

Suivi de réseaux d'observations océanographiques avec le langage de script Perl



Philippe Téchiné, Bruno Buisson, Laurent Testut, Thierry Delcroix
LEGOS / OMP 14 Avenue E. Belin 31400 Toulouse



Réseaux d'observations

Le LEGOS est un laboratoire pluridisciplinaire avec une forte part d'océanographie. Il assure la tâche de 4 Services d'Observations, dont les réseaux de mesure in situ ROSAME et SSS:

ROSAME est le Réseau d'Observation Subantarctique et Antarctique du niveau de la Mer (programme international GLOSS). Il est composé d'une dizaine de marégraphes (Fig. 1) implantés sur les Terres Australes et Antarctiques Françaises. Les applications scientifiques concernent principalement l'étude des marées océaniques et des variations du niveau de la mer, ainsi que la validation d'altimètres satellitaires.

SSS, Sea Surface Salinity, est un réseau d'observation de la salinité de surface océanique (programme international GOSUD). Il est basé sur des thermosalinographes (Fig. 2) embarqués sur une quinzaine de navires marchands sillonnant tous les océans. Les applications scientifiques concernent principalement l'étude de la variabilité du climat et du cycle de l'eau, ainsi que la validation du satellite SMOS.



Figure 1. Marégraphe de Kerguelen.

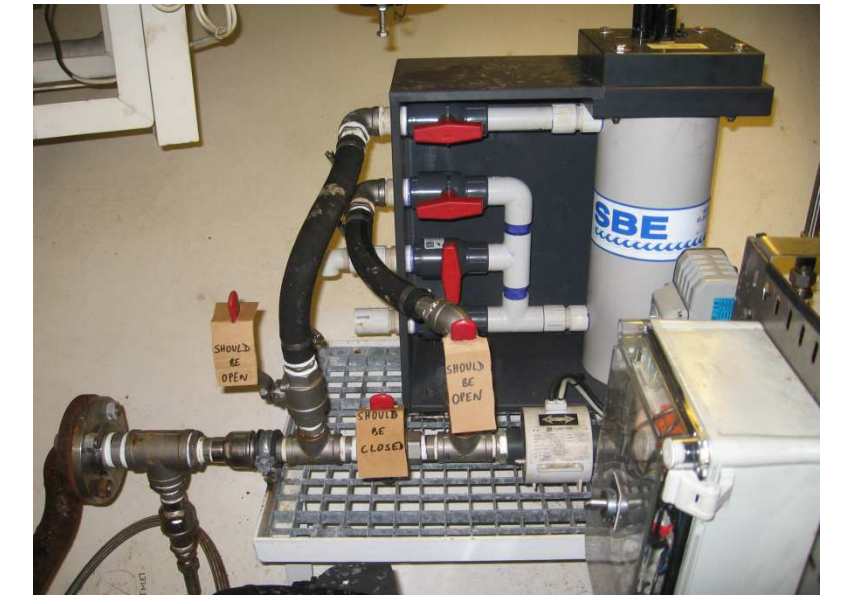


Figure 2. Thermosalinographe du navire Nuka Arctica.

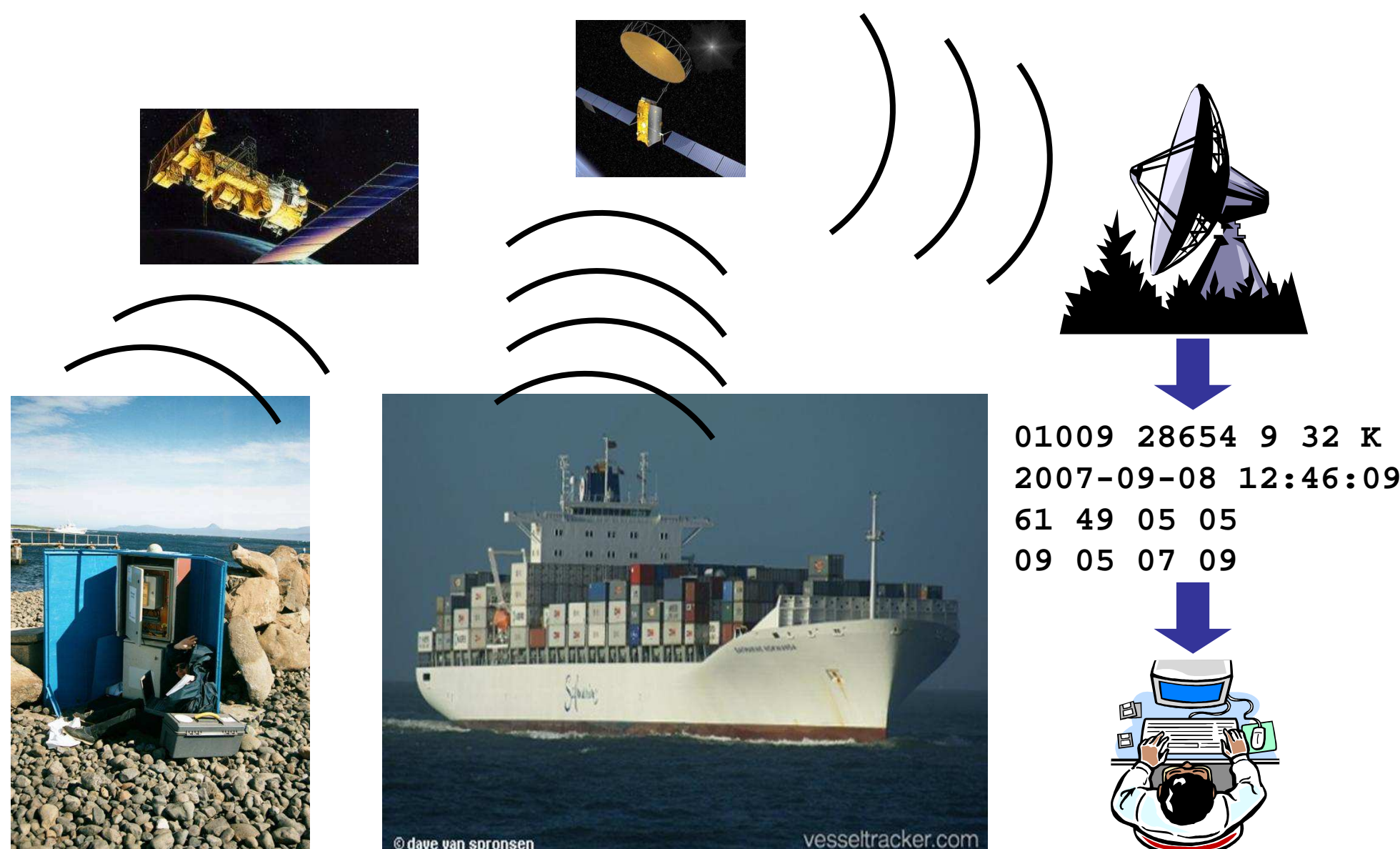


Figure 3. Acquisition et transmission des mesures.

Acquisition et transmission des mesures

Sur chaque site, des stations d'acquisition réalisent automatiquement des mesures: pression atmosphérique, pression de fond, température et conductivité pour ROSAME, température, salinité et position du navire pour SSS. Ces mesures sont regroupées dans des messages transmis au LEGOS via les systèmes satellite Argos pour ROSAME, Inmarsat ou Iridium pour SSS (Fig. 3). Les données transmises sont également enregistrées sur chaque site.

Objectifs du suivi

Conçu au départ pour ROSAME, le suivi a été ensuite appliqué à SSS grâce à un travail de fond sur des modules génériques écrits avec le langage Perl qui permet d'avoir un traitement automatisé, sans intervention humaine, dont voici les principaux objectifs:

- ✓ avoir un code générique pour acquérir, traiter et archiver les messages, qui puisse s'adapter à l'évolution du matériel;
- ✓ avoir des contrôles pendant tout le traitement des données;
- ✓ faire des retours vers les gestionnaires des réseaux d'observations en cas de problème sur un site de mesure;
- ✓ assurer une distribution rapide des données aux utilisateurs;
- ✓ pouvoir faire un suivi des mesures sur Internet.

Choix du langage Perl

Pour atteindre les objectifs du suivi, le langage Perl a été choisi pour les raisons suivantes:

- ✓ c'est un langage multi plateforme et très complet;
- ✓ un très grand nombre de bibliothèques, couvrant un nombre important de domaines d'application, sont disponibles gratuitement sur le site web du CPAN (Comprehensive Perl Archive Network: <http://www.cpan.org/>);
- ✓ c'est un langage modulaire dans lequel les notions d'héritage et de classe existent et ont été utilisées (approche objet) (Fig. 4);
- ✓ c'est un langage polyvalent qui couvre toutes les étapes du traitement:
 - il possède toutes les fonctions nécessaires pour séparer et transformer les données: il est très bien adapté pour scanner les messages (modules Mail::MBoxParser, MIME::Parser), décoder la télémétrie (expressions régulières) et traiter les mesures (fonctions pack/unpack);
 - les mécanismes de gestion des exceptions (fonctions eval et die) ont été utilisés pour les remontées d'alerte par mail;
 - il est très populaire pour la création de code HTML grâce à son module CGI;
 - les feuilles de style (CSS) ont été intégrées facilement dans le code Perl.

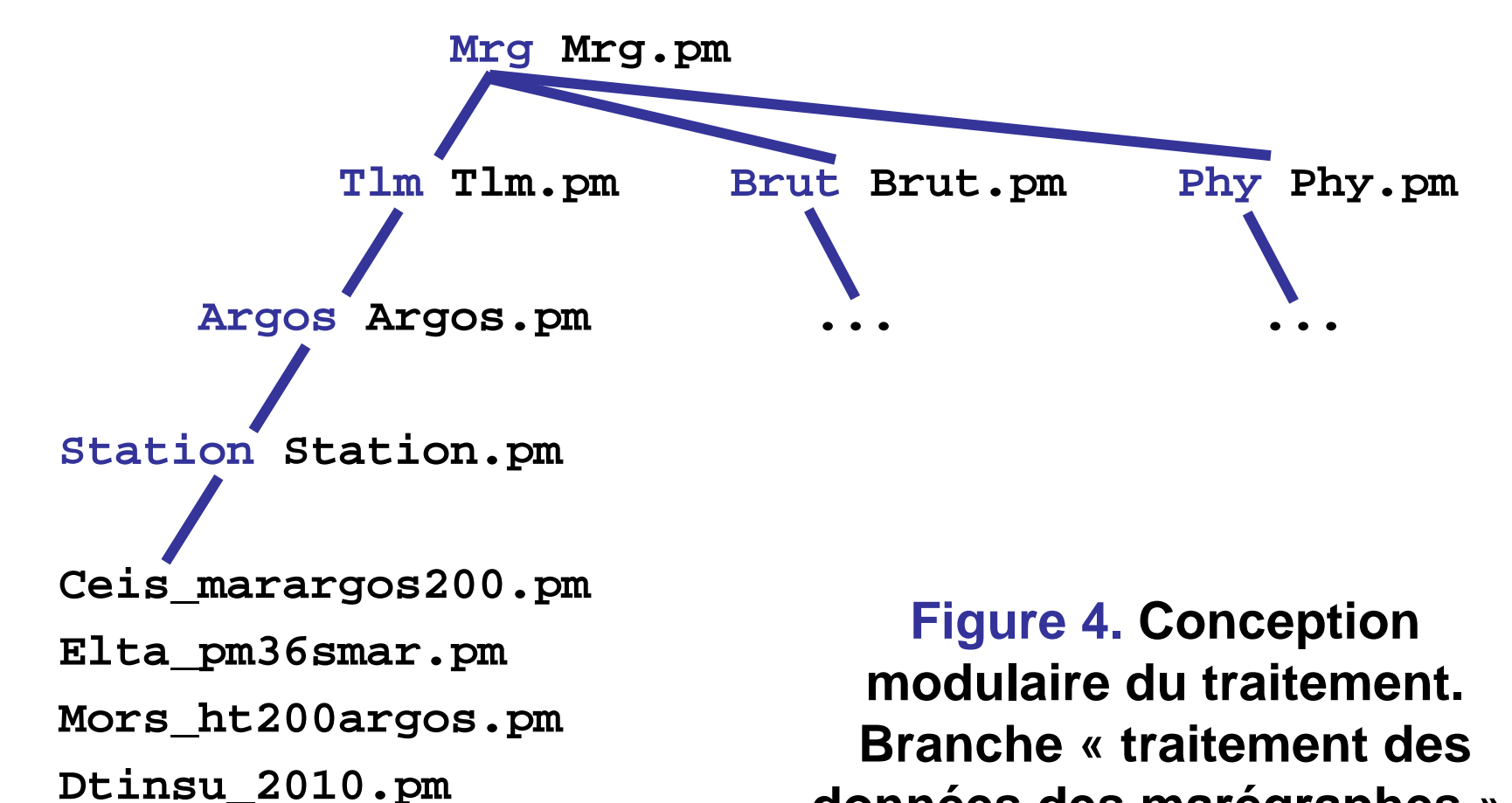


Figure 4. Conception modulaire du traitement. Branche « traitement des données des marégraphes ».

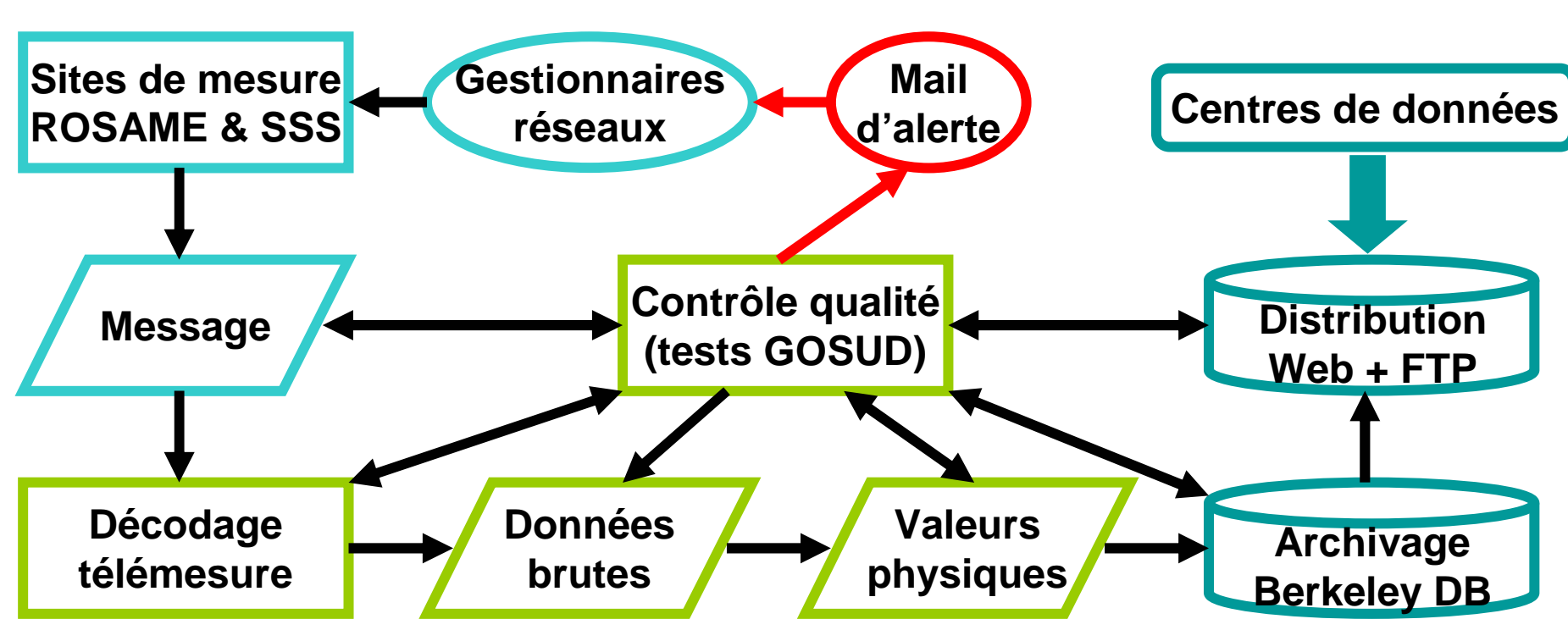


Figure 5. Traitement, contrôle qualité, distribution automatique.

Traitement, contrôle qualité, distribution automatique des données

Chaque jour, plus de 200 messages sont reçus en provenance des sites de mesure et déclenchent automatiquement un logiciel (Fig. 5) basé sur des modules écrits avec le langage Perl. Le logiciel est conçu de façon à ce que l'on puisse ajouter de nouveaux capteurs ou intégrer un nouveau système de transmission sans avoir à changer le cœur du programme.

Le contrôle qualité automatique est effectué à chaque étape du traitement. Lorsqu'un problème survient, la messagerie électronique est utilisée comme un système d'alarme pour avertir automatiquement les gestionnaires des réseaux d'observations afin d'intervenir rapidement sur un site de mesure. Toutes les alertes envoyées par mail sont archivées dans des fichiers.

Les données sont mises à jour automatiquement sur le site FTP du LEGOS pour permettre une distribution rapide aux centres de données: Coriolis, SIMMER, SONEL, BODC, ODINAfrica, UHSLC.

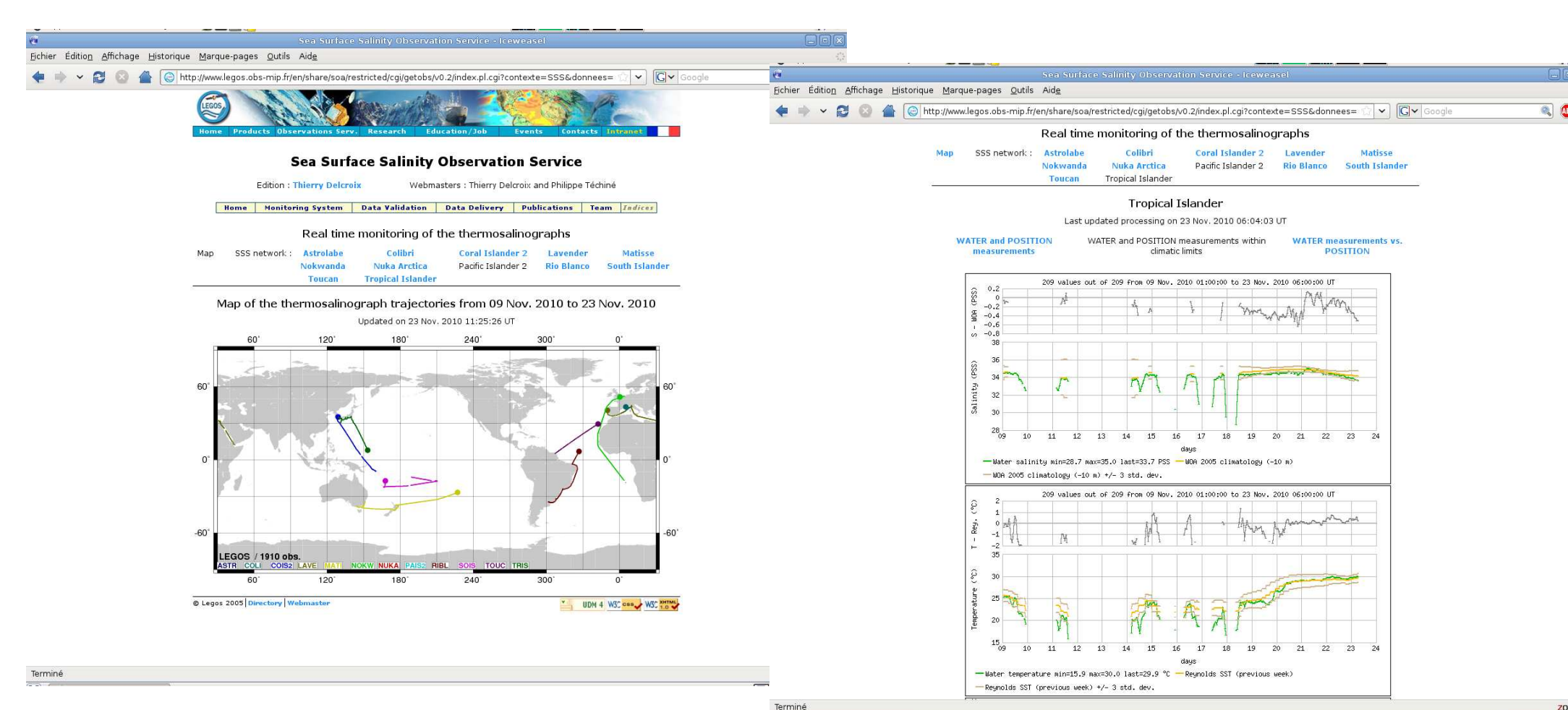
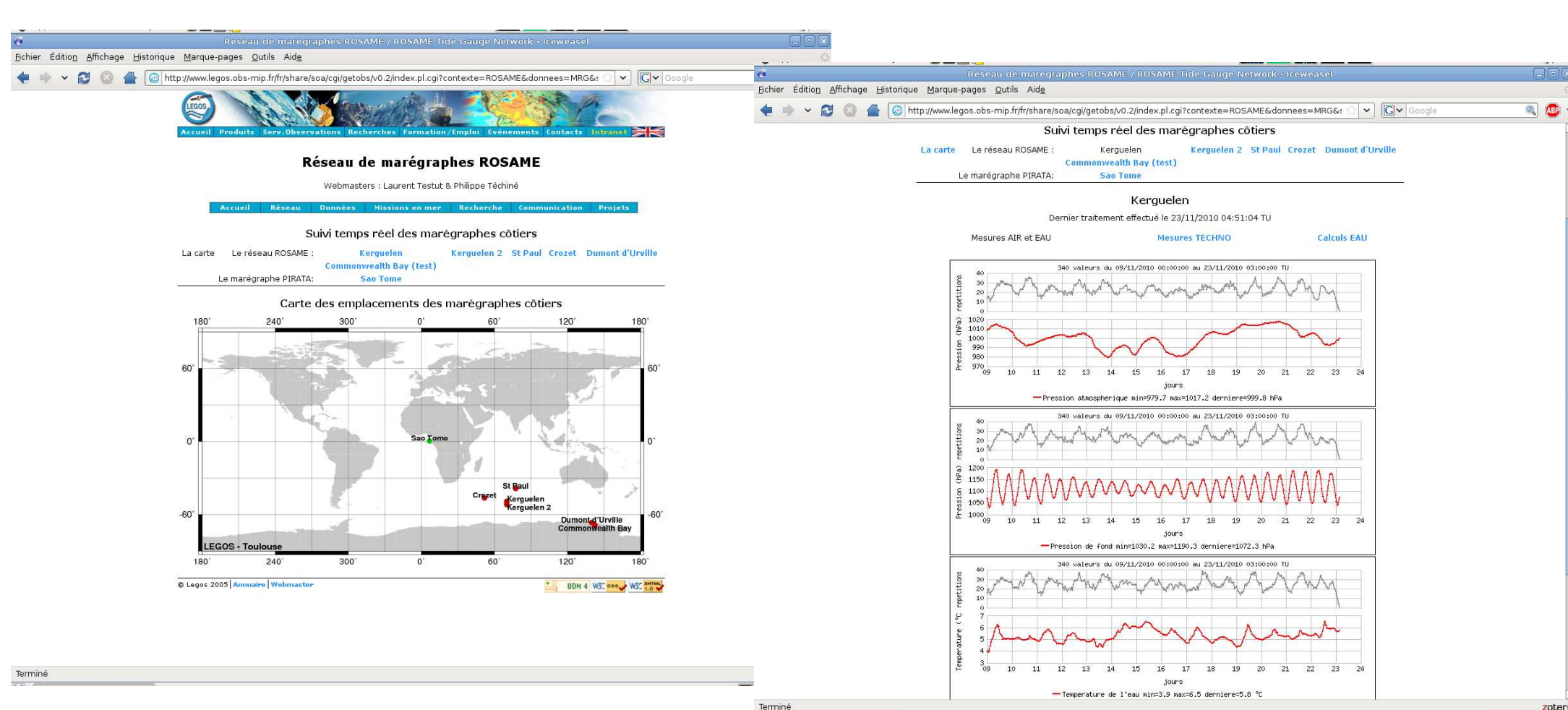


Figure 6. (gauche) Positions des marégraphes et courbes des mesures pour ROSAME. (droite) Trajets des navires et courbes des mesures pour SSS.

Suivi des réseaux d'observations

Le suivi des réseaux d'observations est disponible sur le site web du LEGOS dans des pages web dynamiques (Fig. 6) (module CGI). Celles-ci rassemblent les cartes des positions des marégraphes et des trajets des navires (script Perl appelant les fonctions de la bibliothèque graphique Generic Mapping Tools), les courbes des mesures (module GD::Graph), ainsi que des informations pour un suivi opérationnel des stations d'acquisition.

Plan qualité

Des indicateurs qualité (Fig. 7), mis à jour automatiquement chaque mois sur le site web du LEGOS, permettent de suivre sur le long terme l'évolution de la qualité des mesures reçues.

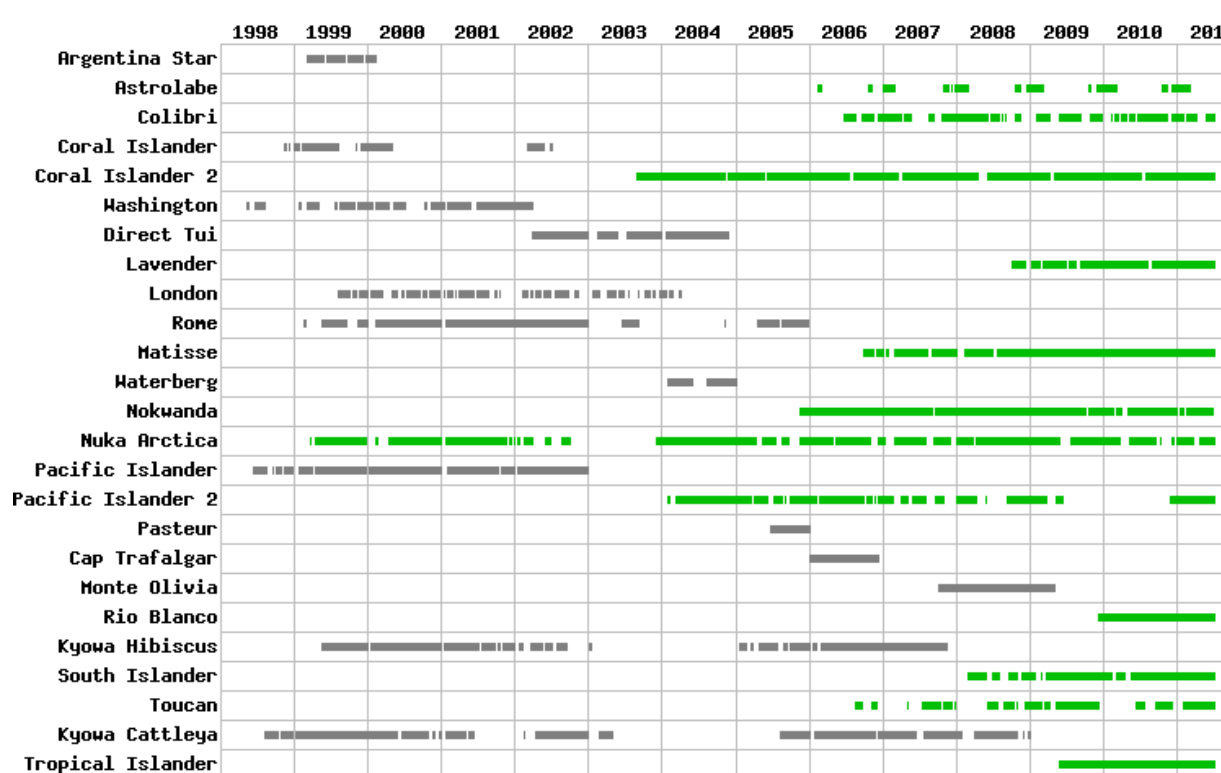
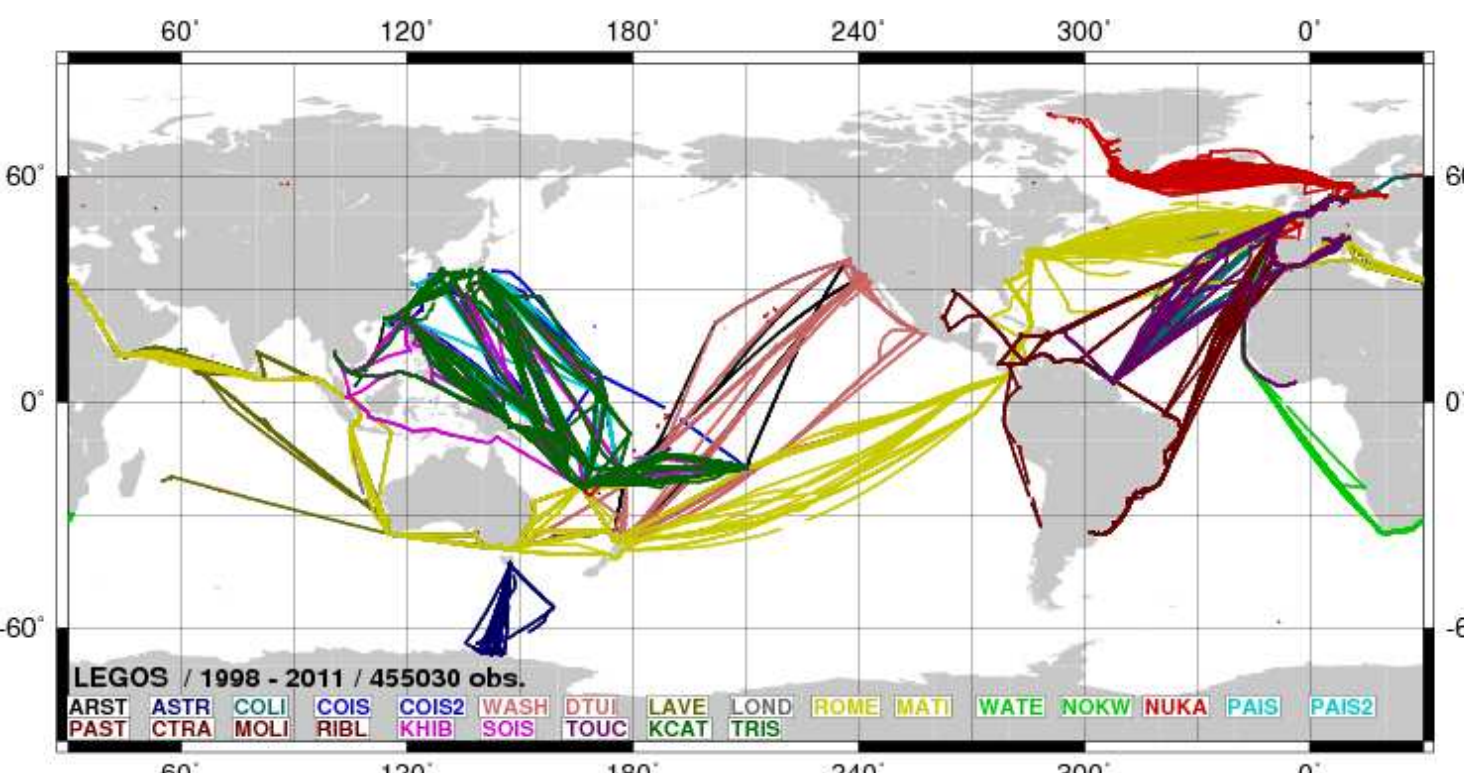
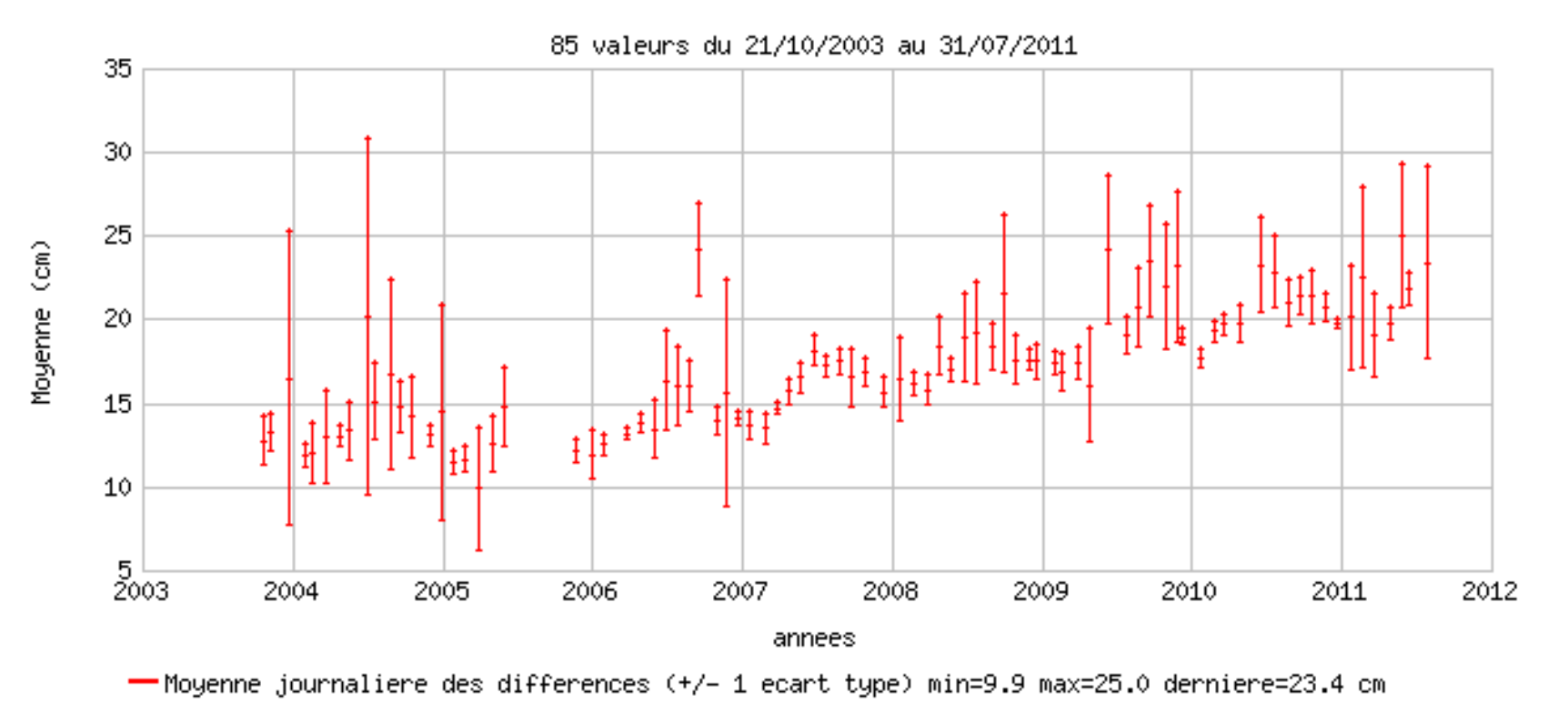


Figure 7. Quelques indicateurs qualité. (gauche et milieu) Distributions spatiale et temporelle SSS 1998-2011. (droite) Suivi des dérives du marégraphe de Kerguelen 2003-2011.



Bilan

L'automatisation du traitement avec le langage Perl a permis de réduire les sources d'erreur et d'augmenter la capacité de traitement. Sur un plan technique, cela améliore la surveillance des capteurs et fiabilise les réseaux de mesure, dans des endroits d'accès difficile pour ROSAME et pour des navires ne restant que quelques heures à quai pour SSS. Sur un plan scientifique, cela permet de minimiser la perte des données et de les distribuer en un minimum de temps aux centres de données impliqués dans l'océanographie opérationnelle. L'interface web permet d'avoir un système de supervision décentralisé: les gestionnaires des réseaux peuvent suivre à tout moment et de partout les capteurs.

Nouveaux enjeux, nouveaux outils

Depuis que le suivi a été mis en place, le nombre de messages reçu a beaucoup augmenté: 39000 en 2005, 47000 en 2007, 75000 en 2009, et déjà plus de 60000 messages reçus cette année (mi-septembre 2011). Le traitement d'un nombre plus important de mesures représente désormais un nouvel enjeu. Les données sont de plus en plus utilisées: le nombre de pages web consultées et de fichiers téléchargés augmente régulièrement. Pour répondre aux besoins futurs avec un nombre de mesures qui va continuer à augmenter, il est prévu de stocker les données dans un système de gestion de base de données PostgreSQL. Cela permettra d'augmenter la volumétrie des données, de standardiser les accès à la base de données en utilisant le langage de requête SQL, et de renforcer la gestion des verrous. Pour la réalisation des pages web dynamiques, le passage au framework Perl de développement web Catalyst (modèle MVC) permettra de continuer à exploiter les modules Perl fournis par le CPAN. Catalyst supporte un grand nombre de base de données grâce au module DBI qui donne la possibilité de programmer en Perl pour générer les accès SQL et de découpler l'accès à la base de données, donc de faire évoluer indépendamment le stockage. Tout ceci devrait permettre d'améliorer encore et de compléter le suivi des réseaux d'observations océanographiques.