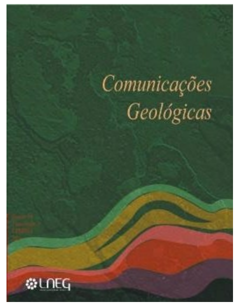
Comunicações Geológicas

Comunicações Geológicas (2024) 111, 1, 5-10
ISSN: 0873-948X; **e-ISSN:** 1647-581X

Levelling geochemical datasets as a tool to overcome boundary features in data applied to mineral exploration

Nivelamento de dados para corrigir efeitos de bordadura, provenientes de diferentes levantamentos de prospeção geoquímica



M. J. Batista[1*], D. de Oliveira[1], R. Salgueiro[1], C. Inverno[1]

DOI: https://doi.org/10.34637/pm5n-kv71
Recebido em 11/09/2023 / Aceite em 28/02/2024
Publicado online em abril de 2024

Artigo original
Original article

Abstract: The existence of a large number of chemical analyses resulting from decades of geochemical exploration campaigns, yields huge numbers of analytical results produced with different analytical methods, sampling protocols and sampling means. Therefore, it is necessary to level these analytical results to make them comparable and thus take advantage of the potential of geochemical information from contiguous areas. In this study, a levelling technique was used on geochemical datasets of tested bands between the 1:25 000 scale maps, 293/294 and 305/306 sheets, respectively, from the Beira Baixa region where stream sediments were sampled during a project for rare earth elements exploration. These samples were analysed for lanthanum by the same analytical methods and in the same laboratory, but in different periods and the border effects were observed. These levelling techniques must be adapted to the data to be compared and in the present case the disappearance of the border effect of the analytical results of 293/294 and 305/306 topographic sheets was observed.

Keywords: Lanthanum, stream sediments, geochemical mapping, Beira Baixa, regression equation.

Resumo: A existência de um grande número de análises químicas resultantes de décadas de campanhas de prospeção geoquímica produz um grande número de resultados analíticos com diferentes métodos analíticos, protocolos e meios de amostragem. Nesse sentido, é necessário nivelar estes resultados analíticos para torná-los comparáveis e assim aproveitar o potencial da informação geoquímica de áreas contíguas. Neste estudo foi utilizada uma técnica de nivelamento de dados de geoquímica baseada em bandas testadas com larguras variadas entre os mapas à escala 1:25 000, 293/294 e 305/306, respetivamente, da região da Beira Baixa. Nestes, foram amostrados sedimentos de corrente durante um projeto de prospeção de Elementos de Terras Raras. Essas amostras foram analisadas para lantânio pelos mesmos métodos analíticos e no mesmo laboratório, mas em períodos diferentes e foram observados efeitos de bordadura. As técnicas de nivelamento foram adaptadas aos dados a comparar e observou-se o desaparecimento do efeito bordadura dos resultados analíticos entre as cartas topográficas 293/294 e 305/306.

Palavras-chave: Lantânio, sedimentos de corrente, cartografia geoquímica, Beira Baixa, equação de regressão

[1] Unidade de Recursos Minerais e Geofísica, Laboratório Nacional de Energia e Geologia (LNEG), Estrada da Portela, Zambujal, Apartado 7586, 2610-999 Amadora, Portugal.

* Corresponding author / Autor correspondente: mjoao.batista@lneg.pt

1. Introduction

Geochemical mapping has been carried out in Portugal since the late 1950's by the Portuguese Geological Surveys (currently Laboratório Nacional de Energia e Geologia: LNEG). Data stored for a long period of time results in an enormous archive of different analytical techniques and geochemical applications in different datasets. To conduct geochemical exploration, it is important to map regional areas to observe regional anomalies before target studies. The first step is to map the existing data. Frequently these geochemical data served specific objectives, analysed by different methods with a variety of chemical extractants resulting in different backgrounds and non-geological effects. Data representation effects such as borders between countries and topographic sheets, when appear in the geochemical maps, put in evidence differences between geochemical campaigns in contiguous areas usually related to analytical or instrumental differences that can be evident of different detection limit of elements (Grunsky, 2002). Sometimes, these are related with the lack of time between campaigns where calibrations of the same instruments may induce differences between sets of data analysis. Sampling protocols for different objectives can also induce different geochemical distribution of elements between campaigns, an example of this is the sampling of dry or wet streams (in dry streams usually is necessary to go deeper to avoid iron oxides and when this procedure is not respected the results may be increased in iron). Levelling different geochemical campaigns facilitates the mapping of old data in larger areas reducing the cost of detailed campaigns. Darnley *et al*. (1995) describe parametric and non-parametric methods of levelling and the restrictions of such procedures, such as the necessity of using comparable media and the same chemical element. For instance, compare cold partial extraction and total extraction, or compare stream sediments collected in the active stream and heavy mineral concentrates. Parametric levelling includes plotting x-y with or without logarithmic transformation although most trace elements need logarithmic transformation to respect the homogeneity of variance. If a logarithmic rule or a Poisson counting process is involved, a logarithmic or square-root transformation may be adequate, respectively, if positive or negative correlated plots are involved (Darnley *et al*., 1995). From the observation of the plots if a diagonal fit (Y=X) is observed there is no need of levelling, if there is a shift in the data only the intercept is significant and a simple positive or negative shift equal to the value of the intercept may be applied, but if only the slope is affected only a multiplier equal to the slope is applied. If both shift and slope are applied both need to be applied and transformed from the regression equation. Non-parametric levelling (Darnley *et al*., 1995) includes: Fractile Normalization, Clarke Normalization and Quantile Regression but details of these techniques will not be further explained. Multiple phases of levelling are sometimes necessary to remove the artificial patterns evident on the raw geochemical images of a regional data compilation as stated in Agnew *et al*. (2004). Levelling can also be made with parts of coincident areas, for example the Geochemical Mapping of Agricultural Soils of Europe (GEMAS) dataset and the Baltic Sea Survey (BSS) dataset were levelled for Fe, V and Y using quantile regression and also Bivariate Least Square regression showing that both methods produce effective equations (Pereira *et al*., 2016). Examples also occur where no correlation exists between the pair of datasets to level and reference data and, is not possible to proceed with levelling. The most frequent cause is the very different resolution between analytical methods.

The objective of this study is to apply quantile regression to boundaries of contiguous sheets with observed artificial effects (Daneshfar and Cameron, 1998). The studied example consists of levelling the lanthanum results in stream sediments in two sampling campaigns in 1:25 000 topographic sheets 293/305 and 293/294/305/306 of the Beira Baixa province. The data were aquired during the Rare Earth Elements Exploration Project (Inverno *et al*, 2007).

2. Geological Setting

The studied area is included in the Rosmaninhal (25-C), Segura (25-D) and Retorta (north sector; 29-A) 1:50 000 scale geological map, belonging to the Beira Baixa province, in central east region of Portugal, included in part of the Central-Iberian Zone Terrane (Fig. 1). The region is mainly composed of the Schist and Greywacke Complex (SGC; Neoproterozoic to lower Cambric) generally oriented WNW-ESE with thin metasiltites interlayered with metagreywacke from lower unit of the Malpica do Tejo formation. The upper unit of the same formation consists of layers of metagreywacke, metasiltite, metapelites and metaconglomerates. Over these formations occurs the Rosmaninhal Formation from the lower Cambric age consisting of metapelites interlayered by metaconglomerates and metagreywackes. The previous formations constitute the *Grupo das Beiras* and were later superimposed discordantly by the quartzitic rocks protoliths of the Armorican Quartzite Formation (Lower-Middle Cambrian to Lower Ordovician) as a consequence of the Sardic deformation. Armorican Quartzite Formation rocks are displayed as a NW-SE Variscan synclinal structure. The former (SGC) and the latter were subject to Variscan low-grade metamorphism. Covering the region occurs the Cenozoic sedimentary Formation of Cabeço do Infante (Romão *et al*., 2010). On a regional scale, mega-structures of interference occur in the Beiras Group, designated as: Malhadio-Serrinha Antiform; Synform of Rosmaninhal-Monforte da Beira. The Cenozoic deposits represented in the region by the Cabeço do Infante, are strongly controlled by the tectonic reactivation of Variscan structures; in this geological context, Ponsul fault, reactivated as reverse fault during Alpine deformation, established several

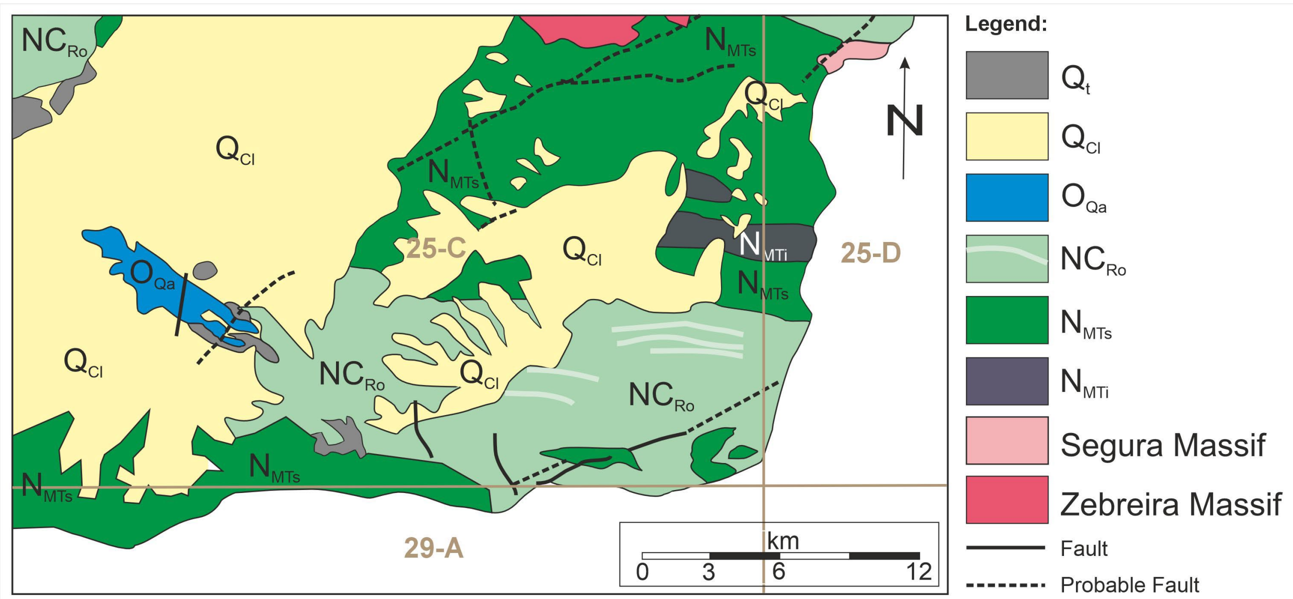


Figure 1. Geological map of the study area, representing the granitic massifs of Zebreira Massif and Segura Massif and, from the older to the more recent, formations Lower Member of Malpica do Tejo formation (NMTi), Upper Member of Malpica do Tejo formation (NMTs), Rosmaninhal formation (NCRo), Armorican Quartzite formation (OQa), Cabeço do Infante formation (QCl) and fluvial terrace sediments of Pleistocene age (Qt) (adaptad from Sheet 25-C Rosmaninhal, 25-D Segura, 29-A Retorta (north sector), 1:50 000 scale Geological Map of Portugal).

Figura 1. Mapa geológico da área de estudo representando os maciços graníticos de Zebreira e de Segura e, das formações mais antigas para as mais recentes, Membro Inferior da Formação Malpica do Tejo (NMTi), Membro Superior da Formação Malpica do Tejo (NMTs), Formação do Rosmaninhal (NCRo), Formação do Quartzito Armoricano (OQa), Formação Cabeço do Infante (QCl) e sedimentos de terraço fluvial de idade pleistocénica (Qt) (adaptado da Folha 25-C Rosmaninhal, 25-D Segura, 29-A Retorta (sector norte), na escala 1:50 000) da Carta Geológica de Portugal.

contacts between the Cenozoic Formations and the SGC metasedimentary rocks (Fig. 1).

3. Materials and Methods

The study focuses on the boundary of 1:25 000 scale sheets 293, 294, and 305, 306 where stream sediment samples were collected between 1996 and 2000 although previous samples were collected during the Gois-Segura project (Viegas, 1987) during the INTERREG Project "Inventariação e Prospecção de Terras Raras nas Regiões Fronteiriças da Beira Baixa e do Norte Alentejo" (Inverno *et al*., 2007). The levelling is based in 340 samples collected in the southern part of 293 and 265 samples collected in the northern part of sheet 305 in the first levelling. In the second levelling, 305 samples were used from the A2 (Fig. 2; East part of sheets 293 and 305 of the topographic map) boundary band and 236 samples from the B2 (Fig. 2; West side of 294 and 306 sheets of the topographic map).

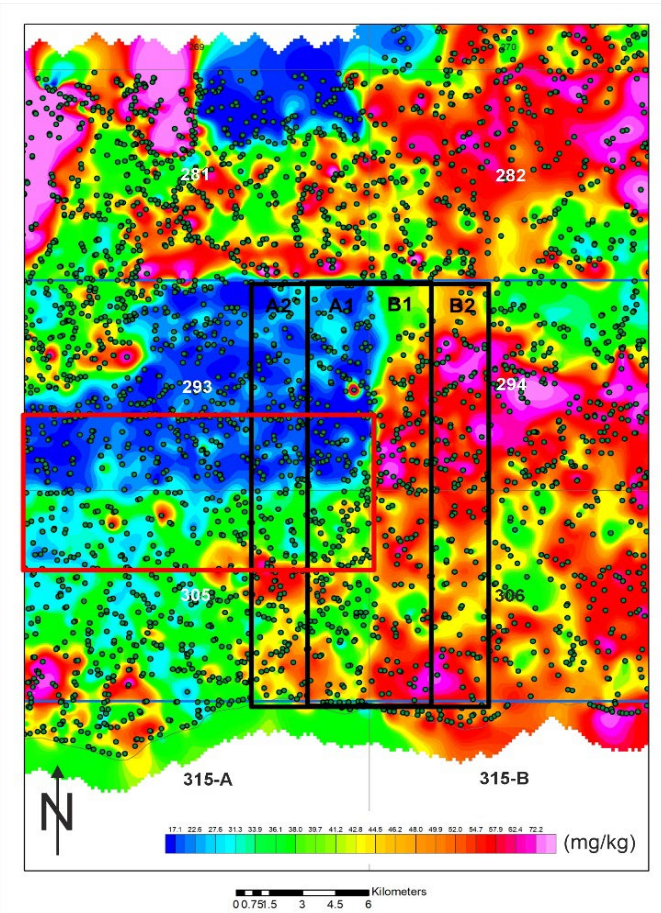


Figure 2. Sample locations and boundary areas to level between 293/305 and between 293-305/294-306 sheets in Lanthanum distribution in the studied area prior to levelling.

Figura 2. Locais de amostragem e áreas limite para nivelar entre folhas 293/305 e entre 293-305/294-306 na distribuição de Lantânio na área estudada antes do nivelamento.

One split of about 30 g of each sample, powdered (-80 mesh) from the stream sediment samples was sent for multi-element chemical analysis (49 elements) in the ACTLABS – Activation Laboratories Ltd. in Ontario, Canada. The project stream sediments were analysed in the same laboratory and with the same method but sent to the laboratory probably in different time periods. Hence, differences appear between the map sheets highlighting artificial effects. This study focuses on lanthanum (La) levelling, a Rare Earth Element analysed by Instrumental Neutron Activation Analysis (INAA) whereas quality control was made with laboratory internal standards, in 5% of the samples were made laboratory replicates differ from the original less than 20%, therefore accuracy is good. International standards of G-2 ad SCO-1 were analysed and differences in instrumental results were within the less than 10% which shows a good precision. The levelling is based on the Daneshfar and Cameron (1998) method, consisting in comparing pairs of quantiles between bands of similar geology. The La results covering most of the area of 293 and 305 topographic sheets (Fig. 2), represented a population of 340 sample results for 293 topographic sheet south border band and 265 sample results in the 305 topographic sheet north border band. The bands of these two sheets (northside of 305 and southside of 293 sheet) were the first to be levelled and 305 considered the sheet with the "true" results and the 293 the sheet to be levelled. Quality control of the process was made by testing the normality of the entire dataset of La 11277 results was determined by visual observation of the histogram showing the lognormality of the data (Fig. 3).

The normality of the bands was tested by Shapiro-Wilks normality test and the results can be observed in table 1.

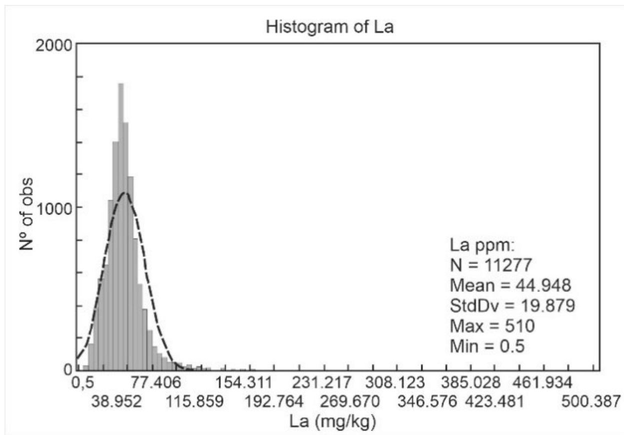


Figure 3. Lanthanum histogram of La (mg/kg) data showing the necessity of log transformation of the data prior to quantile calculations.

Figura 3. Histograma dos dados Lantânio (mg/kg) mostrando a necessidade de transformação logarítmica antes dos cálculos quantis.

Table 1. Shapiro-Wilks normality test results expressed in W and $p < 0.05$.

Tabela 1. Teste de normalidade Shapiro-Wilks expresso em W e p < 0,05.

Bands	W	p < 0.05
LaNside305	0.96624	0.00001
LaSside293	0.78868	0.00000
A1	0.95893	0.00000
B1	0.98322	0.00687
A2	0.9145	0.00000
B2	0.97636	0.00001

Because the method supposes normality and La in these datasets is a lognormally distributed element with all bands with $p < 0.05$, log transformation was necessary.

The quantiles (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95) were calculated after log transforming the geochemical data and quantile-quantile plot adjusting the regression line to a 95% level of confidence, where in the plot, the Y axis represents the levelling quantiles and the X axis the quantiles to be levelled. The equation was used to obtain the levelled 293 data substituted and projected in the regional geochemical map. In the second phase the method consisted in selecting the best possible band width between A1/B1 or A2/B2, meaning that the levelling is made in two sheets each side (Fig. 2) sheets 293-305 as A1 or A2 and294-306 as B1 or B2 (Fig. 2). To ensure good representativity of the geological formations and better correlation of La data between bands in both sides A2/B2 was selected. The procedure in this second levelling was the same as in the first. After levelling both boundaries, all log (La) values were transformed from log scale into the original scale and projected in the map.

Statistical calculations were made with the software STATISTICA 12.0, geochemical mapping was made in Oasis montaj 8.3.5 and the geological overlapping was made in ArcGis 10.8.1.

4. Results and discussion

The La median results of 293 and 305 topographic sheet (Fig. 2), were 22.3 mg/kg and 37.8 mg/kg, respectively, in the original data of the two sheets.

Since the sample medium is the same for every topographic sheet and the analytical method was the same, the reason for the shift observed between sheets can be different calibration of equipment in different periods of analysis. Clearly by observation of the map, the La results in topographic sheet 293 have a different background and is delimited by the boundary of the sheet.

Figure 4 shows the quantile-quantile plot representing the “true” values (topographic sheet 305) in the Y and the levelling results of sheet 293 in the X axis with the interval of confidence of 95%.

According to the criteria of Darnley *et al*. (1995) and Daneshfar and Cameron (1998) the quantile-quantile plot regression model shows that a shift and a multiplier were necessary to level the data.

The regression equation calculated for 293/305 is:

$$\log_{10}La_{(293lev)} = 0.5258 + 0.7621 \times \log_{10}La_{(293)}.$$

Figure 5 shows that between the southern band of sheet 293 and the northern band of sheet 305 the data of La were effectively levelled, where the difference in the median value of La in the sheet 293 changed from 22.3 mg/kg to 35.7 mg/kg after levelling.

Figure 5 further shows that the effect of shift in the boundary disappeared. Nevertheless, it is possible to observe that the boundaries East of the 293 and 305, and West of 294 and 306 topographic sheet still show some shift.

In a second phase, the choice of A2-B2 area instead of A1-B1 area (Fig. 4) was based in the better representativity of the same geological formations between both bands A2 and B2 than in A1-B1. Therefore, the plot of the quantile-quantile pairs was produced for A2 (East side of 293 and 305) band and B2 (West side of 294 and 306).

According to the method it needs to be assigned which sheet have the real values (“true”) where geochemical results represent natural features. The figure 6 is the quantile-quantile diagram with B2 data, considered the “true” data, is represented in the Y axis and A2, the data to level, in the X axis, with the Interval of confidence of 95%.

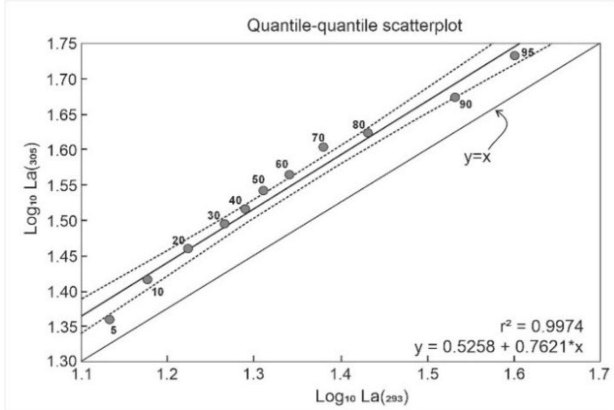


Figure 4. Quantile-Quantile diagram, with R^2=0.9974 and $p<0.05$, of the boundary area (band south of 293 and north of 305) to be levelled showing the straight-line equation used to level the data in that boundary.

Figura 4. Diagrama Quantil-Quantil, com R^2=0.9974 e $p<0.05$, da área limite (banda sul da folha 293 e norte da 305) a ser nivelado, mostrando a equação da reta usada para nivelar os dados nesse limite.

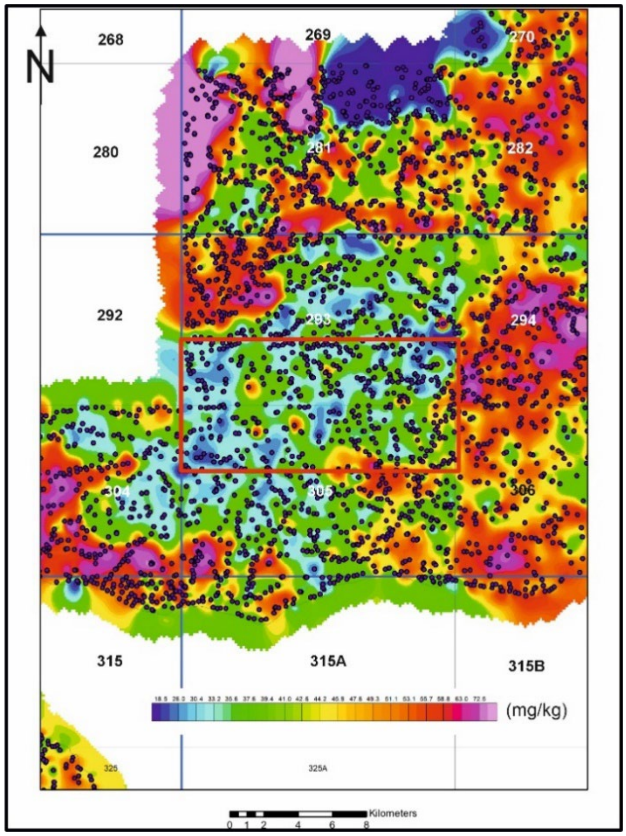


Figure 5. Lanthanum distribution in the studied area after first levelling between 293 and 305 sheets.

Figura 5. Distribuição do Lantânio na área estudada após o primeiro nivelamento entre as folhas 293 e 305.

The regression equation calculated for A2/B2 is:

$$\log_{10}La_{(A2lev)} = 0.627 + 0.6553 \times \log_{10}La_{(A2)}.$$

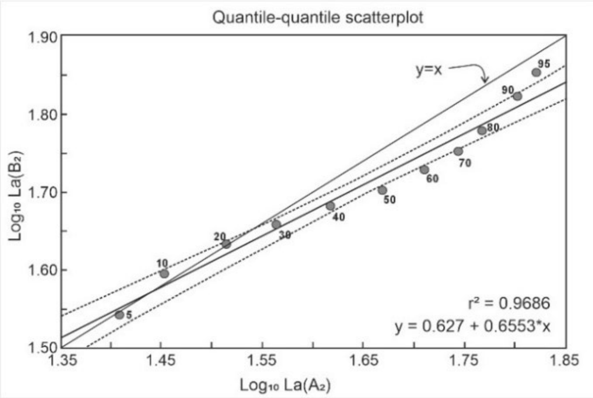


Figure 6. Quantile-Quantile diagram of the boundary area A2/B2 (A2 - 293 and 305; B2 – 294 and 306) to be levelled showing the straight-line equation used to level the data in that boundary.

Figura 6. Diagrama Quantil-Quantil da área limite A2/B2 (A2 - 293 e 305; B2 – 294 e 306) a ser nivelado, mostrando a equação da reta usada para nivelar os dados nesse limite.

Figure 7 shows that the effect of shift disappears. It is now possible to observe the geological distribution of lanthanum.

After leveling the data and overlaying with the 1:50 000 scale geological map (25-C, 25D, 29A) (Romão *et al.*, 2010) the Cabeço do Infante Formation composed of sandstones interbedded with conglomerates that covers the metasediments reveals to contain regional background values of La between 36 and 42 mg/kg lower than the turbiditic metasediments and metaconglomerates of the Malpica do Tejo Formation and metapelites interbedded with metaconglomerates and metagrauvaques from the Rosmaninhal formation (Romão *et al.*, 2010) that have background La contents between 45 and 72 (mg/kg).

It is also important to highlight that La as a REE is usually associated with very resistant minerals such as monazite, with

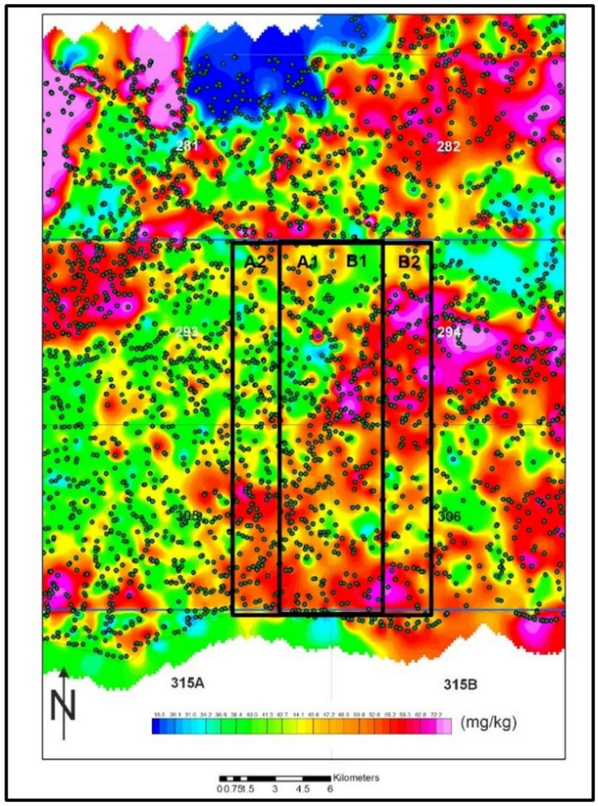


Figure 7. Lanthanum distribution in the studied area after levelling of data.

Figura 7. Distribuição do Lantânio na área estudada após o nivelamento dos dados.

a mechanical mobility throughout the drainage net and this is also reflected in the distribution of the element in the study area (Fig. 8). From a tectonic point of view, the Ponsul fault promoted the superimposition of the Malpica do Tejo

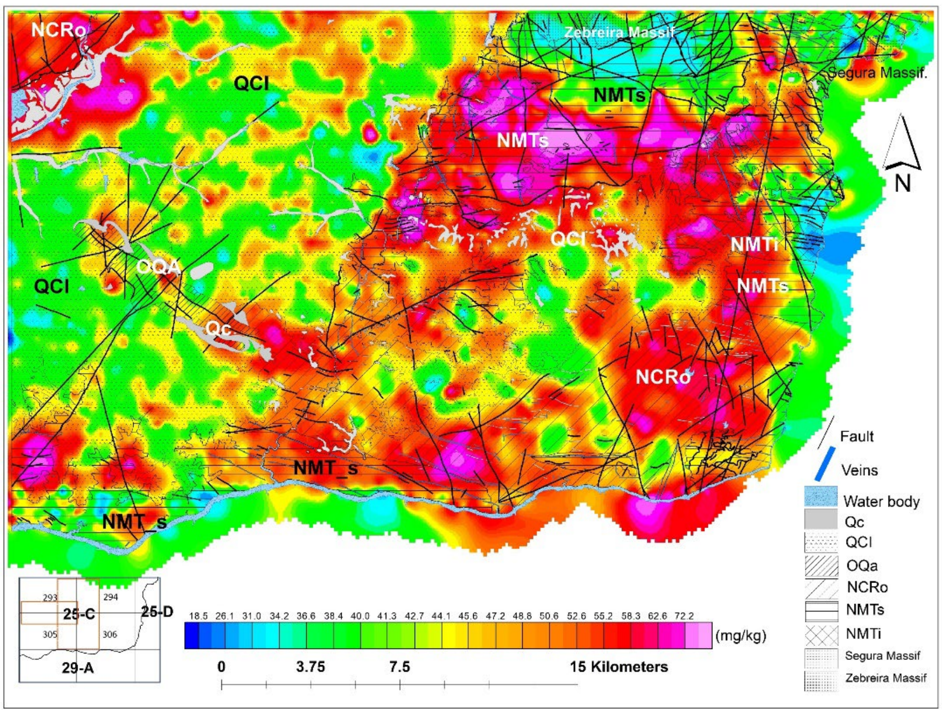


Figure 8. Lanthanum distribution in the 1:50 000 scale, 25-C Rosmaninhal, 25-D Segura, 29-A Retorta (north sector) areas after levelling of data. The scheme of the sheets in the lower left side shows the 1:50 000 and 1:25 000 scale number of sheets; the red areas in the same scheme show the levelled areas.

Figura 8. Distribuição do Lantânio nas áreas da Folha 25-C Rosmaninhal, 25-D Segura, 29-A Retorta (sector norte) na escala 1:50 000 após nivelamento dos dados. O esquema das folhas no canto inferior esquerdo mostra o número de folhas nas escalas 1:50 000 e 1:25 000; as áreas vermelhas no mesmo esquema mostram as áreas niveladas.

Formation over the most recent sediments of the Cabeço do Infante Formation in the NW part of the map influencing the La distribution. It is also possible to observe an elevated La limit in an E-W direction probably controlled by NNE-SSW, E-W; and NNW-SSE faults near the Ribeira do Freixinho (Fig. 8) and to the west where there is not a very clear reason for the limit within the Malpica do Tejo Superior Member.

5. Conclusions

As this study has shown, multiple phases of levelling were necessary to remove the artificial patterns evident on the raw geochemical mapping of a regional dataset compilation.

Levelling was possible and successful in the topographic sheets 1:25 000 scale 293 and 305 using the reference data of the contiguous 294 and 306 sheets.

After levelling the data, the background concentrations of La could be associated with geological formations. Nevertheless, levelling needs to be a cautious exercise with the specific purpose of presenting a larger area of geochemical data. The original data are preserved and should be used whenever local targets are studied. Data with different sampling media, sampling protocols and analytical techniques result in different background levels observable in contiguous areas and in some cases will not be possible to level the data.

Acknowledgements

This research was done using data from the project *Inventariação e prospecção de terras raras nas regiões fronteiriças da Beira Baixa e do Norte Alentejo,* founded by the European Union Cross-border INTERREG II program.

References

Agnew, P. D., Muhling, J., Goldfarb, R. J., 2004. Applications of geochemistry in targeting with emphasis on large stream and lake sediment data compilations. *Predictive Mineral Discovery Under Cover*, 232.

Daneshfar, B., Cameron, E., 1998. Levelling geochemical data between map sheets. *Journal of Geochemical Exploration*, **63**: 189-201.

Darnley, A. G., Björklund, A., Bølviken, B., Gustavsson, N., Koval, P. V., Plant, J. A., Steenfelt, A., Tauchid, M., Xuejing, X., 1995. A Global Geochemical Database for Environmental and Resource Management. Recommendations for international geochemical mapping. Final report of IGCP project 259. UNESCO. Publishing.

Grunsky, E. C., 2002. Statistical analysis in the geosciences. *In*: *Geoinformatics. Encyclopedia of Life Support Systems* (EOLSS) II.

Inverno, C. M. C., Oliveira, D. P. S., Rodrigues, L., (colaboradores: Viegas, L., Matos, J., Martins, L., Salgueiro, R., Lencastre, J., Farinha, J., Rosa, D., Chichorro, M., Santana, H., Oliveira, V., Fernandes, J., Pateiro, R.), 2007. Inventariação e prospecção de terras raras nas regiões fronteiriças da Beira Baixa e do Norte Alentejo. *INETI Report*, Alfragide, 2982.

Pereira, B., Vandeuren, A., Govaerts, B. B., Sonnet, P., 2016. Assessing dataset equivalence and levelling data in geochemical mapping. *Journal of Geochemical Exploration*, **168**: 36-48.

Romão, J., Cunha P. P., Pereira A., Dias, R., Cabral J., Ribeiro A., 2010. Notícia explicativa das folhas 25-C Rosmaninhal, 25-D Segura e 29-A Retorta (Sector Norte) da Carta Geológoica de Portugal na escala 1:50 000. Unidade de Geologia e Cartografia Geológica, Laboratório Nacional de Energia Geologia, 58. ISBN: 978-989-675-009-1.

Viegas, L., 1987. Faixa de Góis-Segura. Relatório anual de atividades de 1986. Divisão de prospeção de minérios metálicos, 116.