

ASTERIX – Réunion T+12 – Rennes, 16 et 17 octobre 2014

Présents : SL, NC, YC (Vannes) ; RT, TC, PG (Rennes) ; AP, JPM, SR (Strasbourg) ; DS, MF, PAH (Toulouse) ; partiellement : JN, ML, PA (Rennes)
Début jeudi 16/10 à 10h30, fin vendredi 17/10 17h

TODO :

- **mettre les slides sur la plateforme web**

11h Accueil

Rappel du programme des 2 journées, point d'étape :

- Accord consortium
- RH : stages, ingénieur/postdoc
- Budget : dépenses prévues 2015
- Web/communication : MAJ partie publique, mise en ligne des supports dans la partie privée ; pas de logo ; rappel de la consigne ANR : mettre en place et tenir à jour un site web public du projet (mise à jour aux jalons du projet et au moins tous les 6 mois)
- Données : présentation des données disponibles, celles à venir ; discussion sur couplage tâches méthodologiques vs tâches applicatives (vs données)
- Réalisation (revue du tableau communiqué dans le CR précédent), discussion sur la mise à disposition des codes
- Calendrier, livrables

Il est important de prendre le temps de discuter des recrutements (stages + profils postes Postdoc)

Retrait de Thomas Guyet du projet, solution de remplacement pour le web/mail à trouver

11h20 (JN) Infrastructure de données spatiales dédiée à la recherche et l'observation scientifique de l'environnement

Présentation d'une plateforme qui pourrait être utilisée pour le projet ASTERIX.

Interrogations sur les efforts à fournir pour mettre à disposition des données, des résultats, des méthodes de traitement. Discussion sur la pérennité du système, et l'ouverture aux utilisateurs. Analyse de la plus-value pour le projet. Ce n'est pas une plateforme de traitement (une autre initiative est menée à l'OSUR dans ce but, avec appel à un prestataire).

Il est possible de créer une session ASTERIX sur Indigéo le temps du projet, avec possibilité de verrouiller les accès. Accord sur l'usage de la plateforme : plutôt en valorisation que dans le cadre d'échanges de données, pour mettre à disposition les résultats du projet auprès de la communauté scientifique.

TODO :

- **Jean crée un espace ASTERIX**
- **Présentation des sites d'études à saisir par les responsables des tâches applicatives**
- **Traitements : attendre que la solution OSUR mûrisse (point à l'ordre du jour dans un an)**

12h10 (TC) POPS

Plateforme de gestion de projet. Connexion basée sur Shibolet. Insertion facile de publis saisies dans HAL. Possibilité d'une petite partie publique. Test de la solution et prévision de remplacement de la solution actuelle Joomla si POPS permet de tout faire.

TODO :

- **Thomas crée quelques comptes de bêtesteurs (TC,RT,JPM,SL) avec rôle de coordinateur pour SL**

14h30 Présentations scientifiques T1.x et T2.x

- 14h40 (SL) Description morphologique supervisée d'images hyperspectrales (ICIP 2014, T1.1)
- 15h15 (MF) Sélection parcimonieuse de variables spectrales (WHISPERS 2014, T1.1)
- 16h10 (ML) Fusion séries temporelles d'images optique et radar (Stage M2)
- 16h50 (SL) Analyse d'image multi-échelle et supervisée pour la classification d'images hyperspectrales (WHISPERS 2014, T1.1)
- 17h30 (PG) Boosted kernels (thèse A. Lechervy)
- 17h55 (SR) apprentissage actif : du bi-classes au multi-classes (thèse)

Vendredi

9h Présentations scientifiques T3.x

- 9h20 (NC) Adaptation de domaine sur des données de télédétection (ECML 2014, T3.3)
- 10h15 (RT) Classification de séries temporelles (stage A. Nguyen)
- 10h40 (TC) Reconstruction de séries temporelles de LAI à partir de séries MODIS (thèse X. Gong)

11h10 (JP) Données IPGS

Mise à disposition d'une série temporelle d'images photo. La technique utilisée par JPM est une corrélation d'image hiérarchique, dans le but d'obtenir la cinématique de l'objet. Un travail est prévu sur une méthode de texture et non directement par analyse de la radiométrie. Il est prévu l'utilisation des outils T1.x sur les images, puis ceux T3.x sur les séquences. Les données à venir sont 50 séries temporelles LIDAR associées (**TODO : dépôt FTP**), 40 séries de photos sont aussi disponibles. Les besoins concernent le suivi d'objets particuliers dans l'image, la recherche de motifs fréquents dans la série temporelle (collaboration N. Méget). TC dispose d'outils d'assimilation de données avec interactions entre échelles (thèse P. Zille). TC propose un stage pour utiliser cette méthode sur les données IPGS, l'outil étant plutôt prévu pour l'estimation et l'analyse de mouvement (le problème de l'assimilation est un peu plus exploratoire). TC maîtrise bien ces outils et peut également réaliser une expérimentation. JPM indique que Floriane Provost (ingénieur géophysique en début de thèse) s'attaquera au problème de déformation/modèle (assimilation) dans un an. Possibilité d'un stage en support de son travail l'an prochain.

11h25 (AP) Données LIVE

L'existant consiste en une analyse bi-date, pas encore d'analyse de la série temporelle. L'objectif est l'obtention de la trajectoire de changement dans la baie. Les objets considérés sont la progression des prés salés, des bancs sédimentaires, le positionnement des chenaux. Les images sur le FTP consistent en 2 séries (cf. description dans le document dédié). La recherche de motifs fréquents fait l'objet d'un stage de 2 mois utilisant la méthode de Méget. La vérité terrain est disponible sur les contours des prés salés et des chenaux (projet ECOSGIL). La classification est numérique à une date, puis on opère une classification symbolique. Il est possible de mettre les données brutes (ou orthorectifiées) et pas seulement les images calculées (indices) (**TODO : dépôt FTP**). L'intérêt méthodologique du LIVE est plutôt sur l'apprentissage actif. Le besoin est un suivi à temps régulier de la progression des prés salés et des autres objets. 1 relevé LIDAR est disponible (une seule date). TC indique qu'un relevé LIDAR est prévu tous les 6 mois. Il existe également des données gratuites (Landsat 8, Spot 4 take 5). Le besoin concerne la

comparaison post-classification, la fouille non supervisée de séries temporelles. L'outil développé par A. Nguyen porte sur la classification supervisée de séries temporelles, il ne semble donc pas adapté au problème.

11h50 (TC) Données COSTEL

L'application concerne uniquement les séries temporelles MODIS pour classer les pratiques agricoles sur des prairies. Les données FTP ne servent qu'à l'apprentissage et la validation, le jeu de données réelles est à venir. Des données THR sont également disponibles (2006 et 2011). Les besoins concernent la classification et le démélangeage.

14h Accord de consortium

Point téléphonique avec Adélaïde Nollet, SATT Ouest Valorisation. Pas d'entreprise donc accord de consortium / collaboration pas obligatoire. L'accord doit traiter au moins des points suivants (voir conventions/décisions ANR) : répartition des tâches, des moyens et des fournitures, partage des droits de PI, publication / diffusion / valorisation.

14h20 (DS) Données DYNAFOR

La tâche (a) concerne l'extraction des couronnes et si possible la classification des espèces (peu réaliste en multispectral). Les données disponibles sur le FTP sont : LIDAR interpolé (nuage brut disponible sur demande), Pleiades à 2m, orthophoto (infrarouge + couleur) à 50 cm. L'ONERA a acquis des données hyperspectrales mais elles ne sont pas encore exploitables. Ces images seront néanmoins disponibles pendant le projet.

La tâche (b) concerne différentes données dont les cartes anciennes. Ces dernières ont déjà été traitées par PAH pendant sa thèse et elles ne sont donc plus d'actualité. Il y a également des photos aériennes (2 anciennes + 2 récentes) (**TODO : mettre les données sur le FTP**). Les méthodes utilisées sont pour l'instant des méthodes classiques (objet). AP précise que le besoin d'extraction de l'arboré à partir de photos aériennes anciennes est fréquent. TC indique qu'A. Lefebvre a réalisé une méthode adaptée (dévignettage, suppression des effets de bord) qui pourrait être utilisée.

La tâche (c) concerne les prairies (cf. cas COSTEL mais avec des données Spot 4 Take 5). Il s'agit de distinguer des fauches et des pâturages (des outils sont disponibles chez TC) (**TODO : mettre les données sur le FTP**).

Passage en revue des différentes tâches et discussion des stages potentiels ou des expérimentations à mener par les membres du projet

- T4.3.3 : stage M2 autofinancé par Dynafor sur les prairies, intégrer le test du code Matlab d'Antoine Lefebvre sur la classification de séries temporelles
- T4.3.2 : stage M1 = test du code Matlab d'Antoine Lefebvre sur le dévignettage, extraction de l'arboré sur les images N&B appliqué aux prairies de Toulouse et Rennes (plutôt stage court car moins coûteux). L'étudiant de M1 OTG sera choisi par AP.
- T1.1 : tester les méthodes de MF et PG sur les attributs extraits par une approche multirésolution (arbres et profils), pour des objectifs de classification. Les jeux de données peuvent être : peupliers (DS), urbain (labels à un seul niveau), des images Quickbird fournis par AP avec des jeux de test
- T1.2 : stage DS+SL en 2015 ou 2016. Camille Kurtz a proposé une approche descendante. On envisage une remise en cause de la segmentation en partant de la résolution plus fine. Explorer également les travaux en adaptation de domaine ?
- T1.3 : stage Marianne réalisé. Publication IGARSS 2014 + soumission revue possible. Marianne reste affectée à cette tâche pendant son début de thèse.
- T1.4 : Yanwei s'intéressera au problème pendant la 2ème partie de sa thèse. Quand les noyaux seront disponibles, ils pourront être transmis à PG pour la tâche

T2.2

- T2.1 : méthode SAGA et travail Sina Nakhostin sur le passage à l'échelle de SAGA. Présenter cette méthode en 2015 pour motiver des collaborations. Travail avec LC sur la détection d'anomalies et la sélection d'attributs.
- T2.2 : 1 miniprojet ENSEA (1 mois de travail perso en M2R), voire 1 stage ensuite si l'étudiant est bon. Sinon envisager une offre de stage (seulement si candidat excellent). Noyaux spatiaux ou spectraux, validation sur Pavia. NC suggère d'améliorer le côté méthodologique et de considérer plutôt la variété, transport de noyaux ? Amorce de collaboration NC/PG sur le volet théorique
- T2.3 : 1 miniprojet ENSEA proposé par AP et encadré par PG pour appliquer le code de Simon à un cas multiclasse à plus de 3 classes (6-7), avec un autre classifieur. Poursuite en stage si bon étudiant (sinon offre de stage possiblement/probablement financée par ENSEA). Travail Yanwei pendant sa thèse. NC ne croit pas à l'utilisation de Perturbo pour l'actif. AP suggère que c'est peut être utile à titre de comparaison (donc intégration dans le stage LIVE : code Matlab, R disponible). Code PG finalement pas disponible.
- T3.1 : Pierre et Anh vont travailler sur des noyaux temporels. Proposition de changement du titre de la tâche.
- T3.2 : thèse Gong Xing, soumission TGRS. Expérimentation à mener sur les données JPGS.
- T3.3 : Application de la méthode de NC pour l'adaptation de domaine à des séries temporelles. Travail de LC pour la détection d'anomalie. Stage à prévoir en 2016.
- T4.1 : La plateforme pour la visualisation pourrait être Indigéo. Les traitements seraient déployables sur la plateforme de l'OSUR mais aucune date de livraison n'est annoncée. Les traitements seraient intégrés sous forme de WPS. Il serait possible d'héberger les traitements sur le cluster Vannes (WPS digérés par Indigéo).

Ressources humaines

- **Stage 2015** : cf discussions ci-dessus
- Ingénieur/Postdoc : proposition de découpage en **1 an de postdoc** (stratégie opportuniste si bon candidat) + **1 an ingénieur** (parallélisation des traitements)

Calendrier

Prochaines échéances : réunions T+18, T+24, tâches/livrables à vérifier.

Rapport intermédiaire à l'ANR en avril 2015.

Prochaine visioconférence 9/10 mars 2015.

Prochaine réunion physique en octobre 2015 à Toulouse.

Rappel des TODO

- **fournir les slides des présentations**
- **initier la zone Asterix sur Indigéo**
- **tester POPS**
- **mettre les données sur le FTP**
- **rédigier et diffuser les offres de stage**
- **mener les expérimentations court-terme**