

C. Boitel (1), M-A. Drouin (1), J-C. Dupont (1), R. Guzman (2), C. Pietras (1)  
(1) Laboratoire de Météorologie Dynamique/SIRTA – (2) Institut Pierre Simon Laplace

## Quelques chiffres

Après 17 ans d'observations, la base de données SIRTA c'est

- 14 To de données
- 15 millions de fichiers
- 80 flux de données
- ~ 600 fichiers par jour
- Bases privées (données brutes, campagnes, projets)
- Base publique
  - ✓ données mises en forme
  - ✓ produits élaborés (filières de traitement)
  - ✓ Quicklooks

C'est aussi un ensemble de systèmes informatiques

- ~ 40 pc d'acquisition
- ~ 20 centrales de mesures
- 4 serveurs de service
- 1 NAS

chargés de la numérisation, de la collecte, du traitement, de la sauvegarde et de la surveillance des mesures issues de 150 capteurs répartis sur 5 zones d'observation.



## L'accès

L'ensemble de cette archive est mise à disposition au travers de comptes ftp spécifiques pour les données brutes privées, et d'un compte ftp générique ou d'un site web<sup>2</sup> pour les données publiques standardisées (format, nomenclature) et issues de traitements (produits dérivés).

par FTP	status/identifiant	Utilisateurs
Données brutes	privé/ compte spécifique <sup>1</sup>	PI, responsable instrumentaux
Données élaborées	public/ sirta_access <sup>1</sup>	communauté

<sup>1</sup> ftp.sirta.ipsl.polytechnique.fr, <sup>2</sup> http://sirta.ipsl.fr

## Le mésocentre de l'IPSL

Implanté sur les campus Pierre et Marie Curie et de Polytechnique, le mésocentre de l'IPSL fournit l'infrastructure informatique et le service dont le SIRTA a besoin pour

- Archivage des données
- Sauvegarde des données
- Hébergement site web
- Traitement
- Partage de la base avec le pôle de données et de service AERIS.

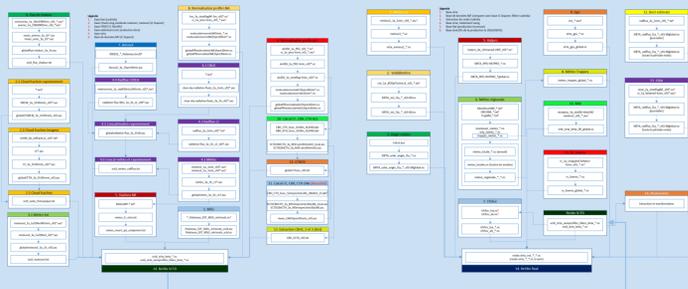


## ReOBS : production d'un fichier NetCDF multiparamètres long terme



Produit ACTRIS, ReOBS est un jeu de données multi-variables colocalisées pour l'étude de processus climatiques à l'échelle régionale sous forme d'un fichier au format NetCDF. Production actuelle très hétérogène et complexe :

- + de 20 filières
- 5 langages de dev.
- NetCDF non standard
- ~ 600 fichiers de codes
- des dizaines de bases temporaires



Objectif des travaux en cours : viser une gestion entièrement paramétrable des filières avec des codes normalisés s'inscrivant dans une architecture logicielle la plus simple possible. Le but :

- Faciliter la maintenance de la gestion de l'ensemble
- Réduire les coûts d'intégration de nouvelles filières
- Avoir une structure de codes modulaire simple à faire évoluer ou à déboguer.



➔ Voir aussi poster ReOBS - R. Guzman

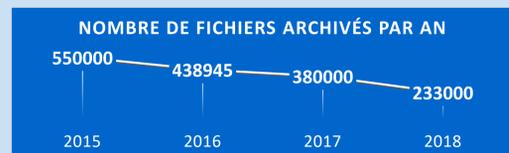
## Outils : la surveillance temps réel

Pour faciliter la surveillance du fonctionnement instrumental, le SIRTA a mis en place autour des outils **InfluxDB** et **Grafana** des flux quasi-temps réel indépendants des flux qui alimentent la base SIRTA.

- Surveillance en temps réel de certains status instrumentaux (à la minute)
- Suivi des produits de données (20 minutes après l'acquisition)

Conséquence positive sur la base :

une baisse très significative du nombre de fichiers archivés quotidiennement par la suppression des fichiers de courte durée (<24 heures) qui servaient au suivi temps réel.



➔ Séminaire de terrain « Base de données SIRTA. Flux de données en temps réel » - M-A. Drouin

## GRUAN : l'IPSL, centre de données



GRUAN (GCOS Reference Upper-Air Network) est la réponse de la communauté scientifique à la question de l'urgence d'établir des mesures de référence dans le cadre du changement climatique L'IPSL, Meteomodem et Météo-France travaille ensemble pour certifier GRUAN la sonde M10.

Rôle du SIRTA et de l'IPSL :

- **Validation** des sondes météo-Modem.
- **Normalisation** du format et des métadonnées.
- Développement de l'algorithme de **correction**
- **Collecte** des radiosondages issus de sondes Meteomodem (Trappes 2018, La Réunion + Faa'a 2019)
- Envoi vers base de données GRUAN

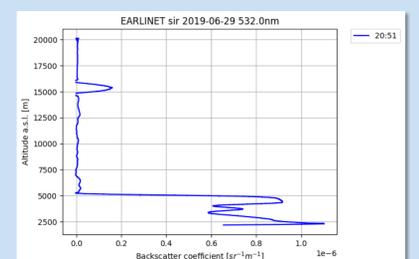
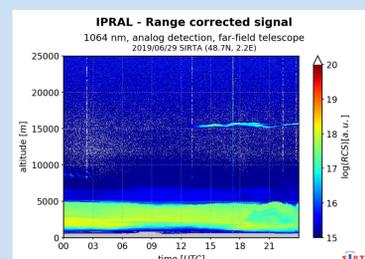


➔ Voir aussi poster GRUAN 1&2 - J-C. Dupont

## SCC : EARLINET Single Calculus Chain



SCC est un outil en ligne pour l'analyse automatique des données lidar aérosols du réseau Earlinet. Le SIRTA est un site référence pour le prototypage de l'automatisation de la soumission des données LIDAR à l'outil SCC. Il est en mesure d'apporter son expertise à d'autres sites désirant soumettre leur données à SCC et est une force de proposition au sein d'ACTRIS pour améliorer les processus d'utilisation de SCC.



Profil d'inversion SCC