

# **Desempenho Escolar e Sua Relação com Dados Socioeconômicos no Município de São Paulo**

**Lino Augusto Sander De Carvalho**

**RESUMO :** O sistema educacional brasileiro, apesar de grande avanço em sua universalização, ainda passa por sérios problemas dentre os quais a segregação pode ser apontada. Como forma de tentar avaliar a o problema, o Estado tem investido na promoção de programas de avaliação de desempenho escolar nos diversos setores da educação nacional. É de grande interesse saber se este indicador esta relacionado a dados socioeconômicos, como forma de direcionar políticas públicas intervencionistas. Ainda assim, existe uma dificuldade em relacionar tais dados devido à complexidade de relacionar o rendimento do aluno com a localidade de sua residência e em decorrência com a condição social de sua família. Neste trabalho se analisou a relação social de dados socioeconômicos (escolaridade e renda) com o indicador de desempenho escolar. Foram adotadas três formas de agregação onde se investigou a dependência espacial em cada uma delas. Alguns resultados demonstraram uma pequena melhora dos modelos explicativos. Outros resultados apresentam a variação da dependência espacial para as formas de agregação encontradas.

## INTRODUÇÃO

O sistema educacional brasileiro tem conseguido significativo avanço na universalização do ensino público (Torres et al 2006). Entretanto ainda existem problemas inerentes à uma dada escola, classes de escolas e regiões de escolas (Torres et al 2002). Desta forma o estudo do processo de segregação educacional pode ajudar a compreender e até mesmo colaborar para a construção de políticas públicas

Segundo Torres e Gomes 2002 um padrão de segregação pode ser identificado na região metropolitana de São Paulo, onde de acordo com o autor as escolas mais próximas ao centro dos municípios apresentaram um melhor rendimento escolar observado através da nota no do exame Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP) do ano de 2001.

Um outro indicador desta segregação esta apresentado no trabalho de Torres et al 2006, que demonstra a diferença entre a distribuição de notas das escolas nas regiões centrais dos municípios e nas regiões de periferia. Nota-se que existe realmente uma diferença, onde a média das notas nas regiões centrais é significativamente mais alta do que nas regiões periféricas.

Vários são os motivos que podem explicar essa segregação. Torres e Gomes 2002 apresentam três motivos gerais. Um deles leva em conta que segmentos menos educados têm menor capacidade de demandar serviços públicos ao governo. Um segundo argumento leva em conta que a burocracia desenvolvida para lidar com questões públicas passa por uma “classe” de servidores que provem das classes mais favorecidas, que não necessariamente compreendem os códigos, linguagens, atitudes, valores e necessidades da população pobre. Ainda há um terceiro argumento ordem econômica, que relaciona a segregação à dificuldade de investimento do estado em regiões pobres, por motivos de irregularidade jurídica.

Os três motivos apresentados têm ligação direta com a escolaridade das regiões onde se localizam as escolas. Entretanto em mesmo estudo, Haroldo e Gomes demonstraram que uma regressão linear entre notas das respectivas escolas aliadas à escolaridade do setor censitário aonde se encontram apresentou um baixo coeficiente de ajuste.

Segundo Camara et al 2001, dados que envolvem áreas podem ser melhor

analisados quando se leva em consideração o efeito de uma dependência espacial. As diversas técnicas de análise exploratória podem ajudar a entender melhor o fenômeno de dependência entre duas grandezas.

Um problema em se tratar a relação entre dados espaciais é a delimitação a área a ser considerada e a forma de agregação dos dados presentes nesta área (Sabatini, F., Cáceres, G., & Cerdá, J. 2001). No caso de estudos utilizando dados de escolas a dificuldade está na forma de agregação que seja mais significativa de forma que se possa supor que a residência do aluno esteja presente na área que se relaciona com a nota da escola.

Desta forma, este trabalho se propõe a estudar a relação entre escolaridade, renda e a nota alcançada pela escola no Saresp. Para isso foram consideradas três diferentes formas de agregação de dados onde, para cada uma delas foi estudada a dependência espacial. Uma comparação entre um modelo espacial e linear simples é feita de forma a se encontra o melhor ajuste.

## DADOS

Neste trabalho, utilizaram-se dados do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP) do ano de 2001, para 581 escolas pertencentes ao município e São Paulo. O dado considerado foi a nota média dos alunos de quinta série diurno no exame de avaliação (NME).

Conjuntamente aos dados de nota média, foram também utilizados os dados de “*Média do rendimento nominal mensal das pessoas responsáveis por domicílios particulares permanentes (MRR)*” e “*Média do número de anos de estudo das pessoas responsáveis por domicílios particulares permanentes (MER)*”, que foram extraídos dos Dados do Censo Demográfico de 2000, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

## METODO

O uso do resultado de avaliação do SARESP como indicador de Escolaridade e Renda seria idealmente estimado utilizando dados de aproveitamento escolar dos estudantes e características específicas de suas famílias. Entretanto esse tipo de dado nem sempre é fácil de obter, o que leva à utilização de um dado agregado por área.

A agregação de dados por setor censitário é a que pode levar a um resultado que se aproxima significativamente do indivíduo. Segundo Haroldo et.al. (2002), a utilização de notas de escola como indicador de parâmetros socioeconômicos, gera um modelo pouco ajustado no caso da *Nota de Escola* como variável resposta e *Escolaridade Média do Responsável* como variável explicativa.

Entretanto, no caso de se considerar o efeito espacial no modelo explicativo uma agregação por setor censitário fica inviável, pois o número de escolas para o qual se tinha dados não chega a 40% do número de Setores Censitários para o município de São Paulo.

O método utilizado neste trabalho buscou estudar a relação entre as variáveis explicativas, *Escolaridade Média por Setor Censitário* e *Renda Média por Setor Censitário*, admitindo uma preferência pela maior correlação encontrada com a variável *Nota Média Da Escola*. O estudo de correlação foi feito para três tipos de área diferentes, sendo estas as subdivisões de “Distritos”, “Microregiões” e “Áreas de Influência”

Após o estudo de cada variável independentemente, foi medido o “Índice de Moran para a Autocorrelação espacial”, como forma de testar a presença de dependência espacial para o indicador “*Nota Média de Escola*”. Uma vez comprovada a existência de autocorrelação espacial, deu-se seguimento à análise estatística dos dados e forma testar requisitos básicos da regressão espacial como normalidade das variáveis resposta; normalidade, homocedasticidade, e autocorrelação espacial dos resíduos para então se aplicar o modelo mais adequado.

Uma vez aplicado o modelo de regressão espacial mais adequado, uma comparação com o modelo linear não espacial e entre os modelos de regressão espacial foi realizada para as três subdivisões propostas. Uma análise dos resíduos de regressão espacial também foi feita.

#### DEFINIÇÃO DAS SUBDIVISÕES DE AGREGAÇÃO

As regiões de agregação foram definidas de acordo com subdivisões pré-estabelecidas como “Distritos” e “Microáreas” que são subdivisões da prefeitura municipal de São Paulo (INFOLOCAL).

Uma outra subdivisão foi estimada a partir dos baseando-se na idéia de proximidade do aluno ao local de estudo.

#### - Distritos

A subdivisão “Distritos” seguiu a delimitação da Prefeitura Municipal de São Paulo totalizando 96 distritos. Os dados de MER e MRR foram agregados, contabilizando a média para cada um desses distritos.

Os dados de NME também foram agregados, contabilizando a média para cada um desses distritos.

Os distritos da República, Sé, Bom Retiro e Barra Funda foram excluídos por não possuírem dados de NME chegando a um número final de 92 distritos.

#### - Microáreas

A subdivisão Microáreas também seguiu a delimitação determinada pela Prefeitura Municipal de São Paulo (INFOLOCAL). A média dos dados de MER e MRR e NME foram agregadas para cada microrregião.

De um total de 456 microáreas, chegou-se a um total de 354 onde 102 microáreas foram excluídas por não possuírem dados de NME.

#### - Regiões de Influência

As regiões de influência foram criadas visando preservar a o dado de NME, para que este não fosse distorcido pela média. A área de influência tenta estimar um círculo em volta do ponto aonde se encontra a escola, criando uma área donde se acredita estarem localizados os domicílios dos alunos que estudam na escola. O raio estipulado tenta se aproximar da idéia de proximidade da residência do aluno ao local de estudo.

O método de criação das regiões de influência se deu através da interpolação dos dados de NME pelo método de Vizinho mais Próximo. Os dados de NME foram transformados e interpolados para um raio máximo de 2000 metros. Cada região possui o dado de NME respectivo à nota de sua escola geradora.

Após a delimitação, os dados de MER e MRR foram agregados para cada uma das regiões de influência criadas. Duas regiões de influência foram excluídas por não possuíram vizinhos.

Das 581 escolas chegou-se a um total de 473 regiões de influência. O menor número de regiões em relação ao de escolas se deve ao fato de que escolas próximas de mesmo valor de NME foram agrupadas na mesma classe gerando uma região única.

## INDICADORES DE ASSOCIAÇÃO ESPACIAL

### - Matriz de Vizinhaça

A matriz de proximidade, ou matriz de vizinhaça busca estimar os vizinhos de uma determinada área de acordo com uma série de critérios (Câmara et. al. 2001). Neste trabalho, o critério utilizado foi o de contigüidade onde polígonos que tocam a área tem peso "1" e os que não tocam tem peso "0"

### - Índice de Moran

O índice de Moran é uma ferramenta de medida de autocorrelação espacial que faz essa estimativa a partir do produto dos desvios em relação à média. Formalmente, pode ser escrito da seguinte maneira (1) :

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2} \quad (1)$$

onde  $n$  é o número de áreas,  $z_i$  o valor do atributo considerado na área  $i$ ,  $\bar{z}$  é o valor médio do atributo na região de estudo e  $w_{ij}$  os elementos da matriz normalizada de proximidade espacial (Câmara et. al. 2001).

Os valores de 0 a 1 indicam correlação direta enquanto os valores de 0 a -1 indicam correlação inversa. O valor 0 indica ausência de autocorrelação espacial.

O índice de Moran tem como requisitos básicos a estacionariedade de primeira e segunda ordem. Em caso de violação dessa condição, outra abordagem pode ser utilizada. Uma dessas abordagens é a aplicação de indicadores locais de associação espacial (LISA). O índice de moran local, que é um dos LISA, é dado pela seguinte formula (2) :

$$I_i(d) = \frac{(x_i - \bar{x})}{s^2} \sum_j w_{ij}(d)(x_j - \bar{x}) \quad (2)$$

onde  $w_{ij}(d)$  é a matriz de vizinhaça para o par  $i$  e  $j$ ,  $d$  é a medida de distancia estipulada

pele modelo de vizinhaça,  $x_i$  e  $x_j$  são os valores encontrados em  $i$  e suas vizinhaças  $j$ ,  $\bar{x}$  é a media amostral e  $s^2$  é a variância amostral global.

O índice de moran local pode ser representado visualmente através de um mapa onde a idéia é comparar a distribuição espacial com a média local de sua vizinhaça. O mapa é construído a partir dos valores dos quadrantes de moran que são estipulados pelo diagrama de espalhamento de moran como representado na figura (1):

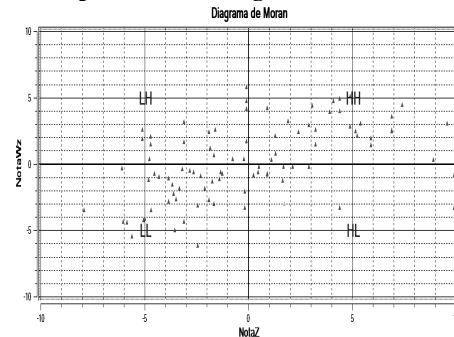


Figura 1 – Diagrama de Espalhamento Moran

Os quadrantes "1" e "2" indicam associação espacial alto-alto (alto valor do índice de moran para o polígono e sua vizinhaça) e baixo-baixo (baixo valor do índice de moran para o polígono e sua vizinhaça) respectivamente, enquanto que os quadrantes 3 e 4 indicam associação espacial alto-baixo (alto valor do índice de moran para o polígono e baixo para sua vizinhaça) e baixo-alta (baixo valor do índice de moran para o polígono e sua vizinhaça) também respectivamente.

### Modelos de Regressão

A regressão linear simples é uma técnica estatística que busca estimar uma função explicativa a partir de dados coletados. O melhor modelo é alcançado pela técnica dos mínimos quadrados, onde a função linear explicativa paramétrica é minimizada. A função, dita, variável resposta pode ser representada da seguinte forma (3):

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (3)$$

Onde  $X$  representa a variável explicativa,  $Y$  a variável resposta,  $\beta_0$  e  $\beta_1$  são os coeficientes de ajuste. O parâmetro  $\varepsilon$  representa o resíduo da regressão.

O modelo de regressão linear tem como condição básica a necessidade de que as observações sejam independentes, ou seja,

que as variáveis sejam variáveis aleatórias normalmente distribuídas. Outra premissa básica é a de que os resíduos devem também, ser independentes e que sigam uma distribuição normal padrão com média zero e variância constante.

Entretanto, efeitos espaciais não são levados em conta na regressão linear simples. Desta forma existem outros modelos de regressão que tentam incorporar a dependência espacial no modelo explicativo. Estes modelos podem tratar a regressão de forma local ou global. Os modelos globais utilizam um único parâmetro para incorporar a estrutura de correlação espacial, enquanto que os modelos locais utilizam uma regressão espacial para regiões determinadas buscando parâmetros de ajuste locais.

Existem dois tipos de modelos de regressão espacial global. Um deles é *Modelo Espacial Auto Regressivo Misto* (“Spatial Autoregressive” - SAR ou “Spatial Lag Models”). Este modelo incorpora toda a autocorrelação espacial na variável resposta. A equação (4) demonstra o calculo do modelo :

$$Y = \rho WY + X\beta + \varepsilon \quad (4)$$

Onde W é a matriz de proximidade espacial, o produto WY expressa a dependência espacial em Y e  $\rho$  é o coeficiente espacial autoregressivo (Camara et al 2001). Outro modelo que modela os efeitos globais é o *Modelo do Erro Espacial* (“Conditional autoregressive”- CAR ou “Spatial Error Model”), que incorpora a autocorrelação espacial no resíduo da regressão, como é demonstrado pela formula abaixo :

$$Y = X\beta + \varepsilon,$$

$$\varepsilon = \lambda W + \xi.$$

Onde  $\varepsilon$  é a componente do erro com efeitos espaciais,  $\lambda$  é o coeficiente autoregressivo e  $\xi$  é a componente do erro com variância constante e não correlacionada.

Neste trabalho apenas os modelos que modelam os efeitos espaciais globais foram utilizados. Apesar de alguns resultados apresentarem constatação do uso de modelos espaciais locais, esta tarefa foi delegada à trabalhos futuros.

## RESULTADOS

O estudo dos dados de NME como indicadores de MER e MRR foram realizados para as subdivisões já descritas na metodologia deste trabalho. O mapa abaixo

demonstra como se dá a distribuição das 581 escolas consideradas na região municipal de São Paulo (figura 2) :

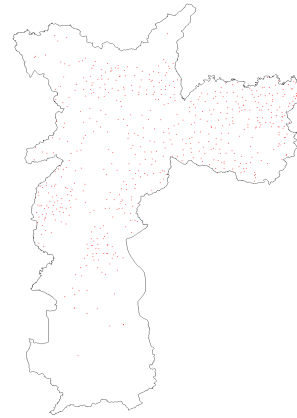


Figura 3 – Distribuição das Escolas em São Paulo

- Distritos

A agregação dos dados por distrito considerou a média dos dados de MER, MRR e NME para cada distrito em questão. Os mapas da distribuição dos dados agregados estão representados nas figuras (3),(4) e (5) :

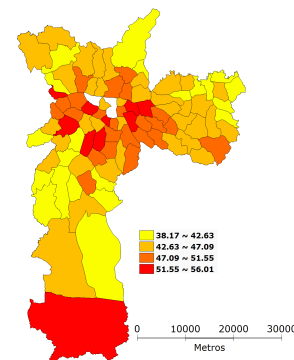


Figura 3 – Dados de NME Agregados

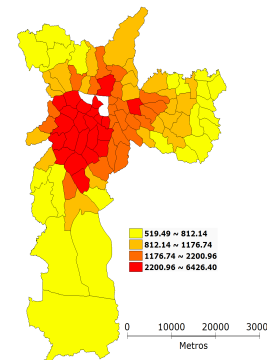


Figura 4 – Dados de MRR Agregados

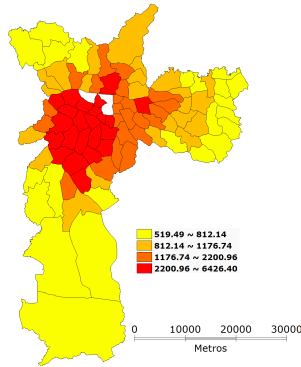


Figura 5 – Dados de MRE Agregados

- Dependência Espacial

Em seguida aos mapas comparativos, foi gerado o índice de moran para os dados de NME. O mapa de moran dos coeficientes de moran esta representado na figura (6):

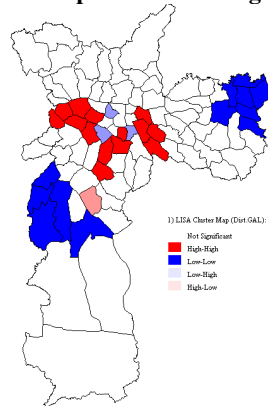


Figura 6 – Mapa de LISA/Moran de NME

O índice de moran global calculado apresentou o valor de 0.3643, o que indica dependência espacial. A variável NME apresentou normalidade segundo o teste de Kolmogorov-Smirnov.

- Análise de Variáveis

Inicialmente as análises estatísticas buscaram determinar a correlação entre as variáveis explicativas. Os dados de NME e MRE apresentaram maior correlação como demonstrado na tabela (1).

Tabela 1- Estudo de Correlação Distritos

	Correlação		
	Nota	Renda	Escol
Nota	1.000000	0.4566167	0.5950300
Renda	0.4566167	1.000000	0.9116292
Escol	0.5950300	0.9116292	1.000000

Um estudo realizado foi o de tentativa de criar um modelo de regressão multivariada onde as variáveis MER e MRR entrariam em conjunto com variáveis

explicativas. Entretanto a alta correlação entre estas variáveis impossibilitou tal feito pois a variável MER “retira” a significância da variável MRR.

Optou-se então pelo descarte da variável MRR e o estudo de regressão linear e espacial foi feito apenas com a variável MER.

- Regressão Linear

O modelo linear para os dados NME e MRE esta representado pela equação 5

$$NME = 36.7 + 1,12MRE \dots(5)$$

A análise estatística do modelo ajustado, esta representada na tabela (2) :

Tabela 2- Regressão Linear Distritos

Regressão Linear				
R-squared :	0.354061	F-statistic :	49.332	
Adjusted R-squared :	0.346884	Prob (F-statistic) :	3.97621e-010	
Sum squared residual:	1026.22	Log likelihood :	-241.488	
Sigma-square :	11.4025	Akaike :	486.975	
Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	36.99596	1.346205	27.48167	0.0000000
ESCOL	1.126762	0.1604234	7.023673	0.0000000
MULTICOLLINEARITY 7.51471				
TEST ON NORMALITY OF ERRORS				
TEST	DF	VALUE	PROB	
Jarque-Bera	2	31.88718	0.0000001	
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.09642926	0.7561575	
Koenker-Bassett test	1	0.04298158	0.8357600	

Através da tabela (2) é possível verificar que os testes F e t são significativos o que considera significativamente a variável escolaridade. Observa-se também que os resíduos não possuem distribuição normal e sua variância é constante.

- Análise dos resíduos de Regressão

Em seguida foi realizada a análise de dependência espacial dos resíduos (tabela 3). Observa-se que o índice de moran é o único que apresenta autocorrelação espacial dos resíduos significativa.

Tabela (3) - Dependência Espacial dos Resíduos - Distritos

Dependencia Espacial dos Resíduos			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.136074	2.2634507	0.0236078
Lagrange Multiplier (lag)	1	2.7241574	0.0988403
Robust LM (lag)	1	0.0922623	0.7613205
Lagrange Multiplier (error)	1	3.6446577	0.0562494
Robust LM (error)	1	1.0127626	0.3142419

Apesar da não significância dos testes de dependência espacial por multiplicadores de Lagrange, deu-se continuidade ao trabalho gerando os modelos Spatial LAG e Erro Model.

- Análise Espacial

Os resultados da regressão utilizando o modelo LAG esta representado na tabela (4) Observa-se que o resultado não é satisfatório pois o testes demonstram a não significância do parâmetro de ajuste do modelo.

Outra constatação é pouca variação do valor de AKAIKE, o que indica que o modelo espacial não ajuda a explicar a relação entre as variáveis NME eMER.

Tabela (4) – Modelo LAG - Distritos

Lag Model				
R-squared : 0.376749				
Log likelihood : -240.28				
Akaike : 486.56				
Sigma-square : 10.7628				
S.E of regression :3.28067				
Variable	Coefficient	Std.Error	z-value	Probability
W_NOTA	0.2018703	0.1305561	1.546234	0.1220481
CONSTANT	29.11077	5.274018	5.519658	0.0000000
ESCOL	0.9455065	0.1934395	4.887868	0.0000010
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.3709349	0.5424951	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	2.415109	0.1201696	

O modelo Error também não se mostrou satisfatório como pode ser observado na Tabela (5).

Tabela 5 – Modelo Error Distritos

Error Model				
R-squared : 0.385340				
Log likelihood : -239.879375				
Sigma-square : 10.614450				
Akaike :483.759				
S.E of regression : 3.25798				
Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	37.02419	1.591728	23.26038	0.0000000
ESCOL	1.121094	0.1895105	5.915737	0.0000000
LAMBDA	0.2489913	0.1363967	1.825494	0.0679265
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.3879225	0.5333935	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	3.216299	0.0729083	

Apesar do valor de AKAIKE ser levemente mais baixo do que ao da regressão linear, o teste t demonstra que seu valor não é significativo.

- Teste de Dependência Espacial dos Resíduos

Observa-se que os valores dos coeficientes de ajuste, tanto do modelo LAG como do modelo Error, não apresentaram significância. O mapa LISA que testa o agrupamento espacial foi construído

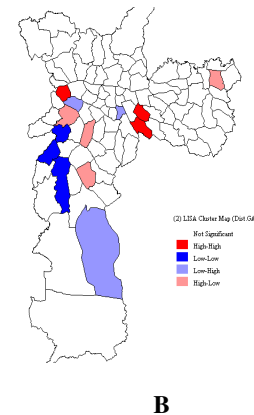
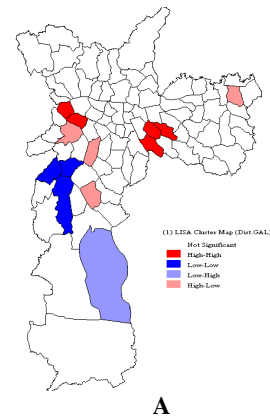


Figura 7 – Mapa LISA dos Resíduos de Regressão Espacial A LAG Model – B Error Model

O valor do índice de moran para os distritos foi de 0.01 e 0.02 para os resíduos do modelo LAG e Error respectivamente. Desta Forma observa-se uma baixa auto-correlação espacial dos resíduos.

Microáreas

A agregação por microáreas seguiu a mesma lógica e agregação dos distrito. Para cada microrregião foi contabilizada a média

dos dados de NME MRR e MER. As figuras (8),(9) e (10) demonstram a distribuição dos dados no município e São Paulo:

**-Dependencia Espacial**

O índice de moran, chegou ao valor de 0,3643 o que indica a presença de autocorrelação espacial. O mapa apresentado na figura (11) demonstra a divisão das microáreas nos quadrantes determinados.

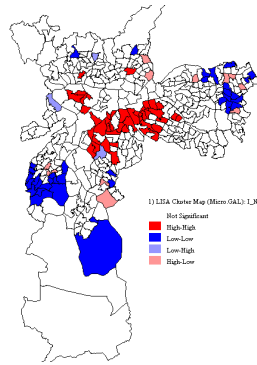


Figura (11) – Índice de Moran de NME

**- Análise de Variáveis**

O estudo das variáveis demonstrou a maior correlação entre a variável NME e MRE como demonstra a tabela (6). A variável NME apresentou normalidade segundo o teste de Kolmogorov-Smirnov.

Tabela (6) – Estudo de Correlação Microáreas

	Correlação		
	Nota	Escol	Renda
Nota	1.0000000	0.53517	0.4601899
Escol	0.5351700	1.00000	0.8738000
Renda	0.4601899	0.87380	1.0000000

Como na divisão de distritos, um estudo realizado foi o de tentativa de criar um modelo de regressão multivariada onde as variáveis MER e MRR entrariam em conjunto com variáveis explicativas. Entretanto a alta correlação entre estas variáveis impossibilitou tal feito, pois a variável MER “retira” a significância da variável MRR.

Desta forma optou-se pelo descarte da variável MRR e o estudo de regressão linear e espacial foi feito apenas com a variável MER.

**- Regressão Linear**

O resultado da regressão linear pode ser visualizado na tabela (7). Observa-se a significância do coeficiente angular encontrado.

No estudo dos resíduos de regressão nota-se a que a normalidade dos resíduos não foi atendida enquanto que a homocedasticidade da variância dos resíduos foi atendida.

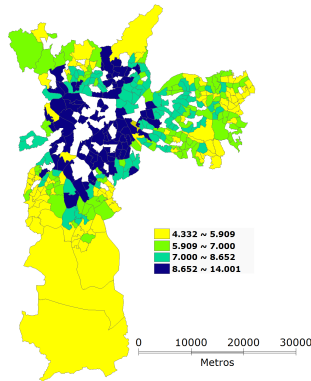


Figura (8)-Agrupamento de dados MER

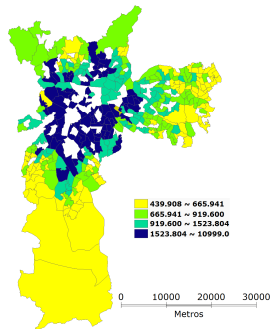


Figura (9)- Agrupamentos de Dados MRR

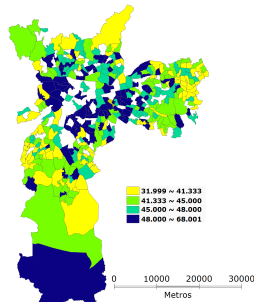


Figura (10) – Agrupamento de Dados NME



O modelo esta representado pela equação abaixo:

$$NME = 34,79 + 1,35MRE \quad (4)$$

Tabela (7) – Regressão Linear Microáreas

Regressão Linear				
R-squared	: 0.286407	F-statistic	: 141.278	
Adjusted R-squared	: 0.284380	Prob(F-statistic)	: 1.27108e-027	
Sum squared residual	: 6910.4	Log likelihood	: -1028.26	
Sigma-square	: 19.6318	Akaike info criterion	: 2060.51	
Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	34.79137	0.8855011	39.29004	0.0000000
ESCOL	1.353172	0.1138453	11.88606	0.0000000
MULTICOLLINEARITY CONDITION NUMBER 7.384981				
TEST ON NORMALITY OF ERRORS				
TEST	DF	VALUE	PROB	
Jarque-Bera	2	54.85768	0.0000000	
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.1904178	0.6625691	
Koenker-Bassett test	1	0.1074377	0.7430804	

#### - Análise dos Resíduos de Regressão

A análise dos resíduos de regressão dada pelos testes de Moran, Multiplicadores de Lagrange está apresentado na tabela 8.

Tabela 8 – Dependência Espacial dos Resíduos

Dependência Espacial Dos Resíduos			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.089368	2.5039505	0.0122816
Lagrange Multiplier (lag)	1	8.5013042	0.0035489
Robust LM (lag)	1	3.9601955	0.0465883
Lagrange Multiplier (error)	1	5.5311842	0.0186805
Robust LM (error)	1	0.9900754	0.3197239

Observa-se que o índice de Moran, apesar de baixo, é significativo, o que indica correlação espacial dos resíduos.

O teste de multiplicadores de Lagrange indica que o melhor modelo a ser aplicado seria o modelo LAG, entretanto os dois modelos foram aplicados buscando uma diminuição do AKAIKE.

#### - Regressão Espacial

Inicialmente o modelo LAG foi aplicado. A tabela (9) demonstra o resultado desta regressão.

Observa-se uma diminuição do AKAIKE em relação à regressão linear simples. O teste para homocedasticidade se apresenta válido. Observa-se também que o coeficiente de regressão é significativo.

Tabela 9 – Modelo LAG Microáreas

Lag Model				
R-squared	: 0.308195	Log likelihood	: -1024.28	
		Akaike	: 2054.55	
Variable	Coefficient	Std.Error	z-value	Probability
W_NOTA	0.1854701	0.06529091	2.840672	0.0045020
CONSTANT	27.9905	2.577262	10.86055	0.0000000
ESCOL	1.148171	0.1305841	8.792574	0.0000000
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.1981256	0.6562381	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	7.961562	0.0047781	

O modelo Error também foi empregado e sua análise esta apresentada na tabela 10.

Tabela 10 –Modelo Error Microáreas

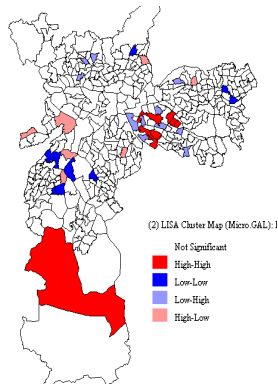
Error Model				
R-squared	: 0.301108	Log likelihood	: -1025.689748	
		Sigma-square	: 19.118746	
		Akaike	: 2055.38	
Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	35.0901	0.9913563	35.39605	0.0000000
ESCOL	1.312328	0.1269091	10.34069	0.0000000
LAMBDA	0.1601715	0.07031835	2.277805	0.0227381
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	0.2760295	0.5993155	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	5.135037	0.0234476	

O modelo Error apresenta também um valor de AKAIKE mais baixo em relação à regressão linear. Da mesma forma que o modelo LAG os testes de ajuste do modelo e de homocedasticidade se apresentaram válidos.

#### - Teste de Dependência Espacial dos Resíduos

Para o teste de dependência dos resíduos foi calculado se através do índice de moran juntamente com o índice foi construído um mapa Lisa dos resíduos da regressão espacial.

O índice de moran dos resíduos teve o valor de 0.0457 o que mostra a ausência de auto correlação espacial.



A

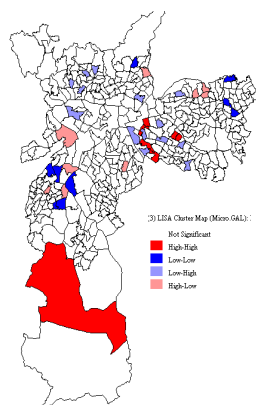


Figura 12 – Mapa Lisa dos Resíduos de Regressão Espacial

O mapa LISA mostra um agrupamento de microáreas. Esse agrupamento é pouco significativo, o que pode indicar ausência de dependência espacial dos resíduos.

### - Áreas de Influência

#### -Forma de Criação

O objetivo da criação das áreas de influência é a preservação dos dados de NME. Desta forma se procurou partir deste dado para a construção das mesmas.

A partir dos pontos, ou escolas, foi interpolada uma grade pelo interpolador de vizinho mais próximo. O raio utilizado tentou aproximar a área de influência ao critério de zoneamento para matrícula.

Desta forma se criou um arquivo de polígonos onde cada área preserva a nota da escola em questão. As áreas de influência criadas estão representadas na figura (13).

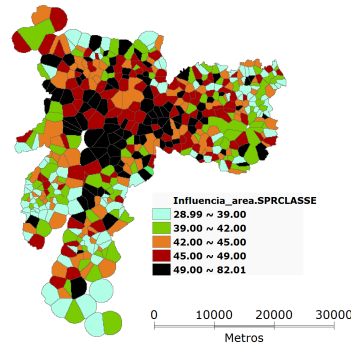


Figura (13) – Áreas de Influência

### - Agregação de Dados

Os dados de MRE e MRR foram agregados para cada região de influência. As figuras (14),(15) e (16) apresentam a distribuição dos dados agregados.

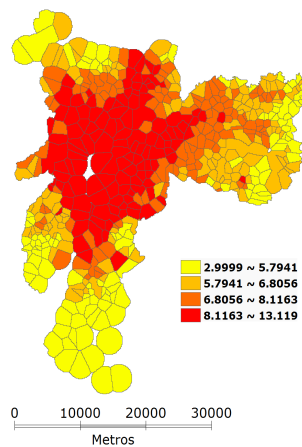


Figura 14 – Dados agregados de MER

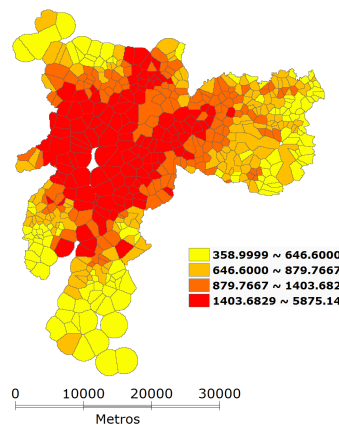


Figura 15 – Agregação de Dados MRR

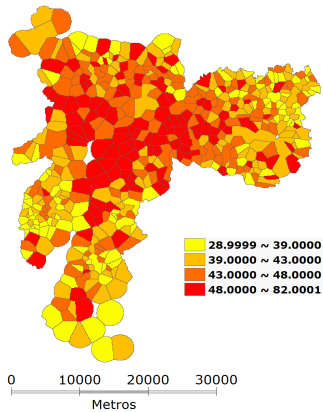


Figura 16 - Distribuição de Notas - Dependência Espacial

O índice de moran para a dependência espacial da variável Nota das Escolas foi gerado. O valor do índice de moran foi de 0.2083, o que indica uma baixa autocorrelação espacial.

O mapa de moran construído está demonstrado na figura 17.

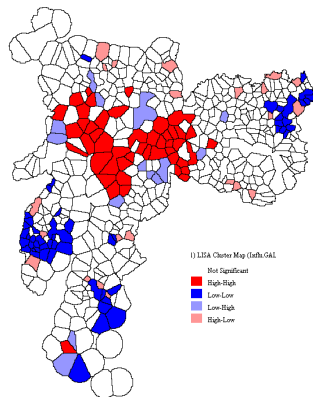


Figura 17 – Mapa de Moran Das Notas de Escola

- Análise de Variáveis

A correlação entre as variáveis foi estudada de forma a determinar qual delas possui a maior relação com a variável nota de escola.

A variável Nota apresentou normalidade segundo o teste de Kolmogorov-Smirnov.

Tabela 10 – Análise de Correlação

	Correlação		
	Nota	Escol	Renda
Nota	1.0000000	0.4657962	0.3983043
Escol	0.4657962	1.0000000	0.9212198
Renda	0.3983043	0.9212198	1.0000000

Como as variáveis são muito correlacionadas, novamente não foi possível envolver as duas variáveis num modelo multivariado

Porém como o neste modelo se preserva o valor da variável nota, as variáveis MRR e MER foram estudadas separadamente onde foi construído um modelo linear para cada uma das variáveis independentemente.

-Regressão Linear Parâmetro MER

Os dados de regressão linear estão apresentados na tabela 11. A equação demonstra o modelo ajustado :

$$NME = 34,79 + 1,35MRE \quad (6)$$

Tabela 11 – Regressão Linear MER

Regressão Linear					
R-squared	:	0.216966	F-statistic	:	130.507
Adjusted R-squared	:	0.215304	Prob(F-statistic)	:	7.56948e-027
Sum squared residual	:	14616.5	Log likelihood	:	-1482.54
Sigma-square	:	31.0329	Akaike	:	2969.09
Variable	Coefficient	Std.Error	t-Statistic	Probability	
CONSTANT	32.73422	0.9947155	32.90812	0.0000000	
ESCOL	1.522633	0.1332843	11.42395	0.0000000	
MULTICOLLINEARITY CONDITION NUMBER 7.635955					
TEST ON NORMALITY OF ERRORS					
TEST	DF	VALUE	PROB		
Jarque-Bera	2	1086.96	0.0000000		
DIAGNOSTICS FOR HETEROSKEDASTICITY					
TEST	DF	VALUE	PROB		
Breusch-Pagan test	1	2.812982	0.0935044		
Koenker-Bassett test	1	0.634627	0.4256632		

Observa-se que o valor dos coeficientes de regressão são significativos, porem os valores dos testes de normalidade e homocedasticidade indicam ausência e normalidade e presença de homocedasticidade.

- Análise dos Resíduos

A tabela 12 indica o estudo de autocorrelação espacial dos resíduos de regressão linear para o parâmetro MER.

Tabela 12 – Dependência Espacial dos Resíduos – Áreas de Influencia

Dependência Espacial			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.055221	2.0830120	0.0372500
Lagrange Multiplier (lag)	1	2.8632912	0.0906219
Robust LM (lag)	1	0.5748933	0.4483210
Lagrange Multiplier (error)	1	3.7238564	0.0536404
Robust LM (error)	1	1.4354585	0.2308757

Observa-se que o índice de moran apresenta valor baixo porem significativo. Entretanto os testes de multiplicadores de Lagrange não foram significativos.

Como não se tem certeza do motivo pelo qual o índice de moran se apresenta maior que zero, os modelos LAG e Error foram construídos.

**- Regressão Espacial**

A tabela 13 apresenta o resultado do modelo LAG construído para os dados de MER.

**Tabela 13 – Modelo LAG**

Lag model				
R-squared	:	0.222737		
Log likelihood	:	-1481.28		
<b>Akaike</b>	:	<b>2968.57</b>		
Sigma-square	:	30.6739		
Variable	Coefficient	Std.Error	z-value	Probability
W_SPRCLASSE	0.1060889	0.06928838	1.531121	0.1257394
CONSTANT	29.01912	2.596701	11.17538	0.0000000
ESCOL	1.396646	0.1587185	8.799512	0.0000000
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	3.11582	0.0775355	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	2.520486	0.1123759	

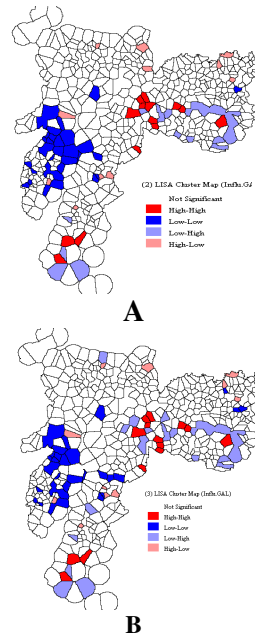
Observa-se a não significância dos parâmetros de ajuste espacial o que indica que o modelo espacial não é adequado.

O modelo Error construído está representado na figura 14. Da mesma forma que o modelo LAG, o modelo error não apresentou significância de seu valor de parâmetro de regressão espacial o que também indica que o modelo espacial não é adequado

**Tabela 14 - Modelo Error Áreas de Influência**

Error model				
R-squared	:	0.224664		
Log likelihood	:	-1480.893220		
<b>Akaike</b>	:	<b>2965.79</b>		
Sigma-square	:	30.597836		
Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	32.66654	1.100937	29.67158	0.0000000
ESCOL	1.534886	0.1472686	10.42236	0.0000000
LAMBDA	0.1252815	0.07072401	1.771414	0.0764917
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST	DF	VALUE	PROB	
Breusch-Pagan test	1	3.22914	0.0723387	
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST	DF	VALUE	PROB	
Likelihood Ratio Test	1	3.302254	0.0691849	

**- Análise de regressão dos Resíduos da Regressão Espacial**



**Figura 15 – Mapa LISA dos Resíduos de Regressão Espacial  
A - LAG Model / B - Error Model**

O índice de moran para o modelo LAG e para o modelo Error são 0.0377 e 0.0387 respectivamente o que indica ausência de agrupamento espacial.

**-Regressão Linear Parâmetro MER**

Os dados de regressão linear estão apresentados na tabela 15.

**Tabela 15 – Regressão Linear MRR**

Regressão Linear					
R-squared	:	0.158646	F-statistic	:	88.8121
Adjusted R-squared	:	0.156860	Prob(F-statistic)	:	1.96204e-019
Sum squared residual:	15705.1	Log likelihood	:	-1499.53	
Sigma-square	:	33.3442	<b>Akaike</b>	:	<b>3003.07</b>
Variable	Coefficient	Std.Error	t-Statistic	Probability	
CONSTANT	40.42066	0.4389337	92.08829	0.0000000	
RENDA	0.002701565	0.0002866681	9.424018	0.0000000	
MULTICOLLINEARITY CONDITION NUMBER 2.969611					
TEST ON NORMALITY OF ERRORS					
TEST	DF	VALUE	PROB		
Jarque-Bera	2	604.4333	0.0000000		
DIAGNOSTICS FOR HETEROSKEDASTICITY					
TEST	DF	VALUE	PROB		
Breusch-Pagan test	1	4.31293	0.0378238		
Koenker-Bassett test	1	1.220358	0.2692904		

A partir dos dados da regressão linear, percebe-se que o parâmetro da inclinação estimado é significativo porém de valor muito baixo.

Os testes de para normalidade e heterocedasticidade demonstram que os resíduos não apresentam normalidade e também não apresentam variância constante.

- Análise dos Resíduos

Tabela 16 – Dependência Espacial dos Resíduos

Dependência Espacial dos Resíduos			
TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.098053	3.5886575	0.0003325
Lagrange Multiplier (lag)	1	12.1225758	0.0004982
Robust LM (lag)	1	0.3944644	0.5299620
Lagrange Multiplier (error)	1	11.7412285	0.0006113
Robust LM (error)	1	0.0131171	0.9088179

A análise dos resíduos demonstra que o índice de moram tem um baixo valor porem significativo, o que pode indicar autocorrelação espacial.

Ambos os testes de multiplicadores de Lagrange, demonstram que há correlação espacial é que os dois modelos podem explicar esta dependência. Porem o teste robusto para autocorrelação espacial dos resíduos não é significativo em nenhum dos casos o que necessita da construção do modelo para que se verifique se algum deles pode ou não melhorar a análise de regressão.

- Regressão Espacial

A tabela 17 demonstra os resultados para o modelo LAG.

Tabela 17 – Modelo LAG para MRR

Lag Model				
R-squared	:	0.182827		
Log likelihood	:	-1494.5		
Akaike	:	2995.01		
Sigma-square	:	32.2489		
Variable	Coefficient	Std.Error	z-value	Probability
W_NOTA	0.204538	0.06673263	3.065037	0.0021766
CONSTANT	32.0265	2.780367	11.5188	0.0000000
RENDA	0.002274944	0.0003159685	7.199907	0.0000000
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST		DF	VALUE	PROB
Breusch-Pagan test		1	1.769152	0.1834871
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST		DF	VALUE	PROB
Likelihood Ratio Test		1	10.06154	0.0015140

Os resultados expressos na tabela demonstram que o parâmetro de regressão

espacial é significativo e que o valor de AKAIKE diminui em relação ao valor na regressão linear.

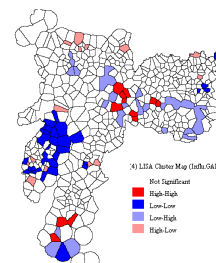
Isto pode indicar que a presença relação espacial no modelo ajuda a explicar a relação entre a nota de uma escola e a renda.

No modelo Error que também apresentou significância de seu parâmetro espacial ( Tabela 18), o valor de AKAIKE foi mais baixo ainda do que ao apresentado no modelo LAG. Isto pode indicar que a relação espacial esta mais presente quando se considera o resíduo de da regressão simples do que à variável resposta, no caso, a Nota de uma Escola.

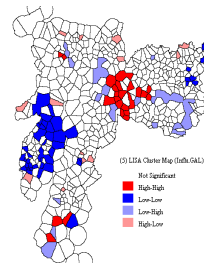
Tabela 18 – Modelo Error para MRR

Error Model				
R-squared	:	0.182936		
Log likelihood	:	-1494.569577		
Sigma-square	:	32.244615		
Akaike	:	2993.14		
Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	40.47604	0.5249647	77.1024	0.0000000
RENDA	0.002687649	0.0003349548	8.023918	0.0000000
LAMBDA	0.2096965	0.0680173	3.082988	0.0020495
DIAGNOSTICS FOR HETEROSKEDASTICITY				
TEST		DF	VALUE	PROB
Breusch-Pagan test		1	2.415484	0.1201408
DIAGNOSTICS FOR SPATIAL DEPENDENCE				
TEST		DF	VALUE	PROB
Likelihood Ratio Test		1	9.928069	0.0016278

- Análise de Regressão dos Resíduos da Regressão Espacial



A



B

Figura 16 – Mapa LISA dos Resíduos de Regressão Espacial  
A - LAG Model / B - Error Model

O índice de moran para o modelo LAG e para o modelo Error são - 0,0466 e - 0.0398 respectivamente o que indica ausência de agrupamento espacial. O s agrupamento apresentados no Mapa Lisa/Moran, podem representar algum agrupamento, porem a contraste com o índice de moran global reforça a idéia de ausência de agrupamento.

## - Discussão Dos Resultados

As análises realizadas nesse trabalho buscaram relacionar os dados de exames de rendimento escolar com a situação socioeconômica da região onde está localizada a escola.

Um ponto relevante é que o número de escolas com o dado “nota” é baixo em relação ao universo de escolas do município de São Paulo.. Outra dificuldade da análise é que não se sabe ao certo se o universo de alunos que realiza a prova do SARESP reside numa região que esta contém sua moradia. Em outras palavras, não se sabe ao certo se o par escolaridade/renda de uma subdivisão contém família do aluno que frequênta a escola.

Inicialmente foi adotada a divisão de distritos. Observa-se através dos mapas de distribuição de dados agregados que renda e escolaridades se apresentam com valores mais altos na região central da cidade. Para os dados de “Nota” a distribuição também se aproxima desta distribuição. Uma limitação deste tipo de análise é que alguns distritos apresentam maior quantidade de escolas faz com que aconteçam aberrações como o distrito localizado mais ao sul do município, que possuía apenas uma escola com uma nota relativamente alta considerando-se a agregação por médias.

Um ponto que também deve ser ressaltado é a necessidade da exclusão de distritos que não possuem dados. Esta alternativa, para a realização das análises por regressão, distorce o dado, pois, principalmente na região central da cidade se encontram os maiores índices de renda e escolaridade.

Os testes de Autocorrelação espacial demonstraram de acordo com o índice de moran, que existe dependência espacial para a variável resposta. Outro teste que apontou nesta direção foi o teste de moran para dependência espacial dos resíduos, que demonstrou autocorrelação espacial dos resíduos de regressão linear simples. Esta correlação levou a crer que a inclusão desta dependência na regressão linear, contribuiria para uma “melhora” do modelo.

Entretanto, os testes de multiplicadores de Lagrange demonstraram não haver possibilidade de aplicação dos modelos espacial LAG e espacial Error.

Para verificação, os dois modelos foram construídos demonstrando que a inclusão da dependência espacial não melhora o modelo linear simples pois os coeficientes para regressão espacial encontrados não apresentaram significância. O estudo dos resíduos da regressão espacial indicou baixa autocorrelação espacial, o que descarta a hipótese do uso de um modelo de regressão espacial local.

De uma forma geral, a análise por distritos já não apresentaria relevância significativa pois dentro de um distrito há uma variação dos dados socioeconômicos que é extremamente suavizada pela média obtida quando se dá agregação. Em outras palavras, poderia se incorrer no erro de falácia ecológica quando se tentasse admitir a correlação entre os dados de NME MRR e MRE para a “escala do indivíduo”.

Desta forma, adotou-se então a subdivisão por microáreas como forma inicial de retirar a suavização dos dados. Um problema que ocorreu foi a grande quantidade de microáreas excluídas o que pode ter distorcido a análise.

A subdivisão por microáreas foi a única que apresentou autocorrelação da variável NME e autocorrelação dos resíduos de regressão linear indicados pelo índice de moran, multiplicador de Lagrange e testes robusto do modelo LAG. A inclusão da dependência espacial no modelo resultou numa leve melhora do modelo de regressão linear com valores de AKAIKE passando de 2060 na regressão linear para 2054 na regressão LAG e 2055 na regressão Error. O estudo dos resíduos de regressão espacial também apresentou baixa autocorrelação espacial.

Porém neste modelo ainda persiste o problema de agregação da variável NME o que torna pouco representativa esse tipo de análise. Desta forma se partiu para a análise em áreas de Influencia. A observação do índice de moran para áreas de influencia para a variavel “nota mostra uma dependência espacial baixa. O mapa de moran entretanto demonstra que existe uma certa relevância na dependência espacial dos valores de nota.

É interessante notar que o valor do índice de moran vai caindo quando se passa de uma delimitação de maiores áreas para menores, o que pode indicar que quando se diminui a área de abrangência o caráter

particular de cada escola vai predominando e diminuindo então sua autocorrelação espacial.

O modelo espacial construído para a variável MER demonstrou, como já se poderia prever que a inclusão da autocorrelação não ajudou a explicar a baixa dependência em relação à variável nota. Tanto o modelo Error como o modelo LAG não se apresentaram significativos. Os teste para autocorrelação espacial do resíduo também não apresentou dependência espacial. De certa forma já se esperava que os modelos de regressão espacial não surtisses efeito visto que os testes de dependência espacial já indicavam a baixa autocorrelação dos resíduos.

Na tentativa de compreender também o que ocorre para variável MRR, foi construído o modelo linear simples. A regressão espacial demonstrou que o ajuste do modelo foi baixo porém significativo. Os testes de dependência espacial dos resíduos demonstraram que existe uma baixa correlação espacial porém significativa.

Os modelos LAG e Error foram construído para esta variável demonstrando que a inclusão da relação espacial no modelo melhora um pouco o valor de AKAIKE, passando de 3003 na regressão linear para 2005 no modelo LAG e 2993 no modelo Error. O estudo dos resíduos de regressão espacial demonstrou que não há correlação espacial.

Por fim é interessante notar que os resultados têm uma forte dependência com a área considerada. Apesar da diminuição do valor do índice de moran para a variável NME isso não indica que o uso de um modelo espacial será adequado.

#### - CONCLUSÃO

Este trabalho procurou estudar a relação entre dados do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP) do ano de 2001, para 581 escolas pertencentes ao município e São Paulo e dados de “*Média do rendimento nominal mensal das pessoas responsáveis por domicílios particulares permanentes (MRR)*” e “*Média do número de anos de estudo das pessoas responsáveis por domicílios particulares permanentes (MER)* extraídos dos Dados do Censo Demográfico de 2000, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

Vários tipos de agregação foram escolhidos de forma a testar qual delas era mais representativa. As regiões de agregação

foram definidas de acordo com subdivisões pré-estabelecidas como “Distritos” e “Microáreas” que são subdivisões da prefeitura municipal de São Paulo. Uma outra subdivisão foi estimada a partir dos critérios de Zoneamento escolar município.

A dependência espacial da variável resposta (NME) foi estudada e apresentou resultados diferenciados para a cada uma das subdivisões. Nota-se através do índice de moran que a autocorrelação espacial vai diminuindo com a diminuição da área de agregação. Isso pode indicar que questões locais, inerentes à cada escola, se mostram mais significativos com a diminuição da área de agregação.

A tentativa de inclusão da dependência espacial no modelo explicativo apresentou –se satisfatória apenas para um tipo de subdivisão. Tanto Distritos como Áreas de Influência, no caso da relação Nota/MRE, demonstraram que a inclusão da variável espacial não apresentou melhora do modelo. Microáreas e Áreas de Influência, na relação Nota/MRR, apresentaram uma pequena mas significativa melhora com a inclusão de variáveis espaciais.

Entretanto, em nenhuma das subdivisões encontradas pode se dizer que houve uma melhora capaz de justificar o uso da relação espacial no modelo explicativo. Tal fato pode ser explicado por dois motivos principais. Um deles é que a condição de subdivisão não reflete a condição social do aluno uma vez que as áreas consideradas não correspondem com certeza à residência do aluno da escola. Outro ponto é admitir que a auto correlação espacial entre a variável NME ou nota é baixa, e portanto um modelo espacial não ajuda explicar relação entre as variáveis propostas.

Por fim, estudos posteriores devem ser feitos buscando diferentes áreas de agregação que possam representar com maior fidelidade a relação espacial das variáveis que foram estudadas neste trabalho.

#### - REFERENCIAS BIBLIOGRÁFICAS

Druck, S.; Carvalho, M.S.; Câmara, G.; Monteiro, A.V.M. (eds) *Análise Espacial de Dados Geográficos*. Brasília, EMBRAPA, 2004 (ISBN: 85-7383-260-6). Disponível em : <<http://www.dpi.inpe.br/gilberto/livro/analise/>> em 11/12/2008.

INFOLOCAL. **Sistema de Informação para o Planejamento Local** - SEMPLA- DEINFO - Departamento de Informações Disponível em < [infolocal.prefeitura.sp.gov.br/](http://infolocal.prefeitura.sp.gov.br/)> em 11/12/2008.

SABATINI, Francisco, CACERES, Gonzalo y CERDA, Jorge. **Segregación residencial en las principales ciudades chilenas: Tendencias de las tres últimas décadas y posibles cursos de acción.** *EURE (Santiago)*. [online]. dic. 2001, vol.27, no.82 [citado 11 Diciembre 2008], p.21-42. Disponible en la World Wide Web: <[http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0250-71612001008200002&lng=es&nrm=iso](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0250-71612001008200002&lng=es&nrm=iso)>. ISSN 0250-7161.

TORRES, Haroldo da Gama, PAVEZ Thais Regina GOMES Sandra, BICHIR, Renata Mirandola. **Educação Na Periferia De São Paulo: Ou Como Pensar As Desigualdades Educacionais?** Texto apresentado no Workshop “Neighborhood effects, educational achievements and challenges for social policies”. Rio de Janeiro, agosto de 2006. Disponível na World Wide Web <[www.centrodametropole.org.br/pdf/2007/Torres\\_\\_Carpim\\_\\_Gomes\\_e\\_Bichir\\_\\_Educacao\\_final.pdf](http://www.centrodametropole.org.br/pdf/2007/Torres__Carpim__Gomes_e_Bichir__Educacao_final.pdf)>

Torres, Haroldo, GOMES, Sandra **Desigualdade Educacional e Segregação Social na região Metropolitana de São Paulo.** Novos Estudos – CEBRAP n 64, PP 132-140 (2002).