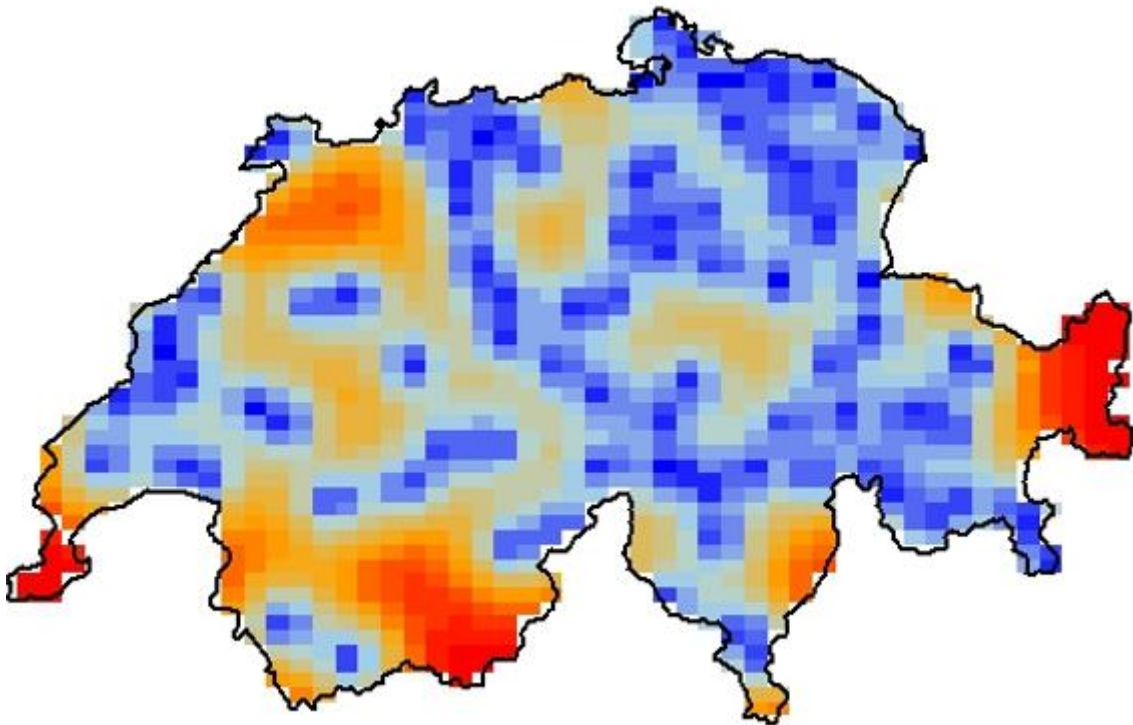




## Rapport du projet Krigeage



Qinyan YANG

Djae Jordy

06/05/2024

# Sommaires

1-	Introduction .....	3
2-	Visualisation des données .....	4
3-	Analyse Variographiques .....	6
4-	Ajustement d'un modèle de variogramme .....	8
5-	Krigeage.....	10
6-	Planification d'expériences .....	11

# 1- Introduction

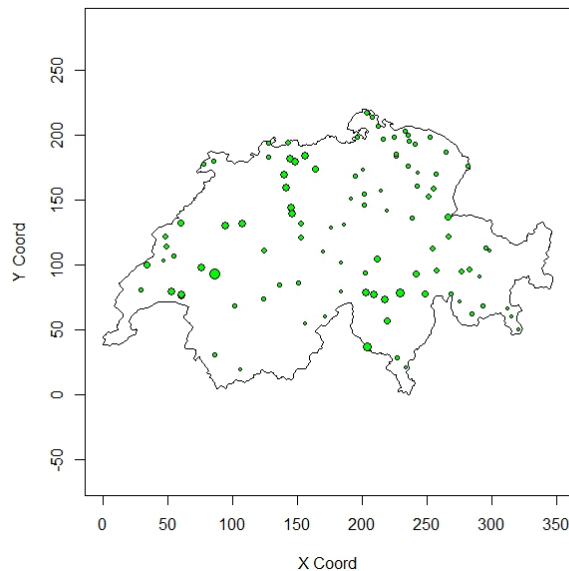
Les précipitations sont un élément fondamental du système climatique, influençant divers aspects de la vie humaine, de l'agriculture à l'approvisionnement en eau. Comprendre les schémas de précipitations est crucial pour la planification et la prise de décision dans de nombreux domaines, notamment l'agriculture, la gestion des ressources en eau et la prévision des catastrophes naturelles telles que les inondations.

La Suisse, avec son relief montagneux et sa position géographique, est particulièrement sensible aux variations des précipitations. La topographie complexe de la région influence la distribution spatiale et temporelle des précipitations, créant des défis uniques pour les scientifiques souhaitant comprendre et prédire ces phénomènes.

Dans ce rapport, nous explorons une visualisation des données, une analyse variographique, un ajustement d'un modèle de variogramme, du Krigeage et une planification d'expériences.

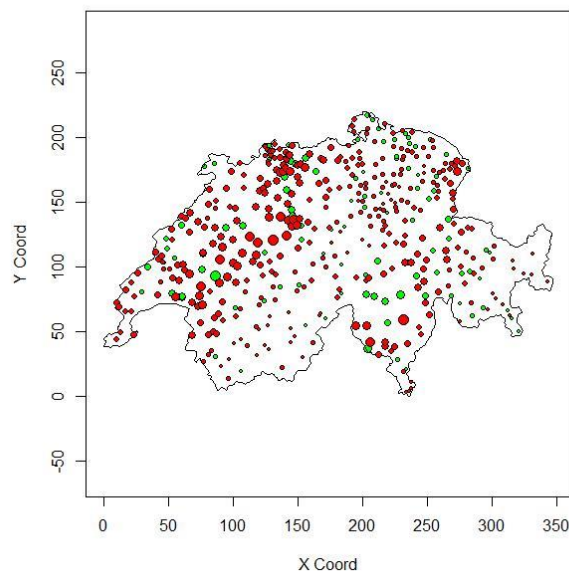
## 2- Visualisation des données

Dans cette section, nous allons examiner les informations quotidiennes fournies par un échantillon de 100 postes météorologiques situés en Suisse. Ces informations portent sur les précipitations et sont distribués par les stations, comme le montre la carte ci-dessous.

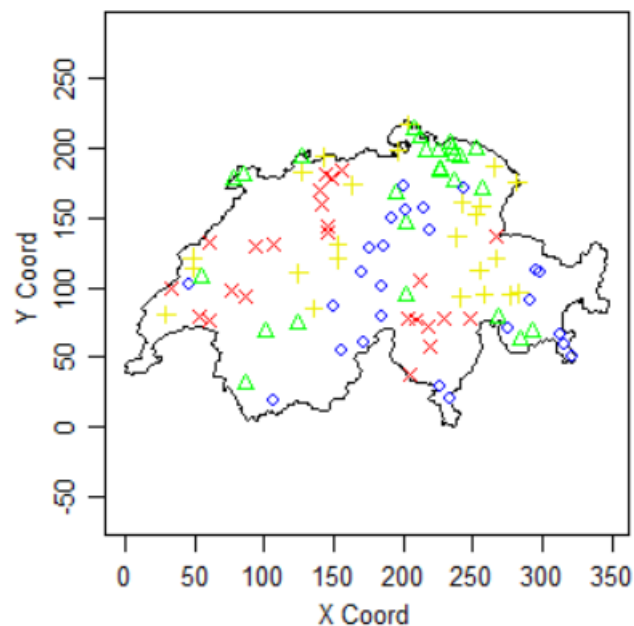


La carte ci-dessus illustre une distribution quasi-uniforme des précipitations à travers la Suisse. Cependant, elle ne donne pas suffisamment de détails sur les régions où les pluies sont les plus intenses ou les plus légères, ni sur la disposition spatiale de la quantité de pluie. Par conséquent, la carte ci-dessus utilise des quartiles pour diviser la quantité de pluie en quatre catégories colorées, allant des faibles précipitations aux fortes pluies.

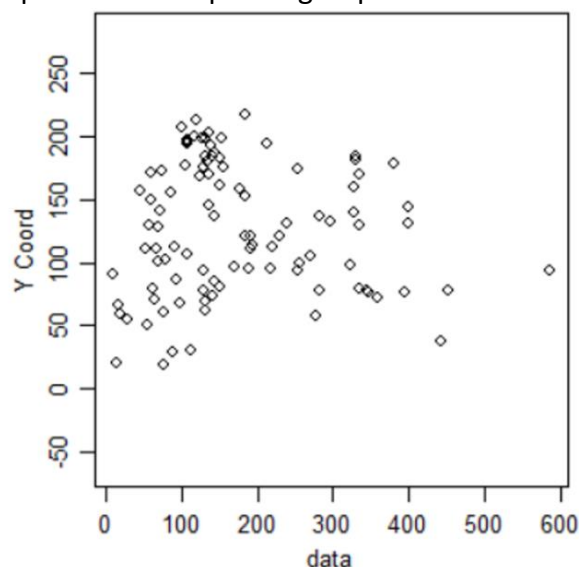
La carte ci-dessous illustre des postes météorologiques non incluses dans l'échantillon qui permettront de comparer les estimations et les observations.

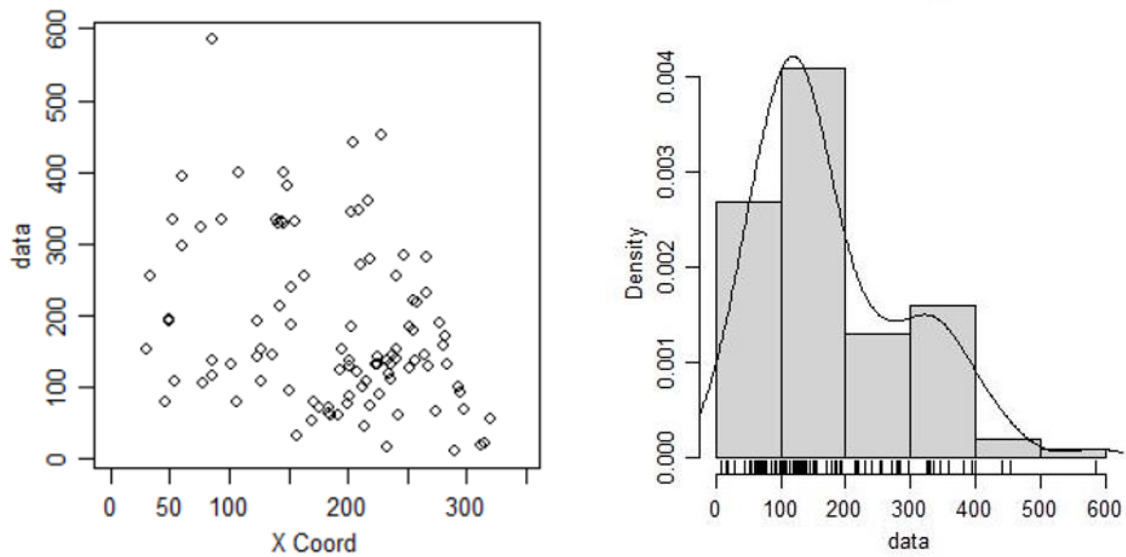


En examinant la carte ci-dessus et en la mettant en parallèle avec la première, il est évident que le bleu symbolise le quartile inférieur, indiquant une faible quantité de pluie, tandis que le rouge représente le quartile supérieur, signifiant une grande quantité de pluie. Les deux autres couleurs, le vert et le jaune, correspondent aux quartiles intermédiaires. On peut également constater que les croix rouges se regroupent, tout comme les losanges bleus, et ainsi de suite. Cela démontre une certaine logique spatiale, indiquant une structuration de l'espace.



Donc pour répondre à la question posée, on peut dire que d'après les différentes observations faites sur les figures ci-dessus, on peut conclure qu'il y a une structuration spatiale de la quantité de pluie ; dans le sens où les zones où il pleut beaucoup se regroupent même choses pour les zones qu'il pleut moins. En poussant un peu plus nos observations, On observe une certaine uniformité des précipitations en se déplaçant du sud vers le nord, ainsi que de l'Est vers l'Ouest, il est raisonnable e de conclure qu'en Suisse, la latitude ou la longitude n'a pas vraiment d'impact sur les conditions météorologiques.

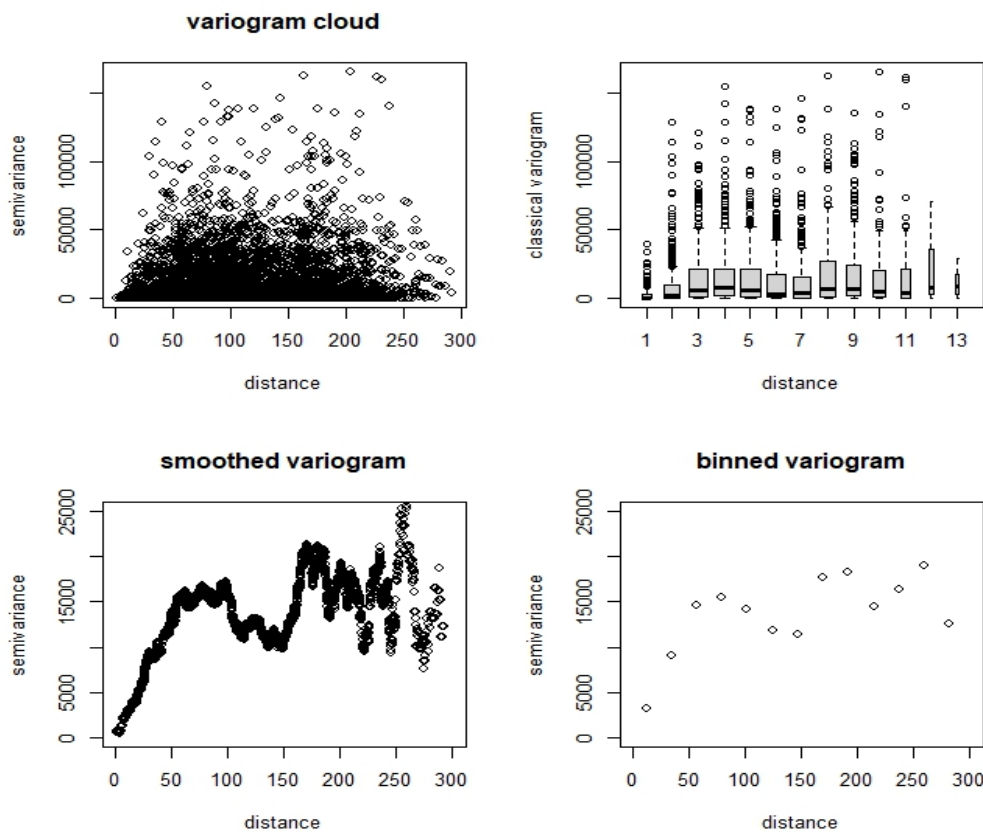




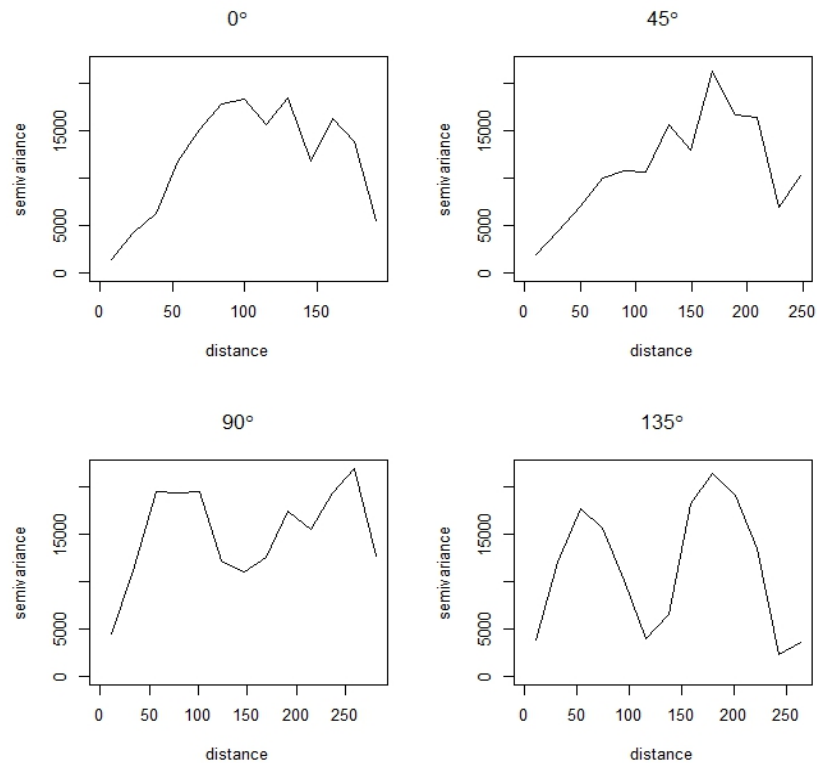
### 3- Analyse Variographiques

- Variogram cloud (ou Nuage de points variogramme) : Ce diagramme présente le variogramme sous la forme d'un nuage de points. Chaque point symbolise la distance entre les paires de données et la variance correspondante. Il est utile pour visualiser la structure de dépendance spatiale des données.
- Clouds for binned variogram (ou Nuages pour le variogramme biné) : Ce diagramme illustre les nuages de points utilisés pour le variogramme biné. Le variogramme biné divise l'espace en cellules et calcule la variance pour chaque cellule. Ces nuages de points représentent les paires de données dans chaque cellule.
- Smoothed variogram (ou variogramme lissé) : Ce diagramme représente le variogramme lissé des données. Le lissage permet de réduire le bruit dans le variogramme en moyennant les valeurs sur des distances voisines. Il est utile pour identifier les tendances générales dans la structure de dépendance spatiale.

- **Binned variogram (ou variogramme Biné)** : Ce diagramme montre le variogramme des données binées, où l'espace est divisé en cellules et la variance est calculée pour chaque cellule. Il permet de visualiser la variance moyenne à différentes distances spatiales.



Les graphiques suivants sont effectivement utilisés pour analyser la structure de dépendance spatiale des données météorologiques représentées par sic.100. Ils permettent d'explorer et de comprendre les variations spatiales et les tendances dans les données, ce qui est crucial pour m'analyse météorologique. La prochaine carte met en évidence une variabilité plus prononcée dans la direction verticale qu'horizontale. De plus, le variogramme atteint des distances plus importantes à 45° par rapport aux autres directions (0°, 90°, 135°). Cela pourrait indiquer une certaine anisotropie dans les données, c'est-à-dire que les propriétés statistiques varient en fonction de la direction.



Pour répondre à la question demandée, on peut dire qu'à partir de 50 km, il n'y a plus de corrélation, puisqu'on a atteint la variance, c'est-à-dire que la différence entre une donnée et une autre séparée de plus de 100 km sera potentiellement grande. Au-delà de 100 km, les données n'ont aucune raison d'avoir des comportements similaires, par contre quand la distance est petite, il y a une ressemblance des données. On remarque que dans une même région, il pleut de même façon. Il y a une variabilité plus forte verticalement qu'horizontalement.

## 4- Ajustement d'un modèle de variogramme

Dans cette partie, on a différentes représentations de variogrammes ajustés avec modèles : Sphérique, Exponentiel et Exponentiel avec pépité.

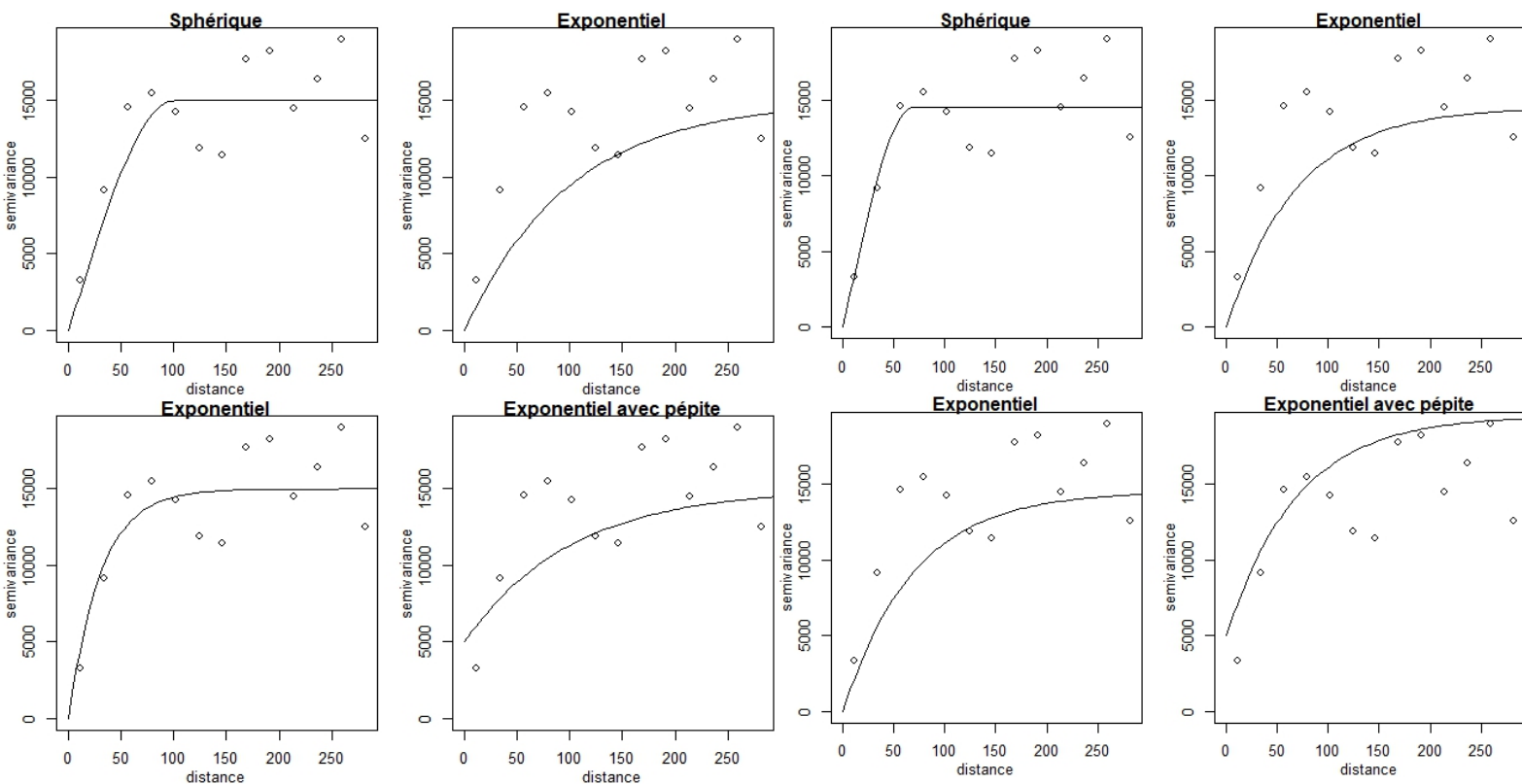
- Sphérique : le graphique montre une courbe qui augmente puis se stabilise à mesure que la distance augmente. Cela peut indiquer un exportée spécifique au-delà de laquelle la dépendance spatiale n'augmente plus.
- Exponentiel : les deux graphiques exponentiels dépeignent des courbes qui augmentent continuellement mais à un rythme décroissant à mesure que la distance



augmente. Cela suggère une dépendance spatiale qui persiste sur de longues distances.

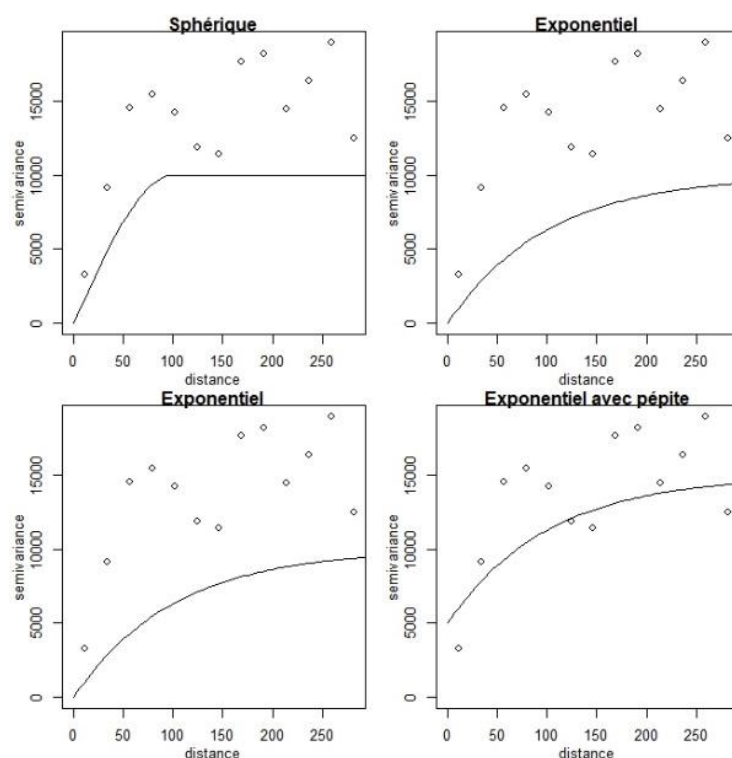
- Exponentiel avec pépite : ce graphique a une courbe similaire aux graphiques exponentiels mais commence à partir d'une valeur de semi variance plus élevée à une distance zéro, indiquant un effet de pépite. Cela signifie qu'il y a une certaine variabilité à petite échelle qui n'est pas capturée par le modèle.

Ces différents graphiques de paramètres différents sont utiles pour comprendre la structure de dépendance spatiale des données et pour choisir le meilleur modèle pour l'analyse géostatistique.



Paramètres utilisés : cov.pars = c(15000,100)

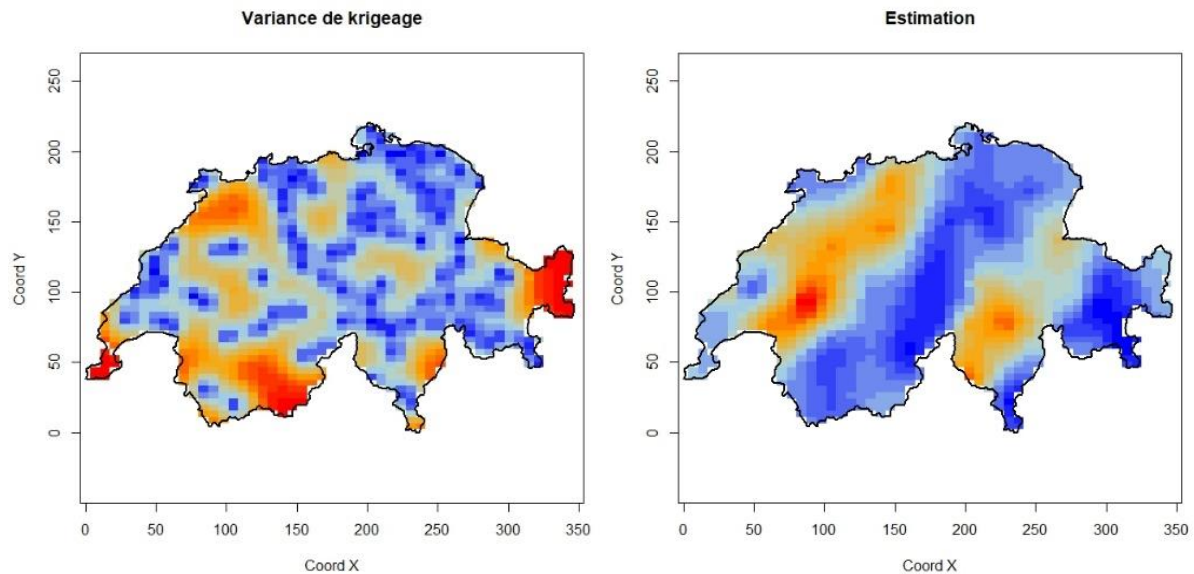
Paramètres utilisés : cov.pars = c(1452269)



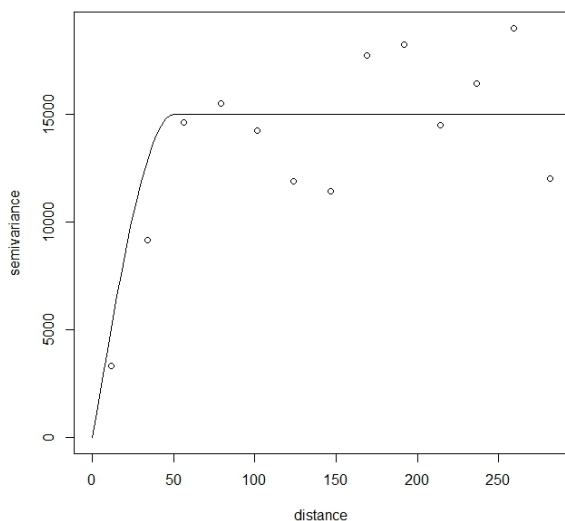
Paramètres utilisés : cov.pars = c(10000,100)

## 5- Krigeage

D'après les cartes ci-après, il semble que la distribution spatiale des données météorologiques en Suisse présente une certaine variabilité. Les zones en rouge/orange indiquent des valeurs plus élevées, ce qui pourrait signifier des précipitations plus importantes, tandis que les zones en bleu représentent des valeurs plus basses, ce qui pourrait indiquer des précipitations plus faibles. Sa pourrait suggérer que certaines régions de la Suisse ont tendance à avoir des conditions météorologiques plus extrêmes que d'autres. Cependant, sans plus d'informations spécifiques sur ce que représentent exactement les couleurs, il est difficile de tirer des conclusions plus précises.



Une chose est sûre, on peut dire que l'erreur de prédiction est plus forte là où il n'y a pas de station météo.



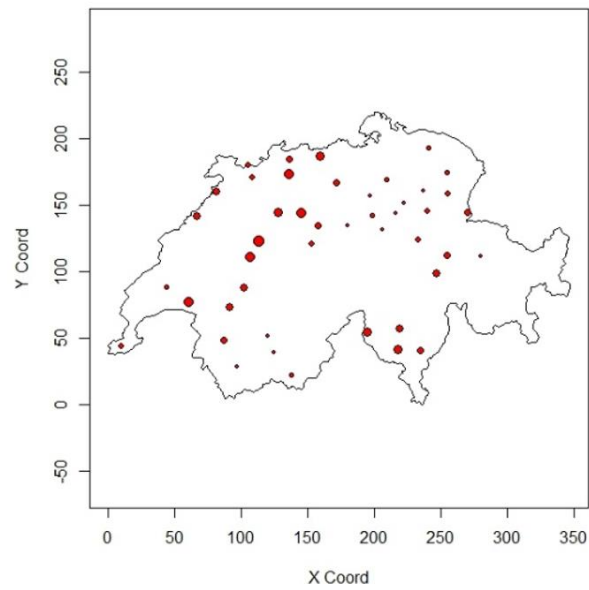
Comme il n'y a pas l'effet de pépité, l'observation est en moyenne moche. Dans le sens où quand l'observation est faible sa prédiction est faible, et quand l'observation est forte, sa prédiction est forte. En moyenne, la tendance est là, mais ce n'est pas terrible.

Erreur de test (appelée aussi erreur de généralisation) vaut  $MSE = 3744,822$ .

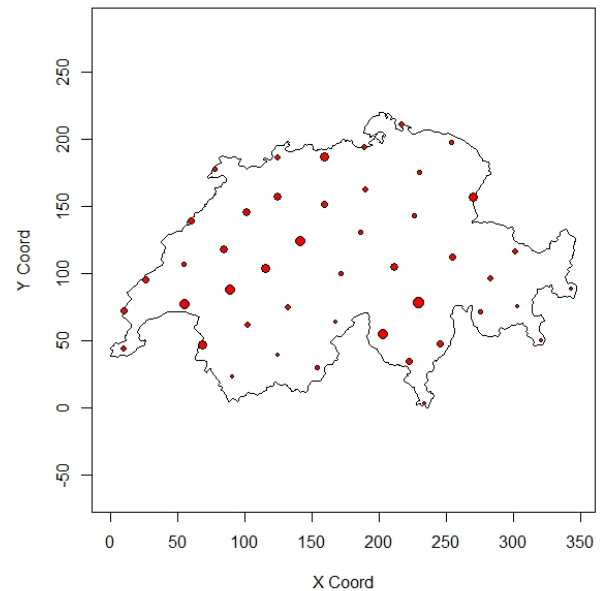
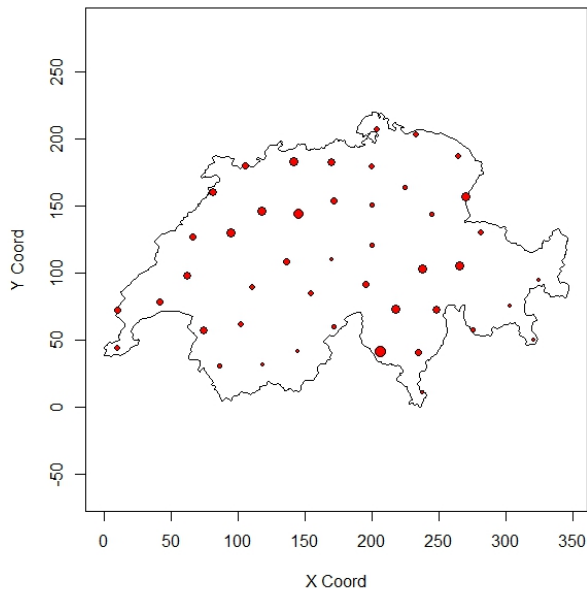
Et pour le coefficient de prédictivité, noté  $Q^2$ , vaut 0.6969142

## 6- Planification d'expériences

Pour la méthode a) qui calcule des distances euclidiennes et qui ne prend pas en compte les bornes du domaine car ce n'est pas convexe, les 45 stations optimales parmi les 467 sont représentées sur la figure ci-après. Le coefficient de prédictivité  $Q^2=0.469295$ .



Pour les méthodes b) et c) les 45 stations optimales parmi les 467 sont représentées sur la figure ci-après de gauche à droite respectivement. Les coefficients de prédictivité pour les méthodes b) et c) sont respectivement,  $Q^2=0,6172156$  et  $Q^2=0,6274703$ .



À la vue de ces différents résultats, on peut en conclure que la méthode a) est la meilleure, car elle a le plus petit coefficient de prédictivité.