

Sujets de mini-projets 2020-21

I Analyse de données : recherche de tendance

On dispose d'une série temporelle longue (presque 20 ans) des concentrations d'ozone sur un site peu pollué (l'Observatoire de Haute Provence, près de Manosque). De telles mesures renseignent sur les niveaux de concentration de fond de l'ozone troposphérique. Le propos est de détecter une éventuelle tendance de ces concentrations sur l'échelle de temps considérée (deux décennies) à partir de ce jeu de données.

La variabilité de ce signal a de multiples origines, et se produit sur des échelles de temps très diverses : diurnes, météorologiques (quelques jours), saisonnières. Il s'agira de mettre en œuvre une ou des méthodes permettant de définir les anomalies relativement à une tendance de long terme.

1. Comment «lisser» les données ? Quelles est la statistique à considérer (médiane, moyenne) ?
2. Quelles sont les distributions des anomalies suivant la statistique choisie ? Ces distributions dépendent des échelles de temps considérées ?
3. Rechercher une tendance linéaire sur la période considérée. Quelle est la distribution des résidus (des écarts au modèle linéaire) ?
4. Comment quantifier un intervalle de confiance sur la tendance ?

II Analyse de données : statistiques météorologiques à Jussieu

La station météorologique Davis installée au sommet de la tour Zamansky du site de Jussieu depuis septembre 2015 mesure les paramètres météorologiques avec un pas de temps de 1 min. La variabilité temporelle des observations (pression, température, humidité, point de rosée, éclaircissement solaire, vent moyen, rafales, ...) peut être imputée à de multiples causes et se produit sur des échelles de temps diverses (horaire, diurne, météorologique, saisonnière).

1. Dans ce projet, on cherchera d'abord à décrire la statistique de la vitesse du vent mesurée à la station Davis. On pourra aussi s'intéresser à décrire sa variabilité mensuelle sur une année.
2. On cherchera ensuite à modéliser la distribution des vitesses de vent obtenues à l'aide de la distribution de Weibull.

La distribution de Weibull est utilisée pour estimer le potentiel éolien d'un site. Elle modélise la probabilité qu'un vent souffle à telle vitesse sur ce site. La densité de probabilité de Weibull d'une variable continue x , représentant ici les données de vitesse de vent, est donnée par

$$p(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, \text{ où } k \text{ et } \lambda \text{ sont les paramètres de forme et d'échelle de la distribution, respectivement.}$$

La fonction de répartition empirique s'écrit de manière très simple sous la forme $F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$.

Lorsqu'on souhaite comparer la prédictibilité des données de vitesses de vent obtenues sur différents sites, le paramètre clé est l'écart relatif Ψ , c'est-à-dire le rapport de l'écart-type à la moyenne de la distribution. De manière remarquable, le paramètre d'échelle λ s'annule dans l'expression de Ψ , qui ne dépend donc que du paramètre de forme k : $\Psi = \frac{\sigma_x}{\bar{x}} = \frac{1}{\Gamma\left(1 + \frac{1}{k}\right)} \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma\left(1 + \frac{1}{k}\right)^2 \right]^{1/2}$.

On cherchera à estimer les paramètres de la loi de Weibull théorique et on comparera les distributions expérimentales et théoriques (histogrammes, qqplot, tests d'hypothèses...).

3. On pourra enfin chercher à déterminer la puissance potentiellement récupérable sur une éolienne placée sur les toits de Jussieu en connaissant son facteur de capacité (puissance fonction de la vitesse du vent).

III Cycle diurne des températures de surface et de l'atmosphère

On cherche à représenter de manière simplifiée l'évolution diurne de la température de surface $T_s(t)$ et de la température de l'atmosphère $T_a(t)$. Dans un modèle purement radiatif, l'évolution diurne $T_s(t)$ est forcée par celle du rayonnement solaire absorbé $S(t)$. On considère que l'atmosphère n'absorbe pas le rayonnement solaire et qu'elle se comporte comme un corps noir dans le domaine de l'infrarouge tellurique. Compte-tenu du rayonnement émis vers l'espace et les échanges radiatifs entre atmosphère et surface, le bilan radiatif local peut être décrit par un couplage de deux équations différentielles ordinaires du premier ordre couplées :

$$\begin{aligned}\frac{dT_s}{dt} &= \frac{1}{C_s} (S(t) - \sigma T_s^4 + \sigma T_a^4) \\ \frac{dT_a}{dt} &= \frac{1}{C_a} (\sigma T_s^4 - 2\sigma T_a^4)\end{aligned}$$

où $C_s = 200 \text{ W h m}^{-2} \text{ K}^{-1}$ et $C_a = 50 \text{ W h m}^{-2} \text{ K}^{-1}$ sont les capacités calorifiques surfaciques de la surface et de l'atmosphère et $\sigma = 5.67 \times 10^{-8} \text{ W K}^{-4} \text{ m}^{-2}$ la constante de Stefan-Boltzmann. $S \geq 0$ est le rayonnement solaire absorbé par la Terre par unité de surface, dont l'évolution diurne peut être modélisée par une fonction sinusoïdale de l'angle horaire $H = 2\pi \frac{t}{T_0}$ ($T_0 = 24 \text{ h}$) d'amplitude maximale $S_m = 600 \text{ W m}^{-2}$.

Le travail consistera d'abord à résoudre numériquement sur une durée d'une vingtaine de jours ce système d'équations différentielles à l'aide de différentes méthodes (Euler, RK4, point milieu), avec différentes températures initiales autour des valeurs asymptotiques attendues. On s'intéressa ensuite plus précisément aux transitions jour/nuit une fois le régime permanent atteint. On cherchera notamment à expliquer l'évolution temporelle des températures en fonction des trois termes de forçages : forçage solaire et émittances de la surface et de l'atmosphère. On pourra ainsi comparer les durées de réchauffement/refroidissement aux durées du jour et de la nuit. Enfin, on ciblera l'évolution de la température de surface pendant un refroidissement nocturne. On cherchera alors à identifier les raisons de l'évolution de l'erreur accumulée au cours d'un cycle diurne dans les différentes méthodes de résolution.

IV Simulation numérique du transport vertical dans l'atmosphère

L'objectif est de modéliser le transport vertical de particules sous l'action du vent vertical $w(z, t)$. Le vent vertical se décompose en la somme d'un profil moyen $\bar{w}(z)$ déterministe et d'une contribution $w'(z, t)$ de moyenne temporelle nulle et fluctuant de façon aléatoire qui peut représenter la turbulence. Le mouvement vertical $z_i(t)$ d'une particule i est déterminé par l'équation différentielle dite stochastique (EDS) du premier ordre :

$$\frac{dz_i}{dt} = w(z_i, t) = \bar{w}(z_i) + w'(z_i, t) \quad (1)$$

que l'on intégrera numériquement pour un ensemble de particules initialement réparties sur l'axe vertical. En construisant les trajectoires de ces particules, on peut modéliser le transport vertical de ce nuage et caractériser statistiquement son évolution.

IV.1 Transport vertical sans turbulence

Dans un premier temps, on étudie le transport sans turbulence, en commençant avec un profil vertical $\bar{w}(z)$ simple, pour lequel des expressions analytiques des trajectoires sont facilement calculables : par exemple vitesse constante positive dans la couche limite et nulle au dessus.

IV.2 Transport vertical avec turbulence

La résolution numérique de l'EDS s'effectue de façon itérative comme pour une EDO classique, mais avec un terme supplémentaire aléatoire dû à la turbulence. On modélise tout d'abord le déplacement vertical élémentaire $w' dt$ par une variable aléatoire gaussienne centrée. On construit donc un générateur aléatoire G de distribution gaussienne centrée réduite et on le caractérise : moments centrés, tests sur la distribution. L'efficacité de la turbulence est représentée par un coefficient de diffusion déterministe $D(z)$ (en $\text{m}^2 \text{s}^{-1}$) qui peut dépendre de l'altitude.

Avec la méthode élémentaire d'Euler, on peut représenter ce déplacement pendant un pas de temps Δt par une variable aléatoire gaussienne centrée de variance $2D(z_i)\Delta t$. Le schéma d'Euler stochastique prend la forme :

$$z_i(t + \Delta t) = z_i(t) + \overline{w}(z_i(t), t)\Delta t + \sqrt{2D(z_i)}G\sqrt{\Delta t} \quad (2)$$

Quand le pas tend vers 0, le terme stochastique décroît en $\sqrt{\Delta t}$ plus lentement que le terme déterministe. On parle de méthode d'ordre 1/2 au sens fort.

On envisagera ensuite le schéma d'ordre 1 au sens fort dit de Milstein :

$$z_i(t + \Delta t) = z_i(t) + \overline{w}(z_i(t), t)\Delta t + \sqrt{2D(z_i)}G\sqrt{\Delta t} + \frac{1}{2} \frac{dD}{dz}(z_i) (G^2 - 1) \Delta t \quad (3)$$

Quelle est l'espérance du terme supplémentaire apporté par Milstein ? On le testera évidemment dans le cas où le coefficient de diffusion dépend de l'altitude.

En résolvant cette EDS, on peut alors étudier l'évolution du nuage de particules : tracé de trajectoires, trajectoire et vitesse du centre de gravité, dispersion, ...