JAMA, Vol. 4, 30-36, 2020

Développement d'un portail web fournissant des produits climatologiques pour le Centre Climatique Régional - Afrique du Nord de l'OMM

Mohamed MEZRED [1]* , Selmane AHMED NOUR EL ISLAM [2]

Abstract
L'Algérie à travers l'Office National de la Météorologie est désignée par l'OMM en 2014 comme un nœud au sein du CCR-Afrique du Nord, chargé de la collecte des données issues des services météorologiques des pays de l'Afrique du Nord pour des fins liées au suivi climatique. Vu l'importance de ces données dans des études des changements climatiques, l'ONM a développé un portail web qui permet de partager les données issues des messages météorologiques SYNOP des pays de l'Afrique du Nord circulant dans le SMT de l'OMM à la disposition de ce Centre qui est basé au Maroc. Cet article décrit le processus mis en œuvre pour élaborer ce portail en allant de la phase décodage du message SYNOP, son contrôle, son archivage dans un système de gestion de la base de données (SGBD) et sa visualisation sur une interface web.
Keywords
SYNOP, CCR, SMT, OMM, SGBD, NOAA, Python, Numpy, Pandas, Ndarray, PsycoPg2, PostgreSQL.

[1] *Office national de la météorologie, Dar El Beida, Alger*
*__Correspondant__:m.mezred@meteo.dz

Contents

Introduction 1
1 Méthodologie 1
1.1 Données utilisées et définition de l'approche pour le décodage du SYNOP 1
1.2 Procédure de contrôle des messages SYNOP . . 1
2 Base de données utilisées 2
3 L'interface de consultation 6
4 Conclusion 6

L'Organisation Météorologique Mondiale (OMM) dispose à travers le monde plusieurs centres climatologiques régionaux. Parmi ces centres : le Centre Climatologique Régional d'Afrique du Nord ou le CCR (en anglais RCC) qui est un centre régional de l'OMM, multifonctionnel et regroupe plusieurs centres ou services climatologiques des six pays membres de la région d'Afrique du Nord : Algérie, Maroc, Tunisie, Lybie, Mauritanie et l'Egypte. Il réalise des activités liées au climat en remplissant collectivement toutes les fonctions requises selon les tâches assignées en commun accord avec l'OMM. Chaque service/centre appartenant au CCR de l'OMM est appelé « nœud ». Un nœud réalisera, pour la région ou la sous-région une ou plusieurs des activités obligatoires que ce soit, les prévisions à long terme (LRF), la surveillance du climat, la collecte des données climatiques (cas de l'Algérie) ou bien la formation. Les CCR de l'OMM sont des centres d'excellence qui établissent également des produits régionaux de prévisions climatiques (notamment saisonnières), et soutiennent les activités climatiques régionales et nationales et renforcent ainsi la capacité des membres de l'OMM dans une région donnée, dans le but d'offrir de meilleurs services climatiques aux utilisateurs nationaux. Le but de ce travail est de faire évoluer le site web RCC (http://rccnara1.marocmeteo.ma/algeriannode.php) par la migration de la base de données de MyQL vers PostgreSQL et de rendre les produits climatiques du RCC plus fiables en appliquant le contrôle qualité selon les recommandations basés par les consignes de l'OMM

1. Méthodologie

Données utilisées et définition de l'approche pour le décodage du SYNOP

Un message SYNOP est un message d'observation élaboré et émis par une station météorologique terrestre et rédigé selon la forme symbolique se référant au Code FM 12 de l'OMM. La diffusion de ce message se fait toutes les trois heures à partir de 00h00 UTC jusqu' à 21H00. Le message SYNOP est composé d'une première ligne qui indique un certain nombre d'informations comme par exemple la date, l'identifiant de la station et puis l'ensemble des paramètres météorologiques observés destinés pour la climatologie et à l'assimilation pour la prévision météorologique. Pour notre cas, nous avons procédé au décodage du SYNOP en provenance des six pays sus-indiqués avec le langage C++ (Stroustrup, 1980) afin d'identifier chaque paramètre météorologique chiffré dans le message SYNOP.

Procédure de contrôle des messages SYNOP

Compte tenu du volume de données accumulées d'une part, et de la complexité du modèle conceptuel d'autre part, il ne peut être envisagé de contrôler les données de façon individuelle, par pointage, de tous les enregistrements dans toutes les tables. Nous avons donc eu recours au contrôle

des données issues du message SYNOP sur la base des consignes élaborées dans le Manuel des codes, Codes Internationaux Volume I.1_Annexe II du Règlement Technique de l'OMM (2018) . La démarche entreprise pour le contrôle consiste à élaborer des programmes en langage Python (Guido Von Rossum, 1990) avec les bibliothèques Pandas et Numpy (Le module Numpy est la boîte à outils indispensable pour faire du calcul scientifique pour modéliser les vecteurs, les matrices, et plus généralement les tableaux à (n) dimensions selon Patrick Fuchs et Pierre Poulain(2019). Le module Numpy fournit le type Ndarray, et Pandas est une librairie python qui permet de manipuler facilement des données à analyser et manipuler des tableaux de données avec des étiquettes de variables (colonnes) et d'individus (lignes). Ces contrôles portent notamment sur les points suivants présentés ici succinctement par ordre croissant de complexité :

- Les recherches des doublons, c'est-à-dire identifiants dupliqués, sont proposées en format standard dans les logiciels modernes de gestion de base de données.
- Contrôles syntaxiques (ex : la température de l'air doit être exprimée par une valeur ayant au plus un chiffre après la virgule) ;
- Contrôle de valeurs à travers la détermination de la fourchette : Pour la détermination de la plage du contrôle de fourchette des trois paramètres (Température maximale, Température minimale, et les Précipitations), nous avons calculé les seuils climatologiques sur une période de 30 années relatifs à ces paramètres au moyen des données d'observation issues de 27 stations des pays membres du CCR complétées par des données issues de la base de données NOAA et ce, pour l'ensemble des stations appartenant à la région. Par exemple les seuils de la température à Dar el Beida-Alger ainsi calculées sont respectivement -3.8°C et +47.9°C, et donc le contrôle de fourchette pour la température pour cette station est basé sur l'intervalle -3.8 et 47.9 pour le paramètre température.
- Calcul du cumul des précipitations annuelles de chaque station.
- Contrôle de l'éventail de climats (ex : est-ce que les données sont cohérentes avec la climatologie).
- Cohérence entre les relevés (ex : la température de l'air ne doit pas être inférieure à celle du point de rosée).
- Cohérence des séries chronologiques (ex : la différence entre deux températures observées successivement sur le même site doit être « plausible »).
- Cohérence spatiale (ex : les limites d'une différence plausible entre les températures d'une station et celle des stations environnantes ne doivent pas être dépassées selon Zahumenský, (2004)).

2. Base de données utilisées

Nous utilisons pour le développement de cette application une base de données PostgreSQL, version_12.0 (Stonebraker, 1985), imposée par la nécessite de pouvoir manipuler des objets géographiques (les sites des stations météorologiques). De plus, la plate-forme utilisant déjà une base de données PostgreSQL avec l'extension PgAdmin4 (version 4.16)(voir fig.3). Cela facilitera l'importation de la base de données ainsi conçue vers cette plateforme. Pour cette phase de modélisation nous avons utilisé l'outil PgAdmin4 qui permet de générer automatiquement le code SQL adapté à une base PostgreSQL. L'outil inclut également une fonction de rétro-conception efficace, ce qui nous a permis, combiné à la génération du code, d'inclure dans le modèle des éléments difficiles à modéliser à partir de l'interface de PgAdmin4 (le code des fonctions servant au contrôle des données). Les fichiers de données seront découpés en ligne par ligne et chaque ligne de données deviendra un tale d'une table. Pour pallier au problème des différents formats de données possibles, nous utiliserons une particularité de PostgreSQL: ce qui nous permet d'utiliser la propriété d'héritage pour faire autant de tables que de formats de données mais en mutualisant certaines informations et contraintes communes à toutes les lignes de données (identifiant, date, relevé associé, identifiant au sein du relevé)(voir fig.2) afin de simplifier la structure de la base et le traitement de ces données. Ces données font donc partie d'un schéma dont l'accès est restreint aux seuls administrateurs de la base afin d'en assurer l'intégrité. La base de données ayant servi de base de test a été hébergée sur un serveur local. Les premières bases de données de test ont été mises en place avant la validation du modèle de base de données afin de réaliser quelques tests avec PostgreSQL. Toujours grâce à des triggers, Nous avons également mis en place un système de contrôle de cohérence des dates (les opérations ne peuvent être effectuées dans le futur, et les lignes de données traitent des données antérieures à leur relevé) et de prévention d'écrasement de données qui vérifie si des données existent déjà pour le couple station/date donnée.

Les scripts d'import ont été réalisés avec python. Nous avons retenu le langage python en raison de sa convivialité et la disponibilité des premières versions des scripts nécessaires pour la réalisation de cette plateforme. Comme nous disposons des concepts avancés sur l'utilisation des bibliothèques Pandas et Numpy permettant ainsi de créer et manipuler aisément différents types de structures de données et une bonne connaissance sur l'usage de PsycoPg2, permettant la communication avec une base de données.

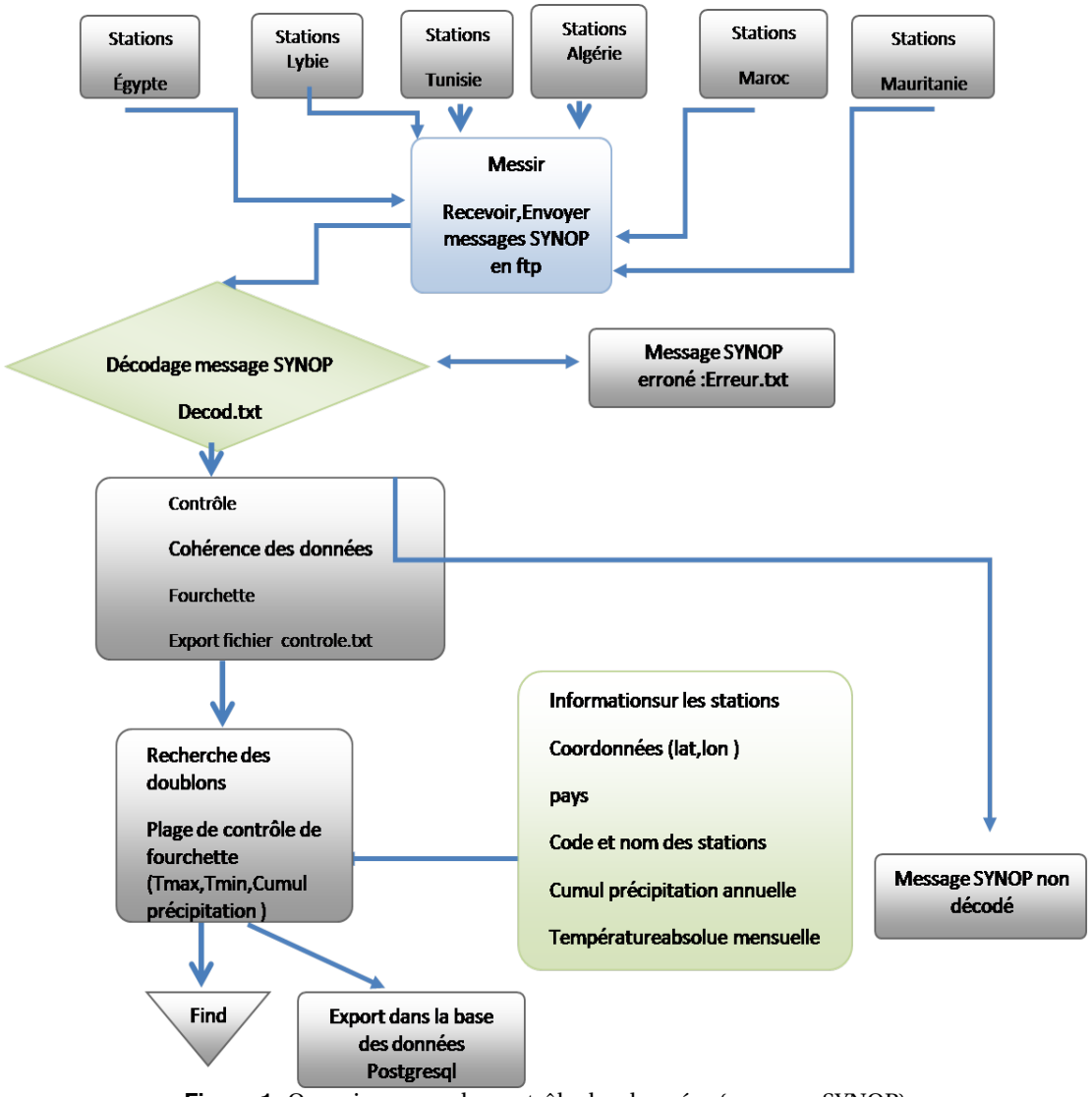

Figure 1. Organigramme du contrôle des données (message SYNOP)

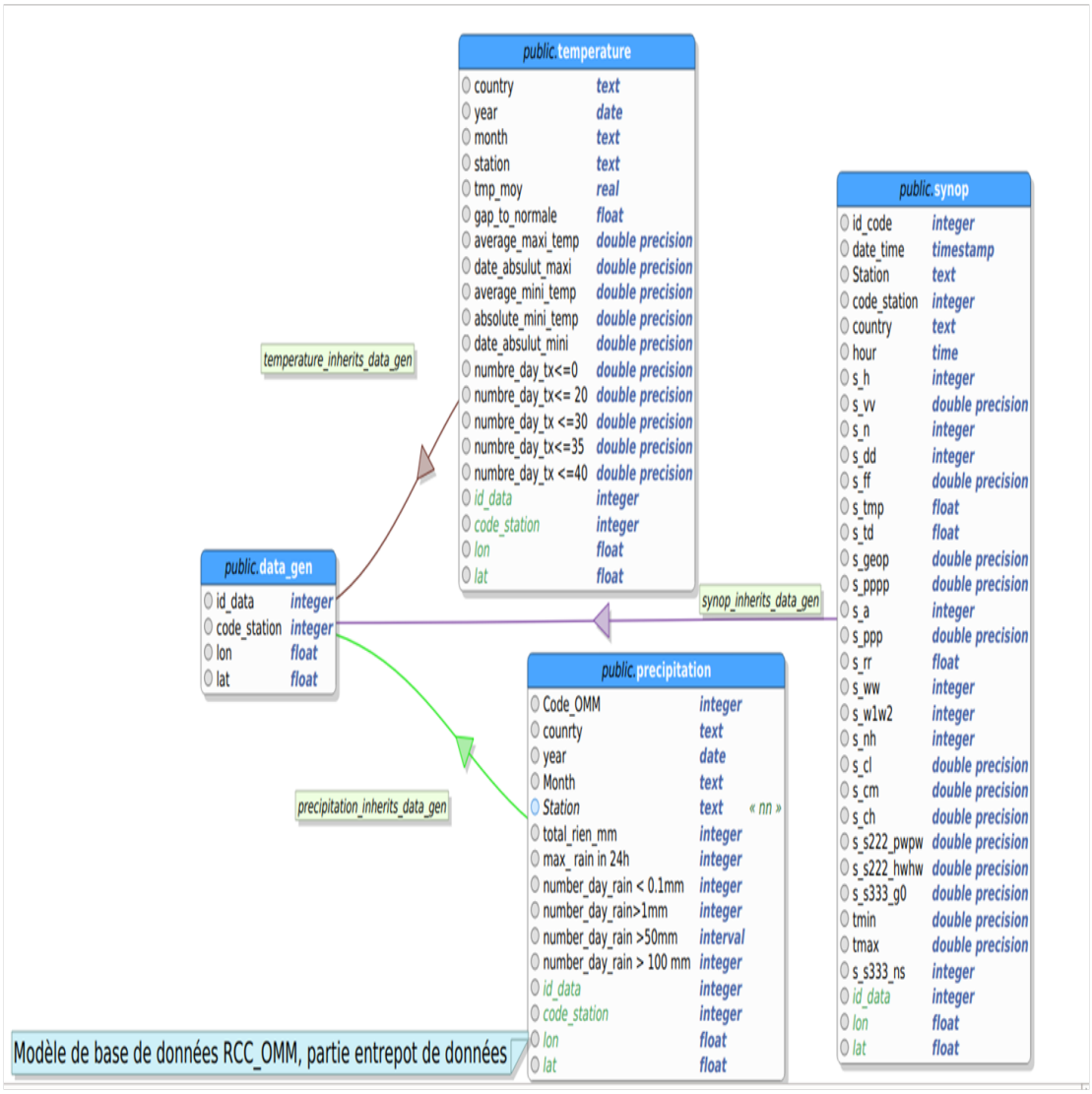

Figure 2. Modèle de base de données

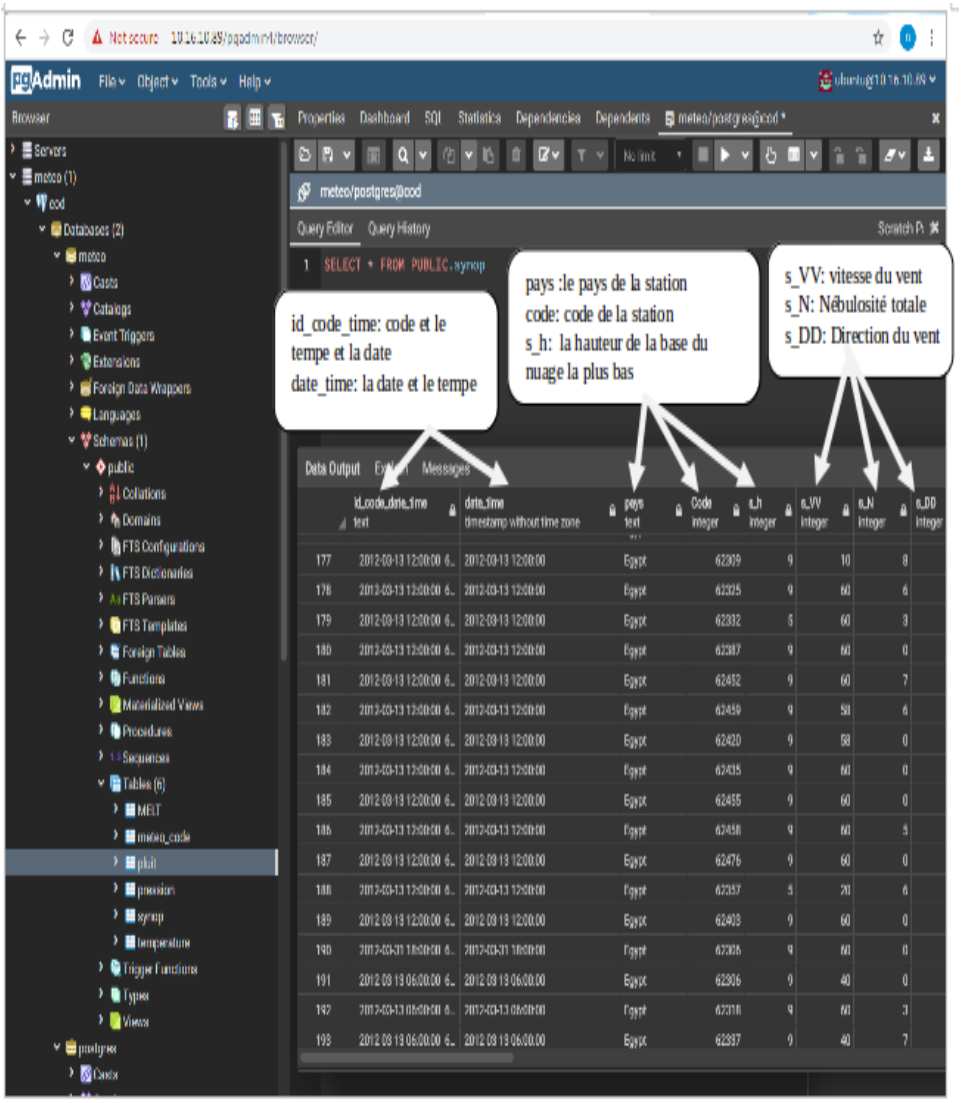

Figure 3. Plateforme PostgreSQL (PgAdmin4)

3. L'interface de consultation

L'aspect graphique de l'interface réalisé en HTML 5 et CSS 3 ainsi que les réactions aux actions de l'utilisateur utilisant JavaScript sont achevées. Nous avons également produit une partie du code dans le PHP (Lerdorf, 1995) relatif à l'interaction avec la base de données. Un ensemble d'informations, un manuel d'administrateur et d'explications sur centre climatologique régional figurent sur la page d'accueil(voir fig.4), et sur le côté gauche de la page il y a aussi une place pour vous de taper le code «Data Access » .

Figure 4. Interface page d'accueil

Après identification de l'invité, la page Data services s'ouvre pour offrir les données de la base , sur la gauche nous trouvons trois choix dans lequel sélectionner le type de données «DATA SERVISES »(voir fig.5) . Après la sélection, d'autres choix sont affichés pour sélectionner le pays «country» sélectionné ainsi que le choix du type de paramètre «parameter», année «years», et mois «month» ;Cliquez sur soumettre «Submit» nous donne un tableau avec les données requises

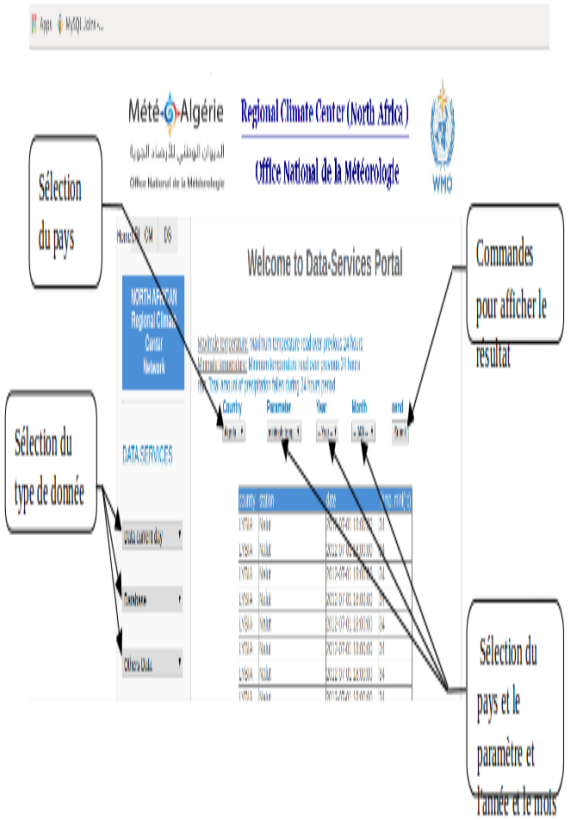

Figure 5. La page de « Data services »

4. Conclusion

Après avoir décodé le message SYNOP nous avons instauré un procédé rigoureux de son contrôle où nous avons eu à appliquer pratiquement l'ensemble des recommandations de contrôle décrites dans le manuel des codes de l'OMM, et ce, afin d'avoir des résultats plus fiables de la donnée issue du message SYNOP à injecter dans la base de données pour un usage partagé entre les services climatologiques des pays membres du CCR - Afrique du nord. Nous avons opté dans la réalisation de cette plateforme et pour la migration des données, pour le choix d'un système moderne de gestion de bases de données climatologiques où l'on a utilisé dans ce système, les programmes et langages de programmation (PostgreSQL, Python, PHP, C++) récents et adaptés à nos exigences en termes d'efficacité et de convivialité. Le processus de contrôle systématique introduit sur les données alimentant la base de données climatologiques du CCR assurera, sans doute, une donnée de qualité fiable pour tout usage scientifique ou de recherche. De par sa mission de nœud chargé de la collecte des données, l'ONM a jugé nécessaire d'améliorer son rôle et son service vis-à-vis du CCR au moyen de cette approche qui constituera indubitablement une contribution importante de l'office à l'égard du CCR et des membres et usagers de la région. Les relevés climatologiques archivées seront dorénavant

beaucoup moins entachés d'erreurs et pourront donc servir une grande communauté tout en ayant une référence qui tire son importance du fait qu'elle soit contrôlée selon les normes et exigences de l'OMM. Elle demeure en plus une ressource pour le présent et pour l'avenir et il est par conséquent plus qu'indispensable d'assurer sa sauvegarde pour les futures générations.

Remerciements

Un remerciement spécial à notre Directeur Mr SAHABI ABED Salah qui nous a aidé à travers ses remarques pertinentes pour l'élaboration de ce travail.

Références

[1] Mehdi Benharrats,Zeste de savoir, La programmation en C++ moderne ,14 mai 2019

[2] Patrick Fuchs et Pierre Poulain , Université Paris Diderot-Paris 7 Paris, France , Cours de Python version du 11 septembre 2019

[3]Sébastien Namèche ,Support de cours PostgreSQL 8 ,installation ,configuration ,exploitation , v_2008-10-02.

[4] Sébastien ROHAUT Cours PHP , 21 juillet 2012.

[5] Organisation météorologique mondial, Manuel des codes, Codes internationaux Volume I.1_Annexe II du Règlement technique de l'OMM Partie A – Codes alphanumériques Mise à jours 2018 .

[6] Igor Zahumenský (Slovakia) Guidelines on Quality Control Procedures for Data from Automatic Weather Stations, 2004

Sites Web:

- http://www.postgresql.org
- https://www.python.org/
- http://pandas.pydata.org/
- https://pypi.python.org/pypi/psycopg2
- http://stackoverflow.com