

**Exercice 9 : Interpolation**

**SOLUTIONS**

## Question 1

Commentez les résultats de l'interpolation.

ME = -2.44

RMSE = 21.83

La moyenne des erreurs est assez éloignée de 0 et indique que l'interpolation sous-estime les valeurs aux points échantillonnés. La RMSE est particulièrement utile pour comparer avec d'autres approches d'interpolation (TIN ou alors IDW avec des paramètres différents). Dans tous les cas, l'échantillonnage n'est pas assez dense pour fournir une bonne interpolation.

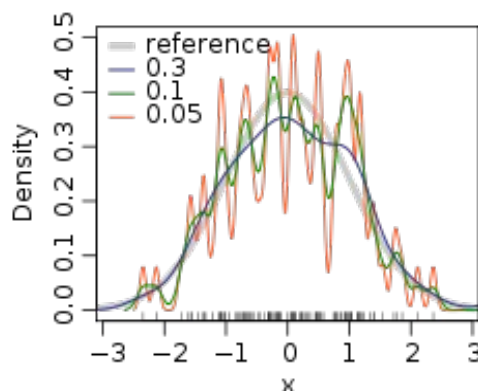
*Note :* Les résultats (ME et RMSE) seront sûrement légèrement différents de ceux obtenus ici car l'interpolation a été effectuée avec un train sample défini de manière aléatoire.

## Question 2

Quel est l'effet d'un rayon plus grand/plus petit ? Quel rayon vous semble plus adapté pour cette analyse ?

La carte de densité peut être vue comme un histogramme 2D de probabilité des séismes. Un bandwidth trop grand lisse les pics de densités et ne reflète pas les discontinuités locales, un bandwidth trop petit génère des petits pics autour de chaque point et n'effectue pas la généralisation que l'on attend. Il existe des méthodes pour évaluer le bandwidth idéal en fonction du nuage de points, mais ceci dépasse le cadre de cet exercice.

Dans notre cas, c'est le bandwidth de 10km qui semble le plus adapté à l'échelle des communes valaisannes. En effet, le bandwidth de 50km lisse l'impact des séismes en Valais qui débordent nettement de l'autre côté de la crête nord des Alpes. A l'inverse, le bandwidth de 1km produit une carte de densité beaucoup trop locale qui reflète les événements individuels mais pas une vue d'ensemble de ceux-ci.



*Figure 1: Wikipedia, Points générés aléatoirement en x en suivant une distribution normale. Le KDE est utilisé pour estimer la distribution normale. Trois bandwidth sont utilisés et approximent plus ou moins bien la distribution normale.*