



# Magest – Importation des mesures

Utilitaire d'extraction des données de sonde dans une base de données

## 1 Présentation

Le logiciel Magest permet d'importer les fichiers générés par les sondes du réseau Magest dans une base de données. Le programme a été écrit pour fonctionner avec une base de données PostgreSQL, mais il devrait pouvoir être facilement adapté à MySQL.

## 2 Installation de PHP

Le logiciel fonctionne en ligne de commande, et utilise PHP comme interpréteur.

### 2.1 Installation Linux

Installez PHP avec la commande :

```
sudo apt install php7.2-cli php7.2-mbstring php7.2-pgsql
```

Vérifiez, dans le fichier de configuration (/etc/php/7.2/cli/php.ini), que le composant *mbstring* est bien activé :

```
extension=mbstring
```

### 2.2 Installation Windows

Téléchargez la version PHP *Non Thread Safe* depuis le site de PHP : <https://windows.php.net/download/>.

## 3 Installation du code

Téléchargez le code depuis Github : <https://github.com/Iristea/magest/archive/master.zip>

Décompressez l'archive dans votre arborescence. Renommez le dossier *magest-master* en *magest*.

Renommez le fichier *param-dist.ini* en *param.ini*.

S'ils n'existent pas, créez les dossiers *magest/import* et *magest/treated*, qui seront utilisés pour déposer les fichiers à traiter ou traités.

### 3.1 Tester l'installation

Linux :

- ouvrez un terminal, et positionnez-vous dans le dossier *magest*
- lancez la commande :

```
php magest.php -h
```

pour afficher le message d'aide.

Windows 10 :

- Ouvrez un terminal *PowerShell* (dans le menu, *Windows PowerShell > Windows PowerShell*). Si *Windows PowerShell* n'est pas installé (ce qui est possible), installez-le depuis le centre de gestion des logiciels de Windows.
- Positionnez vous dans le sous-dossier *magest* :

```
cd .\Documents\magest
```

- lancez la commande :

```
../php/php.exe magest.php -h
```

pour afficher le message d'aide.

## 4 Paramétrage

Les paramètres utilisés pour faire fonctionner le script sont décrits dans le fichier *param.ini*.

Attention : avec une machine Windows, le fichier ne doit être édité qu'avec Notepad++ (<https://notepad-plus-plus.org>), en raison de l'encodage des fins de lignes qui ne sont pas identiques entre Linux et Windows.

Les paramètres sont organisés par section :

### 4.1 general

L'ensemble des paramètres de cette section peuvent être modifiés en ligne de commande.

attribut	Signification
source	dossier contenant les fichiers à traiter
treated	dossier où les dossiers traités sont déplacés en fin de traitement
dsn	paramètres de connexion à la base de données, en suivant les prescriptions de PHP PDO (cf. <a href="https://www.php.net/manual/fr/pdo.construct.php">https://www.php.net/manual/fr/pdo.construct.php</a> )
schema	Nom du schéma contenant la table à alimenter
login	Nom du compte utilisé pour se connecter à la base de données
password	Mot de passe associé au login
separator	Séparateur de champ utilisé dans les données
headerSeparator	Séparateur de champ utilisé dans la partie d'entête des fichiers, pour décrire les colonnes présentes dans le fichier
unitFieldNumber	Dans la partie d'entête décrivant les colonnes, indique le numéro de la colonne où l'unité de mesure est décrite (numérotation à partir de 0)
filetype	extension des fichiers contenant les données
radical	première partie du nom des fichiers à traiter

### 4.2 table

Cette section permet de décrire la table qui recevra les mesures extraites des fichiers.

attribut	Signification
table	Nom de la table qui recevra les données
measure_id	Nom de la colonne qui sert d'identifiant à la table. Elle doit être de type numérique
station	Nom de la colonne qui reçoit le numéro informatique de la station (valeur numérique)
date	Nom de la colonne contenant la date, sous forme <i>timestamp</i>
datefield	Numéro de la colonne, dans les fichiers de sonde, qui correspond à la date. Ce numéro ne doit pas être modifié

### 4.3 stations

Cette section décrit les stations qui sont gérées. Pour chaque station, on retrouve un numéro associé à un libellé, par exemple :

bordeaux=2

Dans ce cas, les mesures concernant la station de Bordeaux seront stockées avec l'identifiant de station n° 2.

### 4.4 fields

Cette section permet de faire la correspondance entre les informations présentes dans les fichiers de sonde et les colonnes de la table. La première partie est construite en extrayant les 4 premiers caractères du champ à traiter, à laquelle est associée l'unité de mesure.

Par exemple, la ligne suivante, dans l'entête d'un fichier :

V:3,Turbidite,NTU

sera référencée en *TurbNTU*. À cette valeur est associée le nom de la colonne utilisée dans la table pour stocker l'information correspondante.

## 5 Exécution

### 5.1 Dépôt des fichiers

Déposez les fichiers à traiter dans le dossier *import*

### 5.2 Lancement du script

Linux :

- ouvrez un terminal, et positionnez-vous dans le dossier *sonde*
- la commande de lancement du script est sous la forme :

```
php magest.php --options
```

Windows :

- Ouvrez un terminal *PowerShell* (dans le menu, *Windows PowerShell > Windows PowerShell*)
- Positionnez vous dans le sous-dossier *magest* :

```
cd .\Documents\magest
```

- la commande de lancement du script est sous la forme (le chemin d'accès à *php.exe* doit être adapté à la situation réelle) :

```
../php/php.exe magest.php --options
```

### 5.3 Options utilisables

Plusieurs options peuvent être ajoutées à la ligne de commande. Elles vont surcharger les paramètres généraux du fichier *param.ini*. Une seule est obligatoire : le nom de la station à traiter.

Paramètre	Fonction
station	nom de la station. Le nom est obligatoire, et doit être référencé dans la section <i>[stations]</i> du fichier de paramètres
dsn	Chaîne de connexion au serveur de base de données
login	Login utilisé
password	Mot de passe associé
schema	Nom du schéma contenant les données dans la base de données
source	Nom du dossier contenant les fichiers à traiter
treated	Nom du dossier où les fichiers seront déplacés après traitement
param	permet de charger un fichier autre que le fichier <i>param.ini</i> pour utiliser des paramètres spécifiques
filetype	extension des fichiers à traiter
radical	racine des fichiers à traiter (début du nom des fichiers)
separator	séparateur utilisé pour séparer les colonnes dans la partie données. Peuvent être utilisés : ; , space, tab
headerSeparator	séparateur utilisé dans l'entête des fichiers pour décrire les colonnes. Les mêmes séparateurs que précédemment peuvent être utilisés
unitFieldNumber	Dans la partie d'entête, indique le numéro du champ contenant l'unité de mesure (numérotation commençant à 0)
noMove	Si vaut 1, les fichiers traités ne seront pas déplacés. À n'utiliser qu'en phase de test ou de mise au point
mode	Si vaut <i>debug</i> , le programme affiche les paramètres lus dans l'entête, ainsi que le tableau d'alimentation de la première ligne de données, puis s'arrête. Cela permet de vérifier que le fichier peut être correctement traité par le programme
help	Le programme affiche un message récapitulant les options, puis s'arrête

### 5.4 Quelques conseils pour l'importation

Le programme traite chaque fichier indépendamment :

- les paramètres des colonnes sont régénérés à chaque fichier ;
- si un problème survient, aucune ligne du fichier n'est importée (fonction *rollback* de la base de données). De plus, le numéro de la ligne qui a déclenché le problème est affiché à l'écran ;
- une fois traité, le fichier est déplacé dans le dossier *treated* pour éviter qu'il ne soit rejoué.

Avant de lancer l'importation d'un nouveau fichier ou d'un nouveau lot, vérifiez systématiquement que la structure est correctement lue, en utilisant l'option `--mode=debug`.

### 5.5 Quelques exemples de lignes de commande

Test du fichier *Pauillac-2017-database.txt* :

```
php magest.php --station=pauillac --separator=tab --radical=Pauillac --  
headerSeparator=tab --unitFieldNumber=2 --noMove=1 --mode=debug
```

Importation du même fichier :

```
php magest.php --station=pauillac --separator=tab --radical=Pauillac --  
headerSeparator=tab --unitFieldNumber=2
```

**Importation de tous les fichiers de Pauillac :**

```
php magest.php --station=pauillac
```

**Importation du fichier *Le-Verdon-2017-database.txt* :**

```
php magest.php --station=leverdon --separator=tab --radical=Le --headerSeparator  
=tab --unitFieldNumber=2
```

**Importation du fichier *Bordeaux-2017-database.txt* :**

```
php magest.php --station=bordeaux --separator=tab --radical=Bordeaux --  
headerSeparator=tab --unitFieldNumber=3
```