International Conference on Applied Informatics

ICAI 2020: **Applied Informatics** pp 281–295Cite as

The Application of DBSCAN Algorithm to Improve Variogram Estimation and Interpretation in Irregularly-Sampled Fields

- O. O. Mosobalaje,
- O. D. Orodu &
- D. Ogbe

- Conference paper
- First Online: 19 October 2020

Part of the Communications in Computer and Information Science book series (CCIS,volume 1277)

Abstract

The empirical variogram is a measure of spatial data correlation in geostatistical modeling and simulations. Typically, the empirical variogram is estimated for some defined lag intervals by applying method of moments on an underlying variogram cloud. Depending on the distribution of pair-wise lag values, the variogram cloud of an irregularly-sampled field may exhibit clusteredness. Issues of noisy, uninterpretable and inconsistent empirical variogram plots are commonly encountered in cases of irregularly-sampled fields with clustered variogram clouds. An insightful diagnosis of these problems and a practical solution are the subject of this paper. This research establishes the fact that these problems are caused by the neglect of variogram cloud cluster configurations when defining lag intervals. It is here shown that such neglect hinders the optimal use of spatial correlation information present in variogram clouds. Specifically, four sub-optimal effects are articulated in this paper as the consequence of the neglect.

Consequently, this research presents an efficient cluster-analysis – driven technique for variogram estimation in cases of irregularly-sampled fields with clustered variogram clouds. The cluster analysis required for this technique is implemented using an unsupervised machine learning algorithm known as Density-based Spatial Clustering of Applications with Noise (DBSCAN). This technique has been applied to a real field to obtain a stable, interpretable and geologically consistent variogram plot. It has also been applied to a synthetic field and was found to give the lowest estimation error among other techniques. This technique would find usefulness in geo-modeling of natural resource deposits wherein irregular sampling is prevalent.

Keywords

- **Cluster analysis**
- **DBSCAN**
- **Variogram cloud**
- **Lag tolerance**
- **Variogram estimation**
- **Irregular sampling**

- DOI: 10.1007/978-3-030-61702-8_20
- Chapter length: 15 pages
- Instant PDF download

- ISBN: 978-3-030-61702-8

- ISBN: 978-3-030-61701-1

Abbreviations

C_{ID}:
: Cluster Identifier

$\vec{h}$ h→ :

Lag vector, m

Δh **:**

Lag distance tolerance, m

***i*:**

Loop counter

***N*:**

Number of clusters

$Z(u_i)$ **:**

Value of a random variable attribute at location u_i

$Z\left(u_i+\vec{h}\right)$ **:**

Value of a random variable attribute at location $u_i+\vec{h}$

$\gamma(\vec{h})$ **:**

Variogram, squared unit of attribute

References

1. Li, M., Zhao, Y.: Geophysical Exploration Technology: Applications in Lithological and Stratigraphic Reservoirs. Elsevier, Amsterdam (2014)

   Google Scholar

2. Zuur, A., Ieno, E.N., Smith, G.M.: Analyzing Ecological Data. Springer, New York (2007)

   MATH Google Scholar

3. Kapageridis, I.K.: Variable lag variography using k-means clustering. Comput. Geosci. **85**, 49–63 (2015). https://doi.org/10.1016/j.cageo.2015.04.004

   CrossRef Google Scholar

4. Deutsch, C.V., Journel, A.: GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York (1998)

   Google Scholar

5. Deutsch, J.L.: Experimental variogram tolerance parameters (2015). http://www.geostatisticslessons.com/lessons/variogramparameters
6. Pebesma, E.J.: Multivariable geostatistics in S: the gstat package. Comput. Geosci. **30**, 683–691 (2004). https://doi.org/10.1016/j.cageo.2004.03.012

CrossRef Google Scholar

7. Yupeng, L., Miguel, C.: A flexible lag definition for experimental variogram calculation. Int. J. Min. Sci. Technol. **21**, 207–211 (2011)

Google Scholar

8. National program on technology enhanced learning (NPTEL). Unsupervised learning (2013). http://14.139.172.204/nptel/CSE/Web/102106023/ch9_Competitive%20Learning_3.pdf
9. Al-Fuqaha, A.: Clustering analysis (2014). https://cs.wmich.edu/alfuqaha/summer14/cs6530/lectures/ClusteringAnalysis.pdf
10. Saraswathi, S., Sheela, M.I.: A comparative study of various clustering algorithms in data mining. Int. J. Comput. Sci. Mob. Comput. **3**, 422–428 (2014)

Google Scholar

11. Hahsler, M., Piekenbrock, M.: DBSCAN: density based clustering of applications with noise (DBSCAN) and related algorithms. R package version 1.1-1 (2017). https://CRAN.R-project.org/package=dbscan
12. Hennig, C.: fpc: flexible procedures for clustering. R package version 2.1-11.1 (2018). https://CRAN.R-project.org/package=fpc
13. R Core Team R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna (2020). https://www.R-project.org/

14. Adegoke, O.S., et al.: Geotechnical investigations of the Ondo State bituminous sands – Volume I: geology and reserves estimate. Geological Consultancy Unit, Department of Geology, University of Ife, Nigeria (1980)

[Google Scholar]

15. Mosobalaje, O.O., Orodu, O.D., Ogbe, D.: Descriptive statistics and probability distributions of volumetric parameters of a Nigerian heavy oil and bitumen deposit. J. Pet. Explor. Prod. Technol. **9**(1), 645–661 (2018). https://doi.org/10.1007/s13202-018-0498-4

[CrossRef] [Google Scholar]

16. Falebita, D.E., Oyebanjo, O.M., Ajayi, T.R.: A geostatistical review of the bitumen reserves of the upper Cretaceous Afowo formation, Agbabu area, Ondo State, Eastern Dahomey basin, Nigeria. Pet. Coal **56**, 572–581 (2014)

[Google Scholar]

17. Mosobalaje, O.O., Orodu, O.D., Ogbe, D.: Estimating and modeling of spatial variability of volumetric attributes of a Nigerian heavy oil and bitumen deposit. In: 2019 SPE Nigerian Annual International Conference and Exhibition (2019)

[Google Scholar]

18. Obaje, N.G.: Geology and Mineral Resources of Nigeria. Lecture Notes in Earth Sciences, vol. 120. Springer, Berlin (2009). http://dx.doi.org/10.1007/978-3-540-92685-6_9
19. Deutsch, C.V.: Geostatistical Reservoir Modeling. Oxford University Press, New York (2002)

[Google Scholar]

Acknowledgements

Inspirations from the only wise God are acknowledged. Dr. A. A. Adepelumi facilitated access to previous report of the Geological Consultancy Unit of Obafemi Awolowo University. Publication support from Covenant University Centre for Research, Innovation and Discovery (CUCRID) is also appreciated.

Author information

Authors and Affiliations

1. **Department of Petroleum Engineering, Covenant University, Ota, Nigeria**
   O. O. Mosobalaje & O. D. Orodu
2. **African University of Science and Technology, Abuja, Nigeria**
   D. Ogbe

Corresponding author

Correspondence to O. O. Mosobalaje .

Editor information

Editors and Affiliations

1. **Universidad Distrital Francisco Jose de Caldas, Bogota, Colombia**
   Hector Florez
2. **Covenant University, Ota, Nigeria**
   Prof. Sanjay Misra

Rights and permissions

Copyright information

About this paper

Cite this paper

Mosobalaje, O.O., Orodu, O.D., Ogbe, D. (2020). The Application of DBSCAN Algorithm to Improve Variogram Estimation and Interpretation in Irregularly-Sampled Fields. In: Florez, H., Misra, S. (eds) Applied Informatics. ICAI 2020.

Communications in Computer and Information Science, vol 1277. Springer, Cham. https://doi.org/10.1007/978-3-030-61702-8_20

- DOIhttps://doi.org/10.1007/978-3-030-61702-8_20
- Published19 October 2020
- Publisher NameSpringer, Cham
- Print ISBN978-3-030-61701-1
- Online ISBN978-3-030-61702-8
- eBook PackagesComputer ScienceComputer Science (R0)