

5

Estudo de Caso

5.1

Introdução

Neste estudo de caso, como meio de avaliar a eficiência do modelo proposto MIP, serão efetuadas previsões baseadas na vazão natural incremental verificada entre as UHE Porto Primavera, Rosana e Itaipu, todas da Bacia do Rio Paraná.

Características físicas da área incremental de Itaipu

A bacia do rio Paraná, no trecho compreendido entre as usinas hidrelétricas de Rosana, Porto Primavera e Itaipu, compreende uma área de drenagem de 150.900 km². A principal contribuição é representada pela confluência na margem esquerda com o rio Paranapanema, situada imediatamente a jusante da UHE Rosana. Outras contribuições significativas correspondem às confluências com os rios Ivaí e Piquiri na margem esquerda e com os rios Ivinheima, Amambaí, e Iguatemi na margem direita.

Este trecho da bacia situa-se entre as latitudes 50°48' e 56°00' sul e longitudes 25°36' e 20°48' oeste, drenando a porção sul do Estado do Mato Grosso do Sul e a porção centro-oeste do Estado do Paraná, além de uma pequena porção do território paraguaio.

Na margem esquerda do rio Paraná, o clima caracteriza-se por verões rigorosos e chuvas bem distribuídas durante todo o ano. Já na margem direita, o clima é caracterizado por invernos secos e verões chuvosos. O período chuvoso tem início entre outubro e novembro e se estende até março. A estação seca tem início em abril e se estende até setembro.

A bacia incremental de Itaipu (representada na figura 5.1) apresenta 68% de sua área de drenagem controlada por cinco postos fluviométricos – Ivinheima, Florida, Balsa Santa Maria, Novo Porto Taquara e Rio dos Patos (há um erro no mapa da figura 1, pois ele mostra o posto pluviométrico de Estrada Iguatemi ao invés do posto fluviométrico de Rio dos Patos). Os restantes 32% da área correspondem ao vale do rio Paraná e às vizinhanças do lago de Itaipu.

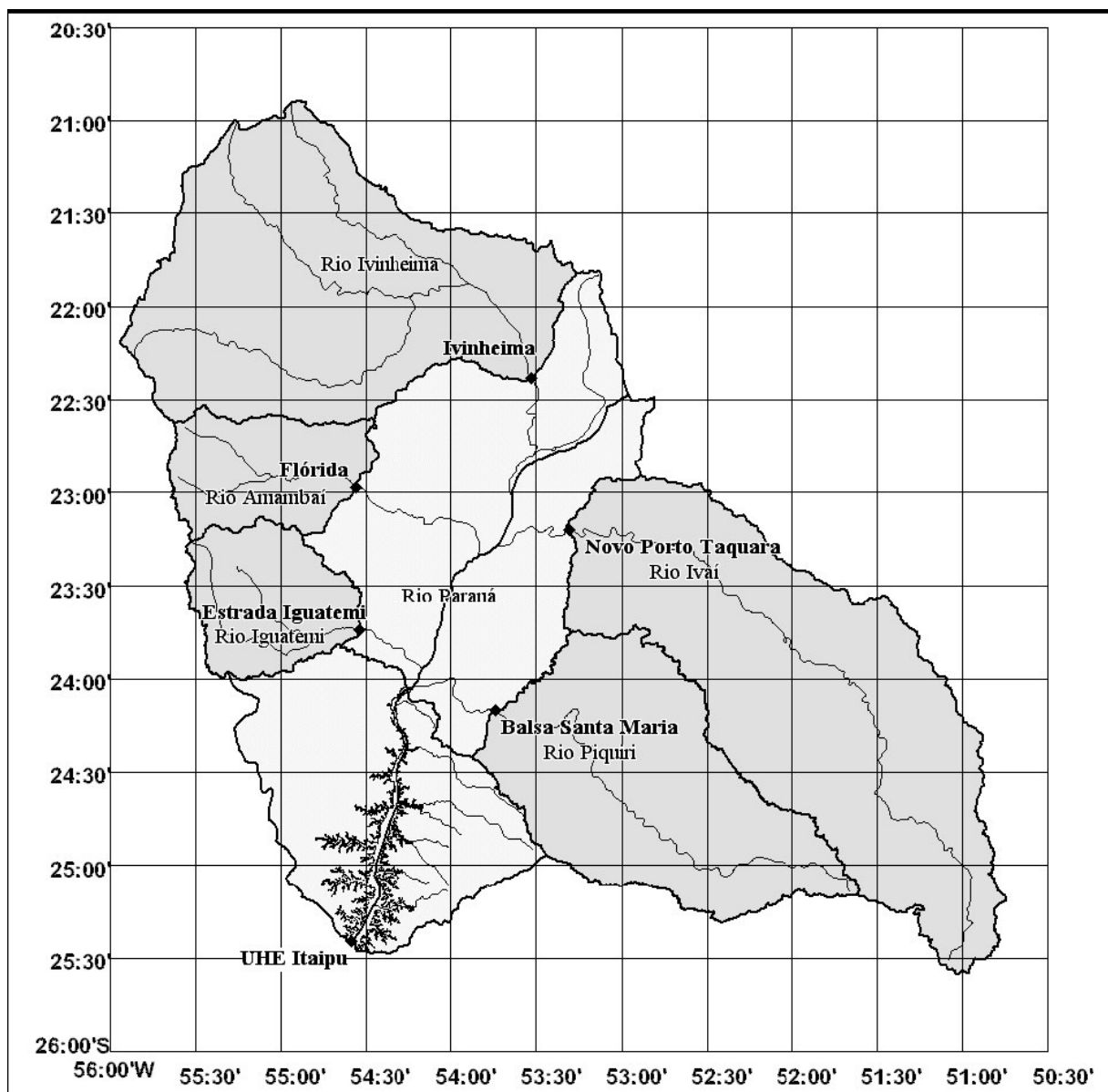


Figura 5.1: Mapa da bacia incremental de Itaipu e sub-bacias

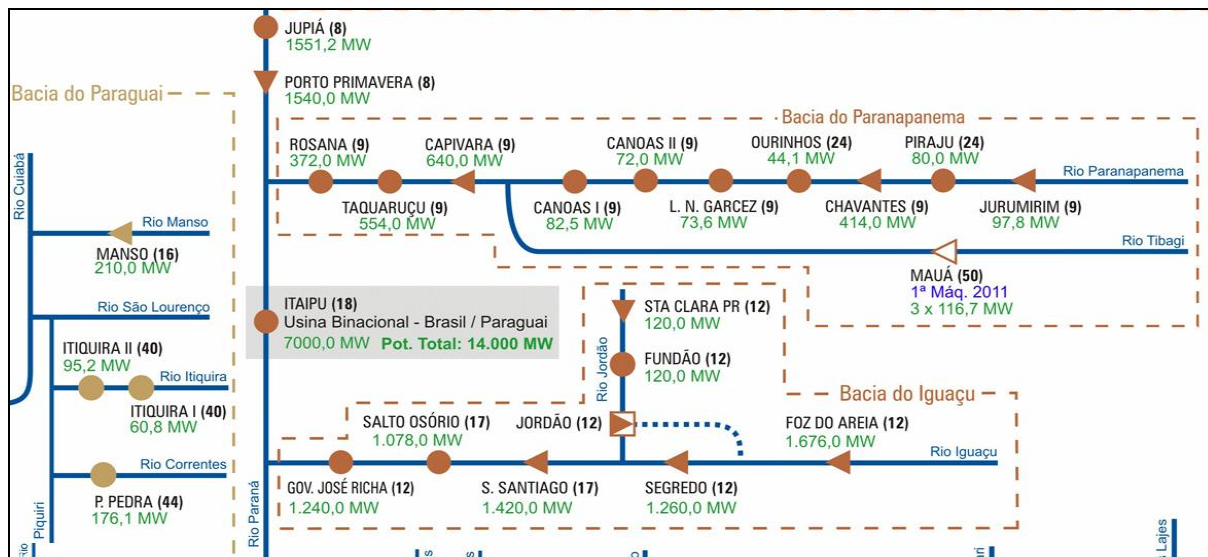


Figura 5.2: Aproveitamentos hidrelétricos onde aparecem as UHE Rosana, Porto Primavera e Itaipu (fonte ONS)

Vazões a serem previstas

As vazões a serem previstas serão as 15 vazões já enunciadas no capítulo 4, ou seja:

- 3 previsões de vazão natural média semanal – entende-se como vazão natural média semanal, uma média das vazões naturais incrementais médias diárias, calculada de acordo com 3 períodos distintos de 7 dias cada, a saber:
 - do 4º ao 10º dia (1ºsábado a 2ªsexta);
 - do 5º ao 11º dia (1ºdomingo ao 2ºsábado);
 - do 6º ao 12º dia (1ªsegunda ao 2ºdomingo).
- Previsões de vazões naturais incrementais médias diárias do 1º ao 12º dia.

Primeiramente, os dados, ou séries históricas, serão arrumados e tratados pelo módulo 1, para que dêem entrada nos módulos 2 e 3. No intuito de chamar a atenção para a melhoria que o módulo 2 traz para o MIP, optou-se por primeiro proceder à previsão das 15 vazões sem que os dados passassem pelo módulo 2 (seção 5.3), ou seja, sem que os dados sofressem a transformação wavelet, indo diretamente do módulo 1 para o módulo 3. Depois, então, na subseção 5.4, dá-se prosseguimento ao fluxo direto do MIP, fazendo com que os dados, após serem tratados no módulo 1, sejam transformados dentro do módulo 2 para depois

constituírem as entradas do módulo 3, onde será feita a previsão MIP propriamente dita.

5.2

Módulo 1 – Tratamento dos Dados

5.2.1

Séries Históricas Disponíveis

Há disponíveis os seguintes conjuntos de dados ou séries históricas para a efetivação das previsões de vazões da Bacia do Rio Paraná – trecho Itaipu:

- Séries históricas de vazões naturais médias diárias incrementais entre as UHE Porto Primavera, Rosana e Itaipu, de 1955 até 2001;
- Séries históricas de pluviometria de 37 postos pluviométricos, razoavelmente bem distribuídos, em ambas as vertentes da bacia, com maior concentração na área próxima ao reservatório de Itaipu. A tabela a seguir apresenta, para cada posto pluviométrico, o tamanho de cada série (definido pelo número de dias decorridos entre o primeiro e o último dia da série), o período histórico coberto por cada série e o percentual de valores ausentes em cada série.

Tabela 5.1: Postos pluviométricos e respectivos percentuais de falhas

#	Identificação do Posto Pluviométrico	#Dias	Período	Percentua
1	02253000 – Ivinhema (PCD INPS)	9904	1974-2001	36,0
2	02253016- Porto São Jose (PCD INPS)	5145	1987-2001	1,8
3	02352000 - Porto Paraíso do Norte (PCD	17824	1953-2001	5,7
4	02354001- Estrada Iguatemi	10823	1972-2001	23,9
5	02354002 – Flórida	10662	1972-2001	17,1
6	02453000 - Balsa do Santa Maria (PCD	10958	1969-2001	1,6
7	02453043 - Novo Porto 2 (PCD INPS)	8740	1978-2001	3,4
8	02454000 - Porto Mendes Gonçalves-Itaipu	22280	1941-2001	0,2
9	02454031 - São Francisco Falso (2454026)	4049	1990-2001	22,5
10	02454032 - ACampamento Central	7519	1981-2001	0,0
11	02454035 - Saltos del Guaíra	3287	1993-2001	8,3
12	02454038 - Corpus Christi	2832	1994-2001	0,0
13	02454039 - Ibel (Itaipu)	2863	1994-2001	1,1
14	02454040 - La Paloma	2802	1994-2001	1,1
15	02554037 - Reserva Itabo	6515	1984-2001	3,5
16	02554040 - Acuicultura	8401	1979-2001	1,5
17	02554042 - Paulistânia	6575	1984-2001	9,2
18	02155000 - Maracajú	10666	1972-2001	23,5
19	02352002 - Quinta do Sol	10015	1974-2001	0,3

#	Identificação do Posto Pluviométrico	#Dias	Período	Percentua
20	02451017 - Porto Espanhol	13302	1965-2001	0,9
21	02451020 - Barbosa Ferraz	10015	1974-2001	1,4
22	02452000 - Balsa do Cantu	12580	1967-2001	1,3
23	02452010 - Janiópolis	12664	1967-2001	0,7
24	02453012 - Corbélia	13394	1965-2001	0,7
25	02550000 - Rio dos Patos	23603	1937-2001	1,1
26	02451010 - Santa Maria	12694	1967-2001	0,7
27	02255004 - Itaum	6467	1984-2001	32,6
28	02254005 - Itaporá	6519	1984-2001	8,9
29	02255003 - Bocaia	6465	1984-2001	11,8
30	02254004 - Porto Wilma	6551	1984-2001	16,7
31	02253015 - Fazenda Jangada	6494	1984-2001	27,2
32	02353003 - Altonia	12694	1967-2001	0,1
33	02353005 - Xambre	12784	1967-2001	1,7
34	02353002 - Balsa do Goio-ere	14551	1962-2001	8,8
35	02453064 - Cascavel-Simepar	1675	1997-2001	4,4
36	02453065 - Toledo-Simepar	1605	1997-2001	5,0
37	02454030 - Santa Helena-Simepar	1625	1997-2001	2,5

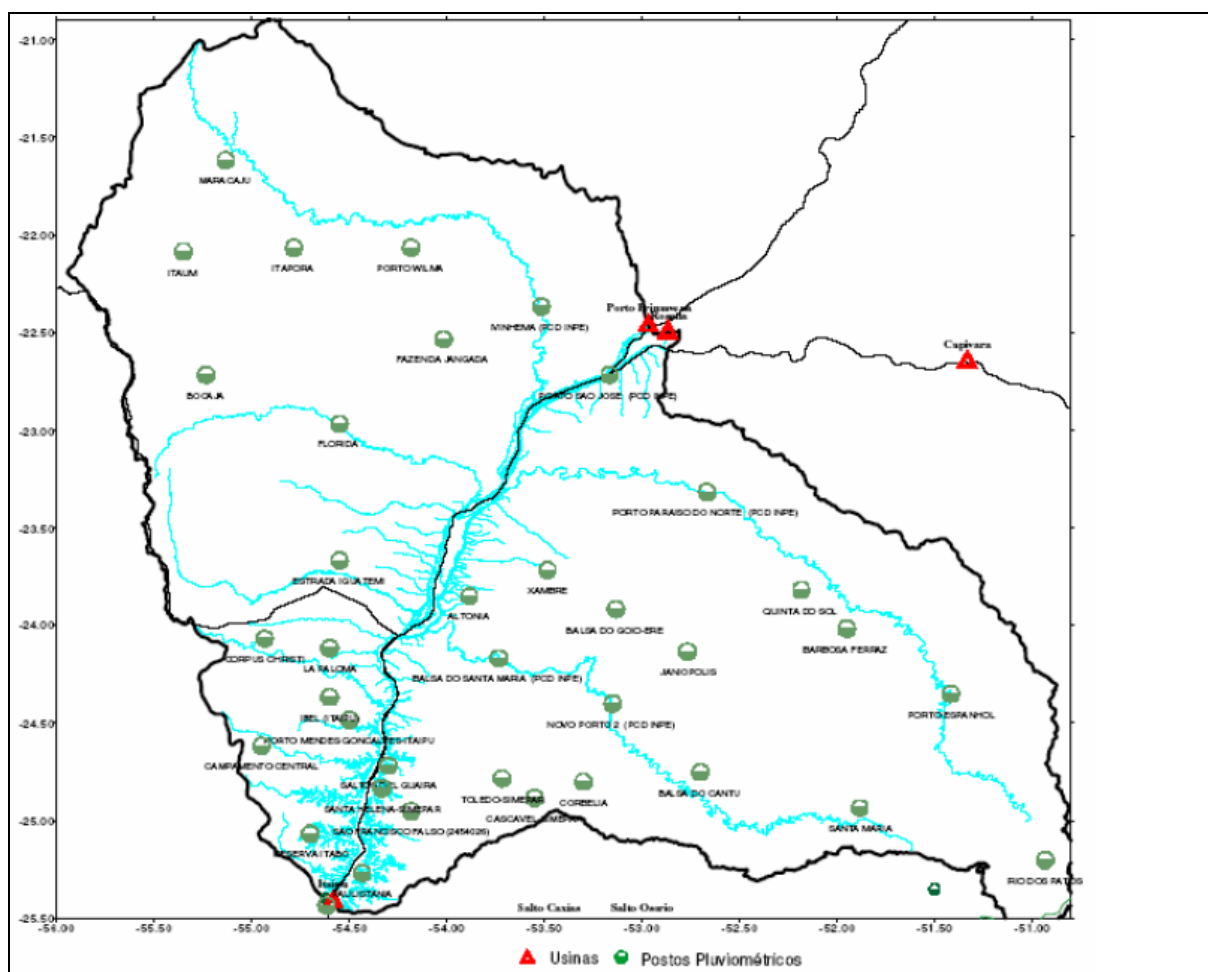


Figura 5.3: Localização dos 37 postos pluviométricos [88]

- Séries históricas de fluviometria de 5 postos pluviométricos. As séries cobrem os principais afluentes do rio Paraná e não apresentam falhas.

Tabela 5.2: Tabela com os dados dos 5 postos pluviométricos da Bacia do Rio Paraná – trecho Itaipu

	Identificação do Posto Pluviométrico	#Dias Série	Período histórico	Percentual falhas %
1	64617000 - Ivinheima - PCD	10767	1972-2001	0,0
2	64620000 - Rio dos Patos	25933	1938-2001	0,0
3	64693000 - Novo Porto Taquara	10029	1974-2001	0,0
4	64715001 - Flórida	10776	1972-2001	0,0
5	64830000 - Balsa Santa Maria - PCD	11796	1969-2001	0,0

Séries históricas de chuvas ETA – o modelo ETA recentemente adquirido pelo ONS realiza previsões do total de chuva acumulada em 24 horas, realizadas semanalmente, às quartas-feiras, para um horizonte de 10 dias, em formato de grade, com resolução espacial de 40 km x 40 km, para os anos de 1996 a 2001.

Tendo em vista que esta dissertação tem como um de seus focos levantar a importância de cada série histórica – principalmente das chuvas ETA – na previsão das vazões, e sendo que elas precisam ter exatamente o mesmo tamanho quando da entrada do módulo 3 (redes neurais), todas as séries serão reduzidas ao tamanho do período histórico das chuvas ETA (pelo fato deste ser o menor período).

5.2.2

Cálculo da chuva de Thiessen para a Bacia e Complemento dos Dados

Conforme já enunciado na seção 4.2, ao invés de se utilizar diretamente as chuvas diárias observadas em cada um dos 37 postos pluviométricos da Bacia do Rio Paraná – trecho Itaipu, o que resultaria em 37 séries históricas de chuvas, será calculada para cada dia a chuva média diária da bacia ou chuva de Thiessen.

A figura 5.1 mostra as 37 áreas de domínio relativas a cada posto pluviométrico.

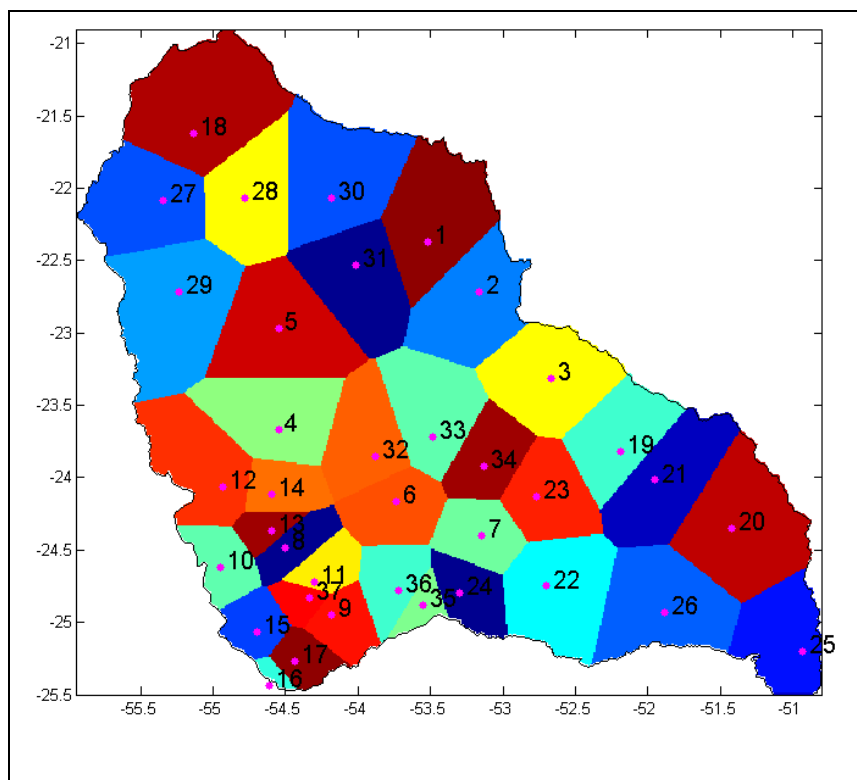


Figura 5.4: Os 37 postos pluviométricos e suas relativas áreas de domínio [88]

Os pesos de Thiessen, como já foi dito, advêm da relação *área de domínio/área da bacia*, podendo ser visualizados na tabela abaixo, juntamente com os dados de longitude e latitude de cada posto.

Tabela 5.3: Tabela de pesos de thiessen relativos aos 37 postos pluviométricos

#	Posto	Lat	Long	Peso
1	02253000 - Ivinhema (PCD INPS)	-22.3667	-53.5167	0.0426
2	02253016- Porto São Jose (PCD INPS)	-22.7167	-53.1667	0.0376
3	02352000 - Porto Paraíso do Norte (PCD	-23.3167	-52.6667	0.0394
4	02354001- Estrada Iguatemi	-23.6667	-54.5500	0.0358
5	02354002 – Flórida	-22.9667	-54.5500	0.0451
6	02453000 - Balsa do Santa Maria (PCD	-24.1667	-53.7333	0.0226
7	02453043 - Novo Porto 2 (PCD INPS)	-24.4000	-53.1500	0.0181
8	02454000 - Porto Mendes Gonçalves-Itaipu	-24.4833	-54.5000	0.0114
9	02454031 - São Francisco Falso (2454026)	-24.9500	-54.1833	0.0131
10	02454032 – ACampamento Central	-24.6167	-54.9500	0.0143
11	02454035 - Saltos del Guaíra	-24.7167	-54.3000	0.0104
12	02454038 - Corpus Christi	-24.0667	-54.9333	0.0341
13	02454039 - Ibel (Itaipu)	-24.3667	-54.6000	0.0069
14	02454040 - La Paloma	-24.1167	-54.6000	0.0139
15	02554037 - Reserva Itabo	-25.0667	-54.7000	0.0119
16	02554040 – Acuicultura	-25.4333	-54.6167	0.0021
17	02554042 – Paulistânia	-25.2667	-54.4333	0.0098
18	02155000 – Maracajú	-21.6167	-55.1333	0.0564
19	02352002 - Quinta do Sol	-23.8167	-52.1833	0.0288
20	02451017 - Porto Espanhol	-24.3500	-51.4167	0.0492

#	Posto	Lat	Long	Peso
21	02451020 - Barbosa Ferraz	-24.0167	-51.9500	0.0366
22	02452000 - Balsa do Cantu	-24.7500	-52.7000	0.0386
23	02452010 - Janiopolis	-24.1333	-52.7667	0.0254
24	02453012 - Corbélia	-24.8000	-53.3000	0.0155
25	02550000 - Rio dos Patos	-25.2000	-50.9333	0.0285
26	02451010 - Santa Maria	-24.9333	-51.8833	0.0415
27	02255004 - Itaum	-22.0833	-55.3500	0.0363
28	02254005 - Itaporá	-22.0667	-54.7833	0.0338
29	02255003 - Bocaja	-22.7167	-55.2333	0.0521
30	02254004 - Porto Wilma	-22.0667	-54.1833	0.0443
31	02253015 - Fazenda Jangada	-22.5333	-54.0167	0.0359
32	02353003 - Altonia	-23.8500	-53.8833	0.0276
33	02353005 - Xambre	-23.7167	-53.4833	0.0325
34	02353002 - Balsa do Goio-ere	-23.9167	-53.1333	0.02
35	02453064 - Cascavel-Simepar	-24.8833	-53.5500	0.0059
36	02453065 - Toledo-Simepar	-24.7833	-53.7167	0.0161
37	02454030 - Santa Helena-Simepar	-24.8333	-54.3333	0.006

Nota-se no gráfico *peso de Thiessen - posto pluviométrico* da figura a seguir, que os postos de maior área (por exemplo: postos 18, 20 e 29) apresentam maiores pesos.

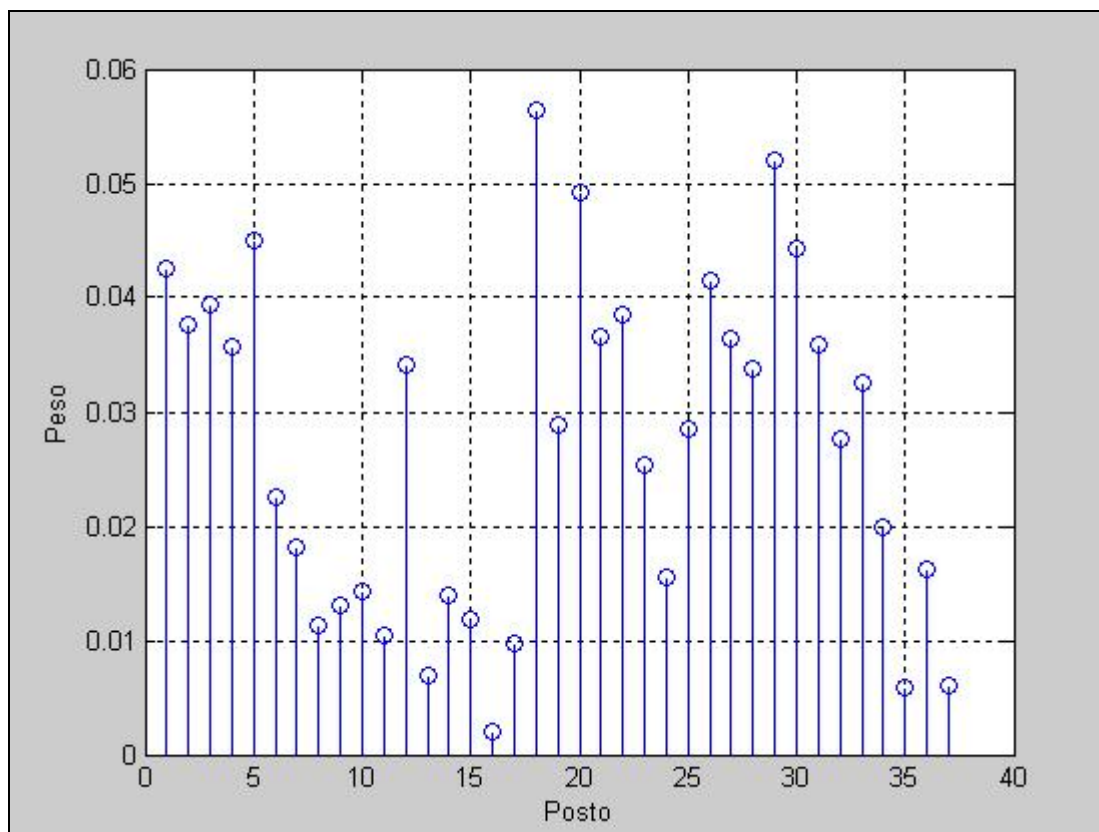


Figura 5.5: Pesos de Thiessen dos 37 postos pluviométricos

Contudo, pelo fato das séries diárias dos postos pluviométricos apresentarem falhas, é necessário preencher os dados faltantes. O método escolhido é aquele descrito no exemplo da subseção 4.2b, que altera os pesos de Thiessen no dia, para os postos que não apresentam falta de registro de chuva.

Para se ter uma idéia da mudança dos pesos, a tabela abaixo mostra os primeiros 20 dias da matriz de chuva média de Thiessen original da bacia e da matriz de chuva média de Thiessen ajustada.

Tabela 5.4: Ajuste da matriz de chuvas médias de Thiessen

Dia	matriz de chuva média de Thiessen original	matriz de chuva média de Thiessen ajustada
1	0,20615	0,233544806
2	1,32151	1,497122465
3	0,37234	0,421819418
4	5,43462	6,15681432
5	10,32953	11,7021978
6	11,01495	12,47870171
7	2,85713	3,236807522
8	1,43091	1,621060383
9	1,57228	1,781216721
10	3,48224	3,944986972
11	0,58394	0,661538462
12	2,78328	3,153143763
13	8,88424	10,06484649
14	12,68453	14,37014841
15	14,99412	16,98665458
16	11,95141	13,53960576
17	12,10793	13,71692534
18	17,25815	19,55154639
19	9,05416	10,25734678
20	3,45856	3,91816019

Ressalta-se que as séries históricas fluviométricas da Bacia do Rio Paraná – trecho Itaipu apresentam-se completas, não necessitando, pois, de nenhum método de preenchimento de dados.

5.2.3

Confecção das Matrizes de Dados e Seleção de Variáveis

Esta subseção se subdivide em 3 etapas. A 1ª etapa corresponde à formação das matrizes (todas de mesmo tamanho) de séries temporais, tendo como base as séries temporais já discriminadas na subseção 5.2.1, que são as que realmente serão analisadas pelos algoritmos computacionais de implementação do método LSE e de Redes Neurais. Para tal, é necessário definir as datas de início e término de cada uma dessas séries, ou melhor, definir as datas de início e término de cada uma dessas matrizes temporais. A tabela abaixo discrimina cada uma dessas matrizes, bem como seus correspondentes períodos históricos.

Tabela 5.5: Tabela de formação das matrizes de dados

	Matriz temporal	Início	Término	Tamanho
1	Matriz Th - Matriz com 7 últimas chuvas Thiessen – de quarta (7 dias atrás) a terça (1 dia atrás)	De 3/1/1996 Chuva de quarta	Até 18/12/2001 Chuva de terça	7 x 311
2	Matriz ETA - Matriz com as 10 chuvas previstas pelo modelo ETA – de quarta (dia atual) a sexta (9 dias à frente)	De 10/1/1996 Chuva prevista de quarta	Até 28/12/2001 Chuva prevista de sexta	10 x 311
3	Matriz Flu - Matriz com a última vazão de cada um dos 5 postos fluviais (vazão de 1 dia atrás - terça)	De 9/1/1996 Vazão de terça	Até 18/12/2001 Vazão de terça	5 x 311
4	Matriz Vinc - Matriz com as últimas vazões naturais incrementais médias diárias (vazão de 1 dia atrás - terça)	De 9/1/1996 Vazão de terça	Até 18/12/2001 Vazão de terça	1 x 311
5	Matriz com as 12 vazões naturais incrementais médias diárias a serem previstas: de quarta (dia atual) ao 2ºdomingo (11 dias à frente)	De 10/1/1996 Vazão de quarta	Até 30/12/2001 Vazão de domingo	12 x 311
6	Matriz Vms1 - Matriz com a vazão natural incremental média calculada do 4º (sábado) ao 10º dia (2ªsexta) de previsão	Da 1ª média – de 13 a 19/1/1996	Até 311ª média – de 22 a 28/12/1996	1 x 311
7	Matriz Vms2 - Matriz com a vazão natural incremental média calculada do 5º (domingo) ao 11º dia (2ºsábado) de previsão	Da 1ª média – de 14 a 20/1/1996	Até 311ª média – de 23 a 29/12/1996	1 x 311
8	Matriz Vms3 - Matriz com a vazão natural incremental média calculada do 6º (segunda) ao 10º dia (2ºdomingo) de previsão – Vms3	Da 1ª média – de 15 a 21/1/1996	Até 311ª média – de 24 a 30/12/1996	1 x 311

Cabe lembrar, que o primeiro dia de previsão, ou melhor dizendo, o 1º dia a ser previsto, é sempre uma **quarta-feira**, estando-se na manhã desta quarta-feira, e tendo, então, como dados mais recentes, os dados do dia anterior, terça-feira. Isso justifica, por exemplo, porque na matriz temporal 1 está estabelecido: “quarta (7 dias atrás) a terça (1 dia atrás)”. A escolha de quarta como o dia atual, justifica-se pelo fato da ONS preparar suas previsões sempre nas manhãs de quarta-feira.

Ressalta-se que cada uma das séries históricas que compõem cada uma das matrizes será tratada como uma variável. Por exemplo, a variável 2ª chuva prevista ETA ou a variável vazão do posto fluviométrico nº 64620000 – Rio dos Patos. Visto isso, deve-se lembrar também que cada uma das 15 variáveis das matrizes 5, 6, 7 e 8 são as variáveis de saída, ou seja, as variáveis que o modelo quer prever. Para simplificar, chamar-se-á as 12 vazões naturais incrementais médias diárias a serem previstas de **Vmd1, Vmd2, ... , Vmd12**.

A 2ª etapa corresponde à seleção de variáveis pelo algoritmo LSE dentro de cada matriz temporal. Como são 15 as vazões a serem previstas, o algoritmo do LSE será rodado 15 vezes dentro de cada uma das matrizes 1, 2 e 3 da tabela 5.5 – a matriz 4 só tem uma única variável – no intuito de definir quais variáveis da matriz mais influenciam cada uma dessas 15 vazões. As variáveis selecionadas dentro de cada matriz, para cada uma das 15 rodadas do LSE, serão aquelas cujo grau de importância (ver algoritmo LSE), em relação àquela determinada vazão, for maior que a média dos graus de importância de todas as variáveis da matriz. Os graus de importância das variáveis de cada matriz em relação a cada uma das 15 variáveis de saída podem ser visualizados em 15 tabelas no Apêndice 6, onde cada uma contém 3 gráficos: 3 matrizes de variáveis x 15 variáveis de saída = 45 gráficos.

A tabela 5.6 a seguir mostra as variáveis de cada grupo (ETA, vazões fluviais e chuvas Thiessen) que foram selecionadas pelo algoritmo do LSE em relação as 15 variáveis de saída.

Cabe, antes, uma pequena explanação sobre a nomenclatura que será usada.

- Variáveis ETA: as 10 chuvas previstas pelo modelo ETA serão designadas por ETA_j ($j = 1,2,3, \dots, 10$), onde ETA_1 é a chuva prevista pelo modelo ETA para quarta-feira, ETA_2 é a chuva prevista para quinta-feira, \dots , e ETA_{10} é a chuva prevista para a 2ª sexta. Logo, $ETA_{2,4,1}$ quer dizer que foram selecionadas com

maior grau de importância (em ordem decrescente de grau), respectivamente, a 2ª, a 4ª e a 1ª chuvas ETA;

- Variáveis vazões fluviais: as 5 vazões fluviais registradas pelos 5 postos fluviométricos, 1 dia antes (terça-feira), serão designadas por Flu₁ (posto 64617000 – Ivinheima), Flu₂ (posto 64620000 - Rio dos Patos), Flu₃ (posto 64693000 - Novo Porto Taquara), Flu₄ (posto 64715001 – Florida), Flu₅ (posto 64830000 - Balsa Santa Maria). Logo, Flu_{4,2} quer dizer que foram selecionadas com maior grau de importância (em ordem decrescente de grau), respectivamente, o 4º e o 2º postos fluviais;

- Variáveis chuvas Thiessen: as 7 últimas chuvas Thiessen calculadas para a bacia serão designadas por Th₁ (chuva de quarta), Th₂ (chuva de quinta), Th₃ (chuva de sexta), ... , Th₇ (chuva de terça). Logo, Th_{7,3} quer dizer que foram selecionadas com maior grau de importância (em ordem decrescente de grau), respectivamente, a 7ª e a 3ª chuvas Thiessen.

Tabela 5.6: variáveis selecionadas pelo 1º LSE

Variáveis selecionadas em relação à								
Grupo	Vms1	Vms2	Vms3	Vmd1	Vmd2	Vmd3	Vmd4	Vmd5
ETA	ETA _{2,1,3}	ETA _{2,1}	ETA _{2,4,1,6}	ETA _{4,1,2,10}	ETA _{4,1,10,3}	ETA _{4,10,3,2}	ETA _{3,10}	ETA _{3,10,2,1}
Vazões fluviais	Flu _{2,4}	Flu _{4,2}	Flu _{4,2}	Flu ₂	Flu _{2,1}	Flu ₂	Flu _{2,4}	Flu _{2,4}
Chuvas Thiessen	Th _{7,5,1}	Th _{5,7,3,1}	Th _{3,6,5}	Th _{4,3,6,1}	Th _{6,4,5,3}	Th _{7,6,3,5}	Th _{7,6,3,5}	Th _{7,1,5}
Variáveis selecionadas em relação à								
Grupo	Vmd6	Vmsd7	Vmd8	Vmd9	Vmd10	Vmd11	Vmd12	
ETA	ETA _{2,3,1,10}	ETA _{2,3}	ETA _{2,4,1,1}	ETA _{4,2,1}	ETA _{4,2,6,3,1}	ETA _{6,3,4,2}	ETA _{3,9,6}	
Vazões fluviais	Flu _{2,4}	Flu _{2,4}	Flu _{2,4}	Flu _{4,2}	Flu _{4,2}	Flu _{2,4}	Flu ₂	
Chuvas Thiessen	Th _{7,1,3,5}	Th _{3,5,7,4}	Th _{4,3,5,6}	Th _{4,6,3}	Th _{6,4,1}	Th _{6,2}	Th _{2,3,4,7}	

A 3ª etapa corresponde à seleção de variáveis pelo algoritmo LSE sobre o conjunto de variáveis selecionadas pelo LSE na etapa anterior. Olhando a tabela acima, este 2º LSE selecionará, por exemplo, dentre as 10 variáveis ETA_{1,2,3,10}, Flu_{2,4} e Th_{1,3,5,7}, aquelas que mais influenciam a saída Vmd6. Como são 15 as vazões a serem previstas, o algoritmo do LSE será rodado, novamente, 15 vezes.

Existem 2 diferenças em relação ao 1º LSE. A primeira é que, neste 2º LSE, está participando da seleção também a última Vazão Incremental (vazão de 1 dia atrás – terça-feira). A segunda diferença é a limitação do número de variáveis selecionadas. Diferentemente do 1º LSE, quando todas as variáveis que tinham grau de importância superior à média, eram selecionadas, neste 2º LSE, são selecionadas as 7 variáveis com maior grau de importância em relação a cada uma das 15 variáveis. Esta mudança de critério de seleção deve-se ao fato que ao se tentar empregar no 2º LSE, o mesmo critério usado no 1º, para a maioria das 15 variáveis a serem previstas, foram selecionadas poucas variáveis (máximo de 5) para a composição das entradas das redes.

A tabela abaixo aponta as variáveis que foram selecionadas pelo 2º algoritmo do LSE em relação as 15 variáveis de saída.

Tabela 5.7: Variáveis selecionadas pelo 2º LSE

Variáveis selecionadas em relação à								
Grupo	Vms1	Vms2	Vms3	Vmd1	Vmd2	Vmd3	Vmd4	Vmd5
ETA	ETA _{2,1}	ETA _{2,1,4}	ETA _{2,4,1,6}	ETA ₁	ETA ₁	ETA ₂	ETA _{3,10}	ETA _{3,2,1}
Vazões Fluviais	Flu _{2,4}	Flu _{4,2}	Flu ₄	Flu ₂	Flu _{1,2}	Flu ₂	Flu _{2,4}	Flu _{2,4}
Chuvas Thiessen	Th _{7,5}	Th ₅	Th ₅	Th _{4,6,3,1}	Th _{5,6,4}	Th _{5,7,6,3}	Th _{7,5}	Th _{7,1}
Vazão Incremental	Vinc	Vinc	Vinc	Vinc	Vinc	Vinc	Vinc	nenhum a
Variáveis selecionadas em relação à								
Grupo	Vmd6	Vmsd7	Vmd8	Vmd9	Vmd10	Vmd11	Vmd12	
ETA	ETA _{2,3,1,10}	ETA _{2,3,1}	ETA _{2,4,1}	ETA _{4,2,1}	ETA _{2,4,6,1,3}	ETA _{3,6,2,4}	ETA _{3,9,6}	
Vazões Fluviais	Flu _{2,4}	Flu _{2,4}	Flu ₄	Flu _{4,2}	Flu ₂	Flu _{2,4}	Flu ₂	
Chuvas Thiessen	nenhuma	Th ₅	Th _{4,5}	Th ₄	nenhuma	Th ₂	Th _{2,4}	
Vazão Incremental	Vinc	Vinc	Vinc	Vinc	Vinc	nenhuma	Vinc	

Há alguns resultados interessantes que podem ser percebidos na tabela acima:

- Em algumas seleções, certas variáveis, apesar de terem apresentado um grau de importância superior ao de outras variáveis selecionadas no 1º LSE, não foram selecionadas pelo 2º LSE, ao passo que as outras com grau inferior no 1º

LSE, foram. Por exemplo, em relação à V_{ms3} , na tabela 5.6, foram selecionadas em ordem decrescente de graus, a 3ª, a 6ª e a 5ª chuvas de Thiessen. No entanto, na tabela 5.7, a chuva de Thiessen selecionada por este 2º LSE foi a 5ª chuva, mesmo tendo apresentado graus menores que as chuvas 3ª e 6ª no 1º LSE (como pode ser confirmado na tabela 4.5 e no gráfico/tabela A.6.3);

- Alguns grupos de variáveis, em relação a algumas vazões, não tiveram nenhuma variável selecionada depois do 2º LSE. É o caso das chuvas de Thiessen em relação à V_{md6} e V_{md10} e até mesmo da Vazão incremental em relação à V_{md5} e V_{md11} .

Há de se ressaltar que estas variáveis finais (tabela 5.7), que foram selecionadas por 2 algoritmos LSE, serão, inicialmente, aquelas que depois de normalizadas constituirão as entradas do módulo 3.

Há, ainda, alguns grupos de variáveis que também passarão pelo menos por um dos módulos 2 e 3, e que foram originadas a partir de algum dos grupos originais (chuvas ETA, vazões fluviais, chuvas Thiessen e última vazão incremental), ou seja:

- Matriz **C** (1 x 311) – matriz que é a soma dos dez dias de previsão de chuvas ETA, ou seja, sua linha representa a soma das dez linhas da matriz **ETA**.

- Matriz **D** (1 x 311) – matriz com os valores das vazões médias semanais atrasadas a partir de duas semanas, ou seja, se está-se querendo prever uma vazão média semanal 1 (sábado a sexta) no tempo t , selecionar-se-ão todas as V_{ms1} de $t-2$ até $t-312$ (311 valores) para usar como padrões de entrada da rede. Não é possível selecionar a partir de $t-1$, pois, como se está prevendo tudo na terça-feira, neste tempo ainda não se tem o valor de $t-1$ exato, pois $t-1$ é a média do sábado passado até a próxima sexta.

- Matriz **F** (1 x 311) – matriz que é a soma das 7 chuvas de Thiessen, ou seja, sua linha representa a soma das 7 linhas da matriz **Th**;

Essas matrizes foram formadas por sugestão de pesquisadores que participaram do Projeto PreVIP [62], sendo assim, com já alguma expertise em relação à modelagem da Bacia de Itaipu. Ressalta-se que as matrizes acima não passaram por nenhum processo de seleção de variáveis.

5.3

Previsão sem Transformação Wavelet

Nesta subseção, conforme já enunciado, se dará a previsão das 15 vazões, sem que os dados tenham passado pelo módulo 2 (transformação wavelet).

5.3.1

Características e Modelagem das Redes Neurais

A topologia e as características de uma Rede Neural Artificial Multi-Layer Perceptron (RNA-MLP) [89], estrutura de rede neural utilizada neste estudo, compreendem o número de entradas, o número de camadas escondidas da rede, o número de neurônios na camada escondida, função de ativação, método/tipo de aprendizado (atualização dos pesos e bias), número de épocas de treinamento.

Segundo os estudos teóricos de Cybenko [72] e Hornik et al. [90], apesar de que Redes Neurais Multi-Layer Perceptron possam ter mais de uma camada escondida, uma simples camada escondida é suficiente para que essas redes aproximem quaisquer funções não-lineares. De acordo com os trabalhos de Coulibaly et al. [91] e Zhang et al. [92], focados no estado da arte da previsão hidrológica através de redes neurais, muitos resultados experimentais confirmam que apenas uma camada escondida é suficiente para a maioria dos problemas envolvendo previsões. Dentre esses resultados estão todos os trabalhos envolvendo previsão de vazões por RNA-MLP, constantes na bibliografia desta dissertação [7-9], [12-14], [16-18], e [20], que ou usam somente uma camada escondida, ou conseguem as melhores previsões com essa estrutura (apenas uma camada escondida), em relação a outras testadas (mais de uma camada escondida).

Segundo alguns autores como Gupta et al. [8] e Kisi [13], há alguns detalhes ainda não resolvidos, em relação a RNA-MLP. Não há, ainda, método automático ou teoria para a determinação da ótima arquitetura (número de entradas, número de camadas escondidas, número de neurônios na camada escondida). A geometria varia muito de acordo com o problema em questão, e o método da tentativa e erro tem sido usado para alcançar uma estrutura aceitável. Além disso, o método de treinamento, habitualmente usado (Back Propagation), é demasiado lento.

Em vista do exposto acima, neste estudo de caso, foram utilizadas RNA-MLP com 3 camadas – uma camada de entrada, uma camada escondida e uma camada de saída.

Em relação ao número de entradas, já foi definido no item **5.2.3**.

Em relação ao número de neurônios na camada escondida, foram avaliadas topologias com número variando num intervalo entre 1 a 19 neurônios, tentando-se seguir a métrica de Hecht-Nielsen [89], segundo a qual o número de neurônios na camada escondida é menor ou igual ao dobro do número de entradas da rede mais um: $N_{\text{escondidos}} \leq 2 \text{ Entr} + 1$.

Para a atualização dos pesos e bias utilizou-se o método de aprendizado Levenberg-Marquardt (LM) – Apêndice 4 – (aprendizado do tipo *batch*). Este método, que atua minimizando a soma dos quadrados de funções não-lineares, tem por característica uma maior velocidade de convergência e substitui o tradicional método do gradiente descendente com termo de momento e taxa de aprendizagem adaptativa. Ainda, sobre treinamento das redes neurais neste trabalho, para se evitar o supertreinamento das redes, usou-se o Método Da Parada Antecipada – Early Stopping [93] (mais detalhes no Apêndice 5).

Em relação às funções de ativação, geralmente, as redes neurais de previsão de vazões utilizam para todos os neurônios, independente da camada, a função *sigmóide* ou *função logística*. No entanto, com o objetivo de acelerar o processo de treinamento, foi usada, neste estudo de caso, a função *tangente hiperbólica* para os neurônios da camada escondida, ficando a função *sigmóide* somente para o neurônio da camada de saída. As respectivas formas das funções podem ser encontradas no Apêndice 7.

Saindo da parte da modelagem propriamente dita, em relação ao uso das séries históricas de dados, foi usado para treinamento e validação o conjunto de dados entre 1996 a 2000 (80% treinamento – 207 padrões e 20% para validação – 52 padrões), e para teste o conjunto de dados relativo ao ano de 2001 (52 padrões). A utilização dos dados somente no período entre 1996 a 2001 se deu em função da existência dos dados de chuva prevista pelo modelo ETA, somente neste período.

5.3.2

Previsão de Vazões

As previsões iniciais foram realizadas somente para a Vazão natural incremental média semanal 1 – Vms1. Tendo em vista a natureza estocástica semelhante de todas as 15 variáveis de saída, levantando-se o conjunto de entradas que resulte na previsão com menor erro MAPE para Vms1, é bem possível que boa parte dessas entradas também resulte em boas previsões para as outras 14 variáveis. Sendo assim, foram realizadas inúmeras previsões, alterando-se o número de neurônios na camada escondida, as variáveis de entrada e até o número de épocas de treinamento. Esse número deve ser entendido como o número de épocas de treinamento, após o qual se apresenta o conjunto de validação à rede.

A escolha do conjunto de variáveis de entrada levou em conta inicialmente as variáveis selecionadas das tabelas 5.6 e 5.7. A partir dessas variáveis foram feitas alterações no sentido de que o resultado da previsão fosse sempre melhorando. É importante lembrar que os métodos de seleção de variáveis dão somente uma noção (a eficiência depende da adequabilidade do método à situação) de quais variáveis têm mais influência sobre uma determinada variável (neste caso a Vms1).

As duas tabelas 5.8 e 5.9, a seguir, mostram vários testes (previsões), realizados com **207** dados para treinamento, **52** dados para validação e **52** para teste. Constam nas tabelas as entradas das redes neurais, as saídas a serem previstas – Vms1, Vms2 e Vms3 na tabela 5.8 e Vmd1 a Vmd12 na tabela 5.9 – os menores erros MAPE encontrados e o respectivo número de neurônios na camada escondida (Nh) para esses erros.

Os 4 primeiros conjuntos de entradas são aqueles selecionados após o 2º LSE (Tabela 5.7). Os demais são algumas variações e algumas combinações possíveis desses 4 conjuntos, e desses com as matrizes **C**, **D** e **F**.

Tabela 5.8: Resultados das previsões semanais sem uso de Wavelets

Previsão sem Wavelet – Previsão Semanal											
Vms1				Vms2				Vms3			
Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh
1	ETA _{1,2}	41.99	9	1	ETA _{1,2,4}	47.51	5	1	ETA _{1,2,4,6}	47.28	7
2	Th _{5,7}	48.21	13	2	Th ₅	47.79	4	2	Th ₅	48.51	5
3	Flu _{2,4}	23.93	19	3	Flu _{2,4}	26.00	13	3	Flu ₄	29.22	8
4	Vinc	23.89	16	4	Vinc	28.58	19	4	Vinc	27.72	7
5	C	48.21	10	5	C	48.08	3	5	C	48.52	18
6	D	33.57	10	6	D	31.77	9	6	D	29.07	18
7	F	45.46	15	7	F	46.21	14	7	F	48.52	8
8	ETA _{1,2,3}	48.21	8	8	ETA _{1,2}	44.29	17	8	ETA _{1,2,4}	48.52	2
9	Vinc; Flu_{2,4}; D	17.892	7	9	Th _{3,5,7}	48.08	2	9	Th _{3,5,6}	48.22	11
10	Vinc; Flu _{2,4} ; D	18.80	10	10	Flu _{2,4} ; Vinc; D	22.34	5	10	Flu _{2,4}	22.84	12
11	Vinc; Flu _{2,4} ; D	22.29	13	11	Flu _{2,4} ; Vinc; D	19.38	7	11	Flu _{2,4} ; Vinc; D	25.08	7
12	Vinc; Flu _{2,4} ; ETA _{1,2} ; F	19.54	14	12	Flu _{2,4} ; Vinc; D	20.35	11	12	Flu _{2,4} ; Vinc; D	22.47	8
13	Vinc; Flu _{2,4} ; D	19.52	12	13	Flu _{2,4} ; Vinc; D	20.88	8	13	Flu _{2,4} ; Vinc; D	22.48	10
14	Vinc; Flu _{2,4} ; ETA _{1,2}	17.896	16	14	Flu_{2,4}; Vinc	19.03	14	14	Flu _{2,4} ; Vinc; D	21.97	15
15	Vinc; Flu _{2,4}	18.03	13	15				15	Flu_{2,4}; Vinc	20.73	15
16	Vinc; D	27.22	14	16				16			

Tabela 5.9: Resultados das previsões diárias sem uso de Wavelets

Previsão sem Wavelet – Previsão diária											
Vmd1				Vmd2				Vmd3			
Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh
1	ETA ₁	44.25	9	1	ETA ₁	48.066	2	1	ETA ₂	46.5	11
2	Th _{1,3,4,6}	47.51	11	2	Th _{4,5,6}	37.68	5	2	Th _{3,5,6,7}	48.41	18
3	Flu ₂	46.73	7	3	Flu _{1,2}	38.81	7	3	Flu ₂	43.03	10
4	Vinc	8.62	11	4	Vinc	14.90	5	4	Vinc	18.20	4
5	C	47.77	2	5	C	48.063	3	5	C	48.41	6
6	F	46.09	10	6	F	31.96	7	6	F	32.87	3
7	ETA _{1,4}	46.39	12	7	ETA _{1,4}	47.48	3	7	ETA _{4,10}	48.39	11
8	Flu _{1,2}	46.16	12	8	Flu _{1,2,3}	18.05	2	8	Th _{6,7}	48.41	11
9	Flu _{2,4}	35.93	14	9	Flu _{2,4}	29.00	2	9	Flu _{2,4}	20.31	15
10	Vinc; Flu _{2,4} F; ETA ₁	9.49	8	10	Vinc; Flu _{1,2,3} F; ETA _{1,4}	11.2	11	10	Vinc; Flu _{2,4} F; ETA ₂	12.14	2
11	Vinc; Flu _{2,4} F	9.48	10	11	Vinc; Flu _{1,2,3} ; F	11.41	15	11	Vinc; Flu _{2,4} F	11.94	15
12	Vinc; Flu _{2,4} ETA ₁	10.80	8	12	Vinc; Flu _{1,2,3} ETA _{1,4}	11.14	11	12	Vinc; Flu _{2,4} ETA ₂	12.49	19
13	Vinc; F; ETA ₁	7.94	4	13	Vinc; F; ETA _{1,4}	12.62	15	13	Vinc; F ETA ₂	18.16	11
14	Vinc; Flu _{2,4}	9.72	8	14	Vinc; F	10.98	6	14	Vinc; Flu_{2,4}	11.74	17
15	Vinc; F	7.86	17	15	Vinc; Flu _{1,2,3}	11.48	8	15	Vinc; F	16.80	12
16	Vinc; ETA ₁	8.36	16	16				16			

Previsão sem Wavelet – Previsão diária											
Vmd4				Vmd5				Vmd6			
Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh
1	ETA _{3,10}	48.35	13	1	ETA _{1,2,3}	46.00	8	1	ETA _{1,2,3,10}	45.86	13
2	Th _{7,5}	48.29	7	2	Th _{1,7}	47.49	15	2	Flu _{2,4}	19.39	16
3	Flu _{2,4}	16.26	15	3	Flu _{2,4}	16.98	13	3	Vinc	30.14	18
4	Vinc	20.65	8	4	Vinc	26.96	7	4	C	48.84	2
5	C	47.85	4	5	C	48.66	4	5	F	45.86	16

6	F	43.71	17	6	F	37.42	10	6	ETA _{2,3}	48.15	17
7	Flu _{2,4} ; Vinc; F; C	14.02	8	7	ETA ₃	46.65	13	7	Th _{1,3,7}	47.35	15
8	Flu _{2,4} ; Vinc; F	13.62	16	8	Flu _{2,4} ; Vinc F; ETA _{1,2,3}	19.42	3	8	Flu _{2,4} ; Vinc ETA _{1,2,3,10} ; F	20.78	10
9	Flu _{2,4} ; Vinc; C	12.78	14	9	Flu _{2,4} ; Vinc; F	16.80	16	9	Flu _{2,4} ; Vinc ETA _{1,2,3,10}	21.25	19
10	Flu _{2,4} ; F; C	18.54	11	10	Flu_{2,4}; Vinc ETA_{1,2,3}	15.70	3	10	Flu _{2,4} ; Vinc; F	23.57	16
11	Flu_{2,4}; Vinc	11.03	7	11	Flu _{2,4} ; F ETA _{1,2,3}	21.90	5	11	Flu _{2,4} ; ETA _{1,2,3,10}	25.14	8
12	Flu _{2,4} ; F	17.27	15	12	Flu _{2,4} ; Vinc	16.32	11	12	Flu_{2,4}; Vinc	15.52	18
13				13	Flu _{2,4} ; F	18.89	18	13	Flu _{2,4} ; ETA _{1,2,3,10}	21.36	12

Previsão sem Wavelet – Previsão diária

Vmd7				Vmd8				Vmd9			
Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh
1	ETA _{1,2,3}	49.19	16	1	ETA _{1,2,4}	48.61	19	1	ETA _{1,2,4}	48.83	4
2	Th ₅	47.97	19	2	Th _{4,5}	48.53	19	2	Th ₄	44.13	12
3	Flu _{2,4}	25.89	14	3	Flu ₄	30.33	7	3	Flu _{2,4}	31.85	9
4	Vinc	30.79	16	4	Vinc	32.26	11	4	Vinc	31.76	5
5	C	49.19	2	5	C	48.61	5	5	C	48.83	12
6	F	46.37	16	6	F	48.50	13	6	F	48.83	10
7	ETA ₂	47.88	7	7	ETA ₂	47.70	19	7	ETA _{2,4}	48.50	18
8	Th _{3,5}	49.06	9	8	Th _{3,4}	46.30	14	8	Th _{4,6}	48.02	15
9	Flu_{2,4}; Vinc	20.86	17	9	Flu _{2,4}	26.67	12	9	Vinc; Flu _{2,4}	28.68	9
10	Flu _{2,4} ; Vinc; F	24.81	15	10	Flu _{2,4} ; Vinc Th _{3,4} ; ETA ₂	29.10	17	10	Vinc; Flu _{2,4} Th _{3,4}	31.00	8
11	Flu _{2,4} ; Vinc ETA ₂	23.68	15	11	Flu _{2,4} ; Vinc Th _{3,4}	31.25	6	11	Vinc; Flu _{2,4} ETA ₂	27.31	9
12	Flu _{2,4} ; F; ETA ₂	25.66	10	12	Flu _{2,4} ; Vinc ETA ₂	25.14	14	12	Vinc ; Th ₄ ETA ₂	32.36	8
13	Flu _{2,4} ; Vinc	21.83	18	13	Flu _{2,4} ; Th _{3,4} ETA ₂	29.15	12	13	Vinc; Flu_{2,4}	24.31	16
14	Flu _{2,4} ; F	26.22	11	14	Flu_{2,4}; Vinc	24.01	6	14	Vinc; Th ₄	35.27	19
15	Flu _{2,4} ; ETA ₂	26.07	15	15				15	Flu _{2,4} ; Th ₄	35.04	6
16				16				16			

Previsão sem Wavelet – Previsão diária

Vmd10				Vmd11				Vmd12			
Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh	Nº	Entradas	MAPE (%)	Nh
1	ETA _{1,2,3,4,6}	49.09	10	1	ETA _{2,3,4,6}	49.09	12	1	ETA _{3,6,9}	48.54	9
2	Flu ₂	48.04	11	2	Th ₂	47.50	17	2	Th _{2,4}	48.07	4
3	Vinc	30.59	10	3	Flu _{2,4}	32.92	12	3	Flu ₂	49.46	11
4	C	48.95	5	4	Vinc	30.75	10	4	Vinc	33.83	14
5	F	45.24	13	5	C	49.09	2	5	C	49.46	19
6	ETA _{2,4,6}	47.80	8	6	F	49.09	2	6	F	49.45	14
7	Th _{4,6}	48.06	2	7	ETA _{3,6}	47.85	12	7	ETA _{3,9}	49.46	7
8	Flu _{2,4}	30.40	3	8	Th _{2,6}	49.09	2	8	Th ₂	48.18	19
9	Flu _{2,4} ; Vinc F; ETA _{2,4,6}	29.74	11	9	Vinc; Flu _{2,4} Th ₂ ; ETA _{3,6}	30.00	14	9	Flu _{2,4}	33.26	4
10	Flu _{2,4} ; Vinc F	30.36	15	10	Vinc; Flu_{2,4} Th₂	29.39	14	10	Flu _{2,4} ; Vinc Th _{2,4} ; ETA _{3,6,9}	36.21	4
11	Flu _{2,4} ; Vinc ETA _{2,4,6}	25.38	18	11	Vinc; Flu _{2,4} ; ETA _{3,6}	32.83	14	11	Flu _{2,4} ; Vinc Th _{2,4}	35.24	10
12	Flu _{2,4} F; ETA _{2,4,6}	33.08	6	12	Vinc; Th ₂ ETA _{3,6}	34.03	19	12	Flu _{2,4} ; Vinc ETA _{3,6,9}	30.55	17
13	Flu_{2,4}; Vinc	25.02	17	13	Vinc; Flu _{2,4}	29.72	16	13	Flu _{2,4} ; Th _{2,4} ETA _{3,6,9}	39.92	3
14				14				14	Flu_{2,4}; Vinc	30.11	10

A tabela 5.10, a seguir, resume os melhores resultados obtidos para as previsões semanais e diárias (tabelas 5.8 e 5.9).

Tabela 5.10: Melhores resultados obtidos e respectivos conjuntos de entrada

Melhores Resultados de Previsão						
Previsão de	Melhor entrada	N _h	MAPE Validação	MAPE Teste	RMSE Teste	
Vms1	Vinc; Flu _{2,4} D; ETA _{1,2} ; F	7	17.892	20.63	0.2477	
Vms2	Flu _{2,4} ; Vinc	14	19.03	20.89	0.2901	
Vms3	Flu _{2,4} ; Vinc	15	20.73	20.75	0.527	
Vmd1	Vinc; F	17	7.86	6.00	0.0841	
Vmd2	Vinc; F	6	10.98	10.60	0.1517	
Vmd3	Vinc; Flu _{2,4}	17	11.74	13.08	0.1738	
Vmd4	Flu _{2,4} ; Vinc	7	11.03	13.25	0.1745	
Vmd5	Flu _{2,4} ; Vinc; ETA _{1,2,3}	3	15.70	16.43	0.2251	
Vmd6	Flu _{2,4} ; Vinc	18	15.52	25.27	0.4211	
Vmd7	Flu _{2,4} ; Vinc; F; ETA ₂	17	20.86	25.35	0.4528	
Vmd8	Flu _{2,4} ; Vinc	6	24.01	27.62	0.3679	

Melhores Resultados de Previsão						
Previsão de	Melhor entrada	N_h	MAPE Validação	MAPE Teste	RMSE Teste	
Vmd9	Vinc; Flu _{2,4}	16	24.31	31.31	0.4432	
Vmd10	Flu _{2,4} ; Vinc	17	25.02	23.57	0.3112	
Vmd11	Vinc; Flu _{2,4} ; Th ₂	14	29.39	36.89	0.5046	
Vmd12	Flu _{2,4} ; Vinc	10	30.11	52.37	0.9904	

5.3.3

Avaliação dos Resultados

Pelos resultados mostrados nas tabelas anteriores, nota-se que o modelo ETA não teve uma influência positiva nos resultados das previsões. Ou seja, nos casos em que a presença das séries históricas das chuvas ETA como entradas das redes neurais melhoraram a previsão da rede (3 em 15 previsões), a diferença no erro MAPE foi pequena. A tabela 5.11 abaixo dá uma idéia dessa diferença.

Tabela 5.11: Tabela comparativa dos resultados com e sem ETA

Previsão	Melhor resultado com ETA		Melhor resultado sem ETA		Diferença absoluta (%)
	variáveis	MAPE (%)	variáveis	MAPE (%)	
Vms1	Vinc; Flu _{2,4} D; ETA _{1,2} ; F	17.89	Vinc; Flu _{2,4}	18.03	0.14
Vmd5	Flu _{2,4} ; Vinc ETA _{1,2,3}	15.70	Flu _{2,4} ; Vinc	16.32	0.62
Vmd7	Flu _{2,4} ; Vinc; F; ETA ₂	20.86	Flu _{2,4} ; Vinc	21.83	0.97

A tabela acima mostra que a inclusão das “chuvas ETA”, só melhora os resultados, em 3 das 15 previsões, e a melhora média nesses 3 casos foi de apenas 0.57%, em termos de erros MAPE (%) de validação. Essa ineficiência advém do fato que o modelo ETA, por exemplo, prevê as 3 primeiras precipitações das semanas de previsão, ou seja, todas as chuvas de quarta a sexta, desde janeiro de 1996 até dezembro de 2000, com um erro médio absoluto de aproximadamente 29%. Prevê as 5 primeiras chuvas – quarta a domingo – no mesmo período, com um erro médio absoluto de aproximadamente 46%. E prevê as 10 primeiras chuvas – quarta a 2ª sexta – com um erro de aproximadamente 101%. Isso faz com que as informações de chuvas previstas pelo modelo ETA, em geral, não agreguem valor significativo às previsões. Como se pode notar pelos números

acima e pela tabela 5.11, somente as primeiras chuvas previstas pelo modelo, em raras vezes, ajudam às previsões.

5.4

Previsão com Transformação Wavelet

Nesta subseção, conforme já enunciado, a previsão das 15 vazões será realizada com a introdução do módulo 2 (transformação wavelet). Os dados então, agora já tratados pelo módulo 1, serão decompostos em coeficientes wavelets.

5.4.1

Características dos Módulos 2 e 3

O módulo de wavelets decompõe os dados ou séries temporais, que vêm do módulo 1, em 2 níveis de coeficientes wavelets. Conforme visto anteriormente (seção 3.3), o nível 1 contém um componente de detalhe – D1 – e o nível 2 contém um componente de aproximação – A2 – e um de detalhe – D2. Serão utilizadas para a decomposição das séries temporais as famílias de wavelets Daubechies, Biortogonal e Coiflets.

Da mesma maneira que na previsão sem o uso de wavelets, em princípio, tentar-se-á descobrir a melhor estrutura de decomposição wavelet para V_{ms1} , entendendo-se como melhor estrutura aquela que implicará em menores erros MAPE no módulo 3.

As características do módulo 3 de redes neurais para essas previsões, com o uso do módulo 2 de wavelets, são as mesmas discriminadas na subseção 5.3.1.

5.4.2

Testes Preliminares

No intuito de verificar a eficácia da utilização da transformada wavelet no processo de previsão, foram realizados vários testes preliminares, todos para previsão de V_{ms1} .

Como em estudos anteriores [84, 87] foram usados como entradas das redes neurais os valores anteriores (decompostos por wavelets) das saídas a serem

previstas, foram testadas ainda como entradas uma janela com 3 valores anteriores de $Vms1 - Vms1_{t-2}$, $Vms1_{t-3}$ e $Vms1_{t-4}$. Não foi usado o primeiro valor anterior de $Vms1$, que é $Vms1_{t-1}$, tendo em vista que na quarta-feira de previsão ainda não se tem disponíveis os valores de vazão do 8º ao 10º dia de previsão da semana anterior, ou seja, quarta, quinta e sexta da presente semana.

Mais uma vez, tendo em vista trabalhos anteriores, como os acima citados, foram testadas a decomposição em 3 níveis e, como uma alternativa mais simples, a decomposição em 2 níveis, usando como primeira alternativa wavelets da família daubechies, mais precisamente db_3 e db_{10} .

Tabela 5.12: Previsões usando wavelets daubechies

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
coeficientes A_3, D_3, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (12 entradas)	db_3	3	17.71	3
coeficientes A_3, D_3 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_3	3	17.59	3
coeficientes A_3 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (3 entradas)	db_3	3	22.42	6
coeficientes A_3, D_3 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_3	3	20.14	6
coeficientes A_3, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_3	3	22.33	4
coeficientes A_3, D_3 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_3	3	21.05	5
coeficientes A_3 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_3	3	21.52	3
coeficientes A_2, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_3	2	20.23	5
coeficientes A_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_3	2	24.25	2
coeficientes A_2 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_3	2	21.74	7
coeficientes A_3, D_3, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (12 entradas)	db_{10}	3	13.57	3
coeficientes A_3, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_{10}	3	14.09	2
coeficientes A_3, D_3 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_{10}	3	12.61	8
coeficientes A_3 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_{10}	3	13.81	5
coeficientes A_2, D_2 e D_1 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (9 entradas)	db_{10}	2	13.26	2
coeficientes A_2 e D_2 de $Vms1_{t-2}, Vms1_{t-3}, Vms1_{t-4}$ (6 entradas)	db_{10}	2	12.9	10

Da tabela acima, pode-se verificar que os melhores resultados ocorreram com o uso de db_{10} , sendo que os níveis de decomposição 2 e 3 alcançaram

resultados bem próximos, 12.9% e 12.61%, respectivamente. Procurando uma melhora ainda maior, testou-se uma previsão tendo como entradas uma janela de 3 valores anteriores da última vazão natural incremental (terça-feira), como indicado na Tabela 5.13 a seguir.

Tabela 5.13: Previsões usando como entradas janela de 3 valores anteriores da última vazão natural

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$ (9 entradas)	db ₁₀	2	11.57	3
coeficientes A_2 e D_2 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$ (6 entradas)	db ₁₀	2	12.68	4
coeficientes A_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$ (6 entradas)	db ₁₀	2	15.34	3
coeficientes A_3 , D_3 , D_2 , D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$ (12 entradas)	db ₁₀	3	12.30	2
coeficientes A_3 , D_3 e D_2 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$ (6 entradas)	db ₁₀	3	15.05	12

Como os últimos resultados com o nível de decomposição 2 (MAPE=11,57%) foram melhores do que os resultados com o nível 3 (12,30%) , passou-se a usar somente o nível 2. Como os resultados melhoraram com o uso de janela de Vinc ao invés de janela de Vms1, optou-se por continuar com Vinc, só que agora aumentando o tamanho da janela. Em vez de 3 dados (vazões incrementais de terça) anteriores, aumentou-se para 5 dados anteriores e testaram-se outras famílias de wavelets.

Tabela 5.14: Previsões usando wavelet coif₅ e como entradas 5 valores anteriores da última vazão natural

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$ e $Vinc_{t-5}$ (15 entradas)	Coif ₅	2	9.59	5

Agora, aumentou-se para 7 dados anteriores e, além das coiflets, usou-se a biortogonal.

Tabela 5.15: Previsões usando wavelet coif_5 e bior e como entradas 7 valores anteriores da última vazão natural

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	coif_5	2	9.28	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{3.1}$	2	17.45	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{3.3}$	2	14.7	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{3.5}$	2	11.59	3
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{3.7}$	2	12.05	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{3.9}$	2	9.73	3
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{4.4}$	2	11.36	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{5.5}$	2	11.74	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{6.8}$	2	10.44	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	$\text{bior}_{2.8}$	2	11.64	2

Voltando, agora, para a família daubechies.

Tabela 5.16: Previsões usando wavelet daubechies e como entradas 7 valores anteriores da última vazão natural

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	db_{10}	2	9.56	3
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	db_{20}	2	8.49	2
coeficientes A_2 , D_2 e D_1 de Vinc_{t-1} , Vinc_{t-2} , Vinc_{t-3} , Vinc_{t-4} , Vinc_{t-5} , Vinc_{t-6} , Vinc_{t-7} (21 entradas)	db_{30}	2	8.82	3

Pode-se notar que, retornando à família daubechies, as previsões melhoraram. Nota-se, também, que o resultado melhora quando aumenta-se a ordem da família de db_{10} para db_{20} , e piora quando se aumenta de novo para db_{30} . Uma hipótese é que a melhor ordem esteja entre db_{20} e db_{30} , o que é confirmado pela tabela 5.17 abaixo.

Tabela 5.17: Previsões usando wavelet db_{25}

Entradas	Wavelet-mãe	Níveis de decomposição	MAPE (%)	Nh
Coefficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	7.52	5

Como db_{25} foi a wavelet-mãe que apresentou o melhor resultado, pelo menos para $Vms1$, esta será usada também para as outras 14 previsões. Tentou-se continuar ainda com a janela de 7 valores anteriores da vazão natural incremental de terça-feira – $Vinc$.

Tabela 5.18: Seis previsões usando wavelet db_{25}

Previsão	Entradas	Wavelet-mãe	Níveis de decomposição	MAPE Validação (%)	MAPE Teste (%)	Nh
$Vms1$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	6.77	8.09	2
$Vms2$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	9.29	12.39	5
$Vms3$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	10.94	12.26	5
$Vmd1$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	7.92	9.26	3
$Vmd2$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	13.72	11.45	3
$Vmd3$	coeficientes A_2 , D_2 e D_1 de $Vinc_{t-1}$, $Vinc_{t-2}$, $Vinc_{t-3}$, $Vinc_{t-4}$, $Vinc_{t-5}$, $Vinc_{t-6}$, $Vinc_{t-7}$ (21 entradas)	db_{25}	2	16.19	16.12	5

Na tabela acima, pode-se notar que, apesar dos bons resultados de Vms1, Vms2 e Vms3, as três primeiras previsões diárias não foram satisfatórias, piores do que as previsões realizadas sem wavelets. Testou-se então, como entradas, uma janela de 7 valores anteriores das próprias vazões a serem previstas. Ressalta-se que, a partir de Vmd8 (8º dia de previsão), a janela de 7 valores anteriores começa em Vmd_{t-2} , pois na quarta-feira (dia em que são feitas as previsões) ainda não se tem os valores $Vmd_{8,t-1}$ a $Vmd_{12,t-1}$ (valores da semana anterior).

Tabela 5.19: Previsões diárias usando wavelet db_{25} e 7 valores anteriores de vazões diárias

Previsão	Entradas	Wavelet-mãe	Níveis de decomposição	MAPE Validação (%)	MAPE Teste (%)	Nh
Vmd1	coeficientes A2, D2 e D1 de Vmd1t-1, Vmd1t-2, Vmd1t-3, Vmd1t-4, Vmd1t-5, Vmd1t-6, Vmd1t-7 (21 entradas)	db25	2	3.47	4.81	2
Vmd2	coeficientes A2, D2 e D1 de Vmd2t-1, Vmd2t-2, Vmd2t-3, Vmd2t-4, Vmd2t-5, Vmd2t-6, Vmd2t-7 (21 entradas)	db25	2	4.40	5.79	3
Vmd3	coeficientes A2, D2 e D1 de Vmd3t-1, Vmd3t-2, Vmd3t-3, Vmd3t-4, Vmd3t-5, Vmd3t-6, Vmd3t-7 (21 entradas)	db25	2	4.06	4.89	3
Vmd4	coeficientes A2, D2 e D1 de Vmd4t-1, Vmd4t-2, Vmd4t-3, Vmd4t-4, Vmd4t-5, Vmd4t-6, Vmd4t-7 (21 entradas)	db25	2	5.12	4.36	2
Vmd5	coeficientes A2, D2 e D1 de Vmd5t-1, Vmd5t-2, Vmd5t-3, Vmd5t-4, Vmd5t-5, Vmd5t-6, Vmd5t-7 (21 entradas)	db25	2	6.64	3.96	3
Vmd6	coeficientes A2, D2 e D1 de Vmd6t-1, Vmd6t-2, Vmd6t-3, Vmd6t-4, Vmd6t-5, Vmd6t-6, Vmd6t-7 (21 entradas)	db25	2	7.01	4.29	1
Vmd7	coeficientes A2, D2 e D1 de Vmd7t-1, Vmd7t-2, Vmd7t-3, Vmd7t-4, Vmd7t-5, Vmd7t-6, Vmd7t-7 (21 entradas)	db25	2	5.37	3.52	2
Vmd8	coeficientes A2, D2 e D1 de Vmd8t-2, Vmd8t-3, Vmd8t-4, Vmd8t-5, Vmd8t-6, Vmd8t-7, Vmd8t-8 (21 entradas)	db25	2	11.59	9.75	3

Previsão	Entradas	Wavelet-mãe	Níveis de decomposição	MAPE Validação (%)	MAPE Teste (%)	Nh
Vmd9	coeficientes A2, D2 e D1 de Vmd9t-2, Vmd9t-3, Vmd9t-4, Vmd9t-5, Vmd9t-6, Vmd9t-7, Vmd9t-8 (21 entradas)	db25	2	11.59	13.99	2
Vmd10	coeficientes A2, D2 e D1 de Vmd10t-2, Vmd10t-3, Vmd10t-4, Vmd10t-5, Vmd10t-6, Vmd10t-7, Vmd10t-8 (21 entradas)	db25	2	12.44	13.31	4
Vmd11	coeficientes A2, D2 e D1 de Vmd11t-2, Vmd11t-3, Vmd11t-4, Vmd11t-5, Vmd11t-6, Vmd11t-7, Vmd11t-8 (21 entradas)	db25	2	13.97	11.10	2
Vmd12	coeficientes A2, D2 e D1 de Vmd12t-2, Vmd12t-3, Vmd12t-4, Vmd12t-5, Vmd12t-6, Vmd12t-7, Vmd12t-8 (21 entradas)	db25	2	15.86	10.89	3

De fato, usando db₂₅, nível 2 de composição e usando tanto a aproximação como todos os detalhes (níveis 1 e 2), os resultados melhoraram consideravelmente para as previsões diárias e semanais. Percebe-se uma queda de rendimento do modelo nas previsões diárias a partir de Vmd8, tendo em vista não se poder usar, conforme mencionado no parágrafo anterior, os valores Vmd8_{t-1} a Vmd12_{t-1} (valores da semana anterior). Mesmo assim, as previsões do modelo para essas 5 saídas, são ainda bem superiores às previsões sem wavelet.

5.5

Previsão sem Wavelet x Previsão com Wavelet

A tabela abaixo resume os resultados das 15 previsões e mostra a grande melhoria obtida pelo uso da análise wavelet.

Tabela 5.20: Previsão sem Wavelet x Previsão com Wavelet

Previsão	Melhor resultado sem wavelet MAPE Validação (%) – coluna 1	Melhor resultado com wavelet MAPE Validação (%) – coluna 2	Melhora absoluta (%)	
Vms1	17.89	6.77	11,12	
Vms2	19.03	9.29	9.74	
Vms3	20.73	10.94	9,79	
Vmd1	7.86	3.47	4.39	
Vmd2	10.98	4.40	6,58	
Vmd3	11.74	4.06	7.68	
Vmd4	11.03	5.12	5,91	
Vmd5	15.70	6.64	9.06	
Vmd6	15.52	7.01	8,51	
Vmd7	20.86	5.37	15.49	
Vmd8	24.01	11.59	12,42	
Vmd9	24.31	11.59	12.72	
Vmd10	25.02	12.44	12,58	
Vmd11	29.39	13.97	15.42	
Vmd12	30.11	15.86	14.25	
MAPE médio	18.94	8.56	Melhora média	10.38

No apêndice 8 estão 6 gráficos, cada um apresentando 52 valores de duas curvas: vazão observada e vazão simulada. Esses 52 valores dizem respeito aos 52 padrões (52 dentre 311 padrões totais – ver subseção 5.3.1) para se testar o desempenho do modelo. Dentre esses 6 gráficos, 3 dizem respeito a vazões simuladas usando wavelet e 3 a vazões simuladas sem o uso de wavelet. A intenção desses gráficos, cujos horizontes de previsão (Vmd2, Vmd4 e Vmd11) foram aleatoriamente escolhidos, é mostrar com clareza a diferença entre as previsões com e sem o uso de wavelet.