

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL REI

Análise de homogeneidade, aleatoriedade, independência e estacionariedade das séries de dados de cota e vazão de estações fluviométricas da bacia do Rio Doce

**Frederico de Souza Rubatino
Pietro Cardoso do Nascimento
Rafael Gherardi de Alencar Roenick
Saymon Henrique Batista**

**Ouro Branco
2017**

**Frederico de Souza Rubatino
Pietro Cardoso do Nascimento
Rafael Gherardi de Alencar Roenick
Saymon Henrique Batista**

**Análise de homogeneidade, aleatoriedade, independência e
estacionariedade das séries de dados de cota e vazão de estações
fluviométricas da bacia do Rio Doce**

Relatório final apresentado a Universidade
Federal de São João del Rei, como parte
das exigências para obtenção do
Bacharelado em Ciência e Tecnologia

Orientador: Emmanuel Kennedy da
Costa Teixeira

Ouro Branco
2017

Sumário

1. Introdução	6
2. Objetivos.....	7
3. Revisão de literatura	7
3.1. A Bacia Hidrográfica do Rio Doce	7
3.2. Hidrologia estatística.....	8
3.3. Estudos Anteriores	15
4. Materiais e Métodos	16
4.1. Materiais.....	16
4.2. Métodos.....	17
5. Resultados e Discussões.....	23
6. Conclusão	25
7. Referências bibliográficas.....	25
Anexo A - Resultados dos testes com os dados de Cota	27
Anexo B – Resultados dos testes com os dados de vazão.....	42
Anexo C – Função para realizar os testes.....	57

Índice de Tabelas

Tabela 1 - Resultados da aplicação dos quatro testes.....	24
--	-----------

Índice de Ilustrações

Figura 1 - Sistema de busca de estações de uma mesma bacia.	17
Figura 2 - Dados assim que foram baixados do HidroWeb em uma planilha do Excel..	18
Figura 3 - Dados após verticalização e substituições.	19
Figura 4 - Importação dos dados para o R.....	19
Figura 5 - Primeiras linhas da função implementada no R.	20
Figura 6 - Resultado da função no R.	22

1. INTRODUÇÃO

Os recursos hídricos têm extrema importância para a sociedade devido à sua necessidade para o abastecimento de cidades, irrigação de plantações, além de ser uma das principais matrizes de geração de energia elétrica no país. Seu estudo, apesar de complexo, gera um impacto direto na capacidade de se prever e gerenciar crises devido a períodos de escassez, como também de excesso.

Os estudos sobre recursos hídricos são de extrema complexidade devido à variação de seus dados durante um período prolongado de tempo, por esse motivo as séries históricas analisadas contêm dados diários referentes a dezenas de anos com intuito de prever o comportamento esperado. A quantidade de dados e a alteração dos padrões a cada ano faz necessária a utilização de métodos estatísticos para realizar as análises, assim como testes para se confirmar e garantir a confiabilidade dos dados utilizados.

Muitos dados utilizados para realizar as análises são obtidos a partir do Sistema de Informações Hidrológicas (HidroWeb), ferramenta online disponibilizada pela Agência Nacional de águas (ANA), que contém dados de estações fluviométricas localizadas em rios de bacias hidrográficas de todo o país. Além dos dados de cota e vazão, da Bacia do Rio Doce, os quais foram utilizados nesse trabalho, o HidroWeb disponibiliza dados sobre diversas outras características dos cursos de água onde as estações estão localizadas.

O Rio Doce é formado pela junção dos rios Piranga e Carmo, cujas nascentes se localizam nas cidades de Ressaquinha e Ouro Preto, respectivamente. Ele percorre uma extensão de 853 km, e sua bacia hidrográfica tem uma extensão territorial de aproximadamente 83.400 km², abrangendo 228 municípios, sendo 202 deles no estado de Minas Gerais e 26 no estado do Espírito Santo, e uma população em cerca de 3,1 milhões habitantes (Companhia Energética de Minas Gerais - CEMIG).

Os testes de hipótese utilizados para verificar a confiabilidade e a tendência dos dados, são de extrema importância para a aplicação da hidrologia estatística, sendo os testes de estacionariedade e homogeneidade pré-requisitos para essa aplicação (NAGHETTINI e PINTO, 2007), além de testes de aleatoriedade e independência que também foram aplicados nesse estudo. A caracterização de estacionariedade é fundamental para a análise e previsão do comportamento das séries históricas, significando que os valores se mantem em equilíbrio em torno de um valor médio e tem uma variância constante com o passar do tempo (MURPHY e ELLIS, 2013). Os testes de homogeneidade aferem se os dados provem de uma mesma população, confirmando a similitude de suas características, sendo esta característica muito

importante para a conferência dos resultados obtidos ao se utilizar esses dados. Outras análises necessárias a serem feitas, são a de que as variações observadas nas amostras são de fato resultado de alterações climáticas naturais, como também a aferição sobre a influência dos dados de determinada amostra sobre observações posteriores, estas análises são feitas pelos testes de aleatoriedade e independência, respectivamente.

Este estudo tem seu foco na análise do comportamento dos dados de cota e vazão relativos às estações fluviométricas presentes na Bacia do Rio Doce. São testados em função de testes de hipótese que confirmam a confiabilidade destes dados para sua posterior utilização em geração estocástica ou outras análises e estudos que necessitem desses resultados.

2. OBJETIVOS

O presente trabalho tem como objetivo central a análise de dados hidrológicos das estações fluviométricas que compõem a Bacia do Rio Doce, verificando aleatoriedade, estacionariedade, homogeneidade e independência dos mesmos.

3. REVISÃO DE LITERATURA

3.1. A Bacia Hidrográfica do Rio Doce

De acordo com a Agência Nacional de Águas (ANA), a bacia hidrográfica do Rio Doce possui 83.500 km², sendo que 86% dessa área se encontra no estado de Minas Gerais e os outros 14% no estado do Espírito Santo. Os recursos hídricos da bacia são de grande importância para a economia, uma vez que diversas atividades como a mineração, a agropecuária, a industrial e a de geração de energia são desenvolvidas ao longo de sua extensão.

O entendimento do comportamento hidrológico de uma bacia hidrográfica é de fundamental importância para o gerenciamento dos recursos hídricos, visto que estes são limitados. Dessa forma, requer-se o conhecimento dos diferentes segmentos do ciclo hidrológico para o planejamento, projeto e operação de qualquer obra relacionada com o aproveitamento e controle dos recursos hídricos (BARBOSA, 2004).

A análise dos dados de uma série hidrológica é um procedimento utilizado para a determinação de algumas características de uma bacia hidrográfica, tais como, a

homogeneidade, a aleatoriedade, a estacionariedade e a independência. Normalmente, os testes estatísticos são utilizados para a análise de tendências temporais nas séries históricas com o objetivo de determinar o melhor modelo probabilístico representativo dos dados. Essa determinação é de grande importância para a hidrologia estatística, uma vez que, na maioria de suas aplicações, as séries hidrológicas reduzidas devem ter como pré-requisito as características de estacionariedade e homogeneidade (NAGHETTINI e PINTO, 2007).

3.2. Hidrologia estatística

Os fenômenos ou processos hidrológicos dificilmente são determinados de forma precisa por equações matemáticas, uma vez que eles estão sujeitos a diversos fatores aleatórios, como, por exemplo, a precipitação, as estações do ano, topografia, mudanças climáticas e diversos outros fatores.

Dessa forma, são raros os casos em que os dados hidrológicos podem ser representados através de métodos puramente determinísticos, ou seja, a caracterização dos fenômenos por meio de equações matemáticas muitas vezes possui erros devido à alta aleatoriedade das variáveis. Sendo assim, pode-se considerar que os processos hidrológicos são estocásticos, uma vez que são regidos pelas leis da probabilidade, ficando a hidrologia estatística responsável pelo estudo da estatística matemática e da teoria de probabilidades aplicadas em dados hidrológicos e a hidrologia estocástica responsável pelo estudo dos processos hidrológicos estocásticos (NAGHETTINI e PINTO, 2007). Simplificadamente, por exemplo, não é possível prever com exatidão a vazão de um rio, mas pode-se determinar a probabilidade de um determinado valor de vazão ocorrer.

As variáveis hidrológicas e hidrometeorológicas têm sua variabilidade registrada por meio das séries temporais, as quais reúnem as observações ou medições daquela variável, organizadas no modo sequencial de sua ocorrência no tempo (ou espaço) (NAGHETTINI e PINTO, 2007). Esses dados normalmente são mensurados em intervalos regulares de tempo ou distância, de forma que as séries temporais podem ser formadas por distintos tipos de dados médios (médias mensais, anuais etc.). Isso ocorre devido ao fato das áreas de estudo possuírem grandes diferenças nas suas características, como por exemplo a área de drenagem, o que influencia diretamente na série temporal que representará de forma mais confiável a variabilidade dos dados.

Para a determinação da distribuição de probabilidades das séries temporais das estações fluviométricas da Bacia do Rio Doce foram analisadas 210 estações fluviométricas utilizando

um procedimento estatístico conhecido como teste de hipótese. Esses testes são úteis nas tomadas de decisões, baseadas em um nível de significância, sobre alguma afirmação de uma população ou parâmetro desta, desejando saber se os resultados de uma amostra contrariam ou não tal afirmação (BUSSAB e MORETTIN, 2010).

Os testes de hipóteses utilizados estão descritos a seguir e são classificados como não paramétricos, uma vez que não dependem que a distribuição da população seja caracterizada por certos parâmetros. O *software* de programação estatística R foi utilizado para a aplicação de todos os testes e as funções empregadas se encontram no Anexo C.

3.2.1. Teste da hipótese de aleatoriedade

O teste de aleatoriedade verifica se as variações na série dos dados hidrológicos são consequências de causas naturais. Para a determinação dessa propriedade nos dados de cotas e vazões das estações, foi aplicado o teste do número de inflexões, sugerido pela NERC (1975), que está detalhado a seguir.

Caso a série histórica seja aleatória, pode-se determinar o número de inflexões (“picos” ou “vales” de um gráfico da variável em questão em função do tempo) através da Equação 1:

$$E(p) = \frac{2(N-2)}{3} \quad (\text{Equação 1})$$

Onde:

$E(p)$ é o número esperado de inflexões;

N é o tamanho da amostra (número de observações).

Com variância dada pela Equação 2:

$$Var(p) = \frac{16N-29}{90} \quad (\text{Equação 2})$$

Onde:

$Var(p)$ é a variância do número de inflexões.

Para que o resultado do teste seja confiável, a amostra deve possuir tamanho $N > 30$, uma vez que, com essa condição satisfeita, a variável p segue aproximadamente uma distribuição normal. Isso acontece porque quando o tamanho de uma amostra aumenta, a distribuição amostral dos valores médios aproxima-se cada vez mais de uma distribuição normal, resultado esse conhecido como Teorema do Limite Central (BUSSAB e MORETTIN, 2010).

Considerando que a hipótese nula H_0 é a da amostra ser aleatória, a estatística T do teste é representada pela Equação 3:

$$T = \frac{p - E(p)}{\sqrt{\text{Var}(p)}} \quad (\text{Equação 3})$$

Onde:

T é a estatística do teste não paramétrico;

p é o número de inflexões observados nos dados.

Com um nível de significância de 5%, a hipótese nula é rejeitada caso $|T| > Z_{97,5}$ (teste bilateral).

3.2.2. Teste da hipótese de independência

O teste de independência verifica se os dados da amostra influenciam em qualquer outra observação posterior. Para a determinação dessa propriedade nos dados de cotas e vazões das estações foi aplicado o teste indicado por Wald e Wolfowitz (1943), que está detalhado a seguir.

Considerando os dados de uma amostra qualquer de tamanho N $\{X_1, X_2, X_3, \dots, X_n\}$, sendo \bar{X} a sua média amostral e X' dado pela Equação 4:

$$X'_i = X_i - \bar{X} \quad (\text{Equação 4})$$

Onde:

X'_i é a diferença entre um dado da amostra na posição i e a média amostral dos dados;

X_i é um dado da amostra na posição i ;

\bar{X} é a média amostral dos dados.

O coeficiente R é calculado pela Equação 5:

$$R = \sum_{i=1}^{N-1} X'_i X'_{i+1} + X'_1 X'_N \quad (\text{Equação 5})$$

R é um coeficiente do teste de Wald e Wolfowitz.

N é o tamanho da amostra (número de observações).

Caso a série histórica seja independente, o coeficiente R segue uma distribuição normal com média calculada pela Equação 6:

$$E[R] = -\frac{s_2}{N-1} \quad (\text{Equação 6})$$

Onde:

$E[R]$ é a média do coeficiente R .

E variância dada pela Equação 7:

$$Var[R] = \frac{s_2^2 - s_4}{N-1} + \frac{s_2^2 - 2s_4}{(N-1)(N-2)} - \frac{s_2^2}{(N-1)^2} \quad (\text{Equação 7})$$

Onde:

$Var[R]$ é a variância do coeficiente R .

As variáveis s_r e m'_r são dadas pelas Equações 8 e 9 respectivamente:

$$s_r = Nm'_r \quad (\text{Equação 8})$$

$$m'_r = \frac{\sum_{i=1}^N (X'_i)^r}{N} \quad (\text{Equação 9})$$

Considerando que a hipótese nula H_0 é a da amostra ser independente, a estatística T do teste é representada pela Equação 10:

$$T = \frac{R - E[R]}{\sqrt{VAR[R]}} \quad (\text{Equação 10})$$

Onde:

T é a estatística do teste não paramétrico.

Com um nível de significância de 5%, a hipótese nula é rejeitada caso $|T| > Z_{97,5}$ (teste bilateral).

3.2.3. Teste de homogeneidade

O teste de homogeneidade verifica se os dados de uma amostra possuem características semelhantes, ou seja, se são provenientes de uma mesma população. Para a determinação dessa propriedade nos dados de cotas e vazões das estações, foi aplicado o teste indicado por Mann e Whitney (1947), que está detalhado a seguir.

Considerando os dados de uma amostra qualquer de tamanho N $\{X_1, X_2, X_3, \dots, X_n\}$, a mesma foi separada em duas sub amostras, $\{X_1, X_2, X_3, \dots, X_{N_1}\}$, de tamanho N_1 , e $\{X_{N_1+1}, X_{N_1+2}, X_{N_1+3}, \dots, X_N\}$, de tamanho N_2 , de forma que $N_1 + N_2 = N$. Os tamanhos das sub amostras devem ser aproximados, de modo que $N_1 \leq N_2$. A amostra completa (tamanho N) foi ordenada de forma crescente e os seus elementos receberam uma classificação de ordem m .

O coeficiente V do teste é dada pelo menor valor entre as Equações 11 e 12, representadas a seguir:

$$V_1 = N_1 N_2 + \frac{N_1(N_1+1)}{2} - R_1 \quad (\text{Equação 11})$$

$$V_2 = N_1 N_2 - V_1 \quad (\text{Equação 12})$$

Onde:

V_1 e V_2 são coeficientes do teste de Mann e Whitney;

N_1 é o tamanho da sub amostra 1 (número de observações);

N_2 é o tamanho da sub amostra 2 (número de observações);

R_1 é o somatório das classificações dos dados da primeira sub amostra (tamanho N_1).

Caso a série histórica seja homogênea e $N_1, N_2 > 20$, o coeficiente V segue uma distribuição normal com média calculada pela Equação 13:

$$E[V] = \frac{N_1 N_2}{2} \quad (\text{Equação 13})$$

Onde:

$E[V]$ é a média do coeficiente V .

E variância é dada pela Equação 14:

$$VAR[V] = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12} \quad (\text{Equação 14})$$

Onde:

$Var[V]$ é a variância do coeficiente V .

Considerando que a hipótese nula H_0 é a da amostra ser homogênea, a estatística T do teste é representada pela Equação 15:

$$T = \frac{V - E[V]}{\sqrt{VAR[V]}} \quad (\text{Equação 15})$$

Onde:

T é a estatística do teste não paramétrico;

Com um nível de significância de 5%, a hipótese nula é rejeitada caso $|T| > Z_{97,5}$ (teste bilateral).

3.2.4. Teste de Estacionariedade

O teste de estacionariedade verifica se os dados, desconsiderando as variações aleatórias, não se alteram ao longo do tempo. Para a determinação dessa propriedade nos dados de cotas e vazões das estações, foi aplicado o teste de Spearman, sugerido pela NERC (1975), que está detalhado a seguir.

Considerando os dados de uma amostra qualquer de tamanho N $\{X_1, X_2, X_3, \dots, X_n\}$, foi analisado o coeficiente de correlação entre a ordem de ocorrência T_t e a ordem de classificação crescente m_t , dos elementos.

O coeficiente r_s do teste de Spearman é calculada pela Equação 16:

$$r_s = 1 - \frac{6 \sum_{t=1}^N (m_t - T_t)^2}{N^3 - N} \quad (\text{Equação 16})$$

Onde:

r_s é um coeficiente do teste de Spearman;

T_t é a ordem de ocorrência dos dados;

m_t é a classificação crescente dos dados.

Caso a série história tenha tamanho $N > 10$, o coeficiente r_s segue uma distribuição normal com média dada pela Equação 17:

$$E[r_s] = 0 \quad (\text{Equação 17})$$

Onde:

$E[r_s]$ é a média do coeficiente r_s .

E variância pela Equação 18:

$$VAR[r_s] = \frac{1}{N-1} \quad (\text{Equação 18})$$

Onde:

$VAR[r_s]$ é a variância do coeficiente r_s .

Considerando que a hipótese nula H_0 é a de que T_i e m_i não tenham correlação, a estatística T do teste é representada pela Equação 19:

$$T = \frac{r_s}{\sqrt{VAR[r_s]}} \quad (\text{Equação 19})$$

Com um nível de significância de 5%, a hipótese nula é rejeitada caso $|T| > Z_{97,5}$ (teste bilateral).

Onde:

T é a estatística do teste não paramétrico;

3.3. Estudos Anteriores

Devida a importância da determinação das características de uma bacia hidrográfica, diversos autores já realizaram estudos investigativos que utilizam dados de séries hidrológicas e testes estatísticos para a verificação das características estatísticas que melhor representam esses dados.

Detzel *et al.* (2011) estudaram a estacionariedade das séries hidrológicas de vazões naturais (desconsidera o uso e ocupação antrópica) de todas as usinas hidrelétricas que compunham o Sistema Interligado Nacional (até o mês de dezembro 2007). A análise dos dados foi realizada utilizando os testes estatísticos paramétricos t-Student, e F-Snedecor e os não paramétricos Cox-Stuart, Wilcoxon, Coeficiente de Correlação de Spearman e Mann-Kendall. Foi concluído que, considerando o total de séries analisadas, a maioria apresentou evidências de não estacionariedade.

Batista *et al.* (2009) verificaram a estacionariedade das séries hidrológicas de chuva e vazão das regiões Sul e Sudeste brasileiras. Para isso, utilizaram os seguintes métodos: (i) Análise de tendência linear; (ii) testes estatísticos t-Student e Wilcoxon; (iii) relação entre escoamento e precipitação; e (iv) curva acumulativa das médias anuais em função do tempo. Foi concluído

que apenas as séries hidrológicas da região Sul poderiam ser caracterizadas como não estacionárias com significância relevante.

Com o objetivo de determinar as distribuições de probabilidades mais adequadas para os dados fluviométricos brasileiros, Costa (2014) verificou a homogeneidade, a estacionariedade e a presença de saltos utilizando os testes de Mann-Whitney, Spearman e Petit, nos dados de vazões médias diárias máximas anuais de 1943 estações fluviométricas do Brasil. Das amostras, 72,5% podem ser consideradas homogêneas, 70,02% estacionárias, 82,86% não apresentam saltos nas médias e 64,49% não tiveram a hipótese nula rejeitada nos três testes. Foi concluído que as distribuições log normal de dois e três parâmetros são as distribuições de probabilidades mais indicadas para os dados brasileiros.

Tröger *et al.* (2004) fizeram um estudo para a aceitação da hipótese de estacionariedade das séries de vazões naturais das usinas hidrelétricas de Três Marias e Sobradinho, no período de 1931-2001. A análise foi feita a partir dos resultados dos testes paramétricos t-Student, e F-Snedecor e os não paramétricos Mann-Kendall, coeficiente de Spearman, Wilcoxon ou Mann-Whitney, Petit e Wald- Wolfowitz. Foi concluído que as séries de dados de vazões médias anuais naturais podem ser caracterizadas como estacionárias.

Torres *et al.* (2016) realizaram um estudo para determinar o tamanho da amostra (em anos) para a estimação da média de precipitação pluvial mensal em locais no Estado de Mato Grosso do Sul. Foi verificada a aleatoriedade dos dados através do *Run Test* e a homogeneidade de variâncias através do teste de Bartlett. Concluiu-se que a estimação do tamanho da amostra para a determinação da média pluvial mensal, nos locais e meses analisados, era de 63 anos.

É possível encontrar muitos outros estudos que envolvem a utilização de testes não paramétricos, uma vez que além da importância de se conhecer as propriedades estatísticas das séries temporais, as mesmas podem se alterar com o passar dos anos.

4. MATERIAIS E MÉTODOS

4.1. Materiais

A principal ferramenta utilizada na análise dos dados foi o *software* de programação estatística R. O Microsoft Excel também foi utilizado principalmente na fase preliminar da análise dos dados. O material analisado foi retirado do Sistema de Informações Hidrológicas (HidroWeb) da Agência Nacional de Águas (ANA).

4.2. Métodos

Inicialmente selecionou-se todas as estações fluviométricas da Bacia do Rio Doce, através do HidroWeb, que continham séries históricas de dados de vazão e/ou cota, num total de 210 estações. Destas, 142 estações possuem dados de vazão e 206 possuem dados de cotas. Os dados são exportados do HidroWeb para Excel em planilhas que contém os dias nas colunas e os meses de todos os anos nas linhas.

The screenshot shows the HidroWeb interface with the following elements:

- Header:** "HidroWeb Sistema de Informações Hidrológicas" and "Fechar Janela Janela Principal".
- Section:** "Sub-bacias".
- Status:** "Consultando o banco de dados..." and "Nova Consulta".
- Message:** "1 registro selecionado."
- Result Table:**

Código	Nome	Jurisdição
56	RIO DOCE	-
- Form:** "Formulário de Consulta" with fields for "Código (2 dígitos): 56 até 56 (Ex.: 10 até 90)" and "Nome: (Ex.: Rio Canoas)".
- Note:** "* A consulta pode ser feita por qualquer campo."
- Buttons:** "Arquivo Access", "Arquivo Texto", and "Listar".
- Instructions:** "Clique em:" followed by a list of actions:
 - Arquivo Access** - para criar arquivo Access compactado com os dados da consulta (esse arquivo pode depois ser importado pelo Hidro).
 - Arquivo Texto** - para criar arquivo texto compactado com os dados da consulta.
 - Listar** - para listar a identificação dos registros consultados na tela.
- Footer:** "© 2001 Agência Nacional de Águas (ANA)" and "Topo".

Figura 1 - Sistema de busca de estações de uma mesma bacia. (Fonte: HidroWeb.)

The screenshot shows an Excel spreadsheet with the following structure:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34

Figura 2 - Dados assim que foram baixados do HidroWeb, em uma planilha do Excel. (Fonte: os autores.)

Para importar tais dados para o R, foi necessário fazer com que os dados fossem verticalizados, ou seja, uma única coluna contendo cada dado de cada dia da série histórica em uma linha diferente, o que no R pode ser chamado de um vetor. A verticalização foi feita através de uma planilha com macro no Excel.

As falhas de dados presentes nas séries históricas são dadas como 0 na planilha exportada do HidroWeb, e conseqüentemente no vetor após a verticalização dos dados. Tais falhas seriam interpretadas como dados pelo R, por isso, ainda no Excel foi feita a substituição de todos os zeros por NA (*not available*), que são interpretados pelo R como falta de dados. Outro ponto importante foi a substituição do separador decimal. Os dados do HidroWeb utilizam a vírgula, no entanto, o R só reconhece o ponto como separador decimal. Após as substituições de 0 por NA e de vírgulas por pontos, a importação dos dados para o R foi feita através da função `scan()`.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Data	Vazão											
2	01/01/1938	NA											
3	02/01/1938	NA											
4	03/01/1938	NA											
5	04/01/1938	NA											
6	05/01/1938	NA											
7	06/01/1938	NA											
8	07/01/1938	NA											
9	08/01/1938	NA											
10	09/01/1938	NA											
11	10/01/1938	NA											
12	11/01/1938	NA											
13	12/01/1938	NA											
14	13/01/1938	NA											
15	14/01/1938	NA											
16	15/01/1938	NA											
17	16/01/1938	NA											
18	17/01/1938	NA											
19	18/01/1938	NA											
20	19/01/1938	NA											
21	20/01/1938	NA											
22	21/01/1938	NA											
23	22/01/1938	NA											
24	23/01/1938	NA											
25	24/01/1938	NA											
26	25/01/1938	NA											
27	26/01/1938	NA											
28	27/01/1938	NA											
29	28/01/1938	NA											
30	29/01/1938	NA											
31	30/01/1938	NA											
32	31/01/1938	NA											

Figura 3 - Dados após verticalização e substituições. (Fonte: os autores.)

```

R version 3.3.3 (2017-03-06) -- "Another Canoe"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes de R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

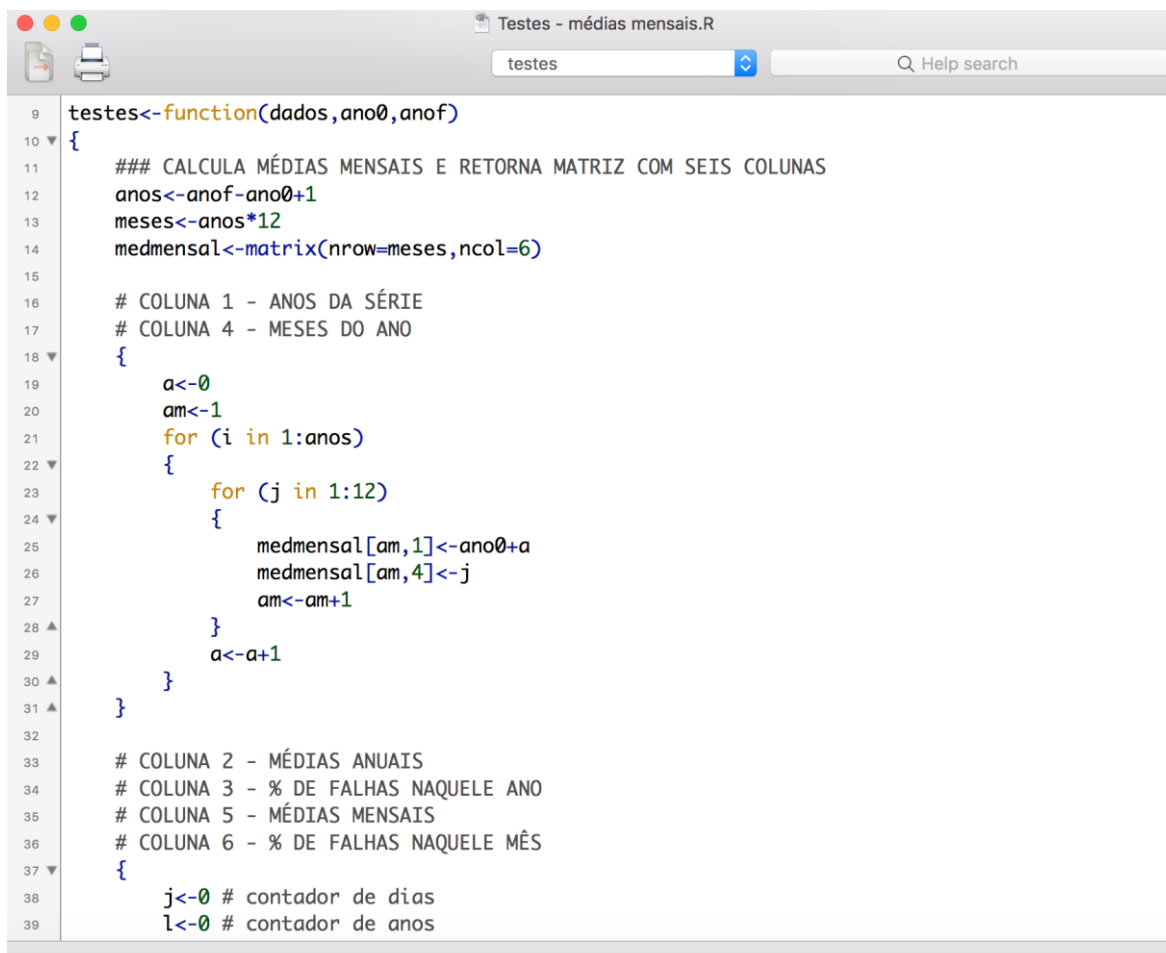
[R.app GUI 1.69 (7328) x86_64-apple-darwin13.4.0]
[Workspace restored from /Users/rafaelgherardi/.RData]
[History restored from /Users/rafaelgherardi/.Rapp.history]

> vazoes<-scan()
1: NA
2: NA
3: NA
4: NA
5: NA
6: NA
7: NA
8: NA
9: NA
10: NA
11: NA
12: NA
13: NA
14: NA
...
3287: NA
Read 3287 items
>

```

Figura 4 - Importação dos dados para o R. (Fonte: os autores.)

Importados os dados para o R, o primeiro passo foi implementar a função para fazer os testes não paramétricos. Tal função deve receber como parâmetros o vetor com os dados, o ano inicial e o ano final da série. Os dados são diários, e os testes foram aplicados para médias mensais, portanto a primeira parte da função faz o cálculo das médias mensais e retorna uma matriz com seis colunas denominada “medmensal” com as médias mensais e outros dados relevantes que serão descritos a seguir.



```

9 testes<-function(dados,ano0,anof)
10 {
11     ### CALCULA MÉDIAS MENSAIS E RETORNA MATRIZ COM SEIS COLUNAS
12     anos<-anof-ano0+1
13     meses<-anos*12
14     medmensal<-matrix(nrow=meses,ncol=6)
15
16     # COLUNA 1 - ANOS DA SÉRIE
17     # COLUNA 4 - MESES DO ANO
18     {
19         a<-0
20         am<-1
21         for (i in 1:anos)
22         {
23             for (j in 1:12)
24             {
25                 medmensal[am,1]<-ano0+a
26                 medmensal[am,4]<-j
27                 am<-am+1
28             }
29             a<-a+1
30         }
31     }
32
33     # COLUNA 2 - MÉDIAS ANUAIS
34     # COLUNA 3 - % DE FALHAS NAQUELE ANO
35     # COLUNA 5 - MÉDIAS MENSAIS
36     # COLUNA 6 - % DE FALHAS NAQUELE MÊS
37     {
38         j<-0 # contador de dias
39         l<-0 # contador de anos

```

Figura 5 - Primeiras linhas da função implementada no R. (Fonte: os autores.)

A primeira coluna da matriz que a função retorna apresenta o ano correspondente aos dados da série histórica. Para isto a função utiliza o ano inicial da série, que é dado como parâmetro, repete o mesmo ano em 12 linhas e então, soma uma unidade ao ano, repetindo o mesmo mais 12 vezes, e assim sucessivamente, até chegar ao ano final da série. Já a quarta coluna apresenta os meses de cada ano, sendo 1 para janeiro, 2 para fevereiro, etc., e para cada vez que a função repete o ano na coluna 1, a mesma adiciona o número correspondente ao mês na coluna quatro.

A função faz a contagem de dia por dia, mês por mês e ano por ano, por isso é necessário que o primeiro dado importado para o R seja do dia 01 de janeiro do ano inicial da série.

Também foram levados em consideração os anos bissextos (nos quais o mês de fevereiro tem 29 dias), e a condição adicionada para isto é de que o ano é bissexto se for divisível por 4 e divisível por 400 ou se for divisível por 4 e não divisível por 100. Se o ano for bissexto, o contador de dias recebe 366 unidades, se não for bissexto, recebe 365 unidades.

A coluna 2 apresenta a média anual dos dados, ou seja, o mesmo valor de média correspondente ao ano da coluna 1 é apresentado 12 vezes. Para cada ano, seja ele bissexto ou não, os dados diários são armazenados em um vetor temporário e a partir deste vetor calcula-se a média anual. A coluna 3 apresenta a porcentagem de falhas naquele ano, valor também apresentado 12 vezes devido à correspondência com o ano da coluna 1. O percentual de falhas em um ano é calculado pela contagem de falhas, apresentadas como “NA”, nos 365 ou 366 dias daquele ano.

Analogamente ao cálculo de médias anuais, o cálculo de médias mensais, apresentado na coluna 5, é feito a partir da contagem dos dias do mês. Para janeiro, conta-se 31 dados, para fevereiro conta-se 28 dados ou 29 dados em casos de anos bissextos, para março conta-se novamente 31 dados, para abril conta-se 30 dados, e assim sucessivamente. Para cada mês os dados diários são armazenados em um vetor temporário e a partir do mesmo calcula-se a média mensal. O percentual de falhas em um mês é calculado da mesma forma que o cálculo é feito para as falhas em um ano, e o resultado é apresentado na coluna 6.

Os testes serão aplicados para as médias mensais pois algumas estações fluviométricas apresentam poucos dados, muitas vezes menos de 10 anos, o que acarretaria em menos de 10 dados para a aplicação dos testes, número de amostras insuficientes para resultados de testes confiáveis, como já apresentado anteriormente. Portanto, a matriz de seis colunas que a função retorna apresenta as médias anuais além das médias mensais, apenas para simples conferência.

O critério utilizado para invalidar um ano inteiro de dados é o de que caso haja algum mês com mais de 10% de falhas, todo aquele ano seria excluído antes da aplicação dos testes. Não se excluiu somente o mês com mais de 10% de falhas, pois assim haveria um desfalque daquele mês em relação ao total de dados, o que poderia impactar no resultado dos testes já que cada época do ano tem uma diferente tendência (épocas de cheias e secas). Nos anos em que haviam falta de dados, considerou-se os meses que não haviam dados como falhas, consequentemente aquele ano inteiro foi excluído da série de dados.

Para o resto da função é criada uma matriz “medmensal1” idêntica à matriz “medmensal”. Primeiramente exclui-se a segunda e terceira colunas da matriz “medmensal1” de seis colunas, passando a ser uma matriz de quatro colunas. Para a exclusão do ano que possui pelo

menos um mês com mais de 10% de falhas, e faz-se uma análise dos dados da quarta coluna de ano a ano. Em cada ano são analisados todos os meses, caso haja uma linha da coluna 4 com valor igual ou superior a 10, uma variável temporária recebe “TRUE”. Se esta variável for igual a “TRUE”, um vetor chamado “linhas” recebe todas as linhas da matriz correspondentes àquele ano. Isso se repete até que todos os anos tenham sido percorridos, e, então, as linhas do vetor “linhas” são excluídas da matriz “medmensal1”. A matriz “medmensal1” é impressa na tela para a conferência dos dados, e a coluna 2 é removida antes de começar a aplicação dos testes para facilitar os cálculos.

A implementação dos testes (aleatoriedade, independência, homogeneidade e estacionariedade, nesta ordem) na função do R foi feita de acordo com a metodologia e as equações apresentadas na Revisão de Literatura. No primeiro teste, criou-se um novo vetor a partir da coluna da matriz (medmensal1) que continha os dados e a partir deste vetor realizou-se os cálculos necessários para encontrar o valor de T, que comparado ao valor z daria o resultado do teste. O resultado do teste é apresentado com o texto impresso na tela: “Rejeita-se H_0 a um nível de significância de 5%”, ou “Não rejeita-se H_0 a um nível de significância de 5%”. Adicionalmente, caso o número de dados não for o mínimo exigido para a aplicação do teste, a função imprimirá o aviso “ATENÇÃO! Os dados disponíveis não são o suficiente para um teste confiável” junto ao resultado do teste.

```

R Console
~/Dropbox/FACULDADE/2017:2/TCIC/Parte prática/
> testes(vazoes,1938,1946)
[1,] [,1] [,2] [,3] [,4]
[1,] 1939 1 27.3577419 0
[2,] 1939 2 6.8889286 0
[3,] 1939 3 2.4209677 0
[4,] 1939 4 2.0953333 0
[5,] 1939 5 1.0756129 0
[6,] 1939 6 0.7940000 0
[7,] 1939 7 0.7542581 0
[8,] 1939 8 0.6671613 0
[9,] 1939 9 0.7167333 0
[10,] 1939 10 1.6112903 0
[11,] 1939 11 2.1175333 0
[12,] 1939 12 5.3716129 0
[13,] 1941 1 8.2129032 0
[14,] 1941 2 4.0907143 0
[15,] 1941 3 5.3722581 0
[16,] 1941 4 5.3943333 0
[17,] 1941 5 1.4880645 0
[18,] 1941 6 1.5700000 0
[19,] 1941 7 1.5377419 0
[20,] 1941 8 1.0667742 0
[21,] 1941 9 2.9113333 0
[22,] 1941 10 2.6548387 0
[23,] 1941 11 3.5486667 0
[24,] 1941 12 16.0887097 0
[25,] 1942 1 19.2764516 0
[26,] 1942 2 8.1217857 0
[27,] 1942 3 8.4635484 0
[28,] 1942 4 3.1410000 0
[29,] 1942 5 1.4116129 0
[30,] 1942 6 1.2496667 0
[31,] 1942 7 1.1705806 0
[32,] 1942 8 1.0519355 0
[33,] 1942 9 1.2030667 0
[34,] 1942 10 3.8564516 0
[35,] 1942 11 9.0723333 0
[36,] 1942 12 3.0555556 0
[43,] 1943 7 1.1170323 0
[44,] 1943 8 0.9823226 0
[45,] 1943 9 0.8340000 0
[46,] 1943 10 2.2601935 0
[47,] 1943 11 5.0426667 0
[48,] 1943 12 18.3935484 0
[49,] 1944 1 4.8680645 0
[50,] 1944 2 11.5034483 0
[51,] 1944 3 6.0925806 0
[52,] 1944 4 3.7410000 0
[53,] 1944 5 1.5629032 0
[54,] 1944 6 1.1042667 0
[55,] 1944 7 0.8867742 0
[56,] 1944 8 0.7470323 0
[57,] 1944 9 0.6672000 0
[58,] 1944 10 0.9201935 0
[59,] 1944 11 3.6004667 0
[60,] 1944 12 9.9629032 0
[61,] 1945 1 29.3906452 0
[62,] 1945 2 9.4042857 0
[63,] 1945 3 10.4509677 0
[64,] 1945 4 7.6973333 0
[65,] 1945 5 1.9729032 0
[66,] 1945 6 1.2713333 0
[67,] 1945 7 0.9405161 0
[68,] 1945 8 0.7310323 0
[69,] 1945 9 0.7326000 0
[70,] 1945 10 1.7880000 0
[71,] 1945 11 7.2036667 0
[72,] 1945 12 32.8864516 0
[1] "Rejeita-se H0 a um nível de significância de 5%, ou seja, a série não é aleatória."
[1] "Rejeita-se H0 a um nível de significância de 5%, ou seja, a série não é independente."
[1] "Não rejeita-se H0 a um nível de significância de 5%, ou seja, a série é homogênea."
[1] "Não rejeita-se H0 a um nível de significância de 5%, ou seja, a série é estacionária."
>

```

Figura 6 - Resultado da função no R. (Fonte: os autores.)

Existem algumas estações que possuem muito poucos dados, algumas de apenas um ano ou dois, e após excluídas as falhas acaba não sobrando nenhum dado. Por questões de algoritmo,

algumas partes da função não funcionam caso não haja um número mínimo de dados, portanto, um aviso será impresso na tela caso este número mínimo não tenha sido satisfeito.

Após finalizada a implementação da função no R, a função foi utilizada para realizar os quatro testes com todas as estações com dados disponíveis, utilizando os dados que já haviam sido importados para o R.

5. RESULTADOS E DISCUSSÕES

Após a organização dos dados e aplicação de ferramentas estatísticas, utilizando o programa R, para análise das 211 estações da Bacia do Rio Doce, verificou-se que para os dados de “Cota”, quatro estações (1,90%) não possuíam dados disponíveis e 12 (5,69%) apresentaram-se insuficientes para análise. Assim, pôde-se avaliar que das 195 estações que tiveram dados para “Cota” a média do período de avaliação foi de 26,91 anos, sendo que o maior período foi de 91 anos na estação de Rio Piracicaba e o menor de apenas um. Já para a análise de “Vazão”, 70 estações apresentaram dados não disponíveis, o que corresponde a 33,18% e cinco estações (2,37%) estavam com dados insuficientes; para essa análise o período médio de coleta de dados foi de 20,42 anos, de forma que o maior período foi de 86 anos e o menor também de apenas um.

As análises de aleatoriedade, homogeneidade, independência e estacionariedade, são realizadas de forma individual para cada estação tendo como a resposta “sim” ou “não”.

Os resultados obtidos tanto para “Cota” como para “Vazão” estão demonstrados na Tabela 1 a seguir e podem ser analisados utilizando a legenda abaixo.

Legenda da tabela:

- **S**: positivo para a análise;
- **S (NC)**: positivo, mas não confiável;
- **N**: negativo para a análise;
- **N (NC)**: negativo, mas não confiável;
- **DND**: dados não disponíveis (inexistentes);
- **DI**: dados insuficientes para executar a função.

Tabela 1 - Resultados da aplicação dos quatro testes

		S	S (NC)	N	N (NC)	DND	DI
Aleatória	Cota	0	12	173	10	4	12
	Vazão	0	3	131	2	70	5
Independente	Cota	3	0	192	0	4	12
	Vazão	1	0	135	0	70	5
Homogênea	Cota	58	20	106	11	4	12
	Vazão	70	6	57	3	70	5
Estacionária	Cota	76	0	119	0	4	12
	Vazão	72	0	64	0	70	5

Considerando apenas os resultados dos testes válidos, isto é, resultados positivos e negativos, é possível observar que a análise de homogeneidade e estacionariedade, para “Vazão”, apresentaram uma pequena maioria das estações com resultados positivos. As demais, e principalmente as de aleatoriedade e independência tiveram ampla maioria de resultados negativos, sendo que a primeira não apresentou nenhum resultado positivo. Se observa também que mesmo estações com períodos longos de acompanhamento tiveram falhas nas coletas de dados não permitindo análises precisas.

Avaliando as estações que apresentaram resultados positivos para os dados de cota e vazão, verifica-se que foram apenas 35 para estacionariedade, correspondendo a 16,59% e 34 (16,11%) para o teste de homogeneidade. No caso de aleatoriedade e independência, nenhuma estação apresentou resultado positivo em nenhuma estação de forma simultânea para cota e vazão. Esta informação leva ao entendimento de que a coleta de dados nem sempre ocorreu de maneira simultânea.

A análise dos dados paramétricos de estacionariedade é primordial para avaliação do comportamento das estações, desta forma, mesmo com falhas de coletas, o resultado dos testes indicou que 36,02% das estações são estacionárias para os dados de cotas e 34,12% para os dados de vazão.

Os resultados das análises para todas as estações da Bacia do Rio Doce se encontram nos Anexos A e B.

6. CONCLUSÃO

O alto número de estações, na Bacia do Rio Doce, com resultados negativos nos quatro testes realizados para os dados de cotas e de vazões, pode indicar a não confiabilidade desses resultados, além de apontar para a direção de dados com constantes variações, não seguindo um padrão, até porque se tratam de amostras de cota e vazão que sofrem consequências de épocas de cheia e seca em diferentes intensidades ao longo dos anos.

O grande retorno de resultados negativos dificulta as pesquisas e acompanhamentos de forma individual das estações e da bacia como um todo. Já o pequeno retorno de resultados positivos em poucas estações, levanta a questão de que forma estão sendo coletados estes dados, incluindo a periodicidade e até mesmo a forma de leitura dos mesmos. Espera-se que seja uma coleta padrão em todas as estações, contudo por serem realizadas manualmente podem apresentar diferenças de um trabalhador para outro da mesma estação.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BARBOSA, S. E S. **Análise dos dados hidrológicos e regionalização de vazões da Bacia do rio do Carmo**, Ouro Preto-MG. / Sylvio Elvis da Silva Barbosa. – Ouro Preto: UFOP, 2004.

BATISTA, A.L.; JÚNIOR, S.A.F.; DETZEL, D.H.M.; MINE, M.R.M.; FILL, H.D.O.A.; FERNANDES, C.F.; KAVISKI, E. Verificação da Estacionariedade de Séries Hidrológicas no Sul-Sudeste do Brasil. **Anais do XVIII Simpósio Brasileiro de Recursos Hídricos**. Campo Grande-MS, 2009.

COSTA, K. T. **Avaliação de distribuições de probabilidades das vazões médias diárias máximas anuais do Brasil**. Belo Horizonte, 2014. 194 p.

DETZEL, D. H. M.; BESSA, M. R.; VALLEJOS, C. A. V.; SANTOS, A. B.; THOMSEN, L. S.; MINE, M. R. M.; BLOOT, M. L.; ESTRÓCIO, J. P. Estacionariedade das Afluências às Usinas Hidrelétricas Brasileiras. **Revista Brasileira de Recursos Hídricos**, v. 16, nº 3, p. 95-111, jul./set. 2011.

MURPHY, K. W.; ELLIS, A. W. An assessment of the stationarity of climate and stream flow in watersheds of the Colorado River Basin. **Journal of Hydrology**, [S.L], n. 509, p. 454-473, dez. 2013.

NAGHETINI, M., PINTO, E. J. A. **Hidrologia estatística**. Belo Horizonte: CPRM, 2007. 552 p.

NERC. **Flood Studies Report, Vol. 1**. London: National Environmental Research Council, 1975.

TRÖGER, F.H.; GONDIM Fº, J.G.C.; NÓBREGA, M.T.; SUGAI, M.V.B. (2004). Verificação da Estacionariedade das Séries Anuais de Vazões Naturais das Usinas Hidrelétricas de Três Marias e Sobradinho. **Anais do VII Simpósio de Recursos Hídricos do Nordeste**, São Luis-MA. nov./dez. 2004.

TORRES, F. E.; FILHO, A. C.; TEODORO, P. E.; CORRÊA, C. C. G.; RIBEIRO, L. P.; CUNHA, E. R. Dimensionamento amostral para a estimação da média de precipitação pluvial mensal em locais do Estado do Mato Grosso do Sul. **Cienc. Rural**, Santa Maria-RS, vol.46 no.1, 2016

ANEXO A - RESULTADOS DOS TESTES COM OS DADOS DE COTA

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Abre Campo	56500000	1939	2016	N	N	N	N
Acacaia	56335000	1939	2016	N	N	N	N
Acacaia - Jusante	56335001	1939	2016	N	N	N	N
Acesita	56695000	1939	2016	N	N	S	S
Afonso Cláudio	56991000	1939	2016	N	N	S	S
Afonso Cláudio Montante	56990990	1939	2016	N	N	S	N
Aimorés	56990710	1939	2016	N	N	N (NC)	N
Aimorés (Casa das bombas)	56990715	1939	2016	N	N	N	N
Alto do Rio Doce	56050000	1938	2016	N	N	S	N
Alvorada de Minas	56760000	1939	2016	N	N	S	N
Ana Matos	56690000	1939	2016	N	N	N	N
Antonio Dias	56681000	1926	2016	N	N	N	S
Assarai	56989500	1939	2016	N	N	S	S
Assarai - Montante	56989400	1939	2016	N	N	S	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Baixo Guandu	56992000	1938	2016	N	N	N	N
Baixo Guandu	56992005	1939	2016	N	N	S (NC)	N
Barra de São Gabriel	56997000	1939	2016	N	N	N	N
Barra do Bananal	56987900	1939	2016	N (NC)	N	S (NC)	S
Barra do Capim	56989900	2015	2016	DI	DI	DI	DI
Barra do Capim	56990500	1939	2016	N	N	N	S
Barra do Córrego da Piaba	56993550	1939	2016	N	N	S	S
Barra do Cuieté	56940000	1939	2016	N	N	N	N
Barra do Cuieté	56941000	1939	2016	N	N	N	N
Barra do Cuieté Jusante	56940002	1939	2016	N	N	N	N
Barra do Mutum	56992500	1939	2016	N	N	S	S
Belo Oriente	56719998	1939	2016	N	N	S	S
Bicas	56336000	1939	2016	N	N	N	N
Bicas - Jusante	56336001	1939	2016	N	N	N (NC)	N
Boa Vista	56993300	1939	2016	N	N	S	S

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Bom Jesus do Galho	56565000	1939	2016	N	N	S	S
Bráz Pires	56055000	1938	2016	N	N	S	S
Brecha	56100000	1939	2016	N	N	N	N
Cachoeira Capivari	56630000	1937	2016	N	N	S	S
Cachoeira da fumaça	56230000	1939	2016	N	N	S	S
Cachoeira da Neblina	56974000	1938	2016	N	N	N	S
Cachoeira da Onça	56997200	1939	2016	N	N	S	S
Cachoeira do Guarara	56012000	1945	1965	N	N	N	N
Cachoeira do Oito	56995000	1939	1966	N	N	S	N
Cachoeira Dona Rita	56780000	1943	1950	N	N	S	S
Cachoeira Dona Rita	56780001	1963	1963	N (NC)	N	N (NC)	N
Cachoeira dos Óculos	56540001	1973	1973	S (NC)	N	S (NC)	S
Cachoeira dos Óculos - Montante	56539000	1975	2015	N	N	N	N
Cachoeira Escura	56720000	1940	1986	N	N	N	N
Cachoeira Suaçuí	56847500	1939	1985	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Campanário	56900000	1940	2013	N	N	N	N
Capim Cheiroso	56635000	1955	1959	N	N	S	S
Caputira	56470000	1944	1957	N	N	S	S
Caratinga	56925001	2006	2006	S (NC)	N	S (NC)	S
Caratinga	56924800	-	-	DND	DND	DND	DND
Caratinga	56925000	-	-	DND	DND	DND	DND
Carrapato (Brumal)	56640000	1955	2015	N	N	N	S
CGH Cachoeira Alta - Jusante	56962080	2012	2013	DI	DI	DI	DI
CGH Matipó - Jusante	56458080	-	-	DI	DI	DI	DI
Chalé	56985000	1939	1956	N	N	N	N
Chapada	56182000	1940	1965	N	N	S	S
Colatina	56994500	1938	2016	N	N	S	S
Colatina	56994505	1960	1986	N	N	S (NC)	S
Colatina Corpo de Bombeiros	56994510	1985	2016	N	N	S	N
Colatina Jusante	56994502	1976	1983	N	N	S	S

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Colégio Caraça	56632000	1937	1964	N	N	N	N
Colônia	56205000	1940	1942	N	N	N (NC)	N
Conceição do Mato Dentro	56750000	1946	2015	N	N	S	N
Conceição do Rio Acima	56631000	1937	1961	N	N	N	N
Conselheiro Pena	56943000	1945	1965	N	N	N	S
Coronel Fabriciano	56700000	1938	1974	N	N	N	N
Cuiabana	56140005	1939	1940	DI	DI	DI	DI
Desterro de Melo	56040000	1941	1965	N	N	N	N
Dom Cavati	56935000	1966	2015	N	N	S	S
Dom Joaquim	56765000	1946	2015	N	N	N	N
Dom Silvério	56340000	1950	1965	N	N	S	S
Dores de Manhumirim	56983000	1940	2015	N	N	N	N
ETA (São Bento Mineração)	56631900	2006	2015	N	N	N	N
Fazenda Aconchego	56846890	2015	2015	S (NC)	S	N (NC)	S
Fazenda Alegria	56590000	1970	1976	N	N	S	S

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Fazenda Bananal	56685000	1946	1947	S (NC)	N	N (NC)	N
Fazenda Barraca	56787000	1966	2015	N	N	N	N
Fazenda Boa Esperança	56986000	1940	1965	N	N	S	N
Fazenda Bragança	56976000	1972	2015	N	N	S	S
Fazenda Bretz	56847000	1972	1977	N (NC)	N	S (NC)	S
Fazenda Bretz - Montante	56846900	1997	2013	N	N	S	S
Fazenda Cachoeira Dantas	56425000	1981	2016	N	N	N	N
Fazenda Caldeirões	56246000	1943	1945	N (NC)	N	N (NC)	S
Fazenda Califórnia	56999005	1965	1966	DI	DI	DI	DI
Fazenda Caporanga	56999000	1969	1971	S (NC)	N	S (NC)	S
Fazenda Cibrão	56215000	1947	1950	DI	DI	DI	DI
Fazenda Cibrão	56220000	-	-	DND	DND	DND	DND
Fazenda Corrente	56845000	1952	2016	N	N	N	N
Fazenda da Vargem	56667000	1945	1965	N	N	N	N
Fazenda do retiro	56005000	1937	1960	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Fazenda Florença	56780005	1951	1957	N (NC)	N	N (NC)	N
Fazenda Maita	56380000	1938	1942	N	N	S	S
Fazenda Nova Floresta	56988000	1938	1956	N	N	N	N
Fazenda Ocidente	56337000	1938	2012	N	N	N	N
Fazenda Olaria	56670000	1941	1965	N	N	N	N
Fazenda Paraíso	56240000	1930	2016	N	N	S	N
Fazenda Urupuca	56880000	1968	1991	N	N	N	N
Fazenda Vargem Alegre	56960005	1983	2016	N	N	N	N
Fazenda Varginha	56090000	1941	2016	N	N	N	S
Ferros	56775000	1940	2016	N	N	S	N
Fojo	56185000	1939	1948	N	N	S (NC)	S
Furquim	56148000	1934	1972	N	N	N	N
Galiléia	56923000	1944	1947	S (NC)	N	S (NC)	N
Gondo	56744500	1954	1965	N	N	S	S
Governador Valadares	56850000	1938	2016	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Igrejinha - Aimorés	56990700	1953	1973	N	N	N	N
Inhapim	56928000	2005	2016	N	N	N	N
Instituto Florestal Raul Soares	56510000	1982	2016	N	N	S	S
Ipanema	56988500	1938	2016	N	N	N	N
Ipatinga	56710000	1974	1986	S (NC)	N	N (NC)	N
Itaguaçu	56993005	1952	1968	N	N	S	S
Itaguaçu	56993000	1965	1977	N	N	N	N
Itaguaçu - Jusante	56993002	1976	2010	N	N	S	S
Itarama	56992900	1958	1966	N	N	N	S
Jampruca	56915500	1973	2010	N	N	N	N
Jequeri	56400000	1938	1965	N	N	S	S
Jusante Corrego da Piaba	56993551	1974	2016	N	N	N	N
Lagoa do Martins	56999050	1964	1981	N	N	N	N
Laranja da Terra	56991500	1967	2016	N	N	N	N
Limoeiro	56155000	1940	1965	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Linhares	56998100	2002	2008	N	N	S	S
Linhares	56998005	1960	1968	N	N	N	S
Linhares	56998000	1940	1965	N	N	S	N
Linhares Cais do Porto	56998200	2003	2016	N	N	N	N
Manhuaçu	56960000	1936	1983	N	N	N	N
Marimbondo	56150000	1940	1966	N	N	N	N
Mario de Carvalho	56696000	1986	2016	N	N	N	N
Matadouro - Aimorés	56990705	1955	1983	DI	DI	DI	DI
Matipó	56460000	1965	2016	N	N	N	N
Mutum	56989001	1973	2016	N	N	S	S
Mutum	56989000	1938	1966	N	N	N	S
Mutum Preto	56350000	2001	2010	N	N	S	S
Naque do Meio	56830000	1938	1961	N	N	N	N
Naque Velho	56825000	1974	2016	N	N	N	N
Nova Era	56660000	1938	1998	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Nova Era - Montante	56659999	1974	1980	N	N	N	S
Nova Era IV	56659998	1989	2016	N	N	N	N
Nova Era Telemétrica	56661000	2002	2016	N	N	N	N
Olaria	56207000	1939	1943	N (NC)	N	S (NC)	S
Ouro Preto	56140000	1942	1956	N	N	N	N
Pai Tomás	56152000	1940	1965	N	N	N	N
Palmital	56020000	1944	1965	N	N	N	N
Parada da Independencia	56963000	1944	1952	N	N	N	S
Paraiso de Ipanema	56982000	1938	1965	N	N	N	N
PCH Benj. Mario Baptista - Mont.	56960080	2008	2013	DI	DI	DI	DI
PCH Neblina - Jusante	56972080	2008	2013	DI	DI	DI	DI
PCH Neblina - Montante	56960180	2008	2013	DI	DI	DI	DI
Pedra Corrida	56835000	1938	1962	N	N	S (NC)	S
Pingo D'Água	56570000	1974	2016	N	N	N	N
Piranga	56028000	1938	2016	N	N	S	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Ponte das Araras	56815000	1944	1955	N	N	N	N
Ponte do Brumado	56633000	1931	1946	N	N	S	S
Ponte do Caboclo	56195000	1938	1943	N	N	S	S
Ponte do Carvalho	56105000	1949	1961	N	N	S	S
Ponte do Pancas	56995500	1965	2016	N	N	S	S
Ponte do Pancas	56995500	1965	2016	N	N	S	S
Ponte do Peres	56430000	1965	1978	N	N	N	S
Ponte do Peres Jusante	56430005	1978	1981	DI	DI	DI	DI
Ponte do São Lourenço	56010000	1938	1965	N	N	N	N
Ponte Itajuru	56631500	1931	1946	N	N	S	S
Ponte Itatiaia	56160000	1940	1966	N	N	N	N
Ponte Itatiáia	56165000	1940	1946	N	N	N	N
Ponte Nova	56110000	1936	1975	N	N	N	N
Ponte Nova Jusante	56110005	1974	2016	N	N	N	N
Ponte Santa Rita	56158000	1940	1965	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Porto da Balsa	56820005	1947	1960	N	N	N	N
Porto da Estrela	56820000	1944	1958	S (NC)	N	S (NC)	S
Porto das Bibocas	56350000	1938	1942	N	N	S (NC)	S
Porto Firme	56075000	1938	2016	N	N	N	N
Porto Santa Rita	56846000	1975	2016	N	N	N	N
Prazeres	56210000	1944	1948	N (NC)	N	S (NC)	S
Presidente Bernardes	56070000	-	-	DND	DND	DND	DND
Raul Soares	56485000	1936	1977	N	N	N	N
Raul Soares Montante	56484998	1976	2016	N	N	N	N
Resplendor	56947000	1945	1961	N	N	S	N
Resplendor	56948000	1938	1989	N	N	N	N
Resplendor Jusante	56948005	1983	2004	N	N	S	S
Rio Casca	56415000	1930	2016	N	S	S	N
Rio Piracicaba	56610000	1925	2016	N	S	N	N
Santa Bárbara do Leste	56923800	2010	2016	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Santa Maria	56994000	1941	1967	N	N	N	N
Santa Maria de Itabira	56782000	1962	1974	N	N	N	N
Santa Maria do Suaçuí	56870000	1965	2016	N	N	N	N
Santa Rita de Minas	56924500	2004	2016	N	N	N	N
Santana de Manhuaçu	56967000	1941	1965	N	N	N	N
Santo Antônio do Munhuaçu	56978000	1965	2016	N	N	N	S
São Caetano de Mariana	56145000	1930	1965	N	N	N	N
São Domingos do Prata	56665000	1948	1962	N	N	N	N
São Miguel do Anta	56385000	1965	2016	N	N	N	N
São Pedro do Avaí	56465000	1958	1961	N	N	S (NC)	N
São Pedro do Suaçuí	56860000	1965	2016	N	N	N	N
São Sebastião da Encruzilhada	56990000	1938	2016	N	N	N	N
Senador Firmino	56065000	1941	2016	N	N	N	N
Senhora do Porto	56800000	1945	2016	N	N	N	N
Seriquite	56085000	1941	2016	N	N	N	N

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Serraria Vivacqua	56990720	1957	1981	N (NC)	N	N (NC)	N
Tabuleiro	56924100	2008	2016	S (NC)	N	S (NC)	S
Tabuleiro	56924000	2004	2008	N (NC)	N	S (NC)	S
Tumiritinga	56919500	1972	1973	S (NC)	N	S (NC)	S
Tumiritinga	56920000	1972	2016	N	N	N	N
Ubaporanga	56926000	2004	2008	N (NC)	N	N (NC)	N
Ubaporanga	56926100	2008	2015	DI	DI	DI	DI
UHE Mascarenhas Montante 1	56992380	2002	2013	N	N	N	N
Usina Ana Florência	56130000	1950	1965	N	N	N	N
Usina do Laje	56927000	1940	1943	S (NC)	N	S (NC)	S
Usina do Pé do Morro	56015000	1941	1949	N	N	S	S
Usina Peti	56650000	1930	1946	N	N	S	S
Vargem do Tijucal	56170000	1938	1957	N	N	S	S
Vermelhor Velho	56520000	1944	2010	N	N	S	S
Vila Matias	56892000	1938	1975	N	N	S	S

ESTAÇÃO	CÓDIGO	COTA					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Vila Matias - Montante	56891900	1974	2016	N	N	N	N
Vila Verde	56996000	1974	1976	S (NC)	N	S (NC)	S

ANEXO B – RESULTADOS DOS TESTES COM OS DADOS DE VAZÃO

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Abre Campo	56500000	1939	2013	N	N	N	N
Acacaia	56335000	1940	1975	N	N	N	N
Acacaia - Jusante	56335001	1975	2015	N	N	N	N
Acesita	56695000	1974	1985	N	N	N	S
Afonso Cláudio	56991000	1965	1976	N	N	S	S
Afonso Cláudio Montante	56990990	1978	2015	N	N	S	S
Aimorés	56990710	-	-	DND	DND	DND	DND
Aimorés (Casa das bombas)	56990715	-	-	DND	DND	DND	DND
Alto do Rio Doce	56050000	1938	1965	N	N	N	N
Alvorada de Minas	56760000	-	-	DND	DND	DND	DND
Ana Matos	56690000	1952	1964	N	N	N	N
Antonio Dias	56681000	1926	1951	N	N	N	S
Assarai	56989500	1938	1977	N	N	N	N
Assarai - Montante	56989400	1976	2015	N	N	S	N

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Baixo Guandu	56992000	1941	2015	N	N	N	S
Baixo Guandu	56992005	-	-	DND	DND	DND	DND
Barra de São Gabriel	56997000	1967	2015	N	N	N	S
Barra do Bananal	56987900	-	-	DND	DND	DND	DND
Barra do Capim	56989900	-	-	DI	DI	DI	DI
Barra do Capim	56990500	1948	1965	N	N	S	S
Barra do Córrego da Piaba	56993550	-	-	DND	DND	DND	DND
Barra do Cuieté	56940000	1938	1975	N	N	S	S
Barra do Cuieté	56941000	1972	1979	N	N	N	N
Barra do Cuieté Jusante	56940002	1975	2016	N	N	N	N
Barra do Mutum	56992500	1962	1965	N	N	N	N
Belo Oriente	56719998	1986	2016	N	N	S	S
Bicas	56336000	-	-	DND	DND	DND	DND
Bicas - Jusante	56336001	1939	1942	N	N	S (NC)	S
Boa Vista	56993300	-	-	DND	DND	DND	DND

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Bom Jesus do Galho	56565000	1941	1982	N	N	S	S
Bráz Pires	56055000	1938	2016	N	N	S	N
Brecha	56100000	-	-	DND	DND	DND	DND
Cachoeira Capivari	56630000	-	-	DND	DND	DND	DND
Cachoeira da fumaça	56230000	-	-	DND	DND	DND	DND
Cachoeira da Neblina	56974000	1938	1965	N	N	N	S
Cachoeira da Onça	56997200	-	-	DND	DND	DND	DND
Cachoeira do Guarara	56012000	1946	1963	N	N	S	S
Cachoeira do Oito	56995000	1944	1966	N	N	S	N
Cachoeira Dona Rita	56780000	1963	1963	N (NC)	N	N (NC)	N
Cachoeira Dona Rita	56780001	-	-	DND	DND	DND	DND
Cachoeira dos Óculos	56540001	1973	1973	S (NC)	N	S (NC)	S
Cachoeira dos Óculos - Montante	56539000	1975	2014	N	N	S	S
Cachoeira Escura	56720000	1941	1985	N	N	S	S
Cachoeira Suaçui	56847500	-	-	DND	DND	DND	DND

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Campanário	56900000	1938	2012	N	N	S	N
Capim Cheiroso	56635000	1955	1959	N	N	S	S
Caputira	56470000	1944	1957	N	N	S	S
Caratinga	56925001	1939	1965	N	N	N	N
Caratinga	56924800	-	-	DND	DND	DND	DND
Caratinga	56925000	-	-	DND	DND	DND	DND
Carrapato (Brumal)	56640000	1955	2013	N	N	N	S
CGH Cachoeira Alta - Jusante	56962080	-	-	DND	DND	DND	DND
CGH Matipó - Jusante	56458080	-	-	DI	DI	DI	DI
Chalé	56985000	1939	1956	N	N	S	S
Chapada	56182000	1940	1964	N	N	S	S
Colatina	56994500	1939	2005	N	N	N	N
Colatina	56994505	-	-	DND	DND	DND	DND
Colatina Bombeiros	56994510	1985	2016	N	N	S	N
Colatina Jusante	56994502	1976	1983	N	N	S	S

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Colégio Caraça	56632000	1937	1954	N	N	S	N
Colônia	56205000	-	-	DND	DND	DND	DND
Conceição do Mato Dentro	56750000	1947	2014	N	N	S	N
Conceição do Rio Acima	56631000	1938	1956	N	N	N	N
Conselheiro Pena	56943000	1945	1965	N	N	N	S
Coronel Fabriciano	56700000	1972	1974	N	N	N (NC)	S
Cuiabana	56140005	-	-	DND	DND	DND	DND
Desterro de Melo	56040000	1941	1963	N	N	N	N
Dom Cavati	56935000	1966	2014	N	N	S	S
Dom Joaquim	56765000	1947	2014	N	N	N	N
Dom Silvério	56340000	1950	1964	N	N	S	N
Dores de Manhumirim	56983000	1939	2004	N	N	S	S
ETA (São Bento Mineração)	56631900	-	-	DND	DND	DND	DND
Fazenda Aconchego	56846890	-	-	DND	DND	DND	DND
Fazenda Alegria	56590000	1972	1976	N	N	S	S

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Fazenda Bananal	56685000	1946	1947	S (NC)	N	N (NC)	N
Fazenda Barraca	56787000	1966	2015	N	N	N	N
Fazenda Boa Esperança	56986000	1940	1965	N	N	S	N
Fazenda Bragança	56976000	1972	2014	N	N	S	S
Fazenda Bretz	56847000	1972	1977	N	N	S (NC)	S
Fazenda Bretz - Montante	56846900	1997	2013	N	N	S	S
Fazenda Cachoeira Dantas	56425000	1981	2016	N	N	N	N
Fazenda Caldeirões	56246000	-	-	DND	DND	DND	DND
Fazenda Califórnia	56999005	-	-	DND	DND	DND	DND
Fazenda Caporanga	56999000	-	-	DND	DND	DND	DND
Fazenda Cibrão	56215000	-	-	DND	DND	DND	DND
Fazenda Cibrão	56220000	-	-	DND	DND	DND	DND
Fazenda Corrente	56845000	1952	2016	N	N	N	N
Fazenda da Vargem	56667000	1945	1965	N	N	N	N
Fazenda do retiro	56005000	1937	1954	N	N	S	S

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Fazenda Florença	56780005	1951	1953	N (NC)	N	S (NC)	S
Fazenda Maita	56380000	-	-	DND	DND	DND	DND
Fazenda Nova Floresta	56988000	1938	1956	N	N	N	N
Fazenda Ocidente	56337000	1938	2015	N	N	S	N
Fazenda Olaria	56670000	1941	1965	N	N	N	S
Fazenda Paraíso	56240000	1930	2016	N	N	N	N
Fazenda Urupuca	56880000	1968	1991	N	N	S	S
Fazenda Vargem Alegre	56960005	1983	2016	N	N	S	S
Fazenda Varginha	56090000	1941	2016	N	N	S	N
Ferros	56775000	1940	2015	N	N	N	N
Fojo	56185000	1939	1948	N	N	S (NC)	S
Furquim	56148000	1934	1972	N	N	N	N
Galiléia	56923000	-	-	DND	DND	DND	DND
Gondo	56744500	-	-	DND	DND	DND	DND
Governador Valadares	56850000	1969	2016	N	N	N	S

ESTAÇÃO			CÓDIGO	VAZÃO					
				Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Igrejinha - Aimorés			56990700	-	-	DND	DND	DND	DND
Inhapim			56928000	-	-	DND	DND	DND	DND
Instituto Soares	Florestal	Raul	56510000	1982	2016	N	N	S	N
Ipanema			56988500	1938	2015	N	N	S	S
Ipatinga			56710000	-	-	DND	DND	DND	DND
Itaguaçu			56993005	1965	1977	N	N	N	N
Itaguaçu			56993000	-	-	DND	DND	DND	DND
Itaguaçu - Jusante			56993002	1976	2010	N	N	N	S
Itarama			56992900	-	-	DND	DND	DND	DND
Jampruca			56915500	1973	2005	N	N	N	N
Jequeri			56400000	1938	1965	N	N	S	S
Jusante Corrego da Piaba			56993551	1974	2005	N	N	N	N
Lagoa do Martins			56999050	-	-	DND	DND	DND	DND
Laranja da Terra			56991500	1967	2005	N	N	S	S
Limoeiro			56155000	1940	2004	N	N	S	S

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Linhares	56998100	1967	1994	N	N	S	S
Linhares	56998005	-	-	DND	DND	DND	DND
Linhares	56998000	-	-	DND	DND	DND	DND
Linhares Cais do Porto	56998200	-	-	DND	DND	DND	DND
Manhuaçu	56960000	1936	1983	N	N	S	S
Marimondo	56150000	1940	1954	N	N	S	S
Mario de Carvalho	56696000	1986	2005	N	N	S	S
Matadouro - Aimorés	56990705	-	-	DND	DND	DND	DND
Matipó	56460000	1965	2005	N	N	N	S
Mutum	56989001	1973	2005	N	N	S	S
Mutum	56989000	1938	1966	N	N	N	S
Mutum Preto	56350000	-	-	DND	DND	DND	DND
Naque do Meio	56830000	-	-	DND	DND	DND	DND
Naque Velho	56825000	1974	2007	N	N	N	S
Nova Era	56660000	1938	1972	DI	DI	DI	DI

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Nova Era - Montante	56659999	1974	1980	N	N	S	S
Nova Era IV	56659998	1989	2005	N	N	S	S
Nova Era Telemétrica	56661000	2014	2016	S (NC)	S	S (NC)	S
Olaria	56207000	-	-	DND	DND	DND	DND
Ouro Preto	56140000	-	-	DND	DND	DND	DND
Pai Tomás	56152000	1940	1954	N	N	N	S
Palmital	56020000	-	-	DND	DND	DND	DND
Parada da Independencia	56963000	1944	1952	N	N	S	S
Paraiso de Ipanema	56982000	1938	1953	N	N	S	S
PCH Benj. Mario Baptista - Mont.	56960080	2008	2010	DI	DI	DI	DI
PCH Neblina - Jusante	56972080	2008	2012	DI	DI	DI	DI
PCH Neblina - Montante	56960180	-	-	DND	DND	DND	DND
Pedra Corrida	56835000	-	-	DND	DND	DND	DND
Pingo D'Água	56570000	1974	2005	N	N	S	S
Piranga	56028000	1938	2005	N	N	S	N

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Ponte das Araras	56815000	1944	1955	N	N	N	N
Ponte do Brumado	56633000	1938	1946	N	N	S	S
Ponte do Caboclo	56195000	1938	1943	N	N	S	S
Ponte do Carvalho	56105000	1952	1961	N	N	N	S
Ponte do Pancas	56995500	1965	2005	N	N	S	S
Ponte do Pancas	56995500	1965	2005	N	N	S	S
Ponte do Peres	56430000	1938	1946	N	N	S	S
Ponte do Peres Jusante	56430005	-	-	DND	DND	DND	DND
Ponte do São Lourenço	56010000	1939	1965	N	N	N	N
Ponte Itajuru	56631500	1937	1946	N	N	S	S
Ponte Itatiaia	56160000	1940	1946	N	N	S	S
Ponte Itatiáia	56165000	-	-	DND	DND	DND	DND
Ponte Nova	56110000	1936	1975	N	N	N	N
Ponte Nova Jusante	56110005	1974	2005	N	N	S	N
Ponte Santa Rita	56158000	1940	1947	N	N	S	S

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Porto da Balsa	56820005	-	-	DND	DND	DND	DND
Porto da Estrela	56820000	-	-	DND	DND	DND	DND
Porto das Bibocas	56350000	-	-	DND	DND	DND	DND
Porto Firme	56075000	1938	2005	N	N	N	N
Porto Santa Rita	56846000	1975	2005	N	N	N	N
Prazeres	56210000	-	-	DND	DND	DND	DND
Presidente Bernardes	56070000	-	-	DND	DND	DND	DND
Raul Soares	56485000	1936	1977	N	N	S	S
Raul Soares Montante	56484998	1976	2005	N	N	N	N
Resplendor	56947000	1938	1982	N	N	N	N
Resplendor	56948000	-	-	DND	DND	DND	DND
Resplendor Jusante	56948005	1984	2004	N	N	S	S
Rio Casca	56415000	1930	2005	N	N	S	N
Rio Piracicaba	56610000	1925	2006	N	N	N	N
Santa Bárbara do Leste	56923800	-	-	DND	DND	DND	DND

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Santa Maria	56994000	1941	1967	N	N	N	N
Santa Maria de Itabira	56782000	-	-	DND	DND	DND	DND
Santa Maria do Suaçuí	56870000	1965	2005	N	N	N	N
Santa Rita de Minas	56924500	-	-	DND	DND	DND	DND
Santana de Manhuaçu	56967000	1941	1965	N	N	N	N
Santo Antônio do Munhuaçu	56978000	1965	2016	N	N	S	N
São Caetano de Mariana	56145000	1930	1965	N	N	N	N
São Domingos do Prata	56665000	1948	1962	N	N	N	N
São Miguel do Anta	56385000	1965	2005	N	N	S	S
São Pedro do Avaí	56465000	-	-	DND	DND	DND	DND
São Pedro do Suaçuí	56860000	1965	2005	N	N	N	N
São Sebastião da Encruzilhada	56990000	1938	1952	N	N	S	S
Senador Firmino	56065000	1941	2005	N	N	S	S
Senhora do Porto	56800000	1945	2005	N	N	N	N
Seriquite	56085000	1988	2005	N	N	S	N

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Serraria Vivacqua	56990720	-	-	DND	DND	DND	DND
Tabuleiro	56924100	-	-	DND	DND	DND	DND
Tabuleiro	56924000	-	-	DND	DND	DND	DND
Tumiritinga	56919500	1972	2003	N	N	N	N
Tumiritinga	56920000	-	-	DND	DND	DND	DND
Ubaporanga	56926000	-	-	DND	DND	DND	DND
Ubaporanga	56926100	-	-	DND	DND	DND	DND
UHE Mascarenhas Montante 1	56992380	-	-	DND	DND	DND	DND
Usina Ana Florência	56130000	1950	1961	N	N	S	S
Usina do Laje	56927000	-	-	DND	DND	DND	DND
Usina do Pé do Morro	56015000	-	-	DND	DND	DND	DND
Usina Peti	56650000	-	-	DND	DND	DND	DND
Vargem do Tijucal	56170000	1938	1954	N	N	S	N
Vermelhor Velho	56520000	1944	2005	N	N	S	N
Vila Matias	56892000	1938	1975	N	N	S	N

ESTAÇÃO	CÓDIGO	VAZÃO					
		Ano Inicial	Ano Final	Aleatória	Independente	Homogênea	Estacionária
Vila Matias - Montante	56891900	1974	2005	N	N	N	N
Vila Verde	56996000	-	-	DND	DND	DND	DND

ANEXO C – FUNÇÃO PARA REALIZAR OS TESTES

```

testes<-function(dados, ano0, anof)
{
  ### CALCULA MÉDIAS MENSAIS E RETORNA MATRIZ COM SEIS COLUNAS
  anos<-anof-ano0+1
  meses<-anos*12
  medmensal<-matrix(nrow=meses,ncol=6)

  # COLUNA 1 - ANOS DA SÉRIE
  # COLUNA 4 - MESES DO ANO
  {
    a<-0
    am<-1
    for (iin1:anos)
    {
      for (jin1:12)
      {
        medmensal[am,1]<-ano0+a
        medmensal[am,4]<-j
        am<-am+1
      }
      a<-a+1
    }
  }

  # COLUNA 2 - MÉDIAS ANUAIS
  # COLUNA 3 - % DE FALHAS NAQUELE ANO
  # COLUNA 5 - MÉDIAS MENSAIS
  # COLUNA 6 - % DE FALHAS NAQUELE MÊS
  {
    j<-0# contador de dias
    l<-0# contador de anos
    m<-1# contador de meses

    for (iinano0:anof)
    {
      if (i%%4==0&& (i%%400==0 || i%%100!=0))
      {
        l<-l+1
        x<-0
        y<-0
        for (kin1:366)
        {
          x[k]<-dados[j+k]
          if (is.na(dados[j+k])==TRUE)
          {
            y<-y+1
          }
        }
      }

      media<-mean(x,na.rm=TRUE)
      falhas<-((y/366)*100)
      for (zinm:(m+1))
      {
        medmensal[z,2]<-media
        medmensal[z,3]<-falhas
      }
    }
  }
}

```

```

}

# Calcula as médias mensais
x<-0
y<-0
for (kin1:31) # Janeiro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (kin1:29) # Fevereiro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/29)*100)
j<-j+29
m<-m+1

x<-0
y<-0
for (kin1:31) # Março
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (kin1:30) # Abril
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)

```

```

medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

x<-0
y<-0
for (kin1:31) # Maio
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

```

```

x<-0
y<-0
for (kin1:30) # Junho
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

```

```

x<-0
y<-0
for (k in 1:31) # Julho
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

```

```

x<-0
y<-0
for (k in 1:31) # Agosto
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}

```

```

    }
  }
  medmensal[m, 5] <- mean(x, na.rm=TRUE)
  medmensal[m, 6] <- ((y/31)*100)
  j <- j+31
  m <- m+1

  x <- 0
  y <- 0
  for (k in 1:30) # Setembro
  {
    x[k] <- dados[j+k]
    if (is.na(dados[j+k]) == TRUE)
    {
      y <- y+1
    }
  }
  medmensal[m, 5] <- mean(x, na.rm=TRUE)
  medmensal[m, 6] <- ((y/30)*100)
  j <- j+30
  m <- m+1

  x <- 0
  y <- 0
  for (k in 1:31) # Outubro
  {
    x[k] <- dados[j+k]
    if (is.na(dados[j+k]) == TRUE)
    {
      y <- y+1
    }
  }
  medmensal[m, 5] <- mean(x, na.rm=TRUE)
  medmensal[m, 6] <- ((y/31)*100)
  j <- j+31
  m <- m+1

  x <- 0
  y <- 0
  for (k in 1:30) # Novembro
  {
    x[k] <- dados[j+k]
    if (is.na(dados[j+k]) == TRUE)
    {
      y <- y+1
    }
  }
  medmensal[m, 5] <- mean(x, na.rm=TRUE)
  medmensal[m, 6] <- ((y/30)*100)
  j <- j+30
  m <- m+1

  x <- 0
  y <- 0
  for (k in 1:31) # Dezembro
  {
    x[k] <- dados[j+k]
    if (is.na(dados[j+k]) == TRUE)
    {

```

```

        y<-y+1
    }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1
}

else
{
  l<-l+1
  x<-0
  y<-0
  for (k in 1:365)
  {
    x[k]<-dados[j+k]
    if (is.na(dados[j+k])==TRUE)
    {
      y<-y+1
    }
  }
  media<-mean(x,na.rm=TRUE)
  falhas<-((y/365)*100)
  for (z in m:(m+11))
  {
    medmensal[z,2]<-media
    medmensal[z,3]<-falhas
  }

  # Calcula as médias mensais
  x<-0
  y<-0
  for (k in 1:31) # Janeiro
  {
    x[k]<-dados[j+k]
    if (is.na(dados[j+k])==TRUE)
    {
      y<-y+1
    }
  }
  medmensal[m,5]<-mean(x,na.rm=TRUE)
  medmensal[m,6]<-((y/31)*100)
  j<-j+31
  m<-m+1

  x<-0
  y<-0
  for (k in 1:28) # Fevereiro
  {
    x[k]<-dados[j+k]
    if (is.na(dados[j+k])==TRUE)
    {
      y<-y+1
    }
  }
  medmensal[m,5]<-mean(x,na.rm=TRUE)
  medmensal[m,6]<-((y/28)*100)
  j<-j+28

```

```

m<-m+1

x<-0
y<-0
for (k in 1:31) # Março
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (k in 1:30) # Abril
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

x<-0
y<-0
for (k in 1:31) # Maio
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (k in 1:30) # Junho
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)

```

```

medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

x<-0
y<-0
for (k in 1:31) # Julho
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (k in 1:31) # Agosto
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (k in 1:30) # Setembro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

x<-0
y<-0
for (k in 1:31) # Outubro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}

```

```

}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1

x<-0
y<-0
for (k in 1:30) # Novembro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/30)*100)
j<-j+30
m<-m+1

x<-0
y<-0
for (k in 1:31) # Dezembro
{
  x[k]<-dados[j+k]
  if (is.na(dados[j+k])==TRUE)
  {
    y<-y+1
  }
}
medmensal[m,5]<-mean(x,na.rm=TRUE)
medmensal[m,6]<-((y/31)*100)
j<-j+31
m<-m+1
}
}
}

# Remove anos que possuem pelo menos um mês com mais de P % de
falhas

P<-10
medmensal1<-medmensal
medmensal1<-medmensal1[, -2]
medmensal1<-medmensal1[, -2]

N<-nrow(medmensal1)
anos<-N/12

y<-1 # conta meses
deleta<-0
linhas<-0
w<-1
for (i in 1:anos)
{
  deleta<- "FALSE"
  for (j in 1:12)
  {

```



```

    if (medmensal1[y,4]>=P)
    {
        deleta<-"TRUE"
    }
    y<-y+1
}
if (deleta=="TRUE")
{
    for (k in 12:1)
    {
        linhas[w]<-(y-k)
        w<-w+1
    }
}
}

medmensal1<-medmensal1[-linhas,]
print(medmensal1)
# Coluna 1 - Ano
# Coluna 2 - Mês
# Coluna 3 - Média mensal
# Coluna 4 - % de falha mensal;
# Exclui coluna 2 para facilitar o resto da função
medmensal1<-medmensal1[,-2]

#####
### TESTE DE ALEATORIEDADE
{
N<-nrow(medmensal1)
dados<-0
for (i in 1:N)
{
    dados[i]<-medmensal1[i,2]
}

p <- 0
N <- length(dados)-1
for (i in 2:N)
{
    if (((dados[i] > dados[i-1]) && (dados[i] > dados[i+1])) ||
        ((dados[i] < dados[i-1]) && (dados[i] < dados[i+1])))
    {
        p <- p + 1
    }
}
p <- p + 2
N <- length(dados)
E.p <- 2*(N-2)/3
Var.p <- (16*N - 29)/90
T <- ((p-E.p)/sqrt(Var.p))
alfa<-0.05
z <- qnorm(1-alfa/2,0,1)

if (N<30)
{
    print("ATENÇÃO! Os dados disponíveis não são o suficiente para um
    teste de aleatoriedade confiável (N<30).")
}
}

```

```

if (abs(T)>z)
{
  print("Rejeita-se H0 a um nível de significância de 5%, ou seja, a
série não é aleatória.")
}
else
{
  print("Não rejeita-se H0 a um nível de significância de 5%, ou
seja, a série é aleatória.")
}
}

#####
### TESTE DE INDEPENDÊNCIA
{
N<-nrow(medmensal1)
dados<-0
for (i in 1:N)
{
  dados[i]<-medmensal1[i,2]
}

s2<-0
s4<-0
dadoslinha<-0
for (i in 1:N)
{
  dadoslinha[i]<-(dados[i]-mean(dados))
  s2<-s2+(dados[i]-mean(dados))**2
  s4<-s4+(dados[i]-mean(dados))**4
}

R<-0
for (iin1:(N-1))
{
  R<-R+(dadoslinha[i]*dadoslinha[i+1])
}

R<-R+(dadoslinha[1]*dadoslinha[N])
E.r<-(-s2)/(N-1)
Var.R<-(((s2**2)-s4)/(N-1))+(((s2**2)-(2*s4))/((N-1)*(N-2)))-
((s2**2)/((N-1)**2))
T<-((R-E.r)/(sqrt(Var.R)))
alfa<-0.05
z<- qnorm(1-alfa/2,0,1)

if (abs(T)>z)
{
  print("Rejeita-se H0 a um nível de significância de 5%, ou seja, a
série não é independente.")
}
else
{
  print("Não rejeita-se H0 a um nível de significância de 5%, ou
seja, a série é independente.")
}
}

```

```
#####
### TESTE DE HOMOGENEIDADE
{
N<-nrow(medmensal1)
N1<-floor(N/2)
N2<-N-N1
dados<-matrix(nrow=N,ncol=3)

for (iin1:N)
{
  dados[i,1]<-i
  dados[i,2]<-medmensal1[i,2]
}

if (N1<20)
{
  print("ATENÇÃO! Os dados disponíveis não são o suficiente para um
teste de homogeneidade confiável (N1,N2<20).")
}

mt1<-order(dados[,2])
dados1<-matrix(nrow=N,ncol=3)

for (iin1:N)
{
  dados1[i,1]<-dados[mt1[i],1]
  dados1[i,2]<-dados[mt1[i],2]
  dados1[i,3]<-i
}

mt2<-order(dados1[,1])
for (iin1:N)
{
  dados[i,1]<-dados1[mt2[i],1]
  dados[i,2]<-dados1[mt2[i],2]
  dados[i,3]<-dados1[mt2[i],3]
}

soma1<-0

for (iin1:N1)
{
  soma1<-soma1+dados[i,3]
}

V1<-((N1*N2)+((N1*(N1+1))/2)-soma1)
V2<-((N1*N2)-V1)
E.V<-((N1*N2)/2)
Var.V<-(((N1*N2)*(N1+N2+1))/12)
if (V1>V2)
{
  V<-V2
}
else
{
  V<-V1
}
T<-((V-E.V)/sqrt(Var.V))
alfa<-0.05
```

```

z<- qnorm(1-alfa/2,0,1)

if (abs(T)>z)
{
  print("Rejeita-se H0 a um nível de significância de 5%, ou seja, a
série não é homogênea.")
}
else
{
  print("Não rejeita-se H0 a um nível de significância de 5%, ou
seja, a série é homogênea.")
}
}

#####
### TESTE DE ESTACIONARIEDADE
{
N<-nrow(medmensal1)

soma<-0

for (iin1:N)
{
  soma<-soma+(((dados[i,3])-(dados[i,1]))**2)
}

rs<-(1-((6*soma)/((N**3)-N)))
E.rs<-0
Var.rs<-((1)/(N-1))
T<-((rs)/(sqrt(Var.rs)))
alfa<-0.05
z<- qnorm(1-alfa/2,0,1)

if (N<10)
{
  print("ATENÇÃO! Os dados disponíveis não são o suficiente para um
teste de estacionariedade confiável (N<10).")
}

else
{
  if (abs(T)>z)
  {
    print("Rejeita-se H0 a um nível de significância de 5%, ou seja,
a série não é estacionária.")
  }
  else
  {
    print("Não rejeita-se H0 a um nível de significância de 5%, ou
seja, a série é estacionária.")
  }
}
}
}

```