Original Article

# Spatial linear regression models in the infant mortality analysis

Lourenço Manuel[1], João Domingos Scalon[2], Denismar Alves Nogueira[3] & Juracy Mendes Moreira[4]

[1]*PhD student in Department of Statistics at UFLA. Author for correspondence Email: zlouremanuel@gmail.com*
[2]*Associate professor at Federal University of Lavras. UFLA. Department of Statistics. Email: scalon@des.ufla.br*
[3]*Associate professor at University of Alfenas – Department of Statistics. Email: denismar@unifal-mg.edu.br*
[4]*Professor at Faculty of Montes Belos, São Luiz de Montes Belos-GO. Email: juraci.moreira@fmb.edu.br*

I N F O   A R T I C L E

A B S T R A C T

Infant mortality is one of the main concerns for governments in programs of public health. It is also an important measure used to evaluate the quality of life in several countries. The aim of this paper is threefold: first, to study the spatial distribution of infant mortality in Alfenas city, Minas Gerais using spatial lattice methods. Secondly, to propose a new method based on the square root transformation in the response variable of the spatial regression models in order to reach residuals with constant variance and normality. Thirdly, to identify the factors affecting infant mortality based on the proposed method. The response variable is "the number of deaths of infants under one-year-old", while the independent variables are "the number of women in fertile age", "the number of women in gestational risk age", "the number of illiterate women", "the monthly income of the woman and the men", "the number of residences with more than six inhabitants" and "the demographic density". All these variables are available for each census sector of Alfenas city, Minas Gerais. The spatial dependence of the number of deaths of infants under one-year-old has been assessed through the global and local Moran indexes. Furthermore, three models have been fitted, namely, the classic regression model, the spatial autoregressive model (SAR) and the conditional autoregressive model (CAR). The application of square root transformation in response variable allowed developing prediction equations for spatial models and reaching residuals with normal distribution and constant variance. The Akaike information criterion (AIC) has indicated SAR model as the best. The "number of women in fertile age" and "monthly income of the women" were identified as the main factors affecting significantly infant mortality in Alfenas city, Minas Gerais.

## 1. Introdução

Infant mortality (IM) is one of the main measures used to assess quality of life in several countries. It refers to deaths of young children, typically those less than one-year age. IM is measured by Infant Mortality Rate (IMR), which is the number of deaths of infants under one-year-old per 1000 live births, during a period of time (Genowska et al., 2015).

There are several factors pointed for high infant mortality rates. Genowska et al. (2015) points out that premature birth is the biggest contributor to the IMR while factors like birth asphyxia, pneumonia, neonatal infection and malnutrition are other causes of IM. The authors also claim that mother´s education level, environment conditions, political and medical infrastructure are important contributors' factors as well. On the other hand, Hall et al. (2016) describes smoking during pregnancy as the most common preventable causes of IM and other measures of public health such as, improving sanitation, access to clean drinking water, immunization against infectious disease can also help to reduce IMR.

Some authors like (Leal & Szwarcwald, 1997 and Menezes et al., 1998) support that infant mortality can be controlled by investments that lead to improved assistance for women and new-borns in the delivery room and nursery, socio-economic conditions of people with low income, education access, sanitation and basic health services. Furthermore, epidemiologic studies promoting knowledge about spatial distribution of infant mortality seems to be useful tools to control infant mortality as well. These studies allow the decision makers to identify critical regions with special needs, where positive interventions must be carried out in order to reduce infant mortality rates improving quality of life.

Silva et al. (2011) claims that neighbour regions are more likely to present similarity in infant mortality rates than neighbours in large distances. Therefore, infant mortality analysis must be carried out taking into account that level of similarity or dissimilarity between neighbour regions using spatial statistics methods. According to Cressie (2015), spatial statistic is divided in three different groups namely random

fields (geostatistics), point pattern process and lattice data. Random fields analyse phenomena distributed continually in space and point pattern processes are expressed by occurrences identified as points located in space. Lattice data analyse aggregated phenomena such as Gross Demand Product (GDP) or count data of occurrence of any human disease by states, counties, census sectors, blocks etc.

Most studies of infant mortality in Brazil are usually analysed by spatial aggregated data in areas because this type of data is often available in terms of count data by regions (Leal & Szwarcwald, 1997; Silva, et. al., 2011). There are two main sources of data to estimate infant mortality, the Birth Information System (SINASC) and Death Information System (SIM). The information for the former system is provided by hospitals where certificates of births are issued while the information for the later system comes from the number of certificates of deaths (D´orsi & Carvalho, 1998). These systems do not provide information about the exact location (coordinates) of the dead. This is probably one of the reasons why infant mortality is often analysed using aggregated data. Moreover, such representation by regions provides a feasible way to represent easily interpretable thematic maps of spatial distribution of infant mortality for decision makers (Andrade & Szwarcwald, 2001).

In the presence of some covariates associated to infant mortality, regression models can be built in order to predict infant mortality in other regions or for forecasting. For spatial data case, where spatial autocorrelation exists, the model must take into account the spatial structure in order to improve the predictions provided by the model. According to Druck (2004) there are two types of spatial regression models with spatial structure in lattice data: the spatial autoregressive (SAR) and the conditional autoregressive (CAR) model. The former incorporates the spatial dependence as a lag variable of the response while the later captures spatial autocorrelation as a nuisance.

This paper aimed to generate predicted maps of infant mortality distribution for a Brazilian city (Alfenas, MG) using spatial regression model which incorporates transformation in the response variable and identify regions of high risks in order to facilitate the decision makers in intervention actions.

## 2. Methods and Material

### Data and study area

The study was conducted in the urban area of Alfenas city, Minas Gerais. The city is located, at $21^o25´45´´$ Latitude South and $45^o56´50´´$ Longitude West (Köppen & Geiger, 1928). Alfenas is geographically divided in 68 census sectors with different shapes and sizes geographically coded by centroids (Figure 1).

The number of deaths of infants under one year old was used as the response variable (dependent) and, in the present paper, represents the infant mortality. These data were collected from Death Information System (SIM) between 2010 to 2014 period. Only cases of the dead geographically able to be coded (those containing address information like avenue, street or neighbourhood) were used. Given the limitation information about the total number of birth per census sectors in the study area, it was not possible to determine infant mortality rates. For this reason the study uses the number of deaths of infants under one year old as infant mortality itself not rates. Furthermore, a set of seven covariates were used as independent variables namely: the number of women in fertile age; the number of women in gestational risk age; the number of illiterate women; the monthly income of the women; the monthly income of the men; the number of residences with more than six inhabitants; and the demographic density of the census sector. All independent variables were collected by IBGE (2010) while the number of deaths of infants under one year old was

collected by one of the author of the study (D.A.N). All variables were collected for each of the 68 census sectors. Thus, the study did not use biological essays with human people, so there was not necessary a certificate from Ethic and Research Committee.
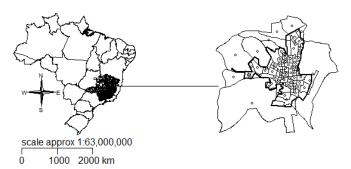


**Figure 1.** Urban area of Alfenas city, MG divided in 68 census sectors geographically coded by centroids. Source: Estaticart, IBGE 2010

### Spatial Dependence Analysis

Spatial dependence between observations of the number of deaths of infants under one-year-old was evaluated using the local and global Moran index. The application of such indexes depends on the definition of a neighbourhood matrix ($W$). This matrix represents the spatial dependence structure of any random variable in lattice data. It gives an idea about the degree of proximity between areas or regions (Werneck, 2008). There are many ways to define $W$. Given $n$ set of areas ($A_1, \dots A_n$), $W$ ($n \times n$) has elements $w_{ij}$ representing any measure of proximity between areas $A_i$ and $A_j$. One criteria used to define $W$ is $w_{ij} = 1$, if area $A_i$ and $A_j$ share the same border between them for $i \neq j$ and $w_{ij} = 0$, otherwise.

According to Waller & Gotway (2004), the Global Moran Index is the main measure used to depict spatial autocorrelation in spatial lattice data. The Global Moran Index is defined by:

$$\hat{I} = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(Y_i - \overline{Y})(Y_j - \overline{Y})}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \qquad (1)$$

where $n$: number of census sectors in Alfenas city; $Y_j$ and $Y_i$ : number of deaths of infants under one year old in census sector $j$ and $i$, respectively; $\overline{Y}$: average of infant mortality in the city of Alfenas ($\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$) and, $w_{ij}$: elements of neighbourhood matrix ($w_{ij} = 1$ if two areas share the same border and $w_{ij} = 0$, otherwise).

The global Moran index has the same interpretation likewise Pearson´s correlation coefficient. However, the index may not be bounded in [-1; 1]. Positive value of Moran index is an indicative of similarity between values of infant mortality in neighbour regions exhibiting clusters in the study area while negative values indicate dissimilarity in neighbour regions, exhibiting a regular pattern in the study area. In the absence of spatial autocorrelation, the Moran index is zero which indicates a complete spatial randomness phenomena (Santos & Souza, 2007). Inference for global Moran index was assessed under the assumption of asymptotic normal distribution (Cliff & Ord, 1981) and a Wald test was carried out defined by:

$$z = \frac{\hat{I} - E[\hat{I}]}{\sqrt{Var[\hat{I}]}}, \qquad (2)$$

where $\hat{I}$ is defined by equation (1), $E[\hat{I}]$ is the expected value of Moran Index and $Var[\hat{I}]$ is the variance (see Cliff and Ord, 1981).

The global Moran index only gives any idea about the presence of spatial autocorrelation. However, it does not identify what areas exhibits similarity or dissimilarity. For instance, what areas can form clusters in the region. For these reasons a Local Spatial Autocorrelation Index (LISA) must be included in the analysis. According to Druck (2004) Moran LISA is defined as

$$\hat{I}_i = \frac{Y_i \sum_{j=1}^n w_{ij} Y_j}{\sum_{j=1}^n Y_j^2}, \qquad (3)$$

where $\hat{I}_i$: local spatial autocorrelation index in area $i$; $n$: number of census sectors in Alfenas city; $Y_i$: number of deaths of infants under one year old in census sector $i$; $Y_j$: number of deaths of infants under one year old in census sector $j$; $w_{ij}$: elements of neighbourhood matrix ($w_{ij} = 1$ if two areas share the same border $w_{ij} = 0$, otherwise).

*Modelling*

According to Draper & Smith (1998) a linear multiple regression model is described as:

$$Y = X\beta + \varepsilon, \qquad (4)$$

where: $Y$ is a vector of observations ($n\times1$) of the dependent variable in $n$ census sectors; $X$ is a matrix ($n \times p+1$) with $p$ independent variables (covariates), measured in $n$ census sectors; $\beta$ is a vector of parameters ($p+1 \times 1$) to be estimated; $\varepsilon$ is a vector ($n \times 1$) of random uncorrelated errors normally distributed with mean zero and a constant variance, $\varepsilon \sim N(\mathbf{0}, I\sigma^2)$.

A stepwise procedure was applied in order to select covariates with significant effect in the model (Charnet et al., 1999). The assumptions of normal distribution, independence and constant variance of the error in classical regression model were analysed by Shapiro-Wilk, Durbin Watson and Breusch Pagan tests, respectively (Draper & Smith, 1998). In addition, a Lagrange Multiplier test was carried out with the residuals of the regression model in order to assess the presence of spatial autocorrelation. The presence of spatial autocorrelation in the residuals is an indicator of lack of independence. Therefore, a spatial component is necessary to be included in the classical regression model in order to reach a high goodness of fit (Anselin, 2005).

Spatial dependence in linear regression models can be incorporated in a single parameter by two ways (Druck, 2004). First, the spatial autocorrelation can be associated to the dependent variable generating a spatial autoregressive model (SAR). Secondly, the spatial dependence parameter is incorporated in the error term, generating a conditional autoregressive model (CAR) also called error model.

According to Anselin (2005), a spatial autoregressive model (SAR) is described as:

$$Y = X\beta + \rho WY + \varepsilon, \qquad (5)$$

where $Y, X, \beta$ e $\varepsilon$ are defined like the classical regression model depicted by equation (4), $\rho$ is the spatial autoregressive coefficient and $W$ is the neighbourhood matrix.

A conditional autoregressive model (CAR) is defined by:

$$Y = X\beta + \varepsilon, \text{ where } \varepsilon = \rho W\varepsilon + u, \qquad (6)$$

and $Y, X, \beta, \rho, W$ e $\varepsilon$ are defined like equation (5) and $u$ is a vector ($n \times 1$) of random uncorrelated errors normally distributed with a vector of mean zero and a constant diagonal covariance matrix, $u \sim N(\mathbf{0}, I\sigma^2)$ (Anselin, 1999).

The null hypothesis for lack of spatial autocorrelation in SAR and CAR models, depicted by equation (5) and (6) was assumed as $\rho = 0$. In other words, if $\rho$ is statistically equal to zero, residuals of the classical regression model are not spatially auto correlated. In CAR and SAR models, the estimates of $\beta$ e $\rho$ were obtained by maximum likelihood method (Bivand, 2010). The adequacy of SAR and CAR models (normality, independence and constant variance of the error) were also evaluated using the same approach applied for the classical regression model. The fitted models were compared by adjusted R², log likelihood and AIC (Akaike information criteria).

A transformation is necessary when the residuals exhibit non constant variance or non-normality. This might be the case when there is count data in the response variable such as infant mortality. In such situation, it is often recommended to take the square root transformation, since it may stabilize the variance and somewhat improves symmetry (Draper & Smith, 1998). The square root transformation is straightforward for the classical regression model but it would be more complicated for the spatial regression models. We advocate using a square root transformation for spatial regression models using the same approach depicted in Miller (1984) for the classical regression model.

For the SAR model, considering a square root transformation in the dependent variable, we have

$$Y^{0.5} = (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\varepsilon.$$

Squaring both sides of the equality, we have

$$(Y^{0.5})^2 = H^2 \text{, where } H = (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\varepsilon.$$

Applying the expectation in both sites, we have
$$E[Y] = E[H^2].$$

Because the second side of the equation is composite by square elements of $H$, this is equivalent to

$$E[Y] = diag(E[HH']),$$

where $diag[.]$ are the diagonal elements of matrix. Thus, we get

$$E[Y] = diag\,(E[(AX\beta + A\varepsilon)(AX\beta + A\varepsilon)']) \text{where } A = (I - \rho W)^{-1}.$$

Applying the transpose, we have

$$E[Y] = diag\,(E[(AX\beta + A\varepsilon)((X\beta)'A' + \varepsilon'A')])$$

which is equivalent to
$$E[Y] = diag\,(E[AX\beta(X\beta)'A' + AX\beta\varepsilon'A' + A\varepsilon(X\beta)'A' + A\varepsilon\varepsilon'A']).$$

It´s known that the expectation of the error term is zero. So, we get

$$E[Y] = diag((AX\beta(X\beta)'A' + AE[\varepsilon\varepsilon']A')),$$

which is equivalent to

$$E[Y] = diag\,(AX\beta(AX\beta)' + AI\sigma^2A').$$

Therefore, the predicted values in the original scale of the dependent variable will be determined by

$$\hat{Y} = diag[(I - \hat{\rho}W)^{-1}X\hat{\beta}((I - \hat{\rho}W)^{-1}X\hat{\beta})'$$
$$+ \hat{\sigma}^2(I - \hat{\rho}W)^{-1}(I - \hat{\rho}W')^{-1}] , \quad (7)$$

where $\hat{\rho}$, $\hat{\beta}$ e $\hat{\sigma}$ are estimates of the unknown parameters obtained by maximum likelihood method using transformed data.

For CAR model it was used the same approach depicted for SAR model. Thus, the predicted values in the original scale of the dependent variable in CAR model will be determined by

$$\hat{Y} = diag[X\hat{\beta}(X\hat{\beta})' + \hat{\sigma}^2(I - \hat{\rho}W)^{-1}(I - \hat{\rho}W')^{-1}]. \quad (8)$$

The predicted values of infant mortality in each census sector were determined by equation (7) and (8). The predictions were used to build thematic maps of infant mortality distribution in Alfenas city in order to identify areas of high risk.

All data analysis were carried out using software R (R-Core-Team, 2016), library "spdep" (Bivand et al., 2011) and Geoda (GEODA, 2010).

## 3. Results and discussion
*Exploratory analysis*

The number of deaths of infants under one-year-old in the city of Alfenas, MG varies from zero to seven and its spatial distribution in quartiles is presented in Figure 2. It shows that

the data have not outliers and infant mortality does not exhibit any trend as well. Thus, the number of deaths of infants under one-year-old presents a first order stationary.
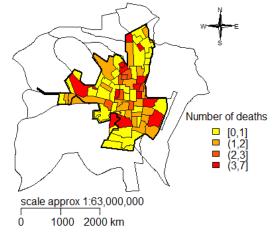


**Figure 2.** Spatial distribution of the number of deaths of infants under one-year-old in quartiles.

The global Moran index (GMI) presented in Table 1 indicates an absence of spatial autocorrelation for the response variable (p value > 0.05).

**Table 1**. Estimates of the global Moran index and its significance

| Index | Estimate | Expetation | Stand. Error | z | p-value |
|---|---|---|---|---|---|
| Moran | -0.081 | -0.015 | 0.051 | -1.30 | 0.903 |

**Table 2.** Results for fitted classical regression model and spatial models (SAR and CAR)

| Covariates | Classical regression model | | SAR model | | CAR model | |
|---|---|---|---|---|---|---|
| | Estimate | Stand. error | Estimate | Stand. error | Estimate | Stand. error |
| Intercept | 0.061 | 0.309 | 1.014** | 0.408 | 0.170 | 0.279 |
| Monthly women income | $-8.7 \times 10^{-6}$** | $2.9 \times 10^{-6}$ | $9.2 \times 10^{-6}$** | $2.6 \times 10^{-6}$ | $7.3 \times 10^{-6}$** | $2.3 \times 10^{-6}$ |
| Number of women in fertile age | $5.4 \times 10^{-3}$** | $1.06 \times 10^{-3}$ | $4.9 \times 10^{-3}$** | $9.6 \times 10^{-4}$ | $4.8 \times 10^{-3}$** | $9.7 \times 10^{-4}$ |
| Rho | 0.00 | - | -0.694** | 0.225 | -0.713** | 0.260 |
| Adjusted R$^2$ | 0.3714 | | 0.4427 | | 0.4280 | |
| Log likelihood | -62.42 | | -58.35 | | -59.32 | |
| AIC | 132.84 | | 126.70 | | 128.64 | |

**Significant at 5% level (p - value < 0.05); AIC – Akaike information criteria

We observe that the GMI is a global index and can mask presence of local spatial autocorrelation. Thus, a spatial local autocorrelation was also carried out using LISA. This index estimates a measure of spatial auto correlation for each area. Figure 3 represents the LISA map where we can see some areas exhibiting a statistically significant local spatial autocorrelation (*p - value* < 0.05). This suggests that the number of deaths of infants under one-year-old in the city of Alfenas is locally spatial auto correlated.

The spatial dependence analysis using local spatial autocorrelation index (LISA) has shown that the number of deaths of infants under one-year-old in neighbour areas are spatially auto correlated. However, there is no trend for such areas to be aggregated in clusters. This can be shown by LISA cluster map depicted in Figure 3. Areas with *High-High* and *Low-Low* designations are those with positive spatial autocorrelation while areas with *Low-High* and *High-Low*

designations belong to regions where the spatial autocorrelation is negative. In Figure 3 there are four areas marked as High-Low. It means that these regions are negatively spatial auto correlated with their neighbours, that is, observations of infant mortality in those four areas are dissimilar between their neighbours.
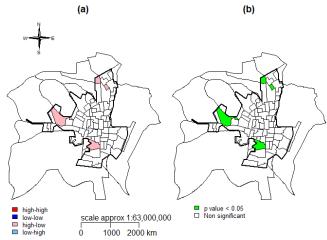
**Figure 3**. Local Index of Spatial Autocorrelation; (a) LISA cluster map; (b) LISA significance map

*Modeling*

Results of classical regression model between transformed data of deaths of infants under one-year-old and the covariates showed that the number of women in fertile age and the monthly income of women are the main variables with statistically significant effects in the model. Table 2 reports the estimates of the parameters of the classical regression model and for spatial models using these two variables.

The Shapiro-Wilk normality test and Breusch-Pagan test showed that the assumptions of normality and constant variance of the error were not violated. However, by the Durbin-Watson test the assumption of non-auto correlated errors was violated. Thus, inferences based on classical regression model are not valid. The lack of independence in the residuals may show evidences of spatial dependence between observations of infant mortality. Therefore, the spatial autocorrelation must be included in the model using SAR or CAR models in order to reach a better goodness of fit.

Likewise, in classical regression model, results of spatial models (SAR and CAR) also reported the number of women in fertile age and the monthly income of women with significant effect in the models (Table 2).

It is noticed that local spatial autocorrelation index between observations (Figure 3) has a great influence while using classical regression model. If spatial auto correlation between observations is ignored, all inferences made by classical regression model may be questionable due the violation of independence assumption. For these reasons, classical regression model presented in Table 2 may not be valid to make accurate inferences. Therefore, a spatial regression model (SAR or CAR) is preferred for inferences. Results for SAR and CAR models showed that transformed data of the number of deaths of infants under one-year-old had an inverse relationship with the monthly income of women, while the number of women in fertile age had a direct relationship with the number of deaths of infants under one-year-old. Therefore, it is expected that areas with more women in fertile age and areas with low monthly income of women will trend to exhibit high values of the number of deaths of infants under one-year-old. In public health programs such areas may deserve special attention for decision makers.

Some authors, as Andrade et al. (2004) and Abreu & Vasconcellos (2010) claimed that it is common to associate socio economic factors with infant mortality. The authors point out that women income is one of the main factors in infant mortality analysis due to the role played by income in purchasing goods and services necessary to achieve high quality of life and health in any country. Menezes et al. (1998), studying risk factors in infant mortality in Pelotas, RS Brazil, reported that infant mortality in families with income less than

the minimum wage was three times greater than families with high income.

The estimates of the covariates parameters are similar in both spatial models and the spatial autoregressive parameter is negative in both models as well. Negative spatial dependence and statistically different from zero, means that values of the number of deaths of infants under one-year-old in neighbour areas are dissimilar among them exhibiting any regular pattern in the study area (Waller & Gotway, 2004).

Shapiro-Wilk normality test and Breusch-Pagan heterokedasticity test for the residuals of SAR and CAR models indicated no violation for normality and constant variance assumptions, respectively. By the three criteria used to choose the best model (adjusted $R^2$, AIC and log likelihood) it is clear that SAR model was the best among the three fitted models (Table 2). The fitted SAR model has the lowest AIC (126.7) and the highest adjusted $R^2$ (44.27%) and log likelihood (-58.35). Therefore, SAR model was used to predict infant mortality in the 68 census sector in Alfenas city. Based on the predicted values a thematic map of the number of deaths of infants under one-year-old distribution was built (Figure 4).

Miller (1984) reports some procedures to predict values of the response variable in classical regression model with transformation in the dependent variable. However, for the spatial regression models such procedures were not found in the literature. Thus, in material and method of this paper it is described a proposal procedure in that sense as depicted by equations (7) and (8). Figure 4 shows the predicted values of the number of deaths of infants under one-year-old in Alfenas city using equation (7) based on SAR model.

The predicted values of the number of deaths of infants under one-year-old in Alfenas city by census sectors varies from zero to four. Based on Figure 4, only few census sector did not present infant mortality, while the majority has in average one and two deaths of infants under one-year-old. We have observed that the census sectors without any occurrence of infant mortality are the noble areas in the city, while census sectors with the highest values of infant mortality (three and four) are those with the worst socio economic conditions in terms of infrastructures and they are the poorest areas in Alfenas city. Thus, in a public health program aiming to reduce infant mortality incidence, these areas shall deserve special attention.
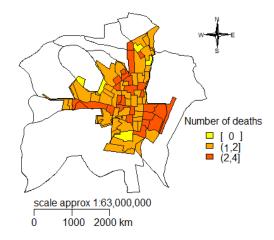


**Figure 4.** Spatial distribution of predicted values of the number of deaths of infants under one-year-old in the city of Alfenas.

## 4. Conclusions

The results obtained in this study showed that infant mortality in Alfenas city, MG has negative local spatial auto correlation. Areas with high values of infant mortality are more likely to be surrounded by areas with low values of infant mortality. The use of a square root transformation in the

response variable allowed developing prediction equations in spatial regression models. The number of women in fertile age and the monthly women income are the main factors affecting infant mortality in Alfenas city, MG.

## 5. Acknowledgements

## 6. References

Abreu, C.D.C.R. & Vasconcellos, E.S. (2010). Análise dos fatores sócio-econômico-ambientais e a sua interferência na mortalidade infantil na favela da Aldeia. Vertices,1, 7-13.

Andrade, C.L.T., Szwarcwald C.L., Gama, S.G.N. & Leal, M.C. (2001). Desigualdades sócio-econômicas do baixo peso ao nascer e da mortalidade perinatal no Município do Rio de Janeiro, Cadernos de Saúde Pública, 20, 44-51.

Anselin, L. (1999). Spatial econometrics. University of Texas: Dallas, USA.

Anselin, L. (2005). Exploring spatial data with GeoDaTM : A Workbook. Urbana-Champaign:
Spatial Analysis Laboratory, Department of Geography. University of Illinois: Illinois, USA.

Bivand, R., et al. (2011). Spdep: Spatial dependence: weighting schemes, statistics and models.
R package version 0.5-31, Disponível em <http://CRAN.R-project.org/package=spdep>. Acesso
em 19/12/2012

Bivand, R.S. (2010). Exploratory spatial data analysis. In: Handbook of applied spatial
analysis: London, Reino Unido.

Charnet, R., Freire C.A.L., Charnet, E.M.R. & Bonvino H. (1999). Análise de modelos de
regressão linear com aplicações: Unicamp.

Cliff, A.D. & Ord, J.K. (1981). Spatial processes: models & applications. Taylor & Francis:
London, Reino Unido.

Cressie, N. (2015). Statistics for spatial data. John Wiley & Sons: New York.

Draper N.R. & Smith H. (1998). Applied regression analysis new york. John Wiley: New York,
USA.

Druck S. (2004). Análise espacial de dados geográficos. Empraba: Brasília, Brasil.

Genowska, A. et al. (2015). Environmental and socio-economic determinants of infant mortality
in Poland: an ecological study. Environmental Health, 14, 1-8.

GEODA (2010). Center for geospatial analysis and computation.

Hall, E.S., Venkatesh, M. & Greenberg, J.M. (2016). A population study of first and subsequent
pregnancy smoking behaviors in Ohio. Journal of Perinatology, 36, 948-953.

IBGE (2010). Censo Demográfico. Características Gerais da População.

Köppen, W. & Geiger, R. (1928). Klimate der Erde. Gotha: Verlag Justus Perthes. Wall-map.

Leal, M.d.C. & Szwarcwald, C.L. (1997). Características da mortalidade neonatal no Estado do
Rio de Janeiro na década de 80: uma visão espaço-temporal. Revista de Saúde Pública, 31, 457
465.

Menezes A. et al. (1998). Fatores de risco para mortalidade perinatal em Pelotas, RS, 1993.
Revista de Saúde Pública, 32, 209-216.

Miller, D.M. (1984). Reducing transformation bias in curve fitting. The American Statistician,
38, 124-136.

R-Core-Team (2016). R: A language and environment for statistical computing. R. Foundation
for Satatical Computing: Viena.

Santos, S.M. & Souza W.V. (2007). Introdução à estatística espacial para a saúde pública:
Ministério da Saúde/Fundação Oswaldo Cruz, Brasília, Brasil.

Silva, S.L.C., Fachel J.M.G., Kato, S.K. & Bassanesi S.L. (2011). Visualização dos padrões de
variação da taxa de mortalidade infantil no Rio Grande do Sul, Brasil: comparação entre as
abordagens Bayesiana Empírica e Totalmente Bayesiana. Cadernos de Saúde Pública, 27, 1423
1432.

Waller, L.A. & Gotway, C.A. (2004) Applied spatial statistics for public health data. John Wiley
& Sons: New Jersey, USA.

Werneck, G.L. (2008) Georeferenced data in epidemiologic research. Ciência & Saúde Coletiva,
13, 1753-1766.