

DAMARES MIRIAN DE SOUZA PAVIONE

**ANÁLISE DE INCERTEZAS E DE SENSIBILIDADE E AVALIAÇÃO  
QUANTITATIVA DE RISCO MICROBIOLÓGICO APLICADAS AO  
DIMENSIONAMENTO DE LAGOAS DE POLIMENTO E À UTILIZAÇÃO DE  
EFLUENTES PARA IRRIGAÇÃO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS – BRASIL  
2010

**Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV**

T

P338a  
2010

Pavione, Damares Mirian de Souza, 1984-  
Análise de incertezas e de sensibilidade e avaliação  
quantitativa de risco microbiológico aplicadas ao dimensionamento de lagoas de polimento e à utilização de efluentes para irrigação / Damares Mirian de Souza Pavione.  
- Viçosa, MG, 2010.  
xvii, 166f. : il. (algumas col.) ; 29cm.

Orientador: Rafael Kopschitz Xavier Bastos.  
Dissertação (mestrado) - Universidade Federal de Viçosa.  
Referências bibliográficas: f. 156-166.

1. Esgotos. 2. Lagoas de estabilização. 3. Águas residuais.  
4. Irrigação com águas residuais. I. Universidade Federal de Viçosa. II. Título.

CDD 22. ed. 628.3

DAMARES MIRIAN DE SOUZA PAVIONE

**ANÁLISE DE INCERTEZAS E DE SENSIBILIDADE E AVALIAÇÃO  
QUANTITATIVA DE RISCO MICROBIOLÓGICO APLICADAS AO  
DIMENSIONAMENTO DE LAGOAS DE POLIMENTO E À UTILIZAÇÃO DE  
EFLUENTES PARA IRRIGAÇÃO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Magister Scientiae*.

APROVADA: 22 de março de 2010

---

Prof.<sup>a</sup> Paula Dias Bevilacqua  
(Coorientadora)

---

Prof.<sup>a</sup> Ana Augusta Passos Rezende

---

Prof. José Ivo Ribeiro Júnior

---

Prof.<sup>a</sup> Silvia Maria Alves Corrêa Oliveira

---

Rafael Kopschitz Xavier Bastos  
(Orientador)

## **AGRADECIMENTOS**

Ao professor e orientador Rafael Bastos, meu agradecimento especial, por suas observações, empenho e dedicação com que me orientou, tornando possível a realização desta dissertação.

À minha família que me apoiou e acompanhou em todas as etapas.

Ao meu amável esposo pelo incentivo e suporte.

A todos os colegas e professores da pós-graduação.

A todos os amigos e ex companheiras de república.

À Universidade Federal de Viçosa.

À Deus pelas misericórdias diariamente renovadas sobre mim.

## SUMÁRIO

LISTA DE TABELAS .....	v
LISTA DE QUADROS.....	viii
LISTA DE FIGURAS.....	x
RESUMO.....	xiv
ABSTRACT .....	xvi
<b>1 – INTRODUÇÃO .....</b>	<b>1</b>
<b>2 – OBJETIVOS .....</b>	<b>4</b>
<b>2.1 - Objetivo geral.....</b>	<b>4</b>
<b>2.2 - Objetivos específicos .....</b>	<b>4</b>
<b>3 - REVISÃO DE LITERATURA .....</b>	<b>5</b>
<b>3.1 - Análise de incertezas em modelos estocásticos .....</b>	<b>5</b>
<i>3.1.1 - Caracterização da distribuição de probabilidades.....</i>	<i>6</i>
3.1.1.1 - <u>Distribuição de probabilidades usuais .....</u>	7
3.1.1.2 – <u>Testes de aderência e verificação gráfica da distribuição.....</u>	14
3.1.1.3 – <u>Estatística descritiva .....</u>	17
3.1.1.4 - <u>Representação gráfica da distribuição de frequências .....</u>	21
3.1.2 - <i>Critérios estatísticos para a análise de frequência.....</i>	21
3.1.3 - <i>Amostragem aleatória das variáveis de entrada em modelos estocásticos .....</i>	22
<b>3.2 - Análise de sensibilidade em modelos estocásticos .....</b>	<b>24</b>
<b>3.3 - Análise de incertezas e de sensibilidade na previsão da qualidade de efluentes de lagoas de estabilização.....</b>	<b>28</b>
3.3.1 - <i>Critérios de dimensionamento de lagoas de estabilização .....</i>	29
3.3.1.1 - <u>Modelo de dispersão - fluxo disperso .....</u>	30
3.3.1.2 - <u>Temperatura do líquido.....</u>	31
3.3.1.3 - <u>Remoção de DBO .....</u>	31
3.3.1.4 - <u>Remoção de coliformes .....</u>	32
3.3.1.5 – <u>Remoção de amônia .....</u>	34
3.3.1.6 – <u>Remoção de ovos de helmintos.....</u>	35
3.3.2 – <i>Caracterização das incertezas associadas aos parâmetros de projeto de lagoas de estabilização.....</i>	36
<b>3.4 - Análise de incertezas e de sensibilidade na Avaliação Quantitativa de Risco Microbiológico aplicada ao uso de efluentes para irrigação.....</b>	<b>38</b>
3.4.1 - <i>Normas e critérios de qualidade microbiológica da água para a irrigação com esgotos.....</i>	38
3.4.2 - <i>Avaliação Quantitativa de Risco Microbiológico (AQRM) .....</i>	42
3.4.3 - <i>Caracterização das incertezas associadas aos modelos de AQRM aplicados à irrigação com esgotos .....</i>	44
<b>3.5 - Normas brasileiras de qualidade da água para irrigação e padrões de lançamento de efluentes .....</b>	<b>46</b>
<b>4 - CARACTERIZAÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADES DOS PARÂMETROS DE QUALIDADE DO ESGOTO BRUTO E DOS EFLUENTES DAS LAGOAS DE POLIMENTO DA ETE VIOLEIRA .....</b>	<b>48</b>

<b>4.1 – Metodologia.....</b>	<b>48</b>
<b>4.2 - Resultados e discussão .....</b>	<b>53</b>
4.2.1 - <i>Verificação de diferenças entre dados dos distintos períodos operacionais .....</i>	53
4.2.2 – <i>Estatística descritiva dos dados do esgoto bruto, reator anaeróbio, biofiltro e dos efluentes das lagoas de polimento agrupados por períodos operacionais com resultados estatisticamente equivalentes .....</i>	55
4.2.3 - <i>Testes de aleatoriedade, estacionariedade, homogeneidade e independência .....</i>	65
4.2.4 - <i>Testes de aderência e gráficos de probabilidade.....</i>	72
4.2.5 – <i>Descrição da distribuição selecionada por meio de hitogramas .....</i>	92
4.2.6 - <i>Medidas de tendência central .....</i>	101
<b>4.3 – Conclusões.....</b>	<b>102</b>
<b>5 - ANÁLISE DE INCERTEZA E DE SENSIBILIDADE NA PREVISÃO DA QUALIDADE DE EFLUENTES DE LAGOAS DE POLIMENTO.....</b>	<b>104</b>
<b>5.1 – Metodologia.....</b>	<b>105</b>
5.1.1 – <i>Modelos de previsão da qualidade dos efluentes .....</i>	105
5.1.2 - <i>Análises de sensibilidade e de cenário.....</i>	109
<b>5.2 - Resultados e discussão .....</b>	<b>110</b>
5.2.1 – <i>Estimativas da qualidade dos efluentes das lagoas de polimento .....</i>	110
5.2.2 - <i>Análise de sensibilidade.....</i>	117
<b>5.3 – Considerações finais, conclusões e recomendações.....</b>	<b>122</b>
<b>6 - ANÁLISE DE INCERTEZAS E DE SENSIBILIDADE NA AVALIAÇÃO QUANTITATIVA DE RISCO MICROBIOLÓGICO APLICADA AO USO AGRÍCOLA DE EFLUENTES DE LAGOAS DE POLIMENTO .....</b>	<b>125</b>
<b>6.1 – Metodologia.....</b>	<b>126</b>
6.1.1 - <i>Modelos dose-resposta e caracterização do risco .....</i>	126
6.1.2 - <i>Modelo de exposição.....</i>	127
6.1.3 - <i>Análises de sensibilidade e de cenário.....</i>	130
<b>6.2 - Resultados e discussão .....</b>	<b>130</b>
6.2.1 - <i>Estimativas de risco de infecção.....</i>	130
6.2.2 – <i>Estimativas de contaminação das hortaliças irrigadas .....</i>	142
<b>6.3 - Análise de sensibilidade .....</b>	<b>145</b>
<b>6.4 – Considerações finais, conclusões e recomendações.....</b>	<b>153</b>
<b>REFERÊNCIAS .....</b>	<b>150</b>

## LISTA DE TABELAS

TABELA 4.1 - Variações operacionais do Sistema de lagoas da Unidade Experimental da Violeira, Viçosa – MG.....	50
TABELA 4.2 - Estatística descritiva das séries de dados de DBO, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	56
TABELA 4.3 - Estatística descritiva das séries de dados de DQO, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	57
TABELA 4.4 - Estatística descritiva das séries de dados de SST, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	58
TABELA 4.5 - Estatística descritiva das séries de dados de Nitrogênio amoniacal, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	59
TABELA 4.6 - Estatística descritiva das séries de dados de Fósforo, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	60
TABELA 4.7 - Estatística descritiva das séries de dados de pH, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	61
TABELA 4.8 - Estatística descritiva das séries de dados de <i>E. coli</i> , Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	62
TABELA 4.9 - Estatística descritiva das séries de dados de coliformes totais, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	63
TABELA 4.10 - Estatística descritiva das séries de dados de ovos de helmintos, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008. ....	64
TABELA 11 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de DBO no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.....	65
TABELA 4.12 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de ovos helmintos no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.....	66
TABELA 4.13 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de DQO no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.....	66
TABELA 14 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de SST nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.....	67

TABELA 4.15 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de nitrogênio amoniacal no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008. ....	67
TABELA 4.16 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de fósforo no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008. ....	68
TABELA 4.17 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de pH nos efluentes das lagoas de polimento, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008. ....	68
TABELA 18 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de <i>E. coli</i> no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008. ....	69
TABELA 4.19 - Resultados (valor <i>p</i> ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de coliformes totais no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008. ....	69
TABELA 4.20 - Síntese dos resultados dos testes de aleatoriedade, estacionariedade, homogeneidade e independência. ....	70
TABELA 4.21 – Número e percentual de séries de dados que atenderam simultaneamente as hipóteses de aleatoriedade, estacionariedade, homogeneidade e independência. ....	71
TABELA 5.1 – Variáveis de entrada dos modelos de previsão da qualidade dos efluentes da ETE Viçosa com as respectivas distribuições de probabilidades e parâmetros que as definem. ....	107
TABELA 5.2 - Estimativa da qualidade dos efluentes das lagoas de polimento (medianas e percentis 95%), <i>E. coli</i> por 100 mL. ....	110
TABELA 5.3 – Percentuais de atendimentos aos níveis de qualidade $10^3$ , $10^4$ e $10^5$ <i>E. coli</i> /100mL. ....	112
TABELA 5.4 - Estimativa da qualidade dos efluentes das lagoas de polimento (medianas e percentis 95%), mg N-NH <sub>3</sub> /L. ....	113
TABELA 5.5 – Percentuais de atendimentos aos níveis de qualidade de 5 e 20 mg N-NH <sub>3</sub> /L. ....	114
TABELA 5.6 – Diferenças (mediana e percentil 95%) entre a previsão das concentrações efluentes de <i>E. coli</i> e de nitrogênio amoniacal utilizando valores de pH, concentrações de <i>E. coli</i> e de nitrogênio amoniacal afluentes segundo distribuição uniforme em relação ao uso da distribuição e lognormal. ....	115

TABELA 6.1 – Parâmetros dos modelos beta-Poisson e exponencial utilizados no modelo de AQRM, valores médios, máximos, mínimos e distribuição de probabilidades adotada. ....	127
TABELA 6.2 – Valores de número de patógenos por $10^5$ <i>E. coli</i> e de redução de patógenos entre a colheita e o consumo utilizados no modelo de AQRM, valores máximos e mínimos e distribuição de probabilidades adotada.....	128
TABELA 6.3 - Concentração de <i>E. coli</i> nos efluente das lagoas de polimento nas séries de dados selecionadas para utilização no modelo de AQRM, medianas e valores-limite (percentis) para a caracterização da distribuição lognormal.....	129
TABELA 6.4 - Consumo anual de hortaliças ingeridas cruas no Brasil, valores médios, mínimos e máximos utilizados para a caracterização da distribuição uniforme e para utilização no modelo de AQRM .....	130
TABELA 6.5 - Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo e que são ingeridas cruas, qualidade dos efluentes das lagoas segundo distribuição lognormal.....	131
TABELA 6.6 – Estimativas de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo e que são ingeridas cruas, qualidade dos efluentes das lagoas segundo distribuição lognormal.....	132
TABELA 6.7 – Percentuais de atendimentos ao nível tolerável de risco de infecção de $10^{-2}$ pppa, cenário de exposição ao consumo de hortaliças irrigadas com efluentes de lagoas de polimento.....	137
TABELA 6.8 – Percentuais de atendimentos ao nível tolerável de risco de infecção de $10^{-3}$ pppa, cenário de exposição ao consumo de hortaliças irrigadas com efluentes de lagoas de polimento.....	137
TABELA 6.9 - Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme. ....	137
TABELA 6.10 – Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme. ....	138
TABELA 6.11 – Diferença em unidades logarítmicas (mediana e percentil 95%) entre o risco de infecção anual estimado com a concentração de <i>E. coli</i> nos efluentes das lagoas segundo distribuição uniforme e lognormal - hortaliças que se desenvolvem rentes ao nível do solo.....	139
TABELA 6.12 – Diferença em unidades logarítmicas (mediana e percentil 95%) entre o risco de infecção anual estimado com a concentração de <i>E. coli</i> nos efluentes das lagoas segundo distribuição uniforme e lognormal - hortaliças que se desenvolvem distantes do nível do solo. ....	139
TABELA 6.13 – Estimativa da contaminação ( <i>E. coli</i> /g) das hortaliças irrigadas, medianas e percentis 95% - qualidade dos efluentes das lagoas segundo distribuição lognormal. ....	142

## LISTA DE QUADROS

Quadro 3.1 – Curtose e assimetria de algumas distribuições de probabilidades .....	20
Quadro 3.2 - Diretrizes da USEPA para o uso agrícola de esgotos sanitários .....	39
Quadro 3.3 – Diretrizes da OMS o uso agrícola de esgotos sanitários.....	40
Quadro 3.4 - Critérios de qualidade da água para irrigação estabelecidos na Resolução CONAMA nº 357/2005 .....	47
Quadro 4.1 - Agrupamento dos períodos com mesmas características de operação das lagoas de polimento.....	53
Quadro 4.2 - Agrupamento dos períodos operacionais das lagoas de polimento de acordo com os resultados do teste de Kruskal-Wallis, com nível de 5% de significância, para identificação de dados de qualidade dos efluentes que não apresentam diferenças estatisticamente significativas.....	54
Quadro 4.3 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de DBO. ....	72
Quadro 4.4 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de DQO. ....	73
Quadro 4.5 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de SST. ....	73
Quadro 4.6 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de pH. ....	74
Quadro 4.7 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de nitrogênio amoniacal. ....	74
Quadro 4.8 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de fósforo. ....	75
Quadro 4.9 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de <i>E. coli</i> . ....	75
Quadro 4.10 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de coliformes totais. ....	76
Quadro 4.11 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de ovos de helmintos. ....	76
Quadro 12 - Resultados dos testes de normalidade das séries de dados ao nível de 5% de significância.....	77
Quadro 4.13 - Distribuições selecionadas para as séries de dados analisadas no esgoto bruto e nos efluentes das lagoas de polimento.....	93
Quadro 5.1 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para a concentração efluente de <i>E. coli</i> . ....	118

Quadro 5.2 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para a concentração efluente de nitrogênio amoniacal .....	119
Quadro 6.1 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por <i>Campylobacter</i> decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo. ....	146
Quadro 6.2 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo. ....	147
Quadro 6.3 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por <i>Cryptosporidium</i> decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo. ....	148
Quadro 6.4 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por <i>Campylobacter</i> decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	149
Quadro 6.5 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	150
Quadro 6.6 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por <i>Cryptosporidium</i> decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	151

## LISTA DE FIGURAS

Figura 3.1 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição normal para $\mu=0$ e diferentes valores de $\sigma$ .....	8
Figura 3.2 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição lognormal para $\mu=0$ e diferentes valores de .....	9
Figura 3.3 – Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição exponencial para diferentes valores de $\lambda$ .....	10
Figura 3.4 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição uniforme .....	11
Figura 3.5 – Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição Weibull para $\lambda=1$ e diferentes valores de $k$ .....	12
Figura 3.6 – Ilustração de funções massa de probabilidades de Poisson para diferentes valores de $v$ . .....	13
Figura 3.7 – Ilustração de funções massa de probabilidades binomial negativa para diferentes valores de $r$ e $p$ .....	13
Figura 3.8 – Ilustração de funções massa de probabilidades geométrica para diferentes valores de $p$ . .....	14
Figura 3.9 – Ilustrações de teste gráfico de aderência às distribuições normal (à esquerda) e lognormal (à direita).....	17
Figura 3.10 – Ilustração de curva leptocúrtica (em vermelho), mesocúrtica (em verde) e platicúrtica (em azul).....	20
Figura 3.11 - Ilustração de um histograma com a respectiva curva de frequência.....	21
Figura 3.12 – Procedimentos da simulação Monte Carlo (à esquerda) e histograma resultante (à direita).....	23
Figura 3.13 - Exemplo de análise de sensibilidade por meio de gráfico de dispersão .....	28
Figura 3.14 - Remoção de coliformes em função do produto $K_b.t$ para diversos valores de número de dispersão $d$ .....	33
Figura 3.15 - Exemplos de combinação de medidas de proteção à saúde para carga de doenças tolerável $\leq 10^{-6}$ DALY pppa.....	41
Figura 4.1 - ETE Viçosa, Viçosa, MG – vista do sistema UASB + BF (à esquerda) e série experimental de lagoas (à direita).....	49
Figura 4.2 - Gráficos para verificação da aderência às distribuições candidatas, dados de DBO, dados de esgoto bruto, reator anaeróbio, lagoa 1, lagoa 2 e lagoa 3.....	78
Figura 4.3 - Gráficos para verificação da aderência às distribuições candidatas, dados de DQO, dados de esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2 e lagoa 3.....	79

Figura 4.4 - Gráficos para verificação da aderência às distribuições candidatas, dados de SST, reator anaeróbico, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo. ....	82
Figura 4.5 - Gráficos para verificação da aderência às distribuições candidatas, dados de pH, lagoa 1, lagoa 2 e lagoa 3. ....	84
Figura 4.6 - Gráficos para verificação da aderência às distribuições candidatas, dados de nitrogênio amoniacal, esgoto bruto, reator anaeróbico, bio filtro, lagoa 3 e lagoas 3 e 4 operando em paralelo. ....	85
Figura 4.7 - Gráficos para verificação da aderência às distribuições candidatas, dados de fósforo, esgoto bruto, reator anaeróbico, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo. ....	87
Figura 4.8 - Gráficos para verificação da aderência às distribuições candidatas, dados de <i>E. coli</i> , esgoto bruto, reator anaeróbico, lagoa 2 e lagoa 3. ....	89
Figura 4.9 - Gráficos para verificação da aderência às distribuições candidatas, dados de coliformes totais, esgoto bruto, reator anaeróbico, biofiltro, lagoa 1 e lagoa 3. ....	90
Figura 4.10 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DBO esgoto bruto, reator anaeróbico, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4. ....	94
Figura 4.11 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DQO esgoto bruto, reator anaeróbico, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4. ....	95
Figura 4.12 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de SST esgoto bruto, reator anaeróbico, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4. ....	96
Figura 4.13 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de pH, lagoa 1, lagoa 2, lagoa 3 e lagoa 3 operando em paralelo com a lagoa 4. ....	97
Figura 4.14 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de helmintos esgoto bruto, lagoa 1 e lagoa 2. ....	98
Figura 4.15 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de <i>E. coli</i> , esgoto bruto, reator anaeróbico, lagoa 1, lagoa 2, lagoa 3 e lagoa 4. ....	98
Figura 4.16 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de coliformes totais, esgoto bruto, reator anaeróbico, biofiltro, lagoa 1, lagoa 2, lagoa 3 e lagoa 3 operando em paralelo com a lagoa 4. ....	99
Figura 4.17 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de nitrogênio amoniacal esgoto bruto, reator anaeróbico, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4. ....	100

Figura 4.18 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de fósforo esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo coma lagoa 4 e lagoa 4. ....	100
Figura 4.19 - Média geométrica e média aritmética em curvas de frequências de séries de dados de <i>E. coli</i> e coliformes totais. ....	102
Figura 5.1 – Histogramas de frequências de variáveis de entrada dos modelos de previsão de concentração efluente de <i>E. coli</i> e nitrogênio amoniacal com distribuição lognormal ou triangular.....	108
Figura 5.2 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das estimativas das concentrações de <i>E. coli</i> efluente a partir das concentrações afluentes segundo distribuição lognormal.....	111
Figura 5.3 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das estimativas das concentrações de nitrogênio amoniacal efluente e pH a partir das concentrações afluentes segundo distribuição lognormal.....	114
Figura 5.4 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada da previsão de concentração de <i>E. coli</i> efluente a partir das concentrações afluentes segundo distribuição uniforme.....	116
Figura 5.5 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das previsões de concentrações de nitrogênio amoniacal efluente a partir das concentrações afluentes e pH segundo distribuição uniforme. ....	117
Figura 5.6 - Gráfico de dispersão, variável de entrada versus concentração efluente de <i>E. coli</i> .....	118
Figura 5.7 - Gráfico de dispersão, variável de entrada versus concentração de nitrogênio amoniacal efluente.....	119
Figura 6.1 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição lognormal.....	133
Figura 6.2 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem distantes do nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição lognormal.....	134
Figura 6.3 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme. ....	140
Figura 6.4 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem	

distantes do nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme. ....	141
Figura 6.5 - Histogramas (distribuição de probabilidades) da contaminação das hortaliças irrigadas ( <i>E. coli</i> por grama) - qualidade dos efluentes das lagoas segundo distribuição lognormal. ....	144
Figura 6.6 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por <i>Campylobacter</i> decorrentes do consumo de hortaliças que se desenvolvem rentes ao solo. ....	146
Figura 6.7 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo. ....	147
Figura 6.8 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por <i>Cryptosporidium</i> decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo. ....	148
Figura 6.9 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por <i>Campylobacter</i> decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	149
Figura 6.10 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	150
Figura 6.11 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por <i>Cryptosporidium</i> decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo. ....	151

## RESUMO

PAVIONE, Damares Mirian de Souza, M.Sc., Universidade Federal de Viçosa, março de 2010.  
**Análise de incerteza e de sensibilidade e avaliação quantitativa de risco microbiológico aplicadas ao dimensionamento de lagoas de polimento e à utilização de efluentes para irrigação.** Orientador: Rafael Kopschitz Xavier Bastos. Coorientadores: Paula Dias Bevilacqua e Paulo Roberto Ceccon.

Este trabalho apresenta a caracterização e análise de um banco de dados relativo à cerca de sete anos de monitoramento de um sistema de tratamento de esgotos sanitários constituído por reator UASB, biofiltro submerso aerado (escala real) e lagoas de polimento (escala piloto). Foram analisadas as seguintes variáveis: DBO, DQO, SST, N-NH<sub>3</sub>, fósforo total, pH, *E. coli*, coliformes totais e ovos de helmintos. As hipóteses de aleatoriedade, estacionariedade e homogeneidade dos dados foram, em geral, confirmadas, mas a de independência não. A maioria das séries de dados apresentou características de assimetria e curtose distintas da distribuição normal. Os resultados sugerem que a função lognormal melhor descreveu a distribuição dos dados de *E. coli*, coliformes totais, pH, SST, DBO, DQO e N-NH<sub>3</sub>. Os dados de fósforo apresentaram melhor adesão ao modelo normal de distribuição. Quanto aos ovos de helmintos, não foi possível identificar uma distribuição que mais nitidamente se ajustasse ao banco de dados disponível. A distribuição lognormal foi aplicada às variáveis de entrada pH, N-NH<sub>3</sub> e *E. coli* em modelagem estocástica (amostragem aleatória por hipercubos latinos com 10.000 iterações) para a previsão da qualidade de efluentes de lagoas de polimento. Constatou-se que uma série de três lagoas parece ser suficiente para o alcance de 20 mg N-NH<sub>3</sub> L<sup>-1</sup> como padrão de lançamento de efluentes, e para atender as recomendações da OMS em termos de concentração de *E. coli* para a irrigação de hortaliças folhosas, raízes e tubérculos. O uso da distribuição lognormal para as concentrações afluentes de *E. coli*, N-NH<sub>3</sub> e valores de pH resultaram em concentrações efluentes de *E. coli* e N-NH<sub>3</sub> significativamente diferentes em relação ao uso da distribuição uniforme. Testes de análise de sensibilidade indicaram que as variáveis de maior impacto na previsão da concentração efluente de *E. coli* foram a concentração afluente de *E. coli* e o coeficiente de decaimento bacteriano. Na previsão da concentração efluente de N-NH<sub>3</sub>, as variáveis de maior impacto foram a concentração afluente de N-NH<sub>3</sub> e os valores de pH. Os resultados indicam, portanto, a necessidade de escolha criteriosa de valores de pH e de coeficiente de decaimento bacteriano para o dimensionamento de lagoas de polimento. Por fim, para efeito de estimativa de riscos de infecção decorrentes da irrigação de hortaliças ingeridas cruas, os dados de qualidade bacteriológica dos efluentes das lagoas foram utilizados em

modelos de avaliação quantitativa de risco microbiológico. Por meio de modelagem estocástica (amostragem aleatória por hipercubos latinos com 10.000 iterações), assumindo distribuição lognormal para a concentração de *E. coli* e nível de risco tolerável de  $10^{-3}$  pppa, padrões de  $10^3$  e  $10^4$  *E. coli*/100 mL pareceram adequados para a irrigação de hortaliças que se desenvolvem, respectivamente, rentes ao e distantes do solo. O uso da distribuição lognormal resultou em estimativas de risco significativamente menores que o uso da distribuição uniforme. Análise de sensibilidade apontou a relação patógenos/*E. coli* e a redução de patógenos entre a colheita e o consumo como as variáveis de maior impacto na estimativa de risco, indicando a necessidade de dados confiáveis desses parâmetros.

## ABSTRACT

PAVIONE, Damares Mirian de Souza, M.Sc., Universidade Federal de Viçosa, March of 2010.

**Uncertainty and sensitivity analyses and quantitative microbial risk analysis applied to polishing ponds design and to wastewater use for irrigation.** Adviser: Rafael Kopschitz Xavier Bastos. Co-Advisers: Paula Dias Bevilacqua and Paulo Roberto Ceccon.

This work presents the characterization and analysis of a database from an approximately seven-year period monitoring of a domestic wastewater treatment system composed of a UASB reactor, a submerged aerated biofilter (field scale) and a polishing pond series (pilot scale). The following variables were analyzed: BOD, COD, TSS, N-NH<sub>3</sub>, total phosphorus, pH, *E. coli*, total coliforms and helminth eggs. The hypothesis of data randomness, stationarity and homogeneity were, in general, confirmed, but data independency was not. Most data presented skewness and kurtosis different from those of the normal distribution. The results suggest that the lognormal function well represented the data distribution of BOD, COD, TSS, N-NH<sub>3</sub>, pH, *E. coli*, and total coliforms. Total phosphorus was better characterized by normal distribution, but the helminth eggs data could not be successfully adjusted to a distribution model. The lognormal distribution was applied to pH, pH, N-NH<sub>3</sub> and *E. coli* data as input variables in stochastic modeling (Latin hypercube sampling at 10,000 iterations) for the prediction of pond effluent quality. It is suggested that a three-pond series is sufficient to achieve an effluent discharge standard of 20 mg N-NH<sub>3</sub> L<sup>-1</sup>, as well as the WHO guidelines, in terms of *E.coli*/100 mL, for the irrigation of leaf and root crops. The use of the lognormal distribution for the pond influent values of *E. coli*, N-NH<sub>3</sub> and pH resulted in effluent concentrations significantly different from those when using the uniform distribution. Sensitivity analysis indicated that *E.coli* concentration and the bacterial decay coefficient were the input variables which had the greatest effect on the output estimates of *E.coli* concentrations. The estimates of N-NH<sub>3</sub> in pond effluents were influenced mostly by the N-NH<sub>3</sub> influent and the pH values. These results point out that pH values and the bacterial decay coefficient should be carefully chosen for an adequate pond design. Finally, in order to estimate infection risks arising from the consumption of wastewater-irrigated salad crops, the pond's bacterial effluent quality database was used in quantitative microbial risk analysis models. Using stochastic modeling (Latin hypercube sampling at 10,000 iterations), assuming the *E.coli* data according to a lognormal distribution, and a tolerable risk level of 10<sup>-3</sup> pppy, effluent quality standards of 10<sup>3</sup> and 10<sup>4</sup> *E. coli*/100 mL were shown to be adequate for the irrigation of, respectively, high- and low-growing crops. The use of the lognormal distribution, instead of a

uniform distribution, resulted in risk estimates significantly lower. Sensitivity analysis indicated that the pathogens: *E.coli* ratio and pathogen reduction between crops harvesting and consumption were the input variables which had the greatest effect on risk estimates, pointing out the need for reliable information on these variables.

## 1 - INTRODUÇÃO

Lagoas de estabilização são sistemas com ampla aplicação para o tratamento de esgotos em países de clima quente, em função de sua estreita dependência das condições climáticas, de sua robustez, simplicidade operacional e baixos custos de implantação e manutenção.

Os critérios de dimensionamento de lagoas encontram-se bem consolidados na literatura nacional e internacional. Tais critérios, bem como os de outros processos de tratamento, são tradicionalmente baseados em modelos determinísticos, isto é, em que as variáveis de entrada dos modelos de dimensionamento são representadas por valores únicos.

Entretanto, vários parâmetros de projeto de estações de tratamento de esgotos apresentam variações, por vezes amplas, devido à variabilidade e incertezas.

Nos modelos estocásticos, as variáveis de entrada são computadas em faixas de valores conforme determinada distribuição de probabilidades, dentre as quais se procede à amostragem aleatória, por exemplo, por meio de simulação de Monte Carlo. Dessa forma, os resultados são também expressos em termos de distribuição de probabilidades, o que proporciona flexibilidade na tomada de decisões. Embora ainda haja muito o que explorar, a abordagem estocástica aplicada ao tratamento de esgotos não constitui propriamente novidade. Porém, trabalhos com lagoas de estabilização são bem mais escassos.

A irrigação com esgotos sanitários é uma realidade e prática regulamentada em diversos países. Lagoas de estabilização são sistemas particularmente adequados à associação com o uso de efluentes tratados para irrigação, pois, além das características acima destacadas, apresentam elevada eficiência de remoção de organismos patogênicos.

A abordagem mais moderna para a formulação de normas de qualidade microbiológica de efluentes para irrigação envolve a metodologia de Avaliação Quantitativa de Risco Microbiológico e a modelagem estocástica para a estimativa de risco, incluindo também a simulação de Monte Carlo. Essa foi, por exemplo, a abordagem adotada na mais recente atualização das diretrizes da Organização Mundial da Saúde para o uso agrícola de esgotos sanitários.

Um aspecto fundamental da modelagem estocástica é o ajuste de funções densidade de probabilidades às variáveis de entrada (nos casos em questão, em modelos de dimensionamento de lagoas de estabilização e de estimativa de risco associado ao uso agrícola

de esgotos sanitários), o que requer a disponibilidade de bancos de dados confiáveis e a aplicação de testes estatísticos adequados.

A literatura internacional não deixa de registrar trabalhos sobre o comportamento (distribuição) de vários parâmetros constituintes dos esgotos sanitários, tanto em afluentes quanto em efluentes de estações de tratamento de esgotos. Entretanto, publicações nacionais - e mais especificamente de lagoas de estabilização - são bem escassas.

Nesta dissertação procurou-se contribuir com tais informações, com base em um vasto banco de dados gerado no monitoramento por cerca de sete anos de um sistema de lagoas em escala piloto em Viçosa - MG, implantado e consolidado como unidade experimental ao longo do Programa de Pesquisa em Saneamento Básico (PROSAB).

No capítulo 3 apresenta-se uma revisão de literatura, onde se procurou cobrir o conteúdo explorado nos capítulos seguintes. Assim, nos capítulos 4, 5 e 6, abaixo descritos, optou-se por incluir uma breve introdução, mas privilegiar a descrição dos aspectos metodológicos pertinentes e a apresentação dos respectivos resultados e discussões.

Além da identificação de distribuições de probabilidades que melhor descrevessem a variação dos constituintes dos efluentes desse sistema específico, o banco de dados foi complementarmente caracterizado por meio de ferramentas da estatística descritiva. Critérios estatísticos de aleatoriedade, independência, homogeneidade e estacionariedade dos dados foram também testados, pois, disso depende o rigor de uma análise de frequência. Esse conteúdo foi desenvolvido no capítulo 4 da presente dissertação.

Tal caracterização do referido banco de dados serviu de base para o desenvolvimento dos capítulos 5 e 6, respectivamente, sobre modelagem estocástica para a previsão de qualidade de efluentes de lagoas e de estimativa de risco em cenários hipotéticos de exposição à utilização agrícola dos efluentes do sistema experimental de lagoas.

No capítulo 5, o monitoramento das lagoas experimentais, e a própria configuração do sistema, forneceu dados de entrada (faixas de valores, distribuições de frequência e de probabilidades) aos modelos consagrados na literatura para a previsão da qualidade de efluentes das lagoas em termos de *E. coli* e amônia.

A modelagem estocástica permitiu a obtenção de resultados também em termos de distribuição de probabilidades. Testes de Análise de Sensibilidade foram realizados com o intuito de identificar as variáveis de entrada (os parâmetros de dimensionamento) que mais nitidamente influenciam a variável de saída (qualidade do efluente). Adicionalmente, testes de Análise de Cenários foram aplicados com o objetivo de identificar combinações de variáveis

de entrada que contribuem preponderantemente para um determinado resultado desejado, ou seja, para o atendimento de padrões de lançamento de efluentes em termos de amônia e de critérios de qualidade microbiológica de efluentes para irrigação.

No capítulo 6, os resultados obtidos no capítulo 4 foram utilizados em modelos de Avaliação Quantitativa de Risco Microbiológico.

Novamente, a modelagem estocástica forneceu resultados de estimativa de risco em termos de distribuição de probabilidades. Nesse caso, testes de Análise de Sensibilidade foram realizados com o intuito de identificar as variáveis de entrada (dos modelos de exposição e de dose-resposta) que mais nitidamente influenciam a variável de saída (estimativa de risco). Testes de Análise de Cenários foram aplicados com o objetivo de identificar quais variáveis de entrada contribuem preponderantemente para um determinado resultado desejado, no caso níveis de riscos tomados como referência.

## **2 - OBJETIVOS**

### **2.1 - Objetivo geral**

Caracterizar estatisticamente os efluentes de um sistema de lagoas de polimento de esgotos sanitários em termos de variáveis de qualidade física, química e microbiológica e testar sua aplicação na previsão da qualidade dos efluentes e em modelos de análise quantitativa de risco microbiológico associados ao uso agrícola dos efluentes.

### **2.2 - Objetivos específicos**

- Caracterizar a distribuição de frequência dos dados de qualidade dos efluentes das lagoas de polimento da Unidade Experimental da Violeira, Viçosa-MG.
- Investigar a aderência dos dados de qualidade dos efluentes das lagoas de polimento da Unidade Experimental da Violeira a diversas distribuições de probabilidades.
- Testar técnicas de modelagem estocástica com simulação de Monte Carlo na previsão da qualidade de efluentes de lagoas de polimento.
- Testar técnicas de modelagem estocástica com simulação de Monte Carlo em modelos de Avaliação Quantitativa de Risco Microbiológico aplicados em cenários de exposição ao uso agrícola de esgotos.
- Identificar variáveis chave de entrada em modelos de estimativa de qualidade de efluentes de lagoas de polimento.
- Identificar variáveis chave de entrada em modelos de estimativa de risco associados ao uso agrícola de efluentes de lagoas de polimento.

### 3 - REVISÃO DE LITERATURA

#### 3.1 - Análise de incertezas em modelos estocásticos

Modelos determinísticos são aqueles cujas variáveis são descritas por valores únicos, como, por exemplo, uma única medida ou a média ou um determinado quantil de uma distribuição de frequência. Modelos estocásticos são aqueles em que são levadas em consideração as incertezas sobre a variabilidade em torno das variáveis de entrada (MORGAN e HERION, 2007; VOSE, 2008).

Variabilidade é uma característica aleatória e resulta da diversidade natural em uma variável ou população. Por sua vez, incerteza é consequência da falta de conhecimento sobre um fenômeno ou uma variável. Portanto, a incerteza pode ser reduzida com melhor controle ou mais informações sobre uma variável (por exemplo, por meio da obtenção de mais dados), mas a variabilidade não, ainda que possa ser melhor caracterizada e também reduzida, mas por meio de melhor controle de processos (MORGAN e HERION, 2007).

Embora variabilidade e incerteza sejam conceitos distintos e expressem fontes de variação diferentes, em modelos estocásticos esse dois aspectos são em geral considerados de forma conjunta, sob a denominação genérica de *incertezas* ou *incerteza total* (VOSE, 2008). Modelos estocásticos são, portanto, aqueles em que os parâmetros de projeto são assumidos em faixas de valores e de acordo com uma determinada distribuição de frequência ou de probabilidades (MORGAN e HERION, 2007).

Adicionalmente, como forma de propagação das incertezas, estas são tratadas por meio de amostragens sucessivas e aleatórias de cada uma das variáveis de um determinado modelo, por exemplo, por meio de técnicas de simulação de Monte Carlo (VOSE, 2008). Dessa forma, os resultados são também expressos em termos de distribuição de probabilidades, o que proporciona flexibilidade na tomada de decisões.

Portanto, uma etapa fundamental na modelagem estocástica é a atribuição a mais criteriosa possível das incertezas associadas às variáveis de entrada, pois afinal, a função ajustada não passará de uma aproximação da real distribuição dos dados.

Para definir-se a distribuição de probabilidades pode-se recorrer aos bancos de dados disponíveis sobre a variável em questão, ao conhecimento sobre os fenômenos que determinam as variações, ou mesmo, no caso de ausência de maiores informações, à ‘opinião

especializada' (VOSE, 2008). No caso de bancos de dados, alguns requisitos adicionais devem ser observados: (i) o banco de dados deve constituir uma amostra representativa e de informações obtidas aleatoriamente; em outras palavras existem dados suficientes e o banco de dados cobre, por exemplo, variações sazonais? e (ii) o banco de dados deve ser confiável do ponto de vista analítico. Como bem observado por Finley *et al.* (1994), é importante que as funções ajustadas não sejam somente “matematicamente elegantes”, mas façam sentido do ponto de vista físico e satisfaçam critérios estatísticos.

Em geral, as distribuições de probabilidades descrevem a variabilidade dos parâmetros de um modelo. No que diz respeito às incertezas propriamente ditas, o ajuste de uma distribuição de probabilidades pode se tornar inviável ou mesmo incabível. A modelação em separado de variabilidade e incerteza é possível por meio dos chamados modelos de segunda ordem, nos quais a incerteza é também descrita por funções densidade de probabilidades, mas esta modelação é bem mais complexa. Por isso, de acordo com Vose (2008), tal esforço extra deve ser bem avaliado tendo em vista os que ganhos adicionais de informação que podem de fato ser obtidos.

### *3.1.1 - Caracterização da distribuição de probabilidades*

Uma investigação organizada de um conjunto de dados na busca de evidências e padrões empíricos da variabilidade pode ser realizada pela aplicação da estatística em um estágio descritivo e/ ou pela aplicação da teoria de probabilidades e dos métodos de inferência estatística que estabelecem um padrão de variabilidade da população onde foi extraída a amostra.

A estatística descritiva de uma amostra extraída da população pode ser sumariada por certas características das distribuições de frequência, denominadas medidas. Essas medidas podem ser de posição, de dispersão, de assimetria e achatamento ou curtose. Por sua vez, a variável aleatória é uma função que associa números reais aos eventos de um espaço amostral. Estas podem ser discretas, contínuas ou mistas.

O comportamento de uma variável aleatória contínua é descrito por sua função densidade de probabilidades (FDP) ou pela sua função distribuição acumulada (FDA), que é uma função contínua, integralmente conhecida sob o ponto de vista estatístico<sup>1</sup>.

---

<sup>1</sup> Em teoria das probabilidades, a ‘função probabilidade’ descreve a probabilidade correspondente a qualquer intervalo, determinando, de maneira única, a distribuição da variável aleatória  $X$  no intervalo  $a, b$ . As distribuições

Analogamente, em uma distribuição discreta o comportamento das variáveis aleatórias é descrito por sua função massa de probabilidades (FMP) ou pela sua função massa acumulada (FMA), onde cada variável é associada a um valor correspondente à sua probabilidade. Essas variáveis são obrigatoriamente um conjunto enumerável (COSTA NETO, 1977).

Conforme exposto por Naghettini e Pinto (2007), quando são conhecidos os modelos distributivos e os valores numéricos dos parâmetros que definem uma variável aleatória, podem ser calculadas as probabilidades associadas a quaisquer eventos definidos pelos valores da variável em questão. Ou seja, o modelo de distribuição de probabilidades pode constituir uma síntese plausível do comportamento de uma variável aleatória e ser empregado para interpolar ou extrapolar probabilidades, ou quantis não contidos na amostra de observações.

#### 3.1.1.1 - Distribuição de probabilidades usuais

Resumidamente, uma distribuição de probabilidades descreve a chance que uma variável pode assumir ao longo de um espaço de valores. Para determinar o comportamento de uma variável aleatória contínua e extrair conclusões de uma amostra de dados ambientais é necessário estabelecer um modelo matemático contendo elementos essenciais do processo que determina a ocorrência das observações. Tal modelo deve ser probabilístico, pois é impossível sintetizar em equações uma lei que descreva a variação de um fenômeno (NAGHETTINI e PINTO, 2007).

As funções de distribuição podem ter curvatura rígida ou flexível (exceto a distribuição uniforme, que não possui curvatura). As de curvatura rígida são aquelas que descrevem tendências definidas e apresentam ponto de inflexão constante. As de curvatura flexível, ao invés de descreverem apenas uma curva característica, descrevem famílias de curvas de diferentes pontos de inflexão.

A seguir é apresentada uma breve descrição de algumas das distribuições de probabilidades mais relevantes para variáveis ambientais contínuas e discretas. São descritas as distribuições normal, lognormal, Weibull, exponencial, uniforme, Poisson, binomial negativa e geométrica, sendo as três últimas distribuições discretas e as demais contínuas. As

---

são ditas contínuas se a função de distribuição correspondente puder ser representada por uma integral, neste caso, utiliza-se a 'função densidade de probabilidade' para representar a distribuição de probabilidade. Variável aleatória discreta é aquela cujo espaço amostral é do tipo numerável, composto por valores isolados, em geral, números inteiros. (CHRISTMANN, 1978)

distribuições normal e exponencial possuem curvas rígidas, enquanto a lognormal e Weibull, curvas flexíveis. As distribuições de curva rígida possuem coeficientes de assimetria e curtose fixos e somente um parâmetro para a definição de sua escala (ver Quadro 3.1).

### Distribuição normal

É a mais importante entre as distribuições de probabilidade, porque muitas variáveis aleatórias de interesse prático são normais ou podem ser transformadas para normais de maneira relativamente simples (CHRISTMANN, 1978).

A distribuição normal é inteiramente descrita pela média e pelo desvio padrão, ou seja, conhecendo-se esses parâmetros consegue-se determinar qualquer probabilidade da distribuição. A função densidade de probabilidades da distribuição normal com média  $\mu$  e variância  $\sigma^2$  (de forma equivalente, desvio padrão  $\sigma$ ) é definida pela Equação 3.1 e sua probabilidade acumulada pela Equação 3.2. Funções densidade de probabilidades (FDPs) e funções distribuição acumulada (FDAs) da distribuição normal estão ilustradas na Figura 3.1.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \infty \geq x \geq -\infty \quad (3.1)$$

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad (3.2)$$

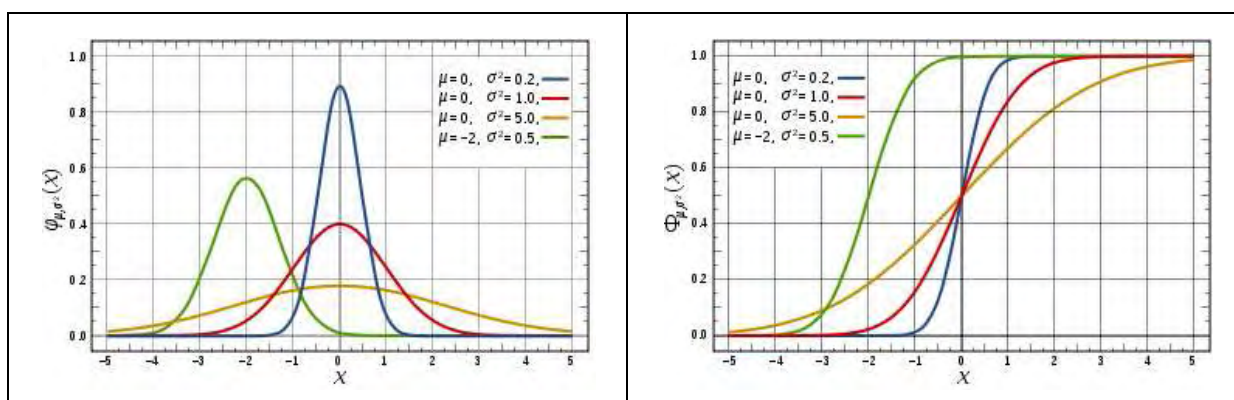


Figura 3.1 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição normal para  $\mu=0$  e diferentes valores de  $\sigma$

Fonte: <http://www.wikipedia.org>

Embora, como referido, várias características naturais sejam frequentemente descritas por distribuição normal, Berthouex e Hunter (1983) destacam que muitos conjuntos de dados ambientais violam condições da distribuição normal (por exemplo, requisitos de que os dados

sejam aleatórios independentes e simetricamente distribuídos), por incluírem amplas incertezas e variabilidade, tais como, valores extremos, problemas da amostragem e medição e complexas relações de causa e efeito.

### Distribuição lognormal

Uma variável aleatória  $X$  tem distribuição lognormal quando o seu logaritmo [ $Y=\ln(X)$ ] apresenta distribuição normal. Portanto, a distribuição lognormal é caracterizada pela média e desvio-padrão de  $Y=\ln(X)$  e sua função de densidade probabilidades e distribuição acumulada (Figura 3.2) são dadas pelas Equações 3.3 e 3.4.

$$f(x; \mu; \sigma) = \frac{1}{x \cdot \sigma \cdot \sqrt{2\pi}} \cdot \exp \left[ -\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right] \quad 0 \leq x \leq \infty \quad (3.3)$$

$$F_X = \Phi \cdot \left( \frac{\ln x - \mu}{\sigma} \right) \quad 0 \leq x \leq \infty \quad (3.4)$$

Onde  $\Phi$  é a função distribuição acumulada normal.

Como o coeficiente de variação é sempre maior que zero, a distribuição lognormal é sempre assimetricamente positiva, com coeficiente de assimetria proporcional ao coeficiente de variação (NAGHETTINI e PINTO, 2007).

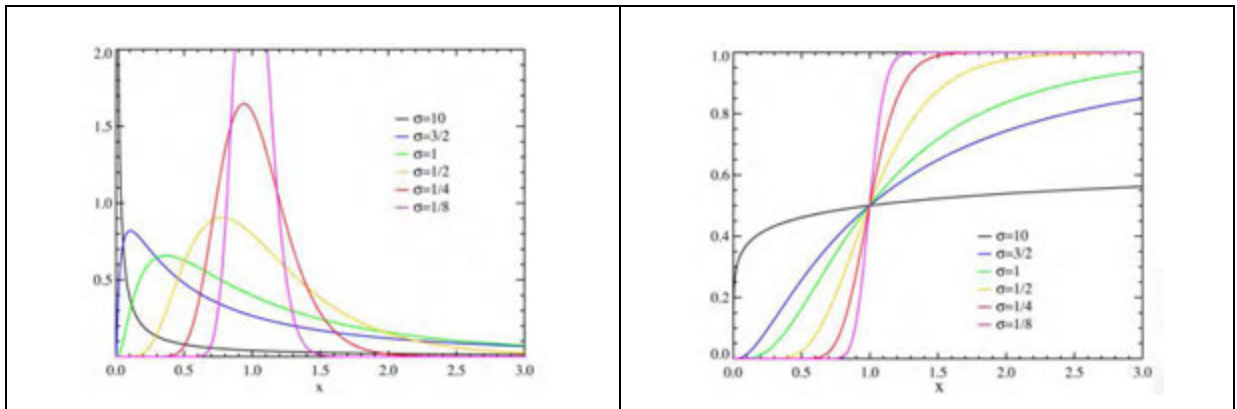


Figura 3.2 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição lognormal para  $\mu=0$  e diferentes valores de  $\sigma$

Fonte: <http://www.wikipedia.org>

A função lognormal, em geral, descreve bem a distribuição de variáveis físicas que não assumem valores negativos e se apresentam positivamente assimétricas, tais como concentrações de poluentes e dados de vazões. É ainda apropriada para descrever variáveis que apresentem incertezas amplas, expressas, por exemplo, em bases multiplicativas ou em

termos de ordens de magnitude<sup>2</sup>. Portanto, não somente é de se esperar, mas de fato tem sido comprovado, que grande parte dos constituintes dos esgotos possa ser modelada pela distribuição lognormal, incluindo variáveis químicas e microbiológicas, tanto no afluente quanto no efluente de ETEs (BERTHOUEX e HUNTER, 1981; TANAKA *et al.*, 1998; METCALF & EDDY, 2003; OLIVEIRA, 2006).

### *Distribuição exponencial*

A distribuição exponencial é um tipo de distribuição contínua de probabilidade, representada por somente um parâmetro ( $\lambda$ ). Suas FDP e FDA são dadas respectivamente pelas Equações 3.5 e 3.6, ilustradas na Figura 3.3.

$$f(x) = \lambda \cdot e^{-\lambda x} \quad 0 \leq x \leq \infty \quad (3.5)$$

$$F(x; \lambda) = 1 - e^{-\lambda x} \quad 0 \leq x \leq \infty \quad (3.6)$$

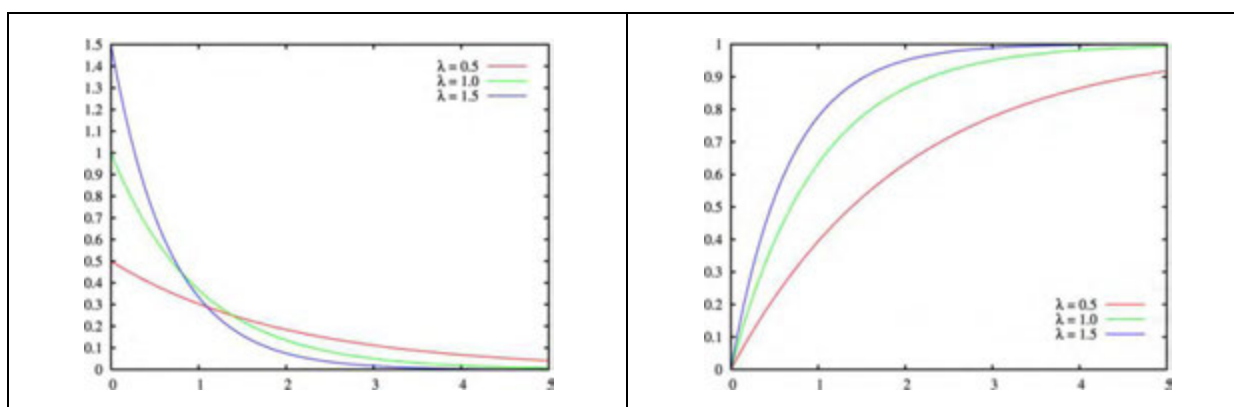


Figura 3.3 – Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição exponencial para diferentes valores de  $\lambda$

Fonte: <http://www.wikipedia.org/>

Se a probabilidade de um evento ocorrer num pequeno intervalo de tempo for muito pequena, mas estatisticamente independente da ocorrência de outros eventos, então o intervalo de tempo entre as ocorrências deste evento será exponencialmente distribuído (por exemplo, intervalos de tempo entre acidentes, e de alguns eventos ambientais, como durações de tempestades). O parâmetro  $\lambda$  é definido como o inverso do tempo médio entre eventos e, portanto, representa a taxa de um processo (CHRISTMANN, 1978).

<sup>2</sup> ‘Ordem de magnitude’ ou ‘ordem de grandeza’ de um número é a potência de dez desse número. A ordem de magnitude de um número pode ser definida como seu logaritmo decimal.

### Distribuição uniforme ou retangular

É a mais simples das funções contínuas de probabilidade. A probabilidade de se gerar qualquer ponto em um intervalo contido no espaço amostral é proporcional ao tamanho do intervalo. Suas respectivas FDP e FDA são representadas pelas Equações 3.7 e 3.8 e ilustradas pela Figura 3.4.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b \\ 0 & \text{para } a < x \text{ ou } x > b \end{cases} \quad (3.7)$$

$$F(x) = \begin{cases} 0 & \text{para } a > x \\ \frac{x-a}{b-a} & \text{para } a \leq x \leq b \\ 1 & \text{para } x > b \end{cases} \quad (3.8)$$

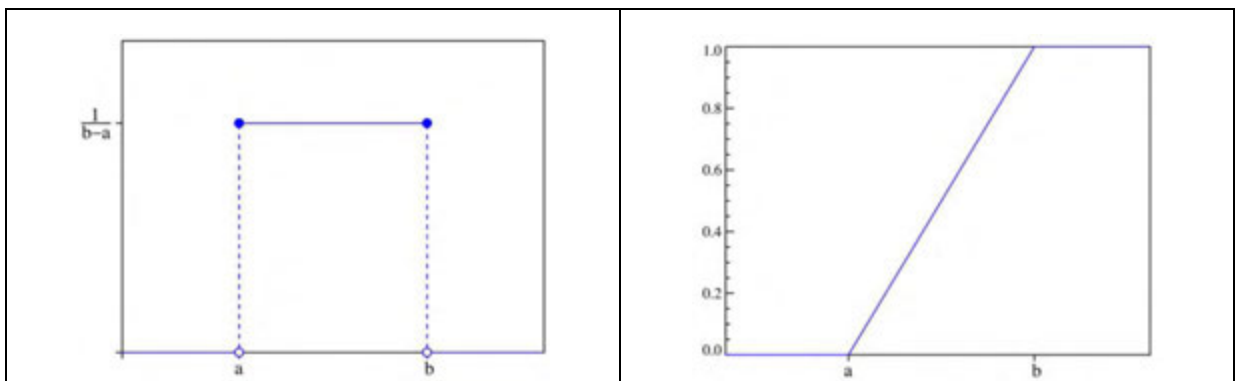


Figura 3.4 - Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição uniforme

Fonte: <http://www.wikipedia.org>

### Distribuição Weibull

Os parâmetros na distribuição Weibull fornecem grande flexibilidade para modelar sistemas em que o número de falhas aumenta, diminui ou permanece constante com o tempo (MONTGOMERY, 2002). As FDP e FDA são apresentadas nas Equações 3.9 e 3.10 e ilustradas na Figura 3.5.

$$f(x; k; \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \cdot e^{-(x/\lambda)^k} \quad x > 0; \lambda > 0; k > 0 \quad (3.9)$$

$$F(x; k; \lambda) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k} \quad x > 0; \lambda > 0; k > 0 \quad (3.10)$$

A distribuição Weibull tem uma importante propriedade, a de se transformar em outras distribuições conforme os valores de seus parâmetros. Por exemplo, se  $k$  for igual a 1, a distribuição passa a ser exponencial e, quanto mais próximo a 3, se aproxima da normal.

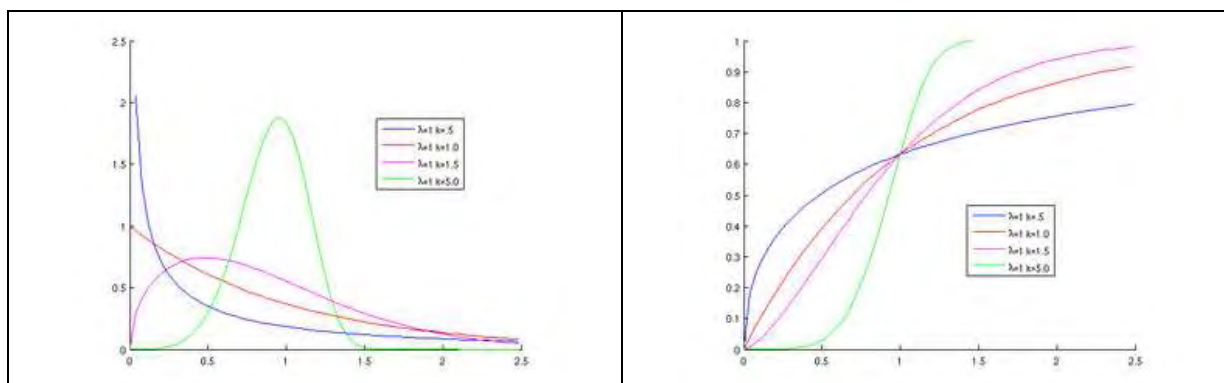


Figura 3.5 – Ilustrações de funções densidade de probabilidades (à esquerda) e funções distribuição acumulada (à direita) da distribuição Weibull para  $\lambda=1$  e diferentes valores de  $k$

Fonte: <http://www.wikipedia.org>

### *Distribuição Poisson*

Os processos de Poisson estão entre os mais importantes processos estocásticos. O parâmetro  $\nu$  representa o número médio de ocorrências de Poisson em um intervalo de tempo  $t$ ; a constante de proporcionalidade  $\gamma$  é denominada intensidade de Poisson e representa a razão média de ocorrência dos eventos por intervalo de tempo. Uma característica dessa distribuição é que a média e a variância de uma variável discreta são iguais (NAGHETTINI e PINTO, 2007). As expressões de FMP e FMA que definem a distribuição de Poisson são apresentadas pelas Equações 3.11 e 3.12 respectivamente. A Figura 3.6 ilustra a função massa probabilidade.

$$p(y) = \frac{\nu^k e^{-\nu}}{y!} \quad (y = 0, 1, 2, \dots, n) \quad (3.11)$$

$$P(y) = \sum_{i=0}^k \frac{\mu^i}{i!} e^{-\mu} \quad (k = 0, 1, 2, \dots, n) \quad (3.12)$$

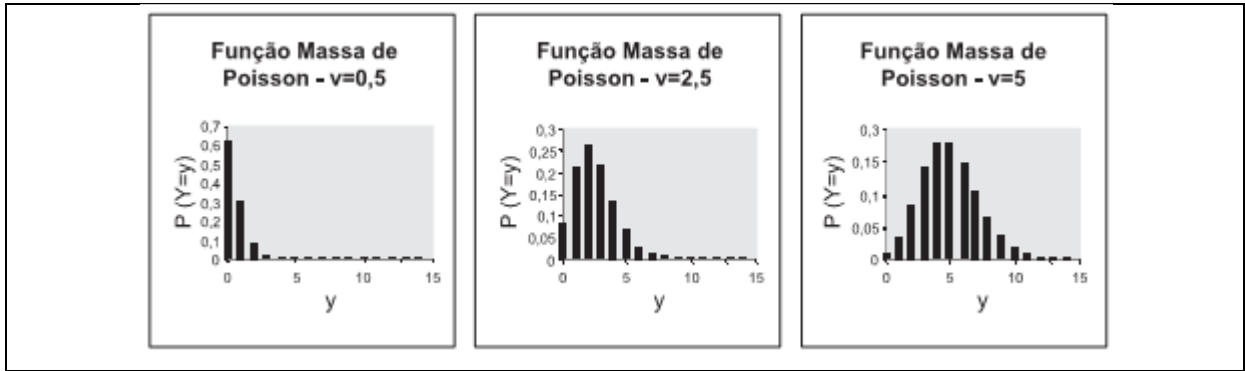


Figura 3.6 – Ilustração de funções massa de probabilidades de Poisson para diferentes valores de  $v$

Fonte: Naghettini e Pinto (2007)

*Distribuição Binomial negativa*

Uma variável apresenta distribuição binomial negativa quando  $Y$  refere-se ao número de repetições independentes necessárias para que um certo número de ‘sucessos’ ocorram. A distribuição é expressa pela FMP e pela FMA representadas pelas Equações 3.13 e 3.14. A Figura 3.7 ilustra a FMP.

$$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r} \quad (y = r, r+1, r+2, \dots) \quad (3.13)$$

$$P(y) = \binom{y-1}{r-1} p^{r-1} (1-p)^{y-r} \quad (y = r, r+1, r+2, \dots) \quad (3.14)$$

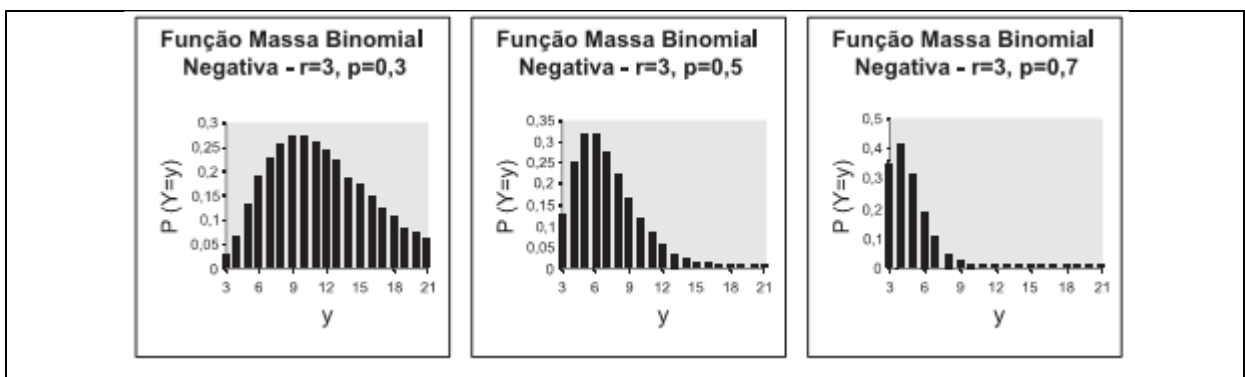


Figura 3.7 – Ilustração de funções massa de probabilidades binomial negativa para diferentes valores de  $r$  e  $p$

Fonte: Naghettini e Pinto (2007)

## Distribuição Geométrica

A variável geométrica  $Y$  está associada ao número de experimentos (ou tentativas) necessários para que um único ‘sucesso’ ocorra. A distribuição é expressa pelas Equações 3.15 e 3.16 e ilustrada na Figura 3.8.

$$p(y) = p(1 - p)^{y-1} \quad (y = 1, 2, 3 \dots \quad 0 < p < 1) \quad (3.15)$$

$$P(y) = \sum_{i=0}^y p(1 - p)^{i-1} \quad (y = 1, 2, 3 \dots) \quad (3.16)$$

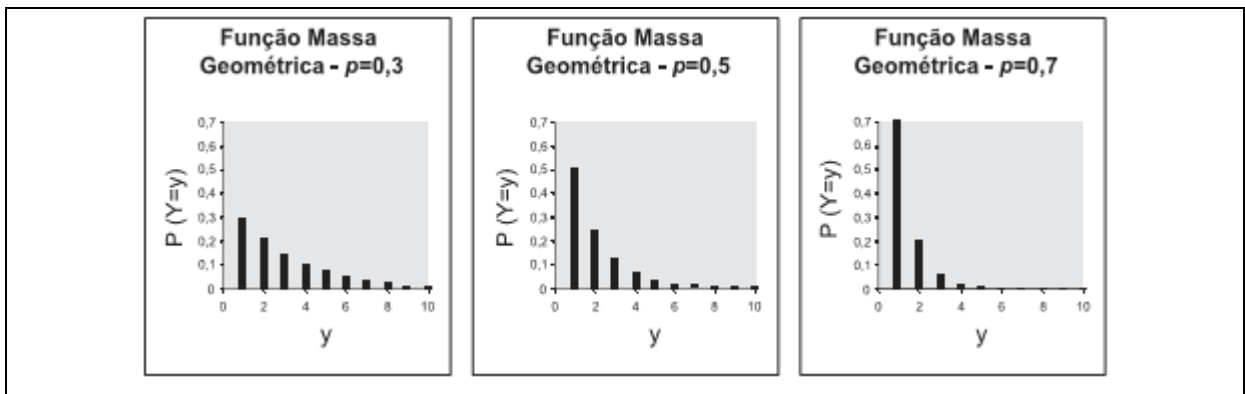


Figura 3.8 – Ilustração de funções massa de probabilidades geométrica para diferentes valores de  $p$

Fonte: Naghettini e Pinto (2007)

Embora, como referido, várias características naturais sejam frequentemente descritas por distribuição normal, Berthouex e Hunter (1983) destacam que muitos conjuntos de dados ambientais violam condições da distribuição normal (por exemplo, requisitos de que os dados sejam aleatórios independentes e simetricamente distribuídos), por incluírem amplas incertezas e variabilidade, tais como, valores extremos, problemas da amostragem e medição e complexas relações de causa e efeito.

### 3.1.1.2 – Testes de aderência e verificação gráfica da distribuição

Os testes de aderência testam hipóteses referentes à forma da distribuição da população. Nesses testes, admite-se, por hipótese, que a distribuição da variável de interesse na população seja descrita por um determinado modelo de distribuição de probabilidades, sendo então verificada a boa ou má aderência da amostra ao modelo. A seguir são descritos alguns testes de aderência usuais.

O teste Qui-quadrado ( $\chi^2$ ) estabelece o grau de correspondência entre as frequências observadas e as esperadas em cada categoria pré-estabelecida. Esse teste é baseado na divisão do banco de dados em intervalos de classes para o cálculo do parâmetro  $\chi^2$  conforme Equação 3.17 (SIEGEL e CASTELLAN, 2006).

$$\chi^2 = \sum_{i=1}^k \frac{(N_i - E_i)^2}{E_i} \quad (3.17)$$

Onde:

$\kappa$  = número de intervalos de classe (NIC);

$N_i$  = número de amostras no intervalo de classes;

$E_i$  = número esperado de amostras no intervalo de classes.

Uma das fragilidades do teste do  $\chi^2$  é a não existência de regra padrão para a determinação do NIC e essa definição influencia os resultados. Alguns *softwares* trazem um NIC pré-definido, por exemplo, próximo da raiz quadrada do número de dados (PALISADE CORPORATION, 2009).

O teste de aderência de Kolmogorov-Smirnov (KS) tem como base a diferença máxima entre as funções de probabilidades acumuladas, empírica e teórica, de variáveis aleatórias contínuas, conforme a Equação 3.18 (SIEGEL e CASTELLAN, 2006).

$$D_n = \text{máximo} |S_m(x) - S_n(x)| \quad (3.18)$$

Onde:

$S_m(x)$  = é a distribuição acumulada de uma amostra de tamanho  $m$ ;

$S_n(x)$  = é a distribuição acumulada da outra amostra de tamanho  $n$ .

O teste KS trata observações individuais separadamente e assim não perde informação devido ao agrupamento, como acontece algumas vezes com o teste Qui-quadrado. É, portanto, um teste de aderência mais poderoso (ou menos arbitrário) que o teste do Qui-quadrado. Por outro lado, o poder dos testes de aderência Qui-quadrado e Kolmogorov-Smirnov de discriminar entre hipóteses falsas e verdadeiras é bastante diminuído nas caudas inferior e superior, tanto em função do reduzido número de observações amostrais, quanto em decorrência dos maiores erros de estimação nessas partições do espaço amostral (SIEGEL e CASTELLAN, 2006).

O teste de aderência de Anderson-Darling (AD) (Equação 3.19) é um teste não-paramétrico que procura ponderar mais fortemente as caudas das distribuições, nas quais, as maiores (ou as menores) observações da amostra podem alterar sobremaneira a qualidade do ajuste (SIEGEL e CASTELLAN, 2006).

$$A^2 = -N - \sum_{i=1}^N \frac{(2i-1)\{\ln f_x(x_i) + [1-f_x(x_{N-1+1})]\}}{N} \quad (3.19)$$

Onde:

$N$  = número de dados;

$f(x)$  = distribuição probabilidades hipotética;

$x_1; x_2; \dots; x_N$  = observações ordenadas de modo crescente.

Assim como o teste KS, o teste de AD não envolve a divisão dos dados em intervalos de classe, mas enquanto o KS tem foco no centro da distribuição, o AD destaca diferenças entre as caudas das distribuições ajustadas e dos dados de entrada.

O teste de aderência de Filliben (FIL) baseia-se na ideia de que a eventual existência de uma forte associação linear entre  $x_i$  (dados da amostra) e  $w_i$  (quantis da distribuição teórica) é um indicador de que as observações podem, de fato, terem sido extraídas de uma mesma população. Apresenta, como principais vantagens, a simplicidade de construção de sua estatística de teste. Formalmente, a estatística do teste de Filliben é expressa pela Equação 3.20.

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(w_i - \bar{w})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (w_i - \bar{w})^2}} \quad (3.20)$$

A decisão de não rejeitar a hipótese nula  $H_0$ , a um determinado nível de significância, não implica estabelecer a ideia de que os dados foram, de fato, amostrados a partir da população hipotética. Essa é, por princípio, desconhecida e pode ser uma entre tantas outras populações incluídas na hipótese alternativa  $H_1$ . Por outro lado, embora os testes de aderência não se prestem rigorosamente à seleção de uma dentre várias distribuições, são instrumentos da estatística que auxiliam a tomada de decisão quanto à adequação ou inadequação de certo modelo distributivo a uma dada amostra.

O teste de Shapiro-Wilk é um teste estatístico cujo objetivo é avaliar a suposta normalidade de uma amostra. Segundo Shapiro e Wilk (1965), é eficaz, mesmo para pequenas amostras ( $N < 20$ ), e sensível a uma ampla variedade de aspectos da não normalidade, principalmente à assimetria e caudas alongadas.

Nesse teste, a rejeição ou aceitação de  $H_0$  está vinculada ao valor de  $W$ , calculado segundo a Equação 3.21.

$$W = \frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.21)$$

Onde:

$N$  = número de dados;

$x_1; x_2; \dots; x_N$  = observações ordenadas de modo crescente;

$a_i$  = constante gerada pela média, variância e covariância de uma amostra de distribuição Normal.

Processos gráficos também podem auxiliar na verificação da aderência dos dados experimentais a certo modelo teórico. São processos simplificados utilizados como reforço aos testes estatísticos para escolha da melhor distribuição de frequência. Nessa análise gráfica, a associação das observações ordenadas às respectivas probabilidades empíricas apresenta consideráveis incertezas que dependem, principalmente, do tamanho e da representatividade da amostra. O teste gráfico de aderência é feito mediante plotagem dos dados em um papel de probabilidades, o qual possui uma das escalas subdividida conforme os percentis da distribuição em questão. Se a série de dados tiver a mesma distribuição de frequência do papel de probabilidades os dados formarão uma reta. A verificação visual da linearidade dos dados amostrais é aceita ou rejeitada empiricamente (SIEGEL e CASTELLAN, 2006; NAGHETTINI e PINTO, 2007; COSTA NETO, 1977; MONTGOMERY, 2002). A Figura 3.9 ilustra o uso do teste gráfico aplicado à verificação de aderência às distribuições normal e lognormal.

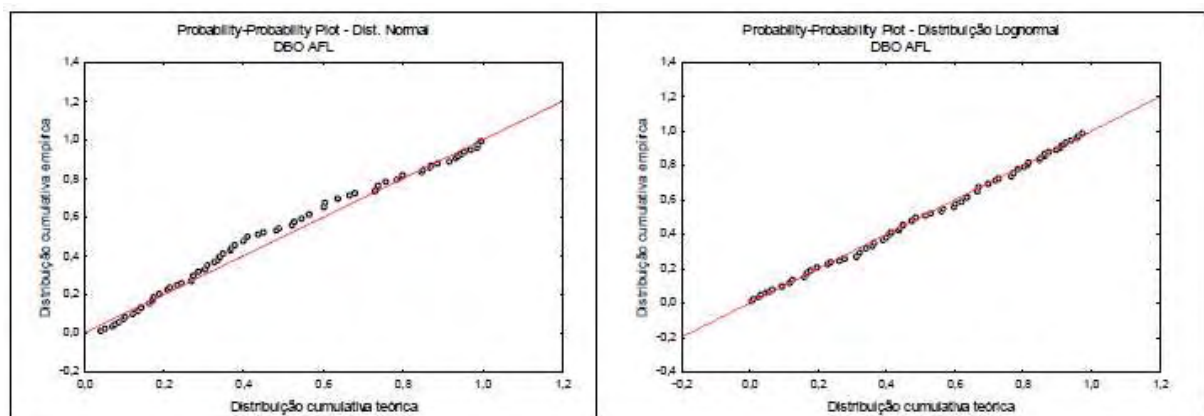


Figura 3.9 – Ilustrações de teste gráfico de aderência às distribuições normal (à esquerda) e lognormal (à direita)

Fonte: Olivera (2006)

### 3.1.1.3 – Estatística descritiva

A estatística descritiva utiliza medidas que sintetizam o padrão de distribuição de uma variável e permitem extrair da amostra as informações necessárias para inferir o comportamento populacional. As medidas podem ser agrupadas em: (a) de tendência central

ou de posição (b) de dispersão e (c) de assimetria e de curtose (NAGHETTINI e PINTO, 2007).

### *Medidas de tendência central ou medidas de posição*

A medida de posição serve para localizar a distribuição de frequência sobre o eixo de variação da variável em questão. As medidas de tendência central mais usuais são a *média*, a *mediana* e a *moda*. Pode-se ainda caracterizar a distribuição acumulada de frequência por meio dos *quantis*, como por exemplo, os *quartis*, *decis* ou *percentis*. A escolha entre tais medidas depende do uso pretendido do valor central (COSTA NETO, 1977).

A média aritmética é a mais utilizada, mas apresenta a desvantagem de ter seu valor afetado pela eventual presença de valores atípicos (extremos); é, portanto, mais apropriada para a descrição de amostras que apresentem distribuição normal. Por sua vez, a média geométrica é mais adequada para estimar o valor central de variáveis que possuem desenvolvimento geométrico, ou seja, aquelas cujos valores sucessivos guardam entre si um fator de crescimento ou decrescimento, como, por exemplo, o aumento populacional ou de cargas orgânicas afluentes à uma estação de tratamento de esgotos.

A média geométrica tem aplicação na determinação do ‘valor esperado’ (tendência central) de dados de amostras ambientais, dentre os quais de afluentes e efluentes de estação de tratamento de esgotos, tais como DBO, SS e coliformes, os quais têm sido descritos por distribuição lognormal. Com relação aos coliformes, o emprego da média geométrica encontra-se bem consolidado, pois as amostras podem apresentar variações de diversas ordens de magnitude e um único valor extremamente elevado pode elevar substancialmente a média aritmética, mesmo que todos os outros valores sejam baixos (BERTHOUEX e HUNTER, 1981, 1983; METCALF e EDDY, 2003; OLIVEIRA, 2006).

Não obstante, em modelos de Análise Quantitativa de Risco Microbiológico (AQRM), tem sido demonstrado que a média aritmética é o melhor descritor da tendência central de banco de dados microbiológicos (organismos patogênicos e indicadores) (HAAS, 1996; BENKE e HAMILTON, 2008).

A mediana representa o valor que separa a metade inferior da amostra, população ou distribuição probabilidade, da metade superior. É uma medida de posição mais ‘resistente’, por ser imune à eventual presença de valores extremos discordantes na amostra.

A moda é o valor que detêm o maior número de observações, ou seja, o valor ou valores mais frequentes. Ao contrário da média e da mediana, a moda não é representada necessariamente por valor único.

### *Medidas de dispersão, assimetria e curtose*

As medidas de dispersão servem para indicar o quanto os dados se apresentam dispersos em torno da região central. Caracterizam o grau de variação existente em um conjunto de valores.

Exemplos de medidas dispersão utilizadas frequentemente são a amplitude, a distância interquartílica, a variância, o desvio padrão e o coeficiente de variação.

As medidas de assimetria, conforme sejam positivas, negativas ou aproximadamente nulas, procuram indicar o afastamento de uma distribuição da simetria. O coeficiente de assimetria assume valores próximos a zero quando a distribuição tem uma curva simétrica. Distribuições alongadas à direita são ditas positivamente assimétricas, e as alongadas à esquerda, negativamente assimétricas.

Em distribuições assimetricamente positivas, os dados encontram-se agrupados mais à esquerda da moda, com a maioria dos valores extremos situados à direita. Nesse caso, o valor da média é superior à mediana, que, por sua vez, excede o da moda. Quando a distribuição é assimétrica à esquerda, pode haver uma tendência de existência de valores negativos (COSTA NETO, 1977; MONTGOMERY, 2002).

Teoricamente, assimetrias positivas são esperadas para dados relativos a processos de tratamento de esgotos, já que não existem valores negativos para concentrações afluentes ou efluentes e não existem limites superiores para estes dados (OLIVEIRA, 2006).

O coeficiente de curtose mede a curvatura ou o grau de achatamento da curva da função de distribuição. Entretanto, a caracterização do achatamento de uma distribuição só tem sentido, em termos práticos, se a distribuição for pelo menos aproximadamente simétrica.

Os valores da curtose são iguais a três para distribuições mesocúrticas, as quais tem o mesmo achatamento que a distribuição normal. Valores maiores que três caracterizam as distribuições leptocúrticas, as quais se apresentam de forma mais alta (afunilada) e concentrada que a distribuição normal. Se o valor da curtose é menor que três, a função é dita platicúrtica, ou mais “achatada” que a distribuição normal (Figura 3.10) (SPIEGEL, 1993).

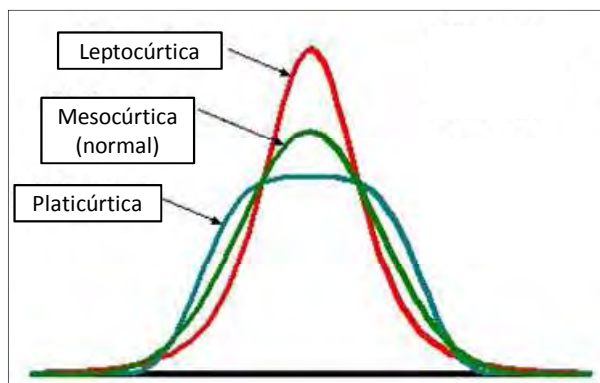


Figura 3.10 – Ilustração de curva leptocúrtica (em vermelho), mesocúrtica (em verde) e platicúrtica (em azul)

Fonte: <http://www.wikipedia.org>

O Quadro 3.1 apresenta as características de curtose e assimetria das distribuições de probabilidades descritas acima. As distribuições com curvas fixas possuem curtose e coeficiente de assimetria representados por valores únicos, enquanto as de curvas flexíveis, valores variados.

Quadro 3.1 – Curtose e assimetria de algumas distribuições de probabilidades

Distribuição	Curtose - $\kappa$	Coefficiente de assimetria - $\gamma$
Normal	$k = 3$	$\gamma = 0$
Lognormal	$k = 3 + (e^{\sigma^2} - 1)(e^{3\sigma^2 \ln x} + 3e^{2\sigma^2 \ln x} + 6e^{\sigma^2 \ln x} + 6)$	$\gamma = 3CV_x + CV_x^3$ $CV_x = \sqrt{\exp(\sigma_{\ln x}^2) - 1}$
Weibull	$\kappa = \frac{\lambda^4 \Gamma(1 + \frac{4}{\kappa}) - 4\gamma \sigma^3 \mu - 6\mu^2 \sigma^2 - \mu^4}{\sigma^4}$	$\gamma = \frac{\Gamma(1 + \frac{3}{\kappa}) \lambda^3 - 3\mu \sigma^2 - \mu^3}{\sigma^3}$
Exponencial	$k = 9$	$\gamma = 2$
Uniforme	$k = 1,8$	$\gamma = 0$
Poisson	$\kappa = 3 + \frac{1}{\mu}$	$\gamma = \sqrt{\frac{1}{\mu}}$
Binomial negativa	$\kappa = 3 + \frac{p^2 - 6p + 6}{r(1-p)}$	$\gamma = \frac{2-p}{\sqrt{r(1-p)}}$
Geométrica	$\kappa = 3 + \frac{p^2 - 6p + 6}{1-p}$	$\gamma = \frac{2-p}{\sqrt{1-p}}$

Fontes: Adaptado Naghettini e Pinto (2007); Haas (1997)

Os coeficientes de assimetria e curtose podem ser utilizados como estatísticas preliminares para a verificação das distribuições dos dados. Oliveira (2006), em seu estudo de análise de desempenho e confiabilidade de ETEs, notou que tanto para as variáveis afluentes quanto para as efluentes a maioria dos valores dos coeficientes de assimetria foi positiva,

indicando distribuições com caudas à direita. Do mesmo modo, grande parte dos valores de curtose foi diferente de três, indicando que o grau de achatamento das distribuições dos dados diferia do observado para uma distribuição normal.

#### 3.1.1.4 - Representação gráfica da distribuição de frequências

O *histograma* é uma representação da distribuição de frequência na forma gráfica por meio de retângulos justapostos em que a base de cada um deles corresponde ao intervalo de classe e a sua altura à respectiva frequência (Figura 3.11). A curva de frequências é formada unindo-se os pontos médios dos patamares dos retângulos do histograma. No caso limite de uma amostra de tamanho infinito, esta curva tornar-se-ia a função densidade de probabilidades da população.

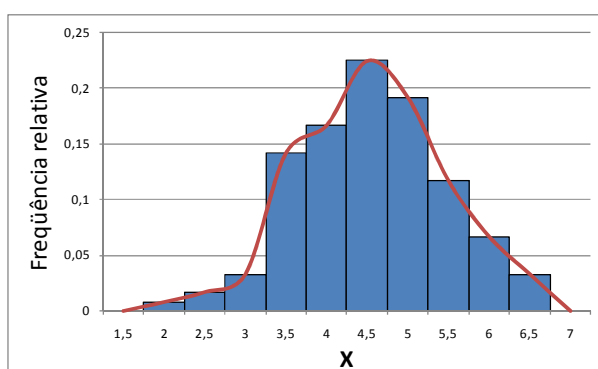


Figura 3.11 - Ilustração de um histograma com a respectiva curva de frequência

Fonte: Adaptado Hoffmann (1991)

#### *3.1.2 - Critérios estatísticos para a análise de frequência*

Segundo Naghettini e Pinto (2007), para que os resultados de uma análise de frequência (incluindo os testes de hipóteses, a estimação de parâmetros, o cálculo de quantis e os testes de aderência) sejam teoricamente válidos, a série de dados analisada deve atender a certos critérios estatísticos, tais como *aleatoriedade, independência, homogeneidade e estacionariedade*. Testar essas hipóteses significa recolher evidências nos dados amostrais que justifiquem a rejeição ou não rejeição de certa afirmação sobre um parâmetro populacional ou sobre a forma de um modelo distributivo, tendo em conta a probabilidade de serem tomadas decisões incorretas.

Por não se conhecer a forma específica da distribuição das séries de dados devem ser utilizados métodos e testes estatísticos que independem das distribuições populacionais e dos parâmetros associados, ou seja, métodos não paramétricos (SIEGEL e CASTELLAN, 2006).

A *aleatoriedade* dos dados é testada para confirmar se as flutuações de certa variável decorrem de causas naturais. O teste de *independência* verifica se qualquer observação presente na amostra influencia a ocorrência de outra observação seguinte (mesmo que uma série seja considerada aleatória, as observações que a constituem podem não ser independentes). O teste de *homogeneidade* visa constatar se todos os elementos da amostra são oriundos de uma única e idêntica população. *Estacionariedade* refere-se ao fato de que, excluídas as flutuações aleatórias, as observações amostrais são invariantes com relação à cronologia de suas ocorrências. A não estacionariedade inclui tendências, ‘saltos’ e ciclos, ao longo do tempo (NAGHETTINI e PINTO, 2007).

No entanto, Oliveira (2006), citando Bobée e Ashkar (1991), relata que o não atendimento a estas hipóteses fundamentais da análise de frequência não invalida o emprego dos testes estatísticos, pois estes não dão respostas inequívocas e podem apenas indicar a probabilidade de satisfação dos critérios.

### 3.1.3 - Amostragem aleatória das variáveis de entrada em modelos estocásticos

Simulação de Monte Carlo é uma técnica de propagação de incertezas agregadas às variáveis de um modelo, com vasto uso nas áreas de engenharia e finanças, particularmente em análise de risco. O método teve origem em pesquisas conduzidas em Los Alamos, EUA, na solução de problemas matemáticos cujo tratamento analítico não era viável, contribuindo com a construção da primeira bomba atômica. O nome Monte Carlo, famoso cassino de Mônaco, foi adotado para sigilo das pesquisas e pela aleatoriedade lembrar os jogos de azar (VOSE, 2008).

A simulação de Monte Carlo é baseada na seleção aleatória de valores dentre as distribuições de probabilidades definidas para cada uma das variáveis de entrada e no cálculo repetido (iterações) da variável de saída de acordo com o modelo em questão (no caso, as equações de dimensionamento de lagoas e modelos de Análise Quantitativa de Risco Microbiológico – AQRM aplicada à irrigação com esgotos). A cada repetição do processo novas combinações aleatórias das variáveis de entrada são selecionadas e o resultado

correspondente é ‘armazenado’. Dependendo da distribuição definida para cada variável, aquelas com maior probabilidade de ocorrência são selecionadas com maior frequência.

É preciso ainda estar atento para a existência de variáveis de entrada correlacionadas, de forma que a simulação aleatória não gere combinações improváveis ou mesmo impossíveis.

O resultado final (a qualidade do efluente tratado ou a estimativa de risco) é então gerado em termos de distribuição de frequência, a qual representa uma combinação das faixas e frequências consideradas para as variáveis de entrada. O número de iterações a ser executada depende do grau de precisão desejado para a estimativa da distribuição de frequência da variável de saída (MORGAN e HERION, 2007; VOSE, 2008). A Figura 3.12 ilustra as etapas da Simulação Monte Carlo e a distribuição de frequência resultante em forma de histograma.

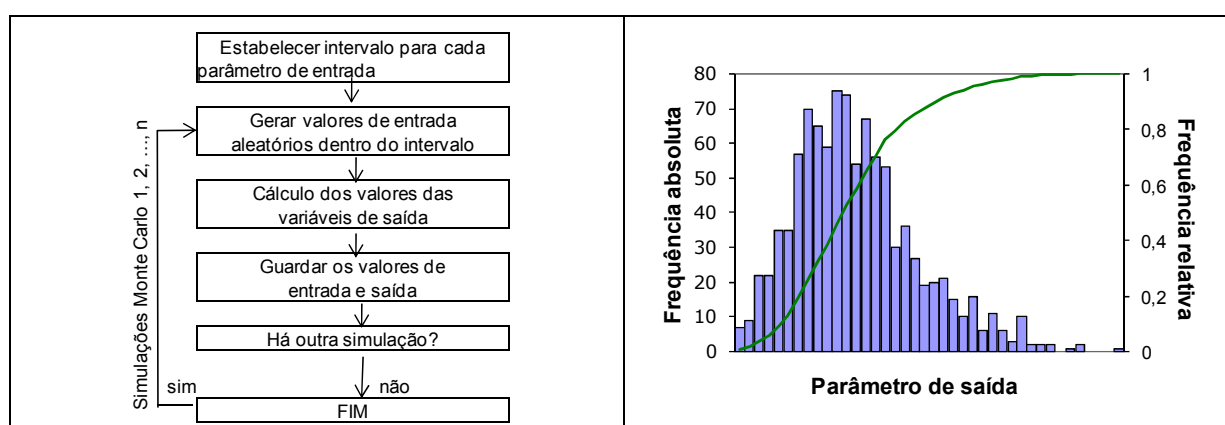


Figura 3.12 – Procedimentos da simulação Monte Carlo (à esquerda) e histograma resultante (à direita)

Fonte: Adaptado von Sperling (1999)

Cabe esclarecer que a Simulação de Monte Carlo não é a única forma de amostragem em modelagem estocástica. Como forma de amostragem aleatória simples, a simulação Monte Carlo seleciona valores aleatoriamente de forma independente, de acordo com a distribuição de probabilidades definida. Em outras palavras, o número aleatório utilizado em uma rodada não influencia os próximos números aleatórios a serem utilizados. Por essa característica, a simulação de Monte Carlo pode incorrer em sub ou super amostragem em diversas faixas da distribuição pré-definida.

Para solucionar esse problema, existem algumas técnicas de redução de variância que aumentam a eficiência da simulação Monte Carlo. A amostragem por ‘Hipercubos Latinos’

seleciona valores aleatoriamente de forma dependente. Tal método divide a distribuição em intervalos com probabilidades iguais de sorteio e seleciona um valor aleatório pertencente a cada um dos intervalos. Os intervalos são ‘marcados’ de forma que não sejam novamente selecionados nas iterações subsequentes. Como o número de intervalos é igual ao número de iterações, ao final cada intervalo será amostrado apenas uma vez, reproduzindo a distribuição de frequência da variável de entrada. A qualidade dos resultados estimados pela amostragem do Hipercubo Latino são superiores aos obtidos pela amostragem aleatória em termos de convergência e de mínima variância. Por isso, a amostragem Hipercubo Latino tem sido privilegiada em estudos de avaliação de risco como forma de garantir que os valores extremos inferiores e superiores da distribuição de frequência das variáveis de entrada não fiquem sub representados (SALIBY, 1989; SALIBY e MOREIRA, 2002; HELTON *et al.*, 2005; VOSE, 2008).

### **3.2 - Análise de sensibilidade em modelos estocásticos**

Essencialmente, a análise de sensibilidade permite determinar a importância de cada variável de entrada do modelo (ou sua variação) sobre a variável de saída e pode ser implementada com objetivos e técnicas diversas (ZWIETERING E VAN GERWEN, 2000; FREY E PATIL, 2002). Por exemplo, a análise de sensibilidade pode indicar quais variáveis de entrada demandam melhor controle ou conhecimento. Se *variabilidade* é a principal fonte de variação de uma variável de entrada com grande impacto na variável de saída, isso pode indicar a necessidade de melhor controle de um determinado processo ou fator. Por outro lado, se a variação de uma variável-chave for devida a *incerteza*, a prioridade pode ser a obtenção de maiores informações sobre a variável (FREY e PATIL, 2002). A análise de sensibilidade permite ainda aferir qualitativamente o resultado das simulações do modelo: se este é particularmente sensível a variáveis de entrada sobre os quais existem poucos dados ou muita incerteza (propriamente dita), o resultado das simulações pode ser considerado ‘frágil’; se, por outro lado, as variáveis de maior importância são aquelas sobre as quais se dispõe de banco de dados confiáveis (inclusive sob o ponto de vista estatístico), o resultado pode ser igualmente considerado confiável.

Frey e Patil (2002) classificam as técnicas de análise de sensibilidade em métodos matemáticos, estatísticos e gráficos. Os métodos gráficos são, em geral, utilizados como complemento, facilitando a visualização dos resultados dos métodos matemáticos e estatísticos.

Os métodos matemáticos resumem-se no cálculo da variável de saída para diferentes valores pontuais de uma variável de entrada. Esses métodos não levam em consideração a variância da variável de saída em função da variância das variáveis de entrada, mas somente permitem aferir o impacto da variação de variáveis de entrada dentre determinadas faixas de valores assumidas como possíveis de ocorrer (MORGAN e HENRION, 2007). São, entretanto, úteis na verificação inicial, permitindo identificar quais variáveis nitidamente demandam mais informações e aquelas que por sua pouca importância não demandam maiores esforços de caracterização das eventuais incertezas em seu redor (FREY e PATIL, 2002).

Um dos métodos matemáticos mais usuais é o ‘fator de sensibilidade’, definido como a relação entre o valor da variável de saída para um valor ‘esperado’ (por exemplo, médio) de uma determinada variável de entrada e aquele correspondente a um desvio da variável de entrada de seu ‘valor esperado’. Zwietering e van Gerwen (2000), em enfoque sobre avaliação de risco microbiológico, sugerem a abordagem do ‘cenário mais desfavorável’, tomando valores extremos para a variável de entrada em análise (Equação 3.22). O ‘fator de sensibilidade’ é calculado separadamente para cada variável de entrada no modelo, mantendo as demais constantes.

$$FS = \frac{N_k \text{extremo}}{N_k \text{médio}} \quad (3.22)$$

Onde:

$FS$  = fator de sensibilidade (%);

$N_k \text{ médio}$  = valor da variável de saída para o valor médio da variável de entrada  $k$ ;

$N_k \text{ extremo}$  = valor da variável de saída para um valor extremo da variável de entrada  $k$ .

Finley e Paustenbach (1994) apresentam uma abordagem similar, tomando um desvio percentual da variável de entrada em torno de seu valor médio, porém levando também em consideração a incerteza em torno das variáveis de entrada (Equação 3.23). Nesse caso, a importância comparativa de cada variável de entrada é expressa em termos de percentual em relação à ‘sensibilidade total’ (soma dos fatores de sensibilidade de cada variável).

$$FS = \left[ \frac{N_k \text{médio} - N_k \%}{X \text{médio} - X \%} \right] \cdot \sigma \quad (3.23)$$

Onde:

$FS$  = fator de sensibilidade;

$N_k \text{ médio}$  = valor da variável de saída para o valor médio da variável de entrada  $k$ ;

$N_k \%$  = valor da variável de saída para um determinado afastamento percentual da variável de entrada  $k$  de seu valor médio;

$X_{\text{médio}}$  = valor médio da variável de entrada  $k$ ;

$X\%$  = valor da variável de entrada  $k$  correspondente a um afastamento percentual de seu valor médio;

$\sigma$  = desvio-padrão da distribuição da variável de entrada  $k$ .

Os métodos estatísticos envolvem simulações em que as variáveis de entrada são consideradas em determinada distribuição de probabilidades, de forma a aferir a variância das variáveis de entrada sobre a distribuição da variável de saída (FREY e PATIL, 2002).

Podem ser empregados testes de regressão ou correlação. Em tese, o teste de regressão é superior ao de correlação, pois expressa relação causal entre as variáveis de entrada e de saída (uma variável de entrada pode provocar pequeno impacto na variável de saída, ainda que as duas se apresentem correlacionadas).

O teste de regressão linear multivariada (*stepwise*) fornece coeficientes de regressão normalizados para cada variável de entrada do modelo, os quais expressam a sensibilidade da variável de saída à cada uma das variáveis de entrada. O resultado é ainda aferido pelo coeficiente de regressão ( $R^2$ ), que expressa o ajuste geral do modelo de regressão - quando o valor do coeficiente de regressão ( $R^2$ ) é baixo, isso significa que o modelo de regressão linear não explica satisfatoriamente a relação entre as variáveis de entrada e de saída e que, portanto, outro tipo de análise deve ser procurada. Entretanto, este é um teste paramétrico, ou seja, dentre outros pressupostos, deve ser utilizado em amostras com população com distribuição normal (SIEGEL e CASTELLAN 2007).

Quando a linearidade do modelo não é comprovada, quando não se conhece a distribuição da população, ou, quando por opção, quer-se aplicar um teste que independa do tipo de distribuição da população da qual a amostra foi retirada, deve-se recorrer aos testes não paramétricos.

Para determinar o grau de relação entre variáveis de entrada e a de saída, podem ser utilizadas medidas de correlação estatística não paramétricas. O coeficiente de correlação posto-ordem  $r_s$  de Spearman é uma dessas medidas não paramétricas. A eficiência desse teste quando comparada com a correlação paramétrica mais poderosa – o coeficiente de Pearson - é em torno de 91%. Esse coeficiente determina a probabilidade associada com a ocorrência de uma correlação utilizando ‘rankings’ de valores e não os valores em si (como no caso do coeficiente de correlação linear). O valor da correlação varia entre -1 e 1. Quanto mais próximo ao valor zero, menor a correlação entre as variáveis. Quanto mais distantes de zero, próximos à -1 ou 1, maior a correlação. Em suma, quanto mais alta a correlação entre uma

variável de entrada e a de saída, mais importante será a variável de entrada na determinação do valor de saída (SPIENGEL, 1993; COSTA NETO, 1977).

Outra verificação possível de sensibilidade é a Análise de Cenários, por meio da qual se procura identificar combinações de variáveis de entrada que provocam determinados valores de saída. Essa análise permite identificar quais variáveis de entrada contribuem preponderantemente para um determinado resultado desejado, como, por exemplo: quais variáveis contribuem para o alcance de um efluente com concentração de coliformes menor que 1.000 org./100 mL? A execução dessa análise envolve a subdivisão das iterações da simulação e a seleção daquelas em que a variável de saída alcança o valor desejado e, em seguida, a análise dos valores amostrados para cada variável de entrada nestas iterações. São encontradas as medianas desse subconjunto de valores amostrais para cada variável de entrada, as quais são comparadas com as medianas relativas a todas as iterações. O objetivo desse processo é encontrar aquelas variáveis de entrada cujo subconjunto, ou ‘mediana condicional’, se diferencia significativamente da mediana geral. Se a mediana do subconjunto é próxima à mediana geral, a variável de entrada em questão é considerada não significativa. Entretanto, se a mediana do subconjunto para a variável de entrada se desviar de forma significativa da mediana geral (isto é, pelo menos  $\frac{1}{2}$  desvio padrão) a variável de entrada é considerada significativa (PALISADE CORPORATION 2009).

A análise de sensibilidade em relação à mediana pode ainda ser conduzida em associação com outros testes não-paramétricos. Em procedimento descrito em von Sperling (1996), a simulação de Monte Carlo é conduzida de forma que uma certa variável de entrada varie de acordo com a faixa e, ou com a distribuição de probabilidades definidas, enquanto as demais variáveis de entrada são mantidas constantes (por exemplo, em seus valores médios). Os resultados são separados em duas classes, tendo como ponto de corte o valor da mediana da variável de saída. Em seguida, as duas distribuições (contendo valores da variável de saída inferiores e superiores à mediana) são comparadas estatisticamente por meio de testes de correlação não paramétricos (por exemplo, Kolmogorov-Smirnof ou Mann-Whitney-Wilcoxon). Se as duas distribuições não diferirem estatisticamente, a variável de entrada é considerada sem importância em termos de impacto no resultado, caso contrário o modelo é considerado sensível às variações da variável de entrada em questão. Von Sperling (1996) e Gawasiri (2003) aplicaram essa técnica em análise de sensibilidade aplicada ao projeto de lagoas de estabilização.

Por fim, a análise gráfica de sensibilidade pode ser feita com o recurso aos gráficos de dispersão, representando os valores calculados de pares de valores de variáveis de entrada e saída em cada iteração. Os resultados dos testes acima descritos de correlação de Spearman e de Análise de Cenários também podem ser apresentados por meio de gráficos.

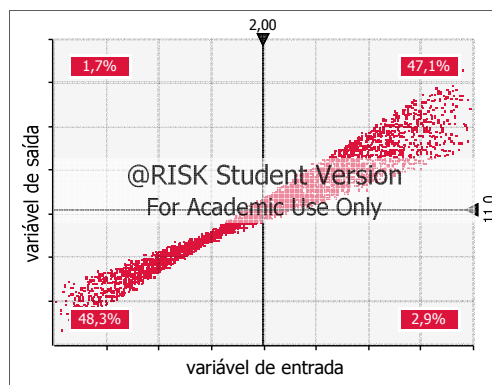


Figura 3.13 - Exemplo de análise de sensibilidade por meio de gráfico de dispersão

No presente trabalho foram empregados todos os testes acima descritos, à exceção do teste de regressão.

### 3.3 - Análise de incertezas e de sensibilidade na previsão da qualidade de efluentes de lagoas de estabilização

Tradicionalmente, o procedimento adotado no projeto de estações de tratamento de esgotos (ETEs) é o uso de valores fixos e únicos para os dados de entrada nas equações de dimensionamento. Essa abordagem *determinística* pressupõe, ainda que não de plena consciência por parte do projetista, certeza sobre os parâmetros de projeto, ou, no mínimo, uma aproximação dos valores esperados desses parâmetros.

Entretanto, a maioria dos parâmetros de projeto de estações de ETEs está sujeita a variações, por vezes amplas, devido a *variabilidade e incertezas*. Mesmo ciente das variações em torno dos parâmetros de dimensionamento, na maioria das vezes e talvez por razões de praticidade, os projetistas adotam modelos determinísticos, assumindo valores de projeto médios ou conservadores. Consequentemente, os resultados, por exemplo, de qualidade do efluente final, de volume ou área da ETE, são obtidos também como valores únicos, o que pode resultar em sub ou superdimensionamento.

Em modelos estocásticos, os parâmetros de projeto são assumidos em faixas de valores conforme uma determinada distribuição de probabilidades. As incertezas são tratadas por meio de amostragens aleatórias de cada variável de um modelo de dimensionamento, por exemplo, por meio de técnicas de simulação de Monte Carlo. Dessa forma, os resultados são também expressos em termos de distribuição de probabilidades, o que proporciona flexibilidade na tomada de decisões, como, por exemplo: qual seria a configuração final do sistema de tratamento (portanto, volume, área e custos) correspondente a determinada frequência de atendimento (inversamente, de violação) do padrão de lançamento de efluentes? Em outra perspectiva, os modelos estocásticos permitem inferências, desde a etapa de projeto, sobre a estabilidade ou confiabilidade dos sistemas de tratamento (NIKU *et al.*, 1979).

A abordagem estocástica, incluindo a simulação de Monte Carlo, aplicada ao tratamento de esgotos não constitui novidade e a literatura registra trabalhos envolvendo diferentes sistemas de tratamento (BUCHBERGER e SHAW, 1995; WERNER e KADLEC, 2000). Porém, trabalhos com lagoas de estabilização ainda são mais escassos (VON SPERLING, 1996; GAWASIRI, 2003), ainda que, dadas as características e a grande demanda de área desses sistemas, sua modelagem dinâmica seja particularmente interessante.

### *3.3.1 - Critérios de dimensionamento de lagoas de estabilização*

A literatura nacional e internacional sobre critérios de projeto de sistemas de lagoas é ampla e atualizada (KELLER e PIRES, 1998; CAVALCANTI *et al.*, 2001; VON SPERLING, 2002, VON SPERLING *et al.*, 2003, MARA, 2004; SHILTON, 2005). Portanto, não se pretende aqui apresentar uma vasta e redundante revisão sobre o tema, mas apenas destacar critérios de dimensionamento atuais, consagrados e específicos sobre o sistema e as variáveis em questão no presente estudo: lagoas de polimento e remoção de DBO, nitrogênio, coliformes e ovos de helmintos.

Na literatura nacional, as lagoas empregadas para o pós-tratamento de efluentes de reatores UASB têm sido denominadas lagoas de polimento com o intuito de distinguí-las das lagoas de estabilização convencionais (facultativas) e das lagoas de maturação. Isso porque as lagoas de polimento, apesar de ainda cumprirem um papel de remoção complementar de matéria orgânica, recebem uma carga orgânica já bastante reduzida e um efluente já bem clarificado. Portanto, dependendo do desempenho do pré-tratamento, o papel principal das

lagoas de polimento pode deixar de ser a estabilização do material orgânico, mas a remoção dos patógenos.

Assim, as lagoas de polimento eram inicialmente projetadas como lagoas facultativas, porém com o reconhecimento de que os problemas de sobrecarga orgânica na primeira lagoa da série são minimizados, verificou-se que as lagoas de polimento podem ser concebidas como lagoas de maturação, ou seja, com baixas profundidades e tempos de detenção hidráulica relativamente curtos. De fato, quanto mais próximas de um reator UASB mais as lagoas de polimento podem se assemelhar às lagoas facultativas e, quanto mais distantes, mais bem funcionam como lagoas de maturação (BASTOS *et al.*, 2009).

### 3.3.1.1 - Modelo de dispersão - fluxo disperso

O regime hidráulico influencia sobremaneira o funcionamento das lagoas de estabilização e, portanto, constitui fator a ser observado com critério na etapa de projeto.

Os regimes hidráulicos de mistura completa e fluxo em pistão constituem, respectivamente, os modelos de menor e maior eficiência, ou sob outro prisma, para uma dada eficiência desejada, o modelo de fluxo em pistão resultaria em dimensionamento mais econômico (menor volume do reator). No entanto, esses regimes são idealizados (fluxo em pistão – dispersão longitudinal do fluxo nula; mistura completa – dispersão do fluxo infinita) e, na prática, as lagoas apresentam regime hidráulico intermediário, de fluxo disperso. Este é, portanto, o critério mais adequado para o dimensionamento de lagoas e será o modelo considerado no presente trabalho.

O grau de dispersão do fluxo é caracterizado pelo número de dispersão ( $d$ ), o qual pode ser definido de acordo com a modelo de Wehner and Wilhelm (Equação 3.24).

$$d = D/uL \quad (3.24)$$

Onde:

$d$  = número de dispersão (adimensional);

$D$  = coeficiente de dispersão longitudinal ( $m^2.s^{-1}$ );

$u$  = velocidade de fluxo ( $m.s^{-1}$ );

$L$  = comprimento do reator (m).

Segundo von Sperling (1996), unidades que apresentam valores de  $d$  em torno de 0,2 ou menos aproximam-se das condições de fluxo em pistão. Por outro lado, unidades com valores de  $d$  em torno de 3,0 ou mais podem ser consideradas como aproximando condições de mistura completa.

O número de dispersão  $d$  pode ser determinado por meio de ensaios hidrodinâmicos com traçadores, ou, na etapa de projeto, ser estimado com o emprego de fórmulas empíricas que o associam à configuração geométrica da lagoa (POLPRASERT, 1983; AGUNWAMBA *et al.*, 1992; YANEZ, 1993; VON SPERLING, 1999). Neste trabalho foi adotado o modelo proposto por von Sperling (1999) (Equação 3.25).

$$d = \frac{1}{(L/B)} \quad (3.25)$$

Onde:

$L$  = comprimento (m);

$B$  = largura (m).

### 3.3.1.2 - Temperatura do líquido

Lagoas de estabilização, como sistemas que envolvem processos biológicos de tratamento, são altamente dependentes das condições climáticas. Portanto, como será visto em itens a seguir, a temperatura constitui importante parâmetro de projeto.

Von Sperling (2002), sistematizando informações de diferentes autores, apresenta a seguinte relação entre a temperatura do ar e da água em lagoas (Equação 3.26).

$$T_{\text{água}} = 12,7 + 0,54T_{\text{ar}} \quad (3.26)$$

Com base em informações da estação meteorológica do campus da Universidade Federal de Viçosa (UFV) e em vasto banco de dados de monitoramento do mesmo sistema de lagoas do presente estudo, Rios (2007) obteve a seguinte relação entre as temperaturas média do ar e do líquido (valores médios à meia profundidade) (Equação 3.27).

$$T_{\text{água}} = 8,17 + 0,73T_{\text{ar}} \quad (3.27)$$

### 3.3.1.3 - Remoção de DBO

Em geral, admite-se o modelo de primeira ordem para descrever a remoção de  $\text{DBO}_5$  em lagoas de estabilização. A Equação 3.28 apresenta o modelo de estimativa de  $\text{DBO}_5$  solúvel em efluentes de lagoas, segundo o regime de fluxo disperso.

$$S = S_0 \frac{4 \cdot a \cdot e^{1/2d}}{(1+a)^2 \cdot e^{a/2d} - (1-a)^2 \cdot e^{-a/2d}} \quad (3.28)$$

$$a = \sqrt{1 + 4K \cdot t \cdot d}$$

Onde:

$S_0$  = DBO total afluente ( $\text{mg/L}^{-1}$ );  
 $S$  = DBO solúvel efluente ( $\text{mg/L}^{-1}$ );  
 $K$  = coeficiente de remoção de matéria orgânica ( $\text{d}^{-1}$ );  
 $t$  = tempo de detenção hidráulica (d);  
 $n$  = número de lagoas em série;  
 $d$  = número de dispersão (adimensional).

O coeficiente  $K$  ( $\text{d}^{-1}$ ) de remoção de DBO para temperatura do líquido de  $20^\circ\text{C}$  pode ser determinado em função da taxa de aplicação superficial  $L_s$  ( $\text{Kg DBO}\cdot\text{ha}^{-1}\cdot\text{d}^{-1}$ ), segundo equação de Arceivala (1981).

$$K = 0,132 \cdot \log L_s - 0,146 \quad (3.29)$$

A taxa de qualquer reação química aumenta com a elevação da temperatura, desde que isso não produza alterações nos reagentes ou no catalisador. Para diferentes temperaturas, o valor de  $K$  pode ser corrigido através da formulação baseada na teoria de van't Hoff-Arrhenius, expressa como:

$$K_T = K_{20}\theta^{(T-20)} \quad (3.30)$$

Sendo:

$K_T$  = coeficiente de remoção à temperatura  $T$  ( $\text{d}^{-1}$ );

$K_{20}$  = coeficiente de remoção à  $20^\circ\text{C}$  ( $\text{d}^{-1}$ );

$\theta$  = coeficiente de temperatura;

$T$  = temperatura ( $^\circ\text{C}$ ).

Embora  $\theta$  seja tratado como uma constante, de fato diferentes valores tem sido atribuídos a esse coeficiente (1,035 a 1,085) (VON SPERLING, 2002; JORDÃO e PESSOA, 2005).

#### 3.3.1.4 - Remoção de coliformes

O decaimento de coliformes em lagoas é também descrito por modelos de cinética de primeira ordem. Portanto, os modelos de remoção para o regime de fluxo disperso (Equação 3.31) apresentam a mesma formulação dos modelos apresentados para a DBO.

$$N = N_0 \frac{4 \cdot a \cdot e^{1/2d}}{(1+a)^2 \cdot e^{a/2d} - (1-a)^2 \cdot e^{-a/2d}} \quad (3.31)$$

$$a = \sqrt{1 + 4K_b \cdot t \cdot d}$$

Onde:

$N_0$  = contagem de coliformes no afluente (org / 100 mL);

$N$  = contagem de coliformes no efluente (org / 100 mL);

$K_b$  = coeficiente de decaimento bacteriano ( $d^{-1}$ );

$t$  = tempo de detenção hidráulica (d);

$n$  = número de lagoas em série;

$d$  = número de dispersão (adimensional).

Como ilustrado na Figura 3.14, o regime hidráulico das lagoas tem grande influência na eficiência de remoção de coliformes (VON SPERLING, 2002).

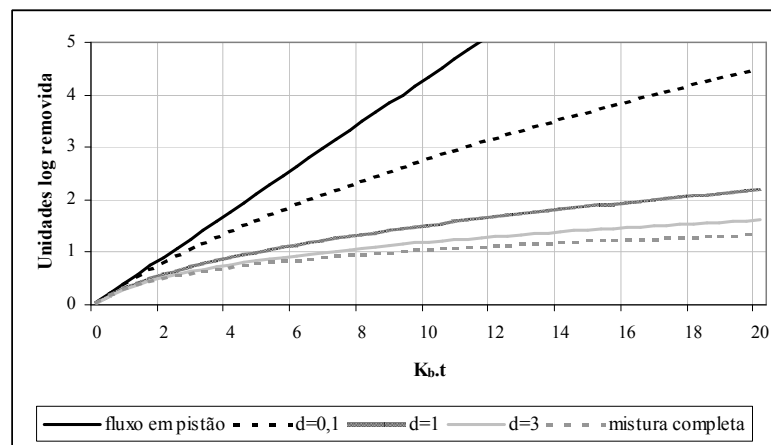


Figura 3.14 - Remoção de coliformes em função do produto  $K_b.t$  para diversos valores de número de dispersão  $d$

Fonte: von Sperling (2002)

A taxa de mortalidade das bactérias é expressa pelo coeficiente de decaimento bacteriano ( $K_b$ ), o qual é uma característica intrínseca ao tipo de organismo (resistência do organismo), à lagoa (configuração geométrica, altura, valores de pH e OD, enfim, às condições ambientais estabelecidas) e à temperatura. Vários modelos disponíveis na literatura procuram expressar a relação dessas variáveis com  $K_b$ , sendo que no Brasil encontram ampla utilização os modelos propostos por von Sperling, como funções da profundidade da lagoa ( $H$ ) e do tempo de detenção hidráulica ( $TDH$ ) (Equações 3.32 a 3.34).

$$K_b = 0,917 \cdot H^{-0,877} \cdot TDH^{-0,329} \quad \text{von Sperling (1999)} \quad (3.32)$$

(33 lagoas no Brasil)

$$K_b = 0,542 \cdot H^{-1,259} \quad \text{von Sperling (1999)} \quad (3.33)$$

(82 lagoas em várias partes do mundo)

$$K_b = 0,628 \cdot H^{-1,286} \cdot TDH^{-0,103} \quad \text{von Sperling (2005)} \quad (3.34)$$

(186 lagoas, a maioria no Brasil)

Rios (2007) testou o ajuste do vasto banco de dados do mesmo sistema de lagoas do presente estudo a dois dos modelos propostos por von Sperling (Equações 3.32 e 3.33). Como estes dois modelos mostraram um ajuste apenas razoável aos dados experimentais, em geral subestimando os valores de  $K_b$ , a Equação 3.35 foi proposta como mais específica para o sistema em questão.

$$K_b = -0,7118 \cdot \ln H + 0,8589 \quad (3.35)$$

Onde:

$H$  = profundidade da lagoa (altura de lâmina d'água) (m).

O coeficiente  $K_b$  deve também ser corrigido para a temperatura do líquido na lagoa, de acordo com a Equação 3.30. Gawasiri (2003) relata vários coeficientes de temperatura encontrados na literatura para o modelo de remoção de coliformes, tais como: 1,19 (MARAIS, 1974), 1,06 (SHERRY e PARKER, 1979), 1,17 (MILLS *et al.*, 1992), 1,07 (YANEZ, 1993) e 1,08 (MAYO, 1995).

### 3.3.1.5 - Remoção de amônia

Um dos modelos consagrados na literatura para a remoção de amônia é o de Pano e Middlebrooks (1982) (Equações 3.36 e 3.37), obtido do monitoramento de três sistemas de lagoas facultativas operando nos EUA com cargas de DBO inferiores a  $40 \text{ kg} \cdot \text{ha}^{-1} \cdot \text{d}^{-1}$  e com base nas seguintes premissas: (i) regime de mistura completa, (ii) cinética de primeira ordem e (iii) a volatilização da amônia como o principal mecanismo de remoção de nitrogênio.

$$C_e = \frac{C_0}{1 + \left[ \left( \frac{A}{Q} \right) \cdot (0,0038 + 0,000134 \cdot T) \cdot e^{(1,041 + 0,0044 \cdot T) \cdot (pH - 6,6)} \right]} \quad (T < 20^\circ\text{C}) \quad (3.36)$$

$$C_e = \frac{C_0}{1 + \left[ 5,035 \times 10^{-3} \cdot \left( \frac{A}{Q} \right) \cdot e^{(1,54 \times (pH - 6,6))} \right]} \quad (T \geq 20^\circ\text{C}) \quad (3.37)$$

Onde:

$C_0$  = concentração afluyente de amônia ( $\text{mg/L}^{-1}$ );

$C_e$  = concentração efluyente de amônia ( $\text{mg/L}^{-1}$ );

$A$  = área superficial da lagoa ( $\text{m}^2$ );

$Q$  = vazão afluyente ( $\text{m}^3 \text{d}^{-1}$ );

$T$  = temperatura do líquido ( $^\circ\text{C}$ );

$pH = \text{pH da lagoa}$ .

Estes modelos evidenciam uma importante relação entre remoção de amônia, pH e carga hidráulica superficial e, portanto, a relação profundidade / tempo de detenção hidráulica, mas por outro lado, não permitem uma solução contínua para temperaturas inferiores e superiores a 20°C; além disso, são dependentes do pH, uma variável não conhecida na etapa de projeto.

Trabalhos mais recentes têm questionado que o principal mecanismo de remoção de nitrogênio seja a volatilização da amônia, sugerindo um papel preponderante de outros mecanismos, tais como a absorção pelas algas e a sedimentação de algas mortas (CAMARGO VALERO e MARA, 2007a, 2007b).

Não obstante, outros trabalhos têm constatado que o modelo de Pano e Middlebrooks bem se ajustou a dados experimentais obtidos no Brasil: Soares *et al.* (1996) na Paraíba, e Rios (2007) no mesmo sistema aqui em estudo.

### 3.3.1.6 - Remoção de ovos de helmintos

A remoção de coliformes não cumpre, adequadamente, o papel de indicador da remoção de helmintos em lagoas de estabilização. Isso porque enquanto a remoção dos coliformes se dá, preponderantemente, por mecanismos de inativação por efeito da radiação solar, a de ovos de helmintos é alcançada, basicamente, por remoção física, resultado da adsorção em flocos ou de sedimentação simples, em virtude de apresentarem maior densidade que a água (CAVALCANTI *et al.* 2001).

Dentre os poucos modelos disponíveis de estimativa de ovos de helmintos em lagoas de estabilização, destaca-se o de Ayres *et al.* (1992), dependente exclusivamente do tempo de detenção hidráulica (Equações 3.38 e 3.39) e válido para lagoas anaeróbias, facultativas e de maturação. As equações devem ser aplicadas sequencialmente em cada lagoa da série, de forma que o número de ovos no efluente final possa ser determinado.

- Eficiência média de remoção, usada para representar condições médias de operação:

$$E = 100[1 - 0,14 \cdot e^{-0,38t}] \quad (3.38)$$

- Eficiência de remoção segundo o limite inferior de confiança de 95%, usada para projeto por se posicionar a favor da segurança:

$$E = 100[1 - 0,14 \cdot e^{(-0,49t+0,0085t^2)}] \quad (3.39)$$

Onde:

$E$  = eficiência de remoção de ovos de helmintos (%);

$t$  = tempo de detenção hidráulica em cada lagoa da série (d).

### 3.3.2 - Caracterização das incertezas associadas aos parâmetros de projeto de lagoas de estabilização

As equações empíricas para a remoção de DBO, amônia, coliformes e ovos de helmintos citadas na seção anterior, incluem diversas variáveis e parâmetros de projeto os quais trazem incertezas em torno de seus valores.

Por exemplo, no modelo de remoção de DBO (Equação 3.28), os seguintes parâmetros estariam sujeitos a variações devidas à incertezas (variabilidade e incerteza propriamente dita):

- (i) A concentração afluente de DBO, devido a variações, por exemplo, de consumo de água residencial e de infiltrações de água na rede coletora de esgotos. No caso específico deste estudo devem ainda ser levadas em conta variações no desempenho do sistema de tratamento anterior às lagoas: reator UASB e biofiltro submerso aerado.
- (ii) O coeficiente de remoção de matéria orgânica ( $K$ ), o qual, como visto na Seção 3.3.1.3, varia em função da taxa de aplicação superficial de DBO e da temperatura. Portanto, também devem ser consideradas incertezas em torno da temperatura e do coeficiente de temperatura ( $\theta$ ).
- (iii) O tempo de detenção hidráulica, como decorrência de variações das vazões afluentes.
- (iv) O número de dispersão, em função de variações no regime de fluxo.

De forma análoga, no modelo de remoção de coliformes (Equação 3.31), além das variações de vazão, tempo de detenção hidráulica e do número de dispersão, devem ser consideradas incertezas em torno dos seguintes parâmetros: (i) concentração de coliformes nos afluentes às lagoas e (ii) coeficiente de decaimento bacteriano ( $K_b$ ), temperatura e coeficiente de temperatura ( $\theta$ ).

Por sua vez, no modelo de remoção de amônia, as seguintes variáveis merecem atenção pelas incertezas associadas (além das já mencionadas variações de vazão e temperatura): (i) concentração afluente de amônia e (ii) pH da lagoa.

Por fim, no modelo remoção de ovos de helmintos as incertezas se dariam em torno (i) do tempo de detenção hidráulica e (ii) do número de ovos de helmintos no afluente às lagoas.

Tais incertezas em torno das variáveis de entrada nos modelos (as equações de dimensionamento) são então definidas por faixas de valores, às quais devem ser ajustadas distribuições de frequência ou funções densidade de probabilidades (ver Seção 3.1).

Metcalf e Eddy (2003) sugerem que, em geral, as variáveis que podem ser modificadas pelo tratamento tendem a ser descritas nos efluentes de ETEs pela distribuição lognormal. Dean e Forsythe (1976a, 1976b), citados por Oliveira (2006), reportam dados de coliformes termotolerantes e turbidez em processos avançados de tratamento se ajustaram à distribuição lognormal. Niku *et al.* (1979, 1981b, 1981c), citados por Oliveira (2006), analisando 43 estações de lodos ativados nos Estados Unidos também encontraram que distribuição lognormal descreveu a qualidade dos efluentes em termos de concentrações de DBO e SST. Charles *et al.* (2005) analisaram dados de qualidade de efluentes (DBO, SS, nitrogênio total, fósforo total e coliformes termotolerantes) de vários tanques sépticos na Austrália e encontraram variações típicas de distribuição lognormal. Por sua vez, Tanaka *et al.* (1998) ajustaram funções densidade de probabilidades lognormais a banco de dados de monitoramento de vírus entéricos em efluentes de ETEs (lodos ativados e filtros percoladores) na Califórnia (EUA).

Dados nacionais e mais especificamente sobre lagoas de estabilização são bem mais escassos. Oliveira (2006) analisou dados de 208 ETEs dos estados de Minas Gerais e São Paulo compreendendo vários processos de tratamento (fossas-filtro, lagoas de estabilização, lodos ativados, reatores UASB) e concluiu que tanto no esgoto bruto como nos efluentes das ETEs a distribuição lognormal mostrou ser a mais adequada para representar o comportamento da maioria das variáveis analisadas (DBO, DQO, SST, Nitrogênio Total, Fósforo Total e coliformes termotolerantes), sendo o Fósforo Total afluente a única exceção. Essa tendência foi confirmada quando os dados dos sistemas de lagoas foram analisados em separado.

Em um dos poucos trabalhos conhecidos sobre modelagem dinâmica para projeto de lagoas de estabilização, Gawasiri (2003) assumiu distribuição uniforme para as seguintes variáveis em faixas de valores pré-definidas: concentrações afluentes de DBO e coliformes, vazão afluente, temperatura, coeficiente de temperatura ( $\theta$ ), número de dispersão. Esse autor argumenta que a distribuição uniforme é adequada para variáveis que possam assumir os valores extremos das faixas consideradas e que esse tipo de distribuição é particularmente

adequado ao projeto de lagoas porque os valores médios determinísticos dos parâmetros de projeto são, em geral, conhecidos com elevado grau de certeza.

Gawasiri (2003) realizou ainda análise de sensibilidade na concepção de lagoas de estabilização (lagoas aneróbia, facultativa e de maturação em série) e concluiu que a concentração efluente de coliformes é significativamente influenciada pela contribuição *per capita* de DBO, contribuição *per capita* de esgoto, taxa de evaporação, temperatura, coeficiente de remoção de coliformes fecais nas lagoas anaeróbias e número de dispersão nas lagoas facultativas, quando os valores de entrada variavam em  $\pm 20\%$ . Já a área total do sistema mostrou-se significativamente influenciada pela contribuição *per capita* de esgoto, DBO *per capita*, população e temperatura.

Em trabalho similar e anterior, von Sperling (1996) também assumiu distribuição uniforme para as seguintes variáveis, dentre faixas de valores entre a média  $\pm 10\%$ : população de projeto, contribuição *per capita* de esgotos, contribuição *per capita* de DBO, temperatura do líquido, coeficiente de remoção de DBO (K), coeficiente de temperatura ( $\theta$ ). Outras variáveis, as quais dependem de escolha do projetista, tais como taxa de aplicação superficial de DBO e profundidade da lagoa, foram assumidos como constantes. O autor concluiu que as variáveis de entrada importantes na determinação da DBO efluente foram o coeficiente de remoção de DBO e o coeficiente de temperatura. Na análise de sensibilidade para a determinação da área superficial necessária, as variáveis importantes foram população e a contribuição *per capita* de DBO.

### **3.4 - Análise de incertezas e de sensibilidade na Avaliação Quantitativa de Risco Microbiológico aplicada ao uso de efluentes para irrigação.**

#### *3.4.1 - Normas e critérios de qualidade microbiológica da água para a irrigação com esgotos*

Estima-se que mais de 20 milhões de hectares em cerca de 50 países sejam irrigados com esgotos sanitários, tratados ou não (SCOTT *et al.*, 2004). Exemplos os mais variados podem ser encontrados em diferentes contextos socioeconômicos, desde sistemas rigidamente controlados por diretrizes e regulamentação governamentais, com elevados níveis de planejamento ou padrão tecnológico, até iniciativas espontâneas envolvendo sérios riscos à saúde (SCOTT *et al.*, 2004). No que diz respeito a critérios para a formulação de normas de qualidade microbiológica de águas residuárias para irrigação, atualmente são reconhecidas as

seguintes abordagens: (i) a ausência de riscos potenciais (perigos), caracterizada pela ausência de organismos indicadores e, ou, patogênicos na água de irrigação; (ii) a medida do risco atribuível à irrigação com esgotos dentre uma população exposta; (iii) a estimativa do risco atribuível mediante o emprego de modelos probabilísticos (BASTOS e BEVILACQUA, 2006).

A primeira tem sido referida como abordagem do ‘risco nulo’ e, embora criticada por rigor excessivo e fragilidade em termos de fundamentação epidemiológica (CARR *et al.*, 2004), constitui a base da regulamentação do reúso da água para fins agrícolas em vários países, com destaque para as normas vigentes nos EUA (USEPA, 2004) (Quadro 3.2).

Quadro 3.2 - Diretrizes da USEPA para o uso agrícola de esgotos sanitários

<b>Tipo de irrigação e cultura</b>	<b>Processo de tratamento</b>	<b>Qualidade do efluente</b>
Culturas alimentícias não processadas comercialmente Irrigação superficial ou por aspersão de qualquer cultura, incluindo culturas a serem consumidas cruas	Secundário + filtração + desinfecção	pH 6 a 9 DB0 ≤ 10 mg L <sup>-1</sup> . Turbidez ≤ 2 uT CRT ≥ 1 mg L <sup>-1</sup> CTer ND Organismos patogênicos ND
Culturas alimentícias processadas comercialmente Irrigação superficial de pomares e vinhedos Silvicultura e irrigação de áreas com acesso restrito ao público	Secundário + desinfecção	pH 6 a 9 DB0 ≤ 30 mg L <sup>-1</sup> . SST ≤ 30 mg L <sup>-1</sup> CRT ≥ 1 mg L <sup>-1</sup> CTer ≤ 200/100 mL
Culturas não alimentícias Pastagens para rebanhos de leite, forrageiras, cereais, fibras e grãos	Secundário + desinfecção	pH 6 a 9 DB0 ≤ 30 mg L <sup>-1</sup> . SST ≤ 30 mg L <sup>-1</sup> CRT ≥ 1 mg L <sup>-1</sup> CTer ≤ 200/100 mL

ND: não detectável; CTer: coliformes termotolerantes; CRT: cloro residual total.

Fonte: Bastos e Bevilacqua (2006), adaptado de USEPA (2004)

Na segunda abordagem, busca-se evidências epidemiológicas de associação entre a irrigação com esgotos e a ocorrência de doença na população por meio de medidas de risco (por exemplo, risco relativo, risco atribuível). Entretanto, segundo Bastos e Bevilacqua (2006), não obstante esse enfoque apresentar dificuldades práticas, por depender de complexos estudos epidemiológicos, foi essa a abordagem principal que em 1989 subsidiou a formulação das Diretrizes da Organização Mundial da Saúde (WHO, 1989).

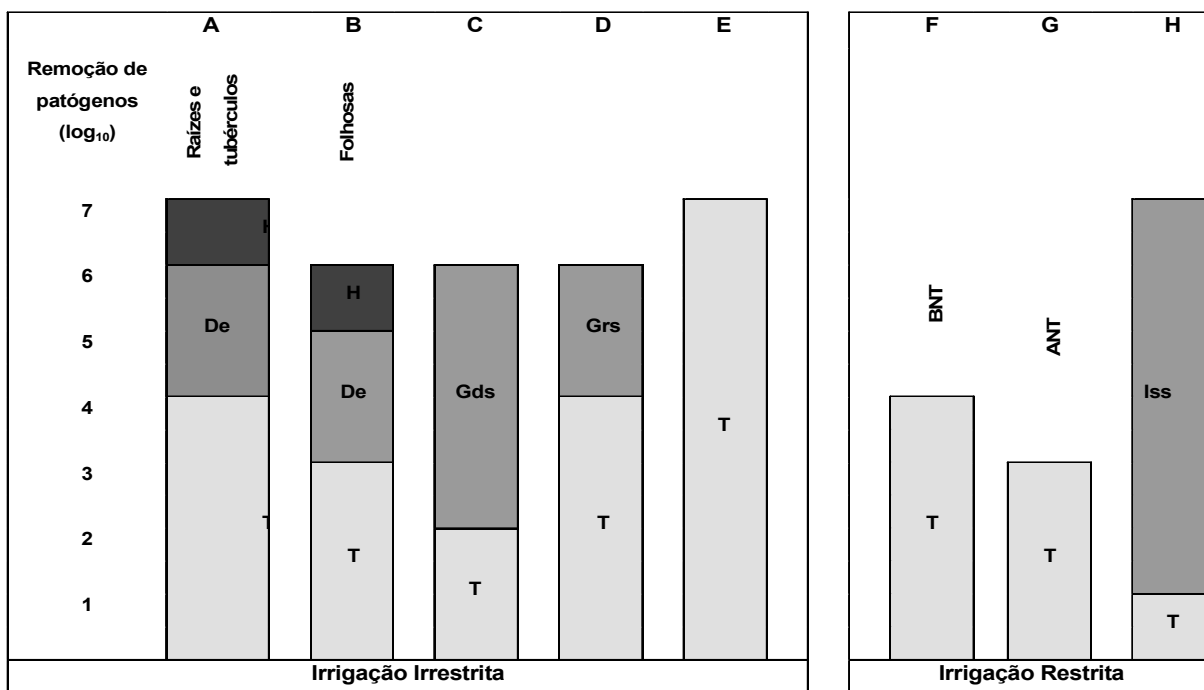
Na terceira perspectiva, a estimativa de risco é realizada com a aplicação da metodologia de avaliação quantitativa de risco microbiológico (AQRM): construído um cenário de exposição (ou seja, conhecida a concentração de organismos patogênicos na água de irrigação, no solo ou nas culturas irrigadas, é estimada a ingestão de patógenos a cada evento de exposição, por exemplo, por meio do consumo de alimentos irrigados ou da ingestão involuntária de partículas de solo), pode-se estimar o risco de infecção em base populacional e temporal mediante o emprego de modelos probabilísticos, os quais levam em consideração informações de dose-resposta de um determinado patógeno (HAAS *et al.*, 1999; WHO, 2006). Esta é a abordagem central que subsidiou a mais recente atualização das Diretrizes da Organização Mundial da Saúde (WHO, 2006) (Quadro 3.3 e Figura 3.15).

Quadro 3.3 – Diretrizes da OMS o uso agrícola de esgotos sanitários.

Categoria irrigação	Opção <sup>(1)</sup>	Tratamento de esgotos e remoção de patógenos (log <sub>10</sub> ) <sup>(2)</sup>	Qualidade do efluente	
			<i>E. coli</i> 100 mL <sup>-1</sup> <sup>(3)</sup>	Ovos helmintos L <sup>-1</sup>
Irrestrita	A	4	≤ 10 <sup>3</sup>	≤ 1
	B	3	≤ 10 <sup>4</sup>	
	C	2	≤ 10 <sup>5</sup>	
	D	4	≤ 10 <sup>3</sup>	
	E	6 ou 7	≤ 10 <sup>1</sup> ou 10 <sup>0</sup>	
Restrita	F	4	≤ 10 <sup>4</sup>	
	G	3	≤ 10 <sup>5</sup>	
	H	< 1	≤ 10 <sup>6</sup>	

(1) Combinação de medidas de proteção à saúde (ver Figura 3.15). (A): cultivo de raízes e tubérculos; (B): cultivo de folhosas; (C): irrigação localizada de plantas que se desenvolvem distantes do nível do solo; (D): irrigação localizada de plantas que se desenvolvem rentes ao nível do solo; (E): qualidade de efluentes alcançável com o emprego de técnicas de tratamento tais como tratamento secundário + coagulação + filtração + desinfecção; qualidade dos efluentes avaliada ainda com o emprego de indicadores complementares (por exemplo, turbidez, SST, cloro residual); (F): agricultura de baixo nível tecnológico e mão de obra intensiva; G: agricultura de alto nível tecnológico e altamente mecanizada; (H): técnicas de tratamento com reduzida capacidade de remoção de patógenos (por exemplo, tanques sépticos ou reatores UASB) associada ao emprego de técnicas de irrigação com elevado potencial de minimização da exposição (irrigação subsuperficial). (2) remoção de vírus que associada a outras medidas de proteção à saúde corresponderia a uma carga de doenças virais tolerável ≤ 10<sup>-6</sup> DALY ppa e riscos menores de infecções bacterianas e por protozoários. (3) Qualidade do efluente correspondente à remoção de patógenos indicada em (2).

Fonte: Bastos e Bevilacqua (2006), adaptado de WHO (2006)



T: tratamento; De: Decaimento no ambiente; H: Higiene dos alimentos (lavagem)

Gds: irrigação por gotejamento de plantas que se desenvolvem distantes do nível do solo; Grs: irrigação por gotejamento de plantas que se desenvolvem rentes ao nível do solo; BNT: agricultura de baixo nível tecnológico (mão de obra intensiva); ANT: agricultura de alto nível tecnológico (mecanização); Iss: irrigação subsuperficial

Figura 3.15 - Exemplos de combinação de medidas de proteção à saúde para carga de doenças tolerável  $\leq 10^{-6}$  DALY pppa

Fonte Fonte: Bastos e Bevilacqua (2006), adaptado de WHO (2006)

Conforme descrito por Bastos e Bevilacqua (2006), nas abordagens que envolvem avaliação de risco, de forma a bem fundamentar a formulação de normas de qualidade da água, a medida e, ou, a estimativa do risco devem ser contrapostas a um determinado nível de risco que se considere ‘tolerável’. Inversamente, estabelecido o risco tolerável, pode-se estimar a concentração admissível de organismos patogênicos na água/efluente tratados e, por conseguinte, o grau de tratamento requerido.

Portanto, uma decisão fundamental ao processo de AQRM é a definição do que seja ‘risco tolerável’. Nos EUA admite-se um risco anual de infecção de  $10^{-4}$  para os diversos organismos patogênicos transmissíveis via abastecimento de água para consumo humano<sup>3</sup> (HAAS *et al.*, 1999). As diretrizes da OMS para a qualidade da água para consumo humano e para a utilização de esgotos sanitários (WHO, 2004; WHO, 2006) encontram-se baseadas no conceito de ‘carga de doença’, medida pelo indicador ‘anos de vida perdidos ajustados por

<sup>3</sup> Um caso anual de infecção dentre 10.000 pessoas

incapacidade' (da sigla inglesa DALYs – *Disability Adjusted Life Years*), o qual permite a transformação de uma 'incapacidade vivenciada' (por exemplo, três dias com diarreia ou óbito devido à diarreia) em 'anos de vida saudáveis perdidos'. A OMS assume como carga de doença tolerável  $1 \times 10^{-6}$  DALY, ou seja,  $1 \mu$  DALY por pessoa por ano (pppa), o que corresponde a níveis de risco toleráveis anuais de  $7,7 \times 10^{-4}$  pppa para rotavírus,  $3,1 \times 10^{-4}$  pppa para *Campylobacter* e  $2,2 \times 10^{-3}$  pppa para *Cryptosporidium*<sup>4</sup>. Portanto, definido o que se queira como DALY tolerável, pode-se estimar requerimentos de remoção de patógenos por meio do tratamento a partir do conhecimento de sua concentração no esgoto bruto (WHO, 2006; MARA, 2007). Essa é a lógica por trás das Diretrizes da OMS, expressa no Quadro 3.3 e na Figura 3.15.

Por fim, cabe registrar, tomando novamente como referência Bastos e Bevilacqua (2006), que as diretrizes da OMS foram desenvolvidas tendo os rotavírus como 'organismo-referência', ou seja, sua remoção por meio do tratamento de esgotos associada a outras medidas de proteção à saúde corresponderia a uma carga de doenças virais  $\leq 10^{-6}$  DALY pppa e a riscos menores de infecções bacterianas e por protozoários. No que diz respeito aos riscos associados a doenças helmínticas, o padrão de qualidade de efluentes permaneceu baseado nas evidências epidemiológicas disponíveis, ao se reconhecer, à época, a insuficiência de informações para a aplicação da metodologia de AQRM e estimativa de DALYs (WHO, 2006).

### 3.4.2 - Avaliação Quantitativa de Risco Microbiológico (AQRM)

Do acima exposto, depreende-se que os modelos de AQRM são constituídos por dois componentes: um modelo de exposição e um modelo dose-resposta.

O modelo de exposição permite a estimativa da dose de patógenos ingerida em cada evento de exposição, podendo ser simplificada apresentada pela Equação 3.40.

$$d = M * D * C \quad (3.40)$$

Onde:

d = dose ingerida de patógenos;

M = concentração de patógenos na superfície das culturas ou em solos irrigados com esgotos;

---

<sup>4</sup> A forma de conversão da medida de risco em DALY e vice-versa pode ser encontrada em WHO (2006), Bastos e Bevilacqua (2009) e Mara (2009).

$D = 10^{-y}$ , que representa a redução da concentração de patógenos entre a última irrigação e o evento de exposição, por decaimento natural nas culturas ou no solo irrigado, ou por remoção durante o processamento ou higiene dos alimentos e;

C = consumo *per capita* diário de alimentos ou ingestão de partículas de solo.

Os modelos dose-resposta são, em geral, obtidos de estudos experimentais com humanos. Os mais utilizados em AQRM têm sido o modelo beta-Poisson para infecção por rotavírus e *Campylobacter* (Equação 3.41) e o modelo exponencial para infecção por *Cryptosporidium* (Equação 3.42) (HAAS *et al.*, 1999).

$$P_I(d) = 1 - [1 + (d/N_{50}) (2^{1/\alpha} - 1)]^{-\alpha} \quad (3.41)$$

$$P_I(d) = 1 - [\exp(-rd)] \quad (3.42)$$

Onde:

$P_I(d)$  = probabilidade de infecção para uma única exposição (risco diário);

d = número de organismos ingeridos por exposição (dose);

$N_{50}$  = dose infectante média,

$\alpha$  e r = parâmetros característicos da interação agente-hospedeiro.

A partir das Equações 3.41 e 3.42 pode-se estimar o risco anual, de acordo com a Equação 3.43. O valor de  $P_{I(A)}$  varia entre 0 e 1, sendo certa a infecção se  $P(A)=1$ .

$$P_{I(A)}(d) = 1 - [1 - P_I(d)]^n \quad (3.43)$$

Onde:

$P_{I(A)}(d)$  = probabilidade anual de infecção decorrente de  $n$  exposições à mesma dose (d),

$n$  = número de exposições por ano.

Vários autores têm recorrido à AQRM para a estimativa de riscos associados à irrigação com esgotos sanitários, incluindo a aplicação de modelos determinísticos (ASANO *et al.*, 1992; ROSE *et al.*, 1996; SHUVAL *et al.*, 1997; BASTOS *et al.*, 2008) ou de modelos estocásticos envolvendo simulação de Monte Carlo (TANAKA *et al.*, 1998; van GINNEKEN e ORON, 2000; PETTERSON *et al.*, 2001; HAMILTON *et al.*, 2006a; MARA *et al.*, 2007). Os trabalhos publicados incluem ampla variedade de cenários de exposição testados, tais como: risco ao consumidor e ao trabalhador (MARA *et al.*, 2007), tipo de cultura irrigada (HAMILTON *et al.*, 2006a), métodos de irrigação (van GINNEKEN e ORON, 2000), estimativas (HAMILTON *et al.*, 2006a; MARA *et al.*, 2007) ou medidas de contaminação das culturas irrigadas (BASTOS *et al.*, 2008; SEIDU *et al.*, 2008) e padrão de consumo de alimentos (HAMILTON *et al.*, 2006b). A maioria dos trabalhos tende a enfatizar riscos associados ao *Campylobacter*, *Cryptosporidium* e principalmente ao rotavírus, tendo sido

estes os organismos tomados como referência na atualização das diretrizes da OMS (WHO, 2006).

### 3.4.3 - Caracterização das incertezas associadas aos modelos de AQRM aplicados à irrigação com esgotos

Tomando como referência os modelos de exposição e de dose-resposta descritos nos itens anteriores, incertezas poderiam ser atribuídas a praticamente todas às respectivas variáveis de entrada.

Por exemplo, nos modelos dose-resposta (Equações 3.41 e 3.42), incertezas podem ser creditadas à dose infectante média ( $N_{50}$ ) e aos parâmetros característicos da interação agente-hospedeiro ( $\alpha$  e  $r$ ). Mara *et al.* (2007) abordaram isso atribuindo variações de  $\pm 25\%$  em torno dos valores de  $N_{50}$ ,  $\alpha$  e  $r$ .

Em relação aos modelos de exposição, muitos trabalhos sobre estimativas de risco para enterovírus, como por exemplo, Petterson *et al.*, 2001; Hamilton *et al.*, 2006a, partem do banco de dados sistematizado por Tanaka *et al.*, (1998) sobre a ocorrência desses organismos em efluentes de estações convencionais de tratamento de esgotos nos EUA (lodos ativados e filtros percoladores), ao qual foi ajustada a função lognormal.

Para a estimativa da concentração de patógenos nos efluentes outros trabalhos assumem uma relação entre patógenos e *E. coli*, inicialmente proposta por Shuval *et al.* (1997) para enterovírus e posteriormente ampliada para *Campylobacter* e *Cryptosporidium* por Mara *et al.* (2007), sendo esta a abordagem assumida nas Diretrizes da OMS (WHO, 2006): 0,1–1 rotavírus e *Campylobacter*, e 0,01–0.1 oocistos de *Cryptosporidium* por  $10^5$  *E. coli*.

Com relação a concentração de patógenos nas culturas ou nos solos irrigados, são poucos os trabalhos que incluem níveis de contaminação de fato medidos. Para a contaminação de culturas irrigadas, a maioria se vale de uma estimativa do volume de esgotos por massa de cultura irrigada, também inicialmente proposta por Shuval *et al.* (1997) para alface (10,8 mL/ 100 g) e pepino (0,36 mL/100g) e posteriormente ampliada por Hamilton *et al.* (2006a) para repolho (3,5 a 8,9 mL/100g) e brócolis (1,85 mL/g) (1,85 mL/g). Hamilton *et al.* (2006a) ainda associaram funções distribuições de probabilidades à esses dados (loglogística para brócolis, normal para pepino e alface e funções empíricas para repolho). Para a contaminação de solos, a maioria dos trabalhos recorre ao assumido nas Diretrizes da

OMS: o mesmo número de *E. coli*/100mL e *E. coli*/100g, respectivamente no esgoto e no solo. Seidu *et al.* (2008) trabalharam com níveis medidos de *E. coli* em culturas e solos irrigados e com as relações *E. coli*/patógenos acima referida, ajustando, entretanto, funções lognormais à maioria desses dados. Bastos *et al.* (2008), com base em dados experimentais, ajustaram modelos de estimativa de contaminação de culturas (*E. coli*/g) em função da qualidade da água de irrigação (*E. coli*/100mL) (Equações 3.44 e 3.45); entretanto, esses autores utilizaram modelagem determinística.

- Plantas que se desenvolvem rente ao solo:

$$\log_{10}(EC_{hortaliça}) = 0,83 \cdot \log_{10}(EC_{efluente}) - 0,73 \quad (3.44)$$

- Plantas que se desenvolvem distantes do solo:

$$\log_{10}(EC_{hortaliça}) = 0,45 \cdot \log_{10}(EC_{efluente}) - 0,67 \quad (3.45)$$

Onde:

$EC_{hortaliça}$ : concentração de *E. coli* por grama na superfície da planta;

$EC_{efluente}$ : concentração de *E. coli* por 100 mL na água de irrigação.

No que diz respeito ao padrão de consumo de alimentos, vários trabalhos recorrem aos pressupostos contidos nas Diretrizes da OMS, formulados por Mara *et al.* (2007): (i) ingestão involuntária de 1-10 mg de solo por evento de exposição para agricultura mecanizada e 10-100 mg para agricultura com mão de obra intensiva e (ii) consumo de 100g de alface a cada dois dias. Outros trabalhos lançam mão de dados locais ou regionais, como por exemplo, Seidu *et al.* (2008) em Gana e Bastos *et al.* (2008) no Brasil. No caso do trabalho de Seidu *et al.* (2008), os valores foram considerados em faixas segundo distribuição uniforme. No trabalho de Bastos *et al.* (2008), como já referido, a abordagem foi a de modelagem determinística.

Outros trabalhos lançam mão de um banco de dados detalhado de consumo de culturas diversas por unidade de peso corpóreo (agregando, assim, outras fontes de variabilidade) nos EUA (portanto, relativo a uma realidade específica e não necessariamente transferível para outros contextos). Não obstante, vale registrar que Hamilton *et al.* (2006a), utilizando esse banco de dados e aplicando análise de sensibilidade em modelagem estocástica, identificou essa variável (padrão de consumo de alimentos) como a mais importante na determinação dos riscos anuais de infecção por rotavírus.

Por fim, para efeito de cômputo da redução da concentração de patógenos em culturas irrigadas, muitos trabalhos usam as faixas recomendadas nas Diretrizes da OMS, propostas por Mara *et al.* (2007): 2-3 log para rotavírus e *Campylobacter* e 0-1 log para oocistos de

*Cryptosporidium*. Outros trabalhos abordam essa questão explicitamente em termos de decaimento natural segundo cinética de primeira ordem assumindo valores para o coeficiente de decaimento de vírus 0,69-1,07 d<sup>-1</sup> (Asano *et al.*, 1992; Petterson *et al.*, 2001). Amoah *et al.* (2007) demonstraram que práticas de higiene dos alimentos podem alcançar 1-2 log de remoção de ovos de *Ascaris*.

Como registrado, alguns dos trabalhos acima referidos procuram ajustar funções distribuição de probabilidades às diversas variáveis de modelos de exposição (algumas vezes de forma bem fundamentada, outras nem tanto). Mara *et al.* (2007) optaram por trabalhar com faixas de valores para todas as variáveis de acordo com distribuição uniforme, sob o argumento de que isso representa uma opção deliberada por um modelo mais simples, pois muito frequentemente não existem dados disponíveis ou adequados para o ajuste de distribuição de probabilidades.

### **3.5 - Normas brasileiras de qualidade da água para irrigação e padrões de lançamento de efluentes.**

Em 2005 o Conselho Nacional de Recursos Hídricos (CNRH) promulgou a Resolução nº 54 que estabelece modalidades, diretrizes e critérios gerais para a prática do reúso direto não potável de água no Brasil (BRASIL, 2005a), remetendo para regulamentação complementar os padrões de qualidade e os códigos de prática para as diversas modalidades de reúso. Esta resolução coloca a atividade de reúso de água como integrante das políticas de gestão de recursos hídricos vigentes no país.

A Resolução CONAMA nº 357/2005 (BRASIL, 2005b) estabelece critérios de enquadramento dos corpos de água superficiais, compatibilizando a qualidade da água com os usos previstos, incluindo a irrigação (Quadro 3.4).

Quadro 3.4 - Critérios de qualidade da água para irrigação estabelecidos na Resolução CONAMA nº 357/2005

Classe	Uso da água	Critério de qualidade
1	Irrigação de hortaliças que são consumidas cruas e de frutas que se desenvolvam rentes ao solo e que sejam ingeridas cruas sem remoção de película	200 coliformes termotolerantes / 100 mL
2	Irrigação de hortaliças, plantas frutíferas e de parques, jardins, campos de esporte e lazer, com os quais o público possa vir a ter contato direto; além da aquíicultura	1.000 coliformes termotolerantes / 100 mL
3	Irrigação de culturas arbóreas, cerealíferas e forrageiras	4000 coliformes termotolerantes / 100 mL

A Resolução CONAMA nº 357/2005 (BRASIL, 2005b) estabelece também padrões de lançamento de efluentes Para nitrogênio amoniacal total foi estabelecido valor máximo de 20 mg/L. A Resolução CONAMA nº 20/86, que precedia a nº 357/2005, estabelecia o limite de 5 mg/L (BRASIL, 1986). Posteriormente, a Resolução nº 397/2008 definiu que o limite de nitrogênio amoniacal total não será aplicável em sistemas de tratamento de esgotos sanitários (BRASIL, 2008).

No estado de Minas Gerais, a Deliberação Normativa Conjunta COPAM/CERH-MG n.º 001/2008 estabelece em seu artigo 19 que efluentes de qualquer fonte poluidora somente poderão ser lançados, direta ou indiretamente, nos corpos de água, após o devido tratamento e desde que obedeçam às seguintes condições (MINAS GERAIS, 2008):

- DBO: até 60 mg/L ou:
  - Tratamento com eficiência de redução de DBO em no mínimo 60% e média anual igual ou superior a 70% para sistemas de esgotos sanitários e de percolados de aterros sanitários municipais; e
  - Tratamento com eficiência de redução de DBO em no mínimo 75% e média anual igual ou superior a 85% para os demais sistemas.
- DQO: até 180 mg/L ou:
  - Tratamento com eficiência de redução de DQO em no mínimo 55% e média anual igual ou superior a 65% para sistemas de esgotos sanitários e de percolados de aterros sanitários municipais;
  - Tratamento com eficiência de redução de DQO em no mínimo 70% e média anual igual ou superior a 75% para os demais sistemas.

## **4 - CARACTERIZAÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADES DOS PARÂMETROS DE QUALIDADE DO ESGOTO BRUTO E DOS EFLUENTES DAS LAGOAS DE POLIMENTO DA ETE VIOLEIRA**

### **4.1 - Metodologia**

O banco de dados analisado, referente ao período de outubro de 2001 a setembro de 2008, é oriundo de trabalhos desenvolvidos na Unidade Integrada de Tratamento de Esgotos e Utilização de Efluentes, localizada no bairro Violeira, Viçosa – MG, instalada como unidade experimental no âmbito do Programa de Pesquisa em Saneamento Básico – PROSAB.

Até o início de 2003, o tratamento de esgotos na ETE Violeira, operada pelo Serviço Autônomo de Água e Esgotos (SAAE) de Viçosa, era realizado em um reator UASB. Após este período foi instalado um biofiltro aerado submerso (BF). O sistema UASB + BF, em escala real, é constituído por unidades pré-fabricadas em aço com as seguintes dimensões: (i) UASB – área = 8,4 m<sup>2</sup> (2,4 m x 2,4 m), altura útil = 5,7 m; (ii) BF - área = 2,9 m<sup>2</sup> (1,7 m x 1,7 m), altura útil = 2,0 m (Figura 4.1). Com população de projeto de 1.500 habitantes, esse sistema recebe hoje a contribuição de cerca de 800 habitantes e com uma vazão média estimada de 115 m<sup>3</sup> d<sup>-1</sup>.

A série experimental de lagoas de pós-tratamento (escala piloto, pré-fabricadas em fibra de vidro), em operação desde outubro de 2001, era constituída inicialmente por três unidades. No início de 2004 foi instalada uma quarta lagoa, com a possibilidade de ser operada em série ou em paralelo à terceira lagoa. Todas as lagoas apresentam área útil de 16,3 m<sup>2</sup> e relação comprimento / largura igual a 2,0 (Figura 4.1). Este sistema contava ainda com flexibilidade de variação da altura da lâmina d'água, o que proporcionou o monitoramento em onze períodos operacionais distintos (Tabela 4.1).



Figura 4.1 - ETE Viçosa, Viçosa, MG – vista do sistema UASB + BF (à esquerda) e série experimental de lagoas (à direita)

As variáveis aqui analisadas foram: demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), sólidos em suspensão totais (SST), nitrogênio amoniacal ( $\text{NH}_3$ ), fósforo total (P), pH, *E. coli* (EC), coliformes totais (CT) e ovos de helmintos (HEL).

Ao longo de todo o período de operação, o sistema foi monitorado em frequência variando entre semanal e quinzenal. Até o quarto período a amostragem dos efluentes das lagoas para a determinação das variáveis físicas e químicas era realizada de forma composta, entre 8 h e 18 h, a cada hora. A partir do quarto período a amostragem passou a ser simples (pontual), por coluna, em torno das 10 h. As coletas para quantificação das variáveis microbiológicas eram realizadas pontualmente nos efluentes de cada unidade de tratamento, na parte da manhã. O pH era determinado em campo com emprego de equipamentos portáteis. Cabe ressaltar que nem todas as variáveis foram determinadas sistemática ou ininterruptamente ao longo de todos os períodos de monitoramento. Portanto, o banco de dados ao longo dos diversos períodos não foi uniforme.

Ao longo do presente trabalho o banco de dados original foi complementado com monitoramento adicional para a pesquisa de ovos de helmintos, realizada de acordo com o método de Bailenger modificado (AYRES e MARA, 1996). As demais variáveis foram determinados em trabalhos anteriores de acordo com o especificado *no Standard Methods for the Examination of Water and Wastewater* (APHA, 1998).

TABELA 4.1 - Variações operacionais do Sistema de lagoas da Unidade Experimental da Viçosa, Viçosa – MG.

PERÍODO	L1			L2			L3			L4		
	Q	TDH	H	Q/A	H	TDH	Q	H	TDH	Q/A	H	Q/A
1 (out 01 - abr 02) <sup>(1)</sup>	1,50	9,40	0,90	0,096	0,90	9,40	1,50	0,90	9,40	0,096	0,90	0,096
2 (jun 02 - nov 02) <sup>(1)</sup>	2,00	7,05	0,90	0,127	0,90	7,05	2,00	0,90	7,05	0,127	0,90	0,127
3 (mar 03 - mai 03) <sup>(2)</sup>	2,00	7,05	0,90	0,127	0,70	5,38	2,00	0,70	2,31	0,127	0,30	0,127
4 (mar 04 - set 04) <sup>(2)(3)</sup>	4,20	3,36	0,90	0,268	0,90	3,36	4,20	0,90	5,14	0,134	0,70	0,134
5 (set 04 - jul 05) <sup>(2)(3)</sup>	3,00	4,70	0,90	0,191	0,90	4,70	3,00	0,90	7,20	0,096	0,70	0,096
6 (out 05 - mar 06) <sup>(2)(3)</sup>	1,50	5,10	0,50	0,096	0,40	4,10	1,50	0,40	18,8	0,048	0,90	0,048
7 (abr 06 - ago 06) <sup>(2)(4)</sup>	1,50	5,10	0,50	0,096	0,40	4,10	1,50	0,40	4,1	0,096	0,40	0,096
8 (set 06 - nov 06) <sup>(2)(4)</sup>	1,50	7,20	0,70	0,096	0,70	7,20	1,50	0,70	4,1	0,096	0,40	0,096
9 (nov 06 - fev 07) <sup>(1)(4)</sup>	1,50	7,20	0,70	0,096	0,70	7,20	1,50	0,70	4,1	0,096	0,40	0,096
10 (mar 07 - ago 07) <sup>(2)(4)</sup>	1,50	9,40	0,90	0,096	0,90	9,40	1,50	0,90	9,4	0,096	0,90	0,096
11 (set 07 - set 08) <sup>(1)(4)</sup>	2,50	5,80	0,90	0,159	0,90	5,80	2,50	0,90	5,8	0,159	0,90	0,159

Q: vazão (m<sup>3</sup> d<sup>-1</sup>); TDH: tempo de detenção hidráulica (d); H: profundidade (m); Q/A: taxa de aplicação hidráulica superficial (m<sup>3</sup> m<sup>-2</sup> d<sup>-1</sup>). (1) o sistema de lagoas recebia o efluente do UASB; (2) o sistema de lagoas recebia o efluente do BF; (3) lagoa 4 operada em paralelo a lagoa 3; (4) lagoa 4 operada em série.

Para efeito de sistematização do banco de dados, inicialmente, os “períodos iguais de operação” (mesma vazão Q, mesmo TDH, mesma taxa de aplicação hidráulica Q/A e mesma profundidade H) foram agrupados (a, b, c, etc.) para os dados da Lagoa 1 (L1), Lagoa 2 (L2), Lagoa 3 (L3), média das Lagoas 3 e 4 quando operadas em paralelo (L3//L4) e Lagoa 4 quando operada em série (L4). Os dados do esgoto bruto, reator anaeróbio e biofiltro não foram agrupados por não apresentarem diferentes períodos de operação.

Desses grupos, foram retirados os pontos atípicos, ou *outliers*. Existem vários métodos de identificação de *outliers*, dentre eles, um amplamente utilizado é a detecção por meio de gráfico box-plot. O diagrama box plot consiste em um retângulo definido pelo primeiro e pelo terceiro quartis ( $Q_1$  e  $Q_3$  respectivamente), contendo a mediana em seu interior. O ponto definido por  $(Q_3 + 1,5(Q_3 - Q_1))$  é considerado limite superior para a identificação de *outliers*. De modo análogo, o ponto definido por  $(Q_1 - 1,5(Q_3 - Q_1))$  é considerado limite inferior para a identificação de *outliers* (NAGHETTINI e PINTO, 2007).

A seguir foi aplicado o teste estatístico não paramétrico, Kruskal-Wallis com nível de significância  $\alpha=5\%$  entre os dados agrupados (a, b, c, etc.) para detectar grupos que não apresentavam diferenças significativas entre si em relação a cada um das variáveis analisadas (isto é, DBO, DQO, etc.), reunindo-os assim em novos grupos (A, B, C). Em seguida, cada um desses novos grupos de dados foi caracterizado por meio de estatística descritiva. Todos esses testes estatísticos foram realizados com o programa SAS.

Histogramas foram realizados para cada grupo de dados com o intuito de auxiliar na escolha da melhor distribuição de frequência. Não foram feitos histogramas para as séries com número de dados (N) menor que 30, por resultarem em histogramas, em geral, não representativos da forma de distribuição de frequência da população da qual foi retirada a amostra (SPIEGEL, 1993). Trabalhou-se com o número de intervalos de classe aproximadamente igual à raiz quadrada do número de observações. Estes histogramas foram realizados através do programa @Risk (PALISADE CORPORATION, 2009).

Em seguida foram efetuados os testes de hipóteses para aleatoriedade, independência, homogeneidade e estacionariedade. O teste de aleatoriedade foi baseado no número de inflexões. Considera-se inflexão um pico ou um vale de um gráfico construído pelos valores de uma determinada variável distribuídos cronologicamente. Um número muito grande ou muito pequeno de inflexões denota não-aleatoriedade (NAGHETTINI e PINTO, 2007). O teste de Wald Wolfowitz foi utilizado na constatação de independência dos dados. A homogeneidade foi verificada por meio do teste proposto por Mann e Whitney. E por fim, o

teste de Spearman foi empregado na verificação da estacionariedade dos dados detectando tendências temporais presentes nas séries de dados. Em todos esses testes foi assumido nível de significância de 5%.

Para testar a aderência das séries de dados a modelos de distribuição de probabilidades foram escolhidas cinco distribuições candidatas - normal, lognormal, uniforme, exponencial e Weibull – com base nos argumentos a seguir. Muitas variáveis são descritas pelo modelo normal ou podem ser transformadas neste de maneira relativamente simples, o que a torna a mais importante entre as distribuições. O modelo lognormal foi escolhido por se mostrar apropriado para descrever variáveis dos constituintes dos esgotos, como descrito por diversos autores e comentado na Seção 3.1.1.1. A distribuição uniforme tem sido normalmente utilizada nas Simulações Monte Carlo, portanto, sua escolha deu-se com o intuito de verificar a coerência de seu uso. A distribuição exponencial é largamente conhecida, de fácil aplicação e tem sido empregada para a descrição de eventos ambientais. A distribuição Weibull foi selecionada como alternativa de distribuição com curva flexível, além da distribuição lognormal.

A aderência dos dados às distribuições acima mencionadas foi verificada por meio dos testes Qui-quadrado ( $\chi^2$ ), Kolmogorov Smirnov (KS), Anderson Darling (AD), além do gráfico de probabilidades. Por terem sido feitas em planilhas simples de Excel, os testes Filliben (FIL) foram aplicados nas verificações de aderências somente às distribuições normal e lognormal. Em seguida, as séries de dados foram testadas quanto à normalidade por meio do teste Shapiro Wilk (SW). O gráfico de probabilidades foi aplicado às séries com o número de dados maior ou igual a 30, pois séries com dados escassos apresentam maior dificuldade para a tomada de decisão quanto à distribuição candidata, somado ao fato de que se trata de uma avaliação visual do gráfico. Todos os testes foram considerados significativos ao nível de 5%.

Os testes de aderência  $\chi^2$ , AD e KS foram realizados com o programa @ Risk, enquanto o teste de normalidade SW pelo programa Stat Tolls 5.5. Os testes de hipóteses, o teste FIL de aderência e o gráfico de probabilidades foram realizados por meio de planilhas do Excel.

Com base em análise conjunta de todos esses procedimentos procurou-se escolher a distribuição que melhor descrevesse cada série de dados considerando as particularidades de cada teste de aderência, os resultados dos testes de hipóteses, as análises dos histogramas e as informações relevantes obtidas pela estatística descritiva, principalmente os valores de

assimetria e curtose. Alguns testes, devido suas peculiaridades, não foram aplicáveis em alguns casos, por exemplo, séries com valor de N muito pequeno.

## 4.2 - Resultados e discussão

### 4.2.1 - Verificação de diferenças entre dados dos distintos períodos operacionais

Os Quadros 4.1 e 4.2 mostram, respectivamente, o agrupamento dos períodos com as mesmas características operacionais e o agrupamento segundo os resultados dos testes de Kruskal-Wallis aplicados para a identificação de dados de qualidade dos efluentes que não apresentam diferenças estatisticamente significativas.

Quadro 4.1 - Agrupamento dos períodos com mesmas características de operação das lagoas de polimento.

Lagoas	Períodos agrupados							
	a	b	C	d	e	f	g	h
<b>L1</b>	1; 10	2; 3	4	5	6; 7	8; 9	11	
<b>L2</b>	1; 10	2	3	4	5	6; 7	8; 9	11
<b>L3</b>	1; 10	2	3	4	5	6	7; 8; 9	11
<b>L3//L4</b>	4	5	6					
<b>L4</b>	7; 10	8; 9	11					

L1: lagoa 1, L2: lagoa 2, L3: lagoa 3; L3//L4: lagoa 3 operando em paralelo à lagoa 4, lagoa 4 operando em série

Quadro 4.2 - Agrupamento dos períodos operacionais das lagoas de polimento de acordo com os resultados do teste de Kruskal-Wallis, com nível de 5% de significância, para identificação de dados de qualidade dos efluentes que não apresentam diferenças estatisticamente significativas.

Variável	L1			L2			L3			L4			L3//L4	
	A	B		A	B	C	A	B	C	A	B		A	B
DBO	1; 2; 3; 10	4; 5; 6; 7; 8; 9		2; 8; 9	1; 3; 4; 5; 6; 7; 10; 11		1; 3; 10; 11	2; 4; 5; 6; 7; 8; 9		7; 8; 9; 10; 11			4; 5; 6	
DQO	1; 2; 3; 6; 7; 10	4; 5; 8; 9		2; 8; 9	1; 6; 7; 10	3; 4; 5; 11	2	3; 7; 8; 9	1; 4; 5; 6; 10; 11	7; 8; 9; 10; 11			4; 5; 6	
SST	2; 3; 6; 7; 8; 9; 10	4; 5; 11		2; 6; 7; 8; 9; 10; 11	3; 4; 5; 10; 11		2	1; 6; 7; 8; 9; 10; 11	3; 4; 5	7; 8; 9; 10; 11			4; 5; 6	
pH	1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11			1; 2; 3; 4; 5; 10; 11	6; 7; 8; 9		1; 10; 2; 4	3; 5; 6; 7; 8; 9; 11		7; 10	8; 9; 11	4	5; 6	
Fósforo	1; 8; 9; 10	2; 3; 4; 5; 6; 7; 11		1; 10	2; 3; 4; 5; 6; 7; 8; 9; 11		1; 3; 10; 11	2; 4; 5; 6; 7; 8; 9		7; 8; 9; 10; 11			4; 5; 6	
Nitrogênio Amoniacal	1; 8; 9; 10	4; 5; 6; 7; 11		1; 2; 10; 11	3; 4; 5; 6; 7; 8; 9		1; 2; 3; 10	4; 5; 6; 7; 8; 9; 11		7; 10	8; 9; 11	4; 5; 6		
<i>E. coli</i>	1; 2; 3; 4; 5; 6; 7; 8; 9			4; 6; 7	1; 2; 3; 5; 8; 9		4; 5; 7; 8; 9	1; 2; 3; 6; 10; 11		7; 8; 9; 10; 11		5	4; 6	
Coliformes totais	1; 2; 3; 4; 5; 6; 7; 8; 9			4; 5; 6; 7; 8; 9	1; 2; 3		4; 5; 6; 7; 8; 9	1; 2; 3; 10		7; 8; 9; 10; 11		4; 5; 6		
Helmintos	1; 2; 3; 4; 5; 6; 7; 8; 9; 10	11		9; 11										

L1: lagoa 1, L2: lagoa 2, L3: lagoa 3, L3//L4: lagoa 3 operando em paralelo à lagoa 4, L4: lagoa 4 operando em série

#### *4.2.2 – Estatística descritiva dos dados do esgoto bruto, reator anaeróbio, biofiltro e dos efluentes das lagoas de polimento agrupados por períodos operacionais com resultados estatisticamente equivalentes*

Nas Tabelas 4.2 a 4.10 constam a estatística descritiva para cada série de dados referentes a DBO, DQO, SST, NH<sub>3</sub>, P, pH, *E. coli*, CT e ovos de helmintos do esgoto bruto, reator anaeróbio, biofiltro e dos efluentes da lagoa 1, lagoa 2, lagoa 3, média das Lagoas 3 e 4 quando operadas em paralelo e lagoa 4 quando operadas em série.

A comparação dos valores de coeficientes de assimetria e de curtose das séries de dados com os das distribuições candidatas auxilia na verificação da aderência dos dados a essas distribuições. Como visto na Seção 3.1.1.3, as distribuições normal, exponencial e uniforme possuem coeficientes de curtose e de assimetria constantes (respectivamente, 3 e 0 para a distribuição normal, 9 e 2 para a distribuição exponencial e 1,8 e 0 para a distribuição uniforme). Já as distribuições lognormal e Weibull apresentam valores de coeficientes de curtose e de assimetria em função de sua variância e dos parâmetros da população.

Oitenta por cento do número total de séries analisadas, ou seja, a grande maioria, apresentou coeficiente de assimetria com valor positivo, denotando distribuições com caudas à direita. Em geral, percebe-se que os valores negativos de assimetria foram encontrados em séries com baixo valor de N, exceto com os dados de nitrogênio amoniacal e fósforo.

A grande maioria dos coeficientes de curtose foi menor que três, próximas a zero, o que caracteriza distribuições platicúrticas, isto é, com topos achatados. Entretanto, alguns dados apresentaram coeficientes de curtose bastante elevados (ovos de helmintos, Lagoa 1, série de dados A; *E. coli*, Esgoto bruto, Lagoa 3, série de dados A; coliformes totais, Lagoa 1), típicos de distribuições leptocúrticas com picos altos. Porém, cabe lembrar que o coeficiente de curtose é elucidativo apenas em distribuições aproximadamente simétricas, pois somente denotam o grau de achatamento da curva.

De toda forma, pode-se dizer, em primeira análise, que a maioria das distribuições parece diferir do esperado para a distribuição normal no que diz respeito à assimetria e curtose.

TABELA 4.2 - Estatística descritiva das séries de dados de DBO, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2		L3		L3//L4		L4	
				A	B	A	B	A	B	A	B	A	B
N	140	133	143	46	89	23	101	36	88	28	25	37	37
Média (mg/L)	303	64	44	46	28	35	26	17	27	23	33	30	30
Mediana (mg/L)	311	60	41	39	27	37	26	15	27	22	32	29	29
Moda (mg/L)	225	72	37	30	20	36	25	12	23	-	32	14	14
1º quartil (mg/L)	230	46	31	30	20	27	21	13	20	18	30	17	17
3º quartil (mg/L)	359	76	58	58	36	43	30	22	32	29	36	36	36
Amplitude (mg/L)	467	130	93	129	47	35	46	23	39	32	17	62	62
Variância (mg/L) <sup>2</sup>	9933	831	411	702	129	101	80	30	65	64	20	229	229
C. Variação (%)	33	45	46	57	40	29	34	32	30	35	13	51	51
Assimetria	0,08	0,67	0,76	1,34	0,48	-0,13	0,62	0,75	-0,01	0,21	0,22	0,69	0,69
Curtose	-0,20	0,36	0,33	2,45	-0,53	-0,80	0,94	-0,18	-0,09	-0,20	-0,26	-0,24	-0,24

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.3 - Estatística descritiva das séries de dados de DQO, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1			L2			L3			L3//L4		L4
				A	B	A	B	C	A	B	C	A	A		
N	139	130	94	73	68	24	53	62	10	44	92	52	52	36	
Média (mg/L)	703	200	150	179	132	196	174	125	277	199	142	157	157	186	
Mediana (mg/L)	694	197	143	164	127	207	179	121	284	189	133	148	148	169	
Moda (mg/L)	512	197	133	103	147	-	156	95	-	-	106	139	139	133	
1º quartil (mg/L)	533	137	109	127	97	142	121	95	249	142	105	132	132	124	
3º quartil (mg/L)	875	245	192	213	167	223	218	150	305	235	167	175	175	249	
Amplitude (mg/L)	1240	407	275	403	197	274	291	176	205	276	317	167	167	348	
Variância (mg/L) <sup>2</sup>	61397	8213	3621	5858	2230	4104	4460	1643	4312	4893	3807	1391	1391	7686	
C. Variação (%)	35	45	40	43	36	33	38	32	24	35	43	24	24	47	
Assimetria	0,28	0,67	0,45	0,92	0,36	0,83	-0,04	0,46	-0,21	0,88	0,89	0,89	0,89	0,15	
Curtose	0,14	0,28	0,14	1,05	-0,55	1,53	-0,49	-0,11	-0,39	0,06	1,07	0,65	0,65	-0,78	

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralela à L4; L4: lagoa 4 operando em série ; A, B e C: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.4 - Estatística descritiva das séries de dados de SST, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008.

Parâmetro	RA	BF	L1			L2			L3			L3//L4			L4		
			A	B	A	B	A	B	A	B	C	A	A	A	A		
N	113	93	67	53	49	76	9	48	57	54	45						
Média (mg/L)	69	48	73	49	103	52	159	83	58	64	81						
Mediana (mg/L)	63	43	66	45	96	48	143	75	56	62	70						
Moda (mg/L)	47	63	30	43	58	68	-	79	31	44	118						
1º quartil (mg/L)	43	31	44	33	66	31	123	50	43	47	52						
3º quartil (mg/L)	85	63	99	60	128	68	200	109	71	79	112						
Amplitude (mg/L)	153	113	131	101	197	119	195	182	92	85	175						
Variância (mg/L) <sup>2</sup>	1086	649	1164	611	1780	752	3540	1953	444	509	1968						
C. Variação (%)	48	54	47	51	41	53	38	53	37	36	55						
Assimetria	0,80	0,82	0,28	0,76	0,69	0,47	0,61	0,50	-0,01	0,25	0,38						
Curtose	0,34	0,28	-1,00	0,14	0,64	-0,20	0,09	-0,33	-0,61	-0,73	-0,51						

N: número de dados; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A, B e C: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.5 - Estatística descritiva das séries de dados de Nitrogênio amoniacoal, Unidade Experimental da Viçeira, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2		L3		L3//L4		L4	
				A	B	A	B	A	B	A	B	A	B
N	131	119	99	41	107	34	110	36	96	52	35		
Média (mg/L)	23	47	47	25	34	9	25	2	15	12	11		
Mediana (mg/L)	22	48	47	24	36	8	27	2	15	11	13		
Moda (mg/L)	21	49	49	25	38	8	32	0	13	1	13		
1º quartil (mg/L)	18	43	43	18	29	7	19	0	8	7	5		
3º quartil (mg/L)	28	51	51	32	40	10	32	3	21	16	15		
Amplitude (mg/L)	35	32	34	38	41	16	36	5	28	28	20		
Variância (mg/L) <sup>2</sup>	52	43	49	112	81	13	86	2	64	45	37		
C. Variação (%)	32	14	15	43	27	43	38	81	54	56	57		
Assimetria	0,26	-0,51	-0,52	0,33	-0,95	0,58	-0,67	0,46	-0,19	0,56	-0,52		
Curtose	0,00	0,53	0,51	-0,83	0,32	0,49	-0,33	-0,67	-1,02	-0,06	-0,97		

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralela à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.6 - Estatística descritiva das séries de dados de Fósforo, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2		L3		L3//L4		L4
				A	B	A	B	A	B	A	B	
N	127	118	94	39	99	22	119	40	97	54	37	
Média (mg/L)	9	9	9	7	9	5	9	4	9	9	7	
Mediana (mg/L)	9	9	10	7	9	5	9	4	9	8	7	
Moda (mg/L)	8	9	-	8	8	-	10	5	7	-	5	
1º quartil (mg/L)	8	8	8	5	8	4	7	3	7	5	5	
3º quartil (mg/L)	11	11	11	9	11	6	11	5	11	12	9	
Amplitude (mg/L)	12	13	12	9	14	4	14	8	14	14	7	
Variância (mg/L) <sup>2</sup>	6,6	7,5	7,7	5,7	8,4	1,0	10,1	4,0	10,4	16,3	4,5	
C. Variação (%)	28	30	30	35	31	20	36	46	37	47	30	
Assimetria	0,13	-0,10	-0,05	-0,53	0,17	1,15	0,17	0,42	-0,07	-0,03	-0,02	
Curtose	-0,17	0,07	-0,34	-0,23	0,00	1,22	-0,43	-0,56	-0,47	-1,22	-1,33	

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.7 - Estatística descritiva das séries de dados de pH, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

Parâmetro	L1		L2		L3		L3//L4		L4	
	A	B	A	B	A	B	A	B	A	B
N	142	97	44	56	84	22	33	24	16	
Média (mg/L)	7,65	8,06	7,68	9,07	8,29	8,92	8,35	7,78	8,44	
Mediana (mg/L)	7,58	8,02	7,68	9,09	8,17	8,65	8,26	7,68	8,53	
Moda (mg/L)	7,6	7,7	7,8	9,2	8,7	8,9	8,2	7,7	8,7	
1º quartil (mg/L)	7,39	7,68	7,44	8,40	7,85	8,33	7,94	7,57	8,15	
3º quartil (mg/L)	7,90	8,41	7,84	9,63	8,71	9,34	8,76	7,92	8,72	
Amplitude (mg/L)	1,80	3,24	1,45	3,81	2,55	3,16	2,06	1,21	2,16	
Variância (mg/L) <sup>2</sup>	0,14	0,30	0,10	0,72	0,32	0,73	0,29	0,09	0,30	
C. Variação (%)	4,9	6,8	4,1	9,4	6,8	9,6	6,4	3,9	6,5	
Assimetria	0,61	0,83	0,53	0,11	0,43	0,66	0,11	1,75	-0,56	
Curtose	-0,17	1,61	0,55	-0,31	-0,40	-0,39	-0,83	3,12	0,56	

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.8 - Estatística descritiva das séries de dados de *E. coli*, Unidade Experimental da Viçeira, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2		L3		L3/L4		L4
				A	B	A	B	A	B	A	B	
N	99	84	46	86	60	21	39	34	10	10	10	15
Média (mg/L)	$7,7 \times 10^7$	$3,4 \times 10^6$	$3,3 \times 10^6$	$1,7 \times 10^5$	$4,5 \times 10^3$	$4,9 \times 10^4$	$4,4 \times 10^3$	$2,9 \times 10^2$	$2,5 \times 10^2$	$1,9 \times 10^2$	$1,9 \times 10^3$	$3,4 \times 10^1$
Mediana (mg/L)	$9,8 \times 10^6$	$3,0 \times 10^6$	$2,7 \times 10^6$	$1,6 \times 10^5$	$3,2 \times 10^3$	$1,9 \times 10^4$	$3,1 \times 10^1$	$4,1 \times 10^2$	$1,8 \times 10^2$	$1,2 \times 10^3$	$1,2 \times 10^3$	$1,0 \times 10^2$
Moda (mg/L)	$2,4 \times 10^7$	$4,9 \times 10^6$	$3,8 \times 10^6$	$2,4 \times 10^5$	$1,0 \times 10^3$	$6,4 \times 10^4$	$2,0 \times 10^0$	$1,0 \times 10^2$	$1,0 \times 10^2$	$1,5 \times 10^2$	$1,5 \times 10^2$	$1,0 \times 10^1$
1º quartil (mg/L)	$3,3 \times 10^6$	$1,7 \times 10^6$	$1,1 \times 10^6$	$5,2 \times 10^4$	$1,0 \times 10^3$	$1,0 \times 10^4$	$1,2 \times 10^1$	$1,0 \times 10^2$	$1,0 \times 10^2$	$1,5 \times 10^2$	$1,5 \times 10^2$	$5,2 \times 10^1$
3º quartil (mg/L)	$2,4 \times 10^7$	$4,8 \times 10^6$	$4,8 \times 10^6$	$2,4 \times 10^5$	$5,6 \times 10^3$	$6,4 \times 10^4$	$1,2 \times 10^2$	$1,2 \times 10^3$	$3,7 \times 10^2$	$2,8 \times 10^3$	$2,8 \times 10^3$	$2,0 \times 10^2$
Amplitude (mg/L)	$2,4 \times 10^9$	$1,1 \times 10^7$	$1,1 \times 10^7$	$7,4 \times 10^5$	$1,8 \times 10^4$	$2,4 \times 10^5$	$1,6 \times 10^3$	$7,8 \times 10^3$	$6,0 \times 10^2$	$5,2 \times 10^3$	$5,2 \times 10^3$	$2,2 \times 10^7$
Variância (mg/L) <sup>2</sup>	$1,2 \times 10^{17}$	$5,6 \times 10^{12}$	$8,4 \times 10^{12}$	$2,7 \times 10^{10}$	$2,1 \times 10^7$	$3,8 \times 10^9$	$9,1 \times 10^4$	$3,4 \times 10^7$	$3,6 \times 10^5$	$3,9 \times 10^5$	$3,9 \times 10^6$	$4,3 \times 10^2$
C. Variação (%)	445	71	87	92	101	125	220	154	77	106	192	192
Assimetria	6,53	1,04	1,25	1,59	1,38	2,00	3,77	2,31	1,14	0,77	0,77	2,38
Curtose	42,96	1,20	1,20	2,99	1,18	4,09	15,58	5,13	0,87	-0,80	-0,80	4,44

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3/L4: lagoa 3 operando em paralela à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.9 - Estatística descritiva das séries de dados de coliformes totais, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2		L3		L3//L4		L4
				A	B	A	B	A	B	A	B	
N	103	105	54	98	40	43	42	36	19		17	
Média (mg/L)	$7,4 \times 10^7$	$1,4 \times 10^7$	$1,2 \times 10^7$	$1,5 \times 10^6$	$1,1 \times 10^6$	$1,2 \times 10^5$	$9,8 \times 10^2$	$8,7 \times 10^4$	$1,2 \times 10^5$		$1,2 \times 10^4$	
Mediana (mg/L)	$5,4 \times 10^7$	$1,2 \times 10^7$	$1,1 \times 10^7$	$5,4 \times 10^5$	$1,5 \times 10^4$	$2,4 \times 10^5$	$6,1 \times 10^2$	$3,7 \times 10^4$	$6,9 \times 10^4$		$1,0 \times 10^4$	
Moda (mg/L)	$2,4 \times 10^7$	$2,4 \times 10^7$	$2,4 \times 10^7$	$2,4 \times 10^5$	$2,4 \times 10^3$	$2,4 \times 10^5$	$2,4 \times 10^3$	$2,4 \times 10^5$	$2,4 \times 10^5$		$2,4 \times 10^4$	
1º quartil (mg/L)	$2,4 \times 10^6$	$5,9 \times 10^6$	$3,7 \times 10^6$	$2,4 \times 10^5$	$2,9 \times 10^3$	$8,2 \times 10^4$	$9,1 \times 10^1$	$2,4 \times 10^4$	$2,1 \times 10^4$		$2,2 \times 10^3$	
3º quartil (mg/L)	$1,3 \times 10^8$	$2,0 \times 10^7$	$1,7 \times 10^7$	$1,2 \times 10^6$	$3,0 \times 10^4$	$4,3 \times 10^5$	$1,6 \times 10^3$	$1,0 \times 10^5$	$2,4 \times 10^5$		$2,4 \times 10^4$	
Amplitude (mg/L)	$2,4 \times 10^8$	$4,6 \times 10^7$	$3,1 \times 10^7$	$2,4 \times 10^7$	$1,6 \times 10^5$	$2,4 \times 10^6$	$3,3 \times 10^3$	$2,4 \times 10^5$	$3,0 \times 10^5$		$2,4 \times 10^4$	
Variância (mg/L) <sup>2</sup>	$4,3 \times 10^{15}$	$9,6 \times 10^{13}$	$7,3 \times 10^{13}$	$1,4 \times 10^{13}$	$2,3 \times 10^9$	$3,6 \times 10^{11}$	$9,2 \times 10^5$	$7,8 \times 10^9$	$1,2 \times 10^{10}$		$9,3 \times 10^7$	
C. Variação (%)	88	70	74	252	140	143	98	102	93		81	
Assimetria	0,94	0,84	0,44	5,22	1,71	2,55	0,66	1,02	0,31		0,20	
Curtose	-0,18	0,48	-1,03	28,09	1,84	5,80	-0,85	-0,58	-1,73		-1,66	

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3//L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

TABELA 4.10 - Estatística descritiva das séries de dados de ovos de helmintos, Unidade Experimental da Viçosa, Viçosa – MG, 2001-2008.

Parâmetro	EB	RA	BF	L1		L2
				A	B	
N	148	141	86	118	10	10
Média (mg/L)	50	18	18	0	9	7
Mediana (mg/L)	40	13	17	0	10	8
Moda (mg/L)	0	0	0	0	0	8
1º quartil (mg/L)	9	3	7	0	0	3
3º quartil (mg/L)	83	27	27	0	13	10
Amplitude (mg/L)	196	75	60	7	20	10
Variância (mg/L) <sup>2</sup>	2158	343	202	1	60	14
C. Variação (%)	92	101	81	3	1	1
Assimetria	0,87	1,19	0,82	3,33	0,07	-0,52
Curtose	0,13	0,84	0,30	12,70	-1,80	-1,23

N: número de dados; EB: esgoto bruto; RA: reator anaeróbio; BF: biofiltro; L1: lagoa 1; L2: lagoa 2; L3 lagoa 3; L3/L4: lagoa 3 operando em paralelo à L4; L4: lagoa 4 operando em série ; A e B: grupos de dados com resultados estatisticamente equivalentes.

#### 4.2.3 - Testes de aleatoriedade, estacionariedade, homogeneidade e independência

As Tabelas 4.11 a 4.19 trazem os resultados dos testes de aleatoriedade, estacionariedade, homogeneidade e independência. Os requisitos dessas hipóteses foram considerados atendidos pelas séries de dados quando os respectivos testes resultaram em valor *p* igual ou superior ao nível de significância de 5%.

TABELA 11 - Resultados (valor *p*) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de DBO no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008

Amostra	Aleatoriedade	Estacionariedade	Homogeneidade	Independência
<b>EB</b>	0,1065 (S)	0,0101 (N)	0,0012 (N)	< 0,0001 (N)
<b>RA</b>	0,1897 (S)	< 0,0001 (N)	< 0,0001 (N)	< 0,0001 (N)
<b>BF</b>	0,8418 (S)	0,7148 (S)	0,5501 (S)	< 0,0001 (N)
<b>L1 A</b>	0,9053 (S)	0,0120 (N)	0,7159 (S)	0,0001 (N)
<b>L1 B</b>	0,1275 (S)	0,9760 (S)	0,7755 (S)	0,0147 (N)
<b>L2 A</b>	> 0,9999 (S)	0,0961 (S)	0,0409 (N)	0,2026 (N)
<b>L2 B</b>	0,2338 (S)	0,2920 (S)	0,9526 (S)	< 0,0001 (N)
<b>L3A</b>	0,1369 (S)	0,5249 (S)	0,7563 (S)	0,0926 (S)
<b>L3 B</b>	0,1730 (S)	0,2911 (S)	0,7989 (S)	0,0021 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,5366 (S)	0,3465 (S)	0,2701 (S)	0,0539 (S)
<b>L3//L4* B</b>	0,8696 (S)	0,0232 (N)	0,0246 (N)	0,0687 (S)
<b>L4 A</b>	0,3509 (S)	0,0277 (N)	0,0306 (N)	0,0352 (N)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; <sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.12 - Resultados (valor *p*) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de ovos helmintos no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>EB</b>	0,0671 (S)	< 0,0001(N)	0,0008(N)	< 0,0001(N)
<b>RA</b>	0,5028 (S)	< 0,0001 (N)	0,0006 (N)	< 0,0001 (N)
<b>BF</b>	0,7960 (S)	0,0796 (S)	0,2217 (S)	< 0,0001 (N)
<b>L1 A</b>	< 0,0001 (N)	< 0,0001(N)	0,7521 (S)	< 0,0001 (N)
<b>L1 B</b>	0,7823 (S)	0,0027(N)	> 0,9999 (S)	0,8834 (S)
<b>L2 A</b>	0,0057(N)	0,0253(N)	0,6857 (S)	0,1877 (S)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.13 - Resultados (valor *p*) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de DQO no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>EB</b>	0,0915 (S)	0,4310 (S)	0,9760 (S)	< 0,0001(N)
<b>RA</b>	0,9443 (S)	< 0,0001 (N)	0,2435 (S)	< 0,0001 (N)
<b>BF</b>	0,8692 (S)	< 0,0001 (N)	0,0910 (S)	< 0,0001 (N)
<b>L1 A</b>	0,3488 (S)	0,2580 (S)	0,8339 (S)	0,0003(N)
<b>L1 B</b>	0,5599 (S)	0,8413 (S)	0,0623 (S)	0,0001(N)
<b>L2 A</b>	0,4014 (S)	0,0406(N)	0,0234(N)	0,0198(N)
<b>L2 B</b>	> 0,9999 (S)	0,0148(N)	0,0023(N)	0,0144(N)
<b>L2 C</b>	0,3591 (S)	0,0344(N)	0,0546 (S)	0,0043(N)
<b>L3 A</b>	0,0531 (S)	0,0230(N)	0,3429 (S)	0,3319 (S)
<b>L3 B</b>	0,7150 (S)	0,0127(N)	0,0244(N)	0,0967 (S)
<b>L3 C</b>	0,6174 (S)	0,1098 (S)	0,4243 (S)	< 0,0001(N)
<b>L3//L4(*) A</b>	0,0256 (N)	0,6629 (S)	0,2903 (S)	0,0232(N)
<b>L4 A</b>	0,0584 (S)	0,8910 (S)	0,7697 (S)	0,0782 (S)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 14 - Resultados (valor  $p$ ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de SST nos efluentes das lagoas de polimento, Unidade Experimental da Viçeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>RA</b>	0,6528 (S)	0,1870 (S)	0,7304 (S)	0,0019 (N)
<b>BF</b>	0,2818 (S)	0,4265 (S)	0,6013 (S)	< 0,0001 (N)
<b>L1 A</b>	0,6244 (S)	<0,0001 (N)	0,0335 (N)	0,5945 (S)
<b>L1 B</b>	0,0203 (N)	0,5198 (S)	0,1679 (S)	0,0005 (N)
<b>L2 A</b>	0,0213 (N)	0,1728 (S)	0,0969 (S)	0,3721 (S)
<b>L2 B</b>	0,2328 (S)	0,8656 (S)	0,7396 (S)	0,3157 (S)
<b>L3 A</b>	0,1548 (S)	0,4795 (S)	0,4000 (S)	0,8681 (S)
<b>L3 B</b>	0,4309 (S)	0,9733 (S)	0,4752 (S)	0,9019 (S)
<b>L3 C</b>	0,1690 (S)	0,1553 (S)	0,8796 (S)	0,0014 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,7681 (S)	0,2111 (S)	0,7557 (S)	<0,0001 (N)
<b>L4 A</b>	0,9042 (S)	0,0441 (N)	0,0129 (N)	0,0322 (N)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.15 - Resultados (valor  $p$ ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de nitrogênio amoniacal no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Viçeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>EB</b>	0,2968 (S)	0,0987 (S)	0,0421 (N)	< 0,0001 (N)
<b>RA</b>	0,1887 (S)	0,0100 (N)	0,8358 (S)	< 0,0001 (N)
<b>BF</b>	0,0651 (S)	0,1825 (S)	0,0100 (N)	< 0,0001 (N)
<b>L1 A</b>	0,0230 (N)	0,7153 (S)	0,0445 (N)	< 0,0001 (N)
<b>L1 B</b>	0,6437 (S)	0,0046 (N)	0,3331 (S)	<0,0001 (N)
<b>L2 A</b>	0,8892 (S)	0,0456 (N)	0,0616 (S)	0,0189 (N)
<b>L2 B</b>	< 0,0001 (N)	0,1399 (S)	0,0028 (N)	< 0,0001 (N)
<b>L3 A</b>	0,5886 (S)	0,0006 (N)	0,1293 (S)	0,0101 (N)
<b>L3 B</b>	0,0003 (N)	0,8817 (S)	0,1122 (S)	< 0,0001 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,4347 (S)	0,8415 (S)	0,0055 (N)	< 0,0001 (N)
<b>L4 A</b>	0,6806 (S)	0,0046 (N)	0,0281 (N)	0,0444 (N)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.16 - Resultados (valor  $p$ ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de fósforo no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

Amostra	Aleatoriedade	Estacionariedade	Homogeneidade	Independência
<b>EB</b>	0,1794 (S)	0,5822 (S)	0,1022 (S)	< 0,0001 (N)
<b>RA</b>	0,2125 (S)	0,0121 (N)	0,0002 (N)	< 0,0001 (N)
<b>BF</b>	0,1177 (S)	0,0043 (N)	0,0001 (N)	< 0,0001 (N)
<b>L1 A</b>	0,0007 (N)	0,4984 (S)	0,4749 (S)	< 0,0001 (N)
<b>L1 B</b>	0,6884 (S)	0,2287 (S)	0,0008 (N)	< 0,0001 (N)
<b>L2 A</b>	0,1592 (S)	0,9320 (S)	0,5787 (S)	0,8808 (S)
<b>L2 B</b>	< 0,0001 (N)	0,7720 (S)	0,033 (N)	< 0,0001 (N)
<b>L3 A</b>	0,0407 (N)	< 0,0001 (N)	< 0,0001 (N)	< 0,0001 (N)
<b>L3 B</b>	0,0059 (N)	0,6908 (S)	0,0030 (N)	< 0,0001 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,2738 (S)	0,5867 (S)	0,5705 (S)	< 0,0001 (N)
<b>L4 A</b>	0,5940 (S)	0,0058 (N)	0,4283 (S)	0,0056 (N)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; <sup>(\*)</sup> período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.17 - Resultados (valor  $p$ ) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de pH nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

Amostra	Aleatoriedade	Estacionariedade	Homogeneidade	Independência
<b>L1 A</b>	< 0,0001 (N)	< 0,0001 (N)	< 0,0001 (N)	< 0,0001 (N)
<b>L2 A</b>	0,8713 (S)	< 0,0001 (N)	0,5009 (S)	0,0017 (N)
<b>L2 B</b>	0,4652 (S)	< 0,0001 (N)	0,1833 (S)	0,0002 (N)
<b>L3 A</b>	0,7473 (S)	< 0,0001 (N)	0,1598 (S)	0,0001 (N)
<b>L3 B</b>	0,4854 (S)	< 0,0001 (N)	0,0033 (N)	0,0010 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,4815 (S)	< 0,0001 (N)	0,4470 (S)	0,0174 (N)
<b>L3//L4<sup>(*)</sup> B</b>	0,7771 (S)	< 0,0001 (N)	0,4068 (S)	0,0161 (N)
<b>L4 A</b>	0,1794 (S)	< 0,0001 (N)	0,0017 (N)	0,0002 (N)
<b>L4 B</b>	0,8338 (S)	0,0002 (N)	0,7382 (S)	0,3962 (S)

EB: esgoto bruto; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; <sup>(\*)</sup> período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 18 - Resultados (valor *p*) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de *E. coli* no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>EB</b>	0,2738 (S)	0,2993 (S)	0,5174 (S)	0,0011 (N)
<b>RA</b>	0,5500 (S)	< 0,0001 (N)	0,9929 (S)	0,5709 (S)
<b>BF</b>	0,5521 (S)	0,8546 (S)	0,4354 (S)	0,0692 (S)
<b>L1 A</b>	0,7960 (S)	0,7261 (S)	0,7391 (S)	0,0025 (N)
<b>L2 A</b>	0,6043 (N)	0,3104 (S)	> 0,9999 (S)	0,0413 (N)
<b>L2 B</b>	0,7181 (S)	0,1531 (S)	0,0050 (N)	0,0757 (S)
<b>L3 A</b>	0,3642 (S)	0,0034 (N)	0,0096 (N)	0,4492 (S)
<b>L3 B</b>	0,0178 (N)	0,3171 (S)	0,1607 (S)	0,0391 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,0271 (N)	0,0586 (S)	0,6857 (S)	0,6706 (S)
<b>L3//L4<sup>(*)</sup> B</b>	0,7823 (S)	0,0242 (N)	0,2000 (S)	0,4127 (S)
<b>L4 A</b>	0,8277 (S)	0,4955 (S)	> 0,9999 (S)	0,0100 (N)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

TABELA 4.19 - Resultados (valor *p*) dos testes de aleatoriedade, estacionariedade, homogeneidade e independência para os dados de concentração de coliformes totais no esgoto bruto e nos efluentes das lagoas de polimento, Unidade Experimental da Violeira, Viçosa – MG, 2001-2008.

<b>Amostra</b>	<b>Aleatoriedade</b>	<b>Estacionariedade</b>	<b>Homogeneidade</b>	<b>Independência</b>
<b>EB</b>	0,00227(N)	0,4296 (S)	0,3040 (S)	0,0012(N)
<b>RA</b>	0,1392 (S)	0,5988 (S)	0,2460 (S)	0,2237 (S)
<b>BF</b>	0,1897 (S)	0,4265 (S)	0,0006 (N)	< 0,0001 (N)
<b>L1 A</b>	0,6286 (S)	0,0427 (N)	0,0710 (S)	0,0018 (N)
<b>L2 A</b>	0,3705 (S)	0,0455 (N)	0,1040 (S)	0,0003 (N)
<b>L2 B</b>	0,0846 (S)	0,2073 (S)	0,1092 (S)	0,0001(N)
<b>L3 A</b>	0,8030 (S)	0,4606 (S)	0,3765 (S)	0,7193 (S)
<b>L3 B</b>	0,0305 (N)	0,3134 (S)	0,5773 (S)	0,0863 (N)
<b>L3//L4<sup>(*)</sup> A</b>	0,8488 (S)	0,3961 (S)	0,7802 (S)	0,0044 (N)
<b>L4 A</b>	0,0679 (S)	0,0044(N)	0,1520 (S)	0,8520 (S)

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo. S: hipótese atendida em nível de significância de 5%; N: hipótese não atendida em nível de significância de 5%.

A hipótese de aleatoriedade dos dados foi confirmada na maioria (80%) das séries de dados analisados. Porém, as hipóteses de estacionariedade e homogeneidade foram verificadas em menor frequência (54 e 68% das séries, respectivamente). A adequação das séries de dados à condição de independência foi a mais baixa (26% das séries). Além disso, 54% das séries apresentadas como independentes possuem menos de 30 dados, ou seja, são amostras pequenas que pouco representam a população da variável em questão.

A baixa adequação à hipótese de independência poderia ser explicada pela frequência de amostragem praticada, a qual variou entre semanal e quinzenal. Oliveira (2006) concluiu que um maior percentual de independência entre os dados dos constituintes afluentes e efluentes de ETEs ocorreu quando a faixa de frequência de amostragens praticada pelas ETEs em estudo era acima de quinzenal. Quando os dados eram coletados em menores intervalos entre as amostragens, era verificado comportamento inverso, predominando o percentual de dependência dos dados. Por outro lado, Berthouex e Hunter (1975), por meio de um programa de monitoramento de pequenas estações de tratamento, concluíram que dados amostrais com intervalos de frequência de 3 a 4 dias já podem ser considerados independentes.

A Tabela 4.20 apresenta uma síntese dos resultados obtidos nos diversos testes para a verificação das hipóteses de independência, homogeneidade, aleatoriedade e estacionariedade considerando o número e o percentual de séries analisadas. Em seguida, a Tabela 4.21 traz o número e o percentual de séries que atenderam simultaneamente a todas as hipóteses fundamentais da análise de frequência.

TABELA 4.20 - Síntese dos resultados dos testes de aleatoriedade, estacionariedade, homogeneidade e independência.

Variável	Total de séries analisadas	Dados aleatórios		Dados estacionários		Dados homogêneos		Dados independentes	
		Nº de séries	%	Nº de séries	%	Nº de séries	%	Nº de séries	%
<b>DBO</b>	12	12	100	7	58,3	7	58,3	3	25,0
<b>DQO</b>	13	12	92,3	6	46,2	10	76,9	3	23,1
<b>SST</b>	11	9	81,8	9	81,8	9	81,8	5	45,5
<b>NH<sub>3</sub>.</b>	11	8	72,7	6	54,5	5	45,5	0	0,0
<b>P</b>	11	7	63,6	7	63,6	5	45,5	1	9,1
<b>pH</b>	9	8	88,9	0	0,0	6	66,7	1	11,1
<b><i>E. coli</i></b>	11	8	72,7	8	72,7	9	81,8	6	54,5
<b>CT</b>	10	8	80,0	7	70,0	9	90,0	3	30,0
<b>Ovos de helmintos</b>	6	4	66,7	1	15,2	4	66,7	2	33,3

TABELA 4.21 – Número e percentual de séries de dados que atenderam simultaneamente as hipóteses de aleatoriedade, estacionariedade, homogeneidade e independência.

Séries de dados	DBO	DQO	SST	NH <sub>3</sub>	P	pH	<i>E coli</i>	CT	Ovos de helmintos
Nº séries	2	1	3	0	1	0	1	2	0
%	16,7	7,7	27,3	0,0	9,1	0,0	9,1	20,0	0,0

Observa-se na Tabela 4.20 que os dados de *E. coli*, SST e coliformes totais foram os que mais consistentemente atenderam às hipóteses de independência, homogeneidade e estacionariedade, apresentando ainda percentuais de atendimento relativamente elevados para os testes de aleatoriedade. A hipótese de aleatoriedade foi atendida de forma mais consistente pelos dados de DBO e DQO. Por sua vez, os dados de SST e *E. coli* foram os que mais bem confirmaram a hipótese de estacionariedade. No outro extremo (baixo atendimento das hipóteses), destacam-se os dados de fósforo e nitrogênio como as variáveis que apresentaram dados menos homogêneos e independentes. Por sua vez, os ovos de helmintos foi a variável que apresentou dados menos aleatórios e estacionários, no primeiro caso juntamente com os dados de fósforo, no segundo com o pH.

Na Tabela 4.21, nota-se que o número de séries que atenderam simultaneamente às quatro hipóteses (independência, aleatoriedade, homogeneidade e estacionariedade) foi bastante reduzido.

No já citado trabalho de Oliveira (2006), a autora concluiu que os dados provenientes de 208 estações de tratamento de esgoto, em geral, não atenderam às hipóteses fundamentais da análise de frequência e que, quando o monitoramento dos sistemas era realizado em elevada frequência, disponibilizando grande número de dados, as hipóteses tendiam a apresentar fração mais acentuada de não atendimento. Essa conclusão pode bem se aplicar às séries de dados aqui estudadas, as quais possuem dados de monitoramento por períodos prolongados de tempo e em geral semanal.

No entanto, assim como discutido por Oliveira (2006), o baixo atendimento às hipóteses fundamentais da análise de frequência não invalida a aplicação de outros testes estatísticos, mas requer maior cautela na interpretação dos respectivos resultados.

#### 4.2.4 - Testes de aderência e gráficos de probabilidade

Os testes estatísticos de aderência fornecem um valor cuja probabilidade associada de ocorrência sob a hipótese nula (aderência à distribuição) é maior ou igual a alguma pequena probabilidade, denotada por  $\alpha$ . Desta forma, o procedimento para definir se a distribuição possui boa aderência é a aceitação de  $H_0$  (hipótese nula) em detrimento a  $H_1$  (hipótese alternativa). Os Quadros 4.3 a 4.11 trazem os resultados dos testes de aderência e normalidade das séries de dados às distribuições candidatas.

Quadro 4.3 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de DBO.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
<b>EB</b>		N; LN; W	N; LN; W	N; LN; W	N
<b>RA</b>		N; LN; W	LN; W	LN	-
<b>BF</b>		-	-	-	-
<b>L1</b>	<b>A</b>	N; LN; W; E	N; LN; W	LN; W	LN
	<b>B</b>	N; LN; W	N; LN; W	LN; W	LN
<b>L2</b>	<b>A</b>	N; U; W	N; U; W	N; LN; U	N; LN
	<b>B</b>	N; LN; W	N; LN; W	N; LN; W	-
<b>L3</b>	<b>A</b>	N; LN; W; E	LN; W	LN; W	LN
	<b>B</b>	N; W	N; W	N; W	N
<b>L3//L4<sup>(*)</sup></b>	<b>A</b>	N; LN; W; U	N; LN; W; U	N; LN; W	N; LN
	<b>B</b>	N; LN; W; U	N; LN; W	N; LN; W	N; LN
<b>L4</b>	<b>A</b>	N; LN; W	N; LN; W	N; LN; W	N; LN

EB: esgoto bruto; RA: Reator anaeróbico; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

(\*) período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.4 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de DQO.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
EB		N; LN; W	N; LN; W	N; LN; W	N
RA		N; LN; W	LN; W	LN; W	-
BF		N; LN; W	N; LN; W	N; LN; W	N
L1	A	N; LN; W	N; LN; W	LN; W	-
	B	N; LN; W; U	N; LN; W	N; LN; W	N; LN
L2	A	N; LN; W	N; LN	N; LN	N; LN
	B	N; W	N; W	N; W	N
	C	N; LN; W	N; W	N; W	N; LN
L3	A	N; E; U; W	N	N	N; LN
	B	LN; E; W	N; LN; W; E	LN; W	LN
	C	LN; W	LN	LN	-
L3//L4 <sup>(*)</sup>	A	N; LN; W	N; LN; W	LN; W	LN
L4	A	N; LN; W; U	N; LN; W; U	N; LN; W	N

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.5 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de SST.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
RA		N; LN; W	N; LN; W	LN; W	LN
BF		-	LN; W	LN; W	LN
L1	A	N; LN; W; U	N; LN; W	LN; W	N
	B	N; LN; W	N; LN; W	LN; W	-
L2	A	N; LN; W	N; LN; W	N; LN; W	N; LN
	B	N; LN; W	N; LN; W	N; LN; W	N
L3	A	N; LN; E; U; W	N; LN; W; U	N; LN; W	-
	B	N; U; W	N; W	N; W	N
	C	N; W; U	N; W	N; W	N
L3//L4 <sup>(*)</sup>	A	N; LN; W; U	N; LN; W; U	N; LN; W	N; LN
L4	A	N; LN; W; U	N; LN; W	N; LN; W	N

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.6 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de pH.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
L1	A	LN	-	LN	-
L2	A	N; LN; W	N; LN; W	N; LN; W	LN
	B	N; LN; W	N; LN; W	N; LN; W	N; LN
L3	A	N; LN; W	N; LN; W	N; LN; W	N; LN
	B	N; LN; W	N; LN; W	N; LN	N; LN
L3//L4 <sup>(*)</sup>	A	N; LN; W	N; LN; W	N; LN; W	N; LN
	B	N; LN; W; U	N; LN; W; U	N; LN; U	N; LN
L4	A	LN; E	LN; E	LN; E	-
	B	N	N	N	N; LN

EB: esgoto bruto; RA: Reator anaeróbico; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.7 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de nitrogênio amoniacal.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
EB		N; LN; W	N; LN; W	N; LN; W	N
RA		W	N; W	W	-
BF		N; W	N; W	N; W	-
L1	A	N; LN; W; U	N; LN; W	N; LN; W	N; LN
	B	-	-	-	-
L2	A	N; LN; W	N; LN; W	N; LN; W	N; LN
	B	-	-	-	-
L3	A	N; LN; E; U	N; LN	N	N
	B	N; W; U	N; U	U	-
L3//L4 <sup>(*)</sup>	A	N; LN; W	N; LN; W	N; LN; W	N
L4	A	W	W	-	N

EB: esgoto bruto; RA: Reator anaeróbico; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.8 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de fósforo.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
EB		N; LN; W	N; LN; W	N; LN; W	N
RA		N; W	-	N; W	N
BF		N; W	N; W	N; W	N
L1	A	N; W; U	N; W	N; W	N
	B	N; LN; W	N; LN; W	N; LN; W	N
L2	A	N; LN; W; E	N; LN; W; E	N; LN; W	LN
	B	LN; W	N; LN; W	N; LN; W	N
L3	A	N; LN; W; U	N; LN; W	N; LN; W	N; LN
	B	N; W	N; W	N; W	N
L3//L4 <sup>(*)</sup>	A	N; U; W	N; U; W	N; U	N
L4	A	N; U; W	N; U	N	N; LN

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.9 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de *E. coli*.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
EB		LN	LN	-	-
RA		N; LN; W	N; LN; W	LN; W	-
BF		LN; W; E	LN; W; E	N; LN; W; E	-
L1	A	E	LN; E	E	-
L2	A	LN; E	LN; E	LN; E	-
	B	E	E	E	LN
L3	A	-	-	-	LN
	B	-	-	-	-
L3//L4 <sup>(*)</sup>	A	N; LN; E; U	N; LN; E	N; LN; E	N; LN
	B	N; E; U	E	U	N; LN
L4	A	N	-	-	LN

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

<sup>(\*)</sup>período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.10 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de concentração de coliformes totais.

Amostra		$\chi^2$	Kolmogorov Smirnov	Anderson Darling	Filliben
EB		-	-	E	-
RA		LN; W	LN; W	LN; W	-
BF		W; E	N; LN; W; E	LN	-
L1	A	-	LN	LN	
L2	A	LN	LN	LN	LN
	B	-	-	-	-
L3	A	-	-	-	-
	B	LN; W; E	LN; E	LN	-
L3//L4 <sup>(*)</sup>	A	LN; N; U	LN; W; E	-	-
L4	A	N; LN; E; U	N; E	-	-

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme.

(\*) período em que as lagoas 3 e 4 foram operadas em paralelo.

Quadro 4.11 - Resultados dos testes de aderência às distribuições candidatas ao nível de 5% de significância para os dados de ovos de helmintos.

Amostra	$\chi^2$
EB	-
RA	-
BF	-
L1 A	-
L1 B	P; G; BN
L2 A	P; G; BN

EB: esgoto bruto; RA: Reator anaeróbio; BF: Biofiltro; Li: lagoas de polimento; A e B: grupos de dados com resultados estatisticamente equivalentes. P: Poisson; G: Geométrica; BN: Binomial negativa.

Quadro 12 - Resultados dos testes de normalidade das séries de dados ao nível de 5% de significância

<b>Amostra</b>	<b>DBO</b>	<b>DQO</b>	<b>SST</b>	<b>pH</b>	<b>NH<sub>3</sub></b>	<b>P</b>	<b><i>E. coli</i></b>	<b>CT</b>	
<b>EB</b>	Normal	Normal	----	----	Normal	Normal	Não normal	Não normal	
<b>RA</b>	Não normal	Não normal	Não normal	----	Não normal	Normal	Não normal	Não normal	
<b>BF</b>	Não normal	Normal	Não normal	----	Não normal	Normal	Não normal	Não normal	
<b>L1</b>	<b>A</b>	Não normal	Não normal	Não normal	Não normal	Normal	Normal	Não normal	Não normal
	<b>B</b>	Não normal	Normal	Não normal	----	Não normal	Normal	----	----
<b>L2</b>	<b>A</b>	Normal	Normal	Normal	Não normal	Normal	Não normal	Não normal	Não normal
	<b>B</b>	Normal	Normal	Normal	Não normal	Não normal	Normal	Não normal	Não normal
	<b>C</b>	----	Normal	----	----	----	----	----	----
<b>L3</b>	<b>A</b>	Não normal	Normal	Normal	Normal	Não normal	Normal	Não normal	Não normal
	<b>B</b>	Normal	Não normal	Normal	Normal	Não normal	Normal	Não normal	Não normal
	<b>C</b>	----	Não normal	Normal	----	----	----	----	----
<b>L3//L4*</b>	<b>A</b>	Normal	Não normal	Normal	Normal	Normal	Não normal	Normal	Não normal
	<b>B</b>	Normal	----	----	Normal	----	----	Normal	----
<b>L4</b>	<b>A</b>	Normal	Normal	Normal	Não normal	Não normal	Não normal	Não normal	Não normal
	<b>B</b>	----	----	----	Normal	----	----	----	----

EB: esgoto bruto; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo.

As Figuras 4.2 a 4.9 representam os gráficos de probabilidades para as séries com 30 ou mais dados para a verificação empírica-visual da aderência às distribuições candidatas que apresentaram aderência por algum dos testes aplicados. O eixo das abscissas refere-se à distribuição cumulativa teórica e o eixo das ordenadas à distribuição cumulativa empírica.

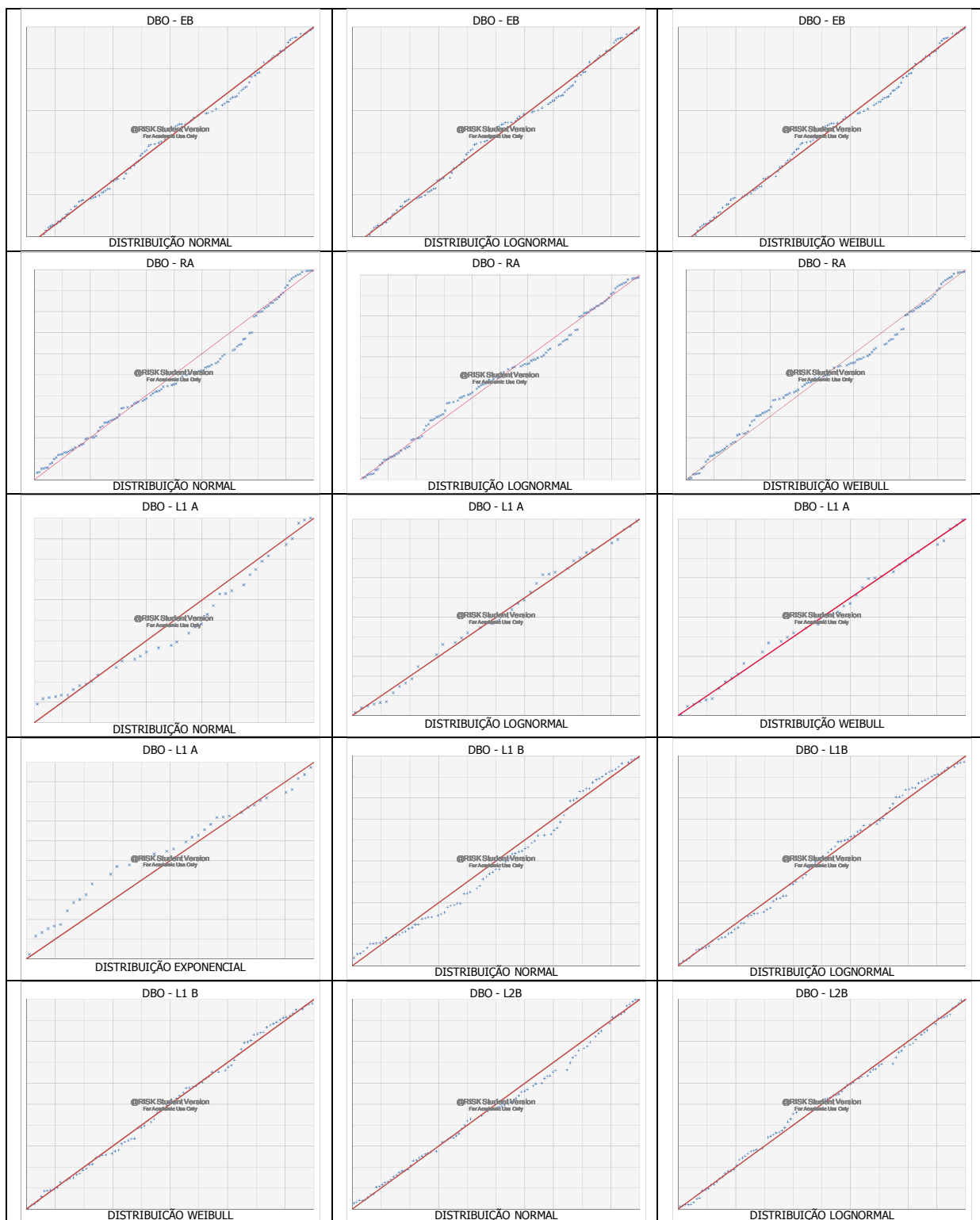


Figura 4.2 - Gráficos para verificação da aderência às distribuições candidatas, dados de DBO, dados de esgoto bruto, reator anaeróbio, lagoa 1, lagoa 2 e lagoa 3 (continua).

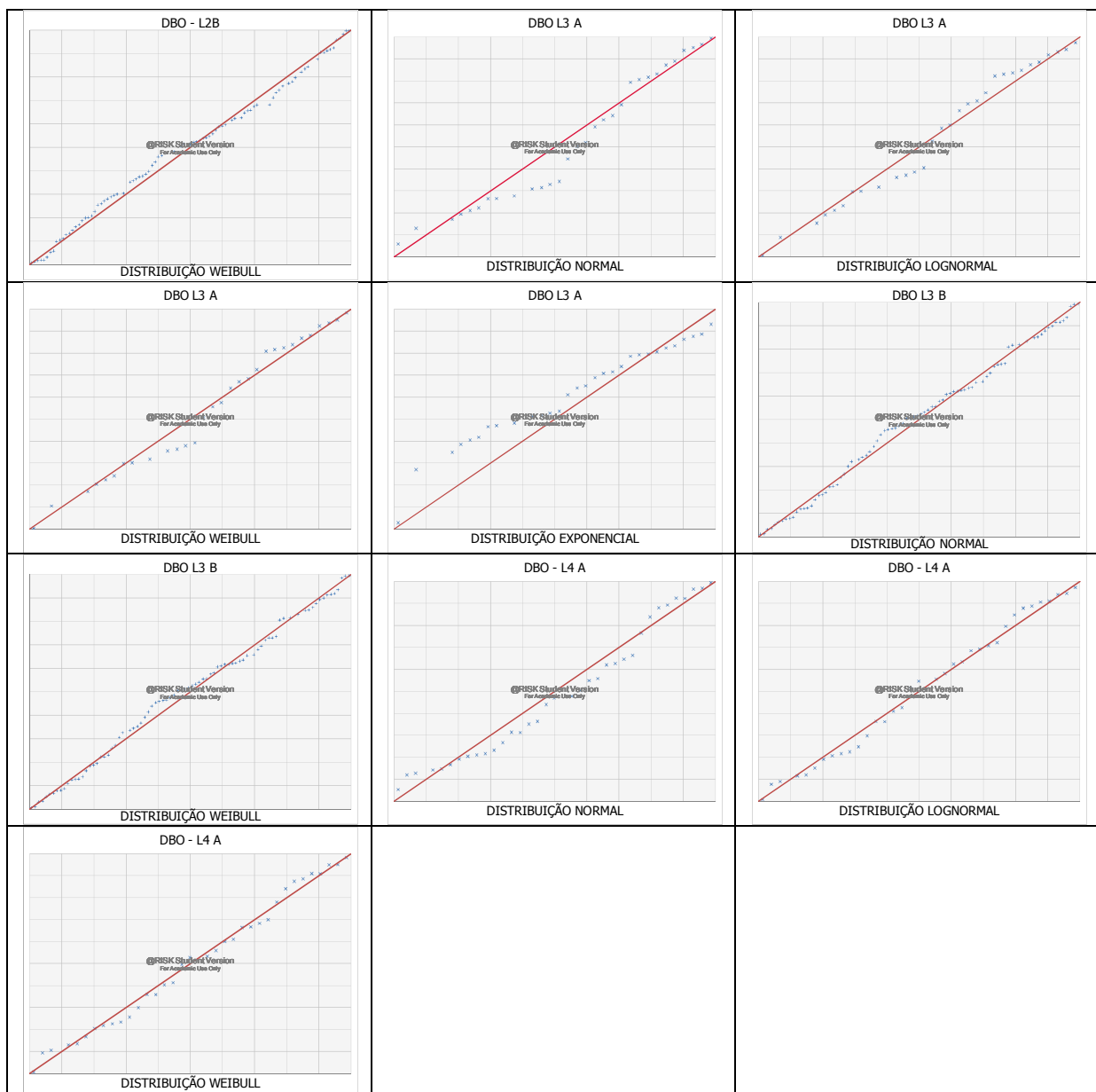


Figura 4.2 - Gráficos para verificação da aderência às distribuições candidatas, dados de DBO, dados de esgoto bruto, reator anaeróbio, lagoa 1, lagoa 2 e lagoa 3 (continuação).

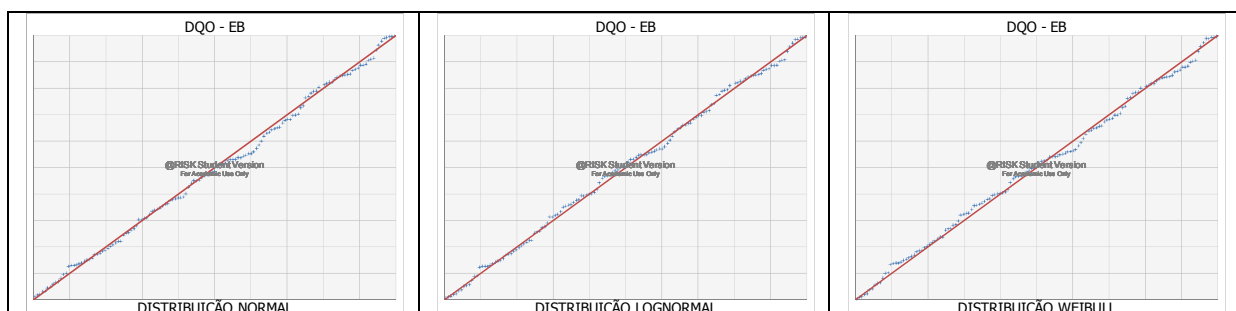


Figura 4.3 - Gráficos para verificação da aderência às distribuições candidatas, dados de DQO, dados de esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2 e lagoa 3 (continua).

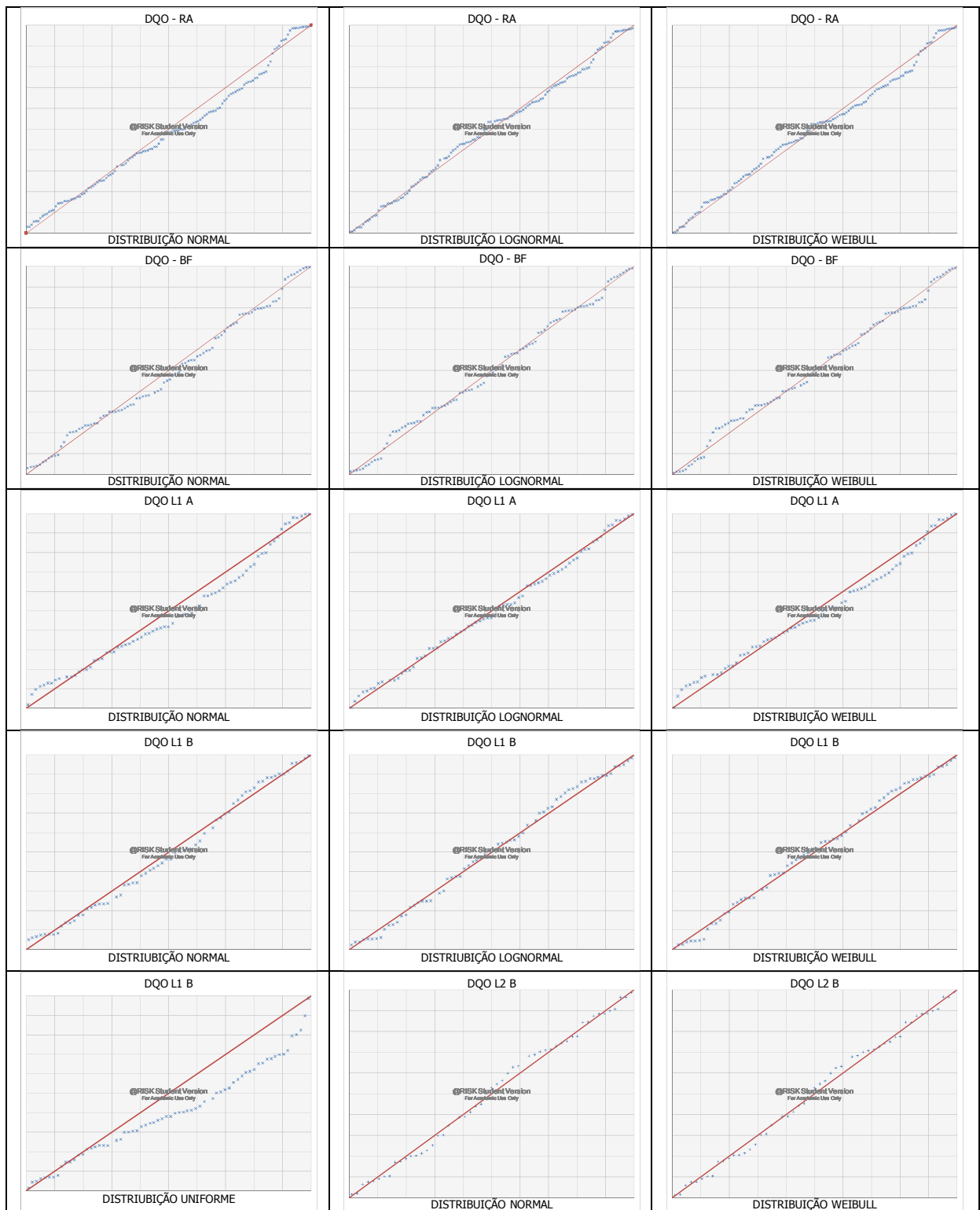


Figura 4.3 - Gráficos para verificação da aderência às distribuições candidatas, dados de DQO, dados de esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2 e lagoa 3 (continua).

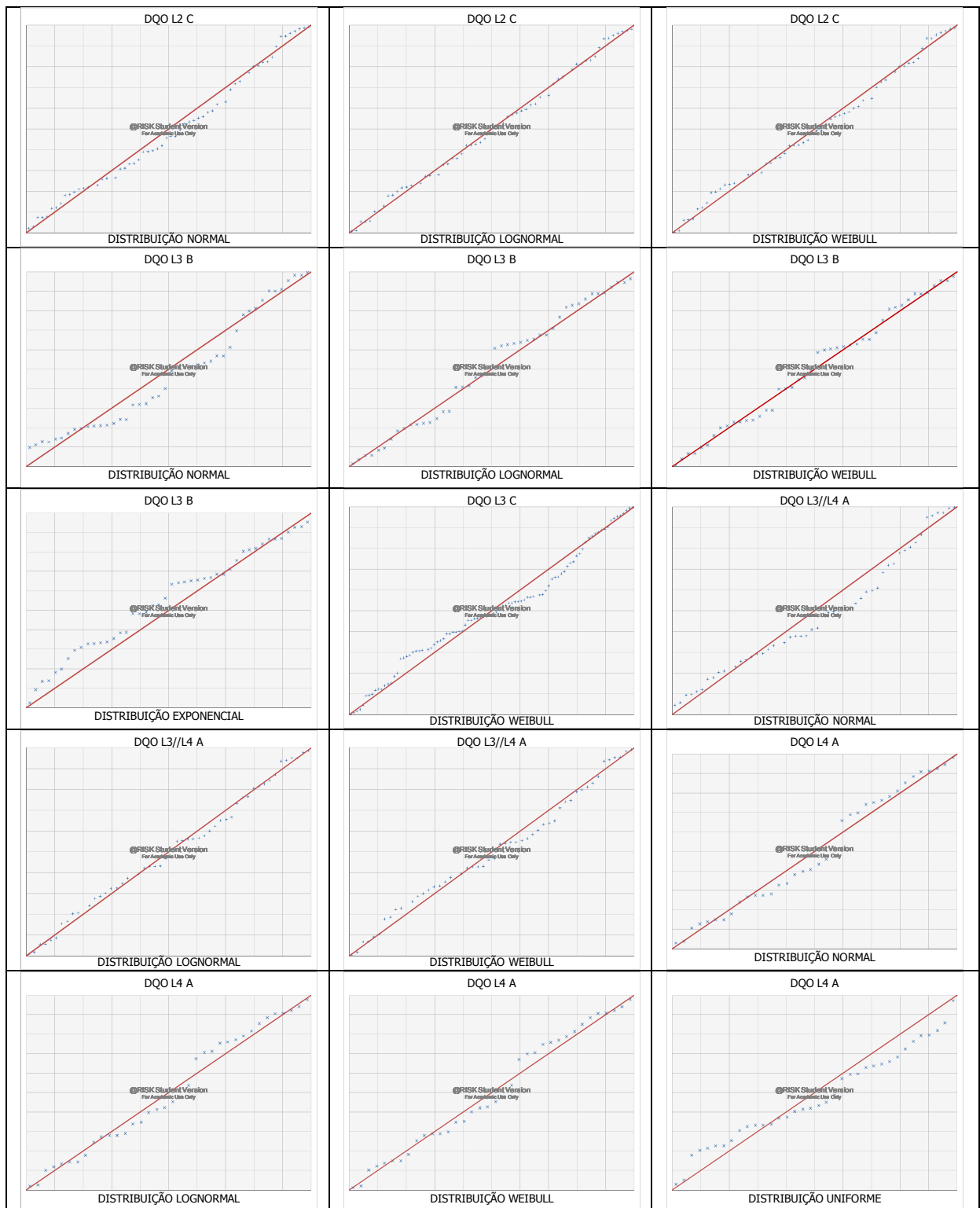


Figura 4.3 - Gráficos para verificação da aderência às distribuições candidatas, dados de DQO, dados de esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2 e lagoa 3 (continuação).

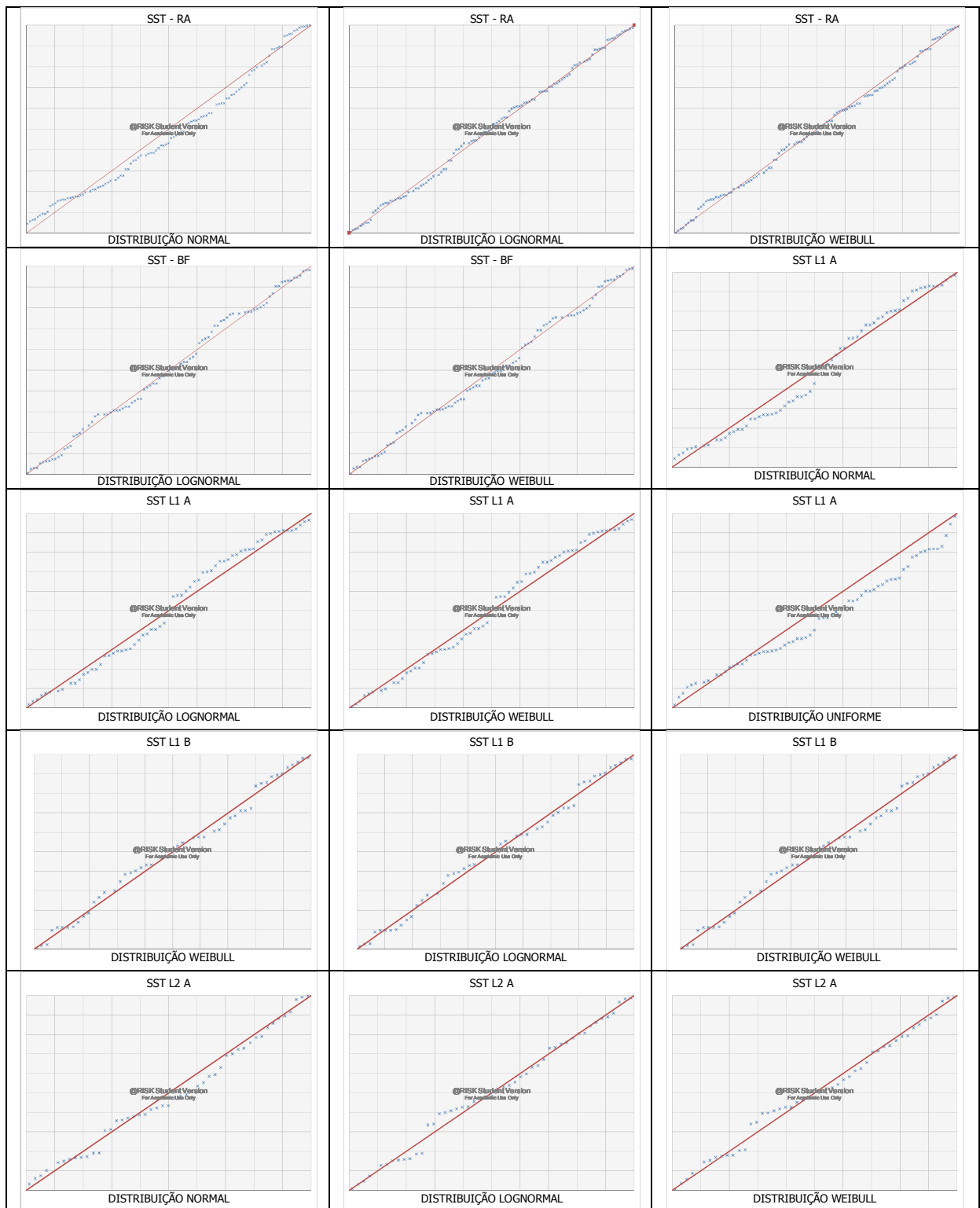


Figura 4.4 - Gráficos para verificação da aderência às distribuições candidatas, dados de SST, reator anaeróbio, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continua).

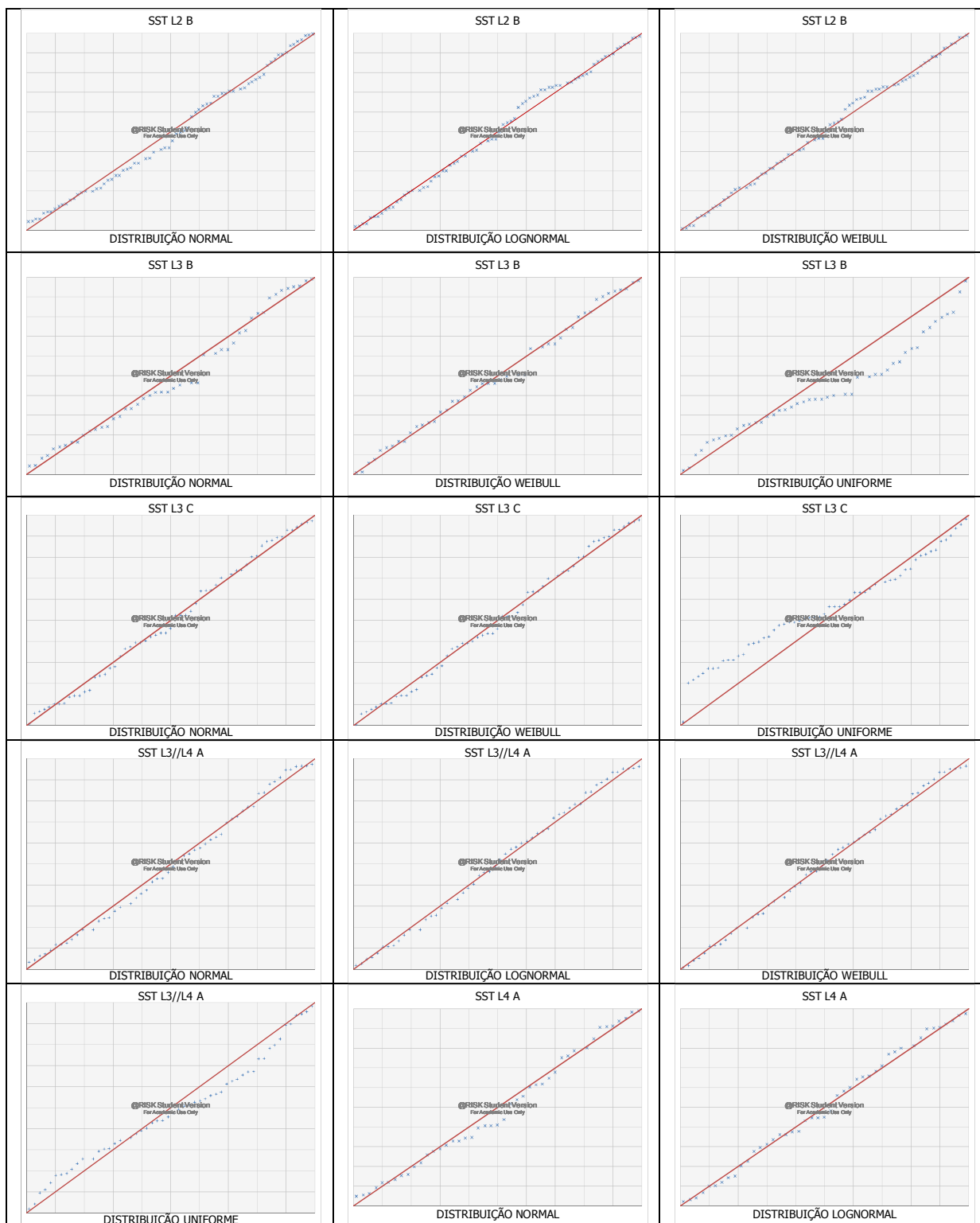


Figura 4.4 - Gráficos para verificação da aderência às distribuições candidatas, dados de SST, reator anaeróbio, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continuação).

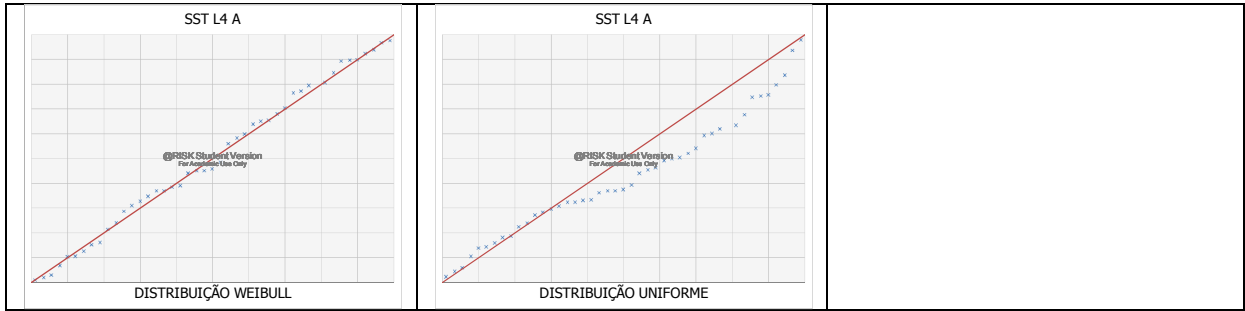


Figura 4.4 - Gráficos para verificação da aderência às distribuições candidatas, dados de SST, reator anaeróbio, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continuação).

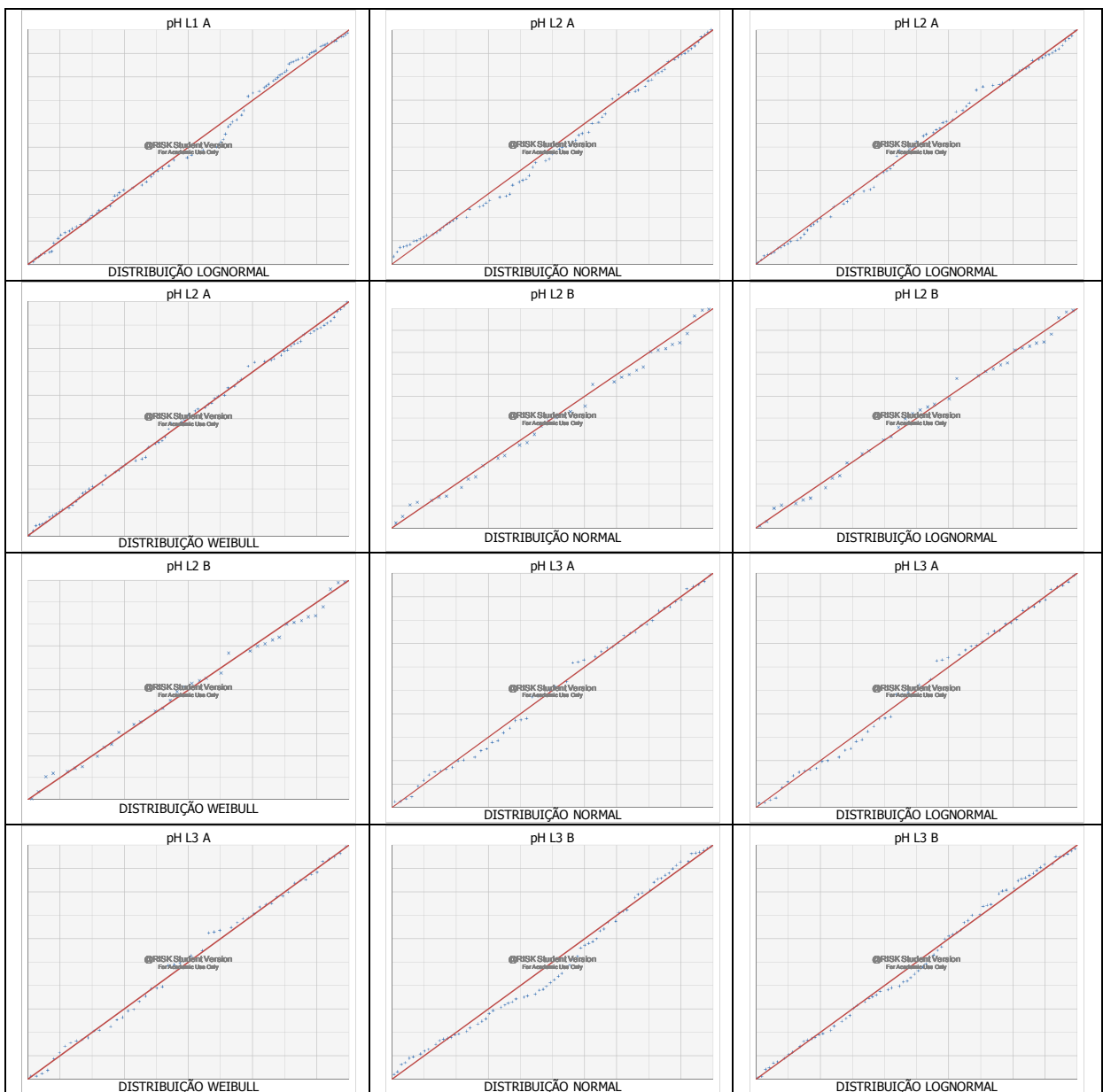


Figura 4.5 - Gráficos para verificação da aderência às distribuições candidatas, dados de pH, lagoa 1, lagoa 2 e lagoa 3 (continua).

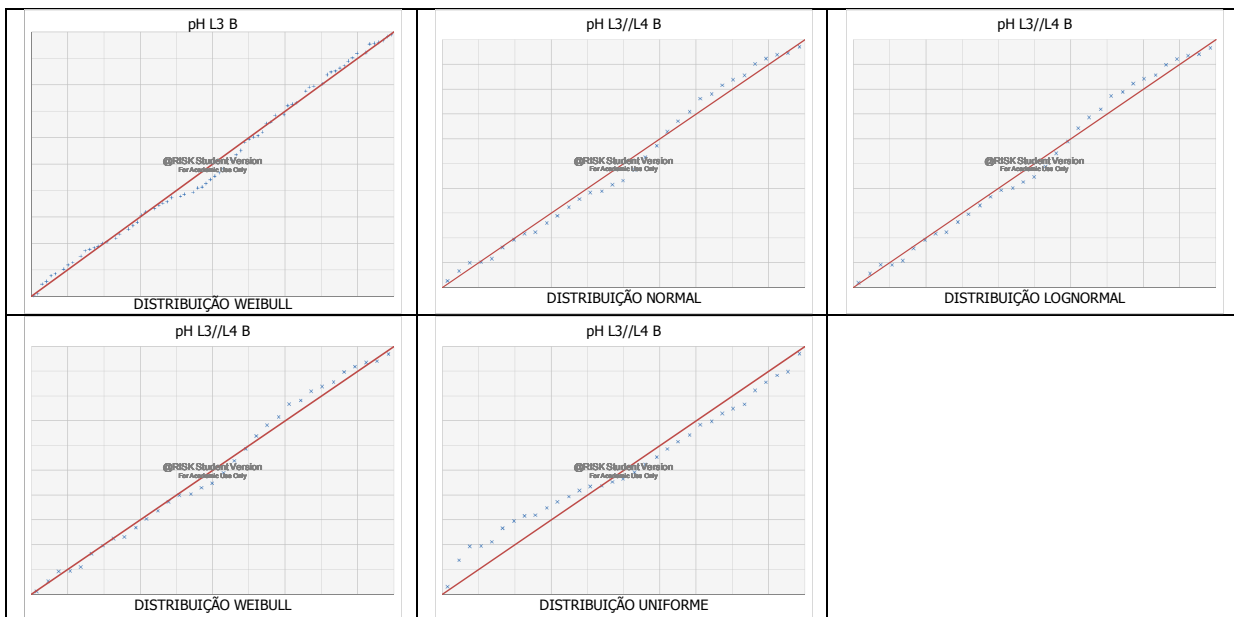


Figura 4.5 - Gráficos para verificação da aderência às distribuições candidatas, dados de pH, lagoa 1, lagoa 2 e lagoa 3 (continuação).

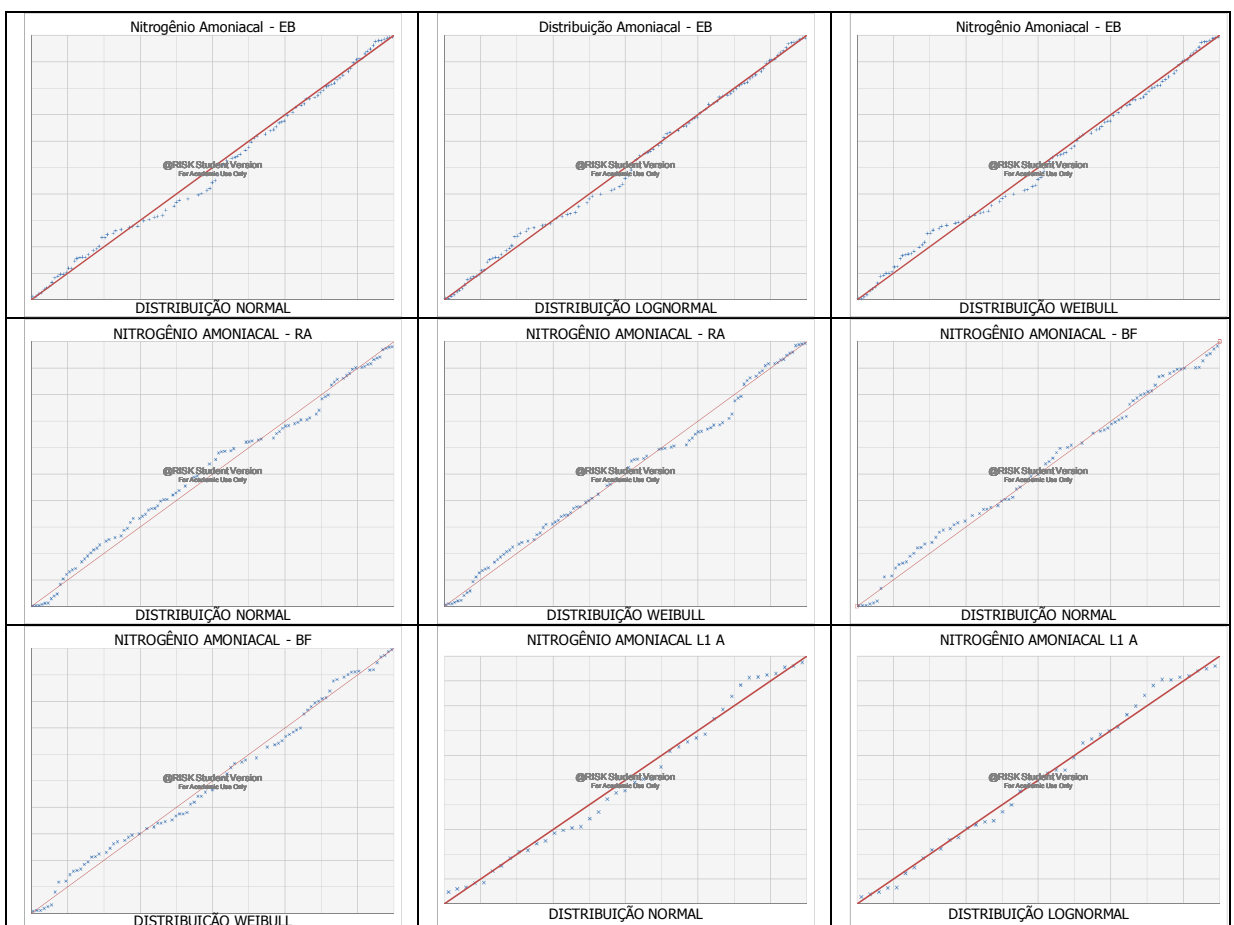


Figura 4.6 - Gráficos para verificação da aderência às distribuições candidatas, dados de nitrogênio amoniacal, esgoto bruto, reator anaeróbio, bio filtro, lagoa 3 e lagoas 3 e 4 operando em paralelo (continua).



Figura 4.6 - Gráficos para verificação da aderência às distribuições candidatas, dados de nitrogênio amoniacal, esgoto bruto, reator anaeróbio, bio filtro, lagoa 3 e lagoas 3 e 4 operando em paralelo (continua).

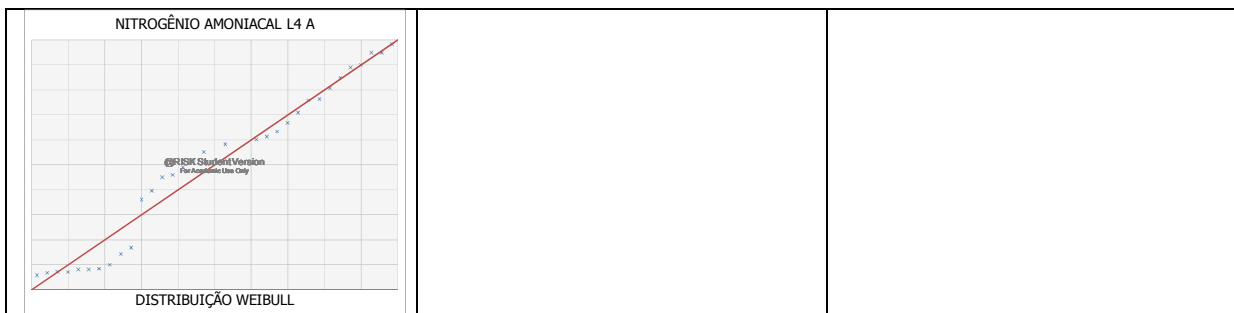


Figura 4.6 - Gráficos para verificação da aderência às distribuições candidatas, dados de nitrogênio amoniacal, esgoto bruto, reator anaeróbico, bio filtro, lagoa 3 e lagoas 3 e 4 operando em paralelo (continuação).

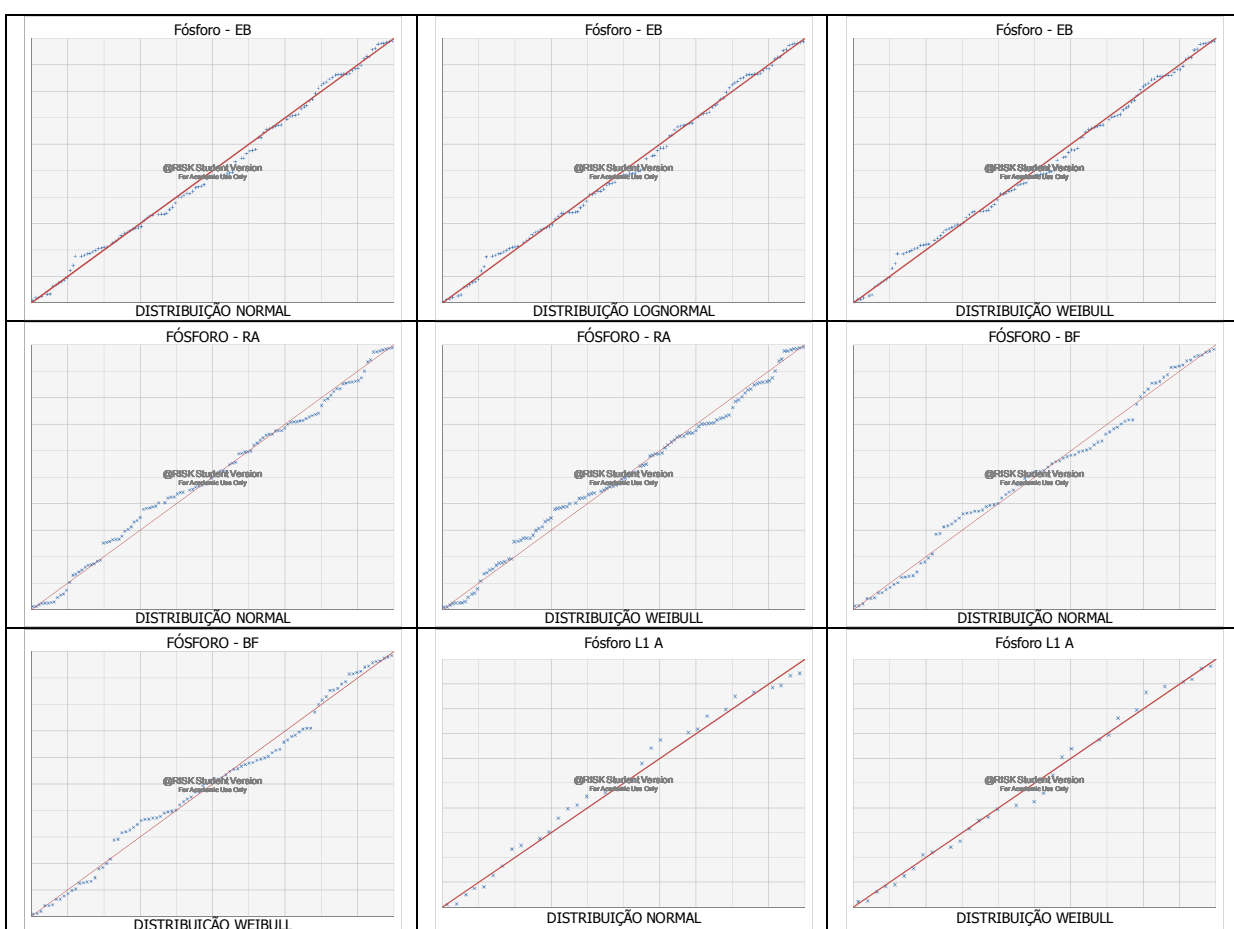


Figura 4.7 - Gráficos para verificação da aderência às distribuições candidatas, dados de fósforo, esgoto bruto, reator anaeróbico, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continua).

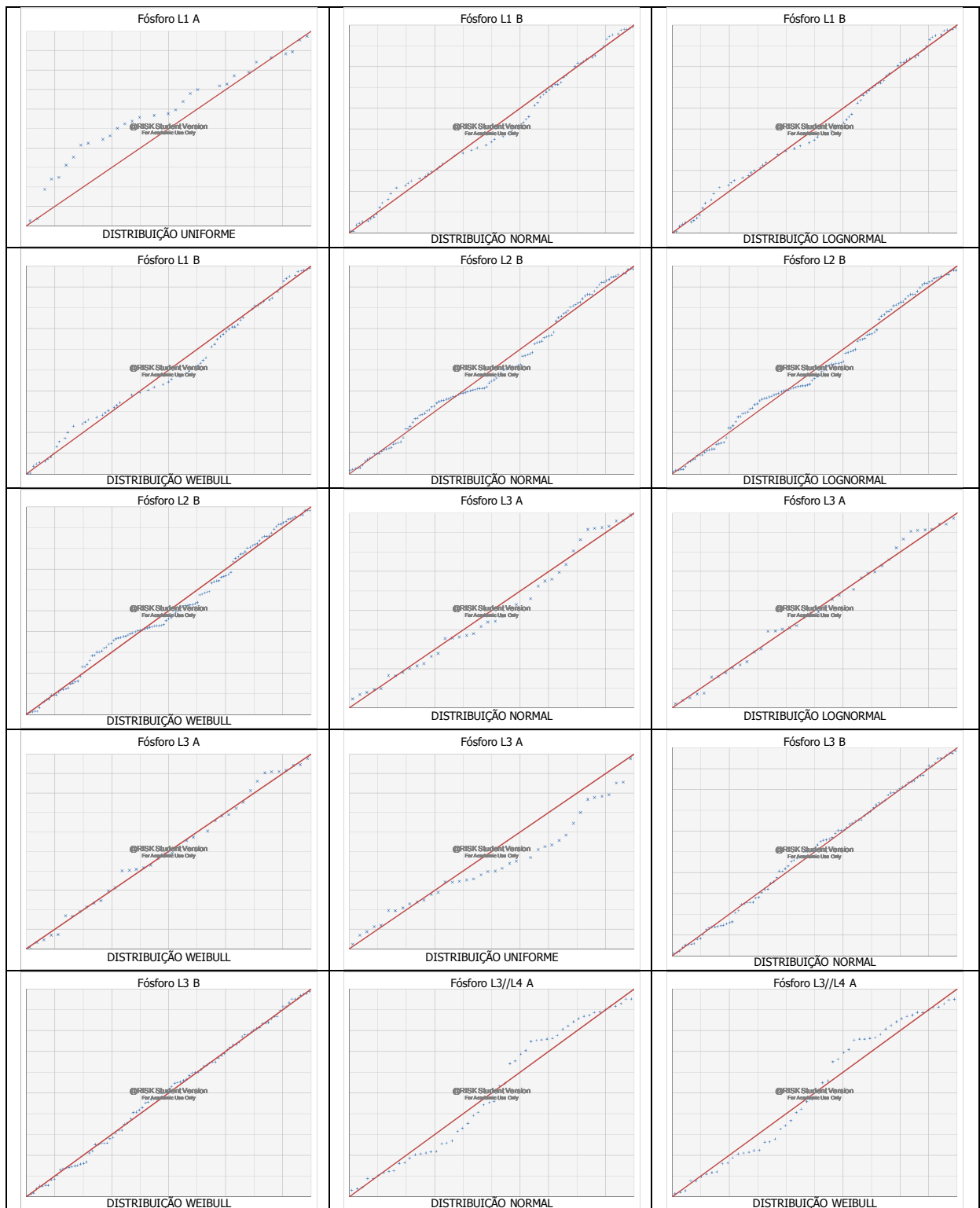


Figura 4.7 - Gráficos para verificação da aderência às distribuições candidatas, dados de fósforo, esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continua).

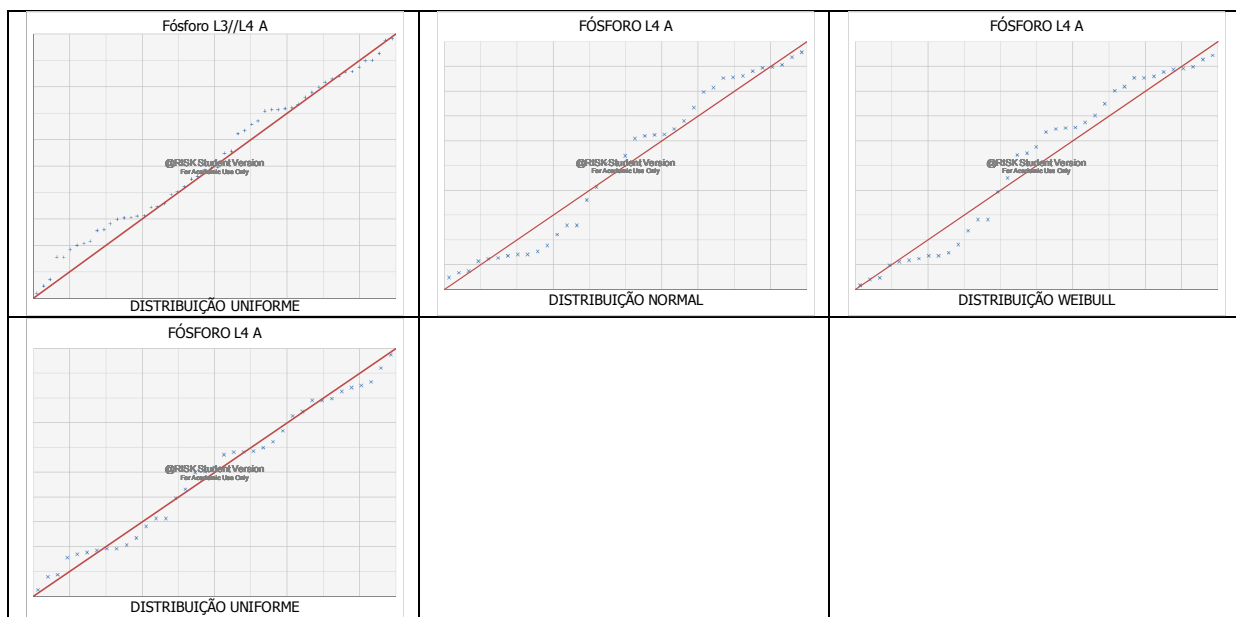


Figura 4.7 - Gráficos para verificação da aderência às distribuições candidatas, dados de fósforo, esgoto bruto, reator anaeróbio, bio filtro, lagoa 1, lagoa 2, lagoa 3 e lagoas 3 e 4 operando em paralelo (continuação).

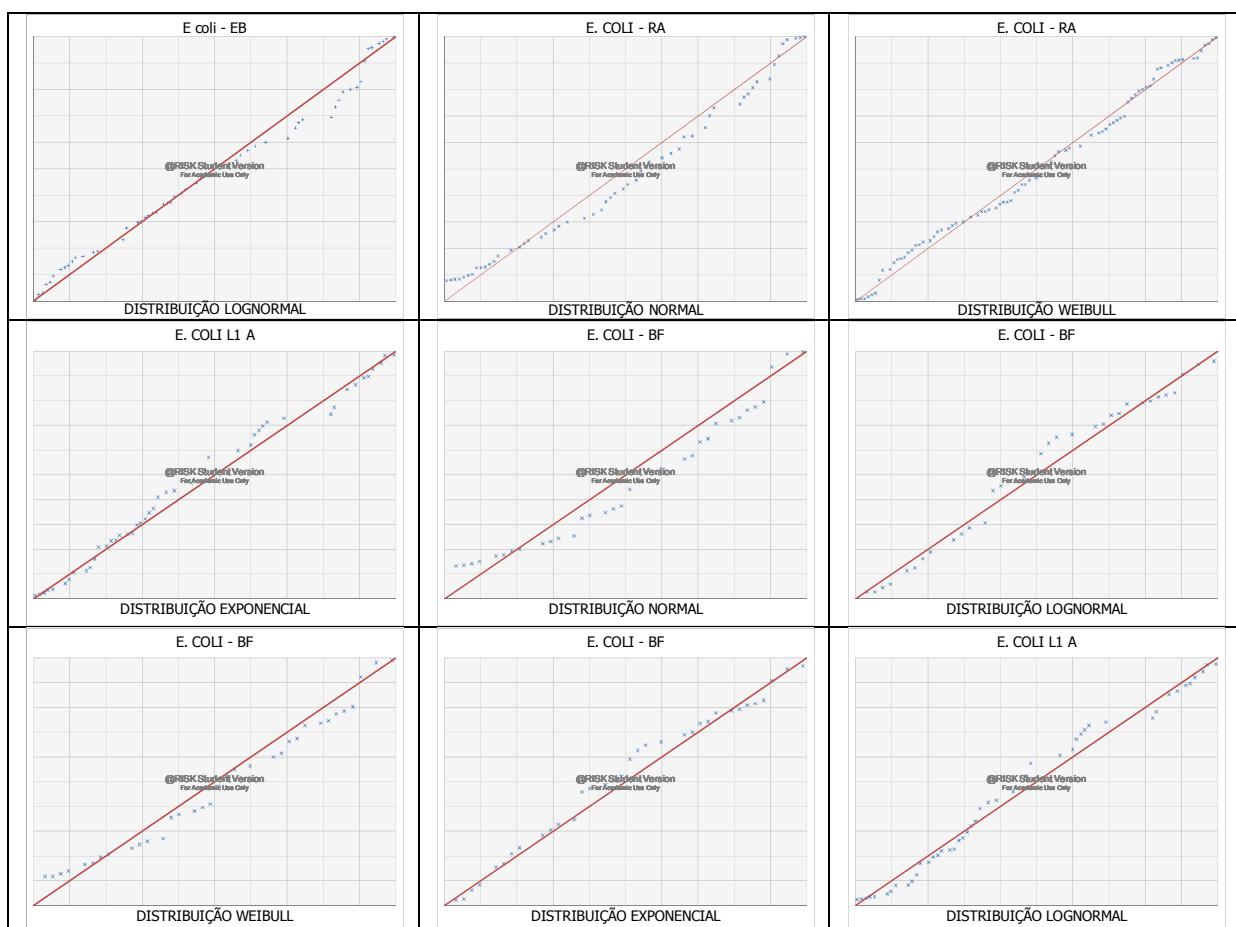


Figura 4.8 - Gráficos para verificação da aderência às distribuições candidatas, dados de *E. coli*, esgoto bruto, reator anaeróbio, lagoa 2 e lagoa 3 (continua).

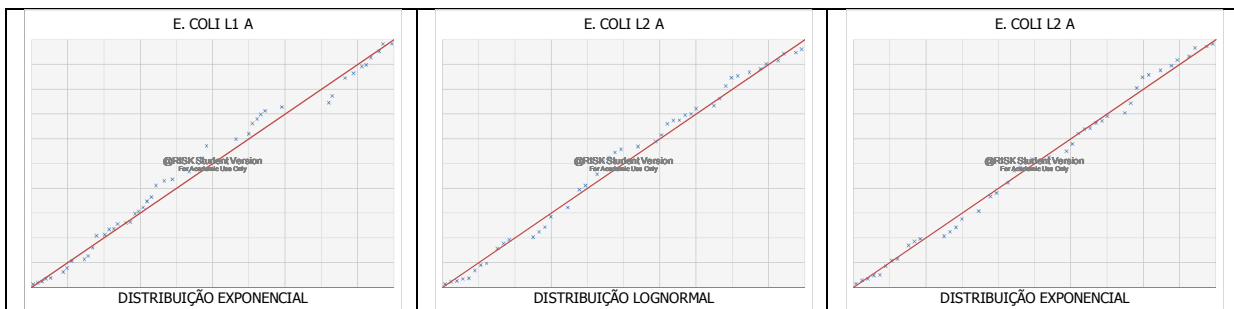


Figura 4.8 - Gráficos para verificação da aderência às distribuições candidatas, dados de *E. coli*, esgoto bruto, reator anaeróbio, lagoa 2 e lagoa 3 (continuação).

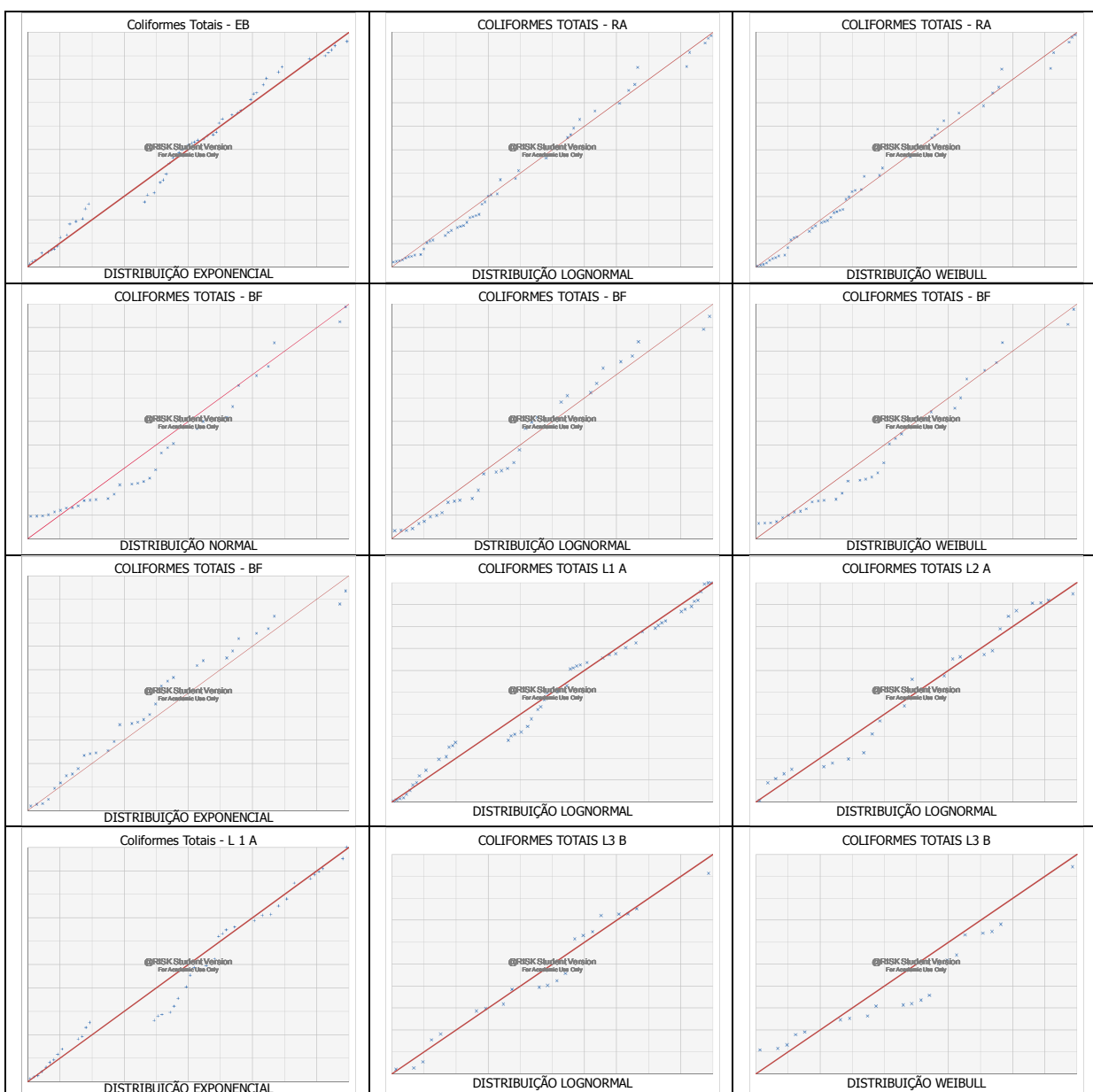


Figura 4.9 - Gráficos para verificação da aderência às distribuições candidatas, dados de coliformes totais, esgoto bruto, reator anaeróbio, biofiltro, lagoa 1 e lagoa 3 (continua).



Figura 4.9 - Gráficos para verificação da aderência às distribuições candidatas, dados de coliformes totais, esgoto bruto, reator anaeróbio, biofiltro, lagoa 1 e lagoa 3 (continuação).

Em geral, os testes de Kolmogorov-Smirnov (KS) e de Anderson Darling (AD) forneceram resultados de certa forma similares para todas as variáveis analisadas. Também de forma geral, o teste Qui-quadrado ( $\chi^2$ ) mostrou-se menos rigoroso (ou seja, ‘aceitando’ um número maior de distribuições) que os testes de KS e AD, principalmente em amostras pequenas.

Como destacado na revisão de literatura, o teste KS é mais poderoso (ou menos arbitrário) que o teste do Qui-quadrado, pois trata observações individuais separadamente e, portanto, não perde informação devido ao agrupamento, como acontece algumas vezes com o teste Qui-quadrado (SIEGEL e CASTELLAN, 2006). Assim como o teste KS, o teste de AD não envolve a divisão dos dados em intervalos de classe, mas enquanto o KS tem foco no centro da distribuição, o AD destaca diferenças entre as caudas das distribuições. Sendo assim, a seleção da distribuição para cada variável analisada teve como critério geral as coincidências de resultados entre todos os testes empregados, com destaque, entretanto, para os testes de KS e AD.

Em análise preliminar e geral, observa-se que as distribuições normal e lognormal foram as mais recorrentes ou que simultaneamente melhor descreveram a distribuição dos dados das diversas variáveis nas diversas amostras analisadas. A distribuição Weibull também apareceu com frequência relativamente elevada, o que se explica pela grande flexibilidade dessa distribuição de probabilidades. Por outro lado, a distribuição uniforme muito raramente foi ajustada aos dados analisados. Já a distribuição exponencial apareceu com relativa frequência apenas nas séries de dados bacteriológicos. As análises gráficas, salvo exceções, forneceram resultados coerentes em relação aos testes estatísticos.

#### 4.2.5 – Descrição da distribuição selecionada por meio de histogramas

Como acima referido, a seleção da distribuição para cada variável analisada teve como critério geral as coincidências de resultados entre todos os testes empregados, tendo em consideração as características de cada teste empregado e, de forma complementar, a análise gráfica. Com base nestes critérios, inicialmente foram selecionadas distribuições para cada variável em cada tipo de amostra (esgoto bruto, reator anaeróbio, biofiltro e efluentes das lagoas de polimento) e, na sequência, a distribuição que melhor representasse o conjunto de dados para cada variável. Os resultados estão resumidos no Quadro 4.13.

Tal procedimento teve como objetivo minimizar arbitrariedades e subjetividades na escolha das distribuições. Por outro lado, como destacado na revisão de literatura, há que se lembrar que os testes de aderência não se prestam rigorosamente à seleção de uma dentre várias distribuições, mas constituem apenas instrumentos da estatística que auxiliam a tomada de decisão. Ou seja, determinado grau de subjetividade ou mesmo de arbitrariedade pode se revelar inevitável. Porém, ainda no intuito de minimizar subjetividades ou arbitrariedades, por vezes optou-se por considerar resultados como inconclusivos.

É preciso ainda esclarecer que a seleção de uma única função como explicativa da distribuição dos dados de cada variável foi realizada tão somente com o objetivo de utilização nas simulações de Monte Carlo nos capítulos seguintes, não se pretendendo aqui fazer maiores inferências definitivas. Por exemplo, há que se destacar que muitas vezes a distinção entre as distribuições lognormal e Weibull mostrou-se difícil. Nesses casos, isso pode ser explicado pela característica comum a esses dois modelos de possuírem curvas flexíveis, ou seja, coeficientes de assimetria não fixos. Outro fator é a versatilidade de formas da distribuição Weibull, em virtude da variação de seus parâmetros  $\kappa$  e  $\lambda$ . Entretanto, por vezes a própria distinção entre as distribuições normal e lognormal também não se mostrou fácil.

Guardadas essas observações, os resultados sugerem que a função lognormal mais nitidamente descreveu a distribuição dos dados de *E. coli*, coliformes totais, pH, SST, DBO, DQO e nitrogênio amoniacal ( $\text{NH}_3$ ). Enquanto o fósforo (P) apresenta melhor adesão ao modelo normal de distribuição. Os resultados relativos aos ovos de helmintos foram inconclusivos, ou seja, não foi possível identificar uma distribuição que mais nitidamente se ajustasse ao banco de dados disponível.

Quadro 4.13 - Distribuições selecionadas para as séries de dados analisadas no esgoto bruto e nos efluentes das lagoas de polimento.

Amostra	DBO	DQO	SST	pH	NH <sub>3</sub>	P	<i>E. coli</i>	CT	Ovos de helmintos	
<b>EB</b>	N	N	-----	-----	N	N	LN	E	I	
<b>RA</b>	LN	LN	LN	-----	W	N	W	LN	I	
<b>BF</b>	I	LN	LN	-----	W	N	LN	LN	I	
<b>L1</b>	<b>A</b>	LN	LN	W	LN	LN	W	E	LN	I
	<b>B</b>	LN	LN	LN	-----	I	N	-----	-----	I
<b>L2</b>	<b>A</b>	N	N	LN	LN	LN	W	E	LN	I
	<b>B</b>	LN	N	W	LN	I	W	E	I	-----
	<b>C</b>	-----	LN	-----	-----	-----	-----	-----	-----	-----
<b>L3</b>	<b>A</b>	LN	N	N	LN / N	LN	LN	LN	I	-----
	<b>B</b>	N	LN	W	LN	U	N	I	LN	-----
	<b>C</b>	-----	LN	N	-----	-----	-----	-----	-----	-----
<b>L3//L4</b>	<b>A</b>	N	LN	LN	LN	LN	W	LN	LN	-----
	<b>B</b>	LN	-----	-----	LN / N	-----	-----	I	-----	-----
<b>L4</b>	<b>A</b>	LN	LN	W	LN	W	U	LN	I	-----
	<b>B</b>	-----	-----	-----	N	-----	-----	-----	-----	-----
<b>Distribuição adotada</b>	<b>LN</b>	<b>LN</b>	<b>LN</b>	<b>LN</b>	<b>LN</b>	<b>N</b>	<b>LN</b>	<b>LN</b>	<b>I</b>	

EB: esgoto bruto; Li: lagoas de polimento; A, B e C: grupos de dados com resultados estatisticamente equivalentes. N: normal; LN: lognormal; E: exponencial; W: Weibull; U: uniforme; I: inconclusivo; (\*) período em que as lagoas 3 e 4 foram operadas em paralelo.

Como já destacado na revisão de literatura, vários trabalhos relatam que a distribuição lognormal bem descreve a variação de dados de qualidade de afluentes e de efluentes de ETEs, incluindo, principalmente, dados de DBO, SST e coliformes (DEAN E FORSYTHE, 1976a, 1976b; NIKU *et al.*, 1979, 1981b, 1981c, citados por Oliveira, 2006; CHARLES *et al.*, 2005). Também como já referido, dados nacionais e mais especificamente sobre lagoas de estabilização são bem mais escassos. Oliveira (2006) analisou dados de 208 ETEs dos estados de Minas Gerais e São Paulo compreendendo vários processos de tratamento e concluiu que tanto no esgoto bruto como nos efluentes das ETEs a distribuição lognormal mostrou ser a mais adequada para representar o comportamento da maioria das variáveis analisadas (DBO, DQO, SST, Nitrogênio Total, Fósforo Total e coliformes termotolerantes), sendo o Fósforo Total afluente a única exceção. Quando os dados dos sistemas de lagoas foram analisados em

separado essa tendência foi de certa forma confirmada, muito embora os resultados tenham indicado também a possibilidade de outras distribuições, por exemplo, normal, ou ainda resultados inconclusivos.

A adesão predominante ao modelo lognormal, tanto no esgoto bruto, no reator anaeróbio, no biofiltro, quanto nos efluentes das lagoas, é confirmada neste trabalho. Por outro lado, deve-se registrar que a não normalidade dos dados mostrou-se menos evidente do que no trabalho de Oliveira (2006).

As Figuras 4.10 a 4.18 mostram os histogramas das séries de dados e as curvas ajustadas das distribuições escolhidas para cada variável e em cada conjunto de dados. A forma do histograma naturalmente mostrou-se influenciada pelo número escolhido de categorias. Sua precisão, quanto à aproximação à curva de distribuição de frequência, parece estar associada ao número de dados.

A maior parte das séries em que a escolha da distribuição candidata não foi inconclusiva apresentou distribuição assimetricamente positiva, ou seja, grande parte dos dados está agrupada à esquerda na curva de densidade probabilidade. Os dados das variáveis *E. coli*, coliformes totais e ovos de helmintos foram os que apresentaram assimetria positiva mais acentuada, o que auxilia na confirmação da adesão ao modelo lognormal, ao menos nos casos de *E. coli* e coliformes totais.

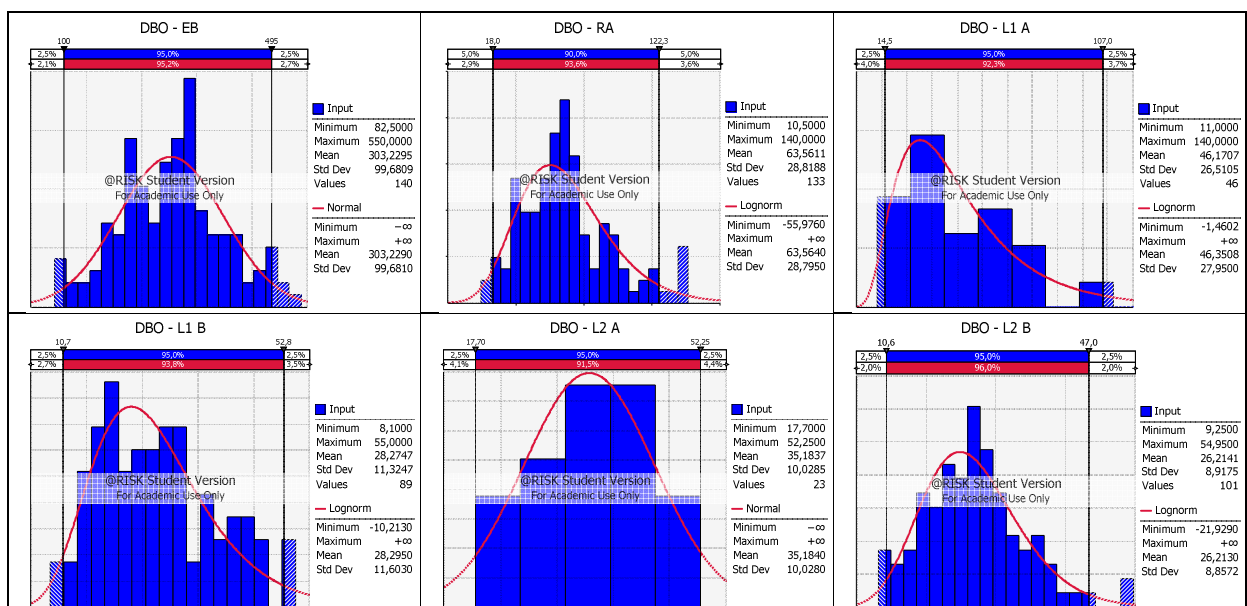


Figura 4.10 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DBO esgoto bruto, reator anaeróbio, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continua).

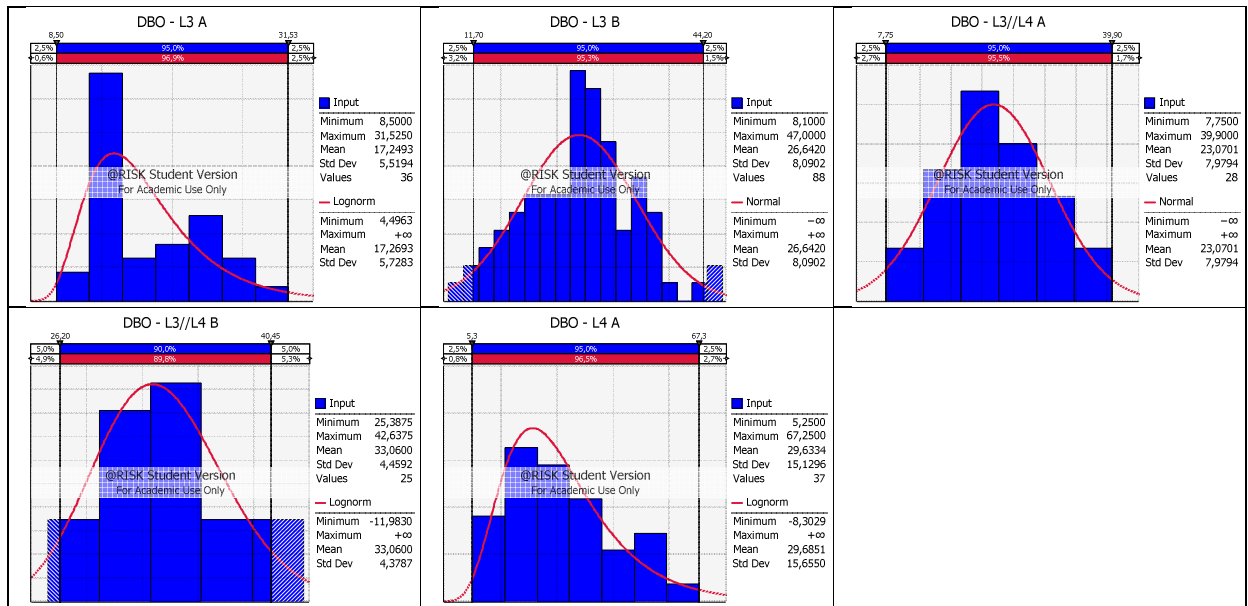


Figura 4.10 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DBO esgoto bruto, reator anaeróbico, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continuação).

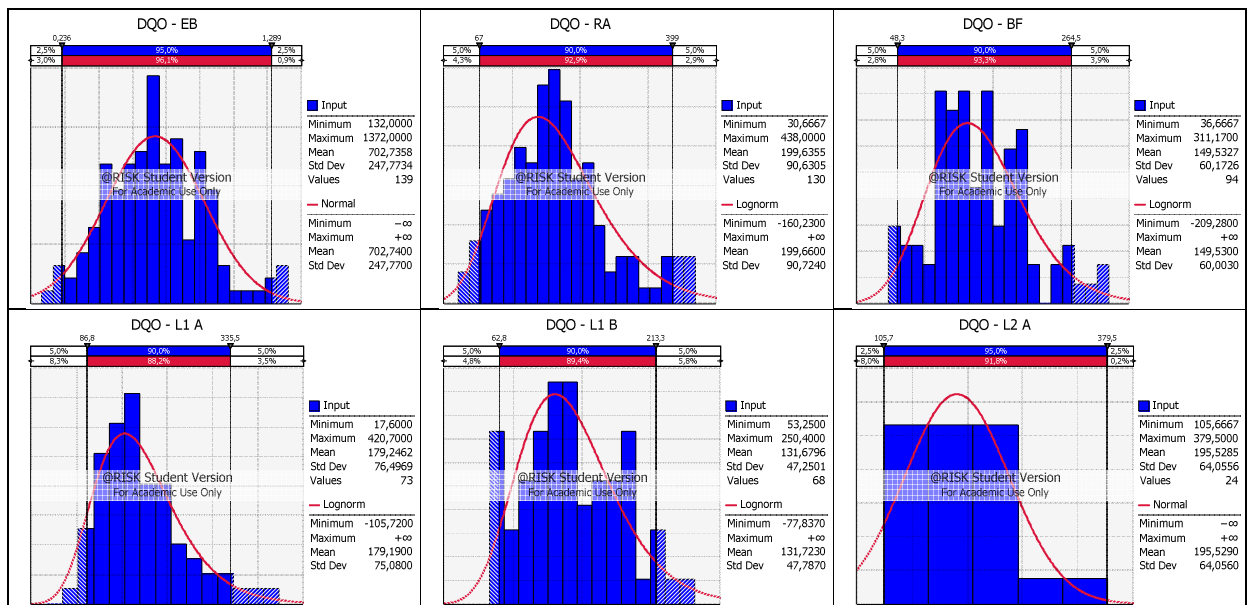


Figura 4.11 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DQO esgoto bruto, reator anaeróbico, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continua).

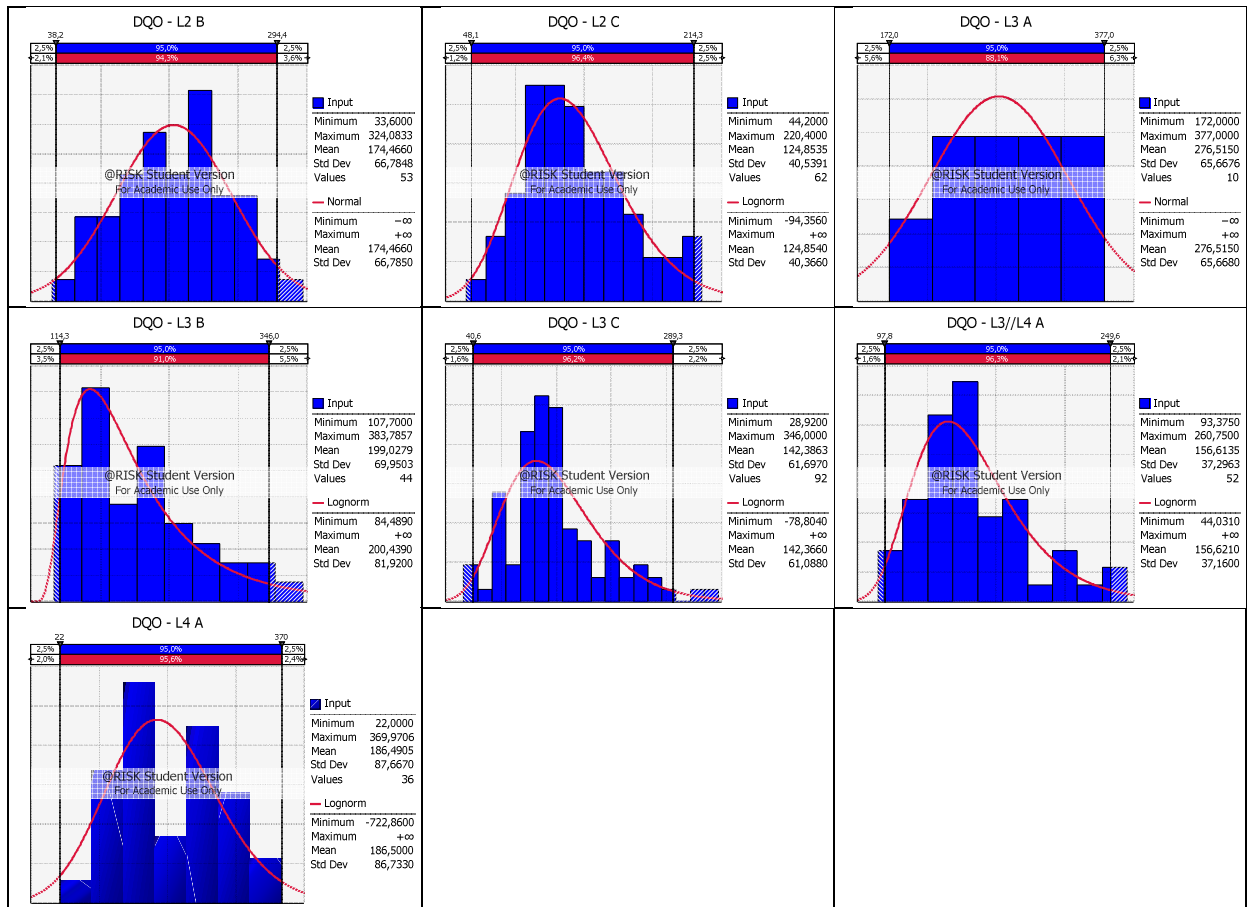


Figura 4.11 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de DQO esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continuação).

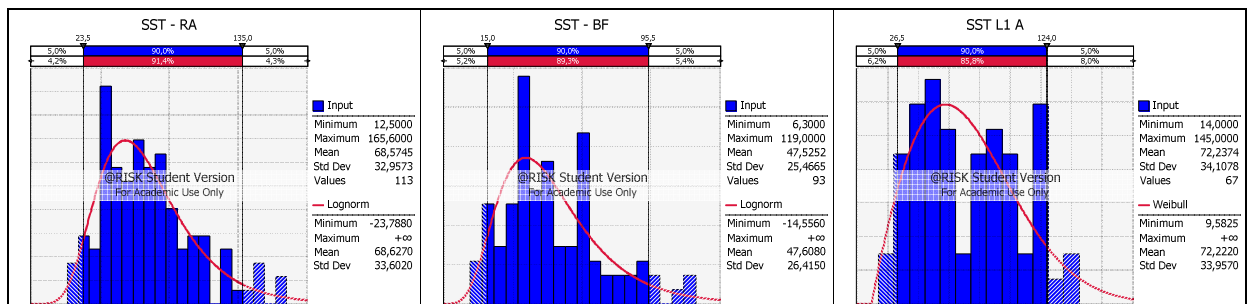


Figura 4.12 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de SST esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continua).

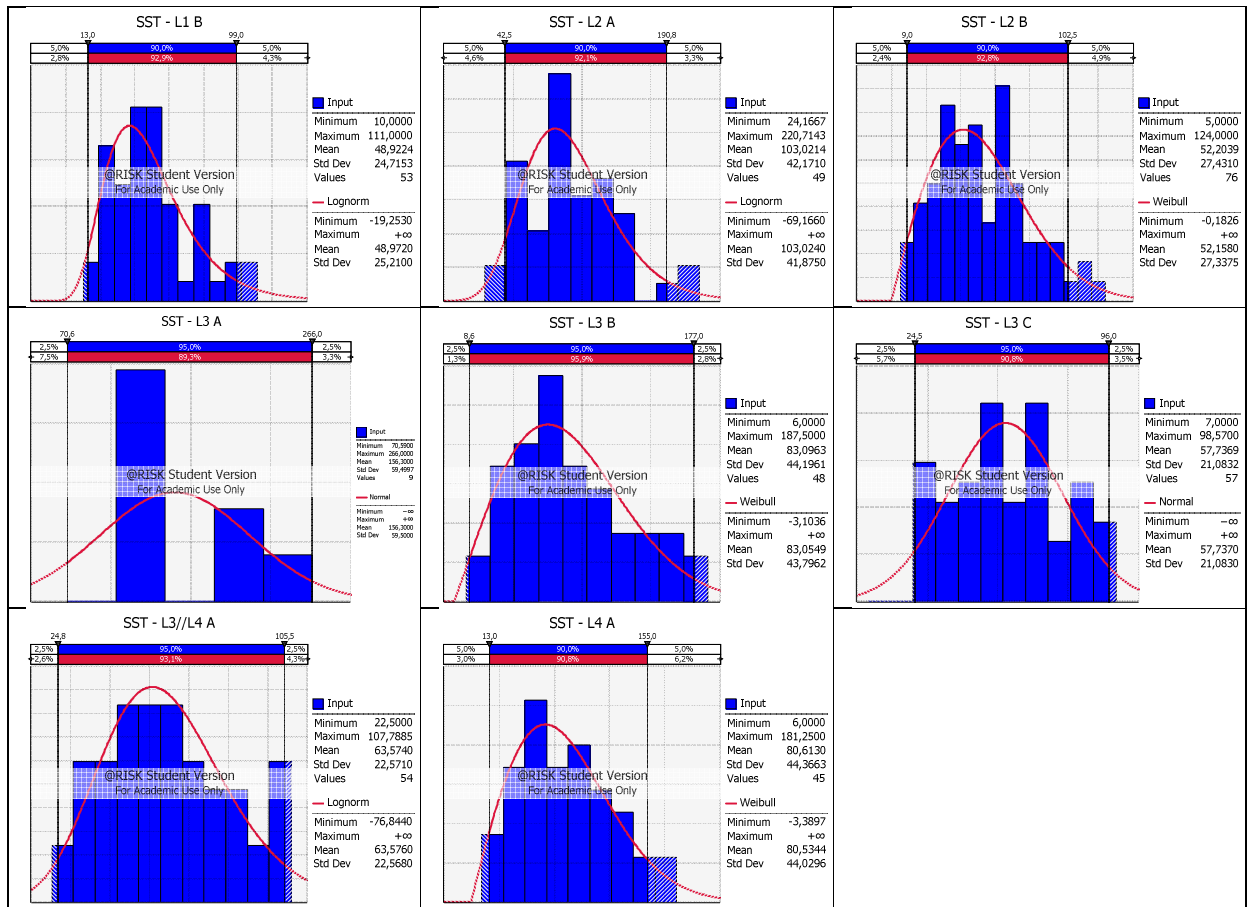


Figura 4.12 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de SST esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo com a lagoa 4 e lagoa 4 (continuação).

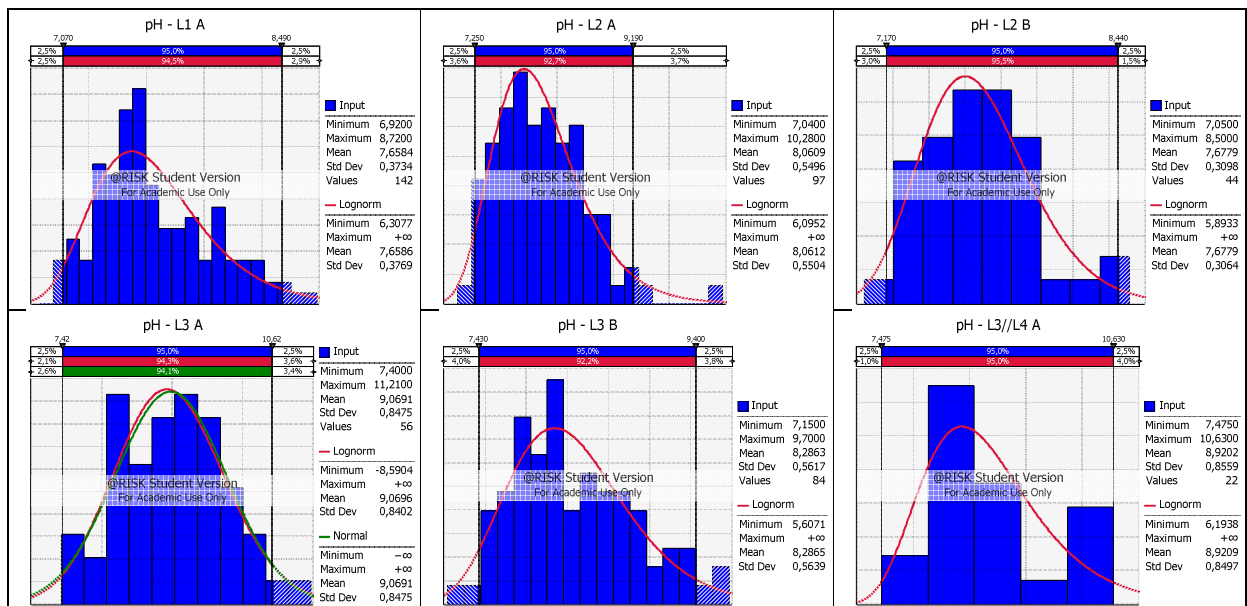


Figura 4.13 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de pH, lagoa 1, lagoa 2, lagoa 3 e lagoa 3 operando em paralelo com a lagoa 4 (continua).

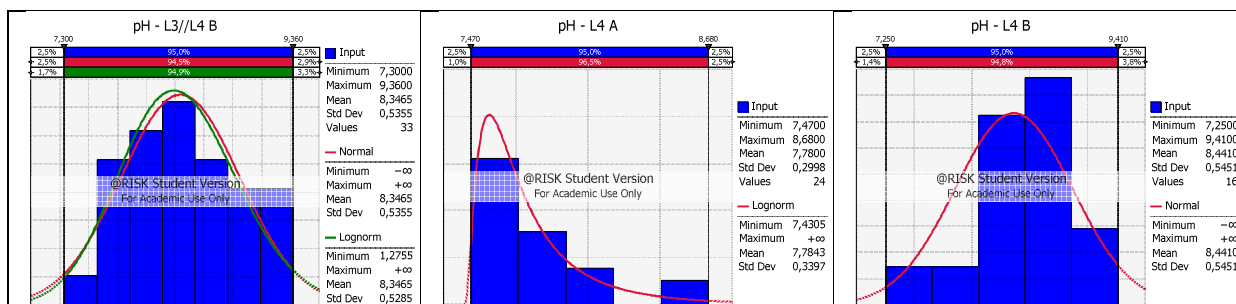


Figura 4.13 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de pH, lagoa 1, lagoa 2, lagoa 3 e lagoa 3 operando em paralelo com a lagoa 4 (continuação).

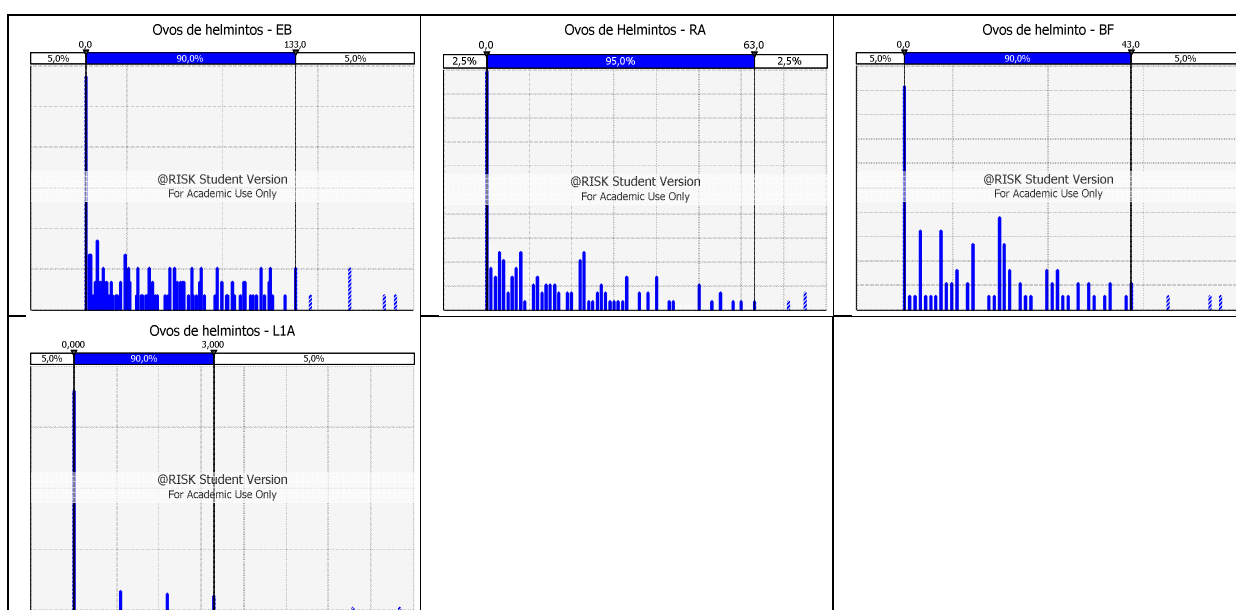


Figura 4.14 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de helmintos esgoto bruto, lagoa 1 e lagoa 2.

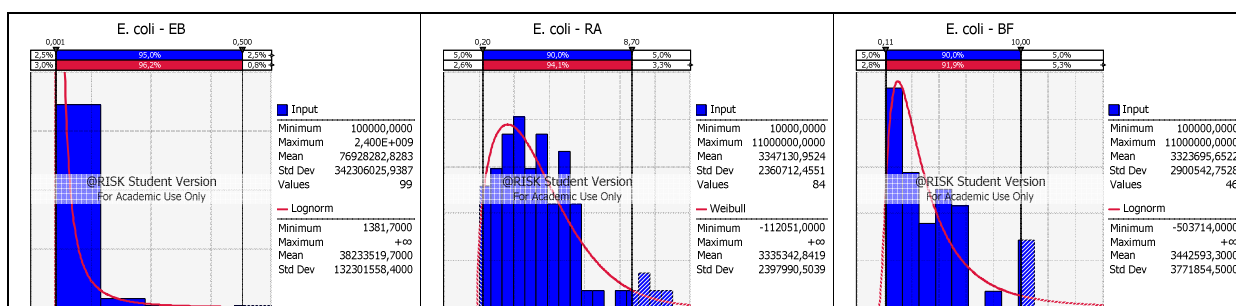


Figura 4.15 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de *E. coli*, esgoto bruto, reator anaeróbico, lagoa 1, lagoa 2, lagoa 3 e lagoa 4 (continua).

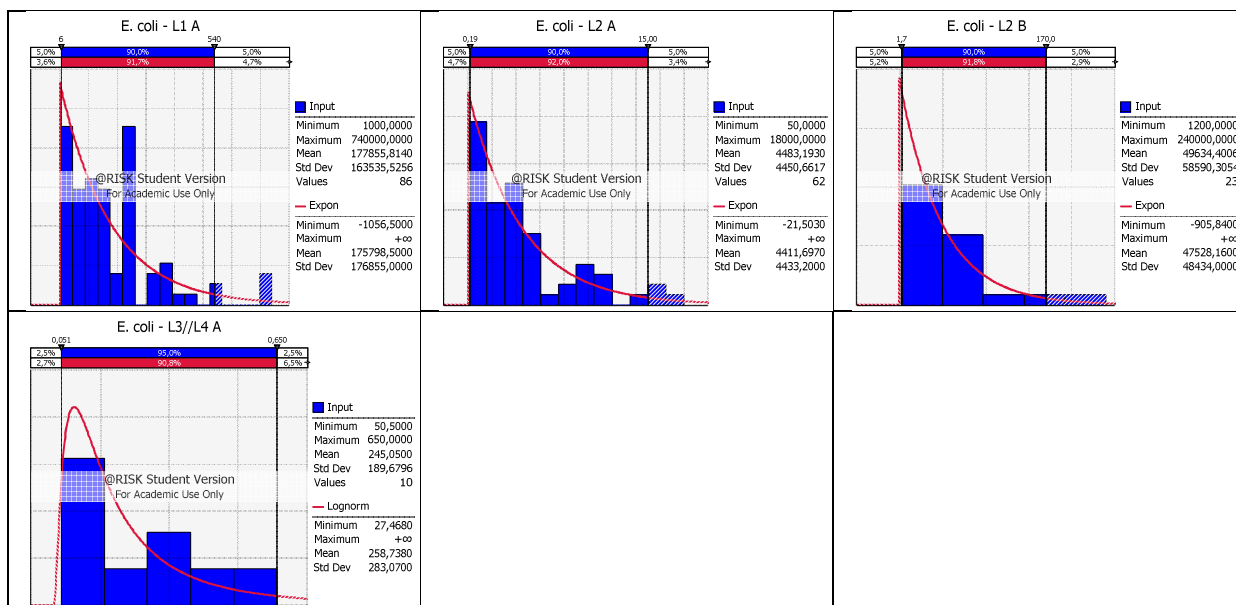


Figura 4.15 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de *E. coli*, esgoto bruto, reator anaeróbio, lagoa 1, lagoa 2, lagoa 3 e lagoa 4 (continuação).

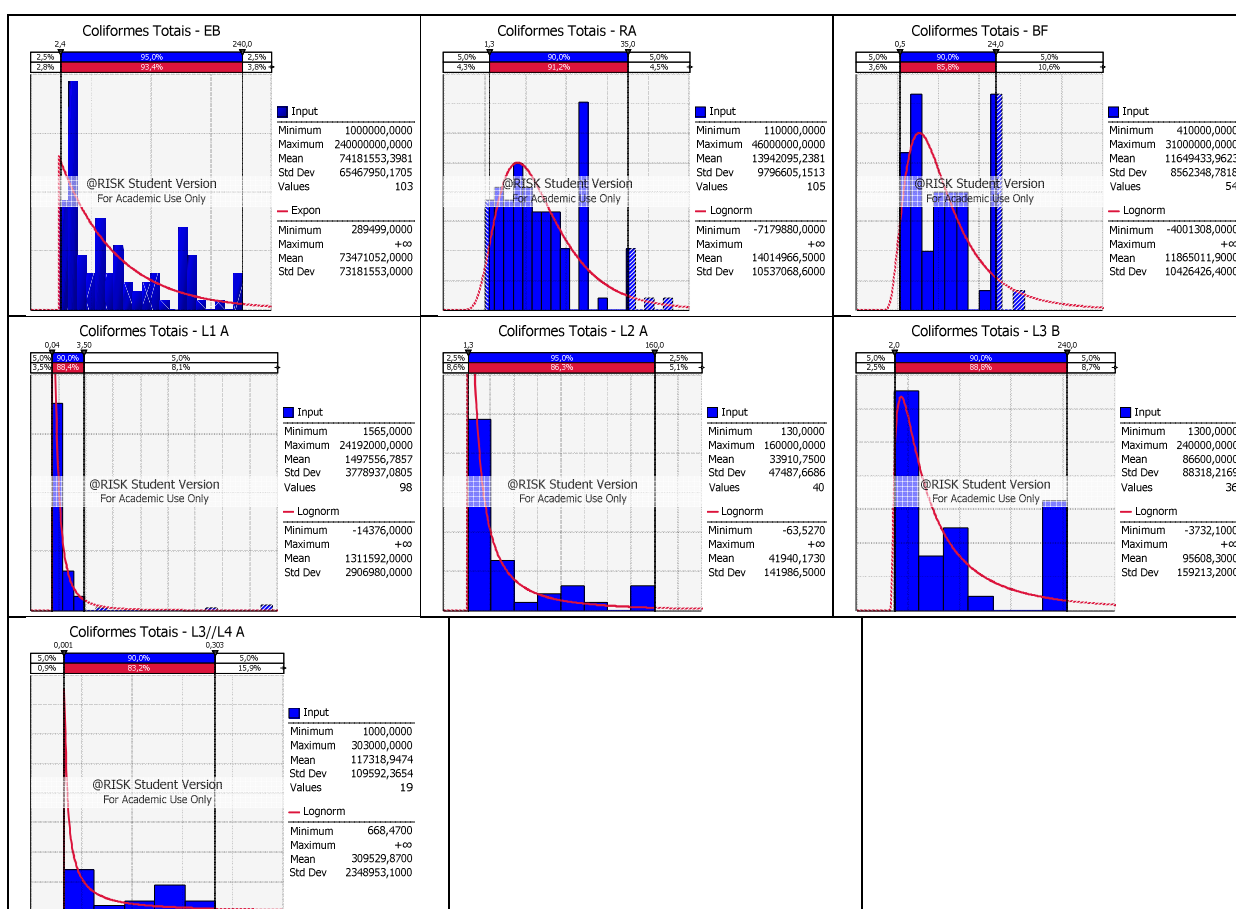


Figura 4.16 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de coliformes totais, esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3 e lagoa 3 operando em paralelo como lagoa.

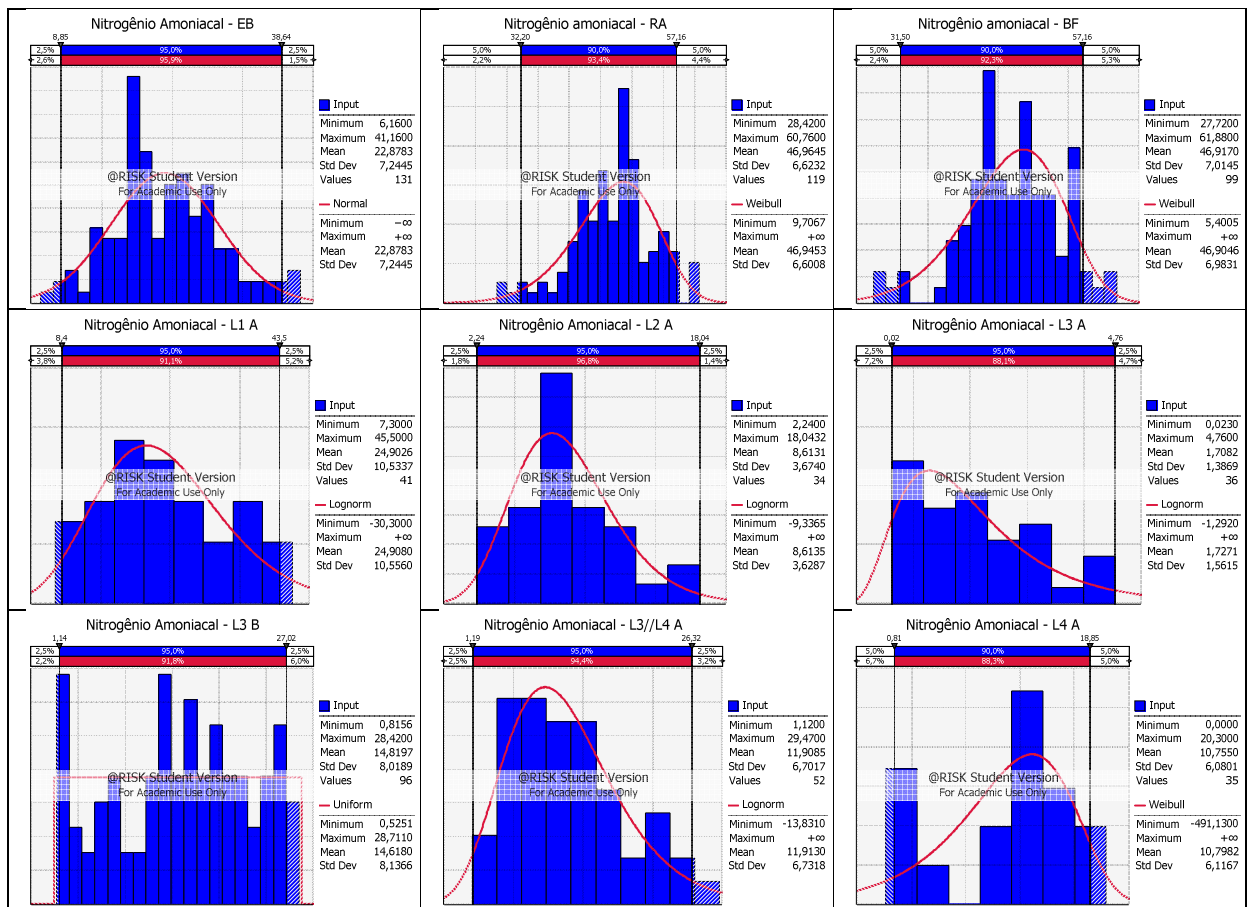


Figura 4.17 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de nitrogênio amoniaco esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo como lagoa 4 e lagoa 4.

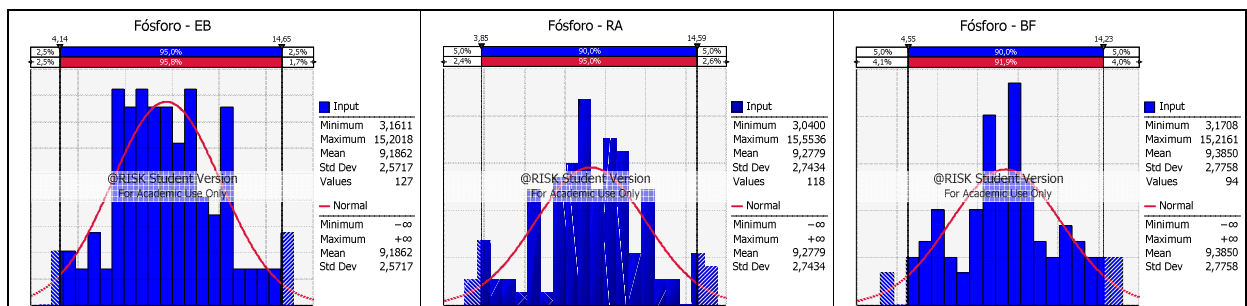


Figura 4.18 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de fósforo esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo como lagoa 4 e lagoa 4 (continua).

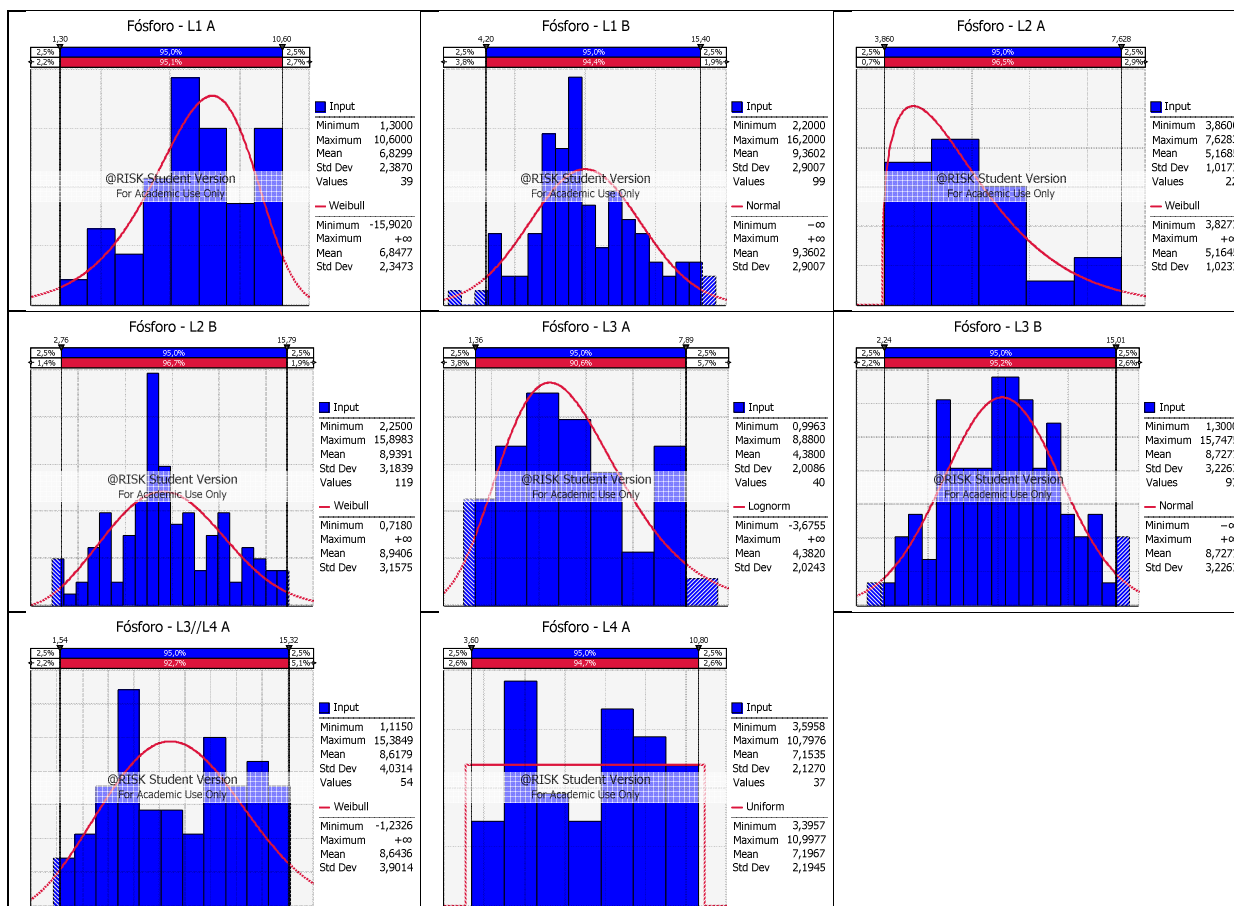


Figura 4.18 - Histogramas e curvas ajustadas das distribuições escolhidas, dados de fósforo esgoto bruto, reator anaeróbio, biofiltro, lagoa 1, lagoa 2, lagoa 3, lagoa 3 operando em paralelo como lagoa 4 e lagoa 4 (continuação).

#### 4.2.6 - Medidas de tendência central

A Organização Mundial de Saúde (OMS), em suas recomendações relativas à qualidade microbiológica para uso agrícola de esgotos, refere-se, nas diretrizes de 1989, à concentrações de coliformes termotolerantes em termos de média geométrica (WHO, 1989). A Resolução CONAMA 357/ 2005 refere-se ao uso da média geométrica na determinação da densidade de coliformes termotolerantes para o cultivo de moluscos bivalves destinados à alimentação humana (BRASIL, 2005). O emprego da média geométrica em bancos de dados bacteriológicos encontra-se de certa forma consolidado na literatura e apresenta consistência, pois tais dados tendem a apresentar grandes variações e valores elevados, o que exerce grande peso na média aritmética, distorcendo o conceito da tendência central. Como visto, os dados bacteriológicos foram os que apresentaram maiores assimetrias positivas, aderindo à distribuição lognormal.

As diretrizes da OMS especificam ainda o uso da média aritmética como tendência central para os ovos de helmintos em efluentes. Contudo, esses efluentes podem apresentar muitos valores iguais a zero, o que torna a média aritmética uma medida de tendência central majorada.

A Figura 4.19 ilustra as médias geométricas e médias aritméticas em curvas de frequência de algumas séries de dados bacteriológicos. Estas curvas, entendidas como aproximação da distribuição de frequências, permitem uma melhor visualização de que o uso da média geométrica é mais adequado em distribuições com grandes caudas à direita.

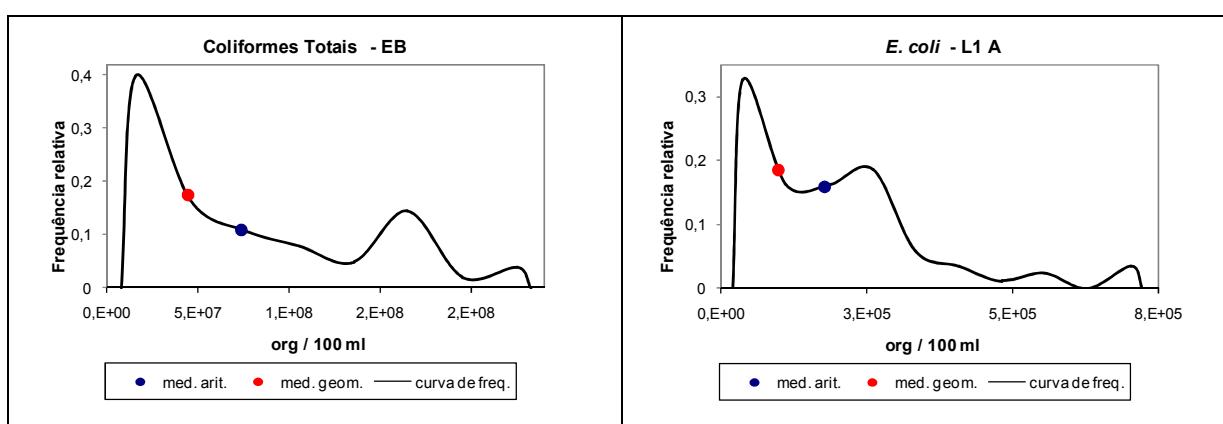


Figura 4.19 - Média geométrica e média aritmética em curvas de frequências de séries de dados de *E. coli* e coliformes totais.

Uma vez que os dados de DBO, DQO, SST e  $\text{NH}_3$  também apresentaram tendência à distribuições com assimetria positiva e, ao menos no caso da DBO, mais nitidamente à distribuição lognormal, os resultados aqui obtidos suscitam dúvidas quanto ao emprego da média aritmética com medida de tendência central para essas variáveis. Para a concentração de fósforo a média aritmética parece bem descrever a tendência central.

### 4.3 - Conclusões

- O número de dados demonstrou exercer grande influência sobre os coeficientes de assimetria e curtose, o teste de hipótese para independência, a forma do histograma e os testes de aderência. Séries muito pequenas fornecem resultados menos confiáveis.
- De modo geral, a maioria dos dados analisados atendeu as hipóteses de aleatoriedade, estacionariedade e homogeneidade, mas não a de independência.

- A maior parte das séries aderiu a distribuições assimetricamente positivas.
- Os dados das variáveis *E. coli*, coliformes totais e ovos de helmintos foram os que apresentaram assimetria positiva mais acentuada.
- A distribuição lognormal demonstrou ser a melhor representante do comportamento de todas as variáveis, exceto para a concentração de fósforo, que mostrou-se ser melhor representada pela distribuição normal. Não foi possível o ajuste de um modelo distributivo da frequência dos ovos de helmintos.

## **5 - ANÁLISE DE INCERTEZA E DE SENSIBILIDADE NA PREVISÃO DA QUALIDADE DE EFLUENTES DE LAGOAS DE POLIMENTO**

Como visto na revisão de literatura (capítulo 3), tradicionalmente, o procedimento adotado na previsão da qualidade dos efluentes de estações de tratamento de esgotos é o uso de valores fixos para os dados de entrada nas equações de dimensionamento, pressupondo, assim, certeza a respeito desses valores. Essa abordagem determinística leva a resultados tratados também como valores únicos. Em contrapartida, a abordagem estocástica leva em consideração incertezas em torno das variáveis de entrada, causadas basicamente por três fatores: variabilidade, erros de medidas e falta de conhecimento. Tais incertezas podem ser definidas por faixas de valores, às quais devem ser ajustadas distribuições de probabilidades e, por conseguinte, os resultados são também assim expressos.

Também como já referido, a abordagem estocástica aplicada ao tratamento de esgotos, incluindo a simulação de Monte Carlo, não constitui novidade e a literatura registra trabalhos envolvendo diferentes sistemas de tratamento, porém, trabalhos com lagoas de estabilização são escassos.

No Capítulo 4, a análise dos 6573 dados, incluindo esgoto bruto e um sistema UASB-biofiltro aerado submerso seguido de lagoas de polimento em série, permitiu a verificação da aderência de vários parâmetros de qualidade de efluentes/afluentes das lagoas a distribuições de probabilidades. Estes resultados são aqui utilizados em modelagem estocástica para a previsão da qualidade dos efluentes das lagoas de polimento. Outras variáveis de entrada, para as quais isso não foi possível, foram assumidas em faixas de valores (e distribuição uniforme) dentre as quais as lagoas em estudo foram operadas e, ou em faixas recomendadas de literatura.

Assim, o que se pretende neste capítulo não é a proposição de critérios de dimensionamento genéricos (tal como nos trabalhos de Von Sperling, 1996 e Gawasiri, 2003), mas a verificação da modelagem estocástica e das variáveis de entrada de maior impacto na previsão da qualidade dos efluentes do sistema de tratamento em questão.

## 5.1 - Metodologia

### 5.1.1 – Modelos de previsão da qualidade dos efluentes

A estimativa da qualidade dos efluentes foi realizada com o programa @Risk 5.5 (PALISADE CORPORATION, 2009). Os modelos consagrados na literatura para remoção de *E. coli* e de amônia, e as incertezas relativas às respectivas variáveis de entrada, foram inseridos no referido programa, o qual realiza modelagem estocástica por meio de simulação aleatória por amostragem por ‘Hipercubos Latinos’ (ver item 3.1.3). A simulação envolveu 10.000 iterações. Os resultados foram expressos por meio de gráficos com histogramas de frequências e distribuição acumulada.

Os dados de qualidade dos efluentes (pH, N-NH<sub>3</sub> e *E. coli*) foram considerados de acordo com as faixas de variação observadas no banco de dados de monitoramento do sistema experimental de tratamento e com as distribuições de probabilidades definidas no Capítulo 4. Para a previsão da concentração de *E. coli* efluente foi utilizado o modelo de fluxo disperso (Equação 3.31). As incertezas das variáveis de entrada foram consideradas como a seguir.

- Concentração de *E. coli* afluente à lagoa 1: distribuição lognormal definida a partir dos percentis de 5, 50 e 95%. A partir desses três valores de percentis o programa @Risk encontra os valores de média e desvio padrão para definição da distribuição lognormal. A concentração de *E. coli* afluente à lagoa 1 refere-se ora à concentração efluente do reator anaeróbio, ora do biofiltro.
- Número de dispersão ( $d$ ): 0,1 a 0,5; de acordo com o sugerido por von Sperling (2002) para lagoas de maturação com relação  $L/B = 2$  a 4, já que as lagoas de polimento do sistema estudado possuem relação  $L/B = 2$ . Por não se dispor de informações suficientes, foi assumido que os valores de  $d$  variassem de acordo com distribuição uniforme.
- Tempo de detenção hidráulica (TDH) e coeficiente de remoção de *E. coli* ( $K_{b20}$ ) em cada lagoa de acordo com as faixas de valores observadas nos diversos períodos de operação do sistema. A essas variáveis também foi atribuída distribuição uniforme.

Para a previsão da concentração de amônia efluente foram utilizados os modelos de Pano e Middlebrooks (1982) (Equações 3.36 e 3.37). As incertezas das variáveis de entrada foram consideradas como a seguir.

- Concentração de nitrogênio amoniacal afluyente: distribuição lognormal definida pelos percentis 5, 50 e 95%. Do mesmo modo que para a *E. coli*, a concentração afluyente de nitrogênio amoniacal refere-se ora à concentração efluyente do reator anaeróbio, ora do biofiltro.
- Valor de pH em cada lagoa: distribuição lognormal definida a partir dos percentis de 5, 50 e 95%. A partir de três valores de percentis, o programa @Risk encontra os valores de média e desvio padrão para definição da distribuição lognormal.
- Taxa de aplicação hidráulica superficial (Q/A) em cada lagoa de acordo com as faixas de valores observadas nos diversos períodos de operação do sistema. À essas variáveis também foi atribuída distribuição uniforme.
- Temperatura do líquido: determinada com a Equação 3.27, que relaciona a temperatura do ar com a do líquido. Esta equação foi ajustada a partir de dados da mesma série experimental de lagoas em questão (RIOS, 2007).
- Temperatura do ar: entre 14,8 a 26,4 °C com média de 19,4 °C. Valores baseados nas normais climatológicas para o município de Viçosa, MG no período de 1960 a 1990, obtidas do Instituto Nacional de Meteorologia (INMET). A faixa adotada variou entre as médias das temperaturas mínimas e máximas absolutas do período em questão. A distribuição de probabilidades adotada foi a triangular, que é uma simplificação da distribuição normal. Caso optássemos pela distribuição uniforme, a temperatura média corresponderia ao valor intermediário entre o valor máximo e mínimo. Porém, como se tinha conhecimento da temperatura média e esta se encontra abaixo do valor intermediário entre a máxima e a mínima, optou-se pela distribuição triangular.

A Tabela 5.1 resume as faixas de valores e as distribuições de probabilidades adotadas para as variáveis de entrada dos modelos de previsão da qualidade dos efluentes das lagoas.

A Figura 5.1 mostra os histogramas de frequências, obtidos por simulação de Monte Carlo, das variáveis de entrada cujas distribuições probabilidades foram assumidas como lognormal ou triangular.

TABELA 5.1 – Variáveis de entrada dos modelos de previsão da qualidade dos efluentes da ETE Violeira com as respectivas distribuições de probabilidades e parâmetros que as definem.

Variáveis de entrada	Distribuição	Percentil			Média	Mínimo	Máximo
		5%	50%	95%			
<i>E. coli</i> afluente ( <i>E. coli</i> /100mL)	Lognormal	$1,08 \times 10^5$	$2,80 \times 10^6$	$8,83 \times 10^6$	-	-	-
pH L1	Lognormal	7,13	7,57	8,34	-	-	-
pH L2	Lognormal	7,29	7,85	8,80	-	-	-
pH L3	Lognormal	7,55	8,48	9,98	-	-	-
pH L4	Lognormal	7,52	7,92	8,79	-	-	-
Amônia afluente (mg/L)	Lognormal	30,90	44,30	57,72	-	-	-
Temperatura do ar (°C)	Triangular	-	-	-	19,4	14,8	26,4
TDH – L1 (dias)	Uniforme	-	-	-	-	3,36	9,40
TDH – L2 (dias)	Uniforme	-	-	-	-	3,36	9,40
TDH – L3 (dias)	Uniforme	-	-	-	-	2,31	9,40
TDH – L4 (dias)	Uniforme	-	-	-	-	0,75	2,50
Q/A – L1 (m/d)	Uniforme	-	-	-	-	0,096	0,268
Q/A – L2 (m/d)	Uniforme	-	-	-	-	0,096	0,268
Q/A – L3 (m/d)	Uniforme	-	-	-	-	0,048	0,159
Q/A – L4 (m/d)	Uniforme	-	-	-	-	0,048	0,159
$K_{b20} - L1$ (d <sup>-1</sup> )	Uniforme	-	-	-	-	0,210	0,840
$K_{b20} - L2$ (d <sup>-1</sup> )	Uniforme	-	-	-	-	0,243	1,702
$K_{b20} - L3$ (d <sup>-1</sup> )	Uniforme	-	-	-	-	0,332	2,324
$K_{b20} - L4$ (d <sup>-1</sup> )	Uniforme	-	-	-	-	0,212	1,482
D	Uniforme	-	-	-	-	0,1	0,5

Li: lagoas; TDH: tempo de detenção hidráulica; Q/A: taxa de aplicação hidráulica superficial;  $K_{b20}$ : coeficiente de decaimento de *E. coli* (20°C)

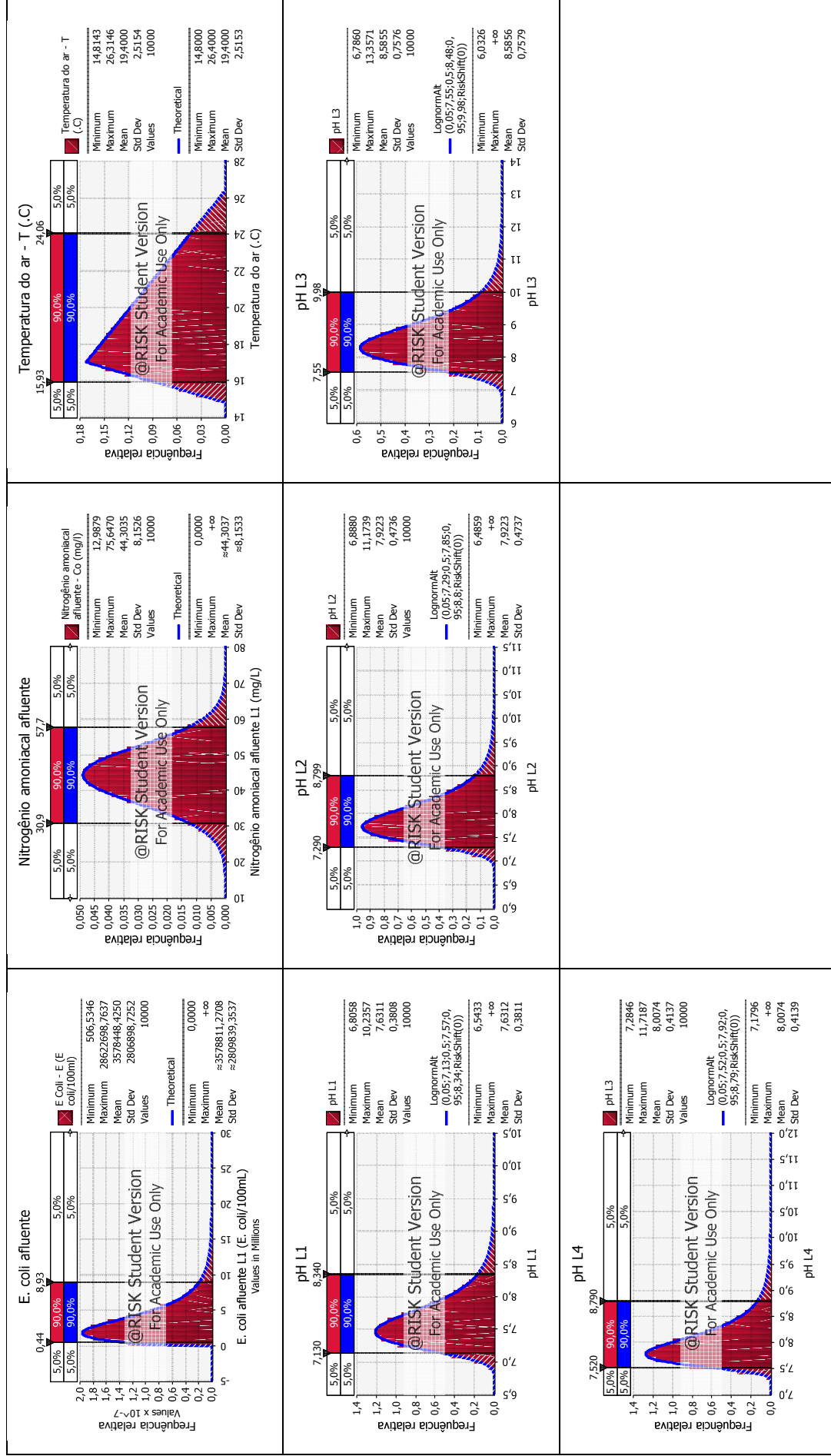


Figura 5.1 – Histogramas de frequências de variáveis de entrada dos modelos de previsão de concentração de nitrogênio de *E. coli* e nitrogênio amoniacal com distribuição lognormal ou triangular.

### 5.1.2 - Análises de sensibilidade e de cenário

A análise de sensibilidade foi realizada, também com o programa @Risk 5.5, utilizando gráficos de dispersão e medida de correlação não paramétrica de Spearman. Adicionalmente foi realizada análise de sensibilidade em relação à mediana, como descrito em von Sperling (1996), utilizando o teste não paramétrico Mann-Whitney-Wilcoxon com o programa Stat Tools 5.5. Nesse caso, a variável de entrada foi considerada crítica na determinação do valor de saída quando os dois grupos da variável de entrada (aquele que conduz a valores de saída menor que a mediana e aquele que conduz a valores de saída maiores que a mediana) eram diferentes ao nível de 1% de significância. Cabe esclarecer que este nível de significância mais rigoroso foi aqui adotado após constatação de que o adotado nos demais testes (5%) não resultou em diferenças significativas.

A análise de cenários foi realizada tomando como referência valores de saída maiores que  $10^3$ ,  $10^4$  e  $10^5$  *E. coli*/100mL, no caso da previsão de qualidade em termos de *E. coli*, e maiores que 5 e 20 mg/L, no caso da previsão da concentração de nitrogênio amoniacal.

Os valores  $10^3$ ,  $10^4$  e  $10^5$  *E. coli*/100 mL são encontrados nas Diretrizes da OMS para o uso de efluentes em irrigação irrestrita (WHO, 2006). O padrão de  $10^3$  *E. coli*/100mL é recomendado para o cultivo de raízes e tubérculos e para a irrigação localizada de plantas que se desenvolvem rentes ao nível do solo. O nível de qualidade  $\leq 10^4$  *E. coli*/100mL é recomendado no cultivo de folhosas; o nível  $\leq 10^5$  *E. coli*/100mL é recomendado para irrigação localizada de plantas que se desenvolvem distantes do nível do solo.

Os limites considerados para nitrogênio amoniacal tiveram como referência padrões de lançamento de efluentes estabelecidos em diferentes versões das resoluções do Conselho Nacional do Meio Ambiente (CONAMA) que regulamentam a matéria. A Resolução CONAMA N° 20/1986 previa um valor máximo de 5 mg/L. Posteriormente, a Resolução CONAMA N° 375/2005 passou a limitar este valor em 20 mg/L. Entretanto, em 2008, a Resolução CONAMA N° 397 excluiu o nitrogênio amoniacal das exigências de lançamento de efluentes de sistemas de tratamento de esgotos sanitários.

Não obstante, a Análise de Sensibilidade foi realizada tomando como referência os valores de 5 e 20 mg/L, levando em consideração que: (i) a referida exclusão do nitrogênio amoniacal tem caráter temporário, (ii) tal decisão tem motivações pragmáticas e não ambientais; (iii) o valor mais restritivo de 5 mg/L aproxima-se de limites recomendados para o uso de efluentes em piscicultura (EDWARDS,1992), ou seja estar-se-ia verificando os

resultados de estimativas de qualidade dos efluentes tendo em vista também o reuso e não somente o lançamento em corpos receptores.

Complementarmente, a Análise de Cenário foi realizada tomando como referências os valores das medianas e dos percentis 95% das distribuições de probabilidades das estimativas de risco.

Como descrito no item 3.2, a execução dessa análise envolve a subdivisão das iterações da simulação e a seleção daquelas em que a variável de saída alcança o valor desejado e, em seguida, a análise dos valores amostrados para cada variável de entrada nestas iterações. São encontradas as medianas desse subconjunto de valores amostrais para cada variável de entrada, as quais são comparadas com as medianas relativas a todas as iterações. O objetivo desse processo é encontrar aquelas variáveis de entrada cujo subconjunto, ou ‘mediana condicional’, se diferencia significativamente da mediana geral. Se a mediana do subconjunto é próxima à mediana geral, a variável de entrada em questão é considerada não significativa. Entretanto, se a mediana do subconjunto para a variável de entrada se desviar de forma significativa da mediana geral (isto é, pelo menos  $\frac{1}{2}$  desvio padrão) a variável de entrada é considerada significativa.

## 5.2 - Resultados e discussão

### 5.2.1 – Estimativas da qualidade dos efluentes das lagoas de polimento

A Tabela 5.2 apresenta os resultados da previsão da concentração de *E. coli* efluente em termos de mediana e percentil de 95% da distribuição de frequência resultante. Na Figura 5.2 apresentam-se os histogramas com as distribuições de probabilidades da qualidade estimada dos efluentes.

TABELA 5.2 - Estimativa da qualidade dos efluentes das lagoas de polimento (medianas e percentis 95%), *E. coli* por 100 mL (\*).

Lagoa	Mediana	95%
Lagoa 1	$3,13 \times 10^5$	$1,67 \times 10^6$
Lagoa 2	$1,05 \times 10^4$	$1,61 \times 10^5$
Lagoa 3	$2,31 \times 10^2$	$8,91 \times 10^3$
Lagoa 4	$7,95 \times 10^1$	$3,59 \times 10^3$

\* qualidade dos afluentes às lagoas de acordo com a Tabela 5.1

Os resultados da Tabela 5.2, de estimativas de qualidade dos efluentes das lagoas, mostram algumas discrepâncias com os dados de campo, de monitoramento – parece haver tendência de superestimativa da qualidade dos efluentes, ou seja, os valores estimados são por vezes inferiores aos de fato produzidos na série de lagoas (Tabela 4.8). Entretanto, é preciso salientar que os dados da Tabela 5.2 foram obtidos da modelagem estocástica levando em consideração variações das variáveis de entrada do modelo de estimativa de *E. coli* efluente que cobrissem todos os onze períodos operacionais das lagoas. Por sua vez, no capítulo 3 os dados de monitoramento foram agrupados por períodos operacionais com resultados estatisticamente equivalentes. Ou seja, uma análise mais rigorosa demandaria a comparação em mesma base temporal.

De toda forma, os resultados aqui apresentados reforçam a importância de verificação de que variáveis de entrada impactam mais nitidamente as estimativas de qualidade dos efluentes das lagoas. Note-se, por exemplo, que Rios (2007) encontrou que os modelos propostos por von Sperling (2002) tendem a subestimar os valores de  $K_b$  para estas lagoas experimentais.

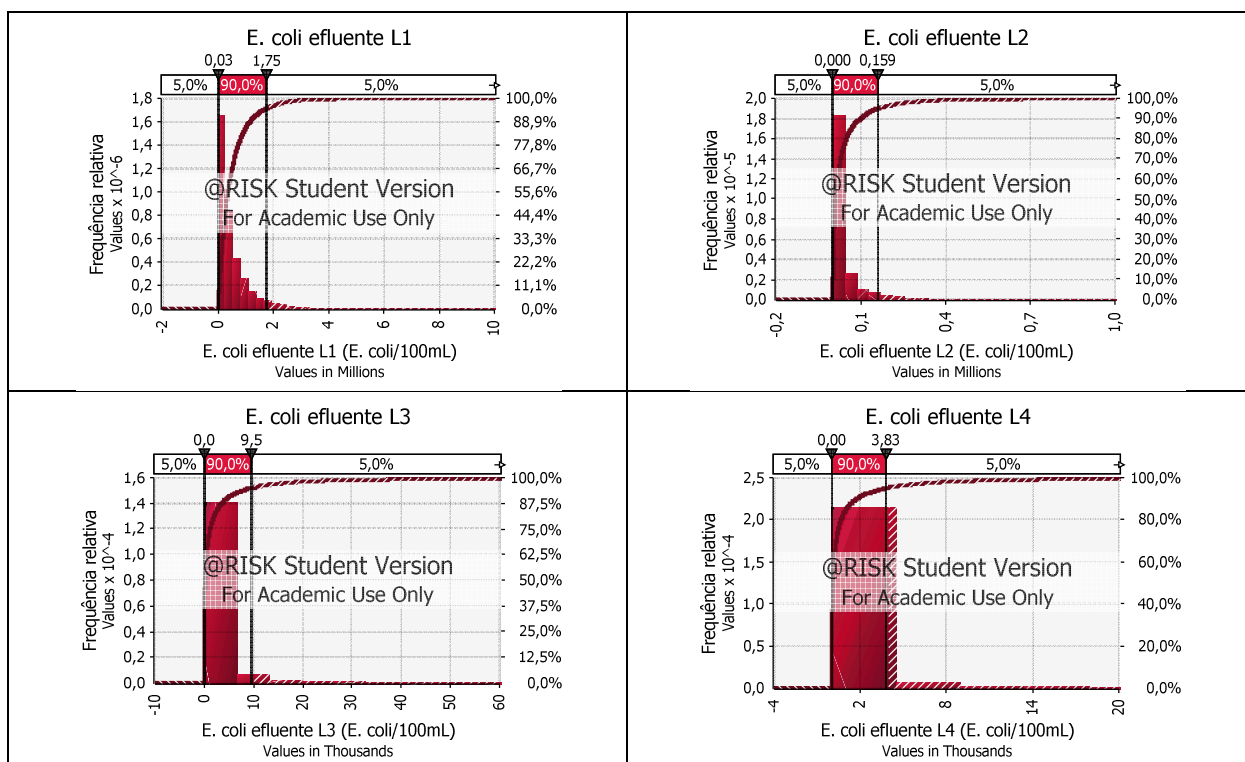


Figura 5.2 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das estimativas das concentrações de *E. coli* efluente a partir das concentrações afluentes segundo distribuição lognormal.

Observa-se na Figuras 5.2 que as qualidades estimadas dos efluentes apresentaram distribuições com grande concentração de dados à esquerda do histograma, o que é esperado, uma vez que distribuição probabilidades para a concentração de *E. coli* nos afluentes às lagoas foi descrita como lognormal no Capítulo 3.

Na Tabela 5.3 são apresentados resultados em termos de percentual de atendimento aos níveis de qualidade de efluentes tomados como referência ( $10^3$ ,  $10^4$  e  $10^5$  *E. coli* por 100 mL).

Tomando como base a recomendação da OMS para irrigação localizada de plantas com desenvolvimento rentes ao solo ou irrigação de tubérculos e raízes ( $\leq 10^3$  *E. coli* / 100 mL) seriam necessárias três lagoas em série para atingir a qualidade com frequência mínima de 50%. Por outro lado, optando-se por um nível de qualidade com percentual de atendimento mais rigoroso, de 95%, a quarta lagoa ainda não seria suficiente.

Para atingir a qualidade de  $\leq 10^4$  *E. coli* / 100 mL, conforme as diretrizes para irrigação de folhosas seriam necessárias três lagoas para atender a frequência mínima de 95%. No entanto, na segunda lagoa a concentração de *E. coli* situa-se muito próxima ao nível de  $10^4$  *E. coli* / 100 mL em 50% dos resultados.

Por fim, considerando o nível de qualidade menos restritivo de  $\leq 10^5$  *E. coli* / 100 mL para irrigação localizada de plantas com desenvolvimento distante do nível do solo, seriam necessárias 2 ou 3 lagoas respectivamente para atender as frequências mínimas de 50 e 95%.

TABELA 5.3 – Percentuais de atendimentos aos níveis de qualidade  $10^3$ ,  $10^4$  e  $10^5$

*E. coli*/100mL<sup>(\*)</sup>

Lagoa	$\leq 10^3$ <i>E. coli</i> /100mL	$\leq 10^4$ <i>E. coli</i> /100mL	$\leq 10^5$ <i>E. coli</i> /100mL
Lagoa 1	0,00%	1,30%	18,50%
Lagoa 2	11,90%	48,80%	90,20%
Lagoa 3	72,30%	95,50%	99,80%
Lagoa 4	85,50%	98,50%	99,90%

\* qualidade dos efluentes de acordo com a Tabela 5.1

Os dados da Tabela 5.3 revelam ainda que uma série de três lagoas seria suficiente para atender todos os requisitos de qualidade da OMS para irrigação irrestrita, assumindo-se a mediana dos resultados como referência. Por outro lado, optando-se por uma referência mais rigorosa, o percentil de 95%, apenas o nível de  $\leq 10^3$  *E. coli* / 100 mL não é atendido na terceira lagoa e nem mesmo na quarta lagoa. Entretanto, percebe-se que o ganho de qualidade a partir da terceira lagoa não é grande. Desta forma, a fixação de um padrão de qualidade tão

rigoroso ( $\leq 10^3$  *E. coli* / 100 mL com frequência de 95%) poderia ser interpretado como injustificável ou inviável, porém uma análise mais completa será feita no capítulo seguinte em termos de risco de infecção.

A Tabela 5.4 apresenta os resultados da previsão da concentração de nitrogênio amoniacal efluente em termos de mediana e percentil de 95% da distribuição de frequência resultante. Na Figura 5.3 apresentam-se os histogramas com as distribuições de probabilidades da qualidade estimada dos efluentes.

TABELA 5.4 - Estimativa da qualidade dos efluentes das lagoas de polimento (mediana e percentis 95%), mg N-NH<sub>3</sub> / L (\*)

<b>Lagoa</b>	<b>Mediana</b>	<b>95%</b>
Lagoa 1	37,90	51,05
Lagoa 2	29,92	43,08
Lagoa 3	14,22	29,04
Lagoa 4	9,17	20,73

\* qualidade dos afluentes às lagoas de acordo com a Tabela 5.1

Assim como observado para a *E. coli*, os resultados estimados de qualidade dos efluentes das lagoas (Tabela 5.4) mostraram discrepâncias com os dados de campo (Tabela 4.5). Mas, nesse caso, dependendo do período operacional considerado, as superestimativas das concentrações efluentes de amônia são marcantes, ou seja a modelagem resultou em concentrações maiores que a de fato produzida na série de lagoas (Tabela 4.5). Entretanto, as mesmas ressalvas feitas no caso da *E. coli* aqui se aplicam: uma análise mais rigorosa demandaria a comparação em mesma base temporal.

Rios (2007), com base em banco de dados destas mesmas lagoas (cobrindo, entretanto um período menor de monitoramento), encontrou que o modelo de Pano e Middlebrooks apresentou um bom ajuste aos dados experimentais para a estimativa de remoção de amônia, embora, de alguma forma superestimando a concentração efluente de amônia.

Apesar de a distribuição lognormal ter sido a que melhor aderiu aos dados de nitrogênio amoniacal, os percentis utilizados para descrever a distribuição de frequência das concentração afluentes à lagoa 1 forneceram distribuição com assimetria próxima à zero (0,13) e curtose próxima à três (2,95), assemelhando-se muito à distribuição normal. A simetria presente na distribuição normal pode ser verificada nos histogramas dos dados de nitrogênio amoniacal efluentes às lagoas 1 e 2. Já nos histogramas dos efluentes das lagoas 3 e 4 percebe-se nítida assimetria, característica da distribuição lognormal. Nota-se, portanto, que

a concentração dos dados à esquerda do histograma é adquirida como a progressão no tratamento do efluente.

Na Tabela 5.5 são apresentados resultados em termos de percentual de atendimento aos níveis de qualidade de efluentes tomados como referência (5 e 20 mg N-NH<sub>3</sub>/L).

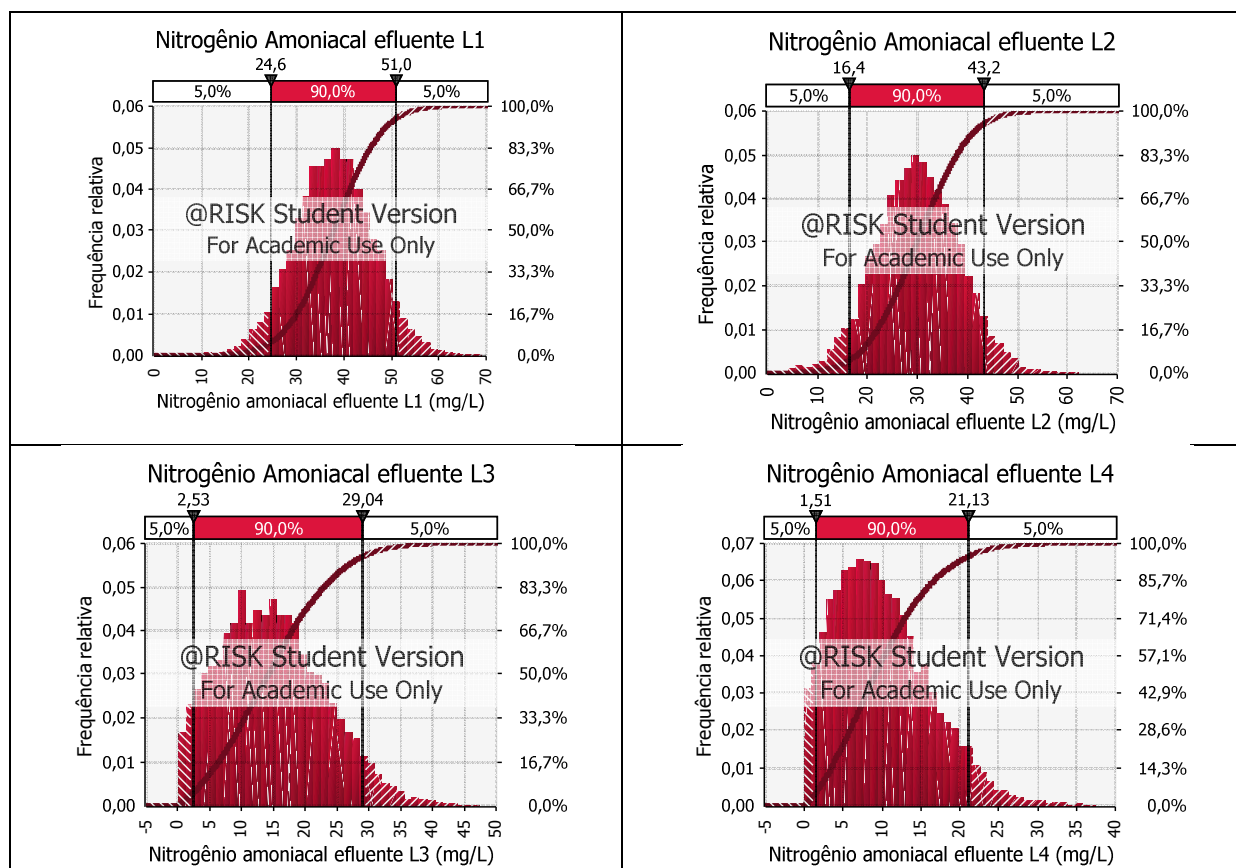


Figura 5.3 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das estimativas das concentrações de nitrogênio amoniacal efluente e pH a partir das concentrações afluentes segundo distribuição lognormal.

TABELA 5.5 – Percentuais de atendimentos aos níveis de qualidade de 5 e 20 mg N-NH<sub>3</sub>/L<sup>(\*)</sup>.

Lagoa	5 mg/L	20 mg/L
Lagoa 1	0,00%	1,50%
Lagoa 2	0,10%	11,20%
Lagoa 3	11,70%	73,70%
Lagoa 4	23,00%	93,50%

\* qualidade dos efluentes de acordo com a Tabela 5.1

O valor de referência da CONAMA 357/2005 (20 mg/L) para lançamento de efluentes foi alcançado no efluente da Lagoa 4 em frequência bastante elevada (mais de 90%) e em frequência um pouco menor, mas ainda elevada, na Lagoa 3 (73,70%). No efluente da Lagoa 4, 95% dos dados de concentração de nitrogênio amoniacal mantiveram-se abaixo de 20,73 mg/L. Por outro lado, o padrão de qualidade mais exigente (5 mg/L) não foi satisfatoriamente atendido em nenhuma lagoa. Porém, cabem aqui as ressalvas já feitas de que a modelagem, por vezes, parece superestimar em muito as concentrações efluentes de amônia (RIOS, 2007).

Para fins de comparação do uso das distribuições lognormal na previsão de qualidade dos efluentes, foram realizadas previsões utilizando a distribuição uniforme para as concentrações de *E. coli* e nitrogênio amoniacal afluentes, bem como para os valores de pH. No caso das concentrações de *E. coli*, as diferenças entre os resultados foram expressas em unidades logarítmicas por se tratar de uma variável com valores muito elevados e expressos em base exponencial. Os resultados são apresentados na Tabelas 5.6 e nas Figuras 5.4 e 5.5.

TABELA 5.6 – Diferenças (mediana e percentil 95%) entre a previsão das concentrações efluentes de *E. coli* e de nitrogênio amoniacal utilizando valores de pH, concentrações de *E. coli* e de nitrogênio amoniacal afluentes segundo distribuição uniforme em relação ao uso da distribuição e lognormal.

Lagoa	<i>E. coli</i> *		Nitrogênio amoniacal	
	Mediana	95%	Mediana	95%
L1	0,15	0,04	0,05	0,50
L2	0,14	0,10	-1,45	0,11
L3	0,13	0,09	-3,06	-1,21
L4	0,14	0,04	-2,62	-1,95

\* Diferenças em unidades logarítmicas

O uso da distribuição lognormal resultou em concentrações efluentes de *E. coli* menores do que quando se utilizou a distribuição uniforme.

Na previsão das concentrações efluentes de nitrogênio amoniacal, o uso da distribuição lognormal, em comparação ou uso da uniforme, resultou em concentrações menores no início do tratamento e maiores no final. Estas diferenças assumem relevância

principalmente quando se considera o padrão de qualidade efluente mais rigoroso (5 mg/L). Nota-se ainda que as diferenças são crescentes ao longo da série de lagoas.

Para verificar a existência de diferenças significativas entre as concentrações efluentes resultantes do uso da distribuição lognormal ou uniforme nas concentrações afluentes aplicou-se o teste não-paramétrico para dados pareados Wilcoxon ao nível de 5% de significância. Embora as diferenças pareçam pequenas as concentrações de todas as lagoas apresentaram diferenças significativas entre o uso das duas distribuições, tanto para a previsão de concentração e *E. coli* como de nitrogênio amoniacal.

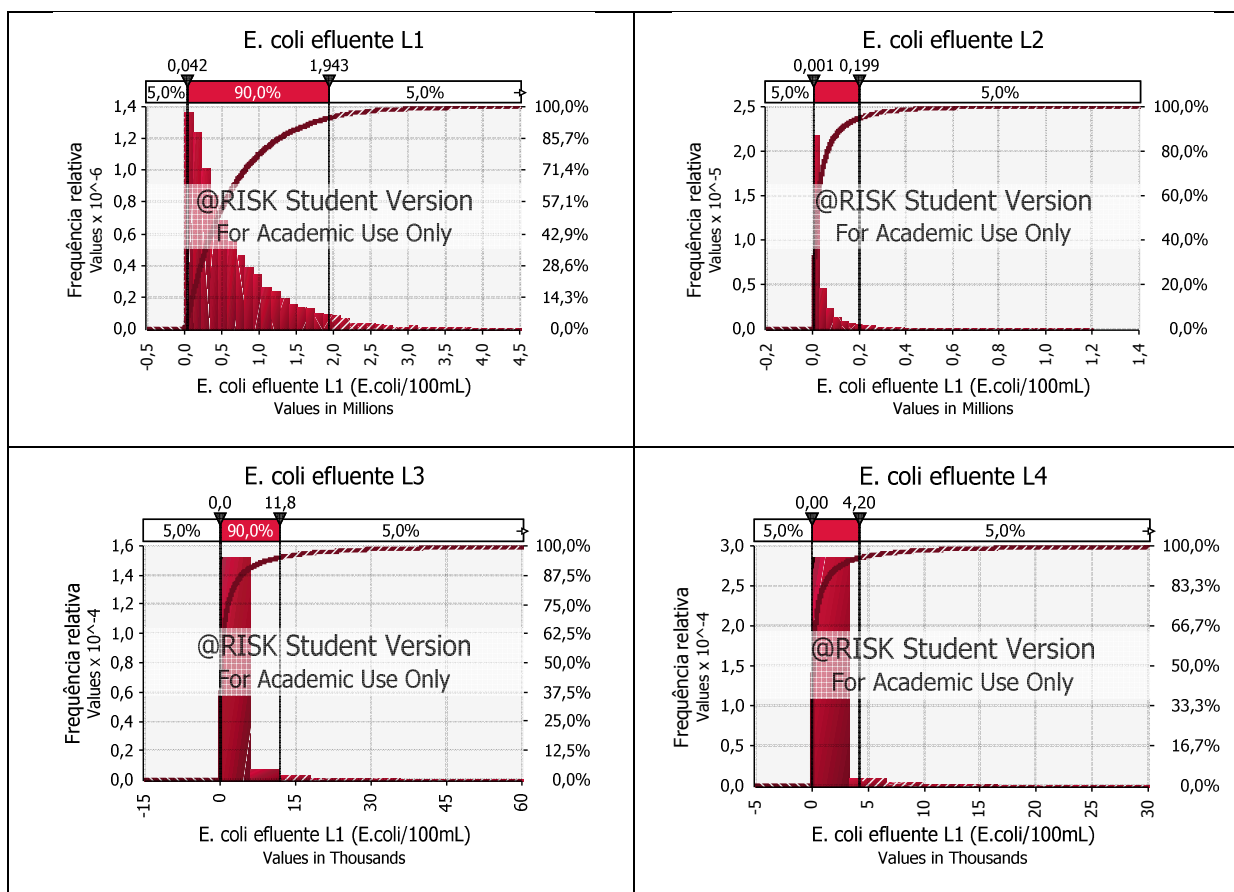


Figura 5.4 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada da previsão de concentração de *E. coli* efluente a partir das concentrações afluentes segundo distribuição uniforme.

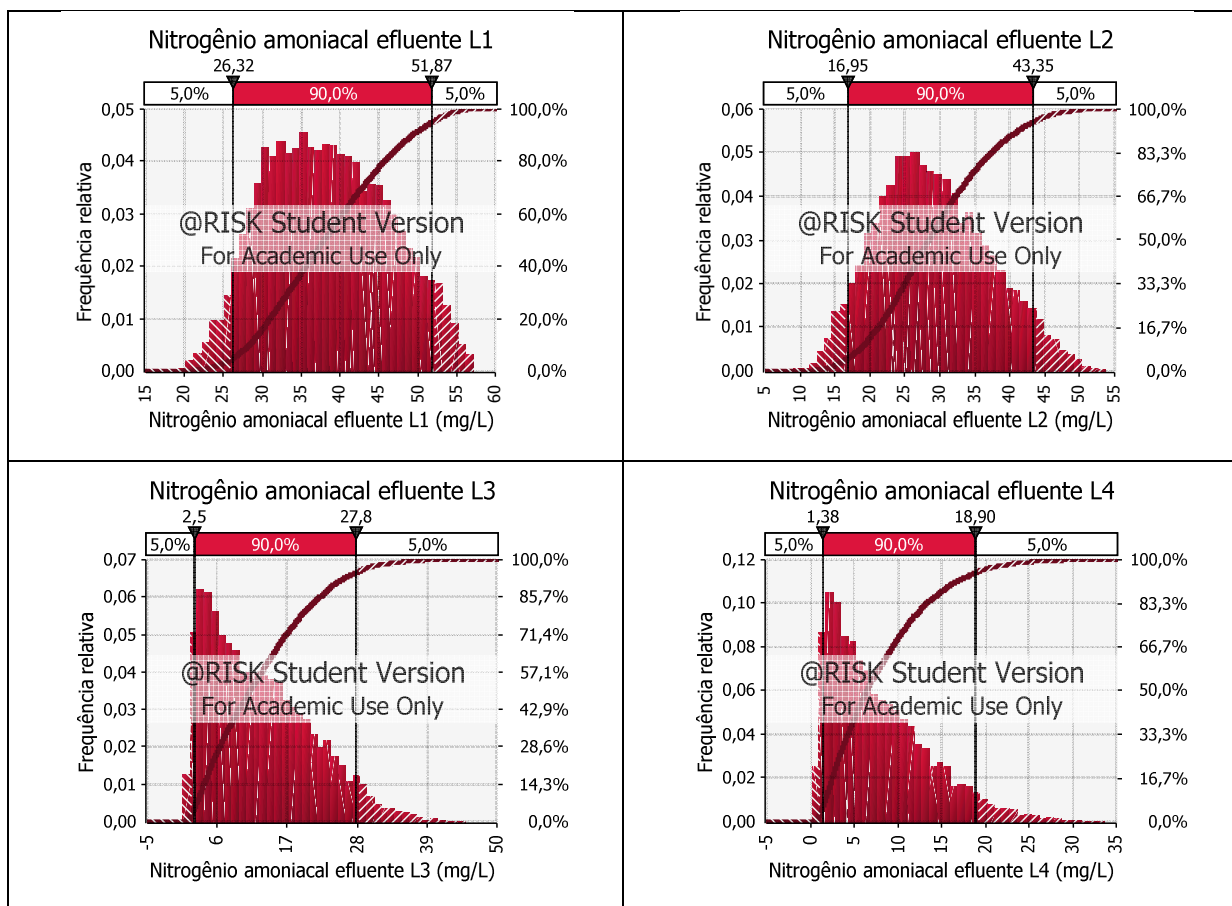


Figura 5.5 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada das previsões de concentrações de nitrogênio amoniacal efluente a partir das concentrações afluentes e pH segundo distribuição uniforme.

### 5.2.2 - Análise de sensibilidade

Os Quadros 5.1 e 5.2 mostram os coeficientes de Spearman, os resultados das análises de sensibilidade com relação à mediana e da Análise de Cenários. As Figuras 5.6 e 5.7 ilustram três casos com maiores coeficientes de Spearman. Quando.

A análise de cenários tem por objetivo indicar quais variáveis de entrada impactam significativamente o resultado para um dado valor de saída, no caso, níveis de risco tomados como referência. Assim, o resultado dessa análise foi expresso em forma de ‘sim’, para as variáveis que contribuem significativamente com os resultados de saída maiores que um determinado valor, e, caso contrário na forma de ‘não’. Naturalmente, quando a maior concentração efluente de *E. coli* ou de nitrogênio amoniacal gerada pelas 10.000 iterações da simulação de Monte Carlo foi menor que a concentração de referência ( $10^3$ ,  $10^4$  ou  $10^5$  *E. coli*/100mL ou 5 e 20 mg de  $\text{NH}_3/\text{L}$ ) não foi possível realizar a Análise de Cenários

Quadro 5.1 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para a concentração efluente de *E. coli*.

Entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>				
				> 10 <sup>3</sup>	> 10 <sup>4</sup>	> 10 <sup>5</sup>	> 50%	>95%
<i>E. coli</i> afluente ( <i>E. coli</i> /100mL)	L1	0,71	Importante	*	Não	Não	Sim	Sim
	L2	0,46	Importante	Não	Não	Sim	Não	Sim
	L3	0,33	Importante	Não	Sim	Sim	Não	Sim
	L4	0,32	Importante	Não	Sim	Sim	Não	Sim
TDH – L1 (dias)	L1	-0,39	Importante	*	Não	Não	Sim	Sim
	L2	-0,25	Importante	Não	Não	Sim	Não	Sim
	L3	-0,17	Importante	Não	Sim	Sim	Não	Sim
	L4	-0,17	Importante	Não	Sim	Sim	Não	Não
TDH – L2 (dias)	L2	-0,35	Importante	Não	Não	Sim	Sim	Sim
	L3	-0,26	Importante	Não	Sim	Sim	Não	Sim
	L4	-0,25	Importante	Sim	Sim	Sim	Não	Sim
TDH – L3 (dias)	L3	-0,38	Importante	Sim	Não	Sim	Não	Sim
	L4	-0,36	Importante	Sim	Sim	Sim	Não	Sim
TDH – L4 (dias)	L4	-0,10	Importante	Não	Sim	Sim	Não	Sim
K <sub>b20</sub> – L1 (d <sup>-1</sup> )	L1	-0,51	Importante	*	Não	Não	Não	Sim
	L2	-0,32	Importante	Não	Não	Sim	Não	Sim
	L3	-0,24	Importante	Não	Sim	Sim	Não	Sim
	L4	-0,23	Importante	Sim	Sim	Sim	Não	Sim
K <sub>b20</sub> – L2 (d <sup>-1</sup> )	L2	-0,60	Importante	Não	Sim	Sim	Não	Sim
	L3	-0,48	Importante	Sim	Sim	Sim	Não	Sim
	L4	-0,42	Importante	Sim	Sim	Sim	Não	Sim
K <sub>b20</sub> – L3 (d <sup>-1</sup> )	L3	-0,48	Importante	Sim	Sim	Sim	Sim	Sim
	L4	-0,47	Importante	Sim	Sim	Sim	Sim	Sim
K <sub>b20</sub> – L4 (d <sup>-1</sup> )	L4	-0,16	Importante	Não	Sim	Sim	Não	Sim
D	L1	0,14	Importante	*	Não	Não	Não	Não
	L2	0,27	Importante	Não	Não	Não	Não	Não
	L3	0,36	Importante	Não	Sim	Não	Não	Sim
	L4	0,37	Importante	Sim	Sim	Sim	Não	Sim

\* o maior valor de concentração efluente de *E. coli* gerado pelas 10.000 rodadas da simulação Monte Carlo é menor que a referência adotada; (1) 10<sup>3</sup>, 10<sup>4</sup> ou 10<sup>5</sup> *E. coli*/100mL; 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades das concentrações efluentes estimadas. TDH: tempo de detenção hidráulica; K<sub>b20</sub>: coeficiente de decaimento de *E. coli* (20°C); d: número de dispersão.

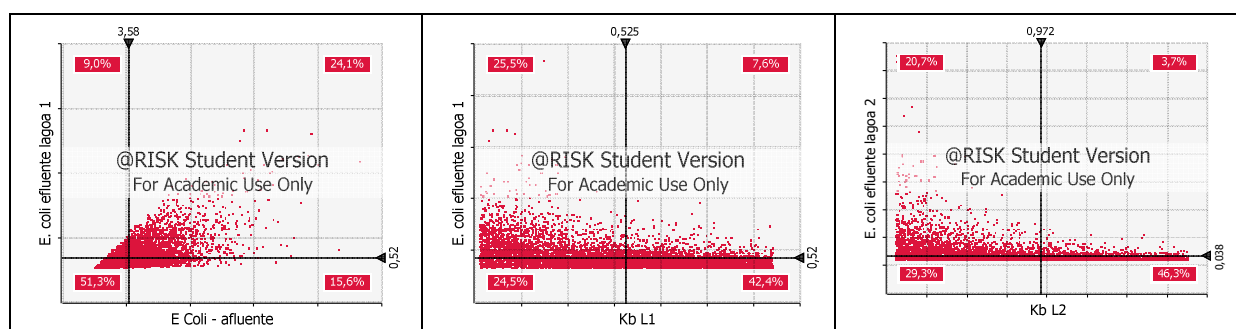


Figura 5.6 - Gráfico de dispersão, variável de entrada versus concentração efluente de *E. coli*.

Quadro 5.2 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para a concentração efluente de nitrogênio amoniacal

Entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 5	> 20	> 50%	> 95%
Nitrogênio amoniacal afluente (mg/L)	L1	0,86	Importante	*	Não	Sim	Sim
	L2	0,67	Importante	Não	Não	Sim	Sim
	L3	0,34	Importante	Não	Não	Não	Sim
	L4	0,31	Importante	Não	Sim	Não	Sim
pH – L1	L1	-0,43	Importante	*	Não	Não	Não
	L2	-0,34	Importante	Não	Não	Não	Não
	L3	-0,16	Importante	Não	Não	Não	Não
	L4	-0,15	Importante	Não	Não	Não	Não
pH – L2	L2	-0,55	Importante	Não	Não	Não	Sim
	L3	-0,27	Importante	Não	Não	Não	Não
	L4	-0,25	Importante	Não	Não	Não	Não
pH – L3	L3	-0,77	Importante	Não	Sim	Não	Sim
	L4	-0,70	Importante	Não	Não	Não	Não
pH – L4	L4	-0,32	Importante	Não	Sim	Não	Sim
Q/A – L1	L1	0,19	Importante	*	Não	Não	Não
	L2	0,15	Importante	Não	Não	Não	Não
	L3	0,07	Importante	Não	Não	Não	Não
	L4	0,07	Importante	Não	Não	Não	Não
Q/A – L2	L2	0,20	Importante	Não	Não	Não	Não
	L3	0,10	Importante	Não	Não	Não	Não
	L4	0,09	Importante	Não	Não	Não	Não
Q/A – L3	L3	0,25	Importante	Não	Não	Não	Sim
	L4	0,23	Importante	Não	Sim	Não	Sim
Q/A – L4	L4	0,16	Importante	Não	Sim	Não	Sim
Temperatura do ar (°C)	L1	0,04	Importante	*	Não	Não	Não
	L2	0,06	Importante	Não	Não	Não	Não
	L3	0,03	Importante	Não	Não	Não	Não
	L4	0,04	Importante	Não	Não	Não	Não

\* o maior valor de concentração efluente de nitrogênio amoniacal gerado pelas 10.000 rodadas da simulação Monte Carlo é menor que a referência adotada; (1) 5 ou 20 mg/L; 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades das concentrações estimadas. Q/A: taxa de aplicação hidráulica superficial.

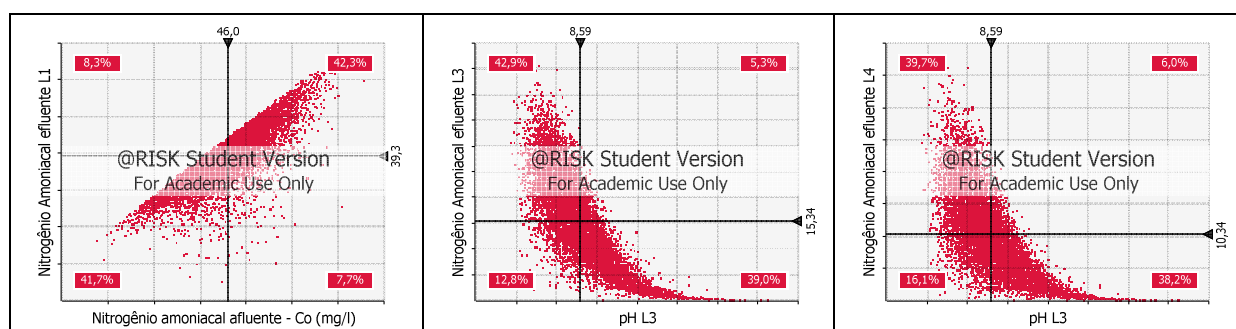


Figura 5.7 - Gráfico de dispersão, variável de entrada versus concentração de nitrogênio amoniacal efluente.

Para as faixas das variáveis de entrada consideradas nos modelos de remoção de *E. coli* e de amônia, a análise de sensibilidade com aplicação do teste de Mann-Whitney-Wilcoxon em torno da mediana mostrou-se pouco seletiva, ou seja, todas as variáveis de entrada foram classificadas como ‘importantes’.

Em relação à análise de sensibilidade por meio da classificação dos coeficientes de Spearman, a concentração de *E. coli* afluente e o coeficiente de decaimento bacteriano ( $K_{b20}$ ) se destacam, impactando mais fortemente a concentração efluente de *E. coli* em relação às demais variáveis de entrada.

A concentração afluente de *E. coli* apresentou correlação positiva com a concentração efluente, sendo decrescente, e de forma marcante, ao longo da série de lagoas. O coeficiente de decaimento bacteriano apresentou correlação negativa com a concentração efluente de *E. coli*. O coeficiente de decaimento de cada lagoa em relação ao seu efluente e aos efluentes das lagoas posteriores também apresentou tendência decrescente ao longo da trajetória do tratamento, todavia, de forma menos acentuada. Em outras palavras, quanto mais ao início da série de lagoas, mais fortemente essas duas variáveis impactam a qualidade do efluente.

É importante destacar o efeito mais pronunciado do coeficiente de decaimento da segunda lagoa em relação aos das demais lagoas, tanto no resultado de saída da própria Lagoa 2, quanto nos resultados de saída das lagoas posteriores. Por exemplo, o coeficiente de Spearman da relação entre o coeficiente de decaimento da Lagoa 1 e a qualidade do efluente a duas lagoas posteriores (Lagoa 3) é -0,24, enquanto que a relação entre o coeficiente de decaimento da lagoa 2 e a qualidade do efluente a duas lagoas posteriores (Lagoa 4) é -0,42.

O tempo de detenção hidráulica (TDH) e o número de dispersão apresentaram menores coeficientes de Spearman em relação às outras variáveis de entrada. O TDH de cada lagoa perde a importância sobre o resultado de saída (concentração de *E. coli*) ao longo da série de lagoas. Por outro lado, o número de dispersão ganha importância na determinação da concentração de *E. coli* efluente ao longo do tratamento.

De modo geral, a Análise de Cenários da previsão de concentração de *E. coli* para resultados extremos de saídas não se mostrou muito seletiva, pois quase todas as variáveis de entrada foram tidas como determinantes nos valores com percentil maior que 95% (à exceção do número de dispersão). Por outro lado, a análise em relação aos valores das medianas dos efluentes de cada lagoa não destacou qualquer variável de entrada como determinante.

Efluentes com  $> 10^3$  *E. coli* / 100 mL são mais frequentes nas primeiras lagoas da série e menos frequentes nas últimas. Por isso, algumas variáveis de entrada se mostraram

determinantes deste nível de qualidade de efluente apenas ao final da série de lagoas, com destaque para o tempo de detenção hidráulica e, principalmente, o coeficiente de decaimento bacteriano. De certa forma essa tendência foi confirmada para os cenários de efluentes com  $> 10^4$  e  $> 10^5$  *E. coli* / 100 mL. Por exemplo, a análise em relação à  $10^5$  *E. coli* / 100 mL forneceu resultados bem semelhantes aos do percentil maior que 95%: quase todas as variáveis de entrada foram consideradas determinantes desse nível de qualidade de efluente.

Em geral, pode-se também dizer que a Análise de Cenários e o teste de correlação de Spearman revelaram resultados equivalentes.

Em relação à análise de sensibilidade para a concentração de nitrogênio amoniacal efluente por meio da classificação dos coeficientes de Spearman, a concentração de afluente de nitrogênio amoniacal e o pH se destacam como as variáveis mais importantes.

A concentração afluente de nitrogênio amoniacal apresentou correlação positiva com a concentração efluente, sendo a correlação alta em relação aos efluentes das lagoas 1 e 2, mas bastante reduzida em relação aos efluentes das lagoas 3 e 4. Significa dizer que quanto maior a concentração de nitrogênio amoniacal afluente maior é sua influência (em si e em relação a outras variáveis) sobre a concentração efluente.

O pH apresentou correlação negativa com a concentração efluente, sendo esse resultado coerente com o reconhecido papel de valores elevados de pH na volatilização da amônia. Além disso, a importância do pH de determinada lagoa sobre a concentração de amônia em seu próprio efluente mostrou-se crescente ao longo das lagoas, até a Lagoa 3. Nesse sentido, o destaque é a própria Lagoa 3, cujos valores de pH impactaram mais fortemente a concentração efluente de amônia, inclusive a da Lagoa 4. Nas faixas de variação consideradas, a taxa de aplicação hidráulica (dados de operação do sistema experimental de lagoas) e a temperatura do ar (variações de temperaturas para o município de Viçosa) pouco impactaram os valores de saída da amônia.

Em geral, a Análise de Cenários não foi capaz de identificar variáveis determinantes na produção de efluentes com concentração  $> 5$  mg/L,  $> 20$  mg/L e com valores superiores às medianas das concentrações de amônia em cada lagoa. O baixo percentual de atendimento ao padrão de qualidade de 5 mg/L (Tabela 5.5), ou seja elevada frequência de valores  $> 5$  mg/L, ajuda a explicar os resultados em relação à este padrão. A ocorrência de valores  $> 20$  mg/L somente se viu determinada nas últimas lagoas da série (onde são raros) e por diversas variáveis (nitrogênio amoniacal afluente, pH e taxa de aplicação superficial).

De certa forma, a Análise de Cenários em relação aos percentis 50 e 95% revelou resultados similares ao da análise para *E. coli*: nenhuma variável pôde ser identificada como determinante na produção de efluentes com valores superiores aos das medianas, porém algumas o foram quando valores extremos (percentil 95%), portanto mais raros, foram tomados como referência. Nesse caso, destaca-se a concentração afluente de amônia (em todas as lagoas), o pH e a taxa de aplicação superficial, estes principalmente nas lagoas mais ao final da série. Percebe-se assim, uma vez mais, boa correspondência entre os resultados da Análise de Cenários e do teste de correlação de Spearman.

### **5.3 – Considerações finais, conclusões e recomendações**

Como referido na introdução deste capítulo, os objetivos aqui não eram a proposição de critérios de dimensionamento de lagoas com o uso de modelagem estocástica, mas a verificação do uso dessa ferramenta na avaliação de um sistema de tratamento específico e, em tal contexto, a identificação das variáveis de entrada de maior impacto na previsão da qualidade dos efluentes.

Em geral, a modelagem estocástica com o uso de Simulação de Monte Carlo parece ter super e subestimado os valores efluentes de, respectivamente, *E. coli* e amônia. Entretanto, como já destacado, isso precisaria ser confirmado com base em análise mais rigorosa, o que é perfeitamente possível com o banco de dados disponível.

Cabe também a ressalva de que apenas para poucas variáveis de entrada nos modelos de estimativa de qualidade de efluentes as respectivas incertezas foram expressas por distribuições de probabilidades (mais especificamente as concentrações afluentes de *E. coli* e de amônia e os valores de pH). Algumas variáveis operacionais (por exemplo, vazão e tempo de detenção hidráulica) foram consideradas dentro faixas relativamente estreitas, as de fato implementadas na operação do sistema de lagoas experimental e, juntamente com outras variáveis, de acordo com distribuição uniforme (opção devida à ausência de informações).

A própria Equação 3.27 de estimativa da temperatura do líquido em função da temperatura do ar poderia entrar no modelo considerando variações, por exemplo, por meio da determinação de intervalos de confiança em torno dos parâmetros da regressão linear (o banco de dados original permite isso). De forma análoga, os coeficientes de decaimento bacteriano foram assumidos como obedecendo à distribuição linear, mas o banco de dados original

permite melhor detalhamento. Enfim, o que se quer dizer é que a modelagem estocástica aqui desenvolvida pode e deve ser refinada.

Não obstante, com relação à previsão da concentração efluente de *E. coli* as seguintes inferências podem ser feitas, limitadas, porém às condições de contorno deste estudo:

- (i) Uma série de três lagoas parece ser suficiente e adequada para atender as recomendações da OMS quanto à qualidade do efluente para irrigação de hortaliças folhosas, raízes e tubérculo.
- (ii) Em geral, a análise de sensibilidade indicou que as variáveis de maior impacto na determinação da concentração efluente foram a concentração afluenta de *E. coli* e o coeficiente de decaimento bacteriano.
- (iii) Em relação à concentração afluenta de *E. coli*, o uso da distribuição lognormal, em lugar da distribuição uniforme, induziu valores mais baixos de concentrações efluentes, com diferenças significativamente diferentes ao nível de 5% de significância pelo teste Wilcoxon. Os resultados enfatizam a importância na determinação de uma distribuição de frequência que bem descreva a variável *E.coli*.
- (iv) Em relação ao coeficiente de decaimento bacteriano, os resultados sugerem a pertinência de melhor apropriação da variabilidade em torno desse parâmetro.
- (v) O tempo de detenção hidráulica e o número de dispersão revelaram pouco impacto na concentração efluente de *E. coli*. Em relação ao primeiro, é de se notar que os modelos de von Sperling (2002) conferem maior importância ao coeficiente de decaimento bacteriano do que ao tempo de detenção hidráulica. Quanto ao número de dispersão, cabe a ressalva de que esse parâmetro é de reconhecida importância, mas foi aqui considerado em faixa estreita de variação.

Com relação à previsão da concentração efluente de nitrogênio amoniacal cabem os seguintes destaques:

- (i) Uma série de três lagoas parece ser capaz de atender o padrão de lançamento de efluentes de 20 mg/L.
- (ii) Em geral, a análise de sensibilidade indicou que as variáveis de maior impacto na determinação da concentração efluente de amônia foram a concentração afluenta e pH, o primeiro mais ao início da série de lagoas, o segundo mais ao final.
- (iii) O uso da distribuição lognormal para a concentração afluenta de nitrogênio amoniacal e pH, em lugar da distribuição uniforme, induziu valores mais elevados de concentrações efluentes de nitrogênio amoniacal no final da série de 4 lagoas, Portanto,

assim como observado para a *E. coli*, os resultados enfatizam a importância de se determinar a distribuição de frequência da variável nitrogênio amoniacal.

(iv) No caso do pH, assume particular relevância o fato de que esta é uma variável de entrada nos modelos de Pano e Middlebrooks, porém desconhecida na etapa de projeto. Os resultados indicam então a necessidade de escolha criteriosa dos valores de pH para efeito do dimensionamento das lagoas, em particular de lagoas em série.

## 6 - ANÁLISE DE INCERTEZAS E DE SENSIBILIDADE NA AVALIAÇÃO QUANTITATIVA DE RISCO MICROBIOLÓGICO APLICADA AO USO AGRÍCOLA DE EFLUENTES DE LAGOAS DE POLIMENTO

Como visto na revisão de literatura (capítulo 3), vários autores têm recorrido à Avaliação Quantitativa de Risco Microbiológico (AQRM) e à modelagem estocástica envolvendo simulação de Monte Carlo para a estimativa de riscos associados à irrigação com esgotos sanitários.

Alguns trabalhos tomam como referência bancos de dados reais de monitoramento de sistemas de tratamento (TANAKA *et al.*, 1998; PETERSON *et al.*, 2001; HAMILTON *et al.*, 2006a)<sup>5</sup>. Outros recorrem a faixas genéricas de valores de *E. coli* possíveis de ocorrer em efluentes tratados (MARA *et al.*, 2007).

A maioria dos trabalhos aborda a produção e consumo, isoladamente, de culturas específicas, com destaque para alface, mas incluindo, em poucos casos, outras culturas tais como cebola (MARA *et al.*, 2007), brócolis, pepino e repolho (HAMILTON *et al.*, 2006a). Diversos trabalhos assumem como cenário hipotético - e desfavorável - a irrigação por aspersão; outros exploram comparações entre métodos de irrigação (VAN GINNEKEN e ORON, 2000).

Entretanto, a grande maioria dos trabalhos publicados é baseada em estimativas de contaminação das culturas irrigadas a partir de pressupostos de volumes de água remanescentes sobre a superfície das culturas após a irrigação. Poucos trabalhos utilizam dados de medidas reais de contaminação como, por exemplo, o de Seidu *et al.* (2000), em Gana, sobre a contaminação de alfaces irrigadas manualmente (com regadores).

Quanto ao padrão de consumo de culturas irrigadas, muitos trabalhos recorrem a meras estimativas, em geral assumindo os pressupostos dos cenários de exposição das Diretrizes da OMS – 100 g de alface a cada dois dias. Outros recorrem a dados sistematizados nos EUA, embora os trabalhos publicados nem sempre tenham origem naquele país (ex: HAMILTON *et al.*, 2006a, HAMILTON *et al.*, 2006b).

Por fim, como também já registrado no capítulo 3, a maioria dos trabalhos tende a enfatizar riscos associados aos rotavírus, *Campylobacter* e *Cryptosporidium* (principalmente o

---

<sup>5</sup> De fato todos esses trabalhos utilizam a mesma fonte de dados – ocorrência de vírus em efluentes secundários (ASANO *et al.*, 1992; TANAKA *et al.*, 1998)

primeiro), tendo sido estes os organismos tomados como referência na atualização das Diretrizes da OMS (WHO, 2006).

Como especificidades do presente trabalho encontram-se: (i) a utilização de um vasto banco de dados de monitoramento (cerca de sete anos) de efluentes de lagoas de polimento, o que permitiu, com alguma segurança, testar a aderência dos dados às distribuições de probabilidades; (ii) a utilização de dados de monitoramento de contaminação de culturas diversas irrigadas manualmente (com regadores), o que possibilitou o ajuste de modelos de estimativa de contaminação em função da qualidade da água de irrigação; (iii) a utilização de estatísticas de consumo de hortaliças no Brasil. Em resumo, a estimativa de risco foi baseada em cenários hipotéticos, porém, mais próximos à realidade brasileira.

## 6.1 – Metodologia

A estimativa de riscos foi realizada com a utilização do programa @Risk 5.5 (PALISADE CORPORATION). Os modelos de exposição e de dose-resposta e as incertezas relativas às respectivas variáveis de entrada foram inseridos no referido programa, o qual realiza modelagem estocástica por meio de simulação aleatória por amostragem por ‘Hipercubos Latinos’ (ver item 3.1.3). A simulação envolveu 10.000 iterações.

Como organismos-referência foram selecionados um vírus, uma bactéria e um protozoário, a saber, rotavírus, *Campylobacter* e *Cryptosporidium*.

O cenário de exposição explorado foi o de irrigação irrestrita, ou seja, de riscos de infecção decorrentes da exposição ao consumo de hortaliças ingeridas cruas.

### 6.1.1 - Modelos dose-resposta e caracterização do risco

Os modelos dose-resposta utilizados foram o modelo beta-Poisson (Equação 3.41) para infecção por rotavírus e *Campylobacter* e o modelo exponencial (Equação 3.42) para infecção por *Cryptosporidium*.

As faixas de valores utilizadas para dose infectante média ( $N_{50}$ ) e os parâmetros característicos da interação agente-hospedeiro ( $\alpha$  e  $r$ ) estão descritos na Tabela 6.1. Foram adotados como valores médios aqueles consagrados na literatura (HASS *et al.*, 1999) e, assim como proposto por Mara *et al.* (2007), variações de  $\pm 25\%$ . Por não se ter informações

suficientes que conduzissem às respectivas distribuições de probabilidades destes parâmetros, adotou-se a distribuição uniforme para os mesmos.

TABELA 6.1 – Parâmetros dos modelos beta-Poisson e exponencial utilizados no modelo de AQRM, valores médios, máximos, mínimos e distribuição de probabilidades adotada.

Parâmetro	Organismo	Distribuição	Valor médio	Mínimo	Máximo
$N_{50}$	<i>Campylobacter</i>	Uniforme	896	672	1120
	<i>Rotavírus</i>	Uniforme	6,17	4,63	7,71
$\alpha$	<i>Campylobacter</i>	Uniforme	0,145	0,11	0,18
	<i>Rotavírus</i>	Uniforme	0,253	0,190	0,316
$r$	<i>Cryptosporidium</i>	Uniforme	0,0042	0,00315	0,00525

$N_{50}$ : dose infectante média;  $\alpha$ : parâmetro característico da interação agente-hospedeiro para *Campylobacter* e rotavírus;  $r$ : parâmetro característico da interação agente-hospedeiro para *Cryptosporidium*

As doses de entrada nos modelos dose-resposta foram estimadas de acordo com o descrito no item a seguir para efeito do cálculo dos riscos diários. Os riscos anuais foram estimados com base na Equação 3.43.

### 6.1.2 - Modelo de exposição

A dose de patógenos  $d$  ingerida a cada evento de exposição foi estimada pela Equação 6.1.

$$d = EC \cdot C \cdot \frac{N}{10^5} \cdot 10^{-D} \quad (6.1)$$

Onde:

$N$  = número de patógenos por  $10^5$  *E. coli*;

$EC$  = concentração de *E. coli* por grama de hortaliça no momento da colheita;

$10^{-D}$  = redução de patógenos entre a colheita e o consumo das hortaliças e;

$C$  = consumo de hortaliças.

A concentração de *E. coli* por grama de hortaliça irrigada ( $EC$ ) foi obtida por meio das Equações 3.44 e 3.45 de Bastos *et al.* (2008), respectivamente para hortaliças que se desenvolvem rentes ao solo (mais especificamente, alface, espinafre e rúcula) e distantes do solo (couve e pimentão). Nos experimentos que geraram essas equações, as hortaliças foram regadas manualmente. Cumpre esclarecer que esses experimentos foram realizados com os efluentes das lagoas aqui em estudo, durante os períodos operacionais 1 a 3 (ver Tabela 4.2)

Para os valores de  $N$  e  $D$  da Equação 6.1, foram assumidas as mesmas faixas de valores adotadas nas Diretrizes da OMS, as quais têm origem em dados de literatura (WHO, 2006) (Tabela 6.2). No valor de  $D$ , está implícita a redução do número de patógenos por decaimento natural no ambiente (superfície das culturas e no solo) e, ou por higiene dos alimentos, essencialmente, lavagem. Novamente, por não se dispor de informações suficientes, foi assumido que os valores de  $N$  e  $D$  variassem de acordo com distribuição uniforme.

TABELA 6.2 – Valores de número de patógenos por  $10^5$  *E. coli* e de redução de patógenos entre a colheita e o consumo utilizados no modelo de AQRM, valores máximos e mínimos e distribuição de probabilidades adotada.

Parâmetro	Patógeno	Distribuição	Mínimo	Máximo
N	<i>Campylobacter</i>	Uniforme	0,1	1
	Rotavírus	Uniforme	0,1	1
	<i>Cryptosporidium</i>	Uniforme	0,01	0,1
D	<i>Campylobacter</i>	Uniforme	2	3
	Rotavírus	Uniforme	2	3
	<i>Cryptosporidium</i>	Uniforme	0	1

N: número de patógenos por  $10^5$  *E. Coli*; D: redução de patógenos entre a colheita e o consumo

Em relação à qualidade dos efluentes das lagoas de polimento, que conduz à qualidade das hortaliças irrigadas por meio das Equações 3.44 e 3.45, para cada lagoa foi selecionada a série com maior número de dados. Para efeito da estimativa de riscos, não foram considerados os períodos em que as lagoas 3 e 4 foram operadas em paralelo. De acordo com os resultados do capítulo 4, as concentrações de *E. coli* nos efluentes das lagoas foram inseridos no modelo obedecendo a distribuição lognormal. Para efeito de comparação, foram também estimados riscos assumindo que os dados de *E. coli* apresentassem distribuição uniforme, tal como se encontra em diversos trabalhos publicados (ex.: MARA *et al.*, 2007).

A Tabela 6.3 apresenta um resumo dos dados de qualidade dos efluentes inseridos no modelo de AQRM. Cumpre esclarecer que o programa utilizado (@Risk) calcula a média e o desvio-padrão para a distribuição lognormal a partir do fornecimento das informações de três percentis selecionados (no caso, os percentis 5% 50% e 95%).

TABELA 6.3 - Concentração de *E. coli* nos efluente das lagoas de polimento nas séries de dados selecionadas para utilização no modelo de AQRM, medianas e valores-limite (percentis) para a caracterização da distribuição lognormal.

Série	Percentil ( <i>E. coli</i> / 100 mL)		
	5%	50%	95%
L1 B	$6,63 \times 10^3$	$1,60 \times 10^5$	$5,20 \times 10^5$
L2 B	$1,91 \times 10^2$	$3,25 \times 10^3$	$1,49 \times 10^4$
L3 B	$2,00 \times 10^0$	$3,10 \times 10^1$	$6,86 \times 10^2$
L4 A	$1,00 \times 10^1$	$1,00 \times 10^2$	$1,89 \times 10^3$

L1 B: série B relativa à lagoa 1, L2 B: série B relativa à lagoa 2, L3 B: série B relativa à lagoa 3, L4 A: série A relativa à lagoa 4

O cenário de exposição foi complementado com dados de consumo de hortaliças fornecidos pela Pesquisa de Orçamentos Familiares (POF) de 2002-2003 (IBGE, 2004). A POF fornece dados de aquisição alimentar domiciliar *per capita* para o Brasil como um todo, por grandes regiões e por estado e, nos dois primeiros casos, por classe de renda. São ainda fornecidos coeficientes de variação para os dados do Brasil como um todo e por regiões, mas não para os dados por classe de renda. Em vista de incertezas ou insuficiências de informações, e também por praticidade, optou-se por utilizar os dados do Brasil como um todo e por assumir que estes seguissem distribuição uniforme, muito embora em análises mais detalhadas se possa no futuro procurar testar a aderência a outras distribuições de probabilidades. Os dados da POF são fornecidos em kg/ano e para efeito de cômputo do *per capita* diário os valores originais foram divididos pelo número de dias no ano (365).

Dentre os dados da POF, foram selecionados os relativos às seguintes hortaliças ingeridas cruas:

- *Hortaliças folhosas e florais* - couve, couve-brócolis, acelga, agrião, alface, cheiro-verde, couve-flor, repolho e outras. Os dados de consumo destas hortaliças foram computados em conjunto (ou seja, foram somados) no modelo de AQRM e foram associados à Equação 3.44 relativa às hortaliças que se desenvolvem rentes ao solo.
- *Hortaliças frutais* - pimentão, pepino e tomate. Estes dados de consumo também foram computados em conjunto, sendo associados à Equação 3.45 relativa às hortaliças que se desenvolvem distantes do solo.

Cumpre, porém, registrar pequenas incongruências nos critérios assumidos: a Equação 3.44 foi obtida com dados de irrigação de alface, espinafre e rúcula e os dados do IBGE para ‘hortaliças folhosas e florais’ incluem couve, que não vem a ser planta que se desenvolve

rente ao solo. De forma análoga, a Equação 3.45 foi obtida com dados de irrigação de couve e pimentão, mas couve não se enquadra em ‘hortaliças frutais’ e, por sua vez, pepino não necessariamente se desenvolve distante do solo. Entretanto, entende-se que estes ‘detalhes’ não venham a comprometer as estimativas de riscos realizadas, uma vez que as diferenças ao se incluir ou não as ‘incongruências’ pouco impactariam o consumo total de cada categoria de hortaliças considerada.

A Tabela 6.4 resume os dados de consumo de hortaliças considerados para a estimativa de riscos.

TABELA 6.4 - Consumo anual de hortaliças ingeridas cruas no Brasil, valores médios, mínimos e máximos utilizados para a caracterização da distribuição uniforme e para utilização no modelo de AQRM

	<b>Hortaliças</b>	<b>Distribuição</b>	<b>Média (Kg/pessoa)</b>	<b>Mínimo</b>	<b>Máximo</b>
<b>Folhosas e florais</b>	Couve	Uniforme	0,127	0,095	0,159
	Couve-brócolis	Uniforme	0,108	0,081	0,135
	Acelga	Uniforme	0,070	0,053	0,088
	Agrião	Uniforme	0,082	0,062	0,103
	Alface	Uniforme	0,643	0,482	0,804
	Cheiro-verde	Uniforme	0,146	0,110	0,183
	Couve-flor	Uniforme	0,175	0,131	0,219
	Repolho	Uniforme	0,993	0,745	1,241
	Outras	Uniforme	0,156	0,117	0,195
<b>Frutais</b>	Pepino	Uniforme	0,484	0,363	0,606
	Tomate	Uniforme	5,000	3,750	6,250
	Pimentão	Uniforme	0,616	0,462	0,770

Fonte: IBGE (2004)

Os resultados são expostos por meio de gráficos com histogramas de frequência e distribuição acumulada (Figura 6.1).

### 6.1.3 - Análises de sensibilidade e de cenário

A Análise de Sensibilidade foi realizada como descrito no capítulo 5, utilizando gráficos de dispersão, medidas de correlação não paramétrica de Spearman, análise em relação à mediana e Análise de Cenários.

A análise de cenários foi realizada tomando como referência valores de saída (risco de infecção) menores que  $10^{-2}$  ou  $10^{-3}$  por pessoa por ano (pppa). O risco de infecção de  $10^{-3}$  pppa é o valor de referência assumido nas Diretrizes da OMS (WHO, 2006) como risco tolerável para infecção por rotavírus e corresponde à um valor de DALY tolerável de  $10^{-6}$  pppa (ver item 3.4.1). Entretanto, Mara *et al.* (2007) consideram este valor como muito rigoroso e propõem  $10^{-2}$  como risco tolerável. Por isso, este valor foi também considerado. A Análise de Cenário também foi realizada tomando como referências os valores das medianas e dos percentis 95% das distribuições de probabilidades das estimativas de riscos.

## 6.2 - Resultados e discussão

### 6.2.1 - Estimativas de risco de infecção

Nas Tabelas 6.5 e 6.6 encontram-se os valores das medianas e dos percentis 95% para os riscos anuais de infecção, calculados de acordo com os modelos de AQRM descritos no item anterior, considerando a qualidade dos efluentes das lagoas (*E. coli*/100 ml) segundo distribuição lognormal. Nas Figuras 6.1 e 6.2 apresentam-se os histogramas com as distribuições de probabilidades de riscos anuais, respectivamente para os cenários de exposição de consumo de hortaliças que se desenvolvem rentes ao e distantes do nível do solo.

TABELA 6.5 - Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo e que são ingeridas cruas, qualidade dos efluentes das lagoas segundo distribuição lognormal

Lagoa*	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	50%	95%	50%	95%	50%	95%
<b>L1</b>	$2,76 \times 10^{-3}$	$2,13 \times 10^{-2}$	$8,01 \times 10^{-2}$	$4,08 \times 10^{-1}$	$5,39 \times 10^{-3}$	$3,26 \times 10^{-2}$
<b>L2</b>	$1,10 \times 10^{-4}$	$9,81 \times 10^{-4}$	$3,35 \times 10^{-3}$	$2,38 \times 10^{-2}$	$2,18 \times 10^{-4}$	$1,51 \times 10^{-3}$
<b>L3</b>	$2,35 \times 10^{-6}$	$5,30 \times 10^{-5}$	$7,16 \times 10^{-5}$	$1,33 \times 10^{-3}$	$4,76 \times 10^{-6}$	$9,05 \times 10^{-5}$
<b>L4</b>	$6,49 \times 10^{-6}$	$1,28 \times 10^{-4}$	$1,90 \times 10^{-4}$	$3,28 \times 10^{-3}$	$1,25 \times 10^{-5}$	$2,16 \times 10^{-4}$

\* qualidade dos efluentes de acordo com a Tabela 6.3

TABELA 6.6 – Estimativas de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo e que são ingeridas cruas, qualidade dos efluentes das lagoas segundo distribuição lognormal.

Lagoa*	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	50%	95%	50%	95%	50%	95%
<b>L1</b>	$8,84 \times 10^{-5}$	$5,86 \times 10^{-4}$	$2,66 \times 10^{-3}$	$1,33 \times 10^{-2}$	$1,74 \times 10^{-4}$	$8,38 \times 10^{-4}$
<b>L2</b>	$1,55 \times 10^{-5}$	$1,04 \times 10^{-4}$	$4,74 \times 10^{-4}$	$2,49 \times 10^{-3}$	$3,05 \times 10^{-5}$	$1,55 \times 10^{-4}$
<b>L3</b>	$1,99 \times 10^{-6}$	$1,86 \times 10^{-5}$	$5,92 \times 10^{-5}$	$4,40 \times 10^{-4}$	$3,98 \times 10^{-6}$	$2,88 \times 10^{-5}$
<b>L4</b>	$3,44 \times 10^{-6}$	$3,04 \times 10^{-5}$	$1,02 \times 10^{-4}$	$7,59 \times 10^{-4}$	$6,77 \times 10^{-6}$	$4,68 \times 10^{-5}$

\* qualidade dos efluentes de acordo com a Tabela 6.3

Como referido no item 3.4.1, a OMS assume como carga de doença tolerável  $10^{-6}$  DALY pppa, o que corresponde a níveis de risco toleráveis anuais de  $7,7 \times 10^{-4}$  pppa para rotavírus,  $3,1 \times 10^{-4}$  pppa para *Campylobacter* e  $2,2 \times 10^{-3}$  pppa para *Cryptosporidium*. Daí o ‘arredondamento’ para  $10^{-3}$  pppa como risco tolerável de referência (ver item 3.4.1).

De acordo com os cenários de exposição considerados, o consumo de hortaliças que se desenvolvem rentes ao nível do solo (cenário mais desfavorável) resultaria em riscos maiores que os ‘níveis toleráveis’ nas seguintes situações:

- (i) *Campylobacter*: lagoa 2, apenas o percentil 95%; lagoa 1, tanto a mediana, quanto o percentil 95%;
- (ii) Rotavírus: todas as lagoas para o percentil 95%; lagoas 1 e 2 tanto as medianas, quanto os percentis 95%;
- (iii) *Cryptosporidium*: lagoa 1, tanto a medianas, quanto o percentil 95%.

No que diz respeito aos riscos associados ao consumo de hortaliças que se desenvolvem distantes do solo, os níveis de risco toleráveis seriam superados nas seguintes situações:

- (i) *Campylobacter*: lagoa 1, apenas o percentil 95%;
- (ii) Rotavírus: lagoa 2, apenas o percentil 95%; lagoa 1, tanto a mediana, quanto o percentil 95%.

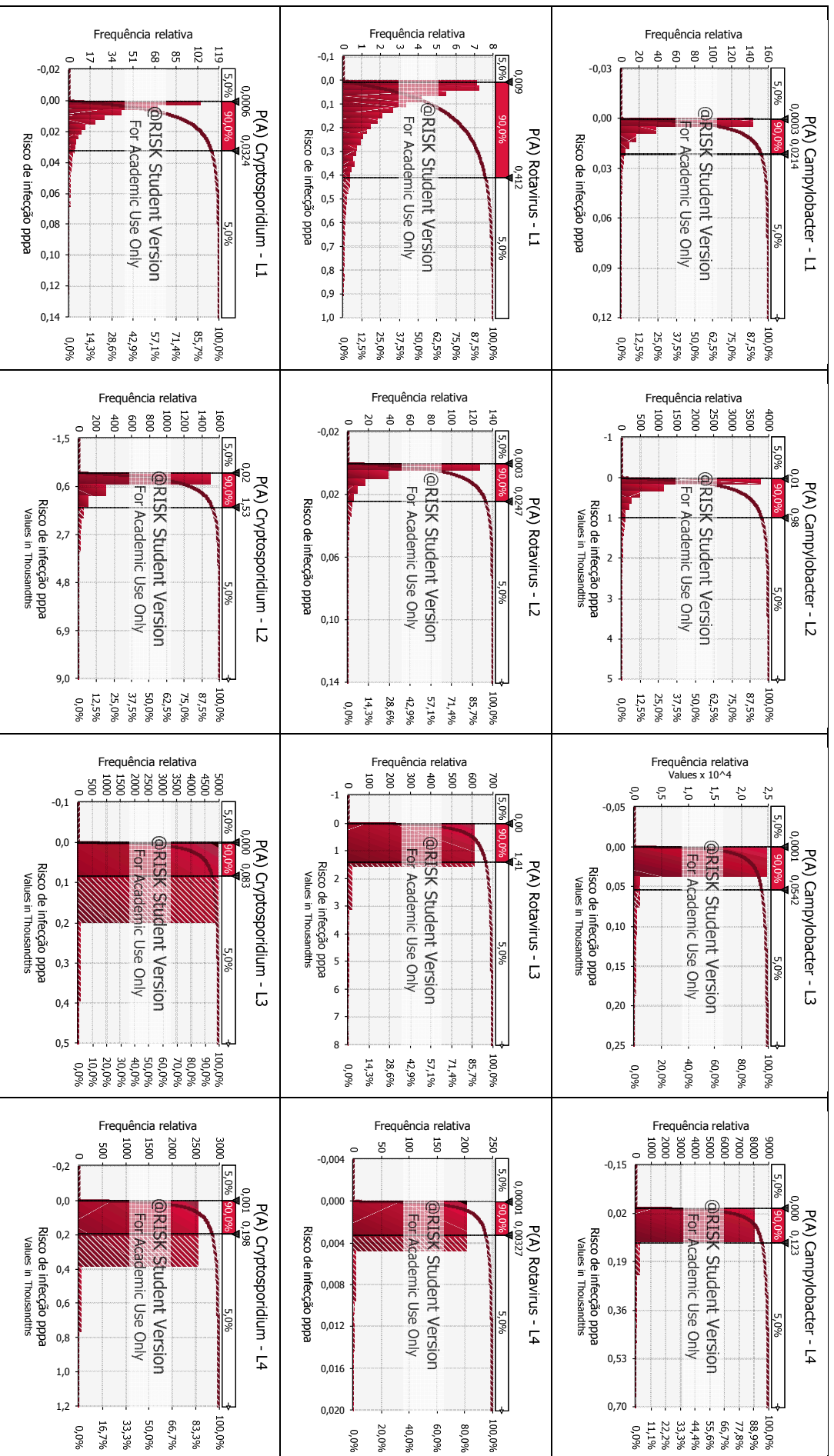


Figura 6.1 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição lognormal

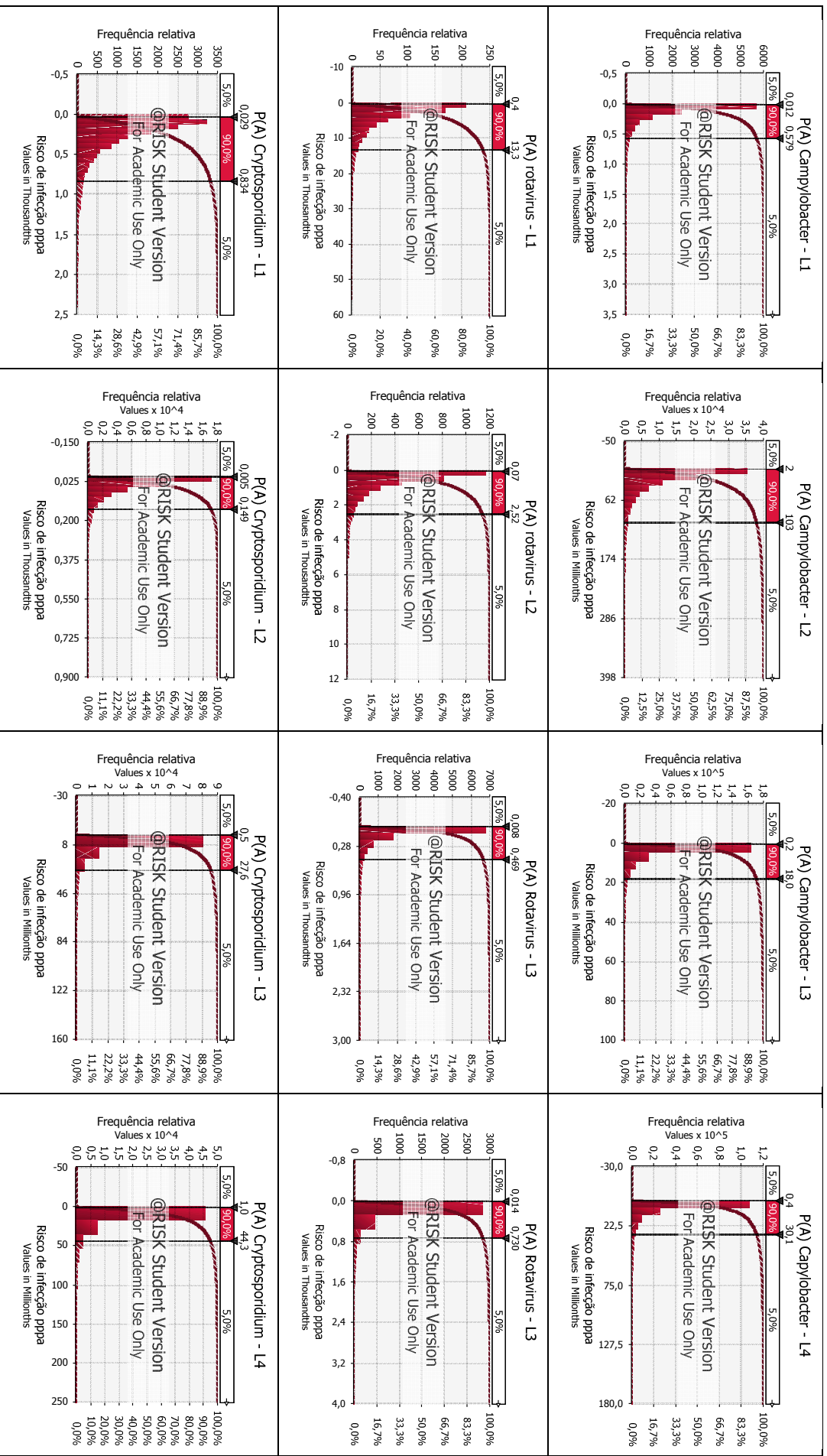


Figura 6.2 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem distantes do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição lognormal.

Observa-se nas Figuras 6.1 e 6.2 que, em geral, as estimativas de risco apresentaram distribuições com grande concentração de dados à esquerda do histograma, o que é condizente com a distribuição lognormal assumida para a concentração de *E. coli* nos efluentes das lagoas.

As estimativas de riscos de infecção por rotavírus apresentaram-se sempre em valores maiores que as estimativas para *Campylobacter* e *Cryptosporidium*, o que é coerente com as características de infectividade dos rotavírus. Na revisão de literatura já foi destacado que por esta razão a maioria dos estudos de AQRM enfatiza os riscos associados aos rotavírus e que este foi o organismo tomado como referência na formulação das Diretrizes da OMS para o uso agrícola de esgotos.

Em análise mais detalhada, as Tabelas 6.7 e 6.8 mostram, respectivamente, os percentuais de atendimento aos níveis de risco de infecção tomados como referência de risco tolerável  $\leq 10^{-2}$  e  $\leq 10^{-3}$  pppa (ver item 6.1.3). Nota-se que o cenário de exposição de consumo de hortaliças irrigadas com os efluentes das lagoas 3 e 4 (mediana das qualidades dos efluentes de  $3,1 \times 10^1$  e  $1,0 \times 10^2$  *E. coli*/100 mL) resultaria em riscos abaixo dos níveis toleráveis em frequência bem elevada, quase integral. Os dados de qualidade dos efluentes dessas lagoas, e os níveis correspondentes de risco confirmam o já observado por Rios (2007) de que uma série com três lagoas seria suficiente para o alcance de efluente final com qualidade adequada para a irrigação irrestrita; ou seja, a inclusão de uma quarta lagoa seria desnecessária.

A irrigação com o efluente da Lagoa 2 (mediana da qualidade do efluente de  $3,25 \times 10^3$  *E. coli*/100 mL), resultaria também em elevado percentual de atendimento em relação aos riscos de infecção por *Campylobacter* e *Cryptosporidium* em quaisquer dos cenários considerados, mas nem sempre em relação aos riscos associados aos rotavírus.

Entretanto, a irrigação com o efluente da Lagoa 1 (mediana da qualidade do efluente de  $1,60 \times 10^5$  *E. coli*/100 mL) resultaria, frequentemente, em níveis acima dos riscos toleráveis, com destaque para a referência mais exigente ( $10^{-3}$  pppa), para a irrigação de hortaliças que se desenvolvem rentes ao solo e para infecção por rotavírus.

As diretrizes da OMS fixam os seguintes valores de qualidade de efluentes, tendo como referência uma carga de doença tolerável  $\leq 10^{-6}$  DALY, ou aproximadamente um risco tolerável de infecção por rotavírus de  $10^{-3}$  pppa (ver item 3.4.1): (i)  $\leq 10^3$  *E. coli*/100 mL para o cultivo de raízes e tubérculos e para irrigação localizada de plantas que se desenvolvem rentes ao nível do solo; (ii)  $\leq 10^4$  *E. coli*/100 mL para o cultivo de folhosas; (iii)  $\leq 10^5$

*E. coli*/100 mL para a irrigação localizada de plantas que se desenvolvem distantes do nível do solo (WHO, 2006).

A qualidade do efluente da Lagoa 2 manteve-se abaixo de  $3,25 \times 10^3$  e  $1,49 \times 10^4$  *E. coli* /100 mL em, respectivamente, 50% e 95% do tempo (Tabela 6.1), mas os riscos de infecção por rotavírus correspondentes ao cenário de exposição de consumo de hortaliças folhosas e florais estiveram abaixo de  $10^{-3}$  pppa apenas em torno de 20% do tempo. Porém, neste mesmo cenário, o percentual de atendimento ao nível de risco de  $10^{-2}$  pppa foi bem mais elevado (cerca de 75%). Por um lado, esses resultados poderiam suscitar dúvidas em relação ao padrão de  $10^4$  *E. coli* /100 mL para a irrigação de hortaliças folhosas e florais que se desenvolvem rentes ao solo, mas, por outro, bem destacam a importância de discussão sobre o que venha a ser risco tolerável.

Por sua vez, o cenário de irrigação de hortaliças folhosas e florais com o efluente da lagoa 4 ( $\leq 1,00 \times 10^2$  *E. coli*/100 mL e  $\leq 1,89 \times 10^3$  *E. coli*/100 mL em, respectivamente, 50 e 95% do tempo), resultou em atendimento do nível de risco de  $10^{-3}$  pppa em mais de 80% do tempo. Assim sendo, um padrão de  $10^3$  *E. coli*/100 mL parecer ser mais adequado para a irrigação de hortaliças folhosas e florais que se desenvolvem rentes ao solo.

Em relação à irrigação de hortaliças frutais (ou que se desenvolvem distantes do nível do solo), o uso do efluente da Lagoa 1 ( $\leq 1,60 \times 10^5$  *E. coli*/100 mL e  $\leq 5,20 \times 10^5$  *E. coli*/100 mL em, respectivamente, 50 e 95% do tempo) resultaria em riscos de infecção por rotavírus superiores à  $10^{-3}$  pppa em elevada frequência (percentual de atendimento de cerca de 20%). Entretanto, o uso do efluente da Lagoa 2 ( $\leq 3,25 \times 10^3$  *E. coli*/100 mL e  $\leq 1,49 \times 10^4$  *E. coli*/100 mL em, respectivamente, 50 e 95% do tempo) resultaria em elevado percentual de atendimento ao nível de risco de  $10^{-3}$  pppa (cerca de 75%). Isso poderia sinalizar que um padrão de  $10^4$  *E. coli*/100 mL talvez fosse mais adequado para a irrigação de hortaliças que se desenvolvem distantes do nível do solo (à exceção naturalmente da irrigação por aspersão). Porém, novamente, cabe destacar a importância da discussão em torno do conceito de risco tolerável, já que mesmo com o efluente da Lagoa 1, o atendimento ao nível de risco de  $10^{-2}$  pppa foi bastante elevado (cerca de 90%).

Por fim, cabe destacar que todas essas observações devem ser relativizadas, tendo em vista que os cenários de exposição considerados aqui e nas diretrizes da OMS guardam diferenças importantes (ver item 6.4).

TABELA 6.7 – Percentuais de atendimentos ao nível tolerável de risco de infecção de  $10^{-2}$  pppa, cenário de exposição ao consumo de hortaliças irrigadas com efluentes de lagoas de polimento.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	RS	DS	RS	DS	RS	DS
1	83,90%	100%	5,90%	89,90%	70,00%	100%
2	100%	100%	80,90%	99,90%	99,90%	100%
3	100%	100%	99,70%	100%	100%	100%
4	100%	100%	98,90%	100%	100%	100%

RS: hortaliças que se desenvolvem rentes ao solo; DS: hortaliças que se desenvolvem distantes do solo.

TABELA 6.8 – Percentuais de atendimentos ao nível tolerável de risco de infecção de  $10^{-3}$  pppa, cenário de exposição ao consumo de hortaliças irrigadas com efluentes de lagoas de polimento.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	RS	DS	RS	DS	RS	DS
1	23,20%	98,70%	0,20%	18,40%	9,20%	96,90%
2	95,20%	100%	18,70%	75,70%	89,70%	100%
3	99,90%	100%	92,90%	99,10%	99,90%	100%
4	99,70%	100%	83,10%	97,10%	99,50%	100%

RS: hortaliças que se desenvolvem rentes ao solo; DS: hortaliças que se desenvolvem distantes do solo.

Para fins de comparação com o uso da distribuição lognormal, foram realizadas estimativas de risco utilizando a distribuição uniforme na concentração de *E. coli* efluente. Os resultados são apresentados nas Tabelas 6.9 e 6.10 e Figuras 6.3 e 6.4.

TABELA 6.9 - Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	Mediana	95%	Mediana	95%	Mediana	95%
L1	$3,98 \times 10^{-3}$	$2,98 \times 10^{-2}$	$1,13 \times 10^{-1}$	$4,93 \times 10^{-1}$	$8,04 \times 10^{-3}$	$4,13 \times 10^{-2}$
L2	$2,05 \times 10^{-4}$	$1,51 \times 10^{-3}$	$6,19 \times 10^{-3}$	$3,56 \times 10^{-2}$	$4,17 \times 10^{-4}$	$2,22 \times 10^{-3}$
L3	$1,57 \times 10^{-5}$	$1,14 \times 10^{-4}$	$4,77 \times 10^{-4}$	$2,79 \times 10^{-3}$	$3,21 \times 10^{-5}$	$1,70 \times 10^{-4}$
L4	$3,70 \times 10^{-6}$	$2,68 \times 10^{-4}$	$1,11 \times 10^{-3}$	$6,52 \times 10^{-3}$	$7,55 \times 10^{-5}$	$3,97 \times 10^{-4}$

TABELA 6.10 – Estimativa de riscos anuais de infecção (medianas e percentis 95%) decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	Mediana	95%	Mediana	95%	Mediana	95%
<b>L1</b>	$1,03 \times 10^{-4}$	$6,52 \times 10^{-4}$	$3,09 \times 10^{-3}$	$1,53 \times 10^{-2}$	$9,29 \times 10^{-4}$	$1,76 \times 10^{-4}$
<b>L2</b>	$2,06 \times 10^{-5}$	$1,29 \times 10^{-4}$	$6,19 \times 10^{-4}$	$3,11 \times 10^{-3}$	$1,88 \times 10^{-4}$	$3,10 \times 10^{-5}$
<b>L3</b>	$5,14 \times 10^{-6}$	$3,15 \times 10^{-5}$	$1,53 \times 10^{-4}$	$7,97 \times 10^{-4}$	$4,62 \times 10^{-5}$	$1,05 \times 10^{-5}$
<b>L4</b>	$8,08 \times 10^{-6}$	$5,08 \times 10^{-5}$	$2,42 \times 10^{-4}$	$1,24 \times 10^{-3}$	$7,36 \times 10^{-5}$	$6,81 \times 10^{-6}$

O uso da distribuição lognormal resultou em estimativas menores de risco para todos os efluentes estudados. A redução ocorrida em termos de unidades logarítmicas foi praticamente a mesma para os três patógenos analisados (Tabelas 6.11 e 6.12). Observa-se também que a redução dos riscos associados ao consumo de hortaliças com desenvolvimento rente ao solo foi maior que a associada com as hortaliças com desenvolvimento distante do solo. Outra constatação é que até a terceira lagoa, a diferença entre o uso da distribuição uniforme e o uso da lognormal na estimativa de risco aumenta gradativamente, reduzindo um pouco na quarta lagoa. No caso de hortaliças folhosas e florais (desenvolvimento rente ao solo), e quando consideradas as medianas referentes à Lagoa 3, essa diferença alcançou valores próximos a uma ordem de grandeza.

As diferenças das estimativas de risco de infecção obtidas por meio do uso da distribuição lognormal e da distribuição uniforme foram verificadas estatisticamente por meio do teste não-paramétrico para dados pareados Wilcoxon ao nível de 5% de significância. As diferenças foram significativas em todas as lagoas e para todos os patógenos, tanto no caso de hortaliças com desenvolvimento rente ao solo como distantes do solo.

TABELA 6.11 – Diferença em unidades logarítmicas (mediana e percentil 95%) entre o risco de infecção anual estimado com a concentração de *E. coli* nos efluentes das lagoas segundo distribuição uniforme e lognormal - hortaliças que se desenvolvem rentes ao nível do solo.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	Mediana	95%	Mediana	95%	Mediana	95%
L1	0,16	0,12	0,15	0,08	0,17	0,10
L2	0,27	0,19	0,27	0,17	0,28	0,17
L3	0,83	0,33	0,82	0,32	0,83	0,27
L4	0,76	0,32	0,77	0,30	0,78	0,26

TABELA 6.12 – Diferença em unidades logarítmicas (mediana e percentil 95%) entre o risco de infecção anual estimado com a concentração de *E. coli* nos efluentes das lagoas segundo distribuição uniforme e lognormal - hortaliças que se desenvolvem distantes do nível do solo.

Lagoa	<i>Campylobacter</i>		Rotavírus		<i>Cryptosporidium</i>	
	Mediana	95%	Mediana	95%	Mediana	95%
L1	0,07	0,05	0,06	0,06	0,07	0,04
L2	0,12	0,09	0,12	0,10	0,13	0,08
L3	0,41	0,23	0,41	0,26	0,41	0,20
L4	0,37	0,22	0,38	0,21	0,37	0,20

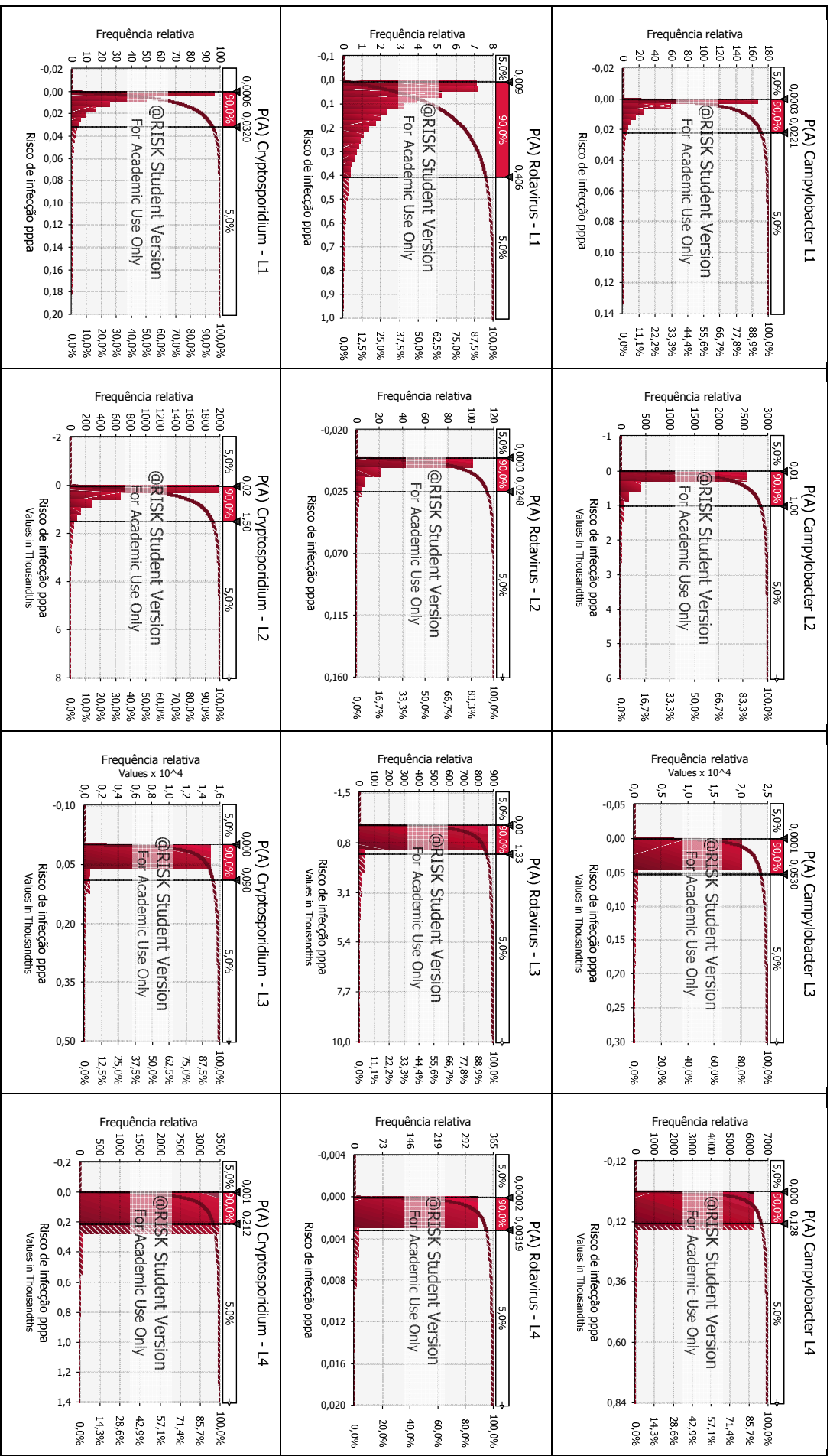


Figura 6.3 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem rentes ao nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme.

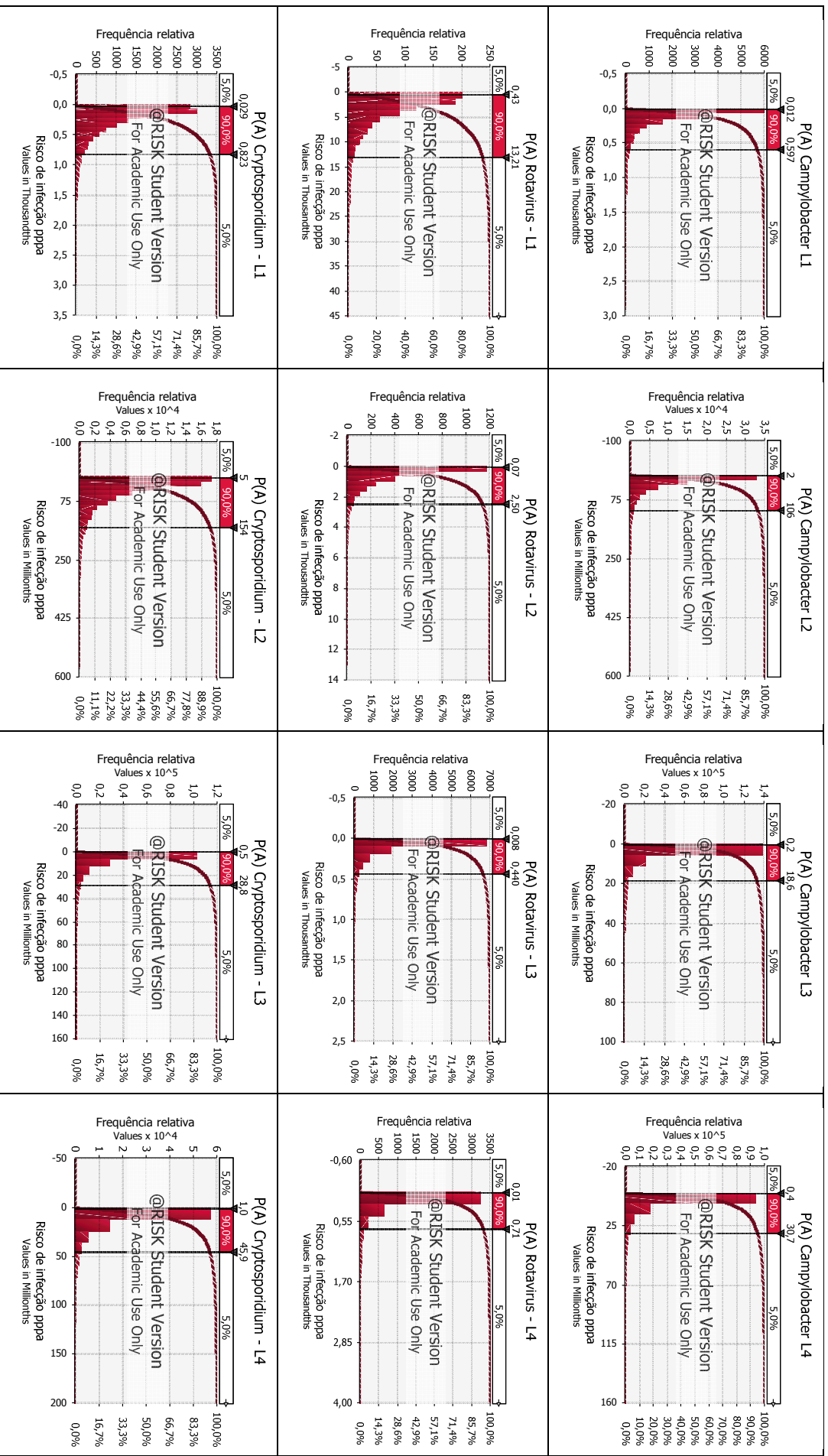


Figura 6.4 - Histogramas das distribuições de probabilidades e curvas de frequência acumulada de riscos anuais de infecção decorrentes do consumo de hortaliças irrigadas que se desenvolvem distantes do nível do solo e que são ingeridas cruas - qualidade dos efluentes das lagoas segundo distribuição uniforme.

### 6.2.2 – Estimativas de contaminação das hortaliças irrigadas

A Figura 6.5 ilustra os histogramas das simulações realizadas para a estimativa da contaminação das hortaliças irrigadas (*E. coli*/g), calculada de acordo com as Equações 3.44 e 3.45 e considerando a qualidade dos efluentes das lagoas (*E. coli*/100 mL) segundo distribuição lognormal. Na Tabela 6.13, estes resultados são resumidos em termos de medianas e percentis 95%.

TABELA 6.13 – Estimativa da contaminação (*E. coli*/g) das hortaliças irrigadas, medianas e percentis 95% - qualidade dos efluentes das lagoas segundo distribuição lognormal.

Lagoa	Desenvolvimento rente ao nível do solo		Desenvolvimento distante do nível do solo	
	Mediana	95%	Mediana	95%
L1	$4,02 \times 10^3$	$1,04 \times 10^4$	$4,79 \times 10^1$	$8,03 \times 10^1$
L2	$1,58 \times 10^2$	$5,47 \times 10^2$	8,28	$1,62 \times 10^1$
L3	3,22	$4,21 \times 10^1$	1,00	4,04
L4	8,51	$9,75 \times 10^1$	1,70	6,37

A Resolução RDC Nº 12 de 2 de janeiro de 2001, da Agência Nacional de Vigilância Sanitária (ANVISA), que aprova o regulamento técnico sobre padrões microbiológicos para alimentos, estabelece um limite máximo de  $10^2$  coliformes termotolerantes por grama para hortaliças frescas, ‘in natura’ (BRASIL, 2001).

Como pode ser observado na Tabela 6.13, a irrigação de hortaliças que se desenvolvem distantes do nível do solo resultou em níveis de contaminação sempre abaixo do padrão estabelecido pela ANVISA, mesmo no caso da Lagoa 1 ( $95\% < 5,2 \times 10^5$  *E. coli*/100 mL – Tabela 6.3), e isso confirma que seleção de culturas (bem como os métodos de irrigação) tem grande impacto sobre os níveis de contaminação e, por conseguinte, sobre os riscos de infecção. Por outro lado, a irrigação com o efluente da Lagoa 1 resultou em estimativas de risco acima dos níveis tomados como referência para risco tolerável (ver Tabelas 6.5 e 6.6). Isso, por sua vez, ressalta a importância da abordagem de AQRM na formulação de normas de qualidade de hortaliças irrigadas.

No cenário mais desfavorável (irrigação de hortaliças que se desenvolvem rentes ao nível do solo), o uso dos efluentes das lagoas 3 e 4 ( $95\% < 6,68 \times 10^2$  e  $1,89 \times 10^3$

*E. coli*/100 mL- Tabela 6.3) resultou em níveis de contaminação muito baixos, bem como em níveis baixos de riscos (ver Tabela 6.5). Isso reforça o entendimento já veiculado de que a inclusão de uma quarta lagoa seria desnecessária e corrobora outro (presente nas Diretrizes da OMS), de que a irrigação com efluentes contendo menos de  $10^3$  *E. coli*/100 mL seria prática segura. No cenário de uso do efluente da Lagoa 2 (95%  $<1,49 \times 10^4$  *E. coli*/100 mL – Tabela 6.3), 50% dos dados de contaminação da hortaliças irrigadas estariam de acordo com o padrão estabelecido pela ANVISA. De certa forma, isso também reforça o entendimento de que o padrão mais rigoroso das Diretrizes da OMS ( $10^3$  *E. coli*/100 mL para o cultivo de raízes e tubérculos e de plantas que se desenvolvem rentes ao nível do solo) é revestido de segurança (ver, porém, discussão no item 6.2.1. sobre os níveis de risco correspondentes).

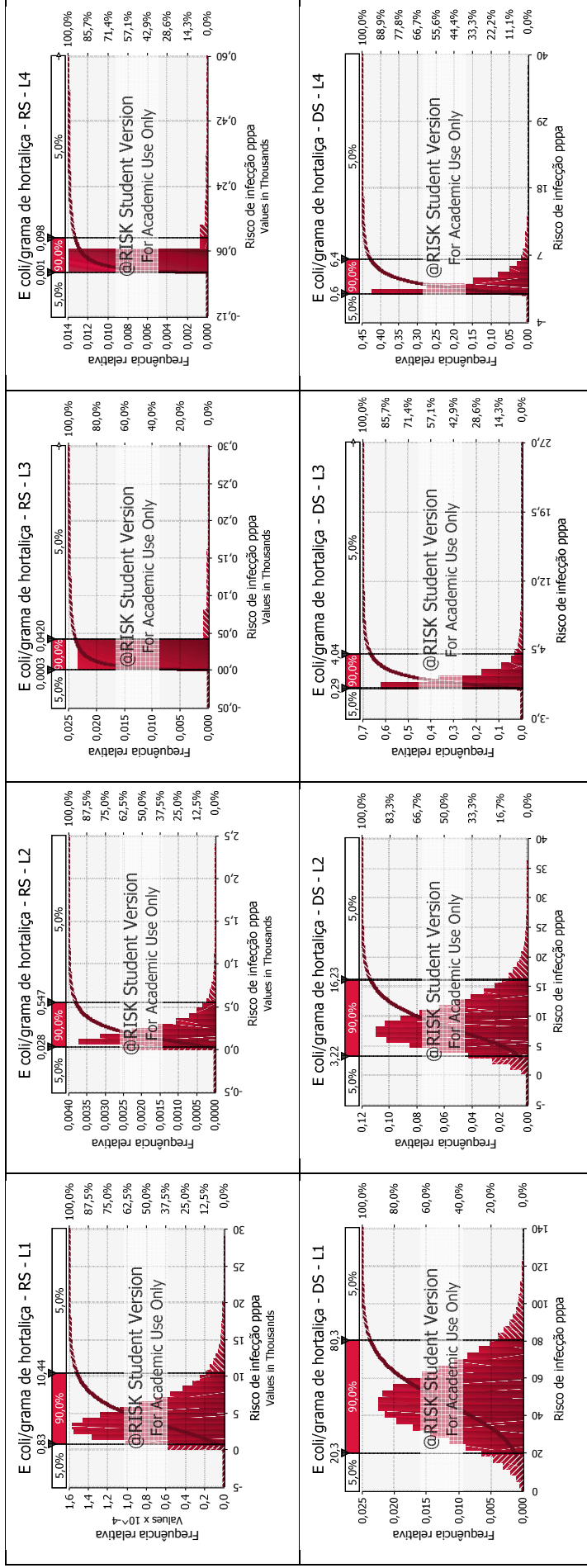


Figura 6.5 - Histogramas (distribuição de probabilidades) da contaminação das hortaliças irrigadas (*E. coli* por grama) - qualidade dos efluentes das lagoas segundo distribuição lognormal.

RS – hortaliças que se desenvolvem rentes ao solo (folhosas e florais); DS – hortaliças que se desenvolvem distantes do solo (frutais); Li: lagoas de polimento.

### 6.3 - Análise de sensibilidade

Os resultados dos testes para a verificação das variáveis de entrada nos modelos de AQRM (modelos de exposição e de dose-resposta) com maior impacto sobre a variável de saída estão apresentados nos Quadros 6.1 a 6.6, de acordo com os coeficientes de Spearman, análises com relação às medianas e análises de cenários. Foram escolhidos, a título de ilustração, gráficos dos três casos com maiores coeficientes de Spearman (Figuras 6.6 a 6.11).

A análise de sensibilidade em relação à mediana revelou-se pouco seletiva na classificação das variáveis como importantes ou não. Portanto, a investigação do coeficiente de Spearman é uma importante análise complementar. Por fim, em geral, os gráficos de dispersão demonstraram coerência com os resultados estatísticos da análise de sensibilidade.

Como descrito na análise de cenários do capítulo 5, o resultado foi expresso em forma de ‘sim’ ou ‘não’, para indicar as variáveis que contribuem significativamente com os resultados de saída maiores que um determinado valor. Novamente, quando o maior valor de risco de infecção anual gerado pelas 10.000 rodadas da simulação de Monte Carlo era menor que os níveis de risco toleráveis ( $10^{-2}$  ou  $10^{-3}$  pppa) não foi possível realizar a análise de cenários.

Como referido anteriormente, para os cenários de exposição e faixas das variáveis de entrada considerados, a análise de sensibilidade com aplicação do teste de Mann-Whitney-Wilcoxon em torno da mediana mostrou-se pouco seletiva, ou seja, todas as variáveis de entrada foram classificadas como ‘importantes’.

Em relação à análise de sensibilidade por meio da classificação dos coeficientes de Spearman, destacam-se as seguintes observações de caráter geral: (i) as variáveis ‘*E. coli* efluente’, ‘redução de patógenos’ (D) e ‘patógenos /  $10^5$  *E. coli*’ (N) revelaram impacto sobre a variável de saída (estimativa de risco). A primeira e a terceira com correlação positiva e a segunda negativa; (ii) as variáveis ‘constante de infecciosidade’ ( $\alpha$  ou r), ‘dose infectante média’ ( $N_{50}$ ) e ‘consumo de hortaliças (C) mostraram pouco impactar as estimativas de risco; (iii) no caso dos riscos de infecção decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo, a concentração de *E. coli* no efluente das lagoas mostrou maior impacto que as variáveis ‘D’ e ‘N’; (iv) a concentração de *E. coli* mostrou importância crescente ao longo da série de lagoas, enquanto a redução de patógenos parece perder importância ao longo do trajeto do efluente, embora de forma menos acentuada.

Quadro 6.1 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por *Campylobacter* decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,54	Importante	Sim	Não	Não	Sim
	L2	0,61	Importante	Sim	Sim	Não	Sim
	L3	0,80	Importante	*	Sim	Não	Sim
	L4	0,77	Importante	*	Sim	Não	Sim
Redução de patógenos – D	L1	-0,51	Importante	Sim	Não	Sim	Sim
	L2	-0,48	Importante	Sim	Sim	Sim	Sim
	L3	-0,37	Importante	*	Não	Não	Sim
	L4	-0,39	Importante	*	Sim	Não	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> - N	L1	0,41	Importante	Sim	Não	Não	Sim
	L2	0,39	Importante	Sim	Sim	Não	Sim
	L3	0,30	Importante	*	Não	Não	Sim
	L4	0,31	Importante	*	Não	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	-0,43	Importante	Sim	Não	Sim	Sim
	L2	-0,40	Importante	Sim	Sim	Não	Sim
	L3	-0,31	Importante	*	Sim	Não	Sim
	L4	-0,34	Importante	*	Sim	Não	Sim
Dose infectante média (N <sub>50</sub> )	L1	-0,10	Importante	Não	Não	Não	Não
	L2	-0,10	Importante	Sim	Não	Não	Não
	L3	-0,08	Importante	*	Sim	Não	Não
	L4	-0,08	Importante	*	Não	Não	Não
Consumo de hortaliças - C	L1	0,06	Importante	Não	Não	Não	Não
	L2	0,06	Importante	Sim	Não	Não	Não
	L3	0,06	Importante	*	Sim	Não	Não
	L4	0,03	Importante	*	Sim	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação Monte Carlo é menor que os níveis de risco assumidos como toleráveis; (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

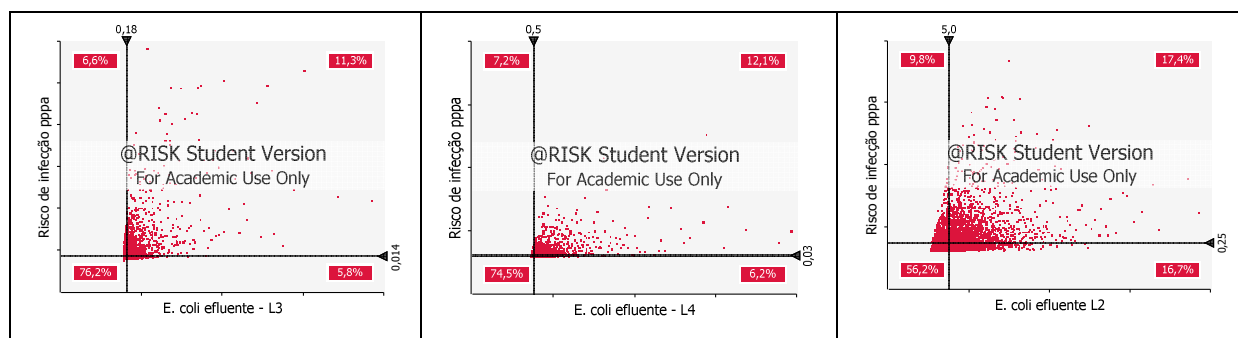


Figura 6.6 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por *Campylobacter* decorrentes do consumo de hortaliças que se desenvolvem rentes ao solo.

Quadro 6.2 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,58	Importante	Não	Não	Não	Sim
	L2	0,65	Importante	Sim	Não	Não	Sim
	L3	0,83	Importante	Sim	Sim	Não	Sim
	L4	0,81	Importante	Sim	Não	Não	Sim
Redução de patógenos - D	L1	-0,56	Importante	Não	Não	Sim	Sim
	L2	-0,52	Importante	Sim	Não	Sim	Sim
	L3	-0,38	Importante	Sim	Sim	Não	Sim
	L4	-0,41	Importante	Sim	Sim	Não	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> - N	L1	0,45	Importante	Não	Não	Não	Sim
	L2	0,42	Importante	Sim	Não	Não	Sim
	L3	0,32	Importante	Sim	Sim	Não	Sim
	L4	0,33	Importante	Sim	Sim	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	-0,23	Importante	Não	Não	Não	Sim
	L2	-0,22	Importante	Sim	Não	Não	Não
	L3	-0,17	Importante	Sim	Sim	Não	Sim
	L4	-0,17	Importante	Sim	Não	Não	Sim
Dose infectante média (N <sub>50</sub> )	L1	-0,12	Importante	Não	Não	Não	Não
	L2	-0,11	Importante	Não	Não	Não	Não
	L3	-0,08	Importante	Não	Não	Não	Não
	L4	-0,09	Importante	Não	Não	Não	Não
Consumo de hortaliças - C	L1	0,05	Importante	Não	Não	Não	Não
	L2	0,04	Importante	Não	Não	Não	Não
	L3	0,04	Importante	Sim	Não	Não	Não
	L4	0,04	Importante	Não	Não	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação de Monte Carlo é menor que os níveis de risco assumidos como toleráveis. (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

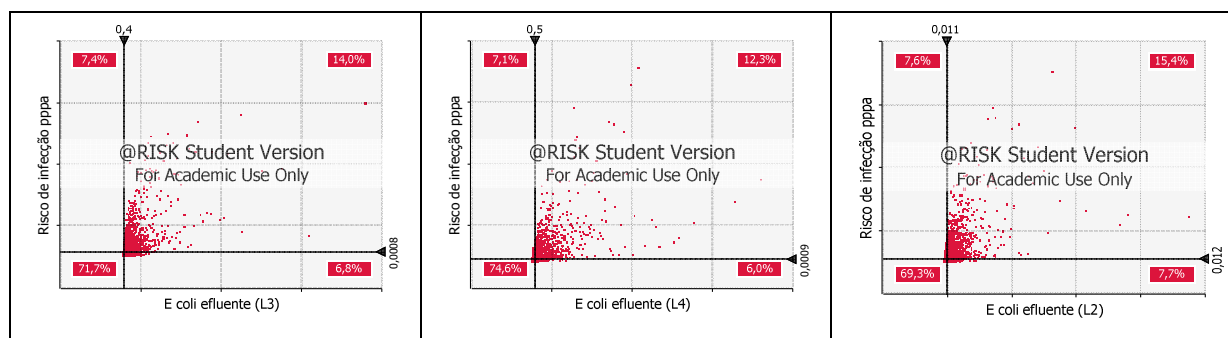


Figura 6.7 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo.

Quadro 6.3 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por *Cryptosporidium* decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,61	Importante	Sim	Não	Não	Sim
	L2	0,67	Importante	Sim	Sim	Não	Sim
	L3	0,85	Importante	*	Sim	Não	Sim
	L4	0,82	Importante	Sim	Sim	Não	Sim
Redução de patógenos - D	L1	-0,57	Importante	Sim	Não	Sim	Sim
	L2	-0,52	Importante	Sim	Sim	Sim	Sim
	L3	-0,37	Importante	*	Sim	Não	Sim
	L4	-0,40	Importante	Sim	Sim	Não	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> - N	L1	0,47	Importante	Sim	Não	Não	Sim
	L2	0,44	Importante	Sim	Sim	Não	Sim
	L3	0,32	Importante	*	Não	Não	Sim
	L4	0,33	Importante	Não	Sim	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	0,11	Importante	Não	Não	Não	Não
	L2	0,10	Importante	Sim	Não	Não	Não
	L3	0,09	Importante	*	Não	Não	Não
	L4	0,09	Importante	Não	Sim	Não	Não
Consumo de hortaliças- C	L1	0,06	Importante	Não	Não	Não	Não
	L2	0,05	Importante	Sim	Não	Não	Não
	L3	0,04	Importante	*	Sim	Não	Não
	L4	0,05	Importante	Sim	Sim	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação de Monte Carlo é menor que os níveis de risco assumidos como toleráveis. (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

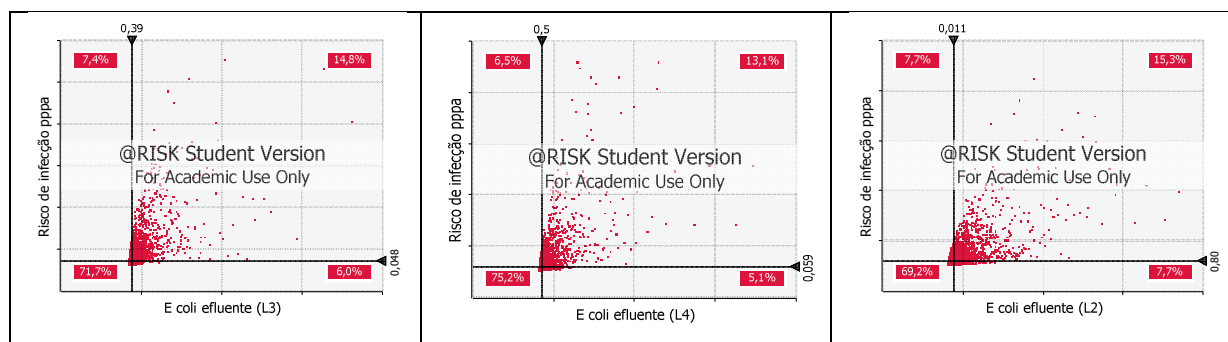


Figura 6.8 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por *Cryptosporidium* decorrentes do consumo de hortaliças que se desenvolvem rentes ao nível do solo.

Quadro 6.4 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por *Campylobacter* decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,34	Importante	*	Sim	Não	Sim
	L2	0,38	Importante	*	*	Não	Sim
	L3	0,59	Importante	*	*	Não	Não
	L4	0,55	Importante	*	*	Não	Não
Redução de patógenos - D	L1	-0,58	Importante	*	Sim	Sim	Sim
	L2	-0,57	Importante	*	v	Sim	Sim
	L3	-0,50	Importante	*	*	Sim	Sim
	L4	-0,51	Importante	*	*	Sim	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> - N	L1	0,46	Importante	*	Sim	Não	Sim
	L2	0,45	Importante	*	*	Não	Sim
	L3	0,41	Importante	*	*	Não	Sim
	L4	0,41	Importante	*	*	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	-0,49	Importante	*	Sim	Sim	Sim
	L2	-0,48	Importante	*	*	Sim	Sim
	L3	-0,42	Importante	*	*	Sim	Sim
	L4	-0,45	Importante	*	*	Sim	Sim
Dose infectante média (N <sub>50</sub> )	L1	-0,11	Importante	*	Não	Não	Não
	L2	-0,11	Importante	*	*	Não	Não
	L3	-0,10	Importante	*	*	Não	Não
	L4	-0,10	Importante	*	*	Não	Não
Consumo de hortaliça - C	L1	0,01	Importante	*	Não	Não	Não
	L2	0,02	Importante	*	*	Não	Não
	L3	0,04	Importante	*	*	Não	Não
	L4	0,01	Importante	*	*	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação de Monte Carlo é menor os níveis de risco assumidos como toleráveis. (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

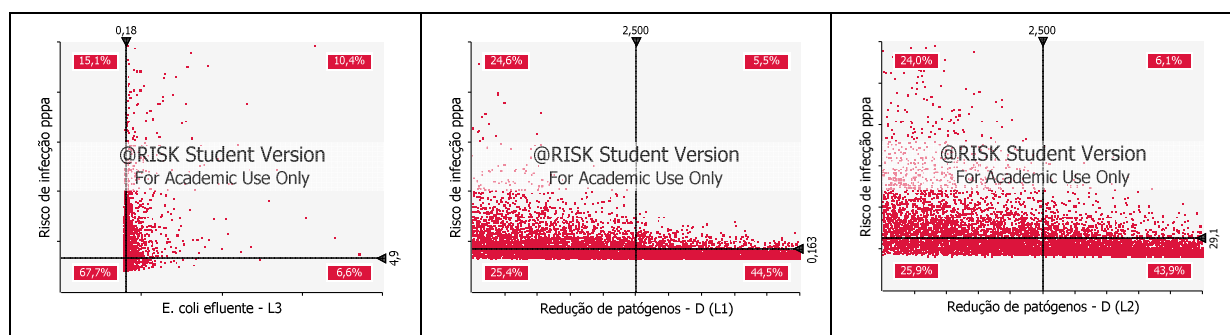


Figura 6.9 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por *Campylobacter* decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

Quadro 6.5 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,37	Importante	Sim	Não	Não	Sim
	L2	0,43	Importante	Sim	Não	Não	Sim
	L3	0,63	Importante	*	Não	Não	Não
	L4	0,59	Importante	*	Sim	Não	Sim
Redução de patógenos - D	L1	-0,64	Importante	Sim	Não	Sim	Sim
	L2	-0,62	Importante	Sim	Sim	Sim	Sim
	L3	-0,53	Importante	*	Sim	Sim	Sim
	L4	-0,56	Importante	*	Não	Sim	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> – N	L1	0,53	Importante	Sim	Não	Sim	Sim
	L2	0,51	Importante	Sim	Sim	Sim	Sim
	L3	0,44	Importante	*	Sim	Não	Sim
	L4	0,46	Importante	*	Não	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	-0,27	Importante	Sim	Não	Não	Sim
	L2	-0,27	Importante	Sim	Sim	Não	Sim
	L3	-0,24	Importante	*	Não	Não	Sim
	L4	-0,24	Importante	*	Sim	Não	Sim
Dose infectante média (N <sub>50</sub> )	L1	-0,14	Importante	Não	Não	Não	Não
	L2	-0,13	Importante	Sim	Não	Não	Sim
	L3	-0,12	Importante	*	Não	Não	Não
	L4	-0,13	Importante	*	Não	Não	Não
Consumo de hortaliça - C	L1	0,01	Importante	Não	Não	Não	Não
	L2	0,02	Importante	Não	Não	Não	Não
	L3	0,02	Importante	*	Não	Não	Não
	L4	0,01	Importante	*	Não	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação de Monte Carlo é menor que os níveis de risco assumidos como toleráveis. (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

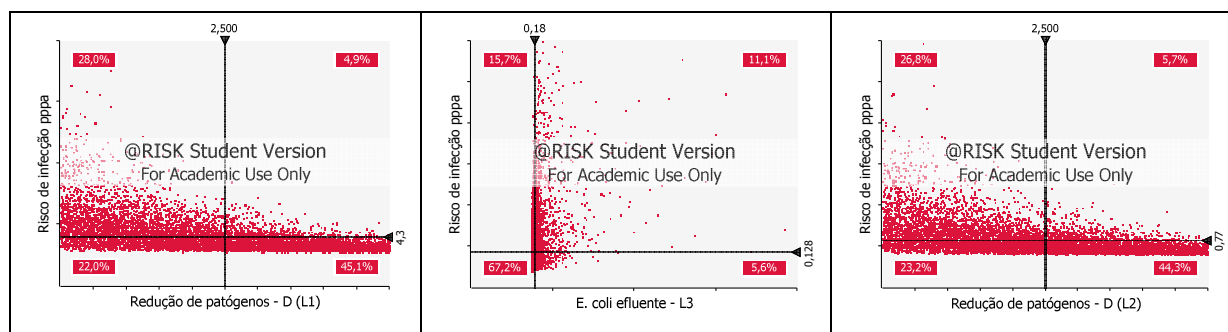


Figura 6.10 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por rotavírus decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

Quadro 6.6 - Coeficientes de Spearman, resultados da análise de sensibilidade em relação à mediana e da análise de cenários para riscos de infecção por *Cryptosporidium* decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

Variável de entrada	Lagoa	Coeficiente de Spearman	Análise da mediana	Análise de cenários <sup>(1)</sup>			
				> 10 <sup>-2</sup>	> 10 <sup>-3</sup>	> 50%	> 95%
<i>E. coli</i> efluente	L1	0,40	Importante	*	Sim	Não	Sim
	L2	0,45	Importante	*	*	Não	Sim
	L3	0,64	Importante	*	*	Não	Não
	L4	0,61	Importante	*	*	Não	Sim
Redução de patógenos - D	L1	-0,67	Importante	*	Sim	Sim	Sim
	L2	-0,64	Importante	*	*	Sim	Sim
	L3	-0,54	Importante	*	*	Sim	Sim
	L4	-0,57	Importante	*	*	Sim	Sim
Patógenos / 10 <sup>5</sup> <i>E. coli</i> – N	L1	0,55	Importante	*	Sim	Sim	Sim
	L2	0,54	Importante	*	*	Sim	Sim
	L3	0,46	Importante	*	*	Não	Sim
	L4	0,47	Importante	*	*	Não	Sim
Constante de infecciosidade ( $\alpha$ ou $r$ )	L1	0,14	Importante	*	Sim	Não	Sim
	L2	0,13	Importante	*	*	Não	Sim
	L3	0,13	Importante	*	*	Não	Não
	L4	0,13	Importante	*	*	Não	Sim
Consumo de hortaliças - C	L1	0,02	Importante	*	Sim	Não	Não
	L2	0,01	Importante	*	*	Não	Não
	L3	0,01	Importante	*	*	Não	Não
	L4	0,01	Importante	*	*	Não	Não

\* o maior valor de risco de infecção anual pppa gerado pelas 10.000 rodadas da simulação de Monte Carlo é menor que os níveis de risco assumidos como toleráveis. (1) 10<sup>-2</sup> e 10<sup>-3</sup>: riscos anuais toleráveis de infecção (pppa); 50% e 95%: valores da mediana e percentil 95% da distribuição de probabilidades dos riscos estimados.

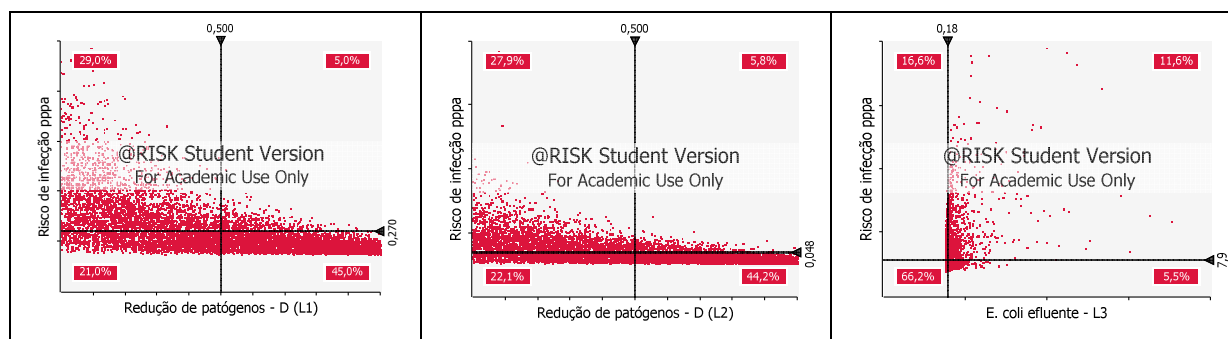


Figura 6.11 - Gráfico de dispersão, variável de entrada versus risco anual de infecção por *Cryptosporidium* decorrentes do consumo de hortaliças que se desenvolvem distantes do nível do solo.

A pouca importância das variáveis ‘constante de infecciosidade’ ( $\alpha$  ou  $r$ ), ‘dose infectante média’ ( $N_{50}$ ) e ‘consumo de hortaliças (C)’ pode ser creditada às faixas de variação relativamente estreitas assumidas, respectivamente, nos modelos dose-resposta e de exposição. No caso das variáveis do modelo dose-resposta ( $\alpha$ ,  $r$  e  $N_{50}$ ) foi assumida variação de  $\pm 25\%$  em torno dos valores médios e de acordo com distribuição uniforme. As variações atribuídas a esses parâmetros responderiam por incertezas. Portanto, a indicação de pequeno impacto de suas variações é importante, tendo em vista que sua melhor apropriação é, de toda forma, difícil.

No caso da variável ‘consumo de hortaliças (C)’, as respectivas variações não se devem exatamente às incertezas, mas ao menos parcialmente à variabilidade, uma vez que a POF do IBGE conta com um banco de dados relativamente vasto e detalhado. O fato de que no presente exercício esta variável não tenha se mostrado determinante na estimativa de risco não significa que o resultado se manteria com outras faixas extraídas da própria POF. Nesse sentido, cumpre registrar que em um dos poucos trabalhos sobre AQRM aplicada ao uso agrícola de esgotos incluindo análise de sensibilidade, a ‘variável consumo de hortaliças’ revelou ser a de maior impacto na determinação dos riscos (HAMILTON *et al.*, 2006).

Por outro lado, a relevância da variável ‘concentração de *E. coli* efluente’ confirma a importância do esforço em bem caracterizar sua distribuição de frequência, cujas variações se devem, tipicamente, mais à variabilidade do que à incerteza. A comparação dos riscos estimados considerando as distribuições uniforme e lognormal reforçam essa observação (ver item .6.2.1).

O fato das variáveis ‘*E. coli* efluente’ e ‘patógenos /  $10^5$  *E. coli*’ (N) serem estreitamente relacionadas (embora a modelagem estocástica não tenha considerado a existência de correlação entre variáveis de entrada), e de que essas duas variáveis tenham revelado impacto na determinação dos riscos, aponta para a necessidade de melhor conhecimento da variável ‘N’. O mesmo se aplica à ‘redução de patógenos’ (D), cujas variações aqui assumidas podem ser consideradas devidas ao desconhecimento (incerteza propriamente dita). Outros trabalhos destacam a importância dessa variável na estimativa de riscos (PETTERSON *et al.*, 2001; HAMILTON *et al.*, 2006).

Os resultados da Análise de Cenários foram bastante variados e este teste ficou por vezes limitado pelo fato de que os valores estimados de risco foram inferiores aos ‘riscos toleráveis’. Não obstante, salvo exceções, os resultados anteriores da análise de sensibilidade viram-se confirmados: as variáveis ‘*E. coli* efluente’, ‘redução de patógenos’ (D) e

‘patógenos /  $10^5$  *E. coli*’ (N) revelaram maior impacto na determinação de riscos superiores aos níveis de referência do que as variáveis ‘constante de infecciosidade’ ( $\alpha$  ou r), ‘dose infectante média’ ( $N_{50}$ ) e ‘consumo de hortaliças (C). No caso das variáveis de pouco impacto, destacaram-se a ‘dose infectante média’ e o consumo de hortaliças. No caso das variáveis de maior impacto, os destaques seriam a ‘redução de patógenos’ (D) e ‘patógenos /  $10^5$  *E. coli*’ (N).

Nota-se ainda que as variáveis de impacto determinam mais nitidamente os valores de risco mais elevados, ou sejam superiores ao percentil 95%. Nota-se também, mais uma vez salvo exceções, que a exemplo dos resultados da análise de sensibilidade, a importância das variáveis ‘concentração de *E. coli*’ e ‘N’ na determinação de riscos superiores à  $10^{-2}$  e  $10^{-3}$  pppa cresce ao longo da série de lagoas.

Entretanto, tomar o percentil 95% como valor de referência de risco seria uma abordagem conservadora. Considerando então a mediana, a variável ‘D’ (redução de patógenos) mostra-se, em geral, mais importante que a concentração de *E. coli* efluente.

#### **6.4 – Considerações finais, conclusões e recomendações**

As estimativas de risco aqui apresentadas são de difícil comparação com resultados de outros estudos devido às especificidades dos cenários de exposição assumidos. Por exemplo, as Diretrizes da OMS estão baseadas no consumo de 100 g de alfaces a cada dois dias e na estimativa de contaminação de hortaliças com base em volumes de água sobre a superfície das culturas irrigadas (pressupostos estes assumidos em vários outros trabalhos publicados).

No que diz respeito ao padrão de consumo de alimentos, cabe destacar que os dados do IBGE adotados neste trabalho são bem menores que os pressupostos das Diretrizes da OMS (100 g de alface a cada 2 dias). Por exemplo, o consumo médio *per capita* anual de hortaliças folhosas e florais no Brasil é de 2,5 kg, ou 6,8 gramas por pessoa por dia. No estado com maior consumo *per capita* (Rondônia), a média anual é de 7,14 kg por pessoa, ou 19,6 g por pessoa por dia.

Por outro lado, o uso da Equação 3.44 para hortaliças que se desenvolvem rentes ao nível do solo eleva a estimativa de contaminação de culturas irrigadas em relação ao critério baseado em volumes da água de irrigação sobre a superfície das plantas. Supõe-se, portanto, que esses dois fatores de certa forma se compensem, mas isso teria de ser verificado

numericamente para efeito de comparações mais objetivas entre os cenários de exposição assumidos aqui e nas Diretrizes da OMS.

Ainda em relação ao padrão de consumo de hortaliças, o critério adotado conduz à estimativas de risco conservadoras, uma vez que se assume o consumo de grupos inteiros de hortaliças folhosas e florais e de hortaliças frutais e que a aquisição domiciliar desses alimentos se restrinja a produtos irrigados com esgotos.

Há também que se observar que, como mencionado anteriormente, o banco de dados da POF do IBGE é vasto e permite abordagens outras que a assumida no presente trabalho, ou seja, as faixas (estreitas) de variação e a distribuição uniforme. Cumpre também destacar que as faixas de valores e distribuições de frequência assumidas para muitas das variáveis de entrada foram baseadas em dados de literatura e muitas vezes revestem-se de incertezas propriamente ditas e não de variabilidade conhecida. As próprias Equações 3.44 e 3.45 podem entrar no modelo de exposição considerando variações, por exemplo, por meio da determinação de intervalos de confiança em torno das variáveis da regressão linear (o banco de dados original permite isso). Em resumo, a modelagem estocástica aqui desenvolvida pode e deve ser sofisticada.

Portanto, mais que os valores de estimativas de risco em si, talvez o mérito maior deste trabalho resida na aplicação mesma da modelagem estocástica e das ferramentas de análise de sensibilidade para a identificação de variáveis-chave que demandem maior esforço para melhor apropriação.

Não obstante, as estimativas de risco permitem as seguintes sugestões, baseadas nos modelos de exposição assumidos:

- (i) Tomando como referência um nível de risco tolerável de  $10^{-3}$  pppa, um padrão de  $10^3$  *E. coli*/100 mL parecer adequado para a irrigação de hortaliças que se desenvolvem rentes ao solo e que são ingeridas cruas.
- (ii) Para o mesmo nível de risco tolerável, um padrão de  $10^4$  *E. coli*/100 mL parece adequado para a irrigação de hortaliças que se desenvolvem distantes do nível do solo e que são ingeridas cruas (à exceção da irrigação por aspersão).

Porém, como discutido anteriormente, em se aceitando um nível de risco tolerável de  $10^{-2}$  pppa, tais padrões poderiam ser flexibilizados.

Em geral, a análise de sensibilidade indicou que as variáveis de maior impacto na determinação dos riscos foram a concentração de *E. coli* efluente, a relação patógenos / *E. coli*

e a redução de patógenos entre a colheita e o consumo. Nesse sentido, conclui-se / sugere-se que:

- (i) O uso da distribuição lognormal, em lugar da distribuição uniforme, para os dados de concentração de *E. coli* efluente leva à estimativas de risco mais baixas e ao mesmo tempo mais realistas, uma vez que a distribuição lognormal bem descreveu a variação da qualidade dos efluentes das lagoas.
- (ii) Para estimativas de risco ainda mais realistas, requerem-se dados mais confiáveis de relação patógenos / *E. coli* ou mesmo de contaminação de culturas irrigadas com base na concentração de patógenos, ao invés de organismos indicadores.
- (iii) De forma similar, tornam-se necessárias informações adicionais sobre a redução de patógenos entre a colheita e o consumo.
- (iv) Em que pesem as incertezas em torno dos parâmetros dos modelos dose-resposta, ao menos dentro da faixa de variação considerada ( $\pm 25\%$  em torno dos valores médios), estas não se mostraram determinantes na estimativa de riscos.

Por fim, sugere-se testar o mesmo modelo de AQRM aqui apresentado, porém utilizando como dados de entrada:

- (i) As médias aritméticas e geométricas das doses de patógenos, bem como as médias de consumo de hortaliças.
- (ii) As médias aritméticas e geométricas das doses de patógenos, porém considerando variações em torno do consumo de hortaliças, de preferência buscando ajustar distribuições de probabilidades aos dados da Pesquisa de Orçamentos Familiares do IBGE.

Isso permitiria a comparação entre as abordagens determinística e estocástica, além da avaliação comparativa do uso de diferentes medidas de tendência central.

## REFERÊNCIAS

AGUNWAMBA, J .C., EGBUNIWE, N.; ADEMILUYI, J .O. Prediction of the dispersion number in waste stabilization ponds. *Water research*, v.26, n.1, p. 85-89, 1992.

AMOAHA, P.; DRECHSEL, P.; ABAIDOO, R. C.; HENSELER, M. Effectiveness of common and improved sanitary washing methods in selected cities of West Africa for the reduction of coliform bacteria and helminth eggs on vegetables. *Tropical medicine and international health*, v. 12, supp.2. p 40-50, 2007.

APHA/AWWA/WEF. *Standard methods for the examination of water and wastewater*, 20. ed. Washington: APHA, 1998.

ARCEIVALA, S. J. *Wastewater treatment and disposal*. Macel Dekker, New York, 892 p. 1981.

ASANO, T.; LEONG, L. Y. C.; RIGBY, M. G.; SAKAJI, R. H. Evaluation of the California wastewater reclamation criteria using enteric virus monitoring data. *Water science and technology*, v.26, n 7-8, p.1513 – 1524, 1992.

AYRES, M. R.; ALABASTER, G. P.; MARA, D. D.; LEE, D. L. A design equation for human intestinal nematode egg removal in waste stabilizations ponds. *Water research*, v. 26, n 6, p. 863 – 865, 1992.

AYRES, M. R.; MARA, D. *Analysis of wastewater for use in agriculture. A laboratory manual of parasitological and bacteriological techniques*. Genova: WHO, 1996. 31p.

BASTOS, R. K. X.; BEVILACQUA. Utilização de esgotos sanitários: Marcos conceituais e regulatórios. In: FLORENCIO, L.; BASTOS, R. K. X.; MANSUR, M. A (Coords) *Normas e critérios de qualidade para reúso da água*. Belo Horizonte: ABES, 2006, p. 17 - 62. (Projeto PROSAB).

BASTOS, R. K. X.; BEVILACQUA, P.D.; KELLER, R. Organismos patogênicos e efeitos sobre a saúde humana. In: GONÇALVES, R. F. (Coord.) *Desinfecção de efluentes sanitários*. Rio de Janeiro: ABES, RiMa, 2003a. p.27-88 (Projeto Prosab).

BASTOS, R. K. X., BEVILACQUA, P. D., SILVA, C. A. B., SILVA, C. V.; Wastewater irrigation of salad crops: further evidence for the evaluation of the WHO guidelines. *Water science & technology*, v. 57, n. 8, p 1213 – 1219, 2008.

BASTOS, R. K. X. ; RIOS, E . N.; BEVILACQUA, P. D.; ANDRADE, R .C. UASB-polishing ponds design parameters. Contributions from a pilot scale study in southeast Brazil. In: IWA SPECIALIST CONFERENCE ON WASTE STABILIZATION PONDS, 8. Belo Horizonte, Brazil. 2009. *Proceedings*....Belo Horizonte: IWA, 2009 (CD ROM) .

BENKE, K. K., HAMILTON, E. A. J., Quantitative microbial risk assessment: uncertainty and measures of central tendency for skewed distributions. *Stochastic environmental research and risk assessment*, v. 22, p 533–539, 2008.

BERTHOUEX, P.M., HUNTER, W.G. How to construct reference distributions to evaluate treatment plant effluent quality. *Journal of water pollution control federation*, v. 55, n. 12, p. 1417-1424, 1983.

BERTHOUEX, P.M., HUNTER, W.G. Simple statistics for interpreting environmental data. *Journal of water pollution control federation*, v. 53, n. 2, p. 167-175, 1981.

BERTHOUEX, P.M., HUNTER, W.G. Treatment plant monitoring programs: A preliminary analysis. *Journal of water pollution control federation*, v. 47, n. 8, p. 2143-2156, 1975.

BRASIL. AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. Resolução nº 12 de 30 de março de 1978. Aprova normas técnicas especiais do Estado de São Paulo, revistas pela Comissão Nacional de Normas e Padrões para Alimentos relativas a alimentos (e bebidas), para efeito em todo território brasileiro. *Diário Oficial da União* 24 de julho de 1987.

BRASIL. CONSELHO NACIONAL DO MEIO AMBIENTE. Resolução nº 375 de 17 de março de 2005. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. *Diário Oficial da União*. 18 de março de 2005, Seção 1, p.58 - 63.

BUCHBERGER, S. G., SHAW, G. B. An approach toward rational design of constructed wetlands for wastewater treatment. *Ecological Engineering*. v. 4, n. 4, p.249-275, 1995.

CAMARGO VALERO, M. A.; MARA, D. D.; Nitrogen removal via ammonia volatilization in maturation ponds. *Water science and technology*, v. 55, n 11, p 87 - 92, 2007a.

CAMARGO VALERO, M. A.; MARA, D. D.; Nitrogen removal in maturation ponds: tracerexperiments with <sup>15</sup>N-labelled ammonia. *Water science and technology*, v. 55,n 11, p 81 - 85, 2007b.

CARR, R.M., BLUMENTHAL, U.J.; MARA D.D. Health guidelines for the use of wastewater in agriculture: developing realistic guidelines. In: SCOTT, C.A.; FARUQUI, N.I. RASHID-SALY, L. (ed). *Wastewater use in irrigated agriculture: coordinating the livelihood and environmental realities*. Wallingford: CAB International, p .45-62, 2004.

CAVALCANTI, P. F. F.; VAN HAANDEL, A.; KATO, M.; VON SPERLING, M.; LUDUVICE, M. L.; MONTEGGIA, L. O. Pós-tratamento de efluentes anaeróbios por lagoas de polimento. In: CHERNICHARO, C. A. (Coord.) *Pós-tratamento de reatores anaeróbios*. Belo Horizonte: ABES, 2001, p.105-170. (Projeto PROSAB).

CHARLES, K. J.; ASHBOLT, N. J.; ROSER, D. J.; MCGUINNESS, R; DEERE, D. A. Effluent quality from 200 on-site sewage systems: design values for guidelines. *Water science and technology*, v. 15, n 10, p 163 - 169, 2005.

COSTA NETO, P. L. O. *Estatística*. São Paulo: Edgard Blücher, 1977, 264p.

CHERNICHARO, C, A, L. *Reatores anaeróbios. Princípios do tratamento biológico de águas residuárias*, Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, 1997, 246p. (Princípios do tratamento biológico de águas residuárias, 5).

CHRISTMANN, R, U. *Estatística aplicada*. 2. ed. São Paulo: Edgard Blücher, 1978, 141p.

EDWARDS, P. *Reuse of human wastes in aquaculture. A technical review* Washington DC: The World Bank, 350p.1992.

FERNANDES, A.; TEJEDOR, C.; CHORDÍ, A. Effect of different factors on the die-off of fecal bacteria in a stabilization pond purification plant. *Water research*, v. 26, n 8, p. 1093 - 1098, 1992.

FINLEY, B.L. AND D.J. PAUSTENBACH. The benefits of probabilistic exposure assessment: Three case studies involving contaminated air, Water, and Soil. *Risk analysis*. v. 14, n. 1, p. 53-73, 1994.

FLORENCIO, L.; BASTOS, R. K. X.; MANSUR, M. A. Utilização de esgotos sanitários- Marcos conceituais e regulatórios. In: FLORENCIO, L.; BASTOS, R. K. X.; MANSUR, M. A (Coords) *Reúso das águas de esgoto sanitário, inclusive desenvolvimento de tecnologia para este fim*. Belo Horizonte: ABES, 2006, p. 1-16. (Projeto PROSAB).

FREY, H.C., PATIL, S.R. Identification and review of sensitivity analysis methods. *Risk Analysis*. v. 22, n. 3, p.553-578, 2002.

GAWASIRI, C. B. *Modern design of waste stabilization ponds in warm climates: comparison with traditional design methods*. 2003. 128f. Dissertation (Master of Science in Environmental Engineering and Project Management) – University of Leeds, School of Civil Engineering. Leeds, UK, 2003.

GRIMASON, A. M.; SMITH, H. V.; THITAI, W. N.; SMITH, P. G.; JACKSON, M. H.; GIRDWOOD, R. W. A. Occurrence and removal of *cryptosporidium ssp.* oocysts and *giardia spp.* cysts in Kenyan waste stabilisation ponds. *Water Science and Technology*, v. 27, n. 3, p. 97–104, 1993.

GUIMARÃES, D., P., *Uma função hiperbólica de distribuição probabilística de alta flexibilidade*. 40p., 2002. Disponível em: <<http://www.cpac.embrapa.br/publicacoes/doc/49>> Acesso em: 2009.

HAAS, C. A., How to average microbial densities to characterize risks. *Water research*. v. 30, n. 4, p. 1036-1038, 1996.

HAAS C. N.; ROSE J. B, GERBA, C.P. *Quantitative microbial risk assessment*. New York: John Wiley & Sons, 1999, 449p.

HAMILTON, A. J., STAGNITTI, F., PREMIER, R., BOLAND, A. M., HALE, G.; Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. *Applied and environmental microbiology*, v. 72, n. 5, p. 3284 – 3290, 2006.

HELTON, J. C.; DAVIS, F. J., JOHNSON, J. D., A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling. *Reliability engineering e system safet*, v. 89, n. 3, p. 305-330, 2005

HINES, W. W., MONTGOMERY, D. C., GOLDSMAN, D. M., BORROR, C. M.; *Probabilidade e estatística na engenharia*. Rio de Janeiro: LTC, 2006. 604 p.

ROFFMANN, R. Estatística para economistas. 3 ed. São Paulo: Livraria Pioneira Editora, 1991, 426p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Pesquisa de aquisição alimentar domiciliar *per capita* - Brasil e grandes regiões. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/pof/2002aquisicao/default.shtm>>. Acesso em: 2009.

JAFFE, J.; ROSS, S. A.; WESTERFIELD, R. W.; *Administração financeira: Corporate finance*, 2ª ed. São Paulo: Atlas, 2002, 7778p.

JORDÃO, E. P.; PESSÔA, C. A. *Tratamento de esgotos domésticos*. 4.ed. Rio de Janeiro: ABES. 2005, 960p.

KAMIZOULIS, G.; *Setting health based targets for water reuse (in agriculture)*. *Desalination*, v. 218, n. 1-3, p. 154-163, 2008.

KARAVARSAMIS, N., HAMILTON, A. J.; Estimators of annual probability of infection for quantitative microbial risk assessment. *Journal of water and health*, 2009.

KATO, M. T.; FLORÊNCIO, L. Pós-tratamento de efluente anaeróbio em lagoa de polimento. In: CHERNICHARO, C. A. (Coord.) *Pós-tratamento de efluentes de reatores anaeróbios*. Belo Horizonte: ABES, 2001, p.59 - 68. (Projeto PROSAB).

KELLER, E.; PIRES, E. C. Lagoas de estabilização: projeto e operação. Rio de Janeiro: ABES, 1998, 244.p.

LEVENSPIEL, O. *Engenharia das reações químicas*. 3. ed. São Paulo: Editora Edgard Blücher LTDA. 1999, 584p.

MANCUSO, P. C. S.; SANTOS, H. F. *Reúso de água*. PHILIPPI JR, A. (Coord.). São Paulo: ABES, 2003, 590p.

MARA, D. D. *Domestic wastewater treatment in developing countries*. Earthscan Publication, London, 2004.

MARA, D. D.; 'A Guide to the guidelines': A numerical guide to the 2006 WHO guidelines on wastewater use in agriculture and practical advice on how to transpose them into national standards. School of Civil Engineering, University of Leeds, Leeds LS2 9JT, UK, November 2006.

MARA, D. D.; Wastewater treatment for wastewater reuse: required effluent qualities and how to achieve them at reasonable cost. *Regional EMWater project conference*, Amman, 2006.

MARA, D. D., SLEIGH, P. A., BLUMENTHAL, U. J., CARR, R. M.; Health risks in wastewater irrigation: Comparing estimates from quantitative microbial risk analyses and epidemiological studies. *Journal of water and health*, v. 5, n. 1, p. 39-50, 2007

MARA, D. D., SLEIGH, A.; Estimation of norovirus infection risks to consumers of wastewater-irrigated food crops eaten raw. *Journal of water and health*, v. 8, n. 1, p. 39-43, 2009.

MARAIS, G. R.; Faecal bacterial kinetics in stabilization ponds. *Journal of the environmental engineering division*, v. 100, n. p. 119-139, 1974.

MENA, K. D., PILLAI, S. D. Developing quantitative risk-based microbial standards for fresh produce. *Journal of water and health*, v. 6, n. 3, p. 359-364, 2008.

MENDES, S. B.; NASCIMENTO, M. J. do; PEREIRA, M. I.; BAILEY, G.; LAPA, N. MORAIS, J; OLIVEIRA, J. S. Efficiency of removal in stabilization ponds II. Statistical analysis of k values. *Water science and technology*, v. 31, n. 12, p. 231 – 238, 1995.

METCALF & EDDY. *Wastewater engineering: treatment, and reuse*. 4<sup>th</sup> ed. New York: Metcalf & Eddy, Inc., 2003, 1819p.

MEYER, P., L., *Probabilidade. Aplicações à estatística*. Rio de Janeiro: Universidade de São Paulo, 1969, 391p.

MINAS GERAIS, CONSELHO ESTADUAL DE POLÍTICA AMBIENTAL; CONSELHO ESTADUAL DE RECURSOS HÍDRICOS, Deliberação Normativa Conjunta nº 01 de 05 de maio de 2008. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. *Diário do executivo de Minas Gerais*. 13 de maio de 2008.

MONTGOMERY, D., C.; *Estatística aplicada e probabilidade para engenheiros*. 2.ed. Rio de Janeiro: LTC Editora, 2002. 462p.

MORGAN, M. G.; HERION, M. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. *Cambridge university press*. v. 28, n 11-12, p. 219-222, 2007.

NAGHETTINI, M., PINTO, E. J. A.; *Hidrologia estatística*. Belo Horizonte: CPRM, 2007. 552 p.

NETO, P. O. C., *Estatística*. 7ª ed. São Paulo: Edgar Blücher Ltda, 1987. 264 p.

NIKU, S., SCHROEDER, E.D., SAMANIEGO F.J. Performance of activated sludge process and reliability-based design. *Journal water pollution control association*, v. 51, n. 12, p. 2841 - 2857, 1979.

NIKU, S., SCHROEDER, E.D. Factors affecting effluent variability from activated sludge processes. *Journal water pollution control association*, v. 53, n. 5, p. 546 - 559, 1981a.

OLIVEIRA, S. M. A. C., *Análise de desempenho e confiabilidade de estações de tratamento de esgotos*. 2006. 214 f. Tese (Doutorado) – Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos.

PALISADE CORPORATION, *Add-In do Microsoft® Excel para simulação e análise de riscos*, Versão 5.5, maio 2009.

PANO, A., MIDDLEBROOKS, E. J. Ammonia nitrogen removal in facultative ponds. *Journal of the water pollution control federation*, v. 4, n. 54, p. 344-351, 1982.

PETTERSON, S. R.; ASHBOLT, N. SHARMA, A. Microbial risks from wastewater irrigation of salad crops: a screening-level risk assessment. *Water environmental resources*, v. 73, n. 6, p. 667–672, 2001.

POLPRASERT, C.; DISSANAYAKE, G. M.; THANH, N. C. Bacterial die-off kinetics in waste stabilization ponds. *Journal water pollution control federation*, v. 55, n. 3, p. 285-296, 1983.

RIOS, E. N., *Caracterização e comportamento de uma série de lagoas de polimento tratando esgotos sanitários*. 2007. 127f. Dissertação (Mestrado) – Universidade Federal de Viçosa, Programa de Pós-Graduação em Engenharia civil.

ROSE, J. B., L. J. DICKSON, S. R. FARRAH, R. P. CARNAHAN. 1996. Removal of pathogenic and indicator microorganisms by a full-scale water reclamation facility. *Water research*. v. 30, n. 11, p.2785–2797, 1996.

SALIBY, E. *Repensando a simulação - a amostragem descritiva*. São Paulo: Atlas; 1989, 182 p.

SALIBY, E.; MOREIRA, F. F. P.; *Estudo comparativo dos métodos de quasi-Monte Carlo, amostragem descritiva, hipercubo latino e Monte Carlo clássico na análise de risco*. 2002. Disponível em < <http://www.virtualbib.fgv.br/dspace/handle> > Acesso em: 2009.

SAQQAR, M. M.; PESCOD, M. B. Modelling coliform reduction in wastewater stabilization ponds. *Water science and technology*, v. 26, n. 7-8, p. 1667 – 1677, 1992.

SAQQAR, M. M.; PESCOD, M. B. Modelling nematode egg elimination in wastewater stabilization ponds. *Water science and technology*, v. 26, n. 7-8, p. 1659 – 1665, 1992.

SEIDU, S., HEISTAD, A., AMOAH, P., DRECHSEL, P., JENSSEN, P. D., STENSTROÖM, T. A. Quantification of the health risk associated with wastewater reuse in Accra, Ghana: a contribution toward local guidelines. *Journal of water and health*. v. 6, n. 4, p. 641–671, 2008.

SERRA NEGRA, C. A., SERRA NEGRA E. M.; *Manual de trabalhos monográficos de graduação, especialização, mestrado e doutorado*. 4ª ed. São Paulo: Atlas, 2009. 218 p.

SCOTT, C.A.; FARUQUI, N.I. RASHID-SALY, L. (ed). *Wastewater use in irrigated agriculture: coordinating the livelihood and environmental realities*. Wallingford: CAB International, 197p., 2004

SIEGEL, S., CASTELLAN, N. J. J.; tradução: CARMONA, S. I. C.; *Estatística não-paramétrica para ciência do comportamento*. 2ª ed. Porto Alegre: Artmed, 2006. 448p.

SHAPIRO, S., S.; WILK, M.,B.; An analysis of variance test for normality (complete samples). *Biometrika*, v. 52, n. 3-4, p. 591-611, 1965.

SHERRY, H.C.; PARKER, C.D. Development of an improved quantitative relation. between bacterial die-off, design and operational factors for anaerobic-aerobi maturation type lagoon systems. *Progress in water technology*, v.11, n.4/5, p. 427, 1979.

SHILTON, A. (Editor). *Pond treatment technology*. London: IWA Publishing, 496p, 2006.

SHUVAL, H. I., LAMPERT, Y., FATTAL, B.; Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. *Water science and technology*, v.35, n.11-12, p.15-20, 1997.

SOARES, J.; SILVA, S. A.; de OLIVEIRA, R.; ARAÚJO, A. L. C.; MARA, D. D.; PEARSON, H. W. Ammonia removal in a pilot-scale WSP complex in northeast Brazil. *Water science technology*, v. 33, n. 7, p. 165-171, 1996.

SPIEGEL, M. R.; tradução e revisão técnica: CONSENTINO P., *Estatística*. 3ª. ed. São Paulo: Makron Books, 1993. 643 p.

SPIEGEL, M. R., SCHILLER, J., SRUNIVASAN, R. A.; *Probabilidade e estatística*. 2ª ed. Porto Alegre: Bookman, 2004. 398 p.

TANAKA, H., ASANO, T., SCHOEDER, E., D., TCHOBANOGLIOUS. G.; Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. *Water enviromental. research*, v. 70, p.39-51, 1998.

USEPA - UNITED STATES ENVIRONMENTAL PROTECTION AGENCY. *Guidelines for water reuse*. USEPA: Washington DC, 2004.

VAN GINNEKEN, M., ORON, G. Risk assessment of consuming agricultural products irrigated with reclaimed wastewater: an exposure model, *Water resources research*, v. 36, n. 9, p. 2691–2699, 2000.

VON SPERLING, M; BASTOS, R. K. X.; KATO, M. T; Removal of *E. coli* and helminthes eggs in UASB: polishing pond systems in Brazil. *Water science*, v.51, n 12. p. 91-97, 2005.

VON SPERLING, M.; Design of facultative ponds based on uncertainty analysis. *Water science and technology*, v. 33, n 7, p 41–47, 1996.

VOSE, D. *Risk analysis*. 3<sup>rd</sup> ed. Chocherter, UK: John Wiley & Sons, 735p. 2008.

VON SPERLING, M. *Introdução à qualidade das águas e ao tratamento de esgotos*. 3. ed. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, 1996, 452p. (Princípios do tratamento biológico de águas residuárias, 1).

VON SPERLING, M.; JORDÃO, E. P.; KATO, M.T.; ALEM SOBRINHO, P.; BASTOS, R. K. X.; PIVELLI, R. P. Lagoas de Estabilização. In: GONÇALVES, R. F. (Coord.). *Desinfecção de efluentes sanitários*. Rio de janeiro: ABES, RiMa, 2003, p. 275-336 (Projeto PROSAB).

VON SPERLING, M. *Lagoas de estabilização*. 2. ed. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, 2002, 196p. (Princípios do tratamento biológico de águas residuárias, 3).

VON SPERLING, M. Performance evaluation and mathematical modelling of coliform die-off in tropical and subtropical waste stabilization ponds. *Water research*, v.33, n. 6, p. 1435-1448, 1999.

VON SPERLING, M. *Princípios básicos do tratamento de esgotos*. Belo Horizonte: Departamento de engenharia sanitária e ambiental, 1996, 211p. (Princípios do tratamento biológico de águas residuárias, 2).

WERNER, T. M., KADLEC, R. H. Stochastic simulation of partially-mixed, event driven treatment wetlands. *Ecological engineering*. v. 14, n 3, p. 253–267, 2000.

WORLD HEALTH ORGANIZATION. *Health guidelines for the use of wastewater in agriculture an aquaculture*. Geneva: WHO. 1989.

WORLD HEALTH ORGANIZATION. *Guidelines for drinking - water quality*. 3<sup>rd</sup> ed. Geneva: WHO, 2006a, 515p. Disponível em: < [http://www.who.int/water\\_sanitation\\_health/dwq/gdwq3rev/en/index.html](http://www.who.int/water_sanitation_health/dwq/gdwq3rev/en/index.html) > Acesso em: 2008.

WORLD HEALTH ORGANIZATION. *Guidelines for the safe use of wastewater, excreta and greywater*. Volume 2. *Wastewater use in agriculture*. Geneva: WHO. 2006b, 213p. Disponível em: <[http://www.who.int/water\\_sanitation\\_health/wastewater/gsuweg2/en/index.html](http://www.who.int/water_sanitation_health/wastewater/gsuweg2/en/index.html)> Acesso em: 2008.

YANEZ, F. *Lagunas de estabilizacion. Teoria, diseño y mantenimiento*. Cuenca, Ecuador: ETAPA, 1993.

ZWIETERING, M. H., VAN GERWEN, J. C.. Sensitivity analysis in quantitative microbial risk assessment. *International journal of food microbiology*. v. 58, n. 3, p. 213–221, 2000.