

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CENTRO DE CIENCIAS EXATAS E TECNOLOGIA
PROGRAMA DE POS GRADUAÇÃO EM TECNOLOGIAS AMBIENTAIS

Luis Antonio Gonçalves

Otimização de Parâmetros de modelos Pluvi-Fluviométricos
Utilizando Algoritmo Genético e Busca Harmônica

CAMPO GRANDE

2008

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CENTRO DE CIENCIAS EXATAS E TECNOLOGIA
PROGRAMA DE POS GRADUAÇÃO EM TECNOLOGIAS AMBIENTAIS

Luis Antonio Gonçalves

**Otimização de Parâmetros de modelos Pluvi-Fluviométricos
Utilizando Algoritmo Genético e Busca Harmônica**

Dissertação apresentada para obtenção do grau de Mestre no Programa de Pós-Graduação em Tecnologias Ambientais da Universidade Federal de Mato Grosso do Sul, na área de concentração em Saneamento Ambiental e Recursos Hídricos.

ORIENTADOR: Prof. Dr. Jorge Luiz Steffen

Aprovada em:

Banca Examinadora:

Prof. Dr. Jorge Luiz Steffen
UFMS

Prof. Dr. Robert Schiaveto de Souza
UFMS

Prof. Dr. Silvio Bueno Pereira
UFGD

Campo Grande, MS
2008

FICHA CATALOGRÁFICA

DEDICATORIA

A minha família e Andréia Nascimento de Andrade,
Que sempre me acompanharam nesta caminhada.

AGRADECIMENTOS

Ao Professor Doutor Jorge Luiz Steffen, pela excelente orientação fornecida durante a elaboração deste trabalho.

À Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pela bolsa de estudo concedida.

A todos os amigos e colaboradores.

SUMÁRIO

LISTA DE FIGURAS	vii
LISTA DE TABELAS	viii
LISTA DE SIMBOLOS	ix
RESUMO	xi
ABSTRACT	xii
1. INTRODUÇÃO	1
2. OBJETIVOS	3
3. REVISÃO BIBLIOGRÁFICA	4
3.1. PROCESSO ESTOCÁSTICO E SÉRIE TEMPORAL	4
3.2. MODELAGEM DE SÉRIES TEMPORAIS.	7
3.3. CARACTERÍSTICAS ESTATÍSTICAS DAS SÉRIES TEMPORAIS.	11
3.4. INTELIGÊNCIA ARTIFICIAL.	12
3.5. ALGORITMOS GENÉTICOS.	13
3.5.1 REPRESENTAÇÃO.	15
3.5.2 SELEÇÃO	16
3.5.3 REPRODUÇÃO	18
3.5.4 PARAMETROS GENÉTICOS	20
3.6. PESQUISA HARMONICA.	21
4. METODOLOGIA	29
4.1. ÁREA DE ESTUDO	29
4.2. PREPARAÇÃO DOS DADOS	30
4.3. MODELO UTILIZADO	31
4.4. ESTIMATIVA DOS PARÂMETROS DO MODELO	33
5. RESULTADOS E DISCUSSÕES	35
6. CONCLUSÕES E RECOMENDAÇÕES	47
7. BIBLIOGRAFIA	48
8. ANEXOS	50

LISTA DE FIGURAS

Figura 3.1 - Exemplo de Sistema.	4
Figura 3.2 - Exemplo de Variável Aleatória.	5
Figura 3.3 - Dois processos estocásticos com valores em pontos discretos (a) e um processo estocástico em Linhas Contínuas (b), Salas & Smith (1980).	6
Figura 3.3 - Caminho sistemático de modelagem de séries hidrológicas, Salas & Smith (1980).	8
Figura 3.4 – Tipos de previsão em tempo real. Base na precipitação, Tucci (1998).	9
Figura 3.5 – Tipos de previsão em tempo real. Base na vazão, Tucci (1998).	10
Figura 3.6 – Tipos de previsão em tempo real. Base na precipitação e vazão, Tucci (1998).	10
Figura 3.7 - Esquema funcional de um Algoritmo Genético.	15
Figura 3.8 - Representação dos Algoritmos Genéticos (<i>AGs</i>).	16
Figura 3.9 - Seleção utilizando o método da roleta.	17
Figura 3.10 - Operação de <i>Crossover</i> de um ponto.	19
Figura 3.11 – Operação de <i>Crossover</i> uniforme.	19
Figura 3.12 – Operação de mutação.	20
Figura 3.13 – Analogia entre improvisação musical e otimização Geem & Lee (2004).	23
Figura 3.14 – Seleção pelo método do sorteio.	24
figura 3.15 – conjunto dos possíveis valores para cada variável de decisão.	25
figura 3.16 – notas armazenadas na memória harmônica.	25
Figura 3.17 – Estrutura da Memória Harmônica.	26
Figura 3.18 - Esquema funcional da Pesquisa Harmônica.	28
Figura 4.1 – Esquema de localização dos Postos Fluviométricos.	29
Figura 4.2 - Localização dos Postos Pluviométricos utilizados, Neto (2004).	30

LISTA DE TABELAS

Tabela 3.1 - Técnicas e Inspirações.	13
Tabela 3.2 - Notas dos cinco indivíduos de uma população.	17
Tabela 3.3 – Comparações entre otimização e performance musical.	22
Tabela 3.3 – Comparações entre otimização e performance musical.	22
Tabela 4.1 – Postos pluviométricos com área de influência, Neto (2004).	31
Tabela 5.1 – Valores de desvio padrão.	46

LISTA DE SIMBOLOS

S	espaço associado ao experimento.
X	função que associa a cada elemento.
s	$\in S$ a um numero real.
μ	média da população.
σ^2	variância.
ξ_t	série independente com média zero.
ϕ	parâmetro do modelo.
$\bar{x}$	média simples.
N	tamanho da amostra e numero de variáveis de decisão.
x_i^k	variável de decisão.
X_i	range das variáveis.
HMS	número de vetores solução na memória harmônica.
$HMCR$	memória harmônica considerando a taxa.
PAR	taxa de ajuste.
NI	número de gerações no vetor solução.
E	número do ensaio.
s^2	variância simples.
E	vazão de entrada.
Q	vazão de saída.
h_i	altura da precipitação.
α_i	porcentagem da precipitação.
P_i	precipitação registrada no posto i num determinado tempo.
A_i	a área de influencia do posto i .
A_b	área da bacia.
np	número de postos.
β	porcentagem do volume armazenado escoado acima da tendência de escoamento sem a ocorrência de chuvas.
$Q(t)$	é a vazão de saída observada no tempo t .
$\hat{Q}(t)$	vazão calculada pelo modelo.

m	número de observações
T	designa a matriz transposta.

RESUMO

GONÇALVES, L. A. (2008). *Otimização de Parâmetros de modelos Pluvi-Fluviométricos Utilizando Algoritmo Genético e Busca Harmônica, Campo Grande, 2008, 55 p. Dissertação (Mestrado) – Universidade Federal de Mato Grosso do Sul, Brasil.*

Durante milhares de anos os rios foram o principal meio de transporte para a população. Com a agricultura e a indústria, a água passou a ser também fator importante na produção, sendo os rios fornecedores naturais de água em quantidade e relativa qualidade.

Mas do mesmo modo que eles proporcionam facilidades, também propiciam catástrofes, ocasionadas principalmente devido aos transbordamentos do seu leito. Com o conhecimento do seu comportamento, aproveita-se melhor suas possibilidades de uso, proporcionando maior segurança à população.

A construção de obras, denominadas medidas estruturais, é um dos meios de minimizar tais riscos, mas de grande vulto quanto a custos financeiros e ambientais. Outro método de diminuir prejuízos é o monitoramento das condições do Recurso Hídrico, sendo chamado de medidas não estruturais, apresentando vantagens em relação às primeiras.

O estudo das séries históricas de dados hidrológicos permite conhecer, modelar e prever condições futuras críticas do rio, maximizando assim as ações no combate as enchentes, tanto nas medidas estruturais como nas não estruturais.

Para tal modelagem existem vários métodos conhecidos, dentre os quais o método dos Algoritmos Genéticos, que imita o método de seleção natural, e o método da Pesquisa Harmônica que deriva do conceito harmônico de banda.

Palavras-chave: precipitação-vazão, algoritmos genéticos, pesquisa harmônica.

ABSTRACT

GONÇALVES, L. A. (2008). *Otimização of Parameters of models Pluvi-Fluviométricos Utilizando Algoritmo Genético e Busca Harmônica, Campo Grande, 2008, 55 p. Dissertação (Mestrado) – Universidade Federal de Mato Grosso do Sul, Brasil.*

For thousands years the rivers were the main transport way for population. With the agriculture and industry, water has also become an important factor in production, and the rivers, natural suppliers of water in quantity and relative quality.

But in the same way they provide facilities, they also provide disaster, mainly caused due to its bed tranship. With the knowledge of rivers behavior, it takes up more possibilities for using, providing greater security for population.

The works construction, called structural measures, is one of the means of minimizing those risks, but of high financial and environmental costs. Other method of reducing losses is monitoring water resource conditions and is called non-structural measures, with advantages in relation to the first.

The study of hydrological data historical series allows knowing, modeling and predicting future critic river conditions, thus maximizing actions in fighting the floods, both in structural measures such as non-structural.

For this modeling there are several known methods, among them, the method of Genetic Algorithms, which imitate the natural selection method, and the Harmonic Search method which is derived from the harmonic concept of band.

Key-words: precipitation-flow, genetic algorithms, harmonic search.

1. INTRODUÇÃO

Durante milhares de anos os rios foram o principal meio de transporte para a população. Com o advento da agricultura e posteriormente da indústria, a água foi um fator importante na produção. Sendo os rios o modo mais fácil de se obter água em quantidade e em até relativa qualidade, suas margens sempre foram causa de disputas territoriais, onde antigamente grande parte dos grandes produtores se localizava nas suas proximidades.

Representaram então, via de transporte e meio de produção entre tantas outras utilidades, sendo as primeiras cidades localizadas sempre as margens dos mesmos, devido às tais facilidades. Mas do mesmo modo que eles proporcionam facilidades, também propiciam catástrofes, ocasionadas principalmente devido aos transbordamentos do seu leito ocasionando inundações.

Com o crescimento das cidades e a falta de um zoneamento urbano adequado, as áreas sujeitas a inundações, antes desabitadas, foram sendo ocupadas, provocando nas épocas de enchentes, prejuízos econômicos e sociais.

O conhecimento do seu comportamento é de suma importância para um melhor aproveitamento das suas possibilidades de uso, proporcionando segurança às populações ribeirinhas e as atividades localizadas às margens dos mesmos.

A cidade de Aquidauana, localizada as margens do rio com mesmo nome, periodicamente sofre com as enchentes, causando enormes prejuízos à população, conforme divulgado em jornais locais. Por isso há grande interesse no desenvolvimento de metodologias que minimizem tais prejuízos.

A construção de obras, denominadas medidas estruturais, é um dos meios de minimizar tais riscos, mas de grande vulto quanto a custos financeiros e ambientais. Alguns conceitos da Engenharia são de dimensionar obras hidráulicas que aumentam o fluxo das ondas de cheia de montante para jusante das cidades mudando o problema de uma região para outra da cidade, Canholi (2005).

Outras medidas estruturais muito utilizadas atualmente são obras que propiciam reter o pico de cheia ou aumentam o processo de infiltração ou retardam o fluxo nas calhas dos córregos e rios.

Outro método de diminuir prejuízos é o monitoramento das condições do Recurso Hídrico, sendo chamado de medidas não estruturais, apresentando vantagens em relação as primeiras, Steffen (1987).

Dentre as medidas não estruturais, destacam-se, Tomaz (2002):

- Programa de educação ambiental;
- Participação publica na gestão dos recursos hídricos da bacia;
- Programa de previsão hidrometereológica;
- Plano de contingência para episódios críticos de inundações;
- Intensificação de medidas de controle na bacia;
- Disposição de resíduos sólidos;
- Coleta e tratamento de efluentes domésticos e industriais; e
- Controle do reuso da água.

O estudo das séries históricas de dados hidrológicos permite conhecer algumas características estatísticas das mesmas e sua posterior modelagem matemática, permitindo prever condições futuras críticas do rio, maximizando assim as ações no combate as enchentes, tanto nas medidas estruturais como nas não estruturais.

Para tal modelagem existem vários métodos conhecidos, dentre os quais o método dos Algoritmos Genéticos, que imita o método de seleção natural e o método da Pesquisa Harmônica que deriva do conceito harmônico de banda.

2. OBJETIVOS

O objetivo deste trabalho é determinar os valores de vazão em uma secção do curso d'água a partir dos dados de chuva registrados em postos pluviométricos de influência na bacia de drenagem.

Para tanto foram ajustados modelos matemáticos relacionando os dados de chuva e vazão que levam em conta as áreas de influência dos postos e coeficientes de redução dos valores de chuva, indicando as perdas por evaporação e infiltração.

No ajuste das séries históricas de precipitação e vazão foram utilizados dois métodos de otimização de parâmetros: os Algoritmos Genéticos (*AGs*) e a Pesquisa Harmônica (*PH*).

Comparar a eficiência de cada método, sua facilidade de aplicação e os tempos de convergência.

Definiram-se assim os objetivos desse trabalho:

- Implementar um modelo de geração de vazão pelos dados fluviométricos e pluviométricos dos postos influentes na bacia em estudo.
- Avaliar o erro na implementação e previsão da vazão no posto e eventos de cheia estudados.
- Utilizar as técnicas de inteligência artificial para ajuste e teste dos modelos.
- Comparar os modelos e técnicas utilizadas.

3. REVISÃO BIBLIOGRÁFICA

A presente revisão tem a finalidade de levantar as informações atualizadas sobre modelagem matemática, os processos de estimativa de seus parâmetros, as restrições e definições necessárias para utilização dos modelos citados.

3.1. PROCESSO ESTOCÁSTICO E SÉRIE TEMPORAL

Qualquer estrutura, dispositivo, esquema ou procedimento, real ou abstrato, que inter-relaciona em um dado tempo de referência, uma entrada, causa ou estímulo a uma saída, efeito ou resposta de informação, energia ou matéria é denominado de sistema, Dooge (1973). Portanto, pode ser considerado como qualquer engenho que através de uma saída, responde a uma entrada, possuindo características próprias, podendo ser dividido em subsistemas onde certas características são mais homogêneas, Tucci (1998). Uma bacia hidrográfica é um sistema onde uma entrada (por exemplo, à precipitação) produz uma saída (vazão), dependendo esta de suas características (solo, vegetação, topografia, etc), Figura 3.1.

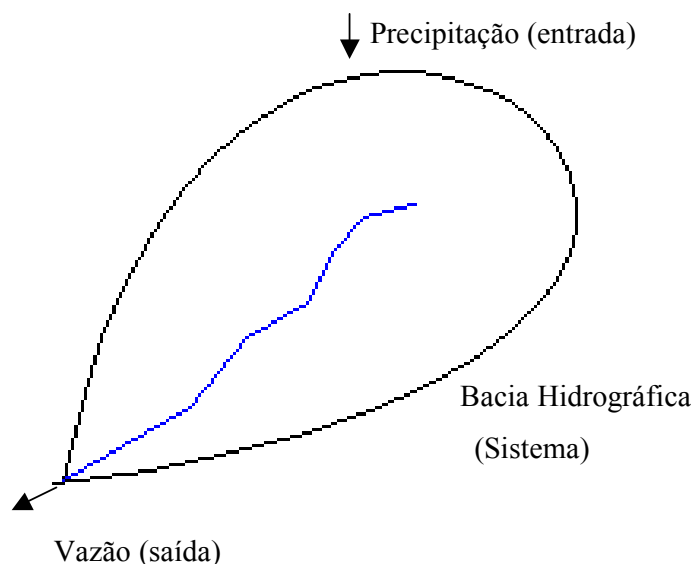


Figura 3.5 - Exemplo de Sistema.

Considerando uma variável de entrada de um sistema denotado por x e que após passar por um processo apresente uma saída e a saída desta variável pode ser predita com

certeza, ela é dita como variável determinística. Porém, se a sua saída não puder ser predita com certeza, ela é denominada de variável aleatória. Neste caso, a saída de x será expressa em probabilidades ou espaço probabilístico, sendo governado por lei ou função de probabilidade.

Variável aleatória é uma variável cujo valor é o resultado numérico de um experimento aleatório. Ela é uma função formada por valores numéricos definidos sobre o espaço amostral de um experimento, Lapponi (2005).

A variável aleatória será discreta se o número de valores possíveis de X (seu contradomínio) for finito (apenas pontos específicos) ou contínua, caso seu contradomínio seja um intervalo ou uma coleção de intervalos (assume qualquer valor em torno de uma linha central ou qualquer valor entre dois valores limites), conforme Figura 3.2 e 3.3.

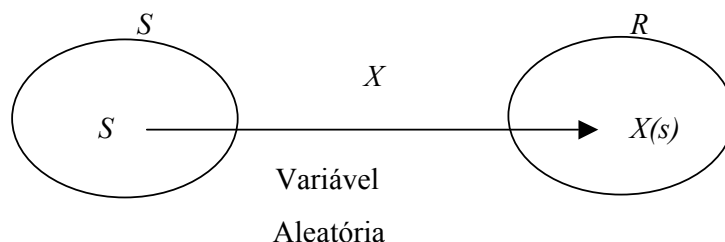


Figura 3.6 - Exemplo de Variável Aleatória.

onde S é o espaço associado ao experimento, X é uma função que associa a cada elemento $s \in S$ a um numero real $X(s)$.

Assumindo como $x_1, x_2, \dots, x_n$ a saída de x , onde o sub-escrito representa o ordenamento de tempo, distância, etc. A seqüência neste intervalo de tempo é denominada de serie temporal, Salas & Smith (1980). São exemplos de série temporal: os valores mensais de temperatura na cidade de Campo Grande, ou a cota do rio Paraguai em Corumbá, Figura 3.3.

Uma serie aleatória é também chamada de série probabilística ou serie estocástica. O processo estocástico pode ser pensado como um conjunto de possíveis trajetórias que poderiam ter sido observadas, ou como um conjunto de variáveis aleatórias, uma para cada tempo t . Cada valor observado de uma trajetória é um dos possíveis valores que poderiam ter sido observados de acordo com a distribuição de probabilidades da respectiva variável aleatória.

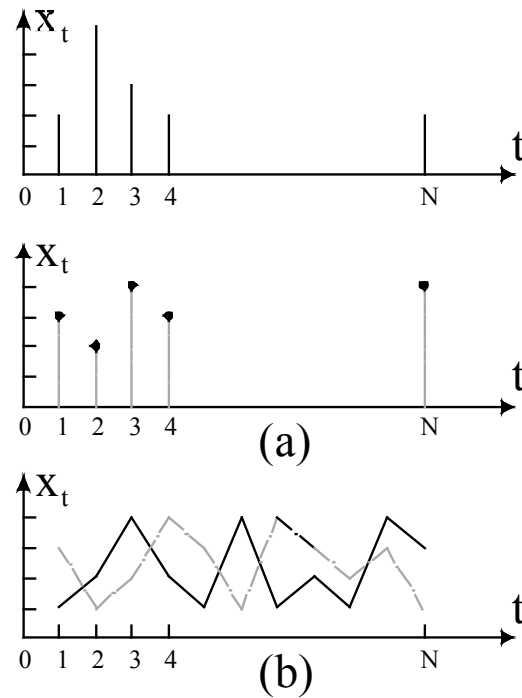


Figura 3.3 - Dois processos estocásticos com valores em pontos discretos (a) e um processo estocástico em Linhas Contínuas (b), Salas & Smith (1980).

Um modelo matemático que representa um processo estocástico é denominado modelo estocástico ou modelo temporal. Ele pode ser representado por uma função única de distribuição de probabilidade $f(x, \theta)$ com os parâmetros $\theta = (\theta_1, \theta_2, \dots)$ válidos para toda posição $t = 1, 2, \dots, N$ e sem qualquer dependência entre $x_1, x_2, \dots, x_N$. Para o instante t , se x é normal com média μ e variância σ^2 , o modelo pode ser convenientemente escrito como:

$$x_t = \mu + \sigma \varepsilon_t \quad (3.1)$$

onde ε_t é também normal com média zero e variância um, e $\varepsilon_1, \varepsilon_2, \dots$ são independentes.

$$\varepsilon_t = \phi \varepsilon_{t-1} + \xi_t \quad (3.2)$$

onde ξ_t é uma série independente com média zero e variância $(1 - \phi^2)$, ε_t é a série dependente, ϕ é o parâmetro do modelo.

As séries hidrológicas podem ser divididas em dois grupos básicos:

- Séries simples ou univariadas para um ponto específico, nos quais as características da série de interesse são explicadas exclusivamente a partir do comportamento da própria série.

- Série múltipla ou multivariada para vários pontos ou séries múltiplas de diferentes tipos para um ponto, na qual as características da série de interesse são explicadas em função não apenas do comportamento da própria série, mas também de outras séries, permitindo, assim, relações de interdependência e causalidade.

Ao conjunto total de todas as saídas, respostas ou estímulos, é dado o nome de população. Entretanto, como na maioria das vezes é não se possui todos os valores de saída, seleciona-se então um subconjunto ou parcela da população, denominada de amostra. Portanto, uma parcela da série contida entre dois pontos é chamada de amostra.

Denomina-se de evento uma amostra ou intervalo da série que representa um evento de cheia em uma dada seção de controle. Um evento de cheia caracteriza-se pela elevação do fluxo na calha do rio em um dado período prolongando-se até o retorno do fluxo ao seu comportamento inicial.

3.2. MODELAGEM DE SÉRIES TEMPORAIS

O entendimento do mundo sempre foi um dos desejos do ser humano, tendo de início a luta pela sobrevivência como sua primeira necessidade, onde a ação sobre o meio prevalecia sobre a compreensão do mesmo.

Na medida em que as necessidades do ser humano foram se tornando mais complexas, cresceram as carências por aperfeiçoar o processo de compreensão do mundo, Goldbarg (2000), sendo essa busca da realidade (esclarecimento) um processo de modelagem.

- Modelo: é um veículo para uma visão bem estruturada da realidade, também podendo ser entendido como uma representação substitutiva da realidade, Goldbarg (2000).

A modelagem de séries temporais é um processo que pode ser simples ou complexo, dependendo das características avaliadas na série. Os modelos para serem implementáveis, devem ser livres de pequenos detalhes onerosos, Goldbarg (2000).

Os modelos hidráulicos podem ser classificados como:

- modelos de comportamento: descrevem o comportamento de um sistema, utilizado principalmente para antecipar como será a sua resposta após mudanças de suas características. Assim, o modelo é utilizado para prever as condições futuras da bacia, Tucci (1998).
- Modelos de otimização: preocupam-se com as melhores soluções, em nível de projeto de um sistema específico. Neste caso, o modelo serve apenas para cálculo dos parâmetros a serem otimizados, Tucci (1998).

- Modelos de planejamento: simulam condições globais de um sistema maior, envolvendo uma região ou bacia buscando não somente a solução hidráulica, hidrológica ou econômica, mas, também, as quantificações socioeconômicas e ambientais, Tucci (1998).

Um caminho sistemático para a modelagem é composto de seis passos, Figura 3.4, Salas & Smith (1980):

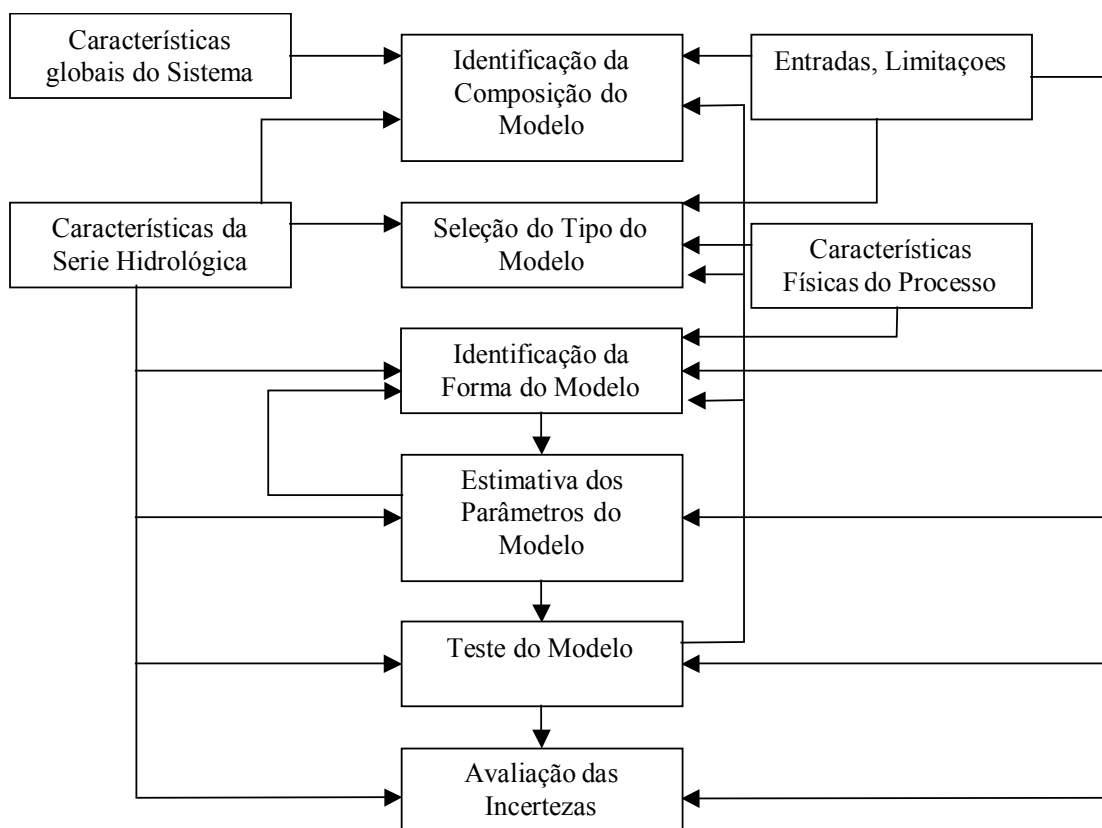


Figura 3.7 - Caminho sistemático de modelagem de séries hidrológicas, Salas & Smith (1980).

- 1- Identificação da composição do modelo: geralmente, em qualquer modelagem, tem-se que decidir se o modelo será univariado ou multivariado, ou uma combinação de univariado e modelo desagregado, ou uma combinação de multivariado e um desagregado, etc. Essa decisão depende das características globais do sistema, das características da série e das entradas do modelo.
- 2- Seleção do tipo do modelo: uma vez identificada a composição do modelo, o tipo pode ser selecionado. Os modelos podem ser classificados como empíricos, conceituais ou combinados. Os conceituais utilizam conceitos hidrológicos para sua formulação,

possuindo assim maior capacidade de extrapolação e tratamento das variantes hidrológicas e menos capacidade de atualização. Os empíricos não utilizam conceitos dos processos hidrológicos tendo como vantagem a fácil atualização dos parâmetros em tempo real. O modelo combinado aproveita o melhor de cada um dos modelos citados, figuras 3.4, 3.5 e 3.6, Tucci (1998).

- 3- Identificação da forma do modelo: uma vez selecionado o tipo, a terceira fase da modelagem é identificar a forma do modelo, isto é, determinar quantos e quais serão os parâmetros utilizados no modelo.

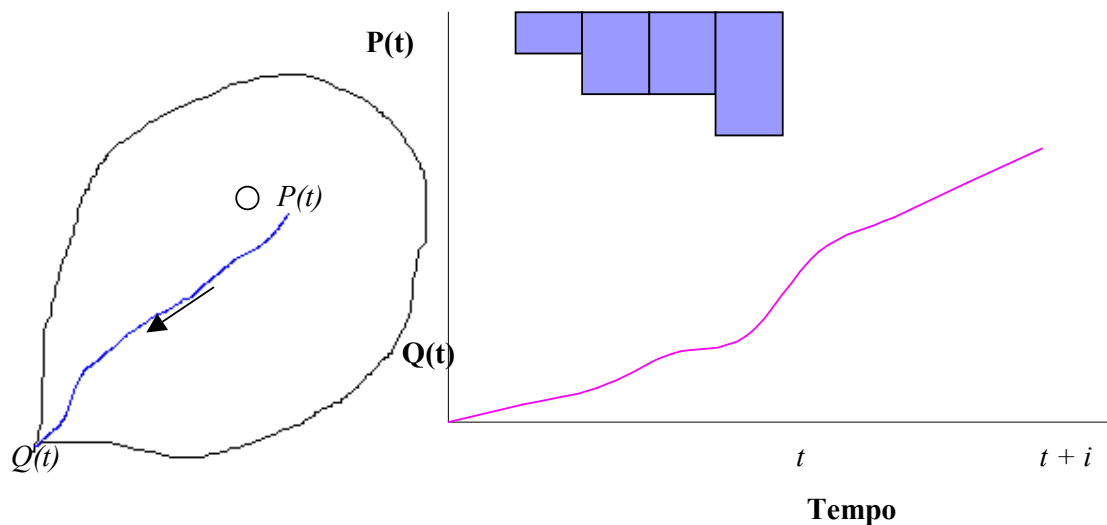


Figura 3.8 – Tipos de previsão em tempo real. Base na precipitação, Tucci (1998).

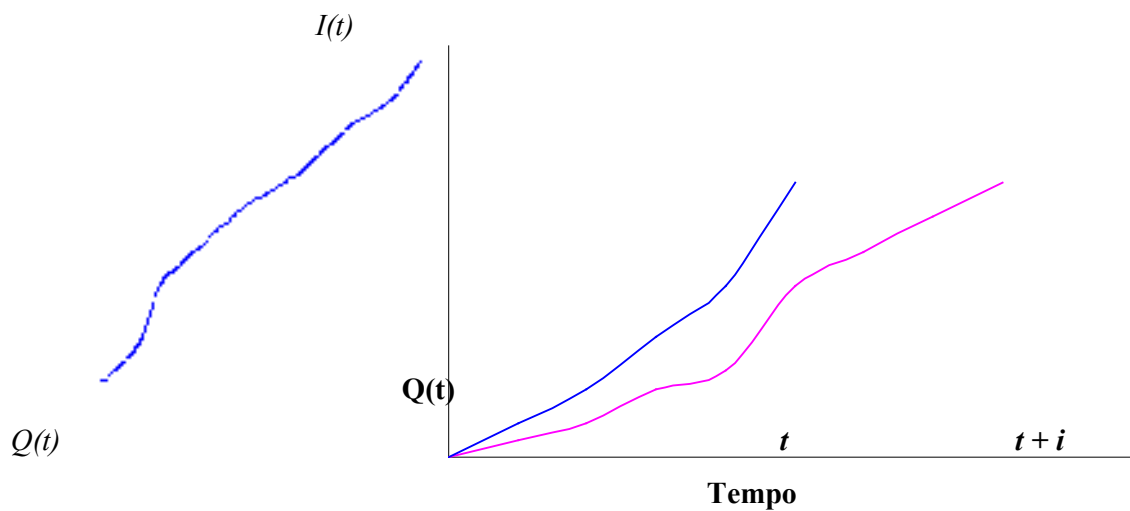


Figura 3.5 – Tipos de previsão em tempo real. Base na vazão, Tucci (1998).

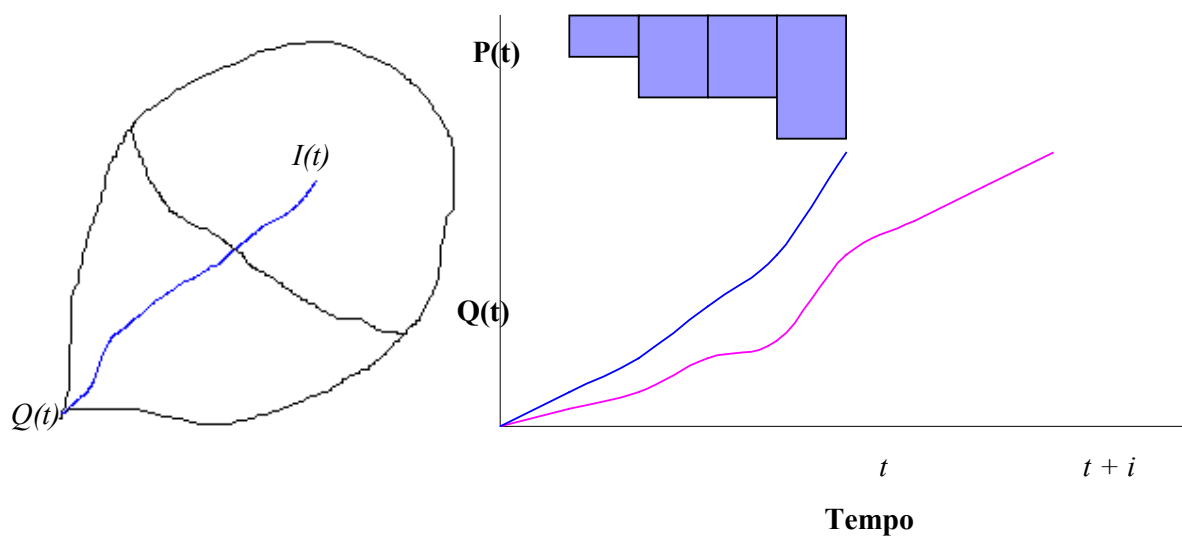


Figura 3.6 – Tipos de previsão em tempo real. Base na precipitação e vazão, Tucci (1998).

- 4- Estimação dos parâmetros do modelo: identificado o modelo, é necessário estimar os parâmetros do modelo. Os parâmetros são influenciados pelas condições iniciais,

como umidade do solo da bacia, perdas iniciais entre outros representando estes processos espaciais de forma concentrada.

- 5- Teste do modelo: verifica-se então a representatividade do modelo. O modelo suposto pode ser verificado através de vários indicadores, sendo o mais usualmente utilizado o erro quadrático médio.
- 6- Avaliação das incertezas: após o modelo ser julgado adequado, resta avaliar as incertezas. Dois tipos de incertezas são encontrados em análise de séries hidrológicas: (a) incertezas do modelo, e (b) incertezas dos parâmetros. Incertezas do modelo porque o verdadeiro modelo não é conhecido e incertezas dos parâmetros pois eles são estimados de uma quantidade limitada de dados.

3.3. CARACTERÍSTICAS ESTATÍSTICAS DAS SÉRIES TEMPORAIS

Os modelos são formulados para reproduzir as características estatísticas da série histórica, mas devido às incertezas, deve-se escolher quais características devem ser reproduzidas e como elas serão interpretadas ou entendidas. Infelizmente, esta não é a única questão. Primeiramente, o modelo matemático exato não é conhecido, pois a verdadeira população é desconhecida, sendo os dados observados apenas um número finito de ocorrências (N) e as características são apenas estimativas. Além disso, tem-se o problema de definição e interpretação. Algumas características importantes são a média e o desvio padrão. Geralmente, a média e o desvio padrão são os que possuem menor incerteza. Outra característica, como a assimetria e a autocorrelação, tem alta incerteza, pois dependem da quantidade de amostras da série. A interpretação destes decide o tipo de modelo a ser usado e sua melhor forma, Salas & Smith (1980).

Para avaliar a qualidade das previsões efetuadas, empregou-se os seguintes indicadores estatísticos:

- a) Média simples: a media simples de uma série x_t , $t = 1, 2, \dots, N$ é dada por:

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t \quad (3.3)$$

onde N é o tamanho da amostra. A média simples é a estimação da media da população μ , e mede a tendência central para x_t ou determina onde a série é locada como um todo.

- b) Variância simples: a variância simples é dada por:

$$s^2 = \frac{1}{N} \sum_{t=1}^N (x_t - \bar{x})^2 \quad (3.4)$$

sendo esta uma estimativa tendenciosa para a variância da população, cuja estimativa não tendenciosa é obtida por

$$s^2 = \frac{1}{N-1} \sum_{t=1}^N (x_t - \bar{x})^2 \quad (3.5)$$

A raiz quadrada de s^2 é denominada desvio padrão. Dividindo a média pelo desvio padrão tem o coeficiente de variação. A média mede a localização para a série, o desvio padrão mede a dispersão ou a extensão da série em torno da média $\bar{x}$. Um pequeno s significa que os valores de x não diferem muito da média $\bar{x}$, enquanto que um grande valor de s geralmente significa que os x 's tem grande dispersão em torno da média.

3.4. INTELIGÊNCIA ARTIFICIAL

A inteligência artificial é uma área da ciência que busca, através de técnicas inspiradas na natureza, o desenvolvimento de sistemas inteligentes que imitam aspectos do comportamento humano, ou seja, a sua capacidade de adaptação às circunstâncias de seu meio, Aurélio *et al* (1999).

Esta inteligência seria a sua capacidade genética de solução de problemas. Por capacidade genética entenda-se todo o conhecimento embutido ao nível de programação, o que permite um determinado conjunto de estados possíveis de funcionamento através de programas. A inteligência de máquina é, portanto um tipo de inteligência construída pelo homem, logo, uma inteligência artificial. Mas o conceito de Inteligência Artificial (IA) abrange mais do que a inteligência de máquina, pretende-se, com ela, capacitar o computador de um comportamento inteligente.

Por comportamento inteligente devemos entender atividades que somente um ser humano seria capaz de efetuar. Dentro destas atividades podem ser citadas aquelas que envolvem tarefas de raciocínio (planejamento e estratégia) e percepção (reconhecimento de imagens, sons, etc.), aprendizado, evolução e adaptação. Tais técnicas podem ser exemplificadas na Tabela 3.1.

Redes Neurais são modelos computacionais não lineares, inspirados na estrutura e operação do cérebro humano, que procuram reproduzir algumas características humanas, tais como: aprendizado, associação, generalização e abstração. Redes Neurais são efetivas no aprendizado de padrões a partir de dados não lineares, incompletos, com ruído ou compostos de exemplos contraditórios, Aurélio *et al* (1999).

Algoritmos Genéticos são algoritmos matemáticos inspirados nos mecanismos de evolução natural e recombinação genética. A técnica de Algoritmos Genéticos fornece um mecanismo de busca adaptativa que se baseia no princípio Darwiniano de reprodução e sobrevivência dos mais aptos, Aurélio *et al* (1999).

Tabela 3.1 - Técnicas e Inspirações.

Técnica	Inspiração
Redes Neurais	Neurônios biológicos
Algoritmos Genéticos	Evolução biológica
Pesquisa Harmônica	Performance musical
Lógica Fuzzy	Processo linguístico
Sistemas Especialistas	Inferência

Pesquisa Harmônica é um algoritmo baseado na performance musical. A técnica de Busca Harmônica fornece resultados em tempos de implementação relativamente pequenos em relação a outros mecanismos de busca.

Lógica Nebulosa (*Fuzzy Logic*) tem por objetivo modelar o modo aproximado de raciocínio humano, visando desenvolver sistemas computacionais capazes de tomar decisões racionais em um ambiente de incerteza e imprecisão. A Lógica Nebulosa oferece um mecanismo para manipular informações imprecisas, tais como os conceitos de muito, pouco, pequeno, alto, bom, quente, frio, etc, fornecendo uma resposta aproximada para uma questão baseada em um conhecimento que é inexato, incompleto ou não totalmente confiável, Aurélio *et al* (1999).

Sistemas Especialistas são programas computacionais destinados a solucionar problemas em um campo especializado do conhecimento humano. Usa técnicas de *IA*, base de conhecimento e raciocínio inferencial, Aurélio *et al* (1999).

No trabalho proposto, pretende-se utilizar as técnicas de Algoritmos Genéticos (*AGs*) e Pesquisa Harmônica (*PH*) para estimativa dos parâmetros dos modelos.

3.5. ALGORITMOS GENÉTICOS

Algoritmos Genéticos são algoritmos de otimização e busca baseados nos mecanismos de seleção natural e genética, Goldberg (1989). As técnicas de busca e otimização geralmente apresentam:

- Um espaço de busca, onde estão todas as possíveis soluções para o problema.
- Uma função de avaliação ou custo que avalia, geralmente através de uma nota, cada possível solução presente no espaço de busca.

Os *AGs* diferem de outros métodos tradicionais de busca e otimização em quatro aspectos principais, Goldberg (1989):

- 1- Trabalham com uma codificação do conjunto de parâmetros e não com os próprios parâmetros.
- 2- Trabalham com uma população de soluções candidatas simultaneamente e não com uma única solução.
- 3- Utilizam informações de custo ou recompensa, e não derivadas de funções.
- 4- Utilizam regras de transição probabilística, e não determinística.

Na terminologia dos *AGs*, uma solução candidata é chamada de indivíduo ou cromossomo. Ao conjunto de indivíduos simultaneamente avaliados é dado o nome de população. A cada indivíduo é associado um grau de adaptação ou aptidão, que mede a capacidade da solução, representada pelo indivíduo para resolver um problema. *AGs* não garantem uma solução ótima, mas geralmente encontram soluções quase ótimas em um tempo aceitável. Esta técnica emprega uma estratégia de busca paralela e estruturada, embora com um forte componente aleatório, que é voltado em direção à busca de pontos de maior aptidão, ou seja, pontos nos quais a função a ser minimizada (ou maximizada) tem valores relativamente baixos (ou altos).

Apesar de aleatório, eles não definem caminhos desordenados em direção a uma solução ótima, uma vez que utilizam informações históricas para definir novos pontos de busca, onde são esperados melhores desempenhos. Isto é realizado através de processos iterativos, em que cada iteração é chamada de geração. Durante cada geração, princípios de seleção e reprodução são aplicados a uma população de candidatos. A população pode variar de tamanho, dependendo da complexibilidade do problema e dos recursos computacionais disponíveis. Espera-se que ao final de um número de gerações o algoritmo venha a convergir para soluções ótimas ou quase ótimas.

AGs são baseados na teoria da evolução natural, proposta por Charles Darwin. Segundo esta teoria, a evolução das populações naturais por várias gerações ocorre de acordo

com os princípios de seleção natural e sobrevivência dos mais aptos. É interessante observar que a teoria da evolução e a computação nasceram praticamente na mesma época.

O primeiro passo é a geração da população inicial. Em seguida deve ser determinado como os indivíduos serão codificados, permitindo sua posterior manipulação pelos operadores genéticos. A este processo dá-se o nome de codificação ou representação. Durante o processo evolutivo, a população é avaliada: para cada indivíduo é dado uma nota, ou índice, refletindo sua habilidade de adaptação a um determinado ambiente. O esquema funcional do Algoritmo Genético é mostrado na Figura 3.7.

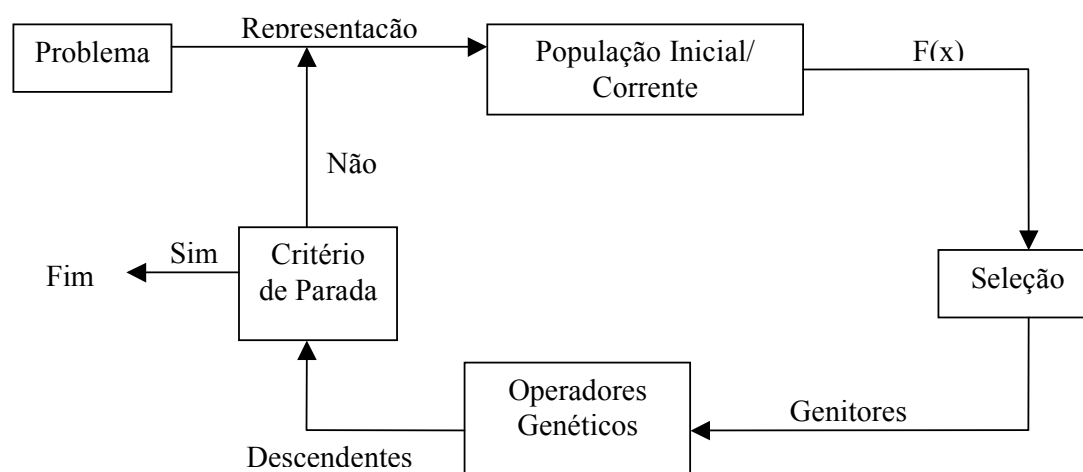


Figura 3.7 - Esquema funcional de um Algoritmo Genético

Uma porcentagem dos mais aptos é mantida, enquanto os outros são descartados (darwinismo). Os membros mantidos pela seleção podem sofrer modificações em suas características fundamentais através de mecanismos como mutações e *crossovers*, gerando descendentes para a próxima geração. Este processo, chamado de reprodução, é repetido até que uma solução satisfatória seja encontrada. Embora possam parecer simplistas do ponto de vista biológico, estes algoritmos são suficientemente complexos para fornecerem mecanismos eficientes e robustos de busca e adaptação, Goldberg (1989).

3.5.1 Representação

O primeiro aspecto a ser considerado na utilização dos AGs como ferramentas para a solução de problemas é a representação destes problemas em uma estrutura, na qual os algoritmos possam trabalhar adequadamente. Dá-se o nome *genótipo* ou *cromossomo* ao indivíduo codificado com esta estrutura, que antes de ser codificado, é chamado *fenótipo*. Para

avaliação de um *cromossomo*, é preciso decodificá-lo para o *fenótipo* correspondente, figura 3.8.

Tradicionalmente os indivíduos são representados por vetores binários, em que cada elemento do vetor denota presença ou ausência de uma determinada característica. Os elementos de um genótipo são combinados para a definição das características reais do indivíduo, ou seja, o seu *fenótipo*. Teoricamente, a representação binária é independente do problema, pois, uma vez encontrada a representação em vetores binários, as operações padrão podem ser utilizadas, facilitando a sua utilização em diferentes tipos de problemas. A utilização de representações em níveis de abstração mais altos também tem sido investigada. Como estas representações estão mais ligadas aos *fenótipos*, elas são mais adequadas para problemas em que a transformação “*fenótipo-genótipo*” é mais complexa. Neste caso, podem ser necessários operadores específicos para utilizar estas representações.

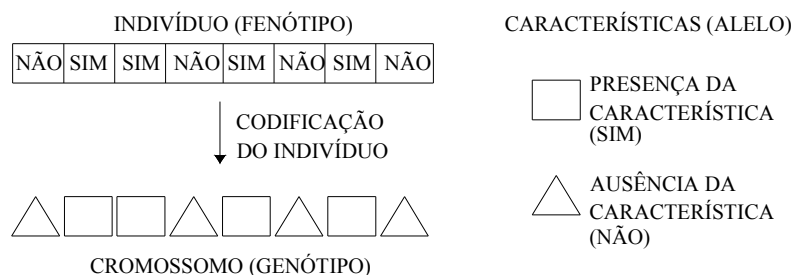


Figura 3.8 - Representação dos Algoritmos Genéticos (*AGs*).

3.5.2 Seleção

O processo de seleção determina quais indivíduos da população podem participar da fase de reprodução. Os indivíduos são selecionados de acordo com uma probabilidade dada pelos seus índices ou notas de aptidão. Assim, os indivíduos com maior adaptação relativa tem maiores chances de participar da reprodução. O princípio básico do funcionamento dos *AGs* é que um critério de seleção faz com que, após varias gerações, o conjunto inicial de indivíduos produza indivíduos cada vez mais aptos, Aurélio *et al* (1999). A maioria dos métodos de seleção é definida para escolher preferencialmente indivíduos com maiores notas de aptidão, embora não exclusivamente (em alguns casos é também desejável manter a diversidade da população).

Um método de seleção muito utilizado é o método da roleta, em que cada indivíduo ocupa, em uma roleta, uma área proporcional ao seu índice de aptidão. Desta forma, aos indivíduos com maior aptidão é associada uma fatia maior da roleta, e vice-versa.

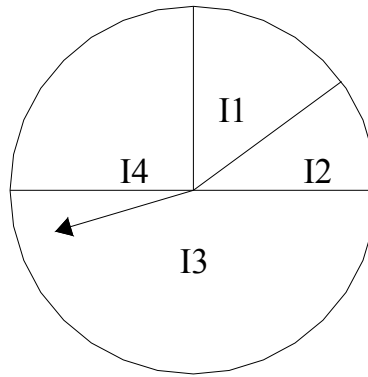


Figura 3.9 - Seleção utilizando o método da roleta.

A Figura 3.9 ilustra a definição das áreas da roleta para um conjunto de indivíduos cujas notas são mostradas na tabela 3.2.

Tabela 3.2 - Notas dos cinco indivíduos de uma população.

Indivíduo	Nota original	Nota normalizada
I1	0.30	0.15
I2	0.20	0.10
I3	1.00	0.50
I4	0.50	0.25

Durante a seleção, a roleta é girada N vezes, selecionando N indivíduos para participarem da fase de reprodução. Indivíduos com maiores notas, por possuírem áreas maiores, possuem maior probabilidade de serem selecionados. A utilização do resultado da função de avaliação para escolher os tamanhos das fatias nem sempre é o método mais adequado. Se a função de avaliação atribuir um valor muito alto a um indivíduo, então este pode monopolizar a seleção, levando a uma convergência prematura do algoritmo genético.

Pode ocorrer também de a função de avaliação atribuir valores de aptidão muito próxima aos indivíduos, tornando sua fatia na roleta aproximadamente igual. Com isto, a

seleção pode não favorecer os indivíduos mais aptos. Uma possível solução para estes problemas é a utilização da técnica de ranking.

Na técnica de ranking, a fatia é definida não pela nota relativa de cada indivíduo, mas pela posição que eles ocupam no ranking de todas as notas. Esta técnica define a fatia de cada indivíduo por meio da escolha de um valor real entre $0,0$ e $1,0$. Cada indivíduo, em ordem decrescente e ranking, ocupa, da área total da roleta que ainda não foi ocupada, uma fatia proporcional a este valor. Por exemplo, se o valor escolhido for $0,6$, o indivíduo de maior ranking ocupará 60% da roleta. O segundo ocupará 60% dos 40% que sobraram, ou seja, 24% , o terceiro ocupará 60% de 16% , que equivale a $9,6\%$.

A função de avaliação define a nota ou índice de aptidão de cada indivíduo da população. A função utilizada depende do problema atacado. Esta função recebe como entrada um genótipo e retorna sua aptidão ou grau de adaptação. Em problemas em que se busca o ponto Máximo de uma função F , por exemplo, a função de avaliação é a própria função F . Para casos mais complexos, a função se torna mais sofisticada. Como exemplo, para o estudo de otimização do projeto de uma ponte, a função de avaliação poderia incluir a carga máxima suportada, o custo da construção e o tempo da construção.

3.5.3 Reprodução

Os indivíduos escolhidos na fase de seleção participam da fase de reprodução, em que podem ser combinados ou modificados, produzindo os indivíduos da próxima geração. Estas combinações e modificações são realizadas por um conjunto de operadores chamados operadores genéticos. Os operadores genéticos são utilizados para transformar uma população através de sucessivas gerações, estendendo a busca até a obtenção de um resultado satisfatório, Aurélio *et al* (1999).

Os principais operadores genéticos são: *crossover* (ou cruzamento) e mutação. Eles são utilizados para assegurar que a nova geração apresente novos indivíduos (soluções), mantendo características desejáveis adquiridas em gerações anteriores. Para prevenir a eliminação dos melhores indivíduos pela manipulação dos operadores genéticos, estes indivíduos podem ser automaticamente colocados na próxima geração através do operador de elitismo.

O *crossover* é o operador responsável pela recombinação de características genéticas dos pais durante a reprodução, permitindo que elas sejam herdadas pelas próximas gerações. Ele é considerado o operador genético predominante, por isso é aplicada com uma probabilidade, chamada taxa de *crossover*, maior que a taxa de mutação. Este operador pode ser utilizado de várias maneiras. As mais frequentes são:

- *Crossover* de um ponto;
- *Crossover* multipontos;
- *Crossover* uniforme.

Para o primeiro caso, *crossover* de um ponto, um ponto é escolhido entre dois genes, onde ocorrerá o cruzamento, ou seja, as informações dos pais serão trocadas. Um dos filhos herda as informações anteriores a este ponto de um dos pais e as informações após este ponto do outro pai. O outro filho herda os genes restantes. Esta operação é ilustrada na Figura 3.10.

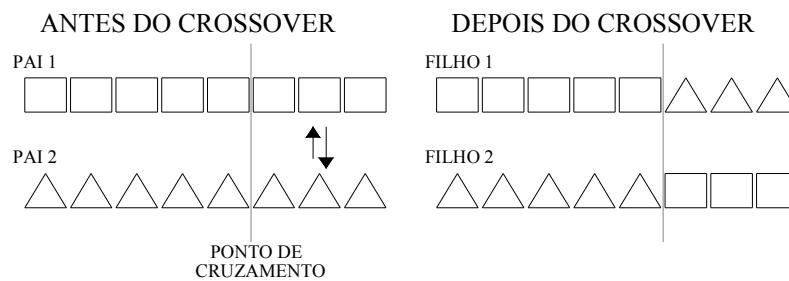


Figura 3.10 - Operação de *Crossover* de um ponto.

O *crossover* multipontos é uma generalização da operação anterior, onde, em vez de um ponto de cruzamento, são utilizados vários pontos.

O outro tipo de *crossover*, o *crossover* uniforme, não utiliza pontos de cruzamento, mas determina por meio de uma máscara, que genes de cada pai serão herdados, por cada um dos filhos. Se o *i*-ésimo elemento da máscara apresentar o valor 1, o valor do *i*-ésimo gene do filho 1 será igual ao valor do *i*-ésimo valor do gene do pai 1. Se o *i*-ésimo elemento da máscara apresentar o valor 0, o valor do *i*-ésimo gene do filho 1 será igual ao valor do *i*-ésimo gene do pai 2. O inverso ocorre na definição do valor dos genes do filho 2. O *crossover* uniforme é apresentado na Figura 3.11.

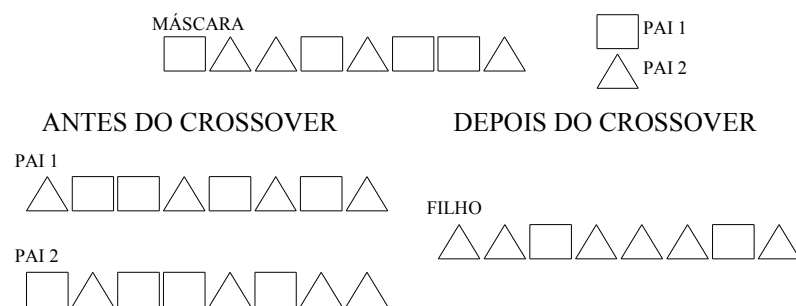


Figura 3.11 – Operação de *Crossover* uniforme.

O operador de mutação é responsável pela introdução e manutenção da diversidade genética da população, alterando arbitrariamente um ou mais genes de um *cromossomo* escolhido aleatoriamente. Este operador fornece meios para introdução de novos elementos na população. A operação de mutação é ilustrada pela Figura 3.12.

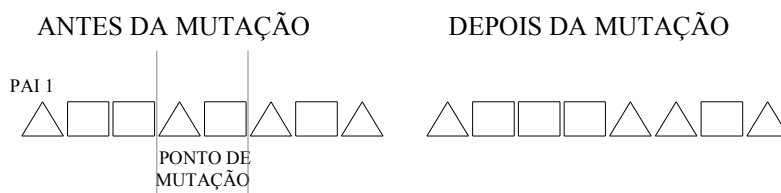


Figura 3.12 – Operação de mutação.

A mutação assegura que a probabilidade de se chegar a qualquer ponto do espaço de busca nunca será zero, além de contornar o problema de mínimos ou máximos locais, por permitir a alteração da direção da busca. O operador de mutação é aplicado aos indivíduos com uma probabilidade dada pela taxa de mutação, que é geralmente pequena. Na maioria das vezes, a mutação é aplicada após a operação de *crossover*. Com base nos conceitos expostos, os seguintes passos devem ser observados pelo Algoritmo Genético.

- 1- Escolher um conjunto de cromossomos iniciais.
- 2- Repetir:
 - 2.1- Definir nota de cada cromossomo.
 - 2.2- Selecionar os cromossomos mais aptos.
 - 2.3- Aplicar operadores de reprodução sobre cromossomos selecionados.
- 3- Até que um cromossomo adequado seja obtido ou sejam realizadas N gerações.

O processo evolucionário termina quando o algoritmo convergir, obtém-se um indivíduo adequado ou for realizado um determinado número de gerações. O algoritmo converge quando pelo menos $X\%$ dos genes possuírem o mesmo valor em pelo menos $Y\%$ dos indivíduos. Os valores de X e Y são definidos pelo usuário de acordo com o problema a ser tratado.

3.5.4 Parâmetros genéticos

Ao utilizar os AGs para solução de um problema, é importante analisar como a escolha dos parâmetros utilizados pode influir no seu comportamento, para que se possa estabelecer

conforme as necessidades do problema e dos recursos disponíveis. Os principais parâmetros e suas influências são os seguintes:

- Tamanho da população: afeta o desempenho global e a eficiência dos AGs. Com uma população pequena, o desempenho pode cair, pois a população pode cobrir apenas uma pequena parte do espaço de busca. Uma grande população geralmente fornece uma cobertura representativa do domínio do problema, além de prevenir contra convergências prematuras para soluções locais, em vez de globais. No entanto, para se trabalhar com grandes populações são necessários maiores recursos computacionais, ou que o algoritmo trabalhe por um longo período.
- Taxa de cruzamento: quanto maior for essa taxa, mais rápida é a introdução de novas estruturas na população. Se esta taxa for muito alta, indivíduos com bons índices de aptidão poderão ser retirados a uma velocidade que supere a capacidade de gerar melhores indivíduos. Se esta taxa for muito baixa, a busca pode estagnar.
- Taxa de mutação: uma baixa taxa de mutação evita que uma dada posição fique estagnada em um valor, além de possibilitar que se chegue em qualquer ponto do espaço de busca. Com uma taxa muito alta, a busca se torna essencialmente aleatória.
- Intervalo de geração: controla a porcentagem da população que será substituída para a próxima geração. Com um valor muito alto, a maior parte da população é substituída, o que pode levar à perda de indivíduos de alta aptidão. Com um valor baixo, o algoritmo pode ser tornar muito lento, pois o número de gerações necessárias pode ser muito grande.

3.6. PESQUISA HARMÔNICA

Pesquisa Harmônica é um algoritmo de otimização e busca baseado em performance musical, por exemplo, um trio de jazz onde o objetivo é a melhor harmonia entre todos os instrumentos, Geem *et al* (2001).

A performance musical busca um estado fixo ótimo, a melhor harmonia, determinada através de uma estimativa estética, melhor som, assim como os algoritmos de otimização e busca buscam também um estado ótimo, custo mínimo e máximo benefício, determinado por uma função objetivo.

Tabela 3.3 – Comparações entre otimização e performance musical.

Fator de comparação	Processo de otimização	Performance musical
Melhor estado	Ótimo global	Melhor harmonia
Estimado por	Função objetivo	Padrão estético
Estimado com	Valor das variáveis	Som dos instrumentos
Numero de	Variáveis	Instrumentos musicais
Processo unitário	Cada iteração	Cada ensaio, prática

Ele faz analogia com a improvisação musical, processo onde músicos num ensaio tocam seus instrumentos experimentando diferentes notas musicais, a fim de obter uma melhor harmonia. O algoritmo *PH* foi aplicado com sucesso para problemas otimização combinatória, como a design de estrutura, Geem & Lee (2004), design rede, Geem (2006), e calibração de parâmetros hidrológicos, Geem *et al* (2001).

A improvisação no jazz busca um melhor estado (melhor harmonia), determinado pela estimativa estética, assim como a otimização de um algoritmo procura um melhor estado (ótimo global) determinado pela avaliação da função objetivo. A estimativa estética é feita pelo conjunto de sons desempenhado por cada instrumento, assim como a avaliação da função objetivo é realizada por um conjunto de valores atribuídos por cada variável de decisão. A melhora na qualidade harmônica é possível pela prática, tal como a melhora na solução é conseguida iteração após iteração.

A analogia entre improvisação e otimização é mostrada na Figura 3.13. Cada instrumento (Saxofonista, baixista e guitarrista) corresponde a cada variável de decisão (x_1 , x_2 , x_3), e a gama de nota de cada instrumento musical (saxofone = (Dó, Re, Mi); baixo = (Mi, Fá, Sol); e violão = (Sol, Lá, Si)) corresponde ao intervalo de cada variável valor ($x_1 = (100, 200, 300)$; $x_2 = (300, 400, 500)$; e $x_3 = (500, 600, 700)$). Se o saxofonista tocar a nota Dó, o baixista Mi, e o guitarrista sol, suas notas juntas fazem uma nova harmonia (Dó, Mi, Sol). Se esta nova harmonia é melhor do que a harmonia existente, a nova harmonia é mantida. Do mesmo modo, o novo vetor solução (100 milímetros, 300 milímetros, 500 milímetros) gerado no processo de otimização é mantido se for melhor do que a harmonia existente, em termos de valor da função objetivo. Assim como a qualidade harmônica é melhorada ensaio após ensaio, a qualidade na solução é melhorada iteração por iteração.

Ele apresenta, assim como os Algoritmos Genéticos:

- Um espaço de busca delimitado pelo usuário, onde estão as soluções possíveis para o problema.
- Uma memória de soluções possíveis que são trabalhadas simultaneamente.
- Uma função que avalia a performance, geralmente através de uma nota, gerada por uma função objetivo.
- Mecanismos de troca de informações entre as soluções candidatas e improvisação para introdução de novas características.

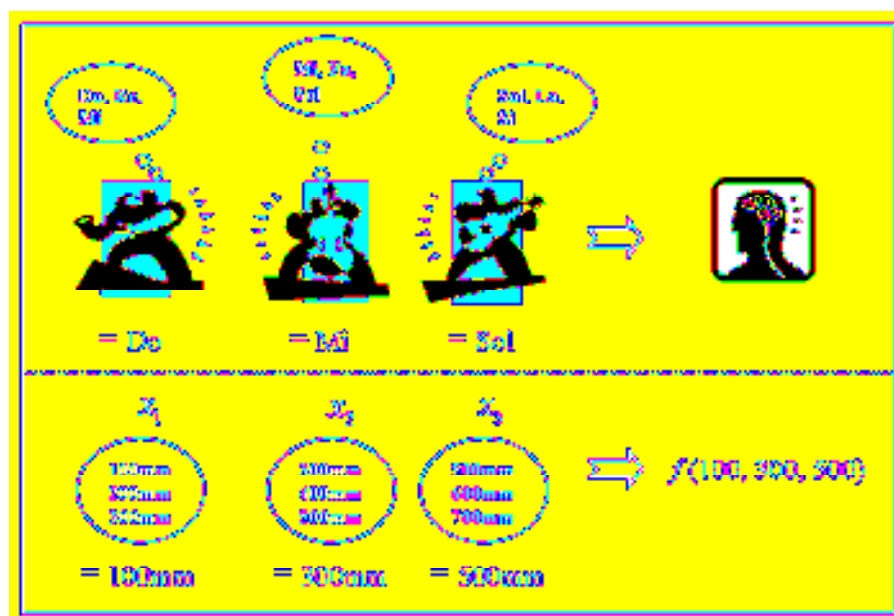


Figura 3.13 – Analogia entre improvisação musical e otimização, Geem & Lee (2004).

Esse método trabalha com os próprios parâmetros e com uma memória de soluções que vão se interagindo após cada iteração produzindo novas harmonias. Devido a sua natureza, apresenta convergência dos valores em um tempo computacional relativamente baixo, com bons resultados.

Ele pode utilizar a própria função a ser otimizada, e não derivadas da mesma, apresentando regras de transição estocástica.

Cada possível solução é denominada de harmonia. Ao conjunto todo de harmonias é denominado memória harmônica sendo associada a cada uma delas uma nota de performance que mede sua capacidade harmônica.

Após cada iteração ou ensaio, novas harmonias produzidas através de improvisação e/ou mesclagem são avaliadas e ranqueadas. As piores harmonias são descartadas e repete-se o ciclo até que se satisfaça o critério de parada.

A representação dos parâmetros pode ser binária ou numérica, dependendo da característica de cada problema.

O processo de seleção determina quais harmonias podem participar da mesclagem. Um método muito utilizado é o do sorteio, onde é gerado um número aleatório dentro da memória harmônica.

A Figura 3.14 representa esquematicamente como é realizado o método do sorteio. Cada harmonia ocupa no sorteio um espaço igualitário.

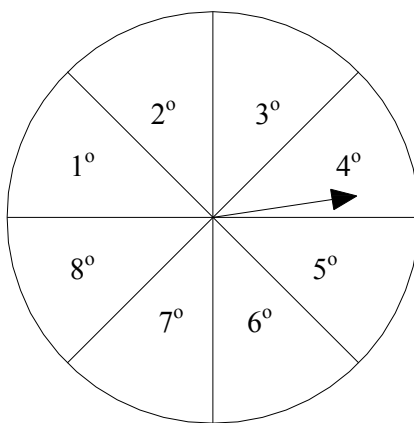


Figura 3.14 – Seleção pelo método do sorteio.

Para cada parâmetro é feito um sorteio para escolha da nova nota musical. Caso o número sorteado esteja fora dos limites estabelecidos, uma nota musical de improvisação é adicionada. Repete-se este procedimento até que todas as notas musicais da harmonia sejam mescladas ou improvisadas.

Definem-se assim os passos do algoritmo:

Passo 1 - inicializar o problema e os parâmetros do algoritmo;

Passo 2 - iniciar uma memória harmônica;

Passo 3 - improvisar uma nova harmonia;

Passo 4 - atualizar a memória harmônica; e

Passo 5 - checar o critério de parada.

No passo 1, o problema de otimização é especificado como segue:

$$\begin{aligned} \min f(x) \\ x \in X_i, i = 1, 2, \dots, N \end{aligned} \tag{3.6}$$

onde $f(x)$ é a função objetivo; x_i^k é cada variável de decisão; N é o número variáveis de decisão (número de instrumentos musicais); X_i é o conjunto dos possíveis valores para cada variável de decisão, Figura 3.15, e k corresponde a posição do vetor.

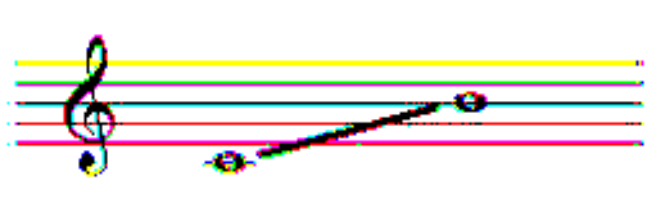


Figura 3.15 – conjunto dos possíveis valores para cada variável de decisão.

Os parâmetros do algoritmo também são especificados nesta etapa. Estes são a memória harmônica (*HMS*), ou o número de vetores solução presentes na memória harmônica; a memória Harmônica considerando a taxa (*HMCR*); a taxa de adaptação (*PAR*) e do número de improvisações (*NI*), e critério de parada.

A memória harmônica (*HM*), é uma memória local onde todas os vetores solução (conjuntos de variáveis de decisão) e os correspondentes valores função objetivo são armazenados, Figura 3.16 e Figura 3.17.



Figura 3.16 – notas armazenadas na memória harmônica.

Os valores da função são utilizados para avaliar a qualidade dos vetores solução. A memória harmônica é parecida com a população do Algoritmo Genético.

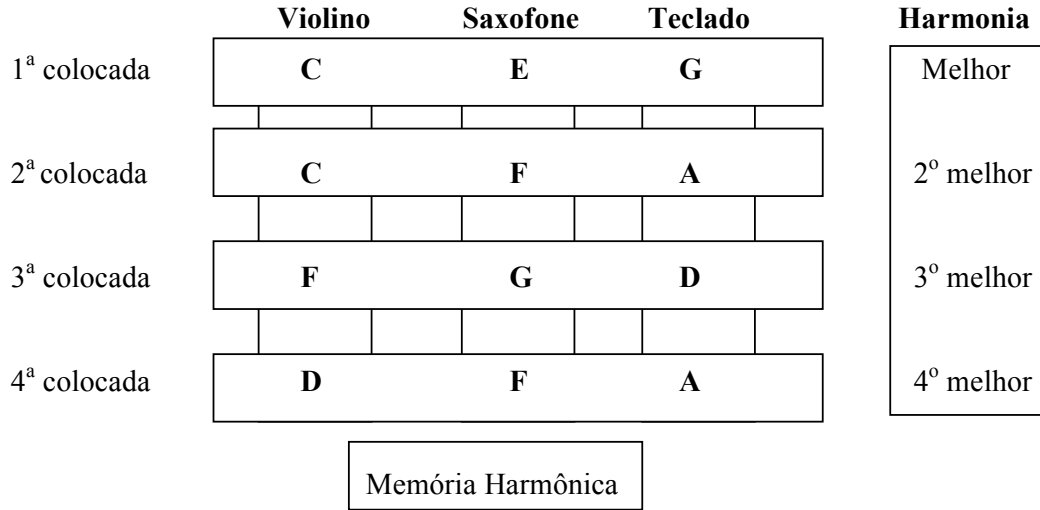


Figura 3.17 – Estrutura da Memória Harmônica.

No passo 2, a matriz HM é preenchida com o maior número vetores gerados aleatoriamente como solução.

$$HM = \begin{bmatrix} x_1^1 & x_2^1 & \cdot & x_{N-1}^1 & x_N^1 \\ x_1^2 & x_2^2 & \cdot & x_{N-1}^2 & x_N^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_1^{HMS-1} & x_2^{HMS-1} & \cdot & x_{N-1}^{HMS-1} & x_N^{HMS-1} \\ x_1^{HMS} & x_2^{HMS} & \cdot & x_{N-1}^{HMS} & x_N^{HMS} \end{bmatrix} \begin{matrix} \Rightarrow f(x^1) \\ \Rightarrow f(x^2) \\ \dots \\ \Rightarrow f(x^{HMS-1}) \\ \Rightarrow f(x^{HMS}) \end{matrix} \quad (3.7)$$

No próximo passo, passo 3, um novo vetor harmonia, $(x_1, x_2, \dots, x_N)$, é gerado com base em duas regras: (1) Considerando a memória harmônica existente e (2) improvisando um novo valor. A Geração de uma nova harmonia é chamada de 'improvisação'.

É feito um sorteio para cada variável de decisão na nova harmonia. Caso o valor sorteado seja maior que PAR , um novo valor é gerado. Caso seja menor, um valor é escolhido a partir dos valores históricos armazenados na HM .

Se o novo vetor, $(x_1, x_2, \dots, x_N)$, for melhor do que o pior vetor na HM , avaliado em termos do valor da função objetivo, a nova harmonia será incluída no HM e a pior é excluída da HM , passo 4.

Caso o critério de parada não seja satisfeito, repete-se os passos 3 e 4. Caso contrario, encerra-se o programa.

Influência dos parâmetros harmônicos: é de suma importância analisar como a escolha dos parâmetros utilizada pode influir no comportamento do algoritmo. Os principais parâmetros são os seguintes:

- Memória Harmônica: dependendo do seu tamanho, o desempenho pode cair. Se for muito pequena, pode ocorrer convergência prematura, pois pouca parte do espaço de busca é avaliada. Caso seja muito grande, pode aumentar em muito o tempo computacional para cálculo dos parâmetros.
- Taxa Harmônica: quanto maior esta taxa, mais harmonias são introduzidas na Memória Harmônica. Usualmente utiliza-se Taxa Harmônica de um.
- Taxa de Improvisação: possibilita a introdução de novas notas musicais na harmonia.

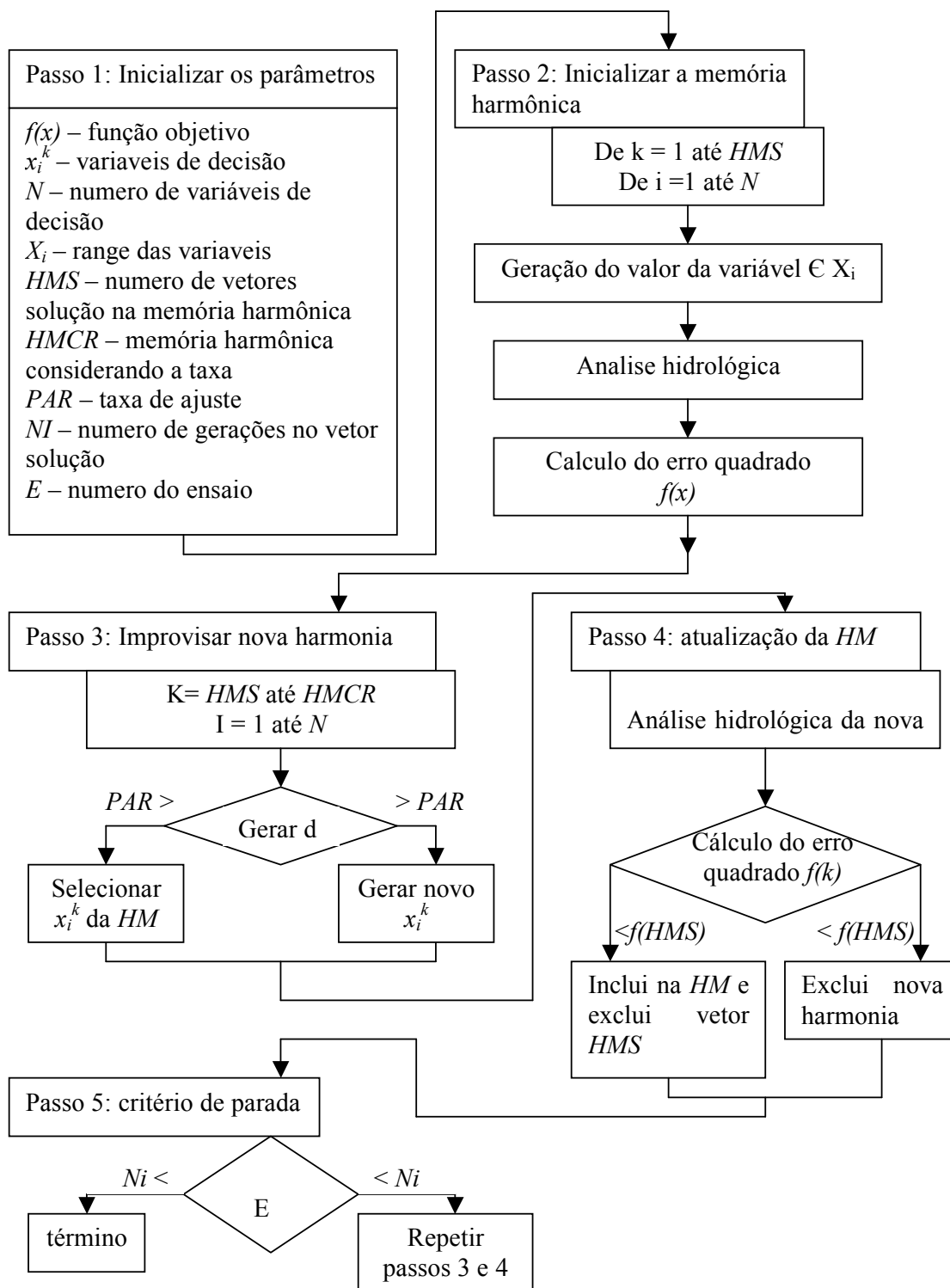


Figura 3.18 - Esquema funcional da Pesquisa Harmônica.

4. METODOLOGIA

4.1.ÁREA DE ESTUDO

A área estudada localiza-se na bacia do Rio Aquidauana, sub bacia do Rio Miranda, no estado de Mato Grosso do Sul. A Figura 4.1 esquematiza a localização da bacia e os postos fluviométricos à montante de Aquidauana. A Figura 4.2 ilustra a localização dos postos pluviométricos na área à montante do posto fluviométrico Ponte do Grego, que serão utilizados no trabalho.

Pode-se dividir a bacia em duas partes diferentes: a norte, drenada pelo curso superior do próprio Rio Aquidauana, revestida, sobretudo, de areias quartzosas com alta susceptibilidade à erosão e a sul drenada pelos afluentes da margem esquerda: Cachoeirão, Dois Irmãos e Taquaruçu, onde predominam solos com textura silto-argilosa. Grande parte da área está ocupada com pastagens para bovinos e também áreas de agricultura potencialmente irrigáveis. Predominam as atividades industriais de laticínios, frigoríficos e, ainda, as águas são utilizadas para o abastecimento público da cidade de Aquidauana, à jusante da bacia.

O Rio Aquidauana, tem sua nascente na Serra de Maracajú com cerca de 1.200 km de extensão, juntando-se ao Rio Miranda 100 Km de sua foz no Rio Paraguai. Sua navegação atende às fazendas com pequenas embarcações nas épocas das cheias, quando as estradas do Pantanal ficam intransitáveis.

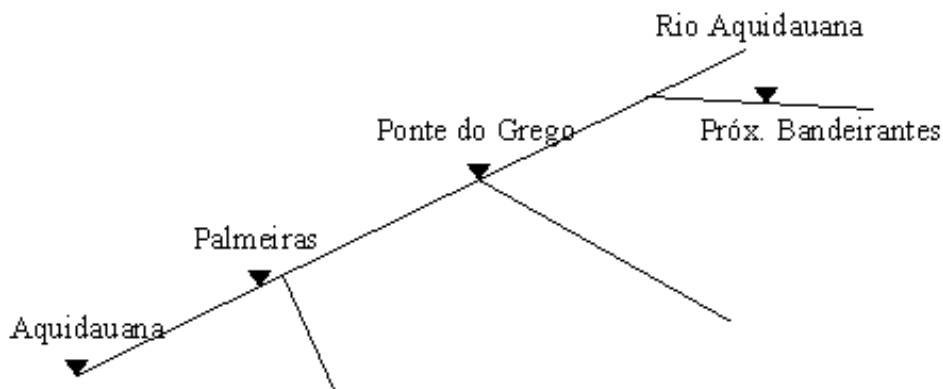


Figura 4.1 – Esquema de localização dos Postos Fluviométricos.

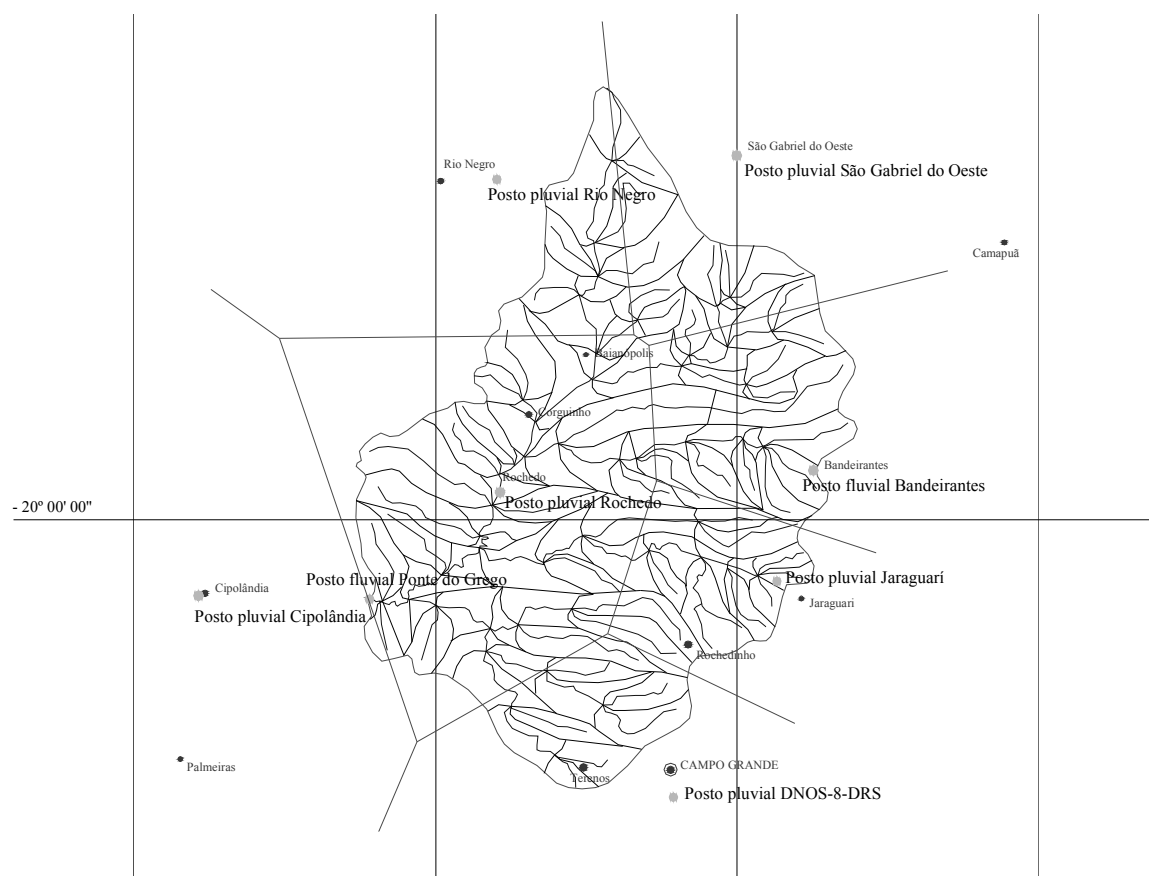


Figura 4.2 - Localização dos Postos Pluviométricos utilizados, Neto (2004).

4.2. PREPARAÇÃO DOS DADOS

Uma análise prévia dos dados disponíveis de vazão e precipitação foi realizada a fim de se eliminar os períodos onde existem dados de vazão, mas faltam os dados de precipitação. Os postos pluviométricos foram selecionados por apresentarem área de influência (calculadas pelo método de Thiessen) na bacia, Neto (2004).

Sete postos foram estudados, tabela 4.1:

Tabela 4.1 – Postos pluviométricos com área de influência, Neto (2004).

Postos pluviométricos	Área de influência (km ²)	Porcentagem em relação a área da bacia (%)
Jaraguari	780,85	12,26
DNOS	641,37	10,07
Bandeirantes	1.109,49	17,42
Rio Negro	549,02	8,62
Rochedo	2.635,53	41,38
São Gabriel do Oeste	629,90	9,89
Cipolândia	22,93	0,36

4.3.MODELO UTILIZADO

Considerando a equação da continuidade.

$$\frac{ds}{dt} = E - Q \quad (4.1)$$

onde E é a vazão de entrada e Q a vazão de saída.

Como a variação temporal do armazenamento de água na bacia é igual à diferença entre a vazão de entrada e a vazão de saída num intervalo de tempo. Quando integrada entre t e $t-l$:

$$\int_{t-l}^t ds = \int_{t-l}^t E dt - \int_{t-l}^t Q dt \quad (4.2)$$

se $l = 2$

$$\int_{t-2}^t ds = \int_{t-2}^t E dt - \int_{t-2}^t Q dt \quad (4.3)$$

$$s_t - s_{t-2} = (a_0 E_{t-2} + a_1 E_{t-1} + a_2 E_t) \Delta t - (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \Delta t \quad (4.4)$$

$$\Delta s = (a_0 E_{t-2} + a_1 E_{t-1} + a_2 E_t) \Delta t - (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \Delta t \quad (4.5)$$

Considerando que E em cada tempo seja

$$E = \frac{hA_b}{\Delta t} \quad (4.6)$$

e que

$$h = \sum_{i=1}^{np} \alpha_i P_i \frac{A_i}{A_b} \quad (4.7)$$

onde a altura $h_i = \alpha_i P_i$ e uma porcentagem α_i da precipitação P_i registrada no posto i num determinado tempo, A_i e a área de influência do posto i e A_b a área da bacia e np e o numero de postos.

$$\Delta s = \Delta t \left(a_0 \frac{h_{t-2} A_b}{\Delta t} + a_1 \frac{h_{t-1} A_b}{\Delta t} + a_2 \frac{h_t A_b}{\Delta t} \right) - (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \Delta t \quad (4.8)$$

$$\Delta s = a_0 \left(\sum_{i=1}^{np} \alpha_i P_{i,t-2} \frac{A_i}{A_b} \right) A_b + a_1 \left(\sum_{i=1}^{np} \alpha_i P_{i,t-1} \frac{A_i}{A_b} \right) A_b + a_2 \left(\sum_{i=1}^{np} \alpha_i P_{i,t} \frac{A_i}{A_b} \right) A_b - \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.9)$$

$$\Delta s = \sum_{i=1}^{np} a_0 \alpha_i P_{i,t-2} A_i + \sum_{i=1}^{np} a_1 \alpha_i P_{i,t-1} A_i + \sum_{i=1}^{np} a_2 \alpha_i P_{i,t} A_i - \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.10)$$

$$\Delta s = \sum_{i=1}^{np} \alpha_i A_i (a_0 P_{i,t-2} + a_1 P_{i,t-1} + a_2 P_{i,t}) - \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.11)$$

Considerando que a_0 , a_1 e a_2 variam de posto a posto com

$$a_0 + a_1 + a_2 = 1 \quad (4.12)$$

e que

$$b_0 + b_1 + b_2 = 1 \quad (4.13)$$

$$\Delta s = \alpha_1 A_1 (a_{01} P_{1,t-2} + a_{11} P_{1,t-1} + a_{21} P_{1,t}) + \alpha_2 A_2 (a_{02} P_{2,t-2} + a_{12} P_{2,t-1} + a_{22} P_{2,t}) + \dots + \alpha_7 A_7 (a_{17} P_{7,t-2} + a_{17} P_{7,t-1} + a_{27} P_{7,t}) - \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.14)$$

considerando que uma porcentagem β do volume armazenado escoe acima da tendência de escoamento sem a ocorrência de chuvas.

$$\beta \Delta s = \Delta V \quad (4.15)$$

$$\frac{Q_{t-1} - Q_{t-2}}{\Delta t} = \frac{Q' - Q_{t-1}}{\Delta t} \quad (4.16)$$

$$Q' = 2Q_{t-1} - Q_{t-2} \quad (4.17)$$

$$\Delta V = \frac{(Q_t - Q') \Delta t}{2} \quad (4.18)$$

$$\Delta V = (Q_t - 2Q_{t-1} + Q_{t-2}) \frac{\Delta t}{2} \quad (4.19)$$

$$\Delta S = \frac{\Delta V}{\beta} = \frac{(Q_t - 2Q_{t-1} + Q_{t-2}) \Delta t}{2\beta} \quad (4.20)$$

então:

$$(Q_t - 2Q_{t-1} + Q_{t-2}) \frac{\Delta t}{2\beta} = \sum_{i=1}^{np} \alpha_i A_i (a_{0i} P_{i,t-2} + a_{1i} P_{i,t-1} + a_{2i} P_{i,t}) - \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.21)$$

isolando Q_t

$$Q_t - 2Q_{t-1} + Q_{t-2} = \frac{2\beta}{\Delta t} \sum_{i=1}^{np} \alpha_i A_i (a_{0i} P_{i,t-2} + a_{1i} P_{i,t-1} + a_{2i} P_{i,t}) - \frac{2\beta}{\Delta t} \Delta t (b_0 Q_{t-2} + b_1 Q_{t-1} + b_2 Q_t) \quad (4.22)$$

$$Q_t (1 + 2\beta b_2) = \frac{2\beta}{\Delta t} \sum_{i=1}^{np} \alpha_i A_i (a_{0i} P_{i,t-2} + a_{1i} P_{i,t-1} + a_{2i} P_{i,t}) + (2 - 2\beta b_1) Q_{t-1} - (1 + 2\beta b_0) Q_{t-2} \quad (4.23)$$

$$Q_t = \frac{2\beta}{\Delta t (1 + 2\beta b_2)} \sum_{i=1}^{np} \alpha_i A_i (a_{0i} P_{i,t-2} + a_{1i} P_{i,t-1} + a_{2i} P_{i,t}) + \frac{(2 - 2\beta b_1)}{(1 + 2\beta b_2)} Q_{t-1} - \frac{(1 + 2\beta b_0)}{(1 + 2\beta b_2)} Q_{t-2} \quad (4.24)$$

com:

$$0 \leq \beta \leq 1$$

$$0 \leq \alpha_i \leq 1$$

com a área em km^2 , Δt em dia e P em mm , tem-se:

$$Q_t = \frac{2\beta}{86,4(1 + 2\beta b_2)} \sum_{i=1}^{np} \alpha_i A_i (a_{0i} P_{i,t-2} + a_{1i} P_{i,t-1} + a_{2i} P_{i,t}) + \frac{(2 - 2\beta b_1)}{(1 + 2\beta b_2)} Q_{t-1} - \frac{(1 + 2\beta b_0)}{(1 + 2\beta b_2)} Q_{t-2} \quad (4.25)$$

4.4. ESTIMATIVA DOS PARÂMETROS DO MODELO

Métodos de ajuste matemático de series numéricas estão presentes em grande numero de referências. Para o problema particular, foram utilizados os métodos dos Algoritmos Genéticos (*AG*) e o da Pesquisa Harmônica.

Os métodos citados estimam quais parâmetros, satisfazendo determinada função, minimize a soma quadrada dos erros residuais entre os valores observados e calculados na saída.

No modelo proposto, a função de avaliação é o erro residual $e(t)$ é dado por:

$$e(t) = Q(t) - \hat{Q}(t) \quad (4.26)$$

onde: $Q(t)$ é a vazão de saída observada no tempo t e $\hat{Q}(t)$ o valor correspondente calculado pelo modelo nas vazões de saída. A equação (4.26) com $Q(t)$ obtido da equação (4.26) pode ser escrito na forma de matriz:

$$e = [e(1), e(2), \dots, e(m)]^T \quad (4.27)$$

$$Q = [Q(1), Q(2), \dots, Q(m)]^T \quad (4.28)$$

$$\hat{Q} = [\hat{Q}(1), \hat{Q}(2), \dots, \hat{Q}(m)]^T \quad (4.29)$$

onde m é o número de observações e T designa a matriz transposta.

No caso, a função de avaliação utilizada é do erro quadrado, dado por:

$$\min e(i) = \sum_{i=0}^m \left[Q(i) - \hat{Q}(i) \right]^2 \quad (4.30)$$

Foi testado um modelo simples e um modelo com uma condição ao erro, onde caso a vazão seja maior que a média do evento, o erro será dado por:

$$e(t) = \left[Q(t) - \hat{Q}(t) \right]^2 \times \left[\frac{Q(t)}{\text{média}} \right] \quad (4.31)$$

Os algoritmos para cálculo foram implementados em VBA Excell, através de macros de funções utilizando os dados digitados na própria planilha. O computador utilizado foi um Notebook, com 60GB de HD, 512 MB de memória e processador Intel Celeron.

Restrições dos parâmetros:

$$a_0 + a_1 + a_2 = 1$$

$$b_0 + b_1 + b_2 = 1$$

$$0 \leq \beta \leq 1$$

$$0 \leq \alpha_i \leq 1$$

com a_0, a_1, a_2, b_0, b_1 e $b_2 \geq 0$.

5. RESULTADOS E DISCUSSÕES

Três eventos de cheia foram escolhidos para cálculo dos parâmetros dos modelos e outro para teste.

O primeiro evento para cálculo foi de 02 de janeiro de 1998 a 31 de janeiro de 1998, tabela 5.1. Para previsão, o evento selecionado foi de 16 de dezembro de 1996 a 07 de março de 1997.

Os métodos utilizados para se estimar os parâmetros foram o Algoritmo Genético (AGs) e a Pesquisa Harmônica.

Para o Algoritmo Genético uma população de cem indivíduos foi utilizada para o cálculo, com uma taxa de mutação e *crossover* de 6% e 90%, respectivamente. O intervalo de variação dos parâmetros foi de zero a um com variação de 0,01.

Taxas maiores de mutação e *crossover* foram testadas, mas não apresentaram grandes melhorias nos resultados e aumentaram o tempo de convergência. Taxa menores também foram utilizadas, mas apresentaram convergência prematura e piores resultados.

Foram testadas também populações com maior número de indivíduos, mas não houve melhoria nos resultados e aumentaram a memória utilizada e o tempo de implementação. Populações menores apresentaram resultados piores.

Para a Pesquisa Harmônica foi utilizada uma memória harmônica de cem harmonias e taxa de improviso de 5%. O intervalo de variação foi de zero a um com variação de 0,01. taxas escolhidas devido ao mesmo exposto anteriormente.

Foram obtidos os tempos de implementação nos gráficos 5.1 e 5.4. Os gráficos das vazões observadas e calculadas no evento, tanto para implementação como para previsão são mostrados nos gráficos 5.2, 5.3, 5.5 e 5.6. Valores com peso são aqueles que possuem a condição de erro nas maiores vazões.

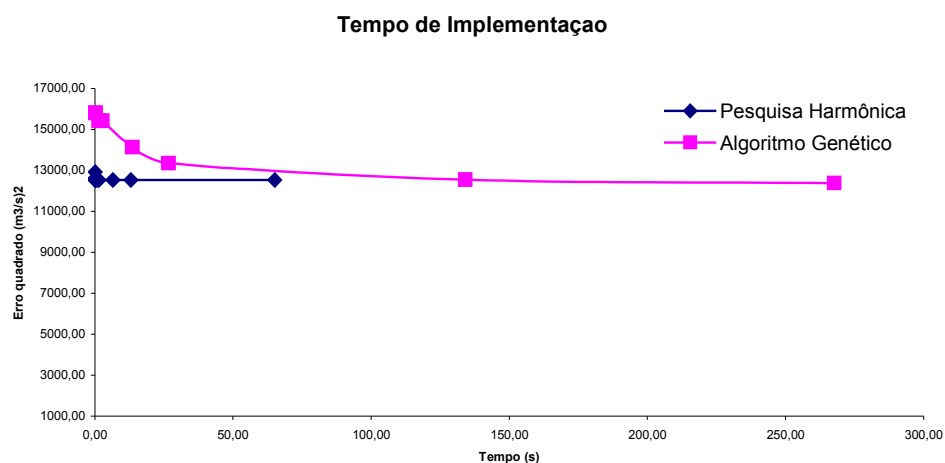


Gráfico 5.1 – tempo de implementação versus erro quadrado, com peso.

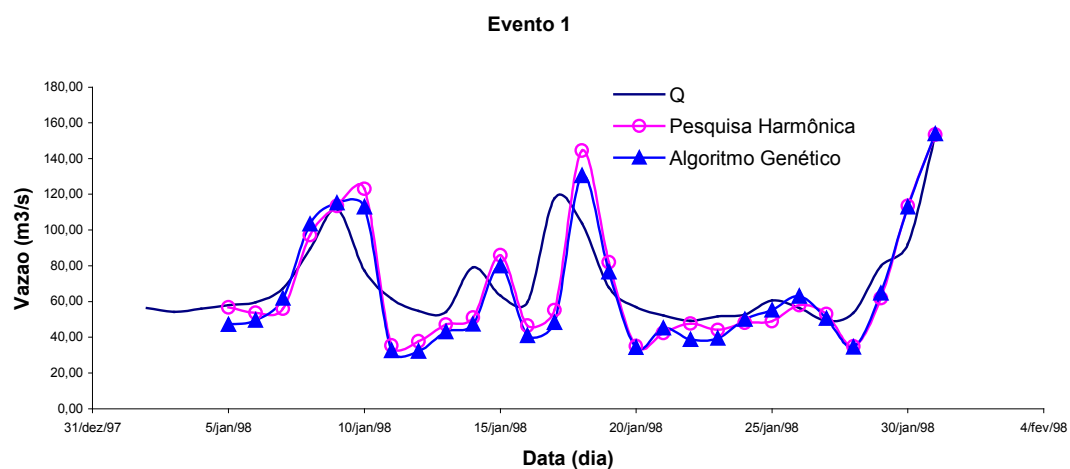


Gráfico 5.2 – Vazão observada e calculada, implementação do evento 1, com peso.

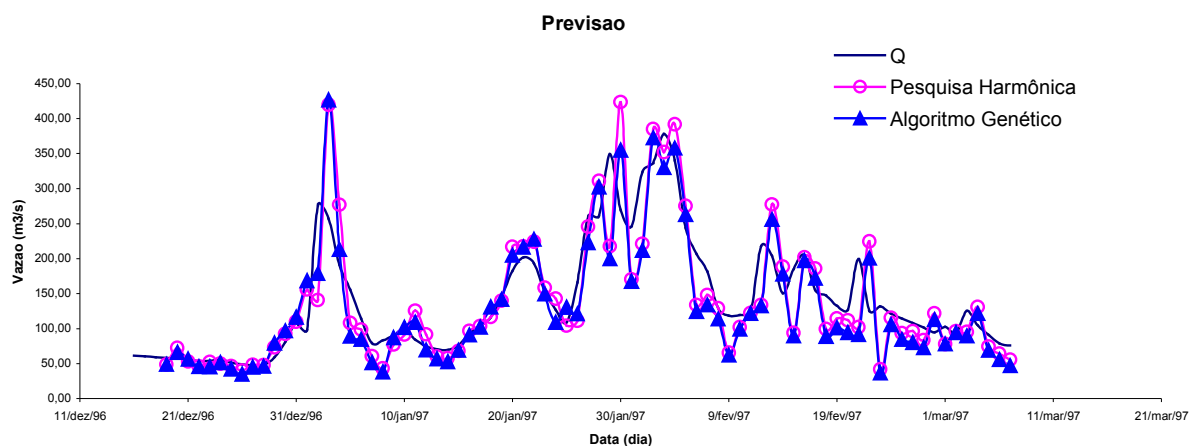


Gráfico 5.3 – Vazão observada e calculada, previsão evento 1, com peso.

Nota-se nesse modelo que os resultados nos picos de vazão se aproximam mais, apresentando certa defasagem, enquanto que para valores menores de vazão existe maior diferença nos valores.

Os valores de desvio padrão foram 21,54 e 50,85 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 21,41 e 48,77 no Algoritmo Genético, na implementação e previsão respectivamente.

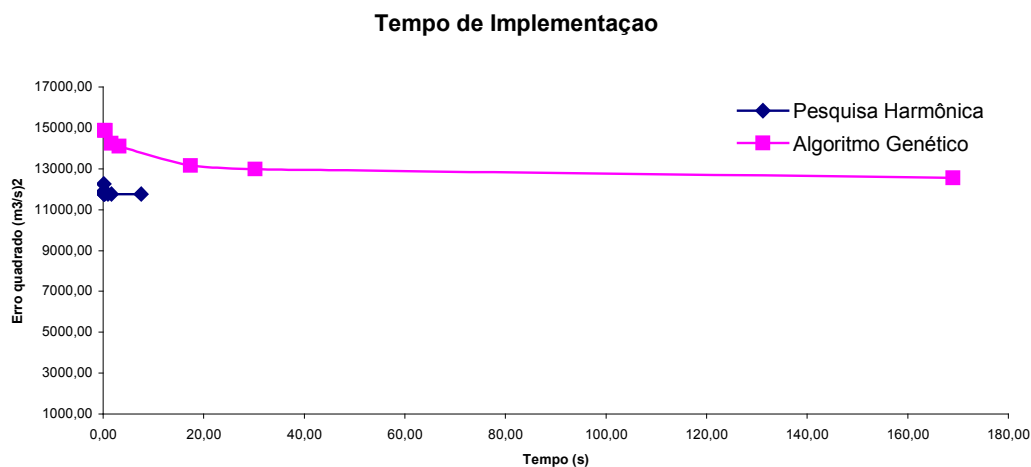


Gráfico 5.4 – tempo de implementação versus erro quadrado, sem peso.

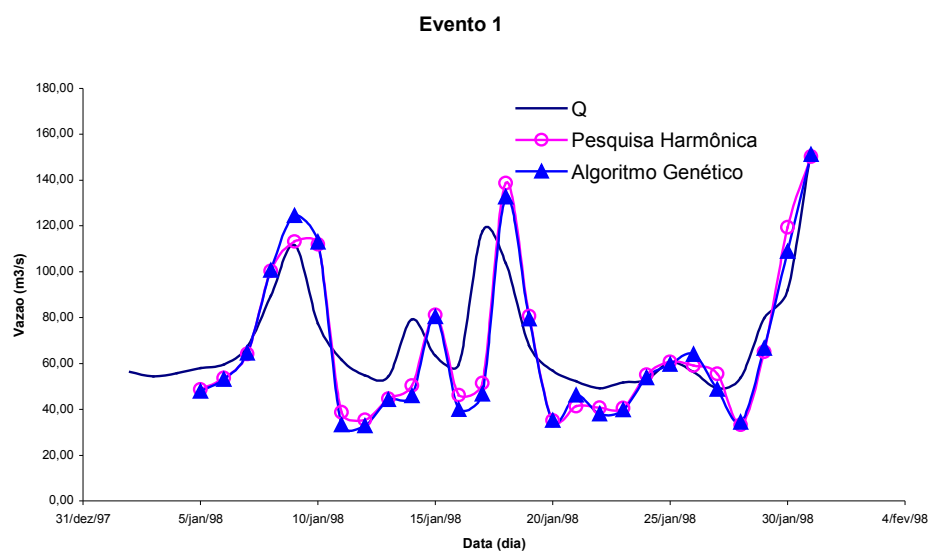


Gráfico 5.5 – Vazão observada e calculada, implementação do evento 1, sem peso.

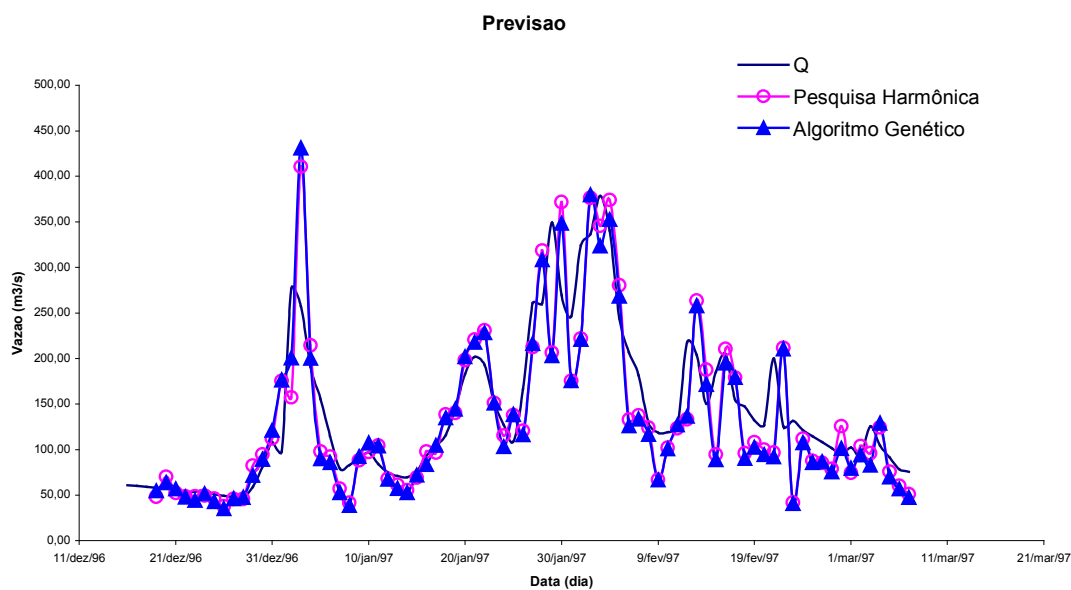


Gráfico 5.6 – Vazão observada e calculada, previsão evento 1, sem peso.

Nota-se nesse modelo que os resultados nos picos de vazão se distanciam, enquanto que para valores menores de vazão existe menor diferença nos valores.

Os valores de desvio padrão foram 20,87 e 48,09 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 21,57 e 48,06 no Algoritmo Genético, na implementação e previsão respectivamente.

O segundo evento para cálculo foi de 01 de novembro de 1994 a 30 de novembro de 1994, tabela 5.3. Para previsão, o evento selecionado foi de 16 de dezembro de 1996 a 07 de março de 1997, o mesmo utilizado para o evento 1.

Os métodos utilizados para se estimar os parâmetros foram o Algoritmo Genético (AGs) e a Pesquisa Harmônica.

Foram utilizados os mesmos parâmetros para os dois algoritmos (população, taxa de mutação, *crossover*, memória harmônica e taxa de improviso). O intervalo de variação dos parâmetros foi de zero a um com variação de 0,01.

Os tempos de implementação são mostrados nos gráficos 5.7 e 5.10. Os gráficos das vazões observadas e calculadas no evento, tanto para implementação como para previsão são mostrados nos gráficos 5.8, 5.9, 5.11 e 5.12.

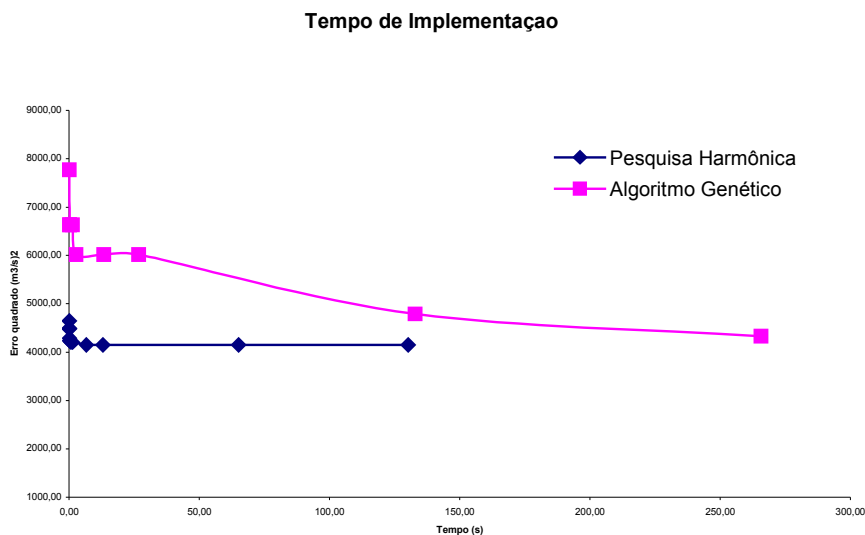


Gráfico 5.7 – tempo de implementação versus erro quadrado, com peso.

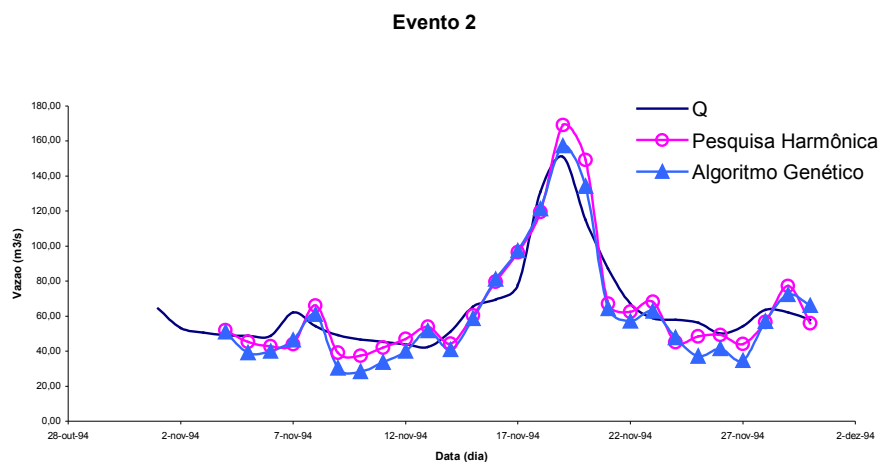


Gráfico 5.8 – Vazão observada e calculada, implementação do evento 2, com peso.

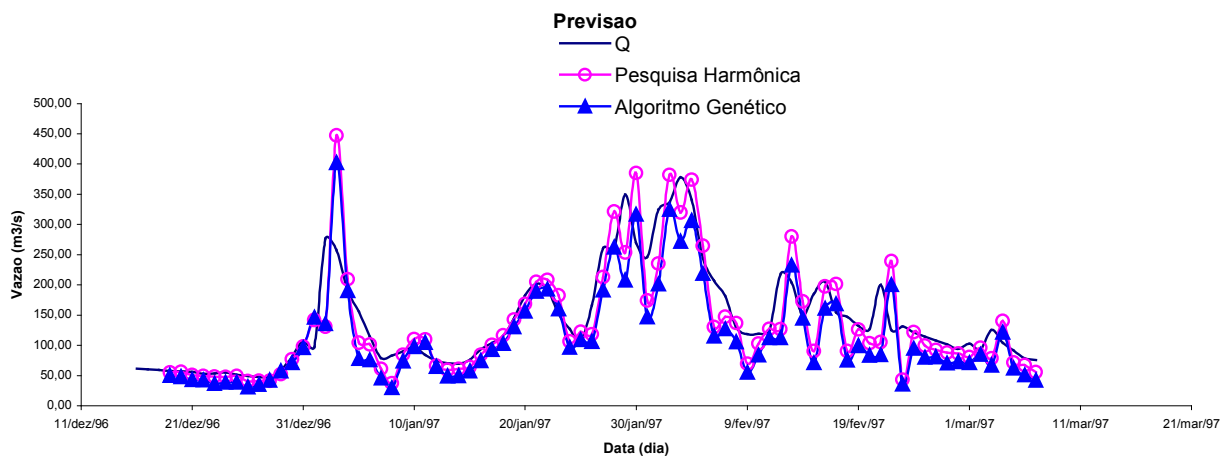


Gráfico 5.9 – Vazão observada e calculada, previsão evento 2, com peso

Nota-se nesse modelo que os resultados nos picos de vazão se aproximam mais, apresentando certa defasagem, enquanto que para valores menores de vazão existe maior diferença nos valores.

Os valores de desvio padrão foram 12,39 e 49,56 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 12,67 e 52,94 no Algoritmo Genético, na implementação e previsão respectivamente.

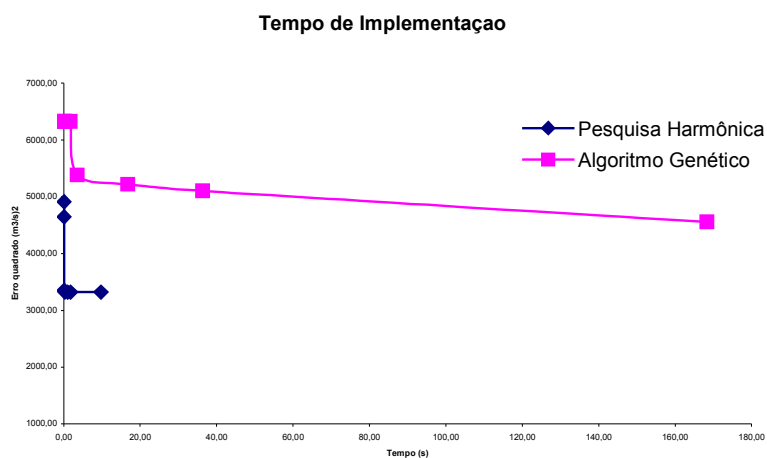


Gráfico 5.10 – tempo de implementação versus erro quadrado, sem peso.

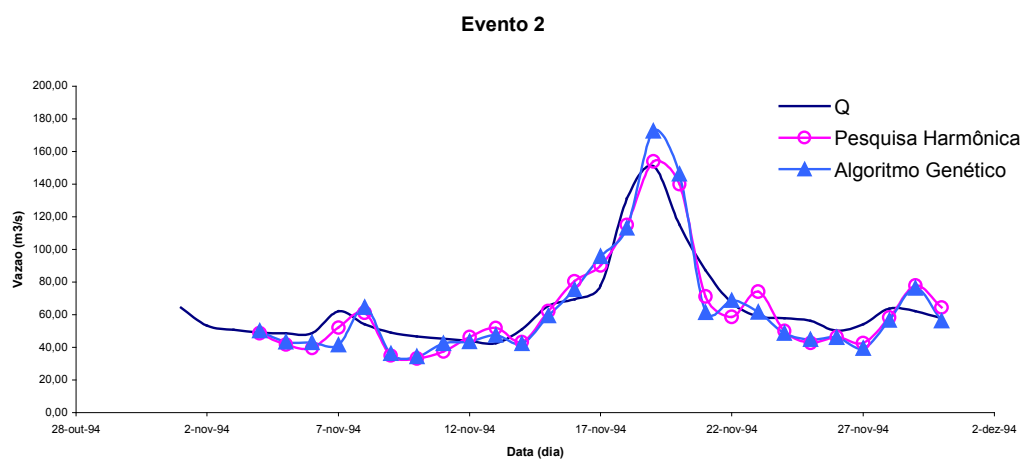


Gráfico 5.11 – Vazão observada e calculada, implementação do evento 2, sem peso.

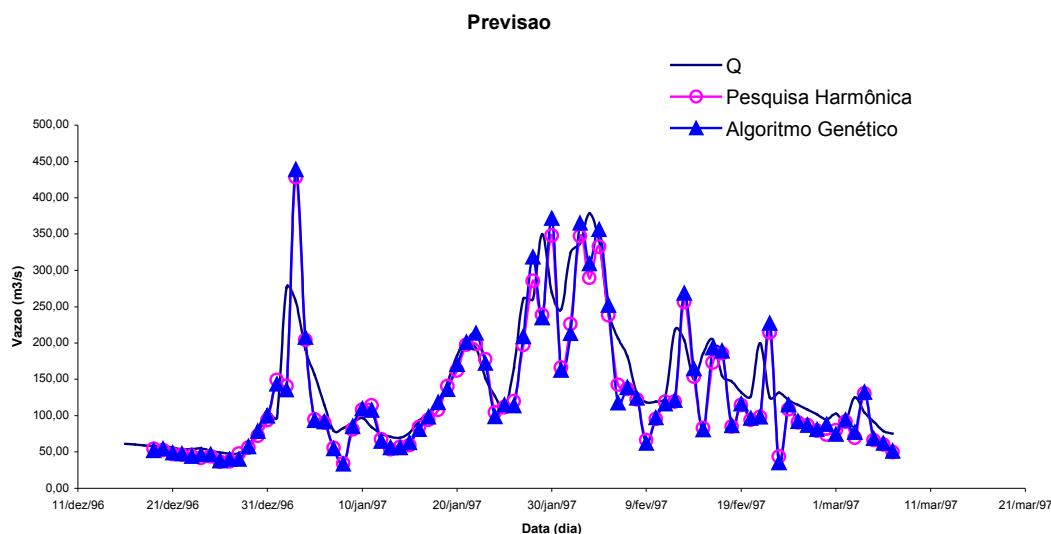


Gráfico 5.12 – Vazão observada e calculada, previsão evento 2, sem peso.

Nota-se nesse modelo que os resultados nos picos de vazão se distanciam, enquanto que para valores menores de vazão existe menor diferença nos valores.

Os valores de desvio padrão foram 11,09 e 48,46 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 12,99 e 51,08 no Algoritmo Genético, na implementação e previsão respectivamente.

O terceiro evento para cálculo foi de 26 de agosto de 1996 a 24 de setembro de 1996, tabela 5.5. Para previsão, o evento selecionado foi de 16 de dezembro de 1996 a 07 de março de 1997, o mesmo utilizado para o evento 1 e 2.

Os métodos utilizados para se estimar os parâmetros foram o Algoritmo Genético (AGs) e a Pesquisa Harmônica.

Para o Algoritmo Genético uma população de cem indivíduos foi-se utilizada para o cálculo, com uma taxa de mutação e *crossover* de 6% e 90%, respectivamente. O intervalo de variação dos parâmetros foi de zero a um com variação de 0,01.

Para a Pesquisa Harmônica foi utilizada uma memória harmônica de cem harmonias e taxa de improviso de 5%. O intervalo de variação foi de zero a um com variação de 0,01.

Os tempos de implementação estão nos gráficos 5.13 e 5.16. Os gráficos das vazões observadas e calculadas no evento, tanto para implementação como para previsão são mostrados nos gráficos 5.14, 5.15, 5.17 e 5.18.

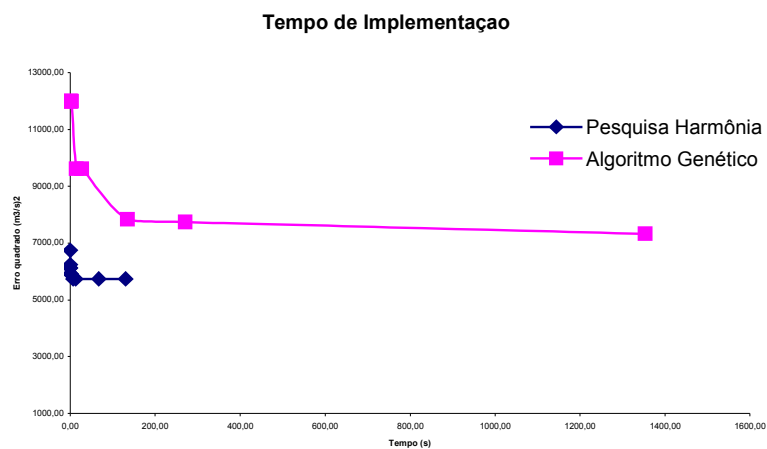


Gráfico 5.13 – tempo de implementação versus erro quadrado, com peso.

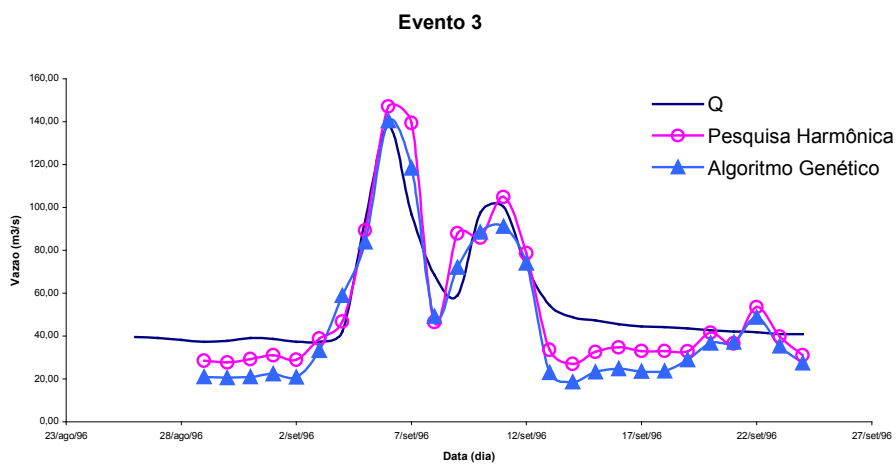


Gráfico 5.14 – Vazão observada e calculada, implementação do evento 3, com peso.

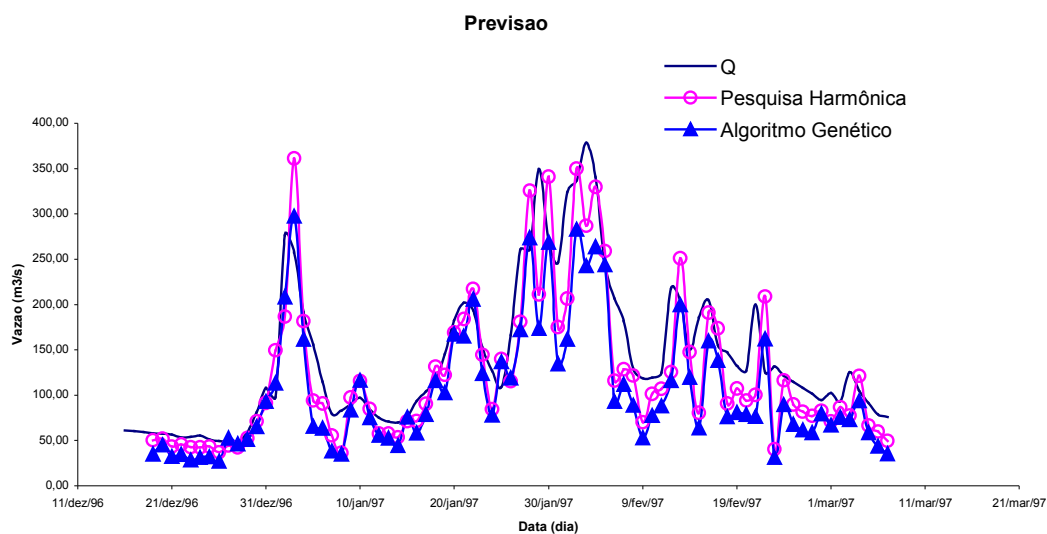


Gráfico 5.15 – Vazão observada e calculada, previsão evento 3, com peso.

Nota-se nesse modelo que os resultados nos picos de vazão se aproximam mais, apresentando certa defasagem, enquanto que para valores menores de vazão existe maior diferença nos valores.

Os valores de desvio padrão foram 14,58 e 47,16 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 16,47 e 56,58 no Algoritmo Genético, na implementação e previsão respectivamente.

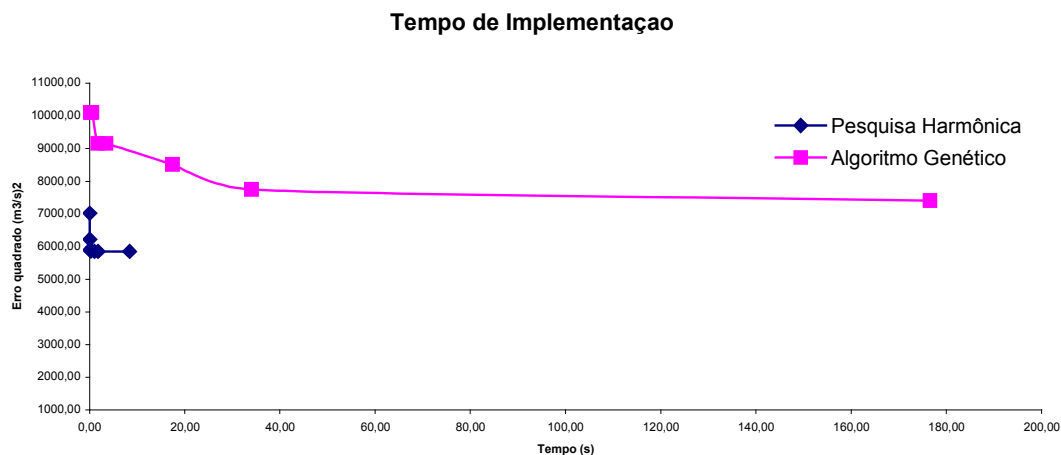


Gráfico 5.16 – tempo de implementação versus erro quadrado, sem peso.

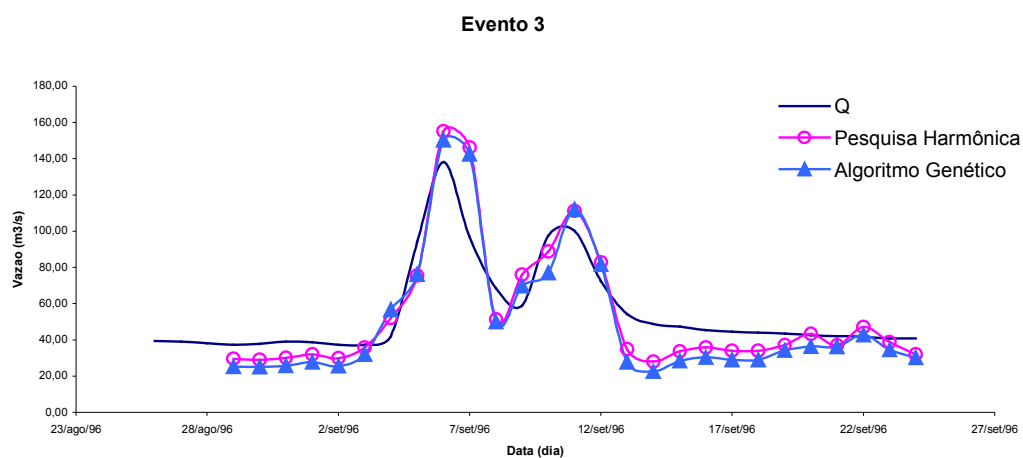


Gráfico 5.17 – Vazão observada e calculada, implementação do evento 3, sem peso.

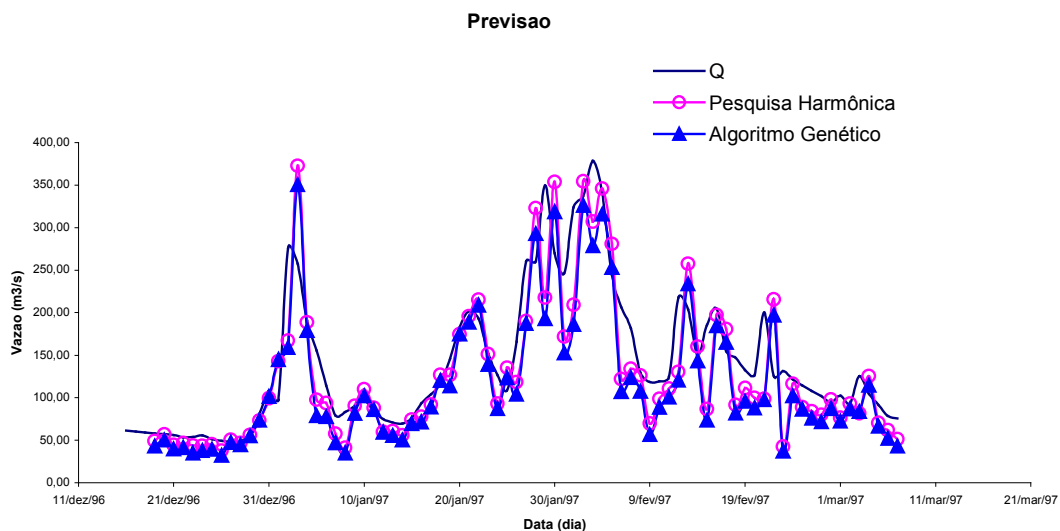


Gráfico 5.18 – Vazão observada e calculada, previsão evento 3, sem peso.

Nota-se nesse modelo que os resultados nos picos de vazão se distanciam, enquanto que para valores menores de vazão existe menor diferença nos valores.

Os valores de desvio padrão foram 14,73 e 48,35 na Pesquisa Harmônica, na implementação e previsão, respectivamente e de 16,35 e 50,41 no Algoritmo Genético, na implementação e previsão respectivamente.

Na tabela 5.1 são mostrados todos os valores de desvio para os três eventos.

Tabela 5.1 – Valores de desvio padrão.

Evento	Pesquisa Harmônica		Algoritmo Genético	
	Com peso	Sem peso	Com peso	Sem peso
Evento 1 (Implementação)	21,54	20,87	21,41	21,57
Evento 1 (Previsão)	50,85	48,09	48,77	48,06
Evento 2 (Implementação)	12,39	11,09	12,67	12,99
Evento 2 (Previsão)	49,56	48,46	52,94	51,08
Evento 3 (Implementação)	14,58	14,73	16,47	16,35
Evento 3 (Previsão)	47,16	48,35	56,58	50,41

6. CONCLUSÕES E RECOMENDAÇÕES

Analisando os resultados obtidos conclui-se que os métodos apresentaram resultados bem parecidos, alternando quanto a menor erro quadrado no modelo com peso. O método da Pesquisa Harmônica obteve melhores resultados nos eventos 2 (dois) e 3 (três), enquanto que o Algoritmo Genético foi melhor no evento 1 (um), no modelo com peso.

Já no modelo sem peso, o método da Pesquisa Harmônica apresentou os melhores resultados em todos os três eventos.

Percebe-se a rápida convergência da Pesquisa Harmônica, com os valores se estabilizando após poucos segundos de implementação enquanto que os Algoritmos Genéticos demoraram certo tempo até atingir os melhores resultados.

Por causa da grande área da bacia estudada e a pequena quantidade de postos, há grandes distâncias entre os postos, ocorrendo chuvas em áreas entre postos sem registro, o que explica o fato de em alguns momentos aumentar a vazão sem a ocorrência de precipitação.

Na implementação dos modelos, foi utilizado o Excell VBA, abrindo mais um leque de possibilidades de uso das planilhas. Devido à facilidade de se encontrar, praticamente em todo computador existe instalado, é de suma importância a elaboração de sistemas mais elaborados implementados em tal ambiente, facilitando o seu uso, mesmo em regiões com poucos recursos.

Recomenda-se um estudo melhor e mais aprofundado sobre os métodos utilizados, principalmente na Pesquisa Harmônica, onde se notou a pouca quantidade de estudos a seu respeito.

Recomenda-se também a pesquisa em um modelo híbrido, utilizando a Pesquisa Harmônica inicialmente até sua estabilização e após o uso do Algoritmo Genético.

7. BIBLIOGRAFIA

- AURÉLIO, M.; VELLASCO, M.; LOPES, C. H. **Descoberta de conhecimento e mineração de dados**, ICA, DEE, PUC-RIO, 1999.
- CANHOLI, A. P. **Drenagem urbana e controle de enchentes**, São Paulo, Oficina de textos, 2005.
- DOOGE, J. C. I. **Linear Theory of Hydrologic Systems**, ARS, US Department of Agriculture, Thechnical Bulletin n. 1948, 1973.
- GEEM, Z. W. **Optimal Cost Design of Water Distribution Networks using Harmony Search**, Engineering Optimization, vol. 38, no. 3, pp. 259-280, 2006.
- GEEM, Z. W.; KIM, J. H. e KIM; E. S. **Parameter estimation of the nonlinear Muskingum model using Harmony Search**, Journal of the American Water Resources Association, vol. 37, no. 5, 2001.
- GEEM, Z. W. e LEE, K. S. **A new structural optimization method based on the harmony search algorithm**, Computers and Structures, 82(9-10), 781-798, 2004.
- GOLDBARG, M. C. **Otimização combinatória e programação linear**, Rio de Janeiro, Ed. Campus, 2000.
- GOLDBERG, D. E. **Genetic algorithms in search, optimization and machine learning**, reading, MA, Addison-Wesley, 1989.
- NETO, J. F. da C. **Distribuição espacial e temporal do escoamento superficial em bacias hidrográficas**, Campo Grande, Dissertação (Mestrado) – Universidade Federal de Mato Grosso do Sul, Brasil, 2004.
- LAPPONI, J. C. **Estatística usando excell**, Rio de Janeiro, Elsevier, 2005.
- SALAS, J. D.; SMITH, R. A. **Uncertainties in hidrologic time series analysis**. Artigo apresentado no ASCE Spring Meeting, Portland, Oregon, Preprint 80-158, 1980.
- STEFFEN, J. L., **Previsão de cheia a curto prazo em trechos de rio: modelos não lineares**, Porto Alegre, Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul, Instituto de Pesquisas Hidráulicas, Brasil, 1987.
- TOMAZ, P. **Cálculos hidrológicos e hidráulicos para obras municipais**, São Paulo, Editora Navegar, 2002.

- TUCCI, C. E. M.. **Modelos hidrológicos**, Porto Alegre, Ed. da Universidade, ABRH, 1998.

ANEXO A – Valores de Vazão e chuva no evento 1

	P1	P2	P3	P4	P5	P6	P7	Q
2/jan/98	0,00	0,00	0,00	22,00	0,00	12,00	0,00	56,34
3/jan/98	0,00	0,00	0,00	0,00	0,00	14,50	0,00	54,27
4/jan/98	0,00	1,20	0,00	0,00	0,00	0,00	0,00	55,82
5/jan/98	0,00	4,20	0,00	0,00	8,10	0,00	0,70	57,92
6/jan/98	22,30	0,00	0,00	0,00	0,00	0,00	0,00	59,53
7/jan/98	10,20	1,80	0,00	0,00	10,00	0,30	0,00	67,29
8/jan/98	0,00	88,00	19,40	60,20	0,00	18,20	87,50	89,41
9/jan/98	0,00	0,00	22,30	0,00	0,00	0,00	0,00	111,67
10/jan/98	0,00	0,00	0,00	0,00	0,00	0,00	0,00	77,35
11/jan/98	0,00	0,00	0,00	0,00	0,00	0,00	0,00	61,70
12/jan/98	0,00	0,00	0,00	0,00	0,00	0,00	0,00	54,78
13/jan/98	0,00	0,00	0,00	30,00	0,00	0,00	0,00	54,27
14/jan/98	0,00	2,30	8,20	0,00	0,00	2,40	0,00	79,20
15/jan/98	0,00	4,40	0,00	0,00	0,00	4,00	0,10	63,35
16/jan/98	0,00	0,00	9,40	0,00	0,00	0,00	0,00	59,53
17/jan/98	0,00	4,50	8,50	5,30	0,00	0,00	0,00	117,62
18/jan/98	0,00	0,00	0,00	0,00	0,00	2,80	0,00	103,74
19/jan/98	0,00	0,00	0,00	31,30	6,40	0,00	2,90	67,86
20/jan/98	0,00	7,40	3,60	0,00	0,00	12,50	0,00	56,87
21/jan/98	0,00	0,00	0,00	6,90	0,00	2,30	11,60	52,23
22/jan/98	0,00	0,20	0,00	0,00	0,00	0,00	0,00	49,25
23/jan/98	0,00	0,00	0,00	17,00	0,00	0,00	0,00	51,73
24/jan/98	49,30	34,80	0,00	0,00	0,00	0,00	0,00	52,74
25/jan/98	0,00	0,00	0,00	0,00	0,00	4,70	0,00	60,61
26/jan/98	15,20	0,00	0,00	0,00	0,00	0,00	0,00	56,34
27/jan/98	0,00	0,00	8,90	0,00	0,00	8,30	0,00	49,25
28/jan/98	0,00	0,10	0,00	0,00	0,00	0,00	0,00	53,76
29/jan/98	0,00	3,90	2,80	22,40	0,80	31,20	19,00	79,82
30/jan/98	10,30	17,00	37,90	0,00	10,00	11,00	45,50	92,06
31/jan/98	0,00	7,10	40,90	53,60	20,30	27,30	3,00	153,03

ANEXO B - Valores de vazão e chuva no evento 2

	P1	P2	P3	P4	P5	P6	P7	Q
1-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	64,40
2-nov-94	42,20	3,40	0,00	0,00	0,00	20,00	0,00	53,20
3-nov-94	22,40	6,50	0,00	0,00	0,00	0,00	0,00	50,60
4-nov-94	12,20	0,00	0,00	0,00	0,00	10,00	0,00	49,10
5-nov-94	0,00	0,00	0,00	0,00	1,00	0,00	0,00	48,70
6-nov-94	0,00	20,30	0,00	42,20	0,00	0,00	0,00	48,70
7-nov-94	0,00	4,80	7,60	0,00	0,00	0,00	0,00	62,20
8-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	54,20
9-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	49,10
10-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	46,70
11-nov-94	0,00	0,00	0,00	0,00	0,00	30,00	0,00	45,30
12-nov-94	0,00	7,00	0,00	0,00	0,00	0,00	0,00	43,90
13-nov-94	25,40	0,00	0,00	0,00	0,00	20,00	0,00	42,50
14-nov-94	0,00	5,60	1,50	4,20	9,30	0,00	0,00	51,10
15-nov-94	0,00	67,80	7,20	0,00	16,80	0,00	0,00	65,50
16-nov-94	0,00	0,10	7,90	0,00	13,80	0,00	0,00	69,50
17-nov-94	22,20	61,20	59,00	0,00	23,40	40,00	0,00	77,80
18-nov-94	24,20	0,40	78,40	0,00	0,00	10,00	15,50	131,00
19-nov-94	0,00	30,00	4,20	14,20	0,00	0,00	33,30	151,00
20-nov-94	0,00	25,50	0,00	22,20	0,00	0,00	0,00	115,00
21-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	87,20
22-nov-94	0,00	0,00	0,00	0,00	0,00	90,00	0,00	67,20
23-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	58,90
24-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	57,80
25-nov-94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	56,30
26-nov-94	0,00	0,00	0,00	0,00	16,00	0,00	0,00	50,10
27-nov-94	0,00	0,10	0,00	16,80	1,30	0,00	0,00	54,20
28-nov-94	0,00	0,00	2,80	19,20	23,80	0,00	9,30	63,80
29-nov-94	0,00	37,40	36,60	40,80	0,00	0,00	26,00	62,20

ANEXO D - Valores dos parâmetros, com peso

	Evento 1		Evento 2		Evento 3	
	PH	AG	PH	AG	PH	AG
β	0,11	0,20	0,08	0,24	0,16	0,4
α_0	0,10	0,26	0,49	0,24	0,34	0,43
α_1	0,01	0,01	0,01	0,12	0,06	0,05
α_2	0,28	0,23	0,04	0,04	0,00	0,03
α_3	0,29	0,28	0,06	0,23	0,16	0,36
α_4	0,15	0,17	0,15	0,08	0,10	0,01
α_5	0,52	0,32	0,46	0,28	0,37	0,29
α_6	0,07	0,30	0,08	0,27	0,00	0,28
$a_{0,0}$	0,30	0,27	0,87	0,75	0,00	0,16
$a_{0,1}$	0,89	0,61	0,73	0,07	0,90	0,41
$a_{0,2}$	0,98	0,70	0,98	0,66	0,36	0,32
$a_{0,3}$	0,80	0,60	0,08	0,22	0,96	0,73
$a_{0,4}$	0,00	0,18	0,23	0,42	0,43	0,42
$a_{0,5}$	0,03	0,49	0,29	0,23	0,90	0,73
$a_{0,6}$	0,60	0,59	0,33	0,14	0,53	0,45
$a_{1,0}$	0,57	0,41	0,08	0,10	0,64	0,40
$a_{1,1}$	0,06	0,14	0,23	0,66	0,05	0,21
$a_{1,2}$	0,01	0,13	0,01	0,11	0,00	0,24
$a_{1,3}$	0,16	0,22	0,65	0,43	0,00	0,03
$a_{1,4}$	0,52	0,22	0,63	0,31	0,55	0,46
$a_{1,5}$	0,00	0,09	0,66	0,53	0,09	0,15
$a_{1,6}$	0,36	0,10	0,31	0,33	0,02	0,14
$a_{2,0}$	0,13	0,32	0,05	0,15	0,36	0,44
$a_{2,1}$	0,05	0,25	0,04	0,27	0,05	0,38
$a_{2,2}$	0,01	0,17	0,01	0,23	0,64	0,44
$a_{2,3}$	0,04	0,18	0,27	0,35	0,04	0,24
$a_{2,4}$	0,48	0,60	0,14	0,27	0,02	0,12
$a_{2,5}$	0,97	0,42	0,05	0,24	0,01	0,12
$a_{2,6}$	0,04	0,31	0,36	0,53	0,45	0,41
b_0	0,00	0,06	0,02	0,06	0,02	0,04
b_1	0,01	0,00	0,27	0,12	0,01	0,03
b_2	0,99	0,94	0,71	0,82	0,97	0,93

ANEXO E - Valores dos parâmetros, com peso

	Evento 1		Evento 2		Evento 3	
	PH	AG	PH	AG	PH	AG
β	0,15	0,19	0,16	0,12	0,14	0,23
α_0	0,10	0,26	0,49	0,24	0,34	0,43
α_1	0,01	0,01	0,01	0,12	0,06	0,05
α_2	0,28	0,23	0,04	0,04	0,00	0,03
α_3	0,29	0,28	0,06	0,23	0,16	0,36
α_4	0,15	0,17	0,15	0,08	0,10	0,01
α_5	0,52	0,32	0,46	0,28	0,37	0,29
α_6	0,07	0,30	0,08	0,27	0,00	0,28
$a_{0,0}$	0,30	0,27	0,87	0,75	0,00	0,28
$a_{0,1}$	0,89	0,61	0,73	0,07	0,90	0,41
$a_{0,2}$	0,98	0,70	0,98	0,66	0,36	0,32
$a_{0,3}$	0,80	0,60	0,08	0,22	0,96	0,73
$a_{0,4}$	0,00	0,18	0,23	0,42	0,43	0,42
$a_{0,5}$	0,03	0,49	0,29	0,23	0,90	0,73
$a_{0,6}$	0,60	0,59	0,33	0,14	0,53	0,45
$a_{1,0}$	0,57	0,41	0,08	0,10	0,64	0,40
$a_{1,1}$	0,06	0,14	0,23	0,66	0,05	0,21
$a_{1,2}$	0,01	0,13	0,01	0,11	0,00	0,24
$a_{1,3}$	0,16	0,22	0,65	0,43	0,00	0,03
$a_{1,4}$	0,52	0,22	0,63	0,31	0,55	0,46
$a_{1,5}$	0,00	0,09	0,66	0,53	0,09	0,15
$a_{1,6}$	0,36	0,10	0,31	0,33	0,02	0,14
$a_{2,0}$	0,13	0,32	0,05	0,15	0,36	0,44
$a_{2,1}$	0,05	0,25	0,04	0,27	0,05	0,38
$a_{2,2}$	0,01	0,17	0,01	0,23	0,64	0,44
$a_{2,3}$	0,04	0,18	0,27	0,35	0,04	0,24
$a_{2,4}$	0,48	0,18	0,27	0,35	0,04	0,24
$a_{2,5}$	0,97	0,42	0,05	0,24	0,01	0,12
$a_{2,6}$	0,04	0,31	0,36	0,53	0,45	0,41
b_0	0,00	0,06	0,02	0,06	0,02	0,04
b_1	0,01	0,00	0,27	0,12	0,01	0,03
b_2	0,99	0,94	0,71	0,82	0,97	0,93

