

**IV-066 - INFLUÊNCIA DO TRATAMENTO ESTATÍSTICO DOS DADOS
ABAIXO DO LIMITE DE DETECÇÃO NA CARACTERIZAÇÃO DE METAIS
NOS CURSOS D'ÁGUA DA BACIA HIDROGRÁFICA DO
RIO DAS VELHAS – MG**

Cristiano Christofaro⁽¹⁾

Biólogo - UFMG. Doutor em Saneamento, Meio Ambiente e Recursos Hídricos (DESA/UFMG). Professor Adjunto da Universidade Federal dos Vales do Jequitinhonha e Mucuri - UFVJM.

Mônica M. Diniz Leão

Engenheira Química - UFMG. Doutora em Engenharia de Anti-Poluição (INSA-FR). Professora Associada da Universidade Federal de Minas Gerais- DESA/UFMG.

Endereço⁽¹⁾: UFVJM – Campus II - Depto. Engenharia Florestal - Rodovia MG 367 n. 5000. Alto do Jacuba - Diamantina, MG; CEP 39100-000 – Brasil.. Telefone: +55 (38) 3532-1248. Fax: +55 (38) 3532-1200 - e-mail: cristiano.christofaro@ufvjm.edu.br

RESUMO

As inerentes limitações dos métodos analíticos de mensuração fazem com que dados de qualidade da água freqüentemente apresentem observações descritas como abaixo de um certo limite de detecção ou dados “censurados à esquerda”. A censura de dados interfere diretamente em quase todos os tipos de análises estatísticas, incluindo o cálculo de parâmetros estatísticos básicos, dentre outros. O objetivo do presente estudo, é avaliar a influência de três métodos distintos de tratamento da censura à esquerda no cálculo da média, mediana e desvio-padrão das concentrações de metais em vinte e nove estações de monitoramento dos cursos d’água da bacia do Rio das Velhas/MG, a fim de permitir a escolha do método mais adequado. Foram utilizados os resultados das análises de 5 metais (As, Cu, Ni, Pb, Zn) coletados de 1998 a 2007 nas 29 estações de amostragem localizadas ao longo do rio das Velhas e afluentes. Três métodos distintos de tratamento da censura das séries de dados foram utilizados: Kaplan-Meier (KM), Estimativa da Máxima Verossimilhança (EMV) e Imputação (ROS). Os valores da mediana e da média calculados pelos três métodos em todas as estações de amostragem apresentaram poucas diferenças. No entanto, o método da EMV apresentou, freqüentemente, os maiores valores de desvio-padrão em todas as séries de metais. Considerando as características dos dados, como percentual de censura e tamanho das amostras, optou-se pela utilização do método ROS para o tratamento da censura dos metais estudados na bacia do Rio das Velhas.

PALAVRAS-CHAVE: Monitoramento, Qualidade da Água, Dados Censurados, Estatística, Metais, Rio das Velhas.

INTRODUÇÃO

Diversos problemas são enfrentados na aplicação de métodos estatísticos em estudos de poluição ambiental. Um desses problemas está associado às inerentes limitações dos métodos analíticos de mensuração, já que dados de exposição freqüentemente apresentam observações descritas como abaixo de um certo limite de detecção. Do ponto de vista estatístico, dados com registros abaixo de um certo limite são denominados “censurados à esquerda” ou simplesmente censurados (Helsel, 1990). Outros termos aplicados para esses limites na literatura incluem “valor crítico”, “limite de detecção do método” ou apenas “limite de detecção” (Helsel, 2005).

A censura de dados interfere diretamente em quase todos os tipos de análises estatísticas, incluindo: parâmetros estatísticos básicos (e.g. média, desvio-padrão, etc.) (Kroll & Stedinger, 1996; Singh & Nocerino, 2002; Antweiler & Taylor, 2008), intervalos de confiança (Singh & Nocerino, 2002; Zhao & Frey, 2004); testes de hipóteses (Helsel, 2005), ajuste de distribuições de probabilidade (Govaerts, 2001; Zolezzi et al. 2005), correlações (Newton & Rudel, 2007), análises de regressão e tendências (Hopke et al., 2001). Dependendo do método utilizado no seu tratamento, os resultados podem sofrer alterações consideráveis, tendo sua interpretação prejudicada. Apesar desses problemas, Helsel (2005) enfatiza que os dados censurados não devem ser eliminados da série estudada pois, nessas situações, distorções ainda piores podem ser geradas.

Uma vez que a presença de dados censurados prejudica a utilização dos testes estatísticos e que esses dados não devem ser eliminados do estudo, técnicas específicas devem ser utilizadas para minimizar a interferência negativa das observações censuradas (Helsel, 1990). Atualmente, considera-se que os métodos de tratamento da censura estão associados a pelo menos quatro abordagens básicas (Helsel, 2005; Baccarelli, 2005): substituição, métodos paramétricos, métodos robustos e métodos não-paramétricos. Cada um desses métodos apresenta vantagens e limitações, assim como situações e critérios específicos para sua aplicação.

O objetivo do presente estudo, é avaliar a influência de três métodos distintos de tratamento da censura à esquerda (e.g. abaixo do limite de detecção) no cálculo da média, mediana e desvio-padrão das concentrações de metais em vinte e nove estações de monitoramento dos cursos d'água da bacia do Rio das Velhas/MG, a fim de possibilitar a escolha do método mais adequado, de acordo com as características dos dados da bacia.

METODOLOGIA

Área de Estudo

A bacia do rio das Velhas está localizada na região central do Estado de Minas Gerais, entre as latitudes 17° 15'S e 20° 25'S e longitudes 43° 25'W e 44° 50'W (Figura 1). Apresenta uma forma alongada na direção norte-sul e corresponde à Unidade de Planejamento e Gestão de Recursos Hídricos SF5 (São Francisco 5). O rio das Velhas tem sua nascente principal na cachoeira das Andorinhas, Município de Ouro Preto, numa altitude de aproximadamente 1.500 m. Toda a bacia compreende uma área oficial de 29.173 Km², onde estão localizados 51 municípios que abrigam uma população de aproximadamente 4,8 milhões de habitantes (destes, aproximadamente 89% residem em distritos e municípios integralmente inseridos na bacia). O rio das Velhas encontra-se com o rio São Francisco em Barra do Guaicuí, Distrito de Várzea da Palma, numa altitude de 478m (IGAM 2008).

Desde 1997, a bacia hidrográfica do Rio das Velhas vem sendo monitorada no âmbito do programa “Águas de Minas” (IGAM, 2008). Nesse programa, 29 (vinte e nove) estações de amostragem são utilizadas para o monitoramento de diversos metais, incluindo o Cobre, na bacia do Rio das Velhas (Figura 1), com frequência de amostragem trimestral ou semestral. Para a análise de tendência executada no presente estudo foram utilizados os dados do Cobre coletados de 1998 a 2007 nas 29 estações de amostragem localizadas ao longo do rio das Velhas e afluentes.

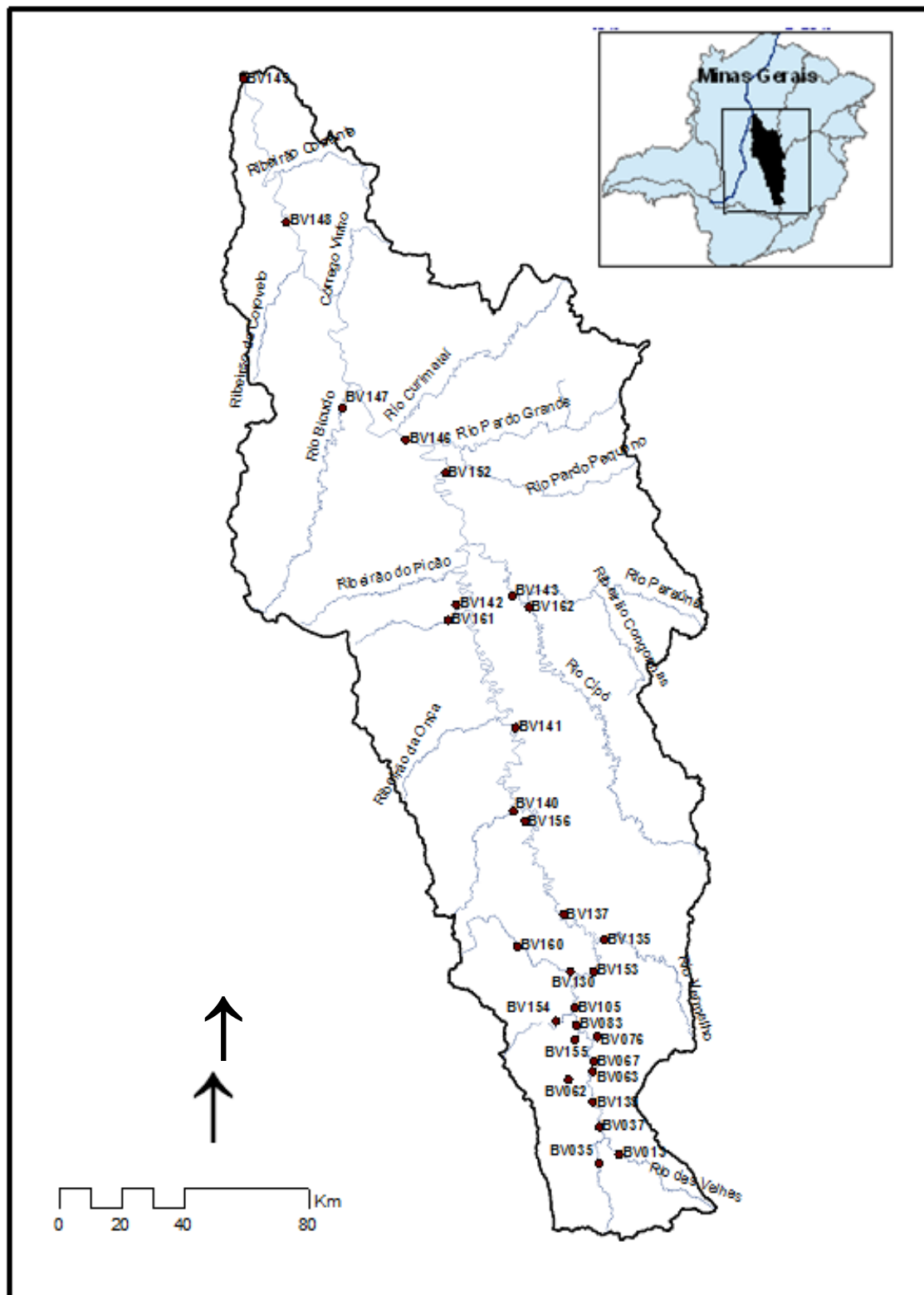


Figura 1: Localização em Minas Gerais e disposição das estações de amostragem do Cobre na bacia hidrográfica do Rio das Velhas – MG, no âmbito do programa “Água de Minas”. Elaborado a partir dos dados disponíveis em www.igam.mg.gov.br e IGAM (2008).

Análises estatísticas

No presente estudo foram utilizados os resultados das análises de 5 metais (As, Cu, Ni, Pb, Zn) coletados de 1998 a 2007 nas 29 estações de amostragem localizadas ao longo do rio das Velhas e afluentes. Três métodos distintos de tratamento da censura das séries de dados foram utilizados (Helsel, 2005): Kaplan-Meier (KM), Estimativa da Máxima Verossimilhança (EMV) e Imputação (ROS).

O método Kaplan-Meier (KM) foi formulado para incluir dados com múltiplos limites de detecção, não requerendo a especificação de uma distribuição de probabilidade (Kleinbaum & Klein, 2005). Na análise KM, um percentil é gerado para cada observação detectada, iniciando do maior valor detectado, seguindo para os demais, tendo como base o número de valores não detectados acima e abaixo de cada observação. Apesar de não serem calculados os percentis para os dados não detectados, os valores censurados influenciam os percentis das observações detectadas. O método, originalmente utilizado em análises de sobrevivência, com censura à direita, pode ser adaptado para tratamento da censura à esquerda após a "inversão" dos dados (Helsel, 2005).

Já no método de Estimativa de Máxima Verossimilhança (EMV), uma distribuição de probabilidades ajustada aos dados é estimada com base nos dados observados acima dos limites de detecção e o percentual de dados abaixo do limite de detecção. Ao considerar o percentual de censura dos dados, o método garante o aproveitamento das informações associadas aos dados censurados (Helsel, 2006). A técnica da EMV tende a gerar pouco desvio nos casos em que as observações não-censuradas apresentam um ajuste satisfatório à distribuição presumida e para amostras suficientemente grandes (Helsel, 2005).

O método de imputação, também denominado "robusto", consiste no preenchimento dos valores censurados ou perdidos sem a determinação de valores repetidos. Nessa abordagem, uma distribuição de probabilidade assimétrica é ajustada ao logaritmo dos dados acima do limite de detecção, por meio de técnicas de plotagem (gráfico QuantilQuantil). Essa distribuição é então utilizada para extrapolar os valores censurados, que passam a fazer parte da amostra. Diferente dos métodos paramétricos, a distribuição é utilizada apenas para a geração de valores das amostras censuradas (Helsel, 2005).

A partir do tratamento da censura, a mediana, a média e o desvio-padrão de cada metal, em cada ponto de monitoramento, foram calculados e comparados entre si, de modo a se caracterizar a influência de cada método. Os tratamentos da censura nos dados foram realizados no programa R (v. 2.8.1) (R Development Core Team, 2008), utilizando o pacote estatístico "NADA" (Lee, 2007).

RESULTADOS E DISCUSSÃO

Os valores da média, mediana e desvio-padrão do Arsênio, Cobre, Níquel, Chumbo e Zinco, por estação de monitoramento, foram calculados considerando três métodos distintos: Kaplan-Meier, Imputação ou Regressão Robusta (ROS) e Estimativa de Máxima Verossimilhança (EMV) (Tabela 1).

Tabela 1: Número de amostras e percentual de valores inferiores ao limite de detecção, por estação de monitoramento no período analisado (1998 a 2007).

Estação	As		Cu		Ni		Pb		Zn	
	n	%	n	%	n	%	n	%	n	%
BV013	21	47,6	20	55,0	21	52,4	36	58,3	20	30,0
BV035	36	55,6	33	57,6	36	58,3	36	58,3	36	30,6
BV037	36	44,4	33	57,6	36	61,1	20	50,0	36	38,9
BV062	40	2,5	37	43,2	40	52,5	40	70,0	24	8,3
BV063	40	5,0	37	40,5	40	35,0	40	65,0	36	8,3
BV067	40	7,5	33	45,5	36	52,8	23	60,9	24	16,7
BV076	20	30,0	20	60,0	20	55,0	20	60,0	39	25,6
BV083	21	14,3	22	31,8	36	41,7	36	55,6	39	7,7
BV105	21	4,8	21	19,0	36	33,3	36	47,2	39	0,0
BV130	20	30,0	21	57,1	20	50,0	23	56,5	23	13,0
BV135	20	40,0	19	73,7	20	65,0	20	70,0	20	30,0
BV137	40	0,0	20	20,0	40	30,0	20	45,0	39	5,1
BV139	36	41,7	33	54,5	40	57,5	20	75,0	36	25,0
BV140	20	45,0	19	78,9	20	65,0	36	75,0	20	30,0
BV141	40	2,5	32	50,0	40	35,0	20	55,0	21	23,8
BV142	40	10,0	20	55,0	40	47,5	20	50,0	21	19,0
BV143	20	55,0	21	71,4	36	66,7	23	78,3	23	47,8
BV146	40	7,5	20	55,0	20	50,0	20	50,0	21	19,0
BV147	20	45,0	19	73,7	20	75,0	20	75,0	20	45,0
BV148	36	2,8	20	60,0	20	60,0	40	65,0	20	25,0
BV149	36	5,6	33	72,7	36	61,1	36	66,7	21	19,0
BV152	40	10,0	19	47,4	20	45,0	20	60,0	20	25,0
BV153	40	2,5	36	25,0	40	45,0	40	50,0	39	5,1
BV154	20	35,0	21	28,6	21	42,9	36	44,4	23	0,0
BV155	20	25,0	37	5,4	20	30,0	40	47,5	39	2,6
BV156	39	5,1	19	42,1	40	30,0	20	50,0	20	15,0
BV160	15	46,7	16	56,3	15	66,7	18	66,7	18	11,1
BV161	31	61,3	17	64,7	31	87,1	15	86,7	15	20,0
BV162	15	53,3	14	85,7	31	80,6	15	80,0	15	33,3
Média		25,4		51,3		52,8		61,1		20,0
Censura										

A tabela 1 indica, para cada metal, em cada estação de amostragem, o número de amostras e o percentual de censura verificado. As menores censuras médias nas estações de amostragem foram verificadas para o Zinco (20%) e para o Arsênio (25,4%). Nenhuma estação de amostragem apresentou censura maior do que 50% dos dados para o Zinco e apenas quatro apresentaram censura maior do que 50% para o Arsênio. Os demais metais apresentaram censura média nas estações de amostragem de 51,3% (Cu), 52,8% (Ni) e 61,1% (Pb). Uma estação de amostragem para o Cu e duas para Ni e Pb apresentaram censura maior do que 80%. Os valores da média, desvio-padrão e mediana calculados para os metais em cada estação de amostragem podem ser visualizados na tabela 2.

Tabela 2: Mediana (Med), média, e desvio-padrão (DP), considerando os métodos Kaplan-Meier (KM), Imputação (ROS) e Estimativa da Máxima Verossimilhança (EMV), para tratamento da censura nas estações de monitoramento da bacia do Rio das Velhas.

Estação	Trat.	As			Cu			Ni			Pb			Zn		
		Med	Média	DP	Med	Média	DP	Med	Média	DP	Med	Média	DP	Med	Média	DP
BV013	KM	0,001	0,004	0,007	-	0,006	0,005	-	0,012	0,011	-	0,015	0,022	0,020	0,033	0,020
	ROS	0,001	0,003	0,006	0,003	0,005	0,005	0,004	0,010	0,012	0,004	0,011	0,023	0,020	0,029	0,023
	EMV	0,000	0,010	0,220	0,004	0,006	0,006	0,004	0,011	0,030	0,004	0,012	0,036	0,023	0,030	0,026
BV035	KM	-	0,009	0,026	-	0,017	0,024	-	0,012	0,017	-	0,016	0,021	0,020	0,058	0,097
	ROS	0,000	0,008	0,026	0,004	0,015	0,024	0,002	0,011	0,017	0,005	0,014	0,022	0,022	0,054	0,098
	EMV	0,000	0,086	33,734	0,003	0,025	0,192	0,003	0,015	0,079	0,004	0,019	0,092	0,029	0,051	0,076
BV037	KM	0,001	0,012	0,041	-	0,016	0,024	-	0,017	0,028	-	0,014	0,011	0,020	0,053	0,072
	ROS	0,001	0,012	0,040	0,003	0,015	0,025	0,004	0,015	0,028	0,008	0,013	0,011	0,020	0,047	0,075
	EMV	0,000	0,043	3,770	0,003	0,024	0,184	0,002	0,030	0,377	0,006	0,015	0,036	0,024	0,049	0,088
BV062	KM	0,071	0,085	0,068	0,004	0,016	0,027	-	0,011	0,025	-	0,007	0,005	0,050	0,068	0,060
	ROS	0,071	0,085	0,068	0,004	0,015	0,027	0,003	0,010	0,025	0,003	0,005	0,006	0,050	0,068	0,060
	EMV	0,054	0,125	0,260	0,005	0,015	0,039	0,004	0,009	0,022	0,003	0,005	0,009	0,050	0,069	0,068
BV063	KM	0,013	0,026	0,061	0,005	0,017	0,032	0,006	0,018	0,042	-	0,012	0,020	0,040	0,071	0,142
	ROS	0,013	0,026	0,061	0,005	0,016	0,032	0,006	0,017	0,042	0,002	0,009	0,021	0,040	0,070	0,142
	EMV	0,010	0,030	0,090	0,006	0,016	0,040	0,006	0,016	0,036	0,003	0,010	0,035	0,043	0,060	0,058
BV067	KM	0,008	0,012	0,011	0,004	0,016	0,023	-	0,014	0,020	-	0,010	0,005	0,040	0,047	0,035
	ROS	0,009	0,012	0,011	0,004	0,015	0,023	0,004	0,013	0,021	0,006	0,008	0,006	0,040	0,046	0,036
	EMV	0,007	0,016	0,036	0,005	0,018	0,055	0,004	0,017	0,072	0,004	0,008	0,013	0,037	0,046	0,035
BV076	KM	0,002	0,005	0,009	-	0,007	0,007	-	0,011	0,011	-	0,009	0,003	0,030	0,036	0,020
	ROS	0,002	0,005	0,009	0,003	0,006	0,008	0,005	0,009	0,012	0,006	0,007	0,004	0,030	0,034	0,022
	EMV	0,001	0,008	0,050	0,004	0,006	0,007	0,004	0,010	0,026	0,004	0,006	0,007	0,026	0,035	0,031
BV083	KM	0,009	0,010	0,008	0,007	0,014	0,014	0,005	0,012	0,013	-	0,031	0,130	0,060	0,084	0,066
	ROS	0,009	0,010	0,008	0,008	0,013	0,014	0,005	0,011	0,014	0,003	0,028	0,128	0,060	0,084	0,066
	EMV	0,005	0,018	0,065	0,008	0,014	0,021	0,005	0,011	0,022	0,004	0,016	0,067	0,062	0,087	0,088
BV105	KM	0,014	0,017	0,012	0,013	0,017	0,017	0,006	0,013	0,021	0,005	0,016	0,042	0,080	0,090	0,053
	ROS	0,014	0,017	0,012	0,013	0,016	0,017	0,006	0,012	0,021	0,006	0,015	0,042	0,080	0,090	0,053
	EMV	0,012	0,022	0,034	0,011	0,016	0,017	0,006	0,011	0,018	0,005	0,012	0,027	0,074	0,093	0,069
BV130	KM	0,001	0,004	0,007	-	0,010	0,019	-	0,009	0,011	-	0,013	0,023	0,050	0,085	0,079
	ROS	0,001	0,004	0,007	0,003	0,008	0,019	0,003	0,007	0,011	0,003	0,010	0,023	0,050	0,084	0,079
	EMV	0,001	0,006	0,037	0,004	0,008	0,015	0,004	0,008	0,013	0,004	0,010	0,022	0,053	0,090	0,122
BV135	KM	0,001	0,002	0,003	-	0,006	0,003	-	0,010	0,007	-	0,009	0,010	0,030	0,036	0,020
	ROS	0,001	0,002	0,003	0,003	0,004	0,004	0,003	0,007	0,008	0,001	0,006	0,010	0,030	0,034	0,022
	EMV	0,000	0,002	0,006	0,002	0,004	0,005	0,002	0,007	0,022	0,002	0,007	0,017	0,026	0,035	0,032
BV137	KM	0,029	0,033	0,023	0,011	0,020	0,020	0,006	0,011	0,010	0,007	0,016	0,014	0,070	0,097	0,100
	ROS	0,029	0,033	0,023	0,011	0,019	0,020	0,007	0,011	0,010	0,008	0,014	0,015	0,070	0,097	0,100
	EMV	0,023	0,039	0,053	0,012	0,020	0,029	0,007	0,012	0,016	0,007	0,017	0,038	0,067	0,099	0,106
BV139	KM	0,000	0,005	0,017	-	0,014	0,023	-	0,009	0,013	-	0,016	0,003	0,030	0,047	0,056
	ROS	0,001	0,005	0,017	0,003	0,012	0,024	0,003	0,008	0,014	0,010	0,011	0,006	0,030	0,044	0,057
	EMV	0,001	0,008	0,121	0,004	0,014	0,055	0,003	0,009	0,023	0,002	0,009	0,051	0,028	0,043	0,048
BV140	KM	0,001	0,002	0,002	-	0,006	0,003	-	0,006	0,004	-	0,006	0,002	0,030	0,047	0,044
	ROS	0,001	0,002	0,002	0,002	0,003	0,004	0,002	0,004	0,005	0,002	0,004	0,003	0,030	0,044	0,046
	EMV	0,000	0,002	0,011	0,002	0,004	0,006	0,003	0,005	0,006	0,003	0,004	0,003	0,028	0,046	0,059

(Continua)

Tabela 2 (Continuação)

Estação	Trat.	As			Cu			Ni			Pb			Zn		
		Med	Média	DP	Med	Média	DP	Med	Média	DP	Med	Média	DP	Med	Média	DP
BV142	KM	0,027	0,043	0,045	-	0,015	0,026	0,004	0,012	0,015	-	0,013	0,017	0,040	0,083	0,087
	ROS	0,028	0,043	0,045	0,004	0,014	0,025	0,005	0,011	0,015	0,005	0,011	0,017	0,040	0,080	0,088
	EMV	0,019	0,099	0,498	0,004	0,019	0,090	0,004	0,012	0,028	0,005	0,012	0,023	0,042	0,091	0,173
BV143	KM	-	0,001	0,001	-	0,006	0,006	-	0,006	0,005	-	0,010	0,004	0,020	0,036	0,030
	ROS	0,001	0,001	0,001	0,001	0,004	0,007	0,002	0,004	0,005	0,003	0,005	0,006	0,020	0,030	0,034
	EMV	0,000	0,002	0,008	0,002	0,005	0,009	0,003	0,005	0,008	0,002	0,005	0,014	0,018	0,032	0,048
BV146	KM	0,016	0,024	0,026	-	0,008	0,007	-	0,008	0,006	-	0,010	0,006	0,030	0,047	0,050
	ROS	0,017	0,024	0,025	0,004	0,007	0,007	0,004	0,007	0,007	0,006	0,009	0,007	0,030	0,045	0,051
	EMV	0,012	0,043	0,150	0,004	0,008	0,013	0,004	0,007	0,011	0,005	0,009	0,011	0,031	0,045	0,046
BV147	KM	0,000	0,004	0,012	-	0,006	0,004	-	0,009	0,005	-	0,007	0,003	0,020	0,034	0,031
	ROS	0,000	0,004	0,012	0,002	0,004	0,005	0,003	0,005	0,006	0,003	0,004	0,004	0,020	0,027	0,034
	EMV	0,000	0,003	0,030	0,002	0,004	0,006	0,001	0,005	0,020	0,003	0,004	0,005	0,017	0,029	0,039
BV148	KM	0,015	0,021	0,024	-	0,008	0,007	-	0,010	0,011	0,002	0,006	0,007	0,030	0,045	0,032
	ROS	0,015	0,021	0,024	0,003	0,007	0,007	0,003	0,008	0,012	0,003	0,006	0,007	0,035	0,043	0,034
	EMV	0,011	0,031	0,082	0,004	0,007	0,013	0,003	0,010	0,034	0,003	0,006	0,011	0,030	0,045	0,049
BV149	KM	0,013	0,019	0,024	-	0,012	0,019	-	0,009	0,012	-	0,009	0,008	0,030	0,051	0,043
	ROS	0,013	0,019	0,024	0,001	0,009	0,020	0,002	0,007	0,012	0,003	0,006	0,009	0,030	0,049	0,044
	EMV	0,010	0,027	0,066	0,001	0,015	0,190	0,003	0,007	0,019	0,003	0,006	0,011	0,034	0,051	0,059
BV152	KM	0,017	0,028	0,028	0,005	0,010	0,008	0,006	0,012	0,009	-	0,011	0,007	0,030	0,056	0,056
	ROS	0,019	0,028	0,028	0,005	0,009	0,008	0,007	0,011	0,009	0,006	0,009	0,008	0,030	0,053	0,057
	EMV	0,013	0,053	0,202	0,005	0,010	0,014	0,005	0,012	0,022	0,004	0,009	0,020	0,033	0,056	0,077
BV153	KM	0,013	0,024	0,065	0,010	0,018	0,021	0,005	0,012	0,016	-	0,011	0,009	0,070	0,100	0,086
	ROS	0,014	0,024	0,065	0,011	0,017	0,021	0,006	0,011	0,017	0,005	0,009	0,010	0,070	0,100	0,086
	EMV	0,011	0,022	0,039	0,010	0,017	0,025	0,005	0,012	0,025	0,005	0,009	0,014	0,071	0,103	0,109
BV154	KM	0,001	0,002	0,004	0,008	0,011	0,008	0,006	0,008	0,010	0,007	0,013	0,021	0,090	0,096	0,039
	ROS	0,001	0,002	0,004	0,008	0,011	0,008	0,006	0,007	0,010	0,007	0,011	0,022	0,090	0,096	0,039
	EMV	0,001	0,003	0,010	0,008	0,011	0,010	0,005	0,007	0,008	0,006	0,010	0,016	0,088	0,096	0,040
BV155	KM	0,001	0,002	0,002	0,013	0,020	0,017	0,009	0,014	0,011	0,006	0,017	0,048	0,140	0,148	0,096
	ROS	0,001	0,002	0,002	0,013	0,020	0,017	0,009	0,013	0,011	0,006	0,015	0,048	0,140	0,148	0,096
	EMV	0,001	0,002	0,006	0,014	0,020	0,020	0,008	0,014	0,020	0,005	0,012	0,024	0,124	0,152	0,109
BV156	KM	0,039	0,051	0,081	0,008	0,012	0,011	0,009	0,015	0,016	-	0,014	0,010	0,060	0,076	0,060
	ROS	0,039	0,051	0,081	0,008	0,011	0,011	0,009	0,015	0,017	0,007	0,011	0,011	0,065	0,075	0,060
	EMV	0,023	0,089	0,325	0,006	0,012	0,021	0,008	0,015	0,025	0,006	0,012	0,024	0,051	0,082	0,102
BV160	KM	0,001	0,002	0,004	-	0,006	0,003	-	0,008	0,003	-	0,007	0,003	0,050	0,063	0,044
	ROS	0,001	0,002	0,004	0,003	0,005	0,004	0,004	0,006	0,005	0,003	0,005	0,004	0,050	0,061	0,045
	EMV	0,000	0,004	0,033	0,004	0,005	0,004	0,003	0,005	0,010	0,004	0,005	0,005	0,049	0,061	0,044
BV161	KM	-	0,002	0,004	-	0,008	0,010	-	0,005	0,004	-	0,008	0,002	0,030	0,048	0,041
	ROS	0,000	0,001	0,004	0,002	0,006	0,010	0,000	0,002	0,004	0,001	0,003	0,004	0,030	0,046	0,042
	EMV	0,000	0,002	0,023	0,003	0,007	0,014	0,001	0,003	0,011	0,001	0,003	0,005	0,035	0,046	0,041
BV162	KM	-	0,001	0,001	-	0,005	0,001	-	0,005	0,003	-	0,005	0,001	0,030	0,034	0,024
	ROS	0,000	0,001	0,001	0,001	0,002	0,002	0,001	0,003	0,004	0,002	0,003	0,002	0,030	0,031	0,026
	EMV	0,000	0,001	0,001	0,002	0,002	0,003	0,002	0,003	0,005	0,003	0,004	0,002	0,025	0,031	0,024

De acordo com a tabela 2, os valores da mediana e da média calculados pelos três métodos, em todas as estações de amostragem, apresentam poucas diferenças entre si para todos os metais, exceto para o Arsênio onde médias calculadas pelo método EMV apresentaram grandes discrepâncias em algumas estações de monitoramento (BV013, BV035 - Rio Itabirito e BV037). No entanto, o método da EMV apresentou, frequentemente, os maiores valores de desvio-padrão em todas as séries de metais.

O método de imputação apresenta algumas vantagens em relação ao método EMV no cálculo de médias e desvios-padrão. Uma delas consiste na menor sensibilidade à distribuição de probabilidade ajustada aos dados observados. Dessa forma, apresentam melhores resultados quando aplicados a pequenas amostras (Helsel 1990; Helsel 2005). Helsel (2005) considera que o método EMV pode ser influenciado por amostras pequenas ($n < 50$), desaconselhando sua utilização nesses casos. Singh *et al.* (2006) consideram que outra vantagem desse método consiste no fato dos valores extrapolados não apresentarem valores negativos, uma vez que são gerados a partir de uma distribuição assimétrica.

Estudos demonstram que métodos de imputação podem gerar estatísticas básicas com elevada acurácia, mesmo em situações em que 60-70% dos dados apresentem censura (Huybrechts *et al.*, 2002; Lubin *et al.*, 2004; Helsel, 2005). Singh & Nocerino (2002) consideraram satisfatórias as médias e variâncias calculadas a partir de métodos de imputação quando os *outliers* foram eliminados. No entanto, uma vez que o método gera, por extrapolação, valores para substituir os dados censurados sua aplicação deve se restringir apenas para o cálculo de parâmetros e testes estatísticos que consideram a amostra como um todo, não podendo ser utilizados em estimativas que considerem especificamente cada amostra (e.g. testes não-paramétricos) (Helsel, 2005). Assim, métodos de tratamento da censura por imputação não devem anteceder a aplicação de correlações, tendências e testes de hipóteses não-paramétricos, uma vez que tendem a criar um ordenamento artificial nas amostras censuradas, podendo alterar significativamente os resultados.

O método de imputação é considerado robusto por enfatizar os dados observados (Helsel, 1990). No entanto, Singh *et al.* (2006) consideram que a robustez desse método foi verificada apenas nos casos de dados ajustáveis a distribuições com assimetria moderada e baixos valores do coeficiente de variabilidade, entendendo que não pode haver uma generalização automática para distribuições que não apresentam tais características.

Diversos estudos consideram o KM a técnica mais adequada para estimativa da média, variância e intervalos de confiança associados a séries de dados com censura Helsel, 2005; Singh *et al.*, 2006; Helsel, 2006; Antweiler & Taylor, 2008). As principais vantagens do método KM incluem o fato do método poder lidar, sem nenhuma adaptação adicional, com limites múltiplos, além de sua imunidade à presença de *outliers*, como todo teste não-paramétrico. Um cuidado a ser enfatizado consiste no fato de que, dependendo do percentual de censura dos dados, a média, mas não os percentis, pode apresentar grandes desvios quando calculadas por esse método (Helsel, 2005).

Para Helsel (2005) a escolha do método de tratamento da censura depende do número de amostra dos dados, do percentual de censura, da quantidade de *outliers* e da qualidade do ajuste a alguma distribuição paramétrica. Para amostras menores que 50, recomenda-se a adoção dos métodos KM e imputação. O primeiro nos casos em que até 50% dos valores mensurados apresentem censura. Já a imputação é recomendada nas situações em que 50 a 80% das amostra estejam abaixo do limite de detecção. Para amostras maiores do que 50, os métodos KM e EMV são recomendados, sendo o método de máxima verossimilhança mais adequado para censuras entre 50 a 80% dos valores mensurados.

Assim, uma vez que todas amostras apresentaram $n < 50$, diversas amostras apresentaram censura em 50 a 80% dos dados (tabela 1) e poucos *outliers* foram verificados nas amostras, optou-se por assumir o método ROS como o mais adequado para o cálculo da média, desvio-padrão e mediana das séries dos metais As, Cu, Pb, Ni e Zn monitorados na bacia do Rio das Velhas.

CONCLUSÕES

Os métodos de tratamento dos dados abaixo do limite de detecção apresentaram resultados pouco discrepantes para os metais analisados no que diz respeito ao cálculo da média e da mediana.

O desvio-padrão foi o parâmetro mais influenciado pelo método de tratamento escolhido, apresentando maiores valores após a aplicação do método EMV.

Considerando as características dos dados, como percentual de censura e tamanho das amostras, optou-se pela utilização do método ROS para o tratamento da censura dos metais estudados na bacia do Rio das Velhas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ANTWEILER, R.C. & TAYLOR, H.E., Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: i. summary statistics. *Environmental Science & Technology*, 42(10), 3732-3738 2008.
2. BACCARELLI, A., PFEIFFER, R., CONSONNI, D., PESATORI, A.C., BONZINI, M., PATTERSON JR., D.G., BERTAZZI, P.A. & LANDI, M.T., Handling of dioxin measurement data in the presence of non-detectable values: overview of available methods and their application in the seveso chloracne study. *Chemosphere*, 60(7), 898-906 2005.
3. GOVAERTS, B.; BECK, B.; LECOUTRE, E.; LE BAILLY, C.; EECKAUT, P. V.; From monitoring data to regional distributions: a practical methodology applied to water risk assessment, Discussion paper 0127, Institut de Statistique, Université Catholique de Louvain, Belgium. 2000.
4. HELSEL, D. R. *Less than Obvious. Statistical Treatment of Data below the Detection Limit*. *Environ. Sci. Technol.* 24, 1766-1774, 1990.
5. HELSEL, D. R. *Nondetects and Data Analysis: Statistics for Censored Environmental Data (Statistics in Practice)*. Wiley-Interscience. 2005.
6. HELSEL, D.R., Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere*, 65(11), 2434-2439 2006.
7. HOPKE, P., LIU, C. & RUBIN, D., Multiple imputation for multivariate data with missing and below-threshold measurements: time-series concentrations of pollutants in the arctic. *Biometrics*, 57(1), 22-33 2001.
8. HUYBRECHTS, T., THAS, O., DEWULF, J. & VAN LANGENHOVE, H., How to estimate moments and quantiles of environmental data sets with non-detected observations? a case study on volatile organic compounds in marine water samples. *Journal of Chromatography A*, 975(1), 123-133 2002.
9. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS - IGAM Monitoramento das águas Superficiais na Bacia do Rio das Velhas 1998-2007. Belo Horizonte: IGAM, 161 p., 2008.
10. KLEINBAUM, D.G. & KLEIN, M. *Survival Analysis-A Self-Learning Text*. Second Edition. Springer. 2005.
11. KROLL, C.N. & STEDINGER, J.R., Estimation of moments and quantiles using censored data (paper 95wr03294). *Water Resources Research*, 32(4), 1005-1012 1996.
12. LEE, L., Nada: nondetects and data analysis for environmental data. 2007. Disponibilizado em: <http://www.R-project.org>
13. LUBIN, J., COLT, J., CAMANN, D., DAVIS, S., CERHAN, J., SEVERSON, R., BERNSTEIN, L. & HARTGE, P., Epidemiologic evaluation of measurement data in the presence of detection limits. *Environmental Health Perspectives*, 112(17), 1691-1696 2004.
14. NEWTON, E. & RUDEL, R., Estimating correlation with multiply censored data arising from the adjustment of singly censored data. *Environmental Science & Technology*, 41(1), 221-228. 2007
15. SINGH, A.; MAICHLE, R. & LEE, S.E. *On the Computation of a 95% Upper Confidence Limit of the Unknown Population Mean Based Upon Data Sets with Below Detection Limit Observations*. EPA/600/R-06/022. 2006.
16. SINGH, A. & NOCERINO, J. Robust Estimation of Mean and Variance Using Environmental Data Sets with Below Detection Limit Observations. *Chemometrics and Intelligent Laboratory Systems*, Volume 60, Number 1, 28 pp. 69-86(18). 2002.
17. ZHAO, Y. & FREY, H. C. Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factor. *Risk Analysis* Vol 24, Nº 4. 2004.

18. ZOLEZZI, M., CATTANEO, C. & TARAZONA, J.V., Probabilistic ecological risk assessment of 1,2,4-trichlorobenzene at a former industrial contaminated site. *Environmental Science and Technology*, 39, 2920-2926 2005.