

Schéma base Climato

1. But

Ce document décrit la structure de la base de données Climato. La base de données est stockée dans une base PostgreSQL avec l'extension TimeScaleDB.

2. Historique des mises à jour

- 1.0 : 29-05-2022 : Version Initiale
- 1.01 : 29-05-2022 : changement de nom dans obs => last_data_date -> last_obs_date, last_data_id -> last_obs_id, et nouveaux champs pour les extremes. Ajout colonne load_json dans postes qui determine si les json doivent etre chargés pour ce poste

3. postes : table des stations

Actuellement seuls les champs suivants sont utilisés par le serveur :

- meteor : code de la station au sein du réseau meteoR.OI
- meteofr : sera utilisé pour les extractions pour Météo France
- last_obs_date : date de l'obs la plus récente. Est maintenue à jour via un trigger sur la table obs.
- last_obs_id : id de l'obs la plus récente
- last_extremes_date : date de l'extrêmes la plus récente. Est maintenue à jour via un trigger sur la table extrêmes.
- last_extremes_id : id de l'extrêmes la plus récente
- load_json : booléen permettant d'activer ou non le chargement des fichiers Json. Dans les deux cas les fichiers Json sont déplacés dans le répertoire archive/{station}...

4. mesures : table des mesures

Cette table sert au paramétrage du serveur, et éventuellement pour récupérer des propriétés des mesures au niveau de l'interface. Voici les champs de la table :

- name : nom de la mesure, seulement pour l'interface utilisateur
- json_imput : nom de la clé utilisé dans le Json de mise à jour venant de la station, qui correspond aussi au nom de la colonne dans la table obs
- json_input_bis : Si le Json ne contient pas la clé json_input, ce synonyme est utilisé (raison de compatibilité principalement)
- archive_col : nom de la colonne dans la table archive de WeeWX (utilisé par le programme de migration seulement)
- archive_table : Forçage du nom de la table pour retrouver les extrêmes dans la base de WeeWX. Par défaut, quand cette colonne est nulle, le nom du champ de WeeWX est concaténé avec « archive_day_ » pour retrouver le nom de la table de WeeWX stockant les valeurs extrêmes journalières. Pour les mesures de type « vent » il faut utiliser une autre valeur indiquée dans ce champ
- val_deca : décalage en heures de la valeur OMM

- max : Calcul de maximum ou pas pour cette mesure
- max_deca : décalage de la valeur max
- min : Calcul de minimum ou pas pour cette mesure
- min_deca : décalage de la valeur min
- is_avg : booléen qui indique si cette mesure doit être agrégée en calculant les moyennes, ou par en faisant la somme des valeurs. Ce champ est utilisé par le programme utilisateur pour utiliser la bonne méthode d'agrégation.
- is_wind : indique s'il faut aller chercher la direction du vent dans la table des extrêmes de WeeWX
- omm_link : indique l'id de la mesure de base associé avec une mesure de type OMM. Null indique que la valeur n'est pas une valeur OMM.
- allow_zero : non utilisé actuellement.
- is_hourly : indique que la mesure est mise à jour qu'une fois par heure, donc qu'une agrégation par minutes ne fait pas de sens

5. exclusion : table des exclusions

non utilisé actuellement, mais le sera bientôt

6. obs : table des observations

Contient toutes les mesures élémentaires remontées par la station, et/ou issu de la base WeeWX de la station. Colonnes de la table :

- time : date et heure en heure local de la station
- id_obs : pour les obs décalées, contient l'id de l'obs principal. Comme cela si l'obs principale est supprimée, les obs liées le sont aussi.
- poste_id : id du poste concerné par la mesure
- duration : durée de la période d'intégration de la station (LOOP). Est en general de 5mn. Une durée égale a zéro est utilisé pour les valeurs d'observation décalée (paramètre val_deca)
- Ensuite une colonne par mesure gérée par le système

```
barometer, barometer_omm, dewpoint, etp, extra_humid1,
extra_humid2, extra_temp1, extra_temp2, extra_temp3, hail,
hail_rate, heatindex, heating_temp, in_humidity, in_temp,
leaf_temp1, leaf_temp2, leaf_wet1, leaf_wet2, out_humidity,
out_humidity_omm, out_temp, out_temp_omm, pressure, radiation,
rain, rain_omm, rain_rate, rx, soil_moist1,
soil_moist2, soil_moist3, soil_moist4, soil_temp1, soil_temp2,
soil_temp3, soil_temp4, uv, voltage, wind,
wind_gust, wind10, wind10_omm, windchill
```

et enfin les colonnes qui seront utilisées quand nous utiliseront des tests de qualité :

- qa_modifications : va stocker le nombre de mise à jour de la mesure
- qa_incidents : va stocker le nombre d'incidents imputés à cette mesure
- qa_check_done : booléen qui indique si le check de qualité a été fait sur cette mesure

7. extremes : table des records

La table des records garde par mesure, les valeurs extremes journalieres relevees. Elles pourront etre agregees a differents niveaux pour indiquer les records sur une periode donnee. Les colonnes sont :

- date : date du jour du record
- poste_id : id du poste concerné par le record
- id_obs : id de l'observation associée au record. Cette zone est simplement utilisée pour effacer les records associés quand une observation est supprimée.
- mesure_id : id de la mesure concernée par le record
- min : valeur minimale du record
- min_time : date et heure du record min
- max : valeur maximale du record
- max_time : date et heure du record max
- max_dir : direction du vent pour les max de type vent

8. incidents : table des incidents

Le programme serveur indique dans cette table tous les incidents qu'il rencontre. Les incidents sont classés par

- dat : date et heure de l'incident
- source : information sur la fonction ayant généré cet incident
- level : exception, error, info
- reason : message de l'incident
- details : Json contenant des informations complémentaires a cet incident. En cas d'exception pendant l'exécution, le stack est stocké dans le json.
- active : booléen pour indiquer si l'incident a été vu et traité (donc active -> false)

9. annotations : table des annotations

Cette table est utilisée par Grafana pour stocker des annotations sur les mesures relevées, comme des événements météorologiques spéciaux

Colonnes :

- time : date et heure du début de l'annotation
- timeend : date et heure de la fin de l'annotation
- text : description de l'annotation
- tags : tag de l'annotation

10. obs_day : vue des observations journalières

Cette vue contient une agrégation journalière des mesures de la table obs. Elle est automatiquement tenue à jour par TimeScaleDB. Elle contient les mêmes colonnes que la table obs.

Les programmes utilisateurs doivent en général parcourir cette table pour retrouver les mesures d'une/plusieurs stations, pour une durée donnée. L'accès à la table obs étant réservé aux demandes d'analyse fine à la suite d'un événement météorologique particulier.