

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM SANEAMENTO,**  
**MEIO AMBIENTE E RECURSOS HÍDRICOS**

**Regionalização hidrológica de vazões mínimas  
por meio do Método dos Mínimos Quadrados  
Generalizados aplicada à bacia do Alto São  
Francisco**

**Sarah Figueiredo Luiz**

**Belo Horizonte**

**2013**

**Regionalização hidrológica de vazões mínimas por  
meio do Método dos Mínimos Quadrados  
Generalizados aplicada à bacia do Alto São  
Francisco**

**Sarah Figueiredo Luiz**

**Sarah Figueiredo Luiz**

**Regionalização hidrológica de vazões mínimas por  
meio do Método dos Mínimos Quadrados  
Generalizados aplicada à bacia do Alto São  
Francisco**

Dissertação apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Recursos Hídricos

Linha de pesquisa: Modelagem de processos hidrológicos

Orientador: Wilson dos Santos Fernandes

Belo Horizonte  
Escola de Engenharia da UFMG  
2013

L953r

Luiz, Sarah Figueiredo.

Regionalização hidrológica de vazões mínimas por meio do Método dos Mínimos Quadrados Generalizados aplicada à Bacia do Alto São Francisco [manuscrito] / Sarah Figueiredo Luiz. – 2013.  
xix., 134 f., enc.: il.

Orientador: Wilson dos Santos Fernandes.

Dissertação (mestrado) Universidade Federal de Minas Gerais, Escola de Engenharia.

Inclui bibliografia.

1. Engenharia sanitária – Teses. 2. Recursos hídricos – Desenvolvimento – Teses. 3. Mínimos quadrados – Teses. I. Fernandes, Wilson dos Santos. II. Universidade Federal de Minas Gerais, Escola de Engenharia. III. Título.

CDU: 628(043)



*“A persistência é o menor caminho do êxito.”*

*Charles Chaplin*

## AGRADECIMENTOS

Agradeço imensamente a Deus por ter me dado a grande oportunidade de fazer o que gosto na vida.

Aos meus pais, a minha mãe pela determinação e amor, ao meu pai pela vontade e a coragem de viver, aos meus irmãos e cunhadas e aos meus sobrinhos (as) por sempre me trazerem a alegria e a magia que só uma criança consegue ter.

Ao meu noivo Hudson, pelo companheirismo, pelo carinho e pela compreensão em todos os momentos.

Ao professor Wilson dos Santos Fernandes por ter sido um excelente orientador, e pelos ensinamentos preciosos, por ter me incentivado em todas as etapas da pesquisa e pela sua motivação em querer sempre aprender mais.

Ao professor Dirceu Reis pela assistência prestada em relação ao método GLS, por ter respondido aos e-mails com muita atenção.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG pelo apoio à pesquisa por meio do projeto APQ 04418-10 e também, pela bolsa a mim concedida durante todo o período em que estive cursando o Mestrado.

Aos professores Mário Cicareli e Éber José de Andrade pela motivação e por serem determinantes para meu crescimento acadêmico.

À Adriana Mendes, aluna de iniciação científica por ter me ajudado na coleta e análise dos dados primários referentes às estações e pelos momentos agradáveis que passamos juntas.

Aos professores e funcionários do EHR.

À minha querida amiga de infância, Vanessa Sampaio e a sua família, por terem sido a minha segunda família e ter me acolhido com amor e carinho.

Aos colegas e amigos que fizeram e fazem parte desta conquista e de todos os outros momentos da minha vida.

## RESUMO

O conhecimento de vazões em locais não monitorados é de grande importância em hidrologia. Daí a utilização de técnicas que possibilitem a transferência de informações de locais monitorados para locais sem monitoramento. Hosking e Wallis (1997) sugeriram a utilização da análise de frequência regional, pois esta incorpora mais informações que a análise local. Além disso, o estudo das vazões mínimas desempenha um importante papel, pois são fundamentais na avaliação da disponibilidade hídrica, na elaboração de projetos hidroelétricos e de irrigação, além de serem utilizadas como critério de outorga. A metodologia empregada neste trabalho foi estruturada com base na análise de estudos de regionalização, principalmente os que envolviam estatísticas de vazões médias mínimas como a  $Q_{7,10}$ . O modelo de regressão a ser avaliado é o Método dos Mínimos Quadrados Generalizados - *Generalized Least Square* (GLS), proposto por Stedinger e Tasker (1985, 1986a, 1986b), juntamente com o Método dos Mínimos Quadrados Ordinários - OLS (*Ordinary Least Square*) e com o Método dos Mínimos Quadrados Ponderados -WLS (*Weighted Least Square*), proposto por Tasker e Stedinger (1986). Tais métodos serão empregados em um estudo de regionalização de vazões mínimas de 7 dias de duração e 10 anos de período de retorno para a bacia do Alto São Francisco, sub-bacias 40 e 41. Inicialmente, a área de estudo foi dividida em 4 regiões supostamente homogêneas (foram atribuídos os códigos BA1, BA2, BA3 e BA4), não tendo sido realizada nenhuma análise estatística para corroborar que estas regiões realmente são homogêneas do ponto de vista hidrológico e estatístico. O critério utilizado para divisão dessas regiões foi através da análise das Unidades de Planejamento e Gestão de Recursos Hídricos (UPGRH) utilizadas pelo Instituto Mineiro de Gestão das Águas (IGAM). Para a determinação dos valores da vazão mínima  $Q_{7,10}$ , realizou-se a análise de frequência local para todas as estações que possuíam mais de 10 anos de dados observados, utilizando a distribuição de probabilidade Weibull 2 parâmetros. Tal distribuição também foi utilizada para a construção da matriz de covariância, ou seja, para determinação da covariância dos erros amostrais. A análise e a determinação do modelo final para cada uma das 4 regiões foi feita mediante a avaliação da variância média da predição (*average variance of prediction* - AVP) e os modelos de regressão gerados foram do tipo potencial (logaritmo do quantil). Os resultados obtidos mostraram que o método WLS teve melhor desempenho que o método GLS em relação à razão da inflação entre as variâncias (*variance inflation ratio* - VIR), pois a VIR para as quatro regiões homogêneas foi inferior a 1. De acordo com Griffis e Stedinger (2007), se a VIR for menor do que 1, então a regressão WLS é suficiente, não

havendo assim necessidade de utilizar o modelo GLS. Com o objetivo de avaliar os critérios relacionados aos resíduos do método OLS, aplicou-se testes de hipóteses de independência, normalidade, média igual a zero e variância constante ao nível de significância de 5%. Essa análise permitiu evidenciar que para as regiões BA2 e BA4, a utilização do método OLS não é recomendada, pois pelo menos uma premissa referente aos resíduos foi violada. Além disso, ao se avaliar a razão entre as variâncias do erro (*error variance ratio* - EVR), verificou-se que a variância do erro amostral é maior do que a variância do erro do modelo, logo o método OLS não deve ser empregado. Salienta-se que o método GLS é mais adequado, pois as vazões mínimas são fortemente correlacionadas, uma vez que método WLS não leva em consideração as correlações entre os fluxos concorrentes. Porém, para o caso das correlações serem pequenas, tal método poderia ser adotado. No entanto, na região estudada, os valores das correlações cruzadas são elevados. Sendo assim, concluiu-se que para o estudo em questão, o método GLS é o mais adequado. Recomenda-se, entretanto, que sejam feitas avaliações concernentes à matriz de covariância dos erros amostrais, pois a metodologia para estimativa da matriz de covariância é baseada nos quantis propriamente ditos, e não no logaritmo dos quantis. Essa transformação não linear certamente causa uma mudança nos elementos da matriz de covariância.

## ABSTRACT

Knowledge of flows at unmonitored sites is of great importance in hydrology. Hence the use of techniques that allow the transfer of information from monitored sites to unmonitored sites. Hosking and Wallis (1997) suggest the use of regional frequency analysis, since it incorporates more information than local analysis. Furthermore, the study of minimum flows plays an important role, since they are fundamental in assessing water availability, in the development of hydroelectric and irrigation projects, as well as being used as an authorization criterion. The methodology used in this study was structured based on an analysis of regionalization studies, primarily those involving minimum average flow rate statistics such as  $Q_{7,10}$ . The regression model to be evaluated is the Generalized Least Squares (GLS) method proposed by Stedinger and Tasker (1985, 1986a, 1986b), in conjunction with the Ordinary Least Squares (OLS) method and the Weighted Least Squares (WLS) method, proposed by Tasker and Stedinger (1986). Such methods will be employed in a study of the regionalization of low flows over a 7-day and 10-year return period for the Alto São Francisco basin, sub-basins 40 and 41. Initially, the study area was divided into 4, supposedly homogeneous, regions (assigned the codes BA1, BA2, BA3 and BA4), although no statistical analysis was performed to corroborate that these regions actually are homogeneous from a hydrological and statistical standpoint. The criterion used to divide these regions was an analysis of the Water Resources Planning and Management Units (UPGRH) used by the Minas Gerais Institute of Water Management (IGAM). To determine the values of the low flow  $Q_{7,10}$ , an analysis was performed of local frequency for all stations with more than 10 years of observed data using 2-parameter Weibull probability distribution. This distribution was also used to construct the covariance matrix, i.e. to determine the covariance of sampling errors. The analysis and determination of the final model for each of the 4 regions was undertaken by assessing the average variance of prediction (AVP) and the regression models generated were of the potential (quantile logarithm) type. The results obtained showed that the WLS method performed better than the GLS method in relation to the variance inflation ratio (VIR), as the VIR for the four homogeneous regions was less than 1. According to Griffis and Stedinger (2007), if the VIR is less than 1, then the WLS regression is sufficient, so there is no need to use the GLS model. In order to evaluate the criteria related to OLS method residuals, hypothesis tests for independence, normality, zero mean and constant variance at a significance level of 5% were applied. This analysis allowed it to be shown that, for the regions BA2 and BA4, the use of the OLS method is not recommended, since at least one

assumption regarding the residuals was violated. Additionally, when evaluating the error variance ratio (EVR), it was found that the sample error variance is greater than the model error variance, therefore the OLS should not be employed. It should be noted that the GLS method is more appropriate, as the low flows are strongly correlated, given that the WLS method does not take into account correlations among concurrent flows. Nevertheless, where correlations are small, such a method could be adopted. However, in the region studied, the values of the cross-correlations are high. Therefore, it is concluded that, for the study in question, the GLS method is most appropriate. It is recommended, however, that the covariance matrix of the sampling errors be evaluated, as the method for estimating the covariance matrix is based on the quantiles themselves, and not on the quantile logarithm. This non-linear transformation certainly causes a change in the elements of the covariance matrix.

# SUMÁRIO

<b>LISTA DE FIGURAS.....</b>	<b>XII</b>
<b>LISTA DE TABELAS.....</b>	<b>XIV</b>
<b>LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.....</b>	<b>XVI</b>
<b>1 INTRODUÇÃO .....</b>	<b>1</b>
<b>2 OBJETIVOS .....</b>	<b>4</b>
2.1 OBJETIVO GERAL .....	4
2.2 OBJETIVOS ESPECÍFICOS.....	4
<b>3 REVISÃO DE LITERATURA.....</b>	<b>5</b>
3.1 REGIONALIZAÇÃO HIDROLÓGICA .....	5
3.1.1 <i>Análise de frequência regional</i> .....	5
3.1.2 <i>Modelos de regionalização</i> .....	6
3.1.3 <i>Seleção e análise dos dados</i> .....	9
3.1.3.1 Determinação do período de dados.....	10
3.1.3.2 Preenchimento de falhas.....	10
3.1.3.3 Ajuste da distribuição de frequência.....	11
3.1.3.4 Definição das características físicas e climáticas.....	13
3.1.4 <i>Regiões homogêneas</i> .....	15
3.2 CARACTERIZAÇÃO GERAL DAS VAZÕES MÍNIMAS .....	16
3.2.1 <i>Características hidrogeológicas</i> .....	19
3.2.2 <i>Vazão média mínima com 7 dias de duração e 10 anos de período de retorno - <math>Q_{7,10}</math></i> .....	21
3.3 MÉTODOS DE REGRESSÃO OLS E WLS .....	22
3.3.1 <i>Método OLS</i> .....	22
3.3.2 <i>Método WLS</i> .....	25
3.4 MÉTODO GLS .....	28
3.4.1 <i>Método GLS Stedinger e Tasker</i> .....	29
3.4.1.1 Matriz de covariância dos erros amostrais.....	30
3.4.1.2 Correlação cruzada entre os fluxos concorrentes.....	32
3.4.1.3 Histórico de utilização do GLS em engenharia de recursos hídricos.....	34
3.4.2 <i>Considerações finais sobre os métodos de regressão</i> .....	37
3.5 CONSIDERAÇÕES FINAIS DA REVISÃO DE LITERATURA .....	38
<b>4 METODOLOGIA.....</b>	<b>39</b>
4.1 LOCALIZAÇÃO E CARACTERIZAÇÃO DA BACIA HIDROGRÁFICA DO RIO SÃO FRANCISCO .....	39
4.1.1 <i>Caracterização da bacia do Alto São Francisco</i> .....	39
4.1.2 <i>Localização e caracterização das sub-bacias 40 e 41</i> .....	39
4.1.2.1 Sub-bacia 40.....	40
4.1.2.2 Sub-bacia 41 .....	41
4.1.3 <i>Caracterização dos aquíferos da bacia Alto São Francisco</i> .....	42
4.2 SELEÇÃO DOS DADOS.....	48
4.2.1 <i>Análise de frequência local</i> .....	49
4.2.1.1 Distribuição de probabilidade Weibull 2 parâmetros.....	50
4.2.2 <i>Estações fluviométricas selecionadas</i> .....	52
4.2.3 <i>Levantamento das características físicas e climáticas</i> .....	54
4.3 DEFINIÇÃO DAS REGIÕES HOMOGÊNEAS .....	55
4.4 CONSTRUÇÃO DO MÉTODO GLS STEDINGER-TASKER .....	57
4.4.1 <i>Variância do erro amostral</i> .....	59
4.4.2 <i>Formulação da equação de covariância dos erros amostrais</i> .....	60
4.4.2.1 Derivadas dos momentos amostrais.....	61
4.4.2.2 Covariância entre os momentos amostrais.....	62
4.4.2.3 Equação de covariância dos erros amostrais.....	65

4.4.3	<i>Correlação cruzada entre os fluxos concorrentes</i> .....	66
4.5	APLICAÇÃO DOS MÉTODOS DE REGRESSÃO .....	67
4.5.1	<i>Elaboração do modelo final</i> .....	68
4.5.1.1	Aplicação do método OLS .....	70
4.5.1.2	Aplicação do método WLS.....	70
4.5.1.3	Aplicação do método GLS .....	71
4.5.1.4	Considerações finais para a elaboração do modelo final .....	71
4.6	PROCEDIMENTOS PARA AVALIAÇÃO DOS MÉTODOS DE REGRESSÃO .....	72
4.6.1	<i>Modelos de regressão</i> .....	72
4.6.2	<i>Resíduos da regressão</i> .....	74
4.6.2.1	Teste de aderência de Anderson-Darling .....	74
4.6.2.2	Teste paramétrico sobre a média .....	75
4.6.2.3	Teste de Durbin-Watson .....	75
4.6.2.4	Teste de Breusch-Pagan .....	76
4.6.2.5	Teste de White.....	77
4.6.3	<i>Análise comparativa entre os métodos OLS, WLS e GLS</i> .....	78
<b>5</b>	<b>RESULTADOS</b> .....	<b>80</b>
5.1	RESULTADOS DO MÉTODO OLS .....	80
5.1.2.1	Análise do teste de Anderson-Darling .....	83
5.1.2.2	Análise da média igual a zero dos resíduos .....	84
5.1.2.3	Análise do teste de Durbin-Watson .....	85
5.1.2.4	Análise dos testes de Breusch-Pagan e de White.....	86
5.1.2.5	Resumo da análise dos resíduos .....	87
5.2	RESULTADOS DOS MÉTODOS WLS E GLS .....	88
5.2.1	<i>Matriz de covariância dos erros amostrais e correlação cruzada</i> .....	88
5.2.2	<i>Resultados do método WLS</i> .....	89
5.2.2.1	Seleção dos modelos de regressão .....	89
5.2.2.2	Análise dos resíduos.....	92
5.2.3	<i>Resultados do método GLS</i> .....	92
5.2.3.1	Seleção dos modelos de regressão .....	92
5.2.4	<i>Análise dos resíduos</i> .....	94
5.3	ESTUDO COMPARATIVO ENTRE OS MÉTODOS OLS, WLS E GLS .....	95
5.3.1.1	Região BA1 .....	96
5.3.1.2	Região BA2 .....	97
5.3.1.3	Região BA3 .....	98
5.3.1.4	Região BA4 .....	99
<b>6</b>	<b>CONCLUSÕES E RECOMENDAÇÕES</b> .....	<b>105</b>
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>110</b>
	<b>APÊNDICE - A</b> .....	<b>119</b>
	<b>APÊNDICE - B</b> .....	<b>124</b>
	<b>APÊNDICE - C</b> .....	<b>129</b>
	<b>APÊNDICE - D</b> .....	<b>134</b>



## LISTA DE FIGURAS

<b>Figura 3.1:</b> Rios efluentes (3.1 (a)) e rios influentes (3.1(b)) de acordo com a posição do lençol freático .....	20
<b>Figura 4.1</b> - Localização da sub-bacia 40 no estado de Minas Gerais .....	40
<b>Figura 4.2</b> - Localização da sub-bacia 41 no estado de Minas Gerais .....	41
<b>Figura 4.3</b> – Mapa de sistemas aquíferos da bacia do Alto São Francisco em Minas Gerais, baseado no mapa de sistemas aquíferos (Mourão <i>et al.</i> , 2001) .....	47
<b>Figura 4.4</b> - Série temporal do ano de 1942 da estação 40025000.....	50
<b>Figura 4.5</b> - Ajuste da distribuição teórica (Weibull 2P) a distribuição empírica.....	50
<b>Figura 4.6</b> - Função densidade da Weibull (2P) mínimos .....	51
<b>Figura 4.7</b> - Delimitação das regiões homogêneas.....	57
<b>Figura 4.8</b> - Correlação cruzada das vazões mínimas anuais ( $Q_7$ ) versus pares de estações com no mínimo 5 anos de vazões concorrentes (2831 dados) .....	67
<b>Figura 4.9</b> - Etapas para determinação dos modelos de regressão .....	69
<b>Figura 4.10</b> - Estatística de teste Durbin-Watson.....	76
<b>Figura 5.1</b> – Resíduos dos modelos OLS .....	82
<b>Figura 5.2</b> – Mapa dos resíduos dos modelos OLS .....	83
<b>Figura 5.3</b> - Ajuste dos resíduos do OLS à distribuição normal .....	84
<b>Figura 5.4</b> - Intervalo de confiança para média dos resíduos igual a zero .....	85
<b>Figura 5.5</b> – Mapa dos resíduos dos modelos WLS .....	92
<b>Figura 5.6</b> – Mapa dos resíduos dos modelos do GLS .....	95
<b>Figura 5.7</b> - Vazões estimadas <i>versus</i> observadas região BA1, com barra da variância do erro amostral (em vermelho).....	97
<b>Figura 5.8</b> - Vazões estimadas <i>versus</i> observadas região BA2, com barra da variância do erro amostral (em vermelho).....	98
<b>Figura 5.9</b> - Vazões estimadas <i>versus</i> observadas região BA3, com barra da variância do erro amostral (em vermelho).....	99
<b>Figura 5.10</b> - Vazões estimadas <i>versus</i> observadas região BA4, com barra da variância do erro amostral (em vermelho) .....	100
<b>Figura 5.11</b> - Intervalos da variância do erro amostral.....	101

<b>Figura 5.12</b> - Variância do erro amostral de todas as estações da região BA1 .....	102
<b>Figura 5.13</b> - Variância do erro amostral de todas as estações da região BA2 .....	102
<b>Figura 5.14</b> - Variância do erro amostral de todas as estações da região BA3 .....	103
<b>Figura 5.15</b> - Variância do erro amostral de todas as estações da região BA4 .....	103

## LISTA DE TABELAS

<b>Tabela 3.1</b> - Classificação dos métodos para determinação de vazões mínimas ambientais ..	17
<b>Tabela 3.2</b> - Resumo de algumas medidas de vazões mínimas .....	18
<b>Tabela 4.1</b> – Unidades hidrogeológicas presentes no Sistema aquífero granular.....	42
<b>Tabela 4.2</b> – Unidades hidrogeológicas presentes no Sistema aquífero cárstico .....	44
<b>Tabela 4.3</b> – Unidades hidrogeológicas presentes no Sistema aquífero fissurado-cárstico ....	44
<b>Tabela 4.4</b> – Unidades hidrogeológicas presentes no Sistema aquífero fissurado .....	45
<b>Tabela 4.5</b> - Principais aspectos hidráulicos dos aquíferos (RURALMINAS, 1999 citado por Mourão <i>et al.</i> , 2001) .....	48
<b>Tabela 4.6</b> - Total de estações nas sub-bacias 40 e 41.....	48
<b>Tabela 4.7</b> - Estações fluviométricas e $Q_{7,10}$ das sub-bacias 40 e 41.....	52
<b>Tabela 4.8</b> - Número de estações por região e estatísticas do número de dados .....	56
<b>Tabela 5.1</b> - Modelos de regressão do método OLS região BA1 .....	81
<b>Tabela 5.2</b> - Modelos de regressão do método OLS região BA2 .....	81
<b>Tabela 5.3</b> - Modelos de regressão do método OLS região BA3 .....	81
<b>Tabela 5.4</b> - Modelos de regressão do método OLS região BA4 .....	82
<b>Tabela 5.5</b> - Análise da estatística de teste de Anderson-Darling .....	84
<b>Tabela 5.6</b> - Análise da estatística de teste (T) sobre média igual a zero .....	85
<b>Tabela 5.7</b> - Análise da estatística de teste de Durbin-Watson .....	86
<b>Tabela 5.8</b> - Regressão auxiliar teste de Breusch-Pagan .....	86
<b>Tabela 5.9</b> - Análise da estatística de teste de Breusch-Pagan .....	86
<b>Tabela 5.10</b> - Regressão auxiliar teste White .....	87
<b>Tabela 5.11</b> - Análise da estatística de teste de White.....	87
<b>Tabela 5.12</b> - Resumo dos testes de hipóteses aplicado aos resíduos do OLS .....	88
<b>Tabela 5.13</b> - Modelos de regressão do método WLS região BA1 .....	89
<b>Tabela 5.14</b> - Modelos de regressão do método WLS região BA2 .....	90
<b>Tabela 5.15</b> - Modelos de regressão do método WLS região BA3 .....	90

<b>Tabela 5.16</b> - Modelos de regressão do método WLS região BA4 .....	91
<b>Tabela 5.17</b> - Modelos de regressão do método GLS região BA1 .....	93
<b>Tabela 5.18</b> - Modelos de regressão do método GLS região BA2 .....	93
<b>Tabela 5.19</b> - Modelos de regressão do método GLS região BA3 .....	94
<b>Tabela 5.20</b> - Modelos de regressão do método GLS região BA4 .....	94
<b>Tabela 5.21</b> – Intervalo de validade das equações regionais .....	96
<b>Tabela 5.22</b> - Resultados dos modelos para região BA1 .....	96
<b>Tabela 5.23</b> - Resultados dos modelos para região BA2 .....	97
<b>Tabela 5.24</b> - Resultados dos modelos para região BA3 .....	98
<b>Tabela 5.25</b> - Resultados dos modelos para região BA4 .....	99
<b>Tabela 5.26</b> - Cálculo da EVR para os modelos WLS .....	100
<b>Tabela 5.27</b> - Cálculo da EVR para os modelos GLS .....	100
<b>Tabela 5.28</b> - Cálculo da VIR .....	104

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

$K$	Fator de frequência
$A$	Área de drenagem da bacia
$A^2$	Estatística de teste de Anderson-Darling
ANA	Agência Nacional de Águas
ASV	Variância média amostral
AVP	Variância média de predição
BGLS	GLS Bayesiano
$C$	Coefficiente de escoamento
$CE$	Capacidade específica ( $m^3/h.m$ )
CPRM	Companhia de Pesquisa de Recursos Minerais
CV	Coefficiente de variação amostral
$Dd$	Densidade de drenagem
$d_{ij}$	Distância Euclidiana entre pares de estações
$d_w$	Estatística de teste de Durbin-Watson
ERM	Erro relativo médio
EVR	Razão entre as variâncias do erro
$f_z(z)$	Função densidade da Weibull
$F_z(z)$	Função de probabilidade acumulada
$F_z(z)^{-1}$	Função inversa de $F_z(z)$
$g$	Coefficiente de assimetria amostral
GLS	Generalized Least Square
<b>I</b>	Matriz Identidade
$I$	Declividade equivalente
IACWD	<i>Interagency Committee on Water Data</i>
IGAM	Instituto Mineiro de Gestão das Águas

$L$	Comprimento do rio principal
LI	Limite inferior
LS	Limite superior
$m'_1$	Momento de primeira ordem em relação à origem
$m_2$	Momento central de segunda ordem
$m_3$	Momento central de terceira ordem
MSE	Erro quadrático médio
$N$	Número de dados
NERC	<i>National Environmental Research Council</i>
$N_i$	Número de dados da estação $i$
$N_{ij}$	Número de dados concorrentes
$N_j$	Número de dados da estação $j$
NRC	U. S. National Research Council
OLS	<i>Ordinary Least Square</i>
$P$	Precipitação média anual
PERH - MG	Plano Estadual de Recursos Hídricos de Minas Gerais
PRT	Técnica regressão dos parâmetros
$Q_{7,10}$	Vazão mínima de 7 dias de duração e dez anos de período de retorno
QRT	Técnica de regressão dos quantis
$R^2$	Coefficiente de determinação
SEMAD	Secretária de Estado de Meio Ambiente e Desenvolvimento Sustentável
SSE	Soma dos quadrados dos resíduos
SSR	Soma dos quadrados da regressão
SST	Soma quadrática total
$T$	Período de retorno
$\hat{u}^2$	Regressão auxiliar dos resíduos ao quadrado Teste de Breusch-Pagan e White

UPGRH	Unidades de Planejamento e Gestão de Recursos Hídricos
USGS	<i>U. S. Geological Survey</i>
USWRC	United States Water Resources Council
<b>V</b>	Matriz que contém pesos os elementos dentro e fora da diagonal principal
VIR	Razão de inflação de variâncias
<b>W</b>	Matriz contém pesos na diagonal principal
WLS	<i>Weighted Least Square</i>
WMO	World Meteorological Organization
<b>X</b>	Matriz que contém as variáveis explicativas
<b>y</b>	Matriz que contém os valores observados
$\alpha$	Parâmetro de escala da distribuição Weibull
$\beta$	Parâmetros a serem estimados pela regressão
$\gamma_1$	Coefficiente de assimetria populacional
$\gamma_2$	Coefficiente de Curtose populacional
$\gamma_3$	Coefficiente de quinta ordem populacional
$\gamma_4$	Coefficiente de sexta ordem populacional
$\delta^2 \varepsilon$	variância do erro do modelo
$\delta^2_{\omega}$	Variância do erro amostral
$\eta$	Erro total, correspondente a soma dos erros do modelo e dos erros amostrais
$\Lambda(\delta^2)$	Matriz de covariância dos erros
$\mu_{11}$	Covariância amostral
$\mu_{13}$	Momento bivariado de ordem um três
$\mu_2$	Momento de segunda ordem populacional
$\mu_{21}$	Momento bivariado de ordem dois um
$\mu_{22}$	Momento bivariado de segunda ordem
$\mu_{23}$	Momento bivariado de ordem dois três

$\mu_3$	Momento de terceira ordem populacional
$\mu_{33}$	Momento bivariado de terceira ordem
$\mu_4$	Momento de quarta ordem populacional
$\mu_5$	Momento de quinta ordem populacional
$\mu_6$	Momento de sexta ordem populacional
$\zeta$	Parâmetro de forma da distribuição Weibull
$\hat{\rho}_{ij}$	Estimador de correlação cruzada dos fluxos concorrentes
$\sigma_i$	Desvio padrão da estação i referente aos anos concorrentes
$\Sigma_{ij}$	Covariância dos erros amostrais
$\sigma_j$	Desvio padrão da estação j referente aos anos concorrentes
$\omega$	Erro dos estimadores amostrais



# 1 INTRODUÇÃO

De acordo com WMO (1974), as vazões mínimas são definidas como sendo “a vazão que escoar em uma determinada seção de um rio durante um período prolongado de seca”, ou seja, a ocorrência das vazões mínimas está associada a um fenômeno sazonal.

Segundo Tucci (2002), a vazão mínima diária é pouco útil, por não ser representativa do período de estiagem. Sendo assim, usualmente utilizam-se vazões médias mínimas com durações entre 7 e 30 dias, associadas a um determinado período de retorno. Por exemplo, a vazão média mínima com 7 dias de duração representa o menor valor da vazão média de 7 dias consecutivos em um ano civil.

A principal justificativa da utilização de vazões mínimas anuais, tal como a  $Q_{7,10}$ , reside no fato destas serem as principais balizadoras no que se refere às licenças ambientais e à avaliação de impacto ambiental de obras hidráulicas. De fato, no estado de Minas Gerais, a Portaria administrativa IGAM nº 010/98 de 30 de dezembro de 1998 em seu artigo 8º regulamenta como vazão de referência a  $Q_{7,10}$  (vazão média mínima de 7 dias de duração e dez anos de tempo de retorno). A Portaria fixa como limite máximo outorgável a vazão de 30% de  $Q_{7,10}$ , ficando garantido, a jusante de cada derivação, um fluxo residual equivalente a 70% de  $Q_{7,10}$ . Quando o curso d'água for regularizado por barramento, o limite poderá ser superior a 30% de  $Q_{7,10}$ , desde que se garanta um fluxo residual igual ou superior a 70% de  $Q_{7,10}$ .

Contudo, em uma resolução conjunta SEMAD-IGAM nº 1548, de 29 de março 2012, seguindo-se recomendações advindas - o Plano Estadual de Recursos Hídricos de Minas Gerais (PERH - MG), regulamentado pelo Decreto Estadual nº 45.565, de 22 de março de 2011- estipula que o máximo limite outorgável de 30% de  $Q_{7,10}$  passará a ser de 50% de  $Q_{7,10}$ , com fluxo residual respectivamente de 70% e 50% de  $Q_{7,10}$ . Contudo, nas bacias hidrográficas dos Rios Jequitaiá, Pacuí, Urucuia, Pandeiros, Verde Grande, Pará, Paraopeba e Velhas, o limite máximo outorgável continua a ser o convencionado anteriormente de 30% de  $Q_{7,10}$ .

Assim, torna-se evidente a importância de estudos no sentido de estimar tais vazões de referência, sobretudo em locais não monitorados.

De acordo com CPRM (2001), “uma rede hidrometeorológica, mesmo que densa, dificilmente atenderá com seus dados às necessidades de informação dos diversos setores que utilizam os recursos hídricos. Sempre haverá a necessidade de se determinar a disponibilidade hídrica

onde se originam as demandas, que muitas vezes são locais sem monitoramento ou com dados constituindo séries de curta duração ou com longos períodos apresentando falhas de observação. Para contornar esta situação, são utilizadas metodologias de regionalização que possibilitem a transferência de informações de uma região monitorada para outra que não possua coleta sistemática de dados, mas consideradas de comportamento hidrológico semelhante.”

Os modelos de regressão são subdivididos em três: Método dos Mínimos Quadrados Ordinários - OLS (*Ordinary Least Square*), Método dos Mínimos Quadrados Ponderados - WLS (*Weighted Least Square*) e Método dos Mínimos Quadrados Generalizados - GLS (*Generalized Least Square*). Neste trabalho será aplicado o método GLS sistematizado por Stedinger e Tasker (1985, 1986a, 1986b). Tal método tem sido utilizado em diversos estudos de regionalização (TASKER e STEDINGER, 1989; MADSEN e ROSBJERG, 1997; KROLL e STEDINGER, 1999; REIS Jr. *et al.*, 2004, 2005; GRIFFIS e STEDINGER, 2007) e os resultados encontrados foram melhores se comparados aos dois outros métodos de regressão. Outro método de regionalização bastante utilizado é o *index-flood* (HOSKING e WALLIS, 1997; MADSEN e ROSBJERG, 1997; CPRM, 2001).

O método GLS sistematizado por Stedinger e Tasker (1985, 1986a, 1986b) é um método de regressão que leva em conta a correlação cruzada entre os fluxos concorrentes e também considera a diferença de tamanho das séries hidrológicas. No método GLS, os erros são descritos como heterocedásticos. Neste trabalho será aplicado o método GLS na bacia do alto São Francisco, sub-bacias 40 e 41, em um estudo de regionalização de vazões mínimas  $Q_{7,10}$  cujo método de regionalização é de um evento com um determinado risco. O método GLS será comparado com os métodos de regressão OLS e WLS, a fim de que se possa avaliar o seu desempenho em relação a estes.

Conforme apontado anteriormente, vários estudos, principalmente para as vazões máximas, demonstraram que a regionalização pelo método GLS oferece resultados superiores àqueles fornecidos por outras metodologias, ou seja, estimativa mais acurada do erro do modelo e dos parâmetros do modelo. No entanto, a literatura envolvendo vazões mínimas ainda é incipiente. Assim, pretende-se avaliar se os resultados para vazões mínimas se assemelharão aos obtidos para vazões máximas e, dessa maneira, seja disponibilizada uma ferramenta para a comunidade científica e técnica no que se refere à estimativa de vazões mínimas em locais não monitorados.

Este trabalho será dividido em 7 capítulos incluindo a introdução. O capítulo 2 abrange os objetivos deste trabalho (geral e específicos). O capítulo 3 compreende a revisão de literatura, a qual possui uma descrição da análise de frequência regional, bem como a apresentação de uma revisão em relação aos métodos de regionalização e as principais etapas construtivas de uma regionalização tradicional. Neste capítulo, também é feita uma descrição das vazões mínimas e do seu emprego em estudos hidrológicos e, além disso, uma breve revisão das características hidrogeológicas e da  $Q_{7,10}$ . Apresenta-se também uma revisão referente aos métodos de regressão (OLS, WLS e GLS) e suas premissas. O capítulo 4 inclui as etapas metodológicas adotadas, como a escolha da área de estudo e sua divisão em regiões homogêneas, a definição das estações e a aplicação do método GLS. No capítulo 5 são apresentados os resultados dos modelos OLS, WLS e GLS para cada uma das 4 regiões. O capítulo 6 contém as conclusões e algumas recomendações com intuito de que sejam feitas melhorias metodológicas do ponto de vista científico. E, por fim, o capítulo 8 contém as referências bibliográficas utilizadas para o desenvolvimento deste trabalho.

## **2 OBJETIVOS**

### **2.1 Objetivo geral**

Avaliar um método de cálculo de vazões mínimas levando em conta a correlação cruzada entre os dados fluviométricos de diferentes postos de monitoramento e, também elaborar uma proposta para a estimativa da matriz de covariância dos erros amostrais.

### **2.2 Objetivos específicos**

- Aplicar o Método GLS em um estudo de regionalização, utilizando a análise de frequência local para cada uma das estações fluviométricas, de modo a estimar os quantis de vazões médias mínimas com duração de 7 (sete) dias associados a um período de retorno de 10 anos;
- Avaliar a dependência das vazões médias mínimas com as características fisiográficas da área em estudo;
- Construir a matriz de covariância dos erros amostrais para a distribuição de probabilidade Weibull (2P);
- Construir uma equação da correlação cruzada em função da distância entre as estações;
- Avaliar o desempenho do método dos Mínimos Quadrados Generalizados (GLS) frente ao OLS e o WLS;
- Avaliar a necessidade da divisão da área de estudo em regiões hidrologicamente homogêneas na aplicação do método GLS.

## 3 REVISÃO DE LITERATURA

### 3.1 Regionalização hidrológica

#### 3.1.1 Análise de frequência regional

As séries de vazões, sejam estas máximas ou mínimas, nem sempre atendem às necessidades de informações requeridas em um estudo hidrológico. Então, para resolver este problema, faz-se frequentemente o uso de equações empíricas ou a extrapolação da curva regional de frequências.

De acordo com Hosking e Wallis (1997), a análise de frequência é a estimativa de quantas vezes um acontecimento especificado ocorrerá. A estimativa de frequência de eventos extremos, frequentemente possui uma importância particular, pois existem inúmeras fontes de incertezas relacionadas aos processos físicos que dão origem aos eventos observados. Assim, uma abordagem estatística para a análise de dados permite que tais efeitos sejam quantificados.

Seja  $Q$  a magnitude de um evento extremo, tal como as vazões médias mínimas anuais. Rigorosamente, pela forçosa existência de componentes aleatórios, não existem relações funcionais e unívocas entre as variáveis características de processos hidrológicos, a vazão média mínima por exemplo. Assim, considera-se  $Q$  uma variável aleatória com distribuição de probabilidades acumuladas dada por:

$$F_x(x) = P(Q \leq x) \quad (3.1)$$

A inversa da função  $F_x(x)$  é denotada por  $F_x^{-1}(x)$ , e equivale ao quantil da distribuição de frequência, o qual, expressa a magnitude de um evento em termos da probabilidade de não excedência de  $F_x$ . O quantil de período de retorno  $T$ ,  $Q_T$ , é um evento extremo que tem probabilidade  $1/T$  de ser igualado ou superado em um ano qualquer.  $Q_T$  é dado por

$$Q_T = F_x^{-1}\left(x\right)\left(1 - \frac{1}{T}\right) \quad (3.2)$$

A regionalização de vazões é a inclusão, na análise de frequência, de dados de locais diferentes daquele onde se deseja obter estimativas de quantis. Conforme apontado por Hosking e Wallis (1997), uma vez que a análise regional incorpora mais dados que a análise

local, ela tem o potencial de fornecer estimativas mais acuradas para os quantis de vazão, mesmo em locais já monitorados

Segundo Tucci (2002), o vocábulo regionalização tem sido utilizado em hidrologia para designar a transferência de dados de uma região para outra.

Com o intuito de aumentar a segurança das estimativas dos parâmetros de uma distribuição de probabilidade, o *U. S. National Research Council – NRC* (1998) recomendou o preceito de substituir o “tempo pelo espaço”, por meio de modelos de regionalização, com a finalidade de melhorar as estimativas de eventos extremos.

Essa técnica de análise tem sido usada em hidrologia há muito tempo, sendo o método *index-flood*, apresentado por Dalrymple (1960), um dos exemplos mais antigos da sua aplicação em estimativa de quantis de vazão. Vários métodos de regionalização têm sido indicados por agências nacionais ligadas a recursos hídricos para uso corrente em hidrologia de enchentes. Hosking e Wallis (1997) citam, por exemplo, o Boletim 17 (USWRC, 1982), nos EUA, e o método do NERC (1975), no Reino Unido.

De acordo com CPRM (2001), análise de frequência regional não se restringe apenas à necessidade de transferência espacial de variáveis hidrológicas, mas também à obtenção de estimativas mais confiáveis de parâmetros e quantis, identificação de regiões com carência de postos de observação e verificação da consistência das séries hidrológicas.

Segundo a Eletrobrás (1985), o método de regionalização mais difundido é o método tradicional, o qual foi elaborado pelo National Environmental Research Center (NERC), a partir análise de métodos estatísticos aplicados à hidrologia (NERC, 1975) e que se baseia nas seguintes etapas:

- Seleção e análise dos dados básicos;
- Identificação de regiões hidrologicamente homogêneas;
- Ajuste do modelo de regressão.

### **3.1.2 Modelos de regionalização**

O estudo e avaliação da adequação de diversos modelos de regionalização, têm sido objeto de pesquisa de diversos pesquisadores.

Pinto (2006) avaliou os métodos de regionalização de vazões mínimas da curva de permanência de vazões para a bacia do rio das Velhas, por meio dos métodos Tucci (1991) (que é um método de regionalização da curva de permanência de longo termo) e o de Clapes e Fiorentino (1997). O autor constatou que ambos os métodos forneceram bons resultados e que o modelo estimado apresentou um bom desempenho.

Silva Júnior *et al.* (2003), aplicaram um estudo de extrapolação espacial na regionalização de vazões específicas médias de longo período, específicas mínimas, específicas máximas e para  $Q_{50}$  e  $Q_{95}$  nas bacias rio Paraopeba em Minas Gerais e na bacia do rio Ijuí no Rio Grande do Sul. Tais autores utilizaram bacias hidrográficas com áreas de drenagem grandes e pequenas, com o intuito de avaliar a extrapolação, e constataram que na extrapolação das vazões mínimas ocorreu uma tendência do modelo em superestimar a referida vazão. Os autores salientam que os modelos de regionalização não devem ser utilizados com finalidade de solucionar falta de dados em algumas regiões. Contudo, a sua essência tem como finalidade a transferência de informações no espaço.

Pinto e Alves (2003) empregaram o método de regionalização da curva adimensional de frequência (método *index-flood*) com uso dos momentos-L, para estimar as vazões máximas, mínimas com durações variadas, e médias de longo termo, para a bacia do alto São Francisco.

Azevedo (2004) realizou a regionalização de vazões mínimas de referência para a sub-bacia do rio Paranã (localizado entre Tocantins e Goiás), o qual avaliou três metodologias, a descrita pela Eletrobrás (1985), a de Chaves *et al.* (2002), e a baseada na interpolação linear de vazões mínimas de referência ( $Q_{7, 10}$ ,  $Q_{90}$  e  $Q_{95}$ ). O autor verificou que a melhor metodologia foi a tradicional, que utiliza como base as equações de regressão regionais.

Novaes *et al.* (2007) avaliaram o desempenho de cinco metodologias de regionalização. Os métodos avaliados foram: método tradicional, interpolação linear, método de Chaves *et al.* (2002), interpolação linear modificado e método de Chaves *et al.* (2002) modificado. As metodologias de regionalização foram aplicadas na estimativa de vazões mínimas e médias de longa duração, na bacia do rio Paracatu. Os autores constataram que todos os cinco métodos apresentaram resultados próximos, o que evidenciou não haver diferenças significativas do ponto de vista estatístico entre tais metodologias.

Chaves *et al.* (2002) propuseram um método para regionalização de vazões mínimas ( $Q_{90\%,7d}$ ), que leva em conta a interpolação e a extrapolação automática, utilizando o ambiente de

Sistema de Informações Geográficas (SIG's). Tal método foi empregado na bacia do rio Itapicuru, localizada no Estado da Bahia. A metodologia proposta por esses autores obtiveram erros relativos médios de 45,1% contra 289,7% do método tradicional (ajuste com o modelo de regressão OLS). Todavia, os autores salientam que o emprego de tal metodologia não pode ser generalizada, aconselhando que o método seja empregado em bacias com características similares a que utilizaram, além de não se aplicar em bacias sem dados fluviométricos (extrapolação).

Mamun *et al.* (2010) desenvolveram um estudo de regionalização para a Península da Malásia. Os autores avaliaram vazões com durações de 1, 3 e 30 dias, com períodos de retorno que variavam entre 1 e 50 anos. Este estudo foi desenvolvido com intuito de atualizar o relatório de Procedimento Hidrológico (HP No. 12) da Malásia de 1976.

Dawdy *et al.* (2012) relataram que apesar de diversas melhorias introduzidas na análise de frequência regional de vazões, elas ainda permanecem restritas, na medida em que não incluem nenhum conhecimento dos processos físicos que produzem as vazões.

De uma forma geral, conforme apontado por Naghettini e Pinto (2007), os métodos regionais podem ser divididos em três classes: métodos de regionalização dos parâmetros da distribuição de probabilidades, métodos de regionalização do evento com um determinado risco e métodos de regionalização da curva adimensional de frequências.

No primeiro caso, pressupõe-se que as observações em cada posto, sejam elas mínimas, médias ou máximas, tenham a mesma distribuição de probabilidades, e que os parâmetros desta distribuição sejam associados ou explicados pelas características fisiográficas do local. Tais características podem ser desde a área de drenagem às características geológicas da bacia. A relação entre os parâmetros e as características da bacia é, em geral, obtida por meio de uma regressão linear múltipla, supondo não haver correlação espacial entre os postos de observação.

O segundo caso consiste, de acordo com Naghettini e Pinto (2007), de uma análise de frequência local para cada amostra de observações hidrológicas, de modo a estimar os quantis da variável hidrológica associados a períodos de retorno previamente especificados. Em seguida, procura-se, por meio de regressão, estabelecer uma relação entre os quantis e as características fisiográficas da região estudada.



O terceiro método de regionalização é mais bem representado pelo denominado método *index-flood*. De acordo com Stedinger (2000), várias pesquisas têm demonstrado as vantagens dos métodos regionais com base no *index-flood* (LETTENMAIER *et al.*, 1987; STENDINGER e LU, 1995; HOSKING e WALLIS, 1997; MADSEN e ROSBJERG, 1997).

A ideia por trás do método *index-flood* é usar dados de bacias hidrologicamente similares para estimar uma distribuição adimensional para as cheias ou para a precipitação. Ou seja, “substituir o tempo pelo espaço” para compensar o fato de cada local possuir registros relativamente curtos.

O conceito do método reside na hipótese de que diferentes locais na mesma região possuem a mesma distribuição para as cheias, a menos de um parâmetro de escala, ou *index-flood*, que reflete a magnitude média dos eventos máximos de precipitação ou de escoamento em cada bacia.

### **3.1.3 Seleção e análise dos dados**

De acordo com ELETBRÁS (1985), os dados necessários para um estudo de regionalização podem ser classificados em:

- Dados descritivos: estão relacionados com a escolha da região de estudo, bem como sua localização geográfica e caracterização hidrológica, geológica, geomorfológica, hidrogeológica e climatológica;
- Dados físicos: Tais dados estão associados à coleta das informações a respeito das características físicas, as quais são realizadas a partir da análise de mapas com detalhamento, ou seja, em escalas que possibilitem a coleta dos dados. Nesta etapa também faz-se o diagnóstico da área de estudo no que se refere à disponibilidade de postos fluviométricos que contenham séries de dados consistentes.
- Dados hidrológicos: Esses dados são levantados na etapa anterior (dados físicos), e incluem as características físicas e climáticas que serão utilizadas como variáveis explicativas na regionalização de vazões, sejam estas mínimas, máximas ou médias.

De acordo com Gustard *et al.* (1992), a seleção dos dados tem que se basear em bacias que possuam registros de boa qualidade e que tenha o regime de escoamento o mais próximo do natural.

Ainda segundo Gustard *et al.* (1992), alguns eventos que ocorrem nas estações fluviométricas podem tornar os dados impróprios. Tais eventos podem incluir desde falhas intermitentes de medição como a manutenção temporária do sistema de recursos hídricos. Contudo, se tais eventos forem de curta duração, não há justificativa para a desclassificação de tal estação, mas simplesmente o ano em que ocorreu tal evento deverá ser retirado das análises dos dados. Estes também salientam que no máximo 24 meses individuais podem ser omitidos da análise das séries, sem que haja o comprometimento dos dados da estação em questão e que determine a sua desclassificação.

Ahearn (2010) desenvolveu um estudo de regionalização para o estado de Connecticut nos EUA. O estudo foi aplicado somente em cursos d'água que não foram impactados pela atividade antrópica (como retirada da água para abastecimento ou para indústria, na bacia de drenagem a montante). Esta medida foi adotada com o intuito de manter a homogeneidade dos dados.

#### 3.1.3.1 Determinação do período de dados

As séries de dados devem ser representativas do fenômeno em estudo. De acordo com Eletrobrás (1985), para a análise de períodos de estiagem, é necessária uma quantidade maior de dados do que para a análise de cheias. Isto decorre da forte correlação entre valores anuais de vazões mínimas, pois um período de estiagem é normalmente comum para diversas estações localizadas numa mesma região.

Contudo, a determinação do período de anos normalmente fica a critério do autor. Além disso, deve-se fazer uma análise das séries das estações, averiguando a quantidade média de anos disponíveis (ELETROBRÁS, 1985).

#### 3.1.3.2 Preenchimento de falhas

Os preenchimentos das falhas nas série hidrológicas, normalmente são realizados mediante o ajuste de um modelo de regressão linear, OLS. Os valores são estimados a partir da análise das séries históricas de postos vizinhos.

De acordo com Bertoni e Tucci (2007), em um estudo hidrológico deve-se efetuar previamente análises referentes ao preenchimento de falhas e também de consistência dos dados hidrológicos.

Tucci (2002) salienta que é necessário determinar um *período base*, que contenha postos com séries históricas completas e coeficientes de correlação aceitáveis.

Zeilhofer *et al.* (2003) destacam que para o preenchimento de falha de séries temporais de vazões ou precipitações, é importante que se faça a escolha de pares de pontos que tenham alta correlação para o emprego da regressão linear.

Oliveira *et al.* (2010) fizeram um estudo comparativo entre as metodologia de preenchimentos de falhas. Os métodos empregados foram: ponderação regional, regressão linear, ponderação regional com base em regressões lineares e vetor regional.

### 3.1.3.3 Ajuste da distribuição de frequência

De acordo com Tucci (2002), os métodos de ajuste da distribuição de frequência podem ser:

- Empíricos: é um método gráfico, onde plotam-se os valores de vazões (sejam elas médias, máximas ou mínimas) em função da duração e, em seguida, ajusta-se uma curva. a vazão desejada com determinado período de retorno é selecionada graficamente.
- Distribuição teórica: é um método no qual faz-se o ajuste de uma distribuição teórica de probabilidade.

De acordo com a Eletrobrás (1985), a prática tem demonstrado que vazões médias mínimas, com quaisquer durações, se ajustam satisfatoriamente à distribuição de probabilidade Weibull.

Kroll e Vogel (2002) utilizaram o diagrama de momentos-L para descrever a capacidade de diversas distribuições de probabilidade em representar séries de vazões mínimas, examinando 1505 estações fluviométricas nos Estados Unidos. Ao comparar locais de rios perenes, observaram uma mudança na proporção dos momentos-L e, em locais onde os rios eram intermitentes, os momentos-L foram relatados como sendo zero. Os resultados indicaram que a distribuições Pearson tipo 3 e a Log-Normal 3 parâmetros são recomendadas para descrever a distribuição estatística de vazões mínimas em rios intermitentes e perenes nos Estados Unidos.

Silva *et al.* (2006) estudaram as vazões mínimas de referência  $Q_{7,10}$  e  $Q_{90}$  na região do alto Rio Grande em Minas Gerais e aplicaram os modelos Log-Normal 3 parâmetros, Weibull e Gumbel.

Kaviski (1983) realizou um estudo de vazões de estiagens em pequenas bacias hidrográficas do Estado de Santa Catarina, e aplicou as distribuições de probabilidades Log-Normal 3 parâmetros, Gumbel e Weibull. A distribuição que melhor se ajustou a seus dados foi a Weibull.

Vogel e Kroll (1989) compararam os estudos de ajuste das distribuições de probabilidades, Log-Normal de 2 e 3 parâmetros, Weibull de 2 e 3 parâmetros e Log-Pearson tipo 3. Eles estimaram os parâmetros para a série anual de vazões mínimas com sete dias de duração e 10 anos de período de retorno. Os autores relataram que as distribuições com 3 ou mais parâmetros tendem a produzir curvas mais aderentes aos dados empíricos, pois permitem maior flexibilidade em termos dos parâmetros de escala, forma e posição da distribuição modelada. No entanto, esta flexibilidade adicional ocorre frequentemente às custas de uma perda de precisão associada aos quantis estimados.

Tasker (1987) utilizou um experimento de reamostragem dos estimadores de quantis de vazões mínimas e comparou o desempenho relativo de três diferentes distribuições, Weibull 3 parâmetros, Log-Pearson tipo 3, logBoughton e também através da transformação de variáveis Box-Cox. Analisando 20 rios na Virgínia, tal autor recomendou a Weibull 3 parâmetros e Log-Pearson tipo 3 para descrever a frequência das  $Q_7$  mínimas anuais.

Vogel e Wilson (1996) realizaram uma análise das distribuições de probabilidade GEV, Log-Normal 3 parâmetros e a Log-Pearson tipo 3, para vazões mínimas em escala nacional, na Nova Zelândia e nos EUA. A Log-Pearson tipo 3 foi considerada pelos autores como sendo a distribuição mais aderente.

Mamun *et al.* (2010), em um estudo de regionalização de vazões mínimas para Malásia, testaram os três tipos de distribuições Generalizada de Valores Extremos (GEV), a Log-Normal e a Log-Pearson tipo 3. A distribuição que melhor se ajustou foi a Generalizada de Valores Extremos III, com parâmetro de forma positivo, ou seja, a Weibull.

Nos estudos de regionalização de vazões utilizando o GLS, Stedinger e Tasker (1985, 1986b), dentre entre outros autores, utilizaram a distribuição de probabilidade Log-Pearson tipo 3. A utilização da Log-Pearson tipo 3 foi recomendada pelas agências federais dos EUA, através do boletim 17, atualmente boletim 17B (IACWD, 1982), intitulado como “Diretrizes para determinação da frequência de vazões de cheia”. Além disso, o terceiro parâmetro desta

distribuição oferece mais flexibilidade para a montagem da distribuição Log-Pearson tipo 3 em diferentes regiões dos Estados Unidos ( DAWDY *et al.*, 2012).

Kroll e Vogel (2002) salientaram que os poucos estudos existentes que avaliaram o ajuste de distribuições de probabilidade às séries de vazões mínimas, não chegaram a um consenso sobre qual distribuição seria a mais adequada.

#### 3.1.3.4 Definição das características físicas e climáticas

Nos estudos de regionalização, usualmente utiliza-se como variáveis explicativas: a área de drenagem, o comprimento do talvegue principal, a declividade equivalente, a precipitação média anual, o coeficiente de escoamento superficial e a densidade de drenagem (TASKER E STEDINGER, 1989; CPRM, 2001).

Silva *et al.* (2003) empregaram um estudo de regionalização de vazões mínimas, máximas e médias e da curva de permanência na bacia do rio Grande, e constataram que a área de drenagem foi a variável mais correlacionada com a vazão, ou seja, a variável que melhor explica seu comportamento.

Catalunha (2004) destaca em seu trabalho, que dentre as variáveis explicativas empregadas em estudos de regionalização, a área de drenagem é a mais utilizada, a qual possui forte correlação com as outras características da bacia.

Vorst e Bell (1981) evidenciaram que apenas as características área de drenagem, comprimento do talvegue principal e declividade equivalente, são significativas para explicar as vazões.

Wright e Ensminger (2004) desenvolveram um estudo de regionalização de vazões mínimas para o estado de Lousiana, nos EUA, e consideraram a área de drenagem, a declividade do canal principal e precipitação média anual, como sendo as características físicas e climáticas mais importantes.

Ries III (1994) utilizou como variáveis explicativas, a área de drenagem, o comprimento de todos os rios da bacia, a área do terreno que apresentavam características de granulação grossa oriundas de depósitos estratificados, a área total de todos os cursos d'água, as elevações máxima e mínima da bacia, as altitudes mínima e máxima oriundas dos depósitos estratificados e a declividade média da bacia.

Tomas e Benson (1970) analisaram diversas características físicas e climáticas, das quais podem ser citadas: área de drenagem, comprimento do rio principal, declividade do rio principal, armazenamento superficial, elevação, área arborizada, índices físicos do solo, largura da bacia, área de aluvião, proporção de meandro, proporção de bifurcação, largura do vale, precipitação média anual, precipitação média mensal, neve, intensidade de precipitação, evaporação média anual e temperatura.

Rameh *et al.* (2003) aplicaram o método de regionalização tradicional nas sub-bacias 46, 47, 48 e 49 do rio São Francisco e propuseram que, além da utilização das variáveis físicas e climáticas, normalmente utilizadas em estudo de regionalização, também houvesse adição de uma variável relacionada com a permeabilidade do solo. Tal variável foi coletada a partir de mapa temático desenvolvido pela Superintendência do Desenvolvimento do Nordeste (SUDENE). Os resultados encontrados, com o acréscimo da variável permeabilidade, possibilitaram uma melhora significativa no desempenho do modelo.

IGAM (2012) desenvolveu um estudo de regionalização de vazões para o aprimoramento do processo de outorga no Estado de Minas Gerais e a variável explicativa utilizada foi a área de drenagem. Foram descartadas as variáveis como o comprimento do rio principal e a densidade de drenagem, pois mostraram pouca interferência no processo de geração de vazões mínimas e médias.

Lemos (2006), em seu estudo de regionalização, utilizou a área de drenagem, a declividade média, o comprimento do rio principal e a densidade de drenagem. Tais características foram extraídas a partir de um modelo digital de elevação hidrologicamente consistente.

Baena (2002), em estudo de regionalização de vazões desenvolvido para bacia do rio Paraíba do Sul, constatou que as escalas cartográficas influenciam de forma expressiva na rede de drenagem, apresentando contudo, pouco efeito na área de drenagem. Assim, comprovou que a utilização da densidade de drenagem em modelos de regionalização pode aumentar as incertezas na estimativa das vazões.

Hiruma e Ponçano (1994) realizaram uma pesquisa, na qual tentaram correlacionar a densidade de drenagem com dados geológico-estruturais, geomorfológicos e pedológicos na bacia do alto rio Pardo, situada entre São Paulo e Minas Gerais. Esses autores compararam mapas temáticos com mapas de isolinhas de densidade de drenagem. Tais autores observaram que as principais anomalias de rarefação da densidade de drenagem foram encontradas em

solos sedimentares da Bacia do rio Paraná, a qual é constituída pela Formação Aquidauana (Subgrupo Itararé). Os valores anômalos superiores de densidade de drenagem foram (densidade de drenagem –  $Dd > 2,50$ ), o valores intermediários ( $2,50 > Dd > 1,25$ ) e valores anômalos inferiores ( $Dd < 1,25$ ). Os autores concluíram que os valores anômalos inferiores de  $Dd$  concentram-se nos latossolos da Depressão Periférica, das quais o substrato são os terrenos sedimentares da Formação Aquidauana e os valores superiores de  $Dd$  estão associados à distribuição dos principais *trends* de estruturas reativadas durante o Mesozóico-Cenozóico.

Diversos estudos de regionalização de vazões mínimas desenvolvidos para diferentes regiões do mundo, utilizaram características geomorfológicas, geológicas, climáticas e topográficas (Riggs, 1972; Vogel e Kroll, 1992; Dingman e Lawlor, 1995).

### **3.1.4 Regiões homogêneas**

As regiões homogêneas são regiões que apresentam características e comportamentos hidrológicos semelhantes. De acordo com Hosking e Wallis (1997), usualmente a determinação das regiões homogêneas é muito subjetiva. Contudo as regiões homogêneas são definidas, a partir da similaridade entre os locais.

Ainda conforme Hosking e Wallis (1997), os métodos para a definição de regiões homogêneas podem ser agrupadas em quatro categorias principais, as quais são:

- **Conveniência Geográfica:** são frequentemente selecionadas com base em áreas administrativas, com intuito de formarem um conjunto de locais contíguos.
- **Agrupamento subjetivo:** normalmente é empregado em estudos de pequenas escalas. A delimitação da região homogênea é feita de forma subjetiva, com base na similaridade de características do local, tais como, relevo, classificação climática dentre outras.
- **Agrupamento objetivo:** agrupa-se os postos em um ou mais conjuntos, de modo que uma dada estatística não supere um valor limite previamente escolhido. Esses limites são escolhidos de modo que minimize a heterogeneidade dentro do grupo.
- **Análise de clusters:** é um método de análise estatística multivariada, no qual um vetor de dados é associado a cada local; os locais podem ser divididos ou reunidos em grupos de acordo com a similaridade se seus vetores de dados.

Hosking e Wallis (1997), também sugerem que sejam utilizadas estatísticas locais, a fim que se possa confirmar a delimitação física inicialmente realizada.

Cunanne (1998) descreveu as principais características de doze métodos de regionalização de vazões. E constatou que, a grande maioria destes métodos requeria dividir a região de estudo em regiões hidrologicamente homogêneas.

Tucci *et al.* (1995) desenvolveram um estudo de regionalização de vazões máximas, mínimas e médias para a bacia do rio Paraguai. Os autores testaram vários critérios para dividir a área de estudo em regiões homogêneas, tais como, tamanho da bacia, talvegue principal e critérios estatísticos, como os que foram propostos por Dalrymple (1960).

Tucci (2001) salienta que a subdivisão da área de estudo pode ser feita mediante a adoção de alguns critérios, tal como, tamanho das bacias hidrográficas (agrupar bacias pequenas, médias e grandes). Tal subdivisão também pode ser feita através da análise dos rios principais e por meio da geografia, ou seja, por meio da avaliação da similaridade das características geográficas da região estudada.

Acreman e Sinclair (1986) e Midgley *et al.* (1994) delimitaram as regiões homogêneas, a partir das características fisiográficas e dos parâmetros climáticos, obtidos a partir de mapas e de dados hidrometeorológicos (precipitação e evaporação).

Rao e Srinivas (2006), Nathan e McMahon (1990) e Souza e Limeira (2006) desenvolveram estudos voltados à melhoria dos critérios que envolvem a delimitação das regiões homogêneas, a partir da análise multivariada das características espaciais e temporais.

### **3.2 Caracterização geral das vazões mínimas**

As vazões mínimas equivalem aos menores valores médios de certa duração registrados a cada ano civil em uma dada seção de um rio, elas são utilizadas na avaliação da disponibilidade hídrica, na elaboração de projetos de irrigação e de energia elétrica, na concessão de outorga e em estudos do habitat aquático e de qualidade de água. De acordo com WMO (1974), as vazões mínimas são definidas como sendo “a vazão que escoar em uma determinada seção de um rio durante um período prolongado de seca”, ou seja, a ocorrência das vazões mínimas está associada a um fenômeno sazonal.



Smakhtin (1997) realizou um estudo de investigação das possibilidades de regionalização de vazões mínimas por meio da análise da curva de permanência para 22 regiões da África do Sul.

Pyrce (2004) apresentou de forma detalhada os índices de vazões mínimas. Este constatou a existência de diversos índices específicos que atendem às necessidades do estudo em recursos hídricos, engenharia e ciências biológicas. Os métodos para estabelecimento de vazões mínimas ambientais podem ser classificados de acordo com a Tabela 3.1.

**Tabela 3.1 - Classificação dos métodos para determinação de vazões mínimas ambientais**

<b>Método</b>	<b>Descrição</b>
1. Hidrológico	<p>-Recomendações de vazão ambiental são feitas usando métodos simples, utilizando principalmente dados hidrológicos (registros de vazões diárias ou mensais).</p> <p>-Considerado adequado ao nível de planejamento para desenvolvimento dos recursos hídricos.</p> <p>- Como exemplo, pode-se citar os métodos onde há inclusão de vários percentuais de excedência de duração de vazões (como <math>Q_{95}</math>, <math>Q_{75}</math>), ou como um único índice de vazões mínimas (como <math>Q_{7,10}</math>).</p>
2. Hidráulicos	<p>-Utiliza mudança nas variáveis hidráulicas (tal como perímetro molhado) para avaliar fatores de habitat conhecidos ou assumidos como limitante em relação à biota. Assim, um valor limite do parâmetro hidráulico selecionado irá sustentar a integridade do ecossistema/biota.</p>
3. Classificação do habitat	<p>-Este método tenta avaliar os requisitos de vazão ecológica com base em análise detalhada da adequação física e ecológica, a partir de descargas de vazões e dados de respostas hidrológicas, hidráulicas e biológicas.</p> <p>-A vazão é tipicamente modelada usando dados de profundidade, declividade, forma da seção transversal, etc., coletados em várias seção transversais na área de estudo.</p>
4. Holístico	<p>-Todas as necessidades do ecossistema são integradas e consideradas (incluindo o canal do rio, zonas ribeirinhas, planícies de inundação, etc).</p> <p>O regime natural do rio é um guia fundamental e deve ser incorporado dentro dos regimes de mudanças de vazões</p> <p>-A base da maioria das aproximações é obtida considerando uma construção sistemática de regime de vazões, modificados mês a mês e de elemento a elemento, na qual irá definir as características do regime de vazão</p>

Continuação da Tabela 3.1

<b>Método</b>	<b>Descrição</b>
4. Holístico	para alcançar os objetivos ecológicos, geomorfológicos, de qualidade da água, social ou outros.

**Fonte:** Pyrcce (2004)

Reis (2007) elaborou um trabalho onde avaliou métodos para determinação de vazões ecológicas, cujo enfoque foi realizar um estudo comparativo entre métodos que fornecessem maior representatividade ecológica, para que os habitats pudessem ser classificados. Os modelos WAIORA e RHYHABSIM foram empregados em sua pesquisa.

Silva *et al.* (2006) realizaram a regionalização hidrológica na região do Alto Rio Grande em Minas Gerais, com base em 7 estações fluviométricas. Além disso, avaliaram previamente a disponibilidade hídrica para a agricultura irrigada, adotando as vazões de referência  $Q_{7,10}$  e  $Q_{90}$ .

Em uma série de relatórios produzidos pelo USGS foram avaliadas as vazões médias mínimas com durações variando de 1 a 183 dias e para períodos de retorno variando de 2 a 50 anos (Armentrout e Wilson, 1987; Zalants, 1982; Giese e Madson, 1993). Este estudos foram desenvolvidos principalmente para avaliação da qualidade das águas dos rios.

Smakhtin (2001) analisou a situação atual das vazões mínimas e a sua variabilidade, bem como os métodos existentes para estimativa das vazões em séries temporais, que incluem curva de permanência, análise de frequência e caracterização do escoamento de base. O autor, contudo, se concentrou em analisar as técnicas de estimativa de vazões mínimas para locais sem monitoramento. Na Tabela 3.2, são apresentadas algumas formas de se obter os valores das vazões mínimas.

**Tabela 3.2 - Resumo de algumas medidas de vazões mínimas**

<b>Medida de vazões mínimas</b>	<b>Descrição das propriedades</b>	<b>Emprego de dados</b>	<b>Fontes</b>
Curva de permanência	Proporção de tempo na qual uma dada vazão é excedida, em dado horizonte temporal.	Vazão diária ou vazão média de vários dias, semanas ou meses	Hugges <i>et al.</i> , 1997 Gordon <i>et al.</i> , 1992 Pitman (1993)
Curva de frequência (mínima anual)	Proporção de anos na qual a vazão média ao longo de um determinado período é inferior a uma determinada magnitude.	Vazão mínima anual - diária ou média sobre vários dias	Zalants, 1982 Midgley <i>et al.</i> , 1994

Continuação da Tabela 3.2

Medida de vazões mínimas	Descrição das propriedades	Emprego de dados	Fontes
Eventos contínuos de baixa vazão e de volume deficitário	Frequência com a qual a vazão permanece, continuamente, abaixo de um limite, para uma dada duração.	Períodos de vazões mínimas extraídos, a partir do hidrograma, seguido de um tratamento estatístico das durações.	Yevjevich, 1967 Çigizoğlu e Bayazit, 1998
Constante de recessão	Índice de decaimento do hidrograma	Fluxo diário durante os períodos secos.	Sugiyama, 1996 Clause, 1992

**Adaptado** de Gustard *et al.* (1992)

De acordo com Singh (1997), as vazões que ocorrem em um período de seca nos rios dependem do escoamento de base, e este depende da quantidade de água precipitada, das condições geológicas e das condições hidrogeológicas, sendo estas as responsáveis para que escoamento de base flua para os rios.

### 3.2.1 Características hidrogeológicas

As vazões mínimas mantêm seu volume no período de estiagem, devido principalmente a recarga dos aquíferos. A capacidade do aquífero em armazenar e transmitir água varia para cada região pois, além dos fatores geológicos.

Uma forma de aferir a produtividade de diferentes aquíferos pode ser feita por meio da capacidade específica - CE ( $m^3/h.m$ ). Tal parâmetro baseia-se em concatenar a vazão de um poço ( $m^3/h$ ) e o decréscimo do nível d'água em metros. Os valores de CE variam, para um aquífero com boa produtividade, cerca de  $3m^3/h.m$ ; e para aquíferos com baixa produtividade o CE médio é inferior ou igual a  $0,5m^3/h.m$  (PINTO, 2006).

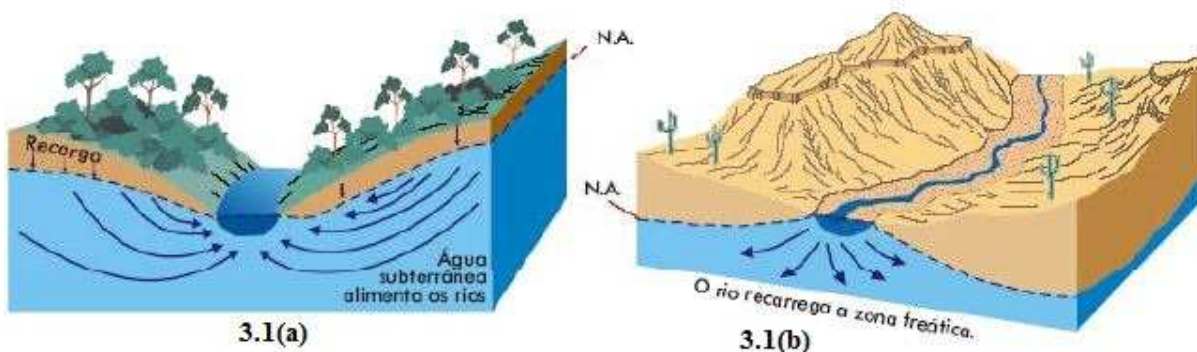
Os aquíferos possuem algumas propriedades, tais como: porosidade, permeabilidade, condições de recarga e capacidade de armazenamento. Estes também podem ser classificados quanto às características dos maciços porosos (CABRAL *et al.*, 2001), sendo dispostos em três sistemas:

- Granulares: Apresentam espaços vazios (são muito porosos) por onde a água circula. Estão associados a rochas do tipo sedimentares e solos arenosos;

- Fraturados e/ou fissurados: Possuem como características a presença de fraturas, onde normalmente a água se acumula. Estão relacionados com rochas do tipo metamórficas ou ígneas;
- Cársticos: Estes apresentam fraturas, as quais podem atingir grandes dimensões, em consequência da dissolução do carbono pela água. São formados por em rochas carbonáticas.

Segundo Press *et al.* (2006), um poço ou reservatório subterrâneo que possui grande disponibilidade hídrica, apresenta altas porosidade e permeabilidade, sendo que uma rocha com alta porosidade e baixa permeabilidade tende a dificultar que água flua.

Os aquíferos ainda podem ser classificados quanto a sua contribuição, sendo que nesta classificação se enquadram os rios efluentes e influentes. Segundo Teixeira *et al.* (2001) os rios efluentes (Figura 3.1 (a)) são aqueles que recebem contribuição contínua pelas águas subterrâneas. São típicos em regiões úmidas. Contrariamente, os rios influentes (Figura 3.1 (b)) são aqueles onde ocorre a diminuição da vazão a jusante, devido à recarga dos aquíferos pelo escoamento superficial. Tais rios são típicos de regiões secas. Neste último caso é comum o rio secar caso o nível freático se rebaixe.



**Figura 3.1:** Rios efluentes (3.1 (a)) e rios influentes (3.1(b)) de acordo com a posição do lençol freático  
 Fonte: Teixeira *et al.* (2001)

As condições de recarga estão associadas à movimentação e circulação de água e armazenamento dentro dos aquíferos. A velocidade na qual a água se desloca no solo compromete o balanço entre a descarga e recarga. Grande parte da água subterrânea se move lentamente, todavia a movimentação do fluxo de água é variável em uma mesma região (PRESS *et al.*, 2006). Sendo assim o engenheiro francês, Darcy, em 1850, propôs um método para poder explicar como os fatores geométricos interferem na vazão. Esta ficou conhecida como a lei de Darcy.

O CPRM disponibiliza alguns mapas hidrogeológicos regionais. Também tem desenvolvido diversos estudos, com intuito de monitorar as águas subterrâneas, principalmente nas regiões do semi-árido brasileiro (CPRM, 1998, 2008).

### **3.2.2 Vazão média mínima com 7 dias de duração e 10 anos de período de retorno - $Q_{7,10}$**

Segundo Tucci (2002), a vazão mínima diária é pouco útil, pois esta não é representativa do período de estiagem. Sendo assim, usualmente utiliza-se vazões médias mínimas com durações médias entre 7 e 30 dias, associadas a um determinado período de retorno. Por exemplo, a vazão média mínima com 7 dias de duração representa o menor valor da vazão média de 7 dias consecutivos em um ano civil.

Nos estados brasileiros de Minas Gerais São Paulo, Rio de Janeiro, Mato Grosso e Rondônia a  $Q_{7,10}$  (ANA, 2005) é empregada como vazão de referência nos critérios de outorga. Recentemente em Minas Gerais, o Instituto Mineiro de Gestão das Águas publicou um estudo de regionalização de vazões mínimas  $Q_{7,10}$ , o qual será utilizado nos procedimentos de concessão de outorga (IGAM, 2012).

De acordo com Harris *et al.* (2000), o uso da vazão mínima de referência é um mecanismo adequado para a proteção dos rios, pois as disposições para as derivações, normalmente são feitas a partir de uma vazão de base de pequeno risco.

Singb *et al.* (1988), determinaram as vazões mínimas ( $Q_{7,10}$ ) em rios do estado de Illinois - EUA. Os valores da  $Q_{7,10}$  foram estimados para diferentes pontos ao longo dos rios. Este estudo teve o objetivo de fornecer diretrizes para o monitoramento dos padrões de qualidade d'água no que se refere ao despejo de efluentes.

Chiang e Johnson (1976) realizaram um estudo para avaliar a qualidade da água em alguns rios na Pensilvânia. Relataram que a qualidade da água do rio era aceitável caso o despejo de efluentes fosse inferior à  $Q_{7,10}$ , sendo que qualquer desvio além deste valor poderia comprometer a qualidade da água do rio.

Flynn (2003) descreveu os resultados do estudo de regionalização de vazões mínimas ( $Q_{7,10}$ ) em New Hampshire. Os modelos de regressão foram ajustados utilizando o aplicativo GLSNET, desenvolvido por Stedinger e Tasker. Este avaliou dados de estações com 10 ou mais anos de registros de vazões, e também analisou o ganho do modelo com a inclusão de

novas estações. Os resultados comprovaram que em geral a inclusão de novas estações proporcionaram maior precisão na informação da vazão mínima regional.

Pyrce (2004) apresentou um resumo dos muitos usos da  $Q_{7,10}$ , os quais são:

- Determinação de vazão mínima/ecológica;
- Regulamentação de qualidade de descargas de águas residuárias ou alocações de resíduos (para prevenir impactos biológicos e ecológicos);
- Alocação da vida aquática, em projetos que envolvem cálculo de fluxo permanente;
- Como critério de manutenção ou proteção do habitat aquático; e
- Como um possível indicador de mortalidade potencial da vida aquática.

Pereira *et al.* (2003) realizaram um estudo da distribuição espacial das variáveis hidrológicas para a bacia do rio São Francisco. Eles concluíram que a  $Q_7$ , em geral, é crescente ao longo do rio São Francisco, atingindo valores inferiores a 100 m<sup>3</sup>/s até a divisa das sub-bacias 40 e 41. Entre as sub-bacias 43 e 47 os valores das  $Q_7$  são sempre superiores a 500 m<sup>3</sup>/s e, a partir da sub-bacia 47, estas são superiores a 1000m<sup>3</sup>/s.

### **3.3 Métodos de regressão OLS e WLS**

Os métodos de regressão podem ser subdivididos em três: Método dos Mínimos Quadrados Ordinário - OLS (*Ordinary Least Square*), Método dos Mínimos Quadrados Ponderados - WLS (*Weighted Least Square*) e Método dos Mínimos Quadrados Generalizados - GLS (*Generalized Least Square*), o qual será abordado no tópico 3.4.

#### **3.3.1 Método OLS**

O Método dos Mínimos Quadrados Ordinário, também conhecido como método dos mínimos quadrados convencional, tem diversas limitações. De acordo com Stedinger e Tasker (1985) o método assume que o resíduo associado às observações individuais são homocedásticos e independentes.

A hipótese de homocedasticidade estabelece que os erros ( $\epsilon$ ) (Equação 3.3) são independentes e normalmente distribuídos, com média zero e variância ( $\delta^2$ ) constante. Tal hipótese requer que

- O valor esperado do erro seja zero,  $E(\epsilon) = 0$ ;

- A  $\text{var}(\boldsymbol{\varepsilon}_i) = \text{var}(\boldsymbol{\varepsilon}_j)$  para todo  $i \neq j$ ;
- A covariância entre os erros residuais seja zero,  $\text{Cov}(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j) = 0$ , por conseguinte, os resíduos possuem correlação serial igual a zero.

$$\boldsymbol{\varepsilon} = \hat{\mathbf{y}} - \mathbf{X}\boldsymbol{\beta} \quad (3.3)$$

O método OLS assume que a variável hidrológica do vetor  $\hat{\mathbf{y}}$  pode ser descrita por uma função linear das características fisiográficas das bacias com uma componente de erro. Considere uma região com  $N$  postos de observação, sendo que cada posto possui  $P$  características fisiográficas (área de drenagem, declividade, precipitação média anual, etc.). O modelo de regressão na sua forma mais genérica para a variável  $\hat{\mathbf{y}}$ , é apresentado na Equação 3.4.

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3.4)$$

Onde  $\hat{\mathbf{y}}$  é um vetor que contém as estimativas do valor verdadeiro de  $\mathbf{y}$  em cada estação,  $\mathbf{X}$  é uma matriz ( $N \times P+1$ ) contendo as  $P$  variáveis explicativas,  $\boldsymbol{\beta}$  é um vetor a ser estimado com os  $P+1$  parâmetros do modelo e  $\boldsymbol{\varepsilon}$  é um vetor contendo os erros para cada um das  $N$  estações fluviométricas. Então, de acordo com o que é estabelecido na hipótese de homocedasticidade, a covariância entre os erros é igual a zero e a média também. Sendo assim, o estimador de  $\boldsymbol{\beta}$  é dado por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{y}} \quad (3.5)$$

De acordo com Stedinger e Tasker (1986b), o modelo básico do OLS assume que  $E(\hat{\mathbf{y}}) = \mathbf{X}\boldsymbol{\beta}$ , e tal hipótese é essencialmente verdadeira, e os resíduos são independentes e homocedásticos, de modo que

$$E[(\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})(\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})^T] = \delta_\varepsilon^2 \mathbf{I} \quad (3.6)$$

onde  $\mathbf{I}$  é a matriz identidade ( $N \times N$ ), e  $\delta_\varepsilon^2$  é a variância do erro do modelo.

A covariância para o estimador  $\hat{\boldsymbol{\beta}}$  é dada pela Equação (3.7) (Stedinger e Tasker, 1986b)

$$\Sigma(\hat{\boldsymbol{\beta}}) = \delta_\varepsilon^2 (\mathbf{X}^T \mathbf{X})^{-1} \quad (3.7)$$

O melhor estimador não enviesado de  $\delta_\varepsilon^2$  (Stedinger e Tasker, 1986b) é dado por

$$\delta_{\varepsilon}^2 = \frac{(\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})}{N - P - 1} \quad (3.8)$$

na qual  $N$  é igual ao número estações e  $P$  o número de variáveis explicativas introduzidas no modelo de regressão.

Stedinger e Tasker (1985) estudaram o estimador  $\delta_{\varepsilon}^2$ . Os resultados mostraram que a Equação (3.8) forneceu uma estimativa da variância altamente enviesada, principalmente quando a correlação cruzada entre dois fluxos quaisquer foi alta, ou quando o erro quadrático médio foi pequeno.

Kroll e Stedinger (1998) realizaram um estudo, onde compararam a eficiência dos estimadores  $\hat{\boldsymbol{\beta}}$  para os métodos OLS e GLS, descrito a frente. Eles concluíram que, para grandes valores da variância do erro do modelo, os estimadores  $\hat{\boldsymbol{\beta}}$ , por ambos os métodos, apresentaram eficiências muito próximas. Contudo, para pequenos valores da variância do erro do modelo, a eficiência do método OLS cai consideravelmente e, para os valores moderados da variância do erro do modelo, a eficiência relativa do estimador OLS diminui conforme as correlações cruzadas aumentam. Em geral, para pequenos períodos de retorno e para valores moderados a altos da variância do erro do modelo, o método OLS foi quase tão bom quanto o GLS, especialmente quando a correlação cruzada entre os fluxos concorrentes foi pequena. Então, para situações como a exposta neste último caso, os autores recomendam a utilização do estimador OLS, pois sua aplicação é mais simples que a do GLS.

Segundo Reis Jr. *et al.* (2005), a hipótese de homocedasticidade é frequentemente violada em problemas hidrológicos, uma vez que a precisão das estatísticas de interesse são diferentes em cada local, já que o tamanho das séries hidrológicas também são diferentes. Além disso, dificilmente as estatísticas de interesse são independentes, já que os dados de vazão possuem, em geral, uma forte correlação espacial. Por outro lado, a correlação serial local é pequena, uma vez que as amostras são formadas por um único valor extremo anual.

Taylor *et al.* (2011) realizaram a análise regional de frequência de vazões para o oeste da Austrália, onde aplicaram o método de regressão OLS para estimativa do quantil de vazões e para estimativa dos parâmetros da distribuição de probabilidade Log-Pearson tipo 3, determinando os quantis de vazões com períodos de retorno variando de 2, 5, 10, 20, 50 e 100. O estudo foi aplicado em 206 bacias hidrográficas. Tais autores, relataram que o resultado de todos os modelos foram bons, com base em testes estatísticos e análise do  $R^2$



ajustado. Contudo, salientam que os resultados encontrados por meio da regressão GLS, no estudo de Haddad *et al.* (2011), são provavelmente mais confiáveis.

Kroll e Stedinger (1999) avaliaram o desempenho de técnicas para estimar os parâmetros de um modelo de regressão regional para a  $Q_{7,10}$ , quando algumas estimativas nos locais são relatadas como sendo zero, ou seja, as técnicas empregadas tinham como objetivo avaliar os dados censurados dentro das séries de vazões. Estes utilizaram a simulação de Monte Carlo para o método OLS. Inicialmente o empregaram utilizando uma pequena constante em todos os quantis do local da estimativa (denotado por OLSC), ou negligenciando todas as observações relatadas como zero (denotado como OLSD). Eles compararam estas duas metodologias com o modelo Tobit, que é um procedimento para determinação da probabilidade máxima do estimador. Os modelos OLSC e OLSD apresentaram desempenho fraco em comparação com o modelo Tobit. Contudo, quando se tem uma pequena quantidade de dados censurados, o método OLSD pode ser aceitável.

### 3.3.2 Método WLS

O método de regressão WLS, assim como o método OLS é descrito pela Equação (3.3). No método WLS, entretanto, os resíduos são heterocedásticos. Assumindo-se que os erros ( $\epsilon$ ) possuem média igual a zero, a matriz de covariância destes erros é dada pela Equação 3.9 (WEISBERG, 2005)

$$E(\epsilon\epsilon^T) = \delta_\epsilon^2 \mathbf{W}^{-1} \quad (3.9)$$

onde  $\mathbf{W}$  é uma matriz ( $N \times N$ ) com a diagonal principal  $w_i$  (pesos), sendo que  $\delta_\epsilon^2$  é igual à variância do erro do modelo. Se  $\mathbf{W}$  for igual à matriz identidade, os erros são homocedásticos e o modelo em (3.5) se reduz ao Método OLS. No caso mais geral,  $\mathbf{W}$  é definida de forma a representar a heterocedasticidade e a correlação entre os resíduos, tal como no método WLS e GLS.

Diferenciando a Equação 3.10 (vetor  $\mathbf{S}(\hat{\beta})$ ) em relação a  $\hat{\beta}$  e igualando a derivada parcial a zero, obtém-se a Equação (3.12). A Equação 3.10 pode ser reescrita de uma forma mais geral, assim como é apresentada na Equação (3.11).

$$\mathbf{S}(\hat{\beta}) = (\mathbf{y} - \mathbf{X}\hat{\beta})^T \mathbf{W}^{-1} (\mathbf{y} - \mathbf{X}\hat{\beta}) \quad (3.10)$$

$$\mathbf{S}(\hat{\boldsymbol{\beta}}) = \sum_{i=1}^N w_i (y_i - \mathbf{X}^T \boldsymbol{\beta})^2 \quad (3.11)$$

onde  $\mathbf{S}(\hat{\boldsymbol{\beta}})$  representa a somatória dos erros quadráticos.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{W}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{-1} \hat{\mathbf{y}} \quad (3.12)$$

A Equação 3.12 refere-se ao estimador dos parâmetros  $\boldsymbol{\beta}$  para o método WLS (WEISBERG, 2005).

Weisberg (2005) esboça a manipulação algébrica apresentada em (3.13), a qual é utilizada na definição do modelo heterocedástico para a regressão WLS, resultando assim na Equação (3.14). Tal autor faz as seguintes considerações: sendo  $\mathbf{W}^{1/2}$  uma matriz  $N \times N$ , o  $i$ -ésimo elemento da diagonal principal é dado por  $\sqrt{w_i}$ . Analogamente, para  $\mathbf{W}^{-1/2}$ , o  $i$ -ésimo elemento da diagonal principal é dado por  $1/\sqrt{w_i}$ . Sendo assim  $\mathbf{W}^{1/2} \mathbf{W}^{-1/2} = \mathbf{I}$ . Utilizando regras da aritmética matricial, a matriz de covariância  $\mathbf{E}(\mathbf{W}^{1/2} (\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T))$  será dada por

$$\begin{aligned} \mathbf{E}(\mathbf{W}^{1/2} (\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T)) &= \mathbf{W}^{1/2} \mathbf{E}(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T) \mathbf{W}^{1/2} \\ &= \mathbf{W}^{1/2} (\delta^2 \mathbf{W}^{-1}) \mathbf{W}^{1/2} \\ &= \mathbf{W}^{1/2} (\delta^2 \mathbf{W}^{-1/2} \mathbf{W}^{-1/2}) \mathbf{W}^{1/2} \\ &= \delta^2 (\mathbf{W}^{1/2} \mathbf{W}^{-1/2}) (\mathbf{W}^{-1/2} \mathbf{W}^{1/2}) \\ &= \delta^2 \mathbf{I} \end{aligned} \quad (3.13)$$

$$\mathbf{S}(\hat{\boldsymbol{\beta}}) = \sum_{i=1}^N (\sqrt{w_i} y_i - \sqrt{w_i} \hat{y}_i)^2 \quad (3.14)$$

Então, de acordo com Weisberg (2005), a matriz de covariância do erro, dado pela Equação 3.9, pode ser reescrita através da Equação 3.15.

$$\text{Var}(\boldsymbol{\varepsilon}) = \delta_{\varepsilon}^2 \begin{bmatrix} w_{11} & 0 & 0 & 0 & 0 \\ 0 & w_{22} & 0 & 0 & 0 \\ 0 & 0 & w_{33} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & w_{NN} \end{bmatrix} \quad (3.15)$$

### 3.3.3 Método WLS Tasker - Stedinger

A principal mudança introduzida no método WLS por Tasker e Stedinger (1986) é a separação da matriz de covariância dos erros do modelo ( $\boldsymbol{\Lambda}(\delta^2)$ ), em variância do erro do modelo e covariância dos erros amostrais (Equação 3.16) (STEDINGER e TASKER, 1986b).

$$\mathbf{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \boldsymbol{\Lambda}(\delta^2) = \delta_{\varepsilon}^2 \mathbf{I} + \boldsymbol{\Sigma} \quad (3.16)$$

onde  $\mathbf{I}$  é a matriz identidade e  $\boldsymbol{\Sigma}$  a matriz de covariância dos erros amostrais.

Segundo Tasker e Stedinger (1986) o estimador dos parâmetros  $\boldsymbol{\beta}$  é dado pela Equação (3.17), sendo que a única diferença é que a matriz de pesos é dada pelas Equações 3.18 e 3.19.

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}^T \boldsymbol{\Lambda}(\delta^2)^{-1} \mathbf{X}]^{-1} \mathbf{X}^T \boldsymbol{\Lambda}(\delta^2)^{-1} \hat{\mathbf{y}} \quad (3.17)$$

$$w_{ij} = [\boldsymbol{\Lambda}(\delta^2)_{ij}]^{-1} \text{ para } i = j \quad (3.18)$$

$$w_{ij} = 0 \text{ para } i \neq j \quad (3.19)$$

De acordo com Stedinger e Tasker (1986), tal como acontece no GLS, uma dificuldade encontrada no procedimento WLS é que o estimador  $\boldsymbol{\beta}$  é definido em termos da variância do erro do modelo  $\delta_{\varepsilon}^2$ , a qual é desconhecida. Esta pode ser encontrada, resolvendo-se iterativamente a Equação 3.17 juntamente com a Equação 3.20 (erro quadrático médio - MSE) (Stedinger e Tasker, 1985).

$$(\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})^T \boldsymbol{\Lambda}(\delta_{\varepsilon}^2 \mathbf{I} + \boldsymbol{\Sigma})^{-1} (\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}}) = N - P - 1 \quad (3.20)$$

Tanto no WLS quanto no GLS é necessário determinar uma matriz de covariância dos erros amostrais. Esta matriz é função da precisão de como o modelo descreve os valores da estatística de vazão, bem como do erro de amostragem nas estimativas. O WLS difere do GLS

principalmente porque não leva em conta a correlação cruzada entre os fluxos (STEDINGER e TASKER, 1985).

Tasker (1980) recomendou a utilização do método WLS em análises regionais de vazões máximas. Os resultados encontrados mostraram que a variância do erro do modelo, para o método WLS foi muito inferior quando comparada com a regressão OLS.

Tasker e Stedinger (1986) empregaram a regressão WLS para obter um estimador generalizado para o coeficiente de assimetria. Os pesos para o método WLS foram determinados separando-se a variância do erro em duas partes, variância do erro do modelo e variância do erro amostral. Tais autores relataram que resultados encontrados através do modelo WLS foram melhores que os do modelo OLS, principalmente no que se refere à variância do erro do modelo.

Wright e Ensminger (2004) utilizaram a regressão logística e o método WLS na estimativa de vazões mínimas  $Q_{7,2}$ ,  $Q_{7,10}$  para locais não monitorados do estado da Louisiana, EUA. Foram atribuídos diferentes pesos, a partir da análise das estações. Eles relataram que o WLS apresentou bom desempenho.

Alexandre e Martins (2005) utilizaram o modelo WLS e GLS em um estudo de vazões médias de longo período para o Estado do Ceará, no Brasil. Eles utilizaram 25 bacias hidrográficas do Estado do Ceará. Os modelos foram comparados através da análise da variância média de predição (AVP) e da variância do erro do modelo ( $\delta_\epsilon^2$ ). Os resultados, tanto para o WLS quanto para o GLS, foram similares, pois a correlação cruzada entre as vazões médias concorrentes apresentaram baixíssimos valores. Logo, o modelo adotado foi o WLS que continha as características físicas e climáticas, declividade média, precipitação média e percentual da bacia situada no cristalino.

### **3.4 Método GLS**

De acordo com o teorema de Gauss-Markiv-Aitken (RAO e TOUTENBERG, 1999), se  $\mathbf{W}$  é conhecido, o estimador não enviesado e de mínima variância para  $\boldsymbol{\beta}$  para o modelo GLS não depende de  $\delta_\epsilon^2$ , sendo dado por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{W}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{-1} \hat{\mathbf{y}} \quad (3.21)$$

na qual  $\mathbf{W}$  é uma matriz de pesos. No método GLS clássico, a matriz de covariância dos erros é descrita pela Equação 3.22, a qual é similar à Equação 3.9, sendo que a única diferença é o fato de  $\mathbf{W}$  ter pesos fora da diagonal principal.

$$E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) = \delta_\varepsilon^2 \mathbf{W}^{-1} \quad (3.22)$$

De acordo com Rao e Toutenberg (1999), a matriz  $\mathbf{W}$  pode ser construída por meio da Equação 3.23.

$$\mathbf{W} = \begin{bmatrix} 1 & \rho & \cdots & \rho^{N-1} \\ \rho & 1 & \cdots & \rho^{N-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & 1 \end{bmatrix} \quad (3.23)$$

na qual  $\rho$  é o coeficiente de correlação considerado em seus valores absolutos.

No entanto, na prática, a matriz  $\mathbf{W}$  é raramente conhecida e deve ser estimada, o que geralmente provoca uma perda de eficiência dos estimadores  $\beta$  e  $\delta^2$  (RAO e TOUTENBERG, 1999).

### 3.4.1 Método GLS Stedinger e Tasker

Stedinger e Tasker (1985, 1986a, 1986b) fizeram uma mudança no método GLS de forma a torná-lo mais aplicável à análise hidrológica. A principal mudança é a separação da matriz de covariância dos erros do modelo. O método assume que o erro total resulta de duas fontes: o erro do modelo  $\boldsymbol{\varepsilon}_i$ , que é identicamente distribuído com média zero e variância constante, e os erros amostrais, advindos do fato do real valor de  $\mathbf{y}_i$  não ser conhecido, sendo que somente as estimativas da variável de interesse são disponíveis. Nesse caso, a Equação 3.4 pode ser reescrita sob a forma.

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\omega} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta} \quad (3.24)$$

onde  $\boldsymbol{\omega}$  é o erro dos estimadores amostrais. Assim, os erros do modelo de regressão  $\boldsymbol{\eta}_i$  são a combinação do erro amostral da estimativa  $\hat{\mathbf{y}}_i$  de  $\mathbf{y}_i$  e do erro do modelo  $\boldsymbol{\varepsilon}_i$ . O erro total  $\boldsymbol{\eta}$  tem média zero e matriz de covariância dada pela Equação 3.25.

$$\mathbf{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \Lambda(\delta^2) = \delta_\varepsilon^2 \mathbf{I} + \boldsymbol{\Sigma} \quad (3.25)$$

sendo  $\Lambda(\delta^2)$  a matriz de covariância dos erros, que depende do erro do modelo, e  $\Sigma$  a matriz de covariância dos erros amostrais, a qual é obtida utilizando os dados de distâncias e das correlações cruzadas entre os fluxos concorrentes.

De acordo com Stedinger e Tasker (1985, 1986a, 1986b), o estimador GLS para os parâmetros  $\beta$  é dado pela Equação 3.26.

$$\hat{\beta} = [\mathbf{X}^T \Lambda(\delta^2)^{-1} \mathbf{X}]^{-1} \mathbf{X}^T \Lambda(\delta^2)^{-1} \hat{\mathbf{y}} \quad (3.26)$$

E, finalmente, a variância do erro do modelo ( $\delta_\varepsilon^2$ ) é dada pela Equação 3.27.

$$\delta_\varepsilon^2 = \frac{(\hat{\mathbf{y}} - \mathbf{X}\hat{\beta})^T \Lambda(\delta^2)^{-1} (\hat{\mathbf{y}} - \mathbf{X}\hat{\beta})}{N - P - 1} \quad (3.27)$$

#### 3.4.1.1 Matriz de covariância dos erros amostrais

Tasker (1980) e Kuczera (1983) elucidaram a respeito da dificuldade associada com a estimativa da matriz de covariância dos erros amostrais. Sendo assim, foram desenvolvidos procedimentos, a fim de estimar a precisão do modelo de regressão subjacente, bem como a variabilidade dos estimadores estatísticos amostrais de vazão e de correlação cruzada.

Stedinger e Tasker (1985), assim como Chow (1964) propuseram a utilização da Equação 3.28, para representar o quantil da variável hidrológica, neste caso a vazão.

$$\hat{y}_i = m_{i'} + K s_i \quad (3.28)$$

na qual  $K$  é o fator de frequência,  $m_{i'}$  e  $s_i$  correspondem a média da variável hidrológica para cada estação e o desvio padrão, respectivamente.

Portanto, segundo Stedinger e Tasker (1985), quando  $m_{i'}$  e  $s_i$  não são correlacionados e as vazões se distribuem de forma simétrica, os elementos da matriz de variância e covariância dos erros amostrais para a Equação 3.28 são dados pelas Equações 3.29 e 3.30.

$$\Sigma(\hat{y})_{ij} = \frac{\mu_2}{N_i} \left[ 1 + K^2 \left( \frac{\kappa - 1}{4} \right) \right] \text{ para } i = j \quad (3.29)$$

$$\Sigma(\hat{y})_{ij} = \frac{\rho_{ij}N_{ij}\sigma_i\sigma_j}{N_jN_i} \left[ 1 + \rho_{ij}K^2 \left( \frac{\kappa-1}{4} \right) \right] \text{ para } i \neq j \quad (3.30)$$

onde  $\mu_2$  é a variância amostral,  $\kappa$  o coeficiente de curtose populacional,  $N_{ij}$  representa o período em comum de dados nas estações  $i$  e  $j$  e  $N_i$  e  $N_j$  o período de dados nas estações  $i$  e  $j$ , respectivamente.

Tasker e Stedinger (1989) propuseram que, para distribuição Log-Pearson tipo 3 os elementos da matriz de covariância são determinados pelas Equações 3.31 e 3.32.

$$\Sigma(\hat{y})_{ij} = \frac{\mu_2}{N_i} \left[ 1 + K_i\gamma_{1i} + \frac{K^2}{2} (1 + 0,75\gamma_{1i}^2) \right] \text{ para } i = j \quad (3.31)$$

$$\Sigma(\hat{y})_{ij} = \frac{\rho_{ij}N_{ij}\sigma_i\sigma_j}{N_jN_i} \left[ 1 + \frac{K_i\gamma_{1i}}{2} + \frac{K_j\gamma_{1j}}{2} + \frac{K_iK_j}{2} (\rho_{ij} + 0,75\gamma_{1i}\gamma_{1j}) \right] \text{ para } i \neq j \quad (3.32)$$

na qual  $\gamma_{1i}$  e  $\gamma_{1j}$  denotam os coeficientes de assimetria das estações  $i$  e  $j$ , respectivamente.

Griffis e Stedinger (2007) estenderam a Equação 3.32 com o intuito dela refletir todo erro associado à estimativa dos quantis amostrais. Tal medida foi adotada para incluir os erros envolvidos na estimativa dos coeficientes de assimetria. Sendo assim, a Equação 3.34 apresenta os novos estimadores para a covariância dos erros amostrais para a distribuição Log-Pearson tipo 3. A Equação 3.33 (variância do erro amostral) foi uma aproximação proposta por Chowdry e Stedinger (1991), com a incorporação do erro associado à estimativa do coeficiente de assimetria.

$$\begin{aligned} \Sigma(\hat{y})_{ij} = & \frac{\mu_2}{N_i} + (1 - W_i)^2 s_i^2 MSE_{Gi} \left( \frac{\partial K_i}{\partial \gamma_{1i}} \right)^2 \left[ 1 + K_i\gamma_{1i} + K^2 \left( \frac{1}{2} + \frac{3}{8} \gamma_{1i}^2 \right) + \right. \\ & \left. + W_i K_i \frac{\partial K_i}{\partial \gamma_{1i}} \left( 3\gamma_{1i} + \frac{3}{4} \gamma_{1i}^3 \right) + W_i^2 \left( \frac{\partial K_i}{\partial \gamma_{1i}} \right)^2 \left( 6 + 9\gamma_{1i}^2 + \frac{15}{8} \gamma_{1i}^4 \right) \right] \text{ para } i = j \quad (3.33) \end{aligned}$$

na qual  $W_i$  é o peso calculado usando  $MSE_{Gi}$ , o qual é estimado a partir de uma equação proposta em Griffis (2006) e  $s_i$  é o desvio padrão amostral.  $\partial K_i / \partial \gamma_{1i}$  é a derivada do fator de frequência em relação ao coeficiente de assimetria populacional, calculada através da aproximação apresentada em Chowdry e Stedinger (1991).

$$\begin{aligned}
\Sigma(\hat{y})_{ij} = & \frac{\rho_{ij}N_{ij}\sigma_i\sigma_j}{N_jN_i} \left[ 1 + \frac{K_i\gamma_{1i}}{2} + \frac{K_j\gamma_{1j}}{2} + \frac{K_iK_j}{2} (\rho_{ij} + 0,75\gamma_{1i}\gamma_{1j}) + \right. \\
& + \frac{1}{2}W_iK_j\gamma_{1i} \frac{\partial K}{\partial \gamma_{1i}} (3\rho_{ij} + 0,75\gamma_{1i}\gamma_{1j}) + \frac{1}{2}W_jK_i\gamma_{1j} \frac{\partial K}{\partial \gamma_{1j}} (3\rho_{ij} + 0,75\gamma_{1i}\gamma_{1j}) + \\
& \left. + W_iW_j\sigma_i\sigma_j \frac{\partial K}{\partial \gamma_{1i}} \frac{\partial K}{\partial \gamma_{1j}} \text{cov}[g_i, g_j] \right] \quad \text{para } i \neq j
\end{aligned} \tag{3.34}$$

na qual  $\text{cov}[g_i, g_j]$  é a covariância entre os coeficientes de assimetria amostrais para as estações  $i$  e  $j$ , e  $\rho_{ij}$  representa a correlação cruzada entre as vazões concorrentes.

Vogel e Kroll (1990) tentaram reduzir o erro associado à variância do erro amostral, aumentando o número de registros de vazões mínimas a serem introduzidos no modelo. Esse procedimento acarretou melhorias modestas no modelo utilizado em seus estudos. Esses salientam que para a diminuição da variância do erro amostral, deve-se avaliar os erros de medição, a forma do modelo e a inclusão de variáveis explicativas relacionadas à bacia. Neste mesmo contexto, Wallis (1965) relatou que não se pode esperar descobrir relações físicas básicas usando procedimentos estatísticos multivariados, sem o conhecimento prévio das relações físicas.

Em relação à variância do erro do modelo ( $\delta_\epsilon^2$ ), sua estimativa é feita através de um procedimento iterativo (TASKER e STEDINGER, 1989; REIS Jr. *et al.*, 2004, 2005).

#### 3.4.1.2 Correlação cruzada entre os fluxos concorrentes

A aplicação do GLS Stedinger e Tasker requer a estimativa de uma matriz de covariância dos erros amostrais (Stedinger e Tasker, 1985) e, para construção desta matriz, é importante que se estabeleça a correlação cruzada entre os fluxos concorrentes.

Hosking e Wallis (1997) e Madsen e Rosbjerg (1997) demonstraram que a correlação cruzada entre os fluxos concorrentes afeta a precisão dos estimadores dos quantis.

A correlação cruzada entre os fluxos concorrentes amostrais é determinada através da Equação 3.35.



$$\rho_{ij} = \frac{\sum_{i=1}^N \left( z_i - \frac{\sum_{i=1}^N z_i}{N} \right) \left( z_j - \frac{\sum_{j=1}^N z_j}{N} \right)}{\sqrt{\left( \sum_{i=1}^N \left( z_i - \frac{\sum_{i=1}^N z_i}{N} \right)^2 \right) \left( \sum_{j=1}^N \left( z_j - \frac{\sum_{j=1}^N z_j}{N} \right)^2 \right)}} \quad (3.35)$$

onde  $z_i$  e  $z_j$  corresponde, por exemplo, aos valores das  $Q_7$  mínimas anuais observadas, das estações  $i$  e  $j$ , respectivamente.

No entanto, ao se utilizar a Equação 3.35, é comum encontrar valores de  $\rho_{ij}$  negativos. De acordo com Tasker e Stedinger (1989), ao se fazer uso dos valores das correlações baseadas nas estimativas amostrais, frequentemente a matriz de covariância dos erros amostrais pode não possuir inversa. Isto é devido ao fato da estimativa de  $\rho_{ij}$  ser imprecisa para séries curtas de vazões usualmente encontradas. Tais séries podem resultar em um conjunto de dados para estimativa  $\rho_{ij}$  que não fazem nenhum sentido hidrológico ou estatístico. Deste modo, para solucionar o problema no que se refere à estimativa de  $\rho_{ij}$ , tais autores recomendaram a suavização da correlação amostral utilizando uma função em relação à distância entre os pares de estações.

$$\hat{\rho}_{ij} = \hat{\rho}(d_{ij}) \quad (3.36)$$

Tasker e Stedinger (1989) determinaram a equação para  $\hat{\rho}_{ij}$  através do ajuste de uma função do tipo potencial ao gráfico da correlação cruzada *versus* distância entre as estações.

No entanto, de acordo com Martins e Stedinger (2000), se a correlação definida através das distâncias não for clara, pode se assumir uma função de correlação homogênea (Equação 3.37).

$$\hat{\rho}_{ij} = \bar{\rho} \quad (3.37)$$

onde  $\bar{\rho}$  corresponde ao coeficiente de correlação médio.

Martins e Stedinger (2002) utilizaram o experimento de Monte Carlo para determinar a relação necessária da função de correlação cruzada entre as vazões máximas anuais concorrentes ( $\hat{\rho}_{ij}$ ). O modelo de correlação cruzada proposto foi

$$\rho(\hat{y}_i, \hat{y}_j) = \text{Sign}(\hat{\rho}_{ij}) cf_{xy} |\hat{\rho}_{ij}|^v \quad (3.38)$$

na qual  $v$  é uma constante,  $\hat{\rho}_{ij}$  é dada pela Equação 3.35 e o fator  $cf_{xy}$  pela Equação 3.39.

$$cf_{xy} = \frac{N_{ij}}{\sqrt{(N_{ij} + N_i)(N_{ij} - N_j)}} \quad (3.39)$$

Vogel e Kroll (1990) relataram que a correlação cruzada média das sequências das vazões com 7 dias de duração foi 0,35. Este valor foi usado para caracterizar a matriz de covariância dos erros amostrais. Tais autores afirmaram que não foi possível encontrar qualquer relação entre a correlação cruzada e a distância.

Veilleux (2011) explorou, para o método GLS Bayesiano, as correlações cruzadas entre os coeficientes de assimetria amostrais, a fim de que se pudesse obter melhores estimativas dos parâmetros do modelo. No entanto, a autora afirmou que se as correlações cruzadas possuírem valores elevados, os estimadores GLS podem se tornar relativamente complicados. A mesma verificou que a precisão da estimativa da correlação cruzada, entre quaisquer duas estações, não é suficiente para justificar os pesos aparentemente incorretos que a análise do GLS Bayesiano gera.

#### 3.4.1.3 Histórico de utilização do GLS em engenharia de recursos hídricos

Segundo Stedinger e Tasker (1985), o método GLS fornece uma melhor aproximação se comparado aos métodos OLS e WLS, ou seja, agrega mais informações ao modelo.

Nesse sentido, o modelo de regressão GLS, proposto por Stedinger e Tasker (1985, 1986a, 1986b) para a estimativa de quantis de vazão, leva em consideração a diferença de tamanho das séries hidrológicas e a correlação cruzada entre dados concorrentes de diferentes locais.

Ainda de acordo com Stedinger e Tasker (1985), o GLS assume inicialmente que as vazões são independentes e identicamente distribuídas e, assim como no WLS, os resíduos são heterocedásticos. Os autores utilizaram simulação de Monte Carlo e demonstraram que o

método GLS fornecia estimadores mais precisos dos parâmetros, melhores estimadores da variância do erro amostral e um estimador não enviesado para a variância do erro do modelo. Em geral, o erro quadrático médio (*mean square error* - MSE) amostral dos estimadores GLS foi menor que o (MSE) do OLS, principalmente quando a variância do erro do modelo foi pequena ou quando a correlação cruzada foi grande.

Embora o GLS tenha sido proposto há quase três décadas, somente recentemente o método começou a ser aplicado em um grande número de estudos. Reis Jr. et al. (2005) lista uma série de estudos que utilizaram o GLS para regionalizar diferentes dados hidrometeorológicos em diferentes locais do mundo. A metodologia tem sido utilizada para estimar parâmetros de modelos de regressão regionais de quantis extremos de vazão máxima anual (TASKER et al., 1986; WMO, 1994; KJELDSEN e ROSBJERG, 2002; FEASTER e TASKER, 2002), vazão mínima anual (VOGEL e KROLL, 1990), chuvas extremas (MADSEN et al., 2002) e parâmetros de distribuições de probabilidades (MADSEN e ROSBJERG, 1997).

Müller e Fill (1997) utilizaram o método GLS na estimativa de dois parâmetros regionais para a determinação dos quantis de cheia, utilizando a distribuição de probabilidade Generalizada de Valores Extremos (GEV).

Ries III (1994) desenvolveu um estudo de regionalização onde utilizou o método GLS na estimativa da vazão média mínima Massachusetts, EUA. Foram atribuídos pesos (valores de entrada na matriz de covariância dos erros amostrais) ao modelo WLS e GLS, de acordo com as diferenças no comprimento dos registros e através da correlação cruzada entre os fluxos concorrentes. Nos resultados encontrados, a variância média dos erros amostrais foi elevada. Contudo os autores ressaltam que a precisão do modelo poderá ser aumentada, caso ocorra uma diminuição na variância do erro amostral.

Haddad *et al.* (2011) desenvolveram um estudo comparativo entre os métodos de regressão GLS e OLS, e os empregaram na regionalização para New South Wales, Austrália. O método GLS apresentou menores valores relacionados ao erro quadrático médio.

Vogel e Kroll (1990) desenvolveram equações regionais de regressão para estimar vazões médias mínimas com durações de 3, 7, 14 e 30 dias. Para isto, presumiram que toda série de vazão mínima é lognormalmente distribuída e utilizaram os métodos de regressão OLS e GLS. Nos resultados encontrados, o método GLS teve estimativa de parâmetros similar à do OLS. Contudo, o método GLS teve vantagens significativas sobre o procedimento OLS,

principalmente em locais onde se dispunha de séries curtas de vazões. Os autores salientam que as vazões médias mínimas com 7 dias de duração,  $Q_7$ , são correlacionadas no espaço.

Griffis e Stedinger (2007) determinaram um modelo de regionalização para 162 estações no estado da Carolina do Sul, EUA, divididas em 4 regiões. Incorporaram na matriz de covariância dos erros amostrais, o erro envolvido na estimativa do coeficiente de assimetria. Acrescentaram mais informações para a realização da análise estatística incluindo relações para verificar a multicolinearidade, além de propor um pseudo- $R^2$  apropriado, para ser utilizado na regressão GLS. O modelo GLS sobressaiu se comparado ao OLS e WLS.

Martins e Stedinger (2002) identificaram, através do experimento de Monte Carlo, uma relação entre a correlação cruzada dos estimadores do coeficiente de assimetria para a distribuição Log-Pearson tipo 3, e para os parâmetros de forma das distribuições Generalizada de Pareto (GP) e Generalizada de Valores Extremos (GEV). A partir desta análise, desenvolveram uma função matemática que descrevia a correlação cruzada entre as vazões concorrentes, onde foi incorporada uma constante relacionada com os parâmetros de forma para Log-Pearson tipo 3, GP e GEV.

Reis Jr. *et al.* (2005) introduziram uma extensão promissora ao método GLS, com aplicação dos conceitos da análise Bayesiana, à regionalização de variáveis hidrológicas. Relatam que nos estudos anteriormente listados, e em diversos outros, o método GLS teve resultados superiores aos demais métodos de regressão. Os autores mostram que as estimativas obtidas pelo GLS têm menor variância e menor erro. Assim, os autores recomendaram o método, sobretudo o GLS Bayesiano, para análises hidrológicas regionais, pois este descreveu de forma mais realista os possíveis valores da variância do erro do modelo, especialmente nos casos em que a variância do erro amostral foram muito grandes.

Haddad e Rahman (2012) propuseram a utilização do método GLS com a inclusão da abordagem Bayesiana (*Bayesian Generalized Least Squares* - BGLS) em um estudo de análise regional de frequência, em regiões de influências, as quais são delimitadas, por exemplo, com intuito de minimizar a variância do erro amostral. Neste estudo foram utilizadas 399 bacias do leste da Austrália. Foram regionalizados os quantis de cheias (técnica de regressão dos quantis - QRT) e os três primeiros momentos da distribuição Log-Pearson tipo 3 (técnica de regressão dos parâmetros - PRT). Para avaliar o desempenho global dos modelos QRT e PRT, foi empregado o procedimento de validação cruzada. Os resultados indicaram que o BGLS, tanto para o QRT quanto o PRT apresentaram bons resultados.

Haddad *et al.* (2012) regionalizaram o quantil (QRT) de 2 a 100 anos de intervalo médio de recorrência e os parâmetros da distribuição Log-Pearson tipo 3 dos logaritmos das vazões máximas anuais. O método de regressão empregado foi GLS Bayesiano. O método foi aplicado em 53 bacias na Tasmânia, Austrália. Para o modelo QRT verificaram que a área de drenagem e a intensidade da chuva de projeto eram as variáveis explicativas mais importantes. Para o PRT, foi adotado um total de quatro variáveis explicativas. Verificaram que o modelo QRT forneceu estimativas mais precisas de quantis das vazões máximas para altos valores de intervalo médio de recorrência, enquanto que o PRT forneceu estimativas melhores para pequenos valores de intervalo médio de recorrência.

Veilleux (2011) apresentou uma extensão da análise Bayesiana quase-analítica para o B-GLS, estudado por Reis Jr. *et al.*, (2005). Foram incorporados o pseudo-R<sup>2</sup> e a pseudo-ANOVA, sendo que seu trabalho também se voltou para a análise de desempenho do B-WLS e B-GLS através de medidas alternativas de influência (Cook's D, DFITS e DFBETAS) e *leverage*. Os modelos de regressão foram aplicados em diferentes partes do Estados Unidos. O objetivo do trabalho era estimar o coeficiente de assimetria regional. Ficou constatado que a versão modificada do B-GLS evitou problemas de instabilidade encontradas na metodologia original do B-GLS, devido às grandes correlações cruzadas entre as vazões de pico.

### **3.4.2 Considerações finais sobre os métodos de regressão**

A fim de que se possa avaliar qual o melhor modelo de regressão para cada um dos métodos (OLS, WLS e GLS), normalmente faz-se uma análise da variância média de predição (TASKER e STEDINGER, 1986; REIS Jr. *et al.*, 2005, GRIFFIS e STEDINGER, 2007). E, para definir qual método é melhor, é comum a análise da variância do erro do modelo. É importante salientar que existem outras maneiras de se verificar se o modelo referente a um dos métodos é mais adequado que o outro. Para o método OLS há de se verificar algumas premissas relacionadas aos resíduos. Sendo assim, deve-se realizar os testes de aderência (distribuição normal), da média dos resíduos igual a zero (NAGHETTINI e PINTO, 2007), de independência e de variância constante (WOOLDRIDGE, 2009; GUJARATI e PORTER, 2008). Na metodologia deste trabalho será apresentada algumas formas para verificar qual o método teve melhor desempenho e também realizar uma análise referente aos resíduos, principalmente os do método OLS.

### **3.5 Considerações finais da revisão de literatura**

A revisão de literatura aqui apresentada foi feita de modo a incorporar os estudos que contribuíram para o desenvolvimento do trabalho. Alguns estudos que repetiam métodos similares e, principalmente, que obtinham resultados muito parecidos, foram suprimidos com intuito de tornar a revisão mais concisa.

Avaliando a literatura estudada, pode-se evidenciar a enorme aplicabilidade do método GLS, principalmente porque agrega mais informações consideradas importantes na análise entre as séries de vazões.

Por fim, o modelo GLS sistematizado por Stedinger e Tasker (1985), objeto de estudo deste trabalho, será detalhado no capítulo 4, o qual envolve as etapas metodológicas adotadas para construção do modelo final de regionalização.

## 4 METODOLOGIA

A metodologia empregada foi dividida em 6 partes. Sendo que, a primeira parte se concentra na descrição da área de estudo, a segunda refere-se à coleta e análise dos dados das estações fluviométricas e das características físicas e climáticas, a terceira corresponde à divisão da área de estudo em regiões homogêneas, a quarta na construção do modelo GLS Stedinger e Tasker, bem como na descrição do processo construtivo da matriz de covariância dos erros amostrais e a quinta descreve a construção dos métodos OLS, WLS e GLS. Por fim, a sexta parte faz alusão aos procedimentos de avaliação dos métodos de regressão.

Para aplicação dos métodos OLS, WLS e GLS, foram escolhidas as sub-bacias 40 e 41, do rio São Francisco, localizadas na região central de Minas Gerais. Essas bacias foram escolhidas, devido ao fato das mesmas terem sido, recentemente, objeto de estudos de regionalização de vazão e precipitação (CPRM, 2001; EUCLYDES *et al.*, 2001, LIMA, 2005 e IGAM, 2012).

### **4.1 Localização e caracterização da bacia hidrográfica do rio São Francisco**

A bacia do rio São Francisco está compreendida entre as latitudes 7° 00' e 21° 00' Sul e as longitudes 35° 00' e 47° 40' Oeste (ANEEL, 1998), possuindo uma área de drenagem de 639.219 km<sup>2</sup>, com vazão média de 2.850 m<sup>3</sup>/s (IGAM, 2012). O rio São Francisco deságua no oceano Atlântico na divisa entre Alagoas e Sergipe.

#### **4.1.1 Caracterização da bacia do Alto São Francisco**

O rio São Francisco possui a extensão de aproximadamente 2.700 km. Nasce na serra da Canastra em Minas Gerais. De acordo com CPRM (1995a citado por CPRM 2001), inicialmente o seu segmento possui a orientação Leste-Oeste e, logo após o sétimo quilômetro, passa a correr seguindo nova direção (Oeste-Leste) até a Lagoa de Martins. Deste ponto até atingir o reservatório de Três Marias assume nova direção, Sudoeste-Nordeste.

#### **4.1.2 Localização e caracterização das sub-bacias 40 e 41**

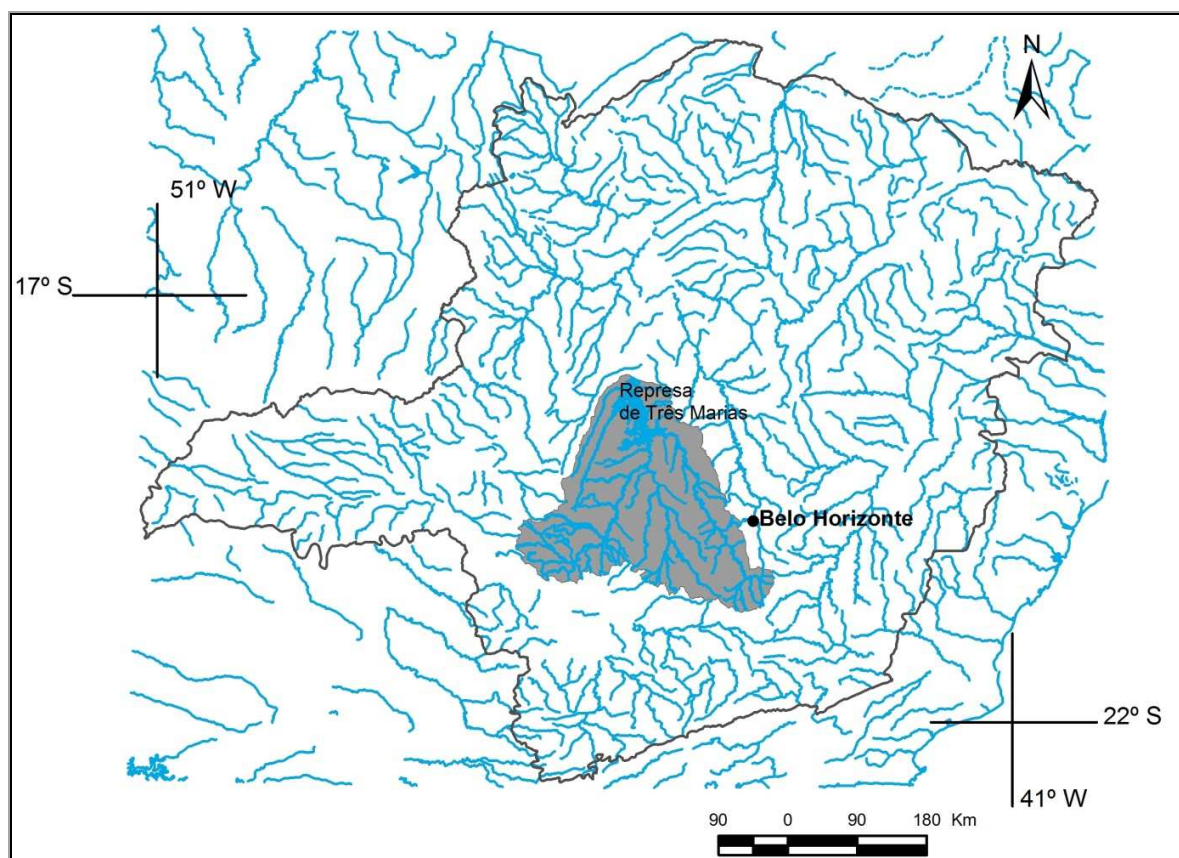
As sub-bacias 40 e 41 (classificação ANEEL) estão localizadas na região denominada como Alto São Francisco, cuja área total é de 90.634,16 km<sup>2</sup>, abrangendo 170 municípios no Estado de Minas Gerais, contendo uma população total de aproximadamente 6.774.302 habitantes (CPRM, 2001).

O alto São Francisco apresenta variabilidade climática relacionado à transição de um clima úmido e subúmido (SILVA e CLARKE, 2004). A vegetação nativa preponderante é o cerrado, caracterizada pela existência de gramíneas, arbustos e árvores de médio porte, as quais possuem como particularidade a presença de troncos e galhos retorcidos e folhas espessas e coriáceas (CPRM, 1995a, 1995b citado por CPRM, 2001).

#### 4.1.2.1 Sub-bacia 40

De acordo com CPRM (2001), a sub-bacia 40 está localizada na região central do Estado de Minas Gerais, possui uma área de cerca de 51.000 km<sup>2</sup>, compreendida entre as nascentes do rio São Francisco e a Barragem de Três Marias, com latitudes 18° 15' e 20° 40' Sul e longitudes 43° 50' e 46° 10' Oeste.

A sub-bacia 40 abrange o alto São Francisco e seus afluentes até a barragem de Três Marias incluindo também os rios Paraopeba, Pará, Borrachudo, Indaiá e Ajudas. A Figura 4.1 apresenta a localização da sub-bacia 40 no estado de Minas Gerais.



**Figura 4.1** - Localização da sub-bacia 40 no estado de Minas Gerais

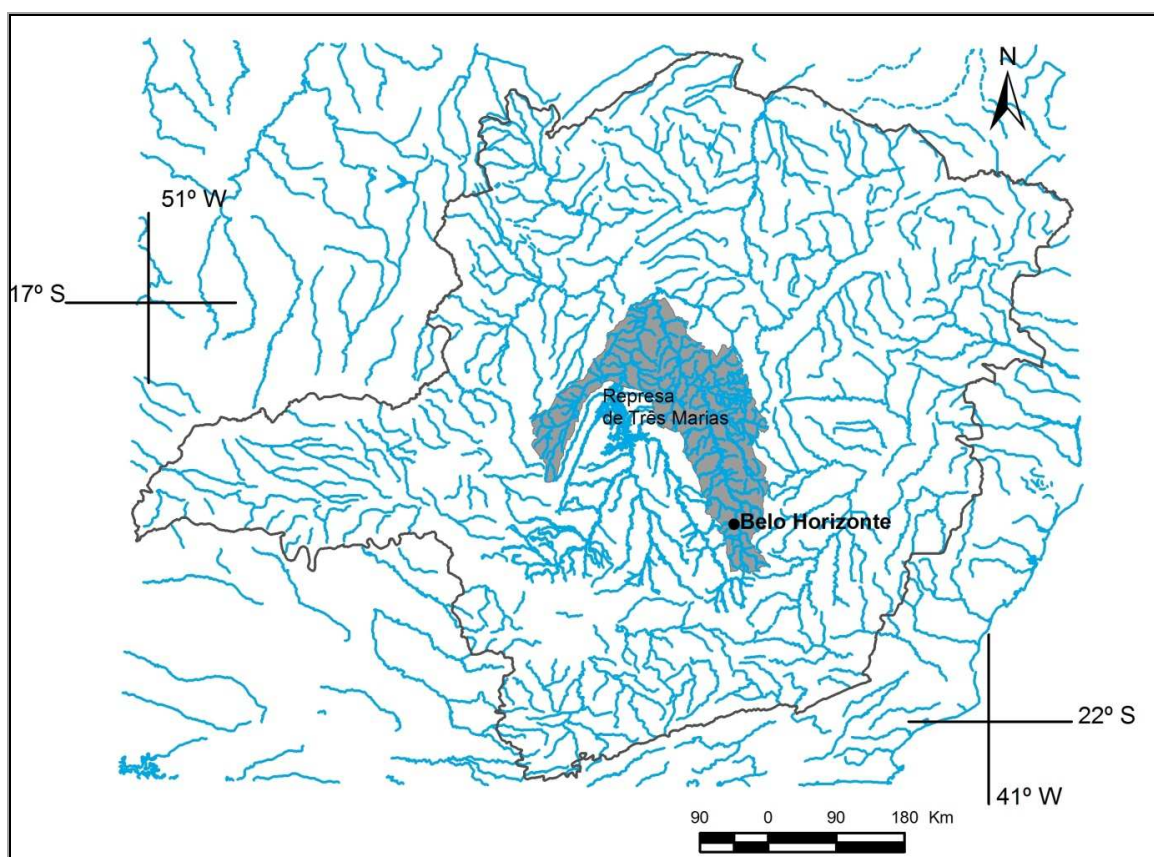
A sub-bacia 40 é constituída fundamentalmente sobre rochas de idade pré-cambriana, desde arqueadas até protozoicas superiores (CPRM, 1995a citado por CPRM, 2001).



O relevo da desta sub-bacia é caracterizado por planaltos, depressões e áreas dissecadas (resultante dos processos morfoclimáticos relacionados ao condicionamento geológico). As unidades encontradas são: Planalto Dissecado do Centro-Sul e Leste de Minas, Depressão São Franciscana e Planalto do São Francisco (CPRM, 1995a *citado por* CPRM, 2001).

#### 4.1.2.2 Sub-bacia 41

A sub-bacia 41, de acordo com CPRM (2001), contém a parcela da bacia do rio São Francisco, com área aproximada de 39.000 km<sup>2</sup>. Compreende o rio das Velhas e áreas incrementais a jusante da barragem de Três Marias. Está localizada no centro do estado de Minas Gerais, entre as latitudes 17° 20' e 20° 30' Sul e longitudes 43° 30' e 46° 20' Oeste. O rio das Velhas e o rio Abaeté são os maiores tributários do rio São Francisco nesta bacia. A Figura 4.2 apresenta a localização da sub-bacia 41 no estado de Minas Gerais.



**Figura 4.2** - Localização da sub-bacia 41 no estado de Minas Gerais

A sub-bacia 41 apresenta, principalmente, formações rochosas do tipo meta-vulcânicas, e meta-vulcano-sedimentares, região drenada pelo rio das Velhas, desde as nascentes do alto rio das Velhas até o norte de Belo Horizonte. Pode-se destacar também rochas dos Grupos: Bambuí, Macaúbas e Espinhaço (CPRM, 1995b *citado por* CPRM, 2001).

O relevo desta sub-bacia é caracterizado por planaltos, unidades morfoestruturais, depressões e áreas dissecadas (decorrente de processos morfoclimáticos). As unidades morfológicas existentes são: O Planalto do São Francisco, a Depressão São Franciscana, o Quadrilátero Ferrífero e a Serra do Espinhaço (CPRM, 1995b *citado por* CPRM, 2001).

#### 4.1.3 Caracterização dos aquíferos da bacia Alto São Francisco

No período de estiagem, que se concentra principalmente entre os meses de junho e meados de outubro, o regime dos rios é controlado, em grande parte, pelos sistemas aquíferos. O sistema de aquíferos da bacia do alto São Francisco (Figura 4.3) será descrito logo a seguir, de forma sucinta.

De acordo com Mourão *et al.* (2001), a separação de forma individual do sistema aquífero, fundamenta-se nos aspectos relacionados à litoestrutura e à natureza da permeabilidade das rochas.

Os sistemas aquíferos presentes na bacia do Alto São Francisco (Mourão *et al.*, 2001), são:

- **Sistema aquífero granular**, o qual é caracterizado pelos depósitos sedimentares ou coberturas detríticas, cuja formação data do período Cretáceo. As unidades hidrogeológicas presentes são: aquíferos em sedimentos aluviais, aquíferos colúvio-eluvionares e Aquíferos areníticos. Na Tabela 4.1, estão apresentadas algumas características das unidades hidrogeológicas encontradas nesse sistema aquífero.

**Tabela 4.1 – Unidades hidrogeológicas presentes no Sistema aquífero granular**

Unidades	Características	litologia ou unidades geológicas associadas
Aquíferos em sedimentos aluviais	<p>“É de grande interesse para exploração de água subterrânea, pois são fáceis de escavar ou perfurar; apresentam níveis potenciométricos rasos e têm, com regularidade, porosidade eficaz e permeabilidade elevadas.”</p> <p>“As aluviões são produzidas pela ação de rios ou demais formas de escoamento superficial e, são encontradas, nos canais fluviais, nas planícies de inundação e nos terraços.”</p> <p>“As ocorrências mais expressivas acontecem ao longo dos rios: São Francisco e Velhas.”</p>	Depósitos aluvionares e de terraços. Areias, siltes, argilas e cascalhos

Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

Continuação da Tabela 4.1

Aquíferos colúvio-eluviais	<p>“Compreendem as unidades aquíferas superficiais relacionadas ao manto de decomposição das rochas (saprolitos, elúvios e colúvios) e aos depósitos detríticos de cobertura do Terciário-Quaternário.”</p> <p>“Geralmente, a espessura não excede 10 metros, porém localmente pode atingir mais de 100 metros. A recarga procede-se quase exclusivamente através da infiltração das águas da chuva. A capacidade de armazenamento é dada por conseguinte, pela permeabilidade das camadas superficiais, pela espessura saturada do aquífero e pela natureza do substrato.”</p> <p>“A heterogeneidade dos sedimentos designa grande variabilidade de valores de porosidade eficaz e transmissividade.”</p>	Eluviões e coluviões, ocasionalmente associados a sedimentos aluvionares de canais suspensos, laterizados ou não laterizados. Areias finas a médias, argila, cascalho e canga.
Aquíferos areníticos	<p>“Envolvem os sedimentos cretáceos dos Grupos Urucuia e Mata da Corda e Areado. Geomorlogicamente, advém como superfícies tabulares resultantes do processo de aplainamento ocorrido no Terciário.”</p> <p>“Grupo Urucuia é constituído por arenitos avermelhados e, de forma subordinada, por conglomerados. Ocorre na margem direita do rio São Francisco.”</p> <p>“Grupo Mata da Corda é composta por rochas piroclásticas (brechas, lapilitos, tufos e cineritos), derrames alcalinos e arenitos cineríticos, localmente conglomeráticos. Possui menor distribuição superficial se comparada às outras unidades cretáceas.”</p> <p>“Grupo Areado é representado por uma sucessão sedimentar constituída por conglomerados, arenitos, siltitos, argilitos, arenito calcíferos, calcário e calcretes. Ocorrem na margem esquerda do rio São Francisco.”</p>	Sequências sedimentares e vulcânicas da bacia cretácea Alto SanFranciscana, composta pelos Grupos Urucuia, Mata da Corda e Areado.

Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

- **Sistema aquífero cárstico**, representados pelas rochas carbonáticas pré-cambrianas do subgrupo Paraopeba - Grupo Bambuí. A unidade hidrogeológica encontrada são os aquíferos de rochas carbonáticas. Na Tabela 4.2, estão apresentas algumas características das unidades hidrogeológicas encontradas nesse sistema aquífero.

**Tabela 4.2** – Unidades hidrogeológicas presentes no Sistema aquífero cárstico

Unidades	Características	litologia ou unidades geológicas associadas
Aquíferos de rochas carbonáticas	<p>“Sistema aquíferos desse tipo apresentam porosidade secundária resultante do alargamento, por dissolução, de superfície de descontinuidade tais como fraturas, planos de acamamento e falhas, gerando meio de circulação altamente anisotrópicos e heterogêneo formado por uma rede intrincada de condutos.”</p> <p>“Os calcários do Grupo Bambuí-Subgrupo Paraopeba ocorrem comumente recobertos pelos sedimentos da Formação Três Marias e unidades cretáceas.”</p> <p>“Os aquíferos carbonáticos representam o sistema de maior capacidade de produção.”</p>	Calcários e dolomitos protezóicos associados às fácies carbonáticas do Subgrupo Paraopeba-Grupo Bambuí, Grupos Paranoá e Vazante e Supergrupo Minas.

Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

- **Sistema aquífero fissurado-cárstico** é o sistema que possui maior representação superficial na bacia, agregando as unidades constituídas pela alternância de rochas pelíticas e carbonáticas, constituídas por margas, siltitos e ardósias com graus de calcário ou dolomito. A unidade presente são os aquíferos de rochas pelito-carbonáticas. Na Tabela 4.3, estão apresentas algumas características das unidades hidrogeológicas encontradas nesse sistema aquífero.

**Tabela 4.3** – Unidades hidrogeológicas presentes no Sistema aquífero fissurado-cárstico

Unidades	Características	litologia ou unidades geológicas associadas
Aquíferos de rochas pelito-carbonáticas	<p>“Apresenta feições estruturais marcadas por falhamentos e dobramentos são observados nos sedimentos situados nas bordas do Cráton e nas faixas móveis. No restante da área mostram-se praticamente indeformados. O maior potencial hidrogeológico é verificado nas porções tectonizadas e com ocorrência de camadas carbonáticas mais espessas.”</p> <p>“As fontes de recarga são correspondentes à infiltração direta de águas pluviais, em áreas de afloramento, ou de água superficial através de drenagens estruturalmente controladas e ainda à drenança dos sistemas granulares sobrejacentes. “O aportes de água superficial nas planícies de inundação dos principais cursos d’água, em períodos de cheia, também deve ser considerado.”</p>	Margas, siltitos, ardósia com intercalações de calcário associados às fácies pelito-carbonatadas do Subgrupo Paraopeba-Grupo Bambuí, Grupo Vazante e Grupo Paranoá.

Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

- **Sistema aquífero fissurado** configura o sistema em que a porosidade está associada à presença de falhas, fraturas e outras superfícies de descontinuidade, oriundas de esforços tectônicos. As unidades encontradas são os aquíferos de rochas metapelíticas e xistosas, aquíferos de rochas quartzíticas, aquíferos de rochas itabiríticas e aquíferos de rochas ígneas e metamórficas de alto grau. Na Tabela 4.4, estão apresentadas algumas características das unidades hidrogeológicas encontradas nesse sistema aquífero.

**Tabela 4.4 – Unidades hidrogeológicas presentes no Sistema aquífero fissurado**

Unidades	Características	litologia ou unidades geológicas associadas
Aquíferos de rochas Metapelíticas e Xistosas	<p>“São compostos de metapelitos e xistos de origem diversa relacionados aos grupos Bambuí, Paranoá, Vazante, Araxá, Canastra, Formação Paracatu, Supergrupo Espinhaço, Supergrupo Minas, Supergrupo Rio das Velhas e sequência metavulcano-sedimentar de Pium-hi.”</p> <p>“Os metassedimentos pelíticos do Grupo Bambuí ocorrem na região cratônica e na faixa móvel a oeste. Em ordem de importância de extensão superficial seguem os metapelitos pertencentes aos Grupos Paranoá, Canastra, Araxá e Vazante. “</p> <p>“A recarga é feita principalmente por meio do fluxo vertical descendente advindo das unidades granulares superiores e de forma significativa, pela infiltração da água superficial, diretamente nas fendas e fraturas, em trechos estruturalmente controlados.”</p>	<p>Filitos, folhelhos, siltitos, argilitos associados ao Grupo Bambuí, Grupo Paranoá e Vazante. Tilitos e diamictitos de matriz pelítica da Formação jequitaiá. Micaxistos dos Grupos Araxá e Canastra. Metalvulcânicas e metassedimentos pelíticos do Supergrupo Rio das Velhas (Grupo Maquiné e Grupo Nova Lima) e do <i>greenstone-belt</i> de Pium-hi.</p>
Aquíferos de rochas quartzíticas	<p>“As rochas quartzíticas geralmente destacam-se no relevo e sustentam as serras da Canastra, Espinhaço. Os metacóseos da Formação Três Marias (Grupo Bambuí) se configuram , geomorfologicamente como superfícies tabulares ou como suaves ondulações.”</p> <p>“A infiltração direta pelos sistemas de fraturas, a percolação através do manto de intemperismo e o aportes advindo dos sistemas granulares sobrepostos constituem os mecanismos de recarga. As nascentes difusas, de ocorrência generalizada, respondem pela manutenção das descargas de base dos rios que cruzam os terrenos quartzíticos.”</p>	<p>Quartzitos e metarenitos e metaconglomerados protezóicos e arqueanos associados aos Grupos Bambuí (Formação Três Marias), Canastra, aos Supergrupos Espinhaço, Minas e Supergrupo rio das Velhas (Grupo Maquiné).</p>

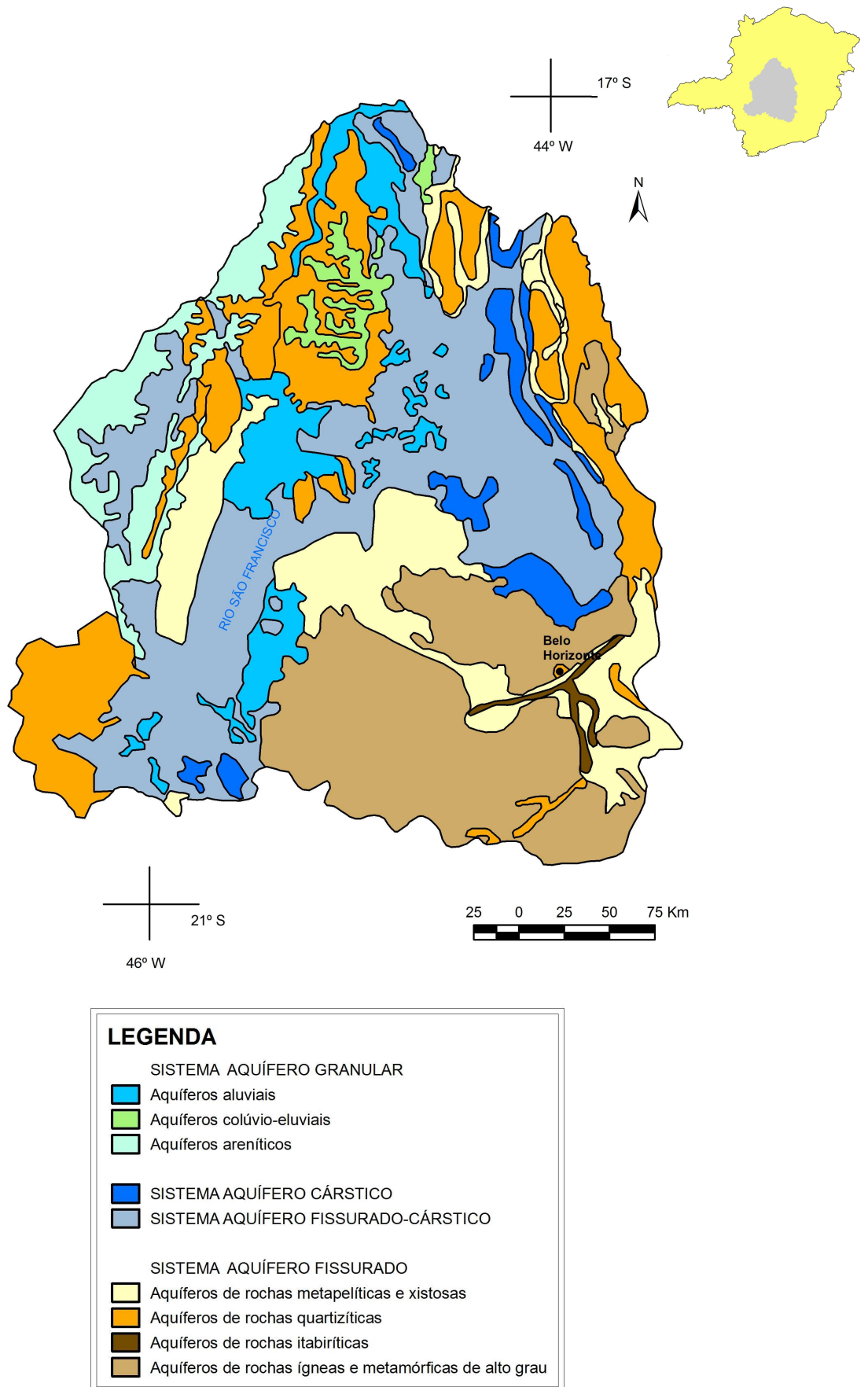
Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

#### Continuação da Tabela 4.4

Aquíferos de rochas itabiríticas	<p>“As rochas itabiríticas correspondem às formações ferríferas bandadas da Formação Cauê (Grupo Itabira do Supergrupo Minas), presentes na região do Quadrilátero Ferrífero. O comportamento hidrogeológico deste litotipo é muito variável e dependente da composição mineralógica e da intensidade de atuação do intemperismo. A forma de interação desses fatores faz com que os aquíferos demonstrem características hidráulicas de sistema fissural ou granular.”</p> <p>“As porções friáveis resultam da dissolução da sílica e / ou carbonato da rocha e formação de um meio altamente poroso e enriquecido em ferro, constituindo o minério comumente explorado.”</p>	Itabiritos do Supergrupo Minas-Grupo Itabira-Formação Cauê.
Aquíferos de rochas ígneas e metamórficas de alto grau	<p>“Aparecem na porção Sul e no limite oriental da bacia como áreas deprimidas de relevo com morfologia de colinas medianamente suavizadas. Por englobar tipos rochosos bastante distintos, essa unidade apresenta grande variabilidade de estruturas de desconuidade e do grau de intemperismo.</p> <p>“O encaixamento de trechos de cursos d’água em sistemas de fraturas e a drenança a partir do manto de cobertura colúvio-eluvionar, que pode em determinados locais ser bastante espesso, promovem a recarga do aquífero. A descarga manifesta-se como fontes pontuais e difusas que brotam de fissuras ou estão condicionadas às zonas de contato do saprolito com a rocha fresca e atuam como reguladores do escoamento superficial.”</p>	Granitos, tonalitos, granitóides, gnaisses charnockíticos, anfíbolitos, granulitos, migmatitos, metavulcânicas, máficas e ultramáficas, metamáficas e meta-ultramáficas associadas aos complexos Barbacena, Mantiqueira. Belo Horizonte. Corpos de rochas ácidas e ultramáficas de idades diversas.

Transcrição adaptada de Mourão *et al.* (2001) e modificado de RURALMINAS (1999) citado por Mourão *et al.*(2001)

As principais características hidráulicas dos sistemas aquíferos, estão descritas na Tabela 4.5.



**Figura 4.3** – Mapa de sistemas aquíferos da bacia do Alto São Francisco em Minas Gerais, baseado no mapa de sistemas aquíferos (Mourão *et al.*, 2001)

**Tabela 4.5** - Principais aspectos hidráulicos dos aquíferos (RURALMINAS, 1999 citado por Mourão *et al.*, 2001)

Unidades Aquíferas	Capacidade específica média (m <sup>3</sup> /h.m)	Transmissividade (m <sup>2</sup> /dia)	Armazenamento específico médio (mm)
<b>Sistema Aquífero Granular</b>			
Aquíferos aluviais	1,64(5)	0,2 a 230	479* <sup>2</sup>
Aquíferos colúvio-eluviais	1,19(8)	33* <sup>1</sup>	113* <sup>2</sup>
Aquíferos areníticos	2,09(21)	58* <sup>1</sup>	278,5-343* <sup>2</sup>
<b>Sistema Aquífero Cárstico</b>			
Aquíferos de rochas carbonáticas	12,74(369)	354* <sup>1</sup>	66,0-458* <sup>2</sup>
<b>Sistema Aquífero Fissurado-Cárstico</b>			
Aquíferos de rochas pelito-carbonáticas	2,23(366)	109* <sup>1</sup>	82,3-385* <sup>2</sup>
<b>Sistema Aquífero Fissurado</b>			
Aquíferos de rochas metapelíticas e xistosas	0,85(22)	233* <sup>1</sup>	79,5-256* <sup>2</sup>
Aquíferos de rochas quartzíticas	1,0(13)	86 a 864	333,5-118* <sup>2</sup>
Aquíferos de rochas itabiríticas	0,44(3)	52 a 3.400	Sem informação
Aquíferos de rochas ígneas e metamórficas de alto grau	0,61(315)	25* <sup>1</sup>	175,5-226* <sup>2</sup>

Valores entre parênteses representam o número de amostras usado para a obtenção da média

\*<sup>1</sup> Valor médio

\*<sup>2</sup> Valor extraído do trabalho de integração, feito por COPASA/Hidrosistemas(1995), para o Estado de Minas Gerais.

## 4.2 Seleção dos dados

As estações utilizadas no presente estudo são as de responsabilidade da *Agência Nacional de Águas (ANA)*. Na Tabela 4.6 são apresentados o número total de estações na sub-bacia 40 e 41, a quantidade de estações em operação e a quantidade de estações desativadas. As 61 estações que estão ativas na atualidade são operadas pela Serviço Geológico do Brasil (CPRM).

**Tabela 4.6** - Total de estações nas sub-bacias 40 e 41

Informações	Total de estações	Estações operando	Estações desativadas
Total de estações fluviométricas nas sub-bacias 40 e 41	446	265	181
Total de estações fluviométricas nas sub-bacias 40 e 41 - responsabilidade ANA	206	61	145

Para a análise de frequência local foram selecionadas as estações fluviométricas que possuíam mais de 10 anos de vazões mínimas observadas. A determinação deste valor foi meramente intuitiva, pois se queria resguardar sobre uma quantidade aceitável de estações para a



regionalização de vazões. Sendo assim foram selecionadas 119 estações fluviométricas com mais de 10 anos de dados observados.

#### 4.2.1 Análise de frequência local

Após a seleção das estações que seriam utilizadas na análise de frequência local, procedeu-se à avaliação da série temporal de vazões mínimas diárias e com a determinação da  $Q_7$  mínima anual. Analisaram-se os períodos de falhas e, para alguns casos, fez-se o preenchimento através de uma regressão linear simples.

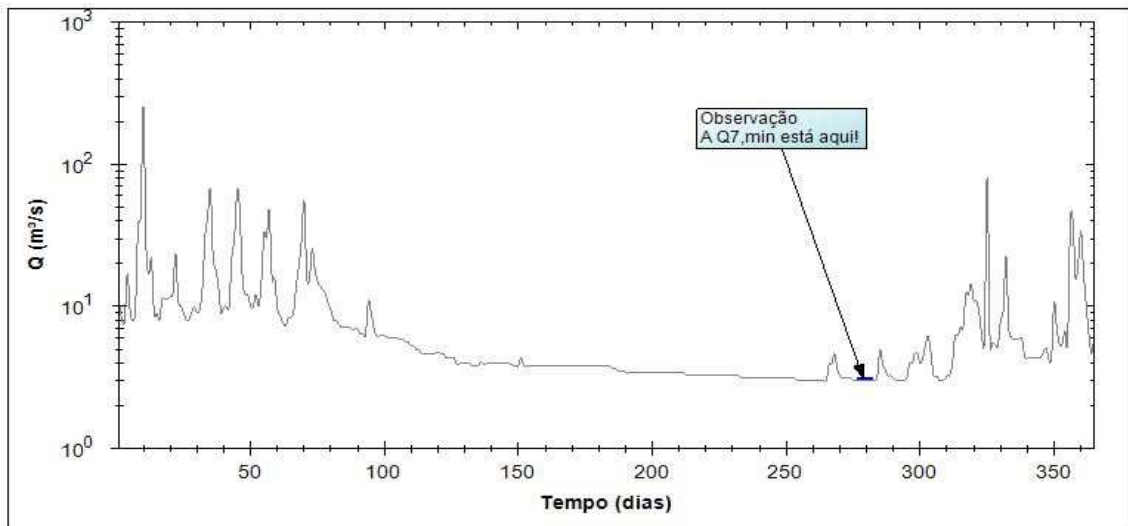
Entretanto, em algumas ocasiões, constatou-se que as falhas ocorriam no segmento de vazões mínimas que determinava o valor mínimo da  $Q_7$  anual. Sendo assim, efetuou-se a eliminação dos anos que continham tais falhas. Tal medida foi adotada a fim de que não houvesse o comprometimento na estimativa da  $Q_{7,10}$ , assim como no modelo de regressão final.

Para a determinação da  $Q_{7,10}$ , construiu-se a curva empírica para as  $Q_7$  mínimas anuais. Neste procedimento, utilizou-se a posição de plotagem de Weibull (Equação 4.1) e a distribuição teórica de probabilidade Weibull de 2 parâmetros.

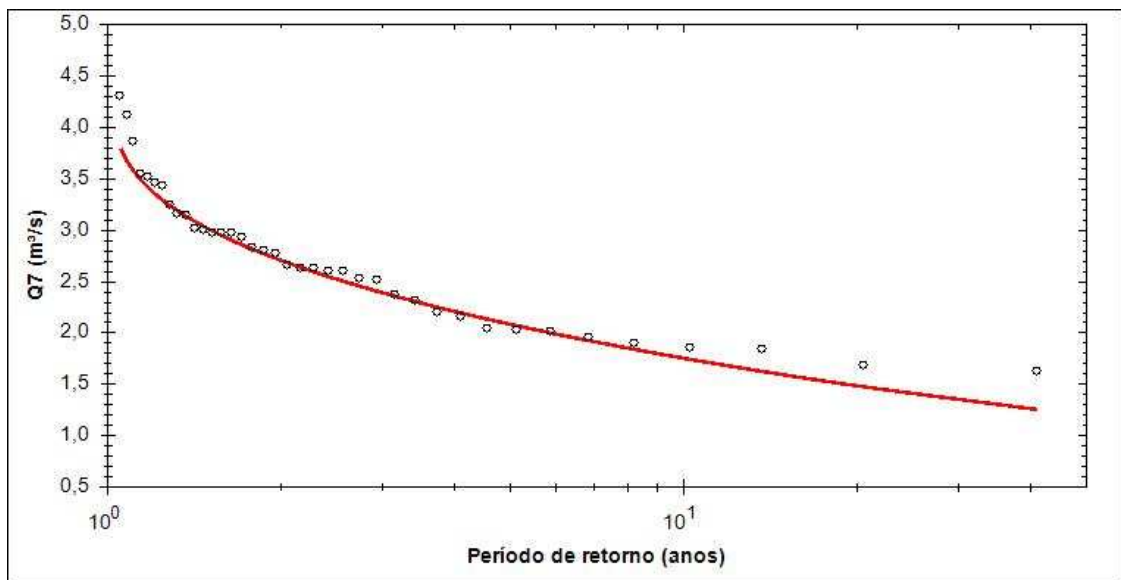
$$pp = \frac{i}{N + 1} \quad (4.1)$$

onde  $i$  é a posição da amostra ordenada de forma crescente e  $N$  é o número de dados.

Nas figuras 4.4 e 4.5 são apresentadas, respectivamente, a série temporal das mínimas anuais e o ajuste da distribuição teórica à curva empírica. Tais figuras referem-se ao ano civil de 1942 da estação fluviométrica 40025000.



**Figura 4.4** - Série temporal do ano de 1942 da estação 40025000



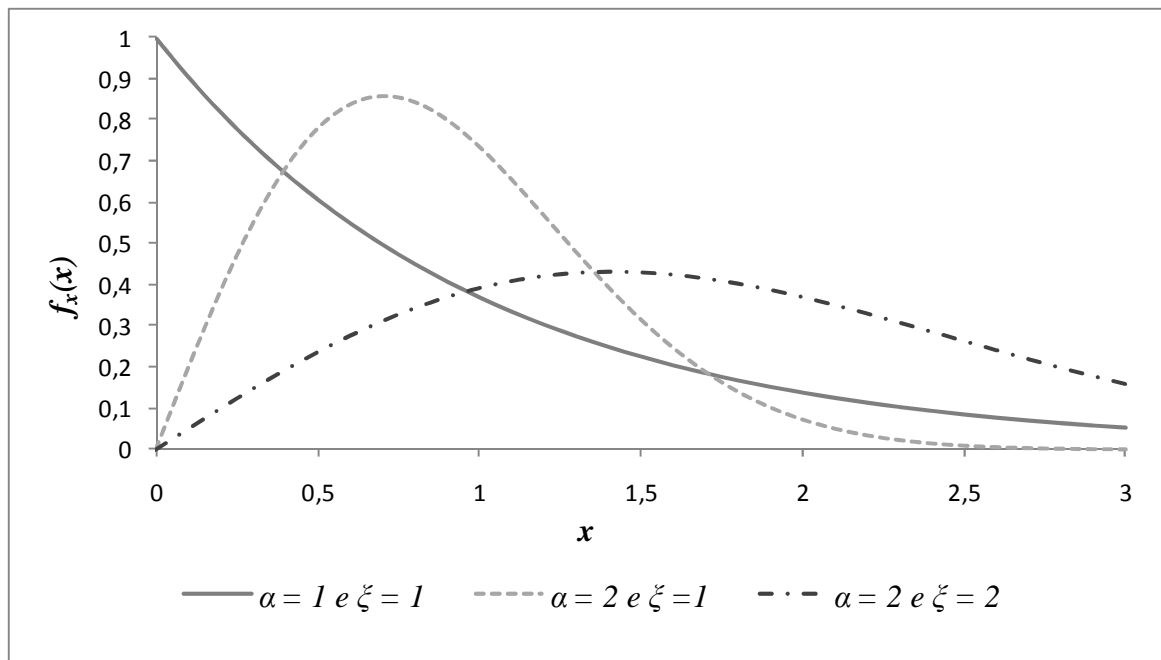
**Figura 4.5** - Ajuste da distribuição teórica (Weibull 2P) a distribuição empírica

Não foram realizados testes estatísticos para confirmar se as séries de vazões eram independentes e estacionárias. Contudo, estas premissas foram admitidas como verdadeiras, pois séries semelhantes foram utilizadas no estudo de regionalização do CPRM (2001), e neste estudo foram realizados testes de independência e estacionariedade.

#### 4.2.1.1 Distribuição de probabilidade Weibull 2 parâmetros

Para a realização da análise de frequência local, determinou-se uma única distribuição de probabilidade a saber, a Weibull de 2 parâmetros (2P) para mínimos. A escolha de tal distribuição é devido ao fato da mesma ter forçosamente um limite inferior igual a zero, como pode ser observado na Figura 4.6. Sendo assim, esta é uma forte candidata na modelagem de eventos hidrológicos mínimos (NAGHETTINI e PINTO, 2007). O modelo que explica a

distribuição Weibull origina-se da teoria assintótica os valores mínimos extremos, especificamente a do tipo III.



**Figura 4.6** - Função densidade da Weibull (2P) mínimos

A função densidade e a função de probabilidades acumuladas da distribuição Weibull (2P) são dadas respectivamente, pelas Equações 4.2 e 4.3 (NAGHETTINI e PINTO, 2007).

$$f_z(z) = \frac{\alpha}{\xi} \left(\frac{z}{\xi}\right)^{\alpha-1} \exp\left[-\left(\frac{z}{\xi}\right)^\alpha\right] \quad (4.2)$$

$$F_z(z) = 1 - \exp\left[-\left(\frac{z}{\xi}\right)^\alpha\right], \text{ para } z \geq 0, \xi \geq 0 \text{ e } \alpha > 0 \quad (4.3)$$

na qual,  $\alpha$  e  $\xi$  são, respectivamente, parâmetros de forma e escala.

Os parâmetros  $\alpha$  e  $\xi$  da distribuição Weibull (2P) foram estimados utilizando-se inicialmente o coeficiente de variação amostral - CV, o qual é resultado do quociente entre o desvio padrão amostral e a média amostral. Em seguida determinou-se o valor de  $\alpha$  por meio da Equação 4.4 (NAGHETTINI e SPERLING, 2007)

$$\alpha = 1,0122 \times CV^{-1,0779} \quad (4.4)$$

Contudo, para proceder aos cálculos do parâmetro  $\xi$ , fez-se o uso da Equação 4.5 (NAGHETTINI e SPERLING, 2007) e o termo  $A(\alpha)$  foi calculado pela Equação 4.6 (NAGHETTINI e SPERLING, 2007).

$$\hat{\xi} = \frac{\bar{Z}}{A(\alpha)} \quad (4.5)$$

$$A(\alpha) = 0,9982 - 0,4419 \times CV + 0,4360 \times CV^2 \quad (4.6)$$

Logo após a determinação dos parâmetros  $\alpha$  e  $\xi$  da distribuição Weibull (2P), calcularam-se os valores dos quantis referentes à  $Q_{7,10}$ , (Equação 4.7) através da função inversa de  $F_z(z)$  (NAGHETTINI e SPERLING, 2007).

$$F_z(z)^{-1} = \xi \left[ -\ln \left( 1 - \frac{1}{T} \right) \right]^{1/\alpha} \quad (4.7)$$

na qual  $T$  é o período de retorno.

#### 4.2.2 Estações fluviométricas selecionadas

Na etapa de análise de frequência local selecionaram-se 90 estações, das quais 59 são da sub-bacia 40 e 31 da sub-bacia 41. As estações referentes às sub-bacias 40 e 41 e os valores das  $Q_{7,10}$  observadas são apresentados na Tabela 4.7.

**Tabela 4.7** - Estações fluviométricas e  $Q_{7,10}$  das sub-bacias 40 e 41

<b>Código</b>	<b>Nome</b>	<b><math>Q_{7,10}</math> observada (m<sup>3</sup>/s)</b>
40025000	VARGEM BONITA	1,269
40032000	FAZENDA SAMBURÁ	5,519
40037000	FAZENDA DA BARRA	3,106
40040000	FAZENDA AJUDAS	0,826
40043001	FAZENDA CAJANGA	1,110
40046000	PORTO SABINO	18,073
40050000	IGUATAMA	20,784
40053000	CALCIOLÂNDIA	0,146
40056002	FAZENDA CAPOEIRÃO	1,020
40056200	MONTANTE DO BOM SUCESSO	1,257
40056500	PONTE CAPOEIRÃO	1,293
40060001	TAPIRAÍ - JUSANTE	3,500
40067000	PONTE OLEGÁRIO MACIEL	24,592
40070000	PONTE DO CHUMBO	33,298
40080000	TAQUARAL	0,447

Continuação da Tabela 4.7

<b>Código</b>	<b>Nome</b>	<b>Q<sub>7,10</sub> observada (m<sup>3</sup>/s)</b>
40100000	PORTO DAS ANDORINHAS	35,977
40102000	PORTO DA BARRA	36,909
40105000	FAZENDA CAMPO GRANDE	1,134
40130000	PONTE DO VILELA	6,685
40150000	CARMO DO CAJURU	10,489
40160000	LAMOUNIER	0,399
40170000	MARILÂNDIA (PONTE BR-494)	1,477
40180000	CARMO DA MATA	0,450
40185000	PARI	2,920
40190002	DIVINÓPOLIS	6,409
40269900	ITAÚNA - MONTANTE	1,700
40300000	JAGUARUNA	2,800
40300001	JAGUARUNA - JUSANTE	3,641
40330000	VELHO DA TAIPA	17,675
40350000	USINA CAMARÃO	0,696
40380000	ARAÚJOS	3,428
40400000	ESTAÇÃO ÁLVARO DA SILVEIRA	3,874
40450000	PORTO PARÁ	36,159
40500000	MARTINHO CAMPOS	1,146
40505000	ABAIXO USINA PARI	0,326
40530000	ABAETÉ	0,032
40549998	SÃO BRÁS DO SUAÇUI - MONTANTE	1,216
40573000	JOAQUIM MURTINHO	1,083
40577000	PONTE JUBILEU	0,966
40579995	CONGONHAS - LINIGRAFO	1,359
40580000	CONGONHAS	2,320
40665000	USINA JOÃO RIBEIRO	0,829
40675000	CACHOEIRA DO GORDO	0,424
40680000	ENTRE RIOS DE MINAS	1,280
40700002	JECEABA	9,104
40710000	BELO VALE	10,069
40720002	MELO FRANCO	13,302
40740000	ALBERTO FLORES	11,529
40770000	CONCEIÇÃO DO ITAGUA	1,541
40790000	BETIM	0,111
40800001	PONTE NOVA DO PARAOPEBA	16,789
40810350	FAZENDA LARANJEIRAS	0,024
40818000	JUATUBA	0,841
40823500	SUZANA	0,316
40830000	FAZENDA ESCOLA FLORESTAL	0,211
40850000	PONTE DA TAQUARA	20,493
40930000	BARRA DO FUNCHAL	1,532
40960000	FAZENDA BOM JARDIM	2,602
40975000	FAZENDA SÃO FÉLIX	1,353
41050000	MAJOR PORTO	2,135
41075001	PORTO DO PASSARINHO	5,761

Continuação da Tabela 4.7

<b>Código</b>	<b>Nome</b>	<b>Q<sub>7,10</sub> observada (m<sup>3</sup>/s)</b>
41151000	FAZENDA ÁGUA LIMPA - JUSANTE	1,037
41160000	GULPIARA	0,670
41180000	ITABIRITO - LINIGRAFO	2,837
41190000	AGUIAR MOREIRA	4,535
41195000	RIO ACIMA	7,746
41199998	HONÓRIO BICALHO - MONTANTE	10,369
41220000	SIDERURGICA	0,740
41250000	VESPASIANO	1,055
41260000	PINHÕES	16,198
41295000	JOSÉ DE MELO	0,880
41300000	TAQUARAÇU	1,306
41340000	PONTE RAUL SOARES	18,512
41380000	PONTE PRETA	0,795
41410000	JEQUITIBA	21,484
41440005	REPRESA - JUSANTE	0,663
41539998	FAZENDA DA CONTAGEM - MONTANTE	1,010
41540000	FAZENDA DA CONTAGEM	0,519
41600000	PIRAPAMA	24,893
41650002	PONTE DO LICÍNIO - JUSANTE	29,861
41685000	PONTE DO PICÃO	0,052
41700001	USINA PARAUNA	1,873
41715000	FAZENDA CACHOEIRA	0,170
41720000	FAZENDA DO CIPÓ	0,469
41780002	PRESIDENTE JUSCELINO - JUSANTE	7,214
41818000	SANTO HIPÓLITO (ANA/CEMIG)	34,687
41880000	BUENÓPOLIS	0,038
41890000	ESTAÇÃO DE CURIMATAI	0,693
41940000	PONTE DO BICUDO	0,142
41990000	VÁRZEA DA PALMA	37,037

A quantidade de anos analisados para cada estação, bem como a relação de anos que foram excluídos na análise de frequência, encontram-se no Apêndice A. Os motivos que determinaram a exclusão de tais anos foram que, alguns deles, continham longos períodos de falhas e o preenchimento das mesmas não foi adequado e/ou as falhas se situavam no trecho que continha a Q<sub>7</sub> mínima anual.

#### **4.2.3 Levantamento das características físicas e climáticas**

No levantamento das características físicas e climáticas, utilizou-se a mesma base de informações desenvolvida no trabalho do CPRM (2001). Contudo, das 90 estações 11 não foram usadas no estudo do CPRM. Sendo assim, fez-se a coleta das características físicas e

climáticas utilizando a base cartográfica disponível pelo CPRM (2001) no formato para *software* MapInfo.

No estudo desenvolvido pelo CPRM (2001), a área de drenagem foi delimitada em mapas na escala de 1:250.000. Contudo, para as estações que possuíam área inferior a 300 km<sup>2</sup> foram utilizadas cartas na escala de 1:100.000 ou 1:50:000. Utilizou-se também estes mapas para elaboração dos perfis longitudinais e, conseqüentemente, para a determinação do comprimento do rio principal.

CPRM (2001) determinou a declividade através do método da declividade equivalente (Equação 4.8). Assim como no método de declividade média, esta leva em consideração todo perfil longitudinal do rio até o ponto desejado, porém considerou-se que este perfil tivesse uma declividade constante.

$$I_{eq} = \left[ \frac{L}{\sum_{k=1}^N \left( \frac{l_k}{\sqrt{d_k}} \right)} \right]^2 \quad (4.8)$$

na qual  $L$  é o comprimento total do curso d'água,  $l_k$  e  $d_k$  são o comprimento e a declividade de cada sub-trecho, respectivamente.

As características físicas e climáticas a serem avaliadas nas regressões foram: área de drenagem ( $A$ ), comprimento do talvegue principal ( $L$ ), declividade equivalente ( $I$ ), precipitação média anual ( $P$ ), coeficiente de escoamento ( $C$ ) e densidade de drenagem ( $Dd$ ), embora se saiba que em estudos de vazões mínimas, seria mais apropriada a incorporação de grandezas hidrogeológicas. Entretanto, tais dados quando existentes não são facilmente disponibilizados.

As características físicas e climáticas para as 90 estações fluviométricas estão dispostas no Apêndice B.

### **4.3 Definição das regiões homogêneas**

A área de estudo foi dividida em 4 regiões supostamente homogêneas (Figura 4.7), não tendo sido realizada nenhuma análise estatística para corroborar que estas regiões realmente foram homogêneas do ponto de vista hidrológico e estatístico. A divisão se justifica pela facilidade e

clareza ao se apresentar os resultados, uma vez que os métodos OLS, WLS e GLS não necessitam dessa premissa, ao contrário do que ocorre no método *index-flood*.

O critério utilizado para divisão dessas regiões foi através da análise das Unidades de Planejamento e Gestão de Recursos Hídricos (UPGRH), utilizadas pelo Instituto Mineiro de Gestão das Águas (IGAM). Sendo assim, as 4 regiões homogêneas são:

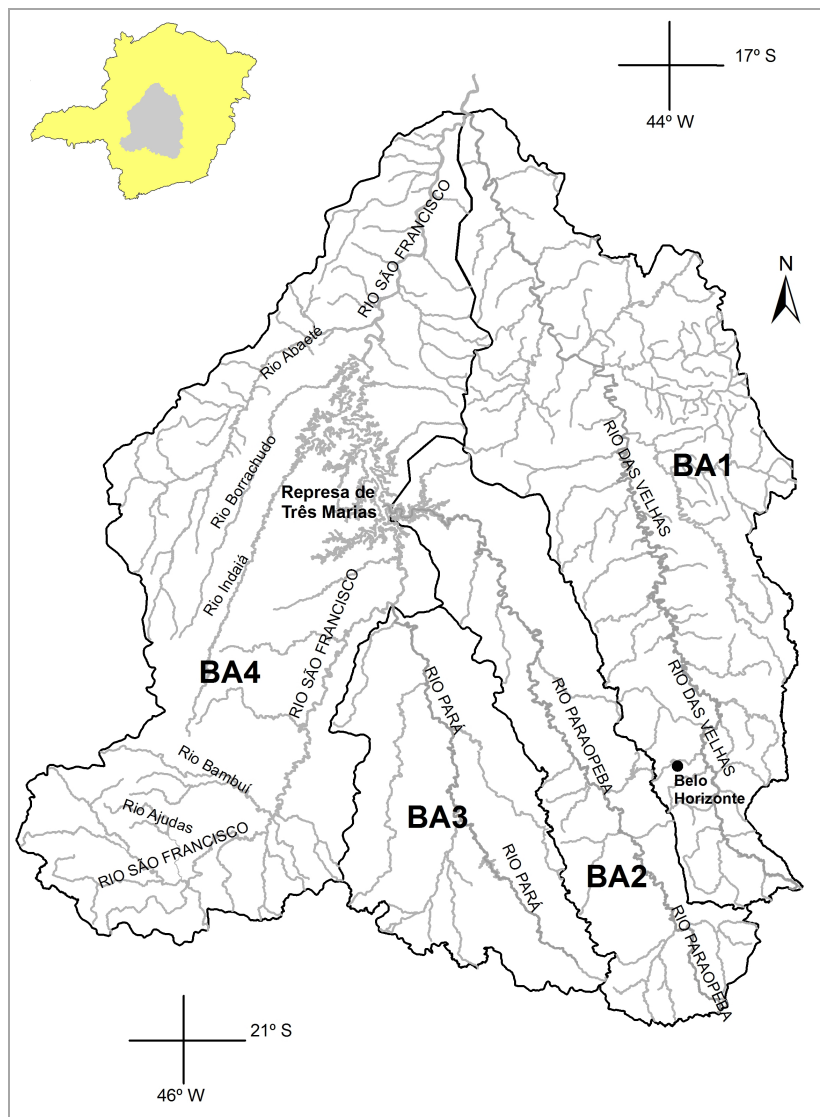
- Bacia do rio das Velhas - BA1;
- Bacia do rio Paraopeba - BA2;
- Bacia do rio Pará - BA3; e
- A união das bacias dos rios Ajudas, Bambuí, Indaiá, Borrachudo, Abaeté e a parcela do rio São Francisco na sub-bacia 40 - BA4.

A Tabela 4.8 contém o número de estações para cada região homogênea, além dos anos de registros mínimos, máximos e médios.

**Tabela 4.8 - Número de estações por região e estatísticas do número de dados**

<b>Regiões</b>	<b>Nº de estações</b>	<b>Nº mínimo de dados</b>	<b>Nº máximo de dados</b>	<b>Nº médio de dados</b>
<b>BA1</b>	29	11	66	33
<b>BA2</b>	20	14	68	29
<b>BA3</b>	18	19	67	35
<b>BA4</b>	23	14	66	32





**Figura 4.7** - Delimitação das regiões homogêneas

No Apêndice D são apresentadas as sub-bacias 40 e 41, assim como a divisão em regiões homogêneas, juntamente com as estações fluviométricas nelas contidas.

#### **4.4 Construção do método GLS Stedinger-Tasker**

Como mencionado anteriormente no item (3.4.1) da revisão de literatura, no modelo GLS Stedinger-Tasker, a matriz de covariância dos erros do modelo é dividida em duas partes, ou seja, o erro associado ao modelo ( $\epsilon$ ) e o erro associado à amostra ( $\omega$ ). Sendo assim, para construção do modelo GLS, é necessário determinar a variância do erro do modelo e a covariância dos erros amostrais ( $\Sigma$ ).

A variância do erro do modelo ( $\delta_{\epsilon}^2$ ) é estimada resolvendo-a iterativamente. Para os elementos de  $i$  diferentes de  $j$  é requerido o cálculo das correlações cruzadas entre os fluxos concorrentes amostrais, ou seja, deve-se estimar uma equação para a covariância dos erros

amostrais. No entanto, assim como relatado por Tasker (1980) e Kuczera (1983), existem dificuldades associadas à determinação desta estimativa.

A matriz  $\Sigma$  é elaborada levando-se em conta a distribuição de probabilidade definida na análise de frequência local (Weibull 2P) e o intervalo de confiança para os quantis. Os estimadores destes quantis podem ser obtidos através de métodos de estimação de parâmetros, como por exemplo, o Método dos Momentos (MOM) e o Método da Máxima Verossimilhança (MVS).

Para o desenvolvimento deste trabalho foram utilizadas as equações de variância e covariância dos erros amostrais para o quantil. Tais equações foram obtidas por meio dos Métodos dos Momentos.

É desconhecida a existência de alguma publicação anterior que tenha proposto uma forma de estimar a matriz de covariância para o caso de um estudo de vazões mínimas empregando a distribuição Weibull 2P. No entanto, Bayazit e Önoz (2004) apresentaram uma forma de cálculo para matriz de covariância entre os coeficientes de variação, a qual se assemelha em partes com a Equação 4.20. Contudo, tais autores assumiram que a covariância entre os momentos amostrais ( $m_1$ ,  $m_2$  e  $m_3$ ) são normalmente distribuída. Sendo assim, os resultados encontrados por esses autores possuem diferenças em relação aos resultados proposto neste trabalho. Logo, uma proposta mais geral para a estimativa da matriz de covariância é a grande contribuição desse estudo, permitindo que os métodos WLS e GLS possam ser aplicados.

Kite (1988) indaga sobre como usar a distribuição de probabilidade selecionada, na análise de frequência. Sendo assim, sugere o uso da equação geral para o quantil de vazão proposta por Chow (1964) (Equação 4.9).

$$\hat{y}_T = \bar{X} + KS \quad (4.9)$$

onde  $y_T$  é a estimativa do quantil associada a um período de retorno,  $\bar{X}$  é a média amostral,  $S$  é o desvio-padrão amostral e  $K$  é o fator de frequência associado à distribuição e ao período de retorno, o qual é calculado para distribuição de probabilidade Weibull, através da Equação 4.10 (RAO e HAMED, 2000).

$$K = \frac{\left[ -\ln\left(1 - \frac{1}{T}\right) \right]^{1/\hat{\alpha}} - \Gamma(1+1/\hat{\alpha})}{\left[ \Gamma(1+2/\hat{\alpha}) - \Gamma^2(1+1/\hat{\alpha}) \right]^{1/2}} \quad (4.10)$$

onde  $\Gamma(\cdot)$  denota a função Gama e  $\hat{\alpha}$  denota a estimativa do parâmetro de forma.

#### 4.4.1 Variância do erro amostral

Segundo Kite (1988), o estimador da variância do erro amostral para a distribuição Weibull 3P é dada pela Equação 4.11. No entanto, para Weibull 2P, utiliza-se somente o termo  $V_1$  desta equação. Isso é devido ao fator de frequência, que depende somente do período de retorno, e parâmetro de forma da distribuição Weibull, que por sua vez depende do coeficiente de variação. Entretanto, a Equação 4.11 foi apresentada em sua forma geral, porque no primeiro instante pensou-se em utilizar a distribuição Weibull 3P, além de estimar o fator de forma através do coeficiente de assimetria. Caso este procedimento tivesse sido adotado, o fator de frequência dependeria da assimetria, e portanto seria avaliado pela Equação 4.11 como um todo.

$$\delta_{\omega ij}^2 = \frac{m_2}{N} \{V_1 + V_2 + V_3\} \quad (4.11)$$

onde  $V_1$ ,  $V_2$  e  $V_3$  são dados, respectivamente, pelas Equações 4.12, 4.13 e 4.14.

$$V_1 = 1 + K\gamma_1 + \frac{K^2}{4}[\gamma_2 - 1] \quad (4.12)$$

$$V_2 = \frac{\partial K}{\partial \gamma_1} \left[ 2\gamma_2 - 3\gamma_1^2 - 6 + K \left( \gamma_3 - \frac{6\gamma_1\gamma_2}{4} - \frac{10\gamma_1}{4} \right) \right] \quad (4.13)$$

$$V_3 = \left( \frac{\partial K}{\partial \gamma_1} \right)^2 \left[ \gamma_4 - 3\gamma_1\gamma_3 - 6\gamma_2 + \frac{9\gamma_1^2\gamma_2}{4} + \frac{35\gamma_1^2}{4} + 9 \right] \quad (4.14)$$

na qual,  $m_2$  é a variância amostral,  $N$ ,  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$  e  $\gamma_4$  correspondem, respectivamente, ao número dados para cada estação, ao coeficiente de assimetria, ao coeficiente de curtose, ao coeficiente de 5ª ordem e ao coeficiente de 6ª ordem populacionais.

A derivada  $\partial K/\partial \gamma_1$  é calculada a partir da Equação 4.15.

$$\frac{\partial K}{\partial \gamma_1} = \frac{\partial K}{\partial(1/\alpha)} \frac{\partial(1/\alpha)}{\partial \gamma_1} \quad (4.15)$$

onde  $\partial K/\partial(1/\alpha)$  é determinado a partir da Equação 4.10 e  $\partial(1/\alpha)/\partial \gamma_1$  pela Equação 4.16.

$$\gamma = [\Gamma(1 + 3/\hat{\alpha}) - 3\Gamma(1 + 2/\hat{\alpha})\Gamma(1 + 1/\hat{\alpha}) + 2\Gamma^3(1 + 1/\hat{\alpha})] \times B_{\hat{\alpha}}^3 \quad (4.16)$$

onde  $B_{\alpha}$  é dado pela Equação 4.17.

$$B_{\alpha} = [\Gamma(1 + 2/\hat{\alpha}) - \Gamma^2(1 + 1/\hat{\alpha})]^{-1/2} \quad (4.17)$$

#### 4.4.2 Formulação da equação de covariância dos erros amostrais

A formulação da equação para a matriz de covariância dos erros amostrais para distribuição Weibull 3P será descrita levando em consideração algumas aproximações, bem como algumas estruturas de cálculos propostas na literatura consultada.

Para elaboração da matriz de covariância dos erros amostrais, fez-se as seguintes proposições:

- As Equações 4.18 e 4.19 referem-se, respectivamente, aos quantis das estações  $i$  e  $j$ ;
- Assim como a variância para um evento  $y_T$ , a covariância para um mesmo evento estimada através do MOM depende dos três primeiros momentos amostrais (média, variância e coeficiente de assimetria);
- A covariância pode ser descrita pela Equação 4.20.

$$\hat{y}_{Ti} = m'_{1i} + \sqrt{m'_{2i}} K_i \quad (4.18)$$

$$\hat{y}_{Tj} = m'_{1j} + \sqrt{m'_{2j}} K_j \quad (4.19)$$

$$\begin{aligned}
\Sigma_{ij} = & \left( \frac{\partial y_{Ti}}{\partial m_{1i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{1j}} \right) \text{cov}(m'_{1i}, m'_{1j}) + \left( \frac{\partial y_{Ti}}{\partial m_{2i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{2j}} \right) \text{cov}(m_{2i}, m_{2j}) + \\
& + \left( \frac{\partial y_{Ti}}{\partial m_{3i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{3j}} \right) \text{cov}(m_{3i}, m_{3j}) + \left( \frac{\partial y_{Ti}}{\partial m_{1i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{2j}} \right) \text{cov}(m'_{1i}, m_{2j}) + \\
& + \left( \frac{\partial y_{Ti}}{\partial m_{2i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{1j}} \right) \text{cov}(m_{2i}, m'_{1j}) + \left( \frac{\partial y_{Ti}}{\partial m_{1i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{3j}} \right) \text{cov}(m'_{1i}, m_{3j}) + \\
& + \left( \frac{\partial y_{Ti}}{\partial m_{3i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{1j}} \right) \text{cov}(m_{3i}, m'_{1j}) + \left( \frac{\partial y_{Ti}}{\partial m_{2i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{3j}} \right) \text{cov}(m_{2i}, m_{3j}) + \\
& + \left( \frac{\partial y_{Ti}}{\partial m_{3i}} \right) \left( \frac{\partial y_{Tj}}{\partial m_{2j}} \right) \text{cov}(m_{3i}, m_{2j})
\end{aligned} \tag{4.20}$$

onde  $m'_1$ ,  $m_2$  e  $m_3$  referem-se, respectivamente, ao momento de primeira ordem em relação à origem (média amostral), ao momento central de segunda ordem (variância amostral) e ao momento central de terceira ordem (coeficiente de assimetria amostral).

#### 4.4.2.1 Derivadas dos momentos amostrais

Para a determinação das derivadas, além das equações 4.18 e 4.19, também deve ser utilizada a função do coeficiente de assimetria amostral para as estações  $i$  e  $j$ , Equação 4.21. As derivadas são calculadas assim como Kite (1988) calculou para a variância do erro amostral.

$$g = \frac{m_3}{\sqrt{m_2^3}} \tag{4.21}$$

As derivadas para cada momento amostral, são dadas pelas Equações 4.22, 4.23, 4.24.

$$\frac{\partial y_{Ti}}{\partial m_{1i}} = 1 \tag{4.22}$$

$$\frac{\partial y_{Ti}}{\partial m_{2i}} = \frac{K_i}{2\sqrt{m_{2i}}} + \sqrt{m_{2i}} \left( \frac{\partial K_i}{\partial m_{2i}} \right) = \frac{K_i}{2\sqrt{m_{2i}}} - \frac{3g_i}{2\sqrt{m_{2i}}} \left( \frac{\partial K_i}{\partial g_i} \right) \tag{4.23}$$

$$\frac{\partial y_{Ti}}{\partial m_{3i}} = \left( \frac{\partial y_{Ti}}{\partial K_i} \right) \times \left( \frac{\partial K_i}{\partial g_i} \right) \times \left( \frac{\partial g_i}{\partial m_{3i}} \right) = \frac{1}{m_{2i}} \left( \frac{\partial K_i}{\partial g_i} \right) \tag{4.24}$$

Os cálculos foram feitos somente para a estação  $i$ , uma vez que para a estação  $j$  os resultados são os mesmos.

#### 4.4.2.2 Covariância entre os momentos amostrais

As covariâncias entre os momentos amostrais foram calculadas mediante auxílio da literatura sobre a teoria de probabilidades (PAPOULIS, 1965; KENDALL e STUART, 1977; DODGE e ROUSSON, 1999; ZHANG, 2007). Esses trabalhos fundamentaram o desenvolvimento das equações que se seguem.

A covariância amostral é dada pela Equação 4.25.

$$Cov(z_i, z_j) = E[z_i z_j] - E[z_i]E[z_j] \quad (4.25)$$

A covariância também pode ser escrita em termos da correlação amostral (Equação 4.26).

$$\rho_{ij} = \frac{Cov(z_i, z_j)}{\sigma_i \sigma_j} \quad (4.26)$$

Sendo assim, a covariância entre os momentos amostrais é dada pelas equações a seguir.

$$Cov(m'_{1i}, m'_{1j}) = E \left[ \left( \sum_{i=1}^{N_i} \frac{z_i}{N_i} \right) \left( \sum_{j=1}^{N_j} \frac{z_j}{N_j} \right) \right] \quad (4.27)$$

$$Cov(m'_{1i}, m'_{1j}) = \frac{1}{N_i N_j} E \left[ \left( \sum_{i=1}^{N_i} z_i \right) \left( \sum_{j=1}^{N_j} z_j \right) \right] \quad (4.28)$$

$$E \left[ \left( \sum_{i=1}^{N_i} z_i \right) \left( \sum_{j=1}^{N_j} z_j \right) \right] = N_{ij} \rho_{ij} \sigma_i \sigma_j \quad (4.29)$$

Logo, substituindo o resultado da Equação 4.29 na 4.28, encontra-se a Equação 4.30.

$$Cov(m'_{1i}, m'_{1j}) = \frac{N_{ij} \rho_{ij} \sigma_i \sigma_j}{N_i N_j} \quad (4.30)$$

onde  $N_{ij}$  é o número de anos em comum (concorrentes) entre as estações  $i$  e  $j$ ,  $\sigma_i$  e  $\sigma_j$  são, respectivamente, os desvios padrões concorrentes das estações  $i$  e  $j$  e  $N_i$  e  $N_j$  são o número de dados, nesta ordem, da estação  $i$  e  $j$ .

$$Cov(m_{2i}, m_{2j}) = \frac{1}{(N_i - 1)} E \left\{ \left[ \sum_{i=1}^{N_i} (z_i - \mu_{1i})^2 - \frac{\sum_{i=1}^{N_i} (z_i - \mu_{1i})^2}{N_i} \right] \times \right. \\ \left. \frac{1}{(N_j - 1)} \left[ \sum_{j=1}^{N_j} (z_j - \mu_{1j})^2 - \frac{\sum_{j=1}^{N_j} (z_j - \mu_{1j})^2}{N_j} \right] \right\} \quad (4.31)$$

Tirando-se o mínimo múltiplo comum dos dois termos da Equação 4.31, têm-se como resultado as Equação 4.32.

$$Cov(m_{2i}, m_{2j}) = \frac{(N_i - 1)}{N_i(N_i - 1)} E \left\{ \left[ \sum_{i=1}^{N_i} (z_i - \mu_{1i})^2 - \frac{\sum_{i=1}^{N_i} (z_i - \mu_{1i})^2}{N_i} \right] \times \right. \\ \left. \frac{(N_j - 1)}{N_j(N_j - 1)} \left[ \sum_{j=1}^{N_j} (z_j - \mu_{1j})^2 - \frac{\sum_{j=1}^{N_j} (z_j - \mu_{1j})^2}{N_j} \right] \right\} \quad (4.32)$$

Resolvendo as multiplicações da Equação 4.32 conjuntamente com análise dos momentos bivariados proposta por Kendall e Stuart (1977), encontra-se a Equação 4.33, a qual pode ser reescrita da forma como apresentado na Equação 4.34.

$$Cov(m_{2i}, m_{2j}) = \frac{N_{ij}}{N_i N_j} (\mu_{22} - \mu_{11}^2) \quad (4.33)$$

$$Cov(m_{2i}, m_{2j}) = \frac{N_{ij}}{N_i N_j} (\mu_{22} - \rho_{ij}^2 \sigma_i^2 \sigma_j^2) \quad (4.34)$$

onde  $\mu_i$  é o momento de primeira ordem populacional e  $\mu_{11}$  e  $\mu_{22}$  são, respectivamente, a covariância amostral (momento bivariado de primeira ordem) e o momento bivariado de segunda ordem.

Nos cálculos das próximas covariâncias serão omitidas as deduções e serão apresentados somente os resultados finais. Os cálculos foram suprimidos porque seguem o mesmo processo construtivo descrito para a Equação 4.31. No entanto, existem diferenças relacionadas às simplificações das relações dos  $N_i$  e  $N_j$  para o estimador não enviesado dos momentos amostrais.

$$Cov(m_{3i}, m_{3j}) = \frac{N_{ij}}{N_i N_j} (\mu_{33} - \mu_{3i} \mu_{3j} - 3\mu_{31} \mu_{2j} - 3\mu_{13} \mu_{2i} + 9\mu_{11} \mu_{2i} \mu_{2j}) \quad (4.35)$$

$$Cov(m'_{1i}, m_{2j}) = \frac{N_{ij}}{N_i N_j} (\mu_{12}) \quad (4.36)$$

$$Cov(m_{2i}, m'_{1j}) = \frac{N_{ij}}{N_i N_j} (\mu_{21}) \quad (4.37)$$

$$Cov(m'_{1i}, m_{3j}) = \frac{N_{ij}}{N_i N_j} (\mu_{13} - 3\mu_{11}\mu_{2j}) \quad (4.38)$$

$$Cov(m_{3i}, m'_{1j}) = \frac{N_{ij}}{N_i N_j} (\mu_{31} - 3\mu_{11}\mu_{2i}) \quad (4.39)$$

$$Cov(m_{2i}, m_{3j}) = \frac{N_{ij}}{N_i N_j} (\mu_{23} - \mu_{12}\mu_{2j} - 3\mu_{21}\mu_{2j}) \quad (4.40)$$

$$Cov(m_{3i}, m_{2j}) = \frac{N_{ij}}{N_i N_j} (\mu_{32} - \mu_{21}\mu_{2i} - 3\mu_{12}\mu_{2i}) \quad (4.41)$$

onde  $\mu_2$  e  $\mu_3$ , são, respectivamente, os momentos de segunda e de terceira ordem populacionais,  $\mu_{12}$ ,  $\mu_{13}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{23}$ ,  $\mu_{32}$  e  $\mu_{33}$  são os momentos bivariados.

Após a determinação das covariâncias entre os momentos amostrais, foram definidos os coeficientes populacionais para os momentos bivariados. É importante salientar que não se criou nenhuma nova relação, elas apenas foram reescritas de forma a incorporar as relações entre os momentos bivariados.

$$\gamma_{2i} + \gamma_{2j} = \frac{2\mu_{22}}{\mu_{11}} \quad (4.42)$$

$$\gamma_{4i} + \gamma_{4j} = \frac{2\mu_{33}}{\mu_{11}\mu_{2i}\mu_{2j}} \quad (4.43)$$

$$\gamma_{2i} + \gamma_{2j} = \frac{2\mu_{31}}{\mu_{11}\mu_{2i}} \quad (4.44)$$

$$\gamma_{2i} + \gamma_{2j} = \frac{2\mu_{13}}{\mu_{11}\mu_{2j}} \quad (4.45)$$



$$\gamma_{1i} \times \gamma_{1j} = \frac{\mu_{3i}\mu_{3j}}{\mu_{11}\mu_{2i}\mu_{2j}} \quad (4.46)$$

$$\gamma_{1j} + \gamma_{1i} = \frac{2\mu_{12}}{\mu_{11}\sqrt{\mu_{2j}}} \quad (4.47)$$

$$\gamma_{1i} + \gamma_{1j} = \frac{2\mu_{21}}{\mu_{11}\sqrt{\mu_{2i}}} \quad (4.48)$$

$$\gamma_{3i} + \gamma_{3j} = \frac{2\mu_{23}}{\mu_{11}\mu_{2j}\sqrt{\mu_{2i}}} \quad (4.49)$$

$$\gamma_{3j} + \gamma_{3i} = \frac{2\mu_{32}}{\mu_{11}\mu_{2i}\sqrt{\mu_{2j}}} \quad (4.50)$$

#### 4.4.2.3 Equação de covariância dos erros amostrais

Após a determinação das derivadas entre os momentos amostrais, das covariâncias amostrais e dos coeficientes populacionais, fez-se a substituição destas na Equação 4.20. Logo, encontrou-se a equação da covariância dos erros amostrais (Equação 4.51).

$$\Sigma_{ij} = \frac{\hat{\rho}_{ij}N_{ij}\sigma_i\sigma_j}{N_iN_j} \left\{ C_1 + \left( \frac{\partial K_i}{\partial \gamma_{1i}} \right) C_2 + \left( \frac{\partial K_j}{\partial \gamma_{1j}} \right) C_3 + \left( \frac{\partial K_i}{\partial \gamma_{1i}} \right) \left( \frac{\partial K_j}{\partial \gamma_{1j}} \right) C_4 \right\} \quad (4.51)$$

na qual  $C_1$ ,  $C_2$ ,  $C_3$  e  $C_4$  são dados respectivamente, pelas Equações 4.52, 4.53, 4.54 e 4.55.

$$C_1 = 1 + \frac{1}{4}K_i(\gamma_i + \gamma_j) + \frac{1}{4}K_j(\gamma_i + \gamma_j) + \frac{1}{8}\rho_{ij}K_iK_j[\gamma_{2i} + \gamma_{2j} - 2] \quad (4.52)$$

$$C_2 = \left[ \frac{\gamma_{2i} + \gamma_{2j}}{2} - 3 - \frac{3}{4}(\gamma_{1i}^2 + \gamma_{1i}\gamma_{1j}) + \frac{K_j}{4}(\gamma_{3i} + \gamma_{3j}) - \frac{3}{8}K_j\gamma_{1i}\rho_{ij}(\gamma_{2i} + \gamma_{2j}) + \frac{3}{4}K_j\gamma_{1i}\rho_{ij} - 2K_j\gamma_{1j} \right] \quad (4.53)$$

$$C_3 = \left[ \frac{\gamma_{2i} + \gamma_{2j}}{2} - 3 - \frac{3}{4}(\gamma_{1j}^2 + \gamma_{1i}\gamma_{1j}) + \frac{K_i}{4}(\gamma_{3i} + \gamma_{3j}) - \frac{3}{8}K_i\gamma_{1j}\rho_{ij}(\gamma_{2i} + \gamma_{2j}) + \frac{3}{4}K_i\gamma_{1j}\rho_{ij} - 2K_i\gamma_{1i} \right] \quad (4.54)$$

$$C_4 = \left[ \frac{\gamma_{4i} + \gamma_{4j}}{2} - \gamma_{1i}\gamma_{1j} - 3(\gamma_{2i} + \gamma_{2j}) + 9 - \frac{3}{4}(\gamma_{1i} + \gamma_{1j})(\gamma_{3i} + \gamma_{3j}) + 6(\gamma_{1i}^2 + \gamma_{1j}^2) + \frac{9}{8}\rho_{ij}\gamma_{1i}\gamma_{1j}(\gamma_{2i} + \gamma_{2j}) - \frac{9}{4}\rho_{ij}\gamma_{1i}\gamma_{1j} \right] \quad (4.55)$$

#### 4.4.3 Correlação cruzada entre os fluxos concorrentes

A correlação cruzada entre as  $Q_7$  concorrentes foi calculada para cada um dos pares de estações. Por conseguinte, determinou-se a combinação simples para 90 estações, ou seja, 90 estações combinadas duas a duas. Os resultados encontrados foram 4005 combinações possíveis para os pares de estações. Desta forma, procedeu-se os cálculos das correlações entre as séries de  $Q_7$  de cada estação. A correlação cruzada,  $\rho_{ij}$ , entre os postos  $i$  e  $j$  foi calculada através da Equação 4.26.

Entretanto, diversos valores da correlações foram iguais a zero e negativos. Sendo assim, seguiu-se as recomendações de Tasker e Stedinger (1989) e fez-se a suavização da correlação amostral utilizando uma função da distância entre os pares de estações.

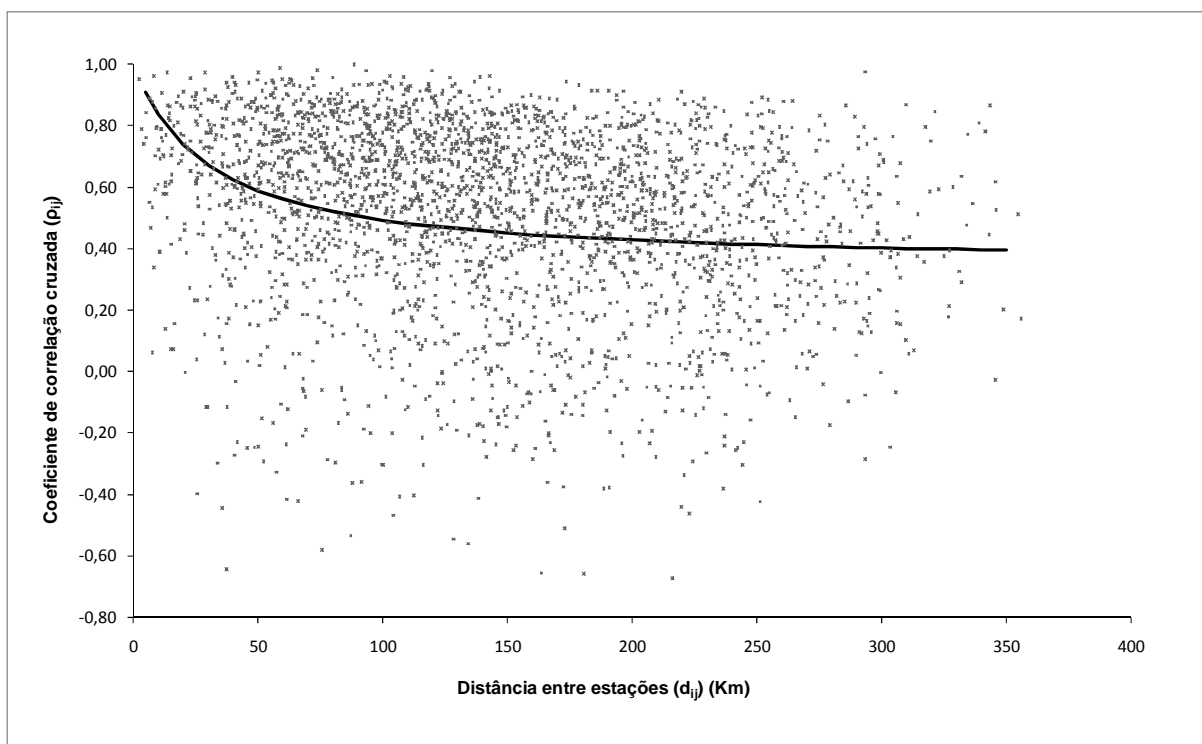
A função de correlação cruzada, Equação 4.56, foi definida após um estudo minucioso de diversas possibilidades de ajustes de funções. O melhor ajuste foi para os pares de estações que tinham no mínimo 5 anos de vazões concorrentes.

$$\hat{\rho}_{ij} = 0,979^{\frac{d_{ij}}{0,020 \times d_{ij} + 1}} \quad (4.56)$$

na qual  $d_{ij}$  é a distância entre os pares de estações, as quais correspondem à distância Euclidiana.

Na Figura 4.8 é apresentado o gráfico contendo a dispersão dos pontos, ou seja, as correlações amostrais para os pares de estações com no mínimo de 5 anos de vazões concorrentes. A razão desta limitação de 5 anos é porque valores inferiores a este acarretam grandes

tendências, uma vez que a correlação entre pequenos números de vazões concorrentes resultariam em 1 ou 0.



**Figura 4.8** - Correlação cruzada das vazões mínimas anuais ( $Q_7$ ) versus pares de estações com no mínimo 5 anos de vazões concorrentes (2831 dados)

É importante salientar que a Figura 4.8 é apenas uma aproximação, sendo que o ajuste de quaisquer outras funções, por exemplo, uma reta, não acarretariam diferenças significativas no valor estimado da correlação cruzada entre os fluxos concorrentes.

#### **4.5 Aplicação dos métodos de regressão**

Os métodos aplicados foram, o OLS tradicional, onde realmente considera a variância do erro do modelo constante, o WLS sistematizado por Tasker e Stedinger (1986), o qual divide a variância do erro total do modelo em duas partes, variância do erro amostral e variância do erro do modelo e, por fim o GLS proposto por Stedinger e Tasker (1985).

Os métodos foram aplicados às 4 regiões homogêneas, sendo que para cada uma destas, construiu-se a matriz de covariância dos erros amostrais. Utilizou-se um algoritmo em linguagem de programação Matlab para análise dos modelos de regressão. O modelo de regressão avaliado foi do tipo potencial. Tal modelo foi escolhido porque em um análise comparativa entre os modelos de regressão polinomial e potencial, o modelo potencial apresentou melhor desempenho, pois no modelo polinomial o erro quadrático médio foi maior. Além disso das 90 estações, 23 tiveram como resultado o quantil de vazão negativo.

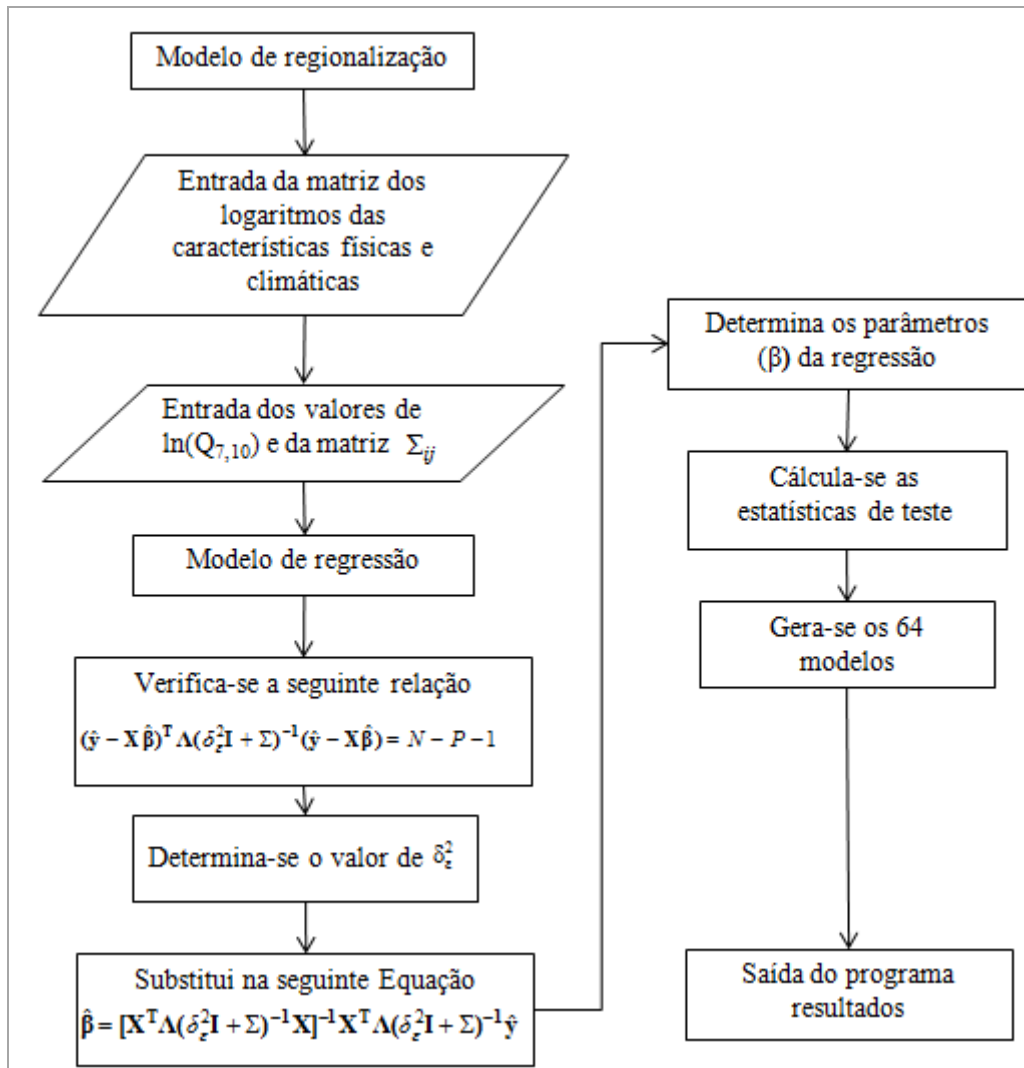
$$\ln(\hat{y}) = \ln \beta_0 + \beta_1 \ln A + \beta_2 \ln L \dots + \beta_p \ln(Dd) + \eta \quad (4.57)$$

#### 4.5.1 Elaboração do modelo final

Para elaboração do modelo final para os métodos WLS e GLS, como já mencionado anteriormente, fez-se o uso de um algoritmo em Matlab. O modelo foi estruturado de acordo com as seguintes etapas:

- Primeiramente, criou-se uma matriz de modelos, a qual continha todas as combinações possíveis de modelos das características físicas e climáticas, juntamente com o intercepto. Sendo assim, fez-se o teste para 6 variáveis fisiográficas, correspondendo a 64 modelos;
- Utilizou-se o Método dos Momentos para determinação do estimador GLS, o qual é definido resolvendo-se, iterativamente, a Equação 4.58, juntamente com equação do erro quadrático médio generalizado (MSE) (Equação 4.59).
- Por fim, determinaram-se os parâmetros  $\beta$  do modelo de regressão.

Os procedimentos descritos acima são sintetizados em um fluxograma, ilustrado na Figura 4.9.



**Figura 4.9** - Etapas para determinação dos modelos de regressão

$$\hat{\beta} = [X^T \Lambda (\delta_\epsilon^2 I + \Sigma)^{-1} X]^{-1} X^T \Lambda (\delta_\epsilon^2 I + \Sigma)^{-1} \hat{y} \quad (4.58)$$

$$(\hat{y} - X\hat{\beta})^T \Lambda (\delta_\epsilon^2 I + \Sigma)^{-1} (\hat{y} - X\hat{\beta}) = N - P - 1 \quad (4.59)$$

onde  $\Lambda$  é a matriz de covariância do erro do modelo, a qual é formada pela soma da variância do erro do modelo e pela variância do erro amostral, e  $I$  a matriz identidade.

A Equação 4.59 é testada de modo a determinar a raiz da função, ou seja, obter um provável valor que possa minimizar a variância do erro do modelo (Equação 4.60).

$$\delta_\epsilon^2 = \frac{(\hat{y} - X\hat{\beta})^T \Lambda (\delta_\epsilon^2 I + \Sigma)^{-1} (\hat{y} - X\hat{\beta})}{N - P - 1} \quad (4.60)$$

onde  $N$  é o número de estações e  $P$  o número das variáveis explicativas.

#### 4.5.1.1 Aplicação do método OLS

No método OLS a variância do erro amostral é igual a zero e a variância do erro do modelo é considerada constante (homocedástica) (Equação 4.61). Sendo assim, a Equação 4.58 reduz-se à Equação 4.62.

$$\delta_\varepsilon^2 = \frac{(\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})^T (\hat{\mathbf{y}} - \mathbf{X}\hat{\boldsymbol{\beta}})}{N - P - 1} \quad (4.61)$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{y}} \quad (4.62)$$

A matriz  $\mathbf{X}$  contém as variáveis explicativas para cada uma das 90 estações, sendo descrita pela Equação 4.63. O vetor  $\hat{\mathbf{y}}$ , que contém as  $Q_{7,10}$ , é descrito pela Equação 4.64.

$$\mathbf{X} = \begin{bmatrix} 1 & \ln(A_1) & \ln(L_1) & \ln(I_1) & \ln(P_1) & \ln(C_1) & \ln(Dd_1) \\ 1 & \ln(A_2) & \ln(L_2) & \ln(I_2) & \ln(P_2) & \ln(C_2) & \ln(Dd_2) \\ 1 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln(A_N) & \ln(L_N) & \ln(I_N) & \ln(P_N) & \ln(C_N) & \ln(Dd_N) \end{bmatrix} \quad (4.63)$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \ln \hat{y}_1 \\ \ln \hat{y}_2 \\ \vdots \\ \ln \hat{y}_N \end{bmatrix} \quad (4.64)$$

#### 4.5.1.2 Aplicação do método WLS

No método WLS, assim como proposto por Tasker e Stedinger (1986), a matriz de covariância é dividida em duas partes (Equação 4.65) (REIS Jr., 2005). Na Equação 4.66 é apresentada, de forma expandida, a matriz de covariância para o método WLS.

$$E(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \begin{cases} \delta_\varepsilon^2 + \delta_{\omega_i}^2 & i = j \\ 0 & i \neq j \end{cases} \quad (4.65)$$

$$E(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \delta_\varepsilon^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \delta_{\omega_1}^2 & 0 & 0 & 0 \\ 0 & \delta_{\omega_2}^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \delta_{\omega_N}^2 \end{bmatrix}$$

$$= \begin{bmatrix} \delta_\varepsilon^2 + \delta_{\omega_1}^2 & 0 & 0 & 0 \\ 0 & \delta_\varepsilon^2 + \delta_{\omega_2}^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \delta_\varepsilon^2 + \delta_{\omega_N}^2 \end{bmatrix} \quad (4.66)$$

#### 4.5.1.3 Aplicação do método GLS

No método GLS, o modelo é construído de forma a incorporar as correlações entre as séries de vazões concorrentes (Equação 4.67) (REIS Jr., 2005). Na Equação 4.68 é apresentada a forma estendida da matriz de covariância da Equação 4.67.

$$\mathbf{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T) = \boldsymbol{\Lambda}(\boldsymbol{\delta}^2) = \begin{cases} \delta_\varepsilon^2 + \delta_{\omega_{ij}}^2 & i = j \\ \delta_{\omega_j} \delta_{\omega_j} \hat{\rho}_{ij} & i \neq j \end{cases} \quad (4.67)$$

$$\begin{aligned} \mathbf{E}(\boldsymbol{\eta}\boldsymbol{\eta}^T) &= \delta_\varepsilon^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \delta_{\omega_1}^2 & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} \\ \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_{22}}^2 & \vdots & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_{NN}}^2 \end{bmatrix} = \\ &= \begin{bmatrix} \delta_{\omega_1}^2 + \delta_\varepsilon^2 & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} \\ \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_{22}}^2 + \delta_\varepsilon^2 & \vdots & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_j} \delta_{\omega_j} \hat{\rho} & \delta_{\omega_{NN}}^2 + \delta_\varepsilon^2 \end{bmatrix} \quad (4.68) \end{aligned}$$

#### 4.5.1.4 Considerações finais para a elaboração do modelo final

Para os métodos WLS e GLS, determinou-se a equação de covariância para cada uma das 4 regiões (Equação 4.51). Considerou-se somente os valores referentes ao termo  $C_I$  e o restante da equação foi desconsiderada. Tal medida foi adotada uma vez que o fator de frequência para distribuição Weibull 2P depende somente do parâmetro de forma. Sendo assim, como o coeficiente de assimetria não foi estimado, não existe a necessidade de avaliar o erro associado a esta estimativa.

Portanto, para a construção da matriz de covariância dos erros amostrais para os métodos GLS, utilizou-se a Equação 4.69. Para  $i$  igual a  $j$  ela se reduz à Equação 4.70 (KITE, 1988), que corresponde à variância do erro amostral para o modelo WLS.

$$\Sigma_{ij} = \frac{\hat{\rho}_{ij} N_i \sigma_i \sigma_j}{N_i N_j} \left\{ 1 + \frac{1}{4} K_i (\gamma_i + \gamma_j) + \frac{1}{4} K_j (\gamma_i + \gamma_j) + \frac{1}{8} \rho_{ij} K_i K_j [\gamma_{2i} + \gamma_{2j} - 2] \right\} \quad (4.69)$$

$$\delta_{\omega_i}^2 = \text{var}(\hat{y}_i) = \frac{m_2}{N} \left\{ 1 + K \gamma_1 + \frac{K^2}{4} [\gamma_2 - 1] \right\} \quad (4.70)$$

A covariância e a variância do erro amostral foram calculadas em relação à série das  $Q_7$  observadas.

#### 4.6 Procedimentos para avaliação dos métodos de regressão

Algumas determinações foram feitas para avaliar e escolher o melhor modelo dentre os 64 modelos calculados, para cada um dos métodos. Fez-se também a análise dos resíduos. Para o método OLS deve-se verificar algumas premissas, como, independência, normalidade, média zero e variância constante (homocedasticidade). Quanto aos métodos WLS e GLS, não existe nenhuma premissa em relação aos resíduos que deve ser verificada. Além disso, realizou-se análise comparativa entre os métodos OLS, WLS e GLS.

##### 4.6.1 Modelos de regressão

A fim de avaliar a precisão dos modelos de regressão, ordenaram-se os 64 modelos das regressões OLS, WLS e GLS em ordem crescente dos valores da variância média de predição (*average variance of prediction* - AVP) (Equação 4.71), selecionou-se os 10 melhores modelos e, em seguida, verificou-se quais destes modelos teriam o menor valor de AVP.

$$AVP = \delta_{\varepsilon}^2 + ASV \quad (4.71)$$

O termo ASV significa a variância média amostral (*average sampling variance* - ASV), dado pela Equação 4.72.

$$ASV = \frac{1}{N} \sum_{i=1}^N x_i [\mathbf{X}^T \mathbf{\Lambda} (\delta_{\varepsilon}^2 + \Sigma)^{-1} \mathbf{X}]^{-1} x_i^T \quad (4.72)$$

De acordo com Tasker e Stedinger (1986), a AVP é uma medida do desempenho médio do modelo regional em estimar os valores de  $Q_{7,10}$  em bacias similares àquelas utilizadas no estudo de regressão. Valores pequenos de AVP correspondem a melhores modelos de regressão (Reis Jr. *et al.*, 2005).



Calculou-se também o  $R^2$  (Equação 4.73) e o  $R^2$  ajustado (Equação 4.74) para os modelos de regressão.

$$R^2 = 1 - \frac{SSE}{SST} \quad (4.73)$$

$$R^2_{ajustado} = 1 - \left( \frac{N-1}{N-P} \right) \frac{SSE}{SST} \quad (4.74)$$

onde SSE é a soma dos quadrados dos resíduos, e SST é a soma quadrática total. O termo SST é representado pela Equação 4.74, a qual é reescrita de forma expandida na Equação 4.75.

$$SST = SSE + SSR \quad (4.74)$$

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 \quad (4.75)$$

onde SSR é a soma dos quadrados da regressão.

No entanto, para os métodos GLS e WLS, a Equação 4.73 não fornece a quantidade desejada, porque ambos os termos SSE e SST levam em conta tanto os erros amostrais como o erro do modelo. Sendo assim, Reis Jr. *et al.* (2004) apresentaram um estimador natural para estes métodos, que descreve a soma do erro quadráticos do modelo como sendo  $N\delta_\varepsilon^2$ . Logo, o estimador que corresponde à real variação de  $y_i$ , é estimado pela Equação 4.76.

$$SST = SSE + SSR = N\delta_\varepsilon^2 + \sum_{i=1}^N (\hat{y}_i - \hat{y}_0)^2 \quad (4.76)$$

Portanto, o  $R^2$  e o  $R^2$  ajustado foram calculados levando em consideração a alteração na Equação 4.76, a qual é apresentada na Equação (4.77)

$$R^2 = 1 - \frac{N\delta_\varepsilon^2}{N\delta_\varepsilon^2 + \sum_{i=1}^N (\hat{y}_i - \hat{y}_0)^2} \quad (4.77)$$

Onde  $\hat{y}_0$  é o valor estimado para a constante do modelo.

## 4.6.2 Resíduos da regressão

Os resíduos dos modelos de regressão foram avaliados segundo alguns critérios. Para os métodos WLS e GLS, plotou-se somente os resíduos, espacialmente, para cada uma das 4 regiões. Contudo para o método OLS, além da análise gráfica, realizou-se alguns testes de hipóteses, como, aderência, média igual zero, independência e variância constante, todos ao nível de significância de 5%. Os testes realizados foram: análise gráfica dos resíduos, teste de aderência de Anderson-Darling, teste paramétrico sobre a média de uma única população normal, teste da independência dos resíduos de Durbin-Watson e testes da variância constante de Breusch-Pagan e White.

Para todos os testes de hipóteses realizados, determinou-se o p-valor, o qual pode ser explicado como o menor valor do nível de significância para o qual rejeita-se  $H_0$ . Logo, se o nível de significância for menor que o p-valor, a hipótese nula não é rejeitada.

### 4.6.2.1 Teste de aderência de Anderson-Darling

O teste de aderência de Anderson-Darling é um teste não paramétrico e baseia-se na diferenças entre as funções de probabilidades acumuladas, empírica,  $F_N(x)$  e teórica  $F_x(x)$ . A distribuição teórica a ser avaliada é a distribuição Normal. Este teste dá mais pesos às caudas inferior e superior (NAGHETTINI e PINTO, 2007). O teste para avaliar a adequação da distribuição tem as seguintes hipóteses:

$H_0$ : a amostra tem distribuição de  $F_x(x)$ .

$H_1$ : a amostra não tem distribuição  $F_x(x)$ .

A estatística de testes proposta por Anderson-Darling é dada pela Equação 4.78.

$$A^2 = \int_{-\infty}^{+\infty} \frac{[F_N(x) - F_x(x)]^2}{F_x(x)[1 - F_x(x)]} f_x(x) dz \quad (4.78)$$

A estatística de teste  $A^2$  pode ser colocada em forma equivalente (Equação 4.79).

$$A_{eq}^2 = -N - \sum_{i=1}^N \frac{(2i-1) \{ \ln F_x(x_i) + \ln [1 - F_x(x_{N-i+1})] \}}{N} \quad (4.79)$$

Para avaliar a hipótese nula da estatística de teste de Anderson-Darling, determinou-se o p-valor através da interpolação da estatística equivalente de Anderson-Darling. Comparando os resultados, se  $A_{eq}^2 < A_{crit,\alpha}^2$ , a decisão é de não rejeitar a hipótese nula. Para a distribuição Normal ao nível de significância de 5%,  $A_{crit,0,05}^2 = 0,752$ . (D'AGOSTINO e STEPHENS, 1986).

#### 4.6.2.2 Teste paramétrico sobre a média

Este teste foi aplicado sobre a média, onde considera-se que os resíduos foram extraídos de uma única população normal. O teste aplicado foi bilateral, onde considerou-se as hipóteses nula e alternativa Equação 4.80.

$$\begin{aligned} H_0 : \mu &= \mu_0 = 0 \\ H_1 : \mu &\neq \mu_0 \end{aligned} \tag{4.80}$$

A estatística de teste é dada pela Equação 4.81, com atributo  $\sigma^2$ : desconhecida e estimada por  $s_z^2$ .

$$T = \frac{\bar{Z} - \mu_0}{s / \sqrt{N}} \tag{4.81}$$

Determinou-se o p-valor através da Equação 4.82 e o intervalo de confiança (Equação 4.83).

$$p - \text{valor} = 2P[t_{\alpha/2} > T] \tag{4.82}$$

$$IC(\mu)_{1-\alpha} = \bar{Z} - t_{\alpha/2} \frac{s}{\sqrt{N}}; \mu; \bar{Z} + t_{\alpha/2} \frac{s}{\sqrt{N}} \tag{4.83}$$

na qual t de Student possui  $N-1$  graus de liberdade,  $\alpha$  é nível de significância e s é o desvio-padrão amostral.

#### 4.6.2.3 Teste de Durbin-Watson

No teste de Durbin-Watson avalia-se a autocorrelação entre os resíduos, ou seja, se estes são independentes ou não. As hipóteses nula ( $H_0$ : os resíduos são independentes) e alternativa ( $H_1$ ) são dadas pela Equação 4.84.

$$\begin{aligned}
 H_0 : \rho &= 0 \\
 H_1 : \rho &> 0
 \end{aligned}
 \tag{4.84}$$

A estatística de teste é dada pela Equação 4.85. Após a determinação da estatística de teste, determina-se os valores de  $d_L$  e  $d_U$  críticos, os quais foram tabelados por Durbin-Watson. Os valores de  $d_L$  e  $d_U$  críticos foram encontrados consultado a tabela de Durbin-Watson disponível em Gujarati e Porter (2008).

$$d_w = \frac{\sum_{i=2}^N (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^N \varepsilon_i^2}
 \tag{4.85}$$

A estatística de teste foi avaliada segundo a análise da Figura 4.10, onde também determinou-se o p-valor.



**Figura 4.10** - Estatística de teste Durbin-Watson.  
 Fonte: Gujarati e Porter (2008)

#### 4.6.2.4 Teste de Breusch-Pagan

O teste de Breusch-Pagan é um teste que avalia a homocedasticidade dos resíduos do método OLS. O teste de Breusch-Pagan consiste em realizar a regressão do quadrado dos resíduos do OLS (WOOLDRIDGE, 2009).

Para análise deste teste estimou-se uma regressão auxiliar referente ao quadrado dos resíduos dada pela Equação 4.86. As variáveis inclusas na regressão auxiliar são as mesmas utilizadas no modelo de regressão OLS que teve o melhor desempenho em relação a AVP.

$$\hat{u}^2 = \beta_0^* + \beta_1^* x_1 + \beta_2^* x_2 + \dots + \beta_p^* x_p + \varepsilon_u
 \tag{4.86}$$

Então, testou-se a significância da regressão auxiliar sob as hipóteses nula ( $H_0$ : os resíduos são homocedásticos) e alternativa ( $H_1$ ), Equação 4.87.

$$\begin{aligned} H_0 : \beta_1^* = \beta_2^* = \dots = \beta_p^* = 0 \\ H_1 : \exists \beta_i^* \neq 0, i = 1, 2, \dots, P \end{aligned} \quad (4.87)$$

Na análise da estatística de teste (Equação 4.86) utilizou-se um teste LM (baseado nos multiplicadores de Lagrange). O teste realizado foi o teste do Qui-Quadrado, o qual é dado pela Equação 4.88. O p-valor foi determinado por meio da Equação 4.89.

$$O = N \times R_{\hat{u}^2}^2 \sim \chi_N^2 \quad (4.88)$$

$$p - \text{valor} = P(\chi^2 > O) \quad (4.89)$$

onde  $R_{\hat{u}^2}^2$  é coeficiente de determinação da regressão auxiliar.

#### 4.6.2.5 Teste de White

No teste de White, assim como no teste de Breusch-Pagan, determina-se uma regressão auxiliar associada aos resíduos. Contudo, o teste de White parte da ideia de que não se conhece o padrão de heterocedasticidade (WOOLDRIDGE, 2009). Sendo assim, inclui-se na regressão auxiliar a variável dependente ( $\hat{y}$ ) e o quadrado da mesma ( $\hat{y}^2$ ).

A regressão auxiliar utilizada é descrita pela Equação 4.90. Tal como no método de Breusch-Pagan, determinou-se os resíduos da regressão OLS e os elevou ao quadrado.

$$\hat{u}^2 = \beta_0^{**} + \beta_1^{**} \hat{y} + \beta_2^{**} \hat{y}^2 + \varepsilon_u \quad (4.90)$$

A significância da regressão auxiliar foi testada sob as hipóteses nula ( $H_0$ : os resíduos são homocedásticos) e alternativa ( $H_1$ ), Equação 4.91.

$$\begin{aligned} H_0 : \beta_1^{**} = \beta_2^{**} = 0 \\ H_1 : \beta_1^{**} = 0 \vee \beta_2^{**} = 0 \end{aligned} \quad (4.91)$$

A avaliação da estatística de teste é similar às utilizadas no método de Breusch-Pagan (Equação 4.88).

### 4.6.3 Análise comparativa entre os métodos OLS, WLS e GLS

Além de utilizar as medidas para avaliar o desempenho do modelo (tópico 4.6.1), utilizou se também medidas para avaliar se uma regressão GLS é necessária ou se, caso contrário, uma regressão WLS ou OLS já seria suficiente. Portanto de acordo com Griffis e Stedinger (2007), a regressão OLS deve ser suficiente no caso da variância do erro amostral ser insignificante em relação à variância do erro do modelo. Então, tais autores sugerem que seja calculada a razão entre as variâncias do erro (*error variance ratio* - EVR), Equação 4.92.

$$EVR = \frac{tr(\Sigma)}{N\delta_{\varepsilon}^2(P)} \quad (4.92)$$

na qual o numerador é o traço da matriz da variância do erro amostral e o denominador corresponde à multiplicação do número de estações ( $N$ ) pela variância do erro do modelo, para um modelo com  $P$  variáveis explicativas.

Portanto, se a variância do erro amostral for insignificante quando comparada com a variância do erro do modelo, o modelo OLS é suficiente. Segundo Griffis e Stedinger (2007), se a EVR for maior que 20%, o modelo WLS ou GLS é recomendado em detrimento ao OLS.

Da mesma forma, a razão de inflação de variâncias (*variance inflation ratio* - VIR) (Equação 4.93) pode ser utilizada para determinar se uma regressão GLS é necessária, ou se a regressão WLS é suficiente (Griffis e Stedinger, 2007).

$$VIR = \frac{w^T \Lambda(\delta_{\varepsilon}^2 + \Sigma)w}{N} \quad (4.93)$$

na qual  $w$  é um vetor ( $N \times 1$ ), dado pela Equação (4.94).

$$w = \frac{1}{\sqrt{\Lambda_{ii}}} \quad (4.94)$$

A VIR fornece uma medida direta da distorção da variância estimada para a constante do modelo regressão WLS ( $\beta_0$ ), quando este é utilizado em vez de uma análise GLS. De acordo com Griffis e Stedinger (2007), se a VIR for significativamente maior do que 1, então a regressão WLS é insuficiente e, o modelo de regressão GLS deve ser adotado. Decidiu-se em

utilizar a VIR, porque segundo Stedinger e Tasker (1985), a covariância entre os estimadores dos valores observados geralmente tem maior impacto sobre a precisão do termo constante.

Realizou-se também uma análise comparativa entre os métodos de regressão quanto ao ajuste linear. Logo, calculou-se o erro relativo médio (ERM), em termos de valores absolutos, Equação 4.95.

$$ERM = \frac{\sum_{i=1}^N \left| \frac{\hat{y}_{observado} - \hat{y}}{\hat{y}_{observado}} \right|}{N} \quad (4.95)$$

Calculou-se os valores de EVR para os modelos estimados através dos métodos WLS e GLS e a VIR para o modelo GLS. Os valores ERM foram calculados para todos os modelos de regressão.

## 5 RESULTADOS

Os resultados para os modelos OLS, WLS e GLS serão descritos separadamente para cada uma das 4 regiões homogêneas. Na aplicação dos métodos foram selecionados os 10 modelos que possuíam os menores valores de AVP.

Para facilitar a análise dos resultados a Equação 4.57 foi reescrita na Equação 5.1, a qual incorpora todas as variáveis explicativas. Decidiu-se reescrevê-la para facilitar a compreensão dos resultados.

$$\ln(\hat{Q}_{7,10}) = \ln \beta_0 + \beta_1 \ln(A) + \beta_2 \ln(L) + \beta_3 \ln(I) + \beta_4 \ln(P) + \beta_5 \ln(C) + \beta_6 \ln(Dd) + \eta \quad (5.1)$$

A utilização da Equação 5.1 resulta numa dificuldade adicional, pois a metodologia para estimativa da matriz de covariância é baseada nos quantis propriamente ditos, e não no logaritmo dos quantis. Essa transformação não linear certamente causa uma mudança nos elementos da matriz de covariância. Porém, não foi possível, até esse momento, avaliar as repercussões desse fato nos resultados finais do estudo de regressão.

Os modelos escolhidos foram delimitados com uma demarcação em vermelho.

### 5.1 Resultados do método OLS

#### 5.1.1 Seleção dos modelos de regressão

Para as regiões BA1 (Tabela 5.1) e BA2 (Tabela 5.2), regiões BA3 (Tabela 5.3) e BA4 (Tabela 5.4), selecionou-se o modelo de número 1. Em todos os modelos, analisou-se o valor da AVP. Conjuntamente com análise da AVP avaliou-se a variância do erro do modelo ( $\sigma^2$ ). De modo geral, o modelo 1 apresentou menor erro e menor AVP. Os  $R^2$  (coeficiente de determinação) encontrados possuem valores acima de 0,90. Então, em relação ao ajuste linear, os modelos estimados apresentam um bom ajuste.



**Tabela 5.1 - Modelos de regressão do método OLS região BA1**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_\varepsilon^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-8,377	1,090	-	-	8,381	-	0,603	0,328	0,045	0,373	0,938
<b>2</b>	-9,014	1,165	-	0,142	8,209	-	0,576	0,337	0,058	0,395	0,938
<b>3</b>	-8,207	1,092	-	-	8,173	0,121	0,599	0,341	0,059	0,400	0,938
<b>4</b>	-8,390	1,114	-0,039	-	8,410	-	0,601	0,342	0,059	0,401	0,938
<b>5</b>	-10,454	1,112	-	-	10,637	-	-	0,401	0,042	0,443	0,921
<b>6</b>	-11,467	1,250	-	0,264	10,126	-	-	0,402	0,055	0,457	0,924
<b>7</b>	-10,134	1,116	-	-	10,248	0,208	-	0,415	0,057	0,472	0,921
<b>8</b>	-10,481	1,225	-0,180	-	10,736	-	-	0,416	0,057	0,473	0,921
<b>9</b>	-7,349	-	1,646	-	6,965	-	0,707	0,484	0,067	0,551	0,908
<b>10</b>	-9,730	-	1,678	-	9,578	-	-	0,584	0,060	0,645	0,884

**Tabela 5.2 - Modelos de regressão do método OLS região BA2**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_\varepsilon^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-7,379	1,283	-	0,555	-	1,710	-0,432	0,066	0,017	0,083	0,985
<b>2</b>	-7,402	1,283	-	0,551	0,184	1,746	-0,432	0,071	0,021	0,092	0,985
<b>3</b>	-7,375	1,286	-0,007	0,553	-	1,707	-0,432	0,071	0,021	0,092	0,985
<b>4</b>	-5,051	1,000	-	-	-	1,637	-0,424	0,102	0,020	0,122	0,976
<b>5</b>	-7,027	1,337	-	0,544	-	1,565	-	0,104	0,021	0,124	0,975
<b>6</b>	-9,083	1,289	-	0,525	-	-	-0,388	0,112	0,022	0,134	0,974
<b>7</b>	-4,750	1,059	-	-	-	1,496	-	0,133	0,020	0,153	0,966
<b>8</b>	-8,633	1,338	-	0,518	-	-	-	0,137	0,021	0,158	0,966
<b>9</b>	-6,806	1,021	-	-	-	-	-0,382	0,139	0,021	0,159	0,965
<b>10</b>	-6,395	1,072	-	-	-	-	-	0,160	0,016	0,176	0,957

**Tabela 5.3 - Modelos de regressão do método OLS região BA3**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_\varepsilon^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-7,462	0,846	0,671	-	5,571	2,295	-	0,186	0,052	0,238	0,925
<b>2</b>	-7,306	1,204	-	-	5,305	1,994	-	0,211	0,047	0,258	0,909
<b>3</b>	-4,453	0,879	0,629	-	-	3,127	-	0,223	0,050	0,273	0,904
<b>4</b>	-4,442	1,214	-	-	-	2,808	-	0,239	0,040	0,279	0,889
<b>5</b>	-4,364	1,272	-	-	-	2,550	0,355	0,235	0,052	0,287	0,899
<b>6</b>	-9,718	1,079	-	-	8,140	-	-	0,252	0,042	0,294	0,883
<b>7</b>	-8,579	1,311	-	0,618	-	-	-	0,312	0,052	0,364	0,856
<b>8</b>	-5,679	1,118	-	-	-	-	0,556	0,329	0,055	0,384	0,848
<b>9</b>	-6,027	0,996	-	-	-	-	-	0,356	0,040	0,396	0,824
<b>10</b>	-5,694	-	1,539	-	-	-	-	0,492	0,055	0,546	0,757

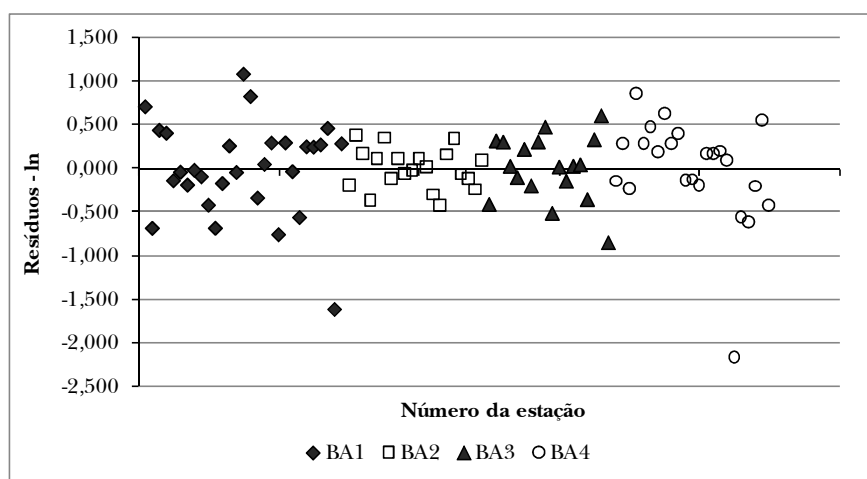
**Tabela 5.4 - Modelos de regressão do método OLS região BA4**

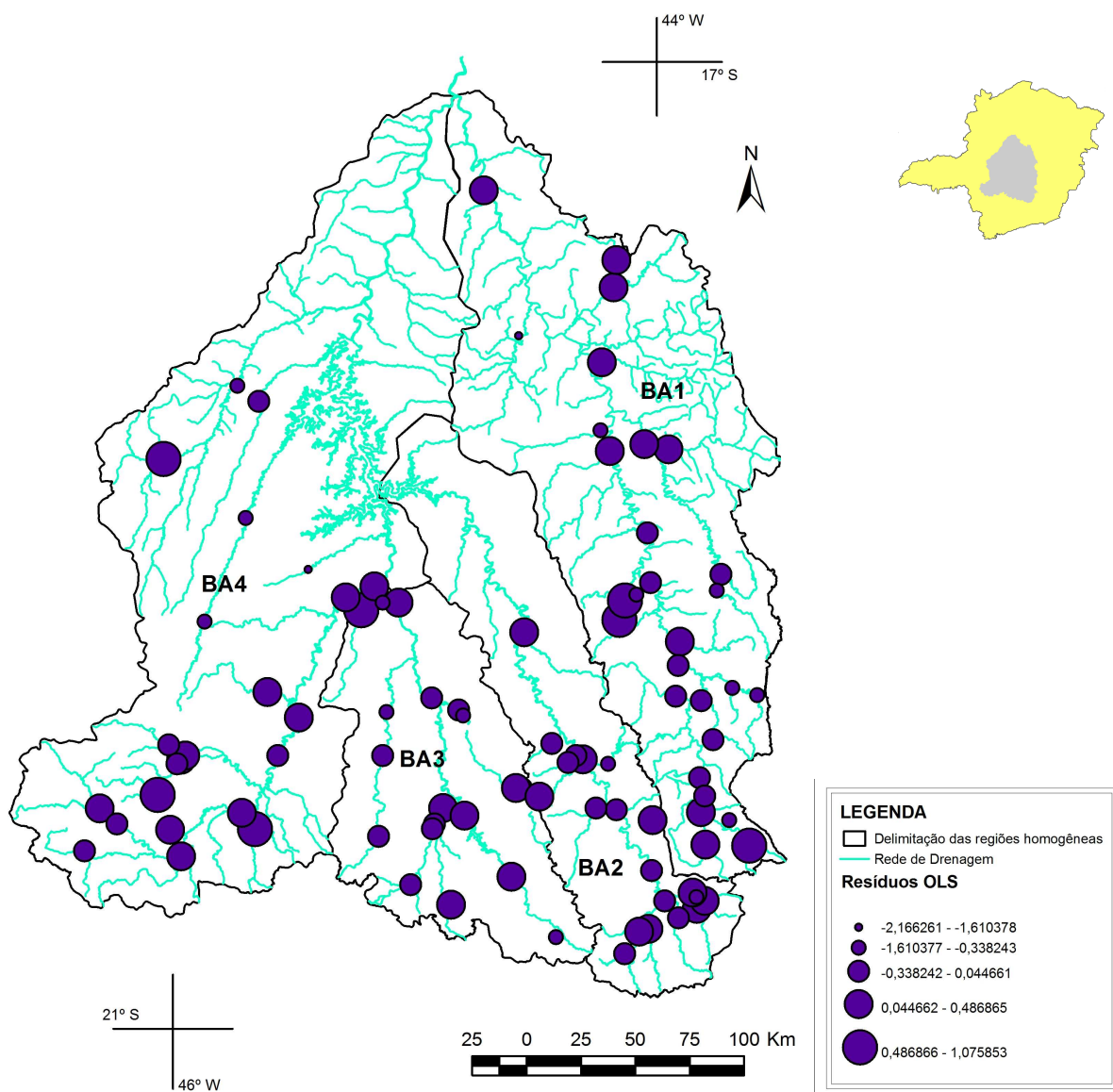
Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_\epsilon^2$	ASV	AVP	R <sup>2</sup>
1	-10,128	1,254	-	-	9,701	2,177	-	0,423	0,074	0,497	0,906
2	-9,771	1,546	-0,502	-	9,535	2,287	-	0,426	0,093	0,519	0,910
3	-9,900	1,257	-	-	9,732	2,083	0,141	0,444	0,096	0,540	0,906
4	-10,153	1,266	-	0,018	9,594	2,198	-	0,447	0,097	0,544	0,906
5	-13,111	1,197	-	-	12,838	-	-	0,489	0,064	0,553	0,885
6	-12,318	1,210	-	-	12,605	-	0,313	0,499	0,087	0,586	0,889
7	-12,955	1,416	-0,379	-	12,832	-	-	0,503	0,088	0,591	0,888
8	-12,451	1,053	-	-0,225	13,814	-	-	0,505	0,088	0,592	0,888
9	-4,273	1,262	-	-	-	4,026	-	0,630	0,082	0,712	0,852
10	-7,051	1,125	-	-	-	-	-	0,989	0,086	1,075	0,756

### 5.1.2 Resíduos dos modelos do OLS

Em algumas análises de resíduos utilizou-se um aplicativo em linguagem R de programação. Os resíduos foram calculados através da subtração do logaritmo da  $Q_{7,10}$  observada pelo logaritmo da  $Q_{7,10}$  estimada. As estatísticas de testes foram avaliadas ao nível de significância de 5% e os resultados referentes à não rejeição (NR- $H_0$ ) ou à rejeição (R- $H_0$ ) da hipótese nula ( $H_0$ ), estão discriminados abaixo.

Primeiramente, plotou-se os resíduos dos modelos ajustados em um mapa (Figura 5.2). Analisando a dispersão dos pontos, percebe-se que, aparentemente, não existe nenhuma autocorrelação entre os resíduos, ou seja, estes parecem ser independentes. Contudo, tal afirmação só pode ser confirmada após um teste de hipótese que avalie a independência dos mesmos. Nas regiões (Figura 5.1), BA1, BA3 e BA4 é perceptível a presença de possíveis *outliers*, o mesmo ocorre nos modelos WLS e GLS, indicando a necessidade de avaliação dos pontos influentes dos modelos de regressão. Esta avaliação pode ser feita através de medidas de influência, tais como *leverage*, Cook's D, DFITS e DFBETAS (Veilleux, 2011).

**Figura 5.1 – Resíduos dos modelos OLS**



**Figura 5.2** – Mapa dos resíduos dos modelos OLS

### 5.1.2.1 Análise do teste de Anderson-Darling

A primeiro teste realizado foi o teste de aderência de Anderson-Darling para distribuição Normal. Determinou-se a estatística de testes equivalentes ( $A_{eq}^2$ ) e o p-valor (Tabela 5.5). Além disso, os valores dos resíduos foram plotados em função da variável padronizada ( $Z$ ) no papel de probabilidade normal (Figura 5.3).

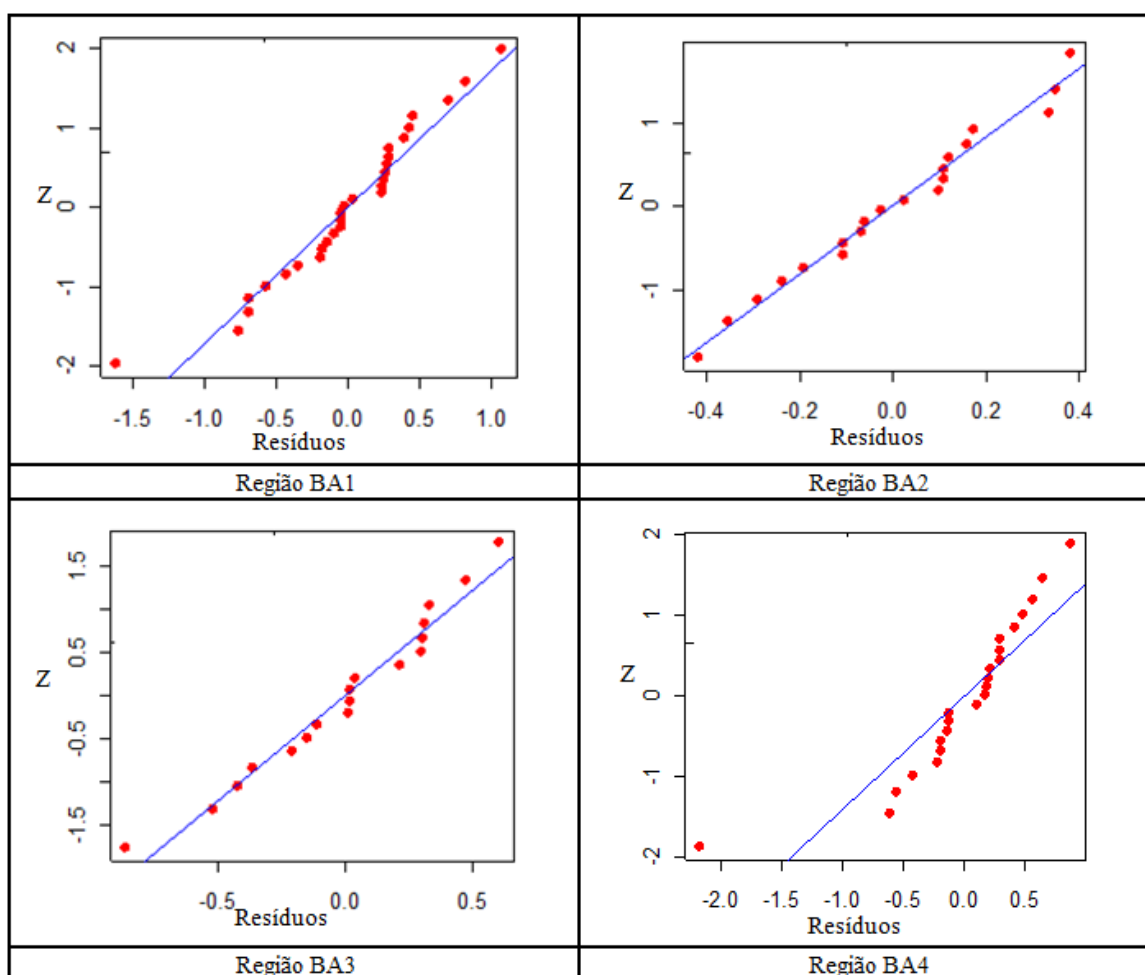
Ao observar a Tabela 5.5, nota-se que somente para região BA4 os resíduos não se distribuem normalmente. Comparando-se os valores de  $A_{eq}^2$  com  $A_{crit,0,05}^2 = 0,752$  (distribuição normal),

confirmou-se que somente para região BA4 é que a estatística de teste de Anderson-Darling não é menor que  $A_{crit,0,05}^2$ .

Portanto, conclui-se que para esta região a utilização de um modelo de regressão OLS não é adequada.

**Tabela 5.5** - Análise da estatística de teste de Anderson-Darling

Regiões	$A_{eq}^2$	p-valor	avaliação da $H_0$
BA1	0,456	0,248	NR- $H_0$
BA2	0,201	0,861	NR- $H_0$
BA3	0,270	0,634	NR- $H_0$
BA4	1,016	0,009	R- $H_0$



**Figura 5.3** - Ajuste dos resíduos do OLS à distribuição normal

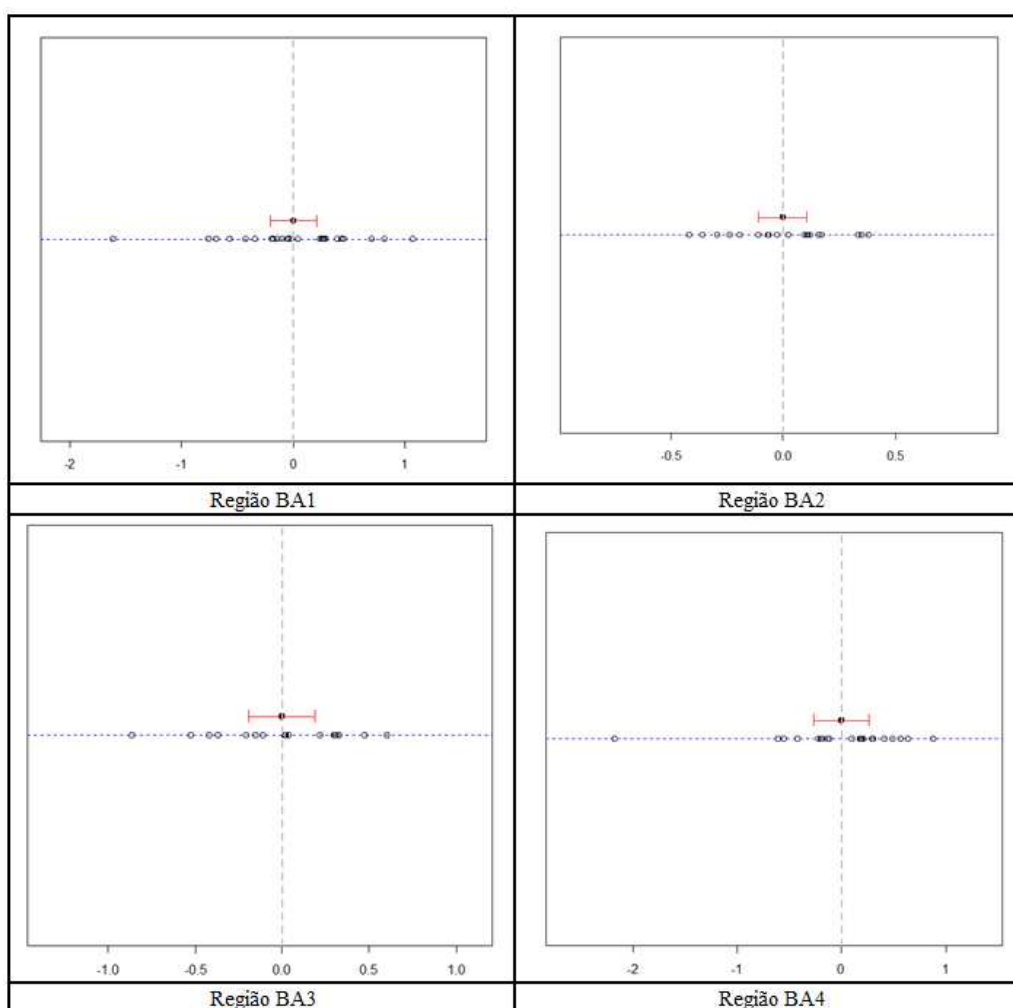
#### 5.1.2.2 Análise da média igual a zero dos resíduos

Após a análise de normalidade, determinou-se que a média dos resíduos deve ser zero. Logo, realizou-se o teste paramétrico (bilateral) sobre a média de uma única população normal (Tabela 5.6), com hipótese nula igual a zero e hipótese alternativa diferente de zero. Determinou-se também o intervalo de confiança (Figura 5.4).

Avaliando a estatística de teste, ficou evidenciado que para todas as 4 regiões a hipótese da média ser igual a zero é aceita. Analisando os intervalos de confiança, a região BA4 foi a que apresentou o maior intervalo, a um nível de confiança de 95%.

**Tabela 5.6** - Análise da estatística de teste (T) sobre média igual a zero

Regiões	T	p-valor	LI	LS	Média dos resíduos	Avaliação da $H_0$
<b>BA1</b>	0,001	0,999	-0,206	0,206	$5,684 \times 10^{-5}$	NR- $H_0$
<b>BA2</b>	0,001	0,999	-0,107	0,107	$6,359 \times 10^{-5}$	NR- $H_0$
<b>BA3</b>	-0,001	0,999	-0,188	0,188	$-8,506 \times 10^{-5}$	NR- $H_0$
<b>BA4</b>	0,003	0,998	-0,261	0,262	$3,380 \times 10^{-4}$	NR- $H_0$



**Figura 5.4** - Intervalo de confiança para média dos resíduos igual a zero

### 5.1.2.3 Análise do teste de Durbin-Watson

Seguindo as análises dos resíduos, examinou-se os resíduos quanto à possibilidade destes serem independentes. Para isso, aplicou-se o teste de Durbin-Watson (Tabela 5.7). Analisando o p-valor percebe-se que para todas as regiões os resíduos são independentes, o qual havia sido verificado através da análise da dispersão dos resíduos (Figura 5.1).

**Tabela 5.7** - Análise da estatística de teste de Durbin-Watson

Regiões	$d_w$	p-valor	Avaliação da $H_0$
<b>BA1</b>	2,224	0,655	NR- $H_0$
<b>BA2</b>	2,025	0,488	NR- $H_0$
<b>BA3</b>	2,226	0,577	NR- $H_0$
<b>BA4</b>	1,538	0,086	NR- $H_0$

#### 5.1.2.4 Análise dos testes de Breusch-Pagan e de White

Por fim, avaliou-se os resíduos em relação à variância constante (homocedasticidade). Para isso, empregou-se os testes de Breusch-Pagan e White.

A regressão auxiliar (Equação 5.2) para o teste de Breusch-Pagan é apresentada na Tabela 5.8 e a estatística de teste na Tabela 5.9. Para a composição da regressão auxiliar, resíduos do modelo potencial foram elevados ao quadrado e fez-se uma regressão utilizando o logaritmo das variáveis explicativas utilizadas nos modelos para o método OLS, definidos para as regiões BA1, BA2, BA3 e BA4.

$$\ln(\hat{u}^2) = \ln(\beta_0^*) + \beta_1^* \ln(A) + \beta_2^* \ln(L) + \beta_3^* \ln(I) + \beta_4^* \ln(P) + \beta_5^* \ln(C) + \beta_6^* \ln(Dd) \quad (5.2)$$

Fazendo a análise do p-valor, nota-se que para as regiões BA1 e BA2 os resíduos não são homocedásticos.

**Tabela 5.8** - Regressão auxiliar teste de Breusch-Pagan

Regiões	$\ln(\beta_0^*)$	$\beta_1^*$	$\beta_2^*$	$\beta_3^*$	$\beta_4^*$	$\beta_5^*$	$\beta_6^*$
<b>BA1</b>	-0,391	-0,042	-	-	0,136	-	-0,461
<b>BA2</b>	1,636	-0,760	-	-0,929	-	-4,960	-0,005
<b>BA3</b>	0,273	-0,041	-0,047	-	-0,966	-0,650	-
<b>BA4</b>	2,393	-0,225	-	-	-4,080	-1,418	-

**Tabela 5.9** - Análise da estatística de teste de Breusch-Pagan

Regiões	Breusch-Pagan	p-valor	Avaliação da $H_0$
<b>BA1</b>	7,898	0,048	R- $H_0$
<b>BA2</b>	11,337	0,023	R- $H_0$
<b>BA3</b>	5,837	0,212	NR- $H_0$
<b>BA4</b>	4,685	0,196	NR- $H_0$

A fim que se pudesse comparar os testes da variância em relação à variável predita ( $Q_{7,10}$ ), realizou-se o teste de variância constante de White, na qual também é feita uma regressão auxiliar dos resíduos ao quadrado, em função dos valores da variável predita (Equação 5.3). A regressão auxiliar é apresentada na Tabela 5.10.

$$\ln \hat{u}^2 = \ln(\beta_0^{**}) + \beta_1^{**} \ln(\hat{y}) + \beta_2^{**} (\hat{y}^2) \quad (5.3)$$

**Tabela 5.10 - Regressão auxiliar teste White**

<b>Regiões</b>	$\ln(\beta_0^{**})$	$\beta_1^{**}$	$\beta_2^{**}$
<b>BA1</b>	0,470	-0,063	0,002
<b>BA2</b>	0,021	-0,001	5,931
<b>BA3</b>	0,203	-0,025	0,001
<b>BA4</b>	0,648	-0,087	0,002

A estatística de teste de White e sua análise são apresentadas na Tabela 5.11.

**Tabela 5.11 - Análise da estatística de teste de White**

<b>Regiões</b>	White	p-valor	Avaliação da $H_0$
<b>BA1</b>	3,940	0,139	NR- $H_0$
<b>BA2</b>	18,893	$7,897 \times 10^{-5}$	R- $H_0$
<b>BA3</b>	1,272	0,529	NR- $H_0$
<b>BA4</b>	1,755	0,416	NR- $H_0$

É notório que para a região BA1 o teste de Breusch-Pagan foi diferente do teste de White, uma vez que no teste de Breusch-Pagan a hipótese nula foi rejeitada, mas no teste de White esta não foi rejeitada. Sendo assim, considerou-se que também para região BA1, os resíduos são homocedásticos.

#### 5.1.2.5 Resumo da análise dos resíduos

Na Tabela 5.12, é apresentado o resumo dos testes e hipóteses aplicados para avaliação dos resíduos dos modelos de regressão para o OLS. Nota-se que nas regiões BA2 e BA4 o modelo de regressão estimado por meio do método OLS não deve ser empregado.

**Tabela 5.12 - Resumo dos testes de hipóteses aplicado aos resíduos do OLS**

Regiões	Testes				
	Normalidade	Média igual a zero	Independência	Variância Constante	Variância Constante
	Anderson-Darling	Teste T da média	Durbin-Watson	Breusch-Pagan	White
<b>BA1</b>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	R-H <sub>0</sub>	NR-H <sub>0</sub>
<b>BA2</b>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	R-H <sub>0</sub>	R-H <sub>0</sub>
<b>BA3</b>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>
<b>BA4</b>	R-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>	NR-H <sub>0</sub>

Diante dos resultados encontrados na análise dos resíduos para as regiões BA2 e BA4, torna-se evidente a necessidade de aplicação de outros métodos de regressão nestas regiões. Contudo, para as regiões BA1 e BA3 também deve-se aplicar outros métodos de regressão, pois há a necessidade de se avaliar a contribuição da variância do erro amostral. Além disso, no método OLS, considera-se que as observações se distribuem com uma mesma variância, ou seja, tal método atribui o mesmo grau de importância a todas às observações.

## 5.2 Resultados dos métodos WLS e GLS

Nos métodos WLS e GLS, fez-se o uso da matriz de covariância dos erros amostrais. As covariâncias e as correlações cruzadas foram calculadas em relação à série das Q<sub>7</sub> mínimas anuais concorrentes. Os resultados encontrados e uma discussão a respeito dos mesmos encontra-se detalhada neste item.

### 5.2.1 Matriz de covariância dos erros amostrais e correlação cruzada

Na Figura 4.7 percebe-se que a nuvem formada pelos pares de estações é muito densa. Tal análise permite confirmar que existe uma forte correlação entre as vazões mínimas. A correlação amostral variou entre -0,67 e 0,98. A correlação média em termos de valores absolutos foi de 0,54. O maior número de anos concorrentes foi 67 anos, entre as estações 40330000 (Velho Taipa) e 40800001 (Ponte Nova do Paraopeba).

Mesmo tendo desenvolvido uma equação da correlação cruzada em função da distância, verificou-se em todas as regiões valores negativos associados à covariância dos erros amostrais. Tais valores são muito próximos de zero, e ocorreram devido ao fato da estimativa dos coeficientes de assimetria e curtose populacionais serem pequenos em algumas estações, o que resultou em valor negativo ao se aplicar a Equação 4.69. Verificou-se também que



aproximadamente 23% dos dados (4005 pares de estações) da matriz de covariância dos erros amostrais possuíam séries anuais concorrentes iguais a zero.

Em relação à distância entre as estações, o valor mínimo foi de aproximadamente de 2 km e o máximo de 350 km. Mesmo evidenciando pequenos valores de distância, percebeu-se que estes não influenciaram a função de correlação cruzada, uma vez que apenas 0,32% dos dados apresentaram pequenos valores de distância e destes, 53% possuíam séries anuais concorrentes iguais a zero.

As matrizes de covariância dos erros amostrais foi suprimida deste trabalho devido a sua grande extensão, mesmo que esteja separada por regiões.

## 5.2.2 Resultados do método WLS

No método WLS levou-se em conta a soma das variâncias do erro do modelo e da variância do erro amostral.

### 5.2.2.1 Seleção dos modelos de regressão

Escolheu-se o modelo 1 para região BA1 (Tabela 5.13), o qual foi selecionado porque apresentou menor valor da AVP e da variância do erro do modelo. Dentre os 10 modelos selecionados, este é o que possui maior coeficiente de determinação. Na análise dessa região, percebe-se que o modelo de número 10 teve  $R^2$  igual a zero. Isto é porque, como foi utilizado  $R^2$  proposto por Reis Jr. *et al.* (2004) (Equação 4.77), a soma dos quadrados da regressão é dada por  $\sum_{i=1}^N (\hat{y}_i - \hat{y}_0)$ . Então, quando não há a inclusão de nenhuma variável explicativa, tem-se

$\sum_{i=1}^N (\hat{y}_0 - \hat{y}_0)$ . Logo, o  $R^2$  é igual a zero.

**Tabela 5.13** - Modelos de regressão do método WLS região BA1

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	$R^2$
<b>1</b>	-6,978	0,899	-	-	7,915	-	0,667	0,268	0,098	0,366	0,921
<b>2</b>	-6,890	0,903	-	-	7,791	0,078	0,662	0,280	0,115	0,396	0,918
<b>3</b>	-7,381	0,947	-	0,073	7,863	-	0,648	0,280	0,116	0,396	0,919
<b>4</b>	-7,341	0,948	-	0,069	7,824	0,032	0,647	0,293	0,134	0,427	0,916
<b>5</b>	-9,654	0,988	-	-	10,403	-	-	0,361	0,106	0,467	0,903
<b>6</b>	-10,863	1,155	-	0,294	9,852	-	-	0,361	0,122	0,483	0,906
<b>7</b>	-9,203	0,990	-	-	9,868	0,273	-	0,373	0,126	0,499	0,900
<b>8</b>	2,681	-	-	-	-	-	1,370	1,345	0,115	1,459	0,330
<b>9</b>	-2,727	-	-	-	8,432	-	-	1,417	0,123	1,540	0,313
<b>10</b>	0,011	-	-	-	-	-	-	2,124	0,095	2,219	0

Para a região BA2 (Figura 5.14) o modelo escolhido foi o de número 1. Este apresentou menores valores de AVP, variância do erro do modelo e ASV. Em contrapartida apresentou um elevado coeficiente de determinação.

**Tabela 5.14 - Modelos de regressão do método WLS região BA2**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-3,931	1,193	-0,747	-	-	2,106	-0,837	0,037	0,031	0,068	0,983
<b>2</b>	-6,079	1,081	-	0,459	-	2,148	-0,639	0,038	0,037	0,075	0,986
<b>3</b>	-5,312	1,244	-0,483	0,285	-	2,114	-0,736	0,035	0,044	0,080	0,986
<b>4</b>	-3,830	0,757	-	-	-	2,142	-0,736	0,055	0,036	0,090	0,976
<b>5</b>	-4,305	1,186	-0,691	-	1,159	2,243	-0,808	0,040	0,050	0,090	0,982
<b>6</b>	-7,716	1,456	-	0,663	-	1,667	-	0,086	0,044	0,130	0,978
<b>7</b>	-6,407	0,884	-	-	-	-	-0,543	0,109	0,051	0,160	0,959
<b>8</b>	-4,657	1,045	-	-	-	1,520	-	0,121	0,044	0,165	0,964
<b>9</b>	-9,333	1,446	-	0,603	-	-	-	0,117	0,047	0,165	0,970
<b>10</b>	-6,421	1,075	-	-	-	-	-	0,142	0,041	0,183	0,958

Na região BA3 (Tabela 5.15) encontrou-se valores da variância do erro igual a zero, o que automaticamente fez com que o R<sup>2</sup> fosse igual a um. No entanto, um valor de R<sup>2</sup> igual a um é irreal. Reis Jr. (2005) também relatou modelos com variância do erro iguais a zero. Tal autor afirma que um modelo com variância igual a zero ou próxima de zero implica em um modelo regional perfeito ou quase perfeito. No entanto, como um modelo com erro igual a zero realmente não parece ser realista e, dado que a estimativa regional não é determinística, em todas as regiões onde ocorreram valores iguais a zero de variância do erro, escolheu-se o modelo seguinte cujo erro não fosse igual a zero. Portanto, devido às considerações anteriores o modelo escolhido para região BA3 foi o de número 7.

**Tabela 5.15 - Modelos de regressão do método WLS região BA3**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-4,676	0,994	-	-	-	-	0,620	0	0,022	0,022	1
<b>2</b>	-5,098	0,985	-	-	1,280	-	0,630	0	0,028	0,028	1
<b>3</b>	-4,714	0,916	0,116	-	-	-	0,588	0	0,029	0,029	1
<b>4</b>	-7,722	0,639	0,908	0,501	-	-	-	0	0,031	0,031	1
<b>5</b>	-4,310	0,953	-	-0,071	-	-	0,644	0	0,031	0,031	1
<b>6</b>	-5,298	0,625	0,868	-	-	1,617	-	0	0,034	0,034	1
<b>7</b>	-5,356	0,484	0,667	-	-	-	-	0,013	0,032	0,045	0,989
<b>8</b>	-4,810	-	1,638	-	-	1,273	-	0,022	0,033	0,055	0,982
<b>9</b>	-5,213	0,879	-	-	-	-	-	0,027	0,029	0,056	0,979
<b>10</b>	-5,013	-	1,350	-	-	-	-	0,041	0,028	0,069	0,962

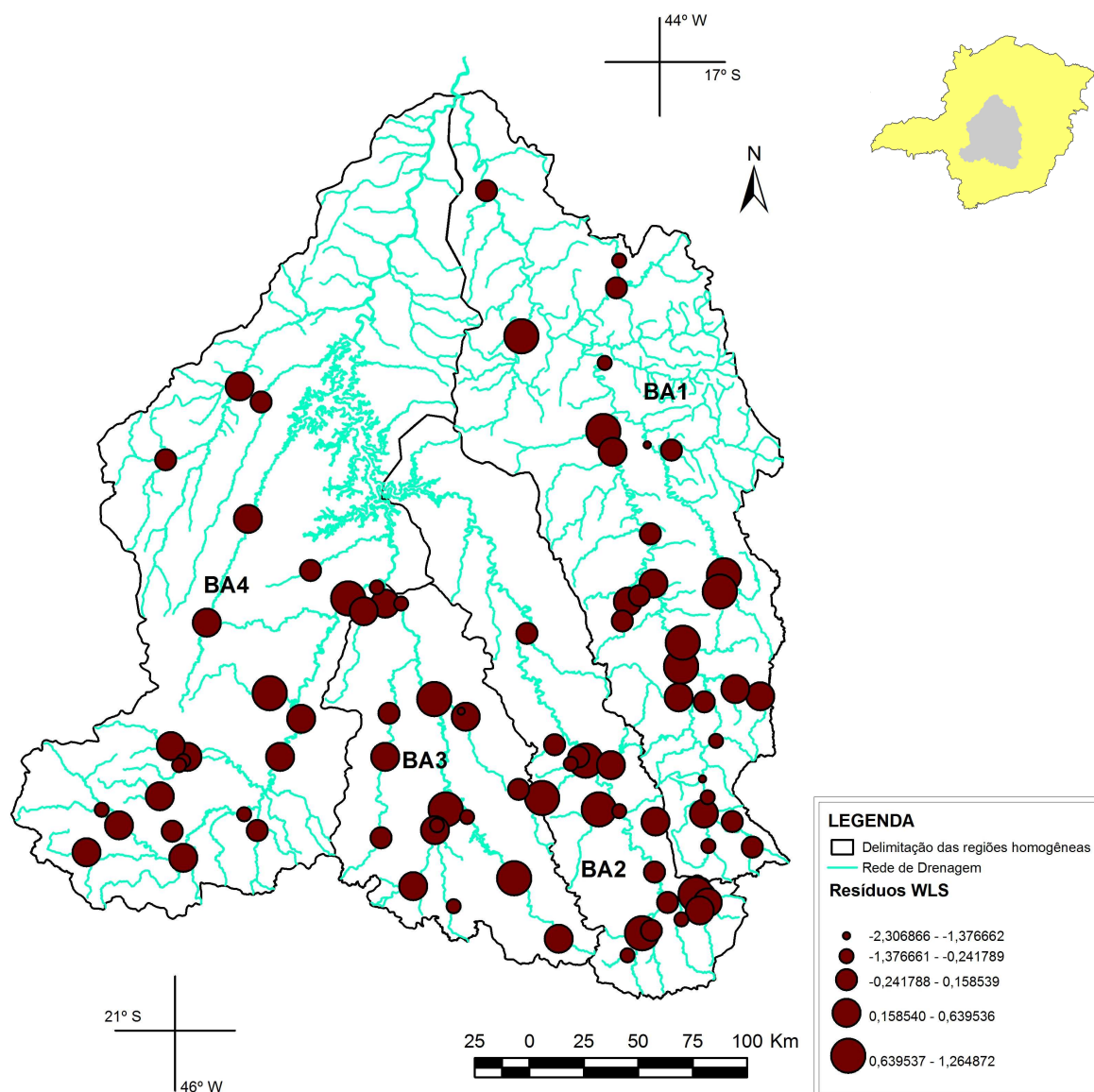
A região BA4 (Tabela 5.16) teve o menor  $R^2$ , quando comparada às outras três regiões onde se aplicou o método WLS. Além do mais foi a região onde se encontrou os maiores valores da AVP e da variância do erro do modelo. O modelo escolhido para esta região foi o de número 1.

**Tabela 5.16** - Modelos de regressão do método WLS região BA4

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_\varepsilon^2$	ASV	AVP	$R^2$
<b>1</b>	-8,297	1,011	-	-	8,971	2,110	-	0,261	0,154	0,415	0,896
<b>2</b>	-8,158	1,244	-0,394	-	9,034	2,157	-	0,270	0,174	0,444	0,895
<b>3</b>	-11,372	0,962	-	-	12,339	-	-	0,331	0,159	0,490	0,870
<b>4</b>	-9,813	-	1,126	-	12,315	-	-	0,542	0,178	0,720	0,698
<b>5</b>	-5,289	-	-	-	12,714	-	-	0,931	0,103	1,034	0,339
<b>6</b>	-3,403	-	-	-	10,920	1,207	-	0,969	0,157	1,127	0,343
<b>7</b>	-6,068	-	-	-	12,937	-	-0,278	0,992	0,155	1,147	0,329
<b>8</b>	3,228	-	-	-	-	3,220	-	1,296	0,144	1,440	0,161
<b>9</b>	0,315	-	-	-	-	-	-	1,525	0,090	1,615	0
<b>10</b>	0,166	-	-	-	-	-	-0,064	1,623	0,169	1,792	0,012

### 5.2.2.2 Análise dos resíduos

Os resíduos dos modelos do WLS para as 4 regiões são apresentados na Figura 5.5. Nota-se que, assim como nos resíduos dos modelos OLS, estes demonstram ausência de correlação entre os mesmos.



**Figura 5.5** – Mapa dos resíduos dos modelos WLS

### 5.2.3 Resultados do método GLS

No modelo GLS considerou-se a correlação existente entre as séries de vazões observadas.

#### 5.2.3.1 Seleção dos modelos de regressão

Para a região BA1 (Tabela 5.17) o modelo escolhido foi o de número 1.

**Tabela 5.17 - Modelos de regressão do método GLS região BA1**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-6,368	0,814	-	-	7,594	-	0,684	0,257	0,120	0,376	0,911
<b>2</b>	-9,226	0,923	-	-	10,196	-	-	0,356	0,135	0,491	0,892
<b>3</b>	0,196	-	-	-0,714	6,914	-	0,942	0,470	0,119	0,589	0,718
<b>4</b>	-0,109	-	-	-	4,789	-	0,878	0,735	0,097	0,832	0,488
<b>5</b>	2,239	-	-	-	-	-	1,286	0,886	0,082	0,968	0,379
<b>6</b>	2,852	-	-	-	-	0,908	1,110	0,856	0,116	0,972	0,414
<b>7</b>	-2,830	-	-	-	7,897	-	-	0,954	0,090	1,044	0,352
<b>8</b>	-2,133	-	-	-	7,120	0,416	-	0,997	0,137	1,133	0,347
<b>9</b>	1,644	-	-	-	-	1,826	-	1,307	0,121	1,428	0,178
<b>10</b>	-0,275	-	-	-	-	-	-	1,588	0,080	1,668	0

Para a região BA2 (Tabela 5.18), definiu-se o modelo 1. Este apresentou uma variância do erro pequena, quando comparada às regiões BA1 e BA4. Logo o R<sup>2</sup> teve um resultado elevado.

**Tabela 5.18 - Modelos de regressão do método GLS região BA2**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
<b>1</b>	-3,949	1,180	-0,749	-	-	2,051	-0,844	0,037	0,033	0,071	0,982
<b>2</b>	-6,082	1,073	-	0,458	-	2,121	-0,642	0,038	0,041	0,079	0,986
<b>3</b>	-5,312	1,234	-0,487	0,283	-	2,075	-0,744	0,035	0,048	0,084	0,985
<b>4</b>	-3,837	0,744	-	-	-	2,099	-0,739	0,055	0,040	0,094	0,975
<b>5</b>	-4,265	1,175	-0,700	-	0,978	2,170	-0,817	0,041	0,053	0,094	0,981
<b>6</b>	-4,627	0,796	-	-	2,512	2,416	-0,697	0,054	0,056	0,110	0,976
<b>7</b>	-7,701	1,457	-	0,657	-	1,676	-	0,087	0,048	0,135	0,978
<b>8</b>	-4,667	1,044	-	-	-	1,514	-	0,121	0,049	0,170	0,963
<b>9</b>	-9,296	1,442	-	0,593	-	-	-	0,118	0,053	0,171	0,970
<b>10</b>	-6,413	1,072	-	-	-	-	-	0,142	0,047	0,189	0,958

Na região BA3 (Tabela 5.19), assim como para o método WLS, observou-se uma convergência na variância do erro para zero. Dos 64 modelos gerados, em aproximadamente 31% a variância do erro do modelo foi igual a zero. Ao classificar a AVP, o modelo visto como sendo o melhor modelo GLS para esta região foi o de número 6. Este modelo foi definido mesmo tendo erro bem próximo de zero, porque ao se testar outros modelos, verificou-se que os quantis de vazões estimados tendiam a ter valores muito discrepantes comparado ao quantil observado.

**Tabela 5.19 - Modelos de regressão do método GLS região BA3**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
1	-4,654	0,989	-	-	-	-	0,626	0	0,033	0,033	1
2	-5,036	0,977	-	-	1,198	-	0,633	0	0,038	0,038	1
3	-4,691	0,882	0,154	-	-	-	0,581	0	0,040	0,040	1
4	-3,956	0,906	-	-0,132	-	-	0,664	0	0,042	0,042	1
5	-5,093	0,860	0,169	-	1,247	-	0,584	0	0,046	0,046	1
6	-7,461	0,553	0,979	0,464	-	-	-	0,004	0,046	0,049	0,997
7	-5,210	0,437	0,692	-	-	-	-	0,017	0,046	0,063	0,984
8	-6,392	-	1,285	-	3,946	-	-	0,027	0,044	0,071	0,967
9	-4,599	-	1,530	-	-	1,130	-	0,027	0,045	0,072	0,975
10	-4,778	-	1,270	-	-	-	-	0,040	0,038	0,078	0,956

Na região BA4 (Tabela 5.20) não se observou nenhuma discrepância dos resultados. Contudo, assim como para o método WLS, foi a região onde foram constatadas os maiores valores da variância do erro do modelo e, conseqüentemente, o menor R<sup>2</sup>.

**Tabela 5.20 - Modelos de regressão do método GLS região BA4**

Modelos	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	ASV	AVP	R <sup>2</sup>
1	-8,242	0,982	-	-	9,081	2,049	-	0,264	0,175	0,439	0,889
2	-11,124	0,923	-	-	12,260	-	-	0,330	0,181	0,512	0,859
3	-9,402	-	1,027	-	12,155	-	-	0,517	0,186	0,703	0,670
4	-5,278	-	-	-	12,396	-	-	0,805	0,095	0,900	0,367
5	-3,345	-	-	-	10,566	1,244	-	0,830	0,141	0,971	0,377
6	-5,926	-	-	-	12,603	-	-0,229	0,861	0,142	1,002	0,354
7	-3,958	-	-	-	10,458	1,552	-0,383	0,872	0,189	1,061	0,374
8	3,096	-	-	-	-	3,221	-	1,145	0,135	1,280	0,183
9	0,189	-	-	-	-	-	-	1,387	0,090	1,477	0
10	0,156	-	-	-	-	-	-0,018	1,478	0,163	1,641	0,004

#### 5.2.4 Análise dos resíduos

Os resíduos dos modelos do GLS (Figura 5.6) estão mais distribuídos que os do WLS.

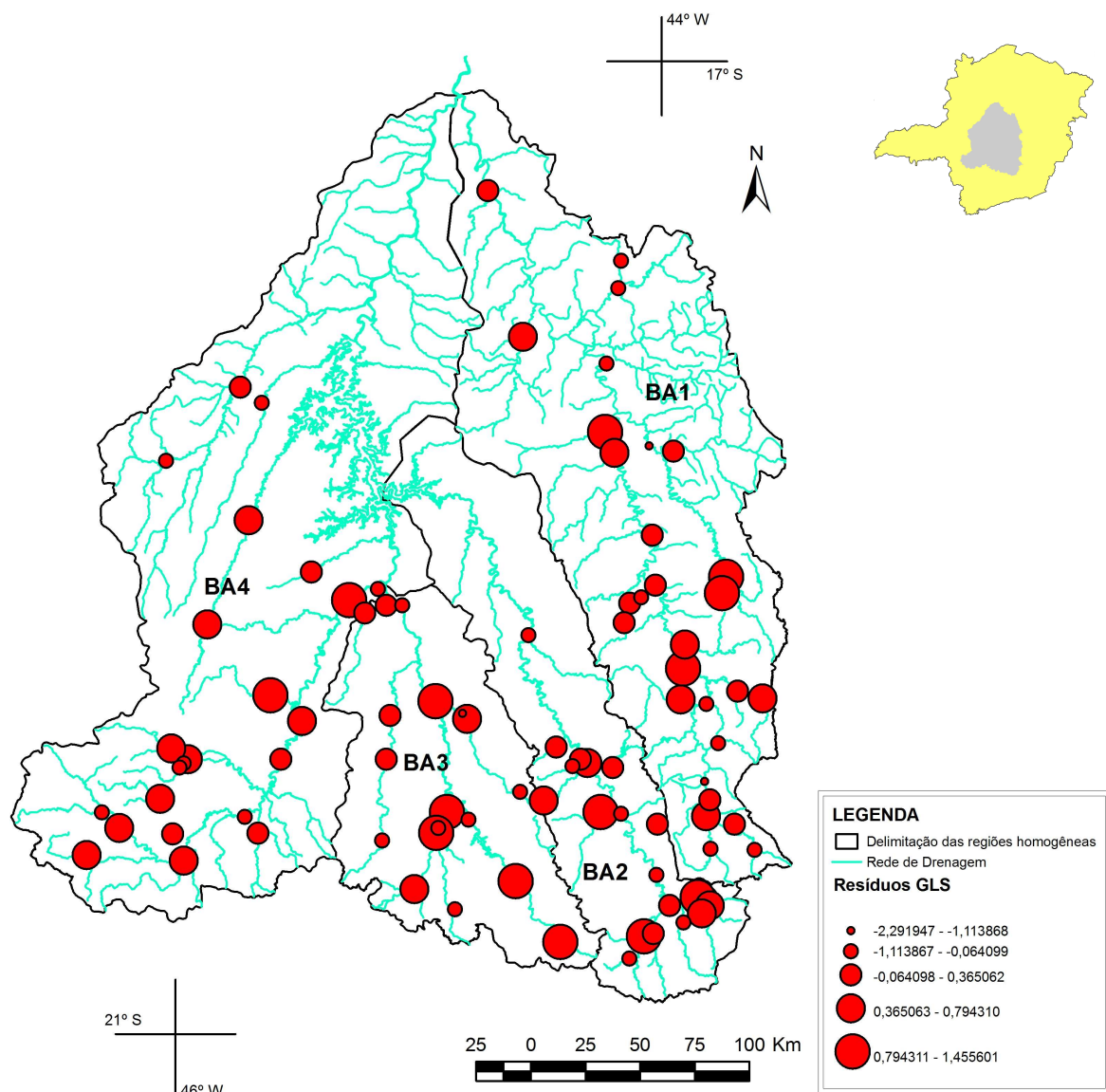


Figura 5.6 – Mapa dos resíduos dos modelos do GLS

### 5.3 Estudo comparativo entre os métodos OLS, WLS e GLS

#### 5.3.1 Modelos selecionados para os métodos OLS, WLS e GLS

Nas Tabelas 5.22, 5.23, 5.24 e 5.25 são apresentados, respectivamente, os modelos referentes às regiões BA1, BA2, BA3 e BA4. Com intuito de se entender melhor cada modelo colocou-se a Equação 5.1 na forma potencial (Equação 5.4). A partir daí, determinou-se os valores do erro relativo médio (ERM).

Plotou-se a vazão estimada versus a observada para cada uma das regiões, na escala logarítmica nos eixos das abscissas e das ordenadas (Figuras 5.7 a 5.10), a fim de que se possa visualizar o ajuste linear dos modelos. Plotou-se somente os modelos OLS e GLS, uma vez que os modelos WLS apresentaram resultados similares aos do GLS. Os valores dos quantis da  $Q_{7,10}$  estimados para cada região homogênea é apresentado no Apêndice C.

$$\hat{Q}_{7,10} = \beta_0 A^{\beta_1} L^{\beta_2} \dots Dd^{\beta_6} \quad (5.4)$$

Na Tabela 5.21, estão apresentados o intervalo de validade, para cada uma das quatro regiões homogêneas.

**Tabela 5.21 – Intervalo de validade das equações regionais**

Região homogênea	Intervalo de validação para cada região A= área de drenagem (km <sup>2</sup> )
BA1	85,0 ≤ A ≤ 26.553,0
BA2	10,2 ≤ A ≤ 8.734,0
BA3	144,0 ≤ A ≤ 11.350,0
BA4	259,0 ≤ A ≤ 14.180,0

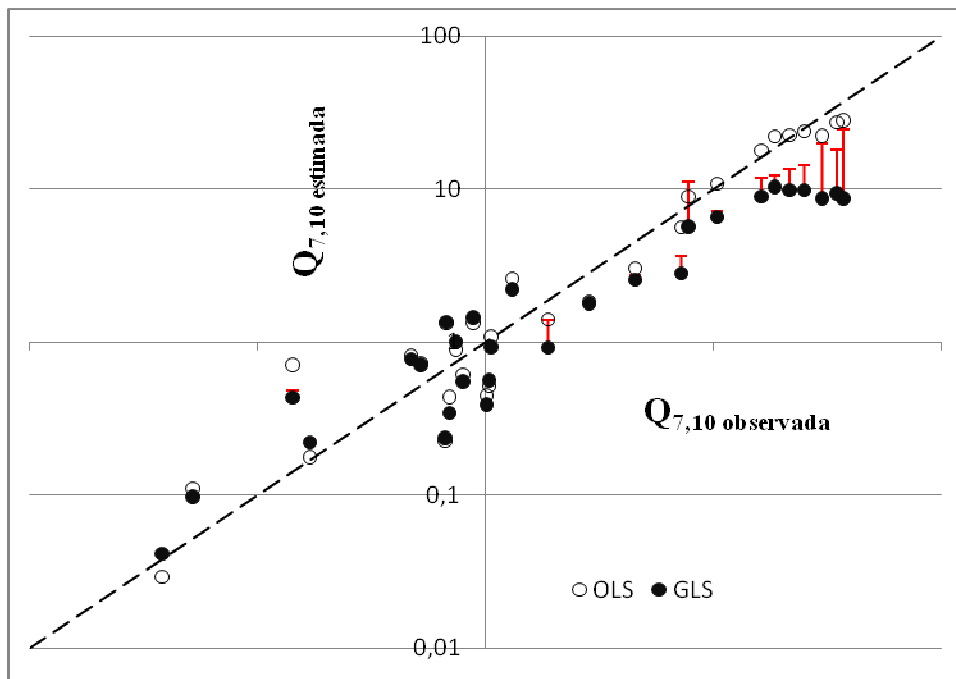
### 5.3.1.1 Região BA1

Em relação à variância do erro do modelo, o método GLS teve melhor desempenho. Contudo o erro relativo médio foi elevado. Os modelos foram bons no ajuste dos valores mínimos. No entanto, para valores máximos o método OLS foi melhor. Analisando os valores das  $Q_{7,10}$  estimadas é perceptível que, o que está fazendo com que o erro relativo médio para os modelos WLS e GLS sejam maiores, é que o modelo não está se ajustando bem aos valores elevados da série das  $Q_{7,10}$  observadas. Nota-se que estes valores possuem valores elevados da variância do erro amostral (Figura 5.7 a 5.10) e, de acordo com Reis Jr. (2005), para o caso em que a variância amostral é elevada, uma das alternativas adotadas seria a utilização do GLS bayesiano.

**Tabela 5.22- Resultados dos modelos para região BA1**

Métodos de regressão	ln β <sub>0</sub>	β <sub>1</sub>	β <sub>2</sub>	β <sub>3</sub>	β <sub>4</sub>	β <sub>5</sub>	β <sub>6</sub>	δ <sub>ε</sub> <sup>2</sup>	R <sup>2</sup>	R <sup>2</sup> ajustado	ERM
<b>OLS</b>	-8,377	1,090	-	-	8,381	-	0,603	0,328	0,938	0,930	0,484
<b>WLS</b>	-6,978	0,899	-	-	7,915	-	0,667	0,268	0,921	0,911	0,559
<b>GLS</b>	-6,368	0,814	-	-	7,594	-	0,684	0,257	0,911	0,900	0,568





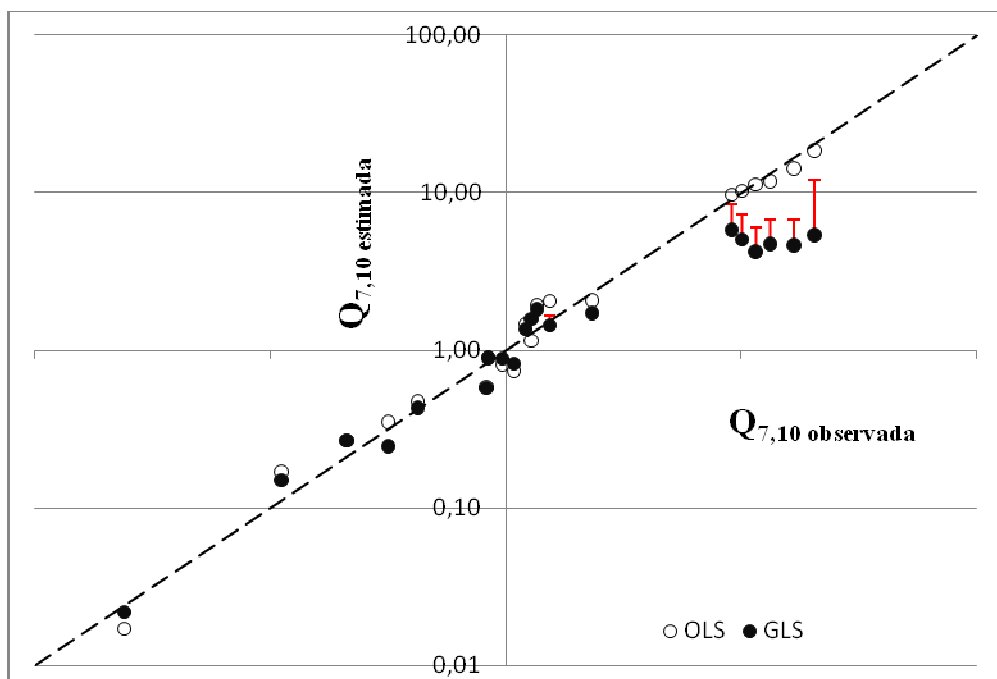
**Figura 5.7** - Vazões estimadas *versus* observadas região BA1, com barra da variância do erro amostral (em vermelho)

### 5.3.1.2 Região BA2

Na região BA2 os valores do erro relativo médio também foram menores para o método OLS. Em termos de erro relativo médio e da variância do erro do modelo, os métodos WLS e GLS foram iguais. Estes modelos também se ajustaram muito bem aos valores mínimos.

**Tabela 5.23** - Resultados dos modelos para região BA2

Métodos de regressão	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	$R^2$	$R^2$ ajustado	ERM
<b>OLS</b>	-7,379	1,283	-	0,555	-	1,710	-0,432	0,066	0,985	0,981	0,189
<b>WLS</b>	-3,931	1,193	-0,747	-	-	2,106	-0,837	0,037	0,983	0,978	0,309
<b>GLS</b>	-3,949	1,180	-0,749	-	-	2,051	-0,844	0,037	0,982	0,977	0,312



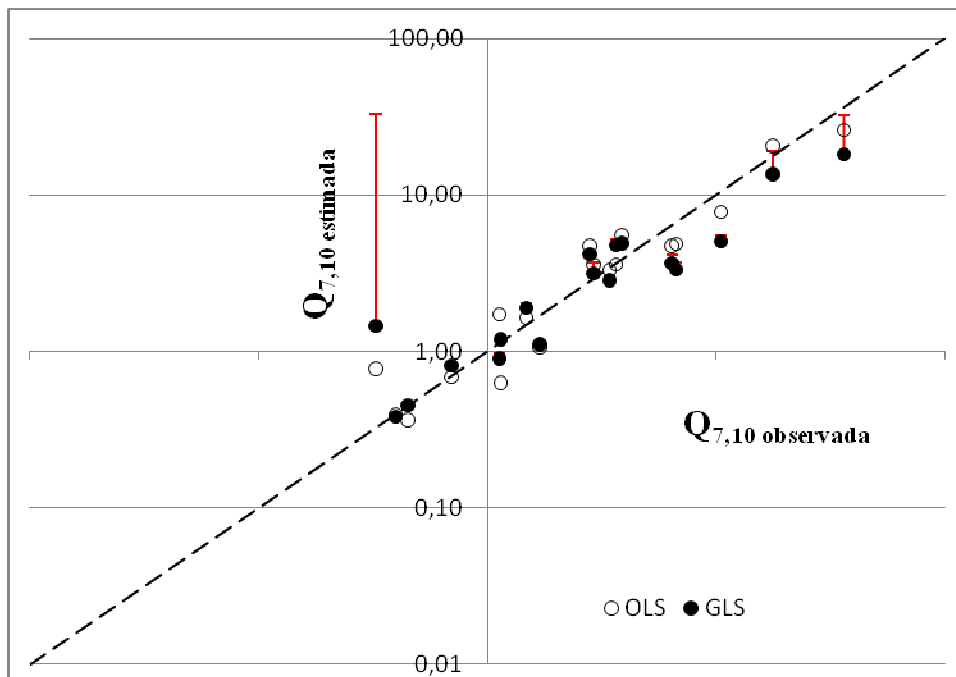
**Figura 5.8** - Vazões estimadas versus observadas região BA2, com barra da variância do erro amostral (em vermelho)

### 5.3.1.3 Região BA3

Na região BA3, o método WLS teve o maior valor do erro relativo médio. E para o método GLS observou-se o maior valor do  $R^2$ . Como já mencionado anteriormente, valores pequenos de variância correspondem a elevados valores de  $R^2$ . No entanto tal fato, está fazendo com que o modelo tenha um maior erro relativo médio. Contudo nesta região, na estimativa da  $Q_{7,10}$  para os valores máximos das séries de mínimas anuais, não se observou grandes diferenças. É importante salientar que esta região é a que possui maiores valores da variância do erro amostral e altas correlações e, de acordo com Kroll e Stedinger (1998), para pequenos valores da variância do erro do modelo, a eficiência do método OLS diminui conforme as correlações cruzadas aumentam. Neste caso, o método GLS torna-se mais adequado.

**Tabela 5.24** - Resultados dos modelos para região BA3

Métodos de regressão	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	$R^2$	$R^2$ ajustado	ERM
<b>OLS</b>	-7,462	0,846	0,671	-	5,571	2,295	-	0,186	0,925	0,904	0,317
<b>WLS</b>	-5,356	0,484	0,667	-	-	-	-	0,013	0,989	0,988	0,565
<b>GLS</b>	-7,461	0,553	0,979	0,464	-	-	-	0,004	0,997	0,996	0,448



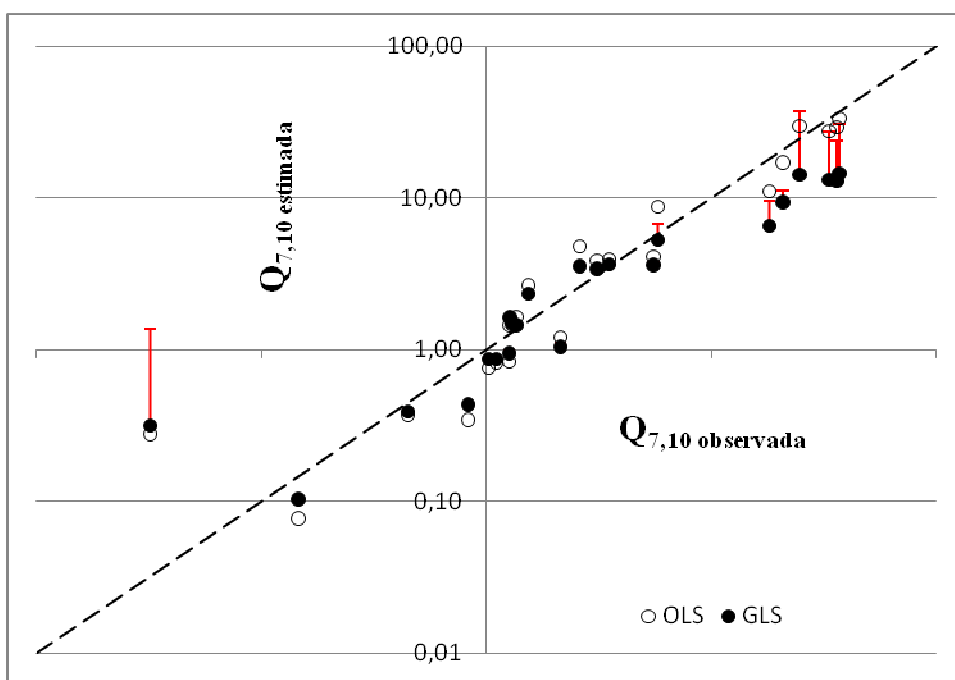
**Figura 5.9** - Vazões estimadas *versus* observadas região BA3, com barra da variância do erro amostral (em vermelho)

#### 5.3.1.4 Região BA4

A região BA4 foi a região onde se observou que para todos os métodos, o valor da variância do erro do modelo foi maior, quando comparada às outras três regiões. O erro relativo médio foi semelhante para os métodos WLS e GLS.

**Tabela 5.25** - Resultados dos modelos para região BA4

Métodos de regressão	$\ln \beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\delta_e^2$	$R^2$	$R^2$ ajustado	ERM
<b>OLS</b>	-10,128	1,254	-	-	9,701	2,177	-	0,423	0,906	0,891	0,638
<b>WLS</b>	-8,297	1,011	-	-	8,971	2,110	-	0,261	0,896	0,880	0,710
<b>GLS</b>	-8,242	0,982	-	-	9,081	2,049	-	0,264	0,889	0,871	0,711



**Figura 5.10** - Vazões estimadas versus observadas região BA4, com barra da variância do erro amostral (em vermelho)

### 5.3.2 Avaliação comparativa entre os métodos OLS, WLS e GLS

Após verificar que os modelos estimados por meio dos métodos WLS e GLS obtiveram elevados valores concernentes ao erro relativo médio, decidiu-se avaliar a razão entre as variâncias do erro - EVR (Tabelas 5.26 e 5.27). No entanto, quando se calculou o traço da matriz de covariância percebeu-se que a variância do erro amostral total tinha valores extremamente altos quando comparados à variância do erro do modelo total. Logo, os valores de EVR foram demasiadamente elevados. Segundo Griffis e Stedinger (2007), se a EVR for superior a 20%, o modelo OLS deve ser rejeitado.

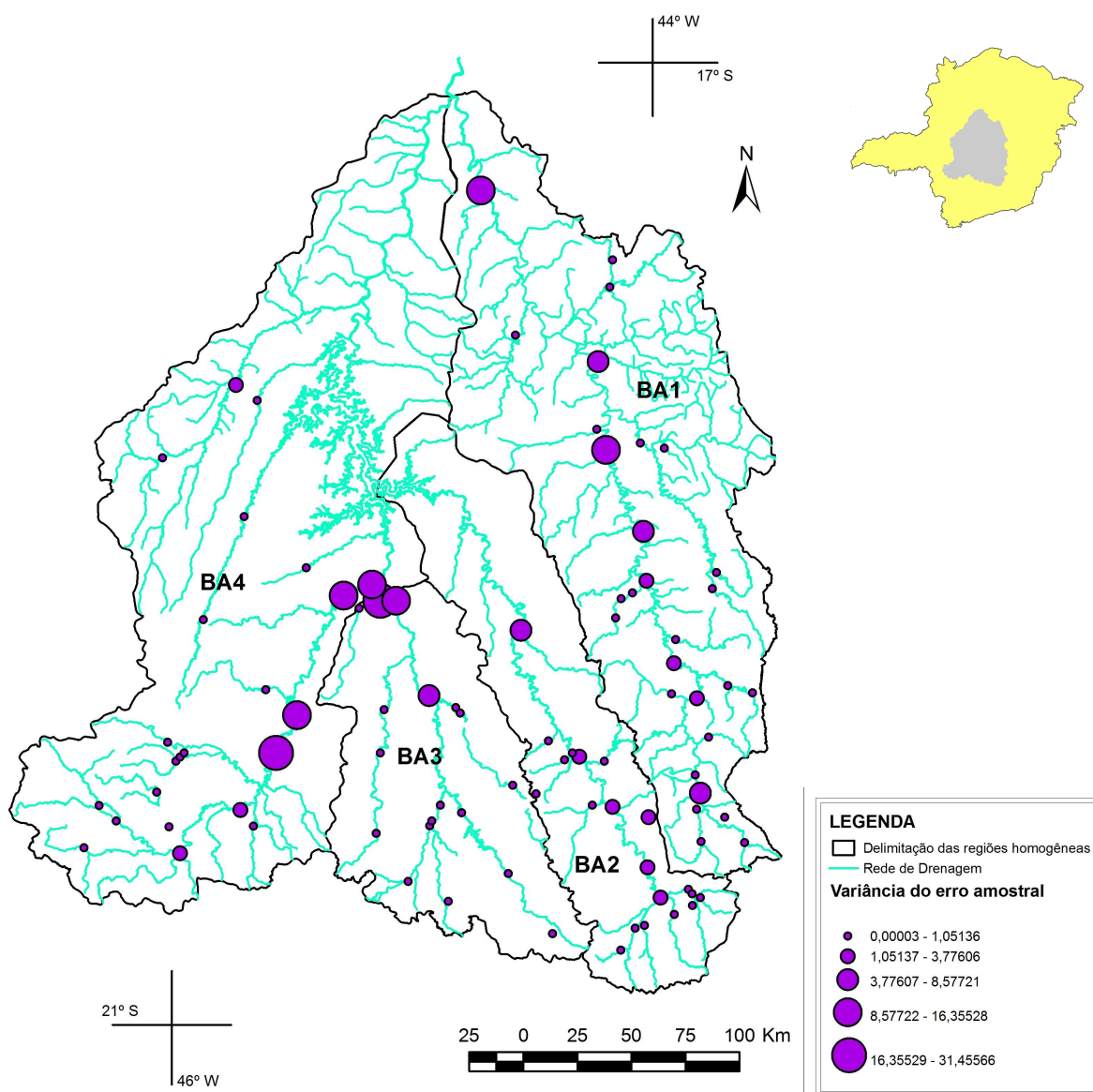
**Tabela 5.26** - Cálculo da EVR para os modelos WLS

Regiões	$tr(\Sigma)$	$N\delta_{\varepsilon}^2(P)$	EVR
<b>BA1</b>	59,941	7,776	7,709
<b>BA2</b>	18,734	0,749	25,022
<b>BA3</b>	30,584	0,237	129,172
<b>BA4</b>	73,392	6,007	12,218

**Tabela 5.27** - Cálculo da EVR para os modelos GLS

Regiões	$tr(\Sigma)$	$N\delta_{\varepsilon}^2(P)$	EVR
<b>BA1</b>	59,941	7,444	8,052
<b>BA2</b>	18,734	0,750	24,991
<b>BA3</b>	30,584	0,065	468,132
<b>BA4</b>	73,392	6,069	12,094

Contudo, analisando a matriz das variâncias amostrais (Figura 5.11), o que está pesando no valor final da somatória das variâncias são algumas poucas estações. Então, a fim de que se pudesse avaliar o efeito dos valores elevados da variância do erro amostral, fez-se uma análise gráfica destas variâncias para cada uma das regiões. Na construção dos gráficos, classificou-se as variâncias dos erros amostrais em ordem crescente e a variância do erro amostral média foi delimitada com uma linha tracejada em vermelho. Considerou-se que os valores da variância do erro amostral que estavam acima do valor médio eram mais significativos, ou seja, tinham mais peso dentro da matriz de covariância dos erros amostrais.



**Figura 5.11 - Intervalos da variância do erro amostral**

A Figura 5.12, refere-se ao gráfico da variância do erro amostral para as 29 estações da região BA1. O valor máximo da variância do erro amostral é 17,27 (m<sup>3</sup>/s)<sup>2</sup>, o mínimo de 0,00638(m<sup>3</sup>/s)<sup>2</sup>, e o médio de 1,99(m<sup>3</sup>/s)<sup>2</sup>. Observa-se que das 29 estações, 8 possuem erros

bastante elevados, as quais são, 41340000, 41260000, 41410000, 41600000, 41195000, 41818000, 41650002 e 41990000.

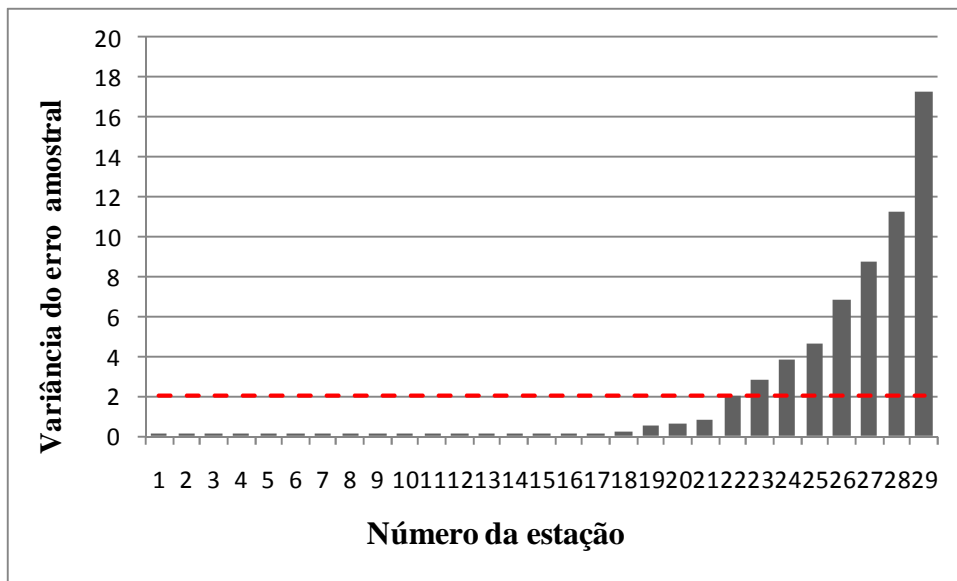


Figura 5.12 - Variância do erro amostral de todas as estações da região BA1

Na a região BA2 (Figura 5.13), verificou-se que das 20 estações, 6 possuem variância do erro amostral elevada, as quais são, 40740000, 40720002, 40800001, 40710000, 40700002 e 40850000. O valor mínimo da variância do erro amostral nesta região foi de  $3,597 \times 10^{-5} (m^3/s)^2$ , o máximo de  $7,108 (m^3/s)^2$ , e o médio de  $0,937 (m^3/s)^2$ .

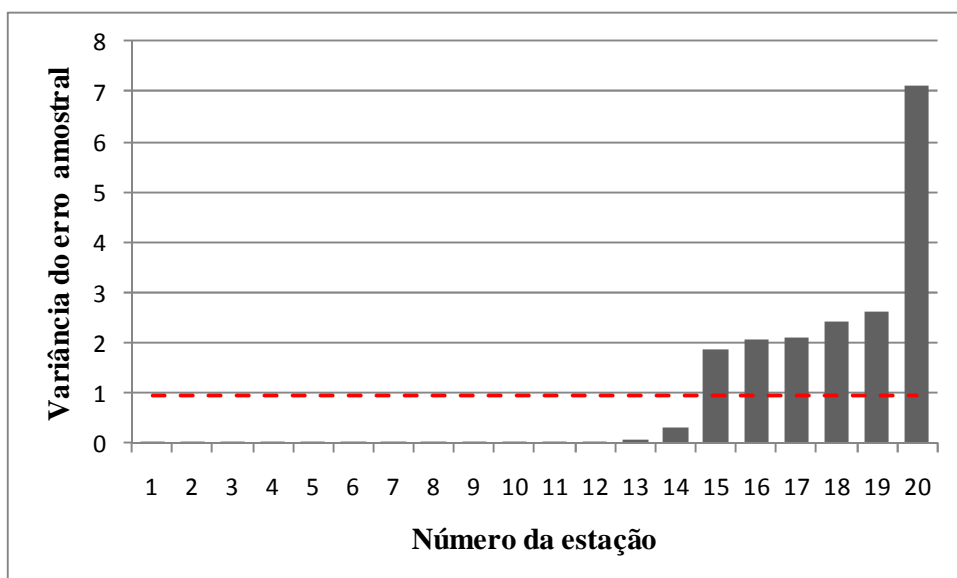
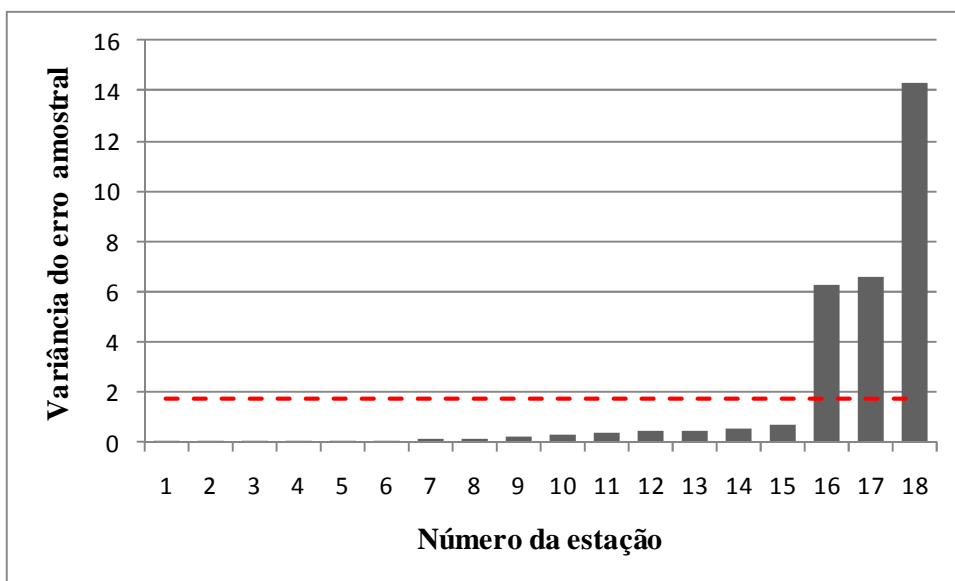


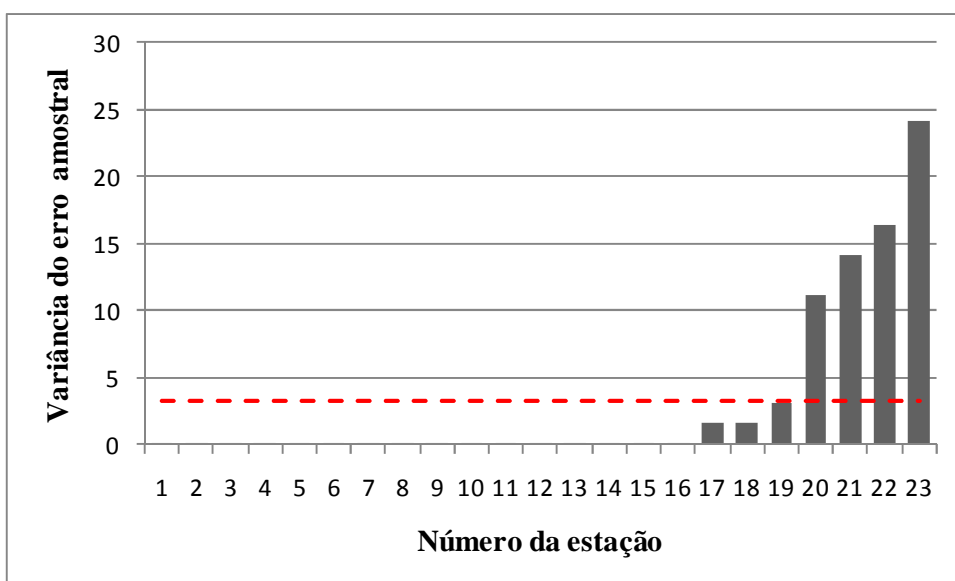
Figura 5.13 - Variância do erro amostral de todas as estações da região BA2

Na região BA3 (Figura 5.14), das 18 estações, as estações 40330000, 40505000 e 40450000 possuem valores elevados da variância do erro amostral, sendo que o valor mínimo nesta região é de 0,0037 (m³/s)², o máximo de 14,331 (m³/s)² e o médio de 1,699 (m³/s)².



**Figura 5.14** - Variância do erro amostral de todas as estações da região BA3

Na região BA4 (Figura 5.15), das 23 estações, as estações 40100000, 40070000, 40102000 e 40067000 apresentaram valores elevados da variância do erro amostral. O valor mínimo da variância do erro amostral dessa região é de 0,002032 (m³/s)², o máximo de 24,183 e o médio 3,190 (m³/s)².



**Figura 5.15** - Variância do erro amostral de todas as estações da região BA4

Constatou-se que todas as estações que possuem valores elevados da variância do erro amostral, possuem valores elevados de variância amostral, os quais são da ordem de 270 (m<sup>3</sup>/s)<sup>2</sup>. Enquanto que para as estações com baixos valores de variância do erro amostral, os valores são da ordem de 8 (m<sup>3</sup>/s)<sup>2</sup>.

Logo conclui-se que o mais adequado seria a rejeição do modelo OLS, pois mesmo que fossem retiradas tais estações, o valor da variância do erro amostral continuaria muito superior ao da variância do erro do modelo. Por exemplo, o traço da matriz de covariância dos erros amostrais passaria a ser, 2,61 (m<sup>3</sup>/s)<sup>2</sup>, 0,537 (m<sup>3</sup>/s)<sup>2</sup>, 3,421 (m<sup>3</sup>/s)<sup>2</sup> e 5,468 (m<sup>3</sup>/s)<sup>2</sup>, respectivamente para as regiões BA1, BA2, BA3 e BA4. Fez-se um teste gerando novos modelos de regressão, sem as estações, que apresentaram valores elevados da variância do erro amostral e constatou-se que o erro do modelo continuou inalterado. Contudo, os modelos apresentaram novos ajustes no que se refere ao R<sup>2</sup>, pois em relação às variáveis explicativas inclusas nos modelos anteriores não ocorreu nenhuma alteração.

### 5.3.3 Avaliação comparativa entre os métodos WLS e GLS

Na comparação dos métodos WLS e GLS, calculou-se a razão de inflação das variâncias (VIR). De acordo com Griffis e Stedinger (2007), se a VIR é significativamente maior que 1, o método GLS deve ser empregado, caso contrário o método WLS é suficiente.

Analisando os resultados da Tabela 5.27. Nota-se que em todas as regiões a VIR foi menor que 1. Logo o método WLS é suficiente e a modelagem por meio do GLS não é necessária.

**Tabela 5.28 - Cálculo da VIR**

<b>Regiões</b>	<b>VIR</b>
<b>BA1</b>	0,128
<b>BA2</b>	0,140
<b>BA3</b>	0,117
<b>BA4</b>	0,136

A VIR, mede somente a distorção do valor constante. Sendo assim, torna-se necessário avaliar se realmente, para o caso da Q<sub>7,10</sub>, a covariância estimada teria maior peso sobre a estimativa do valor constante.



## 6 CONCLUSÕES E RECOMENDAÇÕES

No presente trabalho, propôs-se a realização de um estudo de regionalização de vazões mínimas utilizando o Método dos Mínimos Quadrados Generalizados - GLS (*Generalized Least Square*) e sua aplicação à bacia do Alto São Francisco. Tal bacia compreende as sub-bacias 40 e 41. O método GLS aplicado foi sistematizado por Stedinger e Tasker (1985, 1986a, 1986b).

Para que se pudesse avaliar o desempenho do método GLS, fez-se um estudo comparativo entre este e os Métodos dos Mínimos Quadrados Ordinários - OLS (*Ordinary Least Square*) e o Método dos Mínimos Quadrados Ponderados - WLS (*Weighted Least Square*), sendo este último aplicado como proposto por Tasker e Stedinger (1986), ou seja, a variância total do erro do modelo foi dividida em duas partes, variância do erro amostral e variância do erro do modelo.

Assim como mencionado para o método WLS, o GLS também leva em consideração que o erro do modelo é dividido em duas partes. No entanto, a diferença primordial entre estes dois métodos é que o GLS considera a correlação cruzada entre os fluxos concorrentes. Além desta diferença, o método GLS também leva em consideração a diferença de comprimento das séries de vazões.

Para a construção do modelo regional, inicialmente selecionou-se as estações fluviométricas que possuíam série histórica de vazões mínimas diárias superiores a 10 anos. As estações utilizadas eram as de responsabilidade da Agência Nacional de Águas - ANA. Sendo assim, foram selecionadas 119 estações e após a análise de frequência local restaram 90 estações. A análise de frequência local foi realizada mediante o ajuste da curva empírica da distribuição de probabilidade Weibull 2P. Esta distribuição foi escolhida devido ao fato de apresentar um bom desempenho na modelagem de eventos mínimos.

Após a seleção das estações fluviométricas, fez-se a divisão da área de estudo em regiões homogêneas de forma subjetiva. Mas, para isto, também foram analisados as Unidades de Planejamento e Gestão de Recursos Hídricos (UPGRH), utilizadas pelo Instituto Mineiro de Gestão das Águas (IGAM). Portanto, a área de estudo foi dividida em 4 regiões homogêneas, as quais são, bacia do rio das Velhas, bacia do rio Paraopeba, bacia do rio Pará e a união das bacias dos rios Ajudas, Bambuí, Indaiá, Borrachudo, Abaeté e a parcela do rio São Francisco na sub-bacia 40, as quais são referenciadas, respectivamente com os códigos BA1, BA2, BA3 e BA4.

Na construção do modelo GLS Stedinger - Tasker foi necessário elaborar uma equação de covariância dos erros amostrais para a distribuição Weibull, a qual é a grande contribuição desse trabalho. Por conseguinte, construiu-se uma equação matemática que levasse em consideração a relação entre os momentos bivariados. Tal equação, foi estruturada com base em estudos nos momentos amostrais, média, variância e coeficiente de assimetria. Além da equação de covariância, determinou-se uma função que levasse em conta a relação entre as correlações cruzadas e as distâncias entre os pares de estações, as quais foram ajustadas em uma curva potencial. Recomenda-se que esta equação seja estimada para diferentes funções, além da função potencial.

Foram gerados 64 modelos para 6 variáveis explicativas, as quais são, área de drenagem, comprimento do rio principal, declividade equivalente, precipitação média anual, coeficiente de escoamento e densidade de drenagem. Contudo, para que se pudesse avaliar qual seria o melhor método de regressão, selecionou-se somente 10 modelos que apresentaram os menores da variância média de predição - AVP, ou seja, quanto menor o valor da AVP melhor é o modelo de regressão. O modelo de regressão avaliado foi do tipo potencial.

Para o modelo de regressão OLS, realizou-se testes de hipótese para avaliar as premissas inerentes aos resíduos deste método. Aplicou-se testes para avaliar a normalidade, a média igual a zero, a independência e a homocedasticidade. Os testes aplicados foram, respectivamente, de aderência de Anderson-Darling e paramétrico bilateral sobre média de uma única população normal. Os testes de hipótese foram empregados ao nível de significância de 5%. Também fez-se a avaliação dos resíduos para os métodos WLS e GLS através da análise gráfica.

Os métodos OLS , WLS e GLS foram avaliados de forma comparativa por meio da análise da erro relativo médio -ERM, razão entre as variâncias - EVR e razão de inflação das variâncias - VIR.

Os modelos selecionados para a região BA1 para os métodos de regressão tiveram como inclusão as variáveis explicativas, área de drenagem, precipitação média anual e densidade de drenagem. Na região BA2, para o método OLS, teve-se a inclusão da área de drenagem, da declividade equivalente, do coeficiente de escoamento superficial e da densidade de drenagem e, os métodos WLS e GLS tiveram a inclusão da área de drenagem, do comprimento do rio principal, do coeficiente de escoamento e da densidade de drenagem. A região BA3 teve como inserção as seguintes variáveis explicativas: para o método OLS foram, a área de

drenagem, o comprimento do rio principal, a precipitação média anual e o coeficiente de escoamento. Para o método WLS foram a área de drenagem e o comprimento do rio principal e para o método GLS, foram a área de drenagem, o comprimento do talvegue e a declividade equivalente. Por fim, para a região BA4 os três métodos tiveram a inserção, em seus modelos das variáveis, a área de drenagem, a precipitação média anual e o coeficiente de escoamento.

Na análise dos resíduos dos modelos do OLS por meio dos testes de hipótese, evidenciou-se que para as regiões BA2 e BA4 a utilização do método OLS não é recomendada, pois pelo menos uma premissa referente aos resíduos foi violada. Nas regiões, BA1, BA3 e BA4, para ambos os métodos foi perceptível a presença de possíveis outliers, indicando a necessidade de avaliação dos pontos influentes dos modelos de regressão. Esta avaliação pode ser feita através de medidas de influência, tais como leverage, Cook's D, DFITS e DFBETAS.

Na região BA3 para os modelos do WLS e GLS, observou-se uma forte tendência da variância do erro do modelo em convergir para zero e conseqüentemente o  $R^2$  igual para 1. No entanto, Reis Jr. (2005) relatou que além de um valor de  $R^2$  igual a zero ser irreal, os modelos com variância do erro do modelo também o são. Por mais que pareça que um modelo com erro igual a zero seja perfeito, não faz sentido a inexistência de erro em um modelo, dado que a estimativa regional não é determinística.

A região BA4, para ambos os métodos WLS e GLS, foi a região onde se observou os maiores valores da variância do erro do modelo. Porém, esta teve os menores valores de  $R^2$ .

A região BA3 foi a região onde se observou os menores valores do erro do modelo e, os maiores  $R^2$  e, de acordo com, Kroll e Stedinger (1998) para pequenos valores da variância do erro do modelo, a eficiência do método OLS diminui conforme as correlações cruzadas aumentam. Neste caso, o método GLS torna-se mais adequado. Nesta região, não se verificou grandes distorções para os valores das  $Q_{7,10}$  máximas estimadas e observadas. Como em grande parte, nesta região os valores do erro do modelo convergiram para zero ou próximo de zero e nesta região a variância do erro amostral foi extremamente elevada. Sendo assim, sugere-se a aplicação do GLS bayesiano, assim como apresentado em Reis Jr. (2005).

Os valores do erro relativo médio foi elevado para os métodos WLS e GLS. No entanto, somente a avaliação deste parâmetro não garante que o método OLS tenha sido superior ao WLS e GLS pois, como foi verificado anteriormente, existem outras premissas a serem avaliadas para corroborar que um modelo é melhor que o outro. Todavia os valores elevados

do erro relativo médio para os métodos WLS e GLS podem estar associados ao fato do modelo GLS atribuir ponderações diferentes para cada estação, ou seja, valores variados da covariância amostral.

Ao analisar a razão entre as variâncias do erro (EVR), verificou-se que a variância do erro amostral é demasiadamente elevada quando comparado à variância do erro do modelo. Contudo, realizou-se uma análise das estações que possuíam variância do erro amostral elevada e, mesmo que se fizesse a retirada destas estações, ainda assim a variância do erro amostral foi maior que a variância do erro do modelo, confirmando que a regressão OLS não deve ser empregada. De acordo com Griffis e Stedinger (2007), se a EVR for superior a 20%, o modelo OLS deve ser rejeitado.

A fim de que se pudesse comparar os métodos WLS e GLS, em relação a qual o método seria mais apropriado para ser empregado, calculou-se a VIR e, segundo Griffis e Stedinger (2007), se a VIR é significativamente maior que 1, o método GLS deve ser empregado, caso contrário o método WLS é suficiente. Na aplicação aqui realizada os valores encontrados foram da ordem de 0,12. Logo a regressão WLS é suficiente. Contudo, recomenda-se que seja avaliado se, realmente para o caso da  $Q_{7,10}$ , a covariância estimada teria maior peso sobre a estimativa do valor constante.

Ficou evidenciado que os modelos que possuíam pequenos valores de variância do erro, tiveram valores elevados do erro relativo médio. De acordo com Reis jr. (2005), valores pequenos de variância subestimam o valor da variância de predição e, automaticamente, os resultados finais dos métodos ponderados (WLS e GLS) são afetados.

No entanto, o método GLS se destaca porque este é o modelo mais sensível no que tange à distribuição espacial das estações. Nas regiões onde as estações se encontram próximas, o modelo GLS reduziu significativamente a variância do erro do modelo e forneceu equações com menor erro dos parâmetros.

Por fim, Salienta-se o método GLS é mais adequado, pois as vazões mínimas são fortemente correlacionadas, uma vez que método WLS não leva em consideração as correlações entre os fluxos concorrentes. Porém, para o caso das correlações serem pequenas, tal método poderia ser adotado. No entanto, na região estudada, os valores das correlações cruzadas são elevadas. Sendo assim, conclui-se que para o estudo em questão, o método GLS é o mais adequado. Recomenda-se entretanto, que sejam feitas avaliações concernentes à matriz de covariância

dos erros amostrais, pois a metodologia para estimativa da matriz de covariância é baseada nos quantis propriamente ditos, e não no logaritmo dos quantis. Essa transformação não linear certamente causa uma mudança nos elementos da matriz de covariância. Porém, não foi possível, até esse momento, avaliar as repercussões desse fato nos resultados finais do estudo de regressão. Sendo assim, os resultados apresentados aqui não devem ser vistos como definitivos, mas como ilustração de uma metodologia aproximada para a utilização dos métodos WLS e GLS. Sendo assim, os resultados apresentados aqui não devem ser vistos como definitivos, mas como ilustração de uma metodologia aproximada para a utilização dos métodos WLS e GLS.

Recomenda-se também, que a matriz de covariância seja avaliada através do emprego de outras distribuições de probabilidades que, assim como a Weibull, também sejam boas para modelar vazões mínimas, como por exemplo, as distribuições Gumbel e Log-Normal 2 e 3 parâmetros.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- ACREMAN, M. C.; SINCLAIR, C. D. Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland. *Journal of Hydrology*, n.84, 365–380, 1986.
- AGÊNCIA NACIONAL DE ÁGUAS (ANA). Caderno de Recursos Hídricos. Diagnóstico da outorga de direito de uso de recursos hídricos no país - Diretrizes e prioridades. Brasília: ANA, 2005.
- AHEARN, E. A. Regional regression equations to estimate flow-duration statistics at ungaged stream sites in Connecticut. *U.S. Geological Survey, Scientific Investigations Report 2010-5052*, in cooperation with the Connecticut Department of Environmental Protection, 2010.
- ALEXANDRE, A. M.; MARTINS, E. S. P. R. Regionalização de vazões médias de longo período para o Estado do Ceará. *Revista Brasileira de Recursos Hídricos*, v.10, n.3, p. 93-102, 2005.
- ANEEL. AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. *Atlas Hidrológico Brasileiro – Versão 1.0*. ANEEL, Brasília, CD. ROM. 1998.
- ARMENTROUT, G. W.; WILSON, J. F. Assessment of low flows in streams in northeastern Wyoming. *U. S. Geological Survey, Water Resources Investigations Report*, 85-4246, 1987.
- AZEVEDO, A. A. Avaliação de metodologias de regionalização de vazões mínimas de referência para a sub-bacia do rio Paranã. 2004. Dissertação (mestrado) - Universidade Federal de Viçosa, Viçosa - MG, 2004.
- BAENA, L. G. N. Regionalização de vazões para a bacia do rio Paraíba do Sul, a montante de Volta Redonda, a partir de modelo digital de elevação hidrologicamente consistente. 2002. Dissertação de Mestrado – Universidade Federal de Viçosa, Viçosa, 2002.
- BAYAZIT, M.; ÖNÖZ, B. Sampling variances of regional flood quantiles affected by intersite correlation. *Journal of Hydrology*, v. 291, p. 42-51, 2004.
- BERTONI, J. C.; TUCCI, C. E. M. Precipitação. In: TUCCI, C. E. M. *Hidrologia: Ciência e aplicação*. Porto Alegre: UFRGS, 2007. P. 177-241.
- CABRAL, J. *et al.* Recursos hídricos subterrâneos. In: PAIVA, J. B.; PAIVA, E. M. (coord). *Hidrologia aplicada à gestão de pequenas bacias hidrográficas*. Recope-Rehidro sub rede 1 Finep. UFSM, Santa Maria. ABRH, 2001. Cap. 9, p.237-277.
- CATALUNHA, M. J. Sistema integrado em rede para gestão do uso múltiplo da água e regionalização da  $Q_{7,10}$  para os períodos mensal, bimestral, trimestral e anual. 2004. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2004.
- CHAVES, H. M. L.; ROSA, J. W. C.; VADAS, R. G.; OLIVEIRA, R. V. T. Regionalização de vazões mínimas em bacias através de interpolação em sistemas de informações geográfica. *Revista Brasileira de Recursos Hídricos*, v.7, n. 3, p. 43-51, 2002.
- CHIANG, S. L.; JOHNSON, F. W. Low flow criteria for diversions and impoundments. *Journal of the Water Resources Planning and Management Division*, 102: 227-238, 1976.
- CHOW, V. T. Handbook of applied hydrology. McGraw-Hill, 1964.

- CHOWDHURY, J. E.; STEDINGER, J. R. Confidence interval for design floods with estimated skew coefficient. *Journal of Hydraulic Engineering*, v. 117, n. 7, p. 811-831, 1991.
- CIĞIZOĞLU, H. K.; BAYAZIT, M. Application of gamma autoregressive model to analysis of dry periods. *Journal of Hydrologic Engineering*, v. 3, n. 3, 218–221, 1998.
- CLAUSEN, B. Modelling streamflow recessions in two Danish streams. *Nordic Hydrology*, v. 23, n. 2, p. 73–88, 1992.
- CPRM. COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. *Programa de água subterrânea para a região Nordeste. Projeto aluviões e coberturas recentes no Nordeste oriental*. Fase I Coleta de dados. CPRM/SGB. Recife, 2008.
- CPRM. COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. Projeto Análise de Consistência de Dados Pluviométricos – Bacia do Rio São Francisco. Sub-Bacias 40 a 44: Relatório Técnico Parcial. Sub-Bacia 40. Belo Horizonte, DNAEE/CPRM, 1995a, 3v. Citado por CPRM. Companhia de Pesquisa de Recursos Minerais. *Regionalização de vazões. Sub-bacias 40 e 41: Relatório Final*. CPRM/ANEEL. Belo Horizonte, 2001, 1v.
- CPRM. COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. Projeto Análise de Consistência de Dados Pluviométricos – Bacia do Rio São Francisco. Sub-Bacias 40 a 44: Relatório Técnico Parcial. Sub-Bacia 41. Belo Horizonte, DNAEE/CPRM, 1995b, 3v. Citado por CPRM. Companhia de Pesquisa de Recursos Minerais. *Regionalização de vazões. Sub-bacias 40 e 41: Relatório Final*. CPRM/ANEEL. Belo Horizonte, 2001.
- CPRM. COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. *Projeto para avaliação dos recursos de águas subterrâneas em Iberoamérica*. Informações hidrogeológicas. CPRM/SGB. Rio de Janeiro, 2008
- CPRM. COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. *Regionalização de vazões. Sub-bacias 40 e 41: Relatório Final*. CPRM/ANEEL. Belo Horizonte, 2001.
- CUNNANE, C. Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, v.100, p.269-290, 1998.
- D'AGOSTINO, R. B. e STEPHENS, M. *Goodness-of-fit Techniques*. 1.ed. Dekker, 1986.
- DALRYMPLE, T. Flood-frequency analyses, Manual of Hydrology: Part. 3. Flood-flow Techniques. *Geological Survey Water Supply paper 1543-A*. U. S. Government Printing Office, Washington, D. C., 1960.
- DAWDY, D. R.; GRIFFIS, V. W.; GUPTA, K. V. Regional flood frequency analysis: How We got here and where we are going. *Journal of Hydrologic Engineering*, doi: 10.1061(ASCE)HE.1943-5584.0000584, 2012.
- DINGMAN, S. L.; LAWLOR, S. C. Estimating low-flow quantile from drainage-basin characteristics in New Hampshire and Vermont. *Water Resources Bulletin* 31, p. 243–256, 1995.
- DODGE, Y.; ROUSSON, V. The complications of the fourth central moment. *The American Statistician*, v. 53, n. 2, p. 267-269, 1999.
- ELETROBRÁS. Centrais Elétricas Brasileiras S.A. Metodologia para regionalização de vazões. Rio de Janeiro, v. 1, 1985.
- EUCLYDES, H. P.; FERREIRA, P. A.; RUBERT, O. A. V.; SANTOS, R. M. Regionalização hidrológica na bacia do Alto São Francisco a Montante da Barragem de Três Marias, Minas Gerais. *Revista Brasileira de Recursos Hídricos*, v. 6, n. 2, p. 81-105, 2001.

- FEASTER, T. D.; TASKER, G. D. Techniques for estimating the magnitude and frequency of floods in rural basins of South Carolina, 1999. *U. S. Geol. Surv. Water Resources Research*. Invest. Rep., 02-4140, 2002.
- FLYNN, R. H. A Stream-gaging Network Analysis for the 7-Day, 10-year Annual Low Flow in New Hampshire Streams. *U.S. Geological Survey, Water-Resources Investigations Report* 03-4023, 2003.
- GIESE, G. L.; MASON, R. R. Low-flow characteristics of streams in north Carolina. *U. S. Geological Survey, Water-Supply Paper* 2403, 1993.
- GORDON, N. D.; MCMAHON, T. A.; FINLAYSON, B. L. *Stream Hydrology* (An Introduction for Ecologists). Wiley, New York, 1992, 526 p.
- GRIFFIS, V. W. *Flood Frequency analysis: Bulletin 17, regional information, and climate change*. 2006. Tese (Ph. D. thesis) - Faculty of the Graduate School of Cornell University, Ithaca, New York, 2006.
- GRIFFIS, V. W.; STEDINGER, J. R. The use of GLS in regional hydrologic analyses. *Journal of Hydrology*, v. 344, p. 82-95, 2007.
- GUJARATI, D. N.; PORTER, D. C. *Econometria básica*. 5. ed. McGraw-Hill, 2008.
- GUSTARD, A.; BULLOCK, A.; DIXON, J. M. Low flow estimation in the United Kingdom. *Institute of Hydrology Report No. 108*, 1992.
- HADDAD, K., RAHMAN, A. AND KUCZERA, G. Comparison of Ordinary and Generalised Least Squares Regression Models in Regional Flood Frequency Analysis: A Case Study for New South Wales (technical paper). *Australian Journal of Water Resources*, v. 15, n.1, p. 59-70, 2011.
- HADDAD, K.; RAHMAN, A. STEDINGER, J. R. Regional flood frequency analysis using Bayesian generalized least squares: a comparison between quantile and parameter regression techniques. *Hydrological Processes*, 26, p. 1008-1021, 2012.
- HADDAD, K; RAHMAN, A. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework - Quantile Regression vs. Parameter Regression Technique. *Journal of Hydrology*, 430-41, p. 142-161, 2012.
- HARRIS, N. M.; GURNELL, A. M.; HANNAH, D. M.; PETTS, G. E. Classification of river regimes: a context for hydroecology. In: John Wiley & Sons, Hardcover: *Hydrological Processes*. 2000, v.14, p.2831-2848.
- HIRUMA, S. T.; PONÇANO, W. L. Densidade de drenagem e sua relação com fatores geomorfológicos na área do alto rio Pardo, SP e MG. *Revista do Instituto Geológico*. São Paulo, 15(12), p. 49-57, 1994.
- HOSKING, J. R. M.; WALLIS, J. R. *Regional Frequency Analysis* - an approach based on L moments. Cambridge University Press, Cambridge, 1997.
- HUGHES, D. A.; O'KEEFFE, J. H.; SMAKHTIN, V. Y.; KING, J. Development of an operating rule model to simulate time series of reservoir releases for instream flow requirements. *Water SA*, v.23, n. 1, 21-30, 1997.
- IACWD. Interagency Committee on Water Data. Guidelines for Determining Flood Flow Frequency: Bulletin 17B (revised and corrected), 28 pp., Hydrol. Subcommittee., Washington, D.C., 1982.



- IGAM. Instituto Mineiro de Gestão das Águas. Estudo de regionalização de vazão para o aprimoramento do processo de outorga no Estado de Minas Gerais / Grupo de Pesquisas em Recursos Hídricos da UFV. Belo Horizonte, 2012.
- KAVISKI, E. Vazões de estiagens em pequenas bacias hidrográficas do estado de Santa Catarina. In: Simpósio Brasileiro de Hidrologia e Recursos Hídricos, 5, 1983. Florianópolis. *Anais ...* Florianópolis: ABRH, 1983, p.43-67.
- KENDALL, M.; STUART, A. *The advanced theory of statistics*. 4. ed. Charles Griffin & Company limited, 1977.
- KITE, G. W. *Frequency and Risk Analysis in Hydrology*. Fort Collins (CO): Water Resources Publications, 1988.
- KJELDTSEN, T. R.; ROSBJERG, D. Comparison of regional index flood estimation procedures based on the extreme value type I distribution. *Stochastic Environ. Res. Risk Assess.*, v. 16, p. 358-373, 2002.
- KROLL, C. N.; STEDINGER, J. R. Development of Regional Regression Relationships with Censored Data. *Water Resources Research*, v. 35, n. 3, 775-784, 1999.
- KROLL, C. N.; STEDINGER, J. R. Development of regional regression relationships with censored data. *Water Resources Research*, v. 35, n. 3, p. 775-784, 1999.
- KROLL, C. N.; STEDINGER, J. R. Regional hydrologic analysis: Ordinary and generalized least squares revisited. *Water Resources Research*, v. 34, n.1, p.121-128, 1998.
- KROLL, C. N.; VOGEL, R. M. Probability distribution of low stream flow series in the united states. *Journal of Hydrologic Engineering*, doi: 10.1061/(ASCE)1084-0699(2002)7:2(137), 2002.
- KUCZERA, G. Effect of sampling uncertainty and spatial correlation on an empirical Bayes procedure for combining site and regional information. *Journal of Hydrology*, 65, p. 373-398, 1983.
- LEMOS, A. F. Avaliação de metodologias de regionalização de vazões mínimas de referência para a bacia hidrográfica do rio São Francisco, a montante do reservatório de Três Maria. 2006. Dissertação (mestrado) - Universidade Federal de Viçosa, Viçosa - MG, 2006.
- LETTENMAIER, D. P.; WALLIS, J. R.; WOOD, E. F. Effect of regional heterogeneity on flood frequency estimation. *Water Resources Research*, v. 23, n. 2, p. 313-324, 1987.
- LIMA, A. A. *Metodologia integrada para determinação da enchente de projeto de estruturas hidráulicas por meio de séries sintéticas de precipitação e modelos chuva- vazão*. 2005. Dissertação de Mestrado (Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.
- MADSEN, H.; MIKKELSEN, P. S.; ROSBJERG, D.; HARREMOES, P. Regional Estimation of Rainfall Intensity-Duration-Frequency Curves Using Generalized Least Squares Regression of Partial Duration Series Statistics. *Water Resources Research*, v. 38, n.11, 1239, 2002.
- MADSEN, H.; ROSBJERG, D. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. *Water Resources Research*, v. 33, n.4, 771-782, 1997.
- MAMUN, A. A.; HASHIM, A.; DAOUD, J. I. Regionalisation of low flow frequency curves for the Peninsular Malaysia. *Journal of Hydrology*, 381, p. 174-180, 2010.

- MARTINS, E. S.; STEDINGER, J. R. Cross correlation among estimators of shape. *Water Resources Research*, v. 38, n. 11, 1252-1258, 2002.
- MARTINS, E. S.; STEDINGER, J. R. Generalized Maximum Likelihood GEV Quantile Estimators for Hydrologic Data. *Water Resources Research*, 28(11), p. 3001-3010, 2000.
- MIDGLEY, D. C.; PITMAN, W. V.; MIDDLETON, B. J. Surface water resources of South Africa 1990. *Water Research Commission Report No 298/5.1/94*, Pretoria, África do Sul, 1994.
- MIKKELSEN, P. S.; MADSEN, H.; REOSBJERG, D.; HAREMOËS, P. Properties of extreme point rainfall III: Identification of spatial inter-site correlation structure. *Atmospheric Research*, 40, p.77-98, 1996.
- MOURÃO, M. A. A.; CRUZ, W. B.; GONÇALVES, R. L. F. Caracterização hidrogeológica da porção mineira da Bacia Hidrográfica do São Francisco. In: PINTO C.P.; MARTINS-NETO, M. A. (coord.) *Bacia do São Francisco: geologia e recursos naturais*. Belo Horizonte: SBG-MG, 2001. p. 327-350.
- MÜLLER, F.; FILL, H. D. Quantis de cheia GEV regionais - Uma aplicação prática. *Revista Brasileira de Recursos Hídricos*, v. 2, n. 2, p. 53-64, 1997.
- NAGHETTINI, M. C.; SPERLING, VON M. Características Hidrológicas de Rios. In: SPERLING, VON M. (Org.). *Estudos e modelagem da qualidade da água de rios*. 1ed. Belo Horizonte: DESA, 2007, v. 7, p. 59-96
- NAGHETTINI, M.; PINTO, E. J. A. *Hidrologia estatística*. 1.ed. Belo Horizonte: CPRM, 2007.552p.
- NATHAN, R. J; McMAHON, T. A. Identification of homogeneous regions for the purposes of regionalization. *Journal of Hydrology*, v.121, p.217-238, 1990.
- NERC. National Environmental Research Council. *Flood Studies Report*, Vol. 1. London: National Environmental Research Council, 1975.
- NOVAES, L. F.; PRUSKI, F. F.; QUEIROZ, D. O.; RODRIGUES, R. D. G.; SILVA, D. D.; RAMOS, M. M. Avaliação do desempenho de cinco metodologias de Regionalização de vazões. *Revista Brasileira de Recursos Hídricos*, v. 12, n.2, p. 51-61, 2007.
- NRC - U. S. NATIONAL RESEARCH COUNCIL. *Estimating probabilities of extreme floods*. Washington, D.C.: National Academy Press. 141p. 1988.
- OLIVEIRA, L. F. C.; FIOREZE, A. P.; MEDEIROS, A. M. M.; SILVA, M. A. S. Comparação de metodologias de preenchimento de falhas de séries históricas de precipitação pluvial anual. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 14, n.11, p. 1186-1192, 2010.
- PAPOULIS, A. Probability random variables, and stochastic processes. McGraw-Hill, 1965.
- PEREIRA, S. B.; PRUSKI, F. F.; NOVAES, L. F.; SILVA, D. D.; RAMOS, M. M. Distribuição espacial das variáveis hidrológicas na bacia do rio São Francisco. *Engenharia na Agricultura*, Viçosa, v.11, n1-4, 2003.
- PINTO, E. J. A.; ALVES, M. M. S. *Regionalização de vazões da bacia do Alto do São Francisco*. In: XV SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 2003, Curitiba. *Anais...Curitiba: ABRH, 2003, CD-ROM*.
- PINTO, J. A. O. *Avaliação de métodos para a regionalização de curvas de permanência de vazões para a bacia do rio das Velhas*. 2006. Dissertação de Mestrado (Programa de

- Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.
- PRESS, F.; SIEVER, R.; GROTZINGER, J.; JORDAN, T. H. *Para entender a Terra*. 4. ed. Porto Alegre: Bookman, 2006.
- PYRCE, R. Hydrological low flow indices and their uses. *Watershed Science Centre*, n. 4, 37p., 2004.
- RAMEH, I. M. B., FREITAS, V. A. L., CIRILO, J. A. *Importância da escolha das variáveis explicativas na regionalização de vazões em áreas com escassez de dados*. In: XV Simpósio Brasileiro de Recursos Hídricos. Curitiba. *Anais...*Curitiba: ABRH, 2003. CD-ROM.
- RAO, A. R.; HAMED, K. H. *Flood Frequency analysis*. CRC Press, 2000.
- RAO, A. RAMACHANDRA; SRINIVAS, V. V. Regionalization of watersheds by hybrid-cluster analysis. *Journal of Hydrology*, v. 318, n. 1-4, p. 37-56, 2006.
- RAO, C. R.; TOUTENBURG, H. *Linear Models: Least Squares and Alternatives*. 2. ed. New York: Springer, 1999.
- REIS Jr., D. S. *Flood frequency analysis employing bayesian regional regression and imperfect historical information*. 2005. Tese (Ph. D. thesis) - Faculty of the Graduate School of Cornell University, Ithaca, New York, 2005.
- REIS Jr., D. S.; STEDINGER, J. R.; MARTINS, E. S. Bayesian GLS Regression with application to LP3 Regional Skew Estimation. *Water Resources Research*, v. 41, W10419, 2005.
- REIS Jr., D. S.; STEDINGER, J. R.; MARTINS, E. S. *Operational Bayesian GLS regression for regional hydrologic analyses*. In: World Water Congress, 2004. Formato: <http://www.ascelibrary.org>. Acesso em: agosto de 2012.
- REIS, A. A. Estudo comparativo, aplicação e definição de metodologias apropriadas para a determinação da vazão ecológica na bacia do rio Pará, em Minas Gerais. 2007. Dissertação de Mestrado (Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos) - Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.
- RIES III, K. G. Development and application of Generalized Least Squares Regression Models to estimate Low flow duration discharges in Massachusetts. *U. S. Geological Survey, Water-Resources Investigations Report 94-4155*, 1994.
- RIGGS, H. C. Low-Flow Investigations, U.S. Geol. Surv. Techniques of Water Resource Investigations. Book 4, Chapter B1, *U.S. Geological Survey*, 1972.
- RURALMINAS - Fundação Rural Mineira, Colonização e Desenvolvimento Agrário. Hidrogeologia. In: Plano Diretor de Recursos Hídricos da Bacia de Afluentes do Rio São Francisco, em Minas Gerais. Belo Horizonte, 1999. Citado por MOURÃO, M. A. A.; CRUZ, W. B.; GONÇALVES, R. L. F. Caracterização hidrogeológica da porção mineira da Bacia Hidrográfica do São Francisco. In: PINTO C.P.; MARTINS-NETO, M. A. (coord.) *Bacia do São Francisco: geologia e recursos naturais*. Belo Horizonte: SBG-MG, 2001. p. 327-350.
- SILVA JÚNIOR, O. B.; BUENO, E. O.; TUCCI, E. M.; CASTRO, M. R. Extrapolação espacial na regionalização da vazão. *Revista Brasileira de Recursos Hídricos*, v. 8, n. 1, p 21-37, 2003.

- SILVA, A. M.; OLIVEIRA, P. M.; MELLO, C. R.; PIERANGELI, C. Vazões mínimas de referência para outorga na região do Alto Rio Grande, Minas Gerais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 10, n.2, p. 374-380, 2006.
- SILVA, B. C.; TIAGO FILHO, G. L.; SILVA, A. P. M.; de PAULO, R. G. F. *Regionalização de vazões na bacia do Rio Grande*. In: XV SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 2003, Curitiba. *Anais...Curitiba*: ABRH, 2003, CD-ROM.
- SILVA, B. M.; CLARKE, R. T. Análise estatística de chuvas intensas na bacia do rio São Francisco. *Revista Brasileira de Meteorologia*, v. 19, n.3, p. 265-272, 2004.
- SINGB, K. P.; RAMAMURTHY, G. S.; SEO, I. W. 7-Day 10-year low flows of streams in the Rock, Spoon, La Moine, and Kaskaskia regions. *Illinois State Water Survey Contract Report 440*, 1988.
- SINGH, V. P. *Elementary Hydrology*, 1º ed. 973 p, Prentice Hall. 1997.
- SMAKHTIN, V. U. Low flow hydrology: a review. *Journal of Hydrology*, n. 240, p. 147-186, 2001.
- SMAKHTIN, V. Y.; HUGHES, D. A.; CREUSE-NAUDIN, E. Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa. *Hydrological Sciences*, v. 42, n.6, 1997.
- SOUZA, F. A. S.; LIMEIRA, R. C. Uso de Análise Multivariada na regionalização de cheias do Estado de São Paulo. *Revista Brasileira de Recursos Hídricos*, v. 11, n.4, p. 15-24, 2006.
- STEDINGER, J. R. Flood frequency analysis and statistical estimation of flood risk. In: WOHL, E. E. (Ed.) *Inland flood hazards: human, riparian, and aquatic communities*. Cambridge University Press, Cambridge, UK, p. 334-358, 2000.
- STEDINGER, J. R.; LU, L. A. Appraisal of regional and index-flood quantile estimators. *Stochastic Hydrology and Hydraulics*, v.9, n. 1, p. 49- 75, 1995.
- STEDINGER, J. R.; TASKER, G. D. Correction to Regional hydrologic analysis 1. Ordinary, weighted, and generalized least squares compared. *Water Resources Research*, v. 22, n. 5, p. 844, 1986a.
- STEDINGER, J. R.; TASKER, G. D. Regional Hydrologic Analysis, 2 Model-Error Estimators, Estimation of Sigma and Log-Pearson Type 3 Distributions. *Water Resources Research*, v. 22, n. 10, p. 1487, 1499, 1986b.
- STEDINGER, J. R.; TASKER, G. D. Regional hydrologic analysis 1: Ordinary, weighted, and generalized least squares compared. *Water Resources Research*, v. 21, n.9, p. 1421-1432, 1985.
- SUGIYAMA, H. Analysis and extraction of low flow recession characteristics. *Water Resources Bulletin*, 32 (3), p. 491-497, 1996.
- TALLAKSEN, L. M. A review of baseflow recession analysis. *Journal of Hydrology*, 165, p.349-370, 1995.
- TASKER, G. D. A comparison of methods for estimating low flow characteristics of streams. *Water Resour. Bull.*, v. 23, n.6, p.1077-1083, 1987.
- TASKER, G. D. Hydrologic Regression With Weighted Least Squares. *Water Resources Research*, v. 16, n. 6, p.1107-1113, 1980.
- TASKER, G. D.; EYCHANER, J. H.; STEDINGER, J. R. Application of generalized least squares in regional regression analysis. In: S. S. (Ed.) *Selected papers in Hydrologic Science*. U. S. Geological Survey Water-Supply Paper 2310, p.107-115, 1986.

- TASKER, G. D.; STEDINGER, J. R. An operational GLS model for hydrologic regression. *Journal of Hydrology*, 111, 361-375. 1989.
- TASKER, G. D.; STEDINGER, J. R. Regional Skew with Weighted LS Regression, *Journal of Water Resources Planning and Management*, v. 112, n. 2, p. 225-237, 1986.
- TAYLOR, M.; HADDAD, K.; ZAMAN, M.; RAHMAN, A. *Regional flood modelling in Western Australia: Application of regression based methods using Ordinary Least Squares*. International Congresso on Modelling and Simulation, 19th, 2011, Perth, Austrália. Austrália: MODSIM2011, 2011, p.43-67. Formato: <http://mssanz.org.au/modsim> 2011. Acesso em: setembro de 2012.
- TEIXEIRA, W.; TOLEDO, M. C. M.; FAIRCHILD, T. M.; TAIOLI, F. *Decifrando a terra*. São Paulo: Oficina de Textos, 2001. 558 p.
- THOMAS, D. M.; BENSON, M. A. Generalization of streamflow characteristics, from drainage-basin characteristics. *Geological Survey Water-Supply Paper* 1975, 55 pp., 1970.
- TUCCI, C. E. M. Regionalização de vazões. In: Tucci, C. E. M. (Org.) *Hidrologia: ciência e aplicação*. 2.ed. Porto Alegre: ABRH/UFRGS, 2001. p.573-620.
- TUCCI, C. E. M. *Regionalização de vazões*. Porto Alegre, ANEEL-UFRGS, 2002.
- TUCCI, C.; SILVEIRA, A.; SANCHEZ, J.; ALBUQUERQUE, F. Flow regionalisation in the upper Paraguay basin, Brazil. *Hydrological Sciences*. 40 (4), 485–497, 1995.
- USWRC - United States Water Resources Council. *Guidelines for determining flood flow frequency, Hydrology Committee, Bulletin 17 B (revised)*. U. S. Government Printing Office, Washington DC, 1982.
- VEILLEUX, A. G. *Bayesian GLS regression, Leverage, and influence for regionalization of hydrologic statistics*. 2011. Tese (Ph. D. thesis) - Faculty of the Graduate School of Cornell University, Ithaca, New York, 2011.
- VEILLEUX, A. G.; STEDINGER, J. R.; EASH, D. A. *Bayesian WLS/GLS Regression for Regional Skewness analysis for regions with large Crest Stage gage networks*. World Environmental and Water Resources Congress, 2012, Crossing Boundaries, ASCE, 2012.
- VOGEL, R. M.; KROLL, C. N. Generalized low-flow frequency relationships for ungaged sites in Massachusetts. *Water Resources Bull.*, v.26, n. 2, p. 241-253, 1990.
- VOGEL, R. M.; KROLL, C. N. Low-flow frequency analysis using probability-plot correlation coefficients. *Journal of Water Resources Planning and Management (ASCE)*, v. 115, n. 3, 338–357, 1989.
- VOGEL, R. M.; KROLL, C. N. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research*, 28, p. 2451–2458, 1992.
- VOGEL, R. M.; WILSON, I. Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of Hydrologic Engineering*, v.1, n. 2, 69–76, 1996.
- VORST, P. C.; BELL, F. C. Geomorphic parameters of representative basins and their hydrologic significance. *Australian Water Resources Council Technical Paper* 58:1-131, 1981.
- WALLIS, J. R. Multivariate Statistical Methods in Hydrology - A Comparison Using Data of Known Functional Relationship. *Water Resources Research* 1(4), p. 447-461, 1965.
- WEISBERG, S. *Applied Linear Regression*. 3. ed. New Jersey: Wiley-interscience, 2005.

- WMO - World Meteorological Organization. Guide to Hydrological Practices, Data Acquisition and Processing, Analysis, Forecasting and Other Applications. WMO Rep. 168, Geneva, Switzerland, 1994.
- WMO - World Meteorological Organization. *International Glossary of Hydrology*, WMO, Geneva, 1974.
- WOOLDRIDGE, J. M. *Introductory Econometrics: a modern approach*. 4.ed. South-Western CENGAGE learning, 2009.
- WRIGHT, L. S.; ENSMINGER, P. A. Regionalized regression equations for estimating Low-flow Characteristics for selected Louisiana Streams. *Water Resources Technical Report* No. 75, 2004.
- YEVJEVICH, V. An objective approach to definitions and investigations of continental hydrologic droughts. *Colorado State University*, Fort Collins, Hydrology Paper 23, 18 pp, 1967.
- ZALANTS, M. G. Low-flow frequency and flow duration of selected south Carolina streams through 1987. *U. S. Geological Survey, Water-Resources Investigations Report* 91-4170, 1992.
- ZELHOFER, P.; LIMA, E. B. R.; SANTOS, F. M.; RGO JÚNIOR, L. O. Um ambiente SIG para modelagem integrada da qualidade da água utilizando Qual2e. *Caminhos de Geografia*, v. 8. N.10, p. 107-125, 2003.
- ZHANG, L. Sample mean and sample variance: Their covariance and their (In) dependence. *The American Statistician*, v. 61, n.2, p.159-160, 2007.

# Apêndice - A

## Apêndice A - Anos analisados na análise de frequência local (AF)

<b>Código</b>	<b>Nome</b>	<b>Nº de anos com dados</b>	<b>Nº de anos com falhas</b>	<b>Nº de anos analisados na AF</b>	<b>Anos excluídos na AF</b>
<b>40025000</b>	VARGEM BONITA	57	10	52	5
<b>40032000</b>	FAZENDA SAMBURÁ	38	5	30	8
<b>40037000</b>	FAZENDA DA BARRA	41	1	34	7
<b>40040000</b>	FAZENDA AJUDAS	44	19	42	2
<b>40043001</b>	FAZENDA CAJANGA	25	1	21	4
<b>40046000</b>	PORTO SABINO	18	0	14	4
<b>40050000</b>	IGUATAMA	69	3	66	3
<b>40053000</b>	CALCIOLÂNDIA	40	0	37	3
<b>40056002</b>	FAZENDA CAPOEIRÃO	26	1	23	3
<b>40056200</b>	MONTANTE DO BOM SUCESSO	18	0	16	2
<b>40056500</b>	PONTE CAPOEIRÃO	24	7	19	5
<b>40060001</b>	TAPIRAÍ - JUSANTE	31	0	30	1
<b>40067000</b>	PONTE OLEGÁRIO MACIEL	20	7	16	4
<b>40070000</b>	PONTE DO CHUMBO	36	5	27	9
<b>40080000</b>	TAQUARAL	41	0	36	5
<b>40100000</b>	PORTO DAS ANDORINHAS	49	0	49	0
<b>40102000</b>	PORTO DA BARRA	28	0	26	2
<b>40105000</b>	FAZENDA CAMPO GRANDE	30	11	23	7
<b>40130000</b>	PONTE DO VILELA	21	21	20	1
<b>40150000</b>	CARMO DO CAJURU	71	0	67	4
<b>40160000</b>	LAMOUNIER	26	0	24	2
<b>40170000</b>	MARILÂNDIA (PONTE BR-494)	39	0	36	3
<b>40180000</b>	CARMO DA MATA	37	6	36	1



<b>Código</b>	<b>Nome</b>	<b>Nº de anos com dados</b>	<b>Nº de anos com falhas</b>	<b>Nº de anos analisados na AF</b>	<b>Anos excluídos na AF</b>
<b>40185000</b>	PARI	33	0	33	0
<b>40190002</b>	DIVINÓPOLIS	31	11	27	4
<b>40269900</b>	ITAÚNA - MONTANTE	30	0	27	3
<b>40300000</b>	JAGUARUNA	38	1	35	3
<b>40300001</b>	JAGUARUNA - JUSANTE	28	0	25	3
<b>40330000</b>	VELHO DA TAIPA	68	0	67	1
<b>40350000</b>	USINA CAMARÃO	40	12	39	1
<b>40380000</b>	ARAÚJOS	42	1	39	3
<b>40400000</b>	ESTAÇÃO ÁLVARO DA SILVEIRA	62	9	53	9
<b>40450000</b>	PORTO PARÁ	27	0	19	8
<b>40500000</b>	MARTINHO CAMPOS	39	1	33	6
<b>40505000</b>	ABAIXO USINA PARI	21	1	21	0
<b>40530000</b>	ABAETÉ	41	0	37	4
<b>40549998</b>	SÃO BRÁS DO SUAÇUÍ - MONTANTE	24	0	22	2
<b>40573000</b>	JOAQUIM MURTINHO	21	0	21	0
<b>40577000</b>	PONTE JUBILEU	24	0	23	1
<b>40579995</b>	CONGONHAS - LINIGRAFO	19	0	16	3
<b>40580000</b>	CONGONHAS	41	0	37	4
<b>40665000</b>	USINA JOÃO RIBEIRO	48	0	40	8
<b>40675000</b>	CACHOEIRA DO GORDO	26	0	25	1
<b>40680000</b>	ENTRE RIOS DE MINAS	68	0	65	3
<b>40700002</b>	JECEABA	14	0	14	0
<b>40710000</b>	BELO VALE	41	0	38	3

<b>Código</b>	<b>Nome</b>	<b>Nº de anos com dados</b>	<b>Nº de anos com falhas</b>	<b>Nº de anos analisados na AF</b>	<b>Anos excluídos na AF</b>
<b>40720002</b>	MELO FRANCO	31	0	30	1
<b>40740000</b>	ALBERTO FLORES	43	0	37	6
<b>40770000</b>	CONCEIÇÃO DO ITAGUA	22	3	18	4
<b>40790000</b>	BETIM	26	2	25	1
<b>40800001</b>	PONTE NOVA DO PARAPEBA	68	0	68	0
<b>40810350</b>	FAZENDA LARANJEIRAS	20	10	19	1
<b>40818000</b>	JUATUBA	24	0	23	1
<b>40823500</b>	SUZANA	16	10	15	1
<b>40830000</b>	FAZENDA ESCOLA FLORESTAL	22	2	16	6
<b>40850000</b>	PONTE DA TAQUARA	39	0	37	2
<b>40930000</b>	BARRA DO FUNCHAL	64	3	55	9
<b>40960000</b>	FAZENDA BOM JARDIM	22	0	21	1
<b>40975000</b>	FAZENDA SÃO FÉLIX	37	2	30	7
<b>41050000</b>	MAJOR PORTO	39	0	37	2
<b>41075001</b>	PORTO DO PASSARINHO	28	2	19	9
<b>41151000</b>	FAZENDA ÁGUA LIMPA - JUSANTE	12	0	11	1
<b>41160000</b>	GULPIARA	25	11	22	3
<b>41180000</b>	ITABIRITO - LINIGRAFO	48	2	43	5
<b>41190000</b>	AGUIAR MOREIRA	19	3	13	6
<b>41195000</b>	RIO ACIMA	20	0	20	0
<b>41199998</b>	HONÓRIO BICALHO - MONTANTE	33	0	30	3
<b>41220000</b>	SIDERURGICA	22	3	17	5
<b>41250000</b>	VESPASIANO	64	4	55	9

<b>Código</b>	<b>Nome</b>	<b>Nº de anos com dados</b>	<b>Nº de anos com falhas</b>	<b>Nº de anos analisados na AF</b>	<b>Anos excluídos na AF</b>
<b>41260000</b>	PINHÕES	30	1	28	2
<b>41295000</b>	JOSÉ DE MELO	26	0	25	1
<b>41300000</b>	TAQUARAÇU	64	1	62	2
<b>41340000</b>	PONTE RAUL SOARES	64	5	51	13
<b>41380000</b>	PONTE PRETA	63	0	58	5
<b>41410000</b>	JEQUITIBA	41	0	40	1
<b>41440005</b>	REPRESA - JUSANTE	28	0	26	2
<b>41539998</b>	FAZENDA DA CONTAGEM - MONTANTE	24	0	20	4
<b>41540000</b>	FAZENDA DA CONTAGEM	20	0	18	2
<b>41600000</b>	PIRAPAMA	48	2	46	2
<b>41650002</b>	PONTE DO LICÍNIO - JUSANTE	30	0	28	2
<b>41685000</b>	PONTE DO PICÃO	41	0	38	3
<b>41700001</b>	USINA PARAUNA	50	0	39	11
<b>41715000</b>	FAZENDA CACHOEIRA	24	0	20	4
<b>41720000</b>	FAZENDA DO CIPÓ	27	1	20	7
<b>41780002</b>	PRESIDENTE JUSCELINO - JUSANTE	28	0	28	0
<b>41818000</b>	SANTO HIPÓLITO (ANA/CEMIG)	61	2	57	4
<b>41880000</b>	BUENÓPOLIS	22	0	21	1
<b>41890000</b>	ESTAÇÃO DE CURIMATAI	41	0	36	5
<b>41940000</b>	PONTE DO BICUDO	33	0	26	7
<b>41990000</b>	VÁRZEA DA PALMA	69	0	66	3

# Apêndice - B

## Apêndice B - Características físicas e climáticas das 90 estações utilizadas na regionalização

<b>Código</b>	<b>Nome</b>	<b>Área (km<sup>2</sup>)</b>	<b>L (km)</b>	<b>I equiv (m/km)</b>	<b>P médio (m)</b>	<b>C esc</b>	<b>DD (Junções/km<sup>2</sup>)</b>
<b>40025000</b>	VARGEM BONITA	303,0	37,10	4,550	1,594	0,580	0,079
<b>40032000</b>	FAZENDA SAMBURÁ	763,0	64,00	3,360	1,651	0,470	0,075
<b>40037000</b>	FAZENDA DA BARRA	765,0	72,60	0,600	1,571	0,570	0,081
<b>40040000</b>	FAZENDA AJUDAS	259,0	30,30	2,730	1,508	0,420	0,112
<b>40043001</b>	FAZENDA CAJANGA	502,0	38,00	1,850	1,523	0,410	0,277
<b>40046000</b>	PORTO SABINO	4388,0	115,90	1,310	1,559	0,350	0,085
<b>40050000</b>	IGUATAMA	5426,0	196,00	0,360	1,539	0,400	0,083
<b>40053000</b>	CALCIOLÂNDIA	304,0	40,20	1,790	1,365	0,300	0,063
<b>40056002</b>	FAZENDA CAPOEIRÃO	334,0	59,40	3,220	1,591	0,410	0,096
<b>40056200</b>	MONTANTE DO BOM SUCESSO	334,0	59,60	3,230	1,590	0,430	0,125
<b>40056500</b>	PONTE CAPOEIRÃO	490,0	60,40	3,180	1,640	0,390	0,067
<b>40060001</b>	TAPIRAÍ - JUSANTE	559,0	46,90	3,810	1,795	0,380	0,068
<b>40067000</b>	PONTE OLEGÁRIO MACIEL	8987,0	257,20	0,280	1,535	0,390	0,085
<b>40070000</b>	PONTE DO CHUMBO	9939,0	314,80	0,220	1,522	0,370	0,083
<b>40080000</b>	TAQUARAL	651,0	56,80	1,000	1,425	0,330	0,072
<b>40100000</b>	PORTO DAS ANDORINHAS	13882,0	412,40	0,190	1,494	0,340	0,079
<b>40102000</b>	PORTO DA BARRA	14180,0	433,00	0,200	1,491	0,360	0,079
<b>40105000</b>	FAZENDA CAMPO GRANDE	396,0	25,60	6,590	1,568	0,469	0,152
<b>40130000</b>	PONTE DO VILELA	1726,0	75,80	1,920	1,533	0,330	0,114
<b>40150000</b>	CARMO DO CAJURU	2504,0	121,10	1,120	1,507	0,320	0,107
<b>40160000</b>	LAMOUNIER	159,000	21,700	4,300	1,503	0,400	0,117
<b>40170000</b>	MARILÂNDIA (PONTE BR-494)	1055,0	52,10	2,280	1,479	0,300	0,100
<b>40180000</b>	CARMO DA MATA	144,0	23,60	5,930	1,433	0,440	0,188

<b>Código</b>	<b>Nome</b>	<b>Área (km<sup>2</sup>)</b>	<b>L (km)</b>	<b>I equiv (m/km)</b>	<b>P médio (m)</b>	<b>C esc</b>	<b>DD (Junções/km<sup>2</sup>)</b>
<b>40185000</b>	PARI	1931,0	69,00	1,840	1,460	0,320	0,107
<b>40190002</b>	DIVINÓPOLIS	2000,0	82,70	1,650	1,457	0,340	0,108
<b>40269900</b>	ITAÚNA - MONTANTE	322,0	52,35	2,890	1,515	0,361	0,370
<b>40300000</b>	JAGUARUNA	1275,0	107,85	2,120	1,575	0,307	0,187
<b>40300001</b>	JAGUARUNA - JUSANTE	1558,0	119,10	1,820	1,432	0,310	0,099
<b>40330000</b>	VELHO DA TAIPA	7378,0	219,40	0,750	1,456	0,300	0,036
<b>40350000</b>	USINA CAMARÃO	269,0	40,60	3,160	1,502	0,350	0,130
<b>40380000</b>	ARAÚJOS	1196,0	109,60	0,980	1,445	0,330	0,103
<b>40400000</b>	ESTAÇÃO ÁLVARO DA SILVEIRA	1802,0	148,50	1,030	1,438	0,330	0,100
<b>40450000</b>	PORTO PARÁ	11350,0	292,90	0,460	1,376	0,299	0,114
<b>40500000</b>	MARTINHO CAMPOS	770,0	56,20	1,040	1,393	0,250	0,071
<b>40505000</b>	ABAIXO USINA PARI	881,0	68,10	0,900	1,339	0,271	0,112
<b>40530000</b>	ABAETÉ	481,0	52,70	2,000	1,438	0,330	0,073
<b>40549998</b>	SÃO BRÁS DO SUAÇUI - MONTANTE	461,0	52,00	2,690	1,400	0,380	0,098
<b>40573000</b>	JOAQUIM MURTINHO	291,0	32,70	3,940	1,462	0,300	0,079
<b>40577000</b>	PONTE JUBILEU	244,0	18,30	7,200	1,466	0,330	0,119
<b>40579995</b>	CONGONHAS - LINIGRAFO	579,0	41,60	3,180	1,464	0,360	0,102
<b>40580000</b>	CONGONHAS	627,0	41,60	3,180	1,466	0,384	0,142
<b>40665000</b>	USINA JOÃO RIBEIRO	293,0	45,70	2,440	1,373	0,340	0,123
<b>40675000</b>	CACHOEIRA DO GORDO	189,0	22,95	4,560	1,363	0,427	0,302
<b>40680000</b>	ENTRE RIOS DE MINAS	486,0	47,30	1,250	1,369	0,440	0,136
<b>40700002</b>	JECEABA	2465,0	88,90	1,810	1,141	0,390	0,121
<b>40710000</b>	BELO VALE	2760,0	118,90	1,590	1,408	0,400	0,137

<b>Código</b>	<b>Nome</b>	<b>Área (km<sup>2</sup>)</b>	<b>L (km)</b>	<b>I equiv (m/km)</b>	<b>P médio (m)</b>	<b>C esc</b>	<b>DD (Junções/km<sup>2</sup>)</b>
<b>40720002</b>	MELO FRANCO	3810,0	178,90	1,200	1,472	0,370	0,134
<b>40740000</b>	ALBERTO FLORES	3939,0	187,40	1,210	1,422	0,350	0,134
<b>40770000</b>	CONCEIÇÃO DO ITAGUA	675,0	53,80	2,960	1,485	0,370	0,142
<b>40790000</b>	BETIM	188,0	26,40	3,100	1,460	0,260	0,276
<b>40800001</b>	PONTE NOVA DO PARAOPEBA	5680,0	236,30	1,000	1,449	0,330	0,141
<b>40810350</b>	FAZENDA LARANJEIRAS	10,2	3,46	30,570	1,532	0,347	0,566
<b>40818000</b>	JUATUBA	273,0	40,00	4,520	1,531	0,320	0,064
<b>40823500</b>	SUZANA	156,0	25,10	6,900	1,502	0,344	0,244
<b>40830000</b>	FAZENDA ESCOLA FLORESTAL	84,0	18,00	10,270	1,436	0,350	0,131
<b>40850000</b>	PONTE DA TAQUARA	8734,0	346,30	0,660	1,434	0,320	0,143
<b>40930000</b>	BARRA DO FUNCHAL	876,0	58,70	1,900	1,582	0,430	0,100
<b>40960000</b>	FAZENDA BOM JARDIM	1744,0	162,70	1,120	1,572	0,390	0,088
<b>40975000</b>	FAZENDA SÃO FÉLIX	964,0	142,70	1,380	1,511	0,400	0,076
<b>41050000</b>	MAJOR PORTO	1207,0	60,00	2,290	1,475	0,340	0,031
<b>41075001</b>	PORTO DO PASSARINHO	4143,0	194,60	1,090	1,505	0,380	0,059
<b>41151000</b>	FAZENDA ÁGUA LIMPA - JUSANTE	175,0	26,20	8,590	1,498	0,390	0,115
<b>41160000</b>	GULPIARA	284,0	40,70	5,390	1,498	0,460	0,232
<b>41180000</b>	ITABIRITO - LINIGRAFO	330,0	47,70	5,250	1,518	0,500	0,252
<b>41190000</b>	AGUIAR MOREIRA	539,0	74,90	3,860	1,523	0,430	0,228
<b>41195000</b>	RIO ACIMA	1472,0	83,30	2,930	1,525	0,560	0,218
<b>41199998</b>	HONÓRIO BICALHO - MONTANTE	1698,0	90,30	2,560	1,535	0,360	0,212
<b>41220000</b>	SIDERURGICA	232,0	35,00	7,820	1,446	0,310	0,284
<b>41250000</b>	VESPASIANO	676,0	55,70	2,350	1,329	0,290	0,181

<b>Código</b>	<b>Nome</b>	<b>Área (km<sup>2</sup>)</b>	<b>L (km)</b>	<b>I equiv (m/km)</b>	<b>P médio (m)</b>	<b>C esc</b>	<b>DD (Junções/km<sup>2</sup>)</b>
<b>41260000</b>	<b>PINHÕES</b>	3727,0	156,80	1,420	1,475	0,370	0,204
<b>41295000</b>	<b>JOSÉ DE MELO</b>	291,0	32,60	3,540	1,448	0,360	0,361
<b>41300000</b>	<b>TAQUARAÇU</b>	626,0	46,40	7,020	1,447	0,350	0,272
<b>41340000</b>	<b>PONTE RAUL SOARES</b>	4874,0	200,30	1,130	1,458	0,340	0,209
<b>41380000</b>	<b>PONTE PRETA</b>	553,0	55,40	2,550	1,308	0,310	0,128
<b>41410000</b>	<b>JEQUITIBA</b>	6558,0	288,80	0,760	1,423	0,310	0,178
<b>41440005</b>	<b>REPRESA - JUSANTE</b>	243,0	26,20	5,270	1,349	0,260	0,070
<b>41539998</b>	<b>FAZENDA DA CONTAGEM - MONTANTE</b>	460,0	40,30	1,400	1,346	0,240	0,070
<b>41540000</b>	<b>FAZENDA DA CONTAGEM</b>	420,0	40,30	1,400	1,349	0,247	0,181
<b>41600000</b>	<b>PIRAPAMA</b>	8086,0	360,80	0,610	1,406	0,300	0,158
<b>41650002</b>	<b>PONTE DO LICÍNIO - JUSANTE</b>	10712,0	473,80	0,540	1,360	0,280	0,136
<b>41685000</b>	<b>PONTE DO PICÃO</b>	547,0	53,30	1,160	1,177	0,170	0,033
<b>41700001</b>	<b>USINA PARAUNA</b>	1733,0	80,40	2,750	1,266	0,470	0,100
<b>41715000</b>	<b>FAZENDA CACHOEIRA</b>	85,0	24,30	8,550	1,520	0,550	0,059
<b>41720000</b>	<b>FAZENDA DO CIPÓ</b>	341,0	46,00	4,280	1,454	0,610	0,114
<b>41780002</b>	<b>PRESIDENTE JUSCELINO - JUSANTE</b>	3996,0	100,40	2,480	1,345	0,460	0,096
<b>41818000</b>	<b>SANTO HIPÓLITO (ANA/CEMIG)</b>	16722,0	566,30	0,470	1,330	0,300	0,115
<b>41880000</b>	<b>BUENÓPOLIS</b>	148,0	26,20	11,380	1,108	0,260	0,088
<b>41890000</b>	<b>ESTAÇÃO DE CURIMATAI</b>	1435,0	75,70	1,020	1,108	0,340	0,131
<b>41940000</b>	<b>PONTE DO BICUDO</b>	2023,0	111,50	1,070	1,216	0,250	0,043
<b>41990000</b>	<b>VÁRZEA DA PALMA</b>	26553,0	707,30	0,440	1,271	0,280	0,098



# Apêndice - C

**Apêndice C - Valores estimados da  $Q_{7,10}$**

**Região BA1**

Código	$Q_{7,10}$ observado (m <sup>3</sup> /s)	$Q_{7,10}$ (m <sup>3</sup> /s)		
		Valores estimados		
		OLS	WLS	GLS
41151000	1,037	0,513	0,601	0,563
41160000	0,670	1,328	1,481	1,350
41180000	2,837	1,838	1,986	1,785
41190000	4,535	3,035	2,970	2,548
41195000	7,746	8,925	7,217	5,653
41199998	10,369	10,832	8,482	6,547
41220000	0,740	0,895	1,071	1,005
41250000	1,055	1,079	1,082	0,930
41260000	16,198	17,844	12,327	8,933
41295000	0,880	1,340	1,557	1,439
41300000	1,306	2,588	2,565	2,201
41340000	18,512	22,009	14,582	10,347
41380000	0,795	0,615	0,634	0,552
41410000	21,484	22,518	14,185	9,815
41440005	0,663	0,226	0,257	0,237
41539998	1,010	0,445	0,450	0,391
41540000	0,519	0,728	0,791	0,707
41600000	24,893	23,806	14,422	9,793
41650002	29,861	22,356	12,991	8,632
41685000	0,052	0,111	0,112	0,097
41700001	1,873	1,400	1,171	0,922
41715000	0,170	0,177	0,225	0,222
41720000	0,469	0,823	0,865	0,769
41780002	7,214	5,639	3,884	2,803
41818000	34,687	27,229	14,609	9,336
41880000	0,038	0,029	0,041	0,041
41890000	0,693	0,439	0,418	0,346
41940000	0,142	0,711	0,563	0,433
41990000	37,037	27,975	14,010	8,639

## Região BA2

Código	Q <sub>7,10</sub> observada (m <sup>3</sup> /s)	Q <sub>7,10</sub> (m <sup>3</sup> /s)		
		Valores estimados		
		OLS	WLS	GLS
40549998	1,216	1,473	1,406	1,357
40573000	1,083	0,739	0,836	0,825
40577000	0,966	0,813	0,907	0,890
40579995	1,359	1,940	1,881	1,817
40580000	2,320	2,079	1,797	1,723
40665000	0,829	0,585	0,590	0,575
40675000	0,424	0,472	0,446	0,430
40680000	1,280	1,148	1,664	1,589
40700002	9,104	9,688	6,162	5,800
40710000	10,069	10,315	5,395	5,056
40720002	13,302	11,789	5,049	4,729
40740000	11,529	11,240	4,514	4,238
40770000	1,541	2,061	1,497	1,437
40790000	0,111	0,168	0,151	0,150
40800001	16,789	14,304	4,973	4,659
40810350	0,024	0,017	0,021	0,022
40818000	0,841	0,899	0,910	0,897
40823500	0,316	0,352	0,251	0,246
40830000	0,211	0,267	0,269	0,266
40850000	20,493	18,598	5,785	5,394

## Região BA3

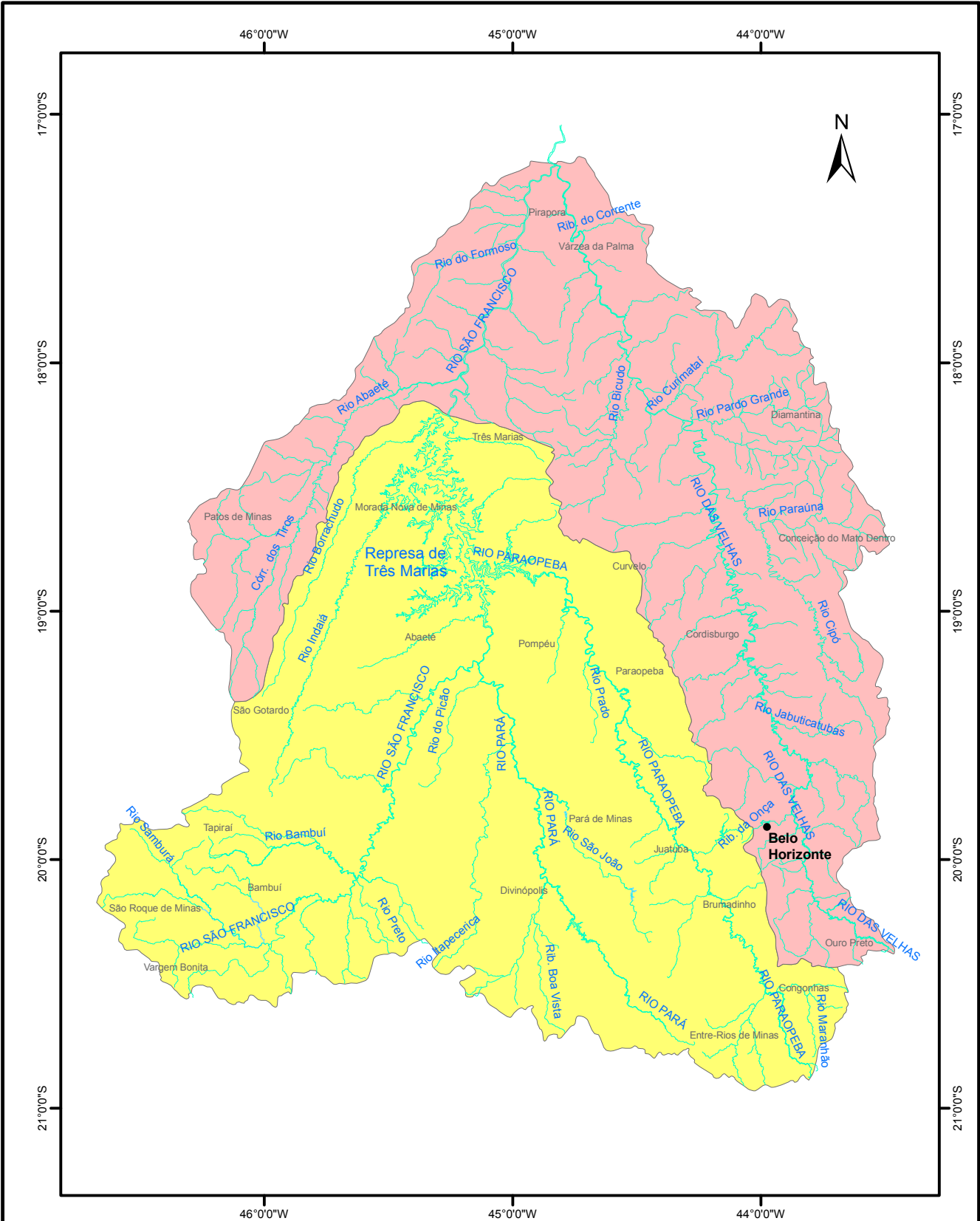
Código	Q <sub>7,10</sub> observada (m <sup>3</sup> /s)	Q <sub>7,10</sub> (m <sup>3</sup> /s)		
		Valores estimados		
		OLS	WLS	GLS
40105000	1,134	1,724	0,741	0,903
40130000	6,685	4,887	3,114	3,325
40150000	10,489	7,769	5,095	5,031
40160000	0,399	0,391	0,427	0,380
40170000	1,477	1,648	1,911	1,901
40180000	0,450	0,363	0,430	0,454
40185000	2,920	3,582	3,088	3,164
40190002	6,409	4,735	3,544	3,662
40269900	1,700	1,059	1,080	1,106
40300000	2,800	4,720	3,402	4,159
40300001	3,641	3,597	4,005	4,770
40330000	17,675	20,566	12,773	13,583
40350000	0,696	0,681	0,835	0,814
40380000	3,428	3,302	3,334	2,850
40400000	3,874	5,575	4,978	4,925
40450000	36,159	26,039	19,075	18,225
40500000	1,146	0,626	1,726	1,194
40505000	0,326	0,771	2,094	1,452

## Região BA4

Código	Q <sub>7,10</sub> observada (m <sup>3</sup> /s)	Q <sub>7,10</sub> (m <sup>3</sup> /s)		
		Valores estimados		
		OLS	WLS	GLS
40025000	1,269	1,451	1,668	1,627
40032000	5,519	4,107	3,730	3,603
40037000	3,106	3,873	3,598	3,416
40040000	0,826	0,345	0,438	0,435
40043001	1,110	0,825	0,888	0,868
40046000	18,073	11,107	7,016	6,521
40050000	20,784	17,102	10,265	9,392
40053000	0,146	0,077	0,104	0,103
40056002	1,020	0,756	0,870	0,865
40056200	1,257	0,834	0,957	0,948
40056500	1,293	1,472	1,515	1,498
40060001	3,500	3,940	3,683	3,670
40067000	24,592	29,707	15,830	14,294
40070000	33,298	27,673	14,531	13,114
40080000	0,447	0,374	0,402	0,392
40100000	35,977	29,227	14,426	12,935
40102000	36,909	33,337	16,331	14,579
40530000	0,032	0,279	0,321	0,317
40930000	1,532	2,659	2,424	2,334
40960000	2,602	4,792	3,738	3,547
40975000	1,353	1,641	1,518	1,457
41050000	2,135	1,208	1,089	1,046
41075001	5,761	8,781	5,740	5,296

# Apêndice - D

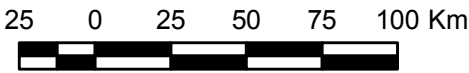
(Mapas da área de estudo e das regiões homogêneas)



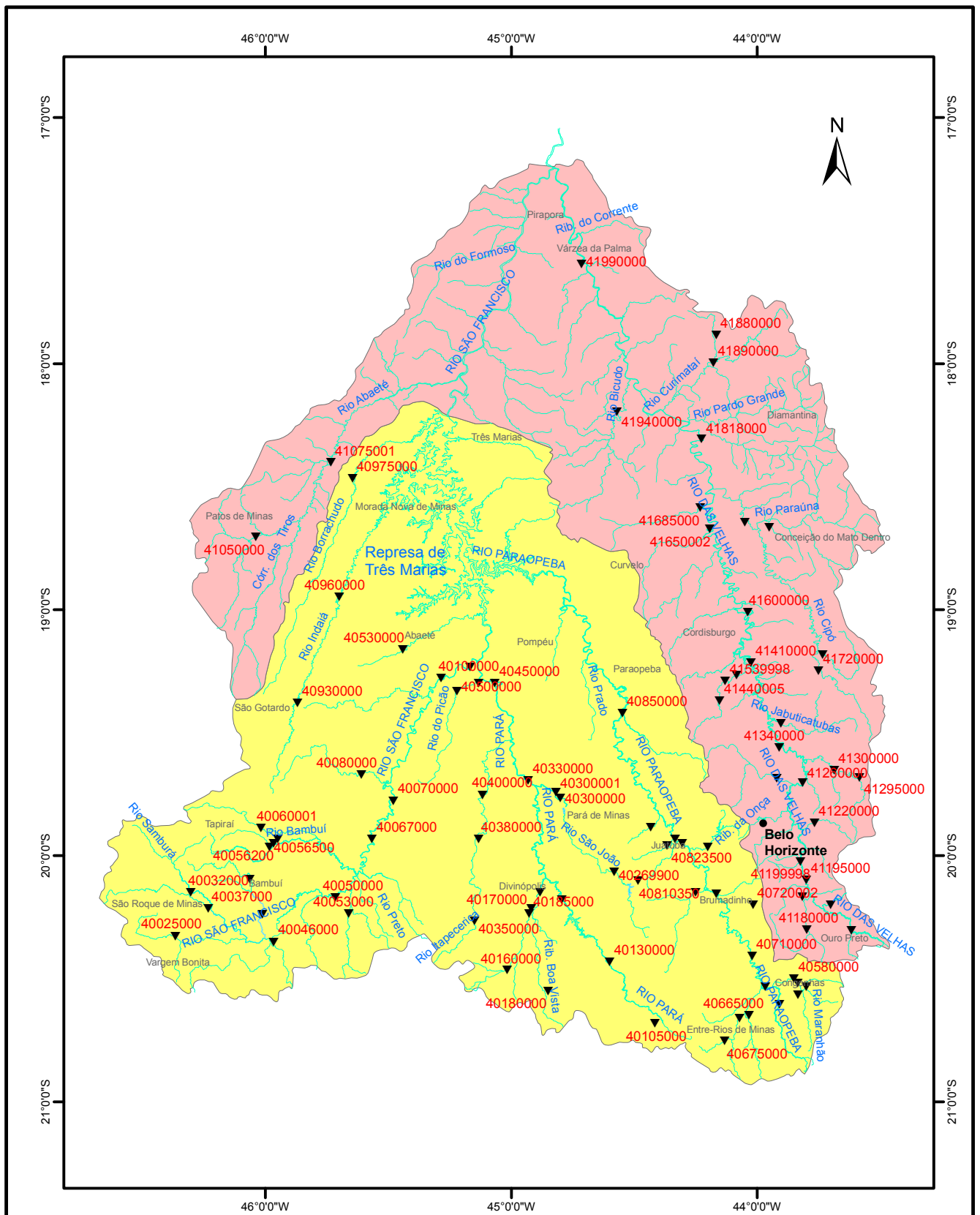
**Sub-bacias 40 e 41 (Classificação ANEEL)**

**LEGENDA**

- Rede de Drenagem
- sub-bacia 40
- sub-bacia 41



Datum: Córrego Alegre



**Mapa de localização da estações fluviométricas nas sub-bacias 40 e 41**

**LEGENDA**

40025000- Código de Estação Fluviométrica

▼ Estações Fluviométricas

— Rede de Drenagem

sub-bacia 40

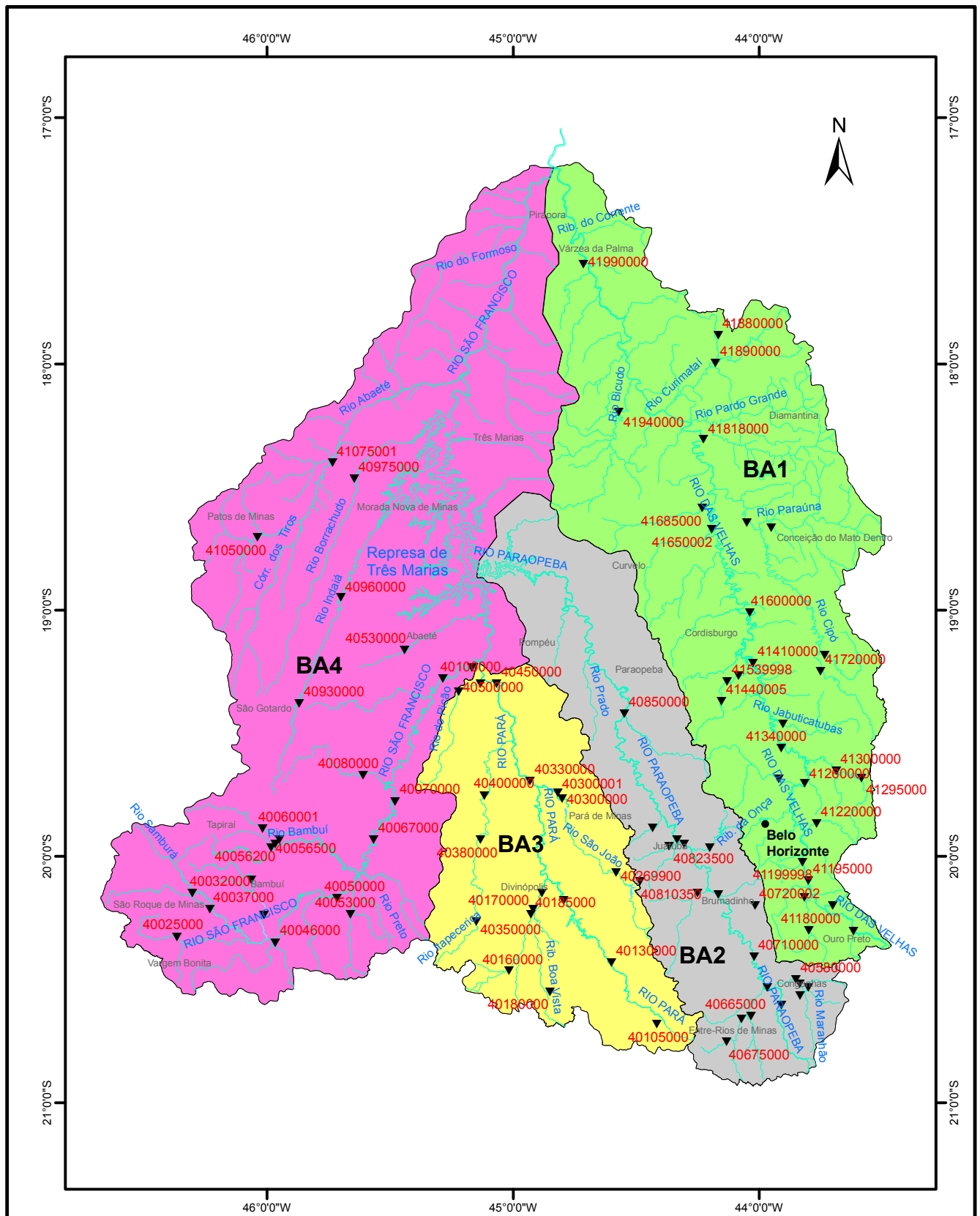
sub-bacia 41

25 0 25 50 75 100 Km



Datum: Córrego Alegre





Mapa de localização da estações fluviométricas nas regiões homogêneas

**LEGENDA**

**BA1** Código atribuído às regiões homogêneas

40025000 - Código de Estação Fluviométrica

▼ Estações Fluviométricas

— Rede de Drenagem

Delimitação das regiões homogêneas

Bacia do rio das Velhas - Código: BA1

Bacia do rio Paraopeba - Código: BA2

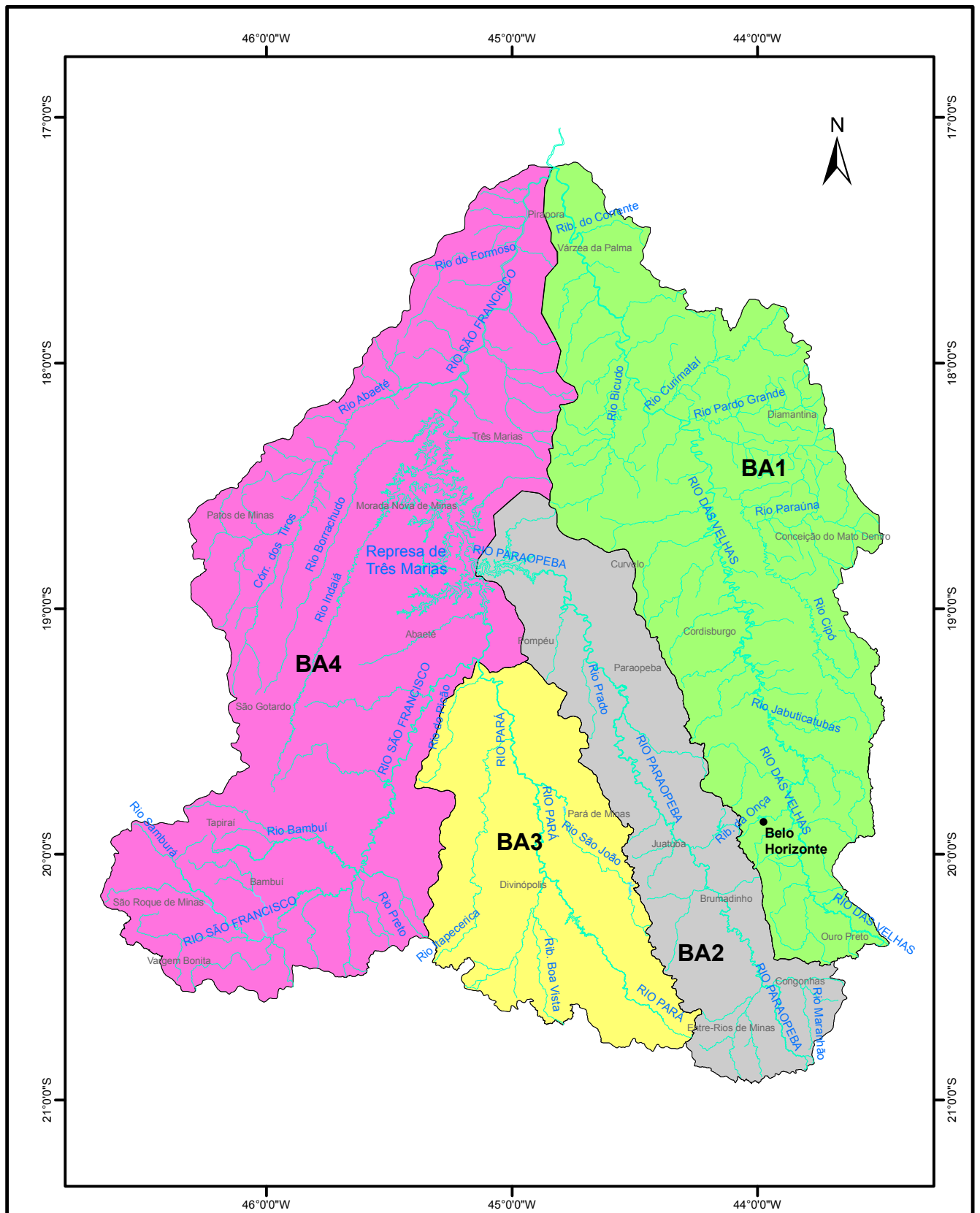
Bacia do rio Pará - Código: BA3

Rio São Francisco, Ajudas, Bambuí, Indaiá, Borrachudo e Abaeté - Código: BA4

25 0 25 50 75 100 Km




Datum: Córrego Alegre





### Regiões homogêneas (agrupamento subjetivo)


#### LEGENDA


**BA1** Código atribuído às regiões homogêneas


 Rede de Drenagem

 Delimitação das regiões homogêneas

 Bacia do rio das Velhas - Código: BA1

 Bacia do rio Paraopeba - Código: BA2

 Bacia do rio Pará - Código: BA3

 Rio São Francisco, Ajudas, Bambuí, Indaiá, Borrachudo e Abaeté - Código: BA4

25 0 25 50 75 100 Km



Datum: Córrego Alegre