



Ministério da
Ciência e Tecnologia



sid.inpe.br/mtc-m19/2010/09.20.14.46-TDI

ASSIMILAÇÃO DE DADOS COM REDES NEURAIS ARTIFICIAIS EM MODELO DE CIRCULAÇÃO GERAL DA ATMOSFERA

Rosângela Saher Corrêa Cintra

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelo Dr. Haroldo Fraga de Campos Velho, aprovada em 27 de setembro
de 2010.

URL do documento original:

<[http://urlib.net/ 8JMKD3MGP7W/389UKPS](http://urlib.net/8JMKD3MGP7W/389UKPS) >

INPE
São José dos Campos
2010

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^a Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr^a 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 e Tecnologia



sid.inpe.br/mtc-m19/2010/09.20.14.46-TDI

ASSIMILAÇÃO DE DADOS COM REDES NEURAIS ARTIFICIAIS EM MODELO DE CIRCULAÇÃO GERAL DA ATMOSFERA

Rosângela Saher Corrêa Cintra

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelo Dr. Haroldo Fraga de Campos Velho, aprovada em 27 de setembro
de 2010.

URL do documento original:

<[http://urlib.net/ 8JMKD3MGP7W/389UKPS](http://urlib.net/8JMKD3MGP7W/389UKPS) >

INPE
São José dos Campos
2010

Dados Internacionais de Catalogação na Publicação (CIP)

Cintra, Rosângela Saher Corrêa.
C493a Assimilação de dados com redes neurais artificiais em modelo
de circulação geral da atmosfera / Rosângela Saher Corrêa Cintra
. – São José dos Campos : INPE, 2010.
xxviii+172 p. ; (sid.inpe.br/mtc-m19/2010/09.20.14.46-TDI)

Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.
Orientador : Dr. Haroldo Fraga de Campos Velho.

1. Assimilação de dados. 2. Redes neurais artificiais. 3. Previsão numérica de tempo. 4. Perceptron de múltiplas camadas. 5. Modelo de circulação geral. I.Título.

CDU 681.3.019

Copyright © 2010 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 © 2010 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. Elbert Einstein Nehrer Macau



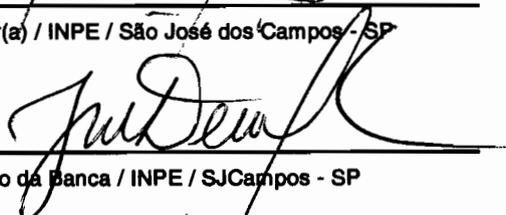
Presidente / INPE / São José dos Campos - SP

Dr. Haroldo Fraga de Campos Velho



Orientador(a) / INPE / São José dos Campos - SP

Dr. José Demisio Simões da Silva



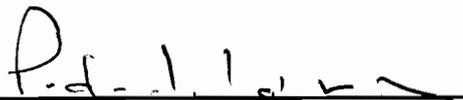
Membro da Banca / INPE / SJCampos - SP

Dr. José Antônio Aravéquia



Membro da Banca / INPE / São José dos Campos - SP

Dr. Pedro Leite da Silva Dias



Convidado(a) / LNCC / Petrópolis - RJ

Dr. Aluizio Fausto Ribeiro Araujo



Convidado(a) / UFPE / Recife - PE

Aluno (a): Rosângela Saher Corrêa Cintra

São José dos Campos, 27 de setembro de 2010

“O êxito ou o fracasso de sua vida não depende de quanta força você põe em uma tentativa, mas da persistência no que fizer”.

J. Y. LELOUP

*Ao orientador e amigo Haroldo, a minha filha e melhor
amiga Luiza, a minha filha que me surpreende a cada dia
Paula e a meu esposo Eleutério.*

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter concedido as graças necessárias para a realização deste sonho.

Ao Instituto Nacional de Pesquisas Espaciais (INPE), em especial ao Programa de Pós- Graduação em Computação Aplicada (CAP), pela oportunidade de crescimento profissional.

Ao Laboratório Associado de Computação e Matemática Aplicada (LAC) pela oportunidade de estudos, apoio profissional e utilização de recursos e instalações. E ainda pelo apoio financeiro em viagens nacionais e internacionais para cursos e apresentação de trabalhos acadêmicos.

Ao Dr. Haroldo Fraga de Campos Velho, pela orientação e pela confiança na execução deste desafio. Agradeço ainda seus valiosos conselhos, ensinamentos, atenção, otimismo, motivação, amizade e por dividir um pouco de sua imensa sabedoria.

Ao professor José Demísio Simões da Silva que recorri quando necessário, pelo apoio e pela confiança no meu trabalho.

A amiga Maria Cristina Braga Peloggia pelo apoio administrativo e pelo constante incentivo. Ao Dr. Airam Jonatas Preto pelo apoio em recursos computacionais.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro que possibilitou a realização de parte da pesquisa em Maryland, Estados Unidos.

Ao pesquisador e amigo Dr. Ricardo Todling, do Global Modeling and Assimilation Office (GMAO/NASA), por ter me recebido em sua instituição, pelo apoio e colaboração na pesquisa aplicada e pelo conhecimento compartilhado.

À Dra. Eugênia Kalnay da Universidade de Maryland/EUA pelo incentivo e contribuição ao trabalho através das enumeráveis informações sobre métodos de assimilação de dados. Agradeço ao Dr. Takemasa Myoshi da Universidade de Maryland/EUA pela concessão dos programas desenvolvidos por ele.

Aos colegas do CPTEC, em especial ao José Roberto Rozante pela amizade, colaboração e troca de conhecimentos em questões de meteorologia.

Aos colegas de curso, pela amizade e por muito colaborarem com o meu crescimento profissional.

Agradeço em especial às minhas filhas Luiza e Paula pelo constante apoio e pela paciência e atenção que sempre tiveram. E ao Eleutério, que tanto me incentivou e apoiou durante todo este trabalho e pelo constante carinho, paciência, compreensão, companheirismo.

Enfim, a todos aqueles que de alguma forma participaram desta etapa da minha vida.

RESUMO

Sistemas de previsão de tempo requerem um modelo para sua evolução temporal e uma estimativa do estado atual do sistema. A previsão numérica de tempo (PNT) integra as equações da dinâmica da atmosfera com os processos físicos e pode prever o estado futuro da atmosfera. Assimilação de dados provê a estimativa inicial da atmosfera onde combina informações de observações e de uma previsão anterior de curto prazo, produzindo uma estimativa de estado atual. Neste trabalho investigou-se a técnica de assimilação de dados com Redes Neurais Artificiais (RNA). As previsões de curto prazo são de um modelo global de equações primitivas. O modelo SPEEDY (*Simplified Parameterizations, primitivE-Equation DYnamics*) é um modelo de circulação geral da atmosfera com grade em coordenadas tridimensionais. Molteni (2003) mostrou que o modelo SPEEDY tem características semelhantes aos modelos atmosféricos do estado-da-arte. Para o esquema de assimilação de dados aplicou-se uma RNA supervisionada (Perceptron de Múltiplas Camadas) para emular os resultados da técnica de Filtro de Kalman por Conjunto Transformado e Localizado (*Local Transform Ensemble Kalman Filter - LETKF*). O esquema LETKF é uma aproximação do Filtro de Kalman onde um conjunto tipo Monte-Carlo de previsões de curto prazo são usadas para estimar as covariâncias do erro do modelo de previsão. O método com RNA neste trabalho pode ser descrito como um processo de assimilação, onde a rede neural após treinada obtém os resultados em função do modelo de estado SPEEDY e de observações sintéticas. A estratégia do treinamento supervisionado da RNA, a implementação das redes, do modelo e das observações são apresentadas. A ênfase principal nesta técnica é a velocidade computacional na obtenção da condição inicial do modelo de estado que acelera todo o processo de previsão numérica do tempo. Os resultados numéricos demonstrem a eficiência da técnica de assimilação de dados atmosféricos utilizando RNA, pois estes apresentam-se muito próximos aos resultados da assimilação de dados realizada com o método LETKF. As simulações demonstram a grande vantagem no uso de redes neurais: o melhor desempenho computacional.

DATA ASSIMILATION WITH ARTIFICIAL NEURAL NETWORKS IN ATMOSPHERIC GENERAL CIRCULATION MODEL

ABSTRACT

Weather forecasting systems require a model for the time evolution and an estimate of the current state of the system. The numerical weather prediction (NWP) incorporates the equations of atmospheric dynamics with physical process and it can predict the future state of the atmosphere. Data assimilation provides such an initial estimate of the atmosphere where it combines information from observations and from a prior short-term forecast producing an current state estimate. This work investigated the approach of data assimilation with Artificial Neural Networks (ANN). The short-term predictions are from a global primitive equation model, the SPEEDY model *Simplified parameterizations, primitive-Equation Dynamics*, simplified physical processes of an atmospheric general circulation with resolution in tridimensional coordinates. Molteni (2003) showed that the SPEEDY model has similar characteristics to the state-of-art atmospheric models. For the data assimilation scheme, it applied a supervised ANN Multilayer Perceptron to emulate the analysis results for *Local Ensemble Transform Kalman Filter* (LETKF). LETKF is an approximation of Kalman filter, with Monte-Carlo ensemble of short-term forecasts to estimate the forecast model error covariances. The method using RNA in this work can be described as a process of data assimilation, where the ANN trained after obtaining the results, like a function of the state model SPEEDY and its synthetic observations. The strategy of ANN supervised training, the implementations of networks and the model and the observations are presented. The main emphasis of this technique is the computational speed in obtaining the initial condition for state model that accelerates the whole process of numerical weather prediction. The numerical results demonstrate the effectiveness of this ANN technique in atmospheric data assimilation because these have been very close to the results compared with LETKF data assimilation results. The simulations demonstrate the great advantage in using neural networks: the best computational performance.

LISTA DE FIGURAS

	<u>Pág.</u>
1.1 História resumida dos principais algoritmos de assimilação de dados utilizados em meteorologia e oceanografia, classificados de acordo com sua complexidade(e custos) de execução e sua aplicabilidade para os problemas em tempo real.	6
2.1 Sistema de Previsão de Tempo.	16
2.2 Divisão vertical da atmosfera.	21
2.3 Coordenadas Esféricas.	28
2.4 Tuncamento do método espectral	35
2.5 Parametrizações físicas usadas em modelos de Tempo e Clima.	41
3.1 Tipos de Assimilação de Dados Sequencial	47
3.2 Ciclo típico de Assimilação de Dados.	52
3.3 Na abordagem local as observações dentro de um raio de influência são considerada. Os pontos de grade do modelo (uniformemente distribuída) e observações (irregularmente distribuídas).	79
4.1 Representação de um neurônio	84
4.2 Modelo simplificado de um neurônio marcado como i	86
4.3 Funções de ativação mais utilizadas.	87
4.4 Perceptron multicamdas com uma camada escondida	91
4.5 Esquema do treinamento supervisionado	93
4.6 Esquemas de assimilação de dados com RNA sendo (a) pesquisa de Nowosad (2001) e (b) pesquisa de Härter (2004).	104
5.1 Simulação rede de estações de <i>observação</i> realista (415 estações).	109
5.2 Divisão do globo em seis regiões iguais conforme tabela 5.3.	114
5.3 Campos de pressão (PS) [hPa] na superfície das análises geradas pelos métodos de assimilação RNA(c) e LETKF(a) e do estado verdadeiro(b). Diferenças entre as duas análises(d) (PMC_{ps} e LETKF). Diferença entre análise da (PMC_{ps} e estado verdadeiro)(e), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação	121

5.4	Campos de componente do vento zonal (u) [m/s] - superfície das análises geradas pelos métodos de assimilação RNA(c) e LETKF(a) e do estado verdadeiro. Diferença entre as duas análises(d) (PMC_u e LETKF)e a diferença entre análise da (PMC_u e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação	122
5.5	Campos de componente do vento meridional (v)[m/s] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro Diferenças entre as duas análises(d) (PMC_v e LETKF) e diferença (e) entre análise da (PMC_v e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação	123
5.6	Campos de temperatura absoluta (T)[K] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_t e LETKF) e diferença (e) entre análise da (PMC_t e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação	124
5.7	Campos de umidade específica(q)[Kg/Kg] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_q e LETKF) e diferença (e) entre análise da (PMC_q e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação	125
5.8	Campos de vento zonal (u)[m/s]- 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_u e LETKF) e a diferença (e) entre análise da (PMC_u e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação	126
5.9	Campos de vento meridional (v)[m/s] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_v e LETKF) e a diferença (e) entre análise da (PMC_v e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação	127

5.10	Campos de temperatura absoluta (T)[K] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_T e LETKF) e a diferença (e) entre análise da (PMC_T e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação	128
5.11	Campos de umidade específica (q)[Kg/Kg] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_q e LETKF) e a diferença (e) entre análise da (PMC_q e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação.	129
5.12	Campo de diferença das Médias dos campos da pressão na superfície das análises RNA e das análises LETKF para o ciclo das 120 execuções. . . .	130
5.13	Campo de diferenças da Média dos campos de umidade específica das análises RNA e das análises LETKF para o ciclo das 120 execuções. . . .	130
5.14	Campo de diferenças da Média dos campos da temperatura das análises RNA e das análises LETKF para o ciclo das 120 execuções.	130
5.15	Média Zonal Global de pressão à superfície [hPa] das análises geradas pelos métodos de assimilação RNA e LETKF e do estado VERDADE de 1 a 31 de janeiro de 1985 em 120 execuções do ciclo de assimilação. As médias foram fixadas na latitude 40N, na latitude (0) equador e na latitude 40S e longitude (0° a 360°).	131
5.16	Média Zonal Global de Temperatura [K] das análises geradas pelos métodos de assimilação RNA e LETKF e do estado VERDADE de 1 a 31 de janeiro de 1985 em 120 execuções do ciclo de assimilação. As médias foram fixadas na latitude 40N, na latitude(0) equador, na latitude 40S e longitude (0° a 360°).	132
5.17	Média meridional de todo o período do ciclo de assimilação do erro médio quadrático da análise em relação ao estado Verdade; os erros das análises do LETKF em linha (azul) e os erros das análises da PMC em linha (vermelha) para a variável pressão à superfície (a) e a diferença entre os erros (b).	133

5.18	Média meridional de todo o período do ciclo de assimilação do erro médio quadrático da análise em relação ao estado Verdade; os erros das análises do LETKF em linha (azul) e os erros das análises da PMC em linha (vermelha) para a variável temperatura absoluta à 500 hPa (a) e a diferença entre os erros (b).	134
5.19	Média temporal do ciclo de assimilação (01 a 31 de janeiro de 1985) do erro médio quadrático em relação ao modelo Verdade; os erros das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), para (a) pressão e (b) temperatura ambos à superfície.	135
5.20	Média temporal do ciclo de assimilação (01 a 31 de janeiro de 1985) do erro médio quadrático em relação ao modelo Verdade; os erros das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), para (a) umidade específica à superfície e (b) umidade específica à 500 hPa.	136
5.21	Perfil do erro médio quadrático das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), no ponto (20S,87W) o perfil da temperatura em níveis sigma (a) e a diferença entre os dois perfis(b).	137
5.22	Perfil do erro médio quadrático das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), no ponto (20S,87W) o perfil de umidade específica em níveis sigma (a) e a diferença entre os dois perfis(b).	138
5.23	Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de pressão [hPa] à superfície.	139
5.24	Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões da componente do vento zonal [m/s] à superfície.	140
5.25	Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões da componente do vento meridional [m/s] à superfície.	141
5.26	Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de temperatura absoluta [K] à superfície.	142

5.27	Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de umidade específica [Kg/Kg] à superfície.	143
5.28	Tempos de execução dos 112 ciclos de assimilação de dados completos com análises geradas pelo método RNA e pelo método LETKF em milésimos de segundos, para o período de 01 de fevereiro de 1982 a 28 de fevereiro de 1982 em quatro horários sinóticos	148
5.29	Tempos de execução dos 112 ciclos de assimilação de dados completos com análises geradas pela PNA/PMC em <i>milésimos</i> de segundos, para o período de 01 de fevereiro de 1982 a 28 de fevereiro de 1982 em quatro horários sinóticos	148
6.1	Crescimento da quantidade dados para a Assimilação	152

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Níveis verticais das saídas do modelo SPEEDY.	40
4.1 Resumo do Algoritmo de Retro-Propagação do erro	99
5.1 Desvio Padrão aplicado aos erros das observações.	108
5.2 Parâmetro Variável na arquitetura das Redes PMC.	111
5.3 Regiões do globo terrestre.	114
5.4 Convergência do Treinamento da PMC_{ps} . Conjunto de dados para (1 nível) a superfície.	116
5.5 Convergência do Treinamento da PMC_T . Conjunto de dados para (7) níveis.	117
5.6 Convergência do Treinamento da PMC_q . Conjunto de dados para 7 níveis.117	117
5.7 Convergência do Treinamento da PMC_u . Conjunto de dados para 7 níveis.117	117
5.8 Convergência do Treinamento da PMC_v . Conjunto de dados para 7 níveis. (*) erro aceitável.	118
5.9 Tempo Total da execução de 112 ciclos de assimilação de dados completos (análise e previsão). Tempo de análise é o tempo de 112 execuções para gerar a análise, excluindo a execução do modelo SPEEDY, para uma previsão ou para o conjunto de previsões de 30 membros.	149
5.10 Tempo médio da execução de cada componente individual do ciclo de as- similação de dados. Tempo computado foi em minutos, segundos e milési- mos de segundos.	149

LISTA DE ABREVIATURAS E SIGLAS

CPTEC	–	Centro de Previsão de Tempo e Estudos Climáticos
NCEP	–	<i>National Center for Environmental Prediction</i>
ECMWF	–	<i>European Centre of Medium-range Weather Forecast</i>
MSC	–	<i>Meteorological Service of Canada</i>
LAC	–	Laboratório Associado de Computação e Matemática Aplicada
CNPq	–	Conselho Nacional de Desenvolvimento Científico e Tecnológico
GFS	–	<i>Global Forecast System Model</i>
INPE	–	Instituto Nacional de Pesquisas Espaciais
GPSAS	–	<i>Global Physical-space Statistical Analysis System</i>
MCGA	–	Modelo de Circulação Geral da Atmosfera
MOCG	–	Modelo Oceânico de Circulação Geral
EKF	–	<i>Extended Kalman filter</i>
FK	–	Filtro de Kalman
PMC	–	Perceptron MúltiCamadas
PMC_u	–	Perceptron MúltiCamadas para vento zonal
PMC_v	–	Perceptron MúltiCamadas para vento meridional
PMC_T	–	Perceptron MúltiCamadas para temperatura
PMC_q	–	Perceptron MúltiCamadas para umidade
PMC_{ps}	–	Perceptron MúltiCamadas para pressão à superfície
RNA	–	Redes Neurais Artificiais
FDP	–	Função de Distribuição de Probabilidade
EnKF	–	<i>Ensemble Kalman Filter</i>
ETKF	–	<i>Ensemble Transform Kalman Filter</i>
EAKF	–	<i>Ensemble Adjustment Kalman Filter</i>
EnSRF	–	<i>Ensemble Square-Root Filter</i>
LEKF	–	<i>Local Ensemble Kalman Filter</i>
LETKF	–	<i>Local Ensemble Transform Kalman Filter</i>
SPEEDY	–	<i>Simplified Parameterizations, primitive-Equation DYNamics</i>
PNT	–	Previsão Numérica do Tempo
CPL	–	camada limite planetária
SCM	–	<i>Sucessive Correction Method</i>
OI	–	<i>Optimal Interpolation</i>
BLUE	–	<i>Best Linear Umbiased Estimator</i>
GrADS	–	<i>Grid Analysis and Display System</i>

LISTA DE SÍMBOLOS

a	– constante responsável pela inclinação da tangente hiperbólica
A	– matriz de covariância do erro de modelagem, conjunto de dados
B	– matriz de covariância do erro de modelagem
∂	– operador diferencial parcial
F	– matriz que representa o sistema dinâmico
P_a	– matriz de covariância do erro da análise
P_f	– matriz de covariância do erro de modelo de previsão
P_b	– matriz de covariância do erro de modelo de referência
R	– matriz de covariância do erro de observação
f	– parâmetro de coriolis
g	– força de gravidade
j	– índice referente a camada oculta das RNA
L	– índice referente a camada de saída das RNA
y	– vetor obtido na saída da RNA
net_i	– sinal de saída do neurônio
u	– vento zonal
v	– vento meridional
θ	– vetor de pesos da RNA
F	– Matriz que representa o sistema dinâmico
W ou K	– Matriz ganho de "Kalman"
H	– Matriz que representa o sistema de observação
I	– Matriz Identidade
J(.)	– Função Custo
j	– Índice referente a camada oculta da Rede neural artificial
k	– k-ésimo instante de tempo discreto
y	– Vetor obtido na saída da Rede neural artificial
P^a	– Covariância do erro de estimativa
P^f	– Covariância do erro de previsão
P^b	– Covariância do erro do modelo de referência
Q	– Covariância do erro de modelagem
r	– quadrante de distância entre os pontos de grade
R	– Matriz de covariância dos erros de observação
u	– Velocidade do vento zonal
v	– Velocidade do vento meridional
x^a, w^a	– Vetor de estado de análise
x^b	– Vetor de estado do modelo de referência
x^t	– Vetor de estado do modelo verdade (sem assimilação)
x^f, w^p	– Vetor de estado previsto pelo modelo numérico

$\bar{\mathbf{X}}$	– média do conjunto de previsões
y_i	– Sinal na saída do neurônio
y^o, w^o	– vetor de observações
δ	– Gradiente local
w	– Vetor de pesos
ϕ	– geopotencial
$\varphi(\cdot)$	– Função de ativação
$(\hat{\cdot})$	– Refere-se a estimativa

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 MODELAGEM ATMOSFÉRICA	13
2.1 Previsão Numérica do Tempo	13
2.1.1 Evolução através dos Tempos	15
2.2 Dinâmica da Atmosfera	19
2.2.1 Formulação Físico-matemática	21
2.2.1.1 Equação do Momentum	22
2.2.1.2 Equação de Continuidade	25
2.2.1.3 Equação de Estado dos gases ideais	26
2.2.1.4 Equação da Energia Termodinâmica	26
2.2.1.5 Equação da conservação da razão de mistura do vapor d'água	27
2.2.2 Equações do Movimento em coordenadas Esféricas	27
2.3 Modelos Numéricos	30
2.3.1 Método espectral	32
2.3.1.1 Transformação de Grade para Espectral	36
2.3.1.2 Transformação de Espectral para Grade	36
2.4 Modelo SPEEDY	36
2.4.1 Revisão das parametrizações físicas	40
2.4.2 Condições de contorno	42
3 ASSIMILAÇÃO DE DADOS	45
3.1 Conceitos em Assimilação de Dados	45
3.1.1 Breve história da Assimilação de Dados Atmosféricos	48
3.1.2 Formulação e Notação	53
3.1.2.1 Matrizes de Erros	55
3.2 Abordagem Bayesiana da Análise	57
3.3 Filtros de Kalman	63
3.3.1 Filtro de Kalman Estendido	63
3.3.2 Filtros de Kalman baseados em conjunto	66
3.3.3 Localização da Covariância em Filtro de Kalman baseados em conjunto	75

3.3.3.1	Local Ensemble Transform Kalman Filter	78
4	REDES NEURAIS ARTIFICIAIS	83
4.1	Conceitos	83
4.1.1	Neurônio Biológico e Neurônio Artificial	84
4.1.2	Funções de ativação	87
4.2	Características Redes Neurais Artificiais	89
4.2.1	Redes Perceptron Multicamadas (PMC)	89
4.3	Aprendizagem em Redes Neurais	92
4.3.1	Algoritmo de Retropropagação do Erro	95
4.4	Aplicabilidade das Redes Neurais Artificiais	100
5	ASSIMILAÇÃO DE DADOS UTILIZANDO PERCEPTRON DE MULTIPLAS CAMADAS	105
5.1	Metodologia	105
5.1.1	Implementação do Modelo SPEEDY	106
5.1.2	Implementação do LETKF	108
5.1.3	Implementação da Rede Neural Artificial	110
5.1.3.1	Treinamento das PMCs	116
5.1.3.2	Ativação/Generalização das PMCs	118
5.2	Resultados da Assimilação de Dados com Redes Neurais Artificiais	119
5.2.1	Análise de Complexidade	144
5.2.2	Avaliação de Tempo	147
6	CONCLUSÕES E TRABALHOS FUTUROS	151
	REFERÊNCIAS BIBLIOGRÁFICAS	157
	GLOSSÁRIO	171

1 INTRODUÇÃO

Na ciência clássica, que nasceu com Descartes e Galileu e se manteve até Einstein, tudo era estruturado pelo determinismo. Na ciência contemporânea, as noções de acaso, incerteza, ruído, caos, desordem, catástrofe, infiltraram-se na representação do mundo. O não determinismo e a contingência estão sempre presentes, ou seja, a ciência atual é uma obra em constante convulsões, trabalhada por novos conceitos que contrastam com a ordem e a simplicidade da ciência clássica. As incertezas estão cada vez mais complexas e aparecem com importância frequente na sociologia, na física quântica, na economia, nas ciências da Terra, na climatologia, na biologia, na meteorologia, etc. Nestes sistemas quanto mais cresce a complexidade mais cresce também a fragilidade. A idéia de vulnerabilidade é introduzida na visão científica do mundo (LACROIX, 1999).

A concepção histórica determinista, fortalecida pela teoria de Newton, possibilitou representar matematicamente a evolução temporal de inúmeros sistemas e prever estados futuros. O determinismo não ficou restrito à física, abrangia todas as ciências naturais até a segunda metade do século XIX. Na citação de Laplace: *"Devemos considerar o estado presente do universo como efeito de seus estados passados e como causa dos que vão seguir. Suponha uma inteligência que pudesse conhecer **todas** as forças pelas quais a natureza é animada e o estado em um instante de todos os objetos, ela englobaria na mesma fórmula os movimentos dos maiores corpos do universo e também dos menores átomos: nada lhe seria incerto e o futuro, assim como o passado, estaria presente ante os seus olhos"* (PROBABILITY, 1990, p. 326), fica claro que se fosse possível observar todas as variáveis em todos os pontos do sistema, poder-se-ia prever o estado futuro deste sistema. Normalmente não se consegue controlar, prever ou entender todas as variáveis, e como elas se interagem. Laplace apresentou uma importante característica das previsões científicas: elas dependem do conhecimento das leis naturais e do conhecimento das condições do sistema em um dado instante.

Entende-se por sistemas dinâmicos todos os sistemas que evoluem no tempo, qualquer que seja a sua natureza. Os sistemas dinâmicos podem ser: *dissipativos* quando implicam na existência de um atrator e *conservativos* quando não têm atrator. Para descrever alguns sistemas dinâmicos, no início do século XIX, Claude-Louis Navier e George Gabriel Stokes estabeleceram as equações que descreviam a evolução do escoamento de fluidos. No início do século XX a busca de soluções exatas destas

equações foi substituída pelo estudo qualitativo das famílias de soluções. Por exemplo, o matemático francês Henri Poincaré utilizou concepções topológicas para analisar características qualitativas de complexidade dos problemas dinâmicos ligadas ao desenvolvimento de computadores (LES..., 1993).

A partir da metade do século XX surgiu a formulação aproximada e numérica onde, de uma lado, estão os movimentos caóticos sensíveis às condições iniciais, e do outro a dinâmica regular dos sistemas ditos integráveis. Com o avanço da tecnologia e, conseqüentemente, com o advento dos computadores com grande capacidade de processamento, têm-se programas que resolvem numericamente as equações não-lineares, de maneira rápida e aproximada. Em conjunto com técnicas matemáticas para o entendimento da complexidade de fenômenos tais como: dinâmica não-linear, análise harmônica, equações diferenciais, matemática computacional e problemas inversos.

Em 1963, Edward Lorenz simulava sistemas meteorológicos em computador quando notou que em duas simulações simultâneas com condições iniciais pouco diferentes, tiveram resultados divergentes no final de certo tempo. Após investigação, Lorenz em suas conclusões, recorre a uma metáfora que se tornou célebre: "o bater de asas de uma borboleta no Brasil, poderia desencadear um tornado no Texas". Isto significa que uma ligeira variação nas condições iniciais pode provocar uma gigantesca disparidade. O artigo clássico de Lorenz (1963) foi um dos primeiros trabalhos a descrever movimentos caóticos sobre um atrator de baixa dimensão e mostra que tais movimentos são caracterizados por uma instabilidade intrínseca devido à sensibilidade do sistema às variações de suas condições iniciais. Sob o nome de "efeito borboleta", o fenômeno caótico tem uma concepção determinista, porém os acontecimentos futuros não são pré-estabelecidos. O efeito borboleta demonstra a impossibilidade de uma previsão perfeita e que sistemas instáveis tem um limite finito de previsibilidade (KALNAY, 2003). Sistemas como o apresentado por Lorenz (1963) não se enquadram como aleatórios ou imprevisíveis, são denominados *caóticos determinísticos* ou simplesmente *caóticos*, apresentam um comportamento aperiódico e a dinâmica é governada por equações diferenciais determinísticas (SILVEIRA, 2008).

Sistemas dinâmicos apresentam processos que permitem compreender e prever a sua evolução. Sistemas imprevisíveis, segundo Kleinke (2007) são caracterizados por flutuação irregulares no tempo, eram conhecidos como aleatórios ou ruidosos. Os sistemas caóticos parecem ser aleatórios, mas são determinísticos, ou seja, são previsíveis. O caos determinista ocorre em sistemas regidos por leis não-lineares. Um

dos aspectos fundamentais do caos determinista é a forte dependência das condições iniciais sobre estados futuros. Os sistemas caóticos são *previsíveis por um tempo curto*. Uma previsão do tempo é válida por alguns dias, a cada dia a mais acrescentado à previsão, maior será a margem de erro. O *tempo* (no sentido meteorológico) não é aleatório, é definido por diversas variáveis pré-definidas e algumas **não** são entendidas e isso o torna *imprevisível* por longo tempo (ALLIGOOD et al., 1997).

A pesquisa científica da atmosfera e as aplicações que dela decorrem, definem o universo e a abrangência da meteorologia. Um dos principais objetivos operacionais da meteorologia é a *previsão do tempo* (um estudo baseado nas condições meteorológicas de um dado momento que busca antecipar com alguma certeza as condições meteorológicas futuras). A previsão dos fenômenos atmosféricos ocorre em um período futuro de até 15 dias. Além da previsão do tempo há a determinação da *tendência da flutuação climática*, em geral referida simplesmente como tendência climática ou previsão climática. Nesse caso, a tendência procura estabelecer as condições da flutuação climática do próximo ano ou da próxima estação. Assim, a previsão do tempo e a previsão climática são definidas por diferentes escalas temporais e espaciais. Os sistemas dinâmicos que regem estas previsões são "caóticos".

Os sistemas de Previsão de Tempo estão constantemente em melhoria devido à importância econômica que abrange áreas como a agricultura, defesa civil e militar, construção civil, transportes, aéreos e marítimos. O *tempo* pode ter impacto na saúde, na mortalidade, no conforto térmico (animal e humano), produção de leite, produção de ovos, pesca, pecuária, indústria têxtil, transportes, esportes, turismo, lazer, mercado financeiro, bolsas, seguros, empréstimos, meio ambiente, poluição, concentração e dispersão de poluentes. O *tempo* tem impactos na geração de energia e no consumo de energia e também tem impacto no consumo, por exemplo de bebidas, sorvetes, remédios, chocolate, roupas, aquecedores, refrigeradores.

Os prognósticos ou previsões dos fenômenos do tempo local, principalmente daqueles fenômenos associados ao tempo severo, como tempestades, ventanias, rajadas, pancada de chuva, granizo, e são muito importantes para uma vasta gama de atividades humanas e para o entendimento das transformações rápidas do ambiente. Por exemplo, nas grandes cidades os fenômenos meteorológicos mais críticos acabam por definir as condições de salubridade e a qualidade ambiental que está sujeita a população. Entre esses fenômenos listam-se as inundações, as estiagens e a disponibilidade de água potável, as condições críticas de temperaturas extremas (ondas

de calor). Em geral, associadas a baixos valores de umidade relativa do ar, prevê-se também os eventos críticos de poluição do ar vinculado à concentrações de poluentes atmosféricos acima de valores aceitáveis à saúde humana, animal e vegetal.

A *Meteorologia* estuda a *atmosfera* em sua interrelação com as outras componentes do planeta: a biosfera, litosfera, criosfera e hidrosfera. A atmosfera é um dos componentes do chamado Sistema Ambiental do Planeta, do qual também participam o oceano e a superfície planetária em geral (solos, rochas etc), onde vive o conjunto dos seres vivos. A previsão de tempo é baseada no estudo e na modelagem da circulação geral da atmosfera.

A circulação atmosférica é o movimento do ar em grande escala, e os meios (juntamente com a circulação do oceano) pelos quais a energia térmica é distribuída sobre a superfície da Terra, representa o escoamento médio do ar ao redor do globo. A circulação geral é função do transporte de calor em direção aos pólos. Os padrões de circulação gerados na atmosfera redistribuem calor, umidade e quantidade de movimento pelo globo de forma não homogênea, podendo aumentar ou diminuir diferenças regionais. As circulações atmosféricas são movimentos do escoamento da atmosfera, onde a conservação de massa e as forças de pressão da superfície da Terra são contraídos em um referencial em rotação (BENNETT, 1992).

Em 1920, Lewis Fry Richardson foi o primeiro a simular movimentos atmosféricos por meio de uma formulação algébrica das equações de Navier-Stokes para auxiliar na previsão do tempo. A metodologia proposta estava correta em linhas gerais, mas a previsão falhou, pois as aproximações consideradas por Richardson não obedeciam certas restrições matemáticas e físicas que só foram formuladas mais de uma década depois. Além disso, a quantidade de cálculos envolvidas tornaram o método impraticável para os anos 1920.

Somente em 1950, quando surgiram os computadores eletrônicos digitais e com o início da análise numérica a "Previsão Numérica do Tempo" (PNT) tornou-se possível por uma questão prática. Na década de 1960, com o aumento da potência dos computadores tornou-se possível executar um modelo de circulação geral da atmosfera (MCGA) global, o que permitiu aos cientistas simular o clima, ou o estado médio da atmosfera por longos períodos. Nos anos 1970, os modelos de circulação geral tinham se tornado instrumentos centrais da ciência do clima e da previsão de tempo tanto em aplicações operacionais como na pesquisa para sua melhoria, desde então,

vêm se aperfeiçoando.

A climatologia - anteriormente centrada nas necessidades locais e regionais de coleta de dados e análise estatística - foi transformada em uma ciência global. Ao mesmo tempo, os cientistas do clima tornaram-se preocupados com os possíveis efeitos do acúmulo de dióxido de carbono na atmosfera. Passaram a utilizar simulações dos MCGAs para analisar os efeitos das alterações climáticas, as chamadas “mudanças globais”. Ao mesmo tempo que os modeladores de oceano começaram a construir simulações similares para a circulação oceânica (MOCGs). Uma vez que os oceanos são importantes componentes do sistema atmosférico global, em meados da década de 1980, os modelos “acoplados”, MOCG com MCGA, tinham estabelecido um novo padrão de modelagem da atmosfera.

Os MCGAs têm desempenhado papel importante no avanço da ciência atmosférica, os modelos atuais vêm melhorando cada vez mais, possuem resolução alta e são capazes de prever fenômenos locais, em prazos maiores, devido aos “supercomputadores” e às técnicas computacionais de alto desempenho. Os modelos são essenciais para a PNT e são resolvidos como problema de valor inicial, ou seja, exibem dependência sensível às condições iniciais. Modelos de PNT são sistemas dinâmicos caóticos determinísticos e a estimativa do estado atual da atmosfera (a condição inicial) do MCGA deve ser a mais precisa possível para aprimorar a qualidade da previsão do tempo.

A estimativa da condição inicial de um MCGA é um processo de fusão de observações meteorológicas e uma previsão de curto prazo, chamada **assimilação de dados atmosférica**. O resultado da assimilação de dados é a **análise**, que deveria ser uma imagem precisa do verdadeiro estado da atmosfera num determinado momento. A análise geralmente é representada em pontos de grade, respeitando a dinâmica do modelo a qual se destina.

A assimilação de dados é um problema inverso, onde um número finito de observações é usado para estimar um campo contínuo (uma função), que tem um número infinito de pontos de dados (isto do ponto de vista teórico, pois na prática os modelos são sempre descritos por equações discretizadas). Em um modelo direto são conhecidas as propriedades do sistema físico e as equações governantes que produzem a previsão (os *efeitos*). Em um modelo inverso, por contraste, a observação ruidosa (medidas) e as equações governantes são usadas para estimar propriedades do sistema físico, ou

seja, obtendo a causa(medida) pode-se estimar o efeito (sistema físico). No caso da assimilação de dados, o modelo direto calcula o estado da atmosfera (a previsão), o dado experimental é a observação do estado da atmosfera e a quantidade a estimar é, de novo, o estado da atmosfera (a análise) (CAMPOS VELHO et al., 2007).

Existem técnicas explícitas de inversão e técnicas implícitas. Quase sempre os dados de observação são imprecisos (ruidosos), incompletos e irregulares e a propriedade a ser estimada não pode ser medida diretamente. A solução computacional do modelo direto requer uma formulação discreta. A metodologia clássica de assimilação de dados parte de uma média entre a previsão e uma ponderação de um incremento de análise. As diferentes metodologias irão variar em como calcular o incremento de análise e sua ponderação, exemplos são método variacional e filtro de Kalman.

Os sistemas ou técnicas abordadas para a assimilação de dados atmosféricos são inerentemente complexos. Muitas técnicas desenvolvidas para meteorologia e oceanografia (Figura 1.1), diferem no custo numérico, na capacidade de encontrar a solução ótima, e na conveniência para abordagem de assimilação em tempo real. A

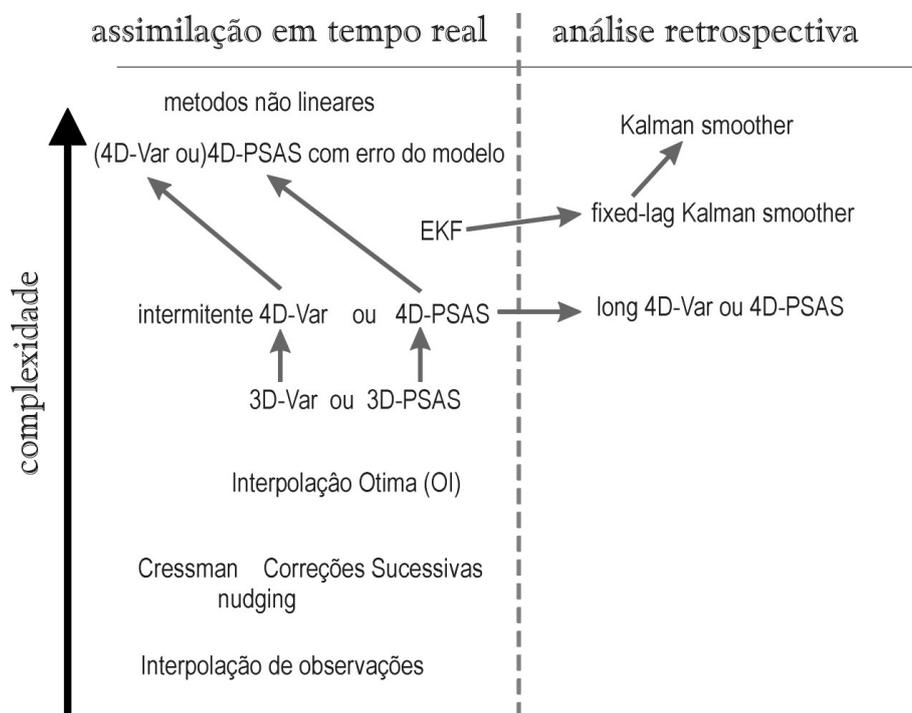


Figura 1.1 - História resumida dos principais algoritmos de assimilação de dados utilizados em meteorologia e oceanografia, classificados de acordo com sua complexidade(e custos) de execução e sua aplicabilidade para os problemas em tempo real.

Fonte: Adaptado de (BOUTTIER; COURTIER, 1999)

motivação deste trabalho é a pesquisa na procura do melhor desempenho em aplicações de sistemas de previsão de tempo, no que se refere à obtenção da análise. A Assimilação de dados pode ser descrita como um processo de dois passos:

$$\text{Passo de previsão: } w_n^p = F[w_{n-1}^p] \quad (1.1)$$

$$\text{passo de análise: } w_n^a = w_n^p + d_n \quad (1.2)$$

onde w_n representa o vetor de estado do modelo no passo de tempo t_n , $F[\cdot]$ é o modelo matemático de previsão, os super-escritos p e a denotam respectivamente os valores preditos e de análise, finalmente d_n é o incremento da análise. O vetor incremento de análise é calculado como sendo um produto entre uma matriz de ponderação e uma função que mede a discrepância entre a previsão do modelo e as observações:

$$d_n = G_n(w_n^o - w_n^p)$$

sendo G_n a matriz de ponderação ou matriz de ganho. Esta matriz pode ser calculada por um estimador ótimo. O estimador poderá ser mínimos quadrados ou a máxima variância, ou ainda, por uma fórmula envolvendo a matriz de covariância de erro do modelo (DALEY, 1991).

No caso de sistemas dinâmicos *lineares*, a solução ótima do problema de estimação do estado é o clássico filtro de Kalman (FK) (KALMAN, 1960; JAZWINSKI, 1970). Para sistemas não-lineares, o *filtro de Kalman Estendido - FKE* tem sido bem sucedida para sistemas de baixa dimensão. A aplicação direta do FK ou FKE para sistemas de alta dimensão é problemática devido às operações com matrizes muito grandes. Por exemplo, para os modelos operacionais globais da atmosfera terrestre usados em previsão de tempo, necessitaria da inversão de uma matriz de dimensão de milhões ou bilhões de pontos que podem ir além da capacidade dos computadores disponíveis. Com o aumento da resolução dos modelos (mais pontos de grade), a dimensão das matrizes também aumentam e a característica atraente do filtro de Kalman, o cálculo das covariâncias de erros de análise e de previsão, deixa de ser viável. O filtro de Kalman produz estimativas compatíveis com a dinâmica do modelo prescrito, contudo, os cálculos da covariância dos erros são difíceis de implementar em sistemas realistas por causa de: (i) o custo computacional, (ii) a não-linearidade da dinâmica, (iii) hipótese de gaussianidade, e (iv) cálculo do erro de modelagem.

Várias técnicas alternativas foram propostas usando a técnica de Filtro de Kalman,

uma delas é o filtro de Kalman por conjunto denominado (*Ensemble Kalman Filter* - EnKF). O EnKF é um método sequencial de assimilação de dados, proposto por (EVENSEN, 1994). Um conjunto de estados do modelo passa a representar as estatísticas de erro do modelo, através de um conjunto de integrações deste modelo no tempo. O EnKF utiliza um esquema de análise que atua diretamente sobre o conjunto de estados do modelo quando as observações são assimiladas. Em esquemas de assimilação de dados por conjunto, o conjunto de previsões (HUNT et al., 2004) é gerado pelo desenvolvimento de um conjunto de condições iniciais distribuídas de acordo com uma análise prévia. Esta abordagem tem a vantagem de não necessitar de desenvolvimento de modelos adjuntos (ferramentas para modelagem reversa de um sistema físico) e nem da manutenção de um modelo de previsão linearizado (modelo tangente linear).

O EnKF provou ser eficiente com sistemas dinâmicos não-lineares e com modelos de grandes espaços de estado (HOUTEKAMER; MITCHELL, 1998). O método é usado em *aplicações realistas*, com modelos de equações primitivas para oceano e atmosfera. Devido a diferentes tipos de perturbações usadas para representar a matriz de covariância da análise, muitos esquemas se derivaram da abordagem do EnKF: o "*Ensemble Transform Kalman Filter*" (BISHOP et al., 2001), o "*Ensemble Adjustment Kalman Filter*" (ANDERSON, 2001), "*Ensemble Square-Root Filter*" (WHITAKER; HAMILL, 2002) e "*Local Ensemble Kalman Filter (LEKF)*" (OTT et al., 2004), neste último, a localização da covariância é feita em pequenos subgrupos de variáveis definidos em pontos de grade do modelo para regiões locais menores que toda a superfície da Terra. A idéia básica do LEKF é executar a análise em cada ponto de grade simultaneamente, utilizando as variáveis do estado e todas as observações na região local centrada naquele ponto.

Segundo CAMPOS VELHO et al. (2007) um dos grandes desafios no século 20 foi desenvolver métodos para resolver equações diferenciais (ou integro-diferenciais), que represente o modelo matemático de um determinado fenômeno físico. No início do século 21, o grande desafio passa ser resumido na expressão da *ciência dos dados*, em várias áreas (como genômica, geofísica, cosmologia, climatologia). Os métodos de assimilação de dados podem ser entendidos como um método da ciência de dados (DALEY, 1991). A metodologia de assimilação de dados dos centros operacionais de PNT trabalha com matrizes envolvidas nos modelos PNT atuais, da ordem de um milhão de equações (o que equivale a matrizes cheias da ordem de 10^{12} elementos!)

e que para serem processadas necessitam de simplificações. É difícil prever como evoluirá a previsão numérica do tempo, mas em um ponto parece haver consenso: haverá necessidade de se aumentar em uma ou duas vezes a ordem de magnitude do número de equações a serem resolvidas. Neste cenário que se insere a busca de novas metodologias para assimilação de dados. Redes Neurais Artificiais tem sido testadas neste contexto e têm apresentado resultados satisfatórios em aplicações de assimilação de dados.

O método de assimilação de dados baseado em redes neurais artificiais (RNA) pode ser descrito como:

$$w_n^a = F_{RNA}(w_n^p, w_n^o)$$

onde F_{RNA} representa o processo de assimilação de dados atmosféricos por RNA.

Este método de assimilação de dados foi pesquisado por em [Nowosad \(2001\)](#), [Härter \(2004\)](#) em sistemas fortemente não lineares: o sistema de Hénon, o sistema de Lorenz em regime caótico e no simulador meteorológico Dynamo (baseado nas equações de águas rasas unidimensional, uma modelagem muito empregada em códigos de previsão ([HALTINER; WILLIAMS, 1980](#)). Em [Nowosad et al. \(2000\)](#) foram apresentados os primeiros resultados utilizando RNA, como método de assimilação de dados utilizando a rede Perceptron Multicamadas (PMC). Em [Härter \(2004\)](#) foram apresentadas outras arquiteturas de RNA, bem como a PMC com nova estratégia de apresentação dos padrões na entrada da RNA. Os resultados demonstraram que a RNA pode ser uma técnica para assimilação de dados e deveria ser investigada para modelos dinâmicos mais próximos da realidade dos centros de previsão de tempo.

Em [Tang et al. \(2001\)](#) e [Tang e Hsieh \(2001\)](#) a RNA é aplicada como um modelo híbrido dinâmico-neural, assim denominado pelos autores, por utilizarem a rede PMC para substituir uma das equações do sistema de Lorenz, num contexto de assimilação de dados com o método variacional em quatro dimensões (4D-VAR). Em [van den BOOGART e Mynett \(2004\)](#), uma rede neural dinâmica (ou construtiva) é aplicada para assimilação de dados em modelo hidrológico. Nestes trabalhos os autores estimam parâmetros da rede neural, parâmetros do modelo dinâmico e condições iniciais para um modelo híbrido. Modelos híbridos são investigados também em [Liaqat A. Fukuhara e Takeda \(2001\)](#).

[Furtado \(2008\)](#) avaliou o desempenho de uma rede PMC treinada para emular: o Filtro de Kalman, o Filtro de Partículas e o Método Variacional; técnicas aplicadas

ao sistema de Lorenz em regime caótico. Em Cintra et al. (2010) a RNA é utilizada na abordagem de assimilação de dados porém para substituir cálculos do FK, a PMC foi implementada para obtenção do incremento de análise no ponto da observação visando a melhoria do desempenho computacional do FK.

As Redes Neurais Artificiais (RNA) são sistemas paralelos distribuídos compostos por unidades de processamento (neurônios) simples que calculam funções matemáticas (normalmente não lineares). Tais unidades podem ser dispostas em uma ou mais camadas e interligadas por um grande número de conexões. Na maioria dos modelos estas conexões estão associadas a pesos, que armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede. A solução de problemas através das RNA é bastante atrativa, pois apresentam a possibilidade de um desempenho computacional superior em relação a solução de problemas baseados na técnicas convencionais. A capacidade de aprender através de exemplos e de generalizar a informação aprendida são os atrativos principais da solução de problemas através de RNA. Esta característica faz da rede uma ferramenta computacional extremamente poderosa para a solução de problemas complexos.

Os resultados da assimilação de dados atmosféricos com redes neurais artificiais publicados e acima citados, tornam a técnica de RNA competitiva. Todavia, os testes com este método, foram realizados em sistemas simples e de baixa dimensão. Tais modelos estão longe da prática operacional.

O objetivo desta tese é desenvolver uma *metodologia* de assimilação de dados atmosféricos utilizando redes neurais artificiais para modelos numéricos de dinâmica atmosférica para todo o planeta (MCGA), que se aproximam da complexidade dos modelos utilizados em centros operacionais de previsão de tempo e que tenha um melhor desempenho computacional em relação aos métodos utilizados atualmente nestes centros.

Dentro da *nova metodologia* desenvolvida a rede neural utilizada emula um método de assimilação de dados baseado em filtro de Kalman por conjunto, o LETKF: *Local Ensemble Transform Kalman Filter*), método atualmente em testes como alternativa para obtenção do campo de análise em centros operacionais de previsão de tempo no Brasil (CPTEC), no Canadá (MSC) e EnKF na Inglaterra (ECMWF). A rede neural investigada nesta metodologia foi uma PMC, que teve como entrada dados

do modelo SPEEDY (um MCGA com um núcleo dinâmico de equações primitivas espectrais e um simplificado esquema de parametrizações físicas) e dados sintéticos de observação. O modelo SPEEDY é um modelo para investigação científicas e tem a resolução mais baixa que os modelos operacionais, utilizado em pesquisas na Universidade de Maryland (EUA), na Universidade de Buenos Aires (Argentina), na Agência Meteorológica do Japão e outros núcleos de pesquisa. Os dados sintéticos de observações são obtidos a partir da modelo SPEEDY contaminados com ruído gaussiano. As posições das observações coincidem com as coordenadas das estações existentes globalmente da Organização Mundial de Meteorologia. E por último, os dados das análises do sistema de assimilação LETKF para o SPEEDY são utilizados como a análise a ser emulada.

O método baseado em RNA apresentou melhor desempenho computacional para obter a uma análise para o modelo SPEEDY. É sabido que assimilação de dados é a etapa mais demorada no processo de previsão numérica do tempo. Por exemplo, por ocasião da implantação do sistema GPSAS, no CPTEC em 2005 (a implementação do GPSAS no CPTEC foi realizada pela autora desta tese), onde o processamento do sistema de análise era aproximadamente de vinte minutos, aproximadamente o mesmo tempo gasto na execução do modelo global do CPTEC para previsão de 15 dias. A resolução horizontal do modelo do CPTEC em 2005 era de 100 quilômetros, ou seja, isto equivale a cerca de 10^6 pontos de grade e as observações (convencionais e de sensoriamento remoto) da ordem de 10^5 . Tanto o sistema de análise e o modelo eram paralelizados e executavam no computador NEC SX6 (CINTRA, 2004). Atualmente, a resolução do modelo aumentou e o número de observações assimiladas é maior em pelo menos uma ordem de grandeza. Os sistemas tem sido atualizados quanto à paralelização e ao hardware, porém o tempo de execução do sistema de análise para obter a condição inicial do modelo, ainda é maior do que a execução do modelo para obter as previsões.

Nos sistemas operacionais atuais de PNT, o número de graus de liberdade do modelo (em pontos de grade) é da ordem de 10^7 a 10^9 e das observações de satélites da ordem de 10^6 a 10^8 . A fusão destes dados nos processos de assimilação de dados atuais demandam grande tempo de processamento para obter a condição inicial para os modelos de previsão atuais, tendo que selecionar observações para que o processo sejam mais ágeis e em tempo hábil de uma previsão de tempo. A metodologia apresentada nesta tese pode ser a solução.

A metodologia desenvolvida com RNA, bem como os conceitos teóricos e a descrição de cada componente necessário à assimilação de dados atmosféricos, são apresentados nos capítulos seguintes desta tese, e estão organizados da seguinte forma:

- *Capítulo 2*: apresenta revisão teórica de modelos atmosféricos de circulação geral (MCGA) como principal instrumento da previsão numérica do tempo. A seção 2.2 descreve a formulação matemática das equações primitivas que regem os MCGA e um método de discretizá-las. Apresenta ainda a formulação e parametrização do modelo SPEEDY;
- *Capítulo 3*: apresenta descrição detalhada da formulação matemática da assimilação de dados e uma rápida descrição dos métodos convencionais utilizados. Apresenta os algoritmos de assimilação de dados baseados em Filtro de Kalman por conjunto, principalmente o LETKF, método de assimilação que será emulado pela RNA;
- *Capítulo 4*: breve descrição das Redes Neurais Artificiais Perceptron de Múltiplas Camadas, do algoritmo de aprendizagem de retropropagação do erro e especificações resumidas dos trabalhos anteriores de assimilação de dados com RNA.
- *Capítulo 5*: descrição da metodologia de assimilação com RNA, descrição da implementação dos componentes do método, o modelo SPEEDY, as observações e a rede. Apresentação dos resultados do treinamento/generalização da RNA. Uma pequena análise de complexidade dos algoritmos e o desempenho computacional do novo método;
- *Capítulo 6*: com base nas análises realizadas no capítulo 5, são apresentadas as conclusões obtidas, como também sugestões para trabalhos futuros.

2 MODELAGEM ATMOSFÉRICA

A atmosfera e os oceanos constituem um invólucro fluido ao redor do planeta Terra, que está em movimento devido ao aquecimento diferencial do Sol. Os movimentos dos chamados fluidos geofísicos se estendem desde as escalas milimétricas até as planetárias de dimensões comparáveis a do próprio planeta, com escalas temporais cobrindo também uma enorme gama de escalas, desde segundos até décadas. Estes movimentos podem ser estudados com as equações da hidrodinâmica e da termodinâmica, usadas para qualquer fluido viscoso, pois a temperatura, variando com as coordenadas espaciais e com o tempo, é influenciada pelo próprio movimento fluido (MARINGOLO; MOURA, 1998).

A meteorologia é uma das áreas da ciência que estuda os fluidos geofísicos e aplica seus conhecimentos em processos como previsão numérica de tempo e clima, por exemplo. Nestas aplicações das ciências geofísicas, são utilizados modelos de circulação geral da atmosfera, onde as equações governantes da dinâmica geofísica são integradas a partir de condições iniciais determinadas por campos meteorológicos em um determinado tempo.

Neste capítulo, serão apresentadas as características da modelagem da atmosfera, considerando as propriedades matemáticas das equações de movimento aplicadas a fluidos geofísicos e à complexidade de processos energéticos em modelos de circulação geral da atmosfera. Os modelos numéricos de previsão de tempo são descritos para apresentação da configuração do Modelo de circulação geral da atmosfera SPEEDY.

2.1 Previsão Numérica do Tempo

Previsão do tempo é uma aplicação da ciência atual para prever o estado da atmosfera para um tempo futuro em uma dada localização. A história da previsão do tempo vem de milênios, desde então as técnicas usadas têm sido modificadas. Atualmente, as previsões de tempo são feitas pela coleta de todos os dados possíveis sobre o estado atual da atmosfera (particularmente a temperatura, umidade e vento) e por processos de entendimento da atmosfera (modelos meteorológicos) para determinar como a atmosfera evolui. A atmosfera tem natureza caótica e o entendimento dos processos é ainda incompleto.

O sistema de previsão de tempo moderno consiste de cinco componentes (Figura 2.1):

- a) *Coleta de dados*: Observações tradicionais são feitas na superfície tais como pressão da atmosfera, temperatura, velocidade e direção do vento, umidade e precipitação coletados em estações meteorológicas convencionais ou automáticas ou bóias. A Organização Meteorológica Mundial (OMM) age para padronizar a instrumentação, as práticas de observação e o tempo de observar, em todo o mundo. Os relatórios das estações são horários no caso do METAR ou cada seis horas no caso de observações SYNOP.

As informações adicionais de ar superior sobre temperatura, umidade e vento são coletadas de radiossondas (balões do tempo), onde dados são geralmente transmitidos desde a superfície até a alta troposfera (camada da atmosfera onde vivem os seres vivos), informando um determinado perfil da atmosfera. Dados de satélite estão sendo usados cada vez mais para (quase) toda a cobertura global. As imagens visíveis são muito úteis para os previsores que acompanham o desenvolvimento de nuvens. Estas informações podem ser usada por modelos de previsão numérica de tempo (PNT). Os dados de sensores de infravermelhos, contudo podem ser usados como informação de temperatura. Também podem seguir a pista de nuvens individuais para um período de tempo e fornecer a informação da direção de vento. Os satélites de órbita polar fornecem perfis de temperatura e umidade da atmosfera. Comparados com dados similares de radiossondas, os dados de satélite tem a vantagem da cobertura global, mas a desvantagem da precisão e definição não serem tão boas. Os radares meteorológicos fornecem informação de localização e intensidade da precipitação. Radares Doppler podem determinar também direção e velocidade do vento. Detalhes em relação à dados e coleta podem ser encontrados em Cintra (2004);

- b) *Assimilação de dados*: Durante um processo de assimilação de dados, a informação pertinente das observações meteorológicas é usada em conjunto com uma previsão recente de um modelo numérico de previsão para o tempo das observações e produz a *análise* meteorológica. Esta deve ser a melhor estimativa da atmosfera e pode ter a representação em três dimensões da distribuição da temperatura, umidade, pressão e vento na configuração do modelo de PNT. Este item será apresentado com maiores detalhes no capítulo seguinte;
- c) *Modelo numérico de previsão de tempo*: Os modelos numéricos de previsão

de tempo (PNT) são simulações computacionais da atmosfera. Tendo a análise como ponto de partida (condição inicial do estado da atmosfera), o modelo evolui no tempo utilizando o entendimento da física e da dinâmica de fluidos. Os modelos integram equações governantes que indicam mudanças no fluido geofísico. São necessários supercomputadores para resolver problemas de alta resolução. As saídas dos modelos fornecem aos meteorologistas os campos básicos para previsão de tempo;

- d) *Pós-processamento das saídas do modelo*: A saída real dos modelos de PNT é sempre modificada antes de ser apresentada como previsão. Esta pode ser na forma de técnicas estatísticas para remover conhecidos "vieses" do modelo, ou pode ter um ajustamento em consenso com outros modelos de previsão. No passado, o previsor humano era responsável por gerar toda a previsão de tempo a partir das observações. Atualmente, os previsores interpretam os dados pós-processados do modelo com previsões que podem ser entendidas pelo usuário final. Os previsores usam o conhecimento adicional de efeitos locais que podem ser menores que o tamanho da célula resolvida pelo modelo, com isso aumenta a precisão da previsão do tempo;
- e) *Apresentação da previsão ao usuário final*: O estágio final do processo de previsão do tempo é mostrar ao usuário final o conhecimento adquirido da previsão de tempo de forma que a informação seja útil e compreendida. A previsão de tempo deve ser pública e estar presente em diversos meios de comunicação nacional e internacional. São informações úteis para o tráfego aéreo, para defesa civil, companhias de fornecimento de energia elétrica e água, indústrias do setor privado, institutos de pesquisas, agências meteorológicas.

2.1.1 Evolução através dos Tempos

O modo de encarar o tempo, primeiramente, era uma curiosa mistura de senso comum e superstição e incluía milhares de regras, ditados esquisitos e provérbios. O senso comum era baseado nas conexões evidentes entre ventos e nuvens. Durante as grandes navegações, no final do século XV, os marinheiros ampliaram esse senso comum para dar conta dos diferentes sistemas de vento e dos padrões de tempo que encontraram ao redor do mundo. A descoberta do caminho das Índias pelos navegantes portugueses no século XV foi o resultado de várias expedições mal sucedidas



Figura 2.1 - Sistema de Previsão de Tempo.

no Oceano Atlântico, lançando os navios ao desconhecido, a um mar com calmarias e tempestades acompanhadas de ventos e agitação marítima intensa (INNOCENTINI et al., 2000).

Na época do descobrimento do Brasil, a frota de Cabral que viajava rumo à Índia, encontrou em seu caminho uma séria adversidade de tempo: um longo período de calmarias, que custou o retardo de sua viagem. Calmaria devido aos ventos que não ofereciam propulsão necessária para a manutenção do curso pré-estabelecido em sua rota. Esta situação perdurou por muito tempo e causou transtorno para toda esquadra na travessia do oceano Atlântico obrigando a frota desviar-se para mais à direita da rota original. Hoje, depois de 500 anos, constata-se que tais condições meteorológicas contribuíram para a descoberta desta Terra de Santa Cruz.

Se havia uma rota prévia da esquadra de Cabral para chegar às novas terras, ou se foram somente condições meteorológicas que condicionaram o desvio, este é um assunto pertinente aos historiadores. Mas, o fato é que as calmarias é um fenômeno real. Como Cabral, através dos séculos, marinheiros, agricultores e outros tentaram fazer previsões baseadas no conhecimento e crenças de sua época e nas suas ob-

servações pessoais. No entanto, essas previsões eram freqüentemente mal sucedidas. Como não havia comunicações adequadas, os observadores não sabiam o que estava acontecendo além do horizonte e normalmente eram surpreendidos por tempestades que chegavam sem muito aviso.

Com a invenção do telégrafo surgiu a previsão sinótica no século XIX. A previsão sinótica consiste na rápida obtenção e análise de observações do tempo feitas no mesmo horário na maior quantidade de localidades possível. Em 1849, foi estabelecida uma rede meteorológica ligada por telégrafo nos Estados Unidos. Os dados eram coletados por voluntários e era preparado um mapa de sinótica, diariamente, com os dados coletados no mesmo horário em todas as localidades observadas. Em 1857, uma rede meteorológica criada na França recebia dados de toda a Europa. A partir de 1860, a previsão sinótica foi ganhando cada vez mais força, culminando na formação de organizações meteorológicas nacionais em vários países.

Já no século XX, as duas grandes guerras mundiais forçaram os governantes a despende grandes esforços para monitorar e prever o tempo, pois as suas variações podiam ter influência no desenrolar das batalhas. O progresso da Meteorologia foi muito favorecido pela tecnologia desenvolvida durante a guerra. As radiossondas, balões carregando instrumentos meteorológicos e transmitindo, via rádio, os dados das camadas de ar acima do solo, e os radares são resultados desse desenvolvimento tecnológico. Após a Segunda Grande Guerra, surgiram também os primeiros satélites artificiais. Com o uso de satélites, foi possível visualizar as nuvens e as tempestades a partir do espaço.

Atualmente, a Meteorologia é uma ciência muito entrosada com a Física, a Química, a Biologia, a Matemática e com a Computação. A evolução da previsão de tempo ocorreu com o surgimento da previsão numérica, baseada em modelos que representam o movimento e os processos físicos da atmosfera. Resolvendo equações iniciadas com valores do estado corrente da atmosfera, podia-se obter projeções para o futuro.

A história da **PNT** se confunde com a dos modelos numéricos. No início do século XX (1904), um meteorologista norueguês, Vilhelm Bjerknes, argumentou que a previsão do tempo pode ser calculada a partir da base física da atmosfera, e que constitui um problema de valor inicial. Desenvolveu um conjunto de sete “equações primitivas” que descreve o comportamento do calor, o movimento do ar e umidade. A solução do conjunto de equações, em princípio, descreve os grandes movimentos atmosféricos.

Em 1922, o matemático e físico britânico Lewis Fry Richardson publicou um sistema numérico para previsão do tempo. Sua idéia era dividir um território em uma grade, cada uma com seu próprio conjunto de números que descrevessem a pressão de ar e temperatura de uma determinada hora. Richardson resolveu as equações do comportamento da atmosfera, usando um método de diferenças finitas. Ele poderia calcular a velocidade e direção do vento, por exemplo, a partir da diferença de pressão entre duas células adjacentes. Richardson levou seis semanas para fazer os cálculos necessários e produzir uma previsão para 24 horas no futuro, com erros maiores do que as que realmente ocorreram.

O trabalho de Richardson, além de pioneiro, revelou os obstáculos que precisavam ser superados: um enorme quantidade de cálculos tinha que ser feitos rapidamente, os dados que representavam o estado inicial da atmosfera eram inadequados, os modelos eram representações da atmosfera muito rudimentares e os problemas com as técnicas matemáticas podiam resultar em pequenos erros que cresciam durante os cálculos.

Os computadores eletrônicos trouxeram a solução para o problema dos cálculos (desenvolvidas durante a Segunda Guerra Mundial para decifrar códigos inimigos e para calcular as explosões de bomba atômica).

O matemático John Von Neumann, fez paralelos entre as simulações das explosões e a previsão de tempo (ambas eram problemas de dinâmica não-linear de fluidos). Em 1946, logo após seu pioneiro computador ENIAC ficar operacional, Neumann defendeu a utilização de computadores para previsão de tempo. A formação do grupo de meteorologistas liderados por Jule Gregory Charney, foi fundamental para que se produzisse a primeira previsão numérica de tempo bem sucedida no computador ENIAC, localizado no U.S. Army Aberdeen Proving Ground, em Maryland. Esse trabalho, em co-autoria com Ragnar Fjortoft, professor visitante norueguês, e von Neumann revolucionou, em 1950, a história da previsão de tempo e deve ser considerado como um marco, um divisor de águas, da Meteorologia moderna. Realmente, dava-se início à fase mais científica desta área da Física aplicada (MOURA, 1996).

A partir de 1955, as previsões por computadores passaram a ser executadas regularmente nos Estados Unidos. Inicialmente, eram no máximo um pouco melhores que as tradicionais.

Mas o grande desafio era o estilo tradicional da teoria da física: combinar e simplificar as equações até obter fórmulas que dessem resultados significativos com uma quantidade limitada de computação. Um sistema inteligente de cálculos poderia ser tão útil quanto um computador que executasse cinco vezes mais rápido. O desenvolvimento de combinações utilizáveis e aproximações das variáveis meteorológicas levaram inúmeras horas de trabalho e uma rara combinação de criatividade e percepção matemática física, como as contribuições de Smagorinsky e Manabe (MOURA, 1996).

A partir do início dos anos 1960 em diante, o trabalho decisivo de modelagem interagiu com os campos da geofísica, tais como hidrologia (umidade do solo e do escoamento), glaciologia (formação de gelo e fluxo), física meteorológica (formação de nuvens e precipitação, intercâmbio entre ventos e ondas, e assim por diante) (KALNAY, 2003).

Atualmente, os fenômenos de pequena escala local são fornecidos como parâmetros básicos para um MCGA e foram melhorando rapidamente graças ao aparecimento de computadores cada vez mais rápidos, que permitiam o uso de modelos mais complexos, representando cada vez melhor a atmosfera. Paralelamente a essa evolução, houve a melhoria no conhecimento do estado inicial com o aumento progressivo da quantidade e qualidade dos dados. Passados cinquenta anos do primeiro modelo computacional, um grande número de documentos técnicos e relatórios têm se dedicado a PNT, o livro de HALTINER e WILLIAMS (1980) é um exemplo da descrição dos métodos numéricos para a modelagem atmosférica. Em todos estes anos, foi enorme o progresso da PNT, incluindo o sucesso na inicialização do modelo e das previsões do conjunto (KALNAY, 2003).

2.2 Dinâmica da Atmosfera

A atmosfera circula. A circulação é global em extensão, transporta ar seco, energia, vapor d'água, vento e também outras quantidades adicionais. A circulação age contra força de atrito, mas é sustentada pela fonte de energia térmica, que vem do Sol. Na média, a Terra intercepta por volta de 340 Wm^{-2} de energia solar incidente, que é convertida em energia cinética da circulação geral (RANDALL, 2009b).

A circulação geral da atmosfera da Terra e da componente atmosférica do ciclo hidrológico são inter-dependentes e os modelos devem abordar os processos úmidos.

Nos modelos é convencional dividir a atmosfera vertical e meridionalmente, e considerar a maior parte da radiação solar absorvida pela Terra na sua superfície.

Vários processos agem para transferir a energia entre os oceanos e a superfície terrestre para a atmosfera. Por este motivo, a circulação geral começa na parte inferior da atmosfera, onde está a camada de ar que interage com a superfície da Terra, chamada "camada limite planetária" (CLP). O ar na CLP é turbulento; a turbulência produz rápida troca de calor "sensível" (principalmente temperatura), umidade e momentum entre a atmosfera e a superfície. As trocas de umidade, acima da CPL na atmosfera se dão através da evaporação. O calor "latente" associado ao fluxo de umidade da superfície é uma das principais fontes de energia para a circulação geral e a força de atrito da superfície é o principal mecanismo que dissipa a energia cinética da circulação geral.

Acima da CLP vem a troposfera *livre*, que caracteriza-se por uma modesta estabilidade estática, ou seja, a temperatura potencial aumenta a uma taxa moderada. A profundidade da troposfera varia com a latitude e a estação. A troposfera é resfriada radiativamente porque emite radiação infravermelha muito mais rápido do que absorve a radiação solar. Esse resfriamento radiativo líquido é equilibrado, principalmente, pela liberação do calor latente de vapor de água, que forma a partir de nuvens e precipitação. A troposfera é a camada da atmosfera de maior interesse da previsão do tempo (Figura 2.2). O estudo da circulação geral da atmosfera reúne conceitos de todas as áreas da ciência atmosférica. Os fenômenos da circulação geral envolvem a dinâmica de grande escala, a convecção, a turbulência, os processos de nuvem e de transferência radiativa (RANDALL, 2009b).

As ferramentas básicas da PNT são os modelos de circulação geral da atmosfera da Terra. Estes modelos consistem de sistemas de equações diferenciais parciais que descrevem fenômenos físicos envolvidos, a partir de variáveis de estado como vento, pressão, temperatura e umidade. A solução destes sistemas requer grande capacidade computacional e atualmente a utilização de modelos cada vez mais detalhados, com resolução cada vez maior, faz com que estes sistemas sejam tratáveis somente com supercomputadores. Os modelos atmosféricos podem ser divididos de acordo com a área de abrangência, em modelos regionais e globais. Em modelos atmosféricos de circulação geral (MCGA), as equações do fluido geofísico são formuladas utilizando, geralmente, coordenadas esféricas, cuja formulação matemática leva à escolha de métodos de discretização do esquema numérico mais apropriado.

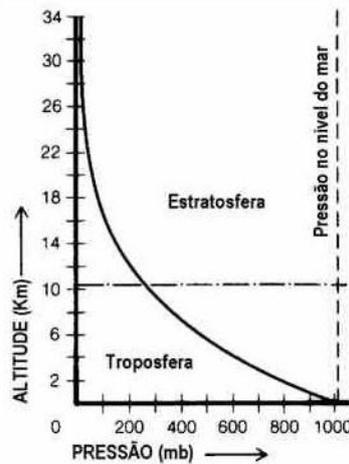


Figura 2.2 - Divisão vertical da atmosfera.

Os componentes da formulação do MCGA são divididos em:

Dinâmica - resolve as chamadas *equações primitivas* da hidrodinâmica da atmosfera. São equações diferenciais parciais não lineares que governam a dinâmica do fluido viscoso, baseados nas leis de conservação de massa, quantidade de movimento, energia e umidade.

Física - resolve processos físicos como a radiação, a formação de nuvens, a precipitação, a umidade da superfície e outros fenômenos de escalas menores e são modelados por parametrizações.

2.2.1 Formulação Físico-matemática

As equações governantes básicas ou "primitivas" de um fluido viscoso que conduz calor, são também conhecidas como equações de *Navier-Stokes*. Estas equações provêm das leis e princípios básicos da Física. Muitas destas equações são *prognósticas*, o que significa que envolvem as derivadas do tempo e são resolvidas por integração no tempo.

Qualquer variável que não é prognóstica é chamada de *diagnóstica*. As variáveis diagnósticas de um modelo podem ser determinadas por variáveis prognósticas e externas que são impostas ao modelo, por exemplo, o raio da Terra. Nesse sentido, as variáveis prognósticas são primárias e as variáveis diagnósticas são secundárias (RANDALL, 2009a). O conjunto de equações primitivas é desenvolvido em MCGAs,

sendo utilizado para previsão atmosférica e também para estudos climáticos de escala global.

O conteúdo aqui apresentado pode ser encontrado em textos de meteorologia tais como: Todling (1999), HALTINER e WILLIAMS (1980), Holton (1992), CAMPOS VELHO (1992), Kalnay (2003) e Satyamurty (2002).

As equações de conservação de massa (ou da continuidade), da conservação da quantidade de movimento e da energia (Primeira lei da termodinâmica), juntamente com a equação de estado ideal dos gases e relações constitutivas, equação da umidade e eventuais definições de variáveis derivadas, no caso de modelos hidrostático, formam um sistema fechado de equações, estabelecido como problema de valor inicial e de condições de contorno. Dessas equações, quatro são prognósticas: duas equações de componentes da velocidade horizontal, uma da continuidade e uma do balanço de energia. As demais equações são diagnósticas: aproximação hidrostática e da componente vertical de velocidade. As variáveis presentes nos modelos atmosféricos são variáveis contínuas.

A representação das equações primitivas dependem também das coordenadas verticais, pode ser em pressão, logaritmo de pressão ou coordenadas *sigma* (σ). Esta última é a relação expressa em coordenadas de pressão p normalizada pela pressão da superfície p_s : $\sigma = p/p_s$ então, a coordenada vertical varia do valor de $\sigma = 1$ na superfície terrestre até o valor $\sigma = 0$ onde a pressão atmosférica desaparece.

No MCGA, admite-se o sistema de coordenadas fixas no centro da Terra em rotação (coordenadas esféricas)(Figura 2.3). Este sistema fixo no espaço pode ser um sistema “inercial” ou sistema absoluto. Quando um movimento aparentemente uniforme no sistema geocêntrico, está sofrendo aceleração, este sistema é “não inercial”. Forças aparentes surgem devido à aceleração das coordenadas. Supõe-se que, em escalas temporal e espacial de interesse, os movimentos estão em balanço hidrostático.

2.2.1.1 Equação do Momentum

A equação vetorial obtida aplicando a segunda Lei de Newton a um elemento fluido pode ser chamada equação de momentum. A equação do momentum é equivalente ao princípio de conservação da quantidade de movimento. A taxa de variação da quantidade de movimento de uma parcela de ar em um sistema inercial é igual a soma de todas as forças atuantes.

As equações matemáticas correspondentes podem ser obtidas usando dois tipos de volume de controle infinitesimal: (1) paralelepípedo de lados $\delta x, \delta y, \delta z$ fixo relativo a sistema de coordenadas, (2) uma parcela infinitesimal "identificada" que desloca com o campo de movimento do fluido. A massa sempre consiste das mesmas partículas do fluido.

Um vetor qualquer \mathbf{A} em um sistema inercial é:

$$\mathbf{A} = A_x i + A_y j + A_z k$$

onde A_x, A_y, A_z são componentes de \mathbf{A} nas direções i, j, k ou nas direções x, y, z , respectivamente. Em um sistema de coordenadas em rotação com velocidade angular $\boldsymbol{\Omega}$, \mathbf{A} é dado por:

$$\mathbf{A} = i' A'_x + j' A'_y + k' A'_z$$

onde i', j', k' são versores nas direções x', y', z' em um dado instante. A'_x, A'_y, A'_z são os componentes de \mathbf{A} nestas direções.

A equação de movimento é dada pela igualdade entre a aceleração medida nas coordenadas não inerciais e o somatório de forças reais por massa unitária atuantes:

$$\frac{d_a \mathbf{U}_a}{dt} = \sum_n \mathbf{F}_n \quad (2.1)$$

o lado direito é a soma de forças reais na parcela de referência, onde \mathbf{U}_a é a velocidade absoluta da parcela de ar nas três dimensões no sistema inercial.

Em um sistema de referência centrado no centro da Terra, a velocidade absoluta \mathbf{U}_a é dado pela soma da velocidade relativa \mathbf{U} mais a velocidade de rotação com a velocidade angular $\boldsymbol{\Omega}$

$$\mathbf{U}_a = \mathbf{U} + \boldsymbol{\Omega} \times r \quad (2.2)$$

onde r é o "vetor posição" da parcela atmosférica. Neste caso particular (para $\mathbf{A} = r$) da fórmula geral que relaciona a derivada de tempo total de qualquer vetor em referencial de rotação $d\mathbf{A}/dt$ à sua derivada total em um referencial inercial $d_a \mathbf{A}/dt$, temos:

$$\frac{d_a \mathbf{A}}{dt} = \frac{d\mathbf{A}}{dt} + \boldsymbol{\Omega} \times \mathbf{A} \quad (2.3)$$

Aplicando nesta formula $\mathbf{A} = \mathbf{U}_a$, tem-se

$$\frac{d_a \mathbf{U}_a}{dt} = \frac{d\mathbf{U}_a}{dt} + \boldsymbol{\Omega} \times \mathbf{U}_a \quad (2.4)$$

Substituindo (2.2) em (2.4) tem-se que as acelerações do sistema inercial e de rotação são relacionados por

$$\frac{d\mathbf{U}_a}{dt} = \frac{d\mathbf{U}}{dt} + 2\boldsymbol{\Omega} \times \mathbf{U} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times r) \quad (2.5)$$

Esta equação indica que o sistema de rotação tem duas forças *aparentes* por unidade de massa: a força de Coriolis (o segundo termo: $2\boldsymbol{\Omega} \times \mathbf{U}$) e a força centrífuga (o terceiro termo: $\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times r)$). O lado esquerdo da equação 2.5 representa as forças reais agindo na parcela de ar: força gradiente de pressão $-\alpha \nabla p$, a força de aceleração gravitacional $\mathbf{g}_e = -\nabla \phi_e$, e a força de fricção \mathbf{F} .

Logo, no sistema de rotação do movimento da Terra, a aceleração aparente é dada por:

$$\frac{d\mathbf{U}}{dt} = -\alpha \nabla p - \nabla \phi_e + \mathbf{F} - 2\boldsymbol{\Omega} \times \mathbf{U} - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times r) \quad (2.6)$$

onde $\alpha = 1/\rho$ é o volume específico (o inverso da massa específica ρ), p é a pressão, ϕ_e é o potencial gravitacional newtoniano da Terra e os dois últimos termos de (2.6), são as forças aparentes: força de Coriolis e força centrífuga respectivamente.

Combinando a força centrífuga com a força da gravidade: $-\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times r) = \boldsymbol{\Omega}^2 \mathbf{l} = \nabla(\Omega^2 l^2/2)$ onde \mathbf{l} é o vetor posição do eixo da rotação da Terra. Definindo como "geopotencial" $\phi = \phi_e - \Omega^2 l^2/2$, então a gravidade aparente é dada por:

$$-\nabla \phi = \mathbf{g} = \mathbf{g}_e + \boldsymbol{\Omega}^2 \mathbf{l} \quad (2.7)$$

Define-se a latitude geográfica φ perpendicular ao geopotencial ϕ . Finalmente, a lei de Newton no sistema de rotação da Terra é escrita como:

$$\frac{d\mathbf{U}}{dt} = -\alpha \nabla p - \nabla \phi + \mathbf{F} - 2\boldsymbol{\Omega} \times \mathbf{U} \quad (2.8)$$

2.2.1.2 Equação de Continuidade

O princípio de conservação de massa fornece a equação da continuidade. Considere um elemento de volume δx , δy , δz , o qual está fixo em relação a um sistema de coordenadas cartesianas. O fluxo de massa em um ponto é dado pelo produto da densidade e a velocidade. Este fluxo através de uma dada área infinitesimal é o produto de densidade do fluido e o componente do movimento perpendicular à área.

A taxa de fluxo do fluido em um volume é dada pelo somatório dos fluxos líquidos nas três direções, x , y e z . Considere a massa M da parcela de ar de densidade ρ

$$M = \rho \delta x \delta y \delta z . \quad (2.9)$$

Se a parcela no tempo conserva sua massa, então, a derivada substancial é igual a zero: $dM/dt = 0$. Se substituirmos a derivada logarítmica da massa

$$\frac{1}{M} \frac{dM}{dt} = 0 \quad (2.10)$$

na equação (2.9) obtém-se a equação da continuidade:

$$\frac{1}{\rho} \frac{d\rho}{dt} + \nabla \cdot \mathbf{U} = 0 \quad (2.11)$$

O significado desta equação é: a razão do aumento de densidade, seguindo a parcela de ar, é igual a convergência (ou divergência negativa) da velocidade ∇ . Desde que, a derivada total de qualquer função $f(x, y, z, t)$ seguindo uma parcela, pode ser expandida como:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial t} = \frac{\partial f}{\partial t} + \mathbf{U} \cdot \nabla f \quad (2.12)$$

A equação (2.12) indica que a derivada total do tempo de uma propriedade é dada pela derivada parcial (no ponto fixo) mais as mudanças devido à advecção. Se expandir $d\rho/dt$ na equação (2.9) usando a equação (2.12) obtém-se uma forma alternativa da equação da continuidade:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{U}) \quad (2.13)$$

2.2.1.3 Equação de Estado dos gases ideais

A atmosfera pode ser assumida como um gás ideal, no sentido termodinâmico, e a pressão p , a temperatura T e a massa específica $1/\rho$ (ou a inversa da densidade) estão relacionadas por

$$p \frac{1}{\rho} = RT \quad (2.14)$$

onde R é a constante dos gases para o ar seco. Para o ar úmido tem-se a pressão parcial do ar úmido, geralmente definida pela temperatura virtual $T_v = 1 + 0.6q$, se a temperatura seca tiver a mesma densidade que o ar úmido na mesma pressão.

2.2.1.4 Equação da Energia Termodinâmica

Esta equação é obtida a partir do princípio de conservação de energia, ou seja, a 1ª lei de termodinâmica. Esta lei diz que a mudança (ou variação) da energia interna de um sistema é igual à diferença entre o calor adicionado ao sistema e o trabalho feito pelo sistema. O calor aplicado a uma parcela, a uma taxa Q por unidade de massa, pode ser usado para aumentar a energia interna C e T produz o sistema:

$$Q = C_v \frac{dT}{dt} + p \frac{d(1/\rho)}{dt} \quad (2.15)$$

Os coeficientes do calor específico a volume constante e a constante de pressão (C_v e C_p , respectivamente) estão relacionadas por $C_p = C_v + R$. Usa-se a equação de estado (2.13) para derivar outra forma da equação da termodinâmica:

$$Q = C_p \frac{dT}{dt} - \frac{1}{\rho} \frac{dp}{dt} \quad (2.16)$$

Define-se que:

- *Processo reversível* é um processo no qual o sistema muda de seu estado e retorna ao estado original sem causar nenhuma mudança na sua vizinhança. Para tais processos ds é um diferencial perfeito e ds/dt é uma derivada total.
- *Processo adiabático* é um processo reversível no qual não ocorre troca de calor entre o sistema e sua vizinhança, ou seja, $Q = 0$.

A taxa de variação de uma parcela de ar é dada pela derivada total $ds/dt = (1/T)\delta Q / \delta T$, ou seja, o calor adiabático dividido pela temperatura absoluta.

Define-se *temperatura potencial* por $\theta = T(p_0/p)^{R/C_p}$, onde p_0 é a pressão de referência. Esta é a temperatura que uma parcela de ar, a uma pressão p e temperatura T , teria se fosse expandida ou comprimida adiabaticamente a uma pressão p . Em meteorologia usa-se $p_0 = 1000$ hPa. Assim, cada e toda parcela de ar, possui uma única temperatura potencial. A temperatura potencial e a entropia específica estão relacionadas por:

$$\frac{ds}{dt} = C_p \frac{1}{\rho} \frac{d\theta}{dt} = \frac{1}{T} \frac{\delta Q}{\delta t} \quad (2.17)$$

Isto mostra que a temperatura potencial é individualmente conservada na ausência de calor adiabático.

2.2.1.5 Equação da conservação da razão de mistura do vapor d'água

Esta equação indica que a quantidade total de vapor d'água em uma parcela de ar é conservada como o movimento ao redor da parcela, exceto quanto existe fontes (evaporação E) e sumidouros (condensação C):

$$\frac{dq}{dt} = E - C \quad (2.18)$$

q é razão de mistura do vapor d'água.

As equações de conservação dos outros constituintes da atmosfera podem ser escritos nos termos de suas fontes e sumidouros correspondentes. Se a equação (2.18) for multiplicada por ρ , expande sua derivada total para $dq/dt = \partial q/\partial t + \mathbf{U} \cdot \nabla q$, e adicionando a equação da continuidade (2.13) multiplicada por q , a equação da conservação da água na alternativa "forma de fluxo", pode ser escrita como:

$$\frac{\partial \rho q}{\partial t} = -\nabla \cdot (\rho \mathbf{U} q) + \rho(E - C) \quad (2.19)$$

O primeiro termo da direita da equação 2.19 é a convergência do fluxo líquido de q .

Assim temos as sete equações com sete incógnitas: $\mathbf{U} = (u, v, w)$, T , p , ρ e q .

2.2.2 Equações do Movimento em coordenadas Esféricas

Como a Terra é esférica, usa-se nos MCGA coordenadas esféricas (λ, ϕ, Z) (Figura 2.3(a)). Próximo à Terra, a gravidade é quase constante, e como a elipticidade da Terra é muito pequena, pode-se aproximar os fatores de escala precisamente por coordenadas esféricas (Philips, 1990a).

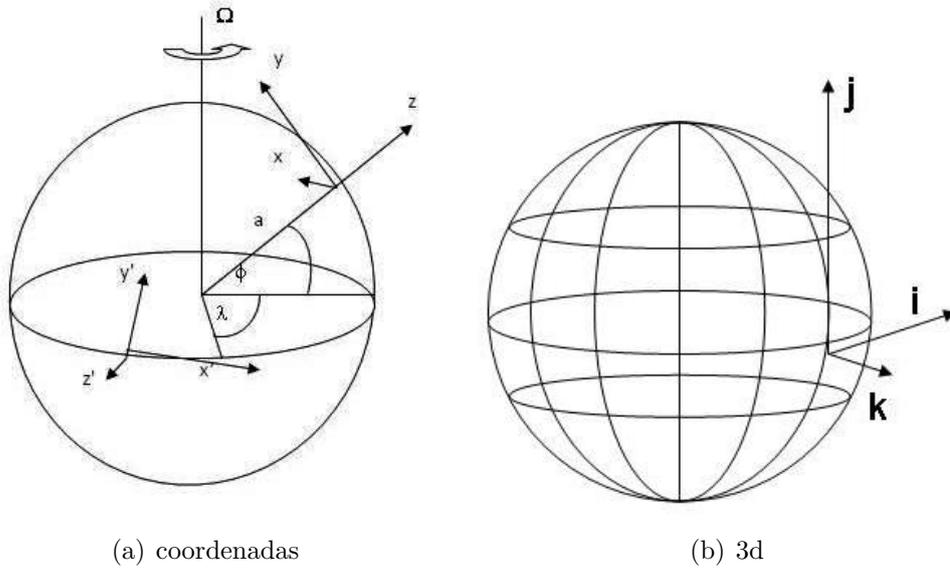


Figura 2.3 - Coordenadas Esféricas.

No que segue \mathbf{i} , \mathbf{j} , \mathbf{k} são vetores unitários (ou versores) direcionados para leste, norte e para cima respectivamente. φ é latitude e λ é longitude da parcela $\mathbf{U} = \mathbf{i}u + \mathbf{j}v + \mathbf{k}w$. Os três componentes de velocidade são escritos como:

$$u = \text{zonal} = r \cos \varphi \frac{d\lambda}{dt} \quad (2.20)$$

$$v = \text{meridional} = r \frac{d\varphi}{dt} \quad (2.21)$$

$$w = \text{vertical} = \frac{dz}{dt} \quad (2.22)$$

Aproximadamente, para movimentos do ar na troposfera $r \approx a$, a é o raio da Terra. Porém, a altura da atmosfera (distância do ponto central da Terra) é bem pequena comparada com o raio da Terra, então se $r = a + z$; $z \ll a$, onde z é a altura acima do nível do mar, pode ser negligenciada ao substituir r pelo raio da Terra, logo, $a = 6371 \text{ Km} = 40.000 \text{ Km}/2\pi$. Assim, temos: $u = a \cos \varphi (d\lambda/dt)$, $v = a (d\varphi/dt)$ e $w = a (dz/dt)$.

\mathbf{i} , \mathbf{j} , \mathbf{k} não são constantes e na expansão do termo de aceleração tem-se que levar em conta a sua variação espacial, então:

$$\frac{d\mathbf{U}}{dt} = \mathbf{i} \frac{du}{dt} + \mathbf{j} \frac{dv}{dt} + \mathbf{k} \frac{dw}{dt} + \mathbf{u} \frac{d\mathbf{i}}{dt} + \mathbf{v} \frac{d\mathbf{j}}{dt} + \mathbf{w} \frac{d\mathbf{k}}{dt} \quad (2.23)$$

Para obter as equações em componentes, é necessário primeiro calcular as taxas de variação dos vetores unitários seguindo a direção do movimento. Primeiro consideremos a derivada total $d\mathbf{i}/dt$. Decompondo-a temos que \mathbf{i} é função de x , isto é, versor apontando para leste e não varia sua orientação nem com latitude nem com altitude: $d\mathbf{i}/dt = u(\partial\mathbf{i}/\partial x)$, o vetor $\partial\mathbf{i}/\partial x$ está direcionado para o eixo de rotação. Então

$$\frac{\partial\mathbf{i}}{\partial x} = \frac{1}{a \cos \varphi} (\mathbf{j} \sin \varphi - \mathbf{k} \cos \varphi) \quad (2.24)$$

Por isso

$$\frac{d\mathbf{i}}{dt} = \frac{u}{a \cos \varphi} (\mathbf{j} \sin \varphi - \mathbf{k} \cos \varphi) \quad (2.25)$$

Considerando $d\mathbf{j}/dt$, tem-se que \mathbf{j} é função somente de x e y . Assim, para movimento na direção leste $\delta\mathbf{j} = \delta x(a \tan \varphi)$. Já que o vetor $\partial\mathbf{j}/\partial x$ está direcionado na direção negativa de x , tem-se então $\partial\mathbf{j}/\partial x = \lim_{\delta x \rightarrow 0} \delta\mathbf{j}/\delta x = -(\tan \varphi/a)\mathbf{i}$.

Para movimento na direção norte, $|\delta\mathbf{j}| = \delta\varphi$. Mas, $\delta\mathbf{y} = a \delta\varphi$ e $\delta\mathbf{j}$ está direcionado verticalmente para baixo, de tal maneira que: $\delta\mathbf{y} = -\mathbf{k}/a$ e $d\mathbf{j}/dt = (u \tan \varphi)\mathbf{i}/a - v \mathbf{k}/a$. Finalmente, por argumentos similares pode-se mostrar que:

$$\frac{d\mathbf{k}}{dt} = \mathbf{i} \frac{u}{a} + \mathbf{j} \frac{v}{a} \quad (2.26)$$

As decomposições das forças nas três direções \mathbf{i} , \mathbf{j} , \mathbf{k} são obtidas da seguinte forma:

- Componentes da força de Coriolis

$$-2\boldsymbol{\Omega} \times \mathbf{U} = -(2\Omega \omega \cos \varphi - 2\Omega v \sin \varphi) \mathbf{i} - 2\Omega v \sin \varphi \mathbf{j} + 2\Omega \omega \cos \varphi \mathbf{k} \quad (2.27)$$

- Componentes da força de gradiente de pressão

$$\nabla p = \mathbf{i} \frac{\partial p}{\partial x} + \mathbf{j} \frac{\partial p}{\partial y} + \mathbf{k} \frac{\partial p}{\partial z} \quad (2.28)$$

- Componentes da gravidade

$$\mathbf{g} = -g\mathbf{k} \quad (2.29)$$

- Componentes da força de atrito

$$\mathbf{F} = \mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z \quad (2.30)$$

Substituindo as expressões 2.24 a 2.30 na equação de movimento e separando os termos, obtém-se a decomposição da aceleração da parcela seguindo o movimento relativo a Terra, em coordenadas esféricas:

$$\frac{du}{dt} = -\frac{\alpha}{a \cos \varphi} \frac{\partial p}{\partial \lambda} + F_\lambda + \left(2\Omega + \frac{u}{a \cos \varphi}\right) v \operatorname{sen} \varphi \quad (2.31)$$

$$\frac{dv}{dt} = -\frac{\alpha}{a} \frac{\partial p}{\partial \varphi} + F_\varphi - \left(2\Omega + \frac{u}{a \cos \varphi}\right) u \operatorname{sen} \varphi \quad (2.32)$$

$$\frac{dw}{dt} = -\alpha \frac{\partial p}{\partial z} - g + F_z \quad (2.33)$$

que possuem o princípio de conservação do momento angular

$$\frac{d}{dt}[(u + \Omega a \cos \varphi)a \cos \varphi] = a \cos \varphi \left(-\frac{\alpha}{a \cos \varphi} \frac{\partial p}{\partial \lambda} + F_\lambda\right) \quad (2.34)$$

Com a "aproximação tradicional" a derivada total do operador em coordenadas esféricas é dada por

$$\frac{d(\cdot)}{dt} = \frac{\partial(\cdot)}{\partial t} + \frac{\partial u}{\partial \lambda} + \frac{\partial(\cdot)}{\partial \varphi} + \frac{\partial(\cdot)}{\partial z} \quad (2.35)$$

e a divergência em três dimensões que aparece na equação da continuidade por

$$\nabla_3 \cdot \mathbf{U} = \frac{1}{a \cos \varphi} \left(\frac{\partial u}{\partial \lambda} + \frac{\partial u \cos \varphi}{\partial \varphi}\right) + \frac{\partial w}{\partial z} \quad (2.36)$$

2.3 Modelos Numéricos

A comunidade de ciência atmosférica inclui um grupo grande de investigadores para realizar medições da atmosfera. Eles fazem o desenvolvimento de instrumentos, desenvolvimento de algoritmos, a coleta de dados, a redução de dados e a análise de dados. Os dados por si só são apenas números. A fim de dar sentido físico aos dados, é necessário colocá-los em um **modelo**. Um modelo pode ser conceitual qualitativo, ou uma teoria analítica, ou ainda, um modelo numérico. Os modelos oferecem uma base para a compreensão dos dados e também para fazer previsões sobre os resultados das medições.

Assim, modeladores em todo o mundo desenvolvem modelos, realizando cálculos e análise dos resultados, em parte, por comparação com as medições. Os modelos por si são "histórias" da atmosfera e devem ser verdadeiras. Em outras palavras, os modelos devem ser compatíveis com todas as medidas pertinentes (RANDALL, 2009a).

Sub-disciplinas da ciência atmosférica, por exemplo: transferência radiativa, química atmosférica e microfísica de nuvens fazem uso de modelos que são essencialmente aplicações diretas dos princípios físicos em fenômenos que ocorrem na atmosfera. Estes modelos são *elementares* pois formam a base conceitual para outros trabalhos de modelagem, que tendem a lidar com os fenômenos micro escala. De modo que a sua aplicação direta a problemas práticos na atmosfera é geralmente negligenciada pela complexidade e tamanho do problema.

Os resultados de um modelo numérico consistem de números e representam "casos" ou "realizações" particulares. Trata-se de exemplos particulares do que o modelo (representando a atmosfera) pode fazer. Por exemplo, "executar" um modelo numérico para criar uma previsão de tempo.

A previsão é composta de um conjunto grande de números e para realizar uma nova previsão, com uma condição inicial diferente, deve-se executar o modelo novamente, gerando um novo conjunto de números. Para se obter todos os acontecimentos da atmosfera, deve-se executar o modelo para muitos casos. Desta forma, os modelos numéricos são diferentes dos modelos analíticos, os quais descrevem todas as possibilidades em uma fórmula. A solução analítica é útil da mesma maneira que a solução numérica. Porém, em alguns casos, a solução de um modelo numérico representa todas as possibilidades (numericamente geradas) na forma de tabela ou gráfico. Os modelos numéricos necessitam de computadores para serem executados.

A modelagem numérica é então uma ferramenta para estudar a atmosfera através de uma abordagem por simulação computacional. Um modelo numérico simula os processos físicos que ocorrem na atmosfera. Existem vários tipos de modelos numéricos, projetados para diversos fins. Uma classe de modelos foi projetado para simular o ambiente real, tanto quanto possível. Exemplos são os modelos de previsão numérica de tempo e modelos de simulação climática. Estes destinam-se a ser substitutos para a atmosfera real e, portanto, as representações de muitos processos físicos.

Modelos numéricos são "discretos", isto significa que um modelo numérico não tem variáveis contínuas. O processo de aproximação de um modelo contínuo por um modelo discreto é chamado de *discretização*, para este processo existem várias abordagens (RANDALL, 2009a).

Os métodos para modelos discretizados em ponto de grade (ponto de junção das

coordenadas) podem ser de vários tipos, como os chamados métodos de diferenças finitas. As derivadas são aproximações em termos de diferenças que envolvem valores vizinhos do ponto de grade. Uma equação de diferenças finitas (ou conjunto de equações) que se aproxima de uma equação diferencial (ou sistema de equações) é chamado esquema de diferenças finitas ou esquema de "ponto-de-grade" e podem ser obtidos por diferentes abordagens. Os campos do modelo (valores do estado do modelo) são definidas em pontos distintos de uma grade, que normalmente evolui no tempo e no espaço.

Outro método importante é o método *espectral*, que implica na expansão dos campos do modelo em termos de somas ponderadas de funções contínuas.

Em um modelo contínuo, muitas funções base são necessárias para representar a distribuição espacial de um campo do modelo. As funções base aparecem dentro de uma soma ponderada por um *coeficiente* que quantifica a contribuição para o domínio em questão. Em um modelo espectral, as funções base são *globais*, o que significa que cada função de base é definida sobre todo o domínio horizontal. Em um modelo computacional, o conjunto infinito de funções base é substituído por um conjunto finito, e a soma infinita é substituída por uma soma finita. Além disso, os coeficientes são definidos em uma grade de tempo discreto. Os modelos espectrais geralmente utilizam métodos de diferenças finitas para representar as estruturas vertical e temporal da solução. As derivadas horizontais são calculadas através da diferenciação das funções base contínuas (RANDALL, 2009a).

Os MCGA geralmente se apresentam no espaço espectral, requerem transformação espectral nas variáveis contínuas que são discretizadas a cada passo de tempo. A discretização pode ser calculada via Transformada Rápida de Fourier e Quadratura Gaussiana. O retorno para o espaço espectral pode ser calculado via Transformada de Legendre e Transformada inversa de Fourier.

2.3.1 Método espectral

A discretização em diferenças finitas pode introduzir erros nas derivadas de espaço. A aproximação de Galerkin pode amenizar este problema (DAHLQUIST; BJORCK, 1969), ao executar a discretização de espaço usando a soma de funções base ou outro tipo de abordagem usando a expansão espectral global, que permite que as derivadas espaciais possam ser calculadas analiticamente, em vez de numericamente.

Até 1970 métodos espectrais eram vistos como interessantes, porém inviáveis para usos em revisão de tempo, devido ao seu custo computacional. Com a introdução do métodos da transformadas por [Gottlieb e Orszag \(1977\)](#), o método se tornou viável computacionalmente e foi aos poucos sendo adotado pela maioria dos centros de previsão de tempo em modelos globais. Métodos espectrais envolvem a busca da solução de uma equação diferencial em séries de funções. Estes métodos são alternativas aos métodos de diferenças finitas e elementos finitos na solução de equações diferenciais parciais.

Em equações diferenciais parciais (EDP) não lineares, as representações em ponto de grade (chamado "espaço físico") ou spectral são úteis; as "derivadas" são computadas eficientemente e precisamente no espaço spectral, enquanto os "produtos não-lineares" são calculados de forma eficiente no espaço físico. O chamado método da transformada ([GOTTSLIEB; ORSZAG, 1977](#)) é utilizado onde a derivada do espaço é calculada no espaço spectral, então \mathbf{U} é transformada de volta para o espaço da grade e o produto $\mathbf{U}_j(\partial\mathbf{U}/\partial x)_j$ é calculado localmente no espaço da grade.

As funções base usadas em métodos espectrais são normalmente soluções de autovalores da equação de Laplace([KALNAY, 2003](#)). Os métodos espectrais podem ser representados por uma função F definida na esfera por:

$$F(\lambda, \varphi) = \sum_{m=-\infty}^{\infty} \sum_{n=-|m|}^{\infty} F_n^m Y_n^m(\lambda, \varphi) \quad (2.37)$$

onde

$$Y_n^m(\lambda, \varphi) = P_n^m(\mu) e^{im\lambda} . \quad (2.38)$$

Os *harmônicos esféricos* Y_n^m são coeficientes de projeção sobre a base de funções ortonormais, de ordem m e grau n , m é o número da onda na horizontal e n é o número "total" das ondas em coordenadas esféricas. Os harmônicos esféricos são produtos de séries de *Fourier* na longitude. Com $\mu = \text{sen}\varphi$, P_n^m são *funções de Legendre associadas* em $\mu = \text{sen}\varphi = \cos\theta$, onde $\theta = \pi - \varphi$ é a co-latidade. O fator $e^{im\lambda}$ descreve a variação das ondas norte-sul e P_n^m descreve a variação sul-norte das ondas dos harmônicos esféricos.

As funções de Legendre de primeira ordem associadas, são polinomiais, satisfazendo:

$$P_n^m(x) = \frac{(2n)!}{2^n n! (n-m)!} (1-x^2)^{\frac{m}{2}} \left[x^{n-m} - \frac{(n-m)(n-m-1)}{2(2n-1)} x^{n-m-2} + \frac{(n-m)(n-m-1)(n-m-2)(n-m-3)}{2 \cdot 4(2n-1)(2n-3)} x^{n-m-4} - \dots \right] \quad (2.39)$$

$n-m$ é o número “nodal meridional”. Os harmônicos esféricos requerem que $n \leq m$.

Os harmônicos esféricos Y_n^m são auto-funções do Laplaciano na esfera:

$$\nabla^2 Y_n^m = \frac{-n(n+1)}{a^2} Y_n^m. \quad (2.40)$$

Aqui, a é o raio da esfera.

A equação de Laplace na esfera pode ser escrita como:

$$\begin{aligned} \nabla^2 Y_n^m &= \frac{1}{a^2} \left[\frac{1}{\cos^2 \varphi} \frac{\partial^2 Y_n^m}{\partial \lambda^2} + \frac{1}{\cos \varphi} \frac{\partial}{\partial \varphi} \left(\cos \varphi \frac{\partial^2 Y_n^m}{\partial \varphi} \right) \right] \\ &= \frac{-n(n+1)}{a^2} Y_n^m \end{aligned} \quad (2.41)$$

onde λ representa a longitude e φ representa a latitude.

Pode-se fazer aproximação de F por uma soma truncada e obter \bar{F} :

$$\bar{F}(\lambda, \varphi) = \sum_{m=-M}^M \sum_{n=-|m|}^{N(m)} F_n^m Y_n^m(\lambda, \varphi). \quad (2.42)$$

Em (2.42) a soma sobre m de $-M$ até M garante que \bar{F} seja real. A escolha de $N(m)$ é um problema de truncamento (Figura 2.4). Duas possibilidades bem conhecidas são o *truncamento triangular*, onde $N = M = \text{constante}$, ou $N - |m| = M - |m|$ e *truncamento romboidal* $N - |m| = M = \text{constante}$. O truncamento triangular representa o espectro da energia cinética “observado”, com um número menor de termos, que o truncamento romboidal (Figura 2.4(b)).

No caso do truncamento triangular (Figura 2.4(a)), o grau mais alto dos harmônicos esféricos Y_n^m é fixo, igual a ordem mais alta destas ondas representado matematica-

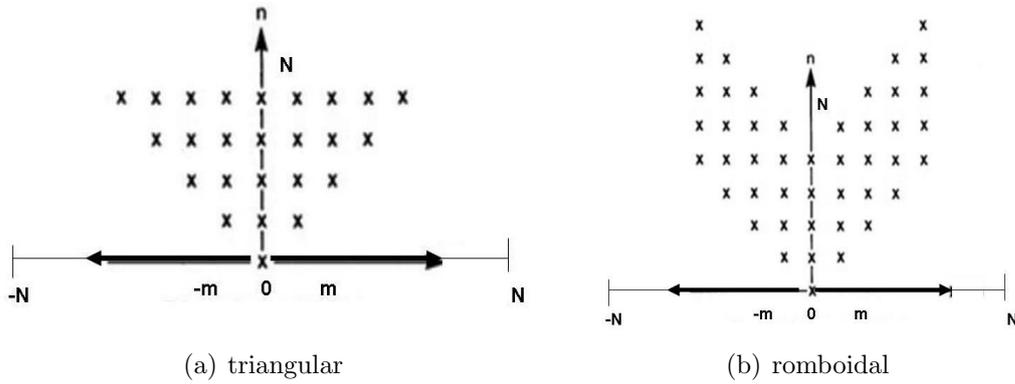


Figura 2.4 - Tuncamento do método espectral

mente por:

$$F(\lambda, \varphi) = \sum_{m=-N}^N \sum_{n=-|m|}^N F_n^m P_n^m(\lambda, \varphi) \quad (2.43)$$

Logo, usando o truncamento triangular temos:

$$\mathbf{U}(\lambda, \varphi, t) = \sum_{n=0}^N \sum_{m=-n}^n U_n^m(t) Y_n^m(\lambda, \varphi) \quad (2.44)$$

a resolução espacial é uniforme por toda esfera. Esta é a maior vantagem em relação ao método de diferenças finitas baseado em uma grade latitude-longitude, onde a convergência dos meridianos nos pólos requerem passos de tempo muito pequenos (KALNAY, 2003, pag.96).

Modelos globais como o modelo SPEEDY usam harmônicos esféricos. Porém, *funções de Hough* também podem ser usadas como funções base para os métodos espectrais, como no estudo da dinâmica de grande escala em resposta à transferência remota de energia associada às fontes de calor tropicais (ANDRADE, 1997). Em Bonatti (1983), é apresentado um modelo espectral inclusive para a dimensão vertical através do uso de solução analítica da equação da estrutura vertical do vento e do geopotencial generalizado. Este tipo de procedimento proporciona a análise das trocas de energia entre os modos verticais e os modos horizontais.

2.3.1.1 Transformação de Grade para Espectral

Para converter valores de ponto de grade para a forma espectral, calcula-se os coeficientes de expansão do campo considerado, isto é, multiplica-se a equação (2.37) por Y_n^m e integra sobre a esfera:

$$F_n^m = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} F(\varphi, \lambda) P_n^m(\sin\varphi) e^{-im\lambda} \cos\varphi \, d\varphi \, d\lambda . \quad (2.45)$$

Para resolver a equação acima, separa-se as partes latitudinal e meridional. A parte latitudinal é obtida utilizando a Transformada de Fourier:

$$P_n^m(\varphi) = \frac{1}{2\pi} \int_0^{2\pi} F(\varphi, \lambda) e^{-im\lambda} \, d\lambda . \quad (2.46)$$

A integração na meridional é obtida a partir da transformada de Legendre, obtendo-se a expressão abaixo para os coeficientes de expansão:

$$F_n^m = \int_{-1}^1 F(x) P_n^m(x) \, dx \quad (2.47)$$

2.3.1.2 Transformação de Espectral para Grade

A conversão dos valores da expansão espectral para valores de grade também é feita em dois passos, na ordem contrária à da seção anterior. Dada uma latitude, são calculados os Polinômios de Legendre e o expansão espectral das variáveis discretizadas na direção zonal, pode ser:

$$F(\varphi_k, \lambda) = \sum_{m=0}^M F^m(x_k) e^{-im\lambda} \quad (2.48)$$

Aplica-se então uma transformada de Fourier discreta às expansões em harmônicos esféricos e obtém-se:

$$F(\varphi_k, \lambda_j) = \sum_{m=0}^M F^m(\varphi_k) e^{\frac{2\pi ijm}{M}} \quad (2.49)$$

2.4 Modelo SPEEDY

A modelagem numérica trouxe maturidade à dinâmica atmosférica. Teorias, simulações e estudos observacionais têm sido desenvolvidos em conjunto, nas últimas

décadas, e assim vai continuar por muito tempo. Estudos observacionais e teóricos orientam o desenvolvimento de modelos de circulação geral da atmosfera e as simulações numéricas sugerem teorias que podem ser usadas para projetar sistemas eficientes.

As comparações diretas com as observações podem ser feitas para a avaliação dos resultados do modelo. O projeto de tais modelos não é um problema puramente matemático. Na prática, os modelos incluem numerosas simplificações e parametrizações, pois têm que ser realistas. Para atender a este requisito, a compreensão física de vários processos físicos e as interações estatísticas da escala de sub-grade, ou ainda movimentos de pequena escala são incluídos através de parâmetros. Experimentos com modelos realistas podem levar à descobertas que não teriam sido possíveis somente com observações. O modelo também pode ser usado como uma ferramenta puramente de simulação. As experiências numéricas de previsibilidade são exemplos.

Modelos numéricos simples são muito úteis para o estudo de fenômenos individuais, na medida em que estes fenômenos possam ser isolados. Simulações com estes modelos podem ser comparados com outros modelos empiricamente a partir de observações, ou com simples modelos teóricos (RANDALL, 2009a).

O modelo SPEEDY (“*Simplified Parametrizations, primitive-Equation Dynamics*”) é um Modelo de Circulação Geral Atmosférica de complexidade intermediária, baseado nas equações primitivas em dinâmica espectral (HELD; SUAREZ, 1994) com um conjunto simplificado de esquemas de parametrizações físicas. Foi desenvolvido no “*Geophysical Fluid Dynamics Laboratory*”, é um modelo hidrostático, de modelo transformada-espectral na forma vorticidade-divergência descrita por Bourke (1974), com tratamento semi-implícito das ondas de gravidade. A resolução do modelo é T30L7, corresponde ao truncamento espectral triangular com 30 número de ondas (T30), com uma grade gaussiana padrão de 96 por 48 pontos e sete níveis verticais (L7). Dos sete níveis verticais, o topo representa a estratosfera, e a base, a camada limite planetária (CLP), com três níveis da troposfera “livre”. Esta é a resolução vertical mínima que permite o desenvolvimento de esquemas de parametrização razoavelmente realistas. Um pacote de parametrizações foi especialmente desenvolvido para trabalhar com modelos com poucos níveis verticais e são baseados nos mesmos princípios físicos adotados em esquemas dos MCGAs mais complexos, tais como: convecção (esquema simplificado do fluxo de massa), condensação em larga-escala, nuvens, radiação de ondas curtas (duas bandas espectrais), radiação de ondas lon-

gas (quatro ondas espectrais), fluxos de superfície de momentum e energia (fórmula aerodinâmica) e difusão vertical. Detalhes do esquema de parametrização física simplificada podem ser encontrados em [Molteni \(2003\)](#).

O modelo SPEEDY é global em extensão com geometria esférica, onde a altitude é suficientemente pequena para que a distância ao centro da Terra seja aproximadamente constante (igual ao seu raio médio), desde que a distância acima da superfície da Terra seja incluída como variável independente. As variáveis prognósticas básicas com os componentes do vento convertidas em vorticidade (Vor) e divergência (Div) (CamposVelho), temperatura absoluta (T) e o logaritmo da pressão à superfície ($\log(p_s)$), o código também calcula a evolução de uma série de outras variáveis do núcleo dinâmico (com fontes e sumidouros especificados pelas parametrizações físicas). No SPEEDY, a única variável adicional utilizado atualmente é a umidade específica (q).

Como foi dito, as equações do modelo são resolvidas por métodos numéricos e como o sistema é altamente não-linear, se discretiza as quatro dimensões. O tempo é dividido em “passos” com espaçamento regular fixo. A integração é feita por um esquema “leapfrog” ([MOLTENI, 2003](#); [ROBERT, 1966](#)) envolvendo as equações da continuidade, de divergência e da termodinâmica, enquanto as equações da vorticidade e da conservação de umidade específica são tratadas explicitamente. As equações do modelo são escritas na forma espectral e as equações do movimento horizontal são transformadas nas equações da vorticidade e da divergência, o que facilita tanto o tratamento espectral quanto a implantação do método semi-implícito de integração no tempo. O modelo utiliza a aproximação hidrostática que representa o balanço entre gradiente de pressão e a força gravitacional na vertical.

A dinâmica do fluido adiabático da atmosfera seca encontrada no modelo pode ser descrita em termos de equações primitivas que governam o fluxo atmosférico, repetidas aqui por conveniência:

$$\frac{d\mathbf{U}}{dt} = -f\mathbf{k} \times \mathbf{U} - \nabla\Phi - RT\nabla \ln p_s + \mathbf{F} \quad (2.50)$$

$$\frac{\partial T}{\partial t} = -\frac{RT}{C_p} \left(\frac{\dot{\sigma}}{\sigma} - \frac{\partial \dot{\sigma}}{\partial \sigma} - \nabla \cdot \mathbf{U} \right) \quad (2.51)$$

$$\frac{d \ln p_s}{dt} = -\nabla \cdot \mathbf{U} - \frac{\partial \dot{\sigma}}{\partial \sigma} \quad (2.52)$$

$$\frac{\partial \Phi}{\partial \sigma} = -\frac{RT}{\sigma} \quad (2.53)$$

O sistema Philips usado em coordenadas verticais como $\sigma = p/p_s$, sendo p_s a pressão na superfície, é empregado no modelo SPEEDY. \mathbf{U} é o vetor vento horizontal com componentes zonal e meridional u , v respectivamente, T é a temperatura absoluta, Φ é altura geopotencial, f é a força de Coriolis, R é a constante dos gases para o ar seco, \mathbf{F} é a força de atrito horizontal, \mathbf{k} é o vetor vertical unitário, ∇ é o operador gradiente horizontal, $\dot{\sigma}$ é a derivada total de σ , C_p é o calor específico à pressão constante para o ar seco e d/dt é a derivada de tempo.

As componentes vertical da vorticidade relativa e a divergência horizontal são (BOURKE, 1974):

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\zeta + f)\mathbf{U} - \mathbf{k} \cdot \nabla \times \left(RT' \nabla lp + \dot{\sigma} \frac{\partial \mathbf{U}}{\partial \sigma} + \mathbf{F} \right) \quad (2.54)$$

$$\frac{\partial D}{\partial t} = \mathbf{k} \cdot \nabla \times (\zeta + f)\mathbf{U} - \nabla \cdot \left(RT' \nabla lp + \dot{\sigma} \frac{\partial \mathbf{U}}{\partial \sigma} + \mathbf{F} \right) - \nabla^2 (\Phi' + RT lp + \frac{1}{2} \mathbf{U} \cdot \mathbf{U}) . \quad (2.55)$$

As equações da continuidade e da energia são analogamente expandidas:

$$\frac{\partial T}{\partial t} = -\nabla \cdot \mathbf{U} T' + T' D + \dot{\sigma} \gamma - \frac{RT}{c_p} \left(D + \frac{\partial \dot{\sigma}}{\partial \sigma} \right) \quad (2.56)$$

$$\frac{\partial q}{\partial t} = -D - \frac{\partial \dot{\sigma}}{\partial \sigma} - \mathbf{U} \cdot \nabla lp \quad (2.57)$$

onde ζ é a componente vertical da vorticidade relativa ($\mathbf{k} \cdot \nabla \times \nabla$), D é a divergência horizontal ($\nabla \times \mathbf{U}$), lp denota ($\ln p_s$), γ é a estabilidade estática, ($\gamma \equiv RT/\sigma C_p - \partial T/\partial \sigma$). O subscrito *zero* denota o valor médio horizontal e o superscrito (*'*) o desvio da média.

Aproximações de diferenças finitas centradas são utilizadas para avaliar as variáveis na vertical e no passo de tempo, calculadas em espaçamento irregular sobre a coordenada de pressão normalizada (coordenada sigma). A coordenada sigma tem a vantagem de permitir uma representação mais simples para a topografia, pois é necessária uma resolução maior próximo à superfície, devido aos processos de camada limite e transferência de fluxos de energia entre solo e ar; também é necessária uma resolução maior na região de interface entre a troposfera e a estratosfera devido a variações verticais rápidas principalmente no campo de temperatura.

A difusão hiper-horizontal tem a forma da equação de Laplace, aplicada nas superfí-

cies sigma. As variações espaciais nas duas dimensões horizontais são representadas por coeficientes de projeção sobre uma base de funções que formam um conjunto ortonormal completo, que são particularmente adequadas para movimentos dos fluidos com geometria esférica (equação 2.41).

O modelo SPEEDY foi desenvolvido para estudos de viabilidade climática para integrações de longo prazo. O código original contém apenas saídas de médias temporais. Miyoshi (2005) desenvolveu processos de entrada e saída de valores nos pontos de grade no passos de tempo intermitentes para capacitar integrações de curto prazo do ciclo análise-previsão.

As variáveis pós-processadas são: componentes de velocidade do vento zonal e meridional (\mathbf{u} , \mathbf{v}), temperatura (\mathbf{T}), umidade específica (\mathbf{q}) e pressão na superfície (\mathbf{ps}). As saídas dos valores em ponto de grade estão no espaço físico de $(96 \times 48 \times 7)$. As entradas são dadas em níveis sigma e as saídas em ambos os níveis: sigma e pressão. As alturas específicas dos níveis verticais são mostradas na tabela 2.1.

Tabela 2.1 - Níveis verticais das saídas do modelo SPEEDY.

Índice dos níveis	Alturas sigma	Alturas de pressão (hPa)
1	0.950	925.
2	0.835	850.
3	0.685	700.
4	0.510	500.
5	0.340	300.
6	0.200	200.
7	0.080	100.

Fonte: (MIYOSHI, 2005)

2.4.1 Revisão das parametrizações físicas

Um conjunto de esquemas de parametrizações físicas (Figura 2.5) foi desenvolvido começando pelos princípios básicos usados nos mais complexos Modelos Atmosféricos de Circulação Geral (MCGA), com um número de suposições de simplificação que são ajustadas a um modelo com resolução vertical pequena. Uma breve descrição dos processos representados pelos esquemas é dada aqui; uma descrição mais detalhada e quantitativa pode ser encontrada em Molteni (2003) e no endereço da internet

(<http://users.ictp.it/~kucharsk/speedy-net.html>):

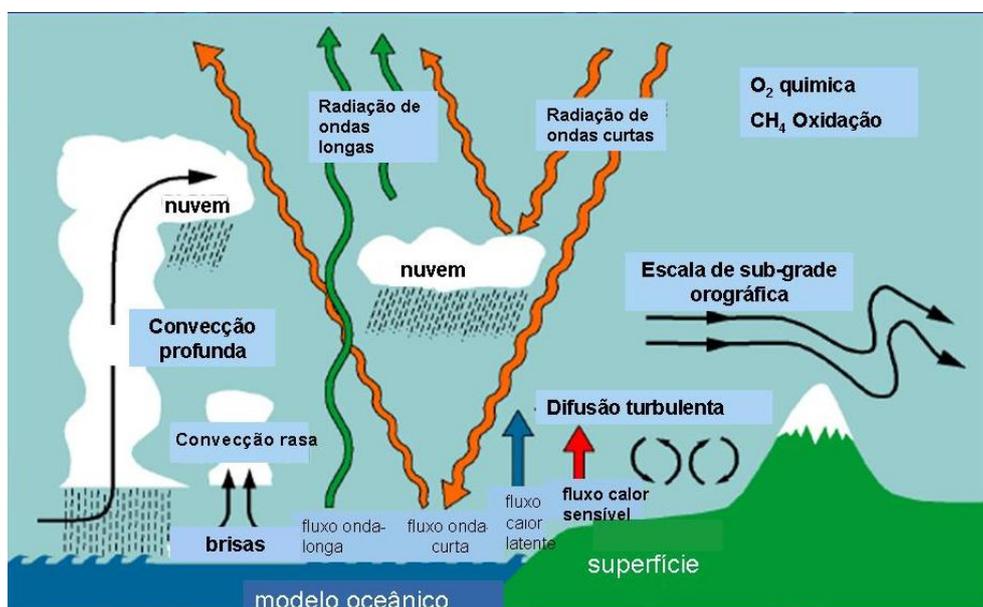


Figura 2.5 - Parametrizações físicas usadas em modelos de Tempo e Clima.

Convecção: Um esquema simplificado massa-fluxo é ativado onde a instabilidade condicional estiver presente (mais precisamente, onde a energia estática da saturação de umidade decresce com a altura entre o mais baixo nível (CPL) e os níveis mais altos da troposfera) e onde a umidade na camada limite planetária exceder o limite prescrito. O fluxo de massa na base da nuvem (no topo da camada limite) é tal que a umidade da camada limite é relaxada através do valor limite na escala de tempo de 6 horas. O decaimento ocorre só ao nível superior de nuvem (determinado pelo critério de instabilidade condicional), enquanto o carregamento ocorre na metade mais baixa da troposfera. O ar nas camadas superiores é assumido ser saturado.

Condensação de grande-escala: Quando a umidade relativa excede o limite que depende da coordenada σ , a umidade específica é relaxada para o valor do limite correspondente na escala de tempo de 4 horas e o conteúdo do calor latente é removido da atmosfera e convertido em energia estática seca;

Nuvens: A cobertura de nuvens e a espessura são definidas por diagnostico a partir dos valores da umidade relativa e absoluta em uma coluna de ar incluindo todos os níveis da troposfera, exceto a CLP;

Radiação de ondas curtas: Os esquemas de radiação usam duas bandas espectrais, uma delas representa a porção próxima do infravermelho. A radiação é refletida no topo da nuvem e na superfície o albedo da nuvem é proporcional a cobertura total das nuvens. A transmissividade de ondas curtas dos níveis do modelo é função do nível de massa, da umidade específica e da cobertura de nuvens.

Radiação de ondas longas: Os esquemas de radiação de ondas longas usam quatro bandas espectrais, uma para cada intervalo espectral atmosférico e outras para regiões espectrais de absorção do vapor d'água e dióxido de carbono. Para cada nível, transmissividade nas quatro bandas são definidas como uma função da camada de massa e da umidade. O efeito das nuvens é representado como um decaimento na transmissividade do intervalo espectral de banda, em função da cobertura de nuvens. A emissão descendente de cada camada é computada como uma função ponderada da temperatura ao nível e da temperatura (interpolada) ao nível abaixo. Para a emissão ascendente, a temperatura é ao contrário.

Fluxos de superfície do momentum e energia: Os fluxos de superfície são definidos pelo volume com fórmulas aerodinâmicas médias (*bulk formulas*) com diferentes coeficientes de troca entre terra e mar. Os coeficientes dos fluxos de calor (sensível e latente) também dependem de um índice de estabilidade simples, enquanto o coeficiente do fluxo de quantidade de movimento sobre a terra é uma função da altura topográfica.

Difusão vertical: Esta é composta de três termos, a redistribuição da energia estática seca e da umidade entre os dois mais baixos níveis do modelo, que simulam a convecção rasa nas regiões de instabilidade condicional; a difusão do vapor d'água em condições estáveis que age na baixa troposfera, dependendo do perfil vertical da umidade relativa; e uma difusão da energia estática seca, no caso do gradiente vertical de temperatura (*lapse rate*) se aproximar (ou exceder) o limite adiabático seco.

2.4.2 Condições de contorno

Como qualquer modelo atmosférico, o modelo SPEEDY requer condições de contorno apropriadas para determinar os fluxos de momentum, calor e umidade na superfície e o fluxo de radiação solar de entrada no topo da atmosfera. Na superfície, inclui a altura da topografia e a máscara oceano-terra, ambas constantes. O modelo requer campos climatológicos das seguintes variáveis:

1. Temperatura da superfície do mar (TSM),

2. Fração de gelo do mar,
3. Temperatura da superfície no topo da camada de solo,
4. Umidade no topo da camada do solo,
5. Profundidade da neve,
6. Albedo de superfície livre (na falta de neve e gelo no mar),
7. Fração da superfície de terra coberta pela vegetação.

Neste trabalho, os dois últimos campos, são usadas valores de médias anuais, enquanto para os outros campos são especificadas médias mensais que são linearmente interpoladas e obtém valores atualizados diariamente.

O albedo de superfície livre é também linearmente combinado com valores atribuídos de gelo no mar e neve. Similarmente, a umidade do solo no topo da camada do solo é linearmente combinada (usando a fração de vegetação); para definir um índice disponível de umidade do solo, é calculada a evaporação sobre a terra. As condições de contorno mais baixas como a TSM, para este trabalho foram obtidas pela reanálise do ECMWF (*European Center Medium Weather Forecast*) durante o período de 1981-90.

3 ASSIMILAÇÃO DE DADOS

Este capítulo apresenta conceitos de assimilação de dados e algumas técnicas de grande relevância nesta área.

Assimilação de dados é uma abordagem para o problema de estimar o estado de um sistema dinâmico por meio de observações atuais, juntamente com um modelo de evolução do sistema de tempo. O processo de assimilação ocorre da seguinte forma: cada vez que novas observações estão disponíveis, estas são usadas em conjunto com a previsão mais recente do estado da atmosfera; estes dados combinados geram uma estimativa da condição inicial do modelo de previsão para a próxima integração temporal, e o procedimento se repete sistematicamente.

3.1 Conceitos em Assimilação de Dados

Como visto no capítulo 2, a previsão numérica de tempo (PNT) é um problema de valor de contorno e de valor inicial. Dada uma estimativa do estado atual da atmosfera (como condição inicial), condições de contorno laterais e de superfície apropriadas, o modelo simula ou prevê a evolução da atmosfera. O sistema que determina as condições iniciais é chamado **assimilação de dados**.

Um problema de condição inicial ou de contorno *bem-posto* deve ter uma única solução que depende continuamente das condições inicial ou de contorno. Claramente, a especificação da própria condição inicial para uma equação diferencial parcial (como as equações primitivas (seção 2.2)), é essencial para se ter um problema bem-posto. Se o número de condições iniciais for correto, mas se estas forem especificadas em tempo ou local impróprio, a solução poderá ser única, mas instável, isto comprova que, pequenos erros na condição inicial podem produzir erros enormes na solução. Este é um problema *mal-posto* (KALNAY, 2003, pag 69).

Quando os parâmetros que descrevem matematicamente um sistema físico não podem ser determinados a partir de observações diretas, mas somente a partir de dados observáveis que a eles se relacionam, diz-se tratar-se de um *problema inverso*. Então, a assimilação de dados pode ser classificada como um problema inverso. Para tal problema considera-se x o vetor que pertence ao espaço dos parâmetros (modelo) e y ao espaço das observações. Algumas dificuldades são típicas de tais problemas: os inevitáveis erros nas observações podem resultar em uma solução que não é estável ou que não tem dependência contínua das observações. Há casos igualmente em que

o problema não tem solução única, quando tem solução.

Pode acontecer uma combinação em que tais propriedades não sejam verificadas, que conforme [Hadamard \(1952\)](#), deixa de ser um problema bem-posto. Para esta classificação, deve-se satisfazer as seguintes condições: (i) existência, (ii) unicidade e (iii) estabilidade da solução.

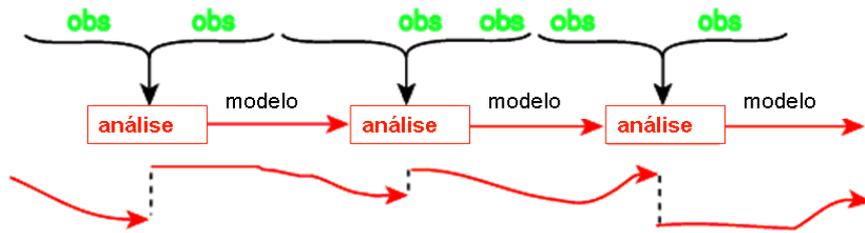
Durante certo tempo admitiu-se que estas condições eram indispensáveis para a solução matemática de um problema físico. Os problemas que não satisfaziam estas condições eram ditos "mal-postos" e considerados irrelevantes para fins práticos. Porém, pode-se transformar um problema mal-posto em um problema "próximo" com soluções suaves e com dependência contínua dos dados, utilizando um método de solução de problemas inversos.

Segundo [van Leeuwen e Evensen \(1996\)](#), os problemas inversos são normalmente formulados, permitindo que um modelo, suas condições inicial e de contorno e um conjunto de medidas pode conter erros. Os quadrados ponderados desses erros são definidos e a solução é minimizar uma estimativa que pode estar próxima das medições e da primeira estimativa inicial e de fronteira, ao mesmo tempo que satisfaz as equações dinâmicas aproximadamente. Por exemplo, pode ser uma solução aproximada que utiliza o método de mínimos quadrados; esta solução tem sido muito aplicada a métodos de assimilação de dados ([HÄRTER, 2004](#)).

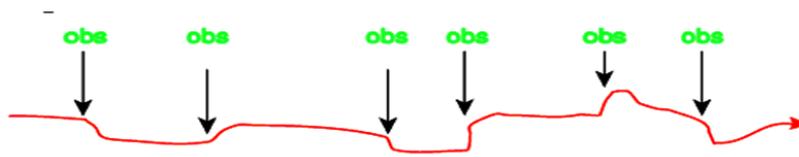
O resultado da *fusão* de dados de observação e dados do modelo matemático (uma previsão) chamado *análise* é o ponto de partida para a integração evolutiva das equações primitivas da dinâmica de fluidos. As equações primitivas são não-lineares e demandam grande capacidade computacional.

O modelo atmosférico é uma representação imprecisa da atmosfera, além disso, tem-se uma compreensão incompleta dos processos, logo as previsões não são exatas. A previsão se afasta da dinâmica real a cada passo de tempo. Deste modo, é necessário inserir informações do sistema atmosférico observado, que irá gerar uma nova análise. Isto caracteriza um ciclo de assimilação de dados e previsão para produzir a análise.

A inserção de dados é um desafio, pois os dados são inexatos e irregularmente distribuídos no espaço e no tempo, e devem ser de alguma forma associados a dados de um modelo matemático imperfeito. Segundo [Daley \(1991\)](#), o objetivo da assimilação de dados atmosféricos é produzir uma representação regular e fisicamente consistente



(a) intermitente



(b) contínua

Figura 3.1 - Tipos de Assimilação de Dados Sequencial

Fonte: Adaptado de (BOUTTIER; COURTIER, 1999)

do estado da atmosfera e é uma tarefa importante na prática operacional da PNT e da Circulação Oceânica.

Assim, a assimilação de dados é uma técnica de análise em que a informação observada é acumulada ao modelo de estado, aproveitando restrições de consistência com as leis de evolução de tempo e propriedades físicas. Existem duas abordagens básicas para assimilação de dados: a assimilação *sequencial*, que considera apenas a observação feita no passado, até o momento da análise, e a *não sequencial* ou assimilação *retrospectiva*, onde a observação do futuro pode ser usado. Outra distinção são os métodos que podem ser *contínuos* ou *intermitentes* no tempo. Em um método intermitente, as observações podem ser processadas em pequenos lotes. Em um método contínuo, os lotes de observação em períodos mais longos são considerados e a correção para o estado analisado é suavizado no tempo (BOUTTIER; COURTIER, 1999). Tipos básicos de assimilação são representadas esquematicamente na Figura (3.1).

Muitas técnicas de assimilação de dados têm sido desenvolvidas para a meteorologia e oceanografia, que diferem em custo numérico, otimização e na sua adequação em tempo real de assimilação de dados. Os detalhes e a descrição de alguns métodos são

explicadas em Daley (1991), Lorenc (1986), Kalnay (2003), Cintra (2004), Furtado (2008), dentre outros.

Uma definição da **melhor análise** é aquela que gera a melhor previsão subsequente. Porém, é necessário considerar o modelo de previsão particular a ser usado, a metodologia de assimilação de dados é independente do modelo, porém alguns ajustes e configurações são próprias do modelo de previsão e devem ser ajustadas na aplicação da metodologia. O modelo geralmente tem uma base discreta que deve ser consistente com os valores observados.

Os modelos globais atuais (alguns operacionais) têm as equações primitivas da ordem de 10^7 a 10^9 graus de liberdade. Em uma janela de tempo de ± 3 horas, o número de observações (convencionais) pode ser até duas ordens de magnitude menor que o número de graus de liberdade do modelos, da ordem de 10^5 a 10^7 . Os dados não são distribuídos de maneira uniforme, assim, se faz necessário, adicionar informações através de um campo de "referência" (também conhecido como **background**) ou suposição inicial ou informação "a priori". Em sistemas operacionais de assimilação de dados é usado como estimativa inicial ou "modelo de referência" uma *previsão de curto prazo*.

3.1.1 Breve história da Assimilação de Dados Atmosféricos

Segundo Barker (2008), as primeiras cartas sinóticas foram desenhadas manualmente, no século XIX, colocando as observações em cartas sinóticas e interpretando uma análise "subjetiva", elaborada por Sir Francis Galton(1822-1911). No século XX, a história da assimilação se confunde com a evolução da previsão numérica de tempo. Ainda com a análise subjetiva, Bjerknes (1911) acrescentou operações matemáticas (como multiplicação, diferenciação, integração em linha, divergência) às cartas. Fjortoft (1952) acrescentou a integração da equação da vorticidade barotrópica e Saucier (1955) o movimento vertical e a tendência de pressão à superfície.

No surgimento da Previsão Numérica do Tempo (PNT), Bjerknes (1904) definiu a previsão de tempo como uma "problema de valor inicial", requerendo: (i) uma estimativa do estado inicial da atmosfera; (ii) condições de contorno, e (iii) as leis da física. Richardson (1922), fez a primeira tentativa de PNT tendo: (i) análise subjetiva de observações; (ii) estado inicial prescrito manualmente e (iii) as equações de movimento integradas numericamente.

Os primeiros experimentos de PNT com Richardson (1922) e Charney (1951) foram baseados em análises elaboradas à mão. Ao mesmo tempo, no Instituto de Princeton para Estudos Avançados, um pequeno grupo foi formado para investigar a estimativa numérica (automática) do estado inicial da atmosfera para PNT. Desta investigação, surgiu a "análise objetiva", não realmente "objetiva" porque as soluções variavam de acordo com o algoritmo escolhido pelo desenvolvedor.

A primeira análise objetiva foi desenvolvida por Panofsky (1949), um algoritmo baseado em interpolação polinomial para ajustar todos os pontos de observação em uma pequena área de domínio de análise (coeficientes polinomiais via ajuste de mínimos quadrados).

Gilchrist e Cressman (1954) desenvolveram o esquema de interpolação polinomial "local" para altura geopotencial, ou seja, consideravam observações dentro de um raio "de influência" do ponto de grade. Bergthorsson e Doos (1955) e Gandim (1963) utilizaram a informação *a priori* ou *referência* proveniente da climatologia ou previsão para calcular o chamado "incremento de observação" (diferenças entre a observação e a referência) e calcularam o "incremento de análise" que são combinações lineares ponderadas do incremento de observação.

Segundo Barker (2008), os primeiros métodos utilizavam erros do modelo calculados de forma empírica: o Método de Correções Sucessivas (SCM) por Bergthorsson e Doos (1955), Cressman (1959), Barnes (1964) e a Relaxação Newtoniana (Nudging) por Hoke e Anthes (1976), Stauffer e Seaman (1990). Estes métodos de assimilação de dados têm forte componente de heurística e os desempenhos na prática produzem uma análise pior do que os métodos mais modernos, além disso: (i) não há como relaxar análise de referência das observações afastadas; (ii) a análise deve respeitar as características da atmosfera, por exemplo, equilíbrio e suavidade; (iii) e há dificuldade para incluir observações "não tradicionais", por exemplo, radiâncias do satélite e do radar.

Os métodos sequenciais de assimilação de dados são feitos em dois passos: (1) assimilação de observações (passo de atualização) e (2) previsão (passo de previsão).

Alguns exemplos são: (i) Assimilação de dados Bayesiana (DALEY, 1991); (ii) Filtros de Kalman - FK (KALMAN, 1960) e (iii) Interpolação Ótima (*Optimal Interpolation - OI*) (LORENC, 1981). Inicialmente a "Interpolação Ótima/Estatística" (O/SI) foi

aplicada por Kolmogorov (1941) em estudos de turbulência; Wiener (1964) mudou o nome para OI; (ELIASSEN, 1954) e (GANDIM, 1963) propuseram um método rigoroso para encontrar a análise "ótima" dada uma estimativa "a priori" do estado da atmosfera.

Inicialmente, a estimativa anterior era a climatologia, mas as previsões de curto prazo deram melhores resultados como estimativa inicial. OI foi o método operacional em muitos centros de PNT entre 1970 e 1990 (LORENC, 1981).

A hipótese fundamental de OI é que para cada variável do modelo, apenas algumas observações são selecionadas para determinação do incremento de análise. Vantagens de OI em relação ao SCM são: poder assimilar as observações não s'mente no espaço do modelo, a simplicidade de implementação dos algoritmos e relativo baixo custo computacional. Porém há desvantagens: a seleção dos dados pode levar a resultados diferentes, se usados em pontos de grade vizinhos, apresentando uma análise não contínua no espaço. Grandes esforços são necessários para estimar covariâncias de erros empíricos do modelo de referência. A análise linear de interpolação estatística, para estes métodos, utiliza a estimação de mínimos quadrados com suposições e simplificações.

O Método dos Mínimos Quadrados (*Least Square Root*) é uma técnica de otimização matemática que procura encontrar o melhor ajustamento para um conjunto de dados. Por exemplo, determinar a e b que minimizam a soma dos quadrados das diferenças (resíduos) entre a curva ajustada e os dados:

$$\sum_n \epsilon_n^2 = \sum_n (a + bx_n - y_n)^2$$

Supondo uma função

$$y = Ax + \epsilon,$$

onde ϵ é um ruído, para achar A que minimiza a média da "soma dos quadrados"

$$\overline{\|\epsilon\|^2} = \overline{|(y - \mathbf{A}x)|^2}.$$

A solução dos mínimos quadrados é

$$\mathbf{A}_{LS} = cov[y, x]cov(x, x)^{-1} .$$

Supõe-se que os dados não sejam igualmente precisos. Se a variância de cada observação for σ_n^2 , então a estimativa apropriada é encontrada pela minimização de:

$$J(a, b) = \sum_n \frac{(a + bx_n - y_n)^2}{\sigma_n^2}$$

chamada de solução de mínimos quadrados ponderada (BARKER, 2008).

Segundo DelSole (2009), Thiele publicou em 1880, uma solução "recursiva" para o problema de estimação de coeficientes de regressão de uma sequencia de dados obtidos no tempo (LAURITZEN, 1981), ou seja, mínimos quadrados recursivos. Kalman encontrou uma solução recursiva para o problema de filtragem e previsão de sistemas dinâmicos lineares com ruído branco .

Com a finalidade de remover a seleção de dados locais no algoritmo OI realizando uma análise global dos campos meteorológicos 3-D, foi introduzido o método variacional tridimensional *3D-Var*, abordagem introduzida por Sasaki (1958), Sasaki (1970). Segundo Lorenc (1986), a essência do algoritmo 3D-Var é reescrever um problema de mínimos quadrados como a minimização de uma função custo.

A técnica tem sido aplicada igualmente a outros problemas. O método variacional de quatro dimensões *4D-Var* é uma generalização do 3D-Var onde as observações são distribuídos no tempo. As equações são as mesmas, os operadores de observação são generalizados para incluir um modelo de previsão que permita uma comparação entre o estado do modelo e as observações no momento exato. Assume-se, como em Filtros de Kalman, que os erros são Gaussianos e com crescimento linear. O método variacional procura um campo de análise x que minimize uma "função custo" (J). Os métodos variacionais (3D/4D) são resolvidos por algoritmos numéricos de minimização que requerem o gradiente da função em direção ao estado x . No método variacional não há seleção de dados, todos os dados são usados simultaneamente. A abordagem quadri-dimensional (4D-Var) implica que é necessário avaliar a derivada da previsão em respeito a condição inicial, integrando o modelo para frente no tempo e então integrando um modelo adjunto no tempo para trás.

Lorenc (1986) mostrou que se:

- (i) a matriz de covariância de erros *a priori* no tempo inicial for correta;
- (ii) se o modelo fosse perfeito e os erros de observações em tempos diferentes

forem não correlacionados;

a análise 4D-Var no tempo final é idêntica ao método de Kalman Filter Estendido. O Centro Europeu de Previsão de Tempo (*European Centre for Medium-Range Weather Forecast - ECMWF*) vem utilizando o método variacional 4D-Var desde novembro de 1997.

Após a abordagem variacional, houve a volta de métodos sequenciais com a assimilação de dados por Filtro de Kalman por conjunto. Evensen (1994) testou um algoritmo de Filtro de Kalman por Conjunto (*Ensemble Kalman Filter - EnKF*) em oceanografia e publicou resultados da aplicação do método de Monte Carlo para aproximar a matriz de covariância de erro. Houtekamer e Mitchell (1998) testaram o EnKF em meteorologia, utilizando um modelo de três níveis, quasi-geostrófico com observações simuladas. O EnKF provê um conjunto de condições iniciais para um conjunto de previsões. Este método atualmente é um dos mais investigados e muitas técnicas derivadas foram implementadas, algumas técnicas são apresentadas posteriormente neste capítulo.

Segundo Ghil (2007) todos os métodos são usados e mais ou menos com sucesso, para fornecer análises de oceano (Bennett (1992); Malanotte-Rizzoli (1996)) e de sistema oceano-atmosfera (JI; LEETMAA, 1997), bem como da atmosfera, em escalas de tempo de horas a meses e anos. Assimilação de dados tornou-se um importante instrumento de investigação para Geociências, estendendo-se à atmosferas planetárias (BANFIELD et al., 1995), clima espacial (ANGELOPOULOS; PANETTA, 1998) e muitos outros.

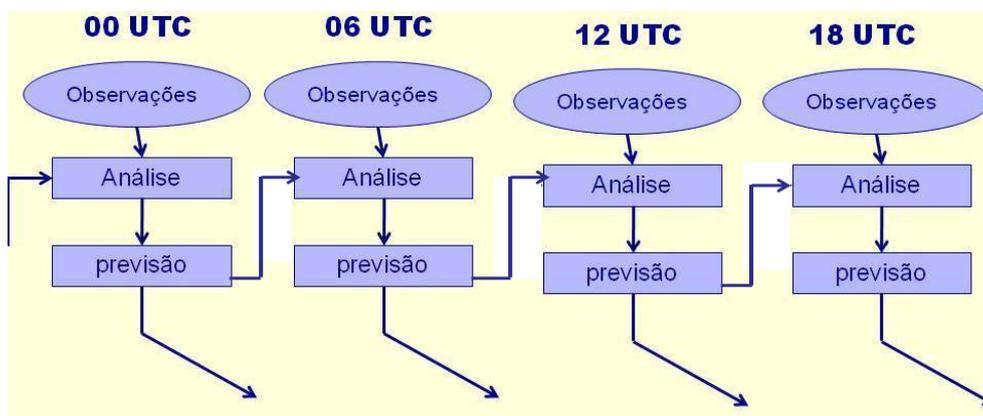


Figura 3.2 - Ciclo típico de Assimilação de Dados.

O ciclo de assimilação de dados é composto de coleta de dados, controle de qualidade das observações, análise objetiva, previsão de curto prazo para prover o modelo de referência para o próximo ciclo. Um ciclo típico, em centros de previsão de tempo, é executado a cada seis horas baseados na hora universal (**UTC**), assim a assimilação é executada quatro vezes ao dia. (ver Figura 3.2).

3.1.2 Formulação e Notação

Neste trabalho utiliza-se a notação proposta por Ide et al. (1997) para sistemas de assimilação de dados. A finalidade da assimilação de dados é reconstruir o mais fielmente possível o estado atmosférico ou fluxo oceânico, utilizando todas as informações adequadas, que consistem em:

observações adequadas, que variam em natureza, resolução e precisão, e são distribuídas mais ou menos regularmente no espaço e no tempo; e

leis físicas que regem a evolução do fluxo atmosférico ou oceânico, disponível na prática sob a forma de uma discretização e necessariamente aproximada, do modelo numérico.

Excelentes textos de introdução à assimilação de dados a partir de diferentes pontos de vista são dadas por [Bennett (1992), Bennett (2002), Bouttier e Courtier (1999), Lorenc (1986), Tarantola e Valette (1987), Tarantola e Valette (1982)]. A formulação discreta é a forma mais comum para o desenvolvimento de algoritmos em modelos numéricos. Diferentes algoritmos de assimilação são derivados de uma fonte comum (a estimação mínimos quadrados), com aproximações particulares a cada algoritmo (OI, 3D-Var, 4D-Var e filtros de Kalman, em particular).

Conforme (Bennett (1992), Bennett (2002)) uma solução inversa generalizada para o problema de assimilação pode derivar muitos dos algoritmos de assimilação como casos especiais.

O problema de assimilação pode ser visto como um problema de valor inicial porém, adicionados de algumas características:

- as equações do modelo para o *modelo de estado* \mathbf{x} são aproximadas com o modelo de erros ϵ_n ;
- a condição inicial \mathbf{x}^b é aproximada com erros ϵ_b (chamados como *erros do modelo de referência*);

- as observações \mathbf{y}_m^o presente com erros de observação ϵ_m^o , de dimensão N_y .

Expressos na forma matemática como:

$$\frac{\partial \mathbf{x}}{\partial t} + \mathbf{M}(\mathbf{x}) = \epsilon_n \quad (3.1)$$

$$\mathbf{x}(0) = x^b(0) + \epsilon_b \quad (3.2)$$

$$\mathbf{y}_m^o = H_m(\mathbf{x}) + \epsilon_m^o \quad m = 1, \dots, N_y \quad (3.3)$$

Na equação (3.3) é introduzido o *operador observação* H_m , que calcula o modelo equivalente a observação \mathbf{y}_m^o . Se \mathbf{y}_m^o é a medida explícita de variáveis de estado (como temperatura), o operador observação necessita ser interpolado para o modelo de estado, que sempre tem dimensão finita, se for na posição da medida observada, ou seja, se as quantidades observadas não são as variáveis do modelo, estas últimas são *convertidas* para as variáveis observadas \mathbf{y}^o .

Quando uma medida é indireta, como, por exemplo, de radiâncias de satélites, o operador de observação H_m pode ser um modelo de transferência radiativa que tem um modelo coluna e calcula a radiância emitida. Os operadores de observação são imprescindíveis para assimilação de dados e podem representar muito no desenvolvimento de algoritmos de assimilação. Na equação (3.1), $\mathbf{M}(\mathbf{x})$ denota o operador total do modelo, mas mantém a derivada de tempo do modelo de estado separado para distinguir a evolução da dinâmica sem assimilação (HOLM, 2008).

Concluindo para melhor entendimento da formulação a seguir tem-se:

campo de referência ou (*background*) é um modelo de previsão ou a previsão de curto prazo x^b .

avaliação inicial das observações denotada como $H(x^b)$ onde H é chamado operador de observação.

incremento de observação é a diferença entre as observações e o modelo de referência ($y^o - H(x^b)$) também chamado de *inovação*.

análise \mathbf{x}^a é obtida pela adição das inovações ao campo de referência com a matriz pesos \mathbf{W} que é determinada baseado na estimação estatísticas das covariâncias dos erros da previsão e das observações.

Com esta notação, a equação da análise pode ser expressa como:

$$x^a = x^b + \mathbf{W}[y^o - H(x^b)] \quad (3.4)$$

onde a matriz de pesos \mathbf{W} (ou matriz "ganho") depende das matrizes de covariância de erros.

$$W = \mathbf{B}H^T(\mathbf{H}\mathbf{B}H^T + \mathbf{R}^{-1})^{-1} \quad (3.5)$$

onde \mathbf{R} é a matriz de covariância dos erros de observação e \mathbf{B} é matriz de covariância de erro do modelo de referência (KALNAY, 2003).

Diferentes técnicas de análises (SCM, OI, 3D-Var, e FK) baseados na equação (3.4), diferem na abordagem dada à combinação do modelo de referência e das observações para produzir a análise. Os métodos mais antigos como SCM utilizam pesos determinados *empiricamente*. Os pesos eram funções da distância entre a observação e o ponto de grade, e a iteração da análise era feita várias vezes.

No esquema de OI, a matriz de pesos \mathbf{W} é determinada pela minimização dos erros de análise em cada ponto de grade. A abordagem variacional tridimensional define uma "função custo" proporcional à diferença quadrática entre análise e ambos, o modelo de referência e as observações. Lorenc (1986) mostrou que as abordagens de OI e 3D-Var são equivalentes se a função custo for definida como:

$$J = \frac{1}{2}[y^o - H(\mathbf{x})]^T \mathbf{R}^{-1}[y^o - H(\mathbf{x})] + (\mathbf{x} - x^b)^T \mathbf{B}^{-1}(\mathbf{x} - x^b) \quad (3.6)$$

A função custo J mede a distância do campo \mathbf{x} e as observações \mathbf{y}^o (primeiro termo) e a distância de \mathbf{x} e a referência (previsão de curto prazo) x^b (segundo termo). O mínimo da função custo é obtida para $\mathbf{x} = \mathbf{x}^a$, que é definida como a análise. A diferença entre a abordagem de OI e a de 3D-Var é o método de solução: (i) em OI, os pesos \mathbf{W} são obtidos para cada ponto de grade ou volume de grade, usando simplificações; (ii) em 3D-Var, a minimização de \mathbf{J} é executada diretamente, permitindo flexibilidade adicional e uso simultâneo de dados globais (KALNAY, 2003, cap.V).

3.1.2.1 Matrizes de Erros

Para representar o fato de que há alguma incerteza no modelo, nas observações e na análise supõe-se um modelo de erros entre os vetores e os vetores "verdadeiros". A

maneira usual de fazê-lo é assumir uma *função densidade de probabilidade*, ou FDP, para cada tipo de erro.

Segundo [Bouttier e Courtier \(1999\)](#), nas aplicações de muitos métodos de ass melhor maneira de se obter as matrizes de erros. Geralmente executa-se um determinado modelo um grande número de vezes com realizações diferentes e erros gerados por causas desconhecidas incluídos nas análises e então, coleta-se as diferenças entre as execuções dos modelos e obtém-se um conjunto de dados para calcular estatísticas. A partir dos conjuntos de dados resultantes, a função de densidade de probabilidades pode ser derivada, bem como todas as estatísticas, inclusive a média e as variâncias. A partir da FDP pode-se obter a matriz de covariância de erro do modelo.

Os erros do modelo de referência são modelados como: $\varepsilon_b = x_b - \mathbf{x}$, com média $\overline{\varepsilon_b}$ e covariância $\mathbf{B} = \overline{(\varepsilon_b - \overline{\varepsilon_b})(\varepsilon_b - \overline{\varepsilon_b})^T}$. Esta é a estimação dos erros do estado do modelo, isto é, as diferenças entre o vetor de estado do modelo e o seus valores verdadeiros.

Os erros das observações são modelados como: $\varepsilon_o = y^o - H(\mathbf{x})$, média $\overline{\varepsilon_o}$ e covariância $\mathbf{R} = \overline{(\varepsilon_o - \overline{\varepsilon_o})(\varepsilon_o - \overline{\varepsilon_o})^T}$. Estes erros devem-se aos processos de observações (erros do instrumento) e erros de representatividade, isto é, erros de discretização que previnem que \mathbf{x} possa ser a imagem perfeita do estado verdadeiro.

Os erros da análise são modelados como: $\varepsilon_a = x_a - \mathbf{x}$, com média $\overline{\varepsilon_a}$. A medida $\|\varepsilon_a - \overline{\varepsilon_a}\|$ destes erros é dada pelo traço da matriz de covariância do erro de análise \mathbf{A} :

$$Tr(\mathbf{A}) = \|\varepsilon_a - \overline{\varepsilon_a}\|^2,$$

que contem os erros da estimação do estado de análise x_a que é o que se procura *minimizar*:

Em um sistema escalar, a covariância do erro de modelo é simplesmente a variância. Em um sistema multi-dimensional, as covariâncias estão em uma matriz quadrada simétrica. Se o vetor de estado do modelo tem dimensão N_x , então a matriz de covariância de erros do modelo é de dimensão $(N_x \times N_x)$. A diagonal da matriz contém variâncias para cada variável do modelo; os termos fora da diagonal são as covariâncias cruzadas entre cada par de variáveis do modelo. A matriz deve ser positiva definida.

As estatísticas de erros (desvios e covariâncias) são funções de processos físicos que regem as simulações meteorológicas da rede de observação. Eles também dependem do conhecimento "a priori" dos erros. As variâncias de erro em particular refletem a incerteza do modelo ou das observações.

Quando da criação de um sistema de assimilação, na prática, as aproximações são inevitáveis, pois é muito difícil obter dados precisos para calibrar as estatísticas: os erros de estimativa não podem ser observados diretamente. Algumas informações sobre os valores das estatísticas podem ser coletados a partir de diagnósticos de um sistema de assimilação de dados existente para obter \mathbf{B} . Porém, pode-se utilizar o *método NMC*, método desenvolvido no antigo *National Meteorological Centers* por Parrish e Derber (1992), onde usa-se as diferenças entre previsões de curto prazo, como substituto para o modelo de referência:

$$B \approx \overline{(\mathbf{x}^{t2} - \mathbf{x}^{t1})(\mathbf{x}^{t2} - \mathbf{x}^{t1})^T},$$

onde x^{t2} é a previsão de 24 horas, x^{t1} é a previsão de 12 horas, por exemplo.

As covariâncias de erro dependem de ciclos de previsão, como descrito acima, ou podem ser estimadas diretamente, como no filtro de Kalman. Finalmente, no senso comum, as estatísticas de erro são usadas, à medida que refletem o conhecimento "a priori" dos processos físicos responsáveis pelos erros (BOUQUIER; COURTIER, 1999).

A equação fundamental para a análise linear em sua forma geral algébrica (equação 3.4) pode ser vista como uma estimativa de mínimos quadrados, caracterizada como *Best Linear Unbiased Estimator (BLUE)*. A estimativa de mínimos quadrados está presente em muitos algoritmos usados em meteorologia e oceanografia. Esta abordagem em relação à assimilação de dados foi discutida em Härter (2004), CAMPOS VELHO et al. (2007), Bouquier e Courtier (1999), Todling (1999) e Kalnay (2003).

3.2 Abordagem Bayesiana da Análise

Nesta seção, a revisão da equação de análise com derivação Bayesiana será discutida em termos da função de distribuição de probabilidade (FDP) multidimensional e por um conjunto padrão de suposições.

Revisando a notação, considere um modelo de previsão particular a ser usado que tem uma base discreta para o modelo de estado denotado por um vetor x com

dimensão N_x . Os dados observados são representados por um vetor y de dimensão N_y . A relação entre a base modelo x e as observações y é representada por um operador explícito H_n tal que

$$y = H_n(x) \tag{3.7}$$

seja a melhor estimativa de y para um determinado x . H pode ser não-linear, embora freqüentemente seja um operador de interpolação simples. O problema de análise é então achar o "melhor" x que inverta a equação (3.7) para um determinado y , permitindo erros de observação e outras informações anteriores; este é um problema "inverso" discreto (MENKE, 1984).

Segundo Lorenc (1986), a abordagem Bayesiana usa probabilidades para descrever a precisão do conhecimento sobre os eventos passados. Temos, então, um formalismo para modificar as probabilidades à luz de novos conhecimentos, o que se faz na assimilação de dados sequencial.

Considere os eventos A e B . $P(A)$ é a probabilidade de A ocorrer, ou uma medida da certeza de que A ocorreu no passado. $P(A|B)$ é a probabilidade condicional de A dado que B ocorreu. Duas maneiras de expressar $P(A \cap B)$:

$$\begin{aligned} P(A \cap B) &= P(B)P(A|B) \text{ ou} \\ &= P(A)P(B|A). \end{aligned} \tag{3.8}$$

Isso leva diretamente ao *Teorema de Bayes*:

$$P(A|B) \propto P(B|A)P(A). \tag{3.9}$$

O teorema de Bayes afirma que a probabilidade posterior de um evento A acontecer, dado que, o evento B é conhecido, é proporcional à probabilidade anterior de A , multiplicado pela probabilidade de B acontecer, dado que A é conhecida. Este teorema é aplicável ao problema inverso da análise.

Se A é o evento $x = x^t$ e B é o evento $y = y^o$, então

$$P(x = x^t|y = y^o) \propto P(y = y^o|x = x^t)P(x = x^t) \tag{3.10}$$

o supercritro t indica o valor verdadeiro, e o indica o valor observado. A equação 3.10 define uma FDP de dimensão N_x denominada de $\mathbf{P}^a(x)$, especificando todo conhecimento da análise. A "melhor" estimativa x pode ser a média, ou pode ser o

módulo de $\mathbf{P}^a(x)$, ou ainda:

$$x^a = \int xP^a(x)dx, \quad (3.11)$$

$$x^a = x \text{ tal que } P^a(x) \text{ seja o máximo.} \quad (3.12)$$

Estas equações (3.11) e (3.12) são respectivamente: a variância mínima e a estimativa de máxima verossimilhança de x^a .

Para uma solução completa do problema inverso generalizado, é necessário saber a precisão de x^a , informação também contida em $P^a(x)$. A probabilidade "a priori" $P(x = x^t)$ encapsula o conhecimento de x antes das observações. A referência deve ser uma função de x em termos de desvios de algum modelo de referência x^b conhecido.

$$P(x = x^t) = P^b(x = x^b)$$

Uma nova informação fornecida da observação com um instrumento perfeito seria uma medida de valor verdadeiro y^t . Os instrumentos reais não são perfeitos, têm erros que são assumidos por uma FDP Gaussiana. Qualquer observação física é sujeita a erros aleatórios de observação. Se os valores verdadeiros de y forem conhecidos, então $y^1 = y^t$, e estes erros podem ser expressos por

$$P(y = y^o | y^1 = y^t) \equiv P^o(y^o - y^t). \quad (3.13)$$

P^o pode ser ainda, uma função de x^t , ou de x^b . A equação 3.10 deve conhecer $P(y = y^o | x = x^t)$ que pode ser $P^{of}(y^o - y^f)$ onde $[y^f = H_n(x^t)]$ é o valor interpolado do valor observado e do modelo, sendo H_n , em geral, inexato. Então:

$$P(y^1 = y^t | x = x^t) = P^f(y^f - y^t). \quad (3.14)$$

Se H_n fosse exato, seria $(y^f = y^t)$ e $(P^{of}(y^o - y^f) = P(y^o - y^f))$. E, como em geral H_n é inexato, estes erros são descritos por

$$P(y^1 = y^t | x = x^t) = P^{of}(y^1 - y^o) \quad (3.15)$$

E $P(A) = \int P(A|B)P(B)db$, onde B é o evento em que um parâmetro tem um valor entre b e $b + db$. Assim

$$P(y = y^o | x = x^t) = \int P(y = y^o | x = x^t \text{ e } y^1 = y^t)P(y^1 = y^t | x = x^t)dy^1, \quad (3.16)$$

$$P^{of}(y^o - y^f) = \int P^o(y^1 - y^o)P^f(y^1 - y^f)dy^1. \quad (3.17)$$

Substituindo a equação (3.10), tem-se

$$P^a(x) \propto \int P^o(y^1 - y^o)[P^f(y^1 - H_n(x))]dy^1 P^b(x - x^b). \quad (3.18)$$

Com esta nomenclatura supõe-se que $P^b(x - x^b)$ e $P^{of}(y^o - H_n(x))$ são independentes, isto é, os erros da referência e os erros de observação são não correlacionados. É possível considerar a seguinte derivação sem essa suposição, por FDP conjunta. Lorenc (1986) especifica as funções de distribuição de probabilidade de P^b , P^o , e P^{of} da equação (3.18) e resolver (3.11) ou (3.12), com a suposição comum que simplifica a solução, onde FDPs são funções multi-dimensionais e Gaussianas:

$$P^b(x - x^b) \propto \exp \left[-\frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b) \right], \quad (3.19)$$

$$P^o(y - y^o) \propto \exp \left[-\frac{1}{2}(y - y^o)^T \mathbf{O}^{-1}(y - y^o) \right], \quad (3.20)$$

$$P^f(y - y^f) \propto \exp \left[-\frac{1}{2}(y - y^f)^T \mathbf{F}^{-1}(y - y^f) \right], \quad (3.21)$$

onde \mathbf{B} , \mathbf{O} , \mathbf{F} , são matrizes de covariâncias, então

$$\mathbf{B} = \overline{(x^b - x^t)(x^b - x^t)^T}, \quad (3.22)$$

$$\mathbf{O} = \overline{(y^o - y^t)(y^o - y^t)^T}, \quad (3.23)$$

$$\mathbf{F} = \overline{(y^t - H_n(x^t))(y^t - H_n(x^t))^T}. \quad (3.24)$$

P^{of} , definido na equação (3.17), é dado por

$$P^{of}(y^o - y^f) \propto \exp \left[-\frac{1}{2}(y^o - y^f)^T (\mathbf{O} + \mathbf{F})^{-1}(y^o - y^f) \right] \quad (3.25)$$

Ainda, segundo Lorenc (1986), as estimativas de mínima variância e máxima verossimilhança (equações (3.11) e (3.12)) são idênticas. Para a estimativa de máxima verossimilhança, maximiza-se:

$$\begin{aligned} P^a(x) &= P^{of}[y^o - H_n(x)]P^b(x - x^b) \\ &\propto \exp \left\{ -\frac{1}{2}[y^o - H_n(x)]^T (\mathbf{O} + \mathbf{F})^{-1}[y^o - H_n(x)] - \frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b) \right\}. \end{aligned}$$

Maximizar P^a é equivalente a minimizar a função custo

$$J = [y^o - H_n(x)]^T (\mathbf{O} + \mathbf{F})^{-1} [y^o - H_n(x)] + (x - x^b)^T \mathbf{B}^{-1} (x - x^b). \quad (3.26)$$

Se \mathbf{H} é a matriz de derivadas parciais de H_n para os componentes de x , então:

$$H_n(x + dx) = H_n(x) + \mathbf{H}dx \quad (3.27)$$

e se a linearização de H_n é válida em todo o intervalo de valores prováveis de x na região de x^b , então uma solução explícita pode ser encontrada pela avaliação de \mathbf{H} em $x = x^b$ é que permite escrever

$$H_n(x^a) = H_n(x^b) + \mathbf{K}(x^a - x^b) \quad (3.28)$$

Se assumir que os erros instrumentais e da aplicação do operador H são independentes, pode-se combiná-los para obter *um único* erro de observação, ou seja, o erro de observação será a soma das covariâncias dos erros instrumentais e representatividade, $F + O$, pode ser escrita como uma única matriz de covariância de erro de observação como \mathbf{R} . Então, a solução explícita pode ser manipulada em diversas formas:

$$x^a = x^b + [\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{I}]^{-1} \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} y^o - H_n(x^b) \quad (3.29)$$

$$x^a = x^b + \mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{B}\mathbf{H}^T \mathbf{R}^{-1} + \mathbf{I}^{-1} [y^o - H_n(x^b)] \quad (3.30)$$

$$x^a = x^b + \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} [y^o - H_n(x^b)] \quad (3.31)$$

A média da covariância de erros da análise, para este caso de suposições Gaussianas linearizadas, é dada por:

$$\overline{(x^a - x^t)(x^a - x^t)^T} = \mathbf{B} - \mathbf{B}\mathbf{H}^T (\mathbf{K}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{B} \quad (3.32)$$

e pode ser minimizada para solução de x^a .

Os métodos de solução iterativa de minimização, geralmente estão relacionados com as equações(3.29) ou (3.30). Ao reavaliar $H_n(x)$ e, se necessário, \mathbf{H} , a cada iteração, é permitida uma não-linearidade em H_n . Ao variar (por exemplo) \mathbf{O} , pode permitir erros não Gaussianos (LORENC, 1995).

Logo, segundo Lorenc (2008) com FDPs conjuntas modelada por Gaussianas multi-

dimensionais temos a equação de análise Bayesiana:

$$p(x|y^o) = \frac{p(y^o|x)p(x)}{p(y^o)} \propto N(y^o|\mathbf{H}(x), R)N(x|x^b, \mathbf{B}) \quad (3.33)$$

Uma propriedade das FDPs Gaussianas é que \mathbf{H} é linearizável, então:

$$\begin{aligned} & Normal(y^o|H_n(x), R) \quad Normal(x|x^b, \mathbf{B}) \\ = & Normal(y^o|H_n(x^b), \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) \quad Normal(x|x^a, \mathbf{A}) \end{aligned} \quad (3.34)$$

onde x^a e \mathbf{P}^a são definidos por:

$$\mathbf{P}^a = \mathbf{B} - \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{B} \quad (3.35)$$

$$x^a = x^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}[y^o - H_n(x^b)] \quad (3.36)$$

Se $p(x|y^o) = N(x|x^a, \mathbf{P}^a)$, então a probabilidade posterior, depois de adicionar a informação da observação, é uma Gaussiana com média x^a e variância \mathbf{P}^a . Assim, a combinação Bayesiana de FDPs Gaussianas provê a mesma análise “ótima” como um método de mínimos quadrados ponderados.

A estimativa da densidade de probabilidade em um espaço de dimensão muito grande, pode ter um custo computacional enorme, a menos que sejam feitas hipóteses simplificadoras sobre a forma da distribuição. Por exemplo, a densidade de probabilidade de uma grade de 100×100 , com uma estrutura “a priori” em 100 dimensões. Então, se as probabilidades conjuntas fossem avaliadas em uma grade semelhante, implicaria em modificar 100^{100} estimativas de densidade. Esses cálculos são proibitivos para um estado de modelo 100-dimensional.

Em consequência do alto custo computacional para estimar as matrizes de covariância de erros, as técnicas de *Monte-Carlo* passaram a ser investigadas e aplicadas em métodos de assimilação. Os cálculos do refinamento da FDP pela atualização de novas observações não são explícitos, nem a subsequente difusão da probabilidade na previsão, devido aos erros do modelo. Como aproximação, obtém-se aleatoriamente a amostra inicial da distribuição de probabilidade $p(x|y)$. Posteriormente, simula-se os efeitos dos erros do modelo e das observações. Os métodos de Monte-Carlo de assimilação de dados são utilizados e descrevem um conjunto de modelos de estado com incerteza inerentes às observações. As probabilidades calculadas passam a ser

computacionalmente possíveis (HAMILL et al., 2001).

3.3 Filtros de Kalman

Nesta seção é feita a revisão bibliográfica de métodos baseados em Filtro de Kalman por conjunto, descrevendo algumas técnicas propostas e publicadas pelos autores referenciados. As técnicas apresentam a derivação na direção do método LETKF, a ser “emulado” pela rede neural artificial, objetivo da presente tese.

A metodologia de assimilação de dados baseada em ‘conjuntos’ vem em parte das técnicas de Monte-Carlo, mas os conceitos subjacentes também são derivados de um método conhecido como *Filtro de Kalman* (Kalman (1960), Kalman e Bucy (1961), Jazwinski (1970, sec. 7.3), Gelb (1974, sec. 4.2), Maybeck (1979, sec. 5.3), Ghil (1997), Daley (1991, sec. 13.3), Talagrand (1997)).

O Filtro de Kalman (FK) é uma aproximação Bayesiana para estimação de estado que assume a linearidade do crescimento de erro e distribuições de erro Gaussianas. Há dois componentes do filtro de Kalman: um passo de *previsão*, onde o estado atualizado e a estimativa das incertezas são propagadas para a frente no tempo e um passo de *atualização*, onde a estimativa de estado e a estimativa das incertezas da previsão são ajustadas quando o próximo conjunto de observações se torna disponível.

3.3.1 Filtro de Kalman Estendido

No caso de dinâmica linear, a técnica matematicamente consistente para definir uma matriz de covariância de erro do modelo de referência é o *Filtro de Kalman* (KALMAN, 1960), que utiliza as equações dinâmicas para evoluir o estado mais provável e a matriz de covariância de erro no tempo. No caso de sistemas lineares com erros imparciais distribuídos, o filtro de Kalman fornece estimativas do estado do sistema que são *ótimas* no sentido dos mínimos quadrados. O método também foi adaptado para sistemas não-lineares, sendo chamado de *filtro de Kalman estendido*, ou (EKF de *Extended Kalman Filter*) (Jazwinski (1970), Kalman e Bucy (1961), Gelb (1974), Ghil e Malanotte-Rizzoli (1991), Bouttier e Courtier (1999)).

No filtro de Kalman discreto, a notação da matriz de covariância de erro do modelo B passa a ser \mathbf{P}^f e x^b passa a ser x^f para o modelo de estado que é a previsão (f - *forecast*). Porém, a notação usada aqui, para o EKF e os filtros baseados em conjuntos (*ensemble*) continuará sendo a do modelo de referência ($b = \textit{background}$),

\mathbf{P}^b para matriz de covariância de erro do modelo e x^b como modelo de estado de referência.

O EKF assume que as distribuições de erro de referência e da observação são Gaussianas: $x_t^b = x_t^t + \epsilon$, onde $\epsilon \approx N(0, \mathbf{P}_t^b)$, ou seja, a densidade de probabilidade anterior é uma distribuição normal multi-variada com média ($N_x \times 1$) no passo t . O modelo de previsão x_t^b conhecido tem uma matriz de covariância de erros $n \times n$ \mathbf{P}_t^b . Da mesma forma, $y = H(x_t^t) + \epsilon$ onde $\epsilon \approx N(0, \mathbf{R})$. H é o operador que mapeia o estado para as observações. $\mathbf{H} = \partial H / \partial x$ representa a matriz Jacobiana ($N_x \times N_y$) $\mathbf{M} = \partial M / \partial x$ representa uma $n \times n$ matriz Jacobiana chamada matriz de transição entre os passos t e $t+1$. N_x é a dimensão do modelo e N_y é a dimensão do vetor observação. \mathbf{M}^T representa um modelo adjunto (DIMET; TALAGRAND, 1986). \mathbf{Q} representa a $N_x \times N_x$ matriz de covariância de erros do modelo ou erros acumulados entre os ciclos de atualização.

As equações do EKF são:

$$x_{t+1}^b = M(x_t^a) \quad (3.37)$$

$$\mathbf{P}_{t+1}^b = \mathbf{M}\mathbf{P}_t^a\mathbf{M}^T + \mathbf{Q} \quad (3.38)$$

$$= \mathbf{M}(\mathbf{M}\mathbf{P}_t^a)^T + \mathbf{Q} \quad (3.39)$$

$$x_t^a = x_t^b + \mathbf{K} [y_t^o - H(x_t^b)], \quad (3.40)$$

$$\mathbf{K} = \mathbf{P}_t^b \mathbf{H}^T (\mathbf{H}\mathbf{P}_t^b \mathbf{H} + \mathbf{R})^{-1}, \quad (3.41)$$

$$\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_t^b \quad (3.42)$$

O conjunto das equações (3.37) - (3.39) propagam a análise com o modelo de previsão não-linear, ou seja, a equação (3.37) é a previsão (feita a partir de uma análise obtida quando novas observações do ciclo estiverem disponíveis). A equação (3.38) calcula a matriz de covariância de erro do modelo. Assume-se que os erros são não correlacionados e o crescimento dos erros de análise é calculado através do modelo de previsão tangente linear. No conceito do filtro de Kalman, as covariâncias de erro da previsão e análise são explicitamente prognosticadas. Estas equações denotam o passo de “previsão” do EKF.

As equações (3.40) a (3.42) denotam o passo de *atualização*. A análise ótima do estado x_t^a é estimada pela correção da previsão x_t^b através do *incremento de observação*

$[y_t^o - H(x_t^b)]$, ponderado pela matriz *ganho de Kalman* denotada por \mathbf{K} . O efeito de \mathbf{K} é aplicar o incremento de observação para corrigir o modelo de referência nos pontos de grade e seus pontos vizinhos. A equação (3.42) indica como atualizar a covariância de erros da análise de modo a refletir a redução das incertezas de assimilar as observações.

A análise reduz as variâncias dos erros em locais onde as observações são abundantes e as covariâncias de erros e a previsão crescem a uma taxa proporcional ao crescimento de erro local. Consequentemente a diferença $(x_t^a - x_t^b)$ depende da estrutura das covariâncias do erro do modelo de referência e do ajustamento das observações, que podem ser complicadas e dependente do fluxo de observação e do tempo (BOUTTIER, 1994).

Os filtros de Kalman assumem crescimento linear e normal dos erros para o problema da assimilação, porém podem ser inapropriados para dados observados atmosféricos e outros aspectos do modelo de estado, que podem ser muito sensíveis a movimentos em pequenas escalas. Onde a escala de tempo de previsibilidade é pequena, os erros crescem e saturam rapidamente. Da mesma forma, se as observações não são regularmente disponíveis, as covariâncias de erro estimadas com o modelo tangente linear podem crescer sem limite (Evensen (1992), Bouttier (1994)).

As estatísticas de erro devem ser cuidadosamente avaliadas e monitoradas, em particular, é importante que a covariância dos erros da matriz do modelo deva ser estimada com precisão. Por exemplo, se as variâncias dos erros da previsão são subestimadas, o EKF assume que as estatísticas de erro do modelo são mais precisas que as observações vizinhas e, portanto, não irá corrigir a previsão (DALEY, 1991, p. 382).

Um dos desafios científicos do século XXI para o EKF, seria a estimativa da matriz \mathbf{Q} , particularmente difícil de ser calculada. A precisão da assimilação pode depender fortemente do modelo assumido para calcular a matriz \mathbf{Q} .

Outra desvantagem dos filtros de Kalman para assimilação de dados atmosféricos é o seu custo computacional. Apesar dos filtros de Kalman fornecerem uma redução no custo computacional em relação à assimilação de dados Bayesiana completa, um vetor de estado de alta dimensão, os custos computacionais em modelos de previsão do tempo podem ser incrivelmente grandes.

Mais considerações sobre o EKF podem ser encontradas em Jazwinski (1970) e Gelb (1974). No caso do filtro de Kalman discreto, com pressupostos de proporcionar a melhor estimativa linear pela (*Best Linear Unbiased Estimate-BLUE*), pode ser encontrado em Talagrand (1997).

3.3.2 Filtros de Kalman baseados em conjunto

O reconhecimento da restrição do EKF (custo computacional) que não permite implementação operacional da previsão do tempo (DALEY, 1991), levou ao desenvolvimento de estimativas reduzidas da matriz de covariância de erros do modelo. Basicamente, existem duas maneiras de obter a estimativa reduzida:

- sistemas baseados em linearização da dinâmica do modelo, projetando o estado do modelo em uma base que é de menor dimensão do que o espaço completo do modelo;
- a utilização de um conjunto de previsões, relativamente pequeno, para estimar a covariância de erros da previsão.

Em esquemas de assimilação de dados baseados em conjuntos, o conjunto das previsões é gerado pela evolução de um conjunto de condições iniciais distribuídas de acordo com o resultado da análise anterior. Esta abordagem tem a vantagem de não exigir o desenvolvimento e manutenção de um modelo linear de previsão.

Os algoritmos de assimilação por conjuntos (*ensemble*) utilizam técnicas do tipo Monte-Carlo e podem fornecer uma análise mais precisa do que a análise de EKF, em situações onde a não linearidade é pronunciada e as FDPs exigem a não-gaussianidade. Estes algoritmos de assimilação podem trabalhar com poucos membros do conjunto de vetores de estado e apresentar menor custo computacional.

Muitos pesquisadores têm proposto uma variedade de métodos de assimilação por conjunto. Apesar das muitas diferenças entre os vários algoritmos baseados em conjunto, todos são formados por um número finito de análises e previsões. As covariâncias de erro do modelo de referência são modeladas usando um conjunto de análises feitas com as observações disponíveis e pelo conjunto de previsões de curto prazo produzidas.

Os algoritmos de assimilação por conjunto também têm uma propriedade desejável, se a dinâmica dos erros é de fato linear e as estatísticas são Gaussianas, à medida que

aumenta o tamanho do conjunto, o estado e a estimativa da covariância convergem para os mesmos do filtro de Kalman estendido (EKF) (BURGERS et al., 1998).

Os conceitos de métodos de assimilação por conjunto têm sido usados em aplicações aeroespaciais e de engenharia, desde os anos 1960. Potter (1964), Andrews (1968), Maybeck (1979, cap. 7), Leith (1983) esboçam a idéia básica para a assimilação de dados atmosféricos. A idéia foi descrita e testada em aplicação oceanográfica por Evensen (1994) e na assimilação de dados atmosféricos por Houtekamer e Mitchell (1998).

Por simplicidade de notação, o subscrito t (o passo de tempo utilizado na seção anterior) será descartado.

Assume-se um conjunto de previsões como amostra aleatória dos erros do modelo de referência. Denota-se \mathbf{X}^b uma matriz cujas colunas são compostas por membros do conjunto de vetores de estado.

A média do conjunto \mathbf{X}^b é definida como

$$\overline{\mathbf{X}}^b = \frac{1}{K} \sum_{k=1}^K x_k^b. \quad (3.43)$$

O subscrito k denota um membro do conjunto e K é número de membros. A perturbação da média para o k -ésimo membro é $x_k^b = x_k^b - \overline{\mathbf{X}}^b$.

Define-se \mathbf{X}'^b como a matriz formada de um conjunto de *perturbações*

$$\mathbf{X}'^b = (x_1^b, \dots, x_K^b), \quad (3.44)$$

e $\hat{\mathbf{P}}^b$ como uma estimativa de \mathbf{P}^b para um conjunto finito de perturbações:

$$\hat{\mathbf{P}}^b = \frac{1}{K-1} \mathbf{X}'^b \mathbf{X}'^{bT}. \quad (3.45)$$

Ao contrário do filtro de Kalman ou do método 3D-Var, a estimativa de covariância de erro de modelo de referência é gerada a partir de um conjunto construído especialmente de previsões não-lineares.

A hipótese de crescimento linear do erro no EKF irá resultar em uma super estimativa da variância do erro do modelo, mas as diferenças entre os membros do conjunto

não crescem sem limites e, portanto, devem fornecer uma forma mais precisa de estatísticas reais do erro do modelo de referência.

Ao contrário de algoritmos de assimilação de dados, por exemplo, o método 3D-Var (a maioria das implementações em análises operacionais), as covariâncias de erro podem variar no tempo e no espaço. Se o modelo de covariância de erro é relativamente preciso, então ele proporcionará uma melhor adaptação às observações (HAMILL et al., 2001).

Considere o passo de atualização (equação (3.42)) em duas classes gerais de filtros por conjunto: filtro estocástico (totalmente Monte Carlo) e filtros deterministas. Ambas as classes, propagam o conjunto de análises com modelos de previsão não-lineares. A principal diferença é o ruído aleatório aplicada durante o passo de atualização para simular a incerteza da observação.

A diferença entre os métodos ou esquemas por conjunto está na geração do conjunto de análises. Uma família de sistemas baseia-se na perturbação de observações como em Evensen (1992); Houtekamer e Mitchell (1998) e Hamill et al. (2001). Nesta abordagem, o conjunto de análises é obtido pela assimilação de um conjunto diferente de observações para cada membro do conjunto do modelo. Os diferentes conjuntos de observações são criados adicionando ruído aleatório às observações reais, onde o componente de ruído aleatório é gerado de acordo com a matriz de covariância dos erros de observação (OTT et al., 2004), conhecidos como algoritmos de atualização estocástica.

A outra família de esquemas por conjunto, os filtros de Kalman raiz quadrada, usa uma abordagem diferente para reduzir o tamanho do conjunto. Essas técnicas fazem a análise apenas uma vez, para obter tanto a média da análise e a matriz de covariância de erro da análise. Logo, as perturbações do conjunto de análise (ou da média da análise) são geradas para transformar o conjunto de perturbações do modelo de referência para um conjunto de vetores, representar a matriz de covariância de erro da análise, os algoritmos são conhecidos como de atualização determinística. Exemplos da abordagem do filtro raiz quadrada são; o *Ensemble Transform Kalman Filter - ETKF* (BISHOP et al., 2001), o *Ensemble Adjustment Kalman Filter - EAKF* (ANDERSON, 2001) e o *Ensemble Square-root Filter - EnSRF* (WHITAKER; HAMILL, 2002).

Algoritmos de atualização Estocástica

O mais conhecido algoritmo estocástico de assimilação de dados baseado em conjunto é o *Ensemble Kalman Filter*, ou "EnKF" (Houtekamer e Mitchell (1998), Burgers et al. (1998), Hamill (2004), Houtekamer et al. (2005)).

O EnKF é um método sequencial de assimilação de dados, baseado na previsão das estatísticas de erro usando métodos Monte Carlo, propondo uma alternativa para resolver as equações usadas nas tradicionais aproximações das covariâncias dos erros usadas no Filtro de Kalman Estendido (EVENSEN, 1994).

O EnKF tem provado ser eficiente com sistemas de dinâmica não-linear forte e os com grandes espaços de estado e agora é usado em aplicações realistas, com modelos de equações primitivas para o oceano e atmosfera, um exemplo é o MSC (*Meteorological Service of Canada*).

O algoritmo do EnKF tem o número de propriedades desejáveis:

- (i) Ele usa o conjunto de estados do modelo integrados no tempo, usando um esquema de análise que atua diretamente sobre o conjunto quando as observações são assimiladas. Para estimar as covariâncias do erro de previsão não requer um modelo de correlação com parâmetros multivariados dos erros;
- (ii) é independente da previsão do modelo, de modo que qualquer modelo de previsão, possivelmente incluindo sofisticadas parametrizações não-lineares, pode ser usado para gerar o conjunto de campos de previsão;
- (iii) é capaz de utilizar as observações não-convencionais, tais como radiâncias de satélite (ARAVEQUIA, 2008);
- (iv) funciona razoavelmente, mesmo se o número de membros do conjunto é modesto. Isso é importante em um contexto operacional.
- (v) produz o incremento de análise espacialmente;
- (vi) pode assimilar de forma eficiente um grande número de observações. Em um ambiente operacional, o número de observações disponíveis pode exceder 10^5 em períodos de 6 h;

(vii) o algoritmo é adequado para computação paralela.

Algumas dessas propriedades (em particular (ii), (iii) e, implicitamente, (v) e (vi)) coincidem com os objetivos de projeto do *Physical-space Statistical Analysis System (PSAS)*, descrito por Cohn et al. (1998) que são discutidos em Houtekamer e Mitchell (1998), Houtekamer e Mitchell (1998), Cintra (2004).

O EnKF foi independentemente proposto pelo oceanógrafo Geir Evensen (JGR, 1994) e Houtekamer e Fillion (NCAR, 1993). As aplicações oceanográficas e atmosféricas visam obter a melhor análise possível, utilizando um conjunto de condições iniciais para um sistema de previsão por conjunto.

O algoritmo do EnKF atualiza cada membro com um conjunto diferente de observações perturbadas com ruído aleatório. Devido a aleatoriedade ser introduzida a cada ciclo de assimilação, a atualização é considerada estocástica.

O EnKF executa um conjunto de ciclos de assimilação de dados em paralelo, $k = 1, \dots, K$, com cada membro atualizado para uma diferente realização das observações:

$$x_k^a = x_k^b + \hat{\mathbf{K}} (y_k^o - H(x_k^b)), \quad (3.46)$$

na equação (3.46), $(y_k^o = y + y_k')$ é o vetor de "observações perturbadas" definidas como $y_k' \approx N(0, \mathbf{R})$:

$$\frac{1}{m} \sum_{k=1}^m y_k' = 0, \quad (3.47)$$

os m grupos de observações perturbadas são criados para atualizar os k diferentes campos do modelo de referência; na equação (3.46) a matriz $\hat{\mathbf{K}}$

$$\hat{\mathbf{K}} = \hat{\mathbf{P}}^b \mathbf{H}^T (\mathbf{H} \hat{\mathbf{P}}^b, \mathbf{H}^T + \mathbf{R})^{-1} \quad (3.48)$$

é similar ao "ganho" de Kalman do EKF em (3.40), porém utilizando o conjunto de erros de previsão para estimar a matriz de covariâncias de erros do modelo de referência, a equação (3.45).

Note que o EnKF assimila observações perturbadas (3.18), em vez das próprias observações. \mathbf{X}^b é uma matriz de análise de desvios do membro do conjunto da média da análise do estado, enquanto (3.44) define os desvios do modelo de referência. $\hat{\mathbf{P}}^a$ provem do conjunto de análises assimilando observações perturbadas usando (2.13).

Então, como o tamanho do conjunto pode ser variado e se a dinâmica é linear, $\hat{\mathbf{P}}^a = \frac{1}{k-1} \mathbf{X}^b \mathbf{X}^{b'} \rightarrow \mathbf{P}^a$, onde \mathbf{P}^a é a covariância dos erros de análise (3.40) do filtro de Kalman estendido (BURGERS et al., 1998).

Diversos algoritmos têm sido propostos para a simplificar e paralelizar a codificação dos métodos de EnKF. Para um modelo operacional complexo de PNT com um vetor de estado de grande dimensão, na forma explícita de \mathbf{P}^b , seria computacionalmente proibitiva, por exemplo, se este modelo complexo for da ordem de 10^7 elementos em seu estado, armazenar e acessar facilmente os 10^{14} elementos de \mathbf{P}^b não é possível.

No entanto, em métodos baseados em conjuntos, $\hat{\mathbf{K}}$ pode ser formada sem calcular explicitamente a $\hat{\mathbf{P}}^b$ total (EVENSEN, 1994). Os componentes da $\hat{\mathbf{P}}^b \mathbf{H}^T$ e de $\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T$ de $\hat{\mathbf{K}}$ são computados separadamente da seguinte forma: calcula-se

$$\overline{H(x^b)} = \frac{1}{K} \sum_{k=1}^n H(x_k^b), \quad (3.49)$$

que representa a média da estimativa das observações interpoladas da previsão. Então, calcula-se

$$\hat{\mathbf{P}}^b \mathbf{H}^T = \frac{1}{K} \sum_{k=1}^K (x_k^b - \hat{x}^b) \left(H(x_k^b) - \overline{H(x^b)} \right)^T, \quad (3.50)$$

e

$$\mathbf{H} \hat{\mathbf{P}}^b \mathbf{H}^T = \frac{1}{m} \sum_{k=1}^{n_x} \left(H(x_k^b) - \overline{H(x^b)} \right) \left(H(x_k^b) - \overline{H(x^b)} \right)^T. \quad (3.51)$$

Algoritmos de atualização Determinística

Vários métodos têm sido propostos para corrigir o conjunto de previsões para novas observações tal que $\hat{\mathbf{P}}^a \rightarrow \mathbf{P}^a$ seja sem adição de ruído aleatório. Os algoritmos que não acrescentam ruído estocástico são chamados *determinísticos*, assim chamados porque se o conjunto dos modelos de referência e as estatísticas de erros associadas são conhecidos, o conjunto de estados da análise será conhecido também.

Esses algoritmos (por exemplo, Bishop et al. (2001), Anderson (2001), Whitaker e Hamill (2002), Hunt et al. (2004)) atualizam a covariância gerando o mesmo erro de análise que seria obtido do filtro de Kalman, admitindo que a covariância de erro do filtro de Kalman do modelo de referência, seja modelada a partir do conjunto de

referências. [Tippett et al. \(2003\)](#) descrevem as semelhanças e diferenças entre vários desses algoritmos.

Nos métodos por conjunto, as covariâncias de erros do modelo de referência, nunca são explicitamente formadas, manipulações são executadas usando a raiz quadrada da matriz (isto é, altera a equação (3.44), a matriz de desvios da média dos membros do conjunto). Conforme [Tippett et al. \(2003\)](#), dada uma matriz \mathbf{U} que represente qualquer $n \times n$ transformação ortogonal tal que $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$, então $\hat{\mathbf{P}}^b$ também pode ser representado como $\hat{\mathbf{P}}^b = \frac{1}{m-1}(\mathbf{X}'^b\mathbf{U})(\mathbf{X}^b\mathbf{U})^T$. Assim, muitos filtros de raiz quadrada podem ser formulados para produzirem a mesma covariância de erros de análise ([HAMILL et al., 2001](#)).

O “Ensemble square-root filter” ou “EnSRF” descrito por [Whitaker e Hamill \(2002\)](#) é matematicamente equivalente ao filtro adaptativo descrito por [Anderson \(2001\)](#), como no EnKF, o EnSRF realiza um conjunto de ciclos de assimilação de dados em paralelo. No EnSRF, é conveniente atualizar as equações da média do conjunto *ensemble* (representado por uma barra superior) e do desvio do k -ésimo membro da média separadamente, $x_k^b = x_k^{\bar{b}} - \bar{\mathbf{X}}^b$. E calcula-se:

$$\bar{x}^a = \bar{x}^b + \hat{\mathbf{K}}(y^o - H(\bar{x}^b)), \quad (3.52)$$

$$x_k^a = x_k^b + \tilde{\mathbf{K}}H(x_k^b) \quad (3.53)$$

onde $\hat{\mathbf{K}}$ é o tradicional ganho de Kalman da equação (3.48), e $\tilde{\mathbf{K}}$ é o ganho “reduzido” usado para atualizar os desvios da média do conjunto.

Quando a sequência do processamento de observações independentes, $\hat{\mathbf{K}}$, $\tilde{\mathbf{K}}$, $\mathbf{H}(\hat{\mathbf{P}}^b)$ e $(\hat{\mathbf{P}}^b)\mathbf{H}^T$ são matrizes n -dimensionais, e $\mathbf{H}(\hat{\mathbf{P}}^b)\mathbf{H}^T$ e R são escalares. Assim, como referido pela primeira vez por [Potter \(1964\)](#) quando as observações são processadas uma a uma:

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}(\hat{\mathbf{P}}^b)\mathbf{H}^T}}\right)^{-1} \hat{\mathbf{K}}. \quad (3.54)$$

A matriz $\hat{\mathbf{K}}$ multiplicada na equação (3.54) torna-se um escalar entre 0 e 1. Isto significa que, a fim de obter a covariância de erros da análise correta com observações não perturbadas, o ganho de Kalman é reduzido em magnitude, em relação ao ganho de Kalman tradicional.

Outro algoritmo de atualização determinística é o filtro de Kalman por conjunto transformado (*Ensemble Transform Kalman filter (ETKF)*) de Bishop et al. (2001). Neste contexto, o ETKF transforma perturbações das previsões em perturbações de análise de uma forma consistente com a equação de atualização da covariância de erro no filtro de Kalman. O processo de transformação no ETKF requer como entrada a localização e covariâncias de erros das observações. É semelhante aos ciclos de “*Bred Vector*” (KALNAY; TOTH, 1994; WEI et al., 2006) onde dois regimes paralelos criam perturbações de análise a partir de perturbações da previsão. Os valores de observação são usados apenas para computar fatores de “inflação” ((HUNT et al., 2004) e para ajustar a magnitude de perturbação da análise.

A formulação ETKF (BISHOP et al., 2001) representa as matrizes de covariância da previsão e análise por perturbações de K -previsão e K -análise (sendo K = número de membros do conjunto). Como:

$$\begin{aligned}\mathbf{X}'^b &= \frac{1}{\sqrt{K-1}}(z_1^b, z_1^f, \dots, z_k^b) \\ \mathbf{X}'^a &= \frac{1}{\sqrt{K-1}}(z_1^a, z_1^a, \dots, z_k^a)\end{aligned}\quad (3.55)$$

onde os vetores de estado de dimensão N_x . ($z_i^b = x_i^b - \bar{x}^f$) e ($z_i^a = x_i^a - \bar{x}^a$)($i = 1, 2, \dots, k$) são perturbações de previsão e de análise do k -membros do conjunto, \bar{x}^b e \bar{x}^a são as médias do conjunto de previsões e análises respectivamente e \mathbf{X}'^b e \mathbf{X}'^a são as matrizes de perturbação (WEI et al., 2006).

As matrizes de covariância da previsão e da análise (dimensão $N_x \times N_x$) são formadas como:

$$\mathbf{P}^b = \mathbf{X}'^b \mathbf{X}'^{bT} \text{ e } \mathbf{P}^a = \mathbf{X}'^a \mathbf{X}'^{aT} \quad (3.56)$$

onde o superscrito T indica matriz transposta. Para prevenir o mal-condicionamento normaliza-se o operador observação $\tilde{\mathbf{H}} = \mathbf{R}^{-1/2} \mathbf{H}$, então

$$\begin{aligned}\mathbf{P} \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H} \mathbf{P}^b \\ &= \mathbf{P}^b \mathbf{H}^T [\mathbf{R}^{1/2} (\mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}^b \mathbf{H}^T \mathbf{R}^{-1/2} + \mathbf{I}) \mathbf{R}^{1/2}]^{-1} \mathbf{H} \mathbf{P}^b, \\ &= \mathbf{P}^a \tilde{\mathbf{H}}^T (\tilde{\mathbf{H}} \mathbf{P}^b \tilde{\mathbf{H}}^T + \mathbf{I}^{m \times m})^{-1} \tilde{\mathbf{H}} \mathbf{P}^b,\end{aligned}\quad (3.57)$$

onde $\tilde{\mathbf{H}}$ é o operador observação normalizado e $(\mathbf{I}^{N_y \times N_y})$ é a matriz identidade de dimensão (N_y) , dimensão do operador observação.

Uma vez que os autovetores $\tilde{\mathbf{H}}\mathbf{P}^b\tilde{\mathbf{H}}^T$ são equivalente aos autovetores de

$$\tilde{\mathbf{H}}\mathbf{P}^b\tilde{\mathbf{H}}^T + \mathbf{I}^{N_y \times N_y} = \mathbf{E}(\mathbf{\Gamma} + \mathbf{I}^{N_y \times N_y})^{-1}\mathbf{E}^T,$$

onde p colunas de \mathbf{E} contêm o conjunto completo dos autovetores ortogonais de $\tilde{\mathbf{H}}\mathbf{P}^b\tilde{\mathbf{H}}^T$, e a matriz diagonal $\mathbf{\Gamma}$ listam os autovalores correspondentes.

A transformação do ETKF é feita a partir das perturbações da previsão nas perturbações da análise, expressas como $\mathbf{X}'^a = \mathbf{X}'^b\mathbf{T}$, transformando

$$\mathbf{P}^a = \mathbf{X}'^b\mathbf{T}\mathbf{T}^T\mathbf{X}'^b{}^T.$$

A matriz transformada \mathbf{T} é dada por:

$$\mathbf{T} = \mathbf{C}(\mathbf{\Gamma} + \mathbf{I}^{K \times K})^{-1/2}$$

onde \mathbf{C} contém a coluna ortonormal direita \mathbf{E} e $\mathbf{\Gamma}$ é a matriz diagonal de autovalores.

Comparando ETKF com EnSRF, a vantagem do ETKF é a velocidade computacional; a desvantagem é que o ETKF não aplica variâncias localizadas, que pode tornar a análise imprecisa, a menos que um grande número de conjuntos seja usado.

Os detalhes anteriores sobre algoritmos baseados em assimilação por conjunto pode parecer mais complexos do que são. Em muitas circunstâncias, o algoritmo básico é fácil de codificar. Um algoritmo para o filtro **EnSRF** acima discutido é descrito com a suposição de que o erro do modelo é tratado através da introdução de ruído de erro aditivo; e que as observações tem erros independentes, o algoritmo se modifica para cada abordagem do filtros EnKF escolhida. Os passos são os seguintes:

1. Construir um conjunto de condições iniciais arbitrárias propagadas, talvez pela amostragem aleatória da climatologia do modelo.
2. Executar o passo de previsão; integrando um conjunto de previsões de curto prazo com o próximo conjunto de observações disponíveis;
3. Realizar o passo de atualização:
 - (a) Acrescentar amostras de erro do modelo a todos os membros;
 - (b) Fazer um laço por todas as observações disponíveis;

- (i) Determinar a média do conjunto do modelo de referência [equação (3.43)] e a matriz de perturbações ao conjunto [equação (3.44)];
 - (ii) Determinar os sub-componentes do ganho de Kalman estimado [equações (3.50) e (3.51)];
 - (iii) Formar o ganho de Kalman [equação (3.48)] e reduzi-lo [equação (3.45)];
 - (iv) Atualizar a média do conjunto e as perturbações individuais [equação (3.52)];
 - (v) Definir a média do modelo de referência e das perturbações para a assimilação das próximas observações ao lado da média da análise e perturbações atualizadas;
4. Adicionar a média atualizada e as perturbações juntos para reformar o conjunto dos estados analisados;
 5. Voltar ao passo 2.

3.3.3 Localização da Covariância em Filtro de Kalman baseados em conjunto

Uma alteração algorítmica para melhorar as estimativas da matriz de covariâncias do erro do modelo é a *localização da covariância*, ou seja, a localização da covariância pode ser um aspecto prático fundamental do filtro de Kalman por conjunto. A localização restringe o uso da informação da covariância do conjunto em pequenos subconjuntos de variáveis definidas em pontos de grade do modelo de regiões locais muito menores que a superfície da Terra. A abordagem local possibilita uma estimativa de alta dimensão das estatísticas da covariância de erros do modelo de referência global com base em um pequeno conjunto.

Nos esquemas de assimilação de dados seqüenciais, as observações são tratadas uma de cada vez. A localização é feita através da atualização da análise de todos os pontos de grade que estão dentro de um raio de correlação pré-definido para cada observação. O esquema proposto por [Ott et al. \(2004\)](#) é um filtro de Kalman raiz quadrada, mas não em esquema sequencial. A idéia básica é executar a análise em cada ponto de grade simultaneamente, usando as variáveis do estado e todas as observações na região local centrada em um ponto de grade. Um modelo computacional similar foi proposto por [Anderson \(2001\)](#), enquanto [Kalnay e Toth \(1994\)](#) usaram o

método local e um método de “vetor criado” (*bred vector*) para incrementar o estado na direção determinada pela diferença entre o *bred vector* e o modelo de referência.

O esquema de assimilação de dados baseado em conjunto que utiliza covariância **localizada** é conhecido como *Local Ensemble Kalman Filter (LEKF)* (OTT et al., 2004). O LEKF calcula a análise de forma independente em cada ponto de grade, usando todas as observações disponíveis dentro de um volume local. LEKF permite uma aplicação paralela muito eficiente e permite a assimilação de um grande número de observações, tais como dados de satélites.

Em Ott et al. (2004) podem ser encontradas as características mais importantes do sistema (LEKF):

- (i) assimila todas as observações que podem afetar a análise em um determinado ponto da rede simultaneamente, e
- (ii) obtém a análise independente para cada modelo em ponto de grade.

O LEKF tem sido testado com sucesso com a implementação no *National Centers for Environmental Prediction Global Forecast System Model* (NCEP GFS), usando observações em ponto de grade simuladas (SZUNYOGH et al., 2008).

O modelo de estado da atmosfera é dado por um campo vetorial $\mathbf{x}(\tau, t)$ onde τ é bidimensional e executado sobre valores discretos τ_{mn} (a grade no espaço físico usado em cálculos numéricos, onde os dois componentes do τ são a longitude e a latitude geográfica).

Técnicas de assimilação de dados tratam $\mathbf{x}(\tau, t)$ como uma variável aleatória caracterizada pela distribuição de probabilidade. A caracterização de \mathbf{x} é atualizada no tempo de duas formas: (i) desenvolvida de acordo com a dinâmica do modelo, e (ii) modificada periodicamente de acordo com as observações atmosféricas recentes.

Cada ponto de grade \mathbf{x} contém a informação do estado do modelo sobre a área local dentro de uma sub-grade centrada em τ_{mn} . A análise é feita localmente no espaço do modelo. Considerando um conjunto de k previsões iniciadas de um passo de análise prévia, é possível construir vetores locais associados com cada membro do conjunto denotado por x_{mn}^b , onde o super-escrito b é novamente do *background* ou modelo de referência.

Nesta técnica de assimilação de dados, a matriz de covariância dos erros de modelo P_{mn}^b e o modelo de estado mais provável \bar{x}_{mn}^b são derivados de $(k' + 1)$ membros do conjunto de vetores de estado globais $(x^{b(i)}(\tau))$, $i = 1, 2, \dots, k' + 1$; $k' \geq k \geq 1$ é o índice de cada membro do conjunto global. O estado mais provável é dado por

$$\bar{x}_{mn}^b = (k' + 1)^{-1} \sum_{i=1}^{k'+1} x_{mn}^{b(i)},$$

onde $x_{mn}^{b(i)}$ é o vetor local centrado em τ_{mn} e do membro do conjunto $x^{b(i)}(\tau)$ (OTT et al., 2004).

A derivação da P_{mn}^b e das perturbações do conjunto de análises LEKF seguem o esquema de Filtro de Kalman por conjunto raiz quadrado, descrito acima. Porém estes parâmetros são executados para cada região local. O algoritmo do LEKF se modifica apenas na “localização da covariância”:

1. Avançar o conjunto de análise “globais” de estados atmosféricos para o próximo tempo de análise para obter o conjunto de previsões globais do modelo atmosféricos;
2. Associar uma região local com cada ponto de grade, para cada região local e para cada membro do conjunto do modelo referência; formar os vetores do estado da atmosfera na região local;
3. Para cada membro do conjunto local obtido no passo 2, obter a perturbação da média do conjunto do modelo, e projetar estas perturbações ao subespaço que melhor represente o conjunto na região;
4. Fazer a assimilação de dados em cada um dos subespaços locais de baixa dimensão, obtendo a média e a covariância da análise em cada região local.
5. A partir da média e covariância da análise local, obtenha um conjunto adequado de análise local dos estados atmosféricos também locais .
6. Usar as análises locais, obtidas no passo 5, para formar um novo conjunto de análises globais. (Aqui retorna a alguns passos do algoritmo do filtro raiz quadrada (EnSRF) (algoritmo 3.3.2);
7. Voltar ao passo 1.

A descrição detalhada do LEKF e do algoritmo pode ser encontrado em Ott et al. (2004).

3.3.3.1 Local Ensemble Transform Kalman Filter

Hunt et al. (2004) propôs uma série de mudanças algorítmicas para o LEKF e o esquema chamado de *Local Ensemble Transform Kalman Filter* (**LETKF**). O LETKF mantém características LEKF (acima descritas) e introduz alterações que melhoram a eficiência computacional do algoritmo e adiciona flexibilidade às observações. A mudança do nome reflete à similaridade ao *Ensemble Transform Kalman Filter* (ETKF) por Bishop et al. (2001). Os cálculos de matrizes são feitos usando o conjunto de perturbações do modelo de referência como na equação (3.56).

O LETKF descreve um meio eficaz de realizar a análise que *transforma* um conjunto do modelo de referência ($x^{b(i)} : i = 1, 2, \dots, K$) em um conjunto de análise apropriado ($x^{a(i)} : i = 1, 2, \dots, k$), utilizando a notação acima definida. Assume-se que o número de membros do conjunto k é menor que o número de variáveis do modelo N_x e que o número de observações N_y , mesmo quando a localização reduz os valores efetivos de N_x e N_y em comparação a uma análise global. A média da análise \bar{x}^a minimiza a função custo do filtro de Kalman (3.44), modificada pelo operador não linear H :

$$J(x) = (x - \bar{x}^b)^T (\mathbf{P}^b)^{-1} (x - \bar{x}^b) + [y^o - H(x)]^T (\mathbf{R})^{-1} [y^o - H(x)]. \quad (3.58)$$

Todavia, a matriz de covariância de erros do modelo de referência $\mathbf{P}^b = (K - 1)^{-1} \mathbf{X}^b (\mathbf{X}^b)^T$ tem o "rank" de no máximo $(k - 1)$ e não é inversível. $(\mathbf{P}^b)^{-1}$ é bem definida em S (matriz simétrica de \mathbf{X}^b E J passa a ser definido para $x - \bar{x}^b$ em S , e a minimização pode ser realizada neste subespaço. Como citado anteriormente, a dimensão reduzida é uma vantagem do ponto de vista da eficiência, embora a restrição da média da análise de S pode ser prejudicial se K é muito pequeno.

Para executar a análise em S é necessário escolher um sistema de coordenada apropriado. Uma abordagem natural é a dos autovetores descrita no ETKF (equação 3.3.2).

A implementação dos algoritmos acima descritos são projetados para ter eficiência e facilidade de implementação. A implementação do LETKF descreve possíveis ganhos de eficiência computacional.

Uma das premissas é fazer a localização da covariância (seção 3.3.3). Neste algoritmo usa-se cada “ponto de grade” para designar uma posição espacial ou (*local*) no modelo de previsão e obter os subconjuntos “locais” para realizar as análises, conforme Figura (3.3).

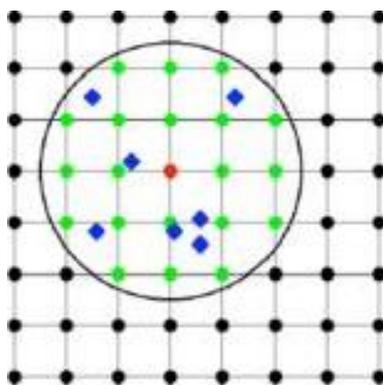


Figura 3.3 - Na abordagem **local** as observações dentro de um raio de influência são considerada. Os pontos de grade do modelo (uniformemente distribuída) e observações (irregularmente distribuídas).

Fonte: Adaptado de (KALNAY, 2008)

A localização é geralmente feita *explicitamente*, considerando apenas as observações de uma região que circunda o local da análise ou *implicitamente*, multiplicando as entradas em P^b por uma função dependente da distância que decai para zero para além de uma certa distância, de modo que as observações não afetam o estado do modelo para além dessa distância. O LETKF utiliza abordagem explícita, fazendo uma análise separada para cada ponto de grade espacial do modelo. Após a divisão em regiões locais, as análises ficam independentes podendo ser calculadas com computação paralela.

As entradas para a análise são um conjunto de modelos de referência de $(N_{x(mn)})$ dimensão de vetores do modelo de estado $(x^{b(i)mn} : i = 1, 2, \dots, k)$; o operador H_{mn} , a partir do $N_{x(mn)}$ –dimensional espaço do modelo para o $N_{y(mn)}$ –dimensional espaço de observação; um $N_{y(mn)}$ –dimensional vetor y^o de observações; e uma matriz de covariância dos erros de observação \mathbf{R}_{mn} de dimensão $(N_{y(mn)} \times N_{y(mn)})$. O subscrito mn denota que as entradas refletem o estado do modelo global e todas as observações disponíveis, a partir de um subconjunto local escolhido para a análise de cada região.

Uma abordagem razoável para selecionar observações, seria escolher as observações feitas a uma certa distância do ponto de grade em que se está fazendo a análise local e determinar empiricamente qual o valor da distância de corte que produz os “melhores” resultados.

No algoritmo abaixo os passos 1 e 2 são operações globais, mas pode ser feito localmente em uma implementação paralela. Passos de 3 a 8, devem ser realizados separadamente para cada análise local (geralmente, para cada ponto de grade). O passo 9 simplesmente combina os resultados das análises locais para formar um conjunto de análises globais ($x_{mn}^{a(i)}$), que é o resultado final da análise (HUNT et al., 2004).

1. Aplicar H_n a cada $x^{b(i)}$ para formar o conjunto global de observação de referência ($y^{b(i)}$), e a média dos últimos vetores para obter o vetor coluna de dimensão $N_y \bar{y}^b$. Subtrair esse vetor de cada $y^{b(i)}$ para formar as colunas $N_y \times K$ da matrix \mathbf{X}^b .
2. Obter a média dos vetores ($x^{b(i)}$) para obter os vetores de dimensão $N_x \bar{x}^b$ e subtrair esse vetor de cada ($x^{b(i)}$) para formar as colunas do $N_x \times K$ da matriz \mathbf{X}^b .
3. Selecionar as linhas de \bar{X}_{mn}^b , e \mathbf{X}_{mn}^b correspondentes ao ponto de grade dado, formando as homólogas locais: o $N_{x(mn)}$ -dimensional vetor \bar{x}^b e a $k \times N_x$ matriz \mathbf{X}^b , que serão utilizados no passo 8. Da mesma forma, selecione as linhas de \bar{x}_{mn}^b e \mathbf{X}_{mn}^b correspondentes às observações escolhidas para a análise no ponto de grade, formando o $N_{y(mn)}$ -dimensional vetor \bar{z}^b e a $k \times l$ matriz \mathbf{X}'^b . Selecione as linhas correspondentes y_{mn}^o e linhas e colunas de \mathbf{R}_{mn} para formar o $N_{y(mn)}$ -dimensional vetor y^o e a matriz \mathbf{R} de dimensão $N_{y(mn)} \times N_{y(mn)}$.
4. Calcular a matriz de dimensão $k \times N_{y(mn)}$ $\mathbf{C} = (\mathbf{X}'^b)^T \mathbf{R}^{-1}$. (Pode ser mais eficiente calcular \mathbf{C} pela solução de um sistema linear $\mathbf{R}\mathbf{C}^T = \mathbf{X}'^b$ ao invés de inverter \mathbf{R} .)
5. Calcule a matriz $\hat{\mathbf{P}}^a$ de dimensão $K \times K$ $\hat{\mathbf{P}}^a = \mathbf{X}'^b \mathbf{T} \mathbf{T}^T \mathbf{X}'^{bT}$ (equação 3.3.2).

6. Calcule a matriz de ganho de dimensão $k \times k$

$$\mathbf{W}^a = [(k - 1)\hat{\mathbf{P}}^a]^{1/2}. \quad (3.59)$$

7. Calcule o vetor K -dimensional $\bar{\mathbf{w}} = \hat{\mathbf{P}}^a \mathbf{C}(y^o - \bar{z}^b)$, e adicione o resultado a cada coluna de \mathbf{W} , formando uma matriz $(k \times k)$ cujas colunas são os vetores de análise ($\mathbf{w}^{a(i)}$).

8. Multiplique \mathbf{X}^b por cada $\mathbf{x}^{a(i)}$ e adicione \bar{x}^b para obter o conjunto de análises dos membros ($x^{a(i)}$).

9. Após executar os passos 3 a 8 para cada ponto de grade, as saídas do passo 8 formam o conjunto de análises globais ($x^{a(i)}$).

O Centro de Previsão de Tempo e Estudos Climáticos – **CPTEC/INPE** visando aprimorar cada vez mais suas análises, está implantando o sistema de assimilação de dados baseado no LETKF que deverá se tornar operacional ([HERDIES et al., 2007](#); [ARAVEQUIA, 2008](#)). Atualmente, está sendo testado em modo de pesquisa, e diversos artigos sobre o desempenho do sistema e suas vantagens são descritos em ([KALNAY, 2007](#)).

Neste capítulo, descreveu-se os conceitos e técnicas para assimilação de dados baseadas em Filtro de Kalman e previsão por conjunto, que são métodos atuais e do estado da arte. Como citado, alguns métodos estão sendo testados em centros operacionais de previsão de tempo por produzirem análises mais adequadas.

O método de assimilação LETKF foi o método escolhido para ser utilizado nesta tese para ser emulado pela Rede Neural Artificial. As análises geradas pelo LETKF já foram avaliadas pela comunidade meteorológica e atestada como de boa qualidade.

4 REDES NEURAIIS ARTIFICIAIS

Este capítulo aborda as principais características de Redes Neurais Artificiais(RNA). Devido a extensão do assunto, as características e conceitos aqui descritos são pertinentes à aplicação do problema de assimilação de dados atmosféricos. Os conceitos apresentados neste capítulo são apenas para referenciar o desenvolvimento das Redes Neurais Artificiais dentro do contexto do trabalho multidisciplinar, computação e meteorologia.

4.1 Conceitos

O estudo das redes neurais artificiais (RNA) é algo fascinante e esse fascínio aumenta à medida que se tem mais conhecimento sobre o assunto. Trata-se de um conceito de extrema importância na computação, responsável pela solução de muitos problemas complexos, inclusive solução de problemas inversos. Redes Neurais Artificiais são baseadas no processamento de dados semelhante ao executado pelo cérebro humano. O cérebro é tido como um processador altamente complexo que realiza processamentos de maneira paralela. Para isso, o cérebro tem uma estrutura organizada, ou seja, neurônios conectados, de forma que realizam o processamento necessário. Isso é feito numa velocidade extremamente alta.

Nas redes neurais artificiais, a idéia é realizar o processamento de informações tendo como princípio a organização de neurônios do cérebro. Como o cérebro humano é capaz de aprender e tomar decisões baseadas na aprendizagem, as redes neurais artificiais procuram fazer o mesmo. Assim, uma rede neural artificial pode ser interpretada como um esquema de processamento capaz de armazenar conhecimento baseado em aprendizagem (experiência) e disponibilizar este conhecimento para a aplicação em questão. Ou seja, RNA são técnicas computacionais que apresentam um modelo inspirado na estrutura neural de inteligência humana.

RNA são sistemas paralelos distribuídos, compostos por unidades de processamentos simples (*nós*) que computam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões. Uma RNA pode ser melhor definida como um grafo direcional, que possui um conjunto de pontos, ao longo de um conjunto de segmentos de linhas direcionadas entre eles. Os nós deste grafo são chamados unidades de processamento ou *neurônios*. Suas arestas são conexões (ou *sinapses*), que funcionam

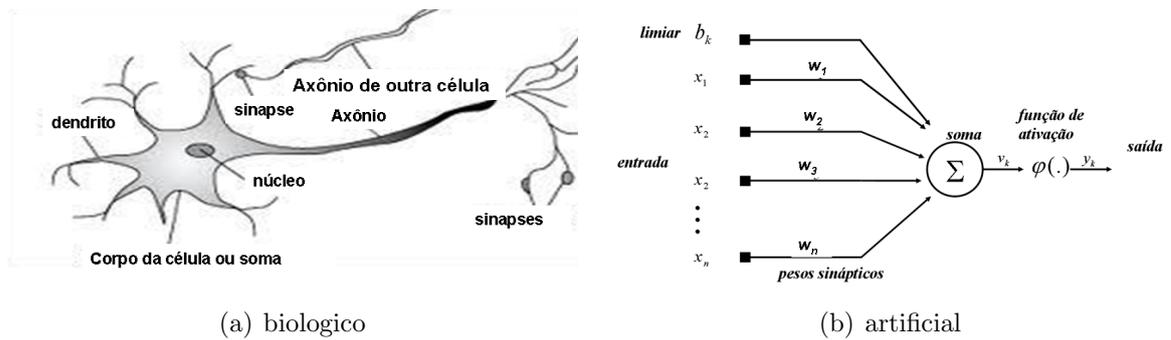


Figura 4.1 - Representação de um neurônio

como caminhos de sinais em uma única direção, de forma que suas unidades de processamento possam receber qualquer número de conexões de entrada. Sinais de entrada em uma rede neural chegam através das conexões de entrada. O resultado do processamento é gerado a partir das unidades de saídas da rede.

4.1.1 Neurônio Biológico e Neurônio Artificial

Nesta seção apresenta-se a analogia entre o neurônio biológico e o neurônio artificial ilustrados na figura 4.1.

O neurônio biológico é uma célula que pode ser dividida em três seções: o corpo da célula, o dendrito e o axônio, cada uma com funções específicas, porém complementares, ilustrado na figura 4.1(a). O corpo do neurônio mede apenas alguns milésimos de milímetros e o dendrito poucos milímetros de comprimento. O axônio pode ser mais longo. O dendrito tem como função, receber informações, ou impulsos nervosos, oriundas de outros neurônios e conduzi-las até o corpo celular. Ali, a informação é processada e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurônios (nós), passando pelo axônio e atingindo o dendrito dos neurônios seguintes.

O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado *sinapse*. É pela sinapse que os neurônios se unem funcionalmente, formando as redes neurais. A sinapse funciona como válvulas, sendo capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os neurônios na rede neural. O efeito da sinapse é variável, é esta variação que dá ao neurônio capacidade de adaptação.

Os sinais oriundos dos neurônios *pré-sinápticos* são passados para o corpo do neurônio, onde são comparados com outros sinais recebidos pelo neurônio. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nós pós-sinápticos). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. A capacidade de solucionar funções complexas surge com a operação em paralelo de todos os 10^{11} neurônios do nosso cérebro (KOVACS, 1996).

O neurônio artificial (Figura 4.1(b)) possui uma analogia com o neurônio biológico. Segundo Haykin (2001), um neurônio artificial é uma unidade de processamento de informações. McCulloch e Pitts (1943) definiram o modelo básico de um neurônio, que é uma simplificação do que se sabia à respeito do neurônio biológico naquela época. A sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n (que representam os dendritos) e apenas um terminal de saída y (que representa o axônio).

Para emular o comportamento das “sinapses”, os terminais de entrada do neurônio artificial têm *pesos* acoplados w_1, w_2, \dots, w_n cujo valores podem ser positivos ou negativos. O efeito de uma sinapse particular no neurônio i pós-sináptico é dado por: $x_i w_i$. Os pesos determinam “em que grau” o neurônio deve considerar sinais de disparo que ocorrem naquela conexão.

Um neurônio dispara quando a *soma* dos impulsos ultrapassa o seu limiar. O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos pelo neurônio (soma ponderada). E decide se o neurônio deve ou não disparar (saída igual a 1 ou a 0) comparando a soma obtida ao limiar do neurônio. A ativação do neurônio é obtida através da aplicação de uma “*função de ativação*”, que ativa a saída ou não, dependendo do valor da soma ponderada das suas entradas. O neurônio de McCulloch e Pitts terá então sua saída ativa quando:

$$\sum_{i=1}^n x_i w_i \geq \theta$$

onde n é o número de entradas do neurônio, w_i é o peso associado à entrada x_i , e θ o limiar do neurônio.

Conforme Figura (4.2), a estrutura do modelo neuronal ou neural, tem o neurônio i marcado para efeitos de referência. O modelo consiste em dois blocos funcionais:

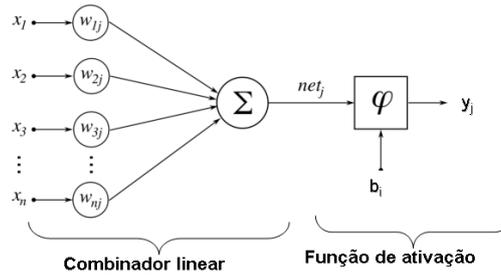


Figura 4.2 - Modelo simplificado de um neurônio marcado como i

1. Um *combinador linear*, que consiste em uma série de pesos (parâmetros ajustáveis) ligado aos terminais de entrada, para simplificar problemas, o peso conectado ao terminal j é representado por uma seta dirigida rotulada com o símbolo w_{ij} . O combinador linear inclui também um viés denotado por b_i que pode ter um valor positivo ou negativo.
2. Uma *função de ativação*, que segue o combinador linear. A função de ativação, denotada por φ é ao mesmo tempo não-linear e sem memória.

A partir do modelo proposto por McCulloch e Pitts foram derivados vários outros modelos que permitem a produção de uma saída qualquer e com diferentes funções de ativação.

As RNA são definidas de acordo com a sua arquitetura, sendo que existem modelos que possuem apenas uma única camada, outras já possuem várias camadas em sua arquitetura.

Concluindo, uma unidade de processamento em uma rede neural é uma combinação linear de várias entradas e pesos, seguido por uma função de ativação. O i -ésimo neurônio apresenta o acoplamento de duas equações:

$$net_i = \sum_{j=1}^L w_{ij}x_j, \quad j = 1, \dots, L, \quad i = 1, \dots, M, \quad (4.1)$$

$$y = \varphi(net_i) + b_i \quad (4.2)$$

em que x_1, \dots, x_n são sinais de entrada; N o tamanho do vetor de entrada; w_{i1}, \dots, w_{ij} são os pesos sinápticos do neurônio i na camada j ; net_i é a saída do “combinador” linear devido aos sinais de entrada; L é o número de camadas descon-

didadas, M , o número de neurônios em cada camada; b_i é o viés; $\varphi(\cdot)$ é a função de ativação e y é o sinal de saída. O uso do viés b_i tem o efeito de aplicar uma transformação afim à saída net_i do “combinador” linear do modelo da figura (4.1(a)). No modelo de um neurônio artificial, (Figura 4.1(b)), o limiar b_i tem o efeito de acrescentar um grau de liberdade a cada neurônio.

4.1.2 Funções de ativação

A função básica do neurônio artificial, ou elemento processador, é realizar o somatório, ponderado por pesos sinápticos, dos elementos do vetor de entrada e aplicar este resultado como entrada de uma função não linear denominada *função de ativação*.

A função $\varphi(\cdot)$ representa a função de ativação, a qual simula características não lineares de um neurônio biológico. A Figura (4.3) ilustra, graficamente, quatro funções de ativação diferentes: a função linear, a função degrau, a função sigmoideal logística, e a função sigmoideal tangente hiperbólica.

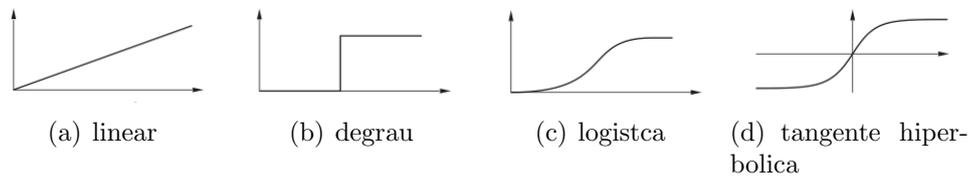


Figura 4.3 - Funções de ativação mais utilizadas.

A função de ativação *linear*, mostrada na Figura 4.3(a), é definida pela equação

$$y = \alpha x$$

onde α é um número real que define a saída linear para os valores de entrada, y é a saída e x é a entrada. Os valores da função linear pode ser restringido para produzir valores constantes em uma faixa $[-\gamma, +\gamma]$, e neste caso, passa a ser a função **rampa**

definida pela notação abaixo.

$$y = \begin{cases} +\gamma & \text{se } x \geq \gamma \\ x & \text{se } x = \gamma \\ -\gamma & \text{se } x \leq \gamma \end{cases}$$

Os valores máximo e mínimo são $-\gamma$, $+\gamma$. A função *rampa* é usada geralmente, como uma função linear simplificada.

A função *degrau*, ilustrada na Figura 4.3(b), é similar a uma função rampa, pois a função produz a saída $+\gamma$ para valores de x maiores que zero, caso contrário a função produz o valor de $-\gamma$, definida por:

$$y = \begin{cases} +\gamma & \text{se } x > 0 \\ -\gamma & \text{se } x \leq 0 \end{cases}$$

A função *sigmoidal* : é a função de ativação mais utilizada em redes neurais artificiais. Definida como uma função que apresenta propriedades assintóticas e de suavidade. Uma função sigmoidal muito comum em modelos neurais é a função *logística* (Figura 4.3(c)) definida por :

$$\varphi(net_i) = \frac{1}{1 + e^{-a(net_i)}} \quad (4.3)$$

onde a é o parâmetro de inclinação da função. E o termo net_i conforme equação (4.1) onde x_i é o sinal aplicado ao i -ésima entrada. A função logística é do tipo “unipolar”, onde a saída do modelo é não negativa, varia entre 0 e 1 . Alternativamente, para se obter uma saída tipo “bipolar”, permitindo valores positivos e negativos, deve-se usar a função sigmoidal, que varia entre -1 e 1, a função tangente hiperbólica (Figura 4.3(d)) dada por:

$$\varphi(net_i) = \tanh\left(\frac{1}{2}net_i\right) \quad (4.4)$$

$$= \frac{1 - e(-net_i)}{1 + e(-net_i)} \quad (4.5)$$

O valores nas entradas devem ser normalizados para valores restritos ao intervalo da função de ativação utilizada e, posteriormente, deve ser feito o processo inverso

para que as variáveis estimadas voltem ao intervalo da função original.

4.2 Características Redes Neurais Artificiais

Uma rede neural artificial, segundo Fausett (1994), é caracterizada por:

- (i) O padrão de conexão entre os neurônios, o que determina a arquitetura da rede;
- (ii) O método de determinação dos pesos sobre as conexões (chamado de treinamento ou algoritmo de aprendizagem);
- (iii) A função de ativação.

Uma das arquiteturas de RNA muito utilizadas, é a rede *multicamada* onde os neurônios são organizados em duas ou mais camadas. As entradas são conectadas aos elementos processadores básicos, que são por sua vez são inter-conectados com elementos de outras camadas e/ou a saída da rede.

A *camada de entrada* é especial, pois não realiza qualquer processamento. Ela apenas distribui os valores de entradas para todos os neurônios da primeira camada de processamento. A camada (ou camadas) intermediária é chamada *camadas escondida* e os neurônios na saída da rede são chamados de *camada de saída*.

A RNA que as entradas projetam-se diretamente sobre a camada de saída é chamada de *feed-forward* ou rede “direta”. Se há uma realimentação entre uma das camadas da rede, esta é chamada de rede “recorrente”.

O processamento de uma RNA é feito em duas etapas: a etapa de *aprendizagem*, ou treinamento, e a etapa de *ativação*. No processo de aprendizagem, um conjunto de exemplos são apresentados à rede e esta extrai as características necessárias para representar a informação fornecida. Essas características são armazenadas em valores (os *pesos sinápticos*), que são utilizados posteriormente no processo de ativação da RNA, que com outras entradas reproduzem respostas baseadas nas características armazenadas.

4.2.1 Redes Perceptron Multicamadas (PMC)

Um modelo de rede neural multicamadas simples é aquele em que a camada de entrada se conecta a camada de saída. Este modelo, com apenas um neurônio na camada de saída recebe a denominação de *Perceptron*. Foi objeto de intensa pesquisa

nas décadas de 50 e 60, porém o modelo Perceptron teve uso limitado. O trabalho original que enfocou o modelamento de um neurônio e sua capacidade computacional de solucionar problemas, teve início com as pesquisas de [McCulloch e Pitts \(1943\)](#). Naquela época surgiram os primeiros computadores digitais e a associação de um modelo artificial de um neurônio com a capacidade de resolver funções Booleanas estava ligada ao pensamento da época, de que era possível construir uma máquina inteligente através de operadores lógicos básicos.

O trabalho de Frank Rosenblatt ([ROSEMBLATT, 1958](#)) introduziu-se o conceito de “aprendizado” em RNA e um modelo composto por uma estrutura de rede tendo como unidades básicas os neurônios e uma regra de aprendizado.

Apesar de ter causado grande euforia na comunidade científica da época, o perceptron não teve vida muito longa, já que duras críticas de [Minsky e Papert \(1969\)](#) foram feitas em relação à sua capacidade computacional, causando, assim, um grande impacto sobre as pesquisas em RNA, fez com que o financiamento para pesquisa na área cessassem durante os anos 70 e início dos anos 80.

A pesquisa nesta área mudou e investimentos voltaram, devido as descrições da rede de [Hopfield \(1982\)](#) e do algoritmo de *retro-propagação de erros* ou *back-propagation*, que será descrito nas próximas seções. Foi em consequência destes trabalhos que a área de RNA ganhou novo impulso, ocorrendo a partir dos anos 80, uma forma de expansão no número de trabalhos de aplicação e teóricos envolvendo RNA e técnicas correlatas.

A interconexão das entradas à camada de saída por pelo menos uma camada de neurônios intermediária, tipicamente denominada de camada escondida com o uso de função de ativação não linear, remove a *limitação* do Perceptron e amplia o leque de problemas solucionados por estas redes.

O *Perceptron Multicamadas* (PMC) é uma rede neural que consiste de uma camada de entrada de nós fontes, uma ou mais camadas escondidas de nós computacionais (neurônios) e uma camada de saída também com nós computacionais. Os nós fontes providenciam os pontos de acesso físico da aplicação com sinais de entrada. Os neurônios na camada escondida agem como “detectores de características”; este neurônio são referidos como escondidos porque são fisicamente inacessíveis às entradas ou às saídas da rede. Finalmente, os neurônios da camada de saída apre-

sentam ao usuário as conclusões encontradas pela rede em resposta aos sinais de entrada (HAYKIN, 2007).

A Figura (4.4) representa um esquema típico de uma rede perceptron múltiplas camadas. Com quatro nós fonte, uma camada com cinco neurônios escondidos e um neurônio de saída.

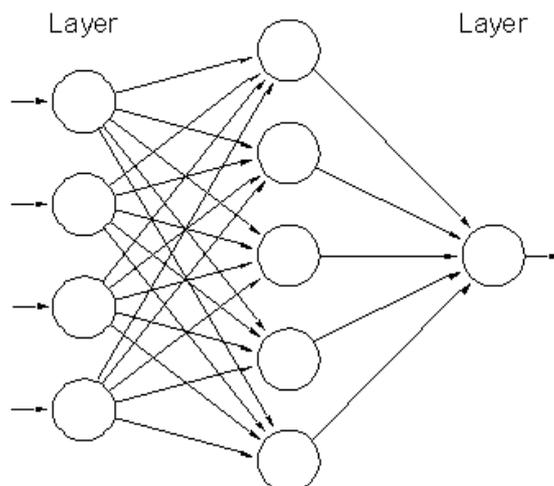


Figura 4.4 - Perceptron multicamadas com uma camada escondida

Duas características desta estrutura são aparentes:

1. - Uma PMC é uma rede direta (*feedforward*), no sentido de que os sinais de entrada produzem uma resposta na(s) saída(s) da rede pela propagação na direção para frente apenas.
2. - A rede deve ser completamente conectada, como na figura 4.4, em que cada nó de uma camada da rede é conectada a todos outros nós na camada adjacente. Alternativamente, a rede pode ser apenas parcialmente conectada, em que alguns dos links de conexão que pode estar faltando. Neurônios conectados localmente representam um tipo importante de rede parcialmente conectada, o “local” indica que um neurônio em uma camada da rede está ligado somente a um subconjunto de possíveis entradas.

A arquitetura da rede PMC apresenta:

- O modelo de cada neurônio ou elemento processador da rede possui uma função de ativação não-linear. A função “sigmoïdal” atende esta exigência e é muito utilizada.
- A rede possui pelo menos uma camada oculta de processamento com neurônios que não fazem parte da entrada ou da saída.
- A rede possui alto grau de conectividade entre seus elementos processadores. Esta conectividade é definida pelos pesos sinápticos.

Definido um vetor de entradas $x = [x_0, x_1, x_2, x_3, \dots, x_N]$ e um vetor de saídas $y = [y_0, y_1, y_2, \dots, y_L]$, uma rede neural multicamadas realiza um mapeamento complexo $y = \Psi(w, x)$, da entrada x na primeira camada, a saída y parametrizado pelos pesos sinápticos w . Redes multicamadas são, portanto, ferramentas poderosas. Dado um número suficiente de neurônios, conseguem aproximar “qualquer” função contínua (suave).

O número de nós fonte na camada de entrada é determinado pela dimensionalidade do espaço de observações (dados) que é responsável pela geração dos sinais de entrada. O número de nós computacionais na camada de saída é determinado pela dimensionalidade exigida pela resposta desejada. Então o projeto de uma PMC requer três premissas: (i) determinação do número de camadas escondidas; (ii) determinação do número de neurônios em cada camada; (iii) a especificação dos pesos sinápticos que inter-conectam os neurônios em diferentes camadas da rede.

O ítem (i) e (ii) se referem à complexidade de uma PMC. Para resolver o ítem (iii), pode-se usar o algoritmo de aprendizagem de “retropropagação” do erro.

4.3 Aprendizagem em Redes Neurais

RNA possuem a capacidade de aprender através exemplos e fazer interpolações e extrapolações do que aprenderam. A capacidade de *aprender* associada a uma rede neural é uma das mais importantes qualidades destas estruturas. Considera-se “aprendizado” o processo que adapta o comportamento e conduz a uma melhoria de desempenho. O aprendizado procura obter regras para determinar a intensidade das conexões entre os neurônios. Um conjunto de procedimentos bem definidos para adaptar os parâmetros de uma RNA, afim de que a mesma possa aprender uma determinada função é chamado de *algoritmo de aprendizado*.

Existem diversos algoritmos com suas vantagens e desvantagens. Estes algoritmos,

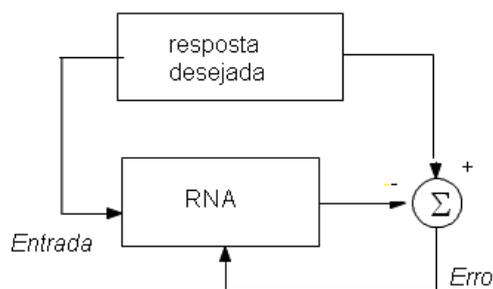


Figura 4.5 - Esquema do treinamento supervisionado

basicamente, diferem pela maneira através da qual o ajuste dos pesos é feito. A utilização de uma RNA na solução de uma tarefa passa, inicialmente, por uma fase de aprendizagem, criando assim uma representação própria para o problema.

A etapa de aprendizagem consiste em um processo iterativo de ajuste de parâmetros da rede, os pesos das conexões entre as unidades de processamento, guardam o conhecimento que a rede adquiriu do ambiente em que está operando (MENDEL; MCLAREN, 1970)

Diversos métodos para treinamento de redes foram desenvolvidos, podendo estes serem agrupados em dois paradigmas principais: Aprendizado Supervisionado e Aprendizado Não Supervisionado.

- a) **Aprendizado Supervisionado** Este método de aprendizado é o mais comum no treinamento das RNA. É chamado aprendizado *supervisionado* porque a “resposta desejada” ou “alvo” que a rede deve resultar é fornecida por um *supervisor* externo. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída desejada fornecidas.

A rede tem sua saída corrente (calculada) comparada com a saída desejada, recebendo informações do supervisor sobre o erro da resposta atual. A cada padrão de entrada submetido à rede compara-se a resposta desejada com a resposta calculada e os pesos das conexões são ajustados para minimizar o erro. A minimização da diferença é incremental, já que pequenos ajustes são feitos nos pesos à cada etapa de treinamento, de tal forma que estes caminhem para uma solução.

A desvantagem do aprendizado supervisionado é que, na ausência do supervisor, a rede não conseguirá aprender novas estratégias para situações não cobertas pelos exemplos do treinamento da rede. Os exemplos mais conhecidos de algoritmos de aprendizado supervisionado são a *regra delta* ou *correção de erros*. Para redes de múltiplas camadas, o algoritmo "retropropagação de erros" ou (*back-propagation*) (WIDROW; HOFF, 1960).

Correção de Erros é o processo de aprendizado por correção de erros procura minimizar a diferença entre a soma ponderada das entradas pelo pesos (saída calculada pela rede) e a saída desejada, ou seja, o erro da resposta atual da rede.

O termo $e(t)$ denota o erro e deve ser calculado através da seguinte expressão: $e(t) = d(t) - y(t)$, onde $d(t)$ é a saída desejada e $y(t)$ é a resposta atual (calculada) da rede no instante t . O ajuste dos pesos deve ser proporcional ao produto do erro pelo valor de entrada da sinapse naquele instante. Este método é utilizado para treinamento e posterior generalização no algoritmo de *retropropagação* (RUMELHART et al., 1986), que será descrito na próxima seção. O método de aprendizado por correção de erros também pode ser chamado de *regra delta* de aprendizado.

- b) **Aprendizado não supervisionado** - é o aprendizado onde não há um supervisor (resposta desejada) para acompanhar o processo de aprendizado. Para este tipo de aprendizado, somente os padrões de entrada estão disponíveis para rede, ao contrário do aprendizado supervisionado, cujo conjunto de treinamento possui pares de entrada e saída. A partir do momento em que a rede estabelece uma harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente. Este tipo de aprendizado, só se torna possível, quando existe redundância nos dados de entrada. Sem redundância seria impossível encontrar quaisquer padrões ou características dos dados de entrada.

Uma das vantagens do aprendizado não supervisionado pode ser a ausência da indicação responsável pela resposta desejada. A desvantagem está ligada ao conjunto de dados para treinamento, que deve ser redundante para que a rede consiga abstrair características em seu treinamento.

Algoritmo como o “*Hebbiano*” e “*Competitivo*” são exemplos de algoritmos não supervisionados e são especificados em Fukushima (1975), Grossberg (1998), Hebb (1949), Kohonen (1982).

Na fase de treinamento da RNA, adotar um critério de parada é importante. Normalmente, a rede é treinada até que um número máximo de “*épocas*” de treinamento, definido anteriormente seja atingido. Uma *época de treinamento* é a apresentação de todo um conjunto (ou padrões) de treinamento da rede. O critério de parada em função do *menor erro objetivo* pode ser aplicado com o paradigma de aprendizagem supervisionada. Uma definição completa dos algoritmos de aprendizagem, no contexto de redes neurais está disponível em (HAYKIN, 2001)

4.3.1 Algoritmo de Retropropagação do Erro

O algoritmo de retropropagação (**RP**) é o processo supervisionado mais utilizado para treinamento de redes neurais PMC, este algoritmo é uma forma alternativa do algoritmo de mínimos quadrados com a escolha de uma função de ativação não linear apropriada. O algoritmo retropropagação utiliza pares (entrada, saída desejada) para, através de um mecanismo de correção de erros, ajustar os pesos da rede.

O treinamento ocorre em duas fases, onde cada fase percorre a rede em um sentido. Estas duas fases são chamadas de fase *forward* e fase *backward*.

Na fase *forward* (para frente), os pesos sinápticos são fixos, e a resposta da rede é calculada por uma conjunto prescrito dos sinais de entrada. A fase para frente do algoritmo RP é análogo ao processo de filtragem do algoritmo de mínimos quadrados (HAYKIN, 2007)

Na fase *backward* (retorno), os ajustes dos pesos sinápticos são calculados com o propósito de minimizar uma função custo definida como a soma dos erros quadráticos (4.9), descrita a seguir. Inicia-se pelo cálculo do erro do sinal na camada de saída e então retorna através da rede, camada por camada, até que a rede toda é coberta. O algoritmo RP deriva seu nome da natureza de retorno do cálculo do erro envolvido na sua implementação. A fase de retorno no algoritmo é análoga ao processo adaptativo do algoritmo de mínimos quadrados (HAYKIN, 2007).

O RP é baseado na regra proposta por Widrow e Hoff (1960), sendo também chamada de *regra delta generalizada*. Este algoritmo propõe uma forma de definir

o erro dos neurônios das camadas intermediárias, possibilitando o ajuste de seus pesos. A derivação da regra delta generalizada é simples e semelhante à derivação da regra delta.

Em uma perceptron multicamadas com L camadas de processamento e N padrões de entrada, a iteração (passo de tempo) n é a apresentação do n -ésimo exemplo de treinamento ao neurônio j . O relacionamento entrada-saída é caracterizada pela equação não linear:

$$y_j(n) = \varphi \left(\sum_{i=1}^L w_{ji}(n)x_i(n) + b_j(n) \right). \quad (4.6)$$

que representa o neurônio j alimentado por um conjunto de sinais funcionais produzidos pela camada i , os sinais se propagam da esquerda para a direita através da rede. O campo local induzido $net_j(n)$ produzido na entrada da função de ativação associada ao neurônio j . O parâmetro b_i é o termo de viés, equivalente ao peso fixo conectado por uma entrada igual a constante $+1$. A equação (4.6) deve ser generalizada para todas as unidades da PMC.

O *signal de erro* é definido como a diferença entre a resposta desejada e a saída corrente da rede. Especificamente, para o i -ésimo neurônio:

$$e_j(n) = d_j(n) - y_j(n), \quad (4.7)$$

onde $d_j(n)$ é a resposta desejada no j -ésimo nó da camada de saída, $y_j(n)$ é a saída calculada pela rede. Define-se o erro para o neurônio j como $\frac{1}{2}e_j^2(n)$. O erro total de todos os neurônios da camada de saída $\xi(n)$ produzido pela soma dos erros quadráticos:

$$\xi(n) = \frac{1}{2} \sum_{j=1}^L e_j^2(n) \quad (4.8)$$

Esta soma que inclui todos os neurônios da camada de saída da rede, definem a função custo normalizada em relação ao conjunto de exemplos (padrões) incluídos no treinamento de tamanho N :

$$\xi_{med} = \frac{1}{N} \sum_{n=1}^N \xi(n) \quad (4.9)$$

O algoritmo retropropagação (RP) minimiza a função custo (4.9) recursivamente

ajustando os pesos da PMC, usando uma aproximação da técnica do *gradiente descendente*. Esta minimização é uma aproximação do métodos dos mínimos quadrados (HAYKIN, 2001).

Aplica-se uma correção $\Delta w_{ij}(n)$ aos pesos definida pela *regra delta*. A equação de atualização dos pesos é:

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n), \quad (4.10)$$

Os pesos são alterados na proporção negativa do gradiente. O termo de atualização é definido por:

$$\Delta w_{ij}(n) = -\eta \nabla_{w_{ij}} \xi(n), \quad (4.11)$$

onde η é um parâmetro taxa de aprendizagem e $\nabla_{w_{ij}} \xi(n)$ é o gradiente da função custo $\xi(n)$ com respeito aos pesos w_{ij} . Primeiramente calcula-se a derivada parcial de $\xi(n)$. De forma similar ao algoritmos de mínimos quadrados, o algoritmo de retropropagação aplica a correção $\Delta w_{ij}(n)$ ao peso sináptico $w_{ij}(n)$, que é proporcional à derivada parcial $\partial \xi(n) / \partial w_{ij}(n)$, que corresponde à:

$$\Delta w_{ij}(n) = \eta \delta_j(n) y_i(n), \quad (4.12)$$

onde o *gradiente local* $\delta_j(n)$ de acordo com a regra da cadeia do cálculo, pode ser expresso como:

$$\begin{aligned} \delta_j(n) &= -\frac{\partial \xi(n)}{\partial net_j(n)} \\ &= -\frac{\partial \xi(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial net_j(n)} \frac{\partial net_j(n)}{\partial w_{ij}(n)} \end{aligned} \quad (4.13)$$

$$= e_j(n) \varphi'_j(net_j(n)) \quad (4.14)$$

$\delta_j(n)$ para o neurônio de saída j é igual ao produto do sinal de erro $e_j(n)$ correspondente para aquele neurônio pela derivada $\varphi'_j(net_j(n))$ da função de ativação associada.

A saída calculada pela rede $y_j(n)$ é dada por

$$y_j(n) = \varphi(net_j(n)) \quad (4.15)$$

onde $net_j(n)$ é dada pela equação (4.1) e se for a ultima camada L tem-se que

$$net_j(n) = \sum_{i=0}^N w_{iL}(n)y_j(n) \quad (4.16)$$

A função de ativação $\varphi(net_i)$ pode ser limitada, diferenciável, função monotonicamente crescente. Por exemplo, a função logística, se esta for a função sigmoidal escolhida.

Quando o neurônio j está localizado em uma camada oculta da rede, não existe uma resposta desejada para o neurônio, considere que o neurônio j como nó oculto da rede e o gradiente local $delta_j(n)$ para este neurônio, como vimos anteriormente:

$$\begin{aligned} \delta_j(n) &= - \frac{\partial \xi(n)}{\partial y(n)} \frac{\partial y_j(n)}{\partial net_j(n)} \\ &= \frac{\partial \xi(n)}{\partial y_j(n)} \varphi'_j(net_j(n)) \end{aligned} \quad (4.17)$$

Para a camada de saída da PMC, utiliza-se o índice k ao invés do índice j .

$$\delta_j(n) = \varphi'(net_j(n)) \sum_k \delta_k(n) w_{kj}(n), \text{ o neurônio } j \text{ é oculto e } k \text{ o último.} \quad (4.18)$$

onde o uso do apóstrofe (do lado direito) significa diferenciação. Outra notação usada para não confundir a computação da última camada é o índice l . Se o neurônio j está na camada de saída, ou seja, $l=L$, onde o L é denominado profundidade da rede. A variável δ é interpretada como um erro retro-propagado, que é recursivamente calculado para cada camada da PMC, começando pela camada de saída, a camada L (HAYKIN, 2007).

A derivação do algoritmo RP é geralmente apresentado para dados em valores reais e pode ser encontrado em Rumelhart et al. (1986), Haykin (2001).

A regra delta generalizada requer que as funções de ativação utilizadas pelos nós da rede sejam contínuas, diferenciáveis, geralmente não decrescente da entrada total recebida pelo neurônio. A equação abaixo ilustra o cálculo do valor de ativação.

$$y_j^p = f_j(net_j^p)$$

Tabela 4.1 - Resumo do Algoritmo de Retro-Propagação do erro

Passos do algoritmo

1. *Iniciação*

Atribua valores reais aleatórios a todos os pesos e vieses;

2. *Apresentação das entradas e saídas desejadas*

Apresente o vetor de entrada $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(N)$

e o correspondente vetor de resposta desejada $\mathbf{d}(1), \mathbf{d}(2), \dots, \mathbf{d}(N)$,

um par a cada instante, onde N é o número de padrões de treinamento;

3. *Cálculo das saídas correntes*

Use a equação (4.3) ou a equação (4.4) para calcular

os sinais de saída y_1, y_2, \dots, y_N ;

4. *Adaptação de pesos e vieses*

$$\Delta w_{ij}^{(l-1)}(n) = \eta x_i(n) \delta_j^{(l-1)}(n)$$

$$\Delta b_j^{(l-1)}(n) = \eta \delta_j^{(l-1)}(n)$$

$$\delta_j^{(l-1)}(n) = \begin{cases} \varphi'(net_j)[d_j - y_j(n)]^L, & \text{para o neurônio } j \text{ da camada de saída } L \\ \varphi'(net_j) \sum_k w_{ji}^{(l+1)}(n) \delta_j^{(l+1)}(n), & \text{para o neurônio } j \text{ na camada oculta } l \end{cases}$$

5. *Iteração*

no qual $x_i(n)$ = saída do nó i na iteração n .

Fonte: (HAYKIN, 2007)

onde:

$$net_j^p = \sum_{i=1}^p x_i w_{ji} + b_i$$

p é o número de padrões, ou seja, representa o número de conexões de entrada do neurônio i e w_{ij} o peso da conexão entre a entrada e o neurônio da camada j . Segundo Haykin (2007), uma rede Perceptron multicamadas treinada com o algoritmo de retropropagação fornece um poderoso instrumento para aproximar um mapeamento de entrada-saída não-linear de caracter geral. Neste contexto, uma questão-chave é o número de camadas ocultas necessárias no projeto do perceptron multicamadas para fazer uma aproximação de maneira uniforme. O *teorema da aproximação universal* (CYBENKO, 1989) pode resolver esta questão.

O teorema de aproximação universal é diretamente aplicável à PMC com as seguintes

características:

- uma camada de entradas de nós N_0 cujas entradas (sinais fonte) são denotados por x_1, x_2, \dots, N ;
- uma única camada escondida de N_1 neurônios sigmoidais, com os pesos do i -ésimo neurônio escondido denotado por $w_{i1}, w_{i2}, \dots, w_{iN_0}$;
- uma camada de saída consistindo de um único neurônio linear.

Nota-se que o teorema de aproximação universal é um teorema de existência, no sentido em que proporciona a justificação matemática para a aproximação de uma função arbitrária contínua, ao contrário de uma representação exata. O teorema afirma que um perceptron de múltiplas camadas com uma única camada intermediária é capaz de realizar uma aproximação uniforme, dado um conjunto de treinamento suficientemente significativo para representar a função. Por outro lado, o teorema não afirma que um PMC com uma única camada é ótimo no sentido de tempo de processamento, facilidade de implementação e eficiência na representação (HAYKIN, 2007).

4.4 Aplicabilidade das Redes Neurais Artificiais

Aplicações de redes neurais são inúmeras. A literatura científica associada a pesquisas que utilizam RNA em aplicações em diversas áreas da ciência apresentam uma ampla aplicabilidade e mostram que a técnica de redes neurais é uma boa alternativa para solução de problemas. Problemas como reconhecimento de padrões, aproximação de funções, previsão de séries temporais, mineração de dados, análise e processamento de sinais, controle de processos, robótica, classificação de dados, filtros contra ruídos eletrônicos, análise de imagens e voz, avaliação de crédito, diagnóstico médico e muitos outros.

A maioria dos trabalhos descreve brevemente o funcionamento de redes neurais em determinada aplicação. A escolha da abordagem mais adequada depende da complexidade do problema a ser tratado e do grau em que o problema é entendido. Assumindo dados suficientes, recursos de computação e se o teórico entendimento do problema estiver disponível, um modelo numérico completo, talvez seja a solução mais desejável.

No entanto, em geral, quando a complexidade de um problema aumenta a compreen-

são teórica diminui, abordagens estatísticas passam a ser necessárias. O uso de redes neurais, em particular o perceptron multicamadas, foi visto como uma alternativa eficaz a técnicas estatísticas mais tradicionais. [Hornik et al. \(1989\)](#) mostrou que a PMC pode ser treinadas para aproximar virtualmente qualquer função computável.

Ao contrário de outras técnicas estatísticas as redes PMC não fazem suposições prévias sobre a distribuição dos dados. É possível modelar funções não-lineares e podem ser treinadas para generalizar com precisão quando apresentam novos dados. Estas características fazem da PMC uma alternativa atrativa ao desenvolvimento de modelos numéricos e também na escolha entre as abordagens estatísticas. A rede perceptron multicamadas tem muitas aplicações no campo das ciências atmosféricas ([GARDNER; DORLING, 1998](#)).

Entre as principais propriedades e capacidades das RNA, estão: a não linearidade, o mapeamento de entrada-saída, adaptabilidade, aprender com exemplos, processamento paralelo e distribuído, capacidade de generalização e tolerância à falhas ([HAYKIN, 2001](#)), propriedades pertinentes para solução de problemas das ciências atmosféricas.

A capacidade de aprender através de exemplos e de generalizar a informação aprendida é o atrativo principal da solução do problema através de RNA. A generalização associada à capacidade da rede aprender com um conjunto relativamente reduzido de exemplos e posteriormente dar respostas coerentes para dados não conhecidos demonstra que a capacidade das RNA vai muito além do que simplesmente mapear relações de entrada e saída.

Não obstante, as redes neurais atuam como mapeadores universais de funções multivariáveis com custo computacional que cresce apenas linearmente com o número de variáveis. Outra característica importante é a capacidade de auto-organização e de processamento temporal que faz das redes uma ferramenta computacional extremamente poderosa e atrativa para a solução de problemas complexos ([BRAGA et al., 2000](#)).

Segundo [Gardner e Dorling \(1998\)](#), apesar do relacionamento de meteorologia e poluição ser complexo, PMCs foram aplicadas em previsão do campo de qualidade do ar, como aplicações para prever a concentração de ozônio e dióxido de carbono. Outra aplicação em ciência atmosférica com redes neurais foi em previsão de tempo

severo, ou para prever a existência de tornados; ainda na área de previsão, pode-se verificar aplicação em previsão de precipitação e outras.

Na abordagem de aproximação de funções, PMCs foram usadas para avaliar os relacionamentos de variáveis da qualidade do ar. E na abordagem da classificação de padrões para a ciência atmosférica, existem aplicações com redes neurais relacionadas às mudanças globais, verificando nuvens, padrões de circulação atmosférica, imagens do sensoriamento remoto e outras (GARDNER; DORLING, 1998).

Outras aplicações das ciências atmosféricas como assimilação de dados têm utilizado redes neurais artificiais a mais de uma década.

O Laboratório Associado de Computação e Matemática Aplicada (LAC) do INPE foi pioneiro na aplicação de RNA para assimilação de dados (NOWOSAD et al., 2000). Na literatura não haviam implementações do método. Nowosad et al. (2000) apresentou o uso de um filtro de Kalman Estendido Adaptativo e Redes Neurais Perceptrons de Múltiplas Camadas, para três modelos de sistemas dinâmicos não lineares: sistema de tempo discreto caótico de Hénon, sistema de Lorenz em regime caótico e modelo computacional DYNAMO¹. A RNA utilizada para fazer assimilação de dados foi treinada para “emular” o filtro de Kalman, com o objetivo de reduzir o custo computacional. Nowosad et al. (2000) concluiu que para sistemas de alta dimensão as redes treinadas podem ser computacionalmente mais rápidas que filtro de Kalman. O algoritmo foi paralelizado em Vijaykumar et al. (2002), havendo uma redução significativa do tempo de processamento durante o treinamento da PMC.

Em Tang e Hsieh (2001), por exemplo, RNA são sugeridas como possível método de assimilação, onde os autores chamam a atenção para as dificuldades em aplicar RNA em meteorologia e oceanografia como a instabilidade não linear, a grande dimensão da aplicação e a dificuldade em interpretar os resultados. Sugeriram controlar a instabilidade com médias de previsão por conjunto (*ensemble*), e aplicar análise de componentes principais para reduzir a dimensão do problema e interpretar o sinal de saída dos neurônios através da análise espectral. Porém, nenhuma publicação com a implementação foi encontrada.

A pesquisa no LAC/INPE teve continuidade com Härter (2004), onde avaliou-se a eficiência das redes: Perceptron Multicamadas, Funções de Base Radial, Rede de

¹Simulador meteorológico baseado na equação de Água Rasa 1D (LYNCH, 1984) e (CAMPOS VELHO, 1992).

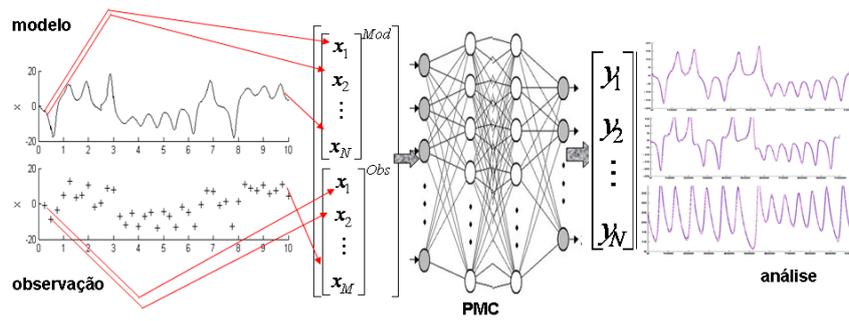
Elman e Rede de Jordan, aplicadas ao mesmo problema de assimilação de dados, em *emular* o Filtro de Kalman. As aplicações foram feitas nos sistemas não lineares caóticos de Lorenz e no modelo DYNAMO, sendo que o PMC também foi aplicado ao modelo de interação não linear entre as ondas de Langmuir, Whistler e Alfvén (modelo de acoplamento de 3 ondas e está ligado à dinâmica da atividade solar)². Este foi um dos resultados inéditos do trabalho de Härter (2004), podendo ser aplicado em clima espacial. No treinamento das RNA, Harter implementou a técnica de validação cruzada (cross-validation), permitindo que se obtenha o melhor conjunto de pesos para o problema estudado.

Uma crítica que foi feita ao uso de RNA-PMC é que, na verdade, a rede não emula propriamente um filtro de Kalman; pois na filtragem de Kalman o ganho varia a cada passo de assimilação. Na PMC os pesos sinápticos, após o treinamento, não variam com o tempo. Deste modo, a assimilação de dados com RNA-PMC não estaria emulando um filtro Kalman padrão, mas, talvez, um filtro de Wiener ou um filtro de Kalman assintótico (onde o ganho é independente do tempo). Uma maneira de se contornar esta restrição foi o emprego de RNA recorrentes (HAYKIN, 2001). Desta forma, o efeito de memória no operador de assimilação estaria presente, emulando de maneira mais fiel um filtro de Kalman (com ganho dependente do tempo). O uso de RNA recorrentes foi explorado na tese de doutorado de Härter (2004) e trabalhos posteriores (HÄRTER; CAMPOS VELHO, 2005).

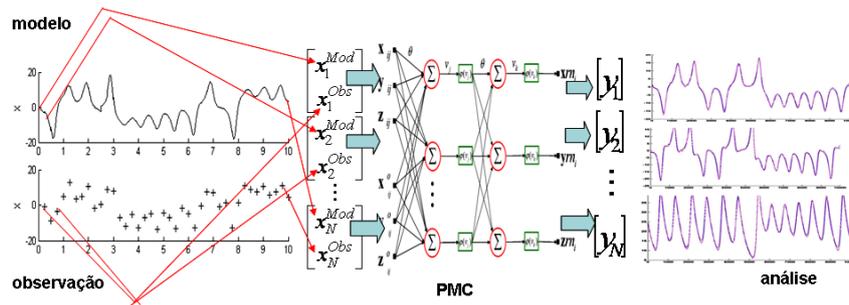
É importante ressaltar que houve uma evolução significativa no desenvolvimento do algoritmo de assimilação de dados com redes neurais artificiais: ao invés de se usar uma rede em que a análise é feita sobre todo o domínio (NOWOSAD, 2001), conforme Figura (4.6(a)). Na proposta de Harter, foi desenvolvida uma estratégia onde a análise é gerada em cada ponto de grade (HÄRTER et al., 2008), conforme Figura (4.6(b)). O ganho com a diminuição do espaço de busca foi um resultado importante da pesquisa de Härter (2004), pois em problemas de minimização de gradiente em aplicações de grande dimensão, diminuir o espaço de busca pode ser o fator a tornar a aplicação operacional viável.

Dando continuidade a essas investigações, a pesquisa de mestrado desenvolvida por Furtado (2008) avaliou o desempenho de uma rede neural PMC treinada para emular: o filtro de Kalman, o filtro de partículas e o método variacional aplicados ao sistema de Lorenz em regime caótico. A técnica obteve sucesso em todas as estimativas.

²Aplicação em assimilação de dados em Clima espacial, veja (HÄRTER, 2004) pag. 93



(a) Rede global para assimilação de todo domínio.



(b) Rede local para assimilação de um ponto de grade de cada vez.

Figura 4.6 - Esquemas de assimilação de dados com RNA sendo (a) pesquisa de Nowosad (2001) e (b) pesquisa de Härter (2004).

Nestas aplicações, analisou-se a diferença absoluta entre as trajetória do modelo obtida pela condição inicial estimada pelas técnicas de assimilação, pela técnica RNA e do modelo verdade (integração do modelo sem assimilação). A assimilação com uma PMC obteve melhor estimativa que a estimativa gerada pelo método variacional. E o Filtro de Partículas e a rede neural foram capazes de assimilar os dados com essa mesma frequência de inserção de observações.

Ainda em aplicação de assimilação de dados Cintra et al. (2010) apresentou resultados de uma aplicação em PMC para emular o *incremento de análise* ao invés da própria análise. O teste aplicado avaliou o desempenho computacional do Filtro de Kalman, generalizando o incremento de análise ao invés de calcular a matriz peso do filtro de Kalman. Nesta tese, vamos avaliar a assimilação de dados em PMC, com a estratégia de rede local em um ponto de grade (Figura 4.6(b) e fazer adaptações para um modelo atmosférico três dimensões.

5 ASSIMILAÇÃO DE DADOS UTILIZANDO PERCEPTRON DE MULTIPLAS CAMADAS

O objetivo da presente tese de doutorado é desenvolver uma *metodologia* de assimilação de dados atmosféricos, que utilize uma rede neural artificial (RNA) e um modelo numérico de três dimensões, *próximo* aos modelos executados em centros operacionais em previsão de Tempo e Clima. O método desenvolvido utilizou a RNA Perceptron Multicamadas (PMC) para obter condições iniciais coerentes com o método de análise do treinamento e para executar ciclos de *previsão numérica do Tempo*, ou seja, a partir de uma análise, executar o modelo realista e obter previsões com desempenho computacional superior ao método de treinamento.

Neste capítulo apresenta-se a *metodologia* utilizada para PMC, descrito no capítulo (4), treinada em pontos de grade do Modelo Atmosférico de Circulação Geral (MCGA) SPEEDY (*Simplified Parametrizations primitivE-Equation DYnamics*), descrito na seção (2.4) do capítulo (2). O método de assimilação de dados utilizado para o treinamento da PMC foi o um Filtro de Kalman por conjunto, o LETKF - *Local Ensemble Transform Kalman Filter* descrito na seção (3.3.3.1) do capítulo (3). O LETKF foi também o método usado para comparação com a análise do método proposto, na apresentação dos resultados deste capítulo. Na descrição do método proposto será apresentado um resumo das técnicas utilizadas que foram completamente descritas em capítulos anteriores.

5.1 Metodologia

O método de assimilação de dados baseado em redes neurais artificiais (RNA) tipo Perceptron Multicamadas (PMC) pode ser descrito como:

$$x_n^a = F_{RNA/PMC}(x_n^f, y_n^o)$$

onde x_n^a representa a análise calculada, $F_{RNA/PMC}$ representa o processo de assimilação por RNA/PMC, x_n^f representa o modelo de estado, y_n^o representa as observações. Como a PMC é um método supervisionado, para o treinamento da PMC é necessário uma outra entrada, o x_n^{as} que representa a análise obtida de outro método de assimilação de dados. Logo para que uma PMC obtenha a análise x_n^a é necessário ter os dados do modelo, das observações e para o treinamento, as análises de outro método de assimilação.

Como foi mostrado no capítulo (4), este método de assimilação de dados foi testado em dois sistemas fortemente não lineares em Nowosad (2001) e Härter (2004). Nestes trabalhos, uma PMC foi empregada, com aprendizado por retropropagação, para simular o filtro de Kalman. Com esta metodologia, uma vez que a rede esteja treinada, o problema de custo computacional foi significativamente reduzido, porém os modelos testados estavam *muito distante* dos modelos numéricos atuais.

Dando continuidade à pesquisa de assimilação de dados utilizando RNA, nesta tese, desenvolveu-se novas PMC(s) utilizando o algoritmo de retropropagação no treinamento. Primeiramente, desenvolveu-se programas para coleta de dados para treinamento da rede, ou seja, os programas de análise do LETKF (análise também do estado da arte em assimilação de dados), foram modificados para aos pré-processar os dados necessários à assimilação, obteve-se as observações nos pontos de grade, a previsão no ponto onde tem observação e a análise resultante do processo de assimilação em arquivos para treinamento. Para esta estratégia de treinamento foi necessário a implementação e conhecimento do modelo SPEEDY, implementação do método de assimilação LETKF além do desenvolvimento dos programas de treinamento e ativação da RNA como método de assimilação de dados. Os programas desenvolvidos e as modificações realizadas estão na linguagem computacional Fortran (90/95).

5.1.1 Implementação do Modelo SPEEDY

O modelo SPEEDY é um MCGA (tridimensional) com um núcleo de equações primitivas (seção 2.4) e um conjunto simplificado de esquemas de parametrizações físicas semelhante aos modelos do estado-da-arte da previsão de Tempo. A versão do modelo, onde os processos de entrada/saída de valores em ponto de grade são executados em intervalos de tempo intermitente, foi desenvolvida a fim de permitir integrações de curto prazo para o ciclo análise/previsão. As alterações no código original (fim climatológico) foram desenvolvidas por Miyoshi (2005). Esta versão do modelo SPEEDY foi concedida para esta pesquisa por ocasião do Curso Intensivo de Assimilação de Dados (KALNAY, 2008) que ocorreu em Buenos Aires em 2008. No curso, o modelo foi utilizado para exercícios de assimilação de dados com o método 3D-Var e LETKF.

A resolução desta versão do MCGA SPEEDY é T30L7, ou seja, o modelo apresenta truncamento horizontal espectral de 30 números de ondas e sete níveis verticais. O pacote de parametrizações é baseado nos mesmos princípios físicos adotados

em esquemas dos MACG mais complexos, tais como: convecção, condensação em larga-escala, nuvens, radiação de ondas curtas, radiação de ondas longas, fluxos de superfície de momentum e energia, e difusão vertical descrito na seção (2.4) deste trabalho. As condições de contorno inferior, tais como TSM (Temperatura à superfície do Mar) são obtidas através de análises do ECMWF, do período 1981-1990. O fluxo de radiação solar são atualizados diariamente.

As variáveis prognósticas são: componentes zonal e meridional da velocidade do vento (\mathbf{u} e \mathbf{v} respectivamente), temperatura(\mathbf{T}), umidade específica (\mathbf{q}) e pressão à superfície (\mathbf{ps}). Os valores de saída estão no espaço físico na grade de tamanho regular ($96 \times 48 \times 7$), ou seja, 96 pontos horizontais (longitude), 48 pontos de latitude e 7 níveis na vertical. As variáveis de entrada do modelo são em níveis sigma (0.08, 0.20, 0.34, 0.51, 0.68, 0.83, 0.95.) e as da saída são em ambos os níveis sigma e pressão (100, 200, 300, 500, 700, 850, 925 hPa).

Os resultados do tempo de execução das integrações do modelo e das análises, se referem à implantação em "cluster" de computadores pessoais (PC) Linux, de processamento de alto desempenho com 27 nós num total de 112 CPUs, instalado no LAC/INPE. Porém as versões das análises/modelo são *sequenciais* e executam *serialmente* em uma CPU, os ciclos de assimilação de dados foram integrações com o LETKF/SPEEDY e RNA/SPEEDY.

Para que o experimento iniciasse o ciclo com o modelo SPEEDY na configuração acima, um período longo de integrações do modelo foi executado para obter o estado *verdadeiro* ou "natureza verdadeira" simulado pela integração do modelo SPEEDY *sem ciclos de assimilação*.

Os experimentos de assimilação de dados nesta tese baseiam-se em sistemas de observação sintéticas, onde a "natureza verdadeira" assume ser conhecida e as observações são obtidas através da adição de ruído aleatório (simulando os erros de observação) às variáveis do modelo do estado "verdadeiro".

Partindo do pressuposto de modelo perfeito, a execução do modelo para simular a natureza é criada com integração do SPEEDY após um período de um (1) ano de *spin-up*¹. A integração do modelo perfeito foi feita para três anos: de primeiro de

¹*Spin-up* é o tempo necessário para que um modelo alcance o estado de equilíbrio, depois de um certo número de várias integrações, supõe-se que os campos simulados não variam significativamente de um tempo para o outro, assim, a partir deste tempo não se observará mudança brusca nas

janeiro de 1982 a 31 de dezembro de 1984, gerando saídas em quatro horários ao dia (00, 06, 12, 18 UTC).

A partir das saídas do modelo "verdadeiro" (chamado **Verdade**) para cada dia e cada horário executou-se o programa para gerar as **observações**, adicionando um ruído aleatório aos valores das variáveis do modelo, com desvio padrão constante na tabela (5.1).

Tabela 5.1 - Desvio Padrão aplicado aos erros das observações.

Variável	desvio padrão
ps pressão à superfície	100.0 [Pa]
u vento zonal	1.0 [m/s]
v vento meridional	1.0 [m/s]
T temperatura	1.0 [K]
q umidade	0.0001 [Kg/Kg] (vapor d'água / ar úmido)

A localização das observações é distribuída de acordo com uma máscara de estações próximo à real rede de estações globais chamada de máscara *realista*, conforme Figura (5.1). A máscara realista, obtida de Miyoshi (2005), simula localizações de **415** estações convencionais de observações globais de radiossondas. A cada ponto da grade latitude/longitude da máscara de observações das variáveis (u, v, t, q) simulou-se uma observação para todos os níveis (superfície inclusive) no ponto escolhido, a variável (ps) foi simulada somente para um nível.

5.1.2 Implementação do LETKF

A idéia básica do LETKF é executar a análise em cada ponto de grade simultaneamente, utilizando as variáveis do estado e todas as observações na região local centrada naquele ponto. Cada membro de um conjunto de estados obtém sua previsão $x_{n-1}^{b(i)}$: ($i = 1, 2, \dots, K$), sendo K o número de membros no tempo n . Para estimar o vetor estado do modelo de referência é usada a *média* deste conjunto de previsões $\bar{\mathbf{X}}^b = 1/K \sum_{i=1}^K x^{b(i)}$.

A partir do modelo de referência inicia-se o LETKF, que será utilizada no passo de

variáveis.

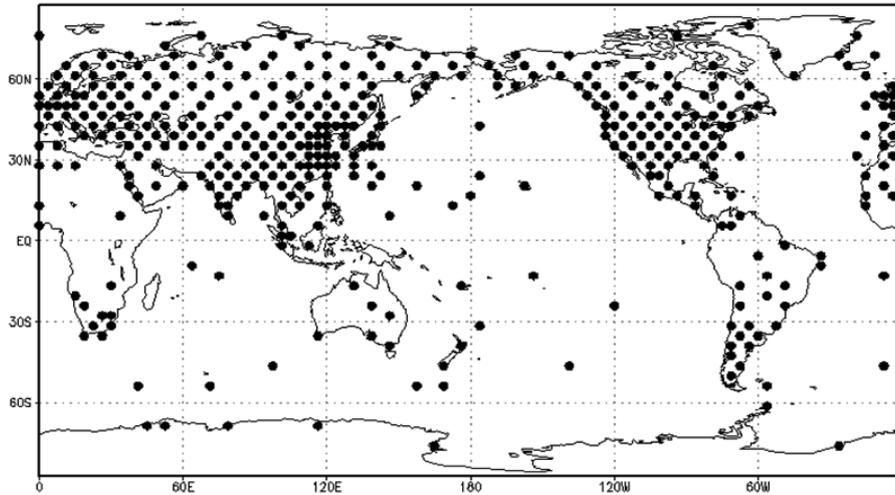


Figura 5.1 - Simulação rede de estações de *observação* realista (415 estações).

previsão do FK, para calcular a matriz de covariância do estado:

$$\mathbf{P}^b = (K - 1)^{-1} \sum_{i=1}^K (x^{b(i)} - \bar{\mathbf{X}}^b)(x^{b(i)} - \bar{\mathbf{X}}^b)^T.$$

No LETKF a assimilação é feita com "localização" espacial, geralmente explícita, então executa-se uma análise para cada ponto da grade do modelo. Para implementação "local", separa-se grupos de observações vizinhas com um ponto central considerando uma raio de distância, para obter uma determinada região do modelo, as observações podem influenciar repetidamente os pontos de grade do modelo.

Na prática, o algoritmo do LETKF foi implementado como segue:

1. Calcula-se um conjunto de previsões globais para obter a média global do conjunto do estado referência (este passo é realizado com K -execuções/integrações do modelo SPEEDY para obter a previsão de curto prazo para cada membro do conjunto);
2. Para cada região local e cada membro do conjunto referência formam-se vetores de estado da região (neste passo é executado no programa que implementa o LETKF, com a máscara realista de observação, considerando uma observação no ponto de grade central de um determinado raio de distância calculando o operador $H_n(x^{b(i)})$ para a grade do modelo);

3. Em cada região local, projeta-se o vetor "local" obtido no passo 2, ao subespaço dimensional que melhor represente o conjunto naquela região;
4. Faz-se assimilação de dados baseada em Filtro de Kalman por conjunto, em cada região local, obtendo a análise (executa-se o passo de previsão e atualização do EnKF para cada região);
5. Com as análises locais obtidas no passo 4, forma-se uma análise global (obtidos os campos $x^{a(i)}$ para cada ponto de grade, retorna os valores para a grade global obtendo x^a e os vetores de "perturbação" (descritos por 3.3.2) para iniciar a previsão do conjunto do próximo horário);
6. Voltar ao passo 1.

Esta implementação foi dividida em dois programas, o *ensemble* (conjunto) de previsões e a assimilação **LETKF**. Os resultados são campos de análises e previsões globais para cada membro do conjunto e um campo de análise para a previsão de curto prazo que será usada no **ciclo** de assimilação de dados do próximo conjunto de observações disponíveis. As saídas foram pós-processadas em níveis *sigma* e em níveis de pressão.

As execuções do LETKF demandam grande espaço em disco, que deve ter monitoramento e transferência de arquivos para outros dispositivos, para preservar os recursos físicos.

5.1.3 Implementação da Rede Neural Artificial

As redes Perceptron Multicamadas (PMC) (GARDNER; DORLING, 1998) são os modelos de RNA mais utilizados em pesquisas científicas, pois são sistemas computacionais com processamento paralelo e distribuído e que apresentam a capacidade de aprender e armazenar conhecimento experimental.

Especificações da arquitetura de uma PMC está descrito na seção (4.2.1) do capítulo (4). Para treinar a rede utilizou-se o algoritmo de aprendizado *retropropagação do erro*. Este algoritmo é um processo supervisionado, onde a rede recebe vetores de entrada com sua respectiva resposta desejada para treinamento. Este processo de treinamento determina os pesos sinápticos descrito na seção (4.3) do capítulo (4).

Após o aprendizado, com o conjunto de pesos sinápticos é capaz de ativar os

neurônios da rede e obter resultados para novas entradas, ou seja, fora do conjunto de dados de treinamento, *generalizando* a informação aprendida. Na realidade a rede produz uma estimativa próxima da resposta desejada, em termos de mínimos quadrados.

Portanto para implementação das PMCs foi necessário o desenvolvimento do algoritmo de retropropagação para treinamento da rede e para posterior ativação. No treinamento há atualização dos pesos sinápticos e na ativação os pesos são fixos, portanto não há o retorno do algoritmo, somente a fase para frente, ou seja, da entrada à saída.

As entradas de cada PMC são: um vetor de cada variável prognóstica do modelo de *previsão* em um ponto de grade e um vetor de uma *observação* simulada para aquele *ponto de grade*. Logo cada PMC tem duas (2) entradas na camada de entrada, um (1) neurônio na camada de saída referente a variável prognóstica da *análise* calculada pela rede para o determinado ponto de grade. A camada escondida tem um número de neurônio para cada rede, conforme Tabela (5.2). O número de neurônios da camada escondida única foram parametrizados por experimentação, a *maioria* das redes obteve melhores resultados com onze (11) neurônios na camada escondida. Definiu-se o número de neurônios em função da dimensão das camadas de entrada e saída da PMC.

Tabela 5.2 - Parâmetro Variável na arquitetura das Redes PMC.

Redes	Variável	unidade da medida	Neurônios camada escondida
PMC_{ps}	pressão	PA	11
PMC_u	vento zonal	m/s	09
PMC_v	vento meridional	m/s	11
PMC_T	temperatura	K	11
PMC_q	umidade	Kg/Kg	11

A utilização de um grande número de camadas escondidas não é recomendado, pois cada vez que o erro médio durante o treinamento é utilizado para atualizar os pesos das sinapses da camada imediatamente anterior, o erro se torna menos útil ou preciso. A única camada que tem uma noção precisa do erro cometido pela rede é a

camada de saída. A última camada escondida recebe uma estimativa sobre o erro. A escolha de uma única camada escondida também foi devido à premissa baseada no teorema de aproximação universal descrito no capítulo (4) e em Haykin (2007) e Nowosad (2001). Muitas camadas escondidas podem dificultar a aproximação, pois os neurônios escondidos tendem a interagir entre si.

Com relação ao número de neurônios nas camadas escondidas, foi definido *empiricamente*, tendo o cuidado de não especificar muitos neurônios, o que pode levar a rede a memorizar os dados de treinamento (*overfitting*), ao invés de extrair as características gerais que permitirão a generalização, ou poucos neurônios, que pode forçar a rede a gastar tempo em excesso tentando encontrar uma representação ótima.

Para este trabalho foram desenvolvidas PMCs para cada variável prognóstica do modelo ($\mathbf{PMC}_u, \mathbf{PMC}_v, \mathbf{PMC}_T, \mathbf{PMC}_q, \mathbf{PMC}_{ps}$), conforme Tabela (5.2). Esta estratégia foi utilizada devido às grandezas numéricas das variáveis serem bem diferentes e a normalização aplicada às entradas ficou comprometida. Devido aos limites da função de ativação, o parâmetros de entrada da rede foram *normalizados*, considerando o *valor máximo* e o *valor mínimo* para cada vetor de entrada da rede para valores limitantes da função de ativação escolhida. Na função sigmóide tangente hiperbólica, os valores devem estar entre (1 e -1). A propriedade multivariável deve ser mantida pois o método que ensina, ou seja, a resposta desejada tem embutida no resultado da análise a propriedade multivariável do LETKF.

A *taxa de aprendizagem* definida como η indica a velocidade de aprendizado da rede e foi parametrizada igualmente para todas as PMCs como $\eta = 0.01$. Uma taxa de aprendizado mais alta provocou oscilações no treinamento e impediu a convergência do processo de aprendizado.

Em todas as redes, a *função de ativação* dos neurônios da camada escondida e do neurônio da camada de saída é a função tangente hiperbólica

$$\varphi(net_i) = (1 - e(-net_i))/(1 + e(-net_i)),$$

equação (4.4) do capítulo (4). Essa escolha foi devido ao fato de ser uma função limitada e monotonicamente crescente para facilitar o objetivo de aproximação de função não-linear e aos limites superior e inferior (1) e (-1).

Os *critérios de parada* utilizados ao obter o melhor conjunto de pesos da rede foi

em função do *menor erro* entre a análise calculada pela rede e a análise do LETKF, comparado a um *erro objetivo*, o erro objetivo foi parametrizado para todas as redes como (0.0001), ou seja, (10^{-4}). O erro da rede é calculado considerando o erro médio quadrático, **emq**, das diferenças entre a entrada da resposta desejada $d = x_{(m)LETKF}^a$ e a saída calculada da rede $y = x_{(m)PMC}^a$ definido como:

$$\mathbf{emq} = \frac{1}{NPAD} \left(\sum_{m=1}^{NPAD} (y - d)^2 \right)$$

sendo *NPAD* o número de padrões de entrada da rede.

Foi definido o *número de épocas* de treinamento como o tamanho do laço de apresentação dos padrões à rede para que haja aprendizado. Caso a rede não alcance um erro menor que o *erro objetivo*, o treinamento acaba quando alcançar o limite de "épocas" especificado. O número máximo de épocas foi fixado em 5000 para todas as redes.

Inicialmente o treinamento foi feito no domínio global, porém para obter os pesos da rede com menor tempo de treinamento, os dados passaram a ser coletados e divididos em regiões, porém, separadas de maneira diferente das regiões no LETKF. Como foi dito, o objetivo desta divisão foi a aceleração do treinamento da PMC, na obtenção de conjuntos de dados menores mas suficientes para o aprendizado da PMC. A análise é feita em cada ponto de grade, independente da divisão em regiões, que como resultado do processamento teremos a análise global. As PMCs utilizam o mesmo algoritmo de retropropagação para treinamento, logo os conjuntos de pesos obtidos são também divididos em região do treinamento.

O globo terrestre foi dividido em seis regiões iguais com $(24 \times 32 \times 7)$ *pontos de grade*: sendo três regiões no hemisfério norte e três regiões no hemisfério sul, conforme Tabela (5.3) e Figura (5.2). O conjunto de dados, separado por região, tem o tamanho diferente, pois as observações estão distribuídas de forma irregular no globo terrestre, simulando a localização real da rede de observações convencionais. Relembrando que a assimilação é feita no ponto de grade que tem observação. Para os horários 00 e 12 UTC, são 415 estações com 5 observações nos sete níveis para as informações de ar superior, e apenas superfície para a pressão. Nos horários de 06 e 18 UTC, as 415 observação tem somente as 5 observações na superfície.

Tabela 5.3 - Regiões do globo terrestre.

Quadrante	latitude (graus)	longitude (graus)
região 1	$(-90^{\circ} a 0^{\circ})$	$(+0^{\circ} a 120^{\circ})$
região 2	$(0^{\circ} a -90^{\circ})$	$(+0^{\circ} a 120^{\circ})$
região 3	$(-90^{\circ} a 0^{\circ})$	$(+123.5^{\circ} a 240^{\circ})$
região 4	$(0^{\circ} a -90^{\circ})$	$+123.5^{\circ} a 240^{\circ})$
região 5	$(-90^{\circ} a 0^{\circ})$	$(+243.5^{\circ} a 360^{\circ})$
região 6	$(0^{\circ} a -90^{\circ})$	$(+243.5^{\circ} a 360^{\circ})$

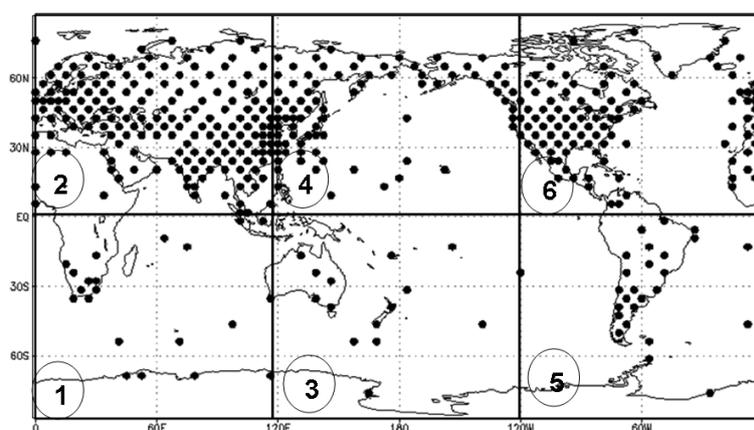


Figura 5.2 - Divisão do globo em seis regiões iguais conforme tabela 5.3.

Foram definidas **30** (trinta) PMCs, ou seja, os pesos fixos são obtidos para cada variável (cinco) em seis regiões. Notem que quanto maior a divisão do espaço físico, maior o controle na manipulação dos arquivos de pesos fixos das conexões da RNA, pois *cada* rede deve ter um conjunto de dados de entrada e *cada* rede gera um conjunto de pesos fixos. A configuração da arquitetura da PMC é feita por variável prognóstica.

Como as regiões foram divididas espacialmente sem considerar a localização das observações, os conjuntos de dados ficaram menores em regiões onde as observações são esparsas e convergiram conforme as Tabelas (5.4) a (5.8).

Neste trabalho, o cálculo do operador H_n , ou seja, a criação das regiões locais considerando as observações fora do ponto de grade central, não foi desenvolvido. As observações sintéticas foram simuladas diretamente no ponto de grade implemen-

tado pelo LETKF. O mesmo conjunto de observações foi utilizado na entrada da rede neural.

Para treinamento das PMCs, o programa que faz assimilação LETKF, foi alterado para gravar os campos necessários para treinamento da PMC. Os conjuntos de dados foram coletados das execuções do ciclo de assimilação/previsão do LETKF/SPEEDY para o período de 01 de janeiro de 1982 a 31 de dezembro de 1984. Os arquivos foram gravados e divididos mês a mês, por variável e por região, contendo os vetores de observação, de previsão e da análise do LETKF, vetores de tamanhos iguais por corresponderem aos pontos de grade com observação e a quantidade pode ser verificada nas Tabelas (5.4) a (5.8).

Na coleta dos dados de entrada, a influência da observação do ponto de grade que a **possui** foi calculada para pontos de grade vizinhos **sem** observação, criando uma *pseudo*-observação denominada \hat{y}^o da seguinte forma:

$$\hat{y}^o = \left(\sum_{p=1}^N y_p^o \frac{1}{(\Delta r)^2} \right) \quad (5.1)$$

onde N é o número de pontos de grade vizinhos dentro de um quadrante de distancia r . Δ é o número de pontos de grade a considerar a influência, r é a distância entre o ponto de grade da observação e o ponto vizinho, p é o ponto de grade dentro do quadrante.

$$r = (x_i^o - x_i^b)^2 + (x_j^o - x_j^b)^2 + (x_k^o - x_k^b)^2,$$

os subscritos i, j, k , são as localizações dos pontos de grade nas dimensões $(x, y, z$, relativos a longitude, latitude, altura respectivamente), x^o é a localização do ponto de grade da observação e x^b a localização do ponto de grade sem observação (somente o valor do modelo). As "pseudo"-observações foram coletadas como observações nas entradas da PMC associado aos pontos de grade da distância Δr com o valor do modelo de previsão e a análise LETKF correspondente.

O parâmetro Δ da equação (5.1) foi definido como 2 (vizinhos), $\Delta = 2$. Portanto o algoritmo verificou doze pontos de grade para incluir a influência da observação, se necessário. Foram dois pontos $j + 1$, dois pontos $j - 1$, dois pontos $i + 1$, dois pontos $i - 1$, dois pontos $k + 1$ e dois pontos $k - 1$. Se no ponto verificado não tiver observação, então foi adicionado a observação de influência \hat{y}^o conformr (5.1). Não foram considerados os pontos na diagonal do ponto de grade nesta implementação.

5.1.3.1 Treinamento das PMCs

Após os dados coletados, arquivos de três meses (janeiro, fevereiro, março) de cada ano foram agrupados para obter os conjuntos de dados de entrada de treinamento para cada PMC. O treinamento foi realizado executando o algoritmo de retropropagação (Tabela 4.1) para cada conjunto de vetores de dados apresentado como padrões de entrada às RNAs. A convergência com o número de épocas e o erro encontrado pode ser verificado nas Tabelas (5.4) a (5.8). O tamanho do conjunto de dados é diferente para cada região, pois as observações estão distribuídas de forma irregular pelo globo terrestre de maneira semelhantes às observações reais. O treinamento passa a ser dividido por regiões com as informações pertinentes a cada região, assim não há necessidade de associar o ponto de grade à sua localização espacial, o que pode ter resultado na rápida convergência do treinamento.

O LETKF foi executado por três anos (1982, 1983, 1984) e obtidos os conjuntos de dados para treinamento das 30 PMCs. Selecionou-se os meses de Janeiro, Fevereiro e Março de cada ano. Assim totalizou-se 9 meses de dados, divididos por região e tipo de variável.

O treinamento demandou o tempo de processamento de aproximadamente 2 : 15 : 00 horas.

Tabela 5.4 - Convergência do Treinamento da PMC_{ps} . Conjunto de dados para (1 nível) a superfície.

Região	Tamanho do conjunto de dados	Número de épocas	erro atingido
região 1	179944	2	2.0218^{-5}
região 2	1203240	2	3.6726^{-6}
região 3	134416	2	2.7949^{-5}
região 4	598368	2	2.5374^{-6}
região 5	215716	2	8.1780^{-6}
região 6	739288	2	2.2814^{-6}

A Tabela (5.6) mostra o número de épocas para a convergência para a unidade. As regiões que têm maior número de observações e previsões, convergem mais rapidamente, enquanto que, regiões com menos observações demandam mais épocas para

Tabela 5.5 - Convergência do Treinamento da PMC_T . Conjunto de dados para (7) níveis.

região	Tamanho do conjunto de dados	Número de épocas	erro atingido
região 1	742540	2	2.2118^{-5}
região 2	5001576	1	4.3505^{-5}
região 3	502810	2	1.1758^{-5}
região 4	2480129	2	3.2969^{-6}
região 5	891048	2	1.4941^{-5}
região 6	3065552	1	4.8132^{-5}

Tabela 5.6 - Convergência do Treinamento da PMC_q . Conjunto de dados para 7 níveis.

região	Tamanho do conjunto de dados	Número de épocas	erro atingido
região1	742540	1425	9.9916^{-5}
região2	5001576	2	8.8069^{-5}
região3	502810	4331	9.9990^{-5}
região4	2480129	2	8.6436^{-5}
região5	891048	1724	9.9995^{-5}
região6	3065552	6	9.7906^{-5}

apresentar convergência. Este padrão de convergência deve estar ligado ao valor numérico para unidade, que apresenta valores muito baixos.

Tabela 5.7 - Convergência do Treinamento da PMC_u . Conjunto de dados para 7 níveis.

região	Tamanho do conjunto de dados	Número de épocas	erro atingido
região1	742540	2	3.5293^{-5}
região2	5001576	1	8.5060^{-5}
região3	502810	2	6.7484^{-5}
região4	2480129	2	3.3624^{-6}
região5	891048	2	4.5183^{-5}
região6	3065552	1	9.5508^{-5}

No treinamento da observação componente do vento zonal, a região 3 apresentou um número baixo de observações, apesar de ser o mesmo período de coleta de da-

dos, porém o treinamento deve ter encontrado um mínimo local e o algoritmo de treinamento não utilizou a variável “de momento” para agilizar esta falha. Porém, a parada do algoritmo se deu pelo número máximo de épocas e o erro conseguido foi aceitável para a convergência.

Tabela 5.8 - Convergência do Treinamento da PMC_v . Conjunto de dados para 7 níveis. (*) erro aceitável.

região	Tamanho do conjunto de dados	Número de épocas	erro atingido
região 1	742540	2	5.6205^{-5}
região 2	5001576	2	4.5512^{-5}
região 3	502810	5000	1.2344^{-4} (*) ^t
região 4	2480129	2	4.7463^{-6}
região 5	891048	2	6.2196^{-5}
região 6	3065552	1	5.7822^{-5}

5.1.3.2 Ativação/Generalização das PMCs

A partir dos conjuntos de pesos fixos encontrados para cada PMC, o algoritmo da RNA para assimilação de dados foi implementado da seguinte forma:

1. Ler as observações e as previsões globais de determinado horário;
2. Separar as entradas por variáveis e por regiões;
3. Para cada tipo de variável prognóstica e cada região local, ativar a PMC específica:
 - (i) ler o vetor de entradas (observação e previsão);
 - (ii) normalizar os vetores de entrada nos limites das função de ativação da rede;
 - (iii) ler o conjunto de pesos fixos;
 - (iv) apresentar o vetor de entradas ao algoritmo de retropropagação de erro; calcular passo 1, 2 e 3 da Tabela 4.1 da seção 4.3.1 para cada linha do vetor de entrada que representa um ponto de grade, obtendo a análise $x^{a(i)}$;
 - (v) retornar o vetor normalizado aos valores originais ao vetor de saída da rede;

4. Com as análises de cada região obtidas no passo 3, retorna uma análise global (obtidos os vetores $x^{a(i)}$ para cada região), retorna os valores para a grade global obtendo x^a para cada variável;
6. Voltar ao passo 3.

Os resultados são campos de análises globais para iniciar o modelo SPEEDY para a obter a previsão de curto prazo que será usada no próximo **ciclo** de assimilação de dados do próximo conjunto de observações disponíveis. As saídas foram pós-processadas em níveis *sigma*.

Esta implementação foi dividida em dois programas, um para *treinamento* das PMCs, executado somente para obter os conjuntos de pesos fixos da rede e outro para executar em ciclo de assimilação de dados com a ativação das PMCs conforme algoritmo acima descrito.

5.2 Resultados da Assimilação de Dados com Redes Neurais Artificiais

O ciclo de assimilação de dados PMC/SPEEDY foi executado de 01 de janeiro de 1985 a 31 de janeiro de 1985 para generalização da RNA e avaliação dos resultados. Para comparação das análises, executou-se o ciclo de assimilação com análises do LETKF implementado no trabalho, para o mesmo período. A previsão executada foi somente para obter o campo do "modelo de referência" para a próxima geração da análise, neste trabalho não foi avaliada a previsão a médio prazo de nenhuma das análises.

Para visualização das figuras, foi utilizado o software livre GrADS (*Grid Analysis and Display System*) ou Sistema de Exibição e Análise de Grade, muito utilizado em aplicações de meteorologia, capaz de visualizar campos meteorológicos de vários formatos (DOTY, 1998). As figuras (5.3) a (5.11) apresentadas a seguir são de campos das análises PMC/LETKF gerados para cada variável prognóstica e campos do "estado verdadeiro" aqui nomeado como "Verdade". Também serão mostradas as diferenças entre as duas análises e a diferença entre a análise gerada pela PMC e o campo do estado verdadeiro. As figuras apresentam os resultados de uma variável para cada horário e nível. Na avaliação visual dos campos, verifica-se a proximidade das análises, os campos de diferenças confirmam a proximidade.

A média mensal das diferenças entre os campos de análises (LETKF - PMC) para

todo o ciclo de assimilação (124 execuções), são apresentadas nas Figuras (5.12), (5.13) e (5.14). As figuras apresentam diferenças pequenas entre as médias, verifica-se que em algumas localizações, como na América do Sul, surgem diferenças maiores.

Para avaliar a trajetória das análises no tempo, calculou-se a "média zonal global" das variáveis \mathbf{T} (temperatura), \mathbf{q} (umidade específica) e \mathbf{Ps} (pressão à superfície) para todo o ciclo de assimilação do experimento. Calculou-se a mesma média dos campos do modelo "verdade" para comparação.

Um determinado ponto (fixo) na latitude do hemisfério norte, do hemisfério sul e do equador, foi escolhido para calcular a média de todos os pontos da longitude dos campos de análises geradas pela PMC. O mesmo cálculo foi realizado para os campos de análise gerados pelo LETKF e para os campos do modelo verdade. Nas Figuras (5.15) e (5.16) verifica-se que as trajetórias das análises estão muito próximas e nota-se que o modelo verdade diverge das duas análises no final do período. O modelo verdade, são integrações do modelo SPEEDY sem assimilação de dados.

O erro médio quadrático (\mathbf{emq}) das análises foi calculado fixando-se um ponto na longitude, obtendo o erro das diferenças entre os campos de análise da PMC e os campos do modelo verdade. Calculou-se da mesma forma o \mathbf{emq} dos campos de análise da PMC e os campos do modelo Verdade.

As Figuras (5.17) e (5.18) apresentam a média meridional do \mathbf{emq} de todo o período do ciclo para um ponto na longitude (87°W), em todos os pontos da latitude. As figuras correspondentes, são as diferenças entre as duas médias meridionais do \mathbf{emq} do LETKF menos o \mathbf{emq} da RNA. Verifica-se as diferenças entre os erros em todo o domínio meridional são da ordem de 10^{-3} no campo da superfície como no campo de nível superior.

Com base no \mathbf{emq} das análises em relação ao estado verdade, a média temporal do erro quadrático médio das análises em relação ao estado verdade, são apresentados nas Figuras (5.19) e (5.20).

Ainda com o \mathbf{emq} calculado, as Figuras (5.21) e (5.22) apresentam o erro do perfil vertical (dos sete níveis sigma) fixando um ponto de latitude (-20°). Os perfis e as diferenças entre eles são de temperatura e umidade. Os resultados confirmam que as saídas da análise do LETKF são similares, pois as diferenças entre os erros dos perfis entre as análises do LETKF e RNA são da ordem de 10^{-3} .

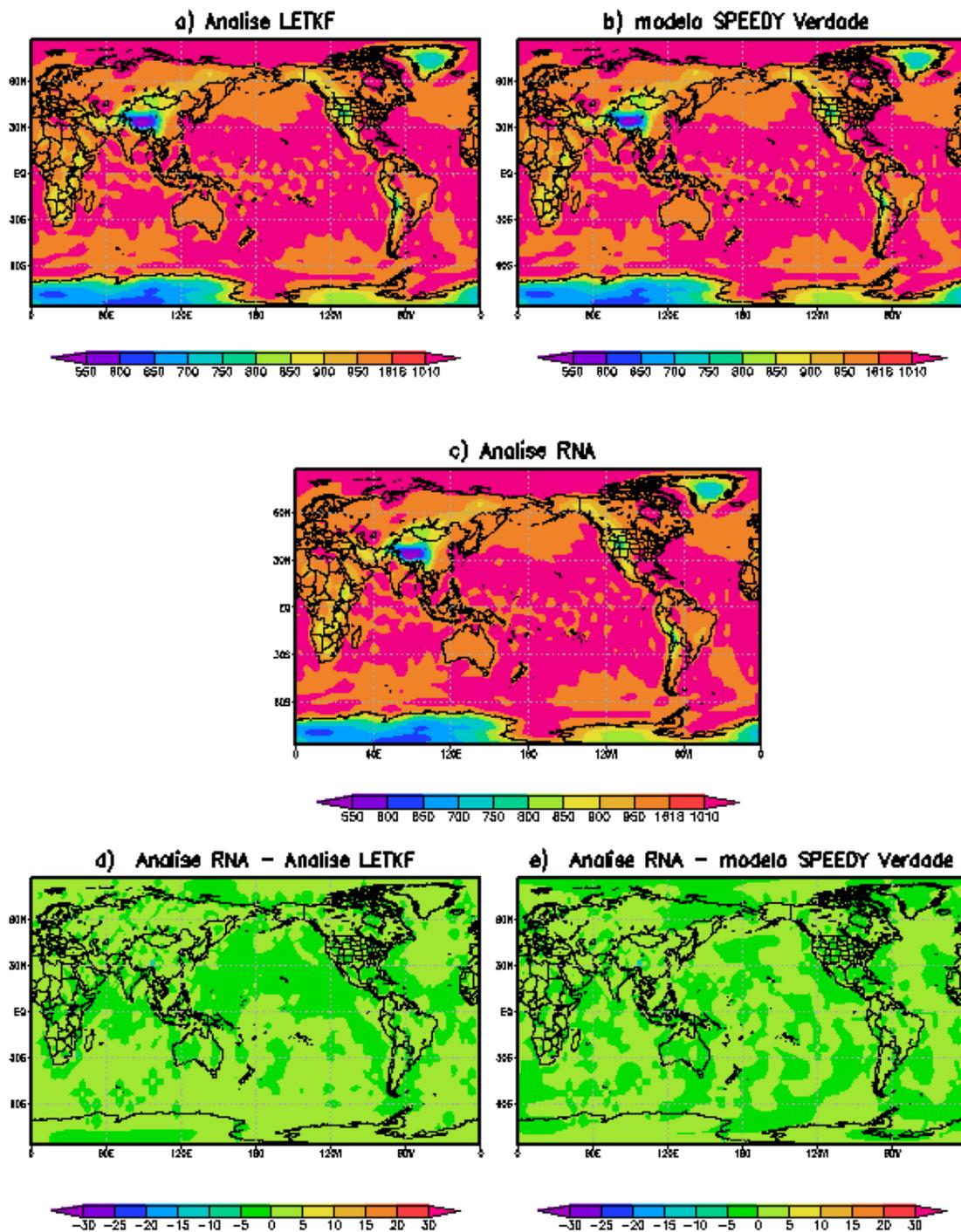


Figura 5.3 - Campos de **pressão (PS)** [hPa] na superfície das análises geradas pelos métodos de assimilação RNA(c) e LETKF(a) e do estado verdadeiro(b). Diferenças entre as duas análises(d) (PMC_{ps} e LETKF). Diferença entre análise da (PMC_{ps} e estado verdadeiro)(e), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação

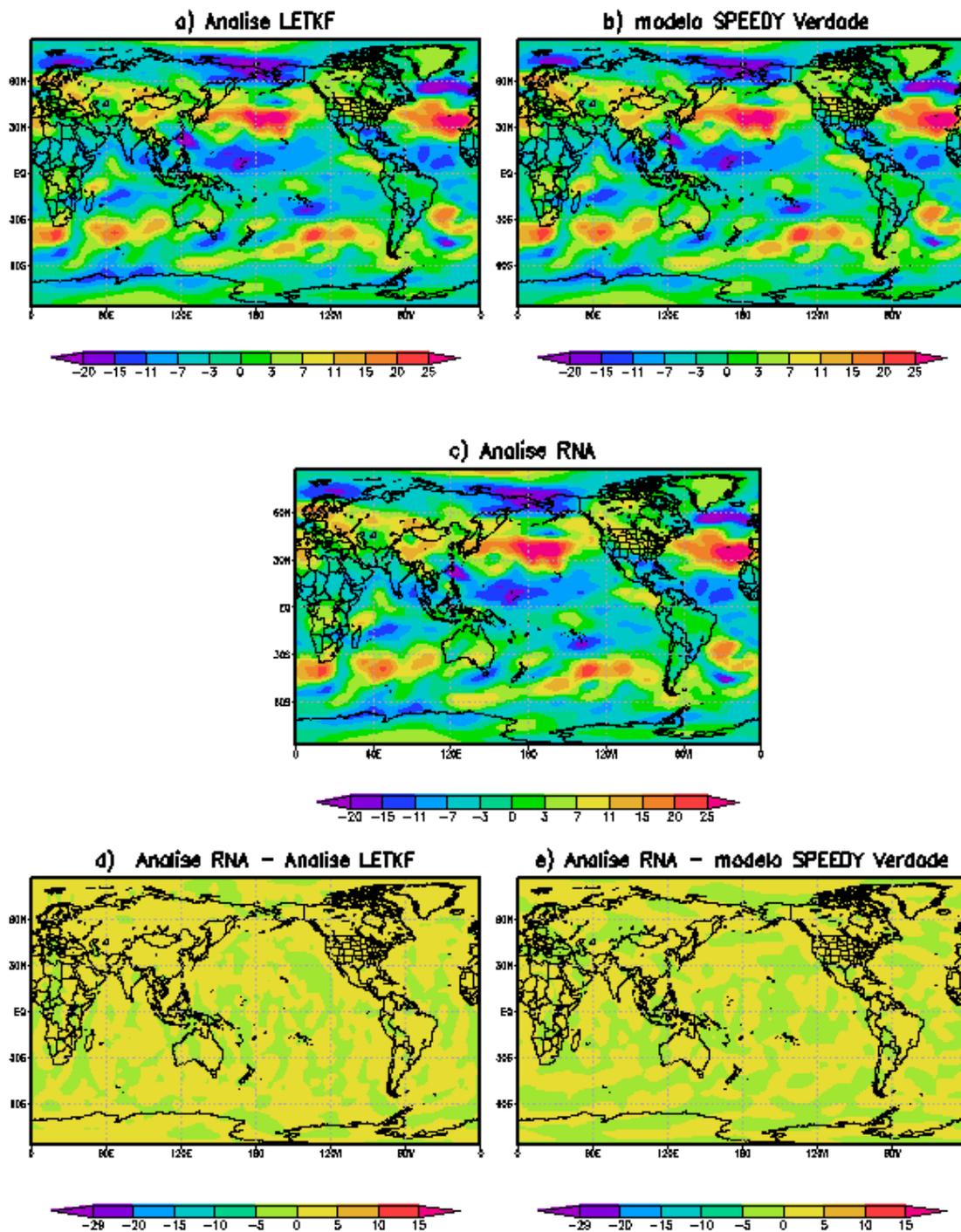


Figura 5.4 - Campos de **componente do vento zonal (u) [m/s] - superfície** das análises geradas pelos métodos de assimilação RNA(c) e LETKF(a) e do estado verdadeiro. Diferença entre as duas análises(d) (PMC_u e LETKF) e a diferença entre análise da (PMC_u e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação

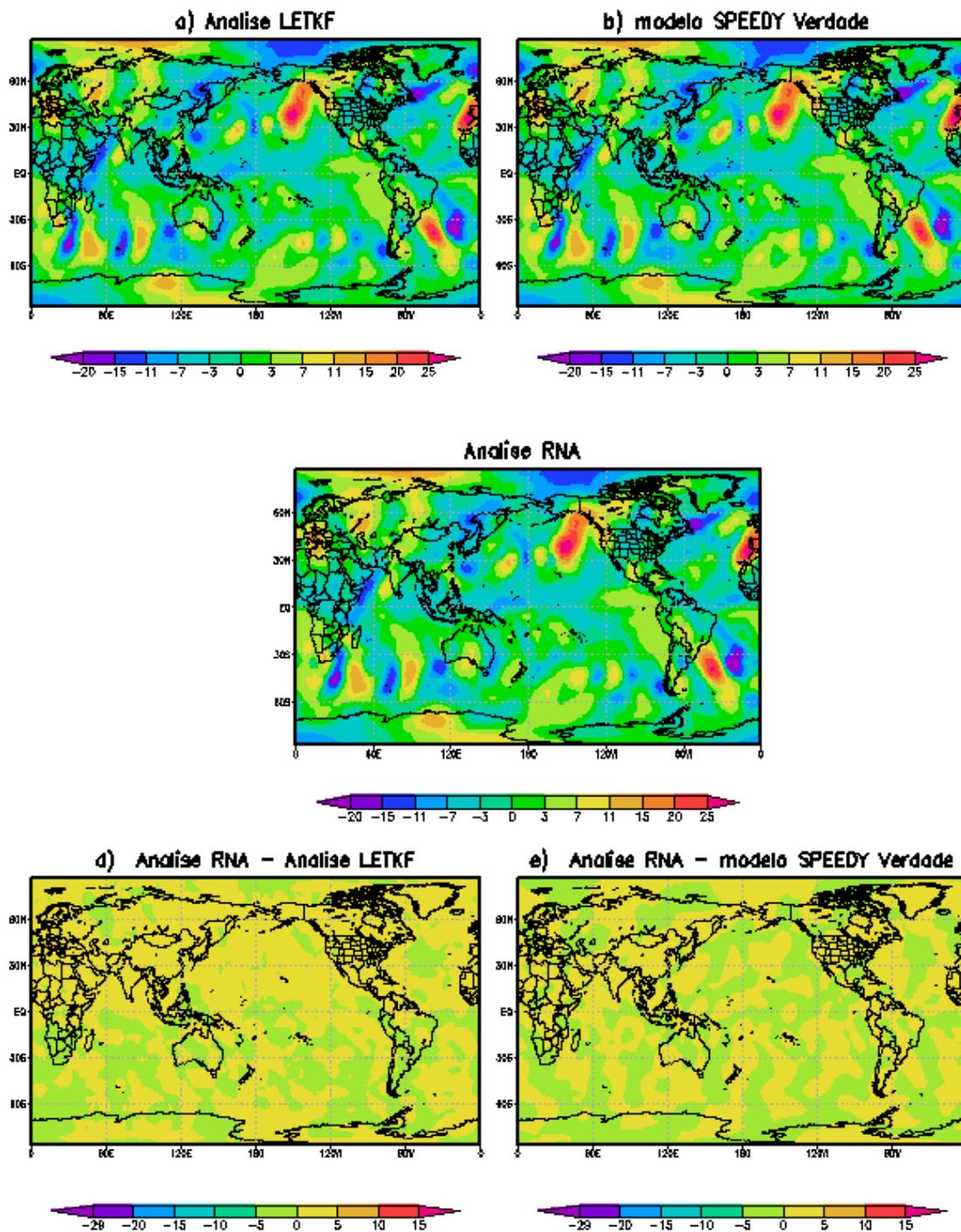


Figura 5.5 - Campos de **componente do vento meridional (v)**[m/s] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_v e LETKF) e diferença (e) entre análise da (PMC_v e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação

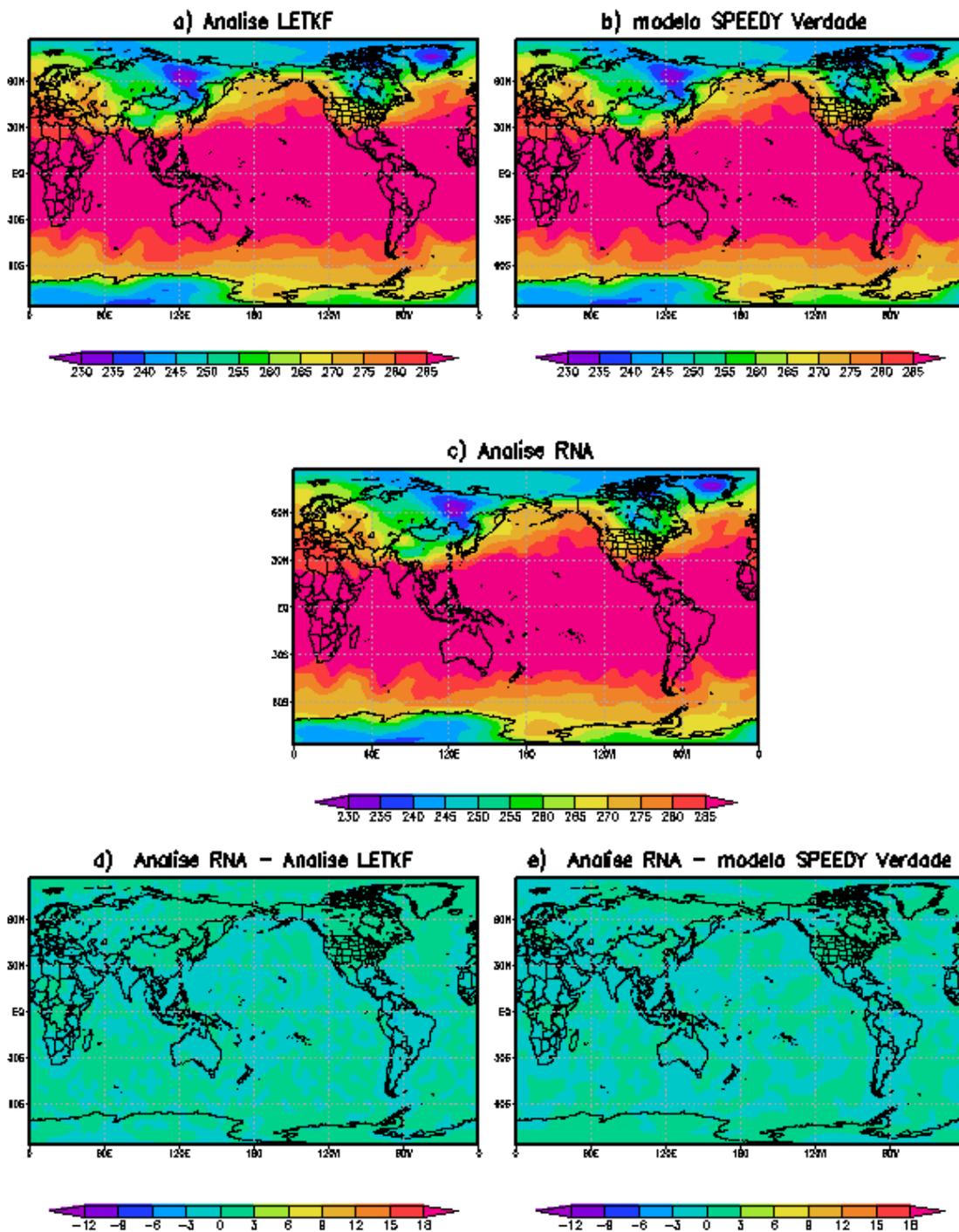


Figura 5.6 - Campos de temperatura absoluta (T) [K] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises (d) (PMC_t e LETKF) e diferença (e) entre análise da (PMC_t e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação

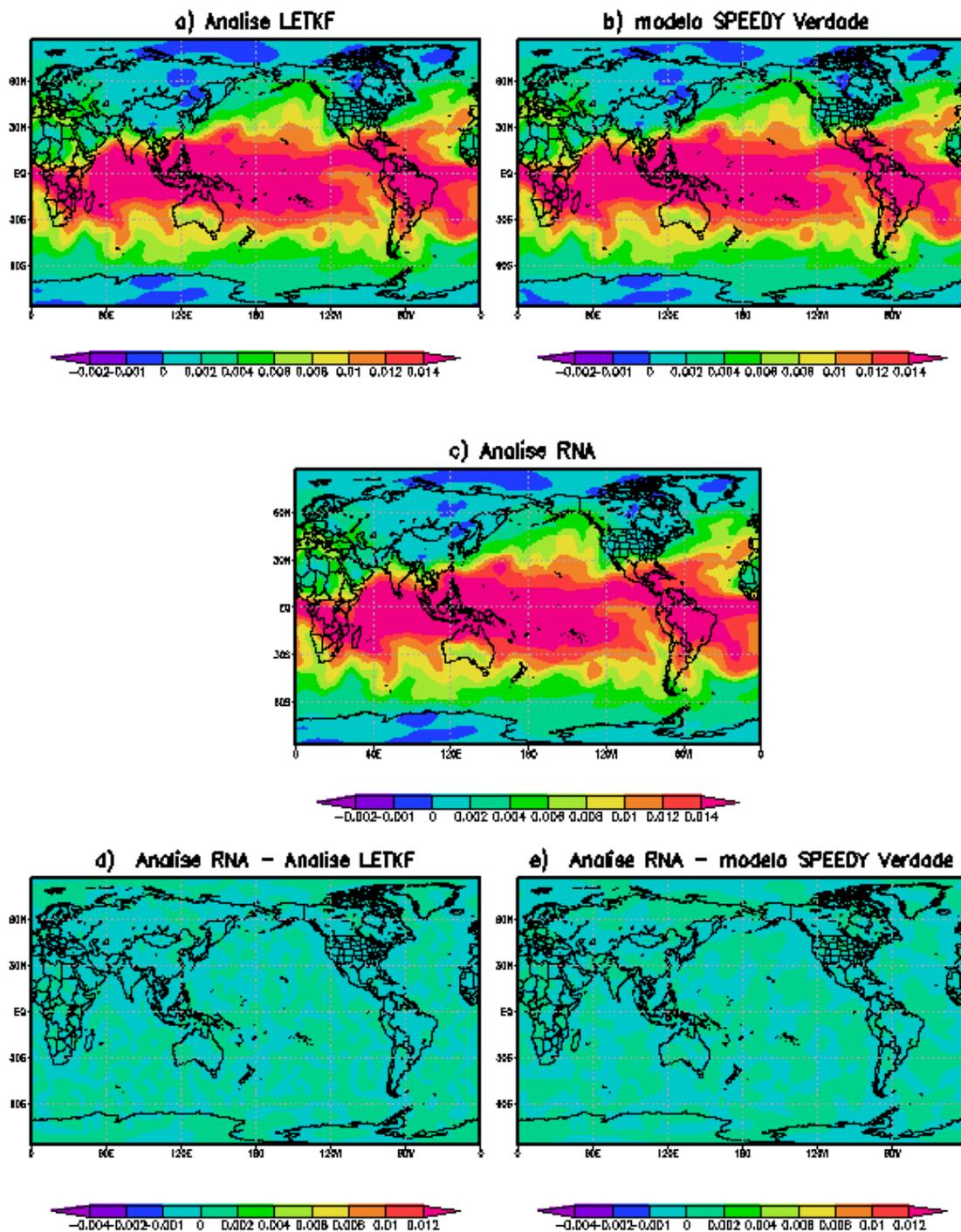


Figura 5.7 - Campos de **umidade específica**(q)[Kg/Kg] - superfície das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_q e LETKF) e diferença (e) entre análise da (PMC_q e estado verdadeiro), no dia 11 de janeiro de 1985 às 18 UTC, ou seja, após 41 execuções do ciclo de assimilação

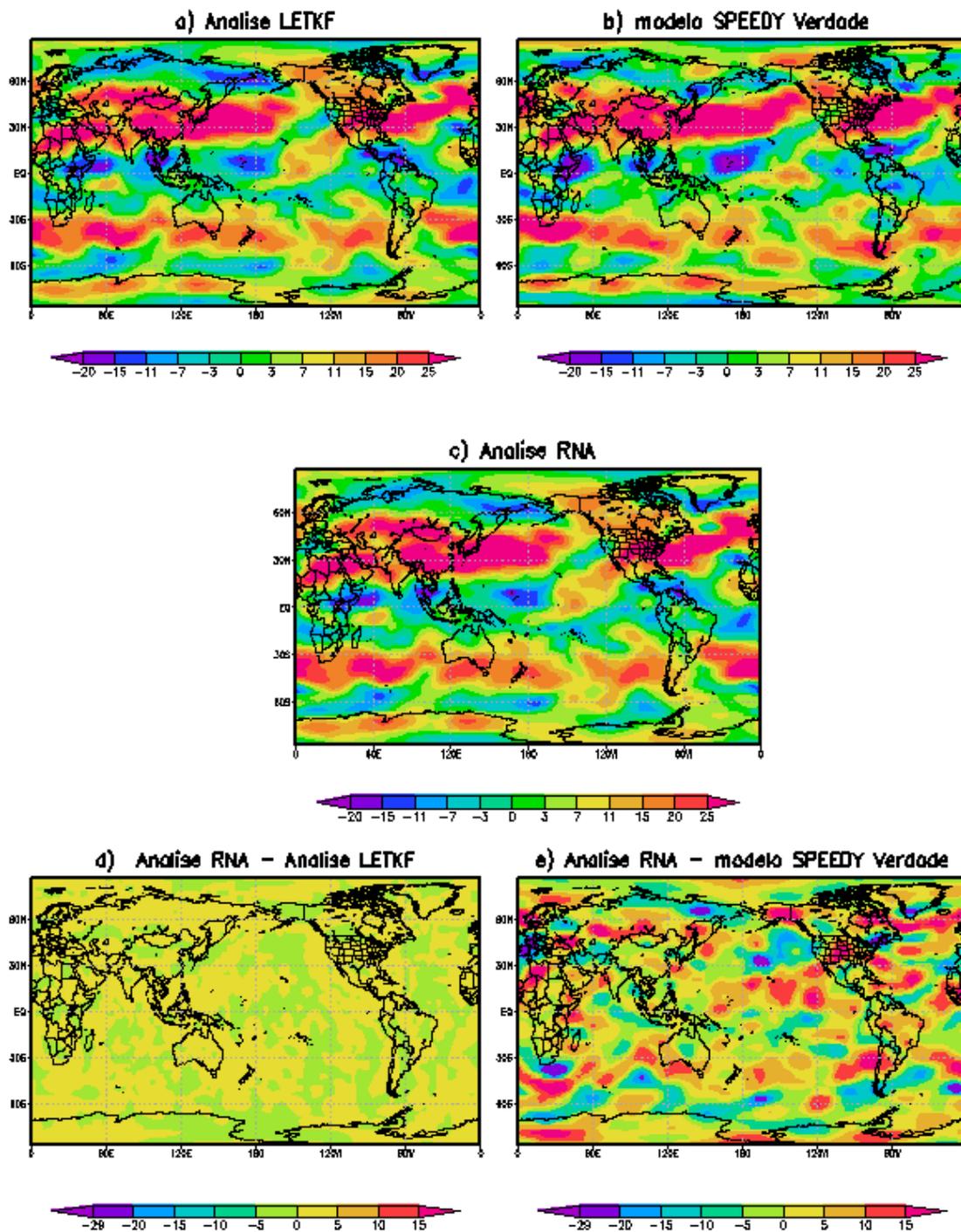


Figura 5.8 - Campos de **vento zonal (u)[m/s]- 500hPa** (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) (PMC_u e LETKF) e a diferença (e) entre análise da (PMC_u e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação

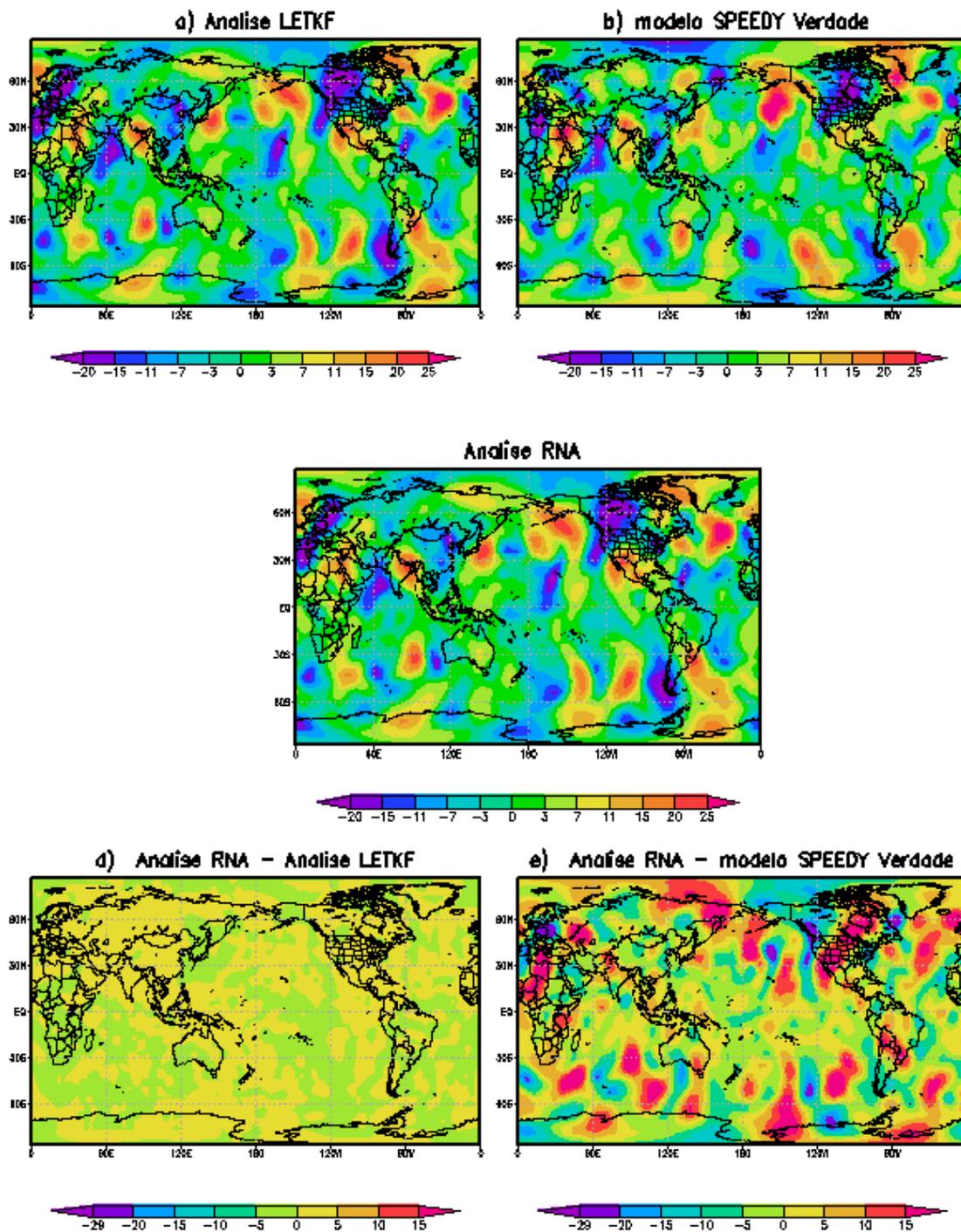


Figura 5.9 - Campos de vento meridional (v) [m/s] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises (d) (PMC_v e LETKF) e a diferença (e) entre análise da (PMC_v e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação

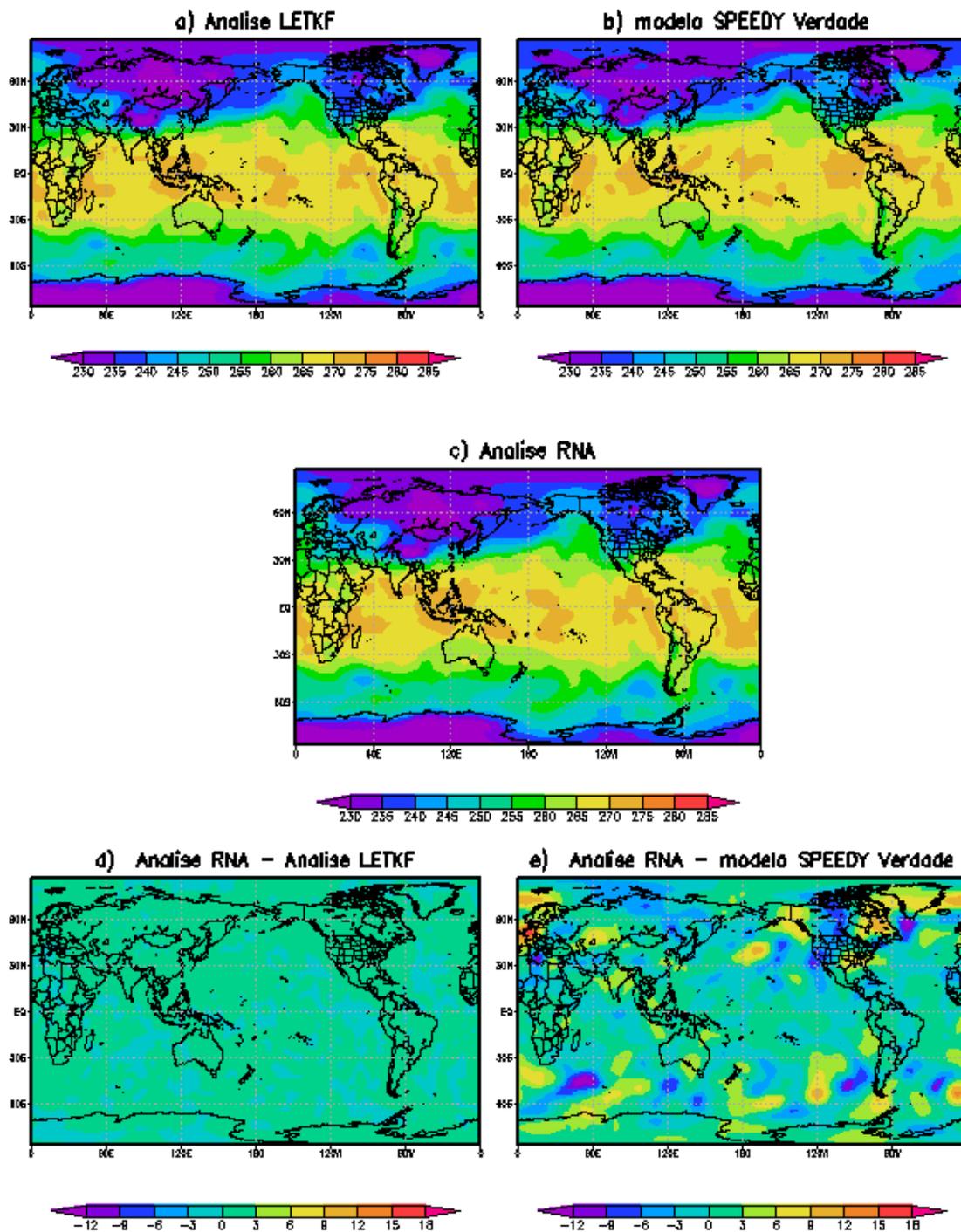


Figura 5.10 - Campos de temperatura absoluta (T)[K] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises(d) ($PMCT$ e LETKF) e a diferença (e) entre análise da ($PMCT$ e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação

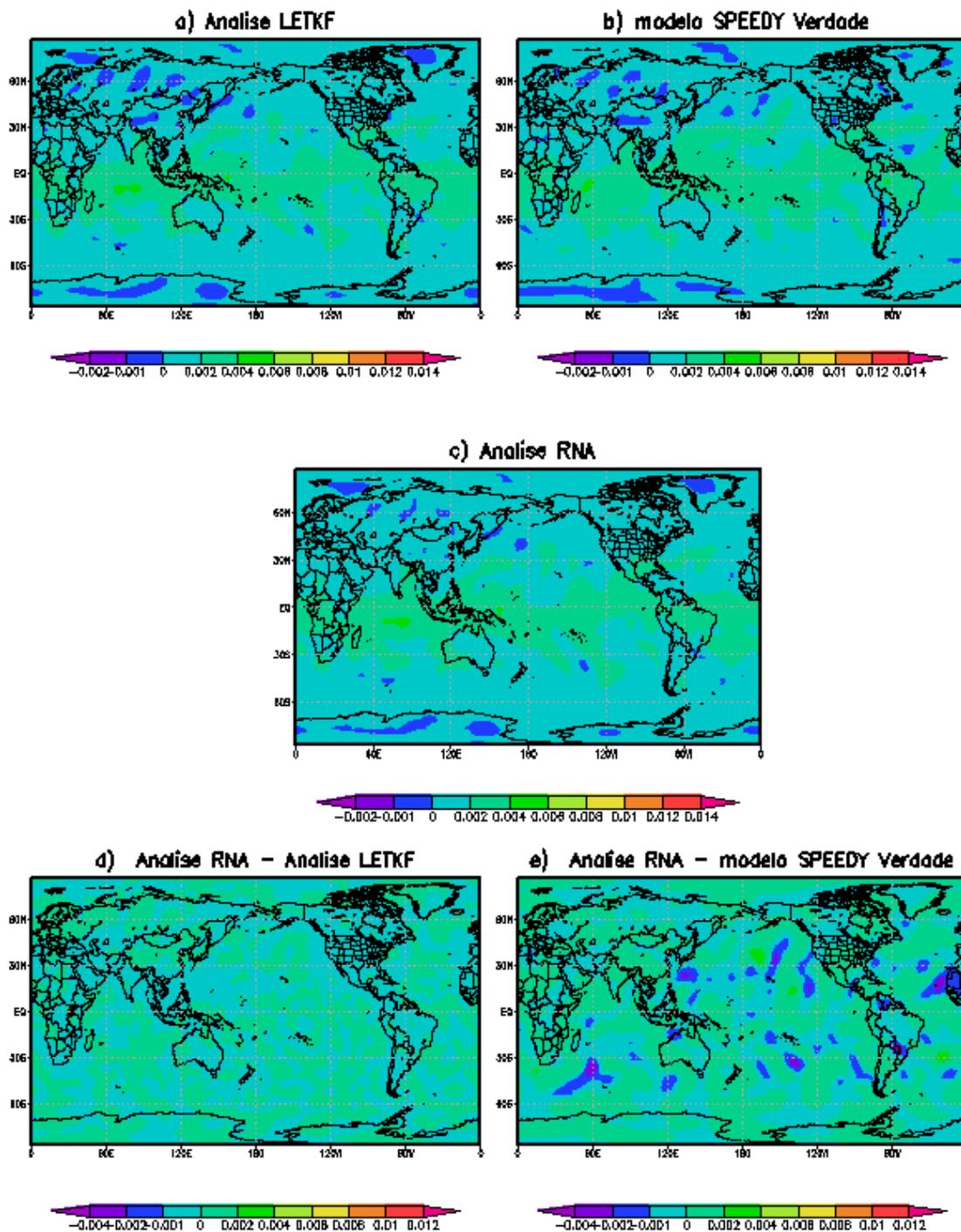


Figura 5.11 - Campos de **umidade específica** (q) [Kg/Kg] - 500hPa (nível 4) das análises geradas pelos métodos de assimilação RNA (c) e LETKF (a) e do estado verdadeiro. Diferenças entre as duas análises (d) (PMC_q e LETKF) e a diferença (e) entre análise da (PMC_q e estado verdadeiro), no dia 31 de janeiro de 1985 às 00 UTC, ou seja, após 120 execuções do ciclo de assimilação.

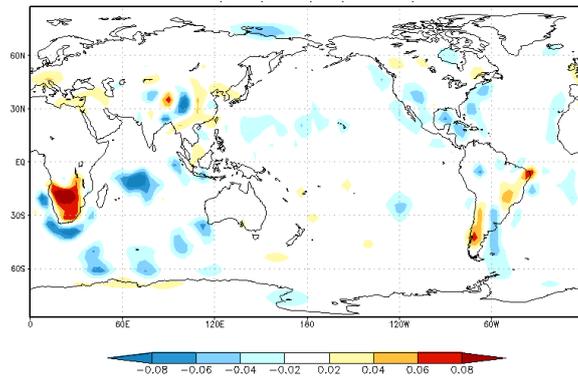


Figura 5.12 - Campo de diferença das Médias dos campos da pressão na superfície das análises RNA e das análises LETKF para o ciclo das 120 execuções.

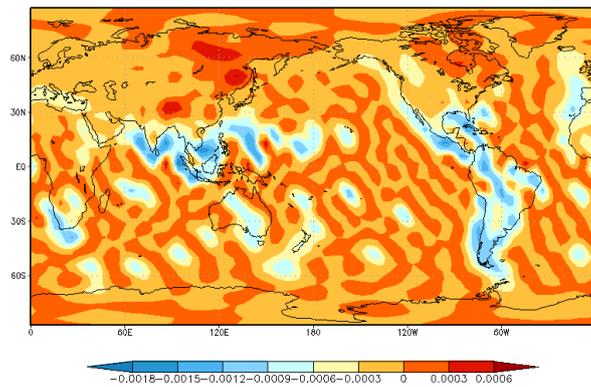


Figura 5.13 - Campo de diferenças da Média dos campos de umidade específica das análises RNA e das análises LETKF para o ciclo das 120 execuções.

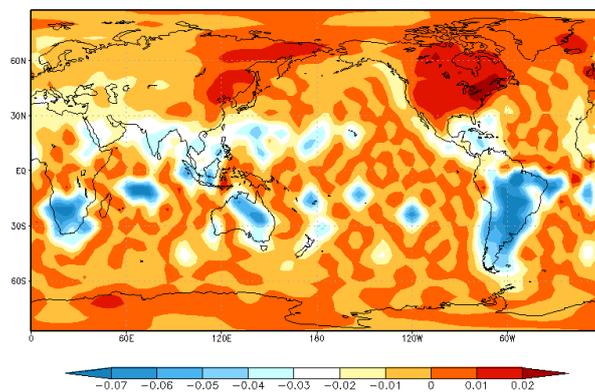


Figura 5.14 - Campo de diferenças da Média dos campos da temperatura das análises RNA e das análises LETKF para o ciclo das 120 execuções.

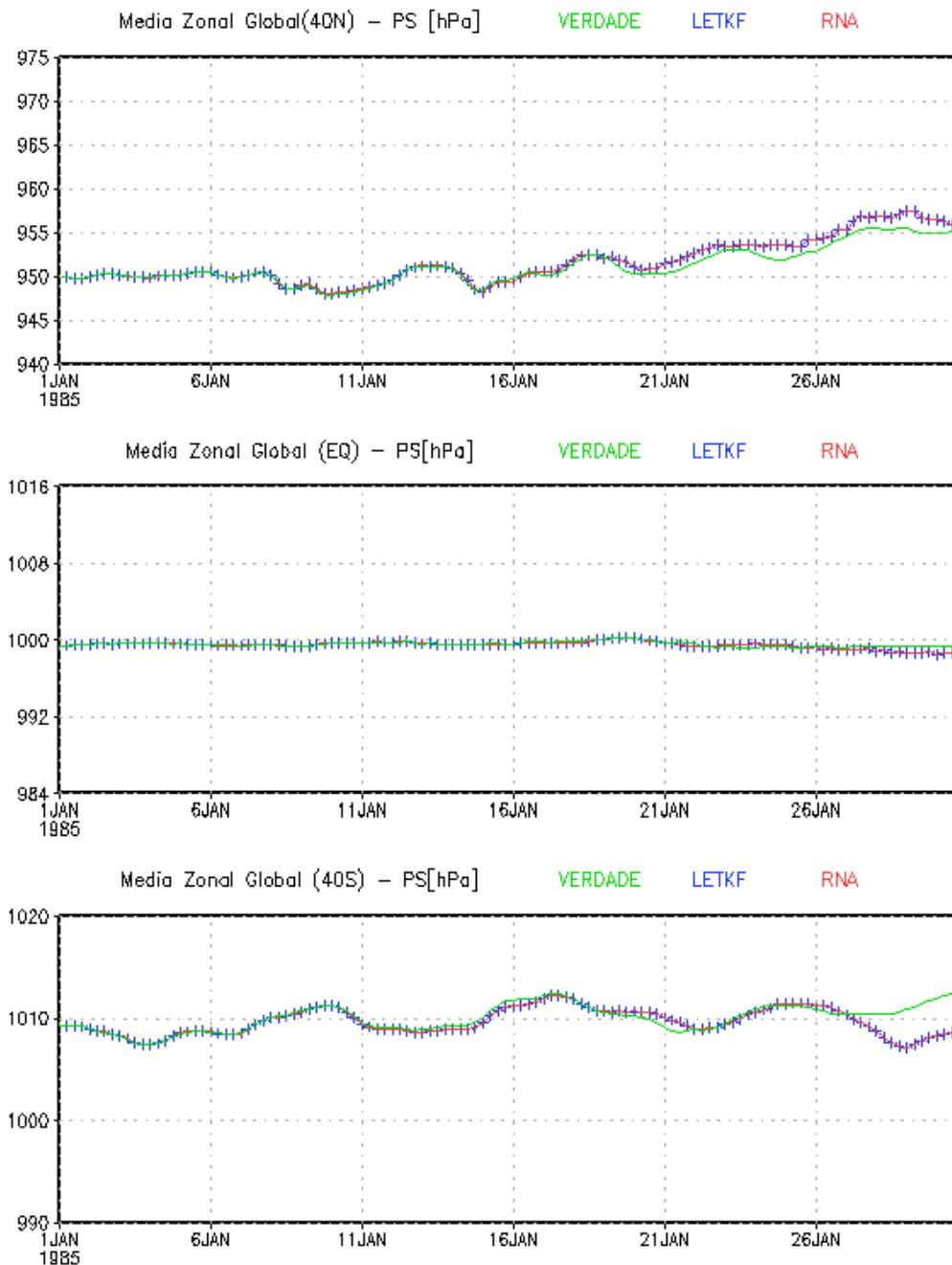


Figura 5.15 - Média Zonal Global de pressão à superfície [hPa] das análises geradas pelos métodos de assimilação **RNA** e **LETKF** e do estado **VERDADE** de 1 a 31 de janeiro de 1985 em 120 execuções do ciclo de assimilação. As médias foram fixadas na latitude 40N, na latitude (0) equador e na latitude 40S e longitude (0° a 360°).

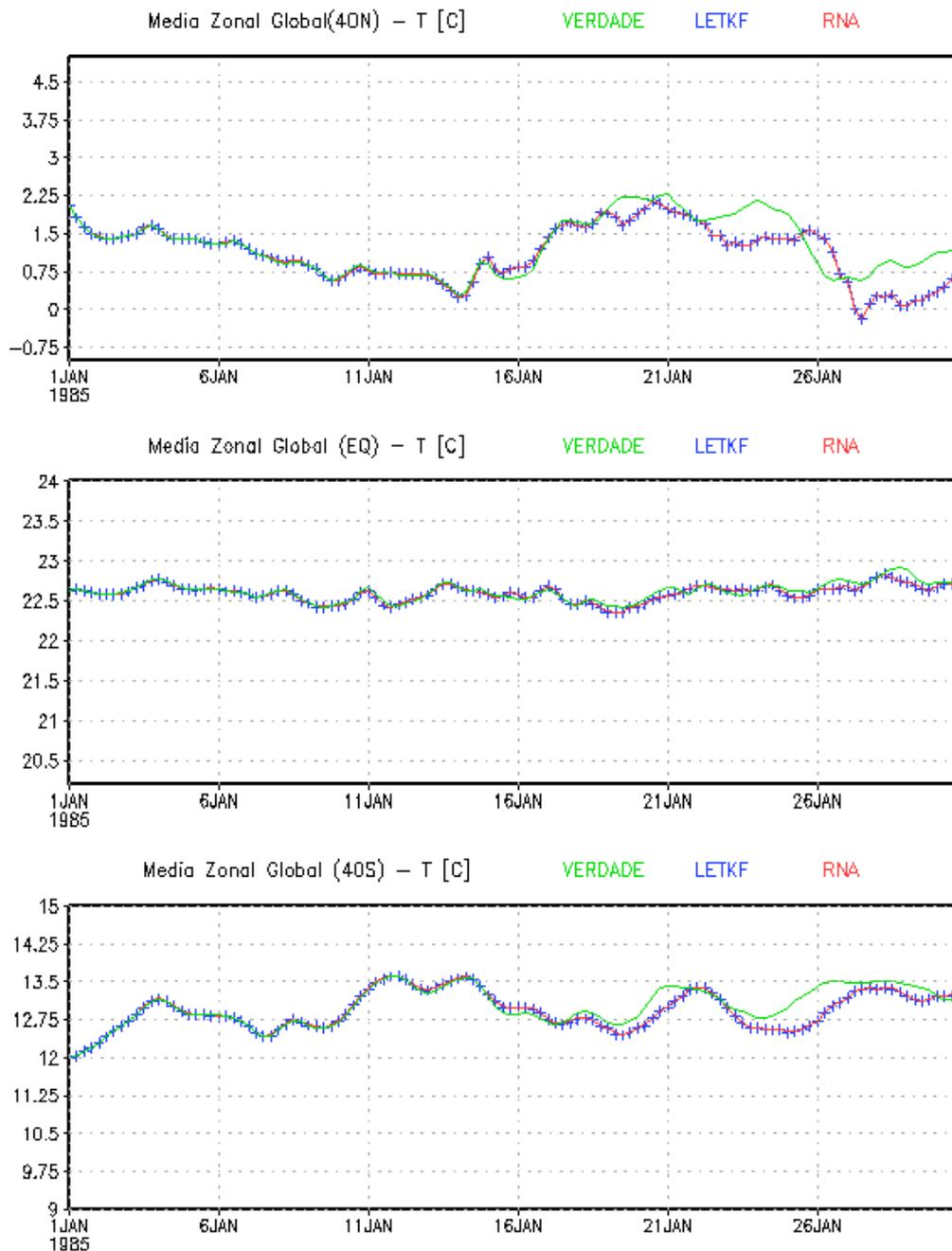
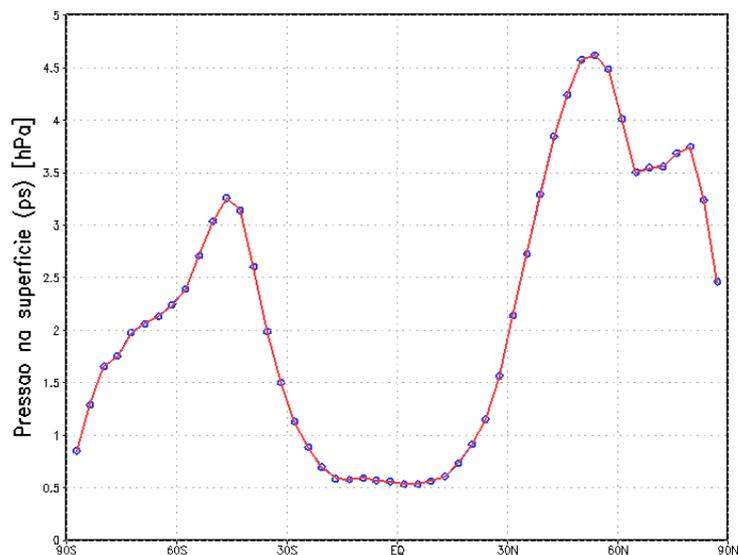
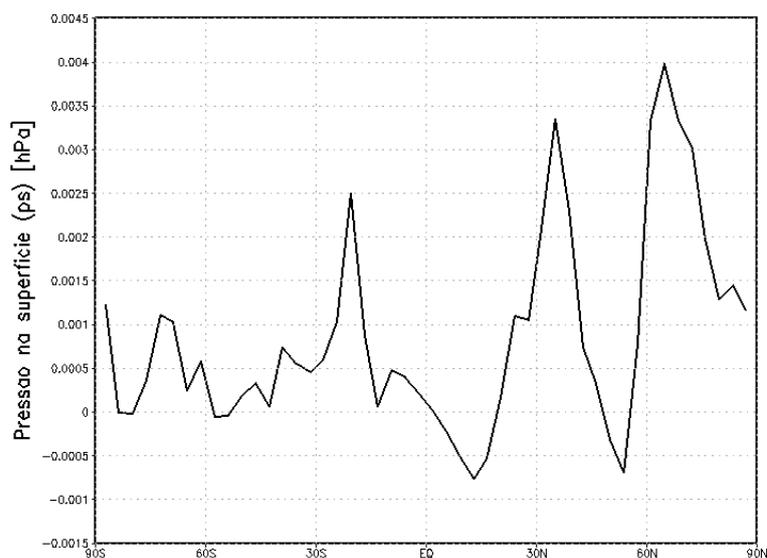


Figura 5.16 - Média Zonal Global de Temperatura [K] das análises geradas pelos métodos de assimilação **RNA** e **LETKF** e do estado **VERDADE** de 1 a 31 de janeiro de 1985 em 120 execuções do ciclo de assimilação. As médias foram fixadas na latitude 40N, na latitude(0) equador, na latitude 40S e longitude (0° a 360°).



(a) erro médio quadrático



(b) diferenças (LETKF-RNA)

Figura 5.17 - Média meridional de todo o período do ciclo de assimilação do erro médio quadrático da análise em relação ao estado Verdade; os erros das análises do LETKF em linha (azul) e os erros das análises da PMC em linha (vermelha) para a variável **pressão à superfície**(a) e a diferença entre os erros (b).



(a) erro médio quadrático

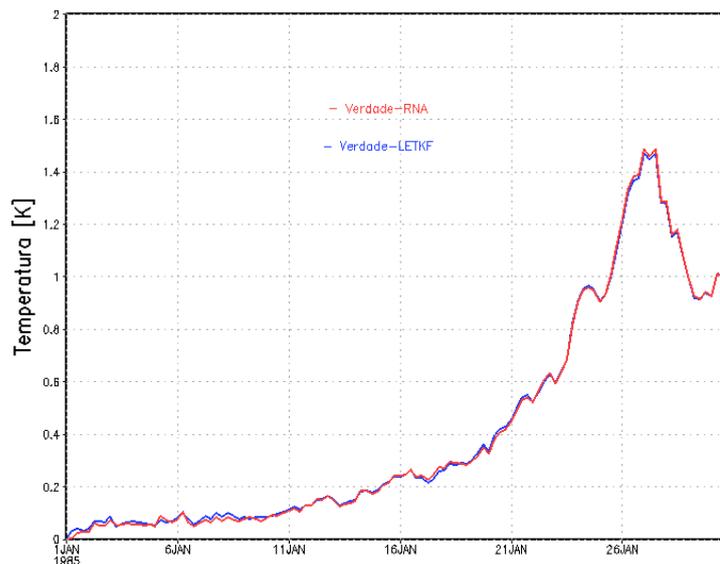


(b) diferenças (LETKF-RNA)

Figura 5.18 - Média meridional de todo o período do ciclo de assimilação do erro médio quadrático da análise em relação ao estado Verdade; os erros das análises do LETKF em linha (azul) e os erros das análises da PMC em linha (vermelha) para a variável **temperatura absoluta à 500 hPa**(a) e a diferença entre os erros (b).



(a) erro médio quadrático

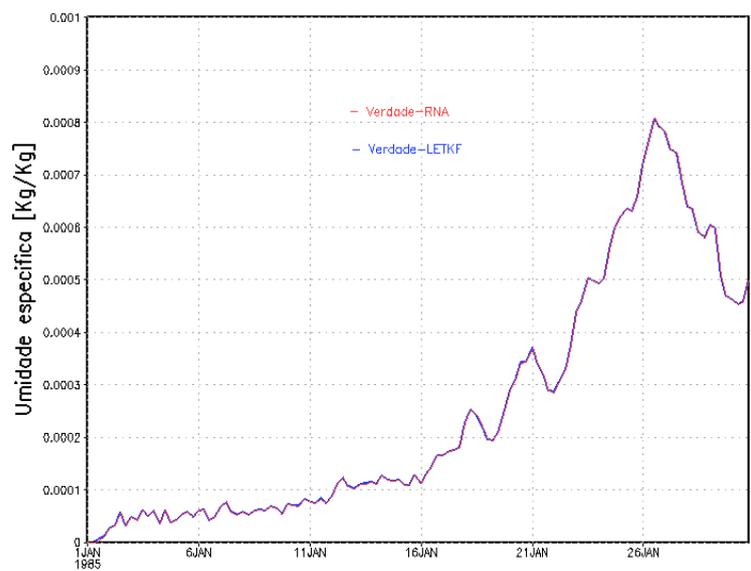


(b) erro médio quadrático

Figura 5.19 - Média temporal do ciclo de assimilação (01 a 31 de janeiro de 1985) do erro médio quadrático em relação ao modelo Verdade; os erros das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), para (a) pressão e (b) temperatura ambos à superfície.

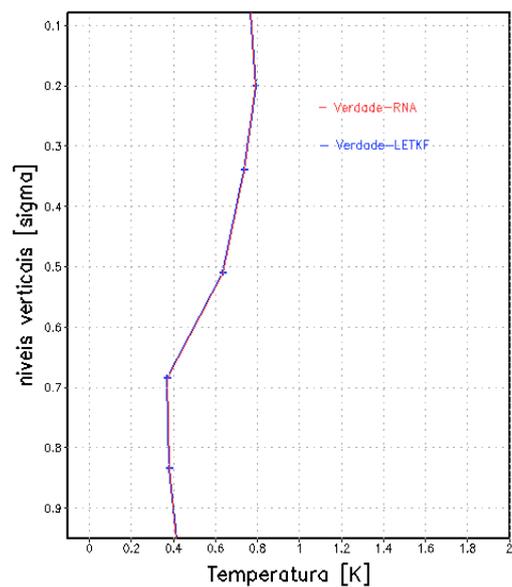


(a) erro médio quadrático

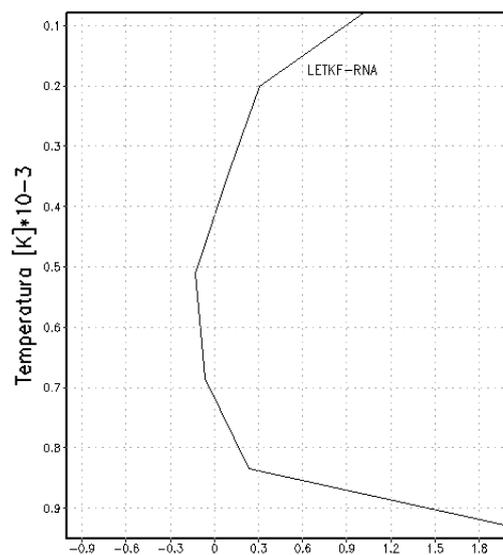


(b) erro médio quadrático

Figura 5.20 - Média temporal do ciclo de assimilação (01 a 31 de janeiro de 1985) do erro médio quadrático em relação ao modelo Verdade; os erros das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), para (a) umidade específica à superfície e (b) umidade específica à 500 hPa.

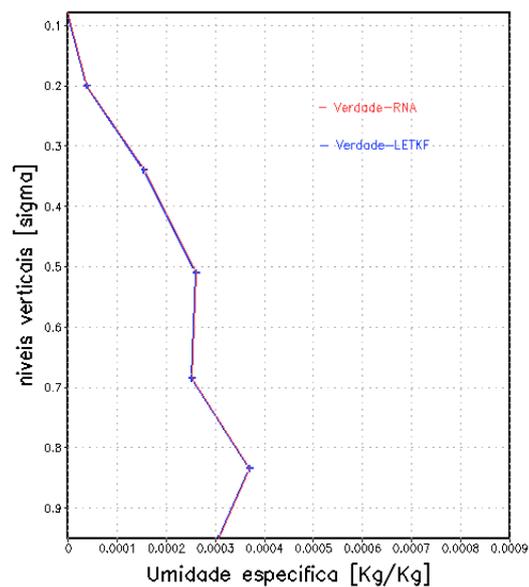


(a) erro médio quadrático

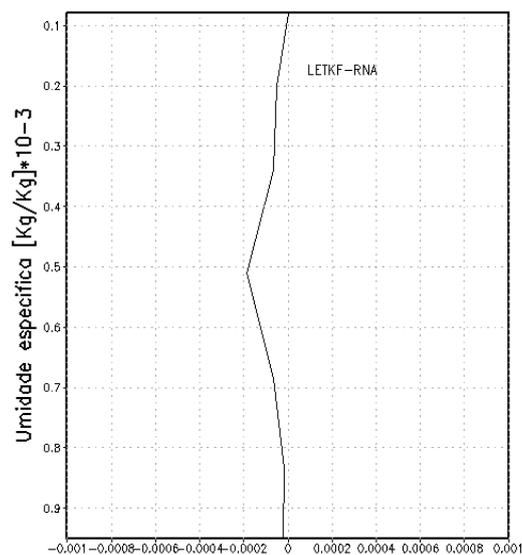


(b) diferenças (LETKF-RNA)

Figura 5.21 - Perfil do erro médio quadrático das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), no ponto (20S,87W) o perfil da temperatura em níveis sigma (a) e a diferença entre os dois perfis(b).



(a) erro médio quadrático



(b) diferenças (LETKF-RNA)

Figura 5.22 - Perfil do erro médio quadrático das análises geradas pelo método LETKF em linha (azul) e os erros das análises geradas pelo método PMC em linha (vermelha), no ponto (20S,87W) o perfil de umidade específica em níveis sigma (a) e a diferença entre os dois perfis(b).

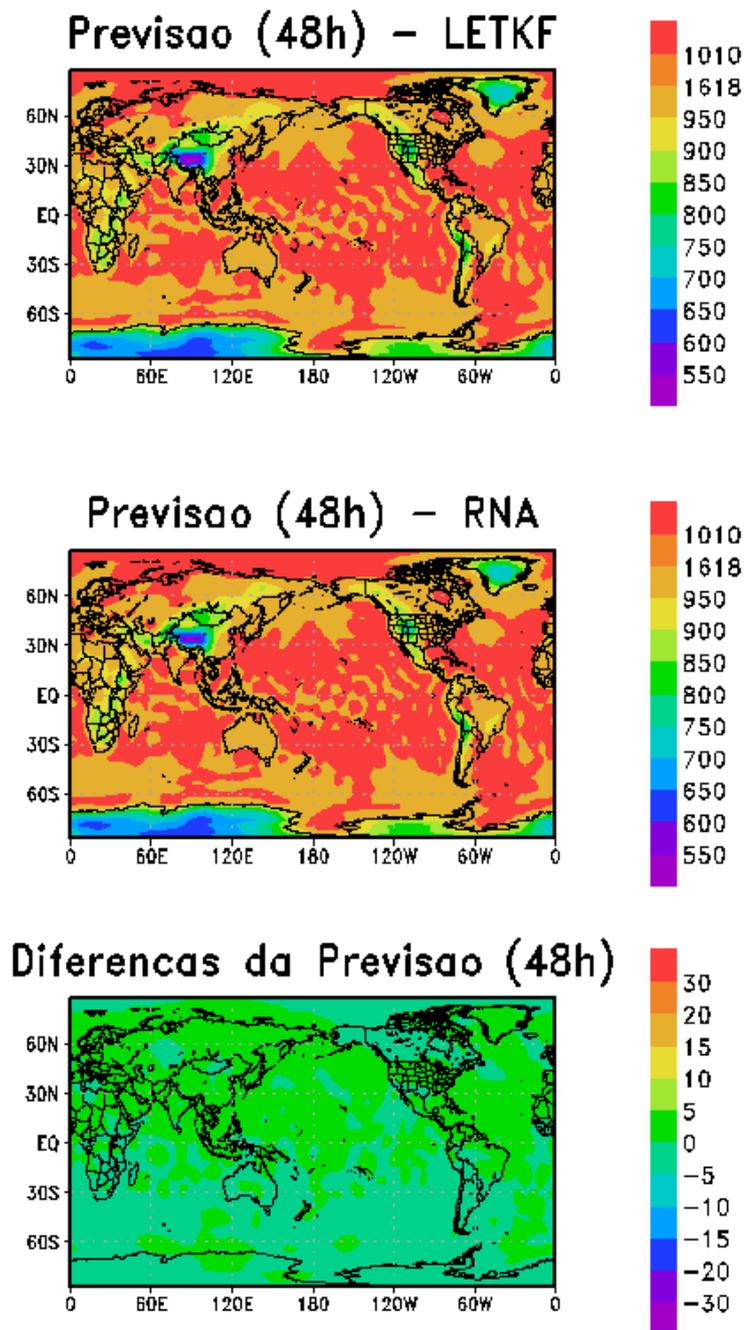
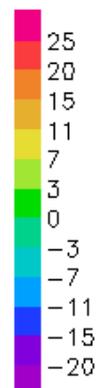
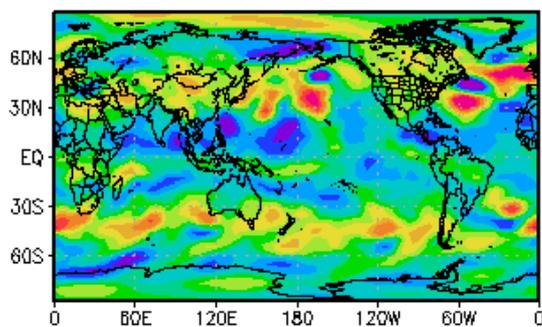
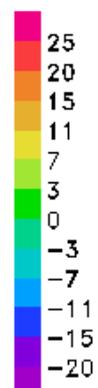
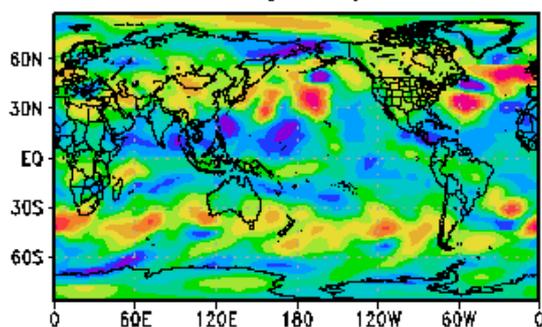


Figura 5.23 - Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de pressão [hPa] à superfície.

Previsao (48h) – LETKF



Previsao (48h) – RNA



Diferenças da Previsao (48h)

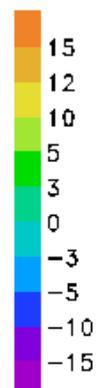
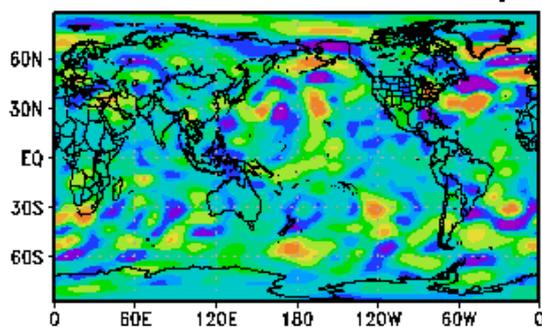


Figura 5.24 - Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões da componente do vento zonal [m/s] à superfície.

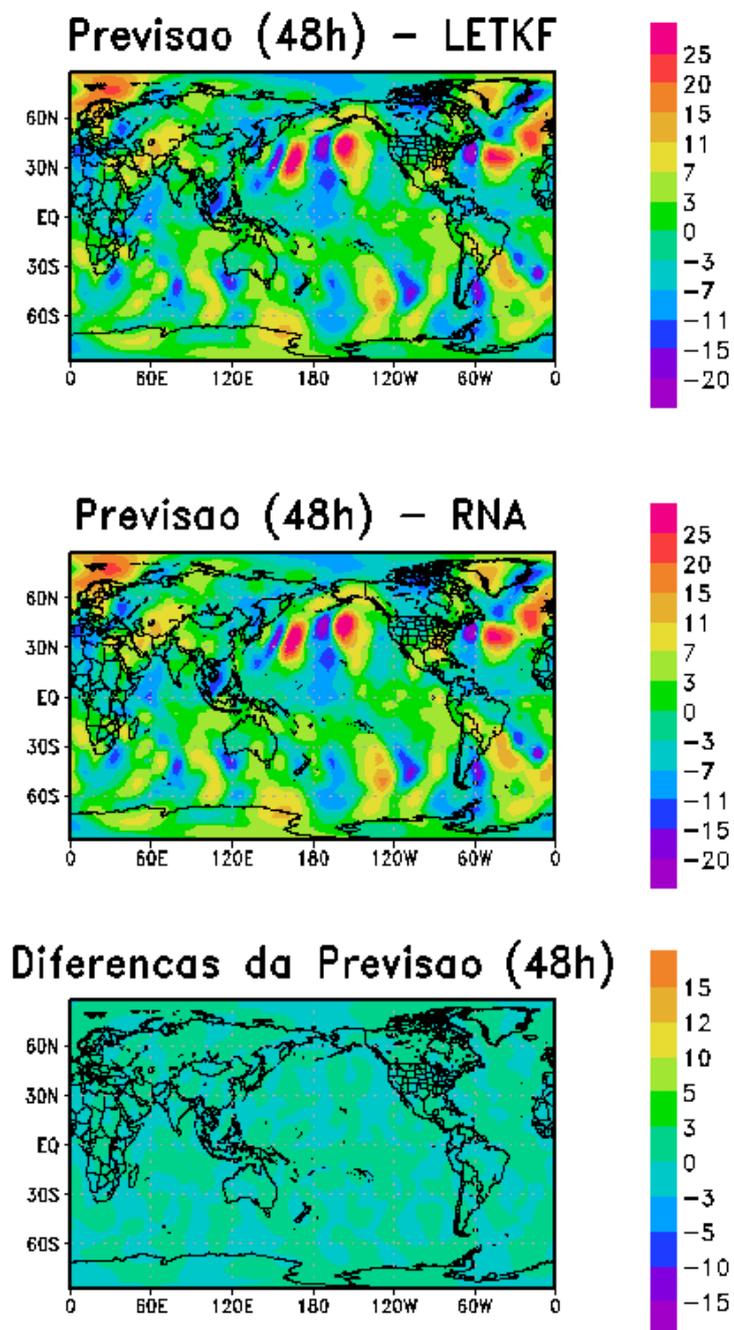


Figura 5.25 - Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões da componente do vento meridional [m/s] à superfície.

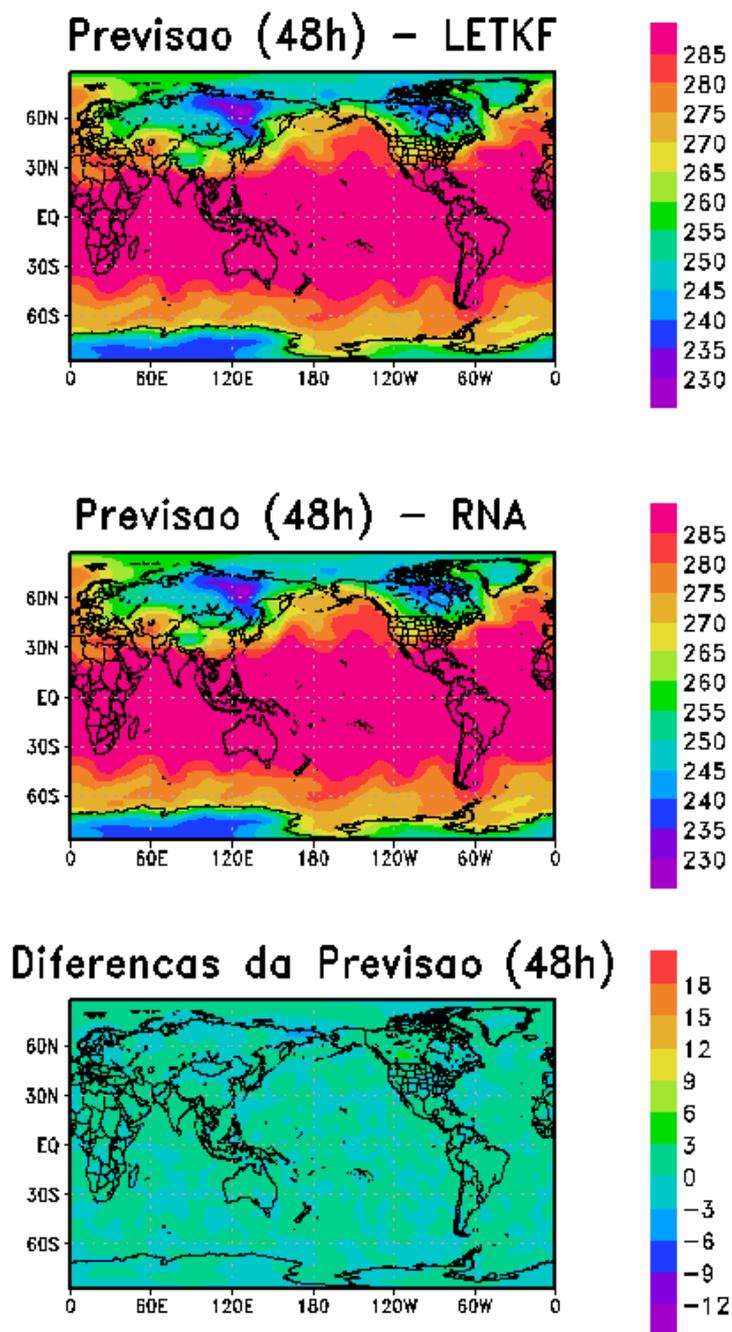


Figura 5.26 - Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de temperatura absoluta [K] à superfície.

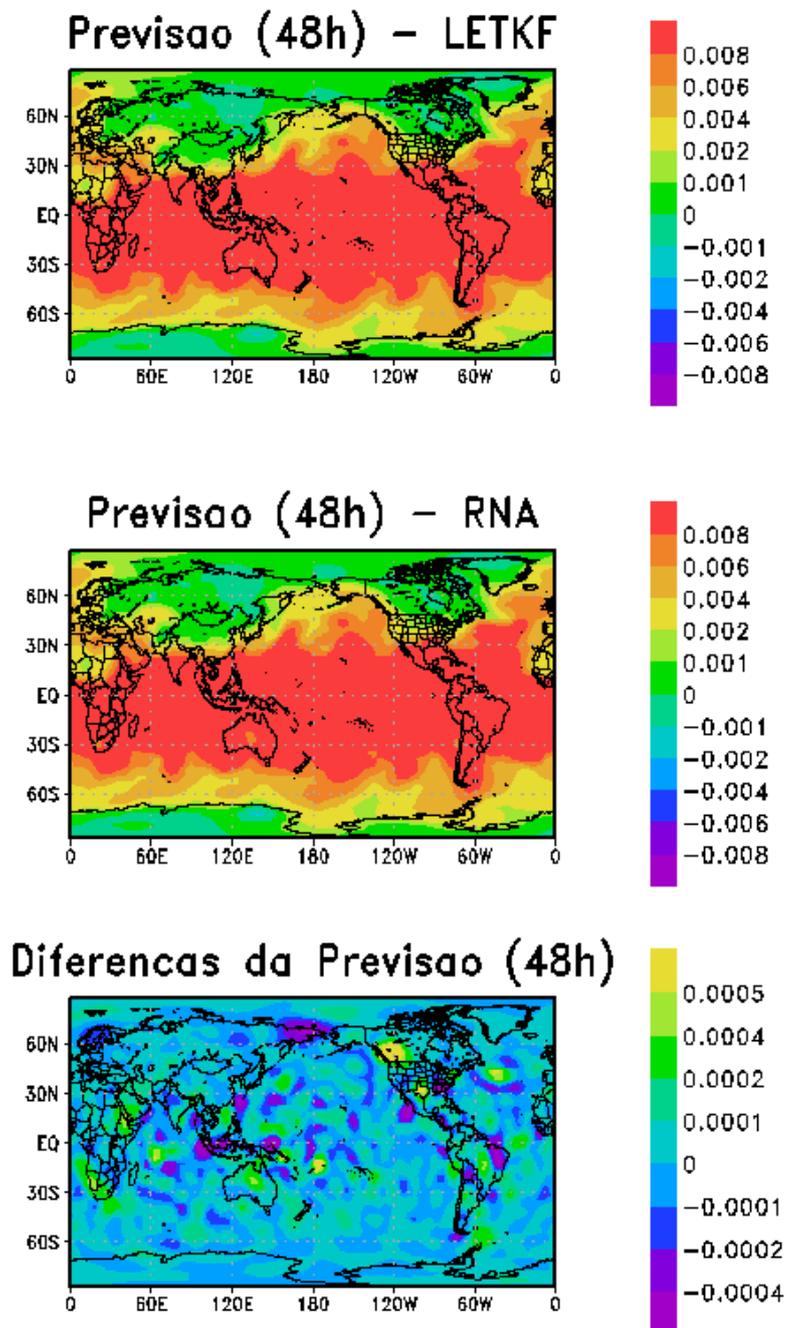


Figura 5.27 - Previsão de 48 horas a partir das análises do LETKF(a) e da RNA(b) do dia 30 de janeiro de 1985, 00 UTC e a diferenças entre as duas previsões de umidade específica [Kg/Kg] à superfície.

Os resultados são semelhantes, a similaridade da assimilação com as redes neurais e o método baseado em Filtro de Kalman por conjunto geraram praticamente o mesmo erro médio quadrático em relação ao estado verdade. As figuras denotam que o **emq** são grande para “ambos” os métodos de assimilação no final do ciclo. A verdade diverge das análises depois de um certo tempo de assimilação.

Na comparação entre as duas análises foi feita a avaliação das previsões de 48 horas, uma vez que a qualidade da análise é medida pela qualidade de suas previsões, verificada na visualização das Figuras (5.23) a (5.27). O modelo speedy executado com a condição inicial gerada pela PMC obteve previsões semelhantes às previsões geradas a partir da condição inicial do LETKF.

Os resultados e as avaliações mostram que o método para assimilação de dados com redes neurais tipo Perceptron Multicamadas *emulam* o método baseado em filtro de Kalman por conjunto (LETKF).

Com base nos resultados apresentados, as análises obtidas pelas PMCs podem gerar de condições iniciais para modelos de PNT, similares as análises obtidas pelo método de assimilação LETKF, porém, a PMC dependente deste método de assimilação toda vez que for necessário retreinar a rede.

Nesta seção avaliaremos o desempenho computacional dos algoritmos apresentados. Os dois algoritmos são sequencias, embora ambos tenham capacidade de computação de alto desempenho, podendo ser paralelizados. As versões implementadas são processadas serialmente.

5.2.1 Análise de Complexidade

A expressão *complexidade* aqui é no sentido de avaliar o número de operações escalares de pontos flutuante dos sistemas de assimilação deste trabalho. O desempenho de um sistema pode ser medido pela complexidade de seu algoritmo. As medidas numéricas de tempo de execução não serão mostradas aqui, pois a aplicação trabalha com modelos simples e o processamento é muito rápido.

A análise de complexidade é uma forma de avaliar o desempenho computacional de seus algoritmos, uma vez que o número de operações a ser executado pódé estimar o tempo a ser gasto na sua execução. Um dos objetivos desta tese é obter para o algoritmo de assimilação com PMC o desempenho computacional *superior*

ao método que está sendo emulado (LETKF), uma vez que a avaliação da análise enquanto produto da PNT é feita pelo método emulado.

Para avaliar a complexidade computacional dos métodos de assimilação de dados, será adotada a hipótese de uma PMC com m_0 entradas, L camadas escondidas de m_1 neurônios e uma camada de saída com m_2 neurônios, *treinada* para emular um Filtro de Kalman. Em cada neurônio, as m entradas alimentam a função de ativação através do produto interno

$$s = \sum_{i=1}^m \theta_i x_i \quad (5.2)$$

cuja complexidade é $O(m_1)$. Qualquer que seja a função de ativação, pode-se supor que esta função de ativação seja representada por uma série de Taylor truncada. O valor funcional de um polinômio pode ser calculado pela regra de Horner (ou *multiplicação aninhada*) (CONTE; de Boor, 1980; STOER; BULIRSCH, 2000) (páginas 33 e 44, respectivamente):

$$y = \varphi(s) = a_0 + a_1 s + a_2 s^2 + \dots + a_{n-1} s^{n-1} + a_n s^n \quad (5.3)$$

$$= a_0 + s(a_1 + s(a_2 + \dots + s(a_{n-1} + a_n s) \dots)) \quad (5.4)$$

Este algoritmo envolve com n multiplicações e n somas, levando a ordem $O(n)$.

A função de assimilação implementada pela rede tem complexidade de ordem

$$O(m_0) + (L - 1) \cdot m_1 [O(m_1) + O(n)] + m_2 [O(m_1) + O(n)].$$

Se $O(n) < O(m_1) = O(m_2)$ e $L = 1$ então o algoritmo para calcular a saída da função de assimilação de dados por redes neurais $x_n^a = F_{RNA}(x_n^f; x_n^o)$ terá complexidade $O(m^2 - 2)$ (NOWOSAD, 2001). Para uma PMC com duas entradas $m_0 = 2m_2 = 2$ e uma camada de saída $m_2 = 1$, uma camada escondida $L = 1$, com onze neurônios $m_1 = 11 \sim 10$ logo, a complexidade de uma rede é $O(m_1^2) = O(10^2)$, e para o método é $30 \times O(10^2)$, isto é, a rede tem $O[10^3]$, mesmo que o tamanho do vetor de entrada seja da ordem de 10^4 .

Segundo Goris et al. (1997), considere para o filtro de Kalman discreto, N_g variáveis do estado x , N_o variáveis do vetor observação y^o . O operados \mathbf{H} de dimensão $N_g \times N_o$, a matriz de covariância do estado \mathbf{P}^b de dimensão $N_g \times N_g$ e a matriz de covariância da observação \mathbf{R} de dimensão $N_o \times N_o$. O número de operações para o filtro de

Kalman é

$$O(2N_g^2N_o) + O(2N_gN_o^2) + (N_g^3) + O(N_o^3).$$

Se $N_g = N_o$ então o número de computações é equivalente a $O(3N_g^3)$, ou seja, para o número de observações igual ao numero de pontos de grade do modelo, a complexidade de um Filtro de Kalman discreto. Este desempenho é devido aos produtos matriciais a cada passo do algoritmo KF. No caso da assimilação com FK com o modelo SPEEDY, se temos que $N_g = N_o = 10^4$ logo, a complexidade do FK com o SPEEDY será de $O[3 \times (10^4)^3] \sim O[10^{13}]$

Para reduzir a carga computacional do filtro de Kalman, um exemplo foi a utilização da rede neural como um *substituto* para o cálculo do ganho de Kalman na equação

$$x^a - x^b = BH^T(HBH^T + R)^{-1}(y - H(x^b)), \quad (5.5)$$

onde as especificações se encontram no capítulo 3 e os resultados e teste desta melhoria do desempenho computacional estão em Cintra et al. (2010).

O ganho de Kalman é expresso por:

$$B_{t+1} W_{t+1} = P_{t+1}^b H_{t+1}, \quad \text{onde: } B_{t+1} = R_{t+1} + H_{t+1} P_{t+1}^b H_{t+1}^T. \quad (5.6)$$

Na equação acima, há 4 produtos de matrizes de dimensões $m \times m$ e de m sistemas para o cálculo das colunas da matriz W_{t+1} . O produto de matrizes tem a mesma complexidade do cálculo das colunas da matriz W_{t+1} (equivalente a complexidade de inversão de matriz): $O(m^3)$. O algoritmo de Strassen tem complexidade $O(N_g^{2.807})$ (STRASSEN, 1969). Em 1990, Coppersmith e Winograd (COPPERSMITH; WINOGRAD, 1990) apresentaram um algoritmo de complexidade $O(N_g^{2.376})$. Embora este algoritmo seja uma referência para análises teóricas, ele não é prático (diferentemente do algoritmo de Strassen), pois só existe vantagem para matrizes muito grandes, que não podem ser processadas pelos atuais computadores (ROBINSON, 2005).

Hunt et al. (2004) apresentou a complexidade computacional total do algoritmo LETKF. Hunt concluiu que o tempo de computação cresce com o número de observações ou número de variáveis do modelo, no máximo, em operações de K membros do conjunto, ou seja, o número de operações requeridas é proporcional ao numero

de K membros. Para um conjunto de previsões de K membros, as matrizes de dimensão $(K \times K)$: matriz de perturbação $\bar{\mathbf{P}}^a$ (3.3.2) e a matriz ganho \mathbf{W}^a (3.59), são da ordem de $O(k^3)$ operações por análise local ou $O(k^3p)$ operações para todo o conjunto; sendo p o número de análises locais realizadas (igual ao número de pontos de grade na abordagem mais básica).

Assim quanto maior o número de membros para o conjunto acarretará em maior custo computacional. Para uma aplicação serial, a abordagem pode ser mais rápida, dependendo da aplicação e do tamanho do conjunto. Experimentos feitos com LETKF demonstram que o método pode produzir resultados de qualidade operacional dentro de poucos minutos em um computador paralelo de tamanho razoável (HUNT et al., 2004)

5.2.2 Avaliação de Tempo

Neste trabalho, coletou-se o tempo de execução de dois ciclos de assimilação de dados. Um ciclo com análises geradas pelo LETKF e outro com análises geradas pela RNA/PMC durante o período de 1 de fevereiro de 1982 a 28 de fevereiro de 1982, quatro vezes a cada dia nos horários sinóticos(00, 06, 12 e 18 UTC). Utilizando o modelo SPEEDY e as observações conforme especificados na implementação. Lembrando que ambas implementações são seriais. Foram 112 execuções de cada análise/previsão.

Para coleta de tempo foi inserido no "script" de execução dos programas LETKF/SPEEDY, um contador de tempo antes da execução do conjunto de previsões e outros no início e final da execução do programa que gera as análises, assim pode-se excluir o tempo de execução do modelo para obter o conjunto de 30 previsões. No "script" de execução dos programas RNA/SPEEDY os contadores de tempo foram inseridos também no início e final da execução do programa que gera análise, pode-se excluir o tempo da execução do modelo para executar uma previsão de curto prazo. Os tempos foram coletados em unidades de minutos, segundos e milésimos de segundo.

Os resultados numéricos estão na Tabela (5.9) e (5.10), os resultados gráficos podem ser verificados na Figuras (5.28) e (5.29).

Avaliando a figura (5.28) verifica-se que em *milésimos de segundos* o tempo de execução do ciclo da análise LETKF com 30 membros do conjunto de previsões é

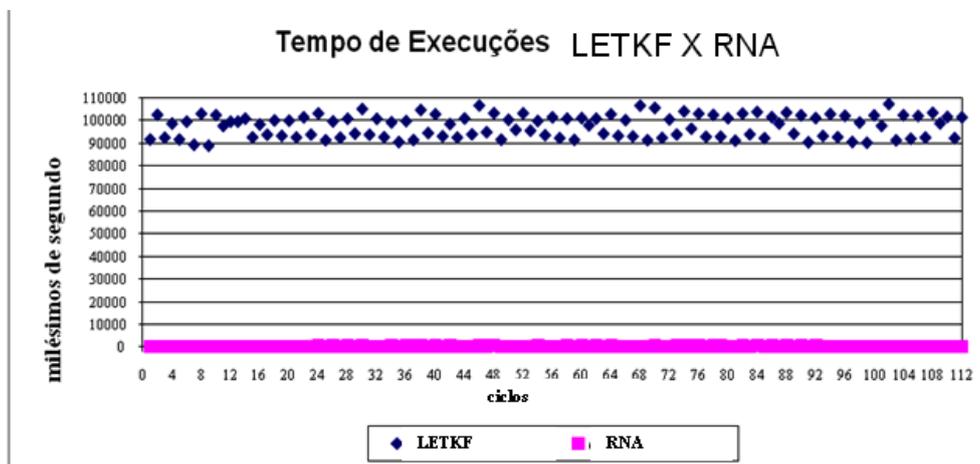


Figura 5.28 - Tempos de execução dos 112 ciclos de assimilação de dados completos com análises geradas pelo método RNA e pelo método LETKF em milésimos de segundos, para o período de 01 de fevereiro de 1982 a 28 de fevereiro de 1982 em quatro horários sinóticos

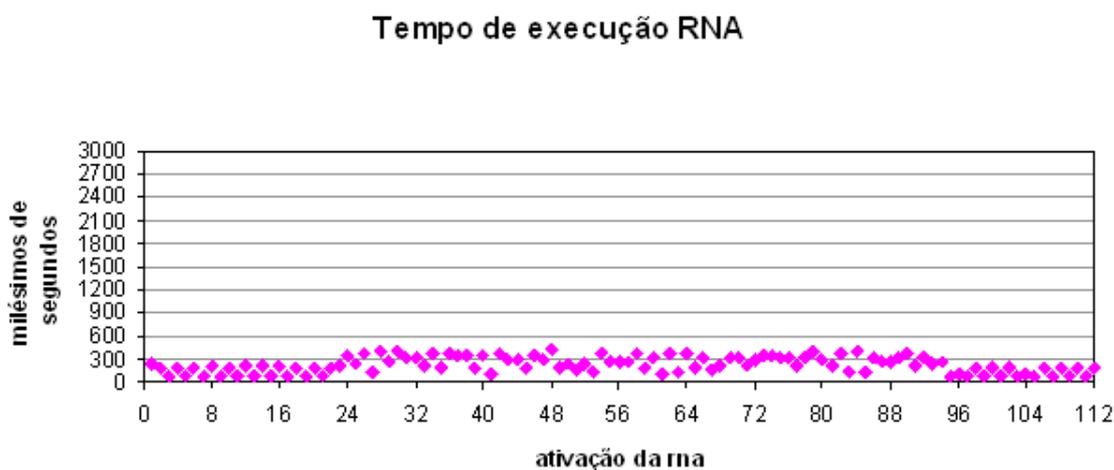


Figura 5.29 - Tempos de execução dos 112 ciclos de assimilação de dados completos com análises geradas pela PNA/PMC em milésimos de segundos, para o período de 01 de fevereiro de 1982 a 28 de fevereiro de 1982 em quatro horários sinóticos

Tabela 5.9 - Tempo Total da execução de 112 ciclos de assimilação de dados completos (análise e previsão). Tempo de análise é o tempo de 112 execuções para gerar a análise, excluindo a execução do modelo SPEEDY, para uma previsão ou para o conjunto de previsões de 30 membros.

Execução	RNA/PMC	LETKF
112 ciclos	(hh:mm:ss)	(hh:mm:ss)
início	11 : 09 : 22	11 : 14 : 41
término	11 : 12 : 15	15 : 35 : 20
tempo		
tempo Total	00:02:53	04:20:39
tempo de		
Análise	00 : 00 : 27	03 : 03 : 33

Tabela 5.10 - Tempo médio da execução de cada componente individual do ciclo de assimilação de dados. Tempo computado foi em minutos, segundos e milésimos de segundos.

Média do Tempo de uma Execução	
(hh:mm:ss:lll)	
previsão por conjunto	00 : 00 : 38 : 880
modelo SPEEDY	00 : 00 : 01 : 700
RNA	00 : 00 : 00 : 380
LETKF	00 : 01 : 49 : 150

100.000 (cem mil) vezes maior que a execução do ciclo da análise RNA/PMC, ou seja, enquanto o primeiro ciclo citado executa em 1 minuto e cinquenta segundos (90000 milsegundos) ou 2 minutos (120000 milsegundos), o segundo ciclo citado executa em menos de 1 segundo (entre 100 e 500 milésimos de segundo).

6 CONCLUSÕES E TRABALHOS FUTUROS

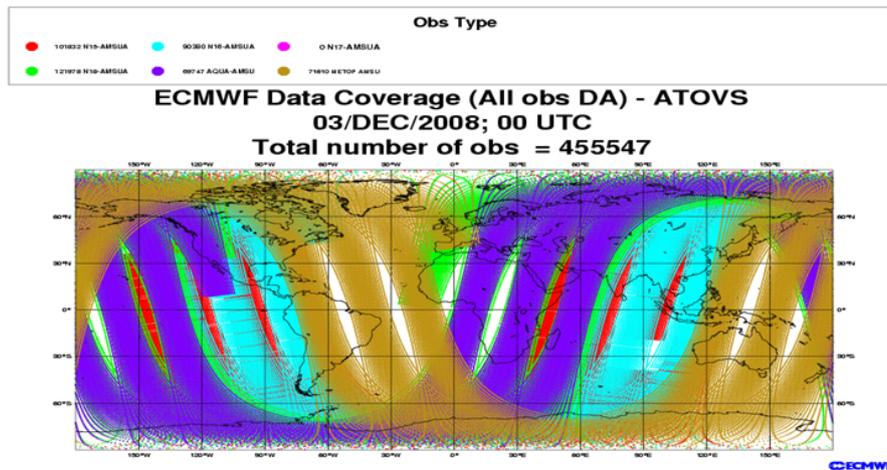
Durante as últimas décadas houve uma evolução significativa na quantidade e qualidade de observações meteorológicas, tanto em estações convencionais como em observações do sensoriamento remoto. Por outro lado, existe um grande avanço na qualidade dos modelos utilizados em previsão numérica de tempo, associada à melhoria nas parametrizações de sub-grade e com resolução cada vez mais detalhada. As previsões estão se tornando cada vez mais precisas devido às melhores análises obtidas pelos sistemas de assimilação de dados.

A assimilação de dados combina as informações de modelos numéricos e as observações meteorológicas, através de um processo matemático-computacional gerando a melhor representação possível do estado da atmosfera em um dado instante de tempo. A utilização de técnicas avançadas de assimilação de dados, como o Ensemble Kalman Filter, têm procurado vencer o desafio de utilizar a quantidade crescente de observações em matrizes da ordem de (10^7) e a quantidade crescente de pontos de grade do modelos de previsão em matrizes da ordem de (10^9) , por exemplo ver Figura (6.1)

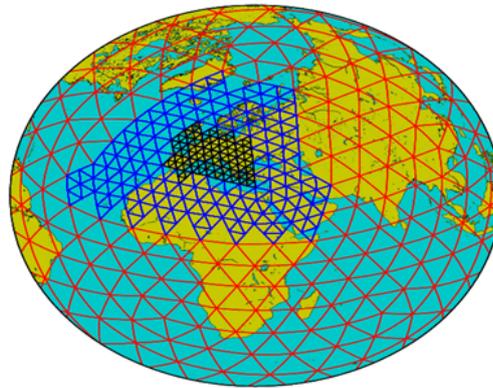
A resolução destes sistemas requer grande capacidade computacional, que faz com que estes sistemas sejam viáveis somente em supercomputadores. O que significa que a habilidade de resolver este problema depende de algoritmos especializados em arquiteturas paralelas e vetoriais que ofereçam crescente velocidade economia computacional.

Neste trabalho foi verificada a eficiência das redes Perceptron Multicamadas (PMC) num contexto de assimilação de dados atmosféricos. Conclui-se que o desenvolvimento de uma metodologia de redes PMCs foi satisfatório para obter condições iniciais de qualidade para o modelo SPEEDY, onde as redes emulam as análises obtidas com o método LETKF sem comprometer a qualidade da análise, com uma notável redução no tempo de processamento para a assimilação de dados.

As redes *aprenderam* todo o processo matemático do método de assimilação através do treinamento com suas análises. Na avaliação dos resultados, a assimilação feita pela PMC, com os mesmos dados de entrada (observação e previsão), obteve análises muito próxima à análise obtida pela assimilação feita pelo LETKF. Porém, o desempenho computacional da técnica que utiliza PMCs foi muito superior. O LETKF é



(a) observações de satélite



(b) modelos de alta resolução

Figura 6.1 - Crescimento da quantidade dados para a Assimilação

uma técnica atual de assimilação de dados, que tem a preocupação obter maior desempenho computacional, em relação ao EnKF. para atender ao desafio da crescente demanda de dados. Esta demanda de dados justifica a utilização de redes neurais artificiais em assimilação de dados atmosféricos.

A característica das redes supervisionadas de obter a aproximação da função de assimilação através do resultado de outro método já avaliado pela comunidade científica, pode ser interessante para obter resultados com desempenho computacional maior. Porém, a desvantagem da técnica é a dependência de outro método de assimilação.

A técnica proposta tem outra vantagem, por exemplo, conhecer o comportamento das previsões do modelo numérico componente do método de assimilação, pois constatou-se que a condição inicial do método proposto e do método emulado são semelhantes. Esta vantagem é correspondente à rede obter todo o conhecimento desenvolvido pelo método de assimilação emulado.

A rede em processamento serial apresentou bom desempenho computacional, porém outras vantagens da RNA/PMC é que os algoritmos são de fácil implementação, intrinsecamente paralelos e podem ser implementados em *hardware*, ou seja, neuro-computadores. A utilização de dispositivos de hardware pode aumentar ainda mais o desempenho para aplicações com modelos atuais de alta resolução e grande quantidade de observações.

O conhecimento do método a ser emulado e do modelo numérico de previsão é importante para metodologia da rede PMC, portanto nesta tese, procurou-se apresentar:

- a) No capítulo (1) e no capítulo (2) a importância da aplicação de previsão numérica tempo para a sociedade, e onde a pesquisa deste trabalho está inserida em sistemas dinâmicos caóticos, como a atmosfera;
- b) No capítulo (2) a descrição resumida da modelagem de circulação geral da atmosfera e do modelo numérico utilizado na aplicação de assimilação de dados deste trabalho;
- c) No capítulo (3) a descrição resumida da teoria de assimilação de dados e revisão bibliográfica de alguns métodos de assimilação baseados em filtro de Kalman por conjunto, que foram precursores do método de assimilação emulado pela aplicação da rede neural artificial, o *Local Ensemble Kalman Filter* (LETKF);
- d) No capítulo (4) a descrição resumida de redes neurais artificiais (RNA), enfatizando somente o método supervisionado do Perceptron Multicamadas (PMC) que foi o método utilizado na aplicação. Neste capítulo apresentou-se também a trajetória da pesquisa de RNA em assimilação de dados atmosféricos realizadas neste Instituto de Pesquisas;
- e) E finalmente, no capítulo (5) os resultados a metodologia aplicada e os resultados da técnica de utilização da RNA/PMC em assimilação de dados para o modelo de circulação geral da atmosfera SPEEDY, emulando o

método de assimilação LETKF. Apresentou-se ainda neste capítulo, uma rápida análise de desempenho computacional das duas aplicações de assimilação.

Assimilação de dados tem grande importância em modelos operacionais de previsão, para meteorologia, circulação oceânica, qualidade do ar, e em clima espacial. A técnica aplicada nesta tese pode ser utilizada nas aplicações citadas.

Com base nas pesquisas realizadas e os resultados obtidos nesta pesquisa, este método pode ser proposto para emular "qualquer" métodos de assimilação de dados avançados e obter bons e rápidos resultados (FURTADO, 2008) e é de fundamental importância para a realização de trabalhos futuros na área.

Finalmente, como linhas de pesquisa para trabalhos futuros sugere-se:

- a) continuar esta pesquisa, utilizando as mesmas PMCs propostas com o modelo SPEEDY, para assimilar observações (sintéticas) em todos pontos de grade, simulando as observações de satélite atuais conforme (6.1(a));
- b) continuar esta pesquisa, utilizando versões dos programas LETKF/SPEEDY (já existentes) paralelos e fazer análise de desempenho completa;
- c) desenvolver a versão paralela dos programas das RNA e fazer análise de desempenho completa;
- d) desenvolver um modelo para implementar observações em grade irregular (reais) para os ponto de grade do modelo (operador (H_n) da notação de assimilação de dados), visto que a técnica proposta utiliza abordagem de diminuição do espaço de busca (HÄRTER, 2004);
- e) com base no método proposto, testar as redes neurais para emular as análises operacionais do CPTEC com observações reais.

A pesquisa RNA e Assimilação é vasta e uma continuação natural deste trabalho seria aplicar redes neurais "não supervisionadas" para obter as análises como condição inicial de um modelo numérico de previsão de tempo, assim eliminaria a dependência da implementação de outro método.

Como conclusão final, mostrou-se que as RNA conseguiram emular a análise obtida com LETKF para o modelo SPEEDY, sem comprometer a qualidade da análise com uma notável redução no tempo de processamento da assimilação de dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLIGOOD, K. T.; SAUER, T. D.; YORK, J. A. **Chaos: an introduction to dynamical systems**. New York: Springer-Verlag, 1997. 602 p. 3
- ANDERSON, J. L. An ensemble adjustment kalman filter for data assimilation. **Monthly Weather Review**, v. 129, p. 2884–2903, 2001. 8, 68, 71, 72, 75
- ANDRADE, C. **Análise das trocas de energia entre modos verticais e horizontais, em resposta a fontes tropicais de calor de grande escala**. 184 p. Tese (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 1997. Disponível em:
<<http://mtc-m05.sid.inpe.br/col/sid.inpe.br/iris@1905/2005/07.29.03.34.53/doc/INPE-5681-TDI-567.pdf>>. 35
- ANDREWS, A. A square root formulation of the kalman covariance equations. **AIAA J**, v. 6, p. 1165–1168, 1968. 67
- ANGELOPOULOS, V.; PANETTA, P. V. **Science closure and enabling technologies for constellation class missions**. Berkeley: University of California, 1998. 151 p. 52
- ARAVEQUIA, J. A. A assimilação de radiâncias no filtro de kalman local por conjunto. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 15., 24-29 ago, São Paulo. **Anais...** 2008. Disponível em:
<<http://urlib.net/sid.inpe.br/mtc-m18@80/2008/07.18.13.38>>. 69, 81
- BANFIELD, N.; INGERSOLL, A.; KEPPELNE, C. A steady-state kalman filter for assimilating data from a single polar-orbiting satellite. **J. Atmos. Sci.**, v. 57, p. 737–753, 1995. 52
- BARKER, D. M. **A Short History of Atmospheric Data Assimilation**. NCAR, 2008. Boulder, Colorado. Disponível em:
<http://weather.snu.ac.kr/da_lecture/Lecture_note/1.1-history.pdf>. 48, 49, 51
- BARNES, S. G. A technique for maximizing details in numerical weather map analysis. **J. Appl. Meteor.**, v. 3, p. 396–4091, 1964. 49

- BENNETT, A. F. **Inverse methods in physical oceanography**. Cambridge, England: Cambridge University, 1992. 346 p. 4, 52, 53
- _____. **Inverse modeling of the ocean and atmosphere**. 1. ed. New York: Cambridge University Press, 2002. 234 p. ISBN 0-521-81373-5. 53
- BERGTHORSSON, P.; DOOS, B. Numerical weather map analysis. **Tellus**, v. 7, p. 329–340, 1955. 49
- BISHOP, H. C.; ETHERTON, B. J.; MAJUMDAR, S. J. Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. **Monthly Weather Review**, v. 129, p. 420–436, 2001. 8, 68, 71, 73, 78
- BJERKNES, V. Das problem der wettvorhersage, betrachtet vom stanpunkt der mechanik und der physik. **Meteor. Zeits.**, v. 21, p. 1–7, 1904. 48
- _____. **Dynamic meteorology and hydrography Part II.: Kinematics**. New York: Cargenie Institute: Gibson Bros., 1911. 48
- BONATTI, J. P. **Funções de Hough: Teoria e Utilização**. 1983. Disponível em: <<http://www.cbmet.com/cbm-files/16-27a4a0f062c989382e03344e11873ce5.pdf>>. 35
- BOURKE, W. A multilevel spectral model: I. formulation and hemispheric integrations. **Monthly Weather Review**, v. 102, p. 687–701, 1974. 37, 39
- BOUTTIER, F. A dynamical estimation of forecast error covariances in an assimilation system. **Mon. Wea. Rev.**, v. 122, p. 2376–2390, 1994. 65
- BOUTTIER, F.; COURTIER, P. **Data Assimilation concepts and methods**. Reading, UK: [s.n.], 1999. 53 p. Meteorological Training Course - ECMWF. 6, 47, 53, 56, 57, 63
- BRAGA, A.; LUDERMIR, T. B.; CARVALHO, A. C. P. L. F. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 2000. 101
- BURGERS, G.; van LEEUWEN, P.; EVENSEN, G. Analysis scheme in the ensemble kalman filter. **Monthly Weather Review**, v. 126, p. 1719–1724, 1998. 67, 69, 71

CAMPOS VELHO, H. F. **Matriz Não-modal em integração e inicialização num modelo barotrópico e um estudo numérico da dispersão vertical turbulenta**. 97 p. Tese (Doutorado em Engenharia) — Universidade Federal do Rio Grande do Sul(UFRG), Porto Alegre, 1992. 22, 102

CAMPOS VELHO, H. F.; CINTRA, R. S.; FURTADO, H. C. **Curso de Assimilação de Dados**. INPE, 2007. 67 transpaências. Disponível em: <http://www.lac.inpe.br/~haroldo/Curso-DataAssimilation/Curso_Assim_Dados-INPE-2.pdf>. 6, 8, 57

CHARNEY, J. G. **Dynamical forecasting by numerical process**. Boston, MA: [s.n.], 1951. Compendium of Meteorology. American Meteorological Society. 49

CINTRA, R. **Implementação do sistema estatístico de assimilação de dados em espaço físico para o modelo global do CPTEC**. 184 p. Tese (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2004. Disponível em: <<urlib.net/sid.inpe.br/jeferson/2004/08.31.13.48>>. 11, 14, 48, 70

CINTRA, R. S.; CAMPOS VELHO, H. F.; TODLING, R. Nova implementação em filtro do kalman estendido para assimilação de dados com redes neurais. **Learning and Nonlinear Model**, v. 7, p. 30–37, 2010. 10, 104, 146

CONTE, S. D.; de Boor, C. **Elementary numerical analysis::** An algorithmic approach. [S.l.]: McGraw-Hill, 1980. 3rd Edition. 145

COPPERSMITH, D.; WINOGRAD, S. Matrix multiplication via arithmetic progressions. **Journal of Symbolic Computation**, v. 9, p. 251–280, 1990. 146

CRESSMAN, G. P. An operational objective analysis system. **Mon. Wea. Rev.**, v. 87, p. 367–374, 1959. 49

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Math. Control Signals Systems**, v. 2, p. 303–314, 1989. 99

DAHLQUIST, G.; BJORCK, A. **Numerical methods**. Englewood Cliffs: Prentice–Hall, 1969. 32

DALEY, R. **Atmospheric data analysis**. New York, USA: Cambridge University Press, 1991. 471 p. ISBN 9780521458252. 7, 8, 46, 48, 49, 63, 65, 66

DELSOLE, T. **A History of Data Assimilation**. Center for Ocean-Land-Atmosphere Studies, 2009. Calverton, MD. Disponível em: <http://mason.gmu.edu/~tdelsole/AdvStat/da_history_multipage.pdf>. 51

DIMET, F. L.; TALAGRAND, O. Variational algorithms for analysis and assimilation of meteorological observations theoretical aspects. **Tellus.**, v. 39A, p. 97–110, 1986. 64

DOTY, B. **GrADS Users Guide**. 1998. Disponível em: <<http://iges.org/grads/>>. 119

ELIASSEN, A. **Provisional report on calculation of spatial covariance and autocorrelation of the pressure field**. 1954. 11 p. Videnskaps-Akademiets Institutt for Vaer-og Klimaforskning, Report No. 5. 50

EVENSEN, G. Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. **J. Geophys. Res**, v. 97, p. 17905–17924, 1992. 65, 68

_____. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. **J. Geophys. Res**, v. 99, p. 10143–10162, 1994. 8, 52, 67, 69, 71

FAUSETT, L. **Fundamentals of neural networks**: architectures, algorithms, and applications. New Jersey: Prentice Hall, 1994. 89

FJORTOFT, R. On a numerical method of integrating the barotropic vorticity equation. **Tellus**, v. 4, p. 179–94, 1952. 48

FUKUSHIMA, K. **Self-organizing multilayered neural network**. 1975. 15–22 p. 95

FURTADO, H. C. **Redes neurais e diferentes métodos de assimilação de dados em dinâmica não linear**. 125 p. Tese (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2008. Disponível em: <urlib.net/sid.inpe.br/mtc-m17@80/2008/02.07.10.49>. 9, 48, 103, 154

GANDIM, L. S. **Objective Analysis of Meteorological Fields**. Jerusalem: Gidrometeorologicheskoe Izdatel'stvo, 1963. 242 p. Translated from Russian in 1965 by Israel Program for Scientific Translations. 49, 50

- GARDNER, M.; DORLING, S. Artificial neural networks, the multilayer perceptron.: A review of applications in the atmospheric sciences. **Atmospheric Environment.**, v. 32 114/15, p. 2627–2636, 1998. 101, 102, 110
- GELB, A. **Applied optimal estimation. the analytic sciences corporation.** Cambridge – Massachusetts: MIT Press, 1974. 374 p. 63, 66
- GHIL, M. Advances in sequential estimation for atmospheric and oceanic flows. **J. Meteor. Soc.**, Japan, v. 75, p. 289–304, 1997. 63
- _____. **The Essence of Data Assimilation or Why Combine Data with Models.** University of California, 2007. Inaugural Lecture WMO. Disponível em: <<http://www.atmos.ucla.edu/tcd/MG/REF/WMOinaug.ps>>. 52
- GHIL, M.; MALANOTTE-RIZZOLI, P. Data assimilation in meteorology and oceanography. **Adv. Geophys.**, v. 33, p. 141-266, 1991. 63
- GILCHRIST, B.; CRESSMAN, G. P. An experiment in objective analysis. **Tellus**, v. 6(4), p. 309–318, 1954. 49
- GORIS, M. J.; GRAY, D. A.; MARCELS, I. M. Y. Reducing the computational load of a kalman filter. **IEEEletronics Leterrs**, v. 33(12), p. 1539–1541, 1997. 145
- GOTTLIEB, D.; ORSZAG, S. **Numerical analysis of spectral methods: theory and applications.** Philadelphia: SIAM–CBMS, 1977. 33
- GROSSBERG, S. **Birth of a learning law.** [S.l.]: INNS/ENNS/JNNS Newsletter, 1998. 1–4 p. 95
- HADAMARD, J. **J. Lectures on Cauchy’s Problem.** [S.l.]: Dover Press, New York, 1952. 46
- HALTINER, G. J.; WILLIAMS, R. T. **Numerical prediction and dynamic meteorology.** 2. ed. [S.l.]: John Wiley and Sons, 1980. 477 p. ISBN 0471059714. 9, 19, 22
- HAMILL, T. **Ensemble–based Atmospheric Data Assimilation.** Boulder, Colorado, USA: [s.n.], 2004. University of Colorado and NOAA–CIRES Climate Diagnostics Center. 69

HAMILL, T. M.; WHITAKER, J. S.; SNYDER, C. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. **Mon. Wea. Rev.**, v. 129, p. 2776–2790, 2001. 63, 68, 72

HÄRTER, F.; CAMPOS VELHO, H. Recurrent and feedforward neural networks trained with cross correlation applied to the data assimilation in chaotic dynamic. **Revista Brasileira de Meteorologia**, v. 20, p. 411–420, 2005. 103

HÄRTER, F. P. **Redes neurais recorrentes aplicadas a assimilação de dados em dinâmica não-linear**. 138 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2004–12–16 2004. Disponível em: <urlib.net/sid.inpe.br/jeferson/2005/03.09.14.05>. xv, 9, 46, 57, 102, 103, 104, 106, 154

HÄRTER, F. P.; REMPEL, E. L.; CAMPOS VELHO, H. F. de; CHIAN, A. Application of artificial neural networks in auroral data assimilation. **Journal of Atmospheric and Solar – Terrestrial Physics**, v. 70, n. 10, p. 1243–1250, março 2008. Disponível em: <[10.1016/j.jastp.2008.03.018](http://dx.doi.org/10.1016/j.jastp.2008.03.018)>. 103

HAYKIN, S. **Redes neurais princípios e prática**. Porto Alegre: Editora Bookman, 2001. ISBN 85–7307–718–2. 85, 95, 97, 98, 101, 103

_____. **Adaptive filter theory**. New Delhi–India: Dorling Kinderley, 2007. ISBN 81-317-0869-1. 91, 95, 98, 99, 100, 112

HEBB, D. O. **The organization of behavior::** A neuropsychological theory. New York: Wiley, 1949. ISBN 978-0805843002. 95

HELD, I.; SUAREZ, M. A proposal for the intercomparison of dynamical cores of atmospheric general circulation models. **Bull. Am. Meteorol. Soc.**, v. 75, p. 1825–1830, 1994. 37

HERDIES, D.; ARAVÉQUIA, J.; FERREIRA, S.; ANDREOLI, R.; SAPUCCI, L.; MATTOS, J. **A Assimilação de dados no CPTEC/INPE**. INPE, 2007.

Disponível em:

<<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/09.03.13.27>>. 81

HOKE, J.; ANTHES, R. The initialization of numerical models by a dynamic relaxation technique. **Monthly Weather Review**, v. 104, p. 1551–1556, 1976. 49

HOLM, E. V. **Lecture notes on assimilation algorithms**. European Centre for Medium-Range Weather Forecasts: [s.n.], 2008. 30 p. 54

HOLTON, J. R. **An introduction to dynamic meteorology**. London, UK: Academic Press, 1992. 497 p. ISBN 0-12-354354355. 22

HOPFIELD, J. Neural networks and physical systems with emergent collective computational abilities. In: NATIONAL ACADEMY OF SCIENCES OF THE U.S.A. **Proceedings...** [S.l.], 1982. p. 2554–2558. 90

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, p. 359–366, 1989. 101

HOUTEKAMER, P. L.; MITCHELL, H. L. Data assimilation using an ensemble kalman filter technique. **Monthly Weather Review**, v. 126, p. 796–811, 1998. 8, 52, 67, 68, 69, 70

HOUTEKAMER, P. L.; MITCHELL, H. L.; PELLERIN, G.; BUEHNER, M.; CHARRON, M.; SPACEK, L.; HANSEN, B. Atmospheric data assimilation with the ensemble kalman filter: results with real observations. **Monthly Weather Review**, v. 133, p. 605–620, 2005. 69

HUNT, B.; KALNAY, E.; KOSTELICH, E. J.; OTT, E.; PATIL, D. J.; SAUER, T.; SZUNYOGH, I.; YORKE, J. A.; ZEMIN, A. V. Four-dimensional ensemble kalman filtering. **Tellus**, v. 56A, p. 273–277, 2004. 8, 71, 73, 78, 80, 146, 147

IDE, K.; COURTIER, P.; GHIL, M.; LORENC, A. Unified notation for data assimilation: Operational, sequential and variational. **J. Meteor. Soc. Japan**, v. 75, p. 181–189, 1997. 53

INNOCENTINI, V.; BRANDÃO, I.; ARANTES, F.; PRADO, S. Possíveis dificuldades que acompanharão a viagem de Vasco da Gama Índia. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 11., Rio de Janeiro. **Anais...** 2000. Disponível em: <<http://www.cbmet.com/index.php>>. 16

JAZWINSKI, A. H. **Stochastic processes and filtering theory**. London, UK: Academic Press, 1970. 400 p. ISBN 0486462749. 7, 63, 66

JI, M.; LEETMAA, A. Impact of data assimilation on ocean initialization and el Niño prediction. **Monthly Weather Review**, v. 125, p. 742–753, 1997. 52

KALMAN, R. E. A new approach to linear filtering and prediction problems. **Trans. of the ASME–Journal of Basic Engineering**, v. 82(Series D), p. 35–45, 1960. 7, 49, 63

KALMAN, R. E.; BUCY, R. S. New results in linear filtering and prediction theory. **Trans. of the ASME–Journal of Basic Engineering**, v. 83(Series D), p. 95–108, 1961. 63

KALNAY, E. **Atmospheric modeling, data assimilation and predictability**. 2. ed. New York: Cambridge University Press, 2003. 2, 19, 22, 33, 35, 45, 48, 55, 57

_____. 2007. Personal Home Page. Disponível em:
<<http://www.umd.edu/caos/~ekalnay>>. 81

_____. **Intensive Course on Data Assimilation**. Buenos Aires, Argentina: School of Sciences of the University of Buenos Aires and Center for Atmospheric and Ocean Research CIMA/CONICET, 27 October-7 November 2008. Disponível em: <<http://4dvarenkf.cima.fcen.uba.ar/course/index.php>>. 79, 106

KALNAY, E.; TOTH, Z. Removing growing errors in the analysis. In: CONFERENCE ON NUMERICAL WEATHER PREDICTION, 10, Portland. **Proceedings ...** [S.l.], 1994. p. 212–215. 73, 75

KLEINKE, M. U. **Sistemas Complexos e o Caos Determinístico**. Unicamp, 2007. Disponível em:
<http://www.ifi.unicamp.br/~kleinke/f082/caos_deterministico.pdf>. 2

KOHONEN, T. Self-organizing formation of topologically correct feature maps. **Biological Cybernetics**, v. 43(1), p. 59–69, 1982. 95

KOLMOGOROV, A. N. The local structure of turbulence in incompressible viscous fluid for very large reynoldsnumbers. In: DOKL. AKAD. NAUK SSSR. **Proceedings....** [S.l.]: USSR Academy of Sciences, 1941. v. 30, p. 301–305. 50

KOVACS, Z. **Redes neurais artificiais: princípio e aplicações**. São Paulo: Collegium Cognitio, 1996. 85

LACROIX, M. **Histórias em Português**. Lisboa: Instituto piaget, 1999. Disponível em: <contadoresdestorias.wordpress.com/2007/11/28/a-ciencia-e-a-fragilidade>. 1

LAURITZEN, S. L. Time series analysis in 1880. **Int. Statist. Rev.**, v. 49, p. 319–331, 1981. A discussion of contributions made by T. N. Thiele. 51

LEITH, C. E. **Predictability in theory and practice**. [S.l.]: Academic Press, 1983. 397 p. Chapter 13 in *Large-scale Dynamical Processes in the Atmosphere*. B. J. Hoskins and R. P. Pearce, Eds. 67

LES méthodes nouvelles de la mécanique céleste. Paris: Ed. Blanchard, 1993. Poincaré, H., Vol 1-3. 2

LIAQAT A. FUKUHARA, M.; TAKEDA, T. Application of neural network collocation method to data assimilation. **Comput. Phys. Commun.**, v. 141, p. 350–364, 2001. 9

LORENC, A. A global three-dimensional multivariate statistical interpolation scheme. **Mon. Weather Rev.**, v. 109, p. 701–721, 1981. 49, 50

_____. Analysis methods for numerical weather prediction. **Quart. J. Roy. Meteor. Soc.**, v. 112, p. 1177–1194, 1986. 48, 51, 53, 55, 58, 60

_____. **Atmospheric Data Assimilation**. Bracknell-England: Meteorological Office, 1995. 61

LORENC, A. C. **A Bayesian view of Data Assimilation and Quality Control of Observations**. Buenos Aires, Argentina: [s.n.], 2008. Intensive Course on Data Assimilation. 61

LORENZ, E. N. Deterministic nonperiodic flow. **Journal of Atmospheric Sciences**, v. 20(2), p. 130–141, 1963. 2

LYNCH, P. **DYNAMO – A one dimensional primitive equation model**. Dublin: Irish Meteorological Service, 1984. Tech. Note No. 44. Disponível em: <<http://mathsci.ucd.ie/~plynch/Publications/DYNAMO.pdf>>. 102

MALANOTTE-RIZZOLI, P. **Modern approaches to data assimilation in ocean modeling**. Amsterdam: Elsevier, 1996. 455 p. 52

MARINGOLO, M. A.; MOURA, A. **Fundamentos de dinâmica aplicados a meteorologia e oceanografia**. São José dos Campos: INPE – UNIVAP, 1998. 484 p. ISBN 9780521458252. 13

- MAYBECK, P. S. **Stochastic models, estimating, and control**. New York: Academic Press, 1979. 63, 67
- MCCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115–133, 1943. 85, 90
- MENDEL, J. M.; MCLAREN, R. W. **Reinforcement learning control and pattern recognition systems**. New York: Academic Press, 1970. 287–318 p. 93
- MENKE, W. **Geophysical data analysis: discrete inverse theory**. New York: Academic Press, 1984. 289 p. 58
- MINSKY, M.; PAPERT, S. **Perceptrons: An introduction to computational geometry**. Cambridge, MA: Atmospheric Environment., 1969. 90
- MIYOSHI, T. **Ensemble Kalman filter experiments with a primitive–equation global model**. 197 p. Tese (Doutorado) — University of Maryland, College Park, Maryland, USA, 2005. Disponível em: <<http://www.atmos.umd.edu/~miyoshi/docs/dissertation.pdf>>. 40, 106, 108
- MOLTENI, F. Atmospheric simulations using a gcm with simplified physical parametrizations. i: model climatology and variability in multi-decadal experiments. **Climate Dyn.**, v. 20, p. 175–191, 2003. xi, xiii, 38, 40
- MOURA, A. D. **Von Neumann e a previsão numérica de tempo e clima**. Estud. av, 1996. Doi: 10.1590/S0103-40141996000100021. Disponível em: <<http://www.scielo.br/scielo.php>>. 18, 19
- NOWOSAD, A.; NETO, A. R.; CAMPOS VELHO, H. Data assimilation in chaotic dynamics using neural networks. In: INTERNATIONAL CONFERENCE ON NONLINEAR DYNAMICS, CHAOS, CONTROL AND THEIR APPLICATIONS IN ENGINEERING SCIENCESA, Campos do Jordão. **Anais...** [S.l.], 2000. p. 212–221. 9, 102
- NOWOSAD, A. G. **Novas abordagens de assimilação de dados meteorológicos**. 120 p. p. (INPE–8481–TDI/779). Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais, 2001–03–14 2001. xv, 9, 103, 104, 106, 112, 145

NOWOSAD, A. G.; NETO, A. R.; Campos Velho, H. Data assimilation using an adaptative kalman filter and laplace transform. In: WORKSHOP ON PHYSICS OF THE PLANETARY BOUNDARY LAYER AND DISPERSION PROCESS MODELLING, 1., 1999, Santa Maria-RS. **Anais...** [S.l.], 2000. p. 212–221. 102

OTT, E.; HUNT, B. R.; SZYNIUGH, I.; ZIMIN, A. V.; KOSTELICH, E. J.; CORAZZA, M.; KALNAY, E.; PATIL, D. J.; YORK, J. A local ensemble kalman filter for atmospheric data assimilation. **Tellus A**, v. 56, p. 415–428, 2004. 8, 68, 75, 76, 77, 78

PANOFSKY, H. A. Objective weather map analysis. **J. Meteor.**, v. 6, p. 386–392, 1949. 49

PARRISH, D. F.; DERBER, J. The national meteorological center's spectral statistical-interpolation analysis system. **Monthly Weather Review**, v. 120, p. 1747–1763, 1992. 57

POTTER, J. **W matrix augmentation**. Cambridge, Massachusetts, 1964. M. I. T. Instrumentation Laboratory Memo SGA 5–64. 67, 72

PROBABILITY. Chicago: Encyclopaedia Britannica, 1990. Laplace, P. S. 1

RANDALL, D. A. **An Introduction to Atmospheric Models**. 2009. Course Outline AT604 – Colorado State University. Disponível em: <<http://kiwi.atmos.colostate.edu/group/dave/at604.html>>. 21, 30, 31, 32, 37

_____. **An Introduction to The General Circulation of the Atmosphere**. 2009. Course Outline AT605 – Colorado State University. Disponível em: <<http://kiwi.atmos.colostate.edu/group/dave/at604.html>>. 19, 20

RICHARDSON, L. **Weather prediction by numerical process**. Cambridge: Cambridge University Press, 1922. Reprinted by Dover(1965). 48, 49

ROBERT, A. The integration of a low-order spectral form of the primitive meteorological equations. **Japan. Meteorol. Soc.**, v. 44, p. 237–245, 1966. 38

ROBINSON, S. Utoward an optimal algorithm for matrix multiplication. **SIAM News**, v. 38(9), 2005. 146

ROSEMBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychol. Rev.**, v. 65, p. 386–408, 1958. 90

RUMELHART, D.; HINTON, G.; WILLIAMS, R. **Learning internal representations by error propagation**. Cambridge, MA: MIT Press, 1986. 318-62 p. 94, 98

SASAKI, Y. An objective analysis based on the variational method. **J. Meteorol. SOC. Japan**, v. 36, p. 2884–2903, 1958. 51

_____. Some basic formalisms in numerical variational analysis. **Mon. Wea. Rev.**, v. 98, p. 875–883, 1970. 51

SATYAMURTY, P. **Rudimentos de meteorologia dinâmica**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais (INPE), 2002. 154 p. p. ISBN 85-17-00019-6. 22

SAUCIER, W. **Principles of meteorological analysis**. Chicago: University of Chicago Press, 1955. 438 p. 48

SILVEIRA, F. L. **Determinismo, Previsibilidade e Caos**. Porto Alegre – RS: [s.n.], 2008. Disponível em:
<www.periodicos.ufsc.br/index.php/fisica/article/viewFile/7279/670>. 2

STAUFFER, D. R.; SEAMAN, N. L. Use of 4-d data assimilation in a limited area mesoscale model. part 1: experiments with synoptic scale data. **Mon. Wea. Rev.**, v. 118, p. 1250–1277, 1990. 49

STOER, J.; BULIRSCH, R. **Introduction to numerical analysis**. [S.l.]: Springer-Verlag, 2000. 3rd edition. 145

STRASSEN, V. Gaussian elimination is not optimal. **Numerische Mathematik**, v. 13, p. 354–356, 1969. 146

SZUNYOGH, Y.; KOSTELICH, E.; GYARMART, G.; KALNAY, E.; HUNT, B.; OTT, E.; SATTERFIELD, E.; YORKE, J. A local ensemble transform kalman filter data assimilation system for the ncep global model. **Tellus**, v. 60A, p. 113–130, 2008. 76

TALAGRAND, O. Assimilation of observations, an introduction. **J. Meteor. Soc. Japan Issue**, v. 1B, p. 191209, 1997. 63, 66

TANG, Y.; HSIEH, W. W. Coupling neural networks to incomplete dynamical systems via variational data assimilation. **Mon. Wea. Rev.**, v. 129, p. 818–834, 2001. 9, 102

TANG, Y.; HSIEH, W. W.; TANG, B.; HAINES, K. A neural network atmospheric model for hybrid coupled modelling. **Climate Dynamics**, Springer Berlin / Heidelberg, v. 17, Numbers 5-6, p. 445–455, March 2001. 9

TARANTOLA, A.; VALETTE, B. Inverse problems = quest for information. **J. Geophys.**, v. 50, p. 159170, 1982. 53

_____. **Inverse problem theory. methods for data fitting and model parameter estimation**. Amsterdam: Elsevier, 1987. 53

TIPPETT, M. K.; ANDERSON, J. L.; BISHOP, C. H.; HAMILL, T. M.; WHITAKER, J. S. Ensemble square root filters. **Mon. Wea. Rev.**, v. 131, p. 1485–1490, 2003. 72

TODLING, R. **Estimation theory and foundations of atmospheric data assimilation**. 1999. Disponível em: <http://citeseer.ist.psu.edu/todling99estimation.html>. 22, 57

van den BOOGART, H.; MYNETT, A. Dynamic neural networks with data assimilation. **Hydrological Processes**, v. 18(10), p. 1959–1966, 2004. Disponível em: <http://www3.interscience.wiley.com>. 9

van Leeuwen, P. J.; EVENSEN, G. Data assimilation and inverse methods in terms of a probabilistic formulation. **Mon. Wea. Rev.**, v. 124, p. 28982913, 1996. 46

VIJAYKUMAR, N. L.; Campos Velho, H. F.; STEPHANY, S.; PRETO, A. J.; NOWOSAD, A. G. A neural network implementation for data assimilation using mpi, application of high performance computing in engineering. In: BREBIA, C. A.; MELLI, P.; ZANASI, A. . (Ed.). **Application of high performance computing in engineering**. Southampton: WIT Press, 2002. p. Section 5, 211–220. Disponível em: <http://urlib.net/sid.inpe.br/mtc-m17@80/2007/08.14.13.09>. 102

WEI, M.; TOTH, Z.; WOBUS, R.; ZHU, Y.; BISHOP, C.; WANG, X. Ensemble transform kalman filter-based ensemble perturbations in an operational global prediction system at ncep. **Tellus**, v. 58A, p. 28–44, 2006. 73

WHITAKER, J. S.; HAMILL, T. M. Ensemble data assimilation without perturbed observations. **Monthly Weather Review**, v. 130, p. 1913–1924, 2002. 8, 68, 71, 72

WIDROW, B.; HOFF, M. . Adaptive switching circuits. **IRE WESCON Conv. Record**, p. 96–104, 1960. Pt. 4. 94, 95

WIENER, N. **Extrapolation, interpolation, and smoothing of stationary time series**. [S.l.]: The MIT Press, 1964. 176 p. ISBN 0262730057. 50

GLOSSÁRIO

Balanco Geostrófico - é um balanço entre as forças de gradiente, pressão e Coriolis. Este balanço é viável somente quando outras forças como centrífuga e atrito são ausentes ou desprezíveis.

Clima - constitui o estado médio e o comportamento estatístico da variabilidade dos parâmetros do tempo (temperatura, chuva, vento, etc.) sobre um período, suficientemente, longo de uma localidade. O período recomendado é de 30 anos.

Força de Coriolis - força aparente que atua sobre um elemento do fluido em deslocamento sobre um planeta em rotação. A força de Coriolis atua na direção perpendicular ao movimento para a esquerda do movimento horizontal no hemisfério Sul. Ela é nula sobre o Equador, negativa no hemisfério Sul e positiva no hemisfério Norte, atingindo valores extremos nos pólos.

Geopotencial - energia potencial de uma parcela do ar de massa unitária, sendo que o seu valor referencial de zero encontra-se, em geral, no nível médio do mar.

Geostrófico vento que sopra paralelo às isóbaras.

Instabilidade Baroclínica - vulnerabilidade de um escoamento planetário representado pela corrente de jato a uma perturbação de escala sinótica e/ou subsinótica. A instabilidade se deve ao gradiente horizontal de temperatura ou cisalhamento vertical do escoamento. O desenvolvimento de sistemas sinóticos de médias latitudes como ciclones extratropicais se deve a este mecanismo.

Instabilidade Barotrópica - vulnerabilidade do corrente de jato para perturbação de escala sinótica devido à variação da vorticidade na zona do jato. A energia cinética da perturbação aumenta recebendo a energia cinética do escoamento básico.

Isóbara - superfície na qual a pressão atmosférica é constante. Em um plano horizontal as isóbaras são linhas que passam pelos pontos de igual pressão. Isto é, um lado da isóbara, a pressão é maior do que no outro lado.

Previsão de Clima previsão numérica das condições meteorológicas futuras para um e seis meses, através de modelagem estatística de grande escala e de baixa resolução, expressa através de desvios positivos ou negativos em relação ao comportamento médio passado.

Previsão de Tempo descrição detalhada de ocorrências futuras esperadas. A previsão do tempo inclui o uso de modelos objetivos baseados em certos parâmetros atmosféricos, a habilidade e experiência de um meteorologista. Também chamada de prognóstico.

Tempo conjunto de condições atmosféricas e fenômenos meteorológicos que afetam a biosfera e a superfície terrestre em um dado momento e local. Temperatura, chuva, vento, umidade, nevoeiro, nebulosidade, etc., formam o conjunto de parâmetros do tempo.

sinótico deriva do grego synoptikos, que significa obter uma visão geral de um local. Para a meteorologia, este termo é utilizado para nomear as cartas elaboradas visando observar fenômenos que possuem grande variação espaço-temporal

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.