

Identificando patrones sinópticos de circulación mediante Mapas Auto-Organizantes

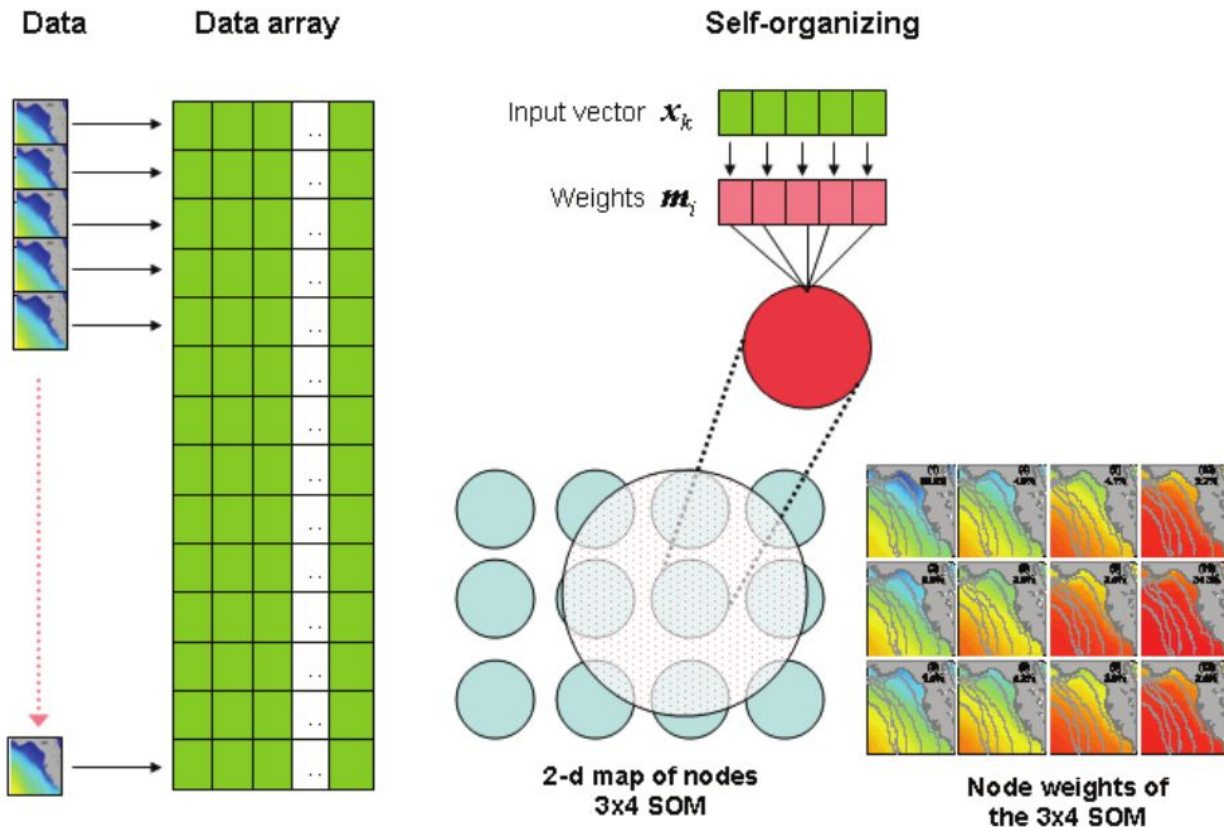
Julián Cañellas
CSC-CONICET - DCAO UBA FCEyN

Motivación:

- La generación de energía fotovoltaica requiere pronósticos de radiación de alta resolución para ser competitiva en el mercado eléctrico.
- Los modelos numéricos no resuelven adecuadamente la escala deseada.
- Es posible hallar patrones en la variabilidad de la radiación solar relacionados con las condiciones de circulación mediante técnicas de machine learning.

Essentials of the self-organizing map

Teuvo Kohonen





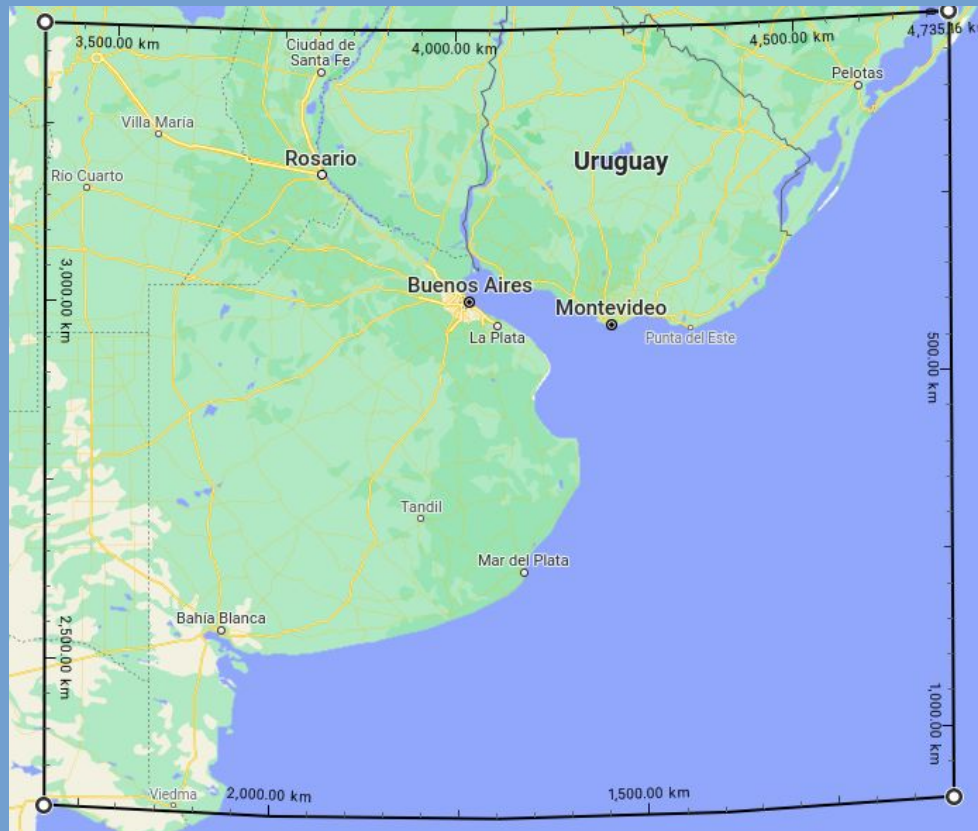
$$1. W_v(s + 1) = W_v(s) + \theta(u, v, s) \cdot \alpha(s) \cdot (D(t) - W_v(s))$$

- s is the current iteration
- λ is the iteration limit
- t is the index of the target input data vector in the input data set $\mathbf{D}$
- $D(t)$ is a target input data vector
- v is the index of the node in the map
- $\mathbf{W}_v$ is the current weight vector of node v
- u is the index of the best matching unit (BMU) in the map
- $\theta(u, v, s)$ is a restraint due to distance from BMU, usually called the neighbourhood function, and
- $\alpha(s)$ is a learning restraint due to iteration progress.

Dominio

El dominio elegido tiene 10x15 grados de latitud-longitud, que teniendo en cuenta la resolución equivale a una grilla de 41x61 puntos (**2501** en total)

Se tomaron 5 años de datos de Presión trihorarios, lo que equivale a un total de **14608** observaciones



Métricas de error empleadas

Quantization Error = Distancia promedio las observaciones y su nodo más cercano (Best Matching Unit, BMU).

Topographic Error = Es una medida de la continuidad de la grilla de nodos. Es cero si todas los pares de observaciones contiguas proyectan a dos nodos contiguos.

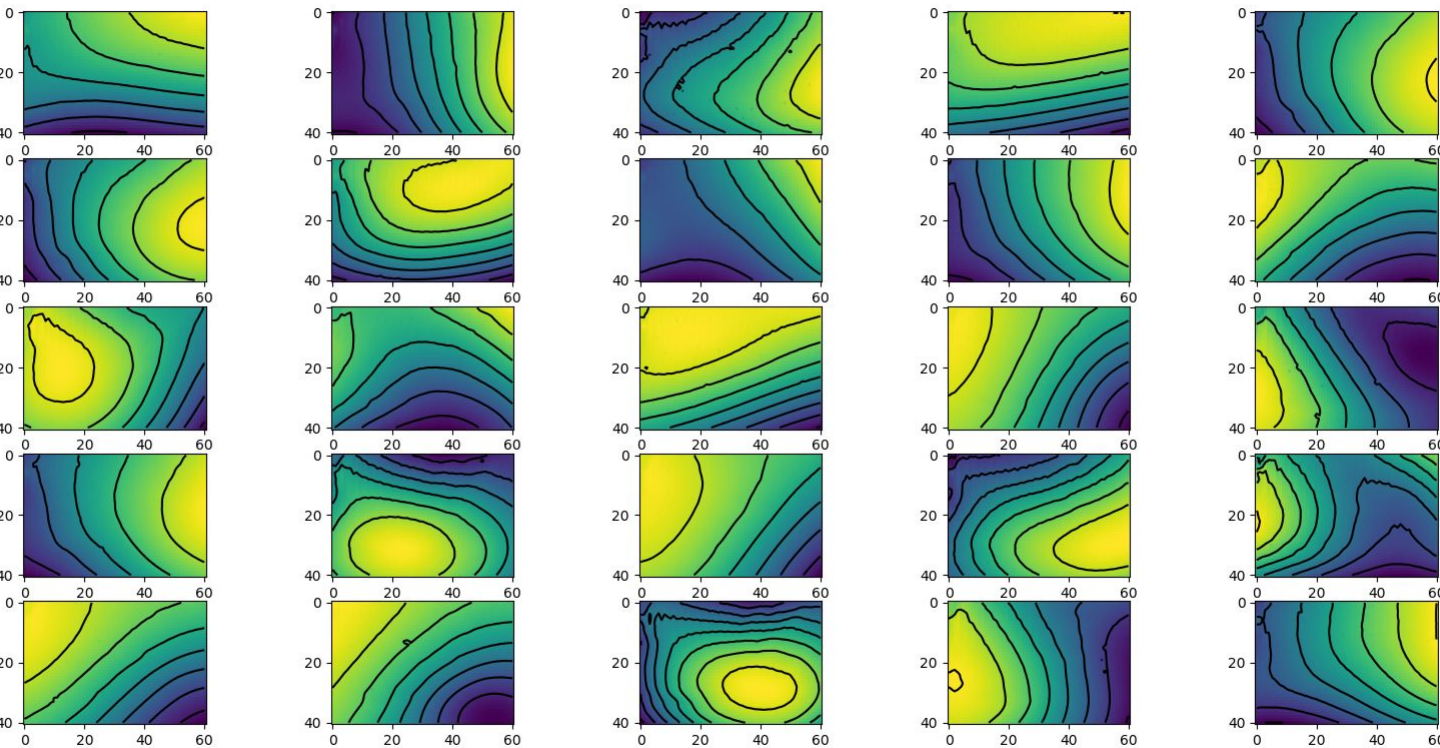
Distribution Matching Error (DME) = Testea si los nodos parecen provenir de la misma población que las observaciones mediante Kolmogorov Smirnov.

Distance Map = Para cada nodo, mide el promedio de la distancia a sus nodos adyacentes.

Activation Response = Para cada nodo, cuenta cuántas observaciones son más próximas.

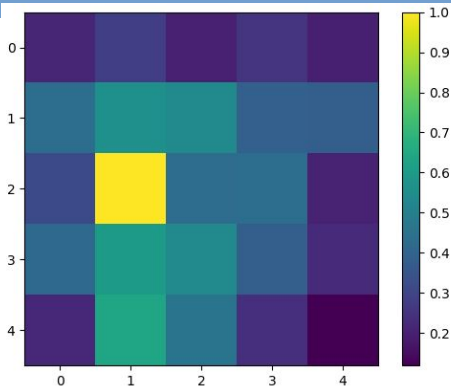
Norma promedio de los datos observados = 501584

Se utiliza k-means para inicializar

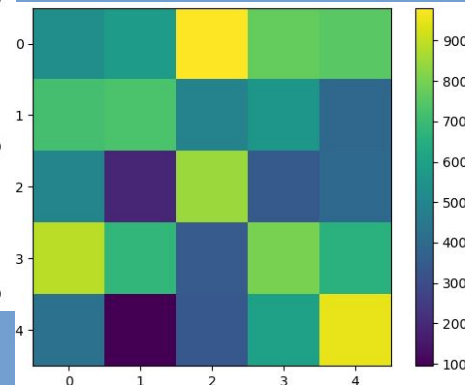


quantization_error = 7429
topographic_error = 0.8328
DME = 0.0

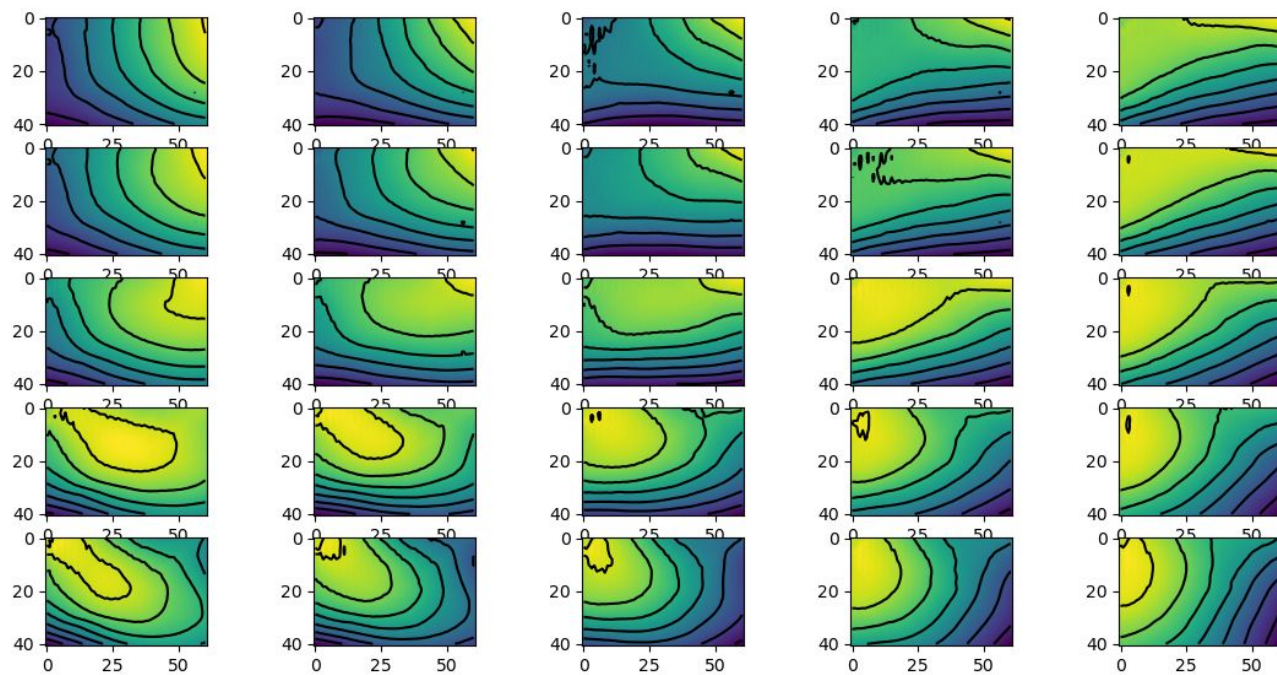
distance map



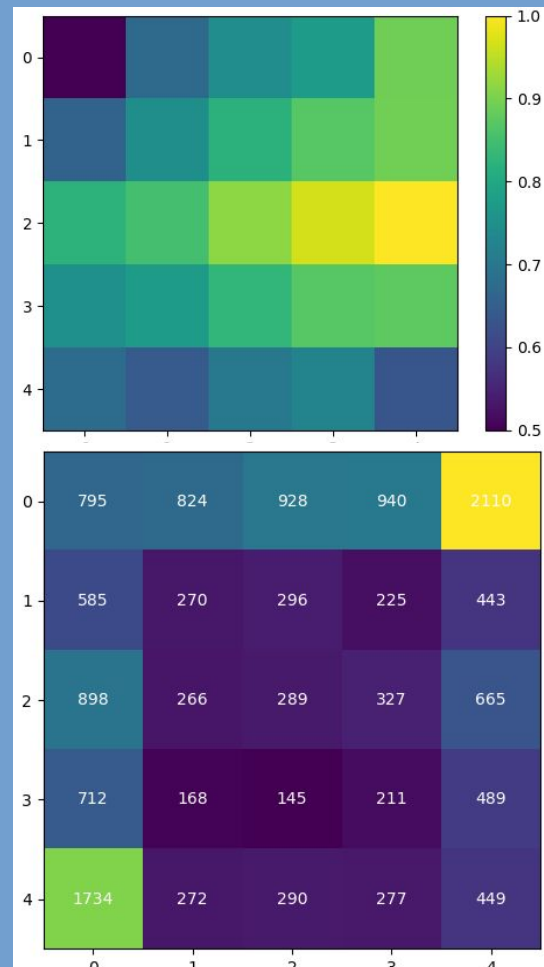
activation response



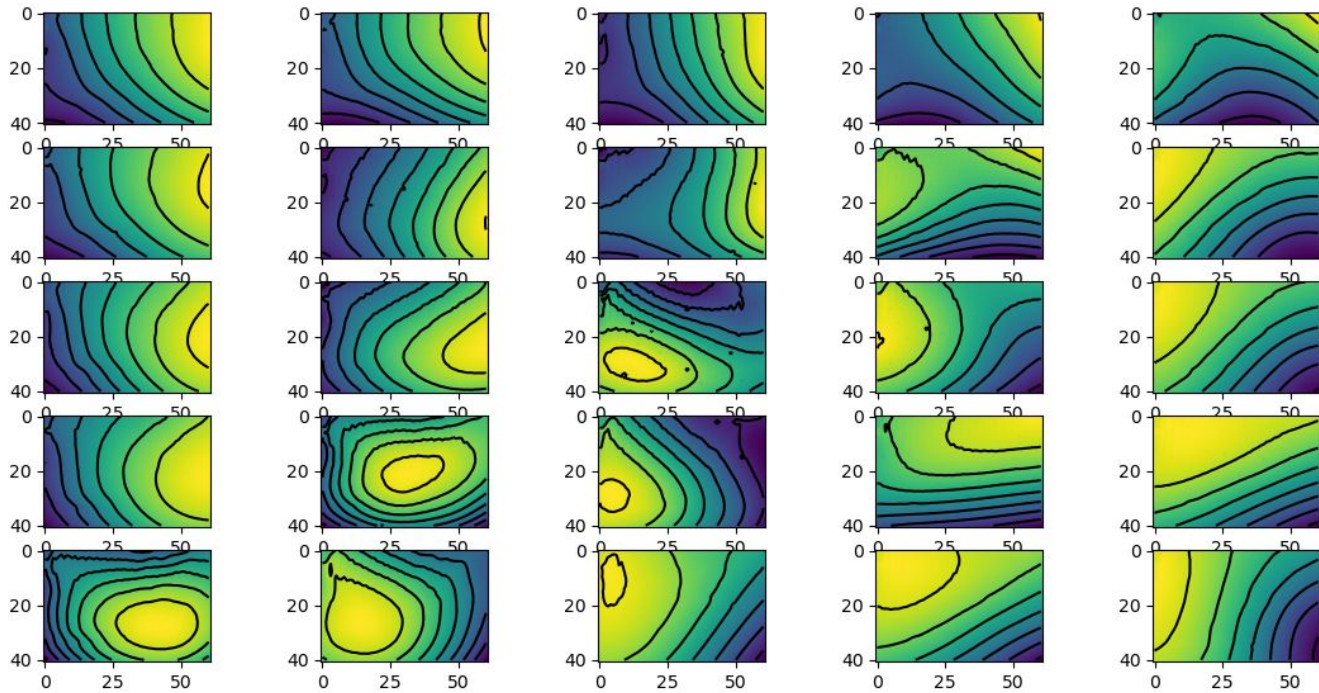
Entreno con sigma alto para acercar los nodos a la nube de datos



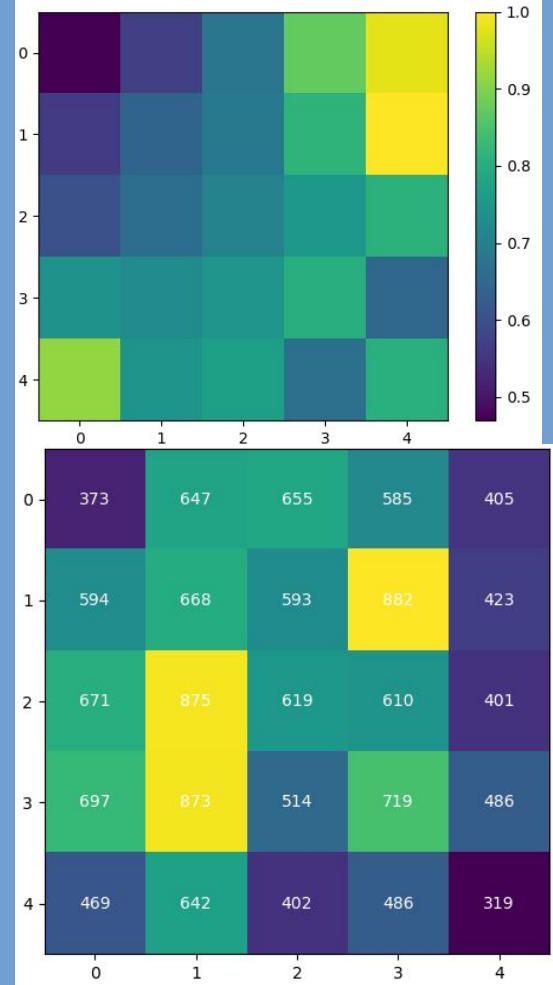
quantization_error = 10022
topographic_error = 0.006
DME = 0.055

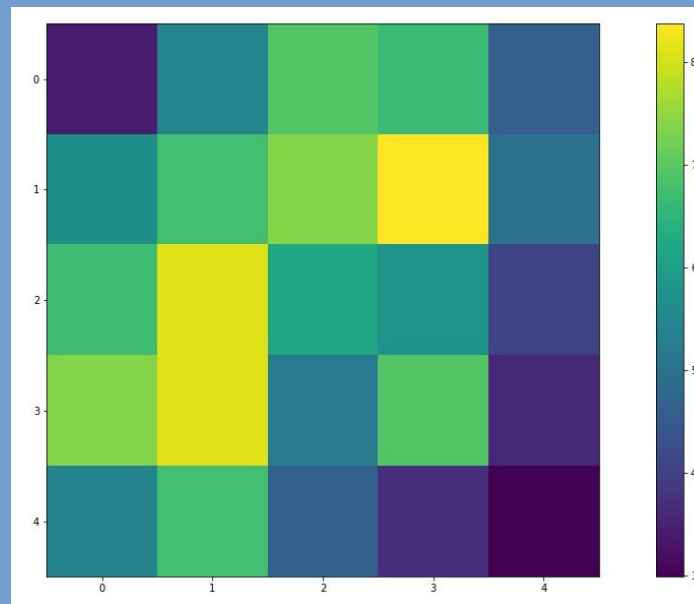
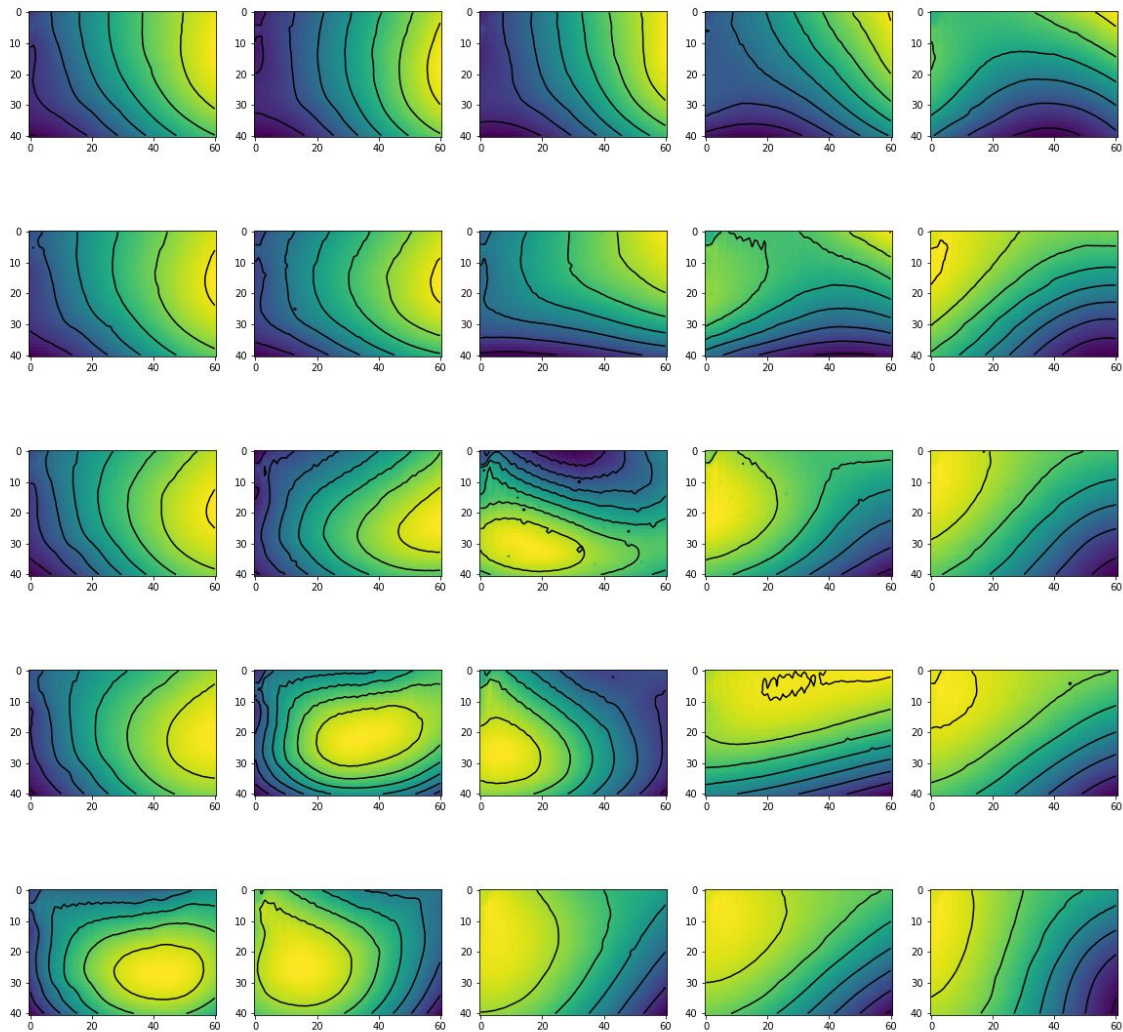


Continúo con sigma bajo por 50 iteraciones y learning_rate 0.001

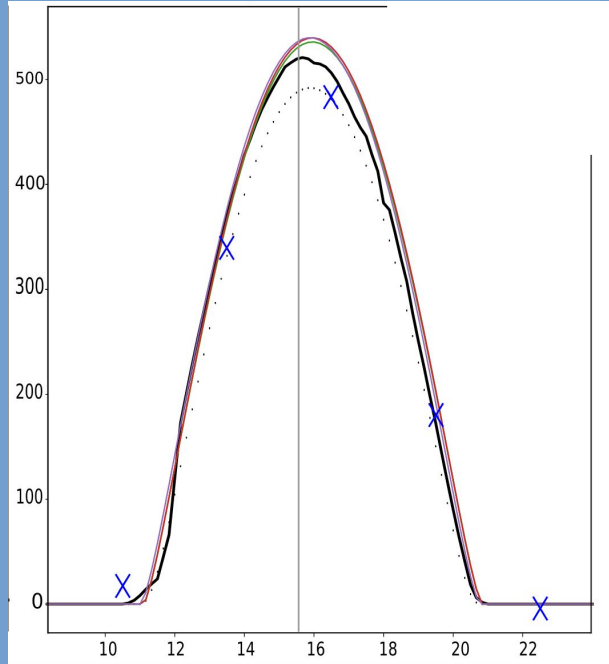


quantization_error = 7456
topographic_error = 0.0535
DME = 0.0

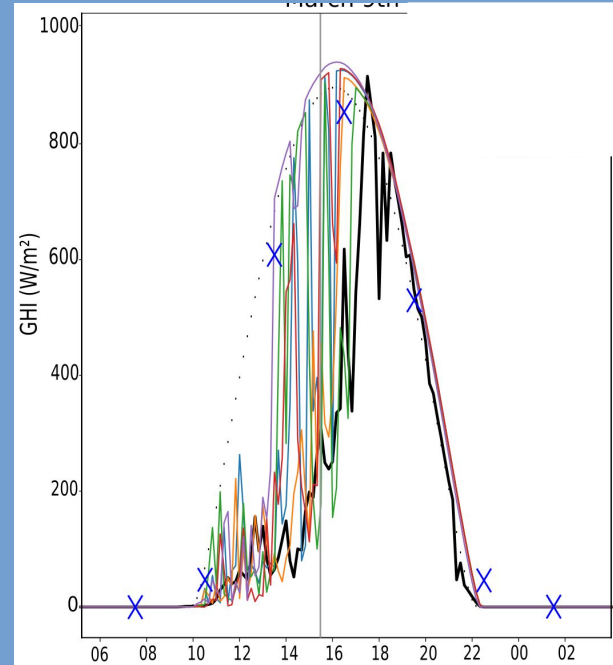




Para relacionar la circulación sinóptica con la radiación solar, se define la intermitencia como la longitud de la curva de la serie diurna de radiación con respecto a la longitud de la curva en un día despejado.



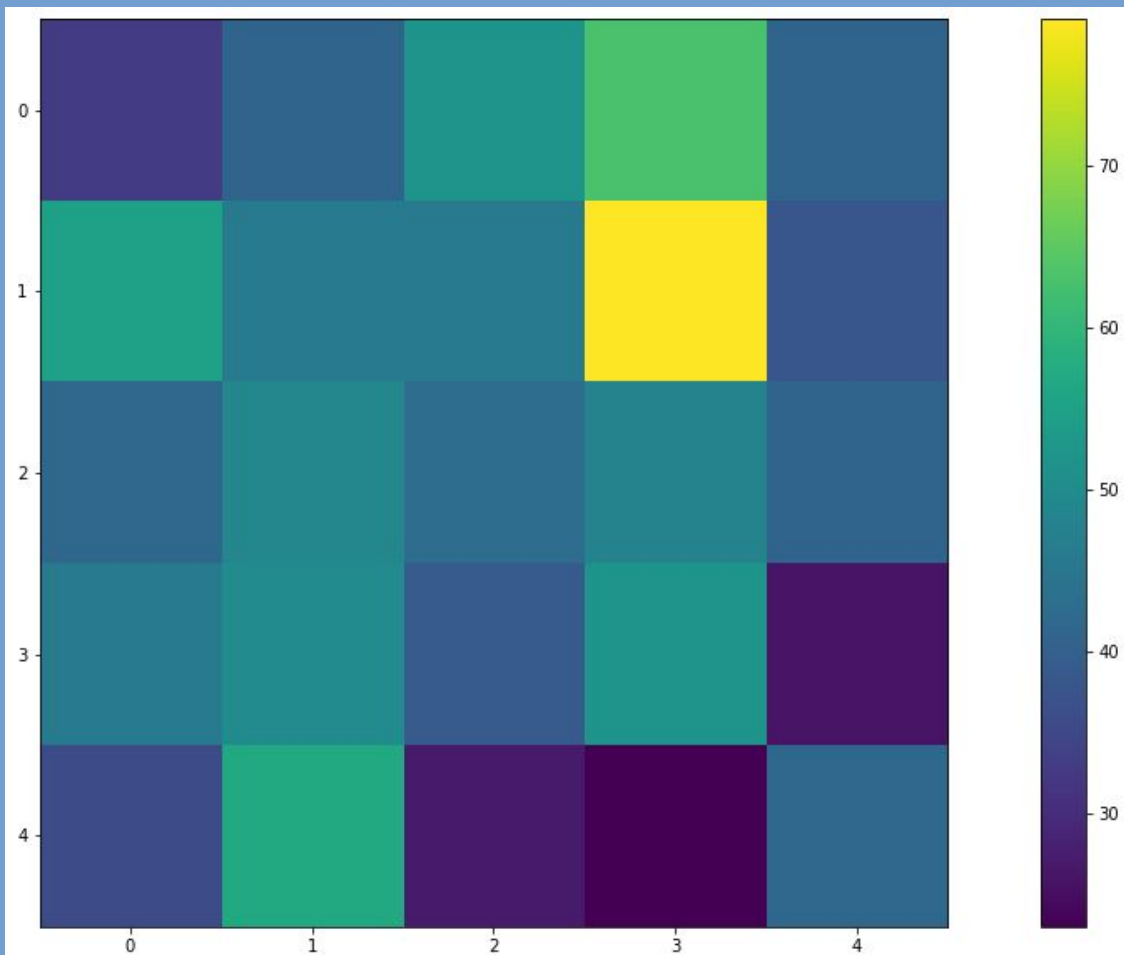
Baja intermitencia



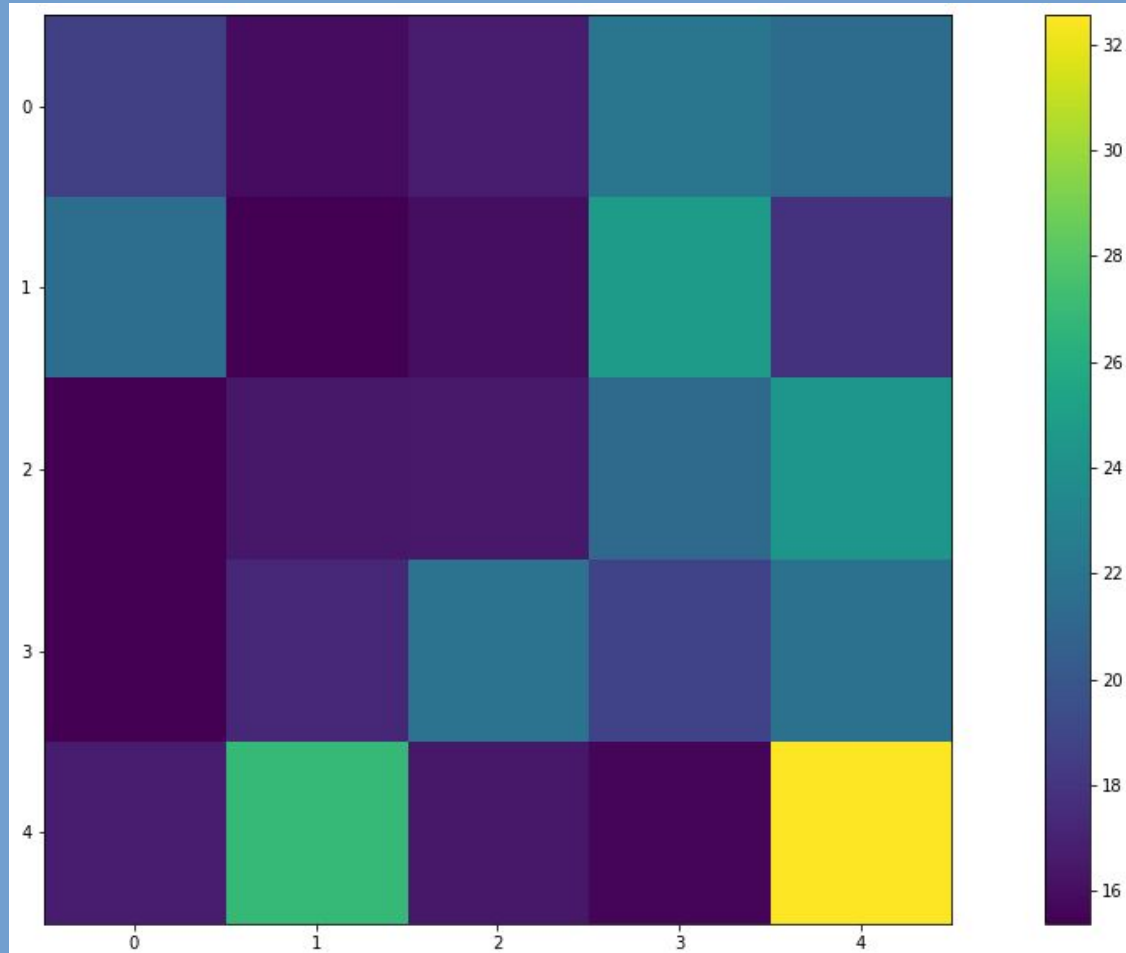
Alta intermitencia

casos con tort > media absoluto

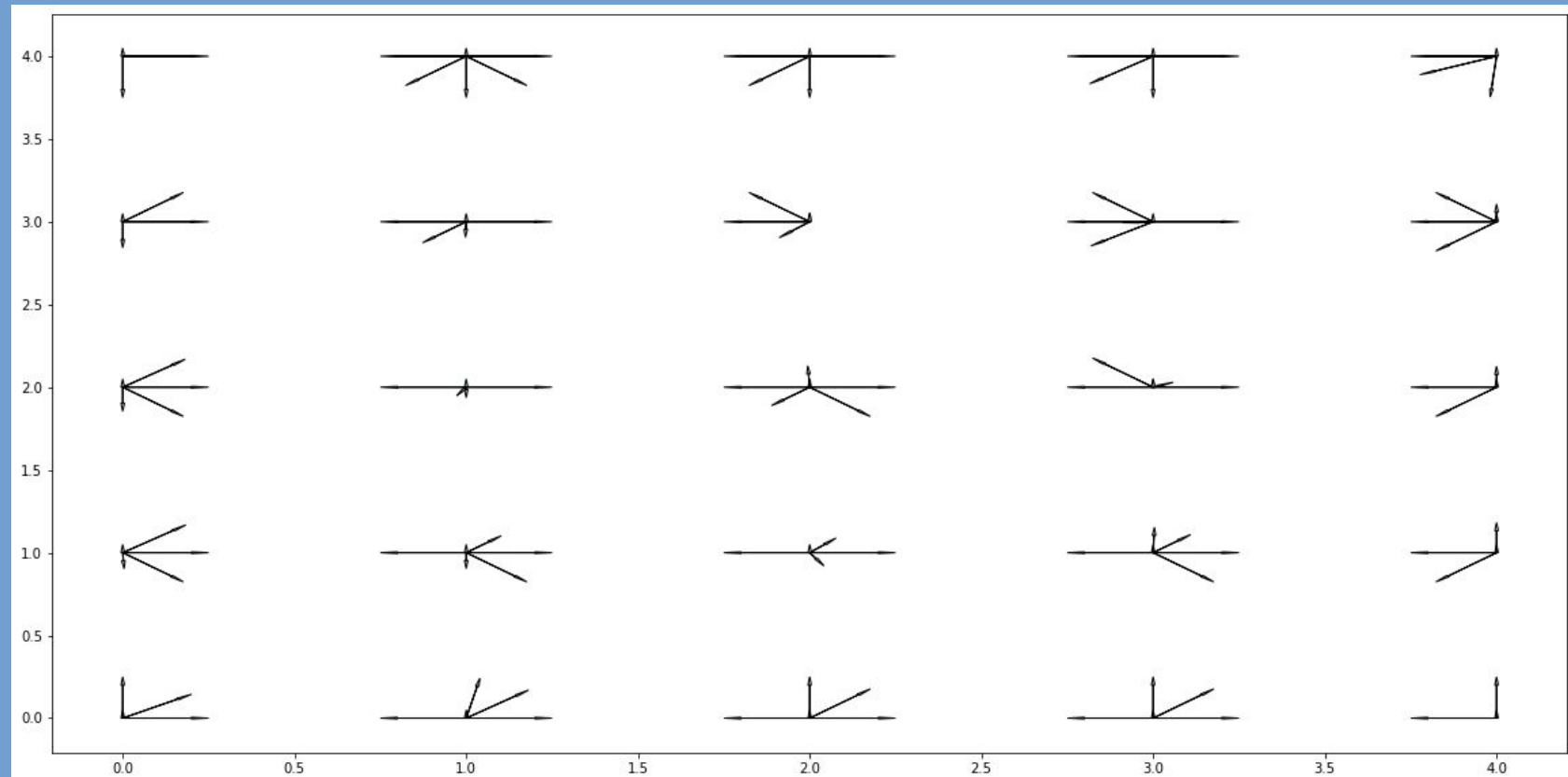
2017 2018



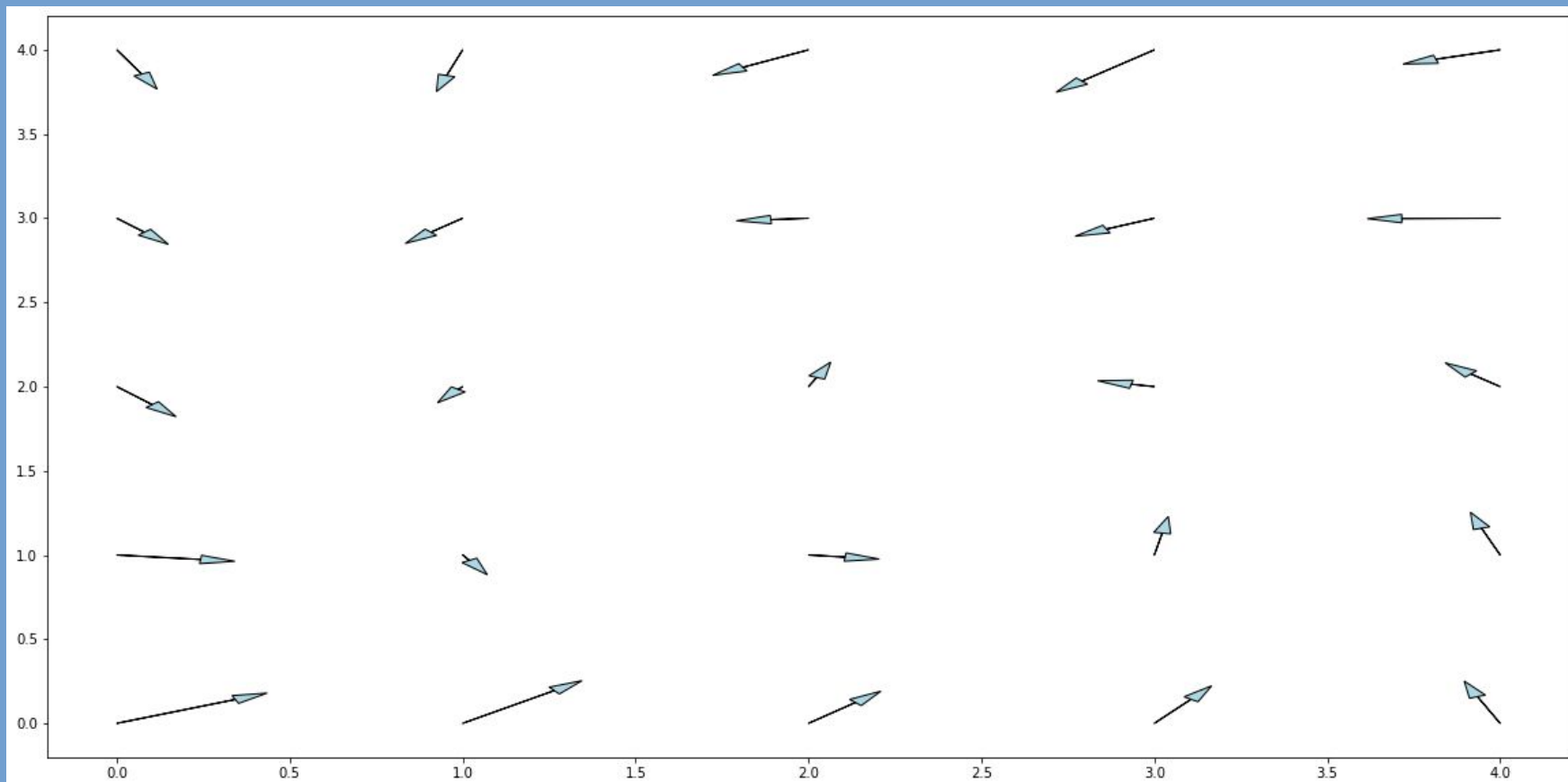
% dias con tort>media relativo a activation response



Una vez determinados los campos asociados a la alta intermitencia, se desea anticipar su ocurrencia mediante la identificación de su antecedente más probable.



Trayectorias medias



MUCHAS GRACIAS