## Analyse des signalements et des données Météo DSK en Île-de-France

In “smartick\_hum\_meteo\_idf\_v6.R”

# 3. Analyse des **Températures** moyennes DSK vs données humaines (méthode Alice Favre)

Nb signalements : 1 746 (n)

Nb de d’enregistrements météo DSK ( ??? mailles \* 1 746 signalements = ???? si pas de jours avec plusieurs piqûres) : 15 936 (n fait seulement 15 936 lignes en ÎdF du 15 juillet 17 au 5 avril 2020)

## 3.1. Histogramme des Températures moyennes "temperature" pour les signalements et "temperature" pour DSK = (tempHigh + tempLow)/2 pour 700 points darksky



### 3.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 1325.8, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median temp

summary(humdataIDF$temperature)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -1.28 15.72 18.57 18.21 21.41 31.30

summary(DSKdataIDF$temperature)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -7.46 7.02 11.62 12.20 17.33 32.22

### 3.2.3. Calcul d'un IC via le test de Wilcoxon sur la médiane (cf. section 3.2 de Poinsot et C:\3VG\MSH\Lyme\Smartick17\data\TCD7\R\IC\_foret\_AL\_CLC5.R)

wilcox.test(humdataIDF$temperature, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [18.2 ; 18.6] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$temperature, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [12.0 ; 12.2] autour de la /!\ médiane /!\

# 4. Analyse des **températures maximales quotidiennes** DSK vs données humaines (méthode Alice Favre)

##.4.1. Histogramme des températures maximales quotidiennes "temperaturehigh" pour les signalements et "temperaturehigh" pour DSK pour 700 points darksky



## 4.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 1269.2, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median temp high

summary(humdataIDF$temperaturehigh)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -0.13 20.49 23.93 23.53 27.15 41.44

summary(DSKdataIDF$temperaturehigh)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -3.26 10.29 15.73 16.57 22.82 42.21

#### 4.2.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$temperaturehigh, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [23.5 ; 24.0] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$temperaturehigh, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [16.20 ; 16.24] autour de la /!\ médiane /!\

# 5. Analyse de l'humidité relative DSK vs données humaines (méthode Alice Favre)

## 5.1. Histogramme de l**'humidité relative** "humidity" pour les signalements et "humidity" pour DSK pour 700 points darksky



## 5.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 385.49, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median humidity

summary(humdataIDF$humidity)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# 34.00 62.00 68.00 68.95 76.00 96.00

summary(DSKdataIDF$humidity)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# 28.00 66.00 76.00 74.29 84.00 99.00

## 5.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$humidity, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [68.5 ; 69.5] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$humidity, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [74.5 ; 75.0] autour de la /!\ médiane /!\

# 6. Analyse des **points de rosée** DSK vs données humaines (méthode Alice Favre)

## 6.1. Histogramme des **points de rosée** "dewpoint" pour les signalements et "dewpoint" pour DSK pour 700 points darksky



## 6.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 1184.9, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median dewpoint

summary(humdataIDF$dewpoint)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -5.400 9.385 12.015 11.690 14.828 22.180

summary(DSKdataIDF$ dewpoint)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# -13.790 3.260 7.180 6.978 11.190 22.320

## 6.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$ dewpoint, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [11.7 ; 12.1] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$ dewpoint, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [7.1 ; 7.2] autour de la /!\ médiane /!\

# 7. Analyse des **pressions atmosphériques** DSK vs données humaines (méthode Alice Favre)

## 7.1. Histogramme des **pressions atmosphériques** "pressure" pour les signalements et "pressure" pour DSK pour 700 points darksky



## 7.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 0.84824, df = 1, p-value = 0.3571 => p-value NON significative, les 2 échantillons ne sont pas significativement différents !

Median pressure

summary(humdataIDF$pressure)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 985.6 1013.4 1016.9 1016.9 1021.2 1045.8 553

summary(DSKdataIDF$ pressure)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 977.1 1011.3 1017.4 1016.9 1022.8 1046.0 182

## 7.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$ pressure, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [1016.8 ; 1017.5] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$ pressure, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [1017.0 ; 1017.3] autour de la /!\ médiane /!\

# 8. Analyse des **vitesses moyennes du vent** DSK vs données humaines (méthode Alice Favre)

## 8.1. Histogramme des **vitesses moyennes du vent** "windspeed" pour les signalements et "windspeed" pour DSK pour 700 points darksky



## 8.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 191.95, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median windspeed

summary(humdataIDF$windspeed)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# 0.220 1.903 2.605 2.700 3.320 10.820

summary(DSKdataIDF$windspeed)

# Min. 1st Qu. Median Mean 3rd Qu. Max.

# 0.070 2.060 3.040 3.369 4.290 12.540

## 8.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$windspeed, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [2.6 ; 2.7] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$windspeed, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [3.16 ; 3.22] autour de la /!\ médiane /!\

# 9. Analyse des **visibilités** DSK vs données humaines (méthode Alice Favre)

## 9.1. Histogramme des **visibilités** "visibility" pour les signalements et "visibility" pour DSK pour 700 points darksky

/!\ GROS PROBLEME DE DONNEES /!\

## 9.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 109.58, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median visibility

summary(humdataIDF$visibility)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 2.558 9.873 10.003 11.675 16.093 16.093 1

summary(DSKdataIDF$visibility)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.188 9.927 14.364 12.791 16.093 16.093 24

## 9.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$visibility, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [11.8 ; 12.4] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$visibility, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [12.9 ; 13.0] autour de la /!\ médiane /!\

# 10. Analyse des **couverts nuageux** DSK vs données humaines (méthode Alice Favre)

## 10.1. Histogramme des **couverts nuageux** "cloudcover" pour les signalements et "cloudcover" pour DSK pour 700 points darksky



## 10.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 140.87, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median cloudcover

summary(humdataIDF$cloudcover)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.00 44.00 66.00 59.32 78.00 100.00 169

summary(DSKdataIDF$cloudcover)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.0 48.0 74.0 65.2 88.0 100.0 992

## 10.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$cloudcover, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [61.0 ; 63.5] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$cloudcover, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [68.0 ; 69.0] autour de la /!\ médiane /!\

# 11. Analyse des **vitesses des rafales** DSK vs données humaines (méthode Alice Favre)

## 11.1. Histogramme des **vitesses des rafales** "windgust" pour les signalements et "windgust" pour DSK pour 700 points darksky



## 11.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 224.31, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median windgust

summary(humdataIDF$windgust)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.500 4.450 6.190 6.806 8.710 27.580 541

summary(DSKdataIDF$windgust)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.180 5.080 8.180 8.943 12.370 28.820 1787

## 11.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$windgust, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [6.3 ; 6.7] autour de la /!\ médiane /!\

wilcox.test (DSKdataIDF$windgust, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [8.7 ; 8.8] autour de la /!\ médiane /!\

# 12. Analyse de l’indice de **rayonnement ultra-violet** (UV) DSK vs données humaines (méthode Alice Favre)

## 12.1. Histogramme de l’indice de **rayonnement ultra-violet** (UV) "uvindex" pour les signalements et "uvindex" pour DSK pour 700 points darksky



## 12.2.2. Test si distribution non normale

## Kruskal-Wallis chi-squared = 1783.2, df = 1, p-value < 2.2e-16 => p-value significative, les 2 échantillons sont significativement différents !

Median uvindex

summary(humdataIDF$uvindex)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.00 5.00 6.00 5.55 7.00 9.00 73

summary(DSKdataIDF$uvindex)

# Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

# 0.000 1.000 3.000 3.129 5.000 9.000 338

## 12.3. Calcul d'un IC via le test de Wilcoxon sur la médiane

wilcox.test(humdataIDF$uvindex, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [5.500009 ; 5.500037] autour de la /!\ médiane /!\ **VALEURS DES BORNES OK ≠ France entière** /!\

wilcox.test (DSKdataIDF$uvindex, conf.int=TRUE)$conf.int

### Soit un IC à 95% = [3.00003 ; 3.00005] autour de la /!\ médiane /!\