



Evaluación de modelos de interpolación para cálculo de aislamiento en región de O'Higgins

Matías Poch Clavero.
Ingeniero Civil en Geografía
icg.poch@yahoo.cl



Introducción

- La Subsecretaría de Desarrollo Regional desarrollo un índice de aislamiento para todas las localidades de Chile.
- Este índice permite identificar cuáles son las localidades que se encuentran en condiciones de aislamiento.
- Cada localidad tiene su índice representado por un punto en el espacio.
- Se plantea generar un raster (continua), que pueda representar de mejor manera este fenómeno medido de manera puntual (discreta)



Introducción

- Con esta representación continua, se pretende clasificar píxeles y determinar zonas de condiciones similares o “territorios aislados”
- La región de O'Higgins cuenta con 2850 localidades, de estas se elegirán 339 como puntos de control, quedando 2511 a interporlar.
- Se probarán 5 métodos de interpolación y rasterización que se analizarán con cuatro medidas de ajuste estadístico.



Objetivos

- General.
 - Determinar que modelo de interpolación es el adecuado para representar el fenómeno de aislamiento para la región de O'Higgins.
- Específicos.
 - Revisión de los métodos provistos por gvSIG 1.11
 - Generación de 5 modelos raster, a partir de métodos de interpolación.
 - Calcular medidas de ajuste, mediante puntos de control.



Metodología

- Métodos de interpolación a aplicar, provistos por gvSIG 1.11
 - Decrecimiento lineal.
 - Densidad.
 - Densidad (kernel)
 - Distancia inversa.
 - Vecindad.



Metodología

- Salidas ráster.
 - Las localidades se encuentran en Datum WGS 84 H19 Sur, por lo que sus salidas serán en ese sistema de referencia.
 - El tamaño de celda es de 500 mts.
 - Los rasters generados serán recortados por el polígono de la región de O'Higgins.
 - Se ajustan los valores de la interpolación al rango del índice $[-1, 2]$, donde los valores menores a 0, según el equipo investigador, son localidades en condición de aislamiento.

-



Metodología

- Puntos de control.
 - Son utilizados para validar los resultados de la interpolación.
 - El tamaño de la muestra de puntos de control depende de los siguiente:
 - Nivel de confianza o grado de error para presentar resultados. El nivel de confianza habitual es de 0,05 ($\alpha=0,05$). El nivel de confianza esta determinado a una probabilidad de ocurrencia en una distribución normal,
 - Varianza estimada de los datos.
 - Margen de error que se está dispuesto a aceptar.



Metodología

- La cantidad de puntos de control está dada por:

$$n = \frac{N}{1 + \frac{(e^2 (N - 1))}{(z^{2pq})}}$$

- Donde:
 - N: Universo de datos.
 - e: error muestral.
 - z: Nivel de confianza ($\alpha=0,05$) corresponde a $z = 1,96$ ó 2 errores típicos.
 - pq: Varianza de datos. Si se supe $p=q$, quiere decir que en los datos existe la máxima diversidad posible y no se corre riesgo en quedar cortos con el número de datos a muestrear.



Metodología

- Medidas de ajuste.
 - Las medidas de ajuste nos permiten determinar que método es el mas idoneo y se basan en la comparación de Iso datos observados y los modelos simulados.
 - Para la formulación se ocupa la siguiente notación.
 - $[X_{ij}^o]$ = Valor estimado en coordenadas i,j .
 - $[X_{ij}]$ = Valor observado en coordenadas i,j .
 - X = Valor medio observado.
 - n = Es el número de puntos de control.



Metodología

- Coeficiente de determinación R^2 .

$$R^2 = 1 - \frac{\sum_{ij} (X_{ij} - X_{ij}^o)^2}{\sum_{ij} (X_{ij} - \bar{X})^2}$$

- Desviación estándar de residuos.

$$A = \sqrt{\frac{\sum_{ij} (X_{ij} - X_{ij}^o)^2}{n^2 - 1}}$$

Metodología

- Error absoluto medio.

$$MABSERR = \frac{\sum_{ij} (|X_{ij} - X_{ij}^{\circ}|)}{n^2}$$

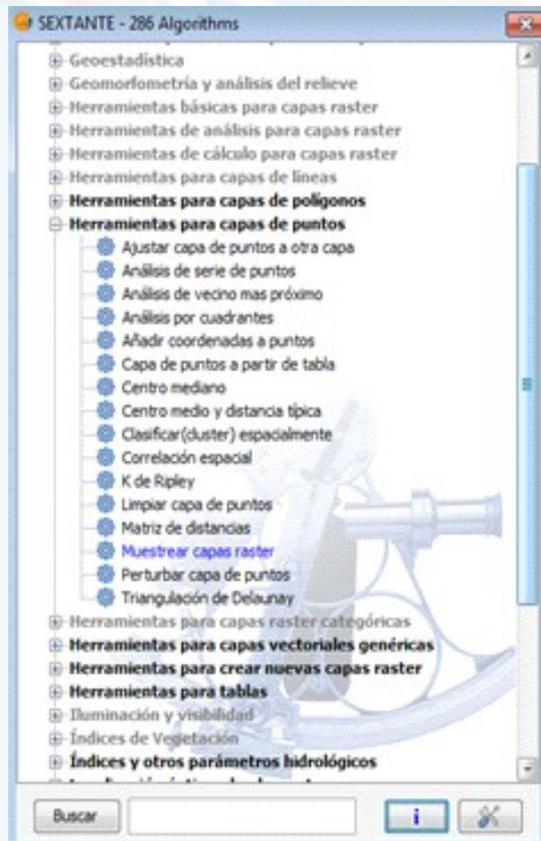
- Valor absoluto medio .

$$MABSERR^{\circ} = \frac{1}{n^2} \sum_{ij} \left(\left| \frac{X_{ij} - X_{ij}^{\circ}}{\bar{X}} \right| \right)$$



Metodología

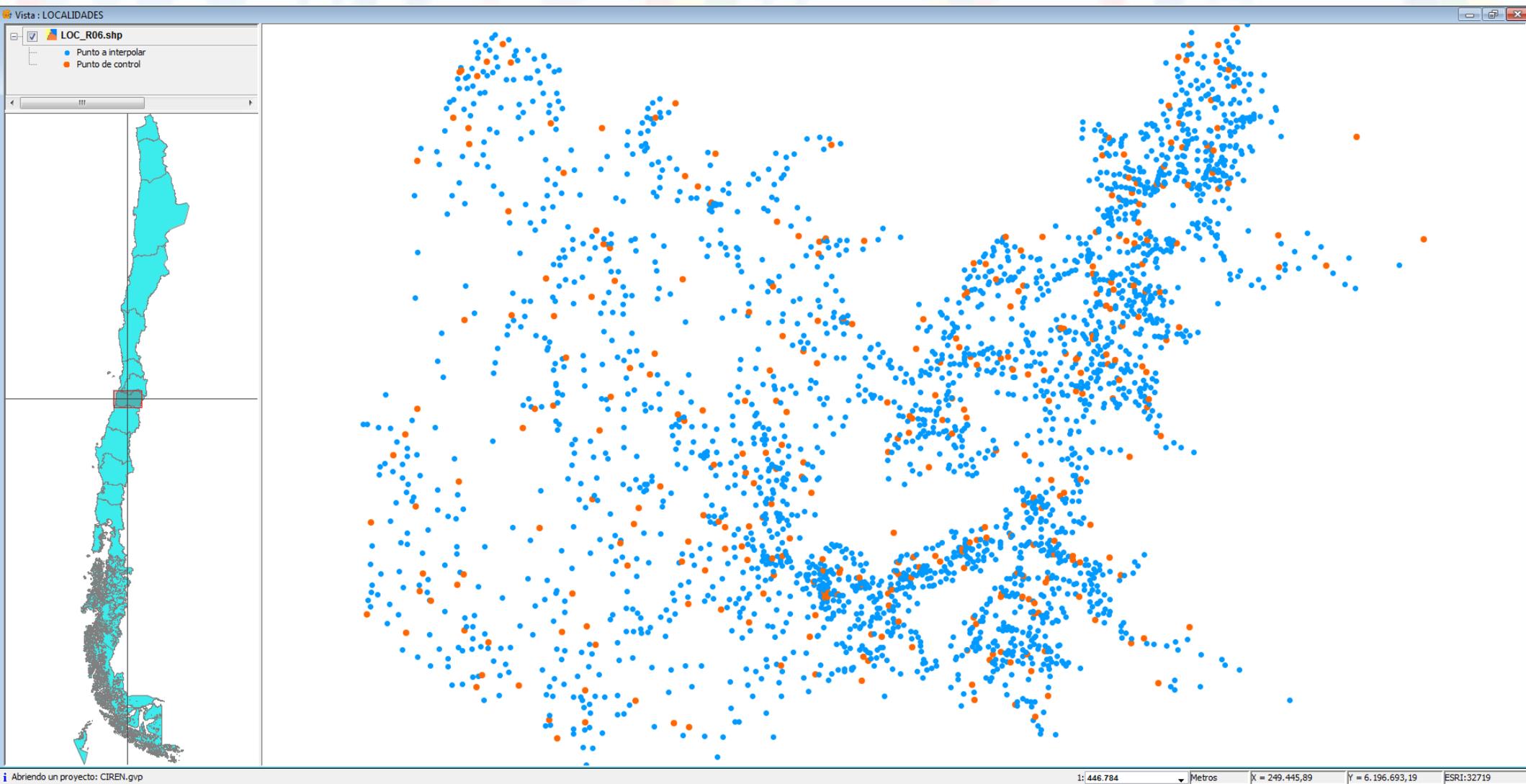
- Muestreo de valores.
 - Se ocupará algoritmo que provee el gestor de geoprosesos SEXTANTE “muestrear capas raster”





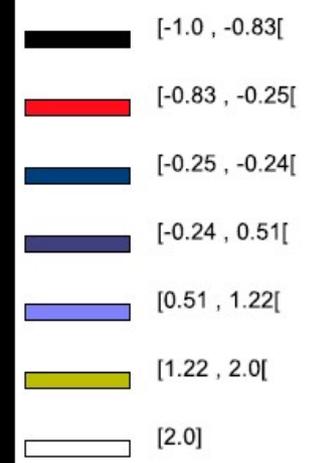
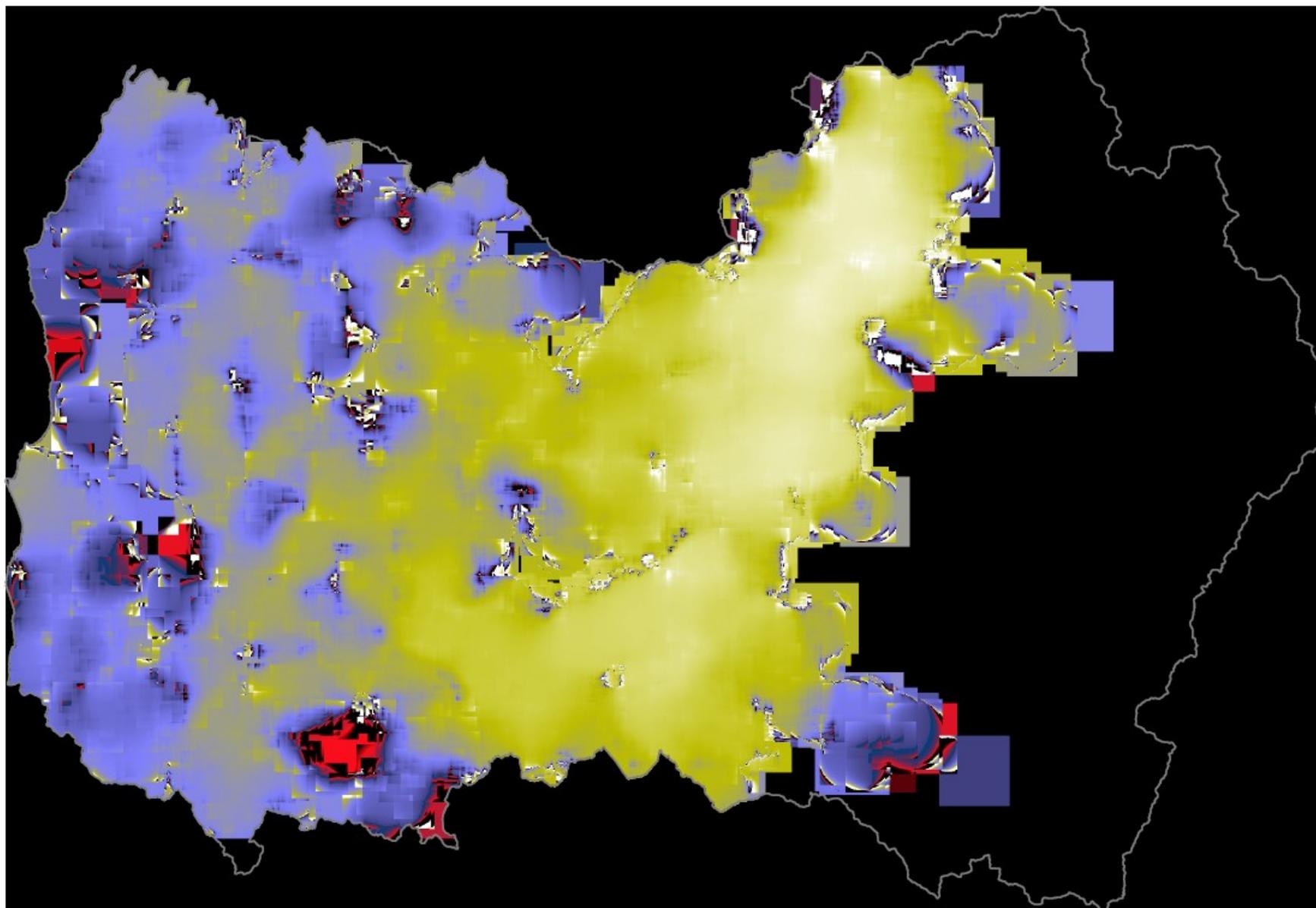
Desarrollo

- 2511 localidades a interpolar y 339 localidades como puntos de control.



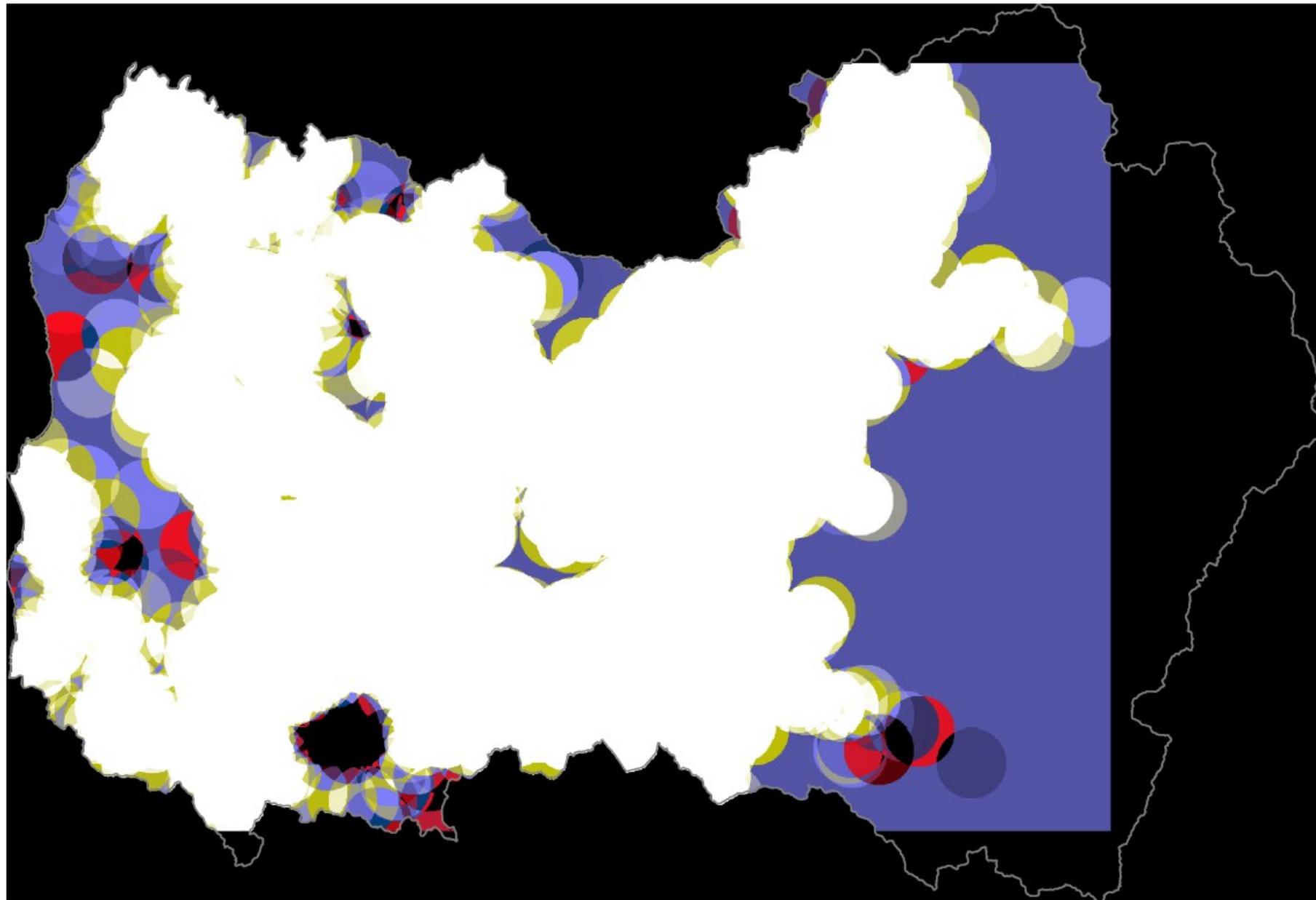


MÉTODO DECRECIMIENTO LINEAL



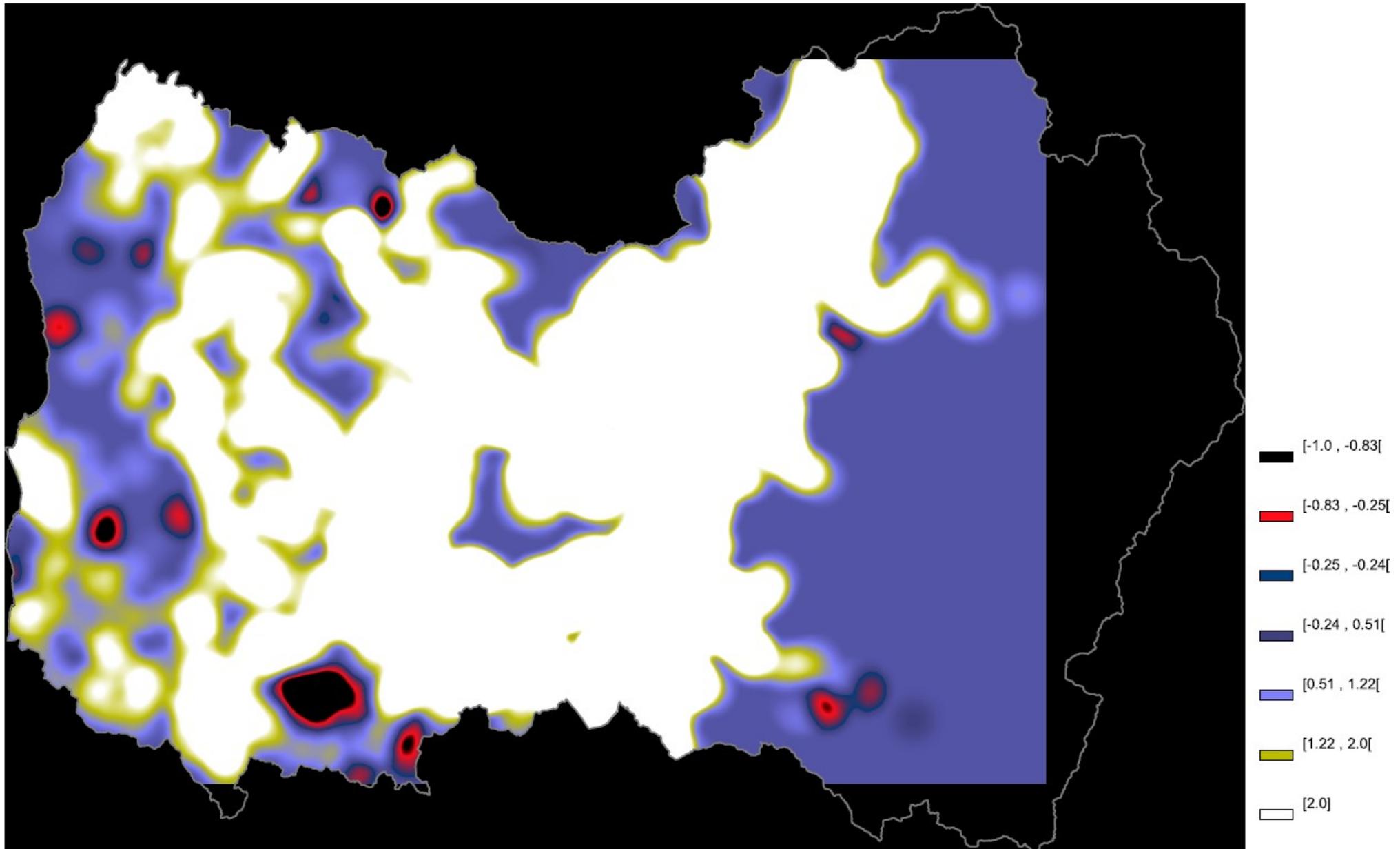


MÉTODO DENSIDAD



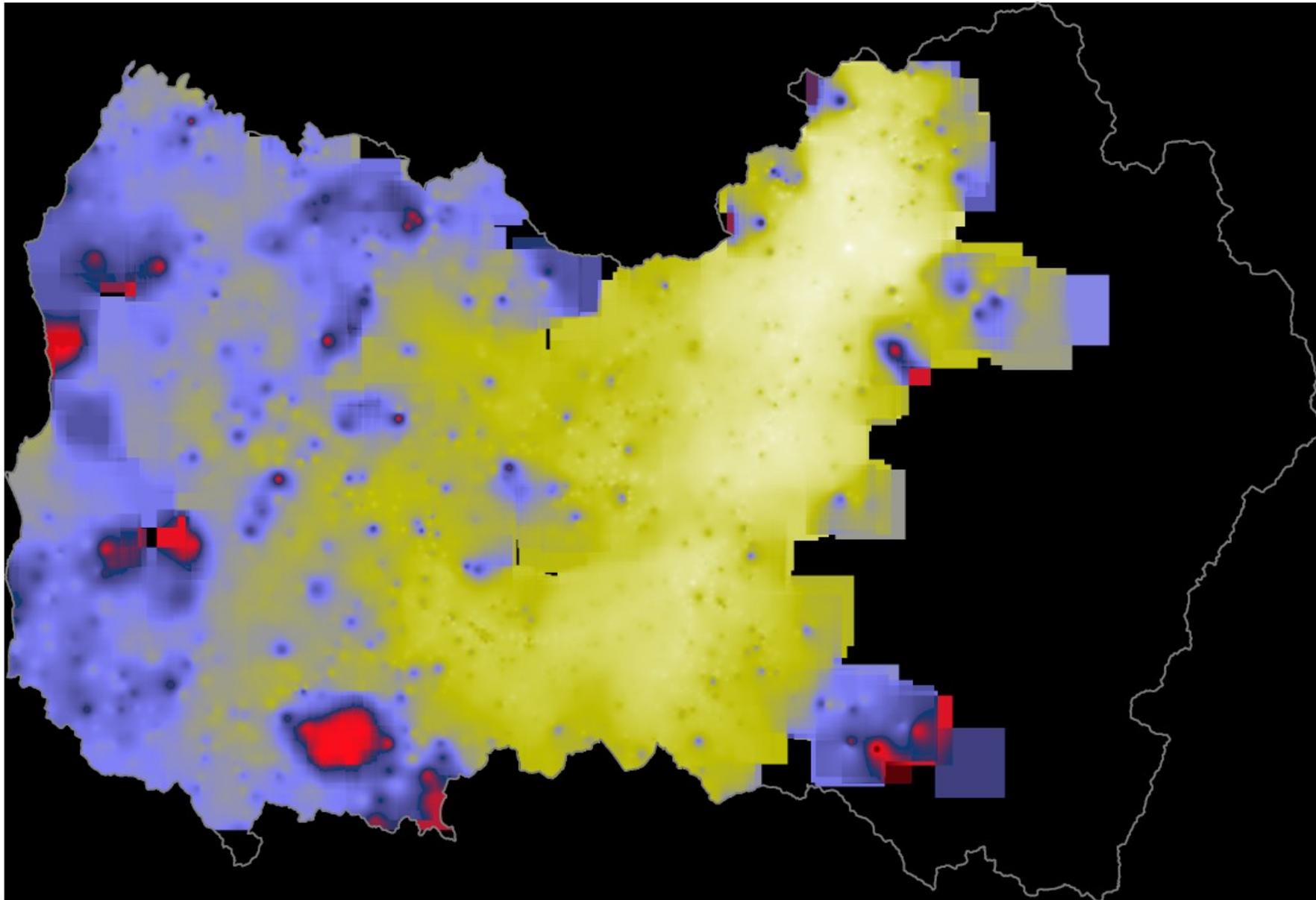


MÉTODO DENSIDAD (KERNEL)



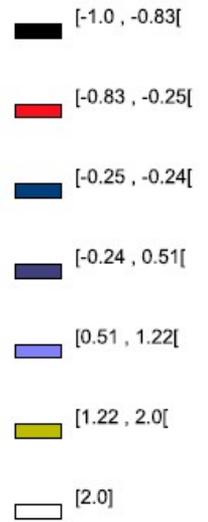
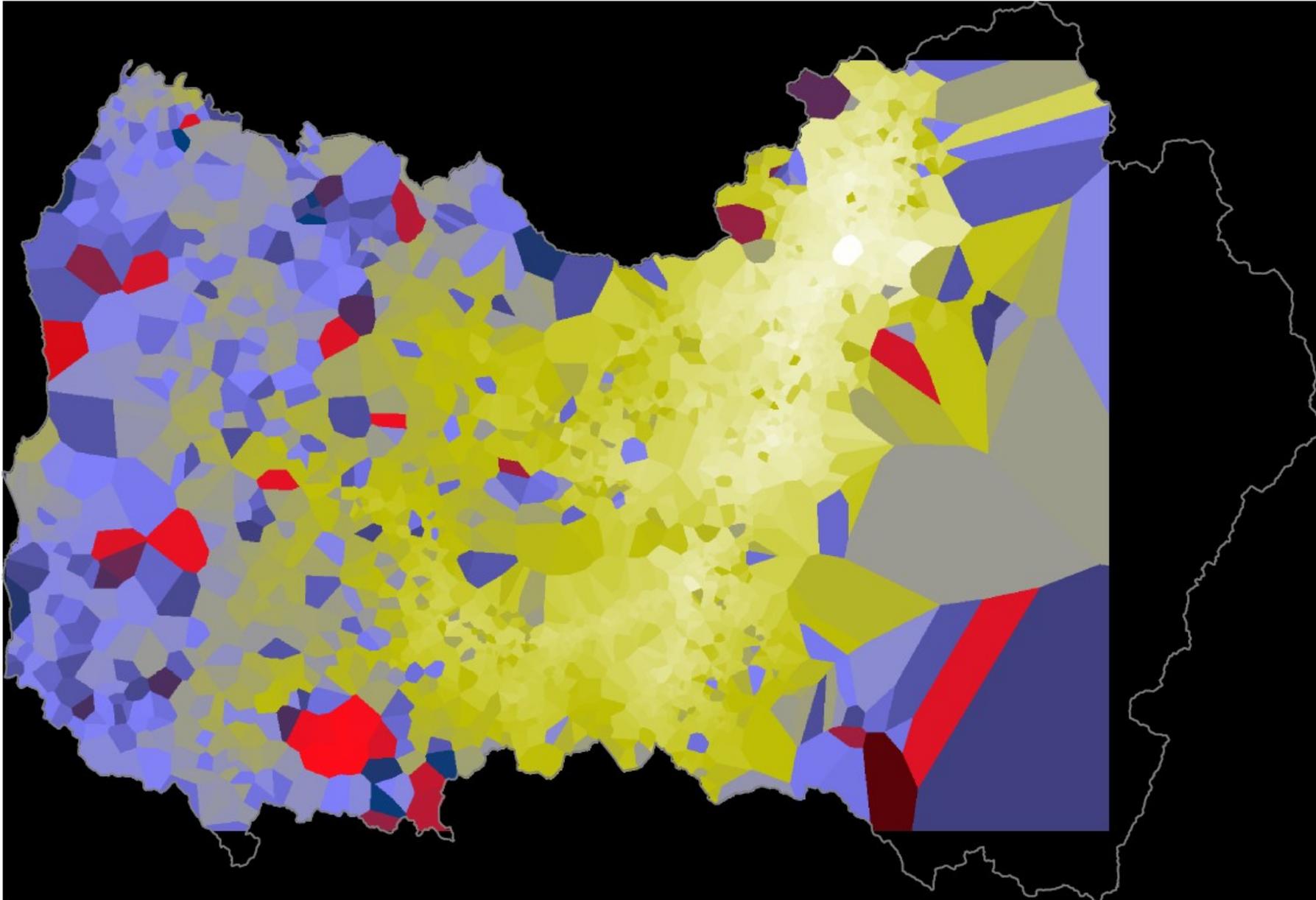


MÉTODO DISTANCIA INVERSA





MÉTODO VECINDAD





Desarrollo

- Medidas de Ajuste.
 - A los 339 puntos de control (aleatorios) se aplica el remuestreo de capas ráster, con cada uno de los métodos de interpolación.

	R^2	A	$MABSERR$	$MABSERR^*$
Decrecimiento Lineal	0,762671291	0,002249778	0,000418129	0,002923542
Densidad	-13483,57701	39,77473651	0,130127042	7,651236206
Densidad Kernel	-1858,412587	5,482066881	0,046504533	1,052710418
Distancia Inversa	0,817117003	0,002410385	0,000349931	0,002923905
Vecindad	0,726043418	0,00214173	0,00038594	0

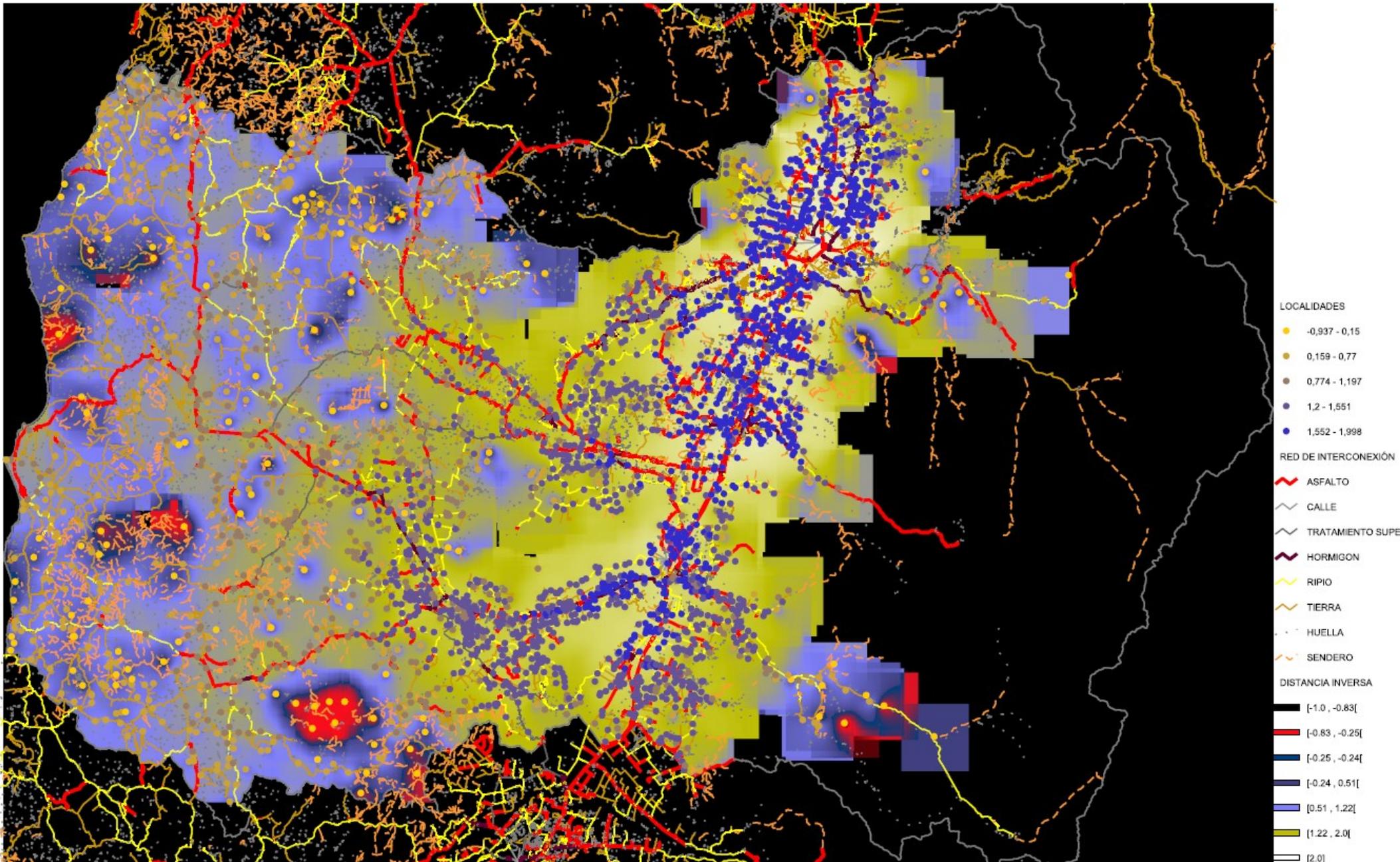


Desarrollo

- De los modelos revisados el que presenta mejores resultados en todas las medidas de ajustes es el método de la distancia inversa.
- En efecto, posee los mejores resultados en R^2 y MABSERR y para las medidas de ajuste de A y MABSRR* se observan valores próximos a los valores mas bajos obtenidos por otros métodos de interpolación.



MÉTODO DISTANCIA INVERSA





Conclusiones

- Con respecto a gvSIG.
 - Tiene un buen comportamiento en el proceso de interpolación.
 - La visualización con paleta de colores presenta algunos imperfectos.
 - Cabe destacar el carácter gratuito de este software, lo cual lo hace atractivo para realizar docencia.



Conclusiones

- Modelos ráster generados a partir de métodos de interpolación.
 - A partir de las medidas de bondad de ajuste el método de interpolación más adecuado sería el de la distancia inversa,
 - Se debe tener en cuenta que estas medidas de ajuste son válidas para la distribución de localidades de la región de O'Higgins y para los parámetros especificados.



Conclusiones

- Modelos ráster generados a partir de métodos de interpolación.
 - Es importante conocer la naturaleza del fenómeno que se está estudiando. En este caso, el índice de aislamiento está calculado en función del tiempo (medido en horas) desde cada una de las localidades a una serie de servicios localizados en el territorio.
 - Se realiza a través del cálculo de ruta mínima de la red de transporte, para esta región caminos de distintas categorías (hormigón, asfalto, tierra, ripio, huellas y senderos) .



Conclusiones

- Modelos ráster generados a partir de métodos de interpolación.
 - Esto es importante, ya que los métodos de interpolación, si bien consideran distancia mediante distintas funciones, esta es, la cartesiana o euclidiana.
 - Se debe tener en cuenta que estas medidas de ajuste son válidas para la distribución de localidades de la región de O'Higgins y para los parámetros especificados.



Consideraciones finales

- Las medidas de ajustes expuestas son de gran utilidad ya que permiten muestrear estadísticamente y validar, no solo métodos de interpolación, si no que cualquier salida raster.
- Los análisis visuales, se corre el riesgo de que la percepción.
- Es importante conocer el fenómeno que se está estudiando, ya que tanto el análisis estadístico y visual a priori, puede inducir a error.