

MEDIÇÕES AUTOMÁTICAS DA COTA DE RIOS: QUÃO CONFIÁVEIS SÃO?

*Thaine Herman Assumpção*¹; Ângelo Breda²*

Resumo – O trabalho realizado teve por objetivo principal identificar e quantificar as diferenças das medições de cota em postos de monitoramento da bacia do rio Iguaçu. Isso foi feito através da comparação entre séries geradas por sensores automáticos e séries de dados medidos convencionalmente, considerados mais fidedignos. Para tal foram estimados parâmetros estatísticos (erro absoluto médio, frequência de erros, coeficiente de Pearson, entre outros) utilizando-se os dados das estações selecionadas. O resultado geral foi de que as séries automáticas e convencionais são fortemente correlacionadas, com a primeira superestimando levemente os valores de cota. Além disso, erros absolutos médios ocorrem em 27% do tempo, e são da ordem de 10 a 30 cm. Séries mensais de erro se mostraram como a sobreposição de diferentes padrões de erro, cuja composição varia de estação para estação, implicando em graus de similaridade definidos individualmente. Concluiu-se que as séries de cota são moderadamente ou bastante similares na maioria dos postos analisados.

Palavras-Chave – Monitoramento fluviométrico; estações automáticas; estações convencionais.

AUTOMATED WATER LEVEL MEASUREMENTS: HOW TRUSTWORTHY ARE THEY?

Abstract – The main goal of this study was to identify and quantify differences between automatic and manual water level measurements at Iguaçu River Basin. For this purpose, statistical parameters (absolute mean error, error frequency, Pearson coefficient, etc) were estimated using the data set from the selected stations. Overall results show that automatic and manual series are strongly correlated and that the automatically generated data tend to overestimate the water level. Moreover, the absolute mean errors take place on 27% of the time, switching its intensity between 10 and 30 cm. Monthly error series revealed themselves as a superposition of different error patterns, whose composition changes from each station, implicating at individual compatibility stats. In conclusion, the water level series are moderately or highly similar in most of the stations.

Keywords – Manual water level measurement; automated water level measurement.

INTRODUÇÃO

A gestão e o planejamento dos recursos hídricos dependem de uma adequada quantificação da disponibilidade hídrica e, por consequência, do conhecimento detalhado das variáveis que compõem o ciclo hidrológico (Monte-Mor, 2012). A realização de estudos hidrológicos surge da necessidade de se entender quais são essas variáveis, como elas se comportam e interagem, entre outros fatores. Para que esses estudos adquiram um caráter quantitativo, é feita a aferição, em laboratório ou em campo, de algumas das grandezas envolvidas.

¹ Bolsista do Instituto Tecnológico SIMEPAR, Centro Politécnico da UFPR, 81531-980 Curitiba – PR. E-mail: thaine@simepar.br

² Pesquisador do Instituto Tecnológico SIMEPAR, Centro Politécnico da UFPR, 81531-980 Curitiba – PR. E-mail: angelo@simepar.br

Sendo assim, dados observados constituem-se como uma importante ferramenta de calibração de estudos hidrológicos, por lhes conferirem confiabilidade de representação de fenômenos reais (Andrade *et al.*, 2013). Nesse âmbito, as principais grandezas são a velocidade, o nível d'água (cota) e a vazão, essa última podendo ser adquirida a partir do registro de cota através de uma curva de descarga.

Considerando apenas dados de cota, esses são obtidos através de medições convencionais ou automáticas realizadas em estações de monitoramento. A diferença entre esses métodos de medição é que no primeiro a coleta dos dados é realizada por um operador que lê e registra de uma régua limnimétrica o valor da cota, enquanto no segundo os dados são medidos por sensores eletrônicos (sensores de célula de pressão ou sistemas de boia e contrapeso), conectados a um sistema de armazenamento de dados (*Datalogger*).

A utilização de sensores automáticos permite a obtenção de séries de dados com maior frequência de registros, fornecendo uma base de informação mais ampla e completa para estudos hidrológicos. Por outro lado, com o uso de tais equipamentos o monitoramento está sempre sujeito à ocorrência de falhas (por mau funcionamento ou interferências naturais) que podem gerar incertezas superiores às esperadas e ocasionar registros errôneos do nível. Por esse e outros motivos, a automatização da coleta de dados fluviométricos vem gerando discussões sobre a sua confiabilidade.

O trabalho realizado por Braga (2005) mostrou que os sensores de nível podem se degradar ao decorrer do tempo (alterando a forma com que relacionam as grandezas medidas e as grandezas desejadas) e que fatores ambientais podem influenciar nas medições. Entretanto, não são encontrados na literatura mais estudos analisando essa questão. É neste contexto que se insere esta pesquisa: de identificar e quantificar as diferenças entre séries de dados de sensores automáticos e séries de dados de medições convencionais, aplicando testes estatísticos para melhor caracterizar a similaridade dessas séries.

MATERIAIS E MÉTODOS

O trabalho envolveu 18 pontos de medição de cota fluviométrica localizados na Bacia do Rio Iguaçu (Figura 1), compreendendo o período de 1998 a 2010. Em cada ponto foram obtidas duas séries de dados:

- Série automática: medida a cada 15 minutos, oriunda da rede de monitoramento automático do Instituto Tecnológico SIMEPAR;
- Série convencional: medida às 07:00 horas e às 17:00 horas diariamente, obtida do banco de dados da Agência Nacional de Águas (ANA).

Devido ao caráter direto das medições para compor a série convencional, supôs-se que tal série está sempre correta, embora não tenham sido realizadas investigações profundas de sua qualidade. A partir disso, refere-se às diferenças entre as séries automática e convencional como a série completa de erro dos dados medidos por sensores.

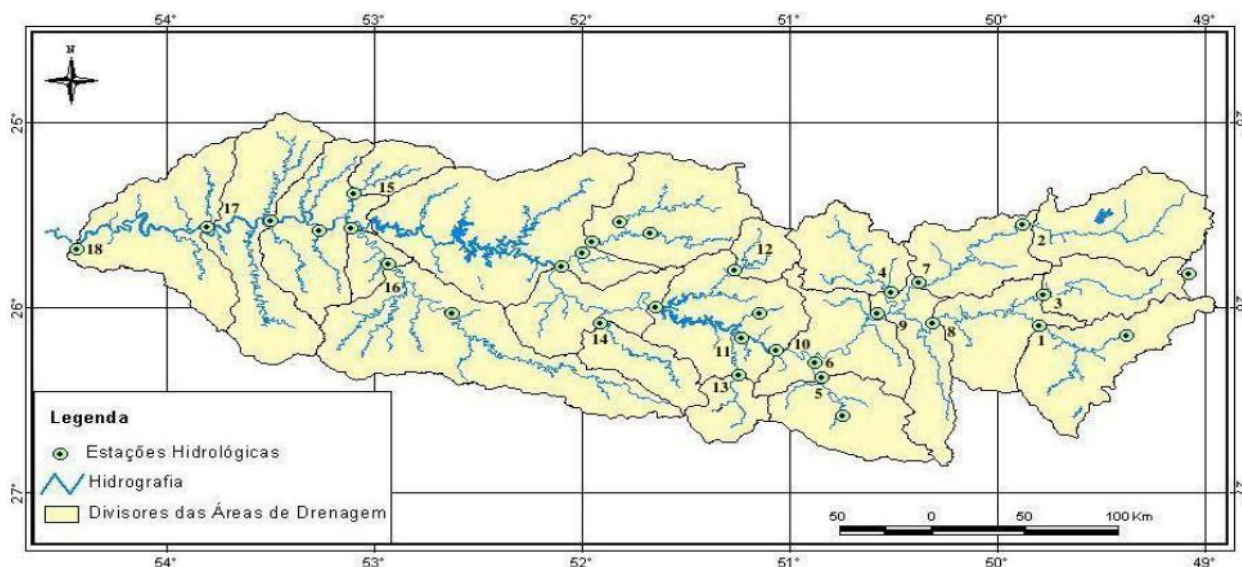


Figura 1 - Mapa da Bacia do Rio Iguaçu contendo as estações hidrológicas da região, entre as quais as estações numeradas são as utilizadas nesse estudo, de montante a jusante

Eventualmente as leituras de régua não são realizadas precisamente no horário definido, podendo ocorrer diferenças significativas entre os registros das duas séries em determinadas situações. Por este motivo, para o cálculo do erro, optou-se por escolher o registro da série automática com menor diferença em relação ao registro convencional, que esteja em um intervalo de uma hora antes ou depois do suposto horário da leitura.

O procedimento de análise foi dividido em duas etapas: análise estatística das séries de erro e análise correlacional das séries de cota fluviométrica. Para a primeira etapa, utilizou-se o fator da tolerância, que busca eliminar a influência de erros não significativos no cálculo dos indicadores, considerando as incertezas relativas a cada tipo de medição. Considerou-se que são aceitáveis até 2 cm de erro nas medições dos sensores automáticos e até 2 cm de erro nas leituras convencionais (causadas por oscilações na superfície da água e erro de paralaxe devido à distância entre o observador e a régua). Desse modo definiu-se uma tolerância de 4 cm, ou seja, foram descartados os erros absolutos de magnitude iguais ou inferiores a 4 cm, quando a tolerância foi usada.

A série completa de erro foi dividida em quatro sub-períodos (1998 a 2001; 2002 a 2004; 2005 a 2007; e 2008 a 2010). Para cada um foram feitos os cálculos do Erro Absoluto Médio (EAM), Erro Absoluto Médio com tolerância (EAM_{tol}), Frequência com tolerância (F_{tol}) e Frequência positiva com tolerância (F_{tol}^+), de acordo com as seguintes equações:

$$E_i = H_{AUTO,i} - H_{CONV,i} \quad (1)$$

$$EAM = \sum_{i=1}^N |E_i| / N \quad (2)$$

$$EAM_{tol} = \sum_{i=1}^N \left\{ \begin{array}{ll} |E_i| & , se |E_i| > 0,04 \\ 0 & , se |E_i| \leq 0,04 \end{array} \right\} / N_{tol} \quad (3)$$

$$F_{tol} = \sum_{i=1}^N \left\{ \begin{array}{ll} 1 & , se |E_i| > 0,04 \\ 0 & , se |E_i| \leq 0,04 \end{array} \right\} \times 100 / N \quad (4)$$

$$F_{tol}^+ = \sum_{i=1}^N \left\{ \begin{matrix} 1 & , se E_i > 0,04 \\ 0 & , se E_i \leq 0,04 \end{matrix} \right\} \times 100 / N \quad (5)$$

Onde $H_{AUTO,i}$ é o i -ésimo registro da série automática, $H_{CONV,i}$ é o i -ésimo registro da série convencional, E_i é o erro da i -ésima comparação entre os registros, N é o número de dados de erro e N_{tol} é o número de dados de erro com magnitude acima da tolerância. A Tabela 1 exemplifica os resultados obtidos com esses indicadores.

Tabela 1 - Estatísticas de erro e índices de correlação em 4 sub-períodos distintos da Estação 18

Período	EAM (cm)	EAM _{tol} (cm)	F _{tol} (%)	F _{tol} ⁺ (cm)	R
1998-2001	6,6	17,6	32	17	0,91
2002-2004	1,8	7,1	7	6	0,99
2005-2007	3,2	9,9	25	7	0,99
2008-2010	6,0	10,8	47	26	0,97

Novamente usando a série completa de erro, foi gerado para cada estação um gráfico da distribuição cronológica dos erros em escala mensal, objetivando uma melhor visualização do comportamento temporal desses. Como pode ser visto na Figura 2, foi calculada uma série de:

- Erro absoluto mensal total: calculado segundo a Equação (3), onde N é o número de dados de erro por mês e N_{tol} é o número de dados de erro acima da tolerância por mês;
- Erro absoluto mensal ponderado: multiplicação do erro absoluto mensal total pela frequência de erros acima da tolerância, calculada de acordo com a Equação (4), onde N é o número de dados de erro por mês.

Com base nessas séries mensais foi calculada a estatística de teste de Spearman (T), conforme o método apresentado por Naghettini e Pinto (2007). O objetivo deste é testar a hipótese de estacionariedade, ou seja, a hipótese de que a série de erro não apresenta variação temporal, com um nível de confiança de 95%. Nesse teste, se T for superior a 1,96, a hipótese de estacionariedade é rejeitada e, se T for igual ou inferior a 1,96, esta é mantida.

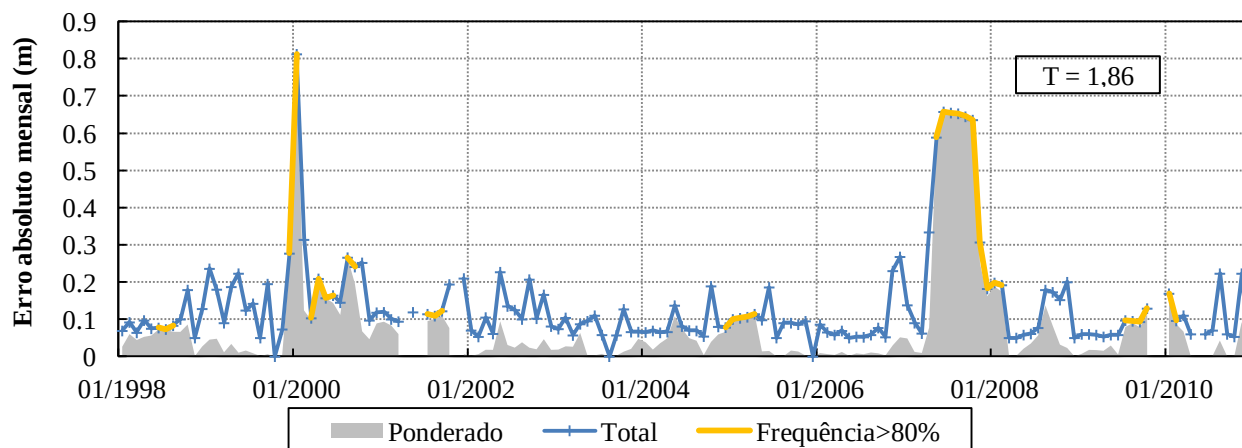


Figura 2 - Gráfico do erro absoluto mensal total e ponderado da Estação 15

Na segunda etapa, buscou-se quantificar o comportamento simultâneo das séries de cota, através do cálculo do coeficiente de Pearson (R), da construção do diagrama de dispersão e da aplicação do método de regressão linear sobre essas séries (Figura 3). O coeficiente de Pearson também foi calculado para os mesmos quatro sub-períodos da etapa anterior.

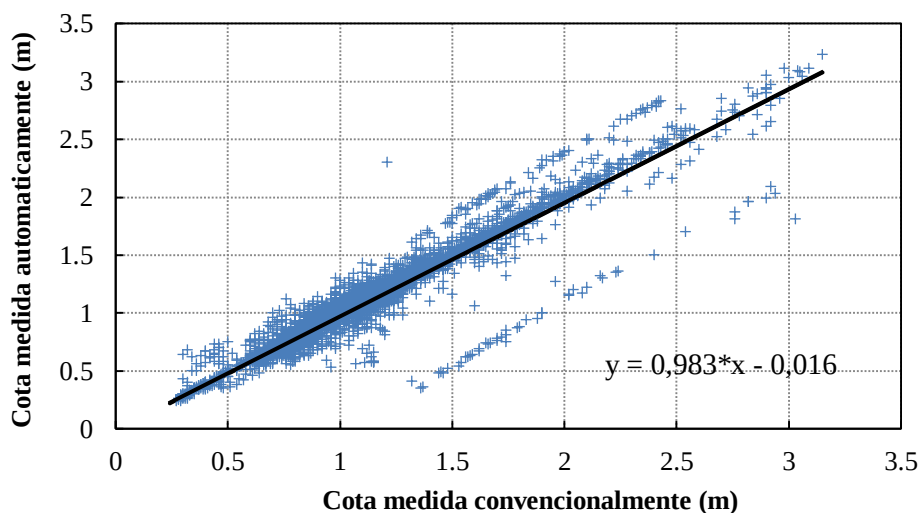


Figura 3 – Diagrama de dispersão com a equação da reta da regressão linear da Estação 18

RESULTADOS E DISCUSSÃO

Os resultados obtidos da análise estatística das séries de erro e da análise correlacional das séries de cota fluviométrica geraram diversas informações, das quais, aquelas relativas ao EAM , EAM_{tol} e F_{tol} foram resumidas nos gráficos da Figura 4, Figura 5 e Figura 6.

Foi observado que 50% dos postos tem seu maior valor de EAM_{tol} ocorrendo no primeiro sub-período, entre 1998 e 2001, assim como 38,9% apontam o maior F_{tol} nesse mesmo intervalo. Entretanto, em nenhum dos quatro intervalos há uma maior ocorrência dos menores valores de F_{tol} . No caso das menores magnitudes do EAM_{tol} , apenas em 11,1% das estações o mínimo ocorre no primeiro sub-período, e em 38,9% dos postos estes acontecem no terceiro sub-período (2005 a 2007). Considerando as estimativas de EAM , as quais não discriminam erros iguais ou abaixo de 4 cm, foi constatado no segundo intervalo (2002 a 2004) o maior valor deste parâmetro em 55,5% das estações, enquanto no último intervalo (2008 a 2010) apenas um máximo foi encontrado. Quanto aos menores valores do índice em questão, estes ocorrem em 16,7% dos postos no primeiro sub-período, enquanto para 38,9% esses foram encontrados no terceiro sub-período.

Na estação 8 verificou-se que em vários momentos os dados convencionais apresentam um comportamento inconsistente, de modo que se optou por descartar os resultados deste posto na contabilidade dos parâmetros estatísticos que representam toda a rede de monitoramento. Assim posto, foi calculada a média dentre todas as estações para os parâmetros EAM_{tol} , F_{tol} e EAM . O resultado está ilustrado na Figura 7. Nessa se percebe que a intensidade e frequência dos diminuem com o passar do tempo, sobretudo comparando os intervalos de 1998 a 2004 e 2005 a 2010. Além disso, os valores de EAM são bastante inferiores aqueles de EAM_{tol} . Não foram detectados padrões de variação temporal nas estimativas de F_{tol}^+ nos valores do coeficiente de Pearson.

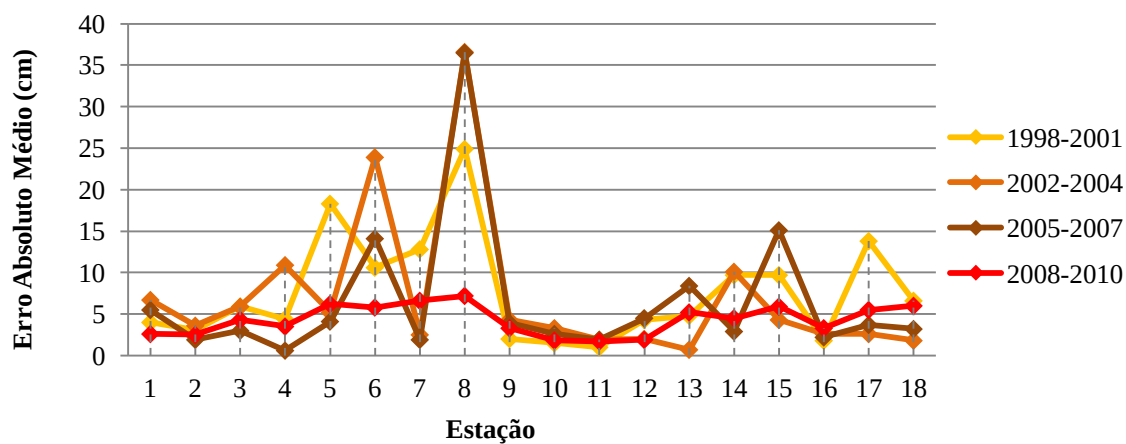


Figura 4 - Gráfico do erro absoluto médio com tolerância por estação em quatro sub-períodos distintos

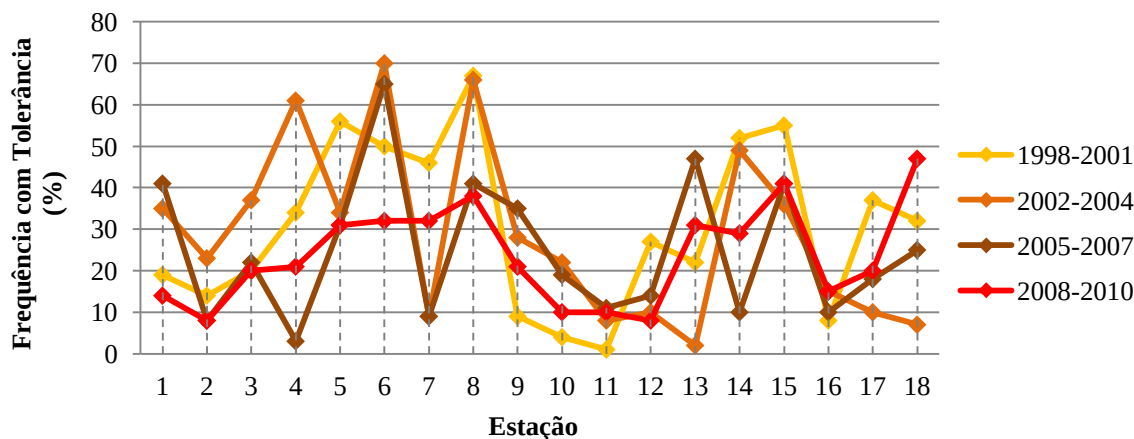


Figura 5 - Gráfico da frequência com tolerância por estação em quatro sub-períodos distintos

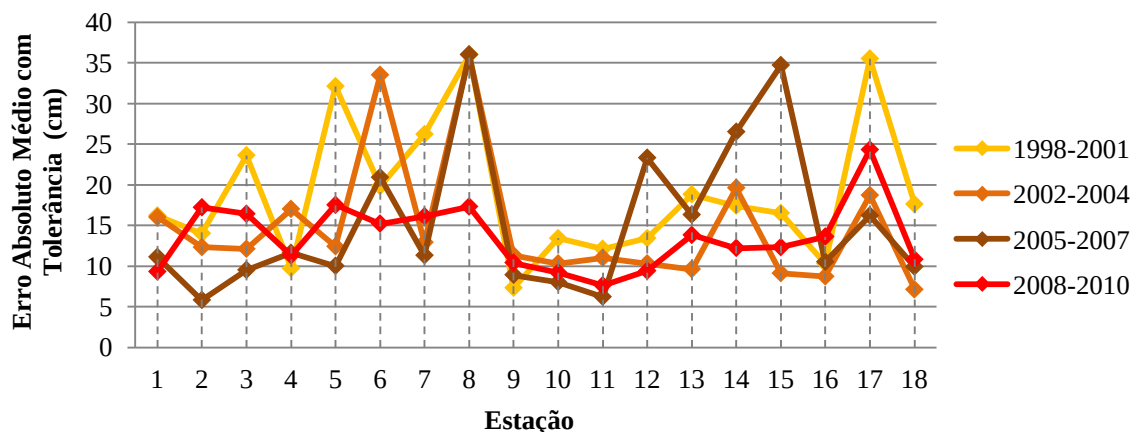


Figura 6 - Gráfico do erro absoluto médio por estação em quatro sub-períodos distintos

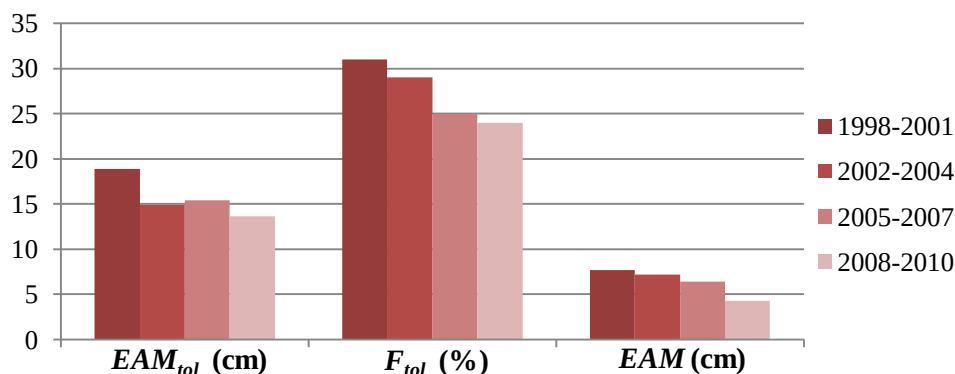


Figura 7 - Gráfico do erro absoluto médio com tolerância, frequência com tolerância e do erro absoluto médio, mediados para todas as estações, em quatro sub-períodos distintos

Considerando o comportamento temporal das séries, as estatísticas de teste de Spearman indicam que em 33,3% dos postos hidrológicos a hipótese de estacionariedade é rejeitada. Foram percebidos, na análise dos gráficos dos erros mensais, alguns padrões de erros que compõem as séries: erro base (erros com maior frequência), picos de erro grandes (superiores a 1 metro), picos de erros moderados (15 cm a 1 m), picos de erro pequenos (5 cm a 15 cm) e blocos de erros (erros praticamente constantes, com duração de 2 meses a 1 ano, sendo encontrados blocos de 15 cm até blocos de 2 m de erro). Na Figura 2, é visível o erro base percorrendo todo o período, assim como um bloco de erro no ano de 2007 e picos pequenos ocorrendo em 1999, 2002 e no final de 2010.

Em 83,3% das estações o erro base é identificável, com valores de até 10 cm e desvio-padrão de 1 a 7 cm. Picos de erro grandes foram detectados em 38,9% das estações, das quais em 33% houve estimativas de F_{tol} associadas superiores a 80%. Os blocos de erro, em geral, também são acompanhados por altos valores do F_{tol} . Pelo menos um bloco foi encontrado em 50% dos postos hidrológicos. Esses blocos elevaram os valores de EAM de seus sub-períodos mais efetivamente que os picos grandes. Todas as estações exibem picos pequenos e moderados ao menos uma vez, porém 50% das estações mostram uma quantidade maior do primeiro tipo e 33,3% do segundo. Avaliando a ausência de erros: 27,8% dos postos somaram até 5 meses sem erros; 44,4% somaram de 16 a 29 meses sem erros; e 27,8% somaram um total de 38 a 63 meses sem erros. Durante as análises por estação, percebeu-se que não há relação entre a quantidade e a intensidade dos erros. Por outro lado, constatou-se que em 22% das estações os erros negativos são predominantes, enquanto em 50% houve predominância positiva.

Na análise correlacional, os coeficientes de Pearson indicam forte correlação entre as séries de cota em todos os postos. Quanto às retas de regressão, é sabido que as séries automática e convencional apresentam máxima similaridade quando os coeficientes angular e linear das retas regredidas são iguais a 1,0 e 0,0, respectivamente. Verificou-se que os coeficientes das retas de regressão são até 11% diferentes desses valores. Os coeficientes lineares são todos negativos, enquanto os coeficientes angulares ficam ligeiramente abaixo de 1,0 em 61,1% dos postos.

No contexto dos resultados expostos, as séries automáticas e convencionais são bastante similares nos postos fluviométricos de número 2, 9, 11 e 16, com índices de correlação quase

perfeitos, erro base de até 6 cm e médias menores que 25% dos valores de F_{tol} . Nos postos 10 e 12 as séries de cota também são similares entre si, apesar de apresentarem um erro base de 6 a 8 cm e coeficientes de Pearson de 0,95. As estações que apresentam pouca similaridade são as de número 6 e 8, nas quais a presença de picos altos e os blocos de erro influenciaram significativamente os indicadores estatísticos. O restante das estações mostrou um comportamento intermediário, nas quais o EAM_{tol} de menor magnitude varia entre 7 e 16 cm e foram detectados estimativas de F_{tol} menores que 39%, assim como altos valores do coeficiente de Pearson.

CONCLUSÃO

Foram realizadas comparações entre os registros de cota obtidos por estações automáticas e por leituras de régua convencionais em 18 postos de monitoramento na bacia do Rio Iguaçu. Os resultados mostraram que, em todo o período analisado de 1998 a 2010, as séries automáticas e convencionais são fortemente correlacionadas, com a primeira superestimando levemente os valores de cota. A estimativa dos erros absolutos, descartando-se diferenças iguais ou menores que 4 cm (tolerância à imprecisão nos registros), exibe valores entre 10 e 30 cm que ocorrem em 27% do tempo. Ainda, esse estudo destaca a heterogeneidade entre as estações, no qual cada série de erro se mostrou como uma sobreposição diferente de padrões de erro, sendo então o grau de similaridade definido individualmente. Concluiu-se que as séries de cota são moderadamente ou bastante similares entre si na maioria dos postos analisados. Ainda, foi averiguado que a magnitude e frequência dos erros decaem com o tempo, do que se pode inferir que os processos relacionados à calibração dos sensores, aprimoramento dos equipamentos de medição e melhora na gestão da rede de monitoramento, vem melhorando a qualidade dos dados registrados em estações automáticas.

A realização do trabalho gerou uma quantidade significativa de informações para o estudo da qualidade dos dados provenientes de sensores automáticos de nível, e evidencia a necessidade de uma investigação mais criteriosa de suas fontes de erro, de um estudo da confiabilidade dos equipamentos em uso e da gestão atual desses, considerando formas de remoção dessas fontes. Estas recomendações certamente auxiliarão na obtenção de séries de nível mais compatíveis com a realidade.

REFERÊNCIAS

- ANDRADE, M. A.; MELLO, C. R. de; BESKOW, S. (2013). Simulação hidrológica em uma bacia hidrográfica representativa dos latossolos da região Alto Rio Grande, MG. *Revista Brasileira de Engenharia Agrícola e Ambiental* v. 17, n. 1, pp. 69-76.
- BRAGA, S. M. (2005). *Análise do potencial de utilização de sensores automáticos hidroambientais: estudo de caso da Bacia do Rio Barigui*. Dissertação de Mestrado. UFPR, 143 p.
- MONTE-MOR, R. C. A. (2012). *Análise de processos hidrológicos em bacias de rios intermitentes no semiárido mineiro*. Tese de Doutorado. UFMG, 307 p.
- NAGHETTINI, M.; PINTO E. J. A. (2007). *Hidrologia Estatística*. CPRM – Serviço Geológico do Brasil, Belo Horizonte - MG, 552 p.