UNIVERSIDADE DE SÃO PAULO–USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Diana Ruth Mejia de Lima

**Previsão de Vazões Afluentes a Usinas
Hidrelétricas Aplicada à Programação
da Operação do Sistema Elétrico
Brasileiro**

São Carlos
2018

Diana Ruth Mejia de Lima

**Previsão de Vazões Afluentes a Usinas
Hidrelétricas Aplicada à Programação
da Operação do Sistema Elétrico
Brasileiro**

Dissertação de mestrado apresentada ao Programa de Engenharia Elétrica da Escola de Engenharia de São Carlos como parte dos requisitos para a obtenção do título de Mestre em Ciências.

Área de concentração: Sistemas Dinâmicos

Orientador: Prof. Dr. Rogério Andrade Flauzino

São Carlos

2018

Trata-se da versão corrigida da dissertação. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

L732p Lima, Diana Ruth Mejia de
Previsão de Vazões Afluentes a Usinas Hidrelétricas
Aplicada à Programação da Operação do Sistema Elétrico
Brasileiro / Diana Ruth Mejia de Lima; orientador
Rogério Andrade Flauzino. São Carlos, 2018.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia Elétrica e Área de Concentração em
Sistemas Dinâmicos -- Escola de Engenharia de São
Carlos da Universidade de São Paulo, 2018.

1. planejamento da operação. 2. sistema
hidrotérmico. 3. previsão de vazão hidrológica. 4.
modelos lineares. 5. sistemas de inferência fuzzy. I.
Título.

FOLHA DE JULGAMENTO

Candidata: Engenheira **DIANA RUTH MEJIA DE LIMA**.

Título da dissertação: "Previsão de vazões afluentes a usinas hidrelétricas aplicadas à programação da operação do sistema elétrico brasileiro".

Data da defesa: 17/09/2018.

Comissão Julgadora:

Resultado:

Prof. Associado **Rogério Andrade Flauzino**
(Orientador)
(Escola de Engenharia de São Carlos/EESC)

Aprovado

Prof. Dr. **Danilo Hernane Spatti**
(Instituto de Ciências Matemáticas e de Computação/ICMC)

Spatti

Prof. Dr. **Ricardo Augusto Souza Fernandes**
(Universidade Federal de São Carlos/UFSCar)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Luís Fernando Costa Alberto**

Este trabalho é dedicado aos meus pais.

Agradecimentos

O desenvolvimento desta dissertação de mestrado não seria possível sem a participação direta e indireta de diversas pessoas das quais recebi constante apoio e motivação. Portanto, gostaria de agradecer a todos que colaboraram de alguma forma na execução desse trabalho.

Agradeço aos docentes da Escola de Engenharia de São Carlos da USP, onde conclui a graduação e, conseqüentemente, adquiri toda base de conhecimento para desenvolvimento não apenas desse como de todos outros trabalhos acadêmicos e profissionais.

Meus agradecimentos, em especial, ao meu orientador Prof. Dr. Rogério Andrade Flauzino, pela oportunidade que me concedeu de fazer parte do programa de mestrado do Departamento de Engenharia Elétrica e de Computação da Escola de Engenharia de São Carlos e pelas orientações no desenvolvimento da presente dissertação.

Agradeço aos meus pais, pelo apoio incondicional via orações. Sou eternamente grata por terem me proporcionado educação e amor, e apesar das inúmeras dificuldades, por sempre me estimularem a continuar.

Ao meu namorado William pelo companheirismo, amizade, paciência, compreensão e apoio nos momentos mais críticos.

Às minhas amigas Maristela Santos, Mirian Wawrzyniak, Jéssica Freire, Mariana Iizuka e Érica Pereira, a quem realmente considero como irmãs. Obrigada por me ajudarem sempre que necessário.

Aos meus amigos e companheiros de trabalho na Compass, Gustavo Arflux, Natalia Biondo, Douglas Canducci, Wendel Loschi, Lucas Picarelli, que me ensinaram muito.

Finalmente, a todos que de alguma maneira contribuíram para minha formação profissional e minha formação como cidadã.

“Todas as vitórias ocultam uma abdicação” - Simone de Beauvoir.

Resumo

Lima, Diana Ruth Mejia. **Previsão de Vazões Afluentes a Usinas Hidrelétricas Aplicada à Programação da Operação do Sistema Elétrico Brasileiro.** 138 p. Dissertação de mestrado – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2018.

Este trabalho aborda o problema de modelagem de séries de vazões afluentes aos aproveitamentos hidrelétricos. A previsão de vazão natural fluvial é realizada semanalmente para 158 usinas hidrelétricas do Sistema Interligado Nacional (SIN), pois trata-se de insumo fundamental para o planejamento e operação do sistema elétrico brasileiro. Diversos modelos são utilizados na determinação destas previsões, entre os quais podem ser citados os modelos físicos, os estatísticos e aqueles que aplicam sistemas inteligentes. Apesar de contínuos aprimoramentos terem sido incorporados ao processo de previsão de vazão, existem alguns aproveitamentos hidrelétricos para os quais os resultados de estimação têm apresentado grandes desvios. Neste contexto, com a motivação de se obter uma resposta acurada, investigam-se os sistemas fuzzy como modelos concorrentes aplicados à previsão de vazões semanais. O objetivo do trabalho é reduzir os erros de estimação para uma usina piloto, incorporando à previsão de vazão os dados de precipitação. Para a construção da série histórica de precipitação média da bacia hidrográfica, fez-se uma exaustiva pesquisa por estações pluviométricas, seguida por tratamento de dados de medição e método de interpolação. Ao final do trabalho, é apresentada uma análise comparativa entre os resultados obtidos com o Modelo Autorregressivo Periódico (PAR) e o sistema de inferência fuzzy. Com base no desempenho observado, superior ao modelo autorregressivo, comprova-se a adequação do modelo proposto para a modelagem do processo hidrológico.

Palavras-chave: sistema hidrotérmico, planejamento da operação, previsão de vazão hidrológica, modelos lineares, sistemas de inferência fuzzy.

Abstract

Lima, Diana Ruth Mejia. **Streamflow Forecasting Applied to the Operation Planning of the Brazilian Electric Power System.** 138 p. Master Thesis – São Carlos School of Engineering, University of São Paulo, 2018.

This work addresses the modelling problem of hydropower plants reservoir streamflow series. The natural streamflow forecasting for 157 hydroelectric power plants of the National Interconnected System - NIS is updated on a weekly basis, which is an essential input for the planning and operation of the Brazilian Electric Power System. Several models are used to determine this prediction, such as physicals, statisticals and the ones that use intelligent systems. Despite the improvements to natural streamflow forecasting, substantial deviation has been found for the expected results of some hydropower plants. Highlighted the importance of this variable, fuzzy systems applied to weekly streamflows forecasts will be investigated as alternative models, in order to obtain better results.

The purpose of this work is to reduce the estimation errors for a pilot hydropower plant, incorporating precipitation data into the forecast. Therefore, an exhaustive research to acquire data from hydrometeorological stations was conducted. After being treated, a variable selection method was applied to the data, defining the most relevant input variables for the prediction model. At the end, a comparative analysis shows that the fuzzy model presents a better performance than the periodic autoregressive model used by ONS to plan the operation of the electric power system.

Keywords: hydrothermal system, planning of the operation, streamflow forecast, linear statistical models, fuzzy inference system.

Lista de ilustrações

Figura 1	Composição da geração de energia elétrica por fonte.	25
Figura 2	Histórico do preço médio mensal calculado para o submercado Sudeste.	35
Figura 3	Processo de decisão operativa com incertezas nas vazões.	38
Figura 4	Cadeia de modelos utilizados no planejamento energético.	39
Figura 5	Representação da Função de Custo Futuro.	40
Figura 6	Representação da árvore de cenários de afluições.	43
Figura 7	Esquema de representação do cenário de vazões afluentes as usinas hidrelétricas.	44
Figura 8	Estágios da metodologia Box e Jenkins	49
Figura 9	Comparação entre vazão observada e prevista com uma semana de antecedência para a usina hidrelétrica de Três Irmãos	62
Figura 10	Representação da variável linguística vazão.	68
Figura 11	Função de pertinência discretizada	69
Figura 12	Diagrama esquemático de um sistema de inferência fuzzy.	70
Figura 13	Representação das entradas do sistema fuzzy.	72
Figura 14	Região fuzzy associada a saída do sistema.	73
Figura 15	Região fuzzy de saída.	80
Figura 16	Algoritmo de extração de regras fuzzy.	83
Figura 17	Sistema fuzzy multicamada.	83
Figura 18	Mapa de localização dos postos pluviométricos na sub-bacia de Camargos.	89
Figura 19	Mapa de localização dos postos pluviométricos na sub-bacia de Três Irmãos.	90
Figura 20	Gráficos de dupla massa para consistência dos dados das estações pluviométricas de Camargos.	92
Figura 21	Gráficos de dupla massa para consistência dos dados das estações pluviométricas de Três Irmãos.	93

Figura 22	Polígonos de Thiessen.	94
Figura 23	Variação da precipitação e da vazão anual padronizadas na UHE Camargos.	95
Figura 24	Correlação cruzada entre a vazão natural semanal e as precipitações médias acumuladas para Camargos.	96
Figura 25	Correlação cruzada entre a vazão natural semanal e as precipitações médias acumuladas para Três Irmãos.	96
Figura 26	Vazão natural mensal ao reservatório de Camargos.	98
Figura 27	Vazão natural mensal ao reservatório de Três Irmãos.	98
Figura 28	Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) para a série de vazão afluente ao reservatório de Camargos.	99
Figura 29	Autocorrelação nos diferentes meses para a série temporal de vazão de Camargos.	99
Figura 30	Autocorrelação parcial nos diferentes meses para a série temporal de vazão de Camargos.	100
Figura 31	Hidrograma da vazão semanal de Camargos (1931 a 2017).	104
Figura 32	Hidrograma da vazão semanal de Três Irmãos (1970 a 2017).	104
Figura 33	Histograma da vazão natural afluente à UHE Camargos (1968 a 2016).	105
Figura 34	Histograma da vazão natural afluente à UHE Três Irmãos (1970 a 2014).	105
Figura 35	Autocorrelação da série de vazão Natural UHE Camargos.	106
Figura 36	Média dos MAPEs dos treinamentos para a Usina Hidrelétrica (UHE) Camargos no período seco.	109
Figura 37	Média dos MAPEs dos treinamentos para a UHE Camargos no período úmido.	110
Figura 38	Média dos MAPEs dos treinamentos para a UHE Três Irmãos no período seco.	113
Figura 39	Média dos MAPEs dos treinamentos para a UHE Três Irmãos no período úmido.	113
Figura 40	Histograma do erro de previsão.	114
Figura 41	Funções de pertinência otimizadas para a primeira entrada associada à precipitação - ajuste para a estação seca da UHE Camargos.	114
Figura 42	Funções de pertinência otimizadas para a segunda entrada - ajuste para a estação seca da UHE Camargos.	115
Figura 43	Funções de pertinência otimizadas para a saída do sistema de inferência fuzzy treinado para o período seco.	115
Figura 44	Funções de pertinência otimizadas para a primeira entrada associada à precipitação - ajuste para a estação úmida da UHE Camargos.	115
Figura 45	Funções de pertinência otimizadas para a segunda entrada - ajuste para a estação úmida da UHE Camargos.	116

Figura 46	Funções de pertinência otimizadas para a saída do sistema de inferência fuzzy ajustado para o período úmido.	116
Figura 47	Base de regras otimizada para o período seco.	116
Figura 48	Base de regras otimizada para o período úmido.	116
Figura 49	Média dos MAPEs com o modelo $PAR(p)$ para a UHE Camargos. . . .	119
Figura 50	Média dos MAPEs com o modelo $PAR(p)$ para a UHE Três Irmãos. . .	120
Figura 51	Hidrograma das vazões observadas e calculadas pelos modelos para a UHE Camargos.	120
Figura 52	Hidrograma das vazões observadas e calculadas pelos modelos para a UHE Três Irmãos.	121
Figura 53	Gráfico de dispersão com os dados observados e previstos pelos modelos para a UHE Camargos.	122
Figura 54	Gráfico de dispersão com os dados observados e previstos pelos modelos para a UHE Três Irmãos.	123

Lista de tabelas

Tabela 1	Desvios médios absolutos do modelo de previsão de vazões Previvaz. . .	57
Tabela 2	Modelos de Previsão.	60
Tabela 3	Conjunto de regras fuzzy.	74
Tabela 4	Descrição dos postos pluviométricos selecionados para UHE Camargos. . . .	89
Tabela 5	Descrição dos postos pluviométricos selecionados para UHE Três Irmãos. . .	91
Tabela 6	MAPEs dos treinamentos para o período seco de Camargos.	107
Tabela 7	MAPEs dos treinamentos para o período úmido de Camargos.	108
Tabela 8	MAPEs dos treinamentos para o período seco de Três Irmãos.	111
Tabela 9	MAPEs dos treinamentos para o período úmido de Três Irmãos.	112
Tabela 10	Ordens dos modelos autorregressivos ajustados para as vazões de Ca- margos.	117
Tabela 11	Ordens dos modelos autorregressivos ajustados para as vazões de Três Irmãos.	117
Tabela 12	MAPE calculado com a aplicação do modelo $PAR(p)$ para UHE Ca- margos (dados de 1933 a 2016).	118
Tabela 13	MAPE calculado com a aplicação do modelo $PAR(p)$ para UHE Três Irmãos (dados de 1970 a 2016).	119
Tabela 14	Modelos autorregressivos ajustados para o cálculo das vazões semanais de Camargos.	137
Tabela 15	Modelos autorregressivos ajustados para o cálculo das vazões semanais de Três Irmãos.	138

Lista de siglas

ACL Ambiente de Contratação Livre

ACR Ambiente de Contratação Regulado

ANA Agência Nacional de Águas

ANFIS Sistema de Inferência Neuro-*Fuzzy* Adaptativa

AR *Autorregressivo*

ARIMA *Autoregressive Integrated Moving Average*

ARMA *Autoregressive Moving Average*

AMO *Atlantic Multi-decadal Oscillation*

CCEE Câmara de Comercialização de Energia Elétrica

CEPEL Centro de Pesquisas de Energia Elétrica

CMO Custo Marginal de Operação

CPTEC Centro de Previsão de Tempo e Estudos Climáticos

DAEE Departamento de Águas e Energia Elétrica

ENA Energia Natural Afluente

FCF Função de Custo Futuro

FAC Função de Autocorrelação

FACP Função de Autocorrelação Parcial

GEFS *Global Ensemble Forecast System*

INMET Instituto Nacional de Meteorologia

INPE Instituto Nacional de Pesquisas Espaciais

MA *Moving Average*

MGB-IPH Modelo de Grandes Bacias

MLP *Multi Layer Perceptron*

MPCV Modelo de Previsão de Classes de Vazões

NAO *North Atlantic Oscillation*

NCEP *National Centers for Environmental Prediction*

NSRBN *Non-linear Sigmoidal Regression Blocks Networks*

ONS Operador Nacional do Sistema Elétrico

PAR Modelo Autorregressivo Periódico

PARX Modelo Autorregressivo Periódico eXógeno

PARMA Modelo Autorregressivo Periódico Média Móvel

PDDE Programação Dinâmica Dual Estocástica

PDO *Pacific Decadal Oscillation*

PLD Preço de Liquidação das Diferenças

PMC *Perceptron* Multicamadas

Previvaz Modelo de Previsão de Vazões Semanais

RNA Redes Neurais Artificiais

SMAP *Soil Moisture Accounting Procedure*

SSARR *Streamflow Synthesis and Reservoir Regulation*

SIN Sistema Interligado Nacional

TS Takagi-Sugeno

TSM Temperatura da Superfície do Mar

UHE Usina Hidrelétrica

Sumário

1	Introdução	25
1.1	Motivação e Relevância do Trabalho	25
1.2	Objetivo do Trabalho	30
1.2.1	Objetivos Específicos	31
1.3	Organização da Dissertação	31
2	Contextualização do Trabalho	33
2.1	Mercado Brasileiro de Energia Elétrica	33
2.2	Formação de Preço de Energia Elétrica	36
2.3	Planejamento da Operação de Sistemas Hidrotérmicos	37
2.3.1	Representação das condições hidrológicas no planejamento de operação de médio prazo	41
2.3.2	Representação das séries de vazão no horizonte de planejamento da operação de curto prazo	43
2.4	Considerações Finais sobre o Capítulo	45
3	Modelos de Previsão de Vazões Naturais Afluentes	47
3.1	Introdução	47
3.2	Modelo Previvaz	48
3.3	Modelos Lineares de Estimação	50
3.3.1	Modelo Autorregressivo	50
3.3.2	Modelo Média Móvel	51
3.3.3	Modelo Autorregressivo de Média Móvel	52
3.3.4	Modelo Autorregressivo Integrado de Média Móvel	52
3.3.5	Modelo Autorregressivo Periódico	53
3.4	Modelos de previsão de vazão que incorporam os dados de precipitação	56
3.5	Considerações Finais sobre o Capítulo	62

4	Aspectos de Sistemas de Inferência Fuzzy	65
4.1	Introdução	65
4.2	Conceitos sobre Conjuntos e Lógica Fuzzy	66
4.3	Sistemas de Inferência Fuzzy	69
4.3.1	Fuzzificação	71
4.3.2	Processo de Inferência Fuzzy	72
4.3.3	Processo de Defuzzificação	79
4.3.4	Arquiteturas Fuzzy	80
4.3.5	Ajuste Estrutural de Sistemas de Inferência Fuzzy	81
4.3.6	Ajustes Paramétricos de Sistemas de Inferência Fuzzy	82
4.4	Considerações Finais sobre o Capítulo	84
5	Análise e Tratamento dos Dados de Entrada	87
5.1	Dados de Vazão Natural	87
5.2	Dados Pluviométricos	88
5.2.1	Análise de Consistência dos Dados Pluviométricos	92
5.2.2	Método de Interpolação dos dados de Precipitação	93
5.3	Correlação entre Precipitação e Vazão Natural Afluente	95
5.4	Análise da Série Temporal de Vazão Natural	97
5.5	Conclusões Finais do Capítulo	101
6	Resultados	103
6.1	Considerações Iniciais	103
6.2	Dados de Entrada do Modelo Preditivo	103
6.3	Resultados da Aplicação do Modelo Fuzzy	106
6.4	Comparação do Modelo Fuzzy com o $PAR(p)$	117
7	Conclusões e Trabalhos Futuros	125
7.1	Trabalhos Futuros	126
	Referências	129
APÊNDICE A	Processo de reconstituição da vazão natural em um trecho de rio	135
APÊNDICE B	Modelos Autorregressivos Periódicos Ajustados	137

Introdução

1.1 Motivação e Relevância do Trabalho

O sistema hidrotérmico brasileiro possui tamanho e características que permitem considerá-lo único em âmbito mundial. É considerado um sistema de grande porte, que abrange aproximadamente 158 aproveitamentos hidrelétricos distribuídos ao longo de 16 bacias hidrográficas nas diferentes regiões do país (MACEIRA et al., 2018).

As condições geológicas contribuem para que o Brasil possua um sistema hidrotérmico de energia com predominância de geração de fonte hídrica. Em termos de produção de energia elétrica, a representatividade das hidrelétricas foi de cerca de 71% do total de energia elétrica média produzida entre os anos de 2015 e 2017, como ilustra a Figura 1.

Figura 1: Composição da geração de energia elétrica por fonte.

Fonte: Elaboração própria a partir dos dados do Operador Nacional do Sistema Elétrico (ONS) para o período de 01/01/2015 a 31/12/2017 (ONS, 2018b).

Em razão da predominância de geração por meio de fonte hídrica, espera-se que a variabilidade das vazões fluviais interfiram significativamente na oferta de geração disponível. Por existir um limite na capacidade de armazenamento dos reservatórios e os níveis dependerem das condições climáticas futuras, existem incertezas relacionadas ao suprimento de energia elétrica, tornando assim o país sensível às condições hidrológicas.

Ao longo dos anos, progressivamente, se verifica uma redução da flexibilidade de operação do sistema elétrico, atribuída à entrada de energia intermitente na matriz energética e a perda gradativa da capacidade de regularização dos reservatórios. Observa-se uma maior variabilidade da energia hídrica disponível entre os períodos úmidos e secos, com afluições em algumas usinas significativas em certos períodos do ano, mas praticamente nula no período complementar.

O crescimento do mercado consumidor de energia combinado a implantação de usinas sem reservatórios tornam o sistema hidrotérmico cada vez mais vulnerável às condições de partida dos reservatórios dos aproveitamentos hidrelétricos, de modo que é necessário aprimorar o gerenciamento dos recursos hídricos. Em razão destas características, a qualidade da previsão das vazões torna-se vital não só para o gerenciamento dos usos múltiplos da água mas também para a programação do despacho de geração do sistema elétrico brasileiro.

A coordenação da operação do sistema elétrico é realizada centralizadamente pelo ONS. Entre suas atribuições está a determinação de metas de geração para cada usina, a cada etapa, que minimize o valor esperado dos custos globais de produção de energia elétrica (JARDIM; MACEIRA; FALCAO, 2001). A definição do melhor arranjo de geração envolve um problema de otimização complexo, pois é de grande porte, dinâmico e não linear, sujeito a incertezas das afluições futuras e as interdependências espaciais e temporais entre as decisões operativas (PEREIRA, 1989).

A interdependência espacial resulta do forte acoplamento hidráulico entre as unidades geradoras construídas em uma mesma cascata de rio, o que faz com que a estratégia de operação de uma usina hidrelétrica interfira na operação das usinas à jusante. Além disso, como em todo sistema hidrotérmico, tem-se o acoplamento temporal da operação, uma vez que uma decisão tomada em um estágio qualquer tem impacto na operação dos estágios subsequentes. Há, portanto, a necessidade de uma cuidadosa coordenação da operação, tanto para que o sistema seja eletricamente seguro quanto para que os recursos sejam aproveitados de forma eficiente. Por conta da complexidade do problema, para auxiliar a programação da operação do sistema elétrico, faz-se uso de modelos computacionais de otimização (CEPEL, 2018).

Os modelos computacionais utilizados no planejamento da operação do sistema elétrico recebem como insumo informações relacionadas as condições futuras de abastecimento do sistema, ou seja, são incorporadas as previsões de carga, a configuração da rede, a disponibilidade dos recursos de geração térmica e hídrica, estando esta última diretamente relacionada as previsões das vazões afluentes aos aproveitamentos hidrelétricos.

A fim de alimentar os modelos computacionais, o ONS utiliza previsões de vazões de curto e médio prazo. No médio prazo, com objetivo de capturar a variabilidade das afluições às usinas e incorporar esta volatilidade nos modelos computacionais, considera-se para propósito de otimização em um horizonte de 5 anos (MACEIRA et al., 2018), cenários

sinéuticos de afluências, gerados através de um modelo autorregressivo periódico (CEPEL, 2006). No horizonte de curto prazo, que corresponde até 6 semanas, consideram-se previsões determinísticas para cada aproveitamento hidrelétrico.

As previsões de vazões de curto prazo são obtidas a partir do Modelo de Previsão de Vazões Semanais (Previvaz) (MACEIRA et al., 1999). O Previvaz é um modelo estocástico que considera como informação apenas a série histórica de vazão semanal. A cada semana, o modelo Previvaz analisa a série de vazão natural de cada aproveitamento hidrelétrico e seleciona um modelo adequado, dentre as 94 alternativas de modelos lineares pré definidos, para realização das previsões de vazões semanais (CEPEL, 2009). Essas alternativas baseiam-se nos modelos de séries temporais propostos por Box e Jenkins, que são modelos autorregressivos com ou sem componente de média móvel (*Autorregressivo* (AR) e *Autoregressive Moving Average* (ARMA), respectivamente), construídos com base na análise estatística do histórico de vazão.

De forma complementar ao processo de previsão de curto prazo, em alguns aproveitamentos hidrelétricos, é realizada a previsão de vazão para a primeira semana por modelos que consideram além dos dados observados as previsões de precipitação no horizonte de 10 dias, fornecidas pelos modelos numéricos regional ETA, do Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) (MOREIRA; GUILHON; ROCHA, 2007) e pelo *Global Ensemble Forecast System* (GEFS) do *National Centers for Environmental Prediction* (NCEP) .

Dentre as abordagens que incorporam a previsão de precipitação, estão os modelos físicos e os sistemas inteligentes (ONS, 2018a). Entre os modelos físicos estão o modelo *Soil Moisture Accounting Procedure* (SMAP), aplicado atualmente nas bacias hidrográficas do rio Grande, Paranapanema, Paranaíba e Paraná, e o Modelo de Grandes Bacias (MGB-IPH), aplicado ao trecho incremental entre a UHE Itumbiara e a UHE São Simão, localizada na bacia do rio Paranaíba (ONS, 2018a). A abordagem dos modelos físicos para representação do comportamento hidrológico segue a mesma linha de pesquisa: ambos consideram o balanço hídrico no solo para a determinação do escoamento superficial e subterrâneo que compõem a vazão incremental natural dos trechos de rios modelados (LOPES; BRAGA; CONEJO, 1982).

Apesar dos modelos físicos apresentarem bons resultados para as bacias hidrográficas da região sudeste do país, observam-se algumas limitações para sua aplicação. Estes modelos exigem o conhecimento sobre as características físicas da sub-bacia a ser modelada, como a delimitação da área de drenagem e totais diários de evapotranspiração potencial. Além disso, seu desempenho para usinas com grande variabilidade de vazões mostra-se baixo, conforme análise feita para os aproveitamentos hidrelétricos do rio Iguaçu (UFF, 2005). No estudo elaborado pelo ONS, observaram-se erros percentuais médios de previsão de vazão entre 31% e 47% com a aplicação do SMAP para a bacia do rio Iguaçu no período entre 2002 e 2003 (UFF, 2005).

Dentre as técnicas que fazem uso de inteligência computacional, citam-se os modelos de Redes Neurais Artificiais (RNA) e sistemas de inferência fuzzy aplicados à previsão de vazão das usinas localizadas nas bacias hidrográficas da região nordeste e sul do país. As RNAs do tipo construtivas *Non-linear Sigmoidal Regression Blocks Networks* (NSRBN) são utilizadas para a previsão de vazão incremental à UHE Três Marias na bacia do rio São Francisco (ONS, 2009), enquanto os sistemas de inferência fuzzy compõem a cadeia de modelos de previsão de vazão das UHE Foz do Areia, UHE Jordão e UHE Salto Osório, na bacia do rio Iguaçu (UFF, 2005).

Uma alternativa híbrida que combina os resultados de sistemas inteligentes com os modelos lineares é fornecida pelo Modelo de Previsão de Classes de Vazões (MPCV), aplicado atualmente a bacia do rio Uruguai. O MPCV, diferentemente dos modelos anteriormente citados, não fornece uma previsão precisa, mas sim, determina limites de classe de vazões esperadas para semana seguinte, de modo a modificar o critério de escolha do melhor modelo linear para previsão de vazão, definido no algoritmo do Previvaz. (ONS, 2007).

Deve-se ressaltar, que apesar dos modelos citados anteriormente terem sido incorporados à previsão de vazão, tradicionalmente a classe de modelos lineares continua sendo aplicada para a modelagem das séries temporais de vazões na programação eletroenergética. Os modelos que integram as previsões de precipitação foram acoplados ao processo de previsão de vazão pelo ONS apenas para a estimativa da primeira semana e em algumas usinas hidrelétricas (ONS, 2018a).

Nos aproveitamentos hidrelétricos para os quais não há um modelo que replique o processo de transformação de chuva em vazão, faz-se uso da classe de modelos lineares construídos no algoritmo do Previvaz. A exemplo, citam-se as usinas da bacia hidrográfica do rio Tocantins, da bacia do rio Paranaíba (jusante de São Simão), bacia do rio Jacuí e bacia do rio Tietê.

A análise de desempenho disponível mostra que o modelo Previvaz não é eficiente, uma vez que seus resultados apresentam grandes desvios com relação aos dados observados (MOREIRA; GUILHON; ROCHA, 2007). Isso ocorre porque os modelos lineares são de estrutura muito simples e não são capazes de representar a não linearidade existente no processo hidrológico. Além da inserção de uma modelagem não linear, há de se destacar a necessidade de incorporação das previsões de precipitação na estimativa de vazão fluvial, dada a relação conhecida entre a altura e duração pluviométrica com a variável hidrológica.

Desde a década de 1990, estudos apontam a melhoria do desempenho de modelos de previsão de vazão por meio da inserção de variáveis exógenas para a descrição do processo hidrológico. Em escala mensal, as pesquisas buscam, por exemplo, a influência de fenômenos meteorológicos no regime de chuvas em determinadas regiões do país com objetivo de antecipar as condições hidrológicas futuras. Trabalhos como de Silveira et al.

(2017) sugerem a inserção de índices climáticos, como *Pacific Decadal Oscillation* (PDO), *Atlantic Multi-decadal Oscillation* (AMO) e *North Atlantic Oscillation* (NAO), para a previsão de vazão em escala mensal nas usinas hidrelétricas localizadas nas regiões norte e nordeste do Brasil. Outros trabalhos como de Lima e Lall (2007) avaliam a inclusão da Temperatura da Superfície do Mar (TSM) em regiões específicas dos oceanos Atlântico e Pacífico para determinação da variabilidade das vazões da UHE Sobradinho. Neste trabalho o Modelo Autorregressivo Periódico eXógeno (PARX) é comparado, em termos de acurácia, com o Modelo Autorregressivo Periódico (PAR) e AR, permitindo com que os autores comprovem uma melhora significativa na previsão de aflúências com até quatro meses de antecedência, a partir da inclusão da variável exógena.

Embora diversas variáveis possam ser avaliadas, a precipitação é a variável exógena mais comum no uso de modelos de previsão de vazões semanais. Para modelar o processo de transformação de chuva em vazão, observa-se na literatura diversos trabalhos que aplicam inteligência computacional. Os estudos existentes mostram resultados promissores com RNAs, inclusive melhores que os obtidos com outros modelos hidrológicos (ONS, 2009) (VALENÇA; LUDERMIR, 2001). Em escala global, destaca-se o trabalho de Anmala, Zhang e Govindaraju (2000), que aplica as redes neurais artificiais *feedforward* e redes recorrentes na estimativa de aflúência em três bacias hidrográficas no Kansas. Em termos de bacias regionais, Sacchi (2009) implementa diferentes modelos de previsão de vazão aplicados a UHE Furnas. Este trabalho testa as redes *Self-Organizing* NARX, SONARX-RBF, rede *Echo State Network* (ESN) e Sistema de Inferência Neuro-Fuzzy Adaptativa (ANFIS) para a previsão de vazão mensal. O autor realiza um pré-processamento da série temporal de vazão e decompõe a série temporal em série de amplitude e série de ângulo de fase, de modo que a previsão é realizada para cada parte da série.

Outros trabalhos como de Valença e Ludermir (2001) avaliam a viabilidade do uso das redes neurais na previsão de vazões semanais para a bacia do rio São Francisco. Neste trabalho, os autores comparam os resultados do modelo linear PARMA com os obtidos com os modelos de redes neurais *Perceptron* Multicamadas (PMC), redes recorrentes Elman e as redes NSRBN para a UHE Sobradinho. Em termos de erro percentual médio absoluto, as redes NSRBN apresentaram melhores resultados, ficando por último os modelos lineares.

Apesar das publicações mostrarem uma boa capacidade preditiva com RNAs, como ponto negativo, destaca-se a ausência de uma relação clara de causa e efeito com esta abordagem. Neste sentido, na busca por uma solução de modelagem mais intuitiva, adotam-se os sistemas de inferência fuzzy para previsão de vazão. Trabalhos como o de Luna, Soares e Ballini (2007) avaliam o desempenho do modelo fuzzy em comparação com os modelos lineares. Os resultados desta pesquisa mostram um desempenho superior do modelo fuzzy tanto para a previsão de uma semana quanto para um horizonte maior de previsão.

Um trabalho semelhante, desenvolvido por Ballini, Soares e Andrade (2001) defende a modelagem de vazão hidrológica a partir de modelos fuzzy, realçando a característica de tolerância à imprecisão, fundamental para a previsão de vazões hidrológicas, que normalmente apresentam elevado ruído de medição.

O acoplamento bem sucedido de sistemas de inferência fuzzy com redes neurais, representado pelo modelo ANFIS (JANG, 1993), também tem sido sugerido em diversos trabalhos que visam a aplicação em sistemas hidrológicos (NAYAK et al., 2004), (HILDE et al., 2005), (Chin-Hui Liu et al., 2006). Em Chin-Hui Liu et al. (2006), por exemplo, se aplica esta ferramenta para previsão de vazão em um intervalo de dez dias na bacia do rio Kao-ping. O autor faz uso de um sistema ANFIS para o mapeamento do processo hidrológico, que recebe como entrada dados de precipitação.

Os trabalhos resumidamente apresentados, apesar de divergirem quanto às técnicas de seleção das entradas, topologias e os aproveitamentos hidrelétricos considerados, chegaram a bons resultados, demonstrando que as RNAs e sistemas de inferência fuzzy fornecem bons resultados na modelagem de vazões fluviais comparando-se aos modelos propostos por Box e Jenkins.

Considerando as limitações dos modelos lineares, acredita-se que possa haver um ganho elevado nos resultados das previsões de vazões semanais aplicadas à programação da operação do sistema elétrico, na medida em que se incorporam as previsões de precipitação e técnicas alternativas para estimação de séries temporais.

Neste contexto, dada a competência de sistemas inteligentes no mapeamento de processos, com o objetivo de obter uma estrutura com alto grau de interpretabilidade da solução fornecida pelo modelo, utiliza-se para a modelagem hidrológica um sistema de inferência fuzzy do tipo *Mamdani*, no qual se aplica o método de ajuste estrutural e paramétrico proposto por Flauzino (2007).

1.2 Objetivo do Trabalho

Tendo como premissa a segurança do fornecimento de energia de forma economicamente viável, a programação da operação do sistema elétrico recebe uma atenção proporcional à sua grande importância e requer o desenvolvimento de mecanismos capazes de descrever o comportamento das vazões hidrológicas nas usinas hidrelétricas. Desta forma, é no contexto de aperfeiçoamento das previsões de vazões em usinas hidrelétricas que fundamenta-se esse trabalho. Tal proposição não apenas tem por objetivo ir ao encontro às necessidades do atual segmento, mas também de aplicar técnicas e ferramentas contemporâneas ao setor elétrico brasileiro.

1.2.1 Objetivos Específicos

- ❑ Analisar o desempenho do sistema fuzzy multicamadas na previsão de vazão natural afluyente às usinas hidrelétricas, com foco na previsão de curto prazo, ou seja, no horizonte de uma semana. Para verificar a precisão e robustez do método desenvolvido, serão realizados testes a partir de dados de vazões naturais da UHE de Camargos e Três Irmãos.
- ❑ Além de inserir um método não-linear, serão incorporados à previsão de vazão os dados de precipitação média na área de drenagem do reservatório de cada sub-bacia.
- ❑ Análise comparativa entre o modelo proposto e o modelo autorregressivo periódico.

1.3 Organização da Dissertação

No capítulo 2, é feita uma breve contextualização sobre o atual desenho de mercado e a cadeia de modelos utilizados no planejamento da operação do sistema elétrico brasileiro. A contextualização tem por objetivo definir o papel das previsões de vazões, sua importância na programação da operação e, consequente rebatimento sobre o mercado de comercialização de energia elétrica.

No capítulo 3, faz-se uma revisão da modelagem de séries temporais a partir de modelos lineares. A preocupação em apresentar tal metodologia vem do fato de ser esta a abordagem adotada pelo setor elétrico para a previsão das séries fluviais. Em complemento, apresenta-se o estado da arte para a previsão de vazões, ou seja, faz-se uma síntese de trabalhos que trouxeram avanços para a modelagem hidrológica.

No capítulo 4, apresenta-se uma revisão conceitual do sistema de inferência fuzzy aplicado ao problema de estimação. Esta seção, ao contrário das anteriores, é mais longa e detalhada, uma vez que o modelo fuzzy é o pilar deste trabalho.

No capítulo 5, são introduzidas as áreas de estudo, a base de dados utilizada e o método empregado para calcular a precipitação média nas áreas de drenagem, bem como uma análise das séries históricas de vazão.

No capítulo 6, serão apresentados os resultados advindos da aplicabilidade da técnica utilizada. Por último, são feitas as considerações finais a respeito deste trabalho, enfatizando quais os ganhos desta pesquisa para a previsão da série de vazões hidrológicas. Ademais, ressalta-se as futuras frentes de pesquisa que podem ser conduzidas a partir das investigações realizadas nesta dissertação.

Contextualização do Trabalho

Este capítulo tem por intuito inserir a proposta de trabalho no contexto vigente. De maneira a definir o papel das previsões de vazões, será apresentada a sua aplicação no planejamento energético e sua influência na formação do preço da energia elétrica comercializada.

2.1 Mercado Brasileiro de Energia Elétrica

O modelo do setor elétrico brasileiro considerava até 1995 empresas verticalizadas, predominantemente estatais, que abrangiam as atividades de geração, transmissão e distribuição de energia elétrica, sendo os recursos para construção de usinas, linhas de transmissão e sistemas de distribuição obtidos a partir de financiamento através de recursos públicos. Toda atividade relacionada à energia elétrica era um monopólio, o mercado era completamente regulado e todos os consumidores eram obrigados a comprar energia elétrica da concessionária que detinha a concessão do serviço de distribuição de energia elétrica (CARNEIRO; ROCHA, 2000).

Esta configuração vigorou por mais de 60 anos e só teve fim após uma forte recessão no país, que levou o Estado a um grande déficit público, tornando insustentável manter a expansão da geração e a qualidade na transmissão e distribuição de energia elétrica. A conjuntura ensejou constantes mudanças na indústria de suprimento de eletricidade, caracterizadas pela busca de uma nova configuração que sustentasse uma trajetória de longo prazo de redução de custos e melhorias dos serviços (LOUREIRO, 2009).

Neste contexto, criaram-se mecanismos para atrair o capital privado. A fim de viabilizar o processo, foi preciso reestruturar o setor elétrico separando as atividades potencialmente competitivas, de geração e comercialização, das atividades de transmissão e distribuição de energia, caracterizadas pela situação de monopólio natural (SILVA, 2001). Este processo foi denominado de desverticalização.

A separação das atividades de geração, transmissão e distribuição permitiu posteriormente, a competição na comercialização de energia elétrica, de modo que é possível

hoje, no ambiente de contratação livre, a compra e venda de energia elétrica a partir de contratos bilaterais livremente negociados entre os agentes.

No atual desenho de mercado, para a celebração de contratos de compra e venda de energia existem dois ambientes distintos de comercialização, o Ambiente de Contratação Regulado (ACR) e o Ambiente de Contratação Livre (ACL). O primeiro é um ambiente estritamente regulado, concebido para o atendimento dos consumidores cativos das distribuidoras, que pagam uma tarifa pela energia consumida (MAGALHÃES, 2009). Já o segundo segmento é um ambiente livre de contratação de energia, onde participam consumidores de grande porte, que tem liberdade de escolha de seu fornecedor de energia elétrica independente se ele está associado ou não à empresa que lhe presta o serviço de distribuição.

No mercado de contratação livre os agentes tomam para si a responsabilidade de gerir suas compras e os riscos associados à contratação de energia. Os agentes negociam bilateralmente volumes e preços de contratos financeiros, de modo a mitigar o risco de exposição à volatilidade dos preços do mercado de curto prazo.

O mercado de curto prazo ou mercado *spot*¹, como é conhecido, funciona como uma câmara de compensação onde a diferença entre o fluxo físico e o fluxo financeiro de energia são contabilizados e liquidados. Mensalmente, a Câmara de Comercialização de Energia Elétrica (CCEE) realiza uma liquidação financeira do mercado de curto prazo, que representa o acerto de eventuais diferenças entre a energia medida e a contratada pelos agentes. Os débitos e créditos, apurados pelo processo de contabilização das operações do mercado, são valorados ao Preço de Liquidação das Diferenças (PLD)².

O PLD é definido semanalmente a partir do valor do Custo Marginal de Operação (CMO), que representa o custo variável da usina mais cara despachada para atender à demanda (CEPEL, 2006). O cálculo do CMO é informado junto com a programação do despacho de geração, ou seja, de forma antecipada à operação real do sistema. Para auxiliar na definição do despacho de geração e, conseqüentemente, na determinação do custo marginal do sistema, consideram-se modelos de otimização que fazem uso das previsões de vazões às usinas hidrelétricas para a determinação da oferta de geração futura.

Assim, fica evidente que as previsões de vazões são importantes não só para a programação da operação do sistema elétrico, como também interferem na formação do preço da energia elétrica comercializada no mercado de curto prazo. Uma das maiores preocupações dos agentes do setor elétrico, em especial dos que o operam no ACL, refere-se à volatilidade e à imprevisibilidade do PLD. Isso porque o PLD apresenta graus de volatilidade que estão relacionados às condições hidrológicas do sistema.

¹O mercado livre difere do *spot*. O primeiro é um ambiente de contratos, prazos, preços e quantidades, enquanto o segundo é um ambiente de liquidação a posteriori entre contratos e consumos medidos.

²O PLD é determinado semanalmente para cada patamar de carga com base no Custo Marginal de Operação, seu valor é limitado por um preço máximo e mínimo vigentes para cada período de apuração.

Historicamente, o preço de curto prazo tem apresentado variações acentuadas ao longo do tempo, conforme mostra a Figura 2.

Figura 2: Histórico do preço médio mensal calculado para o submercado Sudeste.

Fonte: Elaboração própria a partir dos dados da CCEE.

A volatilidade no preço da energia elétrica pode ser atribuída, em parte, à própria característica física do sistema e, portanto, decorre da natureza aleatória da disponibilidade de recursos hídricos, representado no gráfico da Figura 2 pela Energia Natural Afluente (ENA)³ do submercado Sudeste.

Analisando o histórico do PLD, observa-se que as condições hidrológicas instantâneas repercutem rapidamente no preço de curto prazo. Essa volatilidade de preços é atribuída principalmente às previsões de vazões nos modelos computacionais de otimização.

Devido à imprevisibilidade no longo prazo, as afluências futuras são representadas por cenários sintéticos gerados por um modelo autorregressivo. Neste caso, as projeções são influenciadas pelas suas variações momentâneas, de modo que as observações de vazões passadas influenciam diretamente a estimativa para os meses futuros (CEPEL, 2006).

Dado que o preço de curto prazo reflete as condições de abastecimento do sistema, pode-se dizer que seu valor está intimamente relacionado com as estimativas de afluências.

O impacto das condições hidrológicas no preço de energia elétrica fica evidente se analisarmos o cenário de escassez de chuvas no país em 2014, que levou o CMO a valores acima do preço máximo vigente na época. O preço da energia só voltou para patamares mais baixos em 2015, quando houve uma redução artificial do preço por conta de mudanças regulatórias no setor, que reduziram o preço máximo da energia de R\$ 822/MWh para R\$ 388/MWh.

³A ENA representa a transformação das vazões naturais afluentes a cada usina em energia elétrica, MW médios. Seu valor é obtido pela multiplicação da vazão natural pela produtividade da usina em MWmed/m³/s, supondo que todos os reservatórios estejam a 65% de armazenamento. Este artifício matemático permite totalizar os resultados por cada bacia e subsistema do SIN.

Assim, embora o uso mais visível das previsões de vazões esteja na programação da operação do sistema elétrico, não se deve ignorar a sua importância para o mercado de energia.

2.2 Formação de Preço de Energia Elétrica

O Brasil, em razão das características de seu sistema elétrico e da base de recursos do país, adota o modelo de despacho centralizado denominado "*Tight pool*". Este é um modelo no qual o equilíbrio de mercado é determinado indiretamente pela utilização de modelos computacionais que visam otimizar os recursos de geração (ARFUX, 2011).

No modelo incorporado, os geradores hidrelétricos declaram ao ONS a disponibilidade de cada central geradora, enquanto os geradores termelétricos informam, além da disponibilidade das centrais, os respectivos custos variáveis de geração. De posse destes dados e de modelos computacionais de otimização, o operador do sistema define quais usinas devem gerar para abastecer a carga e qual o custo associado a esse despacho, de modo a consolidar o preço da energia a ser comercializada no mercado de curto prazo.

Em sistemas puramente térmicos, o despacho de geração é mais simples sob o ponto de vista energético. Nestes sistemas, o equilíbrio de mercado pode ser determinado diretamente pelas ofertas de preço e disponibilidade dos agentes produtores, que são empilhados em ordem crescente de preço até que a demanda seja totalmente atendida (SILVA, 2001). Neste caso, o custo marginal de operação pode ser associado ao custo do combustível utilizado pela térmica marginal, o que fornece um mecanismo natural para a formação do preço da energia elétrica.

A simplicidade na definição da política operativa dos sistemas puramente térmicos está no fato de que uma decisão operativa de uso do combustível hoje não afeta a capacidade de geração da usina no futuro. Diferentemente do ocorre nos sistemas hidrotérmicos, para os quais se deve considerar no despacho de geração as consequências futuras do uso dos recursos disponíveis nos reservatórios (SILVA, 2001).

No caso de um sistema predominantemente hidráulico como o brasileiro, a água se torna o principal combustível e, apesar de econômico, seu suprimento é incerto, já que existe um limite na capacidade de armazenamento dos reservatórios e seu nível depende das condições climáticas futuras. Se considerássemos para o despacho apenas o critério de menor custo, haveria a máxima utilização da energia hidrelétrica, visto que é a mais econômica. Contudo, esta premissa resulta em maiores riscos de racionamento no futuro. Como resultado, faz-se necessário a utilização de modelos de otimização que definam o valor da água armazenada e o custo benefício de sua utilização comparado com o custo de geração térmica.

Uma correta gestão da produção de energia elétrica só é possível a partir de um planejamento da operação olhando para um horizonte de mais longo prazo, o que por

sua vez só é viável caso existam mecanismos para a determinação das condições futuras de abastecimento do sistema, como por exemplo, as previsões de vazões que auxiliam no conhecimento antecipado da quantidade de água que estará disponível nos reservatórios.

Na próxima seção é feita uma descrição detalhada do planejamento da operação do sistema hidrotérmico brasileiro. Abordam-se os modelos de otimização utilizados para definição do despacho de geração e as diretrizes a serem tomadas nos diferentes horizontes de tempo de planejamento. Serão definidos os insumos utilizados pelos modelos computacionais, dando especial atenção à previsão de vazão, interesse dessa dissertação.

2.3 Planejamento da Operação de Sistemas Hidrotérmicos

O planejamento da operação de sistemas hidrotérmico tem como objetivo determinar uma política de operação que especifique como as usinas térmicas e hidráulicas devem ser operadas, para que os recursos disponíveis para geração de energia elétrica sejam utilizados de forma eficiente. Ressalta-se, entretanto, que o problema de se alocar da melhor maneira possível, no tempo e espaço, os recursos disponíveis é complexo devido aos vários aspectos, tais como: a estocasticidade das vazões afluentes, a interconexão das usinas hidroelétricas em uma mesma cascata de rio, as restrições de transmissão do sistema elétrico e as características operacionais não lineares das usinas hidrelétricas, além das interdependências espaciais e temporais entre as decisões operativas (CEPEL, 2018).

O acoplamento temporal entre as decisões operativas aliado aos longos tempos de maturação exigidos pela expansão da geração e da transmissão, faz com que a operação em tempo real seja precedida de estudos de planejamento e programação para garantir o menor custo de operação a longo prazo (ARFUX, 2011).

De modo a encontrar uma solução que atenda aos critérios de mínimo custo operativo, não só no presente estágio como também para os demais meses, faz-se um planejamento da operação considerando um horizonte de até 5 anos (CEPEL, 2009). De modo, que a otimização da operação do sistema se dá pela minimização do custo total de operação, definido pela soma do custo imediato e pelo custo futuro esperado para os demais estágios do horizonte de planejamento, que podem chegar até 59 meses.

Neste ponto, enfatiza-se a divergência existente entre os objetivos de minimizar o custo operativo e maximizar a segurança de suprimento de energia. Por um lado, atendendo ao critério de mínimo custo, a operação mais econômica é somente acionar as térmicas quando estamos na iminência de um racionamento. Mas, obviamente, isso deteriora a segurança de suprimento de energia. Sob outra perspectiva, o risco de déficit de energia pode ser minimizado com acionamento de térmicas de maneira ininterrupta. No entanto, a energia advinda de usinas termoelétricas tem um elevado custo financeiro, se comparado com as fontes hidráulicas. Logo, isso tornaria a operação mais onerosa. Sendo assim, de modo a

atender ambos os critérios e portanto garantir o suprimento de energia, é incorporado nos modelos de otimização um valor alto de penalidade como restrição ao não atendimento do mercado de energia, o que eleva o custo de operação em cada estágio conforme o corte de carga.

Ao considerarmos essa penalidade, o problema de planejamento pode se resumir a um problema de minimização do custo operativo, que é alcançado através do correto gerenciamento da água armazenada nos reservatórios.

A Figura 3 apresenta de maneira simplificada as principais características do problema operativo resolvido pelos modelos computacionais de otimização para cada estágio do planejamento. O problema envolve três fases distintas: a primeira é a da decisão, a segunda é a observação das aflúências futuras e a terceira o exame das consequências operativas.

Figura 3: Processo de decisão operativa com incertezas nas vazões.

Fonte: Adaptado de Silva (2001).

Na primeira situação, se o operador decidir manter o nível dos reservatórios elevados, usando geração termelétrica preventivamente, e ocorrerem altos índices pluviométricos, haverá vertimento no sistema, o que representa desperdício de recurso e, consequentemente, aumento no custo operativo. Todavia, se esta decisão é tomada e as vazões afluentes forem baixas, a operação terá sido eficiente.

Se por outro lado a decisão operativa for de deplecionar as reservas de energia hidrelétrica com objetivo de minimizar os custos térmicos e, no futuro, um baixo regime pluviométrico vier a acontecer, não haverá água suficiente para o sistema recuperar seus níveis de armazenamento, sendo necessário a utilização de geração térmica adicional com custos mais elevados para atender a carga futura. Em casos extremos, essa situação pode levar à necessidade de corte de carga, o que deixa evidente a importância do planejamento de operação do sistema elétrico de forma eficiente e confiável, não apenas pela necessidade

de se obter uma operação econômica do sistema elétrico, mas também como produto de valor estimável para assegurar o abastecimento do consumo.

Como no exemplo simplificado da Figura 3, a decisão por parte do operador do sistema deve ser tomada sob incerteza, sendo o sucesso da escolha condicionada à qualidade das previsões de afluências. Como em períodos mais longos é praticamente impossível saber quais serão as realizações de vazões, a escolha da melhor decisão pode ser definida com base na probabilidade de ocorrência das afluências.

Devido à incerteza associada às afluências futuras, ao número de aproveitamentos e ao número de estágios, adotam-se simplificações para a solução do problema da operação de sistemas hidrotérmicos. A adoção de simplificações consiste na divisão em subproblemas com diferentes horizontes de estudo, levando-se em conta o nível de detalhamento do sistema elétrico e a representação das afluências.

A desagregação do planejamento da operação em etapas é realizada em horizontes distintos: longo prazo, curto prazo e programação diária (CEPEL, 2006). A cadeia de modelos utilizados nos dois primeiros horizontes de planejamento do sistema elétrico é apresentada na Figura 4.

Figura 4: Cadeia de modelos utilizados no planejamento energético.

Fonte: Adaptado de CEPEL (2006).

Para o horizonte de planejamento de longo prazo e curto prazo adotam-se os modelos computacionais NEWAVE e DECOMP, desenvolvidos pelo Centro de Pesquisas de Energia Elétrica (CEPEL). Para solução do problema de otimização os modelos recebem como parâmetros a demanda, os custos de operação das usinas térmicas, o custo de déficit, o cronograma de novos aproveitamentos, a disponibilidade de equipamentos de geração e transmissão. A previsão destes parâmetros é complexa e sujeita a grandes incertezas. Ainda assim, de modo a viabilizar o cálculo de operação, consideram-se que todos estes dados, à exceção das condições hidrológicas, são variáveis determinísticas, sendo seus valores ajustados mensalmente.

O modelo NEWAVE foi desenvolvido para tratar o planejamento da operação energética do Sistema Interligado Nacional (SIN) quando são considerados horizontes de planejamento plurianuais de até cinco anos. Neste horizonte, as usinas pertencentes aos subsistemas com as mesmas condições hidrológicas são agregadas em reservatórios equivalentes de energia (CEPEL, 2006), de maneira que o menor número de unidades torna viável representar a incerteza relacionada às afluências dos subsistemas através de cenários hidrológicos.

O principal resultado obtido do modelo de longo prazo é uma análise do valor econômico da energia armazenada, que é repassada para o DECOMP através de uma função multivariada, chamada Função de Custo Futuro (FCF), apresentada na Figura 5.

Figura 5: Representação da Função de Custo Futuro.

Fonte: CEPEL (2006)

A FCF informa como varia o valor esperado do custo de operação dos demais estágios, α_T , conforme o volume armazenado nos reservatórios V_T . Observa-se que quanto menor for o volume armazenado nos reservatórios mais alto é o custo associado aos demais estágios.

A otimização do modelo DECOMP consiste em definir as diretrizes para a operação física do sistema elétrico, ou seja, metas de geração térmica e hídrica por usina que minimizem o custo de operação. A decisão ótima global é obtida a partir da decomposição

do problema multi-estágio em soluções recursivas de um único estágio. Deve-se decidir no primeiro mês, com base na energia armazenada inicial e na energia afluenta esperada, a quantidade de energia a ser turbinada que minimize o custo total de operação, calculado pela concatenação do custo da alternativa no próprio mês com o custo dos demais estágios informado pela FCF.

A FCF apresentada na Figura 5 é uma aproximação da função de custo futuro a ser considerada no cálculo da estratégia ótima de operação. Para aplicações realistas, que consideram um grande número de cenários, seria impraticável a solução do problema para a combinação de todo o domínio contínuo das variáveis, devido ao crescimento exponencial do espaço de estados. De modo a evitar o problema de maldição da dimensionalidade, na solução de planejamento da operação hidrotérmica adota-se uma aproximação para a FCF. Calcula-se a solução para um conjunto de valores das variáveis de estado e para cada ponto, ajusta-se uma reta cujos coeficientes angulares são obtidos a partir das taxas de variação do custo futuro nas vizinhanças dos estados (MACEIRA et al., 2018).

Assim, da etapa de longo prazo, o que se busca é extrair um resultado indicativo sobre as condições de atendimento no futuro, com base no armazenamento e nas projeções de carga e oferta de geração disponíveis no horizonte de planejamento.

Na medida em que o horizonte de estudo diminui, a incerteza sobre as afliências futuras também diminui, porém aumenta a necessidade de uma representação mais detalhada das usinas hidrelétricas, térmicas e da rede de transmissão. Assim, no planejamento da operação de curto prazo faz-se a representação do parque de geração de forma individualizada.

O acoplamento entre o DECOMP e NEWAVE se dá ao final do segundo mês, com base nos estados de armazenamento atingidos e nas afliências consideradas nos dois primeiros meses (CEPEL, 2018). Logo, para consultar a função de custo futuro gerada pelo modelo de longo prazo, faz-se necessário conhecer quais as afliências esperadas nos dois primeiros meses, o que é possível a partir da utilização de previsões de vazões para cada usina hidrelétrica do sistema elétrico brasileiro.

Conforme o uso que se faça, trabalha-se com diferentes horizontes e intervalos de discretização. No setor elétrico, trabalha-se com previsões de médio prazo, que corresponde a um horizonte de até 5 anos em base mensal, e curto prazo, que abrange um horizonte de dois meses, com previsões de determinísticas discretizadas em base semanal para o primeiro mês.

2.3.1 Representação das condições hidrológicas no planejamento de operação de médio prazo

No médio prazo, a precisão da representação individualizada de usinas hidrelétricas perde a importância frente às incertezas das vazões futuras. Sendo assim, as 158 usinas hi-

hidrelétricas existentes são agregadas em 12 reservatórios equivalentes de energia (MACEIRA et al., 2018).

A agregação dos reservatórios das usinas em sistemas equivalentes de energia reduz o número de variáveis e a dimensão do problema. Se não fosse adotado esse tipo de simplificação, o problema de otimização tornaria-se custoso do ponto de vista computacional, devido ao número de estágios, que chega até a 60 meses, e à necessidade de se considerar na estratégia de operação um número significativo de combinações de afluições futuras.

Contudo, com o processo de agregação dos reservatórios surge a necessidade de realizar uma adequação nos valores de vazão afluente as usinas hidrelétricas para a compatibilização dessa abordagem com as grandezas verificadas nos reservatórios equivalentes de energia. Dessa forma, a vazão afluente deixar de ser abordada como quantidade de água que chega a cada usina e passa a ser uma quantidade de energia associada a cada reservatório equivalente de energia, conhecida como ENA.

A maneira encontrada pelo setor elétrico para introduzir a estocasticidade das vazões na modelagem computacional foi empregar a técnica de Programação Dinâmica Dual Estocástica (PDDE), desenvolvida por Pereira (1989). Tal metodologia oferece uma abordagem para a incorporação de incertezas das vazões afluentes em problemas de otimização. Considera-se para tanto um conjunto de cenários, que representam as possíveis realizações das vazões hidrológicas futuras. A vantagem deste tipo de formulação está no fato de que o problema de planejamento da operação hidrotérmica pode ser solucionado a partir de técnicas de otimização determinísticas, considerando para isso um grande número de cenários, de maneira a capturar a natureza aleatória da variável hidrológica.

Para o cálculo da política de operação no horizonte de médio prazo, considera-se uma árvore de cenários hidrológicos gerados estocasticamente com base no registro histórico (MACEIRA; DAMAZIO, 2004), como pode ser visto na Figura 6.

Cada caminho na árvore é chamado de cenário hidrológico e cada nó representa uma possível realização da afluição. Como computacionalmente é inviável resolver um problema de otimização para a árvore completa, são sorteados 200 caminhos para representar as energias naturais afluentes futuras.

A geração de séries sintéticas de afluições mensais incorporadas no cálculo de otimização estocástica é feita a partir do modelo PAR(p), onde p é a ordem do modelo autorregressivo para cada período. As séries sintéticas são geradas a partir do sorteio de um número significativo de hipóteses dadas por a_t :

$$Z_t = \underbrace{\mu_m + \phi_1^m \sigma_m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right)}_{\text{Parte determinística}} + \underbrace{a_t}_{\text{Parte aleatória}} \quad (1)$$

em que Z_t é a série sazonal de período m ($m=12$ para série mensal), μ_m é a média sazonal do mês m , σ_m é o desvio-padrão sazonal, ϕ_i^m é i -ésimo coeficiente autorregressivo do mês m , p_m é a ordem do operador autorregressivo no período m e a_t é a série de ruídos.

Figura 6: Representação da árvore de cenários de afluências.
Fonte: Maceira e Damazio (2004)

Esta expressão é conveniente para a geração de séries sintéticas, pois permite um esquema de geração sequencial em que Z_t é obtido em função de seus valores passados. A metodologia de ajuste do modelo $PAR(p)$ será discutida com mais detalhes no capítulo 3.

2.3.2 Representação das séries de vazão no horizonte de planejamento da operação de curto prazo

No modelo de curto prazo, a determinação da estratégia de operação é feita individualmente por usina. Considera-se usualmente um horizonte de planejamento de 2 meses, sendo os dados do primeiro mês discretizados em base semanal e o segundo em base mensal. Os mesmos dados de entrada disponibilizados no modelo NEWAVE em base mensal são discretizados e corrigidos semanalmente para o horizonte de planejamento de curto prazo. A cada semana são atualizadas as informações de disponibilidade de usinas, restrições de geração e transmissão e previsões de carga e vazões.

As previsões de vazões para todos os aproveitamentos hidrelétricos considerados postos base são obtidas pelo modelo Previvaz. Este é um modelo estocástico que analisa a série histórica de vazão de cada usina hidrelétrica e seleciona, para cada semana, uma alternativa dentre as diversas opções de modelos lineares pré definidos para a determinação das próximas realizações.

As alternativas de modelos lineares baseiam-se nos modelos de séries temporais, mais especificamente em modelos ARMA e Modelo Autorregressivo Periódico Média Móvel (PARMA). Estes modelos são construídos como função da informação passada em diferentes *lags*, podendo ou não apresentar estrutura de correlação periódica. Para a obtenção da ordem dos modelos e identificação dos parâmetros, adota-se a metodologia de Box e Jenkins (BOX; JENKINS; REINSEL, 2008). Mais adiante, no capítulo 3, a metodologia de obtenção dos modelos será melhor detalhada.

Como as previsões de vazões afluentes às usinas hidrelétricas são revistas semanalmente, a cada revisão incorporam-se novas observações à série. À medida que as vazões semanais do primeiro mês vão se realizando, seus valores observados são utilizados como tendência para o modelo elaborar as previsões das semanas faltantes do mês corrente. A mesma metodologia é aplicada a cada revisão até que se complete o mês em curso.

O tratamento dado às séries de vazão no curto prazo é resumido na Figura 7. No primeiro mês adota-se uma abordagem determinística para as vazões de todas usinas hidrelétricas e no segundo mês a vazão é considerada uma variável estocástica e segue o tratamento dado no médio prazo, onde as afliências são representadas por uma árvore de cenários.

Figura 7: Esquema de representação do cenário de vazões afluentes às usinas hidrelétricas.
Fonte: CEPEL (2018)

Os cenários sintéticos de vazões no segundo mês são gerados com o programa GEVAZP. Em virtude de restrições de tempo computacional, utiliza-se o método de Amostragem Seletiva para reduzir a amostra de cenários hidrológicos (CEPEL, 2006). O método consiste em aplicar técnicas de agregação aos cenários hidrológicos com características estatísticas similares, de forma a escolher um conjunto representativo a partir da amostra original.

2.4 Considerações Finais sobre o Capítulo

Como apresentado ao longo do capítulo, para o correto gerenciamento dos recursos energéticos consideram-se no planejamento da operação do sistema elétrico brasileiro modelos computacionais de otimização que utilizam como insumo as previsões de vazões hidrológicas. No planejamento de médio prazo, horizonte de até 60 meses, devido à imprevisibilidade na determinação da vazões, são considerados cenários sintéticos de afluições para representar as possíveis realizações das vazões naturais. Conforme reduz-se o horizonte de planejamento, tem-se um maior detalhamento do sistema, de modo que no curto prazo, horizonte de até 6 semanas, as vazões naturais afluentes a cada usina são tratadas no problema de otimização como variáveis determinísticas. Ao considerarmos que esta variável é determinística no primeiro mês, espera-se uma maior confiabilidade das previsões. Neste sentido, no próximo capítulo será feito um levantamento dos métodos de previsão de vazões existentes, de modo a embasar a busca por previsões mais acuradas.

Modelos de Previsão de Vazões Naturais Afluentes

3.1 Introdução

Como apresentado no capítulo 2, as atividades relacionadas com o planejamento de operação do sistema elétrico fazem uso estratégico de previsões de vazões semanais determinísticas para cada aproveitamento hidroelétrico. Neste âmbito se encaixa o objetivo deste trabalho, que é o desenvolvimento de uma ferramenta dedicada à estimativa de vazão natural afluyente aos reservatórios de usinas hidrelétricas. Entretanto, antes de se estabelecer o escopo do desenvolvimento deste projeto, será realizada uma etapa preliminar de conhecimento da oferta tecnológica disponível para a previsão hidrológica. Assim, será apresentada neste capítulo a modelagem linear oficialmente adotada na programação e também outros métodos de previsão de vazão não lineares que incorporam dados de previsão de precipitação, com foco nos modelos físicos, redes neurais artificiais e sistemas de inferência fuzzy. O objetivo é apresentar as vantagens e deficiências de cada técnica, com vistas a contribuir na discussão sobre aperfeiçoamentos na previsão de vazões.

No sistema elétrico brasileiro, para descrever as séries de vazões fluviais utilizam-se modelos lineares univariados, cujas formulações são classificadas em modelos estacionários e modelos periódicos. Na classe de modelos estacionários estão os modelos ARMA, cuja equação de regressão e estrutura de correlação de um instante com os seus valores anteriores são consideradas constantes ao longo do tempo. A aplicação destes modelos em hidrologia, mais precisamente os modelos autorregressivos, para a modelagem de séries de vazões, ocorre desde a década de 60, com trabalhos como o de Thomas e Fiering (1962) e Yevjevich (1977).

Na literatura, a adoção de modelos autorregressivos para a modelagem de séries de vazões é recomendada por considerar um importante fenômeno denominado tendência hidrológica (SOUZA, 2013). A tendência hidrológica é o termo utilizado para se referir à tendência observada nas séries históricas de vazões naturais, de valores superiores a

média do período serem precedidos e seguidos por valores também superiores valor normal (CEPEL, 2009).

A ocorrência de tendência hidrológica é atribuída ao processo de infiltração. A resposta da bacia hidrográfica ao volume precipitado depende do nível de umidade do solo. Para semanas subsequentes com vazões altas, a superfície do terreno se mantém úmida e a infiltração de novos volumes de precipitação é pequena, aumentando o escoamento superficial direto (CEPEL, 2009). Na terminologia de séries temporais, a tendência hidrológica é conhecida como estrutura de dependência temporal, sendo quantificada pela função de autocorrelação estimada do registro histórico.

A partir da identificação de relações de dependência existentes numa série, é possível construir um modelo matemático que represente o comportamento do processo gerador da série, podendo a construção do modelo considerar uma estrutura de correlação constante ou não durante todos os períodos.

Na classe dos modelos periódicos estão os modelos PARMA, que apresentam uma equação de regressão para cada semana. Este tipo de modelo tem se mostrado flexível e popular para modelagem estocástica de vazões fluviais (HIPEL; MCLEOD, 1994). Isso porque é capaz de representar o comportamento periódico na estimação de sistemas que apresentam ciclos sazonais, como é o caso das vazões em base mensal e semanal. Nestes tipos de séries temporais, a cada período sazonal apresenta-se um conjunto de características estatísticas dadas pela média, desvio-padrão e estruturas de correlação periódica, sendo a importância desta sazonalidade dependente da bacia hidrológica (CEPEL, 2009).

É possível definir diferentes modelos de previsão considerando as diversas alternativas de classe de modelos lineares, ordem dos modelos e métodos de estimação dos parâmetros, de modo que a determinação do melhor modelo se faz a partir do processamento do algoritmo do modelo Previvaz.

3.2 Modelo Previvaz

O Modelo Previvaz é um modelo de previsão de vazão semanal constituído por um conjunto que conta com 94 alternativas de modelos lineares estatísticos estacionários ou periódicos (CEPEL, 2009). Para cada semana e cada posto fluviométrico são testadas diferentes alternativas de modelos lineares univariados para a previsão de até seis semanas. Os parâmetros dos modelos são ajustados com base na série histórica de vazão, disponível desde 1931. Logo, não se consideram nas estimativas de vazões naturais informações de estações pluviométricas ou previsões de precipitação na área de drenagem.

A determinação do modelo de previsão é realizado por um esquema de validação cruzada, onde cada série de vazão semanal é dividida em duas partes. Inicialmente, apenas a primeira parte da série é utilizada para a estimação dos parâmetros do modelo, ficando a segunda parte reservada à verificação da acurácia da modelagem. Em uma

segunda etapa, a estimação dos parâmetros passa a ser feita com a outra parte da série, ficando a primeira apenas para o cálculo do erro de previsão. Para cada parte da série é computado o erro médio das previsões. Por fim, com base no desempenho dos modelos para as duas amostras, determina-se, para cada semana, um modelo de previsão. Como não existe apenas um único modelo matemático capaz de descrever o sistema, em geral trabalha-se com uma família de modelos do tipo ARMA e PARMA, cujos parâmetros são definidos a partir da metodologia de Box e Jenkins:

Figura 8: Estágios da metodologia Box e Jenkins
Fonte: Box, Jenkins e Reinsel (2008)

Originalmente, o processo de construção do modelo matemático a partir da metodologia de Box e Jenkins envolve um ciclo iterativo, com três passos distintos, que são repetidos até que se encontre um modelo que represente o processo gerador da série. Para que o modelo construído seja adequado, os resíduos gerados na etapa de diagnóstico devem ser ruídos brancos, ou seja, erros sem nenhuma estrutura de correlação. Se esta suposição não for satisfeita, retorna-se para etapa de identificação para o ajuste de um modelo mais apropriado (BOX; JENKINS; REINSEL, 2008).

Na prática, entretanto, o modelo Previvaz adota uma abordagem diferente, onde as alternativas de modelagem são previamente definidas e incluem diferentes estruturas de modelos, métodos de pré-transformação e métodos de estimação. A alternativa mais adequada é selecionada com base nas melhores estatísticas resultantes do processo de validação cruzada (CEPEL, 2009).

O primeiro processo, descrito no fluxograma da Figura 8, consiste na transformação da

série, feita de modo a atender certas hipóteses assumidas no processo de modelagem por modelos do tipo ARMA. Uma das hipóteses assumidas é que os ruídos da série seguem uma distribuição de probabilidade normal. Para garantir a normalidade da distribuição, consideram-se pré-transformações do tipo Box&Cox ou logarítmica (BOX; COX, 1964).

Para a identificação da ordem do modelo matemático, o método empregado baseia-se no princípio da parcimônia, ou seja, no menor número de parâmetros possíveis para uma representação matemática apropriada do processo. Assim, limita-se a ordem máxima dos modelos AR(p) e PAR(p) até ordem 4 e ARMA(p) e PARMA(p) até ordem 3 (CEPEL, 2009). Para a determinação da ordem dos modelos lineares, analisa-se a Função de Autocorrelação (FAC) e a Função de Autocorrelação Parcial (FACP), que são úteis inclusive na verificação acerca da estacionariedade da série e especificação das sazonalidades existentes que devem ser potencialmente modeladas.

Uma vez determinados os modelos candidatos, deve-se proceder à estimação dos parâmetros, o que pode ser feita por meio do uso do método de máxima verossimilhança e método dos momentos (HIPEL; MCLEOD, 1994).

Posteriormente, com base nos modelos ajustados, realizam-se testes para verificação da acurácia da modelagem para amostras não contempladas no processo de estimação. O modelo com o melhor desempenho no processo de validação cruzada é selecionado para realização das previsões semanais.

3.3 Modelos Lineares de Estimação

Na literatura existem inúmeros modelos de estimação de séries temporais. Dentre os modelos lineares básicos, destacam-se os modelos AR, *Moving Average* (MA), ARMA e os modelos PARMA. Estes são modelos estocásticos que descrevem processos por uma classe de modelos lineares que, por sua vez, calculam o valor do instante presente da série associando-o linearmente, por meio de parâmetros, às informações contidas na própria série temporal.

3.3.1 Modelo Autorregressivo

O modelo autorregressivo é um subconjunto dos modelos ARMA que expressa uma série temporal como uma função linear dos valores passados. Neste caso, assume-se que o valor de cada instante da série está distante de sua média por um ruído, ϵ_t , e que o valor da série temporal para o instante t depende linearmente dos p instantes imediatamente anteriores (BOX; JENKINS; REINSEL, 2008). Formalmente, o modelo AR(p) é definido por (2):

$$Z_t = \mu + \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \dots + \alpha_p Z_{t-p} + \epsilon_t \quad (2)$$

ou

$$Z_t = \mu + \sum_{i=1}^p \alpha_i Z_{t-i} + \epsilon_t \quad (3)$$

em que Z_t é o valor da série temporal no instante t , μ é o nível da série, α_i é o i -ésimo parâmetro do modelo e ϵ_t é o erro associado ao instante t (um ruído de média zero e variância constante).

A identificação da ordem do modelo $AR(p)$ resulta do exame dos coeficientes das funções de autocorrelação (FAC) e autocorrelação parcial (FACP). A análise das funções permite estimar o comprimento de defasagem em que se observam valores significativos de autocorrelações. A função de autocorrelação é definida em (4):

$$\rho(k) = \frac{\sum_{t=k}^N (Z_t - \mu)(Z_{t-k} - \mu)}{\sum_{t=k}^N (Z_t - \mu)^2} \quad (4)$$

em que μ é a média da série temporal e k é a defasagem da autocorrelação.

Esta função estabelece o coeficiente de correlação entre o valor Z_t da série e o Z_{t-k} . Em termos teóricos, caso o coeficiente $\rho(k)$ assuma valor igual a 1, significa que o valor de Z_t pode ser perfeitamente descrito pelo valor de Z_{t-k} .

Uma vez determinada a ordem de p , estimam-se os valores dos parâmetros do modelo, $\alpha_1, \alpha_2, \dots, \alpha_p$. Como se assume uma dependência linear, a técnica dos mínimos quadrados é satisfatória no cálculo dos parâmetros (MORETTIN; TOLOI, 1981).

3.3.2 Modelo Média Móvel

O modelo de médias móveis $MA(q)$ de ordem q assume que o valor da série temporal no instante t é definido pela soma de um valor médio e a ponderação dos erros aleatórios de q períodos passados. O valor da série temporal Z_t é definido a partir da seguinte equação:

$$Z_t = \mu + \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q} \quad (5)$$

em que Z_t é o valor da série temporal no instante t , μ é o nível da série e β é o i -ésimo parâmetro do modelo. Os parâmetros β descrevem como Z_t se relaciona com o valor ϵ_t para $i = 1, 2, \dots, q$. Este modelo é usado quando há autocorrelação entre os resíduos, ou seja, existe uma relação de dependência entre o conjunto de erros em períodos passados.

O modelo $MA(q)$ dado pela equação (5) pode ser reescrito assumindo $\check{Z}_t$ como a diferença entre Z_t e a média μ , $\check{Z}_t = Z_t - \mu$, e usando o operador de retardo $\epsilon_{t-m} = B^m \epsilon_t$.

$$\check{Z}_t = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_q \epsilon_{t-q} = (\beta_0 + \beta_1 B + \dots + \beta_q B^q) \epsilon_t \quad (6)$$

em que $\beta_0 = 1$ e

$$\check{Z}_t = \beta(B) \epsilon_t. \quad (7)$$

Segundo Box, Jenkins e Reinsel (2008), o modelo média móvel supõe que a série temporal é gerada por meio de um filtro linear cuja função de transferência $\beta(B)$ tem como entrada um ruído branco, ϵ_t , e saída $\tilde{Z}_t$.

3.3.3 Modelo Autorregressivo de Média Móvel

Os modelos autorregressivos de médias móveis modelam o grau de autocorrelação entre desvios e observações defasadas. Combinando-se os modelos AR e MA, pode-se obter uma representação adequada com um número menor de parâmetros através do modelo ARMA, que é dado por (8):

$$\tilde{Z}_t = \alpha_1 \tilde{Z}_{t-1} + \alpha_2 \tilde{Z}_{t-2} + \dots + \alpha_p \tilde{Z}_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (8)$$

em que $\alpha_1, \dots, \alpha_p$ são os parâmetros autorregressivos, $\beta_1, \dots, \beta_q$ são os parâmetros de médias móveis e ϵ_t é a série de ruídos independentes com média zero e variância constante.

O modelo em ARMA pode ainda ser representado em termos de dois polinômios, os quais são obtidos a partir da aplicação do operador de retardo:

$$(1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p) Z_t = (1 + \beta_1 B + \beta_2 B^2 + \dots + \beta_q B^q) \epsilon_t. \quad (9)$$

A estimação de parâmetros do modelo ARMA é mais difícil do que o modelo AR, e envolve a resolução de equações não lineares (BOX; JENKINS; REINSEL, 2008). Para a identificação dos valores de p e q , podem ser examinadas determinadas propriedades da função de autocorrelação e autocorrelação parcial da série (MORETTIN; TOLOI, 1981).

3.3.4 Modelo Autorregressivo Integrado de Média Móvel

Os modelos anteriormente discutidos são apropriados para séries temporais estacionárias ¹. A estacionariedade da série, que implica a não alteração das propriedades estatísticas da série como média, variância e autocorrelação ao longo do tempo, é uma importante hipótese para a aplicação de modelos do tipo ARMA.

Contudo, a maior parte das séries apresenta alguma forma de não estacionariedade, como componentes sazonais ou tendências (MORETTIN; TOLOI, 1981). Como existe periodicidade anual dos processos climáticos, sugere-se que as séries de vazões discretizadas em um intervalo inferior a um ano são não-estacionárias. Para viabilizar o uso de modelos do tipo ARMA, utilizam-se transformações estatísticas que permitem remover ou filtrar a parte não estacionária da série.

A fim de tornar uma série estacionária, um tipo especial de filtro consiste em tomar diferenças sucessivas da série original. A nova série temporal, dada por W_t , é obtida a partir da seguinte operação:

$$W_t = \Delta^d Z_t = (1 - B)^d Z_t. \quad (10)$$

¹A série temporal é dita estacionária quando ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável.

Normalmente são necessárias uma ou duas diferenças para que uma série se torne estacionária (BOX; JENKINS; REINSEL, 2008). Considerando uma diferenciação de primeira ordem, teremos:

$$W_t = \Delta Z_t = Z_t - Z_{t-1}. \quad (11)$$

Para o caso de $d=2$, a segunda diferença será dada por (12):

$$W_t = \Delta^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2}. \quad (12)$$

Generalizando, a d -ésima diferença é:

$$\Delta^d Z_t = \Delta\{\Delta^{d-1} Z_t\}. \quad (13)$$

A série original é submetida ao operador de diferenças de modo a se aplicar a classe de modelos lineares estacionários à série diferenciada. Por exemplo, em (14) utiliza-se um modelo ARMA no qual Z_t é substituído pela sua d -ésima diferença.

$$\Delta^d z_t = \alpha_1 \Delta^d z_{t-1} + \alpha_2 \Delta^d z_{t-2} + \dots + \alpha_p \Delta^d z_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q}. \quad (14)$$

Ao considerar $W_t = \Delta^d z_t$, temos o processo *Autoregressive Integrated Moving Average* (ARIMA)(p, d, q), dado por (15):

$$W_t = \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_p W_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (15)$$

$$\alpha_0 W_t + \alpha_1 W_{t-1} + \alpha_2 W_{t-2} + \dots + \alpha_p W_{t-p} = \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (16)$$

ou, equivalentemente

$$\alpha(B)W_t = \alpha(B) \Delta^d Z_t = \theta(B)\epsilon_t. \quad (17)$$

Assim, o modelo ARIMA é uma aproximação para o verdadeiro processo gerador da série. Em (17), a variável Z_t é uma variável integrada de ordem d e sua d -ésima diferença segue um processo autorregressivo de ordem p e médias móveis de ordem q .

Detalhes sobre as rotinas de estimação destes modelos não serão abordados neste trabalho, mas podem ser obtidos em Box, Jenkins e Reinsel (2008).

3.3.5 Modelo Autorregressivo Periódico

Diversos processos físicos em escala mensal, dentre eles as séries de vazões fluviais, apresentam um comportamento periódico descrito pelos ciclos sazonais. Nestas séries temporais a estrutura de autocorrelação depende não somente do intervalo de tempo entre as observações, mas também do período observado.

Para a identificação de modelos apropriados para a representação do comportamento da série, deve-se analisar o conjunto de características estatísticas descritas pela média, desvio-padrão e estrutura de correlações sazonais que se repetem a cada s observações.

Em tais casos, uma abordagem possível para a representação do processo pode ser obtida a partir de uma formulação autorregressiva cujos parâmetros têm comportamento periódico. Estes modelos são chamados de modelos autorregressivos periódicos PAR(p), em que p é um vetor representado por $p = (p_1, p_2, \dots, p_m)$, em que p_i é o a ordem do i -ésimo modelo autorregressivo.

A estratégia deste método é o ajuste de um modelo autorregressivo AR(p) para cada um dos períodos da série em análise, a exemplo, para o caso de periodicidade mensal, um modelo para cada mês. O modelo PAR(p) é descrito matematicamente pela seguinte equação (MACEIRA; DAMAZIO, 2004):

$$\left(\frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \epsilon_t \quad (18)$$

em que Z_t é a série sazonal, μ_m é a média sazonal, σ_m é o desvio-padrão sazonal, p_m é a ordem do operador autorregressivo no mês m , ϕ_i^m é o i -ésimo coeficiente autorregressivo do período m e ϵ_t é a série de resíduos independentes.

Manipulando a equação (18), para isolar a variável no instante t , chega-se à seguinte formulação para a estimação da série temporal:

$$Z_t = \mu_m + \phi_1^m \sigma_m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m \epsilon_t \quad (19)$$

ou,

$$Z_t = \underbrace{\mu_m + \sum_{j=1}^{p_m} \phi_j^m \sigma_m \left(\frac{Z_{t-j} - \mu_{m-j}}{\sigma_{m-j}} \right)}_{\text{Parte determinística}} + \underbrace{\sigma_m \epsilon_t}_{\text{Parte aleatória}}. \quad (20)$$

O modelo PAR(p) é uma reconhecida ferramenta utilizada no setor elétrico brasileiro, além de ser aplicado na previsão de vazões semanais a partir do modelo Previvaz, é também empregado no planejamento de médio prazo para geração das séries sintéticas de afluições mensais incorporadas no cálculo de otimização estocástica (SOUZA, 2013). A expressão (20) é conveniente para a geração de séries sintéticas por permitir um esquema de geração sequencial em que Z_t é obtido como uma função dos valores passados $Z_{t-\tau}$, $\tau = 1, \dots, p_m$ e de um sorteio ϵ_t .

A identificação da ordem do modelo p , ou seja, a ordem do termo autorregressivo de cada mês, geralmente é realizada por meio da análise da FACP, que exige o conhecimento dos coeficientes da FAC. Para a discretização mensal, o conjunto de coeficientes das funções de autocorrelação, ρ_k^m , dos períodos $m = 1, \dots, 12$, que descrevem a estrutura de dependência temporal da série para cada mês é dado pela seguinte formulação (MACEIRA; DAMAZIO, 2004):

$$\rho_k^m = E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (21)$$

em que μ_m é a média sazonal do período m e σ_m é o desvio padrão sazonal do período m .

Os parâmetros do modelo PAR(p) para cada período m são obtidos a partir da resolução de um conjunto de equações comumente denominado de equações de Yule-Walker, que envolve o conhecimento do conjunto de coeficientes ρ_k^m dos m períodos.

Multiplicando ambos os lados da equação (18) por $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right)$ e tomando a esperança dos mesmos, temos a estrutura de dependência temporal do modelo descrita pelo conjunto de coeficientes de autocorrelação ρ_k^m dos m períodos:

$$\begin{aligned} \rho_k^m = & E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] = \phi_1^m E \left[\left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + \dots + \\ & \phi_{p_m}^m E \left[\left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + E \left[\epsilon_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right]. \end{aligned} \quad (22)$$

Ao assumir que o processo ϵ_t se trata de um ruído branco, tem-se a parcela $E \left[\epsilon_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right]$ igual a zero. Fixando-se m e variando k de 1 a p_m em (22), obtém-se para cada período um conjunto de equações como as que seguem:

$$\begin{aligned} \rho_1^m &= \phi_1^m \cdot 1 + \phi_2^m \rho_1^{m-1} + \phi_3^m \rho_2^{m-1} + \dots + \phi_{p_m}^m \rho_{p_m-1}^{m-1} \\ \rho_2^m &= \phi_1^m \rho_1^{m-1} + \phi_2^m \cdot 1 + \phi_3^m \rho_1^{m-2} + \dots + \phi_{p_m}^m \rho_{p_m-2}^{m-2} \\ &\vdots \\ \rho_{p_m}^m &= \phi_1^m \rho_{p_m-1}^{m-1} + \phi_2^m \rho_{p_m-2}^{m-2} + \phi_3^m \rho_{p_m-3}^{m-3} + \dots + \phi_{p_m}^m \cdot 1. \end{aligned}$$

As equações apresentadas são denominadas de equações de Yule-Walker periódicas, sua solução permite a determinação do j -ésimo parâmetro, ϕ_j^m , do modelo autorregressivo de ordem p_m para o mês m .

Para a determinação do j -ésimo parâmetro autorregressivo de um processo de ordem k , ou seja, $p_m = k$, reescrevem-se as equações de Yule-Walker da seguinte forma:

$$\begin{bmatrix} \phi_{k1}^m \\ \phi_{k2}^m \\ \phi_{k3}^m \\ \vdots \\ \phi_{kk}^m \end{bmatrix} = \begin{bmatrix} 1 & \rho_1^{m-1} & \rho_2^{m-1} & \dots & \rho_{k-1}^{m-1} \\ \rho_1^{m-1} & 1 & \rho_1^{m-2} & \dots & \rho_{k-2}^{m-2} \\ \rho_2^{m-1} & \rho_1^{m-2} & 1 & \dots & \rho_{k-3}^{m-3} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{k-1}^{m-1} & \rho_{k-2}^{m-2} & \rho_{k-3}^{m-3} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_{k1}^m \\ \rho_{k2}^m \\ \rho_{k3}^m \\ \vdots \\ \rho_{kk}^m \end{bmatrix}. \quad (23)$$

Para um período m qualquer, o parâmetro ϕ_{kk}^m , chamado de coeficiente de autocorrelação parcial, pode ser interpretado como o último coeficiente autorregressivo de um modelo AR(k). Assim, para calcular a FACP para as defasagens $k = 1, 2, 3, \dots$, o que se faz é ajustar os modelos: AR(1) obtendo-se ϕ_{11}^m , AR(2) obtendo-se ϕ_{22}^m e assim sucessivamente.

A análise do conjunto de valores ϕ_{kk}^m permite definir a ordem mais apropriada dos modelos autorregressivos para cada mês ou semana. Deve-se identificar a maior ordem i tal que todos os valores de ϕ_{kk}^m para $k > i$ não sejam significativos.

No planejamento de médio prazo, para a identificação dos modelos autorregressivos periódicos, adota-se como *lag* máximo o valor igual a 6 e o intervalo de confiança de $\pm 1.96/\sqrt{n}$, em que n é o tamanho da amostra. Neste caso, a identificação da ordem é feita a partir de uma recursão da maior ordem possível até a menor, de modo a identificar o primeiro ϕ_{kk}^m significativo, ou seja, acima de $1.96/\sqrt{n}$ (MACEIRA; DAMAZIO, 2004). Os demais *lags*, maiores que o associado a ϕ_{kk}^m , são considerados insignificantes e são desprezados para a construção do modelo autorregressivo.

Após a etapa de identificação da ordem p_m , procede-se para a estimativa dos parâmetros do modelo usando as equações de Yule-Walker. Mais detalhes relacionados ao método de estimação podem ser encontrados em Sarnaglia, Reisen e Lévy-Leduc (2010).

Nessa linha de pesquisa, uma ferramenta linear híbrida aprimorada a ser mencionada é o modelo PARX. Na literatura, trabalhos como de Lima e Lall (2007) avaliam a inclusão de variáveis exógenas no modelo de previsão de vazão do tipo autorregressivo periódico. Na modelagem, foram incorporados indicadores climáticos como a Temperatura da Superfície do Mar (TSM) em regiões específicas dos oceanos Atlântico e Pacífico e o campo de vento zonal em uma região específica do oceano Atlântico para determinação da variabilidade das precipitações que impactam as vazões naturais afluentes a usina hidrelétrica de Sobradinho. Os autores compararam os resultados do modelo PARX com os modelos PAR e AR que, como já foi mencionado, não fazem uso de nenhuma informação exógena. Em todos os critérios avaliados, o modelo PARX obteve desempenho superior, especialmente na estação chuvosa, quando a memória do processo hidrológico é mais curta e há uma alta correlação da vazão com a precipitação do mês corrente.

3.4 Modelos de previsão de vazão que incorporam os dados de precipitação

Para atender à demanda por previsões de vazões que sejam cada vez mais precisas, o ONS tem buscado continuamente melhorar a qualidade dos modelos de previsão. O processo de evolução abrangeu inclusive a consolidação de dados hidrológicos, de modo a se obter séries de melhor qualidade para o processo de previsão.

Desde o ano de 2000, o operador vem acompanhando o desempenho dos modelos lineares com base na comparação entre as vazões naturais estimadas pelo modelo Previvaz e as vazões naturais observadas nos locais de aproveitamentos do SIN. A análise de desempenho disponível mostra que as vazões previstas por modelos lineares univariados apresentam desvios significativos. No período de 1997 a 2004, considerando a previsão

de uma semana, os erros médios computados para as previsões foram de 12% e 68%, conforme mostra a Tabela 1 (MOREIRA; GUILHON; ROCHA, 2007).

Tabela 1: Desvios médios absolutos do modelo de previsão de vazões Previvaz.

Usina Hidrelétrica	Desvio Previvaz(%)
Tucuruí	12,3
Itaipu	15,5
S. Simão	18,4
Marimbondo	20,9
S. Grande	21,6
Furnas	27,9
Paraibuna	28,6
Três Marias	35,3
Serra da Mesa	35,4
Foz do Areia	38,0
Passo Real	42,0
Itá	67,6

Fonte: (MOREIRA; GUILHON; ROCHA, 2007)

No estudo elaborado por Moreira, Guilhon e Rocha (2007), constatou-se que os maiores erros do modelo Previvaz eram registrados para as usinas localizadas nas bacias da região sul do país, para as quais os modelos lineares não eram capazes de capturar a grande variabilidade das vazões. A dificuldade em modelar o comportamento das vazões nestes aproveitamentos ocorre pela ausência de um padrão hidrológico. As altas declividades dos rios, combinadas com o tipo de solo, fazem com que haja grandes mudanças de uma semana para outra, o que torna as previsões de modelos lineares autorregressivos obsoletas rapidamente (UFF, 2005).

A constatação da necessidade de se aprimorar a previsão de vazão natural semanal, levou o ONS a estudar metodologias alternativas. Visando o aperfeiçoamento do processo de previsão, entre os anos de 2004 e 2006, o ONS estruturou um estudo para o desenvolvimento de novos modelos para a previsão em um horizonte de até 12 dias, considerando, além das vazões verificadas em postos fluviométricos, a disponibilidade de informações de precipitação observada e prevista por modelos numéricos. Os estudos elaborados na época reuniram 41 propostas recebidas de diversos centros de pesquisa. Das alternativas concorrentes foram selecionadas 6 metodologias que incorporam diferentes linhas de pesquisas, dentre as quais estão os modelos estocásticos, os modelos físicos e as técnicas de inteligência computacional (MOREIRA; GUILHON; ROCHA, 2007).

Para a realização dos estudos, foram definidas três bacias hidrográficas, das quais fizeram parte as bacias do rio Iguaçu, Paraná e Paranaíba. Na seleção destes trechos de bacias levou-se em consideração a diversidade de regimes hidrológicos, bem como a relevância

da acurácia da previsão para a programação eletroenergética. Deste trabalho culminou a operacionalização de diferentes modelos de previsão para as bacias selecionadas, cujos modelos foram elegidos com base na melhor resposta para o período de janeiro de 2002 a dezembro de 2003.

Para a bacia do rio Iguaçu foram testados o modelo de previsão SMAP, técnicas estatísticas baseadas em regressão dinâmica, sistemas inteligentes e os modelos estocásticos autorregressivos. Dentre as técnicas de inteligência computacional foram testadas as redes neurais *Multi Layer Perceptron* (MLP) e os sistemas de inferência fuzzy. No confronto entre as técnicas citadas, o modelo fuzzy diário recorrente foi o que apresentou resultados mais promissores no horizonte de uma semana, tendo sido por isso escolhido para integrar posteriormente o processo de previsão de vazões naturais realizados pelo ONS, em substituição aos modelos lineares para as usinas hidrelétricas Foz do Areia, Jordão e Salto Osório (UFF, 2005).

O benefício da mudança de modelo para a previsão de vazão das usinas do rio Iguaçu foi substancial. Ao comparar a resposta do sistema de inferência fuzzy com as saídas do Previvaz, constatou-se uma redução de erro percentual médio absoluto de 17% em relação à aplicação original (UFF, 2005). A eficiência obtida sucedeu-se sobretudo devido à inclusão de dados de precipitação na previsão, o que permitiu capturar com maior acurácia a grande variabilidade das vazões das bacias hidrográficas da região sul.

Em complemento ao projeto, foram testadas também diferentes abordagens para a previsão de vazão da bacia do rio Paraná. Concorreram os modelos estatísticos de regressão dinâmica, o modelo conceitual físico distribuído (MGB-IPH) e o modelo SMAP-MEL (FADIGA et al., 2008). Das técnicas apresentadas, a última foi adotada para compor as previsões hidrológicas para bacia do trecho incremental à Itaipu. Isto porque a sua utilização indicava uma redução de erro médio de previsão de 7,5%, em comparação com o modelo Previvaz.

A previsão de vazão para UHE Itaipu com o modelo SMAP-MEL foi utilizada até novembro de 2017, quando as previsões passaram a considerar apenas as saídas do modelo SMAP, desenvolvido inicialmente por Lopes, Braga e Conejo (1982). O SMAP é um modelo determinístico de simulação hidrológica que fornece uma abordagem conceitual do processo de transformação da chuva em vazão. Para a representação do modelo hidrológico, faz-se uso de reservatórios lineares hipotéticos que representam o solo e o subsolo. Nesta abordagem, a vazão natural do trecho de rio é obtida pela contribuição das parcelas de escoamento superficial e escoamento subterrâneo que provêm de cada reservatório. As principais vantagens do modelo SMAP estão associadas à sua simplicidade e à pequena quantidade de parâmetros a serem utilizados. Dos seis parâmetros do modelo, dois são considerados fixos e estão associados às características da bacia, e os quatro restantes são parâmetros calibráveis com base nas séries históricas de precipitação, evaporação potencial e vazão. Além da bacia do rio Paraná, atualmente, adota-se este modelo na bacia do

rio Grande, Paranaíba e Paranapanema.

Na mesma linha de abordagem, o operador incorporou também um modelo físico para as usinas na bacia do rio Paranaíba. As previsões de vazões para o trecho incremental entre a UHE Itumbiara e a UHE São Simão passaram a ser feitas pelo MGB-IPH. Este modelo, assim como o SMAP, é construído com um forte embasamento físico, ou seja, considera o balanço hídrico no solo para a modelagem hidrológica. A vantagem do MGB-IPH com relação ao SMAP está no fato de ser distribuído, o que permite incorporar a variabilidade espacial da precipitação. No relatório do ONS - Hicon Engenharia Ltda (2005) está disponível a avaliação dos resultados obtidos com cada modelo para a bacia do rio Paranaíba. Em síntese, o MGB-IPH foi superior aos modelos lineares com uma redução dos desvios de previsão de 5,33%.

Na mesma época, teve início a operacionalização de um modelo de previsão de classes de vazão para a bacia do rio Uruguai, o MPCV foi incorporado com o objetivo de modificar o critério de escolha do melhor modelo linear para a previsão de vazões selecionado pelo Previvaz, levando-se em conta informações hidrometeorológicas. Com a finalidade de incluir as informações de previsão de precipitação na estimação da vazão, faz-se uso de técnicas de mineração de dados para definição de limites de classe esperados para a previsão semanal. A obtenção de uma classificação para as previsões de vazões naturais desenvolvida no MPCV emprega as redes neurais com treinamento bayesiano (ONS, 2007). De forma geral, sabendo-se que a vazão da semana em curso está dentro de determinada faixa de valores e a chuva prevista para a semana seguinte se encontra em outra faixa, é possível prever, a partir de uma análise das probabilidades condicionais do histórico de dados, qual será a classe mais provável de vazões da semana seguinte. Se a faixa prevista pelo MPCV contempla a previsão do modelo Previvaz, a saída não é modificada. Contudo, caso a previsão do Previvaz não se encontre dentro da faixa de vazão prevista pelo MPCV, adota-se a saída de um dos 94 modelos lineares que estiver dentro da faixa, respeitando para isso, o modelo com o melhor desempenho na faixa determinada. Se nenhuma previsão do Previvaz estiver dentro da faixa de vazão prevista pelo MPCV, adota-se a vazão prevista que estiver mais próxima. No ensaio realizado para o período de 2001 a 2002, constatou-se ganhos significativos com a aplicação do modelo MPCV. Os erros médios eram de 72%, considerando à utilização isolada do modelo Previvaz, e foram reduzidos para 47% com a aplicação do MPCV para a UHE Barra Grande (ONS, 2007).

Dada a resposta superior dos modelos que incorporam o dado previsto de precipitação em relação ao Previvaz, novos planos de melhoria estudaram novas alternativas para outros aproveitamentos hidrelétricos. No ano de 2011, por exemplo, o ONS desenvolveu um modelo de RNA para a previsão da vazão semanal do aproveitamento hidrelétrico de Três Marias. O modelo Neuro3M foi construído a partir de uma arquitetura neural do tipo NSRBN. Como dados de entrada, foram consideradas as informações dos postos pluviométricos, fluviométricos e a vazão natural observada no reservatório, com diferentes

defasagens. Neste estudo, buscando fazer uma comparação justa, foram consideradas não só as chuvas verificadas para o reprocessamento das previsões mas também a chuva prevista pelo modelo numérico regional ETA (CHOU; NUNES; CAVALCANTI, 2000) no período compreendido entre os anos de 2000 a 2003 e 2006 a 2007. O uso da metodologia proposta proporcionou uma redução do erro de previsão em 11% utilizando a previsão “perfeita” de precipitação e de 7% utilizando a previsão de precipitação do modelo numérico ETA (ONS, 2009).

Seguindo a mesma linha de pesquisa, para o trecho incremental entre o reservatório de Três Marias e o posto hidrométrico de São Francisco também foram incorporadas ao processo de previsão de vazão as redes neurais artificiais. O modelo NeuroSF, assim como o Neuro3M, foi construído a partir de redes neurais construtivas NSRBN. Para o trecho incremental entre Três Marias e São Romão a aplicação da rede neural permitiu uma redução dos erros de previsão de 12% utilizando-se a previsão “perfeita” de precipitação e uma queda de 9% considerando a previsão de precipitação realizada na época pelo modelo regional ETA .

Em resumo, na Tabela 2 apresentam-se os modelos de previsão de vazão atualmente utilizados pelo ONS para a previsão no horizonte de uma semana (ONS, 2018a)² .

Tabela 2: Modelos de Previsão.

Usina/ Trecho/ Bacia	Modelo
Incremental T.Marias - Sobradinho	NEUROSF/CPINS (*)
Uruguai	MPCV (**)
Iguaçu	FUZZY (**)
Incremental Itumbiara - S. Simão	MGB-IPH (**)
Grande	SMAP (**)
Paranaíba (até Itumbiara)	SMAP (**)
Incremental à UHE Itaipu	SMAP (**)
Três Marias	NEURO3M (**)
Paranapanema	SMAP (**)
Demais bacias	PREVIVAZ (***)

O modelo CPINS está na cadeia de modelos utilizados pelo ONS desde o ano de 2006. Ele faz o Cálculo e Previsão de vazões Incrementais e Naturais diárias para a UHE Sobradinho, que compreende o trecho do rio São Francisco a jusante dos aproveitamentos de Três Marias e Queimado. O CPINS é um modelo de propagação baseado na rotina do modelo *Streamflow Synthesis and Reservoir Regulation* (SSARR), que calcula o tempo de

²(*) Previsão semanal para o horizonte de duas semanas. (**) Previsão semanal para o horizonte de uma semana. (***) Previsão semanal para o horizonte de uma semana e para todas as bacias a partir da 2ª semana.

viagem da água que sai das usinas de Três Marias e Queimado e chega em Sobradinho. A partir da informação verificada do aproveitamento ou posto fluviométrico a montante, esse modelo faz a propagação da vazão e determina, pela diferença da vazão verificada no posto fluviométrico de destino, qual a vazão incremental verificada no trecho, permitindo que o operador faça a previsão de vazões incrementais para os demais dias (ONS, 2006).

Vale destacar que tradicionalmente, conforme apresentado na Tabela 2, a classe de modelos lineares continua sendo aplicada para a modelagem das séries temporais de recursos hídricos. As técnicas desenvolvidas e apresentadas anteriormente foram acopladas pelo ONS apenas para a previsão da primeira semana, conhecida como semana chuva-vazão. A previsão das demais semanas faltantes para o mês corrente ainda é obtida pelo modelo Previvaz. Para isso, incorpora-se no histórico dos modelos lineares os resultados dos modelos chuva-vazão na primeira semana de previsão.

Ainda que aprimoramentos tenham sido alcançados para a previsão de vazão de algumas usinas hidrelétricas, existem ainda aproveitamentos que consideram como previsão para a primeira semana o modelo Previvaz. Como exemplo, citam-se as usinas situadas na bacia hidrográfica do rio Tocantins, na bacia do rio Paranaíba (jusante de São Simão), bacia do rio Jacuí e bacia do rio Tietê. Esta última é uma importante bacia da região sudeste, que contempla as usinas hidrelétricas de Barra Bonita, Promissão, Bariri e Três Irmãos.

No último levantamento feito pelo ONS, disponível no Relatório Anual de Avaliação das Previsões de Vazões (ONS, 2018a), constata-se que as previsões para as usinas da bacia do rio Tietê apresentaram um dos piores índices de desempenho em 2017. Os erros médios absolutos considerando apenas a primeira semana de previsão chegam a 39 % no caso da usina de Barra Bonita e 28 % no melhor caso, representado pela usina de Três Irmãos, cujas previsões estão dispostas na Figura 9.

Em linhas gerais, comparando a saída para a primeira semana do modelo Previvaz com o dado observado, nota-se a ocorrência de um efeito sombra nas previsões realizadas ao longo do ano de 2017. As projeções de vazões para usina de Três Irmãos apresentam um atraso em relação à série observada, o que é de se esperar uma vez que a saída dos modelos autorregressivos é influenciada significativamente pelas últimas vazões observadas. Ao analisar a Figura 9, fica evidente que o modelo Previvaz não consegue capturar os picos de vazão com a antecedência que poderia ser fornecida através da incorporação da previsão de precipitação aos modelos hidrológicos. Dado o baixo desempenho, acredita-se que possa haver um ganho bastante elevado nos resultados com a incorporação de tecnologias como os sistemas inteligentes capazes de incluir os dados previstos de precipitação.

O grande apelo para a aplicação de inteligência computacional na modelagem hidrológica está na capacidade destes sistemas de extrair regras e relações matemáticas entre as variáveis do problema de forma automática, não requerendo para isso, *a priori*, informações sobre o sistema. As relações matemáticas entre as variáveis do problema são

Figura 9: Comparação entre vazão observada e prevista com uma semana de antecedência para a usina hidrelétrica de Três Irmãos

Fonte: (ONS, 2018a)

capturadas a partir de um conjunto de dados, o que torna o trabalho de modelagem menos custoso.

Considerando o leque de aplicabilidades em que os sistemas fuzzy são passíveis de serem utilizados, este trabalho foca no problema envolvendo previsões de séries temporais. Espera-se neste trabalho que o sistema de inferência fuzzy seja capaz de melhor representar as vazões, uma vez que este tipo de ferramenta possui uma estrutura matemática flexível capaz de identificar as relações complexas não-lineares entre um conjunto de dados de entrada e saída, o que instiga a sua utilização para previsão de vazões semanais para usinas hidrelétricas.

3.5 Considerações Finais sobre o Capítulo

Neste capítulo foram apresentados alguns dos modelos utilizados pelo setor elétrico brasileiro para a previsão de séries temporais de vazão. Na primeira parte, foram abordados os modelos de Box e Jenkins, que são modelos estocásticos que descrevem o processo hidrológico por uma classe de modelos que calculam o valor do instante presente da série associando-a linearmente, por meio de parâmetros, as informações contidas na própria série temporal. Estes modelos partem do pressuposto de que o sistema gerador da série

pode ser razoavelmente aproximado por um modelo linear. Suposição um tanto restritiva, principalmente se levarmos em conta a não-linearidade existente no processo hidrológico.

Devido à complexidade das dinâmicas envolvidas e à necessidade de uma boa fidelidade do modelo em uma ampla faixa de operação do sistema, a consideração de uma dependência linear parece ser simplória demais. No levantamento feito por Moreira, Guilhon e Rocha (2007) para as previsões compreendidas entre 1997 e 2004, o erro médio absoluto de previsão computado com a aplicação de modelos lineares univariados foi de até 67%, considerando a previsão de uma semana à frente. No último relatório anual de avaliação das previsões de vazões para o ano de 2017, disponibilizado pelo ONS, constata-se erros de previsão da ordem de 39% para as usinas hidrelétricas que fazem uso do modelo Previvaz.

Neste contexto, considerando os desvios elevados com a utilização de modelos lineares, busca-se aprimorar a qualidade das previsões de vazões hidrológicas. Com esse objetivo foi realizado um levantamento das alternativas metodológicas disponíveis para fins de previsão. Dada a importância do tema, verifica-se um grande aporte de pesquisas dedicadas à elaboração de técnicas para a estimativa das vazões naturais no horizonte de uma semana que incorporam métodos não-lineares e consideram a previsão de precipitação. Dentre estas linhas de pesquisa, destacam-se os sistemas inteligentes, aplicados com sucesso em diversas usinas hidrelétricas do SIN (VALENÇA; LUDERMIR, 2001) (NAYAK et al., 2004) (Chin-Hui Liu et al., 2006) (UFF, 2005) (ONS, 2007) (ONS, 2009) (SACCHI, 2009)

Considerando os resultados satisfatórios com a aplicação de sistemas inteligentes e a ordem de grandeza dos desvios existentes nas previsões de vazões naturais a aproveitamentos como da bacia do rio Tietê, acredita-se que há espaço para melhorias no processo de previsão de vazão, o que justifica o desenvolvimento de um modelo de estimação a partir de sistemas de inferência fuzzy voltado para o planejamento da operação do sistema elétrico. O sistema de inferência fuzzy que será empregado é detalhado no Capítulo 4.

Aspectos de Sistemas de Inferência Fuzzy

4.1 Introdução

No capítulo 3 foram apresentados os principais modelos atualmente aplicados nas previsões de vazões naturais afluentes às usinas hidráulicas. Tais métodos consistem na construção de modelos estatísticos que buscam reproduzir o registro histórico a partir de modelos lineares univariados. Constata-se a existência de limitações para este tipo de modelagem, que sob determinadas condições, pode comprometer a correta estimativa desejada destes modelos. A observância de tais limitações fomenta a busca por abordagens atuais que sejam pertinentes para modelar o sistema, e assim, compreender a relação existente entre a precipitação observada em uma sub-bacia e sua vazão.

Com objetivo de melhorar a estimativa de vazão natural afluente a uma usina, adota-se neste trabalho uma abordagem considerando um sistema de inferência fuzzy como ferramenta alternativa para o mapeamento do processo hidrológico. A utilização de tal técnica para aproximação de função se justifica, uma vez que não existe uma equação universal que defina a vazão a partir de dados de precipitação, umidade do solo ou qualquer outra variável. À medida que não existe uma relação matemática bem estabelecida que descreva o comportamento entre as variáveis de entrada e de saída do processo, torna-se atrativa a utilização de modelos empíricos, mais especificamente o uso de sistemas inteligentes, notadamente aqueles baseados em redes neurais e em sistemas fuzzy, aplicados normalmente em processos cuja modelagem por técnicas convencionais são de difícil obtenção.

Nestes casos, podem-se considerar ambas as abordagens, uma vez que estas não requerem, *a priori*, o conhecimento sobre o processo a ser modelado, como relações matemáticas entre as variáveis do problema. Basta, entretanto, que haja um conjunto de informações quantitativas, ou seja, medições que descrevam o comportamento do processo.

A diferença entre os modelos de redes neurais e sistemas fuzzy está principalmente na forma de tratar a informação. As redes neurais artificiais só trabalham com informações

quantitativas, enquanto nos sistemas fuzzy o raciocínio é executado de forma aproximada, incorporando conceitos linguísticos. Por este motivo, geralmente, a interpretação da solução dos problemas acontece de forma mais natural quando o problema é resolvido por meio de um sistema fuzzy do que quando o mesmo problema é resolvido por uma rede neural artificial (TSOUKALAS; R. E. Uhrig, 1996).

Além da interpretabilidade, os sistemas fuzzy possuem um baixo custo de implementação, elevada habilidade em mapear sistemas não-lineares e são capazes de tratar situações em que as informações quantitativas estão imersas em universo de imprecisão, como é o caso da quantificação da vazão natural às usinas hidráulicas. Esta última característica é um atributo que merece destaque quando se pensa em modelar as séries de vazões. Isso porque o dado de vazão natural apresenta elevado ruído oriundo sobretudo do processo de coleta.

Na prática, para muitos postos fluviométricos, visando simplificar o processo de aferição dos dados de vazão, opta-se pela medição do nível de água na seção do rio ao invés do dado de vazão, sendo a determinação da vazão feita a partir de uma curva-chave que relaciona a cota de uma seção do rio com sua correspondente vazão. Esta determinação indireta da vazão pode gerar erros sistemáticos de medição, que são provenientes inclusive da construção e falta de atualização da curva-chave.

Neste sentido, dadas as incertezas no processo de medição de vazão, devem-se considerar as imprecisões na quantificação deste processo hidrológico, o que é possível mediante a utilização dos métodos que aplicam a lógica fuzzy.

4.2 Conceitos sobre Conjuntos e Lógica Fuzzy

Os métodos baseados em lógica fuzzy encontram aplicação em diferentes temáticas envolvendo as engenharias como um todo, é uma vigorosa área de pesquisa multidisciplinar, uma vez que constitui a base para o desenvolvimento de métodos e algoritmos de modelagem. Os sistemas fuzzy podem ser aplicados em áreas como: controle de processos, classificação de padrão, ou ainda, na modelagem não linear de sistemas, por permitir a redução da complexidade de projeto e a extração de informações qualitativas dos sistemas (ROSS, 2004).

A lógica fuzzy é uma extensão da lógica clássica, capaz de lidar, simultaneamente, com dados numéricos e informações linguísticas. A sua implementação permite tratar casos em que a pertinência entre elementos e conjuntos não é precisa, ou seja, não se sabe dizer ao certo se um elemento pertence efetivamente a um conjunto ou não. Para tratar a ausência de definições precisas, substituem-se as funções características bivalentes da lógica tradicional por funções de pertinência, assim como se estendem os conjuntos ordinários para conjuntos fuzzy.

Na teoria de conjuntos clássica, um elemento pertence ou não a uma classe, de forma

que as premissas e conclusões expressem a certeza de inclusão ou não. As fronteiras dos conjuntos são bem definidas e a pertinência de qualquer elemento x a um conjunto A é dado pela função característica $\mu_A(x)$, como descrito a seguir:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A; \\ 0 & \text{se } x \notin A \end{cases} \quad (24)$$

em que a função característica $\mu_A(x) : U \rightarrow \{0, 1\}$ discrimina os elementos do universo de discurso, de forma que os valores de x pertencem ou não ao conjunto A .

Em contraste com a lógica clássica, a caracterização dos conjuntos fuzzy é mais ampla. Na lógica fuzzy se estende a pertinência para o campo da possibilidade¹, propondo através dos conjuntos fuzzy que alguns elementos são mais membros de um conjunto do que outros. A flexibilização do conceito de pertinência é obtida pela substituição da função característica pela função de pertinência, que assume, também, valores intermediários, como apresentado em (25).

$$\mu_A(x) : x \longrightarrow [0, 1], \text{ em que } x \in X. \quad (25)$$

O valor de $\mu_A(x)$ representa o grau de pertinência do elemento x pertencente ao universo de discurso X em relação ao conjunto fuzzy A , quanto mais próximo de 1 for $\mu_A(x)$, mais o elemento se identifica com o conceito representado pelo conjunto fuzzy A .

Dada a caracterização mais ampla, a partir da teoria de conjuntos fuzzy, se fornece uma maneira sistemática para uma descrição aproximada de fenômenos complexos ou mal definidos. Como não há uma disjunção abrupta entre elementos pertencentes e não pertencentes, se torna possível trabalhar matematicamente com variáveis sem uma medida formal, caracterizadas por termos imprecisos como “aproximadamente”, “muito” e “pouco”, descritos em linguagem natural e difíceis de quantificar.

A utilização do tipo de descrição linguística empregada para representação do processo, e não de variáveis quantificadas, permite o tratamento de sistemas que são difíceis de serem analisados através de termos matemáticos convencionais, podendo a variável linguística representar de modo impreciso e, portanto linguístico, um conceito ou um fenômeno complexo.

Para transformar conceitos linguísticos incertos em uma linguagem numérica, associa-se a cada termo linguístico um conjunto fuzzy, que é caracterizado por uma função de pertinência, que descreve a distribuição do conjunto ao longo de um dado universo.

Os conjuntos podem ser dispostos de várias formas no eixo de coordenadas e são nomeados por adjetivos que caracterizam de forma subjetiva o estado de uma variável denominada variável linguística.

Os valores da variável linguística são sentenças em uma linguagem específica, construídas a partir de termos primários. A vazão de um rio, por exemplo, pode ser representada

¹O campo possibilístico indica o quão relacionado está um elemento a um conjunto, de modo que os elementos podem pertencer mais, ou menos, a um determinado conjunto.

de forma aproximada por uma variável linguística como aquela ilustrada na Figura 10, cujos valores (atributos) são expressos por termos linguísticos do tipo vazão “Baixa”, “Média” e “Alta”.

Figura 10: Representação da variável linguística vazão.

No eixo x estão os valores reais da variável que se deseja representar, enquanto no eixo vertical, definido no intervalo unitário, estão os valores de pertinência de cada elemento x às classes. A ideia deste tipo de representação é que, para um dado valor no eixo x , seja normal a existência de mais de uma classe associada a ele. Na Figura 10, por exemplo, o valor de vazão igual a x_6 se encaixa em mais de um conjunto fuzzy, sendo nesta situação particular a sua pertinência em relação aos termos vazão “Baixa” e “Média” iguais.

As funções de pertinência podem assumir diferentes formas, sendo a sua definição feita a partir do conhecimento prévio sobre a grandeza inferida. Como a definição da função de pertinência não é trivial, comumente utilizam-se funções padrões como funções triangulares, trapezoidais e gaussianas.

Na Figura 11, se faz uso de uma função de pertinência do tipo triangular para delimitação do conjunto fuzzy associado à vazão “Média”. O pico em x_8 sugere a ideia de completa inclusão deste elemento no conceito em questão, enquanto valores abaixo de x_4 e acima de x_{12} estão completamente excluídos deste conjunto.

Os conjuntos fuzzy podem ser definidos em universos contínuos ou discretos. Se o domínio X for discreto e finito, consideram-se as funções de pertinência contínuas apenas para produzir os conjuntos fuzzy, que comumente são representados por um conjunto de pares ordenados utilizando a notação em (26):

$$\begin{aligned} A &= \{\mu_A(x)/x\}, \text{ em que } x \in X \\ &= \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \mu_A(x_3)/x_3 + \dots + \mu_A(x_n)/x_n \end{aligned} \quad (26)$$

sendo que o vetor contém os graus de pertinência dos elementos de X no conjunto fuzzy A . O sinal de adição indica a composição dos elementos, e n é o número de elementos que reflete na qualidade de representação da função contínua.

Figura 11: Função de pertinência discretizada

Os conjuntos de termos necessários para descrever a variável linguística, o formato das funções de pertinência e a amplitude do universo de discurso para cada conjunto de dados são parâmetros que precisam ser definidos, de forma que a construção de um sistema fuzzy se inicia pela identificação dos conjuntos que representam as variáveis de entrada e saída do sistema.

A modelagem dos conjuntos pode ser realizada por meio de associações empíricas ou através da prática de uma pessoa especializada com conhecimento a respeito de como o sistema se desenvolve no decorrer de sua observação. Como nem sempre se dispõe de um especialista, ou ainda, o processo de modelagem é complexo, muitas vezes é necessário a aplicação de técnicas de sintonia que sejam eficientes na otimização paramétrica de sistemas de inferência fuzzy.

A sintonização dos parâmetros do modelo é uma das etapas mais custosas do projeto. A grande flexibilidade que decorre da existência de muitos parâmetros exige um grande esforço do projetista, de modo a obter o melhor desempenho do modelo.

Mais adiante será apresentado o método de sintonia aplicado neste trabalho, mas previamente torna-se necessário apresentar a teoria básica de modelagem fuzzy. Assim, na próxima seção, que será inteiramente baseada no trabalho de Flauzino (2007), detalha-se os conceitos envolvidos com os sistemas de inferência fuzzy.

4.3 Sistemas de Inferência Fuzzy

Dentro da teoria compreendida pela lógica fuzzy, uma das implementações mais usuais e que mais se destaca é o sistema de inferência fuzzy. Tal modelo matemático foi concebido por Zadeh e outros pesquisadores, entre os quais E.H Mamdani, que deu início a sua aplicação de caráter prático na década de 70 (MAMDANI, 1974).

Os sistemas fuzzy tiveram sua aplicação inicialmente concentrada na área de Controle, onde se procurava modelar por meio de regras linguísticas o modelo aproximado

de raciocínio humano, capaz de tomar decisões racionais em um ambiente de incerteza e imprecisão. Posteriormente, com o aparecimento de sistemas híbridos como neuro-fuzzy, capazes de lidar com o conhecimento objetivo, a gama de aplicações cresceu e se estendeu à classificação e previsão de séries temporais.

A aplicação dos sistemas fuzzy nas mais diferentes áreas do conhecimento se justifica pela forma sistemática e intuitiva de modelagem de sistemas. O sucesso na aplicação de modelos fuzzy deve-se, sobretudo, às características de interpretabilidade e facilidade de representação do conhecimento, de modo que projetistas e analistas são capazes de traduzir consistentemente, em termos de regras linguísticas, o relacionamento entre as variáveis de entrada e saída dos sistemas. Todavia, apesar de sua concepção parecer mais simples que os modelos convencionais, para que sua atuação seja tão boa quanto uma ferramenta quantitativa, ou seja, capaz de capturar um conjunto de dados de um meio externo e obter uma saída coerente, necessita, na sua fase de desenvolvimento, da especificação de alguns parâmetros relacionados à sua estrutura, dentre os quais se pode citar o número de funções de pertinência para cada entrada e para cada saída, número de regras, bem como a definição de cada uma de suas regras.

A fim de delinear corretamente o tema e apresentar o método de previsão de vazões que será adotado neste trabalho, torna-se necessário um detalhamento prévio do modelo de sistema de inferência fuzzy, o qual pode ser descrito a partir em um diagrama de blocos como o ilustrado na Figura 12.

Figura 12: Diagrama esquemático de um sistema de inferência fuzzy.

Na Figura 12 estão os processos que precisam ser realizados ordenadamente para obter a resposta do sistema fuzzy. Os processos são a fuzzificação das entradas do sistema, a

inferência fuzzy e a defuzzificação. Cada processo possui funções distintas e claramente definidas, de forma que cada processo pode ser representado por uma camada.

4.3.1 Fuzzificação

O primeiro processo, chamado de fuzzificação, consiste no mapeamento de entradas precisas para conjuntos fuzzy. Este mapeamento é imprescindível, uma vez que os sistemas fuzzy processam uma série de regras constituídas por proposições que envolvem termos linguísticos, enquanto na maioria das aplicações práticas consideram-se para os modelos entradas quantitativas, resultantes de medições.

Para permitir o processamento da informação, faz-se necessário transformar o universo de discurso das entradas que são expressas por valores numéricos para conceitos linguísticos. Nesta etapa que é tratada a imprecisão de dados de entrada, pois mesmo que seja fornecido ao modelo um valor numérico, este será fuzzificado por meio de um conjunto fuzzy, de modo que a informação quantitativa perde a importância e a incerteza passa a ser, de certa forma, ignorada, pois uma faixa relativamente grande de valores pode ser toda tratada da mesma maneira segundo um conceito linguístico (BROWN; HARRIS, 1994).

Para fuzzificar uma entrada, mede-se o grau de pertinência de cada entrada em relação aos conjuntos fuzzy associados a cada variável de entrada, podendo a cada variável linguística de entrada estar associada tantos conjuntos fuzzy quanto necessários.

Formalmente, dado um sistema fuzzy com uma única entrada definida pelo escalar x e, a esta entrada, associadas N funções de pertinência, ou seja, N conjuntos fuzzy que definem a referida entrada, tem-se como saída desta camada um vetor coluna $\mathbf{Y}$ com N elementos representando os graus de pertinência da entrada em relação aos respectivos conjuntos fuzzy (FLAUZINO, 2007):

$$\mathbf{Y}(x) = \boldsymbol{\mu}(x) = \begin{bmatrix} \mu_1(x) \\ \mu_2(x) \\ \vdots \\ \mu_N(x) \end{bmatrix} \quad (27)$$

em que $\mu_k(x)$ representa a função de pertinência definida para a entrada x referente ao k -ésimo conjunto fuzzy associado a esta entrada.

De forma análoga, pode-se expandir este conceito para um sistema fuzzy com m variáveis de entrada. Em (28) é apresentada a generalização da saída do processo de fuzzificação:

$$\mathbf{Y}(x) = \begin{bmatrix} \mathbf{Y}^{(1)} \\ \mathbf{Y}^{(2)} \\ \vdots \\ \mathbf{Y}^{(m)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}^{(1)}(x_1) \\ \boldsymbol{\mu}^{(2)}(x_2) \\ \vdots \\ \boldsymbol{\mu}^{(m)}(x_m) \end{bmatrix} \quad (28)$$

em que x_i é a i -ésima entrada do sistema fuzzy e $\mu^{(k)}$ é o k -ésimo vetor de função de inferência fuzzy associado à entrada x_k .

Neste caso, o vetor $\mathbf{Y}(\mathbf{x})$ dado em (28) passa a ser definido como a concatenação vetorial dos vetores $\mathbf{Y}^{(1)}$, $\mathbf{Y}^{(2)}$, ..., $\mathbf{Y}^{(m)}$, os quais de forma individualizada podem ser descritos como uma subcamada da primeira camada do sistema fuzzy.

Como exemplo, do processo de fuzzificação, apresenta-se em (29) a expressão da camada de entrada de um sistema fuzzy constituído de duas entradas, x_1 e x_2 .

$$\mathbf{Y}(\mathbf{x}) = \left[\underbrace{\mu_1^{(1)}(x_1) \quad \mu_2^{(1)}(x_1) \quad \cdots \quad \mu_N^{(1)}(x_1)}_{\mu^{(1)}(x_1)} \mid \underbrace{\mu_1^{(2)}(x_2) \quad \mu_2^{(2)}(x_2) \quad \cdots \quad \mu_M^{(2)}(x_2)}_{\mu^{(2)}(x_2)} \right]^T \quad (29)$$

Considerando um caso particular, como aquele ilustrado na Figura 13, no qual consideram-se duas variáveis linguísticas de entrada, teremos como saída da primeira interface um vetor com 7 elementos, cujos valores correspondem aos graus de pertinência das entradas com relação aos seus respectivos conjuntos fuzzy:

$$\mathbf{Y}(\mathbf{x}) = \left[\underbrace{\mu_1^{(1)}(x_1) \quad \mu_2^{(1)}(x_1) \quad \mu_3^{(1)}(x_1)}_{\mu^{(1)}(x_1)} \mid \underbrace{\mu_1^{(2)}(x_2) \quad \mu_2^{(2)}(x_2) \quad \mu_3^{(2)}(x_2) \quad \mu_4^{(2)}(x_2)}_{\mu^{(2)}(x_2)} \right]^T \quad (30)$$

Figura 13: Representação das entradas do sistema fuzzy.

As funções de pertinência das entradas podem ser definidas por meio de qualquer função, não necessariamente as triangulares e trapezoidais como as da Figura 13. A única restrição é que elas estejam definidas no domínio fechado $[0,1]$.

4.3.2 Processo de Inferência Fuzzy

De posse das entradas fuzzificadas, ou seja, definido o vetor $\mathbf{Y}(\mathbf{x})$, é possível realizar o segundo processo fuzzy, que é denominado inferência. No processo de inferência, as variáveis de entrada fuzzificadas são combinadas entre si segundo as regras linguísticas, utilizando como suporte as operações definidas na teoria de conjuntos fuzzy.

O resultado da aplicação do procedimento de inferência é a produção de uma região fuzzy de saída como aquela ilustrada na Figura 14, a qual estará relacionada com a

resposta final do processo. Como não há dependência entre as regras de inferência, estas são processadas em paralelo e a saída do modelo é obtida a partir da agregação das regiões fuzzy produzidas por todas as regras.

Figura 14: Região fuzzy associada a saída do sistema.

A região fuzzy produzida por cada uma das regras é proporcional à sua força de disparo, valor este que está relacionado com a adequação dos valores apresentados na entrada do sistema com os conjuntos fuzzy componentes de cada uma das regras. Adicionalmente, é possível também valorizar aquelas regras que mais influenciam na solução do problema. Esta ponderação acontece segundo a expressão (31).

$$R_i(Y(x)) = w_i \cdot r_i(Y(x)) \quad (31)$$

em que $R_i(\cdot)$ é a função que representa o valor ponderado da i -ésima regra fuzzy, w_i é o fator de ponderação da i -ésima regra fuzzy e $r_i(\cdot)$ o valor fuzzy da i -ésima regra fuzzy.

Para cada regra fuzzy r_i , é definido um conjunto de condições descrevendo uma parcela observável das entradas que leve a uma saída. As regras de modelagem fuzzy que são declarações condicionais do tipo condição-ação são da seguinte forma:

$$r_i : \text{Se } x \text{ é } A_i \text{ E }^2 y \text{ é } B_i \text{ então } z \text{ é } C_i \quad (32)$$

em que A_i , B_i e C_i são os conjuntos fuzzy associados às variáveis linguísticas de entrada e saída do sistema e x , y e z são os elementos dos universos de discurso de cada variável.

²As proposições fuzzy podem também ser combinadas por meio de outros operadores, como por exemplo o conectivo lógico “OU”.

Tabela 3: Conjunto de regras fuzzy.

	Precipitação (t)			
Vazão (t-1)	Pequena	Média	Alta	Elevada
Muito Baixa	Muito Baixa	Muito Baixa	Baixa	Média
Baixa	Muito Baixa	Baixa	Média	Média
Média	Baixa	Média	Média	Alta
Alta	Média	Média	Alta	Muito Alta
Muito Alta	Média	Alta	Muito Alta	Muito Alta

Tipicamente, a condição é uma proposição linguística sobre o valor de alguma das variáveis de entrada. Nessa formulação as proposições “ x é A_i ” e “ y é B_i ” são conhecidas como antecedentes das regras e expressam as condições a serem satisfeitas que resultam em um desdobramento representado pelo consequente “ z é C_i ”.

A premissa “ x é A_i ” será verdade desde que exista um grau de similaridade diferente de zero entre a entrada x e o conjunto fuzzy A_i . O grau de veracidade da proposição será definido pelo valor de $\mu_{A_i}(x)$, contido no vetor $\mathbf{Y}(x)$.

Matematicamente, a regra fuzzy descrita em (32) pode ser interpretada como uma relação fuzzy r_i definida da seguinte forma:

$$\mu_{r_i} = \mu_{(A_i \text{ and } B_i \rightarrow C_i)}(x, y, z) = [\mu_{A_i}(x) \text{ and } \mu_{B_i}(y)] \rightarrow \mu_{C_i}(z) \quad (33)$$

em que $\mu_{A_i}(x)$ mede o grau de pertinência da entrada x em relação ao conjunto fuzzy A_i , $\mu_{B_i}(y)$ mede o grau de pertinência de y em relação ao conjunto fuzzy B_i e o operador $(\rightarrow)$ denota uma função de implicação fuzzy.

As relações fuzzy, r_i , são a base de conhecimento do modelo, elas armazenam o relacionamento existente entre os antecedentes e consequentes de cada uma das regras.

De modo mais intuitivo, é conveniente representar o conjunto de regras fuzzy em forma de matriz, como a apresentada na Tabela 3. A matriz de relações fuzzy representa linguisticamente o comportamento dinâmico do processo. Na Tabela 3 consideram-se duas variáveis de entrada e uma variável de saída, tem-se nas coordenadas da matriz os estados das variáveis linguísticas de entrada e nas células o valor linguístico da variável de saída.

A leitura das regras é exemplificada para a célula sombreada:

$$\text{Se Vazão}(t-1) \text{ é Baixa } \mathbf{E} \text{ Precipitação}(t) \text{ é Alta } \mathbf{então} \text{ Vazão}(t) \text{ é Média} \quad (34)$$

em que “Alta” é o conjunto fuzzy definido no domínio da variável de entrada (precipitação acumulada em 24 horas) e “Baixa” e “Média” são conjuntos fuzzy definidos no domínio da variável de saída (vazão natural média diária).

A matriz representa as transições possíveis do estado corrente da variável vazão conforme o volume de precipitação. A partir de um conjunto finito dessas regras é possível determinar por meio do processo de inferência o comportamento da variável de saída do sistema. Para isso, é preciso criar uma base de regras fuzzy que contenha as mais importantes ou senão todas as combinações entre as funções de pertinência das entradas.

O número de regras possíveis é função direta do número de conjuntos fuzzy considerados para cada antecedente (no exemplo em questão tem-se $5 \times 4 = 20$ regras possíveis). E quanto maior o número de conjuntos atribuídos aos antecedentes, maior a dificuldade de se estabelecer uma base de regras consistente. Assim, a eficiência no dimensionamento do número de conjuntos e no número de regras deve ser considerada, uma vez que, ao optar por um sistema fuzzy, espera-se obter uma estratégia de ação interpretável, que pode ser comprometida por uma quantidade excessiva de regras.

Para traduzir matematicamente as sentenças das regras fuzzy define-se, em substituição aos conectivos “E” e “OU”, os operadores de intersecção Γ -norma e união S -norma. As operações de intersecção ($A \cap B$) e união ($A \cup B$) são definidas de maneira formal pelas expressões (35a) e (35b), respectivamente.

$$\mu_{A \cap B}(x, y) = \mu_A(x) \Gamma \mu_B(y) \quad (35a)$$

$$\mu_{A \cup B}(x, y) = \mu_A(x) S \mu_B(y) \quad (35b)$$

em que Γ é uma norma triangular (Γ -norma) e S é uma co-norma triangular (S -norma).

Como exemplo de Γ -norma e S -norma têm-se as expressões (36a) e (36b), respectivamente.

$$\mu_A(x) \Gamma \mu_B(y) = \min \{ \mu_A(x), \mu_B(y) \} \quad (36a)$$

$$\mu_A(x) S \mu_B(y) = \max \{ \mu_A(x), \mu_B(y) \}. \quad (36b)$$

Embora a intersecção e a união sejam descritas em (36a) e (36b) com os operadores $\min$ e $\max$, existem ainda outras Γ -norma e S -norma que podem ser empregadas (PEDRYCZ; GOMIDE, 1998). Por exemplo, adota-se neste trabalho como Γ -norma o produto algébrico para a intersecção fuzzy e como S -norma a soma algébrica para a união fuzzy, conforme descrito em (37a) e (37b).

$$\mu_A(x) \Gamma \mu_B(y) = \mu_A(x) \cdot \mu_B(y) \quad (37a)$$

$$\mu_A(x) S \mu_B(y) = \mu_A(x) + \mu_B(y) - \mu_A(x) \cdot \mu_B(y). \quad (37b)$$

Usando como suporte tais operações, é possível combinar os conjuntos fuzzy de diferentes variáveis linguísticas. Os antecedentes das regras fuzzy, por exemplo, podem ser combinados mediante a aplicação da operação de intersecção, $\mu_{A_i \cap B_i}(x, y)$.

Formalmente, a combinação de dois conjuntos fuzzy é representada por uma relação fuzzy $R(x, y)$. Dados dois universos de discurso X e Y , se define uma relação fuzzy como um conjunto fuzzy no subespaço do produto cartesiano, $X \times Y$, que resulta em uma matriz relacional, como a apresentada em (38):

$$R(x, y) = \begin{matrix} & y_1 & y_2 & \dots & y_M \\ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{matrix} & \begin{pmatrix} \mu_R(x_1, y_1) & \mu_R(x_1, y_2) & \dots & \mu_R(x_1, y_M) \\ \mu_R(x_2, y_1) & \mu_R(x_2, y_2) & \dots & \mu_R(x_2, y_M) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_R(x_N, y_1) & \mu_R(x_N, y_2) & \dots & \mu_R(x_N, y_M) \end{pmatrix} \end{matrix} \quad (38)$$

em que os valores de $\mu_R(x, y)$ fornecem o nível de associação entre dois elementos definidos em universos distintos.

A possibilidade de relacionar conjuntos de diferentes variáveis é importante, pois o que se busca em sistemas fuzzy é, justamente, o resultado da relação entre diversas variáveis de um problema.

Além das operações entre relações definidas em um mesmo universo, é possível combinar duas ou mais relações fuzzy definidas em espaços distintos, ou seja, com produtos cartesianos diferentes, por meio de operadores que permitem a composição das respectivas relações.

A composição de duas relações fuzzy $R(x, y)$ e $S(y, z)$, definidas respectivamente nos produtos cartesianos $X \times Y$ e $Y \times Z$, retorna uma relação fuzzy definida no subconjunto $X \times Z$. Para se obter a composição, faz-se uso da seguinte expressão:

$$R \circ S(x, z) = \sup\{\mu_R(x, y) * \mu_{S(y, z)}\} \quad (39)$$

em que “*” simboliza uma Γ -norma e “sup” uma S -norma, dada pelo operador “max”. Para a composição, é possível adotar qualquer uma, embora não somente, das seguintes expressões:

$$\square \text{ composição max-min: } \mu_C(x, z) = \mu_{R \circ S}(x, z) = \max\{\min\{\mu_R(x, y), \mu_S(y, z)\}\}$$

$$\square \text{ composição max-prod: } \mu_C(x, z) = \mu_{R \bullet S}(x, z) = \max\{\mu_R(x, y) \cdot \mu_S(y, z)\}$$

$$\square \text{ composição max-méd: } \mu_C(x, z) = \mu_{R \oplus S}(x, z) = \max\{\frac{1}{2}(\mu_R(x, y) + \mu_S(y, z))\}$$

A matriz resultante da operação de composição será da seguinte forma:

$$R \circ S(x, z) = \begin{matrix} & z_1 & z_2 & \dots & z_M \\ \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{matrix} & \begin{pmatrix} \mu_{R \circ S}(x_1, z_1) & \mu_{R \circ S}(x_1, z_2) & \dots & \mu_{R \circ S}(x_1, z_M) \\ \mu_{R \circ S}(x_2, z_1) & \mu_{R \circ S}(x_2, z_2) & \dots & \mu_{R \circ S}(x_2, z_M) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{R \circ S}(x_N, z_1) & \mu_{R \circ S}(x_N, z_2) & \dots & \mu_{R \circ S}(x_N, z_M) \end{pmatrix} \end{matrix}. \quad (40)$$

Se considerarmos, por exemplo, para o cálculo de $R \circ S(x, z)$ a composição max-min, os elementos da matriz (a seguir representado por $\mu_{R \circ S}(x_1, z_1)$) serão calculados conforme a equação (41):

$$\begin{aligned} \mu_C(x_1, z_1) &= \mu_{R \circ S}(x_1, z_1) = \max \{ \min \{ \mu_R(x_1, y), \mu_S(y, z_1) \} \} \\ &= \max \{ \min \{ \mu_R(x_1, y_1), \mu_S(y_1, z_1) \}, \dots, \min \{ \mu_R(x_1, y_n), \mu_S(y_n, z_1) \} \}. \end{aligned} \quad (41)$$

A composição de relações, tão importante quanto a relação fuzzy, desempenha papel fundamental na implementação computacional. É a partir desta operação que a resposta do sistema de inferência fuzzy é obtida. Basicamente, a saída é calculada mediante a composição de relações fuzzy, em que a primeira relação é um conjunto *singleton*³ e a segunda relação pode ser interpretada como uma relação de implicação.

O processo de inferência segue uma regra de implicação fuzzy baseada na generalização da regra de *Modus Ponens*. Considerando, um sistema com uma entrada x e uma saída y compostas, respectivamente, pelos conjuntos fuzzy $\{A_1, A_2, A_3, \dots, A_n\}$ e $\{B_1, B_2, B_3, \dots, B_m\}$, para um fato observável A' a consequência da regra fuzzy é o consequente B' , descrita da seguinte forma:

$$\begin{array}{ll} \text{Fato:} & x \text{ é } A' \\ \text{Regra:} & \text{se } (x \text{ é } A) \text{ então } (y \text{ é } B) \\ \hline \text{Consequência:} & y \text{ é } B' \end{array}$$

O conjunto fuzzy A' não é necessariamente o mesmo que A (antecedente da regra), assim como B' não é necessariamente o mesmo que o consequente B , mas sim, um consequente com grau de similaridade não nulo em relação ao consequente da regra.

Nesta abordagem a regra, “se $(x \text{ é } A)$ então $(y \text{ é } B)$ ”, é interpretada como a implicação do conjunto A no conjunto B , de forma que se tem uma relação fuzzy do tipo $R_{A \rightarrow B}(x, y)$, calculada a partir de (42):

$$\mu_R(x, y) = \mu_{A \rightarrow B}(x, y) = f_{\rightarrow}(\mu_A(x), \mu_B(y)) \quad (42)$$

em que $\mu_R(x, y)$ é a função de pertinência que mede o grau de verdade da relação de implicação entre $x \in X$ e $y \in Y$. A função de pertinência da relação $R_{A \rightarrow B}(x, y)$ pode ser obtida pelos seguintes métodos:

³*Singleton* é um caso particular de conjunto fuzzy normalizado, cujo suporte é um único ponto, $\mu(x)=1$. Os conjuntos singleton são especialmente utilizados para mapear as grandezas de entrada do sistema fuzzy que geralmente são representadas por valores pontuais.

- *Mamdani* $\Rightarrow \mu_{R_{A \rightarrow B}}(x, y) = \min\{\mu_A(x), \mu_B(y)\}$
- *Zadeh* $\Rightarrow \mu_{R_{A \rightarrow B}}(x, y) = \max\{1 - \mu_A(x), \min\{\mu_A(x), \mu_B(y)\}\}$
- *Larsen* $\Rightarrow \mu_{R_{A \rightarrow B}}(x, y) = \mu_A(x) \cdot \mu_B(y)$
- *Aritmético* $\Rightarrow \mu_{R_{A \rightarrow B}}(x, y) = \min\{1, 1 - \mu_A(x) + \mu_B(y)\}$
- *Booleano* $\Rightarrow \mu_{R_{A \rightarrow B}}(x, y) = \max\{1 - \mu_A(x), \mu_B(y)\}$.

O conjunto B' , que representa a consequência da aplicação do processo de inferência, é calculado pela regra de inferência composicional considerando a composição da relação A' , que é um conjunto fuzzy definido no universo de discurso X , com a relação fuzzy $R_{A \rightarrow B}(x, y)$, conforme definido em (43):

$$B'(y) = A'(x) \circ R_{A \rightarrow B}(x, y). \quad (43)$$

Ao substituir na operação de composição a primeira relação por um conjunto fuzzy definido apenas no universo de discurso X , a operação de composição retorna uma função definida em Y .

Como exemplo, em arquiteturas de implicação *Mamdani* e composição máx-min, o processo de inferência se baseia nas equações (44) e (45):

$$\mu_{R_{A \rightarrow B}}(x, y) = \min\{\mu_A(x), \mu_B(y)\} \quad (44)$$

$$\begin{aligned} \mu_{B'}(y) &= \mu_{A'}(x) \circ \mu_{R_{A \rightarrow B}}(x, y) = \max\{\min\{\mu_{A'}(x), \mu_{R_{A \rightarrow B}}(x, y)\}\} \\ &= \max\{\min\{\mu_{A'}(x), \min\{\mu_A(x), \mu_B(y)\}\}\} \end{aligned} \quad (45)$$

em que $\mu_{B'}(y)$ representa a função de pertinência do consequente da regra fuzzy.

É muito comum, em aplicações práticas, ter como informação uma entrada não fuzzy x_0 , também chamada *crisp*. Nesta situação, o conjunto A' é determinado utilizando o conjunto *Singleton*, um caso particular cujo suporte é um único ponto $x_0 \in X$ com $\mu(x_0) = 1$, isto é, um função de pertinência dada por

$$\mu_{A'}(x) = \begin{cases} 1, & \text{se } x = x_0; \\ 0, & \text{caso contrário} \end{cases} \quad (46)$$

Nos sistemas de inferência a entrada não-fuzzy é representada pela função de pertinência $\mu_{A'}(x) = \{0100\}$, que torna possível simplificar a expressão de composição e, consequentemente, a saída representada por $\mu_{B'}(y)$.

Para a relação de implicação $R_{A \rightarrow B}(x, y)$ representada matricialmente em (47):

$$\mu_{R_{A \rightarrow B}}(x, y) = \begin{matrix} & \begin{matrix} y_1 & y_2 & \dots & y_M \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{pmatrix} \mu_{R_{A \rightarrow B}}(x_1, y_1) & \mu_{R_{A \rightarrow B}}(x_1, y_2) & \dots & \mu_{R_{A \rightarrow B}}(x_1, y_M) \\ \mu_{R_{A \rightarrow B}}(x_2, y_1) & \mu_{R_{A \rightarrow B}}(x_2, y_2) & \dots & \mu_{R_{A \rightarrow B}}(x_2, y_M) \\ \mu_{R_{A \rightarrow B}}(x_3, y_1) & \mu_{R_{A \rightarrow B}}(x_3, y_2) & \dots & \mu_{R_{A \rightarrow B}}(x_3, y_M) \\ \mu_{R_{A \rightarrow B}}(x_4, y_1) & \mu_{R_{A \rightarrow B}}(x_4, y_2) & \dots & \mu_{R_{A \rightarrow B}}(x_4, y_M) \end{pmatrix} \end{matrix} \quad (47)$$

como $\mu'_A(x) \neq 0$ apenas em um ponto ($x = x_0$), a operação de composição será reduzida à operação $\mu'_A(x_0) * \mu_{R_{A \rightarrow B}}(x_0, y)$, tal que:

$$\begin{aligned} \mu_{B'}(y) &= \mu_{A'}(x) \circ \mu_{R_{A \rightarrow B}}(x, y) = \max \{ \min \{ \mu_{A'}(x), \mu_{R_{A \rightarrow B}}(x, y) \} \} \\ &= \left[\begin{array}{c} \max \{ 0, \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_1) \} \}, 0, 0 \} \\ \max \{ 0, \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_2) \} \}, 0, 0 \} \\ \vdots \\ \max \{ 0, \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_M) \} \}, 0, 0 \} \end{array} \right]^T \\ &= \left[\begin{array}{c} \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_1) \} \\ \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_2) \} \\ \vdots \\ \min \{ \mu_{A'}(x_2), \mu_{R_{A \rightarrow B}}(x_2, y_M) \} \end{array} \right]^T. \end{aligned} \quad (48)$$

Na equação (48) a operação *sup* torna-se desnecessária e, em consequência, para $\mu_{A'}(x_2) = 1$, tem-se

$$\begin{aligned} \mu_{B'}(y) &= \max \{ \min \{ 1, \mu_{R_{A \rightarrow B}}(x_2, y) \} \} \\ &= \mu_{R_{A \rightarrow B}}(x_2, y). \end{aligned} \quad (49)$$

Utilizando este resultado e refazendo a expressão (45), verifica-se que o procedimento de cálculo do consequente das regras, $\mu_{B'}(y)$, é mais simples e pode ser definido como:

$$\mu_{B'}(y) = \mu'_A(x_0) * \mu_{R_{A \rightarrow B}}(x_0, y). \quad (50)$$

O resultado advindo de cada regra fuzzy irá compor a região fuzzy de saída do sistema. Em função da multiplicidade de regras ativas, a inferência fuzzy produz diversas regiões $\mu_{B'}(y)$, que devem ser agregadas em apenas uma.

4.3.3 Processo de Defuzzificação

Na camada de saída, após a inferência, as contribuições de cada uma das regras ativas são agregadas em uma única região fuzzy, de acordo com o operador de agregação:

$$B' = \bigcup_i^n B'_i \quad (51)$$

em que $\bigcup$ representa o operador agregação e $B'_1, \dots, B'_n$ são todos os resultados de implicação, derivados das diversas regras acionadas, todos relacionados a uma mesma variável linguística.

Como as saídas das relações fuzzy são conjuntos fuzzy, as operações de agregação podem fazer uso dos mesmos operadores de união *S*-norma anteriormente apresentados. Usualmente, utiliza-se como operador de agregação o operador *max*, embora exista uma ampla classe de operadores de agregação, de forma que se pode considerar sua escolha outro parâmetro do sistema fuzzy a ser definido.

Por fim, obtida a região fuzzy saída, como aquela ilustrada na Figura 15, é feita a interpretação dessa informação para saídas precisas. Isto se faz necessário já que em aplicações práticas são requeridos valores precisos ao invés de quantificações fuzzy.

Figura 15: Região fuzzy de saída.

Para exportar a saída para um ambiente não-fuzzy aplica-se o processo de defuzzificação. A conversão da região fuzzy em valores pontuais é obtida a partir da aplicação dos métodos de centro de área, método da média dos máximos ou método do primeiro máximo. Dentre estes métodos, o mais comum é o centro de área, que consiste em calcular a posição no eixo horizontal do ponto que corresponde ao centro de área da região fuzzy obtida com a agregação. Dadas as possibilidades de escolha, deve-se considerar que o método de defuzzificação também faz parte da parametrização do sistema de inferência fuzzy.

4.3.4 Arquiteturas Fuzzy

Com relação às regras, os sistemas de inferência fuzzy podem ser classificados em:

- Paramétricos - cujos consequentes são funções polinomiais;

$$r_i : \text{Se } x_1 \text{ é } A_{i1} \text{ e...e } x_n \text{ é } A_{in} \text{ então } z_i = f_i(x_1, \dots, x_n) \quad (52)$$

- Não paramétricos - cujos consequentes são conjuntos fuzzy;

$$r_i : \text{Se } x_1 \text{ é } A_{i1} \text{ e...e } x_n \text{ é } A_{in} \text{ então } z_i \text{ é } B_i. \quad (53)$$

Na classe de sistemas paramétricos está o modelo fuzzy Takagi-Sugeno (TS) (TAKAGI; SUGENO, 1985), para o qual apenas os antecedentes das regras são especificados por conjuntos fuzzy, enquanto os consequentes são definidos por funções polinomiais cujos coeficientes são ajustados automaticamente por meio de regressão linear.

As funções polinomiais $f_i(x, y)$ no modelo TS são assumidas funções lineares, o que faz com que o modelo seja composto por regras capazes de descrever, de forma exata ou aproximada, sistemas não-lineares a partir de modelos locais lineares. Nesta abordagem, as regras representam regiões de operação que são consideradas lineares e, através da sobreposição de todas as regiões localmente válidas, obtém-se a dinâmica não linear do sistema.

Apesar deste tipo de modelagem tornar menos custoso o trabalho do projetista de ajustar um sistema fuzzy, já que os consequentes são definidos com base nos dados amostrados, deve-se considerar que a representação de um processo através desta arquitetura acarreta na perda de interpretabilidade junto à base de regras criadas.

Buscando obter uma estratégia mais intuitiva, considera-se a segunda classe de modelos não paramétricos. Nestes, a representação do relacionamento entre o espaço das entradas com o espaço das saídas é realizada de forma verbal, o que oferece um elevado nível semântico. Ademais, o emprego deste tipo de sistema de inferência fuzzy, devido a sua estrutura, vem a oferecer uma robustez quanto à incorporação gradual de regras a sua base de conhecimento, sem que haja perda ou inconsistência no resultado final.

Como ponto contrário, pode-se citar a necessidade de existência de uma base de conhecimento especialista, sendo que a complexidade de modelagem de sistemas de médio até grande porte pode acarretar em uma insuficiência de precisão nos resultados de simulação e, consequentemente, um pobre desempenho dos modelos não-paramétricos com relação aos paramétricos.

Com o objetivo de atender à demanda por um sistema mais intuitivo, torna-se interessante então o desenvolvimento de algoritmos de ajustes de sistemas de inferência fuzzy baseado em arquiteturas *Mamdani*, que vêm a poupar esforço do projetista nas etapas de sintonia das funções de pertinência e inclusive na criação das regras de inferência fuzzy.

Os métodos de ajuste de sistemas de inferência fuzzy podem ser interpretados como ajustes estruturais e paramétricos. A otimização estrutural visa tornar eficiente a geração de regras e envolve normalmente sistemas com uma base muito grande. Nestes casos, é desejável adotar técnicas de indução de regras e ainda otimizar a base de regras induzidas, de modo que o conhecimento armazenado no sistema possa ser melhor compreendido.

A otimização paramétrica envolve o ajuste dos parâmetros livres, que são os parâmetros das funções de pertinência e a ponderação de cada regra fuzzy. A ideia deste tipo de abordagem é definir os formatos das funções de pertinência de entrada e saída dos sistemas de inferência fuzzy de forma eficiente, de modo que o sistema difuso melhor modele a resposta do sistema.

4.3.5 Ajuste Estrutural de Sistemas de Inferência Fuzzy

O algoritmo de criação de base de regras empregado neste trabalho é o mesmo de Flauzino (2007), e portanto baseia-se no algoritmo de otimização Hill Climbing (ROGER

JANG et al, 1997). A otimização da base de regras consiste na busca por uma base de regras que forneça o menor custo. Assim, empregam-se funções custos que expressam o desvio entre os resultados desejados e os resultados fornecidos pelo sistema fuzzy. A função pode ser explicitada em termos do erro quadrático médio entre a saída do sistema fuzzy e os resultados desejados.

O procedimento tem por objetivo realizar uma busca em um determinado espaço válido, dado pelo espaço de busca S . Para um conjunto factível F , $F \subseteq S$, deve-se encontrar r^* tal que a expressão (54) seja verdadeira, ou seja:

$$\xi(r^*) \leq \xi(r); \text{ para qualquer } r \in F \quad (54)$$

em que $\xi(\cdot)$ representa a função custo; r é uma determinada base de regras pertencente ao conjunto de todas as bases de regras válidas; e r^* configura a base de regras que fornece o menor valor para a função custo em relação à região factível F .

Como encontrar a melhor dentre todas as bases não é uma tarefa fácil, já que existem inúmeras possibilidades e a superfície formada pela função custo pode conter vários mínimos locais, torna-se conveniente a adoção de estratégias de busca que operem localmente (FLAUZINO, 2007). Assim, utiliza-se um algoritmo de otimização estrutural como o exposto na Figura 16 que opera em duas fases distintas. Na primeira fase são sorteadas diversas bases de regras, de modo a encontrar a base que forneça o menor custo. Uma vez definida esta base de regras, realiza-se uma busca local restrita a uma determinada região em torno da vizinhança da base escolhida. Por fim, a base de regras do sistema de inferência fuzzy será aquela que apresentar o menor custo depois de realizadas todas as iterações da primeira e da segunda etapa.

4.3.6 Ajustes Paramétricos de Sistemas de Inferência Fuzzy

Assim como no caso do ajuste estrutural, o ajuste paramétrico também tem por objetivo minimizar uma função custo. Para isso, considera-se como conjunto de ajuste os parâmetros livres de cada camada do sistema de inferência fuzzy dispostos na Figura 17. Os parâmetros livres são os parâmetros das funções de pertinência dos conjuntos fuzzy de entrada, o vetor de ponderação das regras e os parâmetros das funções de pertinência dos conjuntos fuzzy do espaço de saída. Assim, considerando que o conjunto de ajuste $\{x, y\}$ é fixo durante todo o processo, pode-se escrever esta função custo ξ responsável pela sintonia da seguinte forma:

$$\xi_p = \xi(\mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}) \quad (55)$$

em que $\mathbf{h}^{(1)}$, $\mathbf{h}^{(2)}$ e $\mathbf{h}^{(3)}$ serão dados por (56), (57) e (58), respectivamente.

Para exemplificar, assume-se um sistema fuzzy multicamadas composto por uma variável com duas funções de pertinência gaussianas na camada de entrada, três regras

```

Inicializar o sistema fuzzy multicamadas

R1 ← Selecionar uma base de regras aleatória

Contador ← 0

REPITA
  R2 ← Selecionar uma base de regras aleatória
  SE custo(R2) < custo(R1)
    ENTÃO
      R1 ← R2
      Contador ← Contador + 1
  FIM SE
ATÉ Contador muito grande

Contador ← 0

REPITA
  R2 ← Inserir uma pequena perturbação em R1
  SE custo(R2) < custo(R1)
    ENTÃO
      R1 ← R2
      Contador ← Contador + 1
  FIM SE
ATÉ Contador muito grande

```

Figura 16: Algoritmo de extração de regras fuzzy.
Fonte: Flauzino (2007)

Figura 17: Sistema fuzzy multicamada.

na camada de inferência e uma variável com duas funções de pertinência gaussianas na camada de saída.

$$\mathbf{h}^{(1)} = \begin{bmatrix} \mu_1^e & \sigma_1^e & \vdots & \mu_2^e & \sigma_2^e \end{bmatrix}^T \quad (56)$$

$$\mathbf{h}^{(2)} = [w_1 \ w_2 \ w_3]^T \quad (57)$$

$$\mathbf{h}^{(3)} = \begin{bmatrix} \mu_1^s & \sigma_1^s & \vdots & \mu_2^s & \sigma_2^s \end{bmatrix}^T \quad (58)$$

em que μ_i^e e σ_i^e são respectivamente a média e a variância da i -ésima função de pertinência de entrada; w_i é o i -ésimo fator de ponderação da regra fuzzy r_i ; μ_i^s e σ_i^s são respectivamente a média e a variância da i -ésima função de pertinência de saída.

Considerando que cada camada do sistema fuzzy tem processos distintos, e o ajuste dos parâmetros de um destes processos pouco influencia os outros, é possível realizar o ajuste paramétrico do sistema de inferência fuzzy camada a camada (FLAUZINO, 2007).

Uma vez auferida a minimização de ξ_p , têm-se os parâmetros que, juntamente com a base de regras encontrada no processo de ajuste estrutural, otimizam o sistema de inferência fuzzy. Esta minimização de ξ_p , que é o passo final para o ajuste completo do sistema de inferência fuzzy, é obtida pelo método do gradiente descendente (FLAUZINO, 2007).

4.4 Considerações Finais sobre o Capítulo

Este capítulo apresentou uma breve revisão da teoria fuzzy. Aprofundou-se nos conceitos de sistemas de inferência fuzzy por conta de ser este o modelo adotado neste trabalho para a previsão de vazão natural afluente às usinas hidrelétricas. Para a escolha desta modelagem considerou-se características importantes que são a tolerância em relação a imprecisão de dados, a capacidade em mapear sistemas não-lineares, baixo custo de implementação e interpretabilidade da solução fornecida por modelos fuzzy. Buscando obter uma estratégia mais intuitiva aplica-se o sistema fuzzy multicamadas com arquitetura *Mamdani*, cujos membros antecedentes e membros consequentes das regras serão definidos por variáveis linguísticas.

A especificação da estrutura e dos parâmetros livres do sistema de inferência fuzzy será feita de forma automática e segue o método proposto por Flauzino (2007). O ajuste estrutural será obtido a partir de uma busca por uma base de regras que permita minimizar a função custo, que expressa o desvio entre os resultados desejados e os fornecidos pelo sistema fuzzy. O número de regras será definido de forma empírica com base nos resultados das simulações.

Para um número de regras qualquer definido, o algoritmo implementado no software Matlab realiza a busca por uma base de regras com o menor custo em duas etapas. Em um primeiro momento são sorteadas diversas bases de regras e dentre estas escolhe-se a que resulta no menor custo. Tendo como ponto inicial a base de regra escolhida, parte-se para uma segunda etapa, que consiste na inserção de pequenas perturbações na base de regra original, de modo a encontrar novas combinações. A base de regras selecionada é aquela que ao final do ajuste estrutural apresenta o menor custo depois de realizadas todas as iterações.

A sintonia fina do sistema de inferência fuzzy será realizada através da adequação dos parâmetros livres do sistema. Consideram-se como funções de pertinência para as variáveis de entrada e saída as funções do tipo gaussiana assimétrica “gauss2mf”. E a sintonia fina se dá pelo ajuste das médias e variâncias das funções de pertinência das entradas e saída, bem como pela definição da ponderação de cada regra. A sintonia é feita com base no algoritmo de retropropagação de erro. O processo é realizado até que por um determinado número de iterações não se observe melhorias significativas.

Para finalizar os ajustes dos parâmetros livres, define-se para este trabalho como Γ -norma o operador “produto” e como S-norma, para a agregação das regiões fuzzy, a soma algébrica conhecida como “ou probabilístico”. O método de defuzzificação utilizado foi o centro de área implementado a partir do operador “centroid”.

Tendo apresentado a técnica a ser utilizada no estudo, o próximo capítulo traz uma descrição da base de dados utilizada e uma breve análise exploratória dos dados relevantes para treinamento do sistema de inferência fuzzy.

Análise e Tratamento dos Dados de Entrada

Para a realização deste trabalho, foram utilizadas informações de estações telemétricas e convencionais de postos hidrométricos localizados na área a montante dos reservatórios das usinas hidrelétricas. Todo o acervo de dados e informações para o desenvolvimento do estudo foram coletados do banco de dados das seguintes entidades: ONS, Agência Nacional de Águas (ANA) e Instituto Nacional de Meteorologia (INMET).

5.1 Dados de Vazão Natural

Para a elaboração do modelo preditivo, serão consideradas as vazões naturais afluentes às usinas hidrelétricas. O histórico de vazão natural para todas as usinas hidrelétricas em operação é disponibilizado pelo ONS em base diária e mensal. Atualmente, a série histórica possui dados consistidos no período entre 1931 e 2016.

Ressalta-se que a vazão hidrológica fornecida pelo operador não corresponde a uma medida física, ou seja, não é compatível com os dados observados nos postos fluviométricos. A vazão natural é uma grandeza calculada considerando o rio na sua condição natural. Faz-se uso deste artifício para isolar do processo de previsão de vazão as decisões operativas de aproveitamentos existentes a montante do reservatório da usina hidrelétrica, que influem no volume que chega nas usinas localizadas a jusante do rio.

Para se obter a vazão natural, realiza-se um processo de reconstituição de vazões naturais incrementais aos reservatórios, desconsiderando as ações antrópicas na bacia contribuinte, tais como regularizações de vazões, desvios de água e usos consuntivos. O método de reconstituição da vazão natural está disponível no Apêndice A

5.2 Dados Pluviométricos

A série temporal de precipitação média na área de drenagem de cada sub-bacia será uma das entradas do modelo fuzzy. Para sua construção, foi necessário o levantamento das estações pluviométricas disponíveis na área de drenagem de cada reservatório. Trabalhou-se com a rede de estações pluviométricas mais densa possível. No período de dados disponíveis de vazão natural (1931 a 2016), foi realizada uma ampla pesquisa de estações de medição existentes e, posteriormente, uma coleta de dados junto à Companhia Energética de Minas Gerais (CEMIG), ao Instituto Nacional de Meteorologia (INMET) e à Agência Nacional de Águas (ANA).

Para a identificação dos postos pertencentes a sub-bacia de interesse, foram consideradas as coordenadas e a delimitação da área de drenagem das bacias hidrográficas realizada no SIG ArcGis 10.6. O critério estabelecido para a seleção das estações pluviométricas foi a posição das mesmas. Foram considerados apenas os postos dentro do contorno da delimitação da bacia. Na Figura 18 é apresentada a distribuição espacial das estações pluviométricas selecionadas, juntamente com a área de drenagem existente a montante do reservatório de Camargos.

Para cada área de drenagem delimitada ¹, selecionou-se o máximo de estações pluviométricas. No inventário disponibilizado pelo Sistema de Informações Hidrológica da Agência Nacional de Águas via HidroWeb (2018) estão disponíveis 18 estações pluviométricas que exercem influência sobre a vazão incremental à usina hidrelétrica de Camargos. Este número posteriormente foi reduzido para 9 estações, em função da identificação de alguns postos com dados inconsistentes, sem histórico no período analisado ou com falhas existentes superiores a 25% do período total do estudo. No caso de Camargos, o período analisado foi de 1967 a 2016, tendo em vista que era o período que continha o maior número de dados para todos os postos pluviométricos. Mais detalhes sobre os postos selecionados são fornecidos na Tabela 4.

¹O mapa da Figura 18 e os similares apresentados ao longo deste capítulo são de elaboração própria através do *software* ArcGis 10.6 para gerenciamento de dados geográficos. Tais informações foram obtidas em plataformas públicas de divulgação de dados do tipo *shapefile* da ANA, IBGE e ANEEL.

Figura 18: Mapa de localização dos postos pluviométricos na sub-bacia de Camargos.
Fonte: Elaboração própria.

Tabela 4: Descrição dos postos pluviométricos selecionados para UHE Camargos.

Código da Estação	Nome do Posto	Latitude	Longitude	Responsável
2244065	Alagoa	-22.17	-44.64	ANA
2144025	Carvalhos	-22.00	-44.46	ANA
2144018	Aiuruoca	-21.98	-44.60	ANA
2144022	Fazenda Paraíba	-21.75	-44.35	ANA
2144019	Andrelândia	-21.73	-44.31	ANA
2144010	São Vicente de Minas	-21.70	-44.44	ANA
2144021	Fazenda Laranjeiras	-21.67	-44.34	ANA
2144006	Luminarias	-21,51	-44.92	ANA
2144007	Madre de Deus de Minas	-21.49	-44,33	ANA
2144038	Carrancas	-21.48	-44.64	ANA
2144005	Itumirim	-21.32	-44.87	ANA

Para a usina hidrelétrica de Três Irmãos foram identificadas no inventário disponibilizado *online* pelo HidroWeb 36 estações pluviométricas localizadas na área de drenagem do reservatório, cuja distribuição espacial é apresentada na Figura 19.

Figura 19: Mapa de localização dos postos pluviométricos na sub-bacia de Três Irmãos.
Fonte: Elaboração própria.

Das 36 estações pluviométricas inicialmente selecionadas, foram consideradas apenas 20 estações para a construção da série histórica de precipitação média. A Tabela 5 fornece mais detalhes sobre as estações pluviométricas selecionadas para a usina hidrelétrica de Três Irmãos, cujo órgão responsável é o Departamento de Águas e Energia Elétrica (DAEE). Os dados fornecidos contemplam o período de 1943 a 2014. Entretanto, o intervalo com maior número de dados para todos os postos pluviométricos e, portanto, o que será considerado no estudo é de 1970 a 2014.

Tabela 5: Descrição dos postos pluviométricos selecionados para UHE Três Irmãos.

Código da Estação	Nome do Posto	Latitude	Longitude	Responsável
2050016	Mennucci	-20.68	-50.92	DAEE
2050019	Major Prazo	-20.87	-50.48	DAEE
2050020	Fazenda Jacarézinho	-20.92	-50.83	DAEE
2150001	Destivale	-21.05	-50.47	DAEE
2150003	Fazenda Santa Terezinha	-21.00	-50.92	DAEE
2150004	Fazenda Santa Cecília	-21.08	-50.38	DAEE
2150006	Fazenda N.S. Perpetuo Socorro	-21.12	-50.83	DAEE
2150017	Água Limpa	-21.33	-50.52	DAEE
2150018	Fazenda São José	-21.25	-50.72	DAEE
2151002	Guaraçai	-21.03	-51.20	DAEE
2151053	Granja Yuba	-21.02	-51.10	DAEE
2050021	Fazenda Aracangua	-20.97	-50.68	DAEE
2050045	Guzolândia	-20.65	-50.67	DAEE
2051004	Bela Floresta	-20.62	-51.23	DAEE
2051007	Lussanvira	-20.73	-51.13	DAEE
2051032	Segunda Aliança	-20.93	-51.12	DAEE
2051033	Terceira Aliança	-20.92	-51.02	DAEE
2051035	Três Irmãos	-20.72	-51.33	DAEE
2051036	Ideal	-20.78	-51.12	DAEE
2150010	Araçatuba	-21.20	-50.45	DAEE

Visando o preenchimento de falhas nas séries diárias de precipitação de cada posto pluviométrico, considerou-se o método de ponderação pelo inverso do quadrado da distância, conforme a equação (59).

$$P_x = \frac{\sum_{i=1}^n \alpha_i \cdot P_i}{\sum_{i=1}^n \alpha_i} \quad (59)$$

em que n é o número de postos pluviométricos com dados disponíveis, P_i é a precipitação observada no i -ésimo posto pluviométrico e α_i é peso do i -ésimo posto pluviométrico, dado por (60)

$$\alpha_i = \frac{\left(\frac{1}{d_i^2}\right)}{\sum_{j=1}^n \left(\frac{1}{d_j^2}\right)} \quad (60)$$

e,

$$d_i^2 = (x_i - x_x)^2 + (y_i - y_x)^2 \quad (61)$$

no qual d_i é a distância entre o posto pluviométrico x e o i -ésimo posto pluviométrico, calculado a partir das coordenadas dos postos.

Optou-se pela aplicação deste método devido a sua capacidade de preservar a importância relativa de cada posto vizinho no preenchimento de dados faltantes, sem exigir um grande esforço computacional, nem a aplicação de metodologias muito complexas.

5.2.1 Análise de Consistência dos Dados Pluviométricos

A análise de consistência dos dados pluviométricos se baseia na metodologia da dupla massa descrita por Tucci (2007). Com o intuito de verificar a homogeneidade das séries pluviais, fez-se a análise de consistência de dados para todas estações das Tabelas 4 e 5.

Em suma, o método consistiu em plotar os registros totais acumulados de cada posto contra uma série acumulada representativa da mesma região, conforme apresenta-se na Figura 20 e na Figura 21, que ilustram o processo para 4 estações da área de drenagem de Camargos e Três Irmãos, respectivamente.

Figura 20: Gráficos de dupla massa para consistência dos dados das estações pluviométricas de Camargos.

O processo de consistência dos dados foi realizado mediante a análise dos gráficos duplo cumulativos, cujas referências corresponderam à média de todos os postos disponíveis para cada região. Segundo Tucci (2007), a identificação de quebras na série e o reconhecimento de mudanças abruptas, como ajustes na medição, pode ser feita mediante a identificação de não linearidades nos gráficos duplo cumulativos.

Para todos os postos analisados, não constatou-se nenhuma irregularidade, nem tampouco ausência de linearidade na série acumulada, quando comparada com a média da

Figura 21: Gráficos de dupla massa para consistência dos dados das estações pluviométricas de Três Irmãos.

região e portanto, seguindo essa métrica, comprova-se a adequabilidade dos dados fornecidos.

5.2.2 Método de Interpolação dos dados de Precipitação

Uma vez feita a análise de consistência das séries de precipitação, parte-se para a etapa de obtenção da precipitação média da sub-bacia. A construção da série de precipitação média na área de drenagem dos reservatórios pode ser obtida de duas formas: através de métodos de interpolação de dados provenientes de estações pluviométricas, como é o caso em estudo, ou por dados de radar. Devendo-se sempre levar em consideração que qualquer um dos meios descritos anteriormente são apenas representações da realidade, sendo extremamente difícil capturar de forma exata toda a variação espacial, temporal e de intensidade de precipitação que ocorre em uma sub-bacia.

Para a construção das séries históricas de precipitação média diária, fez-se uso de um método de interpolação. Dentre os vários métodos existentes, como a média aritmética, método das isoietas e método de Thiessen, o último foi selecionado.

O método de Thiessen consiste em dar pesos aos totais precipitados com base na área de influência definida para cada estação pluviométrica. A aplicação desse método assume que a precipitação em qualquer ponto da bacia é igual ao valor medido no posto mais próximo. A construção do histórico de precipitação média diária na bacia segue a equação

(62)

$$P_m = \frac{1}{A} \sum_{i=1}^n A_i \cdot P_i \quad (62)$$

em que A_i é a área de influência do i -ésimo posto pluviométrico, P_i é a precipitação registrada no i -ésimo posto pluviométrico e A é a área total da sub-bacia.

A determinação da área de influência de cada posto pluviométrico é obtida por meio da aplicação do método dos polígonos de Thiessen, disponível no programa ArcGIS. O *software* ArcGIS permite determinar para cada estação, com base na sua posição, latitude e longitude, a sua zona de influência, dada pela área contemplada pelos polígonos de Thiessen.

Para a criação dos polígonos de Thiessen, primeiramente interligam-se os postos pluviométricos por trechos retilíneos. No ponto médio da linha que interliga dois postos traçam-se linhas perpendiculares que são prolongadas até que interceptem outra mediatriz. A intercepção das linhas perpendiculares define os polígonos, os quais correspondem à área de influência de cada posto pluviométrico, conforme ilustrado na Figura 22.

Figura 22: Polígonos de Thiessen.

A vantagem de se aplicar o método de Thiessen está na facilidade de gerar séries de chuvas médias, visto que o fator de proporcionalidade de cada posto é fixo e fornecido pela geometria obtida pela posição das estações. Deve-se ressaltar, entretanto, que este método apresenta limitações, pois não leva em consideração, por exemplo, a influência do relevo no cálculo dos campos de precipitação.

5.3 Correlação entre Precipitação e Vazão Natural Afluyente

Determinada a precipitação média diária da área de drenagem de cada aproveitamento hidrelétrico, é realizada uma análise estatística da relevância dos dados de precipitação para descrição das variações da vazão natural hidrológica.

Primeiramente, analisam-se os valores anuais de vazão e de precipitação ao longo da série histórica disponível, sendo para Camargos o período compreendido entre 1967 e 2016. De modo a facilitar a comparação, tanto a precipitação quanto a vazão natural anual foram padronizadas e apresentadas na Figura 23.

Figura 23: Variação da precipitação e da vazão anual padronizadas na UHE Camargos.

Nota-se que a variação das grandezas ocorre de forma sincronizada. Os picos e os vales de praticamente toda a série histórica de precipitação coincidem com as vazões superiores e inferiores à média, inclusive no elevado valor de 1983, ano em que houve um intenso evento de El Niño, que elevou de forma atípica a vazão de Camargos (RASMUSSEN; WALLACE, 1983).

Para medir a relação existente entre a série histórica de precipitação e vazão, utilizou-se o coeficiente de correlação de Pearson, que retornou o valor igual a 0,88 para os dados anuais de vazão e precipitação da usina hidrelétrica de Camargos, comprovando a hipótese de que existe uma forte correlação entre a série histórica de precipitação e a de vazão.

No que diz respeito à análise da variabilidade da vazão com relação à precipitação em base semanal, foram considerados diversos arranjos de agrupamento temporal para a precipitação. A correlação da vazão com a precipitação foi medida para o volume

acumulado em diferentes horizontes, considerou-se a precipitação acumulada de uma até dez semanas anteriores à semana correspondente à vazão observada. Nas Figuras 24 e 25, são apresentadas as correlações cruzadas entre as afluências naturais semanais e as precipitações médias acumuladas nas estações úmida e seca.

Figura 24: Correlação cruzada entre a vazão natural semanal e as precipitações médias acumuladas para Camargos.

Figura 25: Correlação cruzada entre a vazão natural semanal e as precipitações médias acumuladas para Três Irmãos.

Por meio do exposto na Figura 24 e 25, conclui-se que a correlação da precipitação com a vazão cresce à medida que acumulam-se chuvas de mais semanas. A correlação da vazão

natural semanal com a série de precipitação acumulada em duas semanas é acima de 0,75 no período úmido para Camargos. Contudo, neste caso, valores acumulados superiores a 3 semanas não fornecem ganho de informação.

Os coeficientes de correlação calculados para as estações úmida e seca, apresentados nas Figuras 24 e 25, comprovam que o valor de correlação da vazão com a precipitação varia conforme o período do ano. Para o período chuvoso, a correlação entre a série temporal de chuva e vazão é forte, enquanto durante a estação seca a correlação torna-se mais fraca. O mesmo comportamento foi detectado para a UHE Três Irmãos.

A identificação de qual é o agrupamento temporal de precipitação mais adequado para a previsão de vazão semanal de cada usina e para cada período será feita de forma automática no algoritmo do modelo proposto. Apesar do usuário fornecer o dados de precipitação média diária da sub-bacia, a entrada do sistema de inferência fuzzy relacionada ao volume de chuva será a precipitação acumulada em n semanas, sendo a quantidade de semanas definida automaticamente, de modo a capturar a melhor grandeza a ser incorporada no treinamento do modelo fuzzy. Optou-se por uma escolha dinâmica, que é feita com base no estudo de correlação, devido à variação do melhor arranjo de agrupamento temporal conforme a bacia hidrográfica e época do ano.

5.4 Análise da Série Temporal de Vazão Natural

Inspirando-se nos modelos lineares de estimação, é comum que sejam utilizadas como entradas dos sistemas inteligentes os valores de alguns instantes da própria série imediatamente anteriores ao instante que se deseja estimar. A fim de extrair a melhor informação da série temporal, será realizada uma análise exploratória, que consiste na identificação de periodicidades e coeficientes de autocorrelação significativos para a previsão.

Neste sentido, ilustram-se nas Figuras 26 e 27 as séries de vazões naturais médias mensais que compreendem os valores observados no período de 1931 até 2016 para as usinas hidrelétricas de Camargos e Três Irmãos, respectivamente.

Nos hidrogramas, observa-se um comportamento periódico das séries de vazões hidrológicas: os períodos de cheias e recessão são bem definidos, os dados de vazões oscilam entre um valor mínimo, que ocorre na estação seca, de maio a outubro, e um máximo, que ocorre nos meses de cheia, de novembro a abril. As maiores vazões são registradas nos meses de janeiro e fevereiro, quando as vazões chegam a duplicar em relação ao período seco. Assim, partindo do pressuposto de que existe um padrão sazonal regular e estável no tempo, com picos e vales frequentemente localizados nas mesmas estações, sugere-se então o ajuste de dois modelos distintos para cada estação.

Para avaliar as relações existentes entre observações atuais e as anteriores, analisam-se as funções de autocorrelação e autocorrelação parcial da série, apresentadas na Figura 28. O interesse nos dois autocorrelogramas está, sobretudo, em descobrir valores de *lag*

Figura 26: Vazão natural mensal ao reservatório de Camargos.

fortemente significativos para a previsão. Destaca-se, entretanto, que valores de autocorrelação altos para *lags* muito espaçados, como o que se observa no *lag* 12, não devem ser considerados, uma vez que resultam da presença de um termo claramente periódico.

Dada a presença de um comportamento periódico na série, repete-se a análise da FAC

Figura 27: Vazão natural mensal ao reservatório de Três Irmãos.

Figura 28: Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) para a série de vazão afluente ao reservatório de Camargos.

Valores abaixo ou acima do intervalo de linha azul estão dentro do intervalo de confiança de 95%.

da série temporal separadamente para cada mês, conforme a Figura 29.

Figura 29: Autocorrelação nos diferentes meses para a série temporal de vazão de Camargos.

Para os meses compreendidos entre maio e outubro, estação “seca”, observa-se uma persistência do estado das vazões. A autocorrelação da vazão com os meses anteriores é

maior do que no período “úmido”. A intensidade dessa persistência varia de acordo com as características da bacia e os meses do ano, sendo maior quando o volume de precipitação é extremamente baixo. Neste caso, grande parte da variabilidade das vazões é explicada por uma componente endógena, relacionada às águas subterrâneas que abastecem o rio com as infiltrações resultantes das precipitações do período úmido (LIMA; LALL, 2007).

Seguindo os métodos estatísticos, para os quais a identificação da ordem dos modelos de previsão é realizado mediante a análise da FACP, identificam-se os *lags* que trazem informação para previsão com base na análise da função de autocorrelação parcial de cada período, ilustrada na Figura 30.

Figura 30: Autocorrelação parcial nos diferentes meses para a série temporal de vazão de Camargos.

A análise da FACP permite obter uma informação sobre autocorrelação na série sem a influência dos atrasos intermediários. Esse tipo de análise é importante, porque se o primeiro mês da série está fortemente relacionado com o segundo, e o segundo com o terceiro, então tem-se uma forte correlação do primeiro mês com o terceiro mês. A simples análise da FAC não permite identificar a relevância isolada de cada mês, uma vez que o coeficiente de autocorrelação de um *lag* k não é independente dos atrasos de 1 a k .

Prosseguindo na análise das funções de autocorrelação parcial para os dados de vazão mensal da UHE Camargos, parece ser adequado supor que um modelo autorregressivo periódico de ordem 1 é apropriado para os meses chuvosos, enquanto de ordem 4 ou

menores são apropriados para a estação seca. Esses resultados sugerem que é adequado adotar uma quantidade de componentes endógenas diferente para cada estação do ano, o que nos leva novamente à necessidade de ajuste de dois modelos de previsão.

5.5 Conclusões Finais do Capítulo

Para se obter um melhor desempenho dos modelos de aprendizado e eliminar atributos redundantes ou irrelevantes, foi realizada uma escolha criteriosa das variáveis representativas na avaliação do comportamento da vazão. Seguindo os modelos estatísticos, propõe-se como critério de seleção da primeira e terceira entrada do modelo fuzzy, a análise da FACP da série temporal de vazão natural. Serão selecionados os instantes que apresentarem os maiores coeficientes de autocorrelação parcial com o valor atual. No Matlab, a função que realiza este cálculo denomina-se “parcorr()”.

Para que o dado de entrada associado à precipitação seja de boa qualidade e representativo das características da bacia hidrográfica estudada, realizou-se uma exaustiva pesquisa na rede hidrometeorológica que compreende a área de drenagem dos reservatórios de Camargos e Três Irmãos. Considerando os dados de precipitação oriundos dos postos pluviométricos, foram realizados pré-processamentos que consistiram no preenchimento de falhas, bem como na análise de consistência dos dados. A chuva média diária das sub-bacias foi calculada considerando os dados consistidos e o método de interpolação de Thiessen.

Para encontrar um conjunto adequado de entradas, foram analisadas as correlações existentes entre a vazão natural e a precipitação média da área de drenagem para diferentes agrupamentos temporais. Para a UHE Camargos e Três Irmãos no período úmido, foram calculados coeficientes de correlação superiores a 0,7 entre a vazão natural e o agrupamento de chuva em duas semanas, demonstrando a pertinência de se incorporar a precipitação no processo de estimação.

Resultados

6.1 Considerações Iniciais

Com o objetivo de obter uma estrutura com alto grau de interpretabilidade, utilizou-se como alternativa para a solução do problema de modelagem da vazão natural afluente às usinas hidrelétricas um sistema de inferência fuzzy do tipo *Mamdani*, cujos ajustes estruturais e paramétricos foram realizados de maneira automática, conforme a abordagem proposta por Flauzino (2007).

Os modelos foram implementados em linguagem de programação MATLAB, totalmente automatizados, permitindo uma operação com mínima intervenção do usuário durante o processo. Com objetivo de demonstrar a acurácia da modelagem, serão apresentados os resultados de previsão de vazões médias semanais para a usina hidrelétrica de Camargos e Três Irmãos. Para comparar a capacidade preditiva do modelo proposto, apresentam-se também os resultados obtidos a partir da aplicação do modelo $PAR(p)$.

6.2 Dados de Entrada do Modelo Preditivo

Para treinamento e teste do sistema de inferência fuzzy, foram utilizadas as séries temporais de vazão natural e precipitação das bacias hidrográficas da UHE Camargos e UHE Três Irmãos. As vazões naturais semanais foram disponibilizadas pelo ONS e as precipitações médias semanais foram calculadas mediante o procedimento descrito no Capítulo 5. Os dados para UHE Camargos contemplam 2.557 amostras, que correspondem ao intervalo entre 1968 e 2016. Para a UHE Três Irmãos foram consideradas 2.341 amostras, que correspondem ao período entre 1970 e 2014.

A fim de evitar erros de sazonalidade, foram ajustados dois modelos para cada usina hidrelétrica, um para o período úmido (novembro a abril) e outro para o período seco (maio a outubro). A justificativa está no fato da série apresentar um comportamento periódico tanto na média, quanto na variância e na função de autocorrelação, o que instiga a busca por um modelo diferente para cada estação. As Figuras 31 e 32 mostram

para cada semana um boxplot, que resume as estatísticas da série de vazão afluyente às usinas hidrelétricas para as 52 semanas do ano.

Figura 31: Hidrograma da vazão semanal de Camargos (1931 a 2017).

Figura 32: Hidrograma da vazão semanal de Três Irmãos (1970 a 2017).

Com o objetivo de reduzir a assimetria na distribuição de probabilidade das vazões, considerou-se no ajuste do modelo proposto o logaritmo natural dos dados de vazões em vez dos valores originais. As Figuras 33 e 34 mostram a distribuição estatística da série de vazão natural semanal antes e depois da transformação.

Observa-se que a distribuição original dos dados de vazão é enviesada à direita. Ao aplicar a função logarítmica natural, tem-se uma distribuição bem mais normal.

Para as variáveis de entrada fluviométricas foram consideradas diferentes defasagens temporais. Os instantes anteriores a serem considerados na estimação foram definidos com base nos maiores coeficientes da FACP determinada para cada estação. Para ilustrar o processo de identificação dos *lags* significativos, apresenta-se na Figura 35 a FAC e FACP calculadas para a série de vazão natural de Camargos.

Figura 33: Histograma da vazão natural afluente à UHE Camargos (1968 a 2016).

Figura 34: Histograma da vazão natural afluente à UHE Três Irmãos (1970 a 2014).

As funções de autocorrelação apresentadas foram calculadas considerando a amostra completa da série temporal de vazão fluvial. Todavia, para o treinamento do modelo fuzzy, adotou-se uma análise individualizada para cada estação do ano. Ainda assim, percebeu-se para as duas amostras que o instante que apresenta a maior autocorrelação com o instante presente é o que tem atraso igual a 1 (equivalente a 1 semana), e esta autocorrelação é superior 0,8. Por isso, a primeira entrada selecionada para o treinamento e teste do modelo fuzzy, independente da estação, é $v(t - 1)$.

A segunda entrada do sistema de inferência fuzzy é a precipitação acumulada em s semanas, sendo o melhor agrupamento temporal estipulado mediante a análise da correlação cruzada entre a série de vazão natural e as séries de precipitação acumuladas em diferentes horizontes. De forma sucinta, o melhor arranjo será aquele que retorne a maior correlação da série de precipitação acumulada com a vazão, desde que

Figura 35: Autocorrelação da série de vazão Natural UHE Camargos.

$correl(v(t), \sum_{i=0}^s p(t-i)) - correl(v(t), \sum_{i=0}^{s-1} p(t-i)) > \varepsilon$, podendo s variar de 1 até um $lag_{máx}$ estipulado. O estudo completo da correlação da série de precipitação acumulada com a vazão para as duas usinas hidrelétricas foi apresentado na seção 5.3 do Capítulo 5.

Adicionalmente, foram construídos também modelos com 3 e 4 entradas. Nestes casos, as demais entradas correspondem a diferentes defasagens temporais da variável vazão natural, cujos *lags* foram definidos com base nos instantes com maior autocorrelação, desconsiderando para isso a primeira semana imediatamente anterior à previsão.

6.3 Resultados da Aplicação do Modelo Fuzzy

Para a sintonia do sistema fuzzy multicamadas, foram informados os vetores com as precipitações acumuladas em uma semana, a vazão média semanal e também, o número de regras para ajuste do modelo. A determinação do número de regras e número de entradas foi feita de forma empírica. Assim, foram feitos diversos treinamentos e testes, variando o número de regras entre 1 e 15 e limitando o número de entradas a 4. A série temporal foi dividida em 2 subconjuntos distintos: 80% da amostra foi alocada para o treinamento e 20% para verificação do desempenho do modelo. Os dados foram embaralhados para evitar que os resultados possam ser prejudicados por comportamentos hidráulicos distintos para o subconjunto de treinamento e teste. Nos gráficos e tabelas a seguir, encontram-se os dados resumidos de teste de todos os sistemas de inferência fuzzy criados. Os erros de previsão foram avaliados na escala da série original, usando-se como referência o erro percentual absoluto médio (MAPE).

Tabela 6: MAPEs dos treinamentos para o período seco de Camargos.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
1	1	23,8	36,2	32,1	31,9	22,2	29,3
	2	28,5	30,7	32,5	37,2	31,9	32,1
	3	25,4	30,0	30,9	33,2	29,8	29,9
	4	32,2	30,9	36,0	26,1	31,7	33,1
2	1	12,9	11,8	10,8	13,9	10,8	12,0
	2	15,2	11,1	11,0	11,8	13,9	12,6
	3	11,2	12,5	9,5	13,2	12,0	11,7
	4	12,5	10,7	13,0	11,2	12,7	12,1
3	1	12,2	11,0	10,9	14,2	11,6	12,0
	2	11,8	10,7	11,5	9,9	14,2	11,6
	3	10,7	10,0	10,0	12,5	11,8	11,0
	4	10,2	9,8	11,9	11,5	11,0	10,6
4	1	12,3	12,0	10,8	11,3	10,9	11,5
	2	10,5	10,1	9,9	12,4	11,3	10,8
	3	10,9	10,8	9,2	12,5	10,7	10,8
	4	10,8	12,3	12,5	10,2	10,2	11,9
5	1	12,7	11,0	10,8	10,4	11,8	11,4
	2	10,3	9,5	10,9	9,8	10,4	10,2
	3	10,9	10,8	8,8	9,1	9,9	9,9
	4	11,3	9,1	10,0	11,8	9,8	10,1
6	1	12,7	11,2	10,8	10,4	10,9	11,2
	2	9,6	10,6	9,9	9,2	10,4	9,9
	3	10,2	10,8	8,0	10,4	9,6	9,8
	4	11,0	10,0	11,2	9,7	10,3	10,7
7	1	12,4	11,4	11,2	10,4	10,9	11,3
	2	10,4	9,5	10,5	9,9	10,4	10,2
	3	11,4	10,4	8,1	10,7	10,1	10,2
	4	11,6	9,2	13,0	9,5	11,0	11,3
8	1	12,5	10,9	11,0	9,9	11,0	11,1
	2	10,5	9,9	9,3	9,9	9,9	9,9
	3	10,1	9,8	7,7	9,7	9,3	9,3
	4	10,0	9,7	11,0	9,1	10,2	10,2
9	1	12,5	11,1	11,0	9,2	10,7	10,9
	2	10,1	9,1	9,5	11,1	9,2	9,8
	3	9,8	9,7	8,0	10,5	9,3	9,5
	4	10,3	9,3	12,5	10,0	9,0	10,7
10	1	12,5	11,3	10,9	10,0	10,9	11,1
	2	9,8	9,5	10,1	10,4	10,0	10,0
	3	9,8	9,5	7,6	10,0	9,5	9,3
	4	9,6	9,4	9,6	9,5	10,2	9,6

Tabela 7: MAPEs dos treinamentos para o período úmido de Camargos.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
1	1	24,9	25,8	26,0	27,9	24,9	25,9
	2	22,9	19,6	23,9	23,4	25,9	23,1
	3	20,2	23,6	23,4	25,4	25,2	23,6
	4	26,6	24,1	21,8	23,2	24,4	24,2
2	1	14,4	15,9	15,5	16,4	15,1	15,5
	2	9,3	15,4	9,7	12,1	14,4	12,2
	3	13,0	10,4	11,0	10,7	13,5	11,7
	4	14,5	12,3	13,2	11,9	11,9	13,3
3	1	14,1	15,4	15,4	14,2	14,6	14,7
	2	8,8	9,4	9,1	10,8	12,2	10,1
	3	9,9	8,7	9,2	10,0	9,5	9,5
	4	11,5	11,8	10,6	9,4	11,0	11,3
4	1	14,3	15,4	15,6	14,8	14,6	14,9
	2	8,7	9,7	9,1	10,6	12,8	10,2
	3	9,1	9,8	9,0	9,2	9,1	9,2
	4	9,8	9,0	9,0	10,1	10,6	9,3
5	1	14,4	15,4	15,4	13,1	14,6	14,6
	2	8,3	9,4	9,4	10,6	11,1	9,8
	3	9,1	9,0	9,5	9,4	8,9	9,2
	4	9,6	8,8	8,9	8,5	10,4	9,1
6	1	14,4	15,4	15,4	13,4	14,6	14,6
	2	8,3	9,2	8,9	10,7	11,4	9,7
	3	8,8	8,5	9,5	10,0	8,8	9,1
	4	9,3	8,0	8,4	9,1	9,9	8,6
7	1	14,4	15,4	15,4	13,1	14,7	14,6
	2	8,5	9,0	9,1	10,7	11,1	9,7
	3	8,7	8,4	8,8	8,7	9,8	8,9
	4	10,0	8,3	9,0	8,8	9,5	9,1
8	1	14,6	15,4	15,4	12,9	14,7	14,6
	2	8,4	9,1	9,0	10,5	10,9	9,6
	3	8,8	8,3	8,7	8,4	8,3	8,5
	4	8,9	9,2	7,9	8,3	8,9	8,7
9	1	14,4	15,5	15,4	12,9	14,6	14,6
	2	8,3	9,3	9,1	10,3	10,9	9,6
	3	8,4	8,1	8,9	8,4	9,2	8,6
	4	9,3	9,5	8,4	8,6	9,6	9,1
10	1	14,4	15,4	15,4	12,9	14,6	14,5
	2	8,2	9,2	9,0	10,6	10,9	9,6
	3	8,1	8,6	8,7	8,4	8,1	8,4
	4	9,7	8,7	8,4	7,8	9,9	8,9

Figura 36: Média dos MAPEs dos treinamentos para a UHE Camargos no período seco.

Com base no exposto na Figura 36 e detalhado na Tabela 6, as estruturas que obtiveram o melhor desempenho no período seco foram os sistemas de inferência fuzzy com 3 entradas e 8 ou 10 regras de inferência. Para a seleção do modelo, além do melhor ajuste, considerou-se também o modelo mais parcimonioso, ou seja, aquele com a estrutura mais simples e portanto, com o menor conjunto de parâmetros possível. Portanto, considera-se como melhor modelo preditivo para o período seco o sistema de inferência fuzzy com 3 entradas e 8 regras, que retornou um erro de previsão médio de 9,3%. Neste caso, as entradas selecionadas foram a precipitação acumulada em 2 semanas e a vazão semanal em $s - 1$ e $s - 3$.

Ainda que a estrutura com 3 entradas tenha retornado o menor MAPE, deve-se ressaltar que os erros com um conjunto menor de entradas não foram muito superiores. Os testes com 2 entradas e 8 regras, por exemplo, que consideram apenas a precipitação acumulada e a vazão semanal $s - 1$, retornaram um MAPE de 9,9%. Para uma estrutura ainda mais simples, com apenas uma entrada, que corresponde a vazão semanal imediatamente anterior à vazão prevista, o MAPE cresce para 11,1%, o que significa que a informação oriunda da precipitação não traz, nesta situação, um ganho significativo para a previsão. Este resultado já era esperado, dado que a correlação entre a vazão semanal e a precipitação acumulada para Camargos é inferior a 0,5 no período seco.

Em compensação, no período úmido, quando a correlação da precipitação acumulada com a vazão natural média semanal é superior a 0,75, observa-se um ganho maior para as previsões de vazões semanais ao introduzir a informação de precipitação. Nos resultados apresentados na Figura 37, evidencia-se uma redução dos erros médios percentuais de previsão de aproximadamente 5%.

O modelo fuzzy construído com 8 regras e 3 entradas retornou um erro médio percentual de previsão no período úmido de 8,5%, ou seja, apesar da maior variabilidade das vazões ser observada neste período, auferiu-se um menor erro de previsão.

Figura 37: Média dos MAPEs dos treinamentos para a UHE Camargos no período úmido.

Observa-se um resultado parecido ao considerar os dados da UHE Três Irmãos. No período seco, o erro médio percentual de previsão de vazões médias semanais foi de 12,7%. O melhor sistema fuzzy, com 8 regras e 4 entradas, obteve um erro médio de previsão considerando os 5 testes de 13,5%. No período úmido, a melhor opção, com 7 regras e 3 entradas, retornou um erro médio de previsão de 10,3%. Para esta usina hidrelétrica, a inserção dos dados de precipitação na estimação da vazão permitiu uma redução do desvio de previsão de aproximadamente 4%.

Considerando os resultados apresentados, parece razoável supor que o dado de precipitação traz um ganho para as previsões de vazões semanais e, certamente, deverá trazê-lo também para as bacias onde se comprove uma forte correlação entre a precipitação e a vazão natural.

Tabela 8: MAPEs dos treinamentos para o período seco de Três Irmãos.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
1	1	28,7	27,0	29,7	29,6	29,8	29,0
	2	24,0	24,9	26,7	26,7	24,1	25,3
	3	22,9	22,8	23,6	25,9	25,6	24,2
	4	26,3	28,2	25,5	26,3	23,3	25,9
2	1	15,4	16,8	15,1	17,1	16,8	16,3
	2	16,0	16,0	15,9	15,9	14,5	15,7
	3	15,9	15,8	16,4	17,9	14,3	16,1
	4	17,6	15,1	18,3	13,9	18,6	17,1
3	1	14,3	16,6	14,1	16,2	15,8	15,4
	2	13,8	15,3	15,8	15,8	14,9	15,1
	3	15,3	14,2	16,5	15,6	14,7	15,3
	4	13,4	15,2	15,0	17,2	16,2	14,9
4	1	14,3	16,6	14,2	16,3	15,8	15,5
	2	13,9	15,8	15,1	15,1	14,8	15,0
	3	14,9	13,7	15,7	14,8	14,4	14,7
	4	14,3	15,8	14,8	13,6	13,7	14,6
5	1	14,5	16,8	14,2	16,4	15,8	15,5
	2	13,9	14,4	14,7	14,7	14,8	14,5
	3	14,1	12,8	14,4	15,3	14,8	14,3
	4	12,8	14,1	15,1	13,5	14,1	13,7
6	1	14,6	16,5	14,3	16,6	15,8	15,6
	2	13,7	14,3	14,9	14,9	13,6	14,3
	3	13,1	12,7	14,9	14,6	14,8	14,0
	4	13,0	14,8	14,4	14,3	13,9	13,9
7	1	14,7	16,6	14,2	16,3	15,9	15,5
	2	13,7	14,0	15,3	15,3	13,9	14,5
	3	13,6	13,3	15,1	14,1	14,5	14,1
	4	13,3	14,2	13,8	14,4	14,6	14,0
8	1	14,5	16,5	14,1	16,3	15,8	15,5
	2	13,8	13,9	14,8	14,8	13,8	14,2
	3	13,0	13,2	13,9	14,5	14,6	13,8
	4	12,9	14,1	14,3	13,1	13,7	13,5
9	1	14,6	16,6	14,3	16,3	15,9	15,5
	2	13,4	14,0	14,5	14,5	13,7	14,0
	3	13,0	12,8	14,4	14,3	14,0	13,7
	4	12,7	13,8	14,1	12,9	14,4	13,6
10	1	14,5	16,6	14,1	16,3	15,9	15,5
	2	13,5	14,0	14,9	14,9	13,3	14,1
	3	13,4	13,4	13,8	14,1	14,8	13,9
	4	14,2	13,7	14,6	13,5	14,3	14,0

Tabela 9: MAPEs dos treinamentos para o período úmido de Três Irmãos.

Número de:		Treinamento					Média
Regras	Entradas	1	2	3	4	5	
1	1	24,2	17,5	26,0	18,0	23,5	21,8
	2	20,8	20,8	16,9	17,2	20,2	19,2
	3	19,8	19,5	19,5	20,0	20,4	19,8
	4	21,5	20,7	19,4	20,8	18,7	20,3
2	1	13,8	14,6	14,5	15,3	13,3	14,3
	2	11,6	10,6	11,0	11,3	11,1	11,1
	3	11,3	12,2	10,7	11,4	12,5	11,6
	4	12,2	11,7	13,3	12,6	12,2	12,0
3	1	14,0	14,1	14,5	15,4	13,2	14,2
	2	10,7	9,9	10,6	10,8	11,1	10,6
	3	10,5	10,7	9,8	10,3	11,4	10,5
	4	12,8	10,7	11,5	10,2	10,4	11,3
4	1	13,7	14,1	14,5	14,8	13,0	14,0
	2	10,5	10,0	11,3	11,6	11,1	10,9
	3	10,6	10,3	10,2	10,4	10,7	10,4
	4	12,3	10,9	11,0	10,6	10,3	11,2
5	1	13,8	14,1	14,5	15,0	13,2	14,1
	2	10,5	10,3	10,7	10,9	11,0	10,7
	3	10,5	10,6	9,7	10,6	10,7	10,4
	4	12,1	10,8	10,9	10,1	10,9	11,3
6	1	13,7	14,1	14,4	14,8	13,0	14,0
	2	10,7	9,8	10,7	10,9	10,9	10,6
	3	11,0	10,6	9,7	10,0	10,6	10,4
	4	11,5	10,9	11,0	10,4	10,6	11,0
7	1	13,8	14,1	14,5	14,8	13,1	14,0
	2	10,5	10,2	10,7	10,9	11,3	10,7
	3	10,4	10,5	9,8	10,0	10,7	10,3
	4	11,6	10,7	10,4	10,5	10,7	11,0
8	1	13,7	14,1	14,5	14,8	13,1	14,0
	2	10,6	10,1	10,7	10,9	11,0	10,7
	3	10,5	10,6	9,9	10,1	10,8	10,4
	4	11,7	10,8	10,6	10,3	10,3	11,0
9	1	13,7	14,1	14,5	14,8	13,2	14,0
	2	10,7	10,1	10,7	10,9	10,9	10,7
	3	10,7	10,6	9,8	10,0	10,8	10,4
	4	11,8	10,4	10,5	10,3	9,8	10,7
10	1	13,7	14,1	14,4	15,0	13,0	14,1
	2	10,5	10,3	10,7	11,0	11,1	10,7
	3	10,3	10,4	10,0	10,2	10,6	10,3
	4	11,6	10,8	10,6	10,1	10,3	10,9

Figura 38: Média dos MAPEs dos treinamentos para a UHE Três Irmãos no período seco.

Figura 39: Média dos MAPEs dos treinamentos para a UHE Três Irmãos no período úmido.

Para demonstrar que os modelos fuzzy ajustados são adequados, analisam-se também os resíduos produzidos nas estimações. Assim, na Figura 40 apresenta-se o histograma dos resíduos provenientes da modelagem com 8 regras e 3 entradas para a previsão de vazão semanal da UHE Camargos, e os resíduos provenientes da previsão com um modelo fuzzy com 7 regras e 3 entradas, ajustado para a estimativa da vazão natural afluente à

UHE Três Irmãos. A distribuição dos erros de previsão assemelha-se a uma distribuição normal com média zero, o que indica que a estimação não está enviesada.

Figura 40: Histograma do erro de previsão.

(a) Resíduos da estimação para UHE Camargos e (b) Resíduos da estimação para UHE Três Irmãos.

Visto que os erros de previsão não variam significativamente com o aumento de 2 para 3 entradas, apesar do melhor resultado ter sido alcançado com o número maior de parâmetros, o sistema de inferência fuzzy do teste 1 com 8 regras e 2 entradas será explorado mais a fundo, unicamente por ser mais simples. As funções de pertinência com os parâmetros otimizados para a primeira e segunda entradas e para a saída aparecem respectivamente nas Figuras 41, 42 e 43.

Figura 41: Funções de pertinência otimizadas para a primeira entrada associada à precipitação - ajuste para a estação seca da UHE Camargos.

Com exceção da última variável linguística apresentada na Figura 46, que corresponde à saída do sistema de inferência fuzzy no período úmido, o modelo ajustado para as demais

Figura 42: Funções de pertinência otimizadas para a segunda entrada - ajuste para a estação seca da UHE Camargos.

Figura 43: Funções de pertinência otimizadas para a saída do sistema de inferência fuzzy treinado para o período seco.

Figura 44: Funções de pertinência otimizadas para a primeira entrada associada à precipitação - ajuste para a estação úmida da UHE Camargos.

variáveis linguísticas do problema considera três conjuntos fuzzy na partição da variável de entrada associada à precipitação e três conjuntos fuzzy para as variáveis linguísticas associadas à vazão natural.

As bases de regras otimizadas para cada período são apresentadas nas Figuras 47 e 48.

Figura 45: Funções de pertinência otimizadas para a segunda entrada - ajuste para a estação úmida da UHE Camargos.

Figura 46: Funções de pertinência otimizadas para a saída do sistema de inferência fuzzy ajustado para o período úmido.

1. Se (entrada_2 é vazao_3) então (saída é vazao_3) (0.8473)
2. Se (entrada_2 é vazao_2) então (saída é vazao_1) (0.95094)
3. Se (entrada_1 não é prec_3) e (entrada_2 é vazao_2) então (saída é vazao_1) (0.10633)
4. Se (entrada_1 é prec_3) e (entrada_2 não é vazao_1) então (saída é vazao_3) (7.7664)
5. Se (entrada_1 não é prec_3) e (entrada_2 não é vazao_3) então (saída é vazao_1) (0.30652)
6. Se (entrada_1 não é prec_1) e (entrada_2 é vazao_2) então (saída não é vazao_3) (1.2437)
7. Se (entrada_2 é vazao_2) então (saída é vazao_2) (1.5562)
8. Se (entrada_2 não é vazao_1) então (saída é vazao_2) (0.26243)

Figura 47: Base de regras otimizada para o período seco.

1. Se (entrada_1 é prec_1) e (entrada_2 é vazao_2) então (saída_1 é vazao_2) (1.1971)
2. Se (entrada_1 é prec_2) e (entrada_2 é vazao_2) então (saída_1 é vazao_3) (2.1158)
3. Se (entrada_1 é prec_1) e (entrada_2 é vazao_1) então (saída_1 é vazao_1) (1.8715)
4. Se (entrada_1 é prec_1) e (entrada_2 é vazao_3) então (saída_1 é vazao_3) (0.86435)
5. Se (entrada_1 é prec_3) e (entrada_2 é vazao_3) então (saída_1 é vazao_4) (0.38142)
6. Se (entrada_1 é prec_2) e (entrada_2 é vazao_1) então (saída_1 é vazao_2) (0.052939)
7. Se (entrada_1 é prec_1) e (entrada_2 é vazao_3) então (saída_1 é vazao_3) (0.924)
8. Se (entrada_1 é prec_3) e (entrada_2 é vazao_1) então (saída_1 é vazao_3) (0.51536)

Figura 48: Base de regras otimizada para o período úmido.

6.4 Comparação do Modelo Fuzzy com o PAR(p)

Para verificar a eficiência do método de previsão desenvolvido é necessário comparar os resultados obtidos com os alcançados por meio de outro modelo. Um modelo frequentemente utilizado pelo setor elétrico para modelagem das séries de vazões hidrológicas é o modelo PAR(p). Assim, seguindo a metodologia descrita em Maceira e Damazio (2004), foram ajustados 52 modelos autorregressivos para a previsão de vazão semanal afluyente aos reservatórios das usinas hidrelétricas de Camargos e Três Irmãos. Os modelos PAR(p) foram ajustados considerando a série completa de vazão semanal, que no caso de Camargos compreende o período de janeiro de 1933 a dezembro de 2016 (83 anos), e Três Irmãos, o período de janeiro de 1970 a dezembro de 2016 (46 anos).

A identificação da ordem dos modelos foi realizada segundo o método clássico de análise da função de autocorrelação parcial periódica. Nas Tabelas 10 e 11 são apresentadas as ordens dos modelos autorregressivos para as vazões de Camargos e Três Irmãos.

Tabela 10: Ordens dos modelos autorregressivos ajustados para as vazões de Camargos.

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
Ordem	1	1	1	1	2	1	1	1	1	1	3	2	1
Semana	14	15	16	17	18	19	20	21	22	23	24	25	26
Ordem	3	2	3	2	1	2	1	3	3	2	1	3	4
Semana	27	28	29	30	31	32	33	34	35	36	37	38	39
Ordem	3	4	1	3	3	3	4	3	4	3	4	4	2
Semana	40	41	42	43	44	45	46	47	48	49	50	51	52
Ordem	3	1	3	1	2	3	4	1	1	1	4	3	1

Tabela 11: Ordens dos modelos autorregressivos ajustados para as vazões de Três Irmãos.

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
Ordem	1	1	1	4	2	1	2	1	1	1	3	1	1
Semana	14	15	16	17	18	19	20	21	22	23	24	25	26
Ordem	3	1	1	1	1	1	1	3	4	4	1	2	3
Semana	27	28	29	30	31	32	33	34	35	36	37	38	39
Ordem	4	4	1	2	1	3	4	1	2	2	3	1	3
Semana	40	41	42	43	44	45	46	47	48	49	50	51	52
Ordem	1	2	1	4	3	1	1	4	2	1	3	4	1

As ordens dos modelos autorregressivos ajustados trazem uma importante informação, que é a relevância das semanas precedentes para a previsão de valores futuros de vazão semanal. Pode-se notar que, para os meses secos (semana 19 a semana 44), a ordem dos modelos selecionados, em geral, é mais elevada que para os meses de cheia, indicando que durante o período seco a correlação com as semanas anteriores é mais forte. Para o período úmido, a ordem dos modelos coincide, em apenas uma unidade.

Os parâmetros dos modelos semanais ajustados estão disponíveis no Apêndice B. Neste capítulo serão apresentados apenas os erros de previsão. Para análise, utiliza-se como referência o MAPE calculado para cada semana. Assim, na Tabela 12 e na Tabela 13 são apresentados os erros de previsão médios calculados, considerando a série temporal de Camargos e Três Irmãos. Os mesmos dados foram utilizados nos processos de estimação e previsão.

Tabela 12: MAPE calculado com a aplicação do modelo $PAR(p)$ para UHE Camargos (dados de 1933 a 2016).

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
MAPE	37,2	31,7	37,3	33,6	31,4	29,7	25,6	25,6	22,2	21,9	20,6	16,7	20,2
Semana	14	15	16	17	18	19	20	21	22	23	24	25	26
MAPE	15,2	13,6	10,2	9,0	9,0	6,8	7,1	6,7	7,2	16,1	7,6	4,9	5,7
Semana	27	28	29	30	31	32	33	34	35	36	37	38	39
MAPE	6,5	4,1	4,1	6,5	5,1	4,5	5,0	4,5	6,8	12,2	13,3	15,4	14,6
Semana	40	41	42	43	44	45	46	47	48	49	50	51	52
Ordem	16,5	16,9	18,3	18,9	24,0	21,2	24,3	28,1	26,1	30,7	29,2	34,6	37,3

Constata-se através da Figura 49, que ilustra os resultados da Tabela 12, que os meses de dezembro e janeiro apresentam maior erro percentual de previsão quando comparados com os demais meses do ano. Estes meses integram o período úmido desta bacia hidrográfica, no qual observa-se uma alta variação das afluências, tornando difícil prever valores futuros por modelos autorregressivos. Em contrapartida, para os meses do período seco, que possuem um baixo coeficiente de variação, observa-se uma diminuição do erro de previsão.

Resultados semelhantes são apresentados na Tabela 13 e na Figura 50, referentes aos erros de previsão de vazão calculados para a UHE Três Irmãos. Novamente se observa um aumento dos erros de previsão na estação com maior variabilidade das vazões. Neste caso, entretanto, ainda que os erros no período seco tenham sido menores, em nenhuma das semanas o erro caiu abaixo de 10%. Analisando os modelos ajustados, disponível na Tabela 15 do Apêndice B, observa-se que, apesar do método selecionar modelos de ordem

Figura 49: Média dos MAPEs com o modelo PAR(p) para a UHE Camargos.

mais elevadas no período seco (a exemplo a semana 23 e 43), alguns dos parâmetros $\phi_{m,j}$ intermediários calculados são nulos, ou seja, na previsão não se considera o dado de vazão observado em alguns *lags* menores do que a ordem ajustada.

Tabela 13: MAPE calculado com a aplicação do modelo PAR(p) para UHE Três Irmãos (dados de 1970 a 2016).

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
MAPE	25,8	29,0	30,2	27,7	29,1	24,1	24,9	24,6	21,4	24,0	19,8	23,6	21,8
Semana	14	15	16	17	18	19	20	21	22	23	24	25	26
MAPE	19,7	17,5	16,1	13,7	16,1	15,3	11,3	12,8	18,5	27,4	16,9	16,5	20,2
Semana	27	28	29	30	31	32	33	34	35	36	37	38	39
MAPE	14,7	10,8	12,8	14,5	12,0	15,0	10,9	13,4	13,9	13,1	15,3	14,2	21,1
Semana	40	41	42	43	44	45	46	47	48	49	50	51	52
Ordem	18,6	23,8	23,5	21,3	17,6	17,0	17,4	23,4	22,7	21,7	20,2	21,6	25,5

Tomando como base para a representação das estimativas do período seco os resultados no intervalo compreendido entre a Semana 19 e a Semana 44, tem-se para a estação seca um erro médio percentual de previsão com os modelos autorregressivos de 10% e 16,4% para a UHE Camargos e Três Irmãos, respectivamente. Para o período úmido, o erro calculado a partir dos desvios médios semanais é de 24,7% e 22,3%.

Ainda que a primeira análise tenha sido realizada com o objetivo de capturar o desempenho do modelo PAR(p) considerando um maior número de dados, a rigor a comparação

Figura 50: Média dos MAPEs com o modelo $PAR(p)$ para a UHE Três Irmãos.

entre os resultados dos dois modelos deve ser feita utilizando a mesma quantidade de dados e a mesma amostra para o modelo estatístico e fuzzy. Assim sendo, para obter uma comparação justa, apresenta-se na Figura 51 e na Figura 52 os resultados de previsão de vazão semanal obtidos com os dois modelos para o período entre 2012 e 2016 para a UHE Camargos e 2012 a 2014 para a UHE Três Irmãos.

Figura 51: Hidrograma das vazões observadas e calculadas pelos modelos para a UHE Camargos.

Os resultados dão uma indicação de que o modelo autorregressivo tem uma baixa capacidade preditiva. O desvio absoluto médio de previsão com o modelo $PAR(p)$ para o período de teste de 2012 a 2016 para a UHE Camargos foi de 24,7% e para o período de 2012 a 2014 para UHE Três Irmãos foi de 28,7%. Enquanto, o resultado combinado

Figura 52: Hidrograma das vazões observadas e calculadas pelos modelos para a UHE Três Irmãos.

dos modelos fuzzy ajustados para a estação seca e úmida retornaram um erro percentual médio absoluto de 12,6% para Camargos e 17,5% para Três Irmãos. Apesar do desvio alto, nota-se uma boa aderência da previsão por modelos fuzzy. Além disso, desaparece o efeito sombra observado com modelos $PAR(p)$, cujas previsões geralmente apresentam-se com um erro de defasagem temporal.

Os gráficos de dispersão com os dados observados e previstos para o período de 2012 a 2014 são apresentados na Figura 54.

Como pode ser visto, os dois modelos tendem a subestimar os valores altos de vazão. Uma inspeção visual desses gráficos de dispersão permite confirmar a melhor adequação do modelo fuzzy quando comparado ao modelo autorregressivo periódico.

Por meio dos resultados apresentados neste Capítulo, pode-se inferir que a estimação de vazão hidrológica semanal a partir de sistemas de inferência fuzzy tende a ser superior à modelagem a partir do modelo $PAR(p)$. Os maiores ganhos de previsão foram observados no período úmido, no qual observa-se uma redução expressiva dos erros de previsão. Isso se deve, sobretudo, à incorporação dos dados de precipitação.

Figura 53: Gráfico de dispersão com os dados observados e previstos pelos modelos para a UHE Camargos.

Figura 54: Gráfico de dispersão com os dados observados e previstos pelos modelos para a UHE Três Irmãos.

Conclusões e Trabalhos Futuros

A partir do processo de revisão, verifica-se que o modelo utilizado pelo operador nacional do sistema elétrico para a previsão de vazões fluviais em alguns aproveitamentos hidrelétricos, como os da bacia do rio Tietê, apresenta resultados de estimação de baixa qualidade. Neste sentido, com o objetivo de promover ganhos ao processo de estimação de vazão, foi estudado o uso integrado de dados de precipitação com um modelo de inteligência computacional.

Neste trabalho, aplicou-se o sistema de inferência fuzzy multicamada do tipo *Mamdani*, em função de sua característica de representar o relacionamento entre o espaço das entradas com o espaço das saídas de forma verbal, tornando possível a interpretação da solução fornecida. Esta característica deve ser pontuada como um dos aspectos motivadores para se empregar este tipo de modelo de estimação.

O ajuste estrutural e paramétrico do modelo preditor foi realizado a partir da apresentação de um conjunto de dados representativos das características da bacia hidrográfica. Considerando que o desempenho do modelo fuzzy está vinculado a qualidade dos dados fornecidos, realizou-se uma exaustiva pesquisa na rede hidrometeorológica da região, na busca por dados de precipitação de grande abrangência temporal. No intuito de encontrar a melhor base para a determinação da precipitação média da área de drenagem, fez-se um tratamento dos dados das estações pluviométricas, que consistiu no preenchimento de falhas, consistência e, posteriormente, agrupamento das séries por meio do método de interpolação de Thiessen.

Para a certificação da eficácia da modelagem proposta, associou-se a saída do sistema de inferência fuzzy aos resultados obtidos com o modelo autorregressivo periódico, amplamente utilizado pelo setor elétrico como modelo preditivo. No estudo comparativo, o sistema de inferência fuzzy obteve um desempenho superior. A redução do erro médio de previsão ocorreu de forma mais acentuada nos meses considerados úmidos, ou seja, entre novembro e abril. Nos meses secos, em geral, os resultados foram próximos. Além disso, o modelo fuzzy se mostrou robusto, pois sua aplicação em diferentes períodos hidrológicos com vazões baixas e elevadas obteve valores de erros médios semelhantes.

Adicionalmente, os resultados apresentados mostram que o modelo foi capaz de captar a variabilidade das séries de vazões estudadas. Percebe-se que as previsões realizadas por meio dos sistemas de inferência fuzzy apresentam um menor erro de defasagem temporal, quando comparado com as previsões realizadas pelo modelo autorregressivo. Isto mostra que o modelo adotado fornece uma melhor aproximação do hidrograma real, prevendo a ascensão, caso haja informação de chuva, com maior sintonia temporal em relação aos dados observados.

Os resultados apreciados vieram a comprovar a adequabilidade de sistemas de inferência fuzzy para estimação de vazão e a importância de incorporar na previsão de vazão a informação de precipitação.

Aponta-se, como limitação desta pesquisa, que tanto a etapa de treinamento quanto a de teste consideram a chuva perfeita, ou seja, aquela medida nos postos pluviométricos. Na prática, entretanto, no processo operacional de previsão de vazão serão adotados os dados de previsão de precipitação. Nota-se, portanto, que o experimento não traz uma condição realista para a utilização desta modelagem em condições operacionais, nas quais a chuva futura não é conhecida. Nesta situação, a estimativa dessa variável é feita a partir da utilização de modelos numéricos de previsão de precipitação, que podem agregar erros ao processo de previsão de vazão, interferindo no desempenho do modelo preditivo.

Partindo do pressuposto de que a estimativa de precipitação seja precisa, o modelo fuzzy é capaz de suprir a necessidade do setor elétrico de prever a vazão natural às usinas hidrelétricas com uma semana de antecedência. Os resultados obtidos mostram uma redução dos desvios de previsão de aproximadamente 11% no período úmido, com relação ao modelo $PAR(p)$, para o caso de Três Irmãos.

Assim, os objetivos traçados para esse trabalho foram atingidos e os resultados obtidos fazem com que novas linhas de pesquisa correlacionadas ao tema abordado sejam delineadas. A fim de sugerir-las, a Seção 7.1 é conduzida.

7.1 Trabalhos Futuros

Diante do que foi apresentado nesta dissertação, existem futuras investigações que podem ser efetuadas.

- Aumentar o número de calibrações: período úmido, seco e de transição.
- Estender a metodologia utilizada no presente trabalho para as demais usinas hidrelétricas.
- Estudar outras técnicas de seleção das variáveis de entrada existentes na literatura, de modo a aumentar a eficiência de estimação.

- ❑ Investir na qualidade dos dados de precipitação para melhorar o entendimento da resposta hidrológica das bacias e tornar mais confiável a aplicação do modelo proposto. Para minimizar os problemas que ocorrem nas interpolações de regiões com baixa densidade, podem ser analisados, por exemplo, os dados de precipitação do produto MERGE (ROZANTE et al., 2010).
- ❑ Para este mesmo caso estudado, seria interessante analisar o desempenho do modelo preditivo, considerando os dados de previsão de precipitação em vez do dado observado.

Referências

ANMALA, J.; ZHANG, B.; GOVINDARAJU, R. S. Comparison of ANNs and Empirical Approaches for Predicting Watershed Runoff. **Journal of Water Resources Planning and Management**, 2000. v. 126, n. 3, p. 156–166, 2000. ISSN 0733-9496.

ARFUX, G. A. B. Definição de estratégia de comercialização de energia elétrica via métodos de otimização estocástica e análise integrada de risco. 2011. p. 129, 2011.

BALLINI, R.; SOARES, S.; ANDRADE, M. G. Multi-step-ahead Monthly Streamflow Forecasting by a Neurofuzzy Network Model. In: **Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference**. [S.l.]: IEEE, 2001. v. 2, p. 992–997.

BOX, G. E. P.; COX, D. R. An Analysis of Transformations. **Journal of the Royal Statistical Society**, 1964. v. 26, p. 211–252, 1964.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis**. 4. ed. Hoboken, NJ: John Wiley & Sons, Inc., 2008. 784 p. ISBN 9781118619193.

BROWN, M.; HARRIS, C. **Neurofuzzy Adaptive Modelling and Control**. Londres, Inglaterra: [s.n.], 1994.

CARNEIRO, F. G.; ROCHA, C. H. Reforma do Setor Público na América Latina: Uma Perspectiva Comparada. In: **A Privatização no Brasil**. Rio de Janeiro: BNDES-OCDE, 2000. p. 26.

CEPEL. **Manual de Referência do Modelo NEWAVE**. Rio de Janeiro: Centro de Pesquisas de Energia Elétrica, 2006.

_____. **Modelo de Previsão de Vazões Semanais Afluentes aos Aproveitamentos Hidroelétricos do Sistema Brasileiro - Modelo PREVIVAZ**. Rio de Janeiro: Manual de Referência, 2009. 32 p.

_____. **Determinação da Coordenação da Operação a Curto Prazo - DECOMP**. Rio de Janeiro: Centro de Pesquisas de Energia Elétrica, 2018. 69 p.

Chin-Hui Liu et al. Forecasting Models for the Ten-day Streamflow of Kao-Ping River. In: **IEEE International Conference on Fuzzy Systems**. [S.l.]: IEEE, 2006. p. 1527–1534. ISBN 0-7803-9488-7.

- CHOU, S. C.; NUNES, A. M. B.; CAVALCANTI, I. A. Extended Range Forecasts over South America using the Regional ETA model. **Journal of Geophysical Research**, 2000. Instituto Nacional de Pesquisas Espaciais, São Paulo, v. 105, n. D8, p. 10147–10160, 2000.
- FADIGA, F. M. et al. Modelos de Previsão de Vazões para a Bacia Incremental à UHE Itaipu. **Revista Brasileira de Recursos Hídricos**, 2008. v. 13, p. 65–75, 2008.
- FLAUZINO, R. A. **Identificação e Localização de Faltas de Alta Impedância em Sistemas de Distribuição Baseadas em Decomposição por Componentes Ortogonais e Inferência Fuzzy**. 316 p. Tese (Doutorado) — Universidade de São Paulo, São Carlos, SP, 2007.
- Hicon Engenharia Ltda. **Análise de desempenho dos protótipos preliminares, Relatório Técnico (ONS-042/05)**. 2005.
- HILDE, V. et al. Comparison of Data-driven Takagi-Sugeno Models of Rainfall-discharge Dynamics. **Journal of Hydrology**, 2005. v. 302, p. 173–186, 2005.
- HIPEL, K. W.; MCLEOD, A. I. **Time Series Modelling of Water Resources and Environmental System**. 1. ed. [S.l.: s.n.], 1994. 1012 p. ISBN 9780080870366.
- JANG, J.-S. ANFIS: Adaptive-Network-based Fuzzy Inference System. In: **IEEE Systems, Man, and Cybernetics Society**. [S.l.]: IEEE, 1993. v. 23, n. 3, p. 665–685.
- JARDIM, D.; MACEIRA, M. E. P.; FALCAO, D. M. Stochastic Streamflow Model for Hydroelectric Systems using Clustering Techniques. In: **IEEE Porto Power Tech Proceedings**. [S.l.]: IEEE Porto Power Tech Proceedings, 2001. vol.3, p. 6. ISBN 0-7803-7139-9.
- LIMA, C. H. R.; LALL, U. Análise e Previsões de Afluências ao Reservatório de Sobradinho utilizando um Modelo Periódico Autoregressivo eXógeno (PARX). 2007. XVII Simpósio Brasileiro De Recursos Hídricos, 2007.
- LOPES, J. E. G.; BRAGA, B. P. F.; CONEJO, J. G. L. SMAP - A Simplified Hydrological Model, Applied Modelling in Catchment Hydrology. **Water Resources Publications**, 1982. Colorado, 1982.
- LOUREIRO, P. G. C. **Custo Marginal do Déficit de Energia Elétrica: Histórico, Avaliação e Proposta de uma nova metodologia**. 113 p. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, COPPE, 2009.
- LUNA, I.; SOARES, S.; BALLINI, R. An Adaptive Hybrid Model for Monthly Streamflow Forecasting. In: **IEEE International Fuzzy Systems Conference**. [S.l.]: IEEE, 2007. p. 1–6. ISBN 1-4244-1209-9. ISSN 1098-7584.
- MACEIRA, M.; DAMAZIO, J. The use of PAR(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the Brazilian hydropower system. In: **2004 International Conference on Probabilistic Methods Applied to Power Systems**. Ames, IA, USA: IEEE, 2004. p. 397–402. ISBN 0-9761319-1-9.

MACEIRA, M. et al. Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil - Main Features and Improvements on the NEWAVE Model. In: **20th Power System Computation Conference**. Dublin, Ireland: [s.n.], 2018.

MACEIRA, M. E. P. et al. Periodic ARMA models applied to weekly streamflow forecasts. **PowerTech Budapest 99. Abstract Records. (Cat. No.99EX376)**, 1999. IEEE, Budapest, p. 86, 1999.

MAGALHÃES, G. **Comercialização de energia elétrica no ambiente de contrato livre: uma análise regulatório-institucional a partir dos contratos de compra e venda de energia elétrica**. 139 p. Tese (Mestrado) — Universidade de São Paulo, 2009.

MAMDANI, E. Application of fuzzy algorithms for control of simple dynamic plant. **Proceedings of the Institution of Electrical Engineers**, 1974. v. 121, n. 12, p. 1585, 1974. ISSN 00203270.

MOREIRA, J. C.; GUILHON, L. G. F.; ROCHA, V. F. Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. **Brazilian Journal of Water Resources**, 2007. p. 7, 2007.

MORETTIN, P. A.; TOLOI, C. M. de. **Modelos para Previsão de Séries Temporais**. Poços de Caldas, MG: [s.n.], 1981. ISBN 9788521203896.

NAYAK, P. C. et al. A neuro-fuzzy computing technique for modeling hydrological time series. **Journal of Hydrology**, 2004. vol. 291, p. 52–66, 2004.

ONS. **Relatório de validação do modelo CPINS (Cálculo e previsão de incrementais e naturais a Sobradinho)**. [S.l.]: Operador Nacional do Sistema Elétrico, 2006. 68 p.

_____. **Alteração do critério de escolha do Previvaz com a utilização de técnicas de "Data Mining" para a bacia do rio Uruguai**. [S.l.]: Operador Nacional do Sistema Elétrico, 2007. 87 p.

_____. **Previsão de Vazões Diárias ao Reservatório de Três Marias usando a técnica de Redes Neurais**. [S.l.]: Operador Nacional do Sistema Elétrico, 2009. 89 p.

_____. **Relatório Anual de Avaliação das Previsões de Vazões e Energias Naturais Afluentes de 2017**. Operador Nacional do Sistema Elétrico, 2018. 100 p. Disponível em: <http://www.ons.org.br/operacao/previsao/_vazoes.a>.

_____. **Resultados da Operação - Histórico da Operação**. Operador Nacional do Sistema Elétrico, 2018. Disponível em: <http://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.a>.

PEDRYCZ, W.; GOMIDE, F. **An Introduction to Fuzzy Sets: Analysis and Design**. [S.l.]: CogNet, 1998. (A Bradford book). ISBN 9780262161718.

PEREIRA, M. F. V. Optimal Stochastic Operations Scheduling of Large Hydroelectric Systems. **International Journal of Electrical Power & Energy Systems**, 1989. 1989.

RASMUSSEN, E. M.; WALLACE, J. Meteorological aspects of the El Niño/Southern Oscillation. **Science**, 1983. p. 1195–1202, 1983.

ROSS, T. J. **Fuzzy Logic with Engineering Applications**. [S.l.]: Wiley, 2004. ISBN 9780470860755.

ROZANTE, J. R. et al. Combining TRMM and surface observation of precipitation: technique and validation over South America. **Weather and Forecasting**, 2010. p. 885–894, 2010.

SACCHI, R. **Política de operação preditiva estabilizada via termo inercial utilizando "Analytic Signal", "Dynamic Modelling" e sistemas inteligentes na previsão de vazões afluentes em sistemas hidrotérmicos de potência**. 233 p. Tese (Doutorado) — Universidade de São Carlos, 2009.

SARNAGLIA, A. J. Q.; REISEN, V. A.; LÉVY-LEDUC, C. Robust estimation of periodic autoregressive processes in the presence of additive outliers. **Journal of Multivariate Analysis**, 2010. v. 101, n. 9, p. 2168–2183, 2010. ISSN 0047-259X.

SILVA, E. L. da. **Formação de preços em mercados de energia elétrica**. [S.l.]: Sagra Luzzatto, 2001. 183 p. ISBN 9788524106460.

SILVEIRA, C. d. S. et al. Monthly streamflow forecast for National Interconnected System (NIS) using Periodic Auto-regressive Endogenous Models (PAR) and Exogenous (PARX) with climate information. **Brasilian Journal of Water Resources**, 2017. v. 22, 2017.

SOUZA, R. M. de. **Modelagem de séries periódicas via estruturas PAR(p) utilizando encolhimento wavelet**. 105 p. Tese (Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro, 2013.

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its application to modeling and control. **IEEE Transactions on Systems, Man & Cybernetics**, 1985. n. SMC -15, p. 116–132, 1985.

THOMAS, A. H.; FIERING, M. B. **Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation**. 459–463 p. Tese (Doutorado) — Cambridge, Massachusetts, 1962.

TSOUKALAS, L. H.; R. E. Uhrig. **Fuzzy and Neural Approaches in Engineering**. Nova York, USA: [s.n.], 1996.

TUCCI, C. E. M. **Hidrologia: ciência e aplicação**. Volume 4. [S.l.: s.n.], 2007. 943 p. ISBN 8570259247, 9788570259240.

UFF. Desenvolvimento de modelo de previsão de vazões com incorporação de informações de precipitação bacia do rio Iguaçu. **Relatório Técnico P4**, 2005. a, 2005. Disponível em: <http://www2.aneel.gov.br/aplicacoes/consulta_publica/documentos-NT_149_2007_Fuzz>.

VALENÇA, M. J. S.; LUDERMIR, T. B. Redes compostas por blocos de regressões sigmóides não-lineares com aplicações na previsão de séries temporais. In: **Anais do V Congresso Brasileiro de Redes Neurais**. Rio de Janeiro: [s.n.], 2001. p. 67–72.

YEVJEVICH, V. Fluctuations of wet and dry years, an analysis by variance spectrum. **Hydrology Papers Colorado State University Fort Collins**, 1977. Colorado, p. 29, 1977.

Processo de reconstituição da vazão natural em um trecho de rio

A vazão natural em um determinado ponto do rio é calculada pela soma das vazões naturais dos reservatórios a montante da usina, devidamente propagadas em condição natural, e a vazão natural incremental calculada para o reservatório, conforme a seguinte expressão:

$$Qtot_{i,t} = Qinc_{i,t} + \sum_{j \in MONT_i} Qtot_{j,t} \quad (63)$$

onde $Qtot_{i,t}$ é vazão natural no aproveitamento de interesse, $Qtot_{j,t}$ é vazão natural do j -ésimo aproveitamento a montante e $Qinc_{i,t}$ é a vazão natural incremental entre os aproveitamentos de montante e o i -ésimo aproveitamento.

Assim, as vazões naturais totais fornecidas pelo ONS para cada usina são calculadas a partir das vazões naturais incrementais somadas de todos os trecho de rio a montante do aproveitamento de interesse.

A vazão natural incremental correspondente à vazão captada entre um posto de medição e todos os postos de medição imediatamente a montante deste. As vazões naturais incrementais relativas a cada reservatório são calculadas a partir das vazões observadas em postos fluviométricos combinadas com as informações de captação de água e vazão defluente a montante da usina. O valor de vazão incremental é obtido a partir da seguinte expressão:

$$Qinc_{i,t} = Qafl_{i,t} - \sum_{j \in MONT_i} Qdef_{j,t} + Quso_{i,t} + Qevp_{i,t} \quad (64)$$

onde $Qafl_{i,t}$ é vazão afluyente ao i -ésimo reservatório, $Qdef_{j,t}$ é a vazão defluente do j -ésimo reservatório a montante, devidamente propagada em condição de reservatório, $Quso_{i,t}$ é a vazão relativa aos usos consuntivos da bacia incremental e $Qevp_{i,t}$ é a vazão relativa à evaporação do reservatório.

A vazão afluyente é definida como sendo a vazão que chega ao aproveitamento hidrelétrico. Ela depende das condições hidrológicas naturais da bacia hidrográfica e dos aproveitamentos existentes a montante. Para o seu cálculo, utiliza-se uma equação de balanço hídrico, como a definida na seguinte expressão:

$$Qafl_{i,t} = Qdef_{i,t} + \frac{\Delta V_{i,t}}{0.0864} \quad (65)$$

em que $Qafl_{i,t}$ é a vazão afluyente ao i -ésimo reservatório, $Qdef_{i,t}$ é a vazão defluente total ao i -ésimo reservatório, dada pela soma das vazões turbinadas e vertidas, e $\Delta V_{i,t}$ é a variação diária do volume acumulado, obtido a partir da leitura da tabela cota-volume do reservatório (hm^3).

Portanto, quando inexistem aproveitamentos hidrelétricos a montante de um reservatório e não há captação de água, a vazão afluyente a uma usina é também sua vazão natural, assim como a vazão natural é igual a sua incremental. Este é o caso da UHE Camargos, primeira usina da cascata da bacia do rio Grande.

Para os demais casos, a determinação da vazão incremental pode ser feita de forma direta a partir da aplicação da equação (63), considerando para isso os dados de vazões naturais disponibilizados pelo ONS.

Modelos Autorregressivos Periódicos Ajustados

Na Tabela 14 e 15 são informados os parâmetros dos modelos autorregressivos semanais para a previsão de vazão semanal da UHE Camargos e UHE Três Irmãos, respectivamente.

Tabela 14: Modelos autorregressivos ajustados para o cálculo das vazões semanais de Camargos.

Semana	ϕ_1^s	ϕ_2^s	ϕ_3^s	ϕ_4^s	Semana	ϕ_1^s	ϕ_2^s	ϕ_3^s	ϕ_4^s
1	0,63	0	0	0	27	0,43	0,09	0,47	0
2	0,67	0	0	0	28	1,20	-0,03	0,07	-0,28
3	0,61	0	0	0	29	0,96	0	0	0
4	0,63	0	0	0	30	0,37	-0,04	0,64	0
5	0,98	-0,32	0	0	31	0,40	0,23	0,36	0
6	0,70	0	0	0	32	0,86	-0,25	0,36	0
7	0,74	0	0	0	33	1,10	-0,37	0,51	-0,32
8	0,74	0	0	0	34	0,75	-0,09	0,32	0
9	0,75	0	0	0	35	0,55	0,21	-0,12	0,29
10	0,69	0	0	0	36	0,68	-0,46	0,60	0
11	0,72	-0,16	0,22	0	37	0,70	-0,26	-0,03	0,45
12	0,63	0,28	0	0	38	0,76	-0,07	-0,06	0,28
13	0,80	0	0	0	39	0,71	0,24	0	0
14	0,76	-0,21	0,29	0	40	0,53	0,09	0,27	0
15	0,56	0,29	0	0	41	0,79	0	0	0
16	0,59	0,04	0,33	0	42	0,77	-0,18	0,32	0
17	0,56	0,39	0	0	43	0,80	0	0	0
18	0,90	0	0	0	44	0,29	0,43	0	0
19	0,71	0,26	0	0	45	0,26	0,20	0,39	0
20	0,89	0	0	0	46	0,42	-0,03	0,06	0,24
21	0,94	-0,29	0,29	0	47	0,56	0	0	0
22	0,59	-0,06	0,43	0	48	0,67	0	0	0
23	1,34	-0,55	0	0	49	0,50	0	0	0
24	0,98	0	0	0	50	0,46	-0,09	0,06	0,26
25	1,57	-0,86	0,24	0	51	0,55	-0,15	0,22	0
26	0,55	0,63	-0,53	0,29	52	0,49	0	0	0

Tabela 15: Modelos autorregressivos ajustados para o cálculo das vazões semanais de Três Irmãos.

Semana	ϕ_1^s	ϕ_2^s	ϕ_3^s	ϕ_4^s	Semana	ϕ_1^s	ϕ_2^s	ϕ_3^s	ϕ_4^s
1	0,70	0	0	0	27	0,57	0,25	-0,49	0,54
2	0,62	0	0	0	28	0,63	-0,07	-0,11	0,41
3	0,71	0	0	0	29	0,87	0	0	0
4	0,88	-0,19	-0,12	0,31	30	1,00	-0,30	0	0
5	0,91	-0,35	0	0	31	0,76	0	0	0
6	0,74	0	0	0	32	0,95	-0,71	0,51	0
7	1,21	-0,60	0	0	33	0,57	-0,18	0,09	0,53
8	0,79	0	0	0	34	0,82	0	0	0
9	0,70	0	0	0	35	1,17	-0,45	0	0
10	0,70	0	0	0	36	0,37	0,43	0	0
11	0,87	-0,44	0,31	0	37	0,66	-0,14	0,34	0
12	0,64	0	0	0	38	0,92	0	0	0
13	0,66	0	0	0	39	0,31	0,30	0,30	
14	0,90	-0,42	0,40	0	40	0,88	0	0	0
15	0,82	0	0	0	41	1,14	-0,48	0	0
16	0,81	0	0	0	42	0,84	0	0	0
17	0,85	0	0	0	43	0,93	-0,35	0	0,35
18	0,86	0	0	0	44	1,03	-0,56	0,36	0
19	0,80	0	0	0	45	0,83	0	0	0
20	0,81	0	0	0	46	0,81	0	0	0
21	0,87	-0,42	0,31	0	47	0,40	0,05	-0,19	0,55
22	0,80	-0,27	-0,09	0,33	48	0,48	0,32	0	0
23	0,80	-0,22	0,01	0,29	49	0,67	0	0	0
24	0,92	0	0	0	50	0,94	-0,38	0,3	0
25	0,36	0,58	0	0	51	0,74	-0,09	-0,20	0,34
26	1,16	0,26	-0,66	0	52	0,74	0	0	0