



AUTARQUIA ASSOCIADA À UNIVERSIDADE DE SÃO PAULO

**MÉTODOS ESTATÍSTICOS PARA ANÁLISE DE DADOS  
DE MONITORAÇÃO AMBIENTAL**

**MARIA ANGÉLICA GARCIA DE CARVALHO**

**Tese apresentada como parte dos  
requisitos para obtenção do Grau de  
Doutor em Ciências na Área de  
Tecnologia Nuclear - Aplicações.**

**Orientador:  
Dr. Goro Hiromoto**

**São Paulo  
2003**

**INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES**  
**Autarquia Associada à Universidade de São Paulo**

**MÉTODOS ESTATÍSTICOS PARA ANÁLISE DE DADOS DE MONITORAÇÃO  
AMBIENTAL**



**MARIA ANGÉLICA GARCIA DE CARVALHO**

Tese apresentada como parte dos requisitos para obtenção do Grau de Doutor em Ciências na Área de Tecnologia Nuclear – Aplicações.

Orientador:

Dr. Goro Hiromoto

**SÃO PAULO**  
**2003**

A Amandina, minha mãe e exemplo;  
A Cláudio e Bruno, meus filhos e minha motivação.

## **AGRADECIMENTOS**

A autora expressa seus sinceros agradecimentos :

Ao Dr. Goro Hiromoto, amigo e orientador, pelo incentivo constante e orientação paciente e segura;

À Dra Linda Caldas, pela atenção e inestimável apoio durante toda a realização deste trabalho;

Ao Centro de Desenvolvimento da Tecnologia Nuclear (CDTN), pela disponibilização do banco de dados;

A todos amigos do IPEN e do CDTN que de alguma forma contribuíram para o sucesso deste trabalho;

Em especial, agradeço a meu marido José Rodrigues Batista e meus filhos, sempre presentes com compreensão, apoio e estímulo durante todo o caminho percorrido.

# MÉTODOS ESTATÍSTICOS PARA ANÁLISE DE DADOS DE MONITORAÇÃO AMBIENTAL

Maria Angélica Garcia de Carvalho

## RESUMO

A execução de programas de monitoração ambiental tem gerado extensos bancos de dados contendo informações valiosas sobre a região de interesse que, muitas vezes, são pouco exploradas devido às dificuldades encontradas em sua análise e interpretação estatística. Um dos fatores determinantes deste problema é a avaliação das incertezas analíticas associadas às medidas ambientais, principalmente porque, em sua maioria, encontram-se na região próxima aos limites de detecção dos métodos de medição empregados. Este trabalho apresenta um método de análise para interpretação estatística de dados de monitoração ambiental, focado principalmente na análise exploratória inicial e na reavaliação das incertezas, utilizando-se o modelo de dois componentes de Rocke & Lorenzato para determinação de erro. A metodologia proposta foi aplicada aos resultados de análises, efetuadas ao longo de vários anos em matrizes ambientais e representando três diferentes protocolos de medida, mostrando-se um instrumento eficaz para uma estimação mais exata das médias e para a avaliação da qualidade dos dados publicados pelo laboratório executor.

# STATISTICAL METHODS FOR ANALYSIS OF ENVIRONMENTAL MONITORING DATA

**Maria Angélica Garcia de Carvalho**

## ABSTRACT

The execution of environmental monitoring programs has generated extensive data sets containing valuable information about the region of interest, many times a few explored due to the complexity of the statistical analysis and results interpretation required. One of these difficulties is the evaluation of the uncertainties associated to the environmental measurements, mainly because the majority of these data lies on the region near to the lower detection limits of the measurements methods. In this study, statistical methods for analysis of environmental monitoring data is presented, mainly focused on exploratory data analysis and re-evaluation of the reported uncertainties, using the Roche & Lorenzato's two-component model for error measurement. The proposed method was applied to the results obtained in the sample analysis representing three different analytical protocols for environmental measurements carried out along of several years. It was shown that this method is an effective tool for more accurate evaluation of the mean value of environmental data sets and for the assessment of the quality of data published by laboratories.

# SUMÁRIO

	<b>Página</b>
<b>1. INTRODUÇÃO.....</b>	<b>11</b>
1.1 Objetivos do trabalho.....	13
<b>2. ESTATÍSTICA AMBIENTAL.....</b>	<b>14</b>
2.1 Dados ambientais : conceituação e características.....	14
2.2 Medições ambientais.....	15
2.2.1 Modelos de medição.....	16
2.3 Protocolo analítico.....	18
2.4 Tratamento estatístico de dados ambientais.....	19
2.4.1 Dados discrepantes.....	20
2.4.2 Dados censurados.....	21
2.4.2.1 Limite de detecção.....	22
2.4.2.2 Métodos estatísticos para conjuntos contendo dados censurados.....	25
2.4.3 Análise exploratória.....	30
2.4.4 Estimação da média.....	30
<b>3. MÉTODO PARA ESTIMAÇÃO DE MÉDIAS EM ANÁLISES AMBIENTAIS.....</b>	<b>32</b>
3.1 Estimação da média de resultados analíticos.....	32
3.2 Resultados analíticos produzidos em PMA's: dados históricos.....	33
3.3 Incerteza "corrigida": aplicação do modelo de dois componentes.....	34
3.4 Metodologia.....	37
3.5 Os parâmetros $\sigma_{\epsilon}$ e $\sigma_{\eta}$ do modelo de dois componentes.....	39
3.6 Avaliação da performance do laboratório.....	39
<b>4. RESULTADOS.....</b>	<b>41</b>
4.1 Análise exploratória.....	41
4.2 Estimação dos parâmetros do modelo de dois componentes.....	43
4.3 Caso 1: matriz água - análise alfa total - laboratório 1.....	44
4.4 Caso 2: matriz água - análise alfa total - laboratório 2.....	49
4.5 Caso 3: matriz água - análise urânio 238.....	54
4.6 Caso 4: matriz aerossol - análise alfa total.....	57

<b>5.</b>	<b>DISCUSSÃO .....</b>	<b>62</b>
5.1	Modelo de Roche & Lorenzato x modelo baseado em Poisson .....	62
5.2	Caso 1: Análise alfa total, matriz água, lab1 .....	63
5.3	Caso 2: Análise alfa total, matriz água, lab2 .....	71
5.4	Caso 3: Análise urânio 238, matriz água.....	75
5.5	Caso 4: Análise alfa total, matriz aerossol .....	77
<b>6.</b>	<b>CONCLUSÕES .....</b>	<b>80</b>
	<b>APÊNDICE A – ANÁLISE EXPLORATÓRIA DOS DADOS.....</b>	<b>82</b>
	<b>APÊNDICE B - RESUMO ESTATÍSTICO E RESULTADOS POR AMOSTRA....</b>	<b>107</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>131</b>

## LISTA DE TABELAS

	<b>Página</b>
TABELA 2-1 Relação entre desvio padrão e concentração: alguns modelos.....	17
TABELA 4-1 Estatística descritiva: análise alfa total, Lab 1, matriz água .....	44
TABELA 4-2 Estatística descritiva: matriz água, alfa total, Lab 1, - sem os dados discrepantes....	45
TABELA 4-3 Valor dos parâmetros e qualidade do ajuste: análise alfa total, água, Lab1.....	46
TABELA 4-4 Comparação das médias: concentração alfa total (Bq/dm <sup>3</sup> ), matriz água, Lab 1 .....	49
TABELA 4-5 Estatística descritiva para resultados alfa total matriz água Lab 2.....	49
TABELA 4-6 Estatística descritiva alfa total matriz água, Lab 2 sem nuvem, sem discrepantes....	50
TABELA 4-7 Valor dos parâmetros e qualidade do ajuste: análise alfa total, água, Lab2.....	51
TABELA 4-8 Estimativas das médias: concentrações alfa total (Bq/dm <sup>3</sup> ), água – Lab 2 s/ ND.....	53
TABELA 4-9 Estimativas das médias: concentrações de alfa total (Bq/dm <sup>3</sup> ), água – Lab 2 .....	53
TABELA 4-10 Estatística descritiva: matriz água, Urânio 238 - todos registros .....	54
TABELA 4-11 Estatística descritiva: matriz água, Urânio 238 conjunto estudado .....	54
TABELA 4-12 Estimativa dos parâmetros e qualidade do ajuste: urânio238, água .....	56
TABELA 4-13 Estimativas das médias: concentrações de urânio 238 (Bq/dm <sup>3</sup> ), água (>LD).....	57
TABELA 4-14 Estimativas das médias: concentrações de urânio 238 (Bq/dm <sup>3</sup> ), água (inclui LD). 57	57
TABELA 4-15 Estatística descritiva: matriz aerossol, alfa total- todos registros.....	58
TABELA 4-16 Estatística descritiva: matriz aerossol, alfa total - conjunto estudado .....	59
TABELA 4-17 Estimativa dos parâmetros e qualidade do ajuste: alfa total, aerossol.....	60
TABELA 4-18 Estimativas das médias: concentrações de alfa total (mBq/m <sup>3</sup> ), matriz aerossol ....	61
TABELA 5-1 Estimativa das médias por faixa de concentração alfa total, águas, lab1 .....	68
TABELA 5-2 Estatística descritiva: concentração alfa total (Bq/dm <sup>3</sup> ) águas, lab1 .....	69
TABELA 5-3 Estatística descritiva: concentração alfa total, águas, lab2 .....	75
TABELA 5-4 Estatística descritiva: concentração urânio 238, águas.....	77
TABELA 5-5 Estatística descritiva: concentração alfa total, aerossol (mBq/m <sup>3</sup> ).....	79

## LISTA DE FIGURAS

	Página
FIGURA 2-1 Comportamento do desvio padrão e desvio relativo em função da concentração.....	16
FIGURA 2-2 Modelo do relacionamento valor verdadeiro - valor observado .....	23
FIGURA 2-3 Métodos de substituição simples.....	27
FIGURA 2-4 Método da Máxima Verossimilhança (MV).....	28
FIGURA 2-5 Método da regressão log-normal.....	29
FIGURA 4-1 Determinação do valor inicial de $\sigma_{\epsilon}$ , alfa total, águas, lab1.....	45
FIGURA 4-2 Determinação do valor inicial de $\sigma_{\eta}$ alfa total, águas, lab1.....	46
FIGURA 4-3 Visualização do ajuste: função erro x concentração, alfa total, água ; Lab1 .....	46
FIGURA 4-4 Comparação do erro publicado x erro previsto modelo R&L alfa total, água, Lab1	47
FIGURA 4-5 Visualização da faixa de concentração $<0,2$ : valor de $\sigma_{\epsilon}$ .....	48
FIGURA 4-6 Detalhe da comparação do erro publicado x erro previsto pelo modelo R&L.....	48
FIGURA 4-7 Erro em função da concentração alfa total, água, Lab2.....	50
FIGURA 4-8 Determinação do valor inicial de $\sigma_{\epsilon}$ : alfa total, água, Lab2 .....	51
FIGURA 4-9 Determinação do valor inicial de $\sigma_{\eta}$ : alfa total, água, Lab2 .....	51
FIGURA 4-10 Visualização do ajuste: função erro x concentração, alfa total, água ; Lab2 .....	52
FIGURA 4-11 Comparação erro publicado x erro previsto modelo R&L, alfa total, água, Lab2 .	52
FIGURA 4-12 Determinação do valor inicial de $\sigma_{\epsilon}$ urânio 238, águas .....	55
FIGURA 4-13 Determinação do valor inicial de $\sigma_{\eta}$ urânio 238, águas .....	55
FIGURA 4-14 Visualização do ajuste: função erro x concentração, urânio 238 , matriz água .....	56
FIGURA 4-15 Comparação erro publicado x erro previsto modelo R&L, urânio 238, água .....	56
FIGURA 4-16 Detalhe : erro alfa em função da concentração, alfa total, aerossol .....	58
FIGURA 4-17 Determinação do valor inicial para $\sigma_{\epsilon}$ alfa total, aerossol .....	59
FIGURA 4-18 Determinação do valor inicial para $\sigma_{\eta}$ , alfa total, aerossol.....	59
FIGURA 4-19 Visualização do ajuste: função erro x concentração, alfa total, aerossol.....	60
FIGURA 4-20 Comparação erro publicado x erro previsto modelo R&L, alfa total, aerossol.....	61
FIGURA 5-1 Comportamento do erro previsto pelo modelo de Poisson e pelo R&L.....	62
FIGURA 5-2 Aderência ao modelo: erro x concentração – concentração $<3,0$ .....	63
FIGURA 5-3 Aderência ao modelo: erro relativo x concentração – concentração $<3,0$ .....	64
FIGURA 5-4 Detalhe: erro x concentração alfa total - modelo Poisson .....	65
FIGURA 5-5 Detalhe: erro relativo x concentração alfa total - modelo Poisson.....	65
FIGURA 5-6 Detalhe: erro x concentração alfa total, modelo Poisson e R&L .....	66

FIGURA 5-7 Detalhe: erro relativo x concentração alfa total, modelo Poisson e R&L.....	66
FIGURA 5-8 Visualização do limite de detecção, alfa total, águas , lab1.....	67
FIGURA 5-9 Detalhe erro x concentração ,alfa total, água, local 049.....	69
FIGURA 5-10 Visualização dos ajustes para cada local amostrado .....	70
FIGURA 5-11 Detalhe -erro x concentração, alfa total, águas,lab2 ( sem nuvem).....	71
FIGURA 5-12 Detalhe –Erro relativo x concentração, alfa total, águas,lab2 (sem nuvem).....	72
FIGURA 5-13 Erro x concentração alfa total, águas, lab2 (subconjunto Poisson).....	72
FIGURA 5-14 Erro x concentração alfa total, água, lab2, (subconjunto Poisson).....	73
FIGURA 5-15 Erro x concentração alfa total, água, lab2 (sem nuvem).....	73
FIGURA 5-16 Detalhe: Erro relativo x concentração alfa total, água, lab2, modelo Poisson.....	74
FIGURA 5-17 Detalhe: erro x concentração U238,água, modelo R&L.....	76
FIGURA 5-18 Detalhe: Erro relativo x concentração urânio 238, águas .....	76
FIGURA 5-19 Detalhe erro x concentração alfa total , aerossol (subconjunto Poisson).....	78
FIGURA 5-20 Erro x concentração alfa total, aerossol, modelo R&L.....	78
FIGURA 5-21 Erro relativo x concentração alfa total, aerossol, modelo R&L.....	79

## 1. INTRODUÇÃO

A preservação do meio ambiente e a busca pelo desenvolvimento sustentável tornaram-se objeto de preocupação para o homem nas últimas décadas e motivaram a realização de um grande número de estudos ambientais, especialmente a implantação de Programas de Monitoração Ambiental (PMA). Estes estudos envolvem a coleta e análise de amostras do meio ambiente que apresentam, na maioria das vezes, baixas concentrações dos constituintes de interesse e são realizados sistematicamente durante vários anos. Constituem extensos bancos de dados contendo resultados referentes a vários tipos de análises e matrizes, que devem ser tratados e interpretados estatisticamente para serem utilizados. Estas medidas, planejadas para atender às finalidades específicas do programa, registram a evolução temporal de cada grandeza estudada e descrevem sua concentração nos vários locais selecionados para coleta. O banco de dados assim formado é uma preciosa fonte de informações sobre a região focalizada e pode ser usado em vários tipos de estudos ambientais. Entretanto, a disponibilização, e conseqüentemente o uso, destes conjuntos de dados são limitados pelas dificuldades em sua análise estatística e interpretação dos resultados analíticos publicados. As particularidades da medição de valores muito próximos ao *background*, que constituem maioria, mas não a totalidade, dos resultados produzidos, são a principal causa destas dificuldades. Os laboratórios encontram problemas na quantificação da concentração nesta região próxima de zero e na avaliação da incerteza da medição afetando a qualidade dos resultados produzidos.

Outro grave problema, encontrado na interpretação dos resultados contidos nestes bancos de dados, envolve conceitos e modelos de medição. A falta de normalização de conceitos como precisão, limite de detecção e métodos de avaliação da incerteza analítica dificulta a interpretação do conjunto de resultados.

Um dos objetivos da análise estatística de um conjunto de dados é buscar o melhor valor para representar este conjunto e a estatística mais usada para esta descrição é a média. No caso de bancos de dados gerados em PMA's, os resultados analíticos dificilmente apresentam o mesmo grau de confiança e a estimação da média considera este fato usando como fator de ponderação para cada medida o inverso do quadrado de sua incerteza. Uma avaliação realista da incerteza analítica, imprescindível na qualificação do resultado analítico individual, é portanto importante na descrição estatística do conjunto de

resultados de medição. Entre as causas de variação da incerteza analítica, dois fatores ligados ao processo de medição devem ser mencionados: as mudanças nas condições operacionais e sua dependência com o valor da concentração medida. Esta dependência, expressa por uma função analítica, é especificada na formulação do protocolo adotado para cada análise. Alguns resultados analíticos, presentes nos atuais bancos de dados e produzidos desde meados da década de 80, foram gerados segundo modelos de medição nem sempre adequados aos valores de concentração encontrados. Os modelos tradicionalmente adotados em análises ambientais (Zorn,1999) descrevem bem o comportamento erro-concentração em faixas restritas dos valores das concentrações. Roche & Lorenzato (1995) apresentaram um modelo de medição que se mostrou adequado para descrever esta relação no domínio pleno das concentrações. A aplicação deste modelo a resultados experimentais (Wilson et al., 2000) mostrou bom desempenho ao descrever o comportamento da incerteza em relação à concentração, para vários métodos analíticos.

Este trabalho apresenta um método de análise para interpretação de resultados analíticos de amostras ambientais gerados em um período longo de tempo. Este método considera o conjunto de resultados produzidos por um mesmo protocolo analítico como uma amostra estatística que representa o procedimento de medição executado pelo laboratório durante o período focalizado. Eles constituem um conjunto de dados que têm em comum o mesmo modelo de medição, mesmos conceitos e foram submetidos aos procedimentos definidos e praticados pelo laboratório executor. O comportamento do erro publicado em função da concentração medida é confrontado com o modelo de medição especificado no protocolo analítico ou analisado para identificação de um padrão. Considerando o modelo de Roche & Lorenzato como o que melhor descreve o comportamento erro x concentração para análises ambientais no domínio pleno das concentrações, este trabalho utiliza sua expressão algébrica para modelar os valores de incerteza a serem associados a cada concentração. Os dois parâmetros do modelo são determinados por ajuste aos resultados analíticos e utilizados para reavaliar as incertezas analíticas. Os valores destas incertezas estarão corrigindo distorções temporais originadas por variações nas condições operacionais e pela não adequação do modelo de medição ao valor da concentração ao qual foi associada. O fator de ponderação, definido como o inverso do quadrado da incerteza e calculado usando estes valores reavaliados, estará incorporando estas distorções à média de cada local.

### **1.1 Objetivos do trabalho.**

Este trabalho tem como objetivo propor um método para analisar e interpretar resultados analíticos de amostras ambientais de baixas concentrações, gerados pela execução de programas de monitoração ambiental. Os objetivos específicos deste método são:

- Verificar a adequação do modelo usado no processo de medição;
- Reavaliar as incertezas associadas a cada concentração medida usando o modelo de Rocke & Lorenzato;
- Estimar a média que melhor representa cada conjunto de resultados referentes a um local amostrado.

## 2. ESTATÍSTICA AMBIENTAL

### 2.1 Dados ambientais : conceituação e características

Dados ambientais são gerados em estudos que visam caracterizar e descrever um processo ou estado do meio ambiente atendendo várias finalidades como monitoração, remediação e pesquisa. O conceito de “soft data”, criado por Patil (1993), para descrever a natureza dos dados ambientais reflete a dificuldade de se tratar estes dados em razão da variabilidade inerente ao processo ambiental e às incertezas acumuladas nas várias etapas de seu processo de geração. De maneira geral, o processo ambiental pode ser caracterizado como complexo, multifacetado, não linear e dinâmico (Ross e Cothorn, 1994). A realização de tais estudos não é portanto simples e a análise dos resultados analíticos produzidos não é trivial mas requer conhecimentos de sua natureza, do processo de medição e de técnicas estatísticas específicas.

Conjuntos de resultados ambientais apresentam algumas características relevantes para sua interpretação e análise estatística (Helsel e Hirsch, 1992). Em geral, eles apresentam:

- distribuição não-normal e assimetria positiva, o que dificulta a aplicação de métodos paramétricos;
- sazonalidade, autocorrelação, dependência com outras variáveis não controláveis;
- pequeno número de resultados, ou seja, amostras pequenas do ponto de vista estatístico;
- dados discrepantes, que podem ser gerados no processo de medição ou ser fruto de sazonalidades, acidentes, variações temporais não esperadas ou outro processo aparentemente não correlacionado ao estudo em questão;
- dados censurados, isto é, resultados com valor abaixo de um limite mínimo pré-estabelecido. A análise destes resultados envolve problemas de conceituação e de terminologia e além disso, eles representam uma importante perda de informação.

Estudos ambientais são dispendiosos e difíceis de serem executados, envolvem problemas de planejamento, logística, freqüentemente requerem instrumentos e métodos sofisticados e longos períodos de monitoração. Por isso a quantidade de resultados é quase sempre pequena em relação ao estatisticamente desejável. O tamanho das amostras

envolvidas em estudos ambientais sempre será um problema a ser resolvido no planejamento do projeto e na análise estatística. Além disto muitos estudos dependem de séries históricas que só podem ser concluídas após anos de metódica coleta e análise.

A característica talvez mais marcante dos conjuntos de dados ambientais é a incerteza, tomada aqui no sentido mais amplo da palavra (Cothorn and Ross, 1994). Além das incertezas quantitativas inerentes a processos de medição e presentes em todas as etapas do estudo, a não-informação é um dos fatores mais críticos que interferem na análise e interpretação de resultados. Dados incompletos ou perdidos, lacunas no tempo ou espaço e que levam ao uso de interpolações para preencher estes vazios são exemplos desta incerteza. Causas como incompatibilidades em virtude do uso não adequado de métodos analíticos ou de amostragem, análises sem correspondência no tempo e espaço, amostragens não estatísticas, são também fontes de incertezas difíceis de serem tratadas, depois que o conjunto de resultados está estabelecido e as informações não foram coletadas adequadamente.

Incertezas, variabilidade e viés (“*bias*”- ABNT, 1988) sempre estarão presentes em qualquer medição de um fenômeno físico e nunca podem ser completamente eliminados (Ross and Cothorn, 1994). Entretanto um dos principais objetivos do analista, seja do ponto de vista do estatístico ou do executor das análises, é minimizar as incertezas e unir esforços para quantificá-las aumentando a qualidade do conjunto de resultados. Decisões envolvendo proteção do público e do meio ambiente, quanto a efeitos nocivos de poluição, são tomadas com base em resultados ambientais (USEPA, 2000a). A qualidade destes dados é, portanto, um atributo crítico para assegurar decisões confiáveis.

## **2.2 Medições ambientais**

A finalidade do processo de medição é determinar o valor da grandeza a ser medida. Este processo envolve a formulação de um modelo, que relaciona o mensurando e as grandezas das quais ele depende, a escolha do método e a especificação do procedimento de medição. A escolha do modelo matemático, que transforma as observações no resultado da medição, é extremamente importante especialmente na avaliação da incerteza (ABNT, 1998). A quantificação desta incerteza faz parte do processo de medição e é imprescindível para que os resultados produzidos possam ser comparados entre si ou com valores de referência, como ocorre na maioria das finalidades e usos de medidas ambientais. A compreensão e a interpretação apropriada dos resultados de medições envolvem o conhecimento do método de avaliação de sua incerteza já que as

medições realizadas são tratadas estatisticamente de acordo com a confiabilidade que cada uma apresenta.

A melhor estimativa da incerteza de uma medição é dada pela variância (ou desvio padrão, que é sua raiz quadrada positiva) de uma série de observações independentes obtidas sob as mesmas condições. Para que esta estimativa seja confiável o número de observações deve ser suficientemente grande do ponto de vista estatístico. A relação entre o valor da grandeza e a incerteza pode ser determinada por uma série de determinações em valores crescentes da medida. No comportamento da incerteza analítica, em relação ao valor da concentração medida, podem ser observadas três regiões distintas: a faixa próxima de zero, a região das concentrações mais altas e uma região intermediária. Nas concentrações mais baixas, onde o sistema de detecção tem dificuldade de distinguir diferenças muito pequenas num sinal fraco, a variabilidade tende a zero e a incerteza mostra-se constante. Nas concentrações mais altas, o erro relativo tende a ser constante. O comportamento do desvio padrão, que é a estatística mais usada para expressar a incerteza associada a um resultado analítico, e também do desvio relativo em função da concentração, (Pritt, 1994) são ilustrados na FIG 2-1. O desvio relativo (também chamado de coeficiente de variação) apresenta-se aproximadamente constante para altas concentrações e cresce rapidamente quando a concentração tende a zero.

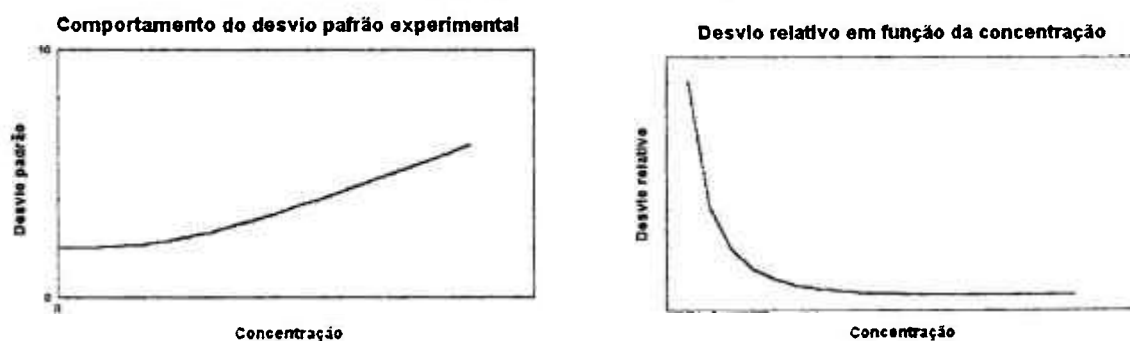


FIGURA 2-1 Comportamento do desvio padrão e desvio relativo em função da concentração

### 2.2.1 Modelos de medição

Um dos modelos mais usados em medições ambientais, sobretudo em virtude de sua simplicidade, é o que assume variância constante. Ele é expresso pela relação

$$x = \mu + \epsilon \quad \text{Eq. 2-1}$$

onde  $x$  é o resultado do processo de medição,  $\mu$  é o valor verdadeiro da concentração e  $\epsilon$  o erro analítico com distribuição normal,  $N(0, \sigma_\epsilon^2)$ . Ele se ajusta bem para valores de concentrações muito baixas porém, supondo que o valor absoluto do erro não é

relacionado com a concentração medida, ele não é uma boa aproximação para descrever o comportamento do erro em outras faixas de concentração.

Outro modelo largamente empregado e que prevê erro relativo constante é expresso por :

$$x = \mu e^{\eta} \quad \text{ou} \quad \ln(x) = \ln(\mu) + \eta \quad \text{Eq. 2-2}$$

onde  $x$  é o resultado do processo de medição,  $\mu$  é o valor verdadeiro da concentração e  $\eta$  o erro analítico com  $N(0, \sigma_{\eta}^2)$ . Este modelo descreve bem o comportamento do erro para concentrações mais altas porém falha ao ser aplicado a concentrações próximas de zero, pois prevê erro absoluto tendendo a zero nesta faixa de resultados.

Análises envolvendo medidas radiométricas usam a estatística de Poisson na estimação do erro associado a cada medida resultando em variância não constante (Knoll, 1987). Esta heterocedasticidade é também encontrada em outros tipos de análises e vários modelos têm sido propostos, em medições ambientais, para descrever o comportamento da incerteza associada a resultados analíticos.

Os modelos mais usados em análises ambientais para descrever a relação entre o desvio padrão e a concentração medida são citados na TABELA 2-1.

TABELA 2-1 Relação entre desvio padrão e concentração: alguns modelos

Modelo	Expressão	Ref
proporcional	$\sigma_x = a_0 C$	Curie, 85
linear	$\sigma_x = a_0 + a_1 C$	Gibbons, 94
quadrático	$\sigma_x = a_0 + a_1 C + a_2 C^2$	Zorn, 99
exponencial	$\sigma_x = a_0 \exp(a_1 C)$	Gibbons, 97
estatística de Poisson	$\sigma_x = (a_0 + a_1 C)^{1/2}$	Zorn, 97

A escolha do modelo apropriado requer a análise do comportamento dos erros a partir dos experimentos realizados na formulação do protocolo analítico a ser adotado em cada análise. Os modelos tradicionalmente usados nem sempre descrevem adequadamente o comportamento da incerteza analítica numa faixa ampla dos valores das concentrações. Este problema tem sido contornado, adotando-se um modelo para concentrações próximas a zero e outro para as concentrações mais altas. Além de não contemplar a faixa intermediária das concentrações, esta solução contém uma decisão arbitrária sobre o valor limite de concentração em que cada um dos modelos deve ser aplicado. O modelo de dois componentes, apresentado por Rocke & Lorenzato (1995), propõe uma combinação dos dois modelos mais usados, o que supõe variância constante e aquele que considera

constante o erro relativo. Ele descreve o comportamento do erro analítico numa ampla faixa de concentrações retratando o comportamento constante da incerteza nas baixas concentrações e também o erro relativo constante para as concentrações mais altas. Uma das vantagens apresentadas por este modelo é sua aplicabilidade à faixa de valores intermediários da concentração, que é importante na análise de dados ambientais.

A idéia básica na formulação do modelo de Rocke & Lorenzato é classificar em dois tipos, aditivo e multiplicativo, as diversas incertezas presentes no resultado do processo de medição. O modelo de medição é descrito pela expressão que relaciona a medida  $x$  com o valor verdadeiro  $\mu$ :

$$x = \mu e^{\eta} + \varepsilon \quad \text{Eq. 2-3}$$

onde  $\eta$  é o erro proporcional exibido em concentrações relativamente altas com distribuição normal  $N(0, \sigma_{\eta}^2)$  e  $\varepsilon$  é o erro apresentado primariamente nas baixas concentrações;  $N(0, \sigma_{\varepsilon}^2)$ .

### 2.3 Protocolo analítico

Ao planejar a execução de uma análise, cada laboratório começa por selecionar um método de medição que atenda à finalidade e requisitos da análise a ser realizada e às disponibilidades do laboratório. O método de medição contém a descrição genérica da seqüência de operações a serem realizadas no processo analítico incluindo a especificação do modelo de medição e avaliação da incerteza (ABNT,98). O processo de implantação da análise em rotina inclui medidas, testes, especificação do procedimento analítico e termina com elaboração do protocolo analítico. Ele é apresentado em um documento que descreve minuciosamente toda seqüência de operações e cálculos necessários à geração do resultado de uma medição e estimação de sua incerteza. Deve ser elaborado com detalhes suficientes para que a execução da análise possa ser realizada sem necessidade de nenhuma informação adicional. É desenvolvido para uma análise específica e para cada matriz a ser analisada e inclui procedimentos de controle e garantia de qualidade.

O procedimento analítico apresenta uma descrição detalhada de todas etapas da marcha analítica, fixa valores, especifica equipamentos e é elaborado de modo que possa atender aos requisitos de qualidade que o laboratório deseja atingir nesta análise. Como todas as condições e parâmetros do processo analítico foram estabelecidas no protocolo, o limite de detecção e a precisão do método ficam especificados.

## 2.4 Tratamento estatístico de dados ambientais

A finalidade primária do tratamento estatístico dos resultados analíticos de amostras ambientais é fornecer informações para um usuário que está fora do sistema de geração de resultados: aqueles que irão tomar decisões gerenciais, elaborar normas e regulamentações e o público em geral (Porter, 88). Ao planejar e executar o tratamento estatístico de um conjunto de resultados deve-se ter sempre presente que um bom tratamento estatístico não melhora dados ruins e que uma boa estatística pode levar a uma má interpretação (Clark e Whitfield, 93).

O tratamento estatístico de um conjunto de dados já estabelecido começa por uma verificação da qualidade e integridade do conjunto. Detectar erros e inconsistências ocorridos nos procedimentos e métodos analíticos ou de amostragem; determinar características e familiarizar-se com o conjunto de dados, são etapas que precedem e definem o procedimento estatístico a ser adotado em cada caso. A não realização deste estudo prévio pode levar a interpretações incorretas ou inconclusivas, por ignorar ou assumir características que os dados não possuem.

Alguns aspectos característicos dos conjuntos de dados ambientais, como valores discrepantes, dados censurados e lacunas de dados, devem ser cuidadosamente analisados no planejamento do método estatístico a ser aplicado uma vez que sua presença pode induzir conclusões não verdadeiras ou não representativas da realidade a ser descrita. Estas características dificultam e, às vezes, impossibilitam a aplicação de técnicas estatísticas usuais. A presença de dados censurados, pode ser minimizada pelo aprimoramento do processo de medição, mas os resultados já existentes não podem ser alterados e continuam presentes nas séries históricas. Outras, como lacunas de dados e valores discrepantes, necessitam de atenção especial do analista no processamento e tratamento destes resultados.

A qualidade de conjuntos de dados contendo resultados analíticos de amostras ambientais pode ser afetada, em cada uma das etapas de produção destes resultados, por várias causas. O termo qualidade aplicado a um conjunto de dados, deve ser definido em relação à finalidade para o qual ele será utilizado; ele pode ser classificado como excelente para um certo uso e não ser adequado a outro (Marker e Ryaboy, 1994). Num sentido mais amplo, a qualidade de dados é uma medida do grau de adequação ou de utilidade dos dados para um determinado objetivo. Os principais atributos de qualidade de um conjunto de resultados especialmente importantes em estudos ambientais são precisão, viés, representatividade, comparabilidade, sensibilidade e completude (USEPA, 2001b). Os

indicadores de qualidade que dizem respeito diretamente ao processo de medição são a precisão, o viés e a sensibilidade; os outros atributos geram indicadores que descrevem a qualidade do conjunto de informações do programa. Os conjuntos de resultados devem ser verificados e validados para que, após o tratamento estatístico adequado, se tornem úteis para o usuário final da informação. O objetivo da verificação é assegurar que os resultados reflitam a realidade do processo de execução do programa; a validação avalia o impacto das não conformidades técnica ou de qualidade sobre o conjunto de dados (USEPA, 2001a).

#### **2.4.1 Dados discrepantes**

Dados discrepantes são resultados consideravelmente mais altos ou mais baixos em relação ao volume dos dados. É uma observação que aparentemente não se adapta ao comportamento estabelecido pelas outras observações. Aparecem freqüentemente em conjuntos de dados ambientais e podem ter várias causas: erros na anotação ou transcrição de resultados; problemas de calibração ou mau funcionamento de equipamentos, resultado pertencente a uma outra população ou um evento raro em uma população com distribuição muito distorcida. Um dos maiores problemas que o analista ambiental tem ao identificar estes dados é que eles podem estar representando a variabilidade espacial ou temporal inerente ao processo ambiental ou algum fator não identificado relacionado à grandeza analisada.

A visualização gráfica do conjunto de dados (usando gráficos de caixas -“*box and whiskers plots*”-, gráficos temporais, de probabilidade normal, de dispersão) é o primeiro passo para a identificação de possíveis discrepâncias (Triola, 1999). Cada um dos dados classificados nesta inspeção como um dado discrepante em potencial deve ser investigado cuidadosamente pelo analista. Uma possibilidade é identificar a origem do problema por meio de outras fontes de informações como a verificação da consistência interna e temporal e comparação com conjuntos de dados paralelos, tomados da mesma população no mesmo período, na mesma região em outras matrizes (Gilbert, 1987). Os resultados assim identificados podem ser submetidos a testes estatísticos para determinar se estes resultados são realmente discrepantes. Estes testes fazem parte do processo de validação dos dados e fornecem evidências probabilísticas que o valor extremo em questão não se ajusta à distribuição do resto dos dados sendo portanto um dado discrepante.

Após a identificação do dado discrepante, o analista deve decidir o tratamento a ser dado a estes dados, sempre tendo cuidado de analisar cada resultado e documentar claramente todo o procedimento. Para erros de anotação e transcrição, uma simples

investigação pode resolver o problema, devendo o analista então anotar o dado corretamente. Entretanto deve-se ressaltar que muitas vezes o conjunto de dados cobre um intervalo de tempo grande e estas informações podem não estar disponíveis. Nos casos em que a origem da discrepância é comprovadamente falha no controle estatístico, o dado deve ser descartado e o processo documentado.

A recomendação de vários autores (Gilbert,87; Helsel e Hirsh, 1992; Meier, 93; USEPA, 2000b) é o uso da experiência e do bom senso ao interpretar o resultado destes testes: investigar cuidadosa e exaustivamente as causas da discrepância e sempre relatar claramente o não uso de algum resultado que for rejeitado. Dados discrepantes não devem ser descartados baseados no fato de parecerem fora do usual. A rejeição de dados para que o restante do conjunto se apresente ajustado a um tipo de distribuição teórica pré-concebida e propiciar o uso de procedimentos estatísticos desenvolvidos para estas distribuições deve ser evitado. Deve-se usar métodos estatísticos mais robustos e que sejam menos afetados por estes valores.

#### **2.4.2 Dados censurados**

Dados censurados representam um grave problema a ser resolvido pelos estatísticos. Envolvem problemas de conceituação, de terminologia e muitas vezes representam uma importante perda de informação numa área em que toda informação é preciosa e não pode ser descartada. Os métodos clássicos da estatística descritiva aplicam-se a conjuntos de dados em que todos resultados sejam valores numéricos; a presença de resultados censurados, como é usual acontecer em estudos ambientais, torna não-trivial o tratamento estatístico destes conjuntos.

A existência de valores publicados como não detectados (ND) nos conjuntos de dados ambientais é um fato freqüente, às vezes inevitável para os analistas que executam as análises e sempre não desejado pelos estatísticos responsáveis pela interpretação destes dados. Resultados são assinalados como ND quando ocorrem dificuldades na quantificação da medida. Tratando-se de valores próximos ao zero, as limitações técnicas devem-se basicamente a um sinal muito fraco para o equipamento distinguí-lo do ruído de fundo ou a um sinal muito baixo e que pode ser confundido com o sinal de uma amostra branco. Em certos tipos de análises, valores marcados como ND podem também ocorrer em virtude da falta de alguma característica de identificação da grandeza analisada. Além destas razões, freqüentemente o analista, ao estabelecer o protocolo de medição, fixa um valor crítico que, na verdade, estabelece o padrão de qualidade exigido por ele e só publica valores maiores que este limite.

São diversas as óticas com que são tratados estes valores marcados como ND: do analista que gera o resultado da medida, do estatístico que analisa o conjunto de dados, e finalmente do usuário da interpretação estatística que irá tomar decisões baseadas nestas análises. O analista publica valores considerados detectados se eles preencherem seus próprios requisitos de qualidade (considerando todo o processo analítico). O estatístico vê neles falta de informação e a invalidação do uso de técnicas convencionais no tratamento estatístico. O usuário final da informação sente insegurança ao tomar decisões baseadas na interpretação destes dados. Esse tipo de dado causa problemas a todos técnicos envolvidos nos estudos ambientais e têm sido feitos grandes esforços para evitar a presença deles nos conjuntos de dados. Agências regulatórias têm recomendado a publicação do valor obtido no procedimento analítico, incluindo valores negativos, e informações adicionais sobre o valor do limite de detecção, sua definição e metodologia de determinação (Gilbert, 1987; USEPA, 2000b). Dados censurados contêm informações relevantes apesar de sua grande incerteza. Adotar a norma de não censurar os resultados analíticos aumenta a capacidade de detectar tendências e amplia os métodos estatísticos disponíveis para utilização nestes conjuntos (Gillion, 1984).

Dados censurados estão presentes em séries históricas e não há como descartá-los ou substituí-los, mas deve-se encontrar um modo de tratá-los estatisticamente. Contrapondo-se ao desenvolvimento tecnológico que possibilita limites de detecção cada vez mais baixos, a busca pela quantificação de concentrações cada vez menores torna-se uma exigência de ambientalistas e do público em geral, de modo que valores menores que o limite de detecção (LD) continuam a ser encontrados e censurados apesar das recomendações para uso dos valores encontrados no processo de medição.

#### **2.4.2.1 Limite de detecção**

A finalidade do processo de medição, expresso pelo resultado do procedimento analítico executado nessa determinação, é encontrar o valor verdadeiro da grandeza em estudo. A relação entre estes dois valores, o valor real e o reportado, é linear num determinado intervalo mas afasta-se da linearidade em virtude de causas diversas. A componente aleatória do erro analítico cresce rapidamente à medida que esta linearidade deixa de reger o comportamento da relação. O modelo desta relação é apresentado na FIGURA 2-2. A região central, onde a linha sólida central representa o melhor ajuste entre o valor observado e o valor verdadeiro, corresponde à região de quantificação. As duas linhas curvas externas mostram o intervalo determinado pelo erro aleatório. A área entre as faixas marcadas B e C (assinalada 2) é a região de detecção, onde o erro aleatório cresce

mais rapidamente e a relação valor observado x valor verdadeiro não é mais linear. Na região abaixo da faixa A o erro analítico é tão grande que não é possível distinguir entre a amostra e o *background* (ou “branco”). As faixas A, B e C correspondem aos critérios de detecção, limite de detecção e de quantificação, respectivamente (Clark e Whitfield, 1994). Estes limites também ocorrem, de maneira inversa, na região de altas concentrações.

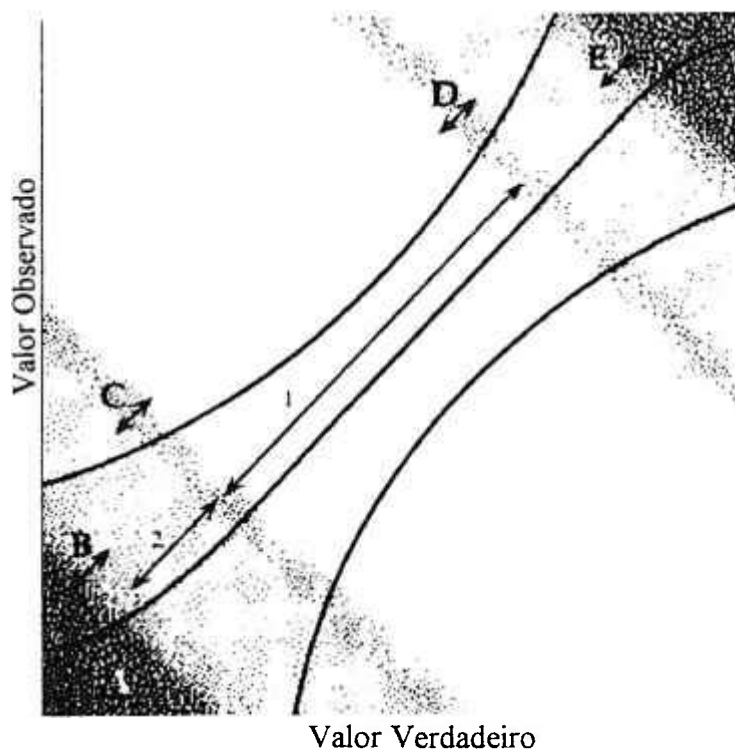


FIGURA 2-2 Modelo do relacionamento valor verdadeiro - valor observado

A região onde a resposta é linear e o erro aleatório é mínimo corresponde ao intervalo no qual são alcançados os melhores resultados analíticos. Percorrendo a escala, na direção dos menores valores ou dos maiores, onde a relação não é mais linear, encontram-se regiões onde a existência de maiores erros aleatórios diminui a qualidade dos resultados chegando às extremidades onde já não se pode mais falar em quantificação. Entretanto há ainda a região de detecção, onde é possível distinguir a amostra branco da amostra contendo o constituinte de interesse e um limite crítico abaixo do qual torna-se difícil distinguir o sinal de sua flutuação estatística. É o ponto de decisão se o sinal foi detectado ou não: o limite de detecção.

O limite de detecção refere-se à capacidade do método analítico de distinguir entre amostras que contem um teor muito baixo e aquelas que não contem, isto é, seu foco é na concentração que pode ser distinguida do ruído de fundo. Existe uma enorme variedade de conceitos e definições de limites de detecção apresentados na literatura,

recomendados por instituições regulatórias e fiscalizadoras sem que se tenha unanimidade (USEPA, 2001a). O conceito de limite de detecção adotado pela USEPA (1997) é fundamentado num trabalho clássico apresentado por Curie (68) e desenvolvido por Glaser et al. (1981) que estabeleceu a definição de Limite de Detecção do Método (LDM), com o seguinte enunciado:

*“O Limite de Detecção do Método é definido como a concentração mínima de uma substância que pode ser identificada, medida e publicada com 99% de confiança de que a concentração do item analisado seja maior que zero e é determinado por análises duplicadas de uma amostra de uma dada matriz contendo esse item.”*

Sua determinação é feita por meio de 7 pseudo-amostras, de concentração presumida próxima ao zero ou amostras-branco, analisadas em condições de repetibilidade. O desvio padrão da distribuição do conjunto de resultados obtidos pela aplicação do protocolo analítico a estas amostras é usado no estabelecimento do LDM juntamente com o valor da estatística t de Student relativo ao nível de confiança estipulado na definição. O LDM é calculado pela expressão:

$$\text{LDM} = s \times t_{(n-1, [1-\alpha]=0,99)} \quad \text{Eq. 2-4}$$

onde s é o desvio padrão das concentrações medidas das determinações repetidas e t é o valor de Student para o nível de confiança  $(1-\alpha) = 0,99$  sendo n o número de determinações realizadas

O Limite de Detecção do Método acrescenta uma padronização ao conceito de limite crítico estabelecendo operacionalmente sua determinação. Várias são as limitações a essa definição: assume distribuição normal e desvio padrão constante; usa uma matriz “limpa” como referência (portanto ela não leva em conta as interferências nas matrizes reais); sua determinação é feita num intervalo de tempo pequeno, não captando a variabilidade temporal do procedimento nem as influências que ocorrem em métodos analíticos executados em laboratórios onde são utilizados múltiplos instrumentos e operadores. Os valores definidos pelo Limite de Detecção do Método são determinados em condições diferentes das amostras analisadas na rotina analítica e muitas vezes resultam em valores muito baixos que não englobam a variabilidade global do procedimento analítico. Alguns laboratórios dedicados a medidas ambientais (Childress, 1999), buscando valores mais realistas, usam um procedimento para a determinação do LDM que leva em conta a variabilidade global do protocolo. Neste conceito, o valor de referência é determinado usando-se medidas tomadas em um intervalo de tempo grande (de 6 meses a 1 ano), um grande número de amostras (ao menos 24 por ano) e submetidas à rotina do protocolo

analítico. Deste modo os valores encontrados capturam a variabilidade típica das análises executadas naquele laboratório.

Um outro método de determinação, sugerido por Osborne e Rocke (2000), é baseado nos resultados produzidos nos estudos de controle de qualidade executados rotineiramente no laboratório. O desvio padrão da concentração zero é determinado usando, na expressão do modelo de dois componentes, os valores do desvio padrão da amostra-branco e o desvio padrão relativo. Este é um método mais robusto e com maior reprodutibilidade na determinação do limite de detecção pois inclui a variância associada com o método analítico completo e utiliza resultados obtidos na rotina diária do laboratório.

Na verdade, o limite de detecção é mais uma estatística que um conceito físico ou químico. Não tem um significado até ser definido operacionalmente em um determinado procedimento de medição e um método estatístico para analisar as medidas produzidas (Berthouex and Brown, 1994). Ele não é uma propriedade fundamental do equipamento, nem é um valor constante e fixo. É definido pelo protocolo de uma rotina analítica que especifica as condições de medida. Pode ser alterado por mudanças neste protocolo, por exemplo, para atender a necessidade de se analisar amostras de teor mais baixo, requisitos legais, exigências do cliente, troca de equipamento. Na etapa de planejamento do protocolo analítico deve-se escolher a definição estatística do LD, definir as regras operacionais para sua determinação e as implicações destas na interpretação dos resultados produzidos.

#### **2.4.2.2 Métodos estatísticos para conjuntos contendo dados censurados**

Os métodos clássicos da estatística descritiva e inferencial aplicam-se a conjuntos de dados em que todos resultados sejam valores numéricos; a presença de resultados censurados, como é usual acontecer em estudos ambientais, torna não-trivial seu tratamento estatístico.

O tipo de dados censurados encontrados em estudos ambientais é aquele classificado como tipo I, censurado à esquerda. O censuramento à direita, isto é, valores observados maiores que um limite máximo de detecção, também pode ocorrer mas é um problema menos crítico. Decisões de remediação tomadas com base em valores “maior que” obviamente não são ambíguas nem contestáveis (Hertzeler et al., 1991).

Visando usar toda informação disponível, diversos métodos foram desenvolvidos para tratar o problema da presença de resultados censurados, mesmo que

eles, reconhecidamente, contenham maior incerteza que os resultados acima do limite de detecção (Davis, 1994). Análises ambientais são difíceis, dispendiosas e não permitem replicação, o tamanho das amostras é quase sempre pequeno: usar toda a informação disponível é fundamental. É consenso entre analistas ambientais que “um resultado ruim é melhor que resultado nenhum”, de modo que resultados censurados não devem ser descartados mas tratados com métodos estatísticos apropriados (Nagaraj, 1994).

Na prática, a natureza destes valores apresentados como não detectados (ND) é desconhecida. O censuramento que produz estes dados é complexo e a relação entre não detectados, detectados e o limite de detecção não é clara (Lambert et al., 1991). Além da grande variedade de conceitos usados na definição do valor de corte, é raro o conjunto de dados que contenha informações de como foram determinados estes limites. O estatístico sabe apenas que o valor medido é menor ou igual a um limite especificado pelo analista. Embora alguns valores marcados como ND ocorram em virtude da incapacidade do sistema analítico de quantificá-lo, outros são assim apresentados por não atingirem um requisito de precisão especificado pelo analista (por exemplo, erro relativo de 10%). Neste caso, existe um resultado numérico que poderia ser recuperado e, apesar de sua imprecisão, compor o conjunto usado na análise estatística eliminando o tratamento necessário aos ND.

Nos métodos para tratamento estatístico de conjuntos de resultados contendo dados censurados existem duas abordagens: a substituição simples por algum valor especificado e o uso de técnicas estatísticas para o tratamento destes resultados. De maneira geral estas técnicas podem ser divididos em paramétricas (distribucionais) e não-paramétricas (livres de distribuição) (Gilbert, 1995). Vários estudos demonstram as grandes diferenças na capacidade destas técnicas de descrever corretamente conjuntos de dados censurados (Hertzeler et al., 1991). A escolha do método de tratamento adequado deve ser determinada por vários fatores como o grau de censuramento (número de dados censurados comparados com o tamanho da amostra), o tamanho da amostra e, sobretudo, a finalidade da análise estatística.

Métodos de substituição transformam o conjunto de dados contendo dados censurados em outro formado pelas observações não censuradas e os dados obtidos da substituição de cada dado censurado por um valor, cuja escolha é arbitrária. Os valores mais usados nestas substituições são o zero, a metade do limite de detecção (LD) ou o próprio LD. A partir do conjunto de valores numéricos assim obtidos (dados acima do limite e os valores substituídos) são calculadas todas as estatísticas necessárias às diversas finalidades da análise ambiental. Embora largamente empregados, estes métodos não têm

base teórica e apresentam bem conhecidos problemas de tendenciosidade. A substituição por zero desvia a média para valores menores que o verdadeiro; na substituição pelo LD (justificada por ser conservativa), a tendência é para valores mais altos que o verdadeiro. Em alguns casos esta tendenciosidade é usada propositalmente, como no caso de agências regulatórias que estipulam a substituição pelo valor LD para produzir uma estimativa conservativa objetivando maior segurança à saúde do público (Gleit, 85).

Estes métodos de substituição resultam em conjuntos de dados cuja distribuição apresenta lacunas na região censurada e não descrevem a distribuição real (Helsel, 1990). Esta situação é ilustrada na FIGURA 2-3 para o caso de um limite único de detecção.

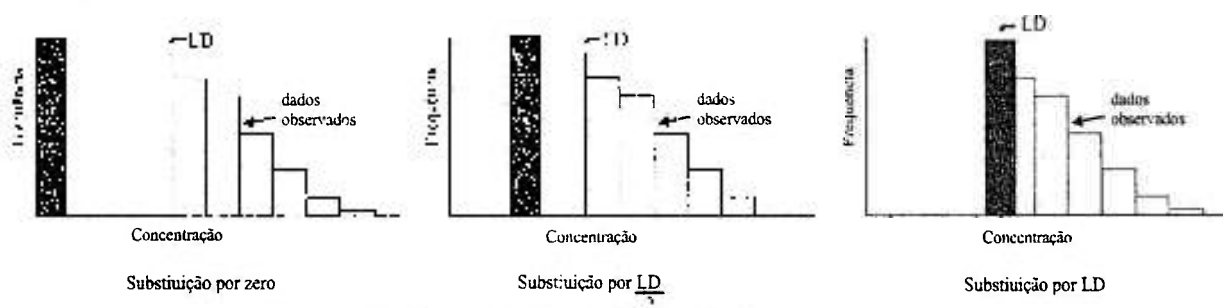


FIGURA 2-3 Métodos de substituição simples

A estatística descritiva é determinada substituindo os valores menores que o LD por zero,  $\frac{1}{2}LD$  ou pelo LD e usando estes dados e os resultados publicados maiores que o LD.

O uso destes métodos na estimação de parâmetros não é recomendado em razão de arbitrariedade da escolha do valor a ser substituído e à tendenciosidade das estimativas obtidas. O argumento de simplicidade não mais se sustenta nos dias de hoje com as facilidades de cálculo disponíveis em qualquer laboratório. Estes métodos são citados como altamente não recomendáveis por vários autores (Helsel e Hirsch, 1992; Akritas et al., 1994; Gilbert, 96). Entretanto alguns autores (Atwood et al., 91 e Davis, 95) sugerem que o valor  $\frac{1}{2}LD$  para substituição dos ND pode ser usado nos testes de hipóteses uma vez que sua performance não é pior de que outras técnicas mais sofisticadas e portanto a simplicidade justificaria seu uso.

Métodos paramétricos ou distribucionais usam as características de uma determinada distribuição para estimar estatísticas: é necessário supor um tipo de distribuição para o conjunto de dados a ser analisado. A distribuição mais usada no estudo de dados ambientais é a log-normal. A teoria da difusão de radionuclídeos ou de contaminantes é uma das teorias citadas para justificar o uso desta distribuição. Mas talvez a principal justificativa para seu uso seja a de que ela é o instrumento mais disponível entre as distribuições estatísticas, que possui as características gerais requeridas para o uso em

estudos ecológicos: dados intrinsecamente positivos, simetria positiva e uma pesada cauda para a direita (Atwood et al., 1991).

As técnicas mais usadas para a estimação da média, desvio padrão e também dos percentis em conjuntos de dados censurados são a da máxima verossimilhança (MV) e os gráficos de probabilidade (Helsel e Hirsh, 92). Estas duas técnicas dependem de um bom ajuste dos dados à distribuição escolhida e do tamanho da amostra. Se o ajuste for bom e o tamanho da amostra for suficientemente grande ( $n > 25$ ), estas estimações são não-tendenciosas (Gleit, 85). Infelizmente, em se tratando de dados ambientais, é raro o conjunto de dados que satisfaz a estas duas condições. Normalmente os conjuntos de dados são pequenos não permitindo facilmente a identificação ou determinação da função distribuição e os estimadores para amostras pequenas são tendenciosos. O limite inferior para o tamanho da amostra seria  $n > 20$  (Cohen, 1991).

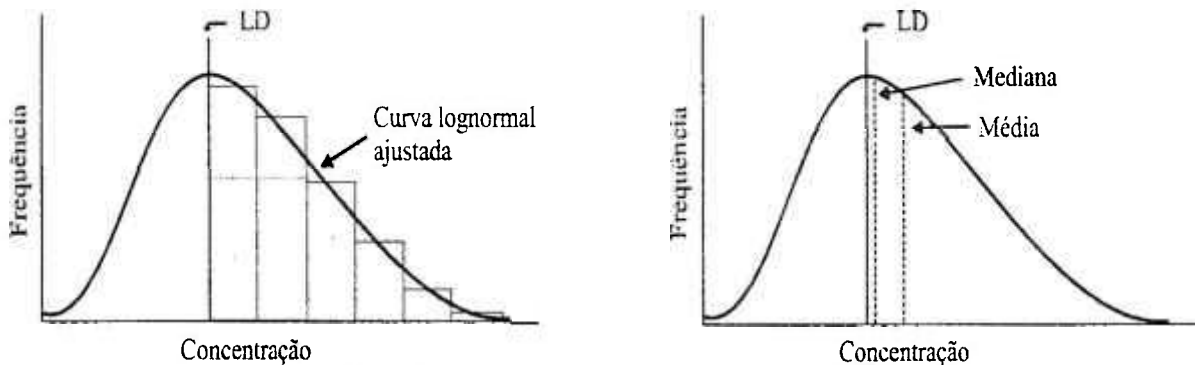


FIGURA 2-4 Método da Máxima Verossimilhança (MV)

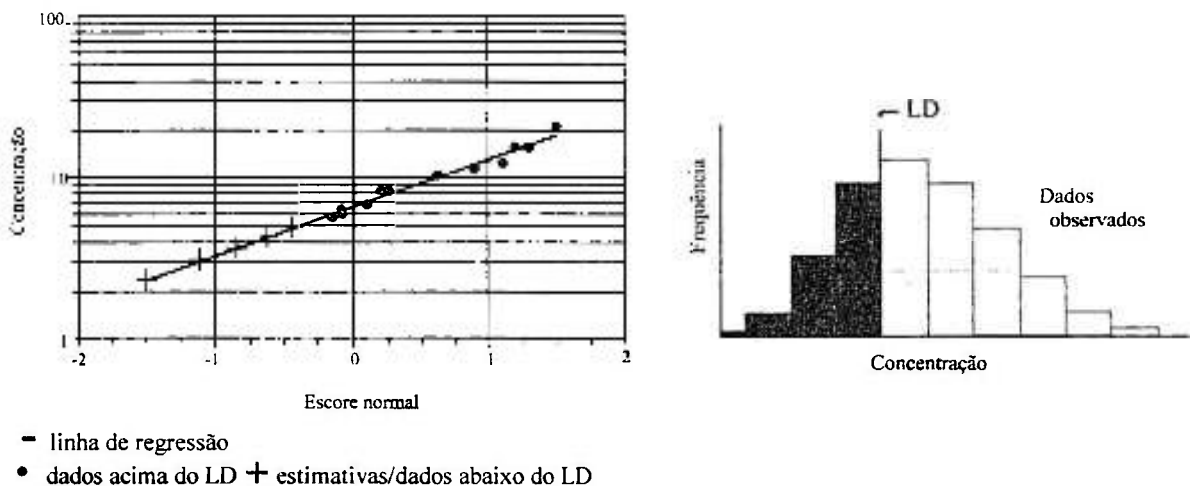
A estatística descritiva é determinada usando os valores obtidos pelo ajuste dos dados acima do LD à expressão da distribuição lognormal.

O método da máxima verossimilhança (FIGURA 2-4) estima os parâmetros por meio de ajuste aos dados originais (ou transformados, como no caso de curva log-normal). A precisão destas estimativas porém depende do tamanho da amostra e da qualidade do ajuste.

Gráficos de probabilidade usados nas estimativas de média, desvio padrão e percentis apresentam as mesmas desvantagens do MV (dependem de um bom ajuste e do tamanho da amostra), entretanto possuem a vantagem de maior facilidade de uso.

Outros métodos paramétricos estimam  $\mu$  e  $\sigma$  à partir dos dados não censurados e calculam os valores esperados das observações censuradas supondo uma distribuição normal (ou log-normal). Usando o conjunto de dados assim formado (observações não-censuradas e aquelas calculadas para substituir os dados censurados)  $\mu$  e  $\sigma$  são recalculados. Gleit (85) apresenta uma técnica para este método que usa a MV e ajustes iterativos.

Métodos robustos para dados censurados combinam os dados observados que estão acima do limite de detecção com valores extrapolados para os dados abaixo do LD para estimar a estatística descritiva. Usando procedimentos de MV ou gráficos de probabilidade é ajustada uma curva de distribuição aos dados observados acima do LD e os valores censurados são encontrados por meio de extrapolações. Estes valores não são considerados como estimativas dos pontos abaixo do limite, mas considerados coletivamente para estimar as estatísticas descritivas do conjunto. Assim, as estimativas são feitas usando diretamente os resultados acima do limite de detecção em lugar da curva ajustada e é evitada a tendenciosidade que seria introduzida pela transformação de unidades já que as estimativas são feitas nas unidades originais.



**FIGURA 2-5 Método da regressão log-normal**

A estatística descritiva é determinada usando os valores publicados acima do LD e os valores estimados por extrapolação através da reta de regressão.

Este método, ilustrado na FIGURA 2-5, combina a técnica do gráfico de probabilidade com a extrapolação e é também conhecido como regressão log-normal. Supondo uma distribuição normal ou log-normal é construído o gráfico de probabilidade usando os dados acima do limite, ajustada a melhor reta e os valores censurados são extrapolados por essa reta. Dessa forma obtém-se um conjunto de resultados numéricos que são usados para determinar a estatística descritiva do conjunto original.

Conjuntos de dados ambientais apresentam muitas vezes distribuições com assimetria elevada de tal forma que, mesmo sem haver censuramento, a estimação pelo método de momentos é tendenciosa, especialmente no caso de amostras pequenas. Nestes casos é recomendável caracterizar tais dados por meio de quartis usando métodos não paramétricos. Quando o grau de censuramento é severo e mesmo quando todos os

resultados estão abaixo do limite de detecção (incluindo múltiplos limites) pode-se determinar a mediana e os interquartis.

### **2.4.3 Análise exploratória**

A análise exploratória dos dados, etapa inicial de qualquer tratamento estatístico, é indispensável para o caso de dados ambientais. O objetivo desta análise preliminar é conhecer a estrutura dos dados disponíveis, para que o analista tenha uma visão mais profunda da natureza destes, possa identificar abordagens mais apropriadas e, sobretudo, as limitações que a base de dados contém (Diggle et al., 1996).

Esta análise, em que não são necessárias hipóteses nem modelos, compreende duas etapas principais: o cálculo das estatísticas básicas e a representação gráfica dos dados. Este sumário estatístico quantifica a locação e a variabilidade dos dados por medidas de tendência central, de dispersão e de distribuição. A representação gráfica dos dados visa identificar padrões e relacionamentos entre os dados.

A execução da análise exploratória é precedida de um manuseio dos resultados de modo a colocá-los em tabelas compatíveis com o recurso computacional a ser usado em seu processamento.

### **2.4.4 Estimação da média**

O resultado de uma medição é uma apenas uma estimativa do valor real da grandeza a ser medida. Esta estimativa será tanto melhor quanto menor for a incerteza da medição. O processo de medição deve, portanto, incluir tanto a determinação do valor do mensurando quanto a quantificação desta incerteza. Operacionalmente, a quantificação de uma grandeza requer uma série de observações obtidas em condições de repetibilidade, gerando um conjunto de resultados. A descrição estatística deste conjunto irá fornecer o valor do mensurando e a incerteza associada a ele (ABNT, 1998).

Um conjunto de dados pode ser caracterizado basicamente por uma medida de tendência central, uma estimativa de variabilidade e pela forma da distribuição de freqüências. A medida de tendência central pode ser definida de várias maneiras: pela média, mediana e moda. No caso de distribuições simétricas e unimodais, estas três medidas são iguais. Cada uma destas medidas apresenta vantagens e desvantagens dependendo da finalidade da estimação e das características do conjunto de dados a ser descrito. A variabilidade em relação à média é expressa pela variância ou pelo desvio padrão (DP), que é numericamente igual raiz quadrada da variância. As medidas mais usadas para descrever um conjunto de dados são a média e o desvio padrão. Quanto à

forma da distribuição, a maioria das técnicas e testes estatísticos é desenvolvida para distribuição normal, que é simétrica e unimodal.

Os parâmetros de uma distribuição podem ser estimados por vários métodos como o de máxima verosimilhança, mais importante e mais usado dos métodos; o dos momentos, interessante por sua simplicidade ou dos mínimos quadrados, o mais conhecido embora seu uso seja restrito a modelos lineares ou generalização destes modelos (Garthwaite, 1995). Todos estes métodos baseiam-se no conhecimento da função distribuição de freqüências do conjunto de dados, o que restringe bastante seu uso no caso de resultados ambientais, que costumam apresentar distribuições fortemente distorcidas para direita.

No processo de medição, a série de observações repetidas segue uma distribuição normal, cujos parâmetros  $\mu$  e  $\sigma^2$  podem ser representados pela média amostral  $\bar{x}$  e pela variância amostral  $s^2$ , se o número de observações for suficientemente grande. Usando qualquer dos métodos de estimação disponíveis, pode-se encontrar as expressões para determinação deste parâmetros:

$$\hat{\mu} = \bar{x} = \frac{1}{N} \sum x_i \quad \text{Eq. 2-5}$$

$$\hat{\sigma}^2 = s^2 = \frac{1}{N-1} \sum (x_i - \bar{x})^2 \quad \text{Eq. 2-6}$$

A média amostral é um dos mais usados e o melhor estimador da média da população. No caso da distribuição normal é um estimador não tendencioso, coerente e de variância mínima, que são as propriedades desejáveis de um bom estimador (Meyer, 1969). Uma das vantagens da média como medida da tendência central é que ela usa todos valores do conjunto porém é fortemente afetada por valores extremos. No caso de distribuições simétricas de conjuntos contendo valores extremos ou dados censurados, pode ser usada a média aparada (Triola, 1999). Esta média elimina valores nas duas extremidades da distribuição segundo critérios pré-estabelecidos porém distribuições simétricas são raramente encontradas em conjuntos de resultados ambientais.

### 3. MÉTODO PARA ESTIMAÇÃO DE MÉDIAS EM ANÁLISES AMBIENTAIS

#### 3.1 Estimação da média de resultados analíticos

As expressões para a estimação da média e do desvio padrão expressos nas eq.2-5 e 2-6 são formuladas sob a hipótese de que a mesma distribuição é válida para cada um dos resultados que formam o conjunto de dados. Isto significa que as observações feitas na determinação de cada medida têm a mesma variabilidade ou seja, têm o mesmo desvio padrão. Estas expressões são apropriadas para conjuntos de resultados analíticos que possuem estas características estatísticas e que foram produzidos segundo um modelo de medição de variância constante (Gilbert, 1987). A relação matemática deste modelo é

$$x_i = \mu + e \quad \text{Eq. 3-1}$$

onde  $x_i$  é o valor observado,  $\mu$  é o valor verdadeiro da grandeza medida e  $e$ , a incerteza analítica (com distribuição normal de média zero e desvio padrão  $s_e$ ).

Freqüentemente os dados contidos em um conjunto de resultados analíticos são produzidos em condições diversas levando a valores diferentes para o desvio padrão associado a cada resultado. Para expressar quantitativamente esta situação, a média deste conjunto de resultados é determinada usando um fator de ponderação. Bevington (1999) aplicou o método da máxima verossimilhança de estimação de parâmetros para mostrar que este fator, para cada medida, é igual ao inverso de sua variância. A expressão para determinar este fator é:

$$a_i = \frac{1}{s_i^2} \quad \text{Eq. 3-2}$$

onde  $a_i$  é o fator de ponderação para a medida  $x_i$  que apresenta um desvio padrão  $s_i$ .

O valor mais provável para a média deste conjunto de resultados, considerando ainda o modelo de medição com variância constante, será dada por:

$$\hat{\mu} = \bar{x} = \frac{\sum x_i a_i}{\sum a_i} = \frac{\sum \left[ \frac{x_i}{s_i^2} \right]}{\sum \left[ \frac{1}{s_i^2} \right]} \quad \text{Eq. 3-3}$$

O desvio padrão da média será dado por:

$$\hat{\sigma}_{\mu} = \sqrt{\frac{1}{\sum a_i}} = \sqrt{\frac{1}{\sum \left[ \frac{1}{s_i^2} \right]}} \quad \text{Eq. 3-4}$$

Esta expressão para a estimativa da média mostra a importância da determinação correta da incerteza analítica no processo de medição de grandezas em amostras ambientais. O valor da incerteza de cada resultado influencia o valor da média representativa do conjunto e portanto sua estimativa é crítica na análise de dados ambientais.

### 3.2 Resultados analíticos produzidos em PMA's: dados históricos

Os dados gerados em um programa de monitoração ambiental são o produto de um processo de medição e portanto são representados por um valor, que expressa o resultado da medição, e uma incerteza associada a esse valor. As expressões para média e desvio padrão formuladas para aplicação em conjuntos de resultados produzidos em condições de repetibilidade, isto é, apresentando a mesma incerteza, não são adequadas a um conjunto de resultados analíticos gerados em PMA's. Estes resultados não apresentam o mesmo valor para a incerteza analítica. Duas causas principais concorrem para isto: a geração dos resultados ao longo de um largo período de tempo e a amplitude dos valores da concentração. Mudanças no sistema operacional do laboratório são inevitáveis ao longo de um período de tempo tão grande como o de execução de um PMA. Incorporação de novas tecnologias, equipamentos com maior automação, renovação da equipe de operadores são alguns motivos destas mudanças. A avaliação da incerteza é, portanto, afetada por causas diferentes ao longo do período de formação do conjunto de dados. No caso do conjunto de dados estudados neste trabalho, por exemplo, para as análises executadas na fase de implantação do programa, a automação de equipamentos era mínima, os resultados eram processados em calculadoras manuais e todos registros feitos manualmente.

A amplitude das concentrações encontradas é o outro fator relevante na variabilidade das incertezas, em virtude da dependência entre a incerteza analítica e a concentração. Resultados de monitoração ambiental situam-se tipicamente na região das

concentrações próximas de zero e numa faixa intermediária de valores, mas concentrações mais altas podem ocorrer. Os modelos de medição tradicionalmente adotados pelos laboratórios na execução de análises ambientais e, em particular, os modelos disponíveis à época da entrada em operação de grande parte dos programas de monitoração, não contemplam o domínio pleno das concentrações. A utilização destes modelos poderá acarretar distorções na avaliação da incerteza de medidas situadas em faixas de valores onde o modelo não é adequado. Uma das soluções mais usadas para este problema é o uso de dois modelos; um apropriado para resultados próximos a zero e outro modelo para concentrações mais elevadas. Neste caso, as incertezas associadas a concentrações intermediárias são arbitrariamente tratadas por um ou outro modelo. No caso da aplicação do modelo baseado na estatística de Poisson pode-se, a priori, observar que nas regiões das baixas concentrações ele prevê valores crescentes para incerteza onde ela deveria ser constante. Nas regiões das concentrações mais altas, a relação entre a incerteza e a concentração para este modelo, tende a ser assintótica a um certo valor, o que contraria o comportamento observado para esta relação.

Outro ponto a ser considerado é que a distinção entre cada faixa não é um limite fixo mas depende do protocolo analítico. As características da instrumentação, os valores fixados no protocolo, como a alíquota analisada, o tempo de contagem nas análises radiométricas, influenciam na delimitação destas faixas, o que é baixa atividade para um protocolo analítico é, às vezes, concentração intermediária em outro. Desta forma é difícil para o laboratório especificar qual modelo se aplica melhor a cada resultado produzido.

### **3.3 Incerteza “corrigida”: aplicação do modelo de dois componentes**

As expressões da média e desvio padrão aplicáveis ao conjunto de dados formado pelos resultados analíticos gerados ao longo da operação de um PMA são certamente as equações 3-3 e 3-4. Elas ressaltam a influência da avaliação da incerteza no resultado da estimação da média representativa do conjunto de resultados analíticos já que seus valores são usados na determinação dos fatores de ponderação para cada medida.

No conjunto de resultados analíticos de um PMA, a avaliação da incerteza de cada medida poderá estar afetada, primariamente, por dois fatores: a variação temporal das condições de operação e a modelagem usada na relação erro-concentração. A variação temporal das condições de operação é traduzida na dispersão dos valores da incerteza para uma mesma concentração. Para um determinado protocolo analítico, em condições operacionais estáveis ao longo do tempo, os resultados analíticos referentes à mesma

concentração apresentariam o mesmo valor para a incerteza segundo a relação erro-concentração adotada. Para medidas de concentração onde o modelo de medição não é adequado, a incerteza pode estar apresentando uma distorção em relação a seu real valor.

A curva erro x concentração apresentada pelos resultados analíticos publicados retrata a função usada na geração destes dados. Para resultados produzidos em condições de controle estatístico, este gráfico apresenta uma linha definida pela expressão algébrica do modelo. A dispersão de resultados em uma mesma concentração representa a variação temporal da incerteza. Uma seqüência de pontos fora da curva projetada pelo modelo usado pode representar uma não adequação do modelo à realidade experimental. A função usada na determinação das incertezas pode ser identificada na análise exploratória de dados e seus parâmetros podem ser estimados. Esta determinação, feita por ajuste aos resultados publicados, garante que estes parâmetros representem as condições de desempenho do laboratório em amostras verdadeiras e não em amostras de controle, pseudo padrões ou amostras branco. Eles representam a efetiva performance do laboratório no período considerado.

A escolha do modelo a ser usado na reavaliação das incertezas deve ser resultado da observação cuidadosa do comportamento do erro em função da concentração; deve ser escolhido o que melhor descreve este comportamento. O modelo de dois componentes, apresentado por Roche & Lorenzato em 1995, é, entre os modelos apresentados, o que melhor descreve o comportamento da incerteza analítica em função da concentração, no domínio pleno das concentrações. Os parâmetros deste modelo são indicadores de desempenho do laboratório, eles representam a sensibilidade e a precisão de um determinado protocolo analítico. A sensibilidade, que pode ser expressa pelo limite de detecção (USEPA, 2001b), é determinada pelo parâmetro  $\sigma_e$  do modelo de dois componentes, numericamente igual ao desvio padrão das concentrações muito próximas de zero. O outro parâmetro do modelo representa o desvio padrão relativo nas concentrações mais altas é, portanto, um indicador de precisão do protocolo analítico.

Os parâmetros do modelo selecionado, que foram determinados por ajuste aos resultados erro-concentração publicados, estarão representando o desempenho do laboratório em relação ao protocolo analítico. As incertezas previstas por meio deste modelo estarão corrigidas das distorções temporais e daquelas relativas à não adequação do modelo usado em sua estimação. A média representativa da concentração do conjunto de resultados apresentará o seu melhor valor por incorporar em sua estimação as informações contidas coletivamente no conjunto de dados formado no período de tempo focalizado.

Considerando o conjunto de resultados das concentrações referentes a amostras coletadas em um mesmo local, a média estimada por este método fornecerá a melhor descrição da grandeza medida.

A aplicação deste método para a reavaliação das incertezas analíticas é limitada pelo número de observações apresentadas no conjunto de resultados a ser analisado. As medidas referentes a cada local e que irão compor a média que descreve a grandeza de interesse, não constituem um conjunto com número de resultados para compor uma amostra de tamanho suficiente do ponto de vista estatístico. A amplitude do valor das concentrações é pequena, ela reflete variações temporais e a dispersão do processo de medição em relação a uma só amostra. Estes dois motivos, número pequeno de resultados e valor de concentração restrito a uma única faixa, podem inviabilizar a aplicação do método aqui proposto. Entretanto a relação erro x concentração refere-se a um protocolo analítico que é especificado para um tipo de análise e matriz. Portanto pode-se usar todos os resultados analíticos produzidos por este protocolo para encontrar os parâmetros que irão descrever o desempenho do laboratório. Estes resultados irão compor uma amostra estatística que será usada no ajuste ao modelo selecionado para descrever o comportamento erro x concentração dos resultados produzidos. Esta amostra representa o produto do processo de medição, não a grandeza física que foi medida.

Em resumo, o método aqui proposto consiste em:

- tratar todos os resultados produzidos por um determinado protocolo analítico como uma amostra estatística;
- verificar a aderência destes resultados a um modelo de medição;
- selecionar o modelo de medição que melhor descreve o comportamento observado, caso o modelo adotado não satisfaça esta condição;
- determinar os parâmetros do modelo por ajuste aos resultados publicados;
- reavaliar as incertezas analíticas usando os parâmetros determinados;
- estimar as médias representativas de cada local amostrado usando as incertezas reavaliadas no cálculo do fator de ponderação.

Considerando o bom desempenho do modelo de Roche&Lorenzato na descrição do comportamento da relação erro x concentração, ele será usado para retratar esta relação nos estudos de caso selecionados para ilustrar o método.

### 3.4 Metodologia

O conjunto de dados gerados pela execução de um Programa de Monitoração Ambiental (PMA) contém resultados de medidas de uma mesma grandeza e de uma mesma matriz, realizadas ao longo de vários anos em locais pré-selecionados. A forma de apresentação destes resultados é especificada no protocolo analítico que descreve o procedimento de análise. De modo geral os resultados analíticos são expressos por:

$$(x_i ; s_i) L \quad \text{Eq. 3-5}$$

onde  $x_i$  é um número que representa o valor da grandeza medida,  $s_i$  representa a incerteza analítica e  $L$  é a unidade de medida no qual o resultado foi expresso.

As duas informações (valor da medida e valor da incerteza) formam um par resultante da aplicação do processo analítico em uma amostra. Todos resultados referentes a um determinado tipo de análise e uma matriz específica, são produzidos pela aplicação de um mesmo protocolo analítico e constituem uma amostra estatística que representa o processo de medição executado pelo laboratório durante o período focalizado. Os valores dos parâmetros do modelo de medição revelam características do laboratório relativas à qualidade. Algumas destas características são determinadas por decisões tomadas na formulação do protocolo analítico como alíquota analisada, tempo de contagem no caso de medidas radiométricas; outras são relativas à instrumentação, como eficiência e *background*. A análise da relação entre o erro e a medida da concentração irá mostrar a fidelidade entre os resultados produzidos e a relação analítica especificada no modelo de medição. A não conformidade entre esta expressão e os resultados produzidos pode significar que as condições operacionais não se mantiveram constantes ou que a relação que descreve este relacionamento não é adequada. Para um laboratório que executou as análises em condição de controle analítico e seguiu rigorosamente o procedimento especificado deve ocorrer um alto grau de aderência entre resultados produzidos e a expressão matemática do modelo de medição. A qualidade do ajuste dependerá também da capacidade do modelo de descrever adequadamente o processo analítico.

Considerando a amostra estatística que contém todos os resultados gerados pela aplicação do protocolo analítico por um laboratório, os parâmetros do modelo de R&L são estimados por ajuste entre a expressão analítica do erro previsto pelo modelo e os resultados produzidos para concentração e incerteza. Estes parâmetros representam os valores integrados no tempo das características do laboratório para o protocolo analítico que foram efetivamente praticadas no período focalizado e incorporam também as diferenças entre o modelo usado no processo de medição e o modelo de R&L.

A expressão para a variância da medida  $V(x)$  segundo o modelo de dois componentes (Rocke e Lorenzato, 1995) é :

$$V(x) = \sigma_{\varepsilon}^2 + \mu^2 e^{\sigma_{\eta}^2} (e^{\sigma_{\eta}^2} - 1) \quad \text{Eq. 3-6}$$

onde  $x$  é o valor observado,  $\sigma_{\varepsilon}$  e  $\sigma_{\eta}$  são os parâmetros do modelo e  $\mu$  o valor verdadeiro da grandeza.

Considerando cada resultado analítico  $x_i$  e representando a incerteza prevista pelo modelo R&L por  $\delta_i$ , temos:

$$\delta_i = \sqrt{V(x_i)} = \sqrt{\sigma_{\varepsilon}^2 + x_i^2 e^{\sigma_{\eta}^2} (e^{\sigma_{\eta}^2} - 1)} \quad \text{Eq. 3-7}$$

A estimativa dos parâmetros  $\sigma_{\varepsilon}$  e  $\sigma_{\eta}$  pode ser feita por esta expressão, uma vez que são conhecidos os valores de  $\delta_i$  e  $x_i$  contidos no conjunto de resultados provenientes da aplicação do protocolo analítico. No caso de resultados analíticos gerados na execução de um PMA, as medidas  $x_i$  são os valores das concentrações e  $\delta_i$  as incertezas analíticas publicadas para cada concentração. Esta estimativa pode ser feita por ajuste da equação 3-7 aos resultados publicados das concentrações e respectivas incertezas. Este conjunto de dados é o resultado da aplicação de um mesmo protocolo analítico a uma mesma matriz, devendo portanto apresentar coerência em seu comportamento com relação à performance do laboratório. Ele não descreve fisicamente nenhuma amostra ou região, é uma amostra estatística.

Conhecendo os valores dos parâmetros do modelo de R&L a previsão da incerteza corrigida  $\delta_i'$ , pode ser feita pela equação:

$$\delta_i' = \sqrt{\hat{\sigma}_{\varepsilon}^2 + x_i^2 e^{\hat{\sigma}_{\eta}^2} (e^{\hat{\sigma}_{\eta}^2} - 1)} \quad \text{Eq. 3-8}$$

onde  $\hat{\sigma}_{\varepsilon}$  e  $\hat{\sigma}_{\eta}$  são as estimativas dos parâmetros e  $x_i$  as concentrações medidas.

O conjunto de dados a ser analisado estatisticamente para estimativa da média será:

$$(x_i; \delta_i') \quad \text{Eq. 3-9}$$

onde  $x_i$  é o valor da concentração publicado pelo laboratório e o valor  $\delta'_i$  é o incerteza corrigida, que será usada no cálculo do fator de ponderação para cada concentração segundo a equação 3-2.

A estimação da média representativa do conjunto de resultados analíticos será então feita pela equação:

$$\hat{\mu} = \bar{x} = \frac{\sum \left[ \frac{x_i}{\delta'_i{}^2} \right]}{\sum \left[ \frac{1}{\delta'_i{}^2} \right]} \quad \text{Eq. 3-10}$$

### 3.5 Os parâmetros $\sigma_\varepsilon$ e $\sigma_\eta$ do modelo de dois componentes

A eq.3.7, que representa o comportamento da incerteza analítica em função da concentração, mostra que quando esta concentração tende para zero a incerteza analítica tende para  $\sigma_\varepsilon$ .

$$\lim_{x_i \rightarrow 0} \delta_i = \lim_{x_i \rightarrow 0} \sqrt{\sigma_\varepsilon^2 + x_i^2 e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1)} = \sqrt{\sigma_\varepsilon^2} = \sigma_\varepsilon \quad \text{Eq 3-11}$$

Este parâmetro portanto, descreve a incerteza na medida da concentração quando esta se aproxima do valor zero. Esta é a condição imposta ao cálculo do desvio padrão quando ele é usado como estatística para determinar limites de detecção de métodos analíticos. Desta forma a estimativa deste parâmetro para um determinado protocolo analítico pode ser usada para a determinação de limites de detecção nas definições mais usadas. O parâmetro  $\sigma_\eta$ , para altos valores da concentração, é aproximadamente igual ao valor do desvio padrão relativo (Wilson et al., 2000) que é um indicador de precisão do método (USEPA, 2001b).

### 3.6 Avaliação da performance do laboratório

O conjunto dos resultados produzidos pode ser usado para avaliar o desempenho do laboratório em relação a todo o processo de medição descrito no protocolo analítico. A interpretação destes dados propicia uma ferramenta de controle de qualidade baseados na produção analítica do laboratório e não em medidas realizadas especificamente para este fim em padrões, pseudo-amostras ou brancos. O conjunto a ser

analisado é formado por todos resultados gerados pelo mesmo protocolo analítico, isto é, referem-se à mesma análise e à mesma matriz.

A visualização do comportamento da incerteza analítica em função da concentração ilustra a aderência da prática analítica ao modelo de medição especificado no protocolo analítico. Para um laboratório trabalhando sob condições de controle estatístico, a reprodutibilidade de seus resultados será visualizada por uma faixa bem determinada em um gráfico que mostre a incerteza analítica em função da concentração medida. Valores fora desta faixa representam resultados atípicos que devem ser investigados: podem representar uma situação de operação fora de controle estatístico, uma amostra apresentando interferências ou dificuldade específica no tratamento químico ou físico da amostra.

O conceito de repetibilidade (USEPA, 2001b) é aplicado aos resultados obtidos por um mesmo operador ao aplicar um protocolo analítico a diversas alíquotas de uma mesma amostra num período curto de tempo. É definida como a probabilidade destes resultados não diferirem entre si de uma determinada quantidade. A qualidade do ajuste, entre os resultados produzidos e o modelo de medição especificado no protocolo analítico, representa este conceito de repetibilidade aplicada ao laboratório como um todo: equipe, instrumentação e práticas analíticas. A aderência dos resultados à expressão matemática do modelo de medição especificado no protocolo analítico mostra boa repetibilidade, isto é, a capacidade do laboratório produzir os mesmos resultados para um mesmo procedimento analítico.

## 4. RESULTADOS

A metodologia proposta neste trabalho foi aplicada a resultados do Programa de Monitoração Ambiental do Centro de Desenvolvimento de Tecnologia Nuclear (CDTN) em operação desde 1986 (Sacramento et al, 1993). O conjunto de dados disponibilizado refere-se a resultados analíticos obtidos pela execução deste programa e produzidos por diversos protocolos analíticos relativos a diversos tipos de análises e matrizes.

A seleção dos casos a serem estudados foi baseada no número de resultados produzidos pela aplicação de cada protocolo analítico, para maior confiabilidade do ponto de vista estatístico, e que pudessem evidenciar as potencialidades do método analítico proposto. Cada caso analisado refere-se a um tipo de análise e matriz, significando resultados referentes a um protocolo analítico específico.

Os protocolos analíticos focalizados foram:

Caso 1: análise alfa total - matriz água - laboratório 1.

Caso 2: análise alfa total - matriz água - laboratório 2.

Caso 3: análise urânio 238 - matriz água.

Caso 4: análise alfa total - matriz aerossol.

Os casos 1 e 2, em que são focalizados protocolos similares e diferentes laboratórios executores, permitem o estudo do mesmo modelo de medição na mesma matriz em amplitudes diferentes de concentração e também observar o desempenho dos laboratórios para um mesmo protocolo. O caso 3 trata da mesma matriz dos casos 1 e 2 mas abordando outro tipo de análise. O caso 4 focaliza um protocolo com o mesmo princípio de medição dos casos 1 e 2 em outro tipo de matriz.

Em cada um dos casos focalizados, são apresentados os resultados da estimação dos parâmetros do modelo de dois componentes. As médias representativas de cada local amostrado são estimadas usando os fatores de ponderação calculados pela Eq 3.2 com as incertezas analíticas previstas pelo modelo de R&L e também com as incertezas analíticas publicadas, para efeito de comparação.

### 4.1 Análise exploratória

A análise exploratória, realizada para cada um destes casos, é apresentada no Apêndice A. É importante assinalar que os resultados apresentados nesta análise

exploratória não têm significado físico, são estimativas feitas para todo o conjunto de resultados de uma mesma análise para uma matriz específica, isto é, representam resultados obtidos pela aplicação de um mesmo protocolo analítico especificado pelo laboratório executor.

Os dados originais foram fornecidos em arquivos no formato de banco de dados do aplicativo Access (Microsoft, 2000) de modo que a preparação dos dados para processamento eletrônico foi realizada utilizando-se este aplicativo e a planilha eletrônica Excel (Microsoft, 2000). Os resultados analíticos foram originalmente apresentados em tabelas separadas, exceto os referentes às análises de alfa e beta total, tabulados em uma tabela única, por apresentarem alguns dados em comum. A estrutura típica destas tabelas contém cinco informações para cada resultado analítico:

- o número de identificação do boletim analítico, que possibilita o rastreamento da análise. Este documento contém informações sobre o laboratório executor, método analítico empregado, limite de detecção e notificação de ocorrências na amostragem ou na análise, nome do responsável e a data de emissão do laudo;
- a identificação da amostra, que é uma composição de códigos: os dois primeiros identificam a instalação a que se refere o programa, os três seguintes referem-se à matriz e os três últimos ao local onde foi coletada;
- a data da coleta;
- a concentração;
- o erro associado ao resultado analítico.

As informações sobre unidades de medida, técnicas analíticas e limites de detecção são publicadas nos boletins analíticos e, juntamente com outras informações gerais utilizadas neste estudo, foram recuperadas em relatórios de avaliação do programa (Sacramento et al., 1993; Peixoto e Pêgo, 1996).

Algumas tabelas apresentam também uma coluna extra para acomodar o sinal gráfico “<” que, nos boletins analíticos, aparecem associados a valores publicados como “menor que o limite de detecção”.

Em um tratamento preliminar, os resultados foram examinados para verificação de falta de informações e erros grosseiros, especialmente em relação a transcrição de dados. Este processo envolve rastreamento e confrontação com resultados originais publicados nos boletins analíticos implicando consultas a arquivos em papéis, especialmente em relação a resultados mais antigos, quando não havia automação em praticamente nenhuma etapa do processo analítico.

O tratamento estatístico dos dados e as representações gráficas requeridas foram realizados usando o pacote estatístico Statistica (Statsoft, 1998).

O conjunto de dados originais foi organizado em tabelas referentes a cada tipo de análise e matriz, isto é, referentes a um mesmo protocolo analítico. Antes de serem submetidos às técnicas tradicionais de análise exploratória, estes registros foram objeto de uma pesquisa minuciosa visando detectar a ocorrência de registros de datas iguais ou superiores ao ano 2000. No arquivo original foram reservados apenas dois dígitos para registrar o valor referente ao ano das datas de coleta. O processamento eletrônico identificou valores relativos a estas datas como referentes aos anos de 1900 e seguintes (bug do milênio). Os dados originais sofreram um tratamento específico para contornar este problema.

Neste estudo, além do objetivo primário de conhecer as características do conjunto de resultados, a análise exploratória foi realizada buscando encontrar o padrão de comportamento para a incerteza publicada em função da concentração medida. Isto significa reconhecer o tipo de expressão analítica do modelo de medição efetivamente praticado pelo laboratório. Gráficos e descrições estatísticas de subconjuntos de resultados foram usados nesta busca e o estudo de dados discrepantes foi realizado com este mesmo objetivo. Os dados considerados discrepantes referem-se a resultados que não se ajustam ao padrão de comportamento do grupo selecionado, considerando as variáveis concentração e incerteza, não significando que estes dados sejam discrepantes no estudo de outras grandezas consideradas.

#### **4.2 Estimação dos parâmetros do modelo de dois componentes**

Os parâmetros do modelo de dois componentes foram estimados por ajuste em regressão não linear. Neste ajuste, entre os resultados selecionados para descrever o desempenho do laboratório para um dado protocolo analítico e a expressão do erro prevista pelo modelo de Rocke & Lorenzato, foi usada a equação 3.7, que relaciona a incerteza analítica com a concentração medida. Foi usado o método de Newton, um processo iterativo onde a função a ser ajustada é avaliada em diferentes pontos, suas derivadas parciais de primeira e segunda ordem são estimadas e usadas no movimento dos parâmetros de iteração a iteração. A determinação dos valores iniciais para serem usados na estimação dos parâmetros do modelo de R&L foi feita, considerando uma relação linear entre erro e concentração, por meio de um ajuste linear. O valor da intersecção com o eixo y para os resultados de concentração próximos a zero foi considerada a estimativa inicial

do parâmetro  $\sigma_\epsilon$  e o valor inicial de  $\sigma_\eta$  foi tomado da inclinação da reta para as concentrações mais altas (Rocke & Lorenzato, 1995).

A qualidade do ajuste é verificada utilizando-se os parâmetros tradicionais: a soma dos quadrados dos resíduos SQR [(valor observado – valor previsto)<sup>2</sup>], dos indicadores R ( $1 \geq R \geq 0$ ) que quantifica o grau de relacionamento entre as duas grandezas (similar ao coeficiente de correlação de Pearson) e o percentual de variância explicada (%V). A variância explicada, calculada pela razão SQR/SQT (SQT = soma dos quadrados dos desvios totais) ilustra a contribuição à variância total relativa ao modelo.

Esses cálculos foram realizados usando o pacote de software Statistica (Statsoft, 1998), na opção para estimação não linear.

### 4.3 Caso 1: matriz água - análise alfa total - laboratório 1

O laboratório, identificado aqui como Laboratório 1, executou análises relativas aos pontos de água de superfície, subterrânea, chuva, potável e de um ponto de esgoto industrial. O protocolo analítico para esta análise prevê filtragem de uma alíquota de 250 ml, deposição em bandejas de aço inox e contagem em contador proporcional. A expressão para cálculo do erro analítico segue uma variação com a raiz quadrada da concentração baseada na estatística de Poisson (CDTN, 1994).

A análise exploratória do conjunto de dados produzidos por este laboratório, pela aplicação do protocolo analítico da análise de alfa total na matriz água, é apresentada com mais detalhes no Apêndice A. A TABELA 4-1 apresenta o resumo estatístico deste conjunto de dados.

TABELA 4-1 Estatística descritiva: análise alfa total, Lab 1, matriz água

Período de Coleta: Jan/86 a Dez/01								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	863	0,108	0,030	-0,6	25	0,80	25	697
Erro publicado (Bq/dm <sup>3</sup> )	863	0,069	0,040	0,002	6	0,064	17	368

Na etapa inicial da operação deste PMA, este laboratório realizou algumas análises introduzindo pequenas modificações no protocolo analítico. Os resultados destas análises, referenciadas aqui como investigativas, não se distanciaram da relação erro-concentração apresentada pelas análises planejadas, de modo que todos resultados foram considerados na determinação dos parâmetros característicos do laboratório para este

protocolo analítico. Na análise exploratória foram identificados os dados discrepantes em relação ao comportamento erro-concentração, sendo excluídos do conjunto dos resultados a serem usados na caracterização do laboratório. As estatísticas básicas que descrevem este subconjunto são apresentadas na TABELA 4-2.

TABELA 4-2 Estatística descritiva: matriz água, alfa total, Lab 1, - sem os dados discrepantes

Período de Coleta: Jan/86 a Dez/01

	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	857	0,106	0,030	-0,03	25	0,80	25	698
Erro publicado (Bq/dm <sup>3</sup> )	857	0,058	0,040	0,002	6	0,045	26	722

Os valores iniciais para  $\sigma_\varepsilon$  e  $\sigma_\eta$ , necessários ao processo de estimação dos parâmetros do modelo de Rocke & Lorenzato, foram determinados usando os gráficos erro x concentração. O primeiro é representado pela intersecção da reta ajustada com o eixo das ordenadas na região das concentrações próximas a zero e o segundo, pela inclinação da reta para a região das concentrações mais altas. Os gráficos usados nestas determinações são apresentados na FIGURA 4-1 e na FIGURA 4-2. Os valores encontrados foram  $\sigma_\varepsilon = 0,02 \text{ Bq/dm}^3$  e  $\sigma_\eta = 0,2$ .

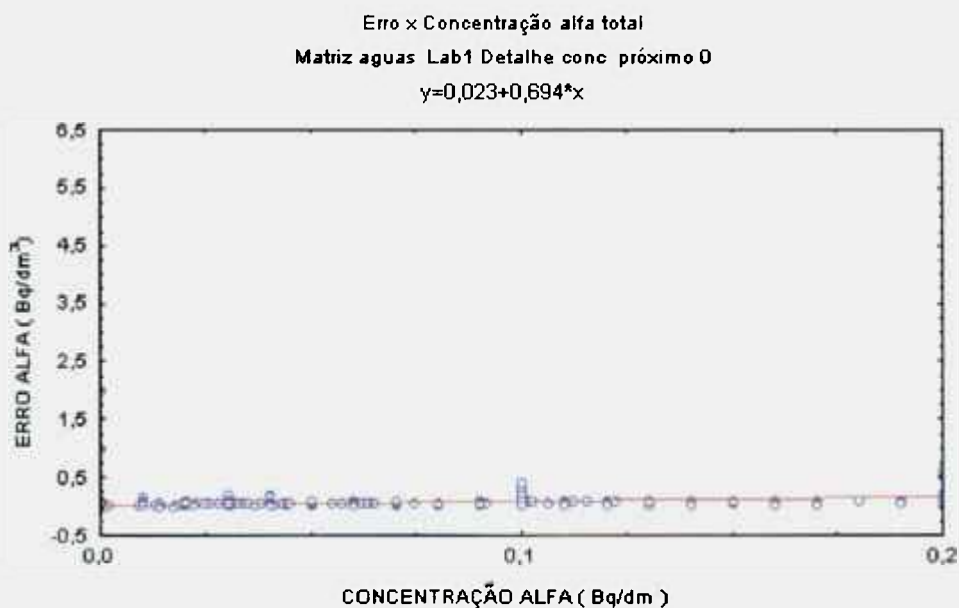


FIGURA 4-1 Determinação do valor inicial de  $\sigma_\varepsilon$ , alfa total, águas, lab1

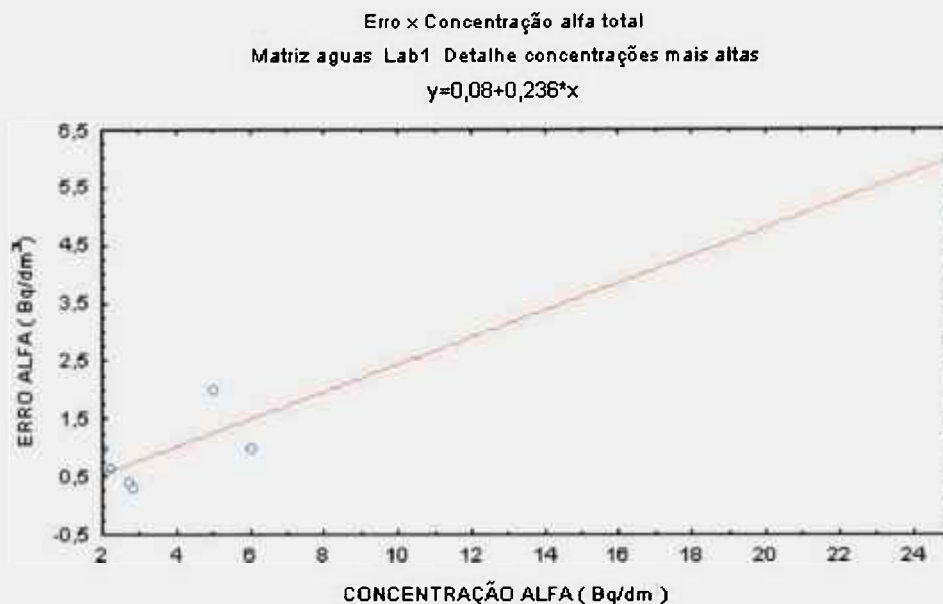


FIGURA 4-2 Determinação do valor inicial de  $\sigma_{\eta}$  alfa total, águas, lab1

Os valores dos parâmetros estimados para este caso, análise de alfa total para a matriz água, são apresentados na TABELA 4-3. Os parâmetros de qualidade do ajuste são apresentados nesta mesma tabela.

TABELA 4-3 Valor dos parâmetros e qualidade do ajuste: análise alfa total, água, Lab1

	N	$\sigma_{\varepsilon}$ (Bq/dm <sup>3</sup> )	$\sigma_{\eta}$	SQR	R	% V
Laboratório 1	857	0,0404	0,2261	1,43	0,98	96,3

A visualização do ajuste, por meio de um gráfico erro x concentração, é apresentada na FIGURA 4-3.

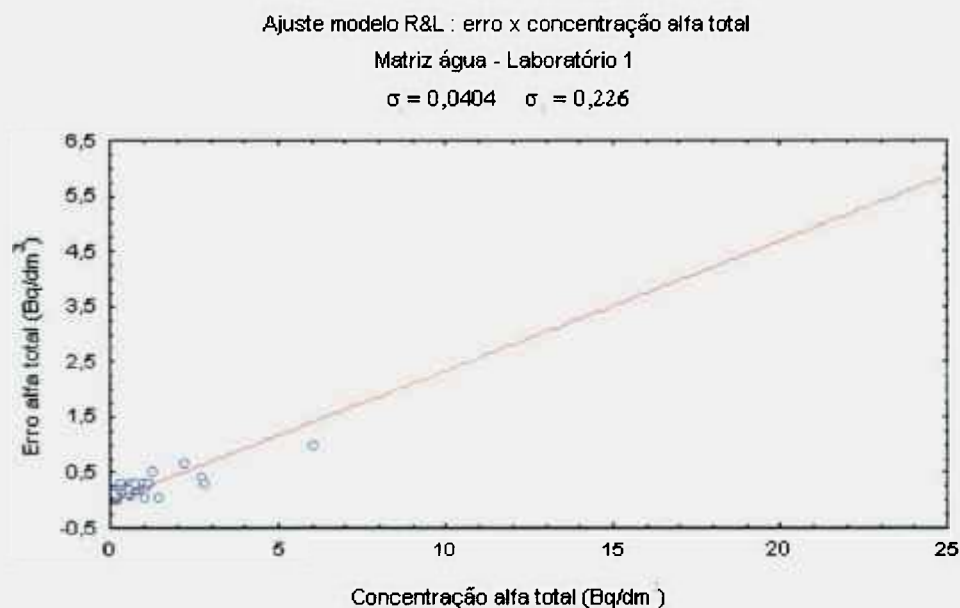


FIGURA 4-3 Visualização do ajuste: função erro x concentração, alfa total, água : Lab1

Os valores estimados para os parâmetros foram usados na equação 3.7 para gerar a os erros associados a cada valor da concentração segundo o modelo de dois componentes. A comparação entre o comportamento dos erros publicados e os previstos pelo modelo de R&L pode ser visualizada na FIGURA 4-4, onde estão representados todos os resultados analíticos e os erros previstos pelo modelo de Rocke & Lorenzato para cada concentração.

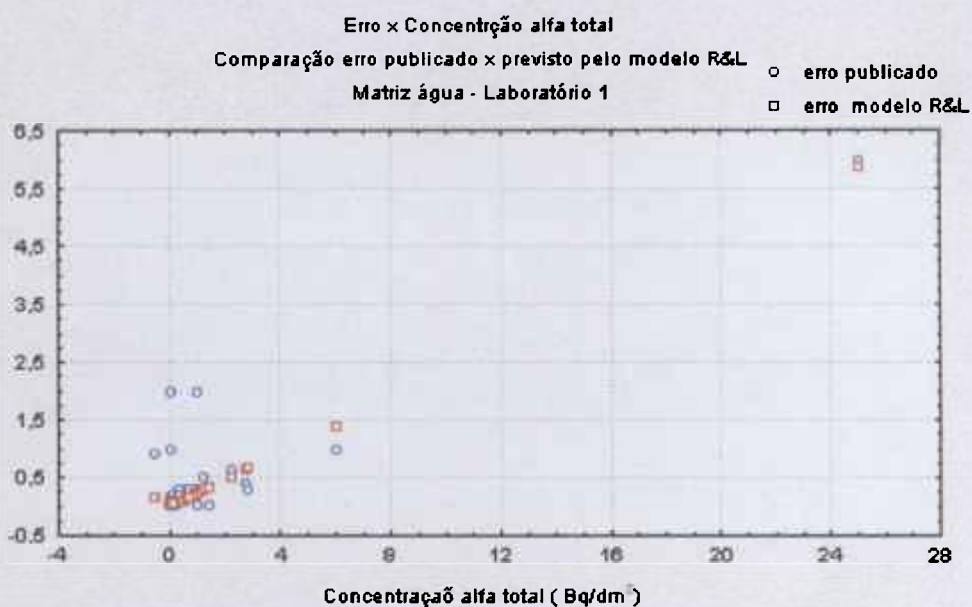


FIGURA 4-4 Comparação do erro publicado x erro previsto modelo R&L alfa total, água, Lab1

Os registros produzidos por este laboratório acham-se fortemente concentrados na região de baixas concentrações, apresentando um padrão bem definido. A região das concentrações próximas a zero é detalhada no gráfico apresentado na FIGURA 4-5, onde pode ser observada a linha paralela ao eixo horizontal indicando valor constante para o erro, o valor de  $\sigma_e$ . A região das concentrações intermediárias é detalhada na FIGURA 4-6, onde é mostrada a curva prevista pelo modelo de dois componentes e os resultados publicados.

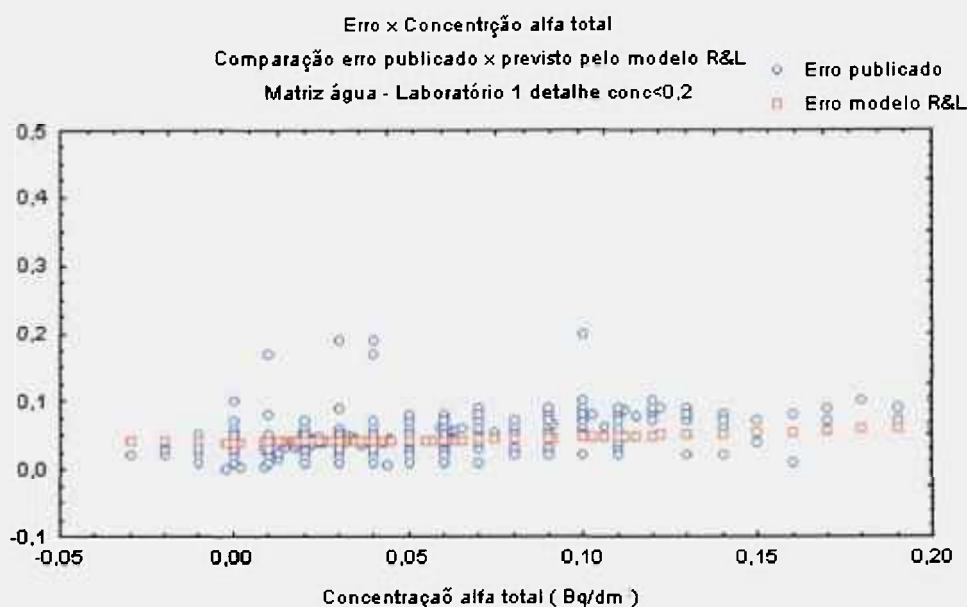


FIGURA 4-5 Visualização da faixa de concentração <0,2 : valor de  $\sigma_e$

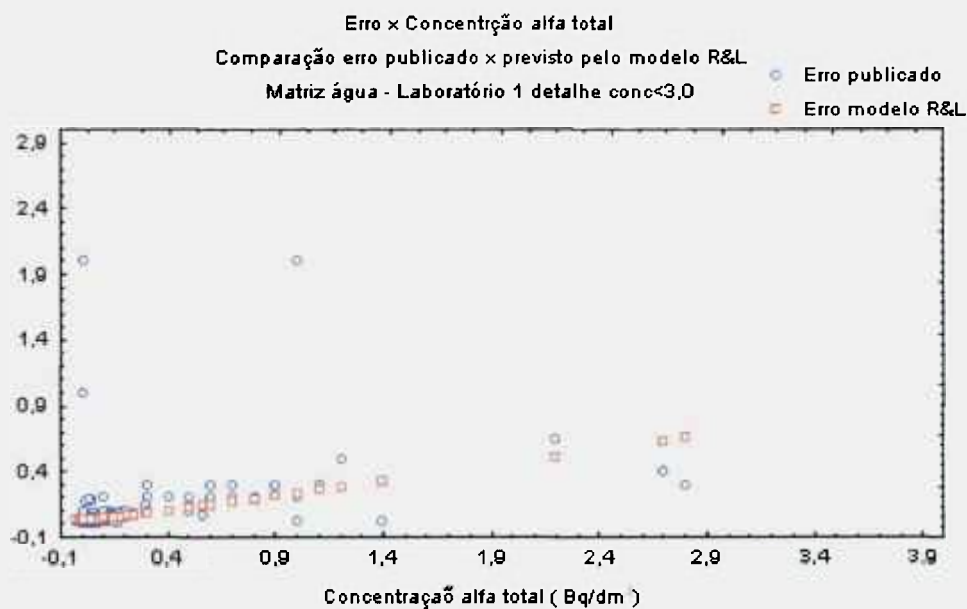


FIGURA 4-6 Detalhe da comparação do erro publicado x erro previsto pelo modelo R&L

Usando os erros previstos pelo modelo de dois componentes foram calculados os fatores de ponderação segundo a equação 3.2 e então, calculadas as médias para o conjunto de resultados referentes a cada amostra escolhida, segundo a equação 3.10. Foram estimadas também as médias ponderadas pelo inverso da variância das incertezas publicadas e as médias aritméticas, somente para efeito de comparação. As estimativas destas médias são apresentadas na TABELA 4-4 e também a razão entre o valor das duas médias calculadas por ponderação.

TABELA 4-4 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>), matriz água, Lab 1

Valores usados para os parâmetros: $\sigma_e = (0,0404 \pm 0,0015) \text{ Bq/dm}^3$ $\sigma_n = 0,226 \text{ l} \pm 0,0014$					
Amostra	N	Média simples	Média ponderada erro publicado	Média ponderada erro modelo R&L	R <sub>MP</sub> =Relação modelo/publicado
A06042	130	0,063	0,0228	0,0329	1,44
A06043	128	0,0459	0,0253	0,0301	1,19
Local 049	149	0,25	0,0145	0,0427	2,94
A08046	28	0,067	0,0370	0,0456	1,23
A08047	32	0,0486	0,0343	0,0428	1,25
A09048	76	0,043	0,0180	0,0283	1,57
A10001	49	0,046	0,0265	0,0333	1,26
A10002	43	0,0370	0,0261	0,0343	1,31
A10003	50	0,0364	0,0226	0,0310	1,37

#### 4.4 Caso 2: matriz água - análise alfa total - laboratório 2

O laboratório, identificado neste trabalho como Lab 2, realizou as análises relativas às águas de esgoto industrial nos pontos 037, 038, 039, 040 e 041. A amostragem nestes locais foi descontinuada sendo substituída por um único local que então passou a ser monitorado sob a codificação 021050.

O protocolo analítico adotado por este laboratório é similar ao utilizado pelo laboratório 1. As práticas analíticas e os procedimentos de calibração e controle adotados em cada laboratório é que acarretam diferenças no resultado produzido.

O resumo estatístico do conjunto de dados a ser usado na determinação dos parâmetros do modelo R&L é apresentado na TABELA 4-5.

TABELA 4-5 Estatística descritiva para resultados alfa total matriz água Lab 2

Período de Coleta: Jan/86 a Mar/02								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	250	3,18	0,20	-18	107	148	6,2	45
Erro publicado (Bq/dm <sup>3</sup> )	222	2,57	0,30	0,001	22	28	2,4	4,6

A análise exploratória deste conjunto de dados mostrou um aglomerado de registros com valores para o erro analítico muito distantes do padrão apresentado pela maioria dos resultados. Este comportamento, identificado no gráfico que mostra a evolução temporal do erro (FIGURA 14, Apêndice A), ocorreu no período de julho de 85 a setembro de 90. No gráfico erro x concentração, FIGURA 4-7, estes resultados aparecem como uma nuvem na região de maiores valores dos erros. Estes pontos diferem do padrão de comportamento em relação à concentração da maioria dos resultados e não se ajustam às

relações propostas pelos modelos de medição conhecidos. Por serem resultados atípicos, provavelmente em razão de alguma falha no controle estatístico do processo, foram excluídos do conjunto a ser usado na determinação dos parâmetros do modelo R&L, já que estes parâmetros irão descrever o desempenho do laboratório.

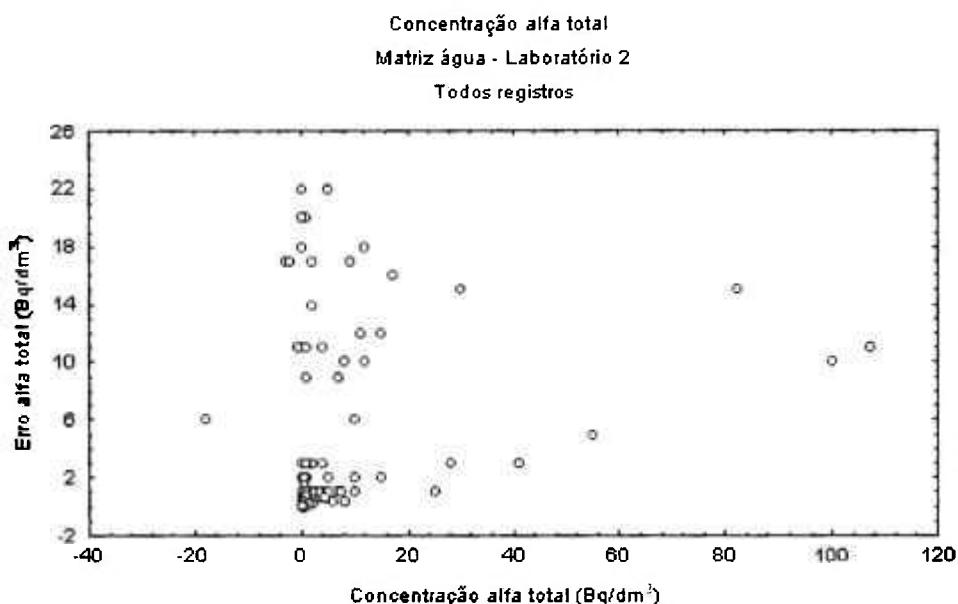


FIGURA 4-7 Erro em função da concentração alfa total, água, Lab2

As estatísticas descritivas básicas, obtidas na análise exploratória, são apresentadas na TABELA 4-6 .

TABELA 4-6 Estatística descritiva alfa total matriz água, Lab 2 sem nuvem, sem discrepantes

Período de Coleta: Jan/86 a Mar/02

Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	158	2,97	0,15	-0,02	107	172	6,5	46
Erro publicado (Bq/dm <sup>3</sup> )	158	0,43	0,10	0,001	11	1,6	6,7	50

Os valores iniciais para  $\sigma_\epsilon$  e  $\sigma_\eta$  foram calculados usando os gráficos erro x concentração apresentados na FIGURA 4-8 e na FIGURA 4-9. Os valores encontrados,  $\sigma_\epsilon = 0,02$  Bq/dm<sup>3</sup> e  $\sigma_\eta = 0,1$  foram usados como estimativas iniciais no processo iterativo de determinação dos parâmetros do modelo R&L.

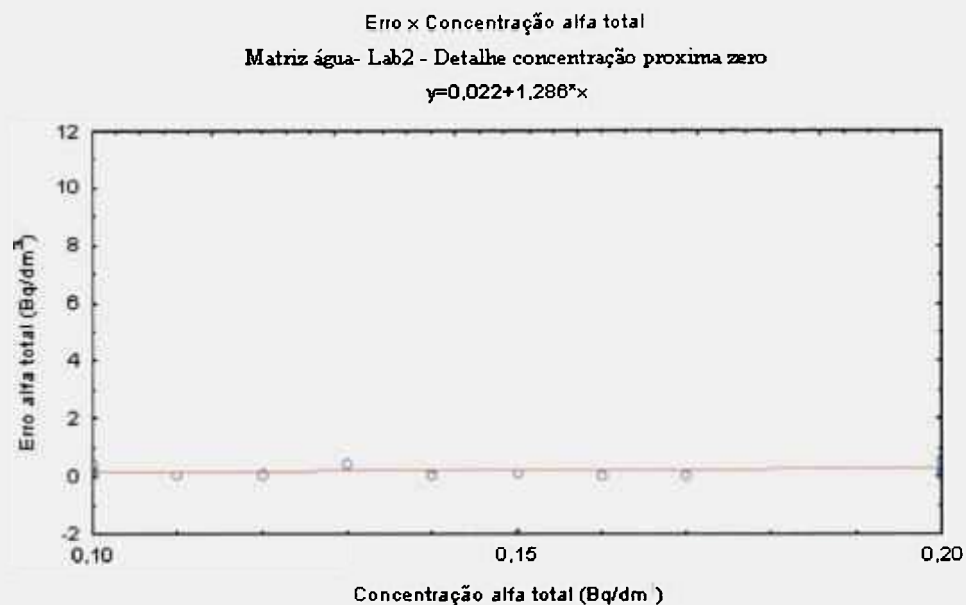


FIGURA 4-8 Determinação do valor inicial de  $\sigma_\epsilon$ : alfa total, água, Lab2

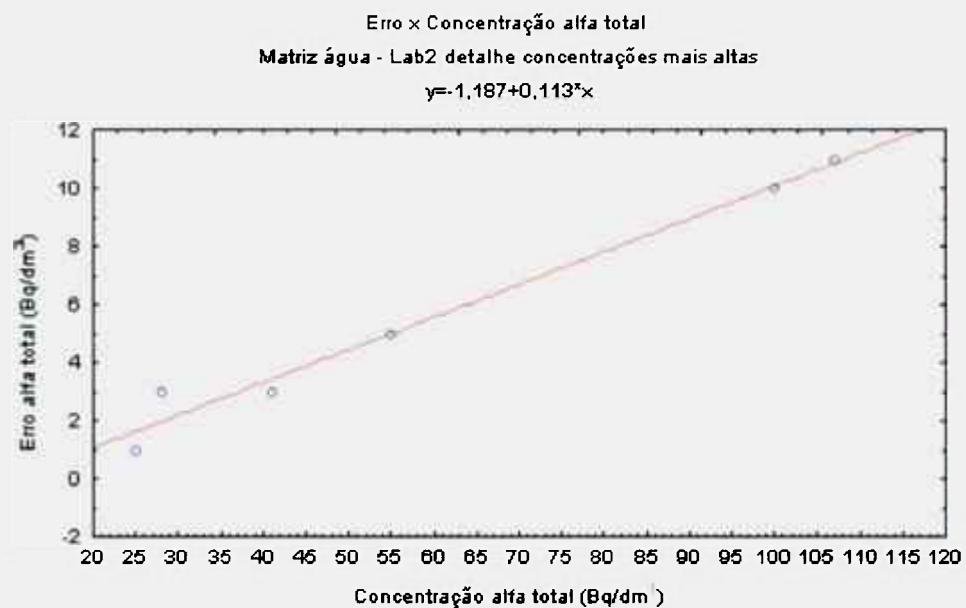


FIGURA 4-9 Determinação do valor inicial de  $\sigma_\eta$ : alfa total, água, Lab2

Os valores das estimativas dos parâmetros do modelo de dois componentes para o laboratório 2, análise alfa total na matriz água, são apresentados na TABELA 4-7, juntamente com os indicadores de qualidade do ajuste.

TABELA 4-7 Valor dos parâmetros e qualidade do ajuste: análise alfa total, água, Lab2

	obs	N	$\sigma_\epsilon$ (Bq/dm <sup>3</sup> )	$\sigma_\eta$	SQR	R	% V
LAB 2	Sem aglomerado	158	0,1712	0,0972	9,903	0,98	96.2

A visualização do ajuste é apresentada na FIGURA 4-10.

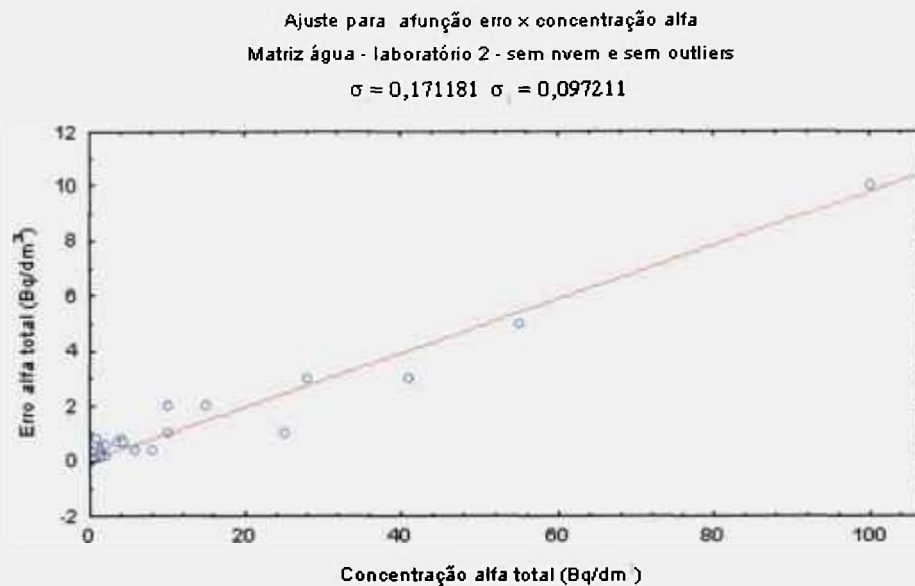


FIGURA 4-10 Visualização do ajuste: função erro x concentração, alfa total, água ; Lab2

Os erros previstos pelo modelo de R&L foram gerados usando, na equação 3-8, os valores estimados para os parâmetros. Na FIGURA 4-11 é apresentado o gráfico do erro em função da concentração mostrando todos resultados publicados e a curva prevista pelo modelo de dois componentes.

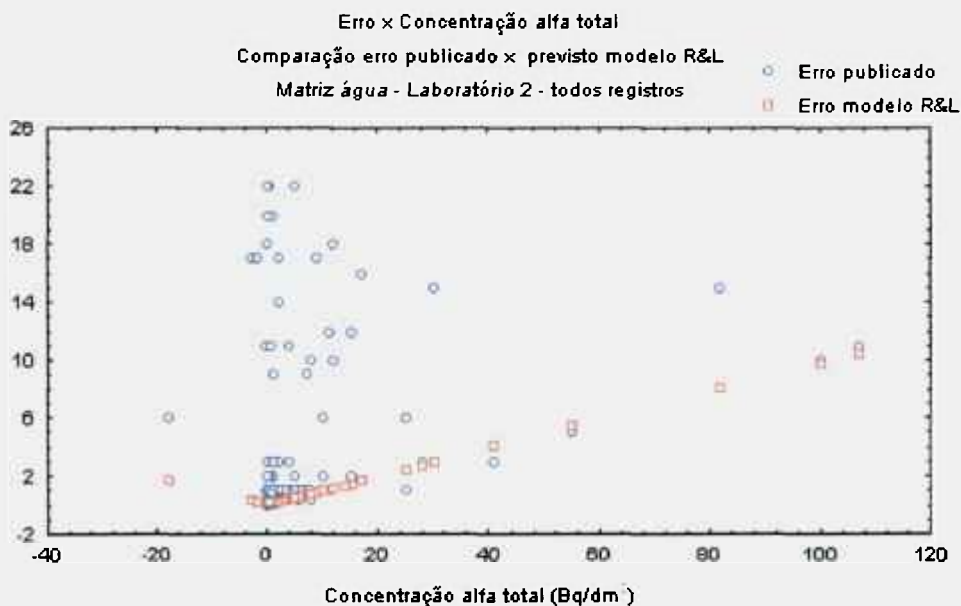


FIGURA 4-11 Comparação erro publicado x erro previsto modelo R&L, alfa total, água, Lab2

Usando os fatores de ponderação calculados pela equação 3.2, as estimativas das médias representativas dos resultados foram estimadas utilizando a equação 3.10 para cada amostra analisada. Os valores obtidos são apresentados na TABELA 4-8. O grau de censuramento apresentado nestes conjuntos de resultados e o número total de registros também são apresentados nesta tabela. Neste cálculo foram usados somente os resultados

publicados como acima do limite de detecção, o número destes resultados é representado por N nesta tabela.

TABELA 4-8 Estimativas das médias: concentrações alfa total (Bq/dm<sup>3</sup>), água – Lab 2 s/ ND

Valores usados para os parâmetros:  $\sigma_e = (0,17 \pm 0,02)$  Bq/dm<sup>3</sup>  $\sigma_n = 0,097 \pm 0,002$

Amostra	NT	N	Grau cen (%)	Média simples (N)	Média ponderada(N) erro publicado	Média ponderada(N) Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A21037		37	2,6	4,1	0,0485	0,348	7,18
A21038	29	23	20,7	11,8	1,20	1,24	1,03
A21040	47	35	25,5	-0,07	0,0096	0,170	17,7
A21041	26	23	11,5	2,6	0,0892	0,498	5,59
A21050	80	80	-	0,277	0,00562	0,195	34,7

Para observar o efeito do censuramento, as médias foram calculadas novamente, usando o método de substituição simples. Os resultados de concentração publicados como censurados foram substituídos pelo valor do limite de detecção. Os resultados destas estimativas são apresentados na TABELA 4-9 onde o número total de registros é especificado na coluna identificada como NT.

TABELA 4-9 Estimativas das médias: concentrações de alfa total (Bq/dm<sup>3</sup>), água – Lab 2

Valores usados para os parâmetros:  $\sigma_e = (0,17 \pm 0,02)$  Bq/dm<sup>3</sup>  $\sigma_n = 0,097 \pm 0,002$

Amostra	NT	N	Grau cen (%)	Média simples NT	Média ponderada(N) erro publicado	Média ponderada(NT) Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A21037	38	37	2,6	4,0	0,0485	0,367	7,57
A21038	29	23	20,7	11,1	1,20	1,04	0,87
A21040	47	35	25,5	-0,28	0,0096	0,221	23,0
A21041	26	23	11,5	2,4	0,0892	0,453	5,08
A21050	80	80	-	0,277	0,00562	0,195	34,7

Nas duas tabelas anteriores, o número de resultados que compõem a média ponderada pelo inverso do quadrado do erro publicado é o mesmo, pois os valores codificados como abaixo do limite de detecção são apresentados pelo laboratório sem registro para o valor do erro. Desta forma a relação R<sub>MP</sub> apresentada na TABELA 4-9 é apenas ilustrativa, já que o número de registros que compõe cada média é diferente: a média ponderada segundo o modelo R&L inclui todos os registros e pondera os valores publicados como abaixo do limite de detecção pelo inverso do quadrado do erro previsto por este modelo.

#### 4.5 Caso 3: matriz água - análise urânio 238

A análise de Urânio 238 na matriz água foi, durante todo o período em estudo, realizada pelo método de nêutrons retardados, por um mesmo laboratório. Os resultados são acompanhados de um valor relativo à incerteza analítica, exceto para aqueles valores publicados como “menor que o limite de detecção” que são publicados precedidos do sinal “<” e sem especificação da incerteza.

O resumo estatístico, que descreve este conjunto de resultados, é apresentado na TABELA 4-10 e os gráficos da evolução temporal da concentração e do erro publicado são apresentados no apêndice A.

TABELA 4-10 Estatística descritiva: matriz água, Urânio 238 - todos registros

Período de Coleta: Abril/86 a Abril/02								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	445	0,523	0,013	-8,000	14,00	7,9	3,1	12
Erro publicado (Bq/dm <sup>3</sup> )	384	0,267	0,010	0,001	8,00	1,1	5,2	28

No período inicial até abril de 92, este conjunto de resultados apresenta uma relação entre a concentração e o erro que se ajusta ao modelo de erro relativo constante (igual a 1) e também uma grande ocorrência de resultados negativos com valores absolutos altos. Todos os resultados censurados, publicados para este conjunto, pertencem a este mesmo período. Este comportamento, diferente do padrão apresentado pelos resultados produzido a partir desta data, foi considerado atípico e os resultados referentes a este período foram excluídos do conjunto a ser usado para descrever o desempenho do laboratório em relação a este protocolo analítico. O modelo de erro relativo constante foi utilizado na faixa de baixas concentrações (<0,2), por isso os registros referentes a este modelo também foram excluídos do subconjunto usado para estimar os parâmetros do modelo de dois componentes.

O resumo estatístico do conjunto de resultados a serem usados na estimação dos parâmetros do modelo dos dois componentes é apresentado na TABELA 4-11.

TABELA 4-11 Estatística descritiva: matriz água, Urânio 238 conjunto estudado

Período de Coleta: Abril/92 a Abril/02								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	117	0,148	0,030	-0,030	3,00	0,16	4,7	27
Erro publicado (Bq/dm <sup>3</sup> )	117	0,027	0,010	0,001	0,60	0,0046	6,0	45

Os valores iniciais para  $\sigma_\epsilon$  e  $\sigma_\eta$  foram calculados usando os gráficos erro x concentração apresentados na FIGURA 4-12 e na FIGURA 4-13. Os valores encontrados,  $\sigma_\epsilon = 0,004 \text{ Bq/dm}^3$  e  $\sigma_\eta = 0,2$  foram usados como estimativas iniciais no processo iterativo de determinação dos parâmetros do modelo R&L.

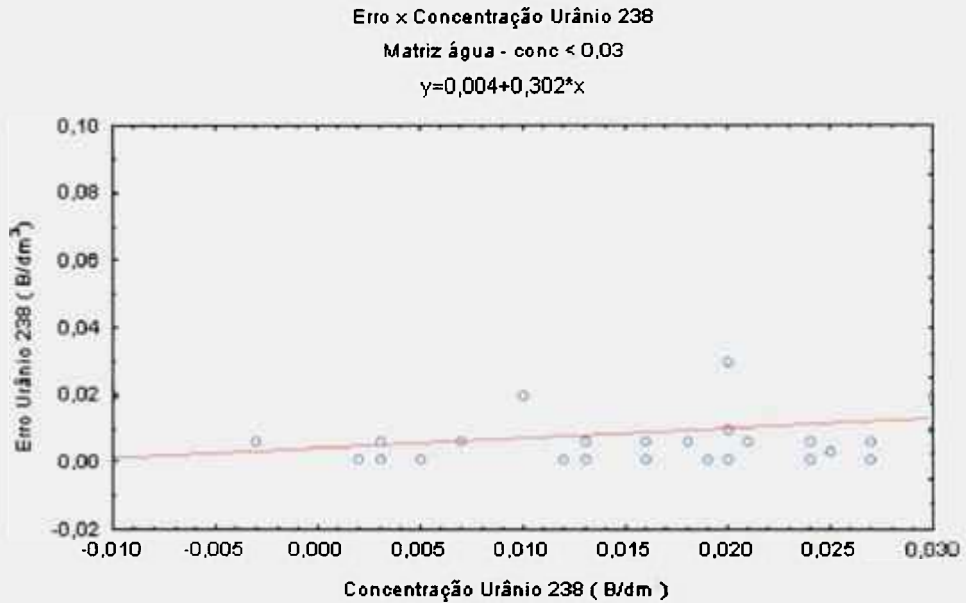


FIGURA 4-12 Determinação do valor inicial de  $\sigma_\epsilon$  urânio 238, águas

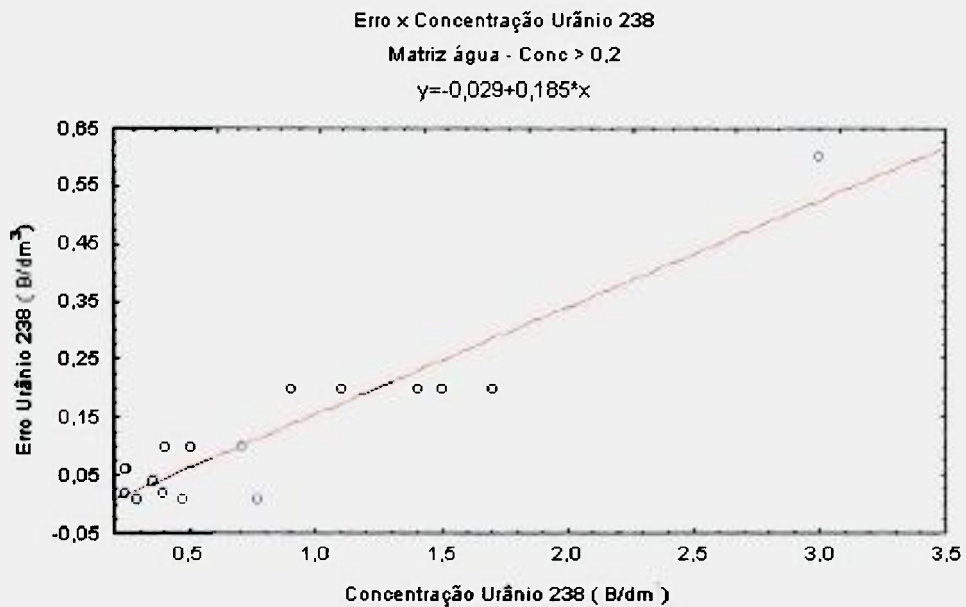


FIGURA 4-13 Determinação do valor inicial de  $\sigma_\eta$  urânio 238, águas

Os resultados obtidos para os parâmetros do modelo de R&L referentes aos resultados da análise de urânio 238 na matriz água, são apresentados na TABELA 4-12. Os indicadores de qualidade do ajuste são mostrados nesta mesma tabela.

TABELA 4-12 Estimativa dos parâmetros e qualidade do ajuste: urânio238, água

Urânio 238	N	$\sigma_c$ (Bq/dm <sup>3</sup> )	$\sigma_n$	SQR	R	% V
Coleta após jan/92	117	0,0076	0,1612	0,0556	0,947	89,6%

A visualização do ajuste é apresentada na FIGURA 4-14.

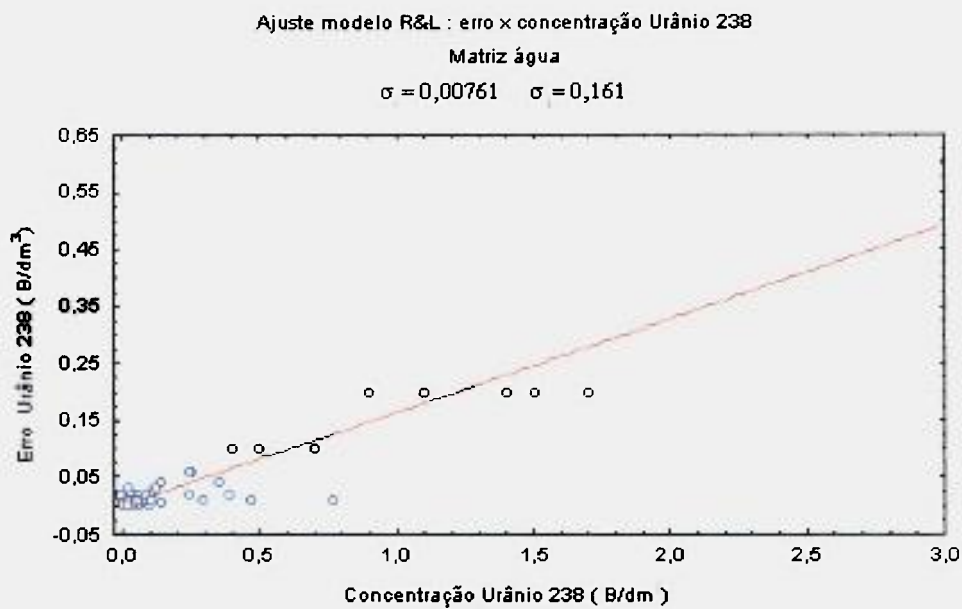


FIGURA 4-14 Visualização do ajuste: função erro x concentração, urânio 238 , matriz água

Usando os parâmetros estimados para o modelo de dois componentes, foram gerados os valores para o erro associado a cada valor da concentração medida usando a equação do modelo de R&L. A comparação entre os erros publicados e aqueles previstos pelo modelo pode ser visualizada na FIGURA 4-15.

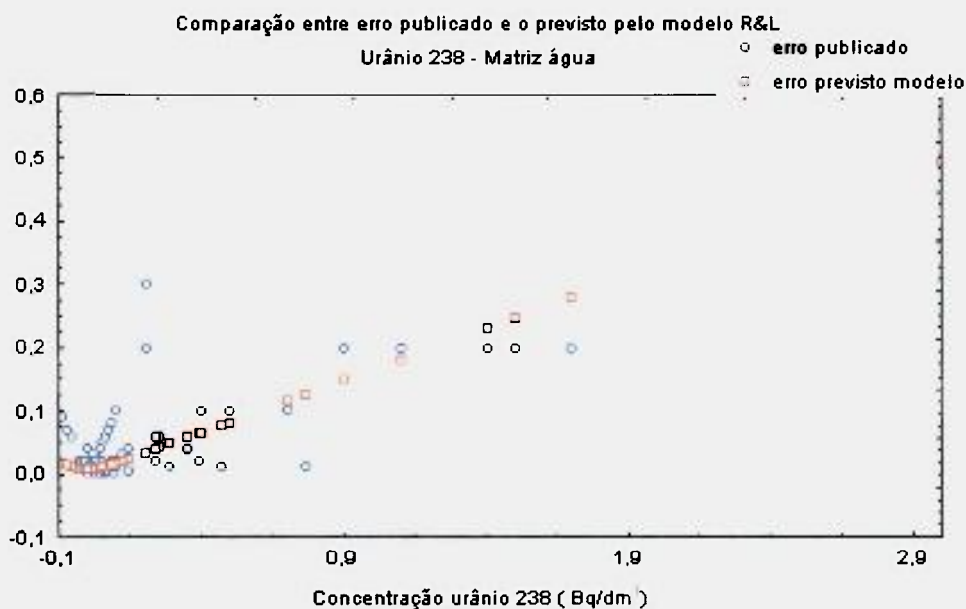


FIGURA 4-15 Comparação erro publicado x erro previsto modelo R&amp;L, urânio 238, água

Usando os valores dos erros previstos pelo modelo de R&L e os valores publicados, foi calculado o fator de ponderação para cada concentração medida segundo a equação 3.2. Foram então estimadas as médias representativas de cada local amostrado. Estas estimativas foram feitas para as amostras que apresentam número significativo de registros. Na TABELA 4-13 abaixo são apresentadas as médias estimadas usando os resultados publicados como valores maiores que o limite de detecção, isto é, foram excluídos os valores considerados abaixo do limite de detecção pelo laboratório executor.

TABELA 4-13 Estimativas das médias: concentrações de urânio 238 (Bq/dm<sup>3</sup>), água (>LD)

Valores usados para os parâmetros:  $\sigma_\epsilon = (0,0076 \pm 0,0025) \text{ Bq/dm}^3$   $\sigma_\eta = 0,1612 \pm 0,0046$

Amostra	NT	N	Grau cen (%)	Média simples (N)	Média ponderada(N) erro publicado	Média ponderada(N) Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A060042	46	45	2,2	0,045	0,017	0,0166	0,98
A060043	47	46	2,1	0,041	0,0037	0,0126	3,38
A060049	42	42	0,0	0,0340	0,0039	0,0094	2,44
A080046	31	30	3,2	0,0198	0,0076	0,0106	1,38
A080047	38	37	2,6	0,0198	0,0056	0,0086	1,54
A210050	37	37	-	0,224	0,0109	0,0193	1,77

Para observar o efeito do censuramento, as médias foram calculadas novamente, usando o método de substituição simples. Os valores de concentração publicados como censurados foram substituídos pelo valor do limite de detecção. Os resultados são apresentados na TABELA 4-14 onde o número total de registros é especificado na coluna identificada como NT.

TABELA 4-14 Estimativas das médias: concentrações de urânio 238 (Bq/dm<sup>3</sup>), água (inclui LD)

Valores usados para os parâmetros:  $\sigma_\epsilon = (0,0076 \pm 0,0025) \text{ Bq/dm}^3$   $\sigma_\eta = 0,1612 \pm 0,0046$

Amostra	NT	N	Grau cen (%)	Média simples (N)	Média ponderada(N) erro publicado	Média ponderada(N) Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A060042	46		2,2	0,044	0,0170	0,0165	0,97
A060043	47		2,1	0,041	0,0037	0,0130	3,48
A080046	31		3,2	0,0196	0,0076	0,0106	1,39
A080047	38		2,6	0,0196	0,0056	0,0087	1,56

#### 4.6 Caso 4: matriz aerossol - análise alfa total

As amostras da matriz aerossol foram coletadas em filtros de fibra de vidro, com periodicidade mensal, em três locais. O local 001 situa-se na direção de menor frequência de ventos e os locais codificados 002 e 003, na direção de maiores frequências de ventos. O protocolo analítico usado para a análise de alfa total e beta total prevê a

retirada de uma alíquota do filtro amostrado, sua transferência direta para a bandeja de contagem feita em inox e contagem em detector proporcional. As estimativas para as estatísticas descritivas do conjunto de resultados para este protocolo analítico são apresentadas na TABELA 4-15.

TABELA 4-15 Estatística descritiva: matriz aerossol, alfa total- todos registros.

Período de Coleta: Jan/86 a Dez/01								
Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (mBq/m <sup>3</sup> )	355	0,476	0,300	0,000	5,000	0,408	3,93	20,25
Erro publicado (mBq/m <sup>3</sup> )	355	0,199	0,100	0,001	5,633	0,241	7,39	64,80

A análise exploratória deste conjunto de resultados apontou a utilização, para a mesma faixa de concentração, de três modelos de medição: erro constante (para o valor 0, 1, 0,2 e 0,3); erro diretamente proporcional à concentração (com intersecção em 0,0) e erro com proporcionalidade à concentração numa forma que sugere a estatística de Poisson. Os gráficos que ilustram o estudo deste comportamento são apresentados no Apêndice A e uma ilustração dos três modelos é apresentado na FIGURA 4-16.

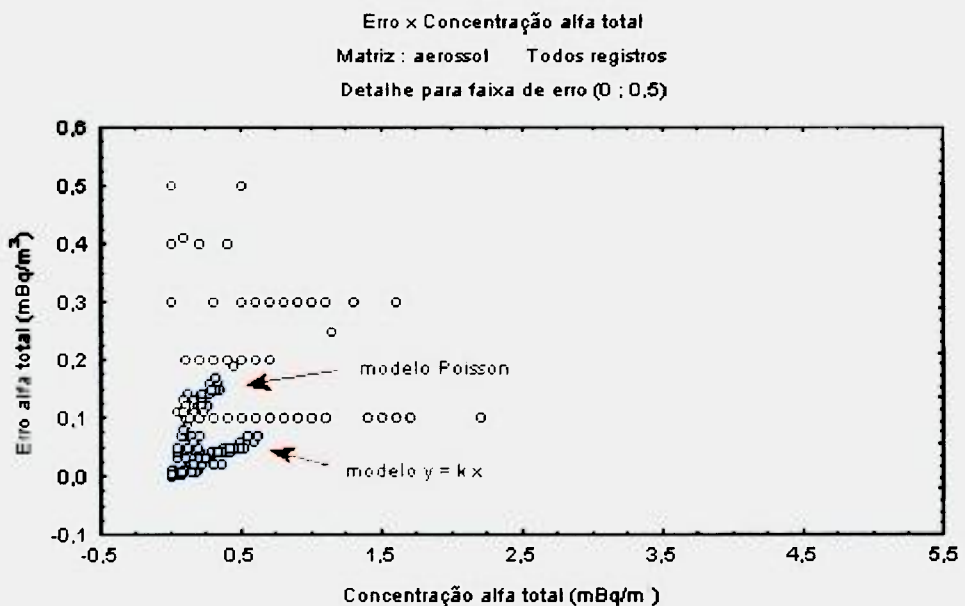


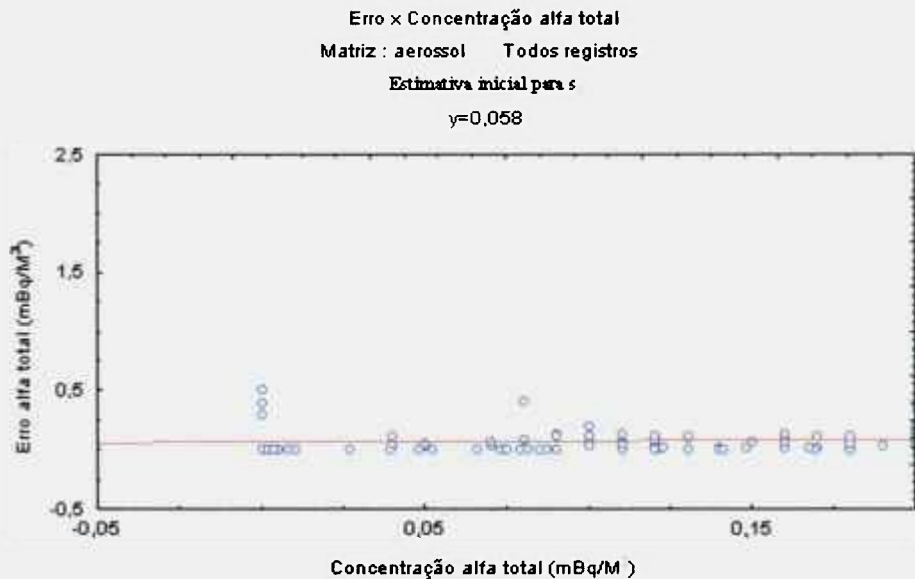
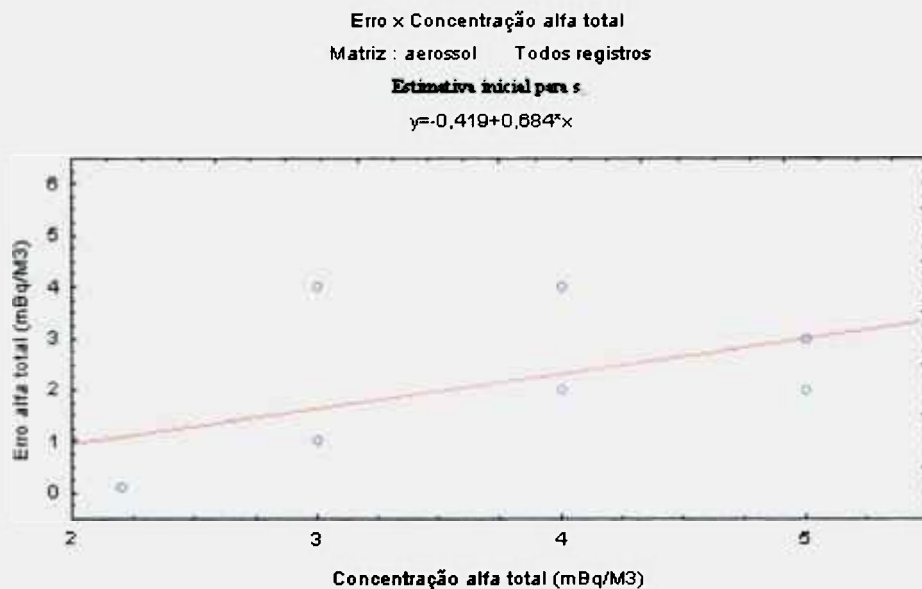
FIGURA 4-16 Detalhe : erro alfa em função da concentração, alfa total, aerossol

Para a caracterização do desempenho do laboratório em relação a este protocolo analítico foram utilizados somente os resultados produzidos segundo o modelo baseado na estatística de Poisson. As estatísticas descritivas básicas para este conjunto de resultados são apresentadas na TABELA 4-16.

TABELA 4-16 Estatística descritiva: matriz aerossol, alfa total - conjunto estudado

Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração (mBq/m <sup>3</sup> )	108	0,404	0,205	0,002	3,00	0,353	0,057
Erro publicado (mBq/m <sup>3</sup> )	108	0,182	0,120	0,001	1,00	0,052	0,022

Os valores iniciais para  $\sigma_\epsilon$  e  $\sigma_\eta$  foram calculados usando os gráficos erro x concentração apresentados na FIGURA 4-17 e na FIGURA 4-18. Os valores encontrados,  $\sigma_\epsilon = 0,06$  mBq/m<sup>3</sup> e  $\sigma_\eta = 0,7$  foram usados como estimativas iniciais no processo iterativo de determinação dos parâmetros do modelo R&L.

FIGURA 4-17 Determinação do valor inicial para  $\sigma_\epsilon$  alfa total, aerossolFIGURA 4-18 Determinação do valor inicial para  $\sigma_\eta$  alfa total, aerossol

As estimativas dos parâmetros resultantes do ajuste pelo método dos mínimos quadrados e os indicadores de qualidade do ajuste são apresentados na TABELA 4-17.

TABELA 4-17 Estimativa dos parâmetros e qualidade do ajuste: alfa total, aerossol

	N	$\sigma_c$ (mBq/m <sup>3</sup> )	$\sigma_n$	SQR	R	% V
	108	0,065	0,348	0,592	0,95	89

A ilustração do ajuste realizado entre o conjunto de pontos selecionado para descrever o desempenho do laboratório para o protocolo analítico referente à análise de alfa total na matriz aerossol e a expressão do modelo de dois componentes é apresentada na FIGURA 4-19.

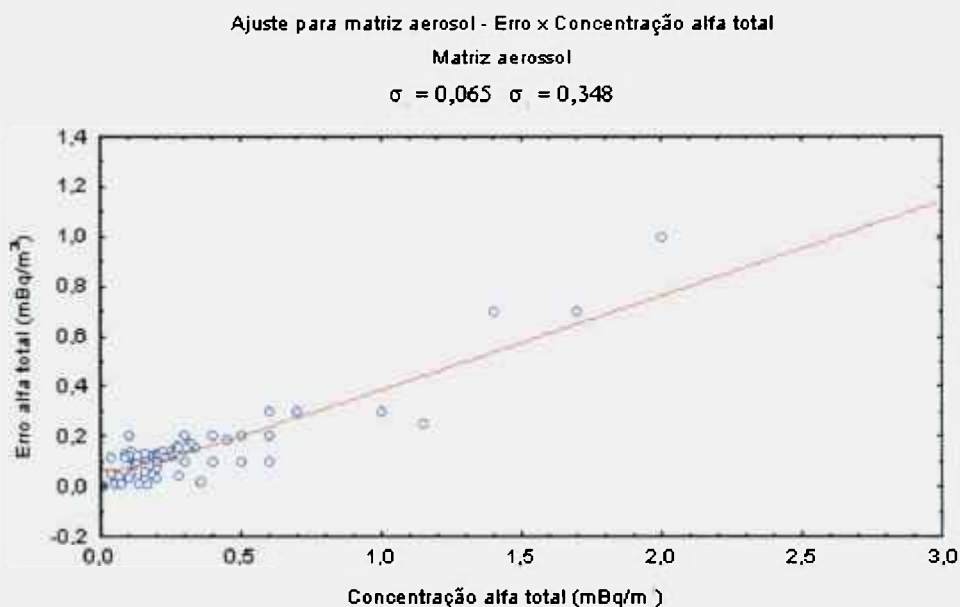


FIGURA 4-19 Visualização do ajuste: função erro x concentração, alfa total, aerossol

Usando a expressão do modelo de R&L foram gerados os erros associados a cada concentração para os valores dos parâmetros estimados. Estes valores foram usados no cálculo do fator de ponderação para estimativa da média do conjunto de resultados referentes à análise alfa total na matriz aerossol para cada amostra. Foram também estimadas a média aritmética e a média ponderada pelo inverso do quadrado dos erros publicados. A comparação do comportamento do erro publicado e do erro previsto pelo modelo é apresentada na FIGURA 4-20.

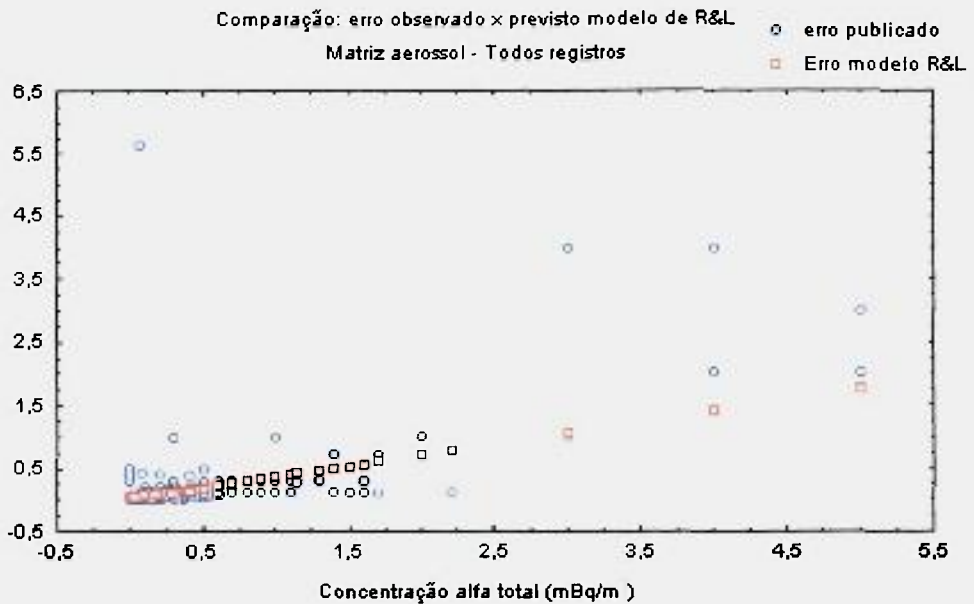


FIGURA 4-20 Comparação erro publicado x erro previsto modelo R&L, alfa total, aerossol

Os valores estimados para as médias representativas de cada local são apresentados na TABELA 4-18.

TABELA 4-18 Estimativas das médias: concentrações de alfa total ( $\text{mBq/m}^3$ ), matriz aerossol

Valores usados para os parâmetros  $\sigma_\varepsilon = (0,065 \pm 0,012) \text{mBq/m}^3$   $\sigma_\eta = 0,348 \pm 0,008$

Amostra	N	Média simples	Média ponderada pelo erro publicado	Média ponderada erro modelo R&L	$R_{MP}$ =Relação modelo/publicado
001	97	0,478	0,00738	0,163	22,1
002	123	0,499	0,01033	0,157	15,2
003	135	0,454	0,02331	0,142	6,1

## 5. DISCUSSÃO

### 5.1 Modelo de Roche & Lorenzato x modelo baseado em Poisson

Muitos métodos analíticos que envolvem medidas radiométricas usam a estatística de Poisson para avaliar a incerteza associada ao resultado da concentração medida. Os protocolos para a análise alfa total, referente a três dos casos focalizados neste trabalho, estão neste caso. A comparação entre este modelo e o de dois componentes é apresentada para fundamentar a discussão destes casos.

A relação entre o erro e a concentração baseada na estatística de Poisson resulta numa expressão regida pela função raiz quadrada. A forma desta curva mostra um comportamento diferente da função erro x concentração prevista pelo modelo de R&L: na faixa próximo a zero as duas curvas se cruzam; na faixa intermediária, a curva do modelo de R&L situa-se acima e, nas regiões das concentrações mais altas, a situação se inverte. Uma simulação entre o comportamento previsto pelos dois modelos é ilustrada na FIGURA 5-1.

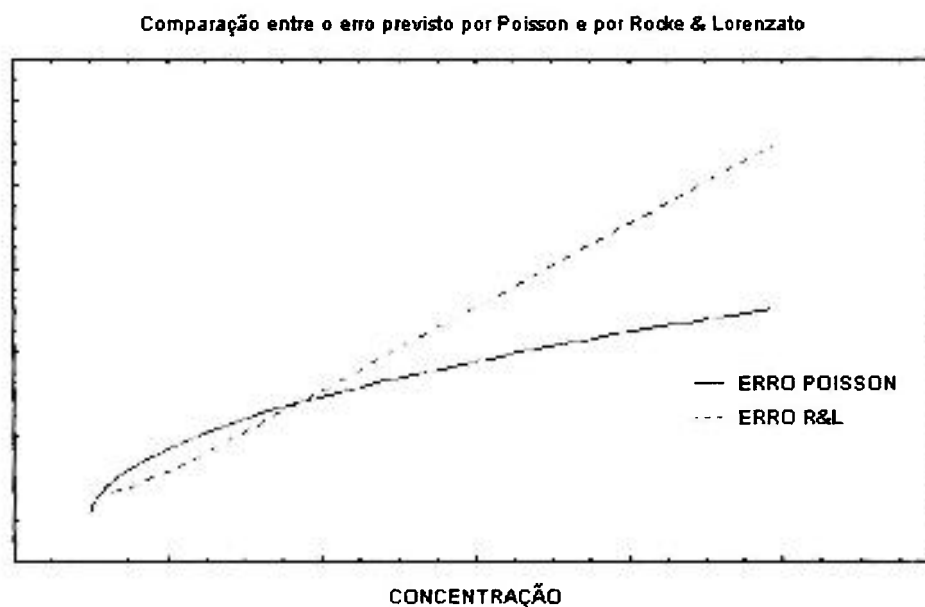


FIGURA 5-1 Comportamento do erro previsto pelo modelo de Poisson e pelo R&L

Especialmente na faixa das concentrações muito próximas ao limite de detecção este modelo não descreve bem o comportamento observado para o erro analítico. Nesta região o método não consegue distinguir pequenas diferenças entre concentrações

distintas e o desvio padrão tende a se tornar constante, entretanto, o modelo baseado em Poisson prevê incertezas crescentes nesta região. Para concentrações mais altas, este modelo prevê valores tendendo para um determinado valor, o que não acontece na realidade.

## 5.2 Caso 1: Análise alfa total, matriz água, lab1

Este conjunto de dados apresenta cerca de 98% dos resultados publicados na faixa de concentração menor que 3 Bq/dm<sup>3</sup> e erro menor que 0,3 Bq/dm<sup>3</sup>: é um conjunto de dados típico de resultados de análises ambientais de atividade baixa.

A análise exploratória deste conjunto de resultados mostrou que os resultados produzidos, durante todo o período estudado, apresentaram grande aderência ao modelo de medição que foi especificado no protocolo analítico. Este fato pode ser ilustrado no gráfico erro x concentração (FIGURA 5-2) que mostra os valores das concentrações menores que 3 Bq/dm<sup>3</sup> e erro menor que 0,5 Bq/dm<sup>3</sup>. Os pontos mostram-se concentrados numa faixa bem definida de valores, podendo ser observado um comportamento próximo à forma da função raiz quadrada, característica do modelo baseado na estatística de Poisson. O comportamento do erro relativo, apresentado na FIGURA 5-3 confirma esta observação: na região das concentrações próxima a zero os erros tendem rapidamente para valores mais altos e à medida que a concentração aumenta, tendem a um valor constante.

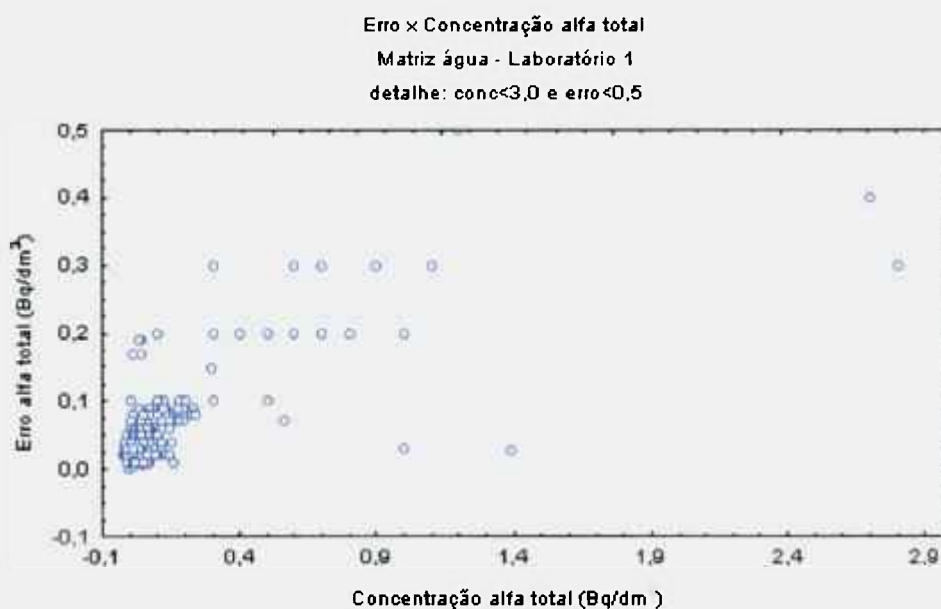


FIGURA 5-2 Aderência ao modelo: erro x concentração – concentração <3,0

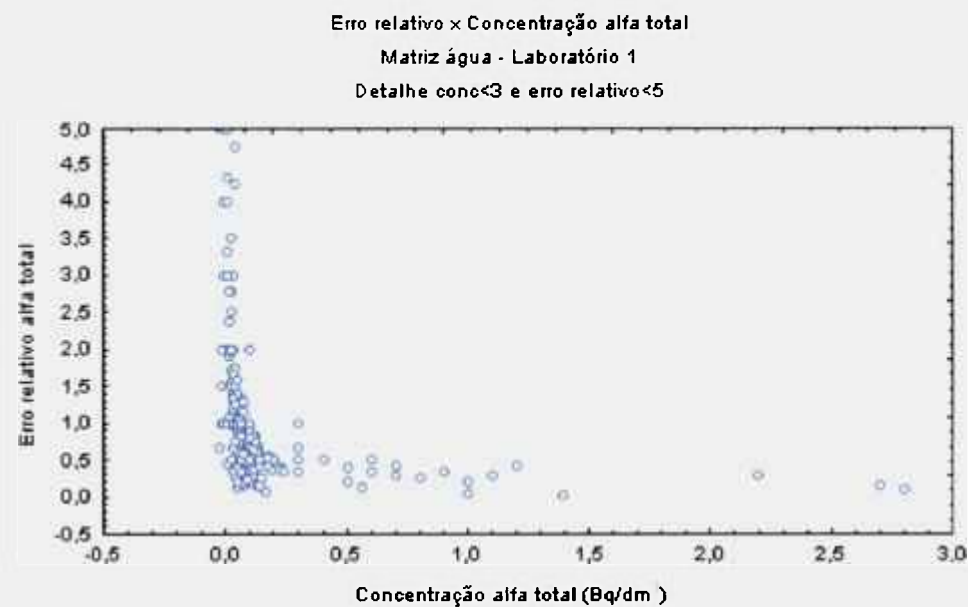


FIGURA 5-3 Aderência ao modelo: erro relativo x concentração – concentração < 3,0

O ajuste entre os resultados produzidos e a expressão prevista pelo protocolo analítico mostra o desempenho do laboratório em relação ao modelo de medição. Especialmente na faixa das baixas concentrações, é possível verificar que os resultados estão alinhados segundo a relação matemática deste modelo. Isto demonstra que, durante o período focalizado, o processo de medição foi executado segundo o protocolo analítico e que as condições de medição não apresentaram grandes variações podendo-se concluir que o processo esteve sob controle estatístico. A comparação, na região das baixas concentrações, entre os resultados produzidos e o modelo de medição adotado, é apresentada na FIGURA 5-4 e na FIGURA 5-5, para o erro publicado e para o erro relativo em função da concentração. Para cada concentração medida, as diferenças entre o erro publicado e o erro ajustado pelo modelo baseado na estatística de Poisson refletem as pequenas variações nas condições de medição que ocorrem na rotina diária do laboratório. Protocolos que requerem múltiplas etapas e vários operadores têm potencialmente mais chances de introduzir variações no resultado final da medição. O protocolo de análise alfa total é um protocolo bastante simples do ponto de vista operacional; as pequenas variações observadas podem ter se originado em mudanças de operador, em processos de automação e atualização que, com certeza, ocorreram durante o período analisado.

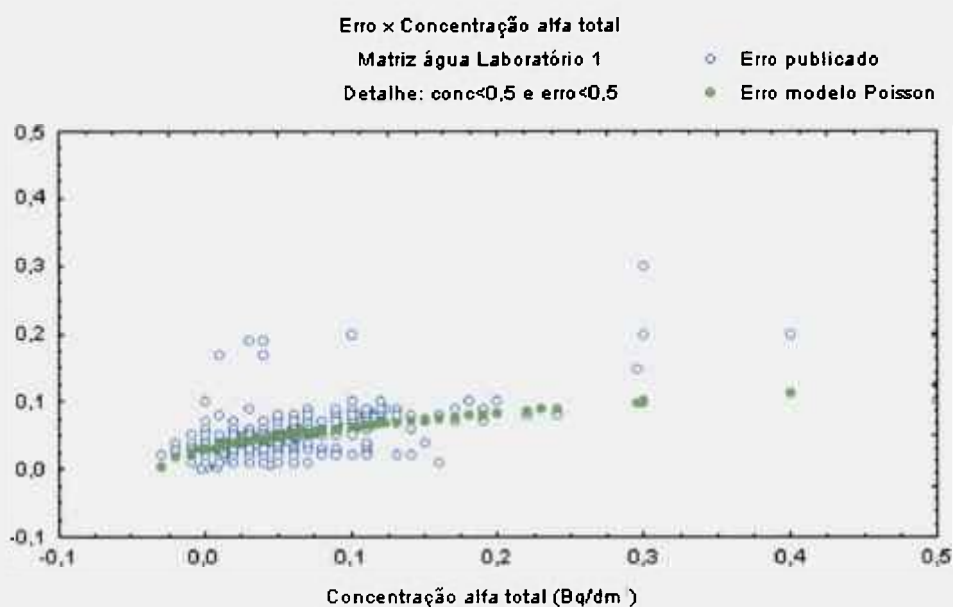


FIGURA 5-4 Detalhe: erro x concentração alfa total - modelo Poisson

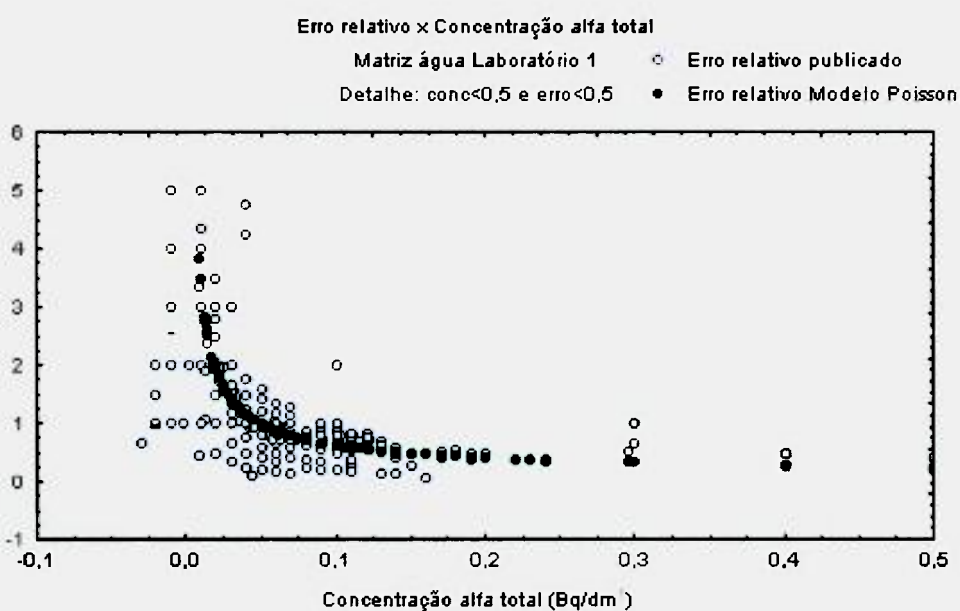


FIGURA 5-5 Detalhe: erro relativo x concentração alfa total - modelo Poisson

A estimação dos parâmetros do modelo de dois componentes por ajuste aos resultados produzidos, e que descrevem o desempenho do laboratório em relação ao protocolo analítico, estará incorporando a correção relativa à diferença entre os dois modelos: o que foi utilizado no processo de medição e o de Rocke & Lorenzato.

A comparação entre o modelo baseado na estatística de Poisson e o modelo de dois componentes é ilustrada na FIGURA 5-6 e na FIGURA 5-7 onde estão também representados os resultados publicados.

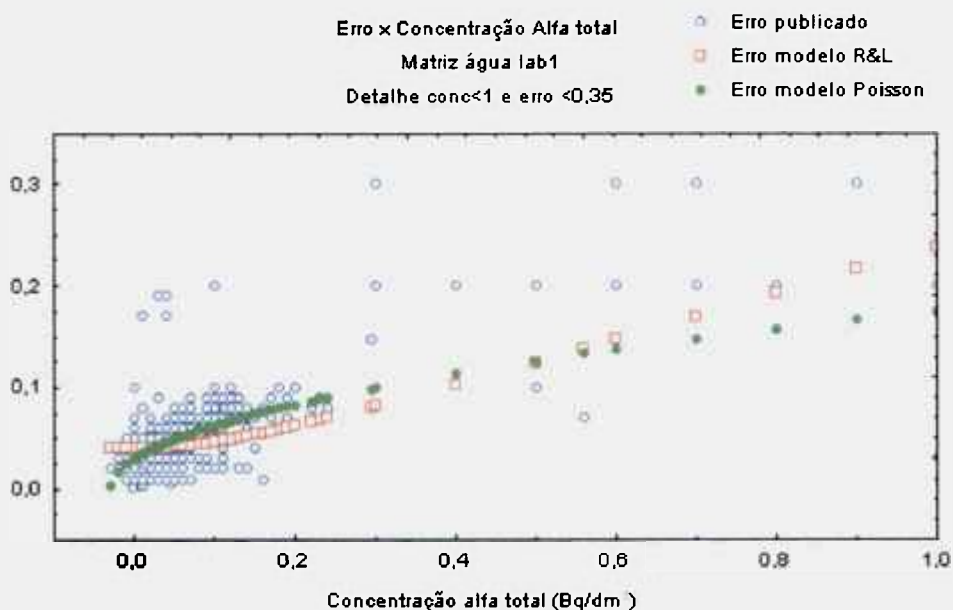


FIGURA 5-6 Detalhe: erro x concentração alfa total, modelo Poisson e R&L

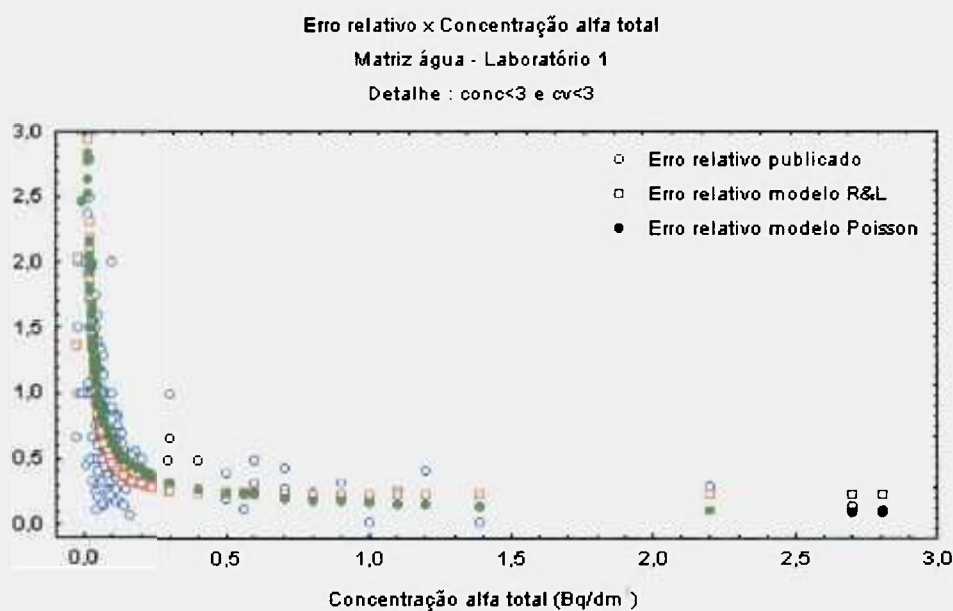


FIGURA 5-7 Detalhe: erro relativo x concentração alfa total, modelo Poisson e R&L

O parâmetro  $\sigma_\epsilon$  do modelo de R&L, que descreve a variabilidade do resultado da medição para as concentrações muito próximas a zero, é determinado pelos valores situados na região das baixas concentrações. Este conjunto de dados contém um grande número de registros nesta faixa de concentração permitindo uma boa confiabilidade no resultado deste ajuste. O valor do erro é constante nesta região sendo, portanto, o valor associado à concentração zero a ser usado em várias das definições propostas de limite de detecção. Considerando a definição do limite de detecção do método (LDM) da USEPA

(citado na equação 2-4), o valor  $0,04 \text{ Bq/dm}^3$  estimado para  $\sigma_e$ , para este caso, projeta uma concentração mínima detectável de  $0,09 \text{ Bq/dm}^3$ . (Considerando que a estimação de  $\sigma_e$  foi feita sobre um número grande de pontos foi usado  $z = 2,3$  da distribuição normal.)

Na FIGURA 5-8, é apresentado o detalhe do comportamento do erro em função da concentração para a região próxima a zero onde o erro é constante e a concentração mínima detectável (CMD) é atingida.

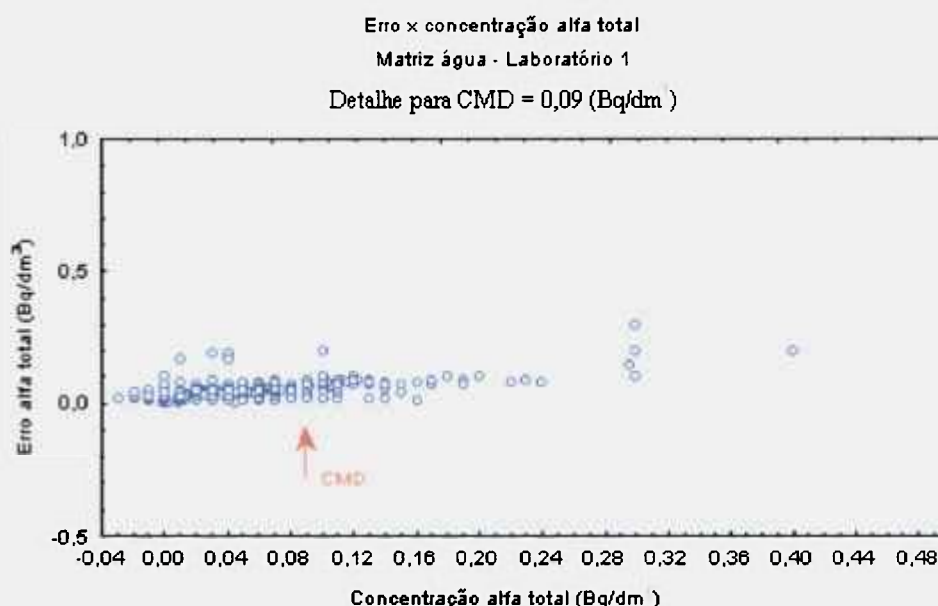


FIGURA 5-8 Visualização do limite de detecção, alfa total, águas, lab1

Qualquer que tenha sido a definição usada para calcular a concentração mínima detectável usando o valor de  $\sigma_e$  (determinado por estimação sobre os resultados de amostras reais), seu valor estará representando o limite efetivamente praticado pelo laboratório ao executar o protocolo analítico ao longo do período de tempo estudado.

As médias estimadas para cada local de amostragem, apresentadas na Tabela 4.4, mostram valores maiores que 1 para a razão  $R_{MP}$ , entre a média estimada usando o fator de ponderação calculado pelo inverso do quadrado das incertezas previstas pelo modelo de Rocke & Lorenzato e a média onde o fator de ponderação foi calculado pelo inverso do quadrado das incertezas publicadas. Isto significa que a média representativa do conjunto de resultados produzidos por este protocolo analítico para cada local e estimada usando os resultados publicados apresenta um valor menor que o verdadeiro. As variações entre os valores desta razão ocorrem, sobretudo, em virtude das diferentes distribuições do valor das concentrações medidas em cada local.

Um estudo da influência da faixa de concentração sobre a razão entre as médias, para este caso, é apresentado na TABELA 5-1. A região das baixas concentrações é onde a relação apresenta seu maior valor. É um resultado esperado em virtude da diferença de comportamento previsto pelos dois modelos para esta região: erro constante pelo modelo de R&L e crescente para o modelo baseado em Poisson. A estimativa das médias não é severamente afetada para valores intermediários da concentração apresentando uma razão 1,07. Nas concentrações mais altas esta razão mostra um comportamento inverso, ou seja a média estimada usando os resultados publicados no cálculo do fator de ponderação é superestimada. Para o conjunto de dados que está sendo analisado, duas considerações devem ser feitas. A primeira diz respeito a estimação do parâmetro  $\sigma_\eta$ , que governa esta região. Como já foi mencionado, este conjunto apresenta poucos registros nesta região e a determinação do parâmetro fica muito dependente destes valores, podendo não corresponder à realidade do processo analítico. A segunda é que, do ponto de vista teórico, a expressão da função raiz quadrada cresce menos rapidamente que a exponencial, projetando valores menores para as incertezas calculadas para um modelo descrito por ela. Na região das altas concentrações, portanto, a média estimada usando como fator de ponderação o inverso do quadrado destas incertezas seria maior que aquela estimada pelas incertezas previstas pelo modelo de R&L .

TABELA 5-1 Estimativa das médias por faixa de concentração alfa total, águas, lab1

Análise Alfa - Total Matriz- água Laboratório 1					
Valores usados para os parâmetros: $\sigma_x = 0,0404 \pm 0,0015$ (Bq/dm <sup>3</sup> ) $\sigma_\eta = 0,2261 \pm 0,0014$					
Faixa de concentração	N	Média simples	Média ponderada publicado	Média ponderada modelo R&L	Razão ( $R_{MP}$ ) modelo R&L/publicado
<0,1	748	0,0312	0,0194	0,0292	1,51
[0,1 – 0,5)	90	0,188	0,144	0,153	1,07
>=0,5	26	2,15	1,10	0,705	0,64

As diferenças observadas entre as razões  $R_{MP}$  para cada local podem ser explicadas pela amplitude e variabilidade de cada conjunto de resultados. O Apêndice B apresenta a descrição estatística e ilustração do conjunto de resultados referentes a cada local de coleta. O resumo estatístico para os resultados de cada local é apresentado na TABELA 5-2.

TABELA 5-2 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>) águas, lab1

Local	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
006042	130	0,063	0,030	-0,03	1,1	0,026	0,014
006043	128	0,0459	0,020	-0,02	1,0	0,011	0,0094
Local 049	149	0,25	0,030	-0,6	25	4,2	0,17
008046	28	0,067	0,040	-0,01	0,56	0,012	0,020
008047	32	0,0486	0,040	-0,01	0,22	0,0019	0,0077
009048	76	0,043	0,029	-0,01	0,70	0,0078	0,010
010001	49	0,046	0,030	0,00	0,50	0,006	0,011
010002	43	0,0370	0,030	-0,01	0,11	0,0011	0,0050
010003	50	0,0364	0,030	-0,02	0,24	0,0020	0,0063

O local 049, que apresenta um valor alto para a razão  $R_{MP}$ , possui um grande número de resultados na região intermediária com valores dos erros publicados maiores que aqueles previstos pelo modelo de R&L. A ponderação pelo inverso do quadrado dos valores previstos pelo modelo de R&L levará a estimativa da média para um valor maior, já que valores menores para os erros serão associados a concentrações maiores, nesta região. Este comportamento é ilustrado na FIGURA 5-9.

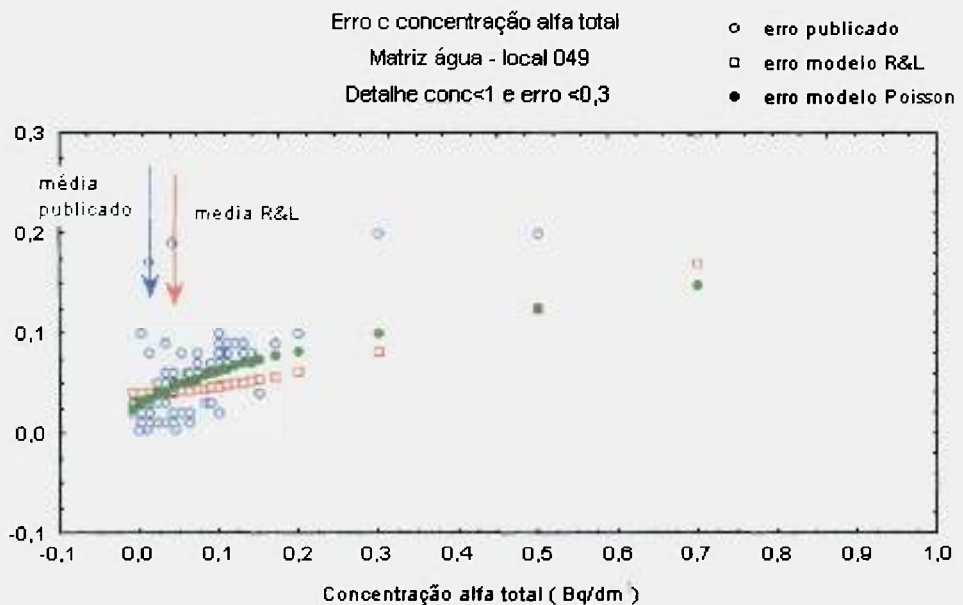


FIGURA 5-9 Detalhe erro x concentração ,alfa total, água, local 049

A influência do modelo de medição sobre a avaliação das incertezas e conseqüentemente sobre a estimação da média está relacionada com a região onde estão concentrados os resultados e com a dispersão destes resultados. Os gráficos que são apresentados a seguir ilustram esta influência para cada local e justificam as diferenças encontradas nas razões entre as médias  $x_P$  e  $x_{RL}$  apresentadas na TABELA 4.4.

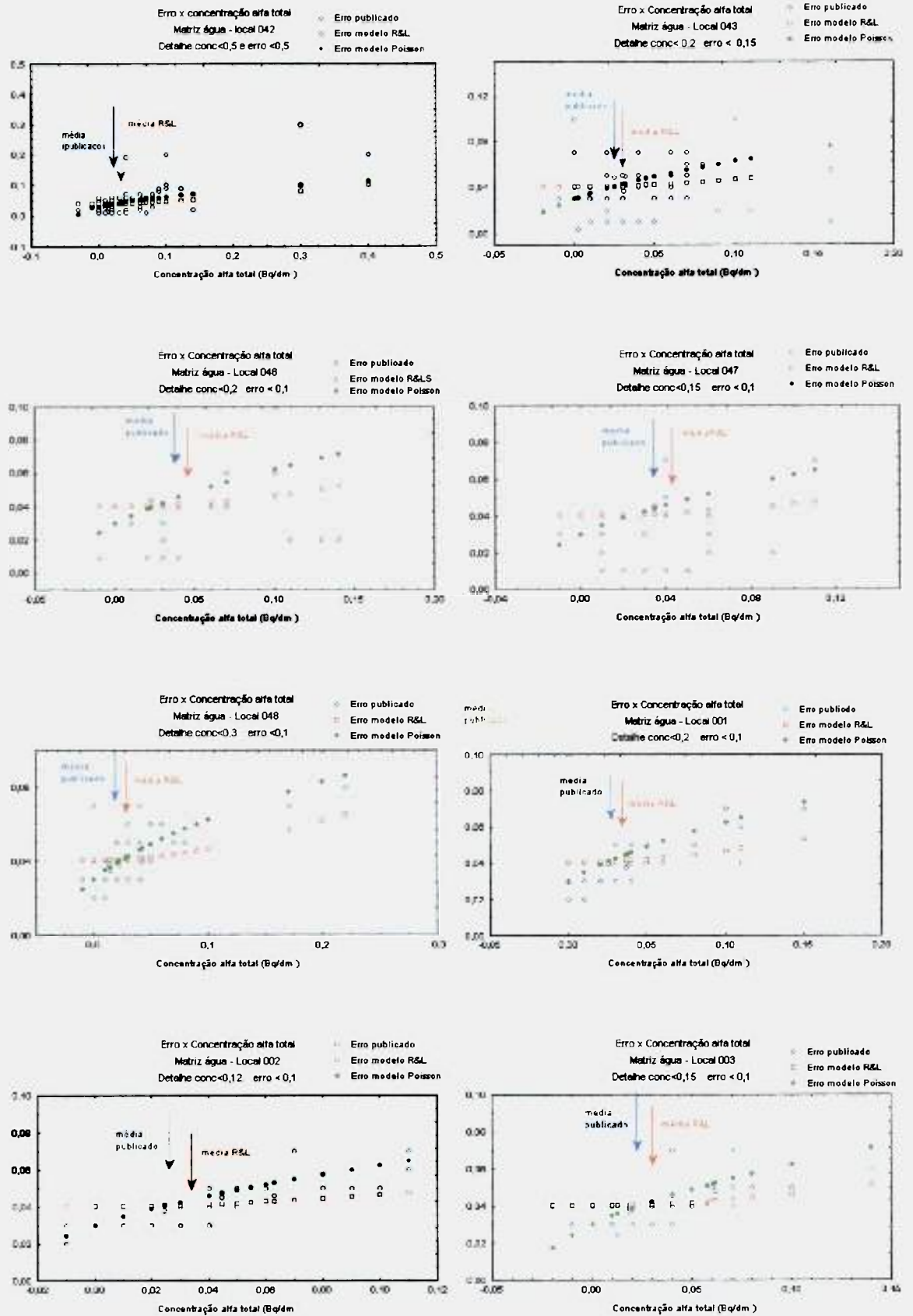


FIGURA 5-10 Visualização dos ajustes para cada local amostrado

### 5.3 Caso 2: Análise alfa total, matriz água, lab2

As amostras analisadas pelo lab2 são coletadas para monitorar a saída de efluentes líquidos dos laboratórios de modo que podem eventualmente conter elementos ou substâncias em teores mais elevados que o background natural. O conjunto de dados referente às análises destas amostras apresenta uma amplitude muito grande no valor dos resultados das concentrações e a variabilidade dos resultados tanto dos valores das concentrações quanto dos erros publicados é bem maior que a variabilidade apresentada pelos registros do laboratório 1.

A análise exploratória deste conjunto mostrou um grupo de resultados com comportamento muito diferente do padrão apresentado pela maioria dos resultados. Este aglomerado de pontos parece refletir uma falha no controle estatístico do processo analítico. Não foram registradas ocorrências relativas a execução destas análises por isso estes resultados foram considerados atípicos e rejeitados com a hipótese de pertencerem a um período de tempo em que o processo esteve fora de controle estatístico.

Os resultados produzidos por este laboratório não apresentam um padrão de comportamento bem definido na faixa de valores da concentração próxima de zero. Na faixa de concentrações menores que  $0,5 \text{ Bq/dm}^3$  ocorrem resultados com erro constante e dois padrões diferentes de comportamento para a relação erro x concentração. A visualização do comportamento do conjunto de resultados para a faixa de concentração próxima de zero é apresentada na FIGURA 5-11 e na FIGURA 5-12, que apresentam o erro publicado e o erro relativo em função da concentração, respectivamente.

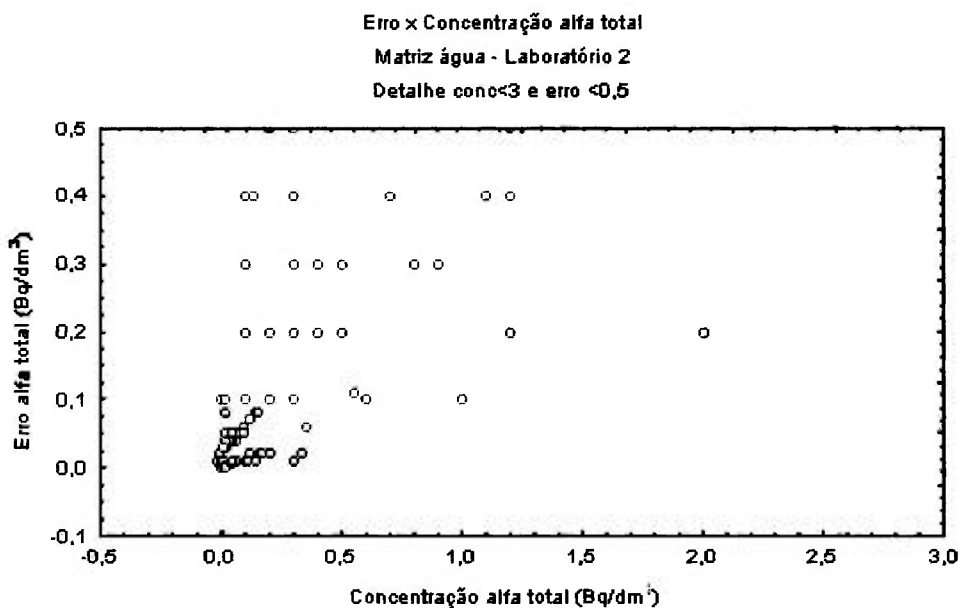


FIGURA 5-11 Detalhe -erro x concentração, alfa total, águas, lab2 ( sem nuvem)

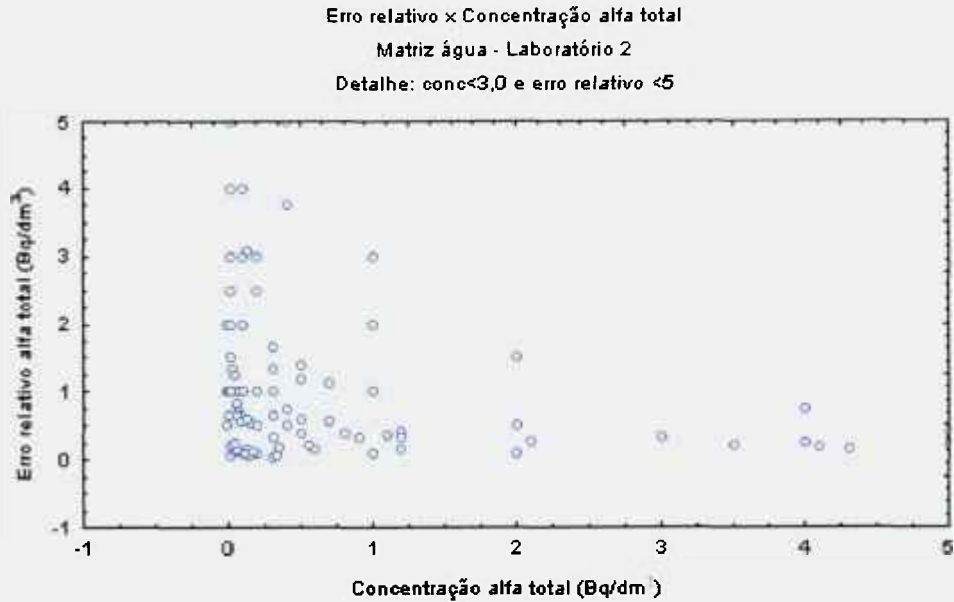


FIGURA 5-12 Detalhe –Erro relativo x concentração, alfa total, águas,lab2 (sem nuvem)

Pode-se observar, para um subconjunto de resultados, um comportamento de acordo com o modelo de Poisson, que é o modelo de medição especificado no protocolo analítico. Uma comparação entre os resultados deste conjunto e a expressão do modelo de Poisson mostra que, exceto para concentrações mais altas, os resultados se ajustam a esse modelo. A visualização deste ajuste é apresentada na FIGURA 5-13 para o conjunto de resultados selecionados. A relação entre o erro relativo e a concentração na região das baixas concentrações é apresentada no FIGURA 5-14, mostrando concordância entre os resultados publicados e a função descritiva do modelo.

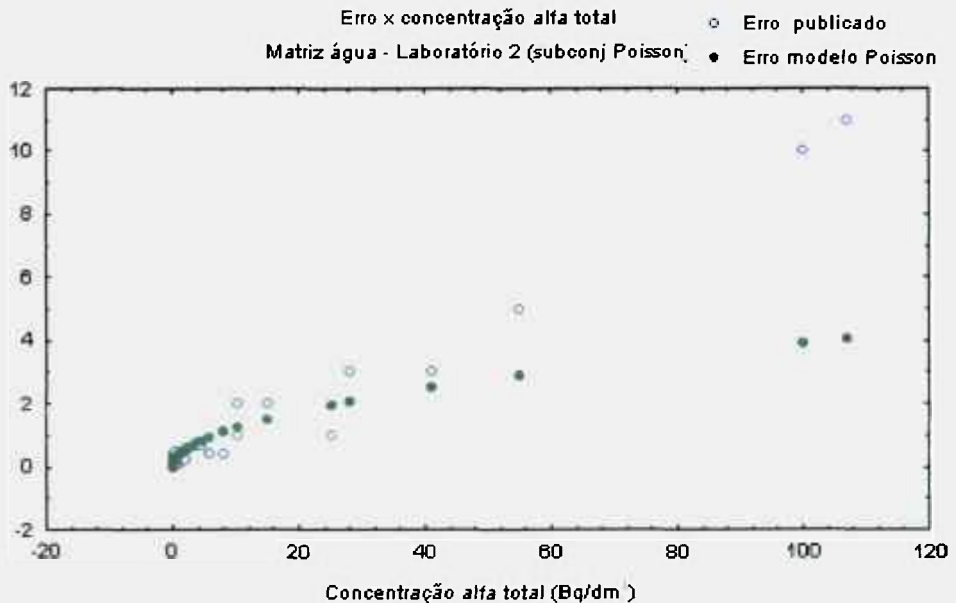


FIGURA 5-13 Erro x concentração alfa total, águas, lab2 (subconjunto Poisson)

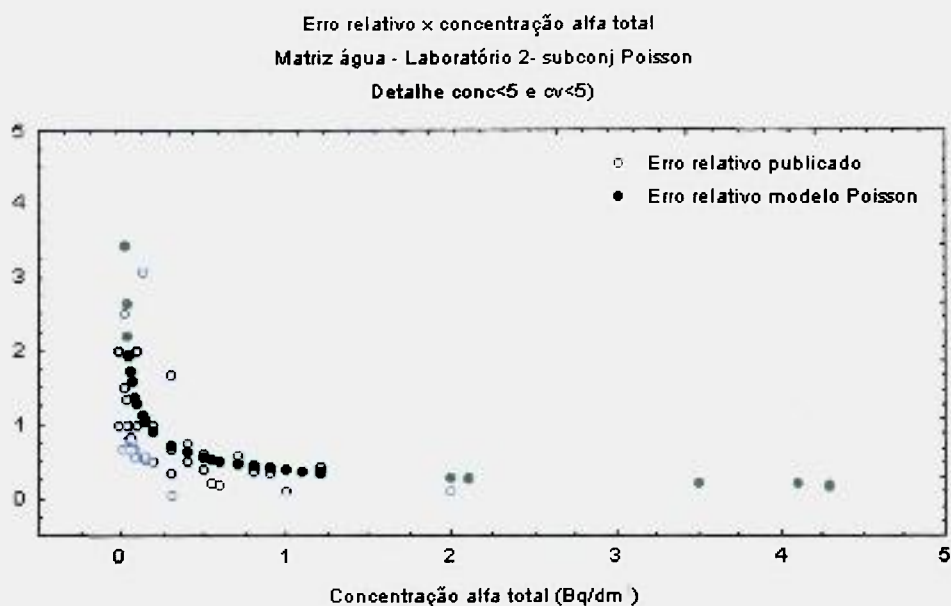


FIGURA 5-14 Erro x concentração alfa total, água, lab2, (subconjunto Poisson)

Para comparar o comportamento de todos os resultados produzidos por esse laboratório (excluídos aqueles pertencentes ao período considerado fora de controle estatístico) foi feita sua representação em função da concentração, juntamente com os resultados previstos em um modelo de Poisson, apresentada na FIGURA 5-15, e o comportamento entre o erro relativo e a concentração, para este mesmo conjunto de dados, mostrado na FIGURA 5-16.

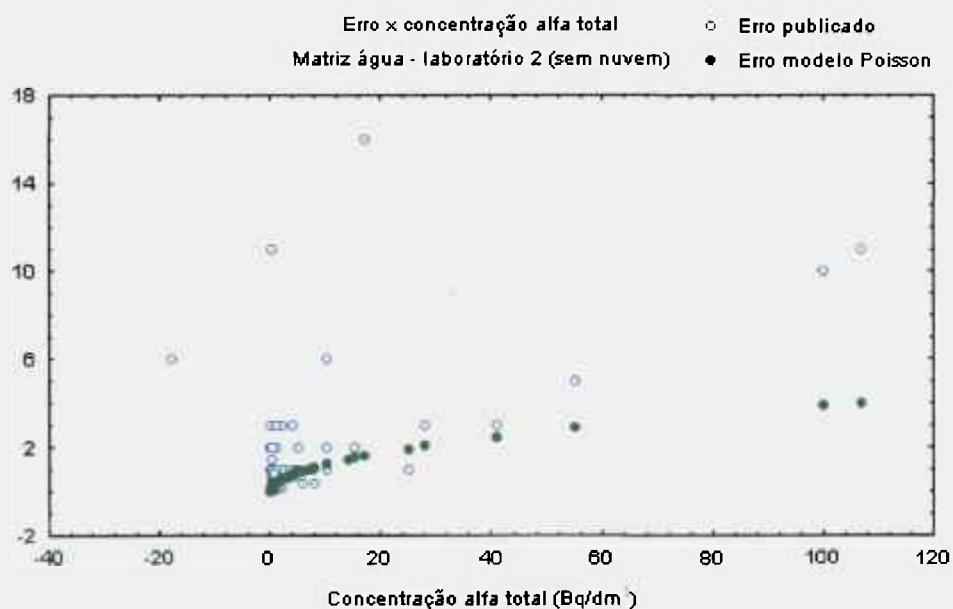


FIGURA 5-15 Erro x concentração alfa total, água, lab2 (sem nuvem)

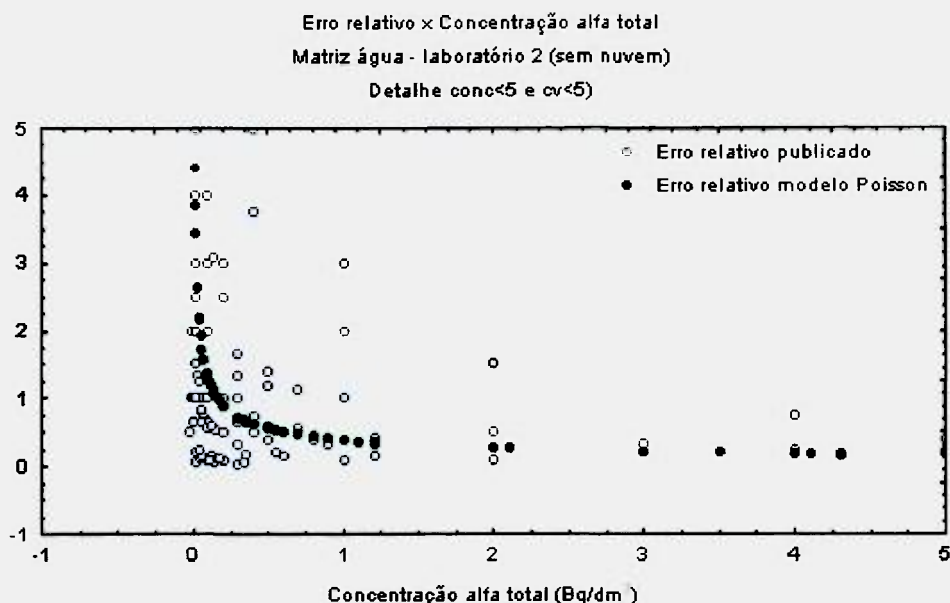


FIGURA 5-16 Detalhe: Erro relativo x concentração alfa total, água, lab2, modelo Poisson

O conjunto de resultados, excluindo o grupo considerado fora de controle estatístico e aqueles que não se ajustam a um modelo de Poisson, foi selecionado para o ajuste ao modelo R&L após terem sido excluídos os dados discrepantes.

A estimativa do parâmetro  $\sigma_\varepsilon$ , estimado pelo ajuste ao modelo de R&L, projeta um valor de  $0,39 \text{ Bq/dm}^3$  para o LDM, segundo a definição da USEPA (2001b). Este valor representaria efetivamente o limite alcançado por este laboratório na aplicação deste método. Entretanto a dispersão de pontos na região próxima a zero prejudica a qualidade do ajuste, o que pode ser notado pelo valor mais elevado da soma dos quadrados dos resíduos (TABELA 4.7). Assim o que se pode concluir sobre este conjunto de resultados é que este laboratório teve realmente dificuldades na prática analítica ao lidar com amostras de muito baixa atividade. Este subconjunto de dados apresentou amplitude maior de concentrações e ocorrência de maior número de registros nas concentrações mais elevadas, de modo que a estimativa de  $\sigma_\eta$  é mais realista que aquela do caso 1. De acordo com esta estimativa a execução do protocolo analítico por este laboratório tem um erro relativo de 0,097, segundo definição da USEPA (2001b).

A descrição estatística, a evolução temporal do erro e da concentração e comparação entre os erros publicados e aqueles previstos pelo modelo são apresentados para cada local de amostragem no Apêndice B. Os valores da razão  $R_{MP}$  para as médias relativas aos resultados das análises do laboratório 2 apresentaram uma grande variação entre os locais focalizados. Estes conjuntos de dados, como é de se esperar em se tratando de amostras de esgoto industrial, apresentam maiores amplitudes no valor das

concentrações. Isto acentua a diferença na estimativa das médias estimadas usando o modelo de dois componentes e aquelas estimadas usando o modelo de Poisson em razão da diferença de comportamento dos dois modelos nas concentrações mais altas. O resumo estatístico do conjunto de resultados para cada local, é apresentado na TABELA 5-3.

TABELA 5-3 Estatística descritiva: concentração alfa total, águas, lab2

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
021037	38	4,0	0,25	-18	41	105	1,7
021038	29	11,1	3,0	0,13	100	533	4,3
021040	47	-0,07	0,15	-18	5	8,2	0,42
021041	26	2,4	0,35	-0,01	25	28	1,0
021050	80	0,277	0,10	-0,02	5	0,39	0,070

Como este conjunto de dados apresentou resultados censurados, as estimativas das médias foram feitas de dois modos: excluindo os resultados censurados e fazendo a sua substituição pelo valor do LD. É interessante notar que, ao fazer esta substituição, é possível prever um valor de incerteza para estes resultados. Desta forma, o par concentração-erro é considerado na estimativa das médias ponderadas, o que não é possível ocorrer quando o laboratório publica um resultado como “<LD”. Os valores  $R_{MP}$  apresentados na tabela 4.8 e 4.9 não podem ser comparados entre si, já que a média ponderada usando como fator de ponderação o inverso do quadrado dos erros publicados é composta sempre pelo número de resultados analíticos publicados como detectados.

#### 5.4 Caso 3: Análise urânio 238, matriz água

O conjunto de todos resultados produzidos por este protocolo analítico, mostra, na fase inicial, um comportamento diferente daquele apresentado no restante do período: resultados negativos com altos valores absolutos e modelo de medição que considera o erro relativo constante. Todos os resultados publicados como “menor que” pertencem a este período. Os gráficos de evolução temporal apresentados no Apêndice A (gráficos 21 e 22) ilustram este comportamento. Estes resultados foram excluídos do subconjunto a ser estudado na estimação dos parâmetros do modelo de dois componentes. Registros representando concentrações com erro relativo constante, que ocorreram nas baixas concentrações, também foram excluídos do subconjunto a ser usado no ajuste para determinação dos parâmetros do modelo de R&L. O conjunto resultante apresenta um padrão muito próximo àquele previsto pelo modelo de R&L: erro constante nas baixas concentrações e uma relação de proporcionalidade nas concentrações mais altas. Por isto

estes resultados apresentaram um ajuste de boa qualidade para o modelo de R&L. A visualização deste ajuste pode ser observado na FIGURA 5-17 e na FIGURA 5-18.

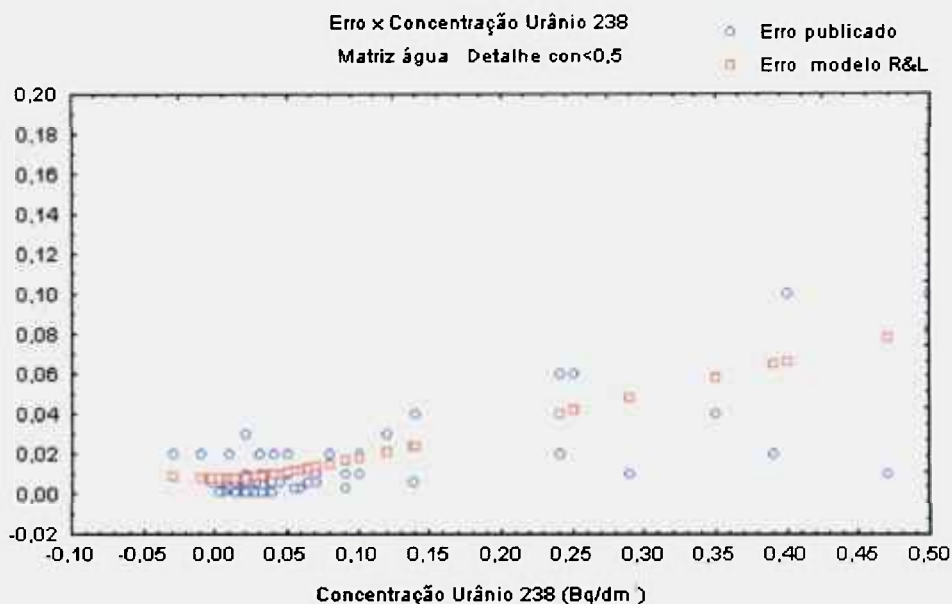


FIGURA 5-17 Detalhe: erro x concentração U238, água, modelo R&L

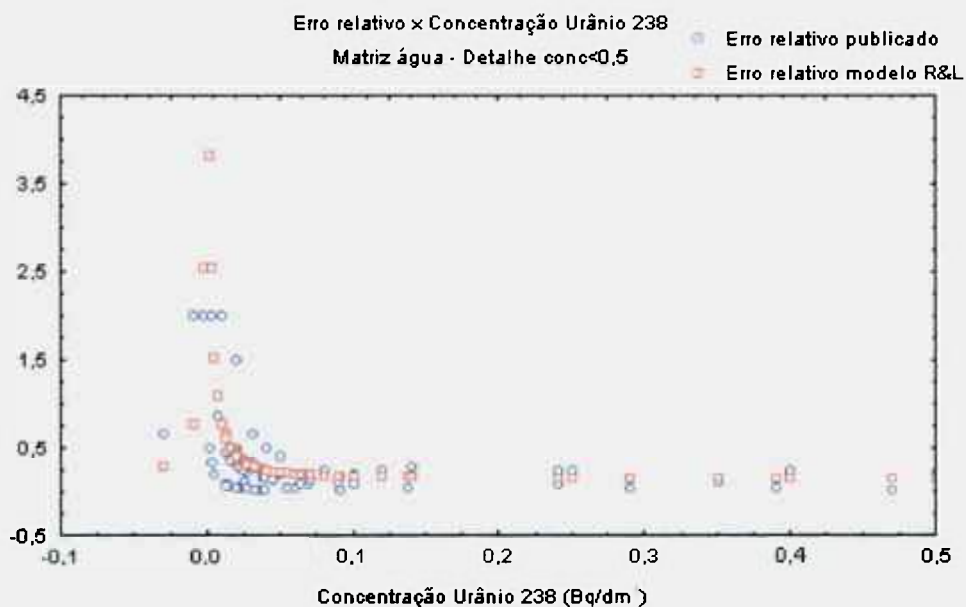


FIGURA 5-18 Detalhe: Erro relativo x concentração urânio 238, águas

Usando os valores dos parâmetros do modelo de dois componentes, o Limite de Detecção do Método (LDM), calculado pela expressão 2.6, para este protocolo analítico é  $0,0017 \text{ Bq/dm}^3$ , e a precisão, expressa pelo parâmetro  $\sigma_\eta$ , é 0,16.

Os resumos estatísticos do conjunto de resultados referentes a cada local são apresentados na TABELA 5-4.

TABELA 5-4 Estatística descritiva: concentração urânio 238, águas

Amostra	Concentração (Bq/dm <sup>3</sup> )						
	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
06042	46	0,044	0,0200	-0,020	0,470	0,0072	0,012
06043	47	0,041	0,016	-0,030	0,90	0,0176	0,019
06049	42	0,034	0,0085	-0,090	0,7700	0,0150	0,019
08046	31	0,0196	0,010	-0,060	0,200	0,0018	0,0076
08047	38	0,0196	0,0080	-0,0600	0,200	0,0019	0,0072
21050	37	0,224	0,0300	-0,0100	1,70	0,206	0,075

A observação dos gráficos elaborados para cada local de amostragem, que ilustram o comportamento do erro em função da concentração medida (Apêndice B), e dos resultados apresentados na tabela 4.1, mostra que o uso de modelos diferentes usados no processo de medição afeta severamente as estimativas das médias de cada local.

### 5.5 Caso 4: Análise alfa total, matriz aerossol

A análise exploratória do conjunto de resultados da análise alfa total para os três locais de amostragem de aerossol mostrou a utilização de três modelos de medição em períodos de tempo distintos. No período de nov/95 a set/99, o comportamento do erro publicado em função da concentração medida mostra um modelo de erro constante, no período dez/87 a mar/91 os resultados apresentam um comportamento semelhante ao modelo de erro relativo constante. Estes modelos descrevem bem a relação erro x concentração em faixas restritas, respectivamente valores próximos ao zero e valores mais elevados de concentração. Neste conjunto de resultados eles foram utilizados em intervalos de tempo especificados, independente do valor de concentração encontrado na análise. Por esta razão foram excluídos na análise estatística. O subconjunto restante mostra um comportamento próximo ao modelo de Poisson. Este subconjunto é mostrado na FIGURA 5-19, onde também é mostrada a curva resultante de um ajuste ao modelo baseado na estatística de Poisson para ilustrar o comportamento deste conjunto. É interessante notar que nas concentrações mais altas o modelo de Poisson prevê erros bem menores do que aqueles que foram publicados.

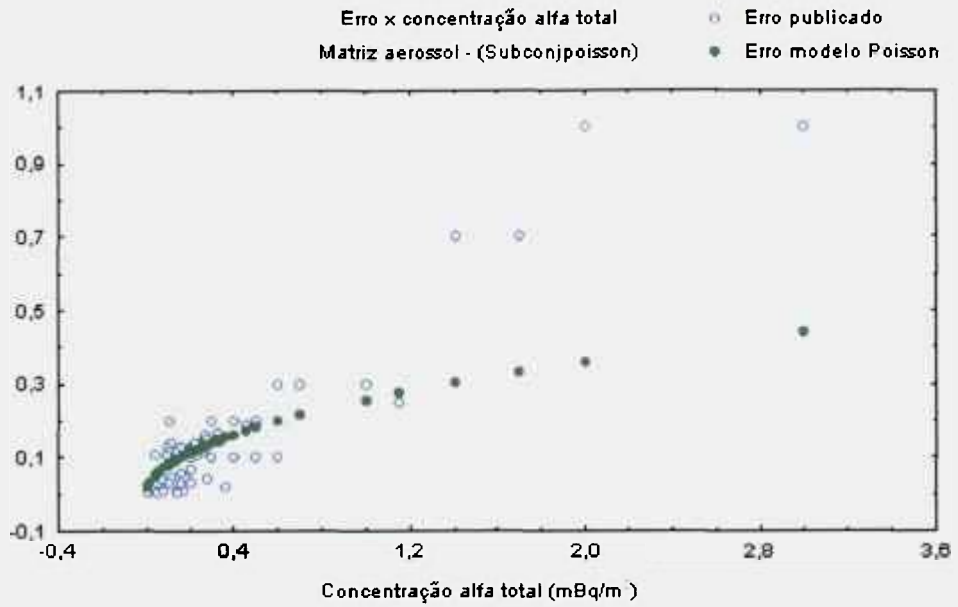


FIGURA 5-19 Detalhe erro x concentração alfa total , aerossol (subconjunto Poisson)

A comparação entre todos resultados publicados para o erro e concentração e os valores previstos pelo modelo de dois componentes é mostrada na FIGURA 5-20 e na FIGURA 5-21.

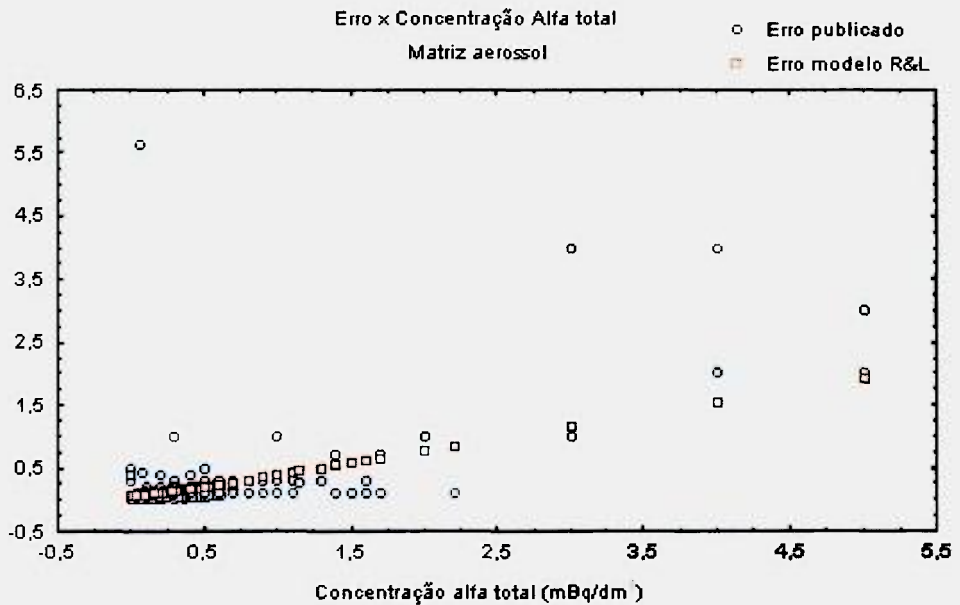


FIGURA 5-20 Erro x concentração alfa total, aerossol, modelo R&L

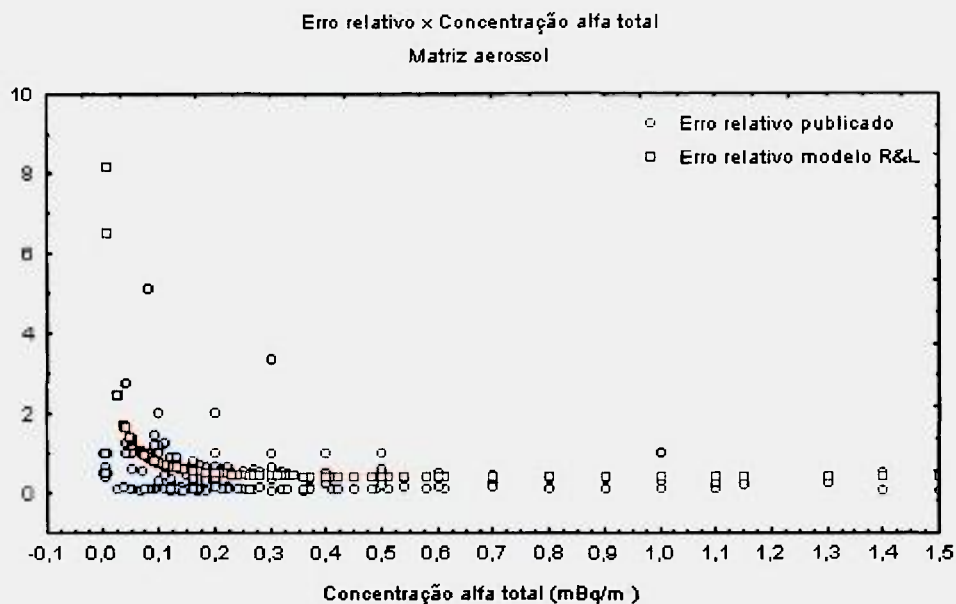


FIGURA 5-21 Erro relativo x concentração alfa total, aerossol, modelo R&L

Para este conjunto de dados o valor do parâmetro  $\sigma_{\epsilon}$ , estimado no ajuste ao modelo de dois componentes, projeta um LDM de  $0,15 \text{ mBq/m}^3$  e o de  $\sigma_{\eta}$  aponta para um erro relativo de 0,35.

As médias estimadas para os três locais de amostragem são bastante diferentes entre si. Observando os gráficos descritivos de cada local (Apêndice B) pode-se notar que foram afetados de maneira diferente pela aplicação dos modelos de medição efetivamente usados neste protocolo. Na TABELA 5-1 é apresentado o resumo estatístico para cada local de amostragem que mostra a grande semelhança apresentada entre o perfil estatístico dos três conjuntos de dados.

TABELA 5-5 Estatística descritiva: concentração alfa total, aerossol ( $\text{mBq/m}^3$ )

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
001001	97	0,478	0,30	0,0	4	0,34	0,059
001002	123	0,499	0,30	0,0	5	0,65	0,073
001003	135	0,454	0,30	0,0	3	0,24	0,042

Os resultados apresentados na tabela 4-18 mostram uma diferença muito grande entre a média estimada usando o fator de ponderação previsto pelo modelo de dois componentes e aquela estimada usando o fator de ponderação calculado pelo erro publicado para cada local. Esta diferença pode ser explicada, observando os gráficos erro x concentração para cada local (Apêndice B) onde pode ser visualizada a influência do modelo de medição sobre o conjunto de resultados. Esta influência sobre a estimativa da média depende de fatores como a amplitude das concentrações, a distribuição destes valores e o número de pontos apresentados pelo conjunto de resultados.

## 6. CONCLUSÕES

- A análise das medidas geradas pela aplicação de um mesmo protocolo analítico, considerado-as como uma amostra estatística, mostrou-se uma ferramenta eficaz na avaliação do desempenho do laboratório em relação ao processo de medição. Foi possível relacionar os padrões de comportamento apresentados pelos resultados publicados aos modelos de medição utilizados, identificar grupos de resultados atípicos em relação ao padrão predominante em um determinado período e distinguir períodos ou faixas de concentrações onde a aderência a um modelo foi deficiente. Em todos os casos focalizados foi possível identificar a conformidade, ou não, de grupos de resultados produzidos a um modelo de medição especificado. Apenas no primeiro caso houve a identificação da aplicação de um único modelo de medição durante todo o período analisado; no caso 4, três modelos foram usados em diferentes períodos de tempo e nos dois outros casos pode ser identificado o uso de mais de um modelo.
- Representando o limite de detecção e a precisão, as estimativas dos parâmetros do modelo de *Rocke & Lorenzato* traduzem a qualidade do processo analítico para cada um dos protocolos focalizados. Determinados por ajuste a resultados referentes a amostras verdadeiras, estes indicadores refletem as condições efetivamente praticadas pelo laboratório no período focalizado, em contraponto aos valores determinados em experimentos planejados e executados em amostras branco ou pseudo-amostras.
- A não-conformidade entre os resultados publicados e o modelo de medição aplicado introduz distorções no valor das incertezas associadas a cada medida que são relacionadas às variações temporais nas condições de operação do processo de medição. São pequenas mudanças que ocorrem na rotina diária e podem escapar aos procedimentos de controle de qualidade estabelecidos pelo laboratório. As distorções introduzidas pela aplicação de um modelo não adequado podem ser bastante severas como ocorreu nos casos em que mais de um modelo foi utilizado na mesma faixa de concentração.
- As distorções entre as incertezas publicadas e aquelas corrigidas para as variações temporais e para as deficiências do modelo de medição afetam a estimativa da

média de cada local de maneiras diversas. Influenciam mais ou menos a estimativa do valor representativo de cada conjunto dependendo do número de pontos que irão compor esta média, da distribuição destes pontos no domínio das concentrações e da intensidade da discrepância entre o modelo utilizado e o modelo de R&L.

É importante ressaltar a contribuição desta abordagem para o aperfeiçoamento dos protocolos analíticos executados e para o aprimoramento das práticas de laboratório. A comparação da relação erro x concentração com a expressão do modelo de medição usando todos os registros produzidos por um mesmo protocolo permite monitorar o desempenho do laboratório em relação aquele protocolo analítico em todos seus aspectos. Possibilita, assim, a descrição efetiva do processo analítico por meio de amostras verdadeiras, propiciando um controle de qualidade sem necessidade de medições específicas que consomem tempo, recursos financeiros e humanos e interrompem a rotina analítica. Monitorando a variação dos parâmetros do modelo (basicamente o limite de detecção e o erro relativo) pode-se conhecer a condição de controle estatístico para o protocolo analítico, avaliando o processo como um todo.

## **APÊNDICE A - ANÁLISE EXPLORATÓRIA DE DADOS**

Este Apêndice apresenta a análise exploratória dos conjuntos de dados, utilizados neste trabalho. Ela foi realizada buscando conhecer a estrutura dos dados e obter uma visão mais profunda da natureza de cada conjunto especialmente em relação às concentrações e incerteza analítica, que são as grandezas de interesse neste estudo. São apresentadas as estatísticas descritivas básicas, a evolução temporal para a concentração e erro publicado. Além disto, este estudo preliminar buscou identificar padrões de comportamento na relação erro x concentração e assim fornecer subsídios para a análise a ser desenvolvida.

Cada conjunto de dados focalizado é produto da execução do protocolo analítico para uma determinada análise e matriz. As estatísticas estimadas para estes conjuntos não têm significado físico, referem-se a diversos locais de amostragem fornecendo a descrição estatística do conjunto de resultados.

Os subconjuntos de resultados, selecionados para estimação dos parâmetros do modelo de Rocke & Lorenzato, são apresentados por seu resumo estatístico e visualizados pelos gráficos ilustrativos da relação erro x concentração.

### **A1. Análise alfa total - matriz água**

Neste programa de monitoração, a análise alfa total foi realizada por dois laboratórios distintos, em face da expectativa de níveis diferentes na atividade das amostras. Elas foram classificadas em dois grupos: aquelas com grande possibilidade de apresentarem baixa atividade e as que poderiam, em potencial, alcançar valores mais altos de concentração. No primeiro grupo encontram-se as águas classificadas como água de superfície, pluvial, potável e subterrânea e analisadas pelo laboratório aqui referenciado como lab1; no segundo grupo, as águas classificadas como esgoto industrial, analisadas pelo laboratório codificado como lab 2 neste estudo. A amostra coletada no ponto 049, classificada como esgoto industrial, foi analisada pelo laboratório 1. Esta amostra foi reclassificada como água de superfície na revisão do PMA em 1992.

O protocolo analítico para a análise de alfa total em amostras de água prevê filtragem de uma alíquota de 250 ml, deposição em bandejas de aço inox e contagem em

contador proporcional. A expressão para cálculo do erro analítico apresenta uma variação com a raiz quadrada da concentração baseada na estatística de Poisson.

### A11 Resultados produzidos pelo Laboratório 1

Este laboratório adotou como norma não censurar os resultados produzidos. Foram publicados os valores encontrados na medição, incluindo valores negativos, portanto não há resultados apresentados como “menor que o limite de detecção”. Os resultados analisados neste estudo correspondem ao período que vai de janeiro/1986 a dezembro/2001.

No período inicial de execução do programa foram executadas algumas análises investigativas em que amostras eram submetidas a pequenas modificações no seu tratamento físico. Os resultados destas análises, que não seguem rigorosamente o protocolo analítico, são diferenciados com as letras S, T, R acrescentados ao código de identificação adotado na formulação do programa.

Na revisão do PMA, realizada em janeiro de 1992, em que a amostra de esgoto industrial no local 049 foi reclassificada para água de superfície, as análises investigativas foram suspensas e foram reduzidos os locais de amostragem.

O número de resultados publicados para cada amostra ao longo do tempo é apresentado na FIGURA 1. A execução das análises de alfa total foi interrompida em janeiro de 1993 e retomada em agosto de 94, como pode ser observado no gráfico.

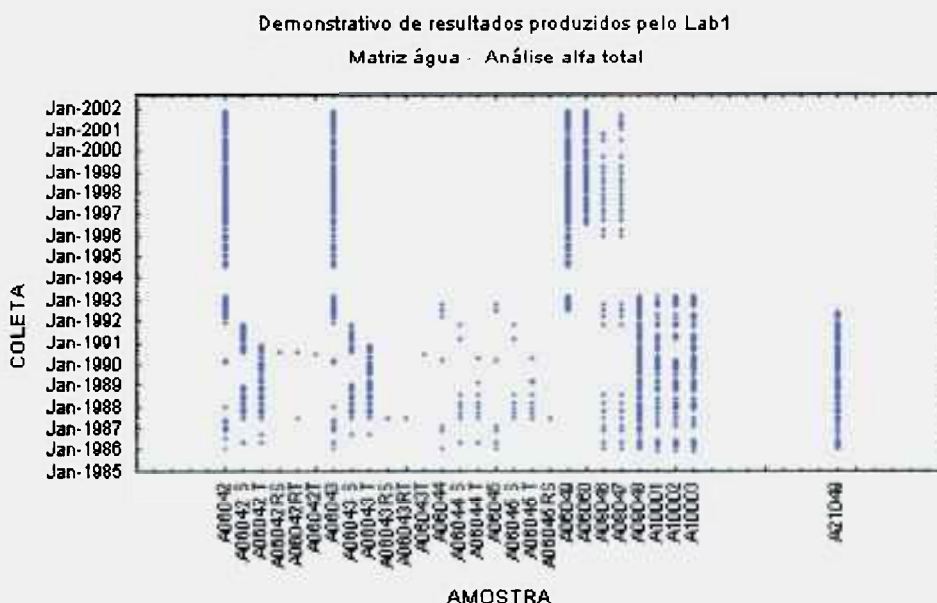


FIGURA 1 Demonstrativo de resultados produzidos, alfa total, matriz águas, lab 1

Para caracterizar estatisticamente este conjunto de resultados são apresentadas na TABELA 1 algumas estatísticas que foram estimadas para todo o conjunto de resultados publicados para análise alfa total por este laboratório. Estas estimativas foram feitas para as concentrações e para a incerteza publicada, que são as grandezas de interesse neste trabalho. É importante assinalar que os resultados apresentados nesta tabela não têm significado físico, são estimativas feitas para todo o conjunto de resultados produzidos pela aplicação do protocolo analítico para a análise alfa total da matriz água por um mesmo laboratório.

TABELA 1 Estatística descritiva: análise alfa total, matriz água, Lab 1

Período de coleta: Jan/86 a Dez/01								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	863	0,108	0,030	-0,6	25	0,80	25	697
Erro publicado (Bq/dm <sup>3</sup> )	863	0,069	0,040	0,002	6	0,064	17	368

A evolução temporal dos resultados da concentração alfa e também do erro associado a estes resultados são apresentados na FIGURA 2 e na FIGURA 3. Os resultados obtidos confirmam a previsão de baixa atividade prevista para estas amostras. Cerca de 99% dos resultados publicados situam-se na faixa de concentração menor que 3 Bq/dm<sup>3</sup> e 98% apresentam erro menor que 0,3 Bq/dm<sup>3</sup>.

Evolução temporal - Concentração alfa total  
Matriz água - Laboratório 1

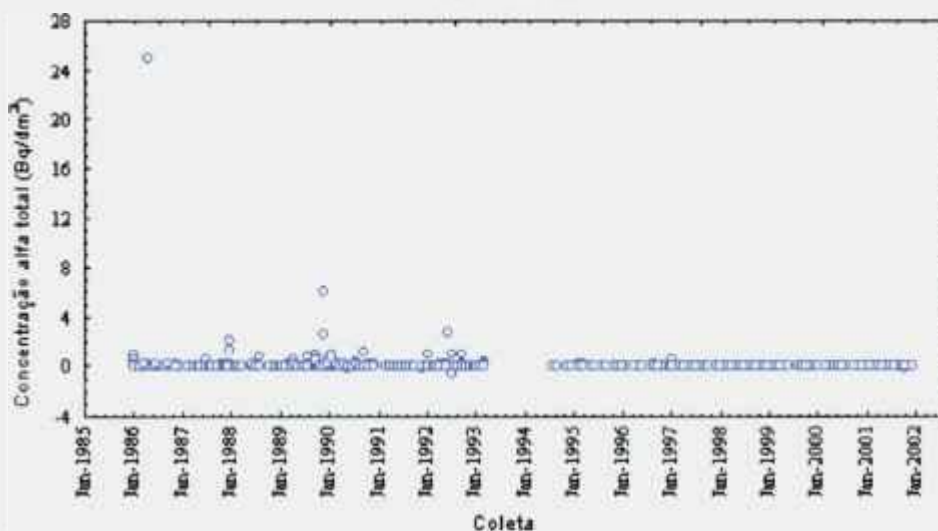


FIGURA 2 Evolução temporal: concentração alfa total, matriz água, Lab1

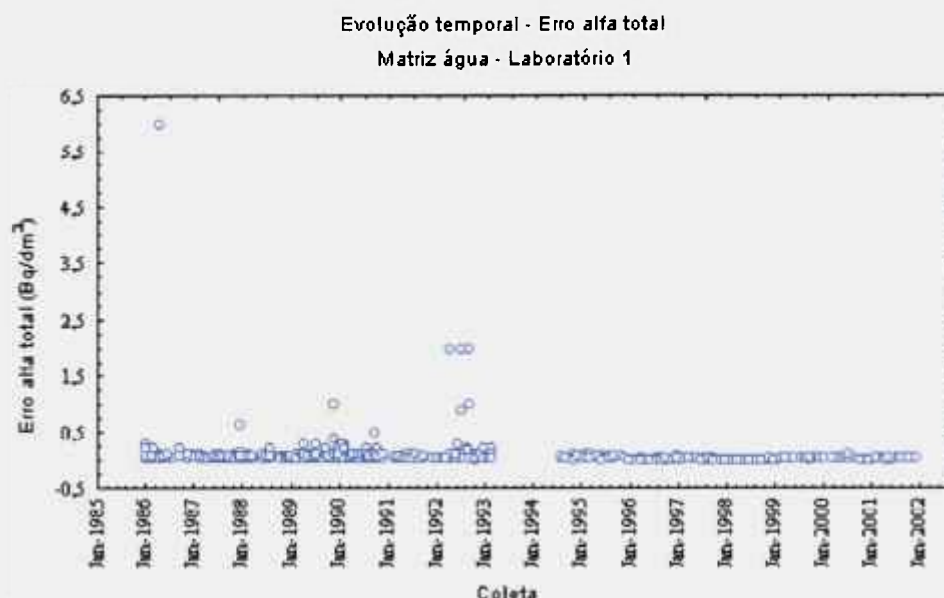


FIGURA 3 Evolução temporal: erro alfa total, matriz água, Lab1

A observação dos gráficos da evolução temporal da concentração e do erro, mostra uma diferença na variabilidade dos resultados referentes ao período anterior e posterior à interrupção da execução das análises. Neste segundo período houve uma redução no número de amostras analisadas e a execução das análises seguiu rigorosamente o protocolo analítico pois não foram mais realizadas análises investigativas.

A visualização do comportamento da concentração e do erro associado em função do tipo de amostra é apresentada na FIGURA 4 e na FIGURA 5. A finalidade destes gráficos é verificar possíveis diferenças na matriz, em virtude dos diversos tipos de água envolvidos, que possam interferir no resultado da análise. Como pode ser observado não há diferenças significativas no comportamento das amostras analisadas; mesmo a amostra classificada como esgoto industrial no local 049 apresenta apenas um ponto discrepante na amplitude tanto da concentração como do erro associado.

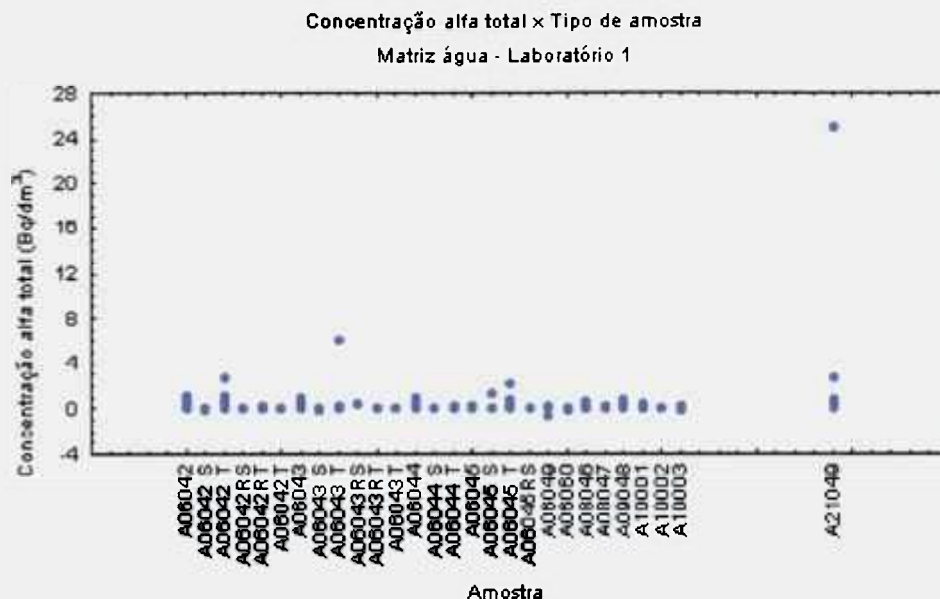


FIGURA 4 Comportamento da concentração x tipo de amostra: alfa total, matriz água, Lab1



FIGURA 5 Comportamento do erro em relação ao tipo de amostra, alfa total, matriz água, Lab1

As análises das amostras que foram realizadas seguindo rigorosamente o protocolo analítico, chamadas daqui por diante de planejadas, foram estudadas separadamente daquelas que tiveram alguma mudança em seu preparo e que serão referidas como análises investigativas. No segundo período de operação do PMA estas investigações não foram mais realizadas e portanto dele só constam resultados analíticos das amostras inicialmente planejadas. A comparação entre os dois conjuntos de resultados é feita para definir se eles podem ser considerados como um único conjunto que represente a aplicação do protocolo analítico.

A estatística descritiva para o conjunto de resultados das amostras planejadas para o período 1 é apresentada na TABELA 2 e a visualização da evolução temporal da concentração alfa total e do erro alfa total é apresentada na FIGURA 6 e na FIGURA 7.

TABELA 2 Estatística descritiva: análise alfa total, matriz água, Lab 1; planejadas

Período de coleta: Jan/86 a Mar/93								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	377	0,15	0,04	-0,6	25	1,7	19	359
Erro publicado (Bq/dm <sup>3</sup> )	377	0,10	0,04	0,01	6	0,14	12	175

Concentração alfa total - Evolução temporal  
Matriz água - Laboratório 1 - Período 1  
Amostras planejadas

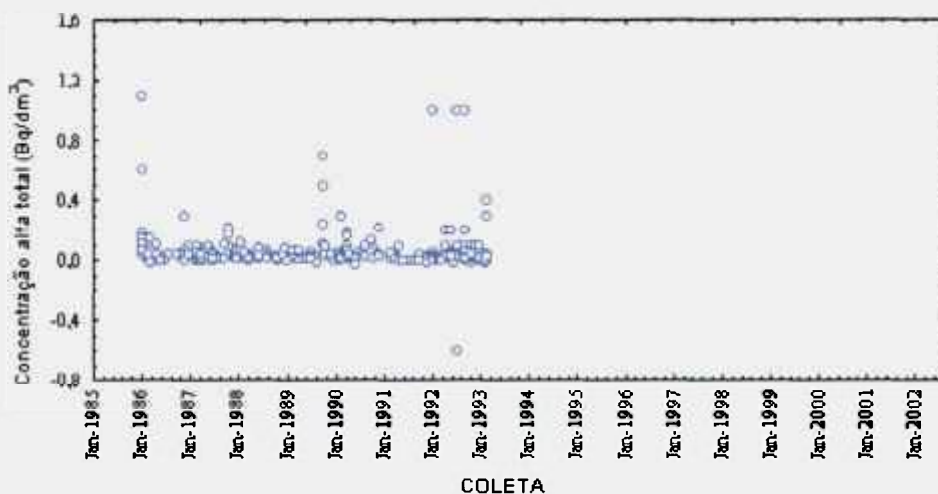


FIGURA 6 Evolução temporal: concentração alfa total, matriz água, Lab1, análises planejadas

Erro alfa total - Evolução temporal  
Matriz água - Laboratório 1 - Período 1  
Amostras planejadas

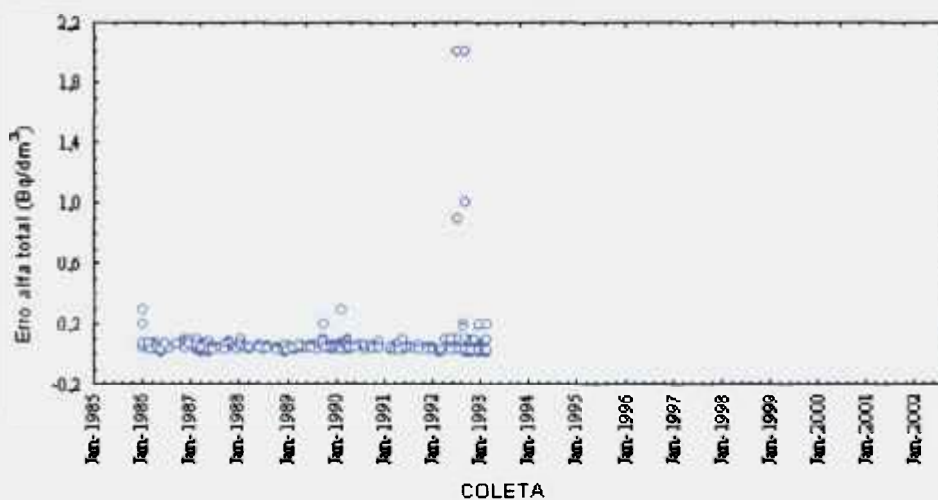


FIGURA 7 Evolução temporal: erro alfa total, matriz água, Lab1, análises planejadas

As análises investigativas foram executadas introduzindo-se pequenas modificações no protocolo analítico como filtração no momento da amostragem ou no laboratório, repetições em um tempo posterior à primeira análise. Estas repetições referem-se a amostras com alguma dificuldade em seu preparo ou que apresentaram resultados atípicos. A estatística descritiva para os resultados referentes a estas análises é apresentada na TABELA 3 e a visualização da evolução temporal da concentração e do erro alfa total é apresentada na FIGURA 8 e na FIGURA 9.

TABELA 3 Estatística descritiva: análise alfa total, matriz água, Lab 1; investigativas

Período de coleta: Abr/86 a Nov/91								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	165	0,17	0,050	-0,03	6	0,32	7,8	72
Erro publicado (Bq/dm <sup>3</sup> )	165	0,083	0,060	0,02	1	0,012	5,7	40

Evolução temporal - Concentração alfa total  
Matriz água - Laboratório 1  
Análises investigativas

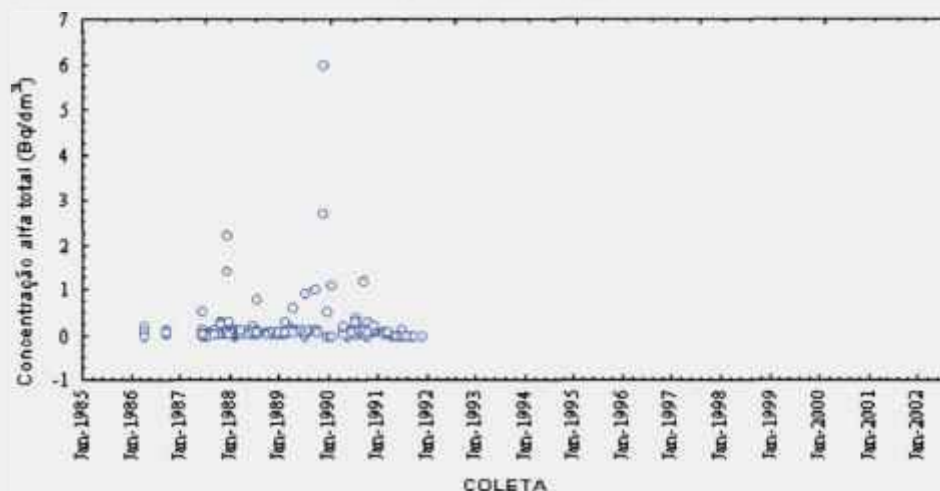


FIGURA 8 Evolução temporal: alfa total, Lab1, matriz água, análises investigativas

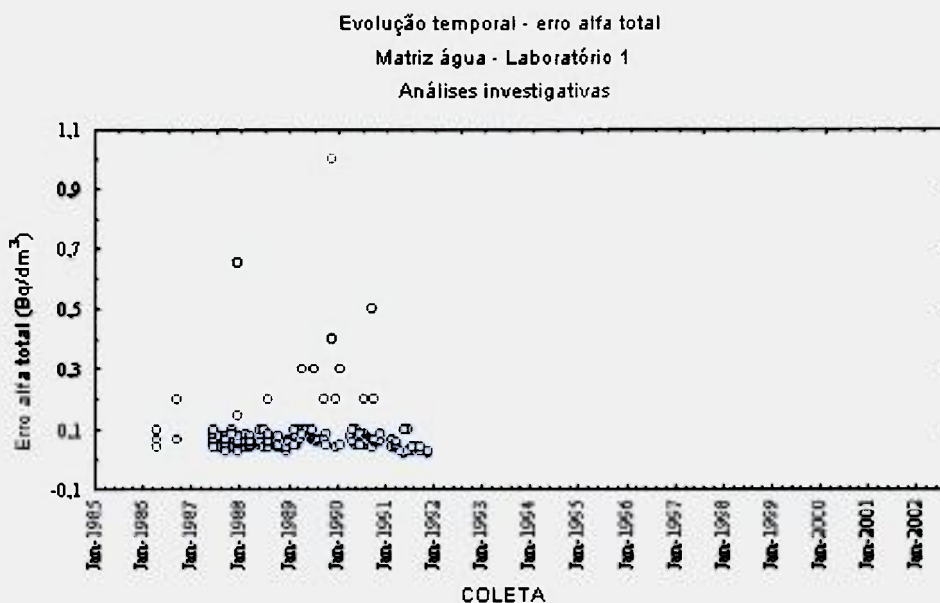


FIGURA 9 Evolução temporal: erro alfa total, Lab1, matriz água, análises investigativas

A amplitude dos resultados apresentados pelas análises investigativas é maior que aquela apresentada pelas análises planejadas por isso foi feita a estimativa das estatísticas descritivas para o intervalo comum (concentração menor que  $1,0 \text{ Bq/dm}^3$ ). Estas estimativas, apresentadas na TABELA 4 e na TABELA 5 mostram que os dois conjuntos têm médias e variâncias semelhantes e portanto os pequenos desvios na execução do protocolo analítico não influenciaram os resultados em relação ao modelo de medição adotado.

TABELA 4 Estatística descritiva: alfa total, matriz água, Lab 1; investigativas (conc<1)

Período de coleta: Abr/86 a Nov/91							
Conc <1	N	Média	Mediana	Mínimo	Máximo	Variância	Desv.Pad
Concentração ( $\text{Bq/dm}^3$ )	159	0,088	0,040	-0,03	1,0	0,022	0,15
Erro publicado ( $\text{Bq/dm}^3$ )	159	0,068	0,060	0,02	0,30	0,0020	0,045

TABELA 5 Estatística descritiva: alfa total, matriz água, Lab 1; planejadas (conc<1)

Período de coleta: Jan/86 a Mar/93							
Conc <1	N	Média	Mediana	Mínimo	Máximo	Variância	Desv.Pad
Concentração ( $\text{Bq/dm}^3$ )	374	0,070	0,040	-0,6	1,0	0,019	0,14
Erro publicado ( $\text{Bq/dm}^3$ )	374	0,084	0,040	0,010	2,0	0,046	0,21

O conjunto de resultados representativo da aplicação do protocolo analítico para a análise alfa total na matriz água será composto por todos os resultados produzidos pelo lab 1. O modelo de medição baseado na estatística de Poisson foi efetivamente seguido durante todo o período focalizado. O comportamento do erro publicado em função da concentração para estes resultados é apresentado na FIGURA 10 onde estão representados todos resultados produzidos pelo laboratório 1.

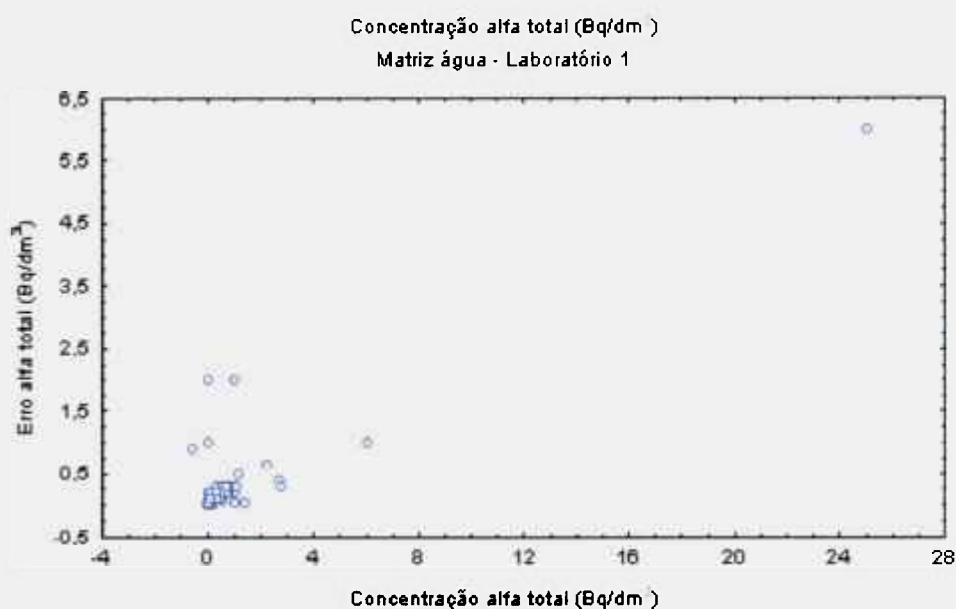


FIGURA 10 Comportamento do erro x concentração alfa total, matriz água, lab 1

A observação deste gráfico mostra alguns pontos fora do padrão apresentado pela maioria dos resultados. Estes pontos discrepantes, situados na região delimitada em concentração < 2 e erro > 0,5 devem ser excluídos do estudo de determinação dos parâmetros de desempenho do sistema. O ponto de coordenadas (25,6) é um resultado discrepante em relação aos resultados de concentração, porém na relação erro x concentração não se pode afirmar que este ponto não segue o padrão seguido pelos outros resultados. O gráfico do comportamento do erro relativo em função da concentração para este conjunto de resultados mostra que este resultado, visualizado na FIGURA 11, se ajusta ao comportamento previsto pelo modelo adotado.

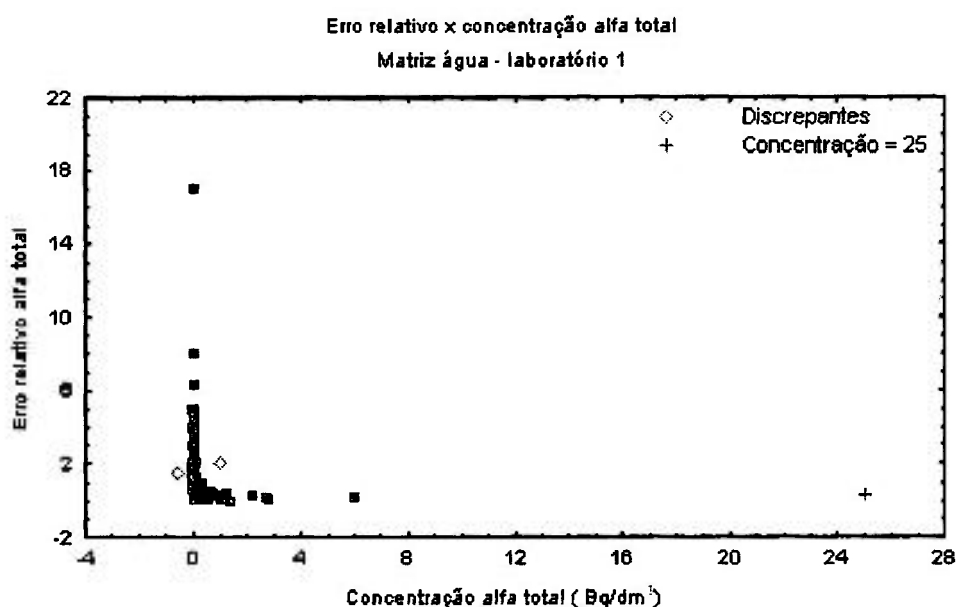


FIGURA 11 Comportamento do erro relativo x concentração alfa total, matriz água, lab 1

Os dados discrepantes a serem excluídos do conjunto de resultados na determinação dos parâmetros estão discriminados na TABELA 6.

TABELA 6 Dados discrepantes: análise alfa total, matriz água, Lab 1

Nº registro	Amostra	Coleta	concentração	erro
IA00410	06049	10/07/92	-0,6	0,9
IA00510	06049	15/09/92	0,0	1,0
IQ00215	21049	28/02/92	0,0	0,010
IA00410	06042	16/7/92	1,0	2,0
IA00513	06042	15/09/92		
IA00410	06044	16/07/92		

A descrição estatística do conjunto de dados a ser usado no estudo do comportamento do erro em função da concentração é apresentada na TABELA 7.

TABELA 7 Estatística descritiva: subconjunto de resultados alfa total, matriz água, Lab 1

Período de coleta: Jan/86 a Dez/01								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	857	0,106	0,030	-0,03	25	0,80	25	698
Erro publicado (Bq/dm <sup>3</sup> )	857	0,058	0,040	0,002	6	0,045	26	722

## A12 Resultados produzidos pelo Laboratório 2

O laboratório 2 executou análises das amostras de água classificadas como esgoto industrial usando protocolo analítico similar aquele praticado pelo laboratório 1. As

diferenças entre os dois protocolos analíticos referem-se aos valores característicos da instrumentação e aos detalhes da rotina analítica própria de cada laboratório.

A partir de abril/96, os resultados apresentados pelo laboratório 2 referem-se apenas à amostra coletada no local 050. O número de resultados produzidos ao longo do tempo para cada local de amostragem é apresentado na FIGURA 12.

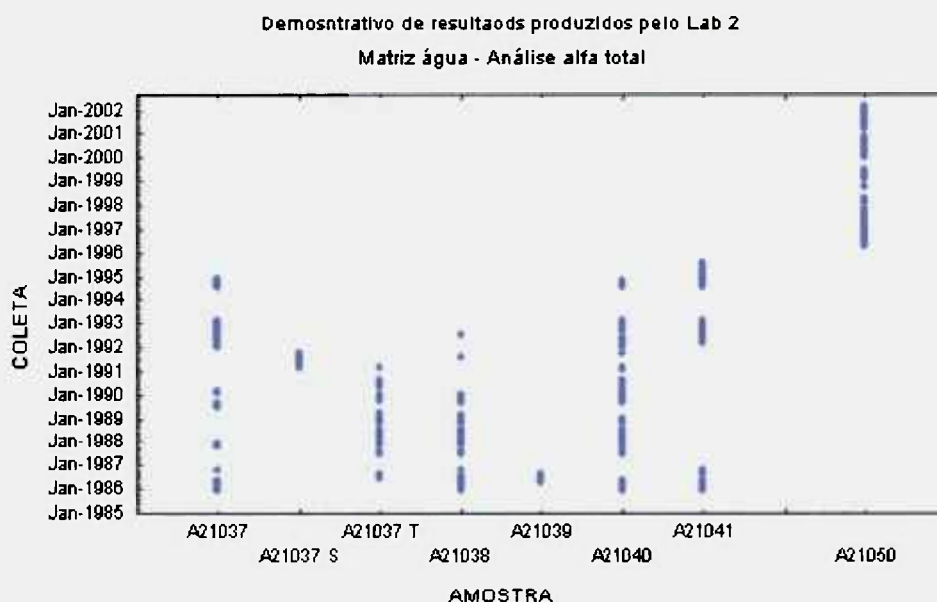


FIGURA 12 Demonstrativo de resultados produzidos pelo lab2; alfa total, matriz água

A descrição estatística do conjunto de todos os resultados publicados para o laboratório 2 é apresentada na TABELA 8. A evolução temporal da concentração e do erro publicado é apresentada na FIGURA 13 e na FIGURA 14 para todos resultados produzidos pelo laboratório 2.

TABELA 8 Estatística descritiva para resultados alfa total, matriz água, Lab 2

Periodo de coleta: Jan/86 a Mar/02								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	250	3,18	0,20	-18	107	148	6,2	45
Erro publicado (Bq/dm <sup>3</sup> )	222	2,57	0,30	0,001	22	28	2,4	4,6

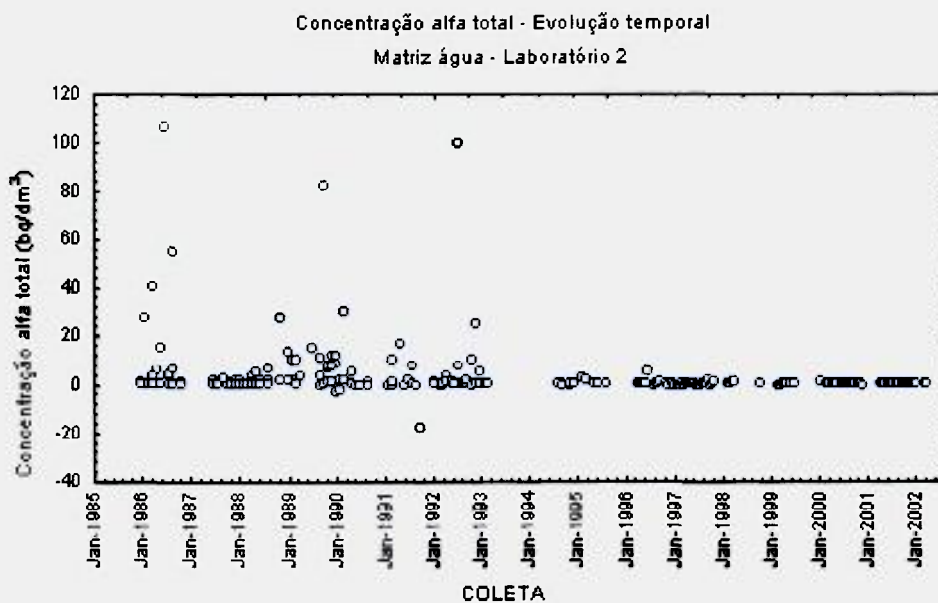


FIGURA 13 Evolução temporal: concentração alfa total, matriz água, Lab2

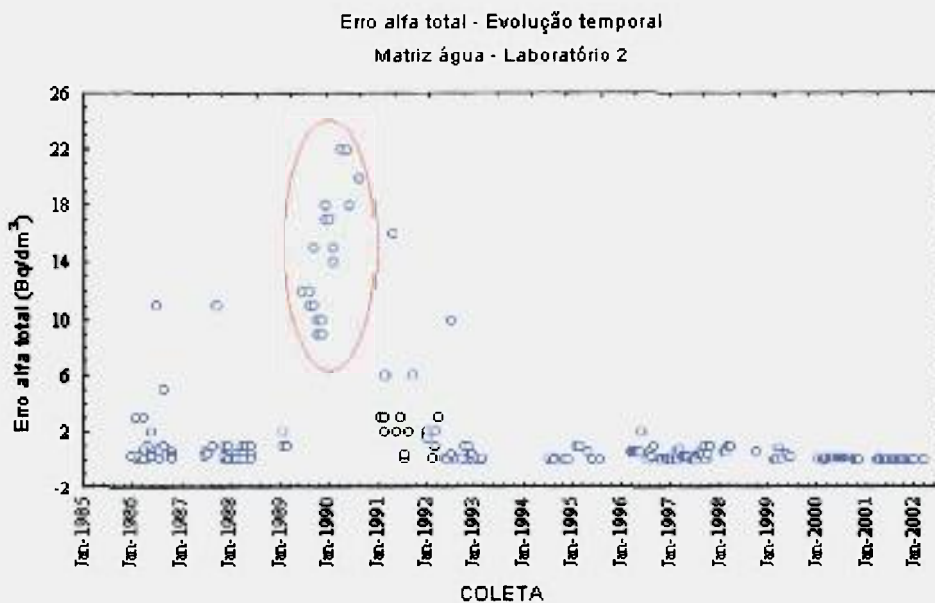


FIGURA 14 Evolução temporal: erro alfa total, matriz água, Lab2

Os erros publicados apresentam valores bastante altos no intervalo de tempo que vai de julho/89 a set/90 e que não estão acompanhando um aumento no valor das concentrações. Este aglomerado de pontos pode ser observado com clareza no gráfico erro em função da concentração apresentado na FIGURA 15 onde aparece destacado do conjunto de resultados das concentrações mais baixas.

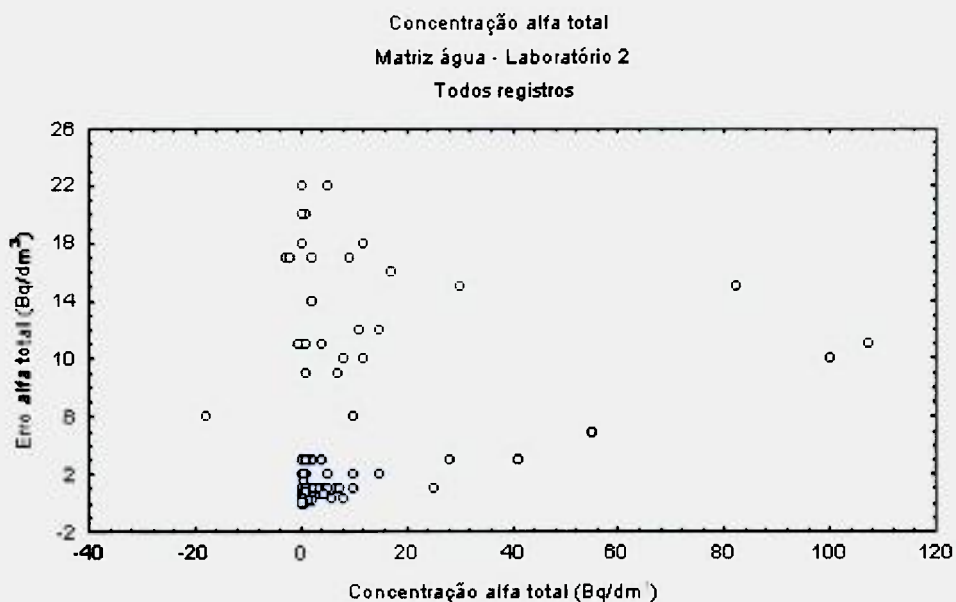


FIGURA 15 Comportamento do erro x concentração alfa total, matriz água, Lab 2

Os resultados que compõem este aglomerado podem refletir uma falha no controle de qualidade do protocolo analítico seja em relação ao preparo das amostras seja no desempenho da instrumentação eletrônica. Os registros relativos a este período serão excluídos do conjunto a ser usado para estimar os parâmetros do modelo de medição. Uma visualização dos resultados deste período é apresentada na FIGURA 16. Pode-se notar que o comportamento do erro em relação à concentração além de não seguir o mesmo padrão que o conjunto dos outros resultados, não apresenta um padrão referente a algum dos modelos já conhecidos.

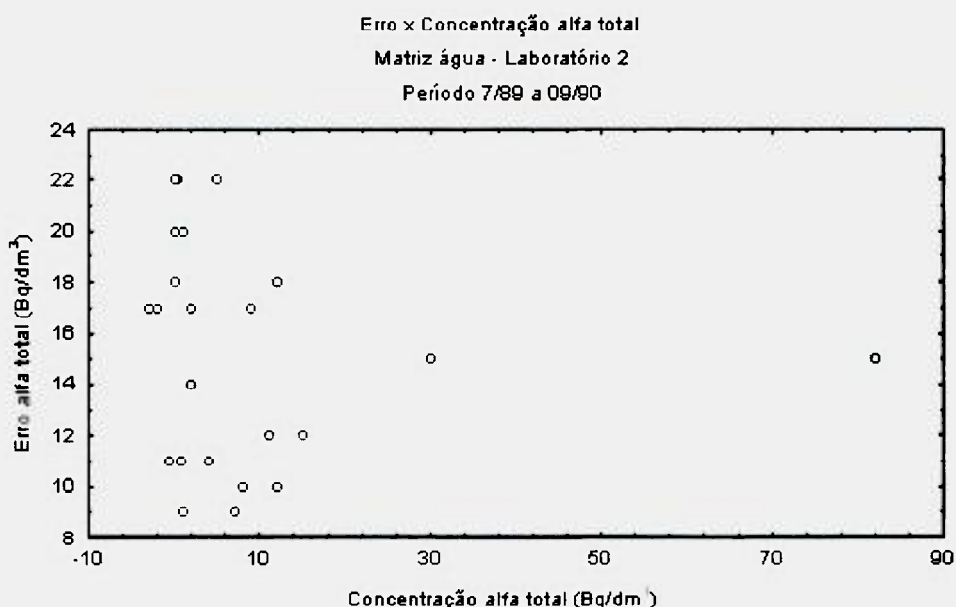


FIGURA 16 Comportamento do aglomerado; erro x concentração alfa total, matriz água, Lab 2

A estatística descritiva para o conjunto de resultados após a exclusão dos registros atípicos é apresentada na TABELA 9.

TABELA 9 Estatística descritiva para resultados alfa total, matriz água, Lab 2; sem nuvem

Período de coleta: Jan/86 a Mar/02								
	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	223	2,64	0,20	-18	107	131	6,9	56
Erro publicado (Bq/dm <sup>3</sup> )	195	0,84	0,20	0,001	16	3,8	4,9	28

O comportamento do erro em função da concentração para este conjunto de resultados é apresentado na FIGURA 17. Neste conjunto pode-se observar, nas concentrações mais altas, um comportamento proporcional ao erro. Nas concentrações próximas de zero entretanto o padrão de comportamento da relação entre o erro e a concentração não é tão bem definido.

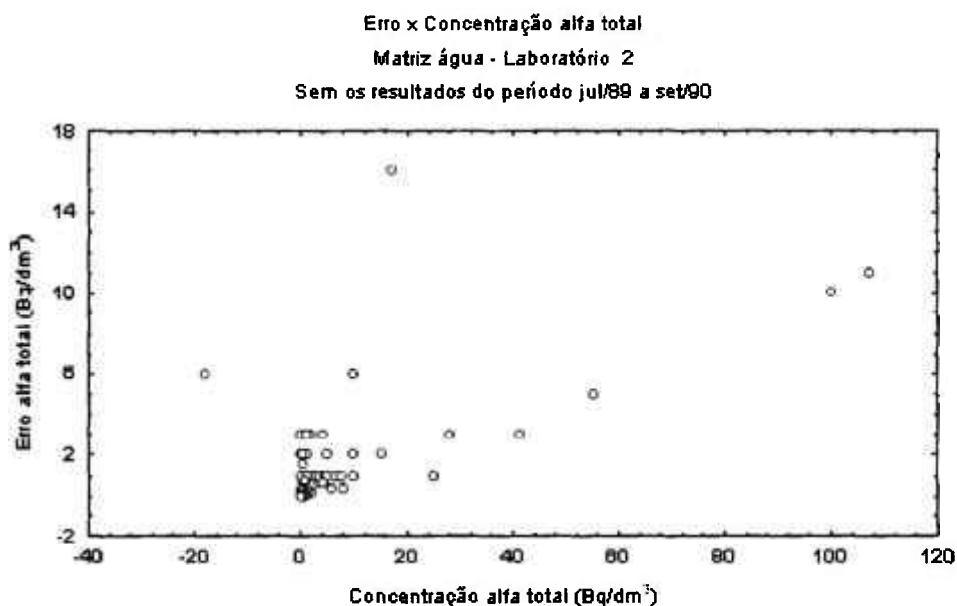


FIGURA 17 Comportamento do erro x concentração, alfa total, matriz água, Lab 2

Um detalhamento do comportamento do erro em função da concentração, apresentado na FIGURA 18, para a região de concentrações mais baixas mostra grupos de resultados com valor constante para o erro em concentrações crescentes (erro = 1,0 a 2,0 e 3,0). Estes resultados serão excluídos do conjunto a ser usado na determinação dos parâmetros do modelo de Rocke & Lorenzato.

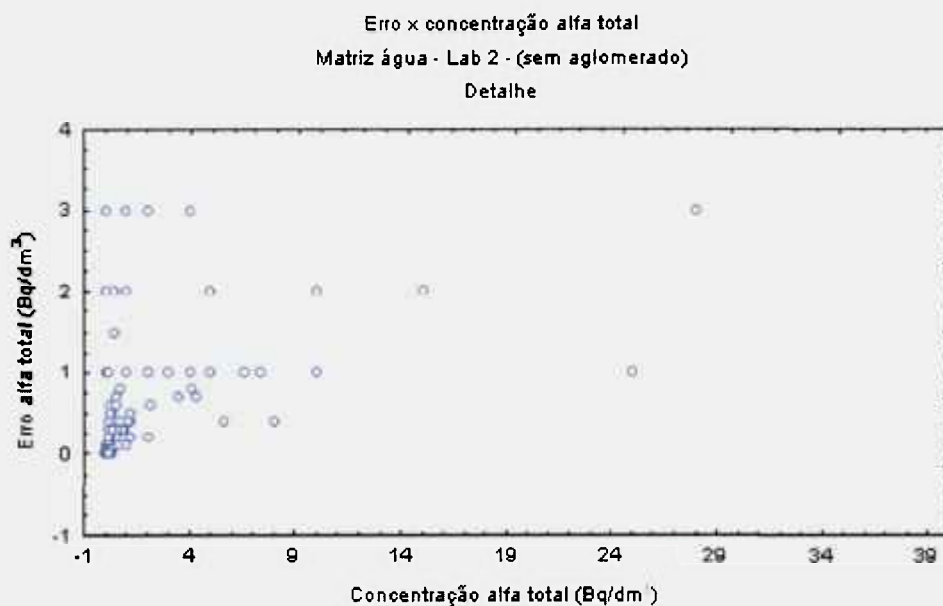


FIGURA 18 Comportamento do erro x concentração: alfa total, matriz água, Lab 2; detalhe

Os resultados a serem usados na determinação dos parâmetros do modelo de medição são apresentados na FIGURA 19 e as estatísticas descritivas destes resultados são apresentadas na TABELA 10.

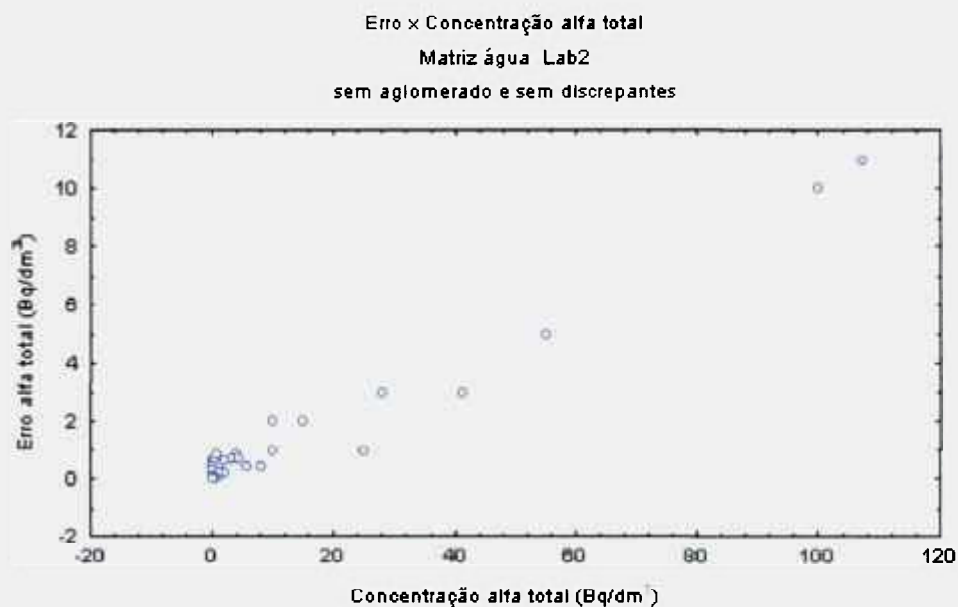


FIGURA 19 Erro x concentração: alfa total, matriz água, Lab 2; subconjunto focalizado

TABELA 10 Estatística descritiva: subconjunto de resultados alfa total, matriz água, Lab 2

Período de coleta : Jan/86 a Mar/02								
Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	158	2,97	0,15	-0,02	107	172	6,5	46
Erro publicado (Bq/dm <sup>3</sup> )	158	0,43	0,10	0,001	11	1,6	6,7	50

## A2. Análise de Urânio 238 - matriz água

A análise de Urânio 238 na matriz água foi realizada por um mesmo laboratório, pelo método de nêutrons retardados. Os resultados da concentração são publicados juntamente com um valor relativo à incerteza analítica, exceto para aqueles valores publicados como “menor que o limite de detecção” que são apresentados com o sinal “<” e sem especificação do erro analítico. O número de resultados produzidos para cada amostra é apresentado na FIGURA 20.

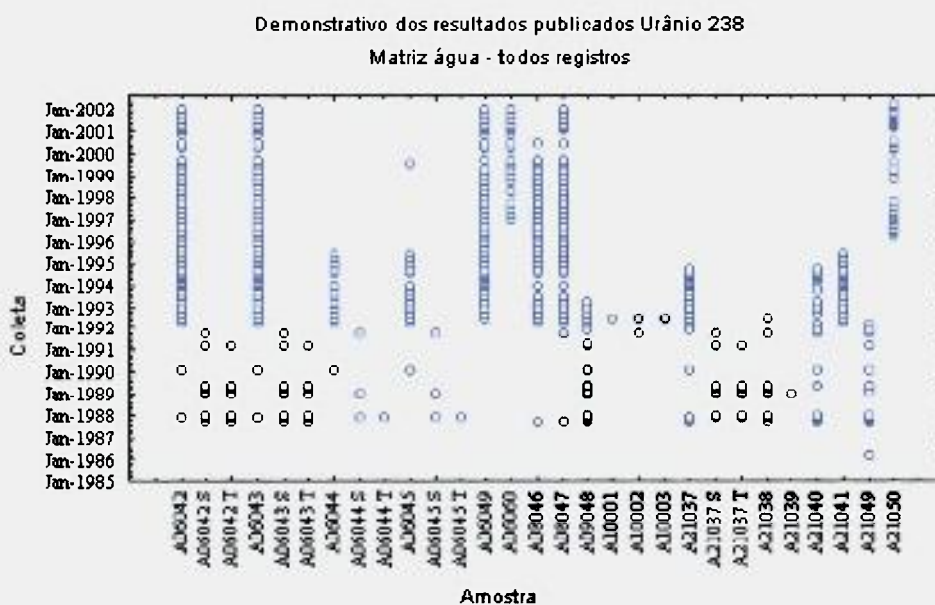


FIGURA 20 Demonstrativo de resultados publicados Urânio 238, matriz água

A caracterização deste conjunto por meio da estimação das estatísticas descritivas para a concentração e o erro publicado é apresentada na TABELA 11.

TABELA 11 Estatística descritiva para todos resultados, Urânio 238, matriz água

Período de coleta: 04/86 a 04/02								
Urânio 238	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	445	0,523	0,013	-8,000	14,00	7,9	3,1	12
Erro publicado (Bq/dm <sup>3</sup> )	384	0,267	0,010	0,001	8,00	1,1	5,2	28

A evolução temporal da concentração e do erro para todos os registros referentes à concentração de urânio 238 em amostras de água é apresentada na FIGURA 21 e na FIGURA 22.

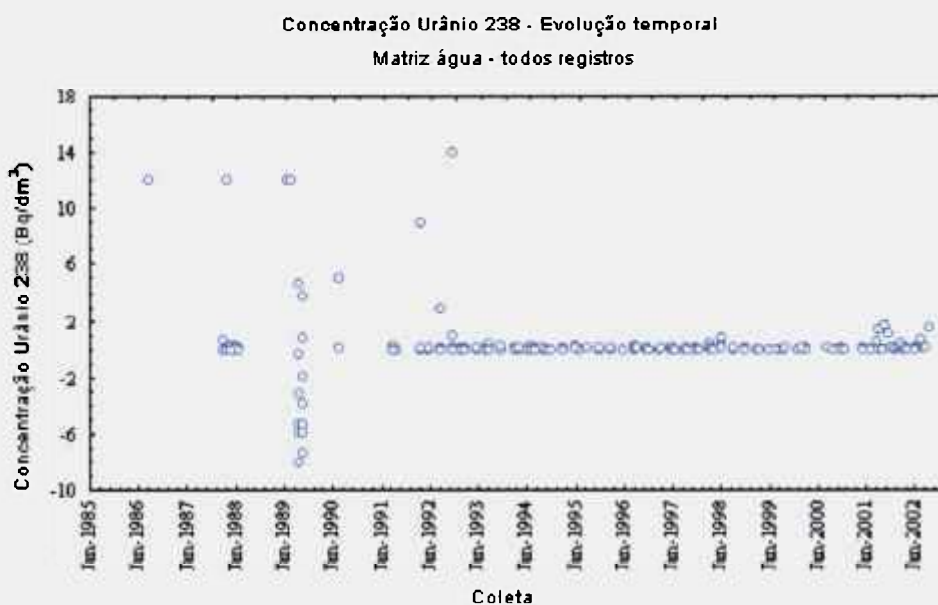


FIGURA 21 Evolução temporal da concentração Urânio 238, matriz água

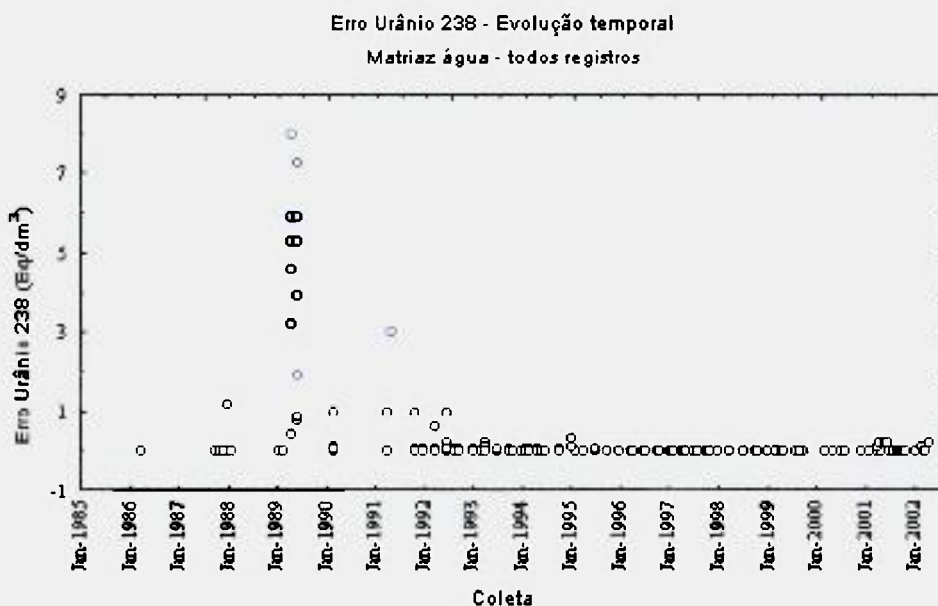


FIGURA 22 Evolução temporal do erro Urânio 238, matriz água

O comportamento do erro publicado e do erro relativo em relação à concentração, para todos registros e mostrado na FIGURA 23 e na FIGURA 24, indicam o uso de um modelo de erro relativo constante (igual a 1) para os valores de concentração negativa alcançando também a faixa positiva das concentrações.

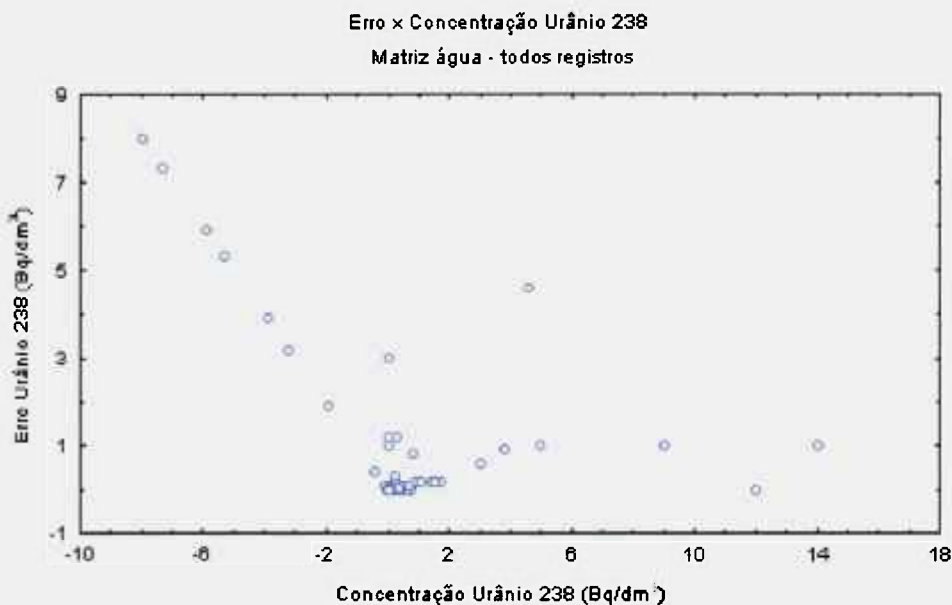


FIGURA 23 Comportamento do erro x concentração, Urânio 238, matriz água

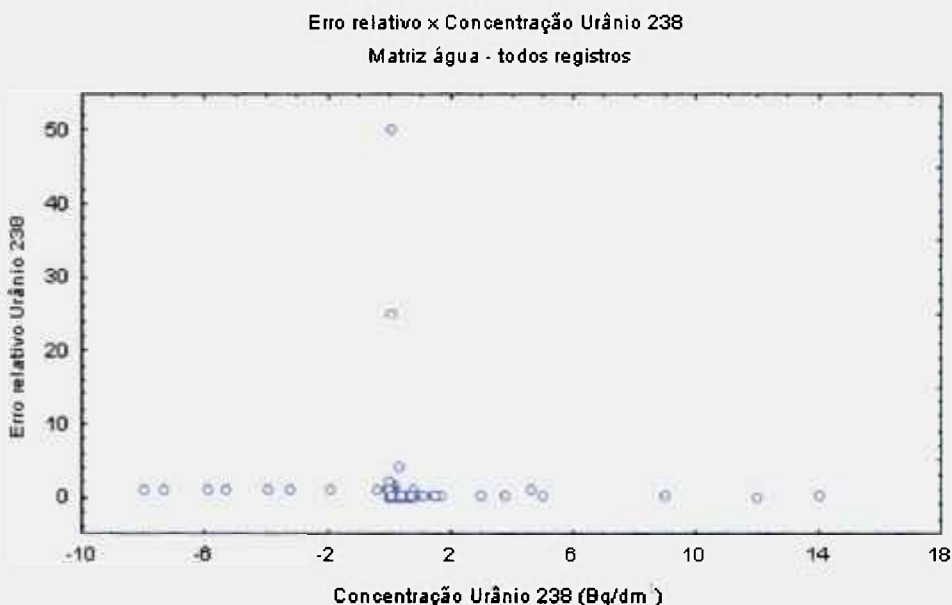


FIGURA 24 Comportamento do erro relativo x concentração, Urânio 238, matriz água

A região do gráfico onde se encontra o maior número de resultados, a faixa de concentração menor que 2, é mostra em detalhe na FIGURA 25. Nesta visualização pode - se observar uma série de valores de concentração onde o erro relativo é constante.

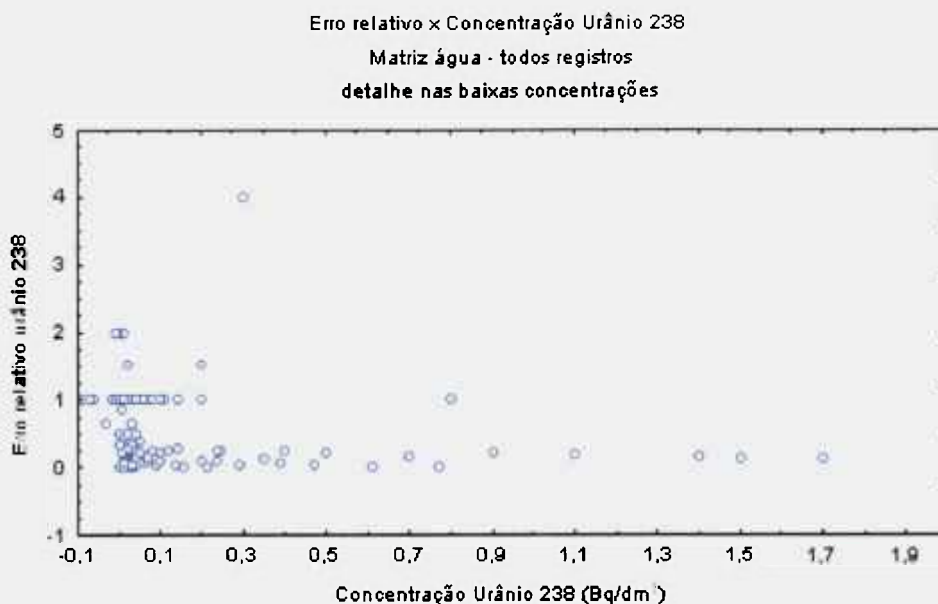


FIGURA 25 Detalhe: erro relativo x concentração, análise Urânio 238, matriz água

O estudo deste conjunto de resultados mostra um período inicial, até abril de 92, em que os resultados analíticos apresentam um padrão de comportamento muito diferente daquele apresentado no resto do período. Os resultados negativos com valores absolutos altos e os resultados censurados que constam do conjunto de resultados da análise de urânio 238 para a matriz água ocorreram todos neste período. Por isso ele foi excluído do subconjunto de resultados selecionado para a determinação dos parâmetros do modelo de dois componentes. O modelo de erro relativo constante foi utilizado na faixa de baixas concentrações (<0,2) intervalo em que ele não se aplica, por isso os registros referentes a este modelo foram também excluídos do subconjunto usado para estimar os parâmetros do modelo de dois componentes.

O resumo estatístico do conjunto de resultados a ser usado na estimação dos parâmetros do modelo dos dois componentes é apresentado na TABELA 12.

TABELA 12 Estatística descritiva para subconjunto resultados de urânio 238, matriz água

Período de coleta : Abr/92 a Abr/02								
Urânio 238	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (Bq/dm <sup>3</sup> )	117	0,148	0,030	-0,030	3,00	0,16	4,7	27
Erro publicado (Bq/dm <sup>3</sup> )	117	0,027	0,010	0,001	0,60	0,0046	6,0	45

### A3. Análise alfa total - matriz aerossol

As amostras da matriz aerossol foram coletadas em filtros de fibra de vidro, com periodicidade mensal, em três locais. O local 001 situa-se na direção de menor frequência de ventos e os locais codificados 002 e 003, na direção de maiores frequências de ventos. O protocolo analítico usado para a análise de alfa total e beta total prevê a retirada de uma alíquota do filtro amostrado, sua transferência direta para a bandeja de contagem feita em aço inox e contagem em detector proporcional. O número de resultados produzidos para cada amostra é apresentado na FIGURA 26.

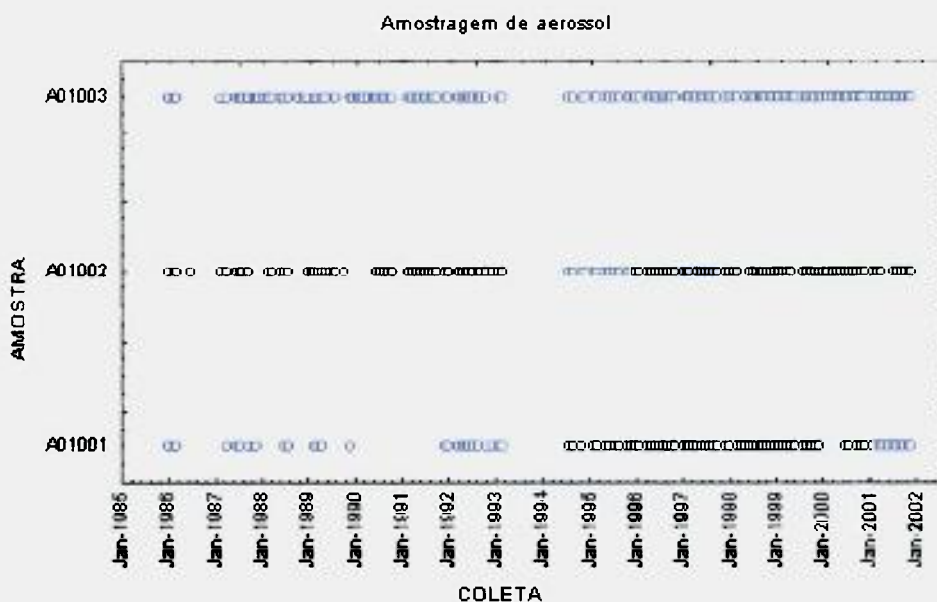


FIGURA 26 Demonstrativo de resultados produzidos alfa total, matriz aerossol

A TABELA 13 apresenta as estatísticas estimadas para todo o conjunto de resultados publicados para análise alfa total. Estas estimativas foram feitas para as concentrações e também para a incerteza publicada, que são as grandezas de interesse neste trabalho. Não há resultados apresentados como “menor que”; todos resultados apresentam uma incerteza associada. Quanto à coleta, houve interrupção da amostragem nos três locais no período de abril de 93 a julho de 94 e falhas ocasionais principalmente no local 001. O conjunto apresenta 355 resultados produzidos de janeiro de 1986 a dezembro de 2001 nos três locais de amostragem.

TABELA 13 Estatística descritiva para os resultados da análise alfa total, matriz aerossol

Período de coleta: Jan/86 a Dez/01								
Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Assimetria	Alongamento
Concentração (mBq/m <sup>3</sup> )	355	0,476	0,300	0,000	5,000	0,408	3,93	20,25
Erro publicado (mBq/m <sup>3</sup> )	355	0,199	0,100	0,001	5,633	0,241	7,39	64,80

A evolução temporal da concentração e do erro é visualizada na FIGURA 27 e na FIGURA 28. A evolução temporal do erro deveria acompanhar a evolução da concentração alfa total, porém aparecem regiões de maior dispersão no comportamento da concentração que não são observadas no comportamento da grandeza erro.

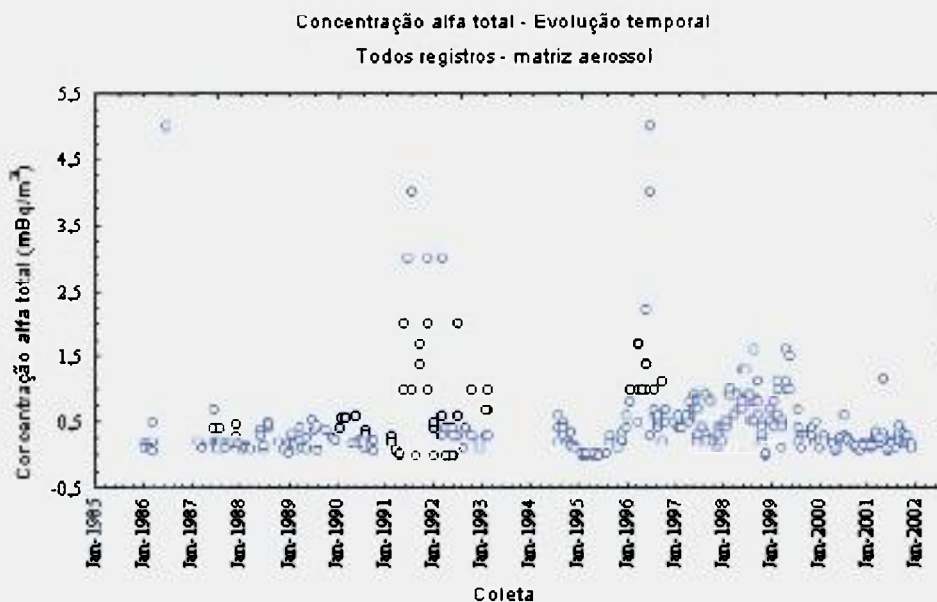


FIGURA 27 Evolução temporal: concentração alfa total, matriz aerossol

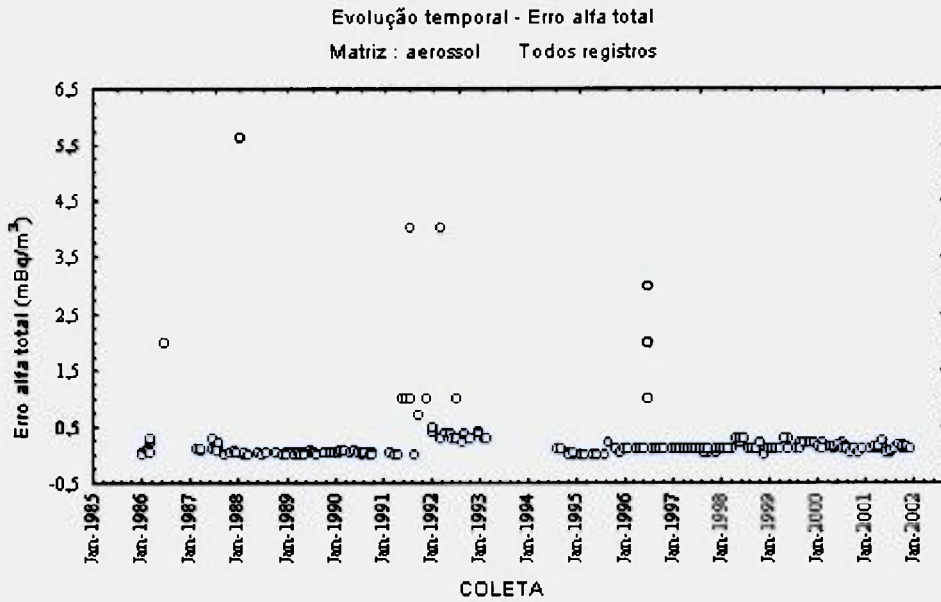


FIGURA 28 Evolução temporal: erro alfa total, matriz aerossol

O comportamento do erro em função da concentração, mostrado na FIGURA 29, fornece indicação do modelo de medição usado no protocolo analítico adotado pelo laboratório. Na FIGURA 30 é mostrado o comportamento do erro relativo em função da concentração.

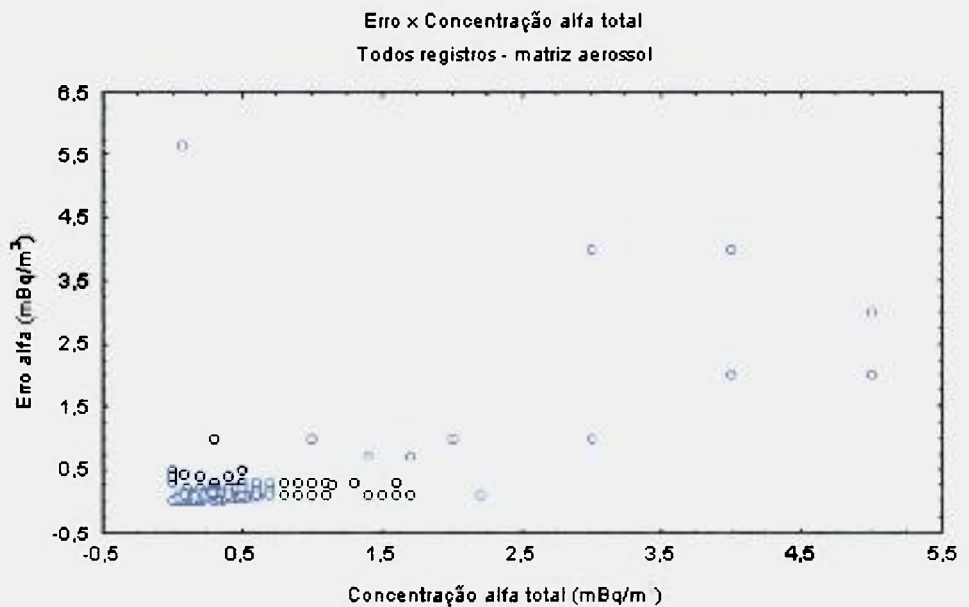


FIGURA 29 Comportamento do erro x concentração alfa total , matriz aerossol

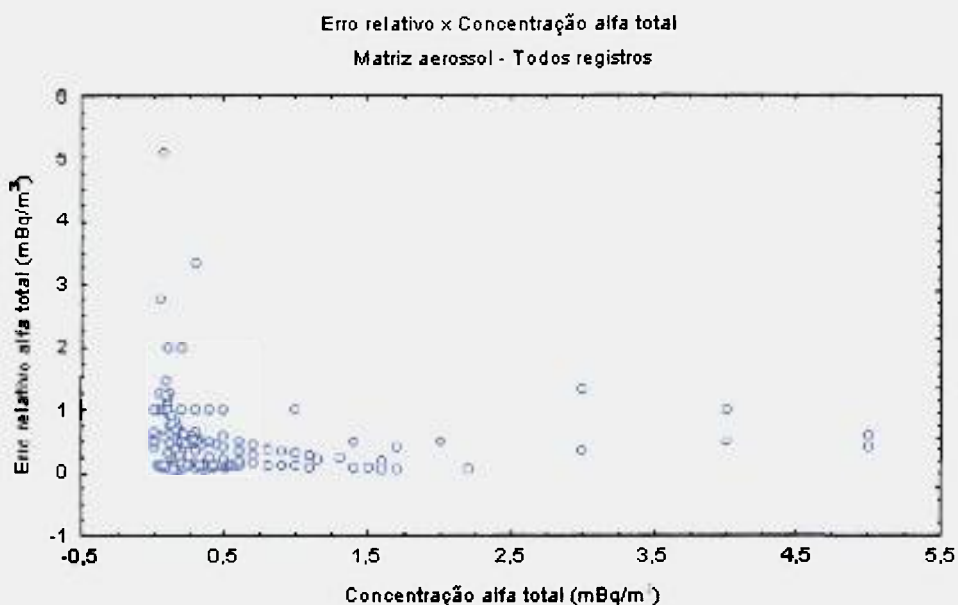


FIGURA 30 Comportamento do erro relativo x concentração alfa total, matriz aerossol

Uma observação mais detalhada do comportamento do erro associado a cada resultado analítico apresenta padrões de comportamento diferentes, sugerindo o uso de modelos de medição distintos. Na FIGURA 31 podem ser observados três modelos de medição: erro constante (para o valor 0,1 ; 0,2 e 0,3); erro diretamente proporcional à concentração (com intersecção em 0,0 e indicado no gráfico como modelo  $y=kx$ ;) e erro com proporcionalidade à concentração numa forma que sugere raiz quadrada (indicado no gráfico como modelo Poisson).

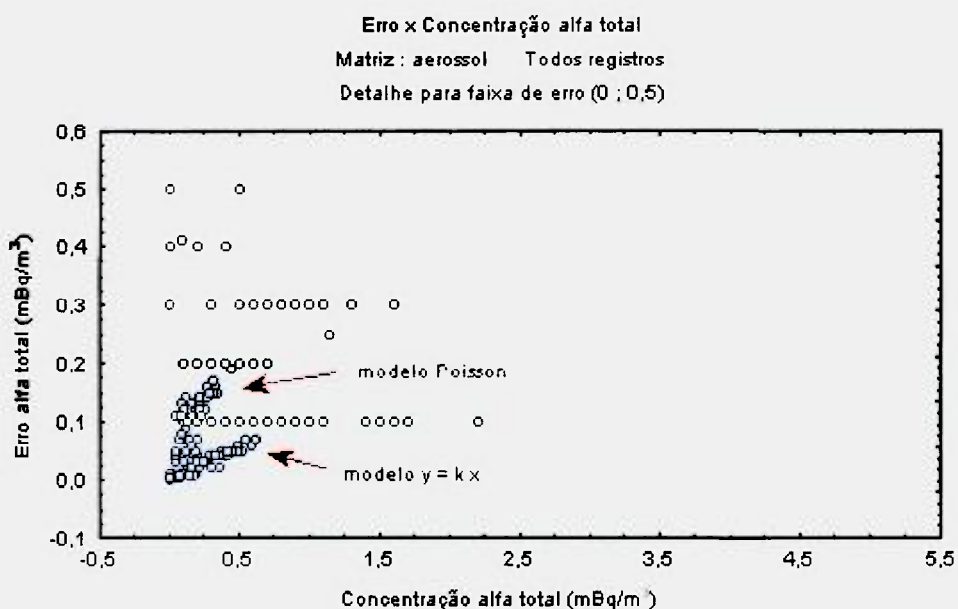


FIGURA 31 Detalhe do comportamento do erro alfa x concentração, matriz aerossol

Na FIGURA 32 é apresentada uma ampliação do gráfico que mostra a evolução temporal do erro publicado em função da concentração medida para os valores do erro publicado na faixa de 0 a 0,5 e que permite determinar o período de tempo em que cada modelo foi utilizado. O período entre nov/95 e set/99, onde o gráfico mostra uma linha paralela ao eixo das concentrações, corresponde ao uso do modelo de erro constante.

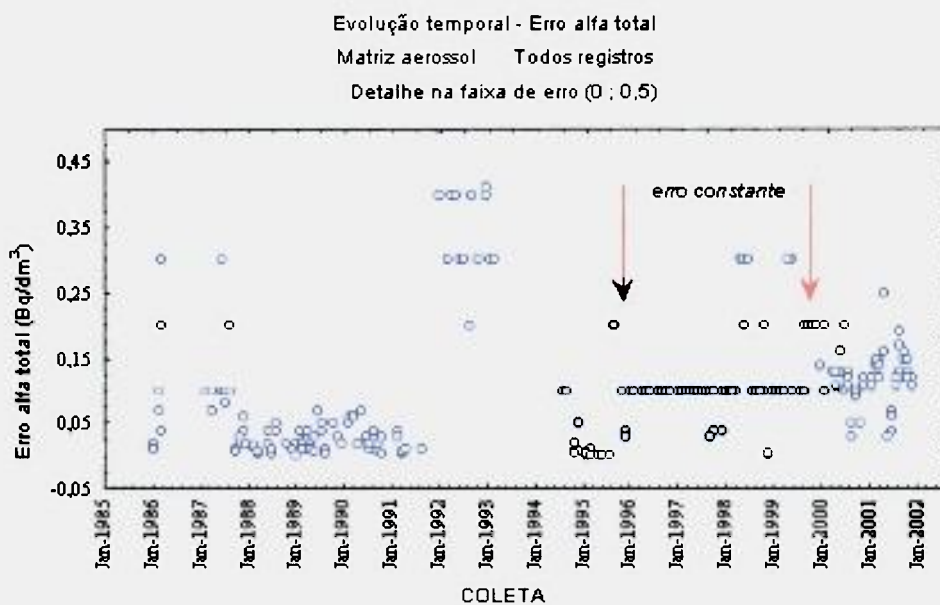


FIGURA 32 Detalhe da evolução temporal do erro alfa total, matriz aerossol

O modelo em que o erro é proporcional à concentração, utilizado no período de dezembro de 87 a março de 91, é ilustrado na FIGURA 33.

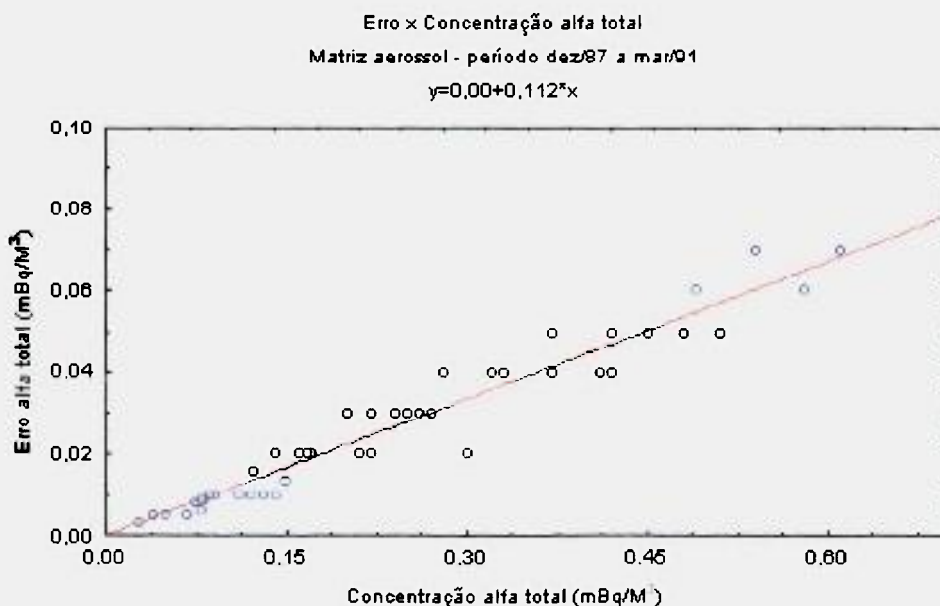


FIGURA 33 Erro x concentração alfa total, matriz aerossol; período dez87 a mar91

O gráfico que mostra o comportamento do erro para o período pleno de resultados (janeiro de 86 a dezembro de 2001) excluindo os dois períodos citados anteriormente e que corresponde a um modelo aparentemente baseado na estatística de Poisson, é apresentado na FIGURA 34.

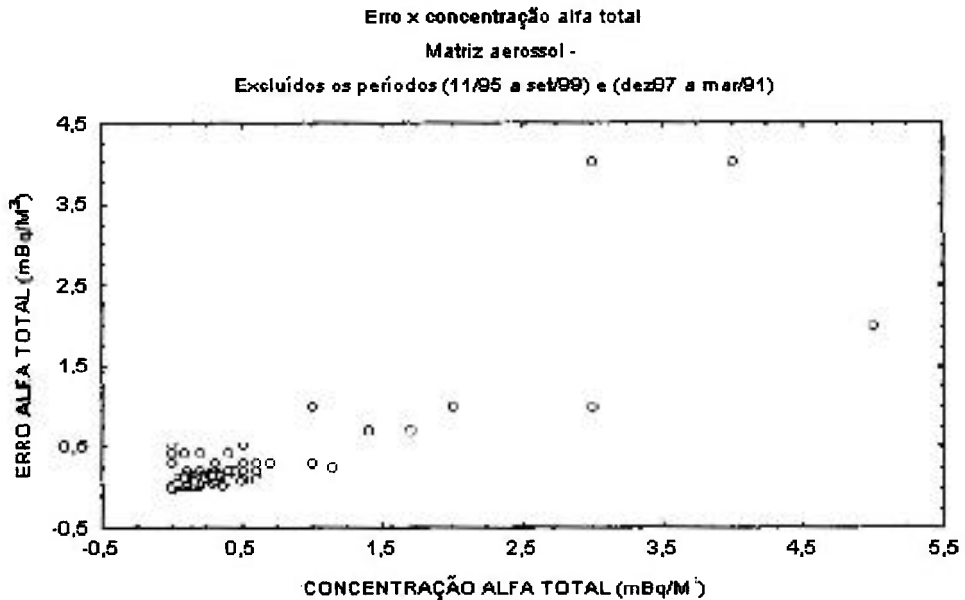


FIGURA 34 Erro x concentração alfa total, matriz aerossol (modelo Poisson)

Os resultados referentes aos períodos em que o erro é constante ou diretamente proporcional à concentração não serão usados na determinação dos parâmetros do modelo de Roche & Lorenzato. O erro constante para todas as concentrações implicaria no parâmetro  $\sigma_{\eta}$  igual à zero e o erro diretamente proporcional acarretaria  $\sigma_{\epsilon}$  nulo. Estes resultados portanto deverão ser excluídos do conjunto a ser usado na determinação dos parâmetros do modelo de R&L.

O conjunto de resultados usados no ajuste para determinação dos parâmetros do modelo de dois componentes é apresentado na TABELA 14.

TABELA 14 Estatística descritiva: subconjunto de resultados alfa total, matriz aerossol

Alfa total	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração (mBq/m <sup>3</sup> )	108	0,404	0,205	0,002	3,00	0,353	0,057
Erro publicado (mBq/m <sup>3</sup> )	108	0,182	0,120	0,001	1,00	0,052	0,022

## APÊNDICE B - RESUMO ESTATÍSTICO E RESULTADOS POR AMOSTRA

Neste Apêndice as amostras cujos resultados analíticos foram focalizados neste estudo são descritas individualmente. São apresentados para cada local de amostragem:

- o resumo estatístico,
- a visualização da evolução temporal da concentração e do erro analítico publicado,
- a visualização do comportamento do erro publicado e do erro previsto pelo modelo R&L em relação à concentração medida,
- as médias estimadas pelo inverso do quadrado erro publicado e pelo inverso do quadrado do erro previsto pelo modelo R&L.

A descrição individual está agrupada de acordo com o protocolo analítico que produziu os resultados. Os casos estudados e as amostras descritas estão relacionadas na TABELA 1.

TABELA 1 Amostras focalizadas

Caso	Análise	Amostras
1	Análise alfa total - matriz água, resultados laboratório 1	Água de superfície: 06042, 06043, local 049 Água subterrânea: 08046, 08047 Água potável: 09048 Água de chuva: 10001, 10002, 10003
2	Análise alfa total - matriz água, resultados laboratório 2	Esgoto industrial: 21037, 21038, 21040, 21041, 21050
3	Análise Urânio 238 - matriz água	Água de superfície: 06042, 06043, 06049 Água subterrânea: 08046, 08047 Esgoto industrial: 21050
4	Análise alfa total - matriz aerossol.	Aerossol: 01001, 01002, 01003

## B1. Análise alfa total - matriz água, resultados laboratório 1

### B1.1. Local 042 – água de superfície

TABELA 2 Estatística descritiva: concentração alfa total ( $\text{Bq}/\text{dm}^3$ ); amostra: 06042

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	130	0,063	0,030	-0,03	1,1	0,026	0,014
Erro publicado	130	0,072	0,040	0,01	2,0	0,061	0,022

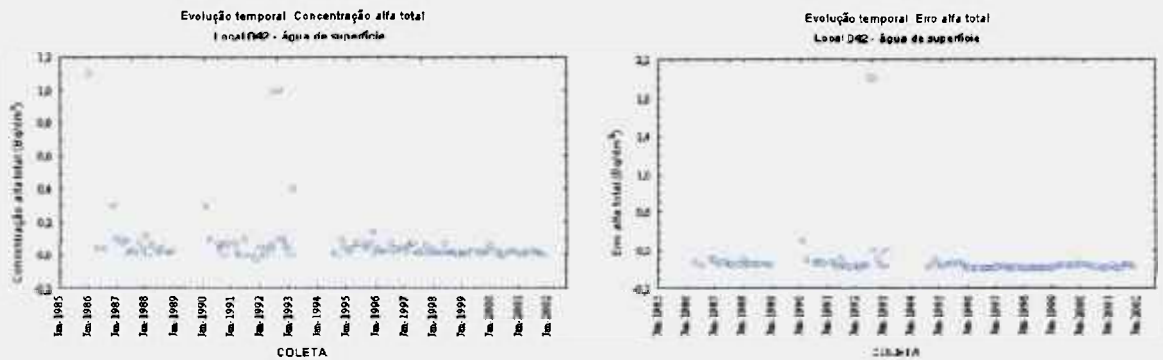


FIGURA 1 Evolução temporal da concentração e do erro alfa total; amostra 06042

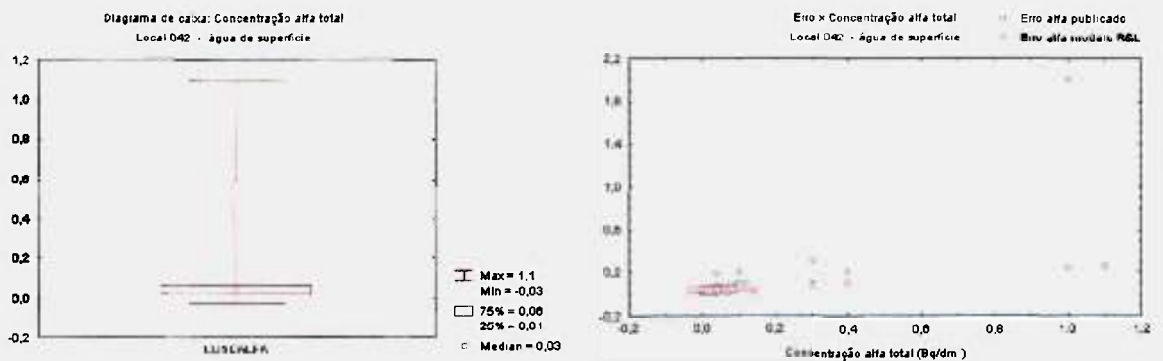


FIGURA 2 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 3 Comparação das médias: concentração alfa total ( $\text{Bq}/\text{dm}^3$ ), amostra: 06042

N	Média simples	Média ponderada erro publicado	Média ponderada erro modelo	$R_{MP}$ =Relação modelo/publicado
130	0,063	0,0228	0,0329	1,44

**B1.2. Local 043 –água de superfície**

TABELA 4 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 06043

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	128	0,0459	0,020	-0,02	1,0	0,011	0,0094
Erro publicado	128	0,039	0,030	0,004	0,2	0,001	0,003

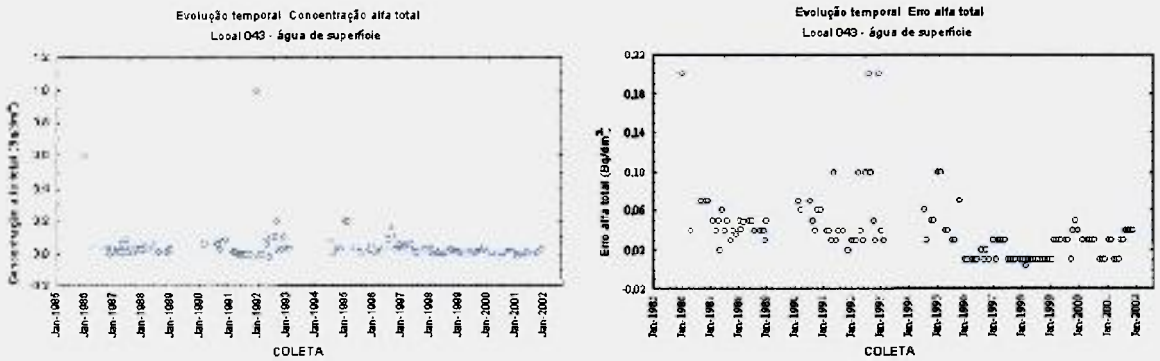


FIGURA 3 Evolução temporal da concentração e do erro alfa total; amostra 06043

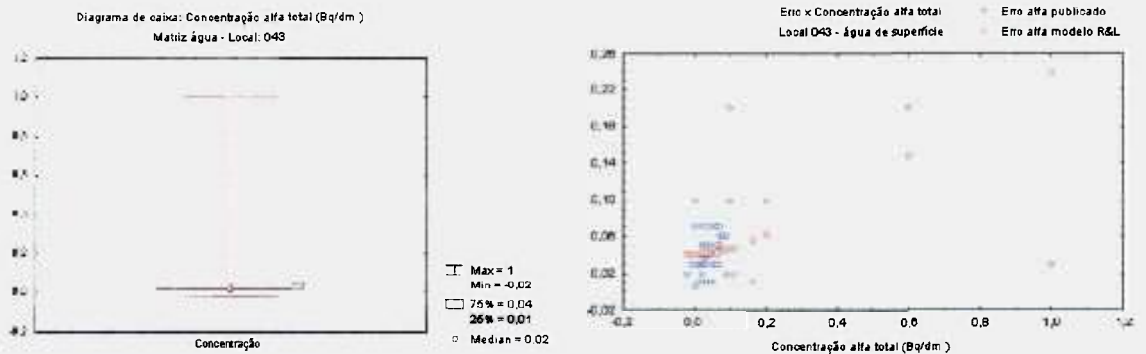


FIGURA 4 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 5 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>), amostra: 006043

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
128	0,0459	0,0253	0,0301	1,19

### B1.3. Local 049 – água de superfície

(Obs: Classificada como esgoto industrial até junho 1992, após esta data água de superfície)

TABELA 6 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 06049 e 21049

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	149	0,25	0,030	-0,6	25	4,2	0,17
Erro publicado	149	0,12	0,040	0,002	6	0,27	0,04

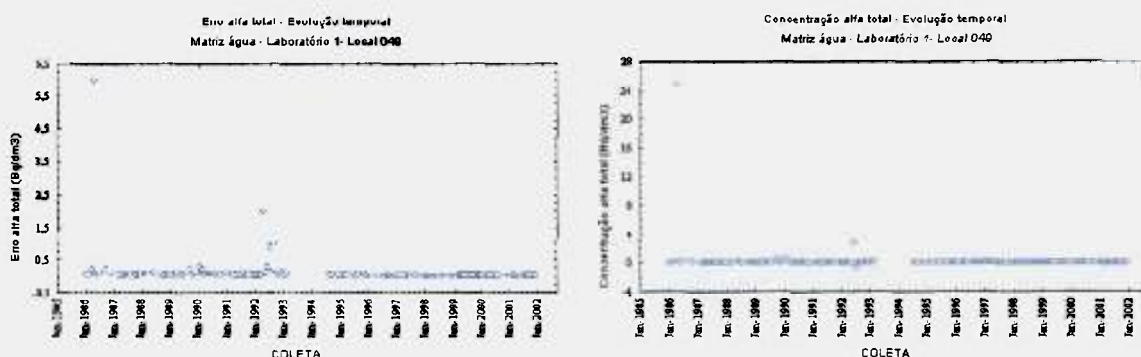


FIGURA 5 Evolução temporal da concentração e do erro alfa total; local 06049

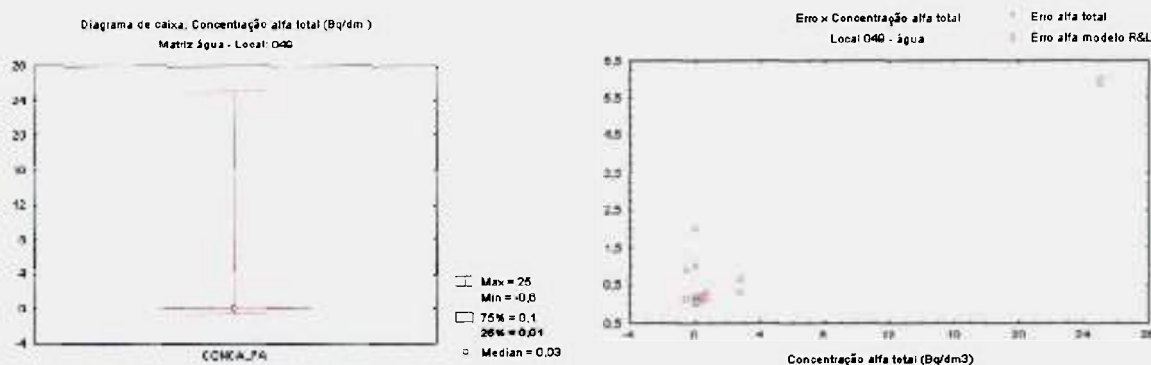


FIGURA 6 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 7 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>), amostra: local 049

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
149	0,25	0,0145	0,0427	2,94

**B1.4. Local 046 –água subterrânea**

TABELA 8 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 08046

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	28	0,067	0,040	-0,01	0,56	0,012	0,020
Erro publicado	28	0,032	0,030	0,01	0,1	0,00054	0,0044

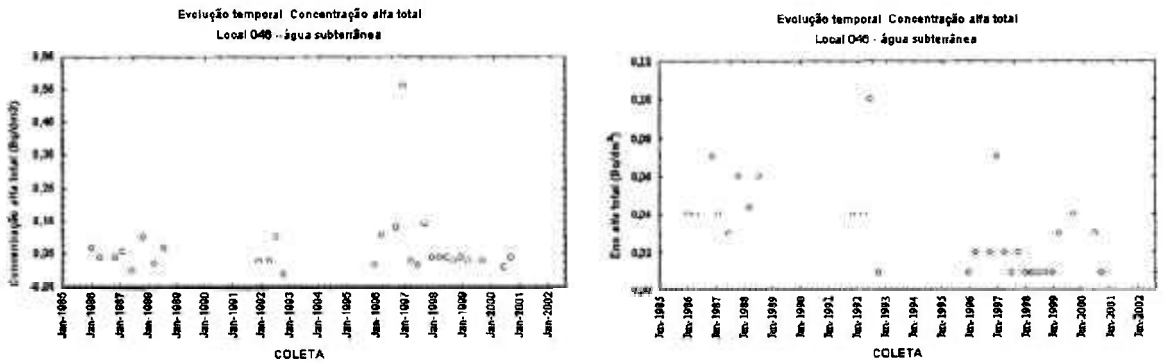


FIGURA 7 Evolução temporal da concentração e do erro alfa total; amostra 08046

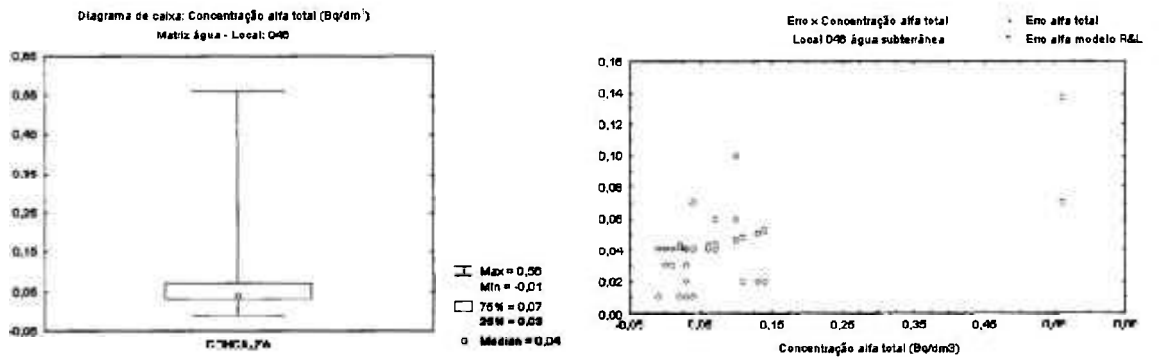


FIGURA 8 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 9 Comparação das médias : concentração alfa total (Bq/dm<sup>3</sup>), amostra: 08046

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
28	0,067	0,0370	0,0456	1,23

**B1.5. Local 047 –água subterrânea**

TABELA 10 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 08047

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	32	0,0486	0,040	-0,01	0,22	0,0019	0,0077
Erro publicado	32	0,038	0,030	0,01	0,10	0,0009	0,0054

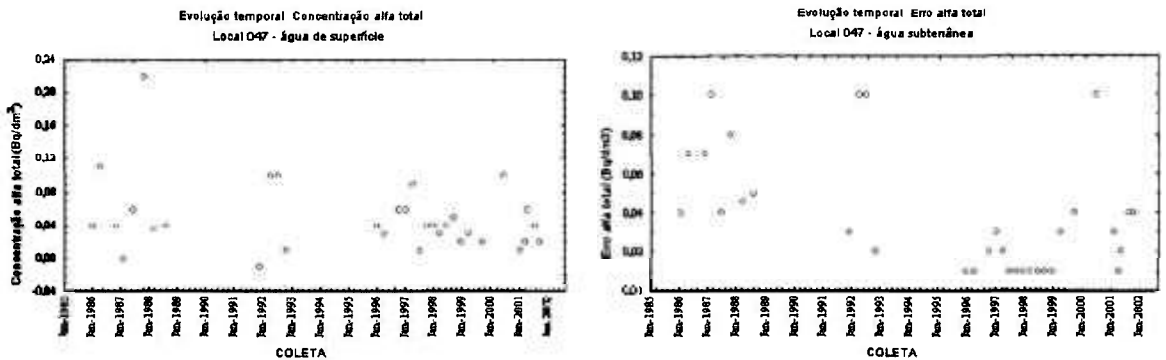


FIGURA 9 Evolução temporal da concentração e do erro alfa total; amostra 08047

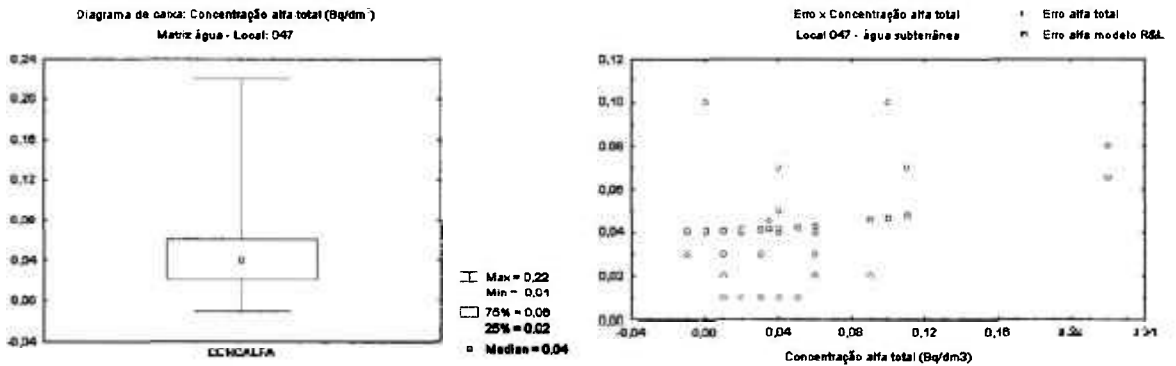


FIGURA 10 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 11 Comparação das médias : concentração alfa total (Bq/dm<sup>3</sup>); amostra: 08047

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
32	0,0486	0,0343	0,0428	1,25

### B1.6. Local 048 –água potável

TABELA 12 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>), amostra: 09048

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	76	0,043	0,029	-0,01	0,70	0,0078	0,010
Erro publicado	76	0,048	0,040	0,02	0,20	0,0011	0,004

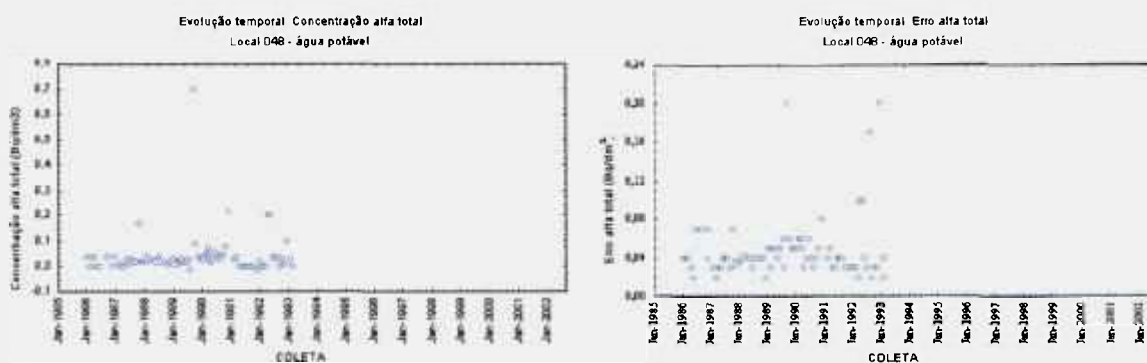


FIGURA 11 Evolução temporal da concentração e do erro alfa total; amostra 09048

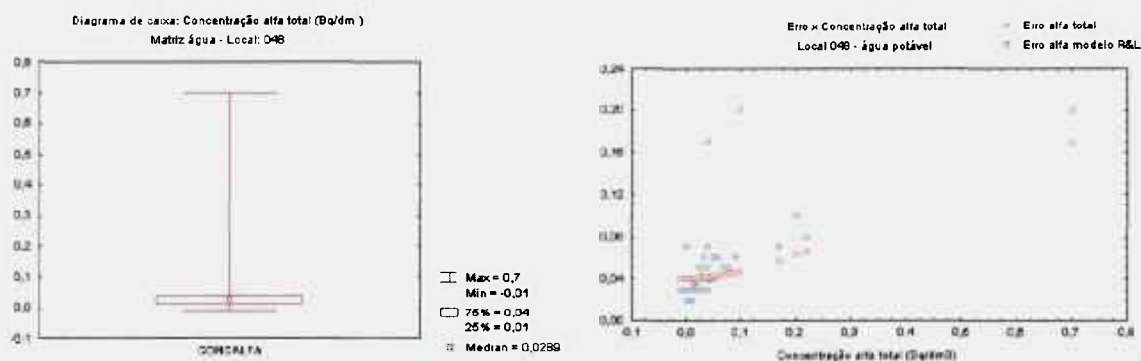


FIGURA 12 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 13 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 09048

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
76	0,043	0,0180	0,0283	1,57

**B1.7. Local 001 –água de chuva**

TABELA 14 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>), amostra: 10001

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	49	0,046	0,030	0,00	0,50	0,006	0,011
Erro publicado	49	0,045	0,040	0,02	0,19	0,0007	0,004

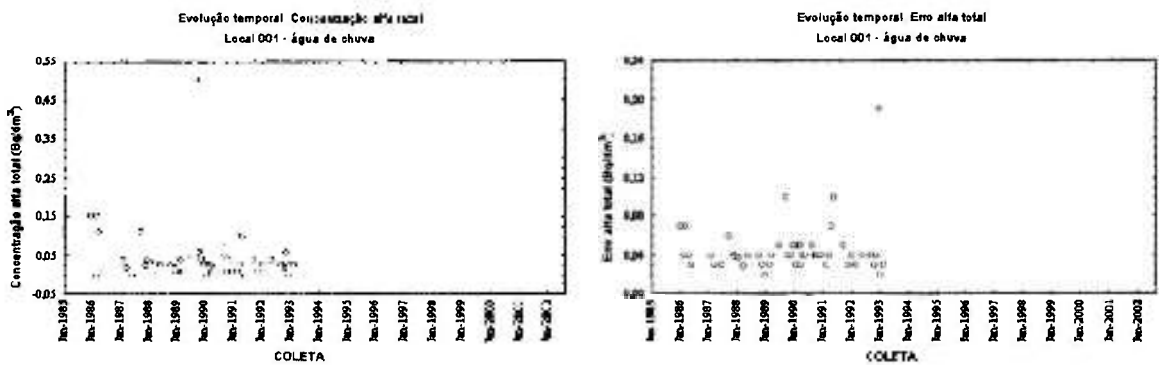


FIGURA 13 Evolução temporal da concentração e do erro alfa total; amostra 10001

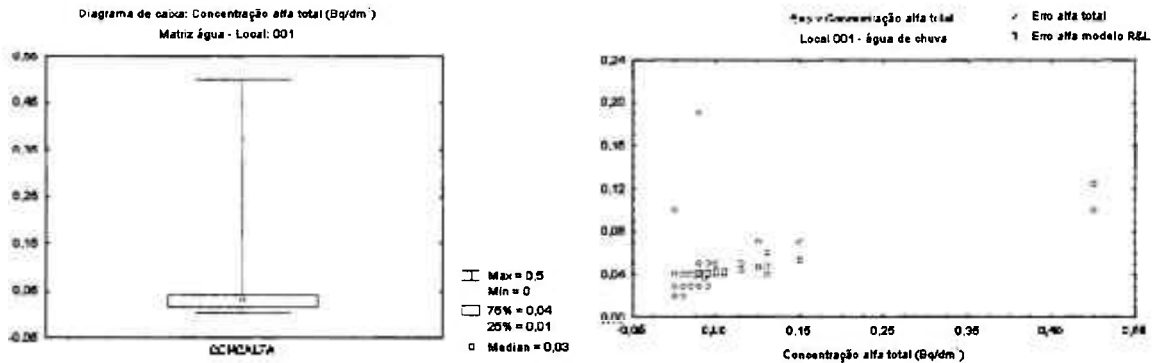


FIGURA 14 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 15 Comparação das médias : concentração alfa total (Bq/dm<sup>3</sup>); amostra: 10001

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
49	0,046	0,0265	0,0333	1,26

### B1.8. Local 002 –água de chuva

TABELA 16 Estatística descritiva: concentração alfa total ( $Bq/dm^3$ ); amostra: 010002

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	43	0,0370	0,030	-0,01	0,11	0,0011	0,0050
Erro publicado	43	0,043	0,040	0,02	0,20	0,0007	0,0041

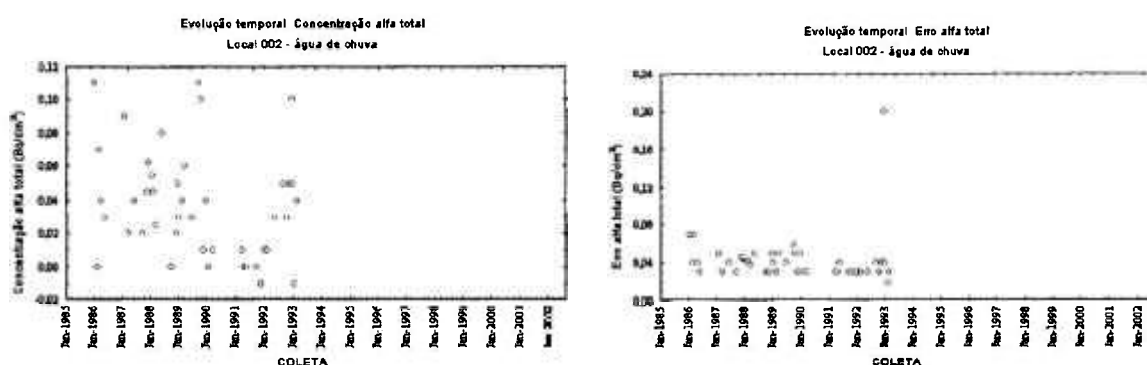


FIGURA 15 Evolução temporal da concentração e do erro alfa total; amostra 10002

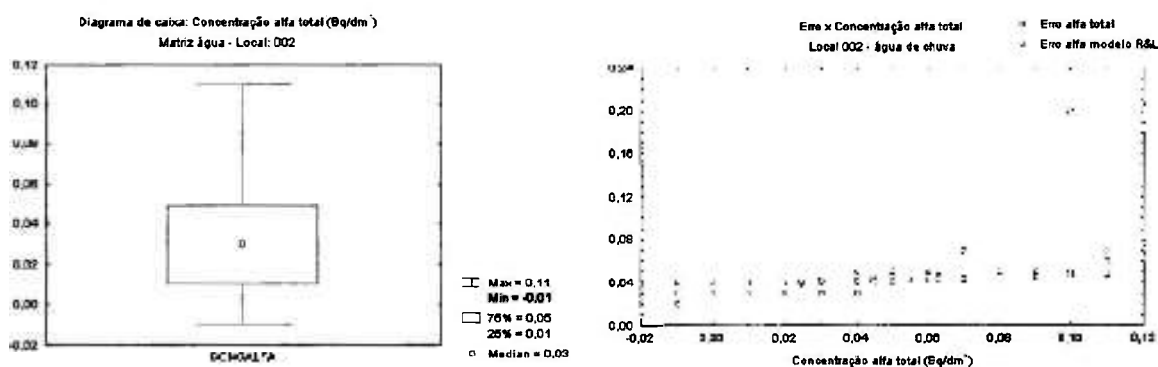


FIGURA 16 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 17 Comparação das médias: concentração alfa total ( $Bq/dm^3$ ); amostra: 10002

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	$R_{MP}$ =Relação modelo/publicado
43	0,0370	0,0261	0,0343	1,31

**B1.9. Local 003 –água de chuva**

TABELA 18 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 10003

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	50	0,0364	0,030	-0,02	0,24	0,0020	0,0063
Erro publicado	50	0,045	0,040	0,024	0,20	0,00077	0,0040

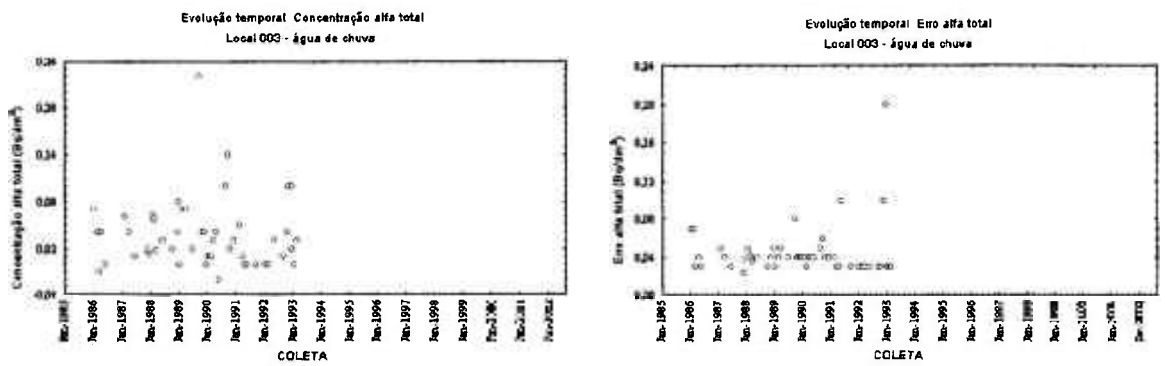


FIGURA 17 Evolução temporal da concentração e do erro alfa total; amostra 10003

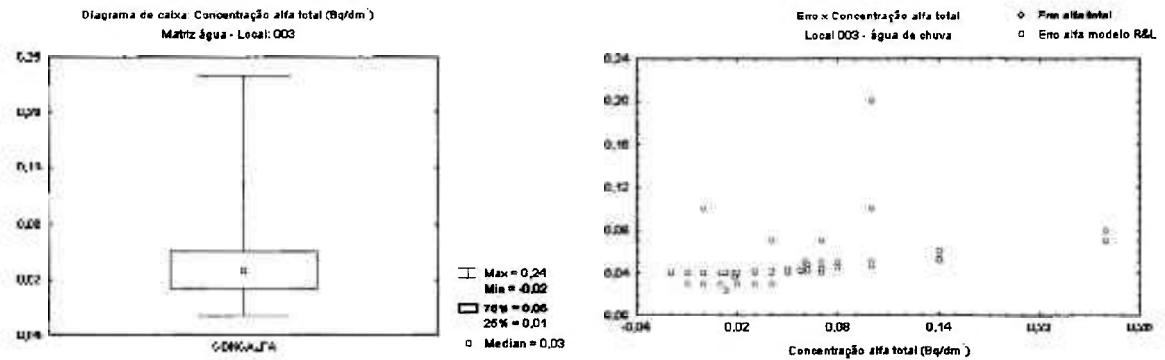


FIGURA 18 Diagrama de caixa e comparação erro publicado e erro modelo R&L, alfa total

TABELA 19 Comparação das médias; : concentração alfa total (Bq/dm<sup>3</sup>); amostra: 10003

N	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
50	0,0364	0,0226	0,0310	1,37

## B2. Análise alfa total - matriz água, resultados laboratório 2

### B2.1. Amostra 021037 –esgoto industrial

TABELA 20 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21037

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	38	4,0	0,25	-18	41	105	1,7
Erro publicado	37	2,8	0,30	0,004	17	23	0,8

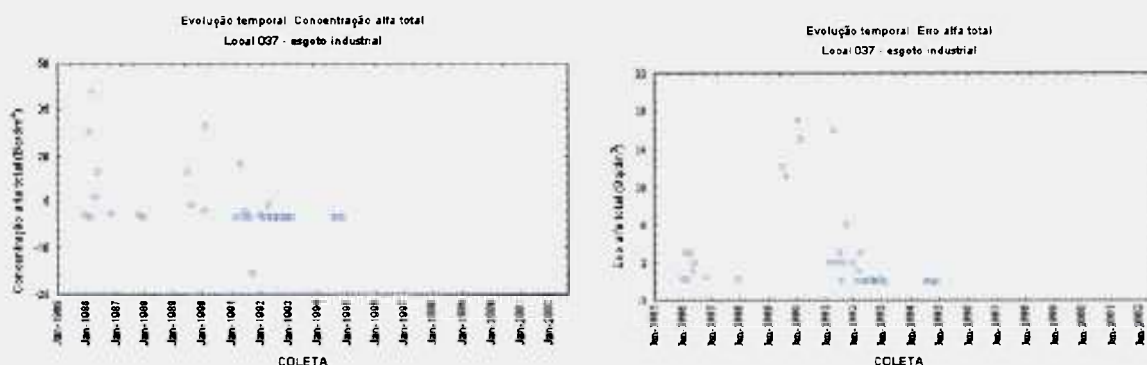


FIGURA 19 Evolução temporal da concentração e do erro alfa total; amostra 21037

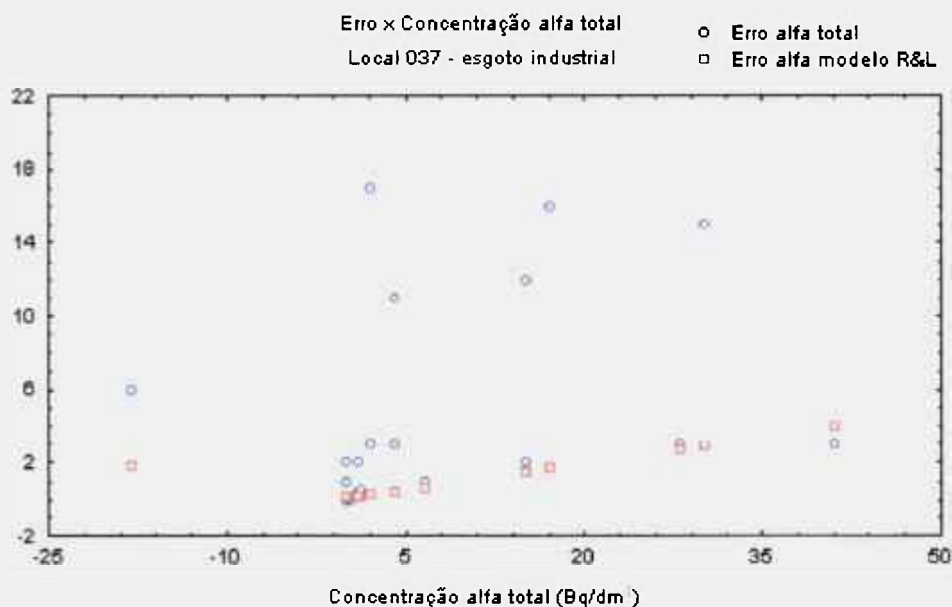


FIGURA 20 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 21 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21037

Amostra: A21037	NT	N	Graucen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
Exclui ND		37	2,6	4,1	0,0485	0,348	7,18
Subst=ND	38			4,0	0,0485	0,367	7,57

**B2.2. Amostra 021038 –esgoto industrial**

TABELA 22 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21038

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	29	11,1	3,0	0,13	100	533	4,3
Erro publicado	23	3,8	1,0	0,2	18	31	1,2

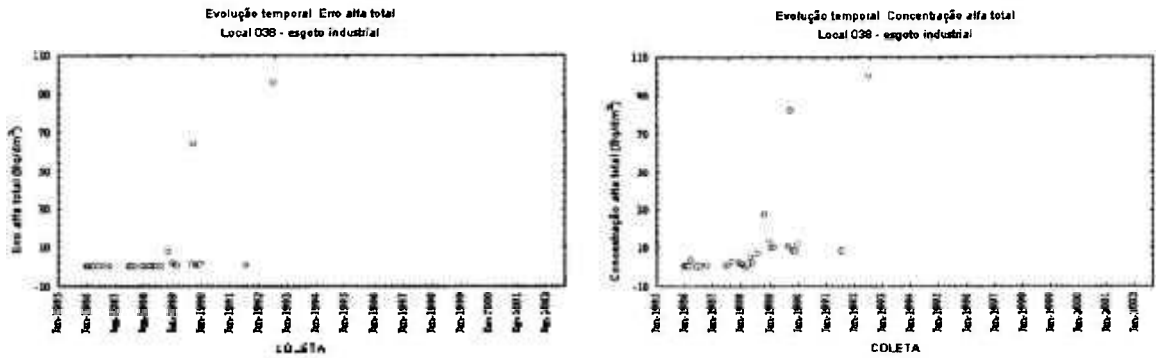


FIGURA 21 Evolução temporal da concentração e do erro alfa total; amostra 21038

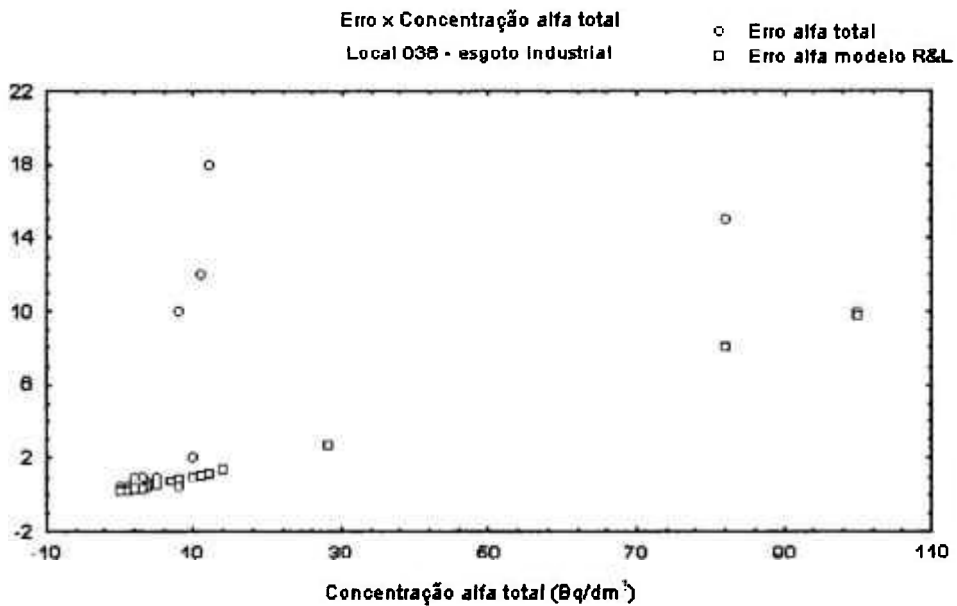


FIGURA 22 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 23 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>), amostra: 21038

Amostra: A21038	NT	N	Graucen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
Exclui ND		23	20,7	11,8	1,20	1,24	1,03
Subst=ND	29			11,1	1,20	1,03	0,87

### B2.3. Amostra 021040 –esgoto industrial

TABELA 24 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21040

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	47	-0,07	0,15	-18	5	8,2	0,42
Erro	35	5,4	0,30	0,01	22	56	1,3

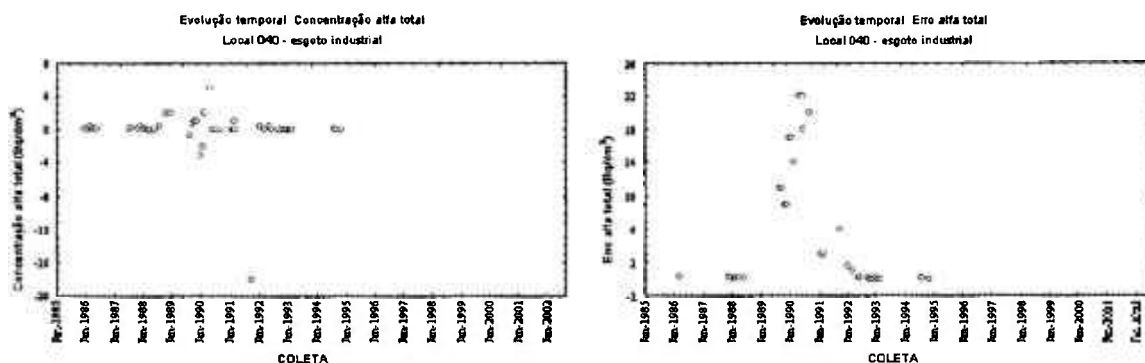


FIGURA 23 Evolução temporal da concentração e do erro alfa total; amostra 21040

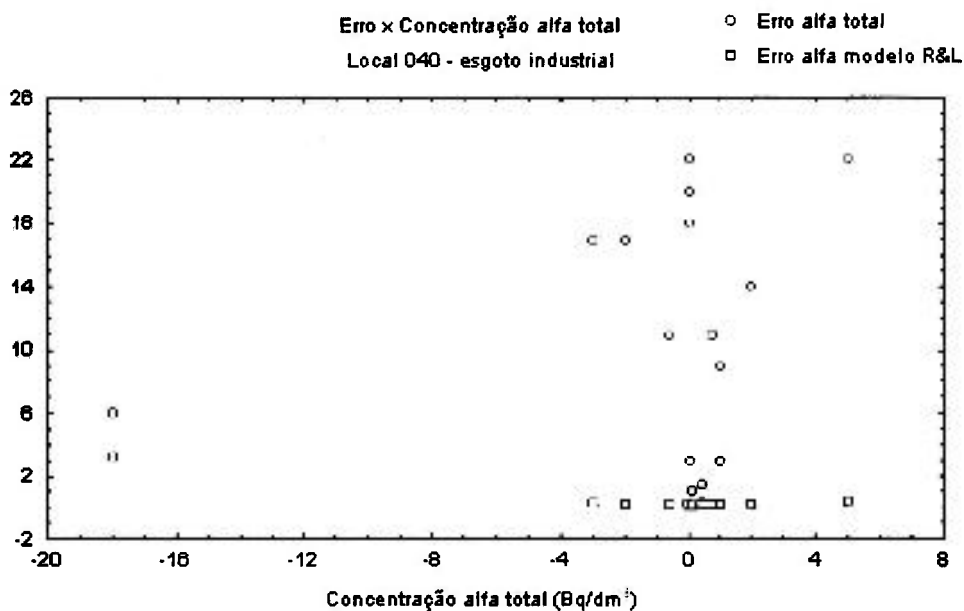


FIGURA 24 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 25 Comparação das médias : concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21040

Amostra: A21040	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>M</sub> P=Relação modelo/publica do
Exclui ND		35	25,5	-0,28	0,0096	0,170	17,7
Subst=ND	47			-0,07	0,0096	0,221	23,0

**B2.4. Amostra 021041 –esgoto industrial**

TABELA 26 Estatística descritiva: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21041

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	26	2,4	0,35	-0,01	25	28	1,0
Erro publicado	23	0,52	0,20	0,02	3	0,5	0,15

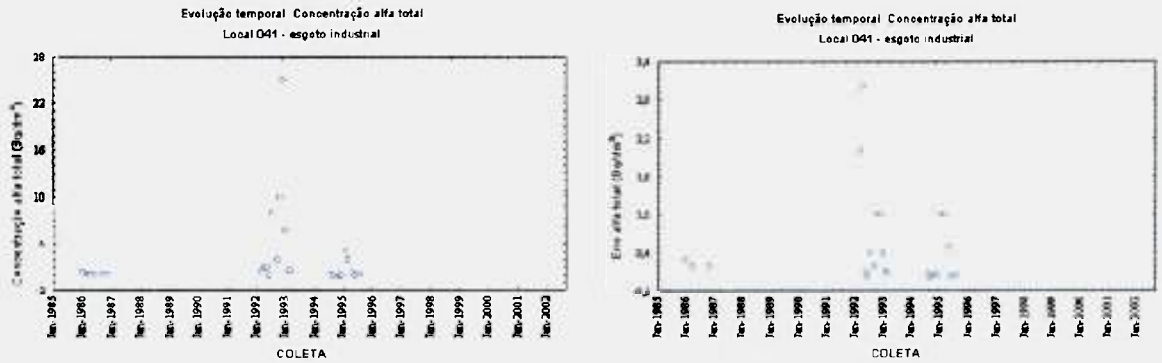


FIGURA 25 Evolução temporal da concentração e do erro alfa total; amostra 21041

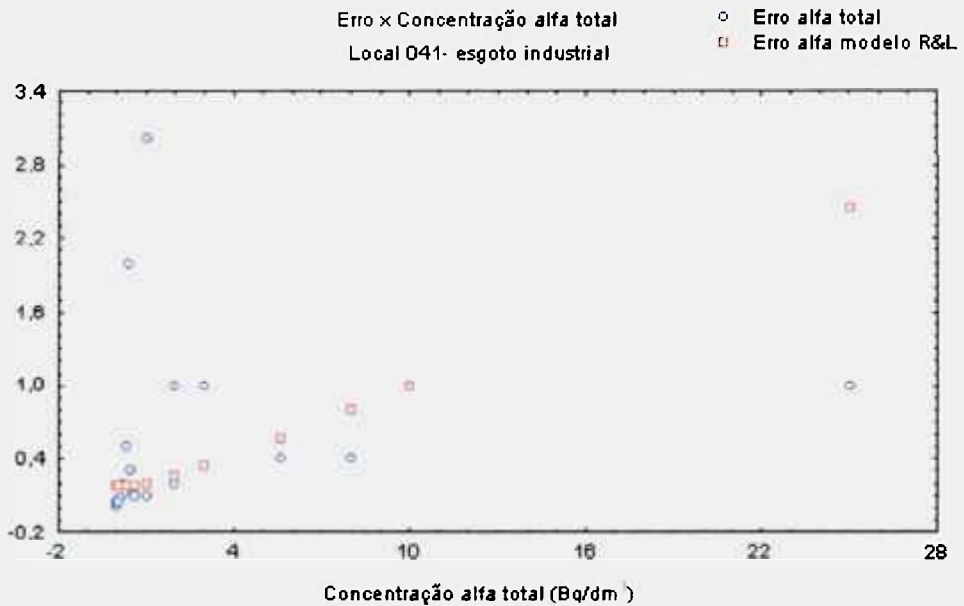


FIGURA 26 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 27 Comparação das médias: concentração alfa total (Bq/dm<sup>3</sup>); amostra: 21041

Amostra: A21041	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publi cado
Exclui ND		23	11,5	2,6	0,0892	0,498	5,59
Subst=ND	26			2,4	0,0892	0,453	5,08

### B2.5. Amostra 021050 –esgoto industrial

TABELA 28 Estatística descritiva: concentração alfa total ( $Bq/dm^3$ ); amostra: 21050

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	80	0,277	0,10	-0,02	5	0,39	0,070
Erro publicado	80	0,26	0,10	0,00	2	0,12	0,039

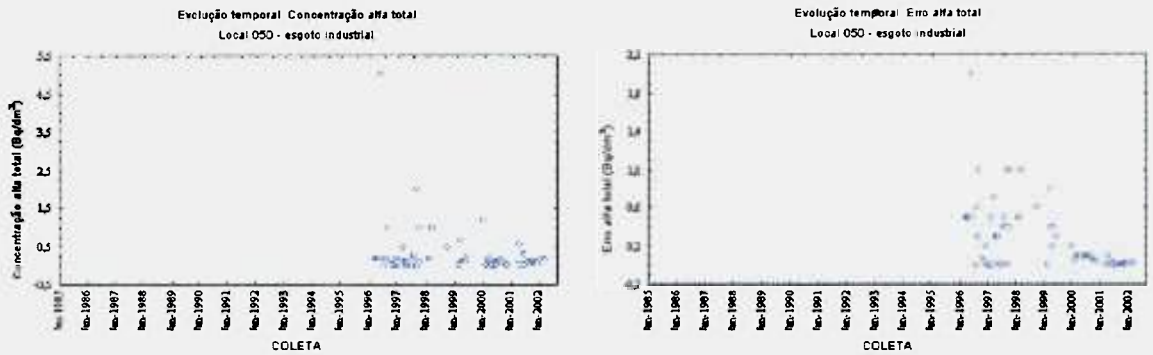


FIGURA 27 Evolução temporal da concentração e do erro alfa total; amostra 21050

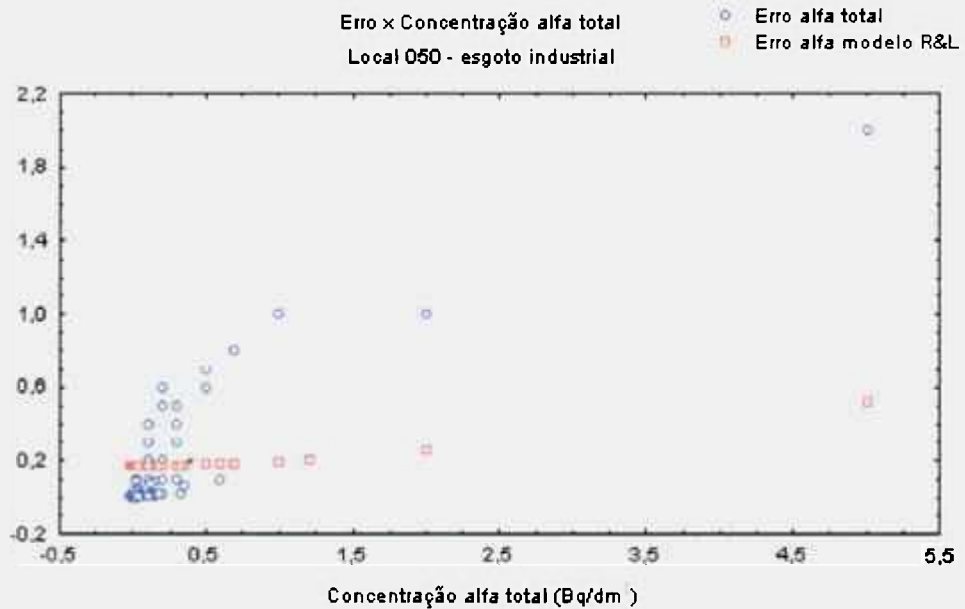


FIGURA 28 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 29 Comparação das médias: concentração alfa total ( $Bq/dm^3$ ); amostra: 21050

Amostra:	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	$R_{MP}$ =Relação modelo/publicado
A21041	80	80		0,277	0,00562	0,195	34,7

**B3. Análise Urânio 238- matriz água**

**B3.1. Amostra 006042 –água de superfície**

TABELA 30 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06042

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	46	0,044	0,0200	-0,02	0,47	0,0072	0,012
Erro publicado	45	0,024	0,0100	0,001	0,30	0,0024	0,0072

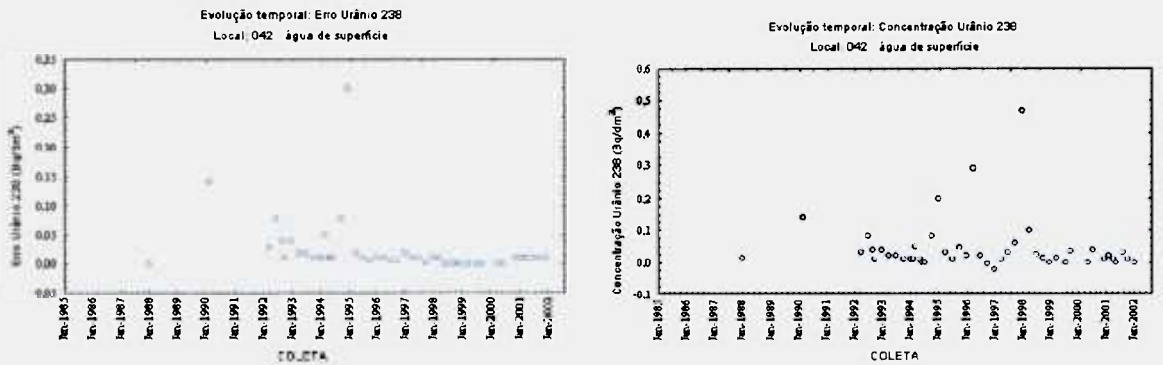


FIGURA 29 Evolução temporal da concentração e do erro Urânio 238; amostra 06042

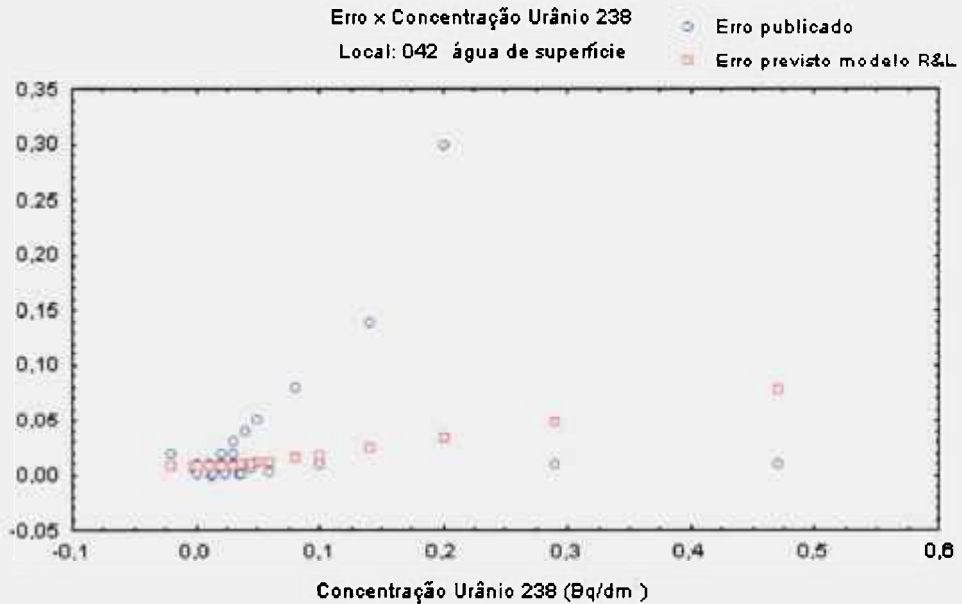


FIGURA 30 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 31 Comparação das médias: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06042

Amostra:	NT	N	Graucen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A21037		45	2,2	0,045	0,0170	0,0166	0,98
Exclui ND				0,044	0,0170	0,0165	0,97
Subst=ND	46						

### B3.2. Amostra 006043 –água de superfície

TABELA 32 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06043

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	47	0,041	0,016	-0,030	0,90	0,018	0,019
Erro publicado	46	0,0267	0,010	0,001	0,30	0,0028	0,0078

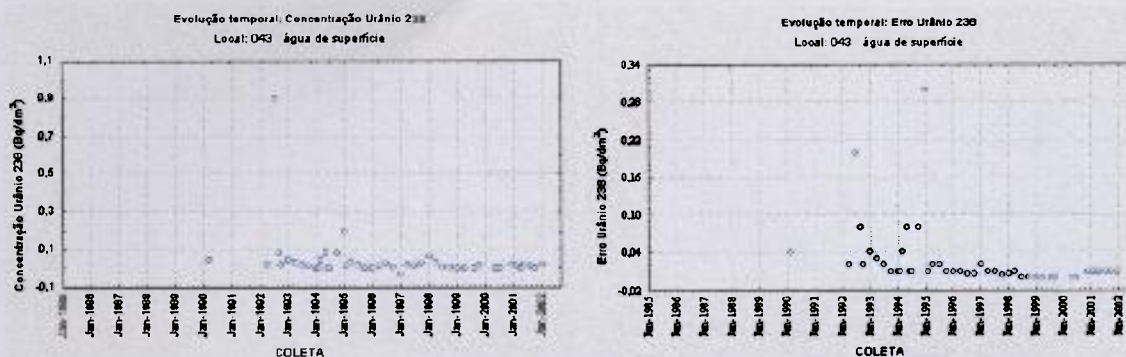


FIGURA 31 Evolução temporal da concentração e do erro Urânio 238; amostra 06043

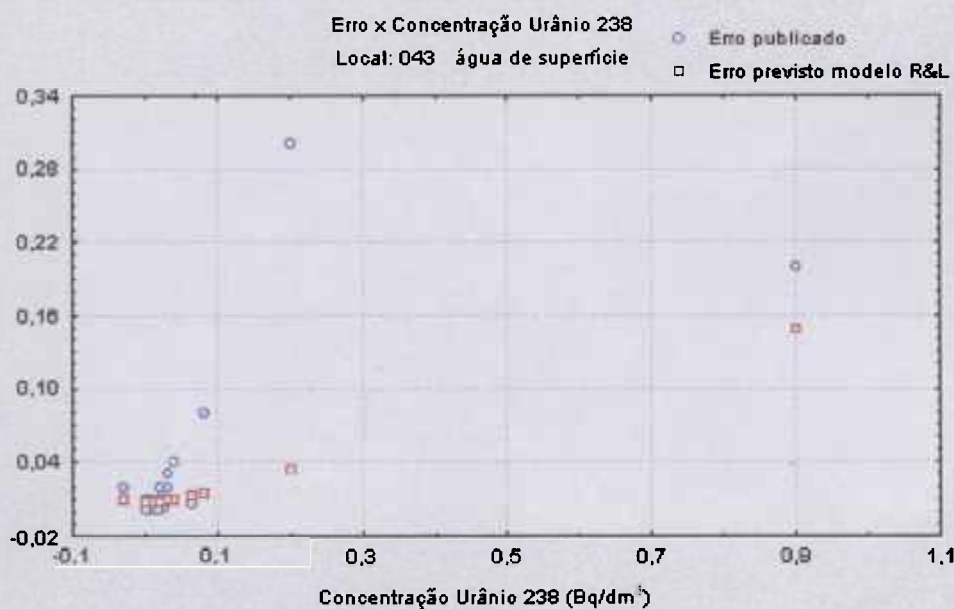


FIGURA 32 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 33 Comparação das médias: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06043

Amostra:	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A21043		46	2,1	0,041	0,00373	0,0126	3,38
Exclui ND				0,041	0,00373	0,0130	3,48
Subst=ND	47						

### B3.3. Amostra 006049 –água de superfície

TABELA 34 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06049

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	42	0,034	0,0085	-0,090	0,7700	0,015	0,019
Erro publicado	42	0,0223	0,0100	0,0010	0,3000	0,0023	0,0074

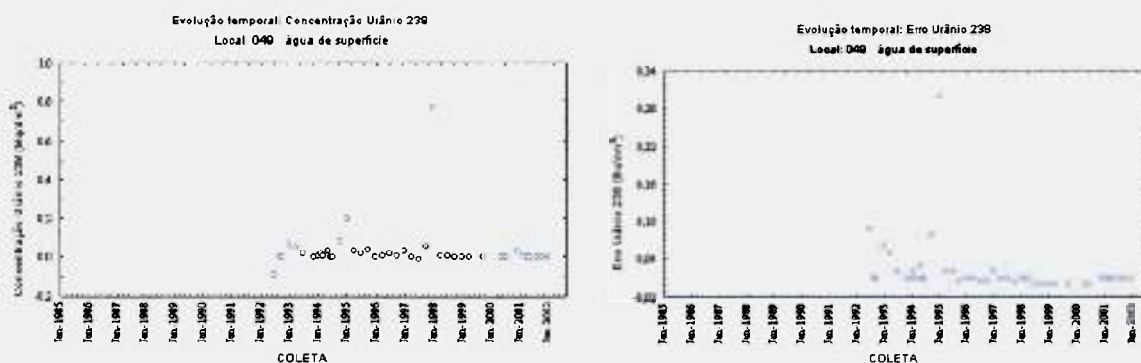


FIGURA 33 Evolução temporal da concentração e do erro Urânio 238; amostra 06049

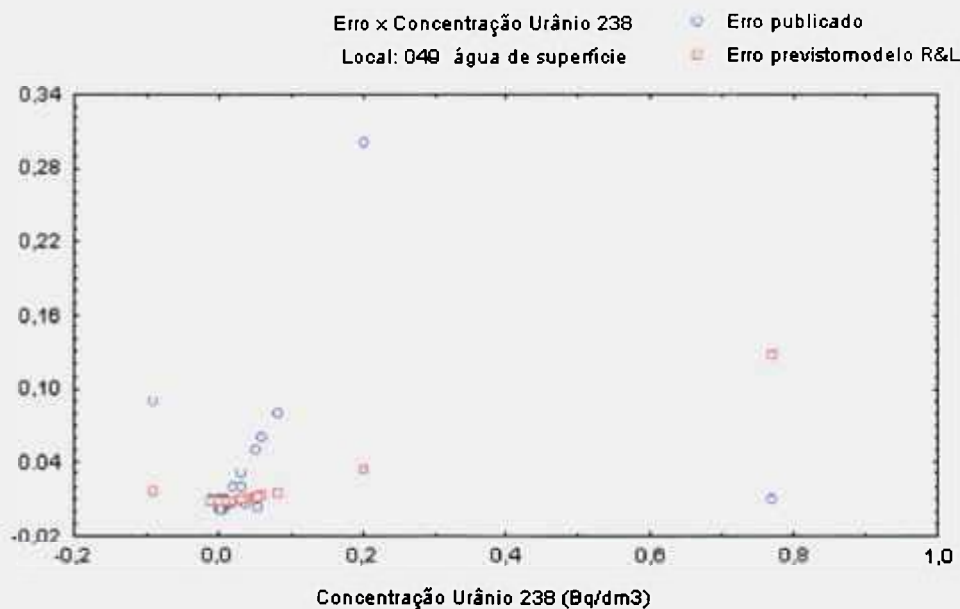


FIGURA 34 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 35 Comparação das médias: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 06049

Amostra:	NT	N	Graucen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
A21049							
Exclui ND	42	-	-	0,034	0,00386	0,00942	2,44
Subst=ND							

### B3.4. Amostra 008046 –água subterrânea

TABELA 36 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 08046

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	31	0,0196	0,010	-0,060	0,200	0,0018	0,0076
Erro publicado	30	0,0240	0,010	0,0010	0,300	0,0030	0,010

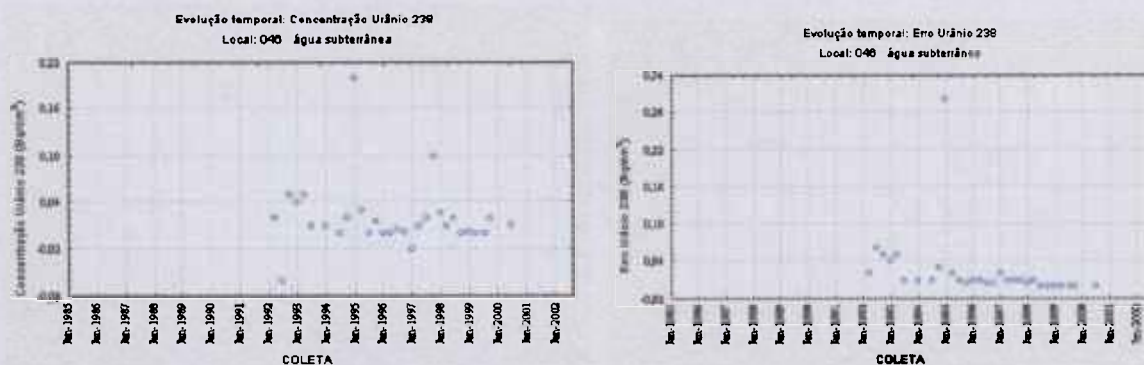


FIGURA 35 Evolução temporal da concentração e do erro Urânio 238; amostra 08046



FIGURA 36 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 37 Comparação das médias: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 08046

Amostra: A21046	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
Exclui ND		30	3,2	0,0198	0,00764	0,01058	1,38
Subst=ND	31			0,0196	0,00764	0,01064	1,39

### B3.5. Amostra 008047 –água subterrânea

TABELA 38 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 08047

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	38	0,0196	0,0080	-0,0600	0,200	0,0019	0,0072
Erro publicado	37	0,0204	0,010	0,0010	0,300	0,0024	0,0080

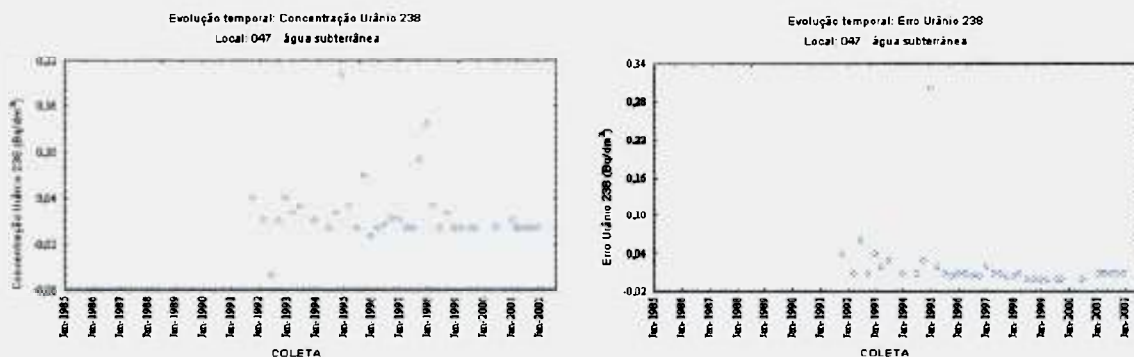


FIGURA 37 Evolução temporal da concentração e do erro Urânio 238; amostra 08047

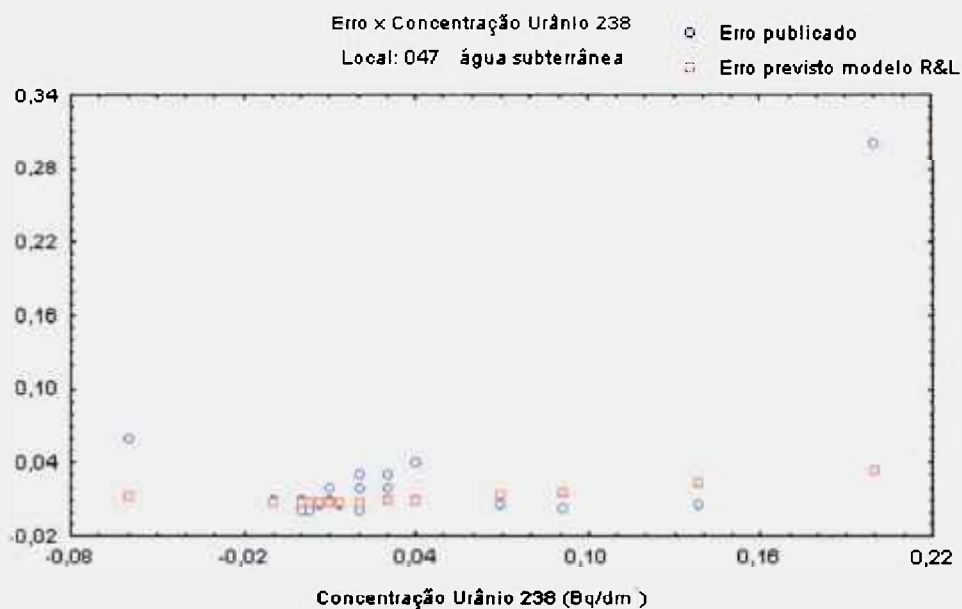


FIGURA 38 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 39 Comparação das médias : concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 08047

Amostra:	NT	N	Graucen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MP</sub> =Relação modelo/publicado
Exclui ND		37	2,6	0,0198	0,00557	0,00860	1,54
Subst=ND	38			0,0196	0,00557	0,00870	1,56

**B3.6. Amostra 021050 –esgoto industrial**

TABELA 40 Estatística descritiva: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 21050

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	37	0,224	0,0300	-0,0100	1,70	0,206	0,075
Erro publicado	37	0,0349	0,0100	0,0010	0,200	0,0039	0,010

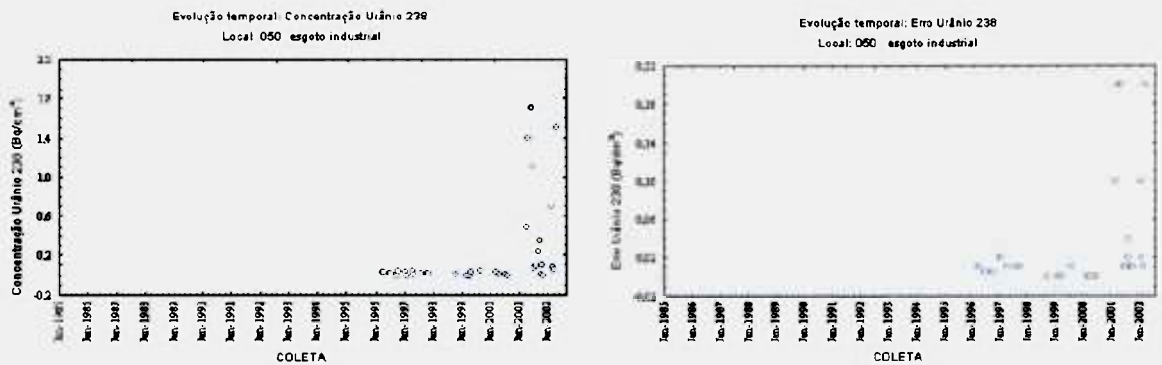


FIGURA 39 Evolução temporal da concentração e do erro Urânio 238; amostra 21050

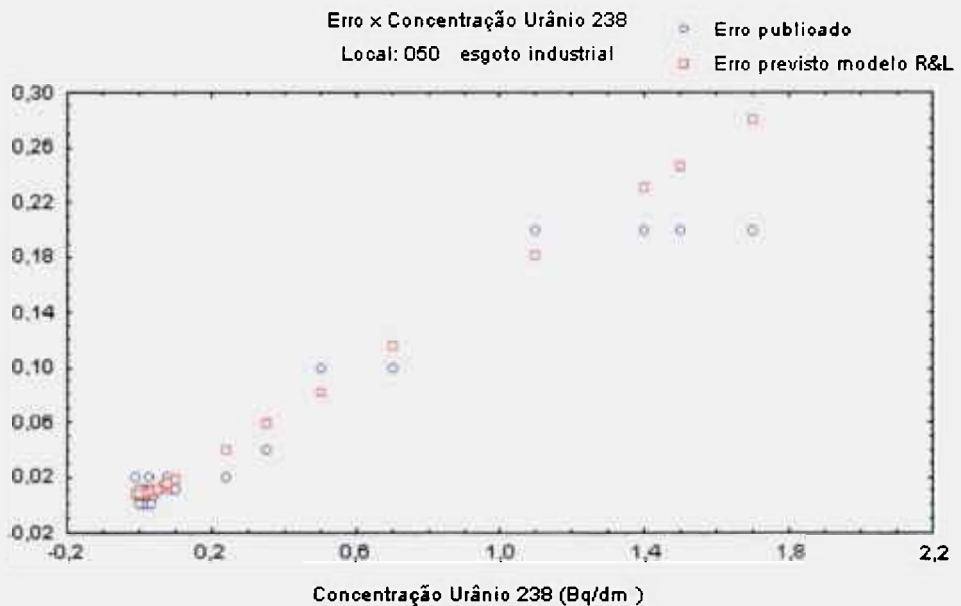


FIGURA 40 Comparação erro publicado e erro modelo R&L, Urânio 238

TABELA 41 Comparação das médias: concentração Urânio 238 (Bq/dm<sup>3</sup>); amostra: 21050

Amostra: A21050	NT	N	Grau cen (%)	Média simples	Média ponderada erro publicado	Média ponderada Erro Modelo	R <sub>MIP</sub> =Relação modelo/publicado
Exclui ND	37	-	-	0,224	0,0109	0,01928	1,77
Subst=ND							

## B4. Análise alfa total - matriz aerossol

### B4.1. Amostra 001001 - aerossol

TABELA 42 Estatística descritiva: concentração alfa total ( $\text{mBq/m}^3$ ); amostra: 01001

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	97	0,478	0,30	0,0	4,0	0,34	0,059
Erro publicado	97	0,20	0,10	0,001	4,0	0,21	0,047

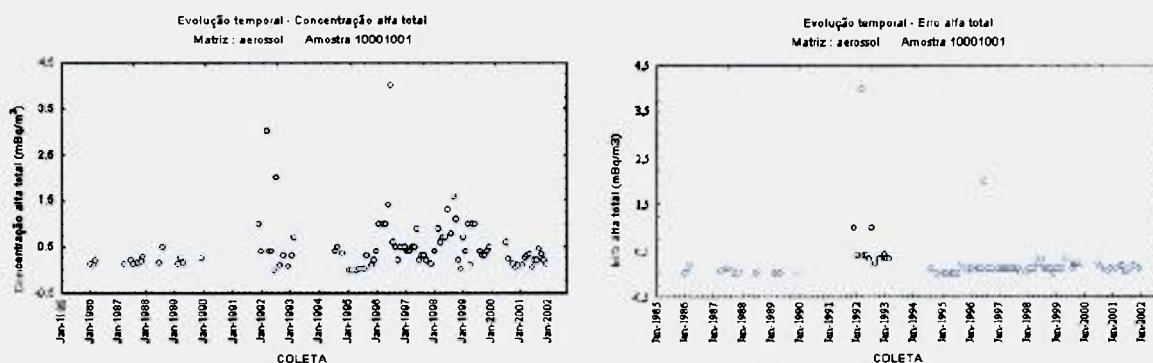


FIGURA 41 Evolução temporal da concentração e do erro alfa total; amostra 01001

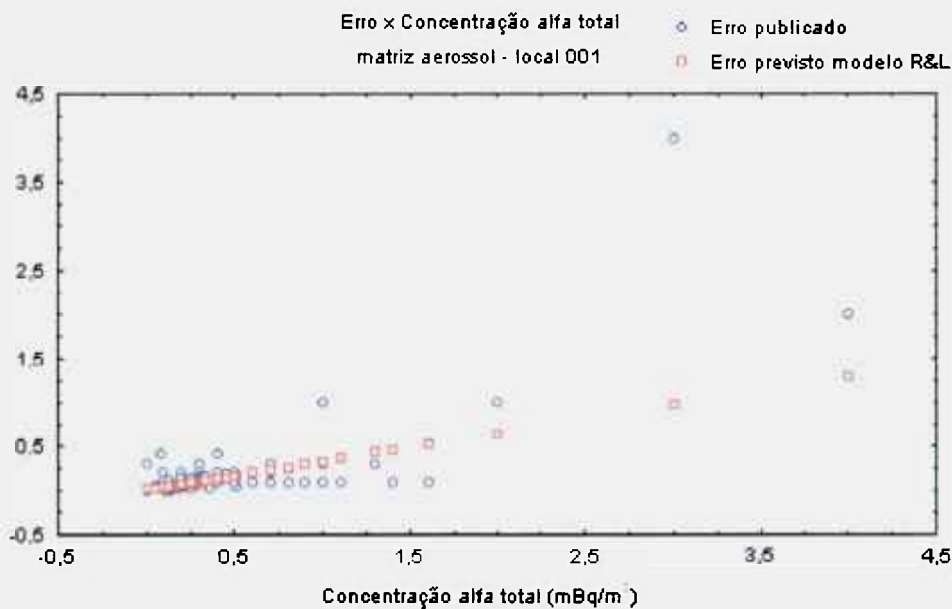


FIGURA 42 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 43 Comparação das médias: concentração alfa total ( $\text{mBq/m}^3$ ), amostra: 01001

N	Média simples	Média ponderada erro publicado	Média ponderada erro modelo	$R_{MP}$ =Relação modelo/publicado
97	0,478	0,00738	0,163	22,1

**B4.2. Amostra 001002 - aerossol**

TABELA 44 Estatística descritiva: concentração alfa total (mBq/m<sup>3</sup>); amostra: 01002

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	123	0,499	0,30	0,0	5	0,65	0,073
Erro publicado	123	0,21	0,10	0,001	4	0,24	0,045

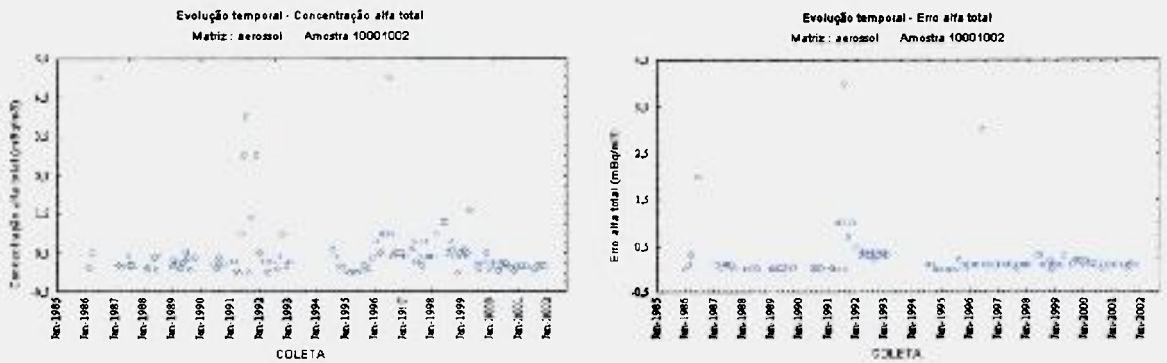


FIGURA 43 Evolução temporal da concentração e do erro alfa total; amostra 010002

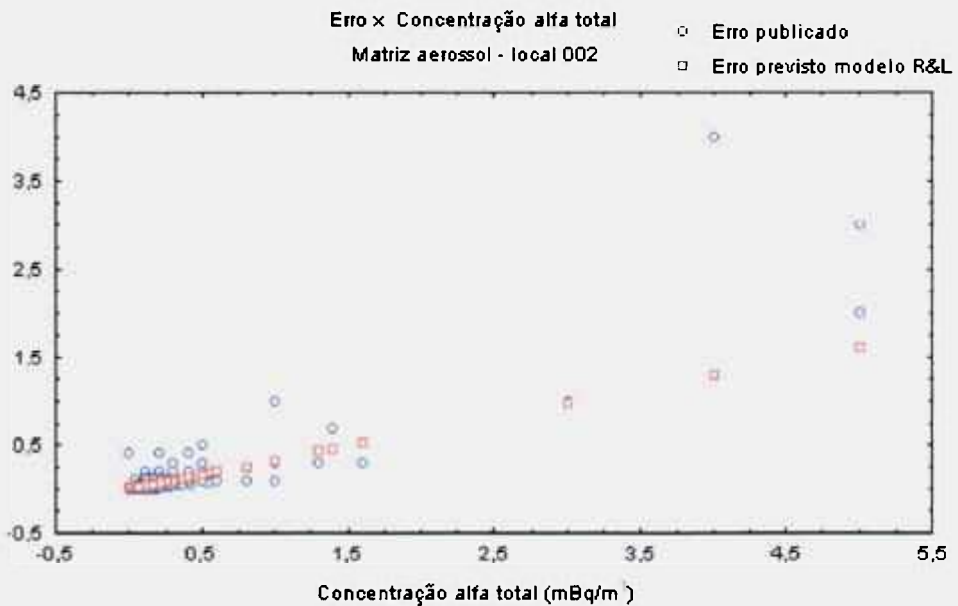


FIGURA 44 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 45 Comparação das médias: concentração alfa total (mBq/m<sup>3</sup>); amostra: 01002

N	Média simples	Média ponderada erro publicado	Média ponderada erro modelo	R <sub>MP</sub> =Relação modelo/publicado
123	0,499	0,0103	0,157	15,2

**B4.3. Amostra 001003 - aerossol**

TABELA 46 Estatística descritiva: concentração alfa total (mBq/m<sup>3</sup>); amostra: 01003

	N	Média	Mediana	Mínimo	Máximo	Variância	Erro da média
Concentração	135	0,454	0,30	0,0	3	0,24	0,042
Erro publicado	135	0,187	0,10	0,002	5,6	0,26	0,044

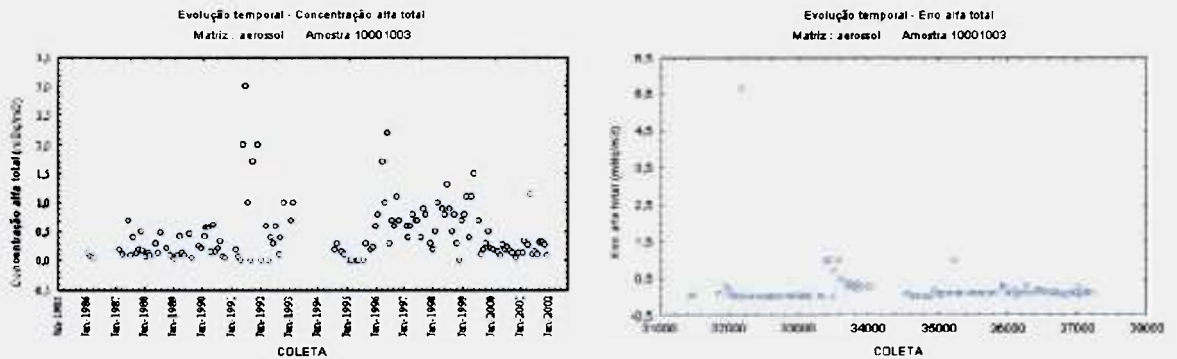


FIGURA 45 Evolução temporal da concentração e do erro alfa total; amostra 01003

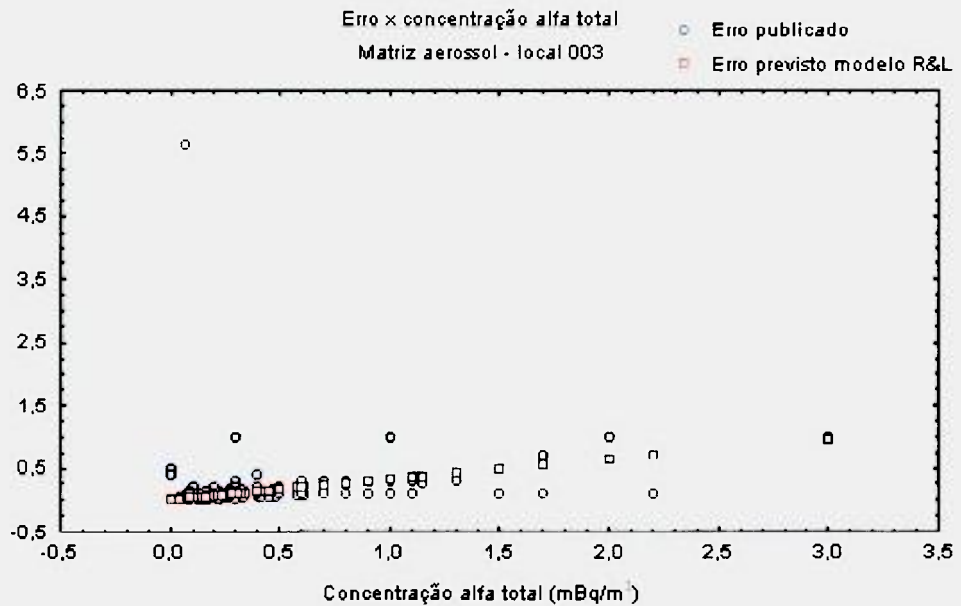


FIGURA 46 Comparação erro publicado e erro modelo R&L, alfa total

TABELA 47 Comparação das médias: concentração alfa total (mBq/m<sup>3</sup>); amostra: 01003

N	Média simples	Média ponderada erro publicado	Média ponderada erro modelo	R <sub>MP</sub> =Relação modelo/publicado
135	0,454	0,0233	0,143	6,1

## REFERÊNCIAS BIBLIOGRÁFICAS

- AKRITAS, MICHAEL G.; RUSCITTI, THOMAS F.; PATIL G. P. Statistical analysis of censored environmental data. In: PATIL, G. P.; RAO, C. R.. Handbook of Statistics. vol 12. Elsevier Science B.V., 1994.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. Estatística .Terminologia. Rio de Janeiro:ABNT,1988. (NBR 10536).
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. Guia para a expressão da incerteza da medição. Segunda edição brasileira do "Guide to the expression of uncertainty in measurement". Rio de Janeiro: ABNT,1998.
- ATWOOD, C. L.; BLACKWOOD, L.G.; HARRIS G. A.; LOWER C.A. Recommended Methods for Statistical Analysis of Data Containing Less-Than-Detectable Measurements. EGG-SARE-9247, Idaho National Engineering Laboratory: Idaho Falls, Idaho, 1991.
- BENVINGTON, P.R. Data Reduction and error analysis for the physical sciences. New York: Mcgraw-Hill, 1999.
- BERTHOUEX, P.M.; BROWN, L.C. Statistics for environmental engineers. Boca Raton: Lewis,1994.
- CENTRO DE DESENVOLVIMENTO DA TECNOLOGIA NUCLEAR (CDTN)  
Determinação de Alfa total e Beta total em águas por contagem alfa e contagem beta.  
RT(IT4) CDTN 0072 rev00 .mar/94.
- CHILDRESS, C.J.; FOREMAN,W.T.O.; CONNOR, B.F.; MALONEY,T.J. New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory. Open file report 99-193. Virginia,US:1999.
- CLARK, M.J.R.; WHITFIELD, P.H. A practical model integrating quality assurance into environmental monitoring. Water Resources Bulletin, 29(1):119 –130; 1993.
- CLARK, M.J.R.; WHITFIELD, P.H. Conflicting Perspectives about Detection Limits and About Censoring of Environmental Data. Water Resources Bulletin, 30(6):1063 –1079; 1994.
- COHEN, A.C. Truncated and censored samples, theory and applications, New York: Marcel Dekker, 1991.
- COTHERN, R.C.; ROSS, P.N. Uncertainties in assessing the state of the environment: an overview of environmental statistics, assessment and forecasting. In: COTHERN, R. C.; ROSS, P.N. Environmental Statistics, Assessment, and Forecasting. Lewis Publishers. 1994.

- CURIE, L.A. Limits for Qualitative Detection and Quantitative determination. *Analytical Chemistry*. Vol40, n<sup>o</sup> 3, USA. 1968.
- DAVIS, C. B. Environmental Regulatory Statistics. In: PATIL, G. P.; RAO, C. R.; *Handbook of Statistics*. vol 12, Elsevier Science B.V. 1994.
- DIGGLE, P.; LIANG, K.; ZEGGER, S. *Analysis of longitudinal data*. Oxford University Press: New York, N. Y., 1996.
- GARTHWAITE, P.H.; JOLLIFFE, I.T.; JONES B. *Statistical inference*. UK: Prentice Hall, 1995.
- GIBBONS, R.D.; COLEMAN, D.E.; Maddalone, R.F. An alternative Minimum level definition for analytical quantification. *Environmental Science Technology*, v. 31, p. 2071-2077, 1997.
- GIBBONS, R.D. *Statistical Methods for groundwater monitoring*. New York: Wiley, 1994.
- GILBERT, R.O. A review on statistical methods for data sets with multiple censoring points PNL-10652. Richland, Washington: Battelle Pacific Northwest Laboratories, 1995.
- GILBERT, R.O. *Statistical Methods for Environmental Pollution Monitoring*. New York: Van Nostrand Reinhold, 1987.
- GILBERT, R.O. "Nondetects: planning and statistical analyses to mitigate their effects." *DQO Statistics Bulletin, Statistical Methods for the Data Quality Objectives Process*, Volume 1, Number 2. PNL-SA-26377-2 Feb, 96. Richland, Washington: Pacific Northwest National Laboratory, 1996.
- GILLIOM, R. J.; HIRSCH, R.M.; GILROY, E.J. Effect of Censoring Trace-Level Water-Quality Data on Trend-Detection Capability. *Environmental Science and Technology* 18:530-535, 1984.
- GLASER, J.A.; FOERST, D.L.; MCKEE, G.D.; QUANE, S.A.; BUDDE, W.L. Trace Analyses for Wastewaters. *Environmental Science and Technology* 15:1426-1435, 1981.
- GLEIT, A. Estimation for small normal data sets with detection limits. *Environmental Science and Technology*. v19, p 1201-1206. 1985.
- HELSEL, D. R.; HIRSCH, R. M.. Summarizing data. In: *Statistical methods in Water resources*. New. York: Elsevier, 1992.
- HELSEL, D.R.. Less than Obvious: Statistical Treatment of Data below the Detection Limit. *Environmental Science and Technology*; 24(12): 1766-1774, 1990.
- HERTZELER, C.L.; ATWOOD, C.L.; HARRIS, G.A. Current methods of handling less-than-detectable measurements and detection limits in statistical analysis of environmental data. EGG-SARE-8609. Idaho Falls, Idaho: Idaho National Engineering Laboratory, 1991.
- KNOLL, Glenn F. *Radiation detection and measurement*. New York: Wiley, 1987.

- LAMBERT, D.; PETERSON, B.; TERPENNING, I. Nondetects, Detection Limits, and the Probability of Detection. *Journal of the American Statistical Association* 86(414):266-277, 1991.
- MARKER, DAVID A.; RYABOY, SVETLANA. The Quality of Environmental Databases. In: COTHERN, R; ROSS C.; PHILLIP, N; *Environmental Statistics, Assessment, and Forecasting*. Lewis Publishers. 1994.
- MEIER, PETER C.; ZÜND, RICHARD. *Statistical Methods in Analytical Chemistry*. New York: Wiley, 1993.
- MEYER, PAUL L. *Probabilidade aplicações à estatística*. Rio de Janeiro: Ao livro técnico, 1969.
- MICROSOFT CORPORATION, Microsoft® Excel 2000. Microsoft Corporation, 2000.
- MICROSOFT CORPORATION. Microsoft® Access 2000 (9.0.2812). Microsoft Corporation, 2000.
- NAGARAJ, N. K.; BRUNENMEISTER S. L. A new approach for accommodation of below detection limit data in trend analysis of water quality. In: COTHERN, R; ROSS C.; PHILLIP, N. *Environmental Statistics, Assessment, and Forecasting*. Lewis Publishers. 1994.
- OSBORNE, K.; ROCKE, D. The calculation of detection limits using a two components error model and laboratory QC Data. *Water Environment Laboratory Solutions*. Vol 7 n<sup>o</sup> 4. 2000.
- PATIL, G. P.; TAILLIE, C.; TALWALKER, S. Encounter sampling and modeling in ecological and environmental studies using weighted distribution methods. In: BARNETT, V.; TURKMAN, K. F.; *Statistics for the environment*. England: John Wiley, 1993.
- PEIXOTO, C. M.; PÊGO, V. D. Relatório de Avaliação dos resultados analíticos do Programa de Monitoração Ambiental do Centro de Desenvolvimento da Tecnologia Nuclear, 1991 e 1992. CDTN-827/96, 1996.
- PORTER, S. P. The detection limit. *Environmental Science and Technology*; vol 22 n<sup>o</sup> 8; 1988.
- PRITT, J. Description and guide for interpreting low level data supplied by NWQL, National Water Quality Laboratory, Technical memorandum 94-12, USA, 1994.
- ROCKE, D. M.; LORENZATO, S. A Two-component model for measurement error in analytical chemistry, *Technometrics*, May 1995, vol 37, n<sup>o</sup> 2, pp. 176-184. 1995.
- ROSS, P.N.; COTHERN, R. C. Environmental statistics: Where do we go from here? In: COTHERN, R. C.; ROSS, P.N. *Environmental Statistics, Assessment, and Forecasting*. Lewis, 1994.

- SACRAMENTO, A. et al. Relatório de avaliação dos resultados analíticos do Programa de Monitoração Ambiental do Centro de Desenvolvimento da Tecnologia Nuclear 1986 – 1990. Supervisão de Meio Ambiente/CDTN. agosto, 1993.
- STATSOFT. Statistica 5.1; Statistica for Windows. Tulsa, OK: Statsoft Inc., 1998.
- TRIOLA, M. F. Introdução à estatística. Rio de Janeiro: Livros técnicos e científicos, 1999.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-USEPA. Guidance for Data Quality Assessment. EPA/600/R-96/084, 2000b.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-USEPA. Guidance on Data Quality Indicators. EPA QA/G-5i. 2001b.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-USEPA. Guidance on Environmental Data Verification and data validation. EPA QA/G8. 2001a.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-USEPA. Guidelines establishing test procedures for analysis of pollutants (App. B, Part 136) U.S. Code of Federal Regulations, Title 40, p 265-267. 1997.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY-USEPA. Quality Manual for environmental programs. EPA 5360A1. 2000a.
- WILSON, M. et al. Application to Environmental Monitoring of a Two-Component Model for Chemical Analytical Error, Application paper. Davis, California: University of California, 2000.
- ZORN, M.E.; GIBBONS, R. D.; SONZOGNI, W.C. Weighted Least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. Analytical Chemistry. 69, p 3069-3075. 1997.
- ZORN, M.E.; GIBBONS, R. D.; SONZOGNI, W.C. Evaluation of approximate methods for calculating the limit of detection and limit of quantification, Environmental Science and Technology, Vol. 33, p 2229 -2295, 1999.