Ocean and Coastal Research

CASE REPORT

http://doi.org/10.1590/2675-2824070.21080lg
ISSN 2675-2824

Fuzzy Logic: adding natural uncertainties into environmental assessment

Leonardo Gripp[1*], Carlos German Massone[1], Renato da Silva Carreira[1], Angela de Luca Rebello Wagener[1]

[1] Pontifícia Universidade Católica do Rio de Janeiro - Departamento de Química (Rua Marquês de São Vicente, 225 Cx. Postal: 38097 - Gávea - Rio de Janeiro - 22451-900 - RJ - Brazil).

* Corresponding author: leogrippbom@gmail.com

ABSTRACT

This research study sought to evaluate aimed at evaluating the possible advantages of using Fuzzy Logic as opposed to Boolean Logic to assess environmental contamination and source appraisal for polycyclic aromatic hydrocarbons (PAH). Results obtained through traditional assessment tools for two different tropical coastal areas through using traditional clustering and principal components analysis were compared with those derived from the Fuzzy Logic, using the by Fuzzy C-means algorithm. The feedings achieved through Fuzzy Logic showed a greater qualitative detail than those derived from traditional tools. The abrupt and unnatural changes obtained from the usual classification methods were avoided by having membership values varying continuously in space, providing a more accurate picture of environmental contamination in complex and multiple sources environments. Furthermore, by not depending on statistic suppositions distribution of data like other methods, becomes more suitable for environmental data. Although Fuzzy Logic does not produce quantitative interpretations, its application generates adequate the data needed to avoid environmental management bias in the inference of contamination sources.

Keywords: Fuzzy C-means, Fuzzy logic, Environmental management, Polycyclic aromatic hydrocarbo.

INTRODUCTION

The Fuzzy Set Theory was proposed by Zadeh (1965) to insert the concept of uncertainty, treating natural phenomena or real situations naturally (Rocha et al., 2012). It can be defined as the part of mathematical logic dedicated to the formal principles of uncertain or approximate reasoning, whereas algebra, using Boolean logic, expresses results in binary form; the "maybe" condition is not possible (Cunha et al., 2011).

A fuzzy set or subset is a collection of ill-defined and indistinct objects (samples) with unsharp boundaries, in which transitions from membership to non-membership in a subset of a reference set are gradual rather than abrupt (Zadeh, 1965). In fuzzy set theory, an element may belong, with appropriate degrees, to different sets defined in the same universe of discourse (Rocha et al., 2012). The interpretation of fuzzy classification is based on membership values that indicate the participation of each sampling point to predefined classes. In classical cluster analysis (hard clustering), each object must be assigned to a single cluster. For hydrocarbons, this means that a sample pertaining to a cluster classified as a "petrogenic source" should not contain pyrolytic contribution. This is an important approach to data analysis, since most authors classify samples by their primary source contribution, refusing intrinsic

Submitted: 01-September-2021
Approved: 21-February-2022

Editor: Rubens M. Lopes

uncertainties (Wagener et al., 2019) . Classifying samples into mutually exclusive groups is a source of ambiguity and error in cases of outliers or overlapping clusters, which implies a loss of information (Sârbu and Einax, 2008).

Historically, there has been a tendency to classify samples by their predominant source, but this can limit the understanding of multiple source contamination in coastal environments, reasonable exposure sites for pollutants such as PAH compounds (Christensen et al., 2010; Baumard et al., 1998; Massone et al., 2013; Wagener et al., 2010; Wagener et al., 2011; Wagener et al., 2012; Yunker et al., 2002). Therefore, the theory of fuzzy sets is an interesting option to classify samples where the signals of different sources are superimposed in unknown proportions, and composition is constantly changing due to weathering (Wagener et al., 2019). Among the challenges in researching contamination evaluation is the complexity of the sources and processes involved in the environment. This issue is underlined by discrepancies observed when applying the so-called diagnostic ratios in PAH source appraisal (Dvorská et al., 2011; Massone et al., 2013) and by statistical methods to characterize the contaminant composition in environmental matrices (Christensen et al., 2010; Hopke, 2015).

Sârbu and Einax (2008) achieved unsatisfactory results applying classical clustering and principal component analysis to results from an extensive monitoring program concerning soil lead content, plant lead content, and traffic density at different sampling locations in East Germany. The authors showed the advantages in Fuzzy C-means (FCM), Gustafson–Kessel (GK), and fuzzy c-varieties (FCV) over traditional clustering methods. Similarly, Tan et al. (2006) applied fuzzy classification combined with spatial prediction to assess the state of soil pollution in the peri-urban Beijing area. The authors reported a prediction model with a quantitative uncertainty evaluation and higher reliability than conventional single geostatistical kriging methods. A similar conclusion was reported by Güler et al. (2012), when performing an impact assessment of anthropogenic activities on the groundwater.

Hu et al. (2016) demonstrated that in risk assessment, better decisions are made when uncertainty and variability are explicitly acknowledged in models, resulting in an effective tool for risk assessments and for managing contaminated sites. In their research, the uncertainties caused by the lack of uniform and scientifically supported environmental quality guidelines and by the variability in degree of exposure of environmental systems to contaminants were incorporated in a stochastic fuzzy environmental risk characterization approach. Since fuzzy classification can deal with abrupt transitions such as single pollution spots (Franssen et al., 1997), it becomes an alternative procedure for interpolating continuous soil pollution spatial data of high variability (Amini et al., 2005; Dobermann and Oberthür, 1997; Franssen et al., 1997; Odeh et al., 1992).

It is challenging for analytical chemists to incorporate uncertainty into their environmental assessments. Researchers attempt to minimize uncertainty to the extent possible to ensure reliable results. Fuzzy Logic is a counterpoint, though not to disqualify analytical effort. The uncertainty in environment assessment associated with non-controllable parameters by analytical chemist is intrinsically much higher than that of controllable ones. Several studies have attempted to remove or reduce this uncertainty in environment assessment, be it through PCA, PMF, hierarchical clustering, or other statistical tools (Christensen et al., 2010; Hopke, 2015; Stanimirova et al., 2011; Yan et al., 2014). This article seeks to corroborate the benefits of incorporating uncertainty into data assessment and the advantages that can result from this approach to unravel the sources of anthropic hydrocarbons to coastal systems.

METHODS

Fuzzy Logic application is intended to improve understanding of different source contributions compared to traditional tools. To achieve this, data from solid published articles about contamination in sediments from tropical estuaries (Massone et al., 2013; Wagener et al., 2010) were reanalyzed using Fuzzy Logic (Figure 1). Several algorithms for Fuzzy Logic application can be found in literature. The choice of algorithm and its application to data is a subject of wide discussion and is not the focus of this article. All analyses herein were performed using the C-means algorithm, known as FCM.

One of the fundamental issues associated with the applications of fuzzy set theory is the determination of membership functions

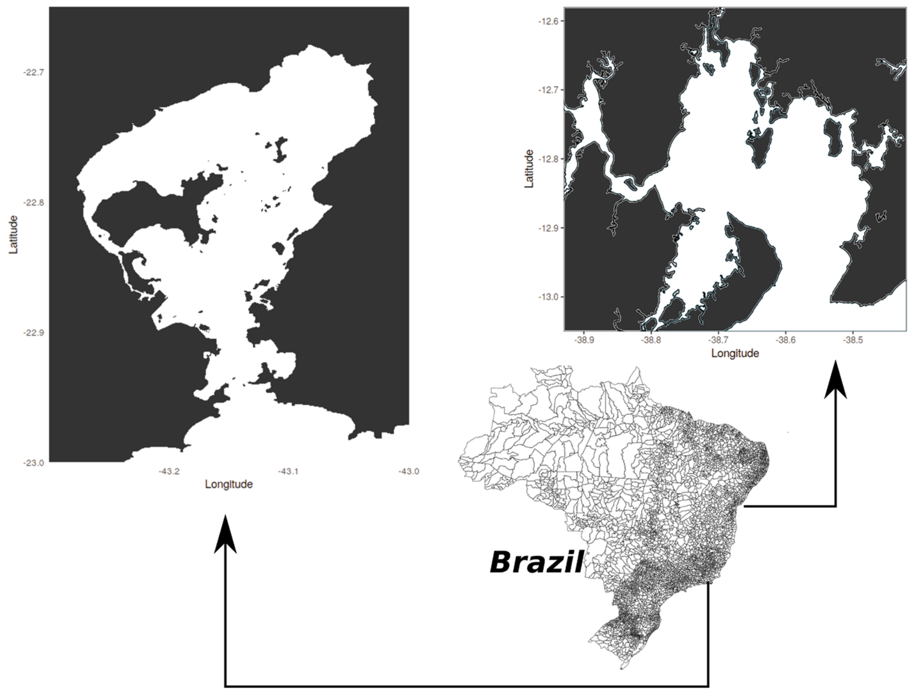

Figure 1. Location of the study areas selected for the present study - Guanabara Bay on the left and Todos os Santos Bay on the right.

(Krishnapuram, 1994). In brief, membership functions of the FCM algorithm (Bezdek et al., 1984) link each sample to all clusters via a real-valued vector of membership degrees, with values between 0 and 1. Values close to 1 indicate a strong association to the cluster, while those close to 0 indicate a lack of association.

Data processing before statistical analysis was maintained to avoid incorporating or minimizing variations beyond Fuzzy Logic. Therefore, the same data matrix used in the original publications (Guanabara Bay (Massone et al., 2013) and Todos os Santos Bay (Wagener et al., 2010)) were preserved. FCM and the feedings derived from the present appraisal were then compared to previous interpretations. This study did not change the number of clusters selected by the authors, as the comparison between different methods of data interpretation is more reliable when maintaining the same number of groups. Laboratory uncertainties between selected studies were similar and not significant for data evaluation, since natural uncertainties in the environmental assessment are much higher than analytical ones.

Data analysis, graphs, and maps were performed in R software (R Core Team, 2013) and related packages: rgdal' (Bivand et al., 2020), 'maptools' (Bivand and Lewin-Koh, 2018), 'maps' (Becker et al., 2018), 'tmap' (Tennekes, 2018), 'ggplot2' (Wickham, 2016) and 'ppclust' (Cebeci et al., 2017). The latter package was used to apply the FCM algorithm, in which cluster partitioning divides objects in a dataset into non-overlapping subsets, or clusters, using prototype-based probabilistic and possibilistic clustering algorithms (Cebeci et al., 2017). In brief, since hydrocarbon sources were the primary objective, concentration bias was removed from the samples prior to statistical analysis by normalization through the total PAH concentration followed by z-score. The number of previously defined clusters was further corroborated by the Simprof analysis, a tool for determining the number of significant clusters produced using the assumption of no a priori clusters.

ASSESSMENT AND DISCUSSION

Guanabara Bay

Massone et al. (2013) applied Principal Component Analysis (PCA) associated to multivariate linear regression (MLR) to attain a quantitative assessment of Policyclic Aromatic Hydrocarbons (PAH) sources in the Guanabara Bay. The authors confirmed the predominance of the pervasive contaminant component superimposed to a generalized petrogenic imprint. The PAH ratios and PCA-MRL analysis were strongly influenced by a common source component, in which 11 sediments spread over this tropical bay shared at least 56% of their relative PAH distribution.

In their research, the PCA feedings were confirmed by cluster analyses identifying three sectors (Figure 2) with distinct PAH source contributions to sediments: (1) an area principally contaminated by petrogenic residues including the west and north bay regions; (2) an area principally contaminated by combustion residues in the southeast region; and (3) an area where natural contributions are evident in the northeast region.

Despite the advanced data treatment carried out by the authors, the Guanabara Bay sediments were classified into regions according to a predominant source. in doing so, the complexity of the bay and of the hydrocarbon sources were not incorporated in the feedings. Guanabara Bay is one of the most important embayments of the Brazilian coast, and various forms of pollution threaten this estuarine environment (Soares-Gomes et al., 2010). The bay's surroundings house 16 oil terminals, 6000 industries, and two shipyards. The bay receives an estimated daily oil input of approximately 9.5 tons, among other pollutants (Francioni et al., 2005). This environment also features intricate hydrodynamics, characterized by the action of

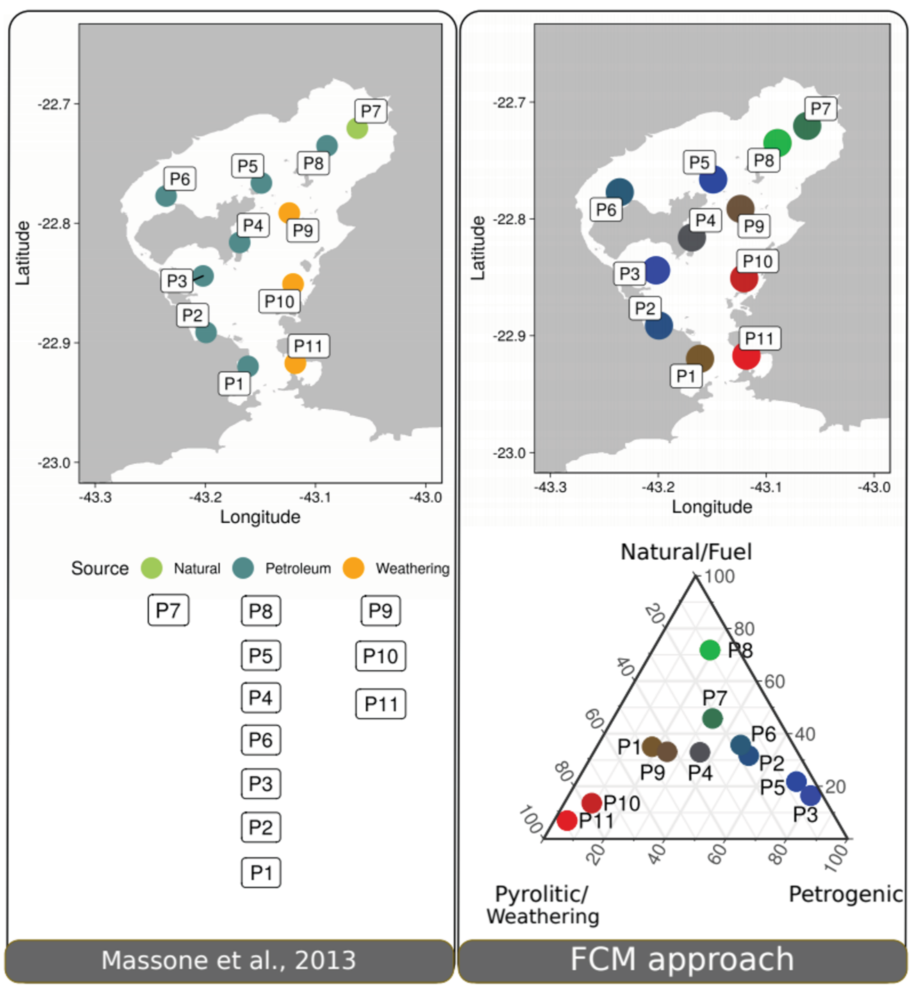

Figure 2. Representation of the degrees of membership of the samples in relation to groups Petrogenic, Natural/Fuel, and Pyrolytic/Weathering throughout Guanabara Bay and the distribution of PAH in the P03, P08, and P11 samples. Data from Massone et al. (2013).

a central channel (Kjerfve et al., 1997) and the efficiency of tidal currents. Therefore, given the complexity of this environment, segmenting the bay in areas of overriding sources does not make environmental sense and weakens management strategies.

Typically, diagnostic ratios have been employed to evaluate sources of PAH in aquatic systems (e.g., Budzinski et al., 1997; Wang et al., 1999; Garrigues et al., 1995; Yunker et al., 2002). However, the original source fingerprinting may be altered by preferential degradation rates among the PAH caused by factors such as temperature, photolysis, and microbiological activity, which is more pronounced in tropical environments. As such, the use of diagnostic ratios for source indication in these environmental conditions was proven to be unreliable (Massone et al., 2013; Wagener et al., 2011; Wagener et al., 2012; Wagener et al., 2010). Dvorská et al. (2011). For instance, suggested that the ratio between the five series of alkylated PAH over 3 to 6 ring PAH, known as the pyrolytic index (Wang et al., 1999), is insufficient for source appraisal and detailed description of mass balance issues not covered by the diagnostic ratios.

Many statistical methods have been applied targeting source discrimination and mass balance calculations in Guanabara Bay (Christensen et al., 2010; Massone et al., 2013; Wagener et al., 2010; Wagener et al., 2012; Wagener et al., 2019). Among them, the FCM approach has gained prominence in environmental sciences because the neglected uncertainty in traditional methods is incorporated into the model (Stanimirova et al., 2011).

The FCM approach applied to Guanabara Bay appears in Figure 2, which shows the degree of membership among groups (Petrogenic, Natural/Fuel, and Pyrolytic/Weathering), highlighting that no sample belongs to a single group. This feature illustrates that, throughout this tropical bay, the PAH sources overlap at different scales. It is important to note that FCM does not produce a quantitative assessment of contributions such as PCA-MLR and, therefore, the degrees of relevance shall be used only for qualitative purposes.

As for major source identification in general, there is consonance between the results reported by Massone et al. (2013) and those reported herein using FCM. The P3 station, pertaining most relevantly to the petrogenic group, has a higher proportion of low molecular weight and alkylated compounds, which are markers of petrogenic origin. PAH in sample P11, pertaining most significantly to the pyrolytic/weathering group, as the abundance of high molecular weight compounds indicates. At this station, there was a difference in the method's classification, and the FCM approach proved to be more consistent considering the sample PAH composition. A sample from the northeast region of Guanabara Bay (P8), a region less polluted and close to an environmental protection area, appear segregated from the others as Natural. This sample, as well as sample P7, has a higher perylene and light alkylated PAH relative contribution, respectively characterizing diagenesis and fuel. In this area, in addition to the Environmental Protection Area, a rail accident in 2005 spilled 60,000 liters of oil directly into rivers, reaching Guanabara Bay. This explains the position of sample P7 in the ternary diagram of Figure 2, as suggested by FCM (and not by the PCA analysis), i.e., a higher influence of anthropic sources of PAH in comparison to P8, reflecting the potential input of diesel from the spill.

The attained gain derived from the analysis by FCM is more relevant for the evaluation of samples permeating different groups of PAH sources. Samples that, according to previous conventional statistical analysis, belong to a single group, now display their complexity by crossing groups. The similarity among samples described by Massone et al. (2013) and not shown in their assessment for group classification is by FCM more prominent.

There is an outstanding overlap of contributions from different components in several samples: P1, P2, P4, P6, P7 and P9. The samples with the highest degrees of membership within the groups are those close to contamination sources. The evaluation by FCM opens new frontiers by raising important aspects in Guanabara Bay, such as the evidence of source overlapping, estuarine mixture, and the influence of the source distance.

Todos os Santos Bay (TSB)

Wagener et al. (2010) studied PAH contamination levels and source identification in sediments of the Todos os Santos Bay (TSB). TSB is the largest coastal bay in Brazil, and its drainage basin houses one of the first areas of inland petroleum exploration in the country as well as an important national petrochemical complex. Source evaluation was principally based on traditional diagnostic ratios (Baumard et al., 1998; Dickhut et al., 2000; Wang et al., 1999; Yunker et al., 2002). Wagener et al. (2010) points out that conditions in tropical regions accelerate degradation of the least persistent compounds and affect the efficiency of PAH diagnostic ratios. The authors also highlight the complexity of oils used in the basin areas as an additional limitation to reliable source evaluation based on diagnostic ratios.

Even though the occurrence of highly alkylated PAH homologous indicating the ubiquitous presence of weathered oil residues, the authors found a strong bias toward pyrolytic source diagnosis in TSB using diagnostic ratios, which masked the relevance of petrogenic contribution. This petrogenic source underestimation may be partly minimized by the FCM approach (Figure 3). The main advantage of this method over diagnostic ratios, whose source assessment is obtained through Boolean Logic (petrogenic or pyrolytic), is highlighted by the degree of membership regarding the pyrolytic to petrogenic sources, corroborated by the respective PAH profiles.

When comparing results from FCM (Figure 3) with those presented by diagnostic ratios (Figure 4), the Boolean Logic limitation becomes noticeable in a qualitative data analysis. Contribution of analytical and natural uncertainties considered by Fuzzy Logic allows for verifying the PAH sources gradient, through which the FCM allowed the determination of two groups ranging from high molecular weight (Pyrolytic) to low molecular weight (Petrogenic) predominance. Figure 3 presents the PAH distributions for more relevant samples pertaining to the predefined petrogenic and pyrolytic groups (I29 and I3, respectively) and two samples with intermediate values (I5 and I11).

There is a major relative contribution of pyrolytic PAH in the northeast/east portion of the TSB, between the mouth and Madre de Deus island. This region houses major pollution sources: an oil refinery, two harbors, an oil terminal, and the urban area of Salvador. Of the 29 sediment samples analyzed thus far, 14 were found in this area, of which 11 showed a greater degree of membership to pyrolytic group (for instance I3 and

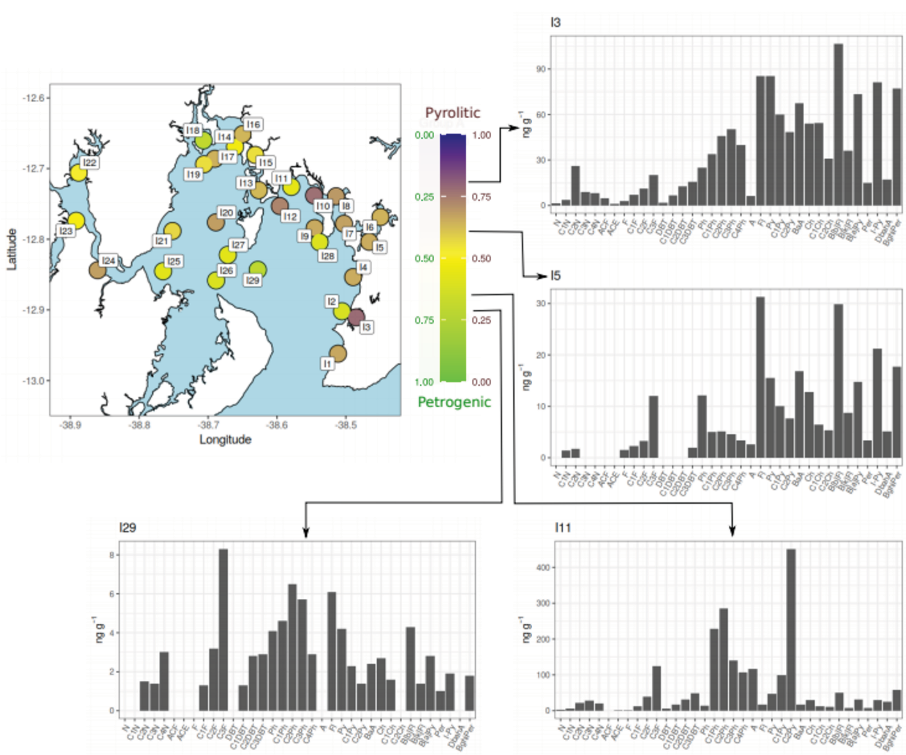

Figure 3. Representation of the degrees of membership of the samples in relation Petrogenic and Pyrolytic groups throughout Todos os Santos Bay and the distribution of PAH in I3, I5, I11, and I29 samples. Data from Wagener et al. (2010).

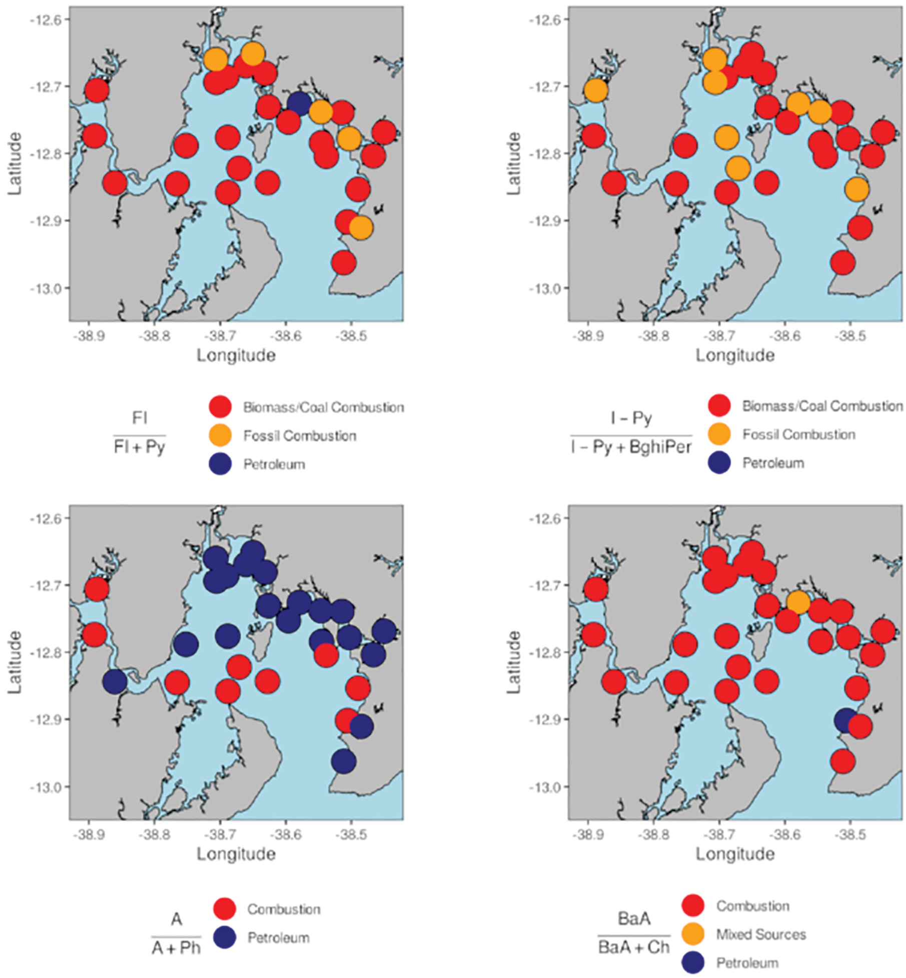

Figure 4. Diagnostic ratios applied to the Todos os Santos Bay samples. Data from Wagener et al. (2010).Ratios: Fl/(Fl+Py) - Fluoranthene (Fl) and Pyrene (Py); (I–Py)/(I-Py+BghiP) - Indeno(1,2,3-cd) pyrene (I-Py) and Benzo(ghi)perylene (BghiPer); A/A+Ph - Anthracene (A) and Phenanthrene (Ph); BaA/BaA+Ch - Benzo(a)anthracene (BaA) and Chrysene (Ch).

I5). It is plausible to also presume a considerable contribution of petrogenic PAH.

Sample I3 lies in the Itapagipe embayment, surrounded by a large urban area and encompassing a significant industrial center. Most diagnostic ratios classified PAH in sample I3 as being of pyrolytic source (Wagener et al., 2010) . In fact, the higher incidence of high molecular weight PAH suggests a predominance of pyrolytic compounds. However, the presence of alkylated naphthalenes and the maximum concentration in C2- or C3-phenanthrene in the phenanthrene series, suggests a recent introduction of oil residues (Tolosa et al., 2004; Varnosfaderany et al., 2015) in addition to pyrolytic PAH. The same conclusion can be drawn from the presence of dibenzothiophenes, which are good markers for diesel oil (MacKenzie and Hunter, 1979; Takada et al., 1991; Williams and Bottrill, 1995). Although sample I3 shows a high degree of membership in pyrolytic group, the FCM evaluation allows for considering the contribution of petrogenic sources, albeit highlighting the predominance of low molecular weight PAH.

For the I5 sample, situated in the Aratu embayment, the diagnostic ratios suggest strong similarities to the I3 sample, whereas in fact there is a greater contribution of high molecular weight PAH in this sample. However, using the Boolean approach, the significance of petrogenic input is not revealed. Aratu is surrounded by terminals, chemical industries, and harbors (Aratu and Naval

Base), allowing for the passage of large draft ships, including those serving the Aratu Industrial Center and the Camaçari Petrochemical Complex (Hatje and Andrade, 2009). Due to these sources of pollution, a higher influence of petrogenic PAH than in sample I3 would be expected. This is evident from the FCM, which shows that I5 sample has a lower degree of membership to the pyrolytic group than the I3 sample. Similarly, I5 has a greater membership in the petrogenic group than I3. For both samples, a portion of the PAH profile constituted by higher molecular weight compounds is probably due to the accentuated degradation of hydrocarbons in this tropical environment.

Sample I11 showed the highest concentrations of PAH among the four highlighted samples, mainly of the phenanthrene (petrogenic marker) and pyrene (pyrolytic marker) series. The region surrounding I11 is crossed by pipelines connecting the refinery to the port terminal (Hatje and Andrade, 2009). Different diagnostic ratios gave controversial indications of the predominant PAH sources (Figure 4), suggesting a predominance of petrogenic or pyrolytic compounds or mixture of sources. Conversely, the FCM identified a stronger influence of petrogenic compounds, while considering a significant pyrolytic contribution, since there was an intermediate degree of membership to the petrogenic group, however smaller than to the pyrolytic.

Sample I29 has both the highest degree of membership to the petrogenic group and the lowest concentration of PAH among the four samples highlighted. Nevertheless, diagnostic ratios suggest a predominance of compounds formed by biomass combustion. However, the degree of membership of this sample reflects distancing from pyrolytic PAH sources and the greater relative predominance of petrogenic compounds. This conflict highlights the advantage of FCM in the qualitative analysis. When used in addition to traditional methods that prioritize the quantitative approach, it offers a detailed profile of the contamination, thereby providing support for appropriate management strategies.

CONCLUSION

FCM enables verify the samples, which show relevance with each classification group. This approach proved efficient for the qualitative analysis of coastal environmental contamination, in which sources are superimposed. The information gathered thus far by FCM combined with data obtained by traditional methods provides an overall characterization of the studied compounds. The FCM applied to data which was assigned membership values for generated classes that varied continuously in space, unlike the abrupt and unnatural breaks from conventional methods. As such, this membership variation allowed new interpretations of the data, mainly in relation to the segregation into areas of influence. The application of this approach proved appropriate as a supplementary tool for environmental management issues. Since FCM does not rely on a normal probability distribution, it is best suited for application to a wide range of environmental data.

ACKNOWLEDGMENTS

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES) - Finance Code 001.

AUTHOR CONTRIBUTIONS

L.G.: Conceptualization, Data curation, Formal analysis, Methodology, Writing - Original Draft, Writing - Review & Editing.

C.G.M.: Conceptualization, Data curation, Formal analysis, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision.

R.S.C.: Writing - Original Draft, Writing - Review & Editing, Supervision.

A.L.R.W.: Writing - Original Draft, Writing - Review & Editing, Supervision.

REFERENCES

AMINI, M., AFYUNI, M., FATHIANPOUR, N., KHADEMI, H. & FLÜHLER, H. 2005. Continuous soil pollution mapping using fuzzy logic and spatial interpolation. *Geoderma,* 124(3-4), 223-233, DOI: https://doi.org/10.1016/j.geoderma.2004.05.009

BAUMARD, P., BUDZINSKI, H. & GARRIGUES, P. 1998. Polycyclic aromatic hydrocarbons in sediments and mussels of the Western Mediterranean Sea. *Environmental Toxicology and Chemistry,* 17(5), 765-776.

BECKER, R. A., WILKS, A. R., BROWNRIGG, R., MINKA, T. P. & DECKMYN, A. 2018. *maps: Draw Geographical Maps. R package version 3.3.0* [online]. Vienna: R package. Available at: https://CRAN.R-project.org/package=maps [Accessed: 15 November 2021].

BEZDEK, J. C., EHRLICH, R. & FULL, W. 1984. FCM: the fuzzy c-means clustering algorithm. *Computers & Geosciences,* 10(2-3), 191-203, DOI: https://doi.org/10.1016/0098-3004(84)90020-7

BIVAND, R., KEITT, T. & ROWLINGSON, B. 2020. *rgdal: Bindings for the "Geospatial" Data Abstraction Library. R package version 1.5-16* [online]. Vienna: R package. Available at: https://CRAN.R-project.org/package=rgdal [Accessed: 17 November 2021].

BIVAND, R. & LEWIN-KOH, N. 2018. *maptools: tools for handling spatial objects. R package version 0.9-4* [online]. Vienna: R package. Available at: https://CRAN.R-project.org/package=maptools [Accessed: 15 November 2021].

BUDZINSKI, H., JONES, I., BELLOCQ, J., PIÉRARD, C. & GARRIGUES, P. 1997. Evaluation of sediment contamination by polycyclic aromatic hydrocarbons in the Gironde estuary. *Marine Chemistry,* 58(1-2), 85-97. https://doi.org/10.1016/S0304-4203(97)00028-5

CEBECI, Z., YILDIZ, F., KAVLAK, A., CEBECI, C. & ONDER, H. 2017. *ppclust: probabilistic and possibilistic cluster analysis. R package version 1.1.0* [online]. Vienna: R package. Available at: https://cran.r-project.org/web/packages/ppclust/ppclust.pdf [Accessed: 17 November 2021].

CHRISTENSEN, J. H., TOMASI, G., SCOFIELD, A. D. L. & MENICONI, M. D. F. G. 2010. A novel approach for characterization of polycyclic aromatic hydrocarbon (PAH) pollution patterns in sediments from Guanabara Bay, Rio de Janeiro, Brazil. *Environmental Pollution,* 158(10), 3290-3297, DOI: https://doi.org/10.1016/j.envpol.2010.07.015

CUNHA, R. C., DUPAS, F. A., PONS, N. A. D. & TUNDISI, J. G. T. 2011. Análise da influência das variáveis ambientais utilizando inferência fuzzy e zoneamento das vulnerabilidades. Estudo do caso da Bacia Hidrográfica do Ribeirão do Feijão, São Carlos - SP. *Geociências,* 30, 399-414.

DICKHUT, R. M., CANUEL, E. A., GUSTAFSON, K. E., LIU, K., ARZAYUS, K. M., WALKER, S. E., EDGECOMBE, G., GAYLOR, M. O. & MACDONALD, E. H. 2000. Automotive sources of carcinogenic polycyclic aromatic hydrocarbons associated with particulate matter in the Chesapeake Bay region. *Environmental Science & Technology,* 34(21), 4635-4640, DOI: https://doi.org/10.1021/es000971e

DOBERMANN, A. & OBERTHÜR, T. 1997. Fuzzy mapping of soil fertility - a case study on irrigated riceland in the Philippines. *Geoderma,* 77(2-4), 317-339. https://doi.org/10.1016/S0016-7061(97)00028-1

DVORSKÁ, A., LAMMEL, G. & KLÁNOVÁ, J. 2011. Use of diagnostic ratios for studying source apportionment and reactivity of ambient polycyclic aromatic hydrocarbons over Central Europe. *Atmospheric Environment,* 45(2), 420-427, DOI: https://doi.org/10.1016/J.ATMOSENV.2010.09.06

FRANCIONI, E., WAGENER, A., SCOFIELD, A. L. & CAVALIER, B. 2005. Biomonitoring of polycyclic aromatic hydrocarbon in Perna perna from Guanabara Bay, Brazil. *Environmental Forensics,* 6(4), 361-370.

FRANSSEN, H. J. W. M. H., VAN EIJNSBERGEN, A. C. & STEIN, A. 1997. Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants. *Geoderma*, 60(2-4), 243-262, DOI: https://doi.org/10.1016/S0016-7061(97)00024-4

GARRIGUES P., BUDZINSKI H., MANITZ M. P. & WISE S. A. 1995. Pyrolytic and petrogenic inputs in recent sediments: a definitive signature through phenanthrene and chrysene compound distribution. *Polycyclic Aromatic Compounds,* 7(4), 275-284.

GÜLER, C., KURT, M. A., ALPASLAN, M. & AKBULUT, C. 2012. Assessment of the impact of anthropogenic activities on the groundwater hydrology and chemistry in Tarsus coastal plain (Mersin, SE Turkey) using fuzzy clustering, multivariate statistics and GIS techniques. *Journal of Hydrology,* 414-415, 435-451, DOI: https://doi.org/10.1016/j.jhydrol.2011.11.021

HATJE, V. & ANDRADE, J. B. 2009. *Baía de Todos os Santos. Aspectos oceanográficos* [online]. Salvador: EDUFBA. Available at: http://www.meioambiente.ba.gov.br/arquivos/File/Publicacoes/Livros/baia2009.pdf [Accessed: 30 October 2022].

HOPKE, P. K. 2015. It is time to drop principal components analysis as a "receptor model." *Journal of Atmospheric Chemistry,* 72, 127-128, DOI: https://doi.org/10.1007/s10874-015-9309-1

HU, Y., WANG, Z., WEN, J. & LI, Y. 2016. Stochastic fuzzy environmental risk characterization of uncertainty and variability in risk assessments: A case study of polycyclic aromatic hydrocarbons in soil at a petroleum-contaminated site in China. *Journal of Hazardous Materials,* 316, 143-150, DOI: https://doi.org/10.1016/j.jhazmat.2016.05.033

KJERFVE, B., RIBEIRO, C. H. A., DIAS, G. T. M., FILIPPO, A. M. & QUARESMA, V. 1997. Oceanographic characteristics of an impacted coastal bay: Baia de Guanahara, Rio de Janeiro, Brazil. *Continental Shelf Research,* 17(13), 1609-1643, DOI: https://doi.org/10.1016/S0278-4343(97)00028-9

KRISHNAPURAM, R. 1994. Generation of membership functions via possibilistic clustering. *In:* IEEE *International Conference on Fuzzy Systems*, 2, 902-908, DOI: https://doi.org/10.1109/fuzzy.1994.343851

MACKENZIE, M. J. & HUNTER, J. V. 1979. Sources and fates of aromatic compounds in urban stormwater runoff. *Environmental Science & Technology,* 13(2), 179-183, DOI: https://doi.org/10.1021/es60150a011

MASSONE, C. G., WAGENER, A. L. R., ABREU, H. M. & VEIGA, Á. 2013. Revisiting hydrocarbons source appraisal in sediments exposed to multiple inputs. *Marine Pollution Bulletin,* 73(1), 345-354, DOI: https://doi.org/10.1016/j.marpolbul.2013.05.043

ODEH, I. O. A., MCBRATNEY, A. B. & CHITTLEBOROUGH, D. J. 1992. Soil pattern recognition with fuzzy-c-means: application to classification and soil-landform interrelationships. *Soil Science Society of America Journal,* 56(2), 505-516, DOI: https://doi.org/10.2136/sssaj1992.03615995005600020027x

R CORE TEAM. 2013. *R: A language and environment for statistical computing* [online]. Vienna: *R* Foundation for Statistical Computing. Available at: http://www.R-project.org/ [Accessed: 15 November 2021].

ROCHA, T., PERES, S. M., BÍSCARO, H. H., MADEO, R. C. B. & BOSCARIOLI, C. 2012. Tutorial sobre Fuzzy-c-Means e Fuzzy Learning Vector Quantization: abordagens híbridas para tarefas de agrupamento e classificação. *Revista de Informática Teórica e Aplicada,* 19(1), 120-163.

SÂRBU, C. & EINAX, J. W. 2008. Study of traffic-emitted lead pollution of soil and plants using different fuzzy clustering algorithms. *Analytical and Bioanalytical Chemistry,* 390, 1293-1301, DOI: https://doi.org/10.1007/s00216-007-1711-3

SOARES-GOMES, A., NEVES, R. L., AUCÉLIO, R., VAN DER VEN, P. H., PITOMBO, F. B., MENDES, C. L. T. & ZIOLLI, R. L. 2010. Changes and variations of polycyclic aromatic hydrocarbon concentrations in fish, barnacles and crabs following an oil spill in a mangrove of Guanabara Bay, Southeast Brazil. *Marine Pollution Bulletin,* 60(8), 1359-1363, DOI: https://doi.org/10.1016/j.marpolbul.2010.05.013

STANIMIROVA, I., TAULER, R. & WALCZAK, B. 2011. A comparison of positive matrix factorization and the weighted multivariate curve resolution method. Application to environmental data. *Environmental Science Technology,* 45(23), 10102-10110, DOI: https://doi.org/10.1021/es201024m

TAKADA, H., ONDA, T., HARADA, M. & OGURA, N. 1991. Distribution and sources of polycyclic aromatic hydrocarbons (PAHs) in street dust from the Tokyo Metropolitan area. *Science of Total Environment,* 107, 45-69, DOI: https://doi.org/10.1016/0048-9697(91)90249-E

TAN, M. Z., XU, F. M., CHEN, J., ZHANG, X. L. & CHEN, J. Z. 2006. Spatial prediction of heavy metal pollution for soils in peri-urban Beijing, China based on fuzzy set theory1 1 project supported by the national natural science foundation of China (Nos. 40571065 and 40235054) and the National Key Basic Research Support. *Pedosphere,* 16(5), 545-554, DOI: https://doi.org/10.1016/S1002-0160(06)60087-8

TENNEKES, M. 2018. tmap: thematic maps in R. *Journal of Statistical Software,* 84(6), 1-39, DOI: https://doi.org/10.18637/jss.v084.i06

TOLOSA, I., MORA, S., SHEIKHOLESLAMI, M. R., VILLENEUVE, J. P., BARTOCCI, J. & CATTINI, C. 2004. Aliphatic and aromatic hydrocarbons in coastal caspian Sea sediments. *Marine Pollution Bulletin,* 48(1-2), 44-60, DOI: https://doi.org/http://dx.doi.org/10.1016/S0025-326X(03)00255-8

VARNOSFADERANY, M. N., BAKHTIARI, A. R., GU, Z. & CHU, G. 2015. Distribution and characteristic of PAHs in sediments from the southwest Caspian Sea, Guilan Province, Iran. *Water Science & Technology,* 71(11), 1587-1596, DOI: https://doi.org/10.2166/wst.2015.091

WAGENER, A. L. R., CARREIRA, R. S., HAMACHER, C., SCOFIELD, A. L., FARIAS, C. O., CORDEIRO, L. G. M. S., LUZ, L. G., BAÊTA, A. P. & KALAS, F. A. 2011. Hydrocarbon composition and distribution in a coastal region under influence of oil production in northeast Brazil. *Marine Pollution Bulletin,* 62, 1877-1882.

WAGENER, A. L. R., FALCÃO, A. P., FARIAS, C. O., MOLINA, F. F., CARREIRA, R. S., MAUAD, C., NUDI, A., SCOFIELD, A. L. & MASSONE, C. G. 2019. Distribution and source apportionment of hydrocarbons in sediments of oil-producing continental margin: a fuzzy logic approach. *Environmental Science and Pollution Research,* 17, 17032-17044, DOI: https://doi.org/10.1007/s11356-019-05103-3

WAGENER, A. L. R., HAMACHER, C., FARIAS, C., GODOY, J. M. & SCOFIELD, A. 2010. Evaluation of tools to identify hydrocarbon sources in recent and historical sediments of a tropical bay. *Marine Chemistry,* 121(1-4), 67-79, DOI: https://doi.org/http://dx.doi.org/10.1016/j.marchem.2010.03.005

WAGENER, A. L. R., MENICONI, M. F., HAMACHER, C., FARIAS, C. O., SILVA, G. C., GABARDO, I. T. & SCOFIELD, A. L. 2012. Hydrocarbons in sediments of a chronically contaminated bay: the challenge of source assignment. *Marine Pollution Bulletin,* 64(2), 284-294, DOI: https://doi.org/http://dx.doi.org/10.1016/j.marpolbul.2011.11.018

WANG, Z., FINGAS, M. & PAGE, D.S. 1999. Oil spill identification. *Journal of Chromatography A,* 843(1-2), 369-411, DOI: https://doi.org/http://dx.doi.org/10.1016/S0021-9673(99)00120-X

WICKHAM, H. 2016. *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.

WILLIAMS, P. T. & BOTTRILL, R. P. 1995. Sulfur-polycyclic aromatic hydrocarbons in tyre pyrolysis oil. *Fuel,* 74(5), 736-742, DOI: https://doi.org/10.1016/0016-2361(94)00005-C

YAN, B., BOPP, R. F., ABRAJANO, T. A., CHAKY, D. & CHILLRUD, S. N. 2014. Source apportionment of polycyclic aromatic hydrocarbons (PAHs) into Central Park Lake, New York City, over a century of deposition. *Environmental Toxicology and Chemistry,* 33(5), 985-92, DOI: https://doi.org/10.1002/etc.2507

YUNKER, M. B., MACDONALD, R. W., VINGARZAN, R., MITCHELL, R. H., GOYETTE, D. & SYLVESTRE, S. 2002. PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition. *Organic Geochemistry,* 33(4), 489-515, DOI: https://doi.org/http://dx.doi.org/10.1016/S0146-6380(02)00002-5

ZADEH, L. A. 1965. Fuzzy sets. *Information and Control,* 8(3), 338-353, DOI: https://doi.org/10.1016/S0019-9958(65)90241-X