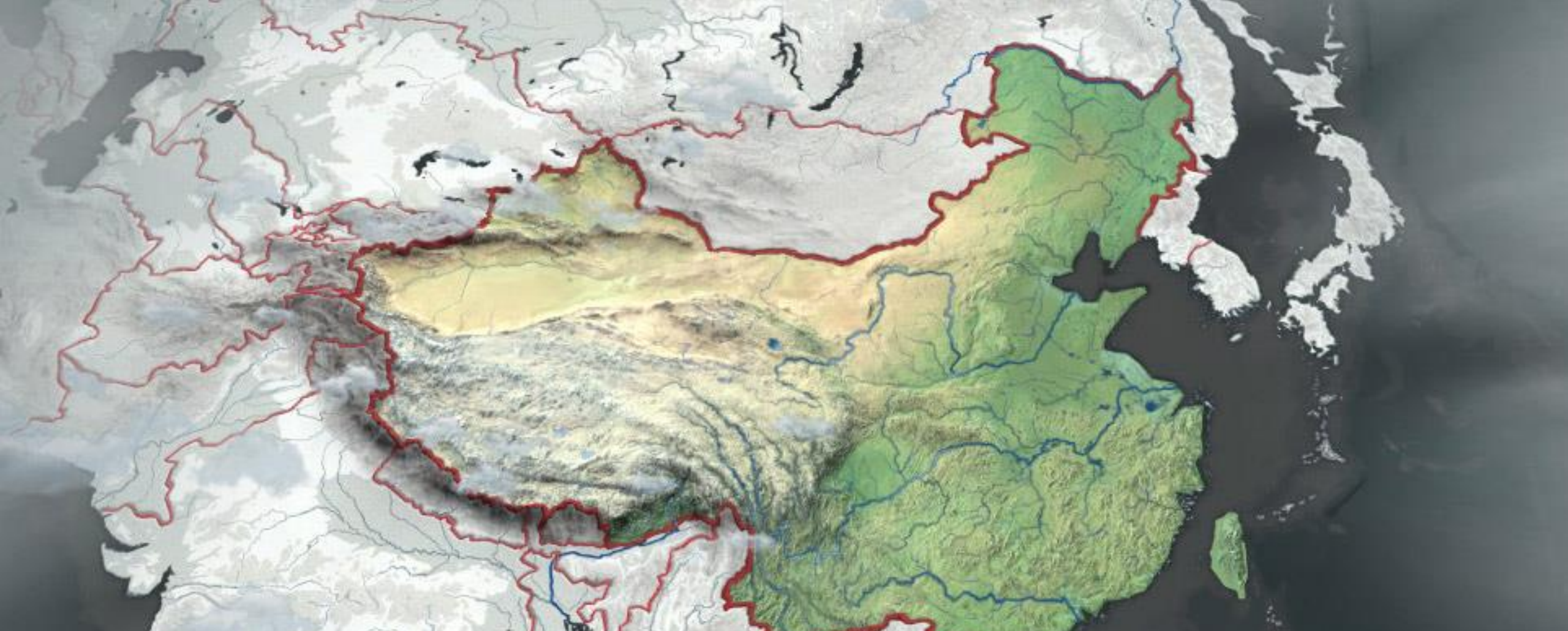




# Socio-Demographic Predictors and Distribution of Pulmonary Tuberculosis (TB) in Xinjiang, China: A Spatial Analysis

Atikaimu Wubuli, Feng Xue, Daobin Jiang, Xuemei Yao, Halmurat Upur, Qimanguli Wushouer (2015)

RAÍSSA CAROLINE DOS SANTOS TEIXEIRA

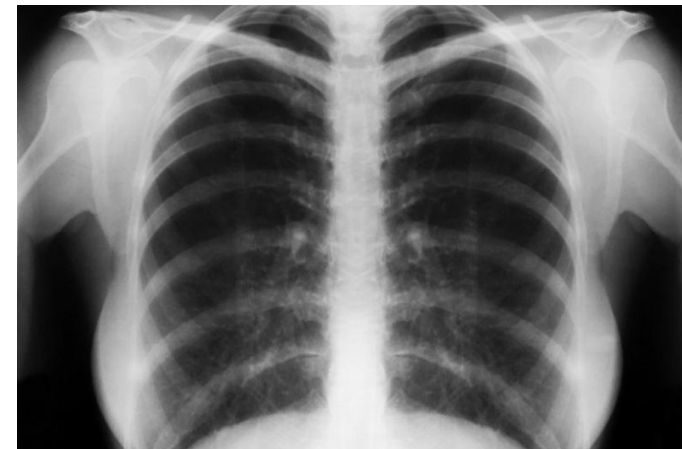
DISCIPLINA: ANÁLISE ESPACIAL – SER 301

DOCENTE: DR. A. MIGUEL VIEIRA MONTEIRO

PG-SER INPE 2019

# Tuberculose

- **Doença infecciosa** que afeta prioritariamente os **pulmões**, embora possa acometer outros órgãos e/ou sistemas.
- Causada por bactérias (gênero *Mycobacterium*), e o bacilo é transmitido por via **aérea na maioria dos casos**.
- É influenciada pela determinação social, apresentando uma relação direta com a **pobreza e a exclusão social**.
- Diagnóstico: exame clínico, exame radiológico de tórax, prova ou reação tuberculínica e **diagnóstico laboratorial** (baciloscopia e a cultura do escarro).



# Justificativa e objetivos do trabalho

A tuberculose (TB) → **problema de saúde global!**

Em 2012 → ~8,6 milhões de novos casos com 1,3 milhão de **mortes** → Muitos pacientes e óbitos para uma **doença curável**.

China → **12% da incidência global**.

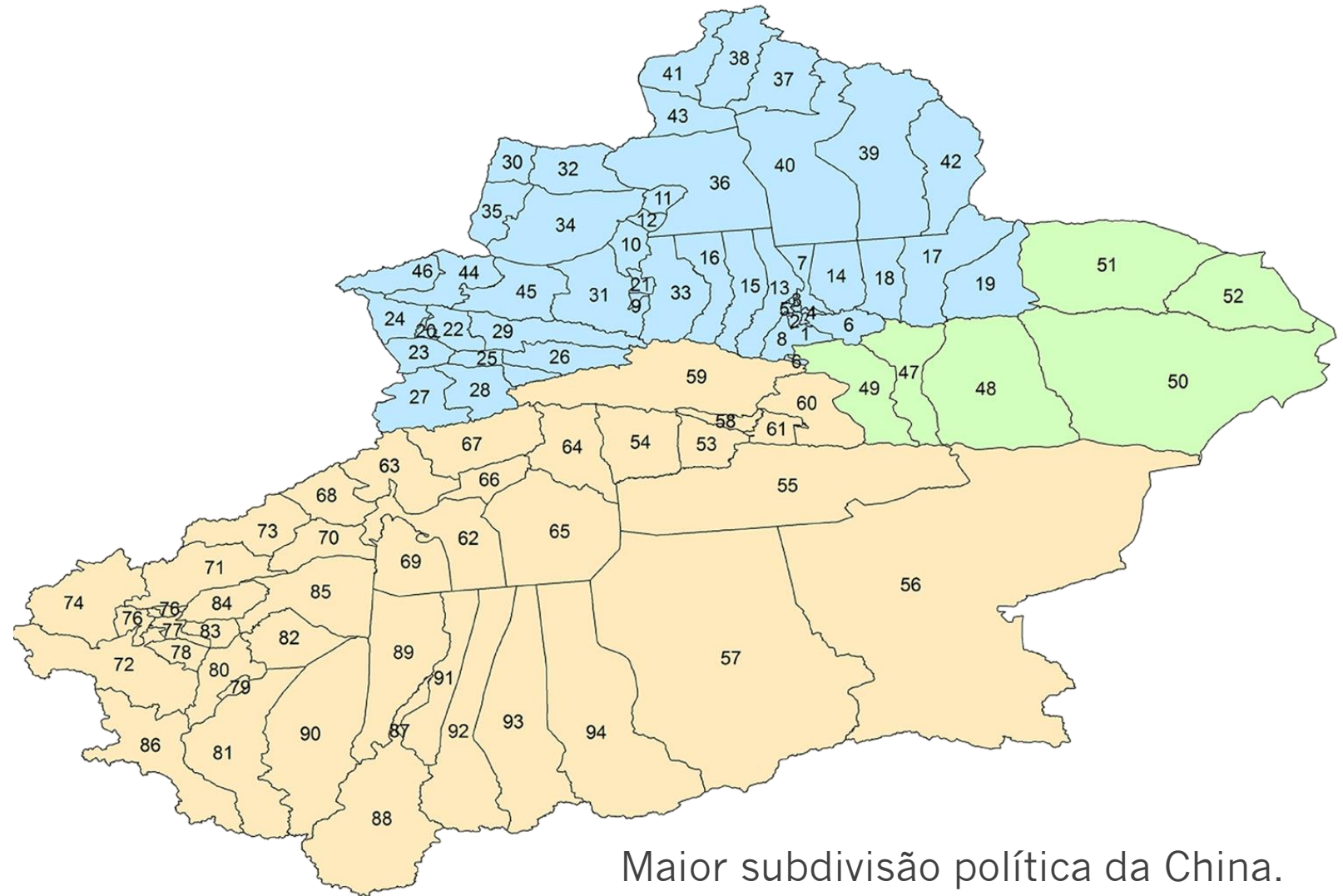
- **Xinjiang** é uma das províncias da China **com maior ocorrência de TB**.

Foi realizada uma **análise espacial** utilizando SIG para melhorar o **entendimento da variação geográfica** da ocorrência da TB pulmonar em Xinjiang, seus determinantes dentro de um contexto social, e compreender o contexto espacial e temporal como forma de melhorar o **direcionamento de intervenções e recursos**.

# Área de estudo



RICH CLABAUGH/STAFF



Maior subdivisão política da China.

14 prefeituras → 94 distritos

# Fontes de dados e definições de variáveis

- *National Infectious Diseases Reporting System (NIDRS), Chinese Center for Disease Control and Prevention* → Número de **casos de Tuberculose pulmonar** (nível de condado/distrito).
- 
- *Xinjiang Statistical Yearbook (2006~2014)* → **Dados populacionais e sociodemográficos** → indicadores/preditores da doença.

## Cálculo da incidência

- Quatro taxas de incidência foram calculadas:
  - TB pulmonar,
  - SS+TB (TB pulmonar confirmada por baciloscopia)
  - Novos SS+TB (casos de SS+TB sem histórico de tratamento para TB)
  - SS-TB (TB com baciloscopia negativa)

# Análise espacial

**Análise de Autocorrelação Espacial** → Avaliar o grau de agrupamento, aleatoriedade ou fragmentação de um padrão espacial.

- **Dependência Espacial Global** → estima o grau geral de autocorrelação espacial de um conjunto de dados.
- **Dependência Espacial Local** → Identifica a localização e os tipos de clusters.

$$\text{Índice global de Moran (I): } I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} z_i z_j}{S_0 \sum_{i=1}^n z_i^2}$$

O valor absoluto do I indica a **força da autocorrelação** espacial.

A significância estatística é testada pelo Z-score e valor de P.

$P < 0,05$  leva à rejeição da hipótese nula e indica a presença de autocorrelação espacial.

# Análise espacial

**Índice local de Moran (I<sub>i</sub>):**

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{ij} (x_j - \bar{X})$$

Os tipos de autocorrelação espacial local incluem:

- HH → “hotspots”
- HL
- LL → “coldspots”
- LH

**Índice local de Getis e Ord (G<sub>i</sub><sup>\*</sup>):**

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{\left[ \sum_{j=1}^n w_{ij}^2 - \frac{\left( \sum_{j=1}^n w_{ij} \right)^2}{n-1} \right]}}$$

Estatística local para avaliar a presença de clusters e fornecer informações adicionais sobre a intensidade e estabilidade dos clusters principais de hotspots/coldspots.

# Análise de Regressão Espacial

- **1ª etapa: OLS - Método dos Mínimos Quadrados (Ordinary Least Squares)**
  - Técnica de otimização matemática que procura encontrar o melhor ajuste para um conjunto de dados tentando minimizar a soma dos quadrados dos resíduos.
- **2ª etapa: SLM - Spatial Lag Model**
  - Este modelo avalia a força da relação espacial entre os índices em todas as áreas contíguas.
- **3ª etapa: GWR - Geographically-weighted Regression (regressão geograficamente ponderada)**
  - Modelo de Regressão com efeitos espaciais contínuos → Variações espaciais modeladas de forma contínua, com parâmetros variando no espaço.

# Análise de Regressão Espacial

- **OLS**

- Ponto de partida para todas as análises de regressão espacial → Limitações para dados espaciais;
- os coeficientes para as variáveis explicativas (isentas de multicolinearidade) do modelo devem ser estatisticamente significativos;
- os resíduos devem ser normalmente distribuídos com uma média de zero e **livres de autocorrelação espacial**.

- **SLM**

- Escolhido como modelo de regressão espacial global com base nas estatísticas de testes multiplicadores de Lagrange (LM);
- O SLM **incorpora diretamente a autocorrelação espacial** no modelo, incluindo um termo de defasagem espacial ( $\rho$ ).
- Formalmente, este modelo é  $y = \rho Wy + X\beta + \varepsilon$ ,
  - onde  $y$ : vetor de observações na variável dependente;
  - $Wy$ : variável dependente espacialmente atrasada para a matriz de pesos  $W$ ;
  - $X$  é uma matriz de observações nas variáveis independentes,  $\varepsilon$  é um vetor de termos de erro,  $\rho$  e  $\beta$  são coeficientes de regressão.

# Análise de Regressão Espacial

- **GWR**

- Usada como modelo de regressão espacial local para explorar a variação espacial nas relações entre a incidência de TB pulmonar e os preditores.
- Este trabalho escolheu o AIC (critério de informação de Akaike), que foi fixado pelo princípio da **máxima verossimilhança** para determinar a largura de banda ideal.
- O Índice de Moran foi utilizado para testar a autocorrelação espacial dos resíduos.
- A verossimilhança de log e o valor de AIC foram usados para comparar a adequação do OLS e SLM, enquanto o valor ajustado de  $R^2$  e AIC foram usados para comparar a adequação do OLS e GWR.

➤ Os softwares **SPSS** 17.0, **ArcGIS** 10.2.2 e **GeoDa** foram utilizados para análise dos dados.

➤ National Geographic Information System database → limite da região de Xinjiang.

# Tendência da incidência de TB pulmonar de 2005 a 2013

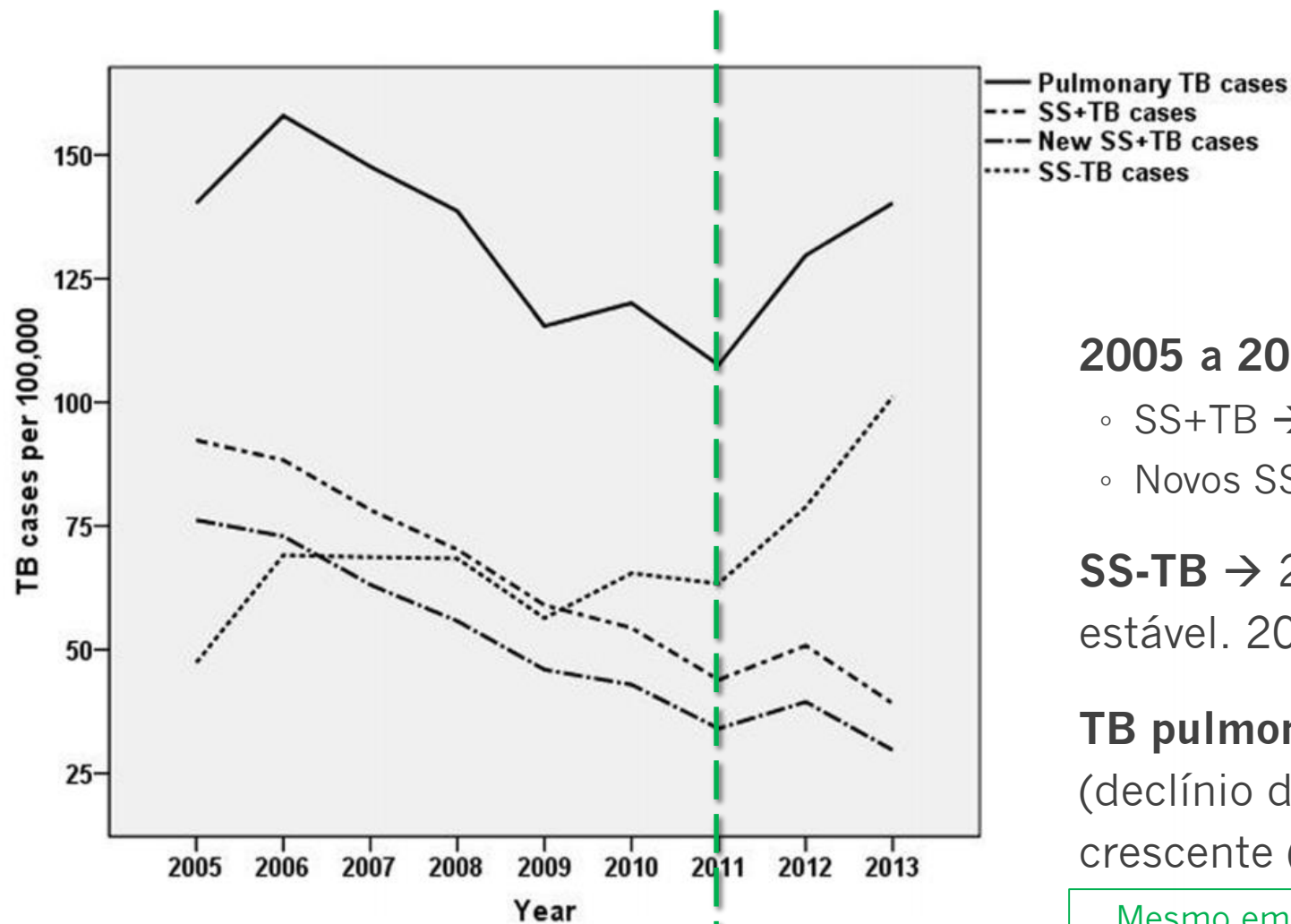


Fig 1. The trend of pulmonary TB incidence from 2005 to 2013.

Incidências médias anuais:

- TB pulmonar → 132,59
- SS+TB → 63,17
- Novos SS+TB → 50,36

**2005 a 2013** - Tendência no declínio:

- SS+TB → 92 para 39
- Novos SS+TB → 76 para 30

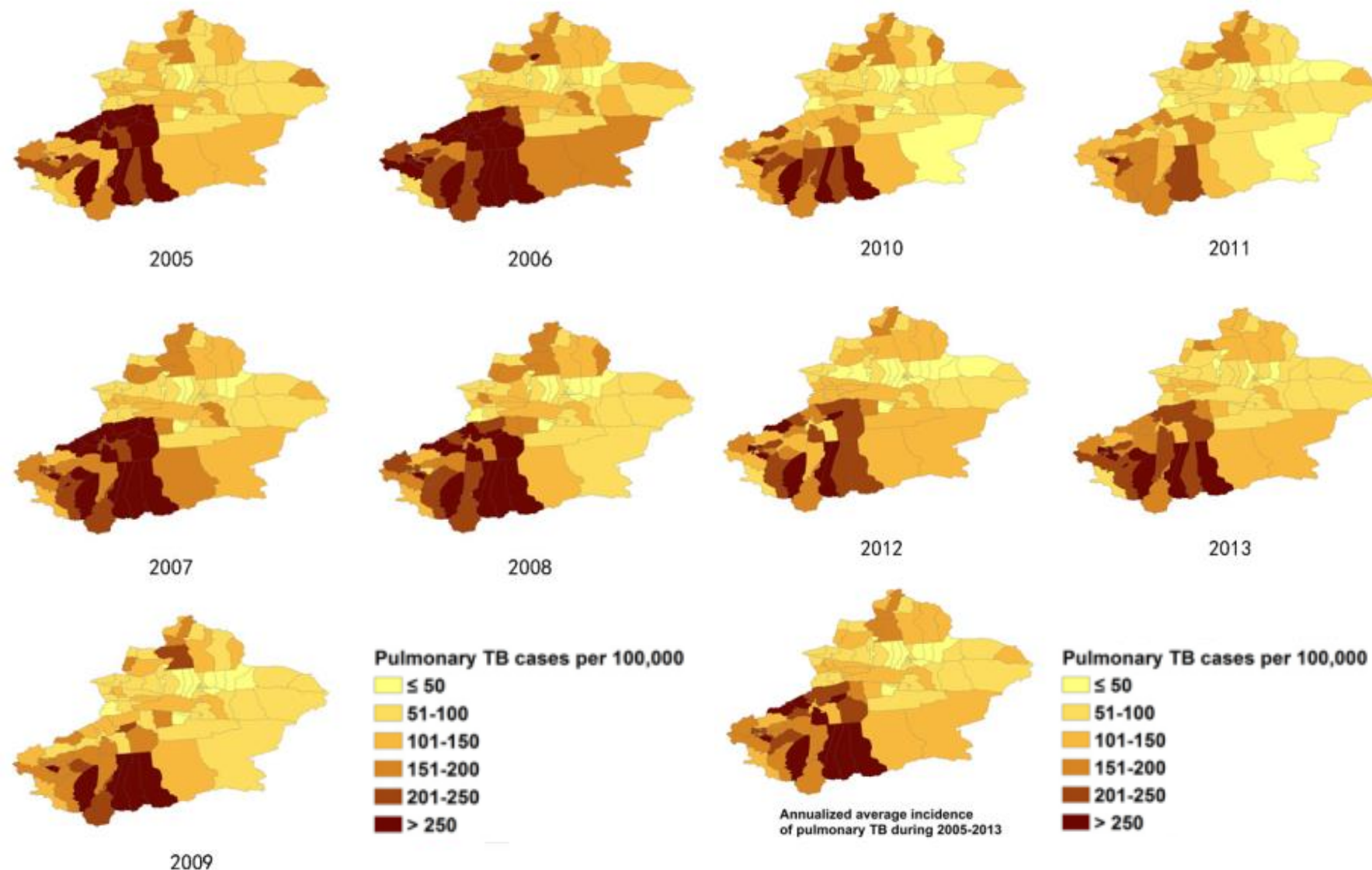
Eficácia do controle da TB nas últimas décadas!

**SS-TB** → 2005 – 2010: Incidência relativamente estável. 2011: Tendência de aumento.

**TB pulmonar** → 2005 – 2011: Tendência de declínio (declínio dos casos de SS+TB). 2011: Tendência crescente (aumento de casos de SS-TB).

Mesmo em declínio, as taxas de incidência de TB de Xinjiang ainda eram mais altas que a média nacional

# Distribuição da incidência de TB pulmonar



- Diferenças regionais.
- Em geral, o **sul de Xinjiang** (parte sul das montanhas de Tianshan), especialmente a parte sudoeste de Xinjiang, teve uma **alta incidência** de TB pulmonar e SS + TB.

# Distribuição da incidência de SS+TB

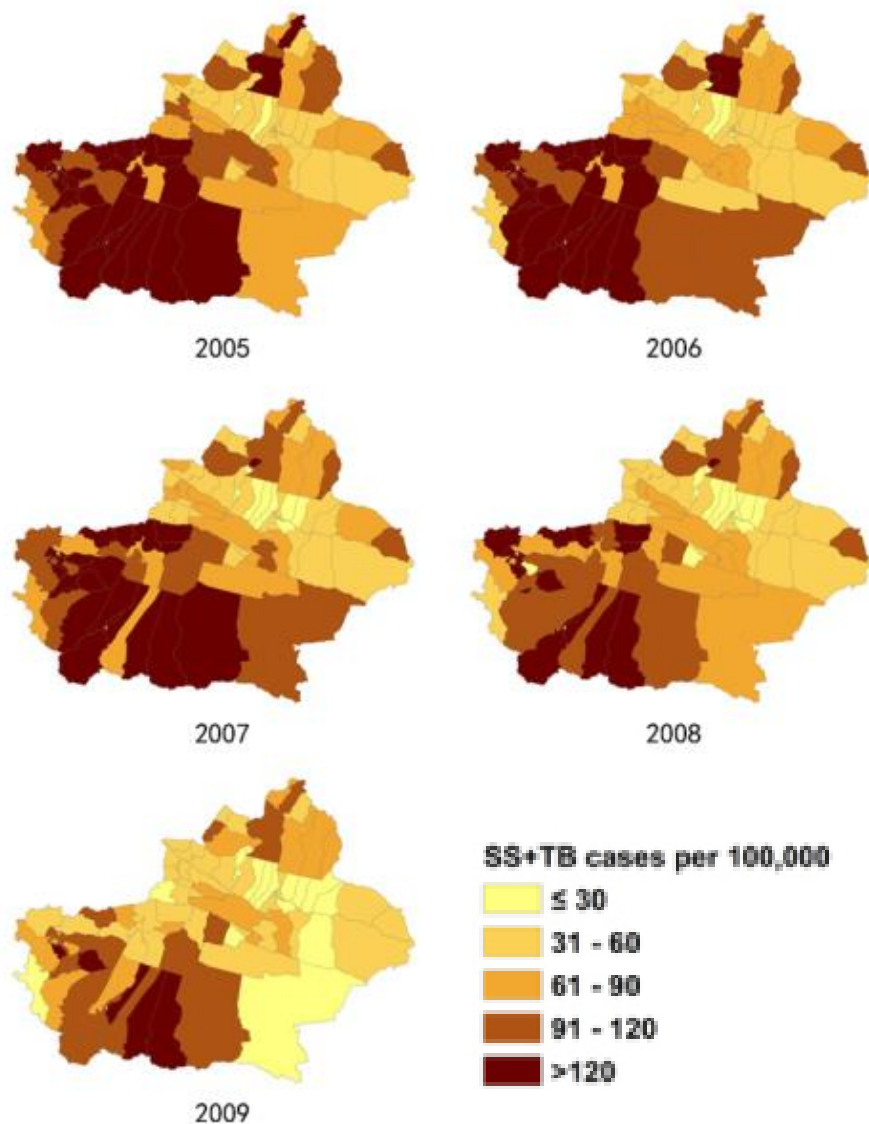


Fig 4. The incidence of SS+TB cases in Xinjiang, from 2005–2009.

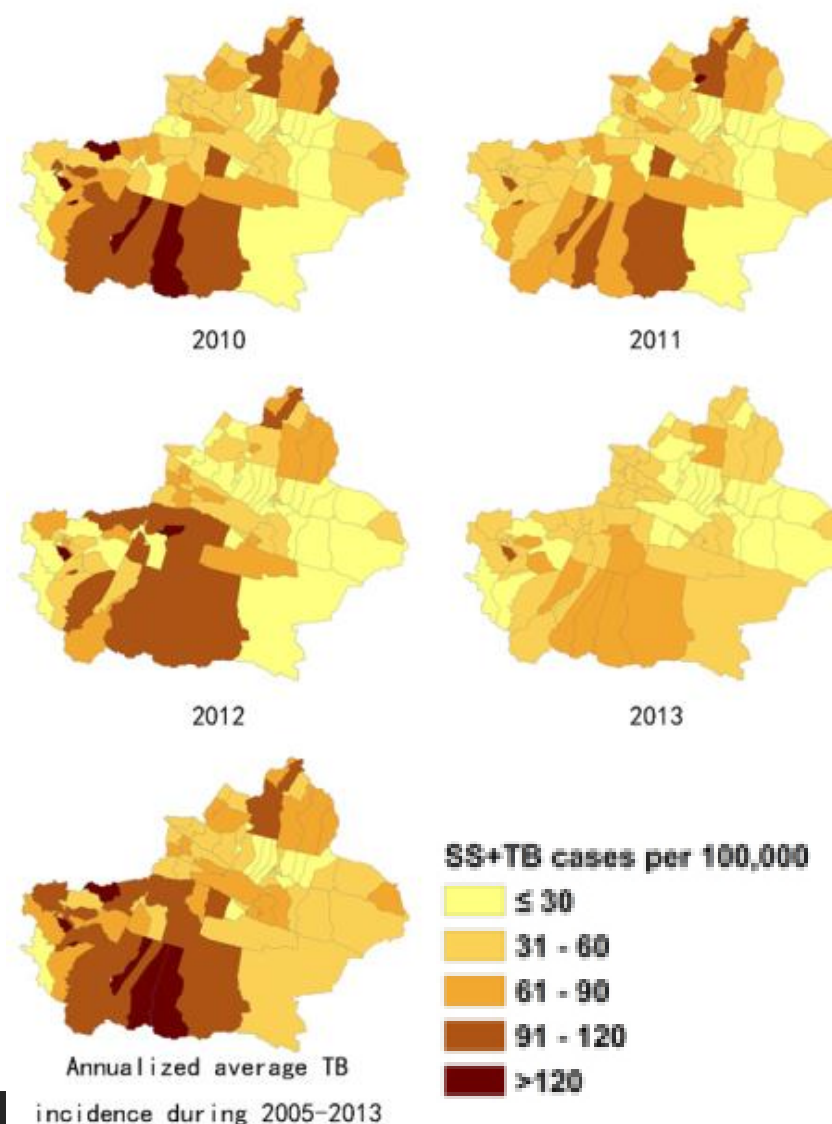


Fig 5. The incidence of SS+TB cases in Xinjiang, from 2010–2013.

# Análises de autocorrelação espacial global

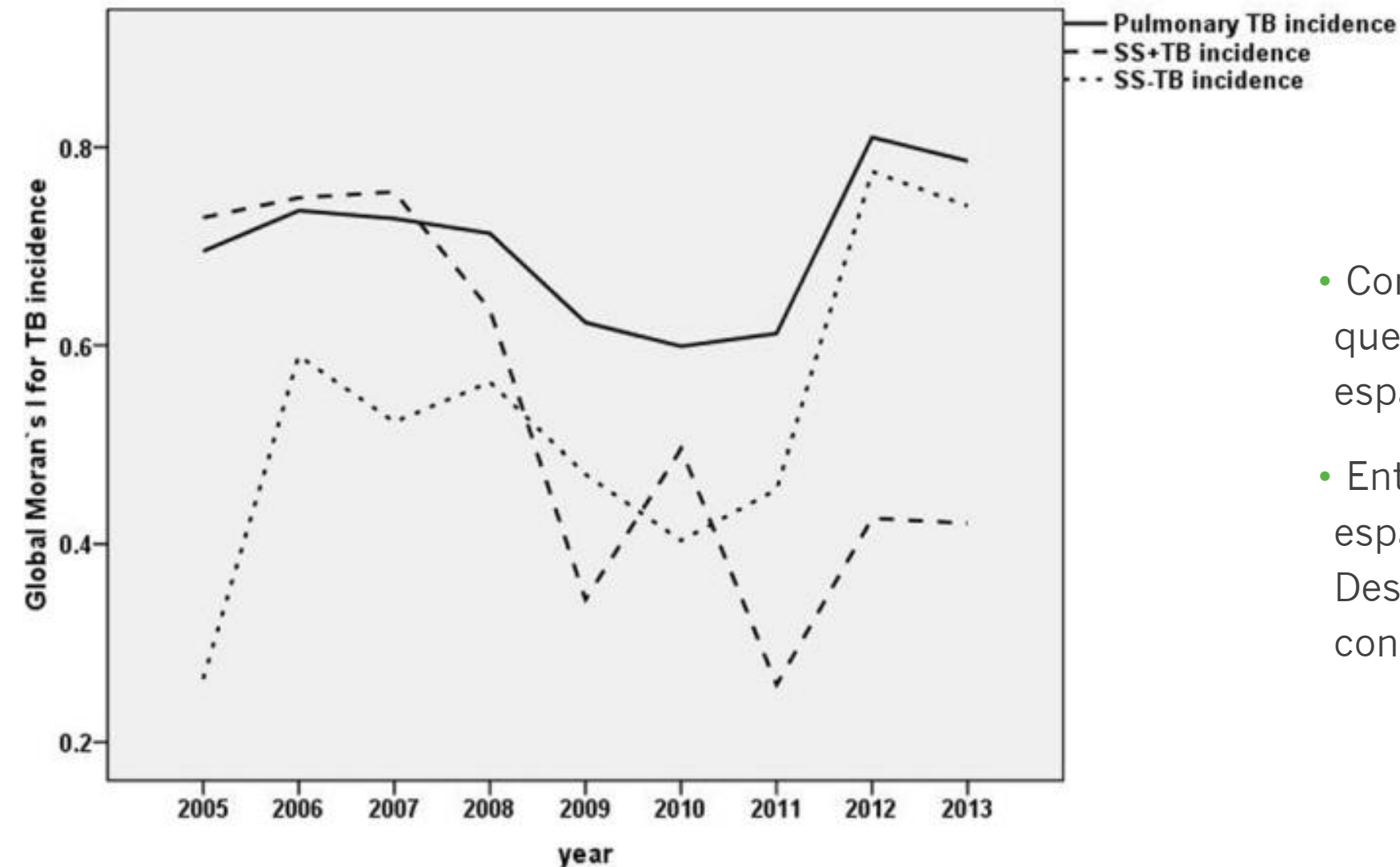
- Presença de autocorrelação espacial **positiva** (Moran global  $I > 0$ ) na incidência de TB pulmonar, SS+TB e SS-TB de 2005 a 2013.

**Table 1. Results of the global spatial autocorrelation analysis of pulmonary TB incidence from 2005–2013.**

| Year                         | Pulmonary TB incidence |        |         | SS+ TB incidence |        |         | SS- TB incidence |        |         |
|------------------------------|------------------------|--------|---------|------------------|--------|---------|------------------|--------|---------|
|                              | Moran's I              | Z      | P       | Moran's I        | Z      | P       | Moran's I        | Z      | P       |
| 2005                         | 0.695                  | 18.497 | <0.0001 | 0.729            | 19.389 | <0.0001 | 0.264            | 7.132  | <0.0001 |
| 2006                         | 0.736                  | 19.328 | <0.0001 | 0.749            | 19.712 | <0.0001 | 0.590            | 15.560 | <0.0001 |
| 2007                         | 0.728                  | 19.130 | <0.0001 | 0.755            | 19.873 | <0.0001 | 0.523            | 13.821 | <0.0001 |
| 2008                         | 0.713                  | 18.746 | <0.0001 | 0.636            | 16.757 | <0.0001 | 0.563            | 14.847 | <0.0001 |
| 2009                         | 0.623                  | 16.429 | <0.0001 | 0.344            | 9.464  | <0.0001 | 0.470            | 12.525 | <0.0001 |
| 2010                         | 0.599                  | 15.818 | <0.0001 | 0.496            | 13.088 | <0.0001 | 0.404            | 10.804 | <0.0001 |
| 2011                         | 0.612                  | 16.108 | <0.0001 | 0.258            | 7.026  | <0.0001 | 0.455            | 12.148 | <0.0001 |
| 2012                         | 0.810                  | 21.228 | <0.0001 | 0.426            | 11.327 | <0.0001 | 0.775            | 20.336 | <0.0001 |
| 2013                         | 0.786                  | 20.634 | <0.0001 | 0.421            | 11.013 | <0.0001 | 0.741            | 19.449 | <0.0001 |
| Annualized average incidence | 0.78                   | 20.733 | <0.0001 | 0.717            | 18.843 | <0.0001 | 0.696            | 18.277 | <0.0001 |

Variação temporal significativa

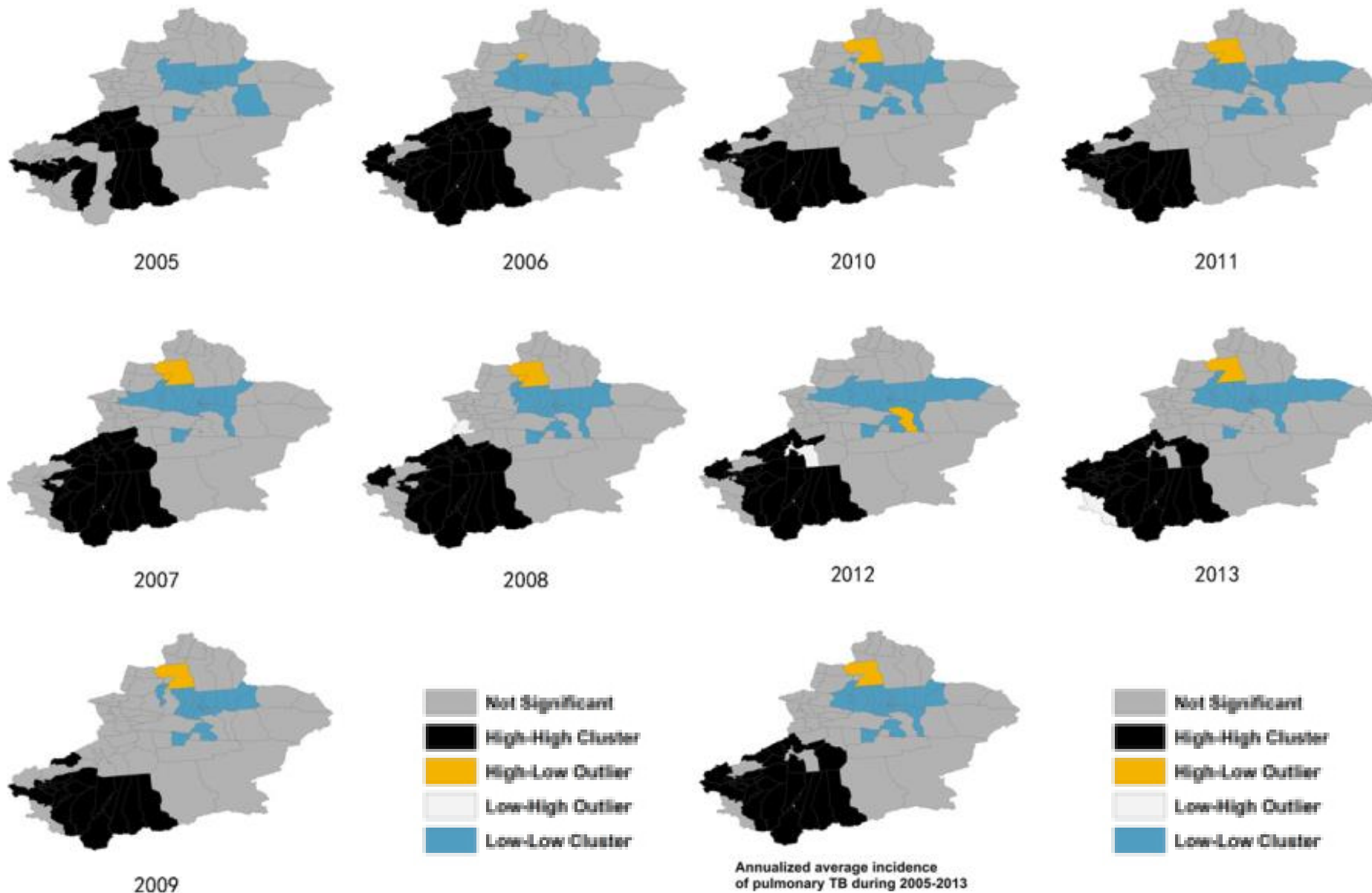
# Análises de autocorrelação espacial global



- Consistente com outros estudos que mostraram autocorrelação espacial significativa.
- Entender as características espaciais da distribuição da TB → Desenvolvimento de políticas de controle de TB mais direcionadas.

Fig 6. The trend of Global Moran's I value for pulmonary TB incidence from 2005 to 2013.

# Autocorrelação espacial local da incidência da TB pulmonar



- A análise espacial local revelou “hotspots” e “coldspots”;
- Mudança significativa ao longo do tempo.

Fig 7. Clusters of the Anselin Local Moran's I analysis, from 2005–2009.

Fig 8. Clusters of the Anselin Local Moran's I analysis, from 2010–2013.

# Análise de *hotspots* com Getis e Ord (Gi\*)

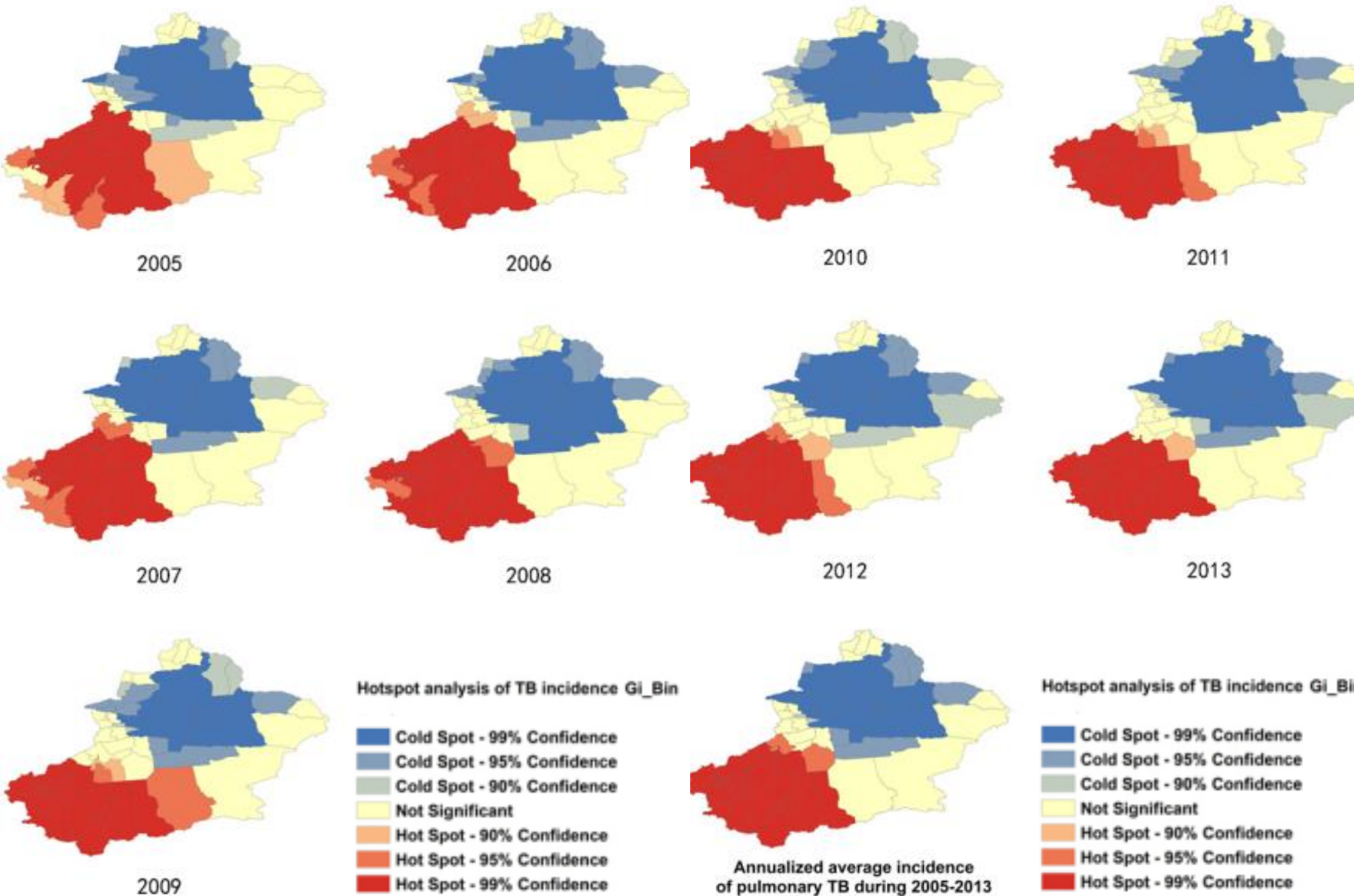


Fig 9. Hotspot Analysis with Getis-Ord Gi\* statistic, from 2008–2009.

Fig 10. Hotspot Analysis with Getis-Ord Gi\* statistic, from 2010–2013.

- A análise com a estatística Getis-Ord forneceu mais informações que indicam a **intensidade e a estabilidade dos clusters** principais de hotspot/coldspot.
- Dois aglomerados parecem separados pelas **montanhas Tianshan** em Xinjiang → diferenças climáticas → influencia agricultura e pobreza.
- Fronteiras com outros países.

# Preditores sociodemográficos da incidência de TB pulmonar

## OLS

- Seis variáveis independentes como “densidade populacional”, “proporção de homens”, “proporção de minorias”, “proporção de população rural”, “taxa de mortalidade” e “PIB per capita” foram selecionadas de acordo com a revisão da literatura e a disponibilidade de dados.
- Duas variáveis estatisticamente significativas:
  - “**proporção de minorias**” → efeito **positivo**
  - “**PIB per capita**” → efeito **negativo**.
- A maioria dos requisitos do método OLS foram atendidos, no entanto, a autocorrelação estatística significativa nos resíduos indicou **autocorrelação espacial das variáveis e natureza não estacionária**.

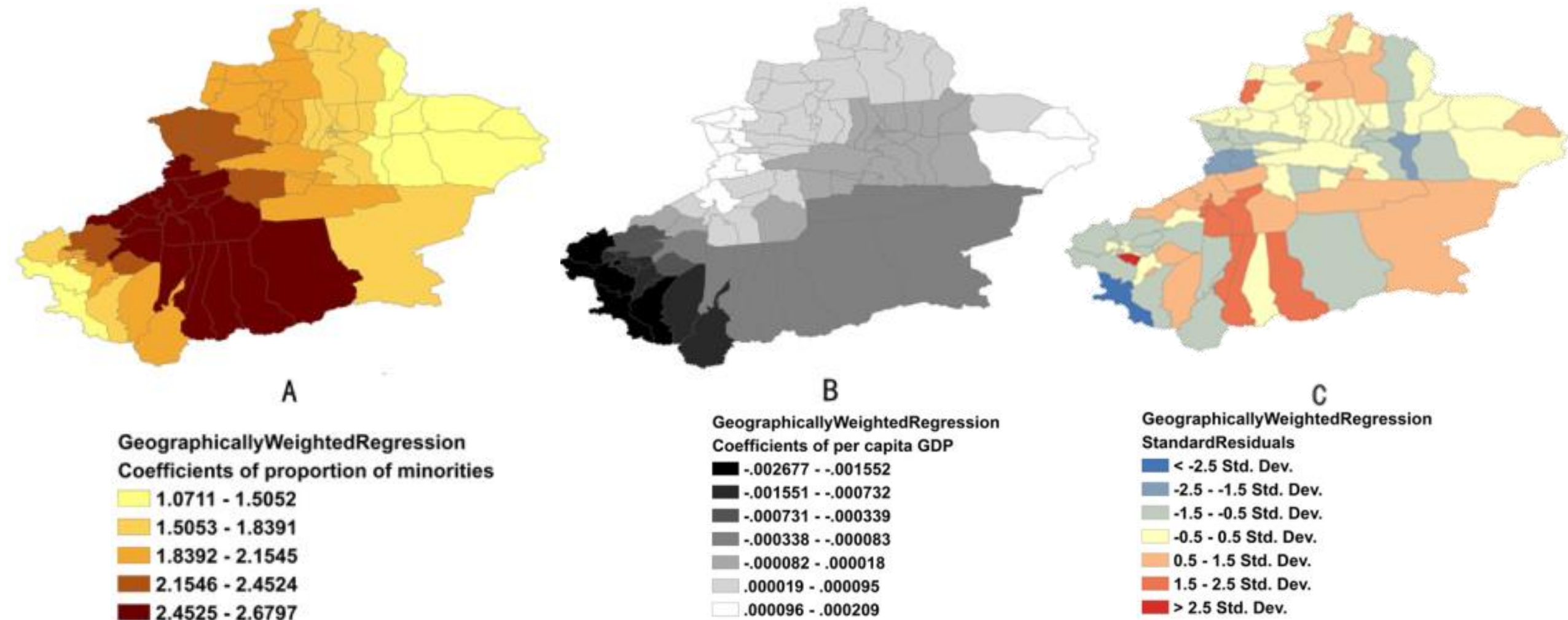
# Preditores sociodemográficos da incidência de TB pulmonar SLM

- O SLM foi utilizado como modelo de regressão espacial global **para lidar com a autocorrelação espacial** de variáveis.
- Os resultados do SLM mostraram autocorrelação espacial positiva da incidência de TB pulmonar.
  - “proporção de minorias” → efeito positivo na incidência de TB, e o valor absoluto do coeficiente de regressão diminuiu em comparação ao modelo OLS.
  - “PIB per capita” → não mostrou mais significância estatística ( $P > 0,05$ ) no modelo SLM.
- Os valores decrescentes de Log-Likelihood e AIC também sugeriram a melhoria do ajuste do modelo no SLM em comparação com o OLS.
- A estatística do teste I de Moran para resíduos do modelo SLM foi praticamente 0 → eliminada a autocorrelação espacial, como esperado.

# Preditores sociodemográficos da incidência de TB pulmonar GWR

- Foi usado como modelo de regressão espacial **local para lidar com a natureza não estacionária** do modelo OLS.
- Melhorou o ajuste do modelo OLS com um aumento do  $R^2$  e uma diminuição do valor de AIC.
- Os coeficientes mapeados para cada município/distrito indicaram onde as variáveis eram preditoras efetivas da incidência de TB pulmonar e onde não eram.

# Coeficientes dos preditores e resíduos padrão da GWR



(A) Coeficientes de proporção de minorias. (B) Coeficientes do PIB per capita. (C) Padrão residual de GWR

# Preditores sociodemográficos da incidência de TB pulmonar

- Os Modelos OLS, SLM e GWR foram utilizados para explorar os preditores de incidência de TB pulmonar.
- SLM e GWR → melhor aptidão do que o modelo OLS.

“**Proporção de minorias** → **relação positiva**

“**PIB per capita**” → **relação negativa** → forte preditor para as regiões sudoeste.

- Os **fatores genéticos** podem ser uma explicação, porque os resultados de alguns estudos mostraram a suscetibilidade genética à TB em minorias.

# Preditores sociodemográficos da incidência de TB pulmonar

**Table 3. Summary of OLS and SLM regression model.**

| Model | Parameter                | Coefficient | Std.Error | t / z | Probability | Log-Likelihood | AIC    |
|-------|--------------------------|-------------|-----------|-------|-------------|----------------|--------|
| OLS   | Constant                 | 11.2189     | 24.8464   | 0.45  | 0.65        | -434.55        | 883.89 |
|       | Proportion of minorities | 2.1573      | 0.2593    | 8.32  | <0.01       |                |        |
|       | per capita GDP           | -0.0005     | 0.0003    | -2.06 | 0.04        |                |        |
| SLM   | Wy                       | 0.3881      | 0.0946    | 4.10  | <0.01       | -427.61        | 864.48 |
|       | Constant                 | -22.5312    | 23.3347   | -0.96 | 0.33        |                |        |
|       | Proportion of minorities | 1.7004      | 0.2583    | 6.58  | <0.01       |                |        |
|       | per capita GDP           | -0.0002     | 0.0003    | -0.49 | 0.62        |                |        |

**Table 4. Summary of GWR model coefficients.**

| Parameter                | Minimum | 25% quartile | 50% quartile | 75% quartile | Maximum |
|--------------------------|---------|--------------|--------------|--------------|---------|
| Intercept                | -32.69  | -14.66       | 1.76         | 11.32        | 139.38  |
| Proportion of minorities | 1.07    | 1.76         | 2.01         | 2.39         | 2.68    |
| Per capita GDP           | -0.0027 | -0.0001      | 0.0000       | 0.0001       | 0.0002  |
| Condition number         | 6.39    | 7.25         | 7.81         | 13.49        | 29.99   |
| Local R <sup>2</sup>     | 0.2506  | 0.5271       | 0.5972       | 0.6366       | 0.6940  |

# Conclusões

- A incidência de SS+TB em Xinjiang → tendência decrescente de 2005 a 2013, mas permanece superior à média nacional da China.
- TB pulmonar, SS+TB e SS-TB → autocorrelação espacial significativa
- A análise detectou dois agrupamentos:
  - “hotspots” nas regiões sudoeste
  - “coldspots” nas regiões centrais-norte.

# Conclusões

- Na exploração de dados sociodemográficos
  - “proporção de minorias” → correlação **positiva** → forte preditor nos municípios no centro-sul.
  - “PIB per capita” → correlação **negativa** → forte preditor na região sudoeste.
- Esforços futuros → prioridades ao sul de Xinjiang.
- Pesquisas sobre a suscetibilidade genética das minorias, em particular a população de Uigur, que é a maior população do sul de Xinjiang.

# Limitações do trabalho

- Dados extraídos da vigilância oficial → possibilidade de **subnotificação de casos** em algumas regiões.
- Estudo ecológico → associação em nível de grupo → **falácia ecológica**
- **Indisponibilidade de dados** no nível de distrito para estudar outros fatores ecológicos relacionados.
- **Mais pesquisas são necessárias** para estudar os fatores socioeconômicos e ambientais.

# Referências

WUBULI, Atikaimu et al. Socio-demographic predictors and distribution of pulmonary tuberculosis (TB) in Xinjiang, China: A spatial analysis. **PloS one**, v. 10, n. 12, p. e0144010, 2015.

Brasil, Ministério da saúde.

Organização Mundial da Saúde – OMS.

A topographic map of China with the Xinjiang region highlighted in red. The map shows the terrain of China, with Xinjiang in the northwest. The word "Obrigada!" is overlaid in the center.

Obrigada!

# Socio-Demographic Predictors and Distribution of Pulmonary Tuberculosis (TB) in Xinjiang, China: A Spatial Analysis

Atikaimu Wubuli, Feng Xue, Daobin Jiang, Xuemei Yao, Halmurat Upur,  
Qimanguli Wushouer (2015)

---

**RAÍSSA CAROLINE DOS SANTOS TEIXEIRA**

PG-SER INPE 2019