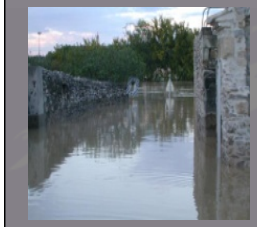
Observatoire de Recherche Méditerranéen de l'Environnement

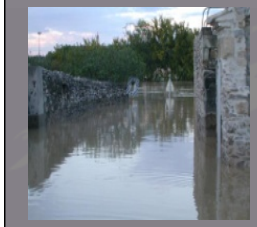
# Outil web interactif de visualisation et validation de séries temporelles

*J. Fabre – O. Lobry – D. Lopa*



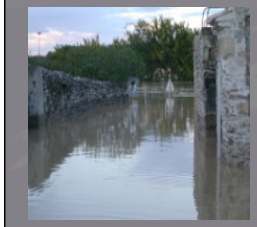
# Problématique

- >80% des données : séries numériques
- Besoins (producteurs et utilisateurs)
  - Visualisation
  - Navigation
  - Export
  - Validation
- Contraintes
  - Facilement accessible (web)
  - Solution réutilisable (mutualisation)
  - Gros volumes de données : 5 ans x Hz x 20 => ~ milliards de mesures, ~ 10 Go



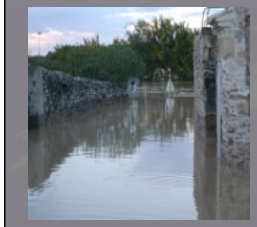
# Choix technologique

- Gestion côté serveur
  - Php (jpGraph, pchart), gplot, GMT, matlab, …
  - Génération à l'avance
  - Images statiques, pas d'interactions
- Côté client (Flash, javascript)
  - Interactif : zoom, déplacement, ...
  - Chargement dynamique
  - Mais transfert des données au client
- Solutions existantes
  - Javascript : Highchart, jqplot, D3, rgraph, zingchart, …
    - Ergonomiques, interactifs, sexy, … mais …
    - peu performants pour les gros volumes de données
  - Flash : délaissé (ex fusionchart), pb support smartphone



# Dygraphs (http://dygraphs.com)

- Dédié aux séries temporelles
- Bon niveau d'interactions (zoom, déplacement, )
- Très bonne gestion du temps (labels)
- Performance : 1 millions de points en moins de 3s
- Intégrable et extensible (callback)



# Présentation structurée des graphiques

## Gravimètre supraconducteur

**Thème (page)**

Accueil / Services d'Observation / SNO H+ / Géodésie en Environnement Karstique / Graphiques /

Intervalle de temps | Données | Lien direct pour cette recherche

**Outils de contrôle**

De 24/06/2015 À 24/09/2015 Resynchroniser

Afficher

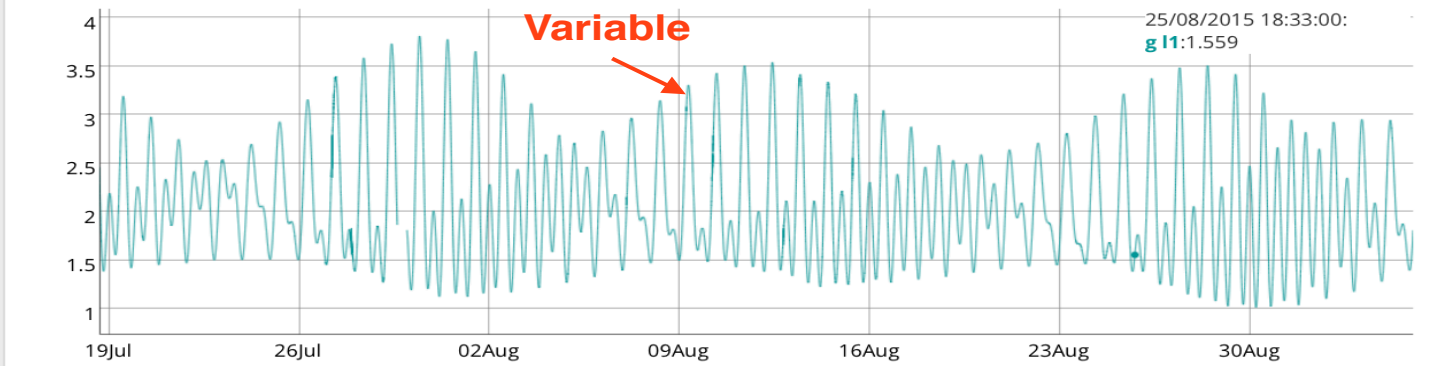
**Groupe (onglet)**

Raw data (L1) | Calibrated data (L2) | Corrected data (L3)

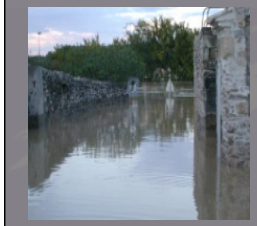
**Graphe**

Gravity (V)

**Variable**




Barometric pressure (mBar)



# Choix des données

## Gravimètre supraconducteur

Accueil / Services d'Observation / SNO H+ / Géodésie en Environnement Karstique / Graphiques / Gravimètre supraconducteur

Intervalle de temps | Données | Lien direct pour cette recherche

☐ Toutes les données

### Raw data (L0)

- ☐ Gravity (V)
- ☐ Barometric pressure (mBar)

### Raw data (L1)

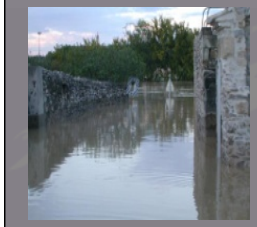
- ☑ Gravity (V)
- ☑ Neck temperature 1 (K)
- ☑ Neck temperature 2 (K)
- ☑ Barometric pressure (mBar)
- ☑ grav bal (V)
- ☑ tilt y bal (V)
- ☑ temp bal (V)
- ☑ grav ctrl (V)
- ☑ tilt x ctrl (%)
- ☑ tilt y ctrl (%)
- ☑ body t (K)
- ☑ belly t (K)
- ☑ dewr pwr (DAC)

### Corrected data (L2)

- ☑ Gravity (nm.s-²)
- ☑ Barometric pressure (mBar)

### Calibrated data (L3)

- ☑ Gravity (nm.s-²)
- ☑ Barometric pressure (mBar)



# Lien direct

Intervalle de temps | Données | Lien direct pour cette recherche

Voici le lien à copier :

http://data.oreme.org/gek/gek/gravity_sup.../?graph-gek_gravi_supra_L1-grav_l1=1&graph-gek_gravi_supra_L1-neckt_1_l1=1&graph-gek_gravi_supra_L1-neckt_2_l1:

Ctr-V Ctr-P

Salut,

J'ai trouvé une corrélation étrange entre ces deux variables…

Tu peux jeter un œil ?

A+
Valentine

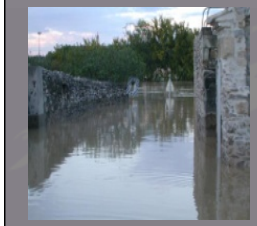
Clic

Ctr-V Ctr-P

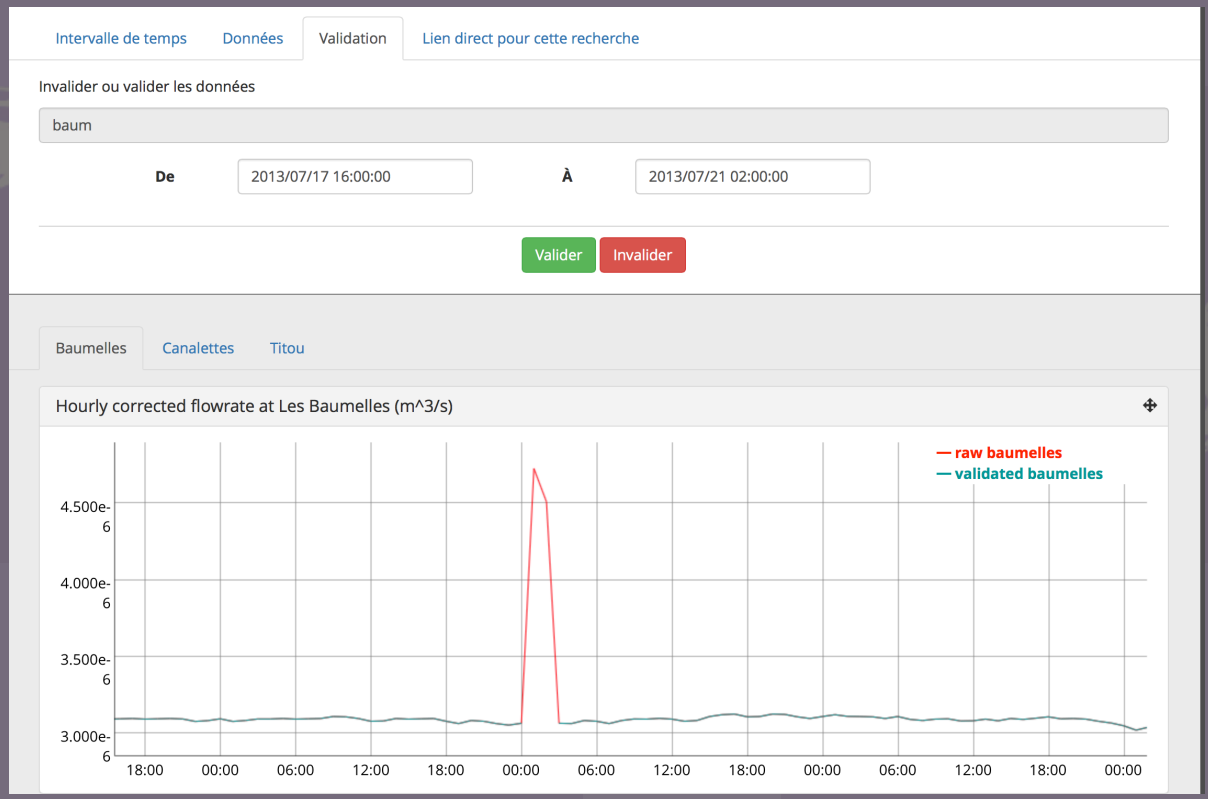
3.5 xxxxx
xxxxxxx.xxxxxxxx
xx. Xxxxxxxxxx
xxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxx
xx.

xxxxxxxx xxx xxxxx

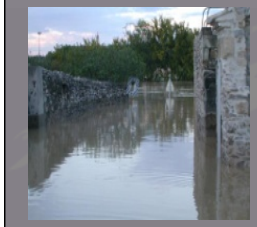
Clic



# Validation des données




- Les données ne sont jamais effacées
- Table d'intervalles de données invalides



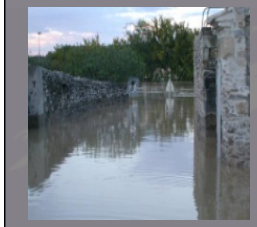
# Export des données

Export des données | Types d'instrument | Campagnes de mesure | Carte des stations

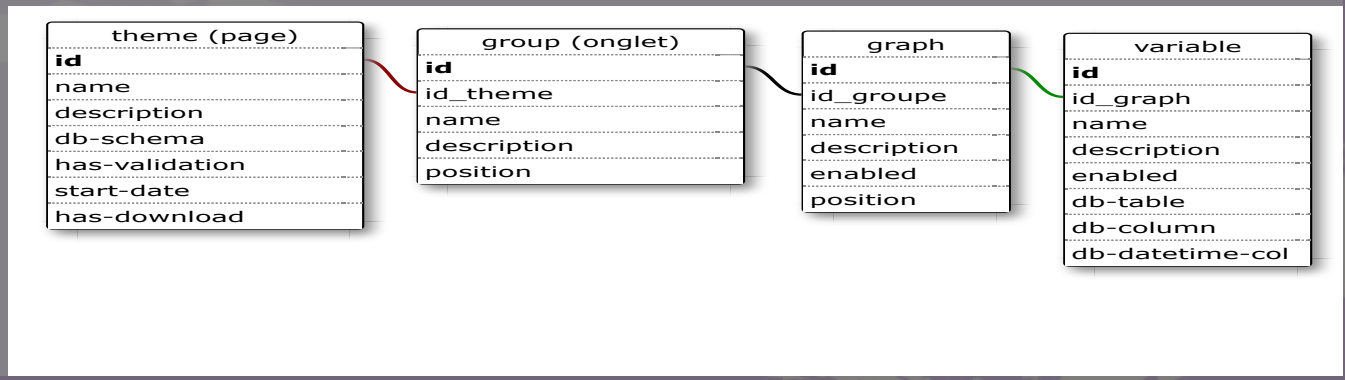
## Conditions d'utilisation

Les organismes et institutions qui financent la collecte des données et les centres de données doivent être remerciés dans toute publication ou communication utilisant les données, avec la phrase suivante :
"Toutes les mesures présentées ici proviennent de l'Observatoire GEK, qui est l'observatoire géophysique des eaux karstiques hébergé par l'OSU OREME (http://www.oreme.org) et le SNO H+ (http://hplus.ore.fr/). L'Observatoire GEK et son instrumentation sont principalement financés par l'INSU, l'ANR, l'UM2, RESIF et la région Languedoc-Roussillon".

| Onglet | Données | |
|---|---|---|
| Raw data (L1) | Gravity | Télécharger |
| Raw data (L1) | Neck temperature 1 | Télécharger |
| Raw data (L1) | Neck temperature 2 | Télécharger |
| Tous | Informations supplémentaires | Télécharger |
| Tous | Conditions d'utilisation | Télécharger |
| Tous | Tous les fichiers | Télécharger |

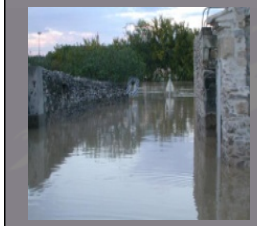


# Présentation modélisée




- Génération automatique des pages de graphiques

# Administration en ligne (CRUD)

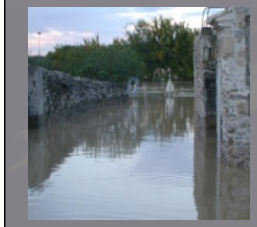


Colonne date | Thème (page) | Groupe (onglet) | Graphique | Variable | Discriminants

Déclarer le thème de la page. Document descriptif

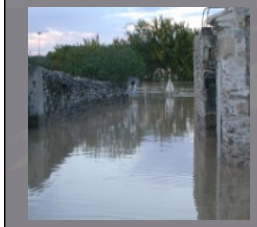
Ajouter Thème (page) | Exporter | Imprimer

| Nom | Schéma BD | Autoriser la validation des données | Autoriser l'export des données | Actions |
|---|---|---|---|---|
| carnoules | carnoules | f | t | |
| cefe_test | cefe | f | t | |
| flowrate | gek | t | t | |
| flowrate_pub | gek | f | t | |
| flux_tower | gek | f | t | |
| gravity_supra | gek | f | t | |
| gravity_supra_pub | gek | f | t | |
| hsm_hydro_chamontin | hsm_hydro | f | t | |
| hsm_hydro_ctd_paiolive | hsm_hydro | f | t | |



# Visualisation de gros volumes de données

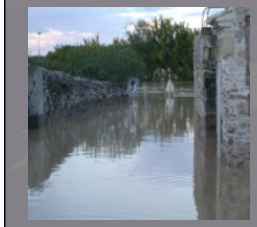
- Stage Daniel Lopa
- Limitations dygraphs
  - Affichage rapide mais chargement des données (très) pénalisant (rapport >10)
  - Latence trop grande au-delà de $10^7$ pts
- Si trop de données
  - Lenteur (expérience utilisateur négative)
  - Surcharge
    - réseau : encombrement non nécessaire
    - Serveur : impact sur la continuité de service
    - Client : risque de plantage (navigateur, voire ordinateur !)

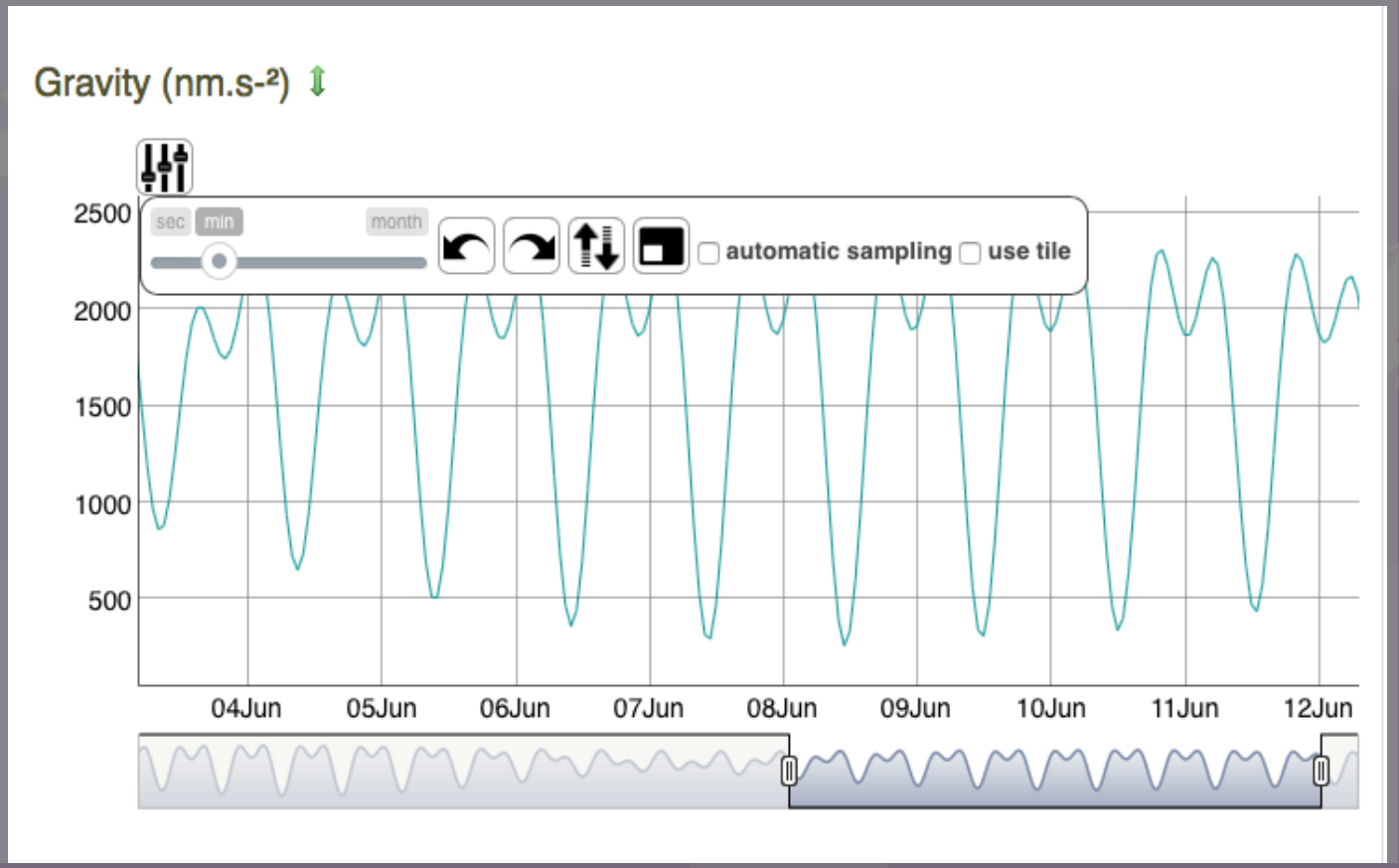


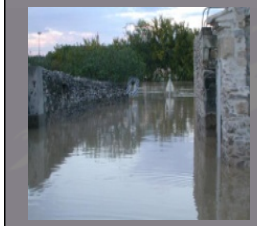
# Chargement progressif des données

- Envoyer des données moyennées (jour, heure, minute, …) en fonction de l'intervalle demandé
- Surcharge à Dygraphs
  - Récupère les évènements de zoom
  - Requête ajax (jQuery) si la résolution change (au moins 1 pt par pixel)
- Similaire au tuilage en webmapping (ex google maps, WMS-T)

# Visualisation de gros volumes de données

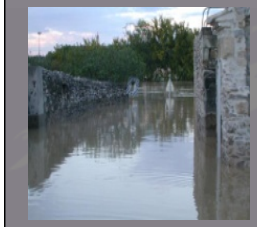






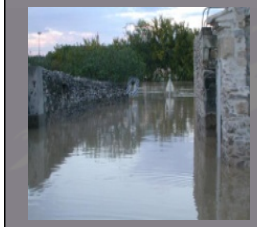
# Calcul des moyennes

- Moyennes calculées sur le serveur à partir des données initiales
  - Automatiquement en php : coûteux
  - Automatiquement en SQL : moins coûteux mais encore trop lent
  - Calculer les moyennes au fur et à mesure
    - Trigger (postgresql)
      - CREATE TRIGGER update_data_day  AFTER INSERT OR UPDATE OR DELETE  ON data FOR EACH ROW  EXECUTE PROCEDURE update_data_day();
    - Vues matérialisées (postgresql >=9.3) ?
      - CREATE MATERIALIZED VIEW data_day AS SELECT date_time, avg(val1) as val1, avg(val2) as val2, … FROM data group by date_trunc('day', date_time) as date_time
  - Le modèle prévoit d'associer une vue pour chaque niveau
  - Surcharge stockage faible (sec => minute : 1,6 %)
  - Surcharge faible compte tenu de l'économie en chargement de données



# Réutilisabilité

- L'approche est complètement réutilisable
  - Base de données relationnelle
  - Php / javascript
  - Dygraphs gratuit et opensource (MIT)
  - Représente env 4 h.m de dev
- Génération d'un nouveau thème en qq mn
- Gestion des gros volumes
  - Nécessite un peu de développement BD (vues matérialisées)
- Réutilisation du code
  - Nettoyage nécessaire
  - Basé sur le framework CodeIgniter (pas très connu) => bundle symfony ?
  - À mettre sur un forge



# Perspectives

- Intégration stage au site de production
- Préchargement des données
- Cache au niveau du navigateur
- Edition intégrée des pages de graphe
- Wizard création de graphes
- Affichage sur tablette / smartphones
- Portage symfony
- Couplage SOS ??