

# TRAVAUX D'INITIATIVE PERSONNELLE ENCADRÉS

# INFORMATIQUE ET PRÉVISIONS

# MÉTÉOROLOGIQUES

CONCOURS CPGE 2021

HÉLÈNE SIBOULET MP  
JUIN 2021

<b>1</b>	<b>Présentation</b>	<b>1</b>
<b>2</b>	<b>État de l'art</b>	<b>2</b>
<b>3</b>	<b>Recherche de données</b>	<b>3</b>
<b>4</b>	<b>Modèles de prévision sur une station : analyse de la température et de l'humidité</b>	<b>3</b>
4.1	Premier modèle : Moyenne de la température et de l'humidité par jour de l'année et par heure . . . . .	4
4.2	Deuxième modèle : Modèle de la persistance journalière . . . . .	4
4.3	Troisième modèle : Approximation de la courbe des températures par une somme de fonctions sinusoïdales	4
<b>5</b>	<b>Exemples Latex</b>	<b>5</b>
5.1	subsection . . . . .	5
5.1.1	q;,c . . . . .	5
5.1.2	Donner des figures . . . . .	5
5.1.3	Citer des figures . . . . .	5
5.2	Placer un tableau . . . . .	5
5.3	Citer un tableau . . . . .	6
<b>6</b>	<b>Références</b>	<b>6</b>

## 1 Présentation

La prévision consiste à prévoir, c'est-à-dire à « voir avant ». Il s'agit donc, à partir de ce que l'on sait à un moment donné, de savoir ce qui se passera à un moment ultérieur, avec la meilleure probabilité. En effet, à part les phénomènes déterministes, comme par exemple le mouvement des solides à court terme, on ne peut avoir que des indications probabilistes. C'est un problème universel, qui s'applique à l'économie, la politique, la vie privée, et aussi bien sûr aux prévisions météorologiques.

Définir l'intelligence en général est très compliqué. Voici une définition simple pour ce rapport : observer-se souvenir-faire un modèle de prévision. Il existe un dicton météorologique : "Quand la neige ne fond pas, c'est quelle en attend d'autre." Ce dicton est confirmé par les observations. Comme les masses d'air en France circulent d'ouest en est, la neige tombe d'abord sur un front d'air chaud -air froid, puis reste sous l'air froid, et tombe encore sur le front air froid - air chaud. On peut dire que ce dicton est une forme d'intelligence. Ce travail de TIPE est semblable, puisqu'il existe des paramètres (température et humidité), ils sont sur des bases, et ce travail développe une utilisation de ces données. On peut parler d'intelligence artificielle.

La prévision repose sur deux éléments : des données disponibles à un moment, et des théories ou des modèles qui permettent d'utiliser ces données.

Les processus météorologiques sont tout à fait chaotiques. Cela signifie que des incertitudes minimales à un instant donné peuvent avoir des conséquences très fortes sur le déroulement. En fait, ces incertitudes n'empêchent pas de prévoir à court terme, par exemple pour le lendemain, mais elles rendent difficiles les prévisions au delà de deux semaines.

Si les processus météorologiques sont chaotiques à moyen terme, ils sont au contraire extrêmement stables dès

lors que on les analyse en moyenne. Par exemple, Milankovitch a prévu que les variations de trois paramètres - excentricité, obliquité et précession - auraient des effets périodiques sur les glaciations. Milankovitch a fait apparaître des périodes de 20 000, 40 000 et 100 000 ans, ce qui a été confirmé expérimentalement par les mesures de  $\delta O_{18}$  dans les carottes glaciaires de Vostock, en 1976. Ces résultats montrent que, en faisant des moyennes sur une centaine d'année, on a des tendances très claires d'évolution des températures, alors qu'il est difficile de prévoir la température dans deux semaines.

Dans ce documents, nous allons faire un bref état de l'art des méthodes de prévision météorologiques utilisées, et proposer une méthode basée sur l'intelligence artificielle. Nous verrons que ces deux approches sont tout à fait différentes. Nous collectons une base de données, nous élaborons une méthode avec de nombreuses variantes. Chaque variante est évaluée numériquement, c'est à dire par comparaison des prévisions avec un ensemble de données qui ont été placées en dehors des données d'entrées.

Les prévisions météorologiques ont beaucoup d'enjeux, elles permettent notamment :

1. aux pilotes d'avion et aux marins de savoir si leur trajet est réalisable, de connaître la quantité de carburant à prendre et de choisir l'itinéraire le plus adapté
2. aux agriculteurs de savoir quand arroser, quelle quantité d'eau est nécessaire, quand labourer, planter ou récolter
3. d'organisation d'activités ou d'évènements en plein air
4. de gérer des risques liés au climat par exemple lorsqu'il y a du verglas, de la neige, une canicule ou une tempête

## 2 État de l'art

Il existe plusieurs sortes de prévisions météorologiques : la prévision du temps qui fait des prédictions sur des périodes de quelques jours, la prévision saisonnière qui fait des prédictions sur l'échelle d'un mois ou de quelques mois et la prévision climatiques prédit l'évolution du climat sur plusieurs années.

Selon [http ://meteocentre.com/intermet/prevision/prevision\\_emp.htm](http://meteocentre.com/intermet/prevision/prevision_emp.htm) les principales méthodes de prévision météorologiques sont :

1. La méthode de la persistance qui prédit que demain il fera le même temps qu'aujourd'hui. Elle fonctionne bien dans certaines régions où les conditions météorologiques varient lentement comme l'Égypte ou les déserts des États-Unis. Elle fait de bonnes prévisions sur des périodes de un ou deux jours ainsi et est efficace pour prédire des tendances sur des durées d'un mois (prévisions saisonnières).
2. La méthode de la tendance qui étudie les déplacements de systèmes météorologiques tels que les dépressions, les anticyclones, les fronts et les zones de précipitation et suppose leurs emplacements futurs.
3. La méthode de l'analogie qui consiste à chercher des cas dans le passé qui ressemblent aux conditions météorologiques actuelles et à supposer que ces conditions évolueront de la même manière.
4. La méthode numérique est la plus utilisée actuellement. On mesure les différents paramètres ce qui nous donne l'état initial. Les lois physiques prédisent ensuite l'évolution du système. Il s'agit d'une simulation physique de l'atmosphère. L'article de Pailleux propose notamment une simulation basée sur les équations hydrodynamiques suivantes :
  - (a) Principe fondamental de la dynamique
  - (b) Équation de continuité
  - (c) Premier principe de la thermodynamique
  - (d) Loi des gaz parfaits
  - (e) Équation bilan des constituants

Cette méthode requiert :

- (a) La découpe de l'atmosphère dans une grille en trois dimensions. Plus la grille est large plus les calculs sont rapides et moins ils sont précis.
- (b) L'intégration des équations précédentes. Le plus souvent en approximant la dérivée des différents paramètres comme constant sur un petit intervalle de temps  $dt$ . Pour que le modèle fonctionne il faut que  $dt$  soit suffisamment petit pour que les paramètres changent peu sur  $dt$  et en particulier que les systèmes météorologiques se déplacent de moins d'une case dans la grille.

- (c) Une puissance numérique très élevée. Par exemple, le calculateur de Météo-France ATOS Sequana XH2000 possède une puissance de 100 pétaFLOPS.

On trouve sur internet d'autres types d'outils de prévision tels que des réseaux de neurones, des algorithmes génétiques, de la logique floue et des réseaux bayésiens (voir Das2017).

Toutes ses méthodes peuvent être exploitées différemment : pour faire des prévisions déterministes ou probabilistes. Étant donné la précision des méthodes actuelles, la méthode déterministe peut être utilisée pour des échéances de trois à quatre jours. Faire des prévisions probabilistes, aussi appelées prévisions d'ensemble, consiste à donner plusieurs prévisions à partir d'états initiaux proches. Ces états initiaux sont choisis de sorte à refléter l'incertitude des prévisions. Il faut ensuite comparer les différentes prévisions pour déterminer quels sont les scénarios les plus probables et en cas de risques liés au climat quels sont les pires scénarios (voir MétéoFrance).

Nous avons choisi pour ce TIPE de privilégier différentes approches de modélisation mathématique.

### 3 Recherche de données

Nous avons trouvé une base de données sur :

[https://donneespubliques.meteofrance.fr/?fond=produit&id\\_produit=90&id\\_rubrique=32](https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32).

Elle contient les relevés de 62 stations en France Métropolitaine et en France d'Outre mer de 1996 à 2020 (avec certaines mesures manquantes : pas plus de 80 en une année soit l'équivalent de dix jours sauf pour 2020) avec un relevé toutes les trois heures soit huit relevés par jour. Les données sont : la température, l'humidité, la direction et la force du vent, la pression atmosphérique, la hauteur de précipitations, le temps sensible, la description des nuages, la visibilité. Chaque fichier couvre un mois et toutes les stations.

Nous avons collecté sur le site avec Webbot.

Nous avons décompressé les fichiers vers le format csv décompressé.

Nous avons trié les données par station.

### 4 Modèles de prévision sur une station : analyse de la température et de l'humidité

Extraction des données de température et d'humidité de Clermont-Ferrand au format json

Les années 2008, 2019 et 2020 sont retirées de la base d'entraînement et serviront de base de test

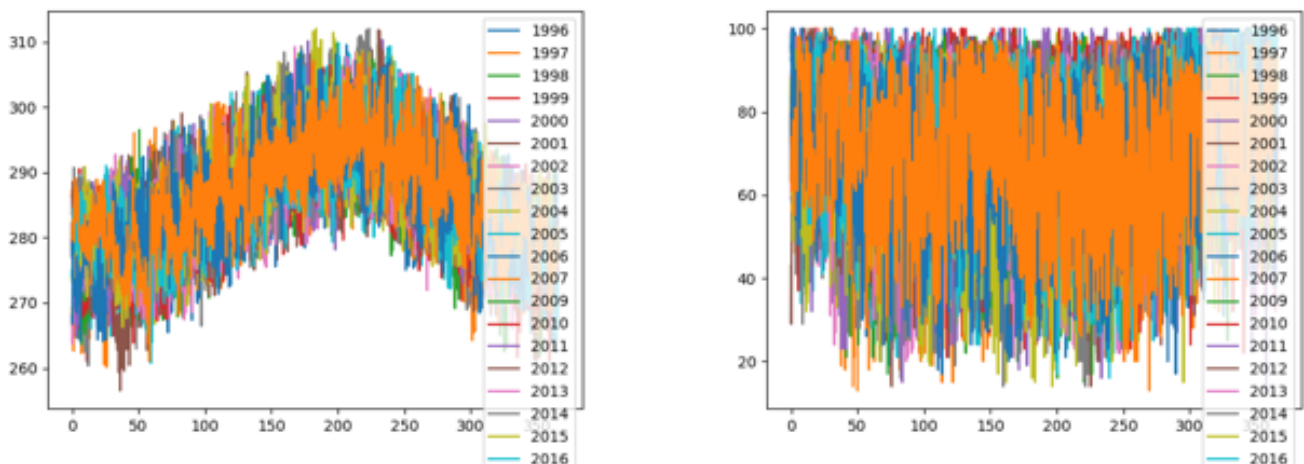


FIGURE 1 – Courbes de température et d'humidité de la base d'apprentissage : Clermont-Ferrand

Comment estimer la valeur d'un ensemble de prévisions ? On calcule l'écart type entre nos prévisions et les températures mesurées. Cela permet d'estimer la valeur des prévisions, mais à condition d'avoir une valeur de comparaison. Nous utilisons pour cette valeur l'écart type des températures.

Calcul de valeurs utiles de la base de donnée : quartiles, medianne, moyenne, écart type.

Température	$Q1 = 6.39^{\circ}C$	$me = 11.9^{\circ}C$	$Q3 = 17.4^{\circ}C$	$m = 12.0$	$\sigma = 7.91^{\circ}C$
Humidité	$Q1 = 59.0\%$	$me = 73.0\%$	$Q3 = 85.0\%$	$m = 71.2$	$\sigma = 17.1\%$

#### 4.1 Premier modèle : Moyenne de la température et de l'humidité par jour de l'année et par heure

Nous considérons dans ce modèle que la température et l'humidité sont cycliques de période un an et nous prévoyons qu'un jour donné à une heure donnée il fera la moyenne des température et humidité qu'il a fait ce jour ci lors des années précédentes. Calcul des moyennes sur la base d'entraînement :

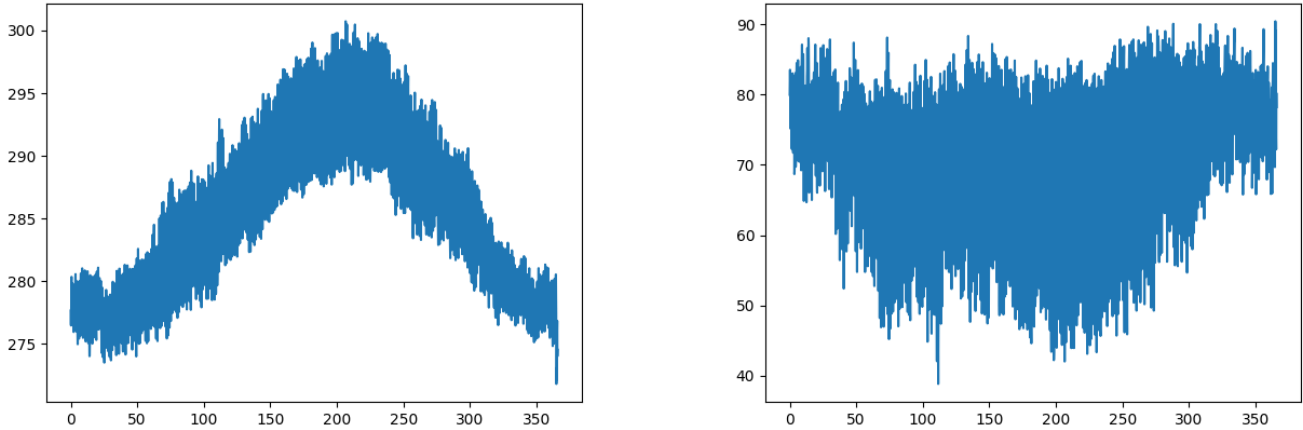


FIGURE 2 – Courbes des moyennes de température et d'humidité de la base d'apprentissage par jour et heure

Écart type de la température	$4.28^{\circ}C$
Écart type de l'humidité	$14.8\%$

#### 4.2 Deuxième modèle : Modèle de la persistance journalière

La température et l'humidité varient selon des cycles jours nuit. Nous considéreront ici que ces paramètres varient peu d'un jour sur l'autre. Nous prédisons donc que dans x jours il fera la même température que x jours

	Prévision à x jour	1	2	3	4	5	6
avant à la même heure.	Écart type de la température	$2.59^{\circ}C$	$11.1^{\circ}C$	$12.3^{\circ}C$	$12.9^{\circ}C$	$13.1^{\circ}C$	$13.0^{\circ}C$
	Écart type de l'humidité	$15.0\%$	$16.8\%$	$17.5\%$	$17.7\%$	$17.6\%$	$17.7\%$

#### 4.3 Troisième modèle : Approximation de la courbe des températures par une somme de fonctions sinusoïdales

À priori les conditions météorologiques sont cycliques par rapport à l'année et au jour. Nous pouvons donc nous demander si il est judicieux de modéliser ses paramètre par une somme de sinusoïdes. Nous réalisons donc une transformation de Fourier pour nous demander si la méthode est appropriée.

Nous concluons que la méthode est appropriée.

Sur les deux courbes on observe des pics aux fréquence de  $1/365$  jours, 1jours et  $1/2$  jours et  $1/3$  de jours.

On approxime donc les paramètres par des fonctions de la forme :

$$f(t) = A\cos(\omega t + \phi_1) + B\cos(\omega 2t + \phi_2) + C$$

$$g(t) = A\cos(\omega t + \phi_1) + B\cos(\omega 2t + \phi_2) + C + D\cos(\omega 3t + \phi_3)$$

$$h(t) = A\cos(\omega t + \phi_1) + B\cos(\omega 2t + \phi_2) + C + D\cos(\omega 3t + \phi_3) + E\cos(\omega 4t + \phi_4)$$

avec  $\omega 1 = 2\pi$  (période journalière) ,  $\omega 2 = 2\pi/365$  (période annuelle),  $\omega 3 = 4\pi$  (période demi-journalière) et  $\omega 4 = 6\pi$  (période d'un tiers de journée).

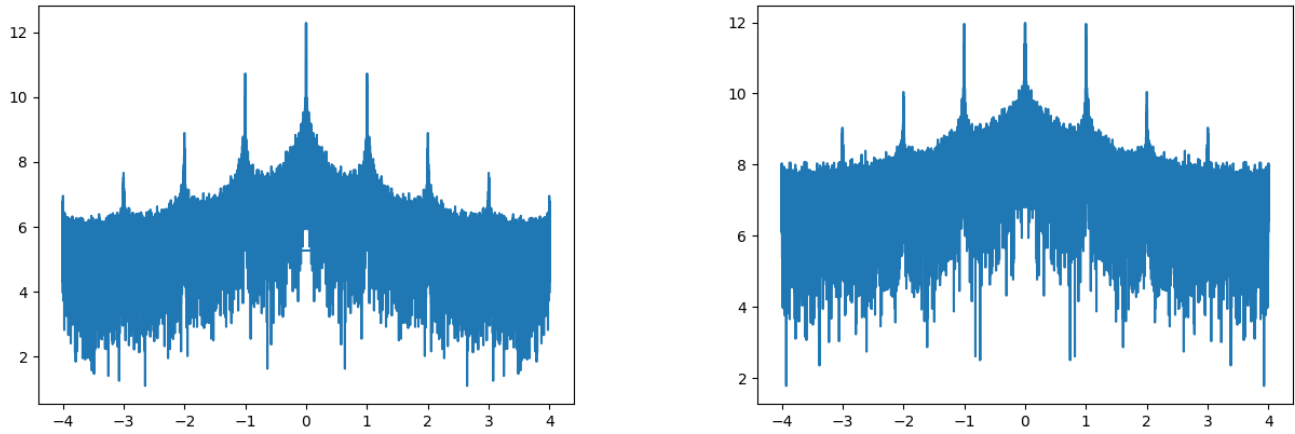
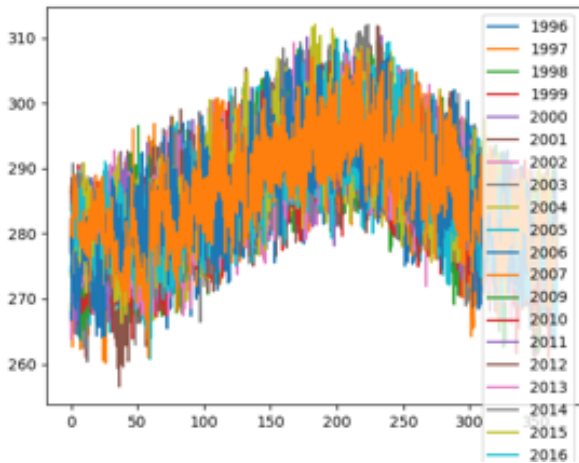


FIGURE 3 – transformation de Fourier discrète de la température et de l’humidité de la base d’entraînement (auquelles ont a enlevé la valeur moyenne moins un)



Il faut ensuite déterminer les valeurs de des coefficients et des déphasage. On utilise alors la méthode du gradient :

- On crée une fonction écart ( $u(t)$ ) qui calcule l’écart entre les prévisions de  $u(t)$  et les valeurs de la base d’entraînement. On cherche alors à déterminer le minimum de écart en fonction des coefficients et des déphasages.
- On initialise aléatoirement ses paramètres
- On dérive écart par rapport à chaque paramètre
- On modifie chaque paramètre  $X$  en faisant  $X = X - dX$
- On répète les deux points précédents un grand nombre de fois

## 5 Exemples Latex

### 5.1 subsection

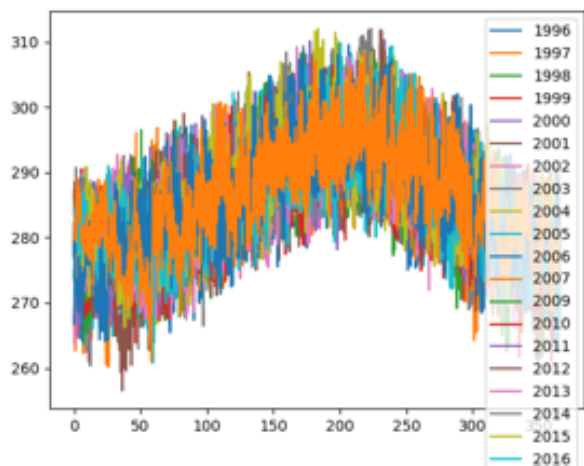
#### 5.1.1 q;,c

#### 5.1.2 Donner des figures

#### 5.1.3 Citer des figures

Je cite la figure 3

### 5.2 Placer un tableau



Ow	15.99945	-0.8476	SPCE
Hw	1.0079	0.4238	SPCE
Na	22.9897	1.0000	
Cl	35.45	-1.0000	
Os	15.9994	-1.000	
Ti	47.8670	2.000	

TABLE 1 – , Charges for TiO<sub>2</sub>/W/Na/Cl

### 5.3 Citer un tableau

Je cite mon tableau 1.

## 6 Références