3
Fundamentação Teórica

3.1
Ajuste de Histórico

O processo de ajuste de histórico tem como objetivo a determinação dos parâmetros do reservatório e do aquífero, bem como do modelo de influxo de água. Uma vez determinados, o modelo de aquífero e os volumes de óleo e/ou gás do reservatório podem ser utilizados doravante para a previsão do comportamento do reservatório [8].

O procedimento consiste no ajuste de uma ou mais propriedades do modelo de simulação, a fim de, ajustar os dados de produção gerados pela simulação aos dados históricos, até que se obtenha uma discrepância aceitável. Durante o processo, os parâmetros do reservatório e/ou do aquífero podem ser alterados para que se obtenha o melhor ajuste possível [1]. Portanto, o ajuste de histórico pode ser interpretado como um processo de otimização onde se procura minimizar as diferenças entre os valores calculados, durante a simulação, e os valores observados no campo. Estes valores podem ser, por exemplo, as pressões, taxas de produção de óleo, água e gás.

Em [9] é apresentada uma definição matemática para o problema, onde o objetivo é determinar uma distribuição espacial *r(x)*, que pode pertencer ao R^2 ou R^3, e um conjunto de parâmetros *P* do modelo, tal que, dado o histórico O_h, $O_s - O_h \approx 0$, onde $O_S = f(R,P)$ são os dados obtidos pela simulação do modelo de reservatório. Em outras palavras, o problema do ajuste de histórico consiste em encontrar um zero para uma função de múltiplas variáveis que variam em função do tempo. Esta função $f(R,P)$ é representada pelo simulador, onde *R* é a malha de blocos do modelo e *P* os valores das propriedades em cada bloco.

O ajuste de histórico é considerado um problema inverso, onde se utilizam os dados observados para ajustar os parâmetros do modelo. Ele também é definido como um problema mal condicionado, uma vez que várias configurações de parâmetros podem resultar em um bom ajuste entre os dados simulados e observados [10]. De acordo com [3], uma boa solução deve considerar, além do ajuste entre as curvas de produção, uma distribuição adequada dos parâmetros no modelo do reservatório, o que torna o problema altamente complexo.

Geralmente estes modelos geológicos são construídos por uma equipe de geologia e o ajuste de suas propriedades é realizado por outra equipe que altera individualmente ou homogeneamente as propriedades de cada bloco do modelo de simulação, podendo assim, resultar em modelos geologicamente inconsistentes.

Segundo [11] o processo de ajuste de histórico pode ser divido em três classes: manual, automático e assistido.

3.1.1 Ajuste de Histórico Manual

Segundo [3] no ajuste manual, o critério para se estabelecer uma configuração adequada de parâmetros é totalmente empírico. Os ajustes são totalmente dependentes do especialista que, com base na sua experiência, tenta encontrar a melhor configuração para os parâmetros. Segundo [12], no procedimento de ajuste alteram-se os parâmetros individualmente, ou seja, caso a alteração em um não seja adequada, escolhe-se outro parâmetro para ser ajustado. Assim, todo o trabalho de acompanhamento das simulações e avaliação dos resultados a cada simulação é feito manualmente pelo especialista. Segundo [13], em muitas das vezes a experiência do especialista contribui para que boas soluções sejam encontradas num curto espaço de tempo, mas com o aumento do grau de complexidade do problema a obtenção dessas soluções tende a se tornar muito trabalhosa.

3.1.2 Ajuste de Histórico Automático

O ajuste automático considera que todas as etapas do processo de ajuste de parâmetros são automatizadas e, portanto, a intervenção do especialista é mínima. Porém, segundo [12], o processo de ajuste é bastante complexo e dificilmente será automatizado em sua totalidade. As principais decisões devem continuar sendo tomadas com cuidado pelo especialista a fim de evitar respostas erradas e simulações desnecessárias.

3.1.3
Ajuste de Histórico Assistido

Segundo [3][13], o ajuste assistido é a tentativa de unir as vantagens do ajuste manual e do ajuste automático. O ajuste assistido combina a automatização de determinadas partes do processo com as tomadas de decisões baseadas na experiência do especialista. Assim, enquanto as técnicas de otimização se encarregam de explorar melhor o espaço de soluções, o especialista tem a liberdade de interferir no processo de otimização a fim de colaborar para a obtenção de resultados melhores.

Em [14] são apresentados alguns fluxogramas com atividades específicas que devem ser realizadas para o ajuste assistido. As instruções presentes nesses fluxogramas variam de acordo com as características particulares do reservatório em que o ajuste está sendo aplicado e também de acordo com as técnicas utilizadas para a realização do ajuste.

Também em [14], é apresentado um exemplo que ilustra como a participação do especialista durante o ajuste de histórico pode ser benéfica. Trata-se de um modelo de reservatório cujo ajuste manual se mostrou muito difícil. A cada ciclo de otimização foi efetuado algum tipo de intervenção do especialista, que permitiu que resultados melhores, em termos de função objetivo, fossem obtidos. O gráfico da Figura 3.1 mostra o comportamento da função objetivo e os dias indicados correspondem às intervenções.

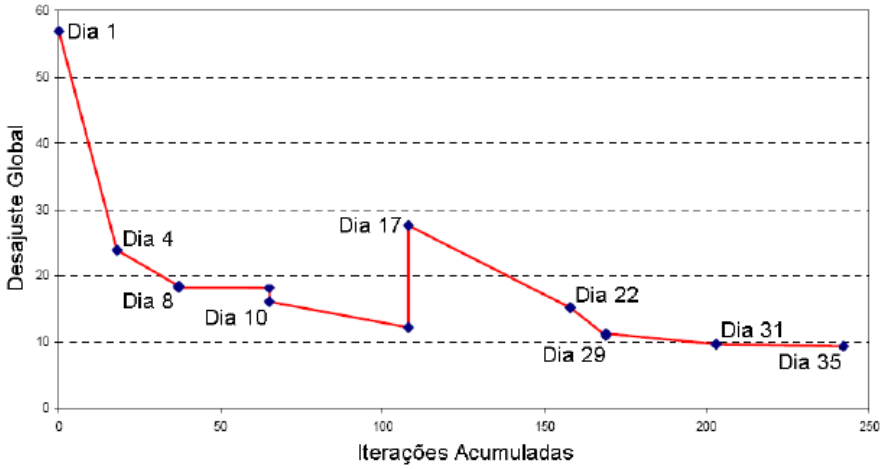

Figura 3.1: Progresso do ajuste de histórico assistido (Fonte, Silva 2011)

3.2 Histórico

Os primeiros trabalhos abordando o ajuste de histórico foram apresentados no início dos anos 60 por Jacquard e Jain.

1964 – Jacquard utilizou o método de zoneamento para dividir um reservatório em regiões uniformes de maneira que a propriedade de interesse tivesse um comportamento homogêneo e através do método de Newton tentou ajustar os parâmetros do modelo [15].

1965 – Jacquard e Jain dividiram um reservatório em regiões onde a permeabilidade era considera homogênea e utilizaram métodos baseados em gradientes para o cálculo dos parâmetros [16].

Desde então, muitas metodologias e técnicas para resolver o problema de ajuste de histórico vem sendo apresentadas. Vários métodos de otimização divididos em dois grandes grupos são apresentados em [13]. Esses grupos foram definidos como métodos locais e métodos globais. Essa classificação é estendida em [3], incluindo métodos baseados em inteligência computacional, linhas de corrente, métodos estocásticos e novas tendências. A seguir serão apresentadas algumas propostas que abordaram o problema utilizando algoritmos genéticos.

1999 – Uma das primeiras abordagens utilizando algoritmos genéticos foi apresentada por [17]. Em seu modelo de solução, cada gene do cromossoma representava uma célula do modelo e armazenava os valores de permeabilidade vertical, horizontal e de porosidade. A simplicidade dos operadores genéticos não impedia a geração de modelos com contrastes geológicos acentuados.

2000/2001 – A fim de reduzir o espaço de busca, [18] apresentou uma nova modelagem para o cromossoma. Esse cromossoma era composto por seis segmentos e seus genes codificavam diferentes tipos de parâmetros do reservatório. Três segmentos multidimensionais utilizavam genes reais que representavam, respectivamente, a porosidade, permeabilidade e fração de óleo. Os outros três segmentos eram unidimensionais com genes binários. Esses segmentos representavam os parâmetros geoestatísticos, parâmetros de falhas e de *skin* dos poços. Devido à modelagem cromossômica, caso os modelos fossem gerados por simulação sequencial, poderiam ser gerados mais de um modelo a partir do mesmo cromossoma, ou seja, a unicidade de soluções não era garantida.

2008 – Uma metodologia que combina a geoestatística com algoritmos genéticos foi proposta por [19]. Inicialmente é gerado um conjunto de realizações geoestatísticas de mapas de distribuição de litofácies, de porosidade e de permeabilidade. Estas propriedades incluindo as litofácies são usadas como parâmetros no ajuste de histórico. O cromossoma que codifica a solução do problema tem representação binária e representa os seguintes parâmetros: os identificadores das realizações dos mapas de litofácies, porosidade e de permeabilidade, a razão entre as permeabilidades horizontal e vertical e o expoente da equação de elevação de escala da permeabilidade horizontal.

2011 – Uma metodologia que combina algoritmos genéticos com geoestatística de múltiplos pontos é proposta por [3]. Nessa abordagem busca-se manter a consistência geológica do modelo através do algoritmo FILTERSIM [20]. No entanto, esta abordagem permite apenas o ajuste de histórico pela otimização de uma única propriedade do modelo de reservatório.

3.3
Algoritmos Genéticos

Os algoritmos genéticos constituem uma técnica de busca e otimização, inspirada no princípio Darwiniano de seleção natural e reprodução genética que privilegia os indivíduos mais aptos com maior longevidade e, portanto, com maior probabilidade de reprodução [21].

Segundo o princípio da sobrevivência dos mais aptos, proposto pelo inglês Charles Darwin em sua *Teoria da Evolução das Espécies*:

> *"Quanto melhor um indivíduo se adaptar ao seu meio ambiente maior será sua chance de sobreviver e gerar descendentes."*

Estes princípios são imitados na construção de algoritmos computacionais que buscam uma melhor solução para um determinado problema, através da evolução de populações de potenciais soluções deste problema. Estas potenciais soluções são codificadas por meio de cromossomas artificiais.

Métodos dessa natureza mostram-se interessantes na resolução de problemas complexos de otimização, pois conseguem um equilíbrio entre a capacidade de exploração do espaço de busca de soluções e o aproveitamento

das melhores soluções ao longo da evolução, tornando-se menos suscetível ao aprisionamento em ótimos locais.

Em algoritmos genéticos um cromossoma é uma estrutura de dados que representa uma das possíveis soluções do espaço de busca do problema. Os cromossomas são submetidos a um processo evolucionário que envolve avaliação, seleção, recombinação (cruzamento) e mutação. Após vários ciclos de evolução a população deverá conter indivíduos mais aptos [22].

Algoritmos genéticos são utilizados para resolver problemas do tipo $f: S \rightarrow R$, onde *S* é o espaço de busca formado pelas soluções do problema. Para todas as soluções existentes no domínio de *S*, um número real é associado, medindo quão adequada, ou apta, é a solução para resolver o problema. Deste modo, a tarefa do algoritmo genético é encontrar, de forma eficiente, em amostras do espaço de busca *S*, soluções satisfatoriamente aceitáveis para o problema.

3.3.1
Estrutura Básica de Um Algoritmo Genético

Um algoritmo genético é basicamente constituído por uma população, ou seja, um conjunto de indivíduos que participam de um processo de evolução. Estes indivíduos possuem uma aptidão e um cromossoma. O cromossoma é a representação genética da solução do problema e este é composto por vários genes, enquanto a aptidão representa quão boa esta solução é para este determinado problema. A Figura 3.2 ilustra uma população contendo três indivíduos.

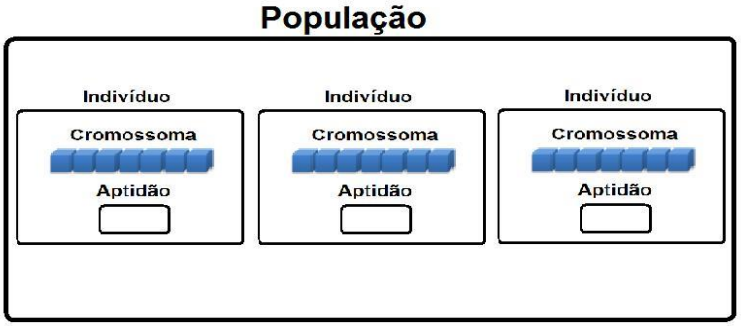

Figura 3.2: Ilustração de população contendo 3 indivíduos

O processo de evolução começa com a criação aleatória dos indivíduos que formarão a população inicial. Essa população inicial também pode ser criada pelo usuário e ser utilizada como semente inicial, e assim, o algoritmo evoluirá a partir desta solução dada. Posteriormente, através de uma função de avaliação, também conhecida como função objetivo, é atribuído um valor de adaptação para cada indivíduo, denominado aptidão, que indica o quanto a solução representada pelo cromossoma deste indivíduo é melhor em relação às outras soluções da população.

Aos indivíduos mais adaptados é dada a oportunidade de se reproduzirem, através de vários critérios de seleção, como por exemplo, a roleta, que permite uma seleção proporcional ao valor da aptidão de cada indivíduo. A partir daí, esses indivíduos selecionados se reproduzem, mediante cruzamentos com outros indivíduos da população, produzindo descendentes com características de ambas as partes. A mutação também tem um importante papel, pois esta altera algumas características do indivíduo de forma aleatória.

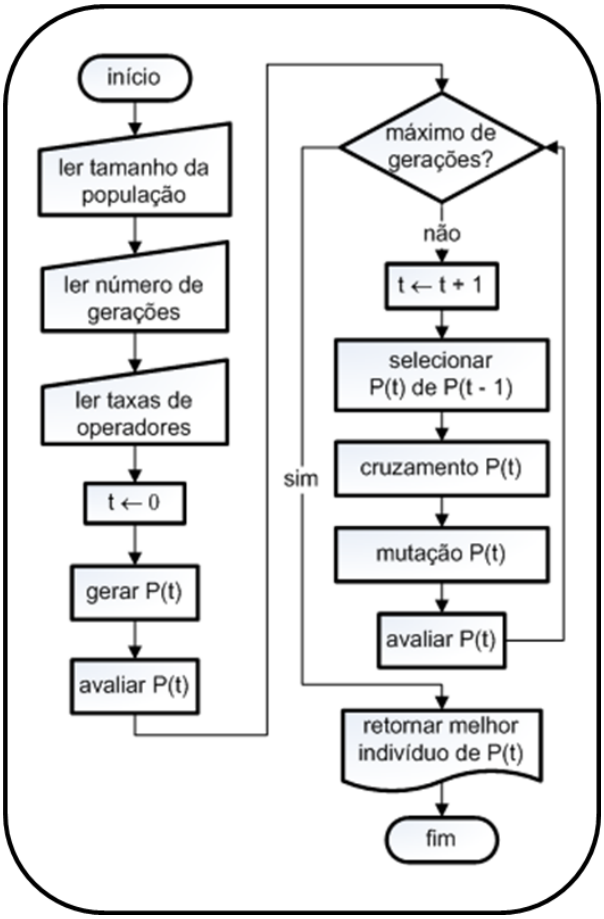

Figura 3.3: Estrutura básica de um Algoritmo Genético

Para determinar o final da evolução pode-se fixar o número de gerações, o número de indivíduos criados, condicionar o algoritmo à obtenção de uma solução satisfatória ou quando atingir um ponto considerado ótimo. Outras condições para a parada incluem o tempo de processamento e o grau de similaridade entre os elementos numa população (convergência). O esquema de um Algoritmo Genético, conforme descrito anteriormente, pode ser visto na Figura 3.3.

As seções seguintes apresentam detalhadamente cada um dos componentes de um Algoritmo Genético.

3.3.1.1
Representação ou Codificação

A representação é um aspecto fundamental na modelagem de um algoritmo genético para a solução de um determinado problema. Ela define a estrutura do cromossoma de um indivíduo, com os respectivos genes que o compõem, de maneira que seja capaz de descrever todo o espaço de busca relevante do problema.

A solução de um problema pode ser representada por um conjunto de parâmetros, denominados genes. Estes genes são unidos para formar uma cadeia de valores, também chamada de cromossoma, a este processo chama-se codificação.

Os cromossomas, ou seja, as soluções do problema são codificadas através de uma sequência formada por caracteres de um sistema alfabético. Originalmente, utilizou-se o alfabeto binário {0, 1}, porém, novos tipos de representações como, números reais, inteiros, grupamento de inteiros e baseados em ordem, vem sendo utilizados para codificar as soluções.

3.3.1.2
Geração da População Inicial

Após a definição da representação, uma população inicial de indivíduos é gerada. Na maioria das vezes, esta população é gerada de forma aleatória. No entanto, existem ocasiões onde é mais apropriada a utilização de uma heurística para a geração da população inicial. Deste modo, é possível introduzir na população inicial indivíduos com características interessantes, como por

exemplo, acrescentar soluções conhecidas, aproximadas, contendo algum tipo de informação prévia.

Segundo [23], a geração da população inicial não é uma fase crítica em Algoritmos Genéticos, no entanto, é necessário que a população inicial contenha indivíduos suficientemente diversificados.

3.3.1.3
Decodificação

Antes de avaliar os indivíduos de uma população é necessário realizar a decodificação do cromossoma. A decodificação consiste basicamente na construção da solução real do problema a partir do cromossoma para que esta seja submetida à função de avaliação.

3.3.1.4
Avaliação da População

A avaliação é a ligação entre o Algoritmo Genético e o problema a ser solucionado. Ela é feita através de uma função, conhecida como função objetivo, que melhor representa o problema e tem por objetivo oferecer uma medida de aptidão de cada indivíduo da população corrente, a fim de guiar o processo de busca.

A função de avaliação calcula o valor da solução, para cada um dos cromossomas, após a decodificação.

3.3.1.5
Seleção dos Indivíduos

Após a etapa de avaliação dos indivíduos de uma população, ocorre o processo inspirado na seleção natural. O processo de seleção está baseado no princípio da sobrevivência dos melhores indivíduos, ou seja, os indivíduos com melhores aptidões possuem maiores probabilidades de fazerem parte de uma nova população.

Existem vários métodos para selecionar os indivíduos sobre os quais serão aplicados os operadores genéticos. A seguir, apresentam-se resumidamente alguns destes métodos:

Seleção por *ranking*: os indivíduos de uma população são ordenados de acordo com seu valor de aptidão e então sua probabilidade de escolha é atribuída conforme a posição que ocupem.

Seleção por roleta: os indivíduos de uma população são distribuídos em uma roleta proporcionalmente ao seu valor de aptidão. Desta maneira os mais aptos ocuparão porções maiores e consequentemente terão mais chances de serem selecionados para compor a próxima população.

Como o tamanho da população é mantido constante ao longo de todo o processo evolutivo, na transição de uma geração para outra, a roleta deve ser girada um número de vezes igual ao tamanho da população.

Devido às características da roleta, espera-se que na geração seguinte, a nova população seja formada pelos indivíduos mais aptos da geração anterior e que os menos aptos sejam eliminados. A Figura 3.4 apresenta uma roleta com cinco indivíduos {“1”, “2”, “3”, “4”, “5”}. Nela é possível observar que o indivíduo identificado como “1” apresenta maior aptidão, uma vez que ocupa um espaço maior na roleta. O contrário acontece com o indivíduo identificado como “2”. Sendo assim, a probabilidade do indivíduo “1” compor a população na geração seguinte é bem maior que a do indivíduo “2”.

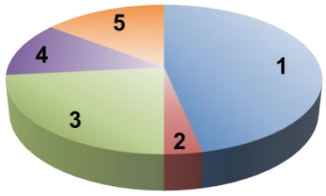

Figura 3.4: Exemplo de seleção por roleta.

Seleção por torneio: a ideia deste método é promover um torneio entre um grupo de n ($n \geq 2$) indivíduos aleatoriamente tomados na população. Assim, o indivíduo com melhor valor de aptidão entre o grupo vencerá o torneio e será selecionado para fazer parte da geração da nova população, enquanto os demais indivíduos do grupo são descartados. O processo de seleção termina quando se realiza uma quantidade de torneios igual ao tamanho da população.

3.3.1.6
Operadores Genéticos

Após a etapa de seleção, alguns indivíduos da nova população sofrem a ação dos operadores genéticos. Os operadores genéticos são responsáveis pela alteração destes indivíduos, gerando assim, novas soluções para o problema. Os principais operadores dos Algoritmos Genéticos são os de cruzamento e mutação. A cada operador genético é atribuída uma probabilidade de aplicação, e a partir dela, é possível selecionar um subconjunto de indivíduos que devem ser modificados.

O operador de cruzamento consiste em recombinar o material genético de dois indivíduos de modo que dois novos sejam gerados. Deste modo, é possível transmitir informação genética de pai para filho. De acordo com a teoria da evolução, este evento representa a hereditariedade. A ideia do operador cruzamento é tirar vantagem do material genético presente na população, e seu funcionamento é apresentado na Figura 3.5.

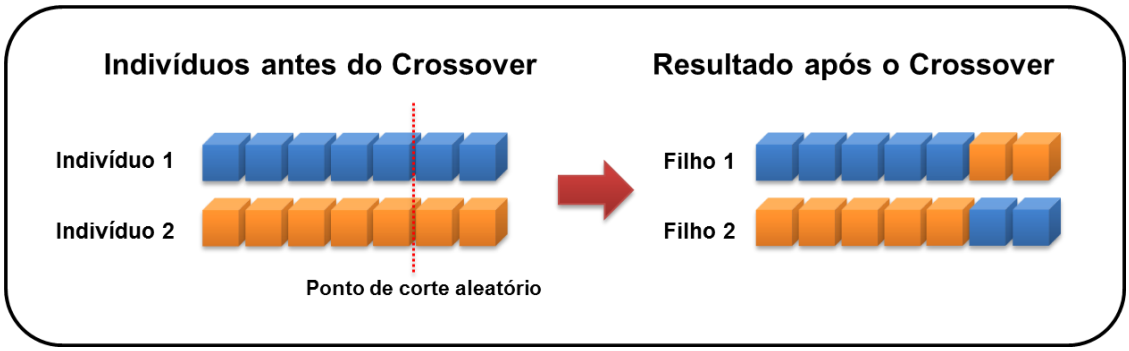

Figura 3.5: Cruzamento de um ponto

O ponto onde ocorre o corte para a realização do cruzamento é escolhido aleatoriamente. No exemplo da Figura 3.5 utilizou-se um único ponto, mas segundo [24][25][26], podem ser realizados cortes em mais de um ponto, caracterizando assim, o cruzamento multiponto.

A Mutação é a troca aleatória do valor contido nos genes de um cromossomo por outro valor válido. Este operador garante a diversidade das características dos indivíduos da população e permite que sejam introduzidas informações que não estejam presentes na população. Além disso, proporciona uma busca aleatória, oferecendo a oportunidade de que mais pontos do espaço de busca sejam avaliados. Um exemplo da operação de mutação pode ser visualizado na Figura 3.6.

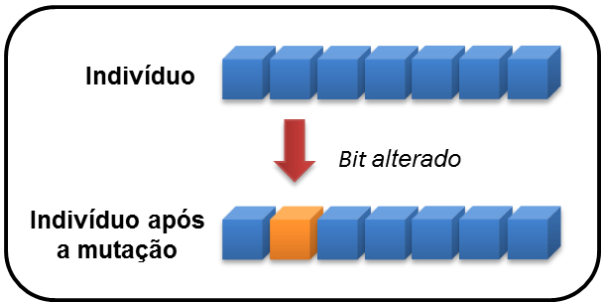

Figura 3.6: Mutação

Após a aplicação dos operadores genéticos tem-se uma nova população e para encerrar um ciclo de evolução, também conhecida como geração, avaliam-se estes novos indivíduos.

3.3.1.7 Parâmetros da Evolução

Os principais parâmetros de um algoritmo genético são:

Tamanho da População: o tamanho da população afeta o desempenho global e a eficiência dos Algoritmos Genéticos. Uma população muito pequena oferece pouca cobertura do espaço de busca, causando uma queda no desempenho. Uma população suficientemente grande fornece uma melhor cobertura do domínio do problema e previne a convergência prematura para soluções locais. Entretanto, com uma população grande tornam-se necessários recursos computacionais maiores, ou um tempo maior de processamento do problema. Logo, deve-se buscar um ponto de equilíbrio quanto ao tamanho da população.

Número de gerações: representa o número total de ciclos de evolução de um algoritmo genético. O número de gerações é considerado um dos critérios de parada do algoritmo.

Taxa de *Steady State*: esta taxa controla a porcentagem da população que é preservada de uma geração para a outra. Este parâmetro garante que os melhores indivíduos de uma geração sejam mantidos em uma próxima geração. O valor desse parâmetro pode ser uma taxa adaptativa onde no início da evolução um número maior de indivíduos é mantido entre gerações e esta taxa decai linearmente ao longo da evolução. O número de indivíduos que será substituído também é conhecido como *gap*.

Taxa de Cruzamento: esta taxa representa a probabilidade de um indivíduo ser recombinado com outro. Quanto maior for esta taxa, mais

rapidamente novas estruturas serão introduzidas na população. Entretanto, isto pode gerar um efeito indesejável, pois a maior parte da população será substituída, ocorrendo assim perda de variedade genética, podendo ocorrer perda de estruturas de alta aptidão e convergência a uma população com indivíduos extremamente parecidos, indivíduos estes de solução melhores ou não. Com um valor baixo, o algoritmo pode tornar-se muito lento para oferecer uma resposta aceitável. Uma solução é utilizar uma taxa adaptativa, onde a taxa é maior no início da evolução e diminui linearmente ao longo da evolução.

Taxa de Mutação: esta taxa representa a probabilidade do conteúdo de um gene do cromossoma ser alterado. A taxa de mutação previne que uma dada população fique estagnada, além de possibilitar que se chegue a qualquer ponto do espaço de busca. Porém, deve-se evitar uma taxa de mutação muito alta, uma vez que pode tornar a busca essencialmente aleatória, prejudicando fortemente a convergência para uma solução ótima. Para evitar este problema, utiliza-se a taxa adaptativa, sendo esta baixa no início da evolução e mais alta no final da evolução.

3.3.2
Algoritmo Genético Coevolutivo

O algoritmo genético coevolutivo é o resultado de alguns aprimoramentos realizados para que se pudessem abordar problemas mais complexos. Segundo [27], para se aplicar algoritmos evolucionários com sucesso em problemas com complexidade cada vez maior, torna-se necessário introduzir noções explícitas de modularidade nas soluções para que elas disponham de oportunidades razoáveis de evoluir na forma de subcomponentes co-adaptados.

Em algoritmos genéticos convencionais, mesmo que se divida um problema complexo em subsoluções, ou subcomponentes, toda a informação ainda estará codificada em um único indivíduo. Supondo que um indivíduo seja composto por um bom subcomponente e outro considerado péssimo, ele receberá uma aptidão média ou inferior, porque durante o processo de avaliação ele será avaliado como um todo. Este fato implica na perda de bons subcomponentes durante o processo evolutivo, resultando em um baixo desempenho do algoritmo.

Em um algoritmo genético coevolutivo, a decomposição de um problema complexo é realizada em subcomponentes interdependentes que evoluirão em

seu próprio espaço de busca, desaclopado dos outros, onde cada subcomponente é representado por um indivíduo. Deste modo, os indivíduos são separados em populações distintas de acordo com as suas características, possibilitando uma interação entre os membros de uma mesma população, ou espécie. Durante o processo de avaliação, uma solução completa é representada pela composição de um indivíduo de cada população.

Segundo [28], a coevolução é definida como a evolução complementar de espécies intimamente associadas. As adaptações inter-relacionadas existentes nas plantas floríferas e seus insetos polinizadores são exemplos claros de coevolução. O relacionamento predador-presa também envolve coevolução, pois um avanço evolucionário no predador estimula uma resposta evolucionária na presa.

Na coevolução competitiva, inspirada no sistema predador-presa, o sucesso de uma das partes é sentido pela outra como uma falha que deve ser corrigida, para que as suas chances de sobrevivência sejam mantidas ou aumentadas. Este tipo de coevolução exerce uma forte pressão evolucionária para que os predadores desenvolvam melhores estratégias de ataques, enquanto às pressas aperfeiçoam as suas defesas. Geralmente, nessa arquitetura os indivíduos são avaliados com uma medida relativa de aptidão, ou seja, quanto mais adaptado for um indivíduo em relação aos seus competidores, maior será sua avaliação. Esta medida é chamada de *competitive fitness* [29].

Em [30], a coevolução competitiva foi utilizada para solucionar problemas de aprendizado de jogos, onde cada espécie representava um jogador. Em [31], também utilizou-se a abordagem competitiva para coevoluírem estratégias para o problema do dilema dos prisioneiros.

A coevolução cooperativa foi inspirada na simbiose, onde duas ou mais espécies colaboram uma com a outra para que ocorra uma evolução conjunta. Ao contrário da coevolução competitiva, onde o aumento da aptidão de um indivíduo ocasiona a queda da aptidão relativa do outro, aqui o sucesso de uma espécie também é o sucesso da outra.

A abordagem coevolutiva cooperativa foi utilizada em [32] para maximizar uma função de *n* variáveis independentes. O problema foi decomposto em *n* espécies e um indivíduo de cada espécie era selecionado para formar uma solução completa a ser avaliada. Em [33], também utilizou-se a coevolução cooperativa para a otimização da função de três bits de Goldberg.

Em [34], utilizou-se a cooperação entre as espécies para evoluir regras de controle de um robô simulado. Nessa abordagem cada espécie continha um conjunto de regras de uma classe de comportamentos. Para a avaliação dos indivíduos de uma espécie era selecionado o melhor da outra, e o conjunto resultante era utilizado para controlar o robô.

Em um problema de planejamento aplicado a descarga de minério em portos, [35] propôs um modelo com duas espécies onde uma evoluía soluções de alocação das tarefas no tempo, enquanto a outra evoluía soluções de alocação de recursos para as tarefas. A avaliação era feita com base no custo de Demurrage, ou seja, a multa por atraso nos carregamentos dos navios.

Uma otimização utilizando a evolução cooperativa foi desenvolvida por [36] para a resolução do *Timetabling* na produção de grades horárias em instituições de ensino. Em seu estudo de caso foram definidas oito espécies, onde cada espécie representava um período do curso de quatro anos. A otimização seria concluída quando todas as turmas apresentassem escalas compatíveis com as restrições impostas.

Recentemente, [37] utilizou um algoritmo coevolutivo cooperativo para configuração de uma rede de sensores sem fio. Seu objetivo era encontrar uma configuração de rede com uma estrutura de comunicação que apresentasse um pequeno comprimento para o caminho médio mínimo e um elevado coeficiente de agrupamento.

3.3.2.1 Modelo Coevolucionário Cooperativo

Em algoritmos coevolucionários cooperativos, novos conceitos como espécies e ecossistema são introduzidos. Nesta arquitetura, cada população representa uma espécie, e duas ou mais espécies diferentes formam um ecossistema. Assim como na natureza, as espécies são geneticamente isoladas, ou seja, os indivíduos que compõem uma população só podem se reproduzir com outros indivíduos da mesma espécie. Isto é feito isolando-se as espécies em populações separadas, de modo que, elas somente interajam umas com as outras através de um domínio compartilhado e tenham uma relação apenas de cooperação.

O algoritmo genético coevolutivo segue a estrutura básica dos Algoritmos Genéticos, apresentada na Seção 3.3.1, mas com algumas peculiaridades, tais como, mais de uma população para reprodução genética e a flexibilidade de

desenvolvimento de operadores genéticos aplicáveis a uma população específica. Como cada espécie representa um subcomponente da solução do problema, durante o processo de avaliação, é necessário selecionar indivíduos das outras espécies para se construir uma solução completa que represente adequadamente o problema. Esse processo de seleção de indivíduos de outra espécie é chamado de colaboração, e será apresentado na próxima seção.

Um modelo coevolucionário cooperativo genérico é exibido na Figura 3.7. Apesar de serem mostradas apenas três espécies, o mesmo pode ser utilizado para *n* espécies. Cada espécie evolui em sua própria população e se adapta ao ambiente através de repetidas aplicações do algoritmo evolutivo. Para o cálculo da aptidão (*fitness*) de um indivíduo de uma determinada espécie, deve-se submetê-lo ao modelo de domínio juntamente com um ou mais colaboradores de cada uma das outras espécies para que se forme uma solução completa a ser avaliada pela função objetivo.

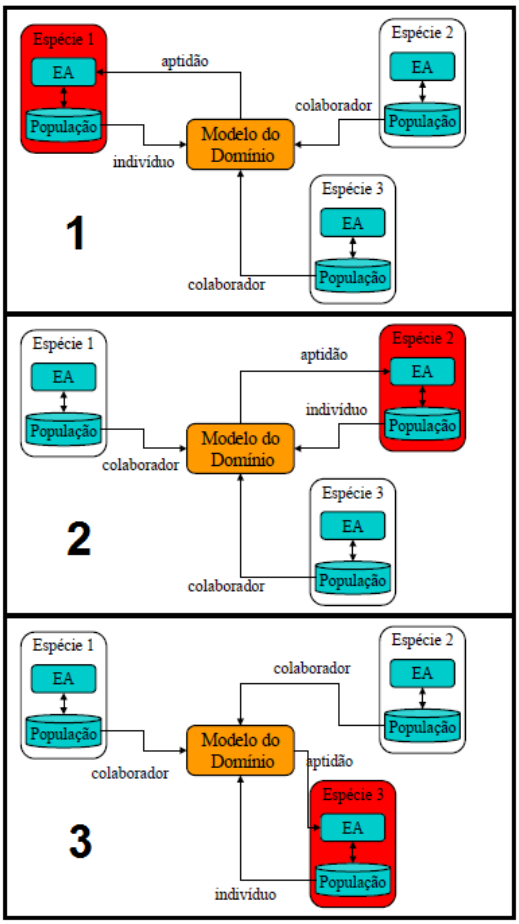

Figura 3.7: Modelo coevolucionário cooperativo genérico (Adaptado de Potter & Jong, 2000)

No primeiro passo do processo de avaliação do algoritmo evolutivo cooperativo exibido na Figura 3.7, é apresentada a avaliação de um indivíduo da população que representa a espécie 1. Outros dois indivíduos provenientes das espécies 2 e 3 são selecionados como colaboradores. Juntos, os três indivíduos formarão uma solução completa a ser avaliada pela função de avaliação. Esta função está definida dentro do modelo de domínio, e seu valor de avaliação é atribuído como aptidão para o indivíduo da espécie 1. Este processo será repetido para todos os indivíduos pertencentes à população da espécie 1.

Nos passos 2 e 3, o processo descrito acima se repete para todos os indivíduos das espécie 2 e 3, respectivamente.

Ao final deste processo de avaliação, inicia-se um novo ciclo de evolução. Os indivíduos de cada população voltam a se reproduzir com os mesmos de sua espécie para posteriormente serem submetidos a um novo processo de avaliação. Este ciclo se repetirá a cada geração até que se alcance uma condição de parada estabelecida previamente.

3.3.2.2 Métodos de Seleção de Colaboradores

Quando se aplica a coevolução cooperativa, a abordagem mais comum é decompor o problema em subcomponentes e associar a cada um deles uma população. A evolução ocorre de forma independente para cada uma das populações, exceto durante a avaliação. Como os indivíduos de cada espécie representam apenas parte da solução, deve-se selecionar indivíduos de outras espécies, segundo algum critério, para formar uma solução completa que represente adequadamente o problema. Este processo é chamado de colaboração.

Segundo [35], o processo de seleção de colaboradores (*collaborator selection*), geralmente, é realizado utilizando-se o último conjunto de avaliações feitas em cada espécie. O peso destas avaliações durante a seleção dos colaboradores é chamado de pressão de seleção de colaboradores (*collaborator selection pressure*). Se apenas um indivíduo de cada espécie for selecionado, ter-se-á apenas um valor de avaliação que poderá ser utilizado como aptidão (*fitness*).

Entretanto, podem-se escolher diferentes combinações de colaboradores das outras espécies. Neste caso, a avaliação do indivíduo de uma determinada espécie vai consistir de múltiplas colaborações. O número de colaboradores de

cada espécie é chamado de tamanho do grupo de colaboração (*collaboration pool size*). Como cada uma dessas colaborações vai ter uma avaliação diferente, é necessário, a partir delas, calcular uma avaliação única para a aptidão (*fitness*). Esta operação chama-se associação de crédito de colaboração (*collaboration credit assignment*) [35].

Segundo [38], existem três métodos para a associação de crédito de colaboração:

- **Otimista**: neste método é associado ao valor de aptidão do indivíduo, o valor da melhor colaboração feita por ele. Este é o método mais tradicional.
- **Média**: associa-se como aptidão, o valor médio das colaborações feitas pelo indivíduo.
- **Pessimista**: a avaliação da pior colaboração é utilizada como aptidão.

Em todos os testes realizados por [38], os resultados apresentados pelos métodos de média e pessimista foram piores em relação ao método otimista.

Quanto à pressão de seleção de colaboradores, [38] apresenta diferentes métodos que variam o grau de pressão exercido durante a seleção. O primeiro método seleciona o melhor indivíduo da população anterior e é considerado muito ganancioso. O segundo busca enfraquecer um pouco esta pressão através da seleção de dois indivíduos da população anterior, sendo o melhor e outro selecionado aleatoriamente. Este método também é considerado ganancioso, uma vez que ainda utiliza o melhor indivíduo. Para enfraquecer ainda mais esta pressão, podem-se estabelecer diferentes combinações entre o melhor indivíduo e algum mecanismo de seleção. Outro exemplo seria a escolha do melhor e do pior indivíduo, ou ainda poderia considerar a inclusão de mais um selecionado aleatoriamente.

Por último, deve-se definir o tamanho do grupo de colaboradores, ou seja, definir o número de indivíduos das outras espécies que participarão da colaboração. Segundo [38], este pode ser considerado o fator mais importante para o sucesso dos algoritmos coevolutivos cooperativos. Em geral, aumentar o número de colaboradores melhora o desempenho do algoritmo, o que também aumenta consideravelmente o custo computacional.

3.4
Geoestatística de Múltiplos Pontos

Segundo [39], desde o início do século a variabilidade espacial de algumas características do solo vem sendo uma das preocupações de pesquisadores. Em 1910, em uma tentativa de eliminar o efeito de variações do solo, Smith estudou a disposição de parcelas no campo em experimentos de rendimento de variedades de milho. Montgomery, em 1913, preocupado com o efeito do nitrogênio no rendimento do trigo, fez um experimento em 224 parcelas, medindo o rendimento de grãos. Outros autores, como Waynick e Sharp, também estudaram as variações do nitrogênio e o carbono no solo em 1919.

Naquela época, os procedimentos eram baseados na estatística clássica e utilizavam um grande volume de dados amostrais para caracterizar ou descrever a distribuição espacial da característica em estudo. Entende-se como estatística clássica aquela que se utiliza de parâmetros como média e desvio padrão para representar um fenômeno baseando-se na hipótese principal de que as variações de um local para outro são aleatórias [39].

No início dos anos 50, ao trabalhar com dados de concentração de ouro, D. G. Krige concluiu que a informação obtida pela variância era insuficiente para explicar o fenômeno estudado. Seria necessário considerar também a distância entre as amostras. Em [40], desenvolveu-se um método empírico para estimar reservas minerais, surgindo então a geoestatística. A geoestatística considera a localização geográfica e a dependência espacial dos dados. Porém, esse método só recebeu um tratamento formal no início dos anos 60, quando G. Matheron elaborou a sua teoria das variáveis regionalizadas. Segundo [41], uma variável regionalizada é uma função numérica com distribuição espacial, que varia de um ponto a outro com continuidade aparente, cujas variações não podem ser representadas por uma função matemática simples.

Segundo [3][42], os métodos geoestatísticos convencionais, baseados no cálculo do variograma, não são os mais adequados para a construção de modelos onde os objetos ou estruturas a serem representados são curvilíneos, pois o variograma calcula a variabilidade espacial entre pares de pontos. A geoestatística de múltiplos pontos torna possível a modelagem com maior fidelidade das não linearidades de determinadas propriedades do reservatório, pois ela calcula esta variabilidade entre vários pontos distribuídos no espaço.

Durante a caracterização de um reservatório, os dados reais disponíveis geralmente não são suficientes para permitir a inferência da estatística de

múltiplos pontos. Em [43], foi proposto um algoritmo de simulação sequencial por indicadores capaz de inferir as estatísticas de múltiplos pontos a partir de imagens de treinamento, em seguida, o modelo geoestatístico numérico é gerado utilizando estas estatísticas. Uma imagem de treinamento é uma representação puramente conceitual dos padrões esperados das heterogeneidades geológicas presentes no modelo do reservatório [3]. A inferência das estatísticas a partir da imagem de treinamento elimina a necessidade de calcular o variograma para identificar a variabilidade espacial dos dados, além de descartar a necessidade da krigagem para derivar as probabilidades condicionais.

Em [42], foi proposto um algoritmo sequencial chamado *SNESIM*. Esse algoritmo foi uma extensão do original proposto por [43] e proporcionou uma redução no custo computacional durante o cálculo da probabilidade condicional associada a cada ponto simulado. Para essa redução, Strebelle propôs um algoritmo que percorria apenas uma vez a imagem de treinamento durante todo o processo de simulação. Quatro anos depois, [20] propôs outro algoritmo de simulação sequencial que foi nomeado de *FILTERSIM*. Esse algoritmo apresentou um baixo consumo de memória a um custo razoável de CPU quando comparado ao *SNESIM*, além de simular tanto propriedades categóricas quanto contínuas.

O algoritmo *FILTERSIM* é o responsável pela aplicação da estatística de múltiplos pontos no modelo de solução proposto por este trabalho. Seu funcionamento erá apresentado na próxima seção.

Segundo [44], uma série de estudos publicados nos últimos anos apresentaram variadas aplicações para a simulação geoestatística condicional. Estas técnicas foram utilizadas para auxiliar na avaliação e classificação de recursos e reservas minerais por (Journel & Kyriakidis, 2004). No Brasil, a simulação condicional também tem sido aplicada na modelagem 3D de depósitos minerais (Souza, 2007), em análises de sensibilidade no sequenciamento de lavra (Peroni, 2002), na avaliação da variabilidade *in situ* de parâmetros físicos e químicos de minérios (Gambin, 2003), na otimização para locação de furos de sondagem em campanhas prospectivas (Koppe, 2009), para quantificar a incerteza das estimativas e auxiliar no dimensionamento de pilhas de homogeneização (Gambin et al, 2005 e Abichequer et al, 2011), assim como no estudo e redução da variabilidade de teores das pilhas e aperfeiçoamento de

estratégias de homogeneização de pilhas (Beretta, 2010; Costa et al, 2008; Marques, 2010).

3.4.1
Algoritmo *FILTERSIM*

O funcionamento básico do algoritmo FILTERSIM, mencionado na Seção anterior, é apresentado a seguir em maiores detalhes, adaptado de [3].

O algoritmo *FILTERSIM* [20] é dividido em duas etapas, uma de classificação dos padrões presentes na imagem de treinamento e uma de simulação da propriedade em questão. Na etapa de classificação, inicialmente o algoritmo recebe uma imagem de treinamento sobre a qual a classificação deve ser realizada. Vale ressaltar que essa imagem de treinamento deve representar adequadamente as heterogeneidades da propriedade em estudo, seja essa uma propriedade categórica ou contínua, Figura 3.8.

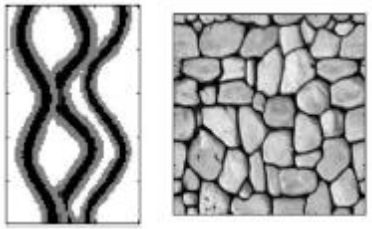

Figura 3.8: Imagens de treinamento para propriedades categórica e contínua (Fonte: Silva, 2011)

O passo seguinte, ainda dentro da etapa de classificação, consiste em atribuir um conjunto de pontuações a cada ponto da imagem de treinamento de acordo com o padrão ao redor daquele ponto. Para isso, seja *X(i, j)* o valor na posição *(i, j)* da imagem de treinamento. A pontuação $S_f(i, j)$ onde $f \in \{1, \dots, 6\}$ para o padrão na vizinhança de *(i, j)* é definida pela aplicação de um filtro *f(u, v)*, Equação 3.1.

$$S_f(i,j) = \sum_{v=-n}^{n} \sum_{u=-n}^{n} X(i+u, j+v) \cdot f(u,v) \tag{3.1}$$

Onde:

a dimensão da vizinhança local é $(2n+1) \times (2n+1)$.

A representação gráfica dos filtros a serem aplicados à imagem de treinamento são apresentadas na Figura 3.9. Nesse exemplo os tamanhos dos filtros são de 15 x 15 prixels (n = 7), mas este valor pode variar.

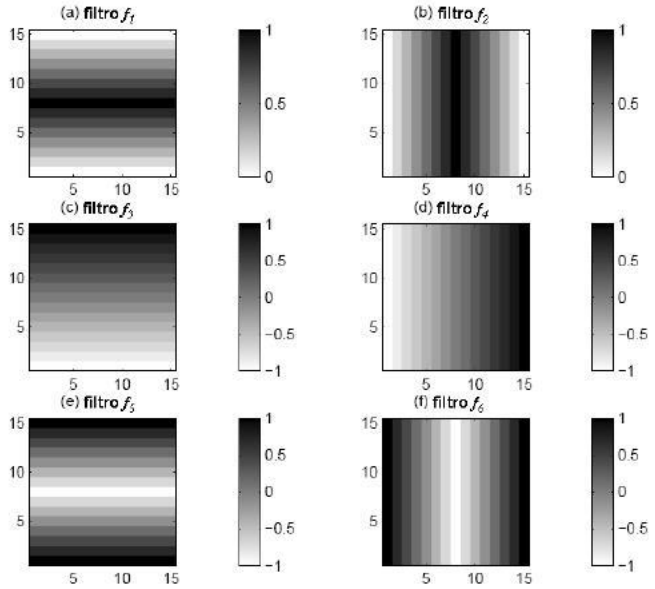

Figura 3.9: Filtros a serem aplicados à imagem de treinamento (Fonte: Silva, 2011)

As definições matemáticas dos filtros *f(u, v)* são dadas pelas Equações 3.2 a 3.7:

- f_1: filtro de média Norte-Sul.

$$f_1(u,v) = 1 - \frac{|v|}{n} \tag{3.2}$$

Onde:

$$v = -n, \ldots n.$$

- f_2: filtro de média Leste – Oeste, que é obtido mediante uma rotação de 90º em f_1.

$$f_2(u,v) = 1 - \frac{|u|}{n} \tag{3.3}$$

- f_3: filtro de gradiente Norte-Sul

$$f_3(u,v) = \frac{v}{n} \tag{3.4}$$

- f_4: filtro de gradiente Leste-Oeste, que é obtido mediante uma rotação de 90º em f_3.

$$f_4(u,v) = \frac{u}{n} \tag{3.5}$$

- f_5: filtro de curva Norte-Sul.

$$f_5(u,v) = \frac{2|v|}{n} - 1 \tag{3.6}$$

- f_6: filtro de curva Leste-Oeste, que é obtido mediante uma rotação de 90º em f_5.

$$f_6(u,v) = \frac{2|u|}{n} - 1 \tag{3.7}$$

O cálculo da pontuação em um dado ponto da imagem é exemplificado na Figura 3.10, onde o filtro f_1 é aplicado a um padrão específico da imagem de treinamento.

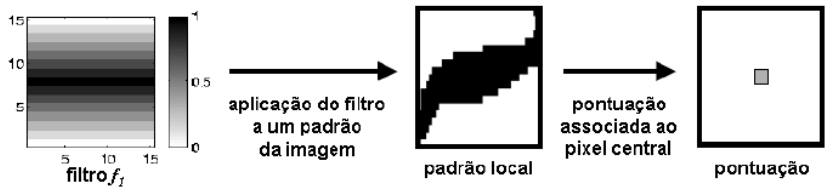

Figura 3.10: Aplicação do filtro à imagem de treinamento (Fonte: Silva, 2011)

Após a aplicação do conjunto de filtros a todos os pontos da imagem de treinamento, classificam-se todos os padrões identificados na imagem de acordo com os seis valores de pontuação atribuídos ao seu ponto central. Para isso, calcula-se a distribuição de frequência dos valores de pontuação de cada filtro. Em seguida divide-se cada distribuição de frequência em *n* partes de forma que todas elas tenham a mesma frequência. Em [45], sugere-se a divisão em cinco partes, conforme ilustra a Figura 3.11. Assim, o espaço hexadimensional de pontuações é particionado em 5^6 células.

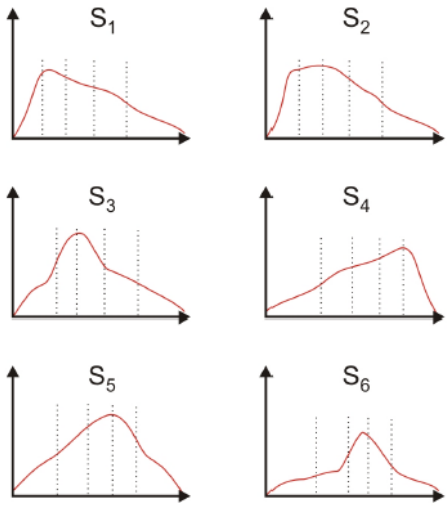

Figura 3.11: Particionamento das distribuições de frequência (Fonte: Silva, 2011)

Uma vez estabelecidas às células, associa-se cada padrão de treinamento à célula que corresponde à combinação das seis pontuações atribuídas ao padrão. Apesar da grande quantidade de células (5^6), na prática muitas dessas células não terão padrões associados devido à inexistência de padrões que apresentem as combinações de pontuações correspondentes. As células que possuem padrões associados representam as classes e, para cada classe, calcula-se um protótipo que é dado pela média de todos os padrões presentes na classe.

Na Figura 3.12 é apresentado um exemplo de um conjunto de padrões pertencentes a uma mesma classe e o seu protótipo correspondente. Na Figura 3.13 é apresentado um fluxograma que resume os passos da etapa de classificação do algoritmo *FILTERSIM*.

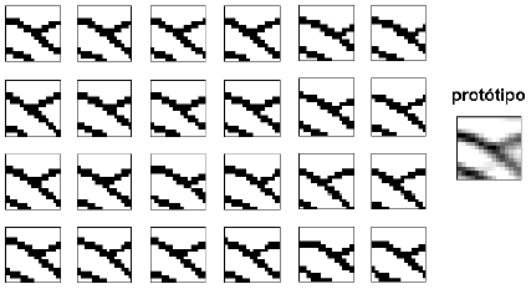

Figura 3.12: Padrões pertencentes a uma classe e seu protótipo correspondente (Fonte: Silva, 2011)

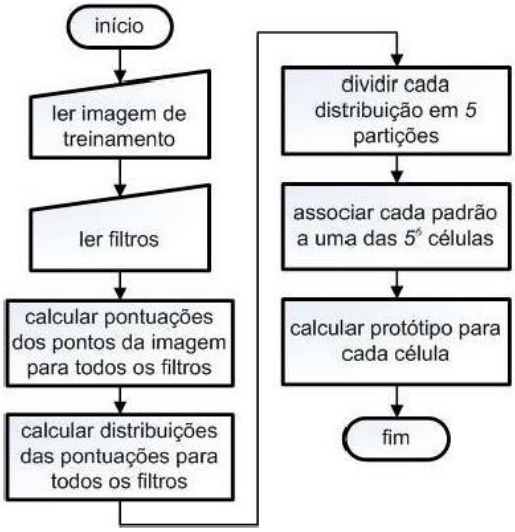

Figura 3.13: Fluxograma da etapa de classificação do algoritmo FILTERSIM (Fonte: Silva, 2011)

Uma vez finalizada a etapa de classificação dos padrões da imagem de treinamento, inicia-se a etapa de simulação dos valores da propriedade em estudo. A etapa de simulação segue os princípios da simulação sequencial e, por isso, inicialmente é preciso estabelecer um caminho aleatório para a visitação de todos os blocos do modelo do reservatório onde o valor da propriedade deve ser estimado.

Para determinar o valor da propriedade em um bloco é necessário, primeiramente, verificar se existem dados condicionantes na região em torno desse bloco. A região a ser verificada tem uma quantidade de blocos igual à quantidade de pontos dos padrões que foram inicialmente classificados, sendo o bloco a ser simulado o centro dessa região. Caso não existam dados condicionantes em torno a ser simulado, seleciona-se um padrão aleatório, pertencente a um protótipo aleatório, e substitui-se a região do modelo pelo padrão selecionado. Caso existam dados condicionantes em torno do bloco, seleciona-se o protótipo mais próximo ao padrão em torno do bloco segundo a função de distância dada pela Equação 3.8:

$$d(R,P) = \sum_{k=1}^{3} w_k \frac{\sum_{i_k=1}^{n_k} |x_k(i_k) - y_k(i_k)|}{n_k} \tag{3.8}$$

Onde:

R é a região em torno do bloco a ser simulado;

P é o protótipo;

k é o tipo de dado condicionante;
w é o peso atribuído a cada tipo de dado;
i é a localização do ponto na imagem;
x é o valor da propriedade no padrão;
y é o valor da propriedade no protótipo;
n é o número de pontos.

Os dados condicionantes podem ser de três tipos:

- valores rígidos, provenientes de medições feitas no próprio reservatório;
- valores estimados em iterações anteriores durante o processo de simulação;
- valores provenientes de partes de padrões associados aos blocos simulados anteriormente.

Uma vez identificado o protótipo mais próximo do padrão em torno do bloco, seleciona-se aleatoriamente um padrão pertencente à classe representada pelo protótipo e substitui-se a região do modelo pelo padrão selecionado. Esse procedimento é repetido até que todos os blocos do reservatório sejam simulados.

A distância calculada segundo a Equação 3.8 é mais adequada para os casos em que o modelo a ser simulado tem apenas duas dimensões. Para modelos com três dimensões o custo computacional aumenta consideravelmente, principalmente em casos em que a imagem de treinamento é mais complexa. Essa limitação é contornada em [46], onde é proposta uma nova maneira de calcular essa distância, que reduz significativamente o seu custo computacional. Na Figura 3.14 é apresentado um fluxograma que resume os passos da etapa de simulação do algoritmo FILTERSIM.

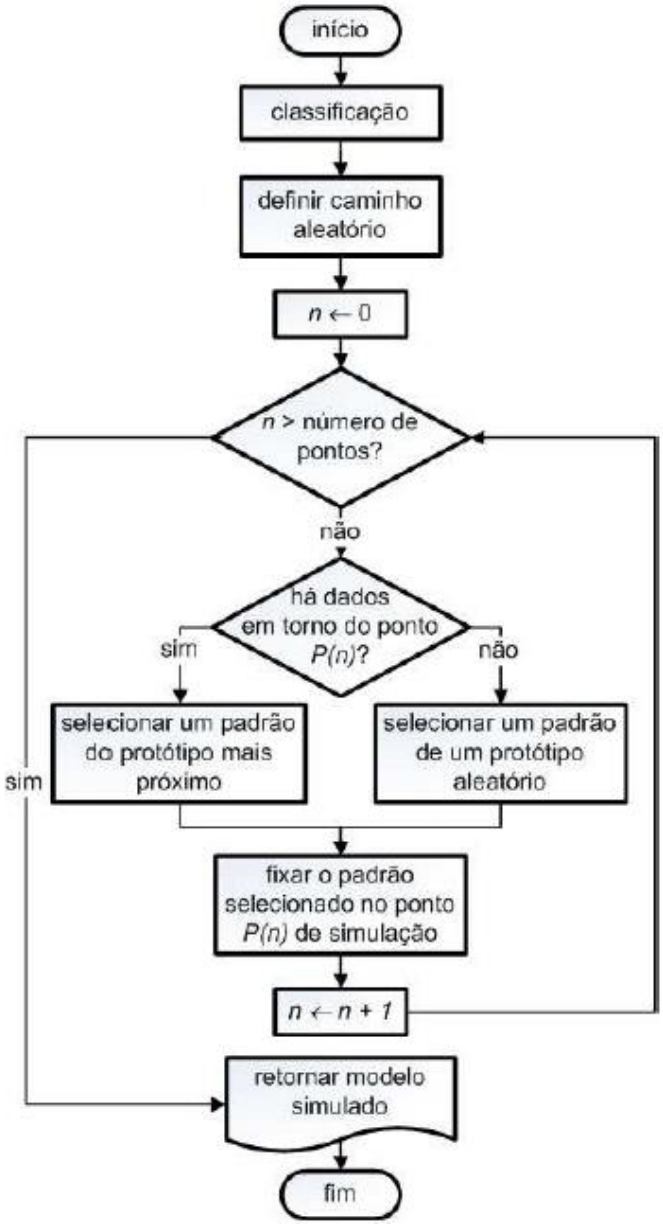

Figura 3.14: Fluxograma da etapa de simulação do algoritmo FILTERSIM (Fonte: Silva, 2011)