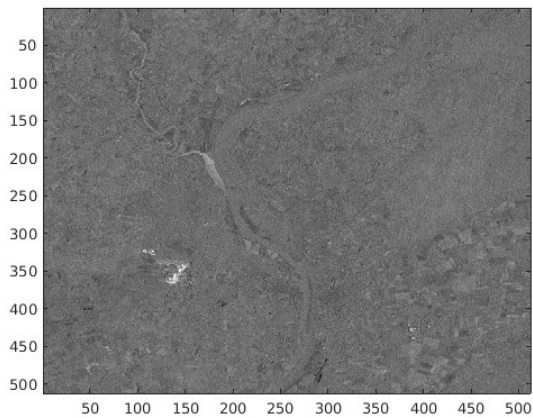


## SI221 : classification automatique

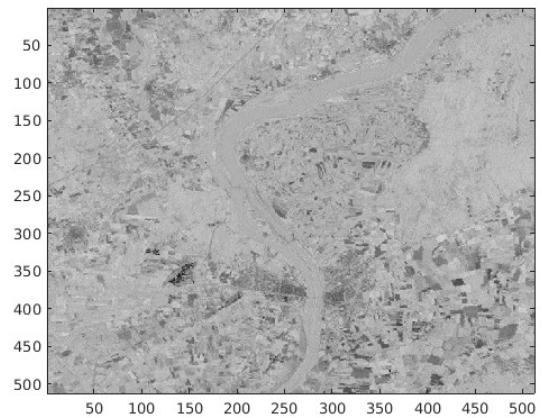
### 3. Analyse en Composantes Principales

#### 3.3 Application aux composantes de LANDSAT sur Tarascon

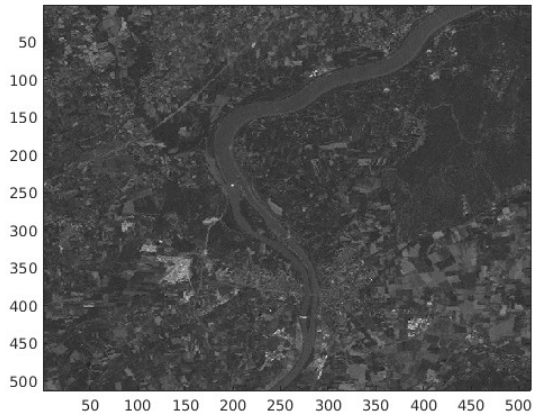
L'application de l'algorithme de l'ACP sur les 8 images LANDSAT de Tarascon donne les 8 nouvelles images ci-dessous, avec leur pourcentage d'inertie, qui correspond qualitativement au pourcentage d'information qu'offre chaque image par rapport à l'ensemble de données initial.



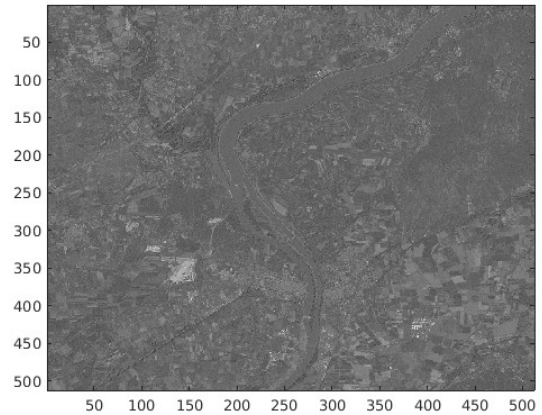
$\lambda = 0.0100$  ; inertie = 0.1249%



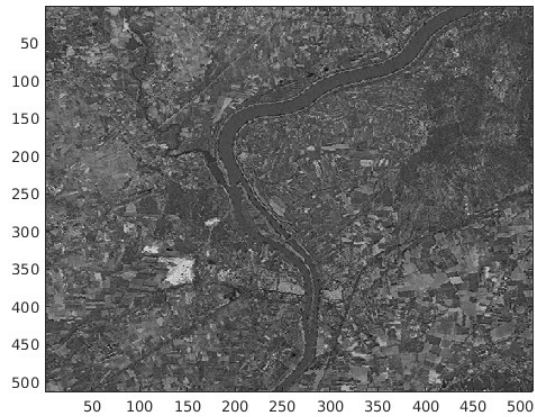
$\lambda = 0.0206$  ; inertie = 0.2573%



$\lambda = 0.0310$  ; inertie = 0.3874%



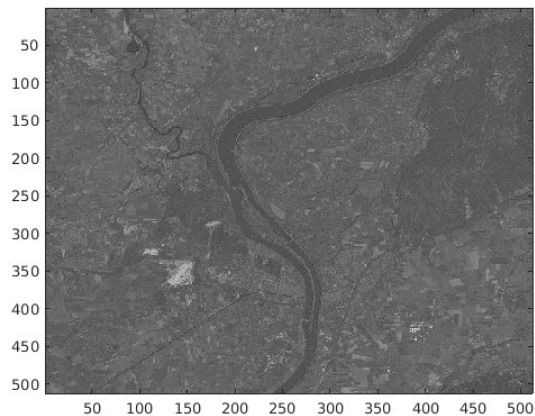
$\lambda = 0.1319$  ; inertie = 1.6484%



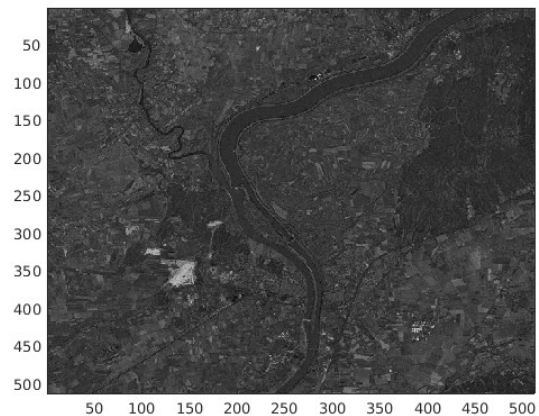
$\lambda = 0.2232$  ; inertie = 2.7896%



$\lambda = 0.7334$  ; inertie = 9.1678%



$\lambda = 1.3258$  ; inertie = 16.5719%



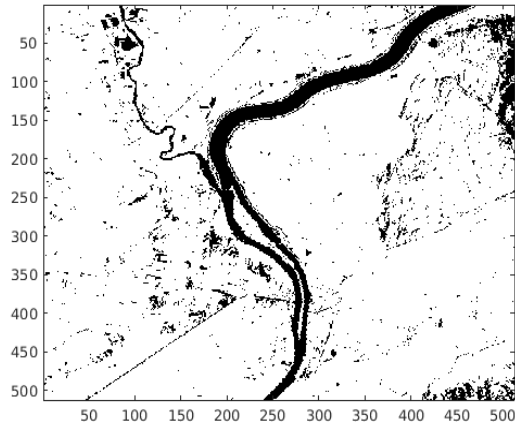
$\lambda = 5.5242$  ; inertie = 69.0528%

À l'aide d'une somme cumulée sur les coefficients d'inertie, on voit qu'il suffit de prendre les trois (94.79%) ou les quatre (97.58%) premières images pour avoir environ 95 % des informations des données précédentes.

## **4 Classification automatique : algorithme des K-moyennes**

### **4.2 LANDSAT Tarascon**

On a choisi comme pixel de terre les coordonnées (277, 229) et comme pixel d'eau les coordonnées (190, 179). Quand on applique l'algorithme des K-moyennes au canal 4 seul, on voit que tous les pixels d'eau sont effectivement reconnus, mais beaucoup de pixels de terre le sont aussi. Quant au canal 1, les valeurs des centroïdes ne permettent apparemment pas du tout une classification adéquate.



*Classification avec le canal 4*



*Classification avec le canal 1*

On pourra noter de plus une tentative d'utiliser les 8 canaux à la fois pour augmenter la précision, ce qui au contraire ne fait qu'empirer la qualité de l'image.



On peut penser à justifier cela par le fait qu'une saturation de données « inutiles » (dans le sens que les données écartent l'apprentissage au lieu de l'orienter) dû à l'utilisation de 8 canaux fausse le résultat, et qu'il aurait été préférable d'utiliser les résultats d'un ACP, réduisant ainsi la dimension de l'espace des paramètres à 3 ou 4 au lieu de 8, tout en concentrant les données importantes. On émettra également l'hypothèse que les centroïdes des classes ne soient pas adaptés pour les 8 images comme ils l'étaient pour le canal 4 seul.