

Centro de Pesquisas de Energia Elétrica - CEPEL



Relatório Técnico

Nº/Ano: 29679/2016 **Nº de Páginas:** 31 **Nº de Anexos:** 1

Título: Estudo de configurações para definição de limites de variação de vazões.

Departamento: Departamento de Otimização Energética e Meio Ambiente - DEA

Área de Responsabilidade: B200 **Conta de Apropriação:** 1335

Cliente: Centrais Elétricas Brasileiras S/A - Eletrobras Av. Presidente Vargas, 409 - 12º andar 20071-003 - Rio de Janeiro/RJ	Equipe de acompanhamento: Jonatan Ross - Eletrobras Lilian L. Sampaio - Eletrobras Marcelo Jaques Martins - Eletrobras Flávio Corga Cardinot - Eletrobras
Atenção: Egídio Schoenberger	

Resumo: Este relatório apresenta resultados da análise do efeito da agregação de períodos semanais na definição da função de distribuição acumulada de probabilidades das variações de vazão entre duas semanas consecutivas, que é utilizada para a definição de limites de variação de vazão no modelo PREVIVAZ.

Autores:
Luciano N. Rodrigues Xavier - Cepel *lx*
Juan Pereira Colonese - Cepel *Juan Pereira Colonese*

Palavras-Chave:
Limites de variação de vazões, previsão de vazões, Previvaz

Classificação: CONTROLADO

Gerente de Projeto
Luciano N. Rodrigues Xavier

Nome: Luciano Nóbrega Rodrigues Xavier
Tel.: 21-2598-6074 **Fax:** 21-2598-6482
E-mail: xavier@cepel.br

Chefe do Departamento de Otimização Energética e Meio Ambiente
Maria Elvira Piñeiro Maceira

Nome: Maria Elvira Piñeiro Maceira
Tel.: 21-2598-6454 **Fax:** 21-2598-6482
E-mail: elvira@cepel.br

Aprovação

RP

Roberto Pereira Caldas
Diretor de Pesquisa, Desenvolvimento e Inovação

16/12/16

CONTEÚDO

1	Sumário.....	2
2	Introdução.....	3
3	Uso de limites de previsão no modelo Previvaz.....	4
4	Caso Tucuruí no PMo de abril/2016.....	8
4.1	Critérios adotados para a definição dos limites e limites adotados: posto 275.....	8
4.2	Modelos escolhidos em cada semana após a consideração dos limites de previsão.....	9
4.3	Conclusões da análise empreendida.....	10
5	Estudo.....	11
5.1	Teste não-paramétrico de Mann-Whitney.....	11
5.2	Postos para estudo.....	13
6	Resultados.....	14
6.1	Comparação entre funções agregadas e originais.....	14
6.2	Mapas com o percentual de casos com funções significativamente diferentes das originais.....	18
7	Conclusões e ações propostas.....	25
8	Referências Bibliográficas.....	27
9	Anexos.....	28

1 SUMÁRIO

O presente Relatório Técnico tem como objetivo a análise do efeito da agregação de períodos semanais para a definição das funções de distribuição acumulada de probabilidades das variações de vazões consecutivas. Essas funções, a partir das quais são definidos os limites de previsão adotados no modelo PREVIVAZ, podem ser construídas a partir da agregação em períodos mensais, trimestrais e/ou semestrais, com o intuito de reduzir o erro em sua estimação. Por outro lado, há o risco de que a função resultante não represente adequadamente a variação de vazões em semanas específicas, como ocorrido em um caso preliminar ao PMO de Abril/2016, e que motivou o estudo cujos resultados encontram-se consubstanciados neste Relatório. A partir de resultados obtidos para 13 postos correspondentes a usinas hidroelétricas pertencentes ao Sistema Interligado Nacional, verificou-se que a agregação em períodos superiores ao mensal pode levar a inconsistências consideráveis na função de distribuição de probabilidades resultante, com conseqüente erro nos limites definidos, o que pode afetar sobramaneira o uso do modelo PREVIVAZ. Após a apresentação dos resultados, indica-se um conjunto de novos estudos e ações a serem empreendidos.

2 INTRODUÇÃO

O modelo estocástico de previsão de vazões semanais PREVIVAZ é utilizado no planejamento da operação do Sistema Interligado Nacional. Baseado na modelagem de séries temporais, o PREVIVAZ avalia um conjunto de alternativas de modelos do tipo ARMA(p,q) e fornece, em um horizonte de seis semanas, o valor esperado e o intervalo teórico de previsões definido a partir de nível de significância definido *a priori* [1].

O PREVIVAZ permite, desde sua versão 5.3, disponibilizada em 2009, o uso de limites de variação de vazões, de modo a evitar a ocorrência de valores atípicos de previsão, de acordo com as características da série histórica de vazões semanais. A definição desses limites, que podem ser adotados tanto na etapa de definição dos parâmetros dos modelos ARMA(p,q) quanto na de previsão, baseia-se na função de distribuição acumulada de probabilidades das variações de vazão entre semanas consecutivas. Essa função é definida a partir do histórico de vazões e percentis de variações mínima e máxima fornecidos pelo usuário definirão os limites de variação de vazões. Para a construção dessa função, permite-se que as variações de vazões sejam consideradas, também, segundo a agregação em períodos mensais, trimestrais e/ou semestrais, de modo a reduzir o erro seu erro de estimação [2].

No presente Relatório Técnico serão apresentados os resultados de um estudo inicial de verificação da influência da agregação das semanas em períodos mensais, trimestrais e semestrais para a definição da função de distribuição acumulada de probabilidades. A motivação central para sua elaboração vem do ocorrido em estudo preliminar ao PMO de Abril/2016 para a previsão de vazões afluentes à UHE Tucuruí, conforme descrito em [3], no qual o uso de determinada configuração em uma tendência hidrológica atipicamente baixa levou à definição de limites inadequados, e estes, a uma previsão incoerente.

Para o presente estudo, foi avaliada a possível ocorrência de diferenças significativas entre as curvas de distribuição acumulada de probabilidades utilizando opções de agregação temporal e separação em níveis de magnitude de variações de vazões, de acordo com as premissas adotadas e as bacias selecionadas apresentadas no item 4. No item 5 serão apresentados os resultados desse estudo, com ênfase no erro médio cometido pela agregação.

3 USO DE LIMITES DE PREVISÃO NO MODELO PREVIVAZ

Com o objetivo de reduzir grandes erros de previsão causados por variações atípicas de vazões, o CEPEL desenvolveu um procedimento para o uso de limites de variação de previsões pelo programa PREVIVAZ. Conforme exposto em [2], essas variações atípicas - isto é, incomuns segundo o registro de vazões semanais - podem ser causadas pela transformação Box-Cox inversa ou pela ocorrência de chuvas "atípicas". Essa metodologia foi implementada na versão 5.3, cujo desenvolvimento deu origem à versão 6 validada em dezembro de 2014 por Força-Tarefa sob comando do ONS.

De forma sucinta, os limites de previsão são definidos de acordo com o seguinte procedimento:

1) Definição da função de distribuição acumulada (*fda*) de probabilidades das variações de vazão entre semanas consecutivas, $Q(t)/Q(t-1)$, segundo a configuração informada pelo usuário que pode considerar que:

- a) as semanas podem ser agregadas em períodos mensais, trimestrais e semestrais para definição da *fda*, de modo a reduzir o erro de sua estimação; conseqüentemente, ter-se-ia a definição de 12, 4 ou 2 funções, sendo adotada a correspondente à semana de previsão.
- b) a *fda* pode ser definida segundo diferentes (2, 3 ou 4) faixas de vazão, segundo distribuição das vazões na semana (ou, em caso de agregação, mês, trimestre ou semestre).

2) De acordo com os percentis admissíveis de variação mínima e máxima, definidos *a priori* pelo usuário, são definidos, a partir da vazão na semana anterior, $Q(t-1)$, os limites mínimo e máximo de variação de vazões na semana, $Q(t)$.

Os limites assim definidos podem ser utilizados na etapa de estimação dos parâmetros dos modelos auto-regressivos e na de previsão, ou apenas para a previsão:

1. Ao se utilizar os limites na etapa de estimação dos parâmetros, as alternativas cujas previsões estiverem fora do intervalo admissível terão suas previsões substituídas pelo valor do limite correspondente (inferior ou superior); conseqüentemente, o erro médio quadrático da alternativa será calculado com essa previsão modificada (isto é, o limite) o

que, naturalmente, poderá levar a uma modificação no ranking que seria definido sem o uso dos limites de previsão. Na Figura 1 abaixo, ilustrativa do uso dos limites na definição do melhor modelo em um caso hipotético, estão indicadas as regiões em que as previsões dos modelos seriam substituídas pelos limites inferior e superior, de acordo com a "posição" da previsão com respeito aos limites, e aquela em que nenhuma modificação seria feita nas previsões originais.

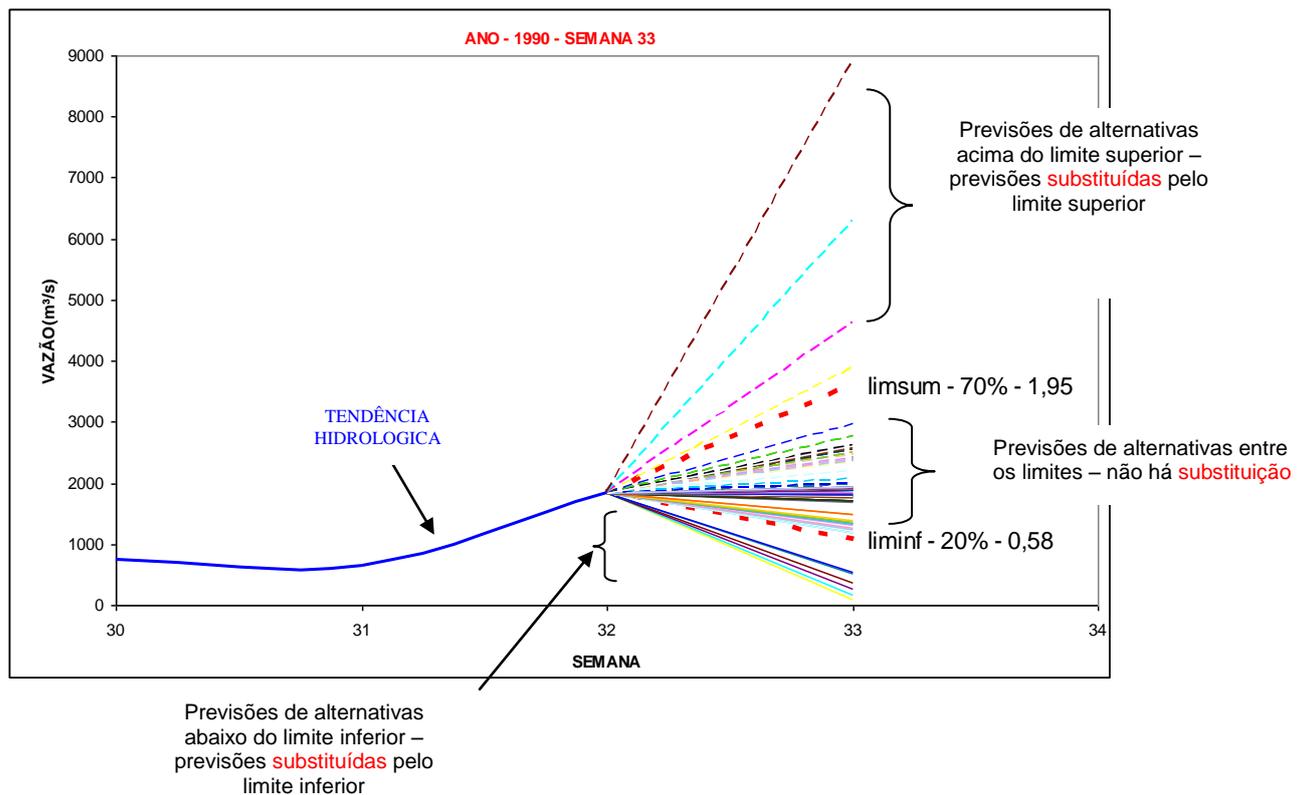


Figura 1: Conjunto de previsões para a semana 33 obtidas em um caso hipotético considerando todas as alternativas; limites inferior e superior associados a probabilidades de não excedência de 20% e 70%, respectivamente. Entre os limites, conforme indicado, região em que não seriam modificadas as previsões de cada modelo (Figura 3 de [2])

- Na etapa de previsão, será considerada a alternativa mais bem classificada a possuir a respectiva previsão contida no intervalo admissível. A Figura 2, retirada de [2], ilustra um caso hipotético em que a 4ª alternativa do ranking original foi escolhida para a previsão.

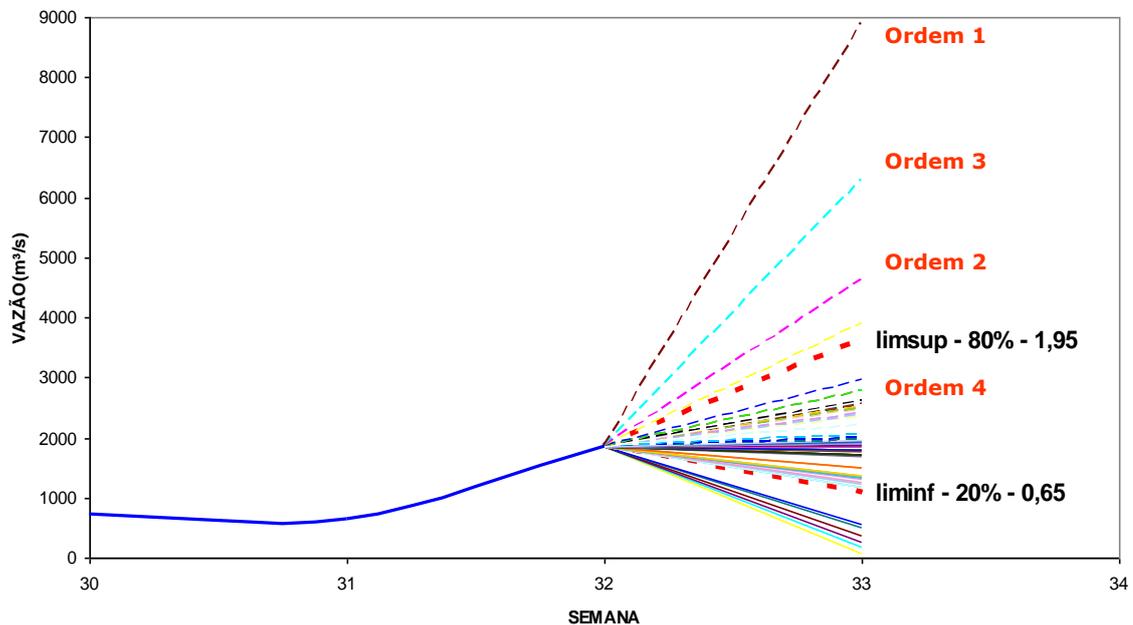


Figura 2: Conjunto de previsões para a semana 33 de um estudo hipotético, e limites inferior e superior definidos segundo probabilidades de não-excedência de 20 e 80%, respectivamente. Indicadas as 4 primeiras alternativas de modelagem segundo o ranking definido na etapa de definição do melhor modelo. Nesta situação, a 4ª alternativa seria a escolhida para a previsão. (Figura 4 de [2])

Tanto na etapa de definição do melhor modelo como na de previsão, pode ocorrer que nenhuma alternativa possua previsão contida no intervalo admissível. Nesta situação, conforme [2], três possibilidades podem ser avaliadas:

- i. caso todas as previsões estejam abaixo do limite inferior, a previsão escolhida será a da alternativa que apresentou o maior valor (mais próximo do limite inferior);
- ii. caso todas as previsões estejam acima do limite superior, será utilizada a previsão obtida pela alternativa que teve a menor previsão (mais próximo do limite superior); e
- iii. finalmente, na situação de todos os valores das previsões encontrarem-se fora dos limites, porém alguns acima do limite superior e os outros abaixo do limite inferior, determina-se, então, os números de previsões acima do limite superior e abaixo do limite inferior e escolhe-se aquele que estiver mais próxima do limite além do qual encontra-se o maior número de previsões.

Ressalta-se que, na primeira semana do horizonte de previsão, t_1 , a vazão $Q(t_1-1)$ corresponderá ao valor informado no histórico de vazões (ou no arquivo vazpast.dat); nas semanas 2 a 6, $Q(t_{2,\dots,6}-1)$ corresponderá ao valor da previsão na semana imediatamente anterior.

Conforme o praticado nos Programas Mensais de Operação, os limites de previsão são utilizados para todos os postos **apenas na etapa de previsão**. Segundo o conjunto de dados adotado no PMO de Dezembro/2016, disponível no endereço eletrônico http://www.ons.org.br/publicacao/PrevisaoVazoes/EntradaSaida/2016/12_DEZEMBRO/, para todos os postos foram adotados seguintes critérios para a definição dos limites de previsão:

- 3 faixas de vazão (33 e 66%)
- Agrupamento trimestral
- Probabilidades de não-excedência: 20 e 80%

Critérios que também foram adotados no PMO de Abril/2016, conforme será verificado no Item a seguir.

4 CASO TUCURUÍ NO PMO DE ABRIL/2016

No que segue, será apresentada uma descrição sucinta do caso que motivou o estudo descrito no presente Relatório; a análise completa desse caso foi apresentada em [3].

4.1 Critérios adotados para a definição dos limites e limites adotados: posto 275

Para a previsão de vazões semanais no posto de Tucuruí, conforme informado nos arquivos de dados gerais e de limites (cuja descrição pode ser encontrada em [4]) correspondentes, respectivamente *275.inp* e *275.lim*, os limites de previsão deveriam ser definidos de acordo com os seguintes critérios:

- Uso de limites apenas na previsão
- 3 faixas de vazão (33 e 66%)
- Agrupamento trimestral
- Probabilidades de não-excedência: 20 e 80%

A partir desses critérios, e considerando a vazão informada na semana 13 (13.832 m³/s), os seguintes limites de previsão foram definidos e impressos no arquivo *Limites.txt* [4]:

Tabela 1 - Limites de previsão para o posto de Tucuruí [3]

Semana	LIMINF	LIMSUP
14	10.888,91	13.859,07
15	10.886,69	13.856,25
16	8.758,67	9.709,01
17	8.758,67	9.709,01
18	8.758,67	9.709,01
19	8.290,73	9.190,29

Cabe observar que, como destacado anteriormente, para a definição dos limites de previsão nas semanas de 15 a 19 foram considerados os valores esperados de previsão determinados nas semanas imediatamente anteriores (14 a 18).

4.2 Modelos escolhidos em cada semana após a consideração dos limites de previsão

De acordo com o impresso no arquivo *275_mod.dat*, cuja descrição pode ser obtida em [4], os seguintes modelos foram escolhidos para a previsão nas semanas 14 a 19:

Tabela 2 - Modelos adotados em cada semana para a previsão em Tucuruí [3]

Semana	Modelo	Posição original
14	TPARMA(3;1)-G4	102
15	CONSTANTE	140
16	CONSTANTE	0
17	CONSTANTE	0
18	TPAR(4)-G1	0
19	TPAR(4)-RO	0

Como pode ser verificado após consulta ao arquivo *Previsao.dat*, em que são impressas as previsões de todos os modelos considerados nas seis semanas do horizonte de estudo [4], correspondente a esse caso, na semana 15 o modelo "Constante" é o único cuja previsão está contida no intervalo definido pelos limites, conforme apresentado na Tabela 1. A partir de consulta ao arquivo *275.exc*, em que são impressos, para cada semana do ano operativo, os respectivos rankings obtidos na etapa de definição do melhor modelo [4], verificou-se que

- Na semana 14, após considerar os limites de previsão, o modelo adotado foi o TPARMA(3,1)-G4 (transformação Box-Cox, função de autocorrelação amostral calculada sem considerar agrupamento)
- Na semana 15, o modelo "Constante" foi o adotado, conforme anteriormente mencionado. Como indicado no arquivo *275_mod.dat* (vide Tabela 2), esse modelo ocupava a última posição (140) no ordenamento original.
- Na semana 16, o uso dos limites não modificou o modelo adotado; entretanto, ao consultar o arquivo *275_mod.dat* pode-se verificar que o modelo adotado foi o "Constante", que possui um valor de previsão superior ao limite superior ("posição original 0"). Este fato pode ser diretamente verificado confrontando as previsões dos 140 modelos, conforme arquivo *Previsao.dat*, com os limites de previsão impressos no arquivo *Limites.txt*, apresentados na Tabela 1.

Ao considerar o agrupamento trimestral, os limites foram calculados a partir da função de distribuição acumulada das variações de vazão observadas de Abril a Junho (semanas 14 a 26). Como pode ser observado a partir de uma análise do histórico, as variações de vazões na metade final deste período, época onde são menos frequentes e menos intensos os aumentos de vazão, apresentam comportamento distinto do início do período, quando ainda ocorrem incrementos de vazão. Assim, ao considerar todas as variações das semanas correspondentes, a função de distribuição acumulada resultante acabou levando à definição de limites superiores muito distintos dos que seriam obtidos ao considerar apenas as variações entre as semanas mais próximas das semanas 13 a 19. Como resultado obteve-se neste período limites superiores muito restritivos.

4.3 Conclusões da análise empreendida

Conforme a análise dos resultados apresentada em [3], as seguintes conclusões puderam ser destacadas:

- 1) Os limites de previsão calculados para a semana 15, definidos a partir da série histórica de Tucuruí e da opção de cálculo dos limites adotada (agrupamento trimestral), não acompanharam as previsões dos modelos mais bem classificados, os quais previam um aumento da vazão. Ressalta-se que a vazão da semana 13, $13.832 \text{ m}^3/\text{s}$, era a segunda menor dessa semana no histórico.
- 2) Considerando a previsão para a semana 15, dada pelo modelo constante (42% da média semanal), os limites de previsão calculados segundo o agrupamento trimestral para as semanas 16 e 17 foram ainda mais restritos, o que tornou impossível que a previsão de qualquer modelo estivesse no intervalo admissível.
- 3) Com respeito ao cálculo dos limites de previsão, sugere-se uma atenção especial ao se adotar um dos agrupamentos possíveis (mensal, trimestral e semestral), sobretudo para postos com marcado comportamento sazonal e em período de transição entre as estações. Na ausência de um estudo mais detalhado para cada posto, pode-se sugerir o uso do agrupamento mensal, como uma forma de compromisso entre (i) a qualidade da estimação da função de distribuição acumulada das variações de vazão e (ii) o uso de semanas nas quais as vazões apresentam comportamento similar.

5 ESTUDO

Neste estudo, foram avaliados os impactos do tipo de agregação adotado (em períodos mensais e trimestrais) na definição das funções de distribuição acumulada de probabilidade, considerando a divisão ou não em faixas de vazão. A análise consistiu no cálculo do percentual de semanas em que a função "agregada" foi diferente da *fd*a semanal, segundo o teste de Mann-Whitney [5] e considerando um nível de significância de 95%.

5.1 Teste não-paramétrico de Mann-Whitney

O teste estatístico selecionado é apropriado para analisar se um conjunto de dados tende a produzir valores maiores (ou menores) do que um segundo conjunto de dados. Assim, a hipótese nula testada tem a forma:

$$H_0: Prob[x > y] = 0,5,$$

onde x e y representam dois conjuntos de dados. No caso de um teste bicaudal a hipótese alternativa será:

$$H_1: Prob[x > y] \neq 0,5.$$

O teste de Mann-Whitney é comparável ao teste t , porém, além da independência a parâmetros e invariabilidade a questões de escalas ou transformações aplicadas aos dados, ele possui a vantagem de não assumir normalidade (nem qualquer outra distribuição específica) da amostra testada. O teste se dá a partir do somatório dos ranqueamentos entre as duas amostras de dados. Em casos, nos quais as distribuições das amostras apresentam forma semelhante, o teste de Mann-Whitney é, simplificada, um teste para o deslocamento das medianas. Assim, considera-se que o teste possua sensibilidade para detecção de diferenças concentradas nos extremos das distribuições amostrais [5].

O procedimento para cálculo da estatística do teste de Mann-Whitney adotado neste relatório foi:

- 1) atribuir a cada valor presente nas amostras um ranqueamento numérico começando em 1 para o menor valor. Caso existam valores iguais, atribui-se para todos estes o valor central dos ranqueamentos não ajustados, por exemplo, dois valores empatados para a posição 2, atribui-se o ranqueamento 2,5 para todos;

2) Soma-se os ranqueamentos provenientes de uma das amostras e obtém-se, R_1 . Como a soma total dos rankings é igual a $N(N+1)/2$, onde N é o número total de dados ($N=n_1+n_2$), obtém-se conseqüentemente R_2 , soma dos ranqueamentos do segundo conjunto de dados;

3) O menor valor obtido pela equação,

$$U_n = R_n - n_n(n_n + 1)/2, \text{ com } n = 1, 2,$$

será a estatística de teste utilizada para consulta a tabelas de significância. Considerando as relações entre as variáveis, $U_1 + U_2 = n_1n_2$, obtém-se rapidamente um valor a partir do outro.

Para amostras grandes ($N>20$, segundo [5]) a distribuição da estatística U pode ser aproximada para uma normal. Assim, é possível obter um valor normalizado,

$$Z = (U - d/2 - \bar{U})/\sigma_U, \text{ sendo } U = \max(U_1, U_2),$$

para o qual o p-valor do teste bicaudal será

$$p = 2(1 - \text{Prob}(Z)).$$

Os valores de média e desvio padrão de U são obtidos, respectivamente, por:

$$\bar{U} = n_1n_2/2;$$

$$\sigma_U = \sqrt{n_1n_2(n_1 + n_2 + 1)/12}.$$

Como sugerido em [5], o termo d utilizado no cálculo do valor padronizado Z é um termo de correção de continuidade. Esta correção é aplicada pois a distribuição normal é ajustada para os pontos centrais das barras que compõem a distribuição da estatística exata de teste. Assim, a correção move as probabilidades de ocorrência do ponto extremo das barras para o centro da mesma. Assim, d tem o valor da menor diferença entre possíveis valores da estatística de teste. No caso do teste de Mann-Whitney este valor é 1, menor diferença entre dois ranqueamentos.

5.2 Postos para estudo

Foram selecionados 13 postos, listados na Tabela 3 a seguir, nas principais bacias hidrográficas com usinas hidroelétricas que constam do Programa Mensal de Operação (PMO). As séries semanais foram construídas a partir do histórico de vazões diárias disponibilizado no site do Operador Nacional do Sistema (disponíveis no endereço eletrônico http://www.ons.org.br/operacao/vazoes_naturais.aspx).

Tabela 3 - Postos considerados no estudo

UHE	Código	Bacia	Início da série
Tucuruí	275	Tocantins	01/09/1969
Boa Esperança	190	Paranaíba	01/01/1931
Sobradinho	168	São Francisco	01/01/1931
Manso	278	Paraguai	01/01/1966
Itumbiara	018	Paranaíba	01/08/1973
Furnas	006	Grande	01/01/1931
Ilha dos Pombos	130	Paraíba do Sul	01/01/1966
Barra Bonita	237	Tietê	01/01/1931
Capivara	061	Paranapanema	01/01/1931
Itaipu	266	Paraná	01/01/1931
Salto Santiago	077	Iguaçu	01/12/1940
Itá	092	Uruguai	01/11/1939
Passo Real	111	Passo Real	01/02/1939

6 RESULTADOS

Os resultados do estudo empreendido serão apresentados em dois grupos:

a) comparação visual das fda's agregadas (em meses e em trimestres) com a função original para os postos de Tucuruí e Salto Santiago, para semanas em que foram verificados os menores e maiores erros médios na função agregada.

b) mapas em que serão indicados o percentual de semanas em que a "fda agregada" (em meses e em trimestres) diferiu da função original, para todos os postos considerados.

6.1 Comparação entre funções agregadas e originais

A seguir são apresentadas as figuras para comparação visual entre as funções de distribuição acumulada de probabilidades para os postos de Tucuruí (Figuras 1 e 2) e Salto Santiago (Figuras 3 e 4). Como o objetivo deste item é apenas permitir uma 'inspeção' visual da possível similaridade entre as funções, escolheram-se apenas as semanas de máxima (semana 34 para Tucuruí e 40 para Salto Santiago) e mínima (semana 1 para Tucuruí e 5 para Salto Santiago) diferença média da função agregada com respeito à original, considerando-se o registro semanal sem divisão de faixas (Figuras 3 e 5) e com divisão em 4 faixas (Figuras 4 e 6).

Figura 3 - Comparação entre as funções de distribuição acumulada de probabilidades com agregação (mensal, em verde, e trimestral, em vermelho) e as fda's da semana correspondente (em azul) - **posto Tucuruí**. Semanas 1 (mês **Janeiro** e trimestre 1) e 34 (mês **Agosto** e trimestre 3). **Sem divisão em faixas.**

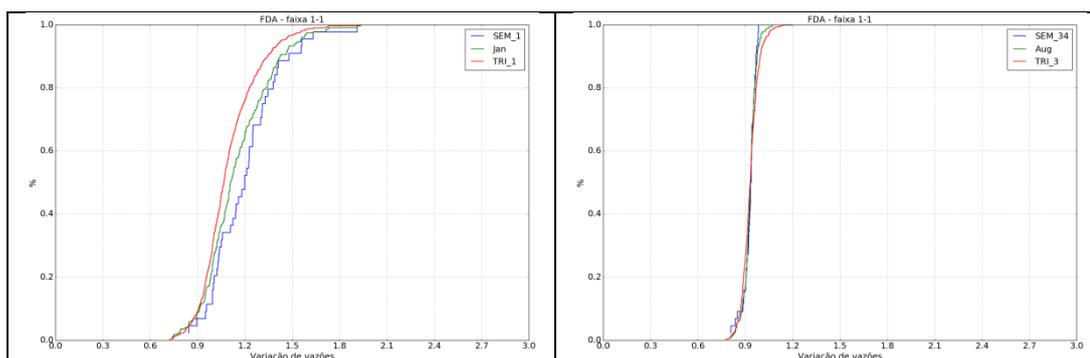
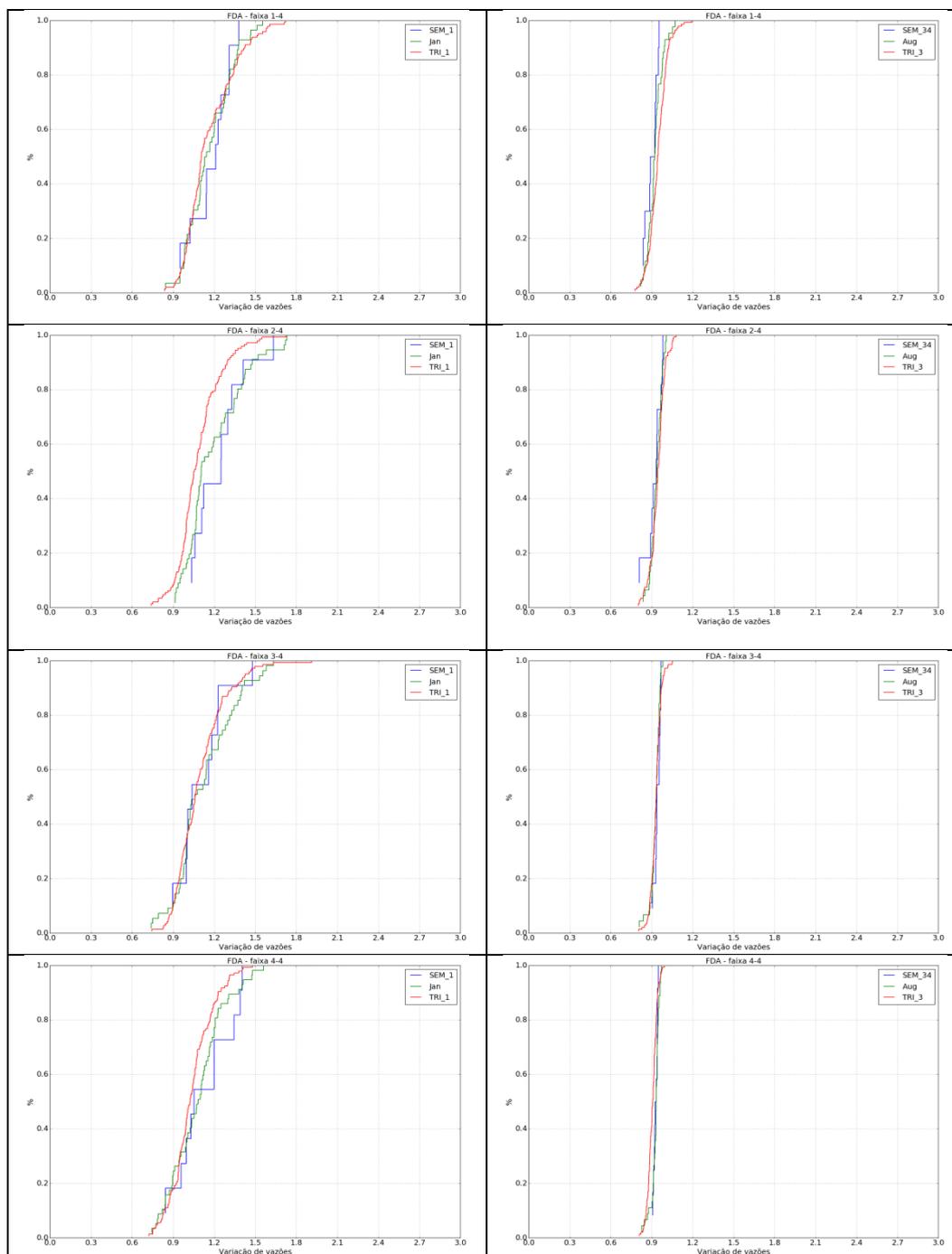


Figura 4 - Comparação entre as funções de distribuição acumulada de probabilidades com (mensal, em verde, e trimestral, em vermelho) e as *fda's* da semana correspondente (em azul) - **posto Tucuruí**. Semanas 1 (mês Janeiro e trimestre 1) e 34 (mês Agosto e trimestre 3). **Divisão em 4 faixas**.



A partir da análise das Figuras 3 e 4 pode-se verificar que:

a) para a semana 1, em período de vazões ascendentes na bacia de Tucuruí, a *fda* resultante da agregação em períodos mensais não difere de modo expressivo da função original, com exceção

dos extremos da distribuição (sobretudo, quando se adota a subdivisão do registro em faixas de vazão); no caso da agregação em períodos trimestrais (no caso, foram considerados os meses de Janeiro, Fevereiro e Março), as diferenças se tornam mais evidentes, especialmente nos extremos. Essas diferenças, entretanto, não foram consideradas significativas a partir do teste e do nível de significância adotados;

b) para a semana 34, em período de recessão na bacia, percebe-se uma grande redução nas discrepâncias, para ambas as formas de agregação (na semana em questão, foram consideradas as vazões semanais verificadas nos meses de Julho, Agosto e Setembro).

Figura 5 - Comparação entre as funções de distribuição acumulada de probabilidades com agregação (mensal, em verde, e trimestral, em vermelho) e as *fda's* da semana correspondente (em azul) - **posto Salto Santiago**. Semanas **5** (mês **Janeiro** e trimestre **1**) e **40** (mês **Outubro** e trimestre **4**). **Sem divisão em faixas**.

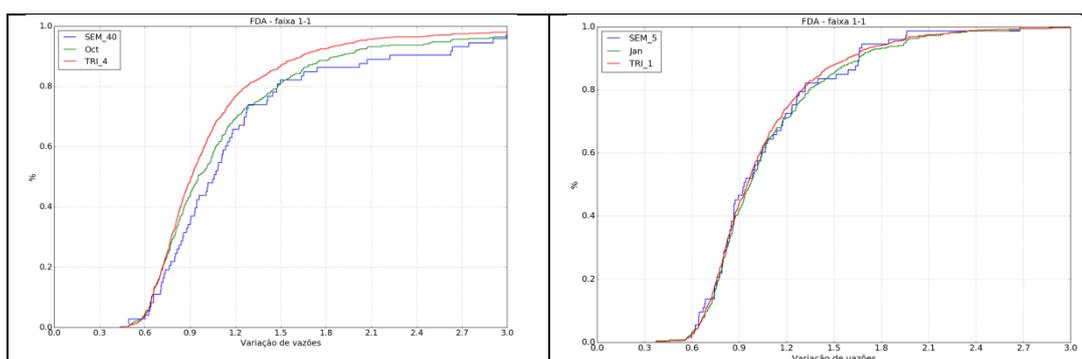
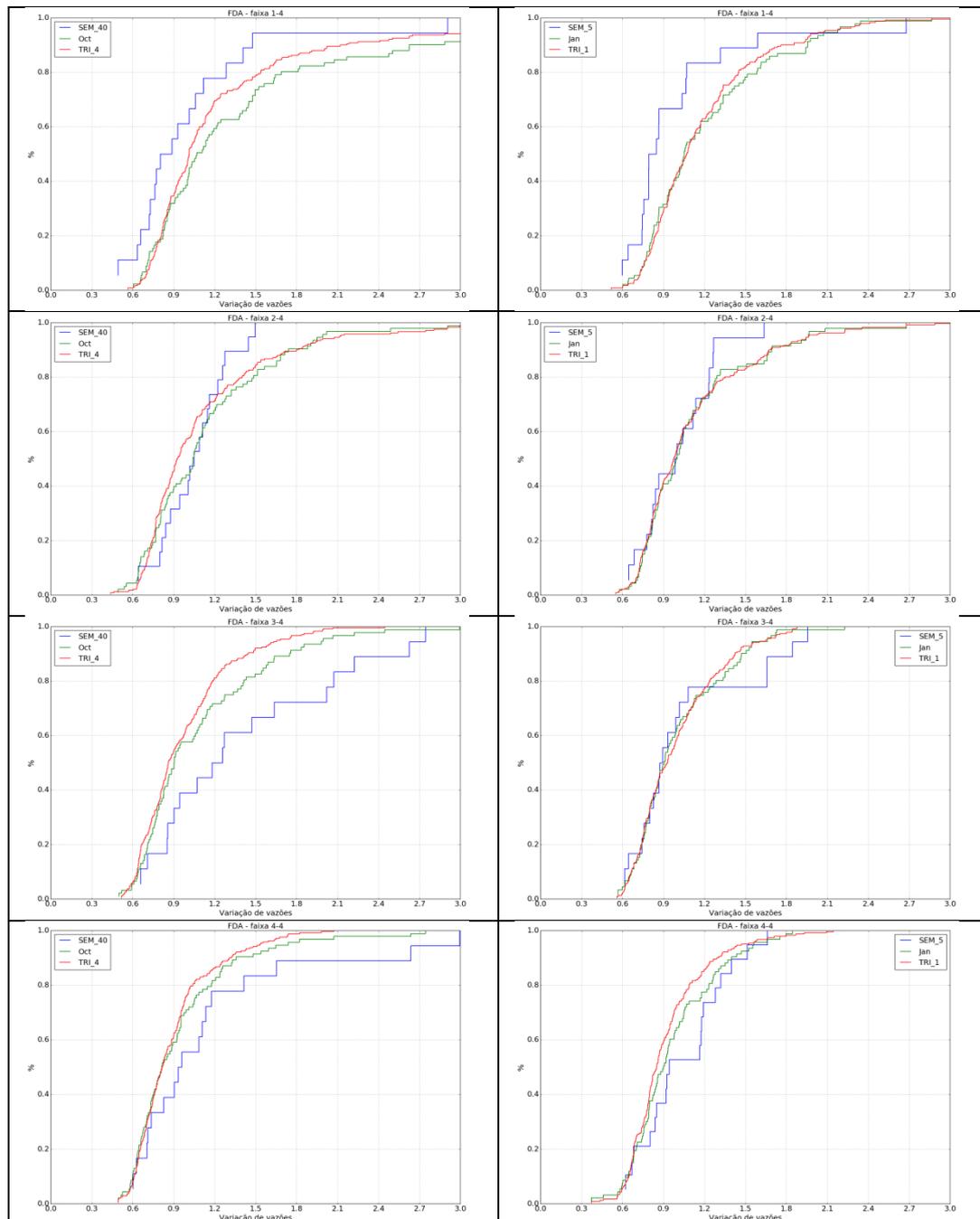


Figura 6 - Comparação entre as funções de distribuição acumulada de probabilidades com (mensal, em verde, e trimestral, em vermelho) e as *fda's* da semana correspondente (em azul) - **posto Salto Santiago**. Semanas 40 (mês Outubro e trimestre 4) e 5 (mês Janeiro e trimestre 1). Divisão em 4 faixas.



No caso de Salto Santiago, conforme Figuras 3 e 4 pode-se perceber que:

- a) para a semana 5, não houve diferença nas *fda's* obtidas com ou sem agregação temporal e sem divisão em faixas de vazão. Entretanto, quando dividido o histórico de variações de vazões em faixas de vazão há maiores discrepâncias entre as *fda's* construídas com agregação

temporal principalmente nas faixas mais extremas (faixa 1 e faixa 4). Nas faixas intermediárias as discrepâncias nas *fda*'s de cada semana com aquelas construídas a partir de agregação mensal ou trimestral são encontradas nos extremos superiores das distribuições.

b) para a semana 40, as diferenças entre as funções são evidentes, com ou sem divisão do histórico em faixas de vazão, mas se tornam maiores com a agregação em períodos trimestrais.

No caso de Tucuruí, em que há uma forte sazonalidade nas séries de vazões semanais, o efeito da agregação em períodos pode se tornar mais relevante, conforme esperado, especialmente em semanas de "transição de estações". Nesta situação, como ocorrido no caso descrito em [3], o efeito de se considerar variações de vazões em diferentes estações, cada qual apresentando diferentes comportamentos estatísticos, tende a contribuir para a definição inadequada dos limites de previsão.

6.2 Mapas com o percentual de casos com funções significativamente diferentes das originais

Seguem abaixo figuras com os mapas indicativos dos percentuais de casos em que a agregação em períodos mensais/trimestrais levou a uma função de distribuição acumulada de probabilidades significativamente diferente da *fda* semanal, segundo as premissas adotadas (uso do teste de Mann-Whitney, nível de significância de 95%). Neste item, foi considerado o registro sem divisão em faixas, e com divisão em 2, 3 e 4 faixas.

No Anexo 1 serão apresentadas as tabelas com os valores obtidos neste estudo.

Figura 7 - Percentual de *fda*'s diferentes (nível de significância de 95%) da *fda* semanal original nos postos considerados - sem divisão em faixas (esquerda: agregação em períodos mensais; direita: períodos trimestrais)

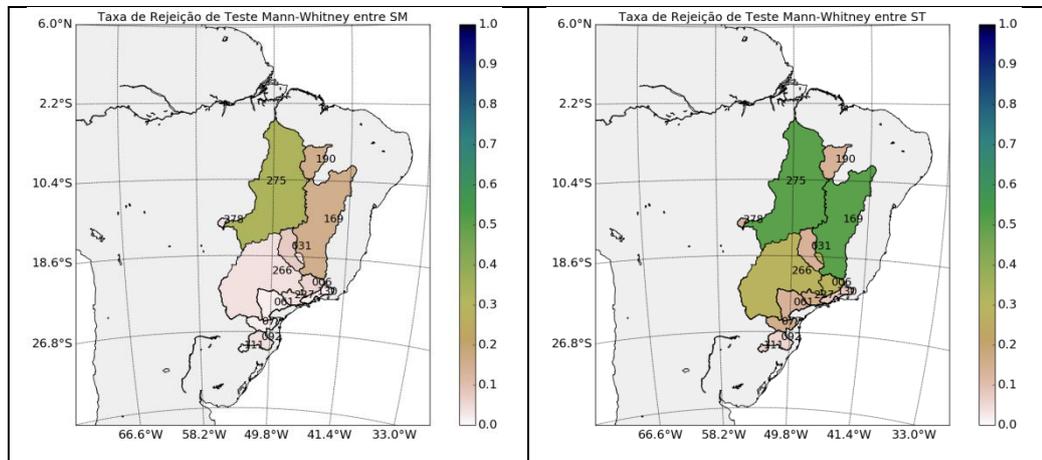


Figura 8 - Percentual de *fda*'s diferentes (nível de significância de 95%) da *fda* semanal original nos postos considerados - divisão em 2 faixas (esquerda: agregação em períodos mensais; direita: períodos trimestrais)

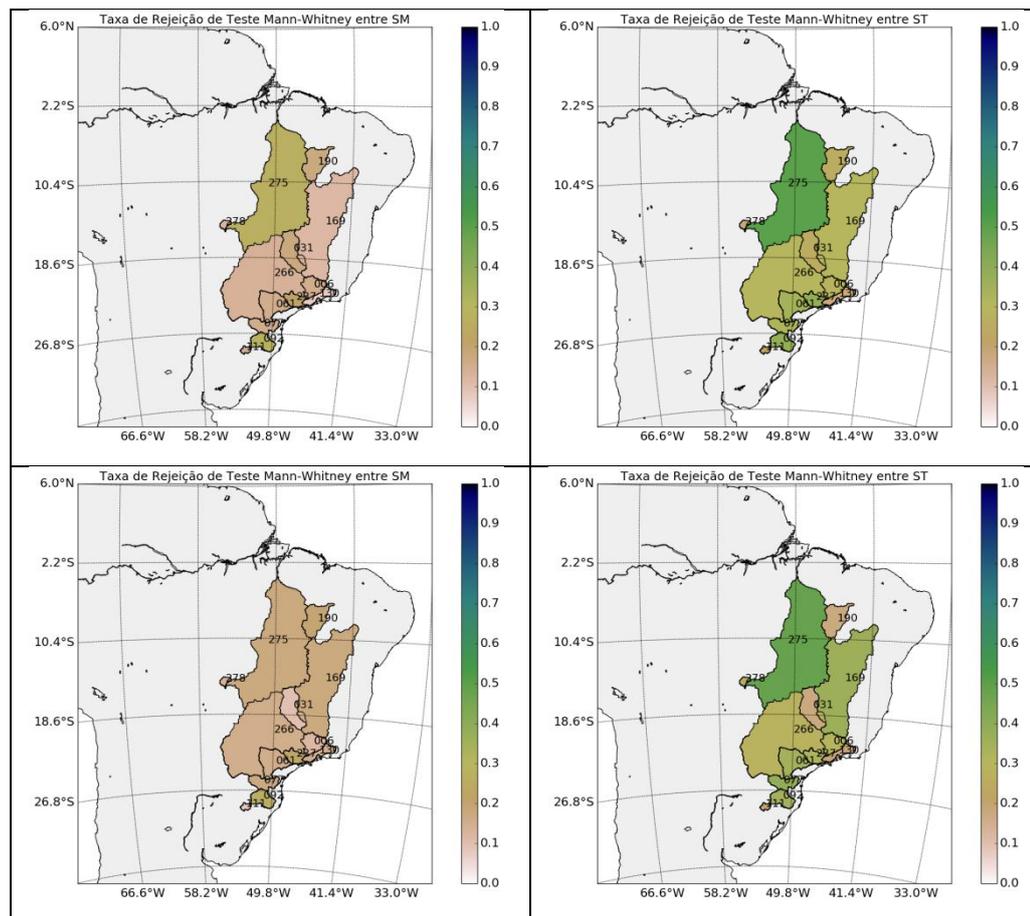


Figura 9 - Percentual de *fda*'s diferentes (nível de significância de 95%) da *fda* semanal original nos postos considerados - divisão em 3 faixas (esquerda: agregação em períodos mensais; direita: períodos trimestrais)

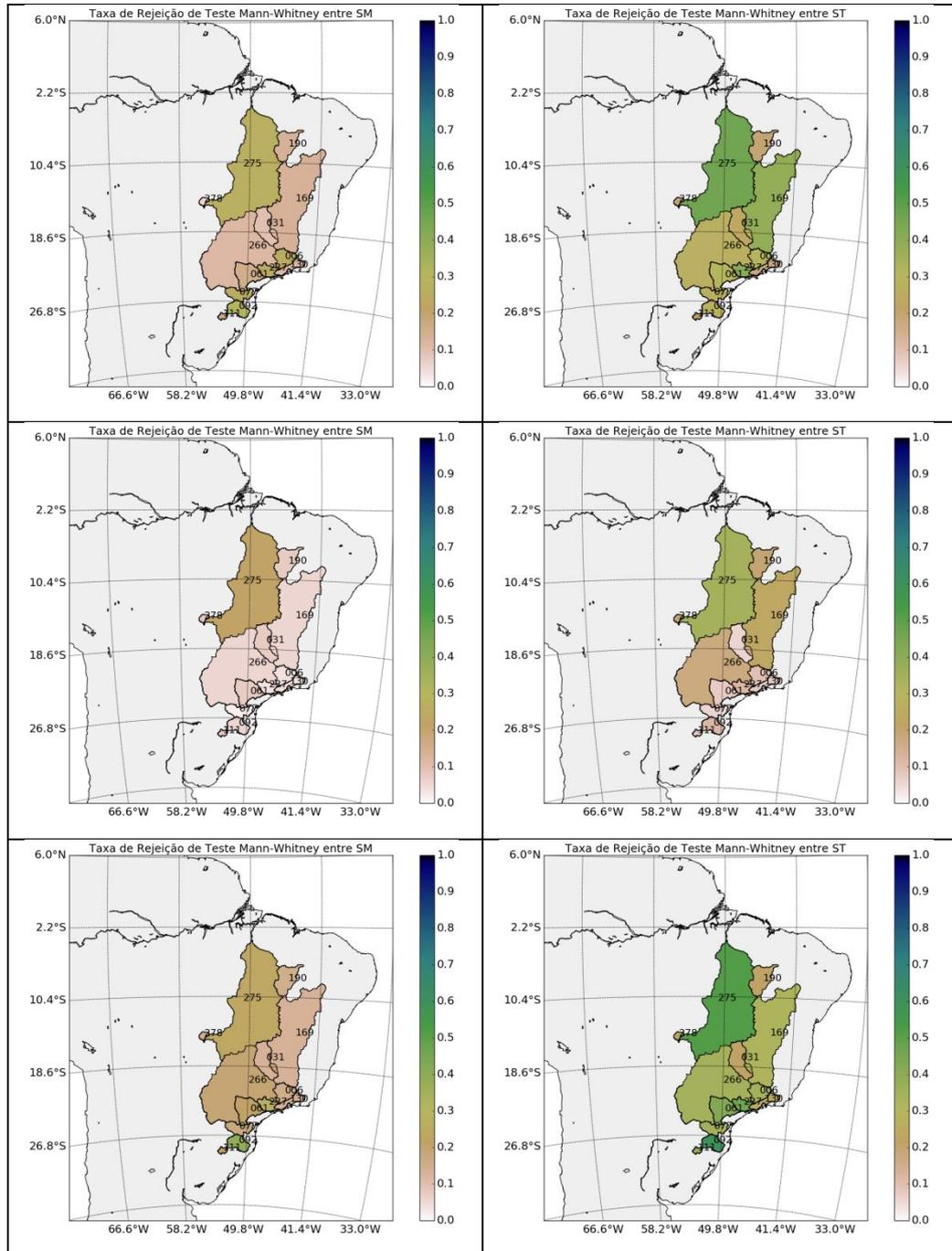
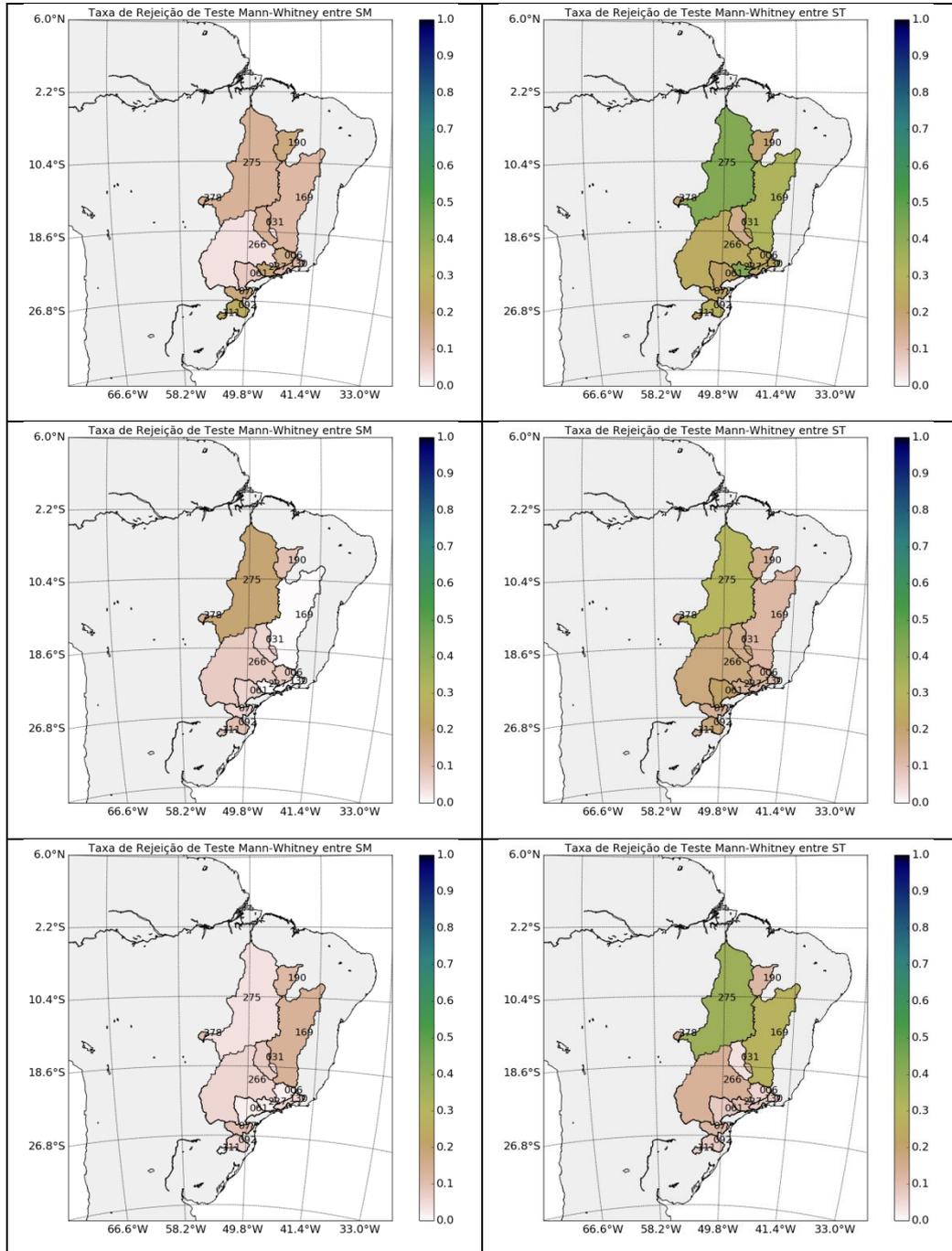
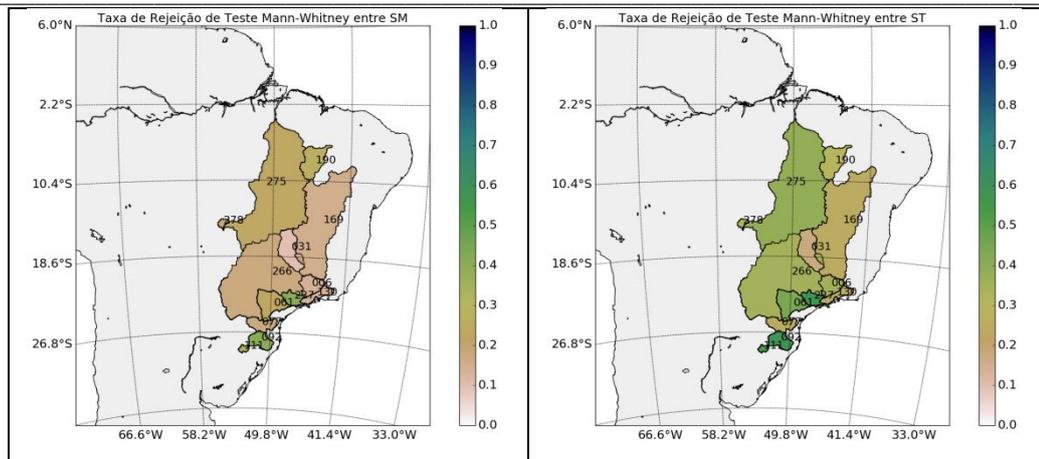


Figura 10 - Percentual de *fdas* diferentes (nível de significância de 95%) da *fdas* semanal original nos postos considerados - divisão em 4 faixas (esquerda: agregação em períodos mensais; direita: períodos trimestrais)





A partir das Figuras acima, e conforme resultados apresentados nas tabelas presentes no Anexo 1, pode-se afirmar que:

a) não considerando a divisão do histórico em faixas de vazão, para maior parte dos casos (10 de 13), a agregação em períodos mensais leva, em média, a diferenças na *fda* em menos de 8% das semanas e em todos os postos analisados, conforme pode ser observado na Tabela A.1; em dois postos da região Nordeste (Sobradinho e Boa Esperança), em média, 15% das semanas do ano apresentam *fda's* diferentes das funções construídas considerando agregação mensal. Para o posto de Tucuruí, localizado na região Norte, 33% das semanas do ano apresentaram distribuições de variações de vazões diferentes quando considerada a agregação mensal. Assim, pode-se afirmar que, para a maioria dos postos, ao se adotar a agregação mensal, os erros cometidos na definição dos limites de previsão não seriam, em média, significativos. Por outro lado, não se poderia recomendar o uso da agregação em períodos trimestrais para os postos de Sobradinho, Boa Esperança e, especialmente, Tucuruí.

b) a partir da divisão do histórico em três e quatro faixas, as maiores diferenças na *fda* resultante de agregação (mensal ou trimestral) são observadas nas "caudas" da distribuição, conforme fora observado no item 4.1 para os postos de Tucuruí e Salto Santiago. Além disso, conforme verificado anteriormente, em média, a agregação em períodos mensais leva a diferenças significativas da *fda* em menos de 10% dos casos para a maioria dos postos, ao se considerar a faixa central (no caso da divisão em 3 faixas, conforme Tabela A.3) e as duas faixas centrais (no caso da divisão em 4 faixas, conforme Tabela A.4); por outro lado, como já destacado, diferenças maiores foram observadas nas faixas correspondentes aos "extremos" da distribuição - em alguns postos, em mais de 20% dos casos, como pode ser observado nas Tabelas A.3 e

A.4. Deste modo, não é evidente o ganho, em termos médios, no uso da agregação em períodos mensais, já que diferenças significativas podem ocorrer em situações nas quais as vazões não se encontrarem na região central da distribuição.

c) com a divisão em duas faixas, observa-se, em média, uma clara tendência à ocorrência de maiores diferenças entre as *fda's* semanais e as obtidas por agregação em períodos mensais, com pelo menos 10% dos casos apresentando distribuições significativamente diferentes, conforme pode ser observado na Tabela A.2. Este comportamento mereceria ser objeto de uma análise mais aprofundada.

d) com respeito à agregação em períodos mensais, na grande maioria dos casos a agregação em trimestres levou a um aumento da diferença entre a *fda* resultante e a função original, conforme pode ser verificado nos casos de melhoria destacados (em amarelo) nas Tabelas A.1 a A.4. Além disso, a agregação em períodos trimestrais, na grande maioria dos casos, acarretou diferenças na *fda* em pelo menos 10% dos casos, sobretudo quando se adotara a divisão em faixas de vazão. Sem essa divisão, os postos de Itá, Passo Real e Ilha dos Pombos foram os únicos nos quais, em média, a agregação em trimestres levou a diferenças em menos de 10% das semanas (6,6 e 8%, respectivamente, conforme resultados apresentados na Tabela A.1 no Anexo).

Deste modo, de acordo com a análise dos resultados exposta anteriormente, na ausência de estudos específicos e mais detalhados para cada posto e suas respectivas séries de vazões com suas particularidades, como diferentes estações do ano, é possível aconselhar o uso da agregação em períodos mensais apenas se não se dividir o histórico em faixas de vazão. Para os postos de Sobradinho, Boa Esperança e, sobretudo, Tucuruí, mesmo nesta situação, deve-se evitar a agregação em períodos mensais. Considerando-se a divisão do histórico em faixas de vazão, estudos mais específicos deveriam ser empreendidos, já que os resultados aqui apresentados não se mostraram conclusivos, apesar do nítido ganho ao se considerar as *fda's* obtidas para vazões na região central da distribuição.

Considerando-se que neste Relatório foram apresentados e discutidos resultados referentes à média nas 52 semanas operativas, os resultados obtidos para a agregação em trimestres e considerando a divisão em três faixas de vazão, critérios usualmente adotados para o uso dos limites de previsão no Programa Mensal de Operação (PMO) e suas revisões, mostram que grandes inconsistências podem ser encontradas, para a grande maioria dos postos. Como pode

ser observado na Tabela A.3, apenas 5 postos apresentaram diferenças entre a *fda* resultante de agregação e a original (isto é, considerando apenas a semana em questão) em menos de 10% (em faixas específicas) das semanas. Naturalmente, de acordo com os resultados aqui apresentados, não se poderia recomendar o uso de tais critérios para a definição dos limites de previsão. Por outro lado, o estudo empreendido concentrou-se em avaliar as diferenças médias em 52 semanas, sendo possível, portanto, que, para semanas (ou meses) específicas, os critérios adotados no PMO não venham a causar erros significativos na definição dos limites de previsão, como no caso apresentado no Item 4. Outrossim, o estudo se concentrou na análise da diferença entre as funções de distribuição acumulada de probabilidades, não levando em conta o impacto desta na definição dos limites de previsão propriamente ditos, e em que medida uma função "similar", segundo uma medida de similaridade, poderia levar à definição de limites idênticos ou similares. Consequentemente, os resultados aqui discutidos não devem ser tomados como conclusivos, mas indicativos de novos estudos a serem efetuados.

7 CONCLUSÕES E AÇÕES PROPOSTAS

De acordo com os resultados apresentados anteriormente, verificou-se que a agregação em períodos mensais leva a diferenças significativas (conforme o nível de significância e o teste adotados neste estudo) na maior parte dos casos analisados, com exceção das situações em que as funções de distribuição foram construídas sem considerar a divisão do histórico de vazões em faixas. Assim, apenas sem a divisão em faixas a agregação em períodos mensais poderia ser recomendada, com a exceção dos postos de Sobradinho, Boa Esperança e Tucuruí. Segundo os resultados aqui apresentados, a agregação em períodos trimestrais não deveria ser adotada para maior parte dos postos, já que levou a diferenças significativas na *fda* resultante na grande maioria dos casos aqui analisados (para apenas três postos menos de 10%, em média, das funções foram consideradas significativamente diferentes em casos sem a divisão em faixas). Mais especificamente, a agregação em períodos trimestrais considerando três faixas de vazão, critérios utilizados para o uso dos limites nas previsões nos Programas Mensais de Operação (PMO), não poderia ser recomendada de acordo com os resultados obtidos neste estudo. Não se analisou a agregação em períodos semestrais, mas pode-se especular que levaria a resultados ainda piores do que os obtidos segundo a agregação em trimestres.

Conforme adotado pelo PREVIVAZ, os períodos trimestrais e semestrais são definidos de acordo com o ano civil, e não o ano hidrológico, adotado para a estimação das funções de autocorrelação amostral [1]. Embora, a princípio, não sejam esperadas mudanças significativas ao se considerar o ano hidrológico para a avaliação da agregação, um estudo posterior poderia avaliar esta premissa, sobretudo no que diz respeito à agregação em períodos mensais. Além disso, poder-se-ia avaliar a agregação em períodos "centrados" na semana de previsão, o que reduziria, a princípio, o risco assumido ao se misturar períodos com características possivelmente diversas para a definição da *fda*.

A análise aqui empreendida não levou em conta o impacto das diferenças na estimação da função de distribuição acumulada de probabilidades na previsão semanal; ainda que erros grandes possam ser cometidos, como demonstrado no estudo apresentado em [3], é possível que, na maior parte dos casos, os erros cometidos não influenciem de forma decisiva os resultados do modelo. Esta avaliação será objeto de estudo futuro.

Outro possível desdobramento do estudo aqui apresentado seria a avaliação da melhor configuração (tipo de agregação, número de faixas) para cada usina, em cada semana operativa (ou mês ou estação), já que aqui os resultados foram considerados em média (ou seja, resultados médios ao longo do período considerado a cada posto como apresentado na Tabela 3). Além disso, esse estudo poderia ser feito ao se considerar (i) a estimação da fda das variações de vazão, e (ii) o impacto das diferenças na previsão de vazões semanais.

8 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Manual de Metodologia do Modelo PREVIVAZ, versão 6.1 (Novembro/2016)
- [2] Relatório Técnico DEA/CEPEL 21418/09 "Aplicação de Limites nas Previsões de Afluências Semanais no Modelo PREVIVAZ", 2009.
- [3] Nota Técnica PREVIVAZ 01/2016 "Análise das previsões e respectivos intervalos de confiança teóricos calculados em estudo preliminar ao PMO de Abril/2016".
- [4] Manual do Usuário do Modelo PREVIVAZ, versão 6.1 (Novembro/2016)
- [5] Helsel, D.R., Hirsch, R.M., "Statistical Methods in Water Resources", In: USGS, **Hydrology Analysis and Interpretation**, USGS, 2002.

9 ANEXOS

Tabela A.1 - Percentual de casos em que a função de distribuição acumulada de probabilidade (*fda*) diferiu (nível de significância: 95%) da *fda* semanal - **sem divisão em faixas de vazão** (em destaque: casos em que a agregação em trimestres levou a uma maior similaridade com as *fda*'s semanais)

UHE	Mês	Trimestre
Tucuruí	0,33	0,50
Boa Esperança	0,15	0,13
Sobradinho	0,15	0,50
Manso	0,04	0,13
Itumbiara	0,08	0,13
Furnas	0,06	0,29
Ilha dos pombos	0,02	0,08
Barra Bonita	0,04	0,15
Capivara	0,02	0,13
Itaipu	0,04	0,29
Salto Santiago	0,02	0,15
Itá	0,04	0,06
Passo Real	0,00	0,06

Tabela A.2 - Percentual de casos em que a função de distribuição acumulada de probabilidade (*fda*) diferiu (nível de significância: 95%) da *fda* semanal - **com divisão em 2 faixas de vazão** (em destaque: casos em que a agregação em trimestres levou a uma maior similaridade com as *fda*'s semanais)

UHE	Mês		Trimestre	
	Faixa 1	Faixa 2	Faixa 1	Faixa 2
Tucuruí	0,27	0,17	0,50	0,48
Boa Esperança	0,17	0,19	0,25	0,17
Sobradinho	0,12	0,17	0,31	0,37
Manso	0,12	0,13	0,17	0,25
Itumbiara	0,17	0,10	0,25	0,17
Furnas	0,15	0,12	0,29	0,31
Ilha dos pombos	0,08	0,15	0,17	0,19
Barra Bonita	0,21	0,23	0,38	0,33
Capivara	0,15	0,15	0,35	0,35
Itaipu	0,13	0,15	0,31	0,29
Salto Santiago	0,15	0,17	0,33	0,38
Itá	0,29	0,27	0,38	0,35
Passo Real	0,13	0,08	0,23	0,19

Tabela A.3 - Percentual de casos em que a função de distribuição acumulada de probabilidade (*fda*) diferiu (nível de significância: 95%) da *fda* semanal - **com divisão em 3 faixas de vazão** (em destaque: casos em que a agregação em trimestres levou a uma maior similaridade com as *fda*'s semanais)

UHE	Mês			Trimestre		
	Faixa 1	Faixa 2	Faixa 3	Faixa 1	Faixa 2	Faixa 3
Tucuruí	0,27	0,21	0,23	0,46	0,35	0,52
Boa Esperança	0,13	0,08	0,15	0,19	0,19	0,21
Sobradinho	0,13	0,06	0,13	0,38	0,23	0,33
Manso	0,06	0,08	0,15	0,19	0,19	0,25
Itumbiara	0,10	0,08	0,13	0,23	0,06	0,21
Furnas	0,25	0,04	0,15	0,31	0,10	0,35
Ilha dos pombos	0,12	0,04	0,13	0,17	0,08	0,25
Barra Bonita	0,25	0,04	0,27	0,40	0,12	0,46
Capivara	0,15	0,08	0,21	0,31	0,08	0,42
Itaipu	0,12	0,06	0,19	0,27	0,17	0,35
Salto Santiago	0,25	0,02	0,17	0,27	0,06	0,35
Itá	0,31	0,06	0,38	0,29	0,12	0,60
Passo Real	0,15	0,08	0,21	0,21	0,08	0,37

Tabela A.4 - Percentual de casos em que a função de distribuição acumulada de probabilidade (*fda*) diferiu (nível de significância: 95%) da *fda* semanal - **com divisão em 4 faixas de vazão** (em destaque: casos em que a agregação em trimestres levou a uma maior similaridade com as *fda*'s semanais)

UHE	Mês				Trimestre			
	Faixa 1	Faixa 2	Faixa 3	Faixa 4	Faixa 1	Faixa 2	Faixa 3	Faixa 4
Tucuruí	0,13	0,19	0,04	0,23	0,42	0,31	0,37	0,38
Boa Esperança	0,17	0,10	0,12	0,27	0,19	0,13	0,12	0,31
Sobradinho	0,12	0,00	0,13	0,15	0,33	0,12	0,29	0,25
Manso	0,13	0,13	0,10	0,19	0,19	0,13	0,13	0,29
Itumbiara	0,13	0,06	0,08	0,10	0,15	0,15	0,04	0,17
Furnas	0,15	0,08	0,02	0,13	0,25	0,13	0,06	0,35
Ilha dos pombos	0,12	0,00	0,06	0,13	0,23	0,10	0,08	0,25
Barra Bonita	0,19	0,00	0,02	0,40	0,42	0,13	0,10	0,58
Capivara	0,08	0,06	0,02	0,23	0,21	0,21	0,08	0,44
Itaipu	0,04	0,08	0,06	0,17	0,25	0,17	0,13	0,35
Salto Santiago	0,17	0,06	0,10	0,17	0,23	0,13	0,12	0,25
Itá	0,25	0,10	0,06	0,40	0,29	0,17	0,08	0,58
Passo Real	0,23	0,08	0,00	0,31	0,23	0,12	0,06	0,52