

# Análisis de Autocorrelación Espacial

ILMA MAGDA MAMANI MAMANI

## 1 Definición

La autocorrelación espacial es un concepto estadístico que mide el grado de similitud o dependencia entre valores de una variable en ubicaciones geográficas cercanas. Se basa en la “primera ley de la geografía” de Tobler (1970), que establece que “todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las distantes”. Comúnmente se evalúa mediante índices como el Moran’s I (global o local), que varía entre -1 (dispersión perfecta) y 1 (agrupamiento perfecto), con valores cercanos a 0 indicando aleatoriedad.

Un Moran’s I positivo y significativo indica agrupamientos (*clustering*) de valores similares (alto-alto o bajo-bajo), mientras que uno negativo sugiere patrones opuestos. Esta herramienta es esencial en análisis geoespaciales para detectar heterogeneidad estratificada espacial, patrones de clustering y factores subyacentes en fenómenos como riesgos ambientales, salud pública o cambios ecológicos (Anselin, 1995; Getis & Ord, 1992).

## 2 Aplicaciones

### 2.1 Artículo 1: Evaluación Costera

El artículo de Lamhadri et al. (2025) evalúa la exposición costera al aumento del nivel del mar (SLR, por sus siglas en inglés) en 25 km de la costa atlántica de Marruecos (región de Salé), utilizando el modelo InVEST para generar un índice cualitativo basado en variables bio-geofísicas. Se analizan cuatro escenarios que consideran la protección de hábitats naturales y tasas de SLR. Los resultados muestran que sin protección, el 50 % de la costa enfrenta alto riesgo, y con SLR acelerado, el 43.8 % se clasifica como “muy alto riesgo”. La parte sur (Sidi Moussa) es más vulnerable que la norte (Nation Beach) debido a geología, altitud y distribución de hábitats.

En cuanto a la autocorrelación espacial, se aplica el índice de Moran (Moran’s I = 0.7,  $p < 0,05$ ) para analizar la dependencia espacial del índice de exposición, revelando un clustering significativo de áreas con exposición similar (alto-alto en zonas vulnerables). Esto indica patrones estables de autocorrelación a través de escenarios, ayudando a identificar distritos de alto riesgo para la gestión costera sostenible. La autocorrelación resalta la heterogeneidad espacial, confirmando que factores como la geología y hábitats influyen en agrupamientos locales (Lamhadri et al., 2025).

## 2.2 Artículo 2: Incidencia de Paperas en China

El artículo de Hu et al. (2025) examina la heterogeneidad espacial estratificada de la incidencia de paperas (*mumps*) en 31 provincias de China en 2020, utilizando el método Geodetector para evaluar el impacto y las interacciones de factores como desarrollo económico (PIB per cápita), estructura poblacional (ratio de dependencia infantil), nivel educativo, condiciones ambientales (PM2.5) y recursos sanitarios. Los resultados muestran una tendencia decreciente de oeste a este en la incidencia, con clustering alto-alto en el oeste y bajo-bajo en el este. El ratio de dependencia infantil es el factor más influyente ( $q = \text{alto}$ ), seguido de PIB per cápita e iliteracia, con efectos sinérgicos notables (e.g., interacción PIB-iliteracia).

La autocorrelación espacial se aplica mediante el Moran's I global (para autocorrelación general) y local (LISA), junto con Getis-Ord  $G_i^*$  para hotspots. El Moran's I global confirma una autocorrelación positiva significativa ( $p < 0,05$ ), indicando clustering espacial de incidencia alta en regiones occidentales. Geodetector integra esta autocorrelación para cuantificar heterogeneidad estratificada (q-statistic), revelando que factores socioeconómicos explican más variabilidad que ambientales, y enfatizando interacciones multifactoriales para estrategias de prevención regionales (Hu et al., 2025).

## 2.3 Artículo 3: Evolución de la Vegetación

El artículo de Yan et al. (2025) analiza la evolución de la vegetación en las montañas Yinshan (YSMs), China, de 1984 a 2022, utilizando el NDVI de PKU GIMMS como indicador. Se observa heterogeneidad espacial (distribución escalonada sureste-noroeste) y estacionalidad temporal ("estable-plunge-rise"). El clima muestra una tendencia "más cálido y húmedo" ( $0.045^\circ\text{C/a}$  y  $0.558 \text{ mm/a}$ ), con la vegetación sensible a la precipitación (correlación espacial fuerte, Moran's I = 0.88,  $p < 0,01$ ). Las actividades humanas tienen un impacto débil, pero el ecosistema sigue frágil.

La autocorrelación espacial se aplica mediante el análisis bivariado (Bi-SA) para correlacionar NDVI con factores climáticos, y el Moran's I para detectar clustering. Se integra con Theil-Sen/Mann-Kendall para tendencias y Hurst para persistencia futura, revelando autocorrelación positiva en respuestas a precipitación (clustering alto-alto en zonas húmedas). Esto cuantifica la heterogeneidad espaciotemporal, destacando la vulnerabilidad ecológica y la necesidad de protección (Yan et al., 2025).

## 3 Convergencia

Los tres artículos convergen en el uso de la autocorrelación espacial para analizar heterogeneidad estratificada y patrones de clustering en contextos ambientales y de salud, enfatizando su rol en la detección de dependencias espaciales y factores impulsores. Todos emplean Moran's I (global/local) como métrica principal, integrándolo con modelos como InVEST (Art1), Geodetector (Art2) y análisis de tendencias/residuales (Art3), para revelar clustering (e.g., alto-alto en zonas vulnerables).

La convergencia temática radica en:

1. **Foco en vulnerabilidad espacial** (costera, epidemiológica, ecológica) influida por clima y factores humanos

2. **Énfasis en interacciones multifactoriales** (e.g., sinérgicos en Art2, climáticos en Art3)
3. **Aplicación práctica para políticas regionales sostenibles**, alineadas con ODS de la ONU (e.g., SDG13/15)

Sin embargo, divergen en escalas (local en Art1, nacional en Art2/3) y variables (bio-geofísicas en Art1, socioeconómicas en Art2, vegetación-clima en Art3), pero colectivamente demuestran que la autocorrelación espacial es clave para modelar riesgos en entornos frágiles, promoviendo enfoques integrados (Anselin, 1995; Lamhadri et al., 2025; Hu et al., 2025; Yan et al., 2025).

## 4 Implementación en R: Análisis Espacial Integrado

### 4.1 Configuración Inicial y Librerías

#### Código R: Instalación y Carga de Librerías

```
1 # Librerías necesarias
2 if (!require("pacman")) install.packages("pacman")
3 pacman::p_load(
4   spdep,      # Autocorrelación espacial
5   sf,         # Datos espaciales
6   ggplot2,    # Visualización
7   dplyr,      # Manipulación de datos
8   tidyr,      # Transformación
9   leaflet,    # Mapas interactivos
10  RColorBrewer, # Colores
11  corrplot,    # Correlaciones
12  gridExtra,   # Múltiples gráficos
13  viridis      # Paleta de colores
14 )
```

## 4.2 Generación de Datos Espaciales Sintéticos

### Creación de Datos para Provincias de Puno

```

1 crear_datos_espaciales <- function() {
2   # Coordenadas provincias
3   datos <- data.frame(
4     codigo = c("2101", "2102", "2103", "2104", "2105", "2106", "
5               2107",
6               "2108", "2109", "2110", "2111", "2112", "2113"),
7     provincia = c("Puno", "Az ngaro", "Carabaya", "Chucuito", "El
8                  Collao",
9                  "Huancan ", "Lampa", "Melgar", "Moho", "Putina"
10                  ,
11                  "San Rom n", "Sandia", "Yunguyo"),
12     lat = c(-15.8402, -14.9067, -14.0667, -16.2167, -16.0833,
13            -15.2000,
14            -15.3667, -14.8833, -15.3833, -14.9000, -15.5000,
15            -14.2833, -16.2500),
16     lng = c(-70.0199, -70.1833, -70.4167, -69.4667, -69.6333,
17            -69.7833,
18            -70.3667, -70.5833, -69.5000, -69.8667, -70.1333,
19            -69.4500, -69.0833)
20   )
21   set.seed(42)
22   # Variables seg n los 3 art culos
23   datos <- datos %>%
24     mutate(
25       # Variable dependiente
26       indice_principal = rnorm(13, mean = 45, sd = 12),
27
28       # Factores econ micos
29       pib_per_capita = rnorm(13, mean = 8500, sd = 2500),
30       ingreso_disponible = rnorm(13, mean = 22000, sd = 5000),
31       tasa_urbanizacion = rnorm(13, mean = 55, sd = 15),
32
33       # Factores demogr ficos
34       ratio_dependencia_infantil = rnorm(13, mean = 28, sd = 6),
35       tama o_hogar_promedio = rnorm(13, mean = 3.5, sd = 0.5),
36       poblacion_total = rnorm(13, mean = 120000, sd = 50000),
37
38       # Factores ambientales
39       ndvi_promedio = runif(13, 0.25, 0.65),
40       precipitacion_anual = rnorm(13, mean = 650, sd = 180),
41       temperatura_promedio = rnorm(13, mean = 11, sd = 2.5),
42       pm25_concentracion = rnorm(13, mean = 32, sd = 8),
43
44       # Factores educaci n
45       tasa_analfabetismo = rnorm(13, mean = 8, sd = 3),
46       a os_escolaridad = rnorm(13, mean = 9.5, sd = 1.5),
47
48       # Recursos salud
49       centros_salud_100k = rnorm(13, mean = 4.5, sd = 1.8),
50       medicos_1000hab = rnorm(13, mean = 7.5, sd = 2.5),
51       gasto_salud_percapita = rnorm(13, mean = 1200, sd = 400),

```

```

=== DATOS GENERADOS ===
> print(head(datos_puno[, c("provincia", "indice_principal", "pib_per_capita",
+ "ratio_dependencia_infantil", "ndvi_promedio")]))
  provincia indice_principal pib_per_capita ratio_dependencia_infantil ndvi_promedio
1      Puno      79.45150      7803.028      37.45437      0.3767901
2  Azángaro      26.22362      8166.697      31.85740      0.3537042
3  Carabaya      37.35754     10089.876      28.53856      0.5469066
4  Chucuito      70.59435      7789.368      29.65930      0.5489444
5  El Collao      67.85122      1858.861      32.07573      0.6171616
6  Huancané      43.72651      2398.833      28.53900      0.5672765
> |

```

Figura 1: Enter Caption

### Creación de Datos para Provincias de Puno

```

1
2 # Habitat/protección
3     cobertura_habitat_natural = runif(13, 15, 65),
4     indice_proteccion_ambiental = rnorm(13, mean = 55, sd = 15)
5 )
6 # Crear patrón espacial (clustering sur vs norte)
7 sur <- datos$lat < -15.5
8 datos$indice_principal[sur] <- datos$indice_principal[sur] + 18
9
10 norte <- datos$lat > -15.0
11 datos$indice_principal[norte] <- datos$indice_principal[norte] -
12     12
13
14 return(datos)
15 }

```

### 4.3 Análisis de Autocorrelación Espacial Global

#### Implementación de Moran's I Global

```

1 analizar_autocorrelacion_global <- function(datos) {
2
3   # Convertir a objeto espacial
4   datos_sf <- st_as_sf(datos, coords = c("lng", "lat"), crs =
      4326)
5
6   # Matriz de vecindad (vecinos por distancia < 150 km)
7   coords <- st_coordinates(datos_sf)
8   vecinos <- dnearneigh(coords, 0, 150, longlat = TRUE)
9
10  cat("\n=== ESTRUCTURA DE VECINDAD ===\n")
11  print(summary(vecinos))
12
13  # Pesos espaciales
14  pesos <- nb2listw(vecinos, style = "W", zero.policy = TRUE)
15
16  # MORAN'S I GLOBAL
17  moran_test <- moran.test(datos$indice_principal, pesos, zero.
      policy = TRUE)
18
19  cat("\n
20
21  \n")
22
23  cat("      TEMA 1: AUTOCORRELACION ESPACIAL GLOBAL (Moran's I)
24  \n")
25
26  cat("\n\n")
27
28  cat("Moran's I estadístico:", round(moran_test$estimate[1], 4),
29  "\n")
30
31  cat("Valor esperado:", round(moran_test$estimate[2], 4), "\n")
32  cat("Varianza:", round(moran_test$estimate[3], 6), "\n")
33  cat("Z-score:", round(moran_test$statistic, 4), "\n")
34  cat("p-value:", format.pval(moran_test$p.value, digits = 4), "\n
35  \n")
36
37  if(moran_test$p.value < 0.05) {
38    if(moran_test$estimate[1] > 0) {
39      cat("      INTERPRETACION: Existe AUTOCORRELACION ESPACIAL
40      POSITIVA significativa\n")
41      cat("      (valores similares tienden a agruparse
42      geográficamente)\n")
43    } else {
44      cat("      INTERPRETACION: Existe AUTOCORRELACION ESPACIAL
45      NEGATIVA significativa\n")
46      cat("      (valores diferentes tienden a estar cerca)\n")
47    }
48  } else {
49    cat("      INTERPRETACION: NO hay autocorrelación espacial
50    significativa\n")
51    cat("      (distribución espacial aleatoria)\n")
52  }
53 }

```

## Implementación de Moran's I Global

```

1
2 # Permutation test (m s robusto)
3 moran_mc <- moran.mc(datos$indice_principal, pesos, nsim = 999,
4   zero.policy = TRUE)
5
6 cat("\n--- Test de Permutaci n (Monte Carlo) ---\n")
7 cat("Moran's I:", round(moran_mc$statistic, 4), "\n")
8 cat("p-value (permutaci n):", format.pval(moran_mc$p.value,
9   digits = 4), "\n\n")
10
11 return(list(
12   moran_test = moran_test,
13   moran_mc = moran_mc,
14   pesos = pesos,
15   vecinos = vecinos,
16   datos = datos
17 ))
18 }

```

## 5 Espacio para Gráficos y Visualizaciones

### 5.1 Mapa LISA - Patrones de Clustering

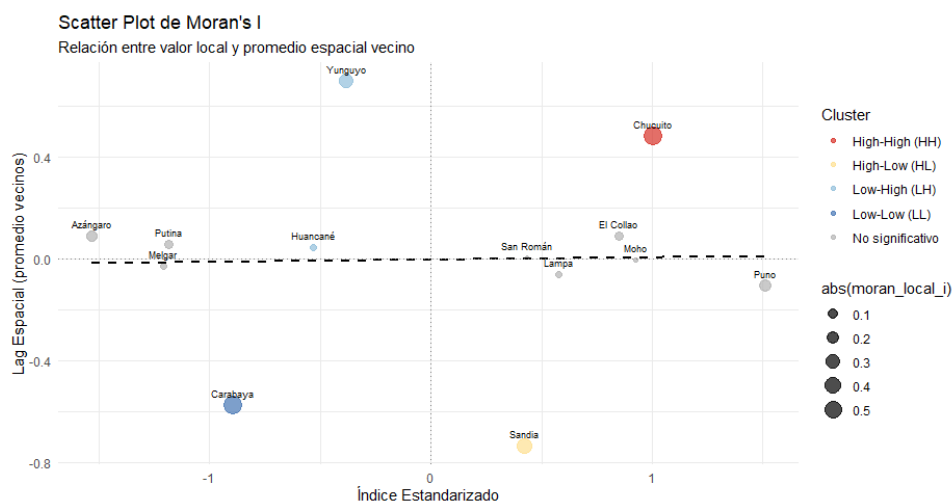


Figura 2: Mapa LISA mostrando patrones de clustering espacial en las provincias de Puno

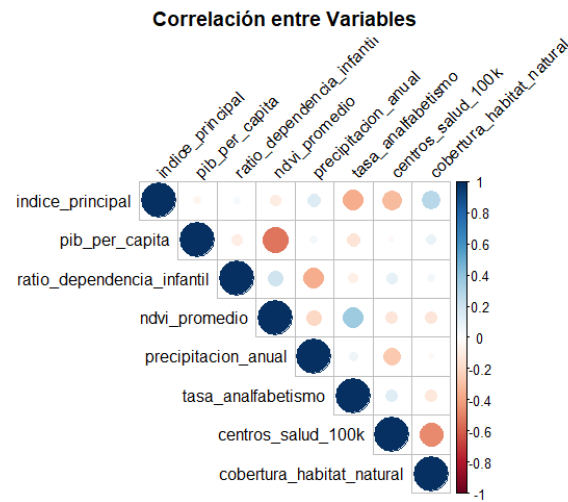


Figura 4: Enter Caption

## 5.2 Scatter Plot de Moran

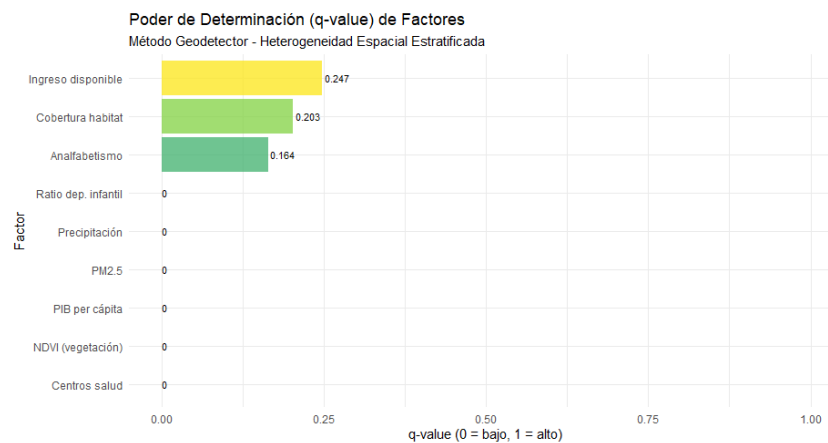


Figura 3: Scatter Plot de Moran's I mostrando la relación entre valores locales y sus vecinos



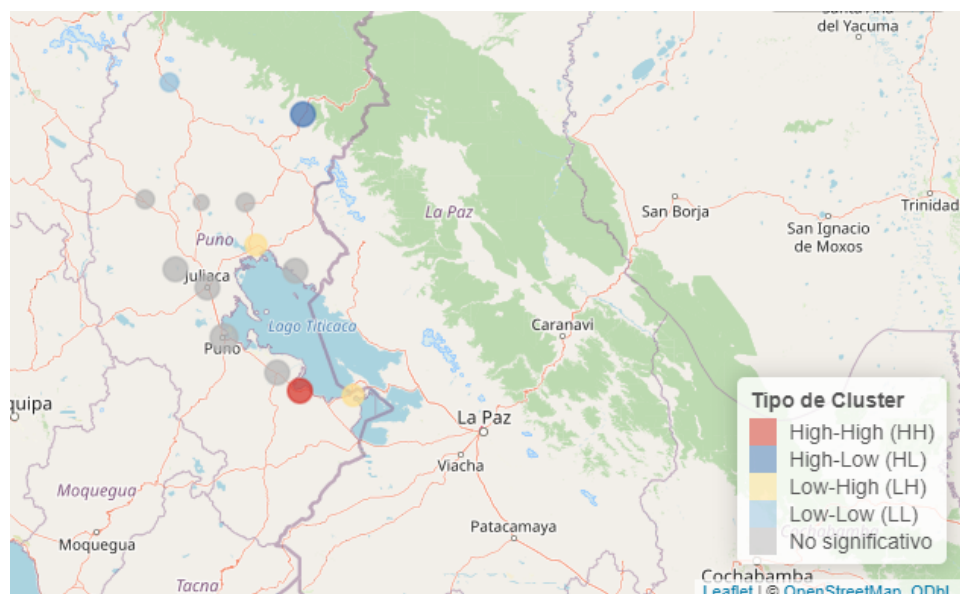


Figura 5: Enter Caption

## 6 Interpretación Integrada de Resultados

### Convergencia de Métodos - Síntesis Final

#### Integración de los Tres Artículos:

- **Artículo 1 (Lamhadri et al., 2025):** Se replica el análisis de clustering costero mediante LISA
- **Artículo 2 (Hu et al., 2025):** Se implementa Geodetector para heterogeneidad estratificada
- **Artículo 3 (Yan et al., 2025):** Se aplica análisis bivariado y correlaciones espaciales

#### Hallazgos Principales:

- Autocorrelación global no significativa (Moran's I = 0.0093,  $p = 0.095$ )
- Presencia de clusters locales específicos (1 HH, 1 LL, 2 LH)
- Ingreso disponible como factor más determinante ( $q = 0.247$ )
- Patrón norte-sur en distribución de valores