



CENTRE NATIONAL D'ÉTUDES SPATIALES

Migrations de catalogues au CNES

CDPP (Centre de Données de Physique des Plasmas)

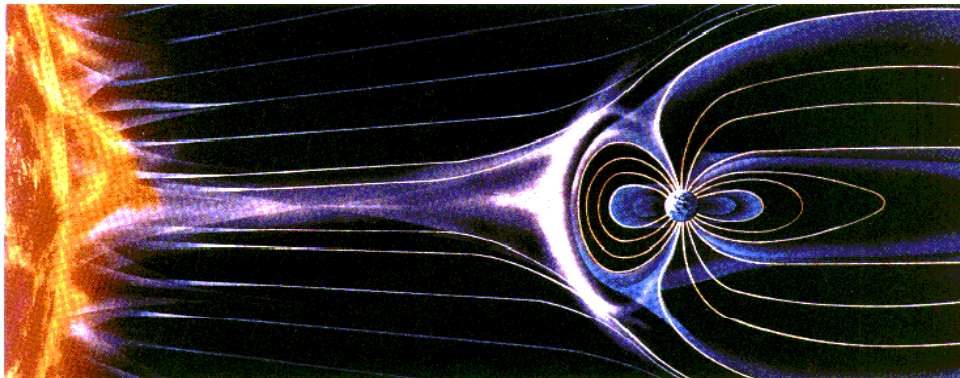
MERCATOR (Océanographie opérationnelle)

AVISO (Altimétrie)

CDPP : Centre de Données de la Physique des Plasmas

■ Qu'est-ce qu'un plasma ?

- Gaz ionisé formé de particules neutres, ions et électrons libres,
- Environnement terrestre (altitude > 80 km) sous forme de plasma (ionosphère, magnétosphère)

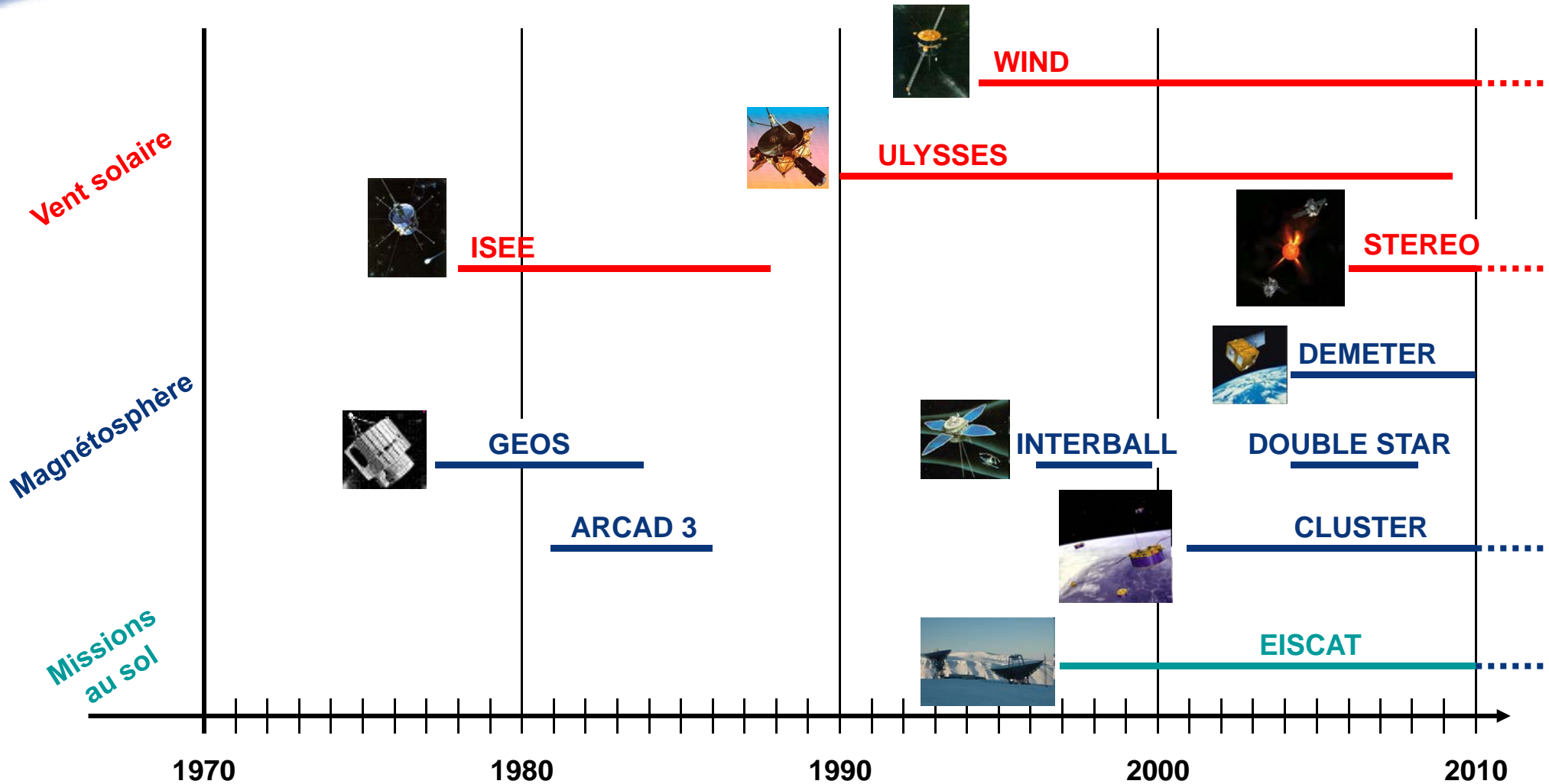


■ Missions du CDPP :

- Assurer la **préservation physique des données** de la physique des plasmas naturels et en **faciliter l'exploitation par la communauté** scientifique nationale et internationale
- Mettre à disposition les **données** ainsi que des **services à valeur ajoutée** (SVA) développés en vue d'enrichir l'utilisation des données archivées

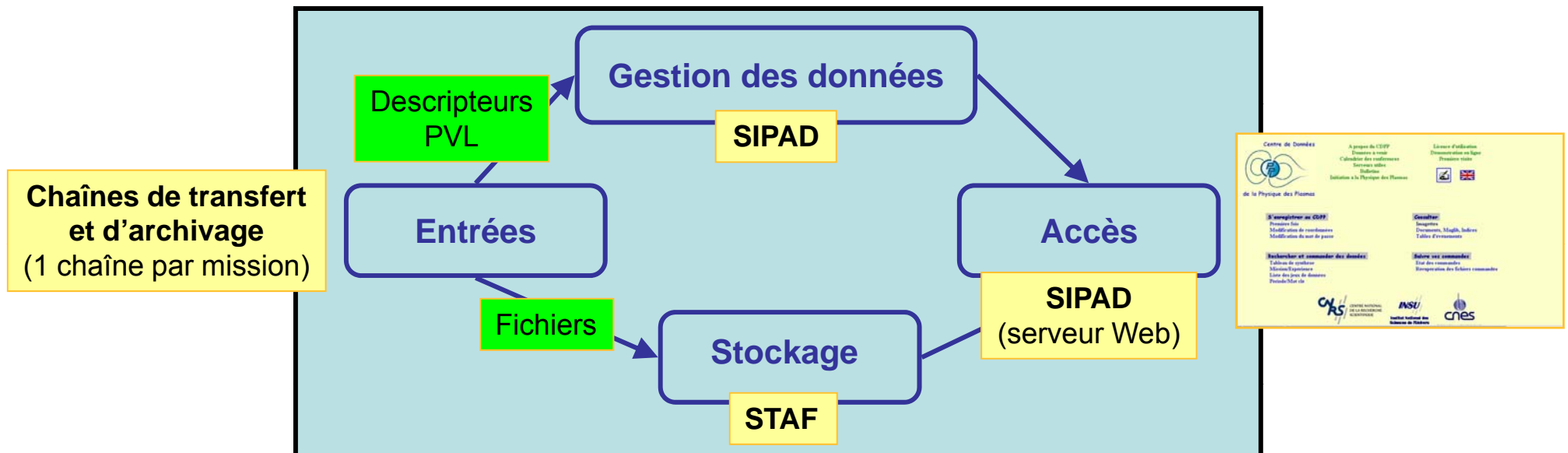
■ Le CDPP est régi par une convention CNES – CNRS

Quelques missions archivées au CDPP



Situation avant migration

■ 1998 – 2006 :



■ Mi-2005 : début des travaux de remplacement du SIPAD par le SIPAD-NG

■ **Objectif initial : migration à l'identique**

- Reconduction du « look & feel » de l'IHM Web Utilisateur
- Migration des métadonnées du catalogue SIPAD vers le catalogue SIPAD-NG
- Adaptation limitée des chaînes pour remplacer PVL par XML
- → **Durée prévue : ~ 2 mois**

Déroulement de la migration

■ Révision des objectifs – Profiter du passage au SIPAD-NG pour :

◆ Refondre l'IHM Utilisateur

- Retours des utilisateurs : IHM du SIPAD obsolète

◆ Enrichir le dictionnaire de données

- Prise en compte de travaux de standardisation en Physique des Plasmas (SPASE)
- Mise en conformité avec le dictionnaire « CLUSTER Active Archive »

■ Démarche :

◆ IHM :

- Prototypage en interne CNES et validation par l'équipe scientifique, puis industrialisation

◆ Métadonnées descriptives des jeux de données :

- Production manuelle des descripteurs XML des jeux
- Développement de formulaires de saisie de métadonnées par les scientifiques

◆ Métadonnées descriptives des objets de données :

- Développement d'un outil d'extraction de la base SIPAD produisant des descripteurs XML pour le SIPAD-NG

Bilan de la migration

■ Effort de migration bien plus important que prévu initialement :

- ◆ **Durée : 1 an** (année 2006)
- ◆ **Charge CNES : ~ 3 hommes*ans** (+ support industriel)
- ◆ **Participation de l'équipe scientifique**

■ Impact sur les différentes fonctions :

- ◆ **Stockage** : aucun impact sur l'archive STAF
- ◆ **Acquisition** :
 - Production des métadonnées associées aux objets de données sous forme de fichiers XML
 - Impact faible car les chaînes génèrent des listes (fichiers catalogues) converties en PVL ou XML par un utilitaire commun
- ◆ **Accès** :
 - Développement d'une IHM Web à partir de l'IHM SIPAD-NG
- ◆ **Gestion des données** :
 - Très gros travail de mise à niveau du catalogue ...
 - ... qu'il faudra entreprendre de nouveau pour intégrer les nouveaux standards SPASE

Mercator-Océan : quel océan fera-t-il demain ?

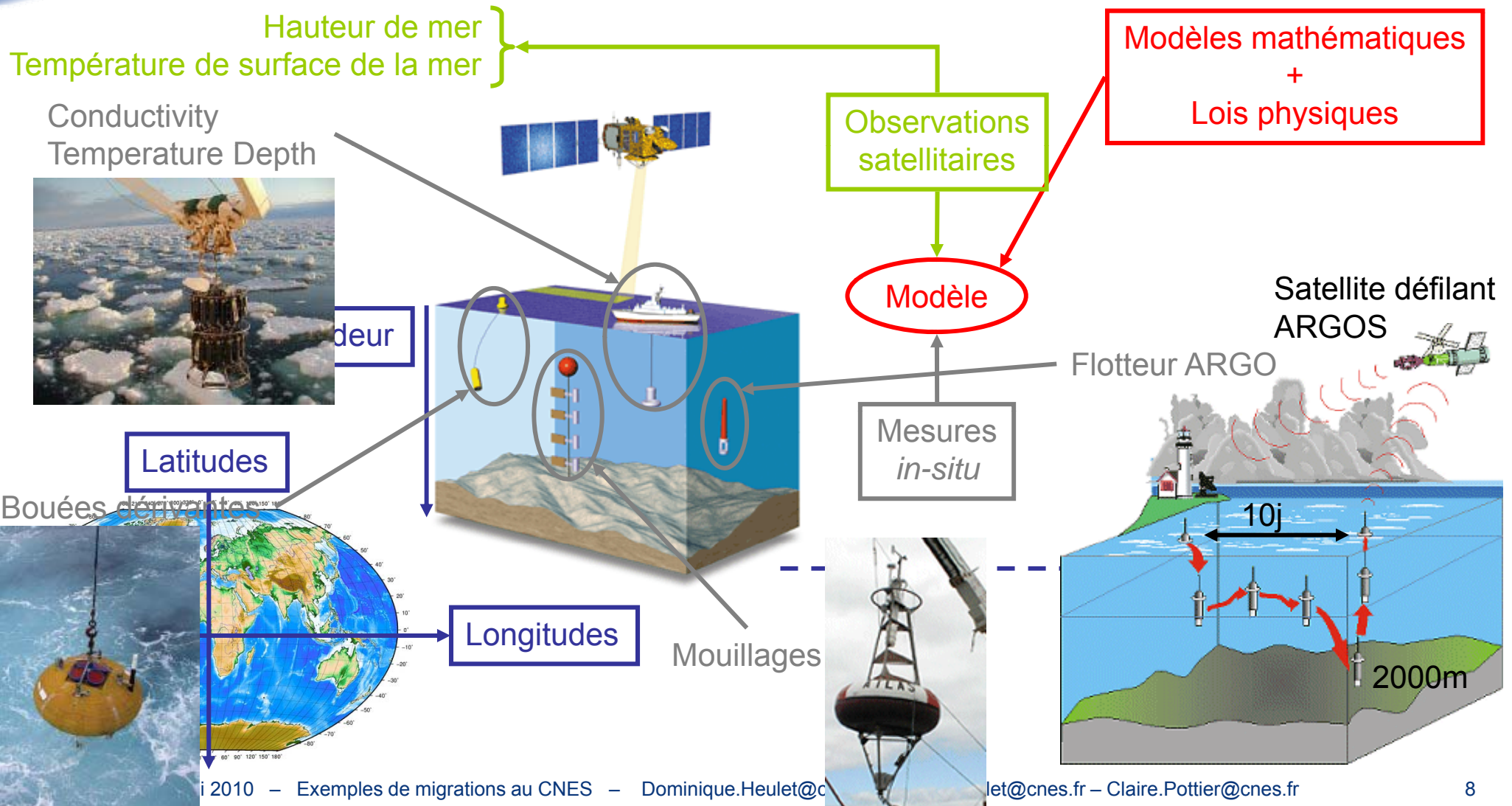


- GIP créé en avril 2002 par :



- Mission: décrire et prévoir en temps réel et à tout moment l'océan sur tout le globe, en surface comme en profondeur

Observer, comprendre et prévoir



Applications

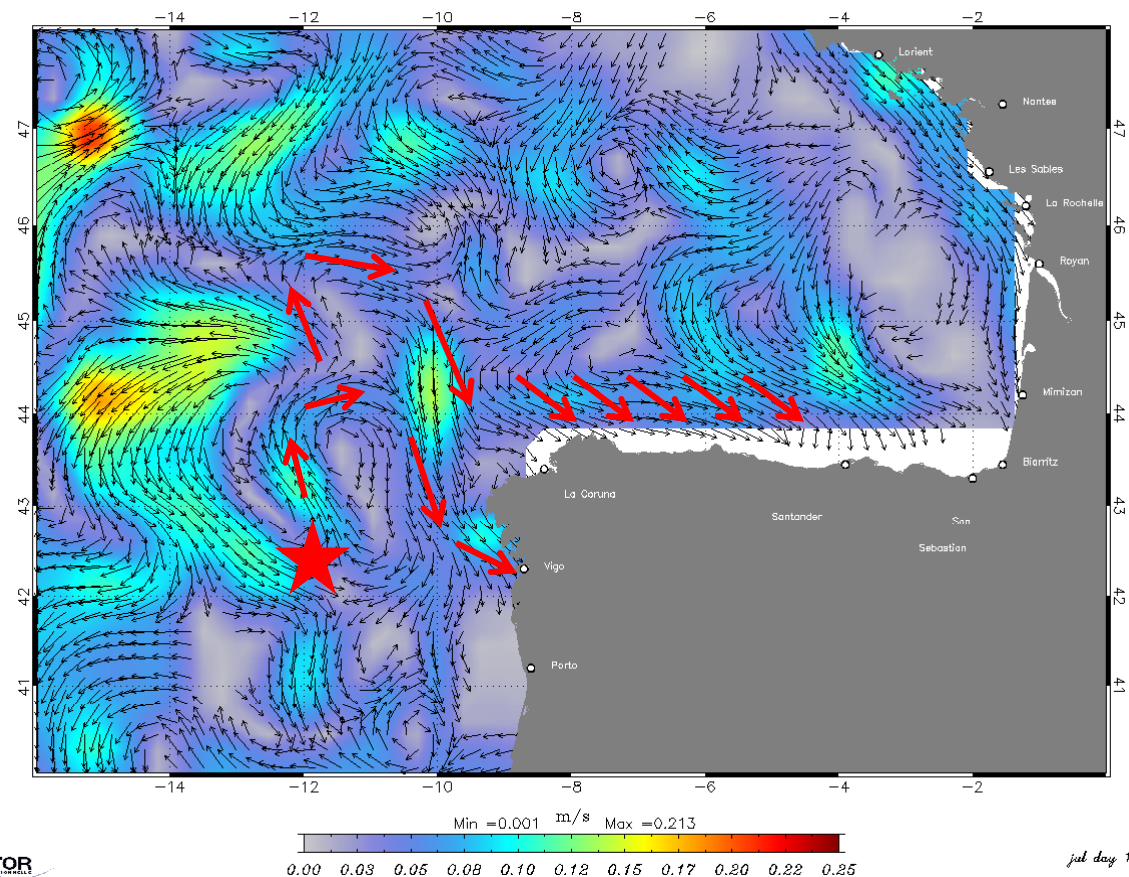
■ Des applications pour tous les spécialistes de la mer:

- ◆ l'environnement et la sécurité
- ◆ la recherche
- ◆ l'industrie
- ◆ la course au large
- ◆ ...

Pollution par le Prestige

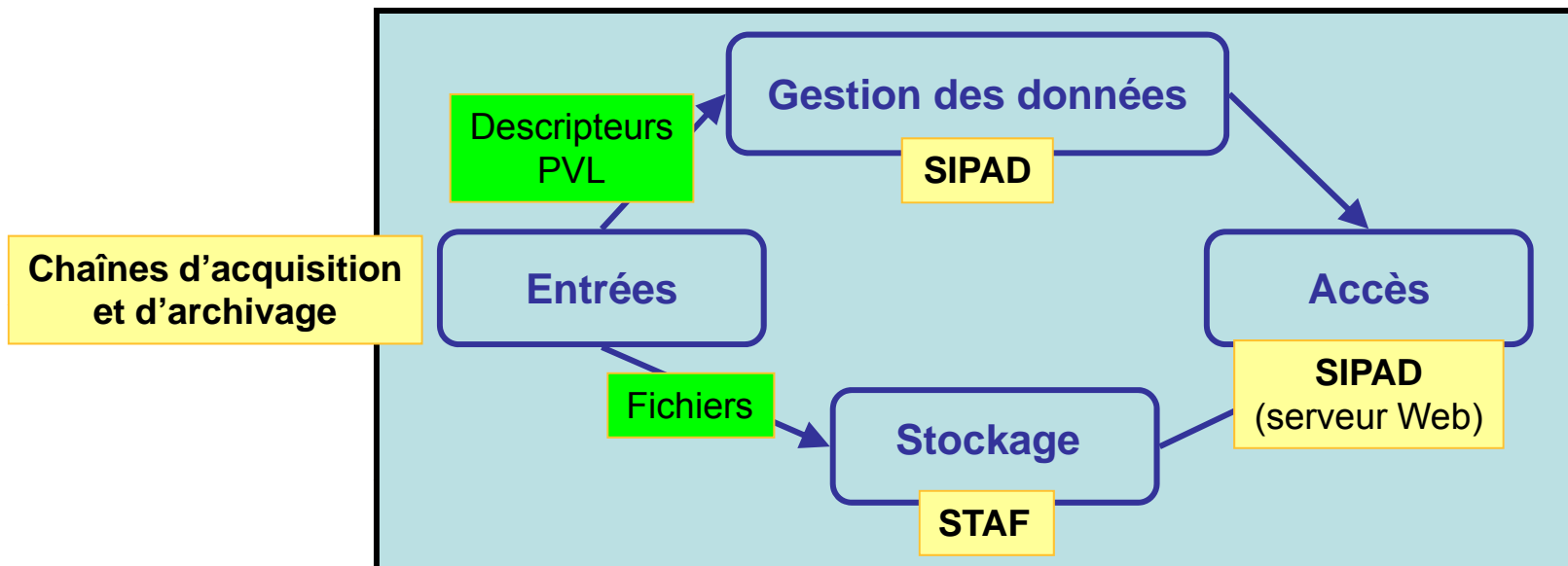
- 13/11/02: avarie au large du Cap Finisterre (Espagne)
- 19/11/02: coule au large des côtes de Galice (3500 mètres de profondeur, ~50 à 60 000 To de fuel lourd dans ses cuves)
- quatre semaines + tard: le fuel s'échappe par les nombreuses fissures de la coque (~125 To/j)

PSY1v1 Mean Surface Current : 12-2002



Situation avant migration

■ 2002 – 2008 :



■ Janvier 2007 : début des travaux de remplacement du SIPAD par le SIPAD-NG

■ **Objectif initial : migration à l'identique**

- Reconduction du « look & feel » de l'IHM Web Utilisateur
- Migration des métadonnées du catalogue SIPAD vers le catalogue SIPAD-NG
- Adaptation limitée des chaînes pour remplacer PVL par XML
- → **Durée prévue : quelques semaines**

Déroulement de la migration

■ Refondre l'IHM Utilisateur

◆ Retours des clients :

- IHM du SIPAD obsolète
- nécessité de 2 IHM distinctes pour 2 communautés

◆ Prototypage en interne CNES et validation par l'équipe scientifique, puis industrialisation

■ Constitution de 2 catalogues (produits grille standard et archive entière)

◆ Définition des dictionnaires de données

◆ Identification des jeux à cataloguer

◆ Production manuelles des descripteurs des jeux de données à partir d'informations fournies par l'équipe scientifique dans des tableaux Excel

◆ Outil de production des descripteurs d'objets à partir du contenu STAF (contenu STAF et catalogue SIPAD incohérents)

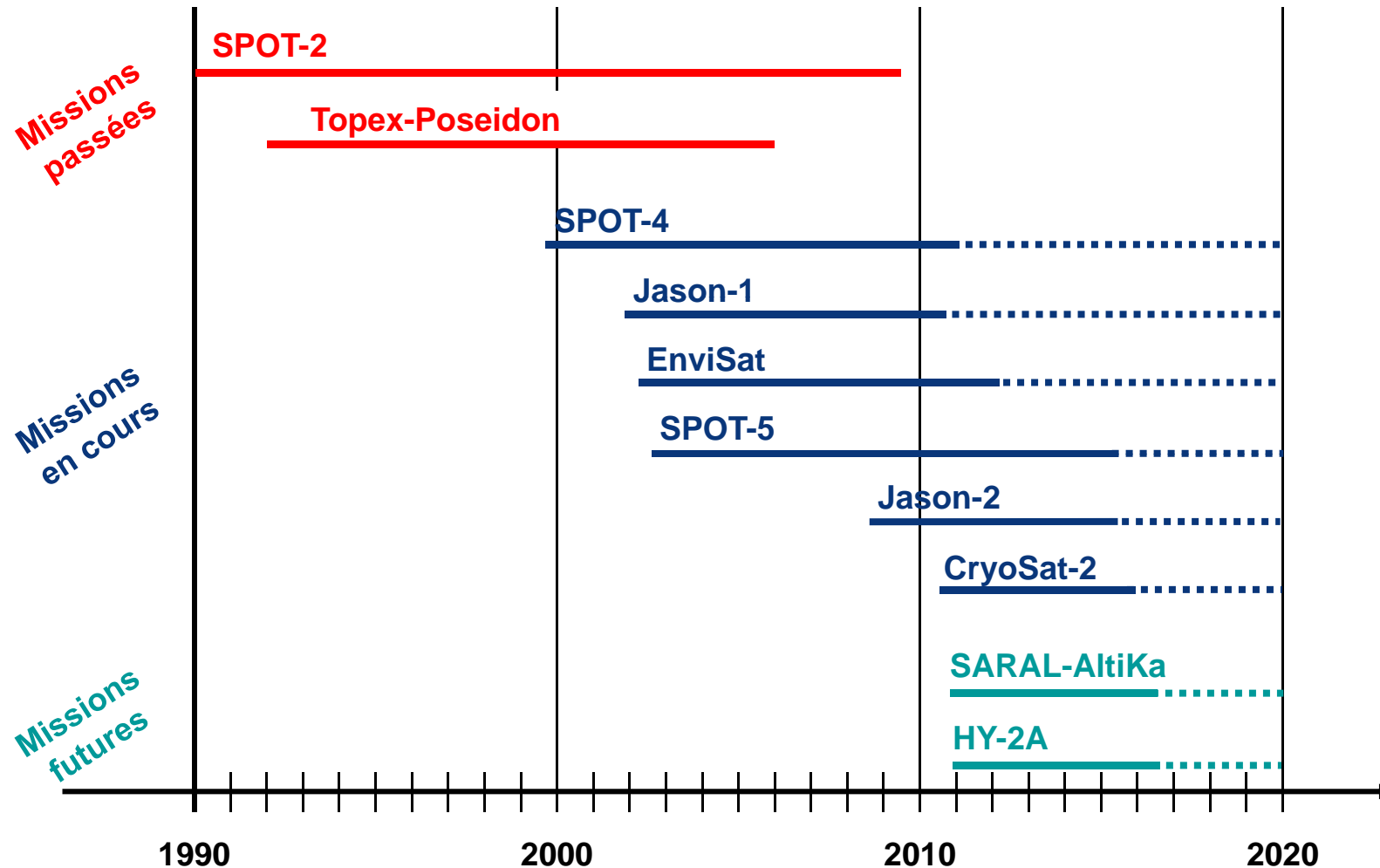
■ Amélioration de l'outil d'extraction

■ Développement d'outils d'automatisation de l'ingestion

Bilan de la migration

- Effort de migration bien plus important que prévu initialement :
 - ◆ Durée : **18 mois** (année 2007 + moitié 2008)
 - ◆ Charge CNES : ~ **1.5 homme*ans** (+ équipe GIP + industriels)
- Impact sur les différentes fonctions :
 - ◆ **Stockage** : occasion de mettre en place une politique de gestion de l'archive
 - ◆ **Acquisition** : production des métadonnées associées aux objets de données sous forme de fichiers XML
 - ◆ **Accès** :
 - meilleure adéquation au besoin des utilisateurs (2 catalogues)
 - développement d'une IHM Web à partir de l'IHM SIPAD-NG
 - amélioration de l'outil d'extraction
 - ◆ **Gestion des données** :
 - Très gros travail de mise à niveau du catalogue ...
 - ... qu'il faut reprendre légèrement suite au retour d'expérience de la part de l'équipe scientifique

Service de diffusion des produits d'altimétrie et d'orbitographie du CNES



DORIS

(Détermination d'Orbite et Radiopositionnement Intégré par Satellite)



DORIS

(Détermination d'Orbite et Radiopositionnement Intégré par Satellite)

Le système DORIS est basé sur l'effet Doppler:

Le satellite est à la **verticale** de la balise, c'est le point TCA (*Time of Closest Approach*)
La fréquence du signal reçu est **égale** à la fréquence du signal émis.

Le satellite **s'approche** de la balise :
La fréquence du signal reçu est **supérieure** à la fréquence du signal émis.

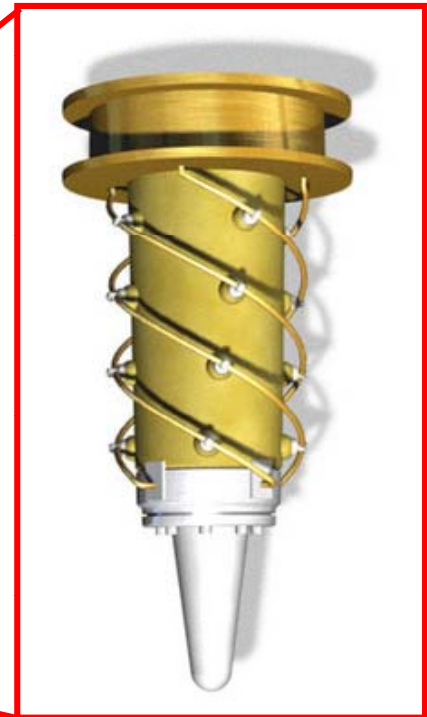
Le satellite **s'éloigne** de la station terrestre :
La fréquence du signal reçu est **inférieure** à la fréquence du signal émis.

Détermination de l'orbite précise du satellite

Antenne réceptrice Doris

Positionnement précis de la balise

Balise terrestre émettrice



Principe de l'altimétrie

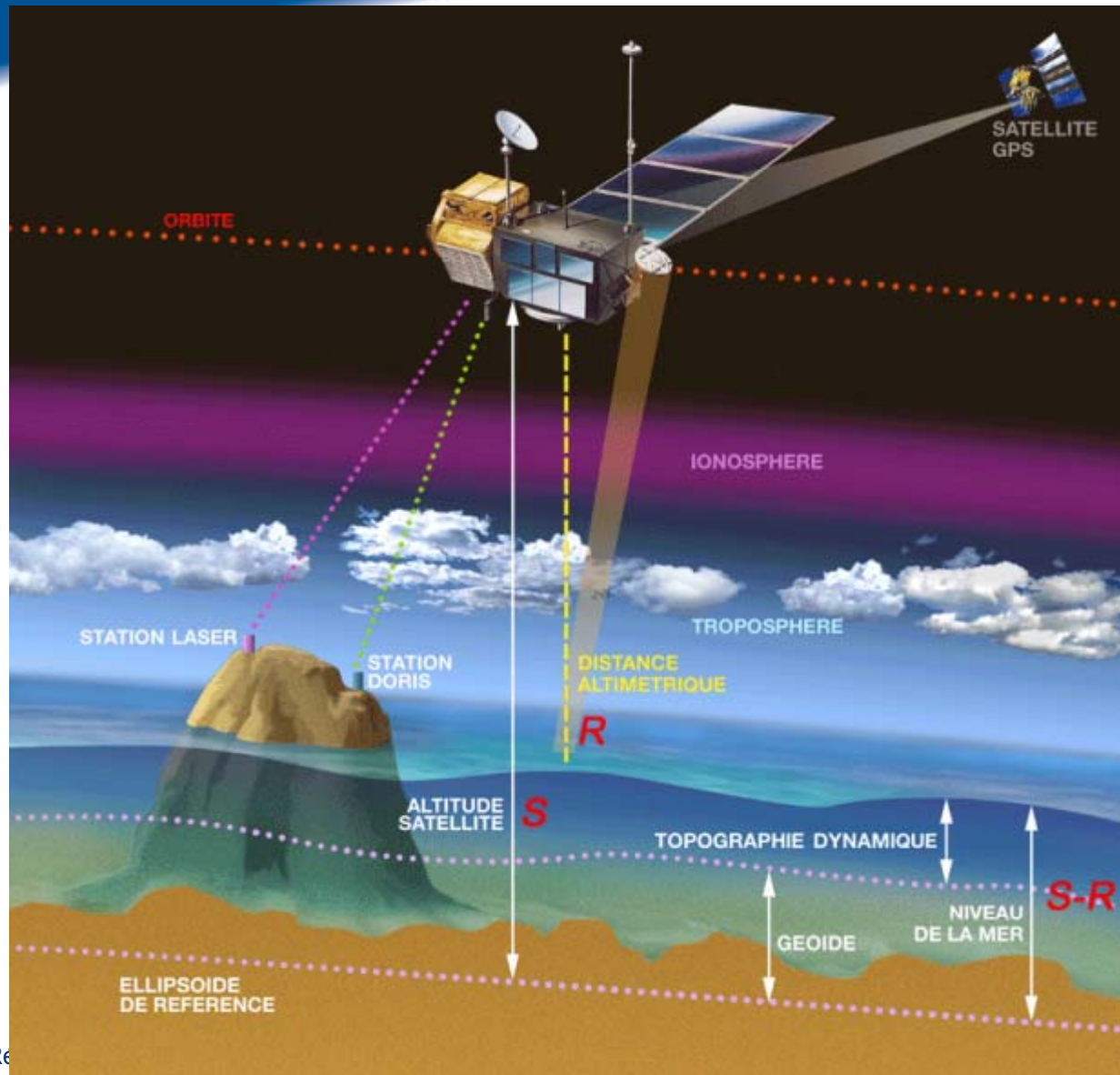
L'altimètre émet une onde radar et l'analyse après réflexion sur la surface:

- temps de trajet aller-retour => distance altimétrique R

S = orbite du satellite

$S-R$ = niveau de la mer:

- géoïde = surface de la mer qu'elle aurait en absence de toutes perturbations (vent, marées, courants, etc.)
- circulation océanique = partie permanente stationnaire (circulation permanente liée à la rotation de la terre, ...) + partie fortement variable (vents, ...)



Situation avant migration

■ 1998 – 2010 : Catalogue AVISO :

- ◆ Pas de fonction d'archivage centralisé dans le système AVISO : chaque sous-système prend plus ou moins en charge l'archivage STAF, le catalogage et la diffusion de ses produits,
 - Archivage en N exemplaires,
 - Perte des données
- ◆ Accès pour la communauté : 1 système d'accès, le Catalogue AVISO (système vieillissant),

■ Septembre 2007 : début de mise en place du SIPAD-NG pour:

- ◆ remplacer la Catalogue AVISO pour la fonction de diffusion,
- ◆ créer une fonction d'archivage commune pour l'ensemble des sous-systèmes AVISO

Déroulement de la migration

■ Développement d'une fonction SIPAD-NG générique supplémentaire

- ◆ Acquisition des données
- ◆ Génération des métadonnées et catalogage
- ◆ Archivage au STAF

■ Constitution du catalogue: démarche analogue à Mercator

- ◆ Définition du dictionnaire de données
- ◆ Identification des jeux à cataloguer
- ◆ Production manuelles des descripteurs des jeux de données
- ◆ Données acquises de 1998 à 2009 (avant mise en service du SIPAD-NG):
production des descripteurs d'objets à partir du contenu STAF et de la base Oracle du Catalogue AVISO

■ Développement de l'IHM Utilisateur

- ◆ Prototypage en interne CNES puis industrialisation

Bilan de la migration

- Effort de migration bien plus important que prévu initialement :
 - ◆ Durée : **toujours en cours** (depuis mi-2007)
 - ◆ Charge CNES : **~ 3 hommes*ans** (+ support industriel important)
- Impact sur les différentes fonctions :
 - ◆ **Stockage** : refonte complète de l'archive STAF
 - ◆ **Acquisition** : développement de nouvelles fonctions du SIPAD-NG
 - ◆ **Accès** :
 - Meilleure adéquation au besoin des utilisateurs (recherche produits, modes de livraison, information produits, ...)
 - Nécessité de faire évoluer le SIPAD-NG en terme de performances compte-tenu du volume et de la quantité de données
 - ◆ **Gestion des données** :
 - Très gros travail sur le catalogue (beaucoup de jeux de données différents et d'objets)

Bilan de ces expériences

■ Migrations rendues nécessaires par :

◆ **Évolution des matériels :**

- Exemple CNES : bandes → 1^{ère} infrastructure STAF → 2^{ème} infrastructure STAF

◆ **Évolution du Web qui oblige à :**

- Faire évoluer la fonction d'accès aux données
- Faire évoluer les catalogues pour rendre les Centres de Données accessibles par des services externes (notions d'interopérabilité, d'Observatoires Virtuels)

◆ **Évolution des disciplines scientifiques :**

- Évolution des thésaurus, des systèmes de classification, etc.
- Apparition de besoins d'interdisciplinarité
- → Nécessité de faire évoluer les catalogues

■ Périodicité :

- Stockage : **tous les 10 à 15 ans**
- Gestion des données (catalogue) : **tous les 5 à 7 ans**
- Accès : **en permanence**

■ Efforts de migration toujours très importants

■ Dans le cas d'une archive OAIS, l'impact de la migration d'une fonction sur les autres fonctions peut être limité