



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE TECNOLOGIA**  
**DEPARTAMENTO DE ENGENHARIA HIDRÁULICA E AMBIENTAL**  
**PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL – SANEAMENTO AMBIENTAL**

**JOÃO IGOR DA ROCHA LEITÃO**

**QUALIDADE DE EFLUENTE – GARANTIA PROBABILÍSTICA, *FUZZY* E  
AVALIAÇÃO DE CAPABILIDADE DO SISTEMA DE TRATAMENTO**

**FORTALEZA**  
**2018**

JOÃO IGOR DA ROCHA LEITÃO

QUALIDADE DE EFLUENTE – GARANTIA PROBABILÍSTICA, *FUZZY* E  
AVALIAÇÃO DE CAPABILIDADE DO SISTEMA DE TRATAMENTO

Tese apresentada ao Programa de Pós-Graduação em Engenharia Hidráulica e Ambiental da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia Civil. Área de concentração: Saneamento Ambiental.

Orientador: Prof. Dr. Fernando José Araújo da Silva

FORTALEZA  
2018

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- L548q    Leitão, João Igor da Rocha.  
          Qualidade de efluente – garantia probabilística, fuzzy e avaliação de capacidade do sistema de tratamento / João Igor da Rocha Leitão. – 2018.  
          124 f. : il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Civil: Saneamento Ambiental, Fortaleza, 2018.  
          Orientação: Prof. Dr. Fernando José Araújo da Silva.
1. Bootstrap. 2. Controle Estatístico de Processos. 3. Estação de Tratamento de Esgoto. 4. Intervalo de Confiança . 5. Risco Fuzzy. I. Título.

CDD 628

---

JOÃO IGOR DA ROCHA LEITÃO

QUALIDADE DE EFLUENTE – GARANTIA PROBABILÍSTICA, *FUZZY* E  
AVALIAÇÃO DE CAPABILIDADE DO SISTEMA DE TRATAMENTO

Tese apresentada ao Programa de Pós-Graduação em Engenharia Hidráulica e Ambiental da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia Civil. Área de concentração: Saneamento Ambiental.

Aprovada em: 06/09/2018

BANCA EXAMINADORA

---

Prof. Dr. Fernando José Araújo da Silva (Orientador)  
Universidade Federal do Ceará – UFC

---

Prof. Dr. Ronaldo Stefanutti  
Universidade Federal do Ceará – UFC

---

Prof. Dra. Marisete Dantas de Aquino  
Universidade Federal do Ceará – UFC

---

Prof. Dr. Emerson Mariano da Silva (Externo)  
Universidade Estadual do Ceará - UECE

---

Prof. Dr. José Maria Brabo Alves (Externo)  
Universidade Estadual do Ceará - UECE

## AGRADECIMENTOS

Agradeço primeiramente a Deus pela oportunidade de estar desenvolvendo este trabalho e por Ele haver disposto todas as ferramentas e conhecimento que tornou este trabalho possível.

Dedico este trabalho a minha família (João Leitão, Ivelise Leitão e Ivna Leitão) pela base sólida e apoio em toda a minha existência.

Agradeço aos meus amigos e discentes do Departamento de Engenharia Hidráulica e Ambiental, como uma menção especial ao Ismael Carlotto Lopes, Mário Barros Ubirajara, Ítalo Lima e Riamburgo Gomes.

Agradeço ao corpo docente do Departamento de Engenharia Hidráulica e Ambiental por tanto conhecimento que estes colocaram a disposição. Com menção especial ao meu orientador o Dr. Fernando José Araújo da Silva e o demais professores do departamento. Um agradecimento especial ao professor André Beserra, Suetônio Mota, José Capelo e Ana Bárbara Araújo.

A equipe do departamento de Estatística com menção especial aos docentes Ana Maria e Silvia Freitas e os alunos João Arrais, Gabriel Fernandez, Rafael Oliveira e Paulo Reineiro.

Concluo os meus agradecimentos a participação especial do docente e amigo dr. Raimundo de Souza (*in memoriam*) a minha formação pessoal e acadêmica. Em meus quase 30 anos de vida nunca cheguei a conhecer homem com tantas virtudes admiráveis que tornaram a minha permanência no Departamento de Hidráulica e Ambiental (DEHA-UFC) tão prazerosa.

“Não a nós, SENHOR, não a nós, mas ao teu nome dá glória, por amor da tua benignidade e da tua verdade.”

Salmo 115.1 – Almeida Corrigida e Revisada Fiel

## RESUMO

A modelagem matemática é uma ferramenta utilizada para prever a ocorrência e a intensidade de fenômenos relevantes. Desta maneira fornecer ferramentas para o responsável por um determinado processo se precaver de situação adversas ou tirar o máximo proveito de situações favoráveis. A performance de Estações de Tratamento de Esgoto (ETE) é comumente susceptível a uma grande aleatoriedade dos dados. Devido, principalmente, a flutuação nos valores de vazão, carga orgânica, parâmetros operacionais, intempéries climáticas, etc. Logo, utilizar a modelagem matemática para prever a performance da remoção de poluentes pela ETE é um trabalho que traz um erro intrínseco da natureza aleatória dos dados. Para isto, utiliza-se Intervalo de Confiança (IC) que delimita uma faixa de valores no qual o valor esperado provavelmente é encontrado. A construção de IC pelo método convencional é demorada e onerosa já que exige a retirada de muitas carreiras de amostras da população em estudo. Torna-se necessário a adoção de ferramentas robustas para prever satisfatoriamente a ocorrência dos fenômenos de remoção de poluentes em ETE. Visando contornar esta dificuldade utiliza-se a Metodologia *Bootstrap* para a construção de IC, com esta ferramenta são levantados cenários antagonicos (bandas de confiança dos IC). O risco de não conformidade do tratamento de esgoto é comumente analisada utilizando o Risco *Fuzzy* associado a um número *Fuzzy* Triangular (NFT) que é uma metodologia simples e de fácil interpretação. É realizado o esforço de integrar o risco probabilístico obtido através da metodologia *Bootstrap* com o risco *Fuzzy* associado ao NFT. Levantado o panorama de confiabilidade relacionada ao tratamento de efluentes através da metodologia probabilística e *Fuzzy* tem-se a necessidade de aferir se o processo é apto a satisfazer um dado objetivo. Neste contexto surge as ferramentas descritas no Controle Estatístico de Processos (CEP). Através desta metodologia foram levantados IC para a garantia da conformidade de parâmetros de qualidade de efluente frente as normas ambientais.

**Palavras-chave:** *Bootstrap*, Controle Estatístico de Processos (CEP), Estação de Tratamento de Esgoto (ETE), Intervalo de Confiança (IC), Reamostragem e Risco *Fuzzy*.

## ABSTRACT

Mathematical modeling is a tool used to predict the occurrence and intensity of relevant phenomenon. In this way it provides tools for the accountable in charge of a given process to guard against adverse situations or to take full advantage of favorable situations. The performance of Sewage Treatment Stations (STS) is commonly susceptible to a large randomness of data. Due mainly to the fluctuation in flow values, organic load, operational parameters, climatic weather, etc. Therefore, using mathematical modeling to predict the performance of STS pollutants is a work that brings an intrinsic error of the random nature of the data. For this, we use a confidence interval (CI) that delimits a range of values in which the expected value is probably found. The construction of CIs by the conventional method is time-consuming and costly since it requires the withdrawal of many rows of samples from the study population. It is necessary to adopt robust tools to predict satisfactorily the occurrence of the pollutant removal phenomena in STS. In order to overcome this difficulty, the *Bootstrap* methodology for the construction of CIs is used, with this tool, the antagonistic scenarios (IC confidence bands) are raised. The risk of non-compliance of sewage treatment is commonly analyzed using *Fuzzy* Risk associated with a Triangular *Fuzzy* Number (TFN) that is a simple and easy to interpret methodology. The effort to integrate the probabilistic risk obtained through the *Bootstrap* methodology with the *Fuzzy* risk associated with NFT is made. Raising the reliability scenario related to the treatment of effluents through the probabilistic and *Fuzzy* methodology has the need to gauge if the process is able to satisfy a given objective. In this context, the tools described in the Statistical Process Control (SPC) emerge. Through this methodology IC was established to guarantee compliance of effluent quality parameters against environmental standards.

**Keywords:** *Bootstrap*, Statistical Process Control (CEP), Sewage Treatment Station (ETE), Confidence Interval (CI), Resampling and *Fuzzy* Risk.

## LISTA DE ILUSTRAÇÕES

Figura 1	Limite inferior controle (LIC), limite superior controle (LSC), valor central (LC), ponto fora de aceitação (vermelho) e pontos com comportamento anômalo amarelo).....	25
Figura 2	Representação de curvas com FDP normal com diferentes valores de médias e desvio padrão ( $\mu, \sigma$ ).....	30
Figura 3	Representação de curvas com FDP lognormal com diferentes valores de logaritmo as médias e do desvio padrão ( $\log\mu, \log\sigma$ ).....	30
Figura 4	Representação de curvas com FDP <i>gama</i> com diferentes valores de $\gamma$ e $\beta$ .	31
Figura 5	Representação de curvas com FDP <i>Weibull</i> com diferentes valores de <i>forma e escala</i> .....	31
Figura 6	Representação de curvas com FDP <i>Gumbel</i> com diferentes valores de <i>locação e escala</i> .....	32
Figura 7	Representação de curvas com FDP <i>Rayleigh</i> com diferentes valores de <i>escala</i> .....	32
Figura 8	Representação de curvas com FDP <i>Frechet</i> com diferentes fatores de <i>escala e forma</i> .....	33
Figura 9	Teste de <i>Kolmogorov-Smirnov</i> entre duas distribuições com o $D_{\max}$ em vermelho.....	34
Figura 10	Simulação Monte Carlo (esquerda) vs Simulação <i>Bootstrap</i> (direita).....	37
Figura 11	Histograma com a estatística de interesse obtido através do método convencional de amostragem.....	39
Figura 12	Histograma com a estatística de interesse obtido através da metodologia <i>Bootstrap</i> de amostragem.....	40
Figura 13	Representação do número difuso triangular (NFT).....	46
Figura 14	Risco Difuso associado ao NFT.....	47
Figura 15	Histograma representando a dispersão de dados para dois processos com valor de $C_p$ elevado (esquerda) e $C_p$ baixo (direita).....	50

Figura 16	$C_{pk}$ em conformidade (valores elevados) e em não conformidade (valores baixos) com os limites a direita e a esquerda da concentração de dados.....	51
Figura 17	Modelo Caixa Preta para ETE.....	53
Figura 18	Modelo esquemático do sistema de lagoas de estabilização.....	53
Figura 19	Representação metodológica para o teste de aderência.....	59
Figura 20	Representação metodológica Bootstrap para o cálculo dos IC para a garantia do tratamento de efluentes.....	60
Figura 21	Representação metodológica para a construção dos índices de capacidade utilizando uma FDP geradora de cenários.....	61
Figura 22	Representação metodológica para o cálculo do risco Fuzzy.....	61
Figura 23	Valores plotados em um gráfico Função Distribuição Acumulada <i>versus</i> DBO efluente para os dados mostrados pela Tabela 8.....	63
Figura 24	IC para a FDA simulada utilizando o Bootstrap paramétrico para o Esgoto Bruto.....	65
Figura 25	IC com base em <i>Bootstrap</i> construídos para a DBO efluente proveniente das lagoas facultativas primárias.....	67
Figura 26	IC construídos para a DBO efluente proveniente das lagoas de maturação primária.....	69
Figura 27	IC construídos para a DBO efluente proveniente das lagoas de maturação secundária.....	71
Figura 28	Valores dos limites inferiores dos IC (pior cenário possível) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> .....	71
Figura 29	Valores dos limites médios dos IC (cenário mais conservador) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> .....	72
Figura 30	Valores dos limites superiores dos IC (cenário mais otimista) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> .....	72

Figura 31	Teste de Kolmogorov-Smirnov realizado através do Software <i>Rstudio</i> comparando as Função de Distribuição Acumulada.....	75
Figura 32	Pares de parâmetros Weibull levantados para cada reamostragem (a esquerda) e IC com as bandas de confiança com 95% das curvas.....	80
Figura 33	IC <i>Bootstrap</i> paramétrico para a DBO (mg/L) do EB afluente as estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> (da esquerda para a direita).....	81
Figura 34	IC <i>Bootstrap</i> paramétrico para a DBO (mg/L) do LFP efluente as estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> (da esquerda para a direita).....	82
Figura 35	IC <i>Bootstrap</i> paramétrico para a DBO (mg/L) do LMP efluente as estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> (da esquerda para a direita).....	83
Figura 36	IC <i>Bootstrap</i> paramétrico para a DBO (mg/L) do LMS efluente as estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> (da esquerda para a direita).....	84
Figura 37	Função distribuição acumulada para o efluente proveniente da lagoas facultativa para a estação PS <sub>5</sub> utilizando a metodologia <i>Bootstrap</i> .....	87
Figura 38	Função distribuição acumulada para o efluente proveniente da lagoas facultativa para a estação PS <sub>6</sub> utilizando a metodologia <i>Bootstrap</i> .....	89
Figura 39	Função distribuição acumulada para o efluente proveniente da lagoas facultativa para a estação PS <sub>9</sub> utilizando a metodologia <i>Bootstrap</i> .....	92
Figura 40	NFT's associado aos intervalos de confiança.....	96
Figura 41	Representação do risco <i>Fuzzy</i> utilizando a metodologia <i>Bootstrap</i> para a construção dos intervalos de confiança para os parâmetros DBO <sub>f</sub> , DQO <sub>f</sub> , AMT e CTT para a lagoa facultativa afluente a estação PS <sub>5</sub> .....	99
Quadro 1	Detalhamento das características do estudo.....	52

## LISTA DE TABELAS

Tabela 1	$C_p$ e a classificação do sistema.....	26
Tabela 2	$C_p$ e $C_{pk}$ para classificação de sistemas.....	27
Tabela 3	Probabilidade relacionado a ocorrência de eventos.....	43
Tabela 4	Severidade de um evento relacionado com a probabilidade.....	43
Tabela 5	Enquadramento e classes de qualidade.....	48
Tabela 6	ETE utilizadas para a confecção deste trabalho (PS <sub>1</sub> a PS <sub>10</sub> ).....	54
Tabela 7	Intervalo de Confiança Inferior (2,5%) e Superior (97,5%) para cada decil.....	63
Tabela 8	Estatística descritiva para o EB afluyente e efluente as lagoas LFP, LMP e LMS para as Estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> .....	75
Tabela 9	Parâmetros das FDP estimadas para a DBO no EB e nos efluentes do estudo para as estações PS <sub>5</sub> , PS <sub>6</sub> e PS <sub>9</sub> .....	76
Tabela 10	Valor das estatísticas (p-value) com efluente tratado na saída das ETE...	77
Tabela 11	Valores de AIC dos modelos para a matéria orgânica em forma de DBO	78
Tabela 12	Melhor Função Distribuição de Probabilidade para cada série de dados..	79
Tabela 13	Valores de referência para avaliação do desempenho dos sistemas de lagoas .....	85
Tabela 14	Teste de aderência usando o teste de KS e Akaike's para a ETE PS <sub>5</sub> .....	86
Tabela 15	Intervalo de confiança para a garantia <i>Bootstrap</i> ETE PS <sub>5</sub> .....	86
Tabela 16	Teste de aderência usando o teste de KS e Akaike's para a ETE PS <sub>6</sub> .....	88
Tabela 17	Intervalo de confiança para a garantia <i>Bootstrap</i> ETE PS <sub>6</sub> .....	88
Tabela 18	Teste de aderência usando o teste de KS e Akaike's para a ETE PS <sub>9</sub> .....	84
Tabela 19	Intervalo de confiança para a garantia <i>Bootstrap</i> ETE PS <sub>9</sub> .....	90
Tabela 20	FDP encontradas usando o teste de KS.....	91
Tabela 21	FDP encontradas usando o teste de AIC.....	91
Tabela 22	Risco <i>Fuzzy</i> associado aos limites de Confiança <i>Bootstrap</i> .....	93

Tabela 23	Quantis dos valores da DBO efluente e seus respectivos valores de IC.	94
Tabela 24	Risco <i>Fuzzy</i> associado aos quantis.....	97
Tabela 25	Risco <i>Fuzzy</i> para o efluente proveniente da LFP da ETE PS <sub>5</sub> .....	98
Tabela 26	Índices de Capabilidade para as saídas ou efluentes das lagoas que compõem a ETE PS <sub>5</sub> .....	103
Tabela 27	Coeficientes de capabilidade dos parâmetros de qualidade (AMT, pH e CTT, DBO, DBOF, DQO e DQOF) efluente as lagoas de estabilização pertencentes a estação PS <sub>5</sub> .....	104

## LISTA DE ABREVIATURAS OU SIGLAS

A	Área
ABNT	Associação Brasileira de Normas Técnicas
AIC	Akaike Information Criterion
AMT	Amônia Total
CAGECE	Companhia de Água e Esgoto do Estado do Ceará
CEP	Controle Estatístico de Processos
C <sub>p</sub>	Índice de Capacidade
C <sub>pk</sub>	Índice de Performance
CTT	Coliformes Termotolerantes
C/L	Razão Comprimento Largura
CECR	Concentração no Corpo Receptor
COEMA	Conselho Estadual de Meio Ambiente
CONAMA	Conselhor Nacional do Meio Ambiente
DBO	Demanda Bioquímica de Oxigênio
DBO <sub>F</sub>	Demanda Bioquímica de Oxigênio Filtrada
DQO	Demanda Química de Oxigênio
DQO <sub>F</sub>	Demanda Química de Oxigênio Filtrada
EB	Esgoto Bruto
ETE	Estação de Tratamento de Esgoto
FDA	Função Distribuição Acumulada
FDP	Função Distribuição de Probabilidade
IC	Intervalo de Confiança
KS	<i>Kolmogorov Smirnov</i>
LFP	Lagoa Facultativa Primária
LMP	Lagoa de Maturação Primária
LMS	Lagoa de Maturação Secundária
LIC	Limite Inferior de Controle
LSC	Limite Superior de Controle
LC	Limite Central
NFT	Número de <i>Fuzzy</i> Triangular
OD	Oxigênio Dissolvido
p	<i>p-value</i>

pH	potencial hidrogeniônico
TDH	Tempo de Detenção Hidráulico
V	Volume

## SUMÁRIO

1	INTRODUÇÃO .....	18
1.1	Importância do controle de efluentes.....	18
1.2	Necessidade ferramentas confiáveis e robustas.....	21
1.3	Objetivos.....	22
1.3.1	<i>Objetivos gerais</i> .....	22
1.3.2	<i>Objetivos específicos</i> .....	22
1.4	Estrutura da tese.....	22
2	REFERENCIAL TEÓRICO .....	24
2.1	Controle Estatístico de Processos (CEP).....	24
2.2	Distribuição de probabilidade em controle de qualidade ambiental.....	28
2.3	Teste de aderência ( <i>Kolmogorov-Smirnov</i> ).....	33
2.3.1	<i>Akaike Information Criterion (AIC)</i> .....	35
2.4	Simulação de dados.....	36
2.4.1	<i>Monte Carlo e Bootstrap</i> .....	36
2.5	Risco e garantia –enfoque sobre a qualidade de efluente.....	40
2.5.1	<i>Abordagem probabilística</i> .....	44
2.5.2	<i>Abordagem Fuzzy</i> .....	45
2.5.3	<i>Grau de conformidade ambiental</i> .....	47
2.5.4	<i>Capacidade de processo</i> .....	49
3.	METODOLOGIA.....	52
3.1	Tipo de estudo.....	52
3.2	Estações de tratamento de esgoto e parâmetros analíticos.....	52
3.3	Ferramentas estatísticas.....	54
3.4	Risco probabilístico e simulação Bootstrap.....	55
3.5	Controle estatístico de processo e o grau de conformidade.....	56
3.6	Risco Fuzzy e garantia probabilística.....	57
3.7	Análise Esquemática Metodológic.....	58
3.7.1	<i>Metodologia Esquemática Bootstrap para a Contrução dos IC de Garantia</i> .....	59
3.7.2	<i>Metodologia Esquemática CEP para o cálculo dos Índices de Capabilidade</i> ....	59
3.7.3	<i>Metodologia Esquemática para o Teste de Aderência</i> .....	60
3.7.4	<i>Metodologia Esquemática para o cálculo do Risco Fuzzy</i> .....	61
4	RESULTADOS E DISCUSSÃO.....	62

4.1	<i>Bootstrap</i> .....	62
4.1.1	<i>Bootstrap não paramétrico</i> .....	62
4.2	<i>Bootstrap paramétrico</i> .....	74
4.2.1	<i>Estimativa e parâmetros e teste de aderência</i> .....	74
4.2.2	<i>IC do bootstra paramétrico</i> .....	79
4.2.3	<i>Teste de aderência para os demais parâmetros e qualidade e garantia Bootstrap</i> .....	85
4.3	<b>Confiabilidade do tratamento com metodologia <i>Fuzzy</i></b> .....	93
4.3.1	<i>Confiabilidade do tratamento com metodologia <i>Fuzzy</i> – demais parâmetros</i>	97
4.4	<b>Controle estatístico de processos</b> .....	101
4.4.1	<i>Função distribuição de probabilidade (FDP) e teste de aderência</i> .....	101
4.4.2	<i>Índice de capacidade do Processo</i> .....	101
4.4.3	<i>Índice de Capabilidade do Processo – Demais parâmetros</i> .....	103
5	<b>CONCLUSÃO</b> .....	106
	<b>REFERÊNCIAS</b> .....	108
	<b>APÊNDICE A - TESTE DE ADERÊNCIA PARA AS FDP DO EFLUENTE AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub></b> .....	115
	<b>APÊNDICE B – IC PARA A GARANTIA BOOTSTRAP PARA OS PARÂMETROS DE QUALIDADE EFLUENTE AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> E PS<sub>9</sub></b> .....	117
	<b>APÊNDICE C - ÍNDICES DE CAPABILIDADE DAS ETE PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub></b>	119
	<b>APÊNDICE D – APÊNDICE D – GARANTIA <i>FUZZY</i> PARA AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub></b> .....	122
	<b>APÊNDICE E – <i>SCRIPT</i> USADO PARA GERAR OS RESULTADOS NO <i>SOFTWARE R</i></b> .....	123

## **1. INTRODUÇÃO**

### **1.1 Importância do controle de efluentes**

Uma forma de mitigar os efeitos deletérios da poluição hídrica é proporcionar coleta e tratamento adequado para os efluentes provenientes das diversas fontes de geração (SILVA et al., 2014). As Estações de Tratamento de Esgoto (ETE) são equipamentos destinados à adequação dos efluentes a níveis apreciáveis à capacidade de suporte dos sistemas ambientais, em consonância com as normas ambientais. Enseja-se assim, a minimização de efeitos nocivos ao meio ambiente e à saúde da população. Este fato alarmante é mostrado por Pimpunchat et al. (2009) que estimaram que cerca de 25 milhões de pessoas morriam por ano devido a doenças vinculadas a contaminação hídrica na primeira década deste milênio.

O sistema de esgotamento sanitário diz respeito a atividades, infraestrutura e instalações operacionais de coleta, transporte, tratamento e disposição final adequado do esgoto sanitário. Isto é definido pela Política Nacional de Saneamento Básico (Lei nº 11445, de 5 de janeiro de 2007). O esgotamento sanitário junto com o abastecimento de água, a limpeza urbana e manejo de resíduos sólidos e a drenagem de águas pluviais formam as atividades e infraestrutura referentes ao saneamento básico.

O estatuto do saneamento – ver Capítulo VII, Artigo 43 – torna claro a necessidade de níveis de tratamento condizentes com os padrões das classes dos corpos hídricos nos quais os efluentes são lançados, de modo a viabilizar o uso múltiplo da água, sempre respeitando os limites financeiros estipulados pelo usuário e pelos agentes da política ambiental (YANG et al., 2007).

Para reduzir os riscos à saúde do homem e ao meio ambiente, as Estações de Tratamento de Esgoto devem ser capazes de garantir qualidade adequada para os efluentes, tanto para o despejos quanto para o reuso, quando esta prática for pretendida. Neste último caso a garantia deve ser tal que os riscos microbiológicos e químicos sejam aceitáveis (SHANNON et al., 2008).

A garantia de qualidade para reuso e despejo depende da eficácia da tecnologia de tratamento utilizada. Taheriyoun e Moradinejad (2015) destacam que a insuficiência no tratamento de efluente traz sérios problemas à saúde pública e ao meio ambiente, como a contaminação com dos recursos hídricos, degradação da qualidade da água e disseminação de doenças. Logo, um processo de tratamento satisfatório, em termos de remoção e

adequabilidade do efluente, envolve a conformidade de diferentes parâmetros de qualidade com um conjunto complexo de padrões, de maneira que o sistema de tratamento seja capaz de satisfazer às necessidades esperadas ao longo do funcionamento.

Em locais sujeitos a estresse hídrico recorrente, como em regiões áridas e semiáridas, o esgoto pode causar impacto mais acentuado. Porém, quando adequadamente tratado, torna-se recurso estratégico (MIZYED, 2013). Entretanto, este recurso, o efluente tratado, sofre com aleatoriedade na remoção de poluentes nas unidades de tratamento. Isto se deve a fatores externos e internos às ETE, que muitas vezes impedem que um determinado padrão seja alcançado. Üstün (2009) afirma que para uma remoção satisfatória de poluentes, além dos fatores biológicos e físico-químicos do tratamento, o *design* da estação e as condições de operação são de grande importância. Esta afirmativa revela a complexidade do fenômeno de remoção de poluentes em ETE, independente da tecnologia empregada.

Guerrini et al. (2016) ressaltam que uma análise adequada da performance de Estações de Tratamento de Esgoto é parte fundamental quanto ao desempenho de subprocessos específicos de serviços de água e esgotos. Esta preocupação levou Hsiao et al. (2007) a estudarem a eficiência de ETE e seus fatores determinantes. Uma das principais fontes de incertezas e desafios a ser enfrentado é a variabilidade intrínseca do afluente (BELIA et al., 2016). É relevante, portanto, que a ETE apresenta certa resiliência, de forma a implicar pouca variabilidade no efluente, sem prejuízo sobre as metas de qualidade.

As ETE devem dispor de flexibilidade operacional, com a adequação do efluente tratado às características desejadas, minimizando o risco e as principais falhas recorrentes do processo de remoção de poluentes. As ETE são susceptíveis a variações decorrentes do clima (destaque para temperatura e precipitação). Essas alterações podem trazer riscos ao bom funcionamento do equipamento, com reflexo sobre a qualidade do efluente final. Para a maior segurança ao tratamento de águas residuárias, Zhang et al., (2012) propuseram um modelo para a avaliação do desempenho de ETE. Os autores traduziram a performance de uma estação de tratamento em termos de confiabilidade, robustez e resiliência, a saber:

- *Confiabilidade*: capacidade de a estação de tratamento desempenhar sua função durante um período de tempo pré-determinado;
- *Robustez*: capacidade de a estação de tratamento funcionar corretamente em situação de estresse/adversa;
- *Resiliência*: capacidade da estação de tratamento em persistir estável frente a uma perturbação externa.

Uma maior atenção deve ser dada ao controle do despejo de efluentes, incluindo contaminantes emergentes, tais como: hidrocarbonetos aromáticos, bifenil policlorados, metais pesados, pesticidas, triclosan, surfactantes, plastificantes, hormônios e disruptores endócrinos (DÍAZ-GARDUÑO et al., 2017; GUO; IWATA, 2017; GARCIA et al., 2013). Estes, somados aos parâmetros normalmente analisados são fontes de não conformidade quanto aos padrões apresentados na legislação ambiental.

Devido à sazonalidade do efluente e a alteração de alguns fatores operacionais é conveniente adotar metodologias que permitam avaliar a performance da remoção de poluentes. A maneira corrente é referida aos parâmetros demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), sólidos suspensos totais (SST), nitrogênio (amônia, orgânico, nitrato e nitrito), fósforo (total, dissolvido, orgânico e inorgânico), organismos indicadores de contaminação microbiana (e.g. coliformes temortolerantes e *Escherichia coli*) e etc.

Em geral a avaliação é feita com base no percentual de remoção e no grau de conformidade com a legislação (UPADHYAYA; MOORE, 2012). Também, é importante realizar um número mínimo pré-determinados de coletas em modelo tipo caixa preta (entrada vs saída) (RODRIGUEZ-GARCIA et al., 2011) e analisar a remoção de poluentes em função dos custos envolvidos (BENEDETTI et al., 2008). Vale lembrar que os parâmetros analisados, número de amostras por ciclo de tempo e a remoção mínima exigida são peculiares de cada arcabouço normativo. Nesse contexto, é importante avaliar, muitas vezes, a performance de ETE com base índices que condensem este conjunto de informações (i.e. remoção de poluentes, atendimento normativo e frequência de monitoramento). Uma abordagem deste tipo é observada nos estudos de Silva et al. (2014), Flores-Alsina et al. (2008) e Oliveira e von Sperling (2008). Entretanto, tal enfoque ainda é insipiente se comparado à aplicação usual quanto à qualidade da água no ambiente ou distribuída por sistemas de abastecimento. Ou seja, comparando as normas de outras águas (bruta e tratada), os efluentes recebem pouquíssima atenção. Isto porque, diferentemente da água de abastecimento ou da água bruta, os efluentes constituem parte de um serviço e não um produto.

## 1.2 Necessidade de ferramentas confiáveis e robustas

O desenvolvimento de modelos de avaliação e previsão da eficiência de ETE são úteis para identificar fatores críticos que influem na remoção de poluentes. Funcionam como ferramentas importantes para o planejamento estratégico e gestão nas estações, ou conjuntos destas. Para que tal enfoque seja satisfatório em sua função, deve-se considerar as seguintes premissas (MCINTYRE; WHEATER, 2003).

- Identificar os principais fatores que afetam a qualidade do efluente, alguns deles normalmente negligenciados, como as variações climáticas;
- Avaliar o risco de não conformidade do efluente tratado associado com intervenções de gestão alternativas;
- Analisar critérios de gerenciamento alternativos e investigar as reais necessidades de qualidade do efluente;
- Estabelecer limites para os modelos e ter ciência de suas limitações;
- Estabelecer modelos computacionalmente viáveis.

O termo “robusto” refere-se a metodologias que necessitam de poucas ou nenhuma premissa para que a sua utilização seja satisfatória (SALSBURG, 2009). A busca por ferramentas com estas características torna-se de fundamental importância devido as variadas características das tecnologias de tratamento de esgoto disponível. Logo, existe apelo para a criação e a adoção de modelos matemáticos e estatísticos que consigam representar mais fielmente os fenômenos envolvidos no desempenho de ETE, sem restrições metodológicas significativas.

Portanto, é importante a utilização de ferramentas robustas para análise de desempenho de ETE e atendimento de norma, de modo a prever situações anômalas e eventos raros. Desta maneira é possível fornecer um panorama geral da ocorrência de falha e os riscos inerentes dos processos de tratamento. Normalmente os sistemas são avaliados utilizando um modelo simplista, sob a hipótese de que a viabilidade é pequena ou nula, não influenciando na performance do sistema como um todo. Dentre as ferramentas estatísticas mais utilizadas para a avaliação de desempenho destacam-se o Controle Estatístico de Processos (CEP) e a abordagem de simulação através da simulação *Bootstrap* e Monte Carlo.

## 1.3 Objetivos

### 1.3.1 Objetivo geral

Demonstrar a aplicação e discutir o uso de ferramentas estatísticas integrativas para o controle da qualidade e conformidade normativa ambiental de efluentes em sistemas de tratamento de esgoto.

### 1.3.2 Objetivos específicos

- Demonstrar um processo mais assertivo para cômputo de Funções de Distribuição de Probabilidade (FDP) dominantes para os parâmetros analíticos caracterizadores de efluentes;
- Demonstrar a utilização do Controle Estatístico de Processos (CEP) como ferramenta para avaliar o grau de adequabilidade do processo de tratamento de esgotos, a partir dos Índices de Capacidade ( $C_p$ ) de Performance ( $C_{pk}$ ) do processo;
- Demonstrar a aplicação da simulação com *Bootstrap* na determinação de intervalos de confiança quanto ao conteúdo de efluentes;
- Avaliar o grau de não conformidade com cenários otimistas de performance através do Risco *Fuzzy* utilizando o Número *Fuzzy* Triangular (NFT) integrado à simulação *Bootstrap*.

## 1.4 Estrutura da tese

O trabalho tem estrutura formal e está dividido em cinco capítulos. O primeiro trata da introdução do geral, da motivação do estudo, trabalho e dos objetivos do mesmo, culminando com sua organização. O segundo capítulo traz a revisão bibliográfica. Este capítulo subsidia o embasamento do estudo, bem como baliza a discussão dos achados.

O Capítulo 3, discorre sobre a metodologia do trabalho, trazendo a indicação de base documental e a abordagem computacional empregada. Os resultados e discussão são apresentados no Capítulo 4.

Encerra-se o trabalho com o Capítulo 5, em que são apresentadas as conclusões da investigação, com algumas sugestões para estudos futuros. Seguem, logo após, as Referências Bibliográficas e conteúdo de apêndice necessário.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 Controle Estatístico de Processo (CEP)

A ISO (Organização Internacional de Normalização) define qualidade como o “grau de satisfação dos requisitos dado por um conjunto de características intrínsecas” (ABNT, 2015). Direcionando esta definição para o escopo do tratamento de efluentes, tem-se o interesse de mensurar o quão “apto” é um processo para satisfazer os requisitos pré-estipulados pela norma ambiental. Logo, existe a necessidade de identificar as não conformidades no processo a medida que estas vierem a surgir, buscando-se a prevenção ou remediação de situações nocivas à qualidade do tratamento de esgoto.

Neste contexto, um importante recurso é a aplicação do “Controle Estatístico de Processos” (CEP). Lima et al. (2007) definem o CEP como um conjunto de ferramentas de monitoramento de qualidade. O CEP permite o registro de não conformidades no processo à medida que ocorrem. Vale destacar que flutuações nos valores de saída de um dado processo podem ter origem tanto em *fatores particulares* como em *causas especiais de variação*. Quando não existem causas especiais de variação o processo é definido como *estável*.

A análise do processo, se este está sob controle estatístico, é realizado através dos intervalos de aceitação de respostas emitidas pelo sistema, a saber: o Limite Inferior de Controle (LIC) e Limite Superior de Controle (LSC). O cálculo desses parâmetros é mostrado pelas equações 1 e 2, que compreendem a aleatoriedade dos dados dentro das causas particulares de um dado processo. Os limites de controle podem ser estipulados pelo gestor das ETE como uma meta a ser alcançada. Consequentemente, pontos que vão além destes limites são provenientes de causas especiais de variação (SKULJ et al., 2013).

$$LIC = \bar{x} - k\sigma_{\bar{x}} \quad (1)$$

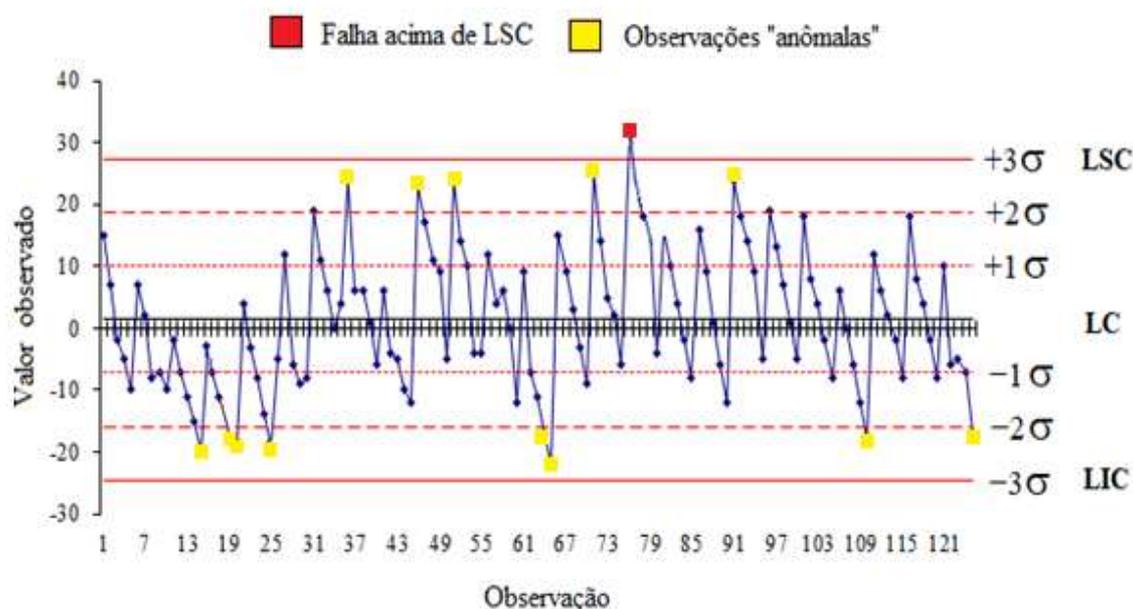
$$LSC = \bar{x} + k\sigma_{\bar{x}} \quad (2)$$

Em que:  $\bar{x}$  é a média;  $\sigma_{\bar{x}}$  é o desvio padrão e  $k$  é o número de desvios-padrão (1, 2 ou 3)

A Figura 1 mostra o comportamento de uma variável com métrica qualquer estimada dentro do intervalo de  $\pm 3\sigma$  na carta de controle. O ponto vermelho representa a

informação/dado observado que está além do limite superior de controle estipulado por  $+3\sigma$ , e os pontos em amarelo são aqueles que possuem algum comportamento anômalo, caso se estipule as zona entre  $-2\sigma$  e  $-3\sigma$  e entre  $+2\sigma$  e  $+3\sigma$  como circunvizinhanças de falha. Através da metodologia de carta de controle de qualidade é possível ter uma representação esquemática da aleatoriedade dos dados à medida que estes ocorrem. Este recurso é amplamente explorado por Berthouex e Hunter (1981) que o utilizam para identificar a ocorrência de sazonalidades, eventos anômalos, tendências, mínimos e máximos em dados afluentes e efluentes de uma ETE.

**Figura 1** – Limite inferior controle (LIC), limite superior controle (LSC), valor central (LC), ponto fora de aceitação (vermelho) e pontos com comportamento anômalo (amarelo).



Fonte: o autor, 2018.

Binny et al. (2017) foram além da utilização convencional do Controle Estatístico de Processos. Os autores analisaram os resultados de saída obtidos utilizando o CEP com alterações premeditadas nos parâmetros intrínsecos do projeto em sistemas de qualidade de serviços médicos. Assim, o CEP é útil para analisar como uma perturbação se propaga no sistema que se deseja analisar.

O CEP é constituído de um ferramental de medição: os índices de capacidade  $C_p$  e  $C_{pk}$  que visam verificar, em que nível, o processo consegue atender ou não um conjunto de especificações buscadas. Ambos os índices possuem uma interpretação intervalar, sendo

o  $C_p$  uma taxa de tolerância à aleatoriedade dos dados. Já o  $C_{pk}$  considera a correta centralização do processo, avaliando se os dados do processo estão inteiramente centrados no intervalo buscado. Logo, o  $C_p$  reflete se a dispersão dos dados pode vir a interferir na conformidade do processo e o  $C_{pk}$  se o intervalo buscado que contém os dados está contido no intervalo desejado. O cálculo do  $C_p$ ,  $C_{pks}$  (superior) e  $C_{pki}$  (inferior) são dados pelas equações 3, 4 e 5 (GEJDOS, 2015).

$$C_p = \frac{LSC-LIC}{6\sigma} \quad (3)$$

$$C_{pks} = \frac{LSC-\bar{X}}{3\sigma} \quad (4)$$

$$C_{pki} = \frac{\bar{X}-LIC}{3\sigma} \quad (5)$$

Gomes (2010) mostra como um sistema é classificado de acordo com os valores de  $C_p$  e  $C_{pk}$ . Vale lembrar que estes limites são dependentes do processo e que os valores mostrados nas Tabelas 1 e 2 são uma orientação geral. Por exemplo, Montgomery (2004) recomenda que para que alguns processos sejam aceitos estes devem possuir  $C_p$  superior a 1,45. Entretanto, estes valores dependem do grau de conformidade que se deseja alcançar. A subjetividade dos valores de referência está na capacidade de se estabelecer metas factíveis e otimizadas pelos custos. Processos com  $C_p$  na faixa de 1,6 costumam apresentar 10 ppm (peças ou unidades por milhão) de amostras em não conformidade em um total de  $10^6$  coletas.

**Tabela 1** –  $C_p$  e a classificação do sistema quanto a dispersão dos dados.

<i>Valores de <math>C_p</math></i>	<i>Interpretação</i>
$C_p < 1$	Processo Incapaz
$1,00 \leq C_p \leq 1,33$	Processo Marginalmente Capaz
$C_p > 1,33$	Processo Capaz

Fonte: Gomes (2010).

**Tabela 2** –  $C_p$  e  $C_{pk}$  para classificação de sistemas quanto a dispersão e localização dos dados.

<i>Valores de <math>C_{pk}</math></i>	<i>Interpretação</i>
$C_{pki} > 1,33$ e $C_{pks} > 1,33$	Processo Capaz
$1,00 \geq C_{pki} \geq 1,33$ e $1,00 \geq C_{pks} \geq 1,33$	Processo Marginalmente Capaz
$C_{pki} < 1,00$ e $C_{pks} < 1,00$	Processo Incapaz
$C_{pki} = C_{pks}$	Processo Centrado
$C_{pki} < C_{pks}$ e $C_{pks} > 1,33$	Fora do Limite Inferior de Especificação
$C_{pks} < C_{pki}$ e $C_{pki} > 1,33$	Fora do Limite Superior de Especificação
$C_{pki} < C_{pks}$	Fora do Limite Inferior de Especificação
$C_{pks} < C_{pki}$	Fora do Limite Superior de Especificação

Fonte: Gomes (2010).

Até o momento os valores dos índices de capacidade foram abordados na sua abordagem bilateral. Ou seja, o intervalo de aceitação está dentro de um intervalo definido e valores fora deste intervalo, superior ou inferior, são tidos como valores aberrantes ou de falha. Entretanto, pode ser de interesse analisar casos em que apenas uma das especificações (inferior ou superior) é importante. Por exemplo, para o oxigênio dissolvido (OD) em água a região de falha é abaixo de um valor determinado e já a matéria orgânica dissolvida (como  $COT_D$ ,  $DBO_D$  ou  $DQO_D$ ) a região de falha é acima de um dado valor (SANCHO et al., 2016). Neste contexto é introduzido o cálculo dos índices de capacidade inferior  $C_{pki}$  e superior  $C_{pkl}$  que são calculados pelas equações (6 e 7) mostradas a seguir:

$$C_{pki} = \frac{\text{Máximo intervalo superior}}{\text{Intervalo superior atual}} = \frac{LCS - \bar{X}}{3\sigma} \quad (6)$$

$$C_{pkl} = \frac{\text{Máximo intervalo inferior}}{\text{Intervalo inferior atual}} = \frac{\bar{X} - LIC}{3\sigma} \quad (7)$$

Somam-se 6 e 7, obtendo a definição do Índice de Capacidade ( $C_p$ ) para ambas as extremidades.

$$C_p = \frac{C_{pki} + C_{pkl}}{2} \quad (8)$$

Através do acompanhamento dos valores do índice de capacidade é possível fazer um planejamento estratégico com uma série de objetivos tangíveis sobre qualidade no tratamento de efluente. As ferramentas disponíveis no CEP tornam possível uma análise acurada do quanto prejudicial a aleatoriedade dos dados pode ser em relação à conformidade com metas pré-estabelecidas. Quanto ao estudo do tratamento de efluentes que apresentam causas aleatórias particulares de cada tecnologia, é possível realizar escolhas com embasamento técnico-científico mais acurado.

## 2.2 Distribuições de probabilidade no controle da qualidade ambiental

As distribuições de probabilidade de estatísticas amostrais possuem utilidade na formulação e construção de intervalos de confiança, estimativa de parâmetros das funções e testes estatísticos de hipóteses (NAGHETTINI, 2007). Descrevem os efeitos aleatórios de variáveis que dependem do acaso. Uma distribuição de probabilidade associa um resultado numérico de um experimento com uma variável aleatória, e torna possível calcular a probabilidade associada a eventos definidos. Esta abordagem é importante, pois com uma variável aleatória *a priori* só é possível mensurar o seu valor após a realização do experimento.

As Funções Distribuição de Probabilidade (FDP) trazem a ideia de medições experimentais aleatórias e imprecisas. Estes conceitos foram introduzidos pelo célebre estatístico Karl Pearson (1857-1936) e representam a “flutuação”, também chamado de “distribuição de números” ou “números dispersos”. Os valores são variáveis independentes em funções matemáticas que informam sobre a probabilidade de um número observado assumir um dado resultado (SALSBURG, 2009). Os resultados dos experimentos ou observações são vistos como valores individuais e aleatórios descritos por uma FDP que podem descrever a natureza e o comportamento dos dados. Ou seja, é um modelo de representação.

A ciência demorou para absorver o conceito de FDP devido, principalmente, a abstração dos conceitos e à limitação de ferramentas computacionais que pudessem ser utilizadas para descrever um dado comportamento de um conjunto numeroso de observações. Logo, este trabalho era demasiadamente demorado e as FDP reduziam-se a uma pequena família de funções que nem sempre conseguiam descrever satisfatoriamente os dados.

Com o advento dos computadores modernos e a rápida evolução dessas máquinas e *softwares*, os estudos de Funções de Distribuição de Probabilidade (FDP) trouxeram funções mais robustas, como a função Gama Generalizada, por exemplo, para a modelagem de fenômenos de Engenharia e Hidrológicos (SILVA et al., 2016). Apesar de resultados satisfatórios, a aderência dos dados obtidos com as FDP deve ser verificada com diferentes funções, sendo usuais as seguintes: *Normal*, *Lognormal*, *Gama*, *Weibull*, *Gumbel*, *Rayleigh* e *Fréchet* (YUAN et al., 2018). Para o contexto aplicado no controle da qualidade tem-se que a a probabilidade de falha pode ser extremamente sensível à função de distribuição que representa os diferentes parâmetros de qualidade do efluente (NIKU; SAMANIEGO, 1983). Logo, a escolha de uma função distribuição de probabilidade deve ser realizada com cuidado para cada série de dados.

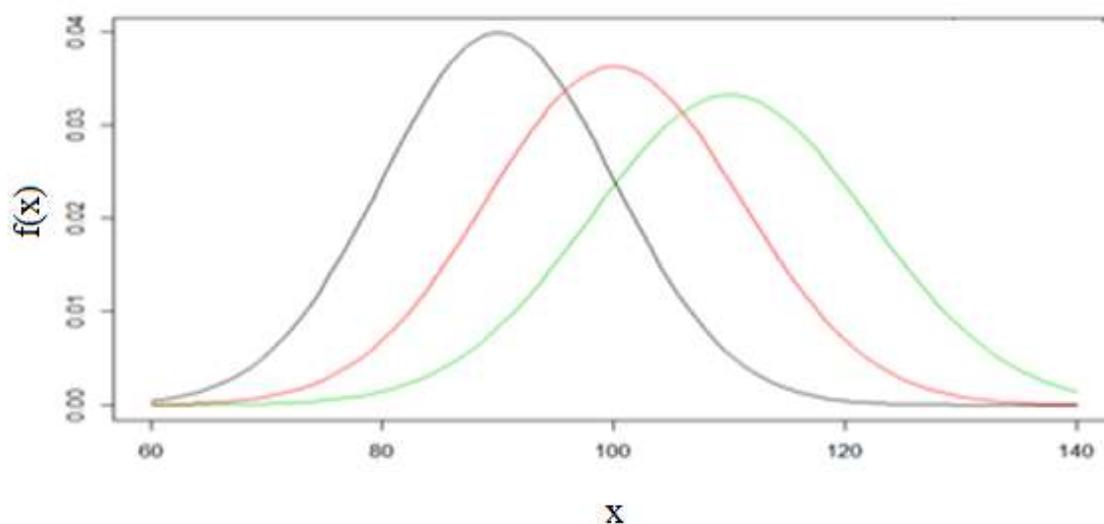
As principais funções distribuição de probabilidade contínuas, referentes aos estudos de fenômenos ambientais, são listadas a seguir com as respectivas curvas, que representam a frequência relativa de um evento (eixo y) com o valor aleatório de ocorrência de resultados por experimento ou observação (eixo x), como mostradas pelas Figuras 2 a 8. As curvas de cada gráfico são geradas da respectiva função distribuição com parâmetros distintos.

### ***Normal***

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (9)$$

Em que:  $\mu$  = média e  $\sigma$  = desvio padrão.

**Figura 2** – Representação de curvas com FDP *Normal* com diferentes valores de médias e desvios padrão ( $\mu$ ,  $\sigma$ ).



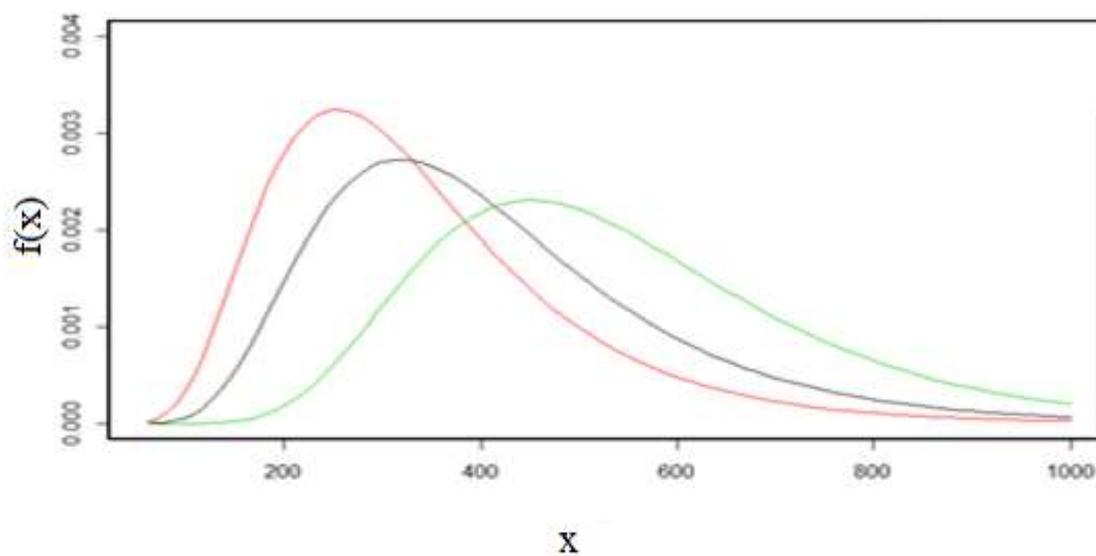
Fonte: o autor, 2018.

### **Lognormal**

$$f(x) = e^{-\frac{(\ln x)^2 / 2\sigma^2}{x\sigma\sqrt{2\pi}}} \quad (10)$$

Em que:  $\ln(\mu)$  = média e  $\ln(\sigma)$  = desvio padrão

**Figura 3** – Representação de curvas com FDP *Lognormal* com diferentes valores de logaritmo as médias e do desvio padrão ( $\log\mu$ ,  $\log\sigma$ ).



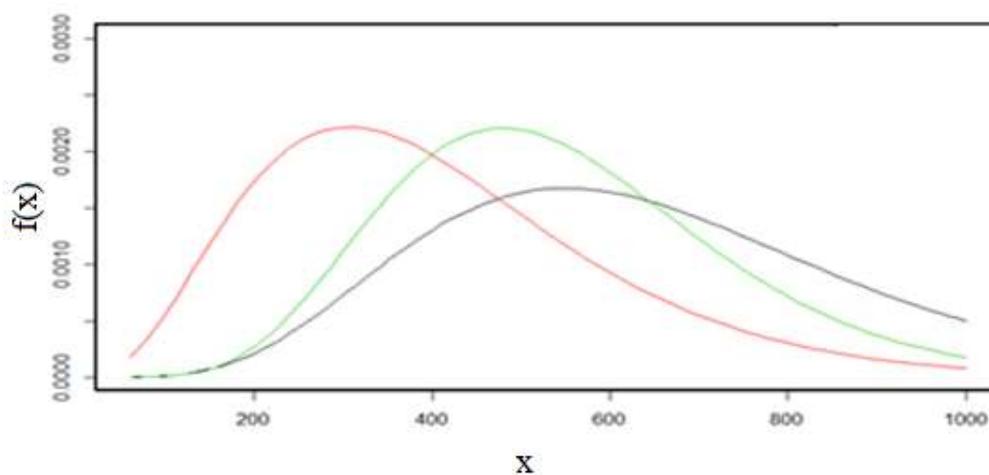
Fonte: o autor, 2018.

### Gama

$$f(x) = \frac{\left(\frac{x-\mu}{\beta}\right)^{\gamma-1} e^{-\frac{x-\mu}{\beta}}}{\beta\tau(\gamma)} \quad (11)$$

Em que:  $\gamma =$  fator de forma e  $\beta =$  fator de escala.

**Figura 4** – Representação de curvas com FDP Gama com diferentes valores de  $\gamma$  e  $\beta$ .



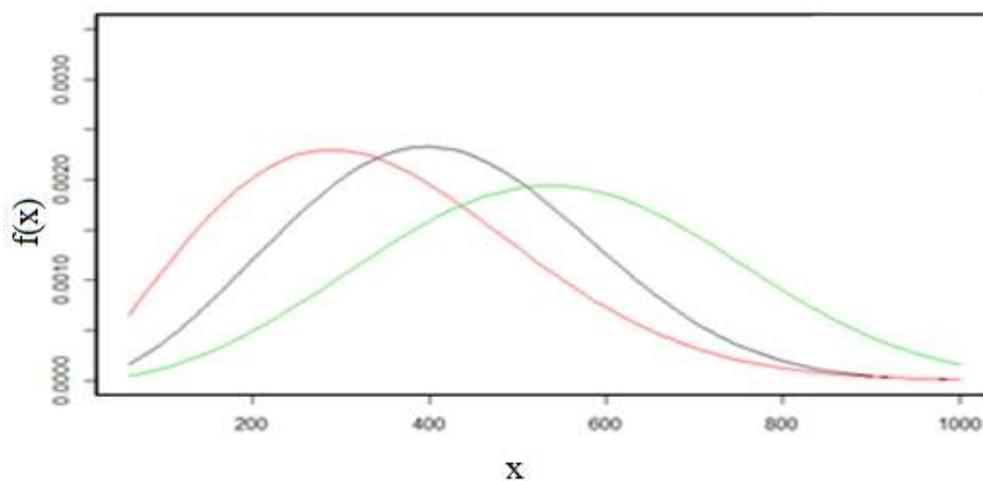
Fonte: o autor. 2018.

### Weibull

$$f(x) = \frac{\gamma}{\alpha} \left(\frac{x-\mu}{\alpha}\right)^{\gamma-1} e^{-\left(\frac{x-\mu}{\alpha}\right)^\gamma} \quad (12)$$

Em que:  $\gamma =$  fator de forma e  $\beta =$  fator de escala

**Figura 5** – Representação de curvas com FDP Weibull com diferentes valores de fator de forma e de escala.



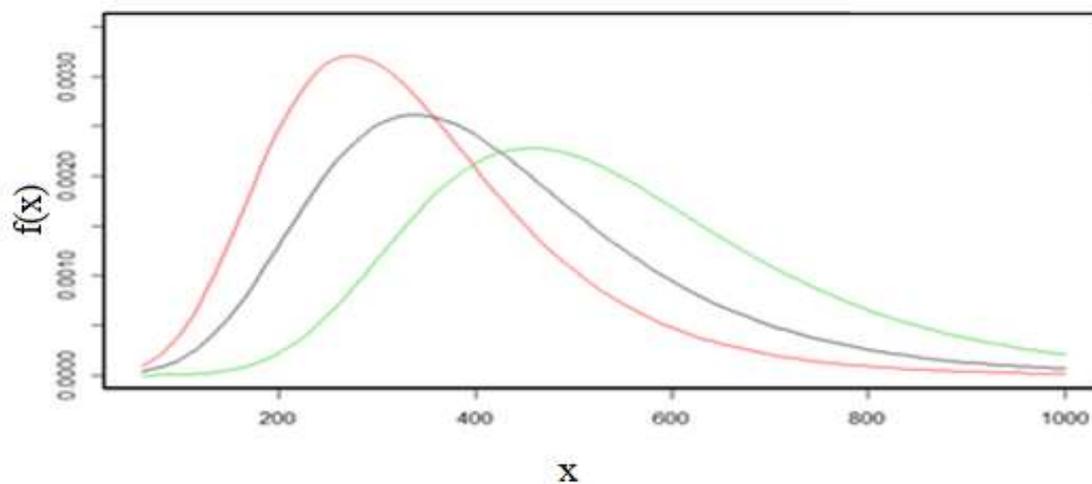
Fonte: o autor. 2018.

### *Gumbel*

$$f(x) = \frac{1}{\beta} e^{\frac{x-\mu}{\beta}} e^{-e^{\frac{x-\mu}{\beta}}} \quad (13)$$

Em que:  $\mu$  = fator de *posição* e  $\beta$  = fator de *escala*.

**Figura 6** – Representação de curvas com FDP *Gumbel* com diferentes valores de *posição* e de *escala*.



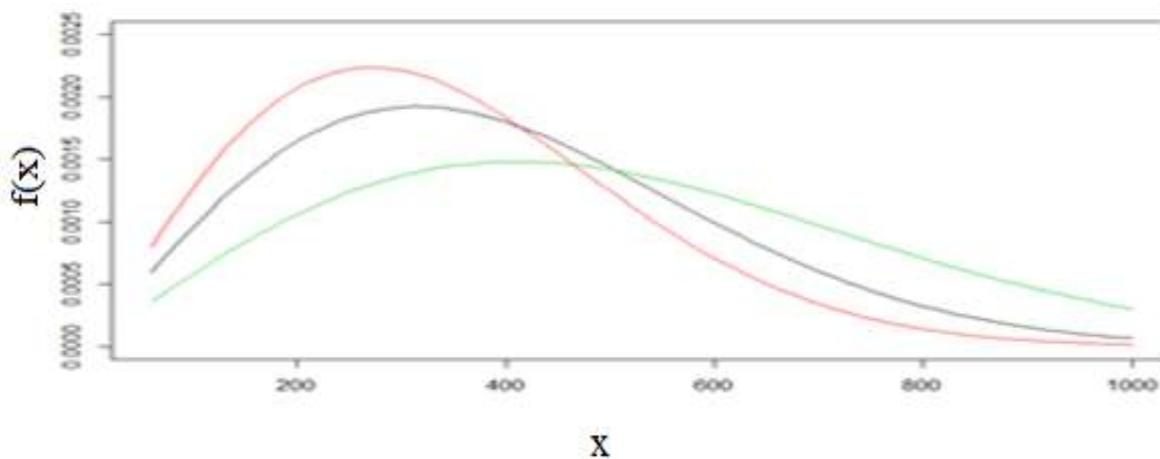
Fonte: o autor, 2018.

### *Rayleigh*

$$f(x) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}} \quad (14)$$

Em que:  $\sigma$  = fator de *escala*.

**Figura 7** – Representação de curvas com FDP *Rayleigh* com diferentes valores de *escala*.



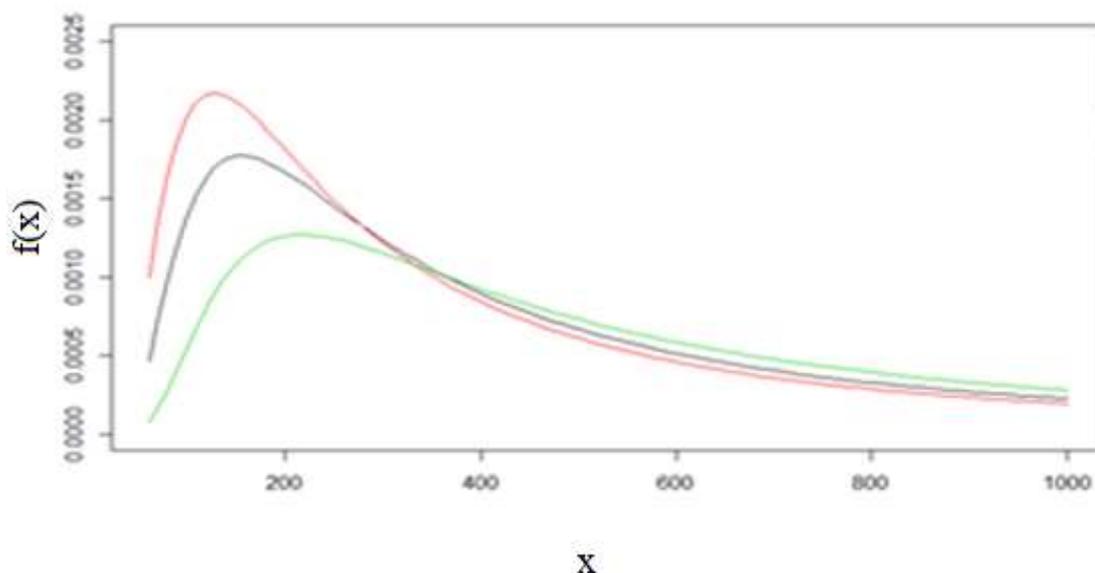
Fonte: o autor, 2018.

**Fréchet**

$$f(x) = \frac{\alpha}{\beta} \left(\frac{\beta}{x-\gamma}\right)^{\alpha+1} e^{-\left(\frac{\beta}{x-\gamma}\right)^{\alpha}} \quad (15)$$

Em que:  $\sigma$  = fator de *escala* e  $\beta$  = fator de *forma*.

**Figura 8** – Representação de curvas com FDP *Frechet* com diferentes fatores de *escala* e de *forma*.



Fonte: o autor, 2018.

### 2.3 Teste de aderência (*Kolmogorov-Smirnov*)

O teste de aderência é um recurso estatístico que busca aferir a concordância de uma função de distribuição de probabilidade teórica com uma determinada série de dados (FERNANDEZ; LIU, 2016). A literatura (e.g. MONTGOMERY; RUNGER, 2014) usa o termo *goodness of fit*, traduzido como a “bondade do ajuste” de um modelo estatístico teórico referente a um conjunto de observações.

O teste de *Kolmogorov-Smirnov* (KS) é não paramétrico e verifica a distância crítica entre duas distribuições acumuladas, como mostrado pela Figura 9. A linha em vermelho na figura representa a maior distância crítica entre as duas funções de distribuição. O teste afere um  $D_{\text{máx}}$  (Equação 16) entre a função de distribuição acumulada empírica  $F(x)$  e outra na qual se deseja testar a aderência  $F(x)^*$  (função teórica). É estabelecido um  $D_{\text{lim}}$  de tolerância como critério de aceitação ou rejeição (valor tabelado). Cabe lembrar que o teste

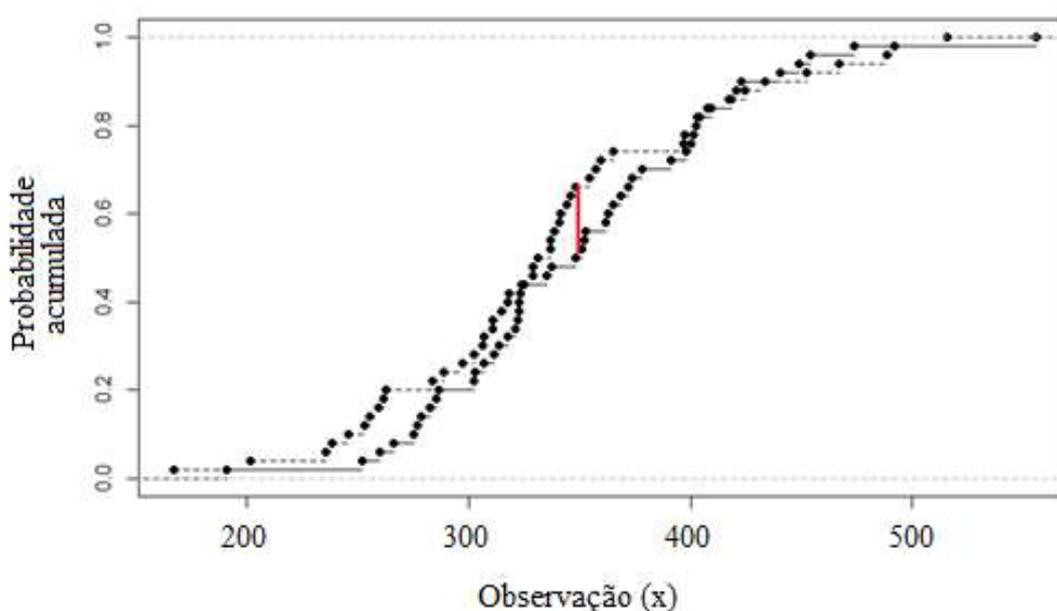
de *Kolmogorov-Smirnov* seleciona se um ou mais modelos conseguem representar satisfatoriamente uma amostra. Para este caso em particular tem-se interesse de analisar se uma amostra é proveniente de uma população gerada por uma FDP qualquer.

$$D_{max} = |F(x) - F(x)^*| \quad (16)$$

O critério de aceitação ou rejeição de hipótese mais utilizado é o *p-value* (valor de p) que analisa o grau e “veracidade” da hipótese nula. Adotado um limite de significância de 5%, tem-se que valores de *p-value* superiores a este mostra que existem forte evidências para aceitar a hipótese nula (i.e. afirmar que a distribuição observada tem aderência a uma determinada distribuição teórica). De maneira análoga, quando o *p-value* é inferior ao limite de significância, existem fortes evidências para rejeitar a hipótese nula e aceitar a hipótese alternativa (MONTGOMERY; RUNGER, 2014).

O teste de *Kolmogorov-Smirnov* é um recurso que testa a aderência do banco de dados para as diversas funções distribuições de probabilidade (FDP). O teste KS é um critério “eliminatório” que analisa se a função distribuição é apta ou não para descrever um conjunto de dados. Posteriormente, esta análise é realizado o teste de AIC, discutido posteriormente, que seleciona o melhor modelo. Ou seja, esta etapa é “classificatória”, dentre aquelas “aprovadas” pelo teste de *Kolmogorov-Smirnov*.

**Figura 9** – Teste de *Kolmogorov-Smirnov* entre duas distribuições com o  $D_{max}$  em vermelho.



Fonte: o autor, 2018.

Diversas técnicas de inferência estatística são conhecidas para analisar a aderência dos dados (*Kolmogorov-Smirnov*, *Cramer-von Mises*, *Anderson-Darling*, *Akaike's* e *Bayesian*). Entretanto, o teste de KS é possivelmente o mais aceito em razão da sua simplicidade e ampla difusão. Logo, o interessado pode ficar a vontade para escolher outro teste, frente aos demais testes disponíveis na literatura.

### 2.3.1 Critério de ajuste de Akaike (*Akaike Information Criterion - AIC*)

O modelo de *AIC* foi desenvolvido por Akaike (1973) como maneira de comparar modelos diferentes para um determinado resultado. Snipes e Taylor (2014) utilizam o teste de *AIC* para identificar quais variáveis influem na classificação do vinho, estimados diversos modelos de regressão estes são avaliadas conforme o critério de *AIC* para determinar qual modelo de regressão consegue se aproximar melhor do resultado exato. A seleção do modelo é extremamente importante já que este deve capturar a verdadeira natureza das variáveis conseguindo uma melhor previsão do resultado almejado. Entretanto, deve-se tomar cuidado com os modelos extremamente simplistas ou complexos. O teste de *AIC* consegue “equilibrar” estes comportamentos extremos.

O teste “classificatório” utilizado para seleção da melhor FDP é de *Aikaike's* (*AIC*). As funções aceitas a partir do teste de aderência devem sofrer um escrutínio mais rigoroso. Como o teste KS elimina as distribuições não aderidas, resta saber qual a função que melhor se ajusta (FILION, 2015). Quanto a isto, dois critérios podem ser empregados: o de maior valor de  $p$  do KS (i.e. quanto maior o valor de  $p$ , mais adequada é a distribuição) e o valor do teste de *Akaike*. O teste de *Akaike* (*AIC*) realiza um “cômputo relativo” entre diversas distribuições simultaneamente e verifica qual função melhor se ajusta (ARNOLD, 2010). O critério *AIC* é uma medida relativa do quão bom é um modelo frente aos demais. Dada pela Equação 17, o menor valor de *AIC* aponta a melhor distribuição.

$$AIC = 2K - 2\ln(L) \tag{17}$$

Em que:  $K$  = número de parâmetros utilizado pelo modelo estatístico e  $L$  = valor máximo da função de verossimilhança.

## 2.4 Simulação de dados

### 2.4.1 Monte Carlo e Bootstrap

A modelagem de fenômenos ambientais fornece subsídio para conhecimento mais amplo do processo. Assim, as tomadas de decisões têm viés prático referente a aspectos diversos, como manutenção, operação e otimização das atividades realizadas. Entretanto, fenômenos que dependem da aleatoriedade dos dados sempre trazem uma flutuação em torno de um valor esperado.

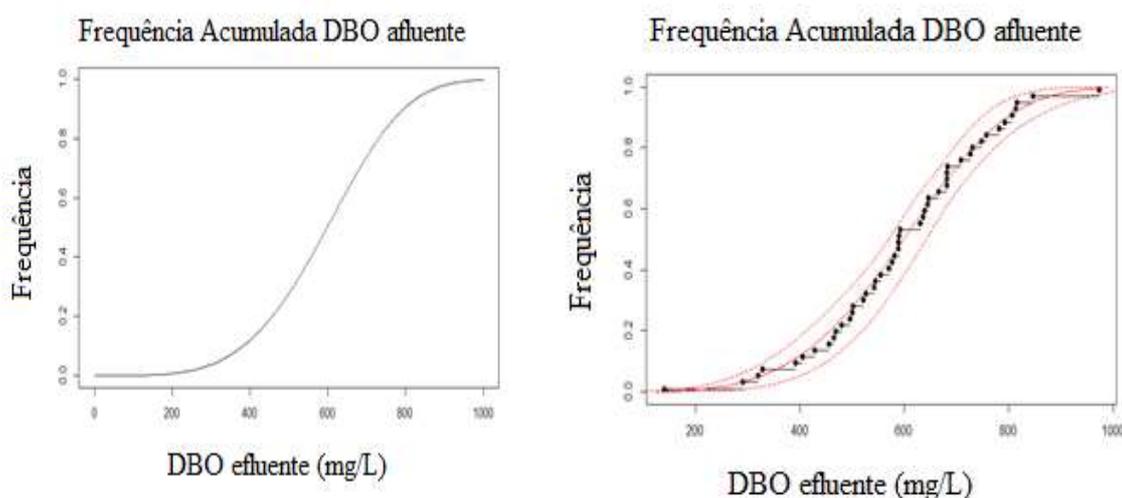
Uma modelagem satisfatória é necessária, devido à natureza aleatória dos dados, já que existe a possibilidade de subestimar ou superestimar um valor obtido através de um modelo matemático determinístico, devido a um elevado conjunto de situações possíveis para um sistema. Assim, os cenários reais tendem a ser distintos daqueles esperados. Por exemplo, Schendel e Thongwichian (2015) estudaram a análise de frequência de inundação para uma dada região. Estes autores enfatizam sobre o quanto pode ser desastroso subestimar ou superestimar o tempo de retorno de um evento hidrológico através dos métodos tradicionais, podendo trazer sérios prejuízos à infra-estrutura da região, tendo em vista que estes dados de projetos são importantes para o projeto de grandes obras hídricas.

Concernente à qualidade de água em geral, um método usual é o de Monte Carlo (KRZYZANOWSKI et al., 2016; POINTON; MATTHEWS, 2016; PEI et al., 2016; JOHNSON et al., 2014). O método, porém, tem caráter paramétrico, exigindo de antemão conhecer uma função distribuição de probabilidade que possua boa aderência com os resultados. Uma alternativa forte é a utilização da metodologia *Bootstrap* para a construção de intervalos de confiança. A metodologia *Bootstrap* pode ser entendida como o resultado da simulação Monte Carlo executado em número elevado de vezes (MULLER et al. 2017). Deste modo, realizadas “*n*” simulações é possível aferir intervalos de flutuação dos valores de saída do sistema estudado.

A Figura 10 mostra a diferença entre a simulação *Bootstrap* e a simulação Monte Carlo. Utilizado os dados de DBO afluente de uma ETE é possível, utilizando ambas as metodologias, aferir a frequência acumulada vs DBO afluente para o intervalo de 0 a 1000 mg/L. Entretanto, como mostrado a seguir, a metodologia *Bootstrap* mostra intervalos de confiança para a probabilidade acumulada, representado pelas linhas vermelhas tracejadas pontilhadas.

Para o exemplo em apreço foi utilizado o Intervalo de Confiança de 95% (chances de que o valor esteja no intervalo compreendido entre as linhas tracejadas). O *Bootstrap* tornou-se difundido em diversas áreas do conhecimento devido a sua simplicidade, poucas premissas e principalmente pelo advento dos supercomputadores. Destaca-se que a metodologia *Bootstrap* “desloca” os intervalos de confiança encontrados pela simulação Monte Carlo dentro de uma faixa de tolerância (limites de confiança).

**Figura 10** – Simulação Monte Carlo (esquerda) vs Simulação *Bootstrap* (direita).



Fonte: o autor, 2018.

Peda et al. (2013) sugerem a técnica *Bootstrap* como ferramenta que pode revelar melhor o impacto das variáveis ambientais e operacionais sobre a eficiência de equipamentos de tecnologia ambiental (e.g. operação de uma ETE). Cabe lembrar que as Estações de Tratamento de Esgoto (ETE) têm o objetivo de adequar as características físicas, químicas e biológicas do esgoto aos padrões previamente estipulado, desta forma tornando-se objeto de estudos para a utilização da metodologia *Bootstrap*.

No estudo do controle da qualidade de efluentes o analista se depara com a flutuação da intensidade dos fenômenos envolvidos, trazendo complexidade ao seu estudo. Este comportamento foi bem descrito por von Sperling (1996) e Silva et al. (2014). Os autores previram a incerteza dos parâmetros de design de uma estação de tratamento, lagoas facultativas, como população, taxa de remoção, matéria orgânica dissolvida, constantes cinéticas, etc.

Sleigh e Mara (2003) desenvolveram uma metodologia que permite realizar milhares de simulações com diferentes valores de entrada. De modo que os dados de saída

da simulação é um histograma que mostra a curva de frequência acumulada dos valores mais prováveis. Este procedimento é particularmente útil, já que mostra a pertinência de um valor de saída dentro de um intervalo esperado, utilizando uma ferramenta gráfica.

Assim, importa estabelecer intervalos de confiança (IC) para as diferentes funções descritivas e ajustes requeridos pela aleatoriedade dos dados (COLIN; VANHOUCHE, 2015). O IC mais utilizado é de 95%, ou seja, o valor correto para uma dada estimativa possui 95% de “chances” de estar neste intervalo. O IC é representado conforme a equação 18 a seguir:

$$IC = EP \pm (VC + \sigma) \quad (18)$$

Em que: IC = Intervalo de Confiança; EP = Estimativa Pontual; VC = Valor Crítico e;  $\sigma$  = desvio padrão

Construídos os valores dos limites de confiança para uma dada distribuição é possível prever a faixa de flutuação no qual estes valores ocorrem. De sorte que cada limite (seja ele inferior ou superior) é um cenário antagônico no qual o fenômeno pode apresentar. Por exemplo, Mirabi et al. (2014) estabeleceram cenários otimistas e pessimistas para uma série de objetivos a serem alcançados por uma ETE e, desta maneira, analisaram o grau de conformidade com uma condição almejada dentro das possibilidades tangíveis.

É interessante mensurar a amplitude das variações da estatística descritiva (média, mediana, desvio padrão, etc) obtidas a partir de amostras retiradas da população, tendo em mente que a estatística populacional almejada está dentro deste intervalo. Entretanto, este método intuitivo e oneroso já que retira “*n*” amostras de uma população até construir intervalos de confiança de algum valor estatístico de interesse coletado das inúmeras amostras, de sorte que o valor procurado está dentro deste intervalo.

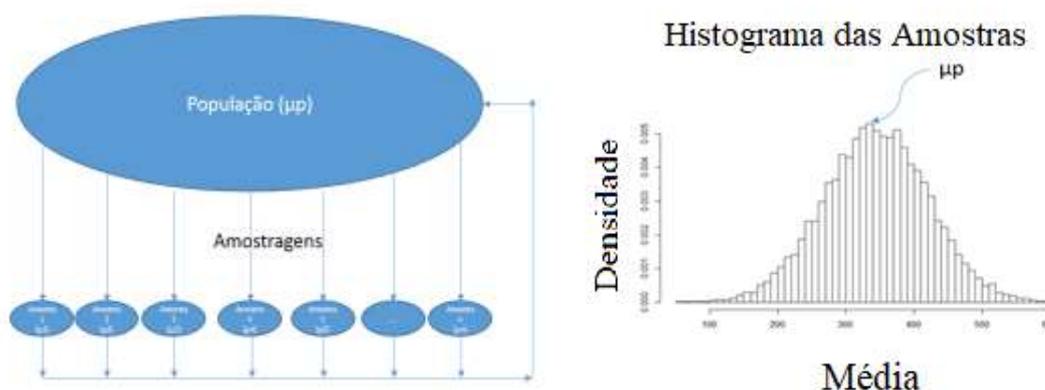
A metodologia *Bootstrap* foi desenvolvida por Efron (1979) e ganhou rapidamente significância acadêmica por sua simplicidade, eficiência e poucas premissas para utilização. A metodologia *Bootstrap* é realizada através de inúmeras reamostragens com reposição de uma amostra inicial, aqui chamada de “amostra mãe”, objetivando formar outras amostras. Este método é utilizado quando o cálculo de estimadores pelo método analítico é complicado, generalidade na qual pode ser aplicado. Fornece respostas precisas e de fácil entendimento. Há ainda a possibilidade de usar o *Bootstrap* não paramétrico (RIZZO; CYMROT, 2006).

A simplicidade da metodologia permite sua utilização em diversos ramos do conhecimento. Possibilita construir intervalos de confiança para uma estatística buscada a partir de uma única carreira de coleta, com base na utilização do esforço computacional que substitui custo e trabalho demasiado do método tradicional de amostragem. O *Bootstrap* é amplamente difundida devido a sua capacidade de analisar a qualidade dos estimadores tais como erro padrão, variância e intervalos de confiança (BASIRI, 2017).

A metodologia convencional de simulação, baseia-se na retirada de “ $n$ ” amostras de uma população para o levantamento da média populacional ( $\mu_P$ ) por exemplo. Para este caso hipotético, mostrado pela Figura 11, é observado que a média possui grande probabilidade de oscilar próximo do valor de 350. Entretanto, são necessários vários cortes amostrais (ou carreiras de amostras), tornando tal prática muito onerosa. O procedimento convencional de construção de IC é mostrado a seguir:

- 1) Coletar “ $n$ ” amostrar de uma população;
- 2) Calcular a estatística de interesse de cada uma das “ $n$ ” amostras;
- 3) Calcular o histograma dos dados obtido no passo 2.

**Figura 11** – Histograma com a estatística de interesse obtido através do método convencional de amostragem.

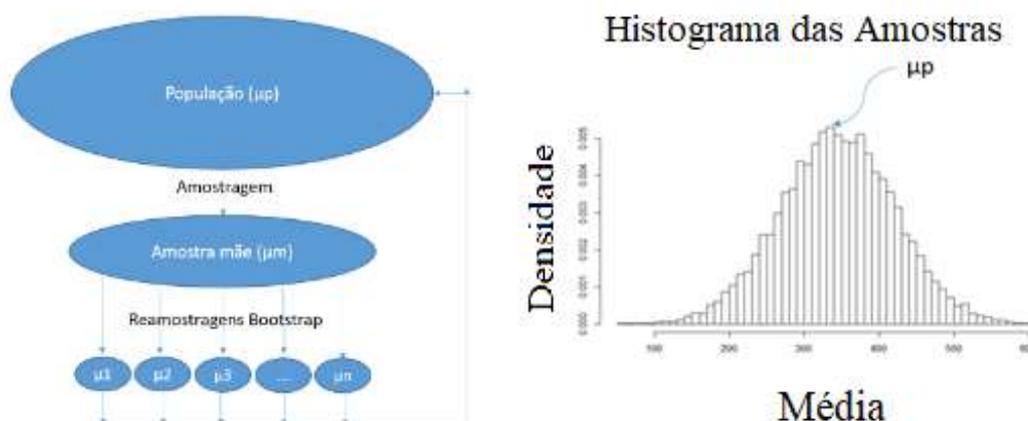


Fonte: o autor, 2018.

A metodologia *Bootstrap* realiza diversas reamostragens, com reposição, a partir de uma única amostra mãe. Dessa forma, quando o número de amostras *Bootstrap* tende a um valor muito elevado o histograma mostrado pela Figura 12 torna-se semelhante ao histograma criado pela metodologia mostrada pela Figura 11. Por exemplo, a metodologia seguida para a construção do histograma mostrado pela Figura 12 foi a seguinte:

- 1) Coletar uma “amostra mãe”;
- 2) Calcular “ $n$ ” amostrar *Bootstrap* da “amostra mãe”;
- 3) Levantar o histograma obtido dos dados retirados na estatística de interesse.

**Figura 12** – Histograma com a estatística de interesse obtido através da metodologia *Bootstrap* de amostragem.



Fonte: o autor, 2018.

## 2.5 Risco e garantia – enfoque sobre a qualidade de efluente

A confiabilidade de um sistema de tratamento de esgoto pode ser definida como a probabilidade de se atingir um desempenho adequado por um período de tempo especificado em condições pré-estipuladas (OLIVEIRA; VON SPERLING, 2008). Para o tratamento de efluentes a literatura mostra uma série de condições que podem ser o resultado de monitoramento de parâmetros físicos, químicos e biológicos e suas conformidades com a legislação (UPADHYAYA; MOORE, 2012). Busca-se, em geral, verificar a percentagem de remoção de poluentes para uma dada tecnologia está dentro das exigências (RODRIGUEZ-GARCIA et al., 2011), a remoção de poluentes em função dos custos envolvidos (BENEDETTI et al., 2008) e os custos de operação da estação (BOURRIER et al., 2010). Paralelo ao conceito de confiabilidade tem-se que o grau de garantia que um determinado processo oferece.

O trabalho pioneiro em relação a este tema foi escrito por Niku et al., (1982) que introduz o conceito de confiabilidade do tratamento de efluente. Este trabalho devido a

escassez de ferramentas computacionais disponíveis na época limita-se a distribuição lognormal para o cálculo do coeficiente de confiabilidade que revela a conformidade do tratamento em satisfazer uma norma ambiental frente a aleatoriedade dos dados. Esta mesma metodologia proposta por Niku et al., (1982) foi utilizada por Oliveira e von Sperling (2007) e Alderson et al., (2015) que utilizaram o coeficiente de confiabilidade para analisar a garantia de várias estações de tratamento de esgoto para diversos parâmetros de qualidade. O trabalho destes autores expandiram o conceito do coeficiente de confiabilidade para diversas tecnologias e parâmetros de qualidade de efluentes. Entretanto, não houve nenhum avanço significativo em relação a técnica elaborada por Niku et al., (1982).

Niku et al., (1983) afirmam que para o contexto aplicado no controle de qualidade tem-se que a probabilidade de falha é extremamente sensível a função distribuição de probabilidade utilizada para modelar o efluente. Este fato revela a necessidade de utilizar a função distribuição que melhor representa o efluente. Frente a esta afirmação são testadas diversas funções e escolhida aquela que possui melhor aderência com os dados buscando aumentar a acurácia do modelo. Niku et al., (1982) conseguiu absorver a natureza aleatória dos dados para o cálculo do coeficiente de confiabilidade. Entretanto, estes autores não conseguiram observar que o próprio coeficiente de confiabilidade sofre com a flutuação dos dados inerentes da tecnologia de tratamento de efluentes. Ou seja, o modelo pode trazer uma previsão fora da realidade por não considerar a faixa de variação do coeficiente de confiabilidade. As limitações do modelo de previsão da confiabilidade criado por Niku et al., (1982) são sanadas neste trabalho através da utilização de intervalos de confiança para a representação mais astuta da garantia fornecida pela estação.

Complementar ao conceito de garantia tem-se o conceito de risco. O risco refere-se à probabilidade conjunta da ocorrência de uma condição adversa, não desejada (SADIQ; KLEINER; RAJANI, 2007). Sob esse conceito, a falha remete ao grau no qual o risco se apresenta. Por exemplo, a gestão dos riscos baseada na toxicidade de constituintes das águas mostra-se ferramenta adequada para aferir a conformidade da utilização de um efluente para uma finalidade específica (ZENG, 2016). Logo, o risco é apresentado como fator importante para analisar a conformidade da utilização de um efluente tratado.

O conceito de risco pode ser expandido, com a inclusão da consequência da falha e se reporta ao grau de risco que um determinado processo oferece e está associado à consequência do evento (ou seja, do prejuízo) (LI; HUANG, 2012), que é mostrado na Equação (19).

$$\text{risco} = \text{probabilidade de um evento acontecer} \times \text{consequencia deste evento} \quad (19)$$

Subestimar a ocorrência de um evento deletério pode causar prejuízos ao reuso de efluentes ou riscos a segurança e a saúde. Por exemplo, Sanctis et al. (2017) afirmam que para a viabilização de um sistema de reuso deve-se analisar o risco referente à contaminação por patógenos. Logo, a gestão dos riscos em níveis adequados é de fundamental importância para que efluentes não venham a causar efeitos negativos sobre a saúde da população ou ao meio ambiente.

Clemente e Oliveira (2015) destacam que o gerenciamento de risco ambiental constitui ferramenta essencial para a categorizar, minimizar, potencialmente e eliminar riscos ambientais das operações em sistemas de produção. Após a estimativa do grau de falha recorrente de uma fonte de risco que um processo apresenta, o responsável pelos equipamentos ambientais (leia-se aqui ETE) deve analisar a necessidade de uma medida paliativa para mitigar o risco, intervenção mais importante ou, se este estiver em intervalos aceitáveis, simplesmente ignorá-lo.

Como já foi sugerido no início deste estudo, a medida de transferir o risco para a realidade das ETE é pouco explorada. Entretanto, vem o questionamento de como estabelecer diretrizes que forneçam indicativos de quais medidas adotar frente à gestão dos riscos. Gargama e Chaturvedi (2011) respondem a esta pergunta analisando a probabilidade de um evento acontecer e a consequência deste, conforme mostrada pelas Tabelas 3 e 4.

Logo, frente as informações mostradas pelas Tabelas 3 e 4 tem-se um indicativo se o risco é aceitável para a operação de uma dada ETE. Por exemplo, se o risco é remoto e possui impacto catastrófico este não pode ser negligenciado. Por outro lado, se o risco possui alta probabilidade e, por outro lado, apresenta um baixo impacto, existe a possibilidade de ser aceito. A gestão dos riscos parte da ideia de que deve ser tirado o máximo proveito quando o risco é baixo e a situação satisfatória. Assim, devem ser tomadas medidas preventivas quando o risco chega a patamares que ofereça cenários que forneçam consequências que vão além do aceitável (GOERLANDT, 2017).

**Tabela 3** – Graus de Probabilidade relacionado a ocorrência de eventos.

<i>Probabilidade categorizada</i>	<i>Frequência de ocorrência</i>
1 – Remota	Conceitualmente possível, mas extremamente improvável de ocorrer durante a vida útil do empreendimento.
2 – Baixa	Possível de ocorrer até uma vez durante a vida útil do empreendimento
3 – Média	Possível de ocorrer até uma vez ao longo do ano operacional
4 – Alta	Possível de ocorrer mais de uma vez ao longo do ano operacional

Fonte: Gargama e Chaturvedi (2011).

A Tabela 4 relaciona o risco associado com a intensidade e com a frequência que o evento pode ocorrer. Observa-se que um evento com severidade catastrófica possui alto impacto no projeto independente da frequência no qual esta pode vir a ocorrer. Por outro lado um evento com consequência baixa não oferece impacto significativo no projeto independente da frequência no qual este pode ocorrer.

**Tabela 4** – Severidade de um evento relacionado com a probabilidade.

<i>Severidade/ Probabilidade</i>	<i>Remota</i>	<i>Baixa</i>	<i>Média</i>	<i>Alta</i>
Catastrófica	Alto Impacto	Alto Impacto	Alto Impacto	Alto Impacto
Crítica	Moderado Impacto	Moderado Impacto	Moderado Impacto	Alto Impacto
Moderada	Baixo Impacto	Moderado Impacto	Moderado Impacto	Alto Impacto
Baixo	Baixo Impacto	Baixo Impacto	Moderado Impacto	Alto Impacto

Fonte: Gargama e Chaturvedi (2011).

Em tecnologia ambiental o risco pode ser computado em relação à eficiência de remoção esperada para um determinado poluente ou quanto ao atendimento a requisitos normativos. Em cenários otimistas, tolerantes ou de alto desempenho os valores de risco são baixos. O risco está diretamente relacionado à capacidade do sistema de tratamento – remoção – em atingir um determinado objetivo.

A exposição dos conceitos de risco e a garantia aplicada ao tratamento de esgoto remete à necessidade da adoção de uma metodologia que venha a mensurar adequadamente

estes valores. Para tanto, deve-se considerar uma abordagem probabilística e/ou *Fuzzy* (CHEN ; SANGUANSAT, 2011).

### **2.5.1 Abordagem probabilística**

A abordagem probabilística é útil por realizar o levantamento de diversos cenários possíveis com a probabilidade de ocorrência recorrente de cada. Em projetos nos quais todas as situações possíveis têm relevância este método é extremamente adequado. A primeira utilização da abordagem probabilística para avaliar o risco de um processo se deu após o desastroso acidente envolvendo a Apollo 1 em 1967 (KELLER; MODARRES, 2005) pela NASA (*National Aeronautics and Space Administration*). A indústria aeroespacial e nuclear foram as pioneiras na utilização da abordagem probabilística para mensurar os riscos decorrentes de um dado processo. A motivação destes estudos decorrem das desastrosas consequências resultantes de eventos críticos.

Martorell et al. (2017) afirmam que análises probabilísticas de segurança são realizadas como objetivo da aprovação de regulamentação da concepção e operação das centrais nucleares. Levantados todos os cenários possíveis torna-se plausível quantificar o risco intrínseco de cada um e, desta maneira, analisar a confiabilidade do sistema como um todo.

A abordagem probabilística é utilizada em diversas linhas pesquisa. Por exemplo, Bagen e Huang (2017) através da abordagem probabilística quantificaram o risco associado à transmissão elétrica. Sendo esta abordagem viável para diversas áreas do conhecimento. Em estudos ambientais é observada uma infinidade de cenários devido às incertezas recorrentes das condições climáticas (insolação, chuva, temperatura, ventos, etc) ou de processo (vazão, características do esgoto, rendimento do processo, etc). Logo, torna-se necessária a adoção de modelos que venham mensurar a significância de cada evento dentro de um conjunto inerente de possibilidades. Prada et al. (2017) avaliaram os impactos relacionados ao manejo de terras agrícolas nos recursos hídricos através de modelagem probabilística hidrológica. Através desta metodologia torna-se possível construir um modelo sensível as incertezas.

Como já relatado a metodologia comumente utilizada para a análise de cenários mais empregada é a simulação *Monte Carlo*. Esta abordagem baseia-se na geração de um grande número de alternativas de diferentes cenários, permitindo encontrar valores de parâmetros realistas (GRAVELINE et al., 2012; PAULI, ARBENZ, 2015). Com o

conhecimento das principais Funções Distribuição de Probabilidade (FDP) que modelam satisfatoriamente os fenômenos de remoção de poluentes é possível gerar uma amostra grande o suficiente, de maneira que esta tenha as mesmas características da população.

A estimativa de valores probabilísticos de ocorrência de um determinado evento deve considerar Intervalos de Confiança (IC). A metodologia de simulação *Bootstrap* permite a construção dessas faixas simuladas, que melhor representam uma dada estatística populacional. Através da abordagem de intervalos de confiança gerados a partir da metodologia *Bootstrap* é possível restringir a população em um intervalo em que existe grande probabilidade do evento esperado estar localizado (FONT, 2011). Logo, restringir o valor em intervalos é uma forma prudente de prever cenários futuros dentro de limites de tolerância. A infinidade de cenários que podem ser gerados pela teoria probabilística e os intervalos de confiança criados pela metodologia *Bootstrap* tornam tais abordagens complementares.

### **2.5.2 Abordagem Fuzzy**

A teoria da lógica difusa fornece uma teoria matemática para capturar as incertezas associadas aos processos cognitivos humanos, como pensamento e raciocínio. O desenvolvimento da lógica *Fuzzy* foi motivado em grande parte pela necessidade de uma estrutura conceitual que possa abordar questões de imprecisão (IANCU, 2018). A abordagem *fuzzy* é extremamente indicada para o estudo do risco, devido a natureza incerta desta variável. Por exemplo, Li et al., (2018) explora a lógica *fuzzy* difusa para avaliar a confiabilidade de sistemas frente o risco de erros dos operadores em salas de controle.

O estudo do risco vem fomentando a utilização e o estudo da lógica *fuzzy* em diversos ramos do conhecimento. Esta medida é fundamentada pela importância de uma medida acurada do risco, valores errôneos podem trazer grandes prejuízos comprometendo todo um sistema. Jiang et al., (2018) estudaram o risco para o projeto de grandes usinas de potência e Skorupski (2016) para a ocorrência de acidentes aéreos envolvendo aeronaves. Estes autores ratificam que o cálculo do risco a partir da metodologia *fuzzy* é calculado de maneira satisfatória para seus casos de estudo em particular. Ou seja, a metodologia *fuzzy* também se destaca por ser robusta para representar diversos sistemas distintos, não só concernente ao tratamento de efluentes.

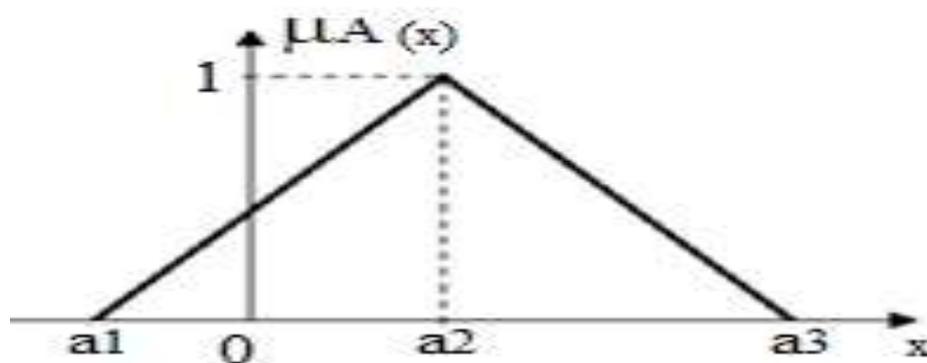
A análise de risco com base em números difusos também pode ser empregada para quantificar riscos associados a incerteza, seja em controle ou tomada de decisões (e.g.

KHORSHIDI; NIKFALAZAR, 2017; MUHAMMETOGLU; YARDIMCI, 2006). A teoria dos conjuntos difusos generalizou a lógica binária (verdadeiro/falso; sim/não; 0/1) da teoria clássica. Com a Teoria *Fuzzy* cria-se uma lógica em que o grau de verdade de uma premissa pode variar de 0 a 1, assumindo valores como “parcialmente verdadeiro” ou “parcialmente falso”. A utilização da metodologia *Fuzzy* é atraente devido a sua simplicidade em computar valores numéricos (WANG; LI, 2016; UPADHYAY et al. 2017).

As representações mais empregadas para números *Fuzzy* estão nas formas triangular (NFT) e trapezoidais (NFT<sub>R</sub>), em que os primeiros são um particularização desses últimos (BODJANOVA, 2005; GIACHETTI; YOUNG, 1997). Um dado NFT é representado por três pontos  $A [A = (a_1, a_2, a_3)]$  e associado a uma função de pertinência  $\mu$  que possui representação funcional (Equação 20) e geométrica como mostrado na Figura 13. O NFT é representado pelos extremos de observação (valores mínimo e máximo) e por um valor mais provável (VMP) A função de pertinência de cada valor associado a um conjunto *Fuzzy* é dada por:

$$\mu_A(x) = \begin{cases} 0, & x < a_1 \\ \frac{x-a_1}{a_2-a_1}, & a_1 \leq x \leq a_2 \\ \frac{a_3-x}{a_3-a_2}, & a_2 \leq x \leq a_3 \\ 0, & x > a_3 \end{cases} \quad (20)$$

**Figura 13** – Representação do número difuso triangular (NFT).

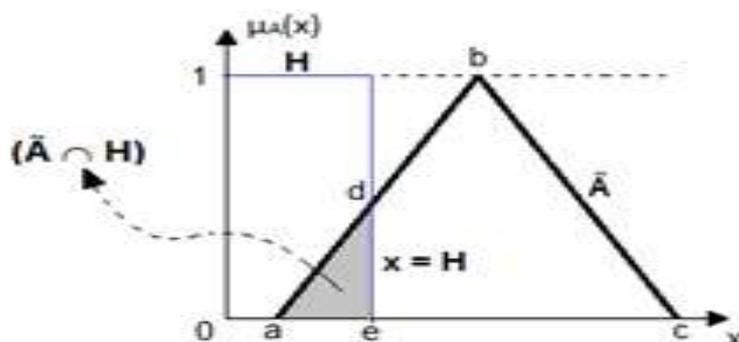


Fonte: Kaufmann e Gupta (1991).

Na Figura 13 acima, os elementos  $a_1$  e  $a_3$  representam o mínimo e o máximo de um conjunto de dados e o  $a_2$  o valor mais provável, relativo a uma medida de tendência central (média aritmética, mediana, média geométrica, moda ou simplesmente um ponto

médio). A partir do NFT é trazido o conceito de risco difuso em concordância ou não com um padrão representado por “e” na Figura 14 (KAUFMANN; GUPTA, 1991).

**Figura 14** – Risco Difuso associado ao NFT.



Fonte: Adaptado de Kaufmann e Gupta (1991).

Através da observação da Figura 13 vem o conceito do *Risco Difuso* ( $R_f$ ). Que é a pertinência de um valor dentro do intervalo de aceitação (sucesso) ou rejeição (falha).

$$R_f = 1 - \frac{\text{Área}_{\Delta ADE}}{\text{Área}_{\Delta ABC}} \quad (21)$$

No risco difuso a não conformidade se baseia na concordância com um valor limite ou outro valor. Não é necessária a adoção de uma função distribuição de probabilidade tornando o método viável para toda e qualquer distribuição. O risco *Fuzzy* varia no intervalo de  $[0,1]$ , em que o limite inferior mostra um processo totalmente apto para satisfazer um dado objetivo e o limite superior mostra um processo totalmente ineficaz. Logo, o risco *Fuzzy* associado a um número triangular traz o grau de falha que um dado processo apresenta, a partir do “mapeamento” da informação. De maneira paralela a teoria probabilística é possível utilizar a abordagem *Fuzzy* como ferramenta para cotejar o grau de distanciamento de um cenário desejado.

### 2.5.3 Grau de conformidade ambiental

Os efluentes tratados podem ter dois destinos: reutilização e disposição no ambiente (OLIVEIRA, 2006). Esta afirmação remete ao seguinte questionamento: quais as

características que o efluente tratado deve possuir para ser apto a um determinado destino final? A CONAMA n°357/2005 introduz o conceito de *padrão* que é o valor limite adotado como requisito normativo de parâmetro de qualidade de água superficial ou efluente. Respeitadas as diretrizes e valores de referência da 357/05, as legislações ambientais que tratam de padrões de lançamento de efluentes ou de reuso adotam seus próprios valores nos quais incorporam as características locais, podendo serem mais restritivas.

A resolução pioneira a referente ao enquadramento dos corpos hídricos foi a CONAMA N° 20, de 18 de junho de 1986. Esta resolução introduz a noção que os níveis de qualidade devem atender às necessidades da população, permitindo assim os usos múltiplos da água. Logo, um tratamento de águas residuais eficiente deve prover ampla variedade de usos tornando a qualidade do efluente obtido satisfatório para o uso preponderante.

A CONAMA n° 357/2005 estabelece diretrizes para o enquadramento do corpo hídricos em águas doces, salobras e salinas de acordo com a qualidade requerida para o seu uso principal em treze (13) classes de qualidade. Onde, os padrões de lançamento de efluente são específicos para cada classe. A Tabela 5 apresenta como estas classes são organizadas dentro do enquadramento do corpo hídrico.

**Tabela 5** – Enquadramento e classes de qualidade.

<i>Enquadramento</i>	<i>Classe</i>
Doce	Especial, Classe I, II, III e IV
Salobra	Especial, Classe I, II e III
Salinas	Especial, Classe I, II e III

Fonte: extraído da resolução CONAMA n° 357/2008.

A Resolução CONAMA n° 430/2011, que altera e complementa a Resolução CONAMA n° 357/2005, traz padrões estipulados para alguns parâmetros de qualidade de efluente ou que agora se dão através da percentagem (%) mínima de remoção. Como por exemplo, a remoção mínima de matéria orgânica em forma de DBO e sólidos em suspensão.

Normalmente, cada esfera administrativa possui seu próprio arcabouço normativo para nortear o despejo de efluentes. Para o caso do Estado do Ceará, região nordeste do Brasil, tem-se a Resolução do COEMA n° 02/2017 que alterou a Portaria da SEMACE n° 154, de 25 de novembro de 2002. A discussão se dá pela necessidade de padrões de despejo de efluentes que se aproxime da realidade regional e se adapte as características geográficas, morfológicas e climáticas. Por exemplo, os padrões de

lançamento de efluentes referente a matéria orgânica dissolvida em países de climas temperados é muito mais exigente que em países de clima tropical.

A resolução CONAMA n° 430/2011 através da vazão do corpo hídrico consegue estabelecer limites de lançamento norteados pela capacidade de diluição dos corpos hídricos mensurado através da Concentração do Efluente no Corpo Receptor (CECR). Dessa forma, a capacidade de autodepuração dos corpos hídricos é levada em consideração quando buscados padrões compatíveis com tais capacidades.

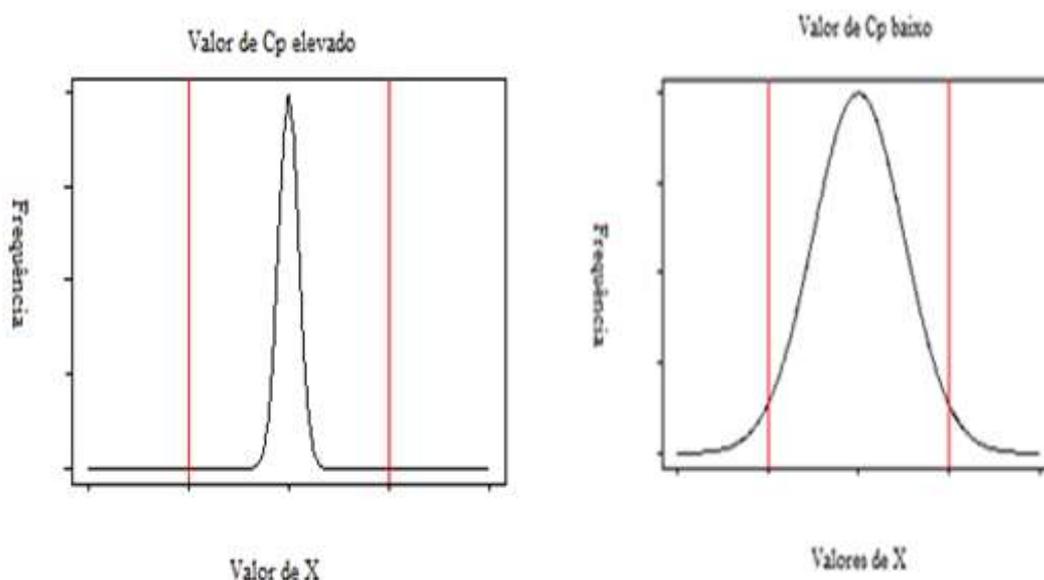
Benedetti et al. (2008) trazem uma abordagem que vai além da esfera ambiental, norteados os padrões de lançamento de efluentes e introduzindo a variável econômica no modelo. Por exemplo, introduzindo o custo total, operacional e variável por unidade monetária por unidade métrica de poluente removido. A análise dos indicadores econômicos possui muita utilidade quando se deseja verificar se os padrões estipulados de lançamento de efluentes estão de acordo com o orçamento dos contribuintes.

Levantado um panorama dos padrões de lançamento de efluente observa-se que estes têm se tornado mais rigorosos com o passar do tempo. Devido principalmente ao avanço das tecnologias de tratamento de esgoto. Isto porque o meio ambiente que vem se fragilizando com o tempo exigindo mais recursos a nível municipal, estadual e federal para atividades relacionadas ao saneamento básico. A Política Nacional do Saneamento Básico (Lei n° 11.445, de janeiro de 2007), objetivando a satisfação de padrões estabelecidos pela legislação ambiental, traz a sugestão de adoção de etapas de eficiência gradual a serem alcançadas pelas ETE respeitando os limites financeiros estipulados.

#### ***2.5.4 Capabilidade de processo***

O estudo da capabilidade de processo busca responder se o processo é apto ou suficiente para desempenhar uma determinada função. Ou seja, deseja-se analisar se mesmo com a aleatoriedade dos dados de qualidade de efluente a estação de tratamento de esgoto consegue valores satisfatórios de adequabilidade. A metodologia apresentada para o cálculo dos índices foi relatada anteriormente, agora é dada atenção para a correta representação de cada índice. As Figuras 15 e 16 ajudam a demonstrar o significado dos índices de capabilidade.

**Figura 15** – Histograma representando a dispersão de dados para dois processos com valor de  $C_p$  elevado (esquerda) e  $C_p$  baixo (direita).



Fonte: o autor, 2018.

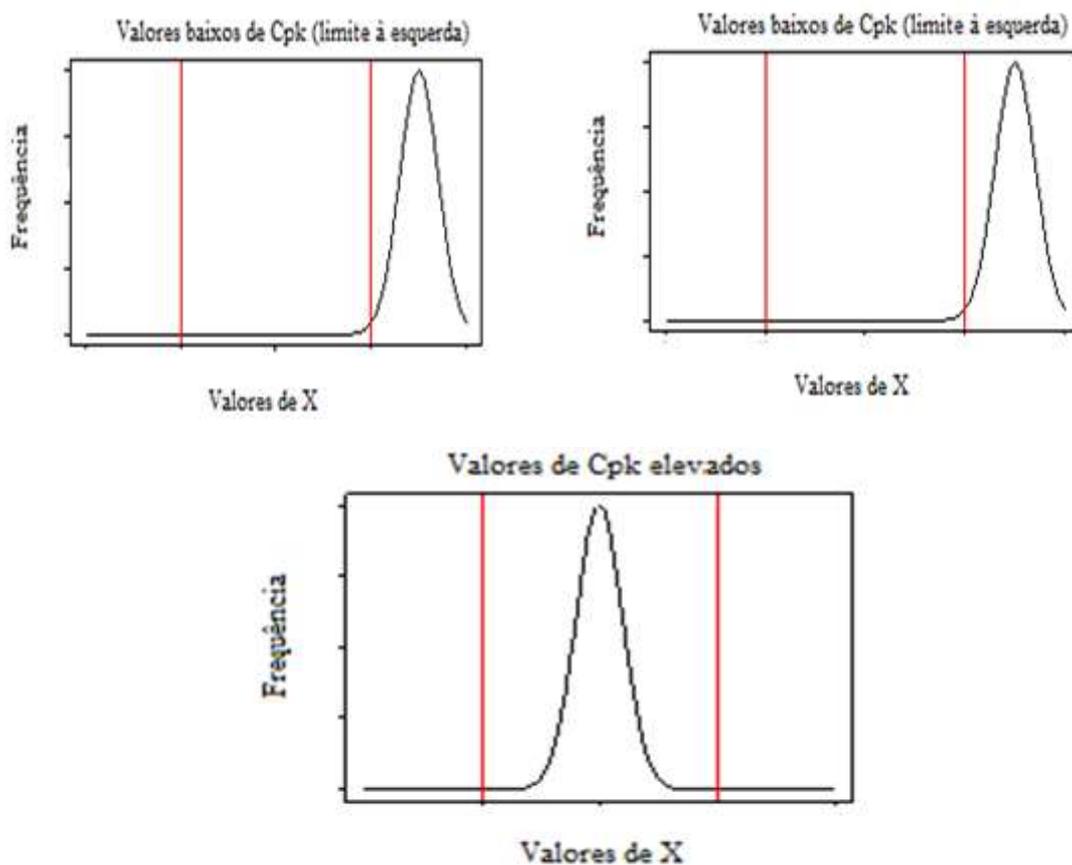
$C_p$  – Considera a não conformidade do processo conforme a taxa de dispersão dos dados. Quanto maior o índice mais apto é o processo em satisfazer as metas estipuladas. Não considera a centralização dos dados em um intervalo, apenas a dispersão dos dados.

$C_{pk}$  – Este índice considera a centralização do processo em torno de um intervalo almejado. Ou seja, busca aferir se os dados estão concentrados entre os limites do intervalo de confiança (mostrados como linhas vermelhas). O valor do  $C_{pk}$  é sensível ao deslocamento da “massa de dados” e dos limites do intervalo de aceitação. Como mostrado anteriormente o valor de  $C_{pk}$  superior a 1,33 mostra que o processo é capaz. A não especificação de um processo conforme a concentração dos dados pode ser devido a não conformidade ao limite inferior (linha vermelha a esquerda) ou ao limite superior (linha vermelha a direita). Para isto é introduzido o conceito do  $C_{pki}$  e  $C_{pks}$ . O valor de  $C_{pk}$  é determinado pelo limite de confiança que mais se aproxima da faixa de variação dos dados do processo.

$C_{pki}$  – Analisa a centralização dos dados conforme o limite inferior de tolerância.

$C_{pks}$  – Analisa a centralização dos dados conforme o limite superior de tolerância.

**Figura 16** –  $C_{pk}$  em conformidade (valores elevados) e em não conformidade (valores baixos) com os limites a direita e a esquerda da concentração de dados.



Fonte: o autor, 2018.

### 3. METODOLOGIA

#### 3.1 Tipo de estudo

O presente estudo tem perfil teórico-aplicativo, conforme o conceito em Prodanova e Freitas (2013). O estudo parte de fonte primária e documental. Sob perspectiva comparativa, coteja realidade e expectativa. De acordo com Fontelles et al. (2009) o detalhamento da pesquisa tem características de acordo com o quadro abaixo:

**Quadro 1:** Detalhamento das características do estudo.

<i>Característica</i>	<i>Explicação sobre a pesquisa</i>
<b>Finalidade</b>	<b>Aplicada</b> (ou <b>tecnológica</b> ) – enseja solucionar um problema concreto, com resultados práticos imediatos
<b>Natureza</b>	<b>Analítica</b> - o investigador deve avaliar com maior profundidade as informações coletadas em um determinado estudo, observacional ou experimental, na tentativa de explicar o contexto de um fenômeno no âmbito de um grupo, grupos ou população
<b>Abordagem</b>	<b>Quantitativa</b> - lida com variáveis expressas sob a forma de dados numéricos. Assim, o caráter analítico é fortalecido, pois tal abordagem oferece maior precisão e confiabilidade, com resultados passíveis de generalização
<b>Objetivos</b>	<b>Exploratória</b> e <b>explicativo</b> - tentar esclarecer melhor a ocorrência dos fenômenos observados
<b>Procedimentos técnicos</b>	<b>Documental</b> – parte de fonte primária, em que a origem dos dados provém de diferentes épocas
<b>Desenvolvimento no tempo</b>	<b>Transversal</b> (ou <b>seccional</b> ), em que a mesma é desenvolvida em um breve espaço de tempo e em um determinado momento

Fonte: adaptado de Fontelles et al. (2009).

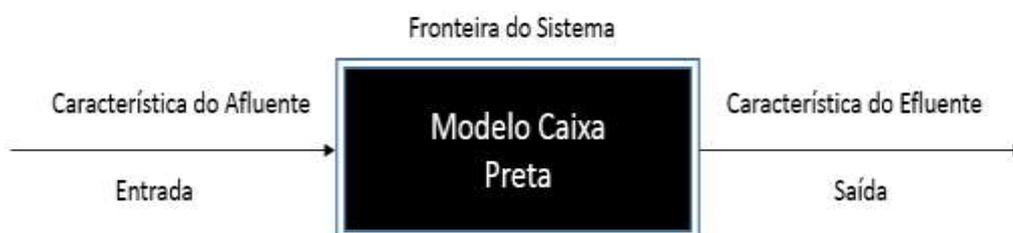
#### 3.2 Estações de tratamento de esgotos e parâmetros de qualidade de efluente

A metodologia proposta no presente estudo considera as estações de tratamento de esgotos como estruturas tipo “caixa preta”, cujos dados são o sinal de entrada e a resposta o de saída. Assim, para o escopo aqui considerado, a tecnologia de tratamento não é

relevante, de forma que somente os sinais de entrada e de saída possam ser considerados (Figura 17).

Apesar da tecnologia não ser o cerne, devido a abordagem “caixa preta”, os dados do estudo são relativos à tecnologia de lagoas de estabilização. Entretanto, a metodologia proposta aqui pode ser replicada para qualquer tecnologia de tratamento de esgoto, bem como qualquer parâmetro de qualidade. Logo, um sistema com possível complexidade, no qual a sua estrutura interna é conhecida ou não, não é levada em consideração na análise. A adoção do modelo caixa preta busca tornar a metodologia robusta para toda e qualquer tecnologia de tratamento de esgoto.

**Figura 17** – Modelo “Caixa Preta” para as ETE do estudo.



Fonte: o autor, 2018.

Foram selecionadas três séries de lagoas de estabilização, com a seguinte configuração: lagoa facultativa primária (LFP), seguida de duas lagoas de maturação (primária (LMP) e secundária (LMS)), conforme mostrado na Figura 18.

A Tabela 6 mostra as coordenadas geográficas de localização, o tempo de detenção hidráulica de cada lagoa e a vazão de cada série. Os dados são relativos ao estudo técnico de Araújo (2011) desenvolvido com lagoas de estabilização no Estado do Rio Grande do Norte.

**Figura 18** – Modelo esquemático dos sistemas de lagoas de estabilização do estudo.



Fonte: adaptado de von Sperling (2002).

**Tabela 6** – Os sistemas de lagoas de estabilização utilizados no estudo.

<i>Designação da ETE</i>	<i>Coordenadas geográficas</i>	<i>Vazão (m<sup>3</sup>/dia)</i>	<i>TDH (dias)</i>		
			<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
PS <sub>5</sub>	6°14'17"S - 35°04'02"W	646	14,24	2,08	2,08
PS <sub>6</sub>	5°53'31"S - 35°11'04"W	7615	13,79	5,52	5,62
PS <sub>9</sub>	5°47'39"S - 35°18'34"W	260	23,26	10,13	10,13

Fonte: Araújo (2011).

As amostras dos efluentes das lagoas e do esgoto bruto foram coletadas no período da manhã, por volta das 10h. As amostras foram analisadas conforme os procedimentos descritos na APHA (1998). Os principais parâmetros considerados para o presente estudo foram: Demanda Bioquímica de Oxigênio (DBO), Demanda Bioquímica de Oxigênio em amostra filtrada (DBO<sub>F</sub>), Demanda Química de Oxigênio (DQO), Demanda Química de Oxigênio em amostra filtrada (DQO<sub>F</sub>), pH, temperatura, Oxigênio Dissolvido (OD), amônia total (AMT) e Coliformes Termotolerantes (CTT).

### 3.3 Ferramenta computacional

No estudo foi utilizado o *software R: A Language and Environment for Statistics Computing* versão 3.3.2, com a interface *Rstudio*. As ferramentas empregadas estão nos seguintes pacotes estatísticos: *Boots (Bootstrap)*, *Fitdistrplus* (seleção de FDP) e *QualityTools* (Controle Estatístico de Processos).

O método utilizado por este *software* para a estimativa dos parâmetros estatísticos foi o da máxima verossimilhança (*mledist*). Entretanto, também estavam disponíveis a estimativa de correspondência de momentos (*mmedista*), estimativa de correspondência dos quartis (*qmedist*) e a estimativa da máxima bondade de ajuste (*mgedista*). Apesar dessas possibilidades de métodos de estimação aquele que conseguiu convergir em 100% dos casos foi o da máxima verossimilhança, logo, este foi utilizado em todo o trabalho.

O teste de *Kolmogorov-Smirnov* (KS) foi realizado utilizando a função *ks.test* do pacote *stats*. O nível de significância para este teste foi de 5% ( $\alpha = 0,05$ ). Posteriormente, ao teste de KS foi utilizado o critério de *Akaike* utilizando a função *AIC* do pacote *stats*.

### 3.4 Risco probabilístico e simulação *Bootstrap*

As incertezas devem ser incorporadas no *design* das estações de tratamento, desta forma, busca-se alcançar efluente com alta qualidade, com o menor custo possível para diversos processos de remoção de poluentes (NIKU et al., 1982). Entretanto, há considerável flutuação nos valores esperados dos parâmetros de qualidade de efluente. Assim, existem incertezas em relação ao risco de não conformidade quanto aos padrões buscados. Portanto, existe a necessidade de uma análise probabilística para mensurar a pertinência de cada cenário em um corte amostral e sua respectiva população. Para isso foram escolhidas ferramentas que pudessem representar uma grande quantidade de eventos.

A seleção de uma FDP que tenha boa concordância (i.e. aderência) com os dados é de vital importância para analisar o grau de confiabilidade da ETE, quanto à qualidade do efluente. A partir da função de probabilidade que tenha aderência com os dados é gerada uma elevada quantidade de números aleatório através do auxílio do *software R*. Dessa forma as grandes quantidades de valores simulados conseguem descrever com acurácia a flutuação dos dados.

A modelagem estatística ou probabilística sempre busca mensurar valores esperados com o máximo de exatidão. Entretanto, como comentado anteriormente neste trabalho existe uma aleatoriedade intrínseca dos dados devido a fenômenos climatológicos, operacionais e ignorância a cerca dos fenômenos envolvidos. Logo, com prudência, foram adotados intervalos de confiança que representaram o valor da informação estatística buscada.

Para o contexto, a metodologia *Bootstrap* constrói intervalos de confiança para uma estatística de interesse. Desta forma o risco é tratado como um parâmetro populacional, e a partir da utilização do *Bootstrap* é possível determinar o Intervalo de Confiança (IC) da probabilidade de adequação do tratamento de efluente a valores limites impostos pela norma. Utilizada a abordagem mostrada na equação 18 o risco toma dimensões intervalares para a qualidade do efluente tratado. Logo, o conceito de risco, através da abordagem adotada para a confecção deste trabalho, foi atualizada para uma interpretação intervalar em substituição a apresentação escalar clássica.

Levantados os intervalos de confiança (superior e inferior) através da simulação *Bootstrap* foi possível analisar a garantia probabilística de conformidade utilizando os padrões de lançamento de efluente. Como exposto anteriormente o conceito de garantia é descrito de maneira intervalar, onde os limites do intervalo de confiança representam

cenários antagônicos. Isso ocorre de tal forma que a proximidade entre os limites dos IC - ou seja dos intervalos – sob determinada probabilidade representa a melhor *performance* alcançável no tratamento do esgoto. Assim, a estação opera próxima da melhor condição que esta consegue atingir sob uma probabilidade computada com base nos resultados da simulação. Por outro lado, a a distância entre os limites de IC revela que o tratamento está ocorrendo de maneira inadequada. Entretanto, tal análise está sujeita aos limites normativos e aos valores de probabilidade da FDA.

Por exemplo, quanto maior o oxigênio dissolvido melhor, mas por outro lado, quanto maior a matéria orgânica dissolvida no efluente pior. Ressalta-se ainda que além da proximidade dos intervalos de confiança (melhores e piores cenários possíveis é possível de serem alcançados por uma dada ETE) a norma ambiental também é utilizada como critério de aceitação do tratamento de efluentes.

### **3.5 Controle estatístico de processos e o grau de conformidade**

Dentro do arcabouço de ferramentas disponíveis no CEP foi utilizado o índice de capacidade ( $C_p$ ) para analisar a conformidade de adequação à legislação para três estações de tratamento. Através deste procedimento foi possível calcular a garantia, em partes por milhão (ppm), de amostras que obtiveram um padrão de despejo satisfatório em relação a um limite balizador. Para isto optou-se por adotar uma abordagem conjunta com as funções distribuição de probabilidade já apresentadas neste trabalho.

O número de amostras coletadas para cada estação é limitado devido ao custo e o tempo envolvido na análise. Entretanto, é preferível utilizar uma amostra que consiga representar fielmente a população da qual esta foi retirada. Para isto é utilizada uma função distribuição de probabilidade que consiga representar todo o espaço amostral. Ou seja, todas as combinações possíveis de valores que podem ser provenientes de uma dada estação. Através da metodologia apresentada é possível “gerar uma amostra artificial” através de uma FDP que consiga representar todo o espaço amostral possível para um dado sistema.

Foi utilizada assim, uma amostra artificial gerada de uma função distribuição de probabilidade que contemplou os cenários referentes à qualidade do efluente. Desta forma a acurácia do cálculo dos índices de capacidade foi incrementado. A metodologia paramétrica do cálculo dos coeficientes de capacidade conforme apresentadas pelas equações (1 a 8) busca uma forma alternativa de alcançar precisão com um número pequeno de amostras. O procedimento proposto seguiu os seguintes passos:

- i) Cômputo da estatística descritivas das amostras coletadas;
- ii) Seleção da função distribuição de probabilidade (FDP), através do conjunto KS-AIC;
- iii) Geração de valores aleatórios em conformidade com a FDP escolhida no passo anterior;
- iv) Estimativa dos valores dos coeficientes de capacidade.

Foram gerados  $10^6$  valores aleatórios para a construção dos índices de capacidade a partir da função distribuição de probabilidade encontrada no passo *ii*. Através de uma função geradora de números aleatórios buscou-se “contemplar” todos os cenários possíveis e aferir o número de amostras em não conformidade dentro de um conjunto de um milhão de coletas (ppm). O pacote empregado nessa etapa foi *QualityTools* do *Rstatistical* (*R*).

Através da análise dos índices de capacidade determinou-se o impacto da variação dos dados em relação à conformidade do tratamento para atender a norma ambiental. Também, foram identificados resultados anômalos que surgiram com as flutuações. Esta análise foi corroborada com representação gráfica na forma de histogramas gerados através dos dados obtidos através das funções distribuição de probabilidade.

### **3.6 Risco *Fuzzy* e a garantia probabilística**

O número *Fuzzy* triangular (NFT) foi empregado para mensurar incertezas integradas à abordagem probabilística empregada com o *Bootstrap*. O risco *Fuzzy* foi calculado através desta representação geométrica triangular e computado conforme a Equação 21. Esta é a maneira mais usual de cálculo do risco *Fuzzy*. Entretanto, introduziu o conceito do risco associado ao distanciamento de uma situação “ótima” ou a proximidade de uma situação “péssima” através do número *Fuzzy* triangular.

A metodologia *Bootstrap* constrói intervalos de confiança através da técnica de reamostragens. A estatística de posição escolhida para este trabalho foi o decil. Deseja-se analisar a porcentagem de amostras com valor inferior ao limite estipulado pela legislação.

A garantia probabilística que um processo oferece é estimado através de um arcabouço de técnicas, dentre elas, a metodologia de *Monte Carlo* é a mais conhecida, como mostrado na Figura 10. Entretanto, este método é limitado a abordagem escalar. Onde,

apenas um cenário (o mais provável) é representado. Este fato compromete o estudo da garantia fornecida por uma ETE devido o amplo intervalo de valores que este parâmetro pode assumir. Busca-se contornar esta limitação introduzindo o conceito de garantia intervalar.

O conceito de garantia intervalar é introduzida através do *Bootstrap*. Este comportamento é revelado a partir da construção dos intervalos de confiança (IC). De forma sucinta os limites de confiança (superior e inferior) revelam os cenários extremos para a garantia e a proximidade ou distanciamento dos limites revelam o desempenho de um processo. A abordagem intervalar tem a grande vantagem em conseguir representar a flutuação dos valores de qualidade de efluente. De sorte que os limites do IC representam a condição “péssima” e “ideal” e os demais cenários possível encontram-se entre estes dois extremos.

Construído os IC para a garantia probabilística tem-se uma infinidade de cenários que são representados entre os limites de confiança, daí surge a necessidade de se calcular onde um cenário está localizado dentro do intervalo. Assim quanto mais próximo do cenário “ideal” melhor e quanto mais próximo do cenário “péssimo” pior. O método encontrado para calcular o distanciamento de um evento em relação os limites de confiança foi o risco *Fuzzy* triangular. Através deste método é possível calcular a garantia associada a um evento dentro de cada conjunto de cenários (intervalos de confiança) para cada ETE.

Construído o IC utilizando a metodologia *Bootstrap* e calculado o risco *Fuzzy* triangular para um cenário dentro de um conjunto de possibilidades. Esta metodologia busca alertar para a necessidade de medidas corretivas para caso a estação esteja longe de um cenário ótimo e/ou buscar otimizar o tratamento dentro da capacidade da estação. As garantias são calculadas através dos padrões de qualidade sugeridas pela legislação competente.

### **3.7 Procedimentos estatísticos**

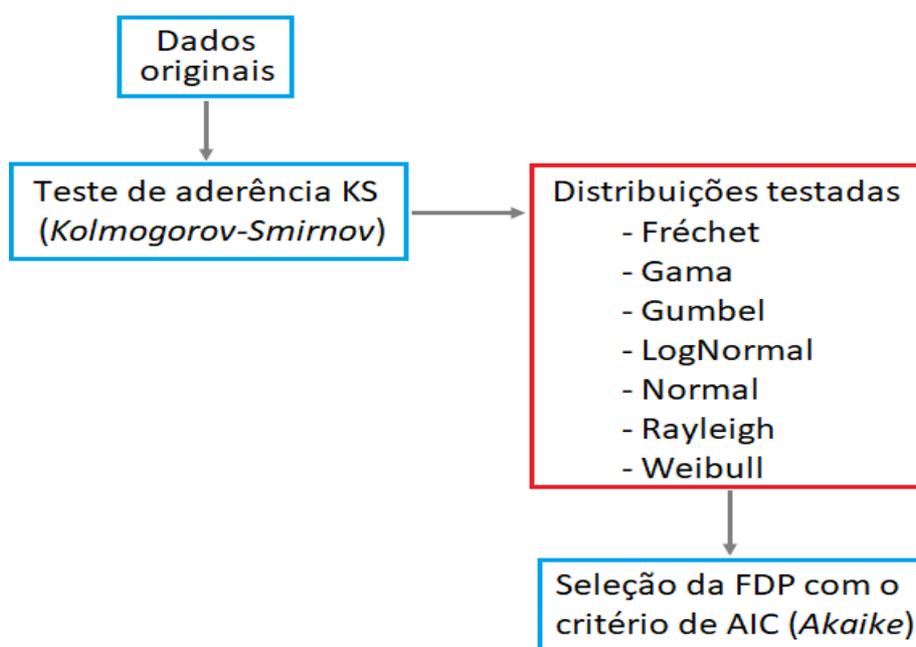
Este tópico busca expor o procedimento estatístico em forma esquemática para favorecer o entendimento das etapas seguidas. Para ficar mais claro o entendimento o *script* do algoritmo executado no *software* R é disponibilizado no apêndice no final deste trabalho.

### 3.7.1 Representação esquemática do processo de seleção das FDP

O teste de aderência para aferir a melhor FDP para cada parâmetro de qualidade é calculado através de duas etapas (conforme mostrada pela Figura 1):

1. Teste KS (seleciona as funções “hábeis” para serem utilizadas);
2. Teste de AIC (seleciona a melhor função dentre aquelas aprovadas na etapa 1.

**Figura 19** - Representação metodológica para o teste de aderência.

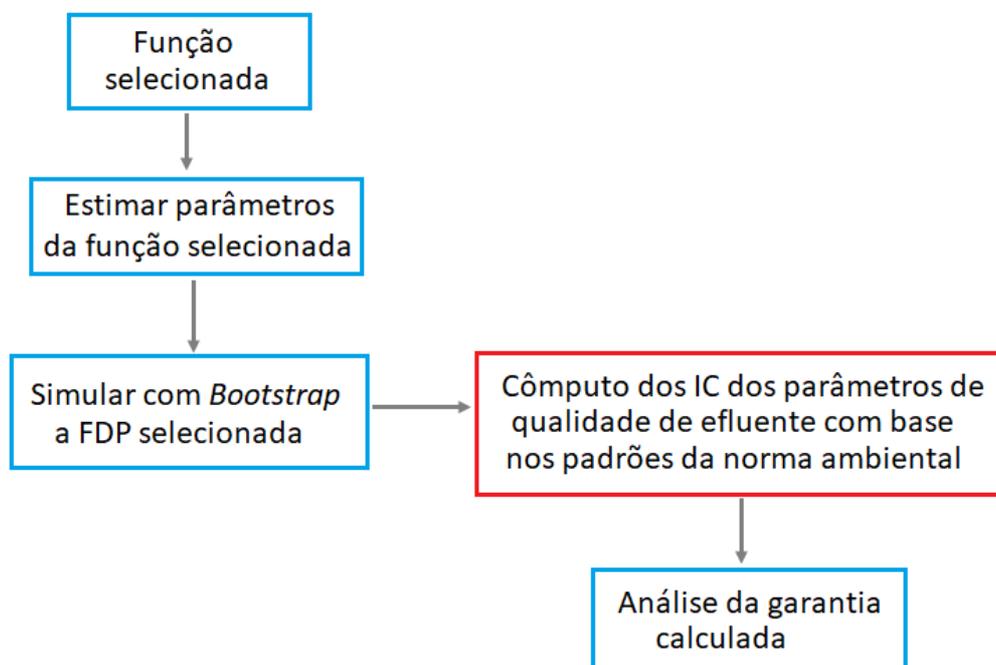


Fonte: o autor, 2018.

### 3.7.2 Representação esquemática com aplicação do Bootstrap e construção dos intervalos de confiança (IC) nos decis

Conhecida a melhor FDP utilizando a metodologia exposta pela Figura 19 a garantia fornecida pelo método *Bootstrap* é calculada seguindo a representação mostrada pela Figura 20.

**Figura 20** – Representação esquemática para Bootstrap e IC no cálculo da garantia na qualidade de efluente

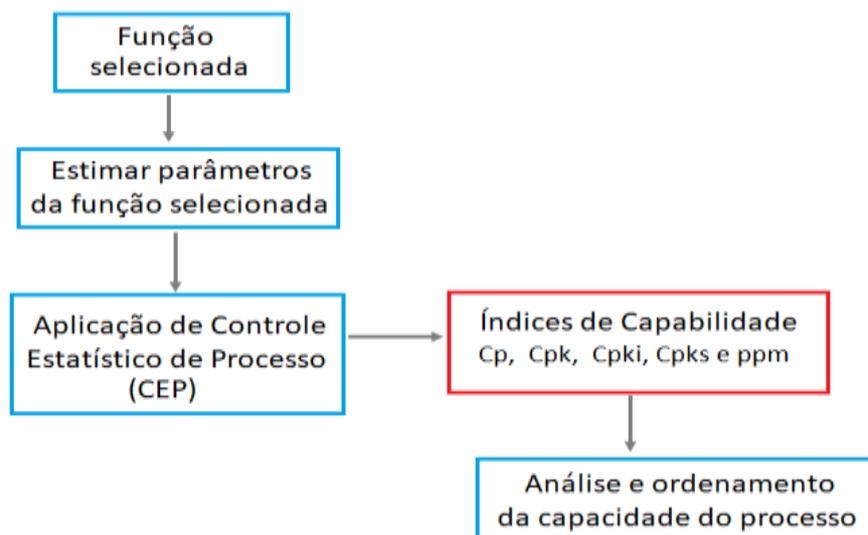


Fonte: o autor, 2018.

### 3.7.3 Representação esquemática para cálculo dos Índices de Capabilidade

Encontrada a melhor FDP para cada conjunto de dados são gerados  $10^6$  números aleatórios. São aplicadas as equações apresentadas no item 2.1 deste trabalho para o cálculo dos índices de capabilidade, conforme as etapas mostradas pela Figura 21.

**Figura 21** – Representação metodológica para a construção dos índices de capacidade utilizando uma FDP geradora de cenários

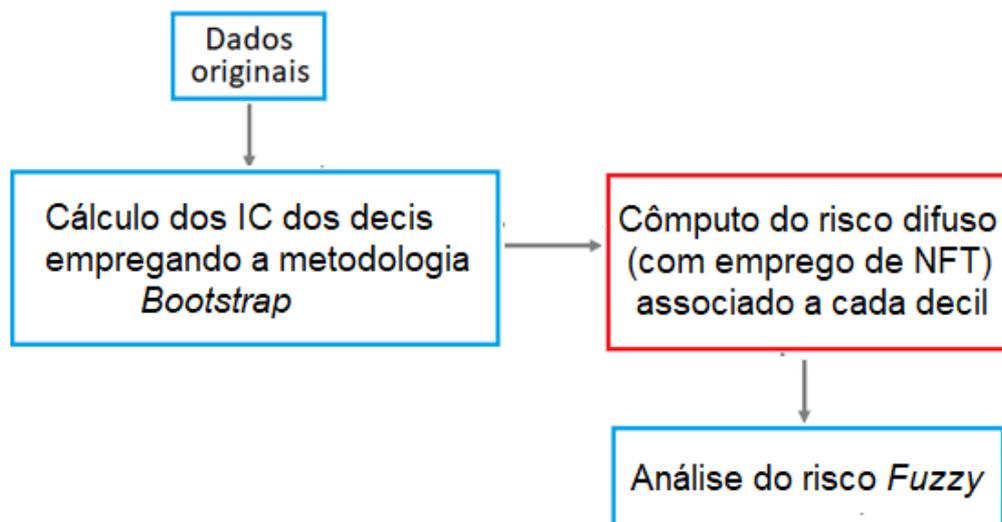


Fonte: o autor, 2018.

### 3.7.4 Representação esquemática para cálculo do risco Fuzzy

Calculado os IC para cada decil utilizando o método *Bootstrap* o risco associado a cada cenário é mensurado utilizando o número Fuzzy Triangular. A Figura 22 mostra este procedimento do cálculo do risco *Fuzzy* para cada decil.

**Figura 22** - Representação metodológica para o cálculo do risco *Fuzzy*.



Fonte: o autor, 2018.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Bootstrap

Conforme expostos anteriormente o *Bootstrap* possui uma forma paramétrica e outra não paramétrica, cada qual é discutida separadamente a seguir. Os Intervalos de Confiança (IC) foram construídos contendo 95% dos dados, para ambas as metodologias. Uma discussão mais detalhada é realizada para a matéria orgânica dissolvida na forma de DBO. Em tópicos posteriores é realizada a análise dos demais parâmetros de qualidade de efluente.

#### 4.1.1 *Bootstrap não paramétrico*

Para o *Bootstrap* não paramétrico os decis foram calculados das reamostras retiradas da “amostra mãe” e com isto levantada a função distribuição acumulada (FDA) com os seus respectivos IC. É possível a construção dos IC para os decis para qualquer parâmetro de qualidade de efluente, neste tópico é explorada a matéria orgânica na forma de DBO.

A Tabela 7 mostra os dados dos IC construídos com a metodologia *Bootstrap* não paramétrico, para facilitar a visualização estes dados foram representados em um gráfico que mostra a função distribuição acumulada vs DBO (Figura 23). Os dados utilizados para construção da Tabela 7 e da Figura 23 foram os dados da DBO afluente a ETE PS<sub>6</sub>. Este procedimento foi seguido para os demais parâmetros de qualidade e ETE consideradas no estudo.

A Figura 23 mostra a FDA para cada decil no intervalo total (0 a 100%), construído com os valores retirados da Tabela 7. Por exemplo, para a concentração afluente de saída de 300 mg/L com o IC contendo 95% dos dados tem-se que em um cenário otimista 37% das amostras são inferiores a este valor e em um cenário pessimista aproximadamente 9% das amostras são inferiores a 300 mg/L. Logo, para esta ETE tem-se que é possível obter apenas [9; 37] % de garantia para amostras com valores inferiores a 300 mg/L. Após a exposição do método utilizado para calcular o IC para a garantia utilizando o *Bootstrap* paramétrico os resultados agora são mostrados preferencialmente utilizando a representação gráfica. As FDA para a matéria orgânica na forma de DBO para cada ETE (PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>)

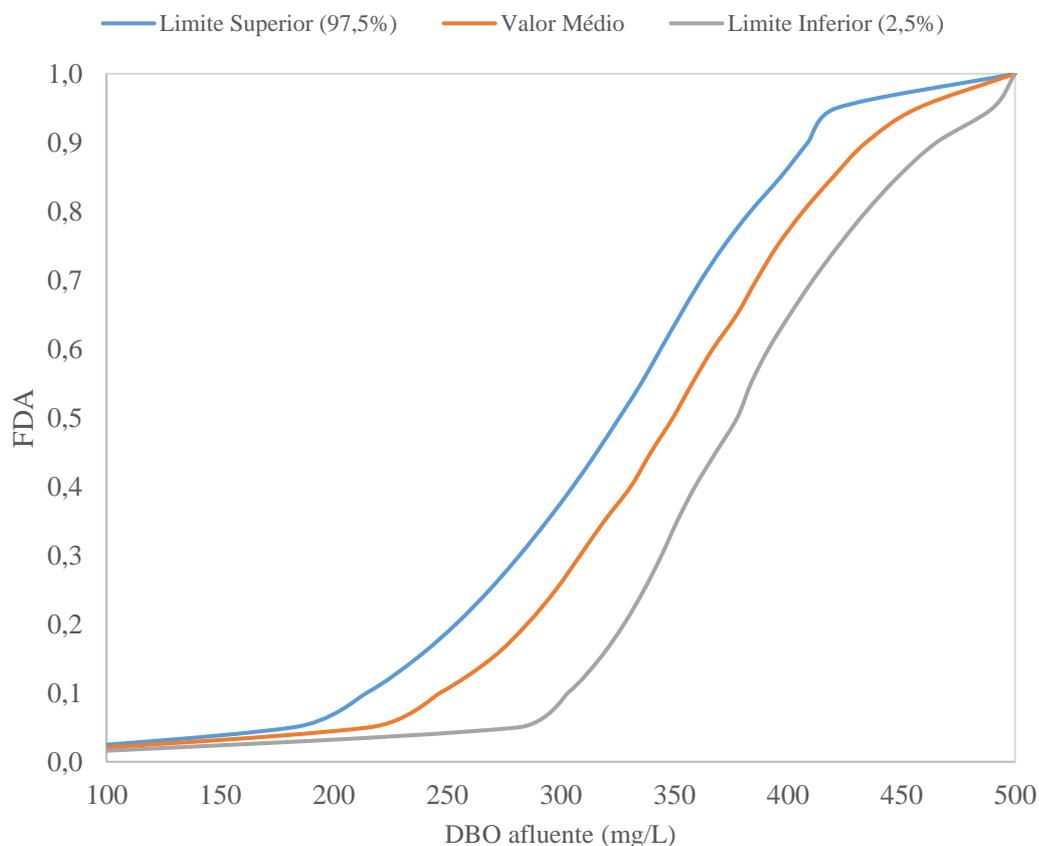
são representados em um mesmo gráfico FDP vs DBO para favorecer a comparação da garantia fornecida para cada ETE.

**Tabela 7** – DBO afluente à PS<sub>6</sub> com Intervalo de Confiança Inferior (2,5%) e Superior (97,5%) nos respectivos decis.

<i>Decil</i>	<i>Limite inferior (2,5%)</i>	<i>Média (50,0%)</i>	<i>Limite superior (97,5%)</i>
10%	214	247	303
20%	254	285	328
30%	282	309	344
40%	305	331	359
50%	326	349	378
60%	344	367	391
70%	362	386	411
80%	384	407	435
90%	409	434	465
100%	500	500	500

Fonte: o autor, 2018.

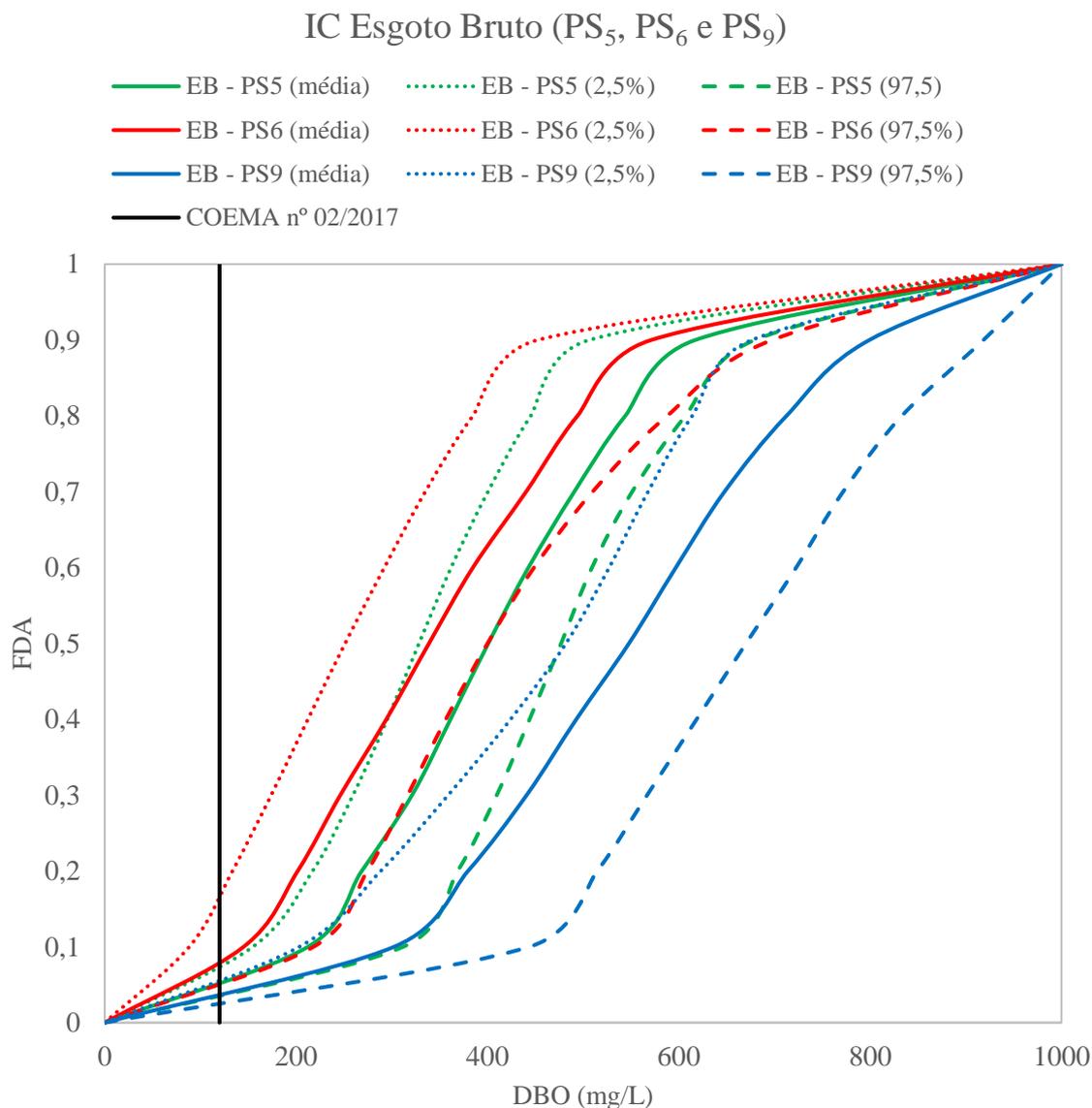
**Figura 23** – Valor da DBO afluente (PS<sub>6</sub>) versus Função de Densidade de Probabilidade Acumulada (FDA).



Fonte: o autor, 2018.

A representação gráfica foi utilizada (Figura 23) para representar a inclinação dos IC em concentrações específicas de DBO, revelando comportamentos que são discutidos adiante. A seguir este procedimento é repetido para demais valores de DBO. Seguida a mesma metodologia apresentada anteriormente para a DBO afluente e efluente as lagoas que constituem as estações de tratamento PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>. Para os valores da DBO a Figura 24 mostra a DBO do EB afluente as estações PS<sub>5</sub> (verde), PS<sub>6</sub> (vermelho) e PS<sub>9</sub> (azul). Onde a linha contínua é o valor médio entre o limite superior (linha pontilhada) onde 2,5% dos dados se encontram acima deste valor e o limite inferior (linha tracejada) onde 97,5% dos dados encontram-se acima deste valor. A linha negra na vertical é referente a resolução COEMA nº 02/2017, que é o padrão de aceitação utilizado neste trabalho para matéria orgânica na forma de DBO.

**Figura 24** – IC para a FDA simulada utilizando o Bootstrap paramétrico para o Esgoto Bruto.



Fonte: o autor.

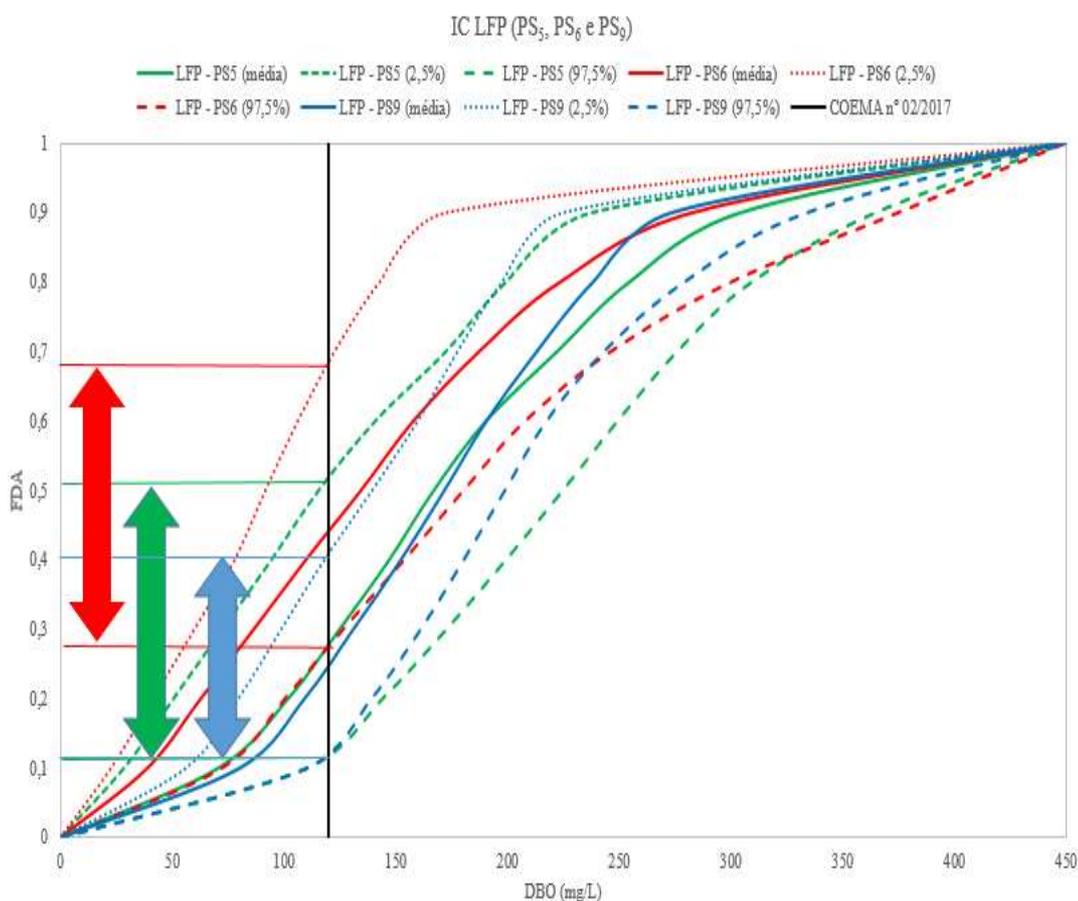
A Figura 24 mostra a faixa de variação do Esgoto Bruto. A região entre 400 e 600 mg/L de DBO no esgoto bruto apresenta Intervalos de Confiança mais extensos. Logo, nesta região é apresentada alta maior variabilidade de valores e, portanto, de melhor caracterização da concentração de matéria orgânica como DBO. Apesar disso, observa-se que a concentração do EB da estação PS<sub>9</sub> apresenta maior intervalo de flutuação, alcançando valores de até 1000 mg/L. Pode, assim, alcançar valores de esgoto mais concentrado em todas as faixas. O esgoto bruto é tratado pela lagoa facultativa primária onde é submetido ao primeiro estágio do tratamento para remoção de matéria orgânica e outros poluentes. A Figura

25 mostra o esgoto tratado efluente a lagoa facultativa e os respectivos IC para os quantis. A linha vertical é o limite de lançamento de efluentes para a DBO estipulado pelo Conselho Estadual de Meio Ambiente (COEMA) nº 02 de 21 de fevereiro de 2017 de DBO de até 120 mg/L.

A garantia para o tratamento de esgoto, conforme o padrão sugerido pela COEMA nº 02 de 2017, torna-se muito clara quando exposta conforme apresentado pela Figura 25. Onde as cetras verde, vermelha e azul mostram o IC para a garantia e a remoção de poluentes. As ETE's PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> foram colocadas em uma mesma figura a título de comparação da garantia para cada lagoa. As cetras foram colocadas apenas na Figura 25 para evitar imagem demasiadamente poluídas.

A decisão de expor os resultados conforme a imagem 25 mostra que os IC's são extremamente sensível a mudança do padrão de qualidade e que o comprimento do intervalo pode variar conforme a lagoa pertencente a cada estação. Este recurso visual pode orientar o gestor se o risco máximo (limite inferior do IC) está de acordo com o almejado ou se a estação opera próximo ao intervalo superior (condição excelente que minimiza o risco e aumenta a garantia do processo).

**Figura 25** – IC com base em *Bootstrap* construídos para a DBO efluente tratado proveniente das lagoas facultativas primárias.



Fonte: o autor, 2018.

Adotado o critério de confiabilidade, a conformidade a COEMA nº 02/2017 o esgoto tratado pela lagoa facultativa pertencente a estação PS<sub>6</sub> mostra que em uma situação otimista 65% das amostras possuem concentração inferiores a 120 mg/L e em uma situação pessimista apenas 31% das amostras possuem concentração inferior a 120 mg/L. Para a lagoa facultativa PS<sub>5</sub> observa-se que o intervalo está entre [0,1; 0,5] e para a lagoa PS<sub>9</sub> este valor encontra-se no intervalo entre [0,13; 0,37]. Como se pode observar a maior quantidade de amostras em conformidade tanto para o limite superior como para o inferior são provenientes da lagoa facultativa pertencente a estação PS<sub>6</sub>. A lagoa PS<sub>5</sub> apresenta o pior resultado em relação a extensão da faixa de variação dos IC em relação os decis, apresentando uma ampla faixa de possibilidades de concentração de DBO efluente.

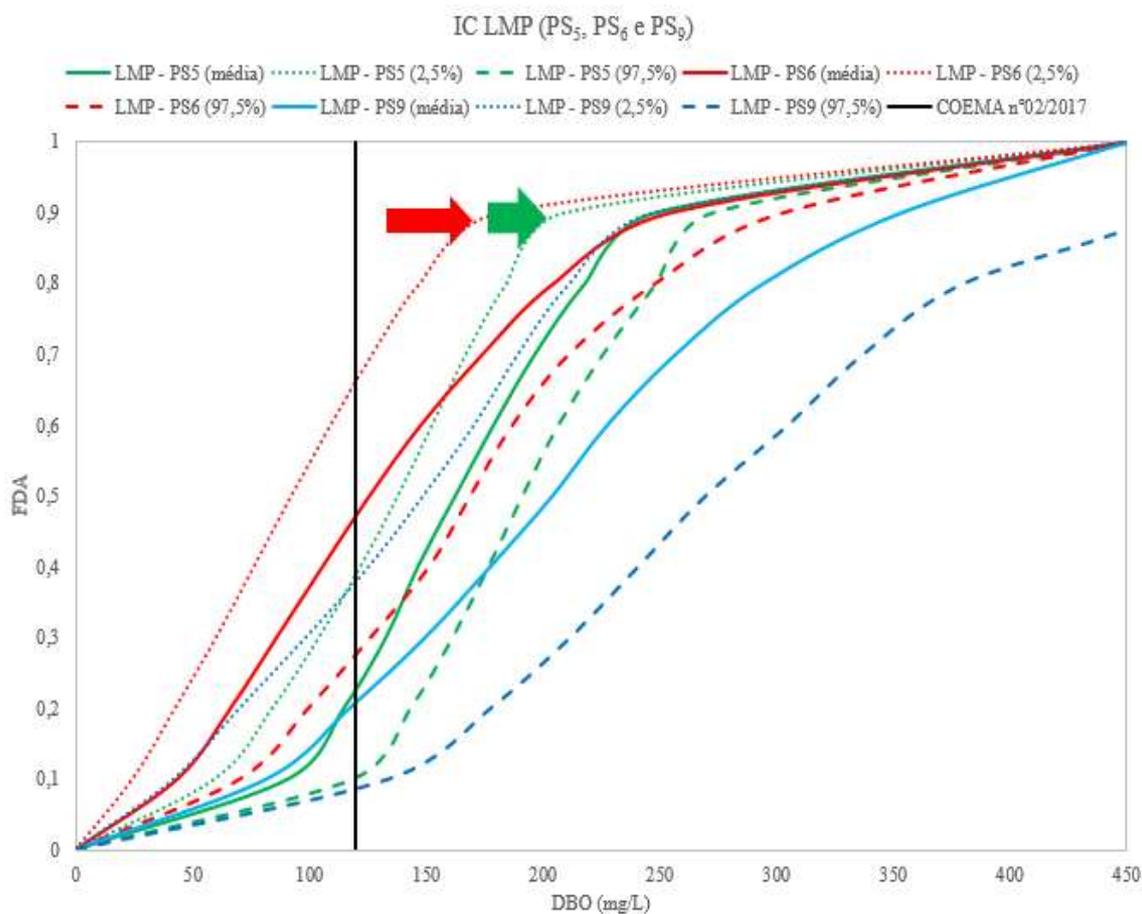
Através da Figura 25 observa-se cenários “ótimos” de cada lagoa, como mostrado pelo limite superior dos intervalos de confiança, não conseguem satisfazer com eficiência o padrão de 120 mg/L. Por exemplo, a lagoa PS<sub>6</sub> consegue no máximo 65% das amostras em conformidade com esta portaria. De início, deve-se somar esforços para que a ETE venha a operar próximo a valores ótimos como mostrados pela linha pontilhada em cada estação. Logo, torna-se claro a necessidade de um tratamento posterior para a remoção de matéria orgânica. O próximo passo do tratamento é a lagoa de maturação primária.

O procedimento similar para a construção dos IC é realizado para a etapa posterior do tratamento que compreende a LMP. Adotado o mesmo critério de confiabilidade exposto pela COEMA n° 02 a Figura 25 mostra os IC para o tratamento. Para a LMP para das ETE PS<sub>6</sub> no máximo 65% das amostras encontram-se inferior ao valor de 120 mg/L e no mínimo 27% das amostras possuem valor inferior a este valor. Para as demais lagoas o limite para a PS<sub>5</sub> está entre [0,1; 0,4] e para a PS<sub>9</sub> está entre [0,04; 0,33]. A Estação de Tratamento que apresenta a melhor LMP em relação ao número de amostras dentro da especificação é a PS<sub>6</sub>. Logo, para o critério de confiabilidade adotando o número de amostras dentro da norma tem-se que a LMP proveniente da ETE PS<sub>6</sub> destaca-se frente as demais.

Um comportamento interessante mostrado pelos intervalos de confiança montados através da metodologia *Bootstrap* para a lagoa de maturação primária (Figura 26) é notado. Que para LMP provenientes das estações PS<sub>5</sub> e PS<sub>6</sub> existe um valor limite no qual o quartil superior mantém-se quase contante. Ou seja, nesta região existe uma matéria orgânica dissolvida tão elevada que compromete a remoção de matéria orgânica pelas lagoas. Note que como mostrado pela Figura 25 o número de amostras em conformidade não cresce a partir de um valor limite. O valor de efluente de 170 mg/L para a estação PS<sub>6</sub> e 200 mg/L para a estação PS<sub>5</sub> mostra um limite no qual a remoção de matéria orgânica pelas lagoas torna-se insatisfatório. Esta mudança pode ser entendida como o limite operacional da lagoa. Mostrando a necessidade de etapas posteriores do tratamento. Sempre quando houver uma mudança abrupta da inclinação das curvas referente a distribuição acumulada o gestor deve estar atento as limitações da estação ao atendimento a alguma meta específica. Algum fator externo à estação pode estar agindo de modo a diminuir a frequência de amostras em conformidade. Ou seja, sempre quando houver uma mudança de inflexão nas curvas é preferível um estudo acurado de modo a analisar os motivos deste comportamento, o ponto no qual é observado este fenômeno é mostrado por cetras na Figura 26. Já que a partir deste limite a estação não consegue reduzir o número de amostras em não conformidade. A

utilização da metodologia *Bootstrap* permitiu analisar o comportamento dos intervalos de confiança em um amplo de intervalo de condições de efluente.

**Figura 26** – IC construídos para a DBO efluente proveniente das lagoas de maturação primária.



Fonte: o autor, 2018.

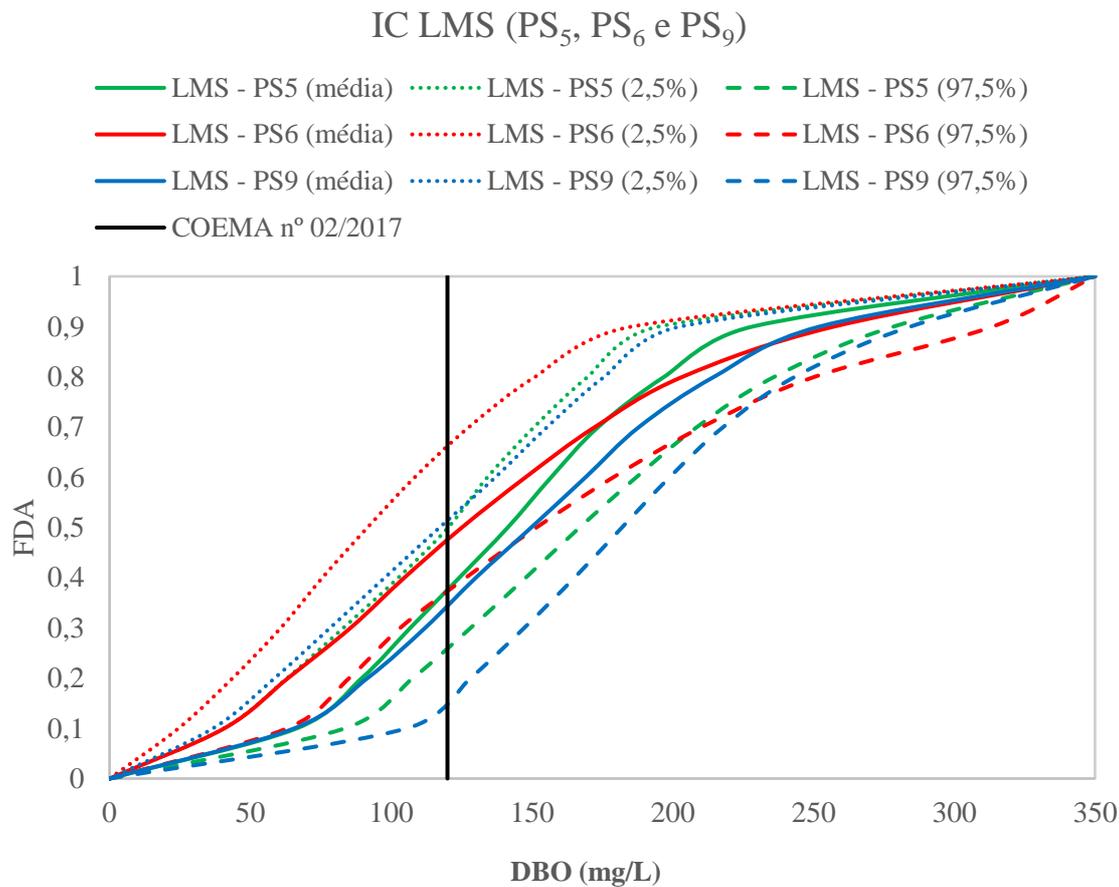
Finalmente, a última etapa do tratamento dos sistemas de tratamento são as LMS, antes do efluente ser despejado no corpo receptor. A Figura 27 mostra os IC construídos com os valores de DBO efluente, das LMS constituintes das estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>. De maneira análoga as demais existe uma região onde as curvas são crescentes e a partir de um dado valor de DBO efluente a inclinação das curvas se mantém constante mostrando que para uma determinada faixa de concentração a estação perde a capacidade de adequação, ou seja, a frequência de amostras em conformidade é prejudicada. Para a lagoa PS<sub>5</sub> o limite no qual o número de amostras em conformidade torna-se constante é por volta de 200 mg/L, para a

lagoa PS<sub>6</sub> é de 150 mg/L e por outro lado este comportamento não é observado para a lagoa PS<sub>9</sub>.

Utilizado o mesmo o limite de 120 mg/L observa-se que os limites expostos estão entre [0,35; 0,70] para a LMS da ETE de PS<sub>6</sub>, [0,2; 0,47] para a LMS da ETE de PS<sub>5</sub> e [0,20; 0,55] para a ETE de PS<sub>9</sub>. Nota-se que a ETE PS<sub>6</sub> é a estação que possui maiores probabilidades de atender a resolução COEMA 02/2017 de 120 mg/L, como mostrado pelo intervalo de confiança superior que consegue obter até 70% das amostras em conformidade. Entretanto, estes valores ainda não são satisfatórios frente a concordância com a portaria. Faz-se necessário um próxima etapa de tratamento.

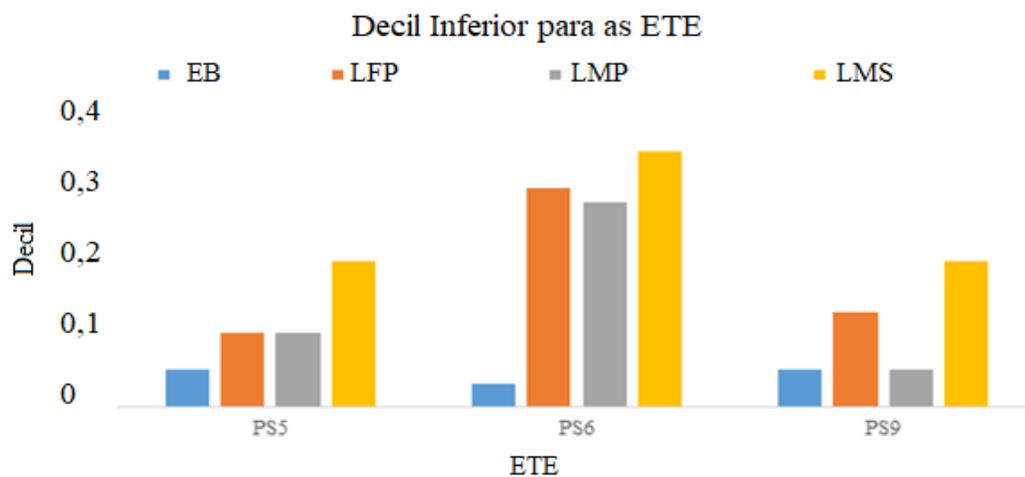
Através desta metodologia é possível avaliar o intervalo de garantia para cada etapa do tratamento conforme a norma de qualidade que deseja-se utilizar. Através da observação das Figuras 25, 26 e 27 observa-se que a Estação PS<sub>6</sub> consegue alcançar probabilidades superiores de amostras inferiores a 120 mg/L (como exigido pela legislação). Outro ponto a ser explorado nesta metodologia é a extensão dos intervalos entre os limites inferiores (2,5%) e superiores (97,5%). Um processo satisfatório deve buscar intervalos mais estreitos entre os limites de confiança tendo em mente que a elevada aleatoriedade dos dados é prejudicial para a obtenção de uma meta. Percebe-se que o intervalo de DBO efluente entre 100 e 200 mg/L nas três ETE os intervalos entre o limite inferior e superior dos IC é demasiadamente extenso. Logo, sempre quando o gestor for trabalhar nesta faixa este deve tomar as devidas precauções referente a aleatoriedade dos dados neste intervalo. Para facilitar a interpretação dos resultados do *Bootstrap* não paramétrico os dados são mostrados através de gráfico de barras. Para cada imagem é mostrado o Intervalo de Confiança (inferior, médio e superior) para a DBO do esgoto bruto afluente e efluente as lagoas facultativas, de maturação primária e secundária.

**Figura 27** – IC construídos para a DBO efluente proveniente das lagoas de maturação secundária.



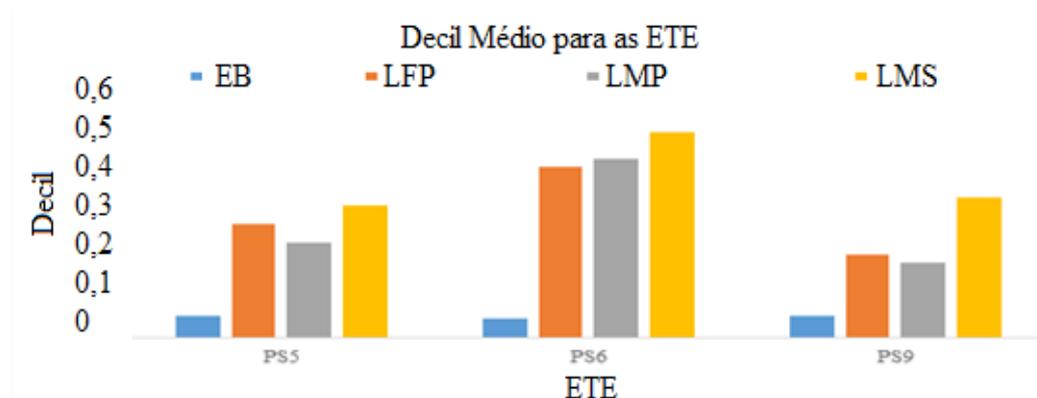
Fonte: o autor, 2018.

**Figura 28** – Valores dos limites inferiores dos IC (pior cenário possível) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.



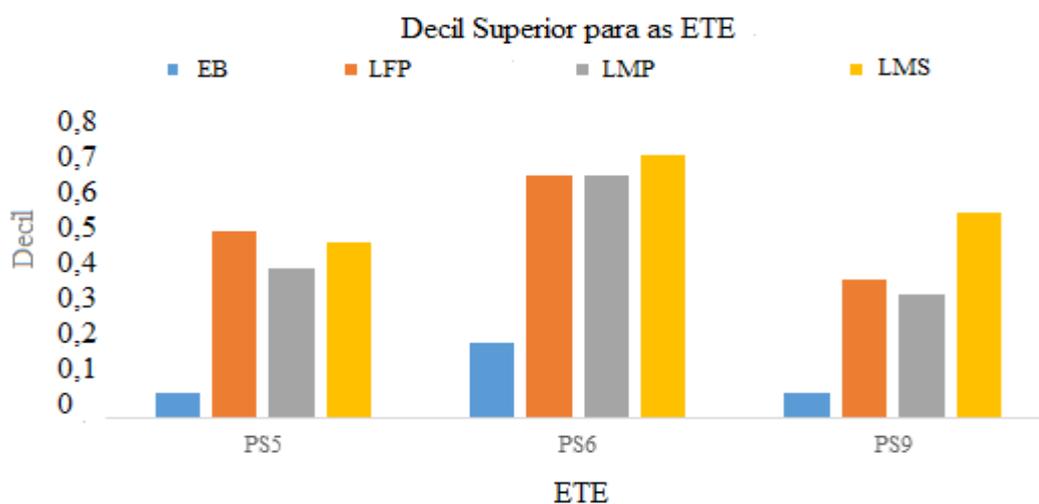
Fonte: o autor, 2018.

**Figura 29**– Valores dos limites médios dos IC (cenário mais conservador) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.



Fonte: o autor, 2018.

**Figura 30** – Valores dos limites superiores dos IC (cenário mais otimista) para os quantis das amostras com valores inferiores a 120 mg/L para as Estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.



Fonte: o autor, 2018.

De início é fácil observar uma situação peculiar entre a lagoa facultativa e a de maturação primária. Em alguns casos tem-se que o número de amostras em não conformidade aumenta após a passagem pela LMP, indicando, provavelmente, o excesso de biomassa algal (COUTO et al., 2018). A análise das Figuras 28, 29 e 30 mostra que a ETE PS<sub>6</sub> se destaca frente às demais, tornando possível diminuir o número de amostras em não conformidade a medida que o processo de remoção de poluentes passa pelas lagoas. Três cenários distintos são mostrados representados pela Figura 28 (pior cenário), Figura 29

(cenário mais conservador) e a Figura 30 (cenário mais otimista), a partir destes é possível analisar as potencialidades e fragilidades de cada sistema e os limites de operação de cada. Conclui-se que uma ETE bem operada tende a ter valores próximos aos mostrados pela Figura 30 e por outro lado quando o processo não está sendo bem operado apresenta valores próximos a aqueles mostrados pela Figura 28.

Através do gráfico de barras mostrados do pelas Figuras 28, 29 e 30 observa-se que a lagoa PS<sub>6</sub> consegue alcançar o maior número de amostras em conformidade nos três cenários (limite inferior, médio e superior). Logo, das três estações a PS<sub>6</sub> consegue oferecer maior confiabilidade ao tratamento. Complementar ao conceito de confiabilidade é analisado o conceito de *robustez* e *resiliência*. De início observa-se que o desprendimento de biomassa algal efluente a lagoa facultativa (evento adverso) traz prejuízos as ETE PS<sub>5</sub> e PS<sub>9</sub> mas o número de amostras em conformidade não é prejudicada, ou este efeito é pouco sentido, na ETE PS<sub>6</sub>. Logo, conclui-se que esta consegue manter-se estável frente ao fenômeno de carreamento de biomassa algal, logo, a ETE PS<sub>6</sub> é a mais robusta frente as demais. Por fim, analisado a resiliência das estações. Esta característica é bastante relacionada com a qualidade do esgoto bruto que é alimentado. Por exemplo, para o limite inferior e médio do esgoto bruto a estação PS<sub>6</sub> é bastante concentrado, devido o pouquíssimo número de amostras em conformidade. Entretanto, a estação PS<sub>6</sub> consegue bons resultados apesar do esgoto bruto bastante concentrado.

Realizada a análise de confiabilidade (número de amostras em conformidade), robustez (permanecer firme frente ao carreamento de biomassa algal) e resiliência (grau de adequação frente a afluentes concentrados) tem-se que a ETE PS<sub>6</sub> é a que se destaca frente as demais. Vale salientar que esta análise foi realizada utilizando uma garantia de 120 mg/L de DBO para todas as lagoas, e as conclusões aqui são sensíveis a este padrão de qualidade adotado. Logo, caso adotado outro padrão de qualidade os resultados serão distintos.

A observação dos três cenários (pior, melhor e cenário mediano) distintos conforme mostrado pelas Figuras 28, 29 e 30 torna possível a adoção de uma política planejamento estratégico. O gestor responsável pela ETE deve canalizar esforços de modo a alcançar o maior número possível de amostras em conformidade próximo do valor limite de confiança superior. Existe a possibilidade da meta buscada estar além da banda superior de confiança. Este fato mostra a ineficácia de uma estação em alcançar os objetivos almejados. Logo, uma alteração substancial deve ser ponderada. Por exemplo, para a ETE PS<sub>6</sub> a melhor situação disponível para o despejo de efluentes alcança 70% de amostras em conformidade com a legislação. É sugerido uma etapa posterior para remoção de matéria

orgânica, já que em condições excelentes é alcançado no máximo 70% de conformidade com a norma ambiental.

## 4.2 *Bootstrap* paramétrico

A metodologia *Bootstrap* paramétrica ocorre da seguinte forma: para cada reamostra tirada da amostra mãe são estimados os parâmetros da função distribuição de probabilidade que melhor tem aderência com os resultados. Os IC dos valores estatísticos procurados são retirados do conjunto de curvas obtidas a partir da função distribuição de probabilidade acumulada construído a partir de cada reamostragem. Para a metodologia paramétrica os IC são retirados a partir da região do gráfico onde 95% das curvas estão localizadas entre os limites de confiança.

### 4.2.1 *Estimativa de parâmetros e teste de aderência*

Para a matéria orgânica na forma de DBO para o Esgoto Bruto (EB) e a efluente proveniente das lagoas facultativas (LFP) e de Maturação Primária (LMP) e Secundária (LMS) são estimados os parâmetros para todas as funções de distribuição apresentadas pelas equações de (9-15). Para que o trabalho não se torne muito extenso é realizada esta análise apenas para a DBO das lagoas PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.

A estatística descritiva para o Esgoto bruto e a DBO efluente das LFP, LMP e LMS encontra-se na Tabela 8. O teste de aderência é fundamental para que a metodologia *Bootstrap* paramétrico tenha bons resultados. Já que a partir das funções distribuição de probabilidade que novas amostras (ou reamostras *Bootstrap*) são geradas. Logo, sempre é preferível seguir esta metodologia, de estimativa dos parâmetros e posteriormente o teste de aderência. Vale lembrar que pode existir mais de uma função de distribuição de probabilidade com boa aderência com os resultados. Entretanto, é feito um esforço para aferir a “melhor função” que gere os melhores cenário a partir de uma amostra mãe.

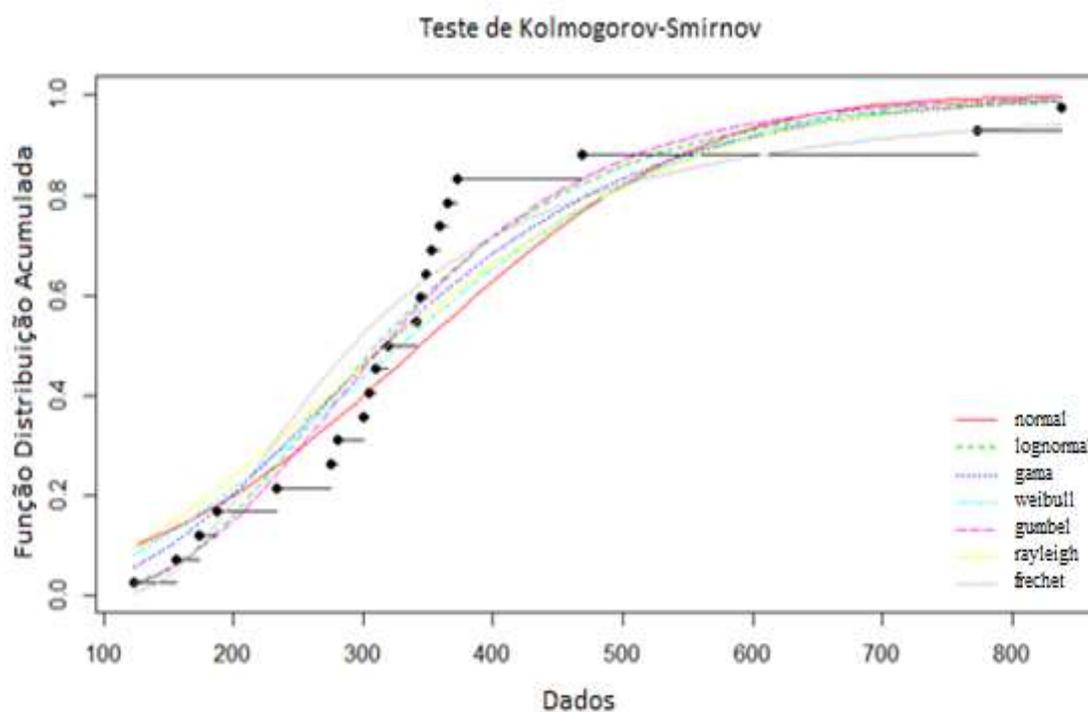
A Figura 31 mostra uma representação gráfica do teste de *Kolmogorov-Smirnov* para diversas FDP para os dados do EB afluente a lagoa PS<sub>6</sub>. Note que são comparadas diversas curvas da distribuição empírica com as teóricas (Normal, Lognormal, *Gama*, *Weibull*, *Gumbel*, *Rayleigh*, *Frechet*).

**Tabela 8** – Estatística descritiva para o EB e efluente as lagoas LFP, LMP e LMS nas séries PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.

Série	EB				
	Mínimo	Média	Máximo	$\sigma$	CV
PS <sub>5</sub>	122	416	752	167	0,40
PS <sub>6</sub>	123	344	838	175	0,51
PS <sub>9</sub>	221	547	991	196	0,36
LFP					
PS <sub>5</sub>	32	179	374	98	0,55
PS <sub>6</sub>	25	148	425	102	0,69
PS <sub>9</sub>	55	180	398	81	0,45
LMP					
PS <sub>5</sub>	63	167	328	64	0,38
PS <sub>6</sub>	34	141	364	89	0,63
PS <sub>9</sub>	50	212	539	116	0,55
LMS					
PS <sub>5</sub>	38	148	316	67	0,46
PS <sub>6</sub>	26	118	259	59	0,50
PS <sub>9</sub>	24	141	211	57	0,40

Fonte: o autor.

**Figura 31** – Teste de Kolmogorov-Smirnov realizado através do Software *Rstudio* comparando as Função de Distribuição Acumulada.



Fonte: o autor, 2018.

Outro fator pertinente ao estudo das FDP é que sempre é aconselhável realizar este procedimento aqui mostrado para aferir a melhor função de distribuição referente a aderência aos dados. Se dados de DBO efluente para uma determinada coleta de dados chegaram na distribuição *Weibull* como aquela que melhor tem aderência aos dados este fato não garante que para outros dados de DBO efluente a distribuição *Weibull* seria a mais indicada. Logo, por segurança sempre é indicado seguir toda a metodologia proposta anteriormente. Os parâmetros mostrados pela Tabela 9 são referentes as FDP estudadas neste trabalho para as ETE para amostras advindas do EB, LFP, LMP e LMS.

**Tabela 9** – Parâmetros das FDP estimadas para a DBO no EB e nos efluentes do estudo para as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>.

<i>Série e Amostra</i>	<i>Função e Parâmetro</i>						
	<i>Normal</i>	<i>Lognormal</i>	<i>Gama</i>	<i>Weibull</i>	<i>Gumbel</i>	<i>Rayleigh</i>	<i>Frechet</i>
<b>EB</b>	$\mu$ e $\sigma$	$\log \mu$ e $\log \sigma$	$\gamma$ e $\beta$	$\gamma$ e $\beta$	$\mu$ e $\beta$	$\sigma$	$\alpha$ e $\beta$
<b>PS<sub>5</sub></b>	416,3	5,94	6,50	2,75	338,39	315,43	2,07
	163,3	0,42	0,01	468,57	140,39		305,44
<b>PS<sub>6</sub></b>	343,9	5,73	4,07	2,13	273,62	271,29	2,32
	170,3	0,44	0,01	389,8	114,6		249,18
<b>PS<sub>9</sub></b>	546,5	6,24	8,19	3,04	456,70	409,29	2,63
	190,8	0,36	0,015	611,70	161,247		426,03
<b>LFP</b>	$\mu$ e $\sigma$	$\log \mu$ e $\log \sigma$	$\gamma$ e $\beta$	$\gamma$ e $\beta$	$\mu$ e $\beta$	$\sigma$	$\alpha$ e $\beta$
<b>PS<sub>5</sub></b>	178,6	4,969	3,49	1,90	131,09	143,21	1,22
	95,5	0,74	0,019	200,50	87,402		96,40
<b>PS<sub>6</sub></b>	147,7	4,76	2,20	1,58	104,30	125,90	1,39
	99,5	0,70	0,014	165,55	69,72		82,07
<b>PS<sub>9</sub></b>	180,0	5,09	5,24	2,40	144,38	138,61	1,97
	78,6	0,45	0,02	203,20	63,62		128,17
<b>LMP</b>	$\mu$ e $\sigma$	$\log \mu$ e $\log \sigma$	$\gamma$ e $\beta$	$\gamma$ e $\beta$	$\mu$ e $\beta$	$\sigma$	$\alpha$ e $\beta$
<b>PS<sub>5</sub></b>	166,85	5,04	7,15	2,85	136,97	125,93	2,23
	62,38	0,41	0,04	187,20	55,28		124,28
<b>PS<sub>6</sub></b>	140,69	4,75	2,74	1,76	102,46	116,17	1,61
	84,85	0,63	0,019	158,93	62,61		84,67
<b>PS<sub>9</sub></b>	212,20	5,21	3,49	1,99	161,90	170,14	1,71
	113,48	0,54	0,016	240,3	84,51		138,58
<b>LMS</b>	$\mu$ e $\sigma$	$\log \mu$ e $\log \sigma$	$\gamma$ e $\beta$	$\gamma$ e $\beta$	$\mu$ e $\beta$	$\sigma$	$\alpha$ e $\beta$
<b>PS<sub>5</sub></b>	147,87	4,878	5,07	2,382	116,68	114,39	1,62
	65,629	0,52	0,034	166,7	57,64		98,80
<b>PS<sub>6</sub></b>	118,13	4,616	4,22	2,16	90,189	92,88	1,43
	57,462	0,614	0,03	133,2	51,33		72,53
<b>PS<sub>9</sub></b>	140,75	4,81	6,532	2,85	111,44	106,72	1,34
	55,06	0,59	0,046	157,1	59,38		88,38

Fonte: o autor, 2018.

Estimados os parâmetros para as funções distribuição de probabilidade o próximo passo é calcular o *p-value* e analisar a “bondade” do ajuste. Lembrando que a significância do teste foi de 5% ( $\alpha = 0,05$ ). Ou seja, para o teste KS o conjunto de dados que obterem um *p-value* superior a 0,05 estão hábil a participar do critério de desempate (teste de AC).

**Tabela 10** – Valor das estatísticas (*p-value*) com efluente tratado na saída das ETE.

<i>Série e Amostra</i>	<i>Função de Distribuição de Probabilidade</i>						
	<i>Normal</i>	<i>Lognormal</i>	<i>Gama</i>	<i>Weibull</i>	<i>Gumbel</i>	<i>Rayleigh</i>	<i>Frechet</i>
<b>EB</b>							
<b>PS<sub>5</sub></b>	0,839	0,994	0,982	0,938	<b>0,998</b>	0,486	0,731
<b>PS<sub>6</sub></b>	0,045	0,331	0,189	0,0933	<b>0,311</b>	0,128	0,252
<b>PS<sub>9</sub></b>	0,871	0,942	<b>0,994</b>	0,859	0,962	0,274	0,587
<b>LFP</b>							
<b>PS<sub>5</sub></b>	<b>0,956</b>	0,234	0,558	0,703	0,644	0,787	0,085
<b>PS<sub>6</sub></b>	0,369	0,906	<b>0,927</b>	0,807	0,777	0,222	0,448
<b>PS<sub>9</sub></b>	0,243	0,191	<b>0,323</b>	0,348	0,300	0,184	0,083
<b>LMP</b>							
<b>PS<sub>5</sub></b>	0,805	0,281	<b>0,473</b>	0,669	0,377	0,139	0,112
<b>PS<sub>6</sub></b>	0,520	0,961	<b>0,978</b>	0,892	0,919	0,582	0,635
<b>PS<sub>9</sub></b>	0,452	<b>0,978</b>	0,917	0,728	0,958	0,720	0,695
<b>LMS</b>							
<b>PS<sub>5</sub></b>	<b>0,604</b>	0,128	0,344	0,404	0,292	0,138	0,032
<b>PS<sub>6</sub></b>	<b>0,904</b>	0,270	0,462	0,668	0,508	0,551	0,196
<b>PS<sub>9</sub></b>	<b>0,587</b>	0,182	0,272	0,465	0,361	0,192	0,096

Fonte: o autor, 2018.

Os dados mostrados pela Tabela 10 revelam que apenas a função distribuição de probabilidade normal para o EB afluente da lagoa PS<sub>6</sub> não possui um modelo estatístico satisfatório conforme o teste de *Kolmogorov-Smirnov* com  $\alpha = 0,05$ . O próximo passo é o teste de *Aikake's* (AIC). Onde o *menor* valor é aquele que relativo aos demais se apresenta como a melhor representatividade da distribuição dos dados. Os dados para o teste de *Aikake's* são mostrados na Tabela 11. As FDP que obtiveram melhor aderência foram usados em negrito tanto na Tabela 10 como na Tabela 11.

**Tabela 11** –Valores de AIC dos modelos para a matéria orgânica em forma de DBO.

<i>Série e Amostra</i>	<i>Função de Distribuição de Probabilidade</i>						
	<i>Normal</i>	<i>Lognormal</i>	<i>Gama</i>	<i>Weibull</i>	<i>Gumbel</i>	<i>Rayleigh</i>	<i>Fréchet</i>
<b>EB</b>							
<b>PS<sub>5</sub></b>	277,589	277,7548	<b>276,6278</b>	276,6564	276,9844	277,7568	284,7902
<b>PS<sub>6</sub></b>	279,3762	<b>270,5295</b>	272,4424	275,5038	270,930	273,6877	273,0644
<b>PS<sub>9</sub></b>	284,1596	282,9532	282,5167	283,7586	<b>282,8708</b>	287,0998	288,0775
<b>LFP</b>							
<b>PS<sub>5</sub></b>	243,1182	247,7053	245,5907	242,2757	244,5575	<b>240,3518</b>	254,3180
<b>PS<sub>6</sub></b>	256,7981	248,9631	<b>248,8979</b>	249,6722	250,4207	249,9686	253,3904
<b>PS<sub>9</sub></b>	235,338	233,4387	232,5774	233,7317	<b>232,5376</b>	232,892	239,7044
<b>LMP</b>							
<b>PS<sub>5</sub></b>	226,0882	226,7182	<b>225,5292</b>	225,5830	226,1140	227,2756	232,8711
<b>PS<sub>6</sub></b>	250,1156	244,1348	244,1361	244,9389	245,390	<b>243,5378</b>	247,4713
<b>PS<sub>9</sub></b>	250,0252	245,4363	245,1336	246,3084	245,2631	<b>244,3092</b>	250,5178
<b>LMS</b>							
<b>PS<sub>5</sub></b>	228,1186	230,1832	227,9938	227,0784	227,8189	<b>226,0351</b>	238,2529
<b>PS<sub>6</sub></b>	222,8029	225,9278	223,6326	221,6692	223,3074	<b>219,8532</b>	233,2419
<b>PS<sub>9</sub></b>	<b>210,2452</b>	221,1388	219,0541	211,2401	215,9470	211,9970	231,0008

Fonte: o autor, 2018.

Estimado os Parâmetros das funções distribuição de probabilidade, realizado o teste de *Kolmogorov-Smirnov* e por fim analisado o melhor modelo através do teste de AIC a Tabela 12 mostra a melhor FDP para cada série de dados. Nota-se que o *menor* valor de AIC revela o *melhor* modelo. Vale ressaltar que não existe uma função que modele preferencialmente a DBO, por exemplo, sugere-se que sempre realize o procedimento mostrado neste item. Aferida qual é a melhor função distribuição de probabilidade que modela particularmente cada série de dados torna-se possível partir para a apresentação da metodologia do *Bootstrap* paramétrico. O procedimento mostrada nesta seção para encontrar as melhores FDP pode ser replicado para qualquer parâmetro de qualidade de efluente ou dados que se desejar analisar.

**Tabela 12** – Função Distribuição de Probabilidade selecionada através do critério de AIC para a matéria orgânica dissolvida na forma de DBO.

<i>Série e Amostra</i>	<i>FDP selecionada</i>
<b>EB</b>	
PS <sub>5</sub>	Weibull
PS <sub>6</sub>	Lognormal
PS <sub>9</sub>	Gama
<b>LFP</b>	
PS <sub>5</sub>	Rayleigh
PS <sub>6</sub>	Gama
PS <sub>9</sub>	Gumbel
<b>LMP</b>	
PS <sub>5</sub>	Gama
PS <sub>6</sub>	Rayleigh
PS <sub>9</sub>	Rayleigh
<b>LMS</b>	
PS <sub>5</sub>	Rayleigh
PS <sub>6</sub>	Rayleigh
PS <sub>9</sub>	Normal

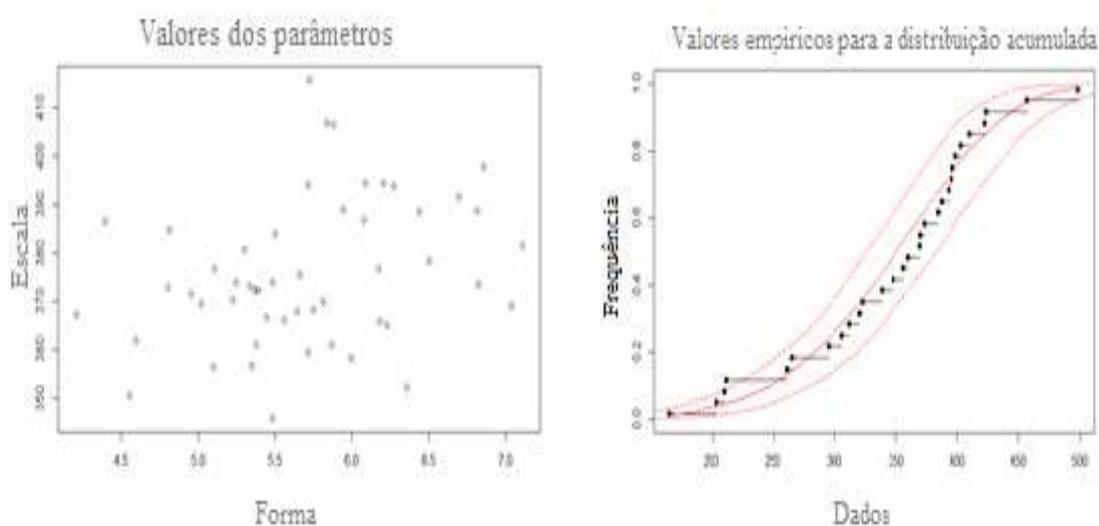
Fonte: o autor, 2018.

#### 4.2.2 IC do Bootstrap paramétrico

Dando procedimento a construção dos IC de confiança através da metodologia *Bootstrap* agora utiliza-se a rota paramétrica. Este procedimento só tornou-se possível após o teste de aderência com as FDP realizado anteriormente. A Figura 32, a esquerda, mostra os pares de parâmetros estimados para 50 reamostragens para uma função distribuição *Weibull*, e a direita mostra a região entre as duas curvas tracejadas (bandas de confiança de 2,5% e 97,5%) no qual 95% das curvas estão presentes. A curva vermelha contínua mostra o valor médio entre os limites de confiança. Concernente o tratamento de Esgoto a curva tracejada superior mostra um cenário otimista, onde 97,5% dos valores estão abaixo de um valor pré-fixado. Já a curva inferior mostra um cenário pessimista, onde 2,5% dos valores estão abaixo de um conjunto de valores representados pela curva tracejada. E a curva cheia mostra um cenário conservados, esperado, para a remoção de poluentes.

Para um exemplo com os dados de DBO efluente proveniente da LFP da ETE PS<sub>6</sub> apresentado pela Figura 32 observa-se que no máximo 35% das amostras coletadas são inferiores a DBO de 300 mg/L (limite superior) e pelo menos 18% das amostras são superiores a 300 mg/L (limite inferior). Desta maneira é construído o IC para este conjunto de dados no plano cartesiano onde os eixos são Função Probabilidade Acumulada (CDF) vs DBO (mg/L) efluente.

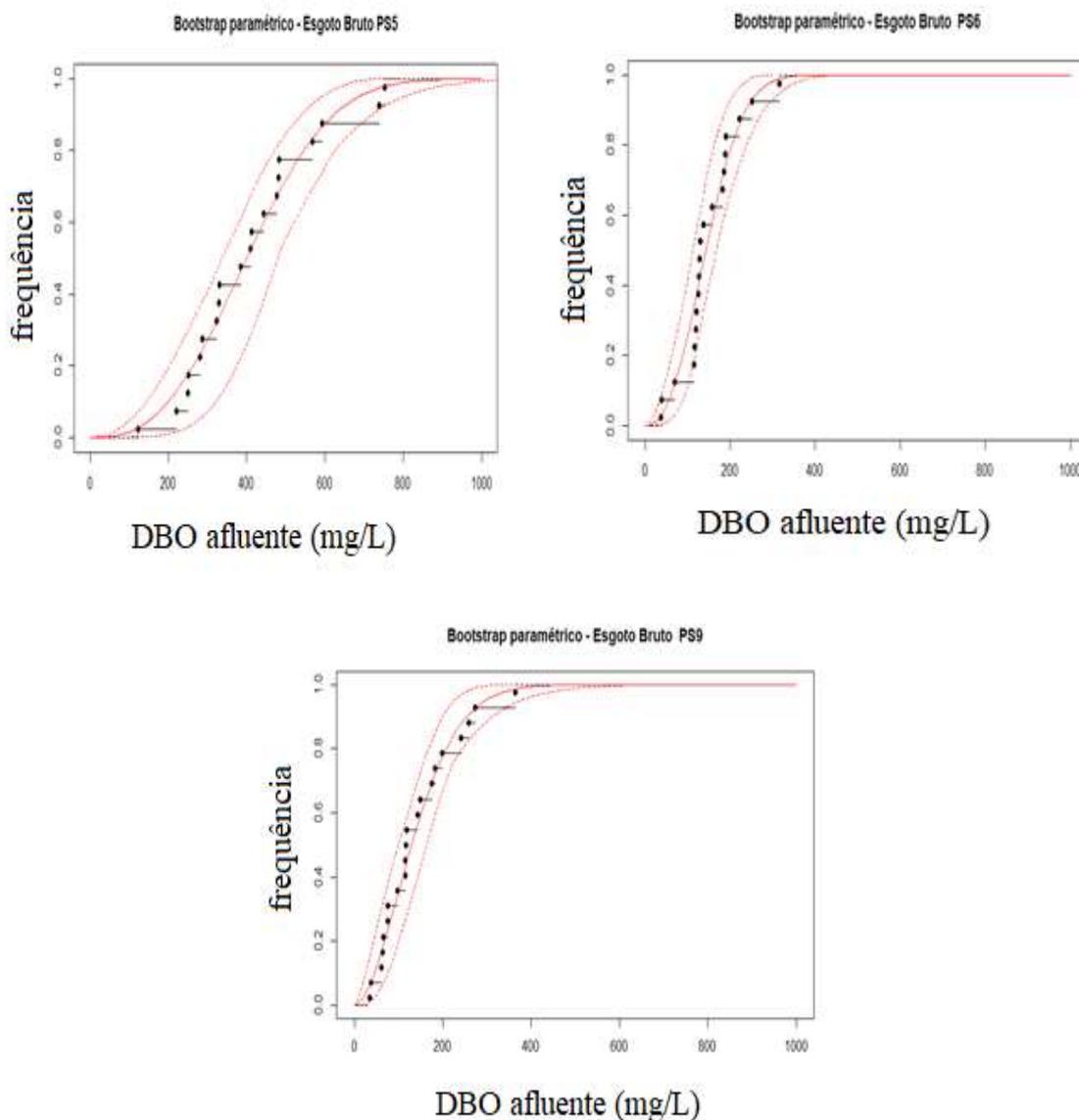
**Figura 32** – Pares de parâmetros Weibull levantados para cada reamostragem (a esquerda) e IC com as bandas de confiança com 95% das curvas.



Fonte: o autor, 2018.

Seguindo esta metodologia é realizado o levantamento de IC para o EB, LFP, LMP e LMS para as lagoas PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> utilizando a matéria orgânica dissolvida na forma de DBO. Para a metodologia paramétrica utilizando o procedimento descrito anteriormente, tem-se os seguintes IC para as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>. Onde a região entre os limites de confiança, mostrado em vermelho, contém 95% das curvas construídas. A Figura 33 mostra os IC da DBO afluente as lagoas PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>. Observa-se que a faixa de variação para a lagoa PS<sub>5</sub> é superior as demais, além de possuir limites de confiança mais extensos. O intervalo de análise para a DBO afluente as lagoas foi usado no intervalo de [0, 1000] mg/L.

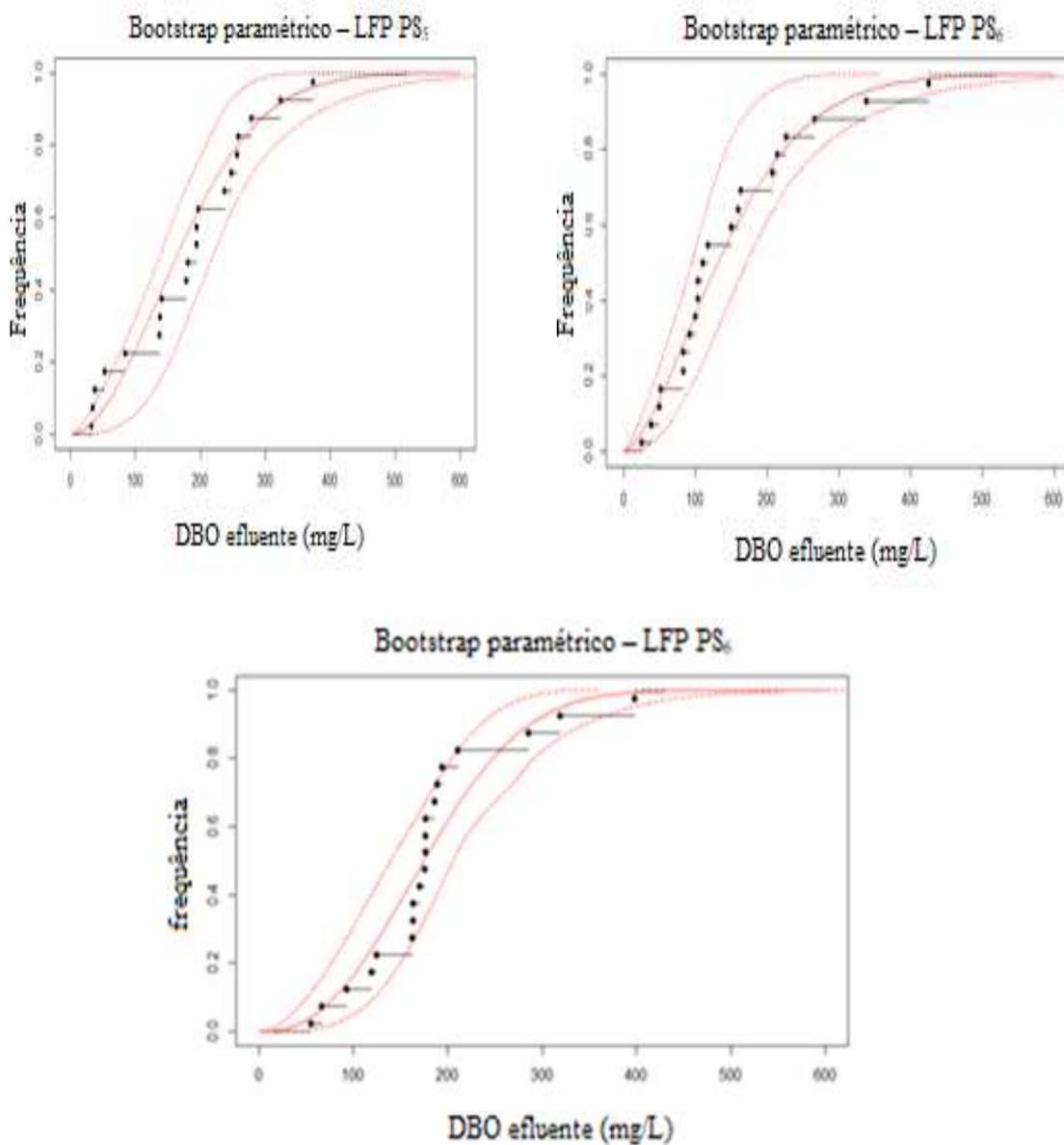
**Figura 33** – IC *Bootstrap* paramétrico para a DBO do EB afluente as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> (da esquerda para a direita).



Fonte: o autor, 2018.

Inicialmente o esgoto bruto é tratado pela LFP e a DBO efluente e os respectivos intervalo de confiança para os quartis (ou função distribuição acumulada) a estas lagoas é mostrado pela Figura 34. Note que cada lagoa possui um comportamento diferente no que tange as regiões nos quais a curva é mais larga ou estreita. Revelando que existem regiões nos quais os dados flutuam em faixas bem mais extensas, como representado pelo IC para o decis.

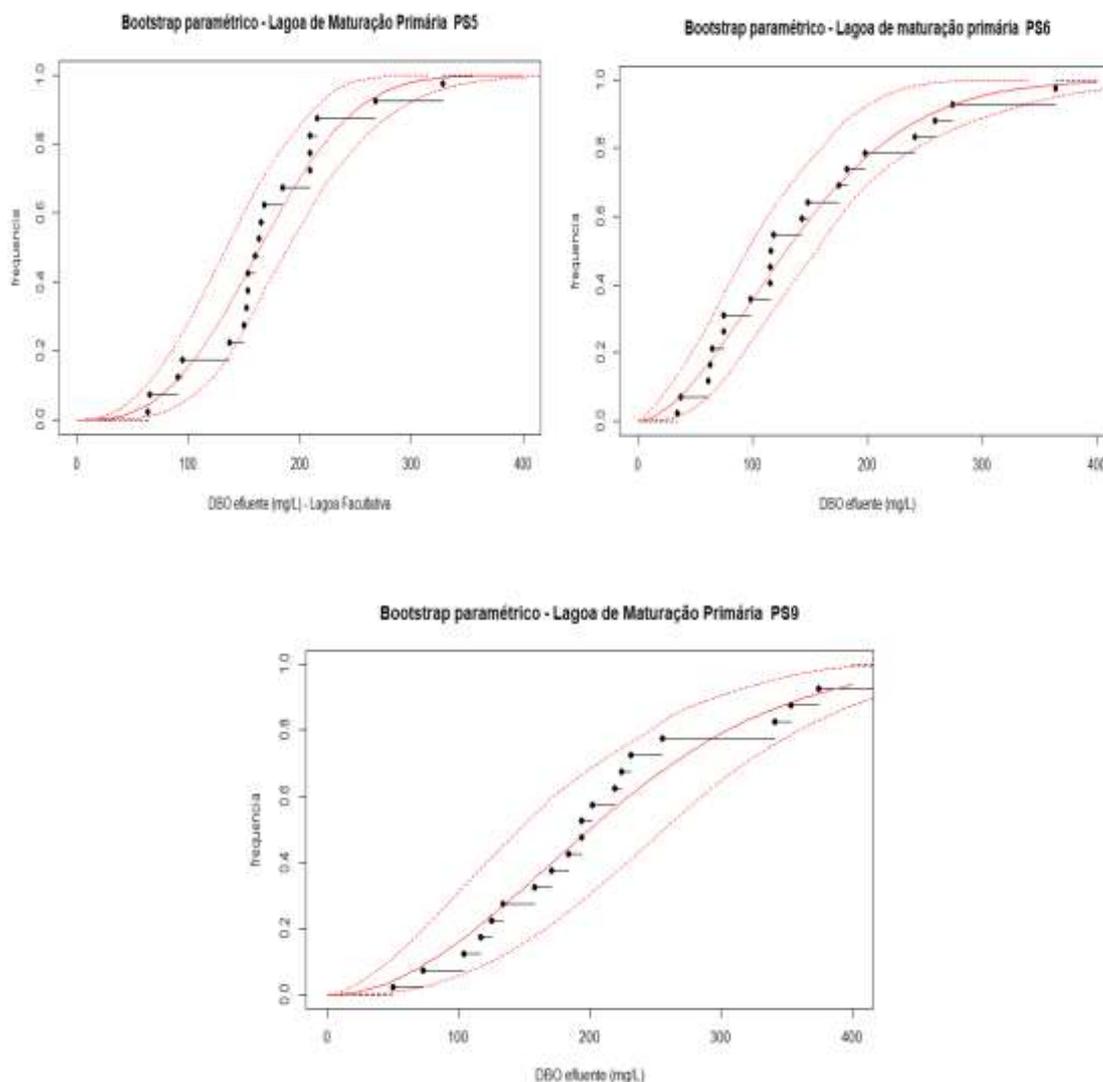
**Figura 34** – IC *Bootstrap* paramétrico para a DBO efluente do LFP nas ETE PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> (da esquerda para a direita).



Fonte: o autor, 2018.

O esgoto tratado nas LFP passa por um novo processo de tratamento na LMP. O mesmo procedimento é utilizado para a construção dos IC referentes ao DBO efluente a LMP conforme mostrado pela Figura 35.

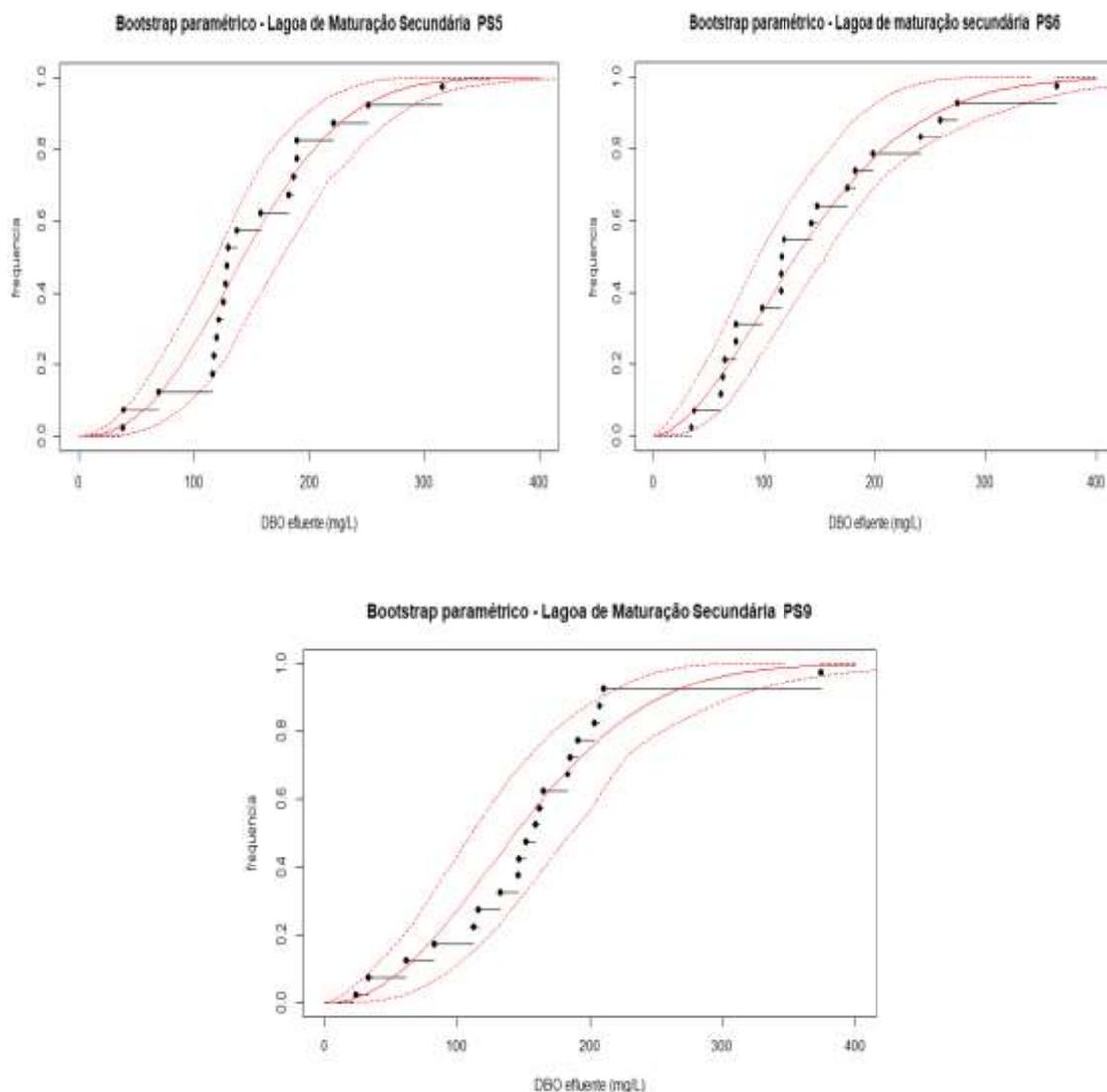
**Figura 35** – IC *Bootstrap* paramétrico para a DBO do LMP efluente as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> (da esquerda para a direita).



Fonte: o autor, 2018.

O efluente tratado da lagoa de maturação primária passa por uma nova etapa de tratamento antes de ser despejado no corpo receptor. O IC efluente da LMS são mostradas pela Figura 36. Através da análise da Figura 34, 35 e 36, nota-se que a medida que a meta torna-se menos exigente o número de amostras em conformação torna-se maior. Sempre que possível deseja-se trabalhar o mais próximo possível do limite superior dos IC, ou seja, o maior número possível de amostras dentro da especificação, isto conforme a capacidade que a estação oferece para remover poluentes do esgoto.

**Figura 36**– IC *Bootstrap* paramétrico para a DBO do efluente da LMS as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> (da esquerda para a direita).



Fonte: o autor, 2018.

Através da construção dos IC para os decis usando o *Bootstrap* paramétrico observa-se que estes possuem concordância com a metodologia não paramétrica, já que os resultados observados nas duas abordagens são muito próximos. Logo, o pesquisador pode escolher uma das rotas sem prejuízo a interpretação de dados. Este trabalho apresentou uma discussão mais detalhada dos fenômenos que ocorrem nas ETE devido a facilidade da sobreposição de figuras em uma mesma imagem.

### 4.2.3 Teste de aderência para os demais parâmetros de qualidade e garantia *Bootstrap*.

A metodologia proposta na seção 4.2.1 torna possível aferir a função que melhor adere a cada conjunto de dados. Neste tópico deseja-se mostrar a FDP que melhor adere aos principais parâmetros de qualidade de efluentes (DBO, DBO<sub>F</sub>, DQO, DQO<sub>F</sub>, AMT, OD e CTT). São estimadas as melhores funções de distribuição para o esgoto bruto, efluente à lagoa facultativa e de maturação primária e secundária para as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>. O critério para escolher a melhor função continua sendo o teste de *Kolmogorov-Smirnov* (*p-value*) e o Critério de *Akaike's*. Os resultados são mostradas no seguinte formato (distribuição predominante/ maior valor do *p-value* do teste KS) e (distribuição predominante/ menor valor do critério de *Akaike's*) conforme mostrado pela Tabela 14. Paralelamente é calculado o intervalo de garantia utilizando o *Bootstrap* paramétrico e a função distribuição de probabilidade encontrada anteriormente. A norma utilizada para o cálculo da garantia é mostrada pela Tabela 13. Por fim, a garantia para os parâmetros de qualidade utilizando os padrões mostrados pela Tabela 13 são mostrados pela Tabela 15.

**Tabela 13** - Valores de referência para avaliação do desempenho dos sistemas de lagoas

<i>Parâmetro</i>	<i>Efluente</i>	<i>Referência</i>
DBO	≤ 120 mg/L	COEMA 02/2017
DBO <sub>F</sub>	≤ 60 mg/L	COEMA 02/2017
DQO	≤ 200 mg/L	COEMA 02/2017
DQO <sub>F</sub>	≤ 200 mg/L	COEMA 02/2017
AMT	≤ 20 mg N/L	CONAMA 430/11
C <sub>TT</sub>	5,000 NMP/100 mL	COEMA 02/2017
pH	7,5 a 10,0	COEMA 02/2017
OD	≥ 3,0 mg/L	COEMA 02/2017

Fonte: o autor, 2018.

**Tabela 14** – Teste de aderência usando o teste de KS e Akaike para a ETE PS<sub>5</sub>.

<i>Parâmetro e Amostra</i>	<i>EB</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
DBO	(gu/0,9978)/ (gu/262,8822)	(n/0,9564)/ (w/242,2757)	(n/0,8046)/ (g/225,4715)	(n/0,6044)/ (g/227,7654)
DBO <sub>F</sub>	(g/0,9847)/ (g/247,8192)	(n/0,9544)/ (w/221,1589)	(gu/0,8616)/ (gu/206,9812)	(w/0,6630)/ (w/203,7665)
DQO	(n/0,9741)/ (g/285,1842)	(gu/0,9290)/ (ln/250,2859)	(w/0,7187)/ (w/249,7708)	(ln/0,8939)/ (w/287,3698)
DQO <sub>F</sub>	(ln/0,9829)/ (g/258,7775)	(w/0,7866)/ (gu/223,1110)	(ln/0,6674)/ (ln/217,6218)	(ln/0,7977)/ (ln/219,5053)
pH	(ln/0,7977)/ (ln/219,5138)	(gu/0,6340)/ (gu/15,1831)	(ln/0,4210)/ (gu/25,7670)	(gu/0,3213)/ (gu/31,3366)
AMT	(n/0,9379)/ (w/165,8398)	(n/0,4870)/ (n/132,7789)	(w/0,9034)/ (n/138,3977)	(gu/0,8489)/ (gu, 140,7966)
OD	-	(gu/0,8564)/ (g/ 91,2899)	(gu/0,9749)/ (w/ 89,4628)	(g/0,9652)/ (g/96,03164)
CTT	(w/0,1366)/ (w/729,2398)	(n/0,8523)/ (w/612,52)	(ln/0,7197)/ (ln/583,7762)	(w/0,7325)/ (ln/532,6295)

Fonte: o autor, 2018.

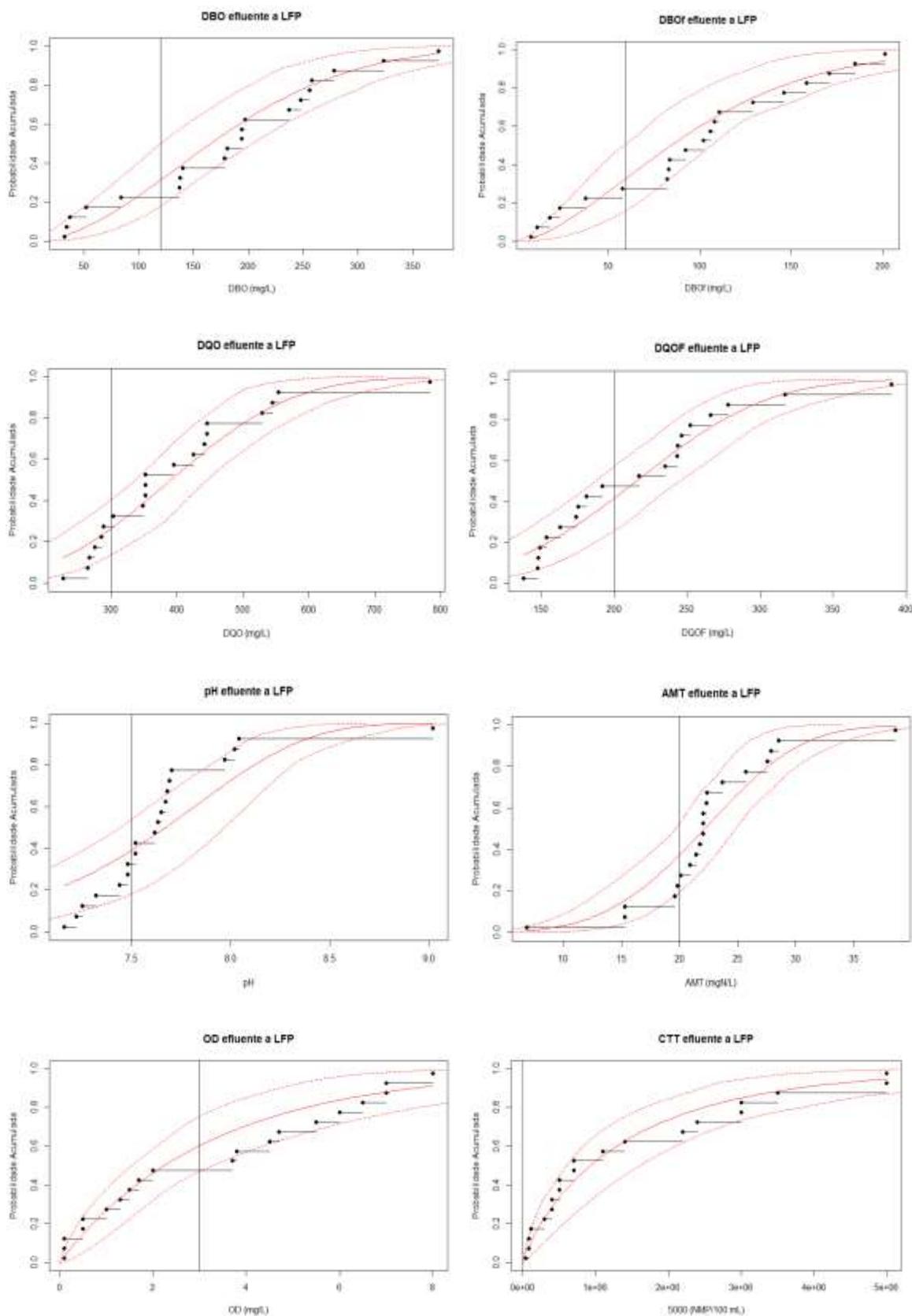
**Tabela 15** – Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>5</sub>.

<i>Parâmetro e Amostra</i>	<i>EB</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
DBO	[0,03;0,05]	[0,18;0,42]	[0,18;0,36]	[0,23;0,53]
DBO <sub>F</sub>	[0;0]	[0,22;0,48]	[0,04; 0,43]	[0,23; 0,53]
DQO	[0,0,08]	[0,18; 0,40]	[0,38; 0,65]	[0,68; 0,95]
DQO <sub>F</sub>	[0, 0,05]	[0,31, 0,65]	[0,75, 0,98]	[0,83; 0,97]
pH	[0,97,1]/[1,1]	[0,18, 0,55]/[1;1]	[0,2, 0,55].[1,1]	[0,15,0,4; 1,1]
AMT	[0; 0,08]	[0,25, 0,65]	[0,45, 0,76]	[0,45, 0,85]
OD	-	[0,45, 0,85]	[0,45, 0,79]	[0,4, 0,75]
CTT	[0,0]	[0,0]	[0,0]	[0,1]

Fonte: o autor, 2018.

Busca-se tornar o resultado obtido pela Tabela 15 mais didático é mostrado a representação gráfica para os intervalos de confiança para a garantia *Bootstrap* para o efluente proveniente da LFP. Optou-se por mostrar apenas estas lagoas facultativas para não tornar o trabalho demasiadamente extenso e repetitivo.

**Figura 37** – Função distribuição acumulada para o efluente proveniente da lagoas facultativa da PS<sub>5</sub> utilizando a metodologia *Bootstrap*.



Fonte: o autor, 2018.

De maneira análoga a apresentada anteriormente é apresentado o teste de aderência para os parâmetros de qualidade da estação PS<sub>6</sub> e os intervalos de confiança para a garantia com utilizando a metodologia *Bootstrap*.

**Tabela 16** – Teste de aderência usando o teste de KS e Akaike's para a ETE PS<sub>6</sub>.

Parâmetro	EB	LFP	LMP	LMS
DBO	(ln/0,3293)/ (ln/270,5196)	(ln/0,9059)/ (g/248,8517)	(g/0,9692)/ (g/244,1043)	(gu/0,8299)/ (g/243,2416)
DBO <sub>F</sub>	(gu/0,8156)/ (gu/212,5283)	(ln/0,9720)/ (ln/225,1494)	(w/0,9544)/ (w/215,3780)	(w/0,7822)/ (w/221,7392)
DQO	(ln/0,9781)/ (gu/281,4716)	(ln/0,8600)/ (ln/259,6499)	(g/0,8629)/ (g/253,2813)	(ln/0,6246)/ (ln/264,8922)
DQO <sub>F</sub>	(ln/0,8271)/ (ln/236,8597)	(g/0,9905)/ (ln/242,4352)	(ln/0,9973)/ (ln/ 211,13)	(g/0,2213)/ (g/232,4569)
pH	(gu/0,6358)/ (ln/17,0127)	(gu/0,5777)/ (gu/13,6789)	(gu/0,4281)/ (gu/0,22075)	(g/0,6141)/ (g/12,6123)
AMT	(gu/0,9683)/ (n/175,4855)	(w/0,8880)/ (w/125,2990)	(w/0,9996)/ (w/124,3301)	(n/0,9254)/ (w/134,6433)
OD	-	(w/0,8255)/ (w/76,4311)	(n/0,8608)/ (w/86,0816)	(n/0,9172)/ (w/ 213,3698)
CTT	(w/0,3047)/ (ln/773,0973)	(ln/0,8931)/ (ln/629,8360)	(w/0,4555)/ (ln/577,5001)	(ln/0,3285)/ (w/520,7322)

**Fonte:** o autor, 2018.

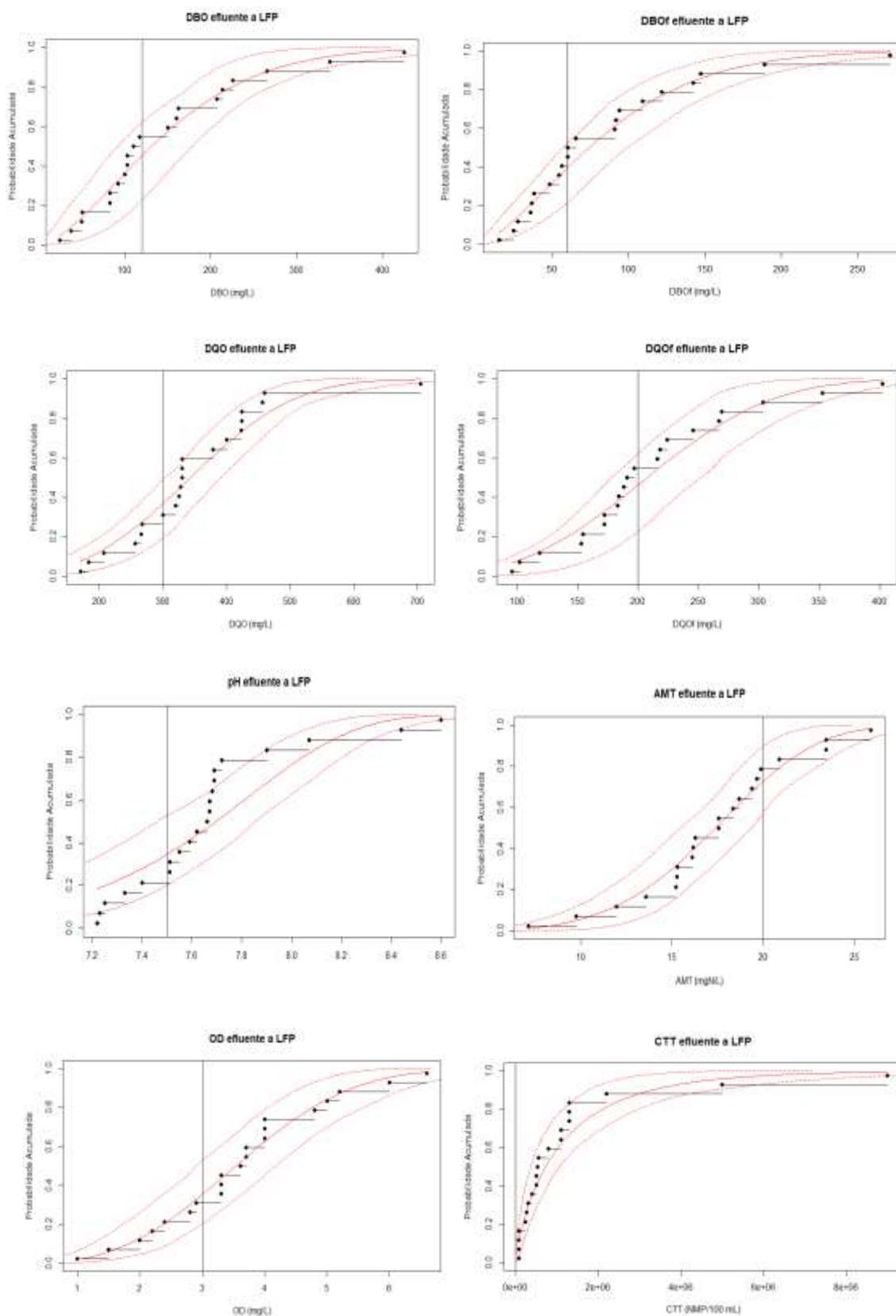
**Tabela 17** – Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>6</sub>.

Parâmetro	EB	LFP	LMP	LMS
DBO	[0,03; 0,17]	[0,25; 0,64]	[0,3; 0,63]	[0,3; 0,67]
DBO <sub>F</sub>	[0,03; 0,1]	[0,23; 0,57]	[0,3; 0,63]	[0,37; 0,57]
DQO	[0,02; 0,07]	[0,25; 0,50]	[0,25; 0,63]	[0,53; 0,85]
DQO <sub>F</sub>	[0,03; 0,1]	[0,3; 0,65]	[0,9; 0,97]	[0,73; 0,97]
pH	[0,43; 0,82]/[1,1]	[0,1/ 0,57]/[1,1]	[0,25; 0,53]/[1,1]	[0,15/0,37]/[1,1]
AMT	[0,1; 0,37]	[0,6; 0,9]	[0,7; 0,9]	[0,75/0,95]
OD	-	[0,15; 0,45]	[0,23; 0,5]	[0,2; 0,55]
CTT	[0,0]	[0,0]	[0,0]	[0,0]

Fonte: o autor

Identificada as distribuições para cada conjunto de dados é mostrado o gráfico com as funções distribuição acumulada utilizando a metodologia *Bootstrap* para o efluente as lagoas facultativas para a estação PS<sub>6</sub>.

**Figura 38** – Função distribuição acumulada para o efluente proveniente da lagoas facultativa para a estação PS<sub>6</sub> utilizando a metodologia *Bootstrap*.



Fonte: o autor, 2018.

Concluído o mesmo procedimento realizado anteriormente para a estação PS<sub>9</sub>, tem-se as funções distribuição de probabilidade, a garantia *Bootstrap* e a representação gráfica para as funções distribuição de probabilidade para o efluente a lagoa facultativa. A Tabela 18 mostra o resultado dos testes de aderência e a Tabela 19 os intervalos de confiança para a garantia do tratamento.

**Tabela 18** – Teste de aderência usando o teste de KS e Akaike para a ETE PS<sub>9</sub>.

Parâmetro	EB	LFP	LMP	LMS
DBO	(G/0,9652)/ (Gu/270,4895)	(w/0,3481)/ (g/232,5431)	(ln/0,9880)/ (g/245,4882)	(w/0,6244)/ (w/231,8462)
DBO <sub>F</sub>	(Gu/0,9987)/ (Gu/237,1699)	(n/0,8606)/ (w/209,8137)	(n/0,9820)/ (w/221,1913)	(n/0,8976)/ (g/208,2259)
DQO	(N/0,9689)/ (G/278,9725)	(n/0,9032)/ (g/252,7365)	(gu/0,9796)/ (gu/249,32)	(gu/0,9816)/ (g/234,96)
DQO <sub>F</sub>	(G/0,9000)/ (G/257,1493)	(w/0,8960)/ (w/229,3265)	(n/0,9093)/ (w/217,1297)	(n/0,9999)/ (w/207,8402)
pH	(Ln/0,8554)/ (N/4,3548)	(w/0,7756)/ (g/6,3045)	(w/0,4698)/ (g/8,0698)	(w/0,6986)/ (w/12,5478)
AMT	(Ln/0,7206)/ (Ln/174,3837)	(w/0,7112)/ (w/124,7112)	(w/0,6497)/ (w/0,1697)	(w/0,5488)/ (w/125,1721)
OD	-	(g/0,8745)/ (g/83,9145)	(n/0,7492)/ (w/87,3698)	(n/0,8719)/ (w/94,6090)
CTT	(Ln/0,8773)/ (Ln/767,9396)	(ln/0,09578)/ (ln/590,0553)	(w/0,9791)/ (w/548,2598)	(ln/0,7057)/ (ln/ 498,5954)

Fonte: o autor, 2018.

**Tabela 19** – Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>9</sub>.

Parâmetro	EB	LFP	LMP	LMS
DBO	[0; 0,03]	[0,1; 0,38]	[0,15; 0,37]	[0,22; 0,6]
DBO <sub>F</sub>	[0; 0,05]	[0,1; 0,45]	[[0,1; 0,4]	[0,23; 0,57]
DQO	[0; 0,1]	[0,27; 0,6]	[0,37; 0,8]	[0,75; 0,97]
DQO <sub>F</sub>	[0,03; 0,2]	[0,5; 0,75]	[0,83; 0,98]	[1,1]
pH	[0,3;0,78]; [1;1]	[0,37; 0,65]; [1;1]	[0,07; 0,37]; [1;1]	[0,07; 0,37]; [1;1]
AMT	[0,05; 0,28]	[0,13; 0,37]	[0,05; 0,3]	[0,3; 0,65]
OD	-	[0,5; 0,85]	[0,3; 0,7]	[0,28; 0,73]
CTT	[0,0]	[0,0]	[0,01]	[0; 0,0]

Fonte: o autor, 2018.

Para facilitar a exposição dos resultados referente as FPD com melhor aderência com os parâmetros de qualidade (DBO, DBO<sub>F</sub>, DQO, DQO<sub>F</sub>, AMT, pH, OD, CTT) para as estações PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub> provenientes das lagoas LFP, LMP e LMS. Para o critério de KS, mostrado pela Tabela 20, e o critério de AIC, mostrado pela Tabela 21.

**Tabela 20** – Sumário das FDP encontradas com melhor aderência pelo teste de KS.

Parâmetro	LFP			LMP			LMS		
	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>
DBO	N	Ln	W	N	G	Ln	N	Gu	W
DBO <sub>F</sub>	W	Ln	N	Gu	W	N	W	W	N
DQO	W	Ln	N	Gu	G	Gu	W	Ln	Gu
DQO <sub>F</sub>	Ln	G	W	W	Ln	N	W	Gu	N
pH	Gu	Gu	W	Ln	Gu	W	Gu	Gu	W
AMT	N	W	W	W	W	W	Gu	N	W
OD	Gu	W	G	Gu	N	N	G	N	N
CTT	N	Ln	Ln	Ln	W	N	W	Ln	W

Fonte: o autor, 2018.

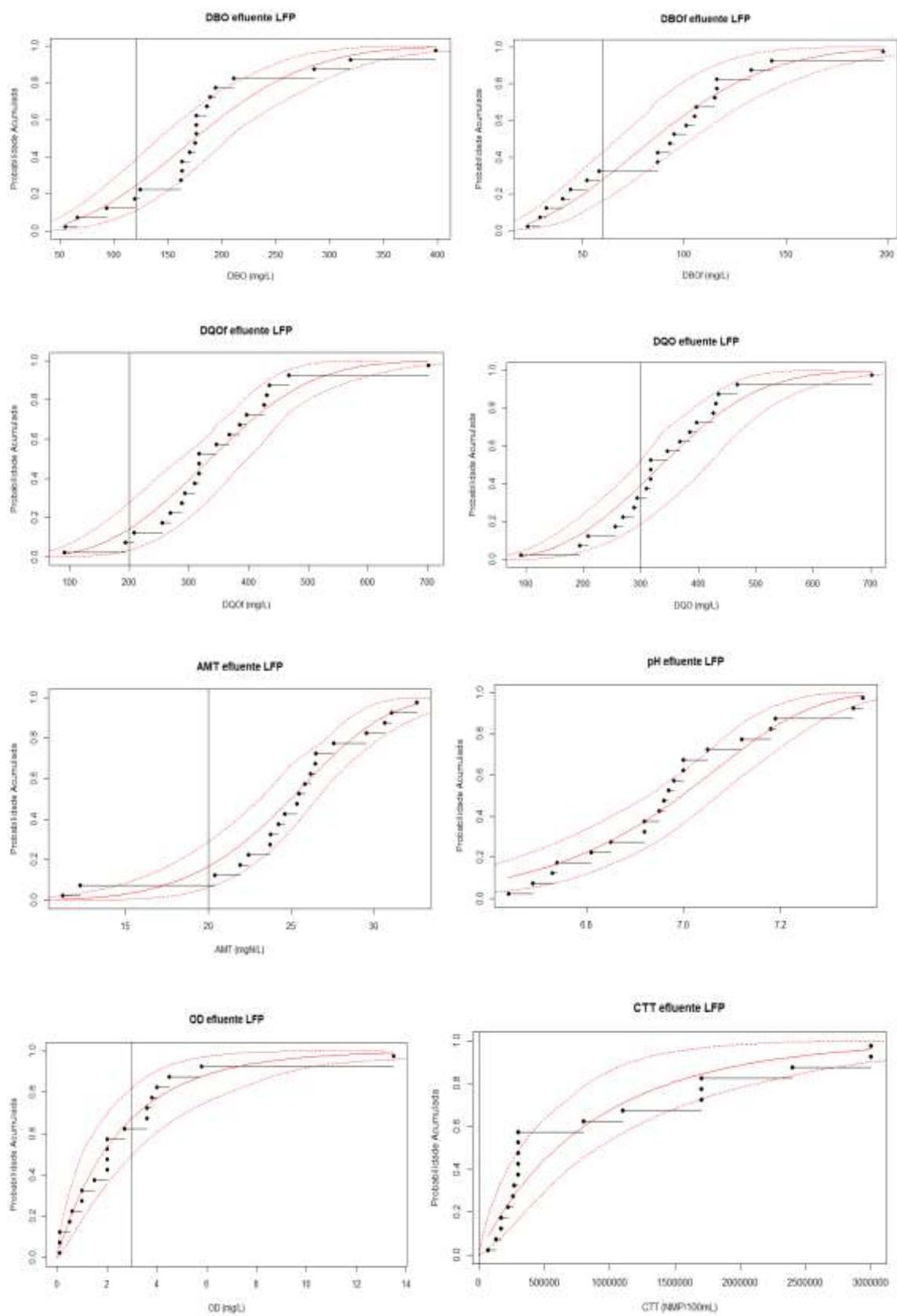
**Tabela 21** – Sumário das FDP selecionadas a partir do teste de AIC.

Parâmetro	LFP			LMP			LMS		
	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>	PS <sub>5</sub>	PS <sub>6</sub>	PS <sub>9</sub>
DBO	W	G	G	G	G	G	G	G	W
DBO <sub>F</sub>	W	Ln	W	G	W	W	W	W	G
DQO	Ln	Ln	G	W	G	Gu	W	Ln	G
DQO <sub>F</sub>	Gu	Ln	W	Ln	Ln	W	Gu	G	W
pH	Gu	Gu	G	Gu	Gu	G	Gu	Gu	W
AMT	N	W	W	N	W	W	Gu	W	W
OD	G	W	G	W	W	W	G	W	W
CTT	Ln	Ln	Ln	Ln	Ln	W	Ln	Ln	Ln

Fonte: o autor, 2018.

De maneira análoga as estações PS<sub>5</sub> e PS<sub>6</sub> são construídas as representações gráficas para a função distribuição de probabilidade para o efluente proveniente a lagoa facultativa. A Figura 39 mostra a representação gráfica para a garantia do tratamento para as lagoas facultativas provenientes da estação PS<sub>9</sub>. Através desta análise sistemática é possível identificar quais são os parâmetros que oferecem a menor garantia e desta maneira investir recursos para mitigá-los os prejuízos de não conformidade.

**Figura 39** – Função distribuição acumulada para o efluente proveniente da lagoas facultativa para a estação PS<sub>9</sub> utilizando a metodologia *Bootstrap*.



**Fonte:** o autor, 2018.

### 4.3 Confiabilidade do tratamento com metodologia *Fuzzy*

Para o cálculo dos intervalos de confiança para os decis em conformidade com a meta estipulada pela legislação ambiental para a DBO efluente a estação PS<sub>6</sub> (por exemplo), quanto mais próximo o valor estiver do limite superior do intervalo de confiança mais próximo este sistema encontra-se de uma situação “ótima”. Tendo isto em mente, é necessário estabelecer uma metodologia para calcular a conformidade do processo através do grau de distanciamento ou proximidade dos limites de confiança. Busca-se introduzir uma metodologia que forneça o grau de distanciamento ou proximidade dos limites de confiança de um sistema optou-se pela utilização simultânea do método *Bootstrap* com Número *Fuzzy* Triangular (NFT).

A Tabela 22 mostra esta relação com os pontos geométricos do NFT e os IC *Bootstrap*. A união da metodologia *Bootstrap*, permite estimar os intervalos de confiança para uma função distribuição de probabilidade acumulada, e o risco associado ao número *Fuzzy* triangular consegue estimar o grau de afastamento de um cenário otimista. Para isto é utilizada a função distribuição acumulada de probabilidade retirada das amostras para analisar o quão distante este cenário está de uma condição “ótima”.

É lembrado que para o exemplo mostrado anteriormente na seção 4.2.2 que analisa a conformidade da concentração do efluente conforme a legislação ambiental pertinente o cenário ótimo é aquele próximo do limite superior de confiança do quartil (elevado número de amostras com valores inferiores a 120 mg/L para a matéria orgânica dissolvida na forma de DBO). De sorte que quanto mais próximo do limite superior melhor é o cenário e recorrentemente menor é o risco *Fuzzy*, conforme mostrado pela equação 21. Logo, sempre é preferível operar próximo do limite superior, indicando que a estação está sendo operada nas melhores condições possíveis.

**Tabela 22** – Risco *Fuzzy* associado aos limites de Confiança *Bootstrap*.

<i>Bootstrap</i>	<i>Fuzzy</i>	<i>NFT</i>
Limite Inferior (2,5%)	Mínimo	a <sub>1</sub>
Valor médio	Número Mais Provável	a <sub>2</sub>
Limite Superior (97,5%)	Máximo	a <sub>3</sub>

Fonte: o autor, 2018.

A metodologia proposta sugere construir um número NFT triangular a partir do limite inferior, valor médio e do limite superior. E com uma nova sequência de coletas em mãos é analisado o risco desta nova coleta estar próxima do cenário ótimo (limite superior do IC) ou do cenário péssimo (limite inferior do IC), isto para a remoção de matéria orgânica em forma de DBO. Logo, tem-se o interesse de analisar se a estação está operando próximo de sua situação ótima, para cada decil do tratamento. De sorte que quanto mais próxima a nova coleta estiver do limite superior, menor o risco *Fuzzy* associado, e mais próximo de uma situação “ideal” a ETE está operando.

Para a ETE PS<sub>6</sub>, na lagoa facultativa especificamente, é realizada uma nova carreira de coletas e com estes dados são calculados os quartis para cada 10% da função probabilidade acumulada. A Tabela 23 mostra os limites *Bootstrap* inferiores, superiores e o número mais provável (médio) e os valores da função distribuição acumulada para os valores obtidos para uma nova coleta de amostras.

**Tabela 23** – Quantis dos valores da DBO efluente e seus respectivos valores de IC.

Decil (%)	Média	2,5%	97,50%	Reamostragem
10	42	27	65	80
20	65	45	94	93
30	84	62	119	108
40	106	82	145	118
50	129	99	172	153
60	152	113	201	184
70	183	129	238	203
80	218	151	283	224
90	269	180	343	242

Fonte: autor, 2018.

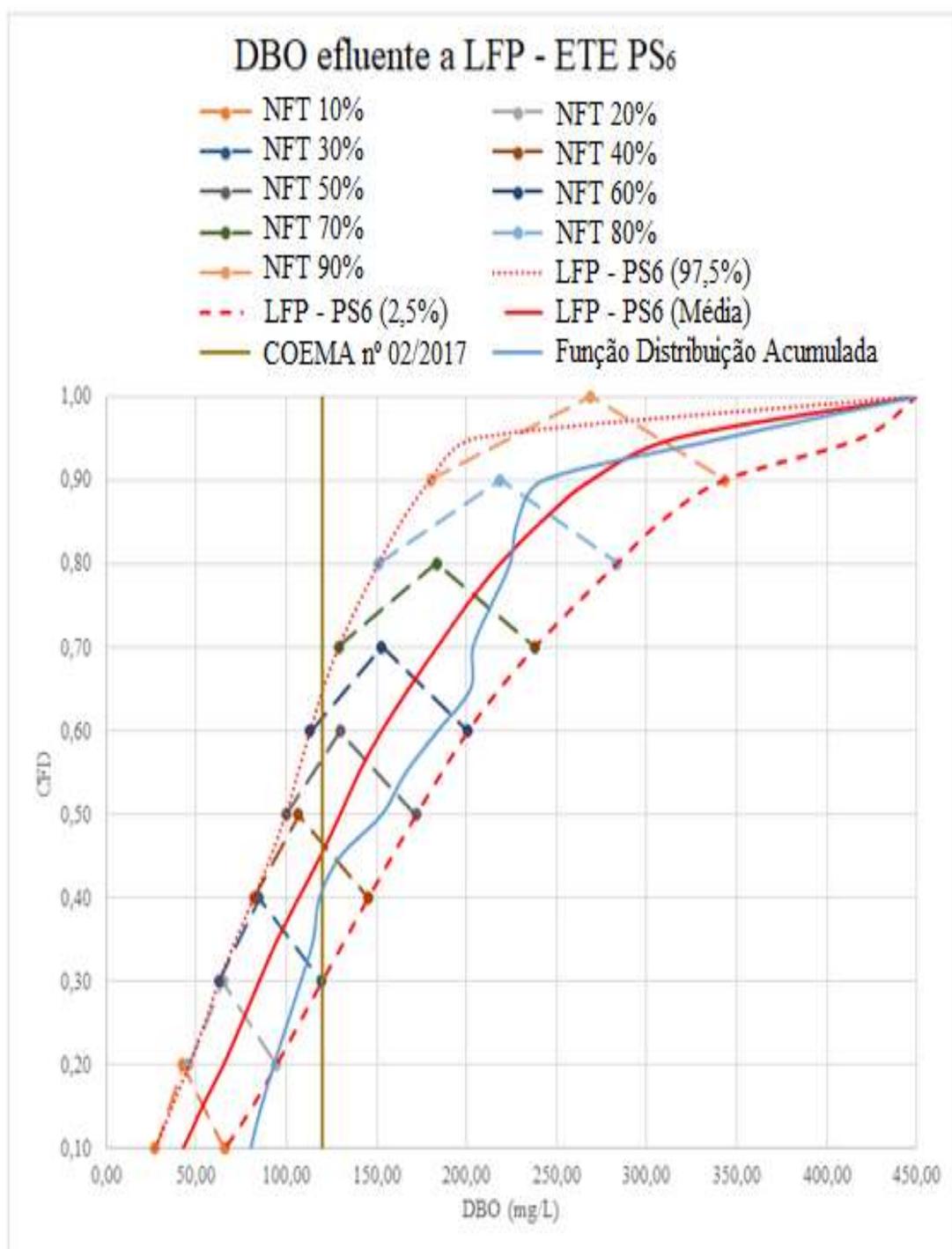
Para os dados mostrados pela Tabela 23 é construída a Figura 40 que explicita os Número *Fuzzy* Triangular (NFT) para os decis de 10, 20, 30, 40, 50, 60, 70, 80 e 90%. A curva vermelha contínua representa o valor médio entre os limites de confiança inferior e superior. A curva azul contínua representa a função distribuição acumulada das amostras. Observa-se de início que a curva azul está quase inteiramente dentro dos limites de confiança, e que o comportamento desta é irregular, hora aproximando-se do limite inferior (alto risco) e hora aproximando-se do limite superior (baixo risco).

Adotado um cenário satisfatório como aquele que aproxima-se do limite superior, ou seja, maior número possível de amostras com o valor de DBO inferior à meta estipulada. O pior cenário é aquele que as amostras aproximam-se do limite inferior, ou seja, poucas amostras com valor de DBO inferior a meta estipulada.

Esta nova metodologia apresentada busca mensurar o “quão próximo de um cenário otimista uma dada estação está sendo operada”. Para isto utiliza-se o conceito de risco *Fuzzy* associado a um número *Fuzzy* triangular, tendo em mente que um evento não desejável, ou seja, que traga efeito prejudicial para o tratamento é quando a estação opera próximo ao quartil de 2,5%. Logo, o distanciamento ou a proximidade de uma situação ótima é um parâmetro que mostra a qualidade do tratamento de esgoto dentro da garantia que esta pode vir a oferecer. A idéia central é trabalhar com objetivos “tangíveis” sabendo que o quartil superior é “ótimo” para uma determinada estação de tratamento, isto para a remoção de matéria orgânica na forma de DBO. Para os diferentes decis calculou-se o risco de ocorrência de pior cenário com base em número *Fuzzy* triangular integrado aos decis da FDA. Logo, para a FDA de cada função de cada parâmetro as menores de distâncias entre o limite inferior e superior determinam o maior risco difuso associado para cada decil.

Através dos NFT mostrados pela Figura 40 é calculado o risco *Fuzzy* associado a não conformidade no IC superior. Ou seja, é preferível que uma reamostragem esteja próximo ao limite superior (97,5%). A lagoa facultativa da ETE PS<sub>6</sub> tem um comportamento interessante. Para uma nova coleta de dados representada pela curva azul é observado que para os primeiros quantis (10, 20 e 30%) o risco de não conformidade é elevado. O risco cai um pouco para o quartil de 40% e após volta a elevar novamente (50,60 e 70%) e finalmente o decil cai acentuadamente até chegar o quartil de 90% (risco *Fuzzy* de 27 %). Esta análise mostra que existem intervalos nos quais a ETE consegue operar com mais eficácia. Logo, o responsável pela ETE deve se precaver e a partir dos dados mostrados pelo risco *Fuzzy* para cada decil procurar operar em regiões com menor risco ou tomar medidas paliativas de modo a diminuir o risco em locais específicos. Entretanto, a maior utilidade do cálculo do risco *Fuzzy* é o distanciamento da operação de uma ETE de uma condição ótima. Através desta análise o gestor pode analisar a necessidade de intervenção na ETE, já que existe a possibilidade de comparar com o distanciamento do intervalo de confiança mostrando se a ETE está operando próxima da condição ótima ou não. A Tabela 24 mostra a área anterior e posterior de concordância para o risco *Fuzzy* associado a cada decil conforme representado pela Figura 14.

**Figura 40** – NFT associado aos intervalos de confiança.



Fonte: o autor, 2018.

**Tabela 24** – Risco *Fuzzy* associado aos decis para a matéria orgânica na forma de DBO

Decil (%)	Área Anterior (Concordância)	Área Posterior (Concordância)	Área total (und.)	Risco <i>Fuzzy</i>
10	0 (0%)	19,41 (100%)	19,41	1
20	0,05 (0,18%)	24,27 (99,82%)	24,32	0,99
30	2,03 (7,05%)	26,63 (92,95%)	28,66	0,92
40	9,51 (29,92%)	22,26 (70,08%)	31,77	0,70
50	4,37 (12,07%)	31,86 (87,93%)	36,23	0,87
60	3,01 (6,87%)	40,81 (93,13%)	43,82	0,93
70	10,89 (19,99%)	43,59 (80,01%)	54,48	0,79
80	27,17 (41,04%)	39,02 (58,96%)	66,19	0,58
90	59,62 (73,06%)	21,98 (26,94%)	81,56	0,27

Fonte: o autor, 2018.

Através da observação dos dados mostrados pela Tabela 24 percebe-se que o risco de não conformidade com um cenário otimista é variável dependendo da região no qual deseja-se operar a ETE. Por exemplo, a estação mostrada neste exemplo oferece risco menor de não conformidade em valores nos quais a DBO aproxima-se do decil de 90%. Ou seja, dependendo da faixa de exigência que o gestor almeja operar, existe um risco relacionado a este intervalo. Para este caso em particular em estudo observa-se que os quartis de 80 e 90% são os que oferecem o menor risco.

A Tabela 24 fornece um panorama de como a remoção de poluentes traduzida em porcentagem de amostras com valores de DBO inferiores a um padrão se comporta em diversas faixas de concentração efluentes. Por exemplo, observa-se que o risco é elevado nos primeiros quantis comportando-se de maneira anômala até reduzir-se a 27% para o último quartil da análise. É interessante observar que em cada região (decil) no qual deseja-se trabalhar existe um risco *Fuzzy*. Logo, existem regiões no qual é preferível operar oferecendo uma garantia maior para o operador.

#### 4.3.1 – Confiabilidade do tratamento com metodologia *Fuzzy* – demais parâmetros

De maneira análoga a aquela apresentada na seção 4.3 deste trabalho tem-se o interesse de calcular o risco *Fuzzy* para as estações de tratamento de esgoto para os demais parâmetros qualidade. De modo a analisar a não conformidade (risco *Fuzzy*) relacionado ao distanciamento dos limites de confiança construídos a partir da metodologia *Bootstrap* para cada decil. A Figura 41 apresenta a representação gráfica do risco *Fuzzy* para a DBO, DQOf, AMT e CTT efluente a lagoa facultativa da ETE PS<sub>5</sub>. Para o trabalho não tornar-se

demasiadamente extenso e cansativo os valores de risco foram representados pela Tabela 25.

Através da construção do número *Fuzzy* triangular entre os decis de maneira similar ao mostrado pela Figura 41 é possível calcular o distanciamento ou proximidade de um cenário ótimo. Através desta metodologia é calculado o risco *Fuzzy* para os parâmetros de qualidade DBO<sub>F</sub>, DQO<sub>F</sub>, AMT e CTT afluente a lagoa facultativa, de maturação primária e secundária para a estação PS<sub>5</sub>.

De maneira análoga a metodologia mostrada para o cálculo dos IC na seção 4.3 foram calculados os limites inferiores (2,5%), superiores (97,5%) e médio. De posse destes valores foi calculado o risco *Fuzzy* para cada decil com o valor dos quartis obtidos para a carreira de coletas e desta maneira foi aferido o grau de distanciamento (risco) de um cenário excelente (limite superior de confiança). Desta maneira é analisado o risco *Fuzzy* utilizando o número *Fuzzy* triangular para cada decil, objetivando aferir em qual “regiões” a estação de tratamento oferece maior garantia para o operador.

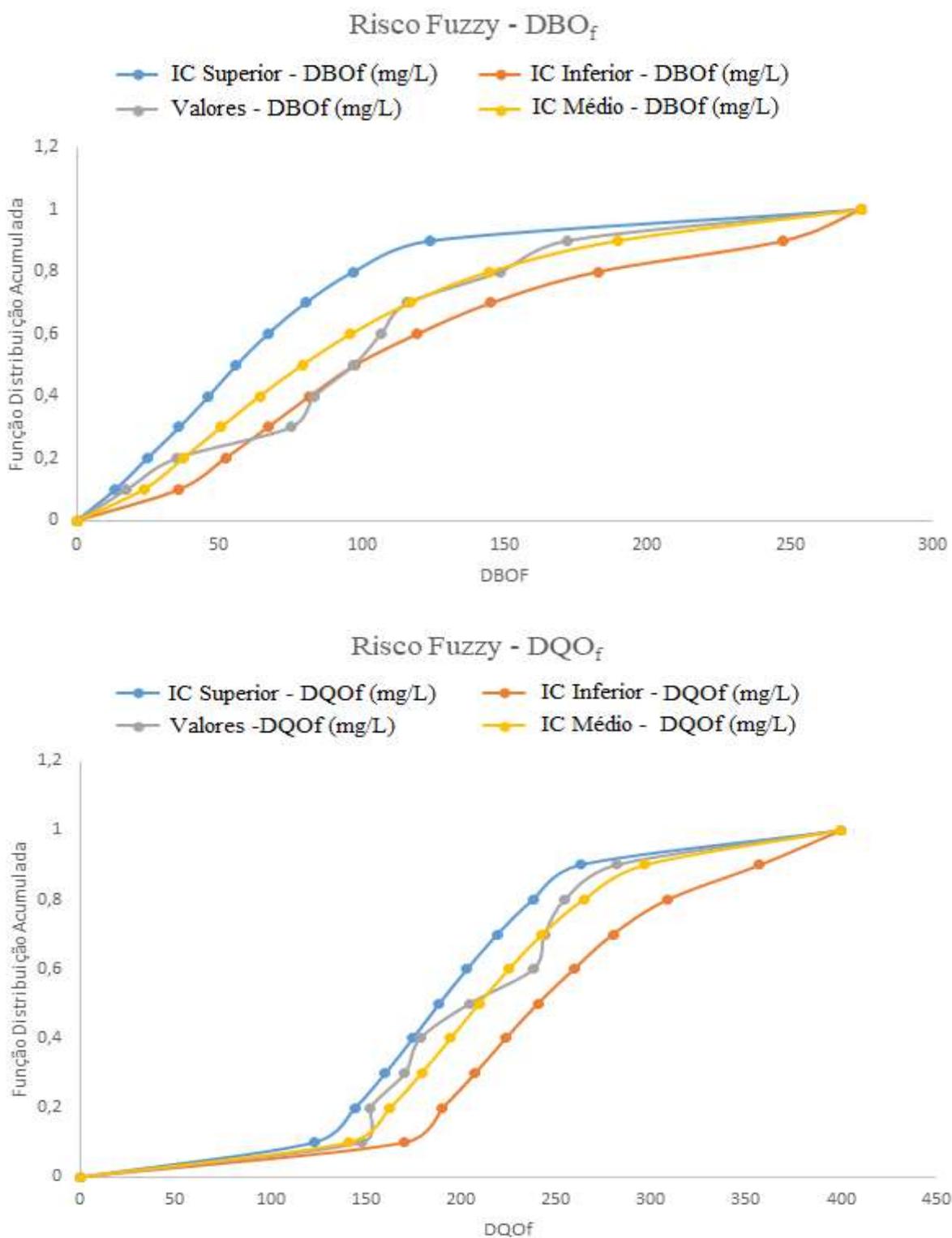
A Figura 41 mostra a representação gráfica para os limites de confiança utilizando a formulação proposta pela Tabela 13. A Tabela 23 mostra o risco associado a cada decil para o DQO<sub>F</sub>, DQO<sub>F</sub>, AMT e CTT a lagoa facultativa pertencente a estação PS<sub>5</sub>.

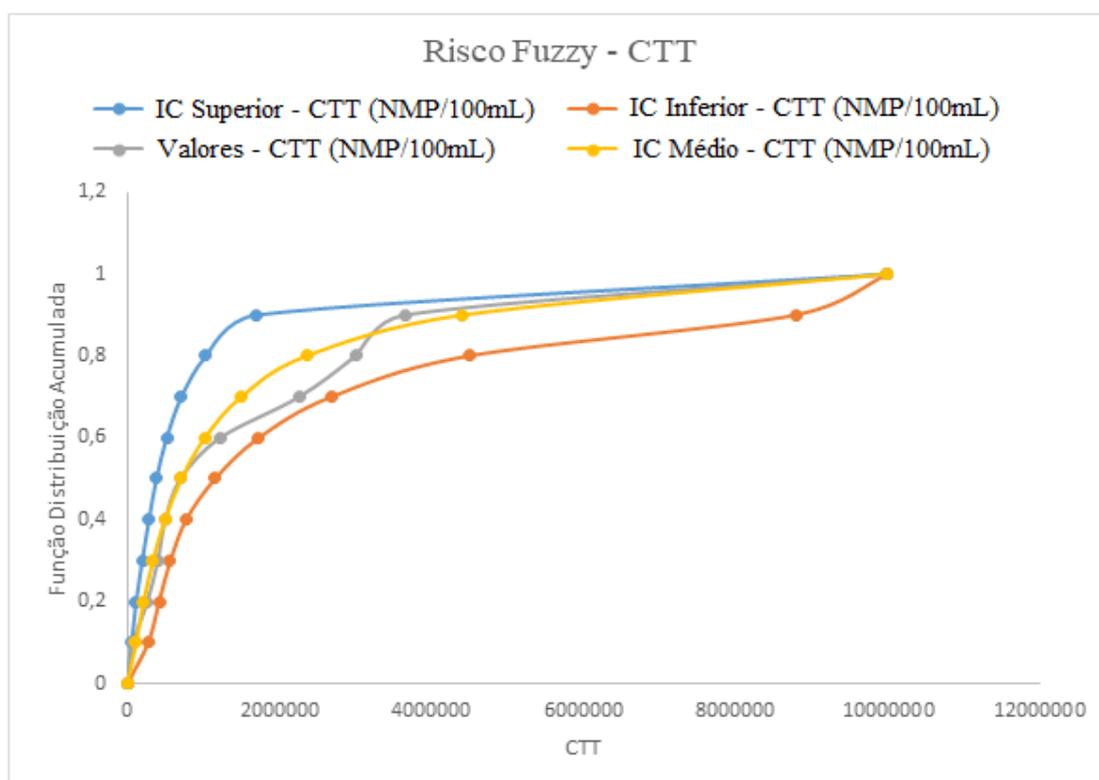
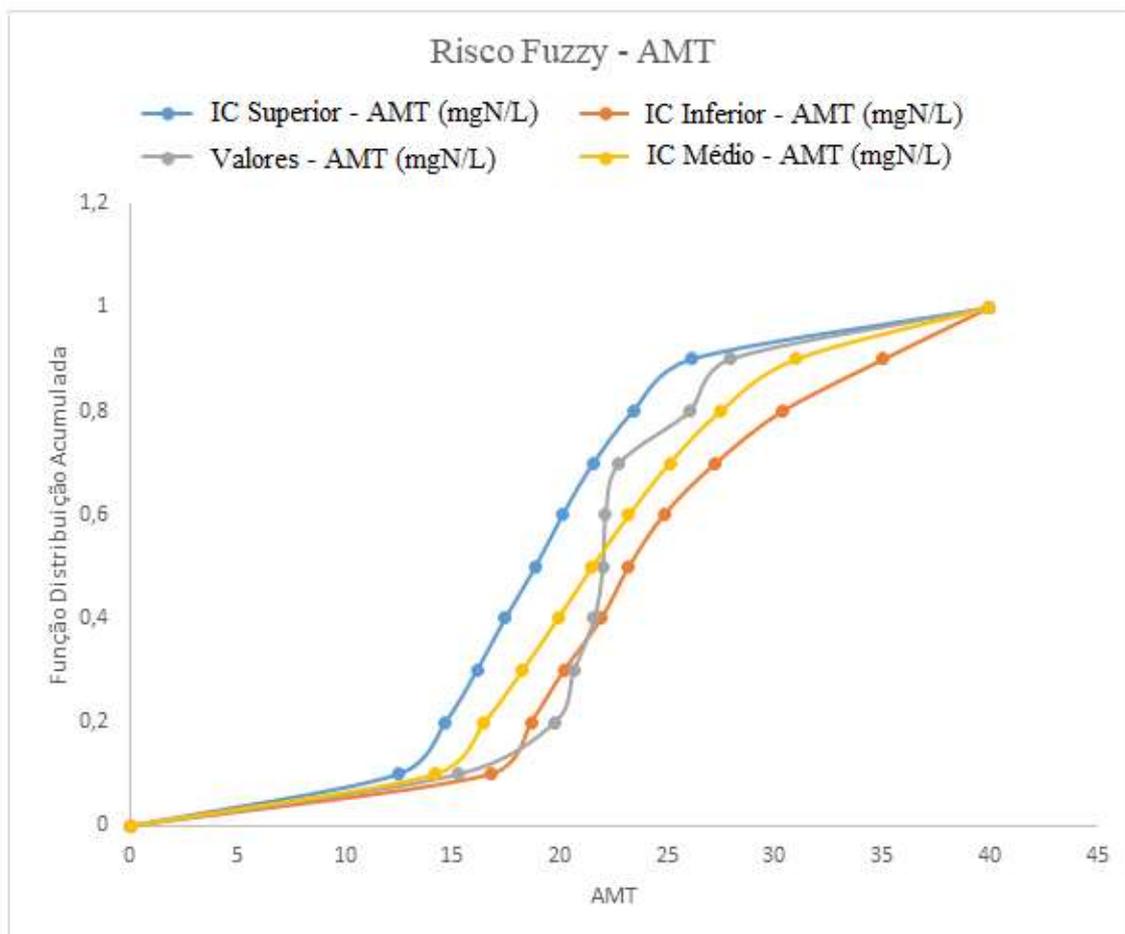
**Tabela 25** – Risco *Fuzzy* para o efluente proveniente da LFP da ETE PS<sub>5</sub>.

<b>Decil (%)</b>	<b>DBO<sub>F</sub></b>	<b>DQO<sub>F</sub></b>	<b>AMT</b>	<b>CTT</b>
10	0,982	0,038	0,180	0,974
20	0,895	0,618	0,010	0,843
30	0,281	0,492	0,105	0,803
40	0,1652	0,934	0,273	0,500
50	0,054	0,784	0,661	0,424
,60	0,180	0,348	0,977	0,670
70	0,034	0,712	0,987	0,865
80	0,056	0,890	0,819	0,620
90	0,158	0,872	0,9082	0,168

Fonte: o autor, 2018.

**Figura 41** – Representação do risco *Fuzzy* utilizando a metodologia *Bootstrap* para a construção dos intervalos de confiança para os parâmetros DBO<sub>f</sub>, DQO<sub>f</sub>, AMT e CTT para a lagoa facultativa afluente a estação PS<sub>5</sub>.





Fonte: o autor, 2018.

## 4.4 Controle estatístico de processos

### 4.4.1 Função distribuição de probabilidade (FDP) e teste de aderência

Adotadas as funções distribuição de probabilidade escolhidas no tópico 4.2.1 os coeficientes de capacidade são estimados para os padrões de qualidade mostrados pela Tabela 13. É realizada uma discussão mais detalhada para a DBO a lagoa facultativa, de maturação primária e secundária utilizando uma tolerância de 10, 20 e 30%. Posteriormente, o procedimento é repetido para os demais parâmetros de qualidade de efluente.

### 4.4.2 Índice de Capacidade do Processo

A Tabela 16 mostra o histograma, coeficientes de capacidade e ppm (partes por milhão em não conformidade) para as lagoas que compõem a ETE PS5 para uma tolerância de 0, 10 e 20% utilizando a COEMA n° 02 de 02/02/2017 como garantia de conformidade de DBO 5 dias de, 20°C de até 120 mg/L.

De início é observado que em todas as lagoas, a medida que a tolerância aumenta, a garantia também aumenta. Tanto relacionado ao número de amostras em não conformidade (ppm) e a dispersão da variação dos resultados ( $C_{pk}$ ), já que o intervalo de aceitação torna-se mais extenso. Um comportamento peculiar já identificado com a metodologia *Bootstrap* é que a matéria orgânica na forma de DBO aumenta após a lagoa facultativa na de maturação primária. Provavelmente, devido a biomassa algal.

O  $C_{pk}$  que recebe o valor de  $C_{pki}$ , por ser menor que o  $C_{pks}$ , recebe valores negativos sempre que a “massa de dados” a direita do limite de garantia é superior a esquerda. Logo, valores positivos e elevados de  $C_{pk}$  são desejáveis indicando que existem muitos valores de DBO efluente que estão à direita do limite de garantia e que a dispersão dos dados não prejudica substancialmente a conformidade dos dados.

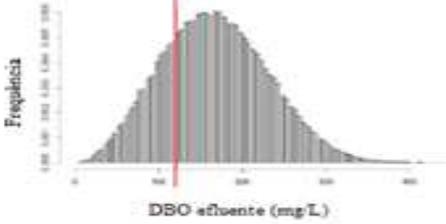
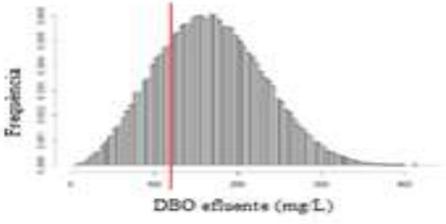
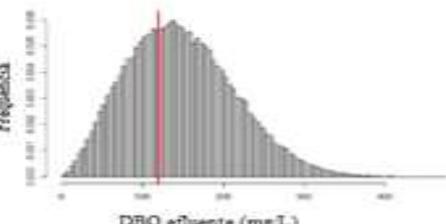
É interessante observar que para uma análise correta do índice de capacidade tanto é importante atentar para a concentração de dados em um intervalo ( $C_p$ ) e tal como a dispersão destes ( $C_{pk}$ ). Não adianta os dados estarem concentrados em uma faixa fora da garantia ou que estes estejam demasiadamente dispersos comprometendo a conformidade dos resultados, já que o preferível é que a massa de dados esteja concentrada dentro do intervalo de aceitação com baixa dispersão.

O histograma foi colocado na Tabela 26 com o intuito de mostrar que a análise do índice de capacidade tanto deve levar em consideração o aspecto da concentração dos dados como da dispersão. Note que o  $C_p$  de 0,41, grifado em itálico, é idêntico para o efluente da lagoa de maturação primária e secundária. Entretanto, o número de amostras em conformidade é muito superior para o efluente da lagoa de maturação secundária. Logo, percebe-se que a dispersão dos dados diminuiu entre estas duas operações.

O índice de capacidade é calculado para o limite de 120 mg/L de DBO e com tolerância de 0, 10 e 20% (120, 132 e 144 mg/L). Com esta abordagem é buscado aferir o “ganho marginal” de estender o intervalo de aceitação tornando a garantia mais branda. Note que sempre o número de amostras em não conformidade diminui a medida que a tolerância aumenta. Entretanto, a proporção no qual o número de amostras em não conformidade reduz é diferente para cada situação, como mostrada pela Tabela 26. Este comportamento revela que a concentração da “massa de dados” é diferente para cada intervalo. Logo, deve-se buscar uma tecnologia que evite a flutuação dos dados quando existe uma diminuição substancial das amostras em não conformidade com um abrandamento da garantia.

Observada as figuras contidas na Tabela 26 nota-se que a dispersão dos dados compromete consideravelmente a conformidade da norma de 120 mg/L para a ETE PS<sub>5</sub>. Existe uma grande massa de dados a direita do limite, representado em vermelho, e além disto os dados se distribuem em uma ampla faixa de variação. Este comportamento é mostrado por valores baixos de  $C_p$  (dispersão de dados em um amplo intervalo) e valores negativos de  $C_{pk}$  (grande volume de dados a direita do limite de aceitação).

**Tabela 26** – Índices de Capabilidade para as os efluentes das lagoas que compõem a ETE PS5, para a DBO.

Estação PS5	[0, 120] mg/L		[0, 132] mg/L		[0, 144] mg/L	
<p><b>LFP</b></p>  <p>DBO efluente (mg/L)</p>	$C_{pki}$	-0,12	$C_{pki}$	-0,1	$C_{pki}$	-0,06
	$C_{pka}$	1,04	$C_{pka}$	1,03	$C_{pka}$	1,05
	$C_{pk}$	-0,12	$C_{pk}$	-0,1	$C_{pk}$	-0,06
	$C_p$	0,22	$C_p$	0,23	$C_p$	0,3
	ppm	687944	ppm	665432	ppm	590644
<p><b>LMP</b></p>  <p>DBO efluente (mg/L)</p>	$C_{pki}$	-0,27	$C_{pki}$	-0,17	$C_{pki}$	-0,11
	$C_{pka}$	1,13	$C_{pka}$	1,13	$C_{pka}$	1,13
	$C_{pk}$	-0,27	$C_{pk}$	-0,17	$C_{pk}$	-0,11
	$C_p$	0,32	$C_p$	0,39	$C_p$	0,41
	ppm	795883	ppm	698650	ppm	626381
<p><b>LMS</b></p>  <p>DBO efluente (mg/L)</p>	$C_{pki}$	-0,09	$C_{pki}$	0,01	$C_{pki}$	0,01
	$C_{pka}$	1,08	$C_{pka}$	1,07	$C_{pka}$	1,08
	$C_{pk}$	-0,09	$C_{pk}$	0,01	$C_{pk}$	0,01
	$C_p$	0,35	$C_p$	0,38	$C_p$	0,41
	ppm	616206	ppm	492730	ppm	487961

Fonte: o autor, 2018.

#### 4.4.3 - Índice de Capabilidade do Processo – demais parâmetros

De maneira análoga ao procedimento mostrado no tópico 4.4.2 são construídos os coeficientes de capabilidade para os demais parâmetros de qualidade. A FDP escolhida para cada parâmetro de qualidade foi a mesma encontrada pela Tabela 27.

**Tabela 27** – Coeficientes de capacidade dos parâmetros de qualidade (AMT, pH e CTT, DBO, DBOF, DQO e DQOF) efluente as lagoas de estabilização pertencentes a estação PS5.

LF		LMP		LMS		Histograma
$C_{pld}$	-0,12	$C_{pld}$	0,01	$C_{pld}$	0,07	
$C_{plk}$	1,25	$C_{plk}$	1,16	$C_{plk}$	1,05	
$C_{pk}$	-0,12	$C_{pk}$	0,01	$C_{pk}$	0,07	
$C_p$	0,56	$C_p$	0,53	$C_p$	0,38	
<b>ppm</b>	656545	<b>ppm</b>	488952	<b>ppm</b>	406494	
$C_{pld}$	2,64	$C_{pld}$	2,63	$C_{pld}$	1,97	
$C_{plk}$	0,06	$C_{plk}$	0,1	$C_{plk}$	0,13	
$C_{pk}$	0,06	$C_{pk}$	0,1	$C_{pk}$	0,13	
$C_p$	0,9	$C_p$		$C_p$	0,73	
<b>ppm</b>	391712	<b>ppm</b>	338510	<b>ppm</b>	298656	
$C_{pld}$	-0,08	$C_{pld}$	-0,04	$C_{pld}$	-0,04	
$C_{plk}$	1	$C_{plk}$	1	$C_{plk}$	1	
$C_{pk}$	-0,08	$C_{pk}$	-0,04	$C_{pk}$	-0,04	
$C_p$	0	$C_p$	0	$C_p$	0	
<b>ppm</b>	992069	<b>ppm</b>	967824	<b>ppm</b>	929507	
$C_{pld}$	-0,1	$C_{pld}$	-0,12	$C_{pld}$	-0,07	
$C_{plk}$	1,02	$C_{plk}$	1,05	$C_{plk}$	1,05	
$C_{pk}$	-0,1	$C_{pk}$	-0,12	$C_{pk}$	-0,07	
$C_p$	0,18	$C_p$	0,25	$C_p$	0,29	
<b>ppm</b>	671385	<b>ppm</b>	681154	<b>ppm</b>	596954	
$C_{pld}$	0,39	$C_{pld}$	0,53	$C_{pld}$	0,47	
$C_{plk}$	-0,34	$C_{plk}$	-0,23	$C_{plk}$	0,03	
$C_{pk}$	-0,34	$C_{pk}$	-0,23	$C_{pk}$	0,03	
$C_p$	0,31	$C_p$	0,42	$C_p$	0,39	
<b>ppm</b>	49165	<b>ppm</b>	27159	<b>ppm</b>	44100	

	LF	LMP	LMS	Histograma
$C_{pki}$	-0,21	$C_{pki}$ 0,03	$C_{pki}$ 0,2	
$C_{pka}$	1,15	$C_{pka}$ 1,12	$C_{pka}$ 1,06	
$C_{pk}$	-0,21	$C_{pk}$ 0,03	$C_{pk}$ 0,2	
$C_p$	0,39	$C_p$ 0,49	$C_p$ 0,5	
<b>ppm</b>	733459	<b>ppm</b> 497918	<b>ppm</b> 235245	
$C_{pki}$	-0,08	$C_{pki}$ 0,33	$C_{pki}$ 0,39	
$C_{pka}$	1,2	$C_{pka}$ 1,11	$C_{pka}$ 1,04	
$C_{pk}$	-0,08	$C_{pk}$ 0,33	$C_{pk}$ 0,39	
$C_p$	0,53	$C_p$ 0,64	$C_p$ 0,57	
<b>ppm</b>	582066	<b>ppm</b> 161506	<b>ppm</b> 99070	

Fonte: o autor

As figuras contidas na Tabela 27 são referentes aos dados de distribuição de frequência proveniente da lagoa facultativa presente na estação PS<sub>5</sub>. É observado que a distribuição de frequência é “particular” de cada parâmetro de qualidade. Através desta abordagem foi construído os coeficientes de capacidade utilizando uma função distribuição de probabilidade adequada para cada parâmetro.

Através desta metodologia é possível priorizar recursos e esforços para os parâmetros que trazem maior prejuízo relacionado a adequabilidade do efluente. É fácil de observar que os CTT é o parâmetro de qualidade mais “problemático” e que existe apenas uma pequena melhora quando o tratamento passa pelo estágio da LMP e LMS. Por outro lado o OD é o parâmetro que ,comparado com os demais, traz menos problemas com a adequabilidade. Conclui-se que a elevada capacidade fotossintética consegue prover oxigênio em abundância para as lagoas. Por outro lado é exposto que a remoção de CTT não atinge níveis satisfatórios, sendo necessário, buscar outros métodos para a remoção de microorganismos antes do despejo final.

## CONCLUSÃO

Aqui são expostas algumas conclusões obtidas através da metodologia proposta para a confecção deste trabalho, que busca analisar a qualidade do tratamento do efluente através de ferramentas robusta.

### *Bootstrap*

#### Método paramétrico

- Foi proposta uma metodologia para identificar a função distribuição de probabilidade que melhor tenha aderência com os dados originais;
- Fora construídos Intervalos de Confiança (IC) para cada concentração de DBO efluente identificando um valor ótimo de máxima performance (quartil de 97,5%) e de pior performance (quartil de 2,5%), referente a percentagem de amostras em conformidade para uma garantia buscada.
- Foi discutida a performance das estações utilizando o conceito de confiabilidade, robustez e resiliência;
- Analisada a garantia probabilística para os principais parâmetros de qualidade para as estações de tratamento aqui estudadas utilizando a metodologia proposta.

#### Método não paramétrico

- Foi identificado a garantia máxima, mínima e média para cada lagoa;
- Foi comparado os IC de lagoas, identificando:
  1. Situações anômalas no qual a faixa de variação de dados é demasiadamente extensa;
  2. Região no qual o número de amostras em conformidade permanece constante;
  3. A lagoa que fornece uma melhor performance frente as demais e em qual faixa de concentração;
  4. Acompanhar a variação dos IC a medida que o tratamento ocorre;

#### Fuzzy integrado à simulação Bootstrap

- Foi calculado do risco *Fuzzy* para cada decil averiguando a região que mais oferece garantia para o intervalo de DBO efluente;

- Levantado o risco associado ao número *Fuzzy* triangular para os parâmetros de qualidade do esgoto efluente as lagoas de estabilização;

#### Controle Estatístico de Processos

- Foi identificado como a aleatoriedade dos dados influi na garantia, relacionando tanto com a faixa de dispersão e com a concentração dos dados em um intervalo;
- Foi mensurado o número de amostras em não conformidade (ppm) para cada lagoa e como a tolerância influi neste valor;
- Diagramar como a aleatoriedade dos dados influi na garantia;
- Levantamento dos coeficientes de capacidade para os demais parâmetros de qualidade e analisar quais são os mais “problemáticos”.

## REFERÊNCIAS

- AKAIKE, H. Information theory as an extension of the maximum likelihood principle. **In: Petrov, B. N., F. (Eds.)**, Second International Symposium of Information. Budapeste, 1973. p. 267-281.
- APHA. Standard methods for the examination of water and wastewater. 20<sup>th</sup> edition. **American Public Health Association**, American Water Works Association and Water Environment Federation. New York, 1998. 1325p.
- ARAÚJO, A. L. C. **Diagnóstico de Sistemas de Lagoas de Estabilização no Rio Grande do Norte**. Fundação Nacional de Saúde – FUNASA. Relatório final. Natal - RN, 2011. p.286.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 9001. **Sistema de Gestão de Qualidade – requisitos**. Rio de Janeiro, 2015. p.5.
- ARNOLD, T. W. Uninformative parameters and model selection using Akaike's Information Criterion. **Journal of Wildlife Management** v. 74, p.1175–1178, 2010.
- BASIRI, S. Enhanced *Bootstrap* method for statistical inference in the ICA model. **Signal Processing** v. 138, p. 53-62, 2017.
- BAGEN, B.; HUANG, D. A. subsystem based probabilistic approach for the assessment of transmission capital project reliability impacts. **Electric Power Systems Research** v. 150, p.76–83, 2017.
- BELIA, E.; VANROLLEGHEM, P.A.; TALEBIZADEH, M. Influent generator for probabilistic modeling of nutrient removal wastewater treatment plants. **Environmental Modelling & Software** v.77, p. 32–49, 2016.
- BERTHOUEX, P. M.; HUNTER, W. G. Simple statistics for interpreting environment data. **Journal WPCF** v.53, p.167-175, 1981.
- BENEDETTI, L.; DIRCKX, G.; BIXIO, D.; THOEYE, C.; VANROLLEGHEM, P.A. Environmental and economic performance assessment of the integrated urban wastewater system. **Journal of Environmental Management** v.88, p.1262–1272, 2008.
- BOURRIER, R.; SATIN, M.; SELMI, B.; **Guide technique de l'assainissement**. Fourth edition. Le moniteur, Paris. 2010.

BRANDÃO, L. P. **Prospects of reusing treated wastewater stabilization ponds for irrigation in the state of Ceará**. Civil Engineering. Federal University of Ceará. Fortaleza, CE., 2000.

BINNY, D.; MEZZENGA, E.; LANCASTER, C. M.; TRAPP, J. V.; KAIRN, T.; CROWE, S.B. Physica medica investigating output and energy variations and their relationship to delivery delivery QA results using statistical process control for helical tomotherapy. **Physica Medica** v. 38, p.105–110. 2017.

BRASIL. Conselho Nacional do Meio Ambiente (CONAMA). **Resolução n° 357, de 17 de março de 2005**. Diário Oficial da União, n° 053, de 18 de março de 2005, p. 58-63, Brasília-DF, 2005.

BRASIL. Conselho Nacional do Meio Ambiente (CONAMA). **Resolução n° 430, de 13 de maio de 2011**. Diário Oficial da União, n° 92, de 16 maio de 2011, p. 89-97, Brasília-DF, 2011.

BRASIL. Conselho Nacional do Meio Ambiente (CONAMA). **Resolução n° 20, de 18 de junho de 1986**. Diário Oficial da União, de 30 de julho de 1986, p. 12-36, Brasília-DF, 1986.

BRASIL. **Lei n° 11.445, de 05 de janeiro de 2007**. Estabelece diretrizes nacionais para o saneamento básico; altera as Leis nos 6.766, de 19 de dezembro de 1979, 8.036, de 11 de maio de 1990, 8.666, de 21 de junho de 1993, 8.987, de 13 de fevereiro de 1995; revoga a Lei no 6.528, de 11 de maio de 1978; e dá outras providências. Disponível em: <[http://www.planalto.gov.br/ccvil\\_03/\\_Ato2007-2010/Lei/L11445.htm](http://www.planalto.gov.br/ccvil_03/_Ato2007-2010/Lei/L11445.htm)>

CLEMENTE, T. C. B.; OLIVEIRA, S. M. A. C. Instrumento para o gerenciamento de riscos ambientais Instrument for the environmental risks management. **Engenharia Sanitária e Ambiental** v. 21, p. 227-236, 2015.

CHEN, S. M.; SANGUANSAT, K. Analyzing *fuzzy* risk based on a new *fuzzy* ranking method between generalized *fuzzy* numbers. **Expert Systems Application** v.38, p. 2163–2171, 2011.

CEARÁ. Conselho Estadual de Meio Ambiente (COEMA). **Portaria de n°21, de 21 de fevereiro de 2017**. Diário Oficial do Estado, série 3, ano IX n° 37.

COLIN, J.; VANHOUCHE, M. Developing a framework for statistical process control approaches in project management. **Internacional Journal of Project Management** v. 33, p. 1289-1300.

COGGINS, L. X.; GHISALBERTI, M.; GHADOUANI, A. Sludge accumulation and distribution impact the hydraulic performance in waste stabilisation ponds. **Water Research** v.110, p.354–365, 2017.

COUTO, E. A.; PINTO, F.; VARELA, F.; REIS, A.; COSTA, P.; CALIJURI, M. L. Hydrothermal liquefaction of biomass produced from domestic sewage treatment in high-rate ponds. **Renewable Energy** v.118, p. 644-653, 2018.

CUELLAR-BERMUDEZ, S. P.; ALEMAN-NAVA, G. S.; CHANDRA, R., GARCIA-PEREZ, J. S.; CONTRERAS-ANGULO, J. R.; MARKOU, G.; PARRA-SALDIVAR, R. Nutrients utilization and contaminants removal . A review of two approaches of algae and cyanobacteria in wastewater. **Algal Research** v. 24, p.438-449, 2016.

DÍAZ-GARDUÑO, B; PINTADO-HERRERA, M. G.; BIEL-MAESO, M.; RUEDA-M, J. J. Environmental risk assessment of effluents as a whole emerging contaminant : Efficiency of alternative tertiary treatments for wastewater depuration. **Water Research** v.119, p.136–149, 2017.

DIAS, D. F. C.; POSSMOSER-NASCIMENTO, T. E.; RODRIGUES, V. A. J.; VON SPERLING, M. Overall performance evaluation of shallow maturation ponds in series treating UASB reactor effluent : Ten years of intensive monitoring of a system in Brazil. **Ecological Engineering** v.71, p.206–214, 2014.

EFRON, B. Bootstrap methods: another look at the jackknife. **Ann Stat** v.7, p.1–26, 1979.

FILION, G. J. The signed Kolmogorov-Smirnov test: why it should not be used. **Giga Science** v. 86, p. 4-9, 2015.

FLORES-ALSINA, X.; RODRIGUEZ-RODA, I.; SIN, G.; GERNAEY, K.V.; Multi-criteria evaluation of wastewater treatment plant control strategies under uncertainty. **Water Research** v.42, p. 4485–4497, 2008.

FERNANDEZ, D.; LIU, I. A goodness-of-fit test for the ordered stereotype model. **Statistics in Medicine** v.35, p. 4660-4696. 2016.

FONT, B.; Bootstrap estimation of the efficient frontier. **Comput Manag Sci** v. 13, p. 541-570, 2011.

GOMES, F. M. **Controle Estatístico de Processos**. Técnico em Qualidade, Lorena, SP. 2010.

GOERLANDT, F.; KHAKZAD, N.; RENIERS, G. Validity and validation of safety-related quantitative risk analysis: a review. **Safety Science** v.99, p.127-139. 2017.

GARGAMA, H.; CHATURVEDI, S. K. Criticality Assessment Models for Failure Mode Effects and Criticality Analysis Using Fuzzy Logic. **IEEE Transactions on Reliability** v. 60, n. 1, p. 102-110. 2011.

GUERRINI, A.; ROMANO, G.; MANCUSO, F.; CAROSI, L. Identifying the performance drivers of wastewater treatment plants through conditional order efficiency analysis. **Utilities Policy** v. 42, p.20–31. 2016.

GUO, J.; IWATA, H. Ecotoxicology and environmental safety risk assessment of triclosan in the global environment using a probabilistic approach. **Ecotoxicology and Environmental Safety** v.143, p.111–119. 2017.

GARCIA, S. N.; CLUBBS, R. L.; STANLEY, J. K.; SCHEFFE, B.; YELDERMAN, J. C.; BROOKS, B. W. Comparative analysis of effluent water quality from a municipal treatment plant and two on-site wastewater treatment systems. **Chemosphere** v. 92, p. 38–44, 2013.

GRAVELINE, N.; LOUBIER, S.; GLEYSES, G.; RINAUDO, J. Impact of farming on water resources : Assessing uncertainty with Monte Carlo simulations in a global change context. **Agricultural Systems** v.108, p.29–41. 2012.

GEJDOS, P. Continuous quality improvement by statistical process control. **Procedia Economics and Finance** v. 34, p. 565-572, 2015.

HSIAO, C. K.; YANG, C. C.; BJORNLUND, H. Performance measurement in wastewater control. Pig farms in Taiwan. **WIT Trans. Ecol. Environ.** v.1, p.467 e 474. 2007.

IANCU, I. Heart disease diagnosis based on mediative fuzzy logic. **Artificial Intelligence in Medicine**. v. 89, p.51-60, 218.

JOHNSON, A. C.; JÜRGENS, M. D.; LAWLOR, A. J.; CISOWSKA, I.; WILLIAMS, R. J. Particulate and colloidal silver in sewage effluent and sludge discharged from British wastewater treatment plants. **Chemosphere** v.112, p. 49–55, 2014.

YUAN, J.; EMAGURA, K. F.C.; ALAM, M. A. Frequency analysis of annual maximum hourly precipitation and determination of best fit probability distribution for regions in Japan. **Urban Climate** v. 24, p. 276-286, 2018.

KAUFMANN, A.; GUPTA, M. M. Introduction to Fuzzy Arithmetic: theory and applications. **Journal of the Franklin Institute** v. 321, p.189-190.1991.

KELLER, W.; MODARRES, M. A historical overview of probabilistic risk assessment development and its use in the nuclear power industry : a tribute to the late Professor Norman Carl Rasmussen. **Fuel and Energy Abstract** v. 89, p. 271–285, 2005.

KOHN, T.; MATTLE, M. J.; MINELLA, M.; VIONE, D. A modeling approach to estimate the solar disinfection of viral indicator organisms in waste stabilization ponds and surface waters. **Water Research** v.88, p.912–922. 2016.

KRZYŻANOWSKI, F.; SOUZA, M.; CÁSSIA, A.; INÊS, M., SATO, Z.; TEREZA, M., RAZZOLINI, P. Assessing the probability of infection by Salmonella due to sewage sludge use in agriculture under several exposure scenarios for crops and soil ingestion. **Science of the Total Environment** v. 568, p. 66–74, 2016.

KHORSHIDI, H.A.; NIKFALAZAR, S. An improved similarity measure for generalized Fuzzy numbers and its application to fuzzy risk analysis. **Applied Soft Computing** v.52 p. 478–486, 2017.

LIMA, A. A. N.; LIMA, J. R.; SILVA, J. L.; ALENCAR, J. R. B.; SOARES-SOBRINHO, J. L.; LIMA, L.G.; ROLIM-NETO, P.J. Aplicação do controle estatístico de processo na indústria farmacêutica. **Revista de Ciência Farmacêutica Básica e Aplicada** v. , p.177-187, 2007.

LI, Y. P.; HUANG, G. H. Risk analysis and management for water resources systems. **Stochastic Environmental Research and Risk Assessment** v.27, p. 593–597, 2013.

LI, P.; ZHANG, L.; DAI, L.; ZOU, Y.; LI, X. An assessment method of operator's situation awareness reliability based on fuzzy logic-AHP. **Safety Science** v. 23, p. 102-106, 2018.

MARTORELL, S.; MARTORELL, P.; MARTÓN, I.; SÁNCHEZ, A. I.; CARLOS, S. An approach to address probabilistic assumptions on the availability of safety systems for deterministic safety analysis. Reliability **Engineering and System Safety** v. 160, p.136–150, 2017.

MAYNARD, H. E.; OUKI, S. K.; WILLIAMS, S. C. Review Paper Tertiary Lagoons : A Review Of Removal Mechanisms And Performance. **Water Research** v.33, p. 1. 1999.

MAYO, A. W.; ABBAS, M. Removal mechanisms of nitrogen in waste stabilization ponds. **Physics and Chemistry of the Earth** v.72, p.77–82, 2014.

MIDDLEBROOKS, E. J.; REED, S. C.; PANO, A.; ADAMS, V. D.; MIDDLEBROOKS, E. J.; REED, S. C.; ADAMS, V. D. Nitrogen Removal in Wastewater Stabilization Lagoons

By Presented at 6 National Drinking Water and Wastewater Treatment Technology Transfer Workshop Kansas City, **Nevada Nitrogen Removal in Wastewater Stabilization Lagoons**, p.1–38, 1999.

MIRABI, M.; MIANABADI, H.; ZARGHAMI, M. Risk-based evaluation of wastewater treatment projects : A case study in Niasar city, Iran. **Resources, Conservation & Recycling** v.93, p.168–177, 2014.

MIZYED, N, R. Challenges to treatment wastewater reuse in arid and semi-arid áreas. **Environment Science & Policy** v. 25, p. 186-195. 2013.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied Statistics and Probability for Engineers**. United States of America, New Jersey, 6<sup>th</sup> edition, 2014.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 2. ed. Rio de Janeiro: LTC, 2004.

MULLER, F.; ZEILER, P.; BERTSCHE, B. Bootstrap Monte Carlo simulation of reability and confidence level with periodical maintenance. **Forsch Ingenieurwes** v.81, p. 383–393, 2017.

MUHAMMETOGLU, A.; YARDIMCI, A. A Fuzzy logic approach to assess groundwater pollution levels below agricultural fields. **Environment Monitoring and Assessment** v.118, p.337-354. 2006.

NAGHETTINI, M.; PINTO, É. J. A. **Hidrologia Estatística**. Serviço Geológico do Brasil. Companhia de Recursos Minerais. Belo Horizonte: CPRM, 2007. 561 p.

NIKU, S.; SAMANIEGO, F.J. Performance of activade sludge process and reability-based design. **Water Envinment Federation**, v. 51, p. 2841-2857, 1983.

NIKU, S.; SCHROEDER, E, D.; HAUGH, R, S. Reliability and stability of trickling filter process. **Journal WPCF** v. 54, p. 129-134, 1982.

OLIVEIRA, S. C.; VON SPERLING, M. Reliability analysis of wastewater treatment plants. **Water Research** v. 42, p.1182–1194, 2008.

OLIVEIRA, S. M. A. C. **Análise do Desempenho e Confiabilidade de Estações de Tratamento de Esgoto**. 2006. 232 f. Tese (Doutorado em Meio Ambiente e Recursos Hídricos) – Programa de Pós Graduação em Saneamento, Universidade de Minas Gerais, Belo Horizonte, 2006.

PAULI, S.; ARBENZ, P. Determining optimal multilevel Monte Carlo parameters with application application to fault tolerance. **Computers and Mathematics with application** v.70, p. 2638-2651, 2015.

PEDA, P.; GROSSI, G.; LIIK, M. Do ownership and size affect the performance of water utilities? Evidence from Estonian municipalities. **J. Manag. Gov.** v.17, p. 237-259, 2013.

PRADA, A. F.; CHU, M.L.; GUZMAN, J.A.; MORIASI, D.N. Evaluating the impacts of agricultural land management practices on water resources: A probabilistic hydrologic modeling approach. **Journal of Environmental Management** v.193, p.512-523. 2017.

PIMPUNCHAT, B.; SWEATMAN, W.L.; WAKE, G. C.; TRIAMPO, W.; PARSHOTAM, A. A mathematical model for pollution in a river and its remediation by aeration. **Applied Mathematics Letters**, v.22, p.304–308, 2009.

POINTON, C.; MATTHEWS, K. Reprint of : Dynamic efficiency in the English and Welsh water and sewerage industry. **Omega** v.60, p.98–108, 2016.

PEI, W.; ZHAN, Q.; YAN, Z.; GE, L.; ZHANG, P.; WANG, Z.; WANG, D. Using Monte Carlo simulation to assess uncertainty and variability of methamphetamine use and prevalence from wastewater analysis. **International Journal of Drug Policy** v.36, p. 1–7, 2016.

R: **A language and Environment for Statistics Computing**, version 3.4.0. R Foundation for Statistical Computing, 2017.

RAGUSH, C. M.; SCHMIDT, J. J.; KRKOSEK, W. H.; GAGNON, G.A.; TRUELSTRUP-HANSEN, L.; JAMIESON, R. C. Performance of municipal waste stabilization ponds in the Canadian Arctic. **Ecological Engineering** v. 83, p. 413–421. 2015.

RIZZO, A. L.T.; CYMROT, R. **Estudo e Aplicações da Técnica Bootstrap**. Universidade Presbiteriana Mackenzie. II Jornada de Iniciação Científica, 2006.

RODRIGUEZ-GARCIA, G.; MOLINOS-SENANTE, M.; HOSPIDO, A.; HERNÁNDEZ-SANCHO, F.; MOREIRA, M. T.; FEIJOO, G. Environmental and economic profile of six typologies of wastewater treatment plants. **Water Research** v. 45, p. 5997–6010, 2011.

ROSS, R. L. G. **Simulação Monte Carlo e número Fuzzy triangular na avaliação de garantia de desempenho de lagoas de estabilização**. Trabalho final de curso – Faculdade de Engenharia Química, Universidade Federal do Ceará, 2016.

SANCTIS, M.; MORO, G. D.; CHIMIANTI, S.; PIERANGELA, R.; LEVANTESI, C.; IACONI, C. D. Removal of pollutants and pathogens by a simplified treatment scheme for municipal wastewater reuse in agriculture. **Science of the Total Environment** v.580 p.17–25, 2017.

SANCHO, J.; IGLESIAS, C.; PIÑERO, J.; MARTINEZ, J, PASTOR, J, J.; ARAUJO, M.; TABOADA, J. Study of Water Quality in a Spanish River Based on Statistics Process Control and Function Data Analysis. **Math Geosci**, v. 48, p.163-186. 2016.

SADIQ, R.; KLEINER, Y.; RAJANI, B. Water quality failures in distribution networks - Risk analysis using Fuzzy logic and evidential reasoning. **Risk Analysis** v.27, p.1381–1394, 2007.

SECRETARIA DO MEIO AMBIENTE DO ESTADO DO CEARÁ (SEMACE). Portaria nº 154, de 22 de Julho de 2002. Publicado no Diário Oficial do Município. 1 de Outubro de 2002.

SILVA, C.; QUADROS, S.; RAMALHO, P.; ALEGRE, H.; ROSA, M. J. Translating removal efficiencies into operational performance indices of wastewater treatment plants. **Water Research** v.57, p. 202–214, 2014.

SILVA, R. B.; BOURGUIGNON, M.; CORDEIRO, G. M. Journal of Computational and Applied A new compounding family of distributions : The generalized gama power series distributions. **Journal of Computational and Applied Mathematics** v.303, p.119–139, 2016.

SHANNON, M.; BOHN, P. W.; ELIMELECH, M.; GEORGIADIS, J. G.; MARIÑAS, B. J.; MAYES, A. M. Science and technology for water purification in the coming decades. **Nature**, v. 452, p.301–310, 2008.

SNIPES, M.; TAYLOR, C. Model selection and Akaike Information Criteria: An exemplo from wine rating and prices. **Wine Economics and Policy** v. 3, p. 3-9, 2014.

SALSBURG, D. **Uma senhora toma chá... Como a estatística revolucionou a ciência do século XXI**. Zahar. Brasil: Rio de Janeiro, 2009.

SCHENDEL, T.; THONGWICHIAN, R. Flood frequency analysis: confidence interval estimation by test inversion Bootstrapping. **Advances in Water Resources** v.83, p.1–9, 2015.

VON SPERLING, M. V. Comparison among the most frequently used systems for wastewater treatment in developing countries. In: **International Symposium on Technology Transfer**, vol. 1, pp. 107–120. 1996.

SLEIGH, P. A.; AND MARA, D. D. (2003) Monte Carlo Program for Facultative Pond Design, available at <http://www.efm.leeds.ac.uk/CIVE/Mcarlo/index.html>.

SKUJL, G.; VRABIC, R.; BUTALA, P.; SLUGA, A. Statistical process control as a service. **Procedia CIRP** v. 23, p. 410-406. 2013.

SUTHERLAND, D. L.; TURNBULL, M. H.; CRAGGS, R. J. Increased pond depth improves algal productivity and nutrient removal in wastewater treatment high rate algal ponds. **Water Research** v.53, p.271–281. 2014.

TAHERIYOUN, M.; MORADINEJAD, S. Reliability analysis of a wastewater treatment plant using fault tree analysis and Monte Carlo simulation. **Environmental Monitoring and Assessment** v.187, p. 4186, 2015.

TCHOBANOGLIOUS, G.; NELSON, K. L.; JIM, B.; DARBY, J. L. Sludge accumulation characteristics and pathogen inactivation in four primary waste stabilization ponds in central Mexico. **Water Research** v. 38, p. 111–127. 2004.

ÜSTÜN, G. E. Occurrence and removal of metals in urban wastewater treatment plants, **Journal of Hazardous Materials** .v.172, p.833–838. 2009.

UPADHYAYA, J. K.; MOORE, G. Sustainability indicators for wastewater reuse systems and their application to two small systems in rural Victoria, **Australia. Can. J. Civ. Eng.** v.39, p.674-688. 2012.

UPADHYAY, R.; SEHWAG, S.; MISHRA, H. N. Electronic nose guided determination of frying disposal time of sunflower oil using Fuzzy logic analysis. **Food Chemistry** v.221, p.379-385. 2017.

VEERESH, M.; VEERESH, A, V., HUDDAR, B, D.; HOSETTI, B, B. Dynamic of Industrial waster stabilization pond treatment process, **Environment Monit Assess**, v.169, p.55-65. 2010.

VON SPERLING, M. Lagoas de estabilização. In: **Princípios do tratamento biológico de águas residuárias**. Volume 3. 2ª Edição. Departamento de Engenharia Sanitária e Ambiental – DESA. Belo Horizonte: Universidade Federal de Minas Gerais, 2002. 196 p.

SPERLING, M. V. Modelling of coliform removal in 186 facultative and maturation ponds around the world. **Water Research** v. 39, p. 5261–5273. 2005.

WANG, G.; LI, J. Approximations of Fuzzy numbers by step type Fuzzy numbers. **Fuzzy Sets and Systems** v.31, p. 47–59, 2016.

WALLACE, J.; CHAMPAGNE, P.; HALL, G. Multivariate statistical analysis of water chemistry conditions in three wastewater stabilization ponds with algae blooms and pH fluctuations. **Water Research** v.96, p.155–165, 2016.

MCINTYRE, H. S.; WHEATER, A. N. R. M. A tool for risk-based management of surface water quality. **Environment Modelling and Software** v. 19, p.1131–1140, 2003.

YI, Q.; HUR, C.; KIM, Y. Modeling nitrogen removal in water hyacinth ponds receiving effluent from waste stabilization ponds, **Ecological Engineering** v.5, p.75–84. 2008.

YANG, T.; LONG, R.; CUI, X.; ZHU, D.; CHE, H. Application of the public – private partnership model to urban sewage treatment. **Journal of Cleaner Production** v. 142, p.1065-1074, 2007.

ZHANG, K.; ACHARI, G.; SADIQ, R.; LANGFORD, C. H.; DORE, M. H. I. An integrated performance assessment framework for water treatment plants. **Water Research** v.46, p.1673–1683. 2012.

ZENG, S.; HUANG, Y.; SUN, F.; LI, D.; HE, M. Probabilistic ecological risk assessment of effluent toxicity of a wastewater reclamation plant based on process modeling. **Water Research** v. 100, p.367–376, 2016.

**APÊNDICE A - TESTE DE ADERÊNCIA PARA AS FDP DO EFLUENTE AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>,**

**Tabela A1** - Teste de Aderência utilizando o Teste de KS e AIC para a Estação PS<sub>5</sub>,

<b>Parâmetro</b>	<b>EB</b>	<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
DBO	(gu/0,9978)/ (gu/262,8822)	(n/0,9564)/ (w/242,2757)	(n/0,8046)/ (gam/225,4715)	(n/0,6044)/ (gam/227,7654)
DBO <sub>F</sub>	(gam/0,9847)/ (gam/247,8192)	(n/0,9544)/ (w/221,1589)	(gu/0,8616)/ (gu/206,9812)	(w/0,6630)/ (w/203,7665)
DQO	(n/0,9741)/ (gam/285,1842)	(gu/0,9290)/ (ln/250,2859)	(w/0,7187)/ (w/249,7708)	(ln/0,8939)/ (w/287,3698)
DQO <sub>F</sub>	(ln/0,9829)/ (gam/258,7775)	(w/0,7866)/ (gu/223,1110)	(ln/0,6674)/ (ln/217,6218)	(ln/0,7977)/ (ln/219,5053)
pH	(ln/0,7977)/ (ln/219,5138)	(gu/0,6340)/ (gu/15,1831)	(ln/0,4210)/ (gu/25,7670)	(gu/0,3213)/ (gu/31,3366)
AMT	(n/0,9379)/ (w/165,8398)	(n/0,4870)/ (n/132,7789)	(w/0,9034)/ (n/138,3977)	(gu/0,8489)/ (gu, 140,7966)
OD	-	(gu/0,8564)/ (gam, 91,2899)	(gu/0,9749)/ (w/ 89,4628)	(gam/0,9652)/ (gam/96,03164)
CTT	(w/0,1366)/ (w/729,2398)	(n/0,8523)/ (w/612,52)	(ln/0,7197)/ (ln/583,7762)	(w/0,7325)/ (ln/532,6295)

Fonte: o autor

**Tabela A2** - Teste de Aderência utilizando o Teste de KS e AIC para a Estação PS<sub>6</sub>,

<b>Parâmetro</b>	<b>EB</b>	<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
DBO	(ln/0,3293)/ (ln/270,5196)	(ln/0,9059)/ (gam/248,8517)	(ga/0,9692)/ (ga/244,1043)	(gu/0,8299)/ (gam/243,2416)
DBO <sub>F</sub>	(gu/0,8156)/ (gu/212,5283)	(ln/0,9720)/ (ln/225,1494)	(w/0,9544)/ (w/215,3780)	(w/0,7822)/ (w/221,7392)
DQO	(ln/0,9781)/ (gu/281,4716)	(ln/0,8600)/ (ln/259,6499)	(ga/0,8629)/ (ga/253,2813)	(ln/0,6246)/ (ln/264,8922)
DQO <sub>F</sub>	(ln/0,8271)/ (ln/236,8597)	(gam/0,9905)/ (ln/242,4352)	(ln/0,9973)/ (ln/ 211,13)	(g/0,2213)/ (gam/232,4569)
pH	(gu/0,6358)/ (ln/17,0127)	(gu/0,5777)/ (gu/13,6789)	(gu/0,4281)/ (gu/0,22075)	(g/0,6141)/ (g/12,6123)
AMT	(gu/0,9683)/ (n/175,4855)	(w/0,8880)/ (w/125,2990)	(w/0,9996)/ (w/124,3301)	(n/0,9254)/ (w/134,6433)
OD	-	(w/0,8255)/ (w/76,4311)	(n/0,8608)/ (w/86,0816)	(n/0,9172)/ (w/ 213,3698)
CTT	(w/0,3047)/ (ln/773,0973)	(ln/0,8931)/ (ln/629,8360)	(w/0,4555)/ (ln/577,5001)	(ln/0,3285)/ (w/520,7322)

Fonte: o autor

**Tabela A3** - Teste de Aderência utilizando o Teste de KS e AIC para a Estação PS<sub>9</sub>,

<b>Parâmetro</b>	<b>EB</b>	<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
DBO	(gam/0,9652)/ (g/270,4895)	(w/0,3481)/ (gam/232,5431)	(ln/0,9880)/ (gam/245,4882)	(w/0,6244)/ (w/231,8462)
DBO <sub>F</sub>	(g/0,9987)/ (g/237,1699)	(n/0,8606)/ (w/209,8137)	(n/0,9820)/ (w/221,1913)	(n/0,8976)/ (gam/208,2259)
DQO	(n/0,9689)/ (gam/278,9725)	(n/0,9032)/ (gam/252,7365)	(g/0,9796)/ (g/249,32)	(g/0,9816)/ (gam/234,96)
DQO <sub>F</sub>	(g/0,9000)/ (g/257,1493)	(w/0,8960)/ (w/229,3265)	(n/0,9093)/ (w/217,1297)	(n/0,9999)/ (w/207,8402)
pH	(ln/0,8554)/ (n/4,3548)	(w/0,7756)/ (gam/6,3045)	(w/0,4698)/ (gam/8,0698)	(w/0,6986)/ (w/12,5478)
AMT	(ln/0,7206)/ (ln/174,3837)	(w/0,7112)/ (w/124,7112)	(w/0,6497)/ (w/0,1697)	(w/0,5488)/ (w/125,1721)
OD	-	(gam/0,8745)/ (gam/83,9145)	(n/0,7492)/ (w/87,3698)	(n/0,8719)/ (w/94,6090)
CTT	(ln/0,8773)/ (ln/767,9396)	(ln/0,09578)/ (ln/590,0553)	(w/0,9791)/ (w/548,2598)	(ln/0,7057)/ (ln/ 498,5954)

Fonte: o autor

**Tabela A4** – FDP recomendadas com base no teste de KS,

<b>Teste KS</b> <b>Parâmetro</b>	<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>
DBO	N Ln W	N G Ln	N Gu W
DBO <sub>F</sub>	W Ln N	Gu W N	W W N
DQO	W Ln N	Gu G Gu	W Ln Gu
DQO <sub>F</sub>	Ln G W	W Ln N	W Gu N
pH	Gu Gu W	Ln Gu W	Gu Gu W
AMT	N W W	W W W	Gu N W
OD	Gu W G	Gu N N	G N N
CTT	N Ln Ln	Ln W N	W Ln W

Fonte: o autor

**Tabela A5** – FDP recomendadas utilizando o teste de AIC.

<b>Parâmetro</b>	<b>Tipo de lagoa e Série</b>		
	<b>LFP</b>	<b>LMP</b>	<b>LMS</b>
	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>	<b>PS<sub>5</sub> PS<sub>6</sub> PS<sub>9</sub></b>
DBO	W G G	G G G	G G W
DBO <sub>F</sub>	W Ln W	G W W	W W G
DQO	Ln Ln G	W G Gu	W Ln G
DQO <sub>F</sub>	Gu Ln W	Ln Ln W	Gu G W
pH	Gu Gu G	Gu Gu G	Gu Gu W
AMT	N W W	N W W	Gu W W
OD	G W G	W W W	G W W
CTT	Ln Ln LN	Ln Ln W	Ln Ln Ln

Fonte: o autor

**APÊNDICE B - IC PARA A GARANTIA BOOTSTRAP PARA OS PARÂMETROS DE QUALIDADE EFLUENTE AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> E PS<sub>9</sub>,**

**Tabela B1** - Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>5</sub>,

<i>Parâmetro</i>	<i>EB</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
DBO	[0,03; 0,05]	[0,18; 0,42]	[0,18;0,36]	[0,23;0,53]
DBO <sub>F</sub>	[0; 0]	[0,22; 0,48]	[0,04; 0,43]	[0,23; 0,53]
DQO	[0; 0,08]	[0,18; 0,40]	[0,38; 0,65]	[0,68; 0,95]
DQO <sub>F</sub>	[0; 0,05]	[0,31; 0,65]	[0,75, 0,98]	[0,83; 0,97]
pH	[0,97; 1]/[1;1]	[0,18; 0,55]/[1;1]	[0,2, 0,55],[1,1]	[0,15,0,4; 1,1]
AMT	[0; 0,08]	[0,25; 0,65]	[0,45, 0,76]	[0,45, 0,85]
OD	-	[0,45; 0,85]	[0,45, 0,79]	[0,4, 0,75]
CTT	[0,0]	[0; 0]	[0,0]	[0,1]

Fonte: o autor

**Tabela B2** - Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>6</sub>,

<i>Parâmetro</i>	<i>EB</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
DBO	[0,03, 0,17]	[0,25; 0,64]	[0,3; 0,63]	[0,3; 0,67]
DBO <sub>F</sub>	[0,03; 0,1]	[0,23; 0,57]	[0,3; 0,63]	[0,37; 0,57]
DQO	[0,02; 0,07]	[0,25; 0,50]	[0,25; 0,63]	[0,53; 0,85]
DQO <sub>F</sub>	[0,03; 0,1]	[0,3; 0,65]	[0,9; 0,97]	[0,73; 0,97]
pH	[0,43; 0,82]/[1,1]	[0,1/ 0,57]/[1,1]	[0,25; 0,53]/[1,1]	[0,15/0,37]/[1,1]
AMT	[0,1; 0,37]	[0,6; 0,9]	[0,7; 0,9]	[0,75/0,95]
OD	-	[0,15; 0,45]	[0,23; 0,5]	[0,2; 0,55]
CTT	[0,0]	[0,0]	[0,0]	[0,0]

Fonte: o autor

**Tabela B3** - Intervalo de confiança para a garantia *Bootstrap* ETE PS<sub>9</sub>,

<i>Parâmetro</i>	<i>EB</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
DBO	[0; 0,03]	[0,1; 0,38]	[0,15; 0,37]	[0,22; 0,6]
DBO <sub>F</sub>	[0; 0,05]	[0,1; 0,45]	[[0,1; 0,4]	[0,23; 0,57]
DQO	[0; 0,1]	[0,27; 0,6]	[0,37; 0,8]	[0,75; 0,97]
DQO <sub>F</sub>	[0,03; 0,2]	[0,5; 0,75]	[0,83; 0,98]	[1,1]
pH	[0,3;0,78]; [1;1]	[0,37; 0,65]; [1;1]	[0,07; 0,37]; [1;1]	[0,07; 0,37]; [1;1]
AMT	[0,05; 0,28]	[0,13; 0,37]	[0,05; 0,3]	[0,3; 0,65]
OD	-	[0,5; 0,85]	[0,3; 0,7]	[0,28; 0,73]
CTT	[0,0]	[0,0]	[0;01]	[0; 0,0]

Fonte: o autor

## APÊNDICE C - ÍNDICES DE CAPABILIDADE DAS ETE PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>

**Tabela C1** – Índices de Capabilidade para o desempenho da ETE PS<sub>5</sub>.

<i>LFP</i>		<i>LMP</i>		<i>LMS</i>		<i>Parâmetro</i>
$C_{pki}$	-0,12	$C_{pki}$	0,01	$C_{pki}$	0,07	AMT
$C_{pks}$	1,25	$C_{pks}$	1,16	$C_{pks}$	1,05	
$C_{pk}$	-0,12	$C_{pk}$	0,01	$C_{pk}$	0,07	
$C_p$	0,56	$C_p$	0,53	$C_p$	0,38	
<b>ppm</b>	656545	<b>ppm</b>	488952	<b>ppm</b>	406494	
$C_{pki}$	2,64	$C_{pki}$	2,63	$C_{pki}$	1,97	pH
$C_{pks}$	0,06	$C_{pks}$	0,1	$C_{pks}$	0,13	
$C_{pk}$	0,06	$C_{pk}$	0,1	$C_{pk}$	0,13	
$C_p$	0,9	$C_p$	0,8	$C_p$	0,75	
<b>ppm</b>	391712	<b>ppm</b>	338510	<b>ppm</b>	298656	
$C_{pki}$	-0,08	$C_{pki}$	-0,04	$C_{pki}$	-0,04	CTT
$C_{pks}$	1,00	$C_{pks}$	1,00	$C_{pks}$	1,00	
$C_{pk}$	-0,08	$C_{pk}$	-0,04	$C_{pk}$	-0,04	
$C_p$	0	$C_p$	0	$C_p$	0	
<b>ppm</b>	992069	<b>ppm</b>	967824	<b>ppm</b>	929507	
$C_{pki}$	-0,10	$C_{pki}$	-0,12	$C_{pki}$	-0,07	DBO <sub>F</sub>
$C_{pks}$	1,02	$C_{pks}$	1,05	$C_{pks}$	1,05	
$C_{pk}$	-0,10	$C_{pk}$	-0,12	$C_{pk}$	-0,07	
$C_p$	0,18	$C_p$	0,25	$C_p$	0,29	
<b>ppm</b>	671385	<b>ppm</b>	681154	<b>ppm</b>	596954	
$C_{pki}$	-0,21	$C_{pki}$	0,03	$C_{pki}$	0,2	DQO
$C_{pks}$	1,15	$C_{pks}$	1,12	$C_{pks}$	1,06	
$C_{pk}$	-0,21	$C_{pk}$	0,03	$C_{pk}$	0,20	
$C_p$	0,39	$C_p$	0,49	$C_p$	0,50	
<b>ppm</b>	733459	<b>ppm</b>	497918	<b>ppm</b>	235245	
$C_{pki}$	-0,08	$C_{pki}$	0,33	$C_{pki}$	0,39	DQO <sub>F</sub>
$C_{pks}$	1,20	$C_{pks}$	1,11	$C_{pks}$	1,04	
$C_{pk}$	-0,08	$C_{pk}$	0,33	$C_{pk}$	0,39	
$C_p$	0,53	$C_p$	0,64	$C_p$	0,57	
<b>ppm</b>	582066	<b>ppm</b>	161506	<b>ppm</b>	99070	
$C_{pki}$	0,39	$C_{pki}$	0,53	$C_{pki}$	0,47	OD
$C_{pks}$	-0,34	$C_{pks}$	-0,23	$C_{pks}$	0,03	
$C_{pk}$	-0,34	$C_{pk}$	-0,23	$C_{pk}$	0,03	
$C_p$	0,31	$C_p$	0,42	$C_p$	0,39	
<b>ppm</b>	49165	<b>ppm</b>	27159	<b>ppm</b>	44100	

Fonte: o autor

**Tabela C2** – Índices de Capabilidade para o desempenho da ETE PS<sub>6</sub>.

<i>LFP</i>		<i>LMP</i>		<i>LMS</i>		<i>Parâmetro</i>
$C_{pki}$	0,32	$C_{pki}$	0,35	$C_{pki}$	0,29	AMT
$C_{pks}$	1,24	$C_{pks}$	1,31	$C_{pks}$	1,53	
$C_{pk}$	0,32	$C_{pk}$	0,35	$C_{pk}$	0,29	
$C_p$	0,78	$C_p$	0,85	$C_p$	1,01	
<b>ppm</b>	193144	<b>ppm</b>	174720	<b>ppm</b>	230614	

$C_{pki}$	2,36	$C_{pki}$	2,29	$C_{pki}$	2,16	pH
$C_{pks}$	3,26	$C_{pks}$	3,31	$C_{pks}$	3,53	
$C_{pk}$	2,36	$C_{pk}$	2,29	$C_{pk}$	2,16	
$C_p$	2,97	$C_p$	2,99	$C_p$	3,1	
<b>ppm</b>	0	<b>ppm</b>	0	<b>ppm</b>	0	
$C_{pki}$	-0,06	$C_{pki}$	-0,04	$C_{pki}$	-0,04	CTT
$C_{pks}$	1,00	$C_{pks}$	1,00	$C_{pks}$	1,00	
$C_{pk}$	-0,06	$C_{pk}$	-0,04	$C_{pk}$	-0,04	
$C_p$	0	$C_p$	0	$C_p$	0,01	
<b>ppm</b>	989761	<b>ppm</b>	946043	<b>ppm</b>	844100	
$C_{pki}$	-0,59	$C_{pki}$	-0,05	$C_{pki}$	-0,05	DBO <sub>F</sub>
$C_{pks}$	1,13	$C_{pks}$	1,03	$C_{pks}$	1,03	
$C_{pk}$	-0,59	$C_{pk}$	-0,05	$C_{pk}$	-0,05	
$C_p$	0,14	$C_p$	0,24	$C_p$	0,26	
<b>ppm</b>	979941	<b>ppm</b>	583911	<b>ppm</b>	571391	
$C_{pki}$	-0,16	$C_{pki}$	-0,11	$C_{pki}$	0,17	DQO
$C_{pks}$	1,17	$C_{pks}$	1,27	$C_{pks}$	1,04	
$C_{pk}$	-0,16	$C_{pk}$	-0,11	$C_{pk}$	0,17	
$C_p$	0,45	$C_p$	0,6	$C_p$	0,44	
<b>ppm</b>	676240	<b>ppm</b>	609823	<b>ppm</b>	278204	
$C_{pki}$	0,01	$C_{pki}$	0,49	$C_{pki}$	0,41	DQO <sub>F</sub>
$C_{pks}$	1,1	$C_{pks}$	1,44	$C_{pks}$	1,09	
$C_{pk}$	0,01	$C_{pk}$	0,49	$C_{pk}$	0,41	
$C_p$	0,45	$C_p$	1,02	$C_p$	0,67	
<b>ppm</b>	481845	<b>ppm</b>	95660	<b>ppm</b>	105119	
$C_{pki}$	-0,17	$C_{pki}$	-0,05	$C_{pki}$	-0,06	OD
$C_{pks}$	1,15	$C_{pks}$	1,03	$C_{pks}$	1,13	
$C_{pk}$	-0,17	$C_{pk}$	-0,05	$C_{pk}$	-0,06	
$C_p$	0,41	$C_p$	0,26	$C_p$	0,44	
<b>ppm</b>	695253	<b>ppm</b>	578823	<b>ppm</b>	572586	

Fonte: o autor

**Tabela C3** – Índices de Capabilidade para a Estação PS<sub>9</sub>.

LFP		LMP		LMS		Parâmetro
$C_{pki}$	-0,56	$C_{pki}$	-0,64	$C_{pki}$	-0,09	AMT
$C_{pks}$	1,68	$C_{pks}$	1,94	$C_{pks}$	1,45	
$C_{pk}$	-0,56	$C_{pk}$	-0,64	$C_{pk}$	-0,09	
$C_p$	0,80	$C_p$	0,99	$C_p$	0,79	
<b>ppm</b>	878186	<b>ppm</b>	895147	<b>ppm</b>	581690	
$C_{pki}$	8,06	$C_{pki}$	3,96	$C_{pki}$	4,15	pH
$C_{pks}$	9,32	$C_{pks}$	5,72	$C_{pks}$	5,92	
$C_{pk}$	8,06	$C_{pk}$	3,96	$C_{pk}$	4,15	
$C_p$	8,96	$C_p$	5,2	$C_p$	5,41	
<b>ppm</b>	0	<b>ppm</b>	0	<b>ppm</b>	0	
$C_{pki}$	-0,07	$C_{pki}$	-0,04	$C_{pki}$	-0,03	CTT
$C_{pks}$	1	$C_{pks}$	1	$C_{pks}$	1	
$C_{pk}$	-0,07	$C_{pk}$	-0,04	$C_{pk}$	-0,03	
$C_p$	0	$C_p$	0	$C_p$	0	

<b>ppm</b>	984817	<b>ppm</b>	945019	<b>ppm</b>	834725	
<b>C<sub>pki</sub></b>	-0,09	<b>C<sub>pki</sub></b>	-0,21	<b>C<sub>pki</sub></b>	-0,04	DBO <sub>F</sub>
<b>C<sub>pks</sub></b>	1,05	<b>C<sub>pks</sub></b>	1,05	<b>C<sub>pks</sub></b>	1,04	
<b>C<sub>pk</sub></b>	-0,09	<b>C<sub>pk</sub></b>	-0,21	<b>C<sub>pk</sub></b>	-0,04	
<b>C<sub>p</sub></b>	0,27	<b>C<sub>p</sub></b>	0,19	<b>C<sub>p</sub></b>	0,29	
<b>ppm</b>	637517	<b>ppm</b>	802997	<b>ppm</b>	565234	
<b>C<sub>pki</sub></b>	-0,13	<b>C<sub>pki</sub></b>	-0,03	<b>C<sub>pki</sub></b>	0,36	DQO
<b>C<sub>pks</sub></b>	1,17	<b>C<sub>pks</sub></b>	1,1	<b>C<sub>pks</sub></b>	1,11	
<b>C<sub>pk</sub></b>	-0,13	<b>C<sub>pk</sub></b>	-0,03	<b>C<sub>pk</sub></b>	0,36	
<b>C<sub>p</sub></b>	0,47	<b>C<sub>p</sub></b>	0,41	<b>C<sub>p</sub></b>	0,67	
<b>ppm</b>	688168	<b>ppm</b>	539311	<b>ppm</b>	143097	
<b>C<sub>pki</sub></b>	0,13	<b>C<sub>pki</sub></b>	0,39	<b>C<sub>pki</sub></b>	0,93	DQO <sub>F</sub>
<b>C<sub>pks</sub></b>	1,09	<b>C<sub>pks</sub></b>	1,07	<b>C<sub>pks</sub></b>	1,13	
<b>C<sub>pk</sub></b>	0,13	<b>C<sub>pk</sub></b>	0,39	<b>C<sub>pk</sub></b>	0,93	
<b>C<sub>p</sub></b>	0,5	<b>C<sub>p</sub></b>	0,64	<b>C<sub>p</sub></b>	1,01	
<b>ppm</b>	339854	<b>ppm</b>	112758	<b>ppm</b>	2787	
<b>C<sub>pki</sub></b>	0,14	<b>C<sub>pki</sub></b>	0,02	<b>C<sub>pki</sub></b>	0,03	OD
<b>C<sub>pks</sub></b>	1,01	<b>C<sub>pks</sub></b>	1,02	<b>C<sub>pks</sub></b>	1,01	
<b>C<sub>pk</sub></b>	0,14	<b>C<sub>pk</sub></b>	0,02	<b>C<sub>pk</sub></b>	0,03	
<b>C<sub>p</sub></b>	0,28	<b>C<sub>p</sub></b>	0,28	<b>C<sub>p</sub></b>	0,21	
<b>ppm</b>	251442	<b>ppm</b>	467825	<b>ppm</b>	438860	

Fonte: o autor

## APÊNDICE D – GARANTIA FUZZY PARA AS ESTAÇÕES PS<sub>5</sub>, PS<sub>6</sub> e PS<sub>9</sub>

**Tabela D1** – Garantia *Fuzzy* para a ETE PS<sub>5</sub> utilizando o padrão de qualidade mostrado pela Tabela 13.

<i>Parâmetro</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
AMT	0,11	0,07	0,83
pH	0,83	0,92	0,94
CTT	0,00	0,00	0,06
DBO <sub>F</sub>	0,07	0,93	0,10
DQO	0,19	0,09	0,24
DQO <sub>F</sub>	0,39	0,43	0,74
OD	0,90	0,83	0,79

Fonte: o autor

**Tabela D2** – Garantia *Fuzzy* na S<sub>6</sub> utilizando o padrão de qualidade mostrado pela Tabela 13.

<i>Parâmetro</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
AMT	0,60	0,66	0,91
pH	0,87	0,90	0,93
CTT	0,00	0,00	0,08
DBO <sub>F</sub>	0,55	0,63	0,54
DQO	0,00	0,00	0,15
DQO <sub>F</sub>	0,35	0,95	0,83
OD	0,30	1,00	1,00

Fonte: o autor

**Tabela D3** – Garantia *Fuzzy* para a ETE PS<sub>9</sub> utilizando o padrão de qualidade mostrado pela Tabela 13.

<i>Parâmetro</i>	<i>LFP</i>	<i>LMP</i>	<i>LMS</i>
AMT	0,0	0,91	0,95
pH	0,85	0,91	0,98
CTT	0,0	0,0	0,1
DBO <sub>F</sub>	0,05	0,43	0,4
DQO	0,01	0,12	0,11
DQO <sub>F</sub>	0,0	0,12	0,0
OD	0,83	0,9	1,0

Fonte: o autor

## APÊNDICE E – SCRIPT USADO PARA GERAR OS RESULTADOS NO SOFTWARE R.

**Apêndice E1** – Script no *software R* referente a Figura 19.

```
library(fitdistrplus)
banco_de_dados<-read.table(file.choose(), header=TRUE)
#teste KS
ajuste_frechet<-fitdist(banco_de_dados, "frechet")
ks.test(banco_de_dados, "pfrechet", parâmetro1, parametro2)
ajuste_gama<- fitdist(banco_de_dados, "gama")
ks.test(banco_de_dados, "gamma", parâmetro1, parametro2)
ajuste_gumbel<- fitdist(banco_de_dados, "gumbel")
ks.test(banco_de_dados, "gumbel", parâmetro1, parametro2)
ajuste_lnorm<- fitdist(banco_de_dados, "lnorm")
ks.test(banco_de_dados, "lnormgumbel", parâmetro1, parametro2)
ajuste_norm<- fitdist(banco_de_dados, "norm")
ks.test(banco_de_dados, "norm", parâmetro1, parametro2)
ajuste_rayleigh<- fitdist(banco_de_dados, "rayleigh")
ks.test(banco_de_dados, "rayleigh", parâmetro1, parametro2)
ajuste_weibull<- fitdist(banco_de_dados, "weibull")
ks.test(banco_de_dados, "weibull", parâmetro1, parametro2)
#AIC
gofstat(list(ajuste_frechet, ajuste_gama, ajuste_gumbel, ajuste_lnorm, ajuste_norm,
ajuste_rayleigh, ajuste_weibull), fitnames=c("frechet", "gamma", "lognormal", "normal",
"rayleigh", "weibull"))
```

**Apêndice E2** – Script no *software R* referente a Figura 20.

```
x1 <- banco_de_dados
f1 <- fitdist(x1, "função_distribuição_escolhida_anteriormente")
b1 <- bootdist(f1, niter=51)
print(b1)
plot(b1)
plot(b1, enhance=TRUE)
summary(b1)
```

```
quantile(b1)
```

```
CIcdfplot(b1, CI.output = "quantile")
```

**Apendice E3** – Script no *software R* referente a Figura 21.

```
x = melhor_função_escolhida_anteriormente(número_de_dados_gerados, parametro1,  
parametro2)
```

```
cp(x, "melhor_função", lsl = limite_inferior, usl = limite_superior)
```