Guide détaillé
Traitement des données NINOX

CNES
CENTRE NATIONAL D'ÉTUDES SPATIALES

Prérequis :

- Avoir effectué au moins une nuit de mesure avec le NINOX ;
- Fiche ou guide pratique « Comment récupérer les mesures NINOX ? ».

Matériel nécessaire :

- Ordinateur ;
- Tableur ou environnement de développement Python

Après avoir laissé le système NINOX relever des mesures de luminosité de ciel nocturne durant au moins une nuit (plus grand est le nombre de nuit de mesure, meilleure sera l'analyse) et les avoir téléchargées depuis la mémoire interne du système (cf. fiche pratique « Récupération des données NINOX »), la prochaine étape consiste à traiter et analyser les mesures afin de caractériser la pollution lumineuse autour du lieu d'installation du NINOX.

Pour cela, il existe de nombreuses solutions. Dans ce projet, deux d'entre elles ont été retenues : l'utilisation d'un tableur (Excel, OpenOffice ou LibreOffice) ou bien passer par la programmation en langage Python.

Si vous choisissez une de ces solutions, vous trouverez sur la page web associée à ce projet, des ressources (fiches pratiques, guides détaillés, mesures d'exemple) explicatives sur chacune des méthodes présentées.

Ce guide contient tout d'abord une description des fichiers contenus dans l'archive contenant toutes les mesures NINOX.

Ce guide détaillé a surtout pour objectif de présenter chacune des solutions retenues pour observer les différentes manipulations et résultats qui peuvent être obtenus. En conclusion de ce document, un comparatif des deux types de solutions sera donné afin de vous permettre de choisir la solution la plus adaptée à votre situation.

1. Présentation de l'archive contenant les mesures NINOX

A la fin de la fiche pratique « Comment récupérer les mesures NINOX ? », une archive contenant les mesures acquises par le système NINOX a été téléchargée. Cette archive contient plusieurs fichiers dont le nom commence toujours par le nome du système NINOX utilisé (ninox<numéro de série>) puis la date de téléchargement et fini par le nom spécifique de chaque fichier.

Tous ces fichiers sont au format CSV (pour *Comma-Separated Values)* qui peuvent être ouverts avec un éditeur de texte comme le bloc-notes.

Voici une description de chaque fichier :

Nom du fichier	Description
location.csv	Contient les coordonnées de chaque lieu d'observation
measure_full.csv	Contient toutes les mesures et leurs informations complémentaires (cf. tableau suivant)
ninox.csv	Contient les détails techniques du NINOX utilisé
nss.csv	Contient les différentes valeurs de NSS (*Night Sky Stability*) calculées durant les nuits de mesure. Pour plus d'informations, se référer au manuel NINOX.
sqm.csv	Contient les détails techniques du capteur SQM (*Sky Quality Meter*) utilisé dans le NINOX

En pratique et dans la suite de ce projet, seul le fichier *measure_full.csv* contient des informations intéressantes pour l'étude de la pollution lumineuse. En outre, ce fichier centralise toutes les informations contenues dans les autres fichiers de l'archive.

Explicitions donc le contenu de ce fichier.

Le fichier *measure_full.csv* contient autant de lignes que le système NINOX a effectué de mesures. Il contient également 22 colonnes, chacune contenant une information particulière :

Nom de la colonne	Description
measure_id	Identifiant de la mesure de NSB
ninox_id	Identifiant du système NINOX (contenu dans ninox.csv)
sqm_id	Identifiant du capteur SQM (contenu dans sqm.csv)
loc_id	Identifiant du site d'observation (contenu dans location.csv)
jd_utc	Date et heure de la mesure en Jours Julien UTC
az	Azimut de visée (en degré) du SQM, **par défaut 0**
alt	Altitude de visée (en degré) du SQM, **par défaut 90**
temp_sensor	Température du SQM (x100, en °C)
temp_ambiant	Température ambiant (x100, en °C), **capteur absent, -10000 par défaut**
humidity	Humidité relative (en %), **capteur absent, -100 par défaut**
pressure	Pression (en hPa), **capteur absent, -100.0 par défaut**
cloud_cover	Couverture nuageuse (en %), **capteur absent, -100 par défaut**
wind_speed	Vitesse du vent (x10, en km/h), **capteur absent, -100 par défaut**
counts	Paramètre « count » du SQM
frequency	Paramètre « frequency » du SQM
sqm_mag	Luminosité du ciel (NSB) mesuré, en mag/arcsec²

sun_alt	Altitude du Soleil (x10, en degré)
moon_alt	Altitude de la Lune (x10, en degré)
moon_phase	Phase de la Lune (x10, en degré)
gal_lon	Longitude galactique du point de visée du SQM (x10, en degré)
gal_lat	Latitude galactique du point de visée du SQM (x10, en degré)
flag_sent	Indique si la mesure à déjà été exportée (1 si oui, 0 sinon)

Ce fichier contient donc de nombreuses informations qui complètent la valeur de NSB mesurée toute les minutes.

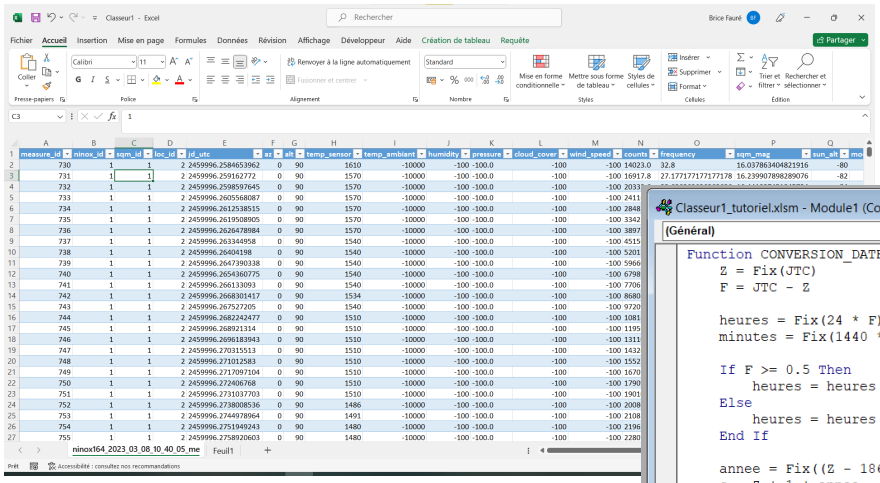
Cependant, en réalité, certaines colonnes ne contiennent pas d'informations exploitables et le début du traitement des données NINOX passera par la simplification des mesures téléchargées en supprimant certaines colonnes inutiles dans notre cas.

Ainsi, une fois ces données récupérées, l'utilisateur peut maintenant passer au traitement de ces données pour les analyser par la suite.

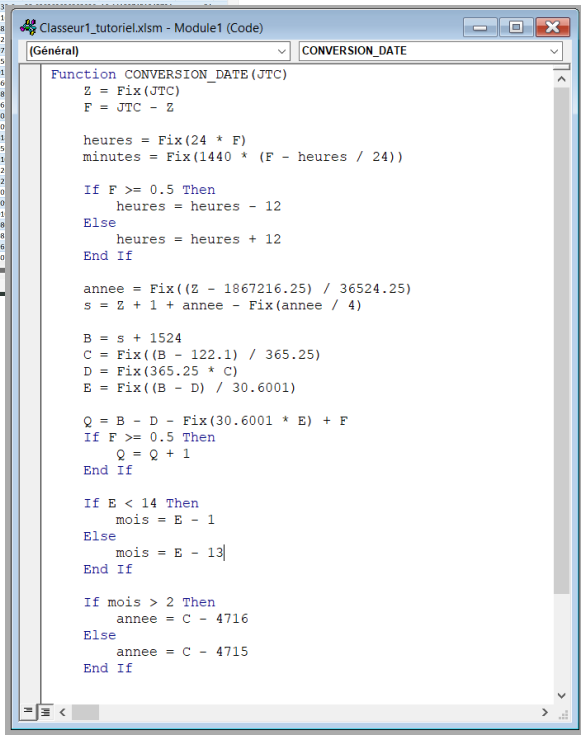
Pour cela, et comme indiquer dans le paragraphe introductif de ce guide, deux solutions ont été retenues : l'utilisation d'un tableur ou la programmation en langage Python.

Dans la prochaine section de ce document, le traitement des données avec un tableur (Excel, OpenOffice ou LibreOffice), sera présenté.

2. Traitement des données avec un tableur

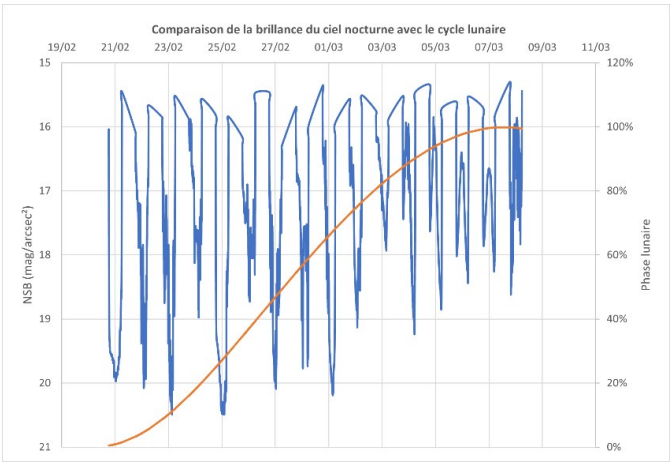


Cela commence par l'importation des données dans une feuille de calcul. Cette étape est facilitée par le format des fichiers, le csv.



La deuxième étape consiste à manipuler la feuille de calcul afin de simplifier les données. L'étape importante est de convertir les dates et les heures représentées en jour julien en dates et heures au format du calendrier grégorien. Cela passe par la création d'une nouvelle fonction dans le tableur afin de réaliser cette conversion.

L'étape suivante est de visualiser les mesures en traçant des courbes comme la suivante :



Si cette solution est retenue, un guide détaillé et des mesures d'exemple sont disponibles en téléchargement sur la page web associée à ce projet.

3. Traitement des données en programmation Python

Une autre solution proposée dans ce projet est l'utilisation de la programmation Python pour effectuer des manipulations semblables à celles des tableurs mais adaptables à de nombreux autres fichiers.C'est aussi l'occasion de découvrir ce langage de programmation. L'utilisateur sera amené à utiliser une bibliothèque Python dédiée à la manipulation d'un grand ensemble de données : la bibliothèque *Pandas*, très utilisée en data science, et d'autres bibliothèques toutes aussi répandues comme *Numpy* ou encore *Matplotlib*.

Pour cela, la première étape consiste à importer les données dans une nouvelle structure de données, les *dataframes* (illustrés ci-dessous), propre à *Pandas* et semblable aux feuilles de calcul d'un tableur est utilisée.

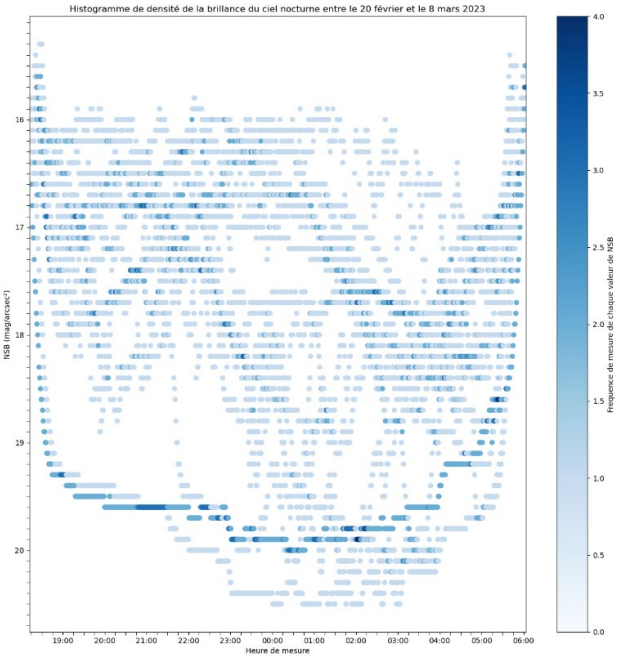
```
Entrée [4]: dataframe_complet # affichage du dataframe
```

Out[4]:

	measure_id	ninox_id	sqm_id	loc_id	jd_utc	az	alt	temp_sensor	temp_ambiant	humidity	...	wind_speed	counts	frequency	sqm_mag	sun
0	730	1	1	2	2.459996e+06	0	90	1610	-10000	-100	...	-100	14023.0	32.800000	16.037863	
1	731	1	1	2	2.459996e+06	0	90	1570	-10000	-100	...	-100	16917.8	27.177177	16.239908	
2	732	1	1	2	2.459996e+06	0	90	1570	-10000	-100	...	-100	20333.6	22.626263	16.441937	
3	733	1	1	2	2.459996e+06	0	90	1570	-10000	-100	...	-100	24116.8	19.161046	16.633952	
4	734	1	1	2	2.459996e+06	0	90	1570	-10000	-100	...	-100	28482.4	16.130711	16.813897	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
11018	11748	1	1	2	2.460012e+06	0	90	930	-10000	-100	...	-100	14674.4	31.000000	16.089945	
11019	11749	1	1	2	2.460012e+06	0	90	930	-10000	-100	...	-100	12923.6	35.400000	15.947974	
11020	11750	1	1	2	2.460012e+06	0	90	930	-10000	-100	...	-100	10968.6	42.000000	15.767974	
11021	11751	1	1	2	2.460012e+06	0	90	930	-10000	-100	...	-100	8982.4	51.000000	15.550000	
11022	11752	1	1	2	2.460012e+06	0	90	930	-10000	-100	...	-100	8096.2	57.000000	15.440000	

11023 rows × 22 columns

La deuxième étape est de manipuler cette structure de données pour afficher des dates et heures compréhensibles ou bien de modifier la forme du dataframe pour ensuite créer des graphiques intéressants. Durant cette étape, l'utilisateur sera amené à coder de nombreuses fonctions. Cette étape est détaillée dans un guide disponible sur la page web de ce projet.



Enfin la dernière étape consiste à créer des graphiques pour visualiser les mesures effectuées. Plusieurs types de courbes sont tracés : des courbes contenant toutes les données mesurées, des courbes unitaires mais aussi un nouveau type de graphique, un histogramme de densité. Cet histogramme permet de visualiser les valeurs mesurées les plus fréquentes à chaque moment de la nuit et cela sur un seul graphique (cf. ci-contre).

Si cette solution est retenue, un guide détaillé et des mesures d'exemple sont disponibles en téléchargement sur la page web associée à ce projet.

4. Comparatif des solutions

Afin de choisir la meilleure solution entre les deux présentées précédemment, voici un tableau comparatif :

Tableur (Excel, OpenOffice ou LibreOffice)	Programmation Python
Peu de connaissances préalables requises	Des connaissances en Python simplifient la découverte
Rapidité d'obtention des premiers résultats	Plusieurs manipulations avant d'arriver aux premiers résultats
Difficulté à manipuler beaucoup de donnée	Manipulations indépendantes de la taille des données
Instable avec beaucoup de données	Stable avec beaucoup de données
Répétition des manipulations pour chaque fichier	Réutilisation des scripts pour d'autres fichiers de mesure

Au final, qu'importe la solution choisie, les connaissances et compétences développées peuvent être réutilisées dans d'autres situations, notamment dans le monde professionnel.

5. Et maintenant ?

Une fois la solution choisie, vous pouvez vous rendre sur la page web associée à ce projet afin de trouver des guides tutoriels pour mettre en place cette solution et l'appliquer soit sur des mesures d'exemple ou bien directement sur vos propres mesures NINOX.