

# APPORT DE L'IMAGERIE PLEIADES POUR LA CARACTERISATION DE LA DENSITE URBAINE

Jacques Bouffier<sup>1</sup>, Dominique Hebrard<sup>1</sup>, Benoit Mingam<sup>1</sup>, Lucie Chadourne-Facon<sup>1</sup>, Didier Treinsoutrot<sup>1</sup>

1 : Cerema, Toulouse,

jacques.bouffier@cerema.fr, dominique.hebrard@cerema.fr, benoit.mingam@cr-mip.fr,

lucie.chadourne-facon@cerema.fr, didier.treinsoutrot@cerema.fr

## Résumé

Les récentes évolutions réglementaires imposent aux services de l'Etat et aux collectivités de mener une politique d'aménagement urbain combinant construction de logements et gestion économe de l'espace. Le portage de ces politiques sensibles nécessite une connaissance fine, fiable, consensuelle et actualisée du territoire. Les bases de données actuellement disponibles, qui présentent par ailleurs des atouts reconnus, montrent cependant certaines limites bloquantes en termes d'homogénéité ou de fréquence de mise à jour. L'objectif de cette étude est d'identifier dans quelle mesure des traitements automatiques sur images satellites à très haute résolution spatiale permettent d'aller au-delà de ces limites et peuvent fournir des informations d'occupation des sols et de densité urbaine. Des couches d'occupation des sols sur l'agglomération toulousaine sont obtenues par classifications supervisées sur images Pléiades. Les surfaces de bâti sont ensuite agglomérées à l'ilot urbain pour définir des indicateurs liés aux documents de planification.

**Mots-clés** : densité urbaine, classifications supervisées, image satellite Pléiades, très haute résolution spatiale, Toulouse

## Abstract

*The French regulations set out the objective of building "sustainable cities" through the redensification of urban centers and the control of urban sprawl. These directives require the introduction of quantitative indicators in planning documents, like the COS (Coefficient of occupation) that determines the regulatory urban density for each area of the cities. In order to determine these mandatory indicators, methodologies are mainly based on the exploitation of existing databases. But currently, obtaining the exhaustive, reliable, consistent, homogeneous and recent mapping of the urban density is a major challenge. The purpose of this study is to identify in which way automatic treatments on VHRS (Very High Spatial Resolution) satellite images can determinate the density information. As a first step, covering layers from a Pleiades image are produced on the Toulouse agglomeration with SVM supervised classifications. The second step is the calculation of density indicators at the adapted scale. The final data can then be linked to planning documents.*

**Keywords:** urban density, supervised classification, Pleiades satellite image, very high spatial resolution, Toulouse

## 1. Introduction

### L'aménagement du territoire urbain

Depuis 2001, la loi SRU<sup>1</sup> impose aux services de l'État et aux collectivités de conduire une politique d'aménagement ambitieuse combinant développement urbain et gestion économe de l'espace dans le double objectif de résorber la crise du logement et de préserver les espaces péri-urbains. Il en résulte une politique de densification des cœurs de ville et de reconquête des dents creuses et des friches urbaines. Les zones en sous densité, qui pourraient faire l'objet d'une mutation de leur forme architecturale, doivent donc être identifiées. Dans la même logique, les niveaux réels de densité de bâti, sous forme de COS (Coefficient d'Occupation des Sols) ou CES (Coefficients d'Emprise au Sol), doivent être reliés à ceux autorisés dans les documents d'urbanisme opposables.

Le cadre législatif et réglementaire des documents de planification a depuis renforcé progressivement les obligations de bilan et d'évaluation des politiques publiques. En particulier, les lois ENL<sup>2</sup> (2006), LMAP<sup>3</sup>

(2010), et Grenelle 2 (2010), imposent désormais une analyse fine de la consommation des espaces naturels, agricoles et forestiers au cours des dix années précédant l'approbation d'un document. Elles prescrivent également la mise en place d'objectifs chiffrés de consommation économe d'espace et un suivi fin des espaces péri-urbains, particulièrement sensibles à l'étalement urbain.

Par ailleurs, depuis 2011, les collectivités compétentes en gestion des eaux pluviales urbaines peuvent mettre en place une taxe sur les surfaces imperméabilisées dans le cadre de la police de l'eau. La détermination des règles d'impositions adaptées et l'estimation *a priori* du potentiel de recette fiscale nécessitent une connaissance fine de ces surfaces.

Ces quelques exemples sont à croiser avec d'autres, notamment dans le domaine de la prévention des risques en secteur sensible, des transports ou des trames vertes et bleues en milieu urbain (SNDD, 2010).

D'une manière générale, les acteurs de l'aménagement doivent disposer, à différentes échelles et sur des niveaux dégressifs de précision, de données urbaines récentes, fiables, homogènes et consensuelles comme

<sup>1</sup> loi n°2000-1208 du 13 déc. 2000 relative à la solidarité et au renouvellement urbain

<sup>2</sup> loi portant engagement national pour le logement

<sup>3</sup> loi de modernisation de l'agriculture et de la pêche

outil d'aide à la décision et de portage de politiques publiques sensibles.

Pour cela, les services institutionnels mobilisent à ce jour les bases de données à leur disposition, comme les fichiers fonciers de la DGFIP<sup>4</sup>, les bases de données de référence de l'IGN ou des enquêtes terrain. Ces données, de qualité reconnue, montrent cependant des limites en termes d'homogénéité ou de fréquence de mise à jour au regard d'une obligation de suivi exhaustif du territoire tout les deux à trois ans (CERTU, 2011).

### Objectif de l'étude

L'observation de la terre par satellite est aujourd'hui dans une phase de maturité nouvelle, tant en termes de performances des satellites qu'en termes d'accès à la donnée et d'outils de traitements. Les programmes actuels permettent ainsi d'obtenir régulièrement des informations précises sur de larges zones.

Les travaux de cartographies automatisées du milieu urbain à partir d'images d'images satellites très haute résolution font l'objet de nombreux travaux mettant en œuvre des méthodes d'apprentissage supervisé (Bauer *et al.*, 2001, Faure *et al.*, 2004), parfois croisées avec des découpages vectoriels (Holland *et al.*, 2006), des approches orienté objet (Blaschke *et al.*, 2004, Walker *et al.*, 2008) ou des méthodes plus complexes de segmentations multi-échelle (Jacquin *et al.*, 2010).

L'objectif de l'étude est d'identifier comment cette approche peut être transposée à des images Pléiades pour répondre de manière opérationnelle aux besoins des aménageurs urbains sur la ville de Toulouse.

## 2. Approche méthodologique

Une fois l'image satellite adaptée identifiée et obtenue, la méthodologie consiste à en extraire une information labellisée caractérisant l'occupation des sols. Cette information brute est ensuite post-traitée et agglomérée afin d'obtenir des indicateurs de densité urbaine directement utilisables par l'aménageur du territoire.

### Données

Le territoire de démonstration choisi est l'agglomération toulousaine du fait des enjeux élevés de cette zone : le double phénomène d'extension et de densification relative est particulièrement marqué, et s'accompagne de l'émergence de nouveaux pôles relais dans la couronne périurbaine (AUAT, 2012). L'objet de l'étude et la précision des informations recherchées nécessitent de disposer d'images submétriques. L'image retenue est une image Pléiades fusionnée (résolution spatiale native de 70 cm), orthorectifiée à 50 cm et fournie en 12 bits dans les quatre canaux (rouge, vert, bleu, proche infrarouge). Sa superficie est d'environ 200 km<sup>2</sup>. L'angle de tangage est de 5°, l'angle de roulis de 25°, ce qui devra être pris en compte ultérieurement. Elle a été fournie dans le cadre de la Recette Thématique Utilisateur Pléiades du CNES et date du 7 mai 2012 (Figure 1).

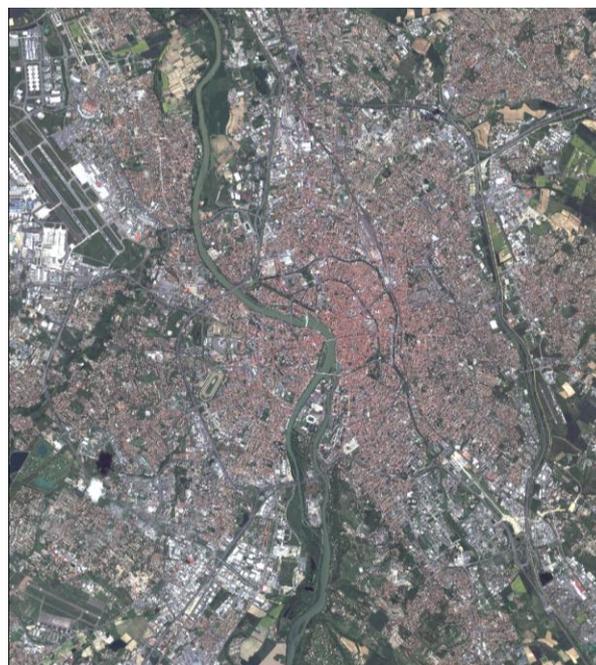


Figure 1 : Visualisation de l'image Pléiades sur la zone d'étude (agglomération toulousaine).

### Méthode de caractérisation de la couverture du sol

Afin d'envisager une mise à jour rapide sur de vastes territoires, l'extraction d'informations se fait par traitements semi-automatiques mettant en œuvre des classifications supervisées. La méthodologie décrite ici est automatisée par des scripts python utilisant les logiciels libres OrfeoToolBox, QGis, Grass et SpatiaLite.

La caractérisation de la couverture du sol se fait en trois étapes. Dans un premier temps, les informations d'entrée (image satellite et échantillons d'apprentissage) sont pré-traitées et enrichies. La seconde étape correspond à la classification proprement dite. Enfin, des post-traitements améliorent la qualité finale du résultat de classification.

Pour exploiter au mieux les informations contenues dans l'image lors d'une classification pixel, la donnée radiométrique initiale (rouges, vert, bleu et proche infrarouge) est enrichie par des informations texturales (textures Haralick principalement), indicelles (NDVI, indice de bâti ISU...), des segmentations ou des informations de sursol. Le choix final des « néo-canaux » retenus est arrêté à l'issue d'une étude paramétrique décrite plus tard dans cet article.

Les échantillons d'apprentissage, obtenus ici par photo-interprétation, sont des polygones labellisés en cohérence sémantique et géométrique avec l'image satellite à classifier. La nomenclature retenue pour l'étude est une nomenclature simple en quatre classes (surfaces bâties, surfaces en eaux, axes de communications et végétation) afin de limiter au maximum les confusions. L'objectif étant la détermination des seules surfaces bâties, une nomenclature plus complète n'est pas nécessaire.

<sup>4</sup> Direction Générale des Finances Publiques

La qualité du résultat de classification dépend directement de « l'homogénéité radiométrique » des classes utilisées. Une classe radiométriquement homogène, finement caractérisée lors de l'apprentissage, entrainera *a priori* moins de confusions qu'une classe radiométriquement hétérogène, grossièrement définie. Les classes étant caractérisées par les échantillons d'apprentissage, l'enjeu est d'identifier des échantillons radiométriquement homogènes, mais représentatifs à eux tous de l'hétérogénéité de l'occupation des sols.

Les échantillons d'apprentissage « bruts » en quatre classes sont, par construction, radiométriquement hétérogènes du fait des objets divers contenus dans une même classe. Les surfaces bâties par exemple se décomposent en toitures grises, toitures orange, toitures blanches, toitures à l'ombre etc...

Une étape importante de la méthodologie consiste à sous-échantillonner les informations d'apprentissage en fonction de critères radiométriques. La nomenclature initiale est ainsi décomposée en sous-classes radiométriquement homogènes, qui n'ont pas nécessairement de signification thématique. Ce découpage peut être obtenu « manuellement » par interprétation visuelle, comme c'est le cas dans cette étude, ou par des procédés automatisés de classifications non supervisées.

Faute d'outils disponibles pour effectuer des classifications objet, la classification retenue est une classification pixel SVM à noyau linéaire (Vapnik, 1999). Pour se rapprocher de la logique des approches objet, des classifications sur images segmentés ont toutefois été intégrées aux tests de calibrage méthodologique. Le résultat est fourni sous forme raster. La nomenclature de sortie, *a priori* identique à la nomenclature d'entrée, contient entre 4 et 18 classes suivant la finesse du sous échantillonnage testé.

Les post traitements sont les opérations effectuées à l'issue de la classification. Ils sont ici de deux types : une fusion des sous-classes, par calculatrice raster, afin de retrouver la nomenclature en quatre classes initialement recherchée, puis un lissage par filtre majoritaire afin de supprimer les pixels isolés et limiter certains problèmes de bordures.

Cette méthodologie met en œuvre de nombreux « paramètres » susceptibles d'impacter la qualité du résultat final : choix et caractéristiques des néocanaux retenus (textures, indices, sursol...), choix des paramètres de l'algorithme de classification, choix de la taille de fenêtre utilisée pour le lissage etc... Différents travaux dans la littérature (Hester *et al*, 2008, Ackermann *et al*, 2004...) ont permis d'identifier 22 combinaisons de ces facteurs, ou scénarios, comme représentatives des configurations possibles.

Par construction, ces scénarios donnent des résultats différents. Le scénario optimal est celui qui maximise la qualité et la stabilité du résultat. Il est identifié en comparant la qualité des classifications obtenues avec les 22 scénarios. En raison de temps de calculs élevés sur la zone complète, qui compte environ 800 millions de pixels, la méthodologie a été calibrée sur deux quartiers tests d'environ 1 km<sup>2</sup> : Saint-Michel au centre-ville (Figure 2), et Rangueil, en périphérie.



Figure 2 : Visualisation de l'image Pléiades sur une des deux zones de paramétrage (quartier Saint-Michel).

La qualité est quantifiée de manière statistique via l'indicateur kappa (Landis *et al*, 1977), qui se calcule au regard d'échantillons de vérification. Aucune occupation des sols suffisamment cohérente avec l'image n'étant disponible, ils sont construits indépendamment des échantillons d'apprentissage, par photo-interprétation d'environ 10% de la zone. Ils sont sous la forme de polygones labellisés en quatre classes (Figure 3).



Figure 3 : Les échantillons de vérification sur la zone Saint-Michel.

Le kappa est calculé pour chacun des scénarios testés. La comparaison des valeurs obtenues (Figure 4) permet de qualifier relativement l'impact des différents paramètres et d'identifier le scénario optimal permettant de produire la meilleure couche d'occupation des sols.

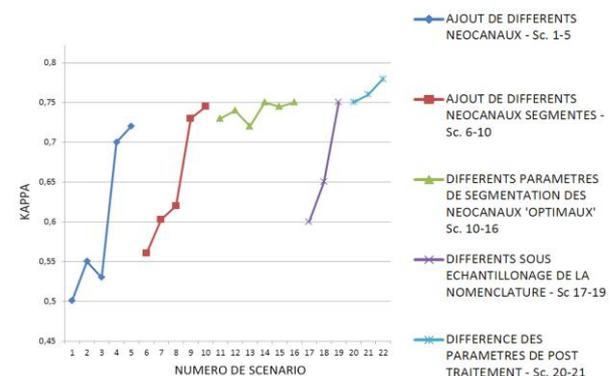


Figure 4 : Coefficient kappa en fonction du scénario.

Le détail des paramètres associés à chacun des scénarios ne sera pas décrit ici. Cependant, l'analyse a montré l'importance du choix des néocanaux, et notamment l'apport d'une couche de sursol dans la limitation des confusions (différence entre scn 1-2-3 et scn 4-5). Le sursol utilisé sur la zone test est obtenu par différence entre un MNE issu d'un couple stéréoscopique et la BD Alti. Enrichir l'image avec des bandes segmentées (algorithme mean shift (Comanico *et al*, 2002)) plutôt qu'avec des néocanaux bruts est également intéressant, ce qui corrobore la littérature sur l'intérêt de l'approche objet pour l'étude des milieux urbains. Les paramètres de segmentation ont quant à eux une influence limitée.

La finesse de la nomenclature d'apprentissage a elle aussi un impact important sur le kappa, le meilleur résultat étant obtenu avec la nomenclature la plus fine.

Au final, le scénario retenu correspond à un traitement des bandes radiométriques enrichies du NDVI, de la variance sur le PIR (radius = 2) et du sursol, toutes les bandes étant segmentées (taille minimale = 20 m<sup>2</sup>, angle spectral = 20° et angle spatial 25°). La nomenclature optimale d'apprentissage compte 18 sous classes, fusionnées en 4 classes à l'issue de la classification. Le kappa associé vaut 0,78, mais cette valeur est délicate à utiliser dans l'absolu dans la mesure où elle dépend d'échantillons de vérification potentiellement subjectifs.



**Figure 5 :** Couche d'occupation des sols optimale sur une des zones de paramétrage (quartier Saint-Michel).

### La détermination de la densité urbaine

L'objectif de l'étude est *in fine* de déterminer une densité de bâti par îlot urbain. L'îlot urbain est une unité urbaine de quelques parcelles cadastrales délimitées par les voies de circulation<sup>5</sup>.

L'indicateur retenu pour caractériser le bâti est la densité surfacique, proche de la notion de « coefficients

<sup>5</sup> En 2006, l'INSEE a remplacé dans ses études l'îlot urbain par l'IRIS (Ilot Regroupé pour des Indicateurs Statistiques), mais pour des raisons de coûts du découpage IRIS, notre étude a retenu l'îlot urbain comme unité de découpe.

d'emprise au sol» des urbanistes. Elle se définit comme le ratio entre la surface de bâti dans un îlot et la surface de cet îlot. Des tests ont été effectués pour calculer une densité volumique (en m<sup>3</sup>/m<sup>2</sup>), définie comme le volume de bâti (en m<sup>3</sup>) par surface d'îlot urbain (m<sup>2</sup>), mais cette piste n'a pas été entièrement exploitée en raison de la difficulté d'obtention de données de sursol à jour. Cette information présente toutefois un intérêt fort pour l'évaluation des « coefficients d'occupation des sols ».

Le bâti extrait de la classification est croisé avec les polygones d'îlots urbains selon des méthodes géomatiques classiques (ici avec le plugin QMarxan de Qgis) afin de fournir une cartographie des densités surfaciques urbaines (Figures 6 et 7).



**Figure 6 :** Surfaces bâties issues de la classification (quartier Saint-Michel).



**Figure 7 :** Densité surfacique de bâti (en %) selon les îlots INSEE (Quartier Saint-Michel).

La méthodologie ainsi mise en place et calibrée sur les zones tests est ensuite appliquée sur l'ensemble de l'agglomération toulousaine (Figure 8).

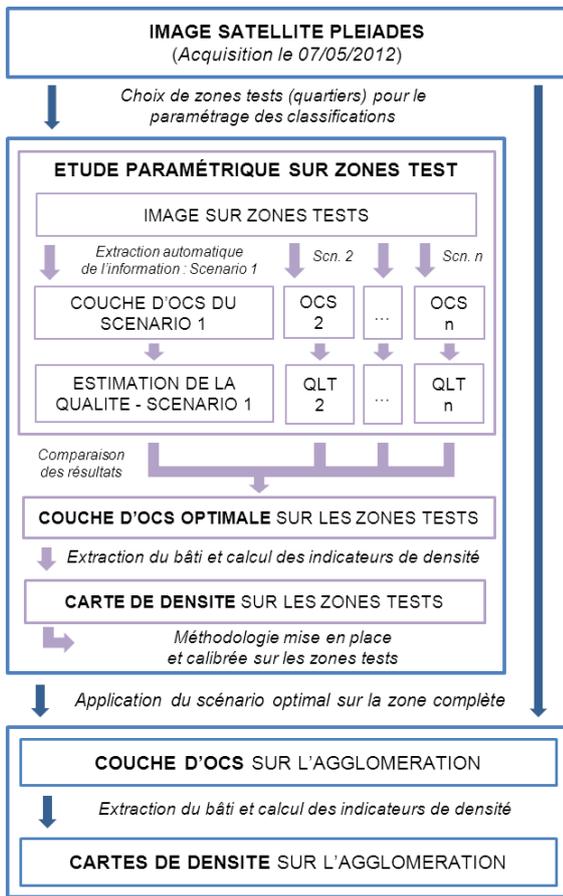


Figure 8 : Principe général de la méthodologie

### 3. Résultats et analyse

Pour des raisons de capacité de calcul, l'image a été découpée en neuf zones. La figure 9 montre le mosaïquage des neuf résultats correspondant.

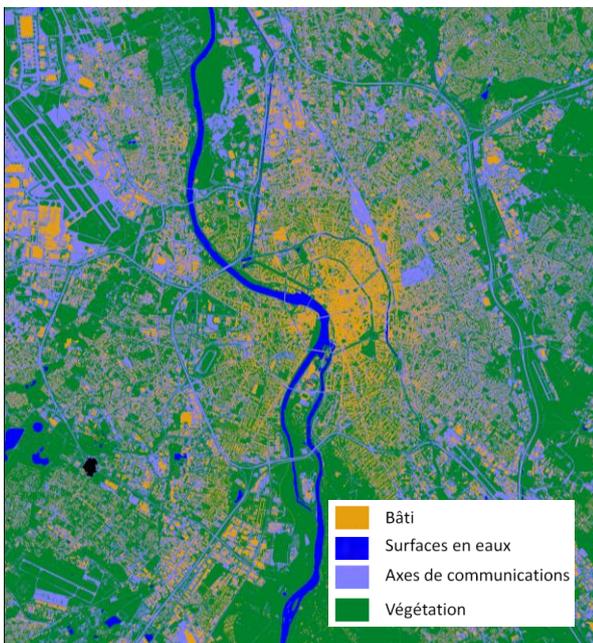


Figure 9 : Couverture du sol obtenue sur la zone d'étude complète (agglomération toulousaine)



Figure 10 : Comparaison entre image Pléiades (gauche) et classification (droite) en zone résidentielle.

Les surfaces de bâti ainsi obtenues diffèrent du bâti d'autres sources d'information, notamment en raison de la mise à jour de l'information. La figure 11 compare le résultat de classification et les bâtiments présents dans la BD Topo dans une zone de changements rapides. Même si le contour des bâtiments y est moins propre, les surfaces réelles de bâti sont statistiquement plus justes sur le résultat de classification.



Figure 11 : Exemple de bâti détecté par télédétection mais non présent dans la BD Topo.

- Bâti
- Surfaces en eaux
- Axes de communications
- Végétation
- Contour du bâti de la BD Topo

La densité de bâti est ensuite déduite sur l'agglomération complète (Figure 12).

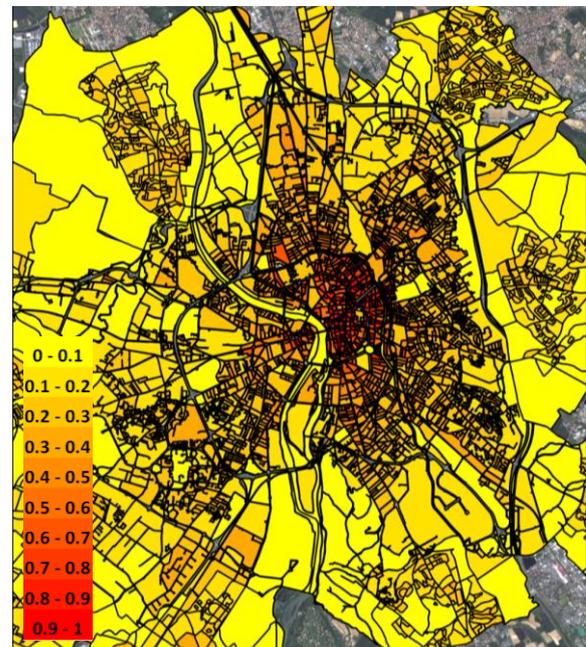


Figure 12 : Densité surfacique de bâti (en %) sur la zone d'étude complète (agglomération toulousaine)

Ce résultat est précieux dans la compréhension de l'étalement urbain et dans l'identification rapide des zones en sous-densité sur un vaste territoire. Le niveau de précision de la donnée rend envisageable son utilisation pour le suivi des politiques d'aménagement, a minima comme premier indicateur pour identifier les zones nécessitant une analyse *in situ* plus fine.

Cette étude montre cependant certaines limites. La qualité de l'image est un premier facteur d'amélioration possible. En particulier, l'angle d'acquisition élevé engendre la présence de façades sur l'image, ce qui a pour conséquence une sur-détection des surfaces de bâti. Par ailleurs, l'accès à un MNE cohérent avec l'image, non disponible au moment de l'étude, permettrait d'accéder à la hauteur des bâtiments, utile pour les analyses de densité volumique.

Pour les étapes de télédétection, le calibrage de la méthodologie, dont l'impact est significatif sur la qualité de reconnaissance du bâti, peut être approfondi. Aujourd'hui, des chaînes de traitements internes permettent d'enchaîner des centaines de scénarios combinant de nombreux paramètres pour un calibrage fin et des études de stabilité complètes.

L'utilisation combinée de bases de données exogènes, pour l'apprentissage ou pour certains post traitements, est une piste en cours d'intégration à la méthodologie. Les étapes de sous-échantillonnage peuvent également être améliorées en définissant de manière plus rigoureuse les nomenclatures intermédiaires sur des critères radiométriques, par l'utilisation par exemple de classifications non supervisées.

Les densités urbaines ont été réalisées selon le découpage des îlots INSEE. D'autres découpages, comme les plans cadastraux, les plans de zonage des PLU, les îlots regroupés pour l'information statistique ou, par défaut, le carroyage INSEE peuvent s'avérer utiles. Généralement, cette surcouche est fournie par l'utilisateur final pour répondre à un besoin précis. Enfin, les indicateurs pourraient être améliorés pour obtenir directement les COS et les CES utilisés pour le droit à construire dans les documents opposables. La qualité du résultat final doit cependant être validée avant d'envisager leur utilisation dans un contexte opérationnel juridiquement sensible.

Au-delà des limites susmentionnées, cette étude met en évidence le potentiel des traitements automatiques sur images Pléiades pour la caractérisation de la densité urbaine, en montrant qu'il est possible d'obtenir des cartes d'indicateurs sur de larges surfaces. La capacité de revisite des satellites Pléiades permet d'envisager des mises à jour annuelles, ce que ne proposent pas les bases de données utilisées à ce jour dans les services institutionnels.

Copyright images Pléiades : ©CNES (2012), distribution Airbus DS/Spot Image

## Références

Ackermann G., Tournaire O., Mering C., 2004. *Etude de l'extension du bâti sur le littoral sénégalais à partir des paramètres texturaux de Haralick*. Revue Française de Photogrammétrie et de Télédétection, n° 173/174, pp.104-112

AUAT, 2012, *Toulouse, quatrième aire urbaine de France*, Perspective Villes, janvier 2012, pp.1-4

Bauer T., Steinnocher K. (2001): *Per parcel land use classification in urban areas applying a rule-based technique*. GeoBIT/GIS 6(2001), pp. 24-27.

Blaschke, T., Burnett, C., Pekkarinen, A., 2004. *New contextual approaches using image segmentation for object-based classification*. De Meer, F., de Jong, S. (Eds.), *Remote Sensing Image Analysis: Including the spatial domain*. Kluwer Academic Publishers, Dordrecht, pp. 211-236

CERTU, 2011, *La consommation d'espaces par l'urbanisation. Panorama des méthodes d'évaluation*, Edition du CERTU, Lyon, ISBN 978-2-11-099573-5

Comanicu D., Meer P., 2002, *Mean shift: a robust approach toward feature space analysis*, IEEE Transactions on pattern Analysis and Machine Intelligence, Volume 24, Issue 5, pp. 603-619

Landis J.R., Koch G.G., 1977, *The Measurement of Observer for Categorical Data*, Biometrics, 1977a, pp.159-174

Comité interministériel pour le développement durable, 2010, *Stratégie Nationale de Développement Durable*

Faure J-F., Tran A., Gardel A., Polidori L., 2004. *Elaboration d'un indice de densité de population et analyse de sa distribution spatiale à Belém (Brésil) et Cayenne (Guyane Française)*, Revue Française de Photogrammétrie et de Télédétection, n° 173/174, pp.135-144

Hester D.B., Cakir H.I., Nelson S.A.C.,Khorram S., 2008. *Per-pixel classification of high spatial resolution satellite imagery for urban land-cover mapping*. Photogrammetric Engineering and Remote Sensing, vol. 74, n°4, pp.463-471

Holland, D.A., Boyd, D.S., Marshall, P., 2006. *Updating topographic mapping in Great Britain using imagery from high-resolution satellite sensors*. ISPRS Journal of Photogrammetry and Remote Sensing, 60 (3), 212\_223.

Jacquin, A., Misakova, L., Gay, M., 2008, *A hybrid object-based classification approach for mapping urban sprawl in periurban environment*. Landscape and Urban Planning 84 (2), pp 152-165.

V. Vapnik, 1999, *The nature of statistical learning theory*, Second edition, Springer Verlag New York, Inc

Walker J.S., Blaschke T., 2008. *Object-based land-cover classification for the Phoenix metropolitan area : optimization vs. transportability*, International Journal of Remote Sensing, vol. 29, n°7-8, pp. 2021-2040