



Classification non supervisée multi-stratégie

Pierre Gançarski
Cédric Wemmert

Le problème

- Données de plus en plus complexes
 - Différents types (numérique, symbolique, hétérogène ...)
 - Différentes origines (sources différentes)
 - Différents domaines d'analyse

Le problème

- Données de plus en plus complexes
 - Différents types (numérique, symbolique, hétérogène ...)
 - Différentes origines (sources différentes)
 - Différents domaines d'analyse
- Méthodes nombreuses
 - Différentes familles de classifieurs
 - Différentes méthodes dans une même famille
 - Nombreux paramètres par méthode

Le problème

- Données de plus en plus complexes
 - Différents types (numérique, symbolique, hétérogène ...)
 - Différentes origines (sources différentes)
 - Différents domaines d'analyse
 - Méthodes nombreuses
 - Différentes familles de classifieurs
 - Différentes méthodes dans une même famille
 - Nombreux paramètres par méthode
- ➔ Comment aider l'utilisateur ?

Un exemple



Image SPOT : 200*250 / 3 bandes sur 1 octet

- Quelle(s) méthode(s) choisir pour extraire les classes de pixels de cette image ?

Le problème

- Comment aider l'utilisateur dans son choix
 - de la méthode en fonction :
 - des données manipulées (symbolique vs numérique)
 - du « type » du résultat (hiérarchie vs partitionnement)

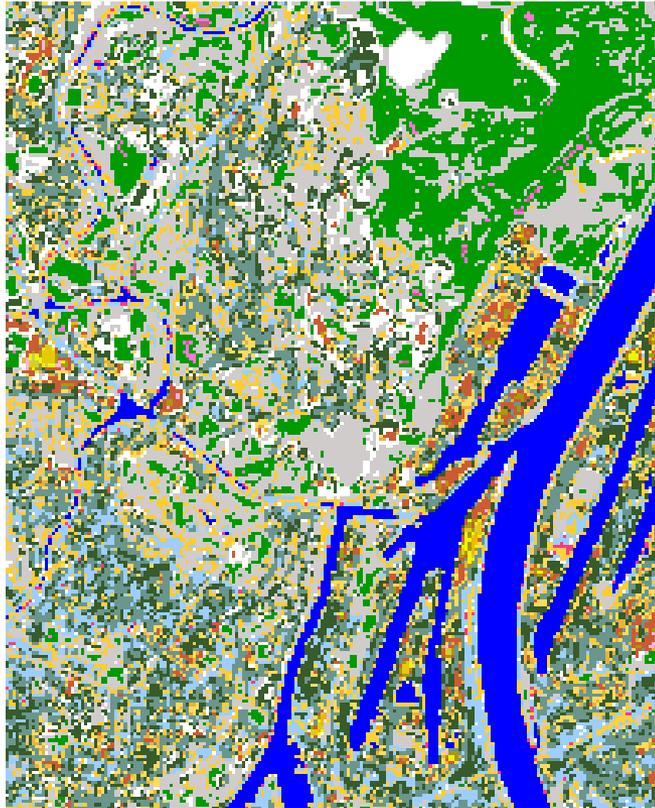
Le problème

Cobweb

ou

Kmeans

- acuité 3,5



→ 23 classes



- K = 20

→ 20 classes

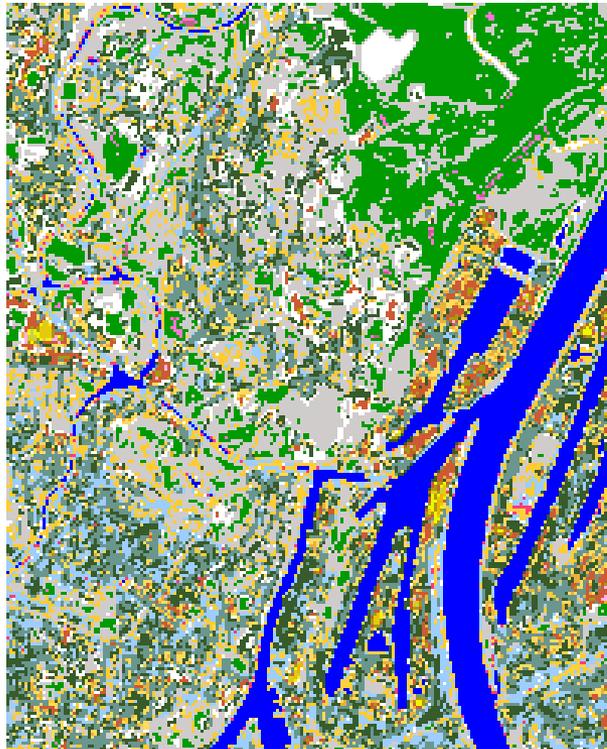
Le problème

- Comment aider l'utilisateur dans son choix
 - de la méthode en fonction :
 - des données manipulées (symbolique vs numérique)
 - du « type » du résultat (hiérarchie vs partitionnement)
 - des paramètres de la méthode
 - exemple : nombre de classes, acuité

Le problème

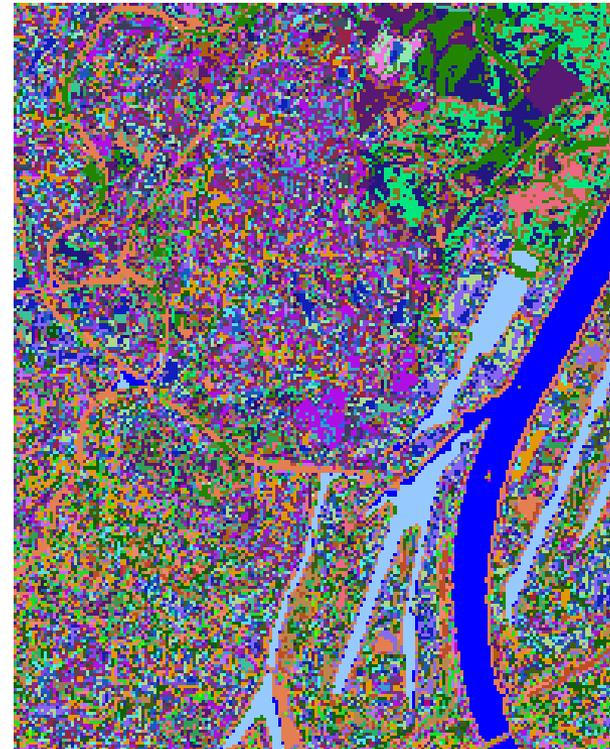
- Cobweb : influence de l'acuité

- acuité 3,5



→ 23 classes

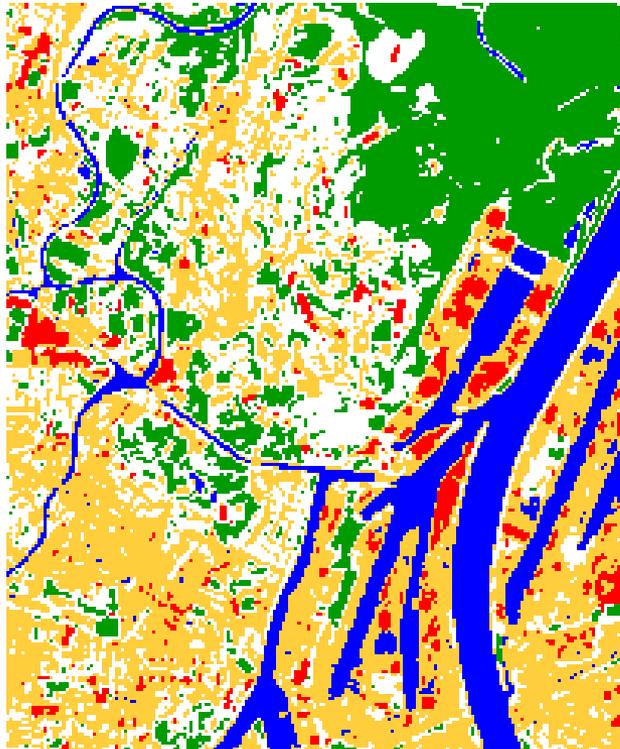
- acuité 1



→ 73 classes

Le problème

- Kmeans : influence du nombre de classes
 - $K = 5$
 - $K = 20$



Le problème

- Comment aider l'utilisateur dans son choix
 - de la méthode suivant :
 - les données manipulées (symbolique vs numérique)
 - le « type » du résultat (hiérarchie vs partitionnement)
 - des paramètres de la méthode
 - Ex : nombre de classes, acuité

➔ Une possibilité : apprentissage guidé

Mais ...

Notre proposition

- Les méthodes sont (dans une certaine mesure) complémentaires
 - Sur leurs types de données de prédilection
 - Sur la structures de leurs résultats
- ➔ Notre proposition : offrir à l'utilisateur le meilleur de chacune des méthodes

Notre proposition

- Offrir à l'utilisateur le meilleur de chacune des méthodes
 - L'utilisateur au lieu d'un seul résultat dispose du résultat de plusieurs méthodes pouvant être différents mais mutuellement « optimisés »
 - Hiérarchie élaguée automatiquement
 - Nombre de classes évoluant dynamiquement
 - L'utilisateur dispose en plus d'un résultat unifiant

Notre proposition

- Offrir à l'utilisateur le meilleur de chacune des méthodes
 - L'utilisateur dispose de résultats mutuellement optimisés et d'un résultat unifiant
- Avec
 - un critère d'optimisation indépendant des méthodes utilisées : intervalle pour K , critère intra-classes

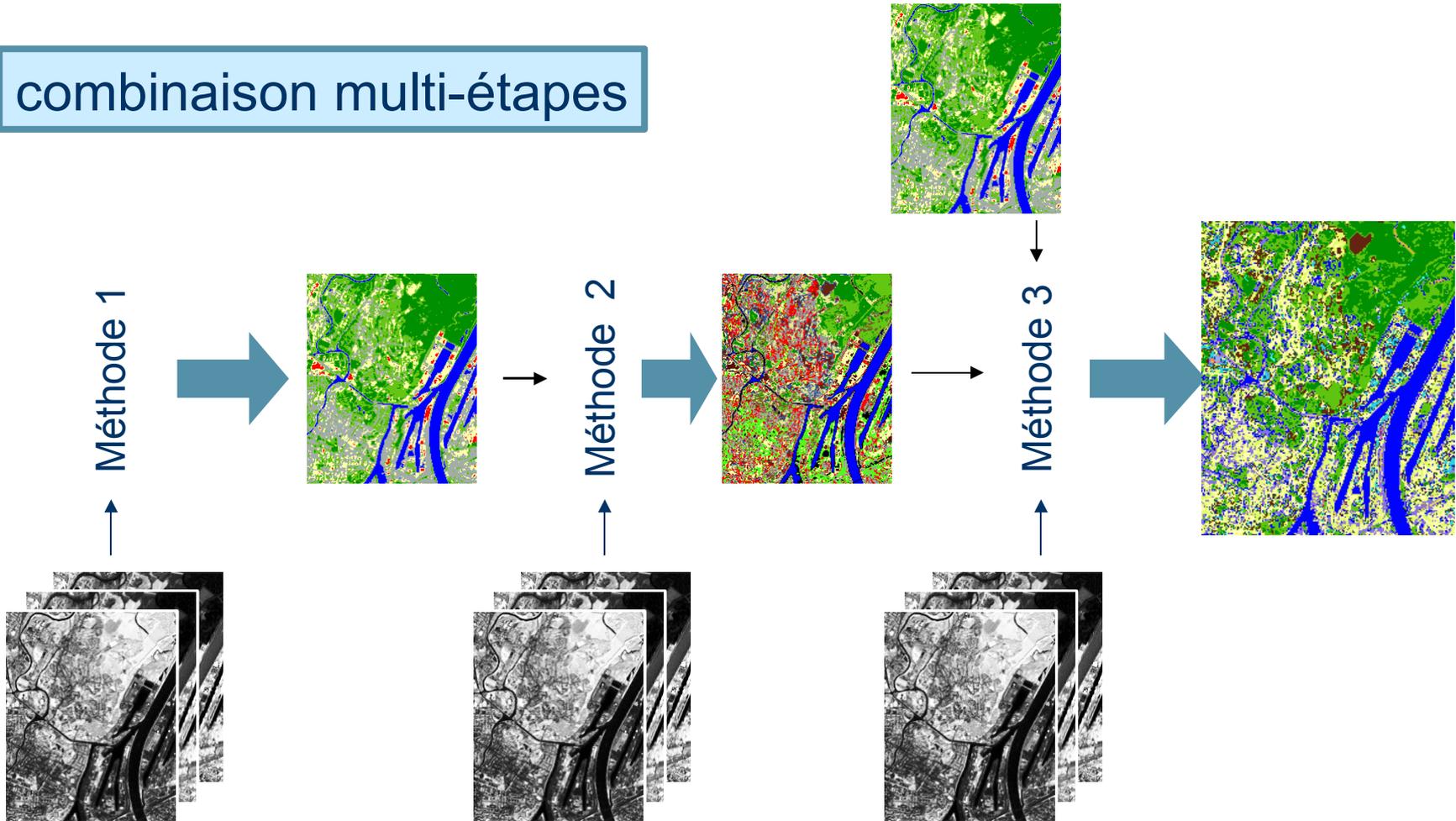
Notre proposition

- Offrir à l'utilisateur le meilleur de chacune des méthodes
 - L'utilisateur dispose de résultats mutuellement optimisés et d'un résultat unifiant
- Avec
 - un critère d'optimisation indépendant des méthodes utilisées : intervalle pour K , critère intra-classes
 - et utilisation possible de plusieurs sources de données :
 - De même nature (ex : images à même résolution)
 - Ou hétérogènes : « découpage » des données et affectation aux méthodes en fonction de leur nature

combinaison de classifieurs

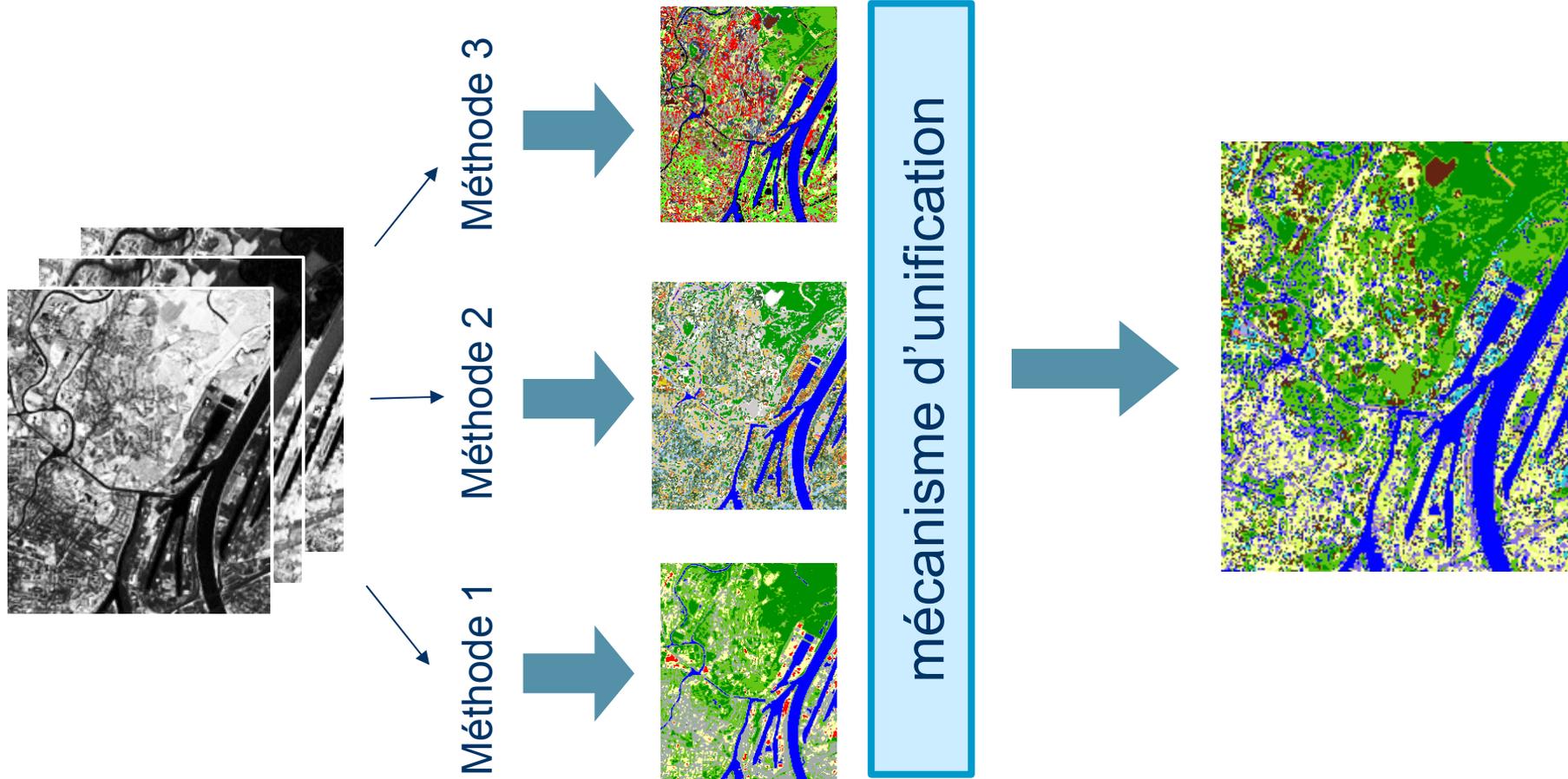
Première approche

combinaison multi-étapes



combinaison de classifieurs

Deuxième approche



combinaison multi-experts

combinaison de classifieurs

■ méthodologie employée

- *multi-experts* : classifications en parallèle puis unification des résultats
- *multi-étapes* : classifications en série

■ méthodes différentes

- *mono-stratégie* : plusieurs occurrences d'une même méthode sont utilisées
- *multi stratégies* : plusieurs occurrences de plusieurs méthodes sont utilisées

combinaison de classifieurs

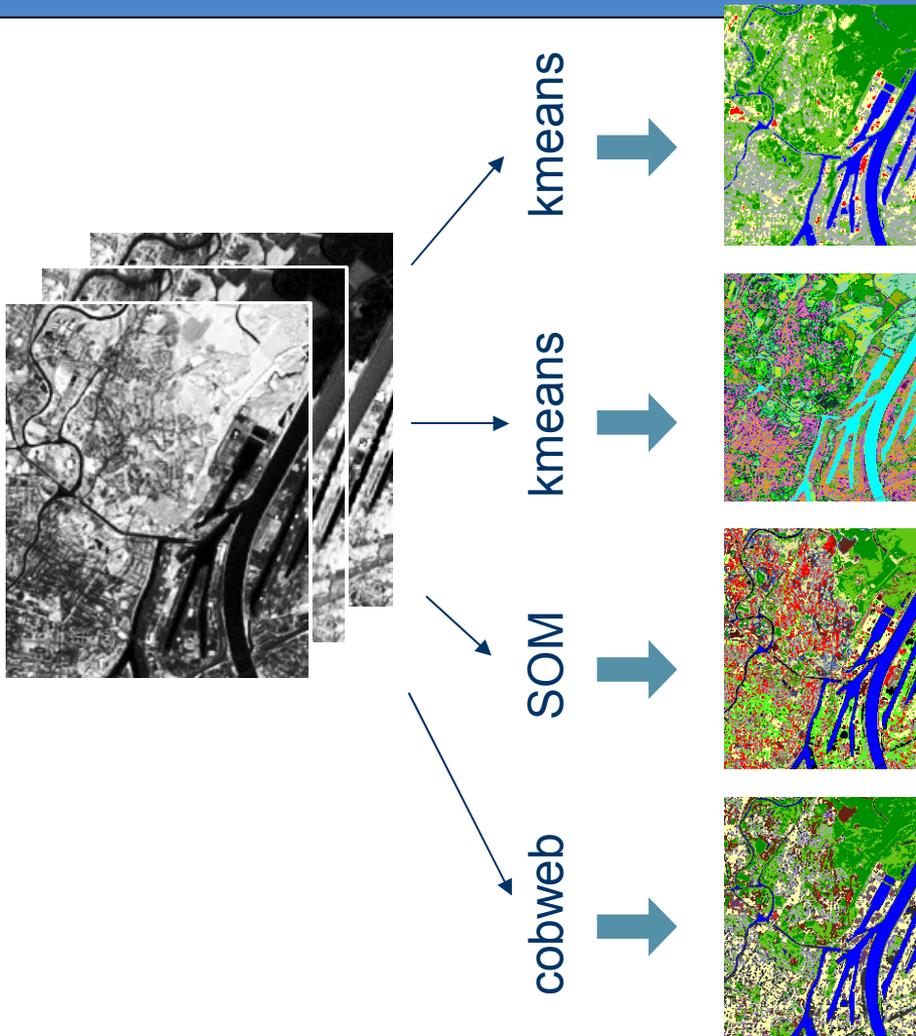
■ nature des données

- ***uni-représentationnelle*** : toutes les méthodes utilisent les mêmes données pour réaliser leurs classifications
- ***multi-représentationnelle*** : les méthodes utilisent des jeux de données différents
 - ex : télédétection (radiométrie, radar, photo, ...)

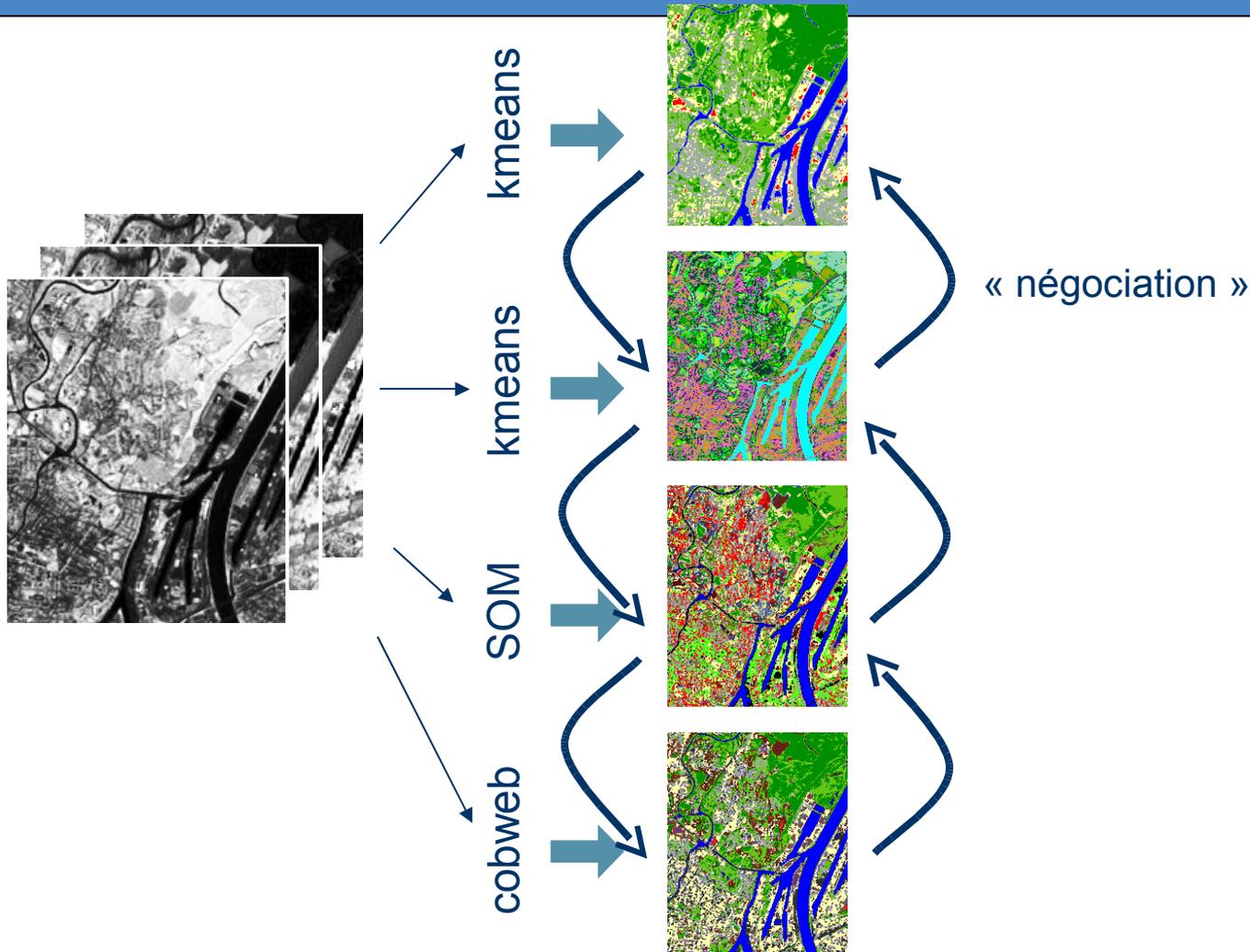
Classification non supervisée multi-stratégie

- Notre proposition : classification multi-experts multi stratégies multi-représentationnelle
- principe retenu : faire collaborer différentes méthodes de classification de natures différentes afin qu'elles proposent des résultats « similaires »
- **s**ystème d'**a**pprentissage **m**ulti-**a**gent par **r**affinement **a**utomatique de **h**iéarchies

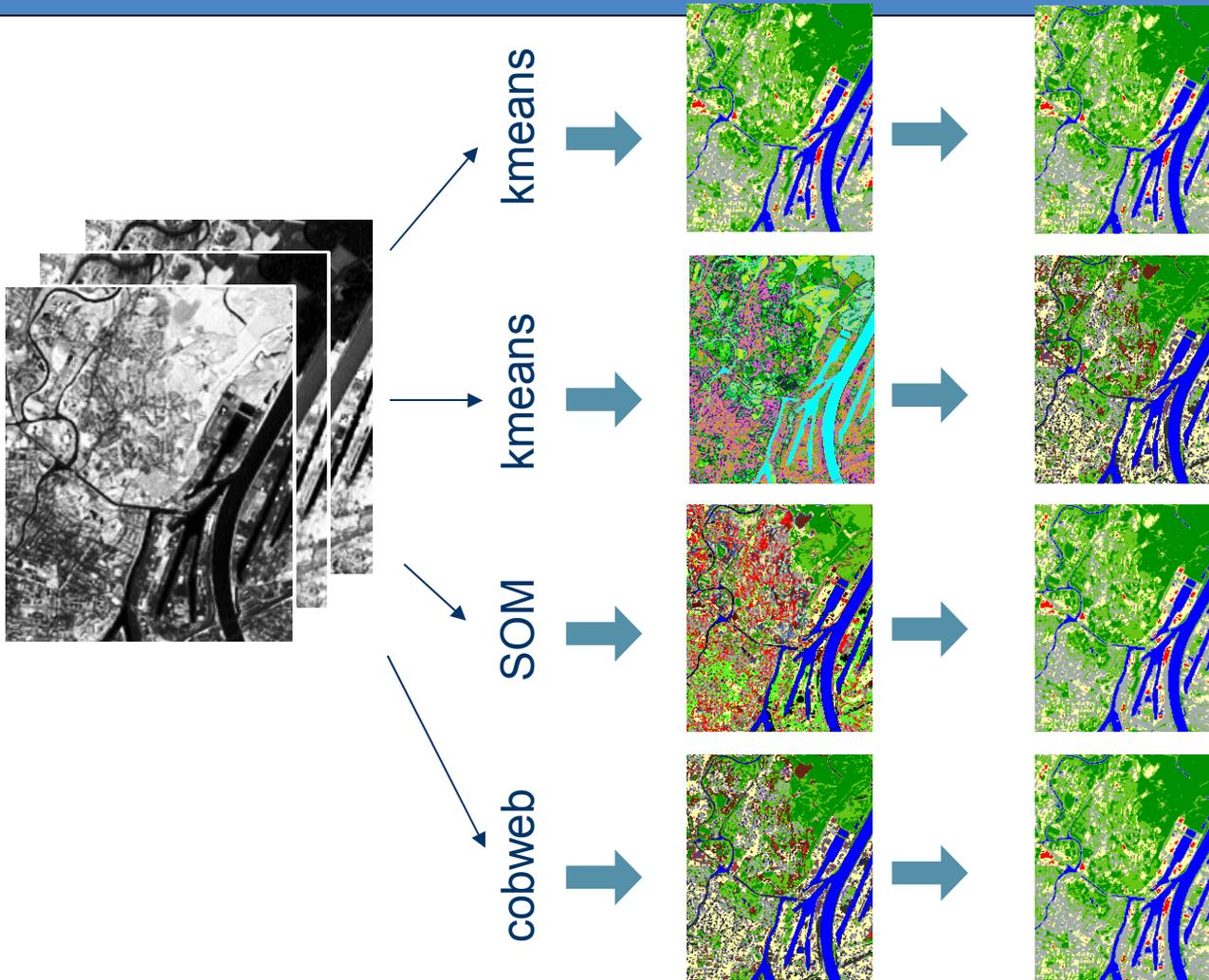
Classification non supervisée multi-stratégie



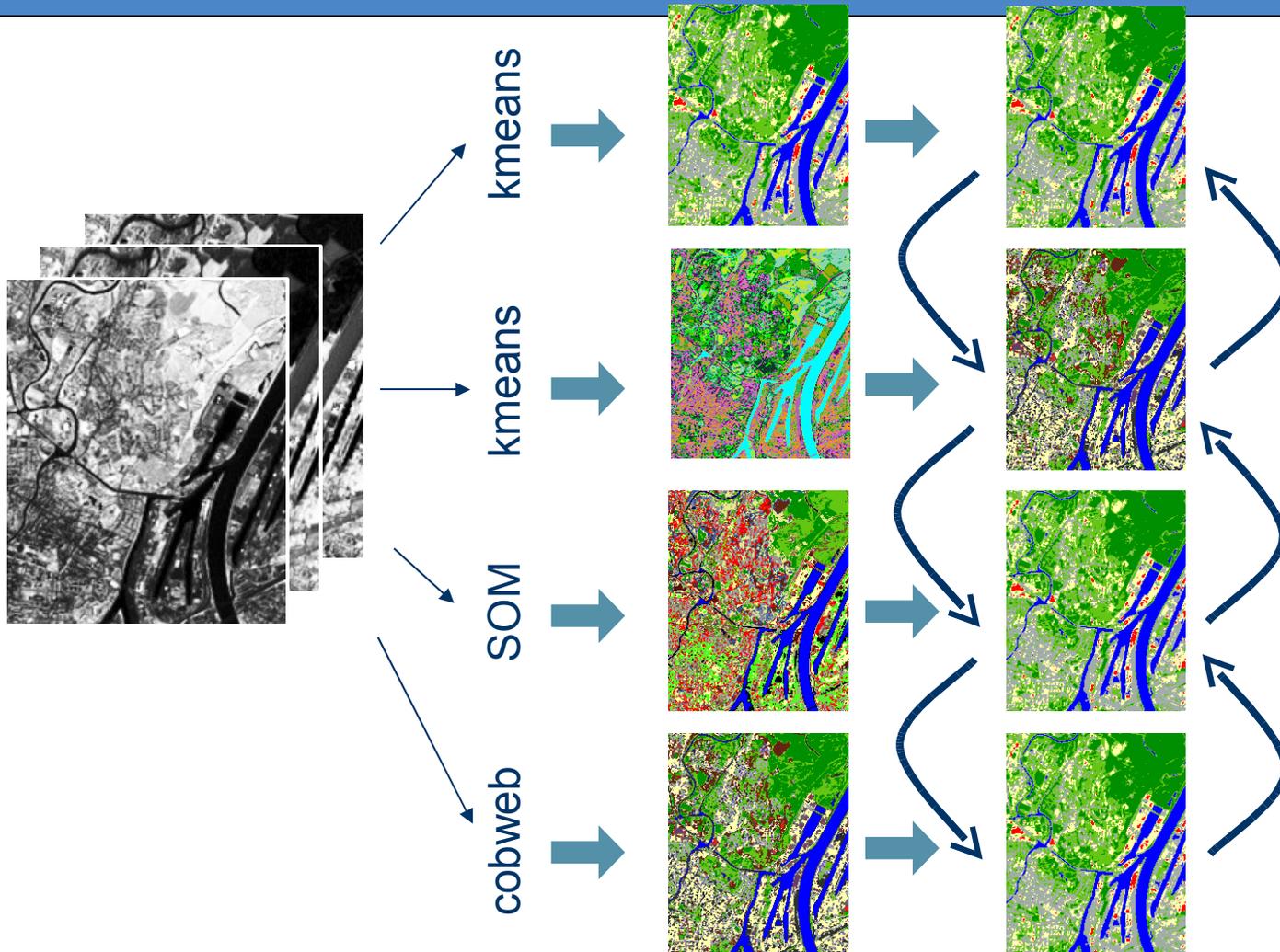
Classification non supervisée multi-stratégie



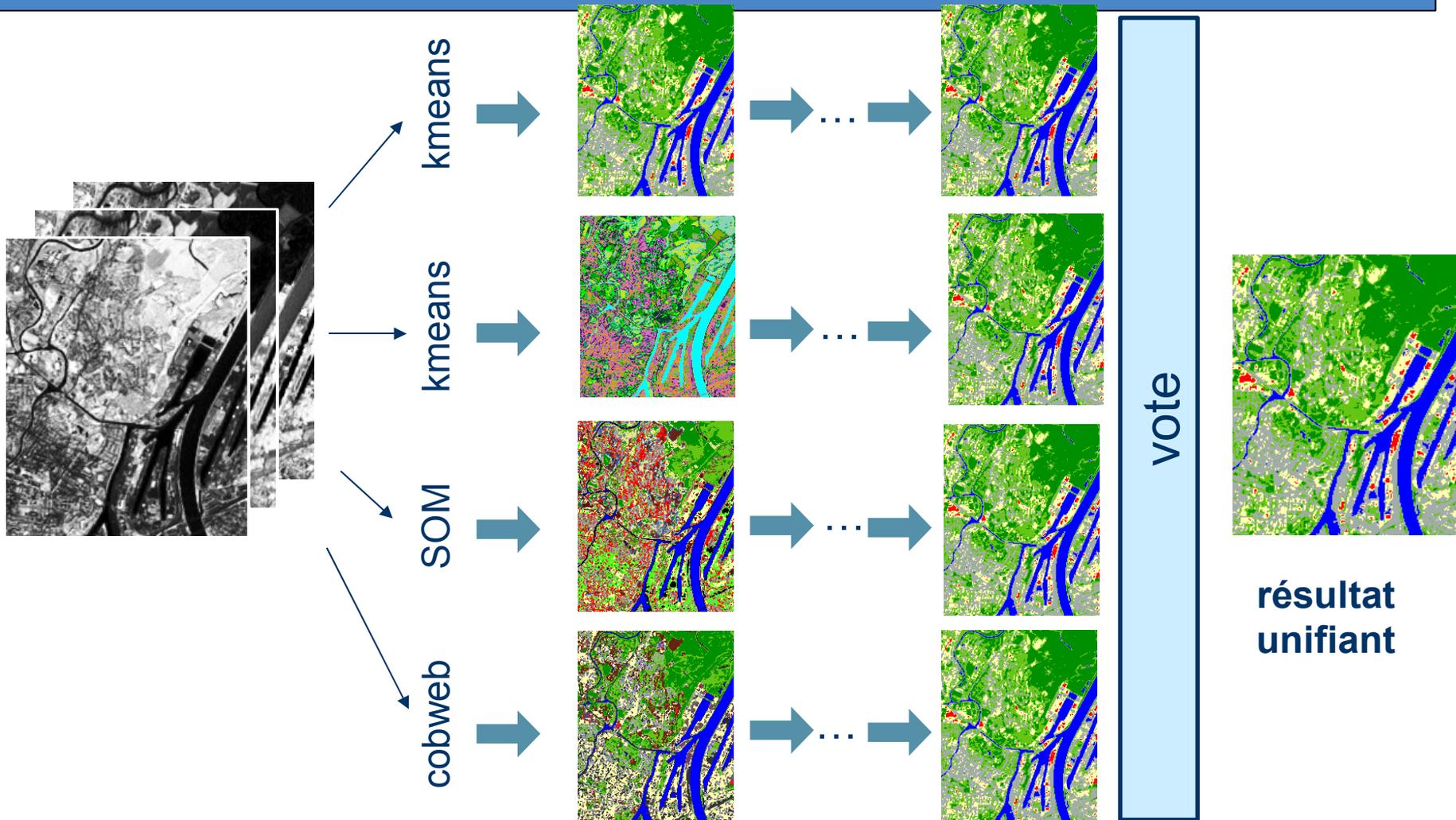
Classification non supervisée multi-stratégie



Classification non supervisée multi-stratégie



Classification non supervisée multi-stratégie



La méthode

- Trois phases
 - Classifications initiales
 - Raffinement
 - Unification

La méthode

- Trois phases
 - Classifications initiales
 - Raffinement :
 - Détection de conflits
 - Résolution locale des conflits
 - Prise en compte globale des résolutions locales
 - Unification

La méthode

■ Trois phases

- Classifications initiales

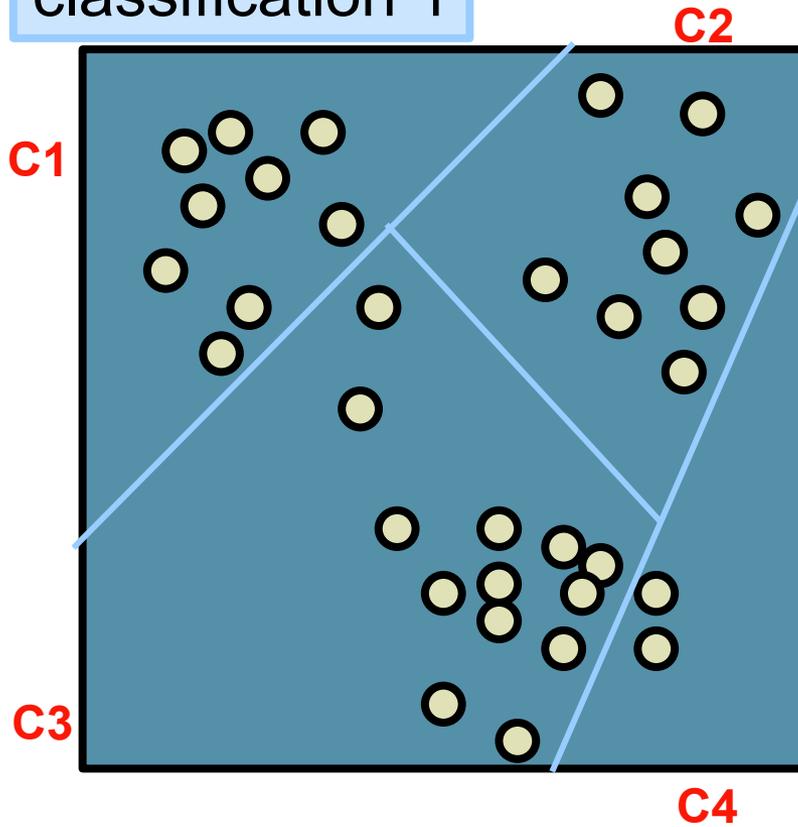
- Raffinement :

- Détection de conflits
- Résolution locale des conflits
- Prise en compte globale des résolutions locales

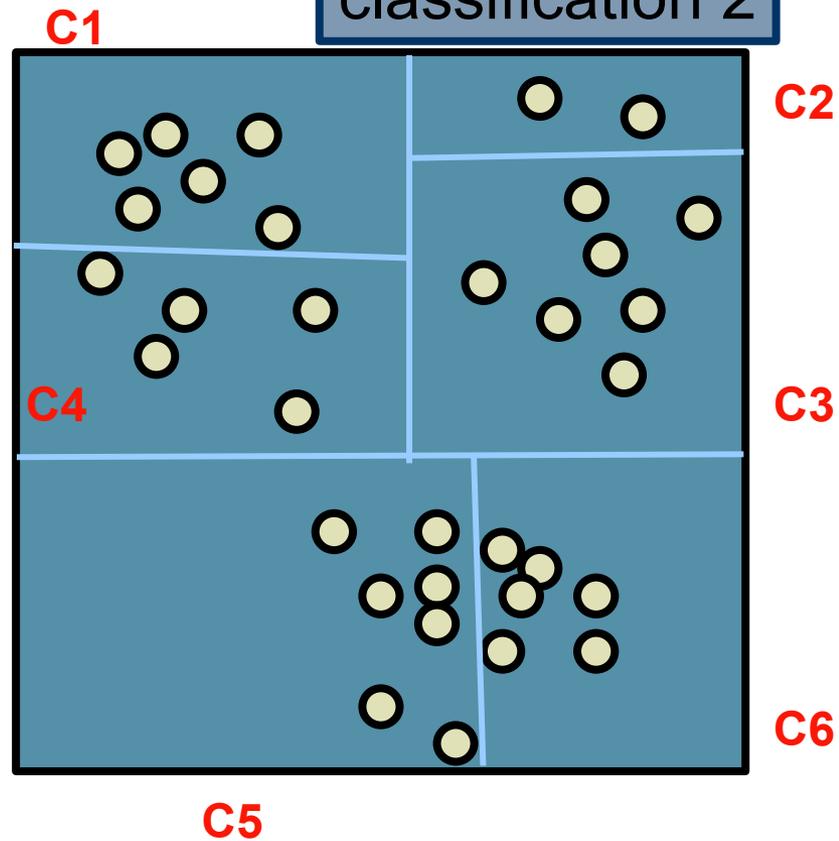
- Unification

Exemple

classification 1



classification 2



La méthode

■ Trois phases

- Classifications initiales
- Raffinement :
 - Détection de conflits :

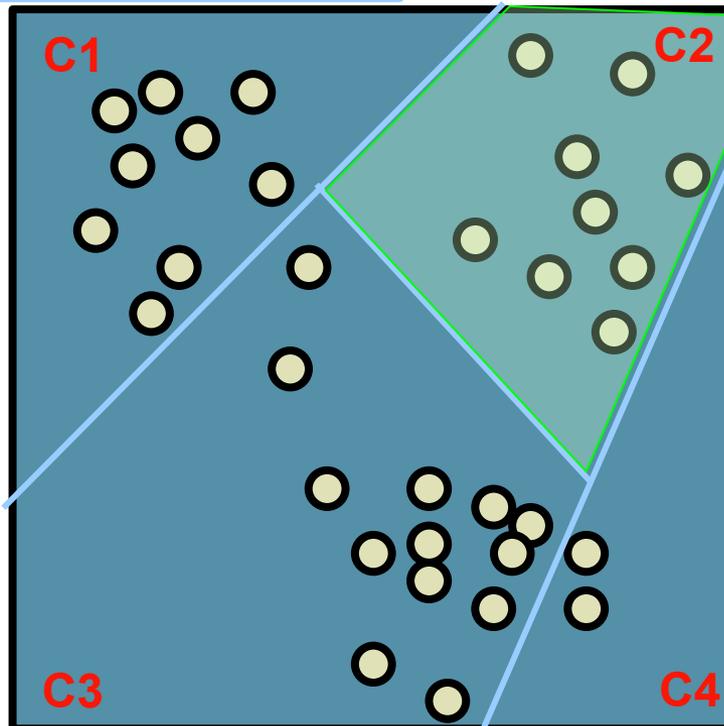
Taux conflit $T(C_k^i, C_l^j) = 1 - \omega_k^{i,j}$ avec $C_l^j = \text{corresp}_j(C_k^i)$

Similarité locale $\omega_k^{i,j} = \rho_k^{i,j} \alpha_{k_m,k}^{j,i}$ avec $\alpha_{k_m,k}^{j,i} = \max\{\alpha_{l,k}^{j,i} : \forall l\}$

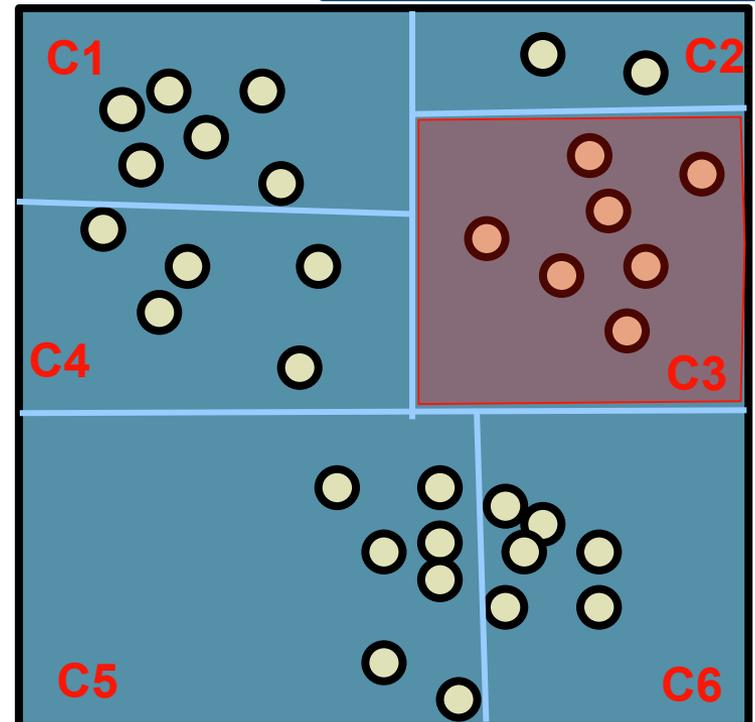
Coefficient de répartition local $\rho_k^{i,j} = \sum_{l=1}^{n_j} (\alpha_{k,l}^{i,j})^2$

La méthode

classification 1



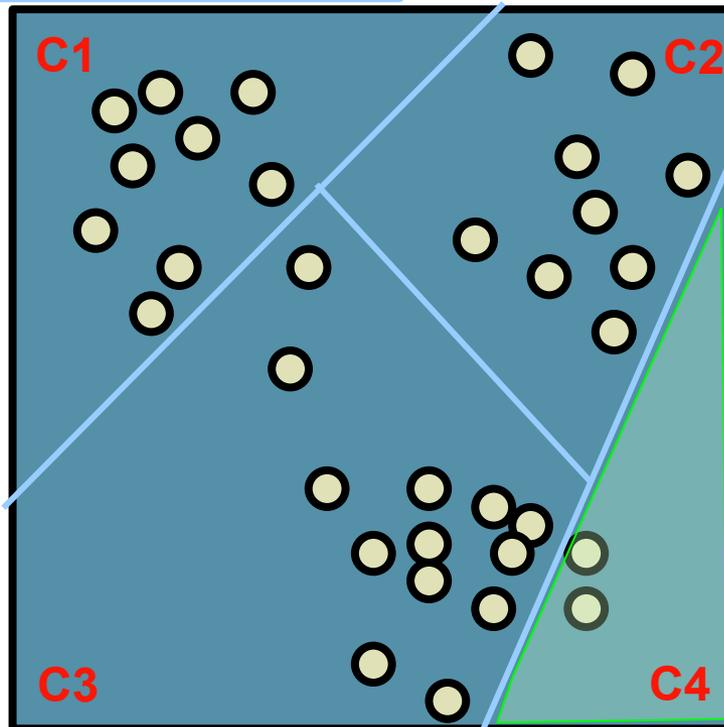
classification 2



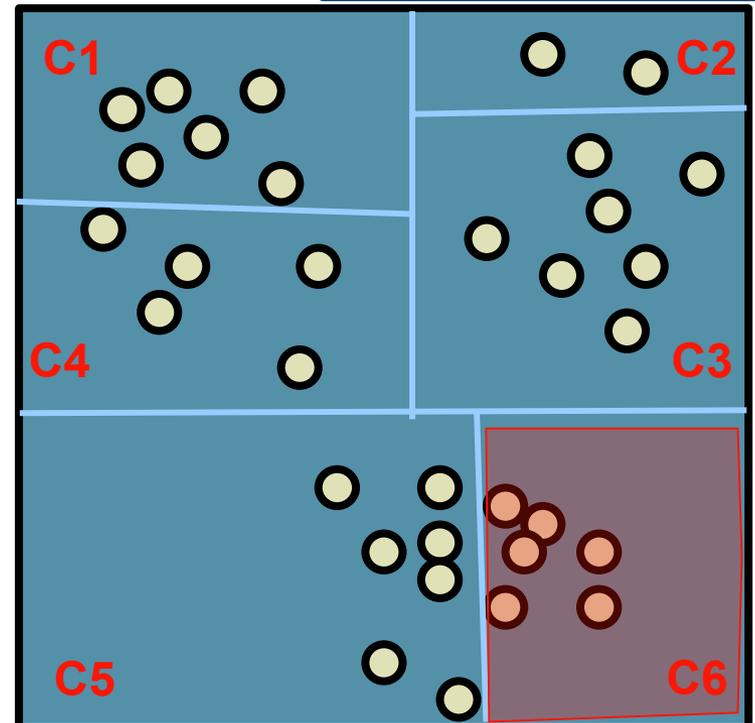
$\omega_3^{2,1} = 0.8$ C3 de R2 est très similaire à sa classe corresp. C2 de R1

La méthode

classification 1



classification 2



$\omega_6^{2,1} = 0.17$ C6 de R2 est peu similaire à sa classe corresp. C4 de R1

La méthode

- Trois phases
 - Classifications initiales
 - Raffinement :
 - Détection de conflits
 - Résolution locale des conflits :
 - scission d'une classe
 - fusion de plusieurs classes
 - suppression d'une classe et reclassement

La méthode

- Mais cette résolution n'a-t-elle pas dégradé la solution locale (effet « red queen ») ?

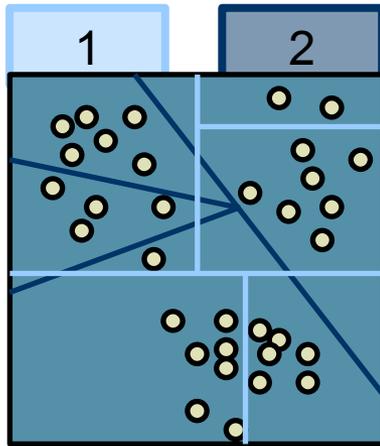
La méthode

■ Trois phases

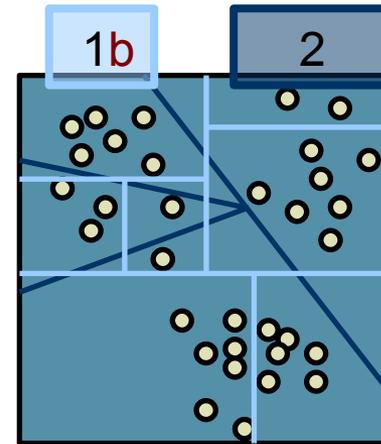
- Classifications initiales
- Raffinement :
 - Détection de conflits
 - Résolution locale des conflits ::
 - application d'opérateurs
 - choix des résultats :
 - qualité « interne » des résultats : δ
 - similitude locale des résultats : ω

$$\gamma^{i,j} = \frac{1}{2} \left(p_s \left(\frac{1}{n_i} \sum_{k=1}^{n_i} \omega_k^{i,j} + \frac{1}{n_j} \sum_{k=1}^{n_j} \omega_k^{j,i} \right) + p_q (\delta^i + \delta^j) \right) \text{ avec } p_s + p_q = 1$$

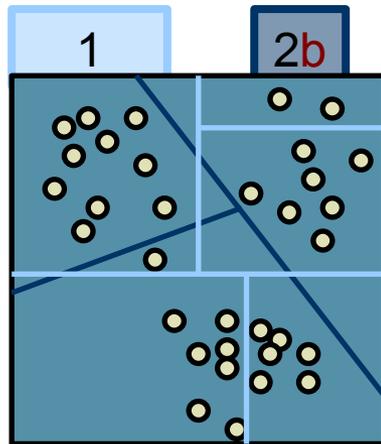
La méthode



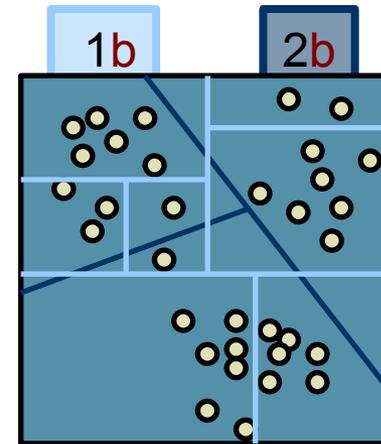
$$\gamma^{1,2} = 0,49$$



$$\gamma^{1b,2} = 0,51$$



$$\gamma^{1,2b} = 0,61$$



$$\gamma^{1b,2b} = 0,47$$

La méthode

- mais cette résolution ne dégrade-t-elle pas le résultat global ?
 - ➔ On ne garde que les résultats de résolution de conflit qui améliore le résultat global
- Deux aspect sà prendre en compte :
 - la convergence des résultats
 - la « qualité » du résultat

La méthode

- Mais cette résolution ne dégrade-t-elle pas la solution globale ?
 - On ne garde les résultats de résolution de conflit qui améliore le résultat global (en admettant des chutes temporaires de la qualité globale)
- Deux aspects à prendre en compte :
 - la convergence des résultats
 - la « qualité » du résultat

La méthode

■ Trois phases

- Classifications initiales
- Raffinement :
 - Détection de conflits
 - Résolution locale des conflits
 - Prise en compte globale des résolutions locales

$$\Gamma = c_{\Gamma} \cdot \Gamma_1 + c_q \cdot Q$$

Avec similitude globale

$$\Gamma_1 = \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{j=1, j \neq i}^M \gamma^{i,j}$$

La méthode

- Trois phases
 - Classifications initiales
 - Raffinement :
 - Détection de conflits
 - Résolution locale des conflits
 - Prise en compte globale des résolutions locales
 - Unification :
 - Une méthode de vote adaptée car il n'y a pas nécessairement bijection entre les classes des résultats

Classification non supervisée multi-stratégie

- Enfin des exemples

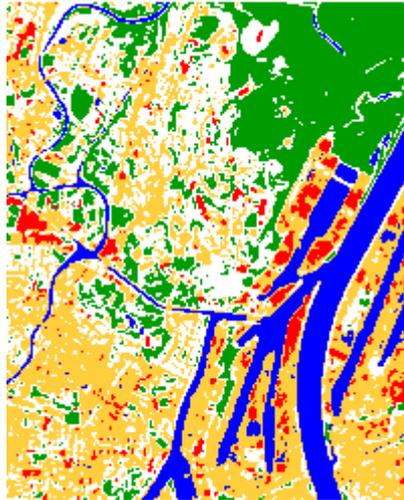
Exemple

Image SPOT : 200*250 / 3 bandes sur 1 octet



Exemple

$$\Gamma = 0,36$$



agent 1

- k-means
- 5 noyaux



- 5 classes

$$\Gamma^1 = 0,35$$



agent 2

- k-means
- 20 noyaux



- 20 classes

$$\Gamma^2 = 0,39$$



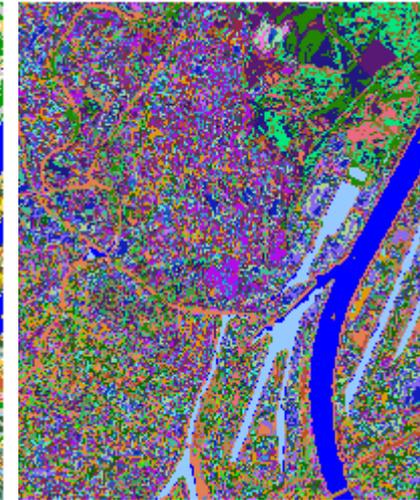
agent 3

- cobweb
- acuité 3,5



- 23 classes

$$\Gamma^3 = 0,38$$



agent 4

- cobweb
- acuité 1



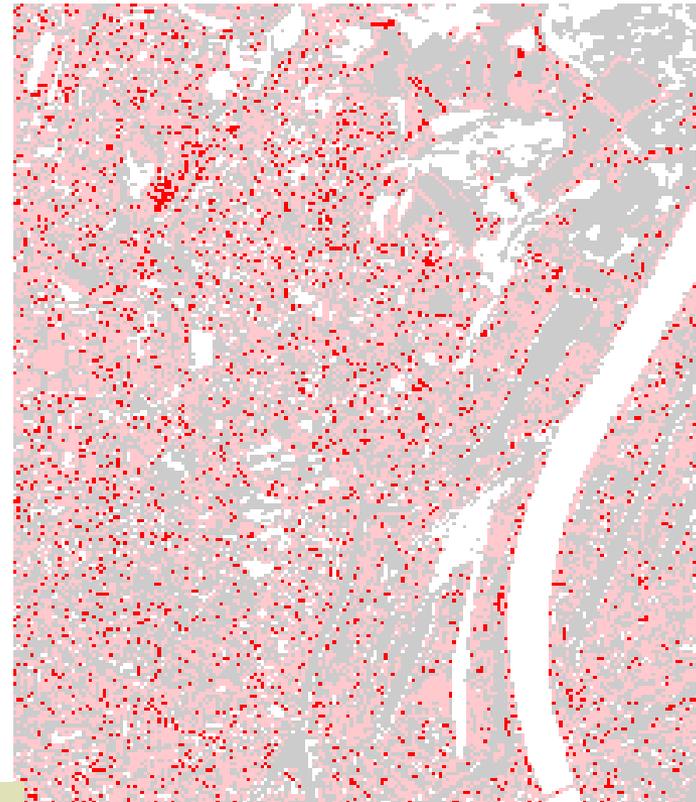
- 73 classes

$$\Gamma^4 = 0,33$$

Exemple



avant raffinement :
15 classes



Pixels non
consensuels

Exemple

$$\Gamma = 0,55$$



agent 1

- k-means
- 13 classes
- $\Gamma^1 = 0,56$

agent 2

- k-means
- 12 classes
- $\Gamma^2 = 0,58$

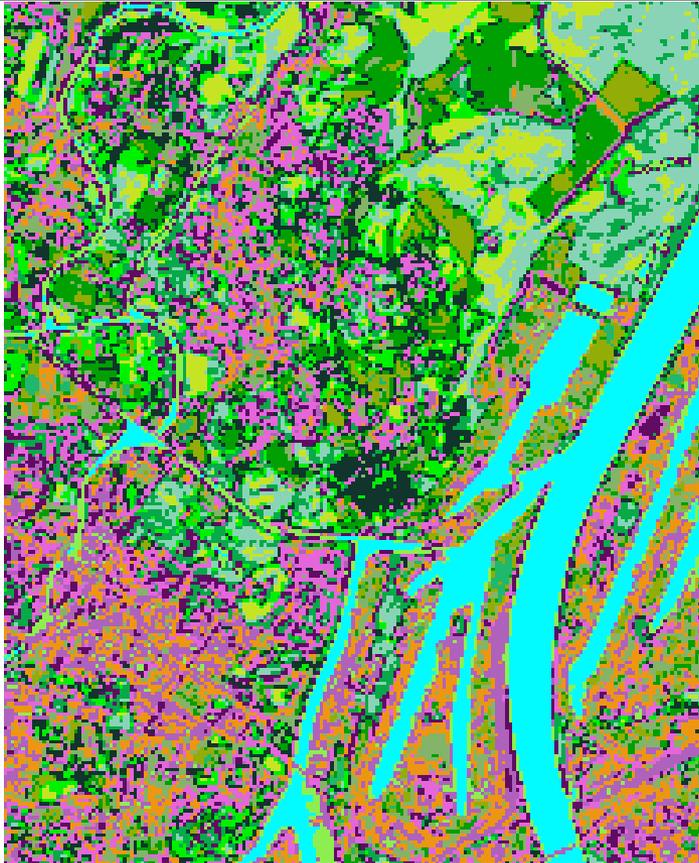
agent 3

- cobweb
- 6 classes
- $\Gamma^3 = 0,48$

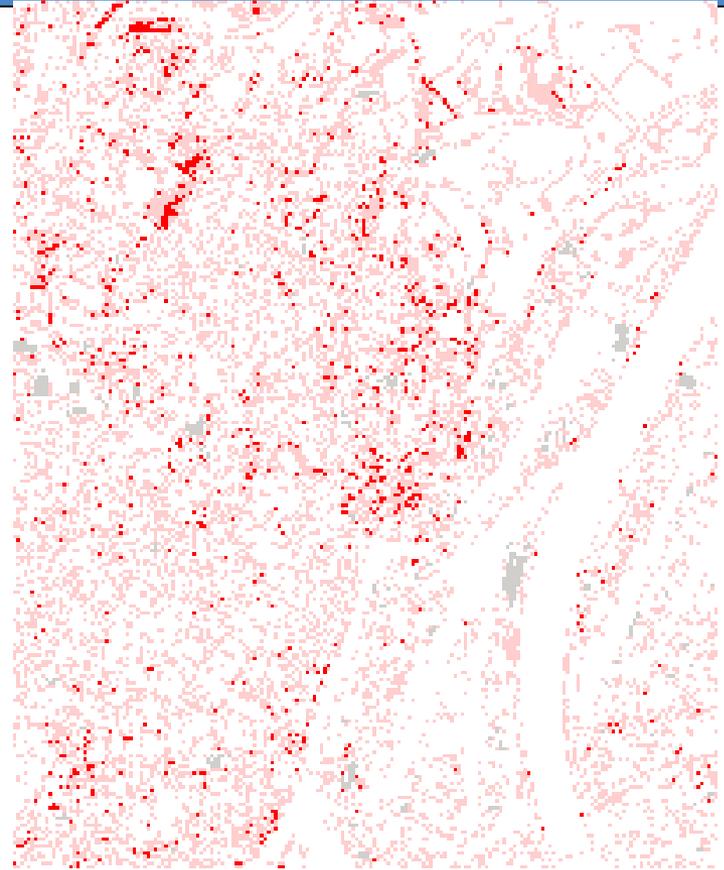
agent 4

- cobweb
- 12 classes
- $\Gamma^4 = 0,59$

Exemple



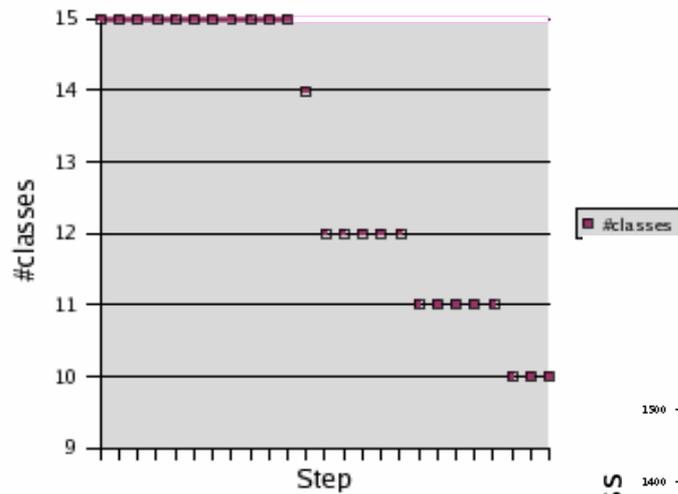
après raffinement :
10 classes



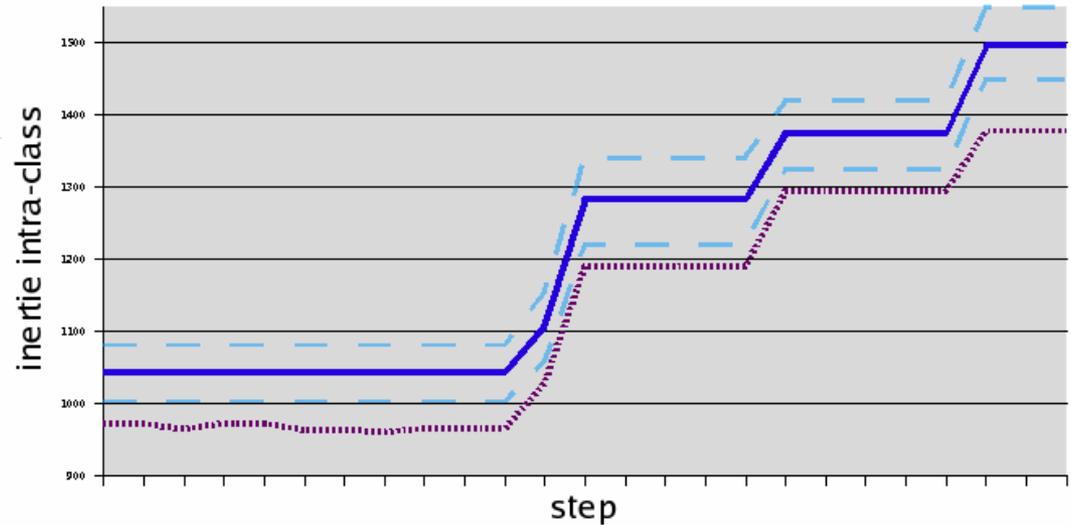
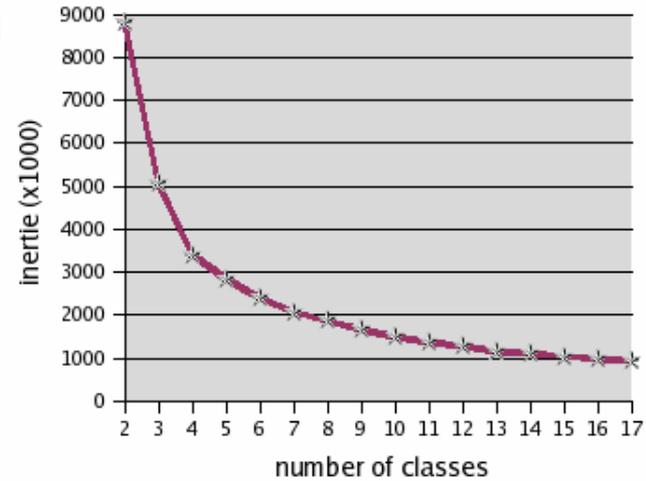
Pixels non
consensuels

Exemple

Classes number evolution

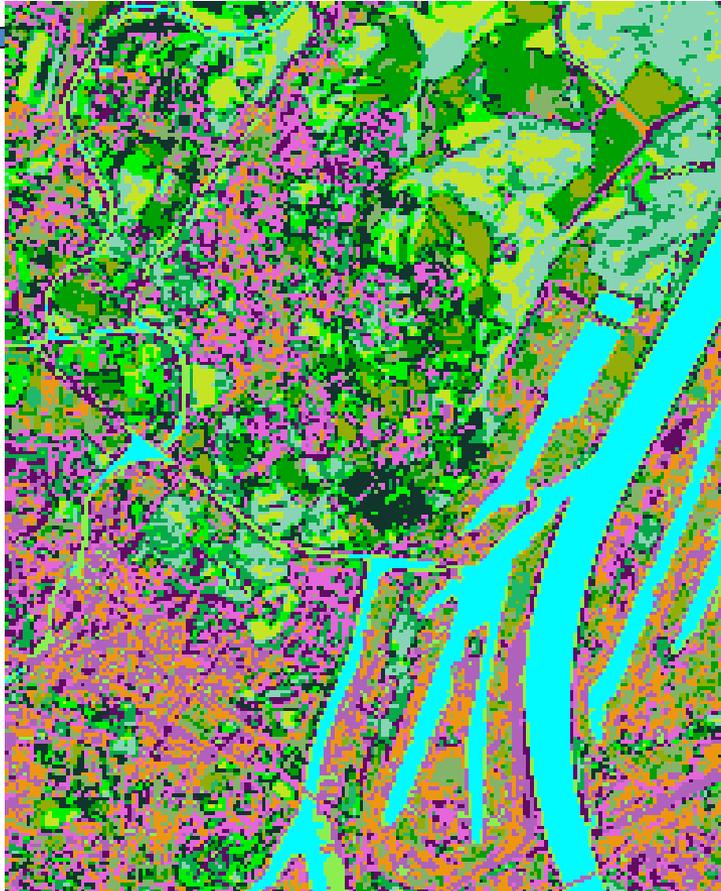


Inertie / number of classes



● Inertie — Empirical average - - -2/3 ecart-type - - +2/3 ecart-type

Exemple



classification initiale



régions construites

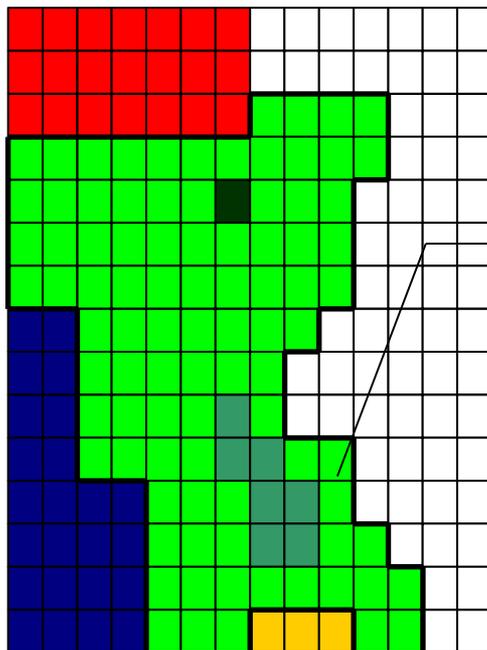
Exemple

- Fusion/agrégation de régions

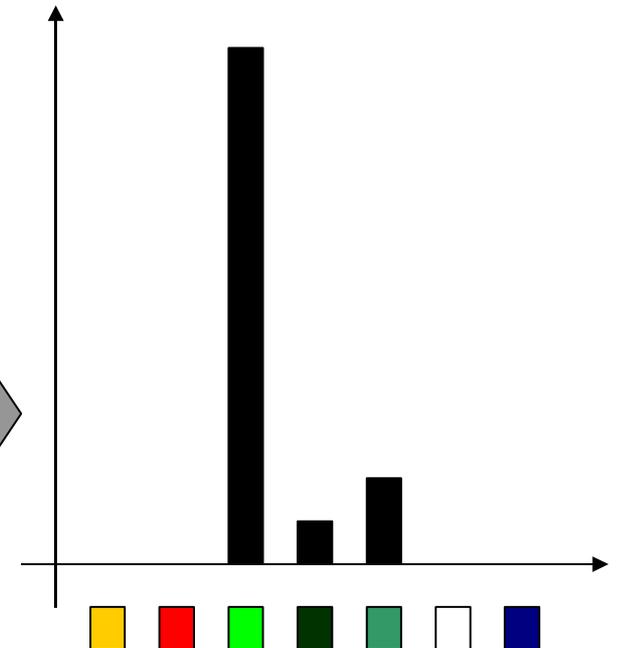
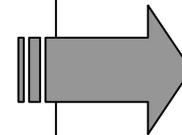
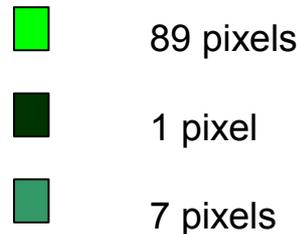


Exemple

- On construit l'histogramme des régions incluses

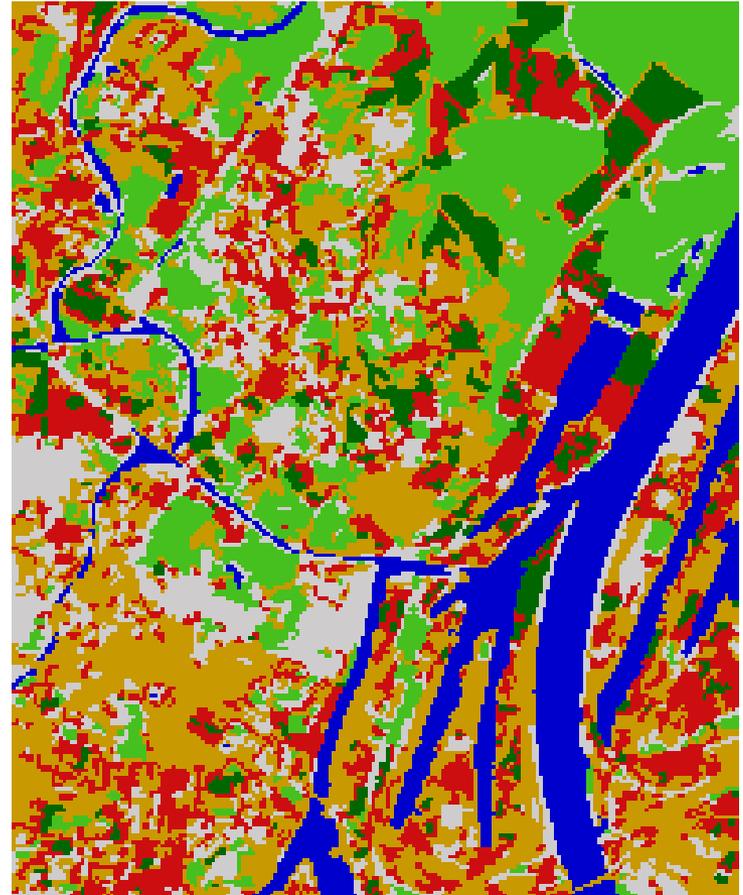


Région verte :
composée de 3
classes différentes



Exemple

- on ré-applique une classification multi-stratégie



Exemple

Classification initiale



Classification des régions



Exemple

Classification initiale



Classification de régions



Exemple

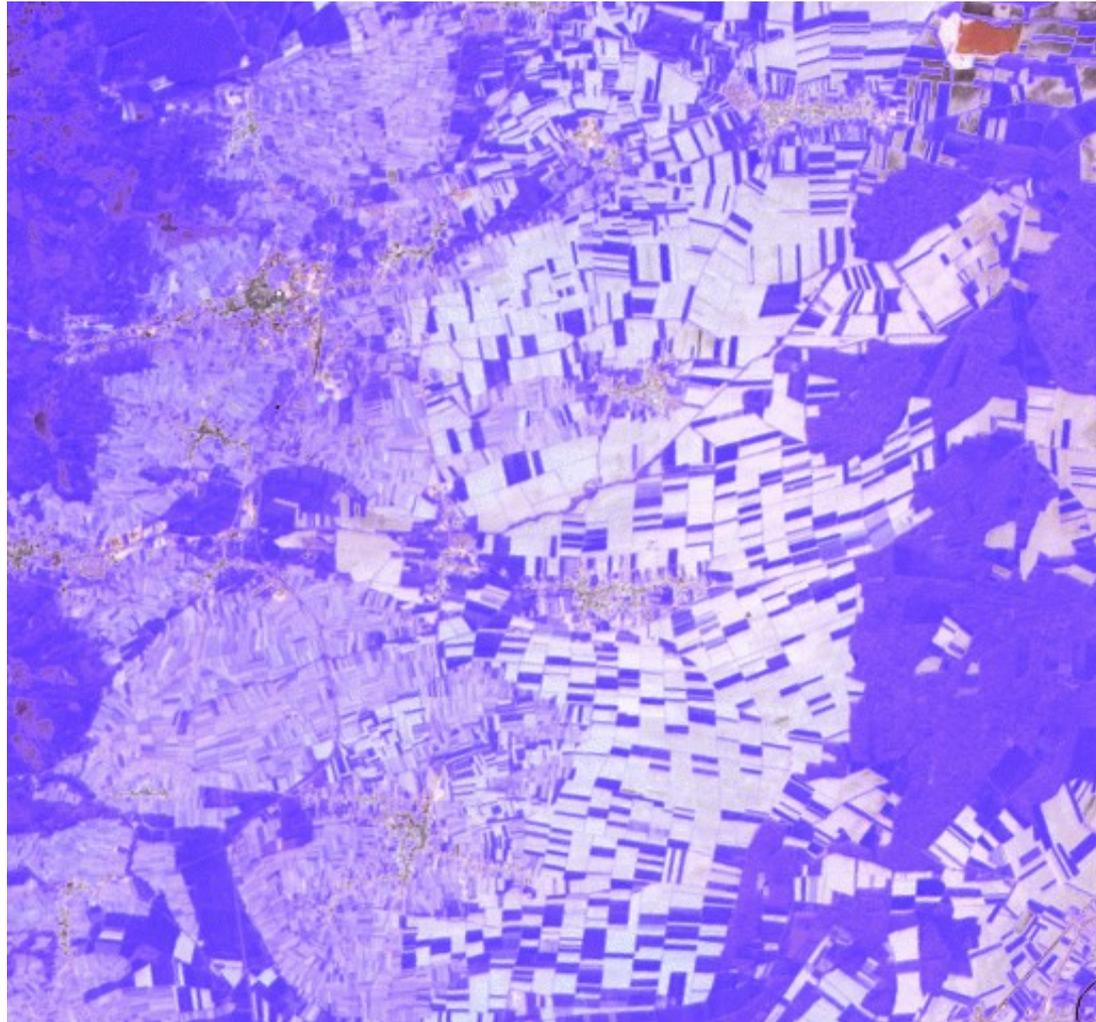
Classification initiale



Classification de régions

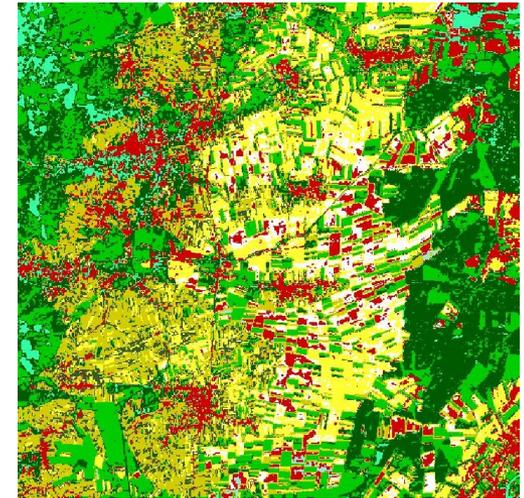
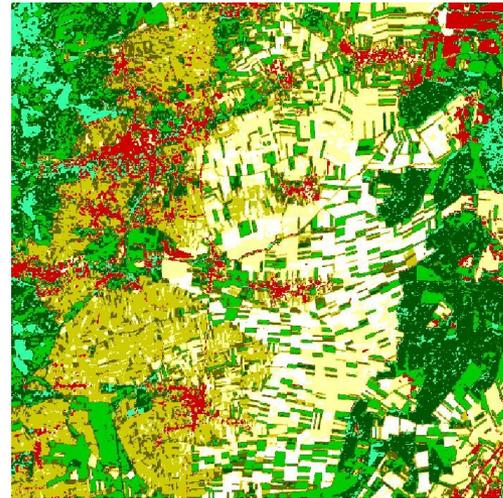
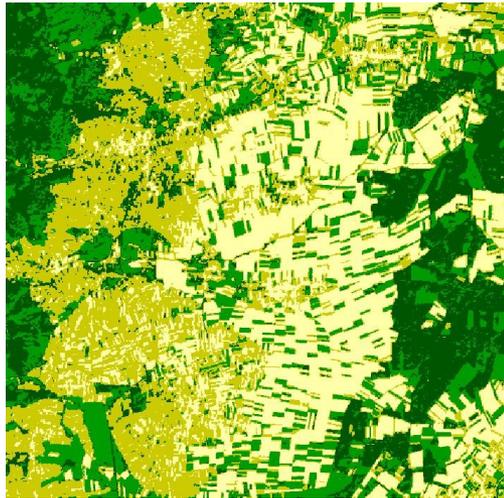


Exemple



Exemple

$$\Gamma = 0,31$$



agent 1

- kohonen
- 4 neurones



- 4 classes
- $\Gamma^1 = 0,35$

agent 2

- k-means
- 8 noyaux



- 8 classes
- $\Gamma^2 = 0,30$

agent 3

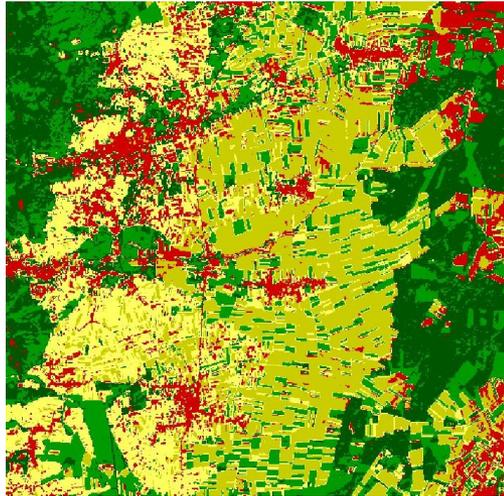
- cobweb
- acuité 8



- 11 classes
- $\Gamma^3 = 0,28$

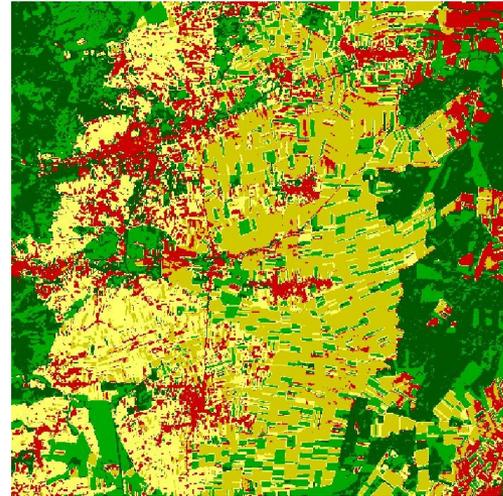
Exemple

$$\Gamma = 0,68$$



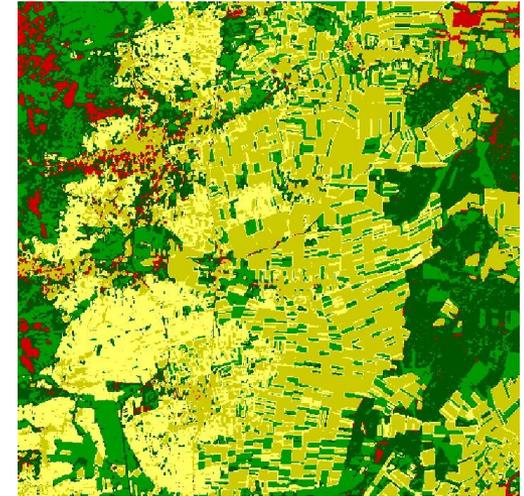
agent 1

- kohonen
- 5 classes
- $\Gamma^1 = 0,73$



agent 2

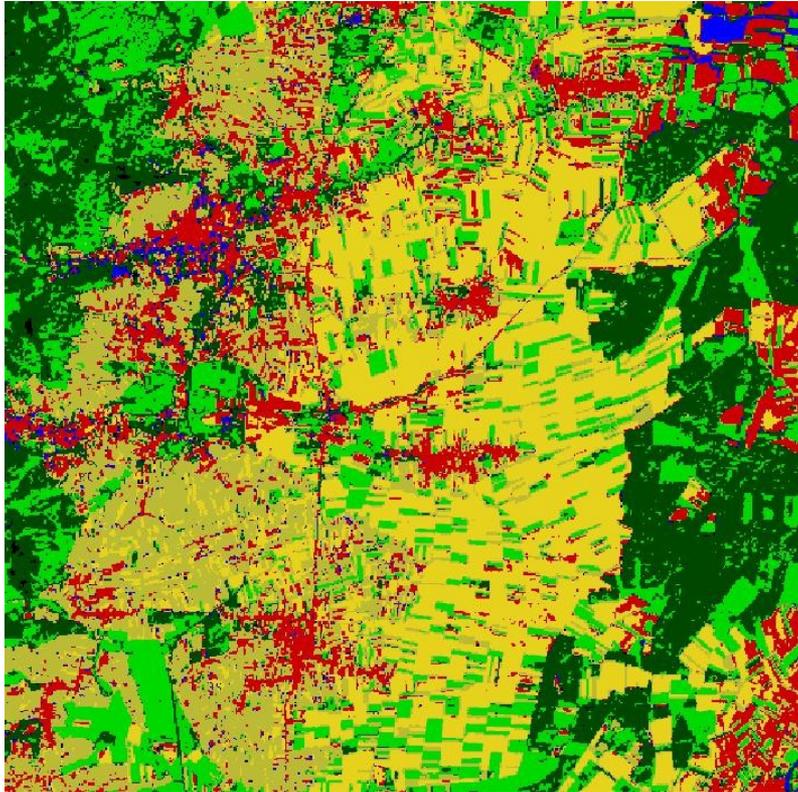
- k-means
- 5 classes
- $\Gamma^2 = 0,72$



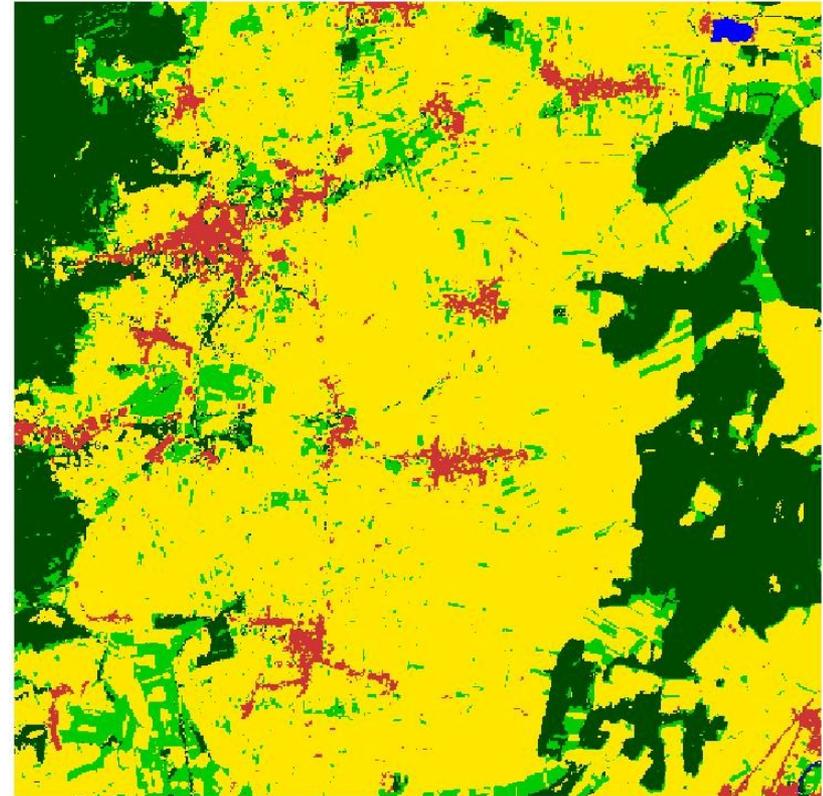
agent 3

- cobweb
- 5 classes
- $\Gamma^3 = 0,60$

Exemple



classification automatique
6 classes



 Cultures  Bâtis  Forêt  Prairies  Eau

classification par l'expert

Conclusion

- Méthode de collaboration entre classifieurs
 - Stabilité des résultats
 - Convergence en fonction d'un critère indépendant de la méthode et de ses paramètres
 - Résultat unifiant avec un « coefficient de confiance »
 - mais
 - Coût CPU
 - Ratio entre :
 - Qualité interne des classes
 - Similitude entre les résultats
 - Qualité globale du résultats
- à déterminer en fonction de l'application

Perspectives

- Intégration d'un critère de qualité dépendant du domaine d'application et des connaissances a priori disponibles
 - Introduction de la sémantique dans le processus de classification
- Au niveau « image »
 - Classification multi-image :
 - plusieurs représentation pour un objet
 - Classification multi-résolution :
 - comparaison entre des classification sur des objets différents mais représentant une même réalité

Classification non supervisée multi-stratégie

- merci
- c'est fini, vous pouvez reprendre une activité normale ...