Geographically Weighted Principal Components Analysis to assess diffuse pollution sources of soil heavy metals. Application to rough mountain areas in Northwest Spain.

Journal: *Land Degradation & Development*

Manuscript ID: LDD-16-0032

Wiley - Manuscript type: Research Article

Date Submitted by the Author: 12-Jan-2016

Complete List of Authors: Fernández, Susana; University of Oviedo, Geology Department
Cotos-Yáñez, Tomas; University of Vigo, Department of Statistics and Operations Research
Roca-Pardiñas, Javier; University of Vigo, Department of Statistics and Operations Research
Ordoñez, Celestino; University of Oviedo, Department of Mining Exploitation and Prospecting, Polytechnic School of Mieres

Keywords: soil heavy metals, Geographically Weighted Principal Components Analysis (GWPCA), diffuse pollution sources, forest fires, Principal Components Analysis (PCA)
Figure 1: The Principality of Asturias location and hillshade map

152x66mm (220 x 220 DPI)
Figure 2: Principality of Asturias geological map
401x271mm (250 x 250 DPI)
Figure 3: The map shows fires identified in Asturias with Landsat images in a long-time period.

338x190mm (96 x 96 DPI)
Figure 4: Sample points and watershed distribution
244x53mm (96 x 96 DPI)
Figure 5: Covariance scores for different bandwidth determination (left). Percentage of total variance for local components 1 to 6 (right).
254x190mm (96 x 96 DPI)
254x190mm (96 x 96 DPI)
Geographically Weighted Principal Components Analysis to assess diffuse pollution sources of soil heavy metals. Application to rough mountain areas in Northwest Spain.

Susana Fernández (1), Tomás Cotos-Yáñez (2) Javier Roca-Pardiñas (2), Celestino Ordóñez (3)

(1) Department of Geology, University of Oviedo, 33004 Oviedo, Spain
(2) Department of Statistics and Operations Research, University of Vigo, 36208 Vigo, Spain
(3) Department of Mining Exploitation and Prospecting, Polytechnic School of Mieres, University of Oviedo, 33600 Mieres, Spain.

Corresponding author: fernandezmsusana@uniovi.es

ABSTRACT

Heavy metals pollution can result in the degradation of the soil, air and water bodies quality affecting the health of all living organisms. We analyse the spatial distribution of the concentrations of several heavy metals and their relationship with natural or anthropogenic origin. The analysis was performed in the Principality of Asturias (mountain region of NW of Spain), where, as in many other parts of the world, soil heavy metal pollution has become a severe problem. First, a standard Principal Components Analysis (PCA) was performed over a population of 334 soil samples taken on slopes and fluvial plains to identify the sources of fourteen soil heavy metals (Ag, As, Ba, Hg, Cd, Co, Cr, Cu, Mn, Mo, Ni, Pb, Sb, Zn). Due to the high geological heterogeneity of the territory, the PCA analysis was improved using a variant of PCA known as Geographically Weighted Principal Components Analysis (GWPCA). The first six principal components in a standard PCA account for about 57% of soil heavy metals variability but when GWPCA is performed this figure increases to more than 80% in some areas. We conclude that WGPC1 corresponds to a geogenetic component.
with changing winning variables adapted to the geological characteristics of the territory (lithology and mining), while WGPC2 corresponds to a factor related to atmospheric pollution including heavy metals released from vegetation cover via forest fires.

Key words: soil heavy metals; Principal Components Analysis (PCA); Geographically Weighted Principal Components Analysis (GWPCA); diffuse pollution sources; forest fires.

INTRODUCTION

Soil heavy metals pollution has become a severe problem in many parts of the world due to the fact that the metal pollution is covert, persistent and irreversible (Bini et al., 2011; Zhang et al. 2009; Zhiyuan, 2014). This kind of pollution not only degrades the quality of the atmosphere, water bodies and food crops, but also threatens the health and well-being of animals and human beings by way of the food chain (Dong et al., 2011; Nabulo et al., 2012; Wang et al., 2012). Non-invasive techniques such as phytoremediation are the most suitable to mitigate the negative effects of soil pollution via uptake or immobilization of soil heavy metals in situ (Mahmoud & Abd El-Kader, 2014; Paz-Ferreiro et al., 2014; Sacristán et al., 2015; Wang et al., 2015). Heavy metals as mercury (Hg); chromium (Cr), cadmium (Cd) or metalloids such as arsenic (As) are present in the environment free or as part of different molecular forms (Chen et al., 1999). In natural soils they are present at a background level and usually occur as cations which strongly interact with the soil matrix (Alloway, 1995). In this way, some physicochemical properties of soils such as pH and organic matter are important parameters that control the accumulation and the availability of heavy metals in the soil environment. Generally speaking, heavy metals are distributed heterogeneously in the Earth’s crust as an effect of geological processes, and the elemental contents of non-
For Peer Review

3 polluted soils are incorporated by rock weathering processes. Among these, the main factor that dictates the elemental content of a soil is the composition of parent material but this can be increased due to anthropic causes (Alloway, 1995; Kabata-Pendias, 2004; Harmanescu et al., 2011).

In industrial areas anthropogenic activities such as agriculture, urbanization, industrialization and mining increase the metals concentration baseline (Adriano, 1992; Sheppard et al., 2000; Facchinelli et al., 2001; Wei & Yang, 2010; Zhong et al., 2012). The geological features such as lithology or mineralized areas associated to faults or thrusts exert a strong control on the heavy metals concentration and its variability in soils (Alloway, 1995; Kabata-Pendias, 2004). Smelting from industrial activities and cities is recognized as the most important source of heavy metals in the environment but little is known about the role of forest fires, which are frequent in this mountainous area. Ash is a key component of the land affected by forest fires (Cerdá 2007; Bodí et al., 2014; Pereira et al., 2013a). Furthermore, the ability of some natural plant species, named metalofitas, to take up, translocate and accumulate heavy metals in their shoots (Nanda et al 1995; Chaney et al. 1997) is well known. The combustion of these plant species could produce smog, necromase and ashes enriched in heavy metals that when deposited in topsoil contributes to raise the concentration of nutrients and pollutants such as heavy metals in soils. Nowadays a research effort concerning the legacy of atmospherically-deposited elements (e.g. heavy metals) in burned soils is needed but during the last decade some authors have shown the role of ash in the Earth and Soil System (Bodí et al., 2012; Pereira et al., 2013c). There are various reports about the elemental composition of the ashes in several burned soils of California, apart from elements like Ca>Al>Fe>K>Mg>Na, high concentrations of heavy metals such as Zn>Ba>Cu>Mn>Ag>As>Cd>Cr>Co were found in these soils (Plumlee et al., 2007;
Hageman et al., 2008a; Hageman et al., 2008b). Recently, a review about the heavy metals composition in wildfire ashes from Australian soils was published (Santín et al., 2015). All these researches seem to indicate that forest fires are an important source of heavy metals in soils.

Often geogenetic and anthropogenic sources of heavy metals are superimposed in the territory and it is very difficult to separate the contribution of each one from the soil heavy metals backgrounds, making difficult the identification of the boundary between natural and contaminated soils. Thus, statistical methods such as Principal Components Analysis (PCA) and clustering have been extensively used to identify sources of heavy metals in the environments (Hu et al., 2013, PCA has demonstrated to be especially useful to discover diffuse pollution sources by analysing the metals association in each principal component (Zhang et al., 2009; Wei et al., 2011). However, PCA method can be improved substantially using a variant method called Geographically Weighted Principal Components Analysis (GWPCA) when there is spatial heterogeneity in the data (Harris et al., 2011; Dempsar et al., 2013). In essence, this method consists in performing a local PCA, that is in the neighbourhood of each observation, instead of a global standard PCA.

The main objective of this research is to find the natural soil heavy metals backgrounds in the Principality of Asturias discovering possible sources of diffuse pollution using PCA over a soil population of 334 taken in the most pristine areas of the territory. The problem of spatial heterogeneity of geochemical background of some metals and its relationship with lithology and human activities was addressed with geographically weighted principal component analysis (GWPCA).
MATERIAL AND METHODS

3.1 Study area

Our research focuses on the Principality of Asturias, a mountain region in the north of the Iberian Peninsula, NW of Spain (Fig.1) that covers an area of about 10,600 km$^2$ on the North face of the Cantabrian Range. It dates from the Hercynian Orogenic cycle, but its relief was rejuvenated during the Alpine cycle and runs parallel to the coast of the Cantabrian Sea following an East-West direction.

The climate of the area is included within the type known as "oceanic cold–temperate domain", with mild temperatures and abundant rainfall, being strongly influenced by its proximity to the Cantabrian Sea. Roughly, the average annual precipitation reaches values ranging between 700 and 1300 l/m$^2$. The temperature has an annual mean of about 13 ± 1 ºC and averages of 9 ± 1 ºC and 19 ± –1 ºC in the coldest and warmest months, respectively.

The vegetation is typical of Atlantic areas, composed mainly of deciduous forests (Quercus petraea subsp. Petraea (Matt.) Liebl., Quercus pyrenaica Willd and Fagus sylvatica L.) which grow on North-facing hillsides. Away from the industrial and urbanized areas, the land has been traditionally used for farming and livestock grazing. Traditional livestock farming has given rise to the repeated burning of extensive sectors (Vélez, 2000a) and the use of fire has caused severe degradation as far as the surface vegetation and soil are concerned (Fernández et al., 2005).

3.2 Geological and geomorphological setting

Geologically, a large part of the study area lies within Variscan Orogen which was divided into five zones according to the nature of the rocks, deformation features and metamorphic grade. Two of these zones are present in the study area: The Cantabrian
zone (CZ) and the West-Asturian Leonese zone (WALZ). The CZ is constituted by a Palaeozoic sequence thrust and folded during the Variscan orogeny. Lower Palaeozoic lithological units are predominantly siliciclastic but the content in carbonates increases substantially in the upper part of the sequence, of Devonian and Carboniferous age (Pérez-ESTÁUN et al., 2004). Also few and very small intrusive bodies are present in the area in relation to Late Variscan magmatism episodes (Fig. 2.). Unconformably over the Palaeozoic rocks lies the Mesozoic materials. This is composed of siliceous conglomerates and carbonate breccias and alternations of argillaceous sandstones, siltstones, clays, marls and limestones. Summarizing, the main lithologies are slate, sandstone, quartzite and several types of limestone; while other sedimentary rocks, clay stones and marls, are limited to low relief tertiary-mesozoic basins in the central part of the region.

The relief is very rough with steep slopes. The highest elevations of the area reach 2500 m (a.s.l.) and geomorphological processes such as fluvial, mass wasting and creep processes are present. At the bottom of the main valleys and covering most of the hillsides in the region there are several quaternary surface deposits which form the parent material in most of the soils researched in this paper. The sharp relief results in very young soils with properties very similar to the parent material.

3.3 Sources of heavy metals in Cantabrian Range

It is well known that soils and sediments contain heavy metals derived from the bedrock weathering or derived from anthropogenic sources. The impact of heavy metal pollution on ecosystems due to anthropogenic activities like smelting or mining activities has been frequently researched (Adriano, 1986; Sheppard et al., 2000; Facchinelli et al., 2001; Wei & Yang, 2010; Zhong et al., 2012). In Asturias, the soil heavy metals background is currently high. The geological features of the area exert a strong control.
on the heavy metals concentration and its variability. In this way, bedrock is the main factor that controls the heavy metals content but there are also mineralized areas near geological faults that are scattered throughout this region causing a high variability in concentration and distribution of the soil heavy metals. (Dallmeyer & Martinez Garcia, 1990; Loghman et al., 2013). The legacy of the historical mining activities in the region remains in the form of old and abandoned industrial installations and mining wastes. The sulphide ore district of Asturias consists of numerous deposits found in mountain areas spread over the Cantabrian Zone of the Iberian Massif. These deposits were mined for decades to obtain Galene, Sphalerite, Cinnabar, Orpiment, Realgar, Barite, etc. In those areas where years of metal sulphide mining has occurred, large amounts of wastes with high contents of trace elements were disposed on the surface (Loredo et al., 2006). These mining areas were avoided in our analysis, but there are numerous mining reports that identify geochemical anomalies associated with certain watersheds or lithologies (Dallmeyer & Martínez García, 1990) in this study the geochemical anomalies were interpreted as natural backgrounds.

In the other hand, atmospheric pollutants associated to industrial activities and mining smelting have affected the soils of the Principality of Asturias during the last century (Loredo, 2006; Ordóñez et al., 2013; Gonzalez-Fernandez et al., 2014). The pollutants have been able to reach very remote areas because long range atmospheric transport of heavy metals can lead to pollutant deposition even in supposedly pristine areas (Steinnes, 1987; Nriagu, 1989; Meyer et al., 2015). In Asturias there are eighty-two potentially polluting industrial installations, six of them correspond to power plants emitting great amounts of heavy metals into the atmosphere (Ordoñez et al., 2003; Gonzalez-Fernandez et al., 2014). The amount of emissions has been annually tested by the Spanish Environment Ministry for the last ten years. Anyway, it is important to note
that industrial emissions are not the only source of the release of heavy metals into the atmosphere. Forest fires remain a common although illegal practice affecting soil organic matter; hydrological properties and nutrients (Fernandez, 2005; Santín et al., 2008; Fernandez et al., 2015). The can also contribute to the suspended heavy metals in the atmosphere and consequently be an important source of soil heavy metals enrichment (Cerdà, 2011; McConnell & Edwards, 2008; Bodí et al., 2011; Pereira et al., 2013a) Figure 3 depicts the frequency of forest fires in the Principality of Asturias over the past 20 years. In this region livestock farming has given rise to the repeated burning of extensive mountain sectors for centuries and nowadays the fire management of the territory role of forest fire ashes as a vehicle of nutrients and contaminants in soils is being researched.

3.4 Sampling design

Soil sampling was designed taking into account the bedrock lithology, the position in the landscape and the vegetation cover. The samples were taken in granite; schists; quartzites; slates; limestones; mesozoic sandstones; mesozoic limestones and marls, claystones; mixed lithology (interbedded limestones; slates; and sandstones); sandstones plus conglomerates and sandstones plus limestone with coal layers from the Central Coal Basin (the coal mining district of Asturias); and fluvial sediments. These lithology classes represent the full range of bedrock lithologies present in Asturias (Fig.2). For population of soils (334) fourteen heavy metals were analyzed: Mn(ppm); Zn(ppm); Cr(ppm); Pb(ppm); Co(ppm); Ni(ppm); Cu(ppm); Ba(ppm); As(ppm); Mo(ppm); Ag(ppb); Cd(ppb); Sb(ppb); Hg(ppb). Also the position in the landscape was taken into account sampling soils both on slopes exposed and unexposed to prevalent winds and upstream of the villages in the floodplains. Finally, the samples were always
drawn from the same type of vegetation cover: heaths and natural meadows avoiding natural forests, agricultural and urban areas.

The sampling design was aimed at collecting the most pristine soils in Asturias and to explore the sources of diffuse heavy metals soil pollution. In each sampling point litter and roots were removed from the top of the soil. Samples were taken using a plastic shovel to avoid contaminating them and they were stored in plastic bags for transport to the laboratory.

3.5 Analytical procedures

The soil samples were dried at room temperature and the fine fraction (< 2 mm) was separated for analysis. The pH and the electrical conductivity were determined with the Multi 340i VWR device, in a mixture of soil and distilled water with a weight ratio 1:2.5 for pH (AFNORM FX standard-31-103, 1998) and 1:5 ratios for Electrical Conductivity (Guitián & Carballas, 1976). Determination of the texture was done using a Beckman Coulter, LS 13320, in a saturated paste prepared with a solution of sodium hexametaphosphate at 1% w/w, after removal of organic matter by adding 7% hydrogen peroxide. Total organic carbon (TOC) was determined by the method of loss of ignition (Pansu & Gautheyrou, 2006).

Heavy metals were extracted from the soil samples by means of a preparation consisting in adding 10 ml of H₂O, 10 ml of aqua regia, and 4 ml of HF to 0.25 g of a homogenized sample in a PTFE vessel. The vessels were sealed and digested in a microwave oven. After completing the digestion program, the vessels were cooled to room temperature and the solutions were filtered through 0.45-µm cellulose paper. Finally, the solutions were diluted with water to a final volume of 50 ml. Heavy metal concentrations were analyzed with a Perkin-Elmer ELAN 6000 Q-ICP-MS.
3.6 Preliminary statistical analysis

Before conducting the WGPCA, an exploratory analysis of the data was carried out in order to find relationships between heavy metal concentrations and soil properties. The linear correlation coefficients between heavy metals and soil properties were calculated. Organic matter percentages; pH; clay; silt and sand percentages and electrical conductivity were the soil properties analyzed.

Also, in order to explore the possibility of atmospheric deposition, a t-test between soils that were taken from slopes exposed and unexposed to prevailing winds was performed. This analysis was conducted with a population of 185 samples corresponding to soil from the slopes.

3.5. Geographically Weighted Principal Components Analysis (GWPCA)

Geographically weighted principal component analysis (GWPCA) is an extension of the classical principal component analysis (PCA) to geographic data that aims to account for a certain spatial heterogeneity in the data (Harris et al., 2011; Dempsar et al., 2013). While PCA analysis can provide information regarding global internal structure, it fails to consider that the covariance structure of the data can change spatially. In essence, GWPCA performs a local PCA analysis by considering a neighborhood around each spatial feature.

In GWPCA the covariances are weighted as a function of the distance between the feature object and the features in the neighborhood (Fotheringham et al., 2002). The GW covariance matrix is calculated as follows:

$$\Sigma(u,v) = X^T W(u,v) X$$

(1)

where $X$ is the $n \times p$ matrix of data, being $n$ the number of observations and $p$ the number of covariates, and $W(u,v)$ a $n \times n$ diagonal matrix of weights that depends on
location \((u, v)\). This weight matrix can be generated using a kernel function, for instance a Gaussian weighting function (Lloyd, 2010):

\[
w_{ij} = \exp\left[-0.5\left(\frac{d_{ij}}{\tau^2}\right)^2\right]
\]

(2)

where \(d_{ij}\) represents the distance between location \(i\) and \(j\), and \(\tau\) the bandwidth which determines the size of the neighborhood. According to (2), \(w_{ij}\) increases when the distance between observations \(i\) and \(j\) decreases. The bandwidth has a heavy influence on the weights, which is larger for small bandwidths. When the bandwidth is very large, the covariance matrix does not depend on the spatial location of the observation and WGPCA reduces to a standard PCA. The bandwidth \(\tau\) can be defined by the user or automatically determined using cross-validation. It can also be constant or variable (adaptive). Estimation of the optimal bandwidth is still a major challenge in GWPCA.

Similarly to what happens in standard PCA, the GW principal can be determined at each spatial location \((u_i, v_i)\) as follows (Joliffe, 2002):

\[
L(u_i, v_i)V(u_i, v_i)L^T(u_i, v_i) = \Sigma(u_i, v_i)
\]

(3)

where \(L(u_i, v_i)\) is a matrix of eigenvectors that represents the loadings of each variable on each principal component and \(V(u_i, v_i)\) the diagonal matrix of eigenvalues that represents the variances of the corresponding principal components. Component scores of the principal components are given by \(Z(u_i, v_i) = X(u_i, v_i)L(u_i, v_i)\).

Unlike PCA, GWPCA provides results for each location, allowing to analyze the spatial variability of the covariance structure of the data. Accordingly, maps of component scores, loadings, explained variance or winning variables (those with the highest absolute values of the loadings) can be represented for each principal component (Harris
et al., 2011). Furthermore, GWPCA allows estimating eigenvalues and their associated eigenvectors at unobserved locations.

3.7 Unit of analysis definition

The watersheds of the Principality of Asturias were used as the basic unit of analysis. The main reason for this election is that the orography of Asturias is very irregular (see figure 1) with steep slopes that are part of small watersheds drained by young river networks developing small fluvial plains. This relief pattern exerts a strong control on the location of industrial installations, villages and urban areas which are preferably located on the fluvial plains. It is expected that this pattern of landscape and uses affects the distribution of diffuse pollution. Moreover, the soil population analyzed in this research corresponds to soils from slopes (187) and from fluvial plains (147). The lithology of the fluvial plains, which controls the geogenetic heavy metals backgrounds, depends on the lithology of the watersheds and is a mixture of them. The territory as a whole can be organized into twenty-four watersheds (Fig.4).

For each watershed, the area occupied by each lithology was calculated and taken into account with validation purposes. In the same way, the watersheds were used to evaluate the atmospheric deposition associated to the industrial release of heavy metals or other possible sources such as the incidence of forest fires, which are common in Asturias.

RESULTS

4.1 Preliminary analysis

Descriptive statistics of the 334 samples studied show average values below normal values on the Earth’s crust shown in the last row of the same table, only Mn; Zn and Hg surpass slightly these thresholds. Nevertheless, it is important to note the high
heterogeneity of the studied soils. However, if attention is focused on the range and maximum values of the soil population Mn; Zn; Cr; Pb; Co; Ni; Cu; Ba; Sb and Hg surpass the crust normal values.

Table I. shows heavy metals vs. soil properties R-Pearson coefficients. It is important to highlight that Organic matter is positively and significantly correlated with Mn (R 0,29); Ba (R 0,33); Ag (R 0,21). Clay percentage and heavy metal contents are positively correlated particularly for Co (R 0,38) and Ni (R 0,35). Other important results are the significant and positive correlation between electrical conductivity and Cr (R 0,43) and Mo (R 0,20).

Regarding the t-test analysis to study relationships between metal concentration and soil exposure to the atmospheric pollution sources, Mn; Cr; Co; Ba and Mo have higher and significant mean values in exposed rather than unexposed slopes. This result seems to indicate that the exposed slopes are enriched in these heavy metals due to atmospheric deposition processes (Zender et al., 2003b; Yong et al., 2014).

4.2 Global Principal Components Analysis

Table II shows the results of PCA analysis. The proportion of variance explained by the first six PCs is approximately 57%. The variables with the highest loadings in PC1, which accounts for about 14% of the total variance, are Hg> Pb> Mo> Cd. In PC2 the metals Cu> Ni> Zn> Co have the highest loading values and explain 11% of variance. All these metals are probably related to mineral deposits and lithology, but the absence of spatial information in PCA does not allow to confirm this hypothesis. Similarly, those metals with the highest loadings in PC3 to PC6 (Mn> Ag in PC3; Ba in PC4; Sb> Co in PC5 and Cr> Cd in PC6) are probably related to different sources of diffuse pollution such as that associated to atmospheric deposition. In this way, the fact that Mn; Ba and Ag are correlated to soil organic matter (see Table I) and that Mn; Cr; Co;
Ba present significant increases in slopes exposed to dominant winds seems to indicate that the origin of this increased concentration is mainly atmospheric deposition.

4.3 Geographically Weighted Principal Components Analysis

An adaptive bandwidth algorithm with robust GWPCA was applied given that it provided the minimum covariance score when it was compared with fixed and adaptive bandwidths and non-robust GWPCA. According to Gollini et al. (2015), in a robust GWPCA each local covariance matrix is estimated using the robust minimum covariance determinant (Rousseeuw, 1985). For each location a matrix of six principal components was obtained as well as the corresponding proportion of variance explained by them and the winning variables (those with the highest loadings).

The majority of the soil samples account for between 55-65% of the variance in the data with an average of 58% reaching up to 80% in the same location. This percentage variance is greater in the center and eastern center of the region as well as at the western edge of the region (see Fig. 5).

In each location the variable with the highest loading (in absolute value), the so-called winning variable, in each GWPCs was determined. The winners (o winning variables) are different depending on the location and are closely related to the watersheds in which they appear. The spatial distribution of the winning variables and the most frequent winning variables associated to each watershed are represented in Figure 6 for GWPC1. The same variables are represented in Figure 7 for GWPC2.

DISCUSSION

Analyzing Table III it is possible to conclude that when the most extensive bedrock is schist the most frequent winning variables in GWPC1 are Cu, Co and Ni. In quartzites the most frequent winning variables are Ba, Hg and Ag, while in the case of limestone,
As and Pb are the winning variables. Similarly, Cr and Mo are present in the watersheds with coal beds. Then, all the metals with the highest loadings in GWPC1 are related to watershed bedrocks and could be interpreted as a geogenetic component. Generally, to evaluate if heavy metals in soils came from anthropogenic or natural sources, the anthropogenic metal enrichments index is used (Shotyk et al., 2000) which compares the relative abundance of a chemical element in a soil to the bedrock in the earth’s crust. The enrichment index was found not applicable to this research because the bedrock is highly fractured resulting in an enrichment of soil heavy metal associated to the presence of metal sulfide ores (Dallmeyer & Martinez Garcia, 1990; Ghrefat et al., 2011; Loghman et al., 2013) Under these conditions increased metal concentration has a geogenic rather than anthropogenetic origin.

We interpret GWPC2 as a variable related to atmospheric deposition. For Nriagu (1989) or Pacyna (1986) the most probable source of the gaseous and particulate atmospheric pollutants is human activity. Nowadays the emissions have fallen significantly but 5 years ago, only in the Caudal Watershed, more than 3000 kg of Zinc were released into the atmosphere according to the data available in the PRTR (Register of Emissions and Sources of Pollutants available on line). Other toxic heavy metals such as As (~1000kg); Cr (~ 2000Kg); Ni(~2500kg); Hg (~400kg) were released into the atmosphere by the power plants in the year 2003 alone. Many studies highlight the importance of industrial emissions in air quality including heavy metals derived from different industrial processes such as As, Cd, Ni, V, Pb and Zn (Moreno et al., 2006; Sánchez de la Campa et al., 2007; Ordoñez et al., 2003; Gonzalez-Fernandez et al., 2014). Taking into account that the atmosphere is an important transport route for natural and polluted compounds from their natural or anthropogenic emission sources to remote areas it is possible to consider atmospheric deposition as the second source of
soil heavy metals just after the geogenetic source. Nevertheless, heavy metals such as Ba, Mn, Sb or Cd which are not registered in the emissions of Asturias’s industries are the winning variables in several watersheds (Table 5). These watersheds correspond mainly to mountain areas where forest fires are very frequent (see Fig.3). In this way metals like Ba, which is accumulated by vegetation (Nanda et al., 1995; Chaney et al., 1997; Lamb et al., 2013), is the first winning value of GWPC2 in five watersheds (Table 5) and is present in six others. Similar behavior is shown by Mn, Cd and Sb. Although there are no analyses about the ash composition for Asturian fires, the analyses on Californian and Australian forest fires show an important concentration, especially of Ba and Mn in ashes (Plumlee et al., 2007; Hageman et al., 2008a; Hageman et al., 2008b). Also substantial redistribution of light white ash by wind has been observed in the first few days after fires in grassland environments (Pereira et al., 2014a). The high dispensability of the ashes could be the main cause of enrichment in heavy metals as Ba or Mn displayed by the soils in watersheds affected by forest fires.

**CONCLUSIONS**

The use of GWPCA presents clear advantages over standard PCA since the former provides information regarding the spatial distribution of the percentage of variance and the variables with most influence in each of the components, while this information is normally obscured using a global analysis. The analysis of the spatial variability of variance and loadings of the components allows a better comprehension of the relationships between the different variables under study across the study area.

The application of GWPCA to the analysis of the location of heavy metals in the Principality of Asturias has allowed to establish a relationship between the concentration of some of these metals and the soil characteristics or the presence of forest fires.
ACKNOWLEDGEMENTS

This work has been partially funded by the Government of the Principality of Asturias (reference SV-PA-10-16. CN-10-089).

REFERENCES


Dallmeyer RD, Martínez García E. 1990. Pre-Mesozoic geology of Iberia : (papers from an international conference hosted by IGCP Project 233 (terrane in the Circum-Atlantic Paleozoic Orogens) in sept. 1986 in Oviedo, Spain). Springer- Verlag:

http://mc.manuscriptcentral.com/ldd
Berlín.


Nanda PBA, Dushenkov V, Motto H, Raskin I. 1995. Phytoextraction: The Use of Plants to Remove Heavy Metals from Soil Environmental Science & Technology **29**: 1232-1238. DOI: 10.1021/es00005a014


Ordonez A, Loredo J, De Miguel E, Charlesworth S. 2003. Distribution of heavy metals in the street dusts and soils of an industrial city in Northern Spain. *Archives of*
Environmental Contamination and Toxicology 44: 160-170. DOI:


DOI:10.1016/j.scitotenv.2013.08.090

### TABLES

#### Table I. R-Pearson coefficients

<table>
<thead>
<tr>
<th></th>
<th>Ba ppm</th>
<th>As ppm</th>
<th>Mo ppm</th>
<th>Ag ppm</th>
<th>Cd ppm</th>
<th>Sb ppm</th>
<th>Hg ppm</th>
<th>Mn ppm</th>
<th>Zn ppm</th>
<th>Cr ppm</th>
<th>Pb ppm</th>
<th>Co ppm</th>
<th>Ni ppm</th>
<th>Cu ppm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>pH</strong></td>
<td>0·06</td>
<td>0·00</td>
<td>0·05</td>
<td>-0·20</td>
<td>0·05</td>
<td>-0·03</td>
<td>-0·02</td>
<td>0·13</td>
<td>0·08</td>
<td>0·27</td>
<td>-0·02</td>
<td>0·28</td>
<td>0·40</td>
<td>0·16</td>
</tr>
<tr>
<td><strong>MO%</strong></td>
<td>0·33</td>
<td>0·01</td>
<td>-0·07</td>
<td>0·21</td>
<td>0·14</td>
<td>-0·05</td>
<td>0·01</td>
<td>0·29</td>
<td>0·11</td>
<td>-0·16</td>
<td>0·05</td>
<td>-0·10</td>
<td>-0·08</td>
<td>-0·03</td>
</tr>
<tr>
<td><strong>Clay%</strong></td>
<td>-0·03</td>
<td>0·07</td>
<td>0·16</td>
<td>-0·13</td>
<td>-0·06</td>
<td>-0·06</td>
<td>0·09</td>
<td>0·19</td>
<td>0·04</td>
<td>0·27</td>
<td>-0·04</td>
<td>0·38</td>
<td>0·35</td>
<td>0·07</td>
</tr>
<tr>
<td><strong>Silt%</strong></td>
<td>-0·11</td>
<td>0·11</td>
<td>0·13</td>
<td>0·02</td>
<td>0·06</td>
<td>-0·06</td>
<td>0·10</td>
<td>-0·18</td>
<td>0·03</td>
<td>0·21</td>
<td>-0·03</td>
<td>-0·16</td>
<td>-0·10</td>
<td>-0·01</td>
</tr>
<tr>
<td><strong>Sand%</strong></td>
<td>0·10</td>
<td>-0·12</td>
<td>-0·18</td>
<td>0·06</td>
<td>0·08</td>
<td>0·08</td>
<td>-0·12</td>
<td>0·03</td>
<td>-0·04</td>
<td>-0·30</td>
<td>0·04</td>
<td>-0·08</td>
<td>-0·11</td>
<td>-0·03</td>
</tr>
<tr>
<td><strong>Cond(μS)</strong></td>
<td>-0·08</td>
<td>-0·03</td>
<td>0·20</td>
<td>-0·02</td>
<td>-0·03</td>
<td>-0·01</td>
<td>0·03</td>
<td>-0·14</td>
<td>0·01</td>
<td>0·43</td>
<td>-0·03</td>
<td>-0·06</td>
<td>0·10</td>
<td>0·13</td>
</tr>
</tbody>
</table>

#### Table II: Summary of the global PCA.

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>PC6</th>
<th>PC7</th>
<th>PC8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard deviation</strong></td>
<td>14·188</td>
<td>12·205</td>
<td>11·151</td>
<td>10·806</td>
<td>10·352</td>
<td>10·085</td>
<td>0·9630</td>
<td>0·9473</td>
</tr>
<tr>
<td><strong>Proportion of variance</strong></td>
<td>0·1438</td>
<td>0·1064</td>
<td>0·0888</td>
<td>0·0834</td>
<td>0·0766</td>
<td>0·0726</td>
<td>0·0662</td>
<td>0·0641</td>
</tr>
<tr>
<td><strong>Cumulative Proportion</strong></td>
<td>0·1438</td>
<td>0·2502</td>
<td>0·3390</td>
<td>0·4224</td>
<td>0·4990</td>
<td>0·5716</td>
<td>0·6378</td>
<td>0·7019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loadings</th>
<th>Mn</th>
<th>-0·03</th>
<th>0·14</th>
<th>-0·62</th>
<th>0·23</th>
<th>-0·03</th>
<th>0·04</th>
<th>0·19</th>
<th>0·45</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zn</td>
<td>-0·14</td>
<td>0·40</td>
<td>0·29</td>
<td>0·08</td>
<td>-0·16</td>
<td>0·35</td>
<td>-0·11</td>
<td>-0·21</td>
</tr>
<tr>
<td></td>
<td>Cr</td>
<td>-0·20</td>
<td>-0·26</td>
<td>-0·21</td>
<td>-0·28</td>
<td>0·11</td>
<td>0·56</td>
<td>0·13</td>
<td>-0·37</td>
</tr>
<tr>
<td></td>
<td>Pb</td>
<td>0·48</td>
<td>0·13</td>
<td>0·27</td>
<td>0·01</td>
<td>-0·03</td>
<td>0·10</td>
<td>-0·12</td>
<td>-0·09</td>
</tr>
<tr>
<td></td>
<td>Ba</td>
<td>0·24</td>
<td>-0·05</td>
<td>0·08</td>
<td>0·61</td>
<td>-0·14</td>
<td>-0·03</td>
<td>0·27</td>
<td>-0·21</td>
</tr>
<tr>
<td></td>
<td>Co</td>
<td>0·09</td>
<td>0·37</td>
<td>0·03</td>
<td>0·25</td>
<td>0·43</td>
<td>-0·12</td>
<td>0·09</td>
<td>-0·26</td>
</tr>
<tr>
<td></td>
<td>Ni</td>
<td>0·16</td>
<td>-0·45</td>
<td>0·14</td>
<td>0·11</td>
<td>-0·16</td>
<td>-0·06</td>
<td>-0·59</td>
<td>0·17</td>
</tr>
<tr>
<td></td>
<td>Cu</td>
<td>0·06</td>
<td>-0·50</td>
<td>-0·12</td>
<td>0·02</td>
<td>0·21</td>
<td>-0·24</td>
<td>0·16</td>
<td>-0·44</td>
</tr>
<tr>
<td></td>
<td>As</td>
<td>0·10</td>
<td>-0·31</td>
<td>0·23</td>
<td>0·21</td>
<td>-0·34</td>
<td>0·28</td>
<td>0·50</td>
<td>0·20</td>
</tr>
<tr>
<td></td>
<td>Mo</td>
<td>0·46</td>
<td>0·08</td>
<td>-0·20</td>
<td>-0·27</td>
<td>-0·04</td>
<td>-0·11</td>
<td>-0·01</td>
<td>0·05</td>
</tr>
<tr>
<td></td>
<td>Ag</td>
<td>-0·02</td>
<td>0·06</td>
<td>-0·47</td>
<td>0·34</td>
<td>-0·38</td>
<td>0·18</td>
<td>-0·40</td>
<td>-0·35</td>
</tr>
<tr>
<td></td>
<td>Cd</td>
<td>0·39</td>
<td>0·03</td>
<td>-0·03</td>
<td>-0·14</td>
<td>0·14</td>
<td>0·53</td>
<td>-0·02</td>
<td>0·16</td>
</tr>
<tr>
<td></td>
<td>Sb</td>
<td>-0·07</td>
<td>-0·18</td>
<td>0·05</td>
<td>0·41</td>
<td>0·63</td>
<td>0·27</td>
<td>-0·23</td>
<td>0·24</td>
</tr>
<tr>
<td></td>
<td>Hg</td>
<td>0·49</td>
<td>-0·00</td>
<td>-0·23</td>
<td>-0·08</td>
<td>0·10</td>
<td>0·01</td>
<td>0·05</td>
<td>-0·14</td>
</tr>
</tbody>
</table>

http://mc.manuscriptcentral.com/ldd
Table III: Winning variables and the percentage of surface covered by each lithology grouped by watersheds

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Agueira</th>
<th>CaresDeva</th>
<th>Caudal</th>
<th>CostaCOR</th>
<th>CostaCOC</th>
<th>CostaOR</th>
<th>Eo</th>
<th>Esva</th>
<th>Gueña</th>
<th>Ibias</th>
<th>NalónH</th>
<th>NalónM-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Ha</td>
<td>26945</td>
<td>40634</td>
<td>94835</td>
<td>45677</td>
<td>17263</td>
<td>33799</td>
<td>26548</td>
<td>46555</td>
<td>14699</td>
<td>38628</td>
<td>27106</td>
<td>68467</td>
</tr>
</tbody>
</table>

**Lithology (%)**

<table>
<thead>
<tr>
<th>Lithology</th>
<th>Granite</th>
<th>Schistos</th>
<th>Cuarzites</th>
<th>Slates</th>
<th>Limestones</th>
<th>Mesozoic sandstones</th>
<th>Marls, Claystones</th>
<th>Mixed</th>
<th>Mixed, Carbon CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>1</td>
<td>88</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>27</td>
<td>52</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>90</td>
<td>6</td>
<td>63</td>
<td>1</td>
<td>37</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td>37</td>
<td>1</td>
<td>37</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>55</td>
<td>8</td>
<td>37</td>
<td>1</td>
<td>37</td>
<td>34</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>28</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>37</td>
<td>34</td>
<td>1</td>
</tr>
</tbody>
</table>

**Winning variables Pc1**

<table>
<thead>
<tr>
<th>wv1</th>
<th>Cu</th>
<th>As</th>
<th>Pb</th>
<th>Mo</th>
<th>Hg</th>
<th>As</th>
<th>Cu</th>
<th>Ag</th>
<th>Pb</th>
<th>Ba</th>
<th>Pb</th>
<th>Hg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>wv2</th>
<th>Co</th>
<th>Hg</th>
<th>Cr</th>
<th>Pb</th>
<th>Ba</th>
<th>Pb</th>
<th>Co</th>
<th>Ba</th>
<th>As</th>
<th>Mo</th>
<th>Ba</th>
<th>Pb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>wv3</th>
<th>Ni</th>
<th>Pb</th>
<th>Hg</th>
<th>Sb</th>
<th>Cr</th>
<th>Hg</th>
<th>Cr</th>
<th>Ni</th>
<th>Hg</th>
<th>Sb</th>
<th>Hg</th>
<th>Cr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Winning variables Pc2**

<table>
<thead>
<tr>
<th>wv1</th>
<th>Ag</th>
<th>Ni</th>
<th>Zn</th>
<th>Co</th>
<th>Co</th>
<th>Ni</th>
<th>Ag</th>
<th>As</th>
<th>Ba</th>
<th>Mn</th>
<th>Ba</th>
<th>Co</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>wv2</th>
<th>Ba</th>
<th>Cr</th>
<th>Cu</th>
<th>Ag</th>
<th>Ag</th>
<th>Cr</th>
<th>Ni</th>
<th>Cr</th>
<th>Sb</th>
<th>As</th>
<th>As</th>
<th>As</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>wv3</th>
<th>Sb</th>
<th>Hg</th>
<th>Hg</th>
<th>Ni</th>
<th>Sb</th>
<th>Hg</th>
<th>Sb</th>
<th>Ba</th>
<th>Ni</th>
<th>Ni</th>
<th>Ni</th>
<th>Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watershed</td>
<td>NarceaH</td>
<td>NarceaM-L</td>
<td>NaviaH</td>
<td>NaviaL</td>
<td>Negro</td>
<td>Nora</td>
<td>Pigueña</td>
<td>Piloña</td>
<td>Porcia</td>
<td>SellaH</td>
<td>SellaL</td>
<td>Trubia</td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Total Ha</td>
<td>90917</td>
<td>52936</td>
<td>48751</td>
<td>28000</td>
<td>17137</td>
<td>38112</td>
<td>40449</td>
<td>51125</td>
<td>20201</td>
<td>44983</td>
<td>14483</td>
<td>47913</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lithology (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Granite</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schistes</td>
<td>46</td>
<td>19</td>
<td>75</td>
<td>59</td>
<td>60</td>
<td>1</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuarzites</td>
<td>47</td>
<td>40</td>
<td>25</td>
<td>41</td>
<td>40</td>
<td>26</td>
<td>21</td>
<td>44</td>
<td>15</td>
<td>32</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Slates</td>
<td>0</td>
<td>17</td>
<td>8</td>
<td>18</td>
<td>25</td>
<td>21</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limestones</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>34</td>
<td>16</td>
<td>1</td>
<td>50</td>
<td>34</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesoic sandstones</td>
<td>2</td>
<td>2</td>
<td>52</td>
<td>14</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marls, Claystones</td>
<td></td>
<td>27</td>
<td>17</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>14</td>
<td></td>
<td>8</td>
<td>31</td>
<td>5</td>
<td>9</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandstones, Carbon CCC</td>
<td>4</td>
<td>1</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed, Carbon CCC</td>
<td>2</td>
<td>1</td>
<td></td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Winning variables Pc1**

<table>
<thead>
<tr>
<th>wv1</th>
<th>Ba</th>
<th>Ba</th>
<th>Cu</th>
<th>Cu</th>
<th>Ni</th>
<th>Mo</th>
<th>Ba</th>
<th>Mo</th>
<th>Cu</th>
<th>Ni</th>
<th>Ni</th>
<th>Cr</th>
</tr>
</thead>
<tbody>
<tr>
<td>wv2</td>
<td>Zn</td>
<td>Hg</td>
<td>Sb</td>
<td>Co</td>
<td>Ag</td>
<td>Hg</td>
<td>Zn</td>
<td>Pb</td>
<td>Co</td>
<td>Pb</td>
<td>Pb</td>
<td>Hg</td>
</tr>
<tr>
<td>wv3</td>
<td>As</td>
<td>Cr</td>
<td>Sb</td>
<td>Ni</td>
<td>Sb</td>
<td>Pb</td>
<td>Mn</td>
<td>Hg</td>
<td>Ni</td>
<td>Hg</td>
<td>As</td>
<td>Pb</td>
</tr>
</tbody>
</table>

**Winning variables Pc2**

<table>
<thead>
<tr>
<th>wv1</th>
<th>Mn</th>
<th>Cu</th>
<th>Ba</th>
<th>Ag</th>
<th>Co</th>
<th>Mn</th>
<th>Ba</th>
<th>Ni</th>
<th>Ag</th>
<th>Ba</th>
<th>Ni</th>
<th>Co</th>
</tr>
</thead>
<tbody>
<tr>
<td>wv2</td>
<td>Cd</td>
<td>Ni</td>
<td>As</td>
<td>Sb</td>
<td>Ba</td>
<td>Zn</td>
<td>Mn</td>
<td>Cu</td>
<td>Ni</td>
<td>Mo</td>
<td>Mo</td>
<td>Ni</td>
</tr>
<tr>
<td>wv3</td>
<td>Ni</td>
<td>As</td>
<td>Cd</td>
<td>Ba</td>
<td>Pb</td>
<td>Sb</td>
<td>Cd</td>
<td>Ba</td>
<td>Sb</td>
<td>Co</td>
<td>Ba</td>
<td>As</td>
</tr>
</tbody>
</table>
FIGURE CAPTIONS

Figure 1: The Principality of Asturias location and hillshade map.

Figure 2: Principality of Asturias geological map.

Figure 3: The map shows fires identified in Asturias with Landsat images in a long-time period.

Figure 4: Sample points and watershed distribution.

Figure 5: Covariance scores for different bandwidth determination (left). Percentage of total variance for local components 1 to 6 (right).