

Protocolos e Métodos para Tratamento Automático de Dados Micrometeorológicos

Andrêza L. de Alencar¹ & José L. Campos dos Santos²

¹ Faculdade de Sistemas de Informação – Universidade Federal do Pará (UFPA)

Av. Marechal Rondon s/n, Caranazal – Santarém – PA – Brasil

² Instituto de Pesquisas da Amazônia (INPA)

Av. André Araújo, 2936, Aleixo – Manaus – AM – Brasil

andreza@lbasantarem.com.br, lcampos@inpa.gov.br

Abstract. *The LBA Program provides a tower network instrumented with sensors, aiming the micrometeorological data sampling. The challenge includes the installation of mechanisms of automatic data sampling and to keep a methodological protocol to control data quality consistently collected. This work describes a set of protocol and methods which were implemented aiming an efficient system, helping directly to the quality control of the data collected. Quality data increases analyses and real scientific results, which can support the decision making process of maintenance and reducing the response time to failure of sensors.*

Resumo. *O Programa LBA mantém uma rede de torres instrumentadas com sensores com a finalidade de coletar dados micrometeorológicos. O desafio inclui instalar mecanismos de coleta automática de dados e manter um protocolo metodológico de controle de qualidade dos dados coletados consistente. Este trabalho descreve um conjunto de protocolos e métodos implementados objetivando um sistema eficiente, colaborando diretamente no controle da qualidade dos dados coletados. Dados de qualidade potencializam análises e resultados científicos verdadeiros, podendo apoiar na tomada de decisão nos procedimentos de manutenção e diminuir o tempo de resposta a falhas de sensores.*

1. Introdução

O Experimento de Grande Escala da Biosfera-Atmosfera na Amazônia (LBA) é uma iniciativa de pesquisa internacional de sucesso que em setembro de 2007 tornou-se um programa científico de governo, renovando a agenda de pesquisas iniciada em 1998 [Batistella *et al* 2007]. O LBA, sob a coordenação científica do Instituto Nacional de Pesquisas da Amazônia (INPA), é uma das maiores experiências científicas do mundo na área ambiental.

Durante os primeiros 10 anos de atividades, o programa teve importante papel na formação de recursos humanos formando mais de 500 mestres e doutores [Batistella *et al* 2007]. Na segunda fase do programa torna-se ainda mais prioritária a implementação de mecanismos para a aplicação dos resultados das pesquisas científicas no desenvolvimento sustentável da Amazônia. Assim, o plano científico corrente

consolidou as sete áreas iniciais de estudo do Programa em três grandes áreas integradas (AIs): (1) a interação biosfera-atmosfera, (2) o ciclo hidrológico e (3) as dimensões sócio-políticas e econômicas das mudanças ambientais.

Na primeira AI, o programa mantém uma rede de sensores de medidas micrometeorológicas através de Plataformas de Coletas de Dados (PCDs), instaladas nas torres de sítios de pesquisas em locais remotos da Amazônia, como: no Pará (Caxiuana e Santarém), no Amazonas (Manaus e São Gabriel da Cachoeira), Rondônia (Ji-Paraná e Ouro Preto do Oeste), e Mato Grosso (SINOP). A Figura 1 apresenta os locais que contam com recursos de torres instrumentadas mantidas pelo LBA.

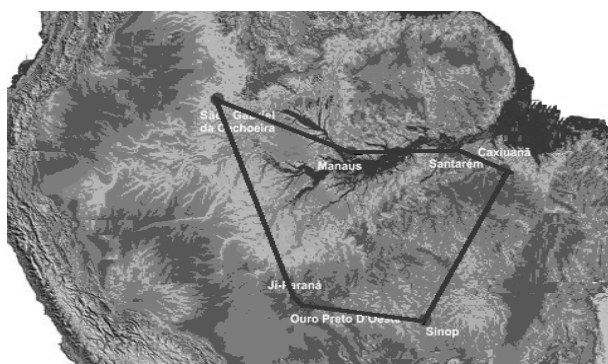


Figura 1: Mapa com a disposição dos sítios com torres instrumentadas de pesquisas do programa LBA.

Para os dados coletados exclusivamente por uma rede de sensores, por longos períodos e continuamente, onde são gerados grandes volumes de dados, enfrentam-se dois grandes desafios: (1) manter os novos dados sincronizados em uma seqüência com os dados já coletados anteriormente e (2) controlar a qualidade dos dados que possam garantir resultados científicos confiáveis além de diminuir o tempo de resposta a falhas de sensores.

As visitas as torres para obter os dados coletados, devido a complexa logística na Amazônia, podem ser efetuadas em intervalos de tempo longos e isso interfere na qualidade dos dados coletados como um todo, tempo de resposta a falhas de equipamentos, a falta de pré-análise dos dados, entre outros. Proporcionar um ambiente de Telemática para o monitoramento remoto e coleta de dados contribuirá para a diminuição do tempo gasto nas atividades de campo e no rápido diagnóstico de problemas proveniente dos equipamentos de coleta, aumentando potencialmente a qualidade dos dados [Serique *et al* 2007]. Assim, dispor de uma estrutura capaz de acelerar e melhorar o controle do monitoramento, coleta, análise e disseminação dos dados e informações dos sistemas de medição de fluxos de carbono torna-se um desafio do grupo envolvido na linha de pesquisa de Armazenamento e Trocas de Carbono do Programa LBA.

Uma rede de medições quando operada sob controle de um ambiente automático integrado, tem demonstrado capacidade para garantir a boa qualidade dos dados [Alencar *et al* 2008]. Como o processo de coleta envolve diferentes tecnologias de *software* e *hardware*, adotam-se protocolos metodológicos para evitar possíveis falhas durante o processo. Tais métodos têm sido extensivamente utilizados, principalmente em procedimentos automáticos de inspeção de séries temporais tendendo a reduzir qualitativamente os esforços de controle e promover produtos científicos consistentes

para a rede de medições [Foken *et al* 2004]. No entanto estes procedimentos automáticos não são desenvolvidos em um ambiente computacional flexível com recursos de programação para implementação de algoritmos de tratamento integrados e, portanto não representam um cenário de soluções ainda considerado ideal.

Desta forma, o desenvolvimento de um sistema computacional, incorporando tecnologia recente aos métodos de tratamento de dados automático promoverá a construção de um ambiente flexível. Este ambiente se torna possível através dos recursos da tecnologia Web 2.0, proporcionando interatividade e desenvolvimento modular ao ambiente, e de pacote de *software* Matlab¹ ou Scilab² como ambiente de implementação dos algoritmos de tratamento por possibilitar a integração de análise numérica, cálculo com matrizes, processamento de sinais e construção de resultados gráficos. Assim, o objetivo deste trabalho é demonstrar, através do tratamento automático, a eficiência dos métodos selecionados e a viabilidade de se implementar um sistema de controle de qualidade de dados ideal, ou próximo, para as torres de medidas micrometeorológicas do Programa LBA.

2. Ambiente de Telemática

Para sustentar a demanda crescente de coleta e análise de dados dos projetos do Programa LBA, foi necessário criar uma infra-estrutura para pesquisa de campo e fortalecer os laboratórios científicos envolvidos, em várias instituições brasileiras, além de estabelecer uma ampla rede de apoio logístico nos seis estados amazônicos e no Distrito Federal. Para o sucesso na gestão de dados e infra-estrutura foi implementada também uma política onde os dados e informações gerados pudessem ser gerenciados no componente de Sistema de Informações e Dados do LBA.

As crescentes necessidades de novas plataformas de coleta de dados automáticas, integrações com outros grupos de trabalho, no âmbito do Programa, e a necessidade de integrar os vários sítios de pesquisa, aumento da demanda de acompanhamento, análise e disseminação dos dados e informações provenientes destas plataformas de coletas de dados. Dentre as sete áreas de conhecimento que o Programa estuda, os estudos de fluxo de carbono têm demonstrado um grande desafio no que diz respeito ao acompanhamento, manutenção, qualidade e de análise dos dados, devido ao grande volume de dados produzidos pelas plataformas de coleta de dados destes experimentos.

No ano de 2005, foi instalado o grupo de telemática e de desenvolvimento de sistemas para monitoramento micrometeorológico em tempo real (telemetria). Essas ações geraram demandas para produção de dados de qualidade para uso em análises e geração de informações científicas produzidas pelos grupos de pesquisa do LBA. Tais dados são obtidos normalmente por diferentes metodologias como: coletas à superfície, em torres instrumentadas e observações em múltiplas escalas espaciais e frequências temporais por sensoriamento remoto envolvendo um amplo conjunto de satélites e

¹ O Matlab é um pacote de software interativo de alta performance que integra análise numérica, cálculo com matrizes, processamento de sinais e construção de gráficos em ambiente fácil de usar [Chapman 2003].

² Scilab é uma plataforma livre para computação numérica desenvolvido na Digiteo pelo Consorcio Scilab – www.scilab.org.

aviões instrumentados; e como as observações em diferentes escalas. A solução recomendada foi criar um Laboratório de Telemática capaz de integrar funcionalidades das torres de medidas de dados, (ver Figura 2), pesquisando e implementando soluções que aumentem a conectividade, a qualidade e a agilidade da disseminação dos dados e informações produzidas pelos grupos de pesquisa que necessitem de infra-estrutura computacional para o gerenciamento e aquisição de dados em tempo real.

Para a construção do ambiente de telemetria utilizamos uma sub-rede local de intranet para conectar as torres de medida e os *dataloggers*³ a um servidor, responsável pela coleta dos dados em tempo real e sua pré-análise. A pré-análise local objetiva satisfazer os requisitos de segurança dos dados primários (transmissão via satélite) e possível geração de dados derivados para alimentação de sistemas de monitoramento online. Esta sub-rede é implementada com cabos par-trançados (UTP) e conexão *wireless* para estabelecer comunicação entre as torres no sítio, alojamento e laboratório central. Conta-se também com uma conexão remota para viabilizar transmissão de dados a distância e promover comunicação entre os sítios, os escritórios regionais e central, através do uso de diversas tecnologias como: rádios para conexão de um servidor FTP através de satélite, antenas parabólicas para uso de internet banda larga por satélite e tecnologia GSM (telefonia celular) sendo cada uma destas aplicada de acordo com as particularidades dos sítios.

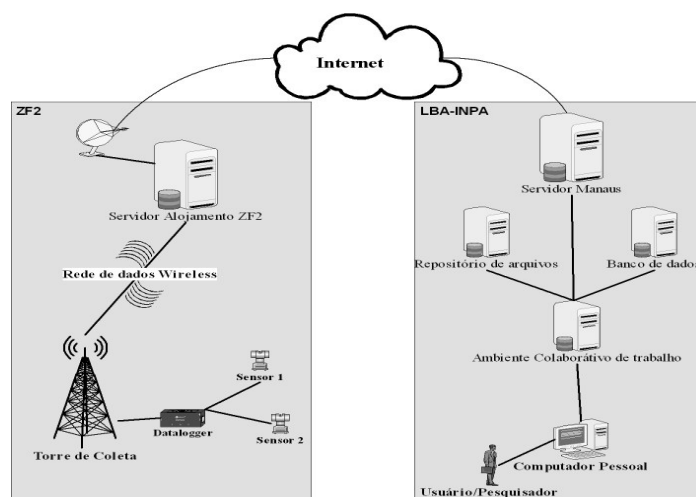


Figura 2: Localização e conexão das redes de dados entre as torres do sítio experimental ZF2 (S -2.60907, W -60.20917) em Manaus [Serique et al 2007].

3. Qualidade de Dados: Protocolos e Mecanismos de Tratamento

Na coleta de dados micrometeorológicos, apesar deste domínio possuir métodos cuidadosos e validados para aquisição de dados, ocorrem erros. O propósito do tratamento de dados é identificar e corrigir esses erros de forma rápida ou ainda minimizar seu impacto sobre o resultado dos estudos [Van de Broeck *et al* 2005].

Diagnósticos e tratamento de dados exigem uma visão detalhada sobre as fontes e tipos de erros em todas as fases de estudo, bem como durante e após a medição. A limpeza de dados deve ser baseada tanto no conhecimento de erros técnicos, como

³ Equipamentos registradores de dados.

falha/descalibragem dos sensores nas PCDs, como na variação de valores normais das observações que estão distante das restantes ou são inconsistentes com as mesmas que são habitualmente designadas por *outliers*.

Para alcançar os objetivos deste trabalho foram estudados os processos de aquisição de dados no contexto do projeto de telemetria e telemática, e protocolos de tratamento confirmando a validade dos mesmos através da implementação e uso de seus mecanismos de rastreamento, detecção, diagnóstico e edição de dados suspeitos de apresentarem anomalias.

3.1 Aquisição, Controle de Qualidade e Cálculo dos Fluxos

O Programa dispõe de vários pacotes de rotinas de aquisição e controle de qualidade de dados e de cálculo de fluxos desenvolvidos no próprio grupo ou adaptados de outros grupos de pesquisa. Para os dados de alta frequência e baixa frequência, o grupo utiliza três pacotes de programas, EddyRe⁴, Alteddy⁵ e Eddyflux_c⁶, dependendo da torre e do grupo de pesquisa que originalmente instalou os equipamentos. As torres possuem uma extensa diversidade de fornecedores de sensores meteorológicos provocando uma variedade de configuração dos sistemas nas torres. Todavia isso exige um maior esforço em treinamento, capacitação e tempo de processamento para sincronismo e cálculos de correção por parte das equipes envolvidas nas análises dos dados visto que estes procedimentos automáticos não são desenvolvidos em um ambiente computacional adequado com recursos de programação para implementação de algoritmos de tratamento integrados.

A. Tratamento de Dados Baseado em Método Semi-automático

A triagem de dados de baixa frequência é composta por procedimentos de recepção dos dados, todos no formato ASCII, organização tabular (nesse caso em tabelas MS Excel), preenchimento de falhas na coluna de tempo, preenchimento de células de dados vazias por “-9999” e exportação para as planilhas para padrão “csv” para que os dados sejam analisados por dois pacotes, um em MatLab e outro em Paradox. Estes terão a missão de identificar os dados espúrios, marcando-os e gerando outra matriz de dados. Esta matriz será tratada a um pesquisador *Ad-hoc* para verificação e certificação se os dados marcados com picos são falhas no processo de medida, equipamento descalibrado ou um evento extremo e natural. Uma terceira matriz com o parecer do *Ad-hoc* será gerada e repassadas para a equipe de processamento. Os dados considerados espúrios serão substituídos por “-9999”, gerando assim uma quarta matrix.

A quarta matrix, deverá ser encaminhada então para o grupo de preenchimento de falhas. Este grupo tem como função preencher todos os valores “-9999” por medidas calculadas através de métodos estatísticos.

⁴ EddyRe é um programa utilizado para processamento de Análise Espectral.

⁵ Alteddy é um programa Visual Fortran desenvolvido e utilizado em Alterra, Holanda, para o processamento de medidas de *eddy correlation* dos dados brutos – www.climatexchange.nl/projects/alteddy/

⁶ Eddyflux é um programa utilizado para processamento de cálculos de fluxos.

B. Tratamento de Dados Complexos por Algoritmos

A análise dos dados de alta frequência exige uma análise espectral dos dados brutos porque a técnica utilizada para os cálculos de fluxos de CO₂, vapor de água e balanço de radiação “eddy correlation”, é estatística e tende a camuflar problemas de picos, principalmente quando não se apresentam em grandes quantidades no intervalo de tempo estabelecido para calcular as médias, normalmente 30 minutos.

Alguns dos pacotes de medidas de *eddy correlation* possuem seus próprios pacotes de análise de dados brutos, todavia são pacotes fechados a alteração de parâmetros ou muito complexos. Já outros pacotes são desprovidos de qualquer ferramenta de análise deste tipo.

- **Análise Espectral dos Dados de Alta Frequência.**

Analisando o comportamento dos sinais no domínio da frequência, o arquivo de dados passa pelo teste de picos onde o programa calcula as relações entre temperatura, dióxido de carbono e vapor de água. Esses três resultados são projetados em um gráfico (Figura 3) e assim podemos comparar as curvas de CO₂ e H₂O em relação ao da Temperatura e fazer as correções para que as duas se aproximem o máximo possível da curva de temperatura, principalmente nas frequências maiores onde ocorrem os menores vórtices. A reta $-2/3$ é a normalização dos dados da temperatura na representação da tendência de sua declinação.

Por vezes, acontece nas mais altas frequências, acima de 1 Hz ou 3Hz, haver escapes ou perdas, o que é representado no gráfico como uma repentina alteração na tendência da curva. Quando essas alterações ocorrem muito próximo da frequência de 1 Hz, pode-se fazer um corte dos dados para que estes não interfiram no cálculo dos fluxos a serem calculados com o Eddyflux_c.

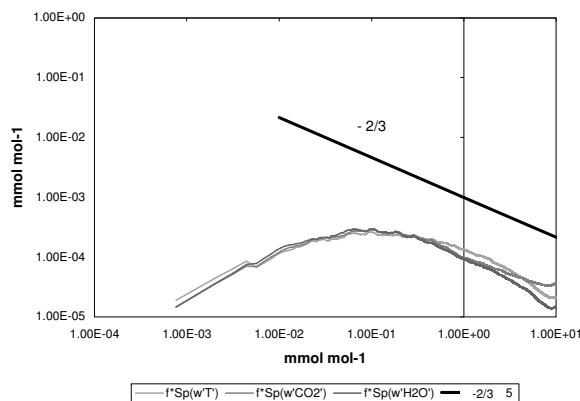


Figura 3: Representação da análise espectral em mmol (milimol) da temperatura (T), do fluxo de CO₂ (c) e do Fluxo de vapor de H₂O (H₂O). A reta $-2/3$ funciona como uma guia para a correção das curvas de CO₂ e H₂O, principalmente para as altas frequências.

C. Cálculo dos Fluxos pelo método da covariância de vórtices

O programa Eddyflux_c calcula os fluxos testando a qualidade dos dados brutos. O programa finaliza sua operação com um relatório em formato reconhecido pelo *software* Excel (Microsoft). Essa facilidade evita uma perda de tempo em relação a outros softwares de cálculo de fluxo existentes.

Suas principais características são: converte dados brutos em dados físicos; calcula as médias e variâncias, filtra as baixas ou altas frequências, usa um algoritmo linear para evitar perdas nas baixas frequências baseado em Gash & Culf (1996), calcula os fluxos de CO₂, de H₂O, de calor sensível e de momentum.

O programa também calcula dados adicionais, tais como: teste de estacionaridade dos dados baseado em Foken & Wichura (1996) considerando as influências das flutuações da densidade, calcula os *foot-prints* de acordo com Schuepp *et al.* (1990), corrige os fluxos de CO₂ e H₂O de acordo com Eugster & Senn (1995), integra as características das turbulências de acordo com Foken & Wichura (1996) e calcula a não-estacionaridade dos dados.

D. Estatística

A utilização de modelos matemáticos e técnicas estatísticas ligada às ferramentas computacionais possibilitam uma melhor análise dos dados. Para análise das variáveis a serem estudadas é realizada a relação dos dados através de regressão linear, regressão linear múltipla e análise de variância (ANOVA).

A regressão linear é um método para se estimar a condicional (valor esperado) de uma variável y , dados os valores de algumas outras variáveis x . Esta é chamada "linear" porque se considera que a relação da resposta às variáveis é uma função linear de alguns parâmetros.

A análise de variância (ANOVA) é uma coleção de modelos estatísticos, e os seus respectivos procedimentos, em que a observada variância é dividida em componentes, devido a diferentes variáveis explicativas. O objetivo da análise de variância é avaliar se as diferenças observadas entre as médias das amostras são estatisticamente significantes.

Para a análise comparativa entre as medidas feitas pelo sistema *eddy correlation* e pelas câmaras de análise de gás são feitos testes não paramétricos.

4. Ambiente de Implementação

Para determinar e entender a necessidade e a viabilidade dos requisitos para o sistema de *Data Cleaning* foram desenvolvidos os algoritmos de tratamento para o controle de qualidade de dados possibilitando a obtenção de soluções consistentes para o problema.

O pacote computacional Matlab foi escolhido para a execução deste trabalho pelas suas vantagens de programação. Dentre as muitas vantagens de se utilizar o ambiente Matlab pode-se citar o uso de linguagens de alto-nível; a implementação rápida e fácil; sintaxe simples e desestruturada; programação na linha de comando; tratamento, visualização e armazenamento de dados permitindo rapidamente implementações iniciais de algoritmos e procedimentos em uma série de aplicações, Figura 4. Tais propriedades fazem com que o Matlab seja um pacote de obtenção de resultados preliminares, e testes de algumas funcionalidades.

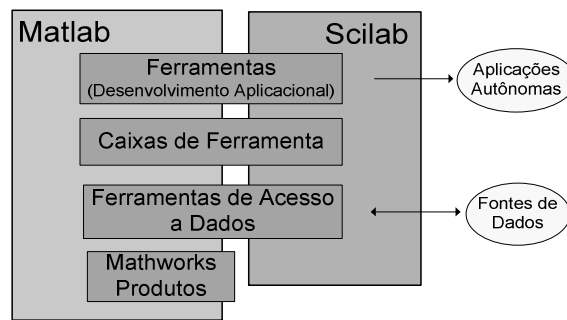


Figura 4: Diagrama comparativo das estruturas e funcionalidades dos pacotes Matlab e Scilab.

O Matlab é um sistema interativo cujo elemento básico de informação é uma matriz, que não requer dimensionamento. Esse sistema permite a integração de análise numérica, cálculo com matrizes, processamento de sinais e construção de resultados gráficos.

No Matlab os problemas e soluções são expressos da mesma maneira como são escritos matematicamente possibilitando efetuar os cálculos estatísticos relacionados às análises das variáveis estudadas.

4.1 ANOVA

Uma análise de variância permite que vários grupos sejam comparados a um só tempo, utilizando variáveis contínuas. O teste é paramétrico (a variável de interesse deve ter distribuição normal) e os grupos têm que ser independentes. Considerando uma variável de interesse com média μ e variância σ^2 tem-se dois estimadores de σ^2 :

$$S_B^2 = \text{dispersão entre os grupos (B)}$$

$$S_W^2 = \text{dispersão dentro dos grupos (W)}$$

e o teste é aplicado com:

$$F = \frac{S_B^2}{S_W^2} \quad (1)$$

A variância das médias amostrais (S^2) pode ser definida por:

$$S_{\bar{x}}^2 = \frac{\sum \bar{x}_i^2 - \frac{(\sum \bar{x}_i)^2}{k}}{k-1} \quad (2)$$

onde, k = número de grupos.

A Regressão linear, por exemplo, é um método para se estimar a condicional de uma variável y , dados os valores de algumas outras variáveis x . Para se estimar o valor esperado, usa-se de uma equação, que determina a relação entre ambas as variáveis.

$$Y_i = \alpha + \beta X_i + \epsilon_i \quad (3)$$

Em termos matriciais, de acordo com organização das tabelas de dados coletados pelas torres dos programa LBA, é preciso definir ainda:

$$\mathbf{Y}_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X}_{n \times p} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1,p-1} \\ 1 & X_{21} & \dots & X_{2,p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{n,p-1} \end{bmatrix} \quad \boldsymbol{\beta}_{p \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} \quad \boldsymbol{\epsilon}_{n \times 1} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Neste caso, o modelo de regressão linear geral é dado por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4)$$

Onde $\boldsymbol{\varepsilon}$ é um vetor de variáveis aleatórias independentes e normalmente distribuídas com esperança (média), $E(\boldsymbol{\varepsilon})=0$ e matriz de variância-covariância dada por:

$$\boldsymbol{\sigma}^2(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma^2 & 0 & \cdot & 0 \\ 0 & \sigma^2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sigma^2 \end{bmatrix}$$

Assim, o vetor das observações \mathbf{Y} tem esperança e variância dadas por:

$$\mathbf{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \quad \boldsymbol{\sigma}^2(\mathbf{Y}) = \sigma^2\mathbf{I} \quad (5)$$

$n \times 1$ $n \times n$

Através do sistema, o grupo de micrometeorologia pode efetuar os procedimentos e operações matemáticas necessários para a análise dos dados de forma automática possibilitando o alcance de resultados certificados e em tempo breve potencializando as análises e publicação de seus resultados. Este permite, ainda, que o pesquisador altere os parâmetros e variáveis a serem utilizadas na análise garantindo resultados de acordo com o esperado pela aplicação dos mesmos.

5. Considerações Finais

O trabalho concentrou no tratamento de dados micrometeorológicos, inserido em um projeto de telemetria e a integração dos ambientes instrumentados das torres nos sítios de pesquisas do LBA que, apesar de possuir métodos consolidados de aquisição de dados, ainda propiciam a ocorrência de erros que se propagam comprometendo resultados das pesquisas devido a natureza da complexidade de cada sítio de pesquisa.

A metodologia adotada neste trabalho incluiu o estudo de vários pacotes de rotinas de aquisição, controle de qualidade de dados e de cálculo de fluxos desenvolvidos no próprio grupo ou adaptados de outros grupos de pesquisa. Assim, observou-se que estes pacotes, apesar de possuírem suas rotinas de análises de dados, não permitem a alteração de parâmetros e outros não dispõem de funcionalidade para análise necessitando assim do desenvolvimento de ferramentas computacionais mais flexíveis a alteração de parâmetros.

No ambiente Malab os problemas e soluções são expressos da mesma forma em que eles são interpretados matematicamente possibilitando a análise dos dados através da utilização de modelos matemáticos e técnicas estatísticas de acordo com os parâmetros e procedimentos desejados.

A implementação dos procedimentos de análise confirmou que os métodos utilizados são eficientes no tratamento dos dados micrometeorológicos além de auxiliar na identificação de falhas dos equipamentos instalados nas torres, otimizando os artifícios de manutenção de sensores, equipamentos registradores de dados e sistema elétrico em tempo hábil de interveniência evitando, com isso, a perda de registros de prováveis fenômenos ambientais que estejam ocorrendo.

Referências

- Alencar, A. L. de. (2008). “Estudo e Implementação de Protocolos de Controle da Qualidade de Dados Micrometeorológicos”. Monografia – Universidade Federal do Pará - Santarém.
- Alencar, A. L. de; Campos dos Santos, J. L.; Maia, J. M. F. (2008). “Uma Prototipação de Sistema para *Data Cleaning* Automático de Dados Micrometeorológicos do LBA”. V Semana de informática, III Semana de Geotecnologias e II Encontro de Software Livre (SIGES 2008): Publicação Eletrônica.
- Batistella, M.; Alves, D.; Artaxo, P.; Bustamante, M.; Keller, M.; Luizão, F.; Marengo, J. A.; Martinelli, L.; Nobre, C. A. (2007) “Plano Científico Lba2 – Programa de Pesquisas Sobre Interações Biosfera-Atmosfera na Amazônia”. <http://www.lbasantarem.com.br/lba/inpa/lba>, Agosto.
- Chapman, Stephen J. (2003). “Programação em Matlab para Engenheiros”. Cengage Learning.
- Eugster, Werner and Senn, Walter. (1995). “A Cospectral Correction Model for Measurement of Turbulent NO₂ Flux”. *Boundary-Layer Meteorology* 74 (4), 321-340.
- Foken T, Göckede M, Mauder M, Mahrt L, Amiro BD, Munger JW. (2004). “Post-field data quality control”. In: Lee X, Massman WJ, Law BE (eds) *Handbook of micrometeorology. “A guide for surface flux measurements”*. Kluwer, Dordrecht, pp 181–208
- Foken, T., Wichura, B. (1996). “Tools for quality assessment of surface-based flux measurements”. *Agric. Forest Meteorol.* 78, 83-105.
- Gash, J.H.C. and Culf, A.D. “Applying a linear detrend to eddy correlation data in real time”. *Boundary-Layer Meteorol.* 79 (1996), pp. 301–306
- Maia, J. M.F.; Sá, M. de O.; Silva, P. R. T. da; Pereira, Q. C. C.; Aguiar, M. J. F. R.; Randow, C. V.; Santos, J. L. C. dos; Silva, J. T. da; Manzi, A. O.; Araújo, A. C. de; Alencar, A. L. de, (2007). “Implementação de protocolos de controle da qualidade de dados micrometeorológicos”. VI Semana de informática, II Semana de Geotecnologias e I Escola de Software Livre (SIGE): Publicação Eletrônica.
- Muller; H., Freytag; J. C. (2003). “Problems, Methods, and Challenges in Comprehensive Data Cleansing”. Humboldt Universitat zu Berlin, Germany.
- Serique, K. J. A.; Campos dos Santos, J. L.; Fortunato, J. M. (2007). “Uma arquitetura de Telemática para o monitoramento remoto de uma rede de PCDs meteorológicas”. VI Semana de informática, II Semana de Geotecnologias e I Escola de Software Livre (SIGE 2007): Publicação Eletrônica.
- Schuepp, P.H. ; Leclerc, M.Y. MacbPherson J.I. and Desjardins, R.L., “Footprint prediction of scalar fluxes from analytical solutions of the diffusion equation”. *Boundary-Layer Meteorol.* 50 (1990), pp. 355–373.
- Van de Broeck, J.; Argeseanu Cunningham, S.; Eeckels, R.; Herbst, K. (2005). “Data cleaning: Detecting, diagnosing, and editing data abnormalities. *PLoS Med* 2(10):e267.