



CLUSTERIZAÇÃO E UTILIZAÇÃO DE AMOSTRAS ESTATÍSTICAS APLICADAS NA OBTENÇÃO DE DADOS DE CAMPO EM SANEAMENTO

Alexandre Arruda Atalla⁽¹⁾

Engenheiro Sanitarista e Ambiental Pleno (UCDB), pós-graduando em Gerenciamento de projetos (PUC-MG)

Bruno Ken Marchezepe⁽²⁾

Engenheiro Civil Júnior e mestrando em Engenharia de Recursos Hídricos pela USP-São Carlos

Matheus Simitan Barros⁽³⁾

Engenheiro Ambiental Júnior e mestrando em Engenharia Hidráulica pela USP-São Carlos

Márcio Donizeti de Barros Júnior⁽⁴⁾

Cientista de Dados Pleno (UNITRI), especialista em Ciência de Dados (Mackenzie) e especialista em business intelligence

Mário Augusto Baggio⁽⁵⁾

Engenheiro Civil Master (UEL), especialista em Engenharia Hidráulica (USP) e Sistemas de Gestão (Instituto Fleming).

Endereço⁽¹⁾: Rua Quintino Bocaiúva, 1.051 – Vila Seixas – Ribeirão Preto - SP - CEP: 14.020-095 - Brasil - Tel: +55 (16) 3610-3069 - e-mail: contato@waterdb.com.br.

RESUMO

A definição de amostras para irregularidades em ligações ativas/inativas e para o perfil de consumo foi realizada através de amostragem estatística, criando clusters através da localização (separando região metropolitana de Vitória e interior do Espírito Santo), consumo e outras características comuns, tendo uma confiabilidade de 95% e precisão amostral de 5%. Os resultados são o número de amostras definidas por grupo (por item) e os resultados provenientes do campo (número de irregularidades, perfil de consumo e IDM médio).

PALAVRAS-CHAVE: Perdas Aparentes, Balanço Hídrico, Mapas de Pressão.

INTRODUÇÃO

Em um Sistema de Abastecimento de Água (SAA), perdas de água, se não controladas, podem acarretar grandes entraves para companhias de saneamento, desde diminuição do faturamento até incapacidade de atender toda a população, gerando problemas sociais e ambientais. A perda de água pode se dividir em duas parcelas: perdas reais, referentes sobretudo a vazamentos no sistema, e perdas aparentes, geradas por consumos não autorizados, ligações clandestinas e erros de submedição em hidrômetros.

Identificar as diversas fontes de perdas de água em sistemas municipais não é tarefa simples. Geralmente, a análise e avaliação de perdas reais e aparentes é realizada percorrendo a rede em busca de vazamentos, e selecionando ao acaso ligações para inspeção (Gumier e Luvizotto Junior, 2007), o que pode ser dispendioso, ineficiente, e muitas vezes, até inviável devido aos recursos disponíveis para tal trabalho e avaliação custo-benefício. A Estatística, então, surge como instrumento para aumento da eficiência no combate às perdas através de técnicas de amostragem. Tais técnicas consistem na seleção de determinado número de amostras que representem com certo nível de confiança a totalidade do sistema.

Este trabalho visou a realização de análise estatística através da análise estatística levando em conta a realidade dos SAA do estado do Espírito Santo, geridos pela Companhia Espírito Santense de Saneamento (CESAN). O trabalho contou com método de determinação de amostras para o estudo de irregularidades em ligações e submedição em hidrômetros, e posteriormente com o cálculo de número de irregularidades em ligações ativas / inativas, perfis de consumo e Índice de Desempenho da Medição (IDM).

OBJETIVOS

O presente trabalho tem como objetivo apresentar o procedimento estatístico de amostragem para determinação de erros em hidrômetros e da taxa de ligações irregulares dentro do contexto de um sistema de abastecimento de água. Mais especificamente, o trabalho visa:

- 1) Estabelecer o tamanho de amostra para avaliação de erro em hidrômetros e existência de fraudes e ligações clandestinas;
- 2) Definir, a partir de análise estatística das amostras previamente estabelecidas, quantidade de fraudes, perfil de consumo e IDM médio



METODOLOGIA

Inicialmente, o trabalho se deu primeiramente através de levantamento de dados preliminares, condições operacionais, verificação da macromedição e micromedição, medições de pressão, em campo e nos sistemas informatizados dos municípios estudados. Em seguida, através dos dados já reunidos, foi realizada a divisão amostral da base de dados para avaliação de fraudes e irregularidades em ligações. Por último, mais uma análise amostral foi feita para determinação de erros de medição de hidrômetros. A descrição da base de dados utilizada e de cada etapa é descrita a seguir.

Dados Utilizados

A base de dados utilizada é de municípios pertencentes ao estado do Espírito Santo, abastecidos pela CESAN. Considera-se assim um total de 52 municípios, correspondentes a cerca de 91 setores de abastecimento ou 89 distritos de medição e controle. Ao todo, os SAA juntos contêm aproximadamente 480 macromedidores, 650 unidades operacionais (reservatórios, elevatórias e válvulas redutoras de pressão), 8800 km de rede, 570 mil ligações ativas, 90 mil ligações inativas e 25 mil ligações clandestinas.

Avaliação de Fraudes e Irregularidades em Ligações

O procedimento amostral para controle de irregularidades em ligações é composto por 6 etapas. Essas etapas estão resumidas a seguir:

1 - Definição de Setores Amostrais. Foram definidos, a partir da técnica estatística de Análise de Agrupamentos, cinco grupos amostrais compostos por municípios com similaridades entre si. As variáveis que resultaram relevantes para a composição dos setores amostrais foram: POPULAÇÃO, CEP, LIGAÇÕES INATIVAS, HIDRÔMETROS, VOLUME NÃO FATURADO. Os grupos foram identificados como grupos A, B, C, D e E.

2 - Seleção dos Tamanhos Amostrais. Para cada grupo selecionou-se a quantidade de cidades a comporem a amostra de acordo com o tamanho de cada grupo e o nível de inspeção requerido, segundo a NBR 5426. No caso, o nível foi considerado padrão regular, como mostra a Tabela 1.

Tabela 1: Escolha da quantidade de cidades amostradas, de acordo com NBR 5426

TAMANHO DE AMOSTRA PARA GRUPOS			
TAMANHO DO GRUPO (Lote)	NÍVEL DE INSPEÇÃO		
	Moderado	Regular	Severo
De 2 a 8	2	2	3
De 9 a 15	2	3	5
De 16 a 25	3	5	8
De 26 a 50	3	8	13
De 51 a 90	5	13	20
De 91 a 150	8	20	32

Para esta etapa, foram escolhidos os parâmetros de confiança e precisão a serem utilizados para a determinação da quantidade de cidades para a amostra. Para os tamanhos amostrais sugeridos neste procedimento foram selecionadas amostras com o padrão regular, com 95% de confiança e 5% de precisão. Em seguida selecionou-se a proporção esperada de ligações com irregularidades de menos de 15%, utilizando-se assim a Tabela 2, com base na metodologia de Arkin (1984), que leva em consideração a faixa de população de ligações que se deseja analisar. O tamanho da amostra encontrada na Tabela 2 deverão ser proporcionalmente distribuídas entre os municípios selecionados para a amostra, de acordo com a população de hidrômetros de cada município. Nos exemplos sugeridos para os grupos de A a E, a amostra encontrada de 192 foi distribuída proporcionalmente à quantidade de hidrômetros de cada município da amostra.

Tabela 2: Proporção esperada de irregularidades em ligações, considerando confiança de 95% e precisão de 5%.

ITEM	POPULAÇÃO DE LIGAÇÕES	AMOSTRA
1	2 a 4	2
2	5 a 7	5
3	8 a 10	8
4	11 a 13	11
5	14 a 16	14
6	17 a 20	17
7	21 a 25	21
8	26 a 30	24
9	31 a 40	30
10	41 a 50	37
11	51 a 60	43
12	61 a 70	49
13	71 a 100	59
14	101 a 200	84
15	201 a 300	109
16	301 a 500	130
17	501 a 1000	150
18	1001 a 2000	171
19	2001 a 5000	184
20	5001 a Mais	192

3 - Seleção dos elementos amostrais. Nesta etapa 3 foram selecionados os elementos a serem investigados em campo. Para a seleção dos elementos procedeu-se de forma aleatória, ou seja, selecionou-se os elementos de modo casual. Esse processo de seleção denomina-se Amostragem Aleatória. Nos municípios selecionados, sorteou-se aleatoriamente uma rua em um bairro típico do município (evitando os bairros muito novos ou muito velhos) e nessa rua realizou-se o procedimento de campo. Se necessário ou oportuno dividiu-se a amostra em duas ou mais ruas. Tendo sido selecionada a rua, quarteirão e bairro a amostrar, foram escolhidas como amostras ligações intercaladas, ou seja, ligações não vizinhas, até completar a quantidade necessária de elementos amostrais. Caso não fosse possível avaliar uma determinada ligação selecionada, adotou-se a ligação imediatamente próxima, seguindo o protocolo amostral.

4 - Vistoria de Campo. De posse dos elementos amostrais, selecionados conforme descrito nas etapas anteriores, procedeu-se à vistoria de campo. Essa vistoria de campo consistiu em observar no local de instalação da ligação e proximidades a existência, ou não, de irregularidades na ligação.

5 - Tabulação dos Resultados. Com os resultados da Vistoria de Campo preencheu-se a guia de controle de irregularidades da planilha de trabalho. Para cada amostra, preencheu-se a planilha com os dados obtidos na Vistoria de Campo de acordo com os tipos de irregularidades observadas.

6 - Estimativa de irregularidades na população. Consistiu na estimativa da quantidade presumida de elementos ou ligações irregulares na população, com base nos resultados da vistoria de campo e parcela de amostras irregulares. Os resultados serão apresentados na seção de Resultados Obtidos.

Identificação de Erro de Medição em Hidrômetros

A análise estatística de erros de medição de hidrômetros através da técnica de amostragem segue método semelhante à avaliação de fraudes. Contudo, algumas diferenças devem ser pontuadas, e se encontram descritas a seguir.

1 – Divisão dos municípios em cinco grupos, de acordo com população, porcentagem de hidrometração, idade de hidrômetros, número de economias e volume consumido. Os grupos estão identificados como Grupo A, B, C, D, E. Determinação de quantidade de municípios a serem avaliados por grupo. Para cada grupo, determinou-se a quantidade de municípios a serem avaliados, de acordo com a Tabela 1.

2 - Determinação da população a ser amostrada, que consiste no total de itens que poderiam compor a amostra. Para a estimação do erro em hidrômetros a população será a quantidade de Ligações Medidas com Hidrômetros.

3 - Determinação da proporção estimada para a ocorrência de hidrômetros com erro. Essa proporção é determinada pela experiência e conhecimento sobre o município em estudo, e baseia na escolha da tabela a ser utilizada. No caso, escolheu-se a mesma proporção de 15% de fraudes, utilizando-se assim a Tabela 2. O nível de confiança escolhido foi 95%, que representa a probabilidade de a amostra refletir a realidade, e precisão de 5%, que indica o erro amostral máximo admitido.

4 – Levantamento de campo. Para cada amostra, foram levantados em campo o consumo semanal de cada imóvel escolhido, e o respectivo hidrômetro retirado para levantamento dos erros por faixa de consumo em laboratório.

5 - Com os dados de campo e bancada em mãos, esses foram inseridos em plataforma BI e calculado o perfil de consumo individual de cada hidrômetro e seu IDM. Os resultados são descritos na seção Resultados Obtidos.



RESULTADOS OBTIDOS

Os resultados são o número de amostras definidas por grupo (por item) e os resultados provenientes do campo (número de fraudes, perfil de consumo e IDM médio).

Avaliação de Fraudes e Irregularidades em Ligações

Foi realizada uma pesquisa amostral de fraudes na Grande Vitória, onde foi definida uma amostra conforme abaixo, para cada grupo. Conforme a Tabela 2, todos os grupos foram definidos com 192 amostras, de acordo com a população de ligações, e dentro de cada grupo, algumas cidades foram escolhidas para serem alvo do estudo amostral, segundo a Tabela 1. Os resultados são apresentados nas figuras a seguir.

ITEM	GRUPO	MUNICÍPIO	MICROREGIÃO	POPULAÇÃO	CEP	LIGAÇÕES_INATIV	HIDRÔMETROS	VOLUME_FAT
1	A	AFONSO CLAUDIO	SERRANA	15,855	29600	573	64,144	789,232
11	A	BREJETUBA	SERRANA	3,416	29630	95	9,741	119,284
16	A	DIVINO DE SAO LOURENCO	SUL	1,742	29590	54	9,324	96,745
18	A	DORES DO RIO PRETO	SUL	3,547	29580	109	13,980	150,000
25	A	LARANJA DA TERRA	SERRANA	3,528	29615	155	13,392	152,030
40	A	SANTA LEOPOLDINA	SERRANA	2,615	29640	100	13,283	173,304
41	A	SANTA MARIA DE JETIBA	SERRANA	11,797	29645	611	47,724	586,400
42	A	SANTA TERESA	SERRANA	11,768	29650	416	54,146	654,257
45	A	SÃO ROQUE DO CANAÃ	SERRANA	11,273	29665	353	25,289	238,478
	9	TOTAL					251,023	

Amostra de cidades conforme os parâmetros:

Confiança 95%, Precisão: 5% e Tabela I.

Quantidades de cidades a amostrar: 3

Variável a ser considerada a População: HIDRÔMETROS

Total População: **251,023**

AMOSTRA SUGERIDA	POPULAÇÃO AMOSTRA	
BREJETUBA	9,741	50
DORES DO RIO PRETO	13,980	73
SANTA LEOPOLDINA	13,283	69
TOTAL	37,004	192

Figura 1: Determinação de amostra de fraudes no Grupo A.



ITEM	GRUPO	MUNICÍPIO	MICROREGIÃO	POPULAÇÃO	CEP	LIGAÇÕES_INATIV	HIDRÔMETROS	VOLUME_FAT
2	B	AGUA DOCE DO NORTE	NOROESTE	6,699	29820	550	29,489	324,063
3	B	AGUIA BRANCA	NOROESTE	3,051	29795	193	15,685	195,464
4	B	ALTO RIO NOVO	NOROESTE	4,256	29760	116	17,453	182,144
9	B	BOA ESPERANCA	CENTRO-NORTE	10,239	29845	364	36,205	405,539
19	B	ECOPORANGA	NOROESTE	14,779	29850	786	66,762	726,010
26	B	MANTENOPOLIS	NOROESTE	8,647	29770	465	37,252	371,904
28	B	MONTANHA	CENTRO-NORTE	13,522	29890	608	68,127	767,158
29	B	MUCURICI	CENTRO-NORTE	3,590	29880	514	18,345	170,010
33	B	PANCAS	NOROESTE	10,099	29750	453	36,633	438,990
37	B	PONTO BELO	CENTRO-NORTE	5,588	29885	376	26,478	253,599
49	B	VILA PAVAO	CENTRO-NORTE	2,998	29843	225	14,835	158,522
50	B	VILA VALÉRIO	NOROESTE	5,042	29785	383	17,871	222,117
	12	TOTAL					385,135	

Amostra de cidades conforme os parâmetros:

Confiança 95%, Precisão: 5% e Tabela I.

Quantidades de cidades a amostrar: 3

Variável a ser considerada a População: HIDRÔMETROS

Total População: 385,135

AMOSTRA SUGERIDA	POPULAÇÃO	AMOSTRA
ECOPORANGA	66,762	99
PANCAS	36,633	54
PONTO BELO	26,478	39
	129,873	192

Figura 2: Determinação de amostra de fraudes no Grupo B.

ITEM	GRUPO	MUNICÍPIO	MICROREGIÃO	POPULAÇÃO	CEP	LIGAÇÕES_INATIV	HIDRÔMETROS	VOLUME_FAT
5	C	ANCHIETA	LITORANEA	18,161	29230	1,103	110,039	1,439,698
8	C	BARRA DE SAO FRANCISCO	NOROESTE	26,357	29800	1,867	119,645	1,430,033
14	C	CONCEICAO DA BARRA	CENTRO-NORTE	22,575	29960	1,356	111,033	1,227,975
17	C	DOMINGOS MARTINS	SERRANA	7,741	29260	398	44,680	613,299
20	C	FUNDAO	GRANDE VITORIA	14,378	29185	1,255	63,674	810,729
27	C	MARECHAL FLORIANO	SERRANA	7,421	29255	347	35,915	445,492
32	C	NOVA VENEZIA	CENTRO-NORTE	30,831	29830	1,360	144,939	1,839,163
34	C	PEDRO CANARIO	CENTRO-NORTE	22,053	29970	1,619	80,114	793,442
35	C	PINHEIROS	CENTRO-NORTE	18,718	29980	896	79,870	887,787
36	C	PIUMA	LITORANEA	17,449	29285	2,128	124,145	1,632,458
39	C	RIO NOVO DO SUL	SUL	5,946	29290	280	26,496	361,134
43	C	SAO GABRIEL DA PALHA	NOROESTE	24,325	29780	1,070	83,618	1,180,571
	12	TOTAL					1,024,168	

Amostra de cidades conforme os parâmetros:

Confiança 95%, Precisão: 5% e Tabela I.

Quantidades de cidades a amostrar: 3

Variável a ser considerada a População: HIDRÔMETROS

Total População: 1,024,168

AMOSTRA SUGERIDA	POPULAÇÃO	AMOSTRA
DOMINGOS MARTINS	44,680	34
PEDRO CANARIO	80,114	62
PIUMA	124,145	96
	248,939	192

Figura 3: Determinação de amostra de fraudes no Grupo C.



ITEM	GRUPO	MUNICÍPIO	MICROREGIÃO	POPULAÇÃO	CEP	LIGAÇÕES_INATIV	HIDRÔMETROS	VOLUME_FAT
6	D	APIACA	SUL	5,211	29450	268	24,264	262,906
7	D	ATILIO VIVACQUA	SUL	6,116	29490	335	27,363	319,868
10	D	BOM JESUS DO NORTE	SUL	8,699	29460	361	43,720	508,954
13	D	CASTELO	SUL	21,817	29360	727	87,468	1,658,887
15	D	CONCEICAO DO CASTELO	SERRANA	5,898	29370	158	28,048	325,650
22	D	IBATIBA	SUL	13,378	29395	573	51,529	562,961
23	D	IRUPI	SUL	4,437	29398	174	16,637	217,773
24	D	IUNA	SUL	15,620	29390	437	64,381	829,988
30	D	MUNIZ FREIRE	SUL	8,670	29380	322	32,814	386,414
31	D	MUQUI	SUL	9,309	29480	426	41,966	482,623
38	D	PRESIDENTE KENNEDY	SUL	3,440	29350	243	14,199	193,305
44	D	SAO JOSE DO CALCADO	SUL	8,350	29470	256	37,380	412,022
47	D	VENDA NOVA DO IMIGRANT	SERRANA	14,809	29375	386	45,099	705,616
	13	TOTAL					514,868	

Amostra de cidades conforme os parâmetros:
Confiança 95%, Precisão: 5% e Tabela I.

Quantidades de cidades a amostrar: 3
Variável a ser considerada a População: HIDRÔMETROS
Total População: 514,868

AMOSTRA SUGERIDA	POPULAÇÃO	AMOSTRA
BOM JESUS DO NORTE	43,720	98
CONCEIÇÃO DO CASTELO	28,048	63
PRESIDENTE KENNEDY	14,199	32
	85,967	192

Figura 4: Determinação de amostra de fraudes no Grupo D.

ITEM	GRUPO	MUNICIPIO	MICROREGIÃO	POPULAÇÃO	CEP	LIGAÇÕES_INATIV	HIDROMETROS	VOL_N_FATURADO
12	E	CARIACICA	GRANDE VITORIA	337,643	29146	18,561	944,028	251,610
21	E	GUARAPARI	LITORANEA	100,528	29200	7,345	374,879	33,660
46	E	SERRA	GRANDE VITORIA	406,450	29176	22,213	719,032	555,999
48	E	VIANA	GRANDE VITORIA	59,632	29135	4,328	174,257	104,704
51	E	VILA VELHA	GRANDE VITORIA	412,575	29100	14,941	641,400	320,964
52	E	VITORIA	GRANDE VITORIA	327,801	29000	7,728	624,509	24,897
	6	TOTAL		1,644,629		75,116	3,478,105	1,291,834

Amostra de cidades conforme os parâmetros:
Confiança 95%, Precisão: 5% e Tabela I.

Quantidades de cidades a amostrar: 2
Variável a ser considerada a População: HIDRÔMETROS
Total População: 3,478,105

AMOSTRA SUGERIDA	POPULAÇÃO	AMOSTRA
VIANA	174,257	42
VITÓRIA	624,509	150
	798,766	192

Figura 5: Determinação de amostra de fraudes no Grupo E.

Para definição da porcentagem amostral de fraudes em ligações ativas e inativas, foram definidos níveis de suspeitas de fraudes, onde zero (0) não há suspeita, um (1) baixa suspeita, dois (2) moderada e três (3) alta suspeita de fraude. Quando zero, não foi considerado como fraude. Acima disso, foi considerado como fraude. Com o resultado em mãos, foi calculada uma porcentagem de irregularidades por ligação total por grupo, como mostra a Tabela 3.

**Tabela 3: Irregularidades por setor amostral.**

TIPO DE IRREGULARIDADE	Grupo A	Grupo B	Grupo C	Grupo D	Grupo E	TOTAL DE IRREGULARIDADES	PERCENTUAL IRREGULARIDADES
Arame trava o medidor	1				2	3	1,9%
By-Pass	1	1	14		3	19	12,0%
Clandestina	2	1	5		8	16	10,1%
Fornecimento visível a terceiros		2	5		6	13	8,2%
Ligação cortada, violada no ramal	13	3			7	23	14,6%
Ligação esgoto conectada sem contrato	1		3		5	9	5,7%
Medidor danificado	3	4		5	9	21	13,3%
Medidor invertido	3			5		8	5,1%
Medidor oxidado		15	4	6	2	27	17,1%
Medidor retirado	1	2		6		9	5,7%
Morsa				2		2	1,3%
Outros tipos	1			7		8	5,1%
Total de Irregularidades	26	28	31	31	42	158	100,0%
Total sem Irregularidades	166	164	161	161	150	802	-
Total Amostra n	192	192	192	192	192	960	-
Percentual Irregulares	13,5%	14,6%	16,1%	16,1%	21,9%	16,5%	-
Percentual Regulares	86,5%	85,4%	83,9%	83,9%	78,1%	83,5%	-

Subsequentemente, através de cálculo de proporções, foi calculado o número de fraudes por grupo, e o resultado é apresentado na Tabela 4.

Tabela 4: Estimativa de Ligações Irregulares nas amostras e na população geral.

DESCRIÇÃO	Grupo A	Grupo B	Grupo C	Grupo D	Grupo E	TOTAL DOS SETORES AMOSTRAIS
POPULAÇÃO TOTAL PARA A AMOSTRA N	37.004	129.873	248.939	85.967	798.766	1.300.549
QUANTIDADE ESPERADA DE EVENTOS IRREGULARES NA POPULAÇÃO N	4.921	18.633	39.630	13.686	173.228	250.098
ESTIMATIVA DE PERCENTUAL DE IRREGULARIDADES NO GRUPO	13%	14%	16%	16%	22%	19%

Posteriormente, através de cálculo em plataforma BI, foi determinado o número potencial de fraudes para as cidades da Região Metropolitana de Vitória, de especial interesse para o estudo.

Tabela 5: Fraudes amostrais detalhados por município.

MUNICÍPIO	ECONOMIAS ATIVAS	POTENCIAL DE FRAUDES (ECONOMIAS)
Anchieta	11.620	1.287



Cariacica	115.661	16.788
Fundão	7.094	1.050
Guarapari	67.338	8.131
Piúma	14.212	2.099
Serra	163.993	24.539
Viana	18.182	3.282
Vila Velha	188.846	20.029
Vitória	143.841	13.878

Identificação de Erro de Medição em Hidrômetros

Como descrito anteriormente, inicialmente determinou-se os níveis de confiança e precisão para a amostra. Para este procedimento define-se: Confiança 95% e precisão amostral 5%. Sob essas condições o tamanho da amostra é de 384 elementos ao todo. A distribuição da amostra foi aleatória sobre os municípios que compõem a base amostral considerando-se 2/3 da amostra na região da Grande Vitória e 1/3 da amostra no Interior do ES. Deste modo, a amostra na região de Grande Vitória é de 256 hidrômetros e na região do Interior, de 128 hidrômetros.

No grupo da região da Grande Vitória estão reunidos 9 municípios, e de acordo com a Tabela 1, adotando o nível de inspeção Moderado, foram selecionados aleatoriamente 2 municípios para compor a amostra, e nesses municípios os elementos amostrais foram distribuídos proporcionalmente à quantidade de ligações medidas com hidrômetro. No Grupo da região do Interior foram reunidos 43 municípios, e de acordo com a Tabela 1, adotando o nível de inspeção Moderado, foram selecionados aleatoriamente 3 municípios para compor a amostra, e nesses municípios os elementos amostrais foram distribuídos proporcionalmente à quantidade de ligações medidas com hidrômetro.

As Tabelas 6 e 7, exibidas a seguir, apresentam uma amostra sugerida, selecionada aleatoriamente dentre os municípios dos grupos da Grande Vitória e Interior. Nos municípios selecionados, sorteia-se aleatoriamente uma rua em um bairro típico do município (evitando os bairros muito novos ou muito velhos) e nessa rua realiza-se o procedimento de campo. Se necessário ou oportuno dividir a amostra em duas ou mais ruas.

Tabela 6: Amostragem sugerida de região da Grande Vitória para identificação de erros em hidrômetros.

ITEM	MACRO REGIÃO	REGIÃO	MUNICÍPIO	POPULAÇÃO	IDADE HIDROMETRO	LIGAÇÃO MEDIDA COM HIDRÔMETRO	Nº AMOSTRAS
4	GRANDE VITÓRIA	GRANDE VITÓRIA	VIANA	76.956	52	14.724	56
6	GRANDE VITÓRIA	GRANDE VITÓRIA	VITÓRIA	358.273	52	52.469	200
	SOMA			435.229		67.193	256

Tabela 7: Amostragem sugerida de região do Interior para identificação de erros em hidrômetros.

ITEM	MACRO REGIÃO	REGIÃO	MUNICÍPIO	POPULAÇÃO	IDADE HIDROMETRO	LIGAÇÃO MEDIDA COM HIDRÔMETRO	Nº AMOSTRAS
5	INTERIOR	CENTRO-NORTE	NOVA VENECIA	49.780	52	12.276	72
14		NOROESTE	ECOPORANGA	23.014	53	5.648	33



29		SERRANA	VENDA NOVA DO IMIGRANTE	24.802	54	3.825	23
	SOMA			97.596		21.749	128

Com os dados da amostragem em mãos, foi calculado o perfil de consumo individual de cada hidrômetro e seu IDM através de plataforma BI, visando atingir um IDM médio mais próximo da realidade, como mostra a Figura 6.

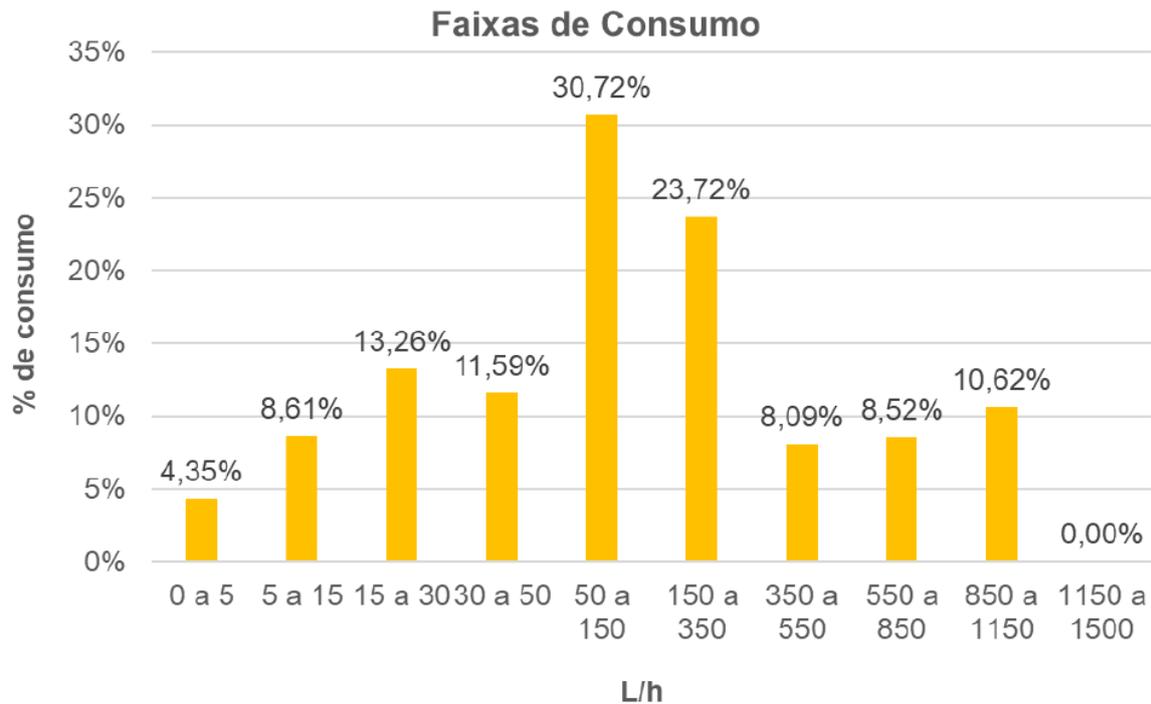


Figura 6: Média das Faixas de consumo amostral da GV.

Com isso, o IDM médio da GV foi encontrado (85,22%), utilizando-o para o cálculo da submedição no BI (14,78%). Foi realizada também uma análise para relacionar o IDM com a idade dos hidrômetros, apresentada no gráfico da Figura 7.

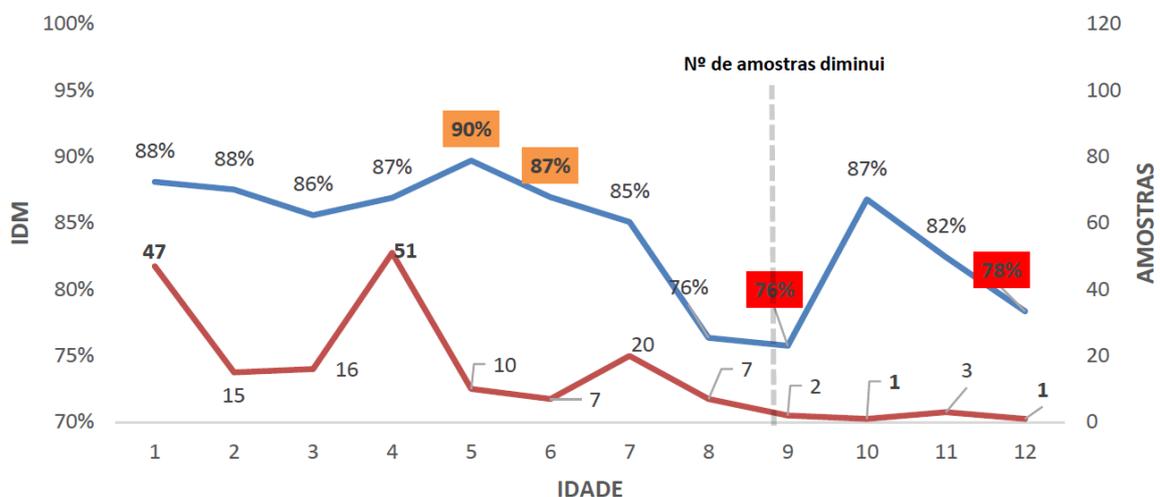


Figura 7: IDM por Idade e nº de amostras.

ANÁLISE E DISCUSSÃO DOS RESULTADOS

Com os resultados em mãos, pôde-se fazer algumas considerações quando ao número de ligações irregulares e erros na hidrometração.

Avaliação de Fraudes e Irregularidades em Ligações

Com elevado grau de confiabilidade e precisão, observou-se que entre os grupos, há relativa semelhança na proporção de ligações irregulares, com variações de 13% a 21% de ligações irregulares, o que corrobora para a confirmação dos resultados.

Identificação de Erro de Medição em Hidrômetros

No gráfico, percebe-se que a maior quantidade de HDs se encontra nos primeiros 4 anos de idade, o que significa que o parque de hidrômetros de Viana e Vitória são jovens e que os resultados são mais confiáveis nessa faixa de idade. A partir de 9 anos de idade, o número fica instável e com poucas amostras. Logo, o dado se mostra menos confiável.

Um aspecto importante que demonstra o gráfico é que os hidrômetros adquiridos nos últimos 4 anos são de menor qualidade se comparado aos hidrômetros adquiridos entre 5 e 8 anos atrás. Essa constatação deve-se ao fato de que o IDM dos hidrômetros jovens (0-4 anos) está menor se comparado à faixa de 5-8 anos. Além disso, o dado se mostra confiável, pois há várias amostras em ambas as faixas de idade.

CONCLUSÕES/RECOMENDAÇÕES

A análise estatística, por meio da técnica de amostragem apresentada, possibilitou a estimativa confiável de número de irregularidades, fraudes e erros de medição de hidrômetros sem tomar tanto tempo ou recursos disponíveis dentro da realidade do sistema de abastecimento. O aumento da eficiência na identificação dessas anomalias é de suma importância para a companhia de saneamento.

Nota-se também a necessidade de, agora dentro de cada município ou região, analisar dados e padrões de consumo únicos para cada localidade, o que poderia acarretar ações de combate a perdas aparentes ainda mais eficazes e assertivas. O trabalho deve ser contínuo, de modo a se ajustar aos cenários e situações identificados com o passar do tempo.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ABNT-Associação Brasileira de Normas Técnicas. NBR 5426 - Planos de amostragem e procedimentos na inspeção por atributos. Rio de Janeiro. 1985.
2. ARKIN. HERBERT. *Handbook of sampling for auditing and accounting*. 3a. Edition. McGraw-Hill. New York. 1984.
3. GUMIER CC. LUVIZOTTO JUNIOR EJ. *Aplicação de modelo de simulação-otimização na gestão de perda de água em sistemas de abastecimento*. Rev. Eng. Sanit. Amb. 12: 32-41. 2007.