1. Introduction

Some diseases such as allergic asthma may be mainly caused by presence in lung of Ascaris lumbricoides (which causes ascariasis), the Ancilostoma, and Strongiloides (causing to Strongyloidiasis) that cause or exacerbate respiratory symptoms, especially in the lung causing coughing, wheezing, and dyspnea (shortness of breath). Some states of allergic contact dermatitis, are also often associated with direct contact with some type of allergen, or substance that the body identifies as dangerous, producing a rash in the local where the contact occurred. Therefore these diseases are important to public health considering the strong social component directly related to poverty and lack of primary health care in areas of high humidity, high concentration of waste and can often be transmitted by direct contact with degraded areas with contaminated soils (Kakkar and Jaffery, 2005).

Soil often acts as a filter to a large part of the impurities deposited in it. However, this capacity is limited, causing an accumulation of material resulting from atmospheric deposition of pollutants, pesticides and fertilizers (Moreira-Nodermann, 1987). Soil contamination is threat to human health and for environmental quality. Among the main pollutants of the soil, heavy metals are very dangerous when in contact with living beings (Lourenço et. al. 2010; Lourenço and Landim, 2005; Alloway, 2001; Franssen et. al. 1977). In fact, several metals are known to be carcinogens, including arsenic, chromium and nickel. (Tang et al., 1999; Winneke et al., 2002; Stein et al., 2002; Yang et al., 2003).

Studies of the spatial distribution of pollutants in air, water and soil, are traditionally carried out by different scientists in different fields of geosciences using different spatial analysis techniques in order to contribute to the understanding of the variability space of certain events that cause damage to the environment and health (Lourenço et. al., 2010; Amini et. al. 2005). Goria et. al (2009) conducted a study in four French administrative departments and highlighted an excess risk in cancer morbidity for residents around municipal solid waste incinerators. The steps to evaluate the association between the risk of cancer and the
exposure to incinerators, was performed by statistical analysis and dispersion modeling using GIS. The study showed that is important to use advanced methods to better assess dose-response relationships with disease risk. Bilancia and Fedespina (2009) studied the triennial mortality rates for lung cancer in the two decades 1981–2001 in the province of Lecce, Italy. The study showed that there is a dramatic increase in mortality for both males and females. Vincenti et. al. (2009) examined the relation between exposure to the emissions from a municipal solid waste incinerator and risk of birth defects in a northern Italy community, using Geographical Information System (GIS). Among women residing in the areas with medium and high exposure, prevalence of anomalies in the offspring was substantially comparable to that observed in the population control, nor dose-response relations for any of the major categories of birth defects emerged. McGrath et al. (2004) produced maps of pollution based on the spatial distribution of Pb in Silvermines, Ireland, where the generated maps serve as valuable information on areas of risk to public health and as decision support and planning. Critto et al. (2003) used geostatistics and the main components of the distribution of chemical contaminants in the soil around a lake near Venice, Italy and evaluated their effects on health. Lin et al. (2002) used the methods to factorial kriging and indicator kriging to analyze the spatial variation of heavy metals in farmland north of Changhua, Taiwan in order to assist in monitoring for environmental remediation proposals and planning. Hills and Alexander (1989) studied surveys which presented the occurrence of leukemia near nuclear plants, and Glass et al. (1995) produced a risk map for the Lyme disease from epidemiological data and from a geographic information system, Mason (1975) presented several field studies conducted as a result of issues related to environmental determinants of cancer that has been raised after the analysis of several atlas published by the American National Cancer Institute. We can also cite important studies oral cancer (Winn et al., 1981), cancer of the bowel (brine et al., 1981), lung cancer (Ziegler et al., 1984), bladder cancer (Hoover and Strasser , 1980) and, finally, studies of associations between sources of contamination and high risk areas, including risk of childhood leukemia in areas near nuclear power plants (Diggle et al., 1990, Elliot et al., 1992) . Given the present discussion, the aim was to study the spatial correlation between the distribution of contaminants in the soil with the spatial distribution of infant morbidity in children under one year of age affected by diseases of the respiratory and intestinal tract in the city of Cubatão, southern coast of São Paulo, Brazil.

2. Material and methods

2.1 Studied area soil sampling

The research was conducted in the city of Cubatão, southern coastal region of the State of São Paulo, Brazil. The studied area has strong industrial activities in the area with a big petroleum refinery and various chemical activities. Small factories are concentrated in the center area near of port region. There are several wastewater treatment plants in the region with high risk of pollution associated to sewage sludge and compounds.

In this study we analyzed the pollutants cadmium-Cd, lead-Pb and mercury-Hg. Cd is a trace element in various industrial uses, such as fungicides, batteries, rubber processing, production of pigments and galvanic industries, among others. Once the Pb is a toxic element and occurs as an environmental pollutant, is given its use in industrial large scale in the petroleum industry, dyes and paints, ceramics, and others. Both Cd and Pb cause serious health problems to people when exposed to them, or by eating contaminated food. Can
cause problems with anemia, infections, headache, sweating (sweat) and various muscle aches. The most serious consequence of chronic exposure to Cd and Pb is cancer, especially cancer of the airways, causing pulmonary emphysema (Okada, et. al, 1997). The Hg is a metal and odorless liquid at room temperature, but when the temperature increases becomes toxic and corrosive vapors denser than air. Integrate the class of transition metals. The risk of disease is high. According to the temperature, the concentration of metallic mercury is changed and when absorbed by the human body tends to accumulate in the brain, liver and kidneys. Because of this, contamination manifests itself by acute problems in the nervous system (sensory and motor disturbances) and deficiencies of bowel function (Zavaris and Glina, 1992).

To determine total metal concentrations Cd, Pb and Hg, the soil samples were sampled with distances from 95 to 650m (Figure 1). After the soil samples were dried and conventionally decomposed by a mixture of nitric acid and concentrated hydrochloric acid according to a standardized procedure (Alloway, 2001). After that, it was weighed (approximately 2.00 g) of pre-dried soil sample that was mixed with 21 mL of 30% HCl (Suprapur) and 7 mL of 65% HNO₃ (Suprapur) in a highly pre-purified quartz vessel (200 mL). The solution was heated first to 100°C and then to 120°C. Subsequently, the samples were digested using 20 mL of concentrated HNO₃ under reflux for 3 hours. Finally, the digested samples were diluted with high-purity water to a final volume of 100 mL. Small undigested soil remainders (approximately 5%) were removed by filtration. Metal determinations were usually carried out with 1/10 dilutions of the digestion solutions. The result of soil digestion by aqua regia was assessed from five replicates and metal determinations were performed by ICP-OES (spectrometer: TJA IRIS AP, Thermo Jarrell Ash, Franklin, MA, USA).
2.2 Mapping soil metal concentrations
To the mapping of total metal concentrations Cd, Pb and Hg of the soil samples it was used geostatistical methods (Van Meirvenne and Goovaerts, 2001; Webster and Oliver, 2001; Lin et. al. 2001; Romic and Romic, 2003; McGrath et. al. 2004). Geostatistics analysis methods are based in the spatial variation of data often distributed irregular, known as a regionalized variable. Therefore, for a geostatistical modeling it is used an interpolation known as kriging. The procedure is similar to that used in weighted moving average interpolation, except that the weights are derived from a variography analysis of the data model. The weights are given by:

\[
\hat{z}(X_0) = \sum_{i=1}^{n} \lambda_i \cdot z(x_i)
\]  

(1)

with \( \sum_{i=1}^{n} \lambda_i = 1 \). The weights \( \lambda_i \) are chosen so that the estimate \( \hat{z}(X_0) \) has less variance \( \sigma_e^2 \) in relation to the sampled values compared to any other linear estimator.

The minimum variance of \([z(X_0) - z(X_0)]\), the prediction error, or ‘kriging variance’ is given by:

\[
\hat{\sigma}^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \phi
\]

(2)

and is obtained when

\[
\sum_{i=1}^{n} \lambda_i \gamma(x_i, x_j) + \phi = \gamma(x_j, x_0) \quad \text{for all } j
\]

(3)

The \( \gamma(x_i, x_j) \) is the semivariance of \( z \) between the sampling points \( x_i \) and \( x_j \); \( \gamma(x_i, x_0) \) is the semivariance between the sampling point \( x_i \) and the unvisited point \( x_0 \). Both semivariance are obtained from the fitted variogram. The semivariance \( \phi \) is a Lagrange multiplier required for the minimization. This method is known as ordinary kriging and it is very well described by many authors (eg. Landim, 2003; Gringarten and Deutsch, 2001; Olea, 1999; Burrough et. al. 1997; Goovaerts, 1997; Isaaks and Srivastava, 1989; Journel and Huijbregts, 1978).

Thus, in this study, variogram models were used to analyze spatial patterns and ordinary kriging to obtain a continue surface of the distribution of soil pollutants Cd, Pb and Hg in the area.

2.3 Mapping morbidity
Morbidity was determined as the health damages to the movement of hospitalization and outpatient care of the study area. The data of the studied area were provided by the Health Brazilian Agency in the year 2007. These data were filtered to obtain only the data of hospital admissions according to the 10th revision of International Classification of Diseases (ICD-10), for hospitalizations related to some kind of disease that can be caused by direct or indirect contact with contaminated soil with high concentration of pollution.
The cases of hospitalization and outpatient care of the study area were used for construction of the discrete map by district and surface map of the morbidity distribution using interpolation of the inverse of the distance with power squared (Burrough, 2004). The equation used for Inverse Distance to a Power (IDP) is:

$$\hat{Z}_j = \frac{\sum_{i=1}^{n} Z_i h_{ij}^{-\beta}}{\sum_{i=1}^{n} 1}$$

where:
- $h_{ij}$ is the effective separation distance between grid node "j" and the neighboring point "i."
- $\hat{Z}_j$ is the interpolated value for grid node "j";
- $Z_i$ are the neighboring points;
- $d_{ij}$ is the distance between the grid node "j" and the neighboring point "i";
- $\beta$ is the weighting power (the Power parameter); and
- $\delta$ is the Smoothing parameter.

This procedure was carried out in order to obtain a surface continue distribution morbidity that could be compared in a space with other maps of soil pollutants.

2.4 Measures of spatial relationship between morbidity and soil pollution

To analyze the relationship between the morbidity spatial maps distribution and the soil maps was used the multiple regression spatial analysis technique. There are many cases where the variation of a variable can be explained by a number of other variables. The variables that help predict the variable of interest are called the independent variable, while the predicted variable is called dependent variable, assuming that a linear relationship exists between them. This is study used the independent variables to the soil pollution and dependent variable to the morbidity by multiple linear regression and the multiple linear regression equation is written as:

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3$$

where $Y$ is the dependent variable; $x_1$, $x_2$, and $x_3$ are the independent variables; $a$ is the intercept; and $b_1$, $b_2$, and $b_3$ are the coefficients of the independent variables $x_1$, $x_2$, and $x_3$, respectively. The intercept represents the value of $Y$ when the values of the independent variables are zero, and the parameter coefficients indicate the change in $Y$ for a one-unit increase in the corresponding independent variable.

In the multiple regression results the $R$ represents the multiple correlation coefficients between the independent variables and the dependent variable. $R$ squared represents the extent of variability in the dependent variable explained by all of the independent variables. The adjusted $R$ and $R$ squared are the R and R squared after adjusting for the effects of the number of variables.
The individual contribution of each independent variable to the individual dependent variable is expressed in the regression coefficients. Its significance is expressed in the form of a \( t \)-statistic. The \( t \)-statistic is the most common test used in estimating the relative success of the model and for adding and deleting independent variables from a regression model. The \( t \)-statistic verifies the significance of the variables departure from zero (i.e., no effect).

Multiple spatial linear regressions allow the construction of spatial maps in predicting morbidity and a map of the residual according to the linear model fitted. As the map of prediction has its variation as a function of predictor variables it can be understood as a risk map from the exposure of pollutants to the occurrence of morbidity, and the residual map as a measure of success of prediction.

3. Results and discussion

Table 1 shows the variogram parameters for soil samples after chemical analysis.

<table>
<thead>
<tr>
<th>Soil attributes</th>
<th>Model</th>
<th>( C_o )</th>
<th>( C + C_o )</th>
<th>( C_o / C + C_o )</th>
<th>Range (m)</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cd</td>
<td>Gaussian</td>
<td>0.15</td>
<td>0.80</td>
<td>0.187</td>
<td>395</td>
<td>0.38</td>
</tr>
<tr>
<td>Hg</td>
<td>Gaussian</td>
<td>0.55</td>
<td>0.95</td>
<td>0.578</td>
<td>1100</td>
<td>0.30</td>
</tr>
<tr>
<td>Pb</td>
<td>Exponential</td>
<td>0.5</td>
<td>0.96</td>
<td>0.520</td>
<td>410</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table 1. Variogram models of heavy metals and their parameters. \( C_o \) = nugget variance, \( C \) = structural variance, and \( (C + C_o) \) = sill variance.

The range values of variograms for Cd and Pb were similar and around 400m, and were lower than those for Hg (around 1100m). The Nug/Sill ratio for the Cd metal was around the 18% (\( C_o / C + C_o \)) showing randomness of the data that is important for good modeling variography while for the values of Hg and Pb variation is more random and unpredictable. The \( R^2 \) was between 0.30 and 0.52 suggesting a good correlation between samples.

The experimental variograms of the heavy metal in soil with the fitted models are presented in Fig. 2 a-c.

![Gaussian variogram of Cd with fitted models](www.intechopen.com)
The results showed that soil with Cd (a), Hg (b) were best fitted with the gaussian model and Pb (c) with the exponential model. The ordinary kriging technique was used here to obtain a surface of the spatial distribution of soil pollution fitted with parameters of the variogram (Figure 3).
Fig. 3a. Prediction mapping of Cd concentration in soil generated by ordinary Kriging

Fig. 3b. Prediction mapping of Hg concentration in soil generated by ordinary Kriging
The Figure 3 a-c presents the spatial patterns of the three heavy metals in soil from the studied area generated from their variograms. The spatial distribution maps showed similar geographical trends, especially for Cd (a) and Hg (b), with higher concentration in the west area and decreasing presence towards northeast. Meanwhile despite Pb (c) showed similar spatial trend, the intensity is higher in west area and also the southern area is emerging as an important local pollution.

3.1 Morbidity map
Two maps of the spatial distribution of cases of hospitalization and outpatient care of the study area were constructed. The first, which contains all the cases for the year 2007 divided by neighborhoods (Figure 4a) and a second map (Figure 4b) constructed with Inverse Distance to a Power technique to obtain a surface of the spatial distribution of soil pollution.

The map of Figure 4a of spatial distribution of morbidity by neighborhood showed that concentrations are localized in the west area. This area is very industrialized beyond to be place with houses of poor and low social standing.

The map of Figure 4b of spatial distribution surface of morbidity showed that concentrations are localized in three different area: a coincident with the map of Figure 4a and two others, one near the central districts and other areas closer to the south of the map.

The southern sector is characterized by areas of proximity to the sea shore, with influences of the waters of the mangroves, which can be further more dangerous for people living on fishing and consumption of other foods from the sea.
3.2 Spatial relationship between morbidity and soil pollution

In order to analyze the relationship between the morbidity spatial distribution maps and the soil maps was used the spatial linear regression analysis. The Figures 5 to 7 show the regression graphs and the statistical parameters of spatial relationship.
Fig. 5. Graph of spatial linear regression between Morbidity and Cd.

Fig. 6. Graph of spatial linear regression between Morbidity and Hg.

Fig. 7. Graph of spatial linear regression between Morbidity and Pb.
The regression trend line shows the stronger linear relationship to the data at soil pollution with morbidity. The correlation coefficient (“r”) next to the equation tells us the same numerically. As can be seen in our data when the morbidity increases, the soil pollution also increases. In this study the correlation coefficients ranged around 0.65 indicating a strong positive relationship between soil pollution, morbidity and the coefficient of determination (r²) around 40%, which leads us to accept that the pollutants from the soil strongly influence the morbidity in the studied area.

However, we see that all pollutants were highly correlated with morbidity and it is unclear which of them would have a greater influence on it. To determine which pollutant found in the soil has a greater influence on the variation of morbidity was performed an analysis known as multiple spatial linear regression. The multiple spatial regression analysis is an important technique that permits the investigation of the relationship of spatial variables over the same sample space. There are many cases where the variation of a variable can be explained by a number of other variables. The variables that help to predict the variable of interest are called the independent variable, while the predicted variable is called dependent variable, assuming a linear relationship existing between them.

This is study used the independent variables to the soil pollution and dependent variable to the morbidity for linear multiple regression and the linear multiple regression equation is written as:

\[
\text{Morbidity} = 0.1326 + 0.4777 \times \text{Cd} - 0.0024 \times \text{Hg} + 0.0027 \times \text{Pb}
\]  

(7)

The regression equation shows coefficients for each of the independent variables and the intercept. The intercept (0.1326) can be thought of as the value for the dependent variable when each of the independent variables takes on a value of zero. The coefficients indicate the effects of each of the independent variables on the dependent variable. For example, if the emission of Cd increases by 100 units, increases the morbidity to 47.77% (i.e., 100 multiplied 0.4777). The multiple correlation spatial coefficient between the independent variables (ie, Cd, Hg and Pb) and the dependent variable (morbidity) was R = 0.91 and the extent of variability in the dependent variable explained by all of the independent variables was R² = 84%, i.e., 84% of the variance in the morbidity is explained by independent variables soil pollution.

The individual regression coefficients express the individual contribution of each independent variable to the dependent variable. The significance of the coefficient is expressed in the form of a \( t \)-statistic. The \( t \)-statistic verifies the significance of the variables departure from zero (i.e., no effect). In this study, the \( t \)-statistic has to exceed the following critical values in order for the independent variable be significant. To 99% confidence level with \( \infty \) degrees of freedom the value is 2.57, and Cd coefficient has a \( t \)-statistic of 8.71, the Pb \( t \)-statistic is 4.75 and the Hg \( t \)-statistic is 3.66 indicating that all variables are highly significant (99%). The \( t \)-statistic is the most common test used in estimating the relative success of the model and for adding and deleting independent variables from a regression model.

Multiple linear spatial regressions allow the construction of spatial maps in predicting morbidity and a map of the residual according to the linear model fitted (Fig. 8a-b). As the map of prediction has its variation as a function of predictor variables it can be understood as a risk map from the exposure of pollutants to the occurrence of morbidity, and the residual map as a measure of success of prediction.
The risk map areas (Figure 8a) with higher risk of morbidity are concentrated in the western area, coincidentally where is the greatest concentration of pollutants in soil. However, this is not the only place that appears morbidity in the area. This suggests that the variation in morbidity may have other factors, or other pollutants causing the variation of morbidity that is not being used in this study.
The analysis of the residual map (Figure 8b) is an important tool for spatial assessment of areas where the prediction obtained the best results depending on the model fitted. Usually the areas closest to the zero values have the best predictions, while the areas with the highest residual would be the worst predictions based on the model fitted. Thus, as seen in this work, the smaller residual occurs in areas with the predictions for higher morbidity risks.

4. Conclusions

The study showed that areas with high concentrations of pollutants in the soil influence the occurrence of morbidity especially those related to the intestinal tract and skin and respiratory allergies. In general, the studied area has serious problems related to the use and occupation by people with low purchasing power and, as a consequence, with little access to public health.

The proposed methodology was efficient for the purpose of showing that there is a degree of relationship between pollutants from soil and some cases of morbidity that can affect the health of people. This methodology may be useful for planning programs and management in promoting the welfare of people. This is possible through the identification of priority areas to assist people beyond the actions of government agencies to control the emission of pollutants into the environment.

Finally it is expected that the results, particularly the maps generated through the techniques of GIS, can be an important tool for urban planning and management, with main purpose to help improving the people quality of life.

5. References


The book entitled Water Quality, Soil and Managing Irrigation of Crops comprises three sections, specifically: Reuse Water Quality, Soil and Pollution which comprises five technical chapters, Managing Irrigation of Crops with four, and Examples of Irrigation Systems three technical chapters, all presented by the respective authors in their own fields of expertise. This text should be of interest to those who are interested in the safe reuse of water for irrigation purposes in terms of effluent quality and quality of urban drainage basins, as well as to those who are involved with research into the problems of soils in relation to pollution and health, infiltration and effects of irrigation and managing irrigation systems including basin type of irrigation, as well as the subsurface method of irrigation. The many examples are indeed a semblance of real world irrigation practices of general interest to practitioners, more so when the venues of these projects illustrated cover a fair range of climate environments.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
