

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

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

**RELATÓRIO TÉCNICO: ESTIMATIVAS DE
VAZÕES DE REFERÊNCIA COM MODELO MGB-
SA**

IPH-ANA-HGE-SR-R9

Porto Alegre - RS

Outubro 2021

Estimativas de vazões de referência com modelo MGB-SA



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: Cléber Henrique de Araújo Gama, Larissa de Castro Ribeiro, Rodrigo Cauduro Dias de Paiva, Walter Collischonn e Mino Viana Sorribas.

COMO CITAR: Gama, C. H. de A., Ribeiro, L. de C., Paiva, R. C. D., Collischonn, W., Sorribas, M. V., 2021. Cooperação em tecnologias para análises hidrológicas em escala nacional: Relatório técnico: Estimativas de vazões de referência com modelo MGB-SA: IPH-ANA-HGE-SR-R9. UFRGS: IPH, [Porto Alegre]. ANA, [Brasília].

Porto Alegre - RS

Outubro 2021

Sumário

1	Apresentação	4
2	Fundamentos	5
2.1	Modelagem Hidrológica chuva vazão	5
2.2	Reanálise hidrológica – Simulações com assimilação de dados	7
2.3	Ensemble Kalman Filter (EnKF) e Local EnKF	10
3	Retrospectiva Hidrológica	13
3.1	Consolidação do modelo MGB-AS para simulação de vazões naturais	13
3.2	Seleção de bases de dados pluviométricas e climáticas	14
3.3	Desenvolvimento de Retrospectiva hidrológica utilizando MGB-AS	14
3.3.1	Análise de erro das vazões de referência geradas através de Modelagem Hidrológica	17
3.3.2	Estimativa do erro da Q95	18
3.3.3	Análise do erro de estimativa das vazões de referência MGB-AS 1990-2010, postos originais	18
3.3.4	Análise da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para o período de 1990-2010, postos originais	21
3.3.5	Análise do erro de estimativa das vazões de referência MGB-AS 1990-2010, postos complementares	23
3.3.6	Análise da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para o período de 1990-2010, postos complementares	26
3.3.7	Análise do erro de estimativa das vazões de referência, MGB-AS 1980-2015	28
3.3.8	Análise da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para o período de 1980-2015	31
4	Reanálise Hidrológica – Simulações com assimilação de dados	34
4.1	Definição de estações fluviométricas para validação da simulação com assimilação de dados	34
4.2	Consolidação do MGB-AS-EnKF: Definição de erros e parâmetros de assimilação	36
4.3	Resultados	36
5	Referências Bibliográficas	50

1 APRESENTAÇÃO

O presente documento apresenta o relatório “IPH-ANA-HGE-SR-R9” do subprojeto “Regionalização 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.

Neste relatório será apresentado o desempenho do modelo MGB-SA para estimativa de vazões de referência, bem como o desempenho a partir da reanálise hidrológica – simulação utilizando uma técnica de assimilação de dados para consideração de vazões observadas e correção do estado do modelo. Por fim, a base de dados final compreende: *shapefiles* com as vazões de referência simuladas, os respectivos erros e uma banda de incerteza, obtida através dos membros da simulação com assimilação de dados, apresentado no volume IPH-ANA-HGE-SR-B1; e arquivo com séries de vazões simuladas, na rede do modelo MGB-SA, com as versões do modelo *open-loop* e com assimilação de dados, disponível na base IPH-ANA-HGE-SR-B3.

2 FUNDAMENTOS

2.1 MODELAGEM HIDROLÓGICA CHUVA VAZÃO

Modelos hidrológicos de bacias hidrográficas são ferramentas criadas para extrapolar medições hidrológicas no tempo e no espaço. Os modelos do tipo conceitual são baseados em alguns princípios físicos fundamentais, como a equação de conservação de massa. Os modelos denominados “de base física” utilizam estes mesmos princípios, além de algumas outras leis físicas, como conservação de energia e quantidade de movimento. Já os modelos empíricos são desenvolvidos buscando apenas uma boa representação das variáveis de saída, sem procurar representar os processos internos do sistema físico.

Mesmo os modelos de base física e conceituais têm embutidos diversos processos que são representados de forma empírica. Dessa forma, existe a necessidade de calibrar os parâmetros destes modelos através da comparação dos valores calculados com os valores observados de variáveis como a vazão. Isto dificulta, obviamente, a aplicação de modelos do tipo chuva-vazão para a estimativa de vazão em locais sem dados.

A aplicação de um modelo do tipo chuva vazão em uma bacia sem dados normalmente é precedida pela calibração anterior em uma ou mais bacias com dados. Depois de estabelecidos os valores dos parâmetros na ou nas bacias com dados, os parâmetros do modelo são transferidos para a bacia sem dados. Há diversas metodologias para esta transferência, conforme mostra uma revisão recente de Guo et al. (2020). Duas estratégias principais são utilizadas para a aplicação de modelos do tipo chuva-vazão em locais sem dados.

A primeira estratégia consiste em calibrar o modelo em diferentes bacias e relacionar os valores dos parâmetros, obtidos por calibração, com valores de outras variáveis que podem ser mais facilmente obtidas, como o percentual de argila no solo, a declividade do terreno, e o percentual de áreas urbanas, através de uma relação empírica. Posteriormente o modelo é aplicado na bacia sem dados, com os valores dos parâmetros estimados pela relação empírica obtida nas bacias com dados. Esta estratégia é muito adotada em modelos hidrológicos do tipo concentrado, em que alguns poucos parâmetros representam toda a área da bacia.

A segunda estratégia consiste em dividir a região em classes de resposta hidrológica, com base em mapas de tipos de solos, tipos de litologia, tipos de vegetação e de características do relevo, e, durante a etapa de calibração, associar os valores dos parâmetros aos valores das classes de resposta hidrológica. Esta estratégia é mais frequentemente adota em modelos hidrológicos do tipo distribuído.

Alguns exemplos da aplicação de modelos hidrológicos chuva-vazão do tipo chuva vazão para a estimativa de vazões em locais sem dados no Brasil são apresentados por Alexandre et al. (2005), Diniz e Clarke (2001), Saraiva et al. (2011), Otsuki e Reis (2011) e Virões e Cirilo (2019).

Alexandre et al. (2005) aplicaram o modelo SMAP e relacionaram dois dos seus principais parâmetros às seguintes variáveis explanatórias: capacidade de armazenamento médio do solo; precipitação média anual e porcentagem da bacia situada sobre rochas cristalinas.

Otsuki e Reis (2011) testaram a aplicação do modelo SMAP em bacias sem dados, e regionalizaram os parâmetros desse modelo considerando cinco variáveis explanatórias: 1) área

da bacia; 2) precipitação média anual; 3) comprimento do rio principal; 4) declividade do rio principal; 5) fator de forma da bacia.

Virões e Cirilo (2019) avaliaram a aplicação do modelo MODHAC com transferência de parâmetros de bacias vizinhas, na região NE do Brasil, e observaram grandes erros na estimativa de vazões baixas.

Segundo o estudo apresentado por Wongchuig et al., (2020), os modelos hidrológicos globais, que são aplicados normalmente com pouca ou nenhuma calibração, estimam vazões com erros nas vazões diárias que podem ser da ordem de 52% a 103% (RMSE) para a região da América do Sul.

Experimentos realizados em outros países também sugerem que a estimativa de vazões mínimas em locais sem dados utilizando modelos hidrológicos do tipo chuva-vazão apresentam erros relevantes, especialmente nas vazões mais baixas.

Gudmundsson et al. (2012) analisaram o desempenho de modelos de grande escala em reproduzir vazões na Europa e concluíram que os modelos hidrológicos são particularmente ruins em prever as vazões mais baixas.

Engeland e Hisdal (2009) compararam estimativas de vazões mínimas obtidas com equações de regressão e com o modelo hidrológico HBV em rios da Noruega, e constataram que as estimativas baseadas no modelo hidrológico eram, em geral, piores do que as estimativas baseadas em equações de regressão.

Petheram et al. (2011) realizaram um estudo semelhante em bacias do Norte da Austrália e chegaram a mesma conclusão: para a estimativa de vazões mínimas em locais sem dados a metodologia baseada na modelagem hidrológica chuva-vazão apresenta erros maiores do que um equação de regionalização baseada em regressão.

Murphy et al. (2012) compararam a modelagem hidrológica chuva-vazão e a regionalização por equações de regressão na estimativa de diversos indicadores hidrológicos com relevância ecológica, incluindo a vazão Q85, na bacia do rio Tennessee, e concluíram que a metodologia baseada no modelo hidrológico chuva-vazão apresenta erros maiores. Na estimativa da vazão Q85 o erro médio em 5 postos fluviométricos foi de 169% com a modelagem hidrológica e de 87% com a equação de regressão. Entretanto, é possível que este mau desempenho da modelagem hidrológica frente à regionalização empírica seja, em grande parte, devido à estratégia adotada para calibração do modelo hidrológico. As métricas utilizadas para a calibração do modelo hidrológico, citadas por Murphy et al. (2012) tendem a valorizar mais os erros em vazões altas do que os erros em vazões baixas.

Situação semelhante é relatada por Farmer et al. (2014) que compararam o modelo hidrológico PMRS a várias abordagens empíricas para estimativa de vazões em locais sem dados. Os erros da estimativa da Q90 do modelo PMRS foram maiores do que dos outros métodos testados. Novamente deve se ressaltar que o modelo hidrológico não foi calibrado com foco em vazões baixas, como a Q90.

Hailegeorgis e Alfredsen (2017) também compararam estimativas baseadas na regressão e na modelagem hidrológica chuva vazão e concluem da mesma forma, isto é, que indicadores hidrológicos ecologicamente relevantes obtidos por modelagem hidrológica chuva-vazão apresentam erros maiores.

Uma das possíveis causas para o mau desempenho dos modelos hidrológicos chuva vazão na representação das vazões mais baixas pode ser a forma com que os modelos hidrológicos são normalmente calibrados. As métricas utilizadas para a calibração de modelos, como o coeficiente de eficiência de Nash-Sutcliffe, podem induzir a um melhor ajuste das vazões altas, em detrimento das baixas.

Pfannerstill et al. (2014) sugerem algumas métricas que poderiam ser utilizadas para uma calibração mais homogênea, entre vazões altas e baixas.

Esta estratégia foi adotada na aplicação do modelo distribuído MGB para a estimativa de disponibilidade hídrica no contexto de alguns planos de bacia elaborados no Brasil ao longo da última década. Por exemplo, no plano de bacia da região hidrográfica dos rios Macaé e das Ostras, no RJ, o modelo MGB foi calibrado buscando minimizar os erros da estimativa da Q90 em cada um dos postos fluviométricos com dados. Posteriormente, os resultados do modelo hidrológico foram utilizados para calcular a vazão Q7,10, utilizada como referência na bacia. No entanto não foram apresentadas estimativas da acurácia do método em prever vazões baixas em locais sem dados.

Acredita-se que o desempenho relativamente ruim de modelos hidrológicos do tipo chuva-vazão na representação de vazões baixas, como a Q90, a Q95 e a Q7,10, possa ser, ao menos parcialmente, melhorado através da utilização de técnicas de assimilação de dados de vazão em alguns postos fluviométricos e pós processamento, conforme descrito no item 2.2.

2.2 REANÁLISE HIDROLÓGICA – SIMULAÇÕES COM ASSIMILAÇÃO DE DADOS

Como abordado no item 2.1, a modelagem hidrológica pode ser uma ferramenta interessante no processo de fornecer informações em locais pouco ou não-monitorados. Sua principal vantagem em relação à métodos clássicos de regionalização de vazão é a possibilidade de geração de uma série temporal de dados e não apenas estimativas pontuais. Contudo, qualquer processo de modelagem está sujeito a erros, refletido em diferenças entre os valores simulados e o que de fato acontece. Estes erros podem ser gerados por diversas fontes, como: estrutura do modelo hidrológico utilizado; dados de entrada utilizados como forçante do modelo; processo de calibração; entre outros.

Com o intuito de diminuir estes erros e reduzir as incertezas associadas ao processo de simulação, as técnicas de Assimilação de Dados (AD) têm sido utilizadas. Estas técnicas baseiam-se em um processo de retroalimentação no qual valores observados de uma ou mais variáveis são utilizados para a correção de variáveis de estado do modelo hidrológico, buscando aproximar as condições simuladas das observações disponíveis.

Desta forma, estes métodos de assimilação, quando aplicados em modelos hidrológicos distribuídos, podem ser utilizados para aprimorar a estimativa de vazões em locais não monitorados (Clark et al., 2008).

A partir da geração de vazões de referência (e.g. vazão média, Q90) os resultados da simulação com assimilação de dados podem também ser utilizados em aplicações como análise de outorga, planos de bacias e estudos ambientais.

Pessoa (2017) utilizou o Modelo de Grandes Bacias (MGB) com uma técnica de assimilação de dados empírica para obtenção de vazões em locais sem dados do Rio Grande do Sul. Ao

comparar informações de vazões de referência (Q10, Q50, Q90, Q95, QmaxTr25(%) e QmaxTr50(%)) estimadas pela modelagem hidrológica com e sem assimilação de dados com métodos tradicionais de regionalização de vazão, foram observados os menores erros para a metodologia com AD e uma grande capacidade desta técnica em corrigir as estimativas a priori geradas pelo modelo, que apresentavam os maiores erros nas vazões de referência. Estes ganhos com a aplicação da técnica de AD foram observados sobretudo para vazões mínimas de referência.

Em uma outra aplicação, uma técnica mais avançada de AD, o Ensemble Kalman Filter (EnKF), foi implementada com o modelo MGB para aplicação na bacia do rio Taquari Antas com o objetivo de obter estimativas de séries temporais de vazão e vazões de referência em locais com poucas informações in-situ (Wongchuig, 2019). A metodologia se demonstrou eficiente em estimar vazões de referência, quando comparadas a métodos simplificados de regionalização. Apesar de apresentar maior complexidade de implantação e utilização, a abordagem com modelo hidrológico e assimilação de dados permite aplicações às quais métodos tradicionais são limitados, por exemplo, a simulação de cenários não registrados como inserção de reservatórios, mudanças climáticas ou mudanças bruscas de uso e ocupação.

Outra grande vantagem da modelagem hidrológica mencionada no item 2.1 é a possibilidade de geração de informação para uma série longa no passado a partir de dados precipitação resultantes de bases de dados de reanálises climáticas.

A Figura 1 apresenta um desenvolvimento recente de Wongchuig et al., (2017) para o estudo de eventos extremos passados na bacia Amazônica, através da fusão do modelo MGB com reanálises climáticas, onde foi possível identificar e estudar as secas e cheias mais extremas do passado e avaliar a sua evolução temporal.

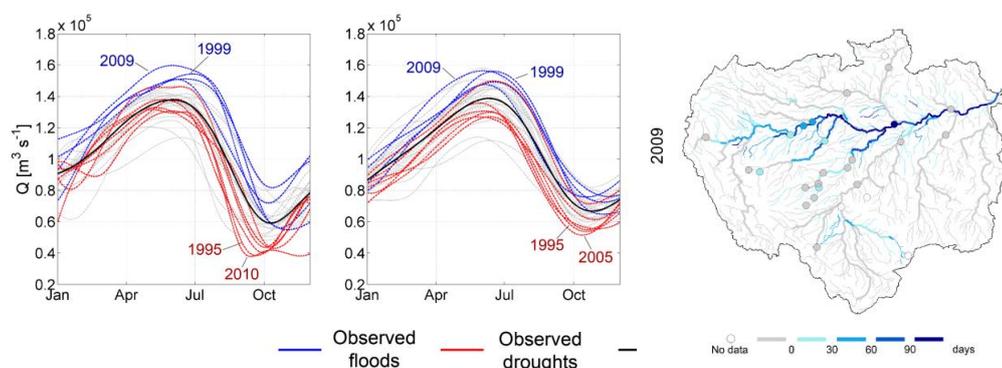


Figura 1 - Hidrogramas de secas (vermelho) e cheias (azul) históricas observadas e simuladas pelo modelo MGB no rio Amazonas e mapa de duração da cheia de 2009 na Bacia Amazônica. (Modificado de Wongchuig Correa et al., 2017).

Este procedimento de simulação, quando combinado com técnicas de assimilação de dados, resulta na metodologia denominada Reanálise Hidrológica (Wongchuig et al., 2019b).

O conceito de reanálise baseia-se em métodos científicos voltados para o desenvolvimento de um registro retrospectivo da evolução de um determinado sistema. Estes métodos foram inicialmente desenvolvidos e aplicados na área de meteorologia e, por isso, conhecidos como reanálises climáticas (e.g. Gibson et al., 1997; Kistler et al., 2001).

Na hidrologia, o conceito de reanálise hidrológica tem a finalidade de permitir um melhor entendimento dos processos hidrológicos a partir de uma série extensa ao passado. Isto

permite, por exemplo, caracterizar ou avaliar: i) o comportamento hidrológico histórico (e.g. multi-decadal, anos recentes, etc.) de certa região, como os seus eventos extremos; ii) as tendências devido a variabilidade climática; iii) melhores estimativas de vazões de referência para gerenciamento dos recursos hídricos, como outorgas para o consumo d'água, planejamento do setor hidrelétrico, e também; iv) fornecer melhores estimativas de outras variáveis hidrológicas como nível d'água, áreas de inundação, umidade do solo ou outras (Wongchuig et al., 2019b).

Esta metodologia de reanálise hidrológica tem sido explorada por algumas iniciativas tais como o GloFAS-ERA5 (Harrigan et al., 2020), base de dados de vazões naturalizados em todo o globo (Lin et al., 2019), PCR-GLOBWB (Wanders et al., 2014), MERRA Land (Reichle e Liu, 2015), ERA-Interim/Land (Balsamo et al., 2015) ou o projeto Earth2Observe. O objetivo principal destes trabalhos é fornecer uma base de dados de reanálises consistentes dos recursos hídricos, muitos deles iniciando no ano de 1979. Em contrapartida, estas reanálises hidrológicas, por serem produtos globais, tendem a usar simplificações nos processos hidrológicos e hidráulicos o que possibilita o aumento das incertezas nos resultados em estudos regionais.

Em uma aplicação regional, Wongchuig et al. (2019b) desenvolveram a reanálise hidrológica do século 20 na bacia Amazônica para o período de 1900 até 2010 mediante o uso do MGB forçado pela reanálise de chuva da base de dados ERA-20CM com remoção de viés e a assimilação de dados (EnKF) de diversas observações in-situ. Este trabalho resultou num registro extenso de variáveis hidrológicas-hidráulicas, denominado de Reanálise Hidrológica do Século 20 (RHXX) e demonstrou melhorias consideráveis na acurácia de estimativas de séries de vazões diárias (Figura 2).

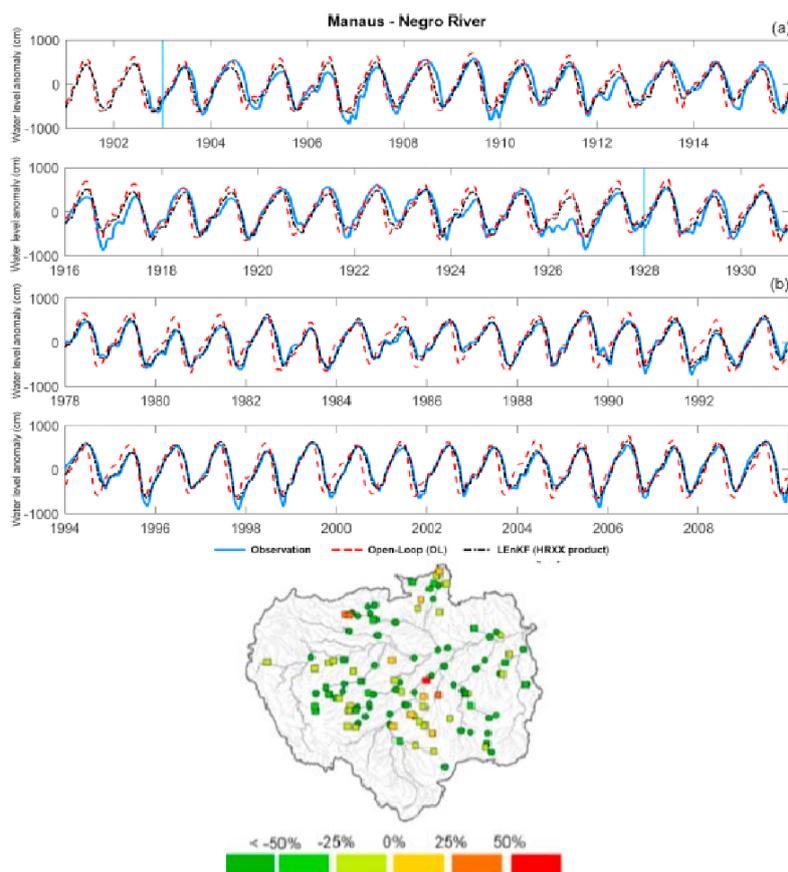


Figura 2 - Validação de série histórica em Manaus com observação de estação fluviométrica (linha azul), simulação com o modelo MGB (linha vermelha) e simulação melhorada com remoção de viés de precipitação e assimilação de observações de vazões (linha preta). Diminuição % do erro RMSE nos postos fluviométricos. Modificado de Wongchuig et al. (2019b).

Este RHXX se diferencia de produtos globais existentes devido à menor simplificação dos processos hidrológicos e hidráulicos (e.g. modelo hidrológico semidistribuído e propagação hidrodinâmica nos rios). No processo foi relatado que a modelagem hidrológica e a AD permitiram melhorar as estimativas de diferentes variáveis hidrológicas-hidráulicas tanto do passado (~100 anos) como do presente.

Neste sentido, a utilização da modelagem hidrológica com assimilação de dados para o estudo de vazões de referência demonstra-se como uma metodologia promissora, seja para o período presente/futuro ou para o passado, a partir do conceito de reanálise hidrológica. Suas principais vantagens estão baseadas i) na geração de uma série temporal de dados; ii) em estimativas mais acuradas, comparadas com modelagem hidrológica sem AD ou métodos tradicionais; iii) flexibilidade para simulação de diferentes cenários, e ainda; iv) a versatilidade para consideração de informações de diversas fontes, podendo considerar, por exemplo, novas estimativas de sensoriamento remoto nas simulações.

2.3 ENSEMBLE KALMAN FILTER (ENKF) E LOCAL ENKF

O objetivo da assimilação de dados é combinar as estimativas incertas do modelo com as informações de dados observados, a fim de fornecer uma otimização dos valores das variáveis de estado do modelo, corrigindo, assim, o estado do modelo em cada instante de simulação. A

função que representa a dinâmica do sistema simulado pode ser indicada pela equação 1 e as observações são dadas pela relação da equação 2.

$$x_{t+1} = M(x_t, u, \theta) + q_t \quad (1)$$

$$y_t = H(x_t) + \varepsilon_t \quad (2)$$

onde, q_t representa os erros, podendo ser referentes a estrutura do modelo M , estimativas de parâmetros θ , forçantes ou estados antecedentes u ; ε_t é o vetor de erros das observações e H é o operador que relaciona as variáveis de estado x_t às observações y_t .

Grande parte das aplicações hidrológicas dos métodos de assimilação de dados utiliza esquemas baseados no Filtro de Kalman (KF, Kalman, 1960). O KF é um método sequencial de assimilação de dados, com solução de mínimo erro, que utiliza técnicas variacionais e também propaga informação dos erros para os passos de tempo seguintes, computando as covariâncias dos erros em uma matriz adicional (Reichle, 2008).

No caso do Filtro de Kalman original, a AD é resolvida usando um estimador linear assumindo que i) os operadores de modelo e observação são lineares; ii) erros de observação são não enviesados e sem correlação temporal e espacial; iii) os erros do modelo não são relacionados e temporalmente não correlacionados e; iv) não há correlação entre modelo e erros das observações.

Para contornar algumas limitações do Filtro de Kalman, como a aplicação a sistemas não-lineares, caso da maioria das relações hidrológicas (Liu e Gupta, 2007), e as desvantagens que, tanto o KF original quanto sua versão não-linear, o KF Estendido (EKF), têm quando aplicados em sistemas grandes e complexos com muitas variáveis de estado (Vrugt et al., 2005), foram desenvolvidas variações do método para adaptar seu uso, sendo um dos mais utilizados o Filtro de Kalman por Ensemble (EnKF), apresentado por Evensen (2003 e 2004).

O método se popularizou por ter uma formulação conceitual simples e por ser relativamente fácil de implementar (Evensen, 2003). O EnKF faz uso do método de Monte Carlo para criar uma função densidade de probabilidade como alternativa à matriz de covariância dos erros (Evensen, 2003).

Em seguida é apresentada a formulação do EnKF de forma simplificada, baseado em Evensen (2003), Evensen (2004) e Katzfuss et al. (2016).

Seja X^b a matriz das variáveis de estado do modelo (equação 3), cuja dimensão é $nestado \times nens$, onde $nestado$ é o número de variáveis de estado e $nens$ é o número de membros do conjunto.

$$X^b = (x_1^b, \dots, x_{nens}^b) \quad (3)$$

A média do conjunto é então definida conforme equação 4.

$$\bar{x}^b = \frac{1}{nens} \sum (x_i^b) \quad (4)$$

A anomalia de cada membro do conjunto, isto é, a diferença de cada membro do conjunto à média é expressa em nova matriz X'^b . Então, uma estimativa da covariância do erro do modelo é calculada a partir da equação 5.

$$P^b = \frac{1}{nens - 1} X'^b X'^b T \quad (5)$$

Assim, as variáveis de estado no modelo são atualizadas, com base nas equações 6a e 6b.

$$x_i^a = x_1^b + K(y_i - Hx_1^b) \quad (6a)$$

$$K = P^b H^T (H P^b H^T + R)^{-1} \quad (6b)$$

$$y = Q_{obs} \quad (7)$$

onde, x_i^a representa a variável atualizada, y_i é o vetor de observações de vazões, H é um operador que leva as variáveis de estado à mesma dimensão das observações, K representa o ganho de Kalman e R é a matriz de covariância das observações.

Para fornecer uma representação adequada da matriz de covariância dos erros, o método EnKF precisa de um número de membros suficiente para evitar correlações espúrias entre grandes rios. Para evitar esse problema, Houtekamer e Mitchell (2001) propuseram ignorar observações por meio do método de “localização de covariância”.

Para tal, utiliza-se uma matriz de correlação, estimada usando uma função de correlação dependente da distância (Gaspari e Cohn, 1999).

As matrizes de correlação são calculadas usando uma função de quinta ordem, conforme determinado por Gaspari e Cohn (1999) e sugerido por Hamill et al. (2001) e Houtekamer e Mitchel (2001), conforme equação 8.

$$\rho(r, e) = \begin{cases} 1 - \frac{1}{4} \left(\frac{e}{r}\right)^5 + \frac{1}{2} \left(\frac{e}{r}\right)^4 + \frac{5}{8} \left(\frac{e}{r}\right)^3 - \frac{5}{3} \left(\frac{e}{r}\right)^2, & 0 \leq e < r; \\ \frac{1}{12} \left(\frac{e}{r}\right)^5 - \frac{1}{2} \left(\frac{e}{r}\right)^4 + \frac{5}{8} \left(\frac{e}{r}\right)^3 + \frac{5}{3} \left(\frac{e}{r}\right)^2 - 5 \left(\frac{e}{r}\right) + 4 - \frac{2}{3} \left(\frac{e}{r}\right)^{-1}, & r \leq e \leq 2r; \\ 0, & e > 2r; \end{cases} \quad (8)$$

onde e representa a distância ao longo da rede de drenagem entre a observação e o ponto analisado e r representa o raio máximo de influência de uma observação.

Então, a incorporação de uma matriz de covariâncias por localização (Sakov e Bertino, 2011) à Equação 7.10, dá origem ao Local Ensemble Kalman Filter (LEnKF) (Houtekamer e Mitchell, 2001, 2005; Ott et al., 2004; Tong, 2018) e a Equação 7.10 passa a ser expressa pela equação 9.

$$x_i^a = x_1^b + \rho(P^f H^T) [\rho_o (H P^f H^T) + R]^{-1} [y_i - H x_1^b] \quad (9)$$

Esse esquema local do EnKF foi aplicado anteriormente ao modelo MGB por **Wongchuig** et al. (2019) e **Wongchuig** et al. (2020) e apresentou melhoria no desempenho do modelo na simulação de séries históricas, bem como melhoria no desempenho em estimativas de vazões de referência.

3 RETROSPECTIVA HIDROLÓGICA

Este capítulo descreve a “Etapa 2 – Desenvolvimento Parte 1” do subprojeto e que remete a consolidação do modelo MGB-AS, considerando o desenvolvimento da retrospectiva hidrológica com o mesmo.

3.1 CONSOLIDAÇÃO DO MODELO MGB-AS PARA SIMULAÇÃO DE VAZÕES NATURAIS

O modelo MGB foi implementado para o domínio da América do Sul (Siqueira et al., 2018), permitindo novas aplicações e estudos usando métodos padronizados e de forma sistemática em todo o território nacional, e denominado MGB-AS. As simulações do modelo continental representam as vazões diárias com acurácia semelhante a aplicações de modelos hidrológicos em escala de grandes bacias hidrográficas (e.g., Amazonas e Prata). São satisfatoriamente representados hidrogramas sazonais de grandes rios, defasados por grande tempo de viagens de ondas de cheia, além de hidrogramas com picos rápidos de rios menores. O coeficiente de eficiência de Kling-Gupta é maior que 0,6 em 70% dos casos, sendo melhor em grandes rios e regiões úmidas.

Nesta implementação a América do Sul foi representada através de cerca de 34 mil minibacias e trechos de rio de aproximadamente 15 km de comprimento. Utilizou-se dados de precipitação da base de dados global Multi-Source Weighted Ensemble Precipitation — MSWEP v1.1 e os parâmetros do modelo MGB foram calibrados a fim de melhorar a acurácia de estimativas de vazões diárias. Os resultados foram validados com mais de 600 postos fluviométricos de instituições nacionais como ANA e ONS e internacionais como IDEAM/Colombia, INA/Argentina, SENAMHI/Peru, SENAMHI/Bolivia, DGA/Chile, ORE-Hybam e GRDC.

Os resultados também foram comparados com modelos hidrológicos globais, concluindo-se que a modelagem continental com o MGB melhorou a acurácia de estimativas de vazões (Siqueira et al., 2018) devido a calibrações manuais, modelagem com física adequada para a região (hidrodinâmica de grandes rios) e experiência a priori sobre a hidrologia da região.

As bases de dados utilizadas para preparação do modelo hidrológico-hidrodinâmico da América do Sul, com as respectivas referências bibliográficas, estão apresentadas na Tabela 1.

Devido à escala espacial do modelo, optou-se por utilizar uma base global de precipitação (com 0.25º de resolução da grade) que combina informações de satélite, reanálise meteorológica e dados observados; Na validação, dados de vazão in situ foram obtidos a partir de instituições de diferentes países, incluindo vazões naturalizadas em diversos pontos controlados por operação de reservatórios. Foram incluídos apenas os postos com mais de 10.000 km² de área de drenagem, totalizando cerca de 600 postos. Parâmetros de geometria dos rios foram extraídos de bases de dados globais e estudos regionais nas bacias Amazônica, Prata e Lagoa dos Patos (RS). Os valores do coeficiente de Manning foram globalmente mantidos em 0.030, com alguns ajustes em tributários específicos da Amazônia conforme Paiva et al. (2013).

Para evitar um excesso de parametrização no modelo devido à resolução mais grosseira dos dados de entrada, foram adotadas sub-bacias delineadas a partir da combinação (intersect) entre um mapa global de litologia / geologia e um mapa das grandes regiões hidrográficas da América do Sul, elaborado a partir de informações da ANA e a Food Agriculture Organization (FAO - Aquastat). Sendo assim, uma calibração tradicional “posto a posto” não foi realizada com

o MGB-SA. Todo procedimento de calibração do MGB-SA foi realizado de forma manual. É importante mencionar que embora uma calibração “posto a posto” permitiria uma melhor acurácia da vazão simulada pelo modelo, outras variáveis relacionadas ao balanço hídrico como evapotranspiração, umidade do solo, e armazenamento terrestre de água potencialmente seriam prejudicadas em detrimento da vazão (dadas as incertezas das bases de dados que são absorvidas no processo de calibração).

3.2 SELEÇÃO DE BASES DE DADOS PLUVIOMÉTRICAS E CLIMÁTICAS

A base de dados pluviométrica utilizada para a atualização do modelo dentro do contexto do subprojeto de regionalização foi a MSWEP v1.1, para o período de Jan/1979 a Dez/2014. A razão desta escolha foi garantir a consistência da base de chuva com àquela adotada no período de calibração do modelo MGB-SA. Por motivo similar, para as demais variáveis climáticas (temperatura, pressão, velocidade do vento, radiação e umidade relativa) foram mantidas as médias de longo-termo do CRU v2.

O MSWEP é uma base de dados global em grade (0.25º de resolução) que combina múltiplas fontes de precipitação de acordo com a qualidade das mesmas, dada em função da escala temporal e posição no espaço. O método desagrega uma precipitação média de longo termo (após correção de viés) com base em uma média ponderada de anomalias de precipitação de sete produtos: dois baseados unicamente na interpolação de observações in situ (CPC Unified e GPCC), três baseados em precipitação de satélite (CMORPH, GsMaP-MVK e TMPA 3B42RT), e dois baseados em reanálise de modelo atmosférico (ERA-Interim e JRA-55). Para cada ponto da grade os pesos atribuídos às observações são calculados em função da densidade de postos, enquanto os pesos atribuídos aos produtos de satélite e reanálises são calculados a partir do desempenho destas bases em relação aos postos mais próximos.

3.3 DESENVOLVIMENTO DE RETROSPECTIVA HIDROLÓGICA UTILIZANDO MGB-AS

A modelagem hidrológica com MGB-AS desenvolvido e demonstrado por Siqueira et al. (2019) considera o período de 1990 a 2010. No processo denominado “Retrospectiva Hidrológica” o modelo foi expandido para realizar a simulação continental, no período de 1980 a 2015.

Em ambas as aplicações, o primeiro ano de cada uma das séries de vazões diárias (1990-2010 ou 1980-2015) obtidas na modelagem hidrológica foram descartados, pois o modelo precisa de um período de aquecimento, ao longo do qual o efeito das condições iniciais ainda é muito alto, podendo gerar maiores incertezas nesse período.

As Figura 3 a Figura 6 demonstram como o modelo agora é capaz de produzir série mais longas, em locais com diferentes regimes hidrológicos.

Tabela 1 - Bases de dados utilizadas no MGB-AS

INFORMAÇÃO	DESCRIÇÃO DOS DADOS	REFERÊNCIA
Modelo digital de elevação (MDE)	Bare-Earth SRTM v 1.0 DEM, upscaled de 3 p/ 15 arcsec (~90m para ~500m)	O'Loughlin et al. (2016)
Mapa de direções de fluxo	Global HydroSHEDS Flow Direction map, 15 arcsec	Lehner et al. (2008)
Precipitação	Global Multi-Source Weighted Ensemble Precipitation v 1.1 (MSWEP), 0.25º	Beck et al. (2017)
Clima	Médias mensais climatológicas do <i>Climate Research Unit</i> (CRU Dataset v. 2.0)	New et al. (2002)
Dados de vazão	Observações in situ - ANA (Brasil), Vazões naturalizadas de reservatórios (ONS - Brasil), INA (Argentina), IDEAM (Colômbia), DGA (Chile), SENAMHI (Peru e Bolívia), HyBAM (Internacional), Global Runoff Data Centre (Internacional)	-
Uso do Solo / Tipo do solo	Mapa de Blocos para América do Sul	Fan et al. (2015)
Regiões hidrográficas	Divisões hidrográficas brasileiras (ANA): http://metadados.ana.gov.br/geonetwork/srv/pt/main.home Mapa Aquastat: http://www.fao.org/geonetwork/srv/en/metadata.show?id=37174	-
Largura / Profundidade dos rios	Relações geomorfológicas de estudos regionais - Amazônia, Prata, bacia da Lagoa dos Patos (RS) Global bankfull width and depth database - demais localidades.	Andreadis et al. (2013), Paiva et al. (2013), Pontes (2016), Lopes (2017), Beighley e Gummadi (2011)

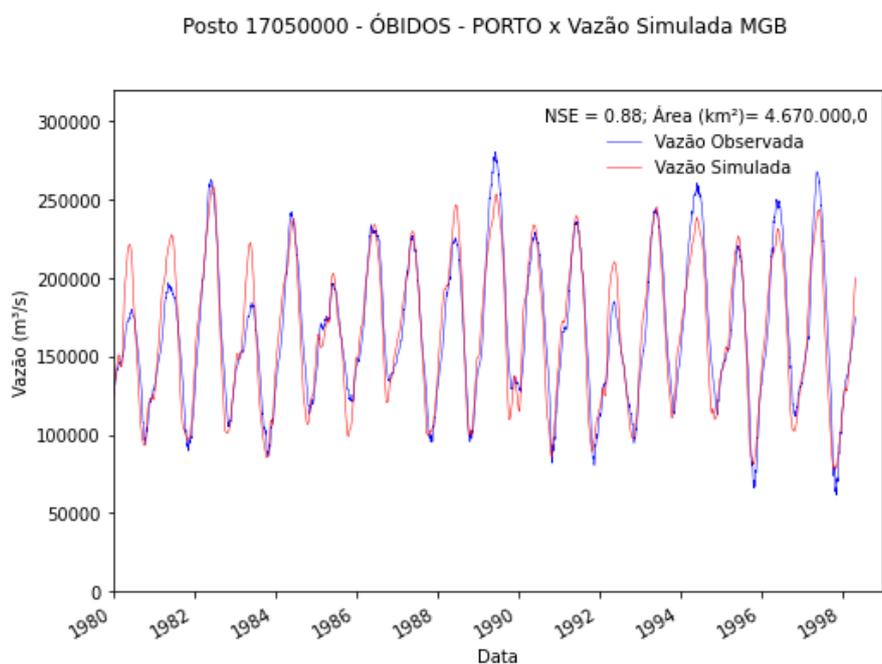


Figura 3. Hidrograma simulado e observado no posto fluviométrico (Rio Amazonas).

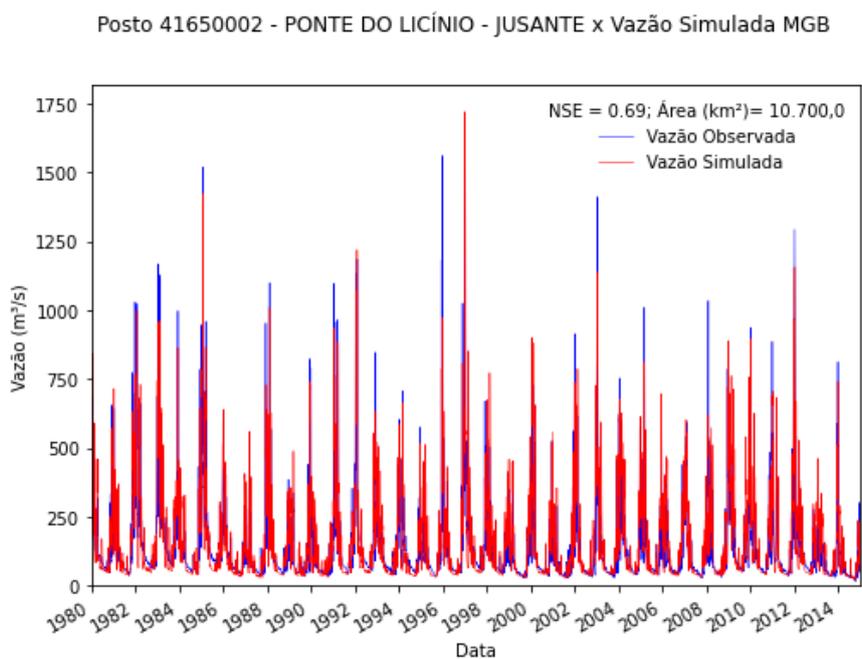


Figura 4. Hidrograma simulado e observado no posto fluviométrico Ponte do Licínio – Jusante Jusante (Rio das Velhas).

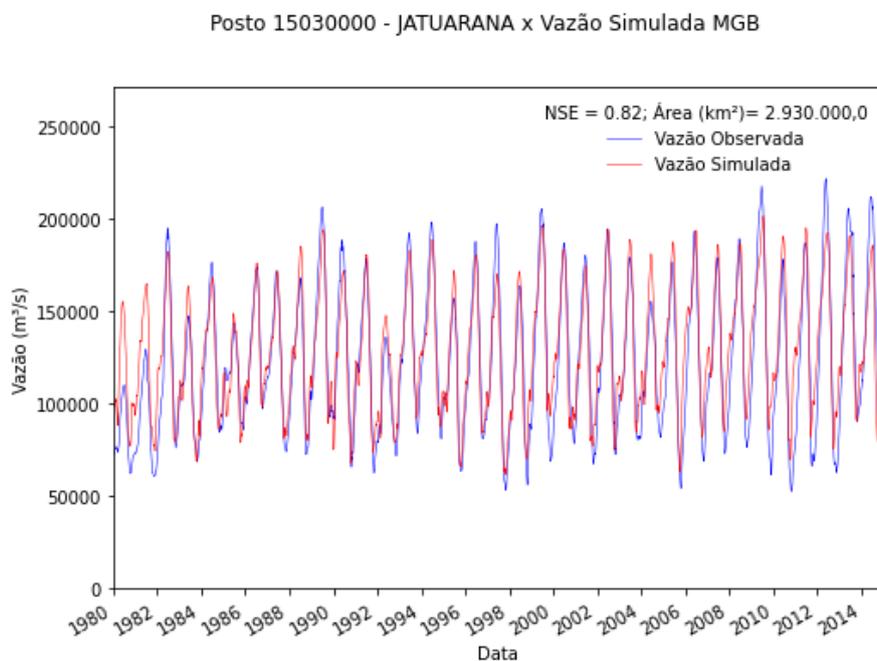


Figura 5. Hidrograma simulado e observado no posto fluviométrico Jatuarana (Rio Amazonas)

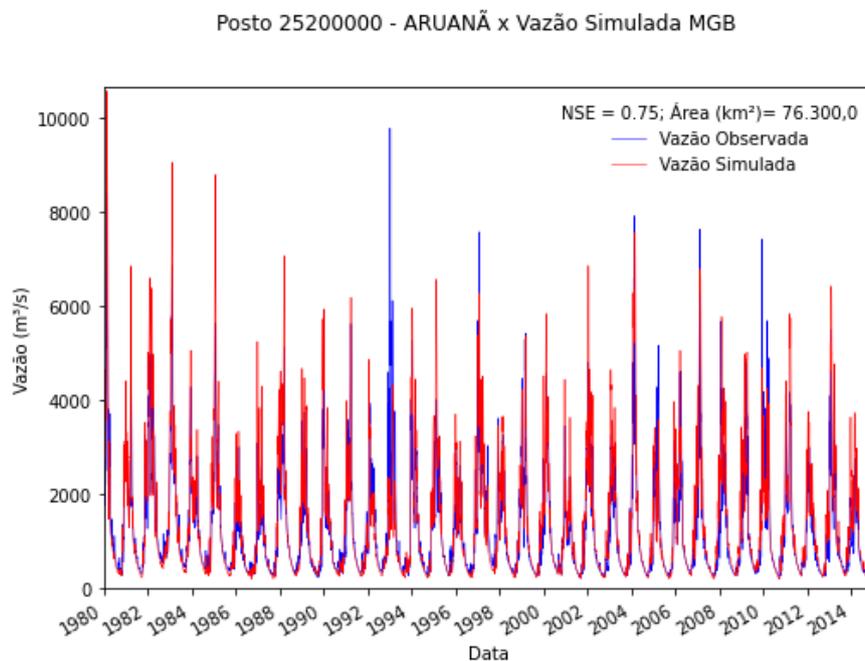


Figura 6. Hidrograma simulado e observado no posto fluviométrico Aruanã (Rio Araguaia).

3.3.1 Análise de erro das vazões de referência geradas através de Modelagem Hidrológica

Foi realizada uma análise preliminar de erro das vazões de referência Q_{95} , onde foi efetuado o cálculo de erro relativo e cálculo da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para a região do Brasil, geradas através de modelagem hidrológica continental do MGB-AS, comparando-as com as vazões referência Q_{95} observadas, calculadas posteriormente para os postos fluviométricos descritos anteriormente na seção 0. Para o cálculo do erro foi relacionado uma minibacia, delimitadas na

modelagem hidrológica, para cada posto fluviométrico, em relação a base hidrográfica do MGB-AS.

3.3.2 Estimativa do erro da Q95

O erro da estimativa da Q95 foi calculado através de erro relativo pela área de drenagem, através da equação 4.

$$ER(\%) = \left(\frac{\left(\frac{Q_{95cotrecho}}{\text{Área}_{cotrecho}} \right) - \left(\frac{Q_{95posto}}{\text{Área}_{posto}} \right)}{\left(\frac{Q_{95posto}}{\text{Área}_{posto}} \right)} \right) \times 100 \quad (4)$$

onde, $ER(\%)$ é o erro relativo; $Q_{95cotrecho}$ representa a vazão com 95% de permanência do cotrecho; $\text{Área}_{cotrecho}$ é a área de drenagem à montante do cotrecho; $Q_{95posto}$ é a vazão de referência Q95 do posto; Área_{posto} é a área de drenagem montante do posto.

A razão $Q_{95modelo}/Q_{95obs}$ foi calculada através da equação 5 descrita a seguir, onde é considerado o sinal da operação da vazão de referência simulada subtraindo a vazão de referência observada e este sinal multiplicando o valor máximo entre a razão da vazão de referência simulada com a observada. Esta métrica indica a proporção em que o modelo estima as vazões em relação as vazões observadas, ou seja, se o valor do RQ95 for de 2 (por exemplo), isto indica que a vazão simulada é 2 vezes maior que a observada.

$$RQ95 = \text{Sinal}(Q_{95modelo} - Q_{95observado}) \cdot \text{Máximo}\left(\frac{Q_{95modelo}}{Q_{95observado}}; \frac{Q_{95observado}}{Q_{95modelo}}\right) \quad (5)$$

onde, $RQ95$ é a razão $Q_{95modelo}/Q_{95obs}$; $Q_{95modelo}$ é a vazão de referência Q95 simulada pelo modelo; $Q_{95observado}$ é a Vazão de referência Q95 observado do posto fluviométrico em análise.

Valores negativos significam que o modelo subestimou as vazões observadas e valores positivos que superestimou. Quando subestimativa será utilizada a razão $Q_{95obs}/Q_{95modelo}$, quando superestimativa será utilizada a razão $Q_{95modelo}/Q_{95obs}$.

3.3.3 Análise do erro de estimativa das vazões de referência MGB-AS 1990-2010, postos originais

Para fins de comparação, apresenta-se o mapa do erro relativo da Q95 (Figura 7) e o mapa do coeficiente de Nash-Sutcliffe (Figura 8) para simulação do período 1990-2010, ilustrando somente postos fluviométricos utilizados no modelo MGB América do Sul original, sendo o total de postos de 524.

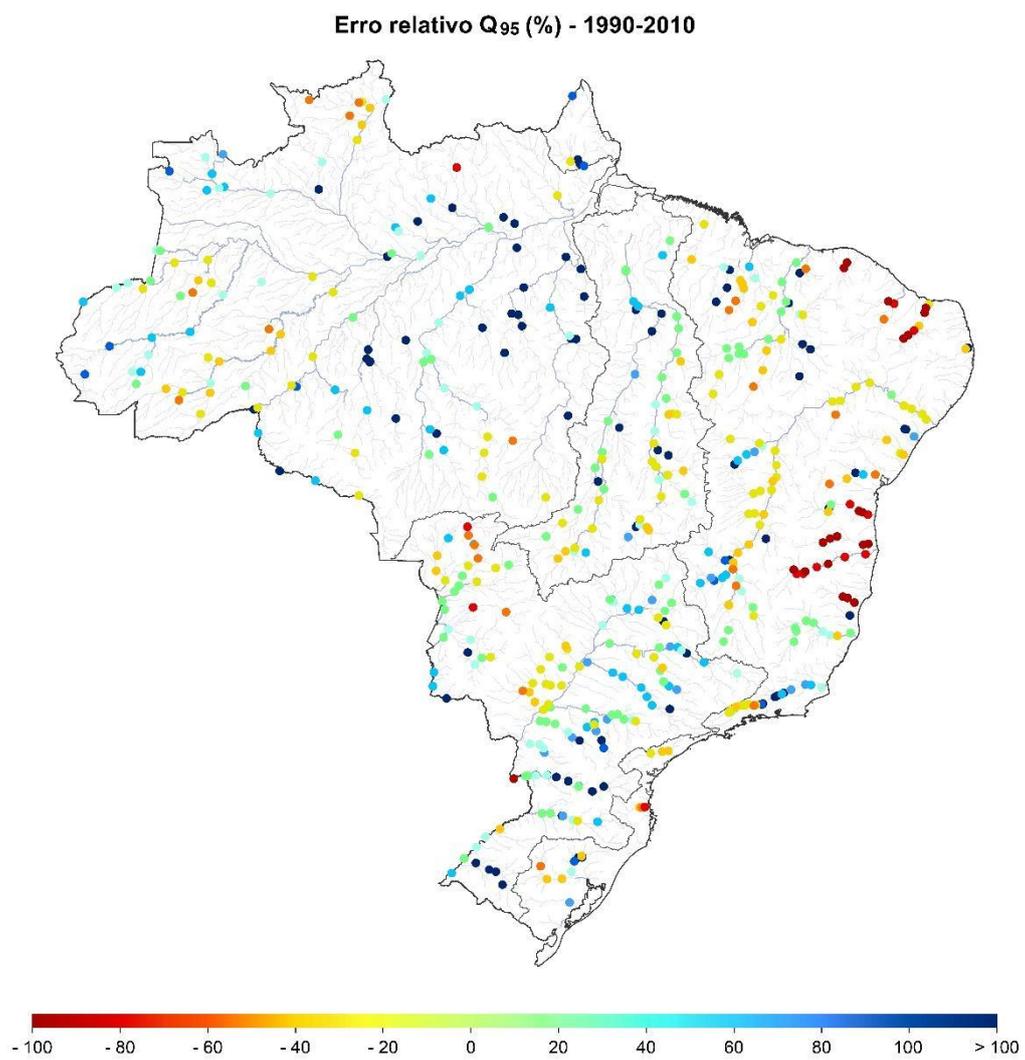


Figura 7. Distribuição espacial do erro relativo da Q_{95} para o período de 1990-2010 (postos originais)

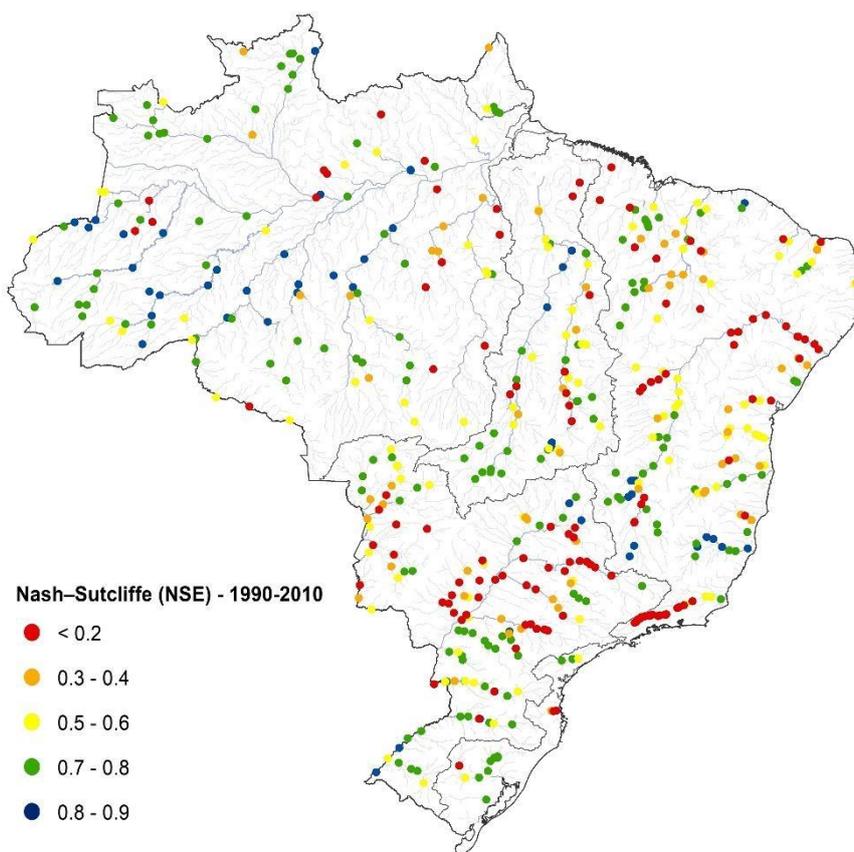


Figura 8. Coeficiente Nash-Sutcliffe (NSE) para o período de 1990-2010 (postos originais)

A Tabela 2 apresenta um resumo dos erros da Q_{95} de todos esses postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores dos erros são maiores para áreas de drenagens menores e os postos se acumulam em maior proporção na faixa de área de drenagem de 25.000 a 50.000 km^2 , o desvio padrão da mesma magnitude da média, além de uma assimetria na distribuição dos erros. A mediana (média) ficou em 18% (51%) para a faixa de 100.000 a 500.000 km^2 e em 45% (253%) para a faixa de 1.000-25.000 km^2 .

Tabela 2 - Análise do erro relativo em relação a área de drenagem para a Q_{95} para o período de 1990-2010 (postos originais)

ÁREA DE DRENAGEM (KM ²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	223	253.14	44.56	363.94
25.000 - 50.000	126	96.14	35.74	104.41
50.000 - 100.000	67	99.91	35.32	112.61
100.000 - 500.000	108	51.52	17.77	57.07

*Erros relativos na Q_{95} em cada intervalo de área de drenagem

A Figura 9 e a Figura 10 apresentam a distribuição de frequência acumulada do valor absoluto do erro relativo da Q_{95} e do coeficiente de Nash-Sutcliffe nos postos analisados. Pode-se

observar que em 50% dos pontos analisados o erro relativo é menor que 33% e que em 83% dos pontos o erro é menor que 100%. Em 50% dos pontos analisados o NSE é maior que 0.55.

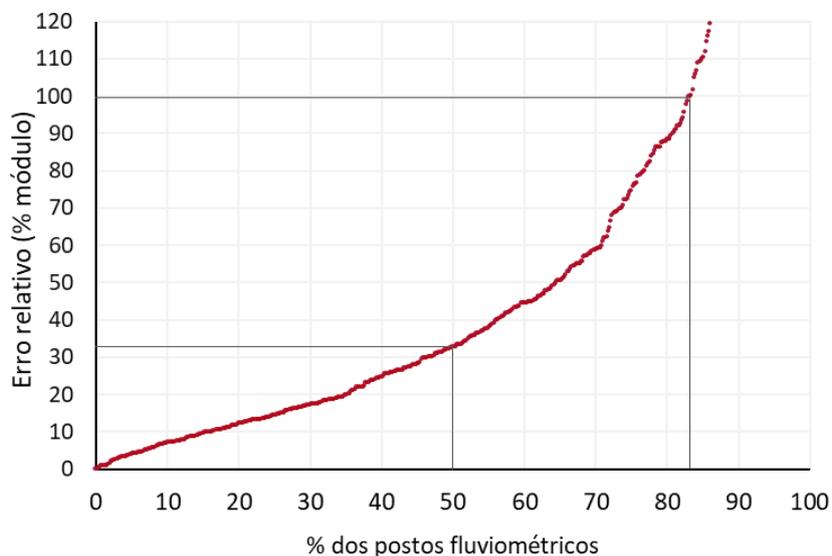


Figura 9. Relação do módulo do erro relativo da Q95 com a quantidade de pontos analisados para o período de 1990-2010 (postos originais)

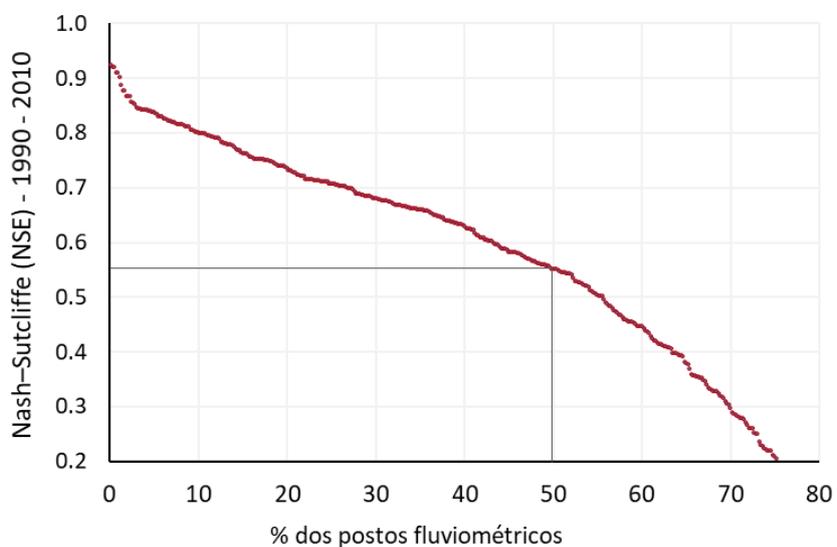


Figura 10. Relação coeficiente Nash-Sutcliffe (NSE) com a quantidade de pontos analisados para o período de 1990-2010 (postos originais)

3.3.4 Análise da Razão $Q_{95\text{modelo}}/Q_{95\text{obs}}$ para o período de 1990-2010, postos originais

A Figura 11 apresenta o mapa da Razão $Q_{95\text{modelo}}/Q_{95\text{obs}}$ para simulação do período 1990-2010 (postos originais), ilustrando somente postos fluviométricos utilizados na calibração do modelo original, sendo o total de postos de 524.

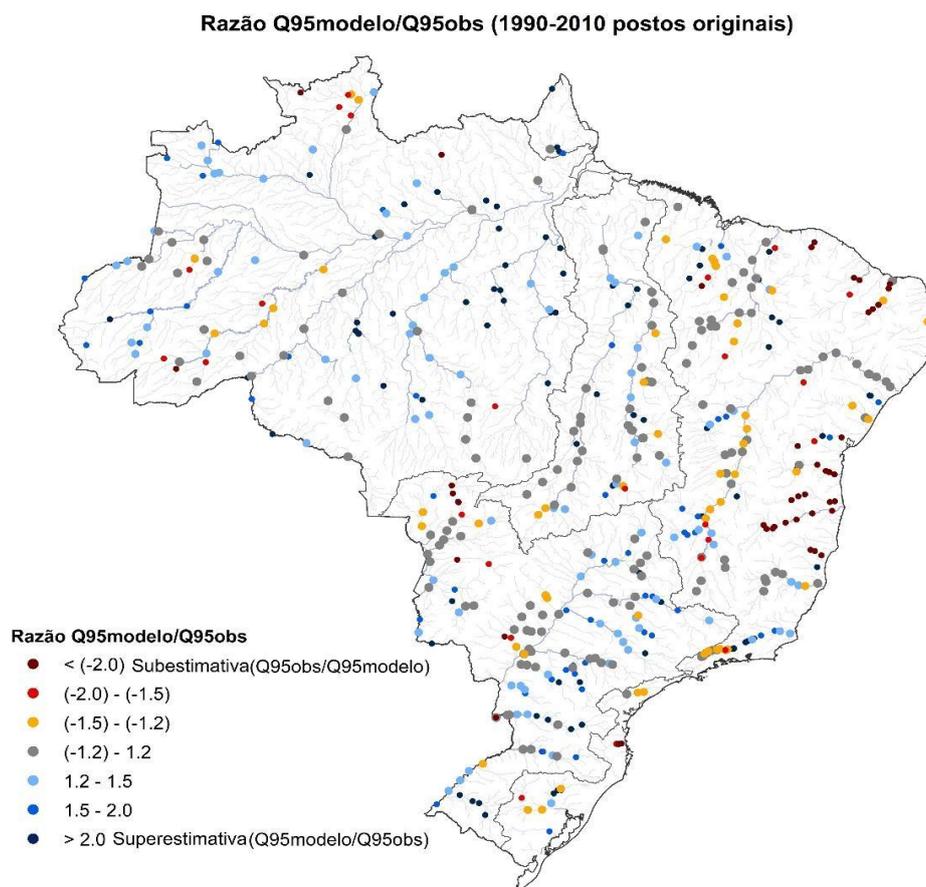


Figura 11. Distribuição espacial da Razão Q95modelo/Q95obs para o período de 1990-2010 (postos originais)

Na Figura 12 podemos observar que 50% dos pontos analisados possuem Razão Q95modelo/Q95obs de até 1.5 e 90% dos pontos de até 4.

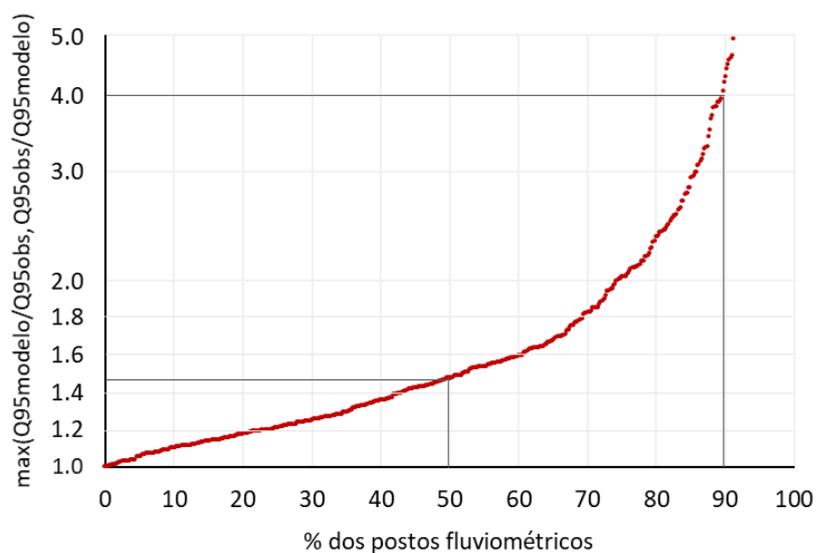


Figura 12. Relação do módulo da Razão Q95modelo/Q95obs com a quantidade de pontos analisados para o período de 1990-2010 (postos originais)

A Tabela 3 apresenta um resumo da Razão $Q_{95}modelo/Q_{95}obs$ de todos esses postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores da Razão $Q_{95}modelo/Q_{95}obs$ são maiores para áreas de drenagens menores.

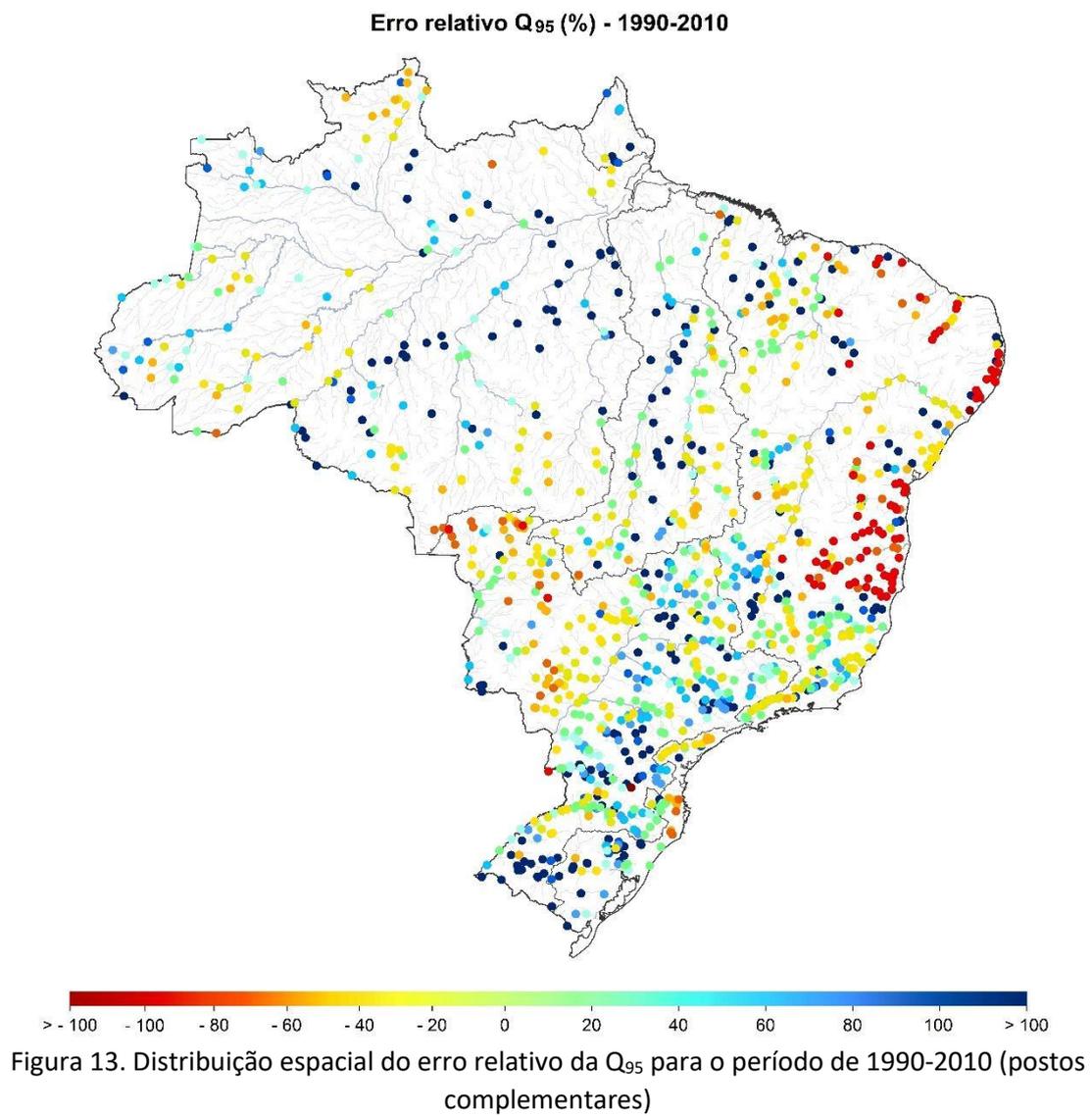
Tabela 3. Análise da Razão $Q_{95}modelo/Q_{95}obs$ em relação a área de drenagem para o período de 1990-2010 (postos originais)

ÁREA DE DRENAGEM (KM ²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	223	10.88	1.54	16.82
25.000 - 50.000	126	2.42	1.44	1.65
50.000 - 100.000	67	2.40	1.40	1.70
100.000 - 500.000	108	1.54	1.20	0.58
*Erros relativos na Q_{95} em cada intervalo de área de drenagem				

3.3.5 Análise do erro de estimativa das vazões de referência MGB-AS 1990-2010, postos complementares

A

Figura 13 apresenta mapa do erro relativo da Q_{95} e o mapa do coeficiente de Nash-Sutcliffe referente à simulação do período 1990-2010, ilustrando agora os postos fluviométricos selecionados conforme o critério de seleção que considera locais com dados complementares não utilizados originalmente no modelo MGB América do Sul, para que a análise fosse realizada em todos os pontos possíveis, totalizando 1243 postos.



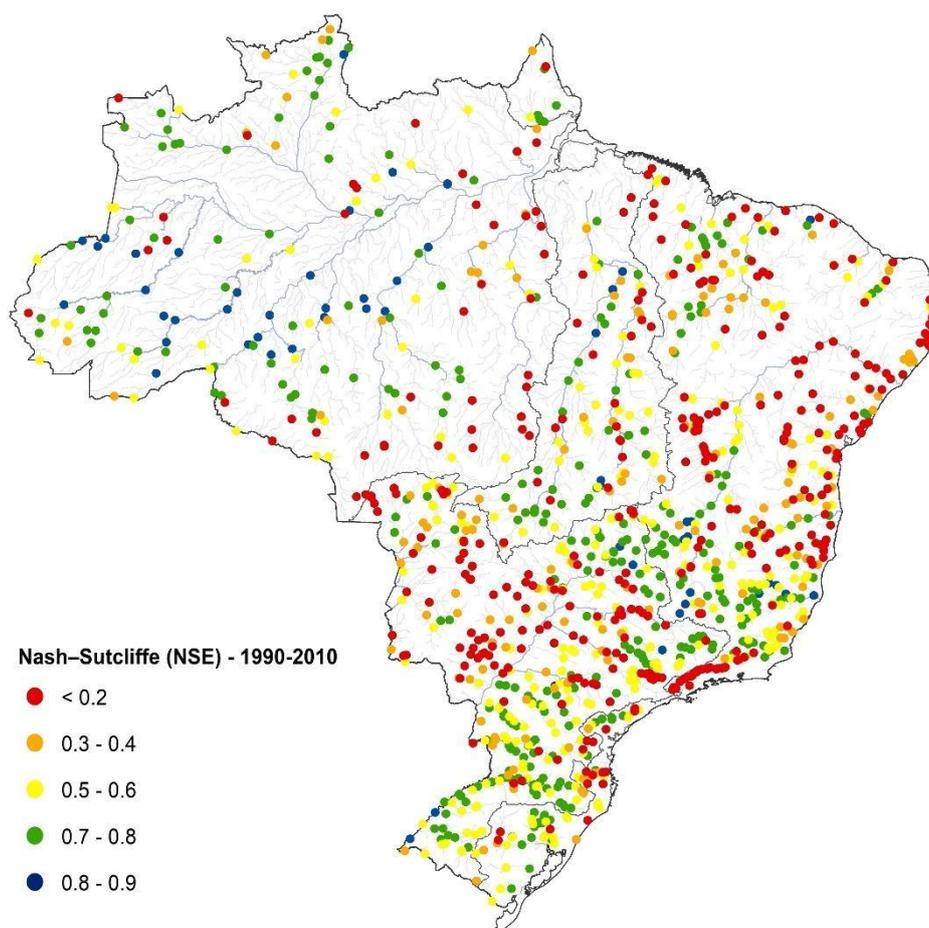


Figura 14. Coeficiente Nash-Sutcliffe (NSE) para o período de 1990-2010 (postos complementares)

A Tabela 4 apresenta um resumo dos erros da Q_{95} de todos os postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores dos erros são maiores para áreas de drenagens menores e os postos se acumulam em maior proporção na faixa de área de drenagem de 50.000 a 100.000 km^2 , o desvio padrão da mesma magnitude da média, além de uma assimetria na distribuição dos erros. A mediana (média) ficou em 18% (56%) para a faixa de 100.000 a 500.000 km^2 e em 55% (1136%) para a faixa de 1.000-25.000 km^2 . Em geral, a inclusão dos postos adicionais aumentou a média dos erros, o que era esperado.

Tabela 4. Análise do erro relativo em relação a área de drenagem para a Q_{95} para o período de 1990-2010 (postos complementares)

ÁREA DE DRENAGEM (KM ²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	273	1136.31	55.37	2007.92
25.000 - 50.000	220	479.33	54.72	768.52
50.000 - 100.000	637	252.89	42.19	366.71
100.000 - 500.000	113	55.72	17.91	62.66

*Erros relativos na Q_{95} em cada intervalo de área de drenagem

A Figura 16 e a Figura 15 apresentam a relação do valor absoluto do erro relativo da Q_{95} , e do coeficiente de Nash-Sutcliffe, em relação a quantidade de postos analisados. Em 50% dos pontos analisados o erro relativo é menor que 42% e em 80% dos pontos o erro é menor que 100%. Em 50% dos pontos analisados o NSE é superior a 0.47.

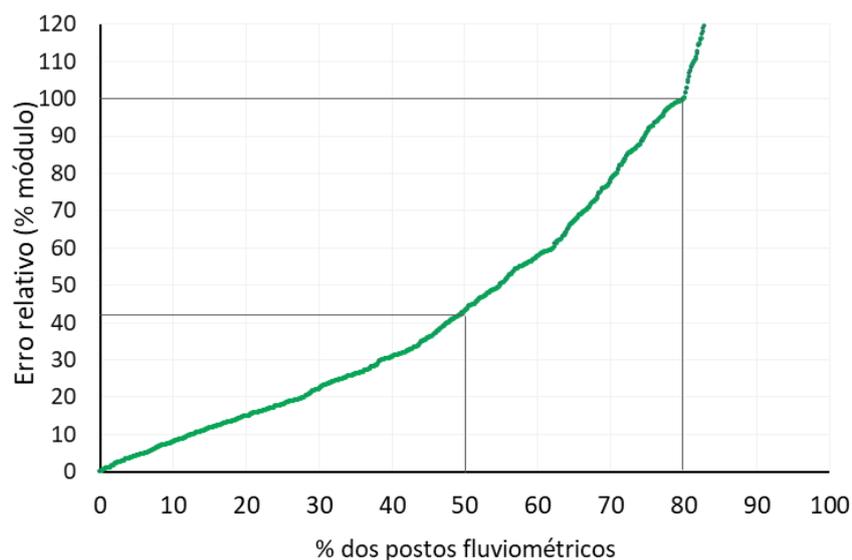


Figura 15. Relação do módulo do erro relativo da Q_{95} com a quantidade de pontos analisados para o período de 1990-2010 (postos complementares)

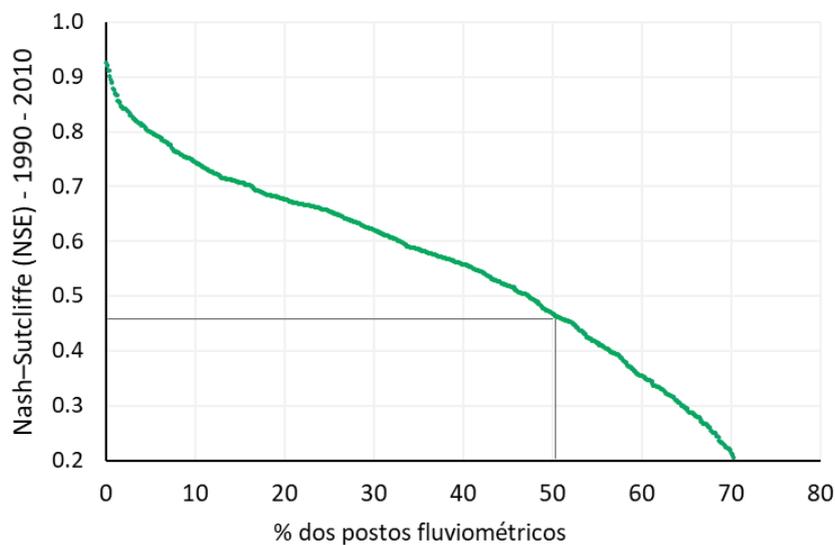


Figura 16. Relação coeficiente Nash-Sutcliffe (NSE) com a quantidade de pontos analisados para o período de 1990-2010 (postos complementares)

3.3.6 Análise da Razão $Q_{95\text{modelo}}/Q_{95\text{obs}}$ para o período de 1990-2010, postos complementares

Na Figura 17 é apresentado o mapa da Razão $Q_{95\text{modelo}}/Q_{95\text{obs}}$ para simulação do período 1990-2010 (postos complementares), ilustrando os postos fluviométricos selecionados conforme o critério de seleção que considera locais com dados complementares, totalizando 1243 postos.

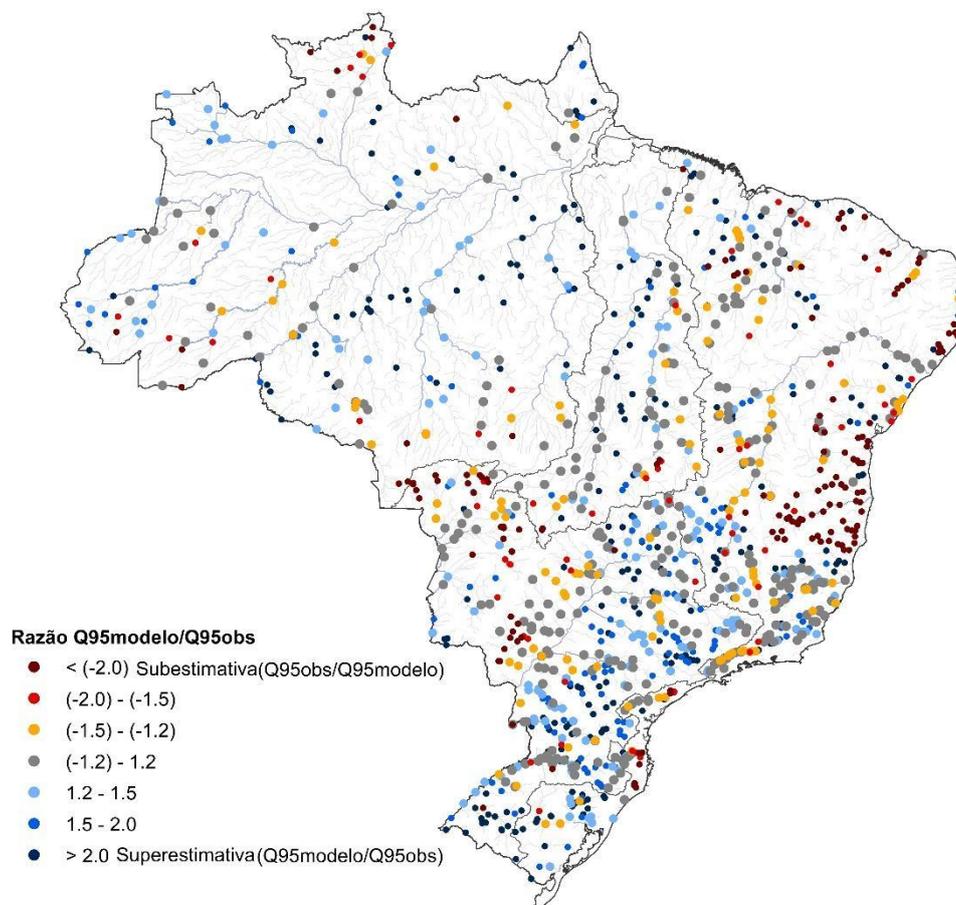
Razão Q95modelo/Q95obs (1990-2010 postos complementares)

Figura 17. Distribuição espacial da Razão Q95modelo/Q95obs para o período de 1990-2010 (postos complementares)

Na Figura 18 podemos observar que 50% dos pontos analisados possuem Razão Q₉₅modelo/Q₉₅obs de até 1.6 e 85% dos pontos de até 4.

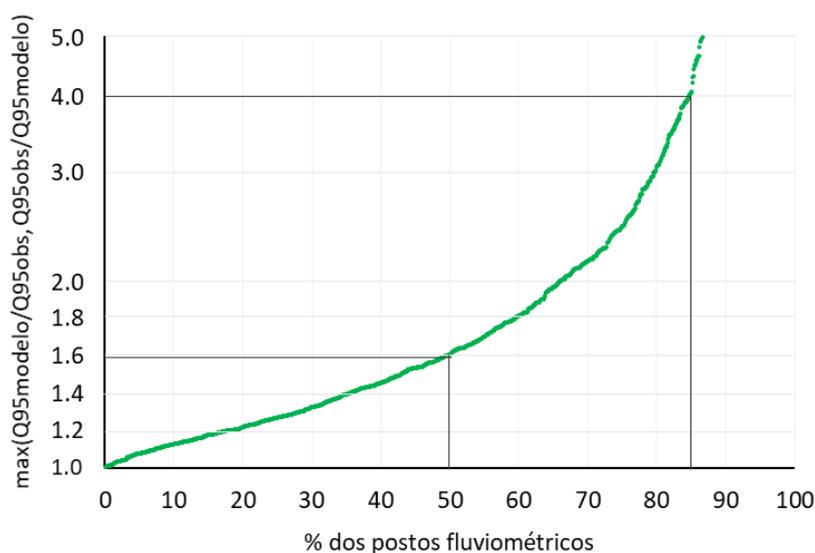


Figura 18 - Relação do módulo da Razão Q95modelo/Q95obs com a quantidade de pontos analisados para o período de 1990-2010 (postos complementares)

A Tabela 5 apresenta um resumo da Razão Q95modelo/Q95obs de todos esses postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores da Razão Q95modelo/Q95obs são maiores para áreas de drenagens menores.

Tabela 5. Análise da Razão Q95modelo/Q95obs em relação a área de drenagem para o período de 1990-2010 (postos complementares)

ÁREA DE DRENAGEM (KM ²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	273	17.81	1.72	28.49
25.000 - 50.000	220	8.27	1.69	11.14
50.000 - 100.000	637	6.75	1.51	8.90
100.000 - 500.000	113	1.58	1.20	0.64

*Erros relativos na Q95 em cada intervalo de área de drenagem

3.3.7 Análise do erro de estimativa das vazões de referência, MGB-AS 1980-2015

A Figura 19 apresenta o mapa do erro relativo da Q95 e na Figura 20 é apresentado o mapa do coeficiente de Nash-Sutcliffe para simulação do período 1980-2015, ilustrando os postos fluviométricos selecionados conforme o critério de seleção que considera locais com dados complementares, totalizando 1371 postos.

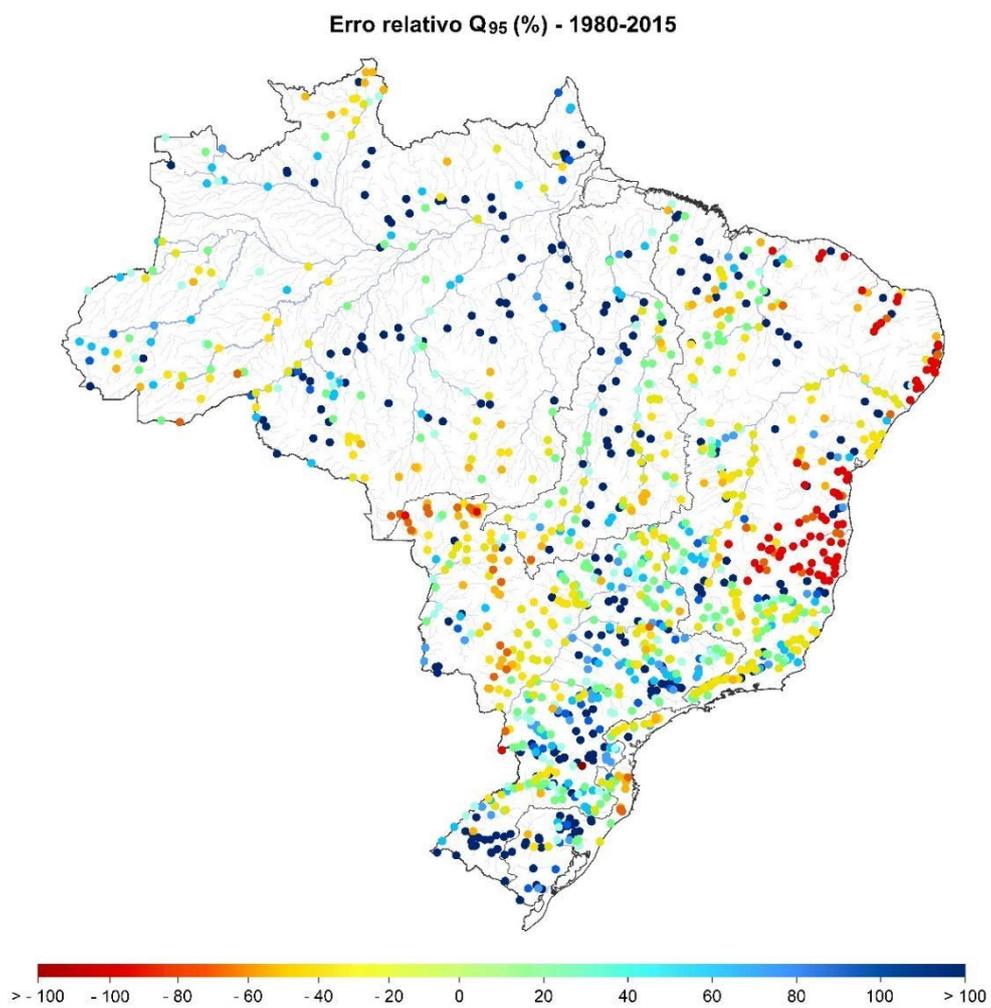


Figura 19. Distribuição espacial do erro relativo da Q_{95} para o período de 1980-2015

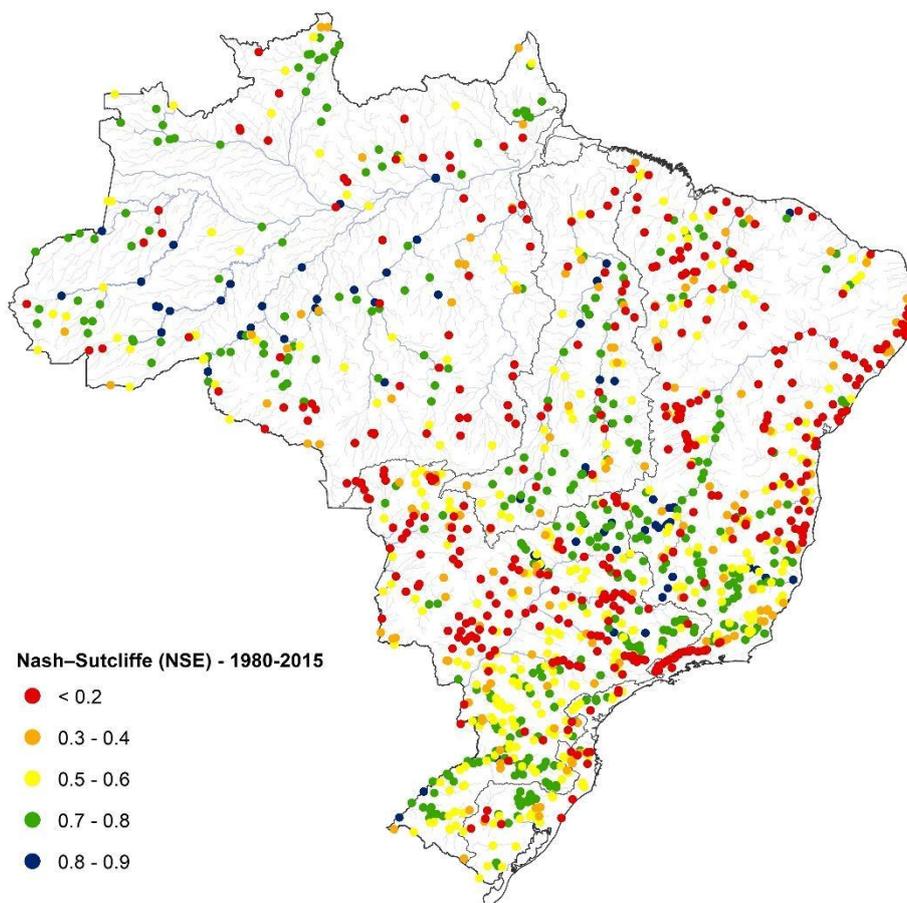


Figura 20. Coeficiente Nash-Sutcliffe (NSE) para o período de 1980-2015

A Tabela 6 apresenta um resumo dos erros da Q_{95} de todos os postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores dos erros são maiores para áreas de drenagens menores e os postos se acumulam em maior proporção na faixa de área de drenagem de 50.000 a 100.000 km^2 , o desvio padrão da mesma magnitude da média, além de uma assimetria na distribuição dos erros. A mediana (média) ficou em 18% (57%) para a faixa de 100.000 a 500.000 km^2 e em 57% (194%) para a faixa de 1.000-25.000 km^2 . Para a faixa de 25.000 a 50.000 km^2 , o erro médio subiu para 2200%. Em geral, a inclusão dos postos adicionais aumentou a média dos erros e a extensão do período considerado também contribuem para a queda de desempenho aparente nos resultados.

Tabela 6. Análise dos erros em relação a área de drenagem para a Q_{95} para o período de 1980-2015

ÁREA DE DRENAGEM (KM ²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	289	194.23	57.50	231.20
25.000 - 50.000	232	2227.87	53.13	4171.73
50.000 - 100.000	725	276.31	41.55	406.54
100.000 - 500.000	125	53.91	17.64	58.71

*Erros relativos na Q_{95} em cada intervalo de área de drenagem

A Figura 21 e a Figura 22 apresentam a relação do valor absoluto do erro relativo da Q_{95} , e do coeficiente de Nash-Sutcliffe, em relação a quantidade de postos analisados. 50% dos pontos analisados possuem erro relativo de até 42% e 78% dos pontos possuem erro de até 100%. 50% dos pontos analisados possuem NSE acima de 0.50.

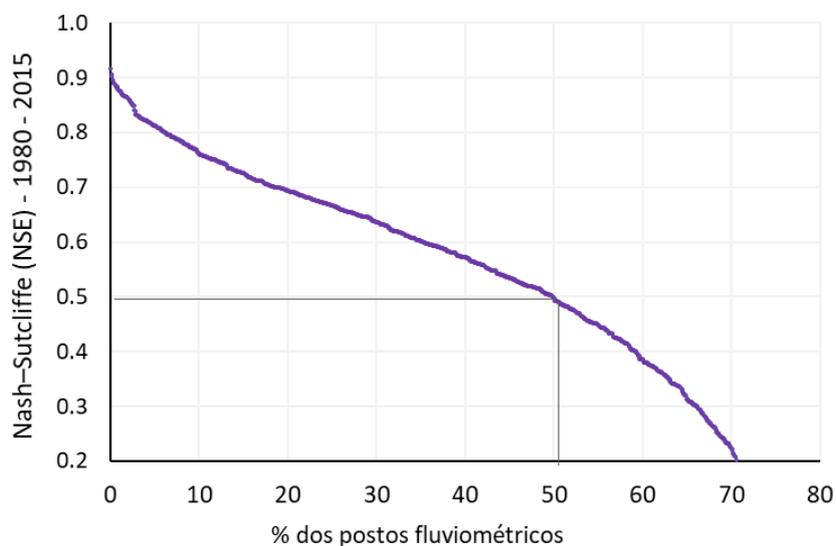


Figura 21. Relação coeficiente Nash-Sutcliffe (NSE) com a quantidade de pontos analisados para o período de 1980-2015

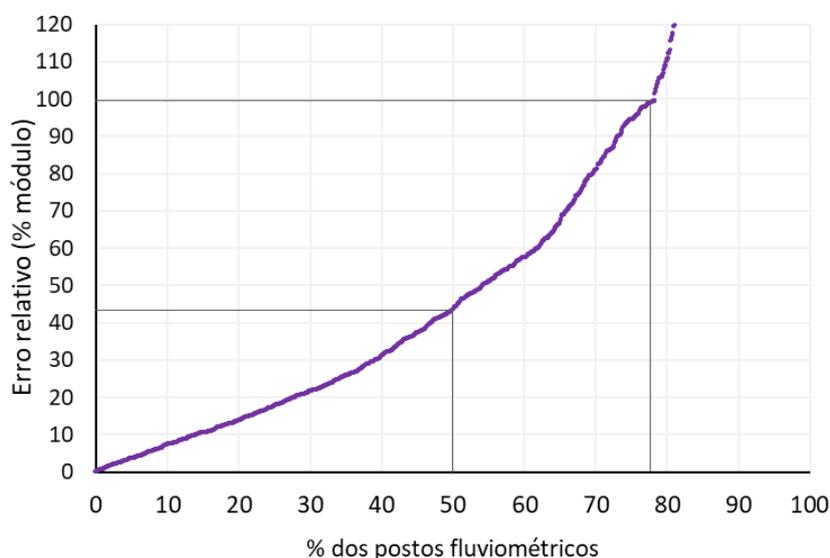


Figura 22. Relação do módulo do erro relativo da Q_{95} com a quantidade de pontos analisados para o período de 1980-2015

3.3.8 Análise da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para o período de 1980-2015

Na Figura 23 apresenta-se o mapa da Razão $Q_{95}^{\text{modelo}}/Q_{95}^{\text{obs}}$ para simulação do período 1980-2015, ilustrando os postos fluviométricos selecionados conforme o critério de seleção, totalizando 1371 postos.

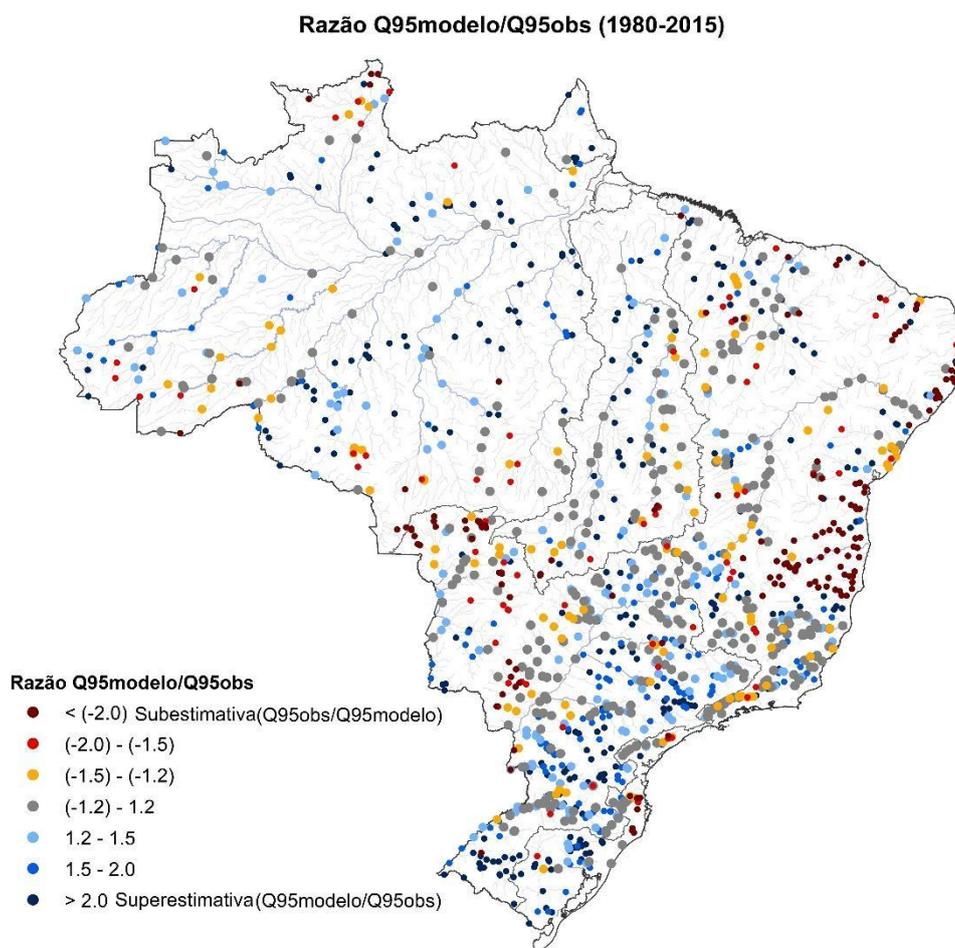


Figura 23. Distribuição espacial da Razão Q95modelo/Q95obs para o período de 1980-2015

Na Figura 24 podemos observar que 50% dos pontos analisados possuem Razão Q95modelo/Q95obs de até 1.6 e 85% dos pontos de até 4.

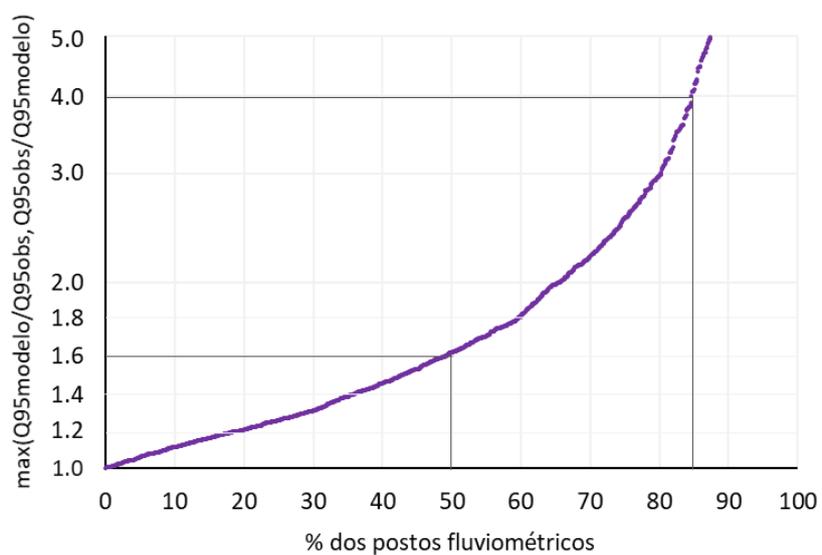


Figura 24. Relação do módulo da Razão $Q_{95}modelo/Q_{95}obs$ com a quantidade de pontos analisados para o período de 1980-2015

A Tabela 7 apresenta um resumo da Razão $Q_{95}modelo/Q_{95}obs$ de todos esses postos fluviométricos, em relação a faixas de área de drenagem. É possível observar que os valores da Razão $Q_{95}modelo/Q_{95}obs$ são maiores para áreas de drenagens menores.

Tabela 7 - Análise da Razão $Q_{95}modelo/Q_{95}obs$ em relação a área de drenagem para o período de 1980-2015

ÁREA DE DRENAGEM (KM²)	QUANTIDADE DE POSTOS	ERRO MÉDIO (MÓDULO)	MEDIANA DO ERRO (MÓDULO)	DESVIO PADRÃO
1.000 - 25.000	289	7.76	1.77	9.88
25.000 - 50.000	232	25.47	1.68	43.78
50.000 - 100.000	725	4.61	1.50	5.13
100.000 - 500.000	125	1.56	1.20	0.60
*Erros relativos na Q_{95} em cada intervalo de área de drenagem				

4 REANÁLISE HIDROLÓGICA – SIMULAÇÕES COM ASSIMILAÇÃO DE DADOS

Nesta etapa do subprojeto, o modelo da América do Sul (Siqueira et al., 2018) foi utilizado juntamente com uma técnica de assimilação de dados para correção das variáveis de estado do modelo ao longo do período de simulação, diante a disponibilidade de dados observados.

4.1 DEFINIÇÃO DE ESTAÇÕES FLUVIOMÉTRICAS PARA VALIDAÇÃO DA SIMULAÇÃO COM ASSIMILAÇÃO DE DADOS

Para adequada verificação da eficiência de uma aplicação de assimilação de dados, se faz necessária a seleção das estações fluviométricas cujos dados serão assimilados e estações que serão utilizadas para validar a aplicação, a partir da análise da alteração do desempenho do modelo.

Em um cenário ideal seria adotado o procedimento denominado “*leave one gauge out*”, isto é, poderiam ser feitas múltiplas simulações, assimilando dados de todas as N estações exceto uma única estação por vez e verificando alterações no desempenho do modelo na estimativa das vazões desta única estação não utilizada na assimilação, emulando um local sem dados. Contudo, isto exigiria a aplicação do método N vezes (N é o número de postos fluviométricos selecionados), o que teria um altíssimo custo computacional. Assim, optou-se por reduzir o número de simulações a partir do agrupamento de estações fluviométricas que seriam assimiladas a cada vez, adotando um procedimento que pode ser denominado “*leave 1/10 gauges out*”, e que precisa ser repetido apenas 10 vezes.

Desta forma, as estações fluviométricas selecionadas para uso neste subprojeto, apresentadas no relatório IPH-ANA-HGE-SR-R3, foram divididas em 10 grupos compostos por 10% das estações. Assim, cada experimento realizado é simulado 10 vezes, onde em cada simulação são assimilados dados de 10% das estações.

Para garantir uma distribuição espacial mais uniforme entre os grupos, ou seja, garantir que um mesmo grupo não possua estações localizadas em uma única região do país, a divisão das estações foi realizada a partir de 11 regiões hidrográficas. Com isso, cada região possui 10% de suas estações em cada um dos grupos. Além disso, na separação dos grupos foi usado o critério de que um posto nunca estaria no mesmo grupo de que o posto mais próximo a ele. Desta forma, temos a garantia de que quando um posto for utilizado para validação, o posto mais próximo a ele está sendo assimilado.

A Figura 25 apresenta as estações fluviométricas que compõem cada um dos grupos (em vermelho), isto é, que serão utilizadas para validação em cada grupo e as demais estações (em azul) que serão assimiladas. Observa-se que há uma homogeneidade na distribuição espacial das estações entre os grupos.

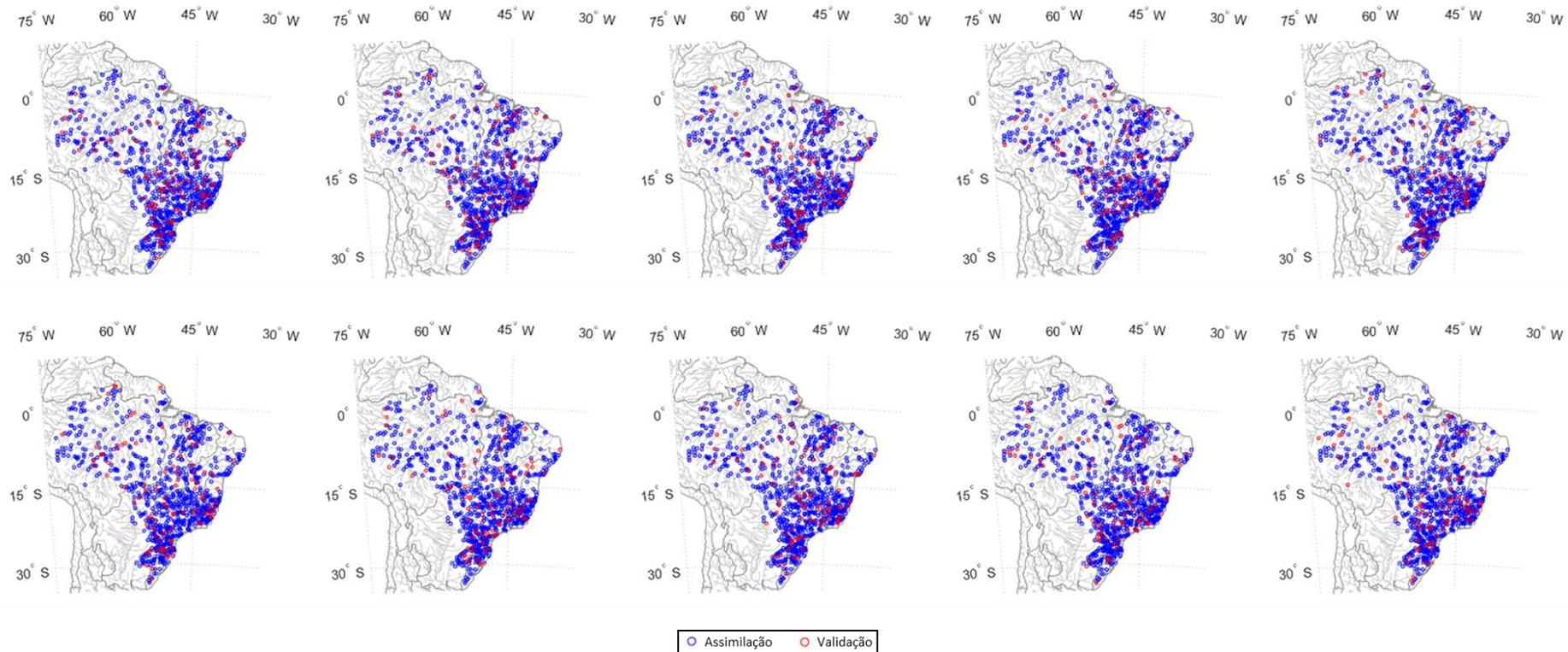


Figura 25. Divisão das estações nos 10 grupos de validação. Em vermelho, as estações fluviométricas que compõem os grupos e serão usadas para validação. Em azul, as demais estações, que serão assimiladas.

4.2 CONSOLIDAÇÃO DO MGB-AS-ENKF: DEFINIÇÃO DE ERROS E PARÂMETROS DE ASSIMILAÇÃO

Para aplicação do EnKF/LEnKF é necessário definir uma estimativa dos erros do modelo e dos dados observados, assim como definir alguns parâmetros de assimilação. Para as aplicações neste subprojeto foram planejados uma série de experimentos para adequar a aplicação do modelo com assimilação de dados à escala continental. Os parâmetros e configurações da assimilação de dados foram definidos com base na aplicação do LEnKF feita por Wongchuig et al. (2020).

Nas simulações deste projeto, serão assimilados os dados de vazões observadas nas estações fluviométricas da base de dados selecionada. O erro estipulado para a vazão observada foi adotado como 10%.

Para assimilação são perturbadas a chuva (P), com erro de 25% e decorrelação temporal de 20 dias, e as variáveis de estado correspondentes ao armazenamento de água no solo (W) e aos volumes armazenados nos reservatórios lineares superficiais (V_{sup}), intermediários (V_{int}) e subterrâneos (V_{bas}), com erros de 10 e 5%, respectivamente. Considerou-se os erros com dependência espacial, com fator de decorrelação espacial igual a 1.5 graus e 50 membros para o conjunto do EnKF.

Com o objetivo de evitar que a assimilação de dados degrade os resultados, e considerando a escala continental de aplicação e a quantidade de estações fluviométricas utilizadas para assimilação, um ponderador de áreas foi adicionado como fator multiplicativo à matriz de correlações, calculadas pela equação 8. Este ponderador permite que um dado observado em um determinado posto fluviométrico tenha pesos diferentes em minibacias vizinhas, mas com área de drenagem diferentes. Desta forma, por exemplo, a assimilação de uma vazão observada em um posto no rio principal apresenta maiores correlações com uma minibacia vizinha também localizada no rio principal do que com uma minibacia vizinha localizada em um afluente com menor área de drenagem. O fator multiplicativo utilizado para esta ponderação é apresentado na equação 13.

$$f_{area} = \min\left(\frac{A_i}{A_j}, \frac{A_j}{A_i}\right)^{0.5} \quad (10)$$

onde A é a área de drenagem de uma minibacia, i é o índice correspondente à minibacia onde se têm a observação a ser assimilada e j é um índice para as minibacias que compõem o modelo MGB-SA.

4.3 RESULTADOS

Os resultados a seguir referem-se a simulações para o período de 35 anos (01/01/1979 – 31/12/2014). As métricas apresentadas a seguir são calculadas sempre em estações que foram deixadas de fora do processo de assimilação de dados e, portanto, representam pontos de validação.

A Figura 26 apresenta uma comparação das métricas KGE, NS, NSLog, BIAS, RMSE e r para versões do modelo MGB-AS com e sem assimilação de dados (Open Loop). Observa-se uma melhoria média do modelo para todas as métricas apresentadas. A valor mediano do NS apresentou um aumento de 0.65 para 0.72 nas estações analisadas, com a utilização da

assimilação de dados. O RMSE mediano do modelo foi reduzido de 59% para 49%. Além disso, é possível observar um estreitamento entre os limites máximos e mínimos do boxplot correspondente ao BIAS com assimilação de dados, demonstrando uma redução no viés do modelo.

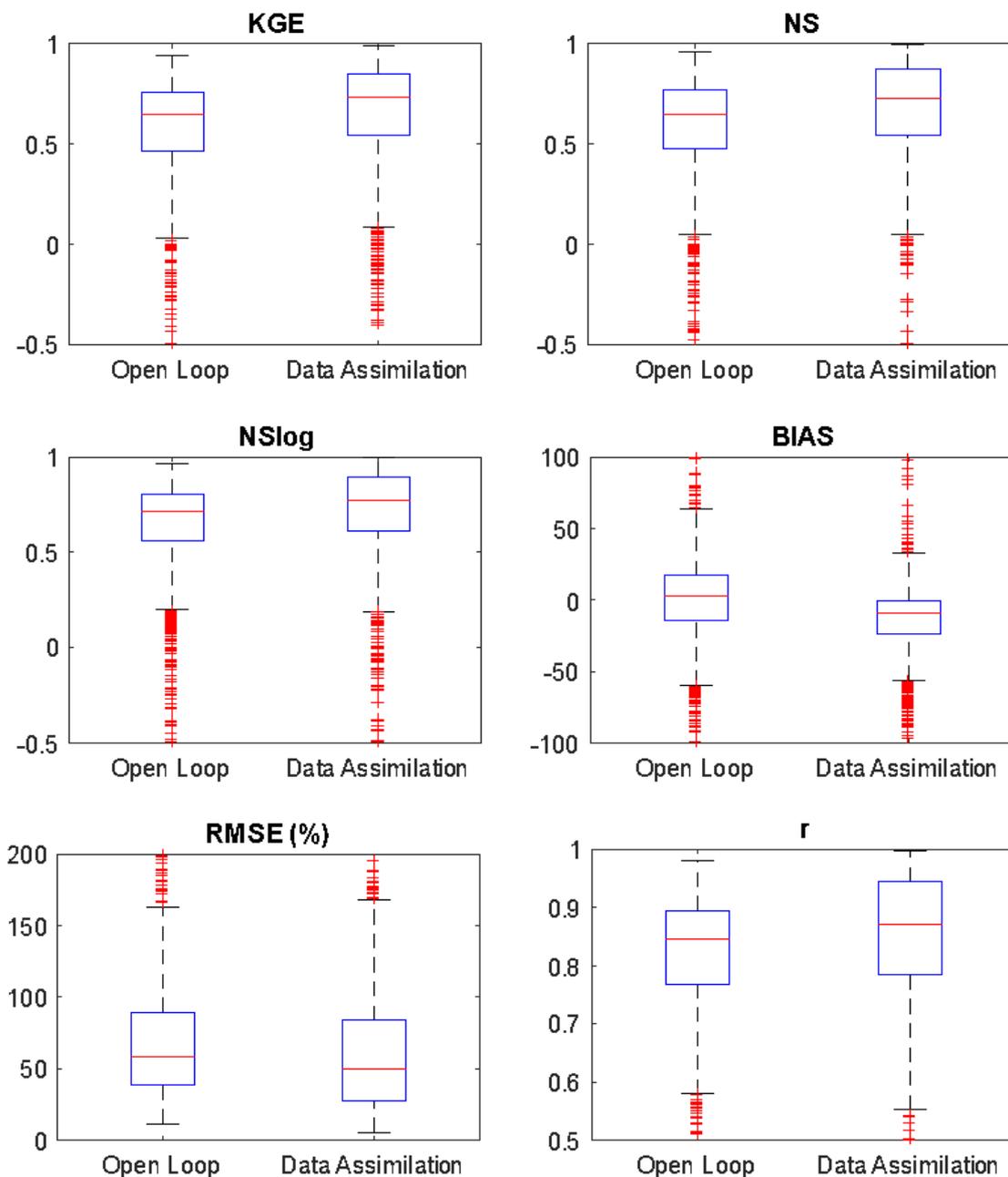


Figura 26 – Boxplots com métricas do modelo MGB-SA com e sem assimilação de dados (Open Loop).

Na Figura 27 são apresentados os boxplots das métricas KGE, NS, NSLog e RMSE para diferentes intervalos de área de drenagem (<5000km², entre 5000 e 10000km² e >10000km²). Observa-se que, para todas as escalas de bacia e todas as métricas analisadas, o uso da assimilação de dados resultou em uma melhoria no desempenho do modelo. Contudo, observa-se maiores melhorias com o aumento da área de drenagem das bacias. Para bacias maiores do que 10000km², a mediana do RMSE teve uma redução de 42% para 29% com uso da assimilação de dados.

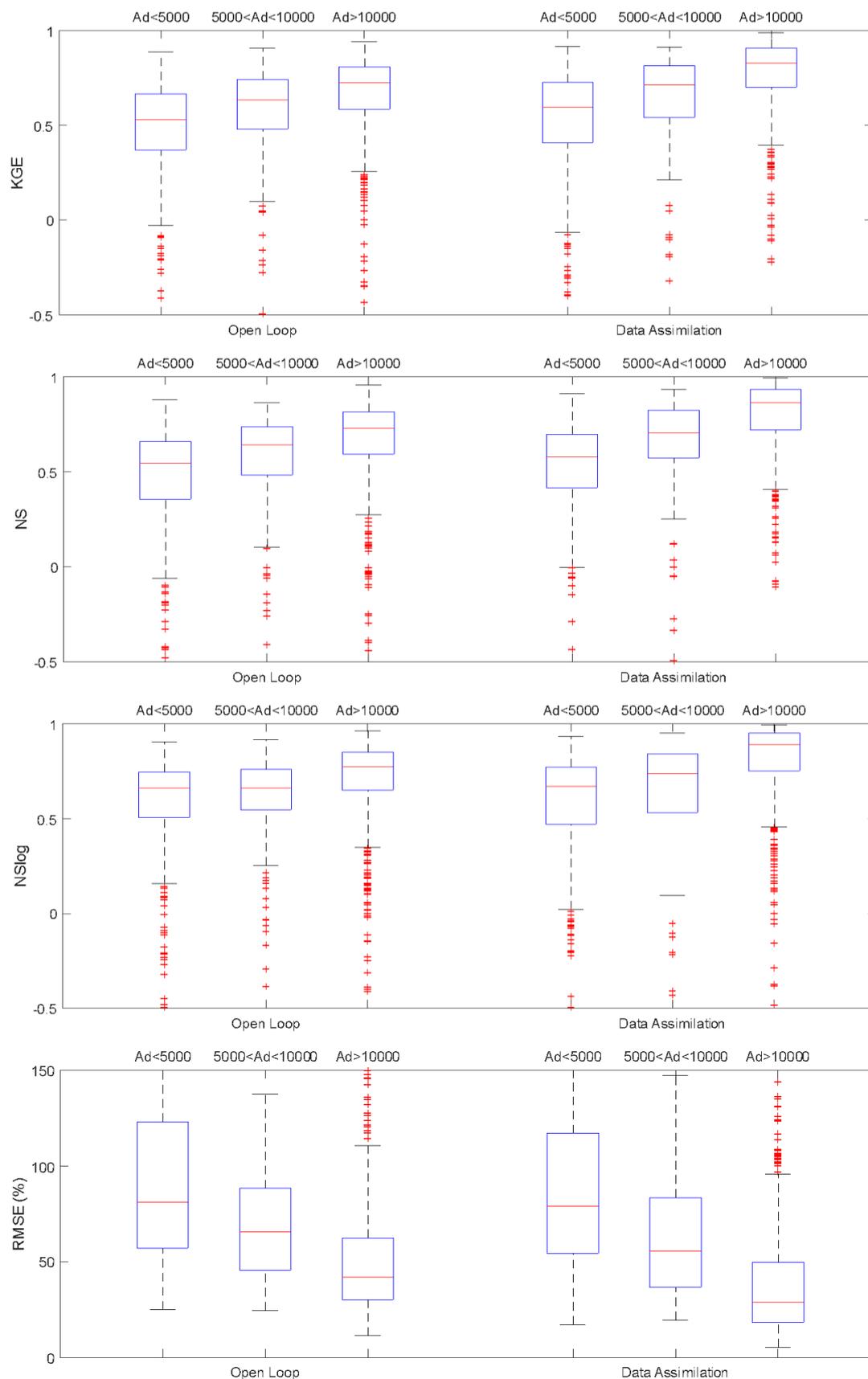


Figura 27 - Boxplots com métricas do modelo MGB-SA com e sem assimilação de dados (Open Loop) para diferentes intervalos de área de drenagem ($<5000\text{km}^2$; entre 5000 e 10000km^2 e $>10000\text{km}^2$).

As Figura 28 apresenta mapas com os valores de NS calculados para cada uma das estações de validação. O mapa da esquerda representa valores de NS sem assimilação de dados e o mapa da direita refere-se a coeficientes de NS para simulações com assimilação de dados. De forma geral, observa-se um aumento dos valores de NS em todas as regiões do país. Contudo, ressalta-se a melhoria em regiões da bacia do Rio São Francisco, Tocantins e em afluentes do Rio Paraná.

Diferenças nos valores de NS e no RMSE são apresentados na Figura 29. Para a maioria das estações, houve uma melhoria de cerca de 0.20 nos valores de NS e uma redução absoluta de 15% no RMSE, conforme já apresentado no boxplot da Figura 26.

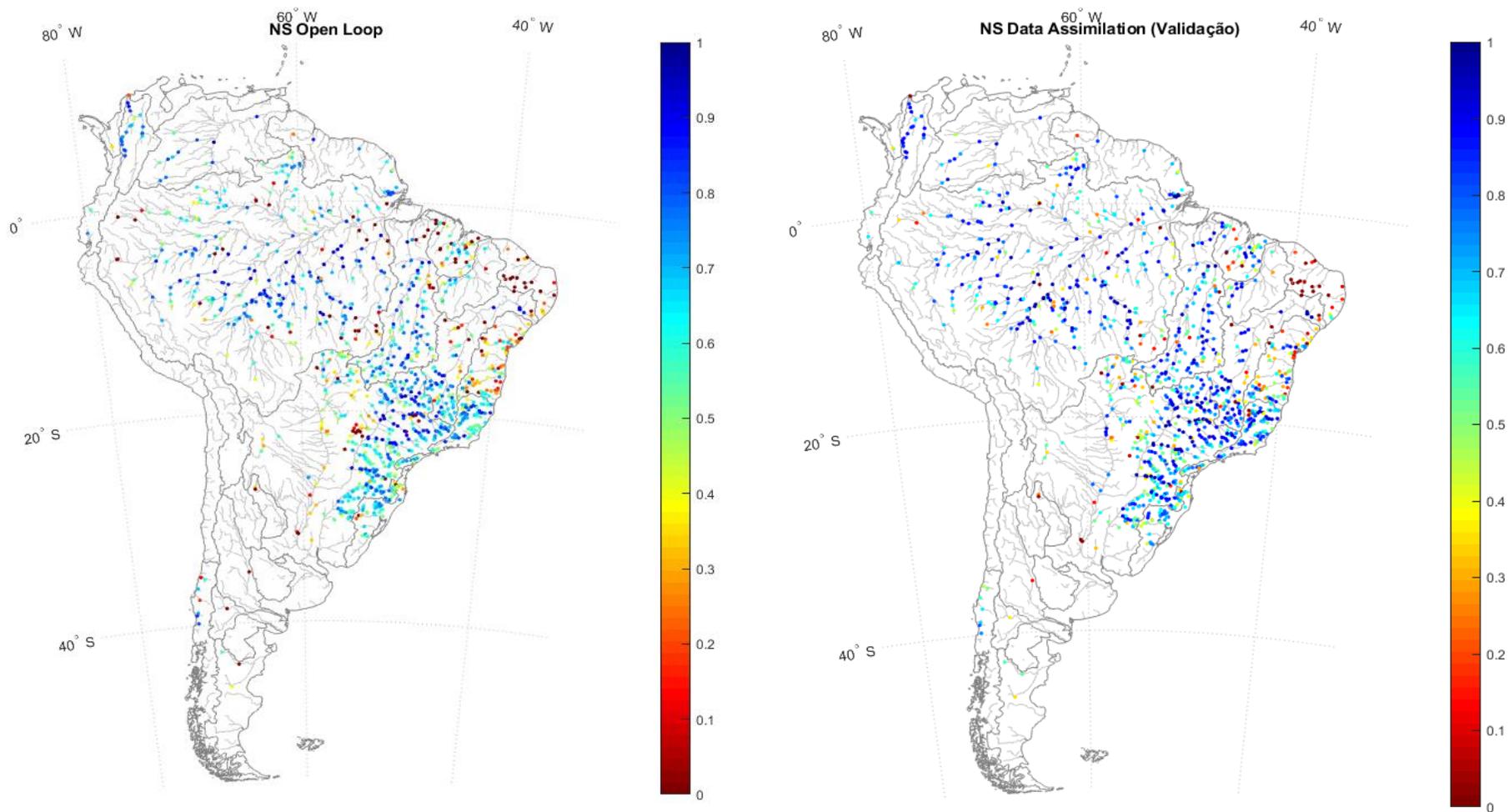


Figura 28 – Valores de NS calculados em estações de validação para versões do modelo MGB-SA sem e com assimilação de dados.

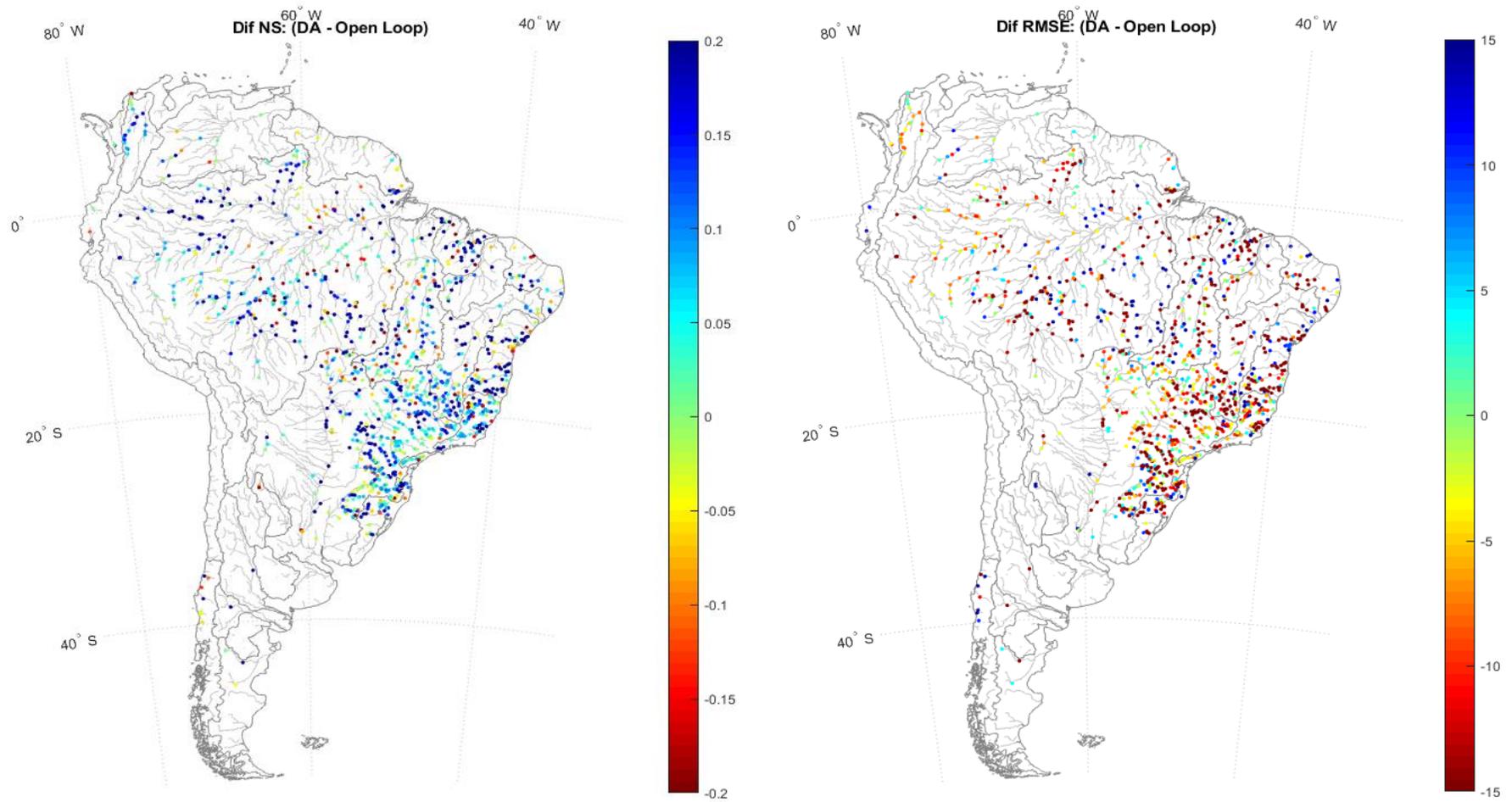


Figura 29 -Diferenças nas métricas NS e RMSE após uso da técnica de assimilação de dados.

Para o mesmo período de simulação, foram calculadas as vazões de referência Q95 para cada estação fluviométrica a partir dos dados observados e dos resultados das simulações com e sem assimilação de dados. Em seguida, foram calculados os erros nas estimativas da Q95, usando como referência a Q95 calculada com dados observados. O erro relativo foi calculado a partir da equação 11 e as razões RQ95 e RQM foram calculadas conforme equações X e XX.

$$ErroQ95(\%) = \frac{Q95_{Simulada} - Q95_{Observada}}{Q95_{Observada}} * 100 \quad (11)$$

A Figura 30 apresenta os boxplots dos erros nas estimativas das vazões Q95 para simulações sem assimilação de dados (open loop) e com assimilação. Observa-se uma significativa melhoria nas estimativas, com mediana dos erros reduzindo de 26% para -2% a partir do uso da assimilação de dados nas simulações. Observa-se também uma diminuição da dispersão dos erros com redução da diferença entre os limites superior e inferior do boxplot. Para as simulações com assimilação de dados, 50% dos postos avaliados tiveram erro na faixa -30 a 28%, enquanto que na simulação original a faixa correspondia de -21 a 104%.

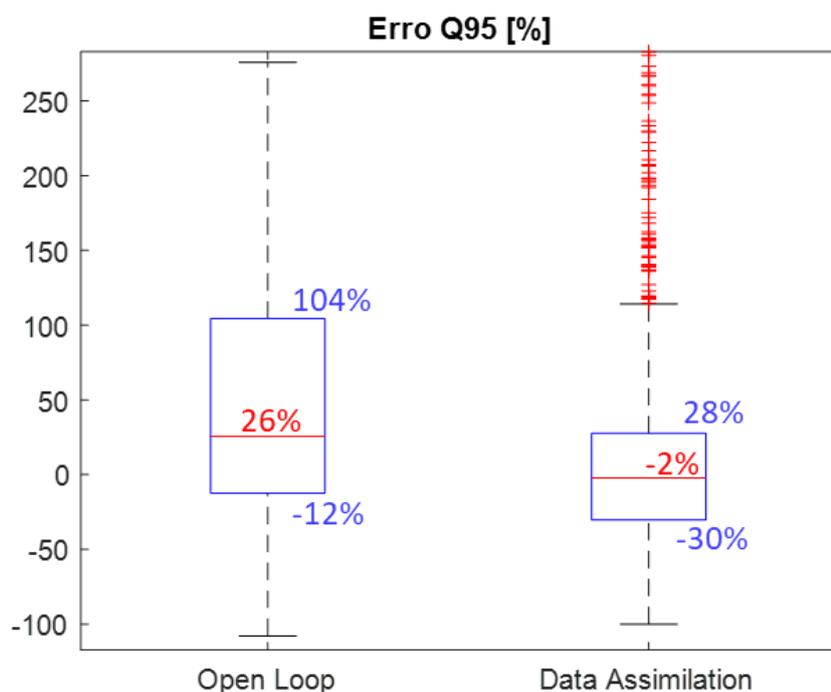


Figura 30 – Boxplot de erros na estimativa das vazões Q95 nas estações a partir de simulações com e sem assimilação de dados.

As Figura 31 e Figura 32 apresentam os boxplots das razões RQ95 e RQM, respectivamente, para as simulações *Open Loop* e com AD. Estes também apresentam uma redução dos erros das estimativas feitas a partir da utilização da assimilação de dados, sendo as mudanças mais significativas nas vazões de referência Q95. A mediana da RQ95 reduziu de 1.62 para 1.35, enquanto que a RQM reduziu de 1.13 para 1.08.

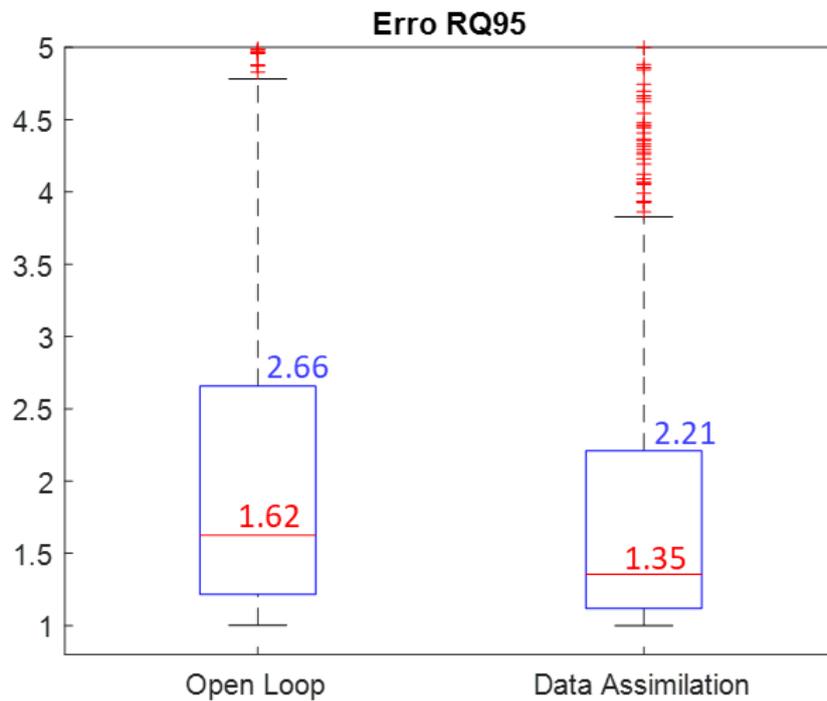


Figura 31 - Boxplot da razão RQ95 nas estações a partir de simulações com e sem assimilação de dados.

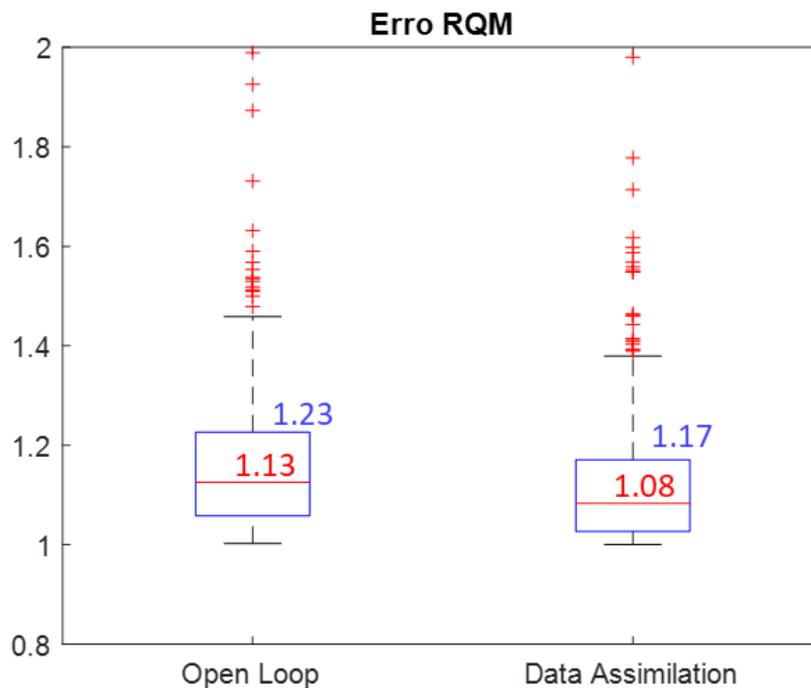


Figura 32 - Boxplot da razão RQM nas estações a partir de simulações com e sem assimilação de dados.

Durante a etapa de separação dos 10 grupos de validação, como explicado no item 4.1, foi definido que ao selecionar um posto o seu vizinho mais próximo automaticamente não poderia pertencer ao mesmo grupo. Contudo, a distância entre um posto e seu vizinho mais próximo pode variar bastante, ou ainda, há muitos casos em que o posto vizinho mais próximo de uma estação localizada em um afluente está localizado no rio principal, com uma área de drenagem possivelmente muito superior.

As Figura 33 e Figura 34 mostram a distribuição dos erros relativos na Q95 e QM, respectivamente, em função da relação entre as áreas de drenagem do posto utilizado para validação e seu vizinho mais próximo – cujas vazões foram assimiladas, segundo metodologia adotada. As curvas superiores correspondem ao membro 48 e as curvas inferiores correspondem ao membro 2 do conjunto da assimilação. Estes membros correspondem à faixa que engloba as vazões de referência Q95 observadas em 75% dos postos.

Conforme esperado, observa-se um estreitamento e minimização das curvas quando a relação entre áreas se aproxima de 1, o que diretamente indica uma proximidade maior entre o posto assimilado e o posto validado. Ainda, tem-se que os maiores erros foram encontrados nos casos em que a relação entre as áreas é inferior a 0.1 (a área do posto assimilado é no mínimo 10 vezes maior que o posto em validação). Nestes casos, a assimilação ainda apresenta uma limitação.

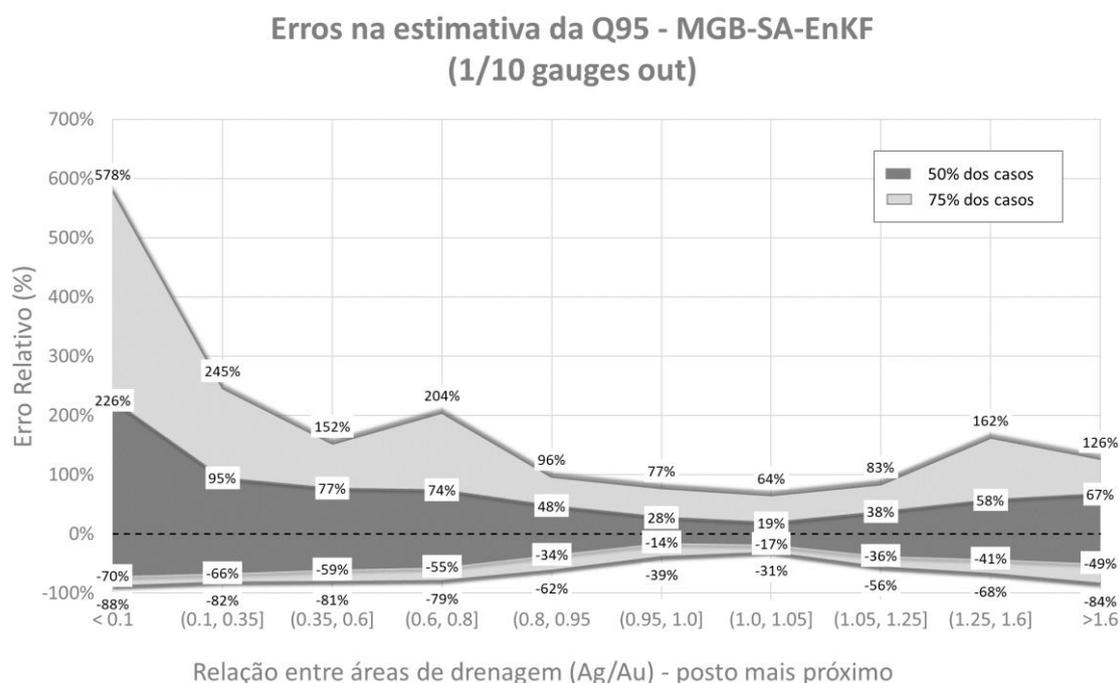


Figura 33 – Erros relativos na estimativa da Q95 em função da relação entre as áreas de drenagem doo posto em validação e do posto assimilado mais próximo.

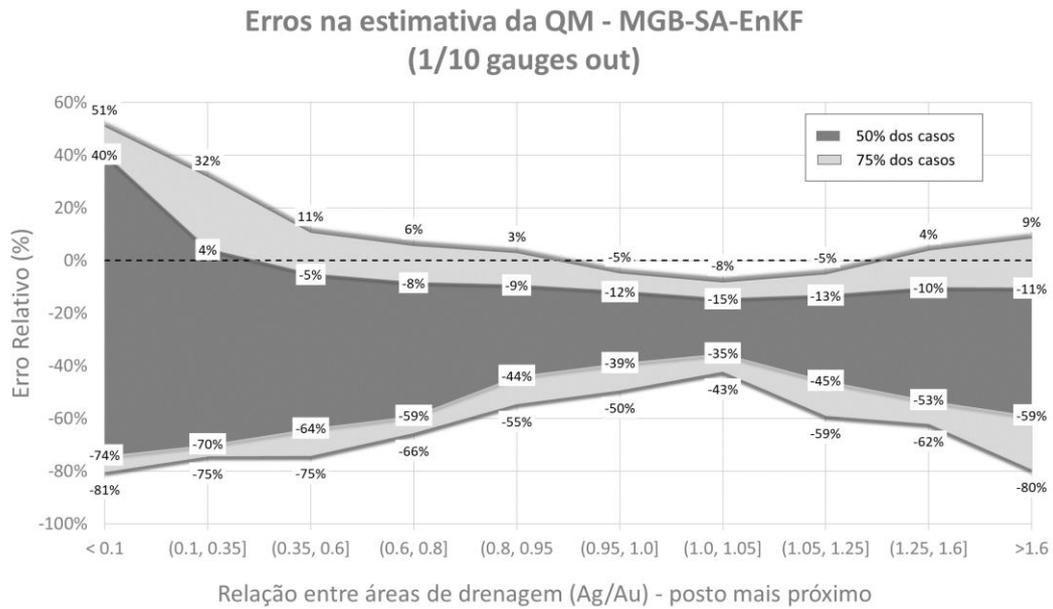


Figura 34 - Erros relativos na estimativa da QM em função da relação entre as áreas de drenagem doo posto em validação e do posto assimilado mais próximo

Ressalta-se que todas as análises acima foram realizadas para postos de validação, isto é, que ficaram de fora do processo de assimilação de dados e o estado do modelo na minibacia correspondente foi resultado das assimilações em estações vizinhas. Contudo, para aplicações futuras, após as análises de validação, é possível utilizar uma simulação com assimilação de todas as estações simultaneamente. Esta simulação, teoricamente, corresponderia à melhor estimativa possível que poderíamos fornecer atualmente com o uso da técnica de assimilação de dados no modelo MGB-SA.

Como resultado deste processo, a Figura 36 apresenta os valores de NS nas estações assimiladas e, de acordo com os boxplots apresentados na Figura 35, observa-se um aumento na mediana dos valores de NS (0.65 – Open Loop; 0.72 – Validação e; 0.86 – Assimilação em todas as estações disponíveis).

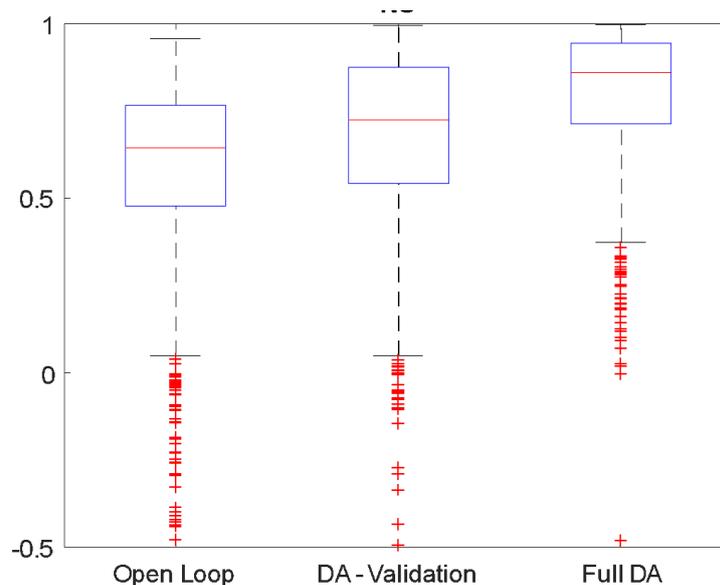


Figura 35 – Boxplots de valores de Nash-Sutcliffe para simulação sem assimilação de dados (Open Loop), com assimilação em postos de validação e com assimilação em todas as estações disponíveis.

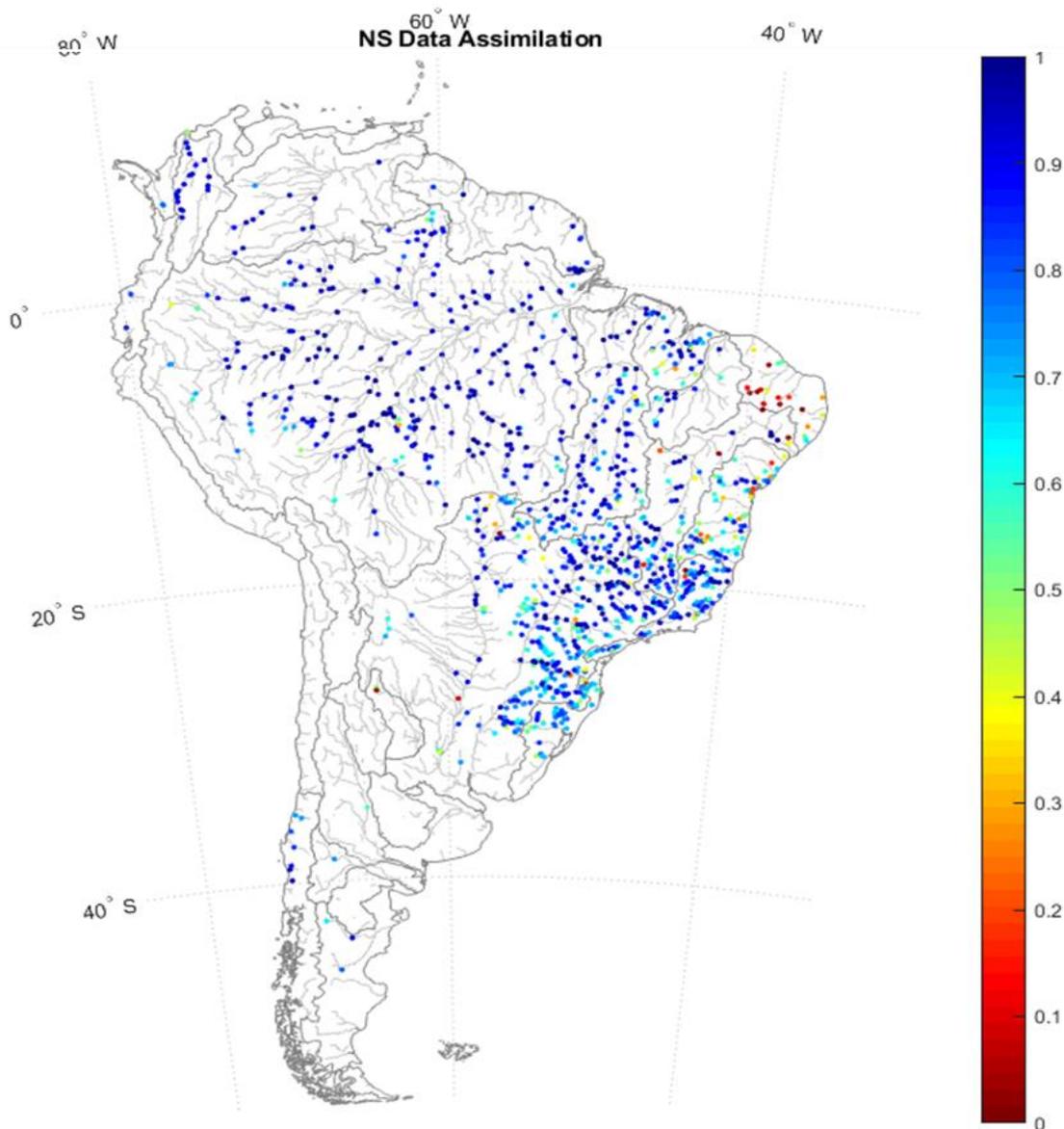


Figura 36 – Mapas com valores do coeficiente de Nash-Sutcliffe para simulação sem assimilação de dados (esquerda) e para simulação com assimilação em todos os pontos calculados (direita).

A Figura 37 apresenta os mapas de erros nas estimativas das vazões de referência Q95 com as versões *Open Loop* e assimilando os dados de todas as estações selecionadas.

Por fim, as Figura 38 e Figura 39 apresentam hidrogramas de vazões calculadas em diferentes regiões do país, a partir de simulações com a versão original do modelo MGB-SA (em vermelho) e a partir de simulações com assimilação de dados (em preto). Em azul, são apresentadas as séries observadas de vazões. Estes hidrogramas são referentes a simulações de validação, ou seja, os postos apresentados não foram assimilados durante a simulação em que são apresentados os resultados.

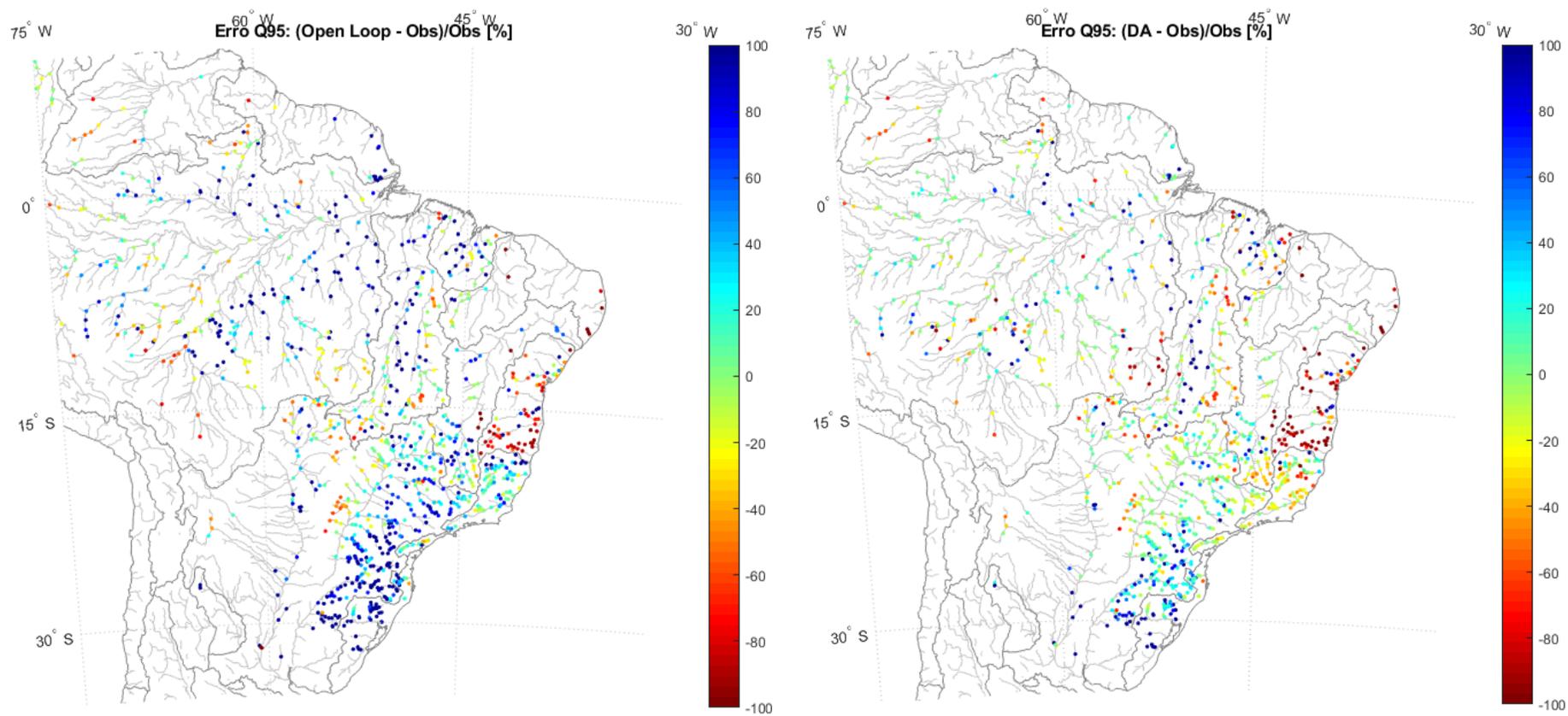


Figura 37 – Mapa de erros na estimativa da Q95 com modelo MGB-SA open loop (à esquerda) e em versão com assimilação de dados de todas as estações.

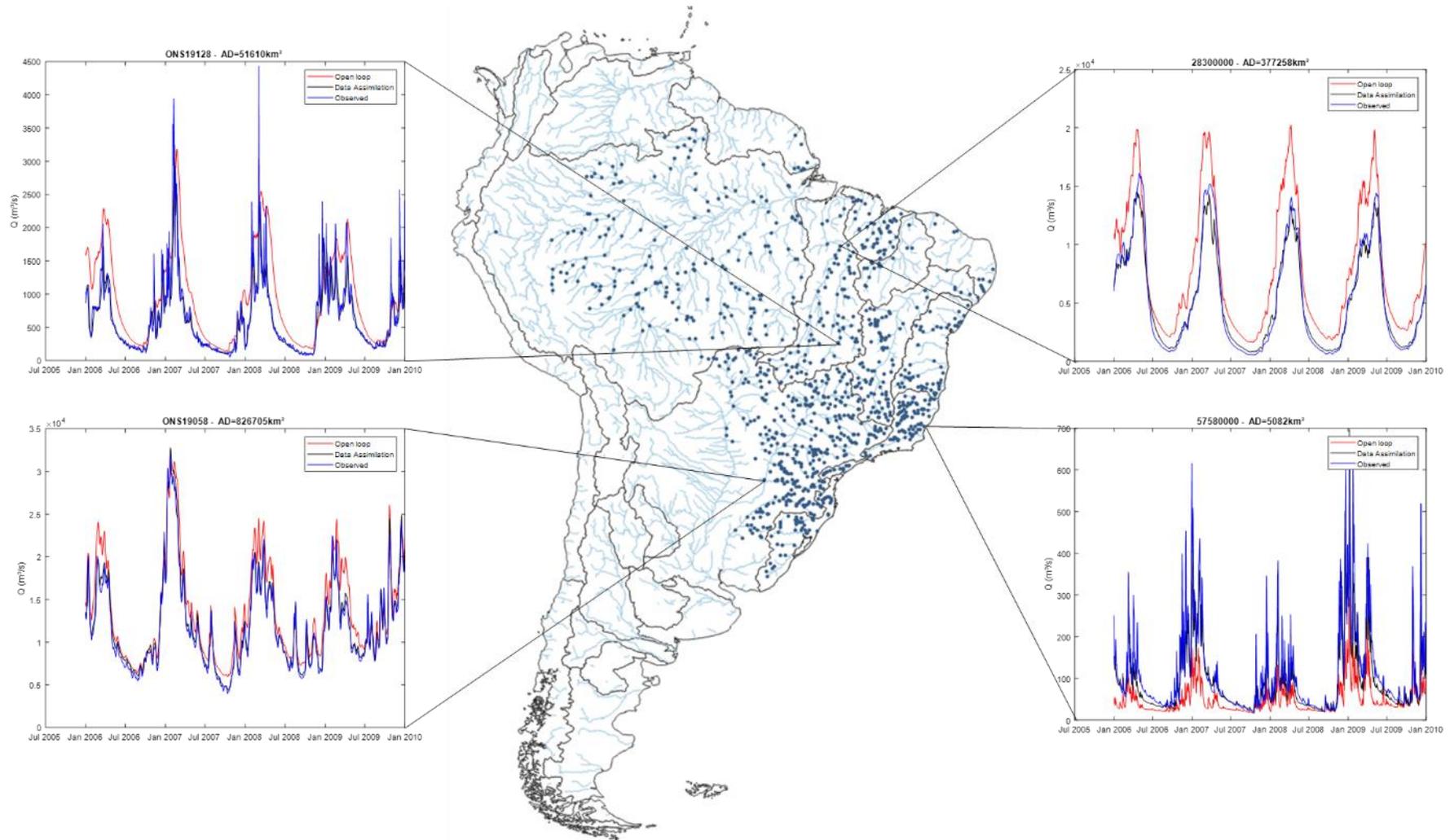


Figura 38 - Hidrogramas com vazões observadas (em azul); vazões calculadas com versão original do modelo MGB-SA (Open Loop, em vermelho) e; vazões calculadas em simulação com assimilação de dados em estações vizinhas.

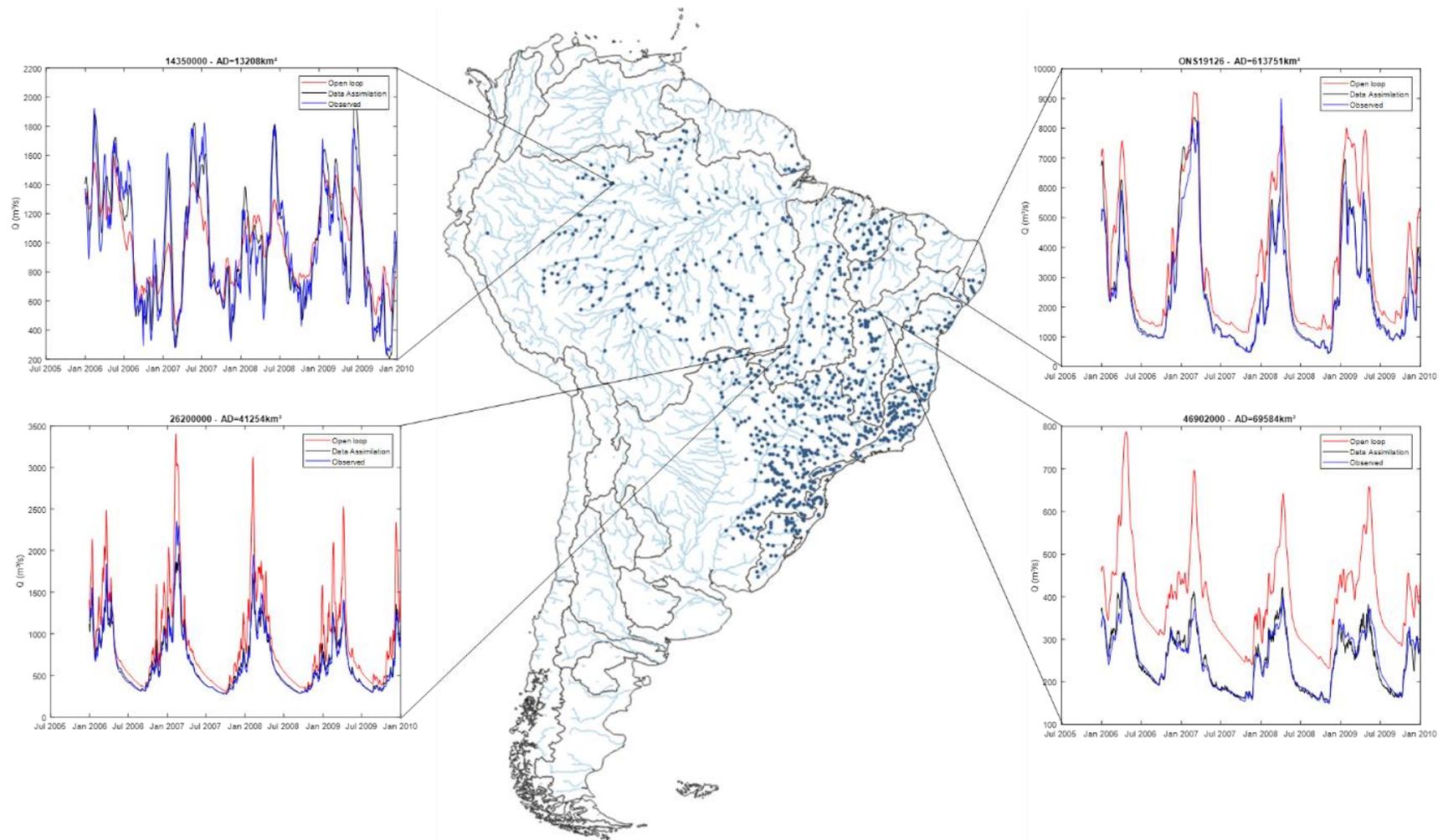


Figura 39 - Hidrogramas com vazões observadas (em azul); vazões calculadas com versão original do modelo MGB-SA (Open Loop, em vermelho) e; vazões calculadas em simulação com assimilação de dados em estações vizinhas.

5 REFERÊNCIAS BIBLIOGRÁFICAS

Alexandre, A. M. B., Martins, E. S., Clarke, R. T., & Reis Jr, D. S. (2005). Regionalização de parâmetros de modelos hidrológicos. Anais do XVI Simpósio Brasileiro de Recursos Hídricos, ABRH, João Pessoa–PB.

Andreadis, K. M., Schumann, G. J. P., and Pavelsky, T.: A simple global river bankfull width and depth database, *Water Resources Research*, 49, 7164-7168, 10.1002/wrcr.20440, 2013.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F., 2015. Era Interim/ land, a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, 19(1), 389–407.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, *Hydrology and Earth System Sciences*, 21, 589-615, 10.5194/hess-21-589-2017, 2017.

Beighley, R. E., and Gummadi, V.: Developing channel and floodplain dimensions with limited data: A case study in the Amazon Basin, *Earth Surface Processes and Landforms*, 36, 1059-1071, 10.1002/esp.2132, 2011.

CLARK, M. P.; RUPP D. E.; WOODS, R. A.; ZHENG, X. IBBITT, R. P.; SLATER, A. G.; SCHMIDT, J.; UDDSTROM, M. J. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advanced in Water resources*. v. 31, n. 10, p. 1309-1324, 2008.

DINIZ, L.S. e CLARKE, R.T. (2001) “Regionalização de parâmetros de modelo chuva-vazão usando redes neurais” In: Anais do XIV Simpósio Brasileiro de Recursos Hídricos e V Simpósio de Hidráulica e Recursos Hídricos dos Países de Língua Portuguesa, Aracajú: Associação Brasileira de Recursos Hídricos. CD-ROM

Engeland, K. and Hisdal, H.: A Comparison of Low Flow Estimates in Ungauged Catchments Using Regional Regression and the HBV-Model, *Water Resour. Manage.*, 23, 2567–2586, doi:10.1007/s11269-008-9397-7, 2009.

Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation *Ocean Dynamics*. 53, pg. 343–367.

Evensen, G. (2004). Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dynam.*, 54, pg. 539–560.

Fan, F. M., Buarque, D. C., Pontes, P. R. M., and Collischonn, W.: Um mapa de unidades de resposta hidrológica para a América do Sul, XXI Simpósio Brasileiro de Recursos Hídricos, Brasília, 2015, PAP019919,

Farmer, W.H., Archfield, S.A., Over, T.M., Hay, L.E., LaFontaine, J.H., and Kiang, J.E., 2014, A comparison of methods to predict historical daily streamflow time series in the southeastern United States: U.S. Geological Survey Scientific Investigations Report 2014–5231, 34 p., <http://dx.doi.org/10.3133/sir2014-5231>

- Gaspari, G., & Cohn, S. E. 1999. Construction of correlation functions in two and three dimensions. *Quarterly. Journal Royal Meteorological Society*, 125(554), 723-757.
- Gibson, J. and for Medium Range Weather Forecasts, E. C., 1997. ECMWF Re-analysis Project Report Series, ERA description. Number v. 1. European Centre for Medium-Range Weather Forecasts.
- Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., ... & Voss, F. (2012). Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. *Journal of Hydrometeorology*, 13(2), 604-620.
- Guo, Y., Zhang, Y., Zhang, L., & Wang, Z. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdisciplinary Reviews: Water*, e1487.
- Hailegeorgis, T.T. and Alfredsen, K. (2017) Regional Flood Frequency Analysis and Prediction in Ungauged Basins Including Estimation of Major Uncertainties for Mid-Norway. *Journal of Hydrology: Regional Studies*, 9, 104-126. <https://doi.org/10.1016/j.ejrh.2016.11.004>
- Hamill, T.M., Whitaker, J.S., Snyder, C. 2001. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, 129, 2776–2790.
- Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present, *Earth Syst. Sci. Data Discuss.*, <https://doi.org/10.5194/essd-2019-232>, in review, 2020.
- Houtekamer, P.L. and Mitchell, 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, 129, 123–137.
- Katzfuss, M.; Stroud, J.R.; Wikle, C.K. (2016). Understanding the Ensemble Kalman Filter. *The American Statistician*, vol. 70 (4), pg. 350-357.
- Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M., 2001. The ncep–ncar 50–year reanalysis, Monthly means cd–rom and documentation. *Bulletin of the American Meteorological Society*, 82(2), 247–267.
- Lehner, B., Liermann, C.R., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., Doll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rodel, R., Sindorf, N., Wisser, D., 2011. High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. *Front. Ecol. Environ.* 9 (9), 494–502
- Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., et al. (2019). Global reconstruction of naturalized river flows at 2.94 million reaches. *Water Resources Research*, 55, 6499–6516. <https://doi.org/10.1029/2019WR025287>.
- Liu, Y.; Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research.*, vol. 43.
- Lopes, V. A. R.: Modelagem hidrológica e hidrodinâmica integrada de bacias e sistemas lagunares com influência do vento. Dissertação de mestrado, Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 215 pp., 2017.

Murphy, J.C., Knight, R.R., Wolfe, W.J., and Gain, W.S., 2012, Predicting ecological flow regime at ungaged sites: A comparison of methods: *River Research and Applications*, doi: 10.1002/rra.2570.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, *Climate Research*, 21, 1-25, 2002.

O'Loughlin, F. E., Paiva, R. C. D., Durand, M., Alsdorf, D. E., and Bates, P. D.: A multi-sensor approach towards a global vegetation corrected SRTM DEM product, *Remote Sensing of Environment*, 182, 49-59, 10.1016/j.rse.2016.04.018, 2016.

Otsuki, R. G., & Reis Jr, D. S. (2011). Análise Comparativa de Metodologias de Estimativa de Séries de Vazões Médias Mensais Aplicadas a Estudos Energéticos de Aproveitamentos Hidrelétricos. XIX Simpósio Brasileiro de Recursos Hídricos, Maceió, AL.

Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A.V., Kostelich, E.J., Corazza, M., Kalnay, E., Patil, D.J., and Yorke, J. 2004. A local ensemble Kalman filter for atmospheric data assimilation, *Tellus*, 56A, 415–428

Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M. P., Frappart, F., Calmant, S., and Bulhões Mendes, C. A.: Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin, *Water Resources Research*, 49, 1226-1243, 10.1002/wrcr.20067, 2013.

PESSOA, Margarita Maria Elisa Pereira. Obtenção De Vazões Em Locais Sem Dados Através De Modelos Hidrológicos Com Assimilação De Dados. 2017. Tese (Doutorado em Recursos Hidricos e Saneamento Ambiental) - Universidade Federal do Rio Grande do Sul, Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Petheram, C., Rustomji, P., Neal, B., & Woodman, A. (2011). Regionalising low-flow responses in large tropical catchments: a comparison of rainfall-runoff modelling and regression approaches. 19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011

Pfannerstill, M., Guse, B., & Fohrer, N. (2014). Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *Journal of Hydrology*, 510, 447-458.

Pontes, P. R. M.: Modelagem hidrológica e hidrodinâmica integrada da bacia do Prata. Tese de Doutorado, Instituto de Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 194 pp., 2016.

Reichle, R. and Liu, Q. 2015. Precipitation and global land surface hydrology in the MERRALand and MERRA- 2 reanalysis datasets. In EGU General Assembly Conference Abstracts, volume 17 of EGU General Assembly Conference Abstracts, page 1838.

Reichle, R. H., 2008. Data assimilation methods in the Earth sciences. *Advances in Water Resources*, Volume 31, pg. 1411-1418.

Sakov, P. and Bertino, L. 2011. Relation between two common localisation methods for the EnKF. *Comput Geosci*, 15: 225. <https://doi.org/10.1007/s10596-010-9202-6>

Saraiva, I., Fernandes, W., & Naghettini, M. (2011). Simulação Hidrológica Mensal em Bacias Hidrográficas sem Monitoramento Fluviométrico. *RBRH—Revista Brasileira de Recursos Hídricos*, 16(1), 115-125.

SIQUEIRA, V. A. et al. (2018). Toward continental hydrologic–hydrodynamic modeling in South America. *Hydrology and Earth System Sciences*. Göttingen: Copernicus. Vol. 22, n. 9 (set. 2018), pp. 4815-4842.

Tong, X.T. 2018. Performance Analysis of Local Ensemble Kalman Filter. *Journal of Nonlinear Sci.*, 1–46. <https://doi.org/10.1007/s00332-018-9453-2>

Virões, M. V., & Cirilo, J. A. (2019). Regionalization of hydrological model parameters for the semi-arid region of the northeast Brazil. *RBRH*, 24.

Vrugt, J. A.; Gupta, H. V.; Ó Nualláin, B.; Bouten, W. (2005). Real-Time data assimilation for operational ensemble streamflow forecasting. *Journal of Hydrometeorology*, vol. 7, pg. 548-565.

Wanders, N., Bierkens, M.F., Sutanudjaja, E., van Beek, R., 2014. The PCR-GLOBWB global hydrological reanalysis product. In *EGU General Assembly Conference Abstracts*, volume 16 of *EGU General Assembly Conference Abstracts*, 16, EGU2014-5369.

Wongchuig, C.S., de Paiva, R.C.D., Espinoza, J.C., Collischonn, W. 2017. Multi-decadal hydrological retrospective: case study of Amazon floods and droughts. *Journal of Hydrology*, 549:667-684. <https://doi.org/10.1016/j.jhydrol.2017.04.01>

Wongchuig, C.S., de Paiva, R.C.D., Siqueira, V., Collischonn, W. 2019b. Hydrological Reanalysis Across the 20th Century: A Case Study of the Amazon Basin. *Journal of Hydrology*, 570:755-773. <https://doi.org/10.1016/j.jhydrol.2019.01.025>

Wongchuig, C.S., Fleischman, A., de Paiva, R.C.D., Fadel, W.A. 2019. Towards the Discharge Estimation for Water Resources Management Uses with a Semi-Distributed Model and Ensemble Kalman Filter Data Assimilation. *Hydrological Science Journal*.

Wongchuig-Correa, S., de Paiva, R. C. D., Biancamaria, S., & Collischonn, W. (2020). Assimilation of future SWOT-based river elevations, surface extent observations and discharge estimations into uncertain global hydrological models. *Journal of Hydrology*, 590, 125473.