

**COOPERAÇÃO EM TECNOLOGIAS PARA  
ANÁLISES HIDROLÓGICAS EM ESCALA  
NACIONAL**

**SUBPROJETO – REGIONALIZAÇÃO DE VAZÕES VIA  
MODELAGEM HIDROLÓGICA**

**RELATÓRIO TÉCNICO: O EFEITO DA  
VARIABILIDADE AMOSTRAL SOBRE A  
ESTIMATIVA DE VAZÕES DE REFERÊNCIA**

**IPH-ANA-HGE-SR-R4**

Porto Alegre - RS

Outubro 2021

# O efeito da variabilidade amostral sobre a estimativa de vazões de referência



**ESTE MATERIAL FAZ PARTE DE UM CONJUNTO DE RELATÓRIOS CRIADOS NO CONTEXTO DO PROJETO DE COOPERAÇÃO EM TECNOLOGIAS PARA ANÁLISES HIDROLÓGICAS EM ESCALA NACIONAL, ENTRE O INSTITUTO DE PESQUISAS HIDRÁULICAS (IPH-UFRGS) E A AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA).**

**AUTORES: Walter Collischonn, Mino Viana Sorribas, Rodrigo Cauduro Dias de Paiva.**

**COMO CITAR: Collischonn, W., Sorribas, M. V., Paiva, R. C. D., 2021. Cooperação em tecnologias para análises hidrológicas em escala nacional: Relatório técnico: O efeito da variabilidade amostral sobre a estimativa de vazões de referência: IPH-ANA-HGE-SR-R4. UFRGS: IPH, [Porto Alegre]. ANA, [Brasília].**

**Porto Alegre - RS**

**Outubro 2021**

## Sumário

1	Apresentação .....	4
2	Introdução .....	5
3	Metodologia .....	5
3.1	Visão geral da metodologia.....	5
3.2	Seleção de postos.....	6
3.3	Cálculo da vazão de referência de série longa.....	7
3.4	Cálculo da vazão de referência de série curta .....	7
3.5	Métricas de erro.....	7
4	Resultados .....	8
4.1	Exemplo ilustrativo dos resultados .....	8
4.2	Efeito do tamanho da série de dados sobre a estimativa da vazão média.....	11
4.3	Efeito do tamanho da série de dados sobre a estimativa da Q95 .....	14
4.4	Efeito de Escala .....	17
4.5	Efeito da localização geográfica das bacias.....	20
5	Conclusões.....	25
6	Referências .....	26

# 1 APRESENTAÇÃO

---

Este relatório refere-se ao produto “IPH-ANA-HGE-SR-R4”, do subprojeto “Regionalizações de vazões via modelagem hidrológica”, no contexto do projeto “Cooperação em tecnologias para análises hidrológicas em escala nacional”, na forma de um Termo de Execução Descentralizada (TED) entre o Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul (IPH/UFRGS) e a Agência Nacional de Águas e Saneamento Básico (ANA).

O projeto está dividido em 4 subprojetos voltados para a (i) capacitação e transferência tecnológica do modelo hidrológico MGB para aplicações e desenvolvimento na ANA, (ii) desenvolvimento e avaliação de métodos para a regionalização de vazões naturais em escala nacional usando modelagem hidrológica (Modelo MGB América do Sul), e desenvolvimento de base de dados de estimativas de vazões e respectivas incertezas em nível nacional, (iii) avaliação de estimativas de áreas inundadas do modelo MGB e (iv) estimativas de fluxos e estoques hidrológicos em escala nacional com base no modelo MGB aplicadas às Contas Econômicas Ambientais da Água (CEAA) no Brasil. Os produtos esperados incluem novas bases de dados, relatórios, manuais técnicos, cursos de capacitação e publicações técnico científicas.

O objetivo proposto para o subprojeto “Regionalização de vazões via modelagem hidrológica” é o desenvolvimento e avaliação de métodos para a regionalização de vazões naturais em escala nacional usando o modelo hidrológico MGB, desenvolvido em escala continental para toda a América do Sul (MGB-AS), além do desenvolvimento de base de dados de estimativas de vazões e estimativa das incertezas em nível nacional em compatibilidade com a base de dados BHO da ANA.

Uma das etapas fundamentais da avaliação é a comparação com métodos clássicos de estimativa de vazões em locais sem dados. Estes métodos clássicos dependem da disponibilidade de séries de dados relativamente longas e simultâneas nos postos fluviométricos, o que nem sempre é possível obter, em muitas aplicações práticas. Neste sentido foi conduzida e está descrita aqui uma análise preliminar com o objetivo de determinar a sensibilidade da estimativa de vazões de referência em locais com dados ao tamanho da série de dados disponível.

## 2 INTRODUÇÃO

---

Vazões de referência são úteis para resumir informações necessárias para a gestão de recursos hídricos. Exemplos de vazões de referência tipicamente utilizadas são a vazão média, e a vazão  $Q_{95}$ , que é a vazão igualada ou superada em um ponto do rio em 95% do tempo.

Em locais monitorados por postos fluviométricos, as vazões de referência podem ser obtidas a partir da série de dados observados. Entretanto, a estimativa da vazão de referência realizada desta forma está sujeita à variabilidade natural da vazão do rio, principalmente em função da variabilidade climática.

A variabilidade climática resulta em períodos de múltiplos anos mais secos ou mais úmidos do que a média de longo prazo. Assim, se a série de dados disponível num local com pouca disponibilidade coincide com um período de anos mais úmidos (secos) do que a média, a vazão de referência possivelmente será ligeiramente maior (menor) do que seria caso houvesse uma série de dados mais longa e mais representativa.

Para minimizar a influência da variabilidade climática sobre a estimativa da vazão de referência, recomenda-se que estimativas de vazões de referência, como a vazão média, a  $Q_{95}$  ou a  $Q_{7,10}$  sejam obtidas com séries de mais de 20 ou 30 anos de dados (Tallaksen e Van Lanen, 2004). Entretanto, nem todos os postos fluviométricos têm séries observadas longas o suficiente para satisfazer esse critério. Além disso, postos fluviométricos com séries de observações de menos do que 20 ou 30 anos podem trazer informações extremamente úteis para a gestão de recursos hídricos, especialmente se estes postos estão localizados em regiões relativamente mal monitoradas (com poucos postos fluviométricos). Assim, em muitos casos a estimativa de vazões de referência é realizada com séries relativamente curtas de dados. Neste sentido, é importante conhecer os potenciais erros da estimativa de vazões de referência utilizando séries de dados observados mais curtas do que o recomendado.

Em um estudo realizado em 132 bacias na Austria, Laaha e Blöschl (2005) observaram um erro médio (RMSE) da estimativa da  $Q_{95}$ , obtida utilizando apenas um ano de dados, de 30% em bacias mais úmidas e de mais de 50% em bacias mais secas, considerando a  $Q_{95}$  obtida com 20 anos de dados como o valor correto. Com 3 anos de dados o erro cai para menos de 20% em bacias úmidas e cerca de 30% em bacias menos úmidas. Finalmente, com 15 anos de dados o erro da estimativa da  $Q_{95}$  é inferior a 10% tanto em bacias secas como bacias úmidas.

Neste capítulo apresentamos a análise do efeito da variabilidade amostral sobre as estimativas de duas vazões de referência (vazão média e  $Q_{95}$ ) a partir de dados de 663 postos fluviométricos do Brasil. O objetivo do trabalho é responder à pergunta: Qual é a acurácia da estimativa da vazão de referência utilizando séries curtas de dados, e como essa acurácia varia em função do tamanho da série de dados?

## 3 METODOLOGIA

---

### 3.1 VISÃO GERAL DA METODOLOGIA

A avaliação do efeito do tamanho da amostra sobre os erros de estimativas das vazões de referência ( $Q_{95}$  e  $Q_M$ ) foi realizada por meio de reamostragem dos dados de postos fluviométricos com disponibilidade de séries longas de observações. Nessa abordagem,

simulações são realizadas de forma a replicar cenários onde cada posto fluviométrico teria apenas uma fração da série longa de dados. Ou seja, a partir de estimativas obtidas de séries curtas extraídas da série original realiza-se uma comparação com a vazão de série longa.

Para isso foram selecionados postos fluviométricos com séries de dados com mais de 30 anos, e com pouca influência de reservatórios. Em cada um destes postos fluviométricos foi calculada a vazão de referência de série longa, obtida com todos os dados disponíveis no período de referência. Posteriormente foram obtidas estimativas da mesma vazão de referência utilizando partes da série original, com apenas N anos de dados, com N variando entre 1 e 30 anos. Para cada valor de N, este procedimento foi repetido 1000 (mil) vezes, com data de início aleatória. Cada uma das 1000 estimativas da vazão de referência para um dado tamanho de série (N) foi comparada com o valor equivalente de série longa através das métricas de erro descritas adiante no texto.

### 3.2 SELEÇÃO DE POSTOS

O estudo foi realizado a partir de dados de séries diárias de vazão de postos fluviométricos com séries longas obtidos da base de dados Hidroweb da ANA. Foram considerados somente postos com séries com mais de 30 anos de dados, ao longo do período entre jan/1980 e dez/2014, sendo no máximo 35 anos de dados. Além disso, considerou-se somente postos sem efeito significativo de regularização artificial e/ou erros grosseiros nas séries de dados, ou de interesse especial para a ANA. Considerando esses critérios, as análises seguiram com uma base de 663 postos fluviométricos, em diferentes regiões do Brasil, ilustrados na Figura 1.

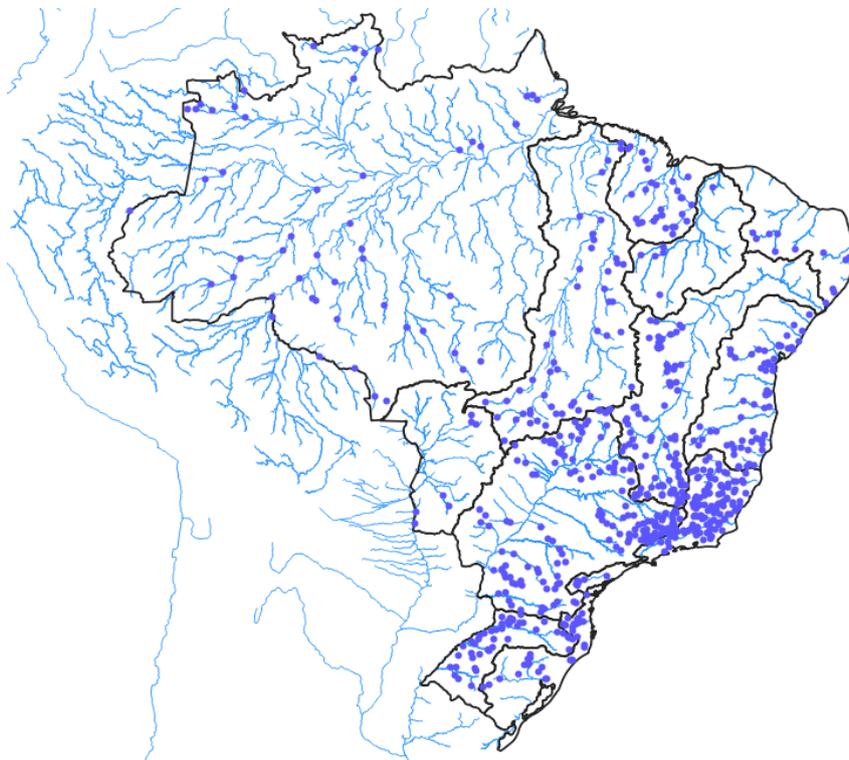


Figura 1: Localização de postos fluviométricos com série longa para análise da variabilidade amostral

### 3.3 CÁLCULO DA VAZÃO DE REFERÊNCIA DE SÉRIE LONGA

Em cada um dos postos fluviométricos selecionados, a vazão de referência “de série longa” foi obtida com todos os dados disponíveis no período de jan/1980 e dez/2014. A vazão de referência “de série longa” foi obtida, portanto, com base em dados de 30 a 35 anos, dependendo da disponibilidade em cada posto fluviométrico. Em cerca de 35% dos postos utilizados na análise a vazão de referência foi obtida com 35 anos de dados, e em 80% dos postos a vazão foi obtida com 32 ou mais anos de dados.

### 3.4 CÁLCULO DA VAZÃO DE REFERÊNCIA DE SÉRIE CURTA

Conforme mencionado antes, o efeito do tamanho da série foi avaliado através da obtenção de estimativas da vazão de referência utilizando apenas partes da série de dados disponível em cada posto fluviométrico. Estas estimativas foram obtidas com N anos de dados, com N = 1, 2, 3 ... até 30 anos. Para cada valor de N, este procedimento foi repetido 1000 (mil) vezes, com data de início aleatória. Testes realizados com um número ainda maior de réplicas aleatórias demonstraram que R=1000 era suficientemente grande.

### 3.5 MÉTRICAS DE ERRO

Existem diferentes métricas para calcular o erro entre os valores estimados e os valores “reais”. A diferença entre a vazão estimada por um modelo e o valor “real” é, tipicamente, denominado de resíduo:

$$RES = QR_{est} - QR_{real} \quad (1)$$

onde: RES é o resíduo;  $QR_{est}$  é a QR estimada por um modelo;  $QR_{real}$  é a QR observada, obtida dos dados do posto fluviométrico.

O resíduo também pode ser calculado em termos de vazão específica:

$$RES = \frac{QR_{est} - QR_{real}}{A} \quad (2)$$

onde A é a área de drenagem da bacia hidrográfica.

O erro relativo entre as estimativas de vazões de referência pode ser calculada por:

$$ER(\%) = 100 \left( \frac{QR_{est} - QR_{real}}{QR_{real}} \right) \quad (3)$$

onde: ER é o erro relativo (%);  $QR_{est}$  é a vazão de referência (pode ser a vazão média ou a Q95) estimada com base na série curta;  $QR_{real}$  é a vazão de referência estimada com base na série longa.

No contexto de disponibilidade hídrica é importante que a métrica de desempenho seja capaz de indicar a proporcionalidade de excesso/déficit hídrico da estimativa. A métrica denominada de Razão entre Vazões Médias (RQM), descrita a seguir, é proposta e utilizada para avaliar o desempenho entre valores observados e estimados das vazões de referência.

$$RQM = Sinal(QM_{est} - QM_{real}) \cdot Máximo \left( \frac{QM_{est}}{QM_{real}}; \frac{QM_{real}}{QM_{est}} \right) \quad (4)$$

onde: RQ é a métrica de razão de vazões;  $QM_{est}$  é a vazão média de série curta;  $QM_{real}$  é a vazão média de série longa. Essa métrica indica a proporção entre vazão estimada e a “real”, de tal maneira que um valor RQ=+2 indica que a vazão estimada é o dobro da “real” e, um valor RQ=-

2 indica que a vazão estimada é a metade da “real”. Um ajuste perfeito fornece um valor de RQM=1.

Ao considerar que a magnitude dos erros de superestimativa e subestimativa (em proporção) são equivalentes e igualmente importantes, a métrica pode ser utilizada ignorando o sinal, critério esse adotado para as análises no presente estudo, ou seja, em valores absolutos, ABSRQM:

$$ABSRQM = \text{Máximo} \left( \frac{QM_{est}}{QM_{real}}; \frac{QM_{real}}{QM_{est}} \right) \quad (5)$$

onde: ABSRQM é a métrica de razão de vazões, em valor absoluto;  $QM_{est}$  é a vazão estimada;  $QM_{real}$  é a vazão média observada, obtida dos dados do posto fluviométrico.

Em geral, as métricas foram aplicadas para as vazões de referência Q95 e QM. A equação 4, adaptada para a vazão de referência Q95, resulta na métrica RQ95, e a equação 5, adaptada para a Q95, resulta na métrica ABSRQ95.

Em locais onde a aridez é maior e as vazões são mais baixas, a aplicação da equação 2 para vazão específica, em mm/dia e L/s.km<sup>2</sup>, fornece uma informação diretamente comparável às medidas típicas de precipitação e evapotranspiração [L/T].

## 4 RESULTADOS

---

Nessa seção apresenta-se uma análise com o objetivo principal de identificar os efeitos do tamanho da série utilizada na estimativa das vazões de referência. O conjunto de amostras obtidas a partir da reamostragem de dados em diferentes postos fluviométricos permite aproximar a distribuição de frequência dos erros e, dessa maneira, medidas de incerteza. Dessa maneira, para cada tamanho da série de dados considerado é possível utilizar parâmetros como a mediana e outros percentis para descrever o erro esperado. Em geral, busca-se verificar a magnitude dos erros em estimativas obtidas a partir de séries mais curtas, em relação ao valor “real” obtido de séries longas.

### 4.1 EXEMPLO ILUSTRATIVO DOS RESULTADOS

Para demonstrar a metodologia, considera-se, como exemplo, a série de dados do posto fluviométrico Lajeado (código 64442800), localizado no rio Imbituva, no Paraná. Utilizando os dados disponíveis (31 anos) neste posto, no período de jan/1980 e dez/2014, a vazão média é 27,9 m<sup>3</sup>.s<sup>-1</sup>. Quando são utilizados apenas períodos de cinco anos (em sequência) para obter a mesma vazão de referência, os valores variam entre 19,2 e 34,6 m<sup>3</sup>.s<sup>-1</sup>. Isto significa que usando apenas 5 anos de dados para estimar a vazão média nesse local é possível encontrar casos em que esta vazão de referência é subestimada em 31% e outros em que a vazão é superestimada em 24%, em relação ao valor que seria obtido com a série longa.

A Figura 2 apresenta o histograma de frequência dos valores de vazão média estimada com apenas 5 anos de dados neste posto fluviométrico, considerando as 1000 estimativas. Observa-se que em 324 das 1000 realizações o valor da vazão média estimada situa-se no intervalo entre 26 e 30 m<sup>3</sup>.s<sup>-1</sup>, que é o intervalo que contém o valor estimado com base na série longa (27,9 m<sup>3</sup>.s<sup>-1</sup>).

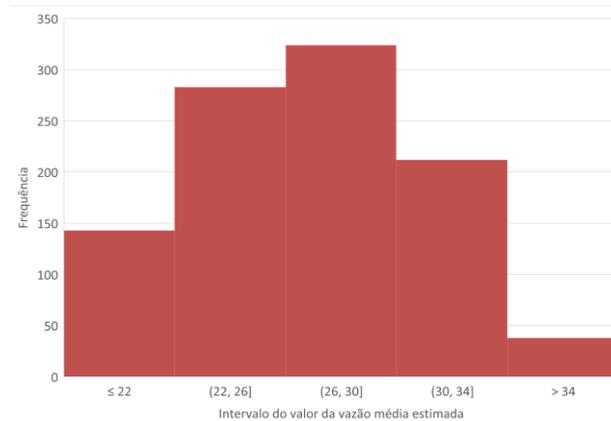


Figura 2: Histograma de frequência de valores de vazão média estimada com base em séries de 5 anos de dados no posto fluviométrico Lajeado (código 64442800), no rio Imbituva (PR).

Os resultados das 1000 estimativas de vazão média obtidas com séries de 5 anos de extensão também podem ser resumidos na forma de um box-plot, como mostra a Figura 3 (no lado esquerdo). Além disso, essa figura também mostra o box-plot das 1000 estimativas de vazão média obtidas com séries de 20 anos de dados (no lado direito).

A Figura 3 mostra, também, o efeito do número de anos da série utilizada para gerar a estimativa de vazão média sobre a incerteza dessa estimativa. O aumento do número de anos de dados reduz a distância entre os valores mínimo e máximo, e reduz a distância entre os percentis 25% e 75%. Utilizando 20 anos de dados o percentil 25% da vazão média é  $26,7 \text{ m}^3 \cdot \text{s}^{-1}$ , e o percentil 75% é  $28,2 \text{ m}^3 \cdot \text{s}^{-1}$ . Portanto, 50% das estimativas de vazão média obtidas com 20 anos de dados estão numa faixa entre  $26,7$  e  $28,2 \text{ m}^3 \cdot \text{s}^{-1}$ . Utilizando apenas 5 anos de dados, a faixa de valores que contém 50% das estimativas vai de  $23,0$  até  $30,0 \text{ m}^3 \cdot \text{s}^{-1}$ .

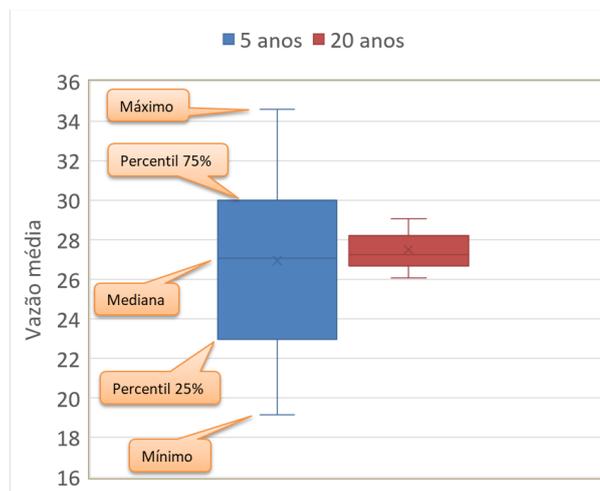


Figura 3: Box-plot das 1000 estimativas de vazão média estimada com base em séries de 5 anos e de 20 anos de dados no posto fluviométrico Lajeado (código 64442800), no rio Imbituva (PR).

Considerando correta a vazão média calculada com a série longa (31 anos), é possível também apresentar os resultados na forma de um box-plot do erro relativo, conforme ilustrado na Figura 4, tanto para 5 como para 20 anos de dados.

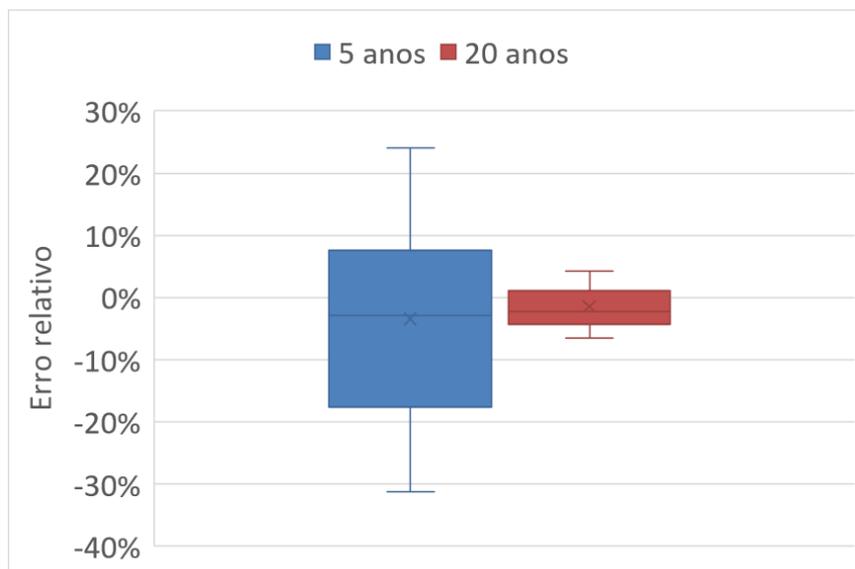


Figura 4: Box-plot dos erros relativos das 1000 estimativas de vazão média estimada com base em séries de 5 anos e de 20 anos de dados no posto fluviométrico Lajeado (código 64442800), no rio Imbituva (PR).

Também é possível analisar os resultados na forma de box-plot da métrica de relação entre vazões ABS(RQM), conforme mostra a Figura 5. Lembrando que o valor de ABS(RQM) é igual a 1 quando o valor estimado é exatamente igual ao valor considerado correto, e igual a 2 quando a vazão estimada é igual ao dobro ou a metade da vazão correta.

O valor da mediana de ABS(RQM) apresentado no box-plot relativo às estimativas baseadas em 5 anos de dados da Figura 5 é de 1,12. Isto significa que em 50% das 1000 estimativas realizadas no posto fluviométrico 64442800 o valor de ABS(RQM) é inferior a 1,12. Em outras palavras, 50% das estimativas estão no intervalo entre 11% a menos e 12% a mais do que o valor real (estimado com a série longa). No caso das estimativas obtidas com 20 anos de dados o valor da mediana de ABS(RQM) cai para 1,03, que significa que 50% das estimativas estão no intervalo entre 3% a menos e 3% a mais do que o valor real.

Os box-plot da Figura 5 também mostram os valores máximos, mínimos e os percentis de 25 e 75% da métrica ABS(RQM) das estimativas. O percentil 75% é particularmente interessante porque pode ser utilizado como um indicador que inclui a maioria (3 em cada 4) estimativas, e que, portanto, pode ser utilizado como um indicador da incerteza da estimativa da vazão de referência baseada em uma série relativamente curta de dados. Os valores do percentil 75% de ABS(RQM) com séries de 5 e 20 anos são, respectivamente, 1,22 e 1,05. Isto significa que, no posto fluviométrico 64442800, no rio Imbituva, as estimativas de vazão média com apenas 5 anos de dados têm uma incerteza tal que, em 75% dos casos, os erros estão em um intervalo definido por 18% a menos e 22% a mais do que a vazão média de série longa. Já no caso das estimativas obtidas com 20 anos de dados, esta incerteza, estimada pelo percentil 75%, é dada por um intervalo definido por 5% a menos e 5% a mais do que a vazão média de série longa.

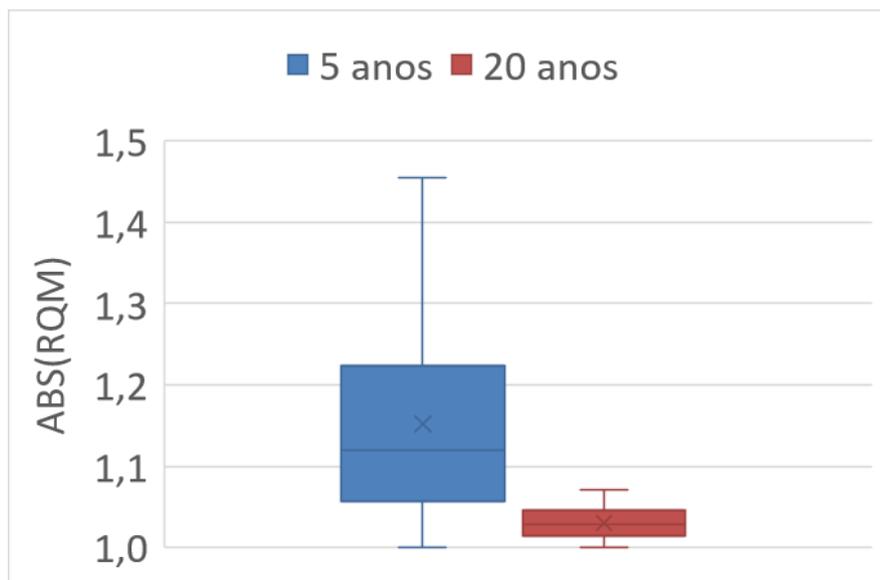


Figura 5: Box-plot do valor da métrica ABS(RQM) das 1000 estimativas de vazão média estimada com base em séries de 5 anos e de 20 anos de dados no posto fluviométrico Lajeado (código 64442800), no rio Imbituva (PR).

#### 4.2 EFEITO DO TAMANHO DA SÉRIE DE DADOS SOBRE A ESTIMATIVA DA VAZÃO MÉDIA

O exemplo ilustrativo anterior trata de um único posto fluviométrico. Análises semelhantes a esta foram repetidas para os demais postos fluviométricos com série longa considerados neste trabalho. Em cada um dos postos fluviométricos o comportamento pode ser diferente. Para sintetizar os resultados, foram geradas curvas que relacionam os valores das métricas com o número de anos de dados utilizados na estimativa da vazão de referência. Estas curvas, no entanto, foram elaboradas agregando os resultados de todos os postos fluviométricos analisados.

A Figura 6 apresenta os percentis 12,5%, 50% (mediana) e 87,5% da métrica de erro relativo da estimativa das vazões médias, considerando tamanhos de série entre 1 e 30 anos. Os erros relativos apresentam a mediana próxima a zero, indicando que em, aproximadamente, 50% dos casos ocorrem superestimativa ou subestimativas. Ao considerar o intervalo entre os percentis 12,5 e 87,5% (ou seja, 75% dos casos em torno da tendência central) observa-se que estimativa da QM baseada em apenas 1 ano de dados teve valores de ERQM entre, aproximadamente, -30% e +30%.

À medida que o número de dados utilizados para realizar a estimativa aumenta, o erro tende a diminuir. Utilizando 10 anos de dados para estimar a vazão média, 75% dos casos apresentam ERQM entre -12% e +10% e, utilizando 20 anos de dados, os erros relativos ficam entre -8% e +5%. Os valores não chegam exatamente a zero porque o número de anos de dados utilizados para estimar a vazão real (“série longa”) foi superior a 30 anos na maioria dos postos fluviométricos, conforme mencionado na metodologia.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Média  
(Percentil 75%)

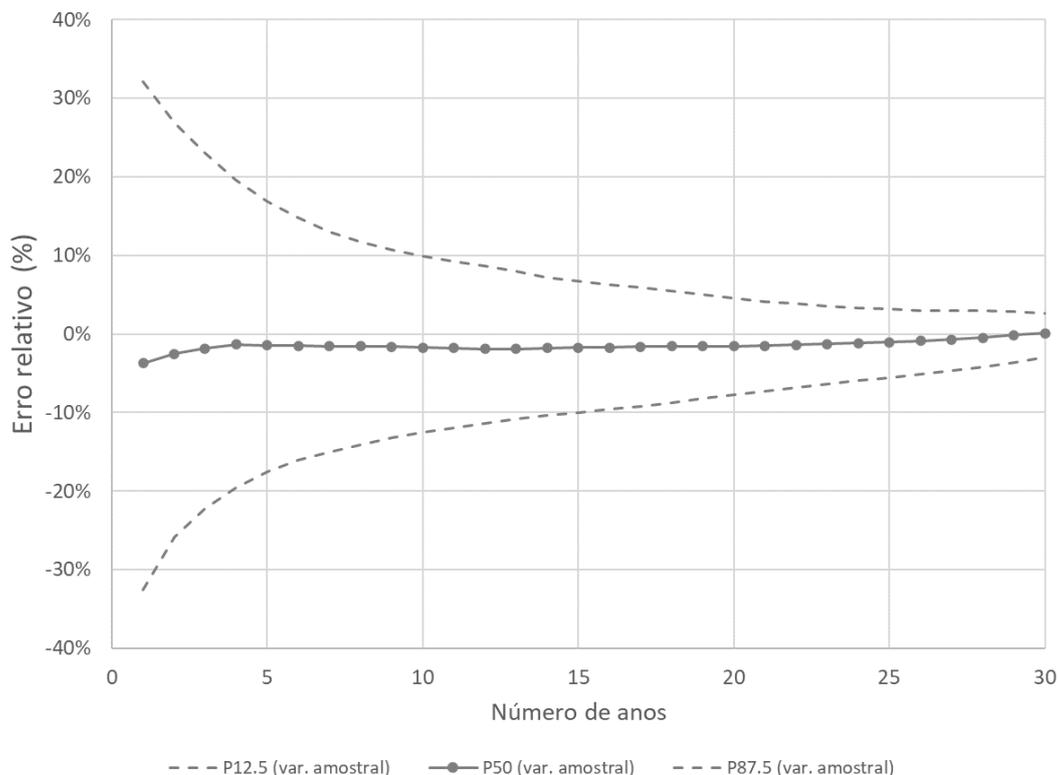


Figura 6: Erro relativo na estimativa da vazão média. A linha sólida representa a mediana, e as linhas tracejadas definem os percentis 12,5% e 87,5% (a região entre as duas linhas tracejadas contém 75% dos casos).

A Figura 7 apresenta o percentil 75% da métrica ABSRQM - razão entre vazões médias, em valor absoluto - para cada tamanho de série curta, obtido pela reamostragem. Dessa maneira, a figura define o limiar de erro que inclui 75% dos casos, quando se considera os erros de superestimativa e subestimativa equivalentes entre si (e.g. metade = dobro). Observa-se que estimativa da vazão média baseada em apenas 1 ano de dados resultou em ABSRQM igual a 1,4. Isto significa que em 75% dos casos analisados, o valor de RQM ficou entre -1,4 e +1,4. Em outras palavras, em 75% dos casos analisados o valor real da vazão média (estimado com 30 a 35 anos de dados) situou-se na faixa definida entre o valor da vazão média estimada MENOS 28% e o valor da vazão média estimada MAIS 40%.

Também se observa na Figura 7 que, com 10 anos de dados observados, o valor de ABSRQM é próximo a 1,12 (no percentil 75%). Isto significa que, em 75% dos casos analisados, o valor real da vazão média (estimado com 30 a 35 anos de dados) situou-se na faixa definida entre 11% a MENOS e 12% a MAIS do que o valor estimado.

A forma da curva da Figura 7 mostra que à medida que o número de dados utilizados para estimativa aumenta, o valor de ABSRQM diminui. Com 20 anos de dados o valor de ABSRQM já é próximo a 1,07, e com 30 anos de dados o valor é inferior a 1,03, o que significa erros da ordem de 3% a mais ou a menos.

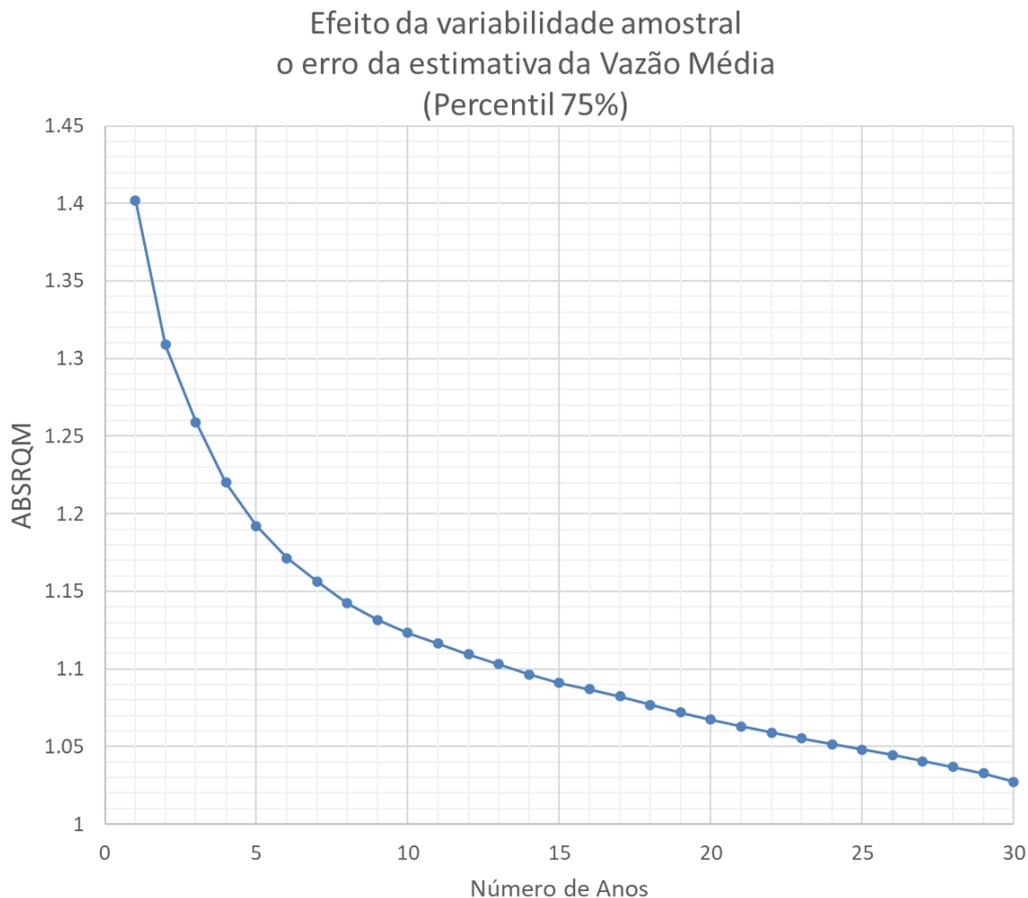


Figura 7: Métrica ABSRQM (75% dos casos) da estimativa da vazão média em função do número de anos de dados (no eixo horizontal) quando comparada à estimativa da vazão média com base em 30 a 35 anos de dados.

Outra métrica importante para a avaliação da incerteza da estimativa das vazões de referência é o resíduo de vazões específicas. Neste caso, o resíduo (equação 1) é dividido pela área da bacia hidrográfica. Esta métrica tende a ser mais adequada em casos em que a vazão real (estimada com a série longa) se aproxima de zero.

A Figura 8 apresenta como o percentil 75% do erro absoluto da vazão específica varia com o tamanho de série curta. Observa-se que com 20 anos de dados o percentil 75% do erro absoluto se aproxima de 1 litro por segundo por  $\text{km}^2$ . Isto significa que estimativas de vazão média obtidas utilizando séries com 20 anos de dados tem erros inferiores a  $1 \text{ l.s}^{-1}.\text{km}^{-2}$  em 75% dos casos analisados. Com 10 anos este erro é praticamente duas vezes maior.

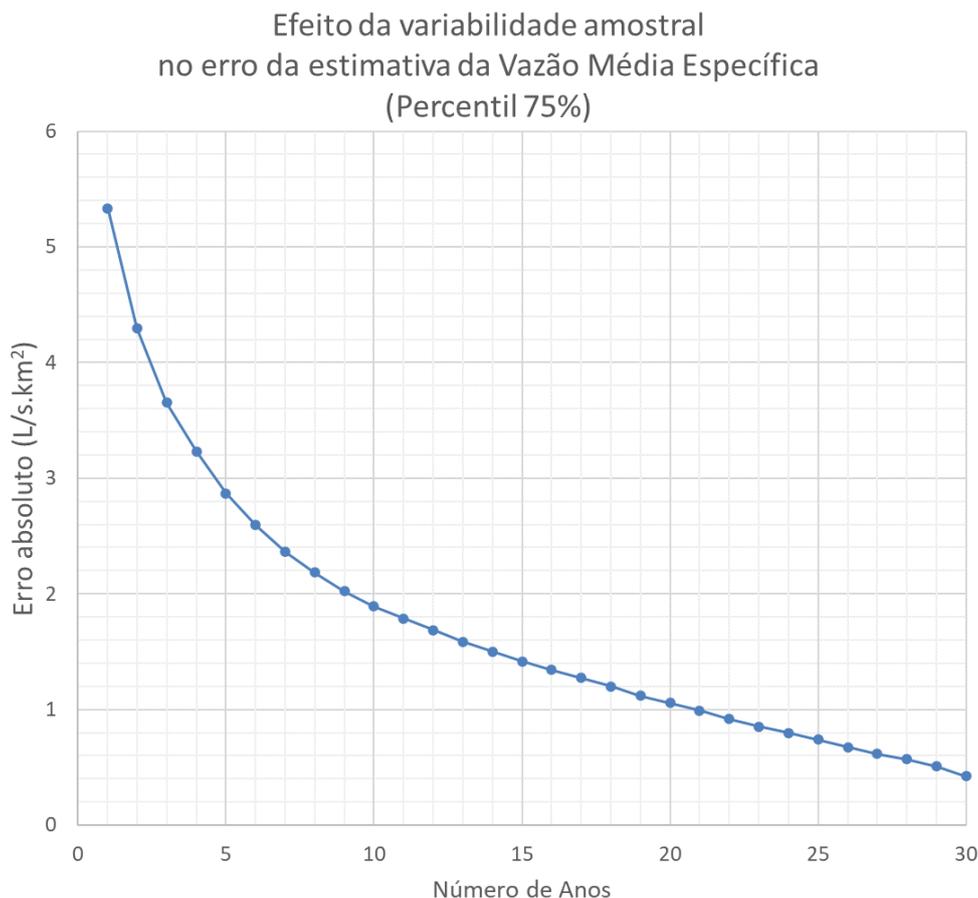


Figura 8: Erro absoluto da vazão específica média estimado com séries curtas de dados (1 a 30 anos) em relação à vazão média estimada com série longa (30 a 35 anos). A linha representa o percentil 75%.

### 4.3 EFEITO DO TAMANHO DA SÉRIE DE DADOS SOBRE A ESTIMATIVA DA Q95

As análises realizadas para a vazão média foram repetidas para a vazão de referência Q95, que é a vazão igualada ou superada em 95% do tempo.

A Figura 9 apresenta os percentis 12,5%, 50% (mediana) e 87,5% da métrica de erro relativo para as vazões Q95. A mediana dos erros relativos indica uma tendência de superestimativas nessa métrica. Ao considerar o intervalo entre os percentis 12,5 e 87,5% (ou seja, 75% dos casos em torno da tendência central) observa-se que estimativa da Q95 baseada em apenas 1 ano de dados teve valores de ERQ95 entre -16% e +91%.

À medida que o número de dados utilizados para realizar a estimativa aumenta, o erro tende a diminuir. Com 10 anos de dados observados, 75% dos casos apresentam ERQ95 entre -12% e +27% e, com 20 anos de dados observados, os erros relativos ficam entre -6% e +13%. Os valores não chegam exatamente a zero porque o número de anos de dados utilizados para estimar a Q95 real (“série longa”) foi superior a 30 anos na maior parte dos postos fluviométricos analisados, conforme mencionado na metodologia.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Q95  
(em 75% dos casos)

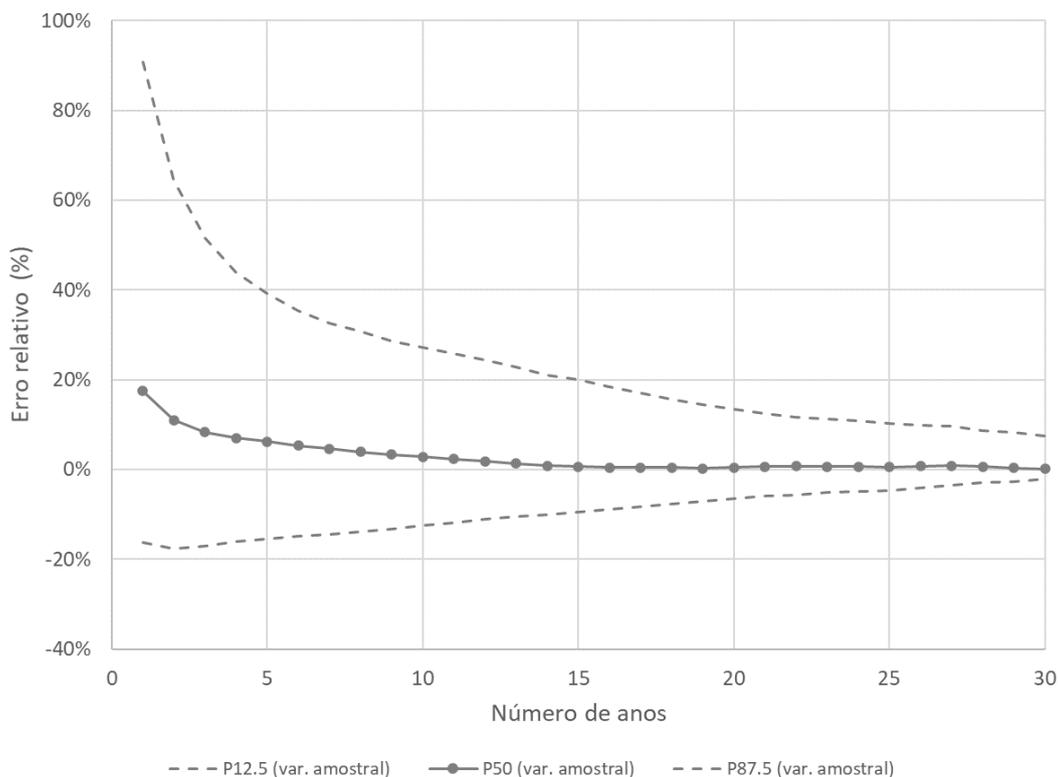


Figura 9: Erro relativo na estimativa da vazão média. A linha sólida representa a mediana, e as linhas pontilhadas definem a faixa central de 75% dos casos.

A métrica de razão entre vazões foi aplicada também para a Q95. Neste caso foi calculado o valor de  $RQ95$ , conforme a equação 4 (adaptada para a vazão Q95), e o seu valor absoluto ( $ABSRQ95$ ). A Figura 10 apresenta o percentil 75% da métrica de  $ABSRQ95$ , obtido pela reamostragem, para cada tamanho de série curta. Dessa maneira, a curva ilustrada na Figura 10 define o limiar de erro que inclui 75% dos 1000 casos simuladas, quando se considera os erros de superestimativa e subestimativa equivalentes entre si (e.g. metade = dobro). Observa-se que estimativa da Q95 baseada em apenas 1 ano de dados resultou em  $ABSRQ95$  igual a 1,6. Isto significa que em 75% dos casos analisados, o valor de  $RQ95$  ficou entre -1,6 e +1,6. Em outras palavras, em 75% dos casos analisados o valor real da Q95 (estimado com 30 a 35 anos de dados) situou-se na faixa definida entre o valor da Q95 estimada MENOS 38% e o valor da Q95 estimada MAIS 60%.

A curva da Figura 10 decresce rapidamente com o aumento do número de anos de dados (eixo horizontal). Com 10 anos de dados observados o valor de  $ABSRQ95$  é próximo a 1,2 (no percentil 75%). Isto significa que, em 75% dos casos analisados, o valor real da Q95 (estimado com 30 a 35 anos de dados) situou-se na faixa definida entre 16% a MENOS e 20% a MAIS do que o valor estimado. Com 20 anos de dados o valor de  $ABSRQ95$  já é próximo a 1,1, e com 30 anos de dados o valor é inferior a 1,1, o que significa erros da ordem de 10% a mais ou a menos.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Q95  
(Percentil 75%)

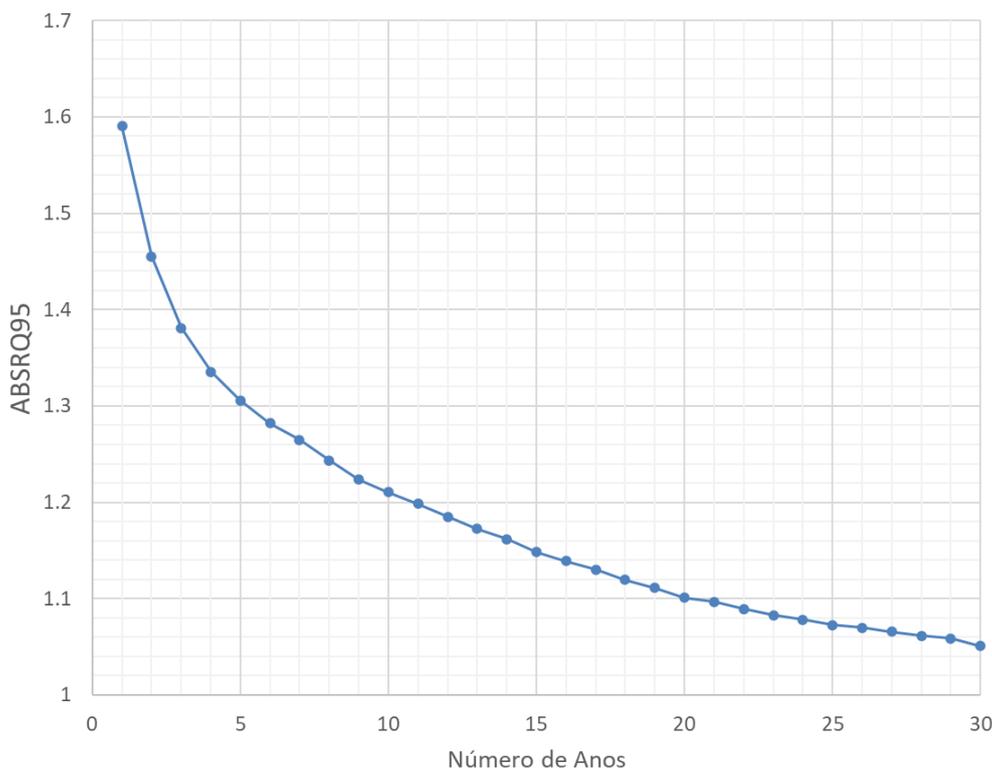


Figura 10: Métrica ABSRQ95 (75% dos casos) da estimativa da vazão média em função do número de anos de dados (no eixo horizontal) quando comparada à estimativa da vazão média com base em 30 a 35 anos de dados.

A comparação entre a Figura 7 e a Figura 10 mostra claramente que os erros de estimativa devidos à variabilidade amostral quando as séries são relativamente curtas são maiores no caso da vazão de referência Q95 do que no caso da vazão média. No caso da vazão média são necessários 13 anos para que o percentil 75% dos valores de ABSRQM atinja um valor igual a 1,1, que corresponde a um erro de aproximadamente 10%. Já no caso da Q95 o mesmo valor do percentil 75% dos valores de ABSRQ95 só é atingido utilizando uma série com 20 anos de dados.

A comparação entre a Figura 6 e a Figura 9 também é interessante, porque mostra que com séries muito curtas de dados há uma tendência clara de superestimar a Q95 e uma ligeira tendência a subestimar a vazão média.

A Figura 11 apresenta o comportamento dos erros da estimativa da Q95 (percentil 75%) em termos de vazão específica. Observa-se que com séries de 15 anos de dados é possível estimar a Q95 específica com um erro absoluto menor ou igual a, aproximadamente, 1/2 litro por segundo por km<sup>2</sup>.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Q95 Específica  
(Percentil 75%)

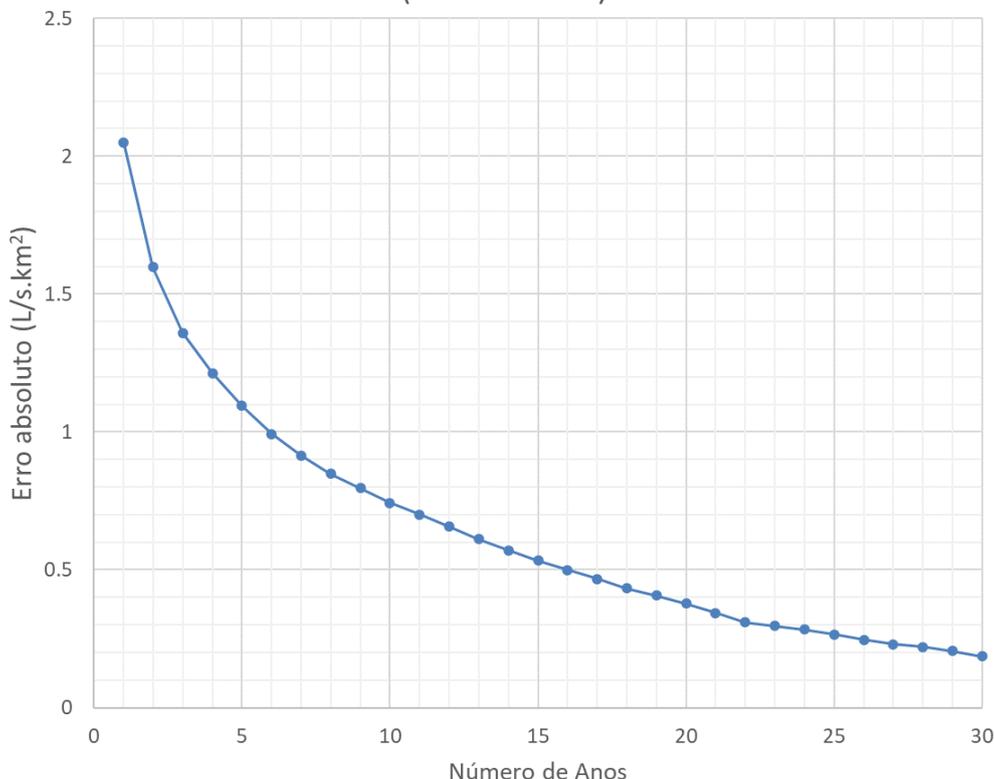


Figura 11: Erro absoluto da Q95 específica (Q95 dividida pela área da bacia hidrográfica) estimado com séries curtas de dados (1 a 30 anos) em relação à Q95 específica estimada com série longa (30 a 35 anos). A linha representa o percentil 75%.

#### 4.4 EFEITO DE ESCALA

Os resultados anteriores foram obtidos sintetizando todos os postos fluviométricos analisados em curvas únicas. Entretanto, é razoável esperar que a variabilidade amostral não seja igual em todos os rios e em todos os locais. Um dos fatores que pode influenciar a sensibilidade das estimativas de vazão de referência ao tamanho da série de dados observados é a escala da bacia, ou o tamanho do rio em que está localizado o posto fluviométrico. Assim, foi realizada uma avaliação adicional considerando uma categorização dos resultados em classes de vazão e em classes de área de drenagem.

A Figura 12 apresenta a mediana dos erros (ABSRQM) na estimativa da vazão média, para diferentes faixas de vazão média e, também, considerando o tamanho da série utilizada. Da mesma forma, aqui as comparações são realizadas em relação às vazões obtidas com séries longas de 30 a 35 anos de dados. Em geral, é possível observar uma tendência de ocorrer erros menores para locais com vazões mais altas. Como a categorização foi realizada homogeneizando o número de casos em cada grupo, a última faixa de vazões representada no eixo das abscissas possui uma banda bastante larga, representando também, rios maiores.

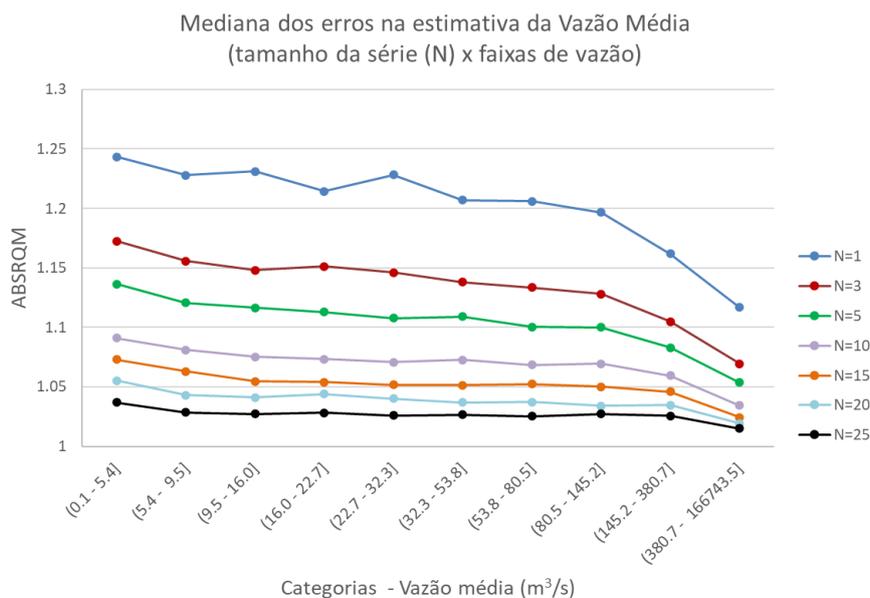


Figura 12: Métrica ABSRQM (50% dos casos) da estimativa da vazão média em função de categorias de vazão média (eixo horizontal) e número de anos de dados (em linhas).

A Figura 13 apresenta avaliação similar para os erros nas vazões Q95 (mediana de ABSRQ95), e mostra que o efeito da escala também é observado nas vazões mínimas, até de uma forma mais pronunciada. Enquanto em rios com Q95 na faixa de 1 a 2 m<sup>3</sup>.s<sup>-1</sup> são necessários 15 anos de dados para que a mediana de ABSRQ95 seja inferior a 1,1, em rios com Q95 superior a 12 m<sup>3</sup>.s<sup>-1</sup> bastam cerca de 10 nos de dados para atingir o mesmo nível de incerteza. Além disso, a figura mostra que com 20 anos de dados o valor mediano de ABSRQ95 é inferior a 1,1 tanto para rios de vazão baixa como em rios de vazão alta.

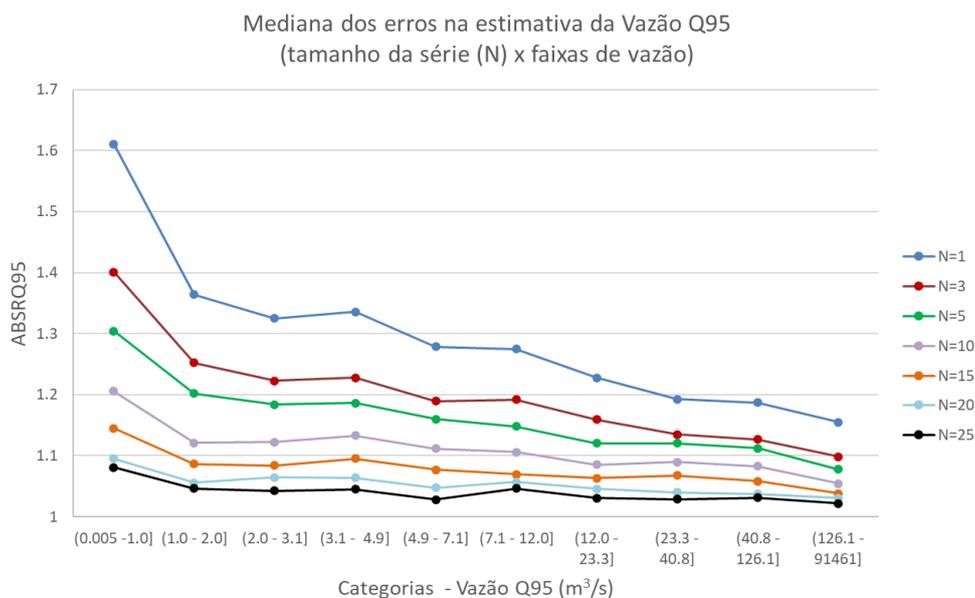


Figura 13: Métrica ABSRQ95 (50% dos casos) da estimativa da vazão Q95 em função de categorias de vazão Q95 (eixo horizontal) e número de anos de dados (em linhas).

Os resultados da Figura 13 podem ser comparados aos resultados obtidos por Laaha e Blöschl (2005) em um estudo na Áustria, conforme comentado na introdução. O erro médio (RMSE) da estimativa da Q95 encontrado por estes autores variou entre mais de 50%, no caso de 1 ano de dados em bacias relativamente secas, e menos de 10%, em bacias úmidas com 15 anos de dados. Esses erros são comparáveis aos encontrados no Brasil, de acordo com a Figura 13, pois com apenas 1 ano de dados o valor mediano de ABSRQ95 varia entre um pouco mais de 1,6 em locais de vazão baixa (seja por pequena área de drenagem ou por baixa precipitação), e menos de 1,2 em locais de vazão alta (mais úmidos ou com maior área de drenagem). E com 15 anos de dados o valor mediano de ABSRQ95 varia entre um pouco mais de 1,1 em locais de vazão baixa (seja por pequena área de drenagem ou por baixa precipitação), e menos de 1,1 em locais de vazão alta (mais úmidos ou com maior área de drenagem). O valor de ABSRQ95 de 1,6 corresponde a uma superestimativa de 60% ou uma subestimativa de 38%, e o valor de ABSRQ95 de 1,1 corresponde a uma superestimativa de 10% ou uma subestimativa de 9%. Embora a mediana dos erros, adotada aqui, e o RMSE, utilizado por Laaha e Blöschl (2005), não sejam exatamente equivalentes, observa-se que os erros da estimativa da Q95 encontrados no presente trabalho tem a mesma ordem de grandeza do que os erros encontrados na Áustria, embora no trabalho atual (Brasil) predominem dados de postos fluviométricos com área de drenagem muito maior do que os analisados na Áustria, por por Laaha e Blöschl (2005).

Outra forma de avaliar a escala é através da área de drenagem da bacia. A Figura 14 e a Figura 15 ilustram o comportamento da mediana dos erros ABSRQM e ABSRQ95, respectivamente, em função de categorias de área de drenagem, e tamanho de série.

Observa-se que a incerteza da estimativa da vazão média (Figura 14) associada às séries curtas de dados é maior em bacias com área de drenagem menor do que, aproximadamente, 10.000 km<sup>2</sup>. Ao mesmo tempo, a incerteza parece ser relativamente independente da área da bacia em áreas de drenagem inferiores a, aproximadamente, 6000 km<sup>2</sup>.

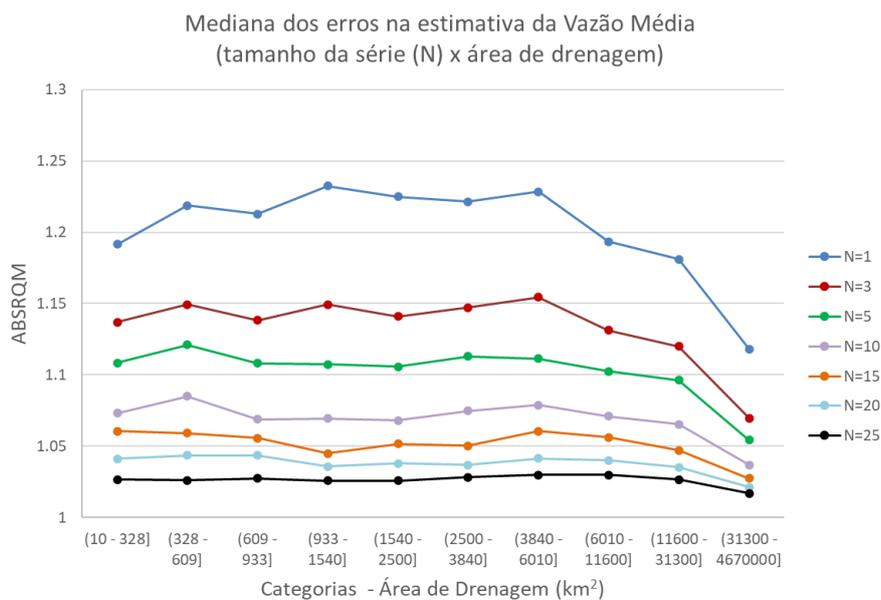


Figura 14: Métrica ABSRQM (50% dos casos) da estimativa da vazão média em função de categorias de área de drenagem (eixo horizontal) e número de anos de dados (em linhas).

A Figura 15 mostra que a incerteza da estimativa da Q95 associada às séries curtas de dados é maior em bacias hidrográficas com área de drenagem intermediária (da ordem de 1000 km<sup>2</sup>). Em bacias com área de drenagem maior do que este valor, a incerteza tende a diminuir com o aumento da área. Surpreendentemente, em bacias com área de drenagem menor do que 1000 km<sup>2</sup> há uma tendência de a incerteza diminuir com a diminuição da área de drenagem. É possível que este resultado inesperado tenha influência da localização geográfica da maioria dos postos fluviométricos cuja área de drenagem é menor do que 1000 km<sup>2</sup>. O efeito da localização geográfica dos postos fluviométricos é avaliado no próximo item.

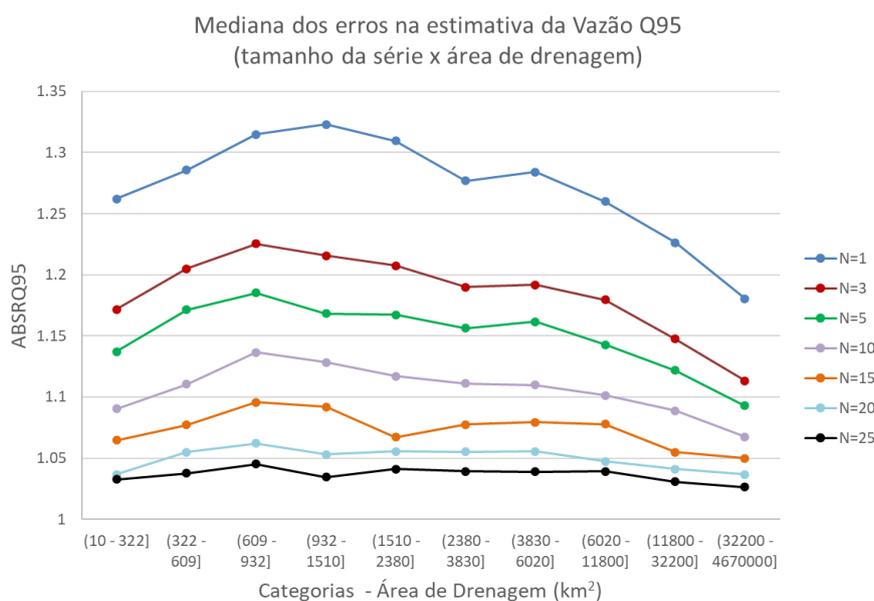


Figura 15: Métrica ABSRQ95 (50% dos casos) da estimativa da vazão Q95 em função de categorias de área de drenagem (eixo horizontal) e número de anos de dados (em linhas).

#### 4.5 EFEITO DA LOCALIZAÇÃO GEOGRÁFICA DAS BACIAS

Devido a escala nacional da distribuição de postos fluviométricos utilizados nesse estudo e a variedade de relações climáticas no país, consideramos uma avaliação similar dos erros, mas considerando agrupamentos por regiões hidrográficas. A Tabela 1 apresenta a relação entre regiões e o código do posto fluviométrico utilizado.

Tabela 1: Regiões hidrográficas do Brasil e algarismo inicial do código dos postos fluviométricos.

Algarismo inicial do Código	Região
1	Amazônica
2	Tocantins-Araguaia
3	Atlântico Nordeste, Parnaíba
4	São Francisco
5	Atlântico Leste, Atlântico Sudeste
6	Paraná-Paraguai
7	Uruguai
8	Atlântico Sul

Analisando a Figura 16 observa-se que os menores valores da incerteza, em qualquer tamanho de série, ocorrem na região R1 (bacia Amazônica). Isto pode ser tanto um efeito do clima tropical e equatorial da região, como um efeito da escala, já que nesta região predominam os postos fluviométricos com grande vazão e grande área de drenagem.

Os maiores valores de incerteza na estimativa da vazão média ocorrem na região R3, que corresponde às bacias do Atlântico Nordeste (Occidental e Oriental) e rio Parnaíba, e incluem grande parte da região semiárida do Brasil. Esta região é caracterizada por uma alta variabilidade climática, quantificada, por exemplo, pelos altos valores do coeficiente de variação da chuva média anual.

Considerando séries de dados muito curtas, com apenas 1 ou 2 anos de dados, a bacia do rio Uruguai (R7) também se destaca pela alta incerteza na estimativa da vazão média.

Considerando séries de dados de 4 a 20 anos, a região R5 (bacias do Atlântico Leste e Atlântico Sudeste) também se destaca pela alta incerteza da estimativa da vazão média, que é somente inferior à verificada na região R3. Esta incerteza também está, possivelmente, relacionada à alta variabilidade climática da região semiárida do país, parcialmente presente na região R5.

As bacias dos rios Paraná (R6) e São Francisco (R4) encontram-se em situação intermediária de incerteza, e a bacia hidrográfica dos rios Tocantins e Araguaia (R2) tem incerteza quase tão baixa como a da bacia Amazônica.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Média por região  
(Percentil 75%)

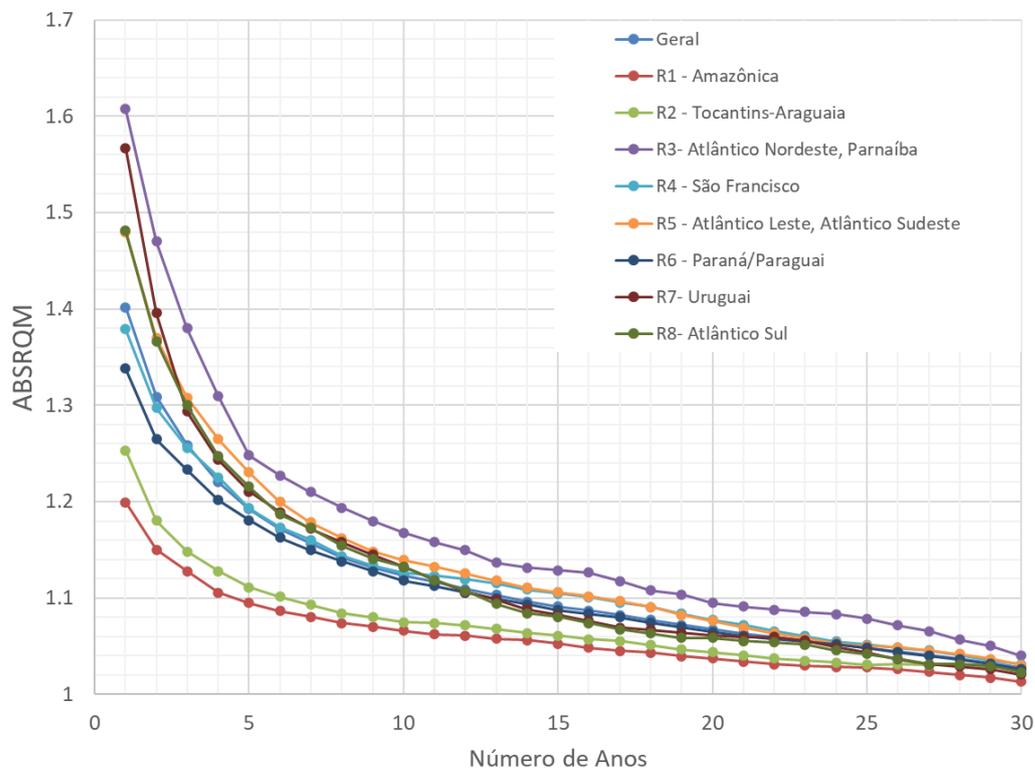


Figura 16: Métrica ABSRQM (75% dos casos) da estimativa da vazão média em função do número de anos de dados (no eixo horizontal) quando comparada à estimativa da vazão média com base em 30 a 35 anos de dados, para diferentes regiões hidrográficas.

A Figura 17 mostra os erros absolutos da estimativa da vazão média específica (percentil 75%) com séries de 1 a 30 anos de dados, considerando os postos fluviométricos de diferentes regiões hidrográficas do país. Neste caso, as regiões que se destacam com maior incerteza em função da variabilidade amostral são as regiões R7 e R8, localizadas no Sul do Brasil. Já a região R3 apresenta valores mais baixos de incerteza, o que ocorre porque as vazões médias específicas na região Nordeste do Brasil são muito menores do que no restante do país.

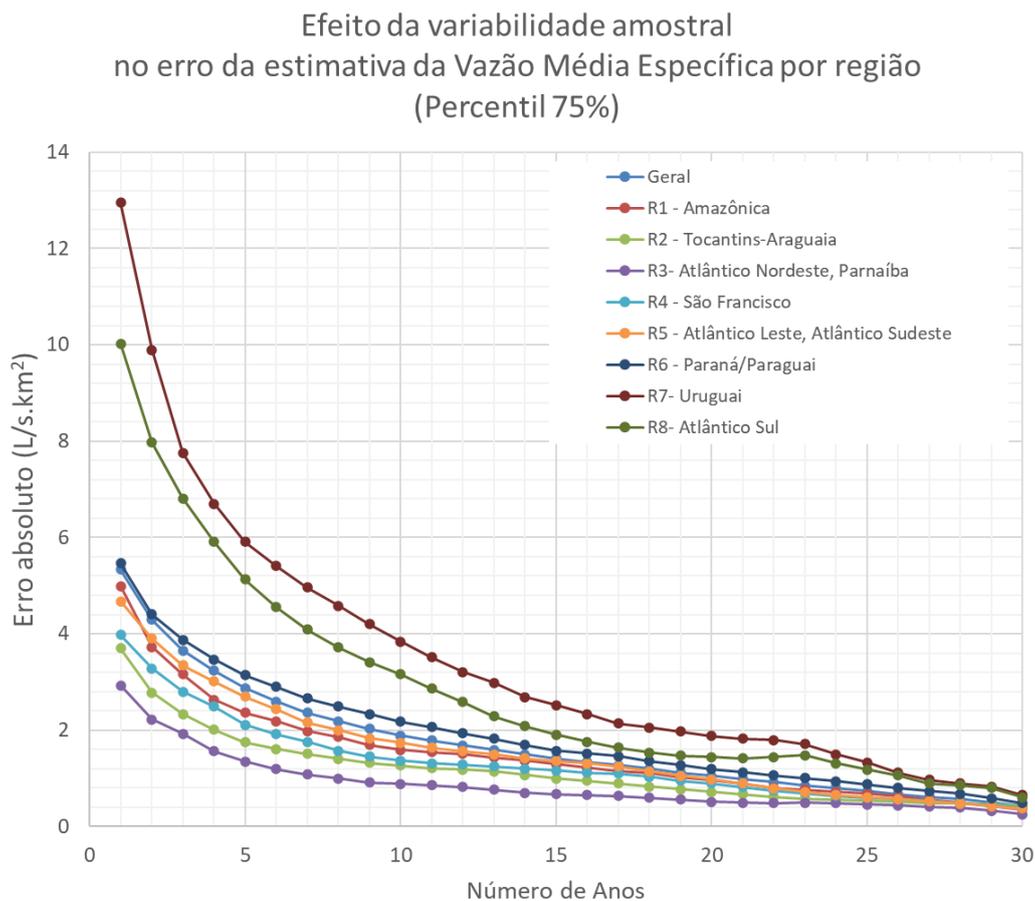


Figura 17: Erro absoluto da vazão específica média estimado com séries curtas de dados (1 a 30 anos) em relação à vazão média estimada com série longa (30 a 35 anos) na avaliação geral e por região hidrográfica. A linha representa o percentil 75%.

No caso da influência da região geográfica sobre a incerteza da estimativa da vazão Q95 (estimada pelo valor de ABSRQ95 - Figura 18) a análise é semelhante à anterior, em vários aspectos. A maior incerteza ocorre na região R3, que inclui grande parte do semiárido. As menores incertezas são observadas nas regiões R1 e R2 (Amazonia e Araguaia/Tocantins). Entretanto, diferentemente do caso da vazão média, a Figura 18 mostra que a incerteza na estimativa da Q95 nos rios das regiões R7 e R8 (rio Uruguai e Atlântico Sul) é alta em quase todos os tamanhos de série, chegando a superar a incerteza observada na região R3 com séries de aproximadamente 20 anos de dados.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Q95  
(Percentil 75%)

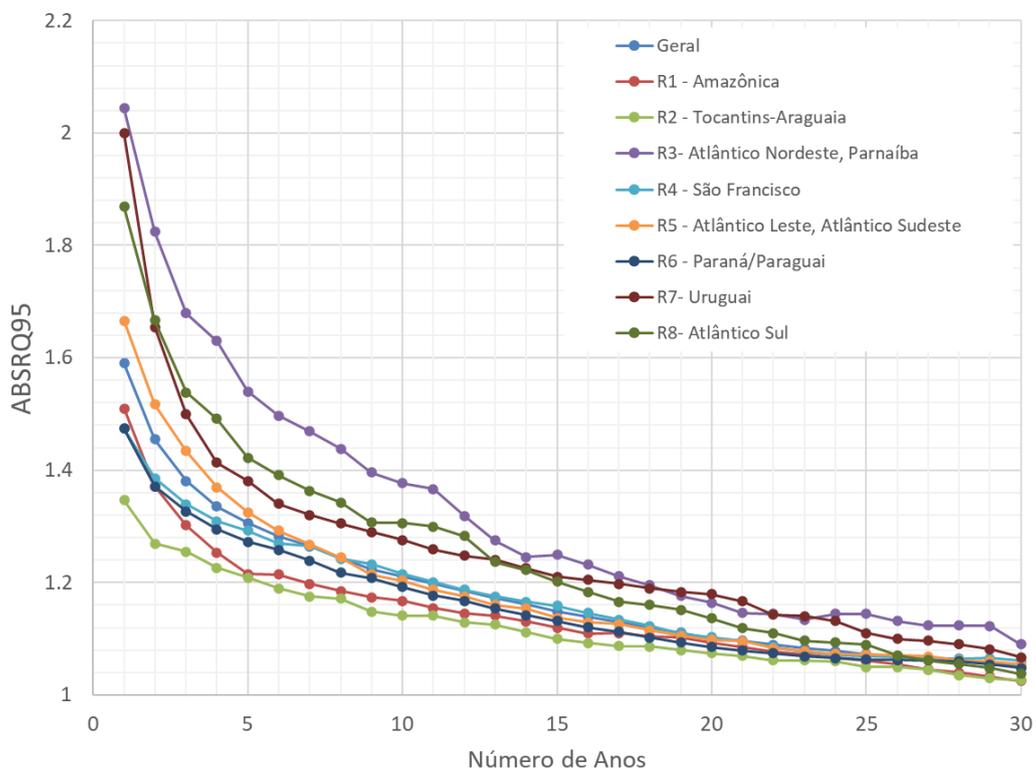


Figura 18: Métrica ABSRQ95 (75% dos casos) da estimativa da vazão média em função do número de anos de dados (no eixo horizontal) quando comparada à estimativa da vazão média com base em 30 a 35 anos de dados, para diferentes regiões hidrográficas.

A Figura 19 mostra os erros absolutos da estimativa da Q95 específica (percentil 75%) com séries de 1 a 30 anos de dados, considerando os postos fluviométricos de diferentes regiões hidrográficas do país. A região que se destaca com menor incerteza em função da variabilidade amostral é a região R3, em que há uma grande proporção dos postos fluviométricos em que a Q95 é igual a zero, ou muito próxima de zero.

Efeito da variabilidade amostral  
no erro da estimativa da Vazão Q95 Específica por região  
(Percentil 75%)

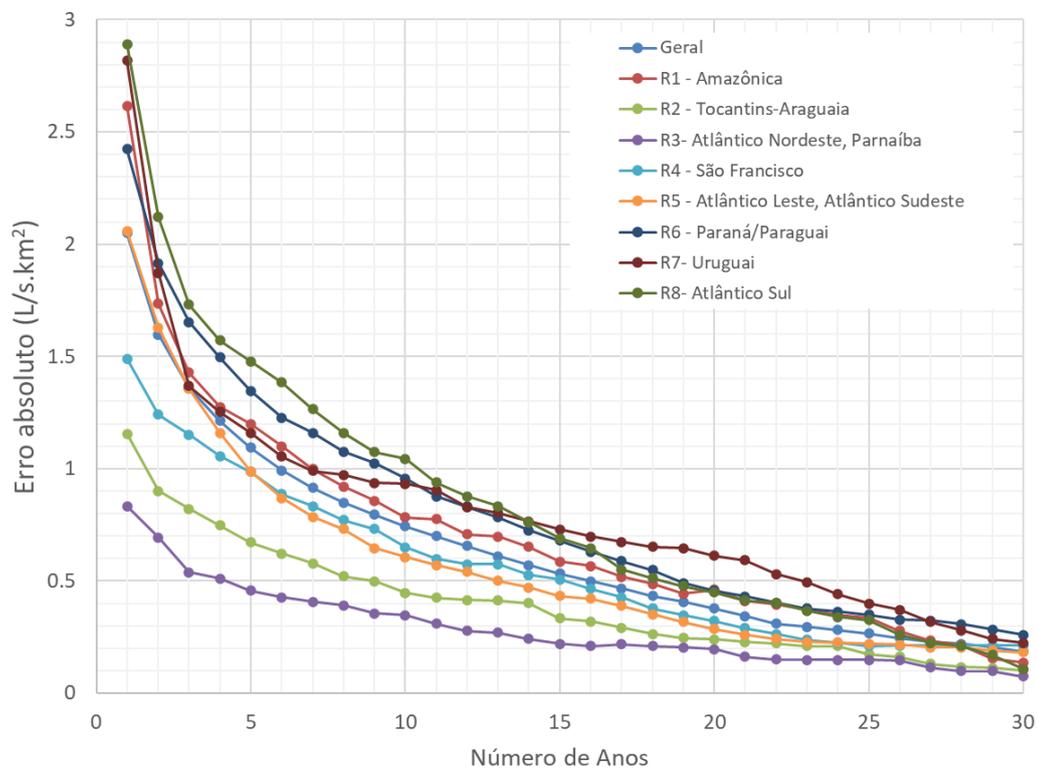


Figura 19: Erro absoluto da Q95 específica estimado com séries curtas de dados (1 a 30 anos) em relação à Q95 específica estimada com série longa (30 a 35 anos) na avaliação geral e por região hidrográfica. As linhas representam o percentil 75%.

## 5 CONCLUSÕES

---

Neste capítulo apresentamos a análise do efeito da variabilidade amostral sobre as estimativas de duas vazões de referência (vazão média e Q95) a partir de dados de 663 postos fluviométricos do Brasil, com o objetivo de avaliar a influência do tamanho da série de dados utilizados na estimativa da vazão sobre a acurácia da estimativa.

A análise dos resultados mostra que a estimativa da vazão de referência sofre com a variabilidade amostral, possivelmente associada com a variabilidade climática, que resulta em períodos de múltiplos anos mais secos ou mais úmidos do que a média de longo prazo. Assim, se a série de dados disponível num local com pouca disponibilidade coincide com um período de anos mais úmidos (secos) do que a média, a vazão de referência possivelmente será ligeiramente maior (menor) do que seria caso houvesse uma série de dados mais longa e mais representativa.

De maneira geral, observa-se que a acurácia da estimativa aumenta consideravelmente com o aumento do tamanho da série de dados.

As vazões de referência mínimas (Q95) apresentam maior variabilidade amostral do que as vazões médias.

Considerando o conjunto dos postos fluviométricos analisados em todo o Brasil, e considerando aceitável uma incerteza igual ou menor do que 10% na maioria (75%) dos casos, os resultados mostram que são necessários 13 anos de dados para uma boa estimativa da vazão média e 20 anos de dados para uma boa estimativa da Q95.

Em termos de erros relativos, a variabilidade amostral tende a ser maior em postos fluviométricos com vazão mais baixa, e tende a ser maior na região hidrográfica R3, que coincide, em grande parte, com áreas de clima semiárido. A variabilidade amostral é maior do que a média também na região R5, que também inclui regiões semiáridas. No caso da vazão de referência Q95 a variabilidade amostral também é maior do que a média no Sul do Brasil, nas regiões hidrográficas R7 e R8. Na Amazônia e na bacia dos rios Tocantins e Araguaia a variabilidade amostral é relativamente mais baixa, sugerindo que estimativas razoáveis de vazões de referência podem ser obtidas com menos anos de dados observados.

Em termos de erros absolutos de vazão específica, os maiores erros associados à variabilidade amostral tendem a ocorrer nos postos fluviométricos dos rios da região Sul do Brasil.

Os valores da incerteza amostral encontrados no presente trabalho têm a mesma ordem de grandeza que os erros encontrados na Áustria, relatados em um estudo publicado por Laaha e Blöschl (2005).

## 6 REFERÊNCIAS

---

Laaha, G., & Blöschl, G. (2005). Low flow estimates from short stream flow records—a comparison of methods. *Journal of Hydrology*, 306(1-4), 264-286.

Tallaksen, L. M., & Van Lanen, H. A. (Eds.). (2004). *Hydrological drought: processes and estimation methods for streamflow and groundwater*.