

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISAS EM SENSORIAMENTO REMOTO
E METEOROLOGIA
CURSO DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

DESENVOLVIMENTO DE ALGORITMO DE AJUSTE DE 'WAVEFORMS'
ALTIMÉTRICAS REFLETIDAS POR LAGOS E RIOS

Autor: FLAVIO PROVITINA

Orientador: Dr. Robin Thomas Clarke

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Sensoriamento Remoto, área de concentração em Recursos Hídricos no Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia da Universidade Federal do Rio Grande do Sul.

PORTO ALEGRE, MARÇO DE 1993

Provitina, Flávio

Desenvolvimento de Algoritmo de Ajuste de 'Waveforms' Altimétricas Refletidas por Lagos e Rios./ Flávio Provitina - Porto Alegre: UFRGS, 1993. ix, 84f.

Dissertação (Mestrado), Sensoriamento Remoto, CEP SRM - Centro Estadual de Pesquisa em Sensoriamento Remoto e Meteorologia.

1. Altimetria.
2. Sensoriamento Remoto Ativo.

AOS

Colegas, hoje amigos.

DESENVOLVIMENTO DE ALGORITMO DE AJUSTE DE 'WAVEFORMS'
ALTIMÉTRICAS REFLETIDAS POR LAGOS E RIOS

Autor : Flávio Provitina ¹

Orientador : Prof. Robin Thomas Clarke

RESUMO

O uso de dados altimétricos para a avaliação e monitoramento da variação dos níveis dos oceanos, teve início com o satélite Seasat e vem desenvolvendo-se com êxito em outros sistemas como o Geosat e o ERS-1. Esta técnica tem se mostrado única em termos de oceanografia e abre caminho para sua utilização em outros campos como a hidrologia.

A possibilidade de avaliar-se os níveis das águas continentais por meio da altimetria, trata-se de uma proposta promissora em termos de recursos hídricos em países de dimensões como o Brasil. O desenvolvimento de um algoritmo capaz de, a partir dos dados do satélite altimétrico Geosat, processá-los de forma mais rápida e econômica, fornecendo informações prévias com as quais se possa efetivamente obter a variação do nível das águas de rios e Lagoas, constitui-se no desafio desta dissertação.

O ajuste matemático proposto, que se utilizaria de um número menor de variáveis, se comparado ao método aplicado pela NASA, não trouxe os resultados esperados, mostrando-se ineficiente. Esta dissertação sugere que, embora este modelo proposto tenha se mostrado inadequado, abre-se um vasto campo de pesquisas para que a altimetria de satélites torne-se uma técnica operacional e viável em termos de recursos hídricos em águas continentais.

¹ - Engenheiro Civil (Escola de Engenharia da Universidade Federal do Rio Grande do Sul).

DEVELOPMENT OF AN ALGORITHM TO FIT ALTIMETRIC WAVEFORMS
REFLECTED FROM INLAND WATERS

Author : Flávio Provitina

Adviser : Dr. Robin Thomas Clarke

ABSTRACT

The use of satellite altimeter measurements to assess and monitor ocean level variability began with Seasat satellite and has developed successfully with others satellites such as Geosat and ERS-1. This technic has shown unique to results in oceanography and its application is suggested for other fields such as hydrology.

The possibility of assessing inland water level changes using satellite altimetry techniques is promising in terms of water resources in large countries such as Brazil. The development of an algorithm which is capable of processing the data obtained from the Geosat altimeter faster and less expensively, providing previous information that effectively yields inland water level variability, can be considered the challenge of this dissertation.

The mathematical fitting proposed, handling a smaller number of variables has not provided the results expected compared to the NASA method and proved inefficient. This dissertation suggests that, although the proposed modelling method has behaved inadequately, a new research field has opened to allow the satellite altimetry technique to become an operating, viable technique in terms of inland water resources.

SUMÁRIO

Resumo	iii
Abstract	iv
Lista de Figuras	vii
Lista de Tabelas	ix
I. INTRODUÇÃO.....	1
1.1. OBJETIVOS.....	3
1.2. HISTÓRICO.....	4
II. REVISÃO BIBLIOGRÁFICA.....	8
2.1. O SISTEMA GEOSAT.....	8
2.1.1. Obtenção das 'waveforms'.....	14
2.2. MEDIDAS ALTIMÉTRICAS.....	16
2.2.1. Processamento básico para os dados altimétricos.....	20
2.2.2. Erros nas medidas.....	21
2.3. MÉTODO DA NASA DE AJUSTE DE 'WAVEFORMS'.....	21
2.3.1. Categorização dos retornos altimétricos.....	23
2.3.2. Ajuste dos retornos especulares.....	24
2.3.2.1. Ponto de meia potência.....	24
2.3.2.2. Parâmetros remanescentes que definem a forma.....	25
2.3.3. Ajuste de retornos difusos.....	27
2.4. PRINCÍPIOS ESTATÍSTICOS.....	29
2.4.1. Distribuição de probabilidades.....	29
2.4.1.1. Variáveis aleatórias.....	29
2.4.1.2. Média e variância de uma distribuição de probabili- dades.....	29
2.4.2. Densidade de probabilidades.....	30
2.4.2.1. Variáveis aleatórias contínuas.....	30
2.4.2.2. Média e variância de uma densidade de probabilidade.....	32
2.4.3. Distribuição log-normal.....	32
2.4.4. Função verossimilhança.....	34

III. MATERIAIS E MÉTODOS.....	37
3.1. DADOS.....	37
3.2. MATLAB.....	40
3.3. PROCEDIMENTOS.....	41
3.3.1. Reordenação das Freqüências.....	41
3.3.2. Avaliação da função log-normal truncada.....	44
3.3.2.1. Verossimilhança.....	45
3.3.2.2. Maximização da verossimilhança.....	50
3.3.2.3. Definição da função log-normal truncada.....	55
3.3.2.4. Maximização da curvatura.....	56
IV. RESULTADOS E DISCUSSÕES.....	59
V. CONCLUSÕES.....	74
VI. BIBLIOGRAFIA.....	76
VII. APÊNDICE.....	79

LISTA DE FIGURAS

Fig.1	Satélites Altimétricos	5
Fig.2	Ilustração de satélite altimétrico	9
Fig.3	Representação esquemática da medida altimétrica.	10
Fig.4	Potência recebida por um altímetro em função do tempo. P_{max} é a potência máxima recebida e P_{noise} é a potência de ruído de fundo	13
Fig.5	Exemplo de 'waveform' do Seasat para 11 metros de SWH (média de 1000) normalizado para ter um valor máximo de 1. A linha pontilhada corresponde a 60 'gates' separados por 3,125 ns, com três 'gates' adicionais localizados no centro da 'waveform' ('gates' 29.5; 30.5; 31.5). O 'half power point' está colocado no 'gate'30.5..	16
Fig.6	Ilustração das duas formas básicas de retorno do radar altimétrico em termos espetaculares....	19
Fig.7	Retorno Oceânico Ideal	22
Fig.8	Ilustração das duas formas básicas de retorno altimétrico em termos difusos	28
Fig.9	Gráfico da distribuição log-normal	34
Fig.10	Altímetro do Geosat em sua órbita sobre a Bacia Amazônica entre o Equador e cinco graus sul	38
Fig.11	Imagem Landsat da região de Manaus, com sobreposição de pistas do altímetro Geosat.....	39
Fig.12	Simplexes regulares para 2 e 3 variáveis independentes	52
Fig.13	Seqüência de simplexes regulares obtidos na minimização de $f(x)$. O traçado indica a projeção..	52
Fig.14	Gráfico da 'waveform'V40433.....	62

Fig.15	Gráfico da log da verossimilhança de V40433.....	63
Fig.16	Gráfico da log-normal truncada para a 'waveform' V40433 com a μ e σ maximizadoras da log da verossimilhança	64
Fig.17	Gráfico da log-normal truncada com as estimativas da média e variância	65
Fig.18	'Waveform' V40428 e gráfico	68
Fig.19	Plotagem da log da verossimilhança para a V40428	69
Fig.20	Plotagem da log-normal truncada para V40428 com μ e σ maximizadoras da log de verossimilhança.	70
Fig.21	Plotagem da log-normal truncada com as estimativas da média e variância	71
Fig.22	Gráfico da função log-normal truncada para V40428, com média e variância calculadas alternativamente	72
Fig.23	Gráfico da função log-normal truncada para V40433, com média e variância calculadas alternativamente	73

LISTA DE TABELAS

Tab.1	Diâmetros de áreas observadas pela antena altimétrica em função da altura significativa de onda $H_{1/3}$ para altitudes de 800 a 1335 Km....	11
Tab.2	Características do radar altimétrico Geosat..	14
Tab.3	Exemplos de alguns 'waveforms' sobre a Amazônia e a Lagoa dos Patos	40
Tab.4	Intervalos das frequências estatísticas.....	43
Tab.5	'Waveform' V40433 escrita sob a forma de vetor	59
Tab.6	'Waveform' V40428.....	66

DESENVOLVIMENTO DE ALGORITMO DE AJUSTE DE 'WAVEFORMS' ALTIMÉTRICAS REFLETIDAS POR LAGOS E RIOS

I. INTRODUÇÃO

O Brasil é um país de dimensões consideráveis, o qual depende imensamente de seus recursos hídricos tanto para o transporte fluvial como para geração de energia. A avaliação deste importante recurso natural pode tornar-se difícil devido a grandes distâncias territoriais, acesso limitado, baixa densidade populacional (caso verificado na região Amazônica), e extrema escassez de observadores confiáveis e hábeis para um trabalho em áreas inóspitas e remotas. Além disto, os instrumentos disponíveis nos dias atuais para medidas hidrológicas, capazes de operar sem manutenção por longos períodos e em condições de alta temperatura e umidade, são limitados.

A avaliação das características do fluxo do rio, requer (a) medidas freqüentes do nível d'água e (b) o conhecimento da relação entre nível d'água e descarga do rio (b). O sensoriamento remoto não pode contribuir para (b), mas é possível que contribua para (a) [Koblinsky, Clarke, Brenner, Frey, 1992]. Neste contexto coloca-se a altimetria de satélites, a qual fora desenvolvida para a medição das declividades da superfície oceânica, e tem uma aplicação potencial em águas continentais.

A obtenção de medidas de variação no nível das águas das bacias continentais através da técnica altimétrica, com precisões e freqüências requeridas para aplicação hidrológica se coloca como um desafio. A operacionalidade desta técnica é de um valor econômico considerável para a administração dos recursos hídricos do país.

Alguns trabalhos têm sido publicados versando sobre a aplicação de dados altimétricos de satélite em recursos hídricos, como Koblinsky et al., (1992).

A técnica altimétrica para bacias hidrográficas ainda é

incipiente, de modo que os materiais disponíveis restringem-se a praticamente trabalhos recentemente publicados por pesquisadores, o que não constitui ainda, uma bibliografia especializada.

Para que os dados gerados via altimetria sejam aplicáveis à hidrologia, necessitam satisfazer três condições básicas :

(a) precisão nas estimativas de variação de nível d'água.

Os maiores fatores de erro em curvas que relacionam a descarga com o nível d'água estão nas incertezas na medição da descarga. Portanto a medição do nível d'água deve ser bastante correta para minimizar os erros decorrentes destas estimativas.

(b) freqüência das estimativas.

Para o monitoramento dos recursos hídricos espera-se medições diárias do nível das águas. No entanto para grandes bacias (que respondem lentamente a precipitação) os dados de medições menos freqüentes poderiam ser interpolados. Estimativas diárias num mesmo local não são possíveis com um único altímetro. O Geosat tem um período de repetição de 17 dias.

(c) o tempo entre a obtenção dos dados e sua disponibilidade para aplicações hidrológicas.

Para algumas aplicações tal como, operação de hidroelétricas, os dados hidrológicos devem estar disponíveis em tempo quase real. Dados de altímetros não podem ser processados nestas condições até o momento. Porém para levantamentos dos recursos hídricos em regiões remotas, esta urgência na disponibilidade é desnecessária. No momento, mesmo em países desenvolvidos, atrasos de alguns meses entre a tomada dos dados de campo e sua disposição em arquivos de qualidade controlada, é comum. Isto deveria ser aceitável em se tratando da avaliação dos recursos hídricos [Koblinsky et al., op. cit.].

Em algumas aplicações hidrológicas, o simples conhecimento do nível das águas é suficiente (previsão de enchentes em áreas urbanas). Para projeto e operação de reservatórios hidroelétricos o conhecimento da descarga é essencial [Koblinsky et al., op. cit.].

1.1 OBJETIVOS

A altimetria satelitária tem sido utilizada com sucesso no campo oceanográfico, fornecendo medidas com precisão satisfatória da variação do nível d'água dos oceanos, o que qualifica os altímetros como um instrumento revolucionário para oceanografia.

A possibilidade de aplicar os dados obtidos por estes satélites às grandes bacias hidrográficas como por exemplo, a bacia do Amazonas, a Lagoa dos Patos entre outras, a manipulação dos dados altimétricos, seu tratamento e análise, pode se tornar alvo de muitas pesquisas. Esta dissertação objetiva contribuir neste campo de pesquisa, mais especificamente no tratamento dos dados obtidos do Geosat.

Objetiva-se obter as mesmas quantidades calculadas pelo método desenvolvido pela NASA utilizando-se dois parâmetros ao invés de até nove parâmetros. Após, estabelecer-se-á uma comparação entre os resultados de ambos os métodos. Isto é, ajustar os retornos altimétricos que constituem-se num conjunto de sessenta e três frequências estatísticas obtidas a partir do satélite altimétrico, a funções que o descrevam matematicamente com menos parâmetros, e que poderão ser ajustadas mais rápida e facilmente. Estará apresentada nesta tese, a tentativa de obter-se uma parametrização alternativa, utilizando uma função log normal truncada. Com esta função, uma vez conhecidos seus dois parâmetros, poder-se-ia obter o tempo de viagem do pulso do radar ao nível d'água com economia computacional e de tempo.

Os retornos altimétricos, referidos como 'waveforms', que são a informação obtida do Geosat, requerem um processamento especial, a partir do qual pode-se calcular as medidas de altura do nível das águas. Existem duas categorias de retornos altimétricos. Um deles é referido como retorno especular, o qual é comumente encontrado sobre gelo marítimo, superfícies planas e desérticas [Satellite Radar Altimetry over Ice, NASA] e pequenos corpos d'água que são altamente refletivos como lagos e rios. Os sinais de retorno

altimétrico recebidos de rios da dimensão do Amazonas, têm características especulares [Koblinsky et al.,1992]. A segunda categoria de retornos, composta pelos remanescentes, chama-se de retorno difuso e parece-se com os retornos oceânicos. Encontram-se sobre gelo continental, calotas polares e superfícies rugosas continentais.

Quando o retorno altimétrico em questão for um retorno especular, ter-se-á vários parâmetros que definem e caracterizam a 'waveform', dependendo se esta for com um ou dois picos. Já o método usado para ajustar um retorno difuso é modelar o retorno com uma função cuja a posição do ponto médio da 'leading edge' seja um parâmetro. Usa-se para isto o método dos mínimos quadrados Baysiano.

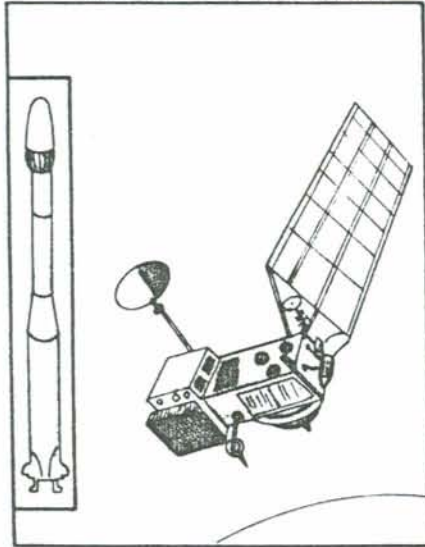
Apresentar-se-á também, a descrição do método de ajuste dos retornos altimétricos especulares e difusos desenvolvido e utilizado pela NASA, para que se estabeleça uma efetiva comparação entre ambos os métodos de parametrizações.

1.2 HISTÓRICO

É cabível de colocar-se entre os tópicos desta tese um breve histórico dos desenvolvimentos da altimetria e dos satélites que vêm permitindo tais avanços.

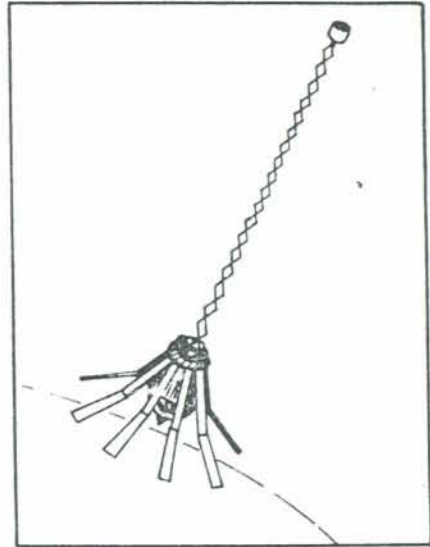
Esta técnica iniciou-se nos anos setenta, quando, com o intuito de avaliar as elevações da superfície média dos mares, a NASA lançou ao espaço três altímetros. O SKYLAB-S193 in 1973 [MacArthur et al.,1987] seguido pelo GEOS-3 em abril de 1975 até dezembro de 1978 e o SEASAT de Julho a outubro de 1978 . Estas missões tinham fins geodésicos e oceanográficos, cujas medições atingiam uma precisão de até 10 centímetros. Presentemente tem-se o altímetro Geosat [MacArthur et. al.,1987] que tem fornecido dados altimétricos úteis desde março de 1985, e o ERS-1 lançado recentemente pela agência europeia espacial. Nos próximos anos a NASA, juntamente com o CNES, deverão lançar os altímetros NASA

TOPEX / POSEIDON



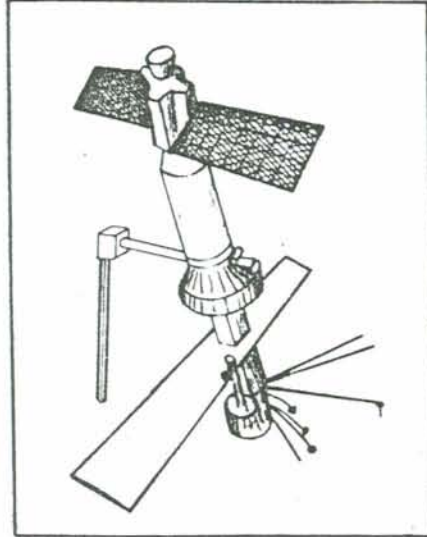
1992

GEOSAT



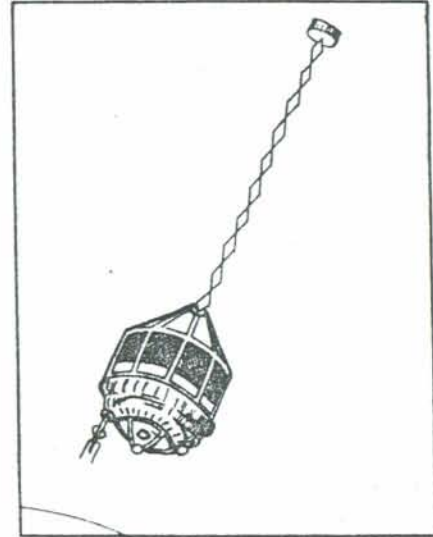
1985-present

SEASAT



1978

GEOS-3



1975 - 1978

Fig. 1 - Satélites Altimétricos

TOPEX e CNES POSEIDON [Chelton et al., 1988]. Veja fig. 1.

O SEASAT cuja a missão foi rapidamente interrompida devido a falhas no sistema de energia, representou um bom avanço em relação aos anteriores. Seu altímetro colheu cerca de 360 horas sobre os continentes de 7 de julho a 7 de setembro, quando operou em uma órbita incluída de 108 graus em relação ao Equador, com um ciclo de repetição inicial de 17 dias que passou para 14 e 8, respectivamente [Koblinsky et al., op. cit.].

O altímetro do Seasat permitiu medir a distância entre o satélite e a superfície do mar, com uma precisão de até dez centímetros a uma altitude de oitocentos quilômetros. Possibilitou aos oceanógrafos medir a declividade absoluta da superfície do mar, tinha a capacidade de medir com exatidão a altura significativa de ondas e estimar a velocidade do vento próximo a superfície do mar.

Cheney et al, estabeleceram medições comparativas obtidas a partir de dados do Geosat e de dados de medição das marés, observando que embora houvesse casos com diferenças entre cinco a dez centímetros, as duas medições geralmente concordam com diferenças de menos centímetros.

O uso de dados altimétricos do Seasat para monitorar águas continentais foi relatado por Rapley et al., 1987. Este foi o primeiro passo rumo a aplicação dos dados altimétricos para avaliação do nível das águas continentais. Concluiu-se então, que seria possível uma precisão final de mais ou menos vinte centímetros, se existisse cerca de dez estimativas de alcance ao longo da trajetória do satélite.

Brenner e Frey, utilizando dados de altimetria do SEASAT (1990), produziram uma topografia continental de alta qualidade para a Austrália e outros países. Isto resultou de um estudo profundo das características dos sinais de altimetria de superfícies não oceânicas, e levou a conclusão que dados derivados dos altímetros Seasat e Geosat, poderiam ser utilizados em regiões

onde os dados de campo são escassos e de má qualidade.

Os estudos persistiram também em relação a área da bacia amazônica, do que concluiu-se que a geração corrente de altímetros de satélite poderia ser usada para monitorar as grandes mudanças sazonais de nível do Amazonas (em torno de dez metros), com uma precisão de cerca de vinte centímetros. Outra aplicação importante da altimetria seria fornecer medidas do perfil de elevação do rio, e sua mudança sazonal.

A missão Geosat da Marinha dos Estados Unidos, lançou em 1985 um altímetro muito semelhante ao do Seasat com algumas melhoras [MacArthur et al., 1989]. Os dados utilizados para a execução desta dissertação são provenientes deste satélite, o qual será abordado com mais detalhes.

Koblinsky et al. (1992), mostraram através de dados do Geosat do período de 1987 a 1988 e medições de nível do rio Negro em Manaus, que é possível um bom ajuste entre estimativas de altimetria e medições de campo. Contudo, antes de se afirmar ser ou não a altimetria de satélite, uma ferramenta para avaliação do nível das águas continentais, resta muito trabalho e pesquisa a ser desenvolvido.

II. REVISÃO BIBLIOGRÁFICA

2.1. O SISTEMA GEOSAT

O radar altimétrico a bordo do Geosat é uma versão aprimorada do altímetro do Seasat. Para um entendimento do processo altimétrico precisa-se conhecer as características do sinal de retorno [MacArthur et al., 1989], o que apresenta-se a seguir.

Quando lançado em 1985, pela marinha norte americana, o Geosat tinha como primeiro objetivo o mapeamento da geóide marinha ao longo de uma densa grade de 'ground tracks', não repetitivos, para propósitos de defesa. Esta operação durou dezoito meses e chamou-se GM (Geodetic Mission). Nesta fase os 'ground tracks' tinham um afastamento médio de quatro quilômetros. Iniciou-se em outubro de 1986, a chamada ERM (Exact Repeat Mission), cujas órbitas eram repetidas a cada dezessete dias e espaçadas de cento e cinquenta quilômetros. Estes 'ground tracks' são quase coincidentes com os respectivos do satélite civil da NASA, o SEASAT. As medidas de elevação são obtidas a intervalos de seiscentos e sessenta e dois metros ao longo da 'ground track'. Veja fig. 2.

A técnica do radar altimétrico, conhecida como compressão de pulso, consiste na emissão de um pulso relativamente longo pelo altímetro. Este processará o sinal de retorno de forma tal, que é equivalente a transmitir um pulso muito curto e medir a série temporal da potência de retorno numa seqüência de 'janelas' [Chelton, Walsh, MacArthur, 1988].

O radar altimétrico mede a distância da antena do radar até a superfície do oceano ou de uma bacia hidrográfica. No ponto de vista oceanográfico, a aplicação primeira do altímetro é medir a

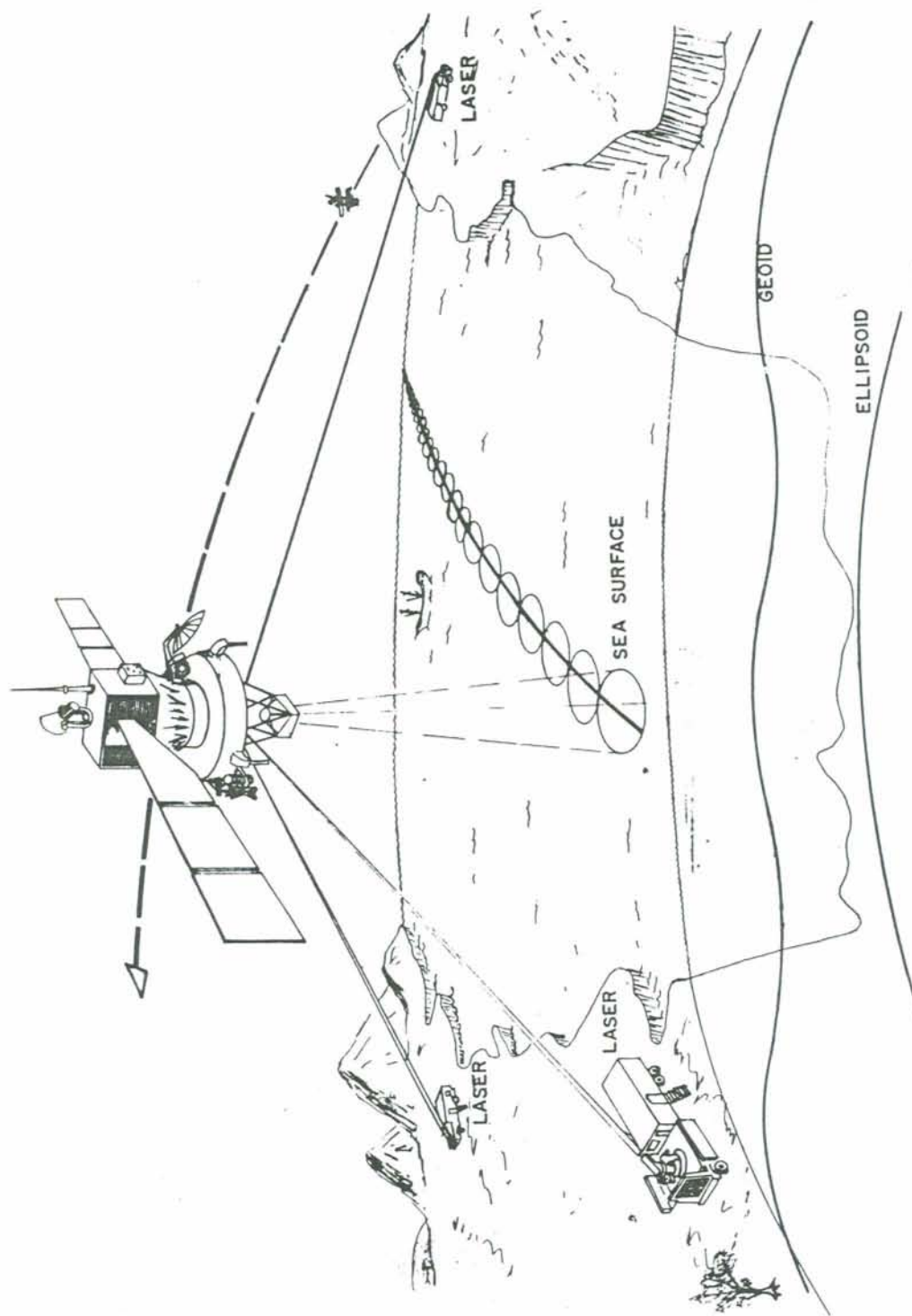
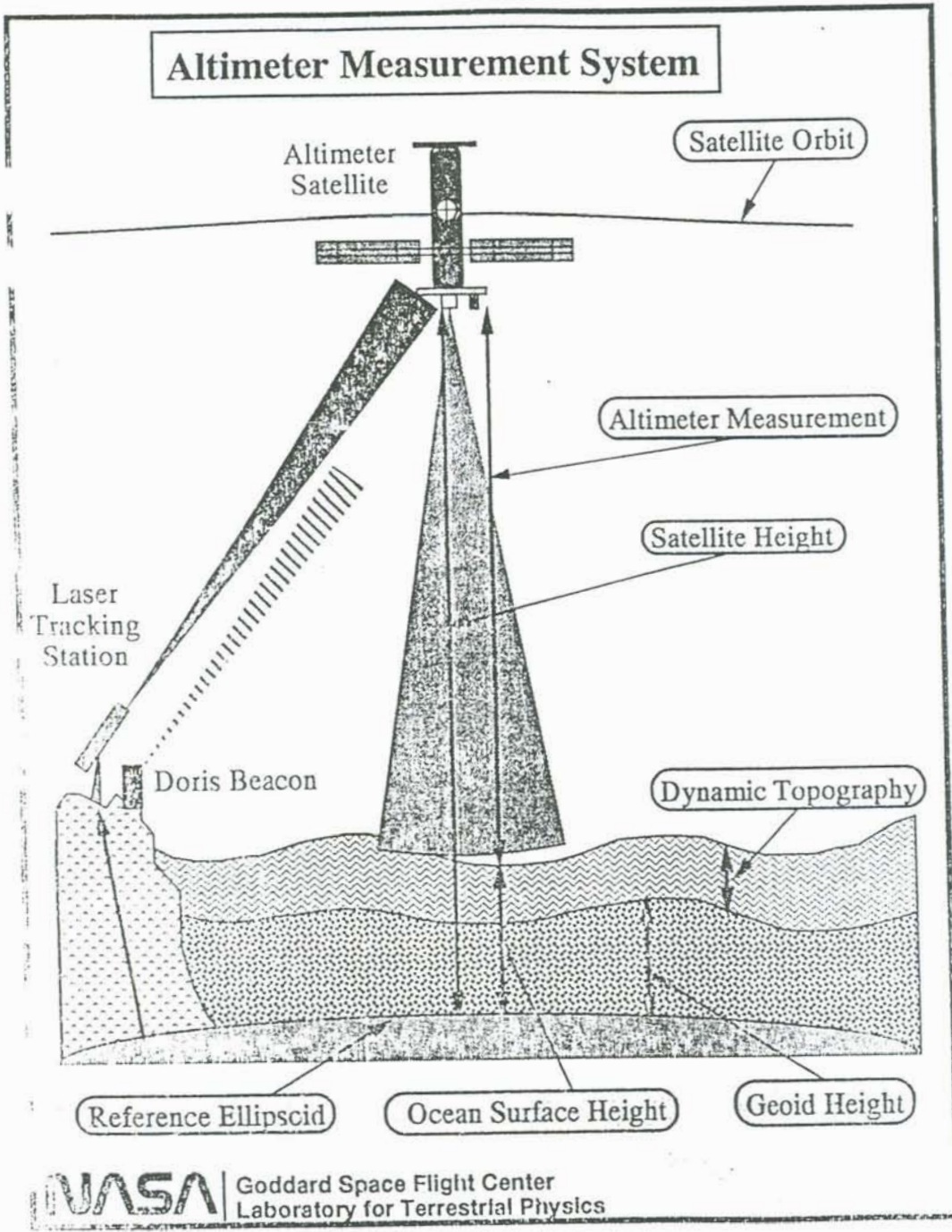


Fig.2 - Ilustração de Satélite Altimétrico (NASA)



NASA | Goddard Space Flight Center
 Laboratory for Terrestrial Physics

Fig.3 - Representação da medida altimétrica

elevação da superfície oceânica, a qual requer uma precisão de poucos centímetros. O altímetro transmite um pulso curto de radiação de micro-onda, cuja a potência é conhecida, em direção a superfície, no nadir do satélite (ponto diretamente abaixo deste). Este pulso interage com a superfície e parte da radiação incidente retorna ao satélite. A distância do satélite até a superfície é determinada a partir do tempo de viagem de ida e volta do pulso. Veja figura 3.

A área observada (footprint) por um radar altimétrico de feixe estreito é a região sobre a superfície iluminada pelo ângulo de feixe da antena. Espera-se num satélite altimétrico, que o 'footprint' seja da ordem de poucos Km de diâmetro para melhor eficiência. [Chelton et al., op. cit.].

Tabela 1.

Diâmetros de áreas observadas pela antena altimétrica em função da altura significativa de onda $H_{1/3}$ para altitudes de 800 e 1335 quilômetros.

$H_{1/3}$ (m)	Diâmetro da área observada (800 Km altitude) Km	Diâmetro da área observada (1335 Km altitude) Km
0	1.6	2.0
1	2.9	3.6
3	4.4	5.5
5	5.6	6.9
10	7.7	9.6
15	9.4	11.7
20	10.8	13.4

Os altímetros Seasat, Topex e Geosat, usam uma duração efetiva de pulso (τ) de 3,125 ns. Para a altitude de 800 Km do Geosat, resulta um 'footprint' de diâmetro máximo de 1,6 Km em superfície de águas calmas. Este diâmetro pode no entanto, incrementar-se até

10,8 Km em mar aberto, cujo SWH seja 20 metros. O SWH (significant wave height $H_{1/3}$) corresponde aproximadamente a distância da crista ao vale da onda de um terço das maiores ondas localizadas dentro do 'footprint'.

A técnica altimétrica é muito complexa, envolve mais de doze correções de efeitos tais como, correções atmosféricas, tendenciosidades instrumentais, e efeitos geofísicos externos (gravidade, maré e pressão atmosférica sobre o oceano). Os componentes de erro na medição da altura menos entendidos, são os efeitos instrumentais. Estes erros estão relacionados à técnica, na qual as distâncias são estimadas a bordo do satélite a partir do sinal de retorno recebido pelo altímetro.

As medidas altimétricas são descritas em termos de altimetria de pulso limitado. A potência do pulso refletido da superfície é medido como uma função de tempo em intervalos de nanosegundos. Em termos oceânicos, o retorno do altímetro, como uma média espacialmente distribuída no nadir do satélite, refere-se como o nível médio do mar. Este nível só pode ser associado a um ponto particular na história temporal das potências de retorno (que é conhecida como 'return waveform'). Este ponto é conhecido como o 'half-power point'. O tempo de viagem de ida e volta do pulso do satélite até o nível médio do mar é determinado pelo rastreamento deste ponto na 'waveform' de retorno.

Observando-se a figura 4, nota-se que a potência medida no sinal de retorno por um altímetro de pulso limitado, comporta-se como uma função de tempo.

' t_c ' é definido como o tempo de viagem ida e volta para o pulso refletir da crista da onda.

' $t_c + \tau + 2 * c^{-1} * H_{1/3}$ ' é o tempo de viagem ida e volta para o pulso refletir do vale das ondas.

A série temporal de potências vista na figura anterior, é conhecida como a forma de onda de retorno ('return waveform'). O nível médio do mar corresponde à metade da distância entre a crista e o vale das ondas. Assim, o tempo de viagem ida e volta t_0 , para o

centro do pulso (meio caminho entre o 'leading edge' e 'trailing edge') refletir do nível médio do mar é dado por ' $t_c + \tau/2 + \bar{c}^{-1} * H_1/3$ '. Este ponto corresponde na 'waveform' a um ponto cuja a potência é a metade da diferença entre Potência máxima (P_{max}) e a Potência de ruído de fundo (P_{noise}).

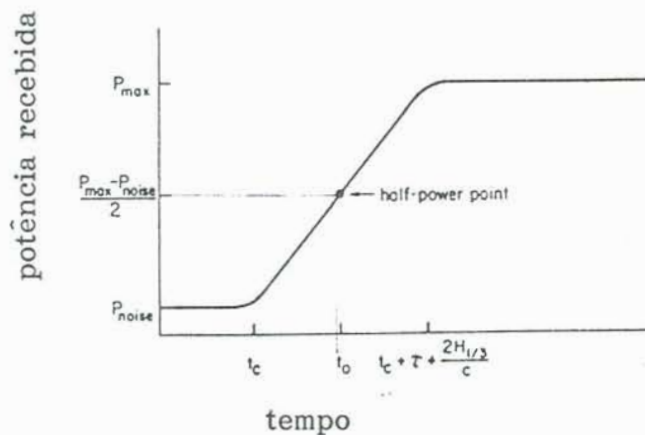


Fig. 4 - Potência recebida pelo altímetro em função do tempo.

De modo que, o tempo de viagem de ida e volta e então, a distância até o nível médio do mar, pode ser determinado pelo rastreamento do 'half power point' na 'waveform'. Para aplicações em águas continentais, nas quais não existem ondas, o tempo de viagem do pulso até a onda é definido localizando-se o ponto do início da curvatura (leading edge) da 'waveform'. A abscissa define o tempo t_c .

A tabela 2 mostra características do radar altimétrico Geosat.

Tabela 2 : Características do Radar Altimétrico.

Waveform:

tipo:pulso linear FM

frequencia central:19.5 gigahertz

largura do pulso:102.4 microsegundos

largura da banda:920 megahertz

periodo do interpulso:980 microsegundos

geracao: sintese digital

Transmissor:

Potencia pico: 20 watta

Potencia de consumo: 70 watta

Processador do sinal:

conversor A/D: 4 bits + sinal

filtro digital: 63

Produtos telemétricos:

Digital(serie)

tamanho do frame:850 bits

razao do frame:10 por segundo

conteudo:63 amostras (range gates)

Antena:

Tipo: 1 metro parabolica

ganho: >37.6 decibeis

largura da feixe: 2 graus

Receptor:

Tipo : conversao dual

Controle de ganho automatico: 0 a 63 decibeis.

Entradas:

frequencia de referencia: 5 megahertz

Peso:

processador do sinal: 47 libras

secao RF: 144 libras

Potencia requerida:

standby: 103 watta

track: 146 watta

2.1.1. OBTENÇÃO DAS WAVEFORMS

A frequência do sinal refletido da crista da onda é menor, e a frequência do sinal refletida do vale da onda é maior, que a refletida do nível médio do mar. A frequência e o tempo de viagem são linearmente relacionados. A técnica de compressão está baseada na transmissão de 'chirps' consistindo de um sinal com uma frequência linear sobre um período de varredura. A forma da distribuição de frequência das potências do sinal de retorno é muito semelhante a distribuição das alturas da superfície do mar [Chelton et al., 1988].

O sinal total retornado do satélite é a soma dos sinais refletidos a partir das ondas orientadas favoravelmente a partir do 'footprint' do altímetro. A forma do espectro de potência em função das frequências é aproximadamente Gaussiana. O pulso do radar é de duração limitada e varia em frequência [Koblinsky et al., 1992].

Uma 'waveform' de comportamento mais regular existe somente a partir da média de muitas 'waveforms' individuais. Para se obter uma 'waveform' relativamente regular em comportamento, deve-se fazer a média entre muitas 'waveforms' individuais. O altímetro do Geosat balanceia dois conjuntos de cinquenta pulsos de retorno cada décimo de segundo para que se produza uma média estimada de potência de retorno nas 63 'janelas' ou 'gates'. Esta distribuição média de potências nos pulsos de retorno distribuídos nos 63 gates define então a 'waveform' de retorno.

A 'waveform' é melhor amostrada digitalmente no domínio de frequências. As frequências discretamente amostradas são mencionadas como 'range gates'.

O altímetro do Seasat e Geosat amostram a 'waveform' em sessenta e três frequências contíguas separadas por 312,5 KHz e 9,766 KHz respectivamente. Estes 63 'gates' são centrados no ponto chamado 'track point' ou ponto rastreador, que corresponde ao retorno do nível médio do mar. Três frequências são amostradas

próximo ao ponto central, nas janelas 29,5 , 30,5 , e 31,5. As frequências estão separadas por 3,125 ns, correspondendo a 46 centímetros em comprimento.

Fig.5 : exemplo de 'waveform' do Seasat para onze metros de SWH (média de mil) normalizado para ter um valor máximo de 1. A linha pontilhada corresponde a sessenta 'gates' separados por 3,125 ns, com três 'gates' adicionais localizados no centro da 'waveform' ('gates' 29.5; 30.5; 31.5).

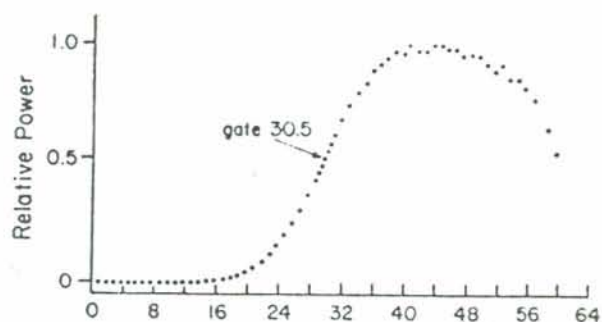


Fig. 5 - Exemplo de waveform de Seasat

As amostras de 'waveforms' são obtidas a cada 662 metros ao longo da trilha do satélite correspondendo a intervalos de 0.1 segundos.

2.2. MEDIDAS ALTIMÉTRICAS

A equação para a medição altimétrica, que define a altura instantânea da superfície d'água é dada por :

$$h' = h_{\text{sat}} - h_{\text{range}}$$

A variável h' é a medida procurada do nível d'água relativo ao centro de massa da Terra. Esta altura é derivada a partir das medidas altimétricas, diminuindo-se h_{sat} , distância entre o centro de massa do satélite e o centro de massa da Terra, de h_{range} , distância do centro de massa do satélite até a superfície terrestre refletora das ondas do radar.

Para estimar h_{sat} , aplicam-se procedimentos precisos de determinação de órbita. Isto é feito pelo ajustamento de um modelo dinâmico da órbita do satélite com medidas diretas da posição do satélite a partir dos sistemas terrestres de rastreamento [Koblinski et al., op. cit.].

A altura do altímetro (h_{range}) é derivada da medida do tempo de viagem ida e volta do pulso do radar. Logicamente esta medição estará sujeita a um conjunto de correções, de modo que :

$$h_{\text{range}} = h_{\text{alt}} + \delta h_{\text{alt}} - \Delta h_{\text{alt}} - h_e - h_b$$

onde : h_{alt} é a metade da distância percorrida pelo pulso do radar (corresponde a metade do tempo de viagem de ida e volta).

δh_{alt} é a correção da distância por efeitos de superfície.

Δh_{alt} são os efeitos de refração atmosférica.

h_e é o ruído do instrumento.

h_b é a tendenciosidade do instrumento.

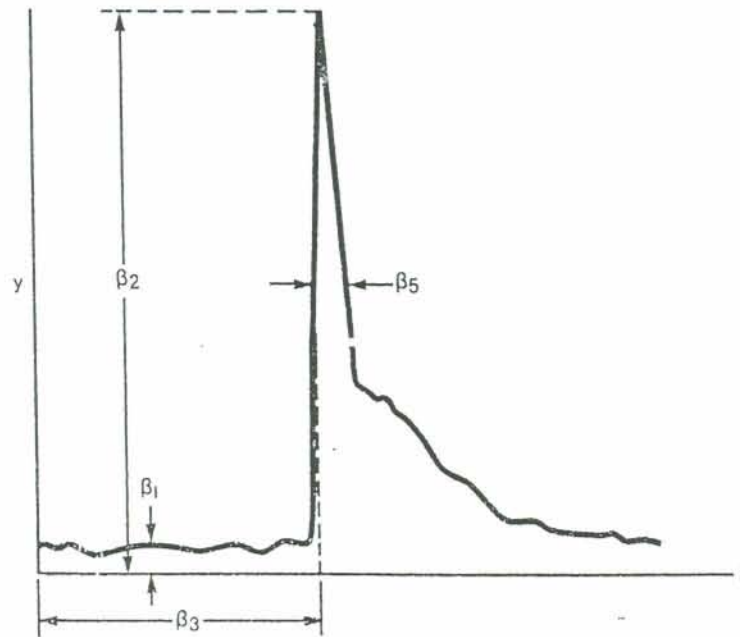
Uma estimativa do tempo de viagem de ida e volta médio h_{alt} é computado a bordo do satélite e telemetrado a Terra.

Como mencionado anteriormente, h_{alt} é baseado no incremento da curva de potência do sinal de retorno sendo centrado no 'gate' intermediário. Sobre superfícies como gelo, terra, e rios, os 'waveforms' de retorno podem ter um padrão difuso complicado ou especular. Para estas superfícies, h_{alt} está errado pois o ponto

médio da 'waveform' não está centrada com o 'gate' do tempo médio (tracking gate). Logo a 'waveform' deve ser individualmente pós-processada para encontrar-se a discrepância δ_{halt} , entre a estimativa de bordo do retorno do radar e o valor real. Este procedimento chama-se 'rastreamento das waveforms' e será abordado com mais detalhes. A figura 6 mostra que o δ_{halt} será obtido do β_3 , isto é, o ponto médio do incremento da curva de potência do sinal de retorno.

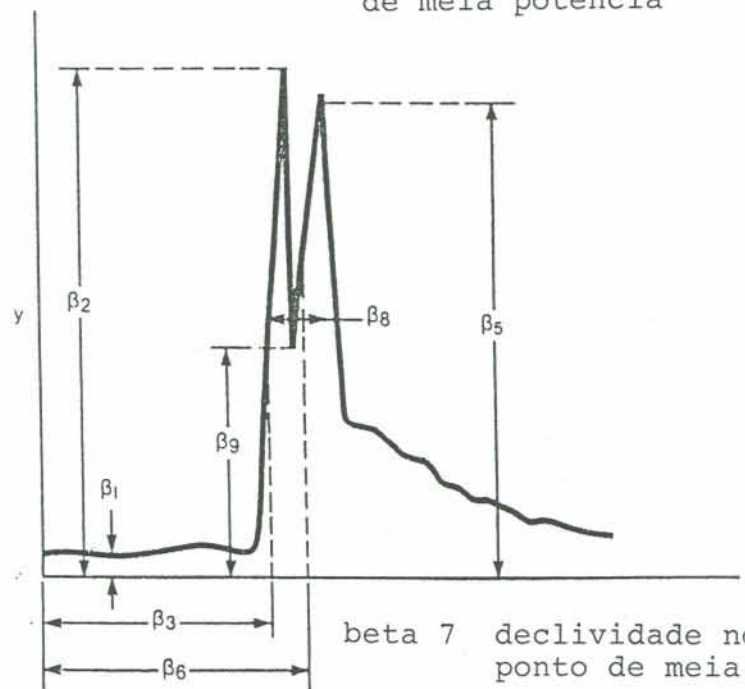
A altura $halt + \delta_{halt}$, estima a distância entre o centro de massa do satélite e a superfície refletora da Terra. Persiste a incerteza quanto ao ruído e a tendenciosidade dos instrumentos eletrônicos do radar. A avaliação destas variáveis está além do escopo deste trabalho.

1 pico



beta 4 declividade do ponto de meia potência

2 picos



beta 7 declividade no segundo ponto de meia potência

Fig.6 - Retorno Altimétrico em Termos Espetaculares

2.2.1. PROCESSAMENTO BÁSICO PARA OS DADOS ALTIMÉTRICOS

As medidas de altura do radar estão sujeitas a várias correções de percurso do pulso, Δh_{at} , devido a refração atmosférica. Esta refração atmosférica é causada devido a presença de partículas carregadas na ionosfera, constituintes neutros na atmosfera, principalmente nitrogênio e oxigênio, e vapor da água na troposfera [Chelton, 1988]. A fim de avaliar-se o nível das águas, as estimativas das correções necessárias são obtidas a partir do NOAA Geophysical Data Record (GDRs). A refração devido ao vapor d'água é corrigida a partir de estimativas de umidade e temperatura da superfície, pois as medidas do conteúdo de vapor d'água colunar não podem ser feita sobre a superfície com micro-ondas passivas [Koblinsky et al., op. cit.].

A medida de altura h' é ajustada para ser relativa a geóide, h_n , e removem-se as variações temporais na geóide causadas pelas marés, h_t . O nível resultante h , ou seja, a estimativa do nível de água é portanto :

$$h = h' - h_n - h_t - h_e$$

O h_n é obtido pela geóide da Universidade do estado de Ohio, a qual utiliza medidas de gravidade de superfície em adição a um campo de gravidade derivado do satélite. Este campo é relativo a uma elipsóide de referência h_e da International Union of Geodesy. O h_t obtém-se a partir de um modelo da NOAA GDR [Koblinsky et al., 1992].

Na missão Geosat Exact Repeat Mission, o satélite repetiu sua trilha a cada 17,05 dias. Consequentemente a geóide e a elipsóide de referência são removidas da computação da mudança de nível do rio em qualquer localização específica, através da comparação da medidas com um nível de referência fixado.

Estas correções têm sido aplicadas nos estudos de medição de nível nas bacias hidrográficas brasileiras, após a parametrização das waveforms segundo o método da NASA. Nesta tese, estes

algoritmos são apresentados a fim de que haja um maior entendimento de todo o processo altimétrico.

2.2.2 ERROS NAS MEDIDAS

O sistema de medição Geosat, tem uma variedade de erros, a maioria deles impossíveis de serem corretamente quantificados, por uma série de razões. Exemplificando, as incertezas nas correções da altura Δh_{alt} , não podem ser quantificadas em qualquer local específico sem a calibração direta pois derivam de modelos atmosféricos de previsão [Koblinsky et al., 1992]. Estes erros são minimizados com o uso de filtros de bandas. Os maiores erros encontram-se no procedimento de determinação de órbita, devido a baixa qualidade das efemérides orbitais fornecidas pela marinha. Estes erros podem ser removidos por meio de um filtro seletivo. Inexiste a bordo do Geosat um radiômetro de micro-ondas, o que incapacita correções precisas para os efeitos refrativos do vapor d'água. Observa-se também a perda de dados do geosat, especialmente em órbitas ascendentes devido ao uso da estabilização do gradiente de gravidade passivo para controlar a postura do 'spacecraft' [Chelton et al., 1990].

2.3. MÉTODO DA NASA DE AJUSTE DA 'WAVEFORM'

Cada retorno altimétrico (waveform), consiste do conjunto de sessenta e três 'gates', tendo cada 'gate' um nível de retorno associado, cujo o valor é medido em contadores (counts). Na figura 7, pode-se observar um retorno oceânico típico procedente do Seasat. O nível de retorno nos primeiros vinte e dois 'gates' está a nível de ruído ou pré-pulso, de quatro a cinco 'counts'. O nível rapidamente cresce a um máximo relativo e então lentamente decresce. Existem três gates no centro que tem um espaçamento de vinte três centímetros ao invés de quarenta e seis. O 'tracking

'gate' é o centro destes 'gates'. O rastreador a bordo do satélite tenta manter o centro do 'leading edge' do retorno posicionado no 'tracking gate' através da previsão do tempo de viagem de cada pulso baseado no retorno prévio.

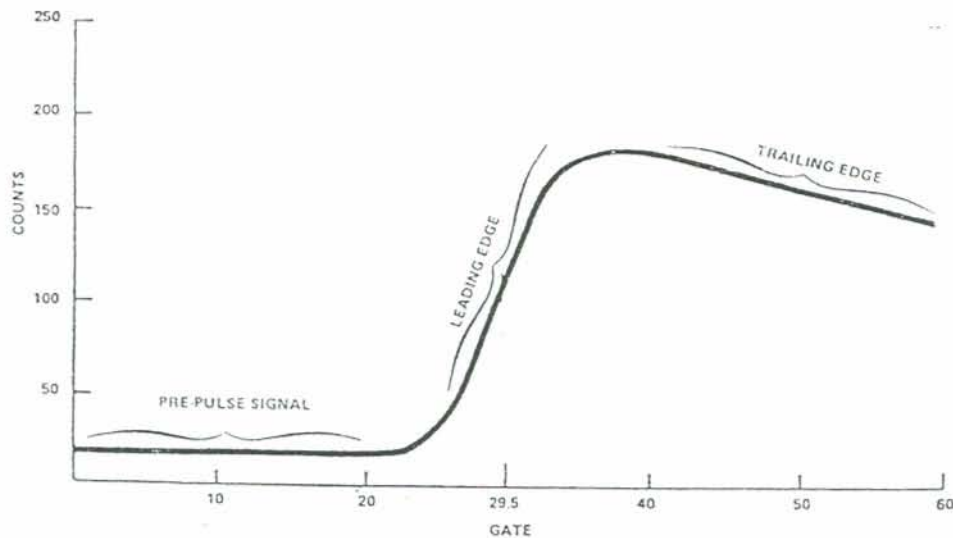


Fig. 7 - Retorno Oceânico Ideal

Os retornos altimétricos sobre superfícies não oceânicas variam muito do exemplo apresentado. Por vezes os retornos não são centrados no 'tracking gate'. A medida telemetrada do rastreador a bordo precisa ser corrigida devido à variação do ponto médio do 'leading edge' do 'tracking gate'. Esta correção de rastreamento é calculada como :

$$\Delta H_{ret} = (G_m - G_t) * g_2m, \text{ onde}$$

G_m = 'gate' do ponto médio do 'leading edge';

G_t = 'tracking gate';

g_2m = conversão de 'gates' para metros = 0.4684375 m/gate;

Segue que:

$$H_{ret}_t = H_{meas}_t + \Delta H_{ret}_{t+1} \text{ onde,}$$

H_{ret_t} = medida altimétrica no tempo t;
 H_{meas_t} = medida calculada pelo rastreador a bordo no tempo t;
 $\Delta H_{ret_{t+1}}$ = correção do rastreamento calculada a partir da
 'waveform' no tempo t+0.98 segundos.

Devido ao fato do retorno ser telemetrado um tempo depois, a correção de rastreamento é calculada para o tempo t+0.098 segundos.

Esta correção de rastreamento, será incorporada no cálculo da variação do nível de água. Apresenta-se a seguir o método de ajuste das 'waveforms' desenvolvido pela NASA, que se utiliza de cinco e nove parâmetros respectivamente.

2.3.1. CATEGORIZAÇÃO DOS RETORNOS ALTIMÉTRICOS

Os retornos são automaticamente categorizados dependendo da existência ou não de picos significantes. Isto é determinado através do seguinte algoritmo.

Y_n = nível de ruído ; calcula-se como o número médio de 'counts' nos primeiros cinco 'gates'.

Y_{max} = é o número máximo de 'counts' em qualquer 'gate'.

$$Y_{med} = \frac{(Y_{max} - Y_n)}{2.0} + Y_n$$

G_{mid} , é o primeiro 'gate' onde o número de 'counts' excede Y_{med} . Formam-se duas somas de 'counts' consecutivos a partir do sinal, Y_{low} e Y_{high} , onde :

$$Y_{low} = \sum_{i=G_{mid}}^{i=G_{mid}+9} Y_i$$

$$Y_{high} = \sum_{i=G_{mid}+10}^{i=G_{mid}+20} Y_i$$

Quando a razão Y_{high}/Y_{low} é menor ou igual a 0.7, o retorno é considerado especular, do contrário ele será difuso.

2.3.2 AJUSTE DE RETORNOS ESPECULARES

Os retornos do tipo especulares são definidos neste procedimento como sendo caracterizados por um ou mais picos extremamente agudos e são rastreados tentando-se localizar o ponto médio ou ponto de meia potência do primeiro pico significativo. Na figura 6 pode-se ver os cinco parâmetros requisitados na definição de um retorno de um só pico. Nesta mesma figura observam-se os nove parâmetros que definem um retorno com dois picos.

2.3.2.1. Ponto de meia potência

Para determinar-se o ponto médio do primeiro pico significativo, deve-se definir primeiro a localização deste pico. Utilizando-se variáveis anteriormente definidas, procura-se um 'gate' cujos 'counts' excedam ou igualem 25 por cento da diferença entre Y_{max} e Y_n . Este 'gate' chama-se G_{rise} .

Se as seguintes condições forem satisfeitas :

$Y_{G_{rise}+1} - Y_{G_{rise}} < 0$ para $Y_{G_{rise}} > Y_{max} \times .3$, onde

$Y_{G_{rise}}$ são os 'counts' para o 'gate' G_{rise} , e

$Y_{G_{rise}+1}$ são os 'counts' para o 'gate' $G_{rise}+1$,

G_{rise} é determinado como sendo o primeiro pico significativo da 'waveform'.

As 'waveforms' menores e mais arredondadas, exigem que se encontrem as seguintes condições :

$Y_{G_{rise}+1} - Y_{G_{rise}} < (Y_{max} - Y_n) \times .05$, para

$Y_{G_{rise}} \leq Y_{max} \times .3$.

O G_{rise} é incrementado um a um, até um máximo número de 'gates', até que uma das condições acima seja satisfeita, após a qual o 'gate' do primeiro pico significativo G_{1st} , e seu correspondente valor, Y_{1st} , serão usados para determinar o ponto de meia potência do pico (half power point). O valor do 'count' no ponto de meia potência, Y_{mid} , é determinado como segue :

$$Y_{mid1} = \frac{(Y_{1st} - Y_n)}{2.0} + Y_n$$

A localização exata do 'gate' do ponto de meia potência, G_{tmid1} , é determinada por meio de uma interpolação linear para o valor do contador Y_{mid1} localizado entre os 'gates' X_1 e X_2 , com os valores respectivos Y_1 e Y_2 .

2.3.2.2. Parâmetros remanescentes que definem a forma.

Para que se possa definir a forma exata do retorno especular, se faz necessário calcular vários outros parâmetros além do nível de ruído, máximos 'counts' do primeiro pico significativo, e a localização do ponto de meia potência. Para os retornos de um só pico e de dois picos, variáveis que definem a largura do pico significativo e a declividade no ponto de meia potência também devem ser avaliados. Um retorno com dois picos têm variáveis adicionais como, valor máximo de 'counts' para o segundo pico significativo, localização do ponto de meia potência para o segundo pico e sua declividade e o mínimo valor de 'counts' entre os dois picos significantes.

As declividades no 'half power point' para ambos os picos Slp_{1st} e Slp_{2nd} , são determinadas pelo seguinte algoritmo :

$$Slp_{1st} = \frac{Y_2 - Y_1}{X_2 - X_1}$$

O Slp_{2nd} utiliza-se da localização dos 'gates' e seus correspondentes valores determinados para avaliar o 'half-power point' do segundo pico significativo. A existência de fato do segundo pico é definida iniciando-se pela localização do 'gate' do primeiro pico, e monitorando-se a diferença entre 'counts' dos 'gates' consecutivos. Quando a mudança em 'gates' sucessivos fica negativa, no local do 'gate' Cnt_{min} , assume-se que outro pico foi

encontrado. Neste ponto uma soma totaliza todos os 'counts' em todos 'gates' que seguem o Cnt_{min} . Quando esta soma iguala ou excede 9 por cento do Y_{1st} então encontrou-se o segundo pico significativo. O 'gate' no qual o segundo pico ocorre, X_{2nd} , é esperado ocorrer quando a diferença nos 'counts' dos 'gates' consecutivos torna-se positiva.

Os 'counts' no segundo pico significativo, Y_{2nd} , são usados para calcular os 'counts' no 'half-power point' do segundo pico, Y_{mid2} , como segue :

$$Y_{mid2} = \frac{(Y_{2nd} - Cnt_{min})}{2} + Cnt_{min}$$

Novamente se faz uma interpolação linear de maneira idêntica a anterior para se determinar a localização exata do ponto de meia potência do segundo pico significativo, G_{tmid2} .

O último parâmetro a ser determinado é a largura total do pico ou picos no primeiro ponto de meia potência. A largura é definida como o número de 'gates' entre G_{tmid1} e a localização, G_{trail} , onde o 'trailing edge' passa através do Y_{mid1} . A largura é computada como segue :

$$Width = G_{trail} - G_{tmid1}$$

Sumariamente os parâmetros para um retorno especular com um só pico significativo são :

$$\beta_1 = Y_n$$

$$\beta_2 = Y_{1st}$$

$$\beta_3 = G_{tmid1}$$

$$\beta_4 = S_{lp1st}$$

$$\beta_5 = width$$

Os parâmetros para um retorno especular com dois picos significantes são :

$$\beta_1 = Y_n$$

$$\beta_2 = Y_{1st}$$

$$\beta_{33} = Gt_{mid1}$$

$$\beta_{34} = Slp_{1st}$$

$$\beta_{35} = Y_{2nd}$$

$$\beta_{36} = Gt_{mid2}$$

$$\beta_{37} = Slp_{2nd}$$

$$\beta_{38} = width$$

$$\beta_{39} = Cnt_{min}$$

2.3.3. AJUSTE DE RETORNOS DIFUSOS

Para rastrear o retorno difuso, precisa-se modelar o retorno altimétrico com uma função que tenha a posição de rastreamento (ponto médio da 'leading edge') como um parâmetro. O método dos mínimos quadrados Bayesiano é usado para estimar os parâmetros da função que melhor se ajustam ao retorno. Neste método, as estimativas iniciais dos parâmetros devem ser fornecidas. São dados pesos a estas estimativas iniciais que designam quão bem cada parâmetro está conhecido em relação aos outros. Resíduos são calculados entre o valor de retorno e o valor da função em cada 'gate'. Estes resíduos são avaliados baseados em sua proximidade com a posição do ponto médio da 'leading edge'. Procura-se um mínimo para a soma destes resíduos ao quadrado através de um método iterativo o qual simultaneamente ajusta todos os parâmetros da função. O processo é repetido até a convergência ou até que um máximo número de iterações seja alcançado. Pelo fato de que métodos lineares são usados para resolver problemas não lineares, o procedimento pode ser numericamente instável, o que requer checagem a fim de que se tenha resultados razoáveis. O caminho para que se faça este método funcionar corretamente está na escolha das estimativas iniciais e das funções pesos. Na figura 8, pode-se observar 'waveforms' ditos difusos.

O tratamento dado às 'waveforms' no ajuste proposto pela NASA é bastante complexo e extenso. Existe no entanto, uma semelhança na forma entre uma 'waveform' (veja figura 7) e uma distribuição

log-normal truncada. Isto sugere que possa ser possível tratar esta 'waveform' matematicamente como sendo uma distribuição log-normal truncada (veja figura 9), com sessenta e três intervalos correspondentes cada qual a uma das sessenta e três frequências estatísticas do retorno altimétrico. A partir disto, apresentar-se-ão princípios estatísticos concernetes a esta hipótese a fim de permitir um melhor entendimento do que se propõe nesta dissertação.

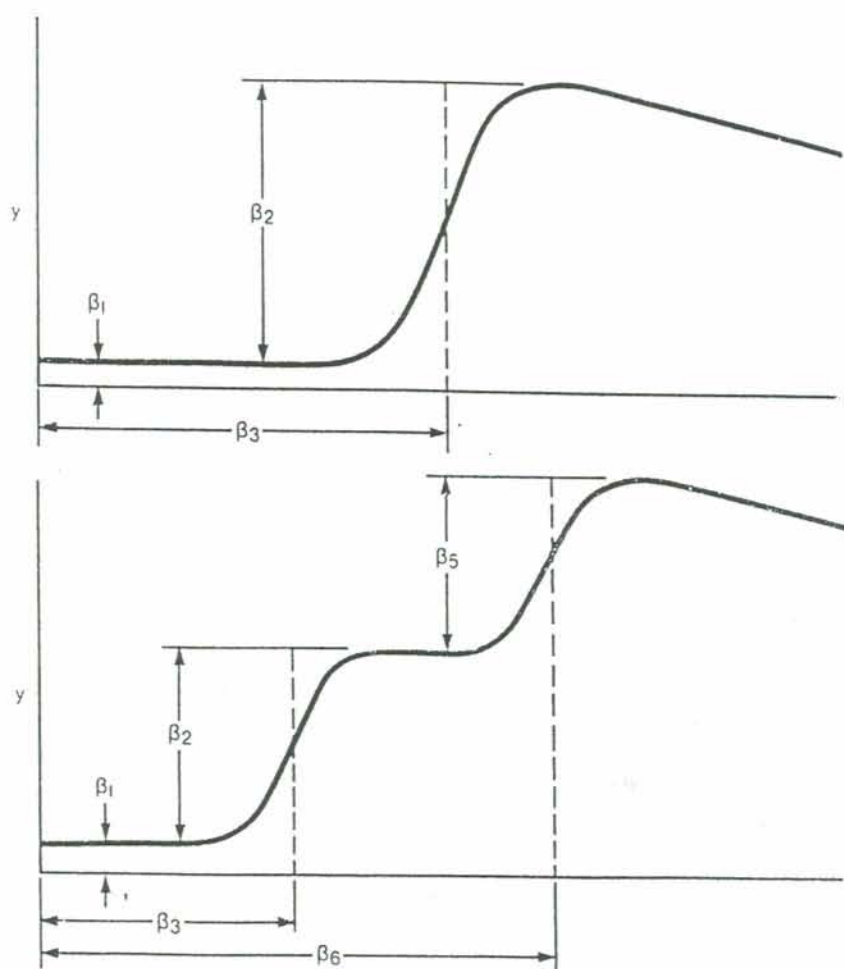


Fig. 8 - Retorno altimétrico em termos difusos

2.4. PRINCÍPIOS ESTATÍSTICOS

2.4.1. DISTRIBUIÇÃO DE PROBABILIDADES

2.4.1.1 Variáveis aleatórias

As variáveis aleatórias podem ser entendidas como sendo uma função definida sobre os elementos do espaço amostral. Estas variáveis aleatórias podem ser discretas, as quais podem assumir somente um número finito ou um número infinito de valores contáveis, ou então podem ser contínuas.

É possível expressar a distribuição de probabilidade por meio de funções, denotadas por $f(x)$. Se $f(x)$ é um valor de densidade de probabilidade, tem-se que :

$$f(x) \geq 0 \text{ para todos os } x \text{ e}$$

$$\sum f(x) = 1 .$$

Muitas vezes, está-se interessado, não só na distribuição de probabilidade $f(x)$ cujo o valor da variável aleatória é x , mas também na probabilidade $F(x)$ cujo o valor da variável aleatória é menor ou igual a x . A função que associa um valor $F(x)$ a cada x num conjunto de variáveis aleatórias, chama-se função de distribuição de uma variável aleatória.

2.4.1.2. Média e Variância de uma Distribuição de Probabilidade

A média e a variância são duas medidas estatísticas muito importantes que descrevem, respectivamente, a localização e a variação de uma distribuição de probabilidade.

A média de uma distribuição de probabilidade é simplesmente a esperança matemática de uma variável aleatória correspondente. Se uma variável aleatória tem valores $x_1, x_2, \dots, \text{ou } x_k$, com as probabilidades $f(x_1), f(x_2), \dots, \text{e } f(x_k)$, a sua esperança matemática

é:

$$x_1 * f(x_1) + x_2 * f(x_2) + \dots + x_k * f(x_k) \text{ ou}$$

$\mu = \sum x * f(x)$, onde μ é a média de uma distribuição de probabilidade discreta. Esta média pode ser dividida pelo $\sum f(x)$, salvo se esta soma se iguale a um, por definição.

A variância de uma distribuição de probabilidades com valores $f(x)$ é :

$$\sigma^2 = \sum (x - \mu)^2 * f(x)$$

Esta medida não está na mesma unidade ou dimensão que os valores da variável aleatória, o que pode ser ajustado tomando-se a raiz quadrada, o que define o desvio padrão :

$$\sigma = \sqrt{\sum (x - \mu)^2 * f(x)}$$

2.4.2. DENSIDADE DE PROBABILIDADES

2.4.2.1. Variáveis Aleatórias Contínuas

No caso contínuo, as variáveis aleatórias podem assumir valores numa escala contínua. Os resultados de um experimento são representados por pontos num segmento de linha ou numa linha, e o valor da variável aleatória é um número apropriadamente associado a cada ponto por meio de alguma regra ou equação. Geralmente escreve-se $P(a \leq x \leq b)$ para a probabilidade associada com os pontos do espaço amostral, para o qual o valor de uma variável aleatória cai no intervalo de a a b . Supõe-se que se divide o intervalo de a até b , sendo estes constantes e $a \leq b$, em n subintervalos iguais de largura Δx , contendo respectivamente os pontos x_1, x_2, \dots, x_n , e que a probabilidade que a variável aleatória assumirá num valor no subintervalo contendo x_i , é dado por

$f(x_i)\Delta x$. Assim, a probabilidade que a variável aleatória assumirá no intervalo de a a b é dada por :

$$P(a \leq x \leq b) = \sum_{i=1}^n f(x_i) \Delta x$$

Se f é uma função integrável definida para todos os valores variável aleatória, define-se a probabilidade fazendo $\Delta x \rightarrow 0$, isto é

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$

Esta definição de probabilidade no caso contínuo pressupõe a existência de uma função apropriada f , a qual, integrada de a até b (sendo $a \leq b$, constantes), dá a probabilidade que a correspondente variável aleatória tem no intervalo de a até b . No caso contínuo, probabilidades são dadas por integrais e não pelos valores de f . A função f pode ser chamada de função densidade de probabilidade ou densidade de probabilidades, as quais são integradas para obter-se probabilidades.

$f(x) \geq 0$ para todos os x dentro do domínio de f .

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Como no caso discreto, escreve-se $F(x)$ como a probabilidade que a variável aleatória com função densidade aleatória f assume num valor menor ou igual a x , e refere-se a função correspondente F como função de distribuição de uma variável aleatória. De modo que, se uma variável aleatória com valores x tem a função densidade de probabilidade f , os valores de sua função distribuição são dados por :

$$F(x) = \int_{-\alpha}^x f(t) dt.$$

Conseqüentemente, a probabilidade que a variável aleatória assumirá no intervalo de α até b é $F(b) - F(\alpha)$, seguindo que a $dF(x)/dx$ é igual a $f(x)$, onde esta derivada existir.

2.4.2.2. Média e Variância de uma Densidade de Probabilidade

As medidas estatísticas usadas para descrever as densidades probabilísticas são muito afins às utilizadas para descrever as distribuições de probabilidade. Substituindo-se os somatórios por integrais tem-se :

$$\mu = \int_{-\alpha}^{\alpha} x f(x) dx, \text{ a média da densidade probabilística.}$$

$$\sigma^2 = \int_{-\alpha}^{\alpha} (x-\mu)^2 f(x) dx, \text{ a variância da densidade probabilística.}$$

2.4.3. DISTRIBUIÇÃO LOG-NORMAL

Esta distribuição contínua ocorre em prática, quando tem-se uma variável aleatória tal, que seu algoritmo tem uma distribuição normal. A sua função densidade de probabilidade é dada por :

$$f(x) = \frac{1}{\sqrt{2\pi\beta}} x^{-1} e^{-\frac{(\ln x - \alpha)^2}{2\beta^2}} \quad \text{para } x > 0, \beta > 0$$

$$f(x) = 0 \quad \text{para demais intervalos,}$$

onde $\ln x$ é o logaritmo natural de x . Para $\alpha = 0$ e $\beta=1$, tem-se um gráfico da log-normal como na figura 9.

Para se encontrar a probabilidade de uma variável aleatória de distribuição log-normal, entre os valores 'a' e 'b' ($0 < a < b$), avalia-se a seguinte integral :

$$\int_{\ln a}^{\ln b} \frac{1}{\sqrt{2\pi\beta}} x^{-1} e^{-(\ln x - \alpha)^2 / 2\beta^2} dx$$

Fazendo-se $y = \ln x$ e identificando-se o integrando como uma densidade normal com $\mu = \alpha$ e $\sigma = \beta$, tem-se que a probabilidade desejada é dada por :

$$\int_{\ln a}^{\ln b} \frac{1}{\sqrt{2\pi\beta}} x^{-1} e^{-(\ln x - \alpha)^2 / 2\beta^2} dy = F(\ln b - \alpha / \beta) - F(\ln a - \alpha / \beta),$$

onde F é a função distribuição da distribuição normal padrão. Para encontrar a fórmula para a média da distribuição log-normal, escreve-se :

$$\mu = \frac{1}{\sqrt{2\pi\beta}} \int_0^{\alpha} x * x^{-1} e^{-(\ln x - \alpha)^2 / 2\beta^2} dx$$

Fazendo-se $y = \ln x$, isto fica :

$$\mu = \frac{1}{\sqrt{2\pi\beta}} \int_{-\alpha}^{\alpha} e^y e^{-(y - \alpha)^2 / 2\beta^2} dy$$

Esta integral pode ser avaliada, obtendo-se um integrando cuja

a forma seja de uma densidade normal, de modo que :

$\mu = e^{\alpha + \beta^2/2}$, é a média da distribuição log-normal.

$\sigma^2 = e^{2\alpha + \beta^2} (e^{\beta^2} - 1)$, é a variância da distribuição log-normal.

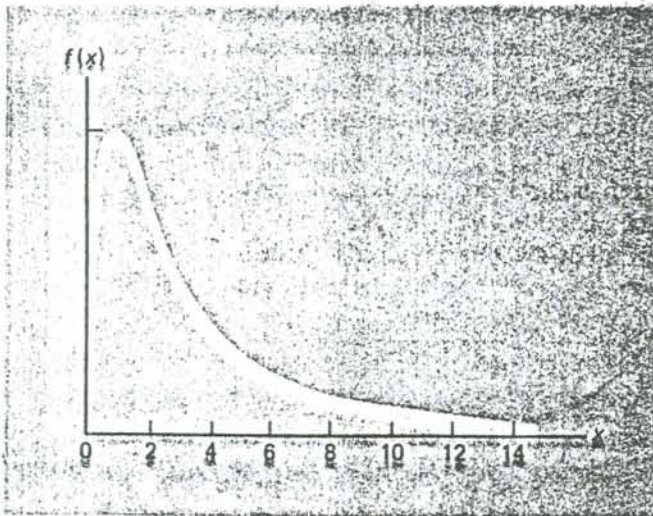


Fig. 9 - Distribuição Log-normal

2.4.4. FUNÇÃO VEROSSIMILHANÇA

Supondo-se que (x_1, x_2, \dots, x_n) seja uma amostra aleatória de uma variável aleatória x , cuja função de distribuição probabilística (no caso discreto) ou a função densidade de probabilidade (no caso contínuo) seja $f(x; \theta)$ o qual depende do parâmetro de população.

A função de distribuição pode depender de mais de um parâmetro, por exemplo, θ pode depender de um par de parâmetros (μ, σ) que é o caso quando x tem uma distribuição normal $N(\mu, \sigma)$ ou

log-normal truncada.

Como (x_1, \dots, x_n) é uma amostra de $f(x, \theta)$ a função densidade de probabilidade da amostra é $f(x_1; \theta) \dots f(x_n; \theta)$, o que se denota por $l(\theta/x_1, \dots, x_n)$ ou $l(x_1, \dots, x_n; \theta)$.

$$l(x_1, \dots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta)$$

Esta função supra mencionada é a função de verossimilhança de θ para uma dada amostra (x_1, \dots, x_n) .

A estimativa da máxima verossimilhança $\hat{\theta}$ para θ é o valor de θ (se existir) tal que :

$$l(\hat{\theta}/x_1, \dots, x_n) > l(\theta'/x_1, \dots, x_n),$$

onde θ' é um outro valor possível de θ . Em outras palavras a estimativa de verossimilhança é o valor de θ que maximiza a função de verossimilhança.

É mais conveniente trabalhar com o logaritmo natural de $l(\theta/x_1, \dots, x_n)$, o que se denota por $\log l(\theta/x_1, \dots, x_n)$ ou $L(\theta/x_1, \dots, x_n)$. Pode-se ver que L tem seu máximo no mesmo valor de θ que o l .

Desda forma o valor de θ para o qual L é máxima é a estimativa de máxima verossimilhança $\hat{\theta}$ para θ .

Naturalmente, $\hat{\theta}$ é uma estatística e, por isso, uma variável aleatória, já que seu valor dependerá da amostra (x_1, \dots, x_n) . Pode acontecer que a distribuição de probabilidade de x dependa de dois ou mais valores de parâmetros, como ocorre com a distribuição normal. Neste caso, θ representará um vetor, $\theta = (\alpha, \beta)$.

A fim de encontrar a estimativa de máxima verossimilhança, deve-se determinar o valor máximo da função, para o que pode se utilizar técnicas usuais de cálculo.

Visto que $\ln x$ é uma função crescente de x ,

$$\ln l(\theta/x_1, \dots, x_n)$$

alcançará seu valor máximo para o mesmo valor de θ que fará com $L(\theta/x_1, \dots, x_n)$.

Admitindo-se que θ seja um número real e que $L(\theta/x_1, \dots, x_n)$ seja uma função derivável em θ , obteremos a estimativa de máxima

verossimilhança pela resolução de :

$$\frac{\delta}{\delta\alpha} \ln l(\theta/x_1, \dots, x_n) = 0$$

que é conhecida como equação de verossimilhança.

Se $\theta = (\alpha, \beta)$, a equação acima deverá ser substituída pelas equações simultâneas de verossimilhança :

$$\frac{\delta}{\delta\alpha} \ln l(\alpha, \beta/x_1, \dots, x_n) = 0 \text{ e}$$

$$\frac{\delta}{\delta\beta} \ln l(\alpha, \beta/x_1, \dots, x_n)$$

Deve-se salientar que a orientação acima nem sempre resulta num sistema de cálculo que converge. No entanto, em um grande número de exemplos importantes, este método fornece, com relativa facilidade, a desejada estimativa de máxima verossimilhança.

Se uma variável x tiver uma distribuição normal truncada, esta distribuição será dita truncada à direita de $x = \tau$ se sua função de distribuição de probabilidade for $f(x) = 0$ se $x > \tau$ e a distribuição será dita truncada à esquerda de $x = \gamma$, se a função de distribuição de probabilidade for $f(x) = 0$ se $x < \gamma$.

Se uma variável contínua com função de frequência $f(x/\theta)$ é duplamente truncada em dois pontos conhecidos a, b com $a < b$, a função verossimilhança para n observações é :

$$l(x/\theta) = \prod_{i=1}^n f(x_i/\theta) / \left(\int_a^b f(x/\theta) dx \right)^n \quad a < x < b.$$

Esta expressão é válida para x perfeitamente definido no intervalo a, b .

III. MATERIAIS E MÉTODOS

Neste capítulo serão apresentados os procedimentos e materiais utilizados para o desenvolvimento deste trabalho de dissertação.

3.1. DADOS

Para a análise proposta nesta dissertação, tem-se como dados os retornos altimétricos tanto da região Amazônica como da Lagoa dos Patos.

Estes retornos altimétricos consistem em um conjunto de sessenta e três frequências estatísticas, sendo cada qual, a amplitude do retorno por 'gate'. Cada 'waveform' é identificada por um número, sendo também informada a sua posição em termos de latitude e longitude.

Na figura 10, pode-se observar a localização das trilhas do Geosat, espaçadas em aproximadamente cento e cinquenta quilômetros. Cada 'waveform' está amostrada na trilha do satélite a intervalos de 662 metros, correspondente à área observada pelo radar no solo, cujo o feixe de abertura é de cerca de catorze quilômetros [Koblinsky et al., 1992].

Na figura 11, observa-se uma imagem landsat da região de Manaus (AM), com a sobreposição das posições que correspondem às 'waveforms'. Na região Amazônica as trilhas do Geosat são 'track' 131, 173, 196 e 238. Sobre a Lagoa dos Patos é o 'track' 45. Estes dados estão armazenados em disquetes e precisam ser reordenados a fim de serem processados matematicamente.

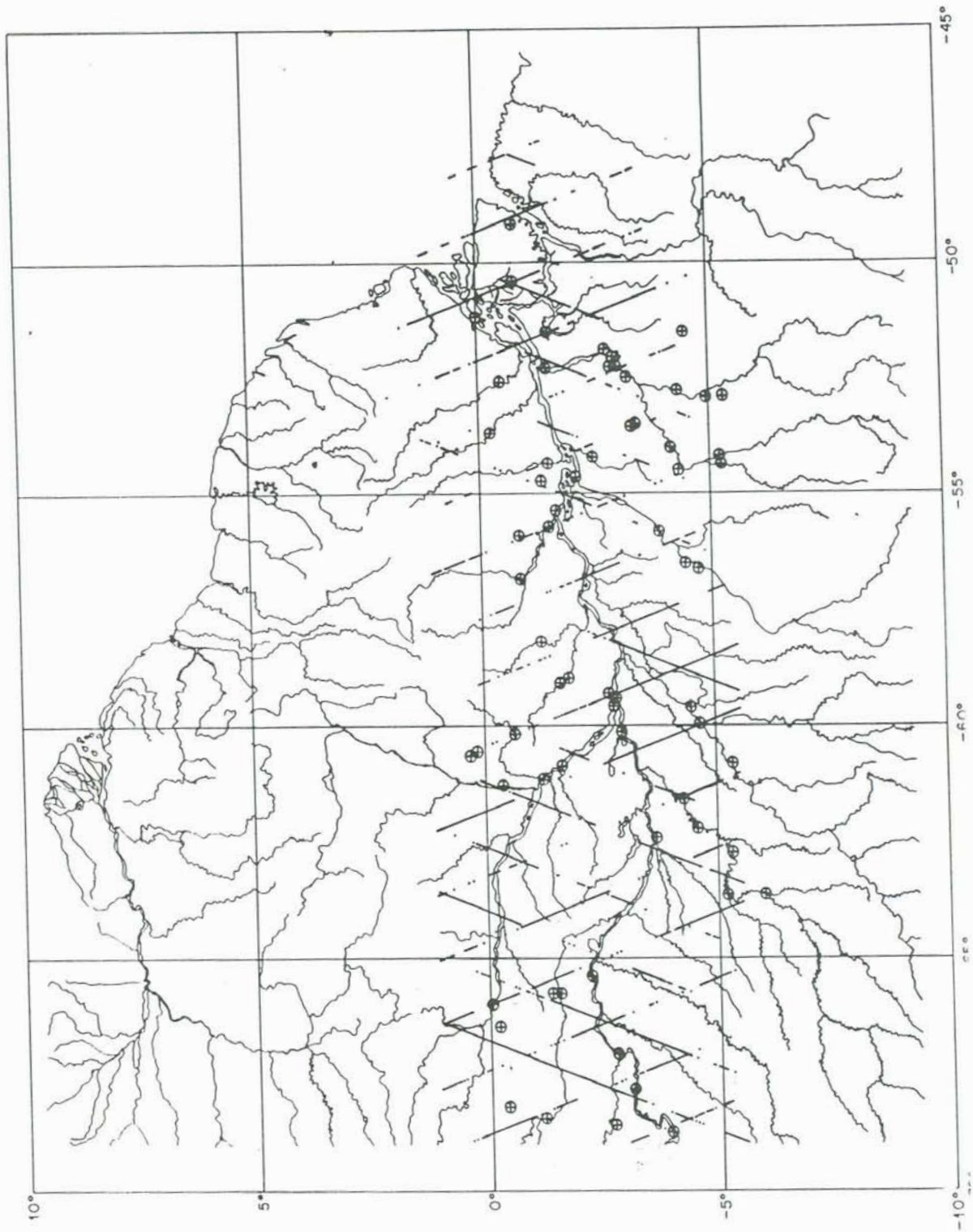


Fig.10 - Órbita do Altímetro Geosat sobre a
Amazônia

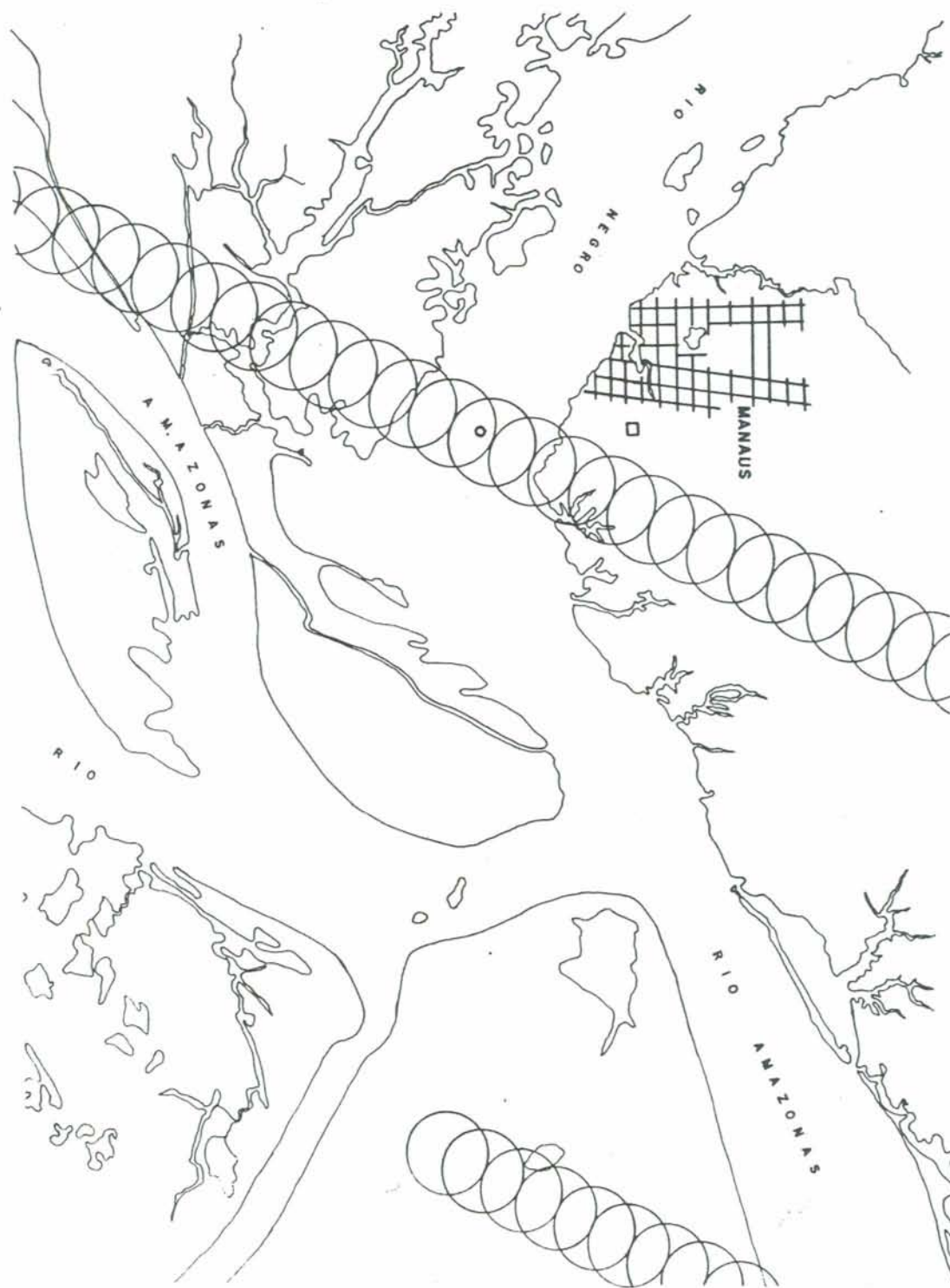


Fig.11 - Imagem Landsat sobre Manaus com as trilhas do Geosat.

Na tabela 3 abaixo, apresentam-se alguns exemplos de 'waveforms' sobre o Amazonas e sobre a Lagoa dos Patos.

TABELA 3:

Exemplo de 'waveforms' sobre a região Amazônica :

WD	355	-3811161(=3.8°S)	300556544(=30.05°W)									*
91	102	92	69	95	96	82	91	85	98	88	91	
111	112	94	72	88	108	121	95	125	128	135	126	
124	92	92	101	118	106	180	132	219	233	185	242	
277	254	210	217	188	176	194	195	207	160	206	209	
201	189	163	168	173	170	147	154	106	114	131	129	
112	161	152										

Exemplo de 'waveform' sobre a Lagoa dos Patos :

WD	41429	-31852720(=31.8°S)	308562432(=30.85°W)									
22	24	16	16	18	16	16	14	18	18	16	14	
14	16	16	18	16	20	16	18	16	16	14	21	
21	17	17	21	30	79	266	325	351	268	322	394	
299	326	367	349	418	364	395	321	335	375	402	310	
353	325	318	309	369	316	445	346	386	368	345	298	
47	155	322										

* Sequencialmente estão o número da waveform e a sua latitude e longitude.

3.2. MATLAB

Para o desenvolvimento dos programas necessários à dissertação, optou-se pelo uso de um pacote matemático, o MATLAB. Este software, requer para seu funcionamento um computador IBM PC, PC/XT, PC/AT, ou um computador compatível MS-DOS. Pelo menos 320K de memória, uma versão 2.0 ou melhor do MS-DOS, e um coprocessador numérico.

No MATLAB trabalha-se com matrizes, vetores e escalares. Permite uma série de operações matriciais e possui recursos que permitem realizar manipulação de dados. A vantagem do Matlab vem de seu extensivo conjunto de funções (em torno de 150), algumas intrínsecas, outras construídas pelo próprio operador, o que fornece ao usuário certa liberdade de adaptar o pacote às suas necessidades. Possui ainda um bom recurso gráfico, incluindo gráficos de três dimensões. Estão disponíveis no pacote comandos de fluxo, o que facilita a programação. Pode ser operado no modo comando, quando comandos isolados entram e ele processa imediatamente fornecendo resultados. Porém, também é capaz de executar uma seqüência de comandos que são armazenados em arquivos. Estes dois modos de funcionamento fornecem um ambiente interpretativo.

Os arquivos que contém as expressões do matlab são chamados M-files por serem um arquivo tipo '.m'. Estes arquivos consistem de uma seqüência de expressões do matlab. Um dos usos dos M-files é automatizar uma longa seqüência de comandos, o que se chama de 'script files'. Outro uso são os 'function files', que permitem que novas funções sejam acrescentadas as já existentes, o que possibilita resolver problemas específicos do usuário.

Tanto os 'script' como os 'function files' são arquivos textos ASCII, e portanto criados usando-se um editor qualquer. No caso utilizou-se o editor Norton, que é facilmente acessado de dentro do próprio Matlab.

Um arquivo 'function' difere do 'script', em que as variáveis definidas e manipuladas dentro do arquivo function são locais e não operam globalmente. Estes arquivos 'function' permitem então criar novas funções do Matlab com a própria linguagem Matlab.

3.3. PROCEDIMENTOS

3.3.1. Reordenação de frequências

Anteriormente foram apresentados exemplos de 'waveforms'.

Nota-se que as três últimas frequências, correspondem de fato as posições dos 'gates' 29.5 , 30.5 , e 31.5, espaçadas de 23 cm ao invés de 46 cm. A fim de tornar os dados matematicamente mais operacionais, reescreveu-se as 'waveforms' utilizadas no trabalho na forma de vetor, legível portanto, pelo Matlab. Deste modo, cada 'waveform', constitui-se de um vetor de 63 linhas e 1 coluna, correspondendo a 63 'gates' consecutivos. As três últimas frequências precisam ser relocadas às posições respectivas dos 'gates' 31, 32 e 33. Cada frequência está distribuída em intervalos de 46 centímetros, salvo as frequências dos 'gates' 31, 32 e 33, que estão distribuídas em intervalos de 23 centímetros. Procedeu-se a um reescalonamento dos intervalos, associando-se os de 46 cm a intervalos de 0.1, e os de 23 cm a intervalos de 0.05.

Isto pode ser melhor descrito na tabela abaixo, onde :

- I - 'gates' correspondentes as frequências de 1 a 63,
- II - limite esquerdo do intervalo,
- III - limite direito do intervalo,
- IV - largura do intervalo,
- V - ponto médio do intervalo,
- VI - limite esquerdo do intervalo modificado,
- VII - limite direito do intervalo modificado,
- VIII - largura do intervalo modificado,
- IX - Ponto médio do intervalo modificado.

Tab 4 - Intervalos de freqüências estatísticas

I	II	III	IV	V	VI	VII	VIII	IX
1	-6.15	-5.95	0.2	-6.05	0.425	0.525	0.1	0.475
2	-5.95	-5.75	0.2	-5.85	0.525	0.625	0.1	0.575
3	-5.75	-5.55	0.2	-5.65	0.625	0.725	0.1	0.675
4	-5.55	-5.35	0.2	-5.45	0.725	0.825	0.1	0.775
5	-5.35	-5.15	0.2	-5.25	0.825	0.925	0.1	0.875
6	-5.15	-4.95	0.2	-5.05	0.925	1.025	0.1	0.975
7	-4.95	-4.75	0.2	-4.85	1.025	1.125	0.1	1.075
8	-4.75	-4.55	0.2	-4.65	1.125	1.225	0.1	1.175
9	-4.55	-4.35	0.2	-4.45	1.225	1.325	0.1	1.275
10	-4.35	-4.15	0.2	-4.25	1.325	1.425	0.1	1.375
11	-4.15	-3.95	0.2	-4.05	1.425	1.525	0.1	1.475
12	-3.95	-3.75	0.2	-3.85	1.525	1.625	0.1	1.575
13	-3.75	-3.55	0.2	-3.65	1.625	1.725	0.1	1.675
14	-3.55	-3.35	0.2	-3.45	1.725	1.825	0.1	1.775
15	-3.35	-3.15	0.2	-3.25	1.825	1.925	0.1	1.875
16	-3.15	-2.95	0.2	-3.05	1.925	2.025	0.1	1.975
17	-2.95	-2.75	0.2	-2.85	2.025	2.125	0.1	2.075
18	-2.75	-2.55	0.2	-2.65	2.125	2.225	0.1	2.175
19	-2.55	-2.35	0.2	-2.45	2.225	2.325	0.1	2.275
20	-2.35	-2.15	0.2	-2.25	2.325	2.425	0.1	2.375
21	-2.15	-1.95	0.2	-2.05	2.425	2.525	0.1	2.475
22	-1.95	-1.75	0.2	-1.85	2.525	2.625	0.1	2.575
23	-1.75	-1.55	0.2	-1.65	2.625	2.725	0.1	2.675
24	-1.55	-1.35	0.2	-1.45	2.725	2.825	0.1	2.775
25	-1.35	-1.15	0.2	-1.25	2.825	2.925	0.1	2.875
26	-1.15	-0.95	0.2	-1.05	2.925	3.025	0.1	2.975
27	-0.95	-0.75	0.2	-0.85	3.025	3.125	0.1	3.075
28	-0.75	-0.55	0.2	-0.65	3.125	3.225	0.1	3.175
29	-0.55	-0.35	0.2	-0.45	3.225	3.325	0.1	3.275
30	-0.35	-0.15	0.2	-0.25	3.325	3.425	0.1	3.375
31	-0.15	-0.05	0.1	-0.10	3.425	3.475	0.05	3.450
32	-0.05	0.05	0.1	0.00	3.475	3.525	0.05	3.500
33	0.05	0.15	0.1	0.10	3.525	3.575	0.05	3.550
34	0.15	0.35	0.2	0.25	3.575	3.675	0.1	3.625
35	0.35	0.55	0.2	0.45	3.675	3.775	0.1	3.725
36	0.55	0.75	0.2	0.65	3.775	3.875	0.1	3.825
37	0.75	0.95	0.2	0.85	3.875	3.975	0.1	3.925
38	0.95	1.15	0.2	1.05	3.975	4.075	0.1	4.025
39	1.15	1.35	0.2	1.25	4.075	4.175	0.1	4.125
40	1.35	1.55	0.2	1.45	4.175	4.275	0.1	4.225
41	1.55	1.75	0.2	1.65	4.275	4.375	0.1	4.325
42	1.75	1.95	0.2	1.85	4.375	4.475	0.1	4.425
43	1.95	2.15	0.2	2.05	4.475	4.575	0.1	4.525
44	2.15	2.35	0.2	2.25	4.575	4.675	0.1	4.625
45	2.35	2.55	0.2	2.45	4.675	4.775	0.1	4.725
46	2.55	2.75	0.2	2.65	4.775	4.875	0.1	4.825
47	2.75	2.95	0.2	2.85	4.875	4.975	0.1	4.925
48	2.95	3.15	0.2	3.05	4.975	5.075	0.1	5.025

I	II	III	IV	V	VI	VII	VIII	IX
49	9.15	9.95	0.2	9.25	5.075	5.175	0.1	5.125
50	9.95	9.55	0.2	9.45	5.175	5.275	0.1	5.225
51	9.55	9.75	0.2	9.65	5.275	5.975	0.1	5.925
52	9.75	9.95	0.2	9.85	5.975	5.475	0.1	5.425
53	9.95	4.15	0.2	4.05	5.475	5.575	0.1	5.525
54	4.15	4.95	0.2	4.25	5.575	5.675	0.1	5.625
55	4.95	4.55	0.2	4.45	5.675	5.775	0.1	5.725
56	4.55	4.75	0.2	4.65	5.775	5.875	0.1	5.825
57	4.75	4.95	0.2	4.85	5.875	5.975	0.1	5.925
58	4.95	5.15	0.2	5.05	5.975	6.075	0.1	6.025
59	5.15	5.95	0.2	5.25	6.075	6.175	0.1	6.125
60	5.95	5.55	0.2	5.45	6.175	6.275	0.1	6.225
61	5.55	5.75	0.2	5.65	6.275	6.975	0.1	6.925
62	5.75	5.95	0.2	5.85	6.975	6.475	0.1	6.425
63	5.95	6.15	0.2	6.05	6.475	6.575	0.1	6.525

tabela 4

Para que todas as 63 frequências fiquem distribuídas num intervalo tal, que varia de 0.425 até 6.575, foram feitas operações aditivas as quais definem as colunas VI, VII, VIII e IX da tabela anterior.

É possível construir um gráfico das frequências estatísticas distribuídas nos 63 'gates', a fim de observar o comportamento deste que é o retorno altimétrico. Dentro do Matlab se processou esta plotagem dividindo-se cada uma das frequências pela soma total de frequências, obtendo-se então, um gráfico em termos relativos.

3.3.2. Avaliação da função log-normal truncada

A distribuição log-normal truncada à qual quer-se ajustar as 'waveforms' é definida por :

$$f(x) = \frac{\frac{1}{\sigma x \sqrt{2\pi}} \exp \left[\frac{-1}{2\sigma^2} (\ln x - \mu)^2 \right]}{\int_{x_0}^{x_n} \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[\frac{-1}{2\sigma^2} (\ln x - \mu)^2 \right] dx}$$

para $x_0 \leq x \leq x_n$.

Esta $f(x)$ é uma função de dois parâmetros μ e σ (média e

variância σ^2), que podem ser avaliadas pelo método da verossimilhança. Isto permite avaliar a 'waveform' em função de duas incógnitas ao invés de cinco ou nove, como é o caso do método de ajuste desenvolvido pela NASA.

3.3.2.1. Verossimilhança

A função verossimilhança para o caso da distribuição log-normal truncada é dada pela fórmula :

$$l(\mu, \sigma^2) = \frac{\prod_{i=1}^n \left[\int_{x_{(i-1)}}^{x_i} \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[\frac{-1}{2\sigma^2} (\ln x - \mu)^2 \right] dx \right]^{f_i}}{\left[\int_{x_0}^{x_n} \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[\frac{-1}{2\sigma^2} (\ln x - \mu)^2 \right] dx \right]^{f_1 + f_2 + \dots + f_n}}$$

- n - 63 frequências estatísticas,
- f_i - valores das frequências,
- x_i - são os intervalos das frequências,
- μ - média,
- σ^2 - variância.

É preciso desenvolver esta fórmula da função de verossimilhança $l(\mu, \sigma^2)$ para as 'waveforms', para que uma vez definidos e avaliados seus parâmetros μ e σ , se possa aplicá-los à log-normal truncada. Esta última por sua vez, se ajustada, deveria definir a 'waveform'.

A função de verossimilhança pode ser calculada como segue :

chamando $I = \frac{1}{\sigma \sqrt{2\pi}} \int_{x_{i-1}}^{x_i} \frac{1}{x} \exp \left[\frac{-1}{2\sigma^2} (\ln x - \mu)^2 \right] dx$

escrevendo : $t = \frac{\ln x - \mu}{\sigma}$

tem-se : $\frac{dt}{dx} = \frac{1}{\sigma x}$, portanto : $dt = \frac{1}{\sigma x} dx$

$$I = \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln x_{(i-1)} - \mu}{\sigma}}^{\frac{\ln x_i - \mu}{\sigma}} \exp \left[-\frac{1}{2} t^2 \right] dt$$

A função cumulativa de distribuição normal é

$$\Phi(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z \exp^{-0.5t^2} dt$$

Assim,

$$I = \Phi \left[\frac{\ln x_i - \mu}{\sigma} \right] - \Phi \left[\frac{\ln x_{i-1} - \mu}{\sigma} \right]$$

O ϕ pode ser calculado por uma aproximação racional e polinomial par $0 \leq x \leq \alpha$.

$$\phi(x) = 1 - Z(x)(b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5),$$

onde :

$$t = \frac{1}{1+px}$$

$$p = 0.2316419$$

$$b_1 = 0.319381530$$

$$b_4 = -1.821255978$$

$$b_2 = -0.356563782$$

$$b_5 = 1.330274429$$

$$b_3 = 1.781477937$$

$$Z(x) = \frac{1}{\sqrt{2\pi}} \exp^{-0.5x^2}$$

Partiu-se então para a construção de duas funções dentro do Matlab. A primeira capaz de calcular a função de distribuição normal $\phi(Z)$, e a segunda que utilizando-se desta, vai avaliar numericamente a função log de verossimilhança.

A função log de verossimilhança para a função log-normal truncada (programada), pode ser descrita como :

$$l(\mu, \sigma^2) = \frac{\left[\phi \left[\frac{\ln(\text{ulim}(1)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right]^{f_1}}{\left[\phi \left[\frac{\ln(\text{ulim}(\sigma_3)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right]} \times$$

$$\frac{\left[\phi \left[\frac{\ln(\text{ulim}(2)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(2)) - \mu}{\sigma} \right] \right]^{f_2}}{\left[\phi \left[\frac{\ln(\text{ulim}(\sigma_3)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right]} \times \dots \times$$

$$\frac{\left[\phi \left[\frac{\ln(\text{ulim}(\sigma_3)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(\sigma_3)) - \mu}{\sigma} \right] \right]^{f_{\sigma_3}}}{\left[\phi \left[\frac{\ln(\text{ulim}(\sigma_3)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right]}$$

onde os valores ulim e llim estão definidos nas colunas VI e VII da tabela 4, f_i são os valores das frequências da 'waveform' analisada e μ e σ^2 são a média e a variância para a qual se está avaliando a $l(\mu, \sigma^2)$.

Se aplicarmos o logaritmo a expressão acima ficará :

$$L(\mu, \sigma^2) = \frac{f_1 \times \left[\phi \left[\frac{\ln(\text{ulim}(1)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right] + \dots}{(f_1 + f_2 + \dots + f_{63}) \times \left[\phi \left[\frac{\ln(\text{ulim}(63)) - \mu}{\sigma} \right] - \phi \left[\frac{\ln(\text{llim}(1)) - \mu}{\sigma} \right] \right]}$$

(1)

A estratégia usada é a avaliação desta função log de verossimilhança (1) para um conjunto de valores de μ e σ^2 , partindo-se de estimativas iniciais de média e variância. Desta forma cria-se uma espécie de rede de médias e variâncias próximas da estimativa inicial, dentro da qual espera-se encontrar os valores que maximizam a $L(\mu, \sigma)$.

Matematicamente, o par (μ, σ^2) que maximiza a função de verossimilhança, isto é, os valores de média e variância para os quais a função é máxima, serão os mesmos que maximizarão a função logarítmica de verossimilhança.

As estimativas iniciais da média e variância são calculadas baseando-se na definição da distribuição log-normal, com os valores das 63 frequências da 'waveform'.

A média μ é:

$$\mu = e^{\alpha + \beta^2/2}$$

e a variância é:

$$\sigma^2 = e^{2\alpha + \beta^2} (e^{\beta^2} - 1)$$

o que pode se escrito como:

$$\mu = e^{\mu + \sigma^2/2} \quad e \quad \sigma^2 = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

A partir disto tem-se :

$$e^{\mu + \sigma^2/2} = \frac{\sum f_i x_i}{\sum f_i} = M \quad (1)$$

$$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} = S^2 \quad (2)$$

Elevando-se (1) ao quadrado e dividindo-se (2) por (1) chega-se a:

$$e^{2\mu + \sigma^2} = M^2$$

$$\frac{e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)}{e^{2\mu + \sigma^2}} = \frac{S^2}{M^2} \quad e^{\sigma^2} - 1 = \frac{S^2}{M^2}$$

$$\sigma^2 = \ln \left(1 + \frac{S^2}{M^2} \right)$$

Após outras operações matemáticas, chega-se a:

$$M^2 = e^{2\mu} \cdot e^{\sigma^2} = e^{2\mu} \left[1 + \frac{S^2}{M^2} \right]$$

$$M^2 = \frac{M^2 e^{2\mu} + e^{2\mu} S^2}{M^2} \quad M^4 = e^{2\mu} (M^2 + S^2)$$

$$2\mu = \ln \left(\frac{M^4}{M^2 + S^2} \right) \quad \mu = \frac{1}{2} \ln \left[\frac{M^2}{1 + \frac{S^2}{M^2}} \right]$$

$$\mu = \frac{1}{2} \left[\ln M^2 - \ln \left[1 + \frac{S^2}{M^2} \right] \right]$$

Esta média μ , e variância σ^2 , são consideradas as estimativas iniciais, calculadas a partir das frequências estatísticas da 'waveform', com as quais inicia-se a avaliação da função log da verossimilhança, em busca do seu valor máximo. Os valores da média e variância para os quais a log da verossimilhança é máxima, são as incógnitas que quando aplicadas a função log-normal truncada estarão definidos esta função ajustada as 'waveforms'.

3.3.2.2. Maximização da Verossimilhança

Uma vez definidas a função log da verossimilhança e as estimativas iniciais da média e variância para um dado conjunto de frequências estatísticas (waveform), parte-se para a maximização da $L(\mu, \sigma)$, a fim de definir os valores de μ e σ^2 , para os quais a L é máxima. Estes valores, uma vez aplicados à função log-normal truncada, deverão permitir que ela se ajuste a 'waveform'.

No princípio da maximização da verossimilhança, onde $\hat{\delta}$ é a estimativa de δ , tem-se que $L(x/\hat{\delta}) \geq L(x/\delta)$, assume-se que δ possa assumir qualquer valor real em um intervalo.

A determinação da forma do estimador de máxima verossimilhança é relativamente simples para uma situação genérica.

Se a função de verossimilhança é uma função duplamente diferenciável em δ , ou seja, tem duas derivadas, os valores de estimativa máxima de $L(\mu, \sigma)$, serão se existirem, as raízes de :

$$L'(x/\delta) = \frac{\partial L(x/\delta)}{\partial \delta} = 0$$

A condição suficiente, mas não necessária, para que qualquer destes valores estimados seja o máximo local, é que :

$$L''(x/\hat{\delta}) \leq 0$$

Uma vez encontrados os máximos locais da função verossimilhança, tem-se encontrado as soluções da equação, isto é, os valores de δ para os quais a $L(x/\delta)$ é máxima.

Neste caso, calcular as derivadas primeira e segunda da função logarítmica da verossimilhança anteriormente definida é inviável. A maneira encontrada para resolver o problema foi a adoção de procedimentos de minimização não restritos sem o uso de derivadas. Como o objetivo é a maximização e não a minimização, basta efetuar a troca de sinais no algoritmo que calcula a função log verossimilhança.

Estes métodos são conhecidos por métodos de otimização por busca, nos quais a direção da minimização são determinados a partir de avaliações sucessivas da função objeto $f(x)$. Estes métodos não requerem que a função tenha continuidade nem que existam as derivadas da função objeto.

Segundo Himmelblau (Applied Nonlinear Programming) Nelder e Mead propuseram um método de busca que tem se mostrado uma estratégia bastante eficiente, facilmente implementada em computadores.

Na técnica Nelder e Mead, o 'simplex' pode alterar-se na forma, por isso, chama-se também de busca de poliedro flexível. Poliedros regulares em E^n (espaço euclidiano n-dimensional), são 'simplexes'. Para duas variáveis, um simplex regular é um triângulo equilátero (três pontos). Para três variáveis, o simplex regular é um tetraedro regular (quatro pontos) e assim por diante.

O método Nelder Mead, minimiza uma função de n variáveis independentes, usando $(n+1)$ vértices de um poliedro flexível em E^n . Cada vértice pode ser definido por um vetor x . O vértice (ponto) em E^n , que gerar o maior valor de $f(x)$ é projetado através do centro de gravidade (centróide) dos vértices remanescentes. Veja figuras 12 e 13. Valores melhorados (mais baixos) da função objeto são encontrados por sucessivas substituições dos pontos com valor mais alto de $f(x)$ por melhores pontos, até que o mínimo de $f(x)$ seja encontrado.

Fig. 12 : Simplexes regulares para 2 e 3 variáveis independentes.

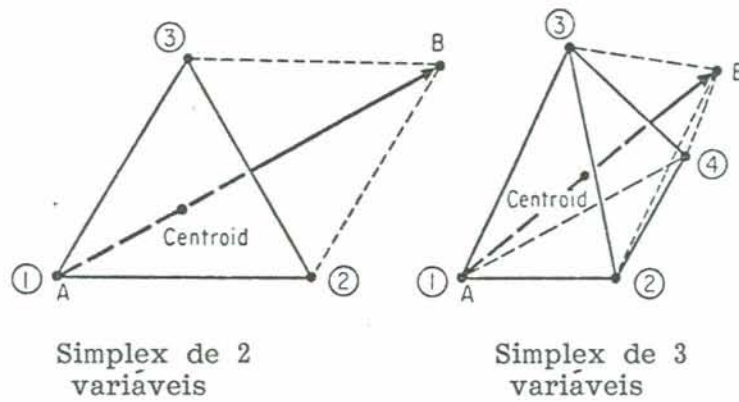
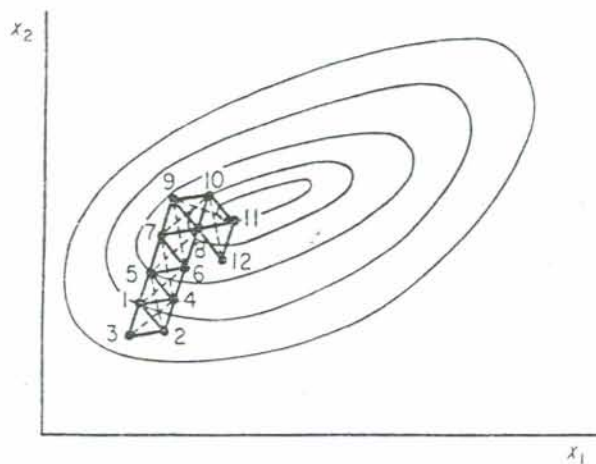


Fig. 13: Seqüência de simplexes regulares obtidos na minimização de $f(x)$. O traçado indica a projeção.



Os detalhes deste algoritmo são os seguintes :

Seja $x_i^k = [x_{i1}^k, \dots, x_{ij}^k, \dots, x_{in}^k]^T$, $i = 1, \dots, n+1$, o i -ésimo vértice (ponto) no E^n , no k -ésimo estágio da busca, $k = 0, 1, \dots$, e seja o valor da função objeto em x_i^k igual a $f(x_i^k)$. Necessita-se nomear os vetores x no poliedro que dão o valor máximo e o valor mínimo de $f(x)$.

Define-se :

$$f(x_h^k) = \max \{ f(x_1^k), \dots, f(x_{n+1}^k) \}$$

com o correspondente $x_i^k = x_h^k$, e

$$f(x_l^k) = \min \{ f(x_1^k), \dots, f(x_{n+1}^k) \}$$

com o correspondente $x_i^k = x_l^k$. Como o poliedro em E^n é feito de $(n+1)$ vértices, x_1, \dots, x_{n+1} , faça-se que x_{n+2} seja o centróide de todos os vértices excluindo-se x_h . As coordenadas da centróide são dadas por :

$$x_{n+2,j}^k = \frac{1}{n} \left[\left[\sum_{i=1}^{n+1} x_{ij}^k \right] - x_{hj}^k \right] \quad j = 1, \dots, n$$

(1)

onde o índice j designa a direção de cada coordenada.

O poliedro inicial é geralmente selecionado como sendo um 'simplex' regular (não precisa necessariamente ser), com o ponto 1 sendo a origem, ou talvez a centróide como a origem. O procedimento para encontrar um vértice em E^n no qual a $f(x)$ tem um valor melhor envolve quatro operações.

1. Reflexão. Reflete x_h^k através do centróide computando-se:

$$x_{n+3}^k = x_{n+2}^k + \alpha (x_{n+2}^k - x_h^k)$$

onde $\alpha > 0$ é o coeficiente de reflexão,

x_{n+2}^k = centróide computada por (1),

x_h^k = vértice no qual $f(x)$ é o maior valor de $(n+1)$ da $f(x)$ no

k-ésimo estágio.

2. Expansão. Se $f(x_{n+3}^k) \leq f(x_l^k)$, expande o vetor $(x_{n+3}^k - x_{n+2}^k)$ computando-se :

$$x_{n+4}^k = x_{n+2}^k + \gamma (x_{n+3}^k - x_{n+2}^k)$$

onde $\gamma > 1$ é o coeficiente de expansão. Se $f(x_{n+4}^k) < f(x_l^k)$, substitue-se x_h^k por x_{n+4}^k e continua-se do passo 1 com $k=k+1$. Por outro lado, substitue-se x_h^k por x_{n+3}^k e continua-se do passo 1 com $k=k+1$.

3. Contração. Se $f(x_{n+3}^k) > f(x_l^k)$ para todos i diferentes de h , contrae-se o vetor $(x_h^k - x_{n+2}^k)$ computando-se

$x_{n+5}^k = x_{n+2}^k + \beta (x_h^k - x_{n+2}^k)$ onde $0 < \beta < 1$ é o coeficiente de contração. Substitue-se x_h^k por x_{n+5}^k e retorna-se ao passo 1 para continuar a busca no $(k+1)$ -ésimo estágio.

4. Redução. Se $f(x_{n+3}^k) > f(x_h^k)$, reduz-se todos os vetores $(x_i^k - x_l^k)$, $i = 1, \dots, n+1$, por uma metade de x_l^k computando-se

$x_i^k = x_l^k + 0.5 (x_i^k - x_l^k)$ $i = 1, \dots, n+1$ e retorna-se ao passo 1 para continuar a busca no $(k+1)$ -ésimo estágio.

O critério usado por Nelder Mead para terminar a busca foi o teste que determina se:

$$\left[\frac{1}{n+1} \sum_{i=1}^{n+1} [f(x_i^k) - f(x_{n+2}^k)]^2 \right]^{1/2} \leq \epsilon$$

onde ϵ é um número arbitrariamente menor, e $f(x_{n+2}^k)$ é o valor da função objeto no centróide x_{n+2}^k .

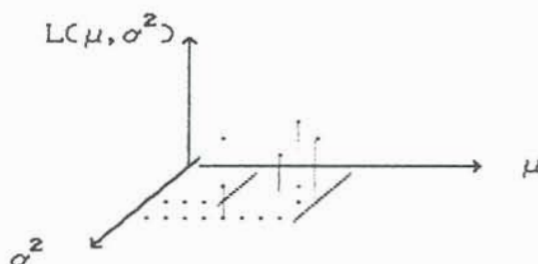
Nelder e Mead recomendam que os valores de $\alpha = 1$, $\beta = 0.5$ e $\gamma = 2$ são valores geralmente satisfatórios para a minimização.

Dentro do Matlab tem-se a função Nelder Mead, a qual minimiza

uma função não linear de várias variáveis. Trocando-se o sinal da função logarítmica da verossimilhança, espera-se que este algoritmo possa maximizá-la ao invés de minimizá-la. O Nelder Mead inicia com um vetor (da média e variância inicialmente estimadas) e produz novos vetores sucessivamente até que se maximize a função log da verossimilhança. Espera-se portanto que este novo vetor seja tal que quando aplicado a função log-normal truncada, permita que esta última se ajuste a 'waveform'.

Concluída esta etapa pode-se partir para a construção do gráfico da função logarítmica de verossimilhança. Para construir este, tomou-se os valores da média e variância que maximizaram a função log da verossimilhança como ponto central, e a partir deles, se construiu uma matriz dez por dez, sendo cada elemento desta, um vetor dois por um, de média e variância. Com estes vetores armazenados em uma matriz e lançando mão a funções gráficas do matlab, construiu-se um gráfico tridimensional da função, calculando-se para cada elemento da matriz, o valor da função log da verossimilhança.

Isto pode ser melhor ilustrado na figura a baixo.



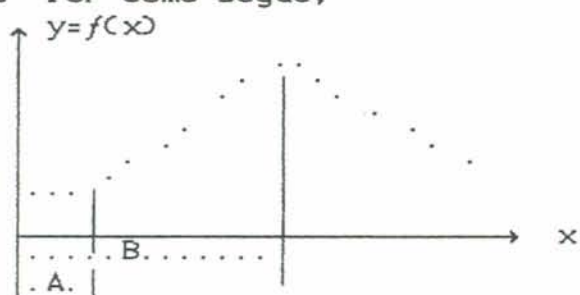
3.3.2.3. DEFINIÇÃO DA FUNÇÃO LOG-NORMAL TRUNCADA

Como fora anteriormente definido, ao encontrar-se os valores μ e σ , maximizadores da log da verossimilhança, aplicam-se estes

mesmos à fórmula da função log-normal truncada, definindo-a e, procede-se a sua plotagem.

Espera-se que esta função se ajuste ao comportamento das 'waveforms', de modo que a partir dela, se possa avaliar parâmetros importantes, tais como o tempo de viagem do pulso do radar à superfície d'água.

Se o comportamento da função log-normal truncada para as 'waveforms' for como segue,



pode-se a partir desta função, procurar os valores das abscissas A e B, correspondentes aos pontos de curvatura máxima da função. A abscissa A define o tempo de viagem ida-e-volta do pulso deste o radar até a superfície d'água.

3.3.2.4. MAXIMIZAÇÃO DA CURVATURA

Uma das maneiras propostas para maximizar a curvatura, isto é, encontrar os pontos onde sua curvatura é máxima, foi a definição dos pontos onde a função tem o seu raio de curvatura mínimo (curvatura máxima) através de fórmulas que definem estes parâmetros em função de derivadas primeira e segunda da função.

A curvatura 'K' num ponto (x,y) é $K=1/R$.

'R' é o raio de curvatura para qualquer ponto de $y=f(x)$.

s é o comprimento de arco medido ao longo da curva $y=f(x)$ a partir de um ponto fixo até qualquer outro ponto $P(x,y)$.

α é o ângulo de inclinação da linha tangente em P.

$$\frac{dx}{dy} = \cos\alpha = \frac{1}{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}$$

$$\frac{dy}{dx} = \text{sen}\alpha = \frac{1}{\sqrt{1 + \left(\frac{dx}{dy}\right)^2}}$$

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy$$

$$R = \frac{ds}{d\alpha} = \frac{\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{3/2}}{dy^2/dx^2} = \frac{\left[1 + (f'(x))^2\right]^{3/2}}{f''(x)}$$

Para resolver esta fórmula de R, ao invés de calcular as derivadas primeira e segunda da função log-normal truncada, optou-se por uma aproximação numérica, capaz de avaliar estas derivadas.

Para melhor ilustrar este procedimento, suponha-se f como sendo o valor da função no ponto médio de cada um dos sessenta e três intervalos da 'waveform'. Seja Δ , o tamanho de cada intervalo. Assim as derivadas serão dadas por :

$$\frac{df}{dx} = \frac{f(i+1) - f(i-1)}{2\Delta}$$

$$\frac{df^2}{dx^2} = \frac{f(i+1) - 2f(i) + f(i-1)}{2\Delta}$$

Assim, um programa pode avaliar as derivadas em cada intervalo e com elas, calcular os raios de curvatura, encontrando-se conseqüentemente os dois menores, correspondentes as duas curvaturas maiores da função log-normal truncada.

A cada um destes pontos de curvatura máxima da log-normal truncada, corresponde um valor de frequência estatística (da waveform), e o valor de sua abscissa, a qual está associada ao tempo

de viagem de ida e volta do pulso deste o radar até o nível d'água.

IV. RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados alguns exemplos de 'waveforms', às quais tentou-se ajustar a função log-normal truncada e seus resultados, através da metodologia anteriormente apresentada. Para uma melhor exemplificação do procedimento sugere-se a observância do apêndice A, onde encontra-se um diagrama de blocos.

Tomou-se a 'waveform' número 40433 da trilha trk45, sobre a Lagoa dos Patos, a qual foi plotada após a reordenação de suas frequências estatísticas e divisão destas pela sua soma total. Veja tabela 5 e figura 14 deste capítulo. No eixo das ordenadas tem-se os valores de cada frequência estatística em termos relativos, e no eixo das abscissas, tem-se os sessenta e três 'gates' correspondentes aos intervalos previamente definidos.

Tabela 5: 'waveform' V40433, escrita sob forma de vetor.

v40433=[26; 22; 23; 18; 18; 18; 16; 16; 18; 20; 16; 16;
18; 14; 14; 18; 14; 18; 18; 22; 16; 18; 16; 19;
19; 21; 25; 23; 30; 83; 264; 317; 339; 359; 304; 293;
333; 324; 383; 373; 305; 372; 339; 331; 378; 381; 302; 356;
325; 319; 334; 361; 445; 342; 343; 350; 355; 303; 307; 286;
46; 165; 320]

Para esta 'waveform' calculou-se as estimativas iniciais da média e variância, cujo os valores resultantes foram respectivamente :

$$\hat{\mu} = 0.4095417 \text{ e}$$

$$\hat{\sigma} = 0.2070859$$

Esta média e variância estão calculadas para o logaritmo natural dos intervalos de frequência da tabela 4, de modo que eles

estão contidos dentro do intervalo que varia de 0.475 a 6.525, correspondendo a média ao 'gate' número doze aproximadamente.

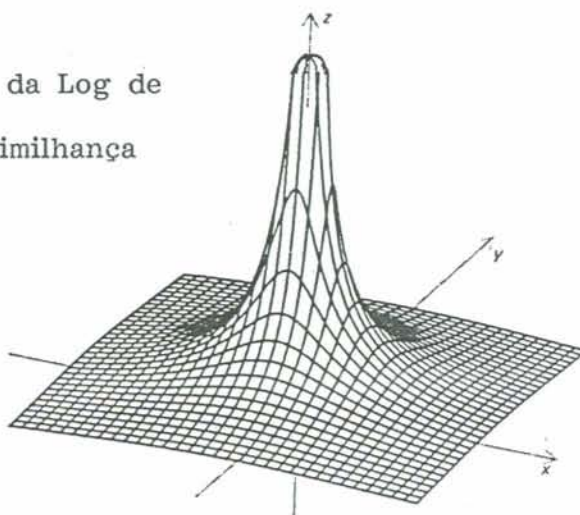
A partir destes valores iniciais, aplicou-se a rotina Nelder Mead, capaz de maximizar a função logarítmica de verossimilhança. Os valores de μ e σ que maximizam esta função para esta 'waveform' foram respectivamente :

$$\mu = 5.2972981 \text{ e}$$

$$\sigma = 1.1826490$$

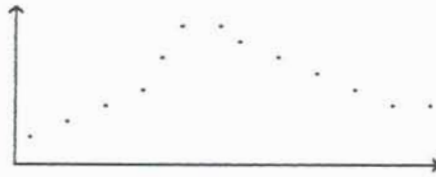
Na figura 15 pode-se observar o gráfico da função log da verossimilhança para esta 'waveform', calculado gerando-se uma grade a partir do ponto cuja a média e variância maximizam a função. Nota-se que seu comportamento diferiu do esperado que deveria ser :

Gráfico da Log de
Verossimilhança



Com estes valores maximizadores partiu-se para a plotagem e cálculo da função log-normal truncada, na expectativa que esta pudesse se ajustar a 'waveform'. Na figura 16, observa-se esta função, a qual foi calculada para cada abscissa que está associada a uma das sessenta e três freqüências da 'waveform'.

Percebe-se que esta função não se comportou como o esperado, ou seja, para que ela estivesse se ajustando a 'waveform', deveria ter um comportamento do tipo :



Uma alternativa para o cálculo e plotagem da função log-normal truncada foi o uso das estimativas iniciais da média e variância, ao invés do uso de seus valores maximizadores. Neste caso o resultado continuou em desacordo com o esperado. Veja figura 17.

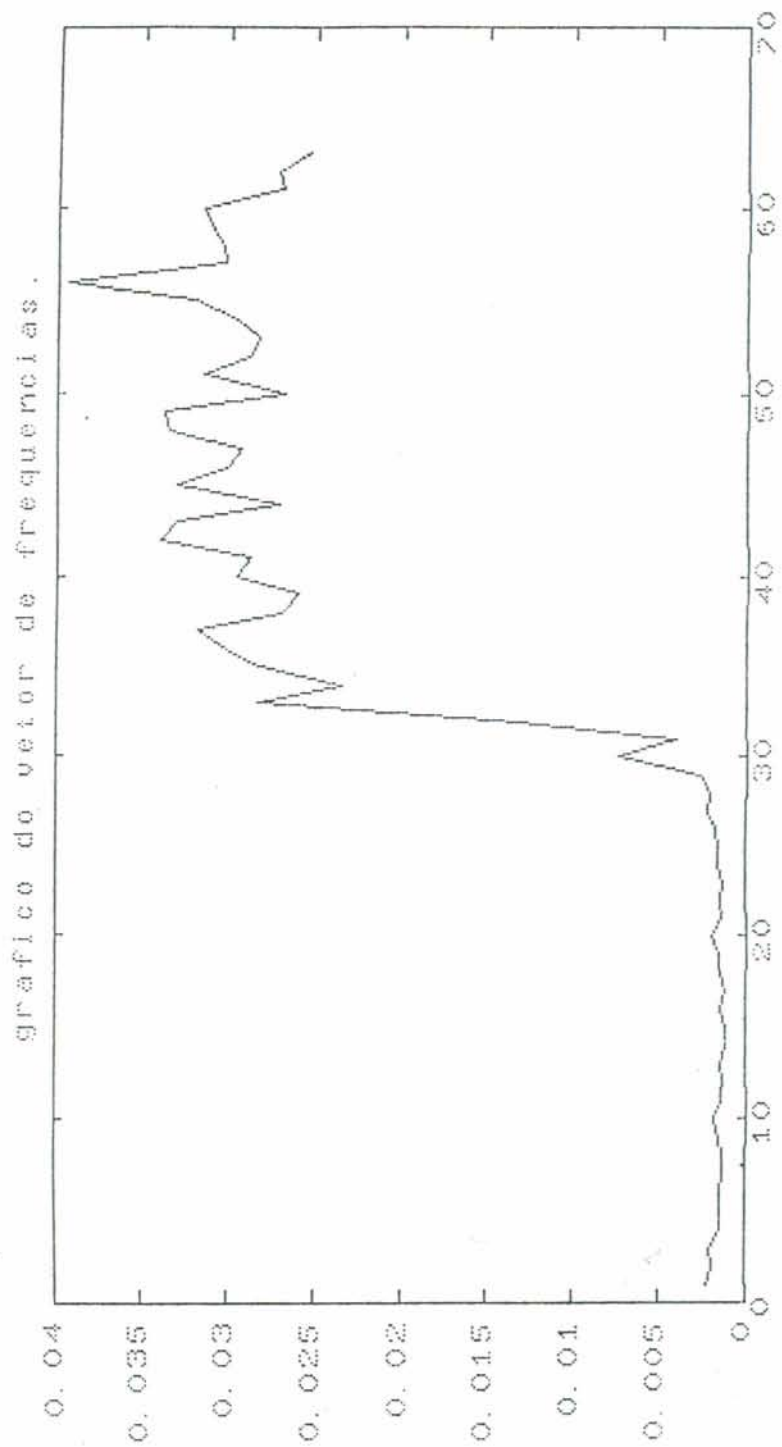


Fig.14 - Gráfico da waveform V40433

grafico da log da verossimilhanca

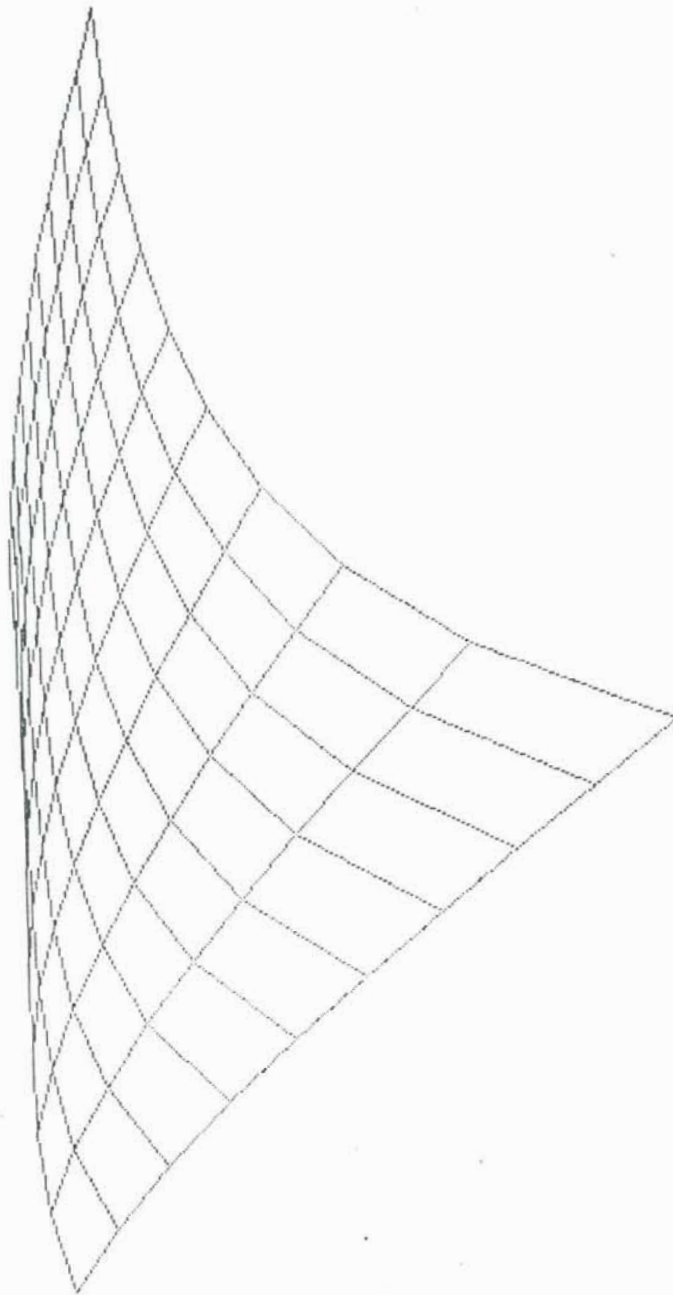


Fig.15 - Gráfico da log de Verossimilhança de V40433

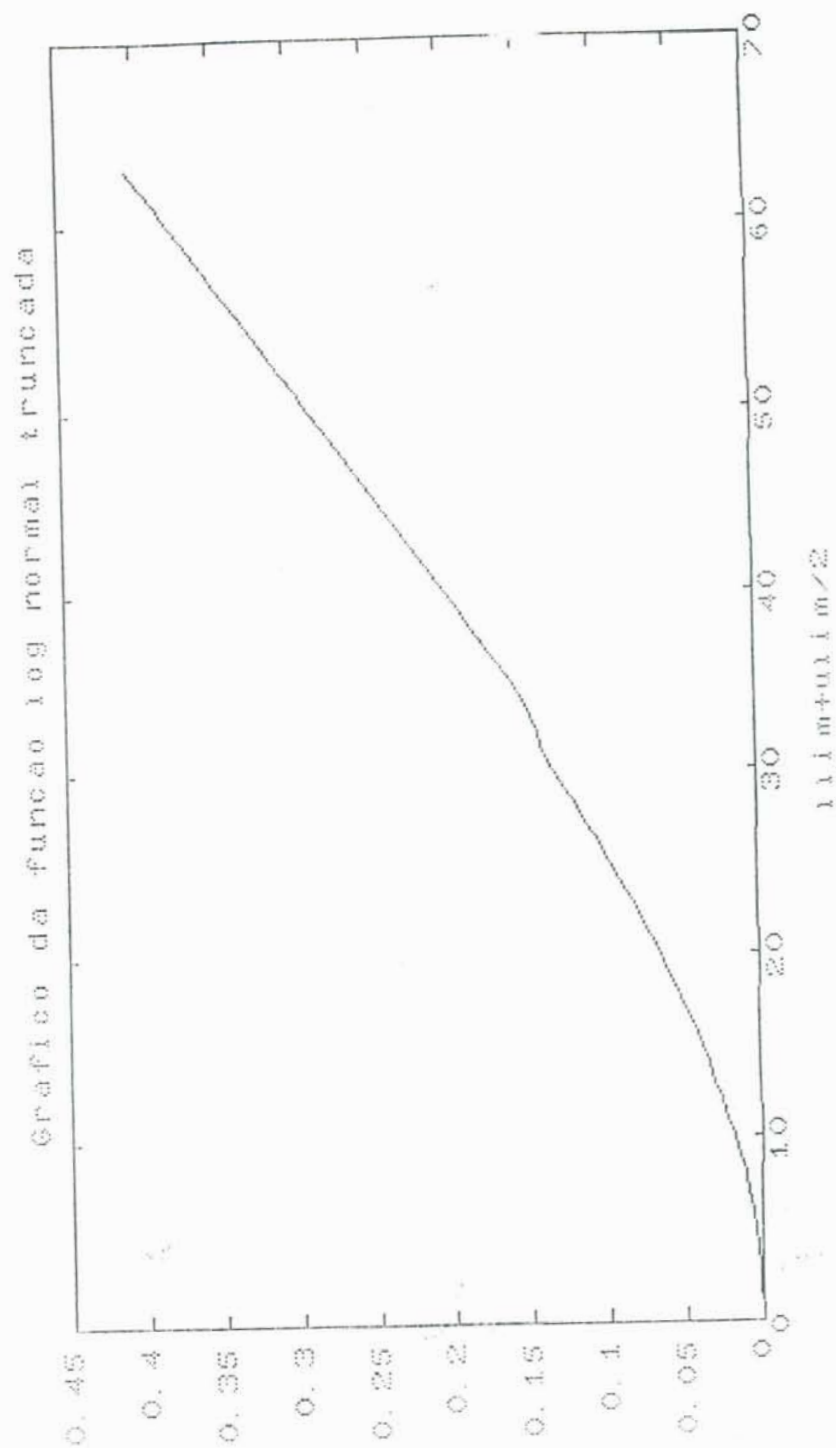


Fig.16 - Gráfico da log normal truncada

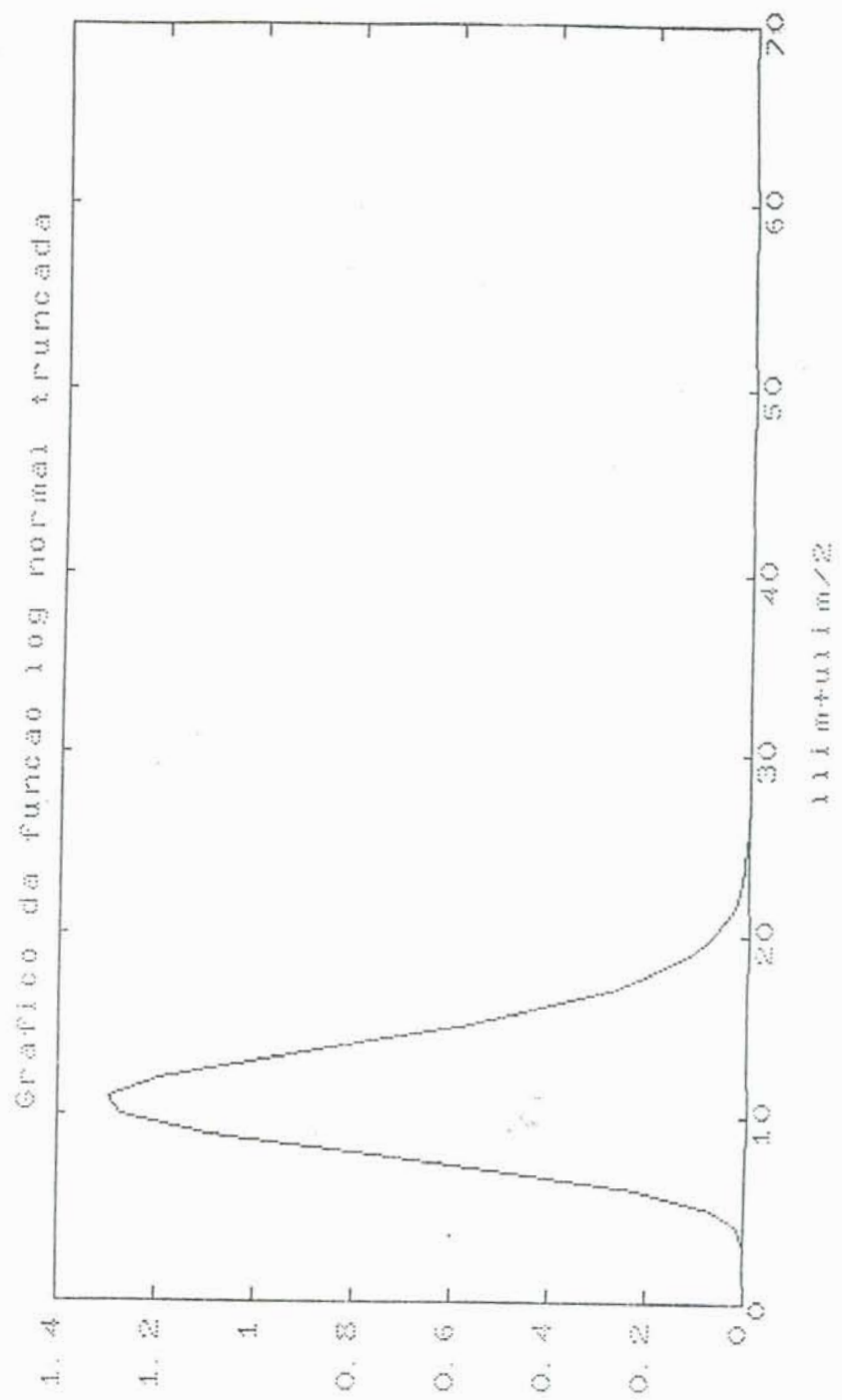


Fig.17 - Gráfico da log-normal truncada

Nestas primeiras tentativas os resultados não foram consonantes com o que se esperava. Mesmo assim aplicou-se o algoritmo que deveria encontrar o ponto de curvatura máxima da log-normal truncada. No caso da log-normal truncada calculada com os valores de μ e σ que maximizaram a log da verossimilhança, os pontos de curvatura máxima correspondem aos 'gates' 59 e 60. Estes valores de curvatura foram calculados para o ponto média de cada intervalo das frequências por meio das derivadas anteriormente apresentadas. Já os pontos de curvatura máxima para a função log-normal truncada calculada com as estimativas iniciais de μ e σ , corresponderam aos 'gates' 7 e 32.

Outro exemplo entre os analisados é o da 'waveform' número 40428, também sobre a Lagoa dos Patos.

Tabela 6: 'Waveform' v49428.

v49428=[24; 15; 14; 20; 14; 14; 14; 16; 16; 14, 14; 18;
 14; 12; 12; 14; 16; 18; 14; 18; 14; 16; 14; 16;
 17; 15; 19; 17; 29; 95; 44; 178; 337; 279; 346; 303;
 283; 361; 407; 345, 344; 304; 330; 386; 322; 327; 357; 288;
 397; 324; 306; 305; 347; 365; 337; 296; 383; 366; 335; 398;
 351; 345; 347]

Procedeu-se a reordenação das frequências e imediata plotagem deste gráfico, como pode ser visto na figura 18.

Para esta 'waveform' as estimativas de média e variância foram respectivamente :

$$\hat{\mu} = 0.4158048$$

$$\hat{\sigma} = 0.1966554$$

Com estes valores iniciais aplicando-se a rotina Nelder Mead, os valores maximizadores da função log de verossimilhança encontrados foram :

$$\mu = 3.7326059$$

$$\sigma = 0.9128292$$

A plotagem da função log da verossimilhança está na figura 19. Partiu-se então, para a plotagem da função log-normal truncada para esta 'waveform' com os valores de μ e σ que maximizam a log de verossimilhança. Neste caso também, como se vê na figura 20, não se alcançou o ajuste desejado. Plotando-se esta mesma função com a média e variância estimadas, tem-se o gráfico da figura 21 como resultado. Para este caso também calculou-se os pontos de curvatura máxima, embora os resultados até então não justificassem pois simplesmente não se tem alcançado o ajuste esperado, que tornaria possível avaliar importantes parâmetros da 'waveform' mais facilmente, quando comparado ao método de ajuste da NASA.

Observou-se que para as 'waveforms' 40433 e 40428, os gráficos da log-normal truncada com média e variância estimadas, forneceram resultados mais próximos do esperado, quando esta média e variância são calculadas como sendo simplesmente :

$$\hat{\mu} = \frac{\sum f_i x_i}{\sum f_i} \quad e \quad \hat{\sigma} = \frac{\sum f_i (x_i - \hat{\mu})^2}{\sum f_i}$$

Isto pode ser observado nas figuras 22 e 23.

Experimentou-se o ajuste a outras 'waveforms', obtendo-se resultados não satisfatórios, de modo que a tentativa de ajustar uma 'waveform' a uma função log-normal truncada, até então, se mostrou inviável, ou pelo menos, seus resultados não provaram ser eficientes.

O modelo estatístico apresentado, poderia estar sujeito a modificações e incrementos, dos quais poder-se-ia atingir resultados dentro das expectativas. Não cabe nesta dissertação desenvolver outros modelos, mas verificar o comportamento do modelo apresentado.

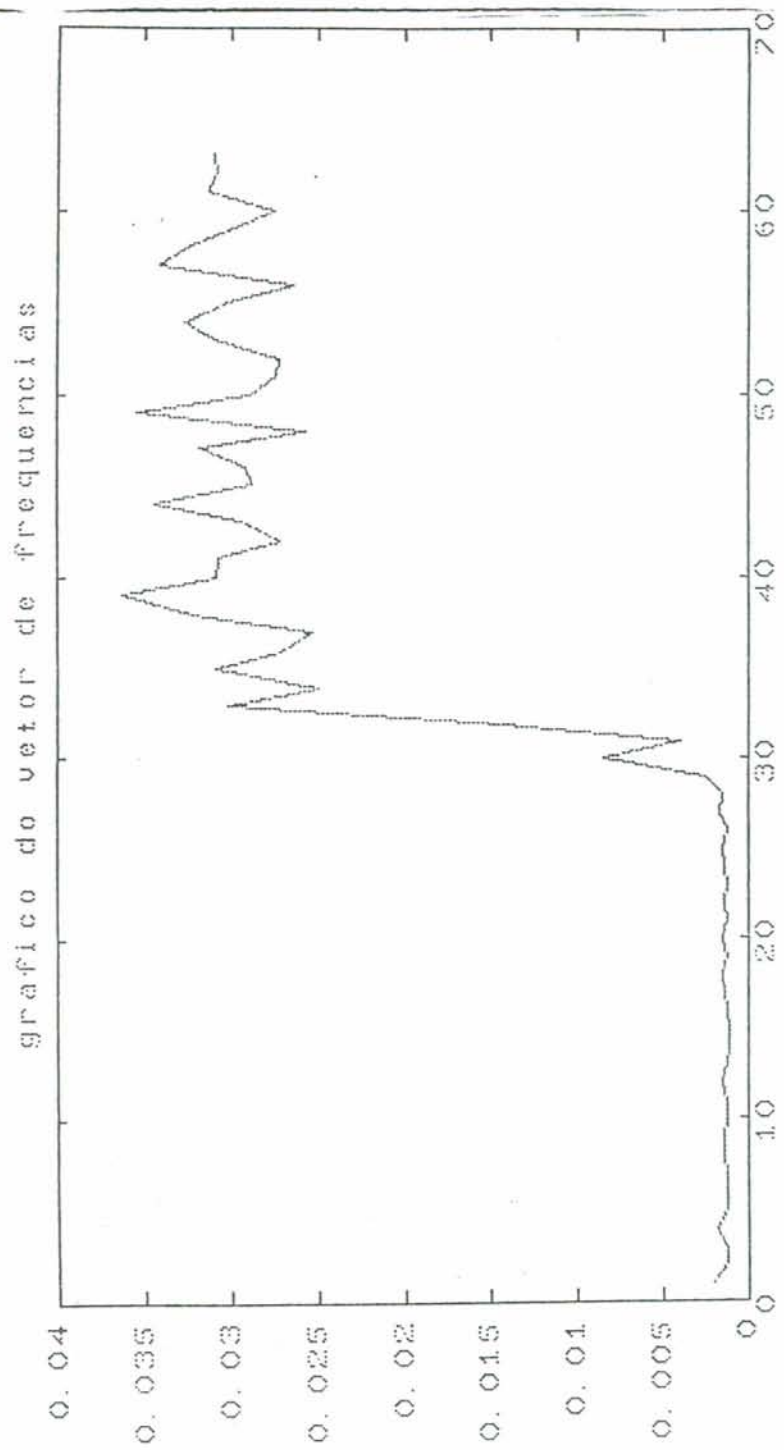


Fig.18 - Waveform V40428 e gráfico

grafico da log da verossimilhanca

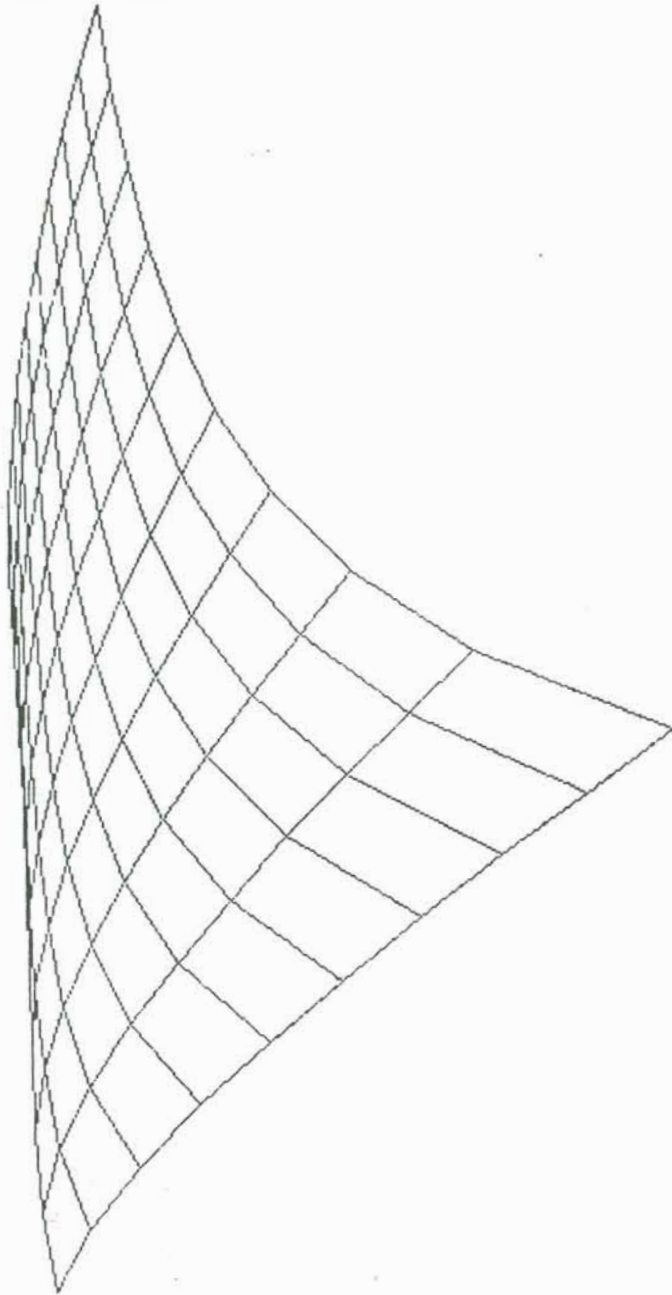


Fig.19 - Plotagem da log de verossimilhanca

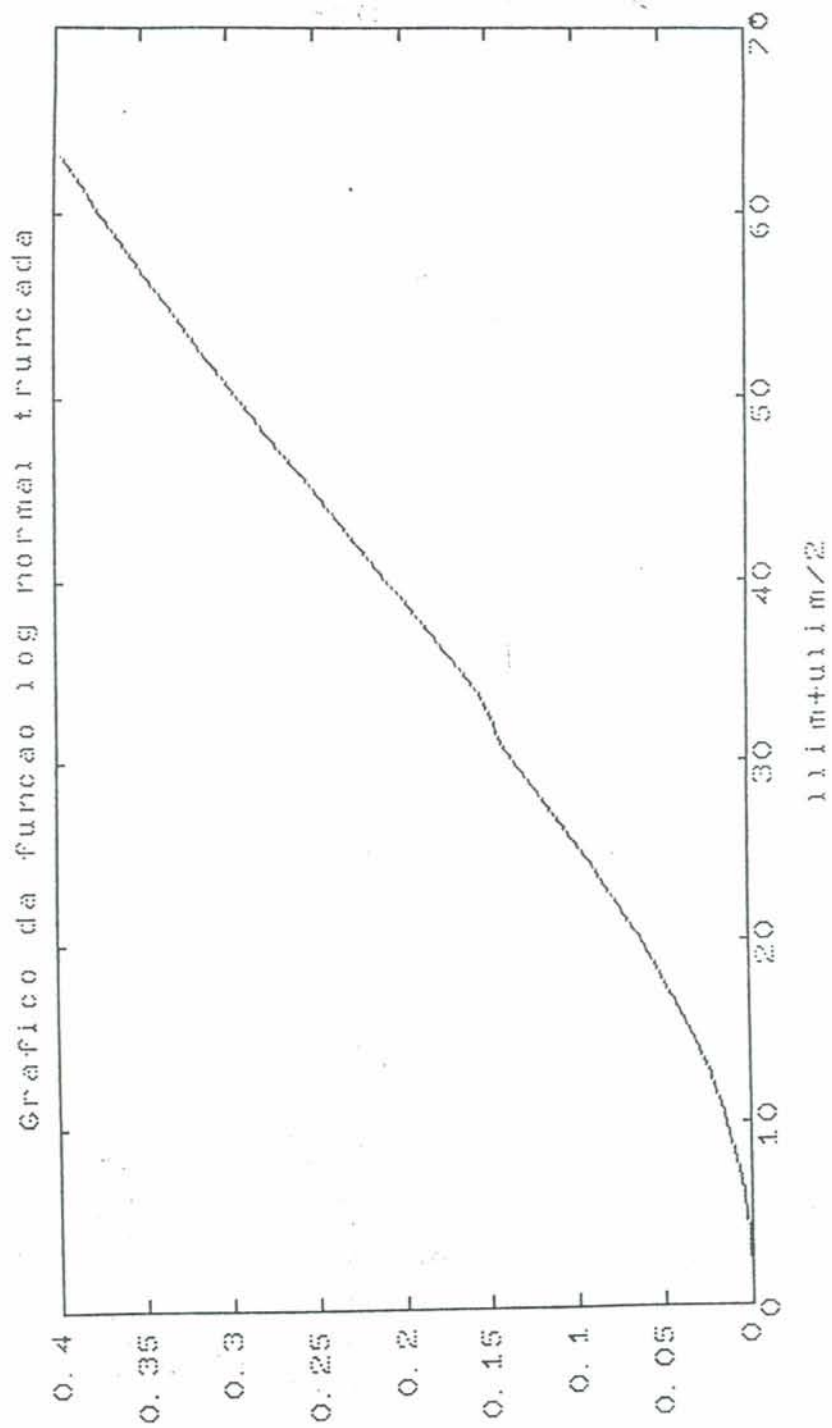


Fig.20 - Plotagem da log-normal truncada para V40428

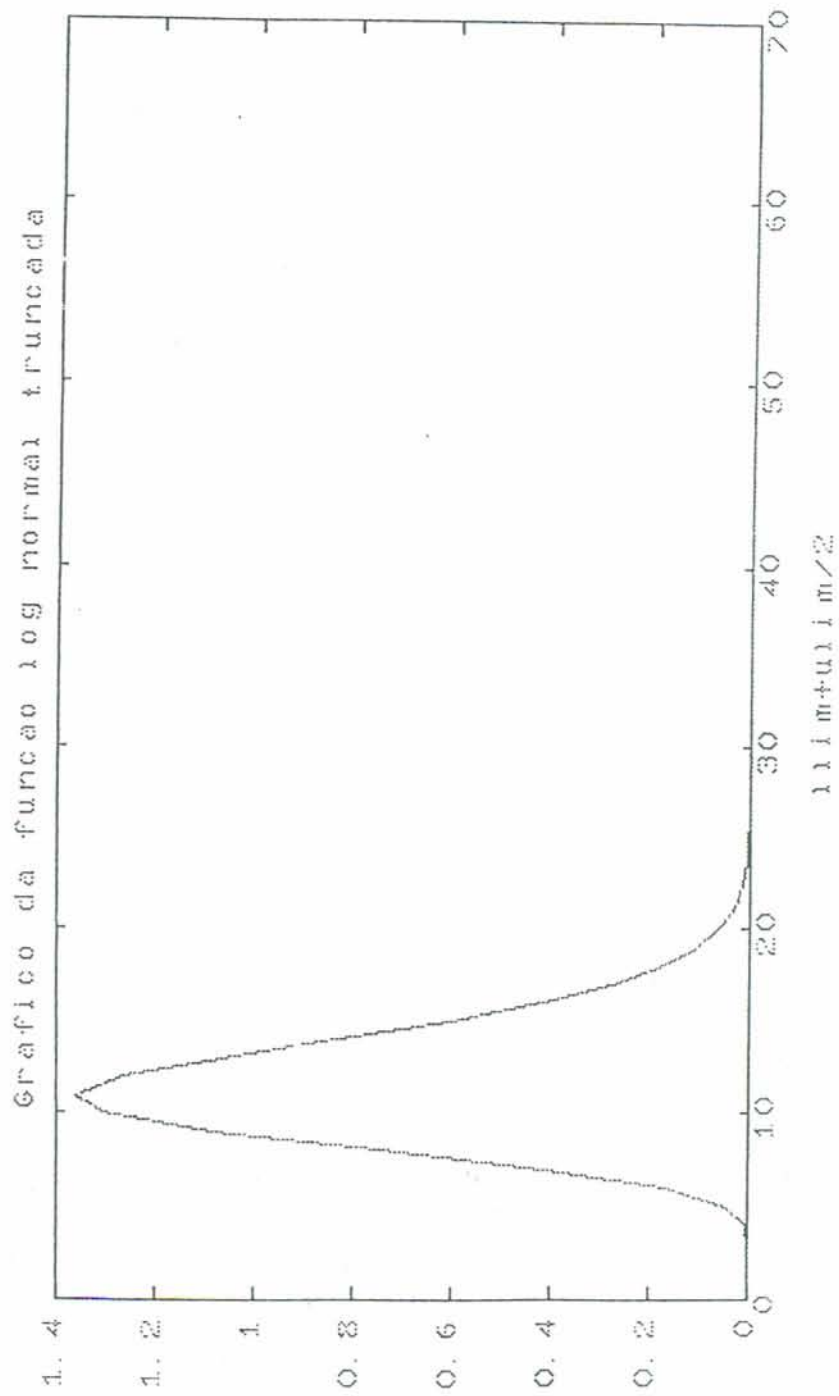


Fig.21 - Plotagem da log-normal truncada

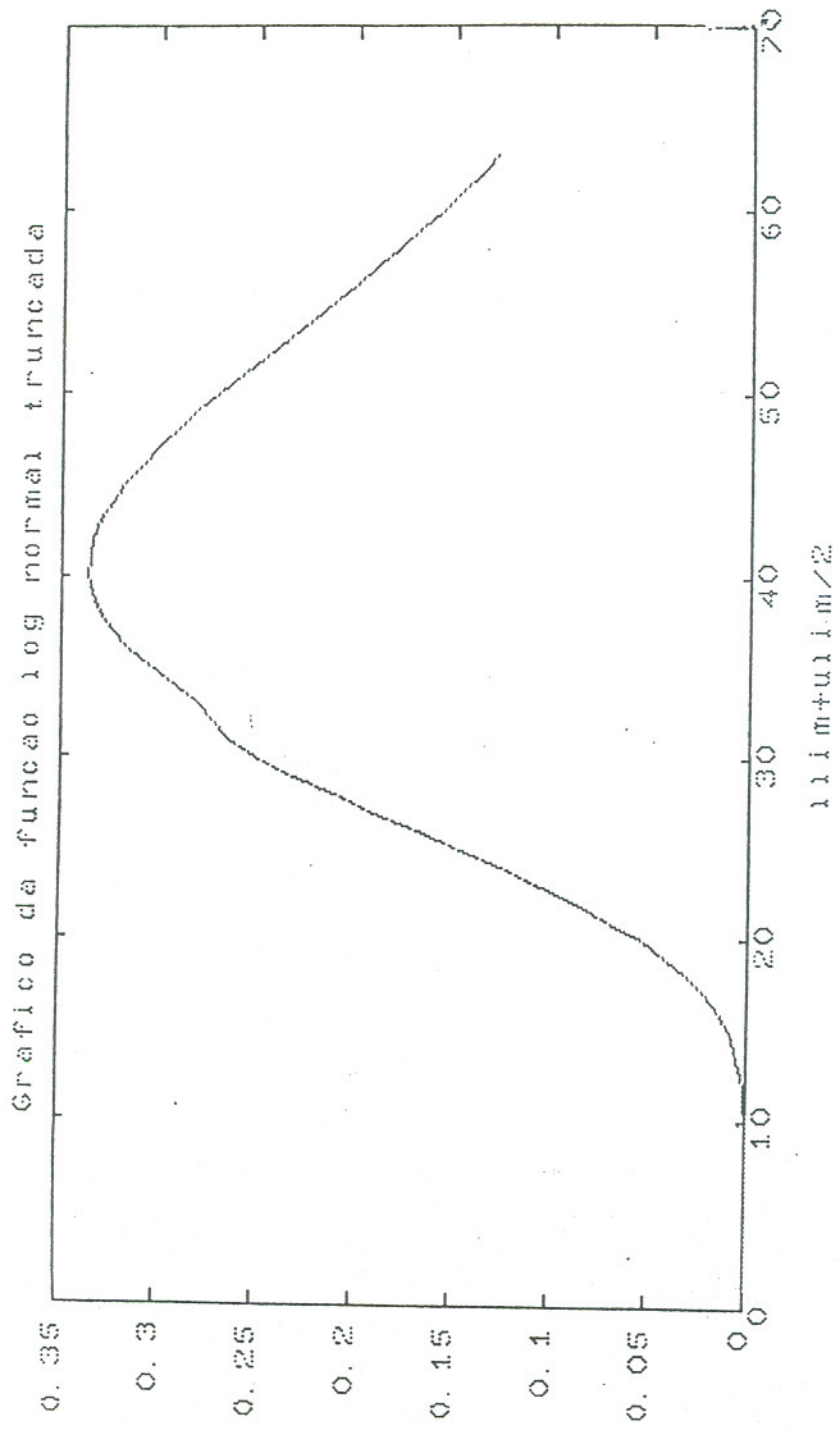


Fig.22 - Gráfico da log-normal

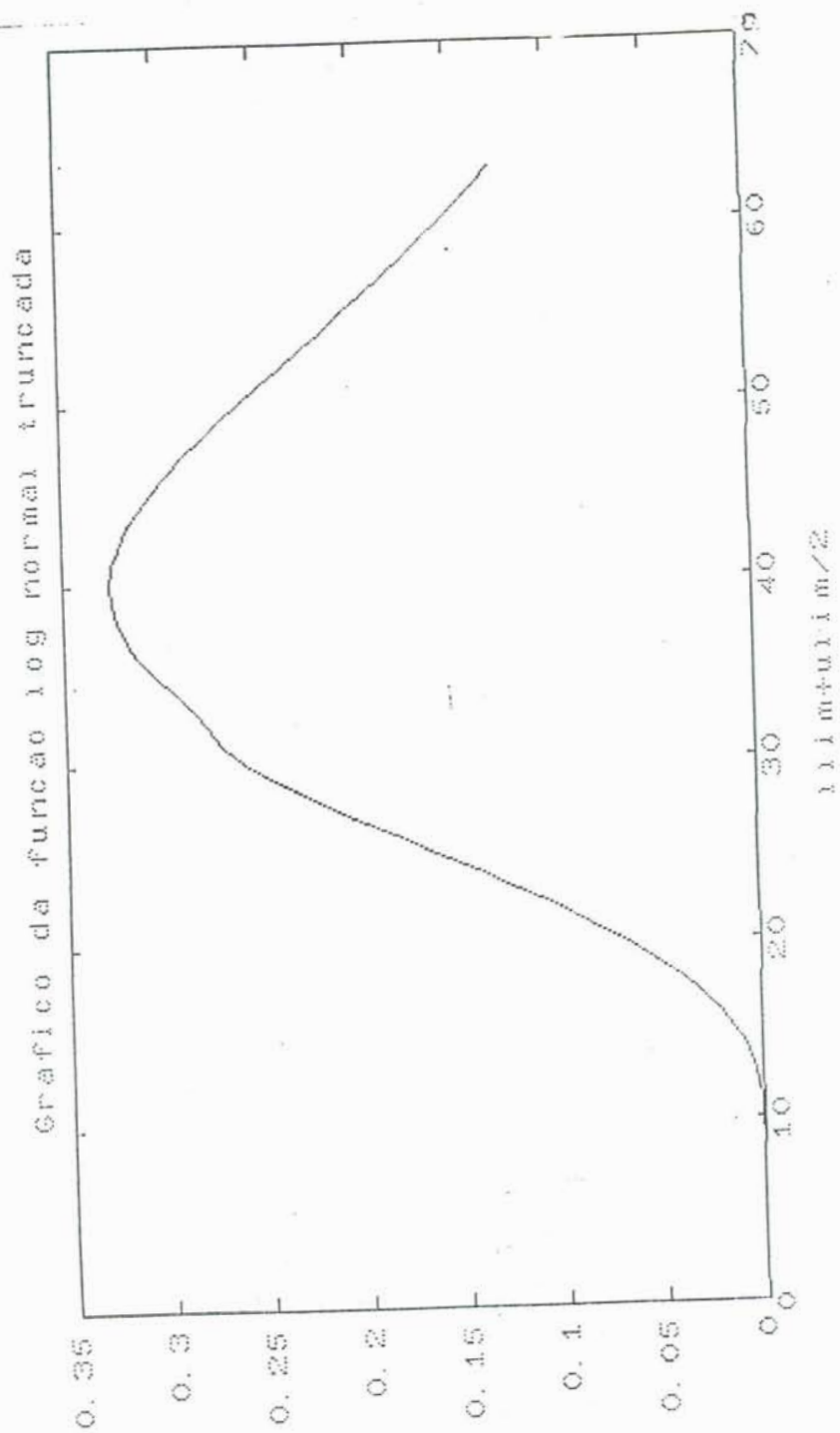


Fig.23 - Gráfico da função log-normal

V. CONCLUSÕES

A técnica que visa utilizar dados de satélites altimétricos e bem aplicá-los à avaliação e monitoramento de recursos hídricos, se coloca como uma ferramenta de grande potencialidade, mas requer no entanto, muitos estudos para o seu aproveitamento.

As primeiras medições altimétricas com precisão e duração suficiente datam de 1985, e só a partir delas se pôde começar a avaliar de fato a aplicabilidade destes dados em águas continentais. Alguns pesquisadores têm voltado bastante atenção para esta matéria, a fim de estabelecer certos parâmetros que permitam definir o nível de precisão da altimetria em grandes bacias hidrográficas. Em termos de Brasil, o desenvolvimento desta técnica e sua conseqüente operacionalidade, constitui um avanço considerável no que concerne medições hidrométricas. Trata-se de mais uma técnica que uma vez associada a outras, possa permitir com economia de tempo e de recursos financeiros, um controle e respectivo aproveitamento mais positivo das potencialidades hidrológicas do país.

Outros satélites têm sido colocados em órbita com instrumentos altimétricos mais aperfeiçoados e existem projetos em andamento. Isto faz crer que, embora ainda um tanto limitada, a altimetria pode vir a se constituir uma ferramenta bastante operacional. Satélites como o ERS-1, já em órbita, ou o Topex/Poseidon, são exemplos disto.

Embora a tentativa deste trabalho de estabelecer uma metodologia capaz de interpretar os retornos altimétricos ('waveforms') utilizando-se dois parâmetros ao invés de nove, tenha se mostrado mais complicada que a metodologia já existente, constituiu-se em mais um passo, no universo de pesquisas que se

impõe a fim de tornar a altimetria uma ferramenta eficiente. O ajuste matemático proposto não se adequou, mostrando-se um caminho ineficiente de análise dos dados oriundos do GEOSAT.

Sem dúvida, outros processamentos, análises, ajustes e estudos estão em aberto a fim de permitir uma resposta mais imediata e computacionalmente menos longa, no que diz respeito aos cálculos necessários para se obter respostas dos dados altimétricos. A expectativa de poder, com o ajuste sugerido, determinar com maior rapidez o tempo de viagem do pulso do radar, embora não alcançada, abre caminho para outras abordagens matemáticas e descarta a aplicação desta sugerida.

A análise destes dados altimétricos se constitui também numa abordagem diferenciada do sensoriamento remoto, na qual o estudo e as conclusões com relação aos alvos e suas propriedades, são feitos através da detecção, registro e análise do pulso do radar refletido da superfície, não limitando-se esta técnica, ao tratamento de imagens simplesmente.

Não resta dúvida que, embora muitas tentativas se façam necessárias para que se tenham repostas e resultados positivos e satisfatórios, o uso de medidas de nível de águas continentais a partir de satélites se coloca como uma alternativa técnica promissora.

VII. BIBLIOGRAFIA

- ABRAMOWITZ, Milton. Handbook of Mathematical Functions, with Formulas, Graphs and Mathematical Tables. Dover Publications, Inc., New York.
- ALLAN, T.D.. Satellite Microwave Remote Sensing, 1983. Ellis Horwood Limited.
- BILHAM, Roger. Earthquakes and Sea Level: Space and Terrestrial Metrology on a Changing Planet. REVIEWS OF GEOPHYSICS, 29. 1. February 1991.
- BRENNER, A.C., FREY, H.V. & ZWALLY, H.. Comparisons between Geosat and Seasat Tracking Over non Ocean Surfaces. Geophysical Research Letters, vol.17, No.10, p.1537-1540.
- BROOKS, Ronald D.L. and LOCKWOOD, Dennis W.. Effects of Islands in Geosat Footprint, Journal of Geophysical Research, Vol. 95, No.C3, March 15, 1990.
- BURLINGTON, Richard S.. Handbook of Mathematical Tables and Formulas. Fourth Edition, 1985, McGraw-Hill Book Company.
- CHELTON, Dudley B., WALSH, Edward J. & MACARTHUR, John L.. Pulse Compression and Sea Level Tracking in Satellite Altimetry, Journal of Atmospheric and Oceanic Technology, Vol.6, No.3, June 1989 (American Meteorology Society).

- CHELTON, Dudley B., SCHLAX, Michel L.G., WITTER, Dona L. & RICHMAN, James G.. Geosat Altimeter Observations of the Surface Circulation of the Southern Ocean. Journal of Geophysical Research, Vol.95, No.C10, October 15, 1990.
- CHENEY, Robert E., MILLER, Lany L., DOUGLAS, Bruce C. & AGREEN, Russel W.. Monitoring Equatorial Pacific Sea Level with Geosat. John Hopkins, APL Technical Digest. Vol.8, No. 2, 1987.
- DEVELOPMENTS in Inland Water and Land Altimetry. Final Report 1990. University College London. Mullard Space Science Laboratory.
- GUTTMAN, Irwin. Introductory Engineering Statistics. John Wiley & Sons Inc., 1985.
- HAAN, Charles T.. Statistical Methods in Hydrology. 1977. The Iowa State University Press.
- HIMMELBLAU, David.. Applied Nonlinear Programming. McGraw-Hill Book Company, 1972.
- KENDALL, Maurice G. & STUART, Allan.. the Advanced Theory of Statistics, Vol.2, 1967. Charles Griffin & Company Limited, London.
- KOBLINSKI, C.J., CLARKE, R.T., BRENNER, A.C. & FREY, H.. The Measurement of Amazon River Level Variations from Satellite Altimetry. NASA/Goddard Space Flight Center & IPH/UFRGS. April 16, 1992.
- KOBLINSKI, C.J., CLARKE, R.T., BRENNER, A.C.. Avanços no Uso de Dados Altimétricos Obtidos por Satélite para Aplicações Hidrométricas na Amazônia.

MACARTHUR, John L., MARTH, Paul C. & WALL, Joseph G.. The Geosat Radar Altimeter. John Hopkins APL, Technical Digest, Vol.8, No.2, 1987.

MEYER, Paul L.. Probabilidade: Aplicações e Estatística. 1969. Livro Técnico S.A. e Ed.da USP, Rio de Janeiro.

MILLER, I., FREUND, J. & JOHNSON, R.. Probability and Statistics for Engineers. 1990, Fourth Edition, Prentice Hall, New Jersey.

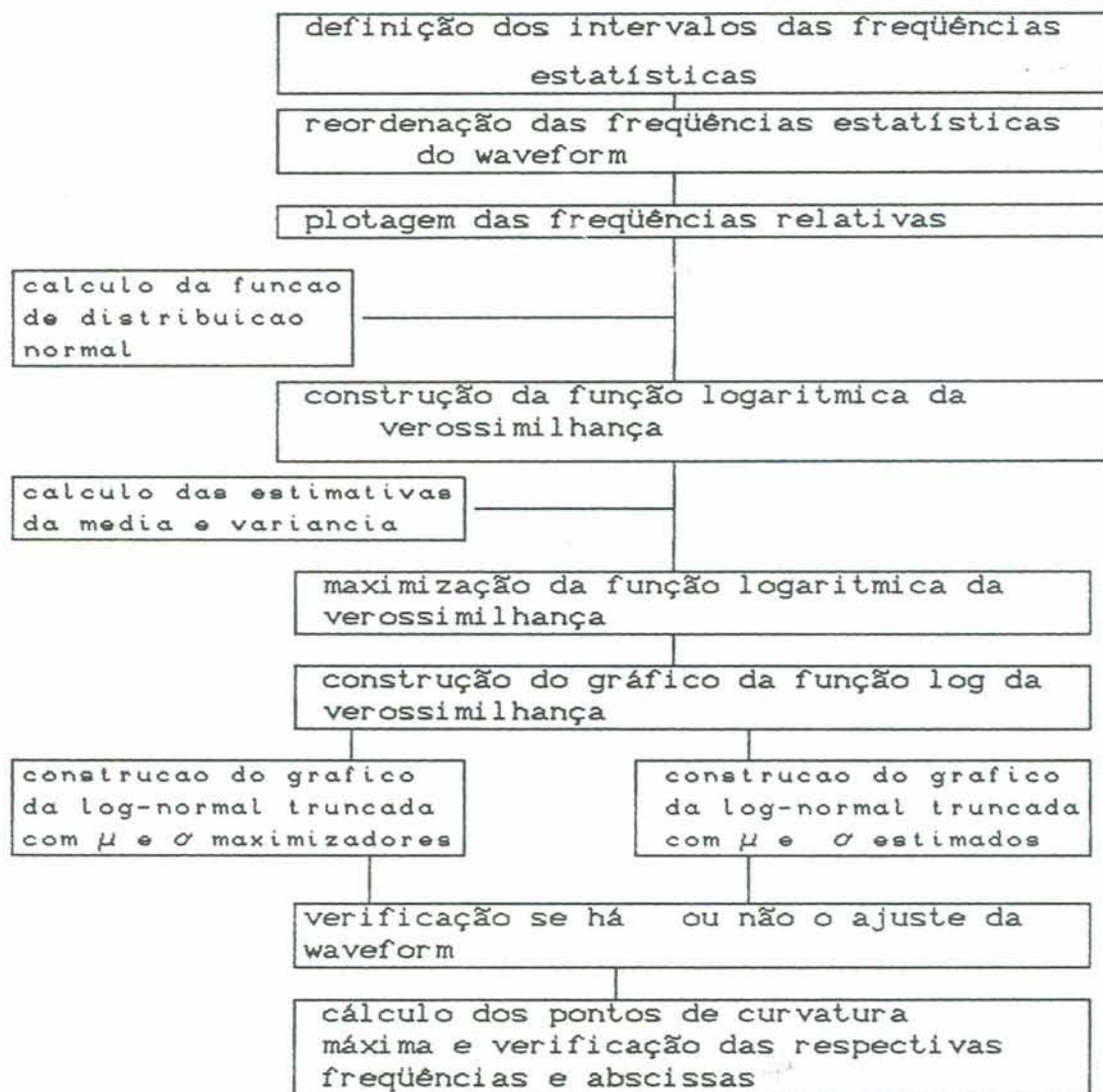
ROBINSON, I.S.. Satellite Oceanography: An Introduction for Oceanographers and Remote Sensing Scientists. Ellis Horwood Limited, 1985.

SATELLITE Radar Altimetry Over Ice, Vol.1: Processing and corrections of Seasat data over Greenland, July, 1989. NASA.

A N E X O

VII. APÊNDICES

VII.1. - APÊNDICE A



VII.2. - APÊNDICE B

Apresentar-se-á neste apêndice, o algoritmo utilizado pela NASA para o ajuste dos retornos difusos.

A teoria que resolve os parâmetros através dos mínimos quadrados Bayesianos utiliza-se das seguintes equações

É dado um conjunto de equações $MX=R$ onde :

M= a matriz das parciais
$$\begin{bmatrix} \partial c_1 / \partial \beta_1 & \dots & \partial c_1 / \partial \beta_n \\ \cdot & & \\ \cdot & & \\ \partial c_m / \partial \beta_1 & \dots & \partial c_m / \partial \beta_n \end{bmatrix} \quad m > n$$

X= vetor coluna =
$$\begin{bmatrix} \beta_{c1} - \beta_1 \\ \cdot \\ \cdot \\ \beta_{cn} - \beta_n \end{bmatrix}$$

R=
$$\begin{bmatrix} m_i & -c_i \\ \cdot & \cdot \\ m_m & -c_m \end{bmatrix}$$

m_i = valores observados ('counts' nos t='gate' i)

c_i = valores calculados de m_i baseados num dado conjunto de parâmetros β .

β_j = melhor estimativa corrente do parâmetro β do modelo.

β_{cj} = melhor estimativa corrigida do parâmetro β do modelo.

i = número do 'gate'

n = número de parâmetros na função.

Pode-se definir a matriz peso, W

$$W = \begin{bmatrix} w_{t1} & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & w_{tm} \end{bmatrix}$$

onde w_{ti} é o peso associado com cada observação i .

Se multiplicarmos ambos os lados da equação por W , obtemos,

$$WMX = WR$$

multiplicando-se por M^t tem-se,

$$M^t WMX = M^t WR.$$

A solução de X é dado por :

$$X = [M^t WM]^{-1} M^t WR$$

onde $M^t WM$ se refere como a matriz normal. Para se acrescentar informação quanto a validade da melhor estimativa corrente dos parâmetros do modelo, uma matriz de covariância V_0 é incluída.

$$V_0 = \begin{bmatrix} w_{t1}\beta_1 & 0 \\ 0 & w_{tm}\beta_n \end{bmatrix}$$

Onde $w_{tj}\beta_j$ = peso associado com um valor a priori de parâmetros j . Esta matriz é então acrescentada a matriz normal antes que ela seja invertida, de modo que a equação fica sendo :

$$X = [M^t WM + V_0]^{-1} M^t WR$$

X = é vetor o vetor dando a melhor nova estimativa dos parâmetros β .

A função representante do retorno altimétrico:

Foi mostrado por Miller e Brown, 1974, que o 'waveform' de retorno médio sobre uma superfície Gaussiana pode ser matematicamente descrito usando-se a função

$$c(t) = \beta_1 + \beta_2 \times P(W)$$

onde

$$P(W) = \int_{-\alpha}^{\alpha} Z(q) dq$$

$$Z(q) = \frac{1}{\sqrt{2\pi}} \exp(-q^2/2)$$

$$W = \frac{t - \beta_3}{\beta_4}$$

A função modificada usada para representar as waveforms do tipo difuso é escolhida como sendo :

$$c(t) = \beta_1 + \beta_2(1 + \beta_5 Q(x)) P(W), \text{ onde}$$

$$Q(x) = 0 \text{ para } t < \beta_3 + 0.5\beta_4$$

$$= t - x \text{ para } t > \beta_3 + 0.5\beta_4$$

veja a figura .

$$x = \beta_3 + 0.5\beta_4$$

As parciais desta função com relação a cada parâmetro são :

$$\frac{\partial c}{\partial \beta_1} = 1.0$$

$$\frac{\partial c}{\partial \beta_2} = P(W) + \beta_5 Q P(W)$$

$$\frac{\partial c}{\partial \beta_3} = \beta_2 \left[\frac{(1 + \beta_5 Q)}{\beta_4} \frac{\partial P}{\partial W} + \frac{P(W)\beta_5}{2} \right]$$

$$\frac{\partial c}{\partial \beta_4} = \beta_2 \left[\frac{(1 + \beta_5 Q)}{\beta_4} \frac{\partial P}{\partial W} W + \beta_5 \frac{P(W)}{2} \right]$$

$$\frac{\partial c}{\partial \beta_5} = \beta_2 Q P(W)$$

onde

$$\frac{\partial P}{\partial W} = \frac{1}{\sqrt{2\pi}} \exp(-W^2/2)$$

O valor de β_3 é o ponto médio da 'leading edge', Gm. Como já fora observado, alguns retornos têm múltiplos 'leading edges'. Uma função de nove parâmetros é usada para representar estes retornos, onde o ponto médio do primeiro 'leading edge' permanece sendo β_3 . O

ponto médio do segundo 'leading edge', β_6 , provavelmente represente um retorno de uma outra superfície no solo, e é armazenada para uso futuro. A função de nove parâmetros é :

$$c(t) = \beta_1 + \beta_2 P(W_1) (1 + \beta_0 Q(x_1)) + \beta_5 P(W) (1 + \beta_8 (Q(x_2))).$$

Veja figura.

$$x_1 = t - \beta_9 - 0.5\beta_4$$

$$x_2 = t - \beta_6 - 0.5\beta_7$$

$$W_1 = \frac{t - \beta_9}{\beta_4}$$

$$W_2 = \frac{t - \beta_6}{\beta_7}$$

As parciais desta função de nove parâmetros são :

$$\frac{\partial c}{\partial \beta_1} = 1.0$$

$$\frac{\partial c}{\partial \beta_2} = P(W_1) [1 + \beta_0] Q_1$$

$$\frac{\partial c}{\partial \beta_3} = -\beta_2 \left[\frac{(1 + \beta_0 Q_1)}{\beta_4} \frac{\partial P}{\partial W_1} + P(W_1) \beta_0 \right]$$

$$\frac{\partial c}{\partial \beta_4} = -\beta_2 \left[\frac{P(W_1) \beta_0}{2} + \frac{(1 + \beta_0 Q_1)}{\beta_4} \frac{\partial P}{\partial W_1} W_1 \right]$$

$$\frac{\partial c}{\partial \beta_5} = 1 + \beta_8 Q_2 P(W_2)$$

$$\frac{\partial c}{\partial \beta_6} = -\beta_5 \left[P(W_2) \beta_8 + \frac{(1 + \beta_8 Q_2)}{\beta_7} \frac{\partial P}{\partial W_2} \right]$$

$$\frac{\partial c}{\partial \beta_7} = -\beta_5 \left[\frac{(1 + \beta_8 Q_2)}{\beta_7} W_2 \frac{\partial P}{\partial W_2} + \frac{P(W_2)}{2} \beta_8 \right]$$

$$\frac{\partial c}{\partial \beta_8} = \beta_5 Q_2 P(W_2)$$

$$\frac{\partial c}{\partial \beta_0} = \beta_2 Q_2 P(W_1)$$