

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Limites estatísticos de controle para instrumento de monitoramento de barragem

Emerson Lazzarotto¹

Centro de Engenharias e Ciências Exatas, UNIOESTE, Foz do Iguaçu, PR

Elias Alberto da Silva²

Universidade Federal do Paraná, Curitiba, PR

Resumo. As leituras da série temporal y_t de um instrumento de monitoramento de uma barragem, mesmo operando em condições normais, podem não estar dentro dos valores determinados no projeto da obra. A variação térmica sazonal é a principal razão desta variabilidade inerente. A literatura de séries temporais sugere a captura desta variabilidade através de um modelo estatístico ARIMA adequado \hat{y}_t e que o controle estatístico se dê sobre a série temporal dos resíduos, $x_t = y_t - \hat{y}_t$. A série dos resíduos é um conjunto de dados, cuja análise pode indicar se os dados reais estão sob controle do ponto de vista estatístico. Desta forma, o principal objetivo deste trabalho é garantir que a característica que um instrumento de monitoramento avalia está sob controle estatístico, além de realizar previsões sobre a sua série temporal de leituras. O ponto de partida é uma amostra da série temporal de um instrumento que monitora uma barragem. Em seguida, uma parte da amostra (fase I) é usada para estimar o melhor modelo ARIMA, que é escolhido com base na sua capacidade preditiva na outra parte da amostra (fase II). Em especial, estabelecem-se limites estatísticos de controle para a média amostral e a amplitude da série temporal dos resíduos x_t . Os resultados mostram um procedimento alternativo para estabelecer limites estatísticos de controle para tais gráficos de controle, mesmo quando algumas hipóteses teóricas não são aplicáveis ao caso real.

Palavras-chave. Gráficos de Controle, Instrumentos de monitoramento de barragem, Modelos ARIMA, não normalidade.

1 Introdução

As barragens de concreto e de terra são utilizadas para diversas finalidades, em especial, barragens de concreto de usinas hidrelétricas tem uma grande importância econômica, social e ambiental. Tais barragens são, em geral, monitoradas por inspeções visuais e instrumentos de monitoramento. A instrumentação pode detectar infiltração nos maciços de terra e nas fundações, trincas no concreto, movimentos entre os blocos da barragem, subpressões na cabeça dos blocos a montante, ou a jusante da Casa de Força [2]. Assim, as

¹emerson.lazzarotto@gmail.com

²elias.s51@hotmail.com

séries temporais de leitura dos instrumentos de monitoramento podem indicar tendências no comportamento da barragem e sugerir ações preventivas da equipe de segurança da barragem.

Em especial, o piezômetro é um instrumento instalado na estrutura de concreto. Sua função é medir a subpressão no contato concreto-rocha e em níveis mais permeáveis do maciço basáltico da fundação [2]. As subpressões e infiltrações nas barragens de concreto não podem ser previstas com acurácia por modelos numéricos que se baseiam nas leis físicas que governam o fenômeno envolvido.

A alta disponibilidade de dados de monitoramento tem difundido a aplicação de modelos estatísticos para a previsão do comportamento da barragem. Modelos estatísticos têm servido como complemento às inspeções visuais e aos modelos numéricos no suporte a tomada de decisões [7]. Em geral, as leituras de instrumentos de monitoramento de barragens são avaliadas com valores definidos em projeto. O domínio de variação (limites de controle) da série de leituras pode não ser mais realístico durante seu período operacional. Deste modo, são necessárias outras formas de análise do monitoramento.

O gráfico de controle é uma ferramenta útil para monitorar a média e a variabilidade de um aspecto de interesse. A qualidade é inversamente proporcional a variabilidade e, quando esta é inerente, deve-se mantê-la dentro de determinados limites probabilísticos de controle para a melhoria de sua qualidade [5].

No caso de instrumentos de monitoramento de barragem de usina hidrelétrica, a variabilidade pode estar relacionada com a temperatura ambiente e isto pode causar subpressões no contato concreto-rocha [6]. Os autores [3] sugerem que essa variabilidade inerente seja capturada e que se estabeleçam limites de controle para os resíduos. Os resíduos são as diferenças entre os valores da série temporal dos dados de leitura do instrumento e os de um modelo adequado encontrado para a respectiva série. Para capturar a variabilidade, os modelos estatísticos ARIMA são previsores lineares consagrados na literatura de análise e previsão de séries temporais [1].

2 Fundamentos Teóricos

Gráficos de controle são uma das principais ferramentas do controle estatístico de qualidade [5]. Dados fora do intervalo dos limites de controle do gráfico de controle são uma evidência, estatisticamente confirmada através de um teste de hipótese, que o processo avaliado está fora de controle. Os limites de controle foram inicialmente estabelecidos para dados provenientes de uma distribuição normal e não autocorrelacionados.

Entretanto, os dados da série temporal de um instrumento de monitoramento, em geral, apresentam autocorrelação e são dependentes da variação ambiental. Neste caso, [4] e [3] sugerem capturar esta estrutura de autocorrelação responsável pela maior parte da variabilidade nos dados através de algum modelo. Sugerem ainda estabelecer limites de controle para o gráfico dos resíduos. O controle estatístico é avaliado mediante o gráfico de controle dos resíduos.

Para amostras com uma unidade, caso da aplicação realizada, os mais importantes gráficos de controle utilizados são da média e da variabilidade do processo [5]. Estas

variáveis são representadas, respectivamente, pela variável $x_t = x(t)$ e pela variável R_t que estima a variabilidade do processo por meio da amplitude móvel de duas observações consecutivas, isto é, $R_i = |x_i - x_{i-1}|$. Os limites de controle para uma probabilidade de 99,73%, presumindo normalidade, do gráfico x de acordo com [5] são dados por:

$$\begin{aligned} LSC &= \bar{x} + 2,6595 \cdot R \\ \text{Linha central} &= \bar{x} \\ LIC &= \bar{x} - 2,6595 \cdot R \end{aligned} \tag{1}$$

e do gráfico R são

$$\begin{aligned} LSC &= 3,267 \cdot \bar{R} \\ \text{Linha central} &= \bar{R} \\ LIC &= 0. \end{aligned} \tag{2}$$

No caso de não normalidade, [5] sugere determinar os limites de controle com base nos percentis de alguma distribuição de probabilidade mais adequada aos dados.

Os modelos autorregressivos integrados de médias móveis (ARIMA) são modelos estatísticos que têm sido empregados, com sucesso, para modelagem de séries temporais que apresentam autodependência [1]. A forma mais simples do modelo ARIMA é expressa por

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} + a_t. \tag{3}$$

A forma mais completa inclui, além de termos autorregressivos e de médias móveis de ordem p e q , termos sazonais de ordem P e Q . O termo 'integração' está relacionado à necessidade de efetuar diferenciações estacionárias (d) ou sazonais (D) para garantir a propriedade de estacionariedade da série. A forma geral é representada por $ARIMA(p, d, q) \times (P, D, Q)_s$ em que s é a ordem de sazonalidade, $s = 12$ no caso de dados mensais.

A 'qualidade' do ajuste de um modelo é medida em termos do erro cometido pelo modelo. Em cada instante de tempo t , o erro e_t , é a diferença entre o valor real e o estabelecido pelo modelo, $e_t = x_t - \hat{x}_t$. Dado um conjunto de n observações, uma das mais importantes formas de avaliar o erro é o erro percentual absoluto médio (MAPE), dado por:

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{e_t}{x_t} \right|}{n}. \tag{4}$$

No diagnóstico de seleção do modelo especificado levam-se em consideração outros testes estatísticos ou gráficos [1]. Neste trabalho, o diagnóstico levou em consideração o gráfico da função de autocorrelação (FAC) e de autocorrelação parcial (FACP); o periodograma integrado dos resíduos e testes de randomicidade dos resíduos.

A hipótese de normalidade foi avaliada por meio do gráfico da probabilidade normal e de testes de normalidade conhecidos da literatura específica como: Shapiro-Wilk, Kolmogorov-Smirnov e Anderson-Darling. No caso da rejeição da normalidade, aplicou-se o teste da razão de verossimilhança para escolher um modelo de distribuição paramétrica que apresente melhor ajuste em relação ao da distribuição normal.

3 Metodologia

Para validar as ideias de controle estatístico de qualidade a série temporal de dados de instrumento de monitoramento de barragem foi aplicada uma metodologia aos dados de um piezômetro, que monitora o bloco 6, do trecho E, da barragem da usina hidrelétrica de Itaipu, localizada no Rio Paraná, entre o Brasil e o Paraguai. Estavam disponíveis a média mensal das leituras do instrumento PSE28, o que compreende 435 observações.

As demais etapas do método desenvolvido são: 1) Seleção dos dados y_t a serem utilizados e separação de parte da amostra usada para estimação (fase I) e validação (fase II) dos modelos ARIMA. 2) Seleção do 'melhor' modelo \hat{y}_t para realizar previsões e obtenção dos resíduos $x_t = y_t - \hat{y}_t$ e das amplitudes dos resíduos $R_t = |x_t - x_{t-1}|$. 3) Verificação das hipóteses teóricas para determinar os limites de controle. 4) Ajustamento dos dados a alguma distribuição de probabilidade alternativa, caso a hipótese de normalidade seja rejeitada. 5) Determinação dos limites de controle via distribuições alternativas. 6) Cálculo dos valores da linha central, do *LSC* e do *LIC* para os gráficos x_t e R_t . 7) Construção dos gráficos de controle x_t e R_t e avaliação do estado estatístico de controle.

4 Resultados

4.1 Seleção dos dados de acordo com o melhor modelo

Dos 435 dados mensais (médias) disponíveis de 11/1981 a 01/2018 selecionou-se as 12 últimas observações para o período de validação (fase II), as demais compreendem o período de estimação (fase I). Na Tabela 1 exibe-se modelos ARIMA $(p, d, q) \times (P, D, Q)_{12}$ escolhidos automaticamente pelo *software* estatístico utilizado e os respectivos valores do MAPE no período de validação (fase II).

Tabela 1: Modelos ARIMA da série temporal do instrumento PSE28.

Modelo ARIMA	$(2, 0, 0) \times (2, 0, 2)_{12}$	$(2, 1, 1) \times (0, 1, 1)_{12}$	$(6, 1, 5) \times (1, 1, 1)_{12}$	$(3, 0, 2) \times (0, 1, 1)_{12}$	$(1, 0, 0) \times (0, 1, 1)_{12}$	$(1, 0, 0) \times (0, 1, 1)_{12}$
MAPE	0,42	0,44	0,37	0,35	0,40	0,40
Data Início	12/88	02/93	03/95	04/97	04/97	06/01

Observa-se na Tabela 1 que o modelo ARIMA $(3, 0, 2) \times (0, 1, 1)_{12}$ é o que melhor se ajusta aos dados pois produz os menores erros. Isto ocorre porque o instrumento tem um comportamento mais regular a partir de 04/97 comparado a toda a série de dados.

Para a seleção do modelo ainda levou-se em consideração: a ausência de autocorrelação residual, conforme se observa na Figura 1 que contém, à esquerda, o gráfico da FAC e, à direita, o gráfico da FACP dos resíduos e testes sobre a randomicidade dos resíduos. Portanto, o modelo selecionado para ajustar a série temporal do instrumento PSE28 foi o ARIMA $(3, 0, 2) \times (0, 1, 1)_{12}$ com dados a partir de 04/1997 até 01/2018 e foram geradas previsões de 02/2018 a 01/2019.

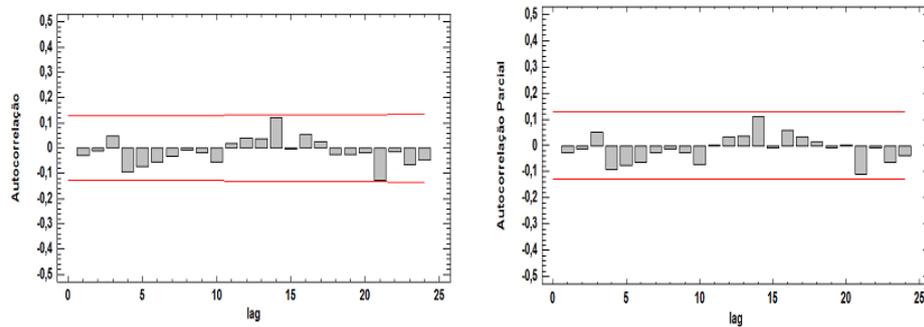


Figura 1: Gráfico da FAC (esq.) e da FACP (dir.) dos resíduos do modelo ajustado.

4.2 Verificação da hipótese de normalidade

Em relação a série temporal dos resíduos x_t e da amplitude R_t , testes de normalidade levaram a conclusão de que não se pode rejeitar a hipótese de que x_t vem de uma distribuição normal e que há indicações para rejeitar a hipótese de normalidade de R_t com 95% de confiança.

Desta forma, para a série das amplitudes, R_t , buscou-se distribuições de probabilidade contínuas alternativas e exibe-se na Tabela 2 o valor da função log-verossimilhança para algumas destas distribuições. A distribuição que melhor se ajustou a esta série é uma Weibull(β, θ) com o parâmetro de forma β estimado em 1,2973 e o parâmetro de escala θ estimado em 0,9358. Testes de ajustamento aplicados a série das amplitudes indicaram que não se pode rejeitar a hipótese de que a série das amplitudes seja proveniente de uma distribuição Weibull(1,2973;0,9358).

Tabela 2: Estatística log-verossimilhança para algumas distribuições alternativas.

Distribuição	Log-verossimilhança
Weibull	-181,277
Gama	-181,861
Normal	-228,938

4.3 Cálculos da linha central e limites de controle dos gráficos x_t e R_t

Para a construção dos gráficos x_t e R_t estavam disponíveis, respectivamente, 238 e 237 observações. As doze últimas observações de cada gráfico compõe a fase II da análise do controle estatístico, as demais, a fase I, de estabelecimento dos limites de controle. Os dados do gráfico x_t seguem uma distribuição normal, assim, fixou-se a probabilidade $\alpha = 0,0027$ correspondente aos tradicionais limites 3σ dos gráficos de controle e obteve-se, de acordo com a equação 1, os valores da linha central do *LSC* e do *LIC*.

Para o gráfico R_t , cujos dados não seguem uma distribuição normal, os limites de controle são definidos pela distribuição Weibull(1, 2973; 0, 9358). A linha central (média das amplitudes) é $LC = 0,864$, calculada conforme a equação 2. O LSC e o LIC foram determinados com base nos valores críticos da Weibull e não conforme a equação 2.

A Figura 2 apresenta o gráfico de controle x_t dos resíduos das observações individuais e a Figura 3 o gráfico de controle R_t das amplitudes dos resíduos. Os gráficos incluem as respectivas linhas centrais e limites de controle. Os gráficos já contêm a totalidade dos dados disponíveis (fases I e II). Observe que para ambos os gráficos não há nenhuma observação fora dos limites de controle.

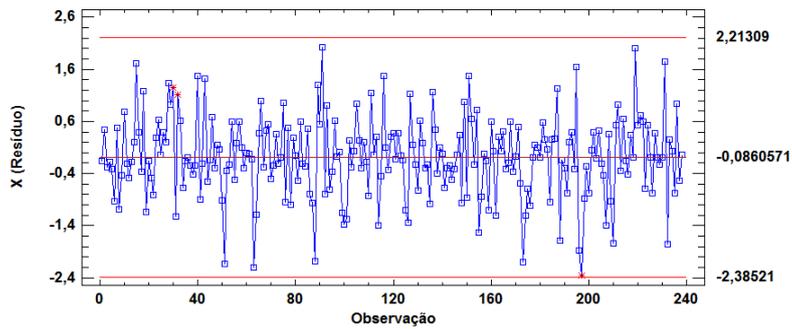


Figura 2: Gráfico x_t dos resíduos.

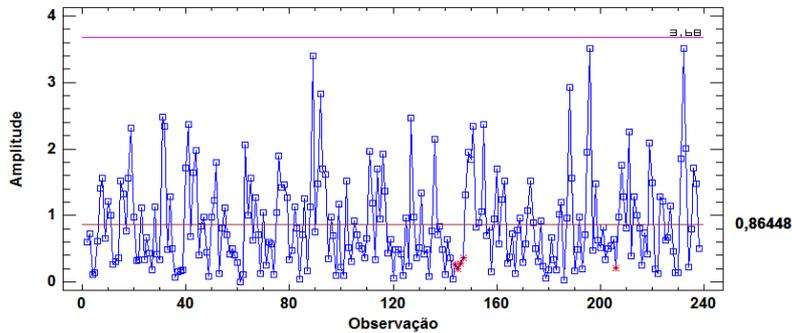


Figura 3: Gráfico R_t das amplitudes dos resíduos.

Usando os limites de controle de 99,73% para o gráfico dos resíduos x_t (distribuição normal) e das amplitudes R (distribuição Weibull), a probabilidade de que nenhuma observação esteja fora dos limites de controle (FLC), se os dados são provenientes das distribuições assumidas, é de

$$P(FLC = 0) = (0,9973)^{(226+225)} = 0,295 > 0,05.$$

Pode-se declarar que o processo está sob controle ao nível de 95% de confiança em todo o período de avaliação. Não há razões para admitir que os dados da leitura do instrumento PSE28 indiquem que seu processo de variação esteja estatisticamente fora de controle.

5 Conclusões

Neste artigo desenvolveu-se um método para estabelecer limites estatísticos de controle para gráficos de controle de dados referentes à série temporal das leituras de um instrumento de monitoramento de barragem. Segundo a equipe de engenharia de segurança da barragem, a característica de qualidade avaliada por este instrumento está sob controle. Entretanto, esse parecer, não se baseia em limites de controle justificados com base em algum conhecimento teórico.

Foi assegurado, no decorrer do trabalho, que o modelo ARIMA selecionado se ajusta bem aos dados das últimas 250 observações, cujo erro médio é inferior a 0,5%. Estatisticamente, o instrumento indica que a subpressão, no local da barragem que ele mensura, apresenta apenas uma variabilidade inerente. Variabilidade esta que, segundo a equipe de engenharia de segurança da barragem, se deve possivelmente à variação térmica do ambiente. Acredita-se, também, que não há causas especiais de variação.

Principais contribuições do trabalho: 1) determina os limites de controle com base em dados históricos que independem de modelos físicos e podem ser mais realísticos; 2) ajusta um modelo, que realiza previsões acuradas, à série de dados de leitura do instrumento; 3) alerta para a seleção prévia dos dados; 4) trabalha com distribuições de probabilidade, alternativas à distribuição normal, o que reduz o número de falsos alarmes.

Referências

- [1] G. Box, G. Jenkins e G. Reinsel. *Time-series analysis - forecasting and control*, 4th. ed., Hoboken, New Jersey: John Wiley and Sons Inc., 2008.
- [2] Itaipu Binacional. *Valores de Controle para os instrumentos integrantes do "ADAS" e critérios para a geração de alarmes*. Foz do Iguaçu, 2002.
- [3] R. Johnson e D. W. Wichern. *Applied Multivariate Statistical Analysis - 6th. ed.*, Upper Saddle River - NJ: Pearson Prentice Hall, 2007.
- [4] R. Mason e J. Young. *Multivariate statistical process control with industrial applications*, Philadelphia: American Statistical Association - Society for Industrial and Applied Mathematics, 2002.
- [5] D. Montgomery, *Introdução ao Controle Estatístico de Qualidade*, 4a ed., Rio de Janeiro: LTC, 2013.
- [6] J. Rosso, A. Fiorini, E. Porto e J. Silveira, Barragem de Itaipu - Lições aprendidas com o comportamento térmico das estruturas tipo gravidade aliviada e contraforte, Relatório Técnico Itaipu Binacional, Foz do Iguaçu, 1995.
- [7] F. Salazar, R. Morán, M. A. Toledo e E. Onate. Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. *Arch Computat Methods Eng.*, 24:1-21, 20017. DOI 10.1007/s11831-015-9157-9.