERAL MEETING, 2005, San Francisco, USA. **Proceedings...** San Francisco: IEEE, 2005. 63
- \_\_\_\_\_. A differential relay for power transformers using intelligent tools. **IEEE Transactions on Power Systems**, n.3, v. 21, p. 1154–1162, 2006. 63
- SERBEDZIJA, N. B. Simulating artificial neural networks on parallel architectures. **Computer**, v. 29(3), p. 56–63, 1996. 8
- SHANNON, C. E. A mathematical theory of communication. **Bell System Technical Journal**, v. 27, p. 379–423, 1948. 73, 75

SHIGUEMORI, E. H.; SILVA, J. D. S.; VELHO, H. F. C.; CARVALHO, J. Neural network based models in the inversion of temperature vertical profiles from satellite data. **Inverse Problems in Engineering, England**, n.5, v. 14, p. 543–556, 2006. 9

SHIGUEMORI E. H.; CAMPOS VELHO, H. F.; SILVA, J. D. S. Atmospheric temperature retrieval from satellite data: new non-extensive artificial neural network approach. In: ACM SYMPOSIUM ON APPLIED COMPUTING, 2008, ACM, New York, NY, USA. **Proceedings...** New York, NY, USA, 2008. p. 1688–1692. 61

SKUPIN, J.; NOEL, S.; WUTTKE, M. W.; GOTTWALD, M.; BOVENSMANN, H.; WEBER, M.; BURROWS, J. P. Sciamachy solar irradiance observation in the spectral range from 240 to 2380 nm. **Advances in Space Research, Issue 3**, v. 35, p. 370–375, 2005. 18

SMITH W. L.; WOOLF, H. M.; HAYDEN, C. M. The tiros-n operational vertical sounder. **Bulletin of the American Meteorological Society**, v. 60, p. 1117–1187, 1979. 12

SOUSA, F.; RAMOS, F.; GALSKI, R.; MURAOKA, I. Generalized extremal optimization: a new meta-heuristic inspired by a model of natural evolution. In: CASTRO, L. N. de; ZUBEN, F. J. V. (Ed.). **Recent developments in biologically inspired computing**. Hershey PA: Idea Group Inc., 2004. p. 41–60. 88

SOUSA, F. L.; RAMOS, F. M.; PAGLIONE, P.; GIRARDI, R. M. New stochastic algorithm for design optimization. **AIAA Journal, Number 9**, v. 41, p. 1808–1818, 2003. 88

SOUSA F. L.; RAMOS, F. M. Function optimization using extremal dynamics. In: INTERNATIONAL CONFERENCE ON INVERSE PROBLEMS IN ENGINEERING (CD-ROM), 4., 2002, Rio de Janeiro. **Proceedings...** Rio de Janeiro, 2002. p. 131–134. 88

SOUSA F. L.; VLASSOV, V.; RAMOS, F. M. Generalized extremal optimization for solving complex optimal design problems. **Lectures Notes on Computer Science**, v. 2723, p. 375–376, 2003. 88



SOUTO, R. P.; VELHO, H. F. C.; STEPHANY, S.; KAMPEL, M. Chlorophyll concentration profiles from in situ radiances by ant colony optimization. **Journal of Physics**, v. 124, 2008. 86

STROW, L. L.; HANNON, S. E.; MACHADO, S. D.-S.; MOTTELER, H. E.; TOBIN, D. C. An overview of the airs radiative transfer model. **IEEE Transactions on Geoscience and Remote Sensing**, v. 41, p. 303–313, 2002. 13, 40, 151

\_\_\_\_\_. Validation of the atmospheric infrared sounder radiative transfer algorithm. **J. Geophys. Res.**, v. 111, 2006. D09S06, doi:10.1029/2005JD006146. 17

SUSSKIND J.; BARNET, C.; BLAISDELL, J. Retrieval of atmospheric and surface parameters from airs/amsu/hsb data in the presence of clouds. **IEEE Transactions on Geoscience and Remote Sensing**, n.2, v. 41, p. 390–409, 2003. 21, 40

TIKHONOV A. N.; ARSEININ, V. Y. **Solutions of ill-posed problems**. USA: Washington, V. H. Wiston & Sons, 1977. 36

TURQUETY S.; HADJI-LAZARO, J.; CLERBAUX, C. First satellite ozone distributions retrieved from nadir high-resolution infrared spectra. **Geophys. Res. Lett.**, 10.1029/2002GL016431, v. 29, 2002. 9

TWOMEY, S. **Introduction to the mathematics of inversion in remote sensing and interactive measurements**. Amsterdam: Elsevier Scientific, 1977. 32, 36

VALKS P.; VAN OSS, R. **TOGOMI algorithm theoretical basis document**. De Bilt, 2003. Relatório técnico. Disponível em: <[http://www.gse-promote.org/services/ozone\\_nrt/togomi\\_ATBD\\_v1.21.pdf](http://www.gse-promote.org/services/ozone_nrt/togomi_ATBD_v1.21.pdf)>. 44

VEEFKIND J. P.; DE HAAN, J. F. **OMI algorithm theoretical basis document**. De Bilt, 2001. Relatório técnico. 44

VELHO, H. F. C. Problemas inversos: conceitos básicos e aplicações. In: ENCONTRO DE MODELAGEM COMPUTACIONAL, RESUMOS (MINI-CURSO CONVIDADO), 4., 2001, São Carlos, SP. **Proceedings...** São Carlos: SBMAC, 2001. p. 63–79. 37

- VELHO H. F.; SHIGUEMORI, E. H. C.; SILVA, J. D. S. Hardware implementation for the atmospheric temperature retrieval from satellite data. In: INVERSE PROBLEMS, DESIGN AND OPTIMIZATION SYMPOSIUM, 2007, Florida, USA. **Proceedings...** Florida, USA, 2007. 8, 156
- WATANABE, H.; ISHIHARA, H.; HIRAKI, K.; MATSUNAGA, T.; YOKOTA, T. Development of gosat ground data system at nies and data processing strategy. **Journal of The Remote Sensing Society of Japan**, v. 28, p. 127–132, 2008. 24
- WHITLEY D.; STARKWEATHER, T.; BOGART, C. Genetic algorithm and neural networks: optimizing connections and connectivity. **Computing**, v. 14, p. 347–361, 1989. 86
- WIDROW B.; HOFF, M. E. Adaptive switching circuits. In: IRE WESCON CONVENTION RECORD, PART 4, 1960, Western Electronic Show and Convention, Los Angeles, California. **Proceedings...** Los Angeles, California: Institute of Radio Engineers, 1960. p. 96–104. 48
- XIONG, X.; BARNET, C.; MADDY, E.; SWEENEY, C.; LIU, X.; ZHOU, L.; GOLDBERG, M. Characterization and validation of methane products from the atmospheric infrared sounder (airs). **Journal of Geophysical Research**, v. 113, 2008. G00A01, doi:10.1029/2007JG000500. 6, 142
- YAO X.; LIU, Y. A new evolutionary system for evolving artificial neural networks. **IEEE Transactions on Neural Network**, v. 8, p. 694–713, 1997. 84, 86, 94, 156
- YOKOTA, T.; OGUMA, H.; MORINO, I.; INOUE, G. A nadir looking swir fts to monitor CO<sub>2</sub> column density for japanese gosat project. In: INTERNATIONAL SYMPOSIUM ON SPACETECHNOLOGY AND SCIENCE, 24., 2004, Miyazaki, Japan. **Proceedings...** Miyazaki: Japan Society for Aeronautical and Space Sciences, 2004. p. 887–889. 17
- ZUPANSKI D.; MESINGER, F. Regional 4-dimensional variational data assimilation in a quasi-operational forecasting environment. **Monthly Weather Review**, v. 121, p. 2396–2408, 1993. Doi:10.1175/1520-0493. 43
- \_\_\_\_\_. Four-dimensional variational assimilation of precipitation data. **Monthly Weather Review**, v. 123, p. 1112–1127, 1995. Doi:10.1175/1520-0493. 43

## **PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE**

### **Teses e Dissertações (TDI)**

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

### **Manuais Técnicos (MAN)**

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

### **Notas Técnico-Científicas (NTC)**

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

### **Relatórios de Pesquisa (RPQ)**

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

### **Propostas e Relatórios de Projetos (PRP)**

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

### **Publicações Didáticas (PUD)**

Incluem apostilas, notas de aula e manuais didáticos.

### **Publicações Seriadas**

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

### **Programas de Computador (PDC)**

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

### **Pré-publicações (PRE)**

Todos os artigos publicados em periódicos, anais e como capítulos de livros.