

Referências Bibliográficas

- [1] ROSA, A. J.; CARVALHO, R. d. S.; XAVIER, J. A. D. *Engenharia de Reservatórios de Petróleo*. Rio de Janeiro: Interciência, 2006. 1.1, 2.2.1, 2.2.2, 2.2.2, 2.4
- [2] TAVASSOLI, Z.; CARTER, J. N.; KING, P. R. Errors in History Matching. *SPE Journal*, v. 9, n. 3, p. 352–361, September 2004. 1.3, 3.1
- [3] THOMAS, J. E. *Fundamentos de Engenharia de Petróleo*. 2ª edição. ed. Rio de Janeiro: Interciência, 2004. 2.1, 2.2, 2.2.1, 2.2.2, 2.2.2, 2.3, 2.4
- [4] VELEZ-LANGS, O. Genetic Algorithms in Oil Industry: An Overview. *Journal of Petroleum Science and Engineering*, v. 47, n. 1-2, p. 15–22, 2005. 2.3, 3.2.2
- [5] ROSA, A. J.; CARVALHO, R. S. *Previsão de Comportamento de Reservatórios de Petróleo: Métodos Analíticos*. Rio de Janeiro: Interciência, 2002. 2.4
- [6] SCHULZE-RIEGERT, R. W. et al. Evolutionary Algorithms Applied to History Matching of Complex Reservoirs. *SPE Reservoir Evaluation & Engineering*, v. 5, n. 2, p. 163–173, April 2002. 3.1
- [7] SUBBEY, S.; CHRISTIE, M.; SAMBRIDGE, M. Prediction under Uncertainty in Reservoir Modeling. *Journal of Petroleum Science and Engineering*, v. 44, n. 1-2, p. 143–153, 2004. 3.1
- [8] SCHIOZER, D. J.; SOUSA, S. H. G.; MASCHIO, C. Ajuste de histórico de produção assistido. *Boletim Técnico da Produção de Petróleo*, Rio de Janeiro, v. 3, n. 1, p. 63–87, 2008. 3.1, 3.1.1, 3.1.3, 3.2
- [9] SANTOS, J. P. M.; SCHIOZER, D. J. Determinação de Metodologia de Ajuste Automatizado de Histórico. In: *Rio Oil & Gas Expo and Conference*. Rio de Janeiro, Brasil: [s.n.], 2000. 3.1.1, 3.1.2
- [10] SCHULZE-RIEGERT, R.; GHEDAN, S. Modern Techniques for History Matching. In: *9th International Forum on Reservoir Simulation*. Abu Dhabi, United Arab Emirates: [s.n.], 2007. 3.1.3
- [11] NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. Second edition. [S.l.]: Springer, 2000. 3.2

- [12] LUKE, S. *Essentials of Metaheuristics*. [S.l.]: Lulu, 2009. Available at <http://cs.gmu.edu/~sean/book/metaheuristics/>. 3.2
- [13] HORST, R.; TUY, H. *Global Optimization: Deterministic Approaches*. 3rd revised and enlarged edition. ed. [S.l.]: Springer, 2003. 3.2
- [14] JACQUARD, P.; JAIN, C. Permeability Distribution From Field Pressure Data. *Society of Petroleum Engineers Journal*, Houston, v. 5, n. 4, p. 281–294, December 1965. 3.2.1, 3.2.4, 7.1
- [15] BISSELL, R. C.; SHARMA, Y.; KILLOUGH, J. E. History matching using the method of gradients: Two case studies. In: *SPE Annual Technical Conference and Exhibition*. New Orleans, Los Angeles: [s.n.], 1994. 3.2.1
- [16] GRUSSAUTE, T.; GOUEL, P. Computer Aided History Matching of a Real Field Case. In: *European Petroleum Conference*. The Netherlands: [s.n.], 1998. 3.2.1
- [17] BRUN, B.; GOSELIN, O.; BARKER, J. W. Use of prior information in gradient-based history-matching. In: *SPE Reservoir Simulation Symposium*. Houston, Texas: [s.n.], 2001. 3.2.1
- [18] OLDENZIEL, T.; DITHUIJZEN, R. van; KRUIJSDIJK, C. van. Geologic Parameterization of Reservoir Model Enhances History-Match Procedure. *The Leading Edge*, Society of Exploration Geophysicists, v. 21, n. 6, p. 544–551, 2002. 3.2.1
- [19] RODRIGUES, J. R. P. Calculating derivatives for history matching in reservoir simulators. In: *SPE Reservoir Simulation Symposium*. Houston, Texas: [s.n.], 2005. 3.2.1
- [20] RODRIGUES, J. R. P. Calculating derivatives for automatic history matching. *Computational Geosciences*, v. 10, n. 1, p. 119–136, 2006. 3.2.1
- [21] LEITÃO, H. C.; SCHIOZER, D. J. Ajuste de Histórico Automatizado através de Otimização Multivariada e Paralelização Externa. In: *Rio Oil & Gas Conference*. Rio de Janeiro, Brasil: [s.n.], 1998. 3.2.2
- [22] LEITÃO, H. C.; SCHIOZER, D. J. A New Automated History Matching Algorithm Improved by Parallel Computing. In: *SPE Latin American and Caribbean Petroleum Engineering Conference*. Caracas, Venezuela: [s.n.], 1999. 3.2.2
- [23] MASCHIO, C.; SCHIOZER, D. J. Ajuste de Histórico Assistido Usando Métodos de Otimização de Busca Direta. In: *Rio Oil & Gas Expo and Conference*. Rio de Janeiro, Brasil: [s.n.], 2004. 3.2.2

- [24] MASCHIO, C.; SCHIOZER, D. J. Proposta de Técnicas Heurísticas para Ajuste de Histórico de Produção. In: *XXV Iberian Latin-American Congress on Computational Methods in Engineering*. Recife, Brasil: [s.n.], 2004. 3.2.2
- [25] SOLENG, H. H. Oil reservoir production forecasting with uncertainty estimation using genetic algorithms. In: *Proceedings of the Congress on Evolutionary Computation*. [S.l.: s.n.], 1999. v. 2, p. 1217–1223. 3.2.2
- [26] ROMERO, C. E. et al. Improved Reservoir Characterization through Evolutionary Computation. In: *SPE Annual Technical Conference and Exhibition*. Dallas, Texas: [s.n.], 2000. 3.2.2
- [27] ROMERO, C. E. et al. A Modified Genetic Algorithm for Reservoir Characterisation. In: *SPE International Oil and Gas Conference and Exhibition*. Beijing, China: [s.n.], 2000. 3.2.2
- [28] ROMERO, C. E.; CARTER, J. N. Using Genetic Algorithms for Reservoir Characterisation. *Journal of Petroleum Science and Engineering*, Elsevier, v. 31, n. 2-4, p. 113–123, 2001. 3.2.2
- [29] MASCHIO, C.; VIDAL, A. C.; SCHIOZER, D. J. A Framework to Integrate History Matching and Geostatistical Modeling Using Genetic Algorithm and Direct Search Methods. *Journal of Petroleum Science and Engineering*, Elsevier, v. 63, n. 1-4, p. 34–42, 2008. 3.2.2
- [30] AL-THUWAINI, J. S.; ZANGL, G.; PHELPS, R. Innovative Approach to Assist history Matching using Artificial Intelligence. In: *Intelligent Energy Conference and Exhibition*. Amsterdam, The Netherlands: [s.n.], 2006. 3.2.3
- [31] RAMGULAM, A.; ERTEKIN, T.; FLEMINGS, P. B. Utilization of Artificial Neural Networks in the Optimization of History Matching. In: *SPE Latin American and Caribbean Petroleum Engineering Conference*. Buenos Aires, Argentina: [s.n.], 2007. 3.2.3
- [32] MOTA, R. O. et al. Utilização de Simulação de Linhas de Fluxo para Ajuste de Histórico do Campo de Barracuda (UN-Rio). In: *SRR2007: Seminário de Reservas e Reservatórios, Petrobras*. Angra dos Reis, Rio de Janeiro, Brasil: [s.n.], 2007. 3.2.4, 7.2
- [33] LISBOA, E. F. A.; CHAVES, R. A. P.; MEZZOMO, C. C. Ajuste Assistido de Histórico do Bloco RJS-360 do Campo de Albacora Leste com o Auxílio de Simulação por Linhas de Fluxo. In: *SRR2007: Seminário de Reservas e Reservatórios, Petrobras*. Angra dos Reis, Rio de Janeiro, Brasil: [s.n.], 2007. 3.2.4, 7.2

- [34] CAERS, J. Geostatistical History Matching Under Training-Image Based Geological Model Constraints. In: *Annual Technical Conference and Exhibition*. San Antonio, Texas, USA: [s.n.], 2002. 3.2.5
- [35] CAERS, J. History matching under training-image based geological model constraints. *SPE Journal*, v. 8, n. 3, p. 218–226, September 2003. 3.2.5
- [36] HOFFMAN, B. T.; CAERS, J. Geostatistical history matching using a regional probability perturbation method. In: *SPE Annual Technical Conference and Exhibition*. Denver, Colorado: [s.n.], 2003. 3.2.5
- [37] HOFFMAN, B. T.; CAERS, J. History Matching with the Probability Perturbation Method: Application to a North Sea Reservoir. In: *IX European Conference on the Mathematics of Oil Recovery*. Cannes, France: [s.n.], 2004. 3.2.5
- [38] HOFFMAN, B. T.; CAERS, J. Regional probability perturbations for history matching. *Journal of Petroleum Science and Engineering*, v. 46, n. 1-2, p. 53–71, 2005. 3.2.5
- [39] ROGGERO, F.; HU, L. Y. Gradual deformation of continuous geostatistical models for history matching. In: *SPE Annual Technical Conference and Exhibition*. New Orleans, Louisiana: [s.n.], 1998. 3.2.5
- [40] CAERS, J. Efficient Gradual Deformation Using a Streamline-Based Proxy Method. *Journal of Petroleum Science and Engineering*, Elsevier, v. 39, n. 1-2, p. 57–83, 2003. 3.2.5, 7.2
- [41] JANSEN, J. D. et al. Closed-loop reservoir management. *First Break*, v. 23, p. 43–48, January 2005. 3.2.6
- [42] AITOKHUEHI, I.; DURLOFSKY, L. J. Optimizing the Performance of Smart Wells in Complex Reservoirs Using Continuously Updated Geological Models. *Journal of Petroleum Science and Engineering*, Elsevier, v. 48, n. 3-4, p. 254–264, 2005. 3.2.6
- [43] CHAMBERS, R. L.; YARUS, J. M.; HIRD, K. B. Petroleum Geostatistics for Nongeostatisticians (Part 1). *The Leading Edge*, Society of Exploration Geophysicists, v. 19, p. 474–479, 2000. 4.1
- [44] KRIGE, D. G. A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society*, v. 52, p. 119–139, 1951. 4.1
- [45] MATHERON, G. *Traité de Géostatistique Appliquée*. Paris: Editions Technip, 1962. 4.1

- [46] DOYEN, P. M. *Seismic Reservoir Characterization: An Earth Modelling Perspective*. Education tour series. [S.l.]: EAGE - European Association of Geoscientists & Engineers, 2007. 4.1, 4.1.1
- [47] LANDIM, P. M. B. *Análise Estatística de Dados Geológicos*. 2ª edição revista e ampliada. ed. São Paulo: UNESP, 2003. 4.1.1, 4.1.1, 4.1.2
- [48] CHILES, J. P.; DELFINER, P. *Geostatistics: Modeling Spatial Uncertainty*. USA: John Wiley & Sons, 1999. 4.1.1
- [49] ISAAKS, E. H.; SRIVASTAVA, R. M. *An Introduction to Applied Geostatistics*. New York, USA: Oxford University, 1989. 4.1.1
- [50] CHAMBERS, R. L.; YARUS, J. M.; HIRD, K. B. Petroleum Geostatistics for Nongeostatisticians (Part 2). *The Leading Edge*, Society of Exploration Geophysicists, v. 19, p. 592–599, 2000. 4.1.2
- [51] STREBELLE, S. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology*, Springer, v. 34, n. 1, p. 1–21, 2002. 4.1.3, 4.1.3
- [52] GUARDIANO, F.; SRIVASTAVA, R. M. Multivariate Geostatistics: Beyond Bivariate Moments. *Geostatistics-Troia*, v. 1, p. 133–144, 1993. 4.1.3, 4.1.3
- [53] SWITZER, P.; ZHANG, T.; JOURNEL, A. Simulating Geological Structures Based on Training Images and Pattern Classifications. Acessado em dezembro de 2008. Disponível em: <www.spatial-accuracy.org/system/files/Switzer2004accuracy.pdf>. 4.1.3, 4.1.3
- [54] LIU, Y. Using the Snesim Program for Multiple-Point Statistical Simulation. *Computers & Geosciences*, Elsevier, v. 32, n. 10, p. 1544–1563, 2006. 4.1.3
- [55] ZHANG, T. *Filter-Based Training Pattern Classification for Spatial Pattern Simulation*. Tese (Doutorado) — Stanford University, 2006. Ph.D. Thesis. 4.1.3
- [56] ZHANG, T.; SWITZER, P.; JOURNEL, A. G. Filter-Based Classification of Training Image Patterns for Spatial Simulation. *Mathematical Geology*, Springer, v. 38, n. 1, p. 63–80, 2006. 4.1.3
- [57] WU, J.; ZHANG, T.; JOURNEL, A. Fast filtersim simulation with score-based distance. *Mathematical Geosciences*, v. 40, n. 7, p. 773–788, 2008. 4.1.3
- [58] ZURADA, J. M. *Introduction to Artificial Neural Systems*. [S.l.]: West Publishing Company, 1992. 4.2

- [59] HAYKIN, S. *Redes Neurais: Princípios e Prática*. 2ª edição. ed. Porto Alegre: Bookman, 2001. 4.2
- [60] BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. *Redes Neurais Artificiais - Teoria e Aplicações*. [S.l.]: LTC, 2007. 4.2
- [61] JANG, J. R.; SUN, C.; MIZUTANI, E. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. [S.l.]: Prentice Hall, 1997. 4.2
- [62] YEN, J.; LANGARI, R. *Fuzzy Logic - Intelligence, Control and Information*. Upper Saddle River, New Jersey: Prentice Hall, 1999. 4.2
- [63] MICHALEWICZ, Z. *Genetic Algorithm + Data Structures = Evolution Programs*. Germany: Springer-Verlag, 1996. 4.2
- [64] LINDEN, R. *Algoritmos Genéticos - Uma Importante Ferramenta da Inteligência Computacional*. [S.l.]: Brasport, 2008. 4.2
- [65] ZADEH, L. A. Fuzzy sets. *Information and Control*, v. 8, p. 338–353, 1965. 4.2.2
- [66] ZADEH, L. A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-2, p. 28–44, 1973. 4.2.2
- [67] ZADEH, L. A. Fuzzy logic and approximate reasoning. *Synthese*, v. 30, p. 407–428, 1975. 4.2.2
- [68] ZADEH, L. A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, vol. 1, p. 3–28, 1978. 4.2.2
- [69] HOLAND, J. H. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan, 1975. 4.2.3
- [70] ALMEIDA, L. F. *Otimização de Alternativas para Desenvolvimento de Campo de Petróleo utilizando Computação Evolucionária*. Dissertação (Mestrado) — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2003. 5.7
- [71] TÚPAC, Y. J. V. *Sistema Inteligente de Otimização de Alternativas de Desenvolvimento de Campos Petrolíferos*. Tese (Doutorado) — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2005. 5.7

- [72] EMERICK, A. A. et al. Well Placement Optimization Using a Genetic Algorithm With Nonlinear Constraints. In: *SPE Reservoir Simulation Symposium*. The Woodlands, Texas, USA: [s.n.], 2009. 5.7
- [73] ICA. *GACOM - Guia de Referência*. jan. 2011. Disponível em: <<http://www.ica.ele.puc-rio.br/Downloads/215/Documentam.pdf>>. 5.7.1
- [74] CMG. *IMEX - Advanced Oil/Gas Reservoir Simulator - Version 2007 - User's Guide*. Calgary, Alberta, Canada: Computer Modelling Group LTD., 2007. 5.7.2
- [75] REMY, N.; BOUCHER, A.; WU, J. *Applied Geostatistics with SGeMS: A User's Guide*. 2ª edição. ed. [S.l.]: Cambridge University, 2009. 5.7.3, A.3
- [76] WU, J.; BOUCHER, A.; ZHANG, T. A SGeMS Code for Pattern Simulation of Continuous and Categorical Variables: FILTERSIM. *Computers & Geosciences*, Elsevier, v. 34, n. 12, p. 1863–1876, 2008. 5.7.3
- [77] FLORIS, F. J. T. et al. Methods for quantifying the uncertainty of production forecasts: A comparative study. *Petroleum Geoscience*, v. 7, p. S87–S96, 2001. 6.2
- [78] EGBERTS, P. J. P.; BROUWER, G. K.; BOS, C. F. M. History Matching and Forecasting with Uncertainty Quantification; A Real Case Study. In: *EAGE 64th Conference & Exhibition*. Florence, Italy: [s.n.], 2002. 6.2
- [79] SCIENCE, D. of E.; LONDON, E. I. C. *PUNQ-S3 Test Case*. February 2011. Disponível em: <<http://www3.imperial.ac.uk/earthscienceandengineering/research/perm/punq-s3model>>. 6.2
- [80] JACQUARD, P. Théorie de L'Interprétation des Mesures de Pression. *Revue de L'Institut Français du Pétrole*, XIX, n. 3, p. 297–334, 1964. 7.1
- [81] ICA. *ANEPI-CI: Análise Econômica de Projetos de E&P sob Incerteza de Campos Inteligentes*. [S.l.], 02 2007. 7.2
- [82] APACHE-HADOOP. *MapReduce*. February 2011. Disponível em: <<http://hadoop.apache.org/mapreduce/>>. 7.2
- [83] CLUSTER-RESOURCES. *Torque Resource Manager*. February 2011. Disponível em: <<http://www.clusterresources.com/products/torque-resource-manager.php>>. 7.2
- [84] MICROSOFT. *Windows HPC Server 2008*. February 2011. Disponível em: <<http://www.microsoft.com/hpc/en/us/default.aspx>>. 7.2

- [85] PLATFORM-COMPUTING. *Platform LSF*. February 2011. Disponível em: <<http://www.platform.com/workload-management/high-performance-computing>>. 7.2
- [86] CRUZ, A. A.; VELLASCO, M.; PACHECO, M. A. Quantum-Inspired Evolutionary Algorithm for Numerical Optimization. In: *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*. [S.l.]: IEEE Press, 2006. p. 2630–2637. 7.2
- [87] MESSER, B. *Projeto de Poços Multilaterais em Reservatórios de Petróleo Otimizados por Algoritmos Genéticos*. Dissertação (Mestrado) — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2009. 7.3
- [88] ALMEIDA, L. F. *Sistema Híbrido de Otimização de Estratégias de Controle de Válvulas de Poços Inteligentes sob Incertezas*. Tese (Doutorado) — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, 2007. 7.3
- [89] TAN, P. N.; STEINBACH, M.; KUMAR, V. *Introdução ao Data Mining - Mineração de Dados*. Rio de Janeiro: Editora Ciência Moderna, 2009. B.1

A Especificação dos Arquivos do SGeMS

A.1 Arquivo *.gslib* para set point

Os arquivos *.gslib* utilizados para a criação de elementos do tipo *set point*, apresentam um *layout* semelhante ao do exemplo a seguir:

```
modelo_1
4
x
y
z
field_property
0 0 0 972.237
1 0 0 944.515
2 0 0 873.251
:
0 0 0 -9966699
1 0 0 583.972
2 0 0 651.773
:
```

Onde:

- *modelo_1* é um título qualquer que, apesar de constar no arquivo, é desprezado pelo SGeMS;
- o valor 4 determina o número de colunas de dados do arquivo;
- *x*, *y*, *z* e *field_property* são os identificadores das colunas de dados;
- os valores atribuídos a *x*, *y* e *z* identificam as coordenadas de um bloco do modelo do reservatório;
- o valor atribuído a *field_property* é o valor da propriedade de interesse no bloco identificado pelas coordenadas *x*, *y* e *z*;

- o valor *-9966699* atribuído a *field_property* anula a propriedade no bloco identificado por *x*, *y* e *z*.

Para a criação de um elemento do tipo *set point* em que todos os blocos sejam nulos, basta que os valores de *field_property* no arquivo *.gslib* sejam iguais a *-9966699* em todos os blocos.

A.2

Arquivo *.gslib* para cartesian grid

Diferentemente do arquivo *.gslib* utilizado na criação de elementos do tipo *set point*, o arquivo *.gslib* para *cartesian grid* apresenta apenas a coluna de dados *field_property*, conforme ilustra o exemplo a seguir:

```

modelo_1
1
field_property
972.237
944.515
732.792
⋮

```

A.3

Script Python para o FILTERSIM

A interação do sistema *SGeMS* com outras aplicações é possível devido à existência de uma biblioteca que contém as funções necessárias, tanto para a importação e exportação de dados de e para o *SGeMS*, quanto para a execução de quaisquer dos métodos geoestatísticos existentes em seu ambiente. Essa biblioteca pode ser importada por um *script* escrito em linguagem *Python* (extensão *.py*) e esse *script* pode ser executado por outra aplicação, sem que haja a necessidade de interação direta com a interface do *SGeMS*. O *layout* do *script* é semelhante ao do exemplo a seguir:

```

01 import sgems
02
03 # carrega o projeto
04 sgems.execute('LoadProject ../SGeMS/campo_xadrez.prj/')
05
06 # abre o arquivo gslib com a propriedade a ser simulada
07 file = open('../Property_0/property.i.gslib', 'r')

```

```

08
09 # retira o cabeçalho do arquivo gslib
10 aux1 = range(7)
11 line = ''
12 for i in aux1:
13     line = file.readline()
14
15 # armazena os dados do arquivo gslib no vetor values
16 aux2 = ''
17 values = []
18 while line != '':
19     aux2 = line.split()
20     values += [float(aux2[3])]
21     line = file.readline()
22
23 # fecha o arquivo gslib
24 file.close
25
26 # cria o conjunto de pontos da propriedade com os valores do vetor values
27 sgems.set_property('conditional_data', 'field_property', values)
28
29 # executa o algoritmo FILTERSIM
30 sgems.execute('RunGeostatAlgorithm
filtersim_cont::/GeostatParamUtils/XML::<parameters>
<algorithm name="filtersim_cont" /> <GridSelector_Sim value="grid" />
<Property_Name_Sim value="field_property" />
<Nb_Realizations value="1" /> <Seed value="241075" />
<PropertySelector_Training grid="training_image" property="field_property" />
<Patch_Template_ADVANCED value="7 7 3" />
<Scan_Template value="9 9 5" /> <Trans.Result value="0" />
<Hard_Data grid="conditional_data" property="field_property" />
<Use_SoftField value="0" /> <Soft_Data value="" />
<Region_Indicator_Prop value="" /> <Active_Region_Code value="" />
<Use_Previous_Simulation value="0" />
<Previous_Simulation_Prop value="" /> <Use_Region value="0" />
<Nb_Multigrids_ADVANCED value="3" /> <Debug_Level value="0" />
<Cmin_Replicates value="10 10 10" /> <Data_Weights value="0.5 0.3 0.2" />
<CrossPartition value="1" /> <KMeanPartition value="0" />
<Nb_Bins_ADVANCED value="4" /> <Nb_Bins_ADVANCED2 value="2" />

```

```
<Use_Normal_Dist value="1" /> <Use_Score_Dist value="0" />
<Filter_Default value="1" /> <Filter_User_Define value="0" />
<User_Def_Filter_File value="Data File with Filter Definition" />
</parameters>')
```

31

32 # salva a propriedade simulada em um arquivo *gslib*

33 *sgems.execute('SaveGeostatGrid*
grid:.../Property_0/property_f.gslib::gslib::0::field_property_real0')

Onde:

- na linha 01 é incluída a biblioteca que fornece os comandos que permitem a interação com o *SGeMS*;
- na linha 04 é carregado o projeto criado no *SGeMS*, antes da execução do *OCTOPUS 2.0*;
- na linha 07 é aberto o arquivo *.gslib* (do tipo *set point*) que pertence ao indivíduo que deve ser reconstruído pelo *FILTERSIM*;
- entre as linhas 10 e 13 são retiradas as 7 primeiras linhas do arquivo *.gslib*, ou seja, o título, o número que identifica a quantidade de colunas de dados e os identificadores das colunas;
- entre as linhas 16 e 21 os dados lidos a partir do arquivo *.gslib* são armazenados em um vetor;
- na linha 24 o arquivo *.gslib* é fechado;
- na linha 27 os dados condicionantes que, inicialmente são todos nulos, são substituídos pelos valores armazenados no vetor;
- na linha 30 são configurados todos os parâmetros para a execução do *FILTERSIM*, dentre os quais se destacam:
 - *Seed* é a semente de geração dos números aleatórios para a simulação sequencial e deve ser alterada a cada execução do *FILTERSIM*;
 - *Patch_Template_ADVANCED* determina as dimensões da seção da imagem de treinamento, cujo ponto central deve ser fixado no ponto a ser estimado;
 - *Scan_Template* determina as dimensões do padrão utilizado para percorrer a imagem de treinamento e, essas dimensões, devem ser maiores ou iguais às da seção da imagem a ser fixada;
 - os demais parâmetros estão detalhados em [75];
- na linha 33 o resultado da simulação com o *FILTERSIM* é gravado em um arquivo *.gslib* (do tipo *cartesian grid*) que é usado para reconstruir o indivíduo.

B

Algoritmo de Lloyd

B.1

Análise de Grupos

A análise de grupos consiste na tarefa de formar grupos de objetos com base apenas nos dados que descrevem os próprios objetos. Especificamente, o que se busca é a construção de grupos em que os objetos pertencentes a cada grupo sejam semelhantes entre si e, ao mesmo tempo, diferentes dos objetos pertencentes aos outros grupos. Segundo [89], um conjunto inteiro de grupos é chamado de agrupamento e, quanto maior a semelhança (ou homogeneidade) dentro de um grupo e maior a diferença entre os grupos, melhor ou mais distinto é o agrupamento. Cada grupo é identificado por um elemento denominado centroide, que é utilizado como referência para a associação de novos objetos ao grupo. Por isso, é fundamental que o centroide caracterize adequadamente o modo como os objetos estão distribuídos dentro de um grupo. Assim, um agrupamento pode ser definido como um conjunto de objetos, tal que cada um desses objetos esteja mais próximo do centroide que define o seu grupo, do que dos centroides de quaisquer outros grupos.

Na análise de grupos, um objeto é qualquer entidade da qual é possível extrair algum tipo de característica. Tomando como exemplo os objetos “laranja” e “melancia”, é possível observar diversas características que definem o padrão de comportamento desses objetos. Dentre elas destacam-se: peso, volume, comprimento da circunferência e cor. A partir dessas características é possível criar um agrupamento com dois grupos bastante distintos que reúnem laranjas e melancias. Isso é possível porque em um conjunto de laranjas, salvo raras exceções, todas apresentam valores semelhantes para as características citadas, ou seja, seguem o mesmo padrão. O mesmo acontece entre as melancias. Porém, entre laranjas e melancias essas características apresentam valores bem diferentes e, conseqüentemente, o padrão de comportamento é bastante distinto. Logo, o que se busca com a análise de grupos é o reconhecimento de padrões e, daqui em diante, um objeto a ser agrupado é simplesmente denominado com um padrão.

B.2**Descrição do Algoritmo**

O algoritmo de *Lloyd* é uma técnica que particiona o espaço dos padrões em k grupos devidamente identificados pelos seus respectivos centroides. Nesse caso, para um conjunto de m padrões P , onde cada padrão é representado por n características, o centroide é dado pela média de todos os padrões pertencentes ao grupo, conforme ilustra a Equação B-1.

$$c = \left(\frac{(p_{11} + p_{21} + \dots + p_{m1})}{m}, \dots, \frac{(p_{1n} + p_{2n} + \dots + p_{mn})}{m} \right) \quad (\text{B-1})$$

Onde:

c é o centroide do grupo;

p_{ij} é o valor da característica j do padrão i .

O critério para associar um novo padrão a um dos k grupos existentes é baseado na distância entre o novo padrão e os centroides de todos os k grupos. Calculadas todas as distâncias, o padrão é associado ao grupo cujo centroide é mais próximo. No caso do algoritmo de *Lloyd*, a distância entre dois padrões P_1 e P_2 é dada pela Equação B-2, que corresponde à distância euclidiana entre dois pontos.

$$d(P_1, P_2) = \sqrt{(p_{11} - p_{21})^2 + (p_{12} - p_{22})^2 + \dots + (p_{1n} - p_{2n})^2} \quad (\text{B-2})$$

Onde:

d é a distância entre os pontos P_1 e P_2 .

Inicialmente o algoritmo escolhe aleatoriamente k padrões, dentre todos a serem agrupados, e os utiliza como centroides. A partir daí, os padrões restantes são associados aos centroides de acordo com o critério de distância da Equação B-2. Após a associação de todos os padrões, os centroides de todos os grupos são recalculados segundo a Equação B-1. Os passos de associação dos padrões e recálculo dos centroides são repetidos por um determinado número de iterações, a fim de que a configuração ideal dos centroides seja encontrada. Ao longo dessas iterações, o algoritmo minimiza a função de erro dada pela Equação B-3. Essa minimização implica na formação de grupos mais compactos, ou seja, grupos em que as distâncias, entre os centroides e seus respectivos padrões, sejam mínimas.

$$SSE = \sum_{i=1}^k \sum_{P \in C_i} d(P, C_i)^2 \tag{B-3}$$

Onde:

SSE é a métrica de erro a ser minimizada;

k é o número de grupos;

P é o padrão;

C_i é o centroide do grupo i .

A Figura B.1 apresenta um fluxograma que resume a maneira como funciona o algoritmo de *Lloyd*. É importante observar que o número k de grupos a serem criados, deve ser informado como entrada para o algoritmo. Essa é uma limitação importante apresentada pelo algoritmo de *Lloyd*, uma vez que nem sempre é possível saber, de antemão, a quantidade ideal de grupos em que os padrões devem ser distribuídos.

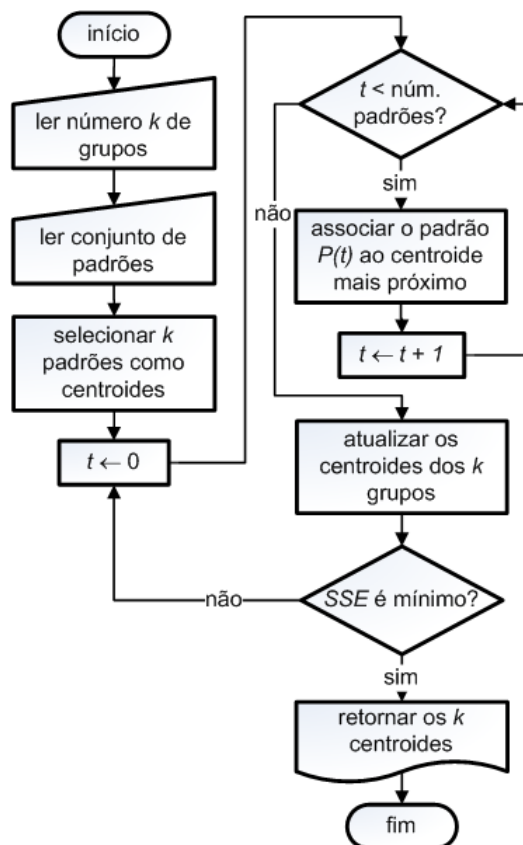


Figura B.1: Fluxograma do algoritmo de *Lloyd*