

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



APLICAÇÃO DE DOIS MÉTODOS NO CONTROLE DE QUALIDADE EM DADOS COLETADOS POR
ESTAÇÕES METEOROLÓGICAS AUTOMÁTICAS DE SUPERFÍCIE

Autor: Luciano Ponzi Pezzi

Orientador: Prof. Dr. Robin Thomas Clarke

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

Porto Alegre, abril de 1993.

CEPSRM
CENTRO ESTADUAL DE PESQUISAS EM
SENSORIAMENTO REMOTO E METEOROLOGIA

Av. Bento Gonçalves, 9500

Campus Vale - Cx Postal 15044

91501 - Porto Alegre - RS

Fones: (051) 336-9827 - 336-8399 - 339-1355

R.: 6221/6264/6665 - Telefax: (IPH) (051) 336-0487

528101
18479
29/15706 *Alph88*

Pezzi, Luciano Ponzi

APLICAÇÃO DE DOIS MÉTODOS NO CONTROLE DE
QUALIDADE EM DADOS COLETADOS POR ESTAÇÕES
METEOROLÓGICAS AUTOMÁTICAS DE SUPERFÍCIE / Luciano
Ponzi Pezzi - Porto Alegre: UFRGS/CEPSRM - Curso de
Pós-Graduação em Sensoriamento Remoto, 1993.

Diss.Mestr.- UFRGS/CEPSRM - Curso de Pós-
Graduação em Sensoriamento Remoto, Porto Alegre, BR-
RS, 1993.

1. Controle de Qualidade

2. Estações Meteorológicas Automáticas de Superfície

UFRGS
BIBLIOTECA IPH

AGRADECIMENTOS

Ao Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia, onde realizei o curso de pós-graduação em Sensoriamento Remoto e desenvolvi meu trabalho de dissertação.

Ao projeto ABRACOS, por intermedio do Dr Robin Thomas Clarke, que forneceu os dados para a realização deste trabalho.

A CAPES pela concessão da bolsa de estudos, que sem a qual não seria possível a realização desse trabalho.

Ao meu orientador Dr. Robin Thomas Clarke, pelas sugestões, colaboração e atenção dispensada durante todas as etapas desse trabalho.

Aos professores do curso de pós graduação em Sensoriamento Remoto, que colaboraram para a minha formação profissional.

Aos amigos, Daniel Caimi que colaborou na elaboração dos programas computacionais e Vinicius Duarte Figueira na revisão do trabalho.

A todos os colegas do pós-graduação, que no decorrer do curso acabaram tornando-se amigos.

Aos amigos funcionários do CEP SRM, pelo apoio técnico, burocrático e amizade.

Principalmente a minha família. Aos meus pais, Terezinha (in memorian) e Cid Pezzi, minhas irmãs Thais e Marta Pezzi, que, ao seu modo, sempre incentivaram, apoiaram meu trabalho e minhas decisões.

E a todas as outras pessoas que de forma direta ou indireta colaboraram para a realização desse trabalho.

APLICAÇÃO DE DOIS MÉTODOS NO CONTROLE DE QUALIDADE EM DADOS COLETADOS POR
ESTAÇÕES METEOROLÓGICAS AUTOMÁTICAS DE SUPERFÍCIE

AUTOR: Luciano Ponzi Pezzi

ORIENTADOR: Dr Robin Thomas Clarke

RESUMO

No início dos anos 90, o Brasil e a Grã-Bretanha, juntos, deram um passo, no sentido de produzir estudos microclimáticos dentro da área que compreende a floresta Amazônica, visando quantificar e verificar o impacto do desmatamento dessa floresta sobre o clima. Para isso escolheu-se três regiões para a instalação de estações meteorológicas automáticas de superfície (Estados do Amazonas, Pará e Rondônia). Cada região recebeu uma estação em uma área florestada e outra em área desmatada. Esse projeto binacional foi chamado de "ESTUDO ANGLO-BRASILEIRO DE OBSERVAÇÃO DO CLIMA DA AMAZÔNIA, (ABRACOS - Anglo-Brazilian Amazonian Climate Observation Study)".

Esse trabalho apresenta a aplicação de dois métodos de controle de qualidade para os dados coletados pelas estações meteorológicas automáticas de superfície. Um desses métodos leva em consideração o comportamento estatístico das variáveis coletadas, e o outro método o comportamento físico e climatológico das mesmas. Por último se faz o uso desses dados com e sem erros no cálculo da evapotranspiração, no sentido de se ver o impacto de uma informação errônea no cálculo de um parâmetro.

O método estatístico demonstrou ter um bom potencial na detecção de erros pequenos (com tamanhos variando de 5 a 20%). Esse método apresentou rendimentos na obtenção de erros que oscilam entre 43% e 80% para a radiação solar global. Quando se testou os valores da temperatura do ar e temperatura do bulbo úmido, obteve-se um dos melhores rendimentos do método estatístico, que chegou a 91%. Os piores desempenhos foram quando se testou a radiação refletida, alcançando-se apenas 17,1%.

O método físico mostrou-se eficiente na detecção de erros maiores (30 a 90%). Os melhores desempenhos aconteceram quando se testaram as temperaturas. Em um caso o rendimento chegou a 100%. Em seu pior desempenho o método detectou apenas 20% (Teste "A"). O método físico somente detectou informações suspeitas, dentro dos erros inseridos nas séries de dados.

¹Dissertação de Mestrado em Sensoriamento Remoto. Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia (CEPSRM).

Abril de 1993

TWO METHODS TO QUALITY CONTROL FOR DATA COLLECTED BY AUTOMATIC
WEATHER STATION ON THE GROUND¹

Author: Luciano Ponzi Pezzi
Adviser: Dr Robin Thomas Clarke

ABSTRACT

In the beginning of the 1990's, Brazil and Great Britain, initiated cooperation to produce microclimate studies in the area covered by the amazon forest, with the purpose of qualifying and observing the impact of deforestation in this region on the climate. Three regions were chosen to set up automatic weather stations on the ground (in the states of Amazonia, Pará and Rondônia). Each region received one station in a forested and another in a deforested area. The binational project was called Anglo Brazilian Amazon Climate Observation Study.

This study shows two quality control methods applied to data collected by automatic weather station on the ground. One of these methods takes into account the statistical behavior of the variables collected, and the other, the physical and climatological behavior. Finally, these data are used, with and without error, to calculate evapotranspiration for the purpose of observing the impact of erroneous information in calculating parameters.

The statistical method showed a good potential for the detection of small errors (with sizes ranging from 5 to 20%). This method achieved a performance, in obtaining errors which varies between 43 an 80% for the global solar radiation. When the values for temperature of the humid bulb were tested, one of the best performances was obtained for the statistical method, reaching 91%. The worst results were found testing reflected radiation and only 17.1% was obtained.

The physical method proved effective to detect larger errors (30 to 90%). the best performances occurred when the temperatures were tested. In one case the performance was 100%. In the worst performance the method only detected suspect information among the errors inserted in the data series.

¹M.Sc. Thesis in Remote Sensing - Centro Estadual de Pesquisas em
Sensoriamento Remoto e Meteorologia. (CEPSRM).
April 1993.

SUMÁRIO

I. INTRODUÇÃO.....	1
1.1. O Problema.....	1
1.2. Objetivos.....	3
1.2.1. Primeiro Objetivo.....	4
1.2.2. Segundo Objetivo.....	4
1.3. Relevância do Trabalho.....	4
II. REVISÃO BIBLIOGRÁFICA.....	6
2.1. Variáveis Meteorológicas.....	6
2.1.1. Radiação Solar.....	6
2.1.1.1. O Espectro Solar.....	7
2.1.1.2. A Constante Solar.....	8
2.1.1.3. Radiação Solar Global.....	9
2.1.1.4. Albedo.....	11
2.1.1.5. Saldo de Radiação.....	12
2.1.2. Temperatura do ar.....	12
2.1.3. Fluxo de Calor no Solo.....	14
2.3. Balanço de Energia.....	15
III. MATERIAL E MÉTODOS.....	18
3.1. Estação Meteorológica Automática de Superfície.....	18
3.1.1. Definição.....	18
3.1.2. A Estação.....	19
3.1.3. Os Sensores.....	22
3.1.3.1. Radiometro Solar.....	23

3.1.3.2. Rad. Radiação Líquida (Saldo Radiômetro)	23
3.1.3.3. Termômetros	23
3.1.3.4. Anemômetro	24
3.1.3.5. Catavento	24
3.2. Material e Equipamento Disponível	25
3.2.1. As Séries de Dados	25
3.2.1.1. Banco de Dados	25
3.2.1.2. Erros Aleatórios	28
3.3. Método Estatístico	29
3.3.1. Observações Suspeitas	29
3.3.2. O Método Estatístico	30
3.3.3. O Programa Computacional do Método Estatístico	39
3.4. Método Físico	42
3.4.1. Descrição dos Testes do Método Físico	43
3.4.2. O Programa Computacional do Método Físico	47
3.5. Método usado no Cálculo da Evapotranspiração	47
 IV. RESULTADOS	 51
4.1. Arquivos Principais	51
4.2. Resultados do Método Estatístico	52
4.2.1. Resultados do Método Estatístico para Fazenda Dimona	56
4.2.2. Resultados do Método Estatístico para Manaus	59
4.2.3. Resultados do Método Estatístico para Reserva Ducke	61
4.2.4. Resultados do Método Estatístico para a série de erros consecutivos crescentes	63
4.3. Resultados do Método Físico	64
4.3.1. Resultados do Método Físico para Manaus	66
4.3.2. Resultados do Método Físico para Fazenda Dimona	68
4.4. Resultados do Cálculo da Evapotranspiração usando o método de Penman-Monteith	70

V. CONCLUSÕES.....	75
5.1. Conclusões para o Método Estatístico.....	75
5.2. Conclusões para o Método Físico.....	76
5.3. Conclusões para o Cálculo da Evapotranspiração.....	77
VI. BIBLIOGRAFIA.....	78
VII. ANEXO I.....	82

LISTA DE FIGURAS

Figura 2.1. Espectro Solar.....	7
Figura 2.2. Componentes da Radiação Solar Global.....	10
Figura 3.1. Diagrama esquemático referente a coleta e transmissão dos dados via satélite.....	19
Figura 3.2. Temperatura do bulbo seco x Temperatura do bulbo úmido.....	31
Figura 3.3. Plotagem do alavancamento (leverage) x Posição da observação (n).....	35

LISTA DE TABELAS

Tabela 3.1. Tabela das estações meteorológicas automáticas de superfície que compõem o projeto ABRACOS	21
Tabela 3.2. Lista de sensores que compõem as estações meteorológicas automáticas de superfície	22
Tabela 3.3. Informações na ordem em que aparecem na mensagem....	26
Tabela 3.4. Mensagem enviada pela estação de Manaus (MS), referente ao dia 100 do ano de 1991.....	27
Tabela 3.5. Dados de temperatura usados no exemplo estatístico..	33
Tabela 3.6. Matriz H para o exemplo ilustrado.....	34
Tabela 3.7. Tabela com todos os r_i^* calculados no exemplo.....	41
Tabela 3.8. Tabela com os r_i^* maiores que os limites testados....	41
Tabela 3.9. Descrição dos testes usados no Método Físico.....	44
Tabela 4.1. Tabela dos Arquivos Principais, gerados a partir do FUSÃO.BAS.....	51
Tabela 4.2. Arquivos derivados da Fazenda Dimona para o método estatístico.....	53

Tabela 4.3. Arquivos derivados da Cidade de Manaus para o método estatístico.....	53
Tabela 4.4. Arquivos derivados da Reserva Ducke para o método estatístico.....	53
Tabela 4.5. Resultados do método estatístico para a Fazenda Dimona.....	58
Tabela 4.6. Resultados do método estatístico para a Cidade de Manaus.....	60
Tabela 4.7. Resultados do método estatístico para a Reserva Ducke.....	62
Tabela 4.8. Arquivos derivados da cidade de Manaus	64
Tabela 4.9. Arquivos derivados da Fazenda Dimona	65
Tabela 4.10. Tabela de resultados do método físico para a Cidade de Manaus.....	67
Tabela 4.11. Tabela de resultados do método físico para a Fazenda Dimona.....	70
Tabela 4.12. Tabela com as informações do arquivo de erros gerados para a Cidade de Manaus (MSE1.DAT)	71
Tabela 4.13. Tabela com os resultados dos cálculos de evapotranspiração, horária e diária, para a cidade de Manaus	72
Tabela 4.14. Tabela com as informações do arquivo de erros gerados para a Reserva Ducke (RDE1.DAT).....	73

Tabela 4.15. Tabela com os resultados dos cálculos de evapotranspiração, horária e diária, para a Reserva Ducke	74
---	----

I. INTRODUÇÃO

1.1. O Problema

A meteorologia é a ciência que estuda a diversidade dos fenômenos físicos e dinâmicos que ocorrem na atmosfera terrestre, investigando as causas das alterações que nela se verificam. Seu estudo é de suma importância, uma vez que os fenômenos por ela estudados interferem diretamente sobre a vida do Planeta em suas mais diversas formas. Nesse sentido é que cientistas brasileiros e britânicos estão empenhados em investigar a Amazônia e preparar modelos climáticos precisos, que tenham habilidade suficiente para descrever situações atmosféricas, bem como prognosticá-las.

Até recentemente, modelos climáticos têm dispensado pouca atenção para com este complexo e importante controle hidrológico exercido pela vegetação. Agora, alguns Modelos de Circulação Global (ou Global Circulation Models-GCMs) incorporam descrições mais realistas, apesar de serem pobremente calibrados com os dados de campo.

Em meio a década de 80, o Brasil e a Grã-Bretanha, juntos, deram o primeiro passo, no sentido de produzir estudos microclimáticos. Foi realizado um estudo micrometeorológico por cientistas do Instituto de Hidrologia (IH-Grã-Bretanha), Instituto Nacional de Pesquisas Espaciais (INPE-Brasil) e Instituto Nacional de Pesquisas da Amazônia (INPA-Brasil), em Manaus, dentro da Reserva Ducke em uma região de floresta. Para isto, foi construída uma torre de 45 metros de altura dentro da floresta. A copa florestal ficou em torno dos 35 metros de altura, e os instrumentos (estações meteorológicas automáticas)

metros de altura, e os instrumentos (estações meteorológicas automáticas) posicionados logo acima da cobertura florestal.

No início da década de 90, este estudo foi expandido para outras localidades da floresta amazônica com objetivos mais ambiciosos, no sentido de produzir estudos, com credibilidade, dos efeitos dos desmatamentos sobre o clima. Esse foi então, o início de um projeto pioneiro de pesquisas na floresta em locais mais afastados da região urbana. Foram escolhidos três regiões. Uma ao norte da cidade de Manaus, uma no estado do Pará (lado este da floresta) e outra no estado de Rondonia (lado oeste da floresta). Este projeto foi chamado de "ESTUDO ANGLO-BRASILEIRO DE OBSERVAÇÃO DO CLIMA DA AMAZÔNIA (ABRACOS - Anglo-Brasilian Amazonian Climate Observation Study)".

Os dados coletados por essas estações meteorológicas automáticas de superfície serão usadas para fornecer as condições de contorno da floresta amazônica com os Modelos de Circulação Global. Os três modelos que geralmente tem sido usados na atualidade são o GISS (Goddard Institute for Space Studies), GFDL (Geophysical Fluid Dynamics Laboratory) e o UKMO (United Kingdom Meteorological Office). TUCCI e DAMIANI (1991), realizaram um estudo na Bacia do Rio Urugai e constataram que os modelos usados, subestimaram as medidas de precipitação para aquela área. O modelo hidrometeorológico que mais se aproximou as medidas feitas, subestimou em 21.6%, GISS, enquanto os outros dois subestimaram em 56.2% e 65.4%, UKMO1 e GFDL respectivamente.

Segundo informe publicado pelo Instituto de Hidrologia (IH-Reino Unido), as últimas previsões são melhores e dão uma indicação da possível mudança climática. Os resultados sugerem um aumento de 2° na temperatura; uma redução de 20% na evaporação e redução de 30% na precipitação da Amazônia. Essas mudanças nos dão razões claras para pensar a respeito de qual é o real valor de uma reserva ecológica nas prováveis mudanças climáticas.

Devido à fraca performance do modelo, no passado, sua credibilidade é, hoje discutível. Atualmente, através de experimentos, podemos substanciar e validar esses modelos, conferindo-lhes maior credibilidade.

Para alimentar esses modelos com a realidade de comportamento da floresta amazônica, foram instaladas estações meteorológicas automáticas de superfície nos locais descritos anteriormente que tem a finalidade de coletar, armazenar e transmitir os dados (via satélite inclusive). As estações podem ser programadas para fazer diversas medidas dentro de um determinado intervalo de tempo, por exemplo: pode-se programa-las para que se fassam coletas de dados de 1 em 1 minuto, e apresentando médias de 30 minutos. Isso gera um grande volume de informações ao final de um determinado período de tempo. Essas informações estão sujeitas a ocorrência de erros que podem ser gerados por uma diversidade de fatores: falhas na transmissão, afastamento da calibração ideal dos sensores, danos causados por animais nativos do local ou até mesmo a própria ação do tempo. A umidade dentro da abóbada que envolve os sensores de radiação pode ser uma fonte de grandes erros para esta informação.

Devido a esta situação se faz necessária a implementação de técnicas de controle de qualidade para estas informações, fazendo-se uso de processos automatizados, uma vez que manualmente despenderia-se de muito tempo devido ao grande volume e rapidez com que as informações chegam ao usuário.

1.2. Objetivos

A estação meteorológica automática de superfície é calibrada ao ser construída e instalada. Seus sensores são devidamente testados e ajustados eletronicamente. A energia que alimenta esta estação é obtida através de baterias carregadas por energia solar. Pelo fato dessa estação provavelmente localizar-se em local de difícil acesso é inviável que um técnico desloque-se a todo momento para junto dos sensores a fim de conferir suas calibrações e aferições.

1.2.1. PRIMEIRO OBJETIVO

- Verificar a qualidade dos dados coletados e transmitidos por uma estação meteorológica automática de superfície, através de dois métodos, e também analisar a eficiência dessas duas técnicas. Um desses métodos é baseado no comportamento físico das variáveis e foi denominado MÉTODO FÍSICO; o outro baseia-se no comportamento estatístico dessas variáveis e é, por conseguinte denominado MÉTODO ESTATÍSTICO.

1.2.2. SEGUNDO OBJETIVO

- Verificar os impactos dos prováveis erros não detectados pelos modelos físico e estatístico, através do cálculo da evapotranspiração potencial, nos locais onde foram coletados os dados, utilizando-se o método do balanço de energia superficial combinado, desenvolvido por PENNMAN-MONTEITH.

1.3. Relevância do trabalho

O clima continental é muito diferente daquele verificado nos oceanos, evidenciando que a atmosfera é largamente influenciada com os acontecimentos ocorridos na superfície terrestre. Recente interesse sobre derrubadas escalonadas da floresta amazônica e os seus prováveis impactos no clima tem sido a base para as previsões feitas pelos "MODELOS DE CIRCULAÇÃO GLOBAL" (GCMs). No entanto, para que esses modelos possam ser rodados e para que resultados confiáveis sejam obtidos, é necessário que se use informações representativas, atuais e seguras, coletadas dentro da área que é compreendida pela floresta amazônica.

Tanto o governo brasileiro quanto o britânico estão ansiosos pela melhoria nas avaliações das previsões feitas pelos modelos climáticos. Para isso se organizou um dos maiores projetos conjuntos, resultante de um MEMORANDO DE ENTENDIMENTO chamado, ABRACOS. Esse estudo é uma colaboração entre o Instituto de Hidrologia-IH (Wallingford,UK), o Instituto de Pesquisas Espaciais-INPE (São José dos Campos-SP), o Instituto de Pesquisas da Amazônia-INPA (Manaus-AM) e o Centro de Energia Nuclear para Agricultura-CENA (Brasil).

Este trabalho de dissertação para a obtenção do grau de mestre, surge como uma pequena parcela de contribuição dentro de um grande acordo binacional entre o Brasil e o Reino Unido, e tem, como benefícios mais evidentes, dentre outros:

- Elaborar um processo de checagem dos dados coletados por uma estação meteorológica automática de superfície, com ênfase na detecção de erros grosseiros a fim de que esses dados tenham a maior confiabilidade possível, para que possam ser usados nos Modelos de Circulação Global.

- Verificar o impacto do erro em outras variáveis, como a evapotranspiração potencial, através do uso de séries de dados supostamente corretas e séries contendo erros.

II. REVISÃO BIBLIOGRÁFICA

2.1. VARIÁVEIS METEOROLÓGICAS

Neste item, foi feita uma breve descrição das variáveis meteorológicas coletadas pela estação meteorológica automática de superfície, e que foram utilizadas neste trabalho. Procurou-se enfatizar as definições e comportamento dessas variáveis meteorológicas.

2.1.1. Radiação Solar

Podemos considerar o sol como sendo uma esfera de gás aquecida por uma reação nuclear, com uma superfície aparente, (chamada fotosfera) e um diâmetro de 1.3914×10^6 Km. A distância média do sol à terra é de 149.6×10^6 Km, muitas vezes chamado de uma unidade astronômica (1 U.A), (SLATER,1980).

O sol é a fonte primária de toda a energia disponível aos processos naturais, que ocorrem na superfície da terra, bem como a todos aqueles processos verificados na atmosfera que envolve o planeta. A irradiância solar é a energia recebida pela Terra na forma de ondas eletromagnéticas, provenientes do sol.

2.1.1.1. O Espectro Solar

Na figura 2.1 podemos verificar a curva da irradiância espectral solar que chega à Terra, plotada em função do comprimento de onda. O gráfico da irradiância ilustra os efeitos atmosféricos na radiação solar que atravessa a massa gasosa que envolve o planeta. A radiação solar sofre a influência desses efeitos atmosféricos até chegar a superfície, sendo os gases constituintes da atmosfera os responsáveis pelos efeitos causados na irradiância solar, quando esta atravessa a atmosfera. Estes efeitos são: espalhamento, reflexão e absorção.

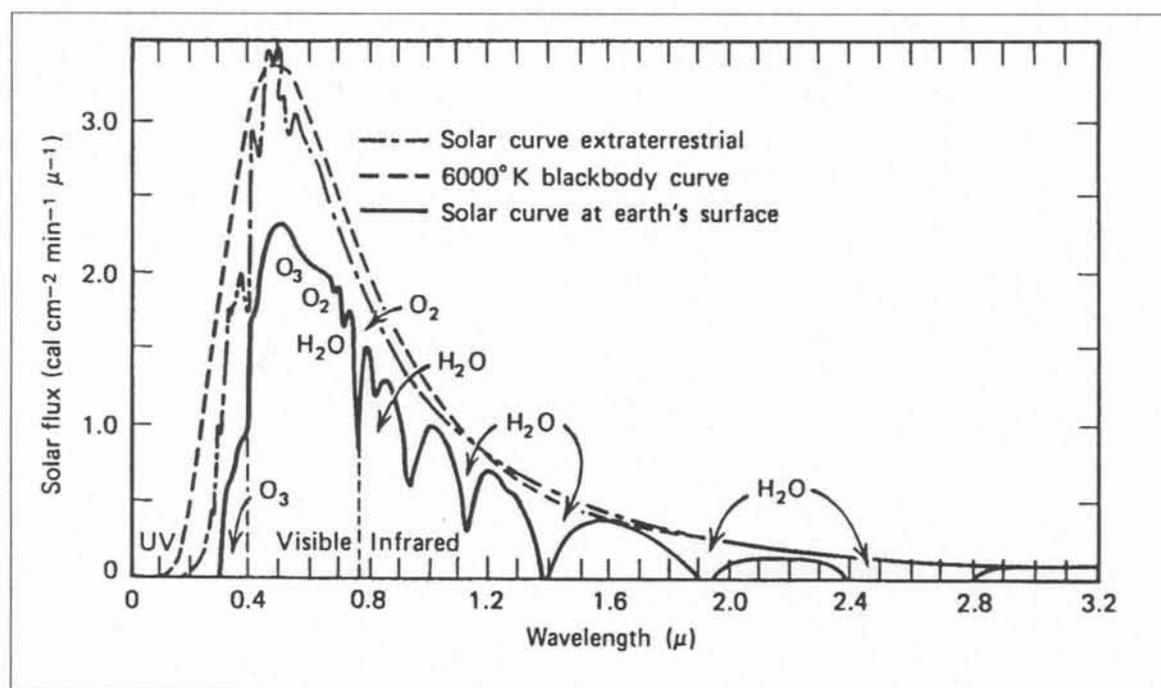


Figura 2.1. Espectro Solar. Fonte (ROSENBERG, 1974)

O espectro eletromagnético solar assemelha-se com aquele de um corpo negro quando aquecido a uma temperatura superficial de aproximadamente 6000 °K. Deste fluxo de radiação que chega do sol, apenas a energia radiante de entrada é que determina o balanço de calor e o regime térmico da terra.

A maior parcela de energia do espectro solar, aproximadamente 99%, encontra-se na faixa do espectro de ondas curtas, com comprimentos de ondas (λ) que variam de 0.3 μm à 4.0 μm . Desta forma, este tipo de radiação é caracterizado como radiação de ondas curtas. Dentro desse espectro existe uma subdivisão, em outras três sub-faixas, que são:

A) A radiação ultravioleta, que se estende desde os 0.4 μm até os menores comprimentos de onda e que, na soma total das sub-faixas, é responsável por 9% da radiação. Sua parcela de contribuição é pequena, porém tem uma grande influência em vários processos biológicos. (BUDYKO, 1974)

B) A radiação visível, que se estende desde os 0.4 μm até os 0.7 μm e, dentro do espectro total, é responsável por 41% da radiação solar total.

C) E, por último a radiação infravermelha que é a responsável pelos 50% restantes do total de radiação solar. Essa radiação se estende desde 0.7 μm até os comprimentos maiores.

2.1.1.2. Constante Solar

Constante solar é definida como sendo a quantidade de energia que incide perpendicularmente sobre uma superfície horizontal e unitária na unidade de tempo, quando o raio vetor terra-sol é unitário e a atmosfera terrestre é considerada ausente. É considerada a quantidade total de energia solar, que chega no topo da atmosfera do planeta. O seu valor absoluto é aceito como sendo 2.0 $\text{cal.cm}^{-2}.\text{min}^{-1}$.

Alguns autores falam à respeito de uma determinada variação na média da constante solar. Essas mudanças podem ser atribuídas, principalmente, as flutuações que ocorrem na porção do espectro referente ao ultravioleta. Essa afirmação foi feita por RENSE(1956) e encontrada em CHANG(1977).

2.1.1.3. Radiação Solar Global

A radiação solar, ao atravessar a atmosfera, tem parte do seu fluxo difundido pelos constituintes fixos e variáveis que compõem a atmosfera terrestre e parte que chega diretamente ao solo. Na superfície da terra, essa radiação que chega é usualmente medida em um plano horizontal. Se R_p é a componente direta da irradiância, em um plano na superfície perpendicular ao feixe solar a uma elevação β , esta componente da radiação solar global é definida como sendo a radiação solar direta.

A outra componente é chamada de radiação solar difusa ou radiação do céu (R_d). Levando-se em conta R_d , que vem do céu e das nuvens, podemos escrever a radiação solar global como sendo:

$$R_G = R_p \cdot \text{sen}(\beta) + R_d = R_D + R_d \quad (2.1)$$

A Figura 2.2 ilustra as componentes de R_G . (MONTEITH,1980). A magnitude, das componentes de R_G , depende da elevação solar, do espalhamento e absorção causados por gases atmosféricos, aerossóis, gotículas d'água e partículas de gelo nas nuvens.

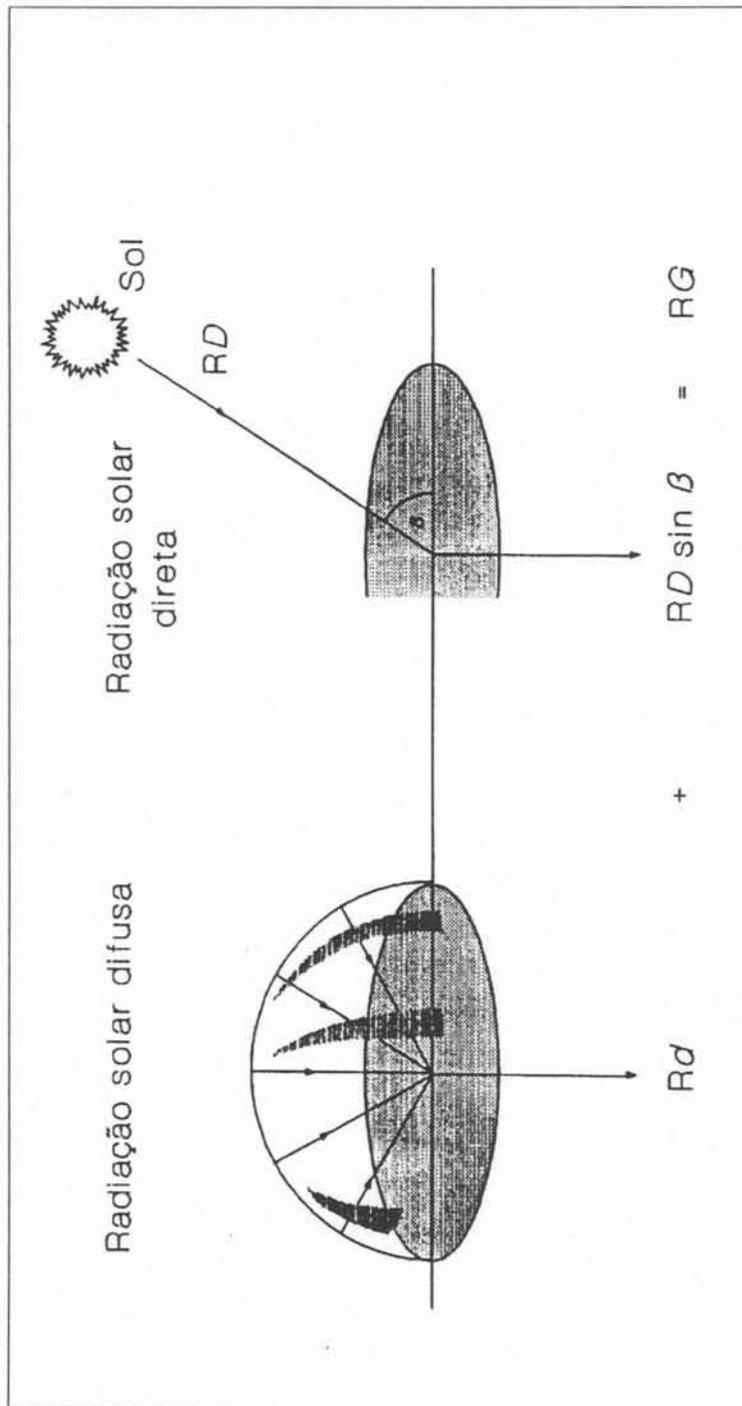


Figura 2.2. Componentes da Radiação Solar Global

2.1.1.4. Albedo

Todos os materiais naturais refletem e transmitem radiação solar nos comprimentos de ondas que vão desde $0.3\mu\text{m}$ à $4.0\mu\text{m}$. Nos curtos comprimentos de onda, o processo radiativo dos materiais é determinado, principalmente, pela presença de pigmentos absorventes de radiação associados com os comprimentos de ondas específicos e transições eletrônicas.

A água líquida é um importante constituinte dos materiais encontrados na natureza, e na banda que estende-se de $1\mu\text{m}$ à $3\mu\text{m}$, se encontra a região de grande absorção deste elemento. Por outro lado, no espectro visível, onde a absorção da água é negligível, a reflexão e transmissão de luz pelos materiais porosos são largamente correlacionados com seus conteúdos de água, (MONTEITH, 1980).

$$a = \frac{SW\uparrow}{SW\downarrow} \quad (2.2)$$

O coeficiente de reflexão ou albedo (como também é chamado) pode ser definido como sendo a refletividade média sobre uma determinada banda.

O termo albedo é usado para descrever a reflexão do feixe solar ($0.3\mu\text{m}$ à $4.0\mu\text{m}$) ou algumas vezes para descrever somente a reflexão na banda do visível ($0.4\mu\text{m}$ à $0.7\mu\text{m}$), (ROSENBERG, 1974). Por convenção, neste trabalho se adotou o termo albedo para ilustrar a primeira situação, descrita no parágrafo anterior.

2.1.1.5. Saldo de Radiação

Saldo de radiação ou radiação líquida (como também é chamada) é definida como, a diferença entre o total dos fluxos que abandonam a superfície da terra, e os que chegam. Essa componente da radiação é de extrema importância, pois esta é a quantidade de energia que fica disponível na superfície da terra e é usada no processo de evaporação, no aquecimento do ar, no aquecimento do solo, bem como outros processos que consomem energia, como por exemplo, a fotossíntese realizada pelos vegetais, (ROSENBERG, 1974).

O saldo de radiação, pode ser melhor definido pela equação do balanço de radiação na superfície dada por:

$$R_n = R_G (1 - a) + (L_{\downarrow} - L_{\uparrow}) \quad (2.3)$$

onde: R_n é o saldo de radiação ou radiação líquida, R_G a radiação solar global, a é o albedo da superfície, L_{\downarrow} é a radiação de ondas longas emitidas da atmosfera para a superfície; e L_{\uparrow} é a radiação de ondas longas emitidas pela superfície para a atmosfera.

2.1.2. Temperatura do Ar

O aquecimento e o resfriamento do ar atmosférico é governado pelo balanço de radiação da superfície do solo. A temperatura possui variações diárias e estacionais. As trocas de calor entre a superfície do solo e do ar ocorrem basicamente por condução do calor, que se verifica na fina camada de ar junto ao solo, e , por convecção, para níveis mais altos. Devido à baixa condutividade térmica do ar, os processos de seu aquecimento e resfriamento não atingem mais de três metros de altura, se ocorrerem exclusivamente por

condução. O fator preponderante de transporte de calor na atmosfera é a difusão turbulenta ou convecção, ocasionada pelos movimentos turbulentos do ar na atmosfera, que transporta, além do calor, outras propriedades importantes como o vapor da água, CO₂, poluentes e outros.

Quando o balanço de energia na superfície do solo começa a ser positivo, o ar em contato com o solo começa a ser aquecido por condução. Esse ar então expande-se, eleva-se e é substituído por ar de maior densidade. Este é o início de um processo de convecção livre, que se intensifica com o aumento do balanço positivo da radiação. Por esse processo o ar vai sendo sucessivamente aquecido, sendo que aquele ar que se encontra junto ao solo se aquece mais rápida e intensamente.

Segundo BUDYKO(1974), as diferenças entre as naturezas dos diversos tipos de superfícies do planeta (solo, oceanos, culturas vegetais e outras) terão temperaturas superficiais diferentes daquelas encontradas em camadas de ar logo acima destas mesmas superfícies. Estas diferenças é que determinarão o aparecimento de fluxos verticais de calor, entre essas superfícies e a camada de ar existente logo acima delas.

A temperatura máxima do ar em contato com o solo ocorre simultaneamente com a temperatura máxima da superfície do solo. À medida que se afasta do solo, a temperatura máxima do ar se atrasa continuamente, indo ocorrer a dois metros de altura, cerca de duas horas depois. Após a temperatura máxima, a temperatura do ar diminui como consequência da temperatura do solo, com o balanço de radiação positivo e decrescente. Quando o balanço de radiação torna-se negativo estabelece-se um fluxo de calor por condução do ar para a superfície. Esse fluxo passa a resfriar o ar, iniciando pelas camadas subjacentes. Outro conceito útil, é a temperatura do termômetro de bulbo úmido. Essa temperatura é determinada através de um termômetro em que o bulbo está envolvido em uma gase, que por sua vez, é embebida em água destilada: Um fluxo de ar é forçado a passar por este bulbo umedecido do termômetro. Com a passagem do ar ocorre a evaporação da água contida

na gase, com a utilização do calor sensível que é retirado do ar passante, de modo que esse termómetro regista a temperatura do búlbo úmido. A diferença da temperatura do ar e da temperatura do búlbo úmido, é denominada depressão psicrométrica.

2.1.3. Fluxo de Calor no Solo

O fluxo de calor no solo, para dentro e para fora do mesmo, é um processo de condução. De acordo com a teoria cinética o calor é uma energia que é proveniente da vibração molecular. A vibração molecular é transmitida para outras moléculas que estejam vibrando mais vagarosamente, e assim um corpo acaba interagindo, e transmitindo energia para outros que estejam inseridos no mesmo meio, (ROSENBERG, 1974).

O fluxo de calor no solo é diretamente proporcional ao gradiente vertical de temperatura existente naquele nível, dessa forma, a equação que fornece esse fluxo para uma determinada profundidade Z, segundo SELLERS(1972), pode ser escrita da seguinte forma:

$$S = \lambda (dT/dZ) \quad (2.4)$$

onde λ é a condutividade térmica, que é a constante física dependente da composição, umidade e temperatura do solo e é quem determina a taxa da transferência de calor, dT/dZ é o gradiente vertical de temperatura do solo e, por fim, S que é o fluxo de calor do solo. Por convenção, diversos autores adotam que todo fluxo de energia que sai da superfície é considerado fluxo negativo ao passo que todo fluxo de calor que entra no solo, é considerado como fluxo positivo, (SELLERS, 1972; ROSENBERG, 1974).

2.3. Balanço de Energia:

Em micrometeorologia, radiação solar é a principal fonte de chegada de energia para o balanço de energia. Por outro lado, condução, convecção e evaporação/condensação são métodos de expedição (MONTEITH,1980).

A técnica do balanço de energia consiste na repartição da energia do saldo de radiação que se encontra junto à superfície, descontadas as trocas de armazenamento de calor no solo, em componentes de calor sensível e de calor latente.

Vários autores e pesquisadores, ROSENBERG(1974), MONTEITH(1980), CUNHA(1988), citam a equação fundamental do balanço de energia como sendo:

$$R_n = H + LE + S + P + M \quad (2.5)$$

onde: R_n é o saldo de radiação, já definido anteriormente, H é o fluxo de calor sensível, LE é o fluxo de calor latente, S é o fluxo de calor no solo, P é representação da fotossíntese e M reúne os diversos tipos de trocas de energia, que podem ser trocas de energia devido à atividade metabólica e armazenamento do calor no tecido das plantas ou no volume do dossel. (ROSENBERG,1974)

O estudo detalhado dos fluxos de energia em um sistema físico engloba aspectos do saldo de radiação e seu particionamento nas formas de calor sensível e calor latente, bem como considerações de fatores de planta e solo que interferem nos mesmos. Na equação (2.5) aparecem dois termos, P e M , que usualmente são menores que os erros experimentais das medidas feitas nas outras componentes, sendo geralmente desprezados. Logo a equação, fundamental e simplificada, do balanço de energia na superfície é escrita da seguinte

forma

$$R_n = H + LE + S \quad (2.6)$$

O processo evaporativo é limitado pela energia disponível ao sistema, e não por fatores de planta ou solo, quando a água estiver disponível a este sistema. A maior parcela da energia transferida para a atmosfera do planeta aparece como calor sensível e calor latente, que irão depender diretamente, em proporções, da água disponível para o processo. As plantas são capazes de utilizar mais energia para a evapotranspiração do que recebem como energia radiante. Deste modo, a quantidade de energia a ser consumida no processo de evapotranspiração pode variar de zero, quando não houver água disponível na comunidade vegetativa, ou até ao outro extremo, onde chega-se aos valores da evapotranspiração potencial. Neste extremo não existe a limitação da água na comunidade vegetativa e a energia consumida no processo pode exceder o saldo de radiação, e até mesmo a radiação solar global, (TANNER e LEMON, 1962).

A taxa de evapotranspiração é governada principalmente pela quantidade de energia disponível e pela quantidade de água presente. Em regiões úmidas a menor parcela de energia usada para o processo de evapotranspiração provém da advecção de calor, enquanto que a maior parcela de energia usada no processo fica por conta da radiação solar incidente, (AUBERTIN e PETERS, 1961). Em seu trabalho sobre evaporação e transpiração, PETERS (1960), afirma que sob determinadas condições, quando se tratam de regiões sem restrição de água onde as fontes de umidade são contínuas, o principal uso do saldo de radiação é na evaporação da água.

Em muitos casos, usa-se o saldo de radiação para se obter a medição precisa do uso da água. Entretanto, essa última afirmação torna-se verdadeira somente quando aplicada a áreas em que não ocorram consideráveis advecções de calor sensível e calor latente ou, ainda, em áreas onde não ocorram freqüentemente períodos secos.

Os fatores da planta que afetam a evapotranspiração são a área foliar, o desenvolvimento das raízes, o tipo e a idade fisiológica das plantas. A área foliar parece ser de grande importância direta, como superfície transpirante, e indireta, como superfície que causa o sombreamento do solo. (TANNER e LEMON, 1962)

III. MATERIAL E MÉTODOS

3.1. ESTAÇÃO METEOROLÓGICA AUTOMÁTICA DE SUPERFÍCIE

3.1.1. Definição

É usual definir-se estação meteorológica de superfície como o local onde se efetua a avaliação de um ou vários elementos meteorológicos. As estações meteorológicas geralmente dispõem dos instrumentos necessários e suficientes à avaliação dos elementos que se pretende conhecer (VAREJÃO, 1979). Neste trabalho a presença de um observador não se faz necessária, uma vez que se trata de uma estação meteorológica automática de superfície que, por suas características de funcionamento, pode ser denominada como um conjunto de sensores remotos.

Sensoriamento Remoto é definido de uma maneira geral como sendo a ciência de aquisição de informações sobre um objeto a partir de medidas feitas a distância desse mesmo objeto, isto é, sem que haja contato físico com ele (LANDGREBE, 1978). Complementando esta idéia, MORAES NOVO (1988) sugere que os sensores são equipamentos capazes de coletar energia proveniente do objeto, e de converter, essa mesma energia, em sinal possível de ser registrado e apresentado de forma adequada à extração de informações.

3.1.2. A estação

As estações usadas neste trabalho de pesquisa, se baseiam em um sistema compacto digitalizador com gravação de dados em memória "solid state", capaz de gravar, em forma digital, até 12 tensões analógicas de entrada, escritas em palavras de "8 bits" sucessivos. Essa estação tem a capacidade de transmitir seus dados em tempo real via satélite, conforme o esquema apresentado a seguir (figura 3.1).

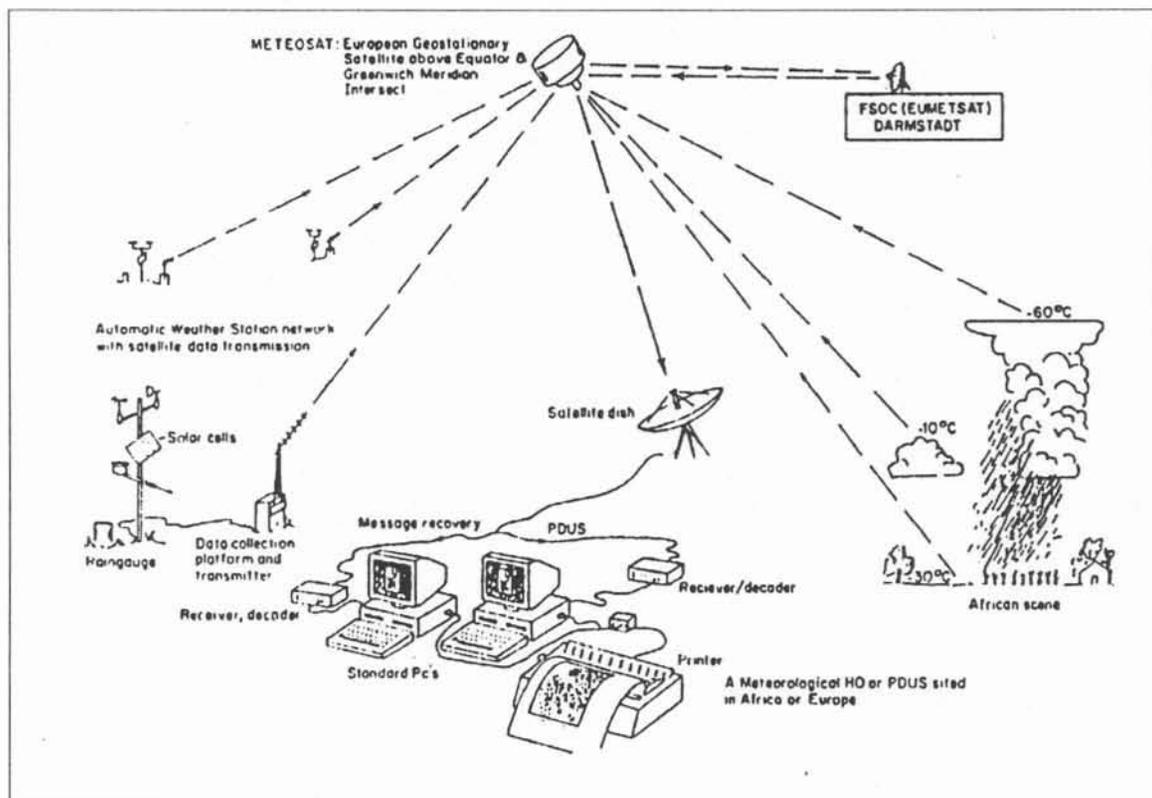


Figura 3.1. Diagrama esquemático referente a coleta e transmissão dos dados via satélite. Fonte (DUGDALE, G. *et al*, 1991).

Os sensores de medição da radiação solar e do saldo de radiação; das temperaturas do ar; do termômetro de bulbo úmido; da velocidade e direção dos ventos, estão montados num mastro simples com braços horizontais, enquanto o pluviômetro fica localizado no solo à cerca de seis metros do mastro. Cada sensor converte o valor medido num sinal elétrico que é então convertido numa tensão compatível para digitalização, por meio de uma unidade compacta de "interface".

Os sinais analógicos são sequencialmente amostrados, digitalizados e gravados pelo digitalizador (operado a bateria). Cada amostragem e sequência de registros constitui uma varredura. Essas amostragens podem ser programadas para ocorrer em intervalos de 10 segundos e a varredura, por sua vez, em 5 minutos.

A fonte de alimentação é um painel de células solares que armazenam energia para o funcionamento da estação. Cada estação é acompanhada de um DCP (Data Collection Platform and Transmitter), cuja finalidade é receber os dados coletados pela estação e enviá-los para o satélite METEOSAT (Satélite Geostacionário Europeu, localizado nas intersecções do Meridiano de Greenwich e a linha do Equador, à 36.000 Km de altura).

Do METEOSAT, essas informações vão para uma estação terrestre em Darmstadt (FSOC/EUMETSAT) na Alemanha, onde todos os dados coletados e transmitidos pelos sensores são processados e decodificados. Após, são reemitidos para o METEOSAT que, por sua vez, redistribui essas informações para os PDUS (decodificador das mensagens) e, finalmente, armazenados nos PC's, onde podem ser manipulados pelo usuário.

Inicialmente, pretendia-se trabalhar com os dados de sete estações meteorológicas automáticas de superfície que são apresentadas na tabela subsequente (tabela 3.1). Essas estações compõem o projeto ABRACOS e estão distribuídas da seguinte maneira: três em Manaus (uma localizada na própria cidade, postada sobre um gramado; uma na Fazenda Dimona, sobre pastagem e

localizada a 70 Km ao norte da cidade; a terceira na Reserva Ducke, que esta instalada sobre a copa das arvores em meio à floresta e localizada à 25 Km ao norte da cidade). As outras 4 estações também já estão instaladas e funcionando em caráter operacional. São duas no Pará e duas em Rondônia. Das duas estações localizadas no estado do Pará, a que está localizada na Reserva do Vale do Rio Doce foi instalada em área florestada e a outra, Fazenda da Boa Sorte, em área desmatada. As duas estações restantes estão no estado de Rondônia. Uma delas se localiza na Reserva Jarú e foi instalada em região florestada; a outra, na Fazenda Nossa Senhora Aparecida, foi instalada em região desmatada. Essas estações realizam suas coletas de dados e enviam ao PDUS instalado no Instituto de Pesquisas Hidráulicas (IPH/UFRGS). Esses dados são recebidos aqui em tempo real, não foi possível utilizar as informações das sete estações pelos motivos expostos, no item 3.2.1.

Tabela 3.1 - Tabela das estações meteorológicas automáticas de superfície que compõem o projeto ABRACOS

N°	CÓD.	NOME	LOCALIDADE	LATITUDE	LONGITUDE
01	MS	MANAUS	AMAZÔNAS	03°00'S	60°00'W
02	RD	R. DUCKE	AMAZÔNAS	02°19'S	59°57'W
03	FD	F. DIMONA	AMAZÔNAS	02°57'S	60°19'W
04	RV	R. V. RIO DOCE	PARÁ	05°45'S	49°10'W
05	BS	F. BOA SORTE	PARÁ	05°10'S	48°45'W
06	RJ	R. JARU	RONDÔNIA	10°05'S	61°55'W
07	NS	F. N. S. APARECIDA	RONDÔNIA	10°41'S	62°22'W

3.1.3. Os Sensores

Na tabela 3.2, a seguir é apresentada uma lista completa dos sensores que estão disponíveis nas estações meteorológicas automáticas de superfície usadas nesse estudo. Essa tabela informa o número de sensores que fazem o registro de cada variável observada e também fornece a unidade de saída para cada uma destas variáveis coletadas. Todas as descrições técnicas dos equipamentos que compõem as estações foram retiradas de um trabalho realizado por CLARKE e SILVEIRA(1979), com uma estação semelhante instalada junto ao 8° DISME/Ministério da Agricultura.

Tabela 3.2 - Lista de sensores que compõem as Estações Meteorológicas Automáticas de Superfície:

SENSORES	UNIDADE
1 RADIÔMETRO SOLAR	W/m ²
1 RADIÔMETRO DE RADIAÇÃO LÍQUIDA	W/m ²
1 RADIÔMETRO DE RADIAÇÃO DIFUSA	W/m ²
1 RADIÔMETRO DE RADIAÇÃO REFLETIDA	W/m ²
1 TERMÔMETRO DE BULBO ÚMIDO	°C
1 TERMÔMETRO DE BULBO SECO	°C
2 SENSORES DE FLUXO DE CALOR DO SOLO	W/m ²
1 SENSOR DE DIREÇÃO DO VENTO	(GRAUS)
1 ANEMÔMETRO	m/s
1 PLUVIÔMETRO	mm

3.1.3.1. Radiômetro Solar

O radiômetro solar tem como principal função medir a radiação solar global. É capaz de medir intensidades de até $8.37 \times 10^4 \text{ J/m}^2 \cdot \text{min}^{-1}$ ($\sim 1400 \text{ W.m}^{-2}$). Esse radiômetro mede a voltagem gerada por uma termopilha cujas junções de aquecimento estão em contato com um disco negro exposto às radiações solares. A caixa do instrumento é protegida por um disco branco maior, que mantém a temperatura ambiente. Esse instrumento está montado no topo do mastro que sustenta todos os outros sensores. Esta posição estratégica conferida ao radiômetro solar se deve ao fato do efeito de sombreamento sobre o instrumento ser indesejado.

3.1.3.3. Radiômetro de Radiação Líquida (Saldo Radiômetro)

Este sensor, como o próprio nome sugere, é responsável pelas medidas do saldo de radiação (ou radiação líquida). Esse instrumento é compreendido por uma termopilha situada entre dois discos negros. Um desses discos está orientado para cima e o outro, para baixo. Ambos os discos estão protegidos por abóbadas confeccionadas em polietileno, que permitem a transmissão da radiação que chega até elas, na banda espectral que se estende desde $0.35\mu\text{m}$ à $60.0\mu\text{m}$.

3.1.3.3. Termômetros

Os dois termômetros que se encontram na estação são os sensores responsáveis pelas medidas de temperatura. O registro efetuado pelo termômetro

de bulbo seco, se refere à temperatura do ar. O outro termômetro é responsável pela medida de temperatura do bulbo úmido.

O sensor de temperatura é confeccionado em resistência de platina, formando um braço de um circuito-ponte na "interface" que admite temperaturas a serem registradas numa gama de -10°C à 40°C , com uma tolerância de 0.6°C . Os termômetros estão inseridos num abrigo compacto, bem ventilado, a fim de que sejam protegidos da exposição direta à radiação solar, que lhes é prejudicial.

A depressão termométrica, é medida por dois termômetros dispostos num abrigo. Um dos termômetros tem seu bulbo envolvido por uma mecha de musselina embebida em água destilada. Esses sensores (termômetros) formam dois braços de um circuito-ponte cuja saída é proporcional à depressão termométrica.

3.1.3.4. Anemômetro

A velocidade do vento é medida por um anemômetro de três conchas que movimentam o cursor de um potenciômetro de giro contínuo, com redução de 3000 para 1. Velocidades do vento de até 110 kmh^{-1} podem ser registradas.

3.1.3.5. Catavento

A direção do vento é sentida por meio de um arranjo de 20 chaves magnéticas dispostas radialmente, ativadas por um ímã ligado ao eixo do leme do catavento.

3.2. MATERIAL E EQUIPAMENTO DISPONÍVEL

3.2.1. As Séries de Dados

Devido ao fato de não estarem em operação desde o início desse estudo, das sete estações meteorológicas automáticas de superfície listadas na tabela 3.1, somente foram consideradas as seguintes localidades: Manaus (MS), Fazenda Dimona (FD) e Reserva Ducke (RD). Hoje, todas as estações estão funcionando em caráter operacional.

3.2.1.1. Banco de Dados

Após serem coletados pelas estações meteorológicas automáticas de superfície os dados são armazenados e transmitidos em pequenos arquivos. Cada estação possui seus próprios arquivos. Os arquivos da Reserva Duke, tem suas iniciais no nome do arquivo, RD001.91. Os três dígitos subsequentes ao nome da estação definem o último dia de informações registradas no arquivo. Os dois dígitos mais à direita, chamados extensão do nome do arquivo, informam ao usuário a qual ano pertencem essas informações. Cada arquivo (mensagem) traz, geralmente, informações de dois a três dias consecutivos, podendo chegar em alguns casos a cinco dias. A tabela 3.3 descreve: o número da coluna a que pertence a informação dentro da mensagem; o nome da grandeza que a informação representa; a respectiva variável adotada para caracterizar a grandeza. Na tabela 3.4, são mostrados os dados observados para o dia de número 100 (10 de abril de 1992) referentes à estação localizada na cidade de Manaus.

Cada arquivo contém valores médios horários, calculados de um intervalo de amostras coletados a cada 10 segundos. O horário é dado em TMG (Tempo Médio de Greenwich), que para as estações de Manaus é de 4 horas à frente. A informação é armazenada em uma coluna, sua respectiva variação temporal é acrescentada na forma de linhas subsequentes. Todos os arquivos têm dezesseis colunas, (esse número é fixo). Por outro lado, a quantidade de linhas será determinada pelo número de observações horárias que forem feitas durante o dia (1 a 24 observações), bem como pelo número de dias, (ver tabela 3.4). Nas linhas finais do arquivo aparece outro tipo de informação: são as médias e os totais diários. Porém, neste trabalho, não foi feito uso desse tipo de informação para qualquer cálculo ou teste (médias diárias). Portanto esse tipo de informação não será apresentada na tabela 3.4. Trabalhou-se somente com as informações horárias, coletadas pelas estações meteorológicas automáticas de superfície.

Tabela 3.3 - Informações na ordem em que aparecem na mensagem

N° da coluna	DESCRIÇÃO DA INFORMAÇÃO	VARIÁVEL
1	Local	LL
2	Dia do ano	Dd
3	Hora da observação	Hh
4	Radiação solar global	Rg
5	Saldo de radiação	Rn
6	Temperatura do bulbo úmido	Tu
7	Temperatura do bulbo seco	Ts
8	Velocidade do vento	Vv
9	Direção do vento	Dv
10	Radiação difusa	Rd
11	Radiação refletida	Rr
12	Fluxo de calor no solo 1	G1
13	Fluxo de calor no solo 2	G2
14	Precipitação	Pp
15	Bateria	Bb
16	Ano	Aa

Tabela 3.2. Mensagem enviada pela estação de Manaus (MS), referente ao dia 100 do ano de 1991

LL	DD	HH	RG	RN	TU	TS	VV	DV	RD	RR	G1	G2	PP	BB	AA
3	100	100	-2.50	-27.02	24.01	25.83	0.08	335.20	-1.97	0.00	-12.92	-21.65	0.0	13.04	1991
3	100	200	-2.38	-19.30	23.78	25.00	0.04	10.29	-1.82	0.00	-12.77	-21.11	0.0	13.04	1991
3	100	300	-2.13	-14.29	23.81	24.85	0.04	5.98	-1.51	0.00	-12.20	-19.26	0.0	13.03	1991
3	100	400	-1.73	-10.24	24.34	25.19	0.08	256.70	-1.45	-0.01	-9.77	-14.71	0.0	13.02	1991
3	100	500	-1.36	-8.36	24.42	25.23	0.04	270.00	-1.45	0.00	-7.97	-11.91	0.0	13.01	1991
3	100	600	-1.31	-7.73	24.35	25.01	0.03	270.00	-1.40	0.00	-8.40	-12.70	0.0	13.01	1991
3	100	700	-1.20	-7.59	24.37	25.02	0.05	239.40	-1.38	-0.05	-8.56	-12.89	0.0	13.00	1991
3	100	800	-1.08	-5.65	24.41	25.05	0.02	234.80	-1.07	-0.05	-7.29	-11.02	0.0	13.00	1991
3	100	900	-0.96	-5.15	24.39	24.96	0.08	198.60	-0.98	-0.01	-6.51	-9.66	0.0	12.99	1991
3	100	1000	-0.83	-6.27	24.34	25.03	0.21	56.81	-1.01	0.00	-6.88	-10.61	0.0	12.99	1991
3	100	1100	70.70	43.27	24.54	25.56	0.10	112.40	42.24	11.60	-3.80	-5.32	0.0	13.03	1991
3	100	1200	224.90	157.80	25.03	27.30	0.17	109.20	113.20	37.44	3.57	5.94	0.0	13.19	1991
3	100	1300	326.50	228.30	25.41	28.97	0.36	264.20	220.40	55.77	10.43	22.11	0.0	13.27	1991
3	100	1400	489.00	343.40	26.03	30.51	0.76	252.30	298.50	85.00	18.99	43.56	0.0	13.32	1991
3	100	1500	497.00	360.50	26.27	30.54	0.98	275.90	304.70	84.70	18.48	39.02	0.0	13.67	1991
3	100	1600	751.00	553.80	26.65	32.30	1.51	276.30	374.40	129.90	33.12	62.42	0.0	13.50	1991
3	100	1700	505.20	350.50	26.02	30.44	2.66	286.60	332.50	88.30	13.91	19.34	0.0	13.40	1991
3	100	1800	514.50	351.80	25.93	30.23	2.62	288.30	282.60	89.60	18.71	24.24	6.4	13.14	1991
3	100	1900	16.94	-10.31	23.46	24.51	2.20	51.90	15.90	3.14	-37.63	-40.98	11.6	13.09	1991
3	100	2000	7.93	-8.66	22.44	23.40	0.81	2.17	7.10	0.92	-20.98	-28.14	0.6	13.06	1991
3	100	2100	16.05	-6.83	22.34	23.30	0.49	323.30	14.22	2.35	-13.65	-24.75	0.2	13.05	1991
3	100	2200	8.44	-11.81	22.65	23.50	0.77	312.70	7.56	1.43	-13.39	-24.30	0.0	13.05	1991
3	100	2300	-1.37	-17.98	22.56	23.42	0.77	310.90	-1.38	0.01	-14.48	-24.76	0.0	13.03	1991
3	100	2400	-1.60	-18.53	22.46	23.29	0.53	312.70	-1.45	0.00	-13.34	-22.94	0.0	13.03	1991

A partir dos dados coletados pelas estações meteorológicas automáticas de superfície (FD,MS,RD), foram gerados três bancos de dados que serviram de base para todo o trabalho realizado. Esses grandes bancos de dados foram gerados a partir do programa computacional, chamado de FUSÃO.BAS. Esse programa tem como objetivo pegar os vários e pequenos arquivos de mensagens com informações de dois ou três dias e agregá-los em um só grande arquivo. Os pequenos arquivos são ordenados de forma diária e sucessiva, compondo, assim, os chamados arquivos principais.

3.2.3. Erros Aleatórios

As séries de dados trabalhadas (neste período de tempo disponível) foram tomadas como sendo informações corretas, ou seja, supos-se que os dados tenham sido previamente checados e por ventura se necessário, corrigidos, antes de chegarem ao PDUS instalado junto ao IPH/UFRGS.

Para gerar os erros aleatórios e incluí-los nas séries dos dados, foram elaborados dois programas para computador, de nomes ALEAPO.BAS e ALEANE.BAS. Esses dois programas têm por finalidade gerar uma série de números aleatórios que são tomados como sendo o número correspondente à linha do arquivo principal ao qual o valor errôneo da informação será inserido. Após a determinação das linhas em que serão inseridos os erros, o programa localiza a coluna (ou seja, a variável), escolhida anteriormente pelo usuário, que receberá o erro. Pode se escolher qualquer variável para se inserir o erro: da coluna 1 à coluna 16 (desde a variável LL até a variável AA). Feito isso, o programa gerador de números aleatórios insere o erro nesta determinada informação. O tamanho do erro fica a critério do usuário e pode ser de 5%, 10%, 15% do valor da própria informação; ou, ainda do tamanho que o usuário julgar mais adequado.

No último passo é que se verifica a grande diferença dos programas ALEAPO.BAS e ALEANE.BAS. O primeiro deles (ALEAPO.BAS) tem a função de gerar os números aleatórios; localizar a informação escolhida; fazer o cálculo da porcentagem do erro e, por fim, somar esse erro à informação, tornando, assim, a informação que anteriormente era tida como correta, em uma informação acrescida da porcentagem sugerida pelo usuário (observação provavelmente suspeita). O outro programa (ALEANE.BAS) segue os mesmos passos do primeiro até o momento onde é feito o cálculo da porcentagem e, conseqüentemente, a determinação do erro a ser inserido. Porém, neste passo a diferença estará no momento de transferir esta porcentagem, que ao contrário do programa anterior, o ALEANE.BAS subtrairá a porcentagem da informação que anteriormente era tida como correta. Os erros gerados por esses dois programas computacionais supracitados, constituem-se em erros isolados. No decorrer do período do dia, escolhido para se inserir os erros, as possibilidades de que esse período receba apenas 1 (um) erro são grandes. Caracterizando assim o chamado "erro isolado". De forma manual foi criada uma pequena série de erros, chamada de "série de erros consecutivos crescentes", para aproximadamente 10 dias. Os erros inseridos, ao contrário dos anteriores, foram colocados de forma sucessiva e crescente dentro da série a ser trabalhada. Por exemplo, tomou-se a série de dados e colocou-se, de forma manual, 10% de erro na primeira observação, 15% na segunda, 20% na terceira e 25% na quarta observação. Isso para vários dias consecutivos. O horário inicial escolhido para se colocar os erros nas séries de dados foi o das 12:00 horas e o final o das 15:00.

3.3. MÉTODO ESTATÍSTICO

3.3.1. Observações Suspeitas

Dentro do método estatístico foi usada a técnica de "ANÁLISE DE REGRESSÃO", apropriada ao ajuste de retas para os dados que possuem resposta

contínua. Segundo BOX *et al*(1978), dois grupos de variáveis podem ser relacionadas entre si, uma sendo chamada de explanatória (variável independente) assumida como correta, e outra chamada de variável de resposta (variável dependente).

A relação entre estes dois grupos de variáveis depende dos valores de um grupo de parâmetros desconhecidos que são estimados por mínimos quadrados, que são, talvez, após a transformação das respostas, aqueles parâmetros que minimizam os desvios das somas dos quadrados, entre as observações e as estimativas do modelo. Na regressão linear simples, o modelo depende linearmente do vetor de parâmetros desconhecidos. Outra abordagem também pode ser feita, como, por exemplo, a regressão linear múltipla, em que a variável de resposta é explicada em cima de duas ou mais variáveis explicativas. Essa é uma ferramenta que tem sido muitas vezes empregada após o advento da programação em computadores eletrônicos (ATKINSON,1987). que possibilitam, através de terminais gráficos com alta resolução, explorar as propriedades e particularidades de muitas medidas. Para cada modelo ajustado uma grande variedade de quantidades podem então ser graficadas. Dentre outros objetivos, o recurso gráfico pode facilitar na observação do comportamento de um grupo de dados, no sentido de se tentar detectar informações suspeitas ou erradas.

As fontes de erros do ponto de vista matemático, podem ser diversas: a simples omissão de um sinal matemático, um ponto decimal, podem acarretar um grande erro, bem como a simples omissão de um dado.

3.3.2. O Método Estatístico

Partindo da idéia de alavancamento (leverage) exposta por HUBER (1981) e encontrada em ATKINSON(1987), pode-se desenvolver a álgebra dos mínimos quadrados e quantificar o alavancamento de dados suspeitos ou errôneos. Na figura 3.2, fez-se uma plotagem com valores colhidos pelos termômetros que medem a temperatura do bulbo seco e a temperatura do bulbo úmido, de horários

simultâneos e é fácil verificar que existe alguma coisa de suspeito na sexta observação pois ela foge à reta que, supostamente, poderia ajustar estes pontos. O resultado disso é a presença de um resíduo junto a esta observação.

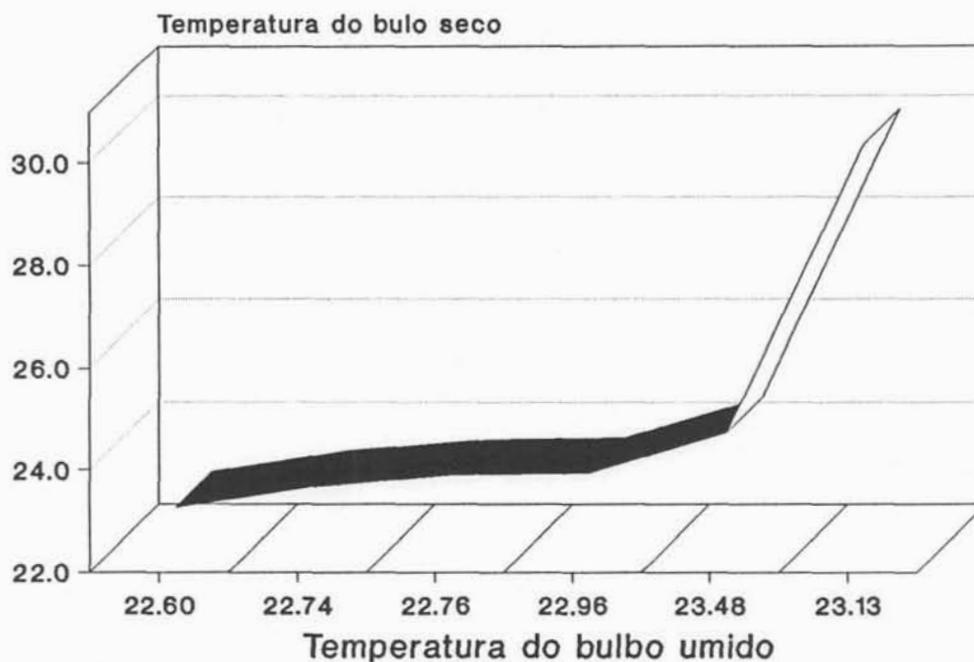


Figura 3.2. Temperatura do bulbo seco x Temperatura do bulbo úmido

Dando continuidade a esta idéia de alavancamento, ATKINSON(1987) sugeriu então uma demonstração matemática para mínimos quadrados (ANEXO I), que parte da seguinte expressão:

$$E(Y) = X\beta \quad (3.1)$$

onde, Y é o vetor de ordem $(n \times 1)$ das variáveis dependentes; $E()$ denota os valores esperados; X é a matriz de ordem $(n \times p)$ de variáveis independentes; β é o vetor de p parâmetros desconhecidos.

A matriz X foi chamada de "Matriz dos Carregadores" e é assumida como não sendo aleatória. Então, após a dedução dos mínimos quadrados, chegou-se à seguinte expressão:

$$\hat{y} = X(X^T X)^{-1} X^T y = Hy \quad (3.2)$$

onde, H é chamada de "Matriz Chapéu" (porque é esta matriz que estima, ou melhor, coloca o chapéu em y); y é a variável de resposta da regressão. Um aspecto importante a ressaltar aqui, é que a matriz H é quem determina o grau de alavancamento em que se trabalha.

A variável h_{ii} representa o i -ésimo elemento da diagonal principal de H (que aqui foi simplificado para h_i) e é expresso por:

$$h_i = x_i^T (X^T X)^{-1} x_i \quad (i = 1, \dots, n) \quad (3.3)$$

onde, h_i é uma medida de quão remota é a i -ésima observação no espaço dos X carregadores. Para observações com um grande alavancamento o valor de h_i é próximo a 1 (um) e o valor de \hat{y}_i é virtualmente determinado pelo valor de y_i .

A fim de que se tornasse mais fácil a compreensão da idéia do alavancamento, h_1 , elaborou-se um pequeno exemplo. Utilizou-se dados de temperatura coletados pelo termômetro de bulbo seco e termômetro de bulbo úmido, da estação meteorológica automática de superfície instalada na Fazenda Dimona, para o dia 19 do ano de 1991. Os dados são apresentados na tabela 3.5, com os respectivos horários locais.

Tabela 3.5. Dados de temperatura usados no exemplo estatístico

N°	HORA	TEMP. B. ÚMIDO	TEMP. B. SECO
1	8:00	22.60	22.95
2	9:00	22.74	23.36
3	10:00	22.76	23.58
4	11:00	22.96	23.62
5	12:00	23.48	23.91
6	13:00	24.13	25.91

Para o exemplo onde foi usado um carregador, pode-se escrever da seguinte maneira:

$$E(y) = \beta_0 + \beta x_1 \quad (3.4)$$

Desenvolvendo-se a equação h_1 , 3.3 chegou-se à

$$h_i = 1/n + (x_i - \bar{x})^2 / \sum (x_i - \bar{x})^2 \quad (3.5)$$

onde, $\bar{x} = \sum x_i/n$ é a média aritmética de x_i , e se $x_i = \bar{x}$, h_i terá o valor mínimo da matriz H, $1/n$. A medida que x_i aumenta em relação a \bar{x} , da mesma forma ocorre um acréscimo em h_i .

Fica claro, através da figura 3.2 que a sexta observação é a mais remota em relação ao resto do conjunto dos dados. A matriz chapéu (H) aparece na tabela 3.6 e se constata que " h_6 " (o quinto elemento que compõe a diagonal principal da matriz H) tem o menor valor dos h_i , e, também que $x_6 = \bar{x}$. Ressalta-se porém que nas observações 4, 3, 2, e 1 os valores de h_i estão crescendo constantemente até chegar em $h_1 = 0.2480$. Sendo o valor de h_6 0.9738, há uma indicação que esta observação tem um grande afastamento em relação as outras.

Tabela 3.6. Matriz H para o exemplo ilustrado.

$$\begin{pmatrix} 0.2480 & 0.2286 & 0.2181 & 0.2162 & 0.1787 & -0.0896 \\ 0.2286 & 0.2137 & 0.2058 & 0.2044 & 0.1758 & -0.0823 \\ 0.2181 & 0.2058 & 0.1992 & 0.1980 & 0.1743 & 0.0046 \\ 0.2162 & 0.2044 & 0.1980 & 0.1968 & 0.1740 & 0.0106 \\ 0.1787 & 0.1758 & 0.1743 & 0.1740 & 0.1684 & 0.1288 \\ -0.0896 & -0.0283 & 0.0046 & 0.0106 & 0.1288 & 0.9738 \end{pmatrix}$$

Uma maneira bastante prática de visualizar-se este comportamento de h_i , é fazer-se o uso de um gráfico em que se pode plotar a medida (h_i) contra sua posição (n_i), como é mostrado a seguir na figura 3.3.

Em alguns casos de dados a serem analisados, esta medida de alavancamento de h_i , mesmo após sua plotagem, pode não ser suficiente para ajudar no discernimento do ponto duvidoso. Isso pode ocorrer quando o maior h_i não se sobressair muito em relação aos demais. Para isso se pode fazer uso de um outro índice de plotagem, $h_i/(1-h_i)$, que é mais sensível a pequenas variações e cujo limite superior não é 1, como ocorre quando se plota com h_i . (ATKINSON, 1987)

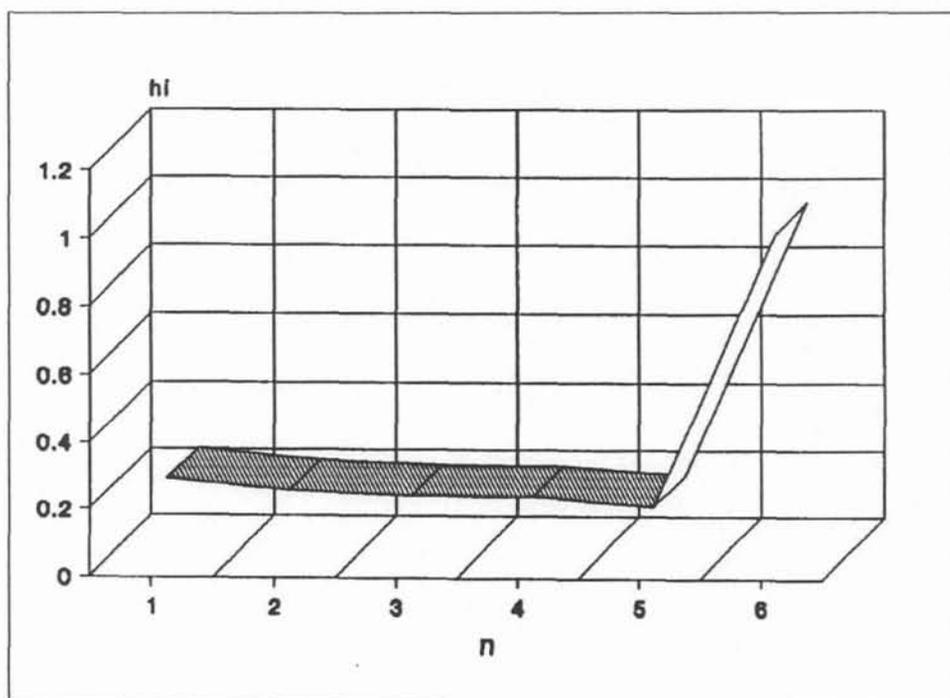


Figura 3.3. Plotagem do alavancamento (leverage) x Posição da observação (n)

O valor estimado \hat{y}_1 como pode ser visto pela tabela 3.6 depende razoavelmente das seis observações, ao passo que \hat{y}_6 depende, quase que exclusivamente, de y_6 . Como se percebe abaixo.

$$\hat{y}_1 = 0,2480y_1 + 0,2286y_2 + 0,2181y_3 + 0,2162y_4 + 0,1787y_5 - 0,0896y_6$$

$$\hat{y}_6 = -0,0896y_1 - 0,0283y_2 + 0,0046y_3 + 0,0106y_4 + 0,1288y_5 + 0,9738y_6$$

Fazendo-se $r_1 = y_1 - \hat{y}_1$, o vetor dos resíduos poderá ser representado por:

$$r = y - \hat{y} = y - X\hat{\beta} = (I - H)y \quad (3.6)$$

onde, I é a matriz identidade. Do mesmo exemplo podemos agora ver o efeito dos resíduos nas duas observações, novamente.

$$r_1 = -0,7520y_1 - 0,2286y_2 - 0,2181y_3 - 0,2162y_4 - 0,1787y_5 + 0,0896y_6$$

$$r_6 = 0,0896y_1 + 0,0283y_2 - 0,0046y_3 - 0,0106y_4 - -0,1288y_5 + 0,0262y_6$$

Em ATKINSON (1987), se encontram alguns comentários a respeito dos valores para h_i que podem ser considerados como "bons valores". SILVEY (1980) disse que os valores de h_i devem ficar próximos a média. Valores de $h_i > 2p/n$ são tomados por BELSLEY *et al.* (1980, p.17) como observações suspeitas, passíveis de uma investigação cuidadosa.

Os resíduos r_1 definidos em (3.6), não têm todos a mesma variância que $(I-H)(I-H)=(I-H)$ e portanto:

$$\text{var}(r) = \text{var} \{(I-H)y\} = (I-H)\sigma^2 \quad (3.7)$$

e, em particular,

$$\text{var}(r_i) = \sigma^2(1-h_i) \quad (3.8)$$

Por esse motivo, um resíduo com um alto alavancamento tem variâncias menores que aqueles pontos com alavancamentos menores. Isso pode ser visto com a comparação das expressões dos resíduos r_i e r_6 dada no exemplo acima.

Para encontrar resíduos com variância constante a fórmula supracitada é dividida pelo seu desvio padrão estimado. Faça-se S^2 a estimativa da média do quadrado dos resíduos de σ^2 , isto é,

$$S^2 = R(\hat{\beta})/(n-p) \quad (3.9)$$

O resíduo padronizado r'_i é definido como

$$r'_i = \frac{y_i - \hat{y}_i}{S \sqrt{(1 - h_i)}} \quad (3.10)$$

A vantagem de r'_i é que se o modelo $E(y) = X\beta$ estiver correto todos os resíduos terão a mesma variância. Eles são, portanto, particularmente apropriados para gráficos de probabilidade normal ou para plotagem, a fim de conferir a homogeneidade da variância. Uma leve desvantagem é que a nomenclatura para estes outros tipos de resíduos não é padronizada. Por isso HOAGLIN e WELCH (1978) e RYAN *et al.* (1986) chamaram r'_i de resíduo padronizado.

A primeira quantidade estatística interessante provém da consideração de que o apagamento da i -ésima informação tem um notado efeito na estimativa. Em particular a y_i observada concorda com a estimativa \hat{y}_i a qual surge quando a i -ésima observação não é usada em ajuste. Como y_i e \hat{y}_i são independentes a diferença,

$$y_i - \hat{y}_i = y_i - x_i^T \hat{\beta}_i \quad (3.11)$$

e tem variância

$$\sigma^2 \{1 + x_i^T (X_1^T X_1)^{-1} x_i\} \quad (3.12)$$

Para estimar σ^2 usou-se a estimativa de apagamento S_1^2 , que também é independente de y_i . O teste para o ajuste dos valores estimados e observados tem uma distribuição t com $n-p-1$ graus de liberdade.

$$r_i^* = \frac{y_i - x_i^T \hat{\beta}_i}{S_1 \sqrt{\{1 + x_i^T (x_i^T x_i)^{-1} x_i\}}} \quad (3.13)$$

Existem muitos nomes para r_i^* , alguns dos quais refletem a distribuição estatística t . BELSLEY *et al.* (1980), encontrado em ATKINSON (1987) usa o nome de RSTUDENT. Em comparação com o resíduo padronizado (r_i'), percebe-se, que a única diferença em relação a r_i^* está somente na estimativa de σ . A exemplo da nomenclatura sugerida para r_i' , para esse resíduo, não houve muito consenso em

ATKINSON (1987). COOK e WEISBERG (1982) chamaram r_1^* de residuo externo de student (externally studentized residual). Em ATKINSON (1981) ele chamou de resíduo "Jack-Knife".

Uma forma alternativa de r_1^* é fazer-se a substituição de S_1 , por outra expressão a qual é definida em ATKINSON (1987,p.21), que resulta na expressão apresentada logo a seguir (equação 3.14)

$$r_1^* = \frac{r_1'}{\left(\frac{n-p-r_1'^2}{n-p-1} \right)^{1/2}} \quad (3.14)$$

O cálculo de r_1^* , referente ao exemplo sugerido neste item, será apresentado em no próximo item (3.3.3)

3.3.3. O Programa Computacional do Método Estatístico

O método estatístico supracitado foi implementado em forma de um programa computacional, escrito em linguagem Basic, chamado de MODEST.BAS. Para o programa ser executado o usuário deve fornecer algumas informações iniciais, tais como:

- A) o nome do arquivo derivado a ser trabalhado.
- B) o dia inicial e o dia final do arquivo.
- C) a hora inicial e a hora final do período do dia cujas prováveis informações suspeitas, interessem ao usuário.
- D) quais as colunas de variáveis que serão trabalhadas (variável explanatória e de resposta).
- E) o valor (do teste t), com que serão testados os resíduos.

Com essas informações iniciais o programa começa sua execução e desenvolve o método estatístico. Para cada dia do arquivo lido pelo programa é calculada a "matriz chapéu" (matriz H) e logo após os resíduos "RSTUDENT", (r_i^*). Após a realização do cálculo dos resíduos para um dia (ou seja, o período do dia escolhido pelo usuário), o programa suprime a primeira hora de observação escolhida pelo usuário e calcula novamente a matriz H e os resíduos r_i , fazendo assim sucessivamente até o último horário do período escolhido dentro do dia. Feito isso, passa ao próximo dia e realiza a mesma operação anterior, até chegar ao último dia do arquivo escolhido pelo usuário.

Este programa fornece dois arquivos de saída, ambos com os resultados de r_i^* . A diferença é que um deles somente contém os valores dos resíduos que foram superiores ao limite superior imposto pelo teste t .

Na tabela 3.7 é ilustrada a saída, com o cálculo de todos os r_i^* . A primeira coluna que aparece é aquela referente ao cálculo dos r_i^* para todos os horários do dia do período escolhido pelo usuário. A segunda coluna, se refere aos resíduos de todo o período, porém com a ausência do primeiro par de observações (vide tabela 3.4): A terceira coluna, se refere aos r_i^* para todo o período com a ausência do segundo par de observações, mas com o primeiro par presente, e assim por diante. Então, para um dia de observações, o usuário terá o número de colunas de resposta para r_i^* (colocadas paralelamente uma em relação à outra) igual ao número de observações, acrescidas de uma coluna. Esses resultados diários aparecem para todos os dias que estão compreendidos entre o dia inicial e o dia final escolhidos pelo usuário, incluindo os dias limites.

Tabela 3.7. Tabela com todos os r_i^* calculados no exemplo

1°	2°	3°	4°	5°	6°	7°
-0,692	-0,975	-0,939	-0,697	-1,380	1,428	-0,886
-0,634	-0,448	-0,468	-0,313	-0,143	-0,256	-0,425
-0,009	0,091	0,100	-0,287	-0,090	-2,702	-0,394
4,686	3,571	3,575	9,385	3,132	0,259	0,212
-6,118	-4,690	-10,340	-4,805	-2,050	1,378	4,497
0,000	0,000	0,000	0,000	0,000	0,000	-5,827

O outro arquivo de saída é aquele que se refere aos r_i^* após terem sido testados pelo teste t. Todos os dados (r_i^*) - maiores que o limite estipulado pelo teste - são armazenados em forma de colunas. Nessa saída existem três colunas (tabela 3.8). A primeira se refere ao dia em que foi localizada a informação suspeita. A segunda coluna informa o horário desta informação suspeita, e finalmente, a terceira fornece o valor de r_i correspondente às duas informações anteriores. Essas colunas são paralelas. Nesse caso o teste foi feito com 3 graus de liberdade ($n-p-1$) e $t_{0,99}(4,54)$.

Tabela 3.8. Tabela com os r_i maiores que os limites testados

DIA	HORA	r_i^*
19	13:00	-5,827

3.4. MÉTODO FÍSICO

O método físico, a propósito do seu próprio nome, foi todo ele fundamentado em alguns princípios do comportamento físico e das variáveis medidas (coletadas) pelas estações meteorológicas automáticas de superfície.

Esse método foi implementado com bases em um outro método descrito por ROBERTS (1989) para ser usado em condições extratropicais, mais precisamente no Reino Unido. O Instituto de Hidrologia (IH/Wallingford-UK) desenvolveu esse método visando a atender sua necessidade de fazer um controle do uso da água, em quantidade e qualidade, daquela região.

O Instituto de Hidrologia desenvolveu esse método de controle de qualidade para os dados coletados pelas estações meteorológicas automáticas de superfície como parte integrante de um grande estudo, em que, a principal ênfase foi dada à quantificação da água disponível nas bacias hidrográficas. As investigações foram feitas para estimar as perdas por escoamento superficial das bacias hidrográficas (levando-se em consideração os diversos usos da terra) e, também, para modelar a contribuição da precipitação em cada uma das bacias estudadas. Para se fazer isso foi usado o princípio do balanço hídrico no qual aparecem quatro variáveis básicas que compõem a seguinte equação;

$$P = Q + E + \Delta S \quad (3.15)$$

onde: P é a precipitação que chega a bacia hidrográfica; Q é perda por escoamento superficial da bacia; E é a perda por evaporação ; ΔS é a variação no armazenamento da água na bacia hidrográfica.

Algumas destas variáveis não podem ser medidas diretamente, mas sim estimadas através de equações que englobam outros parâmetros medidos. É o caso do cálculo das perdas por evaporação de água para a atmosfera, que, dentre vários métodos, pode se calcular teoricamente, usando variáveis medidas, como: radiação; velocidade do vento; umidade. Aqui nesse ponto se faz necessário o uso dos dados coletados pelas estações meteorológicas automáticas de superfície, que proporcionam através de seus dados coletados, o cálculo de outros parâmetros, representado por E na equação 3.15, acima. Foram estudadas diversas bacias hidrográficas com diferentes características de solo, cobertura vegetal, níveis de desmatamento, intensidade de atividade agrícola, bem como particularidades onde além da precipitação em forma líquida foi estudada a contribuição da precipitação em forma de neve. Este último estudo foi feito em BALQHIDDER, Escócia central, no ano de 1981.

3.4.1. Descrição dos Testes do Método Físico

A seguir serão apresentados os testes que compõe o método físico, que foram usados e adaptados para a região da Floresta Amazônica, dentro do território brasileiro.

Em alguns testes, além do princípio físico, se levou em conta, também, o aspecto climatológico das variáveis, para se determinar as constantes limites. Por exemplo, no primeiro teste, onde se avalia a radiação solar global que chega à superfície da terra, para se estipular o limite máximo de 1100 W/m^2 primeiro se fez uma análise nos valores mais altos de radiação solar registrados. Em todas as séries trabalhadas, se separou os valores mais altos e a partir daí se estimou este limite. Com as outras variáveis testadas (onde também aparecem determinadas constantes) procedeu-se da mesma forma. Estipulou-se limites, considerando o seu comportamento climatológico.

É evidente que, a medida em que o número de anos de observações for aumentando e, por consequência disso, aumentar as extensões das séries dos dados geradas por estas estações meteorológicas automáticas de superfície, as probabilidades de se obter uma descrição climatológica mais precisa destes limites impostos à algumas variáveis será bem maior.

Na tabela 3.9 são apresentados os testes realizados pelo método físico. Na coluna da esquerda aparecem os testes físicos propriamente dito. Na outra coluna, a da direita, aparecem as mensagens de erro. Essa coluna da direita contém letras que caracterizam cada um dos testes, como se fossem uma espécie de rótulo. Então a medida em que se realizam os testes esses rótulos, atribuídos a cada um dos testes, somente aparecerão caso as desigualdades apresentadas na coluna da esquerda não forem satisfeitas.

Tabela 3.9. Descrição dos testes usados no Método Físico

TESTE DE CONTROLE	SAIDA
<p>1. Os valores da radiação solar global (R_g) são menores que um limite superior e maiores que um limite inferior</p>	<p>Caso seja falso, imprime "A", Dia: , Hora: .</p>
$-5 \text{ W/m}^2 < R_g < 1100 \text{ W/m}^2$	
<p>2. O valor da radiação solar global (R_g) é maior que o saldo</p>	<p>Caso seja falso, imprime "B", Dia: , Hora: .</p>

de radiação (R_N)

3. 95% da radiação solar solar é maior que a radiação difusa.

Caso seja falso, realiza a segunda parte do teste

$$(0.95 R_G) > R_d$$

A diferença absoluta entre radiação solar e radiação difusa é menor que um ruído

Caso seja falso a segunda parte do teste, imprime "C", Dia: , Hora: .

$$|R_G - R_d| < 10 \text{ W/m}^2$$

4. 25% da radiação solar global (R_G) é maior que a radiação refletida

Caso seja falso, realiza a segunda parte do teste

$$(0.25 R_G) > R_R$$

A diferença absoluta entre radiação solar e radiação refletida é menor que um ruído

Caso seja falso a segunda parte do teste, imprime "D", Dia: , Hora: .

$$|R_G - R_R| < 10 \text{ W/m}^2$$

5. O valor da radiação solar é maior que o valor médio entre os dois níveis de fluxo de calor do solo

Caso seja falso, imprime "E", Dia: , Hora: .

$$R_G > ((G_1 + G_2) / 2)$$

6. O valor da temperatura do bulbo seco é maior que a temperatura do bulbo úmido

Caso seja falso,
imprime
"F", Dia: , Hora: .

$$(T_u - T_s) \leq 0.1$$

7. O valor da velocidade do vento é maior que 0 m/s e menor que um limite superior.

Caso seja falso,
imprime
"G", Dia: , Hora: .

$$0 \text{ m/s} < V_v < 10 \text{ m/s}$$

8. O valor da direção do vento é maior que 0° e menor que 359°.

Caso seja falso,
imprime
"H", Dia: , Hora: .

$$0^\circ < D_v < 360^\circ$$

9. Os valores da precipitação são maiores que 0 mm e menores que um limite superior.

Caso seja falso,
imprime
"I", Dia: , Hora: .

$$0 \text{ mm} < P_p < 80.0 \text{ mm}$$

3.4.2. O Programa Computacional do Método Físico

Foi desenvolvido um "software" alicerçado em desigualdades físicas entre as variáveis que foram trabalhadas, coletadas pelas estações meteorológicas automáticas de superfície, e sub-sequentemente transcrito para uma linguagem computacional (BASIC) a fim de realizar-se os testes com os bancos de dados previamente selecionados. Este algoritmo em forma de programa foi denominado de MEFIS.BAS.

Para cada um dos testes foi atribuído uma letra (de A até I) com a finalidade de distingui-los entre si. Todos os testes foram realizados com dados horários de forma sucessiva, hora após hora.

A Rotina desenvolvida, acessa ao banco de dados previamente estabelecido, lê e realiza os testes de maneira que eles sejam executados um por vez, começando no teste A e terminando em I . Desta forma, se o teste a ser feito é o A, e nenhum erro é detectado automaticamente ele segue para o teste B, ao passo que se a desigualdade em A não for verdadeira, o programa acusa a não veracidade da desigualdade em A, bem como o horário em que ocorreu, e aí sim seguirá para a realização das desigualdades sub-sequentes.

3.5. Método de Penman-Monteith usado no cálculo da Evapotranspiração

A evaporação ocorre quando a água líquida é convertida em vapor de água e transferida nessa forma para a atmosfera. Para o caso de vegetação costuma-se denominar esta perda de água para atmosfera de evapotranspiração, uma vez que a água transferida para a atmosfera é decorrente da própria transpiração dos vegetais mais a evaporação do solo. Para o cálculo da evapotranspiração, foi usado o método de Penman-Monteith, que é apresentado por diversos autores, MONTEITH (1980), ROBERTS (1989) e SHUTTLEWORTH (1979).

Com esse método foi possível fazer o cálculo da evapotranspiração potencial horária, e a fórmula que possibilita esse cálculo é apresentada como

$$\lambda E = \frac{(R_n - G) + \frac{\rho C_p}{r_a} (e_s(T_s) - e)}{\Delta + \gamma (1 + r_s / r_a)} \quad (3.16)$$

onde, segundo ROBERTS (1989), λE é o fluxo de evaporação ($J s^{-1} m^{-2}$), Δ é a variação da pressão de saturação do vapor d'água em função da temperatura do ar ($mb \text{ } ^\circ C^{-1}$), ρ é a densidade do ar ($Kg m^{-3}$), C_p é o calor específico do ar a uma pressão constante, $e_s(T_s)$ é a pressão de saturação do vapor d'água (mb), e é a pressão real do vapor d'água (mb), γ é a constante psicrométrica ($mb \text{ } ^\circ C^{-1}$), r_s é a resistência superficial e r_a a resistência aerodinâmica (sm^{-1}).

O valor da pressão de saturação do vapor d'água no ar (e_s)(mb), foi gerado a partir da equação polinomial sugerida por WRIGHT (1982) e encontrada em CUNHA(1988), que é função da temperatura do ar (T_s), e é dada por

$$e_s = 6,105 + 4,44 \times 10^{-1} T_s + 1,434 \times 10^{-2} T_s^2 + 2,623 \times 10^{-4} T_s^3 + 2,953 \times 10^{-6} T_s^4 + 2,559 \times 10^{-8} T_s^5 \quad (3.17)$$

Para o cálculo da pressão real de vapor da água no ar (e)(mb), GATES (1980) sugere a seguinte equação

$$e = e_s - \gamma (T_s - T_U) \quad (3.18)$$

onde e_s é a pressão de saturação do vapor d'água e é função da temperatura do ar, γ é a constante psicrométrica, T_s é a temperatura do ar (termômetro de bulbo seco) e T_U é a temperatura do termômetro de bulbo úmido.

Para o cálculo da evapotranspiração, usou-se os valores da variação de pressão de saturação do vapor d'água em função da temperatura do ar, Δ ($\text{mb}^\circ\text{C}^{-1}$), tabelados em MONTEITH (1980). As constantes ρ , C_p e λ foram retiradas de ROBERTS (1989).

As resistências superficial e aerodinâmica, também estão descritas em ROBERTS (1989). Para a resistência aerodinâmica usou-se

$$r_a = \frac{[\log_e(Z - d / Z_0)]^2}{k^2 u(z)} \quad (3.19)$$

onde, Z é uma altura qualquer acima da superfície vegetada e d é conhecido como o deslocamento do plano zero. Esses valores de d normalmente variam entre 0.6 e 0.8 da altura do elemento rugoso (culturas). Z_0 é conhecido como o comprimento rugoso. Então, a expressão encontrada no numerador da equação 3.19 descreve o perfil logarítmico de comportamento do vento com a altura. A velocidade média horária do vento (ms^{-1}) é dada por u , e k é a constante de Von Karman. ROBERTS(1989) sugeriu alguns valores para a expressão $(Z - d / Z_0)$, em função da cobertura do solo no local do estudo, e que foram adotados nesse trabalho para o cálculo das resistências aerodinâmicas. São eles,

$$\begin{aligned} (Z - d / Z_0) &= 500 \text{ para a água} \\ &= 100 \text{ para gramado} \\ &= 5 \text{ para floresta} \\ k &= 0,41, \text{ e } u \text{ é a velocidade média horária do vento } (\text{ms}^{-1}). \end{aligned}$$

Para a resistência superficial ROBERTS(1989), sugere os seguintes valores

$r_s = 0$ para a água

= 0 para superfícies úmidas

= 40, 50, 70 (sm^{-1}) as 08:00, 12:00, 16:00 horas para gramado

= 80, 100, 140 (sm^{-1}) as 08:00, 12:00, 16:00 horas para floresta.

IV. RESULTADOS

4.1. Arquivos Principais

Com o auxílio do programa computacional FUSÃO.BAS, descrito anteriormente no item 3.2.1.1 foram gerados três grandes arquivos, denominados de arquivos principais, que serviram de base para a geração dos arquivos derivados que efetivamente testaram os métodos estatístico e físico.

Na tabela 4.1, são apresentados os arquivos principais gerados para as localidades de Manaus, Fazenda Dimona e Reserva Ducke. A primeira coluna contém o nome do arquivo, a segunda e a terceira informam os dias iniciais e finais, a quinta o número total de linhas que compõe o arquivo. A última coluna informa a que ano pertence o arquivo.

Tabela 4.1. Tabela dos Arquivos Principais, gerados a partir do FUSÃO.BAS

NOME	DIA INICIAL	DIA FINAL	N° LINHAS	ANO
MS.DAT	19	185	4008	1991
FD.DAT	20	185	3984	1991
RD.DAT	1	179	4296	1991

4.2. RESULTADOS DO MÉTODO ESTATÍSTICO

A partir dos três arquivos principais de dados, MS.DAT, FD.DAT e RD.DAT, foram gerados outros arquivos derivados desses. Cada um deles com suas características próprias, ou seja com erros controlados, de diversos tamanhos somados ou subtraídos dos valores originais e em variáveis diferentes.

As tabelas 4.2, 4.3, 4.4, mostram os arquivos derivados dos três arquivos principais. Na tabela 4.2, aparecem os arquivos gerados para a estação meteorológica automática de superfície localizada na Fazenda Dimona. Foram gerados seis arquivos a partir do FD.DAT. O tamanho dos erros para essa localidade foram de 10% a mais, em relação aos valores originais (ALEAPO.BAS). Esses erros, foram inseridos nas variáveis R_G, R_N, T_U e T_S, (colunas 4, 5, 6 e 7, respectivamente). Cada vez que um arquivo derivado é gerado, provavelmente esse arquivo, não será igual a outro, mesmo que ocorram as seguintes condições: tamanho do erro seja o mesmo; o período do dia escolhido e a variável a ser tabalhada também sejam as mesmas. A posição em que os erros serão introduzidos sempre será distinta, isto ocorre devido a aleatoriedade proposta pelo programa que gera os erros aleatórios (ALEAPO.BAS e ALEANE.BAS).

A tabela 4.3 mostra os respectivos arquivos gerados para a estação localizada na cidade de Manaus. Para esses arquivos também foram usadas as variáveis R_G, R_N, T_U e T_S (colunas 4, 5, 6 e 7) à exemplo da Fazenda Dimona. Porém as diferenças aqui residem no tamanho dos erros inseridos, que em dois arquivos derivados, MS3.DAT e MS4.DAT, foram de 5%. Outra diferença a ser ressaltada é que se fez uso, para essa localidade, do programa ALEANE.BAS, que tem por finalidade subtrair o erro de tamanho determinado pelo usuário, do valor original aleatoriamente escolhido, como descrito no item 3.2.3. Então aqui os arquivos derivados MS1.DAT, MS2.DAT, MS5.DAT e MS6.DAT, foram gerados

Tabela 4.2 - Arquivos derivados da Fazenda Dimona

ARQ. ORIGINAL	ARQ. ERROS	VARIÁVEL	PORCENTAGEM	ARQ. DERIVADO	PROGRAMA
FD.DAT	FDE1.DAT	Rg	10%	FD1.DAT	ALEAPO.BAS
FD.DAT	FDE2.DAT	Rn	10%	FD2.DAT	ALEAPO.BAS
FD.DAT	FDE3.DAT	Rg	10%	FD3.DAT	ALEAPO.BAS
FD.DAT	FDE4.DAT	Rn	10%	FD4.DAT	ALEAPO.BAS
FD.DAT	FDE5.DAT	Tu	10%	FD5.DAT	ALEAPO.BAS
FD.DAT	FDE6.DAT	Ts	10%	FD6.DAT	ALEAPO.BAS

Tabela 4.3 - Arquivos derivados da cidade de Manaus

ARQ. ORIGINAL	ARQ. ERROS	VARIÁVEL	PORCENTAGEM	ARQ. DERIVADO	PROGRAMA
MS.DAT	MSE1.DAT	Rg	10%	MS1.DAT	ALEANE.BAS
MS.DAT	MSE2.DAT	Rn	10%	MS2.DAT	ALEANE.BAS
MS.DAT	MSE3.DAT	Rg	5%	MS3.DAT	ALEAPO.BAS
MS.DAT	MSE4.DAT	Rn	5%	MS4.DAT	ALEAPO.BAS
MS.DAT	MSE5.DAT	Tu	10%	MS5.DAT	ALEANE.BAS
MS.DAT	MSE6.DAT	Ts	10%	MS6.DAT	ALEANE.BAS

Tabela 4.4 - Arquivos derivados da Reserva Ducke

ARQ. ORIGINAL	ARQ. ERROS	COLUNA	PORCENTAGEM	ARQ. DERIVADO	PROGRAMA
RD.DAT	RDE1.DAT	Rd	10%	RD1.DAT	ALEAPO.BAS
RD.DAT	RDE2.DAT	Rd	15%	RD2.DAT	ALEAPO.BAS
RD.DAT	RDE3.DAT	Rr	10%	RD3.DAT	ALEANE.BAS
RD.DAT	RDE4.DAT	Rr	20%	RD4.DAT	ALEANE.BAS
RD1.DAT	RDE5.DAT	Rr	10%	RD5.DAT	ALEAPO.BAS
RD2.DAT	RDE6.DAT	Rr	15%	RD6.DAT	ALEAPO.BAS

todos com 10% de erro, subtraídos do valor originalmente escolhido pelo programa ALEANE.BAS. Junto com essas tabelas (tabela 4.2, 4.3 e 4.4) aparece também a coluna chamada de "ARQUIVO DE ERROS", que é a segunda coluna das tabelas. Esses arquivos de erros aparecem cada vez que se gera um novo arquivo derivado. Este arquivo de erros informa a localização exata do erro, ou seja, informa o dia e a hora em que o erro foi inserido. Esta informação é de extrema importância, pois com ela podemos saber se os métodos estão detectando os erros com precisão ou não.

A tabela 4.4, caracteriza os arquivos derivados obtidos para a estação localizada na Reserva Ducke. Para essa localidade foram gerados também seis arquivos derivados, mas desta vez fazendo-se uso das variáveis Rd e Rr (colunas 10 e 11), com erros de 10, 15 e 20%. Nessa estação, em dois casos, RD3.DAT e RD4.DAT, fez-se a subtração dos erros (ALEANE.BAS). Nesse ponto, chama-se atenção à respeito da geração dos últimos arquivos da Reserva Ducke, RD5.DAT e RD6.DAT que foram gerados a partir de outros dois arquivos derivados pré-existentes, RD1.DAT e RD2.DAT, que continham erros inseridos na variável Rd (coluna 10). Com isso, então, foram gerados os dois últimos arquivos com erros em duas variáveis, Rd e Rr, diferindo de todos os outros gerados até aqui, que possuíam apenas erros em uma variável. Após a geração de todos arquivos derivados começou-se a aplicação do método estatístico, afim de se verificar qual a precisão e confiabilidade do método implementado. Com as primeiras execuções do programa MODEST.BAS e as respectivas análises dos primeiros resultados, percebeu-se logo a necessidade de distinguir-se os tipos de erros (resultados obtidos com o programa MODEST.BAS) a fim de que se pudesse fazer uma melhor análise. Os resultados foram caracterizados da seguinte forma:

a) ERRO TIPO I: Esse é o tipo de erro que foi inserido no arquivo derivado, e foi detectado pelo método estatístico, no dia e na hora de sua inserção. Esse é o resultado que quantifica realmente a eficiência do método estatístico proposto.

b) ERRO TIPO II: Esse é o tipo de erro que foi inserido no arquivo derivado, e foi detectado pelo método estatístico, somente no dia de sua inserção.

c) ERRO TIPO III: Esse tipo de erro foi considerado para os casos nos quais o método estatístico não detecta os erros associados ao dia e/ou hora de inserção, ou seja, quantifica o número de erros não detectados pelo método estatístico.

d) ERRO TIPO IV: Esse é o tipo de erro que ocorre quando, o método estatístico detecta e quantifica um determinado número de erros, mesmo que esses erros não tenham sido inseridos. Mesmo não havendo inserção de erros nos arquivos derivados o método estatístico acusa a detecção de informações suspeitas.

Desses tipos de erros descritos acima, os erros do tipo I, II e III são os mais importantes e foram os tipos de erros que procurou-se dar maior atenção, uma vez que são eles quem determinam a real eficiência do método estatístico, em função dos erros inseridos. Os erros do tipo IV ocorrem de maneira paralela, ou seja, independente da inserção de erros nas colunas, ele é um tipo de erro importante, e precisa ser monitorado. Porém nesse trabalho as atenções se voltam principalmente para os erros do tipo I.

As tabelas subsequentes 4.5, 4.6 e 4.7, referem-se aos resultados obtidos com o método estatístico (MODEST.BAS), em que foram usados os arquivos derivados apresentados anteriormente. Para tanto, padronizou-se os horários trabalhados. Escolheu-se para ser investigado o período do dia que estende-se das 12:00 até as 20:00 horas (TMG). devido, principalmente, para os valores das radiações que pouco antes das 12:00 horas e após as 20:00 apresentam valores muito pequenos (ver tabela 3.4) em relação aos valores máximos. Em muitos casos estas observações encontradas fora do período escolhido, apesar de corretas, acabariam tornando-se em valores suspeitos ou prováveis erros.

4.2.1. Resultados do método estatístico para a Fazenda Dimona

No caso da tabela 4.5, para a Fazenda Dimona, os arquivos começam no dia 20 e vão até o dia 185 do ano de 1991, com um total de 3984 linhas, mas como o período analisado compreende somente uma parte do dia (12:00 as 20:00 horas TMG), o número total de linhas é de 1485.

Após terem sido calculados os resíduos "RSTUDENT", assim denominado os r_1^* por ATKINSON (1987), foram realizados para todos os dias e os seus respectivos horários, o teste t (teste Student) com o propósito de discriminar os prováveis erros (informações suspeitas) contidas nas séries de dados trabalhados. Usou-se $n-p-1$ graus de liberdade, onde $n=9$ é o número de observações selecionadas dentro do dia, $p=2$ que é o número de variáveis carregadoras (com uma variável explanatória), e $p=3$ quando usou-se três variáveis carregadoras (com duas variáveis explanatórias). Para os testes onde usou-se somente uma variável explanatória, para o teste t de Student com $\alpha = 2.5\%$ e $\nu = 6$ graus de liberdade, usou-se $t\alpha = 2.365$, BOX *et al* (1978). Quando usou-se duas variáveis explanatórias com $\alpha = 2.5\%$ e $\nu = 5$ usou-se $t\alpha = 2.447$. Esses valores ($t\alpha$) foram usados também para as outras duas localidades. As porcentagens de desempenho, que aparecem na sétima coluna, analisadas e apresentadas nas tabelas de resultados (4.5, 4.6, 4.7) referem-se somente aos erros inseridos nas variáveis e detectados exatamente no dia e na hora em que ocorreram (erro tipo I).

Observando-se a tabela 4.5, poderemos analisar o desempenho dos testes realizados. Para o primeiro teste onde se usou 10% para o tamanho dos erros, e inseriu-se 30 erros na variável Rg (coluna 4), usando-se essa como variável explanatória, o desempenho foi de 76,6% de casos suspeitos detectados, tendo o método detectado 23 erros do tipo I, 3 do tipo II e 4 do tipo III. O erro do tipo IV representou 5,4% em relação ao número total de linhas, (1485 linhas) usadas nas análises. Este último erro oscilou para todos os testes entre 4 e 6,5%.

No segundo teste onde inseriu-se os erros de mesmo tamanho e número do anterior na variável RN (coluna 5) e usou-se essa de variável de resposta, o desempenho do método estatístico caiu para 56,6%, para erros do tipo I. Nos terceiro e quarto testes, usaram-se erros na duas colunas trabalhadas. Alternou-se as condições das variáveis Rg e RN de variável de resposta e explanatória, e os resultados permaneceram muito próximos, 62,8% e 60,0% respectivamente, para o erro tipo I.

Com o quinto teste obteve-se o melhor desempenho, do método, com os dados da Fazenda Dimona, Foram inseridos 25 erros na variável Tu (coluna 6), e usou-se essa de variável explanatória. Para a variável de resposta pegou-se Ts. O desempenho do método foi de 88,0% para os erros do tipo I. Por outro lado quando usou-se a variável Ts (coluna 7) com erros, como variável explanatória, no sexto teste, o desempenho caiu para 61,1% com os erros do tipo I. Entretanto, para os erros do tipo IV ocorreu o menor número de erros detectados, caracterizando assim em relação a esse tipo de erro (tipo IV) o melhor desempenho do método estatístico.

Para o sétimo teste, usou-se a variável Ts (coluna 7) com 10% de tamanho dos erros, e esta como variável explanatória. Foram inseridos dentro dessa coluna 19 erros e para testá-la usou-se Rg como variável de resposta. A eficiência do método estatístico para esse teste foi de 21,0% para os erros do tipo I, caracterizando assim o desempenho mais fraco para os teste realizados com os dados colhidos na Fazenda Dimona. A curiosidade verificada nesse teste é que o número de erros do tipo I foi igual ao tipo II.

Até o sétimo teste, percebeu-se que os melhores desempenhos ficaram com os testes onde se usou a coluna com erros de variável explanatória. No segundo teste onde usou-se o contrário, ou seja, usou-se a variável com erros de variável resposta, o rendimento dos testes caíram. Outro ponto a ser ressaltado é que onde se usou para a análise de regressão (com uma variável resposta) somente variáveis de comportamento parecidos (coluna 4 = radiação

solar global, coluna 5 = saldo de radiação, e também coluna 6 = temperatura do bulbo úmido, coluna 7 = temperatura do bulbo seco) o rendimento dos testes tenderam a ser melhores do que aqueles onde isso não ocorreu. A evidência desse comentário foi vista no teste 7, onde usou-se a coluna 7 (temperatura do bulbo seco) com a coluna 4 (radiação solar global) e o rendimento foi muito baixo.

Nos três últimos testes realizados com os dados coletados na Fazenda Dimona se fez o uso da idéia da regressão linear múltipla, onde usou-se três variáveis carregadoras (duas variáveis explanatórias). Nesses testes, usou-se a variável RN, com 30 erros inseridos (10% de tamanho), e a variável Rc, como variáveis explanatórias.

Tabela 4.5. Resultados do método estatístico para a Fazenda Dimona

N°	ARQ. ENT.	ARQ. T	ARQ. r_1 *	VAR. EXPL.	VAR. RESP.	REND. (%)	N° DE ERROS	I	II	III	IV
1	FD1.DAT	FT1.DAT	FR1.DAT	Rc	RN	76,6	30	23	3	4	81
2	FD2.DAT	FT2.DAT	FR2.DAT	Rc	RN	56,6	30	17	11	2	83
3	FD4.DAT	FT3.DAT	FR3.DAT	Rc	RN	62,8	35	22	7	6	83
4	FD4.DAT	FT4.DAT	FR4.DAT	RN	Rc	60,0	35	21	8	6	85
5	FD5.DAT	FT5.DAT	FR5.DAT	Tu	Ts	88,0	25	22	-	3	84
6	FD6.DAT	FT6.DAT	FR6.DAT	Ts	Tu	61,1	18	11	3	4	60
7	FD6.DAT	FT7.DAT	FR7.DAT	Ts	Rc	21,0	19	4	4	11	72
8	FD2.DAT	FT8.DAT	FR8.DAT	Rc-RN	RR	50,0	30	15	7	8	86
9	FD2.DAT	FT9.DAT	FR9.DAT	Rc-RN	Ts	63,3	30	19	6	5	97
10	FD2.DAT	FT10.DAT	FR10.DAT	Rc-RN	Rd	73,3	30	22	5	4	89

No oitavo teste, onde a variável RR (coluna 11) foi usada como variável de resposta, a eficiência do método para os erros do tipo I foi de 50%. No nono teste, usando-se a variável Ts (coluna 7) de variável de resposta o desempenho foi melhor em relação ao teste anterior. Para os erros do tipo I foi de 63,3%. No último teste, onde a variável de resposta foi Rd (coluna 10),

a eficiência do método superou aos dois testes anteriores. Para os erros do tipo I, foi de 73,3%. Com esses últimos testes viu-se também que a idéia de usar-se duas variáveis explanatórias pode ser um bom caminho na detecção dos erros.

4.2.2. Resultados do método estatístico para a cidade de Manaus

Com os dados coletados pela estação meteorológica automática de superfície localizada na cidade de Manaus, usou-se quatro arquivos derivados gerados com o programa ALEANE.BAS e dois arquivos derivados gerados com ALEAPO.BAS, como mostra a tabela 4.3, já citada anteriormente. Esses arquivos começam no dia 19 e vão até o dia 185 do ano de 1991, e tem um total de 4008 linhas.

No primeiro teste realizado com os dados de Manaus usou-se a variável R_c (coluna 4), com erros de 10% (ALEANE.BAS), de variável explanatória e R_n de variável de resposta. O desempenho do método estatístico para os erros do tipo I foi de 43%. No segundo teste usou-se as mesmas variáveis do teste anterior, porém com a variável R_n (coluna 5) contendo os erros e sendo a variável explanatória. Em relação ao teste anterior houve uma melhora sensível no segundo teste, apesar do baixo rendimento de ambos.

Para o terceiro e quarto testes, foram inseridos erros menores, com tamanho de 5% (ALAEAPO.BAS). A diferença entre os resultados desses testes foram bastante acentuadas. A eficiência do terceiro teste foi de 80% para os erros do tipo I, ao passo que o quarto teste caiu para 38,7%. A provável causa do bom desempenho do terceiro teste deva-se atribuir ao fato de terem sido inseridos na variável R_c (coluna 4) somente 5 erros.

No quinto e sexto teste, usou-se as variáveis T_v e T_s (colunas 6 e 7). Tendo a variável T_v, sido usada no quinto teste como variável explanatória e com 10% de tamanho para os seus erros. A variável T_s (coluna 7) apareceu no

sexto teste, também com erros de 10% e sendo a variável explanatória, ambas colunas com erros foram geradas pelo ALEANE.BAS. No quinto teste usou-se a variável Ts (coluna 7) como variável de resposta, e obteve-se o melhor desempenho do método estatístico para os dados da cidade de Manaus. Para os erros do tipo I a eficiência do método foi de 91%, dos 34 erros que foram inseridos o método localizou 31 (erros tipo I), 3 erros do tipo II, e nenhum do tipo III.

A partir dos resultados descritos até aqui já pode-se tentar induzir uma conclusão a respeito da eficiência do método estatístico. Pelos resultados de ambas localidades, Manaus e Fazenda Dimona, os desempenhos do método foram melhores para os testes realizados com uma variável carregadora. Os melhores resultados obtiveram-se com os dados das variáveis Tu e Ts (colunas 6 e 7), ou seja, temperatura do bulbo úmido e temperatura do bulbo seco. Este melhor desempenho, provavelmente deva estar associado as características que os dados dessas colunas apresentam. Os valores das temperaturas apresentam uma pequena variação no decorrer das horas. As variações da observação de uma hora para outra, muitas vezes são de apenas décimos de grau, ou poucos graus. Com isso a série de dados, das variáveis de temperatura do termômetro do bulbo úmido e bulbo seco, tornan-se mais uniformes fazendo com que qualquer acréscimo ou decréscimo nos dados seja percebido pelo método estatístico.

Tabela 4.6. Tabela com os resultados do método estatístico para a cidade de Manaus.

N°	ARQ. ENT.	ARQ. T	ARQ. r ₁ [*]	VAR. EXPL.	VAR. RESP.	REND. (%)	N° DE ERROS	I	II	III	IV
1°	MS1.DAT	MT1.DAT	MR1.DAT	Rg	RN	43,3	30	13	9	8	86
2°	MS2.DAT	MT2.DAT	MR2.DAT	RN	Rg	45,7	35	1	7	13	82
3°	MS3.DAT	MT3.DAT	MR3.DAT	Rg	RN	80,0	5	4	1	-	92
4°	MS4.DAT	MT4.DAT	MR4.DAT	RN	Rg	38,7	31	12	7	11	84
5°	MS5.DAT	MT5.DAT	MR5.DAT	Tu	Ts	91,0	34	31	3	-	89
6°	MS6.DAT	MT6.DAT	MR6.DAT	Ts	Tu	65,6	32	21	4	7	83

4.2.3. Resultados do método estatístico para a Reserva Ducke

Com os dados coletados pela estação meteorológica automática de superfície instalada na Reserva Ducke, foram gerados seis arquivos derivados (tabela 4.4). Os arquivos começam no dia 1 e vão até o dia 179 do ano de 1991. O tamanho dos erros inseridos nesses arquivos foram de 10, 15 e 20%. Quatro desses, arquivos derivados, foram gerados com o programa ALEAPO.BAS e dois com o ALEANE.BAS. Para essa localidade usou-se as variáveis R_c, R_d e R_r (colunas 4, 10 e 11) para se fazer os testes. Outra característica aqui, é que em dois arquivos derivados, RD5.DAT e RD6.DAT, existem duas variáveis com erros, R_d e R_r, com tamanhos de 10 e 15% respectivamente como mostra a tabela 4.4. Os arquivos derivados a partir de RD.DAT tem 4296 linhas, porém como o período trabalhado durante o dia é das 12:00 às 20:00 horas (TMG) o número total de linhas cai para 1611. Os erros do tipo IV para esses testes realizados oscilaram entre 4,9%, no terceiro teste, e 6,9% no segundo teste.

No primeiro e segundo teste, usou-se a variável (coluna 10) como variável explanatória com 10 e 15% de erros, respectivamente, e a variável R_c como variável de resposta. Verificou-se que com o aumento do erro inserido (15%) na variável R_d no segundo teste a eficiência do método também melhorou. Para os erros do tipo I no segundo teste obteve-se 26,9% ao passo que no primeiro somente 17,1%, embora o rendimento em ambos os casos tenha sido muito baixo. Verifica-se com isso que o método estatístico não demonstra ser muito sensível com erros relativamente pequenos inseridos na variável R_d (coluna 10). Notou-se também, que os erros do tipo II nesses dois casos foram maiores que os erros do tipo I.

Para o terceiro e quarto teste, usou-se a variável R_r (coluna 11), com tamanhos de erros de 10 e 20% (ALEANE.BAS), de variável explanatória e a variável R_c (coluna 4) de variável de resposta. A eficiência do método aumentou em relação aos dois primeiros casos. No terceiro teste obteve-se um rendimento de 45,5%, para os erros do tipo I, e o rendimento do teste 4 foi de

46,4%. Comparando-se com os dois primeiros testes, pode-se verificar que a medida em que aumenta-se os tamanhos dos erros da mesma forma obtém-se um aumento na eficiência do método estatístico.

No quinto e sexto teste, usou-se as variáveis R_d e R_R (colunas 10 e 11) como variável explanatória com tamanhos de erros de 10 e 15% respectivamente. Como variável de resposta usou-se novamente R_G (coluna 4). Os rendimentos dos testes, quinto e sexto, aumentaram em relação aos quatro anteriores. O rendimento do quinto teste foi de 51,5% enquanto que o sexto obteve-se 53,3%, ambos para erros do tipo I. Isso demonstra que para as variáveis R_d e R_R (colunas 10 e 11) o método estatístico pode ser mais eficiente usando-se duas variáveis explanatórias.

Tabela 4.7. Tabela com os resultados do método estatístico para a Reserva Ducke

N°	ARQ. ENT	ARQ. T	ARQ r_1 *	VAR. EXPL.	VAR. RESP.	REND. (%)	N° DE ERROS	I	II	III	IV
1	RD1.DAT	RT1.DAT	RR1.DAT	R_d	R_G	17,1	36	6	16	13	86
2	RD2.DAT	RT2.DAT	RR2.DAT	R_d	R_G	26,9	26	7	13	6	112
3	RD3.DAT	RT3.DAT	RR3.DAT	R_R	R_G	45,5	33	15	12	6	79
4	RD4.DAT	RT4.DAT	RR4.DAT	R_R	R_G	46,4	28	13	6	9	95
5	RD5.DAT	RT5.DAT	RR5.DAT	R_d-R_R	R_G	51,5	33	17	7	9	94
6	RD6.DAT	RT6.DAT	RR6.DAT	R_d-R_R	R_G	53,3	30	16	7	7	95

4.2.4. Resultados do método estatístico para a "série de erros consecutivos crescentes"

Neste sub-item se testou o método estatístico com a chamada "série de erros consecutivos crescentes", mencionada anteriormente no item 3.2.3. Criou-se de forma manual, duas séries de dados com erros consecutivos e crescentes dentro de um mesmo dia, diferindo assim da técnica de inclusão de erros usada até aqui.

A partir do arquivo principal MS.DAT, criou-se dois arquivos derivados contendo, em ambos arquivos, quatro informações errôneas (e sucessivas) em cada dia. Tomou-se do dia 50 até o dia 60. Os horários (TMG) e os tamanhos dos erros inseridos, nos dois arquivos, foram os seguintes: as 12:00 horas 10%; as 13:00 horas 15%; as 14:00 horas 20%; as 15:00 horas 25%.

No primeiro arquivo derivado (MS1.DAT) colocou-se os erros sucessivos e crescentes na variável Tu. Para o segundo arquivo derivado (MS2.DAT) a variável receptora dos erros foi Rc.

A) Resultados obtidos com o método estatístico usando MS1.DAT

Neste teste usou-se Tu de variável explanatória e Ts de variável de resposta. Os resultados foram bem animadores. Dos 11(onze) dias investigados, cada dia contendo 4 erros, o método estatístico detectou dez erros. Desses dez erros detectados, 2 pertenceram a um mesmo dia. Os nove erros restantes se distribuíram em um erro por dia. Para todos os dias que foram testados (e acusados o erro) o método indicou como sendo a observação suspeita, o valor referente ao horário das 15:00. Exatamente onde o erro parou de crescer. Somente para o dia onde o método detectou duas informações suspeitas os horários foram das 12:00 e das 15:00.

Apesar do método não ter detectado os 44 erros inseridos em 11 dias (ou seja, 4 erros por dia), dez dias foram localizados. Partindo dessa idéia o usuário tem a chance de localizar o dia com informações errôneas e a partir daí verificar cada horário de observação.

B) Resultados obtidos com o método estatístico usando MS2.DAT

Neste segundo teste usou-se R_c de variável explanatória e R_N de variável de resposta. Dos 11 dias investigados, cada dia contendo 4 erros, o método estatístico detectou 8 dias suspeitos com 9 erros. Desses 9 erros, 2 estão em um mesmo dia. Os outros 7 erros pertencem um a cada dia restante. A exemplo do teste anterior, as informações foram todas detectadas no horário das 15:00 horas. A exceção do dia onde se detectou dois resíduos, que foi no horário das 15:00 e das 16:00 horas.

4.2. RESULTADOS DO MÉTODO FÍSICO

A exemplo do que foi feito para se verificar e testar a eficiência do método estatístico; para o método físico também criou-se arquivos derivados com erros de tamanhos diversos e em variáveis distintas. As variáveis escolhidas para serem inseridos os erros foram as seguintes: variável R_N, R_c, T_u, T_s, R_d, R_R e G₂ (colunas 4, 5, 6, 7, 10, 11 e 13). As tabelas 4.8 e 4.9 mostram os arquivos derivados usados para os testes realizados com os dados da cidade de Manaus e da Fazenda Dimona e que foram gerados a partir dos arquivos originais MS.DAT e FD.DAT respectivamente. As tabelas indicam o nome do arquivo original, o nome dos arquivos onde estão os erros, associados com sua data e hora de ocorrências, a variável onde os erros foram inseridos, o tamanho do erro (%), o nome do arquivo derivado e por fim na última coluna o nome do programa usado para se gerar o arquivo derivado.

O tamanho dos erros usados para se gerar os arquivos derivados com os dados coletados pela estação meteorológica automática de superfície localizada na cidade de Manaus variou de 20 a 65% e para a Fazenda Dimona foi de 15 a 90%

Tabela 4.8. Arquivos derivados da cidade de Manaus

ARQ. ORIGINAL	ARQ. ERROS	VARIÁVEL	PORCENTAGEM	ARQ. DERIVADO	PROGRAMA
MS.DAT	MSE1.DAT	Rr	40	MS1.DAT	ALEAPO.BAS
MS.DAT	MSE2.DAT	Rr	35	MS2.DAT	ALEAPO.BAS
MS.DAT	MSE3.DAT	Rr	45	MS3.DAT	ALEAPO.BAS
MS.DAT	MSE4.DAT	Rn	55	MS4.DAT	ALEAPO.BAS
MS.DAT	MSE5.DAT	Rn	60	MS5.DAT	ALEAPO.BAS
MS.DAT	MSE6.DAT	Rn	65	MS6.DAT	ALEAPO.BAS
MS.DAT	MSE7.DAT	Tu	20	MS7.DAT	ALEAPO.BAS
MS.DAT	MSE8.DAT	Tu	30	MS8.DAT	ALEAPO.BAS

Tabela 4.9. Arquivos derivados da Fazenda Dimona

ARQ. ORIGINAL	ARQ. ERROS	VARIÁVEL	PORCENTAGEM	ARQ. DERIVADO	PROGRAMA
FD.DAT	FDE1.DAT	Rg	45	FD1.DAT	ALEANE.BAS
FD.DAT	FDE2.DAT	Rg	35	FD2.DAT	ALEANE.BAS
FD.DAT	FDE3.DAT	Rg	70	FD3.DAT	ALEAPO.BAS
FD.DAT	FDE4.DAT	Rd	50	FD4.DAT	ALEAPO.BAS
FD.DAT	FDE5.DAT	Rd	35	FD5.DAT	ALEAPO.BAS
FD.DAT	FDE6.DAT	Rr	45	FD6.DAT	ALEAPO.BAS
FD.DAT	FDE7.DAT	Rn	60	FD7.DAT	ALEAPO.BAS
FD.DAT	FDE8.DAT	Rn	55	FD8.DAT	ALEAPO.BAS
FD.DAT	FDE9.DAT	Tu	20	FD9.DAT	ALEAPO.BAS
FD.DAT	FDE10.DAT	Ts	15	FD10.DAT	ALEANE.BAS
FD.DAT	FDE11.DAT	Gz	90	FD11.DAT	ALEAPO.BAS

4.3.1. Resultados do método físico para a cidade de Manaus

Para os dados referentes a cidade de Manaus, foram realizados oito testes fazendo-se uso das variáveis RN, Tu e RR (colunas 5, 6, 11) com vários tamanhos de erros, como mostra a tabela 4.10, a seguir. Nessa tabela a primeira coluna informa o número do teste, a segunda informa o nome do arquivo derivado usado no teste. As três colunas seguintes referem-se ao número de erros inseridos na variável a ser investigada; o número de erros efetivamente detectados pelo método físico; o rendimento do método dentro de cada teste executado. Por último, aparece a coluna com o nome dos arquivos onde os resultados dos testes gerados pelo método físico foram armazenados.

Todos os testes realizados pelo método físico para essa localidade começaram no dia 19 e foram até o dia 185 do ano de 1991. O período do dia escolhido para ser investigado ficou compreendido entre as 12:00 e as 21:00 horas (TMG), inclusive.

Nos três primeiros testes realizados para essa localidade usou-se a variável RR (coluna 11) contendo os erros inseridos. O tamanho dos erros usados nesses testes foram 40, 35, 45% no primeiro, segundo e terceiro testes respectivamente. Para o primeiro teste o desempenho foi de 86,2%, dos 29 erros inseridos foram localizados 25. No segundo teste o rendimento foi de 76,4%, tendo esse 34 erros inseridos e 26 detectados pelo método físico. Com o terceiro teste, apesar desse ter o tamanho dos erros maiores, obteve-se um menor rendimento do método físico para os erros inseridos na variável RR em relação ao primeiro teste. Foram colocados 35 erros na coluna 11 e o método localizou 29 erros. O rendimento para esse último teste ficou em 82,9%. Nos três primeiros casos foi usado o teste "D", do método físico apresentado na tabela 3.9.

Para o quarto, quinto e sexto teste, a variável que recebeu os erros foi RN. Os tamanhos dos erros inseridos foram 55, 60, 65% respectivamente. No

quarto teste, dos 35 erros colocados em RN (coluna 5), somente 25 foram detectados. O rendimento do quarto teste ficou em 71,4%. A medida em que se aumentou o tamanho dos erros na variável RN o desempenho também aumentou. No quinto teste com 31 erros inseridos o método detectou 26 erros elevando o rendimento para 83,8%. Para o sexto teste usou-se 65% de tamanho para os erros e o rendimento foi o melhor deles para a coluna 5. Foram colocados 19 erros e localizados 16, elevando o rendimento do método para 84,21%. Nesses três casos se usou o teste "B" do método físico.

Nos dois últimos testes realizados com dados da cidade de Manaus, usou-se a variável Tu (coluna 6) como sendo a receptora das informações errôneas. No sétimo teste usou-se erros com tamanho de 20% inseridos na coluna 6. Dos 35 erros contidos nessa coluna, 28 foram detectados pelo método físico. O rendimento foi de 80%. No oitavo teste elevou-se o tamanho dos erros para 30% que foram colocados na coluna 6 e o rendimento foi de 100%. Para esses dois últimos casos usou-se o teste "F".

Tabela 4.10. Tabela de resultados do método físico para a cidade de Manaus

N°	ARQ. ENT	N° DE ERROS	ERROS DETECTADOS	RENDIMENTO (%)	ARQ. SAIDA	TESTE
1	MS1.DAT	29	25	86,2	MS1.FIS	D
2	MS2.DAT	34	26	76,4	MS2.FIS	D
3	MS3.DAT	35	29	82,9	MS3.FIS	D
4	MS4.DAT	35	25	71,4	MS4.FIS	B
5	MS5.DAT	31	26	83,8	MS5.FIS	B
6	MS6.DAT	19	16	84,2	MS6.FIS	B
7	MS7.DAT	35	28	80,0	MS7.FIS	F
8	MS8.DAT	13	13	100,0	MS8.FIS	F

4.3.2. Resultados do método físico para a Fazenda Dimona

Com os dados coletados pela estação meteorológica automática de superfície localizada na Fazenda Dimona, foram realizados doze testes com o método físico, como mostra a tabela 4.11. As variáveis investigadas nessa localidade foram RN, RG, TU, Ts, Rd e RR (colunas 4, 5, 6, 7, 10 e 11). Os arquivos derivados para essa localidade começam no dia 20 e vão até o dia 185, do ano de 1991. O período do dia escolhido para ser investigado foi o mesmo usado para a localidade anterior, 12:00 as 21:00 horas TMG.

Para os três primeiros testes realizados com os dados dessa localidade usou-se os tamanhos de erro de 45, 35 e 70%, inseridos na variável RN (coluna 4). Porém para os dois primeiros testes, o arquivo derivado foi gerado a partir do programa ALEANE.BAS, enquanto que o último arquivo derivado foi gerado pelo ALEAPO.BAS. Dos três primeiros testes os dois primeiros tiveram um bom desempenho; principalmente o primeiro teste. Esse teve um rendimento de 91,4%, dos 35 erros inseridos 32 foram detectados. No segundo teste, dos 11 erros inseridos 8 foram detectados pelo método físico, o que lhe valeu um rendimento de 72,7%. Por outro lado com o terceiro teste, arquivo derivado gerado a partir do ALEAPO.BAS, o rendimento foi baixo. Foram inseridos 29 erros e apenas 6 detectados. O rendimento desse teste ficou em 20,6%. Nos dois primeiros casos supracitados, os testes usados pelo modelo físico foram o "B", "C" e o "D", que aparecem na tabela 3.9. No terceiro caso foi usado o teste "A".

No quarto e no quinto teste, colocou-se os erros na variável Rd (coluna 10). Os tamanhos dos erros foram de 50 e 35% respectivamente. O quarto teste teve um melhor desempenho. Dos 29 erros inseridos, 17 foram detectados. O seu rendimento foi de 58,6%. O quinto teste teve um desempenho mais baixo que o anterior. Dos 29 erros introduzidos, apenas 16 foram detectados, ou seja, o método detectou um erro a menos. Seu rendimento ficou em 55,0%. O sexto teste, teve os erros introduzidos na variável RR (coluna 11) com um

tamanho de 45%. Dos 28 erros introduzidos 17 foram detectados. O rendimento para o sexto teste ficou em 60%. Nesses três casos usou-se o teste "C".

O sétimo e o oitavo teste foram feitos com a colocação dos dos erros na variável RN, com os tamanhos dos erros de 60% e 55%. Para o sétimo teste obteve-se o rendimento de 71,4%. Dos 35 erros introduzidos, 25 foram detectados. Com a diminuição no tamanho dos erros, o oitavo teste teve uma queda no desempenho em relação ao anterior. Dos 31 erros inseridos apenas 18 foram detectados. O rendimento para esse teste foi de 58,1%. Nesse caso usou-se o teste "B".

Para o nono e o décimo teste usou-se as variáveis Tu e Ts (colunas 6 e 7), com erros de 20 e 15%, inseridos nelas. Porém na variável Ts (décimo teste) usou-se o programa ALEANE.BAS. Para o nono teste o desempenho foi de 72% ao passo que para o décimo foi de 71,42%. Ambos rendimentos foram muito próximos. Provavelmente isto tenha ocorrido devido ao fato de que o teste que verifica essas variáveis é realizado em função delas mesmas, ou seja, a variável Ts é testada com a Tu e vice-versa. Nesses dois casos usou-se o teste "F".

O décimo primeiro teste foi realizado com a variável G2 (coluna 13). Foram inseridos 18 erros nessa variável com 90% de tamanho, e o método físico apenas detectou 5 desses erros. O seu rendimento ficou em 27,7%. Levando-se em conta o tamanho do erro inserido na coluna 13 e o número de erros detectados pode se considerar que este tenha sido um dos testes com um dos mais baixos rendimentos. Para esse caso usou-se o teste "E".

Tabela 4.11. Tabela de resultados do método físico para a Fazenda Dimona

N°	ARQ. ENT.	N° DE ERROS	ERROS DETECTADOS	RENDIMENTO (%)	ARQ. SAIDA	TESTE
1	FD1.DAT	35	32	91,4	FD1.FIS	B,C,D
2	FD2.DAT	11	8	72,7	FD2.FIS	B,C,D
3	FD3.DAT	29	6	20,6	FD3.FIS	A
4	FD4.DAT	29	17	58,6	FD4.FIS	C
5	FD5.DAT	29	16	55,0	FD5.FIS	C
6	FD6.DAT	28	17	60,0	FD6.FIS	C
7	FD7.DAT	35	25	71,4	FD7.FIS	B
8	FD8.DAT	31	18	58,1	FD8.FIS	B
9	FD9.DAT	25	18	72,0	FD9.FIS	F
10	FD10.DAT	35	25	71,4	FD10.FIS	F
11	FD11.DAT	18	5	27,8	FD11.FIS	E

4.3. RESULTADOS DO CÁLCULO DA EVAPOTRANSPIRAÇÃO USANDO O MÉTODO DE PENMAN-MONTEITH

Com o auxílio da equação de Penman-Monteith, descrita na seção 3.5, foi feito o cálculo da evapotranspiração com os dados coletados por duas estações meteorológicas automáticas de superfície. Uma delas localizada na cidade de Manaus e a outra na Reserva Ducke. Por uma medida de conveniência, a fim de se tentar minimizar "os erros" no uso dos valores, para a resistência superficial, sugeridos por ROBERTS(1989), adotou-se como regra para o cálculo da evapotranspiração, somente escolher os dias onde não ocorreu déficit hídrico no local das estações trabalhadas, ou seja, tomou-se dias em que estivesse chovendo no horário do cálculo ou que poucas horas antes desse

cálculo tenha chovido em uma quantidade razoável. O período do dia usado para se fazer os cálculos foi das 12:00 as 21:00 horas (TMG).

4.3.1. Resultados do cálculo da evapotranspiração para a cidade de Manaus

No cálculo da evapotranspiração para a cidade de Manaus, usou-se além do arquivo principal de dados daquela região (MS.DAT), um outro arquivo derivado. Esse arquivo derivado, contendo erros, foi chamado de MS1.DAT. Usou-se também o arquivo de erros, com a posição dos mesmos, MSE1.DAT que refere-se ao arquivo derivado (MS1.DAT) usado no cálculo da evapotranspiração.

O objetivo de usar-se o arquivo de erros, foi o de localizar dentro da série de dados, os dias em que encontravam-se as informações errôneas a fim de se fazer os cálculos. Na tabela 4.12 são apresentados os dias, em que foram realizados os cálculos, e os respectivos horários que contém as informações errôneas, bem como o tamanho dos erros inseridos na série de dados.

Tabela 4.12. Tabela com as informações do arquivo de erros gerados para a cidade de Manaus (MSE1.DAT).

DIA	VARIÁVEL	PORCENTAGEM	HORÁRIO
14	RN	50	16:00
27	RN	40	13:00
28	Ts	20	15:00
29	Ts	40	14:00

A partir da localização do dia que contém a informação errônea, fez-se o cálculo da evapotranspiração horária, incluindo-se o horário com essa

observação, usando-se o arquivo derivado MS1.DAT. A seguir repetiu-se o cálculo da evapotranspiração para esse mesmo dia porém usando-se a série de dados correta, ou seja, a informação que continha erros foi substituída pela informação correta (MS.DAT). A tabela subsequente (4.13) mostra os resultados obtidos com o cálculo da evapotranspiração para a cidade de Manaus. As colunas que aparecem, nessa tabela, refere-se ao dia escolhido para o cálculo da evapotranspiração e a hora. A terceira e a quarta coluna informam o resultado do cálculo da evapotranspiração horário com erro (mm.h^{-1}) e a evapotranspiração horária sem erro (mm.h^{-1}). A quinta e sexta coluna referem-se aos totais diários evapotranspirados, sendo que a quinta coluna é relativa ao cálculo da evapotranspiração diária incluindo-se o horário errôneo. A sexta coluna é o resultado do total diário calculado porém sem erros.

Tabela 4.13. Tabela com os resultados dos cálculos de evapotranspiração, horária e diária, para a cidade de Manaus

DIA	HORA	ETP C/ERRO mm.hora^{-1}	ETP S/ERRO mm.hora^{-1}	ETP C/ERRO mm.dia^{-1}	ETP S/ERRO mm.dia^{-1}
14	16:00	0,8544	0,6044	4,8036	4,5536
27	13:00	0,0036	0,0034	3,0418	3,0396
28	15:00	0,4317	0,2855	2,9785	2,8320
29	14:00	0,3254	0,2811	3,8144	3,7701

Como se pode ver pela tabela 4.13, para o dia 14, que conteve 50% de erro acrescido ao saldo de radiação (RN) para a observação das 16:00 horas (TMG), cálculo da evapotranspiração para esse horário obteve uma superestimativa de 41%. Porém quando se analisa esta observação no contexto diário o valor, superestimado, cai bastante. Para a evapotranspiração diária, com a inclusão do valor horário errado ocorre um acréscimo de 5% em relação ao cálculo diário com valores corretos.

Com o dia 28, onde usou-se a variável Ts, temperatura do ar, com erros houve também uma superestimativa da evapotranspiração horária em 51%, mas para a diária também caiu para 5%. No dia 29 também usou-se a variável Ts, e o acréscimo no valor do cálculo da evapotranspiração foi de 15%, para o cálculo horário e de apenas 1% para para o diário.

4.3.2. Resultados do cálculo da evapotranspiração para a Reserva Ducke

A exemplo do que foi feito no ítem anterior, para a Reserva Ducke também gerou-se um arquivo derivado do arquivo principal e a partir dele fez-se o cálculo da evapotranspiração horária e diária. Para a Reserva Ducke usou-se o arquivo principal RD.DAT, o arquivo derivado RD1.DAT, com o acréscimo nos valores corretos, e o arquivo de erros RDE1.DAT, que é apresentado, em parte, pela tabela 4.13.

Tabela 4.14. Tabela com as informações do arquivo de erros gerados para a Reserva Ducke (RDE1.DAT)

DIA	VARIÁVEL	PORCENTAGEM	HORÁRIO
8	RN	10	16:00
11	RN	20	15:00
12	RN	30	17:00
21	RN	40	13:00
29	Ts	30	17:00

Os resultados obtidos com o cálculo da evapotranspiração horária e diária para a Reserva Ducke aparecem a seguir na tabela 4.15.

Tabela 4.15. Tabela com os resultados dos cálculos de evapotranspiração, horária e diária, para a Reserva Ducke.

DIA	HORA	ETP C/ERRO mm.hora ⁻¹	ETP S/ERRO mm.hora ⁻¹	ETP C/ERRO mm.dia ⁻¹	ETP S/ERRO mm.dia ⁻¹
8	16:00	1,1054	1,0625	5,1267	5,0838
11	15:00	2,2680	2,1149	9,4887	9,3357
12	17:00	0,5369	0,4650	7,3680	7,2964
21	13:00	0,5368	0,4521	5,3467	5,2362
29	17:00	1,1004	0,9012	6,0122	5,8410

Para os dias 8, 11, 12 e 21, usou-se a variável R_n , saldo de radiação, contendo os erros como especificado pela tabela 4.13. Os erros oscilaram entre 10 e 40% respectivamente. O impacto deles no cálculo da evapotranspiração horária estendeu-se para esses dias de 4% para o dia 8 até 18% para o dia 21. A medida em que o erro aumentou na observações horárias, para esses casos, o erro no cálculo da evapotranspiração horária também aumentou, porém quando se calculou a evapotranspiração diária esses erros quase não se fizeram perceber, a oscilação da superestimativa ficou entre 1 e 2%.

Para o dia 29 usou-se a variável T_s , temperatura do ar, com erro de 30%. O impacto dessa informação errada no cálculo da evapotranspiração horária chegou a 22% a mais que o cálculo feito para o mesmo dia porém usando-se a informação das 17:00 correta. No cálculo do total diário o resultado com erro caiu. Superestimou o cálculo da evapotranspiração diária em 2%.

V. CONCLUSÕES

5.1. Conclusões para o Método Estatístico

O método estatístico demonstrou ter um bom potencial para a identificação das prováveis observações suspeitas. Seu desempenho foi bastante satisfatório no que diz respeito a identificação dos erros tipo I, quando se fez a análise de regressão com as variáveis de temperatura do bulbo seco e temperatura do bulbo úmido. Em seu melhor desempenho, para o caso descrito anteriormente, o método estatístico obteve um rendimento que chegou a 91,0%, tendo detectado 31 erros dos 34 inseridos. Por outro lado, em sua pior performance, o método estatístico obteve o rendimento de 61,0%.

Com as variáveis de radiação seu desempenho foi mediano. Verificou-se que o pior desempenho para o método estatístico foi quando se introduziu erros pequenos, 5%, nos arquivos derivados, (MS4.DAT/MR4.DAT). Para a variável da coluna 10 (radiação difusa) o método também não obteve muito sucesso, quando usada de variável independente no caso de somente uma variável independente, porém quando se usou esta coluna junto com a coluna 11, os resultados dos teste foram um pouco melhores.

A idéia de usar duas variáveis independentes mostrou-se também ser bastante interessante. Para os resultados da Fazenda Dimona onde usou-se os dados ad radiação refletida juntamente com os de radiação solar global e saldo de radiação o método se portou bem, detectando 22 erros de 30 inseridos.

O método estatístico deixou claro, que é um método sensível as pequenas oscilações que por ventura possam ocorrer no comportamento dos dados. Isto foi observado, no início do trabalho quando tentou-se trabalhar com o período de 24 horas do dia. Devido a descontinuidade no comportamento dos dados entre o dia e a noite o número de erros do tipo IV que apareceriam seria muito grande. Esse método pode detectar pequenos erros, de forma isolada.

O método estatístico tornou-se mais sensível, de uma maneira geral, a medida em que aumentou-se os tamanhos dos erros isolados.

Quando usou-se erros consecutivos (4 erros por dia) e crescentes, na série de dados, o método estatístico demonstrou um bom desempenho. Localizou grande parte dos dias (10/11 dias e 8/11 dias) com informações suspeitas, apesar de não ter localizado todas as informações horárias errôneas

5.2. Conclusões para o Método Físico

O método físico demonstrou sua maior eficiência para os casos onde os tamanhos de erros foram grandes. Em comparação com o método estatístico, o método físico, obteve rendimentos semelhantes com os tamanhos de erros maiores que aqueles usados no método estatístico.

O melhor desempenho do método físico foi com os dados de temperatura. Seus maiores rendimentos chegaram a 80 e 100%, quando se usou o teste "F". Que apesar de simples, demonstrou-se muito eficiente.

Com o teste "B", a medida em que se aumentou os tamanhos dos erros inseridos na coluna 5 (saldo de radiação), também aumentou os rendimentos.

Para o teste "C", ocorreu o mesmo, a medida em que aumentou-se o tamanho dos erros melhorou os rendimentos.

Quando se subtraiu os erros da coluna quatro, ou seja da radiação solar global, houve uma repercussão em três testes simultâneos. Nos testes "B", "C" e "D". E a exemplo do ocorrido em outros testes, quanto maior o erro melhor o rendimento.

5.3. Conclusões para o Cálculo da Evapotranspiração

A partir dos dados coletados pelas estações localizadas em Manaus e na Reserva Ducke, ficou claro que os impactos dos erros realmente se fazem sentir no cálculo da evapotranspiração horária. Os resultados horários obtidos em Manaus, superestimaram a evapotranspiração de 4 a 18%. Isso quando usou-se R_n com erros.

Porém quando as integrações diárias são feitas, levando-se em conta a informação errada, essas superestimativas quase não se fazem perceber. Quando se usou R_n essa superestimativa variou de 1 à 2%.

BIBLIOGRAFIA

- ANGUS, D.E. & WATTS, P. J.(1984). Evapotranspiration - How good is the bowen Ratio Method ? Agricultural Water Management, 1984 - 8 : pp 133 - 150.
- ATKINSON, A. C. (1987). Plots, Transformations and Regression: an introduction to graphical methods of diagnostic regression analysis. Imperial College, London. 282p.
- AUBERTIN, G. M. & PETERS, B. D.(1961). Net Radiation Determinations in a Cornfield. Agronomy Journal Madison, 53:269 - 272.
- BERLATO, M. A. & MOLION, L.C.B. 1981. Evaporação e Evapotranspiração. Porto Alegre, IPAGRO, Secretaria da agricultura - RS. 95p. (Boletim Técnico, 7).
- BOX G. E. P., HUNTER W. G. e HUNTER J. S. 1978. Statistics for Experiments and Introduction to Design, Data Analysis and Model Building. Wiley Series in Probability and Mathematical Statistics. J. Wiley e Sons. New York. 603p.
- BUDYKO, M. I. (1974). Climate and Life.Milwaukee. New York Academic Press. 1974 - 508p.
- CHANG, J. H. (1977). Climate and Agriculture An Ecological Survey. Chicago, Aldine Publishing. 1977, pp 163 - 178.
- CLARKE, R. T. & SILVEIRA, R. L. (1979). Instalação e uso de uma estação Meteorológica Automática em Porto Alegre e Análise dos dados Registrados pelos Aparelhos. ABRH, 1979. pp 91 - 99.

- CUNHA, G. R. (1988). Estudo Micrometeorológico da Transferência Vertical de Vapor D'água e Energia em Milho. Porto Alegre, UFRGS, Faculdade de Agronomia. Diss. Mestr. Agronomia, Fitotecnia. 1988 - 142f.
- CUNHA, G. R. (1991). Evapotranspiração e Função de Resposta a Disponibilidade Hídrica em Alfafa. Porto Alegre, UFRGS, Faculdade de Agronomia. Diss. Dout. Agronomia, Fitotecnia. 1991 - 198f.
- DUGDALE, G., HARDY, S. & MILFORD, J. R.(1991). V: Daily Catchment Rainfall Estimated from Meteosat. Hydrological Process. 1991 - 5 : pp 261 - 270
- HOAGLIN, D.C. and WELCH,R.E.(1978). The Hat Matrix in Regression ANOVA. Amer Statist. Assoc. 32, 17-22.
- LANDGREBE, D. A., DAVIS, M. S., PHILLIPS, T. L., SWAIN, P. H., HOFFER, R. M., LINDENLAUB, J. C., SILVA, F. L.(1978). Remote Sensing: The Quantitative Approach. Purdue University. West Lafayette.USA. 1978 - 396p.
- McNAUGHTON, K.G. & JARVIS, P. G. (1984). Using the Penman - Monteith Equation Predictively. Agricultural Water Management. 1984 - 8 : pp 263-278.
- MILLER, D. H. Weather at the Surface of the Earth: An Introduction to Ecosystem Hidrodynamics. Milwaukee, Academic Press. 1977, cap X - XII.
- MONTEITH. J. L.(1975). Vegetation and the Atmosphere. Academic Press Inc (London) LTD. 278p.
- MONTEITH. J. L.(1980). Principles of Environmental Physics. University of Nottingham School of Agriculture, Sutton Bonington. Loughborough, Leics. 1980. 241p.
- MORAES NOVO, E. M. L.(1988). Sensoriamento Remoto: Pricípios e Aplicações. INPE/MCT, São José dos Campos, 1988. 351p.

- PETERS, D. B. (1960). Relative Magnitude of Evaporation and Transpiration. *Agronomy Journal*
- ROBERTS, A. M. (1989). The Catchment Research Data Base at the Institute of Hidrology. Report N°106, May 1989. Institute of Hidrology. Wallingford, Oxon - UK. 52p.
- ROSENBERG, N. J. (1974). *Microclimate: The Biological Environment*. New York, John Wiley. 315p.
- RYAN, T.A., JOINER, B.L., RYAN, B.F. (1976). *Minitab Student Handbook*. North Scituate, Mass. Duxbury Press.
- SELLERS, W. D. (1965). *Physical Climatology*. University of Chicago Press. Chicago. 272p.
- SHUTTLEWORTH, W. J. (1979). Evaporation. Report N°56, July 1979. Institute of Hidrology. Wallingford, Oxon - UK. 61p.
- SLATER, P. N. (1980). *Remote Sensing: Optics and Optical Systems*. Addison-Wesley Publishing Company. Reading, Massachusetts. 575p.
- TANNER, C.B. & LEMON, E. R. (1962). Radiant Energy Utilized in Evapotranspiration. *Agronomy Journal*, Madison. 54:pp 207 - 212.
- TUCCI, C. E. M. (1991). *International Studies on Climate Change Impacts Uruguay River Basin*. Ministério da Educação - UFRGS - IPH. Agosto/1991.
- WMO. (1971). Technical Note N°83. Measurement and Estimation of Evaporation and Evapotranspiration. Geneve. 1971, pp 93 - 119.
- WMO. (1987). *Guidelines on the Quality Control of Data from the World Radiometric Network*. Leningrad, 1987. WCDP - 3. 30p.

WMO.(1989). Clicom Project. Climate Data Management System. April 1989. WMO/TD
- N°299. 14p.

WRIGHT, I. R., GASH, J. H. C., DA ROCHA, H. R., SHUTTLEWORTH, W. J., NOBRE, C.
A., MAITELLI, G. T., ZAMPARONI, C. A. G. P., CARVALHO, P. R. A.(1992). Dry
Season Micrometeorology of Central Amazonian Ranchland. Q. J. R. Meteorol.
Soc. 1992. 118, pp. 1083 - 1099.

ANEXO

$$E(Y) = X \beta \quad (A.1)$$

a i -ésima das n observações é

$$y_i = x_i^T \beta + \varepsilon_i \quad (A.2)$$

onde os erros ε_i tem esperança zero, variância constante e não são correlacionados

$$E(\varepsilon_i) = 0$$

$$E(\varepsilon_i) = \begin{cases} \sigma^2 & i=j \\ 0 & i \neq j \end{cases}$$

A estimativa $\hat{\beta}$ do parâmetro β minimiza a soma dos quadrados

$$R(\beta) = (y - X\beta)^T (y - X\beta) \quad (A.3)$$

e satisfaz a relação

$$X^T (Y - X\hat{\beta}) = 0 \quad (A.4)$$

$$X^T X \hat{\beta} = X^T y$$

então, a requerida estimativa dos mínimos quadrados é

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (\text{A.5})$$

Para este vetor de estimativas, o valor minimizado da soma dos quadrados é

$$\begin{aligned} R(\hat{\beta}) &= (y - X\hat{\beta})^T (y - X\hat{\beta}) \\ &= y^T y - y^T X (X^T X)^{-1} X^T y \\ &= y^T \{I - X^T X (X^T X)^{-1} X^T\} y \end{aligned} \quad (\text{A.6})$$

Uma representação alternativa enfatizando a dependência das soma dos quadrados dos resíduos somente de estimativa dos parâmetros de $y^T y$ e $X^T y$ é

$$R(\hat{\beta}) = y^T y - \beta X^T y \quad (\text{A.7})$$

O cálculo do resíduo da soma dos quadrados não requer cálculo de resíduos individuais

O vetor dos n valores prognosticados do modelo ajustado é

$$\hat{y} = X\hat{\beta} \quad (\text{A.8})$$

Da equação de $\hat{\beta}$ (A.5) isto pode ser escrito como

$$\hat{Y} = X(X^T X)^{-1} X^T y = Hy \quad (\text{A.9})$$