

# CONTROLE ESTATÍSTICO DE PROCESSOS APLICADO AO MONITORAMENTO DO PH E TURBIDEZ DAS ÁGUAS NO ABASTECIMENTO DE CAMPINA GRANDE – PB

David Venancio da CRUZ<sup>1</sup>

Moacyr Cunha FILHO<sup>1</sup>

Ana Patrícia Siqueira Tavares FALCÃO<sup>2</sup>

- **RESUMO:** Os ciclos hidrológicos têm sido alterados de maneira significativa em função de atividades antrópicas. Como consequência, tem-se observado expressiva queda na qualidade das águas superficiais para o consumo humano, o que prejudica a sobrevivência e a qualidade de vida da população. Para conseguir uma avaliação de forma eficiente e rápida dos parâmetros que classificam a água foram utilizadas técnicas de controle estatístico do processo. Este trabalho surgiu da necessidade de monitoramento contínuo de alguns parâmetros relacionados a água fornecida no abastecimento da cidade de Campina Grande. Que devem estar sob legislação do Ministério da Saúde, Portaria Nº 2.914/2011, que estabelece procedimentos de controle e de vigilância da água para consumo humano. Assim, empregou modelos de Series temporais para adquirir uma estrutura sem a presença de autocorrelação, posteriormente aplicando gráficos de controle individuais de Shewhart, média móvel exponencialmente ponderada (EWMA), soma acumulativa (CUSUM). Evitando a ocorrência de falsos alarmes é importante salientar que o gráfico de EWMA apresentou melhor desempenho, no monitoramento das águas, devido a sua eficiência, pois é sensível as pequenas variações na redução ou aumento dos níveis do pH e turbidez. As análises foram realizadas por meio do *software R* com auxílio do pacote *forecast*.
- **PALAVRAS-CHAVE:** controle estatístico; CUSUM; EWMA; qualidade da água.

## 1 Introdução

Com o aumento das atividades econômicas e crescimento populacional numa escala mundial, a água de qualidade tem se tornado um recurso natural escasso. Embora o Brasil apresente uma situação confortável, pois detém aproximadamente 11,6% de toda água doce do mundo, existem problemas de distribuição desse recurso, pois o mesmo não se encontra distribuído de forma homogênea: na Amazônia estão 70% desse recurso, enquanto apenas 3% estão disponíveis no Nordeste brasileiro (BICUDO et al. 2010).

---

<sup>1</sup> Universidade Federal Rural de Pernambuco, UFRPE, Programa de Pós-Graduação em Biometria e Estatística Aplicada, CEP: 52171-900, Recife, Pernambuco, Brasil. E-mail: [davidvenacio@hotmail.com](mailto:davidvenacio@hotmail.com); [moacyr@deinfo.ufrpe.br](mailto:moacyr@deinfo.ufrpe.br)

<sup>2</sup> Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco, IFPE, Pró-Reitora de Pesquisa e Inovação, CEP: 55602-970, Vitória de Santo Antão, Pernambuco. Brasil, E-mail: [apstfalcao@hotmail.com](mailto:apstfalcao@hotmail.com)

Segundo Fay e Silva (2006), no semiárido brasileiro o maior problema é a má distribuição espaço-temporal das águas superficiais para o consumo humano no meio rural, o que prejudica a sobrevivência e a qualidade de vida da população. Os maiores potenciais hídricos disponíveis se encontram nos rios perenes e perenizados, lagos, poços tubulares e nas barragens, das quais a maioria da água armazenada é destinada a irrigação.

A água comporta vários componentes de origem ambiental ou proveniente da atividade humana. A qualidade da água é avaliada a partir de suas características físicas, químicas e biológicas. Estas características são consideradas impurezas quando os valores obtidos são superiores aos valores estabelecidos para determinada utilização (HENNING et al., 2011).

De acordo com Follador (2010), para analisar de forma rápida e eficiente os dados que indicam a qualidade da água, objeto deste estudo, podem ser praticadas técnicas de controle estatístico do processo. Estas técnicas colaboram de forma eficaz para o monitoramento ambiental. São diversos os tipos de análise estatística que podem se utilizar, dentre elas destacam-se ferramentas, como os gráficos de controle para medidas individuais: Shewhart, gráfico de CUSUM, gráfico de EWMA, devido a sua simplicidade operacional e realista com relação à detecção de problemas, resultando em um bom desempenho no monitoramento da água.

Neste contexto, aplicam-se os conceitos de controle estatístico de processos (CEP) na análise dos níveis de pH e turbidez na água fornecida na cidade de Campina Grande. Para atender as suposições de normalidade e independência, procede a estimação do modelo auto-regressivo integrado de média móvel (ARIMA). Posteriormente os gráficos de Shewhart, Soma Acumulativa (CUSUM) e gráfico de controle Média Móvel Exponencialmente Ponderada (EWMA) são aplicados e analisados, com objetivo de preposição para o monitoramento contínuo do pH e Turbidez.

O monitoramento do pH é a turbidez da água potável é justificável pela preocupação atual com o saneamento básico e o uso racional deste recurso. Devido a se tratar da qualidade da água consumida pela população, requer uma necessidade imediata de intervenção quando há uma extrapolação dos valores permitidos pela legislação do Ministério da Saúde, Portaria Nº 2.914/2011, para que a água não seja distribuída em más condições. Neste caso, a utilização de gráficos de CUSUM e EWMA poderá ser uma melhor opção, devido a sensibilidade a pequenas variações na média de um processo, que são absorvidas pelo gráfico de Shewhart.

Finalmente, este artigo encontra-se estruturado cinco seções. A segunda seção traz uma breve revisão de literatura sobre controle estatístico de processos (CEP). A terceira seção mostra conceitos de normalidade e autocorrelação. A quarta seção traz a descrição da área de estudo e dos dados analisados. A quinta seção mostra os resultados e análise, em seguida as considerações finais e recomendações para trabalhos futuros.

## **2 Material e métodos**

### **2.1 Descrição da área de estudo e dos dados analisados**

O objeto desse trabalho foi desenvolvido no sistema de abastecimento de Campina Grande, cidade paraibana de aproximadamente 400 mil habitantes (Figura 1), um dos mais importantes centros urbanos e renomado núcleo de tecnologia do interior do

Nordeste. Esse sistema de abastecimento urbano é administrado pela CAGEPA - Companhia de Águas e Esgotos do Estado da Paraíba, empresa pública estadual.

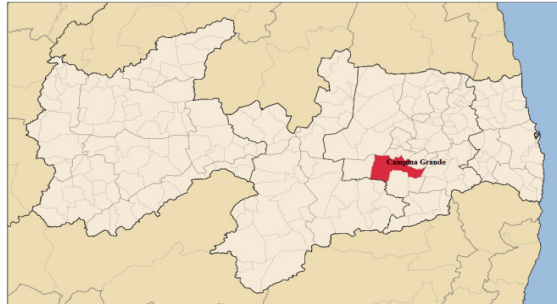


Figura 1 - Localização da Cidade de Campina Grande-PB.

O reservatório localizado no bairro de Santa Rosa em Campina Grande tem capacidade de armazenamento de água tratada de  $40.000^3$ . É o principal centro de reserva e distribuição de água tratada para os demais reservatórios existentes em todo o sistema de abastecimento de Campina Grande, num total de 29 reservatórios elevados e semienterrados, que vão abastecer uma malha de 450 km de rede de distribuição de água com um total de 83.500 ligações domiciliares e uma demanda média de  $60890.4 \text{ m}^3/\text{dia}$  (ALBUQUERQUE, 2004).

A avaliação da água no abastecimento da cidade de Campina Grande foi realizada por meio de uma pesquisa bibliográfica, com os dados disponibilizados pela CAGEPA no site <<http://www.cagepa.pb.gov.br>>, no período de janeiro de 2011 a dezembro de 2013 perfazendo 36 amostras. A companhia - CAGEPA possui para controle da qualidade da água distribuída, Pontos de Controle de Qualidade – P.C.Q's, que são pontos distribuídos no residencial, principalmente em pontos finais de rede e têm como objetivo de favorecer a coleta de amostra de água, para o controle de qualidade da água na rede de abastecimento. Também possibilita a imediata detecção de quaisquer problemas existentes. Nestes pontos são avaliados pH e turbidez. Mensalmente é escolhido um ponto da rede para realização de análises de potabilidade por laboratório externo. As análises foram realizadas por meio do *software* R3.0.2 (<http://www.r-project.org/>) com auxílio do pacote *forecast*.

A turbidez é dada pelo grau de interferência com a passagem da luz através da água apresentando uma aparência turva. A origem dela são os sólidos em suspensão, que podem ser de origem natural, isto é, partículas de rocha, areia e silte, além de algas e outros minerais ou ainda de origem antropogênica, ou seja, despejos industriais, domésticos, microrganismos e erosão. A turbidez de circunstância natural não traz inconvenientes sanitários diretos. Entretanto, é esteticamente desagradável na água potável, servindo de abrigo para microrganismos patogênicos. Nos corpos d'água a turbidez reduz a penetração da luz reduzindo assim a fotossíntese (ARAÚJO, 2010).

O pH é uma característica importante devido a sua influência nos processos biológicos que ocorrem no meio aquático. As águas naturais apresentam um pH entre 4 e 9.5, estando associada pela dissolução de  $\text{CO}_2$  que origina baixos valores de pH, enquanto

que as reações de  $\text{HCO}_3^-$  e  $\text{CO}_3^{2-}$  com água, acarreta em maiores pH. Quando o pH aproxima-se de 9 ocorrem absorção de gás carbônico das águas no processo da fotossíntese (ZUIN, LORIATTI, MATHEUS, 2009).

No Brasil, os padrões de potabilidade foram definidos pelo Ministério da Saúde, através de portaria nº 2.914 de 12 de dezembro de 2011. Esta portaria estabeleceu os procedimentos e responsabilidades relativos ao controle e vigilância da qualidade da água para consumo humano e seu padrão de potabilidade.

## 2.2 Controle estatístico de qualidade

O controle estatístico de qualidade compõe-se de um conjunto de técnicas que permitem a inspeção, identificação, monitoramento, intervenção e aprimoramento das características que refletem a qualidade na produção de bens e serviços. A sua aplicação proporciona uma melhoria contínua em diversos empreendimentos colaborando para que as informações sobre o sistema, ou seja, o estado de equilíbrio e a capacidade de gerar itens possam atender as especificações propostas (FERREIA, 2010). O monitoramento é uma forma de avaliar a qualidade da água e prevê um levantamento de dados em pontos de amostragem, de modo a acompanhar a evolução das condições da qualidade de água ao longo do tempo (FOLLADOR et al., 2012).

O controle estatístico de qualidade é operacionalizado por meio de gráficos de controle que são utilizados para monitorar o desempenho de um processo a partir da definição de uma área de controle aceitável. Os gráficos de controle de qualidade são elementos visuais para o monitoramento de processos, pois conseguem detectar rapidamente alterações desproporcionais e causas especiais, sinalizando a necessidade de intervir no processo (ARAÚJO, 2010).

De acordo com Montgomery (2004) são dois tipos de causas: causas especiais e causas comuns. Causas especiais são eventos que ocorrem casualmente, em períodos irregulares, e que devem ser eliminadas e reduzidas. As causas comuns são relativamente pequenas, em sua maioria alteram a variabilidade do processo, apesar de sua ocorrência a variação é inerente ao processo e são praticamente inevitáveis.

Diversas características de qualidade podem ser mensuradas. Como uma característica de qualidade é uma variável, que em geral se faz necessário monitorar tanto o valor médio da característica de qualidade como sua variabilidade, para isso, gráficos de controle são amplamente utilizados MONTGOMERY (2004).

## 2.3 Gráfico de controle do tipo Shewhart para medidas individuais

O gráfico de controle, numa visão geral, é composto por *LIC* e *LSC*, limites inferior e superior de controle, respectivamente, determinados estatisticamente, a partir da média, e as amplitudes das amostras são função da variabilidade do processo. Em seguida, as equações 1, 2 e 3, que mostram o limite superior, linha média e limite inferior de controle (HENNIG, 2011):

$$LSC = \bar{x} + 3 \frac{\overline{MR}}{d_2}, \quad (1)$$

$$LC = \bar{x}, \quad (2)$$

$$LIC = \bar{x} - 3 \frac{MR}{d_2}, \quad (3)$$

em que:  $LIC$  - linha média;  $\bar{x}$  - média;  $LIC$  - limite inferior de controle;  $MR$  - amplitude da amostra;  $d_2$  - fator de construção de gráficos de controle para variáveis encontrado em Montgomery (2004) e  $LSC$  - limite superior de controle.

Em muitas aplicações dos gráficos para unidades individuais usa-se a amplitude móvel de duas observações consecutivas como base para estimar a variabilidade do processo. A amplitude móvel é definida segundo a equação 4.

$$MR_i = |x_i - x_{i-1}|. \quad (4)$$

Os gráficos de controle de Shewhart fornecem uma análise simples dos dados, sendo de fácil adaptação e integração com práticas de controle já existentes nos laboratórios (SILVA, 2006). Nestes gráficos há baixa probabilidade de falsos alarmes e melhor capacidade de identificação de erros analíticos, por esse motivo são utilizados para uma grande gama de processos (FOLLADOR, 2010).

#### 2.4 Gráfico de controle de soma acumulativa (CUSUM)

Os gráficos de controle de soma acumulativa (CUSUM) são alternativas viáveis aos gráficos de controle de Shewhart. Estes gráficos guardam informações acumuladas das amostras de um processo ponderando-as de forma equivalente, fazendo com que as amostras tenham o mesmo peso. Por esse motivo, esses gráficos se tornam sensíveis para detectar pequenas mudanças na média de um processo (LU, 2001).

A técnica de soma acumulada pode ser aplicada na construção do gráfico CUSUM para observações individuais. Neste caso, a estatística utilizada é a soma acumulada dos desvios de cada valor individual com relação à medida dada pela hipótese que está sendo testada. Esta prática estabelece a hipótese que irá se testar, ou seja, conhecer-se o valor nominal,  $\mu_0$ . Conhecido este desvio, a soma acumulada é iniciada. A soma acumulada  $C_i$ , para o  $i$ -ésimo período é a soma de todos os desvios do valor nominal desde o período 1 até o período  $i$ , dada pela equação 5:

$$C_i = \sum_{j=1}^i (x_j - \mu_0) = (x_i - \mu_0) + C_{i-1}, \quad (5)$$

em que  $x_j$  é a  $j$ -ésima observação individual. O método da soma acumulada pode ser representado de duas maneiras pelo CUSUM Tabular ou pela Máscara V.

O gráfico de CUSUM Tabular é um plano de controle de decisão intervalar bilateral simétrico que utiliza duas estatísticas unilaterais  $C_i^+$  (plano superior) equação 6, para detectar mudanças positivas e  $C_i^-$  (plano inferior) equação 7, para detectar mudanças negativas:

$$C_1^+ = \max[0, x_1 - (\mu_0 + K) + C_{i-1}^+], \quad (6)$$

$$C_1^- = \max[0, (\mu_0 - K) - x_1 + C_{i-1}^-], \quad (7)$$

em que  $K$  é um valor de referência (valor de tolerância ou folga) é aproximadamente a metade do valor que se tem interesse em detectar rapidamente determinado valor entre o valor pretendido  $\mu_0$  e o valor da média fora do controle  $\mu_1$ .

Esta mudança é expressa em unidades de desvio padrão como  $\mu_1 = \mu_0 + \delta\sigma$  logo  $K$  é metade da magnitude dessa mudança segundo a equação 8:

$$K = \frac{\delta}{2}\sigma = \frac{|\mu_1 - \mu_0|}{2}, \quad (8)$$

em que  $\delta$  é o valor da mudança que se deseja detectar em unidades de desvios padrão,  $\sigma$  é o desvio padrão,  $\mu_0$  o valor pretendido e  $\mu_1$  o valor da média fora de controle (HAWKINS, 1993). Quanto menor for este fator, menor será a faixa de variação que o gráfico será capaz de detectar e maior será a sensibilidade do gráfico. Em geral, usa-se  $K = 1/2$  (WALTER et al., 2013)

O sistema estará sob controle se os valores plotados de  $C_i^+$  e  $C_i^-$  estiverem no intervalo de decisão  $H$ . Segundo Montgomery (2004), um valor aconselhável para  $H$  é quatro ou cinco vezes o valor do desvio padrão  $\sigma$ , isto é,  $H = 4\sigma$  ou  $H = 5\sigma$ .

## 2.5 Gráfico de controle média móvel exponencialmente ponderada (EWMA)

Segundo Montgomery (2004), uma alternativa para o gráfico de Shewhart quando está interessado em detectar pequenas mudanças é o gráfico de controle da média exponencial ponderada. O gráfico de controle EWMA apresenta desempenho semelhante ao do gráfico de controle somas acumuladas, e de certa forma, mais fácil de estabelecer e efetuar podendo ser utilizado com amostras de tamanho unitário. Neste caso, o gráfico de controle EWMA para medidas individuais é definido conforme equação 9,

$$z_i = \lambda x_i + (1 - \lambda)z_{i-1}, \quad (9)$$

em que  $z_i$  são valores ponderados da observação  $i$  e  $x_i$  é o  $i$ -ésimo valor observado, o parâmetro  $\lambda$  é uma constante ( $0 < \lambda \leq 1$ ) e o valor inicial  $z_i$  é o valor alvo da média do processo, de modo que,  $z_0 = \mu_0$ .

Ainda a demonstração da expressão 9, constitui uma média ponderada, as amostras pode ser obtida desenvolvendo recursivamente a equação 9. Substituindo o valor de  $z_{i-1}$ , tem-se a equação 10.

$$z_i = \lambda x_i + (1 - \lambda)[\lambda x_{i-1} + (1 - \lambda)z_2], \quad (10)$$

ou desta forma:

$$z_i = \lambda x_i + (1 - \lambda)\lambda x_{i-1} + (1 - \lambda)^2 z_{i-2}, \quad (11)$$

Substituindo  $z_{i-2}$  pelo valor dado pela seguinte equação 9, tem-se que:

$$z_i = \lambda x_i + (1 - \lambda)\lambda x_{i-1} + (1 - \lambda)^2 [\lambda x_{i-2} + (1 - \lambda)z_{i-3}] \quad (12)$$

A expressão 12 aponta a característica de qualidade real  $x_i$  que está presente em todas as expressões de característica da qualidade prevista para um período qualquer  $i$ , indicando que para as amostras mais remotas, a característica de qualidade real

exponencialmente será menor o peso com que aparece na expressão  $z_i$ . Caso seja desenvolvida toda expressão, apenas o valor da característica prevista  $z_0$ , que foi adotado, irá aparecer ao lado de todas as características da qualidade real. De acordo com a equação 13.

$$z_i = \lambda x_i + (1 - \lambda)\lambda x_{i-1} + (1 - \lambda)^2 \lambda x_{i-2} + (1 - \lambda)^3 \lambda x_{i-3} + \dots + (1 - \lambda)^{i-1} \lambda x_0 + (1 - \lambda)^i z_0. \quad (13)$$

De forma geral, ou seja, generalizando a expressão 13 supracitada, obtém-se a seguinte expressão 14:

$$z_i = (1 - \lambda)^i z_0 + \lambda \sum_{j=0}^{i-1} (1 - \lambda) x_{i-j}. \quad (14)$$

Como a média móvel exponencial ponderada (EWMA) considera como uma média ponderada todas as amostras ( $n = 1$ ) remotas e correntes, o gráfico EWMA é insensível à hipótese de normalidade, ou seja, torna o gráfico ideal para ser usado em medidas individuais (SILVA, 2004).

Se as amostras são variáveis aleatórias independentes com variância  $\sigma^2$  então a variância de  $z_i$ , de acordo com a equação 15.

$$\sigma_{z_i}^2 = \sigma^2 \left( \frac{\lambda}{2-\lambda} \right) [1 - (1 - \lambda)^{2i}]. \quad (15)$$

O gráfico EWMA pode ser construído pela plotagem de  $z_i$  versus o número de amostras  $i$ . A linha central e os limites de controle para o gráfico de controle EWMA são apresentados pelas equações 16, 17 e 18.

$$LSC_s = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2-\lambda)}} [1 - (1 - \lambda)^{2i}], \quad (16)$$

$$LC = \mu_0, \quad (17)$$

$$LSC_i = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2-\lambda)}} [1 - (1 - \lambda)^{2i}]. \quad (18)$$

Nas equações 16 e 18,  $L$  representa a largura ou múltiplos de sigma utilizado nos limites de controle, sendo usualmente  $L = 3$ , valores de  $\lambda$ , no intervalo  $0.05 \leq \lambda \leq 0.25$ , funcionam bem na prática, sendo o mais recomendável  $\lambda = 0.2$  (MONTGOMERY, 2004).

## 2.6 Normalidade e autocorrelação

Para a utilização de qualquer tipo de gráfico de controle duas suposições devem ser atendidas. A primeira suposição é que os dados sejam normalmente distribuídos. A segunda é a independência das amostras não apresentem nenhum tipo de correlação.

### 2.6.1 Teste de Kolmogorov-Smirnov

O objetivo desse teste é verificar o grau de concordância entre distribuição de um conjunto de valores e uma distribuição teórica, por exemplo, verificar se os dados seguem uma distribuição normal. O teste de Kolmogorov-Smirnov admite que a distribuição da variável que está sendo testada seja contínua. A distribuição teórica representa o que seria esperado sob  $H_0$ . Então se verifica, as distribuições teórica e observada mostram concordância (SCUDINO, 2008).

Seja  $F_0$  uma função especificada de distribuição de frequências relativas acumuladas a distribuição teórica sob  $H_0$ . Para qualquer valor de  $X$ , o valor de  $F_0(X)$  é a proporção de casos esperados com escores menores ou iguais a  $X$ .

Seja a distribuição de frequências relativas acumuladas observadas,  $S_n$ , de uma amostra aleatória de  $N$  observações. Seja um escore qualquer possível,  $X_i$ , então  $S_n(X_i) = \frac{F_i}{N}$ , onde  $F_i$  é o número de observações menores ou iguais a  $X_i$ . As hipóteses do teste são descritas como:

$H_0$ : A amostra é proveniente de uma distribuição teórica específica;

$H_1$ : A amostra não é proveniente de uma distribuição teórica específica.

Quando  $H_0$  é verdadeira, espera-se que as diferenças entre  $S_n(X_i)$  e  $F_0(X_i)$  sejam pequenas e estejam dentro do limite dos erros aleatórios. A estatística do teste, expressa na equação 19, focaliza o maior dos desvios denominados de desvio máximo:

$$D = \max |F_0(X_i) - S_n(X_i)|, \quad (19)$$

em que  $i = 1, 2, \dots, N$ , verificando-se a hipótese do poder do teste  $p$  – valor, então a normalidade da amostra é verificada. Se  $D < D_{(N,\alpha)}$  não se rejeita a  $H_0$ , ou seja, a amostra é proveniente de uma distribuição normal. Caso contrário  $D > D_{(N,\alpha)}$  é rejeita-se  $H_0$ , ou seja, a amostra não é proveniente de uma distribuição normal.

### 2.6.2 Autocorrelação

Para a aplicação dos gráficos de controle é necessário que os dados atendam a suposição de independência entre si, sendo de extrema importância a investigação da autocorrelação do grupo de informações. A presença da autocorrelação compromete o desempenho dos gráficos de controle, no sentido em que se podem confundir as causas especiais e comuns no processo.

Segundo Montgomery (2004), para contornar o problema da autocorrelação pode-se utilizar uma amostra menos frequente acarretando um aumento no tempo para detectar uma variação real do processo.

Dentro do controle estatístico de processo existem dois princípios básicos para contornar autocorrelação. O primeiro determina uma abordagem dos dados originais, mas devem-se ajustar os limites de controle para absorver a autocorrelação, sendo denominado de alargamento dos limites. O segundo princípio defende uma abordagem de plotagem dos resíduos do modelo de série temporal no gráfico de controle (ARAÚJO, 2010). Esse modelo será adotado neste trabalho, devido a seu desempenho superior, para evitar alarmes falsos e remover a autocorrelação dos dados.



De acordo com Follador (2010), a autocorreção é definida como a medida de dependência de uma observação num estante  $t$ , em relação àquela tomada num estante  $t - k$ , em que  $k$  é a defasagem (*lag*) considerada, isto é, a distância entre duas observações.

Para mensurar a autocorreção ao longo de uma série de observações distribuídas no tempo tem-se uma característica longitudinal, ou seja, uma série temporal. O coeficiente linear é dado pela equação 20, chamada função de correlação.

$$\rho_k = \frac{cov(x_t, x_{t-k})}{V(x_t)}, \quad (20)$$

em que  $cov(x_t, x_{t-k})$  é a covariância de observações separadas por  $k$  períodos de tempo e a variância constante é dada por  $V(x_t)$ . De forma geral estimam-se os valores de  $\rho_k$  com a função de autocorreção explicitada na equação 21.

$$r_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n-k} (x_t - \bar{x})^2}, \quad (21)$$

em que  $x_t$  é observação ou média dos subgrupos no tempo  $t$ ,  $\bar{x}$  é a média das amostras e  $k = 1, 2, 3, \dots$ .

### 2.6.3 Modelo auto-regressivo integrado de média móvel

Se as observações violam a suposição de independência, ocorrem alterações no ajuste das variáveis no controle de processo, pois existindo dependência entre as características de qualidade e as variáveis de controle o modelo assumido não será verdadeiro para os dados que originam do processo.

Nestas situações ocorre um aumento considerável no número de alarmes falsos que produz conclusões erradas em relação à estabilidade do processo. Os modelos de regressão ou de séries temporais têm se revelado úteis para modelar processos que demonstrem uma estrutura correlacional, posteriormente, na utilização de gráficos de controle para os resíduos dos modelos. Se existir um conjunto de variáveis dentro de um modelo de regressão os seus resíduos não serão correlacionados mesmo que a variável resposta esteja correlacionada (PEDRINI; CATEN, 2008).

De acordo com Follador (2010), o modelo *ARIMA* é uma generalização do modelo auto-regressivo de média móvel (*ARMA*). A apresentação *ARIMA(p, d, q)* refere-se, respectivamente, às ordens de auto regressão, de integração e de média móvel:  $p$  é o operador de auto-regressivo,  $d$  é o número de diferenças,  $q$  é o número de termos da média móvel. A escolha do modelo adequado será feito mediante as estatísticas de Akaike Information Criteria (*AIC*). Logo após o modelo ajustado *ARIMA*, verifica-se que os resíduos do modelo são independentes, denominados de ruído branco.

Denomina-se de modelo Auto-regressivo de ordem *AR(p)* o modelo dado pela equação 22.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \quad (22)$$

em que os termos  $Z_{t-1}, Z_{t-2}, + \dots + Z_{t-p}$  são independentes de  $a_t$ ,  $\phi_1$  é o parâmetro (peso) que descreve como  $Z_t$  se relaciona com o valor para  $Z_{t-1}$  (a série de dados) e  $a_t$  é o erro aleatório.

A metodologia implantada de Box e Jenkins e Reinsel (1994) para construção de modelos *ARIMA* é composta por três passos: Identificação; Estimação e Verificação. Logo, o passo 1 é mais relevante pois nele se determina os valores de  $p, d$  e  $q$ . Esta determinação é feita mediante as autocorrelações e autocorrelações parciais estimadas, as quais se espera que represente adequadamente as verdadeiras quantidades teóricas que são desconhecidas.

## 2.7 Critério de Informação de Akaike (AIC)

De acordo com Terra et al (2010), o critério de informação de Akaike (*AIC*) toma por base a teoria de decisão e penaliza os modelos com maior número de parâmetros para evitar excessos de parametrizações. Esse critério é expresso pela seguinte forma:

$$AIC = -2 \times (\log L(\theta; y)) + 2p, \quad (23)$$

em que  $\log L(\theta; y)$  é a função de log-verossimilhança maximizada sob o modelo considerado,  $p$  e o número de parâmetros a serem estimados. Para tomada de decisão em relação a qual o modelo utilizar. Por conseguinte, o modelo selecionado será aquele que apresentar o menor valor de *AIC*.

## 3 Resultados

Um gráfico temporal das medições de pHeo gráfico da função de autocorrelação amostral (Figura 2a e 2b), que evidencia a presença de autocorrelação nos dados.

Da mesma forma, os dados referentes à Turbidez e a função de autocorrelação amostral (Figura 3a e 3b), que indica a presença de autocorrelação.

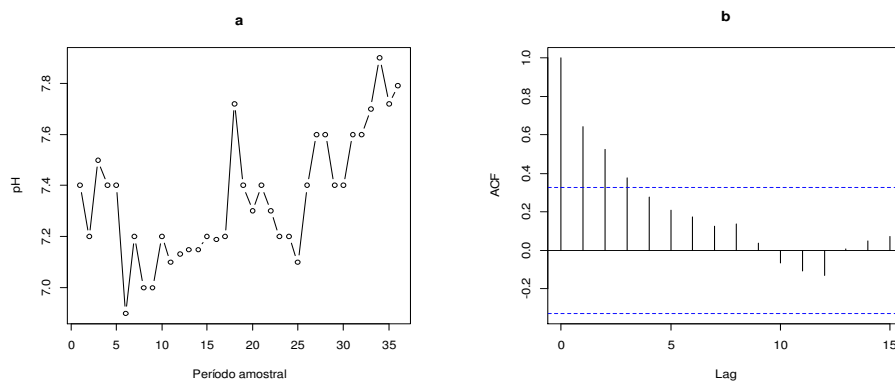


Figura 2 - Série temporal (a) e função de autocorrelação para os dados do pH (b).

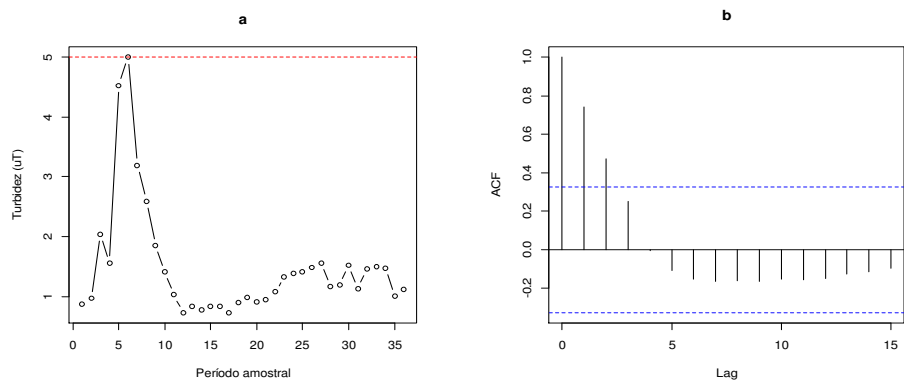


Figura 3 - Série temporal (a) e função de autocorrelação para os dados da turbidez (b)

Os dados originais com relação ao pH (potencial de Hidrogênio) apresentaram uma distribuição normal, ao nível de 5% significância (p-valor 0.1652) para o teste de Kolmogorov-Smirnov. Verifica-se que os dados são autocorrelacionados, não atendendo a suposição de independência das amostras.

Ajustou-se um modelo *ARIMA* (2,0,0), com um  $\sigma^2$  estimado de 0.02958 e critério de seleção  $AIC = -15.79$ . O modelo mostrou-se adequado, pois os resíduos apresentaram uma distribuição normal e não autocorrelacionados (Figura 4).

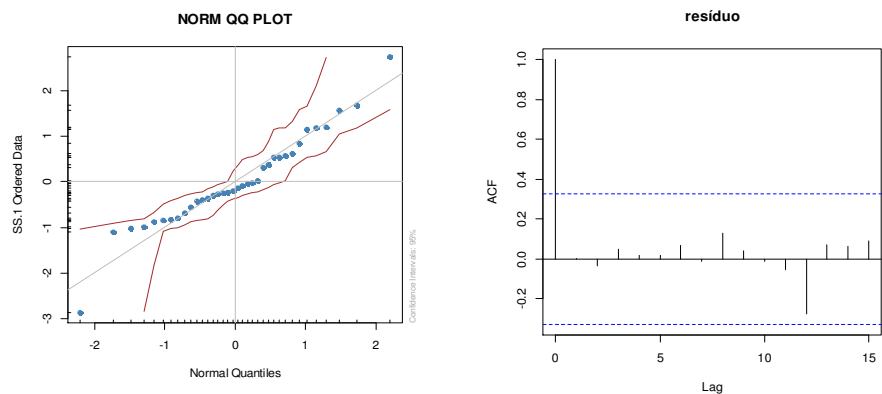


Figura 4 - Gráfico de normalidade e função de autocorrelação para os resíduos do pH.

Continua-se o desenvolvimento do controle estatístico de processo para o monitoramento da variável pH, sem que os resultados sejam alterados pela quebra das

suposições iniciais,(Figuras 5, 6 e 7), encontram-se um gráfico do tipo Shewhart, um gráfico CUSUM e um gráfico EWMA. Estes gráficos foram aplicados aos resíduos do modelo.

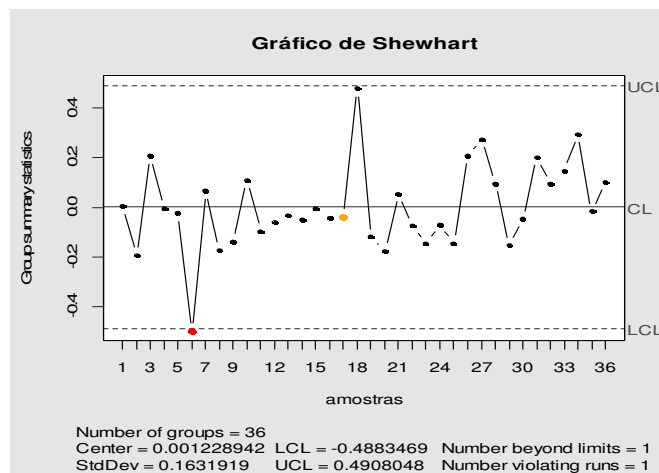


Figura 5 - Gráfico de Shewhart para os resíduos do pH.

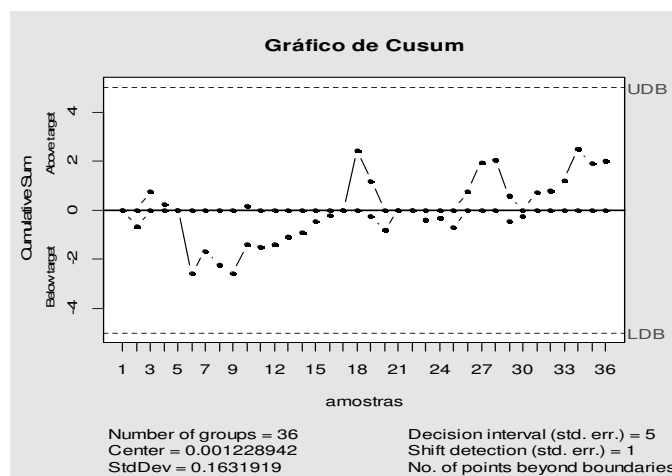


Figura 6 - Gráfico de Cusum para os resíduos do pH.

Os gráficos de controle para medidas individuais aplicados no monitoramento dos parâmetros pH e turbidez na qualidade da água potável em Campina Grande estão representados a seguir.

Verifica-se que os gráfico de Shewhart para o pH com relação ao controle de processo não está sob o controle estatístico, pois amostra de número 6, se encontra fora do

limite de controle e apresenta existência de alguns pontos sequenciais abaixo da linha média do ponto 11 ao 17.

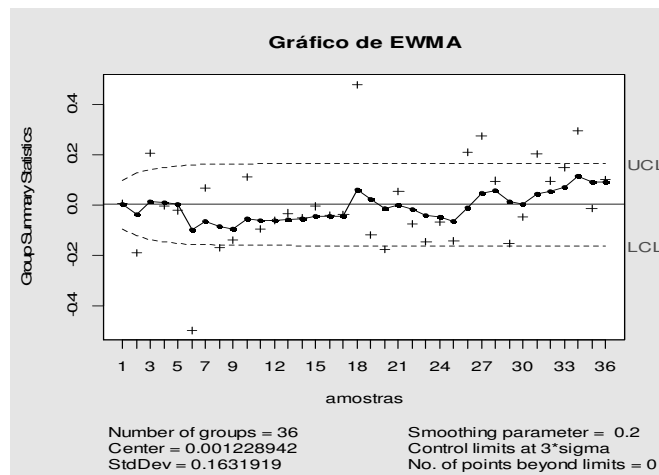


Figura 7 - Gráfico de EWMA para os resíduos do pH.

Os gráficos tipos Shewhart serão convencionais, com limites  $3\sigma$ . O gráfico CUSUM tem limite de decisão igual  $aH = 5$  e  $K = 0.5$ . O gráfico EWMA tem  $\lambda = 0.2$  e  $L = 3.0$ .

No gráfico de CUSUM para pH, verifica-se alguns valores próximos de zero entre as amostras 21 e 23. Atribuindo-se ao fato da média não sofrer alterações, ou seja, os desvios negativos são balanceados pelos positivos, provocando uma pequena oscilação em torno do zero. De acordo com, nos gráficos de CUSUM se existe tendência para se desenvolverem acima ou abaixo, é evidência de que a média do processo mudou e deve-se procurar a causa atribuível a esse fato é observado (Figura 6), onde se encontra alguns pontos sequenciais acima e abaixo da linha zero.

Com relação ao gráfico de EWMA para pH, este apresenta pontos que tem o desempenho que indicam uma possível instabilidade no processo, uma vez que as amostras a partir do ponto 6 a 17 são sequenciais no intervalo de tolerância abaixo da linha média, em seguida, a partir do ponto 30 a 36, acima do intervalo de tolerância

Os dados originais com relação à Turbidez ( $uT$ ) não apresentaram uma distribuição normal, ao nível de 5% significância (p-valor  $7.723e-05$ ) para o teste de Kolmogorov-Smirnov. Verifica-se que os dados são autocorrelacionados, não atendendo a suposição de independência das amostras (Figura 8). O modelo de maior adequabilidade foi  $ARIMA(1,0,0)$  com um  $\sigma^2$  estimado de 0.3907 e critério de seleção  $AIC=75.11$ .

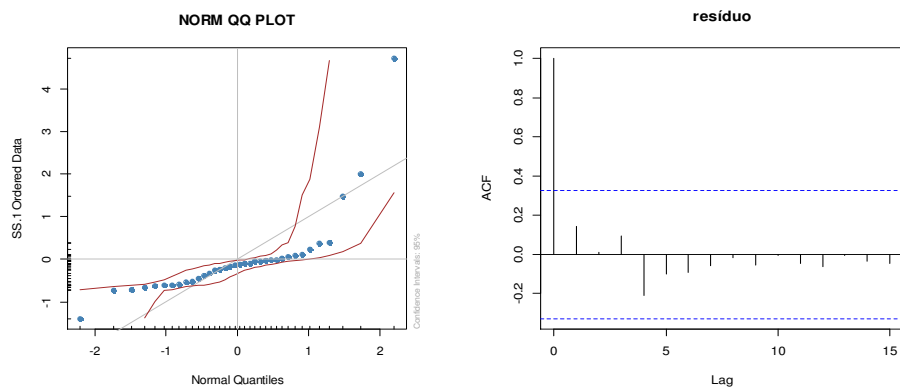


Figura 8 - Gráfico de normalidade e função de autocorrelação para os resíduos da Turbidez.

Continua-se o desenvolvimento do controle estatístico de processo para o monitoramento da variável Turbidez, sem que os resultados sejam alterados pela quebra das suposições iniciais. Estes gráficos foram aplicados aos resíduos do modelo.

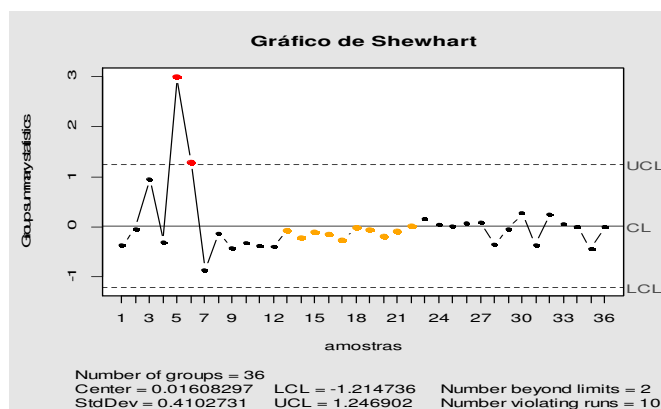


Figura 9 - Gráfico de Shewhart para os resíduos da Turbidez.

No gráfico de Shewhart (Figura 9) para turbidez foram encontrados dois pontos acima do limite superior, evidenciando diferenças de valores de alguns pontos em relação à média do processo, principalmente a partir de amostras coletadas em abril de 2011, caracterizando uma possível falta de controle no processo. Verifica-se que estes pontos acima do limite superior, ou seja, acima de 3 desvios-padrão, podem estar ligados à heterogeneidade dos dados. Apresentam-se pontos sequenciais abaixo da linha média, uma sequência que se encontra entre os pontos 8 e 22, o que caracteriza de acordo com Werkema (1995) uma indicação de falta de controle.

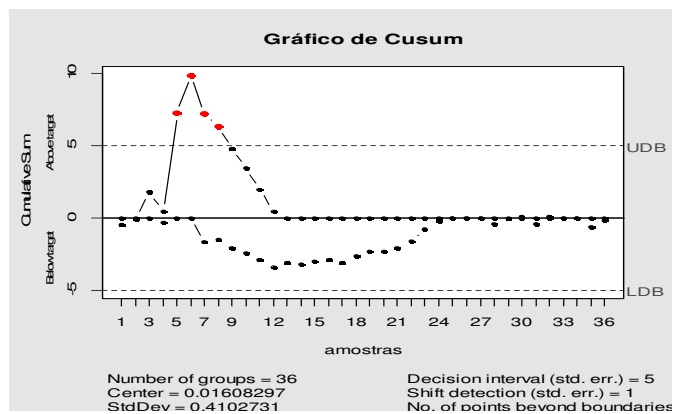


Figura 10 - Gráfico de Cusum para os resíduos da Turbidez.

Com relação ao gráfico de CUSUM para a Turbidez (Figura 10), observa-se que nos pontos 5 a 8, os valores tiveram uma mudança súbita de comportamento, ou seja, uma causa especial está agindo no processo e estar sob a ação de causas especiais, cabendo a investigação da causa desta perturbação no processo (podendo ser atribuído a falha do analista, duplicata da amostra, entre outros fatores), levando à necessidade de ações corretivas com objetivo de levar o processo à estabilidade em seguida após o ponto 28, a uma oscilação menor que esperada em torno do eixo 0.

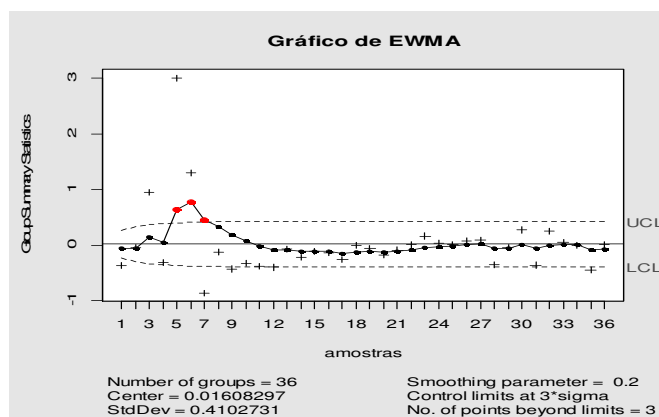


Figura 11 - Gráfico de EWMA para os resíduos da Turbidez.

Com relação ao gráfico de EWMA para a Turbidez (Figura 11) apresentam-se pontos sequenciais do ponto 12 aos 23, que se encontram no intervalo de tolerância inferior, o que indica uma possível instabilidade no processo. Por outro lado, os pontos 5, 6 e 7 são encontrados acima do intervalo superior, o que indica que o processo está fora de controle.

#### 4 Discussões

Os valores do parâmetro pH podem sofrer alterações devido às decomposições ácidas da poluição atmosférica, gás carbônico liberado de forma natural pelo meio ambiente ou oriundos de ações antropogênicas, que, decompostos na água, formam o ácido carbônico que conduz à redução do pH (FOLLADOR, 2010).

No período do estudo, os valores monitorados permaneceram dentro da faixa de valores permitidos pela Portaria Nº 2.914/2011 do Ministério da Saúde, de acordo com Capítulo V, Artigo 39º, Inciso I. Recomenda-se que, no sistema de distribuição, o pH da água seja mantido na faixa de 6.0 a 9.5.

Resultados observados por Campos et al. (2002), a qualidade da água tratada pelo Departamento de Águas e Esgotos da cidade de Araraquara, São Paulo, apesar de ser considerada satisfatória com amostras detectadas que não atendiam ao padrão para o pH, nesse estudo todas as amostras permaneceram dentro do padrão esperado.

A turbidez é caracterizada com relação à qualidade da água, a suspensão de materiais, ou seja, partículas e algas dissolvidas, dificultando a penetração da luz. A turbidez de causa natural não traz inconvenientes sanitários diretos, entretanto os sólidos suspensos podem abrigar microrganismos patogênicos (Von Sperling, 1996).

De acordo com a legislação vigente, o valor permitido para Turbidez no sistema de distribuição da água (em toda a extensão do sistema de distribuição) e de 5uT de acordo com Artigo 30º, Inciso I. Refere-se ao padrão organoléptico de potabilidade. Conforme artigo 5º, inciso IV, o padrão organoléptico é o conjunto de parâmetros caracterizados por provocar estímulos sensoriais que afetam a aceitação para consumo humano, mas que não necessariamente implicam risco à saúde.

Entretanto a Turbidez pode aumentar durante o processo de tratamento da água (em função da adição de produtos químicos pós-filtração) ou na rede de distribuição (devido a interferências como infiltração de águas de chuva ou formação de biofilmes). Nesses casos, deve ser respeitado o valor máximo de 5.0 uT em qualquer ponto da rede.

Em estudo realizado por Correia et al. (2008), Janeiro à Março obteve-se um maior índice de Turbidez, por conta do maior índice de chuvas na região, elevando o processo erosivo, colaborando para a elevação da turbidez. Em nosso estudo os maiores valores para esse parâmetro esta entre as amostras 6 a 9, que corresponde aos meses de Maio a Agosto de 2011, que se dá o período de precipitação da região. Entretanto as amostras não extrapolaram o valor permitido pela legislação vigente.

## Conclusões

O grau de comprometimento da qualidade da água no abastecimento da cidade de Campina Grande é baixo, pois apenas a variável Turbidez apresentou uma única amostra igual ao limite estipulado pela Portaria de nº 2914/2011 do Ministério da Saúde, que dispõe sobre a classificação dos corpos de água potável; A variável Turbidez apresentou-se em alguns pontos de controle elevada variabilidade dos dados, com exceção do pH, que apresentou uma distribuição normal ao nível de 5% de significância. As variáveis pH e turbidez apresentaram-se amostras auto-correlacionadas, não atendendo as suposições de independência das amostras.

Para tratar a autocorrelação foi aplicado um modelo *ARIMA* e posteriormente gráficos de controle aos resíduos do modelo. A metodologia utilizada permitiu definir algumas considerações importantes, tais como uma proposta de monitoramento do nível do pH e Turbidez da água da cidade de Campina Grande, através do Controle Estatístico



de Processo. Além disso, comparou-se a sensibilidade existente entre os gráficos de controle de Shewhart, CUSUM e EWMA na detecção de pequenas mudanças nas variáveis relacionadas à qualidade da água e verificando se realmente estão dentro dos limites de controle.

Os gráficos CUSUM como EWMA, tem como pontos positivos detectar mais rapidamente pequenos desvios e dão informações sobre a tendência do processo, sendo indicados para o monitoramento deste processo. O gráfico EWMA ainda tem uma vantagem adicional: a de ser robusto à hipótese de normalidade. Dentre os gráficos utilizados Shewhart, CUSUM e EWMA. O mais eficiente é o EWMA, pois se mostra uma maior sensibilidade às pequenas variações.

A utilização de gráficos de controle estatístico de processos para monitorar indicadores de qualidade da água em Estações de Tratamento (ETA's) merece atenção especial, uma vez que estes dados geralmente apresentam autocorrelação.

É importante salientar que a utilização de métodos de previsão, como, *ARIMA*, aplicação de gráficos de controle estatístico aos resíduos, são sofisticados, e, portanto, podem ser de difícil interpretação, necessitando de pessoas especializadas para tal fim. Para o monitoramento futuro aconselha-se também a análise da utilização de outras técnicas para tratar a autocorrelação, como limites de controle alargados ou diminuir a frequência de retirada de amostras.

## Agradecimentos

Os autores agradecem à Universidade Federal Rural de Pernambuco pelo apoio.

CRUZ, D. V.; CUNHA FILHO, M.; FALCÃO, A. P. S. T. Statistical process control applied to monitoring and Phwater turbidity no supply Campina Grande- PB. *Rev. Mat. Estat.* São Paulo, v.32, n.4, p.459-477, 2014.

- *ABSTRACT: The hydrological cycles have been significantly altered due to human activities. As a consequence, it has been observed a significant drop in the quality of surface water for human consumption, which impairs the survival and quality of life. To get an assessment so efficient and fast parameters that classify water techniques of statistical process control were used. This work arose from the need for continuous monitoring of some parameters related to water supply provided in the city of Campina Grande. That legislation should be under the Ministry of Health, Ordinance No. 2914/2011 laying down procedures for control and surveillance of water for human consumption. Thus, employed temporal Series models to acquire a structure without the presence of autocorrelation, later applying graphs and individual control Shewhart, exponentially weighted moving average (EWMA), cumulative sum (CUSUM). Avoiding the occurrence of false alarms is important to note that the EWMA chart performed better in water quality monitoring, because of its efficiency, it is sensitive to small variations in the reduction or increase in the levels of pH and turbidity. Analyses were performed using the R software package to aid the forecast.*
- *KEYWORDS: Statistical control; CUSUM; EWMA; water quality.*

## Referências

ALBUQUERQUE, A. A.; FIRMINO, M. B. M.; CURI, W. F - Uso racional de energia no bombeamento de água em sistemas de abastecimento, via programação não-linear. In: IV SEREA-SEMINÁRIO HISPANO-BRASILEIRO SOBRE SISTEMAS DE

ABASTECIMENTO URBANO DE ÁGUA,6., 2004 , João Pessoa. *Anais...* João Pessoa, Brasil, CD-ROM, 2004.

ARAÚJO, N. G. A. *Aplicação de gráficos de controle para monitoramento estatístico da turbidez da água potável*. 2010. 51f. Monografia (Engenharia de Produção) — Centro de Ciências Tecnológicas, Departamento de Engenharia de Produção e Sistemas, Universidade do Estado de Santa Catarina, Joinville, 2010.

BICUDO, C.E.de M.; TUNDISI, J.G.; SCHEUENSTUHLS, M.C.B. *Águas do Brasil: análises estratégicas*. São Paulo, Instituto de Botânica, 2010. 224p.

BOX, G. E. P., JENKINS, G. M., and Reinsel, G. C. *Time Series Analysis, Forecasting and Control*, 3.ed. Prentice Hall, Engle Wood Cliff, New Jersey.1994. 598p.

BRASIL, Ministério da Saúde. Portaria n.º 2.914, de 12 de Dezembro de 2011. Dispõe sobre normas de potabilidade de água para o consumo humano. Brasília: SVS, 2011.

CAGEPA. *Companhia de Águas e Esgotos do Estado da Paraíba*. 2014. Disponível em: <http://www.cagepa.pb.gov.br/portal/>. Acessado em: 02/01/2014.

CAMPOS, J. A. D. B., FILHO, A. F., FARIA, J. B. Qualidade sanitária da água distribuída para o consumo humano pelo sistema de abastecimento público da cidade de Araraquara-SP. *Alim. Nut.*, São Paulo, v.13, n.1, p.117-129, 2002.

CORREIA, Aislán et al. Análise da Turbidez da Água em Diferentes Estados de Tratamento. VIII ERMAC (Encontro Regional de Matemática Aplicada e Computacional). Universidade Federal do Rio Grande do Norte. Disponível em: [http://www.dimap.ufrn.br/~sbmac/ermac2008/Anais/Resumos%20Estendidos/Afinalise%20da%20turbidez\\_Aislán%20Correia.pdf](http://www.dimap.ufrn.br/~sbmac/ermac2008/Anais/Resumos%20Estendidos/Afinalise%20da%20turbidez_Aislán%20Correia.pdf), 2008.

FAY, E. F.; SILVA, C. M. M. S. *Índice de Uso sustentável da água (ISA-ÁGUA) na região do sub-médio São Francisco*. 1.ed. Jaguariúna: Embrapa Meio Ambiente, 2006. 157p.

FERREIRA, G. U. A. *Estudo de testes estatísticos para o vetor de médias em controle de processos multivariados sob amostragem dupla*. 2010, 134f. Dissertação (Mestrado em Estatística) - Instituto de Ciências Exatas, Programa de Pós-graduação em Estatística, Universidade Federal de Minas Gerais, 2010.

FOLLADOR, F. et al. Controle de qualidade da água medido através de cartas de controle de Shewart, CUSUM e MMPE. *Engenharia Ambiental*, v. 9, p. 183–197,2012.

FOLLADOR, F. A. C. *Controle estatístico de processo aplicado à qualidade de águas superficiais*.2010, 118f. Tese (Doutorado em Engenharia de Produção) – Centro de Ciências Exatas e Tecnológicas, Programa de Pós-graduação em Engenharia Agrícola, Universidade Estadual do Oeste do Paraná, 2010.

HAWKINS, D.; OLWELL, D. Cumulative sum control charting: An underutilized spctool. *Quality Engineering*, v. 5, n. 3-4, p. 463–477, 1993.

HENNING, Elisa et al. Um estudo para a aplicação de gráficos de controle estatístico de processo em indicadores de qualidade da água potável. *Sistemas & Gestão*, v. 9, n. 1, p. 2-13, 2011.

- LU, C. W.; JR, M. R. R. CUSUM charts for monitoring an autocorrelated process. *Journal of Quality Technology*, April, v. 33, n.3, p. 1–22, 2001.
- MONTGOMERY, D. *Introduction to statistical quality control*. 4.ed. New York: John Wiley & Sons, Inc., 2004. 513f.
- PEDRINI, D. C.; CATEN, C. S. Comparação entre gráficos de controle para resíduos de modelos. *GEPROS. Gestão da Produção, Operações e Sistemas*, Porto Alegre-RS, v. 4, n.3, p. 123 – 140, 2008.
- R Development Core Team (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- SCUDINO, P. A. A. *A utilização dos alguns testes estatísticos para Análise da variabilidade do preço do mel nos municípios de Angra dos Reis e Mangaratiba, Estado do Rio de Janeiro*. 2008, 51f. Monografia (Bacharel em Matemática) - Instituto de Ciências Exatas, Departamento de Matemática, Universidade Federal Rural do Rio de Janeiro, 2008.
- SILVA, W. Aplicação do gráfico de controle EWMA no processo produtivo de uma indústria de alumínio: Um estudo de caso. In: XI SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO. São Paulo. *Anais...* Bauru, São Paulo, Brasil, 08 a 10 de novembro de 2004.
- SILVA, W. et al. Aplicação do gráfico de controle de média móvel exponencialmente ponderada (MMEP) para monitorar a variabilidade de um processo produtivo. In: SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA, 2, 2006, Resende –RJ. *Anais...* Disponível em: [http://www.aedb.br/anais-seget/ac\\_pos.htm](http://www.aedb.br/anais-seget/ac_pos.htm).
- TERRA, M. F. et al. Ajuste dos modelos logístico e gompertz aos dados de crescimento de frutos da tamareira-anã *Phoenix roebelenii* O'BRIEN. *Magistra*. Cruz das Almas - BA, v. 22, n. 1, p. 01-07, jan./mar., 2010.
- VON SPERLING, M.. *Introdução à qualidade das águas e o tratamento de esgotos*. 2. ed. Minas Gerais: Departamento de Engenharia Sanitária e Ambiental, Universidade Federal de Minas Gerais, 1996. 243p.
- WALTER, O. M. F. C., HENNING, E., CARDOSO, M. E., SAMOBYL, R. W. Aplicação individual e combinada dos gráficos de controle Shewhart e CUSUM: uma aplicação no setor metal mecânico. *Gest. Prod.*, São Carlos, v. 20, n. 2, p. 271-286, 2013.
- WERKEMA, M. C. C. *Ferramentas estatísticas básicas para o gerenciamento de processos*. 1.ed. Belo Horizonte, MG : Fundação Christiano Ottoni, UFMG, 1995. 384p.
- ZUIN, V. G.; LORIATTI, M. C. S.; MATHEUS, C. E. O emprego de parâmetros Físicos e Químicos para a avaliação da qualidade das águas naturais: um proposta para educação química e ambiental na Perspectiva CTSA. *Química e Sociedade*, v. 31, n.1, p. 3–7, 2009.

Recebido em 11.07.2014

Aprovado após revisão em 15.12.2014