



## II-010 - GRÁFICOS DE CONTROLE DA QUALIDADE DE EFLUENTES DE ESTAÇÕES DE TRATAMENTO DE ESGOTOS

**Sílvia Corrêa Oliveira<sup>(1)</sup>**

Doutora em Saneamento, Meio Ambiente e Recursos Hídricos do Departamento de Engenharia Sanitária e Ambiental (DESA) da UFMG. Professora Adjunta do Departamento de Engenharia Sanitária e Ambiental da UFMG.

**Marcos von Sperling<sup>(1)</sup>**

Doutor em Engenharia Ambiental (Imperial College, Universidade de Londres – Inglaterra). Professor Titular do Departamento de Engenharia Sanitária e Ambiental da UFMG.

**Endereço<sup>(1)</sup>:** Departamento de Engenharia Sanitária e Ambiental – UFMG; Av. do Contorno, nº 842 – 7º andar – Centro – Belo Horizonte – MG – Brasil – CEP 30.110-060– Tel: (31) 3409-1935 – e-mail: silvia@desa.ufmg.br; marcos@desa.ufmg.br

### RESUMO

Neste trabalho é avaliada a utilização de gráficos de controle de qualidade na área de tratamento de esgotos, considerando a assimetria acentuada usualmente apresentada pelas concentrações dos diversos parâmetros monitorados. Para verificar se os gráficos podem ser aplicados a dados assimétricos sem um consequente aumento da probabilidade de erros associados às conclusões, este artigo efetua uma comparação entre gráficos de controle convencionais, que supõem a normalidade dos dados, e gráficos modificados, que levam em conta o comportamento lognormal das variáveis. Assim, alguns dados de concentração efluente de DBO de uma ETE que utiliza o processo de lodos ativados, operando na região metropolitana de Belo Horizonte - MG, foram utilizados para tal proposta. Os resultados mostram que, quando um gráfico de controle convencional é utilizado e a suposição de normalidade dos dados não é válida, os resultados podem levar à conclusão errônea de que o processo está fora de controle. Como consequência, falsos sinais de alerta irão acontecer sempre que dados assimétricos alimentem rotinas de construção de gráficos de controle, uma vez que estes são projetados especificamente para detectar qualquer desvio da normalidade.

**PALAVRAS-CHAVE:** Gráficos de controle de qualidade, distribuição lognormal, efluentes, estações de tratamento de esgotos.

### INTRODUÇÃO

Os gráficos de controle são as ferramentas estatísticas mais largamente utilizadas no controle estatístico do processo para detecção de alterações nas características de qualidade, visando subsidiar ações que objetivem a obtenção de estabilidade do processo e a redução da variabilidade. Tais gráficos foram desenvolvidos para auxiliar na distinção entre variações nos processos industriais que eram intrínsecas ao sistema de produção e variações devidas a fatores externos. No entanto, apesar de muito utilizado em controle de processos industriais, o uso de gráficos de controle na área de tratamento de esgotos é relativamente esparso e pouco difundido, ainda que alguns autores já tenham utilizado o conceito ao longo dos últimos anos (BERTHOUEX *et al.* 1978; BERTHOUEX e HUNTER, 1975, 1981 e 1983; BERTHOUEX e FAN, 1986; VAUGHAN e RUSSEL, 1983; WEST *et al.*, 2002).

Esta técnica de monitoramento do processo possibilita a detecção de fontes não usuais de variabilidade nos sistemas, o que permite que ações corretivas possam ser tomadas para remoção das alterações ou falhas presentes. As alterações observadas podem ter duas principais causas de variação, as aleatórias e as atribuíveis. No contexto dos processos biológicos de tratamento, as primeiras são parte inerente ao processo, tais como variações da carga afluente, das condições ambientais, da natureza do esgoto a ser tratado, enfim, da variabilidade inerente aos processos de tratamento biológicos (BERTHOUEX e HUNTER, 1975). Já as causas atribuíveis são aquelas que surgem de maneira súbita ou anormal e que podem ser identificadas e eliminadas do sistema, tais como falhas mecânicas ou erros de operadores. Em geral, apenas as causas atribuíveis são passíveis de intervenção e quando as variações surgem apenas de causas aleatórias, considera-se que os dados estão estatisticamente controlados.

Os métodos de controle estatísticos de qualidade, em geral, partem de dois pressupostos comumente aceitos: (a) os dados analisados são estatisticamente independentes e (b) os dados são originados de populações normalmente distribuídas. No que concerne à área de tratamento de esgotos, a primeira suposição poderá até ser atendida em alguns casos, dependendo da frequência adotada para coleta das amostras. Berthouex e Hunt (1975) comentam que dados de concentrações efluentes coletados entre intervalos de 2 a 3 dias ou mais podem ser considerados independentes o suficiente para permitir que ETEs que adotem tais frequências de monitoramento utilizem os gráficos de controle com relativa segurança. No entanto, a hipótese de normalidade dos dados pode ser descartada para quase toda a totalidade de casos reportados por autores da área, que concluíram que a distribuição lognormal é a mais representativa dos dados (DEAN e FORSYTHE, 1976a e 1976b; NIKU *et al.*, 1979, 1981a, 1981b, 1981c e 1982; BERTHOUEX e HUNTER, 1981, 1983; METCALF e EDDY, 2003; CHARLES *et al.*, 2005; OLIVEIRA, 2006).

Apesar da comprovada violação dos pressupostos exigidos para utilização dos gráficos de controle convencionais em dados efluentes de ETEs, ainda existem dúvidas sobre o impacto de sua utilização sobre os resultados. Para verificar se eles podem ser aplicados a dados assimétricos sem um conseqüente aumento da probabilidade de erros associados às conclusões, este artigo efetua uma comparação entre gráficos de controle convencionais, que supõem a normalidade dos dados, e gráficos modificados, que levam em conta o comportamento lognormal das variáveis. Assim, alguns dados de concentração efluente de DBO de uma ETE que utiliza o processo de lodos ativados, operando na região metropolitana de Belo Horizonte - MG, foram utilizados para tal proposta.

## MATERIAIS E MÉTODOS

A prática usual para elaboração de gráficos de controle convencionais é extrair amostras de processos computadas em subgrupos de igual tamanho, a intervalos regulares. Estes subgrupos, chamados racionais, devem ser selecionados com cuidado, já que a variabilidade dentro dos subgrupos é utilizada para construir os limites de controle. Assim, cada subgrupo deve consistir de dados homogêneos, ou seja, deve-se garantir que apenas causas aleatórias sejam responsáveis pela variabilidade presente. A partir daí, são calculadas as médias ( $\bar{X}$ ), amplitudes ( $A$ ) e desvios padrão ( $s$ ) de cada subgrupo e estas informações são usadas para construir a linha central (LC) e os limites superior e inferior de controle (LSC e LIC), como exemplificado na Figura 1.

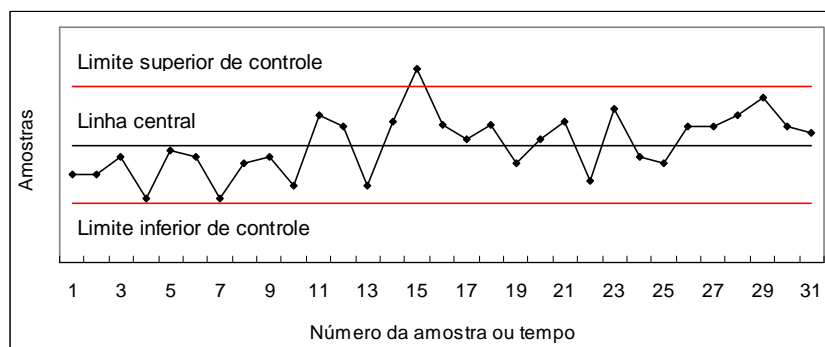


Figura 1: Gráfico de controle típico

A linha central representa o valor médio dos subgrupos de amostras e a suposição de normalidade implica que os limites de controle serão simétricos em torno da LC. Estes limites de controle são determinados a partir de algumas considerações estatísticas simples e, usualmente, são posicionados a 2 ou 3 desvios padrão (2 e 3sigma, respectivamente) em torno da linha central e delimitam regiões que têm probabilidade de 95,4% e 99,7%, respectivamente, de conter qualquer média de cada subgrupo. Assim, para um processo sob controle, praticamente todos os pontos amostrais estarão entre eles. A linha referente a 2sigma pode ser usada como um sinal de alerta, enquanto a de 3sigma pode indicar que alguma ação deve ser tomada para retomada do controle do processo.

Em decorrência do Teorema do Limite Central, a largura dos limites de controle é inversamente proporcional ao tamanho da amostra. Portanto, quanto maior o número de amostras contidas em cada subgrupo, mais próxima será a média deste subgrupo da média de todas as amostras, fazendo com que os limites de controle se aproximem mais da linha central. Este ajuste é efetuado por meio de constantes tabuladas (algumas



desenvolvidas com alto grau de complexidade), que levam em conta o tamanho da amostra contida em cada subgrupo (BURR, 1976, GILBERT, 1987, MONTGOMERY, 2004).

Nos gráficos de controle modificados propostos no presente trabalho foram consideradas as relações entre as distribuições normal e lognormal para cômputo dos parâmetros de controle. Assim, para dados representados pela distribuição lognormal, tem-se a média geométrica ( $X_g$ ) e o desvio-padrão geométrico ( $S_g$ ), e o intervalo  $X_g \times S_g^2$  representa a probabilidade de 95,4% dos valores estarem contidos nesta faixa, enquanto  $X_g \times S_g^3$  compreende 99,7%. A partir desta equivalência, a média geométrica foi proposta para representar a linha central e os limites 2 e 3 sigma foram definidos a partir de uma equação desenvolvida para retratar o estreitamento dos limites de controle em função do aumento do número de dados contidos em cada subgrupo de amostras.

A Tabela 1 apresenta uma comparação entre as metodologias utilizadas para construção dos gráficos de controle convencionais e a proposta efetuada pelo presente estudo.

**Tabela 1:** Fórmulas para gráficos de controle para média, considerando as distintas metodologias

Gráfico	Linha central ou meta	Limites de controle			
		Superior 2 sigma (LS2 $\sigma$ )	Inferior 2 sigma (LI2 $\sigma$ )	Superior 3 sigma (LS3 $\sigma$ )	Inferior 3 sigma (LI3 $\sigma$ )
Convencional	$\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$	$\bar{x} + \frac{2\bar{A}_1}{\sqrt{n}}$	$\bar{x} - \frac{2\bar{A}_1}{\sqrt{n}}$	$\bar{x} + \frac{3\bar{A}_1}{\sqrt{n}}$	$\bar{x} - \frac{3\bar{A}_1}{\sqrt{n}}$
Modificado	$X_g = \sqrt[n]{\prod_{i=1}^n x_i}$	$X_g \times S_g \left( \frac{2}{\sqrt{n}} \right)$	$X_g \div S_g \left( \frac{2}{\sqrt{n}} \right)$	$X_g \times S_g \left( \frac{3}{\sqrt{n}} \right)$	$X_g \div S_g \left( \frac{3}{\sqrt{n}} \right)$

Nota:  $\bar{X}$  = média aritmética;  $X_g$  = média geométrica;

$S_g$  = desvio padrão geométrico:

$$S_g = \exp \left( \frac{\sum_{i=1}^n (\ln X_i - \ln \bar{X}_g)^2}{n} \right)^{1/2}$$

e  $\bar{A}_1$  = amplitude média dos  $k$  subgrupos:

$$\bar{A}_1 = \frac{1}{k} \sum_{i=1}^k \frac{A_i}{d_{2i}}$$

onde  $d_{2i}$  corresponde a constantes tabeladas em função do número de dados dos subgrupos

## RESULTADOS E DISCUSSÃO

Os dados utilizados foram selecionados de uma série histórica de monitoramento de concentrações efluentes de DBO de uma estação de tratamento de esgotos operando na região metropolitana de Belo Horizonte - MG. Houve necessidade de selecionar um período de dados em virtude da falta de monitoramento em algumas épocas, o que dificultaria a demonstração da metodologia desenvolvida. Na prática, o ideal será sempre o desenvolvimento de limites de controle a partir de séries completas de dados, pois o procedimento para subgrupos de tamanhos diferentes é bem mais complexo. Os dados utilizados foram coletados com uma frequência quinzenal de monitoramento, no período de janeiro/1995 a dezembro/2002, e foram disponibilizados pela companhia de saneamento responsável pela operação da ETE.

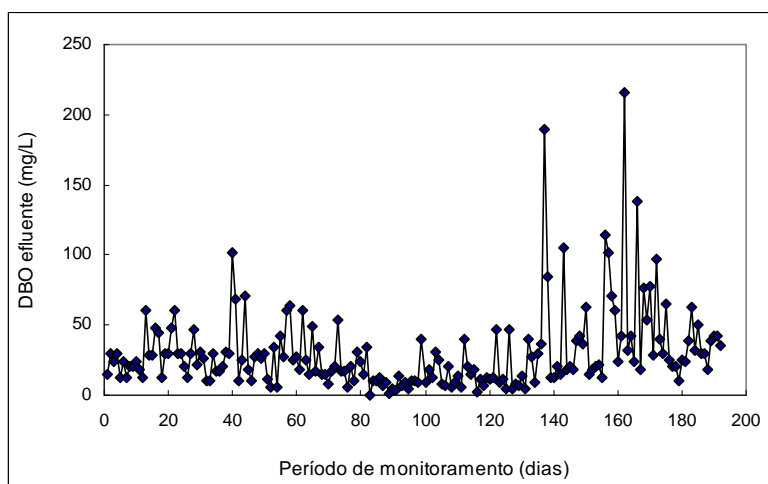


A Tabela 2 apresenta estatísticas descritivas referentes à concentração de DBO efluente, que já permitem uma conclusão preliminar sobre a distribuição dos dados, ficando evidente a assimetria presente. Na distribuição lognormal, os valores da média aritmética são sempre maiores que os apresentados pela média geométrica, o que não acontece na distribuição normal, cujos valores são muito próximos. O coeficiente de assimetria é a medida de desvio ou afastamento da simetria, apresentando valores próximos de zero quando a distribuição tem uma curva simétrica e valores positivos quando existe assimetria à direita. O coeficiente de curtose mede a curvatura ou o grau de achatamento de uma distribuição, apresentando valores iguais a três para uma distribuição normal. Outra consideração importante é a grande influência do coeficiente de variação (desvio padrão dividido pela média aritmética), na forma da distribuição. Ott (1995) comenta que, para valores de CV menores que um sexto (0,16667), a distribuição lognormal apresenta um comportamento muito próximo da distribuição normal, enquanto que elevados valores de CV são associados com expressivas assimetrias.

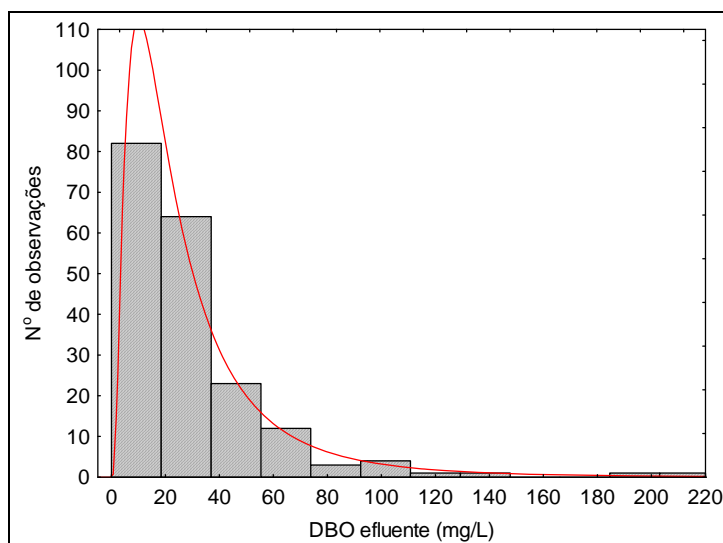
**Tabela 2:** Estatística descritiva referente à concentração de DBO efluente da ETE em estudo

Número dados	192	Média geométrica (mg/L)	21
Valor mínimo (mg/L)	1	Desvio padrão geométrico (mg/L)	13
Valor máximo (mg/L)	216	Mediana (mg/L)	22
Amplitude	215	Percentil 10% (mg/L)	8
Média aritmética (mg/L)	30	Percentil 90% (mg/L)	60
Desvio padrão (mg/L)	29	Coeficiente de assimetria	3
Coeficiente variação	1,0	Coeficiente de curtose	14

A Figura 2 mostra a série de dados de DBO efluente obtidos no período, consistindo de 192 dados, sendo que os primeiros 96 valores foram utilizados para desenvolvimento dos limites de controle. A Figura 3 apresenta o histograma de frequência relativo aos dados analisados, onde fica evidente a assimetria apresentada e o bom ajuste à distribuição lognormal (curva em vermelho no gráfico) dos valores de concentração de DBO efluente.



**Figura 2:** Dados quinzenais de DBO efluente no período de Jan/95 a Dez./02



**Figura 3:** Histograma da concentração efluente de DBO (Média aritmética de 30 mg/L e desvio padrão de 29 mg/L)

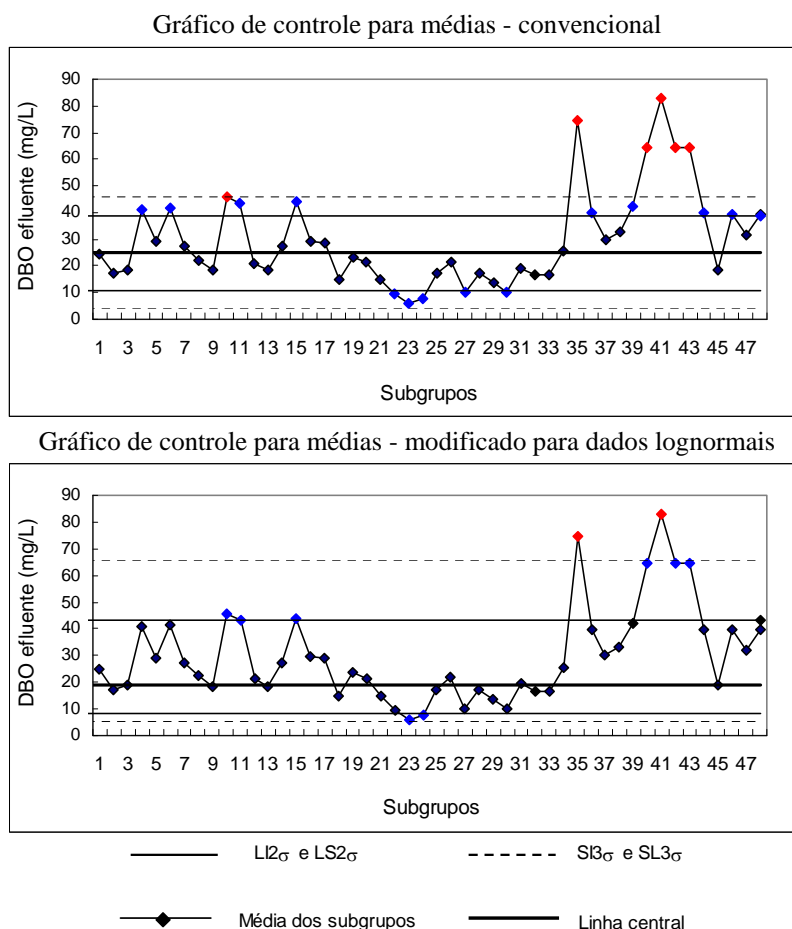
Os dados coletados quinzenalmente foram dispostos em subgrupos bimensais. Considerou-se que as quatro observações quinzenais tomadas dentro do período de dois meses formavam subgrupos racionais e facilmente aplicáveis na prática. A título de ilustração, foram utilizados os 96 dados relativos aos três primeiros anos de coleta (24 subgrupos), enquanto os dados restantes foram utilizados para avaliação da adequação dos parâmetros de controle determinados a partir dos subgrupos selecionados.

A Tabela 3 apresenta os resultados dos parâmetros calculados a partir das equações constantes na Tabela 1 para as linhas centrais e limites de controle, considerando as duas metodologias, convencional e a utilizada para dados lognormais. A Figura 4 mostra os gráficos de controle de médias correspondentes.

**Tabela 3:** Valores calculados para gráfico de controle para médias, considerando distintas metodologias

Parâmetros		Convencional	Modificado
Linha central ou meta		25	19
<b>Limites de controle</b>	Superior 2sigma (LS2 $\sigma$ )	11	8
	Inferior 2sigma (LI2 $\sigma$ )	39	43
	Superior 3sigma (LS3 $\sigma$ )	4	6
	Inferior 3sigma (LI3 $\sigma$ )	43	66

Pode ser verificado que, no gráfico de controle modificado para dados lognormais, foram definidos limites de controle assimétricos em relação à linha central, em virtude da assimetria natural apresentada pelos dados. A utilização do gráfico de controle convencional, que estabeleceu limites de controle simétricos para dados lognormais, aumentou a probabilidade de ocorrência de erro tipo I, ou seja, sinalizou que o processo estava fora de controle, quando na verdade não estava. Isto pode ser observado pela presença de 20 pontos (em azul no gráfico) fora do limite 2sigma e seis valores acima do limite superior 3sigma. No entanto, quando se utiliza o gráfico modificado em função da distribuição efetiva dos dados, o número de valores que ultrapassa os limites 2 e 3sigma cai para 10 e dois, respectivamente. Isto mostra que a utilização de gráficos convencionais levaria à emissão de sinais de alerta em 10 situações que estariam sob controle estatístico do processo, já que os valores de concentrações efluentes de DBO não estariam apresentando uma variabilidade não usual. De forma semelhante e ainda mais importante, o gráfico de controle convencional indicaria a necessidade de intervenção no processo em quatro situações, onde apenas sinais de alerta deveriam ser emitidos.



**Figura 4:** Comparação entre os gráficos de controle convencional e modificado para dados lognormais

Nota: valores em azul correspondem aos dados que extrapolaram os limites de controle 2 sigma e valores em vermelho extrapolaram os limites 3 sigma

## CONCLUSÕES

Quando um gráfico de controle convencional é utilizado e a suposição de normalidade dos dados não é válida, os resultados podem levar à conclusão errônea de que o processo está fora de controle. No entanto, como verificado no exemplo comparativo apresentado, distribuições lognormais podem ocorrer naturalmente em processos de tratamento de esgotos e isto acarretará a presença de valores elevados de concentrações efluentes, não significando, necessariamente, que alguma ação corretiva deva ser tomada. Como consequência, falsos sinais de alerta irão acontecer sempre que dados assimétricos alimentem rotinas de construção de gráficos de controle, uma vez que estes são projetados especificamente para detectar qualquer desvio da normalidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. BERTHOUEX, P.M., FAN, R. Evaluation of treatment plant performance: causes, frequency, and duration of upsets. *Journal of Water Pollution Control Federation*, v. 58, n. 5, p.368 – 375, 1986.
2. BERTHOUEX, P.M., HUNTER, W.G. How to construct reference distributions to evaluate treatment plant effluent quality. *Journal of Water Pollution Control Federation*, v. 55, n. 12, p. 1417-1424, 1983.
3. BERTHOUEX, P.M., HUNTER, W.G. Simple statistics for interpreting environmental data. *Journal of Water Pollution Control Federation*, v. 53, n. 2, p. 167-175, 1981.
4. BERTHOUEX, P.M., HUNTER, W.G. Treatment plant monitoring programs: A preliminary analysis. *Journal of Water Pollution Control Federation*, v. 47, n. 8, p. 2143-2156, 1975.
5. BERTHOUEX, P.M., HUNTER, W.G.; PALLESEN, L. Monitoring sewage treatment plants: some quality control aspects. *Journal of Quality Technology*, v. 10, n. 4, p. 139-149, 1978.





6. BURR, I. W. *Statistical Quality Control Methods*. Marcel Dekker, Inc., New York, 1976, v. 16, 522p.
7. CHARLES, K.J.; ASHBOLT, N.J.; ROSER, D.J.; MCGUINNESS, R.; DEERE, D.A. Effluent quality from 200 on-site sewage systems: design values for guidelines. *Water Science & Technology*, v. 51, n. 10, pp 163-169, 2005.
8. DEAN, R.B., FORSYTHE, S.L. Estimating the reliability of advanced waste treatment. Part 1, *Water & Sewage Works*, 1976a, 87 – 89.
9. GILBERT, R. O. *Statistical methods for environmental pollution monitoring*. New York: John Wiley & Sons, Inc., 1987, 320p.
10. METCALF & EDDY. *Wastewater engineering: treatment, and reuse*. New York: Metcalf & Eddy, Inc., 4 th. Ed., 2003, 1819p.
11. MONTGOMERY, D. G. *Introdução ao Controle Estatístico da Qualidade*. Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A., 1ª ed., 2004, 513 p.
12. NIKU, S., SCHROEDER, E.D. Factors affecting effluent variability from activated sludge processes. *Journal Water Pollution Control Association*, v. 53, n. 5, p. 546 - 559, May, 1981a.
13. NIKU, S., SCHROEDER, E.D., HAUGH, R.S. Reliability and stability of trickling filter processes. *Journal Water Pollution Control Association*, v. 54, n. 2, p. 129 - 134, Feb. 1982.
14. NIKU, S., SCHROEDER, E.D., SAMANIEGO F.J. Performance of activated sludge process and reliability-based design. *Journal Water Pollution Control Association*, v. 51, n. 12, p. 2841 - 2857, Dec., 1979.
15. NIKU, S., SCHROEDER, E.D., TCHOBANOGLOUS, G., SAMANIEGO F.J. Performance of activated sludge process: reliability, stability and variability. Environmental Protection Agency, EPA Grant N° R805097-01, pp. 1 – 124, 1981b.
16. NIKU, S., SCHROEDER, E.D., TCHOBANOGLOUS, G., SAMANIEGO F.J. Performance of activated sludge process: reliability, stability and variability. Environmental Protection Agency, EPA – 600/S2 – 81 – 227., pp. 1 – 11, 1981c.
17. OLIVEIRA, S. M. A. C. *Análise de desempenho e confiabilidade de estações de tratamento de esgotos*. Tese (Doutorado em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 231 f. 2006.
18. OTT, W. R. *Environmental statistics and data analysis*. New York: Lewis Publishers, 1995, 313p.
19. VAUGHAN, W.J. and RUSSELL, C.S. Monitoring point sources of pollution: answers and more questions from statistical quality control. *The American Statistician*, v. 37, Issue 4, Part 2: Proceedings of the Seventh Symposium on Statistics and the Environment, p. 476 – 487. Nov., 1983.
20. WEST, D.; DELLANA, S.; JARRETT, J. Transfer function modeling of processes with dynamic inputs. *Journal of Quality Technology*; 34, 3; ABI/INFORM Global, pg. 315; Jul 2002.