

Cahier des charges pour le projet CO_2 Atmosphérique

Jonathan Schaeffer

20 avril 2015

Résumé

Ce cahier des charges décrit les besoins exprimés dans le cadre du stage de Claude Vro à l'IUEM, pour la gestion des données de CO_2 atmosphérique.

Table des matières

Table des matières	1
1 Import des données	2
1.1 Import manuel des données	2
1.2 Visualisation des données	2
1.3 import de l'étalonnage et modification des mesures brutes	2
1.4 import initial de toutes les données	3
2 Qualification	3
2.1 Modélisation de la qualification	3
2.2 règles de qualification automatique	3
2.3 interface de qualification manuelle	3
3 Analyse	3
3.1 interface avec l'outil d'analyse	3
4 Rapports périodiques	4
5 Déploiement et re-déploiement	4

6	Interfaçages et interopérabilité	4
6.1	système d'import automatisé des données	4
6.2	exploration de la faisabilité d'interopérabilité	4
6.3	affichage dans un site plone	5
6.4	système évolué pour la qualification experte des données? . . .	5

1 Import des données

1.1 Import manuel des données

L'opération de mesure génère un fichier texte pour l'ensemble des mesures. Il s'agit de :

- créer une interface web pour l'importation de ce fichier
- mettre en base de donnée des mesures correspondant aux informations téléchargées

L'opération de chargement doit tenir compte du pas de temps que les administrateurs de la série temporelle ont retenu. Il y aura une opération de moyenne sur ce pas de temps à déterminer avec Peggy Rimmelin.

1.2 Visualisation des données

Il s'agit de créer une interface web pour visualiser les données en base. Cette interface doit permettre de sélectionner le ou les paramètres à afficher et la période de temps.

1.3 import de l'étalonnage et modification des mesures brutes

Les mesures chargées dans le systèmes sont dans un état brut. C'est à dire qu'elles ont besoin qu'on leur applique un coefficient suite à une opération d'étalonnage. L'opération d'étalonnage génère un fichier au même format que les mesures. Les valeurs dans ce fichier permettent d'appliquer un coefficient à un ensemble de mesures brutes. Il s'agit de :

- déterminer avec Peggy Rimmelin comment effectuer l'opération d'étalonnage
- implémenter l'automatisation de l'opération d'étalonnage à partir du chargement d'un fichier d'étalonnage

1.4 import initial de toutes les données

Depuis le début du projet, une grande quantité de données a déjà été collectée. Il s'agit donc d'importer toutes les mesures déjà qualifiées (ou non), déjà étalonnées (ou non) dans la base de donnée en utilisant le système de **seeding** de rails. Il faut déterminer avec Peggy Rimmelin les différents cas de figure des mesures à importées (déjà qualifiées, déjà étalonnées?)

2 Qualification

2.1 Modélisation de la qualification

- une table pour les codes qualité
- une mesure a un code qualité (:has_{one} ou :belongs_{to})

2.2 règles de qualification automatique

Un code qualité est attribué à une mesure en fonction de règles dont certaines sont automatisables. Il s'agit de :

- déterminer avec Peggy Rimmelin les règles automatisables
- implémenter les règles de qualification des données
- déclencher l'application de ces règles au bon moment (à la suite de l'étalonnage ou par une action manuelle?)

2.3 interface de qualification manuelle

Toute l'opération de qualification ne peut pas être automatisée. Il faut développer une IHM web pour l'aide à la qualification :

- mise en évidence des données non qualifiées
- sélection par zone pour qualification

idéalement, cette IHM représente un courbe des valeurs pour l'interaction.

3 Analyse

3.1 interface avec l'outil d'analyse

La phase d'analyse met en jeu des algorithmes statistiques implémentés dans un outil externe (R) et utilisé par Peggy Rimmelin. Il s'agit d'interfacer

cet outil avec le système centralisé, idéalement en utilisant l'API REST et le format d'échange JSON. Cet interfaçage permet de :

1. aspirer les données à analyser
2. stocker les résultats d'analyse

4 Rapports périodiques

Valoriser des résultats d'analyse en créant automatiquement un rapport périodique et des graphs instantanés, à l'image d'un bulletin météo. Il s'agit d'utiliser les résultats d'analyse pour générer des graphiques sur une page web qui puisse être exportée dans quelques formats courants (PDF, ODT).

5 Déploiement et re-déploiement

Il faut pouvoir déployer ce service facilement, afin que ce travail puisse servir :

- pour d'autres projets locaux de gestion des séries temporelles
- pour d'autres sites souhaitant gérer leur série CO2 de la même manière.

6 Interfaçages et interopérabilité

Cette partie sera exploratoire et expérimentale.

6.1 système d'import automatisé des données

Tout d'abord, on pourrait trouver une automatisation totale de l'import des données et de l'étalonnage.

L'import des données reste manuel, mais facilité par l'interface web. Quelles solutions peut-on mettre en oeuvre pour :

- importer automatiquement les données (API REST et automatisations de windows)
- s'assurer que l'import automatisé est toujours opérationnel

6.2 exploration de la faisabilité d'interopérabilité

Ensuite, la série de données peut potentiellement faire partie d'un réseau de séries. Dans cette perspectives, il convient d'explorer les possibilités d'interopérabilité du système pour des systèmes de moissonnage automatique.

6.3 affichage dans un site plone

Afin de valoriser cette série, on voudrait afficher des graphiques de mesures sur le site web de l'observatoire, en Plone. Il s'agirait, via l'interface REST ou via les standards d'interopérabilité, de créer une page web qui permette d'interroger les mesures en temps réel.

6.4 système évolué pour la qualification experte des données ?

Explorer la possibilité de brancher des algorithmes d'intelligence artificielle (réseaux de neurones par exemple) pour faire le travail de qualification de la donnée.