

**UNIVERSIDADE DE CAXIAS DO SUL**  
**Centro de Computação e Tecnologia da Informação**  
**Curso de Bacharelado em Ciência da Computação**

**Marcio Bigolin**

**SISTEMA PARA AVALIAÇÃO DA QUALIDADE DE DADOS ATRAVÉS  
DE PARAMÊTROS FÍSICOS-QUÍMICOS DA ÁGUA APLICADO A UMA  
BACIA HIDROGRÁFICA**

**Caxias do Sul**

**2010**

**Marcio Bigolin**

**SISTEMA PARA AVALIAÇÃO DA QUALIDADE DE DADOS ATRAVÉS  
DE PARAMÊTROS FÍSICO-QUÍMICOS DA ÁGUA APLICADO A UMA  
BACIA HIDROGRÁFICA**

Trabalho de Conclusão de Curso  
para obtenção do Grau de  
Bacharel em Ciência da  
Computação pela Universidade  
de Caxias do Sul.

**Helena Graziottin Ribeiro**  
**Orientadora**

**Vania Elisabete Schneider**  
**Co-orientadora**

**Caxias do Sul**

2010

**Dedico este trabalho a todos  
que de alguma forma me  
ajudaram a chegar até aqui.**

## AGRADECIMENTOS

*Agradeço a Deus, por me dar paciência e sabedoria para ter chegado até aqui.*

*Ao meus pais, pelo apoio e dedicação em todos esses anos de estudo. Sem eles eu jamais chegaria até aqui,*

*À Universidade de Caxias do Sul, especialmente ao ISAM (Instituto de Saneamento Ambiental), por dar a oportunidade de explorar um dos problemas e encontrar soluções,*

*A minha orientadora Dr.<sup>a</sup> **Helena Graziottin**, e co-orientadora Dr.<sup>a</sup> **Vania Elisabete Schneider** pelo apoio desde o início desta jornada e que me proporcionaram muito mais que orientação para este trabalho, mas sim, um grande exemplo de autenticidade, valores éticos e amor ao estudo, que me serviram de espelho para minha própria construção intelectual. Agradeço ao valioso exemplo e amizade.*

*Agradeço a todos os meus professores, pela atenção a mim dedicada durante a realização deste trabalho,*

*Agradeço ao Taison Anderson Bortolin por ter ajudado a montar as regras e levantado os requisitos para o sistema de avaliação de qualidade de dados.*

*A minha namorada Lauren Girardi Cristofoli, e sua mãe Clarice pelo apoio, paciência e incentivo.*

*Enfim, agradeço a todos que, de alguma forma, contribuíram não só para realização deste trabalho, mas para a conclusão de mais uma etapa de minha vida.*

*“Não basta ensinar ao homem uma especialidade. Porque se tornará assim uma máquina utilizável, mas não uma personalidade. É necessário que adquira um sentimento, um senso prático daquilo que vale a pena ser empreendido, daquilo que é belo, do que é moralmente correto. A não ser assim, ele se assemelhará, com seus conhecimentos profissionais, mais a um cão ensinado do que a uma criatura harmoniosamente desenvolvida. Deve aprender a compreender as motivações dos homens, suas quimeras e suas angústias para determinar com exatidão seu lugar exato em relação a seus próximos e à comunidade”.*

*Einstein (1953:29)*

## RESUMO

A busca pela qualidade de um produto normalmente é uma meta para qualquer organização. Com a Era da Informação, os dados que geram essas informações devem possuir um alto grau de confiabilidade e qualidade. O estudo de qualidade de dados não é somente aplicado a competitividade industrial, é importante também para os estudos científicos nas mais diversas áreas. Na área de recursos hídricos não é diferente. A análise da qualidade da água, por exemplo, precisa de qualidade dos dados coletados a campo ou a partir de resultados oriundos de análises laboratoriais. A análise incorreta de um parâmetro pode influenciar totalmente na tomada de decisão referente a um recurso hídrico, e inclusive levar a uma classificação de uso incorreto frente a legislação. O presente trabalho apresenta uma solução para a busca da melhoria da qualidade de dados para recursos hídricos lóticos e lênticos. O trabalho analisa aplica as dimensões de qualidade de dados para o melhoramento dos parâmetros indicadores da qualidade da água, coletados nas mais diversas fontes de informação.

**Palavras-chaves:** Qualidade de dados, Dimensões de dados, Parâmetros indicadores de qualidade da água

## **ABSTRACT**

The search for quality of a product is usually a goal for any organization. With the Information's Age, the data that generate this information should have a high degree of reliability and quality. The study of data quality is not only applied to industrial competitiveness, but also important for scientific studies in several areas. In the area of water resources is no different. The analysis of water quality for example, needs the quality of data collected in the field and from results of laboratory tests. The analysis of an incorrect parameter can fully influence decision-making concerning a water resource and even lead to an incorrect classification of face-to-use legislation. This paper presents a solution to the quest for improving the quality of data for water lotic and lentic. The work applies the intuitive dimensions of data to improve data quality of indicators of water quality collected in a variety of sources.

**Keywords:** Data quality, Dimensions of data, parameters of water quality indicators

## LISTA DE ILUSTRAÇÕES

Ilustração 1 - Qualidade no ciclo de vida do Software.....	18
Ilustração 2 - Modelo de qualidade para a qualidade externa e interna.....	19
Ilustração 3 - Modelo de qualidade para qualidade em uso.....	20
Ilustração 4 - Relação entre atributos, parâmetros e indicadores de qualidade de dados.....	22
Ilustração 5 - Organograma da hierarquia de qualidade de dados.....	23
Ilustração 6 - Diagrama ER Completo do Sistema .....	36
Ilustração 7 - Fluxograma de inserção de dados.....	38
Ilustração 8 - Modelo de dados para as regras de validação.....	39
Ilustração 9 - Equação da média ponderada.....	40
Ilustração 10 - Campos de descrição que ajudam a compreender melhor um dado.....	46
Ilustração 11 - Condição climática atributo transformado em entidade.....	48
Ilustração 12 - Modelagem dimensional do sistema.....	49
Ilustração 13 - Arquitetura cliente/servidor de três camadas .....	50
Ilustração 14 - Execução do comando uname -a mostrando a versão do linux .....	54
Ilustração 15 - Resultado da consulta SELECT version(); no PostgreSQL.....	54
Ilustração 16 - Diagrama de Caso de Uso geral do Sistema .....	55
Ilustração 17 - Descrição do caso de uso avaliar a qualidade do dado.....	55
Ilustração 18 - Descrição do caso de uso Quantificar a qualidade do dado.....	56
Ilustração 19 - Descrição do caso de uso realizar avaliação subjetiva do dado.....	56
Ilustração 20 - Descrição do caso de uso editar dado.....	57
Ilustração 21 - Diagrama de caso de uso do Módulo de Regras.....	57
Ilustração 22 - Descrição do caso de uso Validar o dado.....	58
Ilustração 23 - Descrição do caso de uso Calcular / Recalcular a confiabilidade do dado.....	58
Ilustração 24 - Descrição do caso de uso Aplicar regra ao dado.....	59
Ilustração 25 - Modelo de dados da implementação do módulo de regras.....	60
Ilustração 26 - Diagrama de atividades da trigger executada antes da Inserção.....	61
Ilustração 27 - Diagrama de atividades da trigger executado depois da Inserção.....	62
Ilustração 28 - Diagrama de Componentes demonstrando a arquitetura do módulo de avaliação e controle das notas.....	65
Ilustração 29 - Diagrama de Caso de uso do módulo de avaliação e controle das notas.....	65
Ilustração 30 - Descrição do caso de uso Justificar Dado.....	66
Ilustração 31 - Descrição do caso de uso Descartar Dado.....	66
Ilustração 32 - Descrição do caso de uso Consultar Informações.....	67
Ilustração 33 - Fórmula para calculo da complexidade ciclomática.....	69
Ilustração 34 - Grafo ciclomático das regras estatísticas.....	70
Ilustração 35 - Gráfico da aplicabilidade do sistema aos dados utilizados.....	73
Ilustração 36 - Gráfico da aplicabilidade para grandes quantidades de dados.....	73
Ilustração 37 - Gráfico utilidade para as avaliações comuns aos dados.....	74
Ilustração 38 - Gráfico da facilidade de aprendizado do sistema.....	75
Ilustração 39 - Gráfico da facilidade de utilização do sistema.....	75
Ilustração 40 - Gráfico de facilidade para a interpretação da qualidade do dado.....	76
Ilustração 41 - Gráfico da confiança da avaliação que o sistema realiza sobre os dados.....	76
Ilustração 42 - Gráfico da confiabilidade das regras.....	77
Ilustração 43 - Gráfico sobre a metodologia da avaliação da qualidade dos dados.....	77
Ilustração 44 - Esquema para o melhoramento continuo da qualidade dos dados.....	80



## LISTA DE TABELAS

Tabela 1 - Classificação das características de qualidade de dados.....	21
Tabela 2 - Dimensões de qualidade de dados.....	24
Tabela 3 - Classificação das águas.....	26
Tabela 4 - Classes de qualidade de água doce .....	27
Tabela 5 - Limites do pH em água doce.....	30
Tabela 6 - Limites do Oxigênio Dissolvido em águas doces.....	31
Tabela 7 - Limites do DBO <sub>5,20</sub> em água doce.....	33
Tabela 8 - Pesos dos parâmetros do cálculo do IQA.....	34
Tabela 9 - Série de amostras de um ponto do Rio São Marcos.....	37
Tabela 10 - Itens do sistema que auxiliam a conhecer a reputação de um dado.....	42
Tabela 11 - Medidas de segurança para os dados adotado no sistema.....	51
Tabela 12 - Faixa de qualidade conforme a nota do dado.....	63
Tabela 13 - Consultas do módulo Avaliação e controle das notas.....	67
Tabela 14 - Testes de caixa branca realizados por regra.....	70

## LISTA DE ABREVIATURAS E SIGLAS

<b>SIGLA</b>	<b>SIGNIFICADO EM PORTUGUÊS</b>	<b>SIGNIFICADO EM Inglês</b>
3NF	Terceira forma normal	Third Normal Form
ABNT	Associação Brasileira de Normas Técnicas	
BI	Inteligência de Negócios	Business Intelligence
CONAMA	Conselho Nacional do Meio Ambiente	National Council on the Environment
CSS	Folha de Estilo em cascata	Cascading Style Sheet
DBO	Demanda Bioquímica de Oxigênio	Biochemical Oxygen Demand
DQO	Demanda Química de Oxigênio	Chemical Oxygen Demand
ER	Entidade Relacionamento	Entity Relationship
ETA	Estação de Tratamento de Água	
FAQ	Perguntas Frequentes	Frequently Asked Questions
GUI	Interface Gráfica do Utilizador	Graphical User Interface
HTML	Linguagem de Marcação de Hipertexto	HyperText Markup Language
IQA	Índice de Qualidade da Água	Water Quality Index
ISAM	Instituto de Saneamento Ambiental	Institute of Environmental Sanitation
ISO	Organização Internacional de Normalização	International Organization for Standardization
MVC	Modelo Visualização e Controle	Model View Controller
NSF	Fundação Nacional Sanitária	National Sanitation Foundation
PEGQ	Projeto de Especialização em Gestão de Qualidade	
pH	Potencial Hidrogeniônico	Hydrogen potential
PHP	Pré Processador de Hypertexto	Hypertext Pre Processor
SGBD	Sistema Gerenciador de Banco de Dados	
SIA	Sistema de informação Ambiental	Environmental Information System
SQL	Linguagem Estruturada de Pesquisa	Structured Query Language
UML	Linguagem Unificada de Modelagem	Unified Modeling Language

# SUMÁRIO

1	Introdução.....	13
1.1	O Projeto e o ISAM.....	15
1.2	Objetivos.....	16
1.3	Organização do Trabalho.....	16
2	Qualidade .....	17
2.1	Qualidade de software.....	17
2.2	Qualidade de dados .....	20
2.2.1	Características de qualidade de dados.....	21
3	Parâmetros indicadores da qualidade da água.....	26
3.1	Escolha dos parâmetros.....	29
3.1.1	PH.....	29
3.1.2	Oxigênio Dissolvido.....	30
3.1.3	Temperatura.....	31
3.1.4	Condutividade elétrica da água.....	32
3.1.5	Demanda Bioquímica de Oxigênio (DBO).....	33
3.1.6	Demanda química de Oxigênio (DQO).....	33
3.2	Outros parâmetros da água.....	34
4	As dimensões de qualidade de dados aplicadas ao sistema.....	35
4.1	Qualidade de dados intrínseca.....	37
4.1.1	Acurácia.....	37
4.1.2	Credibilidade.....	39
4.1.3	Objetividade.....	41
4.1.4	Reputação.....	41
4.2	Qualidade de dados no contexto.....	42
4.2.1	Valor acrescentado.....	43
4.2.2	Relevância.....	43
4.2.3	Idade dos dados.....	44
4.2.4	Totalidade.....	44
4.2.5	Quantidade apropriada de dados.....	45
4.3	Representação.....	45
4.3.1	Interpretabilidade.....	46
4.3.2	Compreensibilidade.....	46
4.3.3	Representação consistente.....	47
4.3.4	Representação concisa.....	47
4.3.5	Facilidade de manipulação.....	48
4.4	Acessibilidade.....	49
4.4.1	Segurança .....	50
5	Modelagem e Implementação do sistema de qualidade de dados.....	53
5.1	Seleção do Sistema gerenciador de banco de dados.....	53
5.2	Criação do banco de dados.....	54
5.3	Modelagem do sistema.....	54
5.4	Módulo de regras.....	57
5.4.1	Seleção e execução de regras.....	59
5.4.2	Tipos de regras.....	62
5.4.3	Pesos e notas finais das regras.....	63

5.5 Módulo de avaliação e controle das notas.....	64
5.6 Considerações sobre o sistema.....	67
6 Testes e Implantação.....	69
6.1 Testes para as regras.....	69
6.2 Testes de uso do software.....	72
6.2.1 Avaliação de aplicabilidade e conformidade .....	72
6.2.2 Avaliação de usabilidade.....	74
6.2.3 Avaliação de confiabilidade.....	76
6.3 Implantação.....	78
7 Resultado e discussões finais.....	79
7.1 Objetivos alcançados .....	79
7.2 Trabalhos futuros.....	81
8 Referências.....	82

# 1 INTRODUÇÃO

A água é um recurso vital para a sobrevivência dos seres. Ela representa um elo de ligação entre todos os ecossistemas do planeta, sendo difícil imaginar a vida sem sua presença. Para se avaliar a qualidade ambiental como um todo, é preciso obter informações que estejam integradas entre os fatores bióticos e abióticos que regem o funcionamento do ecossistema.

A constante necessidade de manter e avaliar a qualidade de um recurso hídrico, seja para o lazer, para o consumo, ou manutenção da biodiversidade, torna indispensável que se faça um monitoramento constante e, dessa forma, faça diagnósticos e prognósticos sobre a situação do corpo hídrico.

A água já foi tema principal de muitos livros e artigos científicos e continua sendo discutida entre muitas sociedades e grupos. Isso não é por menos, uma vez que é de vital importância para que qualquer ser vivo possa manter sua existência. Sua utilidade é variada, desde atividades básicas de higiene até geração de energia, e por isso deve ser preservada. Uma das formas de gerenciar este recurso é através do monitoramento de sua qualidade físico-química (BIGOLIN *et al*, 2008).

O monitoramento gera dados que precisam ser armazenados e analisados. Alguns monitoramentos podem gerar rapidamente quantidade de dados na casa de milhares. O volume de dados decorre de três variáveis: a extensão da rede, os parâmetros amostrados e o tempo de operação. Em relação à extensão da rede, quase sempre são amostrados pelo menos dois pontos em cada microbacia do sistema de forma a se ter padrões de comparação. No que se refere ao número de parâmetros, são dependentes do objetivo da rede e do uso do solo, mas sempre envolve variáveis de nível primário, secundário e terciário, no mínimo. O tempo de operação deve ser sempre longo, a fim de permitir avaliações, tendências e alterações nos padrões espaciais e temporais da qualidade das águas (FINOTTI *et al*, 2009).

A avaliação de um recurso hídrico deve contemplar um diagnóstico das condições de entorno, regimes de vazão e chuvas, pontos de lançamento/tipologia de efluentes e sistema de drenagem (urbana/rural). Estes aspectos definem, por sua vez, critérios para a hierarquia fluvial, permitindo uma melhor definição dos pontos e locais de amostragem. Essa hierarquia, como resultado, gera os pontos que devem ser monitorados.

A grande quantidade de pontos e parâmetros monitorados somada a possibilidade de

importação de dados de monitoramento de outras organizações que utilizam o recurso hídrico, traz como consequência um desconhecimento da real qualidade desses dados. Os problemas desses dados podem ser os mais variados, desde confiabilidade, acurácia, representatividade, acessibilidade, entre outros. Os problemas de confiabilidade e acurácia podem ser oriundos do laboratório que efetuou a análise, ou de como foi coletada, armazenada, transportada e conservada a amostra. Além disso, ao passar o dado do laudo analítico para o sistema, o usuário ainda corre o risco de digitar a informação incorreta. Os erros de digitação podem ocorrer em virtude de alguns dados serem descritivos, serem um número ou muito grande, ou muito pequeno.

Segundo Von Sperling (2007), os erros de medição podem ser aleatórios e sistemáticos, e derivam de limitações nos procedimentos de coleta de amostra, análise laboratorial e instrumentos de medição. Estes erros, em maior ou menor escala, são inerentes nas medições das variáveis, e impedem que elas sejam uma representação exata das variáveis de estado.

Contudo, o problema não é exclusivo para a área ambiental ou de recursos hídricos. O problema da qualidade de dados atinge grandes corporações. Segundo (Laudon & Laudon, 2008 apud Forrester Research), “20% das encomendas norte-americanas, despachadas pelo correio ou por empresas de entrega, foram devolvidas por causa de nomes e endereços incorretos. Os consultores da *Gartner Group* relataram que mais de 25 por cento dos dados críticos nos bancos de dados das mil maiores empresas do *ranking* da *Fortune*, estão incorretos ou incompletos; isso inclui erros de descrições e códigos de produtos, descrições de estoque, dados financeiros e informações sobre fornecedores e funcionários”.

A abordagem teórica à qualidade dos dados se concentra em como os dados podem tornar-se deficientes durante o processo de geração ou obtenção dos mesmos. A vantagem de utilizar uma abordagem intuitiva é que cada estudo pode selecionar os atributos mais relevantes para os objetivos particulares deste estudo (WANG et al, 1996).

O presente trabalho, proposto para a melhoria da qualidade de dados, além da abordagem teórica utiliza técnicas da abordagem intuitiva, selecionando as necessidades levantadas junto ao Instituto de Saneamento Ambiental (ISAM).

## 1.1 O PROJETO E O ISAM

A Universidade de Caxias do Sul, atuante em ensino, pesquisa e extensão, tem envidado esforços para a melhoria da qualidade ambiental, tanto internamente quanto para a sociedade como um todo.

O Instituto de Saneamento Ambiental (ISAM), na última década tem sido um dos vetores destes esforços da Universidade. Com um corpo técnico multidisciplinar capacitado técnica e cientificamente, o ISAM desenvolve estudos, pesquisas e atividades de extensão na área ambiental. Colabora ainda com a formação de recursos humanos através do envolvimento de alunos de graduação e pós-graduação em suas atividades. Dentre os objetivos do ISAM destacam-se:

- Contribuir para a melhoria das condições sanitárias e ambientais dos municípios, da região e do país;
- Promover a qualificação de recursos humanos e a socialização do conhecimento na área ambiental;
- Desenvolver, em parceria com organismos públicos e privados, estudos, pesquisas e ações voltados para a prevenção, minimização, tratamento e disposição de resíduos sólidos, líquidos e gasosos;
- Apoiar cursos de graduação, pós-graduação e extensão oferecidos pela Universidade na área sanitária e ambiental.

O projeto Gestão dos Recursos Hídricos na Bacia Hidrográfica do Taquari-Antas, trecho Antas: disponibilidade, qualidade, usos e impactos decorrentes, é um projeto desenvolvido no Instituto e que engloba dentro de si mais 6 subprojetos. O principal objetivo desse projeto é “Avaliar a disponibilidade, usos e qualidade dos recursos hídricos e impactos decorrentes da ocupação da Bacia Hidrográfica Taquari-Antas – Trecho Antas com vistas à modelagem de cenários atuais e futuros como subsídio ao planejamento e à gestão” (ISAM, 2008). Para a tomada de decisão, está sendo desenvolvido um sistema de informação chamado SIA (Sistema de Informação Ambiental). Esse sistema visa dar subsídio para organizar as informações disponíveis e possibilitar o entendimento da região como um todo.

A avaliação da qualidade de dados para sistemas desse porte torna necessário um estudo mais aprofundado das técnicas e controle da qualidade de dados.

## 1.2 OBJETIVOS

O trabalho tem por objetivo desenvolver um sistema para avaliar a confiabilidade dos dados que são coletados em campo, ou gerados por terceiros, no projeto de pesquisa “Gestão dos Recursos Hídricos na Bacia Hidrográfica do Taquari-Antas, trecho Antas: disponibilidade, qualidade, usos e impactos decorrentes”, e dessa forma sugerir a reavaliação ou a refutação de parâmetros ou amostras de qualidade da água.

Dentre os objetivos específicos a serem atingidos estão:

- Definir um processo para a inserção de dados de forma a auxiliar a validação de dados entre eles.
- Criar uma escala numérica de confiabilidade para cada dado e armazenar essa informação associada à ele para obtê-la em consultas posteriores.
- Associar informações descritivas aos dados que possam qualificar os mesmos.
- Definir variáveis externas e geográficas que ajudem a validar ou descartar a informação referente ao recurso hídrico.

## 1.3 ORGANIZAÇÃO DO TRABALHO

O capítulo 2 faz uma revisão sobre as definições de qualidade e como evoluiu a preocupação sobre esta. Apresenta ainda o que já vem sendo estudado em termos de qualidade de *software* e de dados.

O capítulo 3 traz a justificativa da escolha das variáveis de qualidade da água, bem como algumas definições e explicações, que serão avaliadas neste trabalho.

O capítulo 4 apresenta algumas soluções propostas utilizadas para sustentar o sistema de avaliação de qualidade de dados, baseado nas 16 dimensões de qualidade de dados.

O capítulo 5 apresenta o desenvolvimento do sistema, as suas funcionalidades e descreve detalhadamente o processo que o mesmo faz para avaliar a qualidade de cada parâmetro inserido no sistema.

O capítulo 6 apresenta como foi testado e implantado o sistema, assim como demonstra a opinião dos usuários técnicos do sistema.

O capítulo 7 apresenta uma discussão final sobre o trabalho, assim como as tarefas futuras que podem ser feitas para a continuação desse trabalho.



## 2 QUALIDADE

A palavra qualidade é originária do latim "*qualitate*" e significa propriedade, característica. Segundo o dicionário Aurélio(2004): “qualidade é propriedade, atributo ou condição das coisas ou das pessoas, que as distingue das outras e lhes determina a natureza. Ou também pode ser entendido por superioridade, excelência de alguém ou de algo”.

Uma outra definição para qualidade é a apresentada por Montgomery, (2004): “Qualidade é inversamente proporcional à variabilidade”. Porém, aplicada a dados de qualidade da água, a variação pode significar uma nova tendência, ou alguma modificação no meio. Especialmente em um recurso hídrico pode mostrar claramente uma ação antrópica, causando poluição ou então alguma alteração natural no ecossistema.

Laudon (2007) traz a definição de qualidade de um produto em geral através de duas perspectivas: na perspectiva do produtor – qualidade significa conformidade a determinadas especificações, ou ausência de variação em relação a essas especificações. Na perspectiva do consumidor suas preocupações são em relação a durabilidade, segurança, facilidade de uso e a qualidade dos serviços como garantia e suporte ao produto.

Para uma definição precisa de qualidade é necessário uma análise de qual contexto ela está inserida, ou a qualidade pode ser definida de acordo com os requisitos (CROSBY, 1979). Ainda é preciso ressaltar que as exigências e requisitos de qualidade mudam através do tempo, e as soluções e metodologias adotadas anteriormente podem também necessitar de mudanças. Entretanto, todo o investimento feito na busca da qualidade de dados tem o retorno garantido, sendo assim, a qualidade não é um problema e sim uma solução, uma solução à necessidade permanente de redução de custos e de ganho de competitividade (PINHO, 2001).

Acerca de Qualidade de *Software*, tratado no item 2.1, são feitos levantamentos sobre uma possível avaliação do produto de *software* comparado ao “produto dado”. Por último, o item 2.2 Qualidade de Dados, apresenta algumas técnicas de abordagem teórica sobre qualidade de dados que será aplicado no trabalho.

### 2.1 QUALIDADE DE SOFTWARE

Para avaliação da qualidade de dados não basta apenas estudar o dado e as suas relações com o contexto, é necessário que o processo de coleta de informação e o sistema de

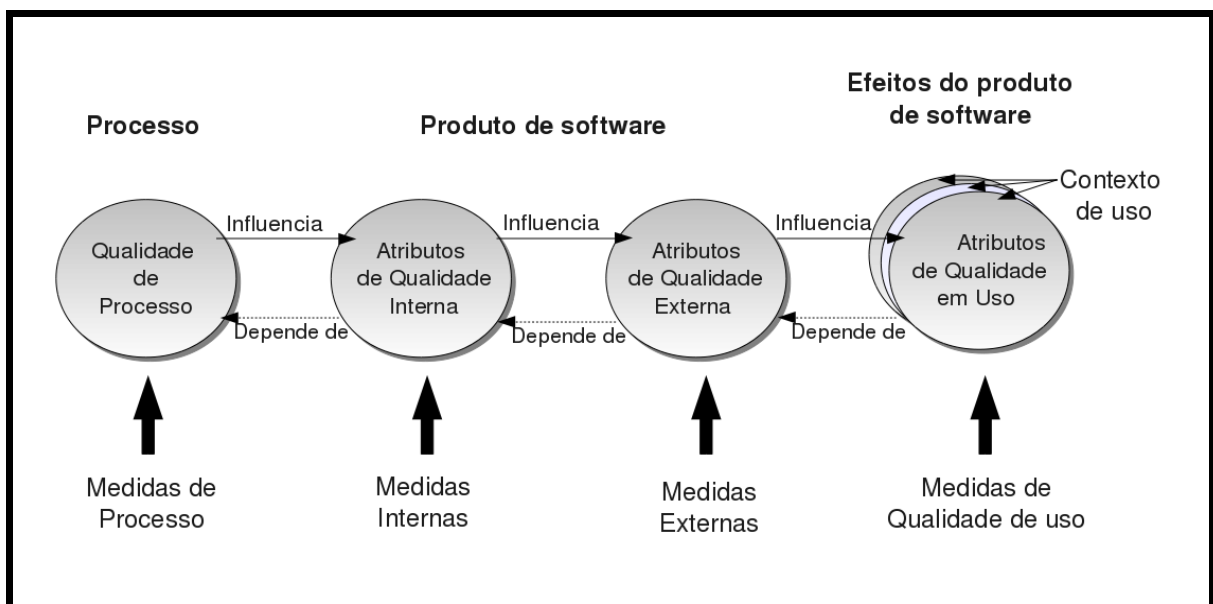
informação também tenha como meta a qualidade. Por isso, avaliar e melhorar o processo é um meio de melhorar a qualidade do produto, assim como avaliar e melhorar a qualidade do produto é um meio de melhorar a qualidade em uso (NBR ISO/IEC 9126-1, 2003).

Um dos estudos realizados em engenharia de *software* é o estudo da qualidade do *software*. A qualidade de *software* é definida por Pressman (2006) como: “Conformidade a requisitos funcionais e de desempenho explicitamente declarados, a padrões de desenvolvimento claramente documentados e a características implícitas que são esperadas de todo *software* profissionalmente desenvolvido”.

Segundo a norma NBR ISO/IEC 9126 a avaliação de produtos de *software* com o objetivo de satisfazer as necessidades de qualidade de *software* é um dos processos no ciclo de vida de desenvolvimento de *software*. A qualidade do produto de *software* pode ser avaliada medindo-se os atributos internos, os atributos externos, ou os atributos de qualidade em uso.

Portanto, o principal objetivo é que o produto (*Software*) tenha o efeito requerido num contexto de uso particular, ou seja, atenda as necessidades do cliente e atenda os objetivos do produtor do *software*.

A Ilustração 1 mostra a qualidade, requerida como produto final, de um *software* durante o seu ciclo de vida.



**Ilustração 1 - Qualidade no ciclo de vida do Software**

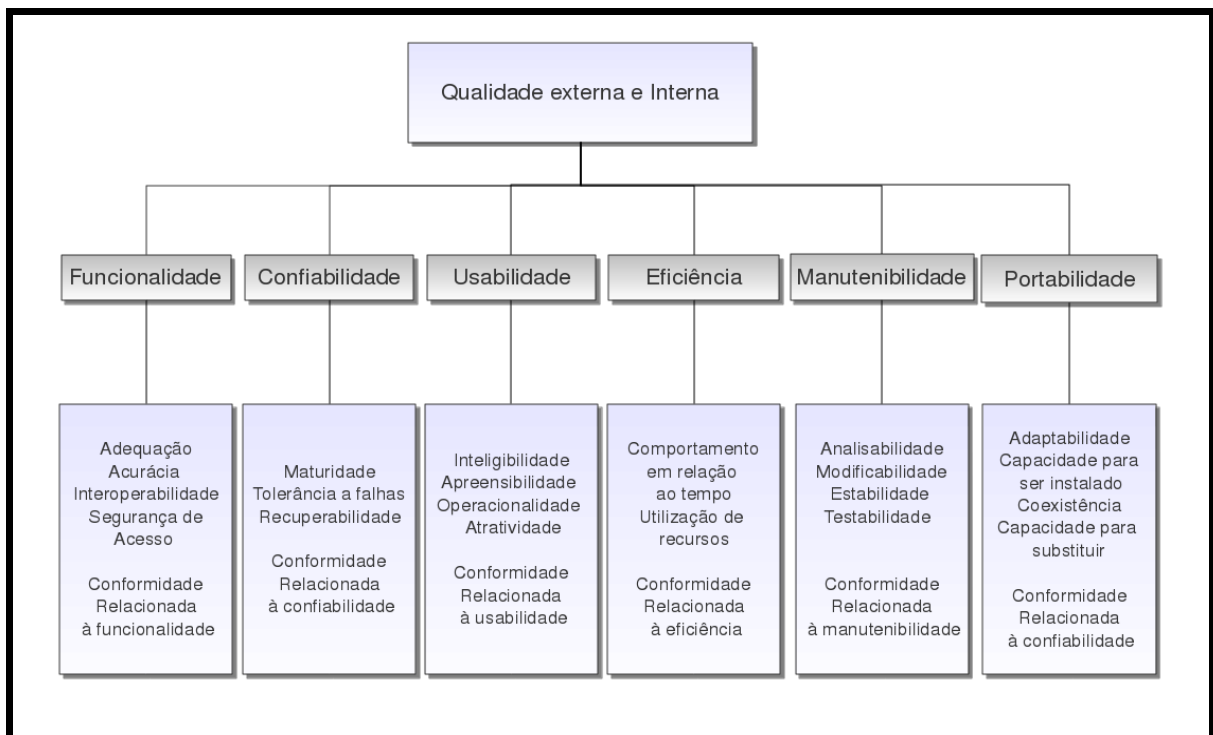
Fonte: Adaptado da (NBR ISO/IEC 9126-1, 2003)

A Qualidade Interna é a totalidade das características do produto de *software* do ponto de vista interno: seus programas, sua documentação e dados. Detalhes da qualidade do

produto de *software* podem ser melhorados durante a implementação do código, revisão e teste. Contudo, a natureza fundamental da qualidade do produto de *software* representada pela qualidade interna mantém-se inalterada, a menos que o projeto seja refeito (NBR ISO/IEC 9126-1, 2003).

A Qualidade externa é a totalidade das características do produto de *software* do ponto de vista externo. É a qualidade quando o *software* é executado, o qual é tipicamente medido e avaliado enquanto está sendo testado num ambiente simulado, com dados simulados e usando métricas externas (NBR ISO/IEC 9126-1, 2003) .

Para os atributos de qualidade interna e externa a norma propõe a hierarquia demonstrada na Ilustração 2.

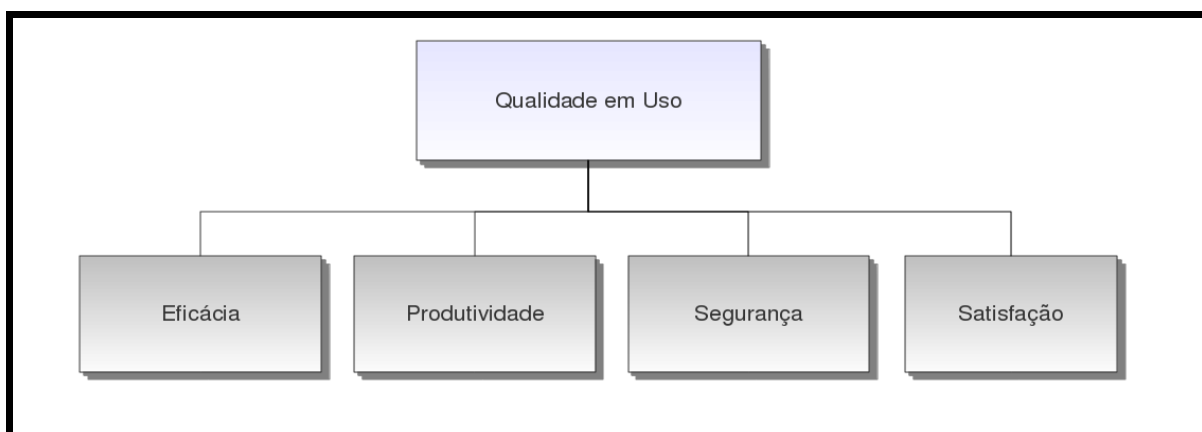


**Ilustração 2 - Modelo de qualidade para a qualidade externa e interna**

Fonte: Adaptado da (NBR ISO/IEC 9126-1, 2003)

A Qualidade em Uso é a visão da qualidade do produto de *software* do ponto de vista do usuário, quando este produto é usado em um ambiente e um contexto de uso especificados. Ela mede o quanto os usuários podem atingir seus objetivos num determinado ambiente, e não as propriedades do *software* em si. Qualidade em uso é, para o usuário, o efeito combinado das seis características de qualidade do produto de *software*: Funcionalidade, confiabilidade, usabilidade, eficiência, manutenibilidade e portabilidade (NBR ISO/IEC 9126-1, 2003) .

Os atributos de qualidade em uso, segundo a norma, são categorizados em quatro características: eficácia, produtividade, segurança e satisfação como mostra a Ilustração 3.



**Ilustração 3 - Modelo de qualidade para qualidade em uso**

Fonte: Adaptado da (NBR ISO/IEC 9126-1, 2003)

Avaliando a qualidade do dado e tratando-o como um produto, pode-se fazer um paralelo a qualidade de *software*: a qualidade em uso de um dado é totalmente diferente na sua qualidade externa ou interna. Classificando um dado nessas três categorias, ele pode ter uma boa qualidade para o usuário ou para o tomador de decisão, mas pode ser extremamente complexo de manipular (Qualidade Externa), ou ser um dado coletado com um processo ruim, e/ou ser incorreto (Qualidade Interna).

## 2.2 QUALIDADE DE DADOS

De forma geral, definir qualidade de dados é um conceito complexo porque possui significados diversos para diferentes pessoas. Uma das razões para isso é o caráter essencialmente subjetivo da qualidade de dados, cuja avaliação pode variar de acordo com a função do observador, do contexto e dos objetivos da avaliação. Além disso, muitas vezes, a qualidade não pode ser expressa quantitativamente, mas precisa ser descrita (Amaral, 2003).

Na NM-ISO 8402 (1997), qualidade de dados é definida como a “totalidade de características que confere ao dado a capacidade de representar o mais fielmente possível a informação, atendendo os requisitos implícitos e explícitos”. Segundo *U. S. Environmental Protection Agency (1997)*, qualidade de dados é a “totalidade das características dos dados que possui a capacidade de satisfazer um determinado propósito, a soma dos graus de excelência para os fatores relacionados aos dados”.

A ISO (*International Organization for Standardization*) publicou em 5 de novembro do ano de 2009 a ISO 8000-110, que trata da especificação de troca de dados característicos: sintaxe, codificação semântica e de conformidade dos dados (ISO 8000-110, 2009). Já a ISO

8000-120, de 23 de junho do ano de 2009, inclui um modelo conceitual em UML (*Unified Modeling Language*) da proveniência dos dados, ajudando o usuário final a avaliar a confiabilidade de dados e fazer uma escolha quando os valores para a mesma propriedade são recebidos de diferentes fontes (ISO 8000-110, 2010).

Para o processo de tomada de decisões, realização de projeções, diagnóstico da realidade ou entendimento do processo histórico-evolucionário, independentemente da área de conhecimento ou aplicação científica ou comercial, são necessário dados. Os dados gerados ou coletados para estas tarefas precisam possuir uma confiança aceitável. Isso quer dizer que, dados de baixa qualidade podem causar prejuízos tanto financeiros quanto em tempo hábil gasto na avaliação das informações geradas por esses dados.

Na seção 2.3.1 Características de Qualidade de Dados serão definidos adjetivos que um dado pode receber e como eles são classificados.

### 2.2.1 CARACTERÍSTICAS DE QUALIDADE DE DADOS

A escolha das características que vão moldar a avaliação da qualidade de um dado normalmente é uma tarefa árdua. Um dado de qualidade pode ser definido como aquele que “satisfaz um determinado propósito”, se avaliado através de uma definição de qualidade mais genérica como Crosby (1979).

Por sua vez, “satisfazer a um propósito” não implica que aquele dado possua uma boa qualidade. Isso porque, na geração de informações a partir de dados que apenas “satisfazem a um propósito”, normalmente há como consequência um acúmulo de erros, gerando tomadas de decisões incoerentes. Para a redução desses erros é necessário a avaliação de algumas características genéricas. Essas avaliações, por sua vez, podem ser aplicadas a qualquer dado que faça parte de um sistema. São exemplo de características genéricas a confiabilidade do dado, a clareza, a atualidade, a disponibilidade, entre outras.

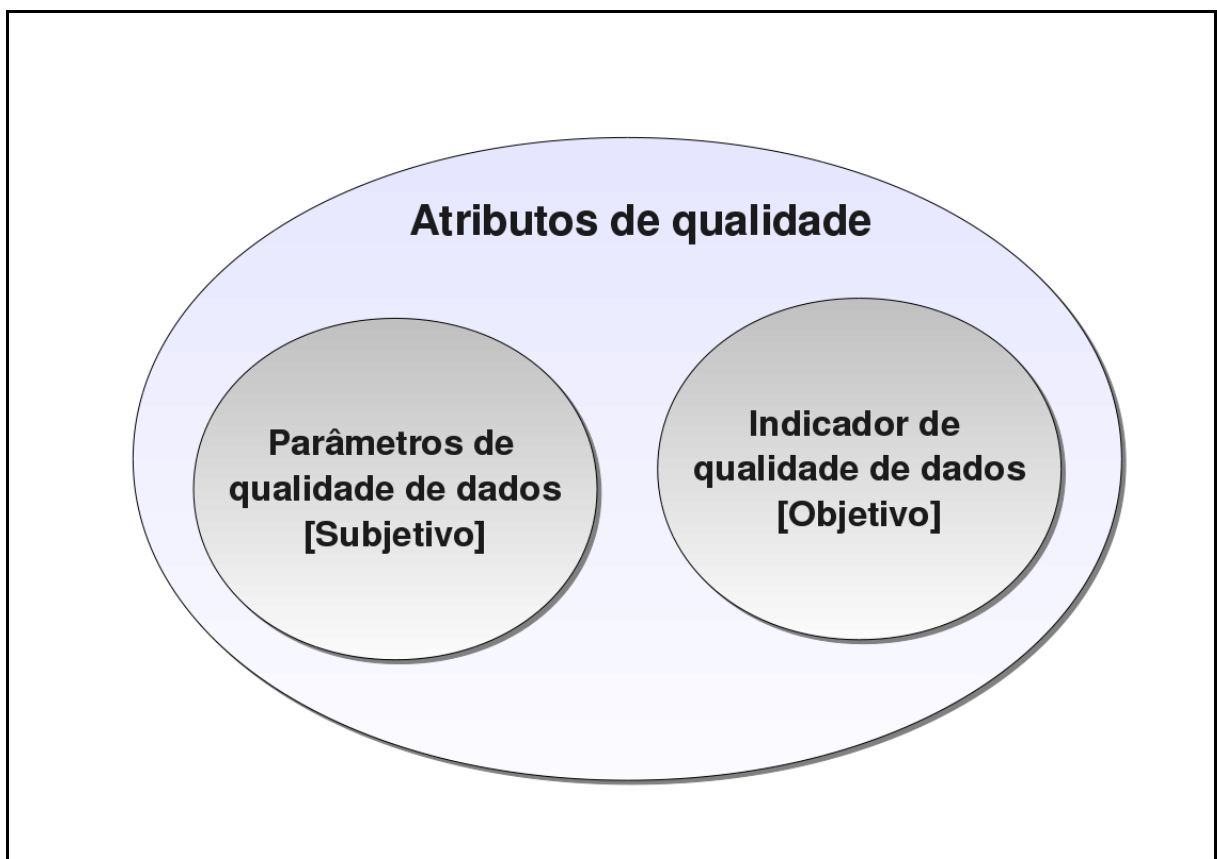
Para operacionalizar as características de um dado Wang (1993) propôs uma classificação das mesmas em grupos, como mostra a Tabela 1. Como resultado foram criados dois grandes grupos: parâmetros e indicadores, sendo que os elementos desses grupos fazem parte de um grupo maior denominado atributos (Ilustração 4).

**Tabela 1 - Classificação das características de qualidade de dados**

Classificação	Descrição	Exemplos
---------------	-----------	----------

Parâmetro de qualidade de dados	Informação qualitativa ou subjetiva pelo qual um usuário avalia a qualidade dos dados	Credibilidade e atualidade
Indicador de qualidade de dados	Informação objetiva sobre os dados	Hora da criação, e método de coleta
Atributo de qualidade de dados	Grupo incluindo os parâmetros de qualidade e indicadores de qualidade	

Fonte: Adaptado de (WANG *et al*, 1993)



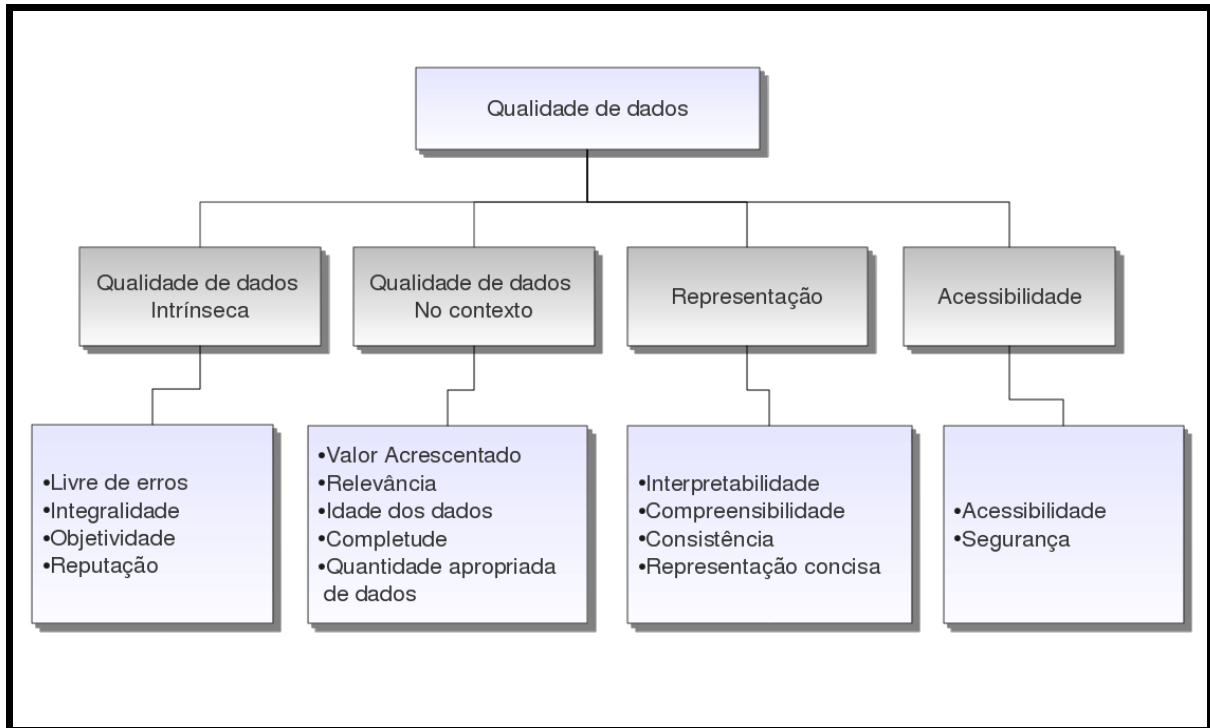
**Ilustração 4 - Relação entre atributos, parâmetros e indicadores de qualidade de dados**

Fonte: Adaptado de (WANG *et al*, 1993)

Obter a lista completa dos atributos para qualificar um determinado dado (devido variação da necessidade do sistema) e o número mínimo de atributos que caracterizam o dado, torna-se uma das tarefas mais complexas no processo de melhoria da qualidade dos dados (ABATE e DIEGERT, 1998).

Ao invés de se avaliar a qualidade de dados através de atributos pode ser realizada a avaliação em conjunto de atributos, que são chamadas por dimensões de qualidade de dados (WANG *et al*, 1996; ABATE e DIEGERT, 1998; PIPINO *et al*, 2002).

A equipe de Wang (1996) em sua primeira pesquisa quanto aos atributos de qualidade de dados elegeu 118 atributos conforme o Anexo I. Esse estudo resultou em 15 dimensões que logo depois foram hierarquizadas conforme mostra a Ilustração 5.



**Ilustração 5 - Organograma da hierarquia de qualidade de dados**

Fonte: Adaptado de (WANG et al, 1993)

A qualidade de dados intrínseca denota que os dados tem qualidade em seu próprio direito. Já a qualidade de dados no contexto denota as exigências que a tarefa que esta sendo executada requer dos dados. A representação por sua vez é um conjunto de atributos que qualificam o dado perante a sua facilidade de interpretação e as suas formas de exibição enquanto que a acessibilidade é o conjunto de atributos que qualificam os dados quanto a sua disponibilidade, sua segurança e privacidade (WANG *et al*, 1993).

A Tabela 2 apresenta as dimensões propostas por Wang (1993) com a adição da dimensão de facilidade de manipulação (PIPINO *et al*, 2002), que é um atributo de quão fácil é realizar trabalhos com esses dados. A dimensão “Fácil manipulação” foi então classificada junto à representação da informação.

Para algumas das dimensões apresentadas na tabela foram colocados sinônimos que outros autores ou obras de qualidade de dados normalmente se referenciam (WANG *et al*, 1993; ABATE e DIEGERT, 1998; PIPINO *et al*, 2002; CHOI *et al*, 2008).

**Tabela 2 - Dimensões de qualidade de dados**Fonte: Adaptado de (PIPINO *et al*, 2002)

<b>Dimensão</b>	<b>Descrição</b>
<b><i>Qualidade de dados Intrínseca</i></b>	
Livre de erros (Acurácia)	Medida em que os dados estão corretos e confiáveis
Credibilidade / Integralidade	A medida em que os dados são considerados como verdadeiros e credíveis
Objetividade	Medida da imparcialidade dos dados, sem ambiguidades
Reputação	A medida em que os dados são altamente considerado em termos de sua origem ou conteúdo
<b><i>Qualidade de dados no Contexto</i></b>	
Valor acrescentado	A medida em que o dado é benéfico e proporciona vantagens da sua utilização
Relevância	A medida em que a informação é aplicável e útil para a tarefa em mãos
Idade dos dados	O dado é suficientemente novo para a tarefa em mãos
Totalidade / Completude	Medida onde os dados não faltantes é de profundidade e largura suficiente para a tarefa em mãos
Quantidade apropriada de dados	Volume de dados correto para a tarefa que está sendo executada.
<b><i>Representação</i></b>	
Interpretabilidade	Medida em que os dados estão na língua apropriada, símbolos e unidades, e as definições são claras
Compreensibilidade	Medida da facilidade de compreensão de um dado



<b>Dimensão</b>	<b>Descrição</b>
Representação consistente	Medida onde os dados são representados sempre no mesmo formato.
Representação concisa	Medida onde os dados são representados de forma compacta e precisa.
Fácil de manipular	Medida em que os dados são fáceis de manipular e aplicar a diferentes tarefas
<b><i>Acessibilidade</i></b>	
Acessibilidade	A medida da disponibilidade de dados ou quanto facilmente e rapidamente recuperável
Segurança	Medida onde o acesso aos dados é segura e permitida e a adesão de dados mantem a mesma segurança

As dimensões de qualidade facilitam uma auditoria de qualidade de dados por orientar uma avaliação. Como metas, essas 16 dimensões podem ser complicadas de se alcançar, ou podem causar alguns problemas na hora de implementá-las. O Capítulo 4 apresenta as soluções obtidas para alcançar algumas dessas metas e melhorar a qualidade dos dados.

### 3 PARÂMETROS INDICADORES DA QUALIDADE DA ÁGUA

Para avaliar a qualidade de qualquer produto utilizam-se parâmetros ou dimensões. A própria qualidade de *software* e de dados também utilizam parâmetros e atributos que a qualificam. Para avaliar a qualidade de um recurso hídrico também são necessários parâmetros ambientais (físico-químicos e biológicos).

Segundo a Resolução do Conselho Nacional de Meio Ambiente (CONAMA) n. 357/2005 as águas são classificadas inicialmente em função de sua salinidade, como mostra a Tabela 3.

**Tabela 3 - Classificação das águas**

Fonte: Adaptado da Resolução CONAMA nº357/2005

Classificação	Salinidade
Doce	Menor 0,5%
Salobra	Entre 0,5% e 30%
Salgada	Maior ou igual a 30%

Essa classificação macro se subdivide em outras 13 classes, sendo que para a água doce, retratada neste trabalho, são propostas 5 classes. Essas classes ditam sobre a destinação e utilização da água. A classe do corpo hídrico é importante para a tomada de decisão e averiguação dos parâmetros medidos, além de determinar limites superiores e inferiores do dado medido. Portanto, uma das necessidades decorrentes é que o sistema gere automaticamente a classificação do corpo hídrico.

O Anexo II apresenta uma tabela com um comparativo dos limites dos parâmetros em cada uma das classes. A Tabela 4 apresenta a destinação possível para a água em cada uma das classes propostas pela resolução CONAMA nº357/2005 .

**Tabela 4 - Classes de qualidade de água doce**

Fonte: Adaptado da Resolução CONAMA nº357/2005

<b>Classe</b>	<b>Pode ser destinada</b>
Classe Especial	<ul style="list-style-type: none"> <li>a) ao abastecimento para consumo humano, com desinfecção;</li> <li>b) à preservação do equilíbrio natural das comunidades aquáticas;</li> <li>c) à preservação dos ambientes aquáticos em unidades de conservação de proteção integral.</li> </ul>
Classe 1	<ul style="list-style-type: none"> <li>a) ao abastecimento para consumo humano, após tratamento simplificado;</li> <li>b) à proteção das comunidades aquáticas;</li> <li>c) à recreação de contato primário, tais como natação, esqui aquático e mergulho, conforme</li> <li>d) à irrigação de hortaliças que são consumidas cruas e de frutas que se desenvolvam rentes ao solo e que sejam ingeridas cruas sem remoção de película;</li> <li>e) à proteção das comunidades aquáticas em Terras Indígenas.</li> </ul>
Classe 2	<ul style="list-style-type: none"> <li>a) ao abastecimento para consumo humano, após tratamento convencional;</li> <li>b) à proteção das comunidades aquáticas;</li> <li>c) à recreação de contato primário, tais como natação, esqui aquático e mergulho;</li> <li>d) à irrigação de hortaliças, plantas frutíferas e de parques, jardins, campos de esporte e lazer, com os quais o público possa vir a ter contato direto;</li> <li>e) à aquicultura e à atividade de pesca.</li> </ul>
Classe 3	<ul style="list-style-type: none"> <li>a) ao abastecimento para consumo humano, após tratamento convencional ou avançado;</li> <li>b) à irrigação de culturas arbóreas, cerealíferas e forrageiras;</li> <li>c) à pesca amadora;</li> <li>d) à recreação de contato secundário;</li> <li>e) à dessedentação de animais.</li> </ul>
Classe 4	<ul style="list-style-type: none"> <li>a) à navegação;</li> <li>b) à harmonia paisagística.</li> </ul>

Diversos parâmetros indicadores da qualidade da água podem ser avaliados. Esses parâmetros são indicadores da situação do recurso hídrico. O Anexo III contém uma lista de parâmetros assim como a unidade padrão. Essa lista foi montada utilizando-se a resolução

CONAMA 357/2005 e a norma da Associação Brasileira de Normas Técnicas - ABNT n. 9897 de 1987. A tomada de decisão referente a um recurso hídrico é baseada nos dados coletados em campo, analisados em laboratório e de acordo com os instrumentos legais, resolutivos e normativos vigentes.

Ao serem medidos, os parâmetros podem sofrer influência direta do equipamento, de sua calibragem, da metodologia de coleta da amostra e análises laboratoriais, podendo resultar em erros analíticos. Esses dados, por vezes, podem retratar algo distante da realidade. Nesse sentido o estudo da qualidade de dados se faz necessário quando se busca dar maior fidedignidade aos diagnósticos, considerando-se que a real situação de um corpo hídrico lótico - ambiente relativo a águas continentais moventes (rios e córregos) - ou lênticos (ambiente que se refere a água parada, com movimento lento ou estagnado - lagos e açudes) (CONAMA 357, 2005), só poderá ser revelada por séries de dados temporais. Dessa forma, é importante e útil o desenvolvimento de sistemas que possam armazenar e avaliar possíveis erros.

Por sua vez, corpos hídricos que possuam barramentos (hidrelétricas ou represas para abastecimento, ou piscicultura) podem apresentar diferentes comportamentos (VON SPERLING, 2007):

- Comportamento lótico: rio corrente nos pontos a montante (antes do barramento) e jusante (após o barramento).
- Comportamento lêntico: comportamento de lago, no próprio reservatório. É importante ressaltar que a dinâmica de um lago é diferente de um rio no que diz respeito aos parâmetros físico-químicos e bióticos em função do processo de estagnação a que o sistema está submetido, influenciando principalmente na dinâmica do ciclo do oxigênio e por consequência todas as reações de oxi-redução que ocorrem ou possam vir a ocorrer no corpo hídrico. Estas mudanças, por se revelarem como fatores limitantes, determinarão, por sua vez, mudanças nas comunidades bióticas, alterando inclusive a cadeia alimentar (ESTEVEZ, 1998; VON SPERLING, 2007).

Neste capítulo, no item 3.1, será apresentada a justificativa da escolha dos parâmetros, e, nos seus subitens, serão descritos os parâmetros escolhidos. Já o item 3.2 apresenta outros parâmetros que poderão ser trabalhados na sequência.

### 3.1 ESCOLHA DOS PARÂMETROS

Para o trabalho foram escolhidos alguns parâmetros que, devido ao seu potencial para classificar o meio, devem ser analisados com mais cuidado para não se realizar um diagnóstico incorreto sobre o recurso hídrico. Esses parâmetros por si só não conseguem indicar a qualidade de um recurso hídrico tanto para o consumo humano, como para a manutenção do ecossistema aquático.

Para que possa abranger mais parâmetros futuramente, o sistema foi desenvolvido de forma a utilizar uma tabela para armazenar parâmetros e um sistema de regras dinâmico. Neste trabalho, foram escolhidos 6 parâmetros para serem analisados e modelados. São eles: pH, oxigênio dissolvido, Temperatura, Condutividade Elétrica, Demanda Bioquímica de Oxigênio (DBO) e Demanda Química de Oxigênio (DQO). Esses parâmetros servirão de base para a modelagem e a criação de regras para validação de outros parâmetros. Estes parâmetros, assim como os outros, necessitam de outras variáveis ambientais para que possam ser validados. Estas variáveis externas necessárias para a validação poderão ser elencadas e monitoradas para um conhecimento melhor do meio.

#### 3.1.1 PH

O potencial hidrogeniônico (pH) consiste na concentração dos íons  $H^+$  e  $OH^-$  (hidroxila) nas águas e representa a intensidade das condições ácidas ou alcalinas do ambiente aquático (LIBÂNIO, 2005). Por influir em diversos equilíbrios químicos que ocorrem naturalmente ou em processos unitários de tratamento de águas, o pH é um parâmetro importante em muitos estudos no campo do saneamento ambiental (CETESB, 2001b).

A influência do pH sobre os ecossistemas aquáticos naturais dá-se diretamente devido a seus efeitos sobre a fisiologia das diversas espécies. Também o efeito indireto é muito importante podendo, determinadas condições de pH contribuir para a precipitação de elementos químicos tóxicos como metais pesados. Outras condições podem exercer efeitos sobre as solubilidades de nutrientes (CETESB, 2001b).

O pH pode ser considerado como uma das variáveis ambientais mais importantes, ao mesmo tempo que uma das mais difíceis de se interpretar. A complexidade dessa interpretação deve-se ao grande número de fatores que podem influenciar no seu resultado (ESTEVEZ, 1998).

O pH pode ser determinado pelo método potenciométrico (eletrométrico) utilizando eletrodo de vidro. Para garantir a precisão da medida é importante verificar e proceder calibração periódica do equipamento usado na medição (FINOTTI *et al*, 2009).

Os limites do pH dentro de um sistema dulcícola, segundo a resolução CONAMA 357/2005, são mostrados na Tabela 5.

**Tabela 5 - Limites do pH em água doce**

Fonte: Adaptado da Resolução CONAMA nº357/2005

Classificação	Limite
Classe 1,2, 3 e 4	6-9

### 3.1.2 OXIGÊNIO DISSOLVIDO

O Oxigênio (O<sub>2</sub>) é um dos mais importantes fatores ambientais na dinâmica e na caracterização de ecossistemas aquáticos (ESTEVES, 1998). Para Libânio (2005) é reconhecidamente o parâmetro mais importante para expressar a qualidade de um ambiente aquático. Sob este aspecto, águas poluídas são aquelas que apresentam baixa concentração de oxigênio dissolvido (devido ao seu consumo na decomposição de compostos orgânicos), enquanto que as águas limpas apresentam concentrações de oxigênio dissolvido elevadas, chegando até a um pouco abaixo da concentração de saturação. No entanto, uma água eutrofizada (grande quantidade de biomassa com aumento significativo de fotossíntese) pode apresentar concentrações de oxigênio bem superiores a 10 mg/L, mesmo em temperaturas superiores a 20°C, caracterizando uma situação de supersaturação. Isto ocorre principalmente em lagos de baixa velocidade onde chegam a se formar crostas verdes de algas à superfície (CETESB, 2001b).

O IQA (Índice de Qualidade da Água), criado pela *National Sanitation Foundation* (NSF) dos Estados Unidos, é um indicador que por meio das características físicas, químicas e biológicas fornece ao público um balizador da qualidade das águas de um corpo hídrico (LIBÂNIO, 2005). No IQA o Oxigênio dissolvido é o parâmetro que possui o maior peso (CETESB, 2001a).

As principais fontes de oxigênio para a água são a atmosfera e a fotossíntese. Por outro lado, as perdas são o consumo pela decomposição de matéria orgânica (oxidação), perdas para

a atmosfera, respiração de organismos aquáticos e oxidação de íons metálicos como, por exemplo, o ferro e o manganês (ESTEVES, 1998).

As principais consequências da redução na concentração de oxigênio dissolvido no meio são: mortandade de organismos aeróbios, solubilização de compostos químicos, aumento da toxicidade, geração de maus odores e inconvenientes estéticos (FINOTTI *et al*, 2009).

Os limites do Oxigênio Dissolvido dentro de um sistema dulcícola, segundo a resolução CONAMA 357/2005, são mostrados na Tabela 6.

**Tabela 6 - Limites do Oxigênio Dissolvido em águas doces**

Fonte: Adaptado da Resolução CONAMA nº357/2005

<b>Classificação</b>	<b>Limite (mg/L O<sub>2</sub>)</b>
Classe 1	Não inferior a 6
Classe 2	Não inferior a 5
Classe 3	Não inferior a 4
Classe 4	Não inferior a 2

### 3.1.3 TEMPERATURA

A temperatura da água e dos fluidos em geral indica a magnitude da energia cinética do movimento aleatório das moléculas e sintetiza o fenômeno de transferência de calor à massa líquida (LIBÂNIO, 2005).

Variações de temperatura são parte do regime climático normal, e corpos de água naturais apresentam variações sazonais e diurnas, bem como estratificação vertical. A temperatura superficial é influenciada por fatores tais como latitude, altitude, estação do ano, período do dia, taxa de fluxo e profundidade. A elevação da temperatura em um corpo d'água geralmente é provocada por despejos industriais (indústrias canavieiras, por exemplo) e usinas termoelétricas (CETESB, 2001).

A temperatura desempenha um papel importante no controle no meio aquático, condicionando as influências de uma série de parâmetros físico-químicos. Em geral, à medida

que a temperatura aumenta, de 0 a 30°C, a viscosidade, tensão superficial, compressibilidade, calor específico, constante de ionização e calor latente de vaporização diminuem, enquanto a condutividade térmica e a pressão de vapor aumentam as solubilidades com a elevação da temperatura. Organismos aquáticos possuem limites de tolerância térmica superior e inferior, temperaturas ótimas para crescimento, temperatura preferida em gradientes térmicos e limitações de temperatura para migração, desova e incubação do ovo (CETESB, 2001).

O aumento da temperatura das águas também pode aumentar a ação tóxica de muitos elementos e compostos. Esta tem sido a razão pela qual há uma maior mortandade de peixes em águas poluídas durante o verão do que no inverno (MOTA, 1995).

### 3.1.4 CONDUTIVIDADE ELÉTRICA DA ÁGUA

A condutividade é uma expressão numérica da capacidade de uma água conduzir a corrente elétrica. A condutividade elétrica não é um parâmetro integrante do padrão de potabilidade brasileiro (LIBÂNIO, 2005). Porém, depende das concentrações iônicas e da temperatura, e indica a quantidade de sais existentes na coluna d'água. Portanto, representa uma medida indireta da concentração de poluentes. Em geral, níveis superiores a 100  $\mu\text{S}/\text{cm}$  indicam ambientes impactados. (CETESB, 2001).

A condutividade elétrica pode fornecer importantes informações tanto sobre o metabolismo do ecossistema aquático, como sobre os fenômenos importantes que ocorrem na sua bacia de drenagem (ESTEVES, 1998).

Esteves ainda levanta as informações que a condutividade elétrica pode fornecer:

- 1º) Magnitude da concentração dos íons conhecidos como macronutrientes (cálcio, magnésio, potássio, sódio, carbonato, sulfato, cloreto e etc).
- 2º) Variação diária da condutividade elétrica da água fornece: informações a respeito de processos importantes nos ecossistemas aquáticos, como produção primária (redução dos valores) e decomposição (aumento dos valores).
- 3º) Ajuda a detectar fontes poluidoras nos ecossistemas aquáticos.
- 4º) As diferenças geoquímicas nos afluentes do rio principal ou de um lago.

Em águas muito puras a resistência é maior, portanto a condutividade é menor (ESTEVES, 1998). Além da temperatura, o pH da amostra pode ter grande influência sobre os valores de condutividade elétricas, especialmente em águas pobres em sais solúveis e baixos valores de pH (<5) (ESTEVES, 1998).



### 3.1.5 DEMANDA BIOQUÍMICA DE OXIGÊNIO (DBO)

A DBO refere-se a matéria orgânica passível de ser estabilizada biologicamente. (LIBÂNIO, 2005; CETESB, 2001b).

A  $DBO_{5,20}$  é normalmente considerada como a quantidade de oxigênio consumido durante um determinado período de tempo, numa temperatura de incubação específica. Um período de tempo de 5 dias numa temperatura de incubação de 20°C é frequentemente usado e referido como  $DBO_{5,20}$  (CETESB, 2001b). A correlação dele com outros parâmetros (principalmente com o Oxigênio Dissolvido somado ao DQO) ajudam a conseguir o real estado de consumo de oxigênio do corpo hídrico.

O fato desse parâmetro ser medido a cada 5 dias pode gerar um problema de representação, já que o valor obtido dele não se refere ao horário da coleta.

Os limites do  $DBO_{5,20}$  dentro de uma água doce segundo a resolução CONAMA 357/2005 são mostrados na Tabela 7.

**Tabela 7 - Limites do  $DBO_{5,20}$  em água doce**

Fonte: Adaptado da Resolução CONAMA n°357/2005

Classificação	Limite (mg/L O <sub>2</sub> )
Classe 1	Até 3
Classe 2	Até 5
Classe 3	Até 10
Classe 4	-

### 3.1.6 DEMANDA QUÍMICA DE OXIGÊNIO (DQO)

É a quantidade de oxigênio necessária para oxidação da matéria orgânica através de um agente químico. Os valores da DQO normalmente são maiores que os da DBO (CETESB, 2001; LIBÂNIO, 2005).

O resultado da DQO é obtido de uma amostra em até 3 horas, e é bastante utilizado para o monitoramento de estações de tratamento de esgoto (LIBÂNIO, 2005).

A diferença entre DBO e DQO está no tipo de matéria orgânica estabilizada: enquanto

a DBO se refere exclusivamente à matéria orgânica mineralizada por atividade dos micro-organismos. A DQO engloba também a estabilização da matéria orgânica ocorrida por processos químicos (FINOTTI *et al*, 2009).

### 3.2 OUTROS PARÂMETROS DA ÁGUA

Para uma avaliação da qualidade real da água são necessários diversos outros parâmetros fundamentais, como o Fósforo - que é um nutriente essencial para a matéria orgânica e, portanto, fundamental para os organismos vivos. A mesma importância atribui-se a parâmetros como o Nitrogênio, Turbidez, Coliformes fecais e Sólidos totais.

Para a qualidade da água quanto a sua potabilidade um dos indicadores é o IQA. O IQA utiliza no cálculo nove parâmetros para avaliação. Cada um desses parâmetros possui um peso, ao qual é calculado o índice. Os parâmetros e os pesos no cálculo do índice estão apresentados na Tabela 8 (CETESB, 2001a).

**Tabela 8 - Pesos dos parâmetros do cálculo do IQA**

Fonte: (CETESB, 2001a).

Parâmetro	Peso
Oxigênio Dissolvido	0,17
Coliformes Fecais	0,15
pH	0,12
DBO	0,10
Fósforo Total	0,10
Temperatura	0,10
Nitrogênio Total	0,10
Turbidez	0,08
Sólidos Totais	0,08

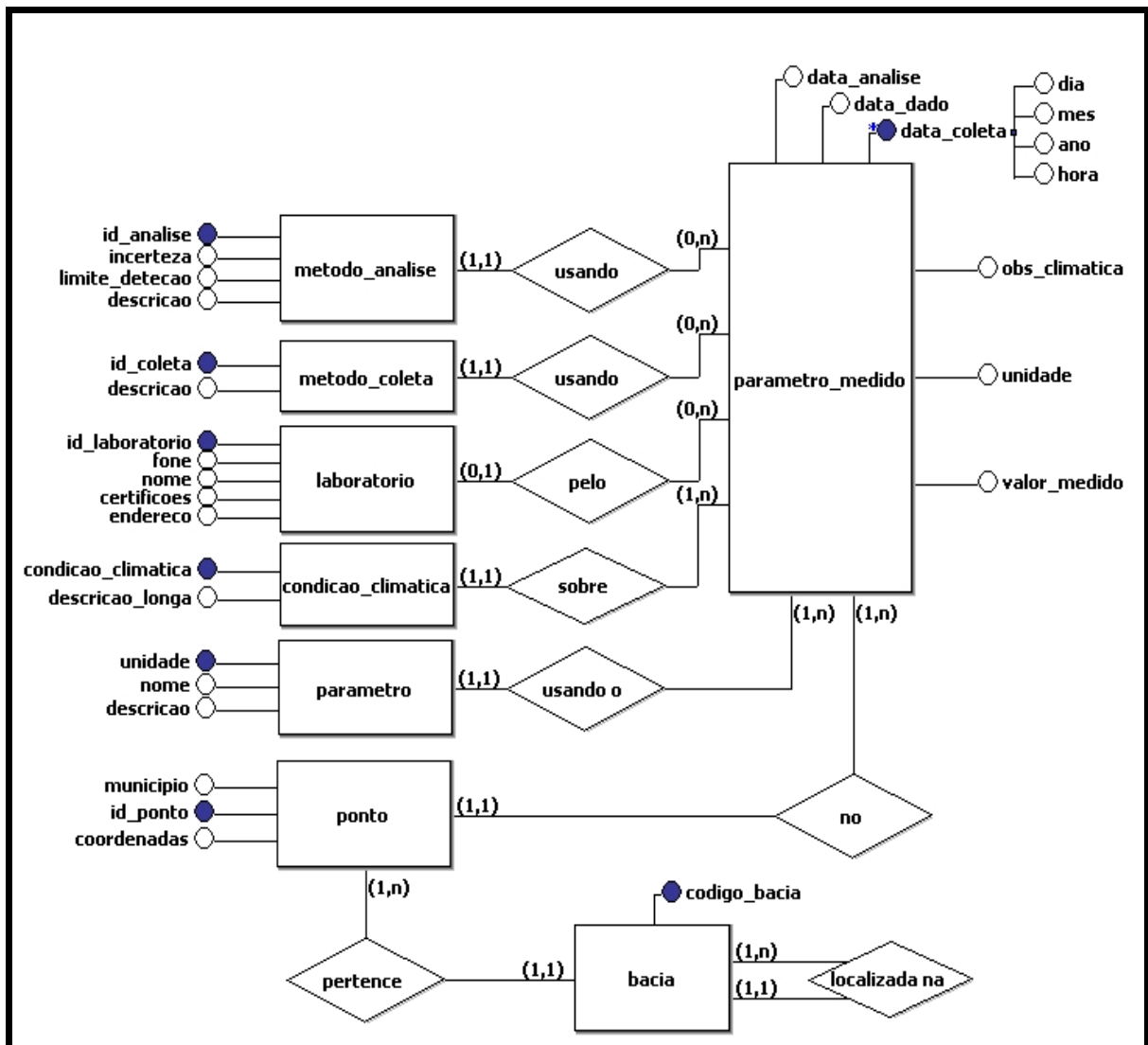
Neste trabalho, será realizada uma análise mais aprofundada dos 6 parâmetros citados nas seções 3.1.1 ao 3.1.6 mas, em continuação, poderá contemplar esses e outros parâmetros físicos, químicos e biológicos da água. Um potencial parâmetro a ser tratado futuramente será Coliformes fecais, que é um parâmetro biológico e possui um grande peso no IQA.

## 4 AS DIMENSÕES DE QUALIDADE DE DADOS APLICADAS AO SISTEMA

Willshire e Meyen em 1997, no artigo “*A Process for Improving Data Quality*”, apresentaram a proposta de um *framework* para a avaliação da qualidade de dados, assim como (WANG *et al*, 1993; PIPINO *et al*, 2002; CHOI *et al*, 2008) apresentam a proposta da avaliação da qualidade dos dados através de técnicas objetivas e outras subjetivas.

Deve-se considerar que os consumidores de dados tem uma melhor ideia da qualidade de dados do que uma empresa ou sistema de validação (WANG *et al*, 1996). Portanto, a avaliação junto ao corpo técnico do ISAM, assim como um constante acompanhamento dos resultados obtidos pelo sistema de avaliação de qualidade dos dados, vai diagnosticar se o sistema está melhorando a qualidade dos dados de forma significativa ou não, e, sendo assim, obtendo uma maior confiabilidade das informações geradas a partir dos dados inseridos na base de informação.

Para a avaliação de qualidade de dados deve-se construir, principalmente, um modelo de dados consistente (CHOI *et al*, 2008). A Ilustração 6 mostra o diagrama ER completo do sistema. A partir desse diagrama será abordado em cada dimensão da qualidade de dados qual as técnicas aplicadas e usadas no modelo de dados que o SIA utilizará.



**Ilustração 6 - Diagrama ER Completo do Sistema**

O modelo de dados apresentado na Ilustração 6 foi montado a partir da análise dos relatórios enviados pelos laboratórios que avaliam os parâmetros físicos e químicos. Alguns dados adicionais foram levantados junto ao corpo técnico do ISAM e respectivamente modelados.

Nesse capítulo serão detalhadas as técnicas utilizadas para obter a melhora da qualidade dos dados em cada uma das 16 dimensões apresentadas por (PIPINO *et al*, 2002).

A organização das seções fica de acordo com a hierarquia proposta no organograma apresentado na Ilustração 5. A seção 4.1 apresenta as dimensões que se enquadram dentro da qualidade intrínseca do dado, ou seja, as dimensões que o dado consegue carregar por si só. A seção 4.2 apresenta a qualidade de dados no contexto.

## 4.1 QUALIDADE DE DADOS INTRÍNSECA

A qualidade de dados intrínseca determina a qualidade que um dado carrega por si só sem a relação desse com o contexto onde o mesmo foi capturado. Estas dimensões por sua vez carregam maior responsabilidade já que todos os dados obrigatoriamente possuem. Dentre as dimensões que essa divisão aborda temos a Acurácia apresentada na seção 4.1.1, a credibilidade apresentada na seção 4.1.2, a objetividade seção 4.1.3 e a reputação seção 4.1.4.

### 4.1.1 ACURÁCIA

Conseguir um sistema onde os dados são livres de erros, se tratando de dados ambientais, é quase uma utopia. Uma das alternativas é a busca contínua da melhora da qualidade dos mesmos. Diferente de um dado de uma venda ou de um estado de algum equipamento ou produto, o dado ambiental sofre influência de diversos fatores externos do ecossistema.

As medições dos parâmetros resultam normalmente em valores numéricos porque em virtude de serem números muito grandes ou muito pequenos ou representarem uma escala, acabam por dificultar o tratamento direto por domínio de dados.

A Tabela 9 mostra uma coleta de dados obtidos de laudos analíticos. Nessa tabela é possível ver, por exemplo, dados como os sólidos suspensos totais, onde ao trocar de laboratório, ao qual possui metodologia diferente (limite de detecção é menor), o dado começou a ser valorado de forma a inutilizá-lo para uma análise. Como as amostras apresentam valores inferiores, não se pode afirmar nada além de que são menores que 10.

**Tabela 9 - Série de amostras de um ponto do Rio São Marcos**

Fonte: ISAM (2009)

Parâmetros	14/12/03	3/12/04	4/8/05	22/3/06	19/9/06	5/11/07*	20/2/08*	12/5/08*	26/8/08*	07/01/09*	04/03/09*	05/05/09*
Temperatura (°C)	19	24	10	22,8	12	15,5	24	12	14	19,3	23,5	16
Condutividade (µS/cm a 20°C)	49,7	142	62,5	134	51,3	39	154	82,7	53,2	91,4	93	229
pH	7	7,1	7	6,9	6,69	7,54	6,6	6,97	6,68	6,43	7,58	6,63
DBO (mg O <sub>2</sub> /L)	2,8	2,2	2	2	6	2	6	1	6	4,7	4,7	6,9
DQO (mg O <sub>2</sub> /L)	69,8	6,81	5	7	18	7	13	5	16	15	12	13
Oxigênio Dissolvido (mgO <sub>2</sub> /L)	7,7	6,9	7	7,3	6,2	9,2	6,1	9,1	7,96	7,38	6,7	6,2

Parâmetros	14/12/03	3/12/04	4/8/05	22/3/06	19/9/06	5/11/07*	20/2/08*	12/5/08*	26/8/08*	07/01/09*	04/03/09*	05/05/09*
Sólidos Suspensos Totais (mg/L)	20	20	11	2	7	<10	<10	<10	<10	<10	<10	<10
Fósforo total (mgP/L)	0,94	0,23	0,83	0,64	0,581	0,5	0,4	0,44	0,38	0,27	0,062	0,948
Nitrogênio total (mgN/L)	1,93	6,31	1,75	1,08	2,95	0,12	1,4	0,74	8,12	<0,5	<0,5	1,36
Coliformes fecais (NMP/100mL)	-	-	-	-	-	70	7	8	2,8x10 <sup>4</sup>	70	7x10 <sup>2</sup>	23

\* Campanha UCS/LASAN

Com o intuito de melhorar a qualidade dos dados (tentando reduzir ao máximo medições incorretas) o sistema para avaliação dos parâmetros e sua relação com o contexto está sendo montado. Porém, nem todo dado que o sistema possa refutar poderá ser considerado como totalmente incorreto. Isso se deve a complexidade entre as correlações ambientais. Para tentar amenizar esse problema, foi montado um fluxo de inserção para os dados. O fluxo para a inserção dos dados está apresentado no diagrama de atividades na Ilustração 7. Este diagrama de atividades vai em encontro aos diagramas de caso de uso apresentados na sessão 5.2 Modelagem do Sistema.

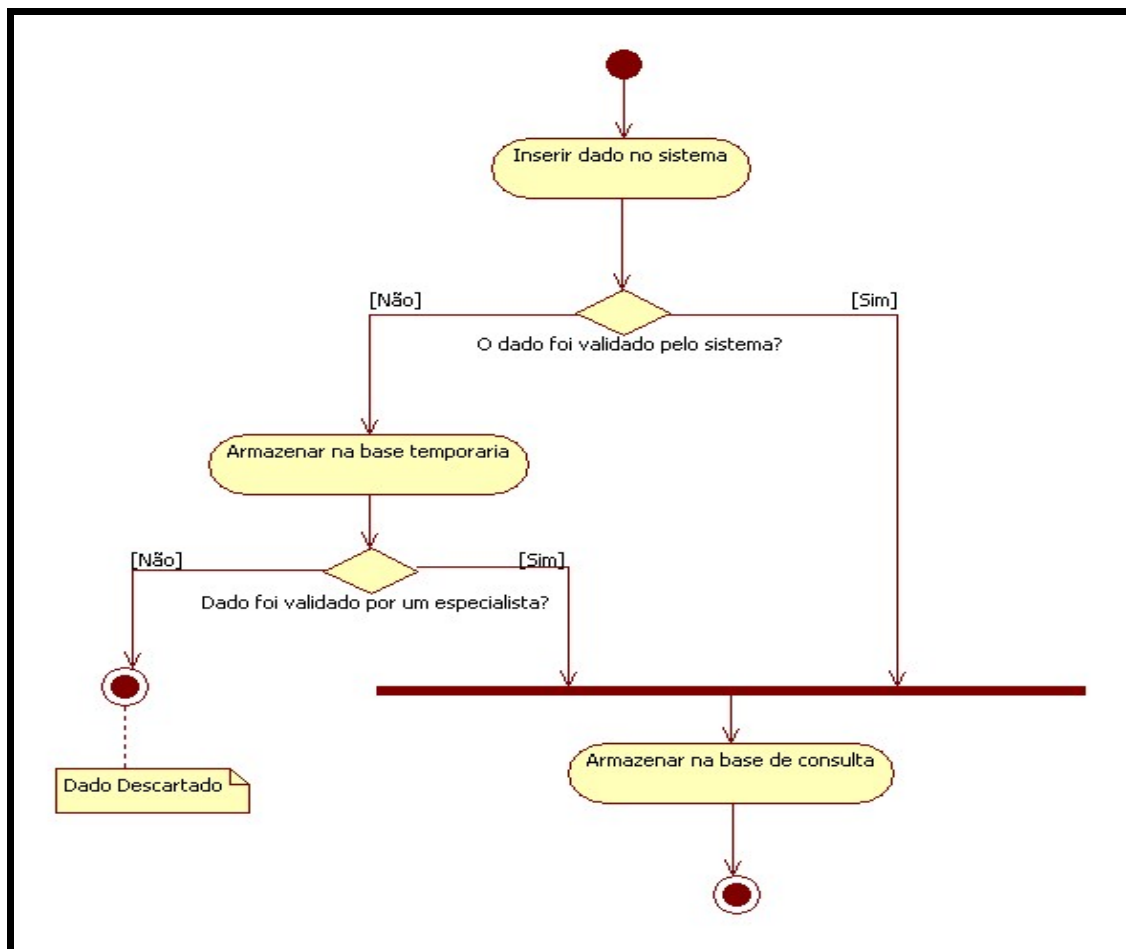


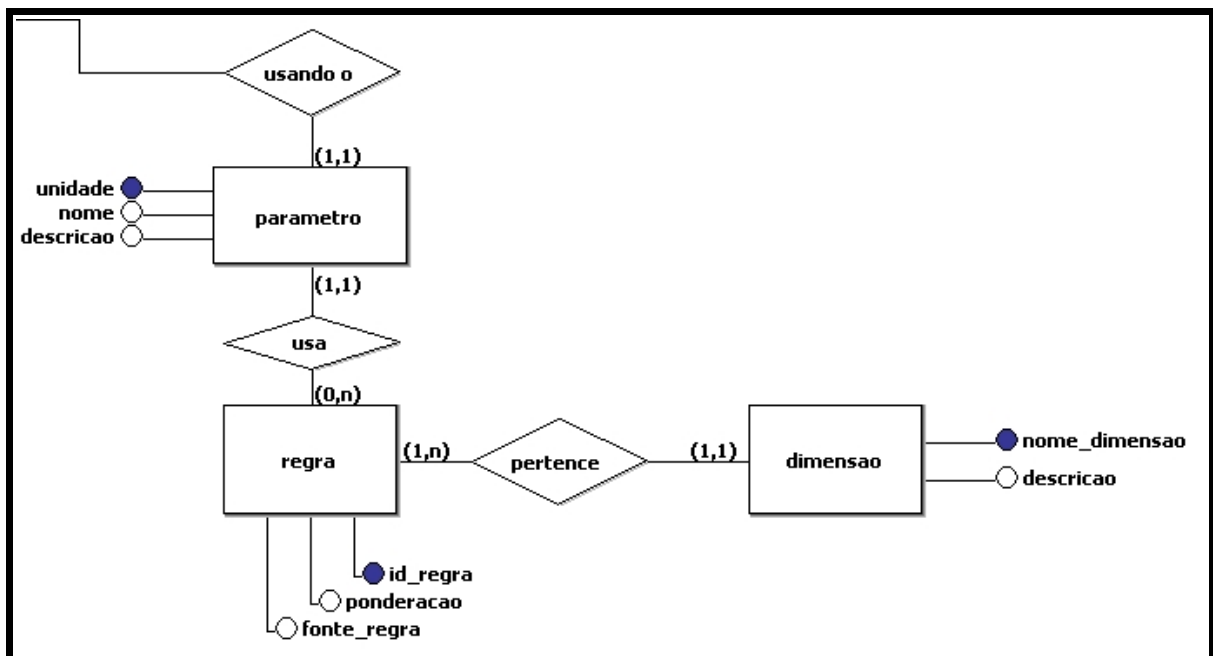
Ilustração 7 - Fluxograma de inserção de dados

Para a detecção de erros nos dados é possível realizar diversos testes inter-

relacionando os dados ou comparando-os com variáveis externas “estáticas”. Outro teste muito comum em qualquer processo de melhoria de qualidade é a detecção de erros utilizando gráficos de controle (MONTGOMERY, 2004).

Todos esses testes podem ser aplicados ao dado que está sendo inserido no momento. A tendência natural do sistema é ir recebendo novos testes e, conseqüentemente, conseguir maior fidedignidade dos mesmos. Para conseguir suprir esse requisito foi pensado em uma base de testes.

A base de testes funciona embasada em regras que poderão ser inseridas conforme a necessidade. Ela é composta por duas tabelas (“regra” e “dimensão”) associadas ao modelo de dados padrão. A Ilustração 8 demonstra as tabelas e suas respectivas relações.



**Ilustração 8 - Modelo de dados para as regras de validação**

O sistema de avaliação, portanto, poderá ir recebendo novas regras, as quais serão aplicadas aos parâmetros que serão medidos. Essa estrutura não é exclusiva para avaliação de dados de qualidade da água, podendo ser utilizada para outras situações que necessitem avaliações.

#### 4.1.2 CREDIBILIDADE

A credibilidade dos dados muitas vezes pode se confundir com outras características como a acurácia, por exemplo. Esta dimensão também pode ser confundida com a reputação do dado. Mas, diferentemente da reputação, a credibilidade significa o quanto o usuário confia

no dado.

O tomador de decisão, ao avaliar diversos dados, imagina que todas as informações coletadas são confiáveis, mas, dados ambientais coletados de fontes externas podem, inclusive, ser alterados por má fé. Nesse sentido, o sistema de regras vai gerar um valor escalar de confiança para os dados.

A escala é um indicador de confiança do dado e ajuda a categorizar diferentes níveis de decisões, inclusive na busca de melhoria dos dados. Ela tem como objetivo ajudar na auditoria dos dados, na escolha do laboratório, escolha de métodos para análise e coleta. Com isso, objetiva a melhoria contínua da qualidade dos dados.

Cada uma das regras aplicadas no sistema gerará uma nota que varia de 0 a 100. Depois desses escores calculados, será feita uma média ponderada para calcular a credibilidade final do dado.

A média aritmética ponderada  $\bar{x}_p$  de um conjunto de números  $x_1, x_2, x_3, \dots, x_n$  cuja importância relativa ("peso") é respectivamente  $p_1, p_2, p_3, \dots, p_n$  (GIOVANNI *et al*, 1994) é calculada como mostra a equação a seguir na Ilustração 9.

$$\bar{x}_p = \frac{p_1 \cdot x_1 + p_2 \cdot x_2 + p_3 \cdot x_3 + \dots + p_n \cdot x_n}{p_1 + p_2 + p_3 + \dots + p_n} = \frac{\sum_{i=1}^n (p_i \cdot x_i)}{\sum_{i=1}^n p_i}$$

Onde:

- n = Total de elementos
- p = Importância relativa
- x = Valor do elemento

**Ilustração 9 - Equação da média ponderada**

Fonte: Adaptado de (GIOVANNI *et al*, 1994)

Essa escala tem como pretensão uma melhor representação e interpretação da qualidade do dado contido no sistema. Os pesos serão determinados em função da importância da regra. São exemplos de regras que se aplicam ao mesmo dado e podem ter pesos diferentes: uma que explora o limite máximo e mínimo do parâmetro terá peso maior que uma que explora estatisticamente o valor máximo e mínimo.



### 4.1.3 OBJETIVIDADE

A objetividade - ausência de ambiguidade nos dados - é uma dimensão importante no sistema. A maioria dos problemas de ambiguidade normalmente são resolvidos com uma boa modelagem (WANG *et al*, 1993). Um mesmo dado não pode apontar para duas possibilidades de interpretação diferentes, pois isso pode gerar tomadas de decisão incoerentes.

Ao analisar apenas os valores dos dados, pode-se chegar a conclusão de que o ambiente está contaminado ou sofreu uma descarga de elementos que poluíram o corpo hídrico. Porém, essa tomada de decisão pode estar incorreta se for levado em consideração a condição climática.

Para ajudar a diminuir a ambiguidade das informações que serão retiradas do sistema é importante estar ciente do que estava ocorrendo no entorno do ponto. Uma técnica que pode ajudar a melhorar a qualidade nesse sentido é a inserção de observações em campos que receberão essas informações na tabela de dados. Os campos da tabela dos parâmetros medidos (parâmetro\_medido) são:

- **Observação da coleta (observacao\_coleta):** Observações sobre como aconteceu a coleta e como estava o local desta.
- **Observação da amostra (observacao\_amostra):** Observação de como estava o estado da amostra e informações sobre a mesma fornecidas pelo laboratório responsável pela análise.
- **Observação climática (observacao\_climática):** Observação com base das condições climáticas que o ponto sofreu nos últimos dias.
- **Observação do entorno (observacao\_entorno):** Observação do que estava acontecendo ao redor do ponto quando foi coletado.

Esses campos ajudam a clarificar alguns detalhes sobre o dado que podem ajudar na melhor interpretação do que o seu valor quer realmente mostrar.

### 4.1.4 REPUTAÇÃO

A reputação do dado medido em campo é diferente da medição de um dado realizado no laboratório. Medir a reputação de um dado pode ser algo muito subjetivo. Diversos fatores influenciam na reputação de um dado e não apenas em qual laboratório foi avaliado ou qual organização o forneceu ou o gerou.

Para conseguir demonstrar essa reputação, são necessários diversos atributos que contextualizam e classificam o dado. Para avaliá-la no contexto da aplicação, foram levantados alguns dados que devem contemplar a informação sobre o parâmetro medido. Estes atributos estão elencados na Tabela 10.

**Tabela 10 - Itens do sistema que auxiliam a conhecer a reputação de um dado**

Item	Descrição
Método da coleta	Entidade de dados que irá representar o método de coleta usado, assim como a bibliografia que explica o seu funcionamento.
Método da análise	Entidade de dados que representa o método utilizado na análise do parâmetro medido
Laboratório onde foi feita a análise	Para parâmetros analisados em laboratório, o dado contará com informações sobre qual laboratório fez a análise bem como as licenças e qualificações que o laboratório possui e quando essas licenças e qualificações foram obtidas.
Responsável pela coleta	Empresa ou pessoa que fez a coleta da amostra da água.
Usuário que inseriu a informação	Essa informação é importante para saber qual foi o usuário que inseriu a informação. Aqui os usuários poderão ser também hidrelétricas, ETAs e empresas que realizam monitoramento.

O principal objetivo do aumento da reputação dos dados é ajudar na ponderação das notas do índice calculado para pontuar a qualidade do dado. A reputação com base nos métodos e responsáveis que estão gerando o dado só poderá ser observada a partir das séries históricas que serão geradas.

## 4.2 QUALIDADE DE DADOS NO CONTEXTO

A qualidade de dados no contexto são as dimensões que levam em consideração a tarefa ao qual o dado está servindo. Essas dimensões são percebidas quando esta se transformando o dado em informação ou se está tomando alguma decisão. Dentre dessas dimensões está o valor acrescentado do dado apresentado na seção 4.2.1, a relevância que o

dado possui seção 4.2.2, a idade que o dado possui para a tomada de decisão seção 4.2.3, a totalidade ou completude do dado seção 4.2.4 e a quantidade apropriada de dados seção 4.2.5.

#### **4.2.1 VALOR ACRESCENTADO**

Valor acrescentado entende-se por dado que gere um potencial competitivo superior a de um concorrente, ou os lucros que esse dado pode gerar. Como o objetivo desse trabalho é a qualidade e a fidedignidade do dado, essa dimensão é consequência direta da obtenção das outras dimensões.

A tomada de decisão correta em relação a um recurso natural é o maior valor acrescentado que um dado ambiental pode fornecer.

#### **4.2.2 RELEVÂNCIA**

A relevância dos dados é outro detalhe importante na hora da modelagem dos dados. Escolher atributos para uma entidade demanda uma análise da real relevância daquele atributo, pois muitas vezes alguns dados são pouco relevantes e difíceis de serem obtidos. A principal consequência de atributos poucos relevantes é uma tabela mais esparsa, ou seja, vários valores nulos, não respondidos ou não pesquisados.

Um problema encontrado na relevância é criar um ponto de equilíbrio entre as informações que ajudam a validar o dado e as informações que não serão de tanta valia. Isso em virtude de que informações como a data do recebimento da amostra pelo laboratório, por exemplo, é uma informação relevante, porém, nem todos os laboratórios informam ou controlam essa informação.

Esses dados sem relevância aumentam o número de campos a documentar para futuras consultas, a complexidade do modelo e não são úteis para a tomada de decisão. Ao contrário, os dados que possuem relevância devem obrigatoriamente ser preenchidos. Uma forma interessante de obrigar o preenchimento é ter no SGBD (Sistema Gerenciador de Banco de Dados) marcada a opção *NOT NULL* para os campos mais relevantes.

### 4.2.3 IDADE DOS DADOS

A idade dos dados é uma dimensão que pode gerar muitos problemas para o sistema. A avaliação da qualidade de um recurso hídrico depende de séries históricas; e a realidade atual depende de campanhas contínuas.

Como garantia de saber a real idade dos dados, o sistema armazena 3 campos de data. Na tabela “parâmetro\_medido” são eles:

- **Data da análise (data\_analise):** Data de quando o dado foi analisado pelo laboratório ou equipamento de campo.
- **Data da coleta (data\_coleta):** Aqui é a data de quando a amostra de água foi coletada. Esse dado é do tipo *timestamp* e possui a hora, que é extremamente importante para a aplicação de testes de qualidade de dados e tomada de decisão, pois a hora influencia diretamente um corpo hídrico em função das variações no fotoperíodo.
- **Data do dado (data\_insercao):** A data do dado é o *timestamp* de quando a informação foi inserida no sistema. Esse é um importante indicador e controlador de séries antigas que poderão ser inseridas no sistema.
- **Data do recebimento da amostra (data\_recebimento\_amostra):** Data de quando o parâmetro foi recebido pelo laboratório. Alguns parâmetros para poderem ser avaliados corretamente tem um tempo máximo desde a coleta até a análise do laboratório. Futuramente algumas regras poderão utilizar essa informação para aumentar ou diminuir a confiabilidade do dado.
- **Data do término das análises (data\_termino\_analise):** Data que armazena a informação de quando o laboratório terminou de analisar o parâmetro.

Os dados deverão permanecer no banco de dados mesmo depois de velhos, pois isso é um requisito do sistema (SIA). Portanto, esses campos deverão ser utilizados em consultas afim de filtrar dados antigos que não deverão aparecer nos cálculos e tomadas de decisões.

### 4.2.4 TOTALIDADE

Essa dimensão é definida como a medida onde os dados não faltantes são de profundidade e largura suficiente para a tarefa em mãos (PIPINO *et al*, 2002). Para conseguir

melhorar essa dimensão é necessário analisar todas as perguntas que um gestor poderá se fazer e assim modelar essas informações.

Para melhorar a totalidade de um dado muitas vezes é necessário criar valores especiais como: “Dado não existe”, “Não foi possível realizar a medição” ou “Dado existe mas não foi informado”. Esses valores especiais ajudam a categorizar e saber a real situação do ambiente.

#### **4.2.5 QUANTIDADE APROPRIADA DE DADOS**

Essa é uma dimensão que, para ser atendida, depende muito da tarefa a ser executada. Se forem dados em excesso o problema pode ser tratado com restrições nas consultas e geração de *views*. Se os dados não forem suficientes a análise não poderá ser feita.

Por isso os pontos de monitoramento devem ser definidos e monitorados segundo a periodicidade necessária para poder gerar uma série histórica com a quantidade de dados necessária e poder ser concisa. A definição da quantidade e dos parâmetros que serão avaliados também é importante.

A tarefa de definição da quantidade apropriada de dados passa por uma análise dos relatórios e modelos matemáticos que os pesquisadores e gestores vão querer usar. Para o sistema construído a exigência mínima dos dados para uma medição em campo é o valor do parâmetro, a data e hora da coleta e o ponto exato (Coordenadas) da mesma. Sem essas informações, o dado não possui serventia nenhuma, podendo assim ser descartado.

### **4.3 REPRESENTAÇÃO**

As dimensões de representação apresentam fatores que determinam a qualidade do dado levando em consideração a forma como o dado está representado e como ele pode ser visualizado ao ser recuperado (PIPINO *et al*, 2002). As dimensões que fazem parte da representação são respectivamente: interpretabilidade seção 4.3.1, compreensibilidade seção 4.3.2, representação consistente seção 4.3.3, representação concisa 4.3.4 e facilidade de manipulação 4.3.5.

### 4.3.1 INTERPRETABILIDADE

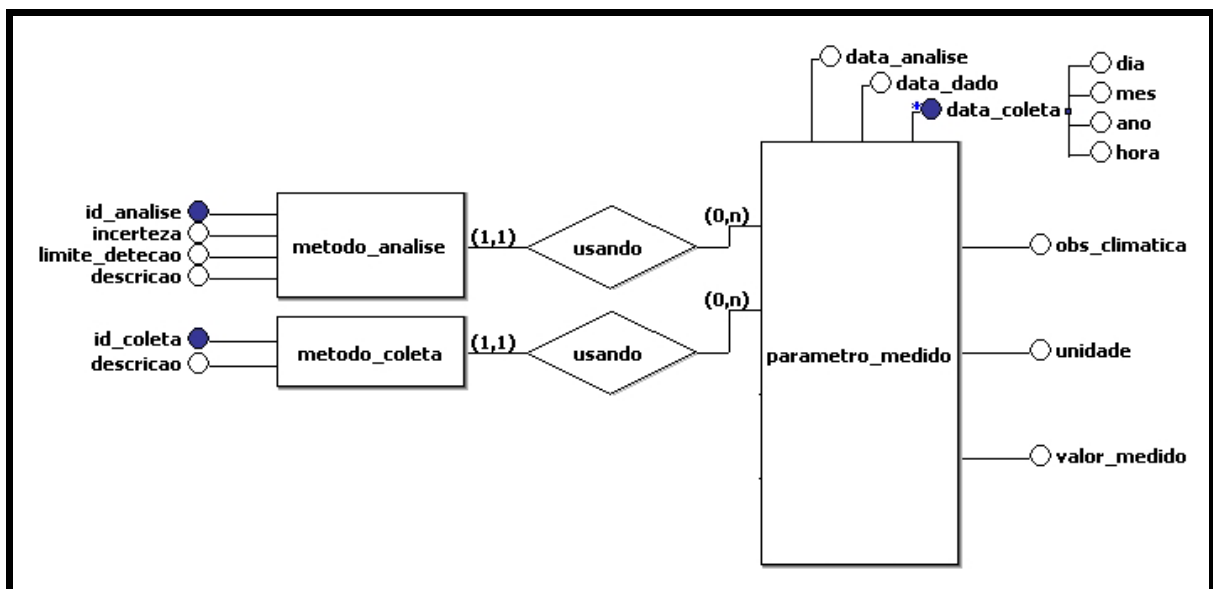
Para interpretar todo um conjunto de dados é extremamente importante armazená-los de forma que seja possível visualizar com clareza dados simples como datas, coordenadas e representações numéricas.

Para essas representações é interessante usar funções do próprio SGBD que mostrem o dado de forma padronizada com o padrão da língua e da região.

A língua em que as informações estão armazenadas também é uma importante forma de medir a interpretabilidade do dado. Os textos explicativos de métodos deverão, portanto, ser traduzidos para o português, melhorando desta forma o entendimento dos métodos como um todo.

### 4.3.2 COMPREENSIBILIDADE

O dado científico é mais compreensível pelo público técnico da área. Contudo, algumas medidas podem ser tomadas para evitar que a compreensão fique restrita a esse público. Uma das técnicas aqui adotadas para aumentar a compreensibilidade dos dados que estão no sistema é a utilização de descrições para informações mais técnicas. A Ilustração 10 mostra as entidades “metodo\_coleta” e “metodo\_amostra” com um campo descrição.



**Ilustração 10 - Campos de descrição que ajudam a compreender melhor um dado**

A descrição de campos ajuda o usuário que não conhece aquela informação entender o que ela significa. Medidas de Interpretabilidade também ajudam na compreensão do dado.

### 4.3.3 REPRESENTAÇÃO CONSISTENTE

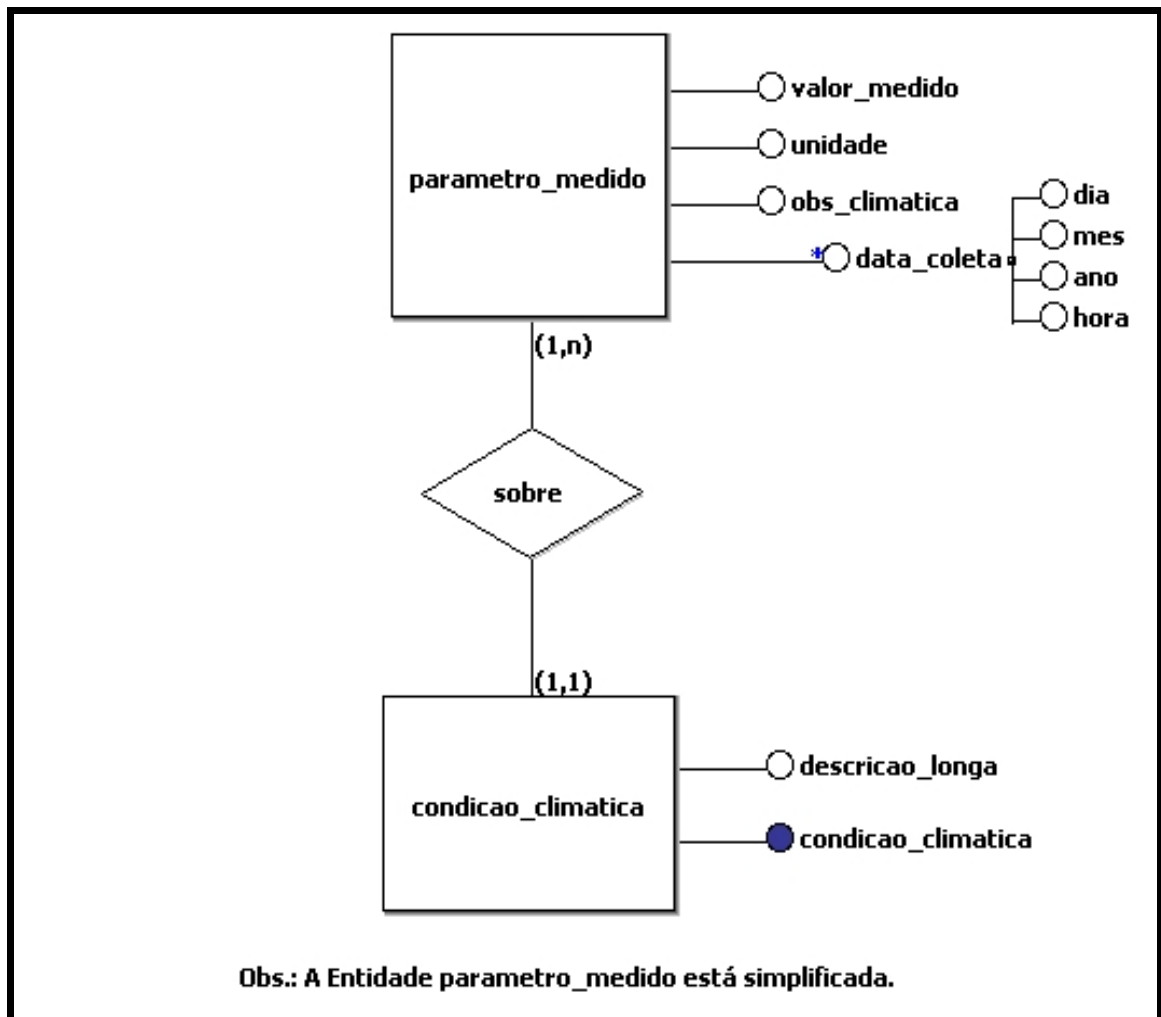
O principal problema de consistência no sistema é a unidade dos parâmetros medidos em campo. Para cada parâmetro terá que ser estipulado uma representação padrão e uma unidade padrão.

As unidades padrão para os parâmetros do sistema estão no Anexo II. Cada uma das unidades deverá possuir métodos implementados dentro do SGBD para que façam conversões automáticas entre elas. Isso ajuda a evitar que quem está inserindo o dado possa fazer uma conversão incorreta.

Além de auxiliar na acurácia dos dados a conversão automática das unidades garante que os parâmetros e seus valores possuam uma representação uniforme e consistente.

### 4.3.4 REPRESENTAÇÃO CONCISA

Essa dimensão deve ser tratada com cuidado para não ferir a dimensão Compreensibilidade. No sistema, para melhorar a concisão, algumas informações foram categorizadas, o que, além de expressar uma ideia com menos palavras, padroniza as informações. Um exemplo disso pode ser visto na criação da entidade *condição\_climática* apresentada na Ilustração 11.



**Ilustração 11 - Condição climática atributo transformado em entidade**

É possível observar na Ilustração 11 que a condição climática poderia ser um atributo de parâmetro medido, entretanto, foi transformado em uma entidade, garantido uma representação mais simplificada. A entidade “condição climática” conta ainda com uma descrição mais detalhada (caso seja necessário).

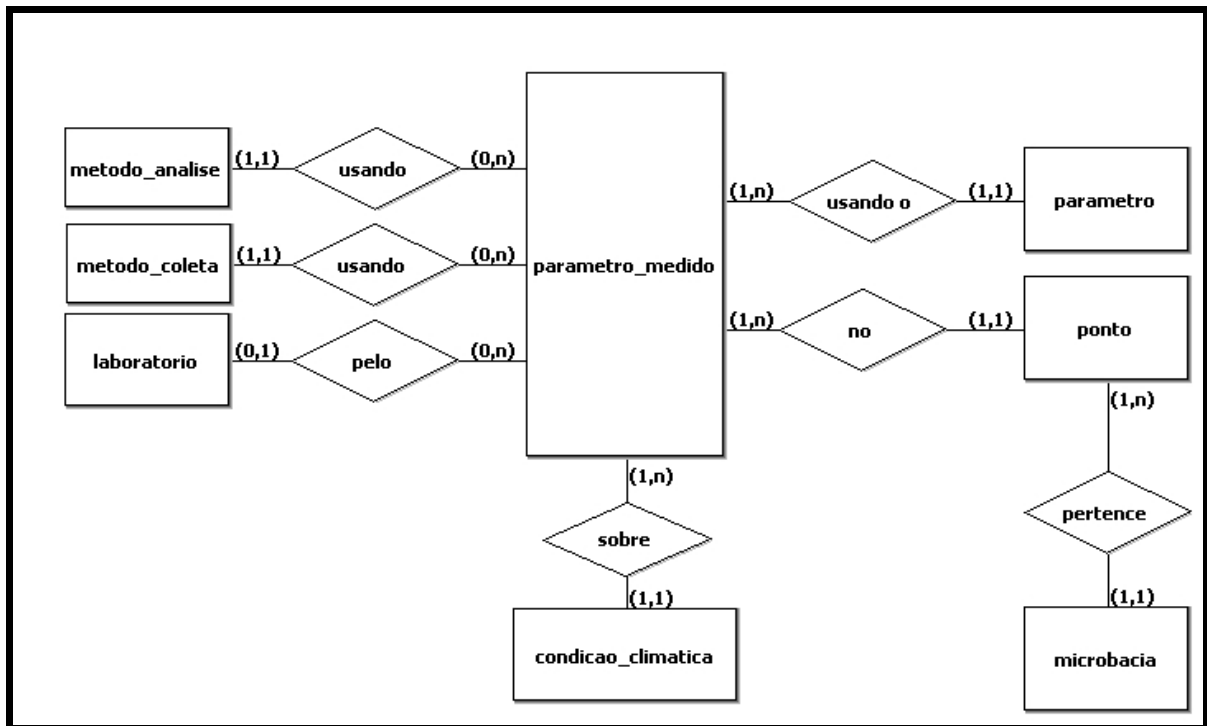
#### 4.3.5 FACILIDADE DE MANIPULAÇÃO

A estrutura para a manipulação dos dados da água foi embasada na modelagem dimensional. A principal vantagem da modelagem dimensional é a simplicidade para a consulta em comparação com a modelagem em 3<sup>a</sup> forma normal (KIMBALL, 2002).

A modelagem simplificada sem os atributos, como pode ser vista na Ilustração 12, que utiliza o modelo conhecido como *Snowflake* (Floco de neve), que é uma variação da *Star schema* (Esquema estrela). Essas modelagens são mais simples, porque a consulta é sempre a



partir de uma entidade de fatos, e as entidades associadas são conhecidas como dimensões. A Ilustração 12 demonstra o modelo *Snowflake* da forma como o sistema foi modelado, onde “parametro\_medido” é a tabela de fatos e as demais entidades são as tabelas de dimensões.



**Ilustração 12 - Modelagem dimensional do sistema**

Kimball recomenda que sempre se use a modelagem no esquema estrela por ser mais simples. Mas, para poder representar de forma mais eficiente a dimensão ponto, foi necessário normalizar essa parte do modelo de dados. A 3NF elimina a redundância de dados. Essa dimensão é um ponto mais crítico e que poderá sofrer mais atualizações.

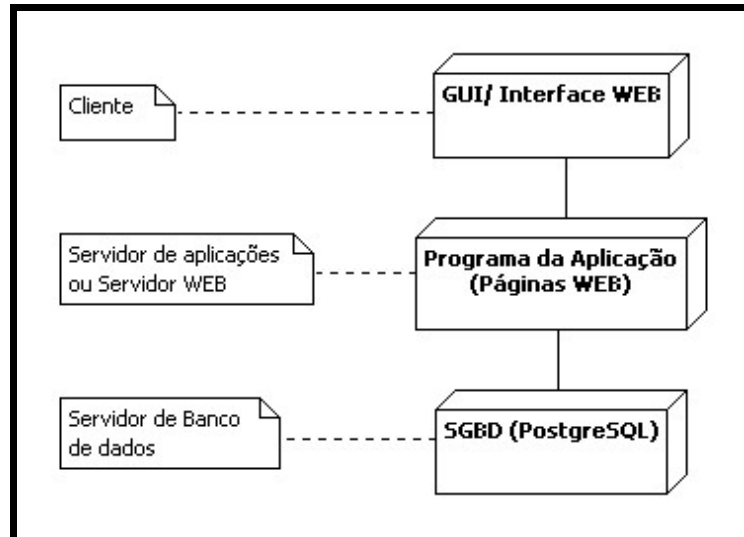
Além da modelagem dimensional, as técnicas utilizadas para a representação consistente e a representação concisa são duas dimensões que somadas contribuem significativamente para a facilidade de manipulação dos dados.

#### 4.4 ACESSIBILIDADE

A dimensão acessibilidade, ou disponibilidade, caracteriza mais uma potencialidade do sistema do que do dado como um todo. Essa dimensão depende de fatores externos como a configuração do servidor, diretrizes de segurança quanto ao servidor e a arquitetura do sistema.

Outra importante característica que deve ser avaliada é a arquitetura do sistema. Este

sistema utiliza a Arquitetura Cliente/Servidor de Três Camadas para aplicações *web* (ELMASRI e NAVATHE, 2005). A Ilustração 13 mostra a comunicação e os componentes das 3 camadas.



**Ilustração 13 - Arquitetura cliente/servidor de três camadas**

Fonte: Adaptado de (ELMASRI e NAVATHE, 2005)

Por definição, as funções de camada são respectivamente: Interface com o usuário, regras de aplicação e acesso aos dados (ELMASRI e NAVATHE, 2005).

Uma boa alternativa para a melhora da acessibilidade é definir e implantar padrões de interoperabilidade, para que os dados fiquem disponíveis para o sistema em si e para outros sistemas.

Para os *softwares* de serviço público, o Governo Brasileiro adota os padrões de interoperabilidade definidos como: “Interoperabilidade não é somente integração de sistemas nem somente integração de redes. Não referencia unicamente troca de dados entre sistemas e não contempla simplesmente definição de tecnologia. É, na verdade, a soma de todos esses fatores, considerando, também, a existência de um legado de sistemas, de plataformas de *hardware* e *software* instaladas” (E-PING, 2008).

A divisão acessibilidade além dela mesma possui como dimensão a segurança de um dado que é apresentada no item 4.4.1.

#### 4.4.1 SEGURANÇA

A dimensão segurança é muito ampla e necessitaria mais estudo do que o trabalho aqui apresentado. Por segurança de dados pode-se abranger problemas envolvendo o servidor,

problemas de *hardware*, *software* mal configurado, problemas de acesso indevido ao equipamento servidor ou por alguma catástrofe natural.

Outra forma de interpretar a segurança de dados é o questionamento de qual é o risco de dados confidenciais de uma organização serem publicados ou acessados por terceiros não autorizados. Nesse sentido, é muito importante a segurança do sistema de informação e a qualidade e satisfação da equipe envolvida com a manipulação dos dados.

Algumas medidas de segurança que estão sendo adotadas no *software* desenvolvido SIA estão demonstradas na Tabela 11.

**Tabela 11 - Medidas de segurança para os dados adotado no sistema**

Item	Descrição
Usuário e senha para cada usuário do sistema	Cada usuário do sistema deverá ter uma identificação privada, o que auxilia a monitorar possíveis problemas que estejam ocorrendo no sistema e poder parar os ataques.
Permissões diferentes para funcionalidade do Banco	O sistema deve tratar funcionalidades diferentes com usuários de SGBD diferentes e com permissões diferentes. Isso ajuda a evitar possíveis ataques conhecidos como <i>SQL Injection</i> . Dessa forma, um usuário de consulta do sistema nunca poderá excluir um dado.
Conexão segura (HTTPS/SSL)	A acessibilidade dos dados vai ser através de um sistema <i>web</i> . Para comunicar essas informações é necessário a criação de uma conexão segura para que os dados sejam trafegados com criptografia pela rede, evitando que algum programa de <i>Sniffing</i> os capture.
Criação de arquivos de <i>log</i> (Auditoria)	Criação de arquivos de <i>log</i> para monitorar o que cada usuário está alterando ou consultando nos dados.
Bloqueio de tentativas de login	A cada 15 tentativas falhas o usuário ficará bloqueado, tendo que solicitar nova senha
Troca de senha	A troca de senha deve ser realizada com certa frequência, o que ajuda a evitar que uma pessoa se passe por muito tempo por outra.
Teste de segurança de senha	Ao criar uma senha de usuário para o sistema, a mesma deverá ter o mínimo de segurança, utilizando ao menos letras e números.

A segurança da informação e dos dados tanto a nível de configuração de servidores, de *hardware* possui padrões e medidas dispostas na apresentadas na ISO 27002(2005) e nos procedimentos de *hardening* em servidores *Linux* (NSA, 2009). Para a segurança do SIA esses

documentos serviram como base e medida. O sistema de avaliação de qualidade de dados, possui algumas medidas de auditoria como as tabelas de *logs* do sistema. A tabela de regras\_aplicadas por exemplo armazena o usuário e a regra que ele aplicou ao dado.

## 5 MODELAGEM E IMPLEMENTAÇÃO DO SISTEMA DE QUALIDADE DE DADOS

O modelo de dados para o sistema de qualidade de dados foi estruturado pensando nas condições e avaliação das 16 dimensões apresentadas no capítulo 4, e divide-se em dois módulos distintos: O módulo de regras implementado direto no banco de dados, e o módulo de avaliação e controle das notas implementado em PHP (*Hypertext Pre Processor*).

### 5.1 SELEÇÃO DO SISTEMA GERENCIADOR DE BANCO DE DADOS

A seleção do banco de dados se baseou em um estudo da viabilidade e necessidade para a implementação do sistema. Devido a questão de não se saber qual a quantidade exata de dados que o banco precisará armazenar e não haver recursos para a compra da licença de um SGBD de grande porte, optou-se por considerar os bancos de dados gratuitos disponíveis.

O SGBDS gratuitos reconhecidos considerados foram: Firebird, MySQL, HSQL, SQLite e PostgreSQL.

Alguns critérios foram utilizados para a escolha do SGBD como:

- Quantidade de documentação.
- Língua da documentação.
- Experiência na utilização.
- Tipo de licença.
- Facilidade na criação de *Triggers* e *Stored Procedures*

Em consideração a estes critérios foram realizados alguns testes, e o SGBD PostgreSQL foi o escolhido. O fator determinante foi a possibilidade de se utilizarem diversas linguagens para a construção de *Stored Procedures* e *Triggers*.

A linguagem utilizada para a criação das *Triggers* e *Stored Procedures* foi o *Ruby* 1.8. Essa linguagem é uma linguagem de *script* orientada a objetos. A documentação para a utilização e instalação, em português, está disponível na tradução do manual do PostgreSQL (OLIVEIRA, 2006).

## 5.2 CRIAÇÃO DO BANCO DE DADOS

O banco de dados foi criado em um servidor com o sistema operacional Linux Ubuntu 9.04. O resultado do comando `uname -a`, apresentando a versão do *linux* e do *kernel* no servidor é mostrado na Ilustração 14.

```
Linux cebolao 2.6.28-18-generic #60-Ubuntu SMP Fri Mar  
12 04:40:52 UTC 2010 i686 GNU/Linux
```

**Ilustração 14 - Execução do comando `uname -a` mostrando a versão do linux**

A versão do *postgresql* utilizada para o projeto foi a versão 8.3 e o resultado do comando `SELECT version();` é apresentado na Ilustração 15.

```
PostgreSQL 8.3.9 on i486-pc-linux-gnu, compiled by GCC  
gcc-4.3.real (Ubuntu 4.3.3-5ubuntu4) 4.3.3
```

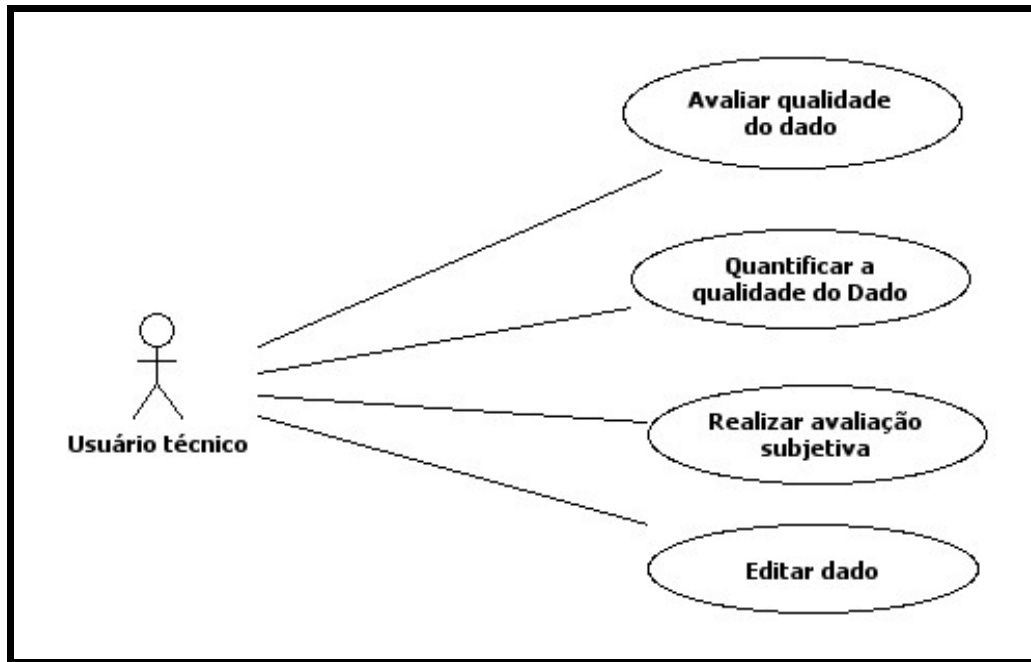
**Ilustração 15 - Resultado da consulta `SELECT version();` no PostgreSQL**

A tradução do diagrama ER gerado para o sistema para a linguagem SQL gerou o código contido no Anexo IV.

## 5.3 MODELAGEM DO SISTEMA

Avaliando os dados das campanhas realizadas juntamente a hidrelétricas da região, os dados de monitoramento realizados pelo Instituto e, através do levantamento dos requisitos com os técnicos e professores do ISAM, foi elaborado o diagrama de caso de uso apresentado na Ilustração 16. Esse diagrama de caso de uso mostra as necessidades que o sistema de avaliação de dado precisaria comportar. A modelagem dos diagramas de Caso de Uso foram elaborados conforme (WAZLAWICK, 2004).

Os casos de uso “Avaliar a qualidade do dado” e “quantificar a qualidade do dado” representam o Módulo apresentado na seção 5.4 Módulos de Regras, já os casos de usos “realizar avaliação subjetiva” e “editar dado” fazem parte do módulo denominado de avaliação e controle das notas apresentado na seção 5.5.



**Ilustração 16 - Diagrama de Caso de Uso geral do Sistema**

As Ilustrações 17, 18, 19 e 20 apresentam a descrição dos casos de uso apresentados na Ilustração 16.

<b>CASO DE USO:</b>	Avaliar a qualidade do dado
<b>VISÃO GERAL:</b>	O usuário técnico ao carregar o dado para o banco de dados, o qual dispara uma <i>Trigger</i> para verificar os dados com possíveis problemas. Para isso vai utilizar regras objetivas fazendo correlações e usando estatística.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	-
<b>PÓS-CONDIÇÃO</b>	O dado está correto ou tende a estar errado.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O banco de dados recebe o dado;</li> <li>2. O sistema seleciona as regras possíveis para avaliar o dado.</li> <li>3. O sistema dá a destinação final do dado.</li> </ol>	

**Ilustração 17 - Descrição do caso de uso avaliar a qualidade do dado**

<b>CASO DE USO:</b>	Quantificar a qualidade do dado
<b>VISÃO GERAL:</b>	O banco de dados deve quantificar a qualidade dos dados, possibilitando dessa forma uma avaliação mais objetiva da qualidade do dado.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	Ter executado as regras objetivas e possuir as notas dessas regras
<b>PÓS-CONDIÇÃO</b>	O dado recebe uma nota da sua confiabilidade.
<b>FLUXO PRINCIPAL</b>	
	<ol style="list-style-type: none"> <li>1. O sistema recebe as notas que o dado obteve executando a regra dos dados;</li> <li>2. O sistema utiliza média ponderada (ver item 4.1.2) para calcular e quantificar a qualidade do dado;</li> <li>3. O sistema aplica essa nota ao dado.</li> </ol>

### Ilustração 18 - Descrição do caso de uso Quantificar a qualidade do dado

<b>CASO DE USO:</b>	Realizar avaliação subjetiva do dado
<b>VISÃO GERAL:</b>	O usuário técnico pode fazer uma avaliação subjetiva do dado, dessa forma ele pode aprová-lo. Essa avaliação subjetiva normalmente apresenta duas situações: aprova o dado (pois não tem justificativas que confirmem que o dado esteja errado) ou justifica o dado com informações .
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	O dado não ter obtido qualidade superior a 80(%) na avaliação objetiva.
<b>PÓS-CONDIÇÃO</b>	O dado com qualidade inferior a 80 (%) vai para a base de consulta. Ou fica marcado como um dado com problemas.
<b>FLUXO PRINCIPAL</b>	
	<ol style="list-style-type: none"> <li>1. O usuário técnico acessa o sistema avalia o dado e informa a justificativa para o dado.</li> <li>2. O usuário aprova o dado.</li> <li>3. O dado vai para a base de consulta.</li> </ol>
<b>FLUXOS ALTERNATIVOS</b>	
	<ol style="list-style-type: none"> <li>2.1 O usuário descarta o dado.</li> <li>2.2 O dado não vai para a base de consulta e fica armazenado como um dado com problema.</li> </ol>

### Ilustração 19 - Descrição do caso de uso realizar avaliação subjetiva do dado

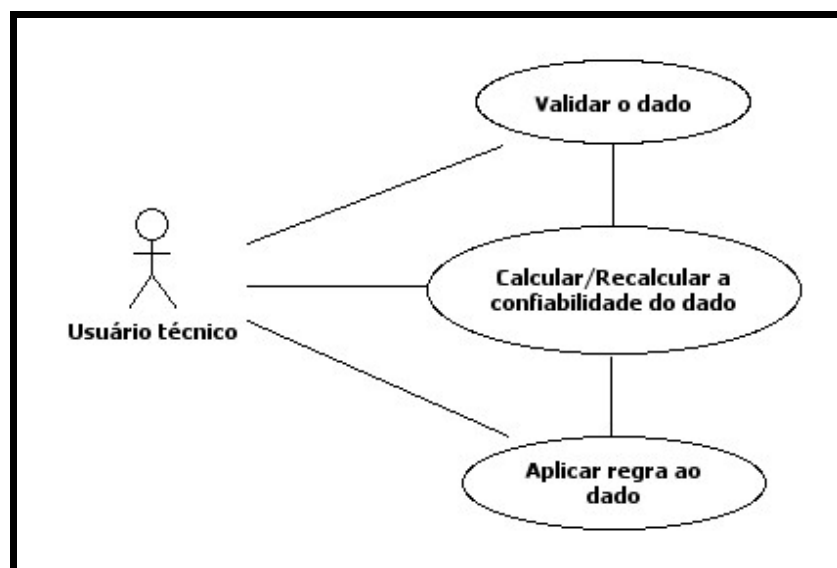


<b>CASO DE USO:</b>	Editar dado
<b>VISÃO GERAL:</b>	Caso o usuário técnico perceba que o dado que foi digitado está incorreto ele pode alterar o seu valor e submeter a uma nova avaliação.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	O dado estar no sistema como um dado com problema
<b>PÓS-CONDIÇÃO</b>	O dado é reavaliado.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O usuário altera as informações necessárias referente ao dado.</li> <li>2. O sistema recalcula as notas desse dado.</li> </ol>	

**Ilustração 20 - Descrição do caso de uso editar dado**

#### 5.4 MÓDULO DE REGRAS

O módulo de regras foi implementado diretamente no banco de dados. As funcionalidades deste módulo são apresentadas no diagrama de caso de uso da Ilustração 21. A descrição dos casos de uso apresentados no diagrama encontra-se, respectivamente, nas Ilustrações 22, 23 e 24.



**Ilustração 21 - Diagrama de caso de uso do Módulo de Regras**

<b>CASO DE USO:</b>	Validar o dado
<b>VISÃO GERAL:</b>	O usuário técnico ao inserir ou atualizar uma informação ativa a <i>Trigger</i> que aplica todas as regras aos dados.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	O dado da campanha não ter sido inserido anteriormente ou ser atualização.
<b>PÓS-CONDIÇÃO</b>	O dado fica validado pelo sistema.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O banco de dados informa o dado ao sistema de regras.</li> <li>2. Seleccionam-se as regras que se aplicam àquele dado.</li> <li>3. Recebe as notas que cada regra retornou.</li> <li>4. Calcula a média ponderada conforme os pesos da regra.</li> <li>5. Armazena a nota final do dado.</li> </ol>	
<b>FLUXOS ALTERNATIVOS</b>	
2.1 Se faltar informação para calcular a regra, marca essa regra para ser validada posteriormente	

### Ilustração 22 - Descrição do caso de uso Validar o dado

<b>CASO DE USO:</b>	Calcular / Recalcular a confiabilidade do dado
<b>VISÃO GERAL:</b>	Calcula a nota final do dado ou recalcula caso seja necessário.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	O dado já estar inserido no sistema.
<b>PÓS-CONDIÇÃO</b>	A nota final do dado é recalculada.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O banco de dados informa uma lista com as notas e os respectivos pesos</li> <li>2. A nota final da confiabilidade do dado é calculada através de uma média ponderada.</li> <li>3. A nota que representa a confiabilidade do dado é armazenada no sistema.</li> </ol>	

### Ilustração 23 - Descrição do caso de uso Calcular / Recalcular a confiabilidade do dado

<b>CASO DE USO:</b>	Aplicar regra ao dado
<b>VISÃO GERAL:</b>	Verifica todas as regras possíveis a ser aplicadas ao dado de entrada
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	O dado da campanha não ter sido inserido anteriormente ou ser atualização.
<b>PÓS-CONDIÇÃO</b>	O dado fica validado pelo sistema.
<b>FLUXO PRINCIPAL</b>	
	<ol style="list-style-type: none"> <li>1. O banco de dados informa o dado ao sistema de regras.</li> <li>2. Selecionam-se as regras que se aplicam àquele dado.</li> <li>3. Recebe as notas que cada regra retornou.</li> <li>4. Calcula a média ponderada conforme os pesos da regra.</li> <li>5. Armazena a nota final do dado.</li> </ol>
<b>FLUXOS ALTERNATIVOS</b>	
	2.2 Se faltar informação para calcular a regra, marca essa regra para ser validada posteriormente

#### **Ilustração 24 - Descrição do caso de uso Aplicar regra ao dado**

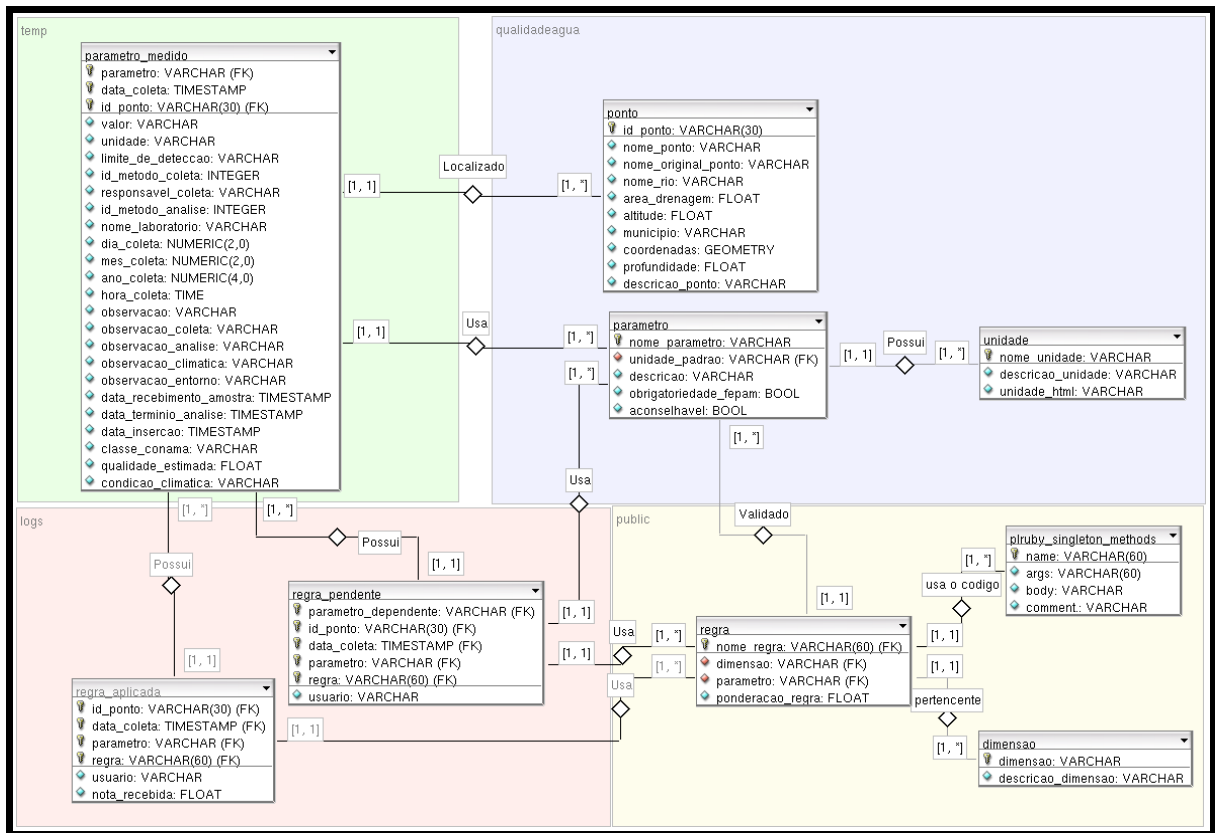
No item 5.3.1 é mostrado como foi implementada a parte de seleção e execução de regras. No item 5.3.2 é apresentado quais são os tipos de regras e como elas trabalham na melhora da qualidade e confiabilidade do dado. Por último, o item 5.3.3 apresenta uma classificação para uniformizar a confiabilidade sobre o dado.

#### **5.4.1 SELEÇÃO E EXECUÇÃO DE REGRAS**

As regras, como apresentadas no item 4.1 Acurácia, mostram que ao diagrama ER foram adicionadas duas entidades: uma para as regras e outra para a dimensão a qual a regra se aplica. A implementação no banco de dados necessitou alguns ajustes: O atributo código de regra foi retirado e foi acrescentada a tabela “*plruby\_singleton\_methods*”. Isso se deve ao fato desta tabela armazenar as funções desenvolvidas pelo usuário na linguagem *Ruby*, (instalada no SGBD).

Na Ilustração 25 podemos observar o modelo de dados implementando as tabelas necessárias para o Módulo de Regras poder ser construído. Os quadros apresentam o esquema ao qual as tabelas foram implementadas. A tabela “*parametro\_medido*”, que fica no esquema “*temp*”, é uma réplica da tabela “*parametro\_medido*” do esquema “*qualidadeagua*” (vide Anexo IV). Além da réplica a tabela “*parametro\_medido*” do esquema “*temp*” que teve o

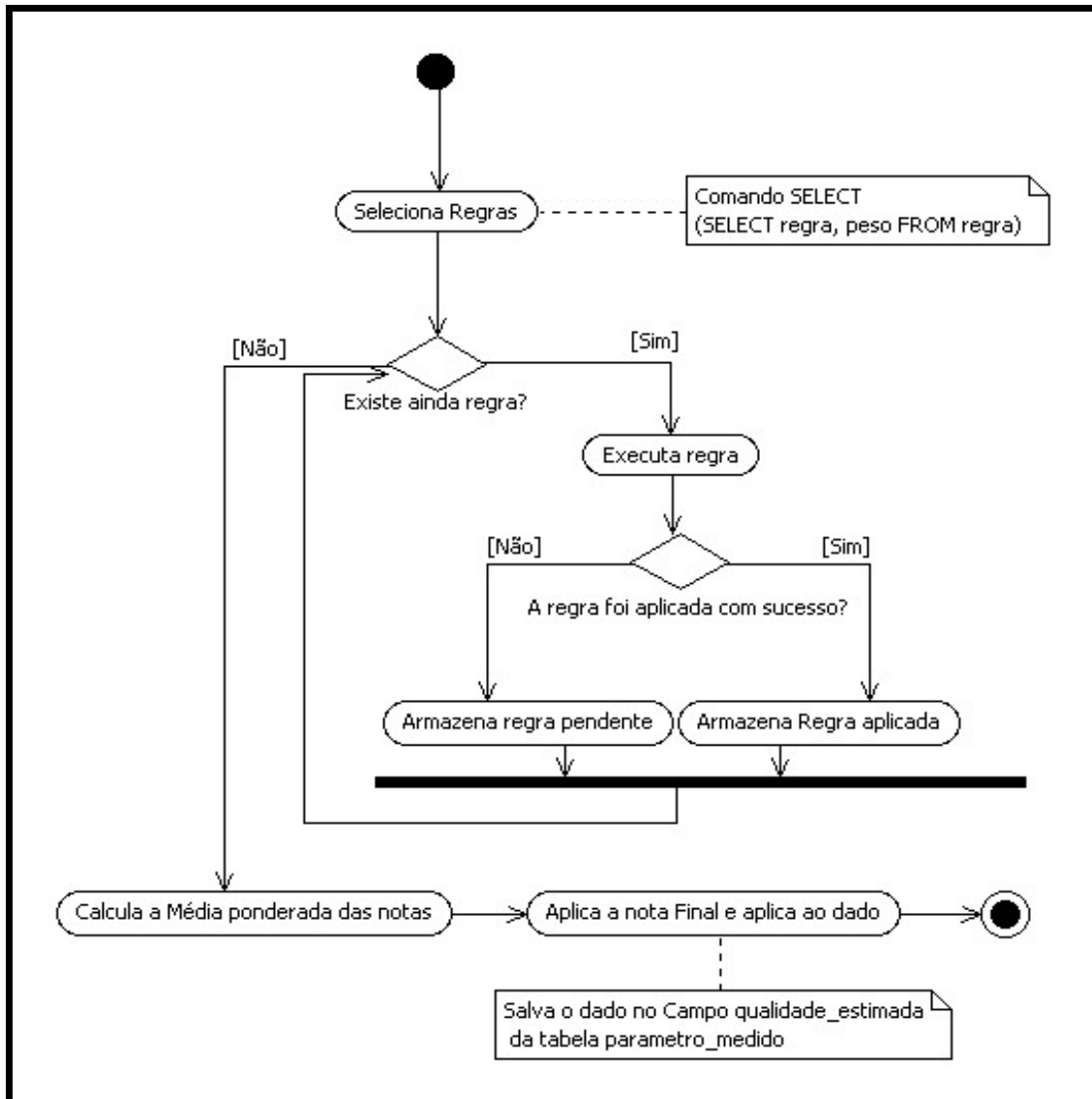
atributo “`dado_avaliado`” acrescentado. Este atributo é um campo booleano simbolizando se o dado foi validado ou não.



**Ilustração 25 - Modelo de dados da implementação do módulo de regras**

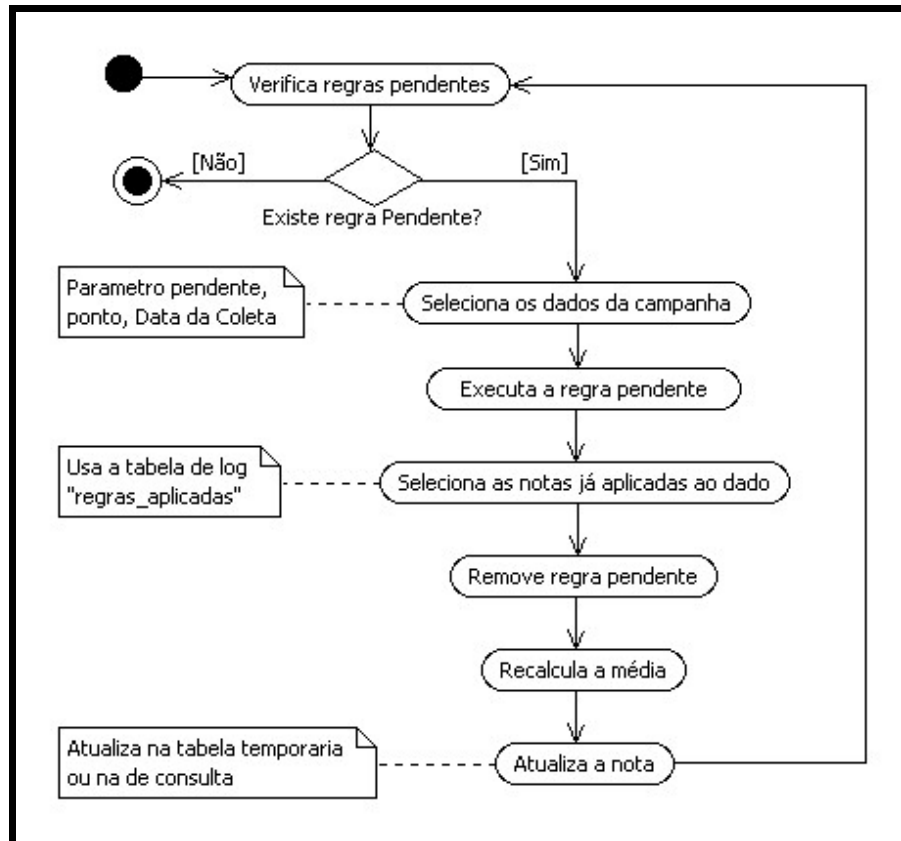
Além dessas tabelas vistas no modelo de dados, as outras entidades que foram modeladas no capítulo 4 foram suprimidas da Ilustração 25 para facilitar o entendimento. Como já citado, a conversão das mesmas para a linguagem do banco de dados SQL está no Anexo IV.

A validação de um dado começa no momento em que o SGBD recebe uma solicitação de inserção ou atualização do dado na tabela “`qualidadeagua.parametro_medido`”. Ele dispara o gatilho (*trigger*) chamado “`valida_dado`” que seleciona as regras e as executa (Ilustração 26).



**Ilustração 26 - Diagrama de atividades da *trigger* executada antes da Inserção**

Após a inserção dos dados na tabela “parametro\_medido”, o banco de dados verifica as regras que ficaram pendentes disparando o *trigger* “regras\_pendentes”, que executa as regras que dependem do dado que acabou de ser inserido. O funcionamento desse *trigger* está apresentado na Ilustração 27.



**Ilustração 27 - Diagrama de atividades da *trigger* executado depois da Inserção**

#### 5.4.2 TIPOS DE REGRAS

O sistema apresenta no total 32 regras (ANEXO V) e são divididas basicamente em 3 tipos de regras distintas, denominadas: Regras simples, Regras compostas e Regras estatísticas.

- Regras simples: São regras que independem dos dados já armazenados, ou seja, utilizam apenas o valor do dado por si só. Essas regras normalmente servem para o limite mínimo e máximo do dado. Algumas dessas regras simples são mais subjetivas, como: “O pH em águas que correm sobre uma geologia que apresentam rochas vulcânicas ácidas como riódacitos e riólitos, característica da bacia (HAZENACK E WEBER, 2007). Devido a esse fator o pH tende a ser menor que 8”. Essa regra está presente no Anexo V.
- Regras compostas: São em maior número e utilizam outros parâmetros ou dados já armazenados no sistema relacionados ao ponto, como, por exemplo, a regra OD2, que está representada no Anexo V. Essas regras podem depender de dados não inseridos no sistema, caso falte algum parâmetro, a regra só será executada quando o mesmo for

inserido. A Ilustração 27 mostra a execução desse tipo de regra.

- Regras estatísticas: Estas regras utilizam como base os dados já aprovados, e usam alguns métodos estatísticos como o gráfico de controle. O gráfico de controle contém uma linha central representando o valor médio da característica da qualidade que corresponde ao estado sob controle. Duas linhas horizontais chamadas de limite superior e inferior de controle. Estes limites de controle são escolhidos de modo que o processo esteja sobre controle. Em alguns casos pode-se usar o desvio padrão para se encontrar os limites superior e inferior. (MONTGOMERY, 2004).

### 5.4.3 PESOS E NOTAS FINAIS DAS REGRAS

Cada regra criada recebeu um peso que leva em consideração a confiabilidade da regra. Ou seja, quanto maior o peso maior é a influência que a regra tem sobre a nota da confiabilidade final. A nota que cada regra retorna toma como base a confiança do dado em relação essa regra.

Para qualificar o dado, de forma mais clara, foram evidenciadas na Tabela 12 faixas de qualidade, sendo que essas receberam os seguintes adjetivos: Péssima, Baixa, Média, Alta e Ótima. Esses adjetivos ajudam a universalizar a confiabilidade do dado. Com essas faixas, a avaliação e as notas da confiabilidade ficam menos subjetiva, tornando-se assim mais objetivas.

**Tabela 12 - Faixa de qualidade conforme a nota do dado**

Intervalo	Qualidade
0-20	Péssima
21-40	Baixa
41-60	Média
61-80	Alta
80-100	Ótima

Não é possível afirmar que um dado está 100% correto, uma vez que as variáveis ambientais são condicionadas por fatores limitantes. Estes fatores condicionam não apenas a existência e a distribuição dos seres vivos como também atuam uns sobre os outros, quer

individualmente, quer em processos sinérgicos que por sua vez interferem igualmente sobre os seres vivos. Os seres vivos, por sua vez, buscando sua adaptabilidade aos fatores ambientais, condicionam igualmente os fatores ecológicos (limitantes) criando microssistemas, microclimas, ambientes que lhe favoreçam a sobrevivência. Neste contexto, elencar os fatores ambientais é uma tarefa quase impossível, considerando-se que praticamente tudo o que acontece no meio sob os aspectos físicos, físico-químicos, químicos, bioquímicos e biológicos interfere na distribuição e sobrevivência dos seres vivos e vice-versa.

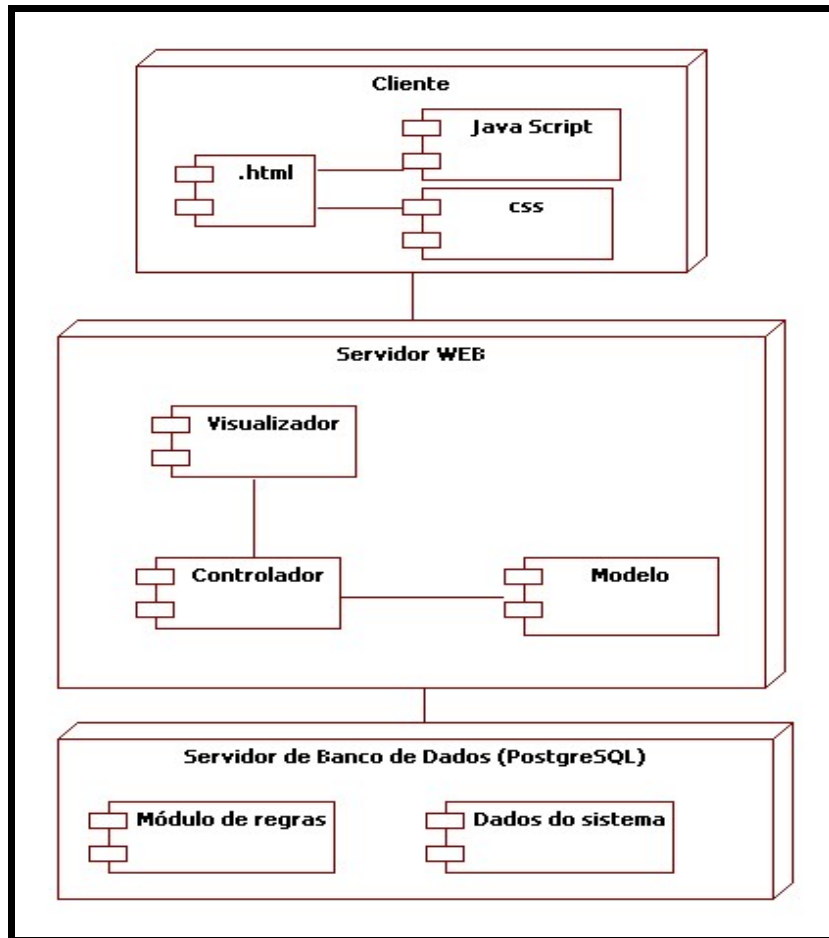
Além do exposto, deve-se considerar ainda a variabilidade dos fatores ambientais decorrentes. Por exemplo, da geografia (latitude e altitude), da pedologia (solos) e, fundamentalmente, dos fatores climáticos que transcendem, inclusive a dimensão planetária, uma vez que são dependentes de uma fonte externa de energia: o sol. Além disto, cabe destacar as inter-relações destes fatores, onde tudo está conectado a tudo, formando uma rede intrínseca de correlações em que a alteração de um fator implica necessariamente em mudanças na teia global de relações, resultando em novos fatores e decorrendo em consequências imprevisíveis a curto e médio prazos (DAJOZ, 2005)..

## 5.5 MÓDULO DE AVALIAÇÃO E CONTROLE DAS NOTAS

Devido a necessidade da interação humana para a avaliação da qualidade de dados, seja a justificativa ou descarte dos mesmos, assim como uma avaliação mais subjetiva deles, foi necessário criar algumas interfaces de interação com o usuário. Para o desenvolvimento das mesmas utilizou-se HTML (*HyperText Markup Language*), CSS (*Cascading Style Sheet*) e *Java Script*, e um sistema em PHP rodando em formato MVC (*Model View Controller*).

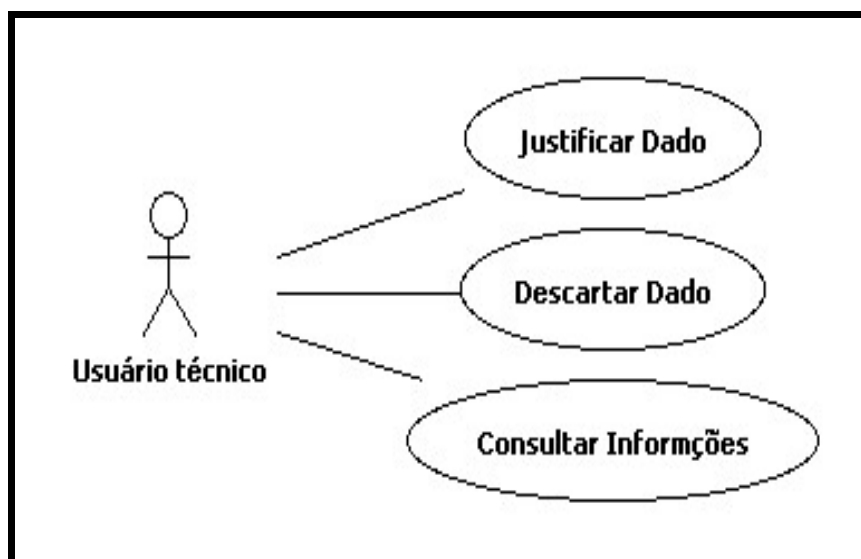
O módulo foi desenvolvido usando componentes do SIA, para que, posteriormente, o protótipo possa ser integrado. O Diagrama de Componentes apresentado na Ilustração 28 ilustra a relação entre os componentes de *software*.





**Ilustração 28 - Diagrama de Componentes demonstrando a arquitetura do módulo de avaliação e controle das notas**

As principais funcionalidades do módulo de avaliação e controle de qualidade de dados possuem estão descritos no diagrama de caso de uso apresentado na Ilustração 29. As descrições dos casos de uso aparecem, respectivamente, nas Ilustrações 30, 31 e 32.



**Ilustração 29 - Diagrama de Caso de uso do módulo de avaliação e controle das notas**

<b>CASO DE USO:</b>	Justificar Dado
<b>VISÃO GERAL:</b>	O usuário justifica o dado, faz as alterações necessárias e libera o dado para ser cadastrado na base de consulta.
<b>ATORES:</b>	Usuário Técnico
<b>PRÉ-CONDIÇÃO</b>	O dado já estar inserido na base temporária.
<b>PÓS-CONDIÇÃO</b>	O dado que estava armazenado na base temporária vai para a base permanente.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O usuário acessa o sistema e escolhe em uma lista qual dado quer justificar;</li> <li>2. O usuário edita as informações necessárias e confirma o dado;</li> <li>3. Independente da nota, o dado vai para a base permanente com a nota.</li> </ol>	

### **Ilustração 30 - Descrição do caso de uso Justificar Dado**

<b>CASO DE USO:</b>	Descartar Dado
<b>VISÃO GERAL:</b>	Elimina um dado que não tem como ser usado.
<b>ATORES:</b>	Usuário Técnico
<b>PRÉ-CONDIÇÃO</b>	O dado já estar inserido na base temporária.
<b>PÓS-CONDIÇÃO</b>	O dado é eliminado do sistema.
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O usuário escolhe em uma lista qual dado deletar</li> <li>2. O usuário técnico confirma a exclusão.</li> </ol>	

### **Ilustração 31 - Descrição do caso de uso Descartar Dado**

<b>CASO DE USO:</b>	Consultar Informações
<b>VISÃO GERAL:</b>	O usuário técnico pode acessar algumas consultas sobre a qualidade do dado ou sobre as regras aplicadas ao dado.
<b>ATORES:</b>	Usuário técnico
<b>PRÉ-CONDIÇÃO</b>	-
<b>PÓS-CONDIÇÃO</b>	-
<b>FLUXO PRINCIPAL</b>	
<ol style="list-style-type: none"> <li>1. O banco de dados executa a consulta.</li> <li>2. O sistema informa os dados dessa consulta e as ações que o usuário pode fazer;</li> <li>3. O usuário pode interagir com as consultas através de uma coluna ações.</li> </ol>	

### Ilustração 32 - Descrição do caso de uso Consultar Informações

As consultas que podem ser implementadas nesse módulo estão descritas na Tabela 13.

**Tabela 13 - Consultas do módulo Avaliação e controle das notas**

<b>CONSULTA:</b>	<b>DESCRIÇÃO:</b>
Visualizar dados da campanha	Visualiza os dados referentes a campanha do dado que está sendo avaliado. O usuário técnico tem a possibilidade de acessar os dados mesmo que estes não tenham sido aprovados para a base de consulta.
Visualizar regras aplicadas	Visualiza as regras que foram aplicadas ao dado assim como a nota atribuída a mesma.
Visualizar regras pendentes	Lista quais regras não puderam ser aplicadas em quais dados.
Visualizar regras	Lista quais as regras e quais os pesos que as mesmas possuem.

## 5.6 CONSIDERAÇÕES SOBRE O SISTEMA

O módulo de regras foi implementado direto no banco de dados. Apesar da linguagem utilizada permitir orientação a objetos, o módulo foi implementado de forma estruturada. O principal motivo foi a simplicidade das ações que o mesmo deveria executar para iniciar as

regras do sistema.

O módulo de avaliação e controle das notas funcionou com algumas funcionalidades básicas para correção do dado, justificativa ou descarte do mesmo, e foi implementado com base na mesma arquitetura e *framework* utilizados no desenvolvimento do SIA. Isso se deve ao fato de que depois que o protótipo estiver bem testado e aprovado poderá ser integrado ao SIA, adicionando, portanto, a funcionalidade da validação de dados para o sistema maior.

## 6 TESTES E IMPLANTAÇÃO

Dentre os principais testes realizados no sistema se destacam-se aqueles dedicados a verificar a acurácia das regras, apresentados no item 6.1, o teste de uso e aceitação do sistema (item 6.2). A implantação do sistema está apresentado no item 6.3.

### 6.1 TESTES PARA AS REGRAS

Antes de implantar o sistema, todas as regras criadas tiveram que ser testadas. Para poder validar as regras implementadas no sistema foram utilizados testes de caixa branca conhecidos como teste do caminho básico. O teste é feito utilizando uma técnica de caixa branca, que calcula a complexidade lógica do software e utiliza esta medida como base para descobrir os caminhos básicos do software, exercendo o teste de modo que todos os caminhos sejam efetuados (PRESSMAN, 2006; SOMMERVILE, 2007).

A partir do código das regras se calculou o número mínimo de testes que precisavam ser feitos. Usando o fórmula conforme é mostrado na Ilustração 33.

$$V(G) = E - N + 2.$$

Onde

$V(G)$  = e a complexidade ciclomática.

$G$  = representa o grafo de fluxo.

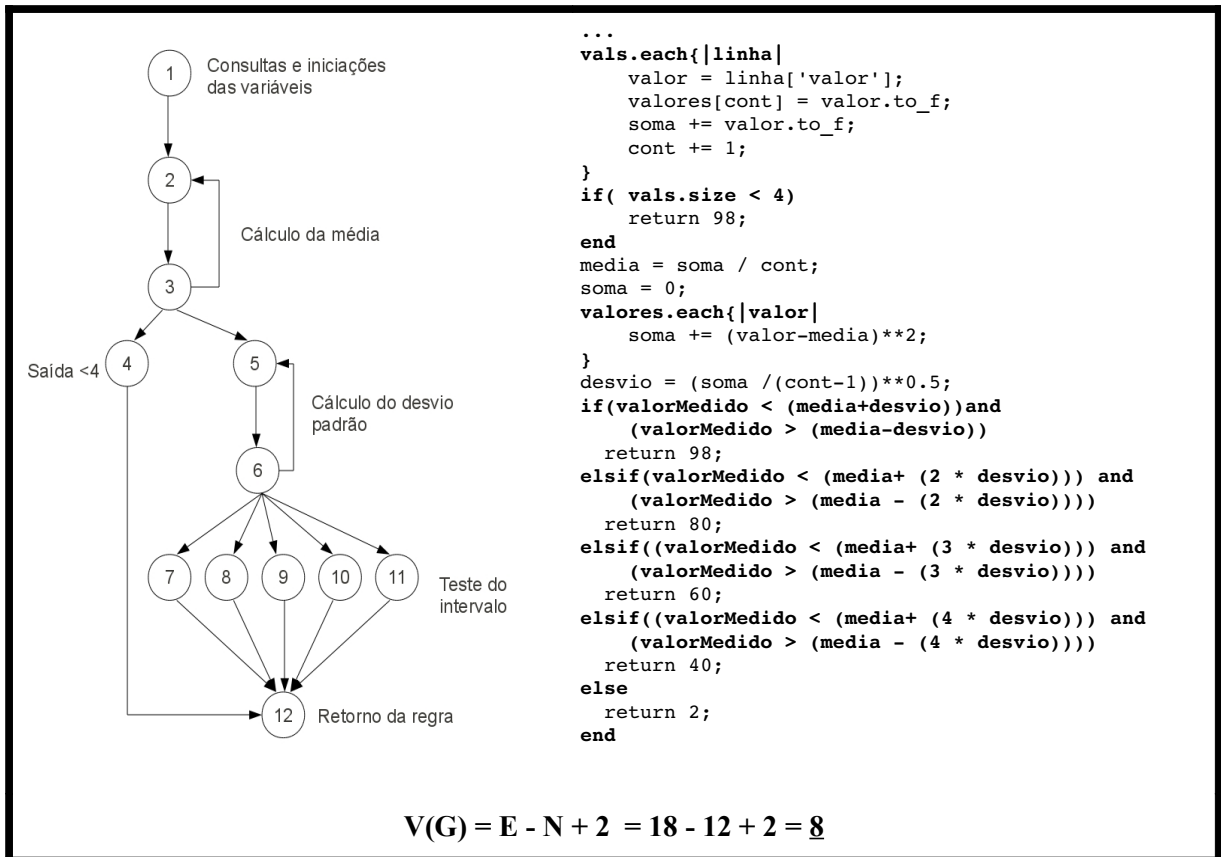
$E$  = representa a quantidade de arestas no grafo.

$N$  = representa a quantidade de ramos no grafo.

#### **Ilustração 33 - Fórmula para calculo da complexidade ciclomática**

Fonte: Adaptado de SOMMERVILE, 2007.

A Ilustração mostra o grafo ciclomático das regras estáticas como a regra od5, junto com o código da mesma, e o cálculo da complexidade ciclomática. Algumas partes do código fonte foram suprimidas o código completo da regra está no Anexo V.



### Ilustração 34 - Grafo ciclômico das regras estatísticas

O número mínimo de testes aplicados a cada regra está na Tabela 14.

Tabela 14 - Testes de caixa branca realizados por regra

Regra	Complexidade Ciclomática
condutividade0	2
condutividade1	8
condutividade2	8
condutividade3	4
dbo0	2
dbo1	2
dbo2	3
dbo3	4
dbo4	8
dbo5	8
dqo0	2
dqo1	3
dqo2	8

dqo3	8
od0	2
od1	2
od2	4
od3	6
od4	8
od5	8
od6	9
ph0	2
ph1	2
ph2	2
ph4	2
ph5	8
ph6	8
ph7	10
temperatura0	2
temperatura1	3
temperatura2	8
temperatura3	8
<b>TOTAL:</b>	<b>164</b>

Os testes aplicados ajudaram a corrigir falhas eventuais. Esses testes foram importantes porque normalmente as regras trabalham com os dados que são exceção e não com os dados corretos. Sem eles um erro de programação poderia levar até alguns anos para ser descoberto em função de dificilmente o dado assumir esses valores. Outro problema caso todos os caminhos possíveis não fossem testados são erros de sintaxe. Estes erros de sintaxe podem ocorrer devido ao PLRuby ser uma linguagem interpretada e não compilada. Ou seja, um erro de programação dificilmente seria encontrado.

Para os testes das regras estatísticas(ver Anexo V) foram necessárias a carga de alguns dados e repetitivas verificações para poder validar a sua consistência.

Devido a novas regras poderem ser inseridas com dados já carregados, podem-se aplicar testes com os dados que estão na base em um ambiente simulado preparado no módulo de regras. Esse ambiente simulado utiliza as tabelas de *log* e usa apenas os dados, sem alterar a base existente.

## 6.2 TESTES DE USO DO SOFTWARE

Para a aprovação do software foi proposto um formulário técnico que foi aplicado aos funcionários e monitores de pesquisa que utilizaram o sistema dentro do ISAM.

O formulário foi desenvolvido em formato *web* e foi acoplado junto com o módulo de avaliação e controle das notas. Para uma avaliação mais imparcial, o mesmo foi construído de forma anônima. O formulário aplicado consta no Anexo VII.

Para estes testes foram utilizados dados fictícios e outros reais dentro de um mesmo ambiente. Esses dados fictícios simularam as possibilidades de entrada e ajudaram a explorar a lógica e o entendimento do sistema.

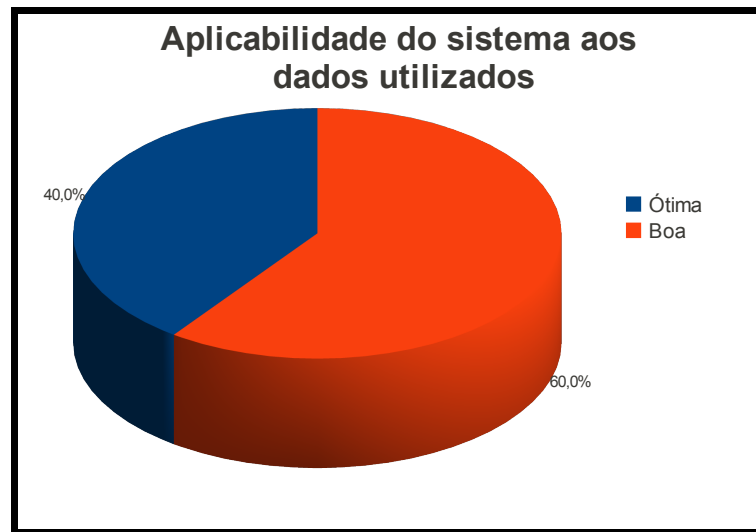
A avaliação foi enviada para 19 usuários potenciais do sistema sendo que destes 15 responderam até o dia 5 de julho de 2010 as 21:13, data do encerramento da coleta de dados para a avaliação. O formulário foi submetido a 5 professores do Instituto, 1 Engenheiro ambiental, 2 biólogas, 10 alunos da engenharia ambiental (bolsistas do ISAM), 2 alunos de biologia também bolsistas do ISAM.

A apresentação dos resultados da avaliação está apresentado nas sessões 6.2.1 referente a avaliação da aplicabilidade e conformidade do sistema, o item 6.2.2 apresenta a avaliação da usabilidade do sistema e por fim o item 6.2.3 apresenta a opinião dos avaliadores quanto a confiabilidade e importância das regras.

### 6.2.1 AVALIAÇÃO DE APLICABILIDADE E CONFORMIDADE

Os resultados da questão 1 sobre aplicabilidade “O sistema tem aplicabilidade prática nos dados que normalmente são utilizados?” são apresentados no gráfico apresentado na Ilustração 35. A aplicabilidade obteve um bom resultado já que 6 pessoas acharam uma ótima aplicabilidade e 9 acharam boa, sendo assim o sistema teve uma boa aceitação quanto a aplicabilidade.

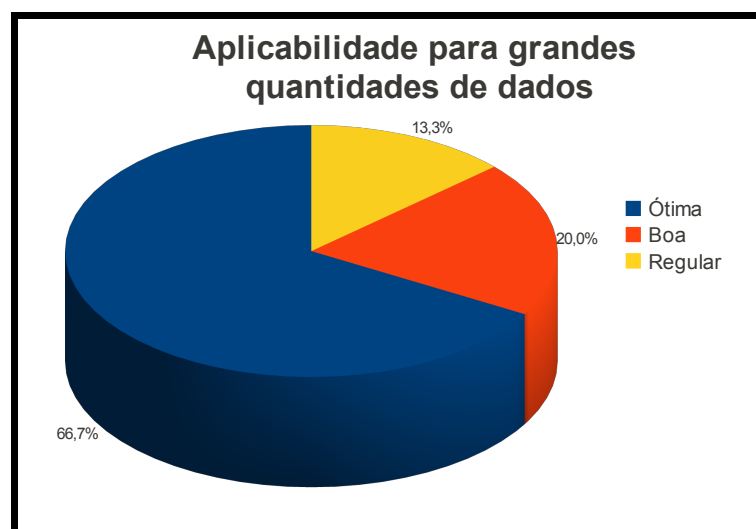




**Ilustração 35 - Gráfico da aplicabilidade do sistema aos dados utilizados**

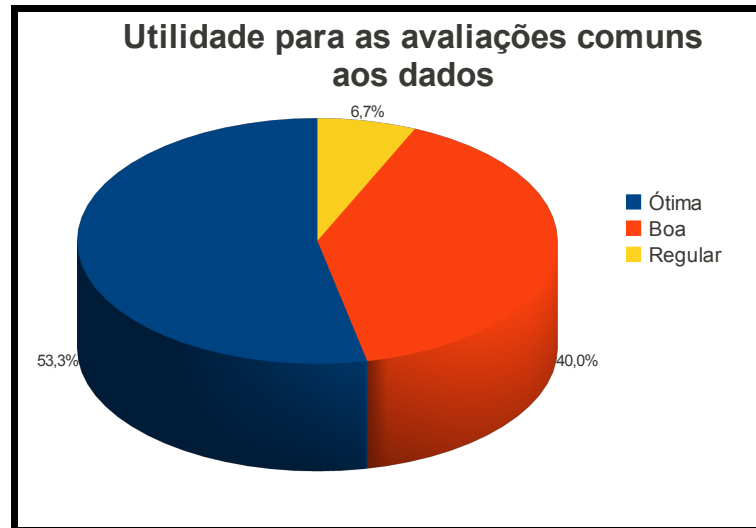
Os resultados da questão 2 sobre aplicabilidade “Para uma grande quantidade de dados classifique a aplicabilidade?” são apresentados no gráfico apresentado na Ilustração 36. Dentre as 15 pessoas que avaliaram o sistema 10 pessoas acharam ótima, sendo assim a solução é considerada aplicável para grande quantidade de dados. Porém em larga escala pode-se encontrar mais problemas para serem avaliados. É importante ressaltar que já se tem como trabalho futura uma separação de dados para cada avaliador diminuindo assim a quantidade de dados para cada pessoa avaliar.

Para essa questão foi sugerido que além de dividir a avaliação de dados, fosse feito um resumo das regras aplicadas ao dado na primeira tela para poder otimizar a tarefa de análise e ganhar tempo.



**Ilustração 36 - Gráfico da aplicabilidade para grandes quantidades de dados**

Os resultados da questão 3 sobre aplicabilidade “Classifique a utilidade para as avaliações que normalmente se faz aos dados?” são apresentados no gráfico apresentado na Ilustração 37. A aplicabilidade quanto as avaliações comuns aos dados também teve uma boa aceitação



**Ilustração 37 - Gráfico utilidade para as avaliações comuns aos dados**

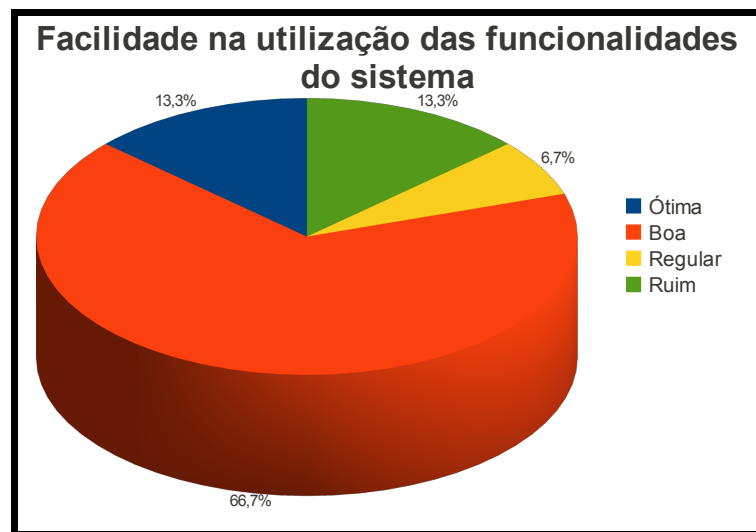
### **6.2.2 AVALIAÇÃO DE USABILIDADE**

Os resultados da questão 1 sobre a usabilidade do sistema “Classifique a facilidade com que você conseguiu aprender as funcionalidades do sistema? ” são apresentados no gráfico apresentado na Ilustração 38. A facilidade do sistema apresentou-se um pouco precária sendo que apenas uma pessoa achou ótima e houve 4 pessoas achando regular, uma péssima e outra ruim. Uma sugestão foi a criação de um menu de ajuda e perguntas comuns (FAQ).



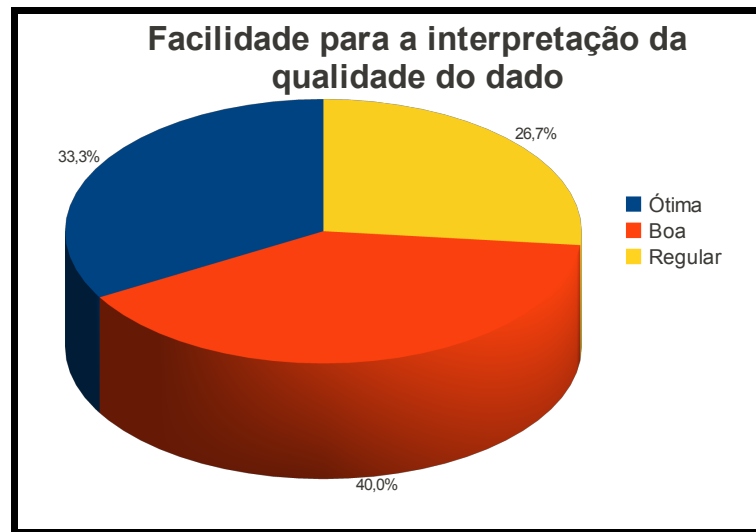
**Ilustração 38 - Gráfico da facilidade de aprendizado do sistema**

Os resultados da questão 2 sobre usabilidade do sistema “Classifique a facilidade com que você conseguiu usar as funcionalidades do sistema?” são apresentados no gráfico apresentado na Ilustração 39. Aqui nessa questão também foi sentida a necessidade de um maior conhecimento do sistema, para poder efetuar as alterações sugeridas pelas regras.



**Ilustração 39 - Gráfico da facilidade de utilização do sistema**

Os resultados da questão 3 sobre usabilidade do sistema “Classifique a facilidade com que você conseguiu entender os resultados da qualidade dos dados?” são apresentados no gráfico apresentado na Ilustração 40. Alguns dos avaliadores aqui sugeriram que o sistema deveria mostrar mais precisamente quais foram as regras e notas dadas a cada dado. Foi sugerido também uma forma de organizar por pior nota e por campanha, criando filtros para a seleção dos dados.



**Ilustração 40 - Gráfico de facilidade para a interpretação da qualidade do dado**

### 6.2.3 AVALIAÇÃO DE CONFIABILIDADE

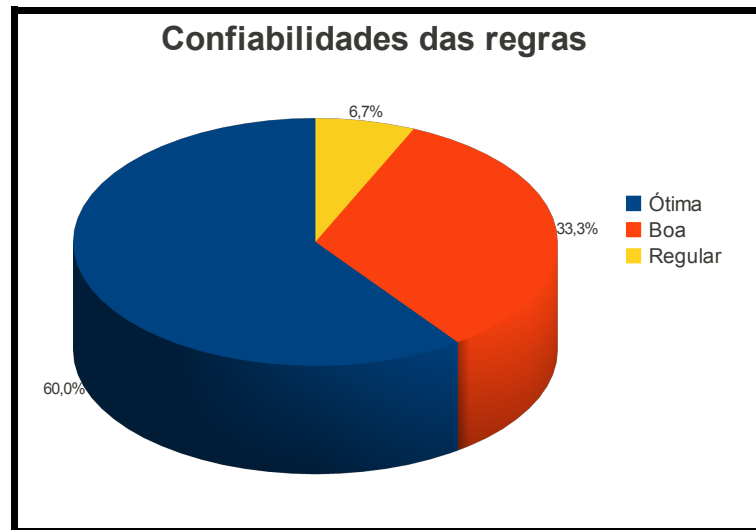
Os resultados da questão 1 sobre a confiabilidade do sistema “Classifique a confiança que o sistema lhe passou em aplicar a nota dos dados” são apresentados no gráfico apresentado na Ilustração 41. Os resultados mostram que a confiança no sistema foi boa, porém ainda assim é necessário fazer algumas melhorias. Como dica foi sugerido explicar um pouco mais as regras do sistema e apresentar o gráfico de controle de forma visual.



**Ilustração 41 - Gráfico da confiança da avaliação que o sistema realiza sobre os dados**

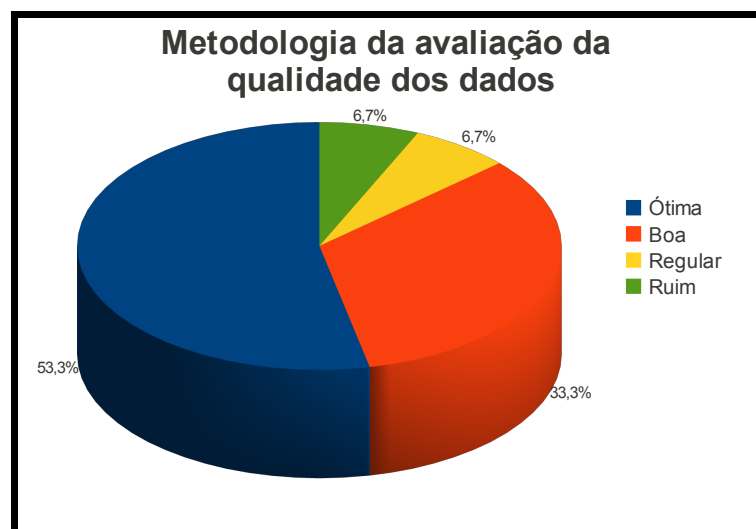
Os resultados da questão 2 sobre a confiabilidade do sistema “Classifique a confiabilidade das regras” são apresentados no gráfico apresentado na Ilustração 42. As regras se mostraram confiáveis, sendo que 60% acharam confiável e apenas uma pessoa achou

regular as regras.



**Ilustração 42 - Gráfico da confiabilidade das regras**

Os resultados da questão 3 sobre a confiabilidade do sistema “Sua opinião sobre a metodologia para avaliar a qualidade do dado” são apresentados no gráfico apresentado na Ilustração 43. Essa questão apresentou um percentual menor de pessoas que acharam ótima a avaliação do sistema, sendo que um dos questionamentos apresentados pela pessoa que deu a nota ruim é que deveria ter alguma forma de explicar que desconhece o motivo pelo qual o dado não atingiu uma boa nota.



**Ilustração 43 - Gráfico sobre a metodologia da avaliação da qualidade dos dados**

### 6.3 IMPLANTAÇÃO

O sistema foi implantado em uma nova base de dados. Esse fato se deu por dois motivos:

1. Para não gerar uma sobrecarga de trabalho para os técnicos avaliadores;
2. Para o sistema ir se adaptando as necessidades ao detectar melhorias para a análise da qualidade dos dados.

Para o sistema implantado no ISAM, os dados estão sendo carregados através de *scripts*. Esses *scripts* carregam dados de planilhas ou fazem a cópia de informações já inseridas em bases de projetos já finalizados ou em andamento.

## 7 RESULTADO E DISCUSSÕES FINAIS

A avaliação da qualidade de dados é uma tarefa que exige o levantamento de requisitos que um dado precisa ter para atingir os objetivos do usuário do dado. Os requisitos básicos foram levantados e, para alcançar os objetivos propostos, foram construídos alguns artifícios que foram adicionados ao modelo de dados. Esses artifícios foram sistematicamente adicionados conforme o estudo das dimensões de qualidade de dados.

### 7.1 OBJETIVOS ALCANÇADOS

A partir da revisão bibliográfica, inicialmente o sistema funcionaria de forma mais simples com implementações das regras direto por *Trigger*. Porém, o estudo dos parâmetros mostrou a necessidade de algo mais dinâmico. Para tal, a proposta apresentada no Capítulo 4 referente ao item livre de erros foi uma das soluções encontradas. Dessa forma, as regras serão armazenadas no sistema e, a cada nova regra, os dados podem ser revalidados e os novos dados já terão uma confiabilidade maior.

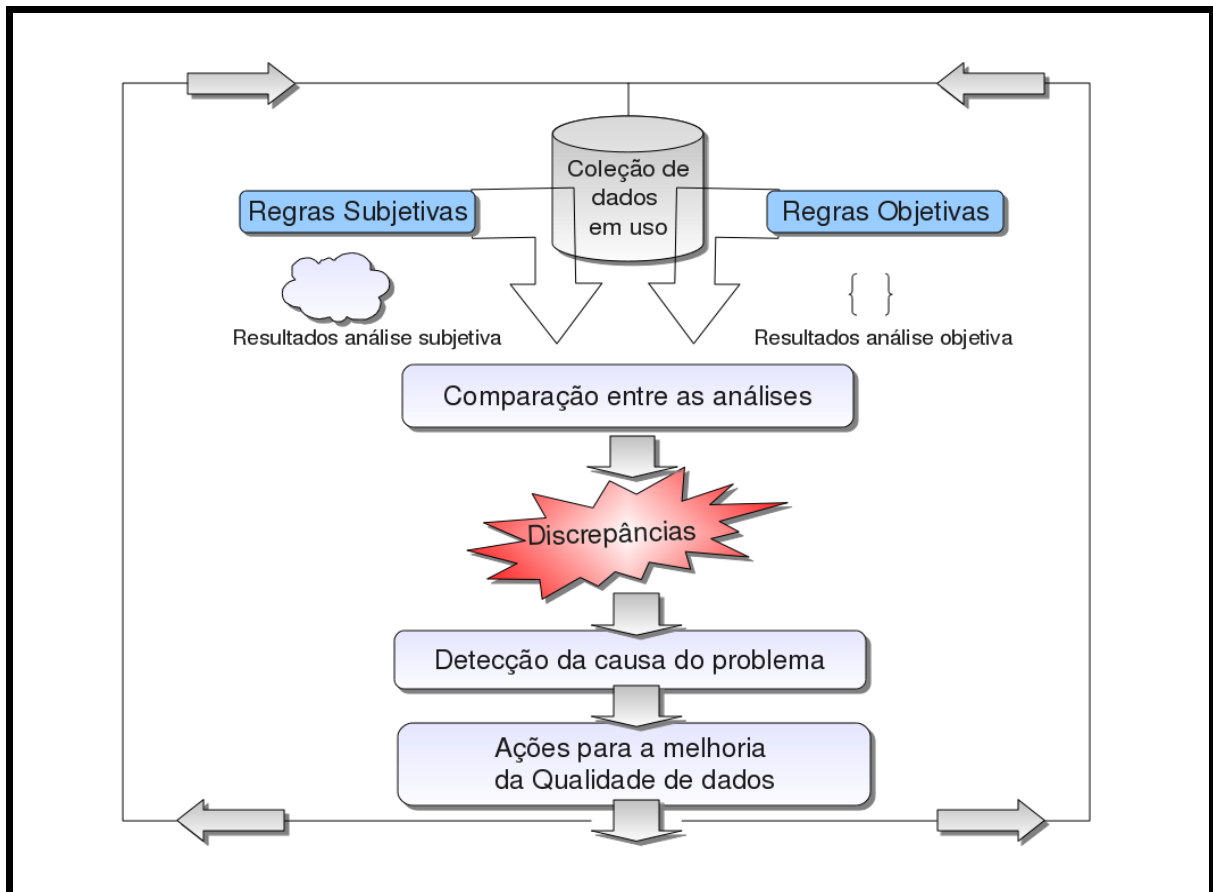
Dentro do cronograma inicial, a definição do fluxo de inserção dos dados, que foi definido ainda na parte 1 do trabalho, foi construída com a necessidade da interação humana caso o dado não passasse numa avaliação pelo sistema. Essa validação humana também será solicitada caso haja um indício de alguma não conformidade no ponto que está sendo avaliado. A validação humana nesses casos é necessária para a explicação do dado, ou porque, naquele tempo, o dado obteve um valor indicando problemas com o corpo hídrico.

Outro objetivo atingido foi definir uma escala/índice para a qualidade do dado. O índice de qualidade do dado permite que se tenha uma noção exata de quão confiável é o dado segundo as regras impostas pelo sistema. Esse índice é, basicamente, uma média ponderada, pois algumas regras poderão ser mais importantes do que outras, e isso a escala deve levar em consideração.

Associar informações aos dados a fim de facilitar uma melhor compreensão do dado também foi determinante para a construção do fluxo de trabalho de um dado resultante. O modelo de dados apresentado possui algumas dessas informações.

A busca da qualidade, porém, não pode ser estanque. A criação de novas regras para o sistema deverá continuar pós trabalho. A Ilustração 44 apresenta um esquema para a melhoria

da qualidade de dados, proposto por Pepino (2002). Ainda para Pepino (2002), a causa dos problemas, sempre que detectada, deve ser documentada e criadas novas regras para que estes problemas não venham a acontecer.



**Ilustração 44 - Esquema para o melhoramento contínuo da qualidade dos dados**

Fonte: Adaptado de (Pipino *et al*, 2002)

A implementação do sistema ainda se manteve como a proposta original, sendo que a principal avaliação da qualidade do dado é realizada direto no SGBD. Esse requisito, citado na proposta, ajudou a garantir que o banco de dados assegure que todos os dados que forem inseridos no sistema sejam auditados.

A necessidade da avaliação subjetiva gerou a construção de uma interface que permitiu uma melhor avaliação e acompanhamento da qualidade dos dados que o banco de dados possui. Essa ferramenta deve ser continuada sendo que novas funcionalidades deverão ser implementadas como propostas futuras.



## 7.2 TRABALHOS FUTUROS

Dentre os trabalhos futuros, pretende-se criar mais um módulo chamado de módulo de auditoria e tomada de decisão, propõe-se algumas funcionalidades e objetivos a serem atingidos:

- Gerar um registro de problemas encontrados em metodologia, laboratórios ou inserção de dados
- Gerar gráficos de controle para os dados inseridos no sistema.
- Avaliação mais apurada de falsos positivos.
- Integrar a sistemas de captura de dados em tempo real, para dessa forma obter relações mais apuradas.
- Integrar a visualização em um *webmapa* de forma a poder avaliar o dado em relação ao ponto.
- Criar uma forma de escalonamento dos dados pendentes por avaliador diminuindo assim a sobrecarga para um avaliador.

Dentre as novas regras a serem implementadas está a modelagem de regras que contemplem:

- Criar regras integradas ao uso e ocupação do solo por microbacia.
- Correlacionar regras com dados pluvio e fluviométricos avaliando o valor do parâmetro em relação ao clima.

## 8 REFERÊNCIAS

ABATE, Marcey L., DIEGERT Kathleen V., **A Hierarchical Approach to Improving Data Quality.** Data Quality Journal, v. 4, n. 1, Set. 1998, Disponível em: <<http://www.dataquality.com/998abate.htm>>. Acesso em: 22 de nov. 2009

AMARAL, G. **AQUAWARE: Um Ambiente de Suporte à Qualidade de Dados em Data Warehouse.** Dissertação de Mestrado. UFRJ, Rio de Janeiro, 2003. 165 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELECTROTECHNICAL COMMISSION. **Engenharia de software: qualidade do produto (NBR ISO/IEC 9126-1)**: parte 1 : modelo de qualidade. Rio de Janeiro: ABNT, 2003. 21 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELECTROTECHNICAL COMMISSION. **Tecnologia de informação – Técnicas de segurança – Código prática para a gestão da segurança da informação (NBR ISO/IEC 27002:2005)**: . Rio de Janeiro: ABNT, 2005. 140 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Planejamento de amostragem de efluentes Líquidos e corpos receptores (NBR 9897)**:.procedimento. Rio de Janeiro: ABNT, 1987. 14p.

ASSOCIAÇÃO MERCOSUL DE NORMALIZAÇÃO. **Gestão da Qualidade e Garantia da qualidade - Terminologia (NM-ISO 8402:97)**:.terminologia. São Paulo: AMN, 1997. 25p.

AZAMBUJA, Telmo Travassos de. **Documentação de sistemas da qualidade: um guia prático para a gestão das organizações.** Rio de Janeiro: Campus, 1996. 283 p. ISBN 8535200770

BIGOLIN, Marcio; BORTOLIN, Taison Anderson; SCHNEIDER, Vania Elisabete; FINOTTI, Alexandra R.. **Programa para cálculo de Índice da Qualidade da Água (IQA)** . 1º Congresso Internacional de Tecnologia para o Meio Ambiente. Bento Gonçalves, RS .out de 2008

BRASIL. Conselho Nacional do Meio Ambiente – CONAMA, **RESOLUÇÃO Nº 357, DE 17 DE MARÇO DE 2005.** Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. Disponível em: <<http://www.mma.gov.br/port/conama/res/res05/res35705.pdf>>. Acesso em: 15 set. 2009.

CETESB, Companhia Ambiental do Estado de São Paulo. **IQA Índice de qualidade das águas.** 24 de nov 2001. Disponível em: <[http://www.cetesb.sp.gov.br/Agua/rios/indice\\_iap\\_iqa.asp](http://www.cetesb.sp.gov.br/Agua/rios/indice_iap_iqa.asp)>. Acesso em: 24 de nov 2009

CETESB, Companhia Ambiental do Estado de São Paulo . **Variáveis de qualidade das águas**. 24 de nov 2001. Disponível em: <<http://www.cetesb.sp.gov.br/Agua/rios/variaveis.asp>>. Acesso em: 24 de nov 2009

CHOY, O., LIM, J., NA, H. S., BAIK, O. **An Efficient Method of Data Quality using Quality Evaluation Ontology**. Third 2008 International Conference on Convergence and Hybrid Information Technology. Disponível em: <<http://ieeexplore.ieee.org>>. Acesso em: 25 de set. 2009.

CROSBY, Philip B.. **Quality is free: the art of making quality certain**. New York: McGraw-Hill, c1979. 309 p. ISBN 0070145121

DAJOZ, Roger. **Princípios de ecologia**. 7.ed. Porto Alegre: Artmed, 2005. 519 p. ISBN 8536305657

ELMASRI, Ramez; NAVATHE, Shamkant B.. **Sistemas de banco de dados**. 4.ed. São Paulo: Addison-Wesley, 2005. 724 p. ISBN 8588639173

ESTEVEZ, Francisco de Assis. **Fundamentos de limnologia**. 2.ed. Rio de Janeiro: Interciencia: 1998. 602p.

FERREIRA, Aurélio B. H, **Miniaurélio Eletrônico versão 5.12**. 7ª ed. Paraná: Positivo, 2004. CD-ROM.

FINOTTI, Alexandra R.; FINKLER, Raquel, SILVA; Maurício D'Agostini; CEMIN Gisele. Colaboradores SHNEIDER, Vania Elisabete; NASCIMENTO FILHO, Irajá; GRACIOLLI, Odacir Deonísio; BIGOLIN, Marcio; BALTAR, Marcos. **Monitoramento de Recursos Hídricos em Áreas Urbanas**. Caxias do Sul, RS: EDUCS, 2009. 272 p. ISBN 978-85-7061-554-1

GIOVANNI, José Ruy; BONJORNO, José Roberto; GIOVANNI JR., José Ruy. **Matemática fundamental**, 2º grau: volume único. São Paulo; FTD, 1994. 560 p. ISBN 85-322-2749-X

GOVERNO BRASILEIRO, **e-PING Padrões de Interoperabilidade de Governo Eletrônico** , Documento de Referência Versão 4.0 , out 2008. Disponível em: <<https://www.governoeletronico.gov.br/acoes-e-projetos/e-ping-padroes-de-interoperabilidade>>. Acesso em: 27 de nov. 2009.

HASENACK, H.; WEBER, E. (Org.). **Base cartográfica digital da Serra Gaúcha**. Escala 1: 50.000. Porto Alegre: UFRGS - Centro de Ecologia, 2007. 1CD-ROM. (Série Geoprocessamento, 2);

ISAM. **Projeto Gestão dos Recursos Hídricos na Bacia Hidrográfica do Rio das Antas: disponibilidade, qualidade, usos e impactos decorrentes**. Universidade de Caxias do Sul – Caxias do Sul. 2008

ISO/TS **8000-120:2009**. Catálogo de compras. Disponível em: <[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=50801](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50801)> . Acesso em: 24 de nov 2009.

**ISO/TS 8000-110:2009.** Catálogo de compras. Disponível em: <[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=51653](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=51653)>. Acesso em: 24 de nov 2009.

KIMBALL, Ralph; ROSS, Margy. **The data warehouse toolkit: guia completo para modelagem dimensional.** Rio de Janeiro, RJ: Campus, 2002. 494 p. ISBN 8535211292

LAUDON, Kenneth C.; LAUDON, Jane Price. **Sistemas de informação gerenciais.** 7.ed. São Paulo: Pearson Prentice Hall, 2007. 452 p.

LIBÂNIO, Marcelo. **Fundamentos de qualidade e tratamento de água.** Campinas, SP: Átomo, 2005. 444 p. ISBN 8576700301

MOTA, S. **Preservação e conservação de recursos hídricos.** Rio de Janeiro (RJ): Associação Brasileira de Engenharia Sanitária e Ambiental, 1995.

MONTGOMERY, Douglas. C., 1996, **Introduction to Statistical Quality Control**, 3 ed. USA. MONTGOMERY, Douglas C.. **Introdução ao controle estatístico da qualidade.** 4.ed. Rio de Janeiro, RJ: LTC, 2004. 513 p. ISBN 8521614004

NATIONAL SECURITY AGENCY (NSA), **Guide to the secure Configuration of The Red Hat Enterprise Linux 5.** Revisão 3. Outubro de 2009, Disponível em: <[http://www.nsa.gov/ia/guidance/security\\_configuration\\_guides/operating\\_systems/linux.shtml](http://www.nsa.gov/ia/guidance/security_configuration_guides/operating_systems/linux.shtml)>. Acesso em: 20 de junho de 2010

OLIVEIRA, Halley Pacheco, **PL/Ruby - Linguagem procedural Ruby**, 2006 Disponível em: <<http://pgdocptbr.sourceforge.net/pg80/plruby.html>>. Acesso em: 1 de fevereiro de 2010

PINHO, Selma Foline Crespio, **Avaliação da Qualidade de Dados pela não Conformidade.** Dissertação de Mestrado. UFRJ, Rio de Janeiro, 2000. 122 p.

PIPINO, Leo L.; LEE, Yang W.; WANG, Richard Y., **Data Quality Assessment.** Communications of the ACM, Abril 2002.

PRESSMAN, Roger S. **Engenharia de Software.** São Paulo, SP. Pearson Makron Books, 1995. 1056pg. ISBN 85-346-0237-9

PRESSMAN, Roger S.. **Engenharia de Software.** 6.ed. São Paulo: McGraw-Hill, 2006. 720 p. ISBN 8586804576

PROTECTION AGENCY, U. S. **Environmental. Glossary of Quality Assurance Terms and Related Acronyms. 10 dez. 1997,** Disponível em: <[http://www.hanford.gov/dqo/glossaries/Glossary\\_of\\_Quality\\_Assurance\\_Terms1.pdf](http://www.hanford.gov/dqo/glossaries/Glossary_of_Quality_Assurance_Terms1.pdf)>. Acesso em: 8 de set. 2009.

SOMMERVILLE, Ian. **Engenharia de software.** 8.ed. São Paulo: Pearson Addison Wesley, c2007. xiv, 552 p. : ISBN 9788588639287

VON SPERLING, M. **Estudos e modelagem da qualidade de água em rios.** 1.ed. Belo Horizonte: UFMG, 2007. 588 p. (Princípios do tratamento biológico de águas residuárias; 7)

ISBN 8588556072.

WANG, Richard Y.; KON, Henry B.; MADNICK, Stuart E. **Data Quality Requirements Analysis and Modeling**. Vienna, Austria, Abr. 1993, Disponível em: <<http://web.mit.edu/tdqm/www/tdqmpub/IEEEDEApr93.pdf>>. Acesso em: 1 de set. 2009.

WANG, Richard Y.; STRONG, Diane; GUARASCIO, Lisa . **Beyond Accuracy: What Data Quality Means to Data Consumers** . Journal of Management Information Systems / Spring 1996, Vol. 12, No. 4, pp. 5-34. ,

WAZLAWICK, Raul Sidnei. **Análise e projeto de sistemas de informação orientados a objetos**. Rio de Janeiro: Elsevier, 2004. 298 p. (Campus/SBC) ISBN 8535215646

WILLSHIRE, J. M.; MEYEN, D., **A Process for Improving Data Quality**, Data Quality Journal, v. 3, n. 1, Set. 1997, Disponível em: <<http://www.dataquality.com/997meyen.htm>>. Acesso em: 10 de out. 2009.

## ANEXO I

### Atributos de qualidade de dados

Capacidade para ser integrado	Capacidade de <i>Upload</i>	Acessibilidade
Habilidade para <i>Download</i>	Aceitabilidade	Precisão
Capacidade de identificar erros	Acesso pela Concorrência	Adaptabilidade
Detalhamento adequado	Volume adequado	Estética dos dados
Idade	Possibilidade de agregação	Alterabilidade
Quantidade de dados	Auditáveis	Autoridade
Disponibilidade	Credibilidade	Amplitude dos dados
Amplitude dos dados	Brevidade	Dados certificados
Clareza	Clareza da origem	Dados livre de responsabilidade
Compacidade	Compatibilidade	Vantagem competitiva
Integralidade	Abrangência	Compressibilidade
Conciso	Concisão	Confidencialidade
Conformidade	Consistência	Conteúdo
Contexto	Continuidade	Conveniência
Exatidão	Sem Corrupção	Custo
Custo de precisão	Custo de cobrança	Criatividade
Crítico	Atual	Personalizável
Hierarquia de dados	Melhoramento da eficiência	Sobrecarregáveis
Definibilidade	Confiabilidade	Profundidade de dados
Dados detalhados	Fonte detalhada	Pouca dispersão
Distinção das atualizações	Dinâmicos	Facilidade de acesso
Facilidade de comparação	Facilidade de correlação	Facilidade de troca de dados
Facilidade de manutenção	Facilidade de obtenção	Fácil de entender
Facilidade de atualização	Facilidade de troca	Fácil de perguntar
Eficiência	Sobrevivência do dado	Esclarecedor
Ergonômico	Livre de erros	Expansibilidade
Pouca despesa	Estendibilidade	Extensibilidade
Extensão	Finalização	Perfeição

Flexibilidade	Forma de apresentação	Formato
Integridade	Amizade	Generalidade
Hábito	Histórico de Compatibilidade	Importância
Livre de inconsistência	Integração	Integridade
Interativo	Interessante	Nível de abstração
Nível de normalização	Localização	Logicamente Conectado
Gerenciável	Manipulável	Mensuráveis
Mediano	Atende aos requisitos	Minimalidade
Modularidade	Estritamente definidos	Sem informações perdidas
Normalidade	Novidade	Objetividade
Otimizados	Ordem	Origem confiáveis
Parcimônia	Particionalidade	Experiência do passado
Pedigree	Personalizado	Pertinente
Portabilidade	Precisão	Integridade proprietária
Propósito	Quantidade	Racionalidade
Exatidão	Sem redundância	Regularidade no formato
Pertinência	Confiabilidade	Sem Repetição
Reprodutibilidade	Reputação	Resolução para Gráficos
Responsabilidade	Recuperabilidade	Auto-críticos
Rigidez	Robustez	Escopo da informação
Sigilo	Seguro	Auto-correção
Interpretação semântica	Semântica	Tamanho
Possui fonte	Especificidade	Velocidade
Estabilidade	Fácil armazenamento	Sincronizável
Atual	Rastreável	Traduzível
Transportável	Sem ambiguidades	Imparcial
Compreensível	Organizado	Útil
Validade	Bem documentado	Bem apresentado

Fonte: Traduzido de (WANG *et al*, 1996)

## ANEXO II

## Parâmetros contidos no sistema

Descrição do Parâmetro	Abreviatura Unidade Medida	Descrição da Unidade de Medida
2,4 D	ug/L	Micro grama por Litro
3,4 DCBTF	ug/L	Micro grama por litro
Acidez	mg /L	Miligrama por litro
Agrotóxicos Organoclorados	ug/L	Micro grama por litro
Agrotóxicos Organofosforados	ug/L	Micro grama por litro
Alcalinidade	mg CaCO <sub>3</sub> /L	Miligrama carbonato de cálcio por litro
Alcalinidade a bicarbonatos	mg CaCO <sub>3</sub> /L	Miligrama carbonato de cálcio por litro
Aldrin	ug/L	Micro grama por litro
Alumínio	mg /L	Miligrama por litro
Amônia não ionizável	Sem Unidade	Sem Unidade
Amônio	mg /L	Miligrama por litro
Antimônio	mg /L	Miligrama por litro
AOX	ug/L	Micro grama por litro
AOX Klabin	Kg/T	kg/tonelada
Arsênio	mg /L	Miligrama por litro
Bactérias Heterotróficas	Sem Unidade	Sem Unidade
Bário	mg /L	Miligrama por litro
Benzeno	ug/L	Micro grama por litro
Benzo-a-pireno	ug/L	Micro grama por litro
Berílio	mg /L	Miligrama por litro
BHC	ug/L	Micro grama por litro
Biocida	Sem Unidade	Sem Unidade
Bismuto	mg /L	Miligrama por litro
Boro	mg /L	Miligrama por litro
Borofórmio	ug/L	Micro grama por litro
Bromoketal	ug/L	Micro grama por litro
Cadmio	mg /L	Miligrama por litro
Cálcio	mg /L	Miligrama por litro
Carbamato	ug/L	Micro grama por litro
Carbono orgânico total	mg /L	Miligrama por litro
Ceriodaphnia dubia-TOX aguda	fat dil (SD)	Fator de diluição - SD
Ceriodaphnia dubia-TOX crônica	fat dil (SD)	Fator de diluição - SD
Chlorpiriphos	ug/L	Micro grama por litro
Chumbo	mg /L	Miligrama por litro
Cianeto	mg CN <sup>1-</sup>	Miligrama de cianeto por litro
Clorato	mg /L	Miligrama por litro
Clordane	ug/L	Micro grama por litro
Cloreto	mg CL <sup>1-</sup> /L	Miligrama de cloreto por litro
Cloreto Klabin	T/dia	Tonelada/dia



Descrição do Parâmetro	Abreviatura Unidade Medida	Descrição da Unidade de Medida
Cloro residual	mg /L	Miligrama por litro
Clorofórmio	ug/L	Micro grama por litro
Cobalto	mg /L	Miligrama por litro
Cobre	mg /L	Miligrama por litro
Coliformes Termotolerantes	NMP/100mL	Número mais provável em 100 ml
Coliformes totais	NMP/100mL	Número mais provável em 100 ml
Compostos clorofenolicos	ug/L	Micro grama por litro
Compostos organoclorados	ug/L	Micro grama por litro
Compostos organofosforados	ug/L	Micro grama por litro
Condutividade	µmhos/cm	Micro mhos/cm
Cor	mg/L Pt	Miligrama por litro de platina
Cor aparente	mg/L Pt	Miligrama por litro de platina
Cromo hexavalente	mg /L	Miligrama por litro
Cromo total	mg /L	Miligrama por litro
Cromo trivalente	mg /L	Miligrama por litro
Daphnia magna - TOX aguda	fat dil (SD)	Fator de diluição - SD
Daphnia similis -TOX aguda	fat dil (SD)	Fator de diluição - SD
DBO5	mg DBO5/L	Miligrama de DBO5 por litro
DCA	mg /L	Miligrama por litro
DDT	ug/L	Micro grama por litro
DDVP + TRICLORFON	ug/L	Micro grama por litro
Dicloroanilina	ug/L	Micro grama por litro
Dieldrin	ug/L	Micro grama por litro
Dinitro (3,5 dinitro paraclorobenzotrifluoreto)	ug/L	Micro grama por litro
Dióxido de enxofre	ug/m3	Micrograma por metro cubico
Dióxido de nitrogênio	ug/m3	Micrograma por metro cubico
Dioxinas	ug/L	Micro grama por litro
DIURON + DCPI	ug/L	Micro grama por litro
DQO	mg DQO/L	Miligrama de DQO por litro
Dureza	mg CaCO3/L	Miligrama carbonato de cálcio por litro
E. coli	fat dil (SD)	Fator de diluição - SD
EH	mV	Mili volts
Endosulfan	ug/L	Micro grama por litro
Endrin	ug/L	Micro grama por litro
EOX	ug/L	Micro grama por litro
Escherichia coli	NMP/100mL	Número mais provável em 100 ml
Espumas	Sem Unidade	Sem unidade
Estanho	mg /L	Miligrama por litro
Estireno	ug/L	Micro grama por litro
Estrôncio	mg /L	Miligrama por litro
Etilbenzeno	ug/L	Micro grama por litro
Ferro	mg /L	Miligrama por litro

Descrição do Parâmetro	Abreviatura Unidade Medida	Descrição da Unidade de Medida
Fluoreto	mg F <sup>1-</sup>	Miligrama de fluoreto por litro
Fosfato orto	mg PO4-P/L	Miligrama de fosforo por litro
Fosfato total	mg P/L	Miligrama de fosforo por litro
Fósforo total	mg P/L	Miligrama de fosforo por litro
Heptacloro	ug/L	Micro grama por litro
Heptacloro epóxido	ug/L	Micro grama por litro
Hidrocarbonetos	ug/L	Micro grama por litro
Hyalala azteca	fat dil (SD)	Fator de diluição - SD
Índice de fenóis	mg /L	Miligrama por litro
Ketone	ug/L	Micro grama por litro
Lindane	ug/L	Micro grama por litro
Lítio	mg /L	Miligrama por litro
Magnésio	mg /L	Miligrama por litro
Manganês	mg /L	Miligrama por litro
Materiais flutuantes	Sem Unidade	Sem unidade
Material particulado	ug/m3	Micrograma por metro cubico
Mercúrio	ug/L	Micro grama por litro
Metoxicloro	ug/L	Micro grama por litro
Molibdenio	mg /L	Miligrama por litro
Mononitro	ug/L	Micro grama por litro
Monóxido de carbono	ug/m3	Micrograma por metro cubico
Níquel	mg /L	Miligrama por litro
Nitrato	mg NO3-N/L	Miligrama de nitrato em nitrogênio por litro
Nitrito	mg NO2-N/L	Miligrama de nitrito em nitrogênio por litro
Nitrogênio amoniacal	mg NH3-N/L	Miligrama de nitrogênio amoniacal em nitrogênio por litro
Nitrogênio orgânico	mg N/L	Miligrama de nitrogênio por litro
Nitrogênio total	mg /L	Miligrama por litro
Nitrogênio total kjeldahl	mg N/L	Miligrama de nitrogênio por litro
Odor	Sem Unidade	Sem unidade
Óleos e graxas minerais	mg/L OG	Miligrama por litro de óleos e graxas
Óleos e graxas totais	mg/L OG	Miligrama por litro de óleos e graxas
Óleos e graxas vegetais e animais	mg/L OG	Miligrama por litro de óleos e graxas
Orgânicos voláteis	ug/L	Micro grama por litro
Ouro	mg /L	Miligrama por litro
Outras substancias / elementos	Sem Unidade	Sem unidade
Óxidos de carbono	ug/m3	Micrograma por metro cubico
Óxidos de enxofre	ug/m3	Micrograma por metro cubico
Óxidos de nitrogênio	ug/m3	Micrograma por metro cubico
Oxigênio	ug/m3	Micrograma por metro cubico
Oxigênio dissolvido	mg OD/L	Miligrama de oxigênio dissolvido por litro
Oxirane	ug/L	Micro grama por litro
Ozônio	ug/m3	Micrograma por metro cubico

Descrição do Parâmetro	Abreviatura Unidade Medida	Descrição da Unidade de Medida
Partículas inaláveis	ug/m3	Micrograma por metro cubico
Partículas totais em suspensão	ug/m3	Micrograma por metro cubico
Pimephales promelas - TOX CRON	fat dil (SD)	Fator de diluição - SD
Pimephales promelas- tox aguda	fat dil (SD)	Fator de diluição - SD
Policlorobenzotricloreto	ug/L	Micro grama por litro
Policlorobenzotrifluoreto	ug/L	Micro grama por litro
Policlorotolueno	ug/L	Micro grama por litro
Potássio	mg /L	Miligramas por litro
Prata	mg /L	Miligramas por litro
Procloraz	ug/L	Micro grama por litro
Profundidade coleta	m	Metro
Profundidade total	m	Metro
Propanil	ug/L	Micro grama por litro
Propiconazol	ug/L	Micro grama por litro
Salinidade	0/00	Partes por mil
Selenastrum capricornutum	fat dil (SD)	Fator de diluição - SD
Selênio	mg /L	Miligramas por litro
Sílica	mg /L	Miligramas por litro
Silício	mg /L	Miligramas por litro
Sódio	mg /L	Miligramas por litro
Sólidos dissolvidos totais	mg/L SDT	Miligramas por litro de sólidos dissolvidos totais
Sólidos sedimentáveis	mL/L	Mililitro por litro
Sólidos suspensos totais	mg/L SST	Miligramas por litro de sólidos suspensos totais
Sólidos suspensos voláteis	mg/L SSV	miligramas por litro de sólidos suspensos voláteis
Sólidos totais	mg/L ST	Miligramas por litro de sólidos totais
Sólidos totais a 180°C	mg/L ST	Miligramas por litro de sólidos totais
Sulfatação total	mg /L	Miligramas por litro
Sulfato	mg SO4/L	Miligramas de sulfato por litro
Sulfeto total	mg S <sup>2-</sup> /L	Miligramas de sulfeto por litro
Sulfito	mg /L	Miligramas por litro
Surfactantes aniônicos	mg/L LAS	Miligramas por litro las
Tebuconazol	ug/L	Micro grama por litro
Temperatura	° C	Graus célsius
Temperatura da água	° C	Graus célsius
Temperatura do ar	° C	Graus célsius
Tetracloro de carbono	ug/L	Micro grama por litro
Tetracloro eteno	ug/L	Micro grama por litro
Titânio	mg /L	Miligramas por litro
Tolueno	ug/L	Micro grama por litro
Toxafeno	ug/L	Micro grama por litro
Toxicidade	Sem Unidade	Sem unidade
Transparência	Sem Unidade	Sem unidade

Descrição do Parâmetro	Abreviatura Unidade Medida	Descrição da Unidade de Medida
Transparência da água	m	Metro
Trifluralina	ug/L	Micro grama por litro
Trióxido de enxofre	ug/m3	Micrograma por metro cubico
TRS	ug/m3	Micrograma por metro cubico
Turbidez	NTU	Unidade nefelométricas de turbidez
Vanádio	mg /L	Miligramas por litro
Varredura de metais(filtro)	ug/m3	Micrograma por metro cubico
Vazão	m3/dia	Metro cubico por dia
Vazão rio	m3/s	Metro cubico por segundo
Xileno	ug/L	Micro grama por litro
Zinco	mg /L	Miligramas por litro
pH	Unid pH	Unidades ph

Fonte: Adaptado da Resolução CONAMA nº357/2005 e NBR 9897, 1989

## ANEXO III

### Comparação entre os limites definidos para classificação de águas doces, segundo a Resolução CONAMA 357 (2005).

Condições de Qualidade	Classe 1	Classe 2	Classe 3	Classe 4
Crítérios de toxicidade	Sem feitos tóxicos crônicos	Sem efeitos tóxicos crônicos	Sem efeitos tóxicos agudos	-
Materiais flutuantes	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes
Óleos e Graxas	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes	Tolera-se iridescência
Gosto ou cor	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes	-
Odor e aspecto	-	-	-	Não objetável
Corantes provenientes de fontes antrópicas	Virtualmente ausentes	Proibidos os não removíveis por coagulação, sedimentação ou filtração	Proibidos os não removíveis por coagulação, sedimentação ou filtração	-
Resíduos sólidos objetáveis	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes	Virtualmente ausentes
Coliformes termotolerantes (colif./100mL)	200	1.000	2.500	-
Demanda Bioquímica de Oxigênio (mgO <sub>2</sub> /L)	Até 3	Até 5	Até 10	-
Oxigênio Dissolvido (mgO <sub>2</sub> /L)	Não inferior a 6	Não inferior a 5	Não inferior a 4	Superior a 2
Turbidez (UNT)	40	100	100	-
Cor (mg Pt/L)	Nível com natural	Até 75	Até 75	-
pH	6-9	6-9	6-9	6-9
<b>Padrões de Qualidade</b>				
Clorofila a (µg/L)	10	30	60	-
Densidade de cianobactérias (células/mL)	20.000	50.000	100.000 50.000	-
Sólidos Dissolvidos Totais (mg/L)	500	500	500	-
Alumínio dissolvido (mg/L)	0,10	0,10	0,20	-
Antimônio (mg/L)	0,005	0,005	-	-
Arsênio (mg/L)	0,01	0,01	0,033	-
Bário (mg/L)	0,70	0,70	1,00	-
Berílio (mg/L)	0,04	0,04	0,10	-
Boro (mg/L)	0,50	0,50	0,75	-
Cádmio (mg/L)	0,001	0,001	0,01	-
Chumbo (mg/L)	0,01	0,01	0,033	-
Cianetos livres (mg/L)	0,005	0,005	0,022	-
Cloretos (mg/L)	250	250	250	-
Cloro residual total (mg/L)	0,01	0,01	-	-
Cobalto (mg/L)	0,05	0,05	0,20	-

Condições de Qualidade	Classe 1	Classe 2	Classe 3	Classe 4
Cobre dissolvido (mg/L)	0,009	0,009	0,013	-
Cromo (mg/L)	0,05	0,05	0,05	-
Ferro dissolvido (mg/L)	0,30	0,30	5,00	-
Fluoretos (mg/L)	1,40	1,40	1,40	-
Fósforo total – ambientes lênticos (mg/L)	0,02	0,03	0,05	-
Fósforo total – ambientes intermediários (mg/L)	0,025	0,05	0,075	-
Fósforo total – ambientes lóticos (mg/L)	0,10	0,10	0,15	-
Lítio (mg/L)	2,50	2,50	2,50	-
Manganês (mg/L)	0,10	0,10	0,50	-
Mercúrio (mg/L)	0,0002	0,0002	0,002	-
Níquel (mg/L)	0,025	0,025	0,025	-
Nitrato (mg/L)	10,00	10,00	10,00	-
Nitrito (mg/L)	1,00	1,00	1,00	-
Nitrogênio amoniacal (mg/L)	2,00 (7,5<pH<8,0) 1,00 (8,0<pH,8,5) 0,50 (pH>8,5)	2,00 (7,5<pH<8,0) 1,00 (8,0<pH,8,5) 0,50 (pH>8,5)	13,3 (pH<7,5) 5,6 (7,5<pH<8,0) 2,2 (8,0<pH,8,5) 1,0 (pH>8,5)	-
Prata (mg/L)	0,01	0,01	0,05	-
Selênio (mg/L)	0,01	0,01	0,05	-
Sulfato (mg/L)	250	250	250	-
Sulfeto (mg/L)	0,002	0,002	0,30	-
Urânio (mg/L)	0,02	0,02	0,02	-
Vanádio (mg/L)	0,10	0,10	0,10	-
Zinco (mg/L)	0,18	0,18	5,00	-
<b>Substâncias Orgânicas</b>				
Acrilamida (µg/L)	0,50	0,50	-	-
Alacloro (µg/L)	20,00	20,00	-	-
Aldrin+Diieldrin (µg/L)	0,005	0,005	0,03	-
Atrazina (µg/L)	2,00	2,00	2,00	-
Benzeno (mg/L)	0,005	0,005	0,005	-
Benzidina (mg/L)	0,001	0,001	-	-
Benzo(a)antraceno (µg/L)	0,05	0,05	-	-
Benzo(a)pireno (µg/L)	0,05	0,05	0,70	-
Benzo(b)fluoranteno (µg/L)	0,05	0,05	-	-
Benzo(k)fluoranteno (µg/L)	0,05	0,05	-	-
Carbaril (µg/L)	0,02	0,02	70,00	-
Clordano - cis+trans (µg/L)	0,04	0,04	0,03	-
2 – clorofenol (µg/L)	0,10	0,10	-	-
Criseno (µg/L)	0,05	0,05	-	-
2,4 – D (µg/L)	4,00	4,00	30,00	-
Demeton (µg/L)	0,10	0,10	14,00	-

Condições de Qualidade	Classe 1	Classe 2	Classe 3	Classe 4
Dibenzo(a,h)antraceno (µg/L)	0,05	0,05	-	-
1,2-dicloroetano (mg/L)	0,01	0,01	0,01	-
1,1-dicloroetano (mg/L)	0,003	0,003	30,00	-
2,4-diclorofenol (µg/L)	0,30	0,30	-	-
Diclorometano (mg/L)	0,02	0,02	-	-
DDT (µg/L)	0,002	0,002	1,00	-
Dodecacloro pentaciclodecano (µg/L)	0,001	0,001	0,001	-
Endosulfan (µg/L)	0,056	0,056	0,22	-
Endrin (µg/L)	0,004	0,004	0,20	-
Estireno (mg/L)	0,02	0,02	-	-
Etilbenzeno (µg/L)	90,00	90,00	-	-
Fenóis (mg/L)	0,003	0,003	0,01	1,00
Glifosfato (µg/L)	65,00	65,00	280,00	-
Gution (µg/L)	0,005	0,005	0,005	-
Heptacloro epóxido + Heptacloro (µg/L)	0,01	0,01	0,03	-
Hexaclorobenzeno (µg/L)	0,0065	0,0065	-	-
Indeno (1,2,3-cd)pireno (µg/L)	0,05	0,05	-	-
Lindano (µg/L)	0,02	0,02	2,00	-
Malation (µg/L)	0,10	0,10	100,00	-
Metolacloro (µg/L)	10,00	10,00	20,00	-
Metoxicloro (µg/L)	0,03	0,03	-	-
Paration (µg/L)	0,04	0,04	35,00	-
PCBs (µg/L)	0,001	0,001	0,001	-
Pentaclorofenol (mg/L)	0,009	0,009	0,009	-
Simazina (µg/L)	2,00	2,00	-	-
Substâncias tensoativas que reagem ao azul de metileno (mg/L)	0,5	0,5	0,5	-
2,4,5-T (µg/L)	2,00	2,00	2,00	-
Tetracloroeto de carbono (mg/L)	0,002	0,002	0,003	-
Tetracloroetano (mg/L)	0,01	0,01	0,01	-
Tolueno (µg/L)	2,00	2,00	-	-
Toxafeno (µg/L)	0,01	0,01	0,21	-
2,4,5-TP (µg/L)	10,00	10,00	10,00	-
Trbutilestanho (µg/L)	0,0063	0,0063	2,00	-
Triclorobenzeno (mg/L)	0,02	0,02	-	-
Tricloroetano (mg/L)	0,03	0,03	0,03	-
2,4,6-triclorofenol (mg/L)	0,01	0,01	0,01	-
Trifluralina (µg/L)	0,20	0,20	-	-
Xileno (µg/L)	300,00	300,00	-	-

Fonte: Adaptado da Resolução CONAMA nº357/2005

## ANEXO IV

### Código SQL gerado para a criação do banco de dados.

```

CREATE SCHEMA qualidadeagua;
-- SCHEMA qualidadeagua
CREATE TABLE qualidadeagua.laboratorio (
  nome_laboratorio VARCHAR NOT NULL,
  cnpj VARCHAR NULL,
  endereco VARCHAR NULL,
  telefone VARCHAR NULL,
  endereco_web VARCHAR NULL,
  municipio VARCHAR NULL,
  estado VARCHAR NULL,
  PRIMARY KEY(nome_laboratorio)
);

CREATE TABLE qualidadeagua.unidade (
  nome_unidade VARCHAR NOT NULL,
  descricao_unidade VARCHAR,
  unidade_html VARCHAR,
  PRIMARY KEY(nome_unidade)
);

CREATE TABLE qualidadeagua.metodo_analise (
  id_metodo_analise serial NOT NULL,
  nome VARCHAR NOT NULL,
  descricao VARCHAR NULL,
  fonte_descricao VARCHAR NULL,
  codigo VARCHAR NULL,
  PRIMARY KEY(id_metodo_analise)
);

CREATE TABLE qualidadeagua.metodo_coleta (
  id_metodo_coleta serial NOT NULL,
  nome VARCHAR NOT NULL,
  descricao VARCHAR NULL,
  fonte_descricao VARCHAR NULL,
  norma_aplicada VARCHAR NULL,
  PRIMARY KEY(id_metodo_coleta)
);

CREATE TABLE qualidadeagua.condicao_climatica (
  condicao_climatica VARCHAR NOT NULL,
  descricao_longa VARCHAR NULL,
  PRIMARY KEY(condicao_climatica)
);

CREATE TABLE qualidadeagua.classe (
  classe VARCHAR NOT NULL,
  descricao_classe VARCHAR NULL,
  fonte_descricao VARCHAR NULL,
  PRIMARY KEY(classe)
);

CREATE TABLE qualidadeagua.bacia (
  nome_bacia VARCHAR NOT NULL,
  bacia_superior VARCHAR NULL,
  codigo_bacia INTEGER NULL,
  descricao_bacia VARCHAR NULL,
  comite VARCHAR NULL,
  comite_site VARCHAR NULL,
  PRIMARY KEY(nome_bacia),
  FOREIGN KEY(bacia_superior)
  REFERENCES qualidadeagua.bacia(nome_bacia)
  ON DELETE NO ACTION ON UPDATE CASCADE
);

CREATE TABLE qualidadeagua.ponto (
  id_ponto VARCHAR(30) NOT NULL,
  nome_ponto VARCHAR NOT NULL,
  nome_original_ponto VARCHAR NOT NULL,
  bacia VARCHAR NOT NULL,

```



```

nome_rio VARCHAR NULL,
area_drenagem FLOAT NULL,
altitude FLOAT NULL,
municipio VARCHAR NULL,
profundidade FLOAT NULL,
descricao_ponto VARCHAR NULL,
PRIMARY KEY(id_ponto),
FOREIGN KEY(bacia)
  REFERENCES qualidadeagua.bacia(nome_bacia)
  ON DELETE NO ACTION ON UPDATE CASCADE
);

CREATE TABLE qualidadeagua.imagens_ponto (
  id_imagem serial NOT NULL,
  id_ponto VARCHAR(30) NOT NULL,
  imagem bytea NULL,
  mime_imagem VARCHAR NULL,
  data_imagem DATE NULL,
  descricao_imagem VARCHAR NULL,
  fotografia VARCHAR NULL,
  PRIMARY KEY(id_imagem),
  FOREIGN KEY(id_ponto)
    REFERENCES qualidadeagua.ponto(id_ponto)
    ON DELETE CASCADE ON UPDATE CASCADE
);

CREATE TABLE qualidadeagua.parametro (
  nome_parametro VARCHAR NOT NULL,
  classe VARCHAR NOT NULL,
  descricao VARCHAR NULL,
  unidade_padrao VARCHAR NULL,
  obrigatoriedade_fepam BOOL NULL,
  aconselhavel BOOL NULL,
  PRIMARY KEY(nome_parametro),
  FOREIGN KEY(classe)
    REFERENCES qualidadeagua.classe(classe)
    ON DELETE NO ACTION ON UPDATE CASCADE,
  FOREIGN KEY(unidade_padrao)
    REFERENCES qualidadeagua.unidade(nome_unidade)
    ON DELETE NO ACTION ON UPDATE CASCADE
);

CREATE TABLE qualidadeagua.parametro_medido (
  parametro VARCHAR NOT NULL,
  data_coleta TIMESTAMP NOT NULL,
  id_ponto VARCHAR(30) NOT NULL,
  id_metodo_coleta integer,
  id_metodo_analise integer NOT NULL,
  nome_laboratorio VARCHAR NOT NULL,
  responsavel_coleta VARCHAR,
  valor VARCHAR NOT NULL,
  unidade VARCHAR NOT NULL,
  dia_coleta NUMERIC(2,0) NOT NULL,
  mes_coleta NUMERIC(2,0) NOT NULL,
  ano_coleta NUMERIC(4,0) NOT NULL,
  condicao_climatica VARCHAR NOT NULL,
  observacao VARCHAR,
  observacao_coleta VARCHAR,
  observacao_analise VARCHAR,
  observacao_climatica VARCHAR,
  observacao_entorno VARCHAR,
  limite_de_deteccao VARCHAR,
  data_recebimento_amostra TIMESTAMP,
  data_termino_analises TIMESTAMP,
  data_insercao TIMESTAMP,
  hora_coleta TIME,
  PRIMARY KEY(parametro, data_coleta, id_ponto) ,
  FOREIGN KEY(parametro)
    REFERENCES qualidadeagua.parametro(nome_parametro)
    ON UPDATE CASCADE,
  FOREIGN KEY(id_ponto)
    REFERENCES qualidadeagua.ponto(id_ponto)
    ON UPDATE CASCADE,
  FOREIGN KEY(nome_laboratorio)
    REFERENCES qualidadeagua.laboratorio(nome_laboratorio)

```

```

        ON UPDATE CASCADE,
    FOREIGN KEY(id_metodo_analise)
        REFERENCES qualidadeagua.metodo_analise(id_metodo_analise)
        ON UPDATE CASCADE,
    FOREIGN KEY(id_metodo_coleta)
        REFERENCES qualidadeagua.metodo_coleta(id_metodo_coleta)
        ON UPDATE CASCADE,
    FOREIGN KEY(condicao_climatica)
        REFERENCES qualidadeagua.condicao_climatica(condicao_climatica)
        ON UPDATE CASCADE),
    FOREIGN KEY(unidade)
        REFERENCES qualidadeagua.unidade(nome_unidade)
        ON DELETE NO ACTION ON UPDATE CASCADE;

CREATE TABLE qualidadeagua.unidades_padrao (
    unidade VARCHAR NOT NULL ,
    parametro VARCHAR NOT NULL ,
    PRIMARY KEY(unidade, parametro) ,
    FOREIGN KEY(unidade)
        REFERENCES qualidadeagua.unidade(nome_unidade)
        ON UPDATE CASCADE,
    FOREIGN KEY(parametro)
        REFERENCES qualidadeagua.parametro(nome_parametro)
        ON UPDATE CASCADE);

--Schema temporario--
CREATE SCHEMA temp;
CREATE TABLE temp.parametro_medido (
    parametro VARCHAR NOT NULL,
    data_coleta TIMESTAMP NOT NULL,
    id_ponto VARCHAR(30) NOT NULL,
    id_metodo_coleta INTEGER,
    id_metodo_analise INTEGER NOT NULL,
    nome_laboratorio VARCHAR NOT NULL,
    responsavel_coleta VARCHAR,
    valor VARCHAR NOT NULL,
    unidade VARCHAR NOT NULL,
    dia_coleta NUMERIC(2,0) NOT NULL,
    mes_coleta NUMERIC(2,0) NOT NULL,
    ano_coleta NUMERIC(4,0) NOT NULL,
    condicao_climatica VARCHAR NOT NULL,
    observacao VARCHAR,
    observacao_coleta VARCHAR,
    observacao_analise VARCHAR,
    observacao_climatica VARCHAR,
    observacao_entorno VARCHAR,
    limite_de_deteccao VARCHAR,
    data_recebimento_amostra TIMESTAMP,
    data_termino_analises TIMESTAMP,
    data_insercao TIMESTAMP,
    usuario VARCHAR,
    hora_coleta TIME,
    PRIMARY KEY(parametro, data_coleta, id_ponto) ,
    FOREIGN KEY(parametro)
        REFERENCES qualidadeagua.parametro(nome_parametro)
        ON UPDATE CASCADE,
    FOREIGN KEY(id_ponto)
        REFERENCES qualidadeagua.ponto(id_ponto)
        ON UPDATE CASCADE,
    FOREIGN KEY(nome_laboratorio)
        REFERENCES qualidadeagua.laboratorio(nome_laboratorio)
        ON UPDATE CASCADE,
    FOREIGN KEY(id_metodo_analise)
        REFERENCES qualidadeagua.metodo_analise(id_metodo_analise)
        ON UPDATE CASCADE,
    FOREIGN KEY(id_metodo_coleta)
        REFERENCES qualidadeagua.metodo_coleta(id_metodo_coleta)
        ON UPDATE CASCADE,
    FOREIGN KEY(condicao_climatica)
        REFERENCES qualidadeagua.condicao_climatica(condicao_climatica)
        ON UPDATE CASCADE,
    FOREIGN KEY(unidade)
        REFERENCES qualidadeagua.unidade(nome_unidade)
        ON DELETE NO ACTION ON UPDATE CASCADE
);

```

```

--SCHEMA PUBLIC
CREATE TABLE plruby_singleton_methods (
  name VARCHAR(60) NOT NULL,
  args VARCHAR(60),
  body VARCHAR,
  comment VARCHAR,
  PRIMARY KEY(name));

CREATE TABLE regra (
  nome_regra VARCHAR(60) NOT NULL,
  dimensao VARCHAR NOT NULL,
  parametro VARCHAR NOT NULL,
  ponderacao_regra FLOAT,
  PRIMARY KEY(nome_regra),
  FOREIGN KEY(dimensao)
    REFERENCES dimensao(dimensao)
    ON UPDATE CASCADE,
  FOREIGN KEY(parametro)
    REFERENCES qualidadeagua.parametro(nome_parametro),
  FOREIGN KEY(nome_regra)
    REFERENCES plruby_singleton_methods(name)
    ON DELETE CASCADE ON UPDATE CASCADE);

CREATE TABLE dimensao (
  dimensao VARCHAR NOT NULL,
  descricao_dimensao VARCHAR NULL,
  PRIMARY KEY(dimensao)
);

--Schema logs--
CREATE SCHEMA logs;
CREATE TABLE logs.regra_pendente(
  parametro_dependente VARCHAR NOT NULL,
  id_ponto VARCHAR NOT NULL,
  data_coleta TIMESTAMP NOT NULL,
  parametro VARCHAR NOT NULL,
  regra VARCHAR NOT NULL,
  usuario VARCHAR,
  regra_aplicada boolean NOT NULL DEFAULT false,
  CONSTRAINT log_regra_pkey PRIMARY KEY (parametro_dependente, id_ponto, data_coleta,
parametro, regra),
  CONSTRAINT log_regra_parametro_dependente_fkey FOREIGN KEY (parametro_dependente)
    REFERENCES qualidadeagua.parametro (nome_parametro) MATCH SIMPLE
    ON UPDATE NO ACTION ON DELETE NO ACTION,
  CONSTRAINT log_regra_parametro_fkey FOREIGN KEY (parametro, data_coleta, id_ponto)
    REFERENCES "temp".parametro_medido (parametro, data_coleta, id_ponto) MATCH SIMPLE
    ON UPDATE CASCADE ON DELETE CASCADE,
  CONSTRAINT log_regra_regra_fkey FOREIGN KEY (regra)
    REFERENCES regra (nome_regra) MATCH SIMPLE
    ON UPDATE CASCADE ON DELETE CASCADE
);

CREATE TABLE logs.regras_aplicadas(
  regra VARCHAR(60) NOT NULL,
  id_ponto VARCHAR(30) NOT NULL,
  data_coleta TIMESTAMP NOT NULL,
  parametro VARCHAR NOT NULL,
  data_aplicada TIMESTAMP NOT NULL,
  nota_recebida DOUBLE PRECISION NOT NULL,
  usuario VARCHAR NOT NULL,
  CONSTRAINT log_regras_aplicadas_pkey PRIMARY KEY (regra, id_ponto, data_coleta, parametro,
data_aplicada),
  CONSTRAINT log_regras_aplicadas_parametro_fkey FOREIGN KEY (parametro, data_coleta,
id_ponto)
    REFERENCES "temp".parametro_medido (parametro, data_coleta, id_ponto) MATCH SIMPLE
    ON UPDATE CASCADE ON DELETE CASCADE,
  CONSTRAINT log_regras_aplicadas_regra_fkey FOREIGN KEY (regra)
    REFERENCES regra (nome_regra) MATCH SIMPLE
    ON UPDATE CASCADE ON DELETE CASCADE
)

```

## ANEXO V

## Regras implementadas no sistema

Regra	Parâmetro	Peso	Descrição
condutividade0	Condutividade	1	Regra que determina qual o valor máximo e mínimo que a condutividade pode obter.
<pre>condutividade = convertef(parametro); if (condutividade &gt; 0 and condutividade &lt; 10000)   return 98; else   return 0; end</pre>			
condutividade1	Condutividade	2	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.
<pre>valorMedido = convertef(parametro); soma = 0; cont = 0; vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro = 'Condutividade' AND id_ponto = '"+ponto+"' "); valores = Array.new; vals.each{ linha    valor = linha['valor'];   valores[cont] = valor.to_f;   soma += valor.to_f;   cont += 1; } if( vals.size &lt; 4)   return 98; end media = soma / cont; soma = 0; valores.each{ valor    soma += (valor-media)**2; } desvio = (soma / (cont-1))**0.5;  #1-sigma if((valorMedido &lt; (media+desvio)) and (valorMedido &gt; (media-desvio)))   return 98; #2-sigma elsif((valorMedido &lt; (media+ (2 * desvio))) and (valorMedido &gt; (media - (2 * desvio))))   return 80; #3-sigma elsif((valorMedido &lt; (media+ (3 * desvio))) and (valorMedido &gt; (media - (3 * desvio))))   return 60; #4-sigma elsif((valorMedido &lt; (media+ (4 * desvio))) and (valorMedido &gt; (media - (4 * desvio))))   return 40; else   return 2; end</pre>			
condutividade2	Condutividade	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
<pre>valorMedido = convertef(parametro); soma = 0; cont = 0; regras = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro = 'Condutividade' AND id_ponto = '"+ponto+"' ORDER BY data_coleta DESC LIMIT 30" );  valores = Array.new; if( regras.size &lt; 4)   return 98; end regras.each{ linha </pre>			

```

    valor = linha['valor'];
    valores[cont] = valor.to_f;
    soma += valor.to_f;
    cont += 1;
  }
  media = soma / cont;
  soma = 0;
  valores.each{|valor|
    soma += (valor-media)**2;
  }
  desvio = (soma / (cont-1))**0.5;

  #1-sigma
  if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
    return 98;
  #2-sigma
  elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
    return 80;
  #3-sigma
  elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
    return 60;
  #4-sigma
  elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
    return 40;
  else
    return 2;
  end
end

```

condutividade3	Condutividade	1	Regra para valores altos de condutividade baseados no pH (ESTEVEES, 2002).
----------------	---------------	---	--

```

valor = convertetef(parametro);
con = PLruby.exec("SELECT valor FROM temp.parametro_medido WHERE id_ponto='"+ponto+"'
                  AND data_coleta = '"+data+"' AND parametro = 'pH'");
if(con[0] == nil)
  return ['pH'];
end

ph = convertetef(con[0]["valor"]);

if(ph < 4 and valor > 1000)
  return 98;
elsif(ph > 4 and valor < 1000)
  return 98;
else
  return 20;
end
end

```

dbo0	DBO	1	Regra que determina o limite mínimo e máximo que pode ser detectado de DBO.
------	-----	---	---

```

dbo = convertetef(parametro);
if(dbo > 0 and dbo < 2000)
  return 98;
else
  return 0;
end
end

```

dbo1	DBO	3	Valores limite do DBO baseado no IQA.
------	-----	---	---------------------------------------

```

dbo = convertetef(parametro);
if dbo > 0 and dbo < 30
  return 98;
else
  return 2;
end
end

```

dbo2	DBO	3	Regra que verifica que se a demanda química de oxigênio é maior que a demanda biológica de oxigênio ou seja se a (DQO > DBO)
------	-----	---	--

```

valor = convertetef(parametro);
con = PLruby.exec("SELECT valor FROM temp.parametro_medido WHERE id_ponto = '"+ponto+"'
                  AND data_coleta = '"+data +"'AND parametro = 'DQO'");
if(con[0] == nil)

```

```

return ["DQ0"];
end
dqi = convertef(con[0]['valor']);

if(valor < dqi)
return 98;
else
return 2;
end

```

dbo3	DBO	3	Valor da qualidade do dado com base no IQA.
------	-----	---	---

```

dbo = convertef(parametro);
if(dbo <= 5)
return 99.96 * (2.71828**(-0.1232728*dbo));
elsif(dbo > 5 and dbo<=15)
return 104.67-(72.5463*log10(dbo));
elsif(dbo > 15 and dbo<=30)
return 4394.91*(dbo**(-1.99809));
else
return 4.789;
end

```

dbo4	DBO	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
------	-----	---	---

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
regras = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido
WHERE parametro = 'DBO' AND id_ponto = '"+ponto+"' ORDER BY
data_coleta DESC LIMIT 30" );
valores = Array.new;
if( regras.size < 4)
return 98;
end
regras.each{|linha|
valor = linha['valor'];
valores[cont] = valor.to_f;
soma += valor.to_f;
cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
return 40;
else
return 2;
end

```

dbo5	DBO	2	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.
------	-----	---	---

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro = 'DBO'
AND id_ponto = '"+ponto+"' );

```

```

valores = Array.new;
if( vals.size < 4)
  return 98;
end
vals.each{|linha|
  valor = linha['valor'];
  valores[cont] = valor.to_f;
  soma += valor.to_f;
  cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
  return 40;
else
  return 2;
end

```

dqo0	DQO	1	Valores limites para o DQO.
------	-----	---	-----------------------------

```

valor = convertef(parametro);
if(valor > 0 and valor < 2500 )
  return 98;
else
  return 0;
end

```

dqo1	DQO	3	Regra que verifica que se a demanda química de oxigênio é maior que a demanda biológica de oxigênio ou seja se a (DQO > DBO)
------	-----	---	--

```

valor = convertef(parametro);
con = PLruby.exec("SELECT valor FROM temp.parametro_medido
  WHERE id_ponto = '"+ponto+"' AND data_coleta = '"+data +"' AND parametro = 'DBO'");
if(con[0] == nil)
  return ["DBO"]
end
dbo = convertef(con[0]['valor']);
if(valor > dbo)
  return 98;
else
  return 2;
end

```

dqo2	DQO	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
------	-----	---	---

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
regras = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido
  WHERE parametro = 'DQO' AND id_ponto = '"+ponto+"' ORDER BY data_coleta DESC LIMIT 30");
valores = Array.new;
if( regras.size < 4)
  return 98;
end
regras.each{|linha|
  valor = linha['valor'];
  valores[cont] = valor.to_f;

```

```

    soma += valor.to_f;
    cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
  return 40;
else
  return 2;
end

```

dqo3

DQO

2

Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido
                   WHERE parametro = 'DQO' AND id_ponto = '"+ponto+"' ");
valores = Array.new;
if( vals.size < 4)
  return 98;
end
vals.each{|linha|
  valor = linha['valor'];
  valores[cont] = valor.to_f;
  soma += valor.to_f;
  cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
  return 40;
else
  return 2;
end

```

od0

Oxigênio  
dissolvido

1

Regra que determina o limite mínimo e máximo do oxigênio dissolvido.

```

od = convertef(parametro);
if(od > 0 and od < 20)
  return 98;
else
  return 0;

```



end			
od1	Oxigênio dissolvido	3	Se o ambiente estiver anóxico.
<pre>od = convertef(valor); if od &lt; 2   return 1; else   return 98; end</pre>			
od2	Oxigênio dissolvido	3	Calcula o valor do Oxigênio Dissolvido se a água estivesse sem nenhuma contaminação
<pre>con1 = PLruby.exec("SELECT valor FROM temp.parametro_medido WHERE id_ponto = '" + ponto + "' AND parametro= 'Temperatura da água' AND data_coleta = '" + data + "'"); con2 = PLruby.exec("SELECT altitude FROM qualidadeagua.ponto WHERE id_ponto = '"+ponto+"'"); valor = convertef(parametro);  retorno = Array.new; if(con1[0] == nil)   retorno[0] = "Temperatura da água";   if con2[0] == nil     retorno[1] = "Altitude";   end   return retorno; elsif con2[0] == nil   retorno[0] = "Altitude";   return retorno; end  temperatura = con1[0]["temperatura"]; altitude = con1[0]["altitude"];  od = interpolaTemperatura(temperatura) * interpolaAltitude(altitude); conf = (parametro - od).abs; return 100 - (5 * conf);</pre>			
od3	Oxigênio dissolvido	2	Calcula se a porcentagem de saturação é menor que 140
<pre>con1 = PLruby.exec("SELECT valor FROM temp.parametro_medido WHERE id_ponto = '" + ponto + "' AND parametro= 'Temperatura da água' AND data_coleta = '" + data + "'"); con2 = PLruby.exec("SELECT altitude FROM qualidadeagua.ponto WHERE id_ponto = '"+ponto+"'"); valor = convertef(parametro);  retorno = Array.new; if(con1[0] == nil)   retorno[0] = "Temperatura da água";   if con2[0] == nil     retorno[1] = "Altitude";   end   return retorno; elsif con2[0] == nil   retorno[0] = "Altitude";   return retorno; end  temperatura = con1[0]["temperatura"]; altitude = con1[0]["altitude"]; od = interpolaTemperatura(temperatura) * interpolaAltitude(altitude); satura = valor *100 / od; if(satura &lt;= 100)   return 98; elsif (satura &lt;= 140)   return 70; else   return 2; end</pre>			
od4	Oxigênio dissolvido	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
valorMedido = convertef(parametro);			

```

soma = 0;
cont = 0;
vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido
                    WHERE parametro = 'Oxigênio dissolvido' AND id_ponto = '"+ponto+"'
                    ORDER BY data_coleta DESC LIMIT 30" );

valores = Array.new;
if( vals.size < 4)
  return 98;
end
vals.each{|linha|
  valor = linha['valor'];
  valores[cont] = valor.to_f;
  soma += valor.to_f;
  cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
  return 40;
else
  return 2;
end

```

od5	Oxigênio dissolvido	2	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.
-----	---------------------	---	---

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro =
                    'Oxigênio dissolvido' AND id_ponto = '"+ponto+"' );

valores = Array.new;
vals.each{|linha|
  valor = linha['valor'];
  valores[cont] = valor.to_f;
  soma += valor.to_f;
  cont += 1;
}
if( vals.size < 4)
  return 98;
end
media = soma / cont;
soma = 0;
valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))

```

```

    return 40;
else
    return 2;
end

```

od6	Oxigênio dissolvido	2	Regra que calcula o Oxigênio dissolvido com base no IQA.
-----	---------------------	---	--

```

con1 = PLruby.exec("SELECT valor FROM temp.parametro_medido WHERE id_ponto='" + ponto + "' AND parametro= 'Temperatura da água' AND data_coleta = '" + data + "'");
con2 = PLruby.exec("SELECT altitude FROM qualidadeagua.ponto WHERE id_ponto = '"+ponto+"'");
valor = convertef(parametro);

retorno = Array.new;
if(con1[0] == nil)
    retorno[0] = "Temperatura da água";
    if con2[0] == nil
        retorno[1] = "Altitude";
    end
    return retorno;
else
    if con2[0] == nil
        retorno[0] = "Altitude";
        return retorno;
    end
end
temperatura = con1[0]["temperatura"];
altitude = con1[0]["altitude"];

od = interpolaTemperatura(temperatura) * interpolaAltitude(altitude);
porsat= (convertef(valor) * 100)/ od;

if(porsat<=50)
    return ((0.34*porsat)+(0.008095*(porsat*porsat))+((1.35252*0.00001)*(porsat** 3))+3);
elsif((50 < porsat) and (porsat<=85))
    return ((-1.166*porsat)+(0.058*(porsat*porsat))-((3.803435*0.0001)*(porsat ** 3))+3);
elsif((85 < porsat) and (porsat<=100))
    return (3.7745*(porsat ** 0.704889))+3;
elsif((100 < porsat) and (porsat<=140))
    return ((2.9*porsat)-(0.02496*(porsat*porsat))+((5.60919*0.00001)*(porsat ** 3))+3);
else
    return 50;
end

```

ph0	pH	1	Regra que calcula os limites do pH
-----	----	---	------------------------------------

```

ph = convertef(parametro);
if ph > 0 and ph < 14
    return 0;
else
    return 98;
end

```

ph1	pH	3	Limites da classificação CONAMA
-----	----	---	---------------------------------

```

ph = convertef(parametro);
if ph > 6.to_f and ph < 9.to_f
    return 1;
else
    return 98;
end

```

ph2	pH	2	Limites apresentados pelo IQA
-----	----	---	-------------------------------

```

ph = convertef(parametro);
if ph > 2 and ph < 12
    return 0;
else
    return 98;
end

```

ph4	pH	1	O pH em águas que correm sobre uma geologia que apresentam rochas vulcânicas ácidas como riolitos e riolitos, característica da bacia (HAZENACK E WEBER, 2007). Devido a esse fator o pH tende a ser menor que 8
<pre> ph = convertef(parametro); if ph &gt; 8   return 10; else   return 98; end </pre>			
ph5	pH	2	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.
<pre> valorMedido = convertef(parametro); soma = 0; cont = 0; vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro = 'pH' AND id_ponto = '"+ponto+"'"); valores = Array.new; if( vals.size &lt; 4)   return 98; end vals.each{ linha    valor = linha['valor'];   valores[cont] = valor.to_f;   soma += valor.to_f;   cont += 1; } media = soma / cont; soma = 0; valores.each{ valor    soma += (valor-media)**2; } desvio = (soma / (cont-1))**0.5;  #1-sigma if((valorMedido &lt; (media+desvio)) and (valorMedido &gt; (media-desvio)))   return 98; #2-sigma elsif((valorMedido &lt; (media+ (2 * desvio))) and (valorMedido &gt; (media - (2 * desvio))))   return 80; #3-sigma elsif((valorMedido &lt; (media+ (3 * desvio))) and (valorMedido &gt; (media - (3 * desvio))))   return 60; #4-sigma elsif((valorMedido &lt; (media+ (4 * desvio))) and (valorMedido &gt; (media - (4 * desvio))))   return 40; else   return 2; end </pre>			
ph6	pH	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
<pre> valorMedido = convertef(parametro); soma = 0; cont = 0; regras = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro = 'pH' AND id_ponto = '"+ponto+"' ORDER BY data_coleta DESC LIMIT 30"); valores = Array.new; if( regras.size &lt; 4)   return 98; end regras.each{ linha    valor = linha['valor'];   valores[cont] = valor.to_f;   soma += valor.to_f;   cont += 1; } media = soma / cont; soma = 0; </pre>			

```

valores.each{|valor|
  soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
  return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
  return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
  return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
  return 40;
else
  return 2;
end

```

ph7	pH	3	Calculo do pH baseado no IQA.
-----	----	---	-------------------------------

```

ph = convertef(parametro);
if(ph <= 2)
  return 2;
elsif(ph > 2 and ph <= 4)
  return 13.6-(10.6 * ph)+(2.4364*( ph * ph));
elsif(ph > 4 and ph <= 6.2)
  return 155.5-(77.36 * ph)+(10.2481*(ph * ph));
elsif(ph > 6.2 and ph<=7)
  return -657.2+(197.38 * ph)-(12.9167*(ph * ph));
elsif(ph > 7 and ph <= 8)
  return -427.8+(142.05 * ph)-(9.695*(ph * ph));
elsif(ph > 8 and ph <= 8.5)
  return 216-(16 * ph);
elsif(ph > 8.5 and ph <= 9)
  return 1415823 * exp(-1.1507 * ph);
elsif(ph > 9 and ph <= 10)
  return 288-(27 * ph);
elsif(ph > 10 and ph <= 12)
  return 633-(106.5 * ph)+(4.5*(ph * ph));
else
  return 3;
end

```

temperatura0	Temperatura da água	1	Limites da temperatura da água
--------------	---------------------	---	--------------------------------

```

temp = convertef(parametro);
if temp > 0 and temp < 100
  return 98;
else
  return 0;
end

```

temperatura1	Temperatura da água	3	Verifica os valores da temperatura da água em relação aos valores que a água em um rio pode assumir levando em consideração o movimento e o volume.
--------------	---------------------	---	---

```

temp = convertef(parametro);
if temp > 4 and temp < 25
  return 98;
elsif temp > 4 and temp < 40
  return 70;
else
  return 2;
end

```

temperatura2	Temperatura da água	3	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta com base nos últimos 30 valores.
--------------	---------------------	---	---

```

valorMedido = convertef(parametro);
soma = 0;

```

```

cont = 0;
vals = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro =
    'Temperatura da água' AND id_ponto = '"+ponto+"' ORDER BY data_coleta DESC LIMIT 30" );
valores = Array.new;
if( vals.size < 4)
    return 98;
end
vals.each{|linha|
    valor = linha['valor'];
    valores[cont] = valor.to_f;
    soma += valor.to_f;
    cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
    soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
    return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
    return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
    return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
    return 40;
else
    return 2;
end

```

temperatura3	Temperatura da água	2	Regra estatística que utiliza o gráfico de controle com base em todos os dados inseridos na base de consulta.
--------------	---------------------	---	---

```

valorMedido = convertef(parametro);
soma = 0;
cont = 0;
regras = PLruby.exec("SELECT valor FROM qualidadeagua.parametro_medido WHERE parametro =
    'Temperatura da água' AND id_ponto = '"+ponto+"' );
valores = Array.new;
regras.each{|linha|
    valor = linha['valor'];
    valores[cont] = valor.to_f;
    soma += valor.to_f;
    cont += 1;
}
media = soma / cont;
soma = 0;
valores.each{|valor|
    soma += (valor-media)**2;
}
desvio = (soma / (cont-1))**0.5;

#1-sigma
if((valorMedido < (media+desvio)) and (valorMedido > (media-desvio)))
    return 98;
#2-sigma
elsif((valorMedido < (media+ (2 * desvio))) and (valorMedido > (media - (2 * desvio))))
    return 80;
#3-sigma
elsif((valorMedido < (media+ (3 * desvio))) and (valorMedido > (media - (3 * desvio))))
    return 60;
#4-sigma
elsif((valorMedido < (media+ (4 * desvio))) and (valorMedido > (media - (4 * desvio))))
    return 40;
else
    return 2;
end

```

## Métodos auxiliares

**convertef(valor)**

```

if var.class.to_s == 'Fixnum'
  return var;
end
var = var.gsub(",", ".");
var = var.gsub(">", "");
var = var.gsub("<", "");

return var.to_f;

```

**interpolaTemperatura(temperatura)**

```

temp = Array.new;
temp[0] = 14.2;
temp[1] = 13.8;
temp[2] = 13.4;
temp[3] = 13.1;
temp[4] = 12.7;
temp[5] = 12.4;
temp[6] = 12.1;
temp[7] = 11.8;
temp[8] = 11.5;
temp[9] = 11.2;
temp[10] = 10.9;
temp[11] = 10.7;
temp[12] = 10.4;
temp[13] = 10.2;
temp[14] = 10.0;
temp[15] = 9.8;
temp[16] = 9.6;
temp[17] = 9.4;
temp[18] = 9.2;
temp[19] = 9.0;
temp[20] = 8.8;
temp[21] = 8.7;
temp[22] = 8.5;
temp[23] = 8.4;
temp[24] = 8.3;
temp[25] = 8.1;
temp[26] = 8.0;
temp[27] = 7.9;
temp[28] = 7.8;
temp[29] = 7.6;
temp[30] = 7.5;
temp[31] = 7.5;
temp[32] = 7.4;
temp[33] = 7.3;
temp[34] = 7.2;
temp[35] = 7.1;
temp[36] = 7.0;
temp[37] = 6.9;
temp[38] = 6.8;
temp[39] = 6.7;
temp[40] = 6.6;
temp[41] = 6.5;
temp[42] = 6.4;
temp[43] = 6.3;
temp[44] = 6.2;
temp[45] = 6.1;
temp[46] = 6.0;
temp[47] = 5.9;
temp[48] = 5.8;
temp[49] = 5.7;
temp[50] = 5.6;

#calculo da temperatura

```

```
x = temperatura;
i = x.to_i;
a = i;
c = (i+1);
fa = temp[i];
fc = temp[i+1];
z = c - a;
y = fc - fa;
w = x - a;
return ((w * y) / z) + fa;
```

### interpolaAltitude(altitude)

```
tabela = Array.new;
tabela[0] = 1;
tabela[1] = 0.998;
tabela[2] = 0.976;
tabela[3] = 0.966;
tabela[4] = 0.954;
tabela[5] = 0.942;
tabela[6] = 0.930;
tabela[7] = 0.920;
tabela[8] = 0.908;
tabela[9] = 0.897;
tabela[10] = 0.887;
tabela[11] = 0.876;
tabela[12] = 0.865;
tabela[13] = 0.855;
tabela[14] = 0.845;
tabela[15] = 0.834;
tabela[16] = 0.825;
x = altitude;
i = x.to_i / 100;
a = i*100;
c = (i+1)*100;
fa = tabela[i];
fc = tabela[i+1];
z = c - a;
y = fc - fa;
w = x - a;
return ((w * y) / z) + fa;
```



## ANEXO VI

### Telas do módulo de avaliação e controle de dados

Tela que apresenta os dados que foram bloqueados pelo sistema e que estão esperando a avaliação de um usuário especialista.

**Sistema de avaliação de qualidade de dados**

Dados Pendentes    Regras Pendentes    Gerenciamento de Regras

**Dados Pendentes**

Parametro	Valor	Unidade	Nota	Data Coleta	Ponto	Ações
Oxigênio dissolvido	5.23	mg/L	22	20/12/2009	2	✓ ✎ ✕ 🔍
pH	5	Sem Unidade	45	25/05/2010 às 19:46	2	✓ ✎ ✕ 🔍
pH	9	Sem Unidade	61.2	09/06/2010 às 14:57	2	✓ ✎ ✕ 🔍
pH	5	Sem Unidade	78.6	19/06/2010 às 16:44	2	✓ ✎ ✕ 🔍
pH	8	Sem Unidade	78.8	19/06/2010 às 17:43	2	✓ ✎ ✕ 🔍
pH	80.2	Sem Unidade	2.6	19/06/2010 às 17:53	2	✓ ✎ ✕ 🔍

A figura a seguir mostra a avaliação de um dado.







Arquivo Editar Exibir Histórico Favoritos Ferramentas Ajuda

http://200.160.134.170/tools/tcc/site/?acao=analizarDado&pe desiclopedia

Gmail - CARTÃO INFO... FINALISTAS.pdf - Pow... Gerenciamento da qu... Analisar dado - Avalia...

## Sistema de avaliação de qualidade de dados

Dados Pendentes Regras Pendentes Avalie o sistema Gerenciamento de Regras

### Analisar dado

**Você esta analisando o dado pH com o valor 8.2 Sem Unidade**

Esse dado foi classificado com a qualidade 70.16%.

Qualidade péssima: entre 0 e 20	Qualidade baixa: entre 21 e 40	Qualidade média: entre 41 e 60	Qualidade boa: entre 61 e 80	Qualidade ótima: entre 80 e 98
---------------------------------	--------------------------------	--------------------------------	------------------------------	--------------------------------

O que você deseja ver.

- [Editar dado](#)
- [Ver regras aplicadas ao dado](#)
- [Ver dados da campanha](#)

Regra	Data aplicada	Nota Recebida
<a href="#">ph0</a>	02/07/2010 às 13:35	98
<a href="#">ph1</a>	02/07/2010 às 13:35	98
<a href="#">ph2</a>	02/07/2010 às 13:35	98
<a href="#">ph4</a>	02/07/2010 às 13:35	10
<a href="#">ph5</a>	02/07/2010 às 13:35	40
<a href="#">ph6</a>	02/07/2010 às 13:35	40
<a href="#">ph7</a>	02/07/2010 às 13:35	84.8

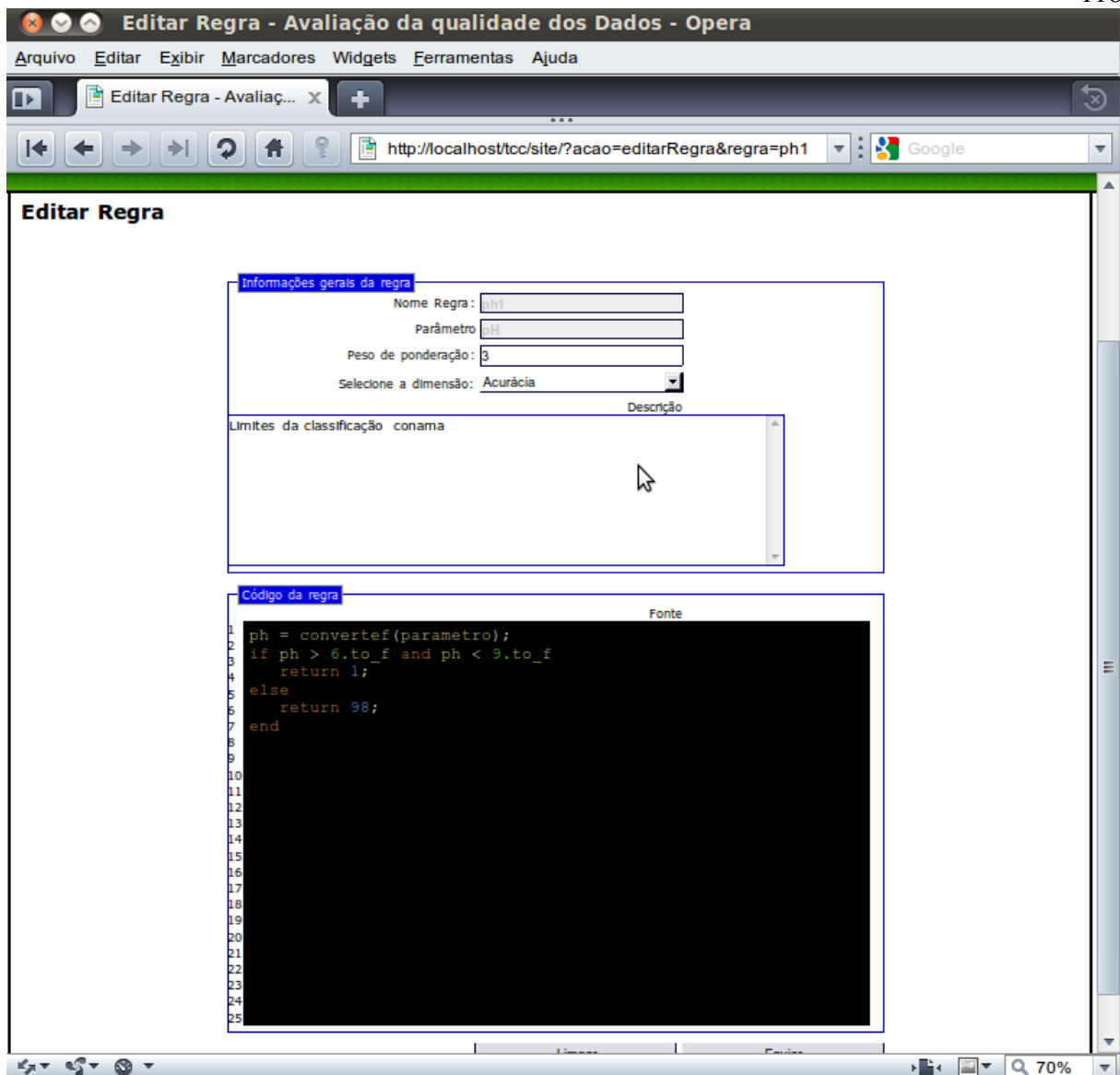
3. [Ver dados da campanha](#)

Data Campanha: 01/01/2005  
Ponto: 1

Parâmetro	Valor	Unidade	Limite de Detecção	Classe Conama	Qualidade Estimado
pH	8.2	Sem Unidade			70,16%
Oxigênio dissolvido	6.35	mg/L			88,00%
DBO	3	mg/L			83,26%

Concluído

A Figura a seguir apresenta a tela para a edição de regras, essa interface automatiza de forma sistemática e facilita o entendimento das regras para usuários mais avançados. Essa tela para a edição de regras será bloqueada para alguns usuários. Essa política de acesso ao sistema de regras ficará como um trabalho futuro.



A tela apresentada na figura a seguir é a tela para o controle e visualização de regras.

Gerenciamento de Regras - Avaliação da qualidade dos Dados - Opera

Arquivo Editar Exibir Marcadores Widgets Ferramentas Ajuda

Gerenciamento de Re... x

http://localhost/tcc/site/?acao=gerenciarRegras

## Gerenciamento de Regras

[\[+\] Criar nova Regra](#)

Selecione as restrições

Exibir regras associadas ao parametro:

Parametro	Nome Regra	Peso	Dimensão	Ações
Condutividade	condutividade0	1	Acurácia	  
Condutividade	condutividade1	2	Acurácia	  
Condutividade	condutividade2	3	Acurácia	  
Condutividade	condutividade3	1	Acurácia	  
DBO	dbo0	1	Acurácia	  
DBO	dbo1	3	Acurácia	  
DBO	dbo2	3	Acurácia	  
DBO	dbo3	3	Acurácia	  
DBO	dbo4	3	Acurácia	  
DBO	dbo5	2	Acurácia	  
DQO	dqo0	1	Acurácia	  
DQO	dqo1	3	Acurácia	  
DOO	dao2	3	Acurácia	  

110%

## ANEXO VII

### Formulário para a validação do protótipo.

#### Aplicabilidade

1. O sistema tem aplicabilidade prática nos dados que normalmente são utilizados?

- Ótima  Boa  Regular  Ruim  Péssima

2. Para uma grande quantidade de dados classifique a aplicabilidade?

- Ótima  Boa  Regular  Ruim  Péssima

3. Quanto a utilidade para as avaliações que normalmente se faz aos dados, qual a aplicabilidade?

- Ótima  Boa  Regular  Ruim  Péssima

Sugestões para melhorar a aplicabilidade do sistema

#### Usabilidade

1. Classifique a facilidade com que você conseguiu aprender as funcionalidades do sistema?

- Ótima  Boa  Regular  Ruim  Péssima

2. Classifique a facilidade com que você conseguiu usar as funcionalidades do sistema?

- Ótima  Boa  Regular  Ruim  Péssima

3. Classifique a facilidade com que você conseguiu entender os resultados da qualidade dos dados?

- Ótima  Boa  Regular  Ruim  Péssima

Sugestões para melhorar a usabilidade do sistema

#### Confiabilidade

1. Classifique a confiança que o sistema lhe passou em aplicar a nota dos dados?

- Ótima  Boa  Regular  Ruim  Péssima

2. Classifique a confiabilidade das regras?

- Ótima  Boa  Regular  Ruim  Péssima

3. Sua opinião sobre a metodologia para avaliar a qualidade do dado?

- Ótima  Boa  Regular  Ruim  Péssima

Sugestões para melhorar a confiabilidade do sistema