

**Detecção de Anomalias e Estimação da Qualidade dos
Dados do Sistema de Telemetria e Monitoramento
Hidrometeorológico**

Roberto River Ferreira

Tese de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais.

Área de Concentração: Engenharia de Potência

Linha de Pesquisa: Sistemas de Energia Elétrica

Orientador: Prof. Peterson de Resende

**PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA – PPGE
UNIVERSIDADE FEDERAL DE MINAS GERAIS – UFMG
BELO HORIZONTE – MG**

2010

RESUMO

O objetivo deste trabalho é a criação de uma ferramenta computacional que forneça uma informação a respeito da qualidade de uma medição realizada remotamente, baseando-se na análise de dados estatísticos históricos do comportamento da grandeza ao longo do tempo.

Como resultado prático, esta metodologia foi aplicada no desenvolvimento de uma aplicação que critica os dados recebidos pelo STH (Sistema de Telemetria e Monitoramento Meteorológico) da CEMIG, baseando-se na análise exploratória de dados, na avaliação dos seus qualificadores, taxas de variação, limites históricos e limites físicos, de forma a aumentar a qualidade e a confiança nos dados recebidos pela plataforma de aquisição de dados.

A aplicação da ferramenta desenvolvida se mostrou de grande utilidade para os processos de tomada de decisão uma vez que fornece de forma direta um índice de qualidade a respeito de uma medição e oferece a capacidade de, indiretamente, criticar a qualidade das demais informações dos sistemas de supervisão e controle, como por exemplo, os dados calculados, que têm como origem as telemedições avaliadas pelo algoritmo de validação aqui proposto. Tudo isso dá a metodologia de tratamento de dados aqui apresentada uma abrangência sistêmica.

ABSTRACT

The objective of this job is to create a computational tool that provides information about the quality of a measurement performed remotely, based on analysis of historical statistics of the behavior of measurement over time.

As a practical result, this methodology was applied in the development of an application that criticizes data received by STH (Telemetry and Tracking System for Weather and Hydrology) of Cemig, based on exploratory data analysis, evaluation of their qualifiers, rates of change and historical and physical limits in order to increase the quality and confidence in data received by the data acquisition platform.

The implementation of the proposed tool proved very useful for decision-making processes as it provides a quality index for the measurements and offers the ability to indirectly inform about the quality of the other information of the system of supervision and control, for example, calculated data, which originate in the measurements criticized by the validation algorithm proposed here. All this gives the method presented here a systemic comprehensiveness.

SUMÁRIO

1.	INTRODUÇÃO	6
1.1.	Relevância e Contexto da Investigação -----	6
1.2.	Objetivo -----	8
1.3.	Metodologia de Desenvolvimento -----	8
1.4.	Organização do Texto -----	9
2.	STH – SISTEMA DE TELEMETRIA E MONITORAMENTO HIDROMETEOROLÓGICO	11
2.1.	Introdução -----	11
2.2.	Histórico -----	11
2.3.	Interface com o Usuário -----	12
2.4.	Base de Dados -----	13
2.5.	Cálculos Periódicos -----	14
2.6.	Monitoria de Elementos Hidrológicos -----	16
2.7.	Simulação -----	20
2.8.	Alarmes -----	21
2.9.	Consultas -----	22
2.10.	Conclusão do Capítulo -----	24
3.	PREPARANDO OS DADOS	26
3.1.	Introdução -----	26
3.2.	Histórico -----	27
3.3.	Primeiras Telemedições -----	28
3.4.	Faltou Sol -----	28
3.5.	Águas Nipônicas Versus Águas Brasileiras -----	29
3.6.	A Remota do STH -----	30
3.7.	Tratando as Inconsistências -----	30
4.	METODOLOGIAS DE QUALIFICAÇÃO DOS DADOS	32
4.1.	Introdução -----	32
4.2.	Regras do Negócio -----	32
4.3.	Análise Macroscópica dos Dados -----	33
4.4.	Análise Microscópica dos Dados -----	35
4.5.	Processamento dos Dados Recebidos -----	40
5.	DESENVOLVENDO OS ALGORITMOS	48
5.1.	Introdução -----	48
5.2.	Condições Existentes -----	48
5.3.	Estratégia de Desenvolvimento -----	49
5.4.	Tarefa de Pós-Processamento -----	51
6.	RESULTADOS INICIAIS ENCONTRADOS	53
6.1.	Introdução -----	53

6.2.	Avaliação das Medições de Temperatura -----	53
6.3.	Avaliação das Medições de Umidade Relativa do Ar ----	57
6.4.	Avaliação das Medições de Pressão Atmosférica -----	61
6.5.	Avaliação das Medições de Radiação Solar Acumulada --	63
6.6.	Avaliação das Medições de Velocidade Média do Vento -	67
6.7.	Avaliação das Medições de Direção do Vento -----	69
6.8.	Avaliação das Medições de Nível d'Água -----	72
6.9.	Conclusão do Capítulo -----	75
7. QUALIFICANDO MILHÕES DE MEDIDAS		77
7.1.	Introdução -----	77
7.2.	Condições Iniciais -----	77
7.3.	Em Busca de Uma "Boa" Amostra Inicial de Dados -----	78
7.4.	Síntese Quantitativa da Qualificação dos Dados -----	82
7.5.	Análise Qualitativa dos Resultados Obtidos -----	85
8. RESUMO ICONOGRÁFICO E RESULTADOS ENCONTRADOS		86
8.1.	Introdução -----	86
8.2.	Resumo do Comportamento das Medições de Temperatura do Ar -----	86
8.3.	Resumo do Comportamento das Medições de Umidade Relativa do Ar -----	92
8.4.	Resumo do Comportamento das Medições de Velocidade Média do Vento -----	94
8.5.	Resumo do Comportamento das Medições de Direção do Vento -----	97
8.6.	Resumo do Comportamento das Medições de Radiação Solar Acumulada -----	99
8.7.	Resumo do Comportamento das Medições de Pressão Atmosférica -----	101
8.8.	Resumo do Comportamento das Medições de Nível d'Água de Rio -----	104
8.9.	Resumo do Comportamento das Medições de Nível d'Água de Reservatório -----	106
8.10.	Resultado Prático da Qualificação em Tempo Real para o Usuário do Tempo Real-----	108
9. CONCLUSÕES E PROPOSTAS DE CONTINUIDADE		111
9.1.	Introdução -----	111
9.2.	Novas Utilidades para o Sistema de Qualificação ----	111
9.3.	A Qualidade do Dado com Foco no Cliente -----	112
9.4.	Propostas de Continuidade -----	112
9.5.	Palavras Finais -----	113
10. BIBLIOGRAFIA		114

1 Introdução

1.1 RELEVÂNCIA E CONTEXTO DA INVESTIGAÇÃO

O controle dos reservatórios das usinas da Cemig (Companhia Energética de Minas Gerais S.A.) é de responsabilidade do COS (Centro de Operação do Sistema).

A operação desses reservatórios tem se mostrado mais complexa a cada dia, sendo necessário lidar com um conjunto cada vez maior de variáveis de controle e exigindo a especialização crescente dos operadores dos sistemas de supervisão e controle. Para acompanhar as novas exigências do setor, as empresas têm buscado o desenvolvimento de ferramentas computacionais que agreguem em um mesmo ambiente todos os requisitos necessários para a operação desses reservatórios de uma forma segura, eficiente e ambientalmente responsável.

Visando uma maior agilidade no processo de decisão, fundamentado na aquisição de dados hidrológicos e meteorológicos, seguida do seu processamento e disponibilização aos usuários, foi desenvolvido na Cemig e inaugurado em 1998 o sistema computacional, STH – Sistema de Telemetria e Monitoramento Hidrometeorológico, com a finalidade de atender a uma crescente demanda interna, reforçada pela Resolução ANEEL 396/98. Desde então, tanto a infra-estrutura quanto o aplicativo vêm sofrendo manutenções evolutivas, buscando a adequação tecnológica, o desempenho e as funcionalidades requeridas pelas necessidades inerentes do próprio sistema e requeridas pelos seus usuários.

Inicialmente o aplicativo STH foi desenvolvido em uma plataforma cliente/servidor, possuía 84 UTR (unidades terminais remotas) coletoras de dados hidrológicos e meteorológicos e possuía em seu cadastro 50 usuários. A sua utilização pelos técnicos dos centros de operação era restrita à visualização de determinadas medições, conforme demanda dos próprios operadores. A partir de 2004, a infra-estrutura foi recebendo ampliações e hoje conta com 99

unidades terminais remotas, 15 concentradoras de comunicação de dados e uma unidade central de comunicação.

Paralelamente ao crescimento da plataforma do STH, o número de usuários cresceu significativamente, causando a necessidade de reestruturação do aplicativo que não havia sido desenvolvido inicialmente com o escopo de se tornar uma ferramenta corporativa. Para atender ao crescimento vertiginoso no número de usuários e para facilitar a administração do próprio sistema, sua arquitetura original foi migrada de uma proposta cliente/servidor para o ambiente da intranet corporativa, baseada em aplicações WEB. Nessa oportunidade foram agregadas novas funcionalidades ao sistema e o número de usuários passou de 500. Com essa evolução de arquitetura, o sistema passou a ter a capacidade de monitorar todos os reservatórios da Cemig que possuem UTR coletoras de dados em suas instalações, permitindo a integração de diversas áreas de planejamento, operação e manutenção, através de uma base única de dados hidrológicos e climatológicos, fundamentais para apoiar o planejamento energético necessário aos negócios da Cemig.

Com o aumento acentuado do número de usuários, incluindo aqui os usuários externos como o ONS (Operador Nacional do Sistema), CPTEC (Centro de Previsão do Tempo e Estudos Climáticos), ANEEL (Agência Nacional de Energia Elétrica), ANA (Agência Nacional de Águas), órgãos de defesa civil e prefeituras, dentre outros, a confiabilidade das informações telemedidas, principalmente no que diz respeito à qualidade da medição, tornou-se uma necessidade primordial. Desta forma, uma ferramenta, objeto foco deste trabalho, capaz de criticar ou pontuar uma telemedida recebida de uma unidade terminal remota, agregaria ao subsistema de telemedição um alto índice de confiança. Como as telemedidas formam a base para todo o processamento matemático realizado pelos processos mais especializados, o ganho alcançado com tal ferramenta teria uma abrangência sistêmica.

Ainda, como fonte de dados para outros sistemas corporativos da Cemig, os resultados desta proposta de trabalho poderão auxiliar e dar suporte de informações a diversos projetos de pesquisa e desenvolvimento (P&D) em curso, já aprovados na empresa, e na ANEEL, tais como o Desenvolvimento da Central de Diagnóstico da Manutenção da Geração e Transmissão e o de Utilização de Limites de Capacidade Dinâmicos em Linhas de Transmissão.

É neste contexto, abordando a importância da análise qualitativa dos dados recebidos pelo Sistema de Telemetria e Monitoramento Hidrometeorológico, que se localiza o desenvolvimento desta dissertação.

1.2 OBJETIVO

O objetivo final desse trabalho é criar uma ferramenta computacional que forneça, em tempo real, um índice que represente a qualidade de uma medição realizada remotamente (telemedição). Espera-se que o tratamento feito sobre o dado em análise permita indicar se tal dado é digno de confiança para ser usado pelos diversos outros processos ou se está tão contaminado por ruídos e erros que não possa ser usado e ainda, que indique se a cadeia de equipamentos de medição associada à telemedida necessita de aferição ou manutenção.

Espera-se também que, da análise dos resultados alcançados, decorram conclusões e propostas de outras aplicações, metodologias e projetos de pesquisa que venham a auxiliar às diversas áreas da Cemig que dependam de dados hidrológicos e meteorológicos para as tomadas de decisões.

Não se espera que o desenvolvimento da ferramenta proposta para a qualificação dos dados telemedidos tenha a complexidade daquelas ferramentas de estimação de estado dos aplicativos de análise de redes dos sistemas de gerenciamento de energia, com seus métodos iterativos de solução de matrizes e equações diferenciais. Nas premissas desse projeto, há uma busca pela simplicidade, a fim de se alcançarem outras como: eficiência computacional; facilidade de compreensão; manutenção simples; facilidade de acesso aos resultados e integração direta com aplicativos e ferramentas de domínio geral.

1.3 METODOLOGIA DE DESENVOLVIMENTO

A princípio, os dados básicos telemedidos, dos quais derivam uma enormidade de outros, e que deverão ser criticados ou pontuados pela ferramenta de qualificação proposta, são: precipitação pluviométrica (chuva), nível d'água de rios e de reservatórios, temperatura e umidade relativa do ar, velocidade média e direção do vento, radiação solar acumulada e pressão atmosférica. Como são grandezas hidrológicas e meteorológicas, associadas a

fenômenos naturais sob os quais não se tem controle, a modelagem matemática se torna extremamente difícil, senão impossível de ser realizada. Sendo assim, torna-se muito atrativa e até mesmo indicada a análise exploratória de dados e a utilização de métodos estatísticos para a solução do problema da qualificação em tempo real dos dados telemedidos pelo sistema de supervisão.

A ferramenta proposta baseia-se na base histórica de telemedidas, que está disponível, em sua maioria, numa base horária desde junho de 1998, e através da análise exploratória dos dados desta base, objetiva inferir em tempo-real, sobre a qualidade da telemedida recebida, criticando os valores medidos pelos sensores e transmitidos ao sistema.

1.4 ORGANIZAÇÃO DO TEXTO

Procurou-se evitar nesta dissertação a citação e a repetição no texto da abordagem de temas exaustivamente já discutidos e detalhados por um grande número de trabalhos da literatura técnica. Estes trabalhos que tratam das ciências das estatísticas, da hidrologia e da meteorologia, serão referenciados no texto quando necessário. Assim sendo, o texto desta dissertação ficará dedicado objetivamente aos desenvolvimentos e resultados da investigação específica em foco.

Nesta perspectiva, o texto desta dissertação foi distribuído em 10 capítulos, incluindo a presente introdução que compõe o **Capítulo 1**.

No **Capítulo 2** é feita uma apresentação sucinta do STH – Sistema de Telemetria e Monitoramento Hidrometeorológico e sua contextualização dentro das atividades do planejamento energético da Cemig.

O **Capítulo 3** apresenta a forma de organização dos tipos de grandezas disponíveis na base de dados histórica do STH e a forma de tratamento inicial dos mesmos para sua utilização no desenvolvimento da ferramenta de qualificação das telemedições.

O **Capítulo 4** aborda a estratégia a ser usada para a qualificação das telemedições, indicando os mecanismos a serem utilizados pelos algoritmos que serão desenvolvidos para a ferramenta de qualificação em tempo real.

O **Capítulo 5** apresenta os desenvolvimentos dos algoritmos, baseados nas estratégias traçadas conforme descritas no Capítulo 4. Espera-se construir inicialmente uma ferramenta *off-line* para a confirmação da eficácia dos mecanismos de tratamento de dados e posteriormente, validados os mecanismos, construir uma ferramenta que funcione de maneira *on-line*, em tempo-real.

O **Capítulo 6** apresenta alguns resultados encontrados durante os desenvolvimentos dos trabalhos de preparação inicial da base de dados histórica de grandezas hidrológicas e meteorológicas sobre a qual foram realizados os estudos.

O **Capítulo 7** apresenta os resultados encontrados após o desenvolvimento dos algoritmos de qualificação de dados, proposta principal desta dissertação.

O **Capítulo 8** apresenta de forma iconográfica alguns perfis de comportamento das grandezas avaliadas pelo sistema de qualificação a fim de permitir sua comparação com aqueles verificados no Capítulo 6, de forma a validar a metodologia proposta nesta dissertação.

As conclusões finais desta dissertação, decorrentes das avaliações realizadas, assim como os possíveis desdobramentos relativos às propostas de continuidade, são apresentadas no **Capítulo 9**.

Ao final do texto é apresentada a **Bibliografia**.

2 **STH – Sistema de Telemetria e Monitoramento Hidrometeorológico**

2.1 INTRODUÇÃO

O sistema de telemetria da Cemig existe desde 1998 e foi criado para atender a uma demanda interna, reforçada pela Resolução ANEEL 396/98. Desde então, tanto a infra-estrutura quanto o aplicativo vêm sofrendo manutenções evolutivas, buscando as adequações tecnológicas e o desempenho requeridos pela sua inerente responsabilidade e pelos seus usuários.

2.2 HISTÓRICO

Conforme citado anteriormente no Capítulo 1, o STH, sigla original do sistema, era baseado numa arquitetura cliente/servidor, possuía 84 unidades terminais remotas, 10 concentradoras de comunicação de dados e 50 usuários. A sua utilização pelos técnicos dos centros de operação era restrita a visualização de determinadas medições, conforme demanda dos próprios operadores.

De 2004 em diante, a infra-estrutura vem sendo ampliada e hoje conta com 99 remotas e 15 concentradoras de comunicação de dados. Além disso, o aplicativo foi reestruturado e migrado para o ambiente WEB, quando, então, seu nome original ganhou três letras e foi rebatizado STHWeb. Foram agregadas novas funcionalidades e o número de usuários passou de 500, o que deu ao sistema uma abrangência corporativa. Com as ampliações, o sistema passou a ter a capacidade de monitorar todos os reservatórios da Cemig que estivessem equipados com sub-sistemas de medição e transmissão de dados, integrando diversas áreas da empresa através de uma base única de dados.

A seguir serão descritas cada uma das principais funcionalidades desenvolvidas para o sistema.

2.3 INTERFACE COM O USUÁRIO

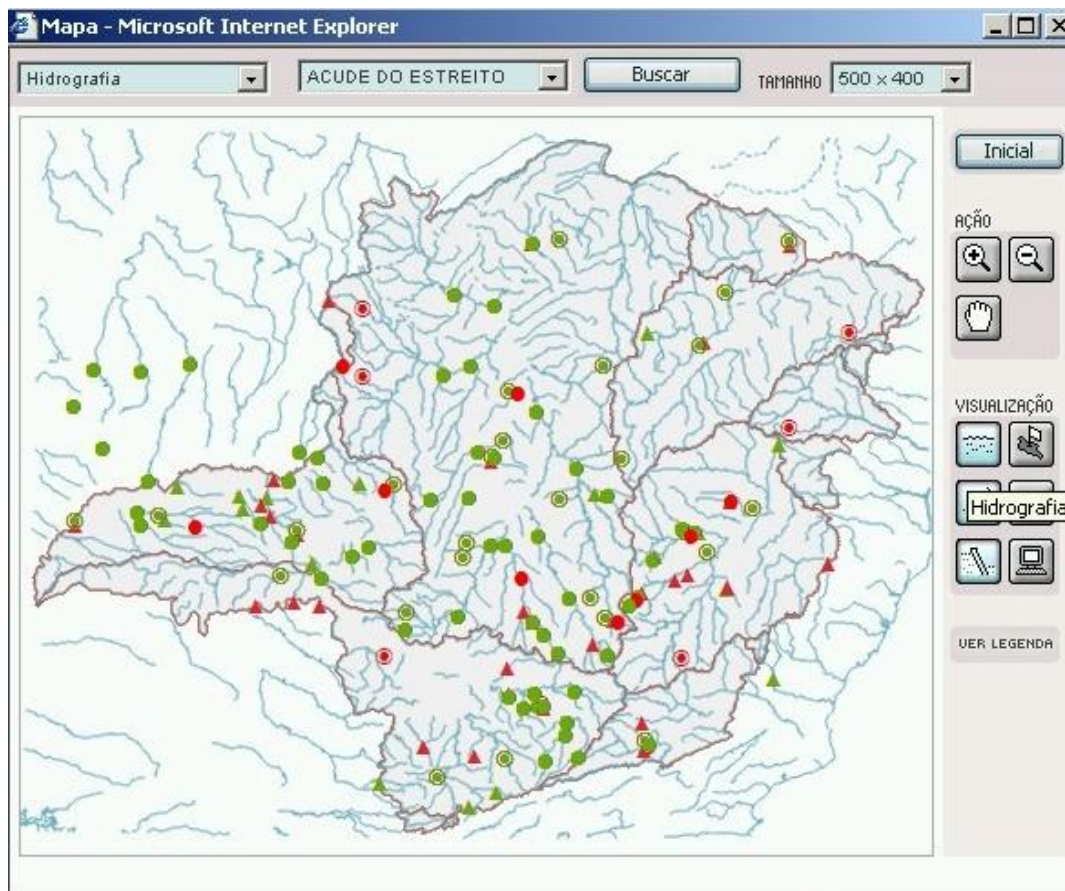
O sistema atual foi resultado de uma série de reuniões entre a equipe de desenvolvimento e os usuários em potencial, onde se buscou a integração dos interesses das principais áreas envolvidas: planejamento energético, pré-operação e procedimentos operativos, equipe de tempo real e equipe de pós-operação.

A navegação pelas páginas do sistema foi concebida para permitir o agrupamento de todas as funcionalidades em uma barra de menu com um *help on-line* associado e em um mapa de referência conforme ilustram as Figuras 2.1 e 2.2 a seguir.

Figura 2.1 – Tela Inicial do Sistema de Telemetria e Monitoramento Hidrometeorológico



Figura 2.2 – Mapa georeferenciado do Sistema de Telemetria e Monitoramento Hidrometeorológico



2.4 BASE DE DADOS

O gerenciador de banco de dados do STH é o Microsoft SQL Server. O aporte de dados ao sistema é feito por meio de diversas formas, a saber:

- Unidades terminais remotas hidrológicas (coletores de dados de chuva acumulada e nível d'água de rios e reservatórios);
- Unidades terminais remotas climatológicas (coletores de dados de chuva acumulada, temperatura e umidade relativa do ar, direção e velocidade média do vento, radiação solar acumulada e pressão atmosférica);
- Unidades terminais remotas mistas ou hidroclimatológicas (coletores de dados hidrológicos e climatológicos);
- Conexão com o Sistema de Supervisão e Controle Distribuído – SSCD do COS da Cemig, para intercâmbio de medições;
- Digitação dos dados na própria interface com o usuário do STH;

- Sistemas associados ao STH, como por exemplo, o Sistema de Gerenciamento da Operação – SIGO, instalado nas usinas.

Os dados são recebidos em tempo real das UTR e são utilizados e trabalhados pelos diversos módulos que compõem o sistema. São então exibidos na interface gráfica ficando disponíveis ao usuário.

Todo valor de uma determinada medida armazenada na base de dados permanente está sempre vinculado a um elemento hidrológico, a saber: Bacia Hidrográfica, Rio, Reservatório, Posto ou Usina. Por sua vez, cada elemento hidrológico está vinculado a um número restrito de tipos de informação. Uma associação válida entre um elemento hidrológico e um tipo de informação caracteriza uma grandeza. O STH possui 525 elementos hidrológicos, que representam, de maneira geral, as localizações nas quais há alguma medição sendo realizada; 76 tipos de informação, que representam cada uma das unidades de medida supervisionadas e calculadas, além de 3993 grandezas (vide Tabela 2.1 no final do capítulo).

Para a manutenção do banco de dados e administração de usuários, foi construído um aplicativo chamado *STHWebAdmin* visando facilitar as tarefas de cadastro e manutenção do sistema pelo administrador (*back office*), não exigindo conhecimentos específicos de gerenciamento e manipulação de banco de dados.

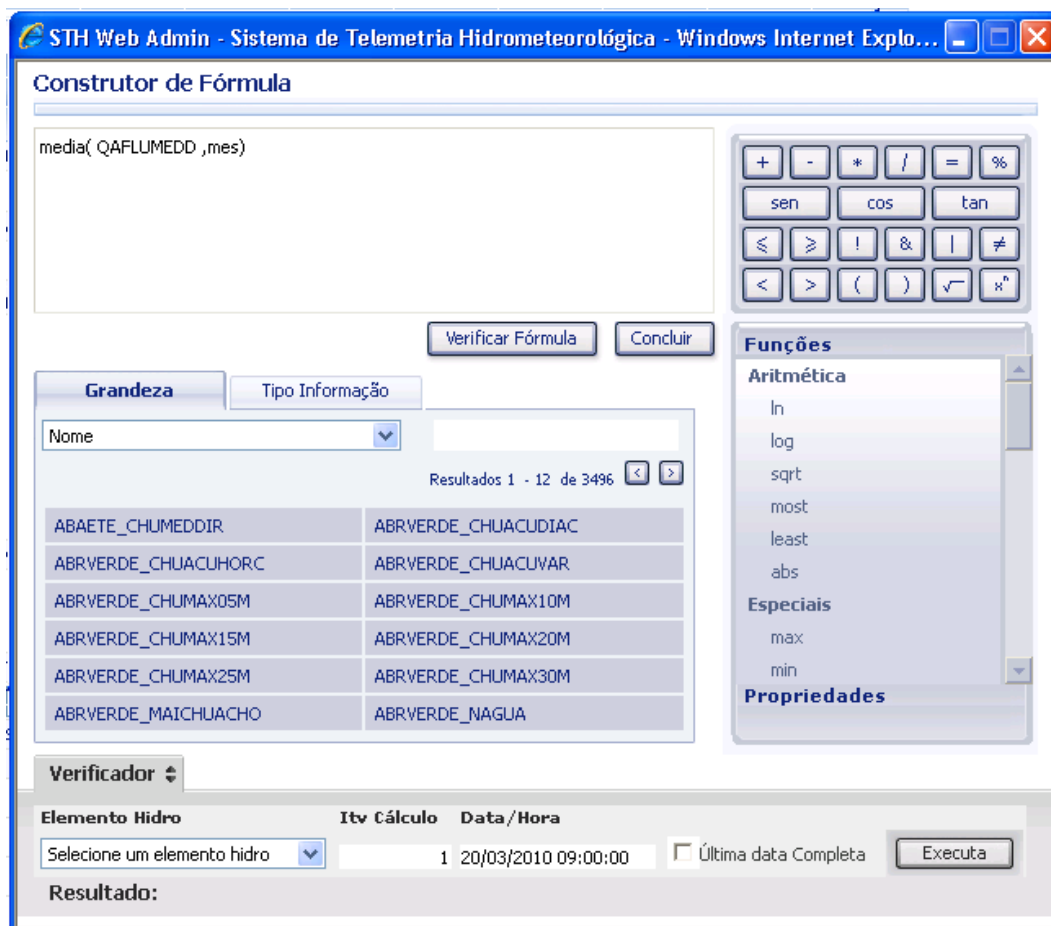
2.5 CÁLCULOS PERIÓDICOS

O sistema possui um módulo específico para a realização de cálculos que são totalmente parametrizáveis. O administrador pode interferir manualmente na frequência de sua realização, na sua fórmula, nas fontes de medida dos dados que correspondem às parcelas integrantes do cálculo, na especificidade de execução para um ou vários elementos hidrológicos ao mesmo tempo. Além disso, o sistema pode interferir de forma inteligente e automática na frequência de realização de um cálculo relativo a um reservatório em função de seu estado operativo, ou seja, se for detectado que o reservatório está em um estado que mereça uma maior atenção por parte dos operadores e usuários do sistema, a frequência de cálculo será alterada automaticamente, a fim de refletir essa necessidade.

Outro grande avanço na nova versão do sistema foi o retrocálculo, que permite que um cálculo que já tenha sido efetuado possa ser refeito caso alguma parcela seja considerada inconsistente e o usuário decida informar um novo valor. O retrocálculo corrigirá não somente o cálculo em questão, mas também todos os demais vinculados a ele, formando um processamento em cascata de todos os cálculos interdependentes.

A inserção de uma nova fórmula ou a verificação daquelas existentes é feita através de um construtor de equações que permite ao administrador cadastrar e testar as mais diversas possibilidades. A tela que ilustra esse construtor de equações é mostrado na Figura 2.3 a seguir.

Figura 2.3 – Construtor de fórmulas do STH



2.6 MONITORIA DE ELEMENTOS HIDROLÓGICOS

A monitoria de elementos hidrológicos, ou postos de medição, é dividida de acordo com perfis sendo que cada operador visualizará apenas os elementos hidrológicos (reservatórios, rios, postos de medição) sob sua responsabilidade, podendo os mesmos ser agrupados por Bacia Hidrográfica ou exibidos de forma específica.

A monitoria é baseada nas instruções de operação dos reservatórios, as quais estabelecem, dentre inúmeras outras diretrizes, a periodicidade de coleta de dados para verificação das seguintes grandezas hidrológicas: níveis d'água a montante e jusante do reservatório, suas vazões afluente, vertida, turbinada e defluente além do seu volume de espera.

Chegado o momento da monitoria de um reservatório, o sistema emitirá um alerta, através do mapa georeferenciado (ver Figura 2.2) ou da própria interface, por meio de sinalizadores de alarme de cores diferenciadas, conforme ilustra a Figura 2.4 a seguir.









Figura 2.4 – Painel de sinalização de alarmes de monitoria

The screenshot displays the STHWEB interface for monitoring hydrological elements. The main header reads "SISTEMA DE TELEMETRIA E MONITORAMENTO HIDROMETEOROLÓGICO" with the STHWEB logo. A navigation menu includes "Monitoria", "Supervisão", "Medidas", "Ferramentas", "Mensagens", and "Relatórios". The left sidebar shows "Monitoria BACIAS HIDROGRÁFICAS" with a dropdown menu set to "Paranaíba" and buttons for "Filtrar" and "Exibir Mapa".

The main content area is titled "Paranaíba" and contains a table of reservoir data. The table has columns for "Reservatório", "Data - Hora", "NA(m)", "VU(%)", "Qa(m³/s)", "Qt(m³/s)", "Qv(m³/s)", and a set of four colored status icons (red, green, blue, yellow). The data rows are as follows:

Reservatório	Data - Hora	NA(m)	VU(%)	Qa(m³/s)	Qt(m³/s)	Qv(m³/s)	Status
Capim Branco 1	27/11/2006 18:05	623,48	21,48	297,97	223,46	0,00	[Red][Green][Blue][Yellow]
Emborcação	27/11/2006 18:05	643,14	48,19	458,55	854,48	0,00	[Red][Green][Blue][Yellow]
Miranda	27/11/2006 18:05	695,59	85,52	370,86	320,74	0,00	[Red][Green][Blue][Yellow]
Nova Ponte	27/11/2006 18:05	810,60	82,53	267,79	289,69	0,00	[Red][Green][Blue][Yellow]
São Simão	27/11/2006 18:05	392,98	20,73	2.828,23	2.156,81	1.304,82	[Red][Green][Blue][Yellow]

At the bottom left, there is a yellow "OBSERVAÇÕES" box with the text: "Anil. O STH ainda não carregado com as tabelas de descarga dos órg...". Below this is a pagination indicator: "Página 1 de 3 << >>".

Os botões sinalizadores     mostrados na Figura 2.4 indicam respectivamente:  horário de coleta e verificação obrigatória das grandezas hidrológicas;  existência de observação temporária para o reservatório;  o reservatório encontra-se fora da região normal de operação;  existência de alarme gerado pela violação de alguma inequação do reservatório.

As cores dos botões indicam: **Vermelho** – evento gerado pelo sistema e ainda não reconhecido pelo operador; **Amarelo** – evento gerado pelo sistema, reconhecido pelo operador, mas com a permanência da situação que o causou; **Verde** – situação normal.

Foi solicitado pelos operadores, também conhecidos como despachantes do COS da Cemig, a integração do STHWeb com o SSCD (Sistema de Supervisão e Controle Distribuído), para permitir a visualização de informações e alarmes dos dois sistemas em uma única tela. Assim, o mapa de bacias hidrográficas disponível no STH (ver Figura 2.5) foi inserido e fica ativo dentro da interface gráfica do SSCD e, no caso da ocorrência de algum dos eventos citados acima, a bacia hidrográfica em questão muda de cor, alertando ao despachante sobre a necessidade de uma ação de sua parte, referente à monitoria da bacia hidrográfica em questão no STH.

Figura 2.5 – Mapa informativo da situação do monitoramento das bacias hidrográficas



Para subsidiar o controle e as decisões operativas o usuário do sistema conta ainda com o auxílio de diversas telas que apresentam as informações recebidas em tempo real vindas dos postos de coleta de dados hidroclimatológicos, conforme ilustram as Figuras 2.6 e 2.7 a seguir.

Figura 2.6 – Exemplo de tela de monitoria de posto de medição

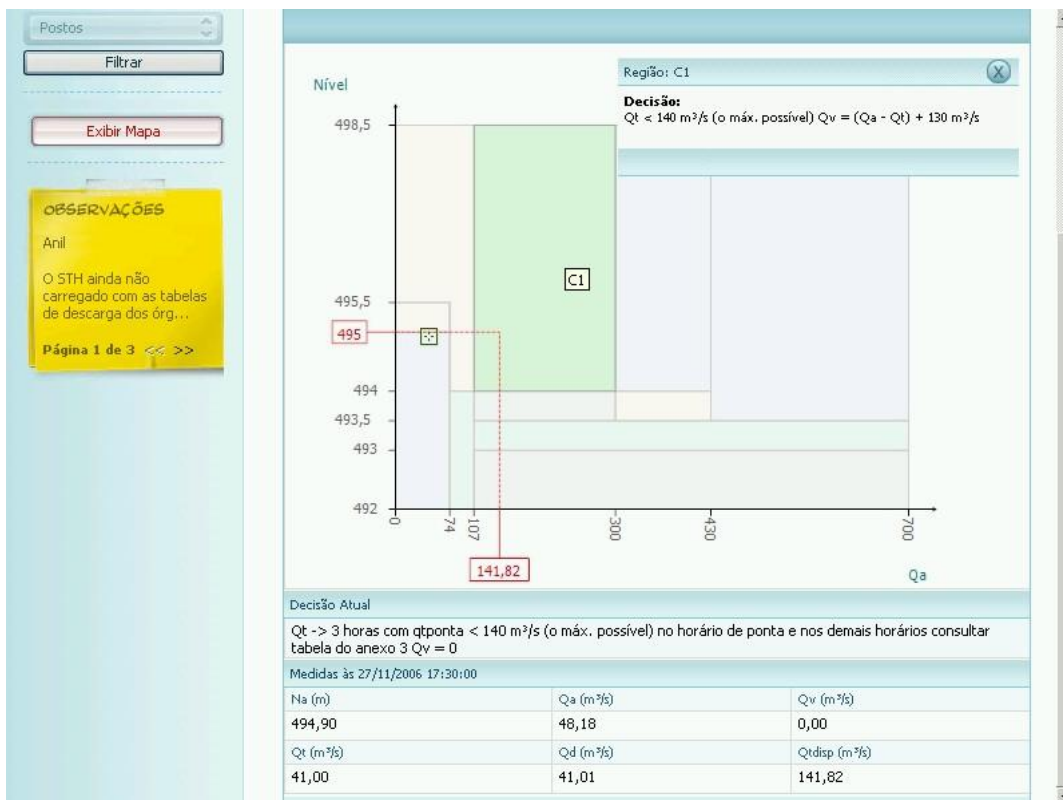


Figura 2.7 – Exemplo de tela de monitoria de posto de medição



As grandezas hidrológicas também podem ser visualizadas de forma gráfica mais elaborada, facilitando a sua compreensão e a interpretação do estado operativo em que se encontram cada um dos reservatórios. O estado operativo atual do reservatório bem como os demais estados possíveis, além das suas respectivas decisões operativas constantes das IO (Instruções de Operação) são “traduzidos” e então, apresentados em tempo real através do diagrama de operação do reservatório que também pode ser exibido na interface com o usuário. Esta funcionalidade é uma excelente ferramenta de suporte à decisão conforme ilustra a Figura 2.8 a seguir.

Figura 2.8 – Diagrama de Operação de um reservatório identificando suas regiões de operação e decisões operativas



2.7 SIMULAÇÃO

O sistema dispõe também de ferramentas de simulação, onde é possível realizar simulações diárias e horárias a fim de se verificar, de forma virtual, o comportamento de um determinado reservatório, ou de uma cascata deles, para determinadas circunstâncias hidrológicas.

Além disto, partindo-se de um estado de vazão afluente conhecido, pode-se aplicar um coeficiente de recessão à mesma e simular o comportamento do reservatório ou então utilizar os valores previstos pela área de hidrologia ou a MLT (Média de Longo Termo) e seus percentuais ou ainda a digitação de valores de vazão afluente, conforme mostra a Figura 2.9.

Figura 2.9 – Tela de simulação do comportamento de reservatórios

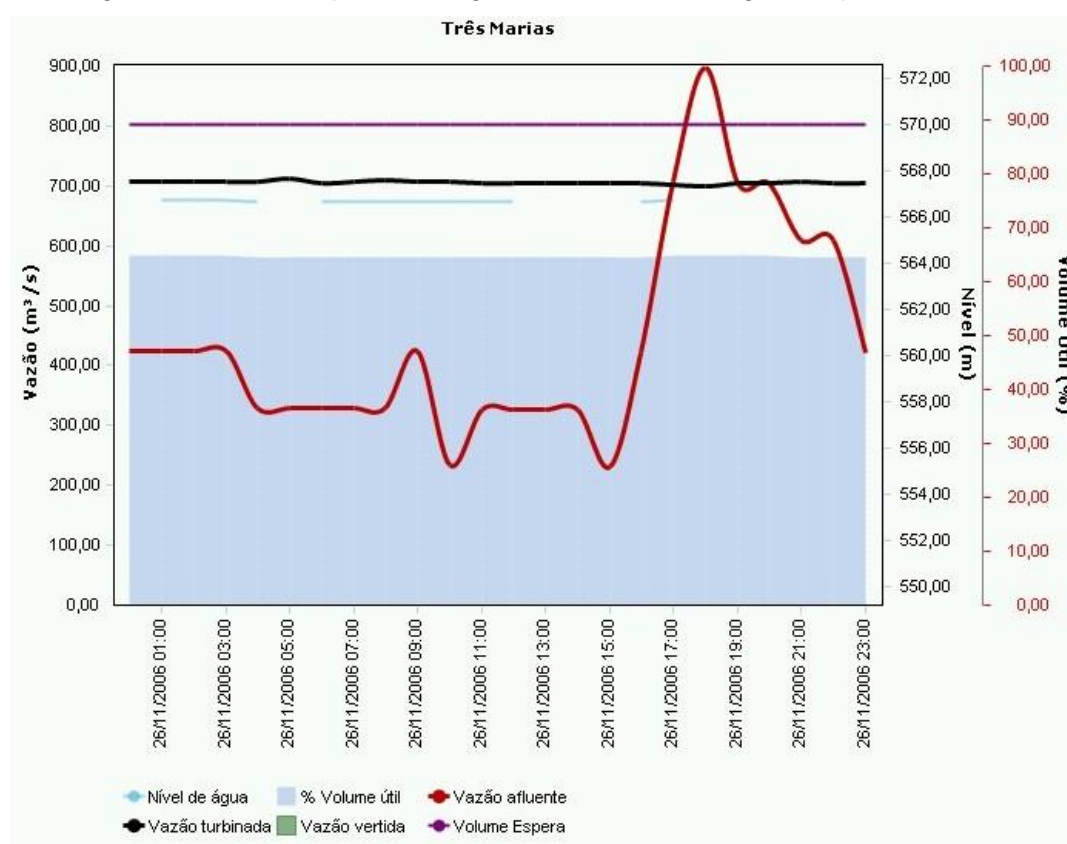
The screenshot displays the STHWEB interface for simulating the behavior of the Jaguará reservoir. The main navigation bar includes 'Monitoria', 'Supervisão', 'Medidas', 'Ferramentas', 'Mensagens', and 'Relatórios'. The 'Ferramentas' section is expanded to show 'SIMULAÇÃO'. On the left, there are dropdown menus for 'Grande', 'Jaguara', and 'Postos', along with a 'Filtrar' button. Below these is a 'Lista de Reservatórios' with entries for Camargos, Itutinga, Funil, Jaguará, Igarapava, and Volta Grande. The Jaguará entry is selected, showing its maximum and minimum values. The simulation parameters are set to 'Hórraria' (Hourly) with a 'Fonte de dados' of 'Prev. de Afluência', an 'Intervalo' of '12 horas', and a 'Data início' of '27/11/2006 16:00'. A 'Simular' button is present. The resulting data is shown in a table with the following columns: Hora, Na inicial (m), Qa(m³/s), Geração (MW), Qt(m³/s), Qv(m³/s), Qd(m³/s), Na final (m), Vol Esp (m), and VU final (%).

Hora	Na inicial (m)	Qa(m³/s)	Geração (MW)	Qt(m³/s)	Qv(m³/s)	Qd(m³/s)	Na final (m)	Vol Esp (m)	VU final (%)
16:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
17:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
18:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
19:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
20:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
21:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
22:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
23:00 27/11/2006	558	1000	300	745,9	250	995,9	558		83,33
00:00 28/11/2006	558								
01:00									

Uma nova funcionalidade foi desenvolvida e incorporada na última versão do sistema, que é a possibilidade de realizar as simulações em cascata, ou seja, a simulação do comportamento de um reservatório servirá como entrada para a simulação do comportamento do reservatório à jusante e assim por diante. Além

disto, várias simulações podem ser efetuadas sem que se percam as simulações anteriores e é possível visualizar as mesmas através de gráficos (ver Figura 2.10), ou exportá-las para planilha do Microsoft® Excel®. Nas simulações também se pode optar pela verificação do resultado do monitoramento das inequações, visando informar aos usuários e operadores sobre os efeitos e as conseqüências da simulação que está sendo efetuada, comparando os resultados com os limites operativos do reservatório disponíveis no cadastro do sistema.

Figura 2.10 – Exemplo de um gráfico de tendência gerado pelo sistema



2.8 ALARMES

O sistema classifica os alarmes emitidos pelo sistema, de acordo com a sua origem:

Supervisão da infraestrutura

- as UTR (Unidades Terminais Remotas) coletoras de dados, instaladas em campo, são capazes de enviar em tempo real qualquer ocorrência

que mereça a intervenção da equipe de manutenção. Exemplos dessas ocorrências são: porta do abrigo aberta; falha de corrente contínua; falha de corrente alternada e falha de comunicação. Estes alarmes são exibidos na interface gráfica com o operador, nas telas do subsistema de mensagens do sistema ou podem ser verificados através do mapa georeferenciado, na camada de “postos telemididos”. As cores que identificam as UTR no mapa são dinâmicas, indicando o tipo e a gravidade da ocorrência verificada através da alteração da cor do símbolo que a identifica.

Violação de inequações

- o STH possui um subsistema de inequações, onde os usuários e operadores podem construir as mais diversas combinações entre as variadas grandezas do sistema e colocar estas inequações e fórmulas sob monitoramento. Caso a comparação dos valores dessas grandezas, recebidos em tempo real, viole o determinado pela inequação cadastrada, o subsistema de inequações emitirá alarmes que serão exibidos na interface de monitoria de reservatórios (botão A da Figura 2.4 mostrada anteriormente) e nas camadas de bacias ou reservatórios existentes no mapa georeferenciado. Estes alarmes também podem, se configurados para tal, alertar o usuário através do envio de uma mensagem SMS, via telefone celular ou através do envio de uma mensagem de e-mail.

Todos os alarmes emitidos pelo sistema buscam garantir a integridade e a segurança na operação dos reservatórios supervisionados sob todos os aspectos, seja no acompanhamento das variáveis dos reservatórios, seja na confiabilidade da fonte geradora de tais dados, ou seja, as unidades terminais remotas (UTR).

2.9 CONSULTAS

O volume e a variedade de informações geradas pelo STH são imensos, e atendem não só às equipes do COS e dos centros de operação regionais, mas também às áreas de hidrologia, meteorologia, hidrometria e as mais diversas equipes de engenharia da Cemig. O sistema também possui vários clientes

externos, como o ONS, ANEEL, ANA, CPTEC, órgãos de defesa civil, prefeituras, universidades e outras instituições.

As consultas a esta enorme base de dados podem ser construídas pelos próprios usuários de forma interativa e dinâmica e, além disso, também podem ser armazenadas para execução posterior. O usuário pode definir as grandezas, a frequência de coleta dos dados, o período de abrangência da pesquisa e as fontes das medidas, dentre outros parâmetros configuráveis.

O sistema disponibiliza os resultados das consultas através de gráficos e planilhas, sendo que estas são exportáveis para uso em outros aplicativos. A Figura 2.11 ilustra uma destas possibilidades.

Para os usuários externos à Cemig, o sistema é parametrizado para gerar consultas na frequência e formato desejados e disponibilizar os arquivos em locais pré-definidos, como por exemplo, servidores FTP (*File Transfer Protocol*), tudo isto de forma automática.

Figura 2.11 – Relatório baseado em consulta realizada na base de dados

SISTEMA DE TELEMETRIA E MONITORAMENTO HIDROMETEOROLÓGICO

Monitoria Supervisão Medidas Ferramentas Mensagens Relatórios Help About

Relatórios CONSULTA

Dados para FEAM - Baía SFrancisco

20/11/2006 00:00 26/11/2006 00:00 Filtrar Exportar Voltar

Data/Hora	NAGUA	PIRAPORAPTE		
		QAFLUPTCT	CHUACUHORC	
25/11/2006 23:00:00	2,79	888,50	0,00	1,87 ▲
25/11/2006 22:00:00	2,80	900,00	0,00	1,87
25/11/2006 21:00:00	2,80	900,00	0,00	1,87
25/11/2006 20:00:00	2,80	900,00	0,00	1,87
25/11/2006 19:00:00	2,80	900,00	0,00	1,87
25/11/2006 18:00:00	2,80	900,00	0,00	
25/11/2006 17:00:00	2,80	900,00	0,00	1,87
25/11/2006 16:00:00	2,80	900,00	0,00	1,87
25/11/2006 15:00:00	2,80	900,00	0,00	1,87
25/11/2006 14:00:00	2,80	900,00	0,00	1,87
25/11/2006 13:00:00	2,80	900,00	0,00	1,87
25/11/2006 12:00:00	2,80	900,00	0,00	1,87
25/11/2006 11:00:00	2,80	900,00	0,00	1,87 ▼

2.10 CONCLUSÃO DO CAPÍTULO

Desde a sua inauguração em 1998, o STH apresentou-se como um sistema pioneiro no setor elétrico brasileiro, na área de monitoramento em tempo real de grandezas hidrológicas e meteorológicas. Atualmente, com o desenvolvimento e implantação da versão STHWeb, pode-se afirmar que este sistema de supervisão representa um esforço pioneiro da Cemig em tornar as decisões operativas necessárias à gestão de seus reservatórios, mais dinâmicas e fundamentadas em uma base de dados única e segura além de ser um sistema ágil e amigável.

Todo o esforço é feito para tornar o STH uma ferramenta altamente customizada que atenda à diversidade de demandas dos seus usuários.

A seguir, a Tabela 2.1 enumera os principais tipos de informação disponíveis na base de dados do STH. Vale ressaltar que a lista de tipos de informação é uma lista dinâmica, uma vez que novas informações podem ser incorporadas, sejam elas resultados de um cálculo periódico criado ou de uma nova telemedição implantada.

Tabela 2.1 - TIPOS DE INFORMAÇÃO

CHUVA ACUMULADA DIÁRIA CALCULADA	ENERGIA MÉDIA DIÁRIA
CHUVA ACUMULADA HORÁRIA CALCULADA	ENERGIA MÉDIA HORÁRIA
CHUVA ACUMULADA MENSAL CALCULADA	ENERGIA MÉDIA MENSAL
CHUVA ACUMULADA POR TEMPO VARIÁVEL	ENERGIA VERTIDA TURBINÁVEL
CHUVA MÁXIMA DE 10 MINUTOS	EVAPOTRANSPIRAÇÃO
CHUVA MÁXIMA DE 15 MINUTOS	MAIOR CHUVA OCORRIDA NO DIA
CHUVA MÁXIMA DE 20 MINUTOS	MEGAWATT DISPONIVEL
CHUVA MÁXIMA DE 25 MINUTOS	NÍVEL DE ÁGUA
CHUVA MÁXIMA DE 30 MINUTOS	NÍVEL DO CANAL DE FUGA
CHUVA MÁXIMA DE 5 MINUTOS	PERCENTUAL DE VOLUME ÚTIL
CHUVA MÉDIA DIÁRIA	POTÊNCIA GERADA
CHUVA MÉDIA HORÁRIA	POTÊNCIA GERADA DIÁRIA
CHUVA MÉDIA MENSAL	PRESSÃO ATMOSFÉRICA INSTANTÂNEA
DIREÇÃO INSTANTÂNEA DO VENTO	PRESSÃO ATMOSFÉRICA MÁXIMA DIÁRIA
DIREÇÃO MÉDIA DO VENTO	PRESSÃO ATMOSFÉRICA MÉDIA

PRESSÃO ATMOSFÉRICA MÍNIMA DIÁRIA	VAZÃO DEFLUENTE MÉDIA DIÁRIA
RADIAÇÃO SOLAR GLOBAL	VAZÃO DEFLUENTE MÉDIA HORÁRIA
TEMPERATURA INSTANTÂNEA	VAZÃO EM PONTOS DE CONTROLE
TEMPERATURA MÁXIMA DIÁRIA	VAZÃO MÉDIA DIÁRIA EM PONTOS DE CONTROLE
TEMPERATURA MÉDIA	VAZÃO MÉDIA MENSAL EM PONTOS DE CONTROLE
TEMPERATURA MÍNIMA DIÁRIA	VAZÃO NO TUNEL
UMIDADE RELATIVA DO AR INSTANTÂNEA	VAZÃO PELO DISPOSITIVO AUXILIAR
UMIDADE RELATIVA DO AR MÉDIA	VAZÃO PELO DISPOSITIVO AUXILIAR
VAZÃO AFLUENTE	VAZÃO PELO TÚNEL MÉDIA DIÁRIA
VAZÃO AFLUENTE HORÁRIA	VAZÃO TURBINADA
VAZÃO AFLUENTE MÉDIA DIÁRIA	VAZÃO TURBINADA DISPONÍVEL
VAZÃO AFLUENTE MÉDIA MENSAL BH	VAZÃO TURBINADA MÉDIA DIÁRIA
VAZÃO AFLUENTE NATURAL MÉDIA DIÁRIA	VAZÃO TURBINADA MÉDIA MENSAL
VAZÃO AFLUENTE NATURAL MÉDIA DIÁRIA PRELIMINAR	VAZÃO VERTIDA
VAZÃO AFLUENTE NATURAL MÉDIA MENSAL	VAZÃO VERTIDA MÉDIA DIÁRIA
VAZÃO DE EMERGÊNCIA	VAZÃO VERTIDA MÉDIA MENSAL
VAZÃO DE RESTRIÇÃO	VELOCIDADE INSTANTÂNEA DO VENTO
VAZÃO DE RESTRIÇÃO A JUSANTE	VELOCIDADE MÉDIA DO VENTO
VAZÃO DEFLUENTE	

3 PREPARANDO OS DADOS

3.1 INTRODUÇÃO

Considerando-se as grandezas para as quais será desenvolvida a ferramenta de qualificação proposta nesse trabalho, que são: precipitação pluviométrica (chuva), nível d'água de rios e de reservatórios de usinas, temperatura e umidade relativa do ar, velocidade média e direção do vento, radiação solar e pressão atmosférica, temos que até 01/01/2010 foram recebidas e armazenadas nas bases de dados do STH 67.607.330 (sessenta e sete milhões seiscientos e sete mil trezentos e trinta) medidas. Destas medidas, derivaram uma infinidade de outras, como resultado de cálculos realizados pelos diversos módulos que compõem o sistema. Sendo assim, até aquela data o sistema possuía em sua base 126.532.505 (cento e vinte e seis milhões quinhentos e trinta e dois mil quinhentos e cinco) medidas armazenadas sem terem recebido nenhum tipo de tratamento que avaliasse a qualidade das mesmas.

Como mostra o parágrafo anterior, o número de medidas a serem tratadas é extremamente elevado e, sendo assim, um tratamento manual das grandezas se tornou impossível, sendo necessário construir pequenos aplicativos com os mais diversos fins onde um dos principais é aquele responsável pela filtragem inicial dos dados. O principal motivo para a necessidade destes aplicativos auxiliares no tratamento inicial dos dados é o grande número de medidas inconsistentes recebidas nas fases iniciais de implantação do sistema, quando ocorrem as instalações em campo das unidades terminais remotas e seus diversos sensores, seguindo-se com as calibrações de todos os equipamentos e testes de aceitação.

Para garantir um tratamento estatístico adequado às medidas, estas inconsistências verificadas foram filtradas, o que é conhecido como “análise de *outliers*”. Os motivos principais das inconsistências e os mecanismos de filtragem serão detalhados no decorrer deste capítulo.

3.2 HISTÓRICO

Toda a plataforma de aquisição de dados, no que diz respeito aos seus aspectos, tais como sistemas de alimentação de energia, seja por painel solar ou pela rede comercial de distribuição; unidades coletoras de dados e suas capacidades de processamento de dados; sistemas de transmissão de dados, sejam por *links* de rádio VHF/UHF (*Very High Frequency/Ultra High Frequency*), microondas ou satélite de comunicação, foram especificados pelos especialistas da Cemig.

Foi feita então uma licitação internacional para definição da empresa que desenvolveria os equipamentos, conforme as especificações técnicas, e integraria todos os elementos da plataforma de *hardware* e *software* de dispositivos (*firmware*), formando um completo sistema de aquisição de dados hidrológicos e meteorológicos. A empresa vencedora da licitação internacional foi a JRC – *Japan Radio Company* que tem larga experiência na implantação de sistemas com estas características seja no Japão ou ao redor do mundo.

Toda a infra-estrutura de *software* do aplicativo que compunha o STH na sua primeira versão foi especificada e desenvolvida internamente por especialistas da Cemig. O sistema inicial era baseado na arquitetura cliente/servidor, suportada por um gerenciador de base de dados relacional e com a interface gráfica com o usuário baseada em telas no padrão MS Windows®. Por conta desta decisão, toda a tecnologia e o conhecimento adquirido com o desenvolvimento e a evolução do sistema ficaram preservados dentro da empresa, o que garante uma continuidade e manutenção do sistema ao longo dos anos e independência com relação a desenvolvimentos externos. Foi isso que permitiu a migração da arquitetura inicial, para a arquitetura atual, baseada em interface WEB.

O início da implantação da infra-estrutura de aquisição de dados ocorreu em 1998. Este período inicial já contava com um protótipo do SCADA – *Supervisory Control and Data Acquisition*, o que possibilitou desde as primeiras instalações em campo, popular a base de dados do sistema com os dados telemedidos pelas unidades terminais remotas. Esta precocidade na recepção de dados foi muito útil durante a fase inicial de implantação e, entretanto, trouxe com ela alguns problemas que dizem respeito à qualidade das medições recebidas. Estes problemas serão comentados nos tópicos que se seguem.

3.3 PRIMEIRAS TELEMEDIÇÕES

As primeiras medidas recebidas pelo sistema não podiam ser consideradas como confiáveis por diversas razões. Dentre elas destacam-se: O sistema estava em sua fase inicial de implantação, o que envolvia obras civis, montagens e desmontagens de equipamentos de sensoriamento, de comunicação de dados e de alimentação elétrica; os sensores estavam com calibração “de fábrica” e suas medições não haviam sido confrontadas com aquelas de medidores padrão; os protocolos de comunicação entre as unidades coletoras de dados e a central estavam em fase de validação. Além disso, os *links* de dados estavam na fase inicial de alinhamento, calibração e adequações de potências de transmissão dos rádios e equipamentos afins.

Sendo assim, um grande número de telemidas inconsistentes foi recebido pelo sistema preliminar de aquisição de dados e estas telemidas, mesmo estando contaminadas pelos mais diversos ruídos, foram armazenadas na base de dados relacional.

Essa situação perdurou por alguns meses, visto que quando se terminava o “comissionamento” e implantação de um posto de medição, iniciava-se a implantação de outro. Vale ressaltar que a implantação de postos de medição envolveu um número elevado de viagens por todo o estado de Minas Gerais e em parte do estado de Goiás, como se pode verificar no mapa apresentado no Capítulo 2 que mostra a distribuição geográfica dos postos de medição.

3.4 FALTOU SOL

Desde a implantação dos primeiros postos, percebeu-se que era comum a ausência de recepção de dados durante grande parte da noite e até o nascer do sol. Rapidamente foi constatado que o problema era causado por incapacidade de carga das baterias que alimentavam as estações coletoras de dados que não dispunham de fornecimento de energia convencional, através da rede de distribuição. Estas estações se localizam em lugares de difícil acesso e não dispõem de nenhuma infra-estrutura. Sendo assim, o posto ou estação de medição tem como fonte de energia painéis solares acoplados a bancos de baterias. Durante o dia os painéis solares fornecem alimentação aos

equipamentos de medição e simultaneamente carregam os bancos de baterias que, à noite, são os responsáveis pela alimentação da estação coletora de dados. Porém, em dias mais nublados, a carga não era suficiente para manter a operação durante todo o período noturno e o posto literalmente “era colocado para dormir”.

Este desligamento dos equipamentos de medição trazia consigo, além da ausência dos dados “noturnos”, falhas nos dados que eram acumulados durante o dia, a saber, os dados de chuva acumulada e a radiação solar global. Estas medidas eram “resetadas” para o zero, a cada inicialização da estação que acontecia com o carregamento das baterias ao nascer do sol.

3.5 ÁGUAS NIPÔNICAS VERSUS ÁGUAS BRASILEIRAS

Com a intenção de facilitar a implantação e manutenção futura do sistema de medição de nível d’água, o sensor de nível d’água foi especificado como sendo do “tipo pressão”. Sua instalação é muito mais simples do que o sensor de nível tradicionalmente usado, do tipo bóia e contrapeso.

O sensor de nível d’água do tipo “pressão” é baseado numa câmara de pressão conectada a um cabo de interligação composto por um duto de ar que tem a capacidade de medir a altura da coluna de água sobre o mesmo. Sua instalação é muito simples, uma vez que o mesmo é lançado nas águas do rio ou reservatório e é ancorado no leito dos mesmos, dispensando as obras civis demandadas pelo sensor do tipo bóia e contrapeso para instalação do “tubulão” e dos instrumentos de medição. Além disso, esperava-se que a sua necessidade de manutenção, ou em outras palavras, que seu MTBF (*Mean Time Between Failures*) fosse muito menor do que o apresentado pelo sensor de bóia e contrapeso que possui muitas peças móveis.

Entretanto, após alguns meses de operação de diversos desses sensores baseados na câmara de pressão, inúmeras medições de nível d’água começaram a apresentar erros grosseiros. Verificou-se que os sensores estavam sendo rapidamente atacados por um processo de corrosão que chegava a alcançar a câmara de pressão, alterando significativamente as medidas. Esta corrosão, entretanto, não ocorria de forma tão acentuada nas águas dos rios japoneses, o que levou a se acreditar que a composição das águas brasileiras,

mais especificamente das águas mineiras (talvez pela grande presença de minerais), estava “fora das especificações adequadas” para uso do sensor baseado em pressão.

Por conta deste acontecimento, todos os sensores de pressão acabaram sendo substituídos pelos “bons e velhos” sensores com bóia e contrapeso.

3.6 A REMOTA DO STH

O equipamento coletor de dados, mais conhecido nos sistemas de supervisão como UTR (Unidade Terminal Remota), ou simplesmente Remota, foi especificado pela Cemig e desenvolvido pela JRC no Japão. A especificação deste equipamento previu diversas capacidades de processamento a serem implementadas no *firmware* (micro código operacional do equipamento) da remota, dentre as quais, a capacidade de autosupervisão do estado dos sensores conectados. Este monitoramento do estado dos sensores inclui:

- Sensor defeituoso;
- Sensor desconectado;
- Sensor desligado;
- Medida inválida;
- Remota em manutenção;
- *After start* (inicialização).

Esta supervisão da situação dos sensores já fornece às telemedidas uma idéia inicial da qualidade da mesma. Estes qualificadores são enviados juntamente com o valor da medida, atrelados a ela. Nos quatro primeiros casos, a medida vem com o valor igual a zero, no quinto caso (em manutenção) é informado o valor de engenharia apropriado, mas o “*flag*” qualificador indica que é necessário levar em conta que a remota estava em manutenção, o que indica que a medição pode estar contaminada por algum procedimento de teste. No último caso, a indicação informa que a remota sofreu uma inicialização, ou “*reset*”, seja por falta de energia no posto de medição ou por solicitação de um operador do sistema que tenha os privilégios necessários para esta tarefa.

3.7 TRATANDO AS INCONSISTÊNCIAS

O número de medições inadequadas devido aos erros de medição e aos problemas citados durante a implantação do sistema mostrou-se muito elevado e

certamente essas inconsistências acabariam influenciando negativamente no tratamento estatístico dos dados, tais como médias e desvio padrão, fundamentalmente necessários à elaboração da ferramenta de qualificação proposta nesta dissertação. Sendo assim, uma filtragem foi feita de forma a criar uma base de dados inicial adequada para o início das pesquisas.

Esta filtragem inicial, baseada na identificação de “*outliers*” ou pontos fora da curva, se baseia no tratamento de limites físicos das grandezas e no tratamento de violações de taxas de variação da grandeza. Estes valores limites foram conseguidos junto aos hidrólogos e meteorologistas da Cemig. Os resultados iniciais encontrados bem como o processo de filtragem inicial das inconsistências verificadas nos dados serão apresentados no Capítulo 6 desta dissertação.

Para tratar toda a massa de dados disponível, filtrando as inconsistências primárias, foram desenvolvidos alguns aplicativos que ao final de sua execução possibilitaram levantar com um adequado nível de qualidade as médias, desvios padrão e taxas de variação consideradas normais para cada tipo de grandeza telemedida. Com esses dados disponíveis, criou-se no banco de dados relacional do sistema um modelo de dados composto por tabelas nas quais eram armazenados estes dados notáveis, que posteriormente são usados em processos de confrontação de valores e seus limites, numa primeira etapa de qualificação da telemedida recebida.

Os mecanismos subseqüentes de qualificação das medidas serão discutidos em mais detalhes no Capítulo 4.

4 METODOLOGIAS DE QUALIFICAÇÃO DOS DADOS

4.1 INTRODUÇÃO

Após as filtragens iniciais, para a eliminação de erros grosseiros como os descritos no Capítulo 3, devidos principalmente à fase de instalação dos equipamentos de coleta de dados, iniciaram-se então as análises dos dados disponíveis na base de dados históricos do sistema.

Baseando-se nestas análises, alguns mecanismos de qualificação das medições recebidas foram propostos, a fim de se desenvolver os algoritmos computacionais. Estas propostas de mecanismos para qualificação são discutidas ao longo deste capítulo.

4.2 REGRAS DO NEGÓCIO

Após algumas entrevistas com usuários do sistema, em sua maioria hidrólogos, percebeu-se a grande importância e a grande preocupação com a qualidade das séries históricas de dados hidrológicos. Algumas destas séries remontam ao ano de 1930 e são importantes até hoje, para a realização de previsões de afluência d'água aos reservatórios das usinas e índices de precipitação pluviométrica nas bacias hidrológicas do estado de Minas Gerais.

De forma geral, essas séries são de dados horários que posteriormente eram convertidos em diários e mensais através de cálculos simples. Como o STH é uma enorme fonte de dados hidrometeorológicos, se não a maior disponível atualmente na Cemig, estabeleceu-se como frequência mínima de recepção de dados telemídidos a frequência horária, a fim de manter a coerência com as séries históricas já existentes e de forma a dar consistência aos dados disponíveis para avaliação e tratamento numérico. Opcionalmente, ou

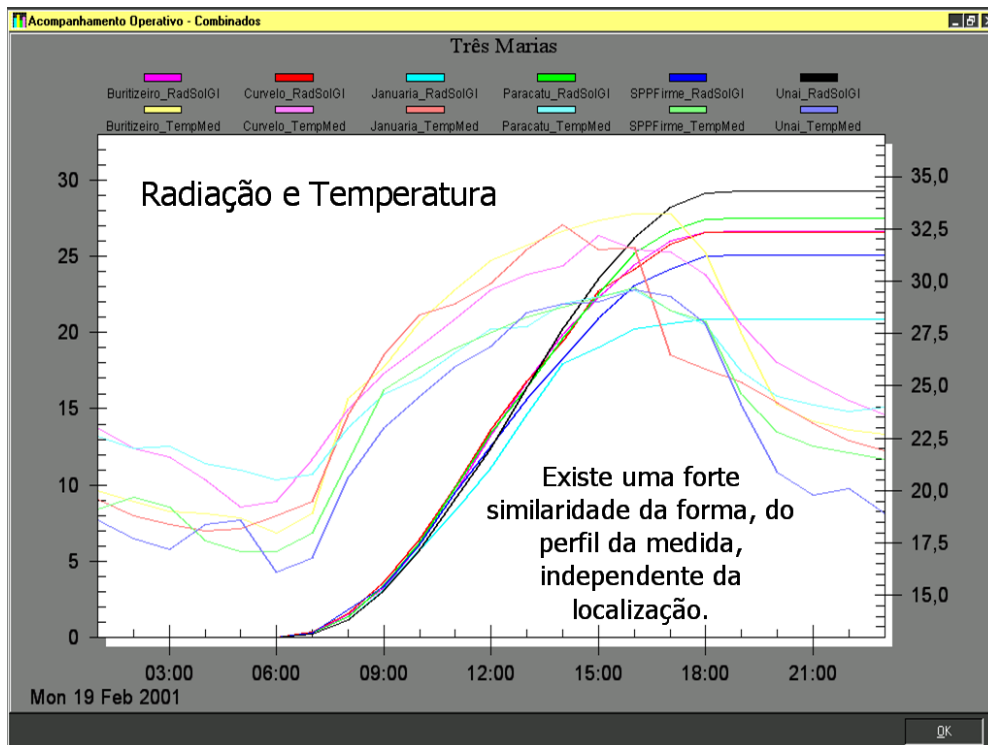
devido a alguma necessidade operativa, esta frequência pode ser aumentada, obtendo-se assim intervalos menores de recepção de dados.

4.3 ANÁLISE MACROSCÓPICA DOS DADOS

Tendo sido estabelecidas as regras para a avaliação das séries de dados históricos, iniciaram-se as primeiras análises dos dados disponíveis no sistema.

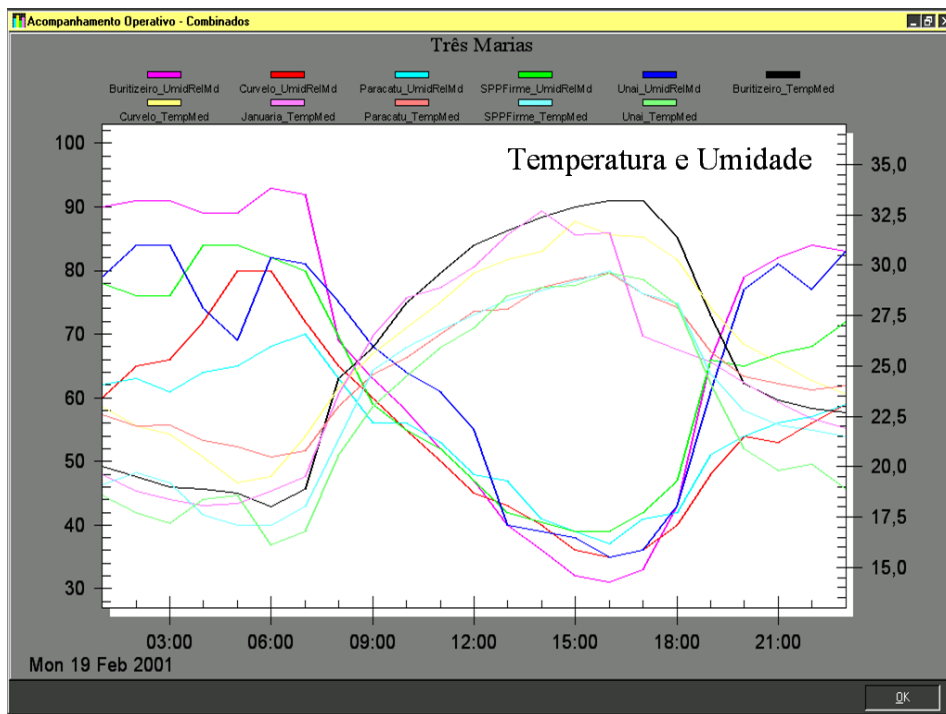
Com foco numa data específica, aleatoriamente escolhida, foi feita uma macro análise dos dados disponíveis para aquela data e verificou-se que existe para a maioria das grandezas uma grande similaridade da forma, do perfil, do comportamento da variação da grandeza ao longo do dia, independente da localização geográfica da estação de medição. Esta similaridade pode ser verificada na Figura 4.1, a seguir, onde são apresentados os dados de temperatura do ar e radiação solar acumulada, coletados pelo sistema no dia 19 de fevereiro de 2001, em diversas estações de medição que compõem o subsistema de supervisão da bacia do Rio São Francisco, onde se encontra o reservatório da usina de Três Marias.

Figura 4.1 – Similaridade de comportamento das grandezas



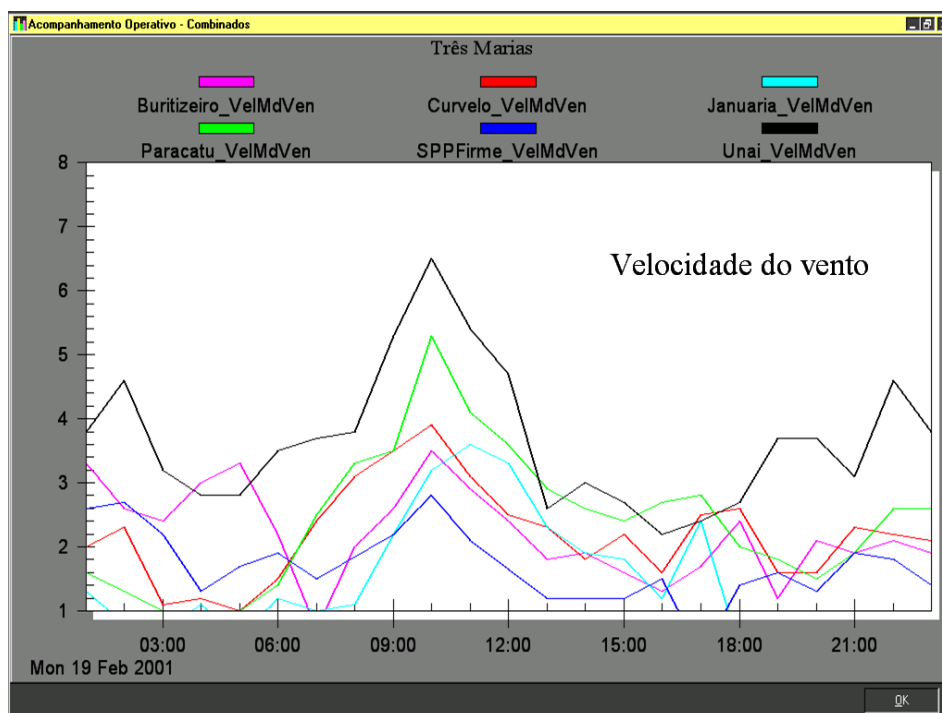
A similaridade de comportamento continua a ser verificada para outras grandezas representativas dos mesmos postos de coleta de dados da bacia do Rio São Francisco, conforme mostra a Figura 4.2, além de ficar também evidenciada certa relação de proporcionalidade entre algumas grandezas, caso claro de temperatura e umidade relativa do ar.

Figura 4.2 – Similaridade de comportamento das grandezas e indicadores de correlação entre comportamentos



Até mesmo a velocidade média do vento apresenta um comportamento cujo perfil diário, baseado em medições horárias, é consistente no que diz respeito à sua forma, conforme ilustra a Figura 4.3.

Figura 4.3 – Similaridade de comportamento das medições de velocidade média do vento



4.4 ANÁLISE MICROSCÓPICA DOS DADOS

A macro análise descrita na seção anterior focou o perfil da grandeza, para uma determinada data, independente da sua localização geográfica, enfatizando a similaridade no comportamento da medição, a despeito de sua localização dentro do estado de Minas Gerais.

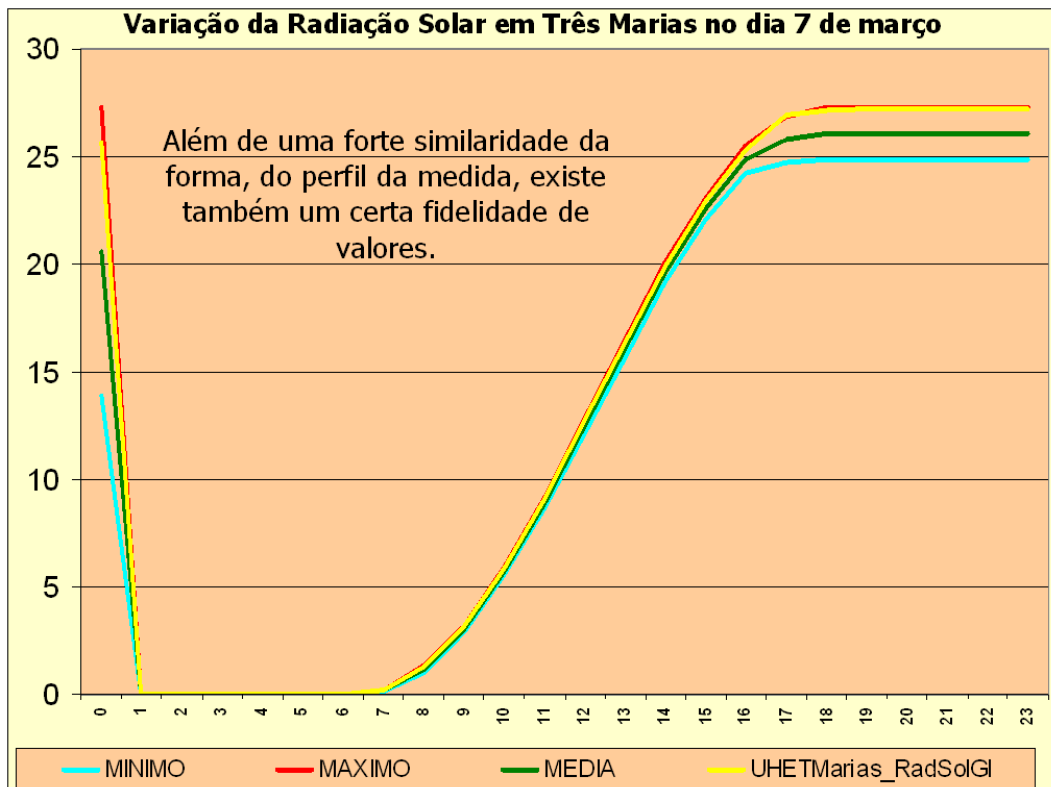
Agora, levando-se em consideração uma única localização geográfica, verifica-se que, além da forte similaridade da forma, do perfil da grandeza, existe também certa fidelidade de valores ao longo do dia, tendo como base de avaliação uma data específica do ano, escolhida aleatoriamente, para diferentes anos em questão, disponíveis na base de dados do sistema. Pode-se verificar que estes valores não se dispersam muito de um ano para o outro, indicando uma razoável repetibilidade, ficando os valores mais atuais na circunvizinhança

de uma faixa relativamente estreita margeada por um valor máximo e um mínimo histórico para o dia específico em questão.

A seqüência de figuras apresentadas a seguir (Figuras 4.4, 4.5, 4.6 e 4.7) mostram este comportamento histórico das grandezas, tendo como dia de referência o dia 7 de março de todos os anos que compõem o histórico do posto de medição. Como exemplo, as medidas a seguir foram coletadas pela unidade terminal remota do reservatório da usina de Três Marias desde o ano de 1988.

Dado este comportamento, qualquer desvio grosseiro pode ser facilmente detectado por algoritmos de comparação de valores notáveis.

Figura 4.4 – Repetibilidade de valores das medições de radiação solar acumulada para o dia 7 de março na Usina de Três Marias



Esse comportamento também pode ser percebido para as outras grandezas coletadas pelo sistema, conforme ilustram as Figuras 4.5, 4.6 e 4.7 a seguir.

Figura 4.5 – Repetibilidade de valores das medições de temperatura do ar para o dia 7 de março na Usina de Três Marias

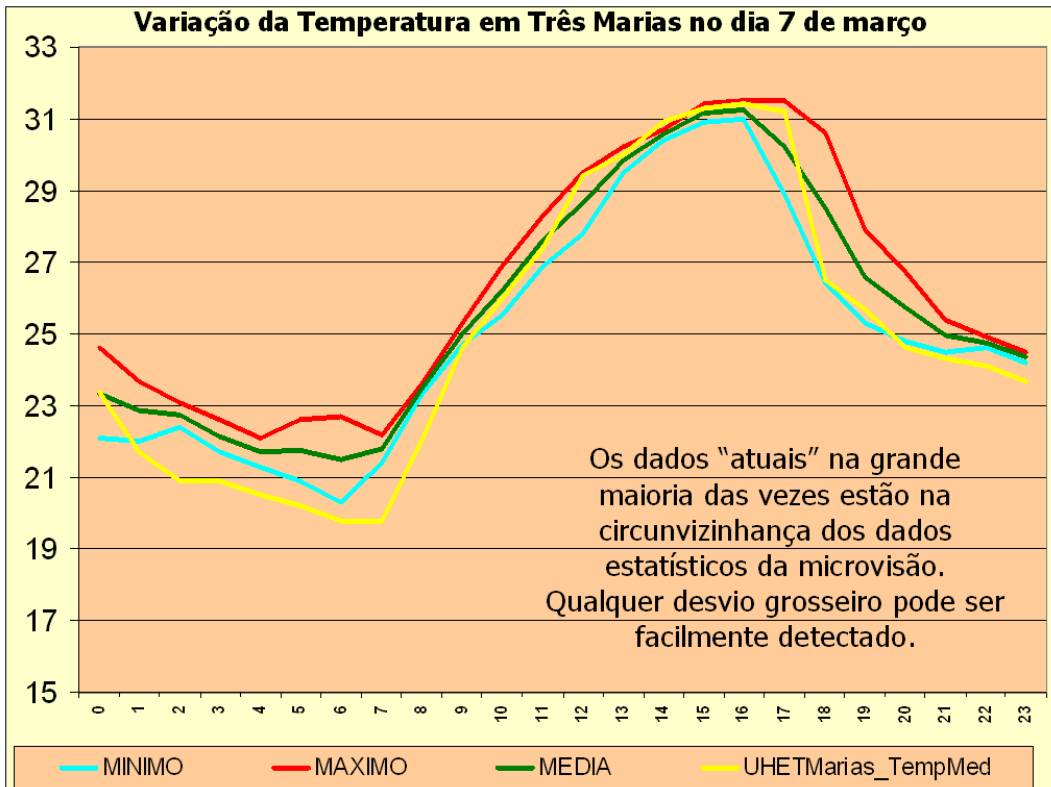


Figura 4.6 – Repetibilidade de valores das medições de pressão atmosférica para o dia 7 de março na Usina de Três Marias

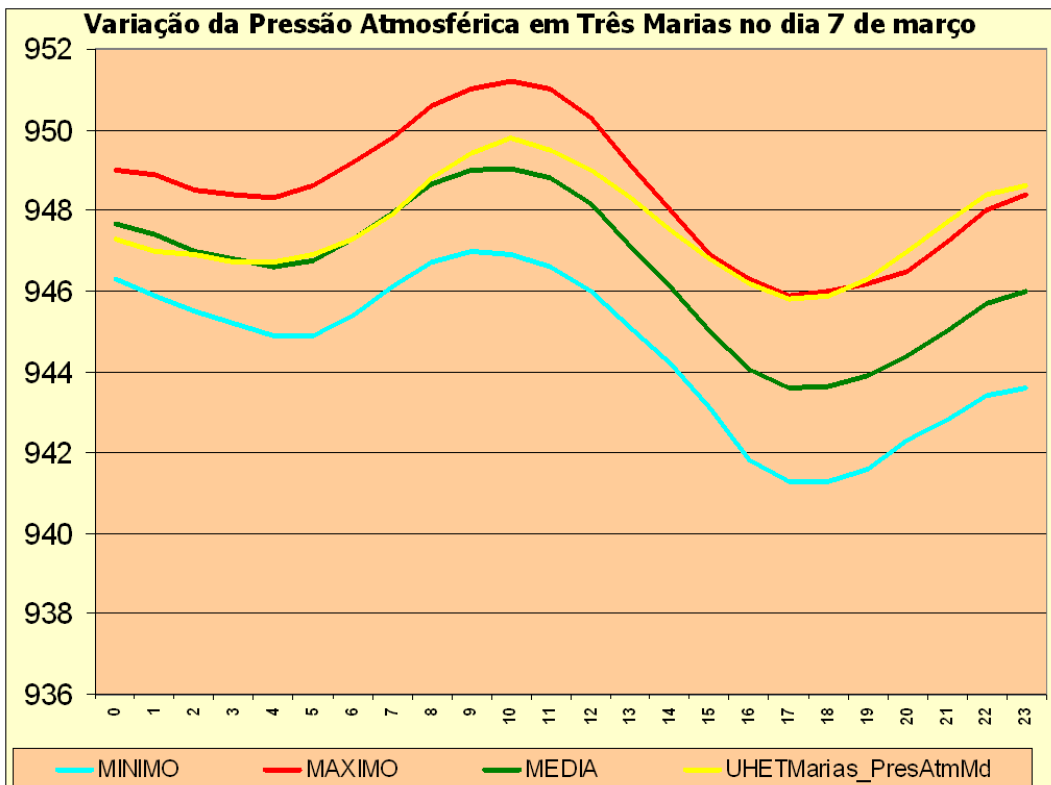
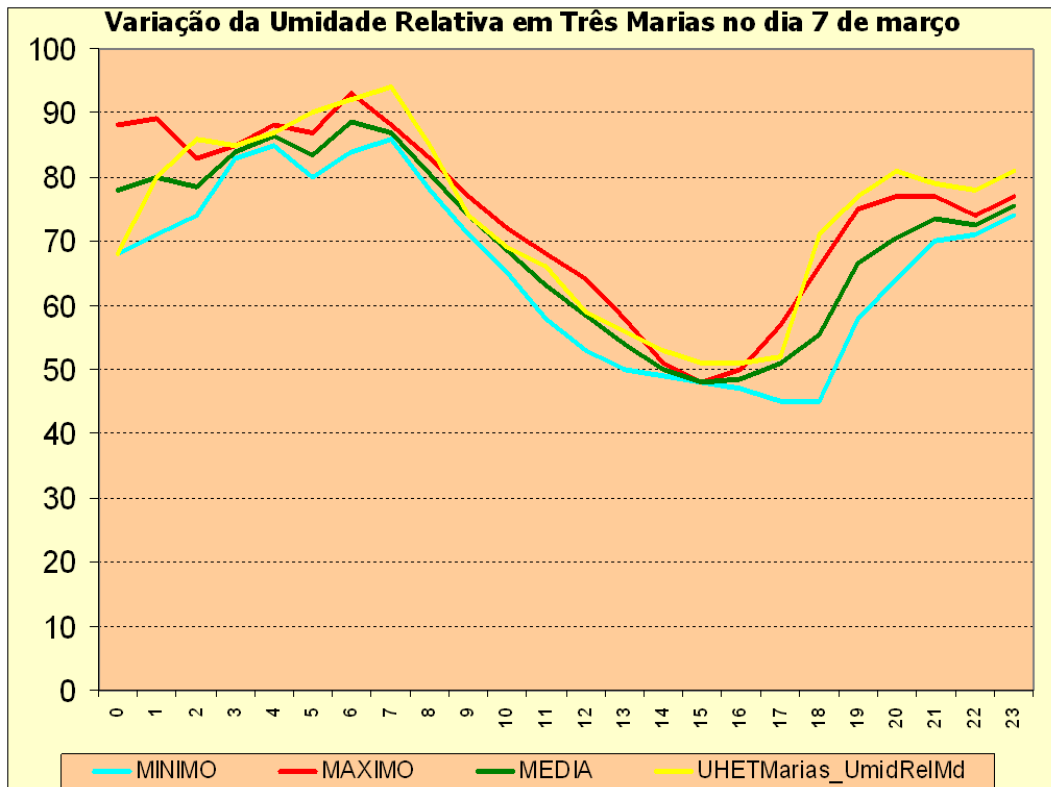


Figura 4.7 – Repetibilidade de valores das medições de umidade relativa do ar para o dia 7 de março na Usina de Três Marias



Este comportamento previsível nem sempre ocorre para todos os tipos de grandezas em análise, tendo algumas delas um comportamento bastante intempestivo, caso da velocidade e principalmente da direção do vento como pode ser visto na Figura 4.8 e Figura 4.9 a seguir. Não é sem fundamento que o indicador de direção do vento é popularmente conhecido como “biruta”.

Nestes casos, o tratamento baseado em valores históricos de caráter comportamental é menos efetivo e a eficiência da análise fica então comprometida. Para estas grandezas de comportamento mais nervoso e imprevisível, acaba-se contando apenas com as verificações de limites físicos e verificação de violações das taxas de variação.

Figura 4.8 – A medida de direção do vento possui comportamento mais nervoso e não apresentam grande repetibilidade de valores

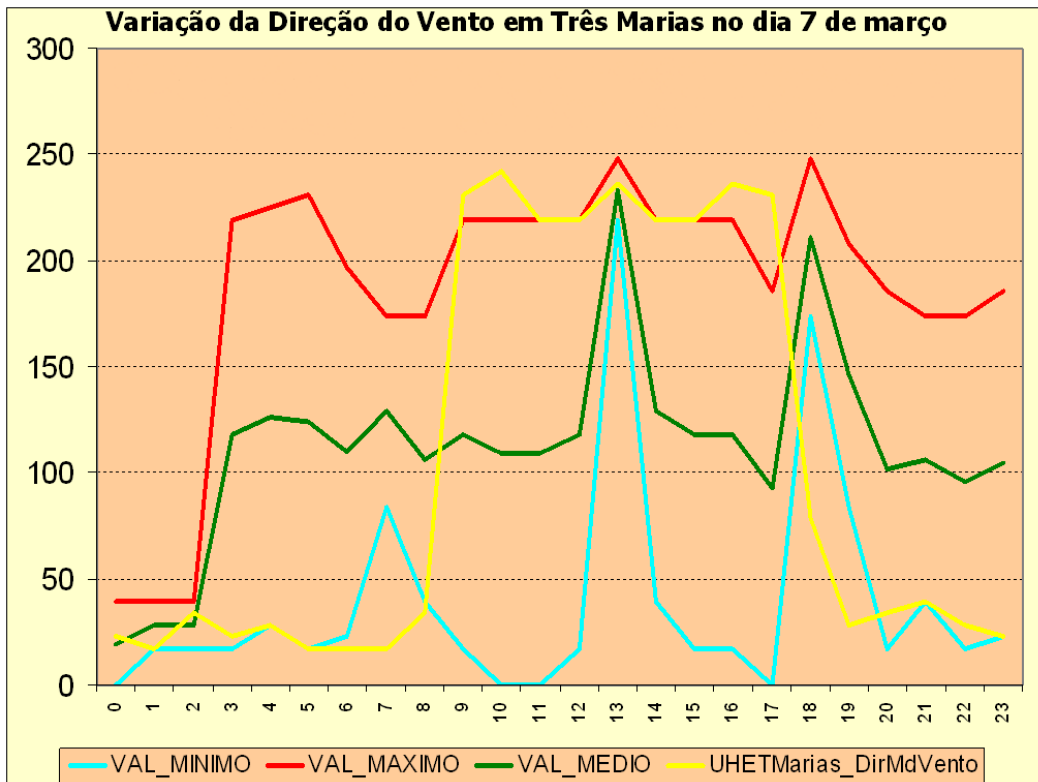
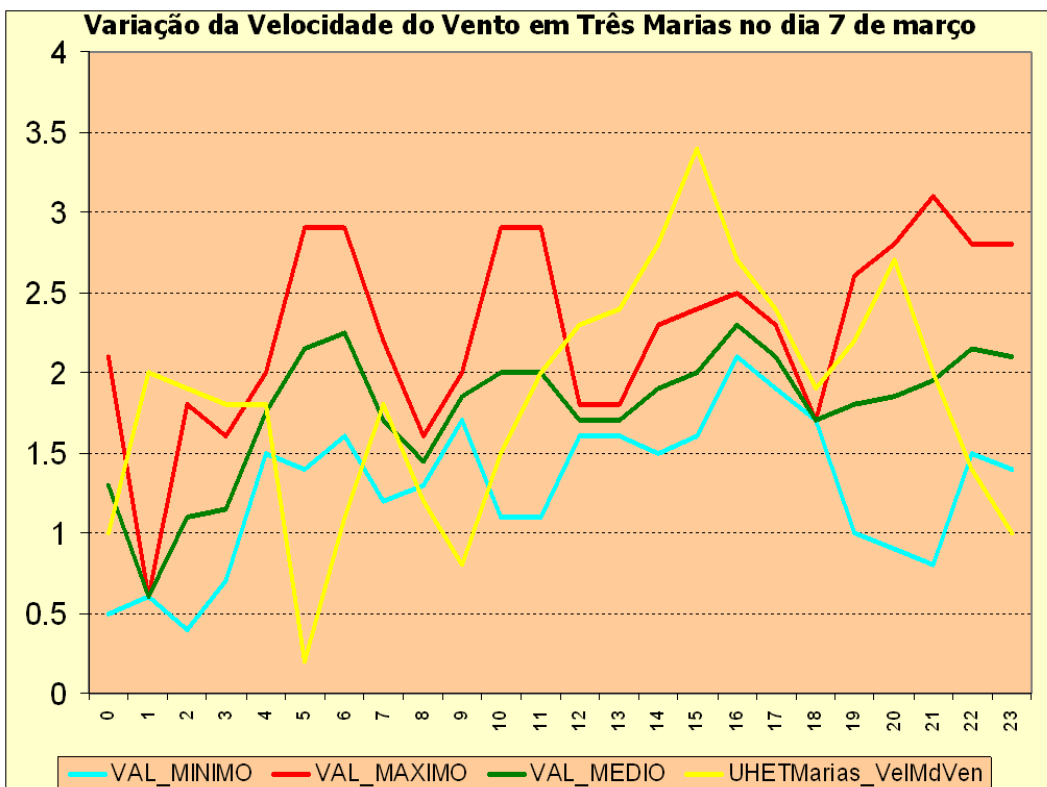


Figura 4.9 – A medição de velocidade do vento também apresenta pouca repetibilidade de valores



4.5 PROCESSAMENTO DOS DADOS RECEBIDOS

Conforme descrito anteriormente no Capítulo 3, seção 3.6, a unidade terminal remota faz uma supervisão automática do estado operacional de seus sensores e das condições do posto de medição. Sendo assim, definiu-se para o primeiro nível de tratamento que uma medida receberá uma nota de valor igual a zero (“nota zero”) se vier acompanhada de qualquer um dos seguintes (des)qualificadores:

- Medida inválida;
- Sensor defeituoso;
- Sensor desligado;
- Sensor desconectado;
- Remota em manutenção.

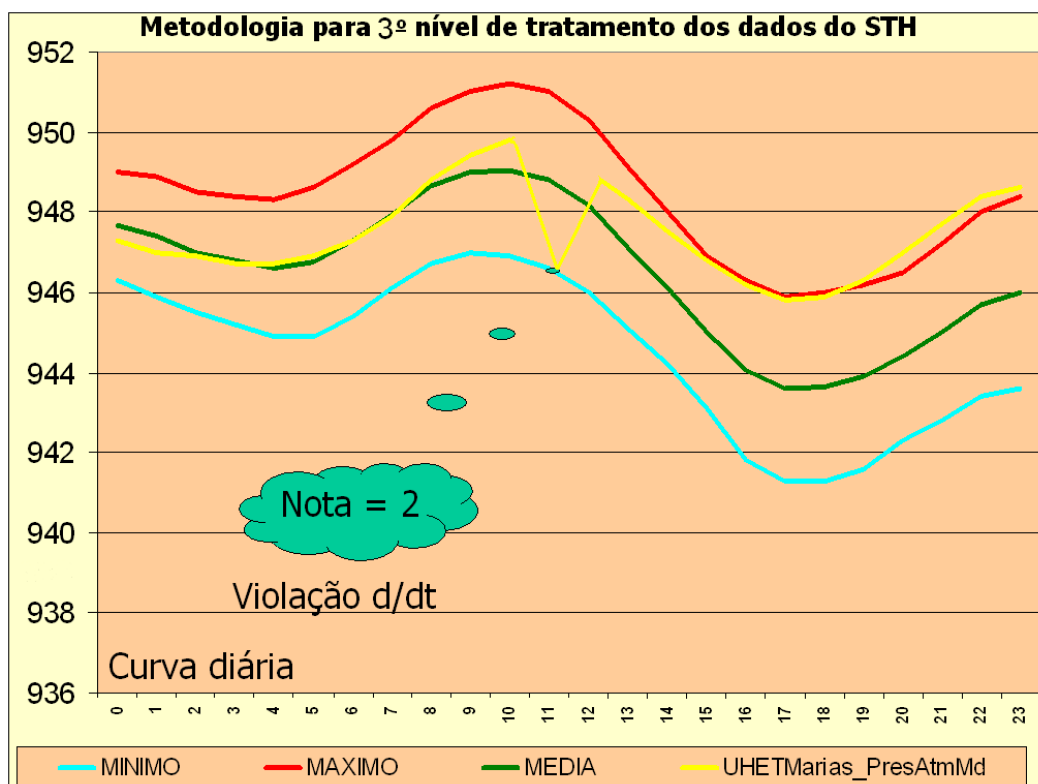
Num segundo nível de tratamento, a medição receberá uma nota de valor igual a um (“nota 1”) se tiver violado um de seus limites físicos seja o inferior ou o superior.

No terceiro nível de avaliação, receberá uma nota de valor igual a dois (“nota 2”) aquela medição que tenha violado sua taxa de variação admitida, a priori, como normal.

Estes limites usados na segunda e terceira fases de avaliação foram definidos por grandeza e por localização do posto, nas entrevistas com os hidrólogos e com os meteorologistas da Cemig, e estão armazenados numa tabela do banco de dados do STH.

A Figura 4.10 a seguir ilustra a avaliação feita no terceiro nível de tratamento da grandeza em análise, que trata de verificar se o valor recebido está violando a sua taxa de variação admissível (dv/dt).

Figura 4.10 – Processo de verificação de violação de taxa de variação



De forma diferente que os três primeiros, o quarto nível de tratamento baseia-se nos estudos estatísticos do comportamento histórico das grandezas conforme descrito anteriormente, levando-se em conta o comportamento da medição para a data e hora em avaliação. Para tal, foi criada uma grande tabela com a base de dados histórica horária para cada grandeza de cada posto de medição. Esta tabela armazena os dados estatísticos básicos (máximo, mínimo, média, número de elementos da amostra e desvio padrão) em base horária, verificados para cada grandeza ao longo do tempo de existência do posto de medição. Sendo assim, para cada tipo de informação a ser qualificado pela ferramenta proposta nessa dissertação, existem 8784 registros de valores notáveis nesta tabela. Nela foi arbitrado como ano de referência o ano 2000, apenas para compor uma data de referência que faz parte da “chave primária” da tabela, a fim de auxiliar na localização dos valores representativos da medida sob avaliação. Nesta tabela estão representadas cada hora do ano, de 1º de janeiro a 31 de dezembro, o que totalizou 3.015.447 registros estatísticos, avaliados desde 1998 até 01/01/2010, para um total de 99 localizações geográficas diferentes.

Um exemplo da estrutura desta tabela é mostrado na Tabela 4.1 a seguir. A proposta desta tabela é criar uma imagem de sazonalidade anual, do comportamento de cada uma das grandezas supervisionadas para cada um dos postos de medição que integram o sistema de aquisição de dados. A primeira coluna contém o código que identifica o posto de medição, a segunda coluna tem o código que identifica o tipo de informação dado pela medição, a terceira coluna indica a data e hora da medição em questão, onde o ano 2000 foi arbitrado apenas para fins de consulta à base. As quarta, quinta e sexta colunas indicam respectivamente os valores mínimo, máximo e médio, verificados para a grandeza ao longo dos anos. A sétima e oitava colunas indicam o número de elementos na amostra avaliada e a data da última atualização de dados.

Tabela 4.1 – Estrutura da tabela de dados estatísticos

COD_ELEMENTO_HIDR	COD_TIPO_INFO	DAT_REFERENCIA	VAL_MINIMO	VAL_MAXIMO	VAL_MEDIO	NUM_ELEM_AMOSTRA	DAT_ATUALIZACAO
33	12	22/12/2000	17	39	33,5	11	01/01/2010
33	16	22/12/2000	554	560,1	557,7	11	01/01/2010
33	18	22/12/2000	945,5	949,5	947,2	11	01/01/2010
33	21	22/12/2000	4,83	7,37	6,2	11	01/01/2010
33	22	22/12/2000	25,1	27,7	26,52	11	01/01/2010
33	25	22/12/2000	61	79	67,75	11	01/01/2010
33	32	22/12/2000	1,6	5,8	3,7	11	01/01/2010

Quando uma medição é recebida pelo sistema de aquisição de dados ela é então confrontada com os dados estatísticos armazenados na tabela de estatísticas, para a data e hora da medição, desprezando-se aqui o ano em questão.

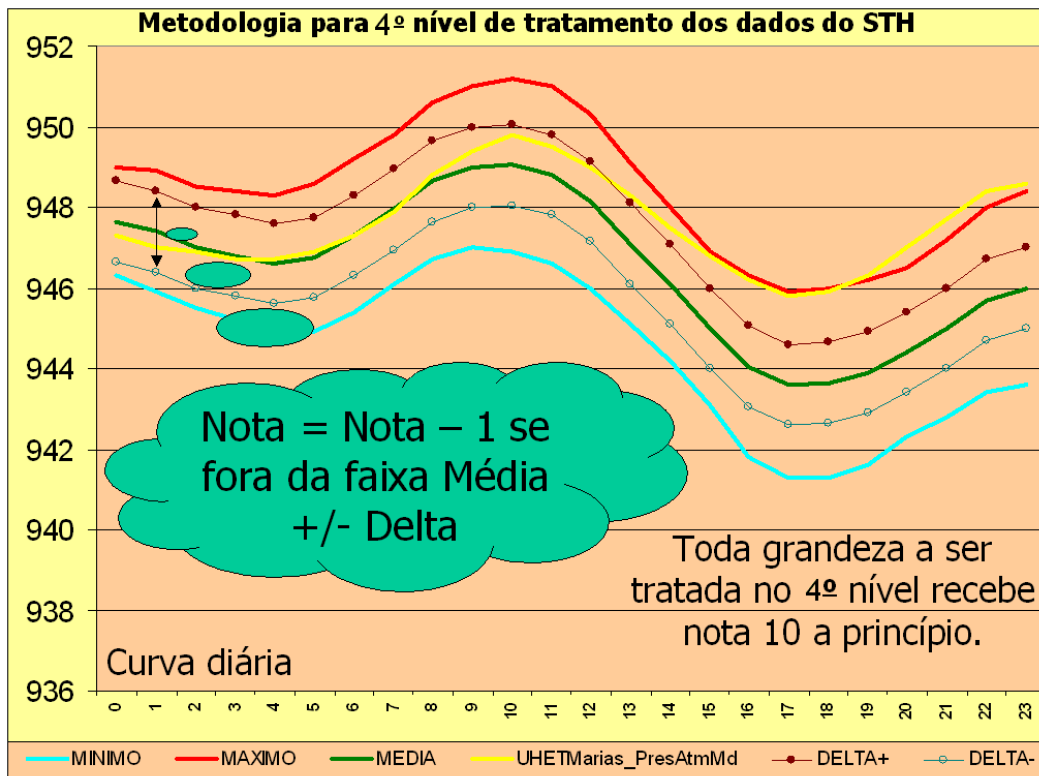
Este quarto nível de tratamento da grandeza prossegue por várias fases dependendo da qualidade com o qual a medida foi recebida, conforme descrito nos parágrafos seguintes.

Se uma medida passou pelos testes de verificação do primeiro, segundo e terceiro níveis de tratamento atribui-se a ela a nota máxima, de valor igual a dez (“nota 10”). Iniciam-se então os testes do quarto nível de tratamento. Nesta fase, para cada um dos testes no qual a grandeza seja rejeitada, a nota é subtraída de 1 ponto, assim sucessivamente, até que se chegue ao final da cadeia de testes estatísticos do quarto nível de avaliação.

No primeiro passo do quarto nível de tratamento verifica-se se a medida recebida se encontra dentro de uma faixa de valores limitada pela média histórica dos valores para aquela “estampa de tempo” da grandeza (hora, dia e mês), mais ou menos um “delta”. Se ela se encontra nesta faixa, o tratamento se encerra com a medida recebendo “nota 10” sem ter perdido nem um ponto. Este valor do delta, inicialmente teve como proposta o desvio padrão do conjunto de medidas disponíveis para o dia, hora e mês associado à grandeza. Porém o número de elementos de cada uma das amostras ainda é muito pequeno, sendo no máximo 12, uma vez que cada ano de vida do posto de medição contribui com uma única e específica medida, identificada pela sua hora, dia e mês de ocorrência. Essa situação torna o desvio padrão pouco representativo e, sendo assim, decidiu-se então usar inicialmente um valor de delta fornecido pelos hidrólogos e meteorologistas da Gerência de Planejamento Energético da Cemig.

A Figura 4.11 a seguir ilustra este primeiro passo de tratamento da medida recebida, lembrando-se que toda medida que chega ao quarto nível de tratamento possui inicialmente a nota máxima (“nota 10”).

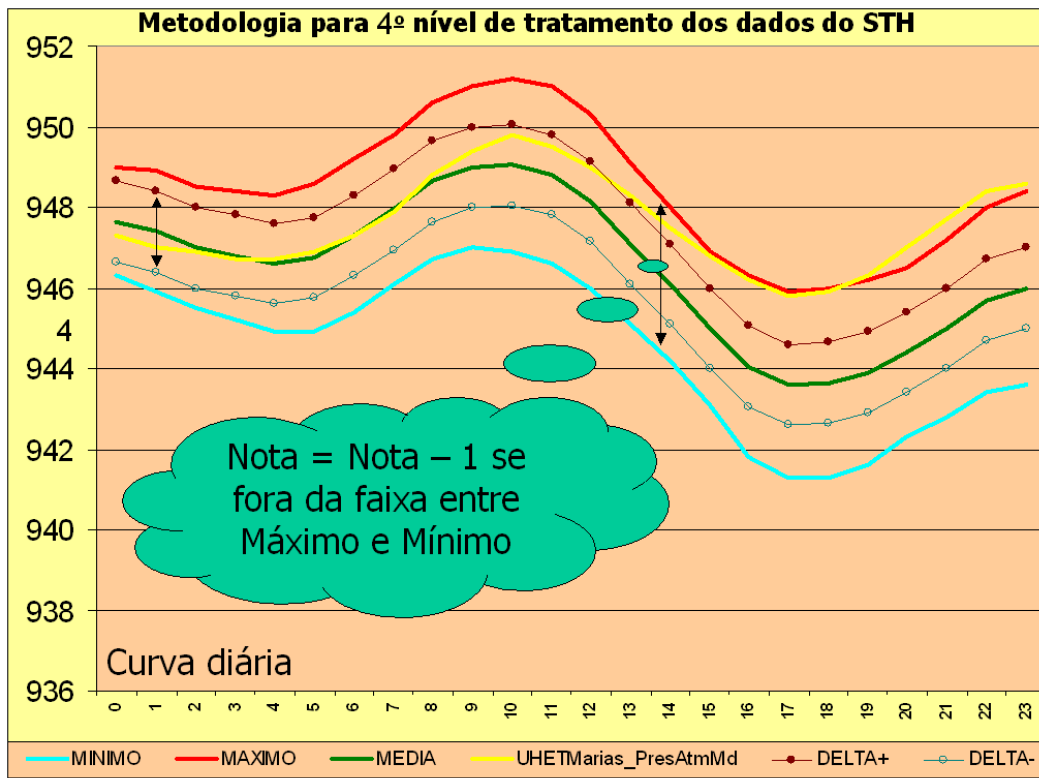
Figura 4.11 – Primeiro teste da quarta fase de tratamento de uma medida



No segundo passo do quarto nível de tratamento verifica-se se a medida recebida se encontra dentro de uma faixa de valores limitada pelos valores históricos de máximo e mínimo verificados para aquela “estampa de tempo” da grandeza (hora, dia e mês). Se a seqüência de testes alcançou este passo, significa que a medida não passou pelo passo anterior e se encontra com “nota 9”. Estando a medida dentro da faixa de limitação, entre o máximo e mínimo históricos para a hora, dia e mês em questão, o processamento se encerra com a medida sendo pontuada com a “nota 9” recebida anteriormente, caso contrário ela segue no processo de qualificação.

A Figura 4.12 a seguir ilustra este segundo passo de tratamento da medida recebida, lembrando-se que toda medida que chega ao segundo passo de avaliação possui inicialmente “nota 9”.

Figura 4.12 – Segundo teste da quarta fase de tratamento de uma medida



Do terceiro passo do quarto nível de tratamento até o final da cadeia de testes, verifica-se se a medida recebida se encontra dentro de uma faixa de valores limitada pelos valores históricos de máximo e mínimo verificados para uma janela de tempo, que tendo como ponto central a hora, o dia e o mês da medida recebida se estende por 15, 30, 90, 180 e 366 dias. Estando a medida dentro da faixa de limitação, entre o máximo e mínimo históricos para a janela de tempo em análise, o processamento se encerra com a medida sendo pontuada com a resultante das subtrações sucessivas de 1 ponto da nota em relação àquela nota recebida no passo anterior. No pior caso, que diz respeito à janela de 366 dias (considerando aqui também os anos bissextos), ou seja, uma janela que cobre todo o histórico anual das estatísticas da medida, a mesma receberá nota 3, indicando que se trata de um novo valor limite, seja um máximo *maximorum* (máximo dos máximos) ou um mínimo *minimorum* (mínimo dos mínimos), e que neste caso, o valor deve ser observado com maiores critérios pelos especialistas, hidrólogos e meteorologistas.

As Figuras 4.13 e 4.14 a seguir ilustram estes passos de tratamento da medida recebida, que se baseiam numa janela de tempo que se expande continuamente.

Figura 4.13 – Terceiro teste da quarta fase de tratamento de uma medida

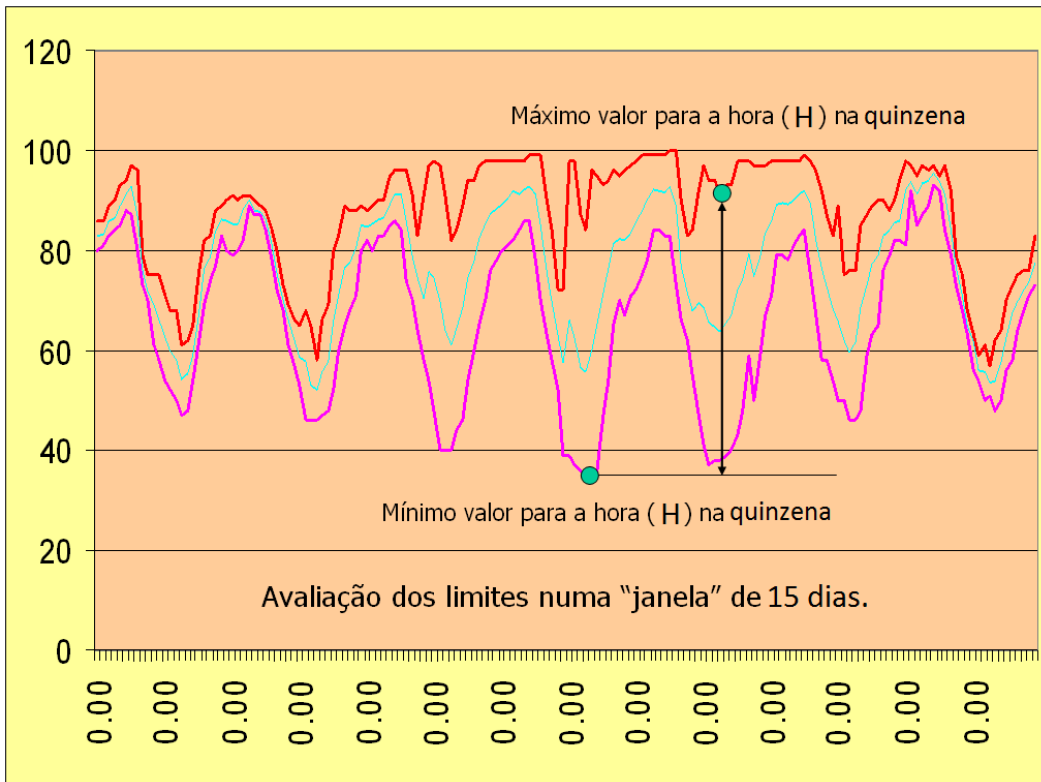
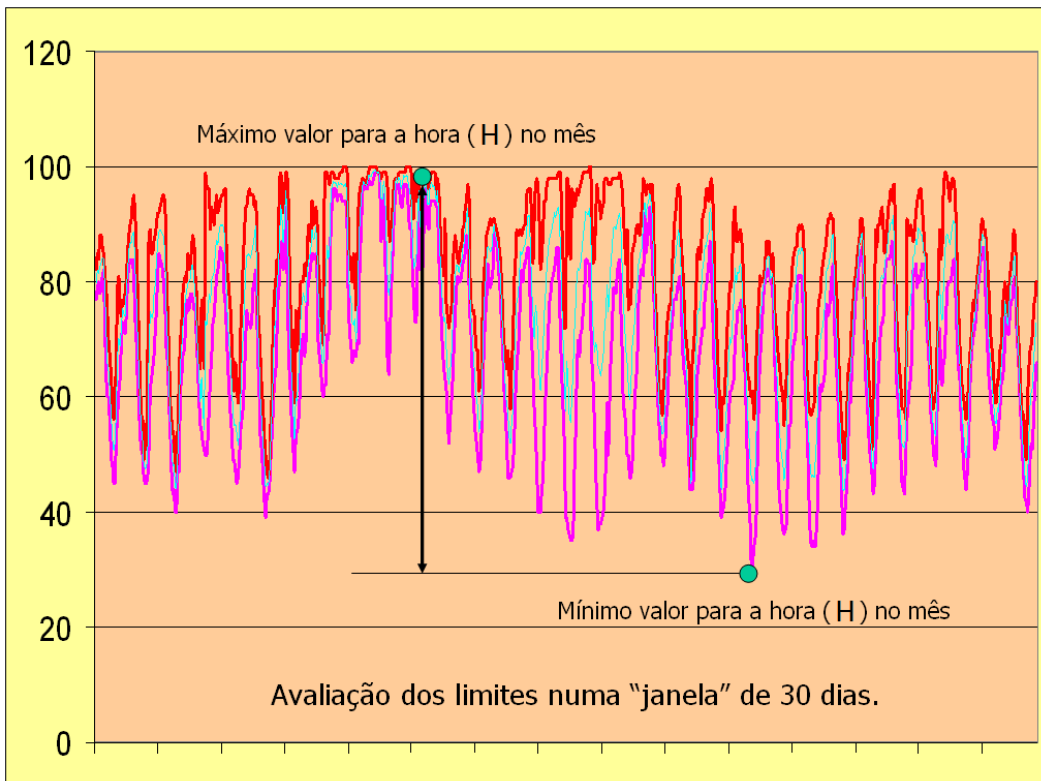


Figura 4.14 – Quarto teste da quarta fase de tratamento de uma medida



Arbitrou-se a princípio como sendo uma boa medida aquela que obtivesse uma nota igual ou superior a 7 (nota azul) e para as outras menores que sete a nota seria vermelha, indicando que a medida está muito afastada de seus valores históricos, devendo ser usada com certa cautela ou até mesmo desconsiderada como no caso de nota zero (qualificadores vindos da remota que invalidam a medição), nota 1 (medida viola limite físico) ou nota 2 (medida viola taxa de variação admitida).

Esta cadeia de testes sucessivos pode ser visualizada resumidamente através da Figura 4.15 a seguir, que sumariza a nota da medida baseada na sua “localização” em relação à sua média histórica.

Figura 4.15 – Resumo ilustrado da cadeia de testes sucessivos da quarta fase de avaliação de uma telemedição

Resumo	...Maximorum	Nota=Nota -1
	Lim. Sup. Hora (90)...	Nota 5
	Lim. Sup. Hora (30 dias)	Nota 6
	Lim. Sup. Hora (15 dias)	Nota 7
	Lim. Sup. Hora	Nota 8
	Média + delta	Nota 9
Média		Nota 10
Média		Nota 10
	Média - delta	Nota 9
	Lim. Inf. Hora	Nota 8
	Lim. Inf. Hora (15 dias)	Nota 7
	Lim. Inf. Hora (30 dias)	Nota 6
	Lim. Inf. Hora (90)...	Nota 5
	...Minimorum	Nota=Nota -1

Como pode ser visto pela Figura 4.15, a quarta fase de tratamento de uma telemedição pode realizar no mínimo 1 e no máximo 7 testes de comparação da medida recebida com seus dados históricos. A nota de uma medida decresce de 1 ponto de um teste para o outro e pode alcançar, no pior caso, a nota 3.

5 Desenvolvendo os Algoritmos

5.1 INTRODUÇÃO

Nos capítulos anteriores foram apresentada a descrição do STH, os trabalhos realizados para a preparação inicial dos dados e a metodologia proposta para o tratamento das medições recebidas. É chegada a hora de se desenvolver um *software* que realize todo o trabalho de “mastigação numérica” dos dados a fim de apresentar os resultados esperados.

Este *software* em sua versão final será integrado aos atuais módulos do STH, sem comprometer o funcionamento e o desempenho do sistema em operação. Para que isso ocorra sem surpresas desagradáveis, todo o cuidado na sua elaboração deverá ser tomado e exaustivos testes deverão ser realizados antes de sua integração definitiva no sistema de produção. Para tanto, deverá existir uma fase de homologação que garanta o perfeito funcionamento do produto final, objetivo deste trabalho.

5.2 CONDIÇÕES EXISTENTES

O STH, antes de sua migração para o ambiente WEB, foi originalmente desenvolvido utilizando-se como plataforma o sistema operacional Microsoft Windows NT 4.0 (SP6a). Era baseado numa arquitetura cliente-servidor e na análise e programação orientada a objetos. Os códigos-fontes dos programas desenvolvidos foram escritos em três diferentes linguagens, C++, Fortran e Visual Basic 6.0® e o gerenciador da base de dados era o Oracle®.

Com o passar do tempo, a plataforma inicial foi sofrendo alterações e atualizações de seu *software* básico (sistemas operacionais e gerenciadores de base de dados relacionais), sem se alterar o *software* original do STH. O sistema foi suportado pelo *Microsoft Windows 2000®* e atualmente está no *Microsoft Windows 2003 Standard Server®*. O gerenciador de base de dados é agora o *Microsoft SQL Server®* em substituição ao *Oracle®* devido a seus elevados custos de licenciamento e manutenção de contratos de suporte.

Em 2006, a interface homem-máquina do sistema foi migrada para o padrão WEB, e o sistema passou a ser acessado através da intranet da Cemig via *Microsoft Internet Explorer*. Essa interface gráfica foi totalmente reescrita na linguagem C# (*code behind*), *ASP.NET*, *HTML*, *AJAX* e *JAVA Script*.

5.3 ESTRATÉGIA DE DESENVOLVIMENTO

Para validar a idéia da metodologia para a qualificação dos dados, detalhada no Capítulo 4, foi construído um programa de computador que executa todo o procedimento de validação proposto naquele capítulo. Este programa inicial, entretanto, funciona de forma *off-line*, desacoplado dos aplicativos que compõem os diversos módulos computacionais do STH.

A linguagem de programação escolhida para o desenvolvimento deste programa foi o *Microsoft Visual Basic 6.0®*, por ser uma das linguagens usadas no desenvolvimento de diversos programas que compõem o sistema original como um todo.

O desenvolvimento deste protótipo inicial foi fundamental, pois permitiu a realização de testes de desempenho e a mensuração do impacto causado no desempenho do sistema. Além disso, possibilitou medir o consumo de recursos do servidor do sistema e do servidor do banco de dados relacional, necessários à realização dos processamentos de validação da grandeza recebida.

Após diversos testes com o protótipo desenvolvido, verificou-se que a metodologia proposta se mostrou bastante acertada, uma vez que conseguiu indicar um grande número de anomalias nas grandezas armazenadas. A partir desse protótipo, iniciou-se o desenvolvimento da ferramenta definitiva a ser integrada ao sistema, fazendo a qualificação dos dados recebidos de forma *on-line* e em tempo real.

A estratégia adotada para o desenvolvimento do aplicativo definitivo foi a de criação de uma DLL (*Dynamic Link Library*) de forma que a mesma possa ser acoplada aos diversos módulos executáveis do STH que necessitam validar uma telemedição recebida, como por exemplo, o SCADA, principal usuário da DLL proposta, uma vez que é o módulo responsável pela recepção das mensagens de comunicação que transportam os dados telemedidos, sua decodificação e disponibilização das telemedições no banco de dados do sistema. Além do

SCADA, outros módulos podem se beneficiar do tratamento de dados, como por exemplo, o módulo de cálculos periódicos, o sistema de monitoramento de inequações e o módulo de interface de comunicação com outros sistemas.

A figura 5.1 a seguir mostra a tela principal do protótipo *off-line* desenvolvido para a validação inicial da idéia apresentada nesta dissertação, onde se percebe a estratificação das grandezas avaliadas baseando-se nas notas recebidas por cada uma delas.

Figura 5.1 – Tela principal do aplicativo de qualificação “off-line”

	Temperatura	Umidade	Pressão	Radiação	Dir. Vento	Vel. Vento	N. Água	Chuva
Nota 10	62,56%	63,77%	40,98%	58,56%	83,28%	95,07%	50,57%	97,05%
Nota 9	10,87%	11,32%	22,69%	13,92%	2,89%	0,03%	23,40%	0,00%
Nota 8	17,48%	14,04%	22,61%	15,37%	7,92%	0,16%	8,64%	0,00%
Nota 7	2,52%	2,31%	3,05%	2,35%	0,72%	0,04%	4,51%	0,00%
Nota 6	1,69%	2,58%	2,18%	2,33%	0,43%	0,04%	5,06%	0,00%
Nota 5	0,16%	1,02%	0,48%	0,65%	0,07%	0,02%	2,00%	0,00%
Nota 4	0,03%	0,11%	0,14%	0,59%	0,06%	0,01%	0,74%	0,00%
Nota 3	0,06%	0,08%	0,16%	1,61%	0,01%	0,03%	0,99%	0,00%
Nota 2	0,03%	0,01%	3,08%	0,01%	0,00%	0,00%	0,11%	0,01%
Nota 1	0,00%	0,14%	0,00%	0,00%	0,00%	0,00%	0,04%	0,00%
Nota 0	4,60%	4,64%	4,63%	4,60%	4,62%	4,60%	3,93%	2,94%
N. Med	495254	493144	493146	495515	493151	495761	1603405	1567454
1425 - 51 - 31/12/2009 23:45:00 - 570 - 10						Total de Medidas:		6139201
Data Inicial:	01/07/2009	Data Final:	01/01/2010	Data Atual:	31/12/2009			
Progresso Data Atual								
Progresso Geral								
Início Processamento:			30/12/2009 10:43:20	Fim Processamento:		09/01/2010 16:51:03		
Partir			Parar			Sair		

O desenvolvimento do aplicativo para funcionamento *off-line* permitiu verificar a validade da metodologia proposta para a qualificação dos dados telemidados além de permitir a realização de testes de desempenho e de medições de consumo de recursos do sistema computacional. Pôde-se observar que o tempo médio para qualificação de uma grandeza fica em torno de 130 milésimos de segundo, considerado adequado para a proposta do mesmo.

Dado os resultados positivos verificados com o sistema de qualificação que opera em modo *off-line*, foi construído então a DLL que acoplada ao módulo SCADA passou a fazer a qualificação das medidas recebidas em tempo-real.

5.4 TAREFA DE PÓS-PROCESSAMENTO

À medida que os dados vão sendo recebidos pelo sistema de telemetria, qualificados pela metodologia proposta e armazenados no banco de dados, um grande volume de medidas passa a estar disponível, tornando possível a utilização de novas amostras de dados para a atualização dos dados estatísticos que representam as mesmas. Desta forma foi necessário desenvolver uma outra ferramenta computacional que atualizasse a tabela de estatísticas recalculando os valores de máximo, mínimo e de média aritmética devido às novas amostras recebidas.

Como um dado estatístico atualizado só servirá para a validação do dado de mesma data do ano seguinte, não se mostrou necessário o desenvolvimento de uma aplicação de atualização da tabela de estatísticas que funcione em tempo real. Sendo assim, essa ferramenta é usada sob demanda, dentro dos procedimentos de manutenção e atualização do sistema. Foi então desenvolvido um programa que é executado para realizar a atualização da tabela de estatísticas. A fim de possibilitar a interferência do especialista, hidrólogo ou meteorologista, este programa permite definir a nota mínima, separadamente por tipo de grandeza, que será admitida como adequada para a realização da atualização da tabela principal de estatísticas. Dessa forma, pode-se filtrar dados considerados incoerentes, não permitindo que os mesmos incorporem erros grosseiros aos cálculos estatísticos.

A Figura 5.2 a seguir mostra a tela principal do programa utilizado para a atualização das estatísticas das grandezas. Neste exemplo, a nota de corte usada foi a nota 3, o que permite ao processo atualizar os dados de mínimos *minimorum* e máximos *maximorum*.

Figura 5.2 – Tela principal do aplicativo de atualização dos dados estatísticos das grandezas

Atualização das estatísticas das grandezas

Indique no quadro abaixo as notas mínimas válidas para atualização

Temperatura	Umidade	Pressão	Radiação	DVento	VVento	NÁgua
3	3	3	3	3	3	3

Início: 01/07/2007 Fim: 01/08/2007

74 - 18 - 1/7/2007 01:45:00 - 933,2 - 8

Iniciar

6 Resultados Iniciais Encontrados

6.1 INTRODUÇÃO

A fim de ser possível verificar a eficiência da ferramenta de qualificação de dados proposta nesta dissertação, era necessário se obter um panorama da situação dos milhões de medidas disponíveis na base de dados do sistema, no que diz respeito à qualidade da medida antes da mesma ter recebido qualquer tipo de tratamento. Esse panorama inicial tornou possível uma comparação com aquele que se observou após a execução das avaliações qualitativas das medições, feitas pelos algoritmos desenvolvidos para esse fim. Essa comparação dos resultados é que, de forma efetiva, pôde dizer se os tratamentos propostos são realmente eficazes, validando o modelo desenvolvido.

Este capítulo descreve algumas ocorrências interessantes, percebidas durante as análises iniciais das grandezas e apresenta alguns dos resultados encontrados.

6.2 AVALIAÇÃO DAS MEDIÇÕES DE TEMPERATURA

A Tabela 6.1 a seguir mostra um resumo do universo de dados de temperatura do ar, sem nenhum tipo de tratamento, que representam todos os dados recebidos pelo sistema durante a fase de avaliação inicial das grandezas, que foi feita a fim de se localizarem os erros grosseiros presentes na amostra de dados. O número de elementos da amostra é de 2.209.165 (dois milhões duzentos e nove mil cento e sessenta e cinco) medidas, recebidas até o dia 01/01/2010 e de forma geral representam a situação inicialmente verificada dentro do estado de Minas Gerais.

Os dados referentes à temperatura do ar produziram algumas ocorrências inusitadas, dentre elas a ocorrência de temperaturas abaixo de 0° (zero graus) em localidades onde normalmente temos um calor intenso e mais curioso ainda, temperaturas da ordem do ponto de ebulição da água. É claro que se devem à medições equivocadas, oriundas de sensores defeituosos ou de manutenções realizadas nas unidades terminais remotas sem a devida sinalização para o sistema de recepção de dados. Essa possibilidade de ocorrência é justificável durante a fase de implantação e montagem dos equipamentos no campo.

Fazendo-se uma estratificação dos dados, que representam todo o estado de Minas Gerais, avaliando-se os mesmos por sua localização geográfica, pudemos perceber que vários postos de medição registraram temperaturas incompatíveis com sua localização, conforme tabela a seguir, sendo necessário para alguns desses casos uma avaliação mais rigorosa, a fim de decidir pelo descarte das medidas.

Tabela 6.1 - Dados de Temperatura coletados pelo STH revelando algumas inconsistências

LOCALIZAÇÃO	MINÍMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	13,0	38,4	24,0	3,8	26.617
Bocaiúva	9,1	35,9	21,6	3,9	41.211
Botelhos	3,2	32,8	19,8	4,3	26.790
Buritizeiro	0,0	38,4	23,6	5,0	81.871
Camargos	0,0	33,5	4,8	8,8	84.254
Carmo de Minas	-1,4	33,6	17,7	6,8	90.964
Coa - Belo Horizonte	7,3	33,1	21,3	3,6	28.503
Curvelo	10,1	35,8	23,0	4,3	25.030
Diamantina	0,0	29,2	18,0	3,5	26.543
Fazenda Campo Grande	0,0	31,6	4,4	8,4	57.369
Fazenda Energetica Uberaba	1,2	35,5	21,9	4,9	27.841
Felisburgo	10,9	33,3	21,2	3,5	25.265
Gafanhoto	6,6	33,6	21,4	4,5	17.796
Hotel Fazenda Curva do Rio	0,0	54,3	23,7	7,0	17.894
Igarapé	10,8	33,2	22,6	4,0	2.650
Igicatu	0,0	35,9	21,0	3,9	34.091
Itambacuri	0,0	35,6	21,7	3,9	25.965
Januária	0,0	39,0	24,2	5,0	28.162
Juiz de Fora	0,0	32,7	19,8	4,3	24.017
Machado Mineiro	0,0	35,8	21,3	4,9	52.806
Nova Ponte	0,0	40,9	21,7	4,7	84.632
Paracatu	7,4	34,0	21,9	3,6	22.578
Passos	0,0	35,3	20,5	5,9	46.295

LOCALIZAÇÃO	MINÍMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Patos de Minas 1	6,7	34,7	20,7	3,9	57.946
Peti	5,5	35,0	20,4	4,2	23.410
Prata	4,2	37,1	23,3	4,9	27.916
Quartel Geral	0,0	37,5	22,2	4,5	71.774
Salinas	0,0	38,0	23,0	4,5	47.123
São Pedro da Ponte Firme	4,7	33,2	20,9	3,9	24.505
São Roque de Minas	0,9	31,0	18,3	3,6	85.904
São Simão	0,0	39,1	24,7	5,3	90.293
SE Alpinópolis	5,8	100,1	21,0	3,8	36.136
SE Araçuaí	17,5	39,4	26,6	4,1	3.072
SE Barreiro	0,0	31,4	18,9	5,3	55.070
SE Curvelo	0,0	36,3	22,8	4,1	57.154
SE Diamantina	0,0	30,6	18,5	3,3	51.222
SE Governador Valadares	0,0	38,3	24,5	3,9	60.274
SE Ituiutaba	0,0	37,3	22,7	6,3	36.011
SE Januária	11,3	38,9	24,9	4,0	52.202
SE Juiz de Fora	0,0	34,2	18,0	7,2	60.211
SE Mesquita	0,0	37,0	23,2	3,6	60.803
SE Montes Claros	0,0	36,8	18,7	10,0	25.823
SE Ouro Preto	0,0	29,8	13,2	8,7	8.819
SE Pouso Alegre 1	0,0	33,4	18,0	7,2	54.462
SE Teófilo Otoni	0,0	37,7	23,0	4,3	34.539
SE Uberaba	0,0	35,8	19,9	7,9	56.080
Sede	-1,3	33,5	20,6	5,7	13.760
Sete Lagoas - EFAP	10,4	34,2	22,3	3,5	11.344
Três Marias	0,0	51,6	10,1	12,0	90.821
Turmalina	0,0	35,9	22,0	4,6	23.218
Unai	9,3	34,6	21,7	4,0	22.215
Vargem Alegre	0,0	37,5	23,3	4,3	24.158
Viçosa	0,0	34,2	19,9	4,2	43.756

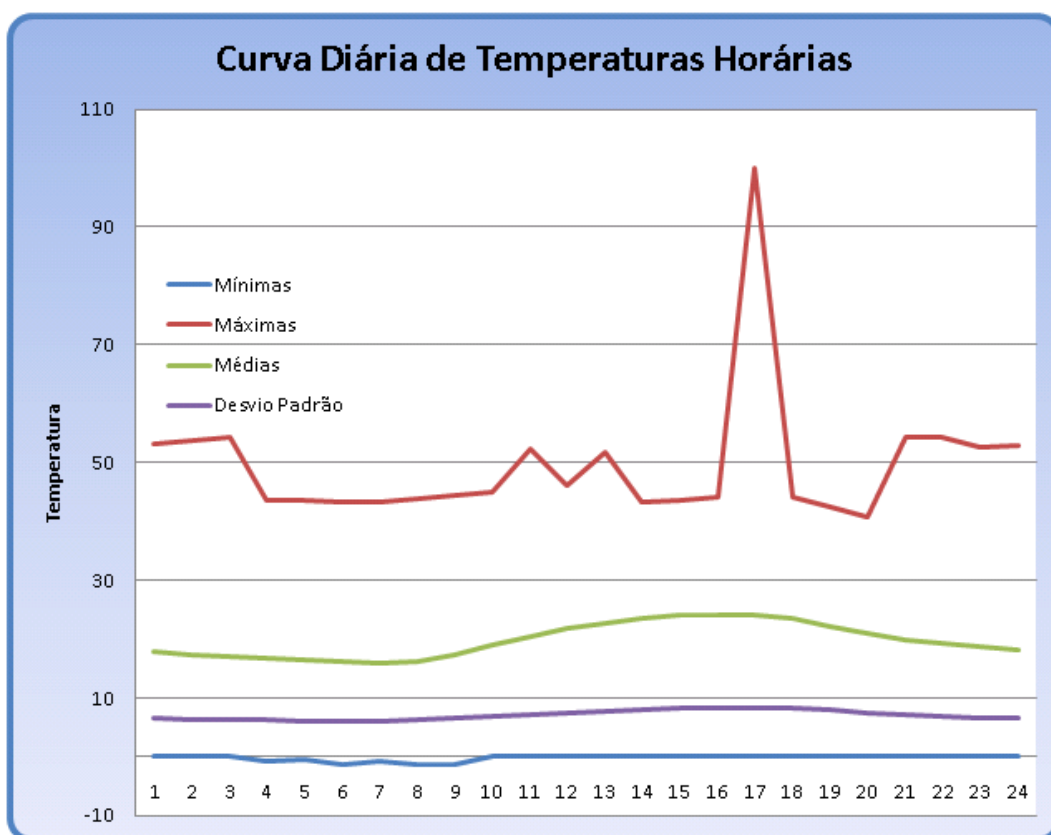
Resumindo a tabela de valores de temperatura, mostrada anteriormente numa tabela menor (Tabela 6.2) mostrada a seguir, que mostra para todo o estado de Minas Gerais, os valores extremos verificados, suas médias e desvios padrão, verifica-se a seguinte representação.

Tabela 6.2 – Resumo dos Dados de Temperatura coletados pelo STH revelando algumas inconsistências

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	-1,4	29,2	4,4	3,3
Máximo	17,5	100,1	26,6	12,0
Médio	2,9	37,1	20,5	5,1
Desvio Padrão	4,5	9,7	4,2	1,8

Uma imagem diz mais que mil palavras. Sendo assim, foi feita uma estratificação dos dados mostrados nas tabelas anteriores, avaliando-se os mesmos mais detalhadamente no que diz respeito ao seu perfil de comportamento horário. Esta avaliação resultou numa imagem, representada pelo Gráfico 6.1 mostrado a seguir, que ilustra a curva típica diária, das medidas de temperatura mínima, máxima e média, em graus Celsius, verificadas no estado de Minas Gerais como um todo, para cada uma das 24 horas do dia.

Gráfico 6.1 – Perfil do comportamento das medidas de temperatura do ar coletados pelo STH revelando algumas inconsistências



Mesmo para os não meteorologistas o gráfico anterior faz perceber verdadeiros absurdos, principalmente nas medidas que se referem às temperaturas horárias máximas e mínimas. Espera-se que a ferramenta de qualificação das medidas, proposta nesta dissertação seja capaz de filtrar e identificar todas as inconsistências encontradas e claramente evidenciadas nas tabelas e no gráfico anterior. Isso é o que se discute no próximo capítulo desta dissertação.

Vale ainda comentar a influência do tamanho de uma amostra de dados nos resultados obtidos para seus dados notáveis. Observando no gráfico anterior a curva das temperaturas médias horárias, percebe-se que, mesmo com as grandes inconsistências evidenciadas para os valores máximos e mínimos, a mesma apresenta um perfil e uma seqüência de valores compatível com o que se esperaria para o estado de Minas Gerais, evidenciando o fenômeno de regressão à média. Vale lembrar que esta amostra de dados é composta de 2.209.165 (dois milhões duzentos e nove mil cento e sessenta e cinco) medidas, o que torna os valores médios horários calculados altamente representativos, uma vez que cada uma das amostras horárias tem em média 84.000 (oitenta e quatro mil) medidas.

6.3 AVALIAÇÃO DAS MEDIÇÕES DE UMIDADE RELATIVA DO AR

Temperatura e umidade relativa do ar estão sempre relacionadas entre si e, como foi verificado para a temperatura, mesmo com a presença de inconsistências, os níveis médios da umidade relativa do ar são bastante adequados e contribuem para um bem estar e conforto dos indivíduos. Porém, nos meses de baixa precipitação pluviométrica (estiagem), os valores podem alcançar níveis alarmantes (menores que 30%), chegando em alguns casos à níveis desérticos (em torno de 10%), cujos valores provavelmente estão incorretos. Noutros casos é clara a inconsistência do dado, como por exemplo, umidades relativas nulas, indicando 0%, além de outros absurdos numéricos. A Tabela 6.3 a seguir, mostra um resumo de uma amostra de dados de umidade relativa do ar verificados em alguns postos sem nenhum tipo de tratamento aplicado às grandezas. O número de elementos da amostra completa é de 2.208.635 (dois milhões duzentos e oito mil seiscentos e trinta e cinco) medidas.

Tabela 6.3 - Dados de Umidade Relativa do Ar coletados pelo STH revelando várias inconsistências

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	19	100	71	18	26617
Bocaiúva	0	100	69	23	41211
Botelhos	0	100	68	19	26790
Buritzeiro	0	100	68	23	81871
Camargos	0	100	19	36	84254
Carmo de Minas	0	100	77	27	90963
Coa - Belo Horizonte	13	100	71	18	28503
Curvelo	8	100	63	20	25030
Diamantina	0	100	76	18	26543
Fazenda Campo Grande	0	100	18	34	57371
Fazenda Energetica Uberaba	0	100	66	22	27841
Felisburgo	26	100	80	16	25265
Gafanhoto	0	100	76	22	17796
Hotel Fazenda Curva do Rio	0	100	72	25	17894
Igarapé	29	99	77	19	2650
Igicatu	0	100	80	19	34091
Itambacuri	0	8288	81	54	25965
Januária	0	100	65	23	28162
Juiz de Fora	0	100	79	18	24017
Machado Mineiro	0	21248	81	95	52806
Nova Ponte	0	100	72	21	84632
Paracatu	13	100	66	20	22578
Passos	0	100	66	25	46295
Patos de Minas 1	0	100	74	21	57946
Peti	13	99	76	18	23410
Prata	9	100	63	23	27916
Quartel Geral	0	100	73	20	71774
Salinas	0	100	68	21	47123
São Pedro da Ponte Firme	14	100	67	20	24505
São Roque de Minas	0	100	72	24	85904
São Simão	0	100	69	21	90293
SE Alpinópolis	16	260	74	19	36136
SE Araçuaí	11	100	65	22	2545
SE Barreiro	0	100	68	22	55070
SE Curvelo	0	100	68	21	57154
SE Diamantina	0	100	78	20	51221
SE Governador Valadares	0	100	73	18	60274
SE Ituiutaba	0	100	71	25	36011
SE Januária	0	100	63	23	52201
SE Juiz de Fora	0	100	77	30	60211
SE Mesquita	0	100	78	16	60802
SE Montes Claros	0	100	40	38	25823
SE Ouro Preto	0	100	53	43	8819
SE Pouso Alegre 1	0	100	72	29	54462
SE Teófilo Otoni	0	100	80	19	34538

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
SE Uberaba	0	100	61	30	56079
Sede	0	100	63	23	13760
Sete Lagoas - EFAP	7	100	70	20	11344
Três Marias	0	100	29	36	90822
Turmalina	0	100	71	21	23218
Unaí	12	100	67	21	22215
Vargem Alegre	0	100	73	18	24158
Viçosa	0	100	64	35	43756

Resumindo a tabela de valores de umidade relativa do ar, mostrada anteriormente numa tabela menor (Tabela 6.4) com os valores extremos verificados para todo o estado de Minas Gerais, suas médias e desvios padrão, temos a seguinte representação:

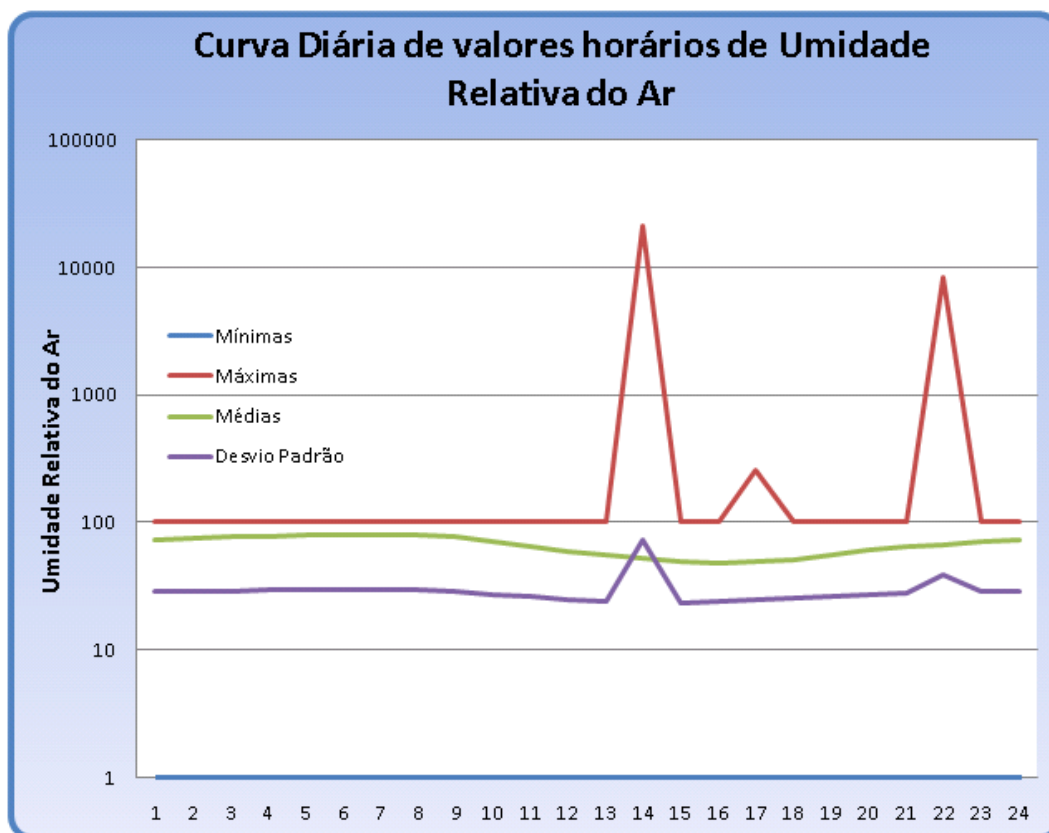
Tabela 6.4 – Resumo dos dados de umidade relativa do ar coletados pelo STH revelando inconsistências

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	0,0	99,0	18,0	16,0
Máximo	29,0	21248,0	81,2	94,5
Médio	3,6	656,5	67,5	25,1
Desvio Padrão	7,0	3065,0	13,3	12,0

Também nas medidas de umidade relativa do ar se verificam algumas inconsistências, e até mesmo alguns absurdos físicos, conforme destaca a tabela anterior, que podem ser atribuídos às causas já citadas anteriormente.

Mais uma vez foi feita uma estratificação dos dados mostrados nas tabelas anteriores, avaliando os mesmos mais detalhadamente no que diz respeito ao seu comportamento horário. Esta avaliação resultou no Gráfico 6.2, que ilustra a curva típica diária das medidas de umidade relativa do ar mínima, máxima e média, verificadas no estado de Minas Gerais, para cada uma das 24 horas do dia.

Gráfico 6.2 – Perfil do comportamento das medidas de umidade relativa do ar coletados pelo STH, revelando inconsistências



Similarmente ao gráfico elaborado para as medidas de temperatura, o gráfico das medidas de umidade relativa do ar faz perceber verdadeiros absurdos, principalmente nas medidas que se referem às umidades relativas máximas que obrigaram o uso de uma escala logarítmica no eixo das ordenadas. Também aqui, espera-se que a ferramenta de qualificação das medidas seja capaz de filtrar e identificar todas as inconsistências encontradas.

Mais uma vez pode-se perceber que, mesmo com as grandes inconsistências evidenciadas, a curva de umidade relativa média apresenta um perfil e uma seqüência de valores compatível com o que se esperaria para o estado de Minas Gerais, evidenciando novamente o fenômeno de regressão à média. Esta amostra é composta de 2.208.635 (dois milhões duzentos e oito mil seiscentos e trinta e cinco) medidas, tendo também cada uma das amostras horárias em média 84.000 (oitenta e quatro mil) medidas.

6.4 AVALIAÇÃO DAS MEDIÇÕES DE PRESSÃO ATMOSFÉRICA

Similarmente ao que foi feito para as medições de temperatura e umidade relativa do ar, as análises seguem agora com os resultados iniciais encontrados para as medições de pressão atmosférica (unidade hPa). A Tabela 6.5 a seguir apresenta os dados recebidos pelo sistema, e da mesma forma que nas análises anteriores, sem nenhum tipo de tratamento. O número de elementos da amostra completa é de 2.208.636 (dois milhões duzentos e oito mil seiscentos e trinta e seis) medidas.

Tabela 6.5 - Dados de Pressão Atmosférica coletados pelo STH revelando algumas inconsistências

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	969,5	996,7	984,1	4,0	26.617
Bocaiúva	852,7	959,1	910,6	10,9	41.211
Botelhos	889,4	911,0	901,6	3,1	26.790
Buritzeiro	0,0	967,1	952,8	13,1	81.871
Camargos	0,0	918,3	103,8	289,2	84.254
Carmo de Minas	0,0	962,1	914,0	6,3	90.961
Coa - Belo Horizonte	900,6	921,9	912,2	3,1	28.503
Curvelo	924,2	946,6	936,9	3,4	25.030
Diamantina	0,0	875,0	867,3	9,6	26.543
Fazenda Campo Grande	0,0	919,7	213,3	385,2	57.371
Fazenda Energetica Uberaba	916,8	939,3	930,2	3,1	27.841
Felisburgo	930,6	951,1	941,9	3,2	25.265
Gafanhoto	895,1	953,3	929,9	5,9	17.796
Hotel Fazenda Curva do Rio	839,4	958,2	940,5	24,0	17.894
Igarapé	921,5	936,7	928,2	2,7	2.650
Igicatu	0,0	944,0	919,0	12,5	34.091
Itambacuri	0,0	976,8	956,0	7,2	25.965
Januária	0,0	970,7	960,6	6,5	28.162
Juiz de Fora	0,0	942,5	929,3	9,6	24.017
Machado Mineiro	0,0	949,5	939,2	19,0	52.806
Nova Ponte	0,0	941,2	924,7	10,1	84.632
Paracatu	896,8	917,2	908,8	2,8	22.578
Passos	0,0	936,2	922,7	53,2	46.295
Patos de Minas 1	887,9	907,6	898,0	2,8	57.946
Peti	912,4	934,4	924,6	3,3	23.410
Prata	927,6	956,7	941,1	3,6	27.916
Quartel Geral	0,0	950,8	938,1	4,8	71.774
Salinas	0,0	984,2	944,1	26,9	47.123
São Pedro da Ponte Firme	862,7	936,3	906,1	6,7	24.505

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
São Roque de Minas	860,0	880,3	871,0	2,7	85.904
São Simão	0,0	988,7	961,7	38,4	90.293
SE Alpinópolis	328,8	924,3	915,3	4,3	36.136
SE Araçuaí	968,7	984,2	976,2	2,6	2.545
SE Barreiro	0,0	914,3	849,6	193,9	55.070
SE Curvelo	0,0	964,7	939,1	7,5	57.154
SE Diamantina	0,0	881,6	873,2	4,7	51.221
SE Governador Valadares	0,0	1.032,6	992,2	9,2	60.274
SE Ituiutaba	0,0	1.026,6	945,0	28,2	36.011
SE Januária	906,3	987,8	958,0	8,5	52.202
SE Juiz de Fora	0,0	953,2	935,8	62,2	60.211
SE Mesquita	963,3	990,0	977,0	3,9	60.803
SE Montes Claros	0,0	1.006,8	943,5	7,7	25.823
SE Ouro Preto	0,0	901,7	707,0	352,2	8.819
SE Pouso Alegre 1	0,0	930,5	918,7	10,2	54.462
SE Teófilo Otoni	0,0	982,6	971,8	9,8	34.538
SE Uberaba	0,0	934,5	910,9	112,1	56.080
Sede	0,0	939,9	863,3	194,7	13.760
Sete Lagoas - EFAP	880,0	946,7	918,8	8,5	11.344
Três Marias	0,0	1.045,6	403,1	468,3	90.822
Turmalina	0,0	949,6	939,4	7,0	23.218
Unaí	806,5	917,4	881,3	35,0	22.215
Vargem Alegre	0,0	995,0	982,3	7,4	24.158
Viçosa	-3.084,1	968,5	929,2	42,1	43.756

Percebe-se claramente um sem número de inconsistências e mais uma vez, resumindo a tabela anterior numa que represente Minas Gerais, ficam evidentes na Tabela 6.6, os valores extremos verificados, suas médias e desvios padrão.

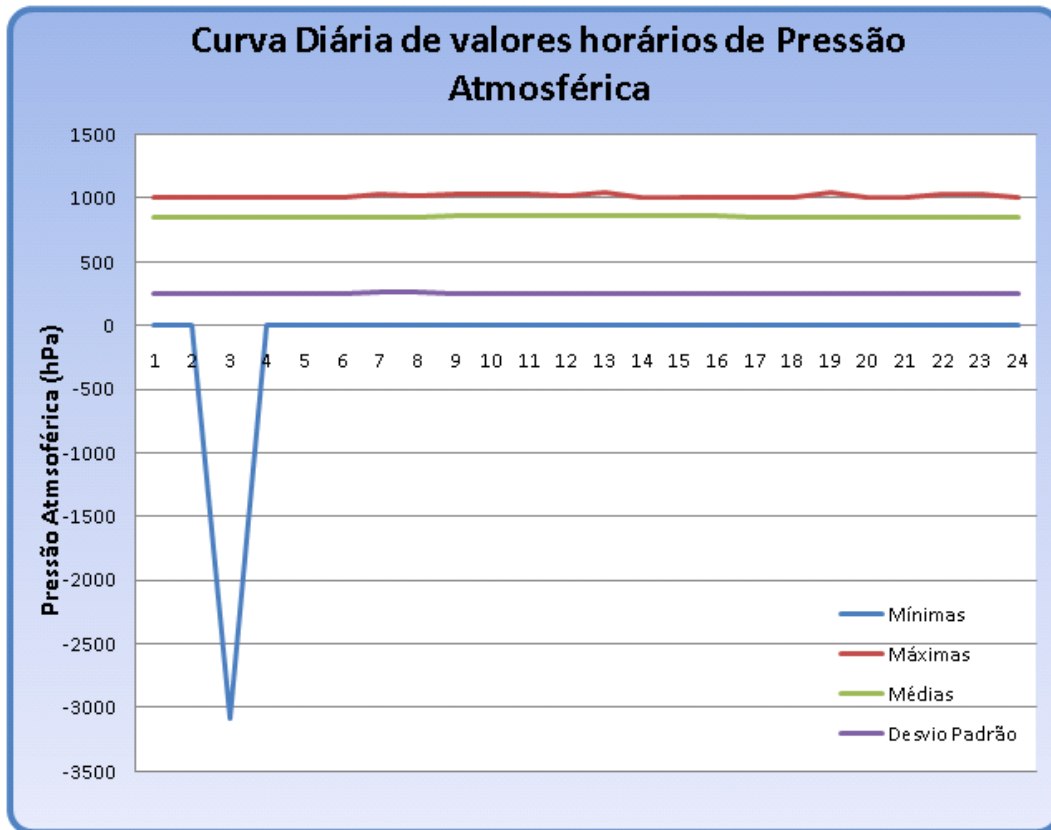
Tabela 6.6 – Resumo dos dados de pressão atmosférica coletados pelo STH revelando inconsistências

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	-3084,1	875,0	103,8	2,6
Máximo	969,5	1045,6	992,2	468,3
Médio	304,8	951,2	885,7	48,3
Desvio Padrão	640,8	36,1	166,5	102,4

Mais uma vez pode-se verificar a existência de verdadeiros absurdos físicos e, da mesma forma que anteriormente, segue o Gráfico 6.3 com a representação de uma curva típica diária, com os dados notáveis hora a hora de

pressão atmosférica, a fim de ilustrar os resultados encontrados e permitir sua comparação após o uso da metodologia de qualificação.

Gráfico 6.3 – Perfil do comportamento das medidas de pressão atmosférica coletados pelo STH revelando inconsistências



Vale lembrar novamente que o número de elementos da amostra completa é de 2.208.636 (dois milhões duzentos e oito mil seiscentos e trinta e seis) medidas, tendo cada uma das amostras horárias em média 84.000 (oitenta e quatro mil) medidas.

6.5 AVALIAÇÃO DAS MEDIÇÕES DE RADIAÇÃO SOLAR ACUMULADA

As mesmas análises anteriores foram feitas para as medidas de radiação solar acumulada, recebidas pelo sistema. Sendo assim, apresenta-se a seguir a Tabela 6.7 com os dados verificados sem que os mesmos tenham recebido

qualquer tratamento. O número de elementos da amostra é de 2.201.145 (dois milhões duzentos e um mil cento e quarenta e cinco) medidas.

Tabela 6.7 - Dados de Radiação Solar Acumulada coletados pelo STH revelando algumas inconsistências

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	0,0	100,0	9,1	9,6	26.540
Bocaiúva	0,0	100,0	9,8	9,4	41.060
Botelhos	0,0	32,1	9,4	8,7	26.702
Buritzeiro	0,0	51,9	10,0	9,1	81.809
Camargos	0,0	30,6	2,2	5,7	84.238
Carmo de Minas	0,0	49,0	9,2	8,7	90.710
Coa - Belo Horizonte	0,0	34,1	9,9	9,0	28.477
Curvelo	0,0	31,1	10,4	9,3	24.967
Diamantina	0,0	31,3	9,5	8,6	26.428
Fazenda Campo Grande	0,0	30,5	2,1	5,7	57.145
Fazenda Energetica Uberaba	0,0	31,8	9,5	8,7	27.810
Felisburgo	0,0	30,8	9,1	8,3	25.209
Gafanhoto	0,0	26,8	8,9	8,3	17.793
Hotel Fazenda Curva do Rio	0,0	41,2	9,7	8,8	17.894
Igarapé	0,0	31,2	10,0	9,6	2.650
Igicatu	0,0	30,9	8,8	8,4	34.082
Itambacuri	0,0	30,2	8,8	8,2	25.841
Januária	0,0	30,2	10,4	9,3	28.120
Juiz de Fora	0,0	31,0	8,8	8,2	23.914
Machado Mineiro	0,0	204,8	9,0	8,5	52.650
Nova Ponte	0,0	32,2	9,5	8,8	84.436
Paracatu	0,0	30,5	10,2	9,3	22.543
Passos	0,0	46,4	9,6	8,9	46.248
Patos de Minas 1	0,0	31,4	9,4	8,7	57.911
Peti	0,0	30,7	8,5	8,2	23.410
Prata	0,0	31,9	9,7	9,0	27.877
Quartel Geral	0,0	100,0	10,0	10,4	71.670
Salinas	0,0	96,5	9,5	8,7	47.001
São Pedro da Ponte Firme	0,0	31,5	9,9	8,8	24.475
São Roque de Minas	0,0	48,6	9,1	8,6	85.814
São Simão	0,0	37,3	9,3	9,0	88.183
SE Alpinópolis	0,0	100,0	9,0	8,6	36.106
SE Araçuaí	0,0	23,9	8,1	7,3	2.545
SE Barreiro	0,0	30,3	8,7	8,5	54.993
SE Curvelo	0,0	35,4	9,9	9,0	57.144
SE Diamantina	0,0	31,6	8,7	8,4	51.120
SE Governador Valadares	0,0	30,1	8,6	8,1	60.092
SE Ituiutaba	0,0	29,9	9,2	8,6	36.008
SE Januária	0,0	40,5	10,2	9,3	52.180
SE Juiz de Fora	0,0	30,4	7,9	7,8	59.988
SE Mesquita	0,0	31,0	8,3	8,0	60.624
SE Montes Claros	0,0	26,5	6,6	8,0	24.957
SE Ouro Preto	0,0	31,9	6,6	8,0	8.233

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
SE Pouso Alegre 1	0,0	31,8	8,9	8,6	54.359
SE Teófilo Otoni	0,0	29,9	8,6	8,3	34.526
SE Uberaba	0,0	29,8	8,7	8,3	56.024
Sede	0,0	27,7	7,8	7,6	13.759
Sete Lagoas - EFAP	0,0	27,3	8,6	7,9	11.344
Três Marias	0,0	39,1	4,2	7,8	90.738
Turmalina	0,0	31,2	10,0	8,7	23.012
Unaí	0,0	30,9	10,2	9,3	22.197
Vargem Alegre	0,0	30,4	9,0	8,2	24.055
Viçosa	-131,2	31,8	8,6	8,2	43.534

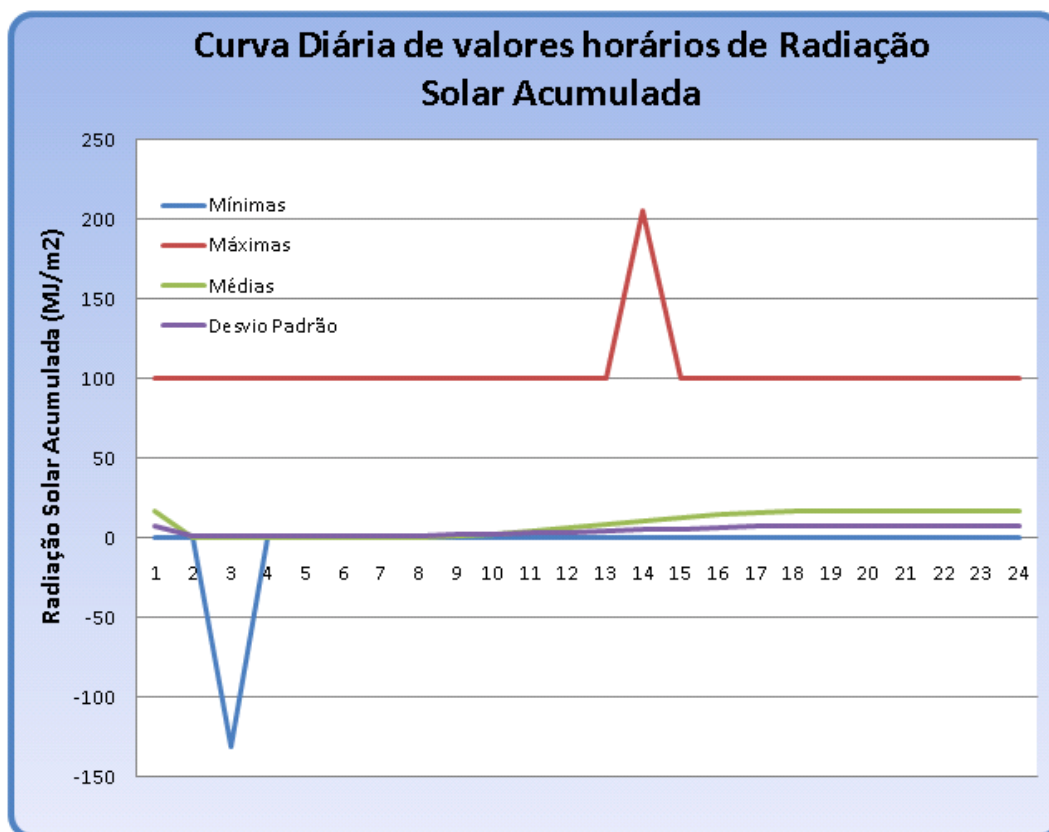
Resumindo-se a tabela anterior em uma mais macro que retrata Minas Gerais, a Tabela 6.8 ilustra, similarmente às outras medidas avaliadas, grandes inconsistências.

Tabela 6.8 – Resumo dos dados de radiação solar acumulada coletados pelo STH, revelando inconsistências

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	-131,2	23,9	2,1	5,7
Máximo	0,0	204,8	10,4	10,4
Médio	-2,5	42,4	8,8	8,5
Desvio Padrão	17,8	30,2	1,7	0,8

Da mesma forma, ilustrando-se a situação por meio do Gráfico 6.4 mostrado a seguir, percebem-se os absurdos encontrados, evidenciados pelos picos nas curvas de máximo e mínimo. Entretanto, a curva de valores médios de radiação solar acumulada, apesar de visualmente afetada pela escala, mostra uma tendência aparentemente normal para o comportamento da radiação solar, indicando zero para as primeiras horas do dia, passando então a crescer com o amanhecer e depois se estabilizando num valor máximo após o pôr do sol. Isso mais uma vez indica a influência do tamanho da amostra na avaliação de dados notáveis. Maiores detalhes poderão ser observados após o tratamento dos dados, quando então, espera-se remover todas as inconsistências verificadas.

Gráfico 6.4 – Perfil do comportamento das medidas de radiação solar acumulada coletados pelo STH revelando inconsistências



6.5.1 UM BUG DETECTADO

Durante o tratamento de amostras de dados de radiação solar acumulada foram verificadas algumas anomalias que ocorreram devido a um problema em alguns *firmwares* ou micro-códigos de *software* instalados na remota. A radiação solar que deveria ser acumulada até a meia-noite e “resetada” ou, levada a zero, neste horário, em alguns casos continuava acumulando após o pôr do sol, alcançando valores impraticáveis, como os mostrados na análise anterior. Este *bug* foi então corrigido pelo fornecedor do *hardware* que gerou uma nova versão do *firmware* a ser instalado nas unidades terminais remotas.

6.6 AVALIAÇÃO DAS MEDIÇÕES DE VELOCIDADE MÉDIA DO VENTO

As análises iniciais feitas nas medições de velocidade média do vento (m/s) revelaram um comportamento diferente das demais análises anteriores, se mostrando muito mais estáveis e com um comportamento que não apresenta grandes inconsistências que possam ser facilmente identificadas. O tamanho da amostra de dados de velocidade média do vento é de 2.212.993 (dois milhões duzentos e doze mil novecentos e trinta e três) medidas. A Tabela 6.9 a seguir mostra os resultados inicialmente encontrados.

Tabela 6.9 - Dados de Velocidade Média do Vento coletados pelo STH sem tratamento de valores

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	0,0	11,2	2,4	1,5	26.623
Bocaiúva	0,0	16,6	3,0	1,5	41.213
Botelhos	0,0	9,9	2,3	1,4	26.793
Buritzeiro	0,0	15,7	2,0	1,4	81.873
Camargos	0,0	13,7	0,7	1,5	84.255
Carmo de Minas	0,0	15,2	1,5	1,4	90.965
Coa - Belo Horizonte	0,0	10,9	2,4	1,2	28.504
Curvelo	0,0	8,6	2,1	1,1	25.032
Diamantina	0,0	12,8	3,4	1,5	26.546
Fazenda Campo Grande	0,0	10,0	0,5	1,2	57.376
Fazenda Energetica Uberaba	0,0	10,7	2,8	1,4	27.843
Felisburgo	0,0	11,0	2,9	1,6	25.270
Gafanhoto	0,0	9,8	1,4	1,1	17.847
Hotel Fazenda Curva do Rio	0,0	8,1	1,1	1,2	18.434
Igarapé	0,0	9,6	1,4	1,5	3.515
Igicatu	0,0	9,0	2,4	0,9	34.092
Itambacuri	0,0	9,9	1,9	1,5	25.968
Januária	0,0	7,5	1,7	1,1	28.166
Juiz de Fora	0,0	12,8	2,7	1,6	24.024
Machado Mineiro	0,0	7,0	1,5	1,3	52.807
Nova Ponte	0,0	14,5	1,8	1,7	84.634
Paracatu	0,0	9,6	3,1	1,4	22.579
Passos	0,0	11,5	2,5	1,6	46.294
Patos de Minas 1	0,0	12,2	2,7	1,9	57.949
Peti	0,0	15,2	1,7	1,3	23.684
Prata	0,0	12,0	3,0	1,6	27.918

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Quartel Geral	0,0	15,4	1,8	1,3	71.772
Salinas	0,0	9,6	2,1	1,4	47.129
São Pedro da Ponte Firme	0,0	10,0	3,0	1,3	24.509
São Roque de Minas	0,0	15,0	3,7	1,8	85.905
São Simão	0,0	22,4	2,0	1,9	90.299
SE Alpinópolis	0,0	13,0	2,5	1,5	36.174
SE Araçuaí	0,0	0,4	0,1	0,1	2.545
SE Barreiro	0,0	16,1	4,6	3,0	55.070
SE Curvelo	0,0	14,5	2,5	1,5	57.250
SE Diamantina	0,0	13,2	2,4	1,3	51.284
SE Governador Valadares	0,0	12,1	1,7	1,4	60.777
SE Ituiutaba	0,0	11,0	1,7	1,4	36.012
SE Januária	0,0	13,7	2,2	1,4	52.202
SE Juiz de Fora	0,0	9,8	1,1	1,2	60.897
SE Mesquita	0,0	9,3	1,2	1,2	61.101
SE Montes Claros	0,0	7,1	1,2	1,1	26.064
SE Ouro Preto	0,0	13,4	2,7	2,0	8.896
SE Pouso Alegre 1	0,0	13,6	1,5	1,4	54.463
SE Teófilo Otoni	0,0	10,9	1,6	1,2	34.539
SE Uberaba	0,0	10,5	2,0	1,3	56.101
Sede	0,0	6,2	1,1	0,8	14.268
Sete Lagoas - EFAP	0,0	10,1	1,9	1,3	11.344
Três Marias	0,0	21,0	1,1	1,6	90.828
Turmalina	0,0	8,4	2,1	1,2	23.223
Unaí	0,0	12,4	3,4	1,9	22.215
Vargem Alegre	0,0	8,5	1,6	1,1	24.162
Viçosa	0,0	9,1	2,3	1,4	43.760

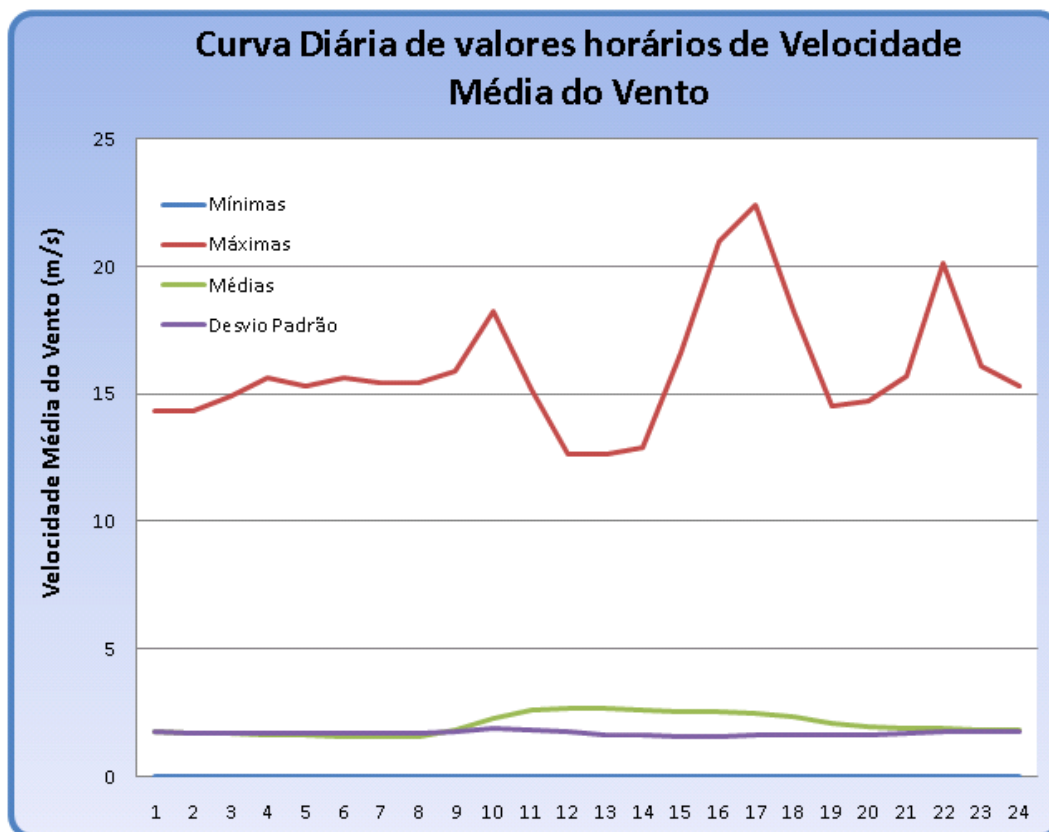
O resumo da tabela acima, apresentado na Tabela 6.10, mostra a grande consistência dos dados encontrados, diferente do que foi verificado para as outras grandezas, e não revela, pelo menos a princípio, nenhuma inconsistência.

Tabela 6.10 – Resumo dos dados de velocidade média do vento coletados pelo STH sem tratamento qualitativo dos valores

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	0,0	0,4	0,1	0,1
Máximo	0,0	22,4	4,6	3,0
Médio	0,0	11,5	2,1	1,4
Desvio Padrão	0,0	3,6	0,8	0,4

Isso pode ser visualmente verificado quando se traça o Gráfico 6.5 a seguir, que representa a curva típica diária da medida, fornecendo uma macrovisão do comportamento da velocidade do vento no estado de Minas Gerais.

Gráfico 6.5 – Perfil do comportamento das medidas de velocidade média do vento coletados pelo STH



6.7 AVALIAÇÃO DAS MEDIÇÕES DE DIREÇÃO DO VENTO

Da mesma forma que a análise anterior, as análises iniciais feitas nas medições de direção do vento também revelaram um comportamento diferente das demais, sendo muito mais estáveis e com um comportamento que não apresenta grandes inconsistências facilmente identificáveis. É provável que isso se deva ao projeto do sensor de direção do vento, que é integrado juntamente com o sensor de velocidade do vento, sendo ambos acoplados opticamente à unidade terminal remota. Isso torna o conjunto bastante robusto e imune a interferências externas. O tamanho da amostra de dados de direção do vento é de 2.208.549 (dois milhões duzentos e oito mil quinhentos e quarenta e nove) medidas. A Tabela 6.11 a seguir, mostra os resultados inicialmente encontrados,

onde o valor 0° (zero grau) representa a direção Norte, 90° a Leste, 180° a Sul e 270° a Oeste.

Tabela 6.11 - Dados de Direção do Vento coletados pelo STH sem tratamento de valores

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Alpercata	0	354	134	73	26.617
Bocaiúva	0	354	118	64	41.211
Botelhos	0	354	135	79	26.790
Buritzeiro	0	354	191	108	81.776
Camargos	0	354	34	75	84.254
Carmo de Minas	0	354	172	104	90.964
Coa - Belo Horizonte	0	354	118	76	28.503
Curvelo	0	354	132	73	25.030
Diamantina	0	354	130	75	26.543
Fazenda Campo Grande	0	354	37	80	57.371
Fazenda Energetica Uberaba	0	354	117	89	27.841
Felisburgo	0	354	133	53	25.265
Gafanhoto	0	354	176	97	17.796
Hotel Fazenda Curva do Rio	0	354	157	107	17.894
Igarapé	0	354	163	103	2.650
Igicatu	0	354	153	71	34.091
Itambacuri	0	354	112	70	25.965
Januária	0	354	196	82	28.162
Juiz de Fora	0	354	126	102	24.017
Machado Mineiro	0	354	115	55	52.806
Nova Ponte	0	354	182	112	84.632
Paracatu	0	354	130	79	22.578
Passos	0	354	159	76	46.295
Patos de Minas 1	0	354	119	90	57.946
Peti	0	354	164	89	23.410
Prata	0	354	130	90	27.916
Quartel Geral	0	354	146	94	71.774
Salinas	0	354	141	86	47.123
São Pedro da Ponte Firme	0	354	135	88	24.505
São Roque de Minas	0	354	123	88	85.904
São Simão	0	354	176	112	90.294
SE Alpinópolis	0	354	171	100	36.136
SE Araçuaí	0	354	110	102	2.545
SE Barreiro	0	354	126	73	55.070
SE Curvelo	0	354	147	77	57.154
SE Diamantina	0	354	157	108	51.222
SE Governador Valadares	0	354	141	114	60.275
SE Ituiutaba	0	354	138	122	36.011
SE Januária	0	354	151	120	52.202
SE Juiz de Fora	0	354	156	108	60.211
SE Mesquita	0	354	133	99	60.803
SE Montes Claros	0	354	181	98	25.823

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
SE Ouro Preto	0	354	206	137	8.819
SE Pouso Alegre 1	0	354	153	107	54.463
SE Teófilo Otoni	0	354	182	104	34.539
SE Uberaba	0	354	133	91	56.080
Sede	0	354	159	136	13.760
Sete Lagoas - EFAP	0	354	132	96	11.344
Três Marias	0	354	57	92	90.822
Turmalina	0	354	157	81	23.218
Unaí	0	354	116	81	22.215
Vargem Alegre	0	354	188	100	24.158
Viçosa	0	354	119	92	43.756

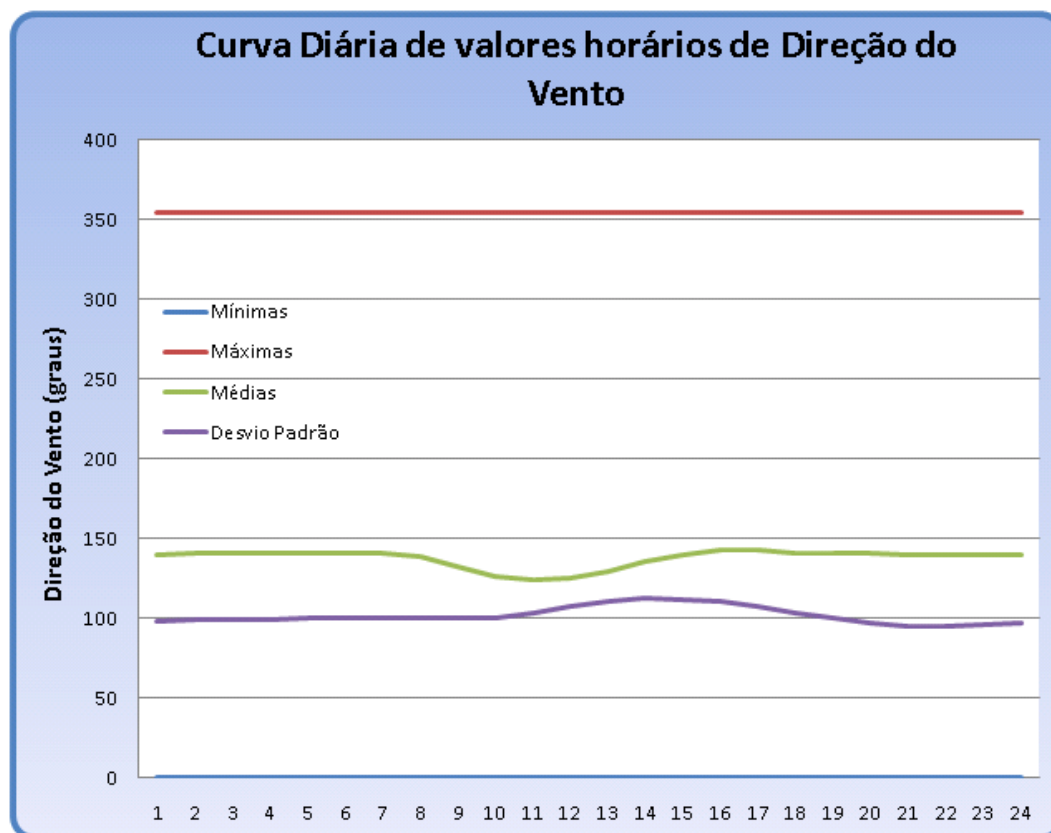
Resumindo a tabela anterior numa menor (Tabela 6.12) verifica-se a grande consistência dos dados encontrados, conforme comentado anteriormente, diferentemente das inconsistências que foram verificadas para as outras grandezas. Esse resumo dos dados de direção do vento, da mesma forma que o resumo dos dados de velocidade média do vento, também não revela pelo menos a princípio, nenhuma inconsistência.

Tabela 6.12 – Resumo dos dados de direção do Vento coletados pelo STH sem tratamento qualitativo de valores

	Dos Mínimos	Dos Máximos	Das Médias	Dos Desvios
Mínimo	0	354	34	53
Máximo	0	354	206	137
Médio	0	354	141	92
Desvio Padrão	0	0	34	18

O Gráfico 6.6 a seguir ilustra os dados de direção do vento.

Gráfico 6.6 – Perfil do comportamento das medidas de direção do vento coletados pelo STH



6.8 AVALIAÇÃO DAS MEDIÇÕES DE NÍVEL D'ÁGUA

Conforme descrito anteriormente, os sensores de nível d'água do tipo pressão não se deram bem em águas mineiras e tiveram que ser substituídos pelos do tipo bóia e contrapeso. Esses problemas com os sensores do tipo pressão foram facilmente percebidos durante os trabalhos de tratamento inicial dos dados, conforme ilustra a Tabela 6.13 a seguir. Foram avaliados 5.351.547 (cinco milhões trezentos e cinquenta e um mil quinhentos e quarenta e sete) medidas de nível de água de rios e de reservatórios.

Tabela 6.13 - Dados de nível d'água coletados pelo STH sem tratamento qualitativo de valores

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Abadia dos Dourados	0,00	25,16	1,43	0,87	49.942
Abaixo da Barra do Rio Verde	0,00	5,02	2,62	0,72	73.393
Aimorés	0,00	807,93	87,07	15,09	42.837
Aiuruoca	0,00	602,29	4,12	37,70	70.259
Anil	0,00	847,61	826,53	128,95	506
AVH 40	0,00	4,16	1,52	0,34	22.083
Bananal	0,00	598,00	517,88	200,99	93
Barra do Escuro	0,00	27,09	3,19	1,79	33.478
Bom Jardim de Minas	0,00	4,63	1,07	0,52	70.774
Bom Jesus do Galho	473,57	474,67	474,12	0,78	2
Caatinga	0,00	8,78	2,65	1,47	42.396
Cachoeira do Gambá	0,00	92,87	1,70	1,02	74.439
Cajuru	700,00	763,86	752,27	2,92	95.764
Calhauzinho	87,60	364,46	347,41	25,55	106
Camargos	13,92	998,94	908,84	5,12	82.457
Capim Branco 1	565,67	624,94	623,52	0,67	37.013
Capim Branco 2	0,00	654,91	564,73	7,49	27.652
Caraíbas	97,17	498,00	446,90	35,47	111
Carrapato	0,00	12,08	1,48	0,64	34.174
Cemitério dos Peixes	0,00	13,55	3,09	1,33	23.928
Dona Rita	534,27	536,52	536,11	0,57	13
Emborcação	600,00	661,02	644,39	12,00	79.286
Fazenda Boa Vista de Goiás	0,00	22,29	1,51	2,15	53.312
Fazenda Boa Vista de Minas	0,00	36,01	3,86	1,11	66.139
Fazenda Bonita de Baixo	0,00	30,46	1,96	2,42	40.494
Fazenda Cambauba	0,00	4,53	1,56	0,53	36.209
Fazenda Canadá	0,00	3,17	1,97	0,37	17.544
Fazenda Cigano	0,00	4,12	1,91	0,75	16.740
Fazenda das Vertentes	0,00	3,30	1,53	0,21	19.375
Fazenda Guariroba	0,00	15,70	2,26	0,72	30.403
Fazenda Jardim	0,00	8,10	1,02	0,82	1.508
Fazenda Laranjeiras	0,00	6,24	1,55	0,85	94.421
Fazenda Letreiro	0,00	2,75	0,97	0,37	30.327
Fazenda Maracuja	0,00	165,00	1,80	3,82	1.844
Fazenda Meloso	0,00	301,65	2,10	5,30	71.238
Fazenda Nova do Turvo	0,00	90,65	1,63	1,41	34.955
Fazenda Ouro Fino	0,00	301,10	0,78	5,21	70.456
Fazenda São Domingos	0,00	6,34	1,28	0,86	82.957
Fazenda Tamboril - SAAE	0,50	10,49	3,43	1,34	1.209
Funil	0,00	907,93	807,88	3,33	60.370
Gafanhoto	0,00	710,22	707,86	5,27	18.207
Grão Mogol	0,00	6,06	1,01	0,88	22.726
Guilman Amorim	16,00	739,61	494,78	2,75	54.794
Hotel Fazenda Curva do Rio	0,23	6,49	1,45	0,51	16.127
Ibituruna	0,00	19,66	2,01	1,42	96.277
Igarapava	495,20	515,42	512,06	0,11	79.696
Iguatama	0,00	15,68	3,59	2,61	92.405
Irapé	460,00	520,00	497,10	13,69	34.596

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Itumirim Jusante	0,00	27,19	1,91	1,32	97.381
Itutinga	0,00	885,96	885,51	11,89	60.775
Itutinga Jusante	0,00	5,72	2,44	0,81	35.431
Jacutinga	787,00	788,00	787,50	0,41	8
Jaguara	530,00	559,50	557,20	0,48	92.000
Joasal	595,30	893,00	694,77	171,68	3
Juiz de Fora Jusante	1,24	4,90	1,77	0,43	23.548
Luiz Carlos Barreto	622,00	622,00	622,00		1
Luiz Dias	895,00	895,00	895,00		1
Macaia	0,00	6,00	1,22	0,71	36.368
Machado Mineiro	0,00	693,20	684,36	18,80	82.625
Madre de Deus de Minas	0,00	20,93	1,28	1,15	92.263
Major Porto	0,00	91,10	1,74	3,02	43.241
Marmelos	646,00	646,00	646,00		1
Martins	702,00	709,28	708,63	0,61	1.815
Miranda	0,00	696,21	692,73	39,23	83.471
Mosquito	646,19	650,09	649,73	0,59	82
Nova Ponte	775,30	81.504,00	807,44	345,12	109.424
Paciência	505,00	505,00	505,00	0,00	2
Pai Joaquim	801,10	931,70	831,49	0,99	17.265
Pandeiros	494,00	495,00	494,50	0,71	2
Pandeiros Montante	0,00	1,57	0,45	0,15	19.077
Paraúna	640,00	640,00	640,00	0,00	2
Pari	0,85	8,00	1,50	0,70	1.980
Pari-Copasa	1,16	3,15	1,74	0,49	1.834
Patos de Minas	0,00	7,57	1,11	1,22	45.873
Peti	0,00	716,64	709,87	3,28	94.292
Piau	0,00	893,05	632,20	11,81	19.280
Picada	0,00	659,51	656,52	40,32	14.686
Piçarrão	807,00	809,90	808,45	2,05	2
Pirapora Ponte	0,00	15,69	2,57	0,89	85.245
Poço Fundo	1.155,00	1.156,26	1.156,13	0,06	1.520
Ponte Br 040	0,00	4,87	0,96	0,46	93.913
Ponte Br 146	0,00	6,46	3,22	0,76	76.679
Ponte BR 153	0,00	3,22	1,62	0,35	35.143
Ponte do Prata	0,00	23,36	0,82	0,59	70.673
Ponte do Rio Verdão	0,00	23,53	0,99	1,83	43.532
Ponte dos Vilelas	0,00	12,09	1,36	0,75	94.939
Ponte Goiás 206	0,00	18,61	1,95	1,14	52.607
Ponte Nova do Paraopeba	0,00	91,55	0,52	1,40	62.083
Ponte São Marcos	0,00	25,09	2,04	4,08	27.077
Porto Andorinhas	0,00	7,32	2,38	1,29	83.244
Porto Cavalo	0,00	10,85	3,59	1,86	37.810
Porto dos Pereiras	0,00	18,68	1,26	1,33	77.163
Porto Estrela	248,70	25.688,00	257,94	231,95	60.111
Porto Indaiá	0,00	18,17	2,85	0,80	87.055
Porto Mesquita	0,00	11,28	2,51	1,31	79.898
Porto Pará	0,00	6,04	1,68	0,89	82.035
Porto Tiradentes	0,00	98,10	2,09	8,51	96.572
Povoado dos Gomes	0,00	6,77	2,06	0,50	32.588
Queimado	521,90	836,70	824,95	3,45	49.739

LOCALIZAÇÃO	MÍNIMO	MÁXIMO	MÉDIO	DESVIO	AMOSTRA
Ribeirão Vermelho	0,00	25,62	0,62	1,26	87.347
Rio de Pedras	0,00	89.286,00	894,99	313,71	79.604
Rosal	2,78	555,00	287,60	293,55	6
Sá Carvalho	0,00	387,81	368,80	31,78	65.370
Salinas P	463,51	494,89	493,89	3,56	77
Salto do Paraopeba	832,00	832,00	832,00		1
Salto Grande	35,17	654,58	353,80	4,78	81.254
Salto Grande - SA	0,00	36.160,00	356,07	179,70	82.472
Salto Morais	500,00	509,64	508,69	0,40	2.148
Samambaia	0,00	746,99	737,30	81,92	83
Santa Clara	60,00	100,00	85,49	1,27	60.089
Santa Luzia	456,00	457,00	456,50	0,71	2
Santa Marta	799,00	863,39	863,06	3,41	698
Santo Hipólito	0,00	6,66	1,86	1,13	48.228
São Bernardo	1.554,00	1.555,38	1.555,17	0,37	19
São Sebastião da Encruzilhada	0,00	4,34	1,46	0,41	1.513
São Simão	0,00	401,44	397,88	4,22	91.834
Sobragi	437,00	437,20	437,09	0,07	32
Sumidouro	413,89	416,00	414,95	1,49	2
Três Marias	0,00	594,23	533,76	124,49	106.725
Três Marias Jusante	0,00	19,34	1,48	0,77	92.247
Tronqueiras	0,00	455,15	367,87	180,15	37
Tronqueiras Jusante	0,00	4,79	2,08	0,45	22.299
Tumiritinga	0,00	6,02	2,16	0,91	1.866
Várzea da Palma	0,00	8,73	2,28	1,10	84.681
Vila Terra Branca Jusante	0,18	6,87	0,76	0,75	16.207
Vila Urucuia	0,00	256,23	2,37	2,13	34.269
Volta Grande	474,91	494,79	494,08	0,46	85.038
Xicão	992,23	996,00	994,10	1,77	5

Os dados disponíveis foram confrontados com os limites mínimos e máximos conhecidos historicamente para cada posto e uma rápida observação da tabela anterior mostra a presença de inconsistências em praticamente todos os postos de medição de nível d'água. Espera-se que as medições inconsistentes sejam detectadas pela ferramenta de qualificação de dados proposta nesta dissertação.

6.9 CONCLUSÃO DO CAPÍTULO

O trabalho para tratamento inicial dos dados forneceu uma antecipação da visão do aspecto hidrológico e meteorológico do estado de Minas Gerais. Essa imagem inicial está contaminada de milhões de inconsistências nas medidas disponíveis e esse fato permitirá de forma clara e eficiente realizar

comparações e avaliações com os resultados obtidos após a utilização da ferramenta de qualificação de dados, proposta neste trabalho de dissertação. Essas comparações permitirão comprovar o desempenho e a eficiência da metodologia para o tratamento em tempo real, das medições recebidas pelo sistema de supervisão, proposta nessa dissertação.

Além disso, foram verificados diversos aspectos típicos da implantação de uma plataforma de coleta de dados do porte do STH, evidenciados pelo grande número de inconsistências encontradas durante as fases iniciais de implantação da plataforma de coleta de dados.

7 Qualificando Milhões de Medidas

7.1 INTRODUÇÃO

Neste capítulo serão abordados os trabalhos realizados a fim de se construir uma ferramenta de *software* capaz de realizar o tratamento de dados proposto pela metodologia de avaliação descrita anteriormente e serão apresentados os resultados encontrados após a aplicação da metodologia.

Ao longo do desenvolvimento deste capítulo será descrita também, a metodologia científica utilizada que permitiu a tomada de algumas decisões iniciais, necessárias para permitir a construção dos algoritmos de tratamento de dados.

7.2 CONDIÇÕES INICIAIS

No capítulo anterior foram descritas as condições iniciais encontradas no que diz respeito aos milhões de medições disponíveis nas bases de dados do STH. Pode-se perceber que há um número considerável de inconsistências que podem afetar consideravelmente os dados notáveis de cada uma das amostras.

Por outro lado, pode-se também observar que o número de elementos de cada uma das amostras é gigantesco (da ordem de milhões), o que faz com que os valores médios encontrados tenham grande significância, devido à característica de regressão à média “quase real” típica das grandes amostras de dados. Isso permite usar os valores médios como ponto de partida para os procedimentos de validação das grandezas.

Com relação aos dados verificados de valores mínimos e máximos, ficou explicitado no capítulo anterior que os mesmos estão extremamente contaminados com os mais diversos tipos de ruídos e por conta disso uma

análise especializada se fez necessária, a fim de se estabelecerem de forma adequada os limites físicos, mínimos e máximos de cada uma das grandezas avaliadas para que fosse possível a realização da filtragem dos valores discrepantes, conhecidos como “*outliers*”. Esta análise foi feita com o apoio dos hidrólogos e meteorologistas da Gerência de Planejamento Energético da Cemig.

7.3 EM BUSCA DE UMA “BOA” AMOSTRA INICIAL DE DADOS

Considerando-se as grandezas para as quais foi desenvolvida a ferramenta de qualificação proposta nesse trabalho, que são: precipitação pluviométrica (chuva), nível d'água de rios e de reservatórios de usinas, temperatura e umidade relativa do ar, velocidade e direção do vento, radiação solar e pressão atmosférica, o STH recebeu e armazenou até 01/01/2010 67.607.330 (sessenta e sete milhões seiscentos e sete mil trezentos e trinta) medidas. Esse número é muito maior que o total de medidas horárias usadas nas amostras descritas nos capítulos anteriores, por que grande parte dessas medidas é recebida pelo sistema, por conveniência de operação, a cada 15 minutos (dados meteorológicos) e uma outra parte é recebida a cada 5 minutos (dados hidrológicos). Destas medidas, derivaram uma infinidade de outras, como resultado de cálculos realizados pelos diversos módulos que compõem o sistema. Sendo assim, até aquela data o sistema possuía em sua base 126.532.505 (cento e vinte e seis milhões quinhentos e trinta e dois mil quinhentos e cinco) medidas armazenadas sem terem recebido nenhum tipo de tratamento que avaliasse a qualidade das mesmas.

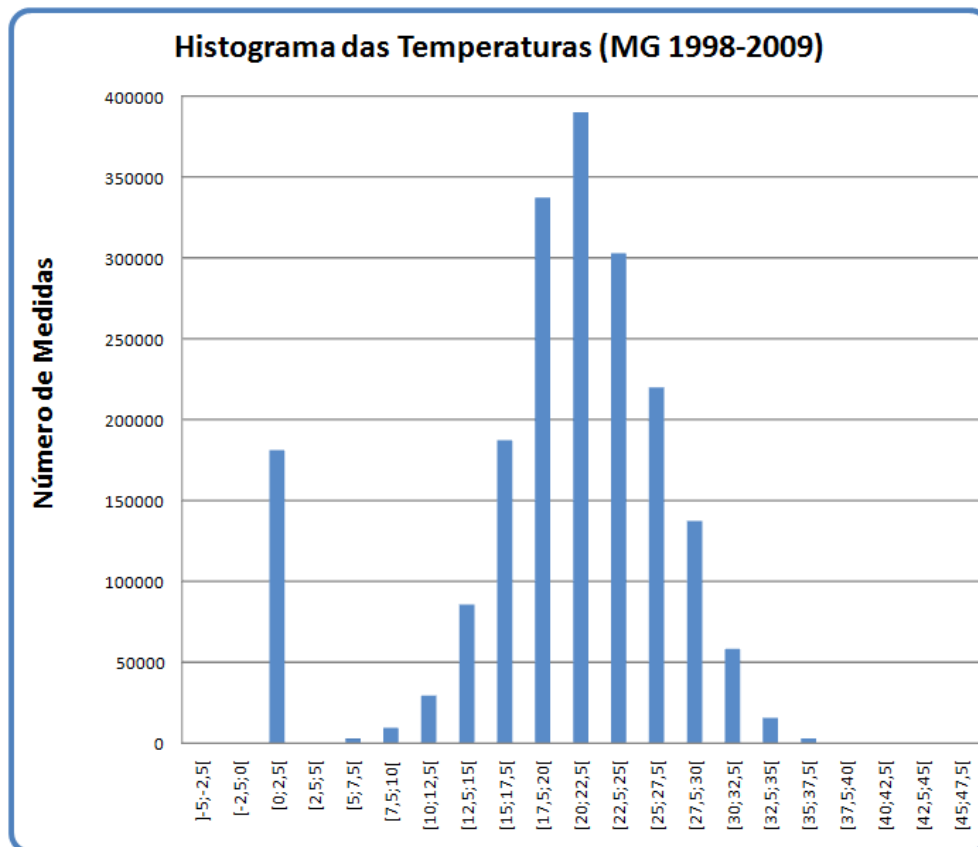
Como mostra o parágrafo anterior, o número de medidas a serem tratadas era extremamente elevado e sendo assim, um tratamento manual das grandezas se tornou impossível, tornando necessário construir pequenos aplicativos com os mais diversos fins onde um dos principais é aquele responsável pela filtragem inicial dos dados. O principal motivo para a necessidade destes aplicativos auxiliares no tratamento inicial dos dados é o grande número de medidas inconsistentes recebidas nas fases iniciais de implantação do sistema, quando ocorrem as instalações em campo das

unidades terminais remotas e seus diversos sensores, seguindo-se com as calibrações de todos os equipamentos e testes de aceitação.

Para garantir um tratamento estatístico adequado às medidas, estas inconsistências verificadas foram filtradas. Certamente essas inconsistências acabariam influenciando negativamente no tratamento estatístico dos dados, tais como médias e desvio padrão, fundamentalmente necessários à elaboração da ferramenta de qualificação proposta neste trabalho de dissertação.

Para se ter uma idéia da forte influência dos “*outliers*” verificados durante a análise exploratória de dados realizada sobre os dados históricos disponíveis nas bases de dados do STH, ilustra-se através da Figura 7.1, a seguir, um histograma representativo das probabilidades de ocorrência de valores de temperatura medidos pelos diversos postos de coleta de dados do sistema. Nele percebe-se facilmente uma quantidade exageradamente grande de valores de temperatura muito baixas: são 181.189 (cento e oitenta e um mil cento e oitenta e nove) medidas na faixa entre 0° e 2,5° Célcius.

Figura 7.1 – Histograma das temperaturas verificadas antes de qualquer tratamento



O comportamento da distribuição das medidas que se verificou para as temperaturas foi também verificado para todas as outras grandezas monitoradas pelo sistema de supervisão. Numa análise posterior, em busca da causa de tal comportamento, verificou-se que essas medidas eram aquelas correspondentes às feitas por um sensor com algum tipo de problema físico (danificado ou desconectado) ou elétrico (desligado). Nessa situação, o valor de medida enviado pelo coletor de dados era sempre igual a 0 (zero), o que causou a descontinuidade e o desequilíbrio na distribuição das probabilidades de ocorrência dos valores.

Conforme Rousseuw e Zomeren (1990), “*outliers are observations that do not follow the pattern of the majority of the data*”, ou seja, “*outliers*” são valores observados que não seguem o padrão de comportamento da maioria dos dados. Segue então um questionamento: O que fazer com eles? Trata-se de uma questão controversa e as opiniões a respeito não são unânimes. Uns defendem a sua rejeição e outros rezam que isso não deve ser feito nunca. Ainda, um “*outlier*” é um evento esporádico, considerado obra do acaso, da aleatoriedade, causados em sistemas de supervisão por um erro de medição ou um erro de transmissão.

Dada a descrição no parágrafo anterior e considerando a quantidade de ocorrências e o fato de serem fruto de um defeito ou falha permanente de um sensor do sistema, a grande quantidade de dados “fora da curva” verificados não podem ser considerados nas análises “*outliers de fato*”, devendo ser então rejeitados.

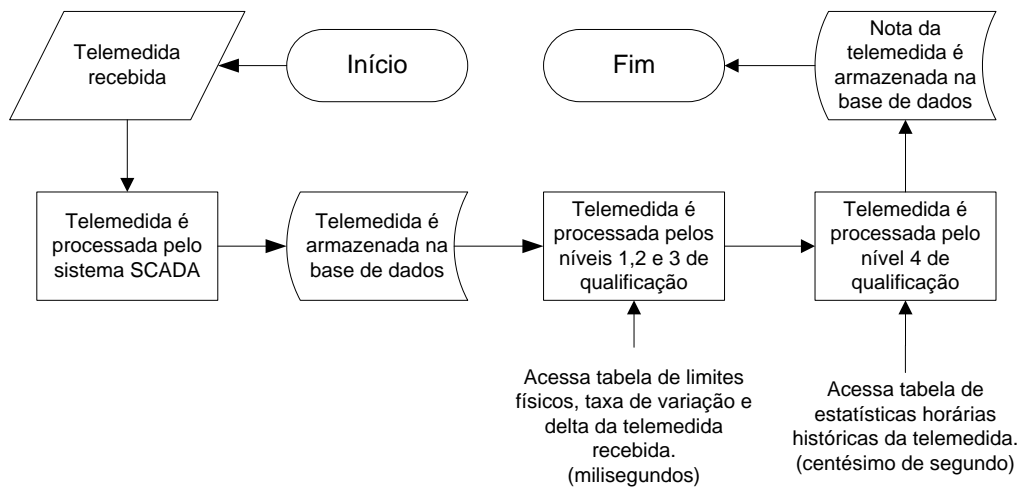
Esta filtragem inicial, foi realizada pelas fases 1, 2 e 3 da metodologia de tratamento, descritas no Capítulo 4. Esses filtros, realizados por aplicativos computacionais, se basearam no tratamento de limites físicos e no tratamento de violações de taxas de variação das grandezas (dv/dt). Estes valores limites e taxas de variação foram conseguidos através de entrevistas e avaliações junto aos hidrólogos e meteorologistas da Cemig. Essa filtragem, baseada em limites e taxas de variação pré-estabelecidas, foi então executada sobre as séries históricas disponíveis no sistema, totalizando 67.607.330 de medidas avaliadas. Durante esta fase, verificou-se que o tempo médio necessário para confrontar a medida em análise contra seus limites era da ordem de alguns milissegundos, indicando que o impacto desse método sobre o sistema de supervisão não seria significativo, quando do processamento da medida em tempo-real.

Com a remoção dos “*outliers*”, realizada pelos filtros das fases 1, 2 e 3 de tratamento das medidas, foi possível se obter com um razoável nível de qualidade as médias, medianas, desvios padrão, além das taxas de variação (dv/dt) e valores máximos e mínimos considerados normais, sob o aspecto físico, para cada um dos tipos de grandeza. Essa filtragem possibilitou a criação de um modelo inicial no banco de dados relacional, onde foram armazenados estes dados notáveis iniciais, de frequência horária, que posteriormente foram usados em processos de confrontação de valores recebidos pelo sistema durante a fase 4 de tratamento das grandezas, também descrita no Capítulo 4. Durante esta última fase de tratamento, verificou-se que o tempo médio necessário para confrontar a medida em análise contra o seu comportamento histórico era da ordem de centésimos de segundo, também indicando que o impacto sobre o sistema de supervisão não seria significativo, quando do processamento em tempo-real.

Esse modelo de dados mais consistentes, inicialmente criado a partir da aplicação das filtragens descritas nos dois parágrafos anteriores, permitiu então a criação de um processo computacional e de um modelo relacional no banco de dados que, incorporando todas as fases de tratamento de dados descritas no Capítulo 4 e propostas pela metodologia, é capaz de realizar a pontuação de uma telemedição recebida em tempo real, sem causar impacto no desempenho das outras ferramentas e processos do sistema de supervisão. Verificou-se que o tempo médio para a qualificação de uma grandeza é de 125 milissegundos. Essa capacidade de qualificação de dados em tempo real forneceu ao sistema uma considerável elevação no nível de confiança em relação aos dados recebidos.

Um resumo da seqüência de tratamento aplicada às telemedições é ilustrado conforme mostra o fluxograma apresentado na Figura 7.2 a seguir:

Figura 7.2 – Fluxograma básico do processo de qualificação de uma telemedição



Após a criação e integração ao STH dos processos responsáveis pela qualificação em tempo-real das telemedições recebidas, decidiu-se pela qualificação de todos os dados armazenados na base de dados histórica do sistema, ou seja, decidiu-se qualificar os 67.607.330 (sessenta e sete milhões seiscentos e sete mil trezentos e trinta) de medidas, disponíveis até 01/01/2010. Considerando-se um tempo médio de 0,125 segundos para cada medida, o tempo gasto para essa tarefa foi de 97 dias de processamento ininterrupto, realizado durante o primeiro semestre de 2010, entre os meses de março e junho. Paralelamente durante todo esse período de processamento, o STH permaneceu em produção, já com o sistema de qualificação em tempo-real em funcionamento, sem ter seu desempenho afetado de forma significativa.

Esse processamento dos dados disponíveis na base histórica permitiu a realização de alguns estudos, cuja síntese é apresentada na seção a seguir.

7.4 SÍNTESE QUANTITATIVA DA QUALIFICAÇÃO DOS DADOS

Após os trabalhos iniciais de tratamento dos dados para a eliminação das inconsistências evidentes, foram desenvolvidos os algoritmos descritos no Capítulo 5 baseados na metodologia de qualificação de dados proposta e descrita no Capítulo 4 e então o aplicativo de qualificação *off-line* foi executado para efetuar a verificação de toda a base de dados histórica disponível desde 1998 até o final do ano de 2009, o que totalizou 67.607.330 (sessenta e sete

milhões seiscientos e sete mil trezentos e trinta) medidas. Os resultados obtidos são mostrados nas Tabela 7.1 a seguir.

Tabela 7.1 – Resultado geral da qualificação das medidas históricas disponíveis

	Temperatura	Umidade	Pressão	Radiação	DVento	VVento	Nível d'água	Chuva
Nota 10	69,68%	70,35%	45,56%	60,33%	82,86%	89,51%	51,05%	99,18%
Nota 9	12,27%	11,49%	26,59%	14,80%	2,26%	0,02%	30,32%	0,00%
Nota 8	6,63%	5,78%	11,59%	7,73%	3,82%	0,09%	3,68%	0,00%
Nota 7	0,99%	0,94%	1,23%	1,40%	0,42%	0,03%	1,12%	0,00%
Nota 6	0,71%	0,89%	1,13%	1,96%	0,28%	0,04%	1,88%	0,00%
Nota 5	0,16%	0,24%	0,31%	0,80%	0,07%	0,02%	0,74%	0,00%
Nota 4	0,12%	0,22%	0,13%	0,82%	0,03%	0,03%	0,86%	0,00%
Nota 3	0,18%	0,20%	0,44%	1,83%	0,01%	0,04%	0,92%	0,00%
Nota 2	0,09%	0,05%	1,63%	0,06%	0,00%	0,00%	0,14%	0,00%
Nota 1	0,00%	0,70%	0,20%	0,05%	0,00%	0,00%	0,58%	0,00%
Nota 0	9,16%	9,15%	11,18%	10,23%	10,25%	10,23%	8,70%	0,81%
Soma:	100%	100%	100%	100%	100%	100%	100%	100%

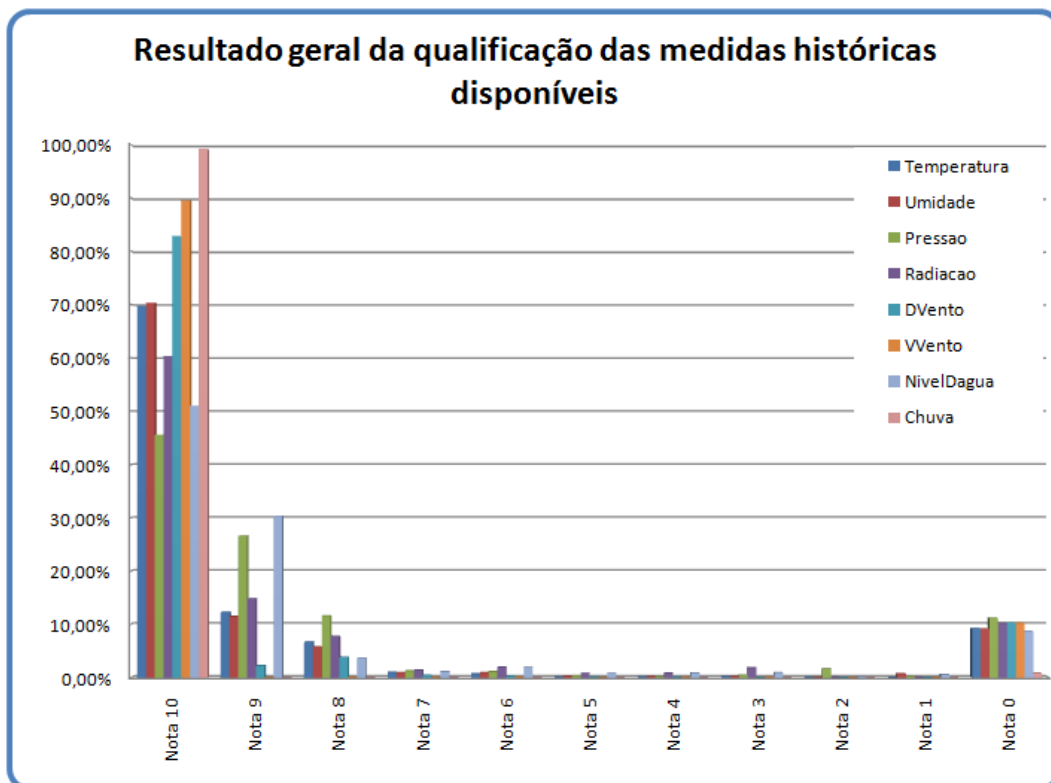
Vale lembrar aqui como é a estratificação das notas que uma telemedida pode receber:

- Nota 10 se está dentro de uma faixa determinada pela sua média histórica mais ou menos um “delta” inicialmente arbitrado pelos meteorologistas e hidrólogos.
- Nota 9 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para aquela dia do ano e hora específicos.
- Nota 8 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para uma “janela” de tempo de 15 dias, tendo como ponto central aquela data e hora.
- Nota 7 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para uma “janela” de tempo de 30 dias, tendo como ponto central aquela data e hora.
- Nota 6 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para uma “janela” de tempo de 90 dias, tendo como ponto central aquela data e hora.
- Nota 5 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para uma “janela” de tempo de 180 dias, tendo como ponto central aquela data e hora.
- Nota 4 se está dentro da faixa limitada pelos seus limites históricos máximo e mínimo para uma “janela” de tempo igual a um ano completo.

- Nota 3 se, apesar de não violar seu limite físico, seu valor supera o máximo ou é inferior ao mínimo histórico verificado até então, se constituindo num novo máximo *maximorum* ou mínimo *minimorum*.
- Nota 2 se a medida viola a sua taxa de variação admissível, estipulada pelos hidrólogos e meteorologistas.
- Nota 1 se a medida viola seus limites físicos, também estimados pelos hidrólogos e meteorologistas.
- Nota 0 se a medida já é recebida com algum *flag*, informado pelo algoritmo interno da remota coletora de dados, indicando alguma falha no processo de sensoriamento.

Uma representação gráfica da Tabela 7.1 é mostrada a seguir através do Gráfico 7.1.

Gráfico 7.1 – Resultado geral da qualificação das medidas históricas disponíveis



7.5 ANÁLISE QUALITATIVA DOS RESULTADOS OBTIDOS

A situação ilustrada pelo Gráfico 7.1, da seção anterior, mostra que sob o ponto de vista de qualidade da telemedição, o STH vem se comportando muito bem. Apesar disso, um número elevado de medições classificadas com a nota 0 (zero) e um número menor de medidas com notas 1 (um) e 2 (dois) mereceu atenção por parte das equipes responsáveis pela manutenção dos sistemas de medição.

A maior intensidade na ocorrência de notas tão baixas foi verificada no segundo semestre de 1998, o que é justificado por este ser o período quando ocorreu o maior volume de implantações de estruturas para coleta de dados hidrometeorológicos nos mais diversos pontos de Minas Gerais e alguns de Goiás. Nesta fase inicial de implantação do sistema, muitos testes foram realizados para a validação de todas as funcionalidades das partes integrantes do projeto e, nestes testes, muitas medidas fictícias foram enviadas para o sistema central de processamento de dados.

Posteriormente à fase inicial de implantação da plataforma de coleta de dados, ocorreu uma estabilização na quantidade de medidas criticadas como inadequadas para utilização pelos outros processos que compõem o STH.

Apesar disso, ocorreram alguns picos isolados de indicações de falhas na medição de nível d'água quando da ocorrência do problema, já comentado anteriormente, associado aos medidores de nível baseados na pressão da coluna de água. Com a substituição dos sensores do tipo pressão pelos sensores do tipo bóia e contrapeso os problemas de medição de nível foram solucionados e as falhas constantes facilmente identificadas pelo processo de qualificação de dados foram reduzidas ao mínimo.

De maneira geral, o sistema de medição tem se comportado muito bem, tendo em vista que praticamente 90% das telemedições obtiveram nota superior a sete, índice considerado adequado para os propósitos do sistema.

8 **Resumo iconográfico e resultados encontrados**

8.1 INTRODUÇÃO

A proposta deste capítulo é apresentar de forma visual, através de gráficos e tabelas, alguns dos resultados encontrados após aplicação da metodologia de qualificação às séries históricas de medidas disponíveis na base de dados do STH. Será feita a apresentação dos resultados encontrados por tipo de grandeza, a fim de permitir a comparação dos seus perfis de comportamento, após o tratamento realizado pelo processo de qualificação, com aqueles sem o tratamento dos dados mostrados no Capítulo 6 dessa dissertação.

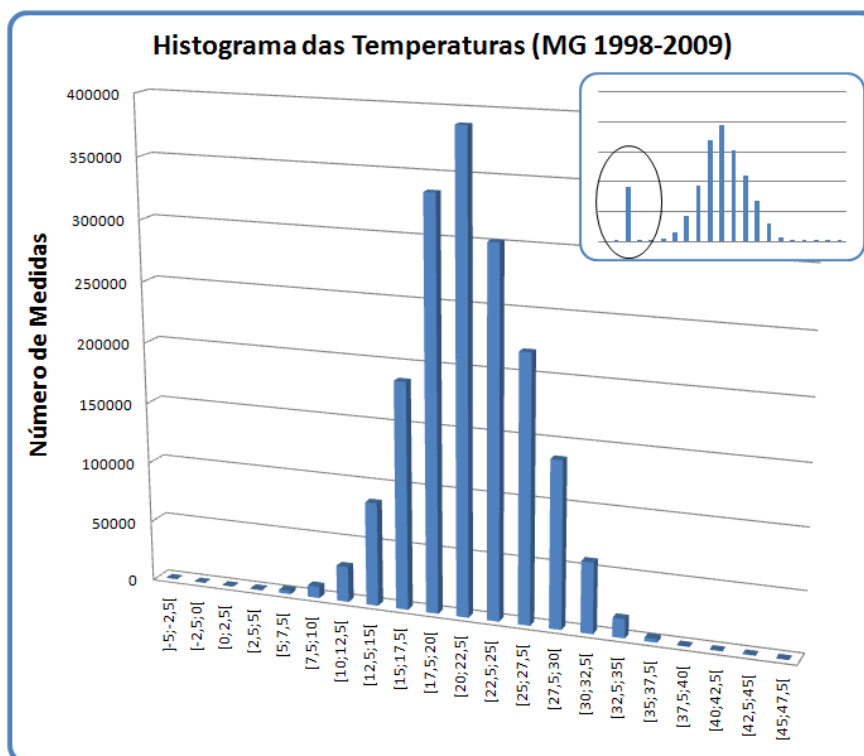
8.2 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE TEMPERATURA DO AR

Verificaram-se para as medições de temperatura do ar, resultados muito bons sob o aspecto qualitativo; mais de 88% das medidas de temperatura receberam uma nota maior ou igual a 8 (oito), indicando uma alta repetibilidade dos valores no entorno de uma média verificada historicamente. Poucas foram as ocorrências de notas entre 1 (um) e 7 (sete), ou seja, não se verificou um grande número de medidas violando os limites físicos inicialmente estabelecidos, caso em que receberiam nota 1 (um), como também não se verificou um grande número de medidas violando suas taxas de variação horária consideradas normais. Nos casos em que se verificou a ocorrência da nota 2 (dois), constatou-se que, quando não se tratava de alguma manutenção sendo executada na unidade terminal remota, a causa da violação da taxa de variação era decorrente da ocorrência de chuva de granizo sobre o posto de medição, o que ocasiona uma redução abrupta da temperatura do ar.

O Gráfico 8.1 apresenta um histograma dos valores de temperatura recebidos pelo sistema de supervisão por todos os postos de coleta de dados disponíveis, que receberam notas iguais ou superiores a 3 (três), desta forma eliminando da amostra de dados aqueles que violaram seus limites físicos, suas taxas de variação admitidas normais e aqueles que foram recebidos com alguma indicação de má qualidade registrada pela unidade coletora de dados. Esses dados extirpados da amostra deformam o perfil do histograma, como pode ser visto no detalhe anexado à figura (canto superior direito), e com certeza impactam de forma significativa no cálculo dos valores notáveis das temperaturas registradas.

Sendo assim, este gráfico, de certa forma, representa o histograma de temperaturas registradas no estado de Minas Gerais de 1998 até 2009. Pode-se verificar a semelhança do perfil do histograma com o perfil da curva normal, o que corrobora na confirmação da adequação da metodologia estatística para o tratamento dos dados recebidos.

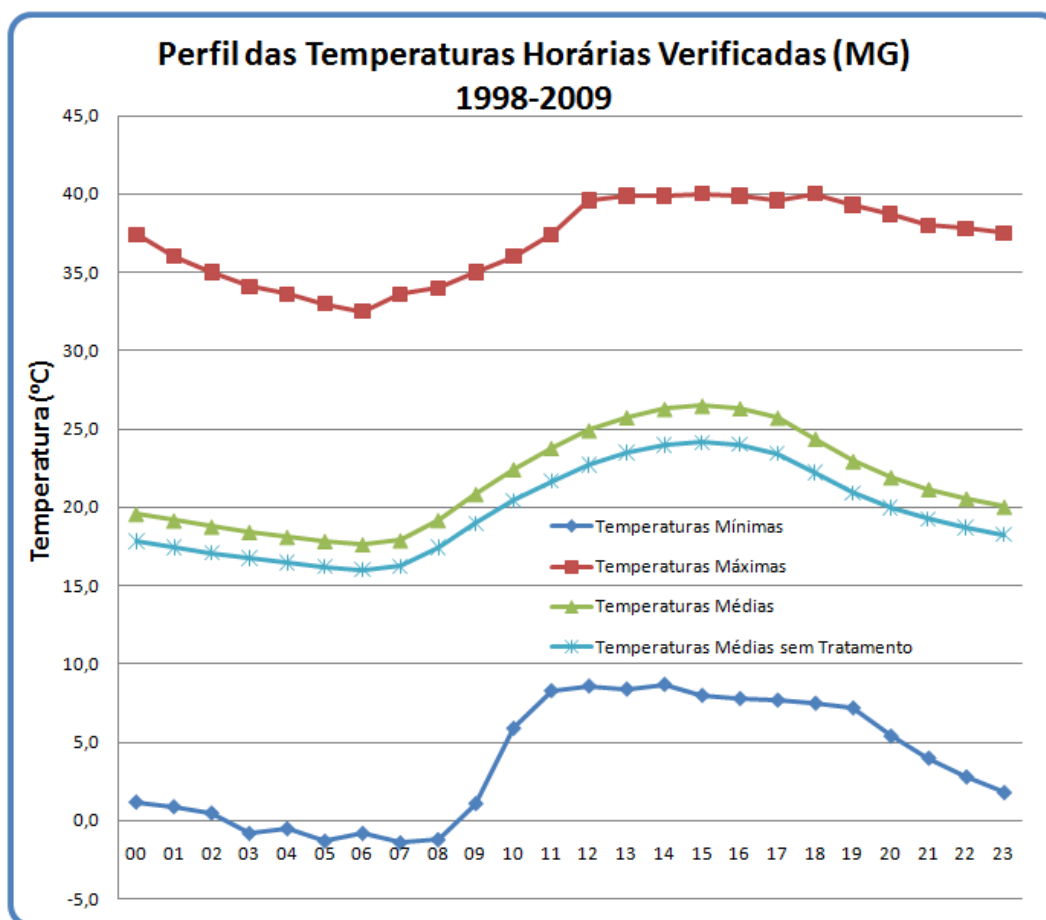
Gráfico 8.1 – Histograma das temperaturas já tratadas pela metodologia



Deixando de lado o histograma e indo em busca de um perfil do comportamento dos valores horários de temperatura recebidos, foi realizada uma

pesquisa que, conforme feito para a criação do histograma, leva em conta todas as medidas recebidas por todos os postos de medição. Com isso foi possível traçar o perfil horário representativo das temperaturas no estado de Minas Gerais no período de 1998 a 2009. O Gráfico 8.2 ilustra esse perfil. Para fins de comparação, vale lembrar que na seção 6.2 (vide Gráfico 6.1), durante a análise de “*outliers*” dos valores de temperatura, foram observadas medidas reconhecidamente absurdas para o estado de Minas Gerais, com pico máximo de 100° Célsius, média das máximas horárias, para quaisquer das 24 horas do dia, acima de 37° Célsius e mínimas de 0° Célsius. Conforme mostra o Gráfico 8.2, esses valores inconsistentes foram devidamente pontuados pelo processo de qualificação, com notas muito baixas, e dessa forma puderam ser facilmente filtrados a fim de não serem considerados para análises qualitativas. Ainda vale ressaltar que a presença de um grande número de “*outliers*”, principalmente aqueles cujos valores se encontram na faixa de temperaturas que vai de 0° a 2,5° graus Célsius, impactam significativamente no cálculo da temperatura média, seja para Minas Gerais como um todo, como mostrado no Gráfico 8.2, seja para uma região em estudo ou para a localização onde se encontra o posto de medição.

Gráfico 8.2 – Perfil histórico horário das temperaturas após tratamento das medidas pela metodologia de qualificação de dados



Com a imensa base de dados devidamente tratados e, considerando-se que todos os postos são localizados no sistema, não apenas por seus nomes mas principalmente por sua posição geográfica (latitude e longitude), podem ser feitos estudos estratificados por regiões. Como exemplo, foram escolhidos três postos de medição que se localizam dentro da RMBH (Região Metropolitana de Belo Horizonte), a saber: Edifício Sede da Cemig, no bairro Santo Agostinho, região centro-sul da capital; subestação Barreiro ao sul da RMBH e o COA no extremo oeste da região no centro de uma grande área verde. Esses três postos de medição estão identificados na Figura 8.1, formando entre eles um triângulo de aproximadamente 54km² de área. Realizando-se o estudo do perfil das temperaturas verificadas para os três postos identificados na figura, obtém-se o resultado mostrado no Gráfico 8.3 e na Tabela 8.1.

Figura 8.1 – Área de 54km² delimitada por 3 postos de medição do STH

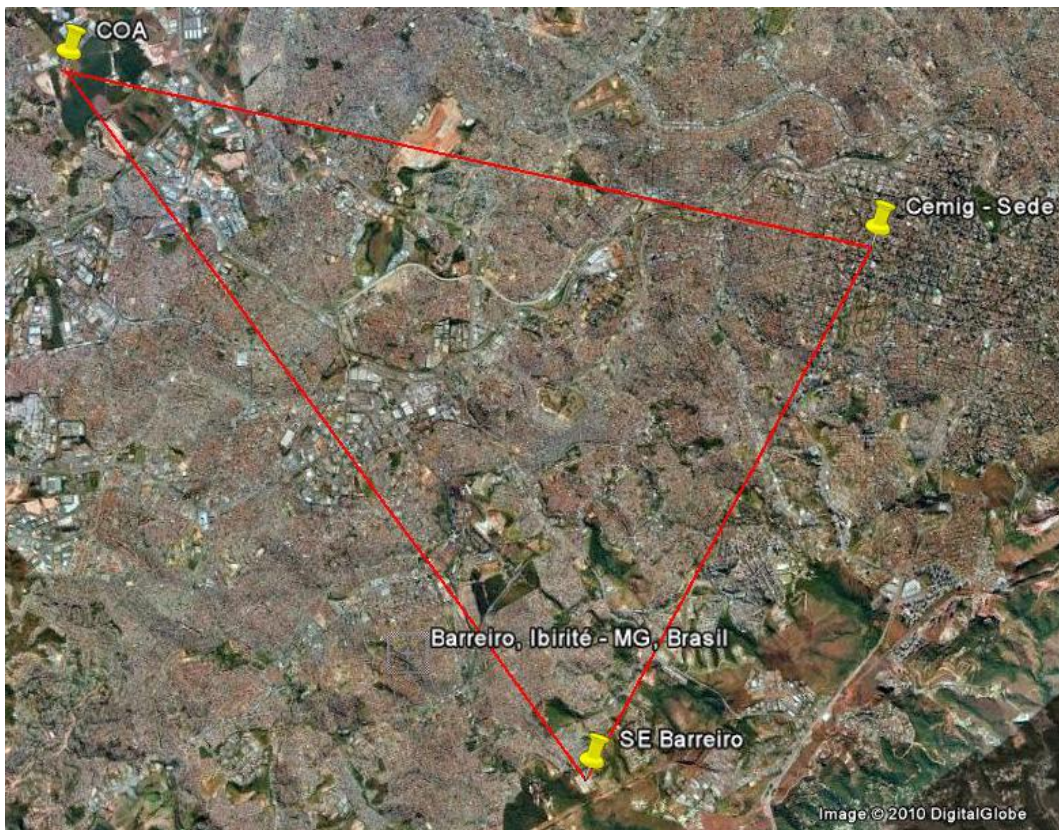


Gráfico 8.3 – Perfil de temperatura delimitada por 3 postos de medição do STH

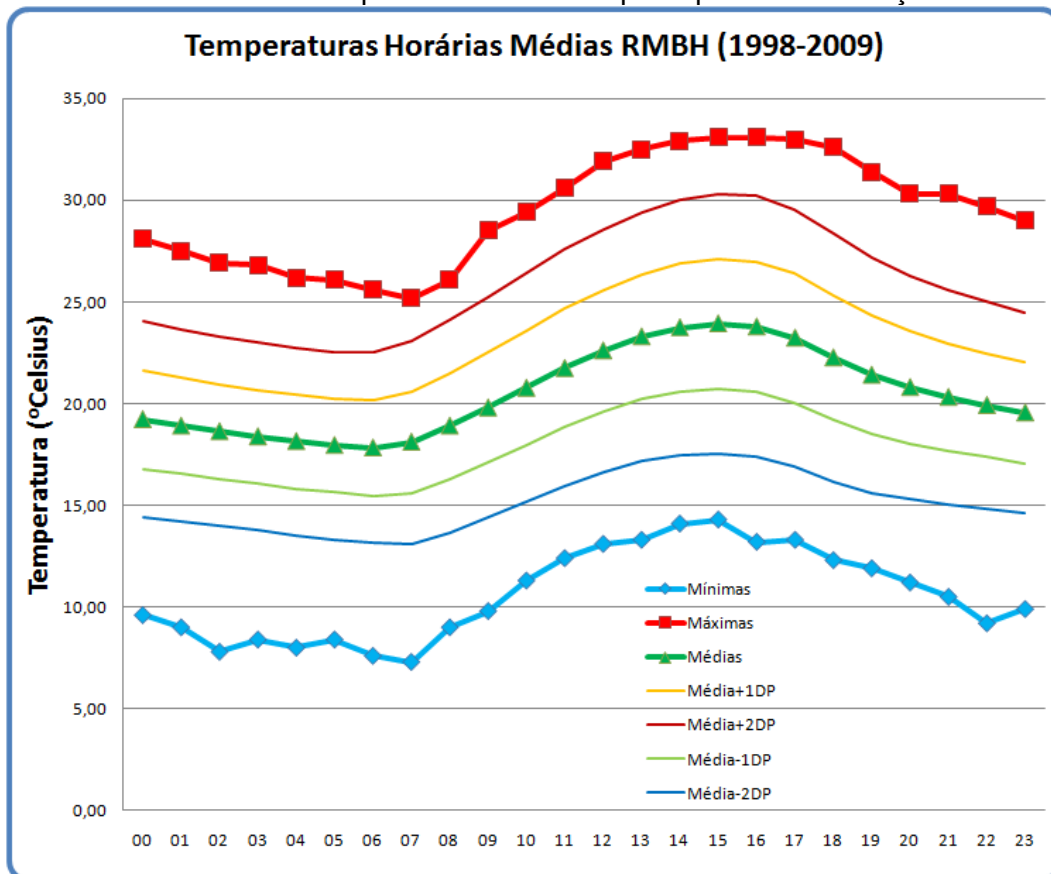


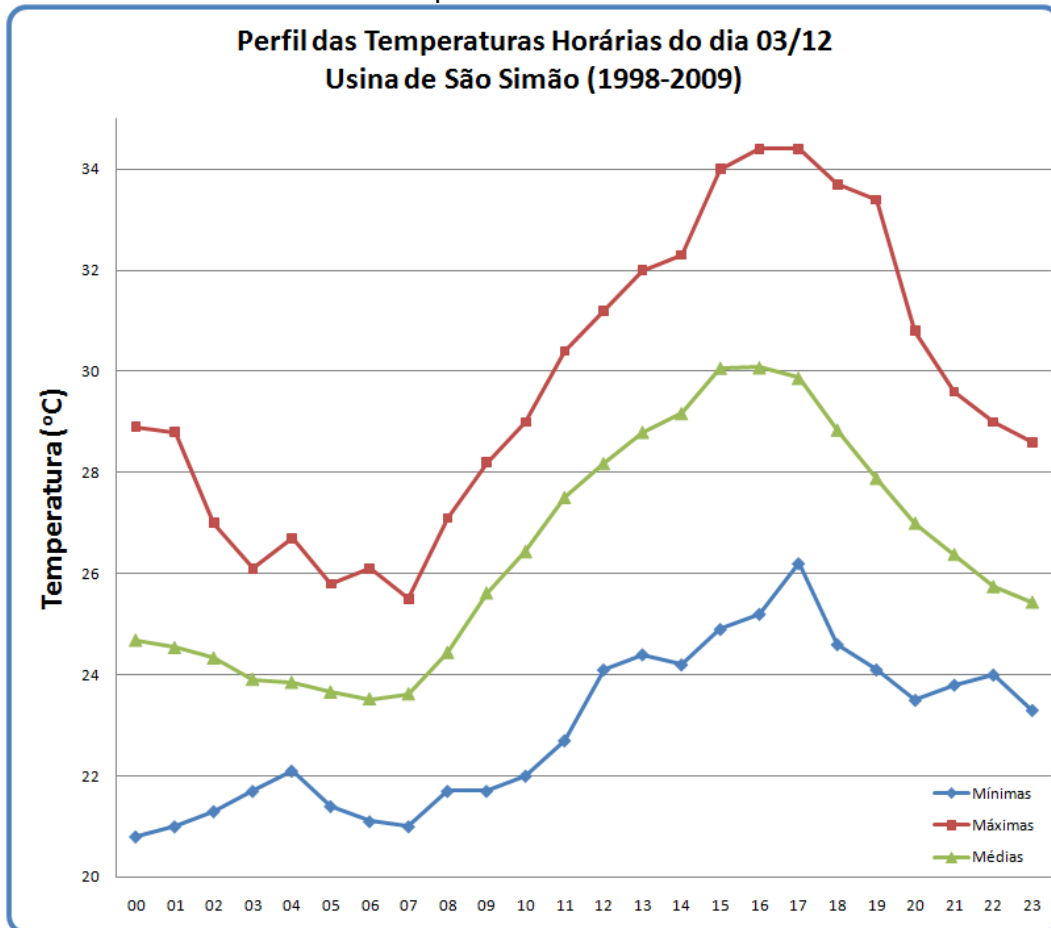
Tabela 8.1 – Temperaturas verificadas por 3 postos do STH na RMBH

Hora	Mínimas	Máximas	Médias	Desvio Padrão	Amostra
00	9,60	28,10	19,22	2,42	3889
01	9,00	27,50	18,91	2,36	3894
02	7,80	26,90	18,63	2,33	3877
03	8,40	26,80	18,38	2,31	3865
04	8,00	26,20	18,14	2,31	3870
05	8,40	26,10	17,95	2,31	3846
06	7,60	25,60	17,83	2,34	3844
07	7,30	25,20	18,12	2,50	3845
08	9,00	26,10	18,91	2,61	3895
09	9,80	28,50	19,80	2,71	3922
10	11,30	29,40	20,77	2,81	3932
11	12,40	30,60	21,74	2,91	3934
12	13,10	31,90	22,61	2,99	3936
13	13,30	32,50	23,29	3,07	3928
14	14,10	32,90	23,74	3,14	3936
15	14,30	33,10	23,93	3,19	3940
16	13,20	33,10	23,78	3,21	3944
17	13,30	33,00	23,23	3,16	3936
18	12,30	32,60	22,26	3,06	3945
19	11,90	31,40	21,41	2,90	3932
20	11,20	30,30	20,80	2,75	3936
21	10,50	30,30	20,32	2,65	3929
22	9,20	29,70	19,92	2,55	3925
23	9,90	29,00	19,55	2,47	3921

As pesquisas dos perfis de comportamento das temperaturas horárias realizadas até aqui levaram em conta todo o histórico de dados existente na base de dados do sistema para o posto ou postos em questão, e traçaram o perfil levando em conta todos os dias do ano como período de pesquisa. A fim de se demonstrarem as outras diversas possibilidades, o Gráfico 8.4 mostra, como exemplo, o comportamento das temperaturas horárias registradas e tratadas pelo algoritmo de qualificação na Usina de São Simão especificamente para o dia 3 de dezembro. Esta usina está localizada no Rio Paranaíba, na divisa entre os estados de Minas Gerais e Goiás. Percebe-se que as pesquisas podem ir mais além, e chegar a um nível de estratificação horário, que é uma das condições de contorno estabelecidas para o desenvolvimento dessa proposta de qualificação. Sendo assim, as possibilidades de estratificação de

dados são muitas, permitindo ao cliente das informações, pesquisa baseadas em perfis que vão do comportamento anual de uma grandeza até o comportamento para uma hora específica, passando pelos perfis de comportamento mensais, semanais e diários ou por períodos estabelecidos arbitrariamente.

Gráfico 8.4 – Perfil histórico diário das temperaturas na Usina de São Simão verificado para o dia 3 de dezembro



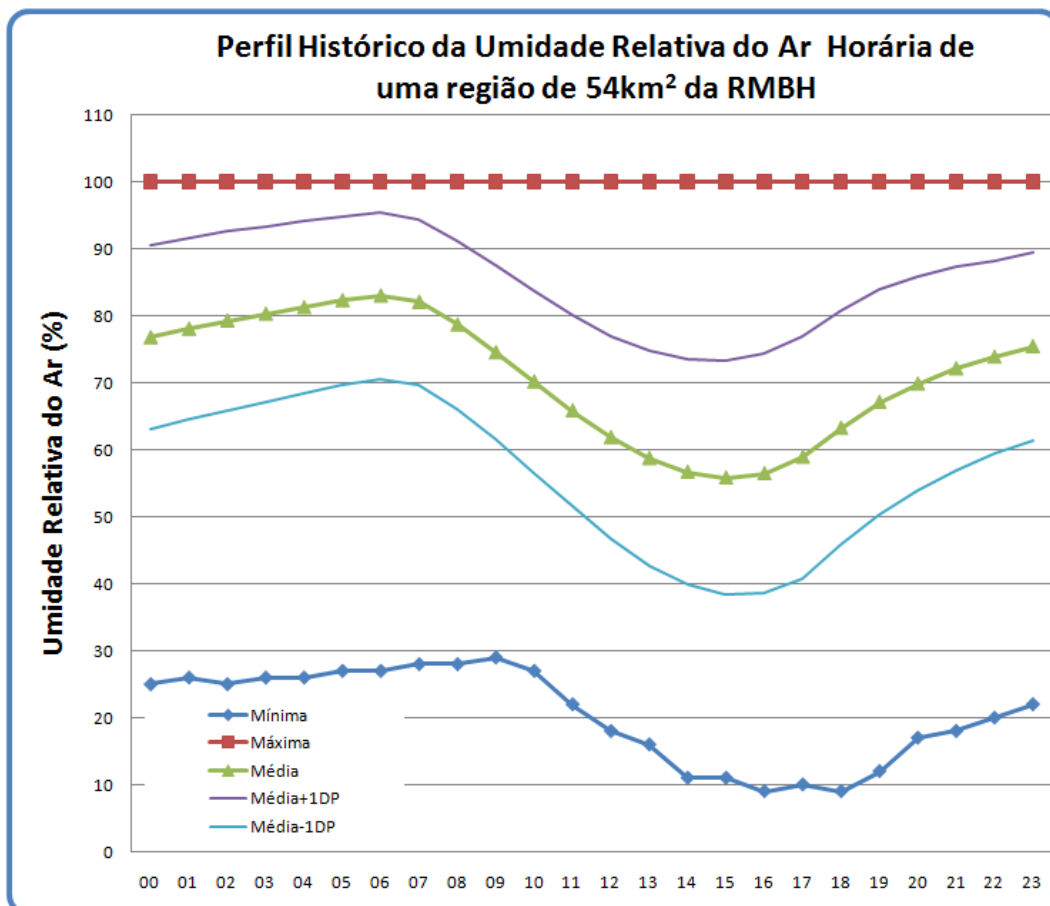
8.3 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE UMIDADE RELATIVA DO AR

Sob o aspecto qualitativo das medições de umidade relativa do ar recebidas pelo STH, as mesmas repetiram quase que literalmente o comportamento verificado para as medidas temperatura do ar, com mais de 85% das medidas recebidas recebendo notas iguais ou superiores a 8 (oito). Serão apresentadas nessa sessão alguns perfis de comportamento dos valores de umidade relativa do ar. Vale lembrar que todas as formas de pesquisa

realizadas, na sessão anterior, para os valores de temperatura podem ser aqui repetidas. Entretanto para mostrar outras possibilidades, alguns perfis apresentados terão uma estratificação diferente dos anteriores.

Tendo como referência a área triangular de 54km² mostrada na sessão anterior (Figura 8.1), e realizando-se uma pesquisa do comportamento dos valores históricos horários de umidade relativa do ar, a partir da amostra com todos os dados recebidos pelos três postos de coleta localizados nos vértices do triângulo, obtém-se o Gráfico 8.5 mostrado a seguir. Percebe-se no gráfico uma coerência de valores bem diferente da incoerência registrada no gráfico da sessão 6.3 que, onde tão absurdos eram os valores, fez-se necessário utilizar uma escala logarítmica. Vale ressaltar que, mesmo se estabelecendo um filtro para a seleção de medidas com notas maiores ou iguais a três, obteve-se amostras horárias com aproximadamente 4.000 elementos.

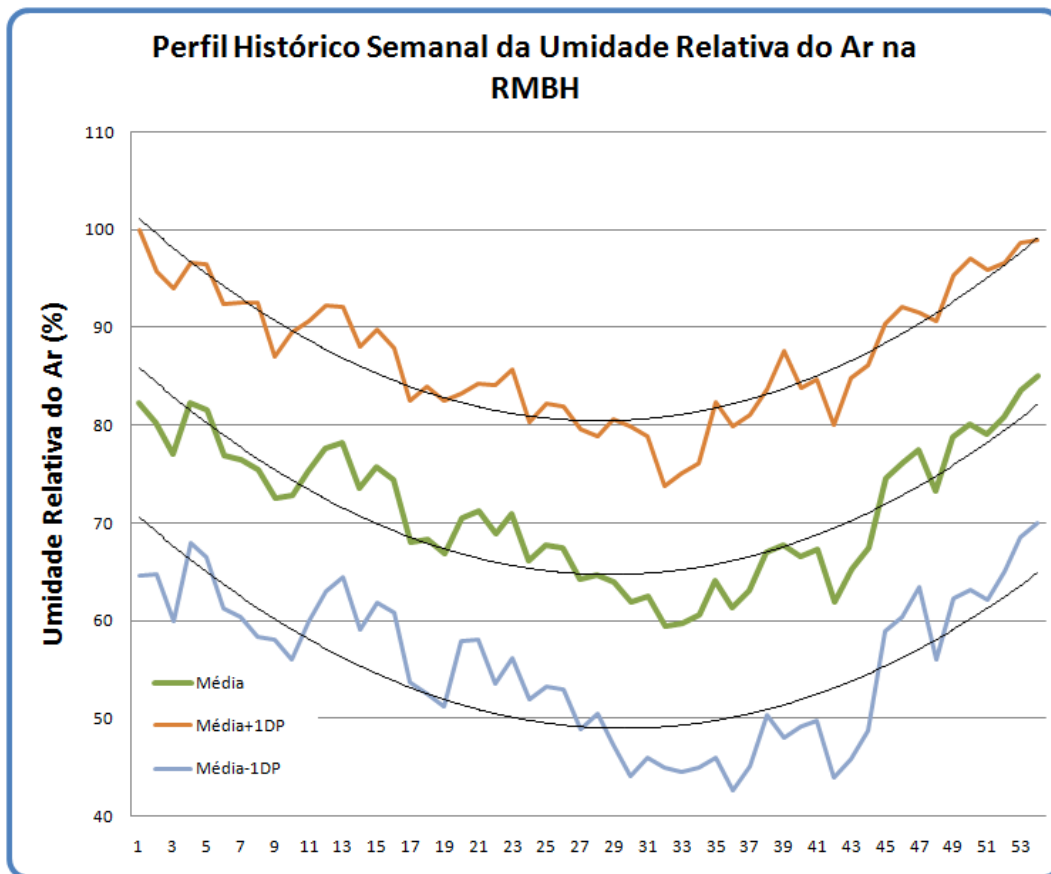
Gráfico 8.5 – Perfil histórico da Umidade Relativa do Ar Horária na RMBH



Pode-se perceber no Gráfico 8.5 um afastamento em relação à média, bem mais acentuado para os valores mínimos. Essa característica chamou a atenção dos meteorologistas que pretendem confirmar, através de testes de calibração, a suspeita de imprecisão do sensor de umidade para valores inferiores a 30% de umidade relativa do ar.

Uma pesquisa do comportamento dos valores históricos médios semanais de umidade relativa do ar, para a mesma região considerada anteriormente gerou como resultado o Gráfico 8.6 mostrado a seguir onde foram inseridas linhas de tendência de 2ª ordem apenas para harmonizar as curvas apresentadas, a saber: média histórica semanal dos valores de umidade relativa do ar (curva interior), média histórica semanal dos valores de umidade relativa do ar somadas a 1 (um) desvio padrão (curva superior) e média histórica semanal dos valores de umidade relativa do ar subtraídas de 1 (um) desvio padrão (DP).

Gráfico 8.6 – Perfil histórico anual da umidade relativa do ar estratificado por semana na RMBH



8.4 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE VELOCIDADE MÉDIA DO VENTO

Como mostraram as análises iniciais dos dados de velocidade média do vento, descritas no Capítulo 6 (sessão 6.6) e ocorridas antes do processamento qualitativo dos valores pela metodologia aqui proposta, os valores recebidos se mostraram bastante consistentes e aparentemente sem nenhuma anomalia. Esse fato foi confirmado pelo mecanismo de qualificação, quando foram processados os milhões de valores relativos à velocidade média do vento e verificou que praticamente 90% deles receberam nota 10, ou seja, se encontravam numa faixa de valores delimitada pela sua média histórica mais ou menos um delta definido pelos meteorologistas.

Os Gráficos 8.7, 8.8 e 8.9 ilustram o perfil de comportamento da velocidade média do vento estratificados numa base horária, numa base semanal e numa base mensal com amostras respectivas contendo em média 7.775, 1.775 e 3.921 elementos.

Gráfico 8.7 – Perfil histórico diário da velocidade média do vento na RMBH

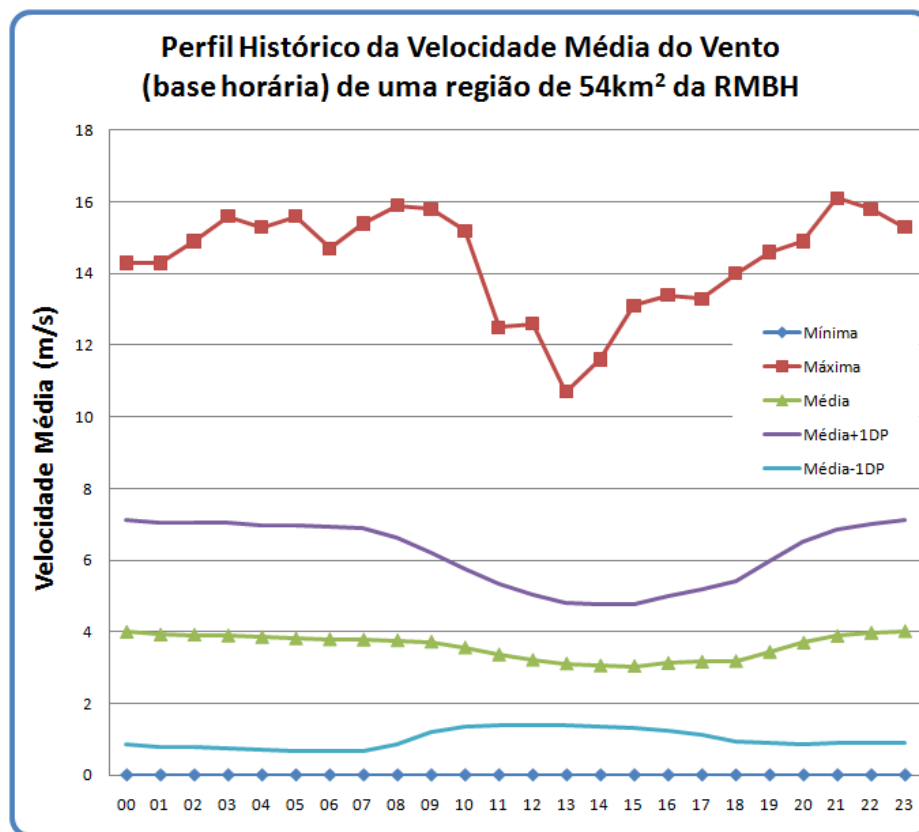


Gráfico 8.8 – Perfil histórico anual da velocidade média do vento na RMBH estratificado por semana

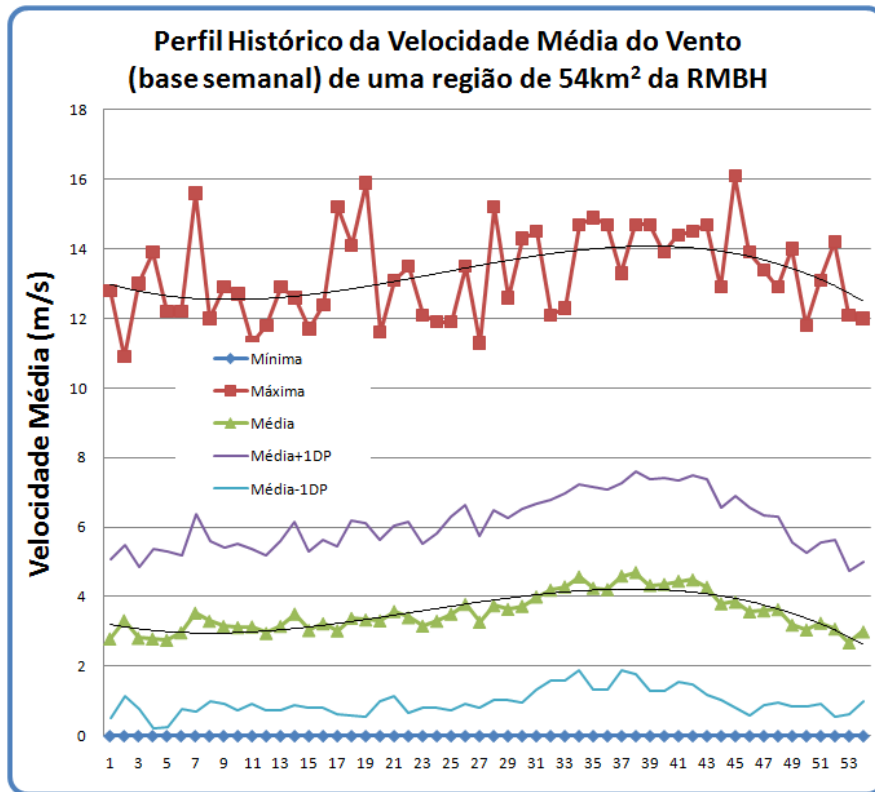
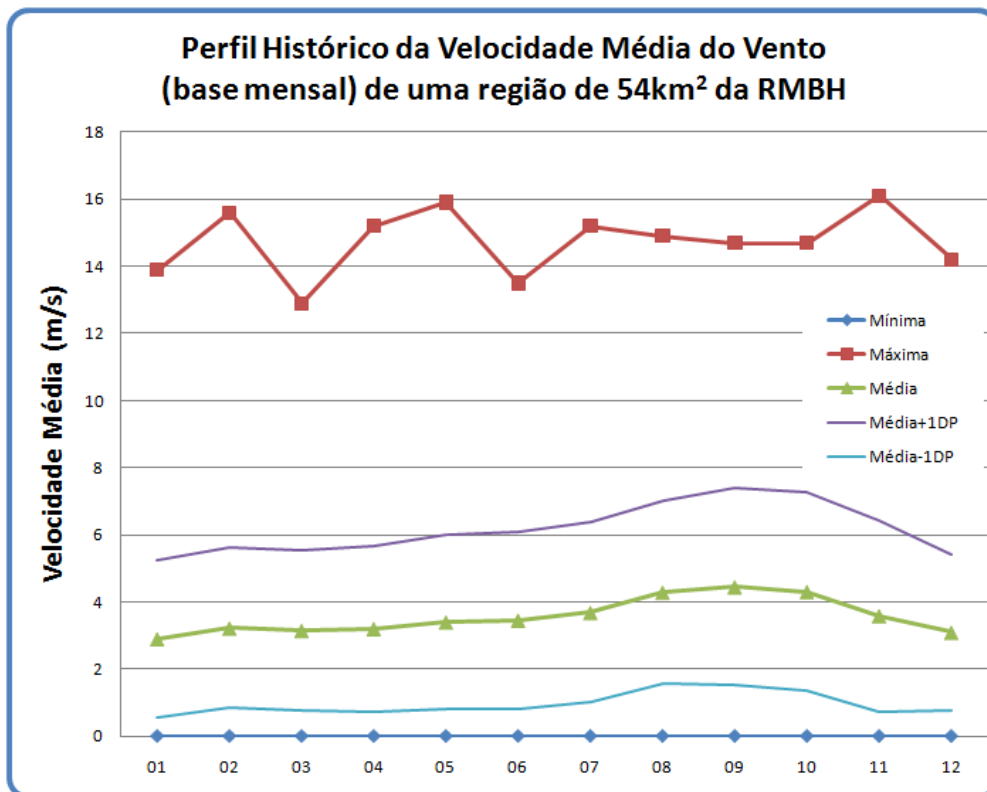


Gráfico 8.9 – Perfil histórico anual da velocidade média do vento na RMBH estratificado por mês



8.5 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE DIREÇÃO DO VENTO

Similarmente aos valores de velocidade média, os valores de direção do vento recebidos se mostraram bastante consistentes e aparentemente sem nenhuma anomalia. Esse fato também foi confirmado pelo mecanismo de qualificação, quando processados os milhões de valores relativos à direção do vento. Verificou-se que mais de 80% deles receberam nota 10, ou seja, se encontravam numa faixa de valores delimitada pela sua média histórica mais ou menos um delta definido pelos meteorologistas.

Como apresentado para a velocidade média, os Gráficos 8.10, 8.11 e 8.12 a seguir, ilustram os perfis de comportamento da direção do vento estratificados numa base horária, numa base semanal e numa base mensal com amostras respectivamente contendo em média 7.745, 1.740 e 3.872 elementos.

Gráfico 8.10 – Perfil histórico diário da direção média do vento estratificada por hora na RMBH

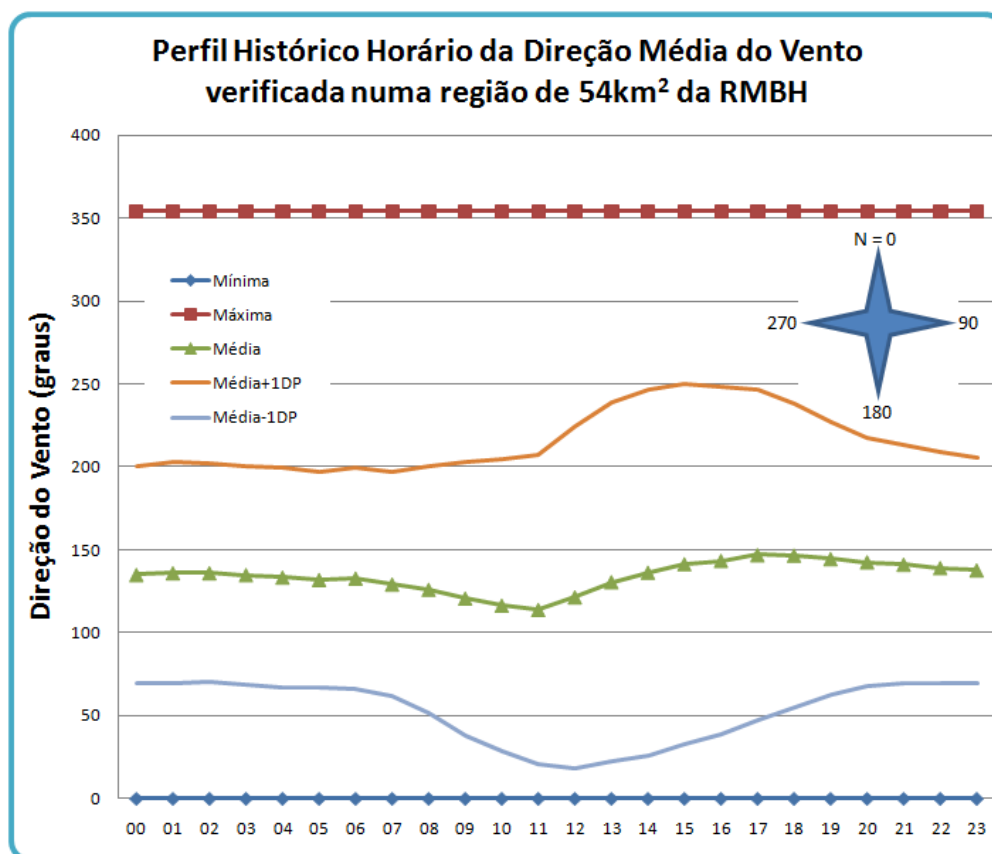


Gráfico 8.11 – Perfil histórico anual da direção média do vento na RMBH (dados estratificados por semana)

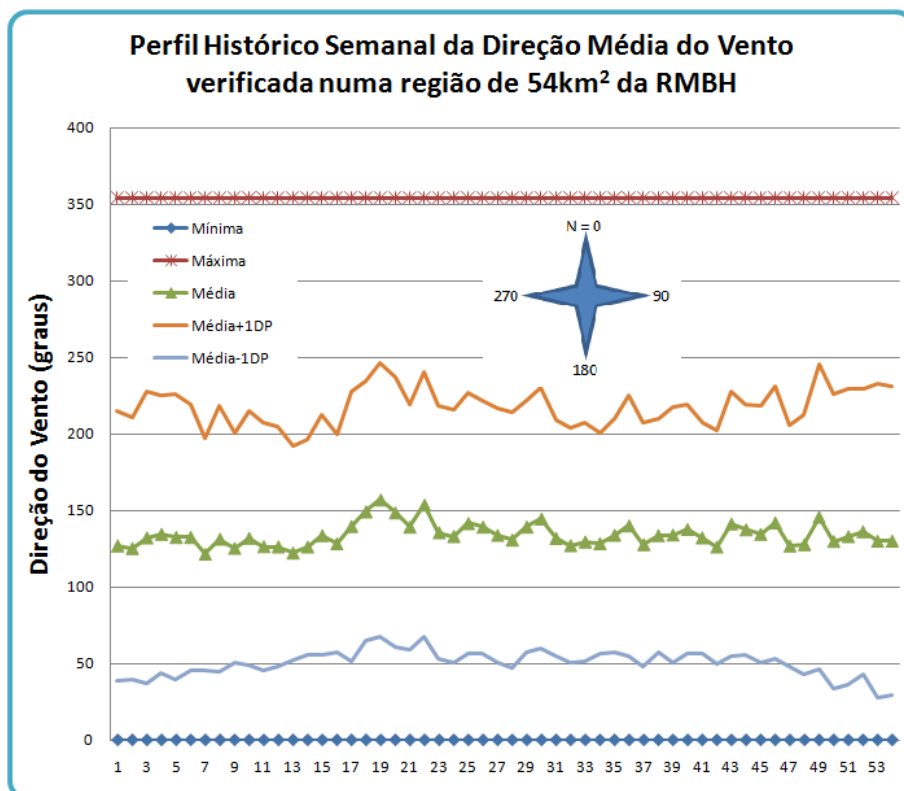
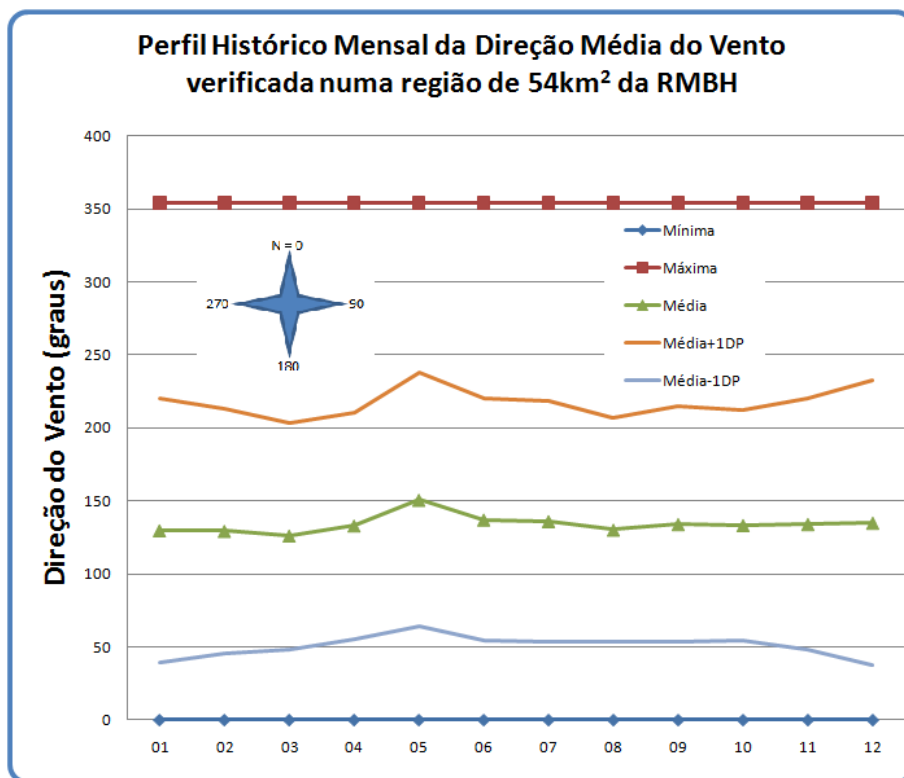


Gráfico 8.12 – Perfil histórico anual da direção média do vento na RMBH (dados estratificados por mês)



8.6 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE RADIAÇÃO SOLAR ACUMULADA

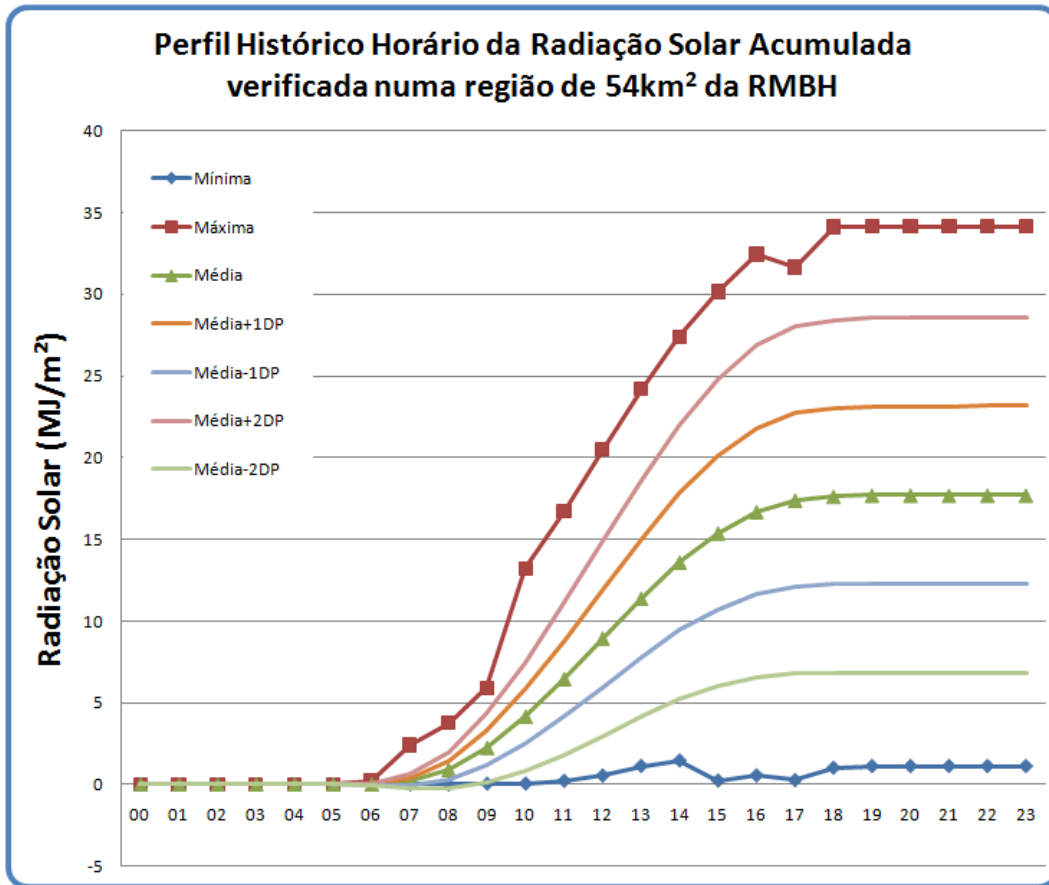
Com um comportamento mais diversificado em relação à distribuição das notas recebidas pelas telemedidas, as medidas de radiação solar acumulada apresentaram valores muito baixos para os valores mínimos *minimorum* horários recebidos, como mostra o Gráfico 8.13 a seguir, que apresenta o perfil de comportamento diário das medidas, estratificadas numa base horária.

Uma avaliação realizada pelos meteorologistas indicou que os valores mínimos estão errados, mesmo sendo encontrados historicamente na base de dados do sistema. Sendo assim, foi feita uma análise técnica a fim de se descobrir a causa de tantas falhas para medidas apresentando pequeno valor de radiação solar acumulada, mesmo muitas horas após o nascer do sol.

A causa foi descoberta e identificada como sendo fruto da própria insuficiência de radiação solar nos meses de inverno, principalmente. A situação é causada porque diversas estações coletoras de dados são alimentadas por painéis solares durante o dia e por baterias durante a noite. Essas baterias são carregadas durante o dia pelos painéis solares e ocorria que nos dias de pouca insolação, algumas estações de aquisição de dados simplesmente se desligavam algumas horas após o nascer do sol devido à incapacidade do painel solar em alimentar os equipamentos e simultaneamente carregar as baterias já descarregadas durante o período de operação noturna. Como o contador de radiação solar acumulada é um acumulador, a cada desligamento da estação seu valor era levado à zero, reiniciando a acumulação. Essa situação explica os baixos valores de mínima radiação solar presentes na base de dados do sistema. Por conta disso, os valores de mínimos *minimorum* horários de radiação solar acumulada devem ser usados com ressalvas.

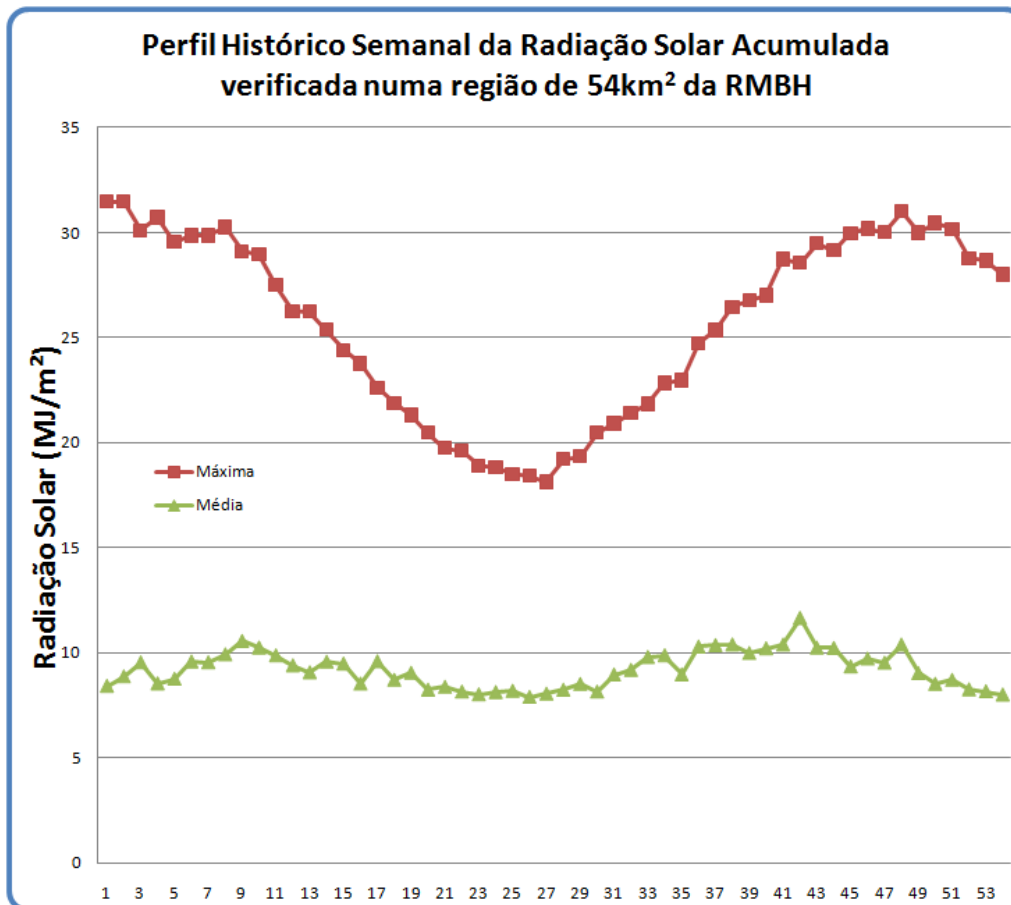
Esse problema pôde ser percebido devido a um importante número de medidas sendo avaliadas com nota 3 (três) pelo sistema de qualificação (mais de 1,8%), ou seja, identificadas como novos mínimos *minimorum*. Foi resolvido com a troca das baterias da estação coletora por outras de maior capacidade e pelo aumento da área de captação do painel fotovoltaico. Verificou-se posteriormente que esses valores, apesar de inadequados, não chegaram a afetar de forma significativa o cálculo da média da radiação solar acumulada.

Gráfico 8.13 – Perfil histórico diário da radiação solar global acumulada



O Gráfico 8.14 a seguir apresenta o perfil de comportamento anual das medidas de radiação solar acumulada, estratificadas numa base semanal. Os valores mínimos não foram apresentados devido aos fatores descritos nos parágrafos anteriores.

Gráfico 8.14 – Perfil histórico anual da radiação solar global acumulada



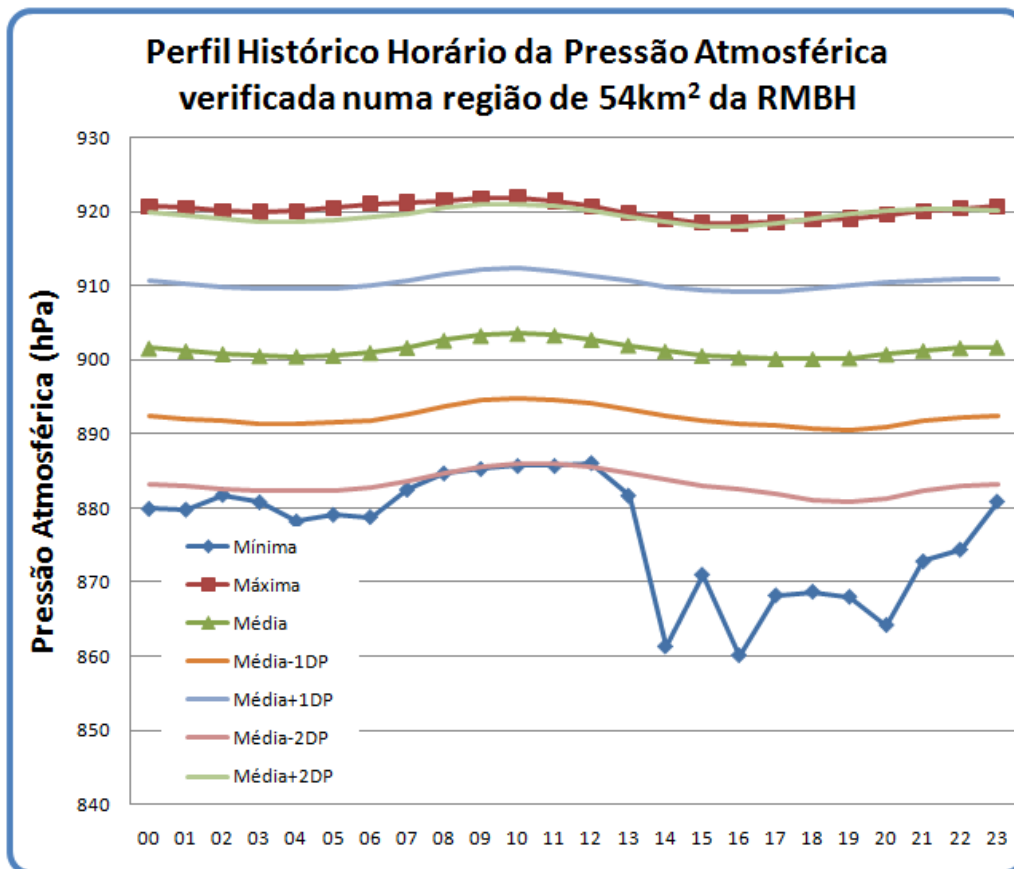
8.7 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE PRESSÃO ATMOSFÉRICA

Praticamente repetindo o comportamento das medições de radiação solar, as medidas de pressão atmosférica também apresentaram um comportamento mais diversificado em relação à distribuição das notas recebidas pelas telemedidas. Algumas medidas de pressão atmosférica apresentaram valores muito baixos para os valores mínimos *minimorum* horários recebidos e outras apresentaram taxas de variação relativamente elevadas para essa grandeza que é considerada pelos meteorologistas como uma das mais bem comportadas no que diz respeito à estabilidade. O Gráfico 8.15 a seguir, que apresenta o perfil de comportamento diário das medidas estratificadas numa base horária, ilustra essa situação.

Uma avaliação realizada pelos meteorologistas indicou que esses valores mínimos *minimorum* não poderiam ser levados em consideração, mesmo sendo encontrados historicamente na base de dados do sistema. Da mesma forma como foi feito para as medidas de radiação solar, foi realizada uma análise técnica a fim de se descobrir a causa de tantas falhas para medidas apresentando pequeno valor de pressão atmosférica e grandes taxas de variação horárias.

A causa foi descoberta e identificada como sendo o mau funcionamento de alguns sensores instalados na plataforma de aquisição de dados e principalmente, para o sistema como um todo, de um defeito persistente no sensor de pressão atmosférica do posto instalado na cidade de Unaí, o qual se mostrou responsável pelo grande número de medidas pontuadas com a nota 2 (dois), indicativa de violação de taxa de variação admitida como normal (dv/dt).

Gráfico 8.15 – Perfil histórico diário da pressão atmosférica na RMBH



O Gráfico 8.16 e o Gráfico 8.17 a seguir apresentam os perfis de comportamento anual das medidas de pressão atmosférica, estratificadas numa base semanal e numa base mensal respectivamente.

Gráfico 8.16 – Perfil histórico anual da pressão atmosférica na RMBH

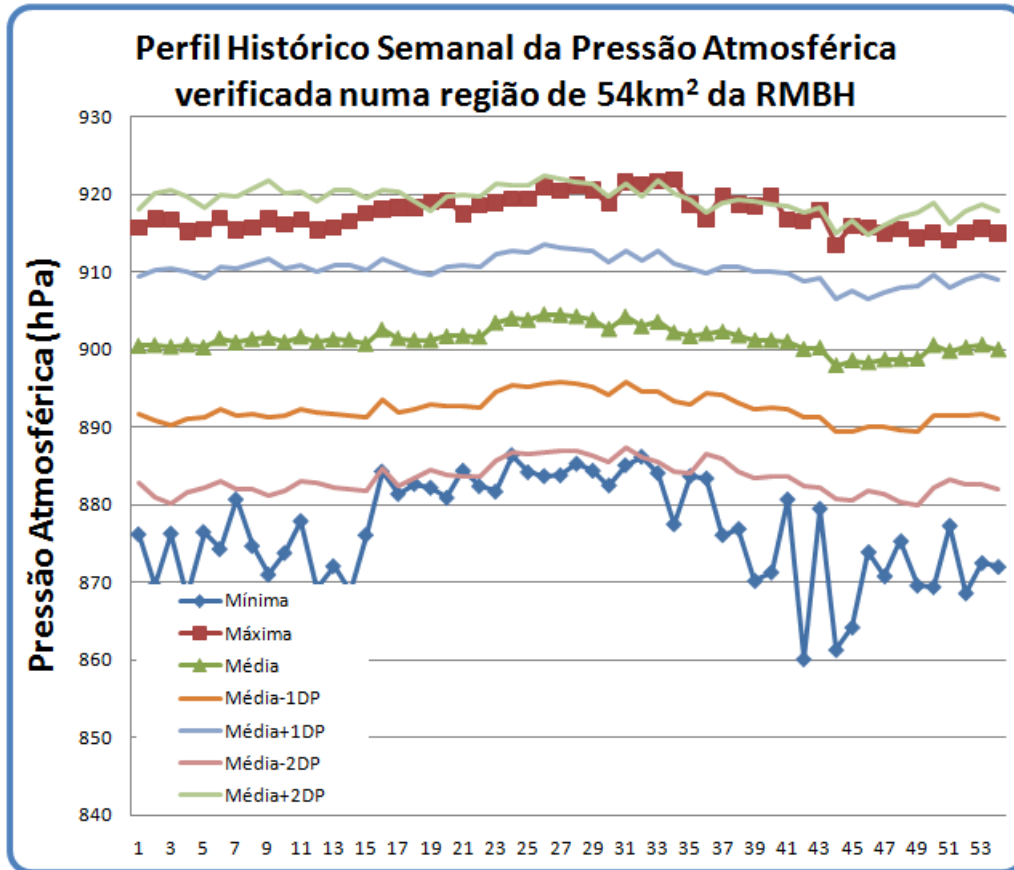
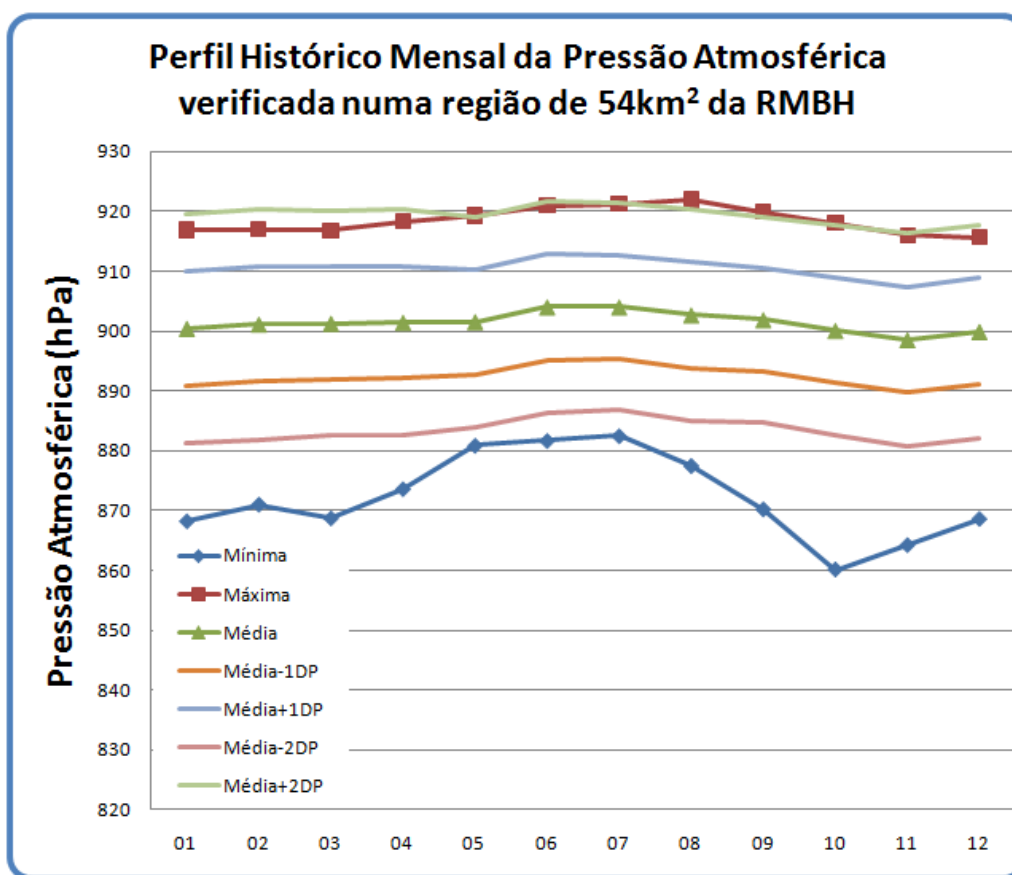


Gráfico 8.17 – Perfil histórico anual da Pressão Atmosférica na RMBH



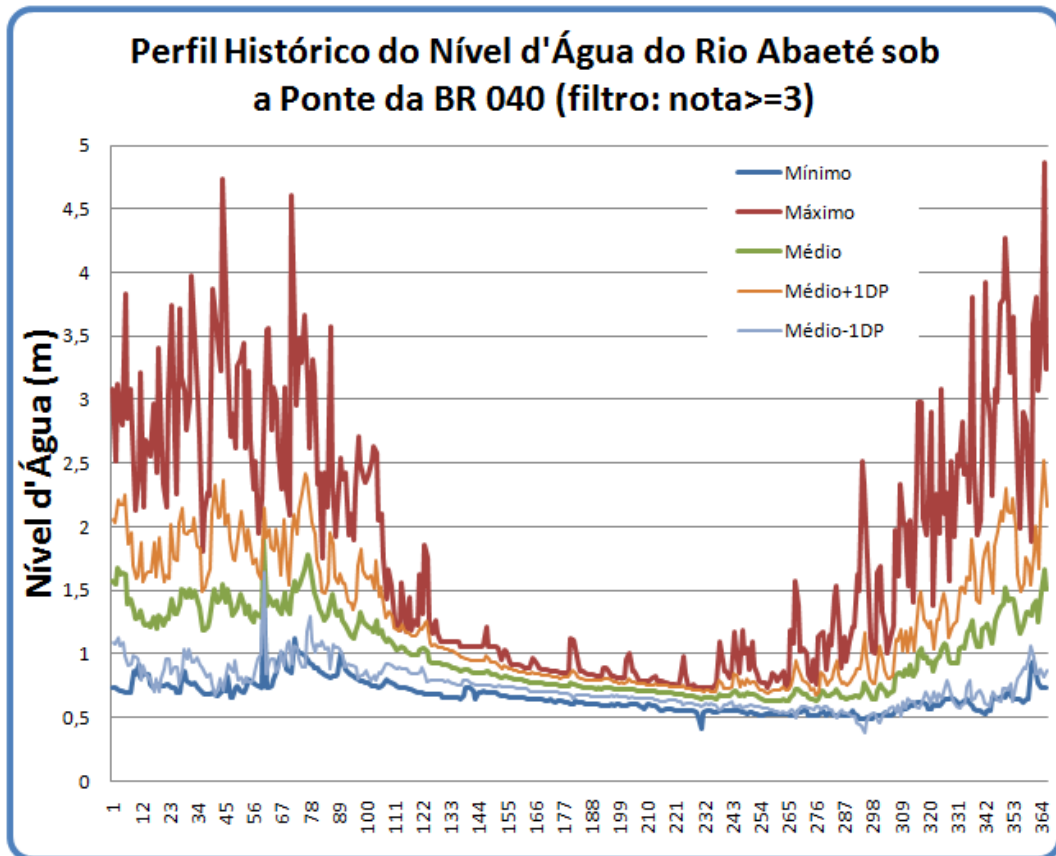
8.8 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE NÍVEL D'ÁGUA DE RIO

Dadas as particularidades da medição de nível d'água, não se pode apresentar aqui um perfil geral ou um perfil por área de uma determinada região. As variações de nível d'água de um rio estão intimamente relacionadas com o perfil topográfico de sua calha e com o seu nível de assoreamento. Sendo assim, a fim de ilustrar os benefícios trazidos pela metodologia de qualificação de grandezas, será apresentada a seguir uma comparação do perfil de comportamento do nível d'água do Rio Abaeté, coletado sob a Ponte da BR 040, no município de São Gonçalo do Abaeté em Minas Gerais.

O Gráfico 8.18 a seguir ilustra o perfil anual do comportamento dos valores de nível d'água já tratados pela metodologia de qualificação de dados. Os dados mostrados estão estratificados por dia. O filtro aqui utilizado foi o de medições com notas maiores ou iguais a 3 (três). O gráfico apresenta cinco

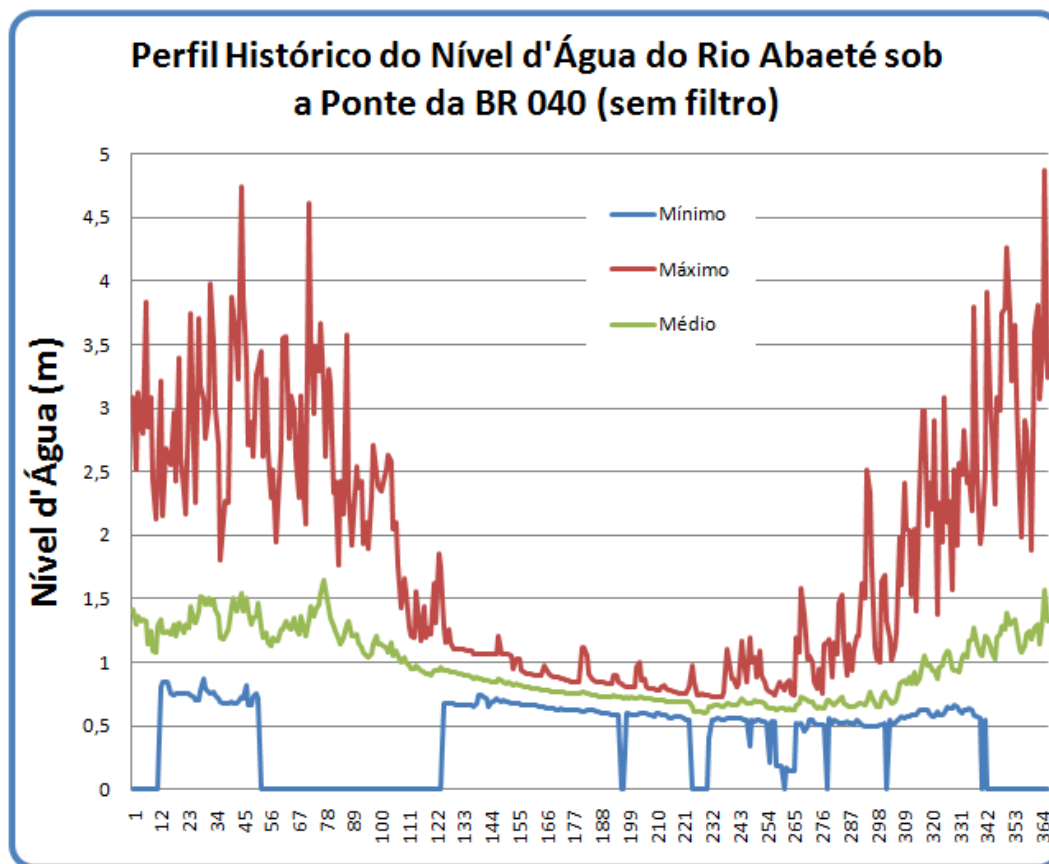
curvas, identificadas de cima para baixo, a saber: Nível d'água máximo, nível d'água médio acrescido de 1 DP (desvio padrão), nível d'água médio, nível d'água médio subtraído de 1 DP e nível d'água mínimo.

Gráfico 8.18 – Perfil histórico do nível d'água do Rio Abaeté



Já o Gráfico 8.19 a seguir ilustra o mesmo perfil anual do comportamento dos valores de nível d'água sem qualquer forma de tratamento qualitativo. Pode-se perceber facilmente pela curva de níveis mínimos diversas indicações de níveis iguais a 0 (zero) o que indicaria, se fosse verdade, que o rio simplesmente secou, o que literalmente ainda não aconteceu. Para fins de simplificação, o gráfico a seguir apresenta três curvas, identificadas de cima para baixo, a saber: Nível d'água máximo, nível d'água médio e nível d'água mínimo (o mais afetado por medições errôneas).

Gráfico 8.19 – Perfil histórico do nível d'água do Rio Abaeté, medido sob a Ponte da BR 040 (valores sem filtragem qualitativa)



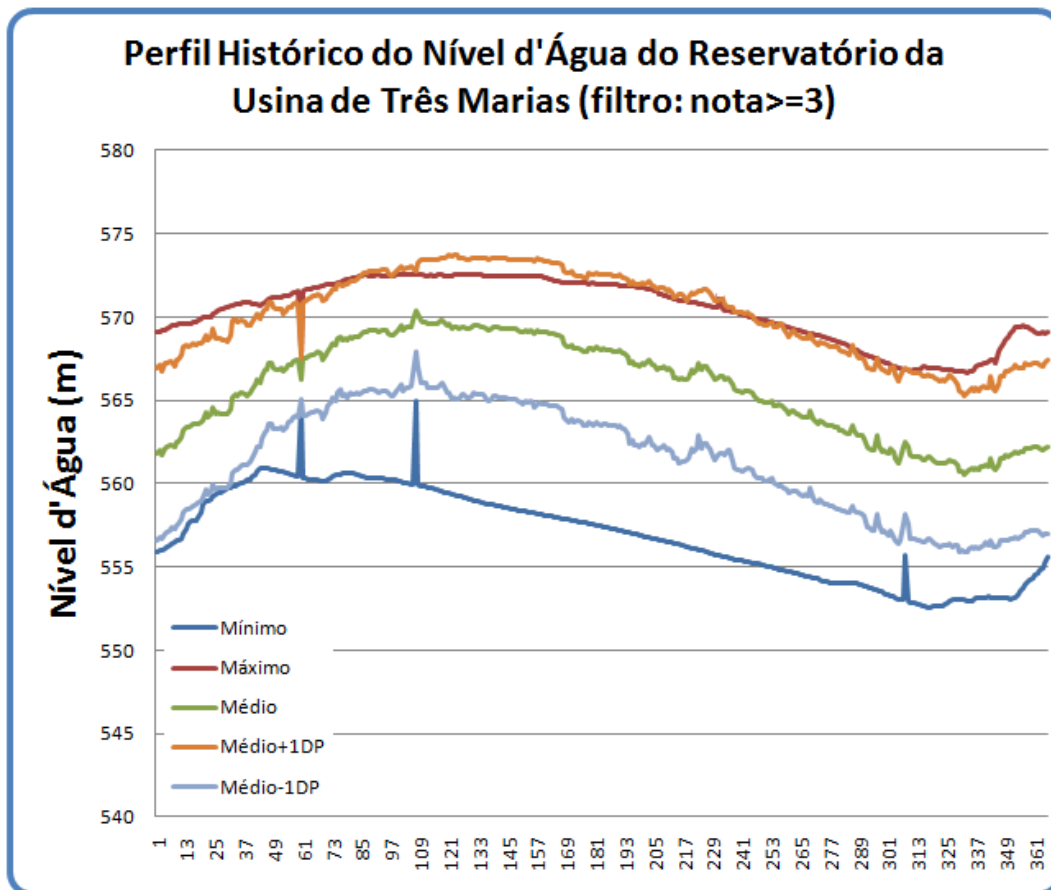
8.9 RESUMO DO COMPORTAMENTO DAS MEDIÇÕES DE NÍVEL D'ÁGUA DE RESERVATÓRIO

Pelos mesmos motivos apresentados para a medição de nível d'água de rio, serão apresentados nessa sessão gráficos comparativos de medição de nível d'água de reservatório, a fim de se evidenciar os benefícios trazidos pela metodologia de qualificação de grandezas. Sendo assim, será apresentada a seguir uma comparação do perfil de comportamento do nível d'água do reservatório da Usina de Três Marias.

O Gráfico 8.20 a seguir ilustra o perfil anual do comportamento dos valores de nível d'água, referenciados ao nível do mar, já tratados pela metodologia de qualificação de dados. Os dados mostrados estão estratificados por dia. O filtro aqui utilizado foi o de medições com notas maiores ou iguais a 3 (três). Conforme o apresentado para o nível de rio, o gráfico do reservatório

apresenta cinco curvas, identificadas de cima para baixo, a saber: Nível d'água máximo, nível d'água médio acrescido de 1 DP (desvio padrão), nível d'água médio, nível d'água médio subtraído de 1 DP e nível d'água mínimo.

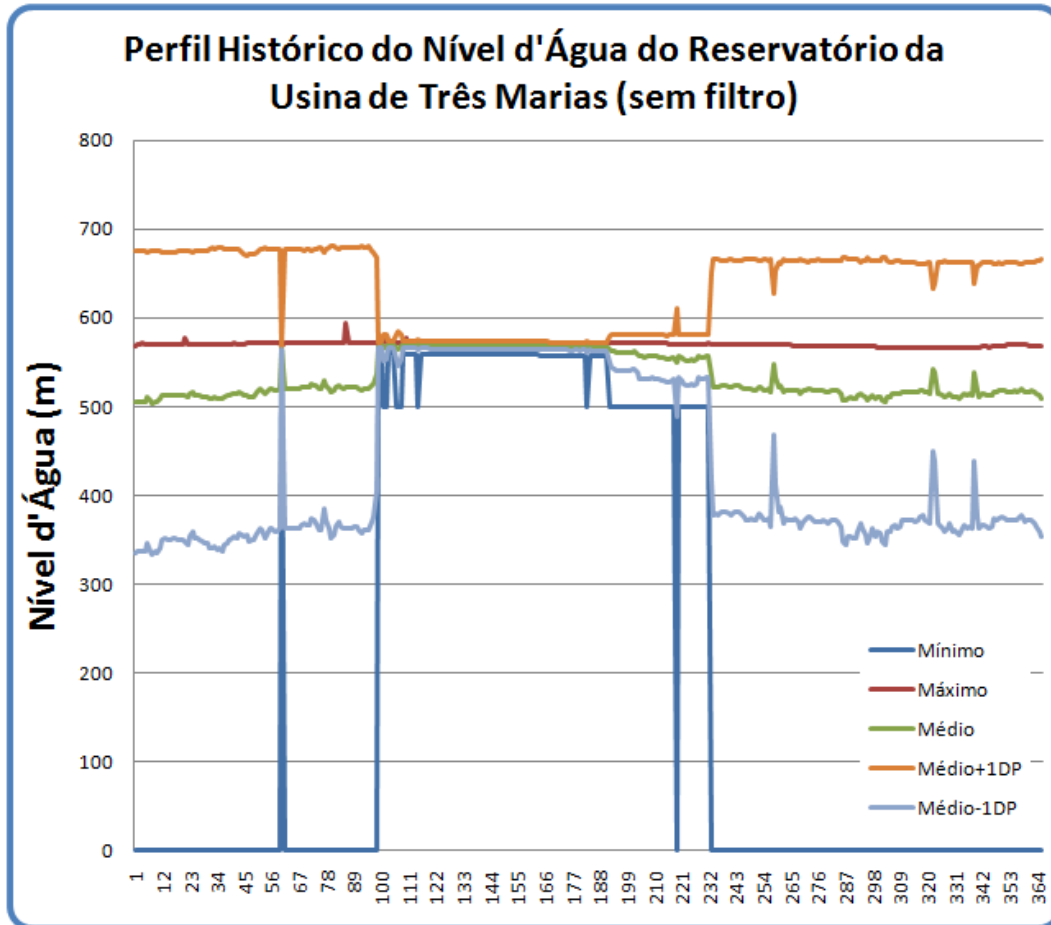
Gráfico 8.20 – Perfil histórico do nível d'água do reservatório da Usina de Três Marias (filtragem qualitativa para nota ≥ 3)



O Gráfico 8.21 a seguir ilustra o mesmo perfil anual do comportamento dos valores de nível d'água do reservatório sem qualquer forma de tratamento qualitativo. Pode-se perceber facilmente pelas curvas de níveis diversas indicações de níveis iguais a 0 (zero), para a curva de nível mínimo e valores praticamente constantes ao longo de todo o ano para a curva de nível máximo. Essas inconsistências afetam de forma significativa os cálculos estatísticos de média e desvio padrão. O gráfico a seguir apresenta cinco curvas, identificadas de cima para baixo, a saber: Nível d'água máximo, nível d'água médio acrescido de 1 DP (desvio padrão), nível d'água médio, nível d'água médio subtraído de 1 DP e nível d'água mínimo. No caso apresentado, todas as curvas estão afetadas

por grandes inconsistências que foram perfeitamente filtradas pela metodologia, conforme apresentado anteriormente.

Gráfico 8.21 – Perfil histórico do nível d'água do reservatório da Usina de Três Marias (sem filtragem qualitativa)



8.10 RESULTADO PRÁTICO DA QUALIFICAÇÃO EM TEMPO REAL PARA O USUÁRIO DO TEMPO REAL

Todos os resumos de comportamento apresentados nos itens anteriores desse capítulo ilustram o resultado alcançado, deixando claro o impacto trazido e o acréscimo de qualidade dado aos resultados das consultas realizadas nas bases de dados do sistema de supervisão. O que se poderia esperar então como resultado de tempo real? A resposta original era apresentar ao usuário do sistema, que tenha necessidade de informação em tempo real a respeito da qualidade do dado, uma diferenciação entre o dado bom e o dado ruim.

O resultado final dos algoritmos que realizam o trabalho proposto pela metodologia de qualificação de dados é uma nota entre 0 (zero) e 10 (dez) para cada uma das telemedições primárias realizadas pelo STH. Esta informação a respeito da nota fica persistida na base de dados, juntamente com a medição em questão. Faltava então, apresentá-la aos usuários nas telas do sistema de supervisão e não apenas como resultado de consultas para a geração de relatórios e gráficos desacoplados da IHM (Interface Homem-Máquina). Para resolver essa questão, foram realizadas diversas alterações nos processos principais de exibição de dados do STH, a fim de apresentar nas telas do próprio sistema a indicação de qualidade da telemedição. A Figura 8.2 ilustra essa nova funcionalidade dada ao sistema de supervisão. Quando se repousa o cursor sobre uma medida surge na tela uma caixa de informações com um resumo a respeito da medição em questão, incluindo ali a nota que a mesma recebeu do sistema de qualificação de dados. No caso ilustrado, a medida está alarmada (fundo rosa) e sua nota é muito boa (10), o que corrobora para a atenção ao alarme apresentado.

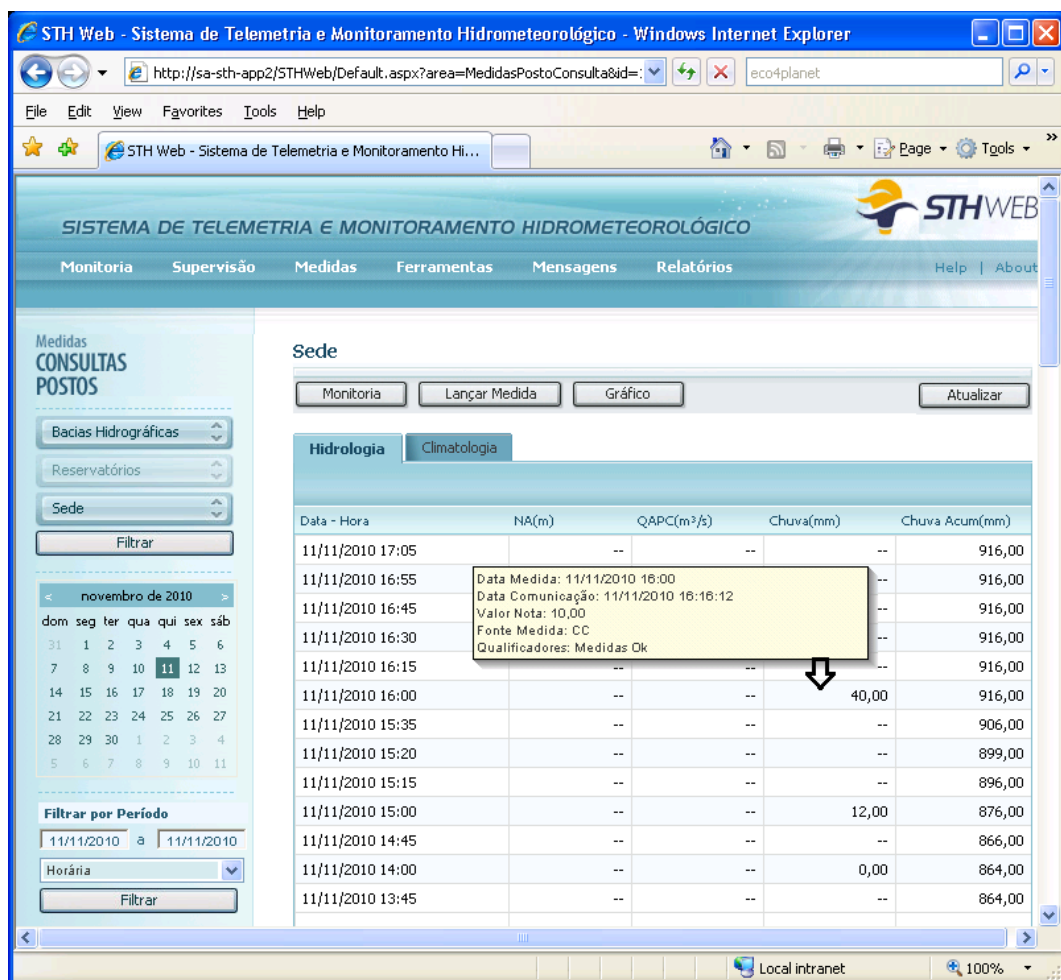
Figura 8.2 – Tela do sistema apresentando em tempo real a nota da telemedição

The screenshot shows the 'Sede' (Station) monitoring page in the STH Web system. The page displays a table of meteorological data for the 'Climatologia' (Climatology) section. The table columns include Date-Time, Wind Speed (VelVento), Air Temperature (TempAr), Solar Radiation (RadSol), Relative Humidity (UmidRel), Wind Direction (DirVento), and Atmospheric Pressure (PrAtm). A tooltip is visible over the row for '14/03/2010 21:15', which has a pink background indicating an alarm. The tooltip provides details: 'Data Medida: 14/03/2010 21:15', 'Data Comunicação: 14/03/2010 21:15:32', 'Valor Nota: 10,00', 'Fonte Medida: CCU', and 'Qualificadores: Limite Inferior Violado'. The interface also includes a navigation menu, a calendar for March 2010, and various filters.

Data - Hora	VelVento(m/s)	TempAr(°C)	RadSol(MJ/m²)	UmidRel(%)	DirVento(°)	PrAtm(hPa)
14/03/2010 23:45	0,10	22,8	14,35	80	124 SE	907,7
14/03/2010 23:30	0,50	22,8	14,35	78	84 NE	907,7
14/03/2010 23:15	0,50	22,8	14,35	80	96 SE	907,7
14/03/2010 23:00	0,60	22,6	14,35	83	174 SE	908,0
14/03/2010 22:45	0,20	22,4	14,35	85	124 SE	908,1
14/03/2010 22:30	0,20	22,3	14,35	87	107 SE	908,1
14/03/2010 22:15	0,00	22,4	14,35	85	158 SE	908,3
14/03/2010 22:00	0,00	22,6	14,35	86	84 NE	908,1
14/03/2010 21:45	0,00	22,6	14,35	86	0 N	907,8
14/03/2010 21:30	0,00	22,6	14,35	86	17 NE	907,6
14/03/2010 21:15	0,00	22,4	14,35	84	338 NO	907,4
14/03/2010 21:00	0,00	22,6	14,35	80	135 SE	907,3
14/03/2010 20:45	0,20	22,5	14,35	79	174 SE	907,3
14/03/2010 20:30	0,20	22,2	14,35	80	158 SE	907,2

De forma a dar uma abrangência sistêmica aos benefícios trazidos pelo sistema de qualificação, foram também alterados vários dos muitos algoritmos do STH responsáveis pelos cálculos periódicos, a fim de que esses passassem a levar em consideração a nota recebida pelas parcelas que o compõem. Nesse caso decidiu-se que, mesmo que indiretamente, o resultado de um cálculo também seria pontuado, com o mesmo recebendo como nota, a menor nota dada a uma das suas parcelas. A forma de apresentação desse resultado da qualificação indireta do resultado de um cálculo é a mesma descrita no parágrafo anterior, conforme ilustra a Figura 8.3, onde se pode notar o cálculo de chuva horária recebendo uma nota baseada nas notas da chuva acumulada que é telemedida através de um pluviômetro. Tanto o valor acumulado pelo pluviômetro às 15hs (876mm) quanto o das 16hs (916mm) receberam nota 10, tendo o cálculo da chuva horária das 16hs herdado essa nota.

Figura 8.3 – Tela do sistema apresentando em tempo real a nota dada a um cálculo realizado pelo sistema de supervisão



9 Conclusões e Propostas de Continuidade

9.1 INTRODUÇÃO

Este capítulo encerra esta dissertação e apresenta algumas propostas de continuidade do trabalho com vistas para o maior aproveitamento das séries de dados tratadas pelo sistema de qualificação desenvolvido como atividade fim desse trabalho. O aplicativo desenvolvido para a qualificação de dados em tempo real foi integrado ao STH, colocado em operação e isso faz com que estas séries de dados qualificadas cresçam continuamente.

9.2 NOVAS UTILIDADES PARA O SISTEMA DE QUALIFICAÇÃO

O uso do aplicativo de qualificação de dados e da base estatística gerada pelo mesmo acabou mostrando outras utilidades tais como alertar os responsáveis pela manutenção da plataforma sobre a necessidade de manutenções preventivas e corretivas nos coletores de dados, quando para determinado posto, passa-se a receber de forma contínua telemedições com notas muito baixas.

Além da utilidade descrita no parágrafo anterior, graças ao sistema de qualificação de dados em tempo real, uma nova funcionalidade foi acrescentada ao STH, que permite fazer uma qualificação para os dados resultantes de cálculos realizados por outros módulos do sistema. Nesta qualificação, a menor nota dada a uma parcela que compõe o cálculo é atribuída ao resultado do cálculo, indicando ao cliente da informação que aquele valor resultante deve ser usado com precaução.

Ainda, este novo atributo da medida é também levado em consideração quando da realização de simulações do comportamento de reservatórios, do

monitoramento de inequações e da geração de mensagens para os usuários do sistema e plantonistas, dentre outros.

9.3 A QUALIDADE DO DADO COM FOCO NO CLIENTE

O STH, de forma direta e indireta, é uma fonte de dados hidrológicos e meteorológicos para um grande número de clientes internos e externos à Cemig, que extraem destes dados as mais diversas informações. Dentre os principais clientes externos podemos citar a ANA – Agência Nacional das Águas, a ANEEL – Agência Nacional de Energia Elétrica, o ONS – Operador Nacional do Sistema, o CPTEC – Centro de Previsão de Tempo e Estudos Climáticos, o INPE – Instituto Nacional de Pesquisas Espaciais, Órgãos de Defesa Civil de inúmeras prefeituras e órgãos de imprensa, dentre outros. Além destes existem inúmeros clientes internos das mais diversas áreas da Cemig que se utilizam dos dados das mais diversas formas, seja para o suporte de informações aos mais diversos projetos até o suporte a teses acadêmicas e projetos de pesquisa, desenvolvimento e inovação.

O sistema de qualificação de dados desenvolvido, de forma simples, eficiente, transparente e quase anônima, ajuda a garantir um nível e qualidade no que diz respeito à informação fornecida a todos os clientes do sistema, assegurando ou informando ao cliente sobre o grau de confiança a ser depositado no dado e possibilitando a ele, caso deseje, a elaboração de filtros para dados considerados errados quando da criação de relatórios e consultas a partir dos dados disponibilizados pelo STH.

9.4 PROPOSTAS DE CONTINUIDADE

Acredita-se que esse trabalho não se encerra aqui. Esta dissertação apresentou uma idéia e o produto dela originado com certeza não é intocável e pode ser melhorado e enriquecido em diversos aspectos, através da incorporação de novas funcionalidades. Na medida em que as bases de dados históricos vão crescendo e conseqüentemente vão crescendo os tamanhos das amostras de dados, os dados estatísticos das mesmas vão se tornando cada vez mais consistentes e acabarão permitindo novas possibilidades de

processamento e avaliação dos mesmos, além da realização de correlações entre eles.

Uma proposta possível é a da utilização dos dados históricos para a realização de testes de hipótese com diferentes níveis de incertezas, a fim de se construir um previsor *on-line* de valores futuros para as grandezas hidrológicas e meteorológicas, através do qual se poderia traçar curvas indicativas das faixas limites de variação mais prováveis para determinada grandeza para uma data específica em estudo.

Outra possibilidade é a de utilização desta idéia em outros sistemas de medição que tenham, similarmente ao STH, uma grande massa de dados históricos e que necessitem de uma forma de se ter uma idéia sobre a qualidade dos dados armazenados.

Ainda sob o aspecto de continuidade, seria muito interessante a integração da imensa base de dados gerada com aplicativos e sistemas de visualização e análise gráfica de informações.

9.5 PALAVRAS FINAIS

Esta dissertação apresentou uma idéia extremamente simples e deu origem a um produto que agregou valor a um sistema de supervisão e monitoramento de dados em tempo real. Sem o uso de algoritmos complexos e computacionalmente exigentes ofereceu-se aos usuários e clientes do sistema um excelente nível de confiança nos dados apresentados aos mesmos.

10 Bibliografia

Braga Jr., B. P. F. ; Tucci, C. E. M. ; Pompeu, C. T.; Barth, F. T.; Fill, H. D.; Kelman, J. - Modelos para Gerenciamento de Recursos Hídricos, Nobel/ABRH, 1987.

Chaves, A. G. M.; Setti, A. A. ; Pereira, I. C. ; Lima, J. E. F. - Introdução Ao Gerenciamento De Recursos Hídricos, ANEEL-ANA, 2001.

Costa Neto, P. L. O. - Estatística, Edgard Blucher, 1977.

El Tiempo Y El Agua - Organizacion Meteorologica Mundial, Secretaría De La Organización Meteorológica Mundial, OMM, 1977.

Gomes, F. P. - Curso de Estatística Experimental, Nobel, 1982.

Gray, D. M. - Handbook on the Principles of Hydrology, McGraw-Hill, 1970.

Jeffrey, C.; Downing, D. - Estatística Aplicada, Saraiva, 2001.

Leite, E. A.; Pessoa, M. L. ; Mine, M. R. M. - Meteorologia e Hidrologia - Aspectos e Considerações no Contexto Brasileiro, COPEL, 1994.

Martins, G. A.; Fonseca, J. S. - Curso de Estatística, Atlas, 1982.

McFedries, Paul – Formulas and Functions with Microsoft Office Excel 2007, QUE, 2007

Meet Minitab v14 – Minitab Inc., 2004

Memória, J. M. P. – Breve História da Estatística, Embrapa, 2004

Mlodinow, L – O Andar do Bêbado: Como o acaso determina nossas vidas, Zahar, 2008

Rousseuw, P. J. e Zomeren, B. C.- Unmasking multivariate outliers and leverage points. Journal of the American Statistical Association, 1990

Spiegel, M. R. - Estatística, Schaum Publishing Company, McGraw-Hill, 1984.

Taleb, N. N. – A Lógica do Cisne Negro: o impacto do altamente improvável, BestSeller, 2007

Toledo, L. G.; Martins, G. A.; Fonseca, J. S. - Estatística Aplicada, Atlas, 1980.

Triola, M. F. - Introdução à Estatística, LTC, 1999.

Walkenbach, John – Excel 2007 Bible, Wiley Publishing Inc., 2007