

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ELETRÔNICA E COMPUTAÇÃO
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**DESENVOLVIMENTO DE UMA
FERRAMENTA DE
ASSIMILAÇÃO DE DADOS PARA O
SUPIM-DAVS**

TRABALHO DE GRADUAÇÃO

Cristiano Reis dos Santos

Santa Maria, RS, Brasil

2011

DESENVOLVIMENTO DE UMA FERRAMENTA DE ASSIMILAÇÃO DE DADOS PARA O SUPIM-DAVS

Por

Cristiano Reis dos Santos

Trabalho de Graduação apresentado ao Curso de Ciência da
Computação da Universidade Federal de Santa Maria (USFM, RS),
como requisito parcial para a obtenção de grau de
Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Benhur de Oliveira Stein
Co-Orientador: Dr. Adriano Petry

Trabalho de Graduação N. 326

Santa Maria, RS, Brasil

2011

**Universidade Federal de Santa Maria
Centro de Tecnologia
Curso de Ciência da Computação**

A Comissão Examinadora, abaixo assinada,
aprova o Trabalho Final de Graduação

**DESENVOLVIMENTO DE UMA FERRAMENTA DE
ASSIMILAÇÃO DE DADOS PARA O SUPIM-DAVS**

elaborado por
Cristiano Reis dos Santos

como requisito parcial para obtenção do grau de
Bacharel em Ciência da Computação

COMISSÃO EXAMINADORA:

Prof. Dr. Benhur de Oliveira Stein (UFSM)
(Presidente/Orientador)

Dr. Adriano Petry (INPE)
(Co-Orientador)

Prof^a.Dr^a. Andrea Schwertner Charão (UFSM)

Prof^a. Dr^a. Juliana Kaizer Vizzotto (UFSM)

Santa Maria, 15 de Dezembro de 2011

*"Se você for tirar a vida de um homem, você deve olhar em seus olhos e ouvir suas palavras finais. Se você não puder fazer isso, então talvez esse homem não mereça morrer."
- Eddard Stark.*

AGRADECIMENTOS

Agradeço à minha família pelo apoio e o amparo nos momentos difíceis. Mãe, irmã e minha linda sobrinha Vitória. À minha amada Ana Maria, pelo apoio e incentivo recebido todos os dias. Aos colegas de laboratório pelo apoio e contribuições dadas ao longo do ano, em especial ao André e à Miriam. Ao Instituto Nacional Pesquisa Espacial (INPE) pela infraestrutura e oportunidade de pesquisa a mim proporcionada. Ao CNPq pelo apoio financeiro. Ao meu orientador, Dr. Adriano Petry pelo incentivo, conhecimento compartilhado e a oportunidade de pesquisa proporcionada, que é de grande importância para minha formação acadêmica.

Agradeço aos meus amigos de infância e as amigas que construí na vida acadêmica. Aos grandes amigos do INFCOOLGUYS: André, Bernardo, Cícero, Frederico e Daniel que preferiu virar médico e atender todos nós de graça.

Agradeço aos meus professores pelo conhecimento passado e pelos conselhos dados. Agradeço a professora Lisandra, minha primeira orientadora, pelo conhecimento passado. Agradeço ao professor Benhur, por ter aceitado o convite de orientar esse trabalho. Agradeço as professoras Andrea e Juliana por terem aceitado o convite para comporem minha banca avaliadora. Agradeço a Janice, que sempre foi muito atenciosa quando precisei de sua ajuda.

RESUMO

Trabalho de Graduação
Curso de Ciência da Computação
Universidade Federal de Santa Maria

DESENVOLVIMENTO DE UMA FERRAMENTA DE ASSIMILAÇÃO DE DADOS PARA O SUPIM-DAVS

Autor: Cristiano Reis dos Santos
Orientador: Prof. Dr. Benhur de Oliveira Stein
Co-Orientador: Dr. Adriano Petry

Data e Local da Defesa: 15 de dezembro de 2011, Santa Maria.

Alguns sistemas físicos podem ser descritos por modelos matemáticos. Esses modelos podem ser usados para prever a dinâmica esperada para um determinado sistema real, a partir de um estado inicial. Muitas vezes, no entanto, o comportamento do sistema modelado acaba divergindo do comportamento do sistema real. Nesse momento, dados observacionais oriundos de instrumentos de medição podem ser apresentados ao modelo matemático, que os utiliza para que sua dinâmica seja aproximada ao verdadeiro estado do sistema. Esse processo é conhecido como Assimilação de Dados.

Este trabalho de conclusão de curso teve por finalidade o desenvolvimento de uma ferramenta de assimilação de dados. Essa ferramenta foi acoplada ao sistema já desenvolvido, o SUPIM-DAVS. O sistema deverá ser utilizado para assimilar valores de densidade eletrônica a partir de observações provenientes do ambiente em análise, a ionosfera, e valores de densidade eletrônica presentes no modelo SUPIM-DAVS, atualizando o estado do modelo.

Palavras-chave: Assimilação de Dados, Ionosfera, Densidade Eletrônica.

ABSTRACT

**Undergraduate Final Work
Graduation in Computer Science
Federal University of Santa Maria**

DEVELOPMENT OF A TOOL FOR DATA ASSIMILATION TO SUPIM-DAVS

Author: Cristiano Reis dos Santos
Advisor: Prof. Dr. Benhur de Oliveira Stein
Co- Advisor: Dr. Adriano Petry

Date and Local: December 15 2011, Santa Maria.

Some physical systems can be described by mathematical models. These models can be used to predict the expected dynamics for a given real system from an initial state. Often, however, the behavior of the modeled system does not agree with the real system behavior. At this time, observational data from the measuring instruments can be presented to the mathematical model, which uses them to approximate its dynamics to the true state of the system. This process is known as data assimilation.

This conclusion of course work aims to develop a tool for data assimilation. This tool was coupled to the system already developed, SUPIM-DAVS. The system should be used to assign values of electron density from observations in the environment under analysis, the ionosphere, and values of electron density in the model SUPIM, updating the model state.

Keywords: Data Assimilation, Ionosphere, Electron Density.

LISTA DE FIGURAS

Figura 2.1- Representação superficial de um processo de assimilação de dados (Lahoz et al., 2010).....	19
Figura 2.2 - Métodos de assimilação de dados classificados de acordo com sua complexidade (Bouttier; Courtier, 1999).....	20
Figura 2.3 - Pontos em grade representados pelos círculos e observações representadas pelos quadrados escuros (Kalnay, E. 2003).	24
Figura 2.4 - Representação gráfica do processo da Análise de Cressman (Bouttier; Courtier, 1999).....	25
Figura 2.5 - Representação da área de abrangência da IO (Bouttier; Courtier, 1999).	27
Figura 2.6 - Passos de execução do Filtro de Kalman e Filtro de Kalman Estendido (Bouttier; Courtier, 1999).	29
Figura 2.7 – Mapa de TEC para América do Sul as 20:00 UT (http://200.132.24.196/site/supim/20.0h.gif).....	31
Figura 2.8 - Representação das camadas da ionosfera (Dave, A; Rowell, F.T, 1999).	33
Figura 2.9–Ilustração da análise gerada por uma assimilação de dados sendo utilizada como condição inicial de uma PNT(Kalnay, E. 2003).	35
Figura 2.10 - Esquema do Sistema de Observação Global (http://www.wmo.int/pages/prog/www/OSY/GOS.html).....	37
Figura 2.11 – Estações de ionossondas utilizadas pelo SPIDR (SPIDR, 2011).....	38
Figura 3.1 – Principais classes do programa.	40
Figura 3.2 – Exemplo da estrutura do arquivo de parâmetros.....	41
Figura 3.3 – Exemplo de um arquivo no formato SAO.....	42
Figura 3.4 – Horário de busca para 15h30min e intervalo de 15 minutos.....	43
Figura 3.5 – Parte de um arquivo utilizado pela assimilação do SUPIM-DAVS.....	44
Figura 3.6 – Diagrama de classe do SUPIM-DAVS. A classe OI está destacada pelo retângulo vermelho.	46
Figura 3.7 – Fluxograma que representa de forma simples o funcionamento do método.....	49
Figura 4.1 – Representação do processo de assimilação na geração dos mapas de TEC.....	50
Figura 4.2 – Perfil eletrônico para Jicamarca com raio de 50.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação.	53

Figura 4.3 - Perfil eletrônico para Jicamarca com raio de 100.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação	53
Figura 4.4 - Perfil eletrônico para Jicamarca com raio de 150.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação	54
Figura 4.5 - Perfil eletrônico para Cachoeira Paulista com raio de 50.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação.....	54
Figura 4.6 - Perfil eletrônico para Cachoeira Paulista com raio de 100.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação.....	55
Figura 4.7 - Perfil eletrônico para Cachoeira Paulista com raio de 150.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação.....	55
Figura 4.8 – Análise para a ionossonda de Cachoeira Paulista utilizando raio de 50.000 metros.	57
Figura 4.9 - Análise para a ionossonda de Cachoeira Paulista utilizando raio de 100.000 metros.	57
Figura 4.10 - Análise para a ionossonda de Cachoeira Paulista utilizando raio de 150.000 metros.	58
Figura 4.11 - Análise para a ionossonda de Jicamarca utilizando raio de 50.000 metros.	59
Figura 4.12 - Análise para a ionossonda de Jicamarca utilizando raio de 100.000 metros.	59
Figura 4.13 - Análise para a ionossonda de Jicamarca utilizando raio de 150.000 metros.	60

LISTA DE TABELAS

Tabela 1- Configuração do Ambiente de Execução. 51

LISTA DE SIGLAS E ABREVIATURAS

INPE	INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS
PNT	PREVISÃO NUMÉRICA DE TEMPO
GPS	GLOBAL POSITIONING SYSTEM
TEC	TOTAL ELECTRON CONTENT
LABCCE	LABORATÓRIO DE COMPUTAÇÃO PARA O CLIMA ESPACIAL
CRS	CENTRO REGIONAL SUL DE PESQUISAS ESPACIAIS
BLUE	BEST LINEAR UNBIASED ESTIMATOR
OI	INTERPOLAÇÃO ÓTIMA
NOAA	NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
SPIDR	SPACE INTERACTIVE DATA RESOURCE
SUPIM	SHEFFIELD UNIVERSITY PLASMASPHERE IONOSPHERE
DAVS	DATA ASSIMILATION AND VISUALIZATION SYSTEM
SCM	SUCCESSIVE CORRECTIONS METHOD
GrADS	GRID ANALYSIS AND DISPLAY SYSTEM
GOS	GLOBAL OBSERVING SYSTEM
NOAA	NACIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
WMO	WORLD METEOROLOGICAL ORGANIZATION
NGDC	NATIONAL GEOPHYSICAL DATA CENTER
STP	SOLAR AND TERRESTRIAL PHYSICS DIVISION
FTP	FILE TRANSFER PROTOCOL
SÃO	STANDART ARCHIVING OUTPUT
IDW	INVERSE DISTANCE-WEIGHTING

LISTA DE SÍMBOLOS

x	Vetor do modelo de estado
xb	Modelo de estado de <i>background</i>
xt	Estado verdadeiro
y	Vetor de observação
Δ_x	Incremento da análise
B	Matriz de covariância dos erros de <i>background</i>
H	Operador de observação
K	Matriz de ganho
R	Matriz de covariância dos erros de observação
$Q(i)$	Matriz de covariância de erros do modelo do Filtro de Kalman
ϵ^2	Coefficiente de erro do Successive Corrections Method
P_f	Matriz de covariância dos erros de <i>background</i> do Filtro de Kalman
P_a	Matriz de covariância dos erros da análise do Filtro de Kalman
x_f	Previsão do filtro de Kalman
$M_{i \rightarrow i+1}$	Modelo de previsão do Filtro de Kalman
O^+	Átomo de Oxigênio
H^+	Átomo de Hidrogênio
He^+	Molécula de Hélio
N_2^+	Molécula de Nitrogênio
O_2^+	Molécula de Oxigênio
NO^+	Nitrogênio Óxido
UV	Ondas Solares Ultravioletas
D	Camada D da ionosfera
E	Camada E da ionosfera
F	Camada F da ionosfera
fp	Frequência de plasma

n Densidade ou concentração eletrônica

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	Justificativa.....	16
1.2	Objetivos	16
1.2.1	Objetivos específicos	17
1.3	Estrutura do texto.....	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Assimilação de Dados	18
2.1.1	Conceitos Básicos de Assimilação de Dados	19
2.1.2	Técnicas de Interpolação Estatística.....	22
2.1.3	Métodos de Assimilação de Dados.....	23
2.2	Sheffield University Plasmaphere Ionosphere Model	29
2.3	SUPIM-DAVS	30
2.4	Ionosfera	31
2.4.1	Camadas da Ionosfera.....	33
2.4.2	Ionossonda	34
2.5	Previsão Numérica do Tempo.....	35
2.6	Observações	36
2.6.1	Sistema de Observação Global	36
2.6.2	SPACE PHYSICS INTERACTIVE DATA RESOURCE.....	37
3	DESENVOLVIMENTO.....	39
3.1	Dados observacionais ionosféricos	39
3.1.1	Programa de obtenção dos dados.....	40
3.1.2	Estações de ionossondas	41
3.1.3	Obtenção dos arquivos de densidade eletrônica	42
3.1.4	Obtenção dos dados relevantes para a assimilação	43
3.1.5	Conversão entre plasma e concentração eletrônica	44
3.2	Método de Assimilação de Dados Ionosféricos.....	44
3.2.1	Busca dos pontos mais próximos	46
3.2.2	Definição das matrizes de covariância de erros.....	47
3.2.3	Operador de Observação	47
3.2.4	Cálculo da Matriz Inversa.....	48
3.2.5	Etapas do processo de assimilação	49
4	RESULTADOS	50
4.1	Ambiente de Execução.....	50
4.2	Parâmetros de execução	51
4.3	Perfil eletrônico nas estações de ionossonda.....	51
4.4	Análise interpolada para a dimensão das observações.....	56
5	CONCLUSÃO.....	61
	REFERÊNCIAS	63

1 INTRODUÇÃO

Alguns sistemas físicos podem ser realisticamente descritos por modelos matemáticos. Esses modelos são usados para prever o comportamento esperado para um determinado sistema, a partir de um estado inicial. O estado perfeito do ambiente em análise não é passível de representação, então, o que tem se tentado fazer é descrever a melhor representação possível de ser gerada matematicamente e computacionalmente.

O modelo que representa o sistema e sua condição inicial contêm incertezas em suas representações. Por esta razão, observações provenientes de diversas fontes são utilizadas com a finalidade de ajustar o modelo de previsão, diminuindo os erros e aumentando a exatidão. Isso é conhecido atualmente como assimilação de dados, que vem sendo desenvolvida e aprimorada a cada dia.

Uma boa previsão depende de dois fatores importantes: observações de boa qualidade, que nos dirão como o sistema parece estar se comportando, e um modelo matemático confiável, responsável por combinar as observações com o estado passado do ambiente em análise, para gerar a provável situação futura, conforme podemos encontrar em Lahozet.al(2010).

A maioria dos ambientes simulados se comporta de maneira pouco previsível, modificando-se constantemente. A ionosfera terrestre apresenta esse tipo de comportamento, ocasionando uma dificuldade a mais no processo de assimilação de dados.

O sistema SUPIM-DAVS (DAVS - DATA ASSIMILATION AND VISUALIZATION SYSTEM) (Petry, Souza e Velho, 2011) é um sistema para previsão da dinâmica ionosférica, sendo que seu desenvolvimento foi baseado no modelo físico SHEFFIELD UNIVERSITY PLASMASPHERE IONOSPHERE MODEL (SUPIM) (BAILEY, G, 1990). O sistema SUPIM-DAVS executa diariamente, gerando mapas com a previsão da dinâmica da ionosfera para um período futuro. As imagens de conteúdo eletrônico total (TEC - Total Electron Content) são disponibilizadas diariamente no sítio do Programa de Clima Espacial do INPE (INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS).

Este trabalho de conclusão de curso tem por finalidade o desenvolvimento de uma ferramenta de assimilação de dados. Essa ferramenta deverá ser acoplada ao sistema já

desenvolvido, o SUPIM-DAVS. O sistema deverá ser utilizado para assimilar valores de densidade eletrônica a partir de observações provenientes do ambiente em análise, a ionosfera, e valores de densidade eletrônica presentes no modelo SUPIM, atualizando o estado do modelo.

1.1 Justificativa

Muitas áreas são altamente dependentes do poder de previsão da assimilação de dados. Entre elas está a Previsão Numérica de Tempo (PNT) que utiliza modelos de previsão e detalhes sobre a atmosfera para simular seu estado de algumas horas até várias semanas à frente. A Previsão Numérica de Tempo está diretamente relacionada à economia. A capacidade de prever acontecimentos climáticos pode evitar grandes perdas na área agropecuária. Governos realizam ações preventivas baseados na previsão de seca ou chuvas em grandes quantidades. Muitas escolhas para o futuro são baseadas na previsão que se espera encontrar no futuro.

O conhecimento sobre a densidade eletrônica presente na ionosfera terrestre é de grande importância para comunicações via radiofrequência, como por exemplo, o sistema global de posicionamento (GPS – Global Positioning System) que tem seu sinal afetado por mudanças do conteúdo eletrônico total (TEC – Total Electron Content).

Para aprimorar a capacidade do Laboratório de Computação para o Clima Espacial (LABCCE) do Centro Regional Sul de Pesquisas Espaciais (CRS – INPE) de realizar previsões de conteúdo eletrônico total, julgou-se necessário acrescentar dados assimilados na obtenção da previsão do TEC.

1.2 Objetivos

O principal objetivo desse trabalho de conclusão de curso é desenvolver e auxiliar a construção de algoritmos de assimilação de dados para serem usados para assimilação de dados da ionosfera terrestre. O processo de assimilação será concentrado principalmente para correção da densidade eletrônica presente na ionosfera, utilizada pelo sistema SUPIM-DAVS, para geração de mapas de TEC.

1.2.1 Objetivos específicos

O trabalho pode ser dividido inicialmente em alguns objetivos específicos, que poderão mudar durante a realização do trabalho. A seguir serão listados estes objetivos:

- Realizar uma revisão bibliográfica sobre conteúdos referentes à assimilação de dados e ionosfera terrestre;
- Pesquisar e obter informações mais detalhadas sobre fontes de informações sobre densidade eletrônica da ionosfera;
- Programar uma ferramenta para obtenção de dados observacionais de densidade eletrônica. Codificar, utilizando a linguagem de programação C++, algoritmos de assimilação de dados;
- Usar efetivamente a ferramenta de assimilação de dados, auxiliando na previsão dos mapas de TEC pelo SUPIM-DAVS.

1.3 Estrutura do texto

O presente texto está estruturado como segue. No Capítulo 2 é realizada uma fundamentação teórica sobre os assuntos abordados no trabalho, como assimilação de dados, onde são citados alguns dos principais métodos utilizados no passado e atualmente, e uma descrição da ionosfera terrestre, ambiente de trabalho do projeto. No Capítulo 3 é realizada a descrição das atividades desenvolvidas, apresentando todo o processo de realização do projeto, como a ferramenta de obtenção e tratamento das observações, e o processo de assimilação de dados. No Capítulo 4 serão apresentados os resultados obtidos com a execução do trabalho e no Capítulo 5 será apresentada a conclusão e as contribuições do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

A seguir, será apresentada uma revisão sobre os conteúdos considerados importantes para a realização deste trabalho.

2.1 Assimilação de Dados

Uma análise é a produção de uma imagem precisa do estado verdadeiro da atmosfera ou sistema físico em um determinado tempo, representado em um modelo como uma coleção de números (Bouttier; Courtier, 1999). A análise pode ser utilizada para compreensão do estado do sistema, como dados de entrada para outras operações, como entrada para uma previsão numérica do tempo, entre outras finalidades.

A assimilação de dados tem por finalidade combinar observações com um modelo de sistema dinâmico, para estimar o atual e futuro estado do sistema, como podemos encontrar em Lahoz et al.(2010).

A assimilação de dados tenta manter o modelo o mais realístico possível. O comportamento caótico visto na atmosfera, por exemplo, nos impede de confiar em uma previsão inicial. Por esse motivo, as observações são obtidas por diversos meios e combinadas com a versão mais nova do modelo gerando uma nova previsão, mais condizente com a realidade do estado. Essa nova previsão é utilizada futuramente como o novo estado do modelo para realizar uma nova previsão.

Atualmente os modelos de assimilação de dados trabalham com uma grande quantidade de dados. A dimensão utilizada em modelos de previsão meteorológicos é de aproximadamente da ordem de 10^7 a 10^8 para as variáveis do modelo, e o número de observações está na ordem de 10^5 a 10^6 (Lahoz et al., 2010). Os dados não são distribuídos de forma uniforme no espaço, na maioria dos casos os dados são esparsos, com buracos onde não há observações, ocasionando erros de previsão mais visíveis.

O estado verdadeiro dos sistemas não é passível de representação. Para contornar esse problema, é utilizada uma representação do sistema que se julga a melhor possível da realidade. Os algoritmos de assimilação de dados necessitam de alguma informação confiável do estado do sistema, essa importante informação utilizada é conhecida como *background* ou *first guess*. Uma fonte de dados observados do ambiente em análise deve disponibilizar dados

sobre o estado verdadeiro do sistema, dados observacionais têm por finalidade ajustar o modelo de previsão, tentando corrigir as discrepâncias do modelo. Todo sistema de assimilação faz uma previsão com uma margem de erros. A minimização dos erros é muito importante, levando muitos centros de pesquisa a desenvolver diferentes abordagens de tratamento de erros.

A Figura 2.1 é uma representação superficial de como o processo de assimilação de dados é realizado. A figura demonstra o método de assimilação de dados recebendo informação proveniente de observações feitas no ambiente em análise, uma informação confiável sobre o estado do sistema no momento da análise, descrito na figura como o *Forecast*, e uma estimativa de erro, tanto das observações quanto do modelo de previsão. O método de assimilação utiliza todas as informações anteriormente descritas para gerar um novo estado para o sistema representado na figura pela *Analyses*.

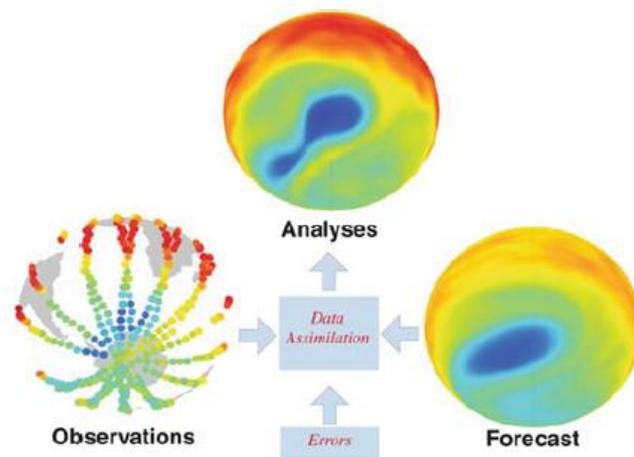


Figura 2.1- Representação superficial de um processo de assimilação de dados (Lahoz et al., 2010).

2.1.1 Conceitos Básicos de Assimilação de Dados

Assimilação de dados é uma técnica onde a informação observada é acumulada dentro do estado do modelo, também conhecido como o estado do sistema, tendo como vantagem a consistência das restrições de evolução do tempo e propriedades físicas do sistema.

Em seu curso de treinamento (Bouttier; Courtier, 1999), definem duas abordagens básicas para assimilação de dados:

- **Assimilação sequencial**, que somente considera observações obtidas até o momento da análise, a qual é utilizada em sistemas de tempo real;

- **Assimilação não sequencial ou retrospectiva**, onde observações do futuro também podem ser usadas, por exemplo, em uma reanálise;

Outra distinção que pode ser feita é entre os métodos intermitentes e contínuos no tempo:

- Em um **método intermitente** observações são processadas em pequenos lotes, é usualmente a técnica mais conveniente.
- Em um **método contínuo** são utilizadas observações em um longo período de tempo, e a correção para o estado analisado é suavizada no tempo, é o método fisicamente mais realista.

Muitas técnicas têm sido desenvolvidas para meteorologia e oceanografia. Elas diferem no custo numérico, proximidade da análise do estado ótimo, e em sua adequação a sistemas de tempo real de assimilação de dados. A Figura 2.2 apresenta alguns desses métodos classificados de acordo com sua complexidade e custo de implementação. A figura demonstra os principais métodos de assimilação de dados, as setas indicam que determinados métodos foram implementados utilizando a ideia já utilizada em outro método, com modificações que tornaram o método mais indicado para determinados sistemas do que o seu anterior.

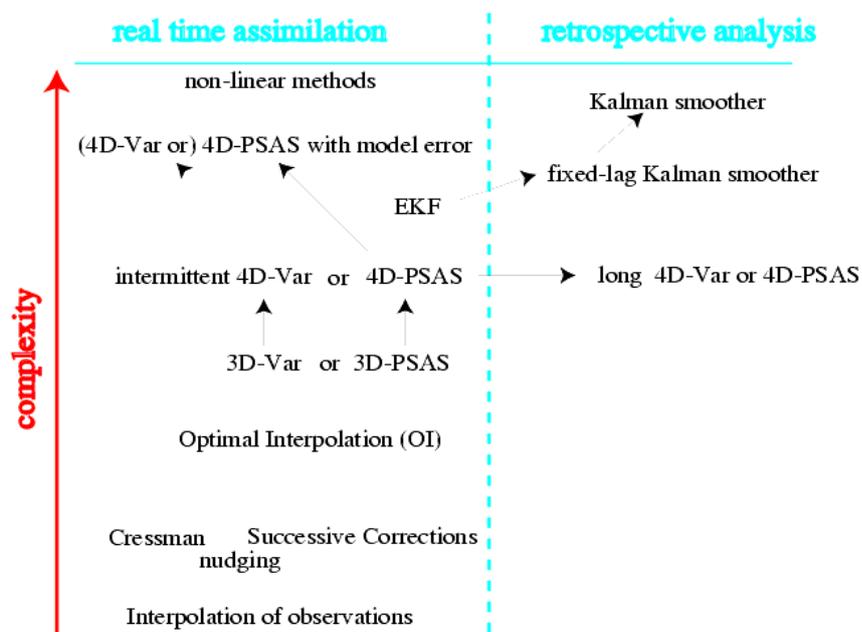


Figura 2.2 - Métodos de assimilação de dados classificados de acordo com sua complexidade (Bouttier; Courtier, 1999).

Os critérios para uma boa análise segundo Bouttier e Courtier (1999) são:

- A estimativa inicial deve ser de boa qualidade;
- Se as observações são densas, então se assume que a probabilidade verdadeira encontra-se próxima da média. A análise deve aproximar-se dos dados considerados mais confiáveis, enquanto dados suspeitos devem ter menor peso;
- Quando estamos longe de uma observação, a análise deve relaxar suavemente em direção à estimativa inicial;
- A análise deve respeitar as características físicas do sistema, reconhecendo casos excepcionais, pois esses são muito importantes.

Os métodos de assimilação de dados apresentam conjuntos de informações utilizados em suas formalizações matemáticas, que estão presentes na maioria dos sistemas aqui apresentados. Algumas informações serão denotadas abaixo.

O espaço do modelo da assimilação é definido como uma coleção de números representada como uma matriz coluna, chamada de *state vector* x . A realidade é mais complexa do que pode ser representada pelo *state vector*, assim utiliza-se a melhor representação possível no momento da análise, conhecida como *true state*, representado por x_t , o x_t no processo de assimilação geralmente não é utilizado, devido à dificuldade em defini-lo. Outra importante informação é o vetor de *background*, representado por x_b , ele é uma estimativa do estado verdadeiro antes da análise ser realizada, o x_b geralmente é oriundo de uma previsão anterior. A análise é representada pelo vetor x_a .

Algumas vezes não é possível trabalhar com o espaço do modelo, pois o poder computacional não é suficiente, sendo assim, o domínio da região analisada deve ser reduzido. O problema da análise é encontrar uma correção δ_x (ou incremento da análise) que aproxime x_a o máximo possível de x_t (Bouttier; Courtier, 1999), como mostrado na Equação 2.1.

$$x_a = x_b + \delta_x \quad (2.1)$$

2.1.2 Técnicas de Interpolação Estatística

Um dos mais importantes avanços da área de assimilação de dados foi o desenvolvimento das técnicas de interpolação estatística. Os pesos dados às observações e aos dados de background passaram a ser baseados nas características de seus erros.

Em seu livro (Swinbank et al., 2010) ilustra os princípios da assimilação de dados baseada em métodos de interpolação estatística. Considere duas variáveis independentes não relacionadas, de estimativa x_1 e x_2 de mesma grandeza. Assume-se que os erros são conhecidos e representados por α_1 e α_2 respectivamente, podemos realizar uma combinação para estimar x^a representado pela Equação 2.2.

$$x^a = \frac{\alpha_2^2 x_1 + \alpha_1^2 x_2}{\alpha_1^2 + \alpha_2^2} \quad (2.2)$$

A estimativa ótima x^a é obtida pela ponderação entre x_1 e x_2 inversamente proporcional aos seus respectivos erros. A Equação 2.2 pode ser escrita de outra forma, representada na Equação 2.3.

$$x^a = x_1 + \frac{\alpha_1^2}{\alpha_1^2 + \alpha_2^2} (x_2 - x_1) \quad (2.3)$$

A Equação 2.3 é generalizada por (Swinbank et al., 2010) para uma abordagem onde tem-se um conjunto de dados de observações e um conjunto de dados de *background*, cada um com suas específicas características de erros. A notação padrão para a interpolação ótima é mostrada na Equação 2.4, onde x^a representa a análise, x^b é a informação de *background* com matriz de covariância de erros B , y^o é o conjunto de dados de observação com matriz de covariância de erros R , e H denotado como *observation operator*, tendo a finalidade de levar as observações que geralmente se encontram em menor quantidade para o espaço do *state vector*, para que possam ser comparados. Usualmente H é uma coleção de operadores de interpolação, onde existe uma linha para cada observação (Bouttier; Courtier, 1999).

$$x^a = x^b + BH^T (HBH^T + R)^{-1} (y^o - Hx^b) \quad (2.4)$$

A Equação 2.4 corresponde exatamente a Equação 2.3, a diferença encontra-se na escrita da equação que é feita em função de vetores, em vez de escalares. A Equação 2.4

demonstra que o incremento da análise $\&x(x^a - x^b)$ é encontrada calculando-se $(HBH^T + R)^{-1}(y^o - Hx^b)$ no espaço de observação. A solução encontrada é então transformada de volta para o espaço do modelo, multiplicando o resultado por BH^T .

2.1.3 Métodos de Assimilação de Dados

Os métodos de assimilação de dados são baseados em técnicas de interpolação, que são processos que utilizam valores de dados conhecidos para estimar valores de dados desconhecidos. Várias técnicas de interpolação são utilizadas atualmente nas ciências atmosféricas e oceânicas. Entre esses métodos podemos encontrar desde abordagens simples, até métodos muito complexos de estimativa de valores. Entre as técnicas mais simples está a Interpolação Linear, que requer o conhecimento de dois pontos e uma constante de variação entre eles. Com essas informações, podem-se interpolar valores em qualquer região entre os pontos conhecidos. Outro método um pouco mais complexo, mas também simples de ser entendido é conhecido como Análise de Cressman.

Entre as técnicas mais avançadas para interpolação de dados encontram-se o Best Linear Unbiased Estimator (*BLUE*), a Interpolação Ótima (*OI*), o Filtro de Kalman, o *3D - VAR* e o *4D - VAR*, todos descritos em (Bouttier; Courtier, 1999). Esses métodos permitem que atualmente seja possível prever fenômenos climáticos com um alto grau de exatidão, que não seria possível apenas utilizando equipamentos de medição, seja pelo valor de aquisição desses equipamentos, pela grande área que deveria ser coberta por eles ou pelo tempo para analisar a informação e distribuí-la para os interessados com um alto grau de confiança.

2.1.3.1 Análise de Cressman

A Análise de Cressman foi desenvolvida em 1959 por George Cressman. É uma técnica de interpolação de estações de dados definidas através de altitude e longitude, sendo uma técnica empírica que confia muito nas observações. Muitos passos são realizados através de dados em grade, utilizando um raio para influenciar no aumento da precisão (Bouttier; Courtier, 1999).

O raio de influência é definido como o máximo raio de um ponto da grade para uma estação de dados ao qual foi observado um valor. Estações de dados além do raio de influência não tem qualquer influência no valor da grade. Para cada ponto da grade um novo

valor é calculado baseado em um fator de correção. Esse fator de correção é determinado analisando cada estação dentro do raio de influência. Observações próximas ao ponto de grade carregam maior peso. Com a distância aumentando o peso de influência diminui. Na Figura 2.3, é apresentada uma grade de pontos representados por círculos e observações representadas por quadrados pretos. Para o ponto de grade j apenas as observações dentro do seu raio de influência afetarão no novo valor do ponto j .

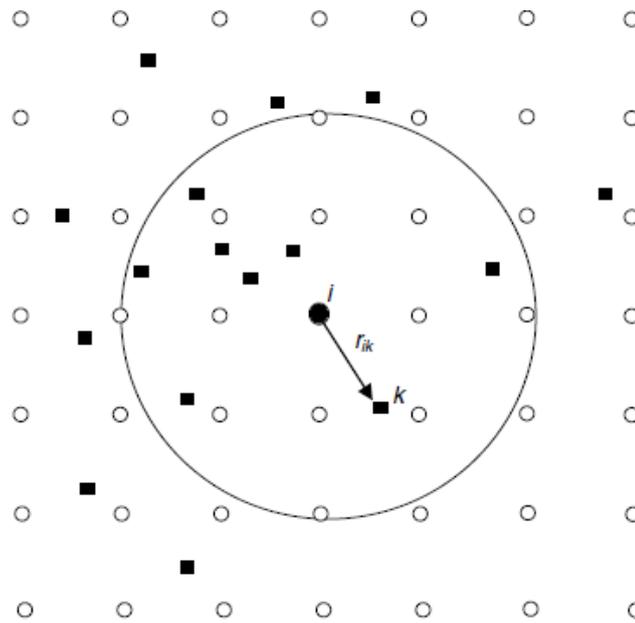


Figura 2.3 - Pontos em grade representados pelos círculos e observações representadas pelos quadrados escuros (Kalnay, E. 2003).

$$x_{a(j)} = x_{b(j)} + \frac{\sum_{i=1}^n w(i,j) \{y(i) - x_{b(i)}\}}{\sum_{i=1}^n w(i,j)} \quad (2.5)$$

$$w(i,j) = \max\left(0, \frac{R^2 - d_{i,j}^2}{R^2 + d_{i,j}^2}\right) \quad (2.6)$$

A Equação 2.5 encontrada em (Bouttier; Courtier, 1999), representa a formulação da Análise de Cressman, onde x_a representa a nova análise, ou seja, a nova estimativa de valores para os pontos do modelo. Assume-se que x_b é o *first guess* ou *background*. Os valores $y(i)$ das observações e $w(i,j)$ representado pela Equação 2.6, é o peso dado à observação, que dependerá da distância entre os pontos de grade e a observação. R é uma constante definida pelo usuário, conhecida como raio de influência e $d(i,j)$ é a distância entre os pontos de

grade e a observação. A Figura 2.4 ilustra o resultado esperado para a Análise de Cressman aplicada a um *background* ilustrativo.

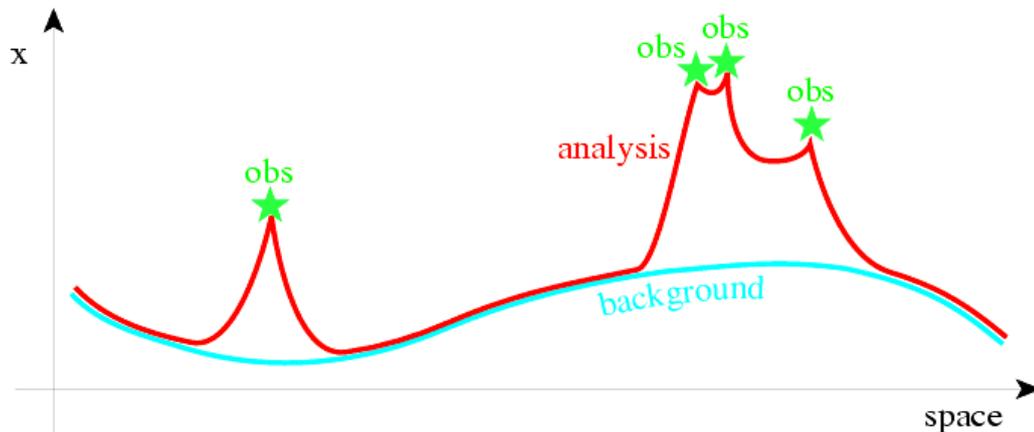


Figura 2.4 - Representação gráfica do processo da Análise de Cressman (Bouttier; Courtier, 1999).

Na Análise de Cressman, a análise converge fortemente para as observações presentes no raio de influência. Uma versão chamada Successive Corrections Method (*SCM*) utiliza um coeficiente representado por ϵ^2 , para representar o erro presente nas observações em relação ao *background*. Assim um peso maior é dado aos valores de *background*, fazendo com que a análise convirja de forma mais suave entre as observações e o *background*. Essa fórmula é apresentada na Equação 2.7.

$$x_a(j) = x_b(j) + \frac{\sum_{i=1}^n w(i,j)\{y(i) - x_b(i)\}}{\sum_{i=1}^n w(i,j) + \epsilon^2} \quad (2.7)$$

2.1.3.2 Best Linear Unbiased Estimator

O Best Linear Unbiased Estimator (BLUE), também conhecido como estimador dos mínimos quadrados, fornece a melhor ponderação estatística possível, considerando os erros de observação e de *background*. O BLUE realiza a análise considerando todos os pontos presentes no modelo. Sua utilização para sistemas que contêm muitos dados pode ser proibitiva, pois o processamento requerido é muito elevado.

Serão apresentadas a seguir, algumas notações utilizadas pelo BLUE, e pequenos comentários sobre essas notações, que foram adaptadas de (Bouttier; Courtier, 1999).

Nas equações 2.8 a 2.13 o x_t é a melhor representação possível da realidade. O x_b é denominado como *background*. O x_b representa uma estimativa do estado verdadeiro antes da

análise ser realizada. O resultado de uma nova análise é x_a , que representa a nova previsão para o tempo atual e informações utilizadas na análise. A análise é obtida através da Equação 2.13.

Para realizar uma análise, necessita-se de um número de valores observados do ambiente. Isso é conhecido como vetor de observações y , que tem por finalidade ajustar o modelo de previsão. A matriz H é conhecida como operador de observação e tem por finalidade conduzir o espaço do modelo para a mesma dimensão das observações, para que possam ser comparadas. O operador K representado na Equação 2.12 é chamado de ganho, ou matriz de peso da análise, ele definirá qual será o incremento para cada valor analisado. As matrizes B e R representam os erros do modelo. B é uma matriz de covariância que representa os erros de *background* e R é a matriz de covariância que representa os erros de observação. As matrizes são definidas a partir da Equação 2.8 e da Equação 2.10, respectivamente.

$$B = \overline{(e_b - \bar{e}_b)(e_b - \bar{e}_b)^T} \quad (2.8)$$

$$e_b = x_b - x_t \quad (2.9)$$

$$R = \overline{(e_o - \bar{e}_o)(e_o - \bar{e}_o)^T} \quad (2.10)$$

$$e_o = y - H(x_t) \quad (2.11)$$

$$K = BH^T(HBH^T + R)^{-1} \quad (2.12)$$

$$x_a = x_b + K(y - H[x_b]) \quad (2.13)$$

2.1.3.3 Interpolação Ótima

A Interpolação Ótima (*OI – Optimal Interpolation*) é uma variação BLUE. Ela se diferencia do BLUE por utilizar apenas observações que se encontram em uma região próxima aos pontos de análise.

Para cada variável do modelo de análise, o incremento é dado por uma linha correspondente na matriz K vezes o vetor de inovações $y^o - Hx^b$. A principal hipótese na Interpolação Ótima é: para cada variável do modelo, somente algumas observações são importantes na determinação do incremento da análise (Bouttier; Courtier, 1999).

Segundo Bouttier e Courtier (1999) é possível ganhar algum tempo computacional utilizando a Interpolação Ótima, decorrente da utilização de algumas observações selecionadas, e não de todas, como ocorre com o BLUE, como podemos visualizar na Figura 2.5.

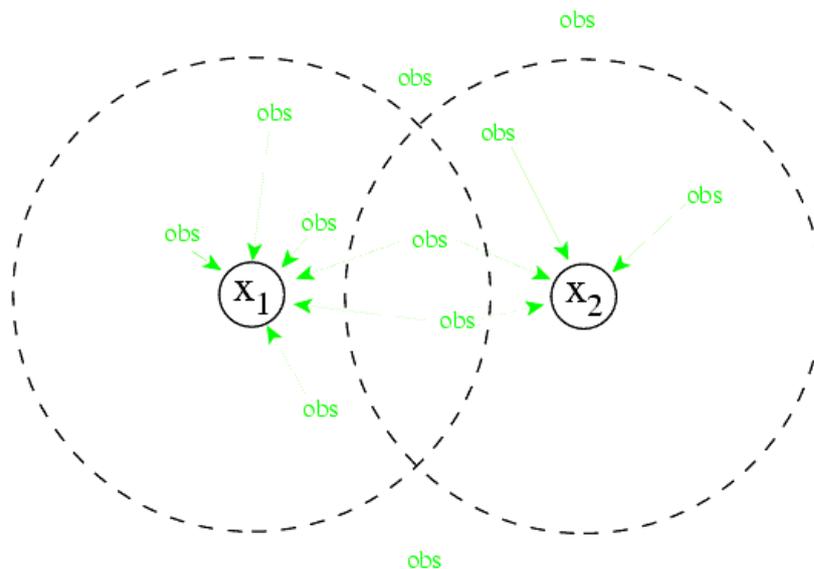


Figura 2.5 - Representação da área de abrangência da IO (Bouttier; Courtier, 1999).

A seleção das observações deve em princípio prover todas as observações as quais possam ter um peso significativo na análise ótima. Uma das desvantagens do OI são os ruídos produzidos na análise devido aos diferentes conjuntos de observações usados em diferentes partes do modelo (Bouttier; Courtier, 1999).

2.1.3.4 Filtro de Kalman

O Filtro de Kalman para sistemas lineares foi desenvolvido a partir do método dos mínimos quadrados, no qual os dados de *background* são obtidos de uma previsão oriunda de uma análise prévia (Bouttier; Courtier, 1999).

As equações de análise para a versão linear do Filtro de Kalman são exatamente as mesmas descritas para o Best Linear Unbiased Estimator. A notação utilizada é a mesma, exceto para as matrizes de covariância de erros do *background* e da análise que agora são representadas por P_f e P_a respectivamente. O *background* é x_b e a previsão é x_f (Bouttier; Courtier, 1999).

O fato de utilizarmos P_f para representar a matriz de covariância de erros em vez de B é para demonstrar que agora o *background* é uma previsão.

O filtro de Kalman utiliza uma estratégia conhecida como “análise circular”, onde as observações são utilizadas apenas uma única vez e após isso são descartadas como é encontrado em Kalnay (2003). Assume-se que uma análise completa no tempo i é disponibilizada, e o objetivo é processar uma análise em $i + 1$. Isso é realizado em duas fases: a primeira é a fase de previsão e a segunda é a atualização da análise (Kalnay, E. 2003).

A previsão $x_f(i + 1)$ é obtida pela Equação 2.14 a partir de uma análise prévia $x_a(i)$. $M_{i \rightarrow i+1}$ é o modelo de previsão responsável por levar as informações do tempo i para $i + 1$. A nova previsão da matriz de covariância de erros representada por $P_f(i + 1)$ é “atualizada” pela Equação 2.15 para a análise no tempo $i + 1$. $Q(i)$ representa a matriz de covariância de erros do modelo que são acumulados entre os passos de tempo. Os passos executados pela equação 2.14 e 2.15 representam a fase de previsão.

$$x_f(i + 1) = M_{i \rightarrow i+1} x_a(i) \quad (2.14)$$

$$P_f(i + 1) = M_{i \rightarrow i+1} P_a M_{i \rightarrow i+1}^T + Q(i) \quad (2.15)$$

As Equações 2.16, 2.17 e 2.18 são as mesmas equações utilizadas pelo *BLUE* e representam a fase de atualização.

$$K(i) = P_f(i)H^T(i)[H(i)P_f(i)H^T(i) + R(i)]^{-1} \quad 2.16$$

$$x_a(i) = x_f(i) + K(i)[y(i) - H(i)x_f(i)] \quad 2.17$$

$$P_a(i) = [I - K(i)H(i)]P(i) \quad 2.18$$

Quando utilizamos H ou/e M não lineares o Filtro de Kalman é chamado de *Extended Kalman Filter* (Bouttier; Courtier, 1999). O custo para aplicar o Filtro de Kalman é significativo. Assim o algoritmo deve ser utilizado como uma referência para desenvolver outros algoritmos de assimilação de dados usados atualmente.

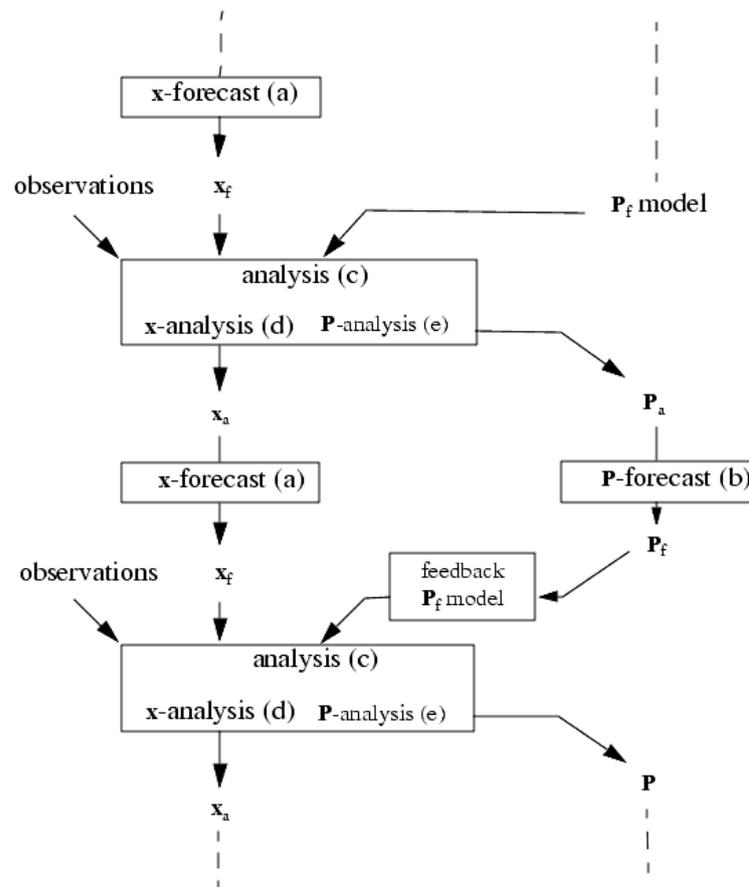


Figura 2.6 - Passos de execução do Filtro de Kalman e Filtro de Kalman Estendido (Bouttier; Courtier, 1999).

2.2 Sheffield University Plasmaphere Ionosphere Model

O modelo físico atmosférico SHEFFIELD UNIVERSITY PLASMAPHERE IONOSPHERE MODEL (SUPIM) encontrado em (BAILEY, G, 1990), é composto por um

sistema de equações da continuidade, do movimento e da conservação de energia, as quais descrevem os processos físicos e químicos da ionosfera terrestre (SANTOS, 2005). Ele é resolvido separadamente para os principais íons constituintes da ionosfera, O^+ , H^+ , $He^+N_2^+$, O_2^+ e NO^+ , e os elétrons, ao longo de linhas fechadas de campo magnético, fornecendo uma previsão para seu comportamento dinâmico no tempo (BAILEY, G, 1990).

O ponto de referência utilizado na solução dessas equações é o centro do dipolo que representa o campo magnético, sendo os cálculos efetuados ao longo das linhas de força deste campo. As equações são resolvidas ao longo de linha fechadas do campo magnético dipolar entre altitudes de aproximadamente 130 km. Tais equações fornecem valores para as densidades, para as velocidades alinhadas ao campo e para as temperaturas dos elétrons e desses seis íons segundo encontramos em SANTOS (2005).

2.3 SUPIM-DAVS

O sistema SUPIM-DAVS (DAVS - DATA ASSIMILATION AND VISUALIZATION SYSTEM) é um sistema para previsão da dinâmica ionosférica, sendo que seu desenvolvimento foi baseado no sistema SUPIM descrito na seção 2.2.

O sistema SUPIM-DAVS executa diariamente, gerando mapas com a previsão da dinâmica da ionosfera para um período futuro. O sistema atual está em operação desde março de 2011 rodando em um cluster de processadores, disponível junto ao Centro Regional Sul de Pesquisas Espaciais (CRS/INPE). As imagens de mapas de TEC são disponibilizadas diariamente no sítio do Programa de Clima Espacial do INPE - <http://www.inpe.br/climaespacial/index.php> segundo PETRY, SOUZA e VELHO (2011).

O sistema DAVS objetiva, basicamente, utilizar os resultados gerados pelas várias rodadas do código SUPIM em uma simulação para construir uma grade homogeneamente espaçada em latitude, longitude e altitude geográficas com estimação de valores de concentração iônica/eletrônica em cada ponto como proposto por Petry, Souza e Velho (2011). Para a interpolação dos dados é utilizada uma técnica chamada “explosão de perfil”, na qual os pontos simulados são associados aos pontos da grade próximos a ele. A ponderação para estimar a concentração para cada ponto da grade então utilizará apenas alguns dos pontos simulados identificados como próximos, de acordo com o alcance desejado (relacionado à “ordem de explosão”) segundo Petry, Souza e Velho (2011).

O DAVS gera como saída um pacote de arquivos no formato utilizado por um sistema de visualização de dados geofísicos, chamado Grid Analysis And Display System (GrADS, 2011). Onde a partir da grade tridimensional gerada pelo DAVS, podem ser calculados mapas bidimensionais com a estimativa do TEC, como pode ser visualizado na Figura 2.8 (PETRY; SOUZA; VELHO, 2011).

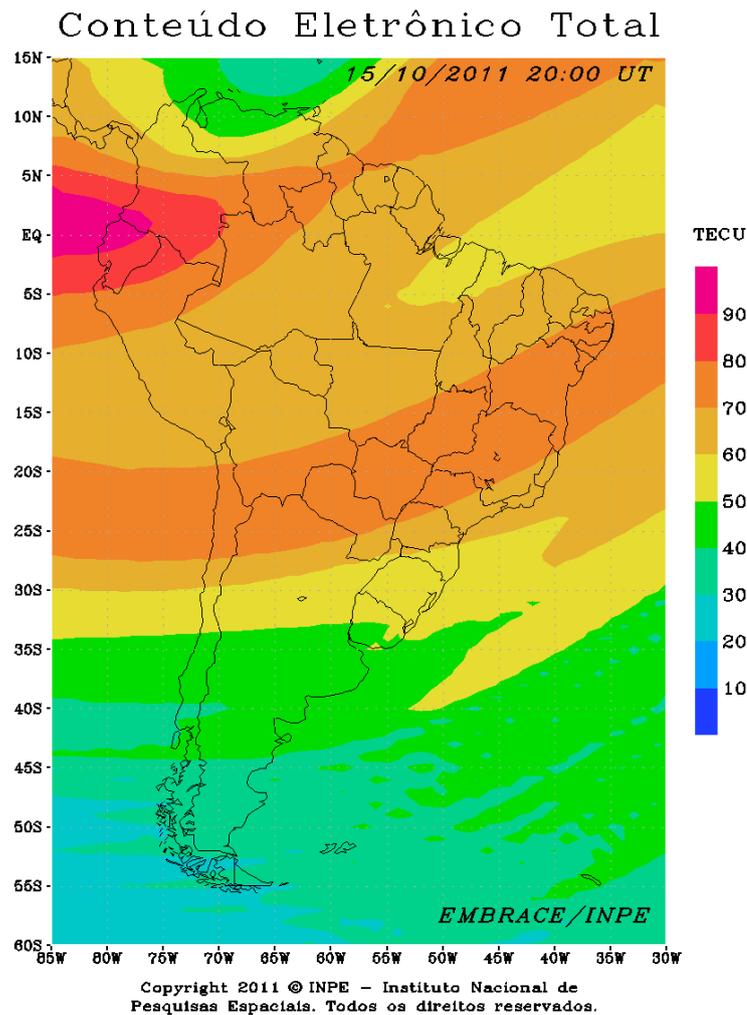


Figura 2.7 – Mapa de TEC para América do Sul as 20:00 UT (<http://200.132.24.196/site/supim/20.0h.gif>).

2.4 Ionosfera

A ionosfera é usualmente definida como região acima da atmosfera onde as partículas estão ionizadas em um número suficiente para afetar a propagação de ondas de rádio (BAILEY, 2008). A região ionosférica encontra-se aproximadamente entre altitudes de 60 e 600 km. Segundo Santana (2010), a ionosfera é definida como um meio eletricamente condutor atribuído à radiação solar sobre as camadas superiores da atmosfera. A radiação solar incidente nessa região é absorvida com diferentes taxas e frequências nas diferentes

altitudes da ionosfera. Isso permite que ocorra a formação de regiões ionizadas com diferentes características.

Devido à influência da gravidade, a atmosfera e a ionosfera apresentam uma configuração horizontal estratificada, subdividida em camadas. A estrutura da atmosfera pode ser organizada a partir da representação do seu perfil de temperatura, enquanto a organização da ionosfera tem por base a distribuição do perfil de densidade (também chamada de concentração) do plasma como podemos encontrar em KELLEY (1989). A ionosfera é composta por um estado da matéria denominado plasma, obtida a partir de outro estado da matéria, o gás.

- **Gás:** é um estado da matéria, sem forma e volume. Consiste em uma coleção de partículas cujos movimentos são particularmente aleatórios;
- **Plasma:** é um estado da matéria similar ao gás onde as partículas encontram-se ionizadas, é formado por íons positivos e elétrons.

As ondas solares ultravioletas (*UV*) são responsáveis pelo aumento de temperatura da atmosfera com o incremento de altitude. Essa mesma energia é responsável pela produção do plasma, desde que prótons solares tenham suficiente energia para ionizar a atmosfera neutra segundo KELLEY (1989), produzindo um número equivalente de íons positivos e elétrons. A densidade do plasma é representada por n e encontra seu pico na presença da luz do dia, pois é nesse período que a radiação do espectro ultravioleta é mais intensa.

A classificação das camadas ou regiões ionosféricas é baseada na radiação solar incidente entre as diferentes altitudes. Essas camadas diferem entre si basicamente por processos físicos e químicos que controlam o comportamento de cada camada. As camadas são chamadas de *D*, localizada entre 60 e 90 km, *E*, localizada entre 90 e 150 km, e *F*, sendo subdividida em *F1* e *F2*, localizada entre 150 e 500 km, como encontrado em (BAILEY, 2008). A ionosfera é um sistema dinâmico, controlado por muitos parâmetros, como emissões eletromagnéticas e variações do campo eletromagnético, sendo é extremamente sensível às mudanças atmosféricas.

2.4.1 Camadas da Ionosfera

Como já citado anteriormente a ionosfera é dividida em três camadas, denominadas *D*, *E* e *F*. As camadas da ionosfera não são estáticas e variam conforme a hora do dia, estações do ano e condições solares. Podemos visualizar as camadas da ionosfera na Figura 2.9.

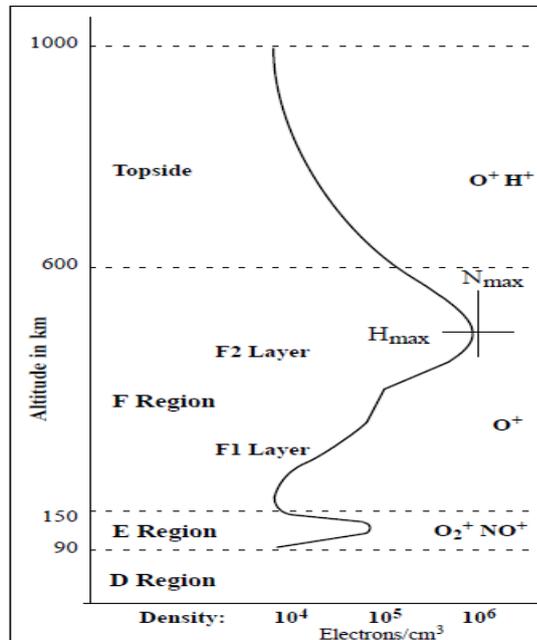


Figura 2.8 - Representação das camadas da ionosfera (Dave, A; Rowell, F.T, 1999).

2.4.1.1 Camada D

A camada *D* é a região mais inferior da ionosfera, localizada aproximadamente entre 60 e 90 km de altitude da superfície terrestre. Apresenta a menor densidade eletrônica entre todas as regiões da ionosfera. É a região onde a ionização é principalmente responsável pela absorção de ondas de rádio de alta frequência. A camada sofre uma redução substancial após o por do sol (SANTANA, 2010).

2.4.1.2 Camada E

A camada *E* é a região da ionosfera localizada entre 90 e 150 km de altitude. O principal íon nessa região é o oxigênio molecular (O_2) e ela é caracterizada por uma densidade eletrônica crescente, atingindo o seu valor máximo em torno de 120 Km de altitude (SANTANA, 2010). Assim como na camada D, sua densidade eletrônica é bastante reduzida durante a noite, pois sua fonte primária de ionização não está mais presente.

É na região *E* que se encontra a máxima condutividade elétrica da ionosfera. Essa alta condutividade é importante, pois ela permite a presença de correntes elétricas que nela fluem através da interação com o campo magnético terrestre, como encontra-se em SANTANA (2010).

2.4.1.3 Camada F

A camada *F* é a região da ionosfera localizada aproximadamente entre 150 e 600 km de altitude. É na camada *F* que se encontra o maior volume de densidade eletrônica. A camada *F* pode ser subdividida em camadas distintas chamadas *F1* e *F2*.

A camada *F1* situa-se aproximadamente entre 150 km e 200 km de altitude. O pico da camada encontra-se em torno de 180 km, onde predominam processos fotoquímicos cujo elemento majoritário é o íon de oxigênio atômico (O^+) (SANTANA, 2010).

A camada *F2* engloba toda a camada superior da ionosfera, nela encontra-se a região do pico entre 250 e 300 km aproximadamente. A densidade eletrônica na região *F2* é cerca de 10 vezes maior que a da região *E* (SANTANA, 2010). É na camada *F2* que ocorrem as bolhas de plasma ionosférico, que são regiões onde ocorrem irregularidades na densidade do plasma. Essas irregularidades podem causar degradação nos sinais de ondas de rádio que utilizam a ionosfera como meio de propagação, provocando grandes danos nas telecomunicações, principalmente via satélite, prejudicando, por exemplo, o Sistema Global de Localização (GPS-Global Positioning System) segundo Mascarenhas (2011).

2.4.2 Ionossonda

Ionossonda é um instrumento bastante representativo para medidas de parâmetros ionosféricos e tem sido a mais importante técnica desenvolvida para a investigação da estrutura global da ionosfera (SANTANA, 2010).

O princípio de funcionamento das ionossondas tem como base a reflexão e refração das ondas eletromagnéticas no plasma ionosférico (MASCARENHAS, 2011). Utilizando o princípio da reflexão e da refração da onda eletromagnética na ionosfera, pode-se calcular e observar muitas propriedades das camadas da ionosfera, através dos ionogramas obtidos com as ionossondas, segundo proposto em SANTANA (2010).

2.5 Previsão Numérica do Tempo

A Previsão Numérica do Tempo (*PNT*) implica em usar modelos atmosféricos computacionais para simular o provável estado da atmosfera para um determinado período de tempo. Segundo (Lahoz et.al., 2010) os passos essenciais na *PNT* são:

- Obter observações do estado do tempo atual;
- Assimilar as observações em um modelo numérico, para representar o estado corrente da atmosfera;
- Integrar o modelo numérico para simular a evolução futura da atmosfera;
- Geração de produtos para informar os usuários sobre a previsão do tempo;

O conjunto de dados de análise gerados em uma assimilação de dados pode ser utilizado como informação de entrada para outros processos. Entre esses processos está a previsão numérica do tempo, onde a análise representa um estado inicial de entrada da previsão, como podemos visualizar na Figura 2.10.

A qualidade da análise é um dos fatores que irão definir o grau de precisão da previsão, a previsão vai se degradar com o tempo, mesmo que a análise represente o estado preciso da atmosfera, e, como já foi antes mencionado, não é passível de representação. Assim quanto mais realística a análise, menor e mais lenta será a divergência da previsão com sua representação real.

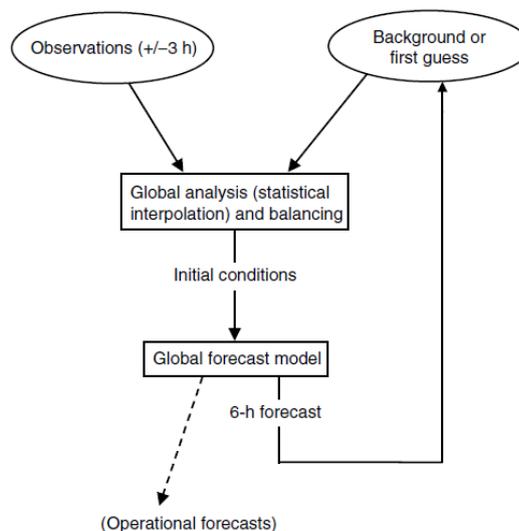


Figura 2.9–Ilustração da análise gerada por uma assimilação de dados sendo utilizada como condição inicial de uma PNT (Kalnay, E. 2003).

2.6 Observações

A informação primária de entrada para sistemas de PNT é provida de medições regulares feitas por instrumentos meteorológicos. Esses dados são conhecidos como observações. A assimilação de dados necessita que um conjunto de observações tenha sido obtido do ambiente em que se deseja realizar a análise.

As informações atmosféricas usadas em assimilação de dados não são perfeitas, elas contêm vários tipos de erros, incluindo erros dos instrumentos de medição e erros de origem humana.

As observações podem ser obtidas da atmosfera de diversas maneiras. A próxima seção apresentará o Sistema de Observação Global (*GOS – Global Observing System*), o principal mecanismo de obtenção e distribuição de informação observacional. O GOS é responsável por captar e distribuir informações principalmente meteorológicas da atmosfera. Outros centros, como os disponibilizados pela NACIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA) fornecem informações mais abrangentes do espaço, como dados da região ionosférica.

2.6.1 Sistema de Observação Global

O Sistema de Observação Global (*GOS – Global Observing System*) disponibiliza observações do estado da atmosfera e da superfície dos oceanos para a preparação de análises do tempo, previsões, avisos e alertas, monitoramento climático e atividades ambientais realizadas por programas do World Meteorological Organization (*WMO*) e outras relevantes organizações internacionais. As fontes de captura de dados observacionais são bastante variadas, como podemos visualizar na Figura 2.11.

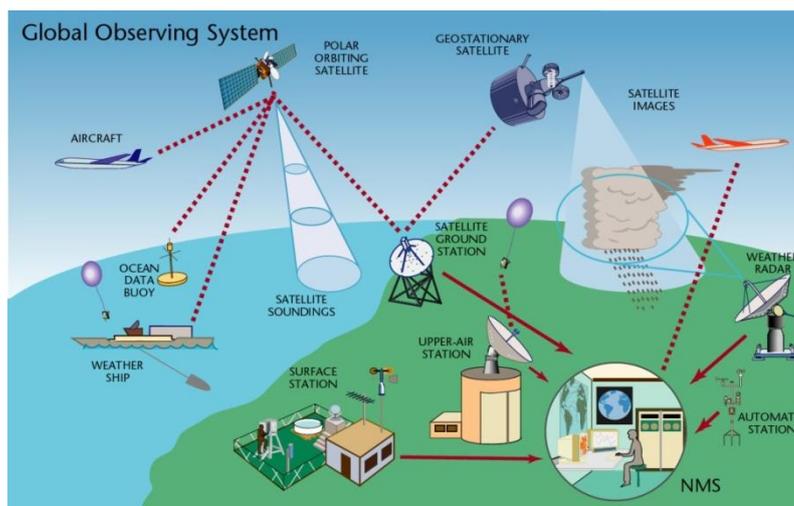


Figura 2.10 - Esquema do Sistema de Observação Global (<http://www.wmo.int/pages/prog/www/OSY/GOS.html>).

2.6.2 SPACE PHYSICS INTERACTIVE DATA RESOURCE

O SPACE PHYSICS INTERACTIVE DATA RESOURCE (*SPIDR*) foi desenvolvido para permitir acesso e gerenciamento de dados físicos do espaço. O SPIDR é uma rede distribuída e sincronizada de dados, desenvolvida pela NOAA, que pode ser acessado no seguinte sítio (SPIDR, 2011).

O SPIDR possui uma base de dados composta por estações de ionossondas espalhadas pelo planeta. Essa base de dados é composta por uma coleção de arquivos que armazenam um conjunto de informações medidas pela ionossonda, em diversos horários do dia. Entre esses dados, encontra-se a distribuição da densidade eletrônica em diferentes altitudes da ionosfera, sendo esta uma informação relevante para o trabalho.

Diariamente estações de ionossondas espalhadas globalmente medem a concentração eletrônica para sua região e atualizam a base de dados, possibilitando que usuários possam obter informações sobre o estado da ionosfera em determinado momento. Podemos visualizar na Figura 2.11 as estações de ionossondas que compõem a base de dados do SPIDR. Muitas dessas estações não estão mais ativas, o que compromete a obtenção de dados para a região onde a ionossonda está localizada, visto que há um número pequeno de estações para o espaço territorial do planeta.

● ≤ 1 year, ● > 1 year, ≤ 2 years, ● > 2 years, ≤ 5 years, ● > 5 years, ● status unavailable



Figura 2.11 – Estações de ionossondas utilizadas pelo SPIDR (SPIDR, 2011).

3 DESENVOLVIMENTO

O Capítulo 3 destina-se a apresentação das atividades realizadas para execução do trabalho final de graduação. Aqui serão apresentados os passos executados para que o objetivo do trabalho fosse alcançado. Primeiramente serão apresentadas as tarefas realizadas para obtenção dos dados utilizados no processo de assimilação da densidade eletrônica. Uma ferramenta específica para esta finalidade foi desenvolvida. Na segunda parte desse capítulo serão apresentadas as etapas realizadas para desenvolvimento da ferramenta de assimilação de dados, e a integração da mesma com o SUPIM-DAVS.

3.1 Dados observacionais ionosféricos

Para que um processo de assimilação de dados seja realizado, um conjunto de dados observacionais do ambiente em análise deve estar disponível. Essa informação observacional informará ao modelo matemático como o ambiente está se comportando para que uma correção seja realizada. No caso do projeto é necessária a utilização de valores de concentração eletrônica como dados observacionais.

Foi realizada uma pesquisa em busca de fontes de dados ionosféricos. A principal referência encontrada foi a NACIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA) encontrada no sítio (NOAA, 2011). A NOAA é uma agência do departamento de comércio dos Estados Unidos da América, que realiza pesquisas desde a superfície do sol até a superfície terrestre. A NOAA possui diversas agências especializadas, entre elas encontra-se a NATIONAL GEOPHYSICAL DATA CENTER (NGDC) encontrada no sítio (NGDC, 2011). O NGDC fornece informações de dados geofísicos através de diversos departamentos, como a Divisão Física Solar e Terrestre (STP – Solar and Terrestrial Physics Division), onde encontra-se o SPIDR. O SPIDR é uma importante fonte de dados ionosféricos utilizada no trabalho.

O SPIDR disponibiliza ionogramas gerados pelas estações de ionossondas para diversos horários. Ionogramas contêm diversas informações sobre a ionosfera. A partir dos dados ionosféricos disponibilizados pelo SPIDR foi desenvolvido um software para obtenção e tratamento desses dados. A seguir será apresentado o processo de desenvolvimento do programa de obtenção dos dados.

3.1.1 Programa de obtenção dos dados

O programa de obtenção de informações ionosféricas, mais precisamente a densidade ou concentração eletrônica, foi o primeiro passo do trabalho. Para desenvolvimento do programa foi utilizada a linguagem de programação C++. Os dados são obtidos através de arquivos obtidos por servidores que utilizam o File Transfer Protocol (FTP).

Para incluir a funcionalidade de transferência de arquivos por FTP no programa, foi utilizada a biblioteca cURL disponível em (cURL, 2011). A cURL é uma ferramenta para transferência de dados com sintaxe URL.

As classes do programa podem ser visualizadas na Figura 3.1.

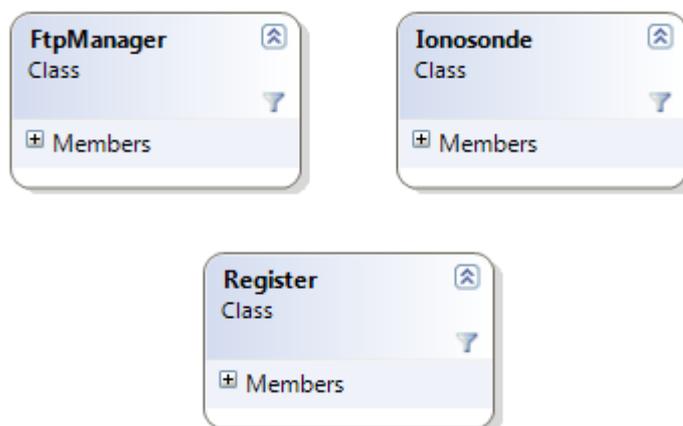


Figura 3.1 – Principais classes do programa.

- **FtpManager:** Possui os principais métodos utilizados;
- **Ionosonde:** Armazena o nome das ionossondas utilizadas pelo programa;
- **Register:** Armazena informações do arquivo, como dia, ano e horário;
- **Main:** Arquivo principal. Controla a execução do programa.
- **Util:** Não é uma classe, mas um arquivo contendo métodos de conversão entre diferentes tipos, através de templates.

O arquivo de parâmetros utilizado pode ser visualizado na Figura 3.2. O arquivo de parâmetros é composto por um cabeçalho e pela informação para execução do programa. As informações são as seguintes: ano, dia Juliano, horário, intervalo de busca, formato do arquivo buscado, nome no arquivo de ionossondas, diretório dos arquivos buscados e diretório dos arquivos gerados para a assimilação.

```
1 Year Day Target(21.0) Interval(60) Format FileIonosondas DirSAO DirConvert
2 2011
3 270
4 21.0
5 60
6 SAO
7 Ionosonde.txt
8 SAO
9 DATA
```

Figura 3.2 – Exemplo da estrutura do arquivo de parâmetros.

O programa possui dois tipos de busca. A primeira forma é a busca individual, onde a busca é realizada para uma ionossonda específica, passada como parâmetro ao executar o programa. A segunda forma de busca é a partir de um arquivo contendo uma lista de ionossondas, quando não é passado o nome da ionossonda como parâmetro. O programa carrega os parâmetros para execução do programa de um arquivo. Se a escolha de execução for a partir de um arquivo de ionossondas o programa verifica se já existe o arquivo com os nomes das ionossondas, se existir usa esse arquivo, se não, gera uma nova lista de nomes de ionossondas acessando a base de dados da NOAA e obtendo todos os diretórios de prováveis estações de ionossondas.

Após os parâmetros de execução terem sido definidos, e quais ionossondas serão usadas, é estabelecida uma conexão com o servidor FTP para cada ionossonda e é realizada a busca do arquivo conforme os parâmetros passados. Realiza-se o download dos arquivos no formato SAO. Depois de obter os arquivos de densidade eletrônica, é realizada a conversão e tratamento dos dados para o formato aceito pelo SUPIM-DAVS.

3.1.2 Estações de ionossondas

As estações de ionossondas utilizadas pelo programa estão armazenadas em um arquivo de texto. As estações de ionossonda utilizadas são as disponibilizadas pela NOAA e três estações brasileiras disponibilizadas pelo INPE, localizadas nos seguintes municípios: São Luiz no estado do Maranhão, Fortaleza no estado do Ceará e Cachoeira Paulista no estado de São Paulo.

A busca pode ser realizada em todas as estações ou em estações específicas. A Figura 2.11 apresenta as possíveis estações da NOAA. No entanto, cabe ressaltar que muitas não estão mais produzindo dados.

parâmetro de busca e o arquivo para o determinado horário têm nomes diferentes, separados algumas vezes por segundos. Para que o programa possa localizar os arquivos que satisfazem os parâmetros, uma estratégia baseada em busca por intervalos foi utilizada.

A primeira etapa da busca foi a leitura de todos os nomes dos arquivos presentes no servidor FTP da ionossonda para o dia especificado. Assim foi possível tomar conhecimento dos nomes de todos os arquivos. Após a obtenção dos nomes dos arquivos é realizada a seleção do arquivo que melhor satisfaz o horário. No arquivo de parâmetros um intervalo de busca foi definido, assim o arquivo que foi gerado mais próximo do horário de busca é selecionado. Podemos ter uma ideia do funcionamento da seleção na Figura 3.4. Definiu-se o horário de busca como sendo às 15h30min e o intervalo de busca de 15 minutos.

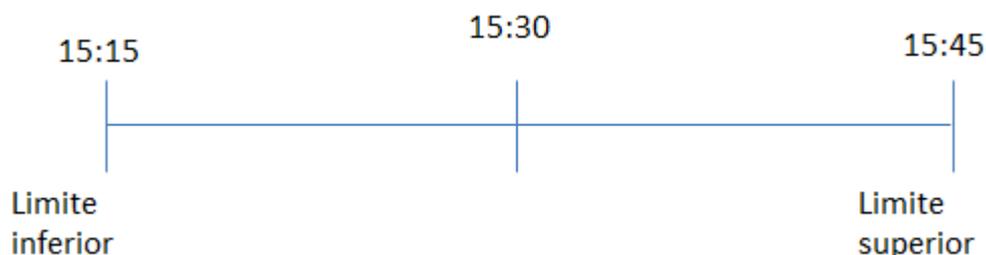


Figura 3.4 – Horário de busca para 15h30min e intervalo de 15 minutos.

Assim, o problema da nomeação dos arquivos foi resolvido. Realizando a busca pelo arquivo mais próximo do horário, e não mais pelo horário específico. Após a seleção do nome do arquivo é realizada a “montagem” do endereço FTP para realização do download.

3.1.4 Obtenção dos dados relevantes para a assimilação

Conforme mencionado anteriormente os arquivos SAO contêm diversos tipos de valores para diferentes dados ionosféricos. Para o propósito do projeto apenas a informação da densidade eletrônica e sua alturas são relevantes. Assim uma etapa para tratamento dos dados obtidos foi realizada. Essa etapa consiste na leitura dos arquivos SAO e obtenção da altitude e da densidade eletrônica. Para obter a informação, a estrutura do arquivo SAO teve que ser compreendida. O arquivo SAO é composto por um índice de dados e pelos dados. O índice de dados descreve quais informações estão presentes no arquivo. Cada índice representa um grupo, de um total de 80 diferentes grupos. Se o valor do índice for 0 (zero) significa que a determinada informação não está presente no arquivo.

A partir dos valores presentes no índice, é encontrada a posição para a densidade eletrônica para cada altitude dentro do arquivo. Os dados de densidade eletrônica são lidos para outro arquivo, para serem utilizados pela assimilação de dados do SUPIM-DAVS. A Figura 3.5 apresenta a estrutura do novo arquivo gerado.

```

1 Time 21
2 Day 270
3 Year 2011
4 LAT 0.583464
5 LON 2.20435
6 ALT 90612
7 NE 4.95e+08
8 LAT 0.583464
9 LON 2.20435
10 ALT 95000
11 NE 3.22e+09
12 LAT 0.583464
13 LON 2.20435
14 ALT 100000
15 NE 5.48e+09

```

Figura 3.5 – Parte de um arquivo utilizado pela assimilação do SUPIM-DAVS.

3.1.5 Conversão entre plasma e concentração eletrônica

Em alguns arquivos SAO não é apresentada a informação de densidade eletrônica, mas outra informação chamada de frequência de plasma. Segundo encontrado em KELLEY (1989) uma boa aproximação para a densidade eletrônica a partir do valor da frequência de plasma pode ser obtida pela Equação 3.1. Onde n representa a densidade eletrônica e fp representa a frequência de plasma.

$$n \approx \left(\frac{fp}{9000} \right)^2 \quad (3.1)$$

Assim podemos converter a frequência de plasma em densidade eletrônica, obtendo um número maior de arquivos de densidade eletrônica.

3.2 Método de Assimilação de Dados Ionosféricos

O método de assimilação de dados codificado para o trabalho foi a Interpolação Ótima. Um dos principais fatores que levaram a escolha da Interpolação Ótima foi que ela

utiliza um número menor de pontos do modelo com o objetivo de diminuir o tempo de processamento e o consumo de memória. Assim, a Interpolação Ótima encaixou-se no projeto no qual esse trabalho está inserido, pois o sistema SUPIM-DAVS trabalha com um número elevado de valores, impossibilitando o uso de vários métodos de assimilação de dados.

Para o desenvolvimento do método foi utilizada a linguagem de programação C++. Duas bibliotecas externas foram utilizadas para codificação do método. A biblioteca para busca de dados n-dimensionais Approximate Nearest Neighbor (ANN) (ANN, 2010) e a biblioteca para cálculos numéricos uBLAS (BOOST, 2011).

Como anteriormente mencionado, esse trabalho está incluído em um projeto mais amplo, sendo assim desenvolvido com base nas características do software SUPIM-DAVS. O método de assimilação de dados foi definido por uma classe denominada **OI**. Essa classe pode ser visualizada na Figura 3.6, destacada no diagrama de classes do sistema SUPIM-DAVS. Esse diagrama de classes representa todo o sistema de previsão da dinâmica da ionosfera, a execução do sistema é todo controlado pela classe **Kernel**. A classe **OI** faz parte da família de objetos **DataAssimilation** que foi implementada com o padrão de projeto de software **Abstract Factory** que possibilita a definição de famílias de objetos relacionados ou dependentes, através de uma única interface e sem que a classe concreta seja especificada.

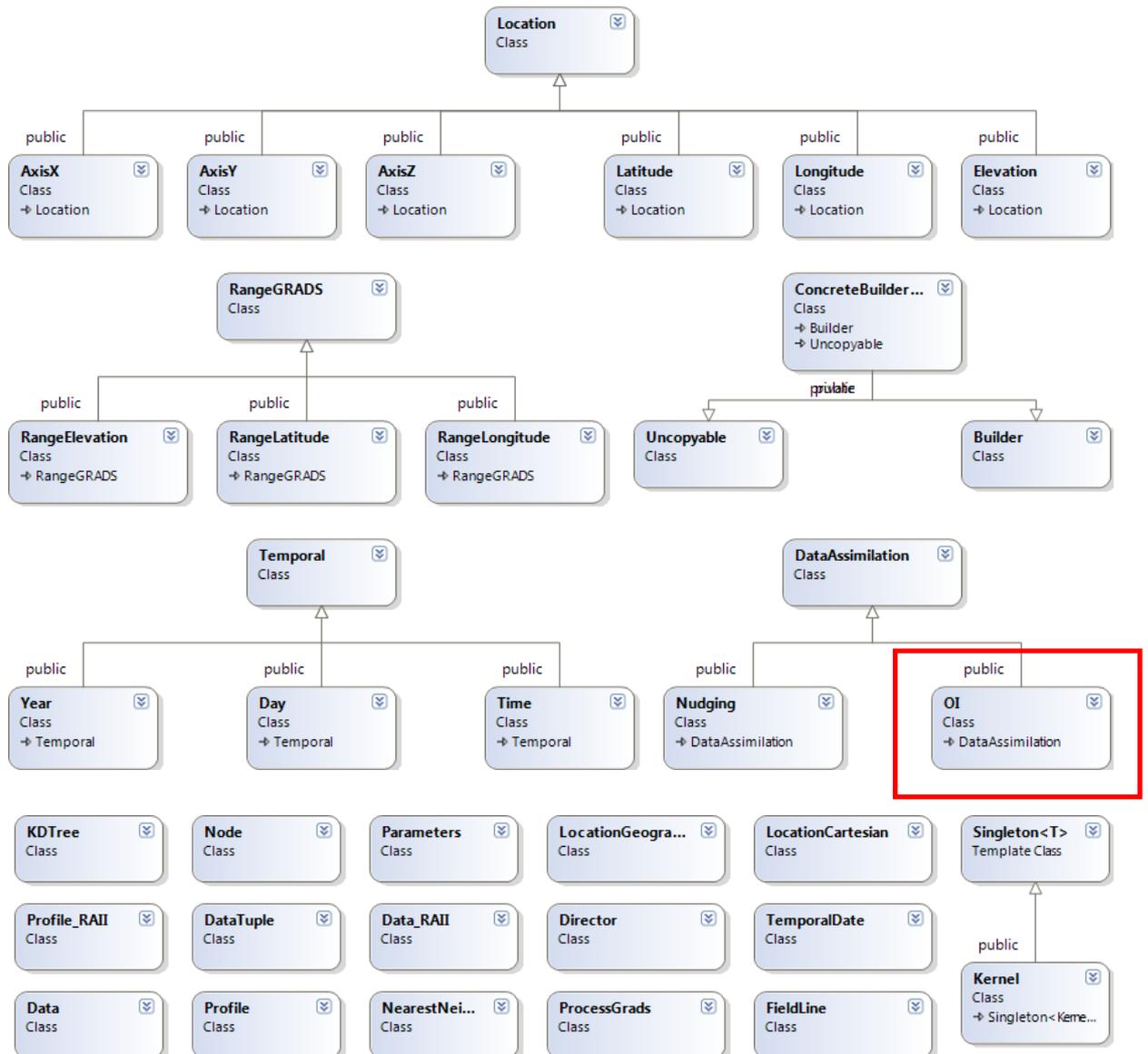


Figura 3.6 – Diagrama de classe do SUPIM-DAVS. A classe OI está destacada pelo retângulo vermelho.

3.2.1 Busca dos pontos mais próximos

Como já comentado no texto o método OI trabalha com um número menor de pontos do que outros métodos de assimilação. Não é possível realizar uma assimilação completa para os valores modelados pelo SUPIM, assim um número menor de pontos foi selecionado. Os pontos do SUPIM, que são utilizados no processo de assimilação, são definidos com base no arquivo de ionossonda que é composto por medições de densidade eletrônica para diversas altitudes.

Para cada ponto do arquivo de ionossonda é selecionado um conjunto de pontos de densidade eletrônica provenientes do SUPIM. Os pontos presentes nesse conjunto, serão

influenciados pelos valores das densidades eletrônicas do arquivo de observação. Isso é feito para cada arquivo de ionossonda, ou seja, cada arquivo de ionossonda vai possuir um conjunto de valores que serão afetados por ele no processo de assimilação. Para encontrar os valores mais próximos para cada valor do arquivo de ionossonda foram utilizados métodos disponibilizados pelo ANN.

3.2.2 Definição das matrizes de covariância de erros

Para o cálculo das matrizes de covariância de erros B e R foi utilizado o método NMC descrito em Kalnay (2003). O método NMC utiliza a diferença entre pequenos intervalos de previsão para o mesmo instante de tempo. A Equação 3.2 representa o método NMC.

$$B \approx E \left\{ [x_f(48 h) - x_f(24 h)][x_f(48 h) - x_f(24 h)]^T \right\} \quad (3.2)$$

É calculada uma matriz B e R para cada arquivo ionossonda. Para o cálculo da matriz B não é possível manter em memória os valores de densidade eletrônica dos arquivos do SUPIM para o dia atual e para o dia anterior. Primeiro é carregado os arquivos para o dia anterior e são buscados os pontos mais próximos para cada arquivo de ionossonda. Após esse procedimento a memória é liberada e os arquivos para o dia atual são carregados para a memória e onde são buscados os pontos mais próximos para cada arquivo de ionossonda. Com as informações do dia anterior e do dia atual é possível estimar a matriz B .

Para o cálculo da matriz R todos os arquivos de ionossonda para o dia atual e para o dia anterior são carregados para a memória, e suas respectivas matrizes R são estimadas.

As matrizes de covariância de erros são muito importantes, pois elas são construídas de forma estatística e definirão a direção da análise gerada.

3.2.3 Operador de Observação

A matriz H é conhecida como operador de observação e tem por finalidade conduzir o espaço do modelo para a mesma dimensão das observações, para que possam ser comparadas.

Trabalhamos com um número maior de valores de densidade eletrônica, provenientes do SUPIM em relação ao número de valores de observação. Assim o operador H é utilizado para que seja possível comparar essas informações com diferentes dimensões.

Para cada valor de observação, há uma linha em nosso operador H . Essa linha funciona como um operador de interpolação, e cada linha gera um novo valor a partir das variáveis do modelo que pode então ser comparada com a observação referente a linha.

Para construir o operador H usou-se o método do inverso do quadrado da distância (IDW – Inverse Distance-Weighting) (AVELLANEDA, 2007), que trabalha com a ideia de que os pontos do SUPIM mais próximos da observação têm maior influência na estimativa do valor a ser comparado com a observação.

Para cada observação $y(i)$, para i variando de $i = 1$ até $i = p$, onde p é a dimensão do vetor de observações, e para cada valor de estado $x(j)$, com $j = 1$ até $j = n$, onde n é a dimensão do vetor de valores do SUPIM utiliza-se a Equação 3.3 para gerar os valores para cada posição da matriz H utilizando a distância do valor do SUPIM para a observação que representa a linha i . Após isso, cada valor é dividido pelo somatório das distancias de todos os valores de sua linha, utilizado a Equação 3.4.

$$H(i, j) = \frac{1.0}{dist} \quad (3.3)$$

$$H(i, j) = \frac{H_i(i, j)}{\sum_{l=1}^n H_i(i, l)} \quad (3.4)$$

3.2.4 Cálculo da Matriz Inversa

Em determinada etapa das operações matriciais realizadas pelo método, necessita-se de uma inversão de matriz. Essa inversão causou bastante problema. Os valores gerados pela inversão da matriz não eram os esperados, quando comparados com a mesma inversão utilizando o MATLAB. Acredita-se que a imprecisão no resultado na inversão esteja ligada aos valores elevados que ela possui. Os primeiros métodos de inversão utilizados não apresentaram bons resultados. Para solucionar o problema buscou-se implementações alternativas de inversão de matrizes, encontrando-se uma codificação em C++ do algoritmo de eliminação de Gauss–Jordan, codificado e disponibilizado por colaboradores da uBLAS.

3.2.5 Etapas do processo de assimilação

Os valores utilizados no processo de assimilação pelo método são obtidos de arquivos. Toda informação que não necessita mais ser utilizada é liberada da memória, pois a quantidade de memória utilizada é elevada.

Os primeiros cálculos realizados pelo método são as definições das matrizes B e R . Após esse passo é construída a matriz H utilizando o método IDW .

As operações matriciais são realizadas até que se obtenha a análise. A Figura 3.7 representa um fluxograma simples com as etapas realizadas.

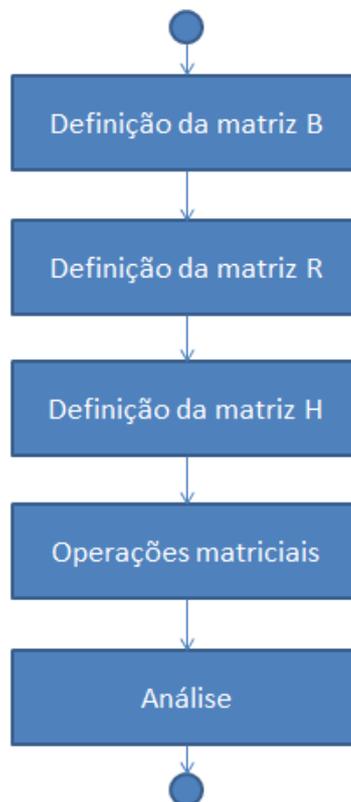


Figura 3.7 – Fluxograma que representa de forma simples o funcionamento do método.

4 RESULTADOS

Este capítulo é destinado à apresentação dos resultados obtidos com um caso de teste selecionado para demonstrar o processo de assimilação. Os resultados são provenientes da versão atual do software, que por se tratar de um projeto do grupo de pesquisa do clima espacial, ainda está em desenvolvimento. Modificações e aprimoramentos devem ser realizadas para aumentar a qualidade de processo de assimilação.

Para demonstrar os resultados obtidos pelo processo de assimilação, foi definida uma data específica para realizar a assimilação nos valores de densidade eletrônica, fornecidos pelo SUPIM. Definiu-se o dia juliano 303 do ano de 2011 (equivalendo ao dia 30/10/2011). A hora definida para a assimilação foi as 21h. A Figura 4.1 representa de forma simplificada esse processo.

O teste aqui especificado é inviável de ser realizado em arquiteturas de computadores comuns devido a grande necessidade de processamento requerida. Assim, ele foi realizado em um conjunto de nós que formam um *cluster* de computadores.

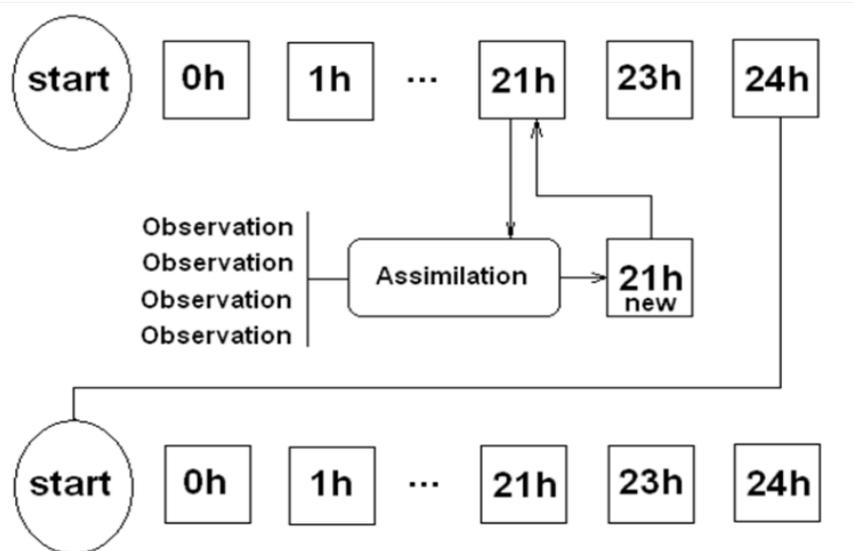


Figura 4.1 – Representação do processo de assimilação na geração dos mapas de TEC.

4.1 Ambiente de Execução

A execução do caso de teste foi realizada no cluster EMBRACE localizado no Centro Regional Sul de Pesquisas Espaciais (CRS/INPE), utilizado nas atividades do Laboratório de

Computação para o Clima Espacial (LABCCE). A configuração do *cluster* pode ser visualizada na Tabela 1.

Tabela 1- Configuração do Ambiente de Execução.

TIPO DO NÓ	PROCESSADOR	NÚCLEOS	RAM	CACHE
HP DL360 G7 – 5 UNIDADES	2 Processadores Xeon 2.53GHz	8	16 GB	12 MB L3
HP DL 380 G7 – 2 UNIDADES	2 Processadores Xeon 2.93GHz	12	32 GB	12 MB L3
HP P4300 G2 - 2 UNIDADES	1 Processador Xeon 2.27GHz	4	4 GB	8 MB L3

4.2 Parâmetros de execução

A execução do método é controlada por um *script* que controla a execução completa de todos os programas utilizados na simulação da dinâmica da Ionosfera. Esse script trabalha com um arquivo de configuração onde estão definidos os parâmetros que se deseja utilizar na simulação.

O processo de assimilação é realizado para a região da América do Sul, e foram utilizadas duas ionossondas localizadas em Jicamarca e Cachoeira Paulista. O número de ionossondas é menor do que as ionossondas disponíveis, mas até o momento são as que não apresentaram problemas. Usaram-se diferentes raios de influência para demonstrar o comportamento do método.

4.3 Perfil eletrônico nas estações de ionossonda

Com o objetivo de apresentar a situação dos valores de densidade eletrônica após a assimilação, utilizou-se o software GrADS para plotar o perfil eletrônico antes e após a assimilação nas coordenadas exatas onde se encontram as estações de ionossondas.

As Figuras 4.2, 4.3 e 4.4 representam o perfil eletrônico obtido na localização da estação de ionossonda de Jicamarca. A curva representada pelos pontos verdes representa os

valores de densidade antes do processo de assimilação. A curva representada pelos pontos brancos representa os valores de densidade após o processo de assimilação.

Os perfis de densidade eletrônica gerados para Jicamarca apresentam um resultado semelhante aos valores originais do SUPIM. Mas é bastante visível que ao aumentar o raio de influência da estação de ionossonda a análise começa a divergir dos valores do SUPIM. Esse comportamento não era esperado, e uma possível explicação esteja relacionada ao fato de que a técnica de interpolação de dados utilizada para gerar os pontos do SUPIM, antes ou após a assimilação, é bastante diferente da técnica utilizada para escolha dos pontos a serem assimilados, baseada na definição de um raio de influência. Notou-se também, que com o aumento do raio, pontos de análise negativos começam a surgir. Isso não deveria acontecer, podendo ser ocasionada pela inversão da matriz ou algum tipo de erro conceitual, ficando a solução, se ela existir, como um trabalho ainda a ser desenvolvido.

Nas Figuras 4.5, 4.6 e 4.7 está representado o perfil eletrônico obtido na localização da estação de ionossonda de Cachoeira Paulista. Os resultados da análise para essa região divergiram bastante dos resultados originais do SUPIM antes da assimilação, principalmente no pico da densidade eletrônica. Uma das possíveis razões para esse fato pode ser que essa ionossonda possui uma distribuição de pontos de observação concentrada nas menores e maiores altitudes, sendo muito esparsa na região central, o seu pico. Isso pode estar ocasionando que a análise para a região central esteja bastante modificada em relação aos valores originais.

Ao demonstrar os perfis o objetivo era que fosse possível visualizar as modificações antes e após o processo de assimilação. Com essas imagens podemos visualizar possíveis problemas que possam estar ocorrendo, para futuras correções do método.

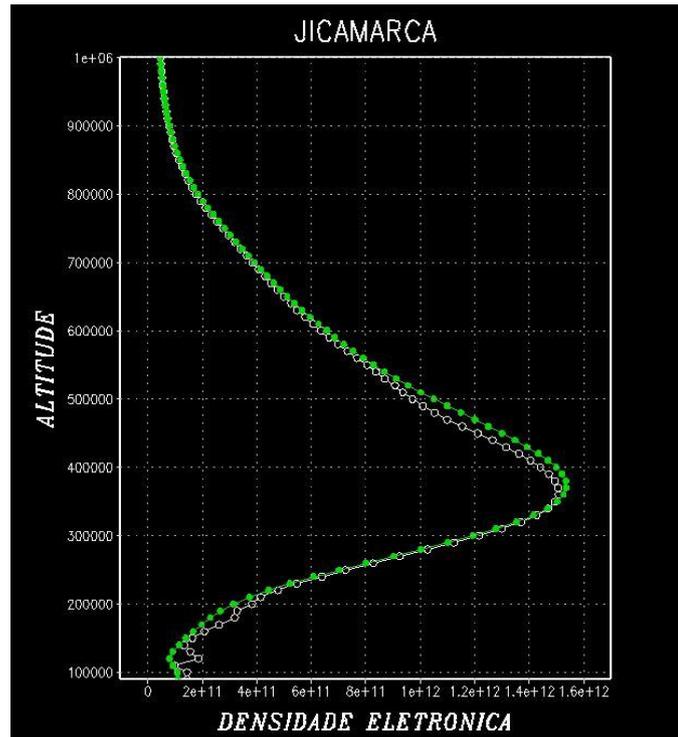


Figura 4.2 – Perfil eletrônico para Jicamarca com raio de 50.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação.

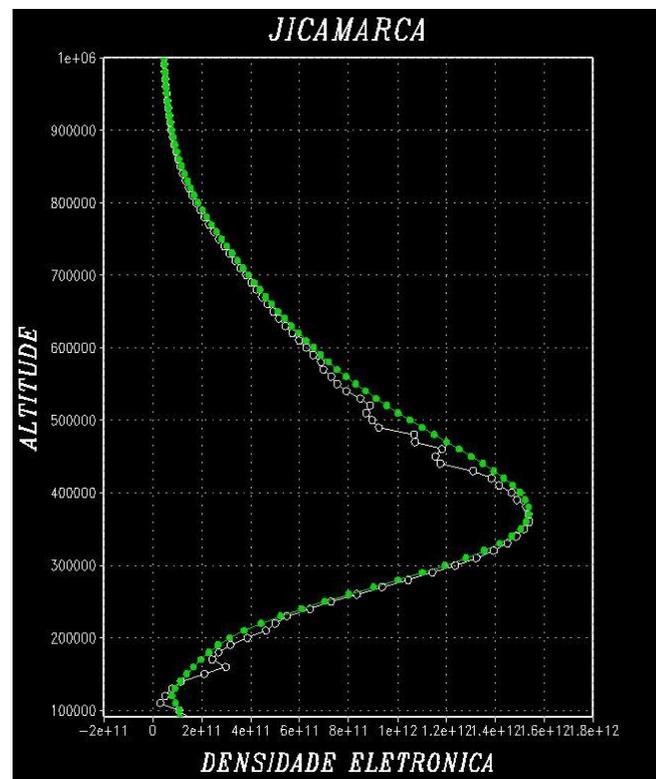


Figura 4.3 - Perfil eletrônico para Jicamarca com raio de 100.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação

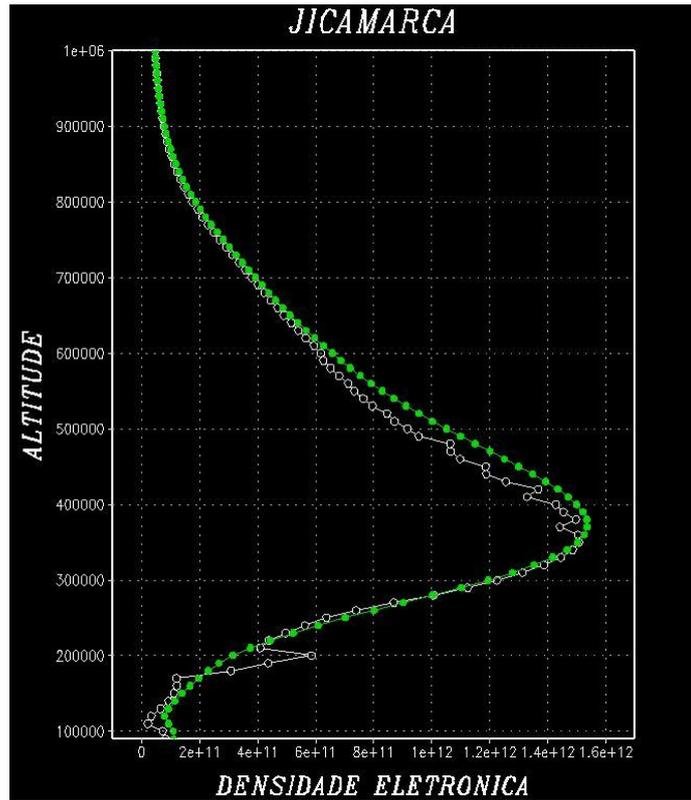


Figura 4.4 - Perfil eletrônico para Jicamarca com raio de 150.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação

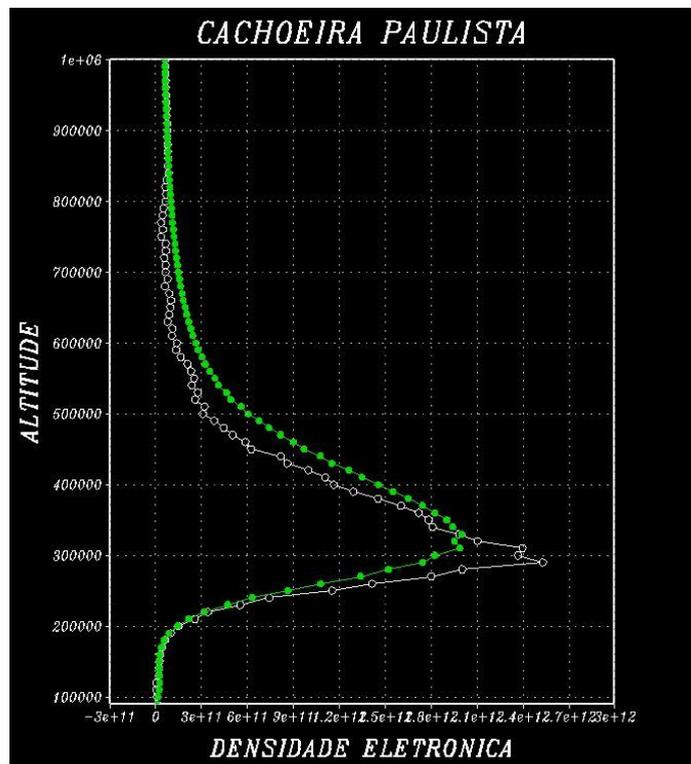


Figura 4.5 - Perfil eletrônico para Cachoeira Paulista com raio de 50.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação

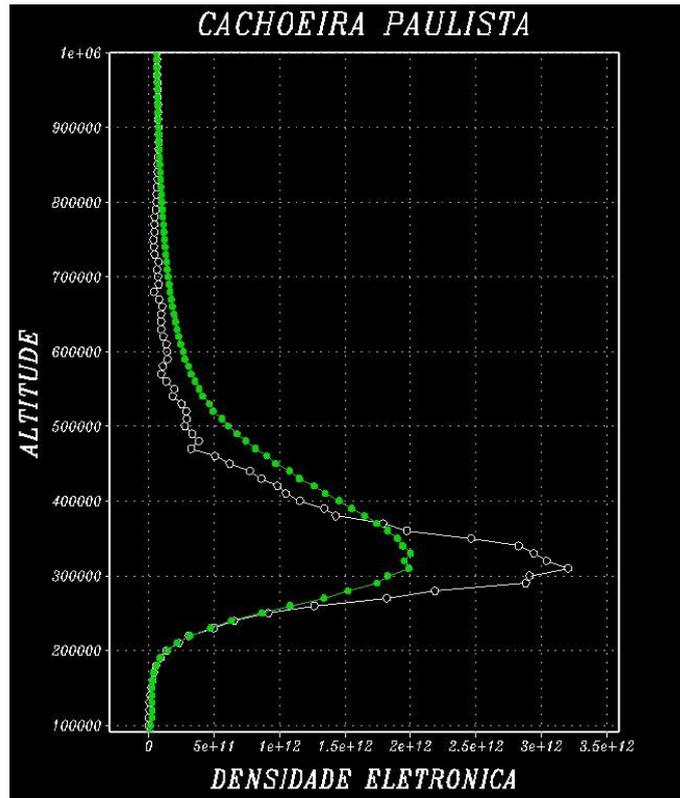


Figura 4.6 - Perfil eletrônico para Cachoeira Paulista com raio de 100.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação

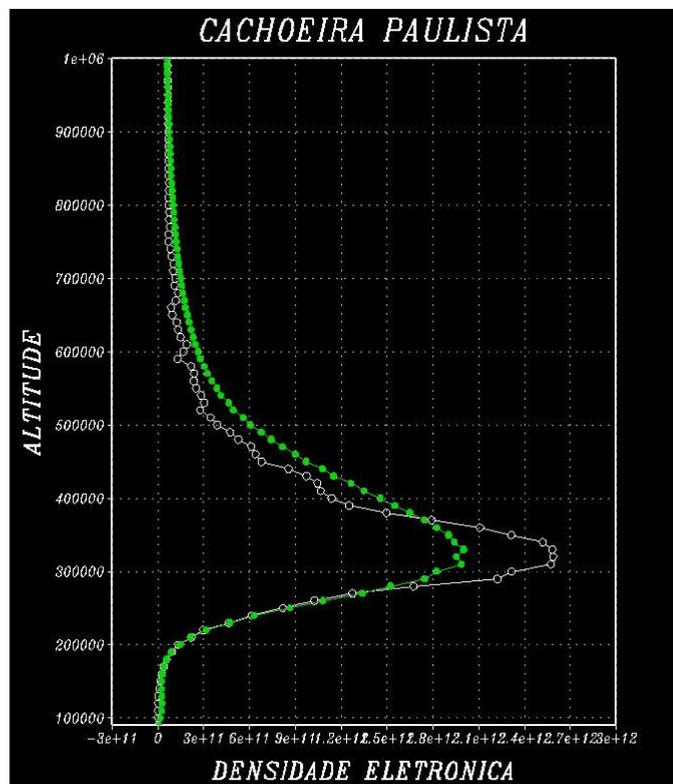


Figura 4.7 - Perfil eletrônico para Cachoeira Paulista com raio de 150.000 metros. Os pontos verdes representam os valores da densidade antes da assimilação e os pontos brancos indicam os valores da densidade após a assimilação

4.4 Análise interpolada para a dimensão das observações

Após a geração da análise, ela foi comparada com os pontos do SUPIM e as observações. Como a dimensão das observações é menor, utilizou-se o operador H para levar a análise e os pontos do SUPIM para a dimensão das observações.

Utilizou-se o software MATLAB para gerar os gráficos. Os gráficos gerados pelo MATLAB apresentam o valor da densidade eletrônica das observações, da análise interpolada com o operador H e dos pontos do SUPIM interpolados com o operador H , ordenados pela altitude dos pontos de observação, assim, enquanto maior o índice do ponto, maior sua altitude. Todas as informações plotadas foram geradas pelo programa em C++, sendo a única funcionalidade do MATLAB a geração dos gráficos. Esse tipo de análise é importante para eliminar possíveis problemas oriundos da aplicação de técnicas de interpolação para geração do perfil eletrônico.

As Figuras 4.8, 4.9 e 4.10 representam a análise para a ionossonda de Cachoeira Paulista em São Paulo. Pode-se perceber ao analisar as imagens que com o aumento do raio de influência da ionossonda tanto os valores assimilados, quanto os valores do SUPIM começam a se afastar dos valores da densidade eletrônica da ionossonda. Esse comportamento não era esperado, e pode estar associado a problemas na implementação da rotina de inversão matricial, ou estimação das matrizes de erros B e R .

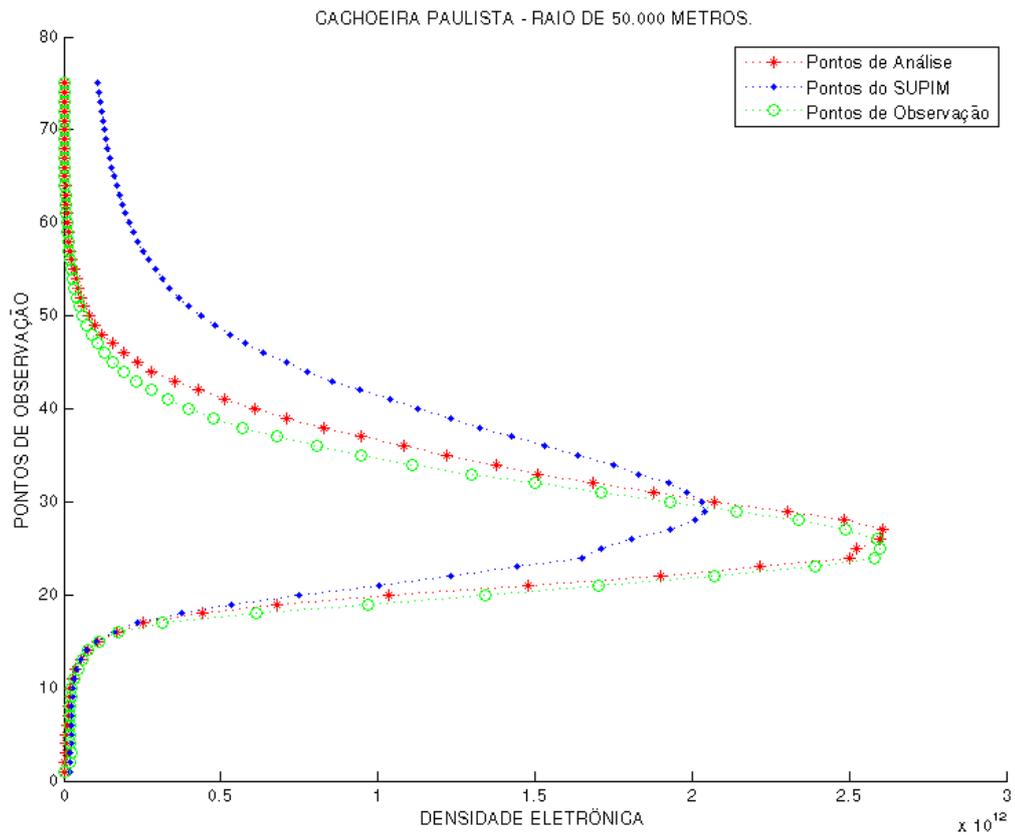


Figura 4.8 – Análise para a ionossonda de Cachoeira Paulista utilizando raio de 50.000 metros.

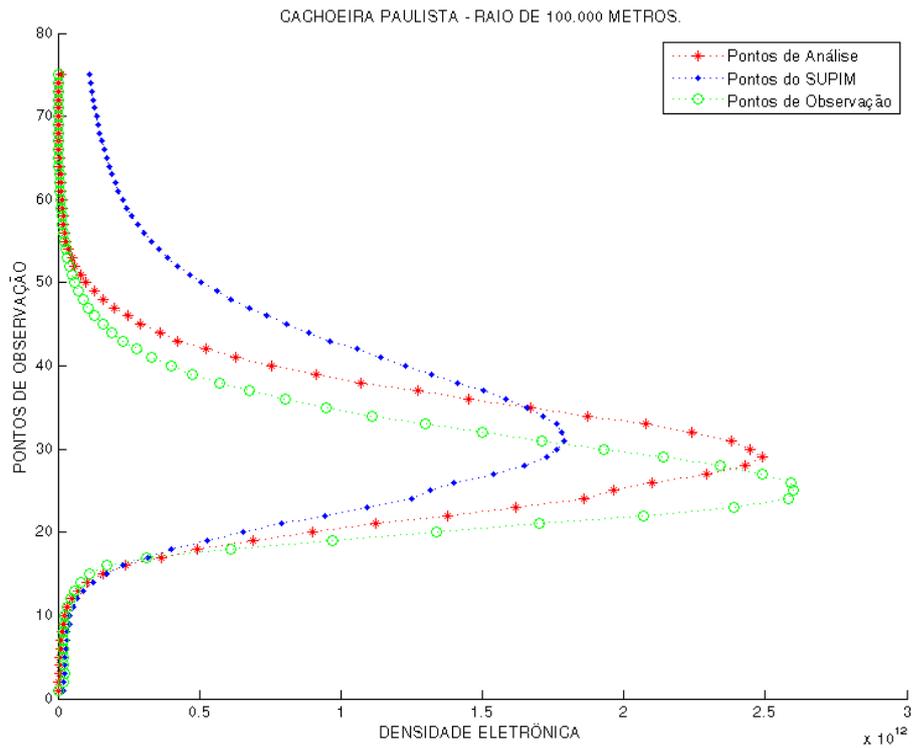


Figura 4.9 - Análise para a ionossonda de Cachoeira Paulista utilizando raio de 100.000 metros.

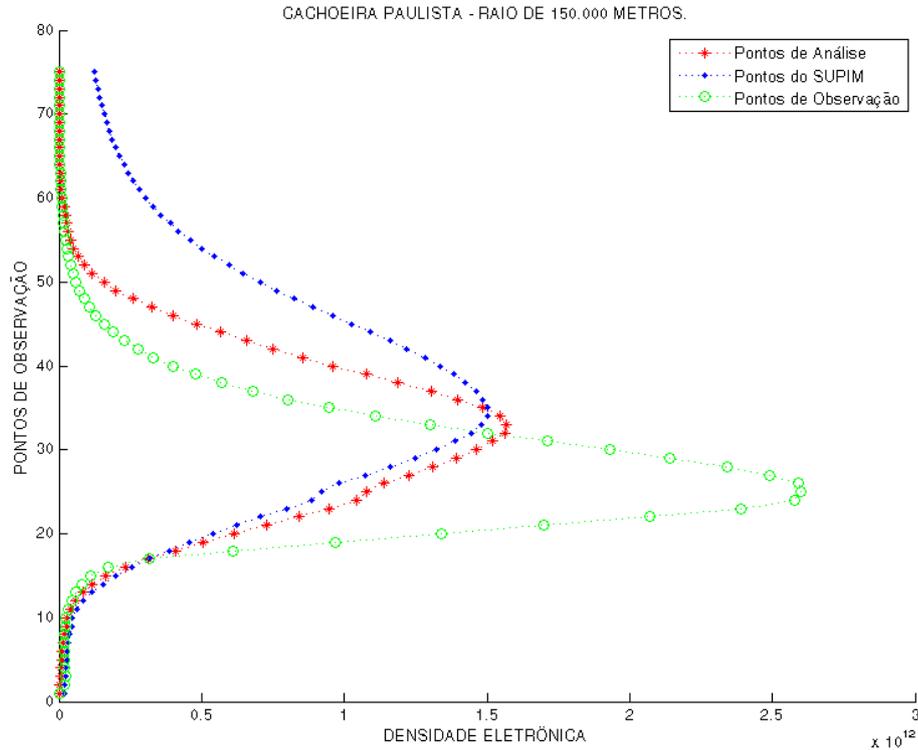


Figura 4.10 - Análise para a ionossonda de Cachoeira Paulista utilizando raio de 150.000 metros.

As Figuras 4.11, 4.12 e 4.13 representam a análise para a ionossonda de Jicamarca no Peru. Essa ionossonda, ao contrário de Cachoeira Paulista apresenta resultados mais consistentes. Podemos perceber que para os três raios de influência, as observações e os pontos do SUPIM foram considerados para geração da análise.

É importante mencionar que o comportamento mais irregular mostrado nas figuras 4.2, 4.3, 4.4, 4.5, 4.6, e 4.7 não foi verificado aqui. Isso indica que esse tipo de anomalia é provavelmente decorrente da técnica de interpolação de dados utilizada pelo SUPIM-DAVS, uma vez que os resultados mostrados nas figuras 4.8, 4.9, 4.10, 4.11, 4.12 e 4.13 não mostraram o mesmo tipo de irregularidade. A ideia é construir a matriz R , a partir de um vetor com a diferença de densidade eletrônica para os mesmos pontos em diferentes dias. Mas notou-se que os arquivos não apresentam medições para as mesmas coordenadas entre diferentes dias. Sendo uma possível causa das imprecisões da análise.

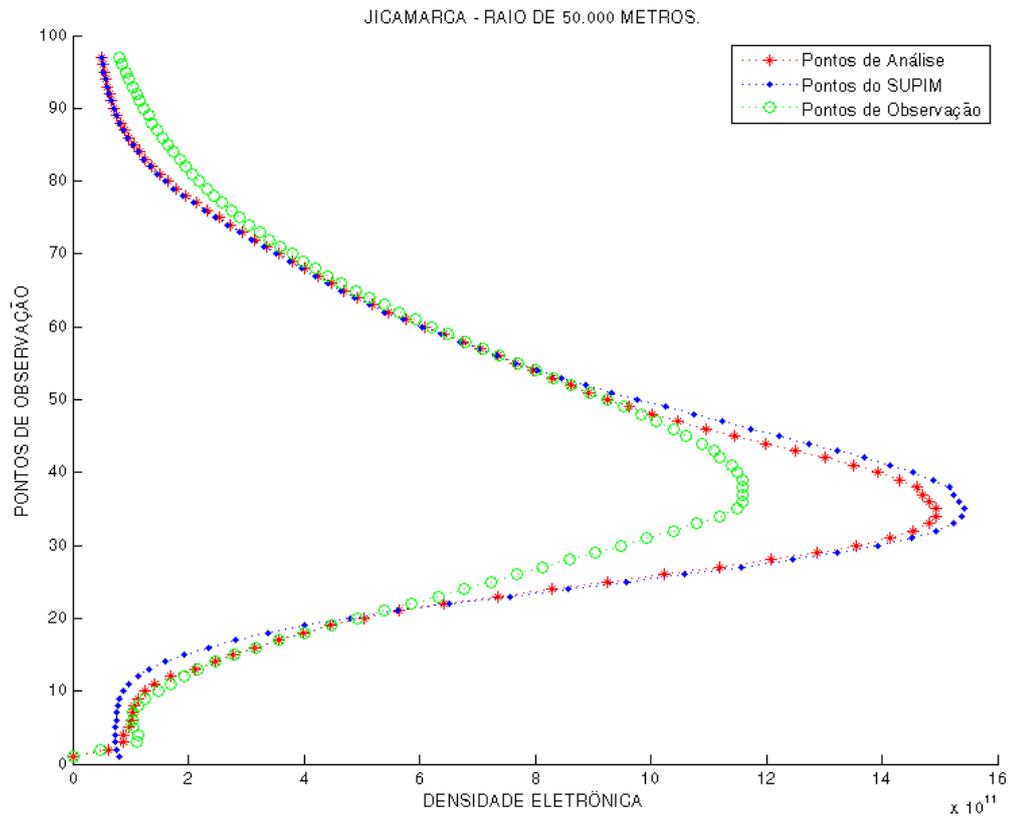


Figura 4.11 - Análise para a ionossonda de Jicamarca utilizando raio de 50.000 metros.

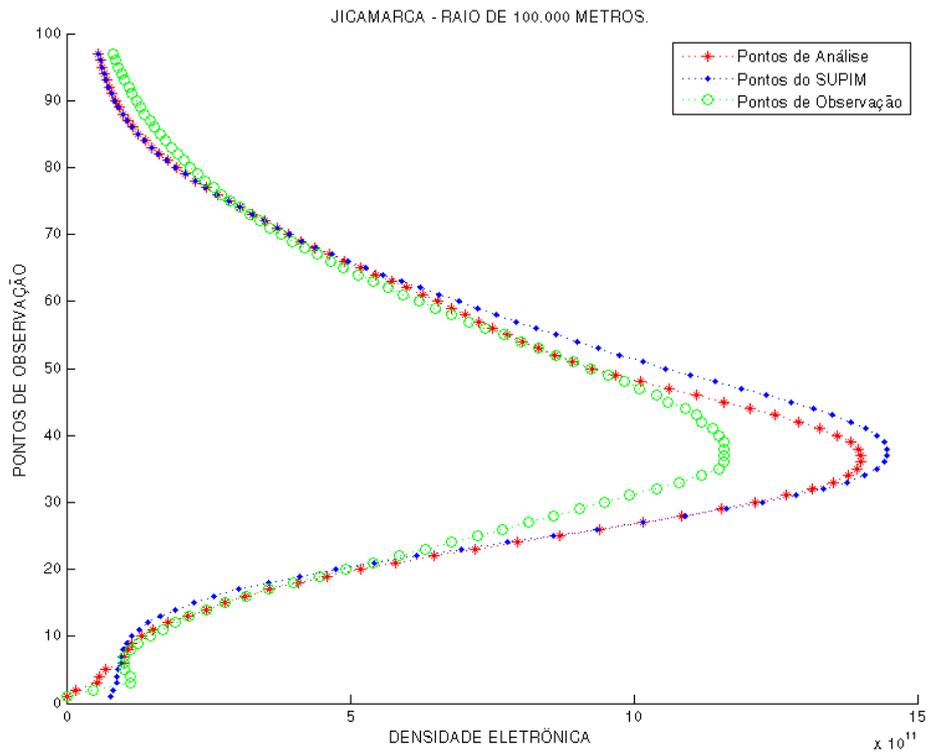


Figura 4.12 - Análise para a ionossonda de Jicamarca utilizando raio de 100.000 metros.

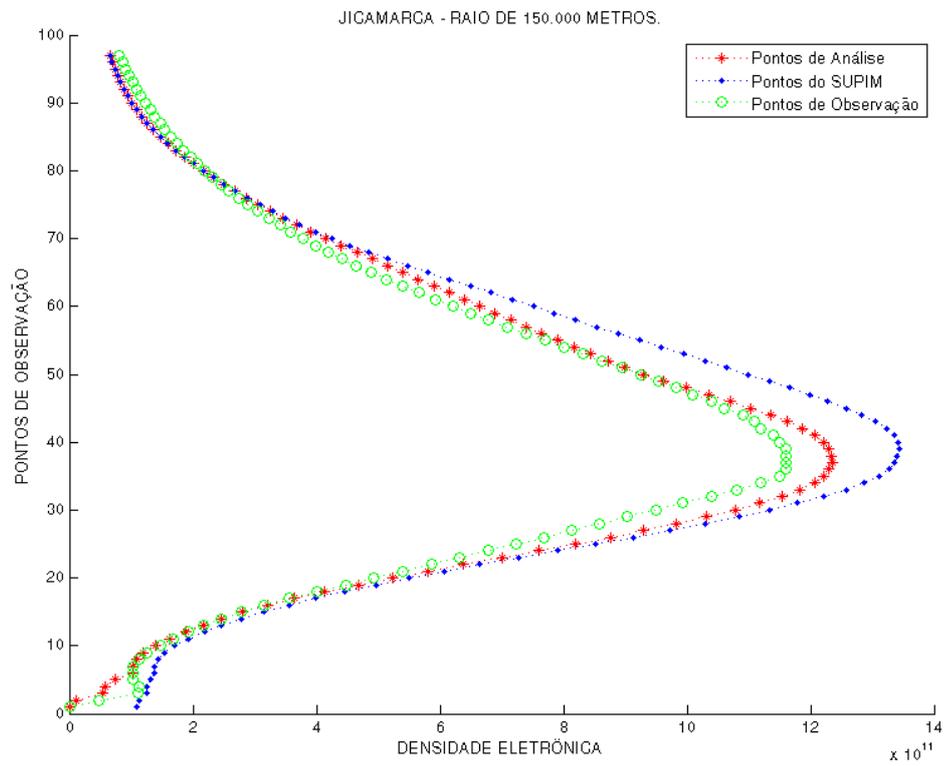


Figura 4.13 - Análise para a ionossonda de Jicamarca utilizando raio de 150.000 metros.

5 CONCLUSÃO

Esse trabalho faz parte do projeto SUPIM-DAVS, ele têm por finalidade adicionar funcionalidade de assimilação de dados ao software de previsão da dinâmica da ionosfera. O SUPIM-DAVS já possui um método de assimilação de dados, mas com uma abordagem diferente. Codificou-se a técnica Interpolação Ótima para assimilar os valores de densidade eletrônica. Essa técnica é utilizada quando se têm muitos valores a ser assimilados, o que dificulta o processo de assimilação, seja por falta de memória ou pelo elevado tempo de processamento.

Anteriormente à implementação da técnica de assimilação de dados, um sistema de obtenção de dados observacionais foi desenvolvido. Esse sistema objetiva a busca de dados em base de dados de estações de ionosondas. Isso é importante, pois sem essa informação a segunda parte do trabalho não teria como ser desenvolvida. A informação obtida deve ser tratada e convertida para o formato que o SUPIM-DAVS utiliza.

Dificuldades conceituais sobre o funcionamento do método e a estratégia de definição de matrizes utilizadas no processo de assimilação foram bastante recorrentes. Há dúvidas sobre algumas estratégias utilizadas para o desenvolvimento que devem ser solucionadas com o andamento do projeto, como a definição da matriz R. Pelo fato da assimilação de dados ser algo recente na previsão da dinâmica da Ionosfera, não há um conhecimento prévio para servir de base.

Nas implementações iniciais, problemas de consumo de memória ocorreram, impossibilitando o uso de raios de influência muito grandes. Raios de influência muito grandes também ocasionaram um elevado tempo de processamento nas operações matriciais. Tentou-se utilizar estratégias no algoritmo utilizado para minimizar esses problemas.

A quantidade pequena de fontes de dados observacionais a que o grupo de pesquisa tem acesso, especialmente para a América do Sul, região para a qual está sendo realizada a previsão do conteúdo eletrônico total, é uma dificuldade. Essa escassez de dados observacionais acarreta em imprecisões no processo de assimilação, degradando a análise quando se tem um raio de influência muito elevado.

Acredita-se que o trabalho de conclusão aqui apresentado seja útil na previsão da dinâmica da ionosfera. Sendo uma fonte de consulta para pessoas interessadas nessa área, pois alguns conceitos podem ser úteis para resolver problemas de outros pesquisadores da dinâmica da ionosfera, além de serem apresentados problemas enfrentados no desenvolvimento do programa. O trabalho utilizou técnicas de assimilação de dados em uma área onde ainda não havia sido aplicada, assimilação de dados provenientes do software SUPIM, tornando-se uma contribuição para quem deseja realizar pesquisas nessa área.

Os trabalhos futuros estarão concentrados em melhorar a precisão da assimilação, diminuir o consumo de memória, diminuir o tempo de processamento da assimilação, que atualmente é uma parte bastante relevante do tempo de execução do SUPIM-DAVS.

REFERÊNCIAS

- ANN - Approximate Nearest Neighbor. Disponível em <<http://www.cs.umd.edu/~mount/ANN/>>. Acesso em novembro de 2011.
- AVELLANEDA, D. R. Spatial interpolation techniques for estimating levels of pollutant concentrations in the atmosphere, 2007. Disponível em: <<http://www.scielo.org.mx/pdf/rmf/v53n6/v53n6a4.pdf>>. Acesso em: novembro de 2011.
- BAILEY, G. SHEFFIELD UNIVERSITY PLASMASPHERE IONOSPHERE MODEL (SUPIM), 1970. Disponível em < <http://gbailey.staff.shef.ac.uk/supim.html>>. Acesso em setembro de 2011.
- Boost C++ Libraries, 2011. Disponível em < <http://www.boost.org/>>. Acesso em outubro de 2011.
- cURL. Disponível em < <http://curl.haxx.se/>>. Acesso em agosto de 2011.
- Dave, A; Rowell, F, T. Space Environment TOPICS: The Ionosphere. Disponível em: <<http://www.swpc.noaa.gov/info/Iono.pdf>>. Acesso em: Outubro de 2011.
- European Centre for Medium-Range Weather Forecast, ECMWF. Bouttier, F; Courtier, P. Data assimilation concepts and methods. 1999. Disponível em: < http://www.ecmwf.int/newsevents/training/rcourse_notes/pdf_files/Assim_concepts.pdf >. Acesso em: setembro de 2011.
- Grid Analysis and Display System, 2011. Disponível em: <<http://www.iges.org/grads/>>. Acesso em setembro de 2011.
- Lahoz, W; Khattatov, B; Ménard, R. Data Assimilation: Making Sense of Observations. New York: Springer, 2010. 718 p.
- Kalnay, E. Atmospheric Modeling Data Assimilation And Predictability. Cambridge University Press, 2003, 341 p.
- KELLEY, C.M. The Earth's Ionosphere: Plasma Physics and Electrodynamics. Academic Press 1989, 484 p.

MATLAB , 2009. Disponível em <<http://www.mathworks.com/products/matlab/>> Acesso em dezembro de 2011.

MASCARENHAS, M. Estudo de irregularidades ionosféricas através de simulação numérica e dados experimentais. 2011. 195 p. (sid.inpe.br/mtc-m19/2010/11.26.19.57-TDI). Tese (Doutorado em Geofísica Espacial) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010. Disponível em:<<http://urlib.net/8JMKD3MGP7W/38LSM62>>. Acesso em: setembro de 2011.

NGDT – Nacional Geophysical Data Center. Disponível em <<http://www.ngdc.noaa.gov/ngdc.html>>. Acesso em: Outubro de 2011.

NOAA - NACIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION. Disponível em <<http://www.noaa.gov/>>. Acesso em: outubro de 2011

SANTANA, W. M. Métodos não lineares aplicados ao estudo da dinâmica ionosférica. 2010. 222 p. (sid.inpe.br/mtc-m19/2010/10.06.17.24-TDI). Tese (Doutorado em Geofísica Espacial) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010. Disponível em: <<http://urlib.net/8JMKD3MGP7W/38CJC52>>. Acesso em: 15 out. 2011.

SANTOS, P. M. T. Estudo da dinâmica do sistema termosfera - ionosfera por modelo teórico e observações experimentais. 2005. 134 p. (INPE-13270-TDI/1032). Tese (Doutorado em Geofísica Espacial) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005. Disponível em: <<http://urlib.net/sid.inpe.br/iris@1913/2005/04.28.13.11>>. Acesso em: 15 out. 2011.

SPIDR - SPACE PHYSICS INTERACTIVE DATA RESOURCE. Disponível em <<http://spidr.ngdc.noaa.gov/spidr/>>. Acesso em: outubro de 2011.

Swinbank, R; Shutyaev, V; Lahoz, W.A. Data Assimilation for the Earth System. London: Kluwer Academic Publishers, 2002. 377 p.

PETRY, A.; SOUZA, J. R.; VELHO, H. F. C. Sistema para previsão operacional da dinâmica da ionosfera baseado no modelo SUPIM. São José dos Campos: INPE, 2011. 40 p. (sid.inpe.br/mtc-m19/2011/06.03.18.08-MAN). Disponível em: <<http://urlib.net/8JMKD3MGP7W/39QK26P>>. Acesso em: 15 out. 2011.

uBLAS - Basic Linear Algebra Library. Disponível em <
http://www.boost.org/doc/libs/1_48_0/libs/numeric/ublas/doc/index.htm>. Acesso em
novembro de 2011.