PRINCIPAL COMPONENT ANALYSIS APPLIED TO STREAM SEDIMENT DATA IN PASSO DO SALSINHO, RS, BRAZIL

José Leonardo Silva Andriotti [(1)]

(1) CPRM – Serviço Geológico do Brasil

RESUMEN

Análisis en Componentes Principales (ACP) es una técnica de tratamiento multivariable de datos muy utilizada en la Exploración Geoquímica en los últimos años. Por medio de ACP se puede identificar asociaciones importantes de variables en el conjunto de datos y relacionar estas asociaciones con procesos geológicos que actuaron sobre las rocas estudiadas. Una de sus ventajas es la posibilidad de cartografiar los procesos geológicos identificados. ACP fue aplicada en Passo do Salsinho (Rio Grande do Sul, Brasil) utilizándose 441 muestras de ríos analizadas para 20 variables, y mostró mineralizaciones y variables con poca contribución al conocimiento de la geoquímica de la región.

ABSTRACT

Principal Component Analysis (PCA) is a multivariate technique that has had large uses in Geology since long ago, and its uses have been increased a lot in the last two decades. Geochemical surveys, as they are multivariable studies, must consider the approaches that treat all the variables under consideration at the same time, taking into account their relationships together with the individual characteristics of each variable. PCA was used to treat 441 stream sediments data analysed for 20 variables in an area of the Rio-Grandense Shield, south of Brazil (Passo do Salsinho area), and the results were considered satisfactory as they showed the contour of the main mineralized zones known in the area, and it has also showed the variables that were not contributing to the knowledge of the geochemical pattern of the area.

PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis is a multivariate technique that aims at studying several variables together. This type of treatment allows us to take into account the relationships (represented by the correlation matrix) that exist among all studied variables. As a result of its improvement PCA enables the creation of new variables – the principal components (PCs) – that are linear combinations of the original ones. These new variables are uncorrelated and they are nothing than the eigenvectors of a correlation matrix, and they are derived in decreasing order of importance so that, for example, the first principal component accounts for as much as possible the variation of the original data. The new uncorrelated variables form a set with the same quantity of the original variables, but the reducing of dimensionality of the problem is obtained because one study only the PCs that explain an important portion of the total variability of the data. Many of the operations associated with PCA are distribution-free. In PCA the goal is to maximize variances, and the standard procedure for maximizing a function of several variables subject to one or more constraints is the method of Lagrange multipliers.

Each PC explains a percentage of the total variability of the original data and this percentage accounted for is interpreted as the portion "explained by" the process represented by the PC. The first PC is more highly correlated with the original variables than the second, the second is more correlated with them than the third, and so on. The individual transformed observations are called scores, and PCs interpreted as representing geological processes can be mapped by the scores. One of the important points of PCA is that the PCs can be associated to geological processes and they reveal groupings of variables which would not be found by other means. The interpreter is asked to look at the groups of variables suggested by the PCs and consider whether the components have some meaningful interpretation. It is rather dangerous to read too much meaning into principal components.

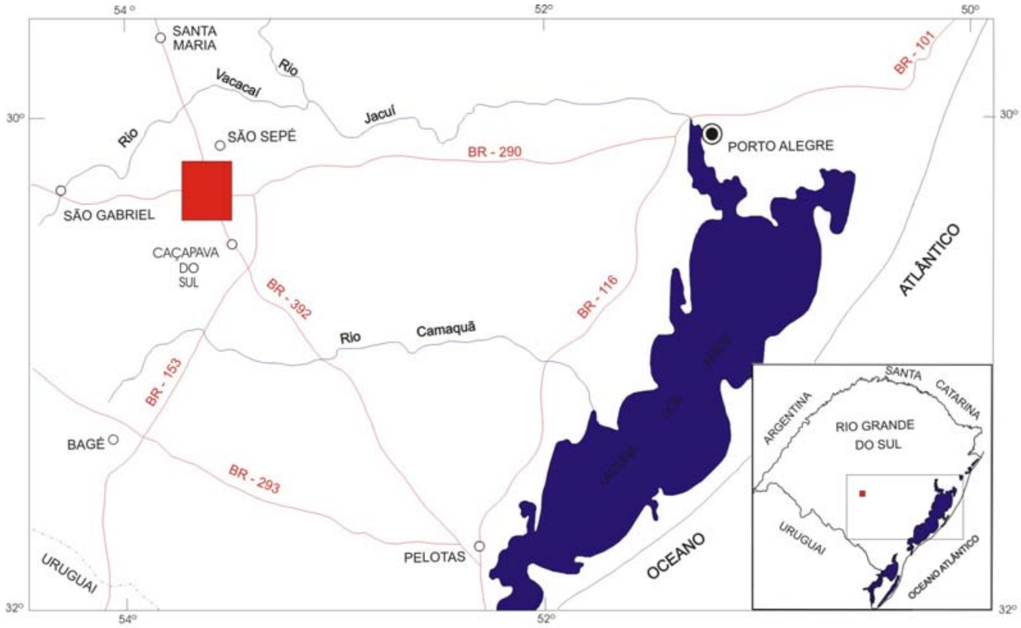
Geologists have been rather confused in the use of correct terminology in the case of PCA, mixing the use of PCA and Factor Analysis and their results, the PCs and the factors. Among the advantages of PCA one can cite that it's possible to determine the correlation of each PC with each of the original variables, it enables us to find outliers and groups of variables and allows us to reduce the dimensionality of the problem by the elimination of some variables in the next steps of the mineral exploration, if we consider that they are not helping to explain the processes interpreted via PCs. By the time simple vectors are obtained they probably will not be not well correlated with the original vectors. To alleviate this it is useful performing a rotation on the characteristic vectors producing new components that may be useful (rotations to a nonorthogonal or to an orthogonal solution are permissible, and the VARIMAX type is the more used rotation method, it has been used in this paper).

GEOLOGY

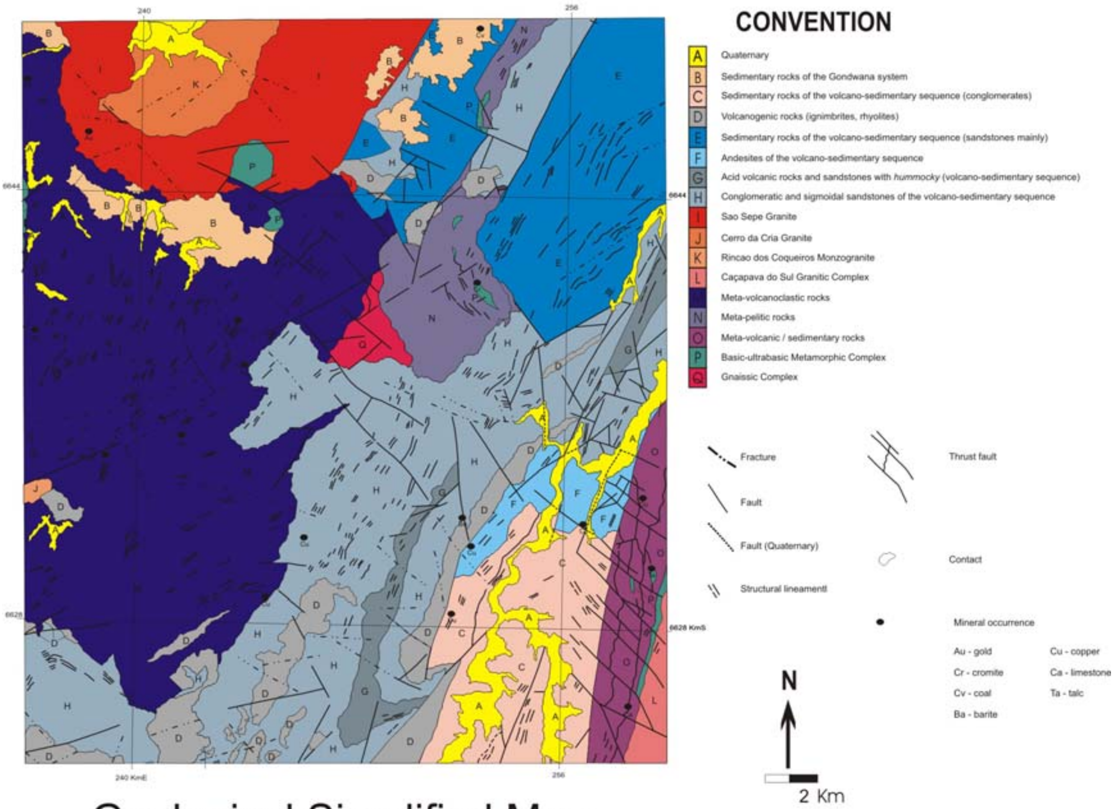
Passo do Salsinho area is situated in the south of Brazil (Rio Grande do Sul State), between 53° 30' and 53° 45' W and 30° 15' and 30° 30' S (figure 1) and is situated within the Sul-Riograndense Shield, a region that includes some areas that have been worked on for several years by geologists because they have some important mineral occurrences. Some of these mineral occurrences have been mined and others are in exploitation now. Passo do Salsinho presents some mineral occurrences like gold, copper, cromite, barite, talc, limestone and coal within 696 square kilometers. Passo do Salsinho is included in the Sao Gabriel portion of the

Sul-Riograndense Shield. This portion includes dominantly Proterozoic/Eo-Paleozoic volcano-sedimentary rocks of green–xists metamorphic facies and also stratiform basic–ultrabasic complexes, mafic and ultramafic sequences, granitic stocks with ages varying from 670 to 460 million years (Soliani Jr., 1986). Sedimentary rocks of the Gondwana system are also represented in the area; within these rocks there are two coal deposits. The main lithological units are represented in figure 1, according to the geological mapping due to Porcher *et al*. (1995), scale 1:50,000. The main fault systems observed in the area of Passo do Salsinho have dominant directions N30° – 40°E, N30° – 50°W and ENE –WSW to EW. Soliani Jr. (1986) studying this region stated the main ages K/Ar and Rb/Sr of the area. The Mata Grande Gabroic Complex has 2.2 billion years, and Sao Sepe and Cerro da Cria granites between 530 and 550 million years.

The gold mineral occurrences of the area are mainly emplaced in quartz veins and also in colluvial zones derived from these veins and from the metamorphic host rocks. Copper occurs as oxidated minerals (malachite and azurite togheter with primary minerals like chalcopyrite, bornite and chalcosine) or as hydrothermal remobilizations in fault zones, like the barite occurrence. Talc is associated to serpentinites and magnesian xists, cromite is associated with ultrabasic rocks in small and isolated bodies. Limestones occur as lenses in the meta-volcano sedimentary sequence, and coal seams occur in the permo-carboniferous sedimentary rocks of the Gondwana system.



Location map



Geological Simplified Map

Figure 1

Figure 1 – Geological Simplified Map

RESULTS

Twenty variables analysed by spectrometry emission for 441 stream sediments samples were used in this paper. The variables are Fe, Mg, Ca and Ti in % and Mn, B, Ba, Co, Cr, Cu, La, Ni, Pb, Sn, Sr, V, Y, Zr and Be in ppm, and analysed by atomic absorption for As, resulting one sample to each 1.5 square kilometers (Andriotti, 1999). This campaign was accompanied by other type of data, like rocks, soils, heavy mineral concentrates, geophysics and LANDSAT images.

Six PCs accounted for about 65% of the total variability of the original data. There were used for interpretation only the PCs whose eigenvalues were higher than one, and VARIMAX rotation was performed to best fit the variables to the axes that represent the PCs. The first PC (figure 2) is composed by Ca, Mn, Mg and Sr as important constituents, and three concentrations of high values of the scores were found. The first one is situated in the southeastern portion of the area and includes two copper and the talc mineral occurrences, that are within rocks owning to meta-volcano sedimentary rocks and to the basic-ultrabasic complex. The second one shows the area where two gold mineral occurrences are. The last concentration doesn't show relantionship with known mineral bodies, but it fits well with the volcano sedimentary rocks. It is important to remark that this component shows, as was expected, coincidence with the maps of its principal constituents, but also with high values of Cu and Ni, elements of no importance in the constitution of the first PC.

The second PC (figure 3) is represented by Y, La, Sn, Pb and Be, all of them with negative weights, ie, this component is more important where these elements don't present high values. Co and As, of no importance in this PC, show similarity of distribution of its scores. This PC is particularly important because it shows the contour of five gold occurrences, existing two others just at the side of its external contour. This PC can be associated with gold occurrences, and also shows the contour of the granitic rocks that outcrops in the area (low values). The third PC is based on Ba, Zr and, with secondary importance, Pb and Fe, all of them with negative weights, and has lithological significance only. The fourth and the fifth PCs are also associated with gold, copper and cromite occurrences, repeating with some approximation the results of the two first PCs. The sixth PC also shows the contour of the regions that encompass three mineral occurrences and the proximity of three others; its distribution is very similar to that of Ba in the area.

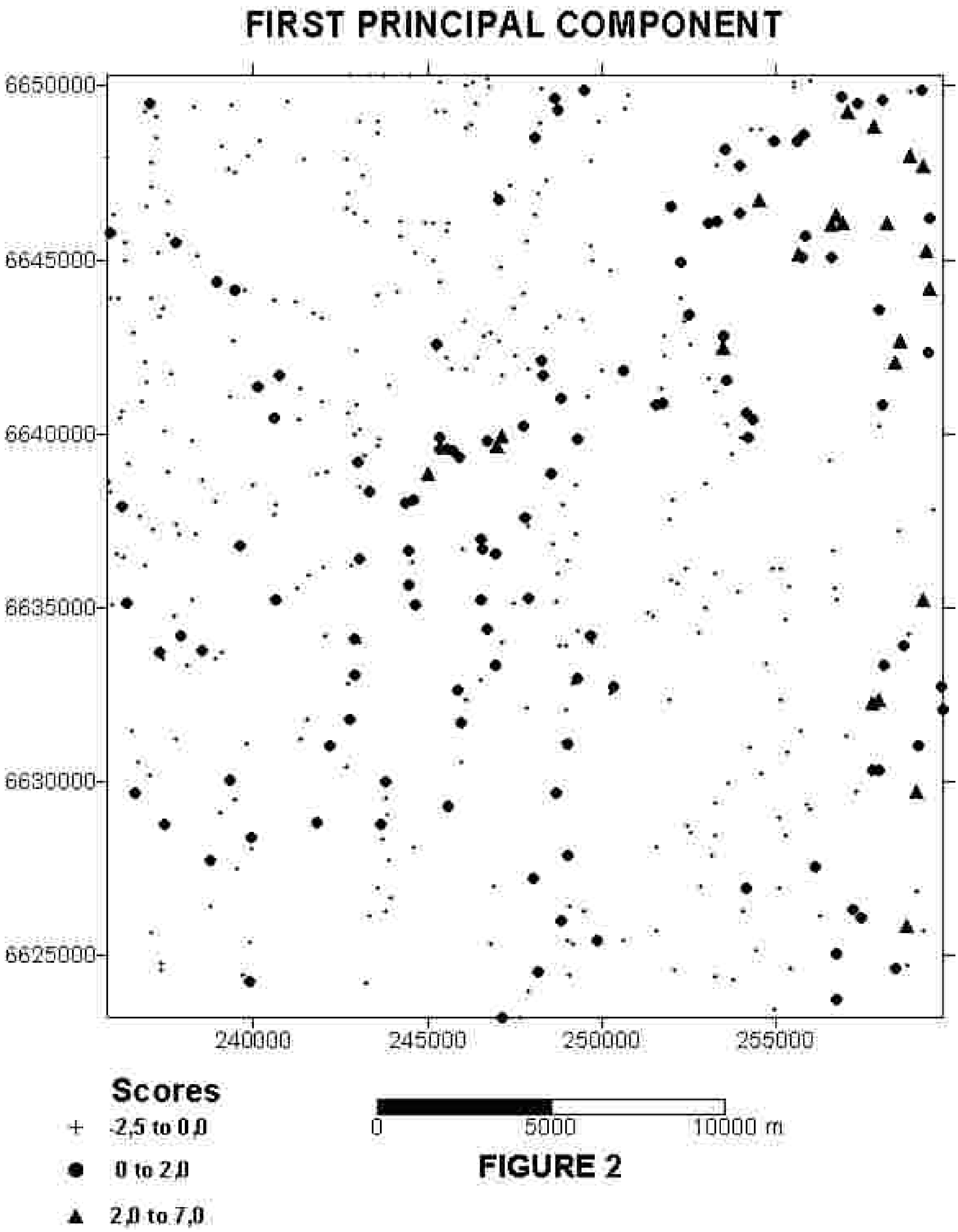


Figure 2 – First Principal Component

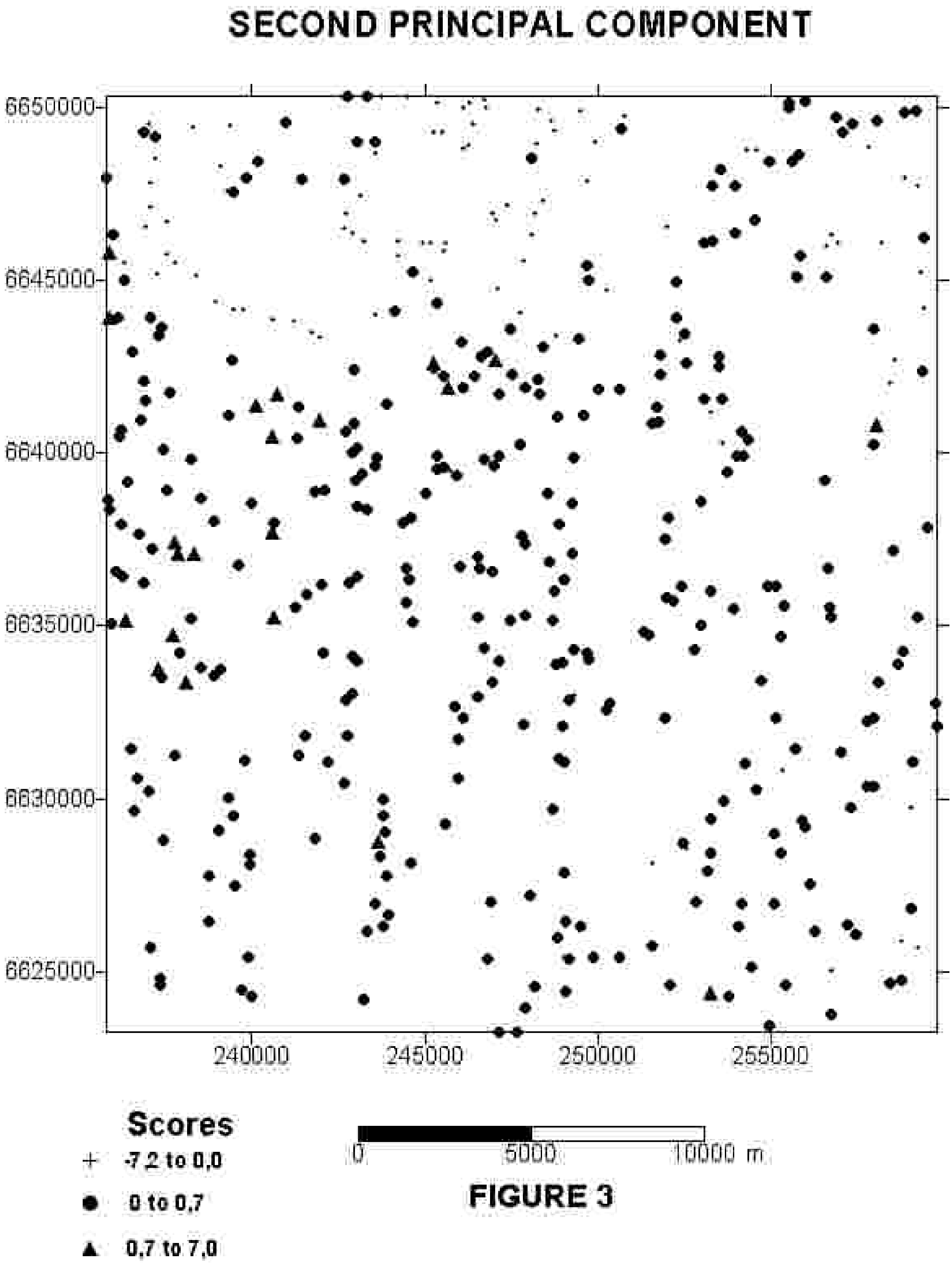


Figure 3 – Second Principal Component

CONCLUSIONS

The Principal Component Analysis (PCA) applied to stream sediments data showed results considered satisfactory because the maps of scores showed the contour of the main mineral occurrences of the area, and also some lithological contours. Its importance is also linked to the fact that some variables that were not contributing to the elucidation of the geochemical pattern of the rocks that outcrop in the area or to identification of the known mineral occurrences were identified, and they can be ignored in the next steps of the geochemical exploration of this area.

PCA confirms in the Passo do Salsinho area the good results and conclusions achieved by many authors all over the world when they used this multivariate technique to interprete stream sediments data.

REFERENCES

Andriotti, J. L. S. – 1999 - Interpretação e Modelamento de Dados Geoquímicos e de Sensoriamento Remoto por meio da Análise de Componentes Principais: Implicações na Exploração de Depósitos de Metais no Escudo Sul-rio-grandense - Tese de Doutorado, UFRGS, 227 p.

Porcher, C. A.; Leites, S. R.; Ramgrab, G. E. & Camozzato, E. – 1995 – Folha Passo do Salsinho – Folha SH.22-Y-A-I-4. CPRM, Programa Levantamentos Geológicos Básicos do Brasil, 1:50,000, 372p.

Soliani Jr. – 1986 – Os dados geocronológicos do Escudo Sul-Riograndense e suas implicações de ordem geotectônica. PHd thesis, USP, 425p.