

Annexes

Annexe 1 : détail de la revue de littérature scientifique mondiale

- I. Excess mortality in England and Wales, and in Greater London, during the 1995 heatwave (Rooney et al., 1998)

L'étude de Rooney et al. offre une perspective précoce sur les conséquences mortelles des vagues de chaleur, se concentrant sur l'impact dévastateur de la canicule de 1995 en Angleterre et au Pays de Galles. Cette recherche révèle une augmentation significative de la mortalité, avec un excès estimé de 8,9 %, équivalent à 619 décès supplémentaires pendant la période caniculaire. Les maladies respiratoires et cérébrovasculaires ont été particulièrement importantes, soulignant une vulnérabilité accrue des systèmes respiratoire et cardiovasculaire pendant les vagues de chaleur. L'étude met en évidence une corrélation entre la surmortalité et l'augmentation de la pollution atmosphérique, suggérant que jusqu'à 62 % des décès supplémentaires pourraient être attribués à ce facteur. Elle a aussi mis en évidence l'impact négatif qu'a l'effet de chaleur urbain sur la santé des habitants du Grand Londres. Cette constatation illustre la complexité des interactions entre les vagues de chaleur, la pollution atmosphérique et les caractéristiques urbaines.

En termes de méthodes épidémiologiques, l'étude a été menée de la manière suivante :

1. Extraction des données de mortalité pour les années 1993–95, incluant le nombre quotidien de décès par sexe, groupe d'âge, et cause.
2. Obtention des données météorologiques, notamment la Central England Temperature (CET) et les températures quotidiennes de Londres pour 1995.
3. Analyse temporelle en traçant le nombre quotidien de décès et la CET moyenne pour mai à septembre 1995.
4. Calcul de la mortalité excédentaire en comparant les décès observés avec les valeurs moyennes mobiles, séparément pour 1993–94 et 1995.
5. Estimation de la mortalité liée à la pollution de l'air en utilisant des coefficients de risque spécifiques à la saison pour Londres.
6. Considération des effets indépendants des polluants, avec l'utilisation de coefficients de mortalité de Londres pour estimer l'impact probable de la pollution accrue sur la mortalité quotidienne associée à la vague de chaleur en Angleterre et au Pays de Galles.

Cette étude de 1998 est une base de travail intéressante pour la santé publique, préconisant des améliorations des conditions de logement, une protection augmentée des populations vulnérables et l'importance d'alertes météorologiques lors d'événements extrêmes. Selon nous, les résultats, même deux décennies plus tard, restent pertinents dans le contexte du changement climatique mondial. Ils mettent en évidence la nécessité de stratégies d'adaptation face à des vagues de chaleur qui s'annoncent inévitablement plus fréquentes à l'avenir.

II. The role of increasing temperature variability in European summer heatwaves (Schär et al., 2004)

Cette étude explore l'impact des vagues de chaleur pendant l'été 2003 en Europe. Elle met en lumière l'augmentation de la variabilité des températures, soulignant que la croissance des vagues de chaleur est due non seulement à une élévation des températures moyennes mais également à une variabilité accrue. Les projections basées sur un scénario à concentrations atmosphériques accrues de gaz à effet de serre prévoient une augmentation significative de la variabilité de la température, particulièrement en Europe centrale et orientale. Les résultats soulignent que l'année 2003 peut être expliquée par une augmentation de la variabilité d'une année à l'autre, résultant du forçage des gaz à effet de serre. L'article conclut en disant que cette tendance pourrait influencer l'incidence future des vagues de chaleur et des sécheresses, comme le montre la représentation imagée des scénarios anticipés par l'équipe de chercheurs (Fig. 1).

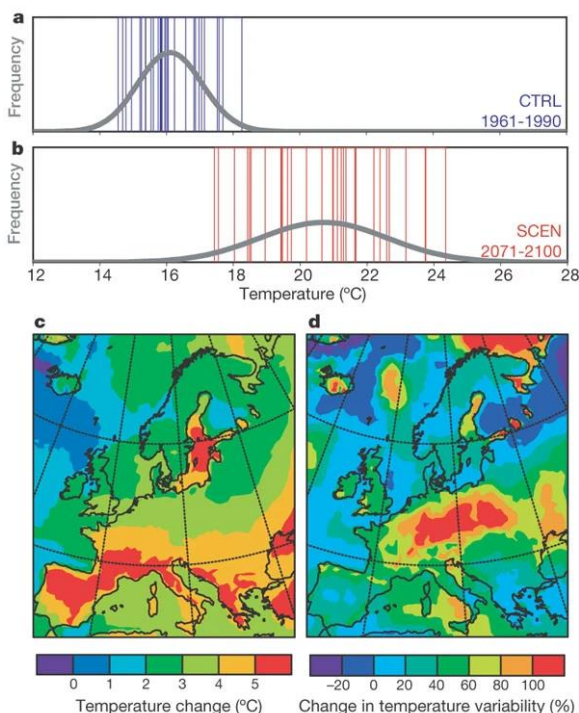


Figure 1. *a, b, Distribution statistique des températures estivales dans le nord de la Suisse pour CTRL et SCEN, respectivement. c, Changement de température associé (SCEN-CTRL, °C). d, Changement de variabilité exprimé comme changement relatif de l'écart-type des moyennes estivales ((SCEN-CTRL)/CTRL, %).*

En termes de méthodes épidémiologiques, l'étude a été menée de la manière suivante :

1. **Analyse de l'été 2003** : Utilisation de données de réanalyse ERA-40 et d'analyses opérationnelles de l'ECMWF pour évaluer l'anomalie de température continentale pendant l'été 2003.
2. **Estimation du temps de retour** : Adoption d'un concept supposant des températures JJA¹ identiquement distribuées. Les paramètres de distribution sont estimés à partir des données de la période de référence, et le temps de retour de l'événement est calculé avec des intervalles de confiance obtenus par échantillonnage paramétrique.
3. **Simulations du changement climatique** : Trois modèles numériques différents sont employés pour simuler les conditions de contrôle et de scénario. La variabilité de la température simulée est influencée par le type de sol composant le modèle considéré.

Nous observons au travers de ce travail scientifique que la pollution atmosphérique semble, une nouvelle fois, jouer un rôle important dans les problèmes de santé publique causés par de telles vagues de chaleur.

III. Climate change and extreme heat events (Luber et al., 2008)

En 2008, Luber et al. se concentrent sur l'ampleur des impacts des épisodes de chaleur extrême (EHE pour Extreme Heat Events) sur la santé mondiale. Leur analyse examine les facteurs épidémiologiques de risque liés à la mortalité et qui résultent de l'exposition à une chaleur excessive, ainsi que les futurs éléments de mortalité associés à la chaleur, tels que le réchauffement climatique, l'effet d'îlot de chaleur urbain et le vieillissement démographique.

Les EHE sont identifiés comme la principale cause de mortalité liée aux conditions météorologiques à l'échelle mondiale, surpassant le nombre annuel de décès combinés dus aux ouragans, éclairs, tornades, inondations et tremblements de terre. Ce document souligne que l'exposition à une chaleur extrême constitue un enjeu majeur de santé publique à l'échelle mondiale, en particulier dans les zones moins adaptées en termes d'infrastructures, qui se trouvent bien souvent dans les pays en voie de développement.

Le changement climatique a déjà influencé les vagues de chaleur, avec une tendance à la hausse observée dans leur fréquence, indiquant une potentielle augmentation globale. Les

¹ JJA : Juin-Juillet-Août

modèles climatiques prévoient une intensification future de ces épisodes en termes d'ampleur et de durée, touchant diverses régions du globe. Les populations mondiales pourraient ainsi subir davantage de maladies et de décès en réaction aux variations des températures.

Les tendances démographiques, telles que le vieillissement de la population et la migration vers les zones urbaines, augmentent la vulnérabilité aux EHE à l'échelle mondiale. Le document souligne l'impact significatif de ces conditions sur les populations vulnérables, notamment les personnes âgées, handicapées, sans accès à la climatisation et souffrant de problèmes de santé préexistants.

Cette recherche plutôt rétrospective que prospective s'est basée sur un ensemble de données disponibles pour en tirer des conclusions :

1. **Analyse des vagues de chaleur (EHE) :** L'étude évalue l'impact des vagues de chaleur en se basant sur des événements passés, tels que les canicules de 1995 à Chicago et de 2003 en Europe.
2. **Analyse de la morbidité et mortalité aux États-Unis :** L'étude examine les impacts sur la santé, en se concentrant sur la mortalité et la morbidité associées aux vagues de chaleur, en mettant en évidence les événements de 1995 à Chicago et de 2003 en Europe.
3. **Projection des futurs scénarios :** L'impact du changement climatique sur les vagues de chaleur est évalué en utilisant des modèles climatiques et des scénarios de l'IPCC². Les projections suggèrent une augmentation de l'intensité et de la durée des vagues de chaleur.

En conclusion, l'article publié par l'équipe américaine, plus concentré sur un constat épidémiologique que sur de nouvelles méthodes de collecte de données, met en avant la nécessité d'efforts de santé publique pour prévenir les maladies liées à la chaleur lors des périodes de chaleur extrême, en prêtant une attention particulière aux facteurs de risque au niveau individuel et communautaire. Il propose également des stratégies de conception environnementale, telles que les "villes fraîches", comme moyen d'atténuer les effets des températures extrêmes à l'échelle mondiale. Les auteurs soulignent que les praticiens de la santé publique ont un rôle actif à jouer dans l'élaboration de mesures d'adaptation, notamment des plans de réponse à la chaleur basés sur les meilleures preuves issues d'études épidémiologiques et de modèles écologiques sur la relation entre les risques et les résultats pour la santé.

² IPCC : Intergovernmental Panel on Climate Change

IV. The hot summer of 2010: redrawing the temperature record map of Europe (Barriopedro et al., 2011)

L'article scientifique explore l'impact des méga-vagues de chaleur en Europe pendant les étés 2003 et 2010. Les chercheurs démontrent que la méga-vague de chaleur de 2010 a dépassé celle de 2003 (sur laquelle les auteurs des deux premiers travaux de cette section avaient déjà travaillé), établissant des records de température saisonnière sur 500 ans dans 50 % de l'Europe. Selon leur scénario de projection, la probabilité d'occurrence de méga-vagues de chaleur devrait augmenter de 5 à 10 fois au cours des 40 prochaines années.

Il met en avant les effets dévastateurs de ces vagues de chaleur, incluant les décès liés à la chaleur, les mauvaises récoltes et des pertes économiques substantielles. Il examine l'évolution spatiale et temporelle de l'événement de 2010, le comparant avec les vagues de chaleur historiques, soulignant l'ampleur exceptionnelle de celle de l'été 2010. L'étude prédit une augmentation future de la fréquence de ces événements extrêmes, présentant des risques significatifs pour des zones densément peuplées sans stratégies d'adaptation.

En termes de méthodes épidémiologiques, l'étude a été menée de la manière suivante :

1. **Analyse temporelle** des zones avec des températures record sur base de données quotidiennes moyennes à long terme provenant de réanalyses remontant à 1871.
2. **Analyse statistique** de la distribution des températures estivales européennes sur une période de 510 ans (données reconstruites via des instruments d'extrapolation statistique) et mise en évidence des étés extrêmes.
3. **Projections multi-modèles** des périodes de retour des températures maximales estivales régionales basée sur 11 modèles climatiques régionaux différents
4. **Analyse de la concordance** entre les modèles pour évaluer les projections futures des vagues de chaleur.

Les chercheurs utilisent donc des données de réanalyses et des projections de plusieurs modèles pour offrir un aperçu de la fréquence et de l'ampleur potentielles des futures vagues de chaleur en Europe. Ils soulignent l'importance de faire face à l'augmentation potentielle de la fréquence de tels événements extrêmes et de mettre en place des stratégies adaptatives pour atténuer leurs impacts négatifs. L'impact de l'augmentation des concentrations de gaz à effet de serre sur l'amplification de la variabilité des températures estivales en Europe est également mis en évidence.

Comme dit précédemment, les résultats suggèrent une probabilité accrue de subir des méga-vagues de chaleur sur de vastes zones, avec des conséquences potentiellement graves sans gestion efficace. En conclusion, ce travail fournit des informations sur les impacts à long terme et la fréquence potentielle des méga-vagues de chaleur, soulignant encore une fois la nécessité de mesures proactives pour relever les défis posés par ces événements extrêmes en Europe.

V. Increasing trends in regional heatwaves (Perkins-Kirkpatrick et al., 2020)

Cette recherche, la plus récente de celles considérées dans cette section, se concentre sur l'impact des vagues de chaleur, en définissant ces périodes comme des épisodes prolongés de chaleur excessive, utilisant la définition de l'OMM que nous avons explicitée précédemment.

Partant des données de température de la Terre, les résultats révèlent que la fréquence des vagues de chaleur connaît le changement le plus rapide et le plus significatif dans presque toutes les régions. De plus, l'étude souligne que l'intensité, la fréquence et la durée des vagues de chaleur ont augmenté, avec une tendance qui s'accélère depuis les années 1950. Les tendances croissantes en matière de fréquence, de durée et d'intensité cumulée des vagues de chaleur historiques se sont également accélérées dans de nombreuses régions. Les chercheurs mettent en avant l'importance d'évaluer ces tendances sur plusieurs décennies en raison de l'influence significative de la variabilité climatique.

L'étude évalue l'inégalité des impacts des vagues de chaleur, révélant que les pays en développement sont plus fortement touchés en raison d'un manque de capacité d'adaptation et de contraintes culturelles variables, comme l'avait illustré Luber et al. De plus, elle souligne que les impacts des vagues de chaleur sont susceptibles d'augmenter en raison d'un réchauffement climatique accru, entraînant des conséquences plus graves, voire irréversibles, dans certains secteurs. La chaleur cumulée ressentie pendant les vagues de chaleur a nettement augmenté à l'échelle mondiale et régionale en raison de l'augmentation du nombre total de jours de canicule. L'étude suggère également la nécessité de mesurer l'intensité des vagues de chaleur en évaluant la chaleur supplémentaire générée, car cela pourrait être permettre d'évaluer plus complètement les impacts de celles-ci.

En termes de méthodes épidémiologiques, l'étude a été menée de la manière suivante :

1. Utilisation des données de température quotidienne maximale provenant du jeu de données Berkeley Earth, une grille mondiale de température de surface terrestre
2. Utilisation de la définition de vague de chaleur basée sur la température maximale quotidienne (CTX90), où une vague de chaleur est détectée lorsqu'au moins 3 jours consécutifs dépassent le 90e percentile de la température maximale quotidienne pour chaque jour du calendrier, se basant sur une fenêtre mobile de 15 jours des températures maximales quotidiennes sur la période de référence 1961-1990
3. Évaluation des vagues de chaleur selon 4 caractéristiques : l'intensité, la fréquence, la durée et la chaleur cumulative (qui représente la chaleur supplémentaire produite par les vagues de chaleur au cours d'une saison donnée)
4. Calcul des tendances décennales à l'aide de l'estimateur de pente de Sen's Kendal, un outil non paramétrique robuste contre les valeurs aberrantes.
5. Analyse réalisée à l'échelle mondiale et régionale, les changements dans la chaleur cumulative et la fréquence des vagues de chaleur sont évalués et exprimés en pourcentage pour permettre une comparaison directe entre les deux métriques.

L'étude fournit la première analyse mondiale et régionale complète des changements observés lors des vagues de chaleur. Les résultats démontrent le rythme inquiétant de l'augmentation des vagues de chaleur depuis le milieu du XXe siècle, avec des tendances inégalement réparties et plus prononcées dans les régions sujettes à des impacts négatifs plus importants, que nous avons mentionnés ci-dessus. Ces résultats ont des implications majeures pour tous les systèmes exposés de manière chronique à la chaleur, soulignant une énième fois la nécessité de mettre en place des stratégies de gestion et d'adaptation appropriées. La recherche recommande une étude plus approfondie des répercussions des changements dans la chaleur cumulée et des tendances globales des vagues de chaleur, ainsi qu'un examen attentif de l'accélération de ces tendances dans le contexte du changement climatique anthropique croissant.

Annexe 2 : historique des vagues de chaleur en Belgique (1992 à nos jours)

Année	Période	Durée (j)	Poids (DJU)³	Intensité (DJU/j)	Total j >30°C	T°C max
1994	19/07 → 06/08	19	59.6	3.14	8	33.9
1995	07/07 → 13/07	7	22.7	3.24	3	33.0
1995	20/07 → 12/08	24	47.8	1.99	8	34.0
1997	05/08 → 25/08	21	55.7	2.65	4	31.9
1998	09/05 → 14/05	6	15.1	2.51	3	31.0
2003	01/08 → 13/08	13	58.2	4.47	6	33.8
2005	18/06 → 25/06	8	28.4	3.55	4	32.9
2006	09/06 → 13/06	5	10.5	2.10	3	31.3
2006	15/07 → 30/07	16	66.3	4.14	7	36.2
2010	23/06 → 14/07	22	42.0	1.91	6	33.9
2013	21/07 → 27/07	7	25.7	3.67	3	32.4
2015	30/06 → 05/07	6	24.2	4.03	3	34.5
2016	23/08 → 27/08	5	18.3	3.66	3	32.5
2017	17/06 → 22/06	5	22.6	4.52	4	32.4
2018	13/07 → 07/08	26	82.2	3.16	9	35.4
2019	23/06 → 30/06	8	23.5	2.94	3	32.6
2019	22/07 → 26/07	5	34.7	6.94	4	39.7
2019	23/08 → 28/08	6	19.7	3.28	3	33.3
2020	05/08 → 16/08	12	57.2	4.77	8	35.9
2022	09/08 → 16/08	8	32.1	4.01	5	32.1
2023	08/06 → 17/06	10	26.8	2.68	3	31.2
2023	04/09 → 11/09	8	23.5	2.94	6	31.9

NB : il s'agit des valeurs journalières 0-24 enregistrées à Uccle. Les auteurs se sont efforcés de ne pas couper des vagues de chaleur de façon artificielle et ont donc admis un jour sous le seuil des 25°C si, et seulement si, un jour au-dessus de 30°C était encore présent par la suite. (Deux jours sous 25°C pour deux jours au-dessus de 30°C, etc.).

³ Le DJU (degré jour unifié) est une valeur représentative de l'écart entre la température moyenne d'une journée donnée et un seuil de température préétabli (ici 20°C). On additionne les DJU d'une même période pour obtenir le poids de celle-ci

Annexe 3 : code R (graphes, modèles statistiques)

#GRAPHE MORT/JOUR DE LA SEMAINE

```
# Installer et charger les packages nécessaires
```

```
install.packages("viridis")
install.packages("ggplot2")
library(ggplot2)
library(viridis)
```

```
# Tableau de données codes postaux et noms de province
```

```
postal_code_mapping <- data.frame(postal_code = c("4000", "10000", "20001", "20002", "30000",
"40000", "50000", "60000", "70000", "80000", "90000"),
                                province = c("Bruxelles-Capitale", "Anvers", "Brabant flamand", "Brabant wallon",
"Flandre occidentale", "Flandre orientale", "Hainaut", "Liège", "Limbourg", "Luxembourg", "Namur"))
```

```
mortality_province <- mortality %>%
  group_by(Localisation, Day_of_week = weekdays(Date)) %>%
  summarize(total_deaths = sum(Nombre_de_décès))
```

```
# Fusionnez données avec table de correspondance
```

```
mortality_province_code <- merge(mortality_province, postal_code_mapping, by.x = "Localisation",
by.y = "postal_code", all.x = TRUE)
mortality_province_code <- mortality_province %>%
  left_join(postal_code_mapping, by = c("Localisation" = "postal_code"))
```

```
# graphique avec les ajustements
```

```
ggplot(mortality_province_code, aes(x = Day_of_week, y = total_deaths, fill = province)) +
  geom_bar(stat = "identity", position = "stack") +
  scale_fill_viridis(discrete = TRUE, option = "D") +
  labs(
    title = "Mortalité par jour de la semaine et par province",
    x = "Jour de la semaine",
    y = "Nombre de décès",
    fill = "Province"
  ) +
  theme_minimal(base_size = 15)
theme(
  plot.title = element_text(face = "bold", size = 20, hjust = 0.5),
  axis.title = element_text(face = "bold", size = 15),
  axis.text = element_text(face = "bold", size = 14)
  legend.position = "right",
  legend.title = element_text(face = "bold", size = 15),
  legend.text = element_text(size = 12)
)
```

#GRAPHE MORT/JOUR/AGE

```
library(dplyr)
library(ggplot2)
library(zoo)
library(viridis)

# colonne pour le jour de l'année
mortality <- mortality %>%
  mutate(Day_of_Year = as.numeric(format(Week, "%j")))

# moyenne mobile centrée et extrapoler valeurs manquantes
calculate_rolling_average <- fonction(data, k) {
  rolling_average <- zoo::rollmean(data$total_deaths, k = k, align = "center", fill = NA)

  # Extrapolation des valeurs manquantes aux extrémités
  na_indices <- which(is.na(rolling_average))
  if (length(na_indices) > 0) {
    if (min(na_indices) == 1) {
      fit_start <- lm(total_deaths ~ as.numeric(Date), data = data[1:k, ])
      rolling_average[1:(k/2)] <- predict(fit_start, newdata = data.frame(Date =
as.numeric(data$Date[1:(k/2)])))
    }
    if (max(na_indices) == length(rolling_average)) {
      fit_end <- lm(total_deaths ~ as.numeric(Date), data = data[(nrow(data) - k + 1):nrow(data), ])
      rolling_average[(length(rolling_average) - k/2 + 1):length(rolling_average)] <- predict(fit_end,
newdata = data.frame(Date = as.numeric(data$Date[(length(rolling_average) - k/2 +
1):length(rolling_average)])))
    }
  }

  return(rolling_average)
}

# classe d'âge 0-24 ans
mortality_0_24 <- mortality %>%
  filter(Classe_d_age_combinee == "0-24")

# Agréger données par jour
mortality_daily_0_24 <- mortality_0_24 %>%
  group_by(Date) %>%
  summarise(total_deaths = sum(Nombre_de_décès)) %>%
  ungroup()

# Calculer moyenne mobile 1 an et extrapoler les valeurs manquantes
mortality_daily_0_24 <- mortality_daily_0_24 %>%
  mutate(rolling_average = calculate_rolling_average(mortality_daily_0_24, 1825))

# Filtrer données 25-64 ans
mortality_25_64 <- mortality %>%
  filter(Classe_d_age_combinee == "25-64")
```

```

# Agréger
mortality_daily_25_64 <- mortality_25_64 %>%
  group_by(Date) %>%
  summarise(total_deaths = sum(Nombre_de_décès)) %>%
  ungroup()

# Calculer moyenne mobile 1 an et extrapoler les valeurs manquantes
mortality_daily_25_64 <- mortality_daily_25_64 %>%
  mutate(rolling_average = calculate_rolling_average(mortality_daily_25_64, 1825))

# Filtrer 65-84 ans
mortality_65_84 <- mortality %>%
  filter(Classe_d_age_combinee == "65-84")

# Agréger
mortality_daily_65_84 <- mortality_65_84 %>%
  group_by(Date) %>%
  summarise(total_deaths = sum(Nombre_de_décès)) %>%
  ungroup()

# Calculer moyenne mobile 1 an et extrapoler les valeurs manquantes
mortality_daily_65_84 <- mortality_daily_65_84 %>%
  mutate(rolling_average = calculate_rolling_average(mortality_daily_65_84, 1825))

# Filtrer 85 ans et plus
mortality_85_plus <- mortality %>%
  filter(Classe_d_age_combinee == "85+")

# Agréger
mortality_daily_85_plus <- mortality_85_plus %>%
  group_by(Date) %>%
  summarise(total_deaths = sum(Nombre_de_décès)) %>%
  ungroup()

# Calculer moyenne mobile 1 an et extrapoler les valeurs manquantes
mortality_daily_85_plus <- mortality_daily_85_plus %>%
  mutate(rolling_average = calculate_rolling_average(mortality_daily_85_plus, 1825))

# graphe combiné avec moyennes mobiles pour chaque classe d'âge
combined_plot <- ggplot() +
  geom_line(data = mortality_daily_0_24, aes(x = Date, y = total_deaths, color = "0-24 ans"), alpha =
0.5) +
  geom_line(data = mortality_daily_25_64, aes(x = Date, y = total_deaths, color = "25-64 ans"), alpha
= 0.5) +
  geom_line(data = mortality_daily_65_84, aes(x = Date, y = total_deaths, color = "65-84 ans"), alpha
= 0.5) +
  geom_line(data = mortality_daily_85_plus, aes(x = Date, y = total_deaths, color = "85 ans et plus"),
alpha = 0.5) +
  geom_line(data = mortality_daily_0_24, aes(x = Date, y = rolling_average), color = "black") +
  geom_line(data = mortality_daily_25_64, aes(x = Date, y = rolling_average), color = "black") +
  geom_line(data = mortality_daily_65_84, aes(x = Date, y = rolling_average), color = "black") +
  geom_line(data = mortality_daily_85_plus, aes(x = Date, y = rolling_average), color = "black") +

```

```
scale_color_viridis(discrete = TRUE, option = "D") +
labs(title = "Évolution journalière de la mortalité par classe d'âge",
      x = "Date",
      y = "Nombre de décès") +
theme_minimal() +
guides(color = guide_legend(title = "Classe d'âge"))
```

```
# Afficher graphe combiné
print(combined_plot)
```

GRAPHE MORT/JOUR + SINUSOÏDALE

```
install.packages("minpack.lm")
```

```
library(dplyr)
```

```
library(ggplot2)
```

```
library(minpack.lm)
```

```
# Calculer la mortalité totale par jour de l'année
```

```
total_daily_mortality <- mortality %>%
```

```
  group_by(Date) %>%
```

```
  summarise(total_deaths = sum(Nombre_de_décès))
```

```
# Supprimer les valeurs à l'extrémité
```

```
cutoff <- max(total_daily_mortality$Day_of_Year) - 1
```

```
total_daily_mortality <- total_daily_mortality %>%
```

```
  filter(Day_of_Year <= cutoff)
```

```
# fonction sinusoidale
```

```
sinusoidal_function <- function(x, A, B, C, D) {
```

```
  A * sin(B * (x - C)) + D
```

```
}
```

```
# Ajuster fonction sinusoidale aux données
```

```
fit <- nlsLM(total_deaths ~ sinusoidal_function(Day_of_Year, A, B, C, D),
```

```
  data = total_daily_mortality,
```

```
  start = list(A = 1000, B = 2 * pi / 365, C = 0, D = mean(total_daily_mortality$total_deaths)))
```

```
# Extraire les coefficients ajustés
```

```
coef_fit <- coef(fit)
```

```
# Extraire les paramètres
```

```
params <- coef(fit)
```

```
A_final <- params["A"]
```

```
B_final <- params["B"]
```

```
C_final <- params["C"]
```

```
D_final <- params["D"]
```

```
# colonne pour la courbe sinusoidale ajustée
```

```
total_daily_mortality <- total_daily_mortality %>%
```

```
  mutate(sinusoidal_fit = sinusoidal_function(Day_of_Year, coef_fit['A'], coef_fit['B'], coef_fit['C']),
```

```
coef_fit['D']))
```

```
# graphique avec la courbe ajustée
ggplot(total_daily_mortality, aes(x = Day_of_Year, y = total_deaths)) +
  geom_line(color = "blue") +
  geom_line(aes(y = sinusoidal_fit), color = "red", size = 1.2) +
  labs(title = "Total de la mortalité journalière en Belgique (1992-2020)",
       x = "Mois de l'année",
       y = "Nombre total de décès") +
  theme_minimal() +
  scale_x_continuous(breaks = cumsum(c(0, 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31)),
                    labels = c("", "Janvier", "Février", "Mars", "Avril", "Mai", "Juin", "Juillet", "Août",
"Septembre", "Octobre", "Novembre", "Décembre")) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

#GRAPHES T°C-MORT/SEMAINE 2004-2010;2010-2015;2015-2020

```
# packages
install.packages("viridis")
install.packages("lubridate")
library(ggplot2)
library(dplyr)
library(lubridate)
library(viridis)

# données de mortalité 2004-2010
mortality_filtered_2004_2010 <- mortality %>%
  filter(Date >= as.Date("2004-01-01") & Date < as.Date("2010-01-01"))

# Agréger par semaine
mortality_weekly_2004_2010 <- mortality_filtered_2004_2010 %>%
  mutate(Week = floor_date(Date, "week")) %>%
  group_by(Week) %>%
  summarise(weekly_deaths = sum(Nombre_de_décès))

# données de température 2004-2010
meteo_data_filtered_2004_2010 <- meteo_data %>%
  filter(timestamp >= as.Date("2004-01-01") & timestamp < as.Date("2010-01-01"))

# Agréger
meteo_weekly_2004_2010 <- meteo_data_filtered_2004_2010 %>%
  mutate(Week = floor_date(timestamp, "week")) %>%
  group_by(Week) %>%
  summarise(avg_temperature = mean(air_temperature_avg, na.rm = TRUE))

# Fusionner les deux dataframes sur la colonne 'Week'
combined_data_2004_2010 <- merge(mortality_weekly_2004_2010, meteo_weekly_2004_2010, by
= "Week")
```

```

# Filtrer les extrémités
combined_data_2004_2010 <- combined_data_2004_2010 %>%
  filter(Week > as.Date("2004-01-07") & Week < as.Date("2009-12-25"))

# Tracer le graphique
p_2004_2010 <- ggplot(combined_data_2004_2010, aes(x = Week)) +
  geom_bar(aes(y = avg_temperature * 100), stat = "identity", fill = viridis(1, option = "G"), alpha =
0.3) + # Température en barres, multipliée par 100 pour l'échelle
  geom_line(aes(y = weekly_deaths), color = viridis(1, option = "B"), size = 1) +
  scale_y_continuous(
    name = "Nombre de décès",
    sec.axis = sec_axis(~./100, name = "Température moyenne (°C)", breaks = seq(0, 300, by = 20))
  ) +
  labs(title = "Comparaison de la mortalité et de la température moyenne en Belgique (2004-2010)",
    x = "Date") +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5, face = "bold", size = 16),
    axis.title = element_text(face = "bold"),
    legend.position = "bottom"
  )
)

# Afficher le graphique
print(p_2004_2010)

# mortalité 2010-2015
mortality_filtered_2010_2015 <- mortality %>%
  filter(Date >= as.Date("2010-01-01") & Date < as.Date("2015-01-01"))

# Agréger
mortality_weekly_2010_2015 <- mortality_filtered_2010_2015 %>%
  mutate(Week = floor_date(Date, "week")) %>%
  group_by(Week) %>%
  summarise(weekly_deaths = sum(Nombre_de_décès))

# température 2010-2015
meteo_data_filtered_2010_2015 <- meteo_data %>%
  filter(timestamp >= as.Date("2010-01-01") & timestamp < as.Date("2015-01-01"))

# Agréger
meteo_weekly_2010_2015 <- meteo_data_filtered_2010_2015 %>%
  mutate(Week = floor_date(timestamp, "week")) %>%
  group_by(Week) %>%
  summarise(avg_temperature = mean(air_temperature_avg, na.rm = TRUE))

# Fusionner les deux dataframes
combined_data_2010_2015 <- merge(mortality_weekly_2010_2015, meteo_weekly_2010_2015, by
= "Week")

```

```

# Filtrer les extrémités
combined_data_2010_2015 <- combined_data_2010_2015 %>%
  filter(Week > as.Date("2010-01-07") & Week < as.Date("2014-12-25"))

# Tracer le graphique
p_2010_2015 <- ggplot(combined_data_2010_2015, aes(x = Week)) +
  geom_bar(aes(y = avg_temperature * 100), stat = "identity", fill = viridis(1, option = "G"), alpha =
0.3) + # Température en barres, multipliée par 100 pour l'échelle
  geom_line(aes(y = weekly_deaths), color = viridis(1, option = "B"), size = 1)
  scale_y_continuous(
    name = "Nombre de décès",
    sec.axis = sec_axis(~./100, name = "Température moyenne (°C)", breaks = seq(0, 300, by = 20))
  ) +
  labs(title = "Comparaison de la mortalité et de la température moyenne en Belgique (2010-2015)",
    x = "Date") +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5, face = "bold", size = 16),
    axis.title = element_text(face = "bold"),
    legend.position = "bottom"
  )
)

# Afficher le graphique
print(p_2010_2015)

# mortalité 2015-2020
mortality_filtered <- mortality %>%
  filter(Date >= as.Date("2015-01-01") & Date < as.Date("2020-01-01"))

# Agréger
mortality_weekly <- mortality_filtered %>%
  mutate(Week = floor_date(Date, "week")) %>%
  group_by(Week) %>%
  summarise(weekly_deaths = sum(Nombre_de_décès))

# température 2015-2020
meteo_data_filtered <- meteo_data %>%
  filter(timestamp >= as.Date("2015-01-01") & timestamp < as.Date("2020-01-01"))

# Agréger
meteo_weekly <- meteo_data_filtered %>%
  mutate(Week = floor_date(timestamp, "week")) %>%
  group_by(Week) %>%
  summarise(avg_temperature = mean(air_temperature_avg, na.rm = TRUE))

# Fusionner les deux dataframes
combined_data <- merge(mortality_weekly, meteo_weekly, by = "Week")

# Filtrer les extrémités
combined_data <- combined_data %>%
  filter(Week > as.Date("2015-01-07") & Week < as.Date("2019-12-25"))

```

```

# Tracer le graphique
p <- ggplot(combined_data, aes(x = Week)) +
  geom_bar(aes(y = avg_temperature * 100), stat = "identity", fill = viridis(1, option = "G"), alpha =
0.3) + # Température en barres, multipliée par 100 pour l'échelle
  geom_line(aes(y = weekly_deaths), color = viridis(1, option = "B"), size = 1) + # Mortalité en ligne
  scale_y_continuous(
    name = "Nombre de décès",
    sec.axis = sec_axis(~./100, name = "Température moyenne (°C)", breaks = seq(0, 300, by = 20)) #
Diviser l'axe secondaire par 100 et définir les breaks
  ) +
  labs(title = "Comparaison de la mortalité et de la température moyenne en Belgique (2015-2020)",
    x = "Date") +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(hjust = 0.5, face = "bold", size = 16),
    axis.title = element_text(face = "bold"),
    legend.position = "bottom"
  )

# Afficher le graphique
print(p)

```

#NUAGE DE POINTS + COURBE DE TENDANCE

```

# Charger les packages nécessaires
library(dplyr)
library(ggplot2)
library(minpack.lm)
library(lubridate)

# Fusionner les dataframes
combined_data_full <- bind_rows(combined_data, combined_data_2004_2010,
combined_data_2010_2015)

# Ajouter une colonne pour le jour de l'année
combined_data_full <- combined_data_full %>%
  mutate(Day_of_Year = yday(Week))

# Définir la fonction sinusoidale
sinusoidal_function <- function(x, A, B, C, D) {
  A * sin(B * (x - C)) + D
}

# Ajuster la fonction sinusoidale aux données de mortalité
fit <- nlsLM(weekly_deaths ~ sinusoidal_function(Day_of_Year, A, B, C, D),
  data = combined_data_full,
  start = list(A = 1000, B = 2 * pi / 365, C = 0, D = mean(combined_data_full$weekly_deaths)))

```

```

# Résumé du modèle de régression non-linéaire
summary(fit)

# Calculer les valeurs prévues et les résidus
combined_data_full <- combined_data_full %>%
  mutate(fitted_deaths = predict(fit, newdata = combined_data_full),
         residuals = weekly_deaths - fitted_deaths)

# Calculer corrélation résidus et température
correlation <- cor(combined_data_full$avg_temperature, combined_data_full$residuals)

# analyse de régression linéaire
regression_model <- lm(residuals ~ avg_temperature, data = combined_data_full)

# Résumé du modèle de régression
summary(regression_model)

# graphe résultats
ggplot(combined_data_full, aes(x = avg_temperature, y = residuals)) +
  geom_point(color = "black", alpha = 0.5, size = 2) +
  labs(title = "Relation entre la température et les résidus de mortalité (2004-2020)",
       x = "Température moyenne (°C)",
       y = "Résidus de mortalité") +
  theme_minimal(base_size = 15) +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", size = 20),
    plot.subtitle = element_text(hjust = 0.5, size = 15),
    axis.title = element_text(face = "bold", size = 15),
    axis.text = element_text(face = "bold", size = 14)
  )

# Normalisation des résidus
combined_data_full$residuals_normalized <- scale(combined_data_full$residuals)

# graphique avec les nouvelles limites de l'axe y
ggplot(combined_data_full, aes(x = avg_temperature, y = residuals_normalized)) +
  geom_point(color = "black", alpha = 0.5, size = 2) +
  geom_smooth(method = "loess", color = "red") +
  scale_y_continuous(limits = c(-1, 1), name = "Résidus normalisés") +
  labs(title = "Relation entre la température et les résidus de mortalité (2004-2020)",
       x = "Température moyenne (°C)") +
  theme_minimal(base_size = 15) +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold", size = 20),
    axis.title = element_text(face = "bold", size = 15),
    axis.text = element_text(face = "bold", size = 14)
  )

```

```

# BOXPLOT résidus de mortalité entre jours de vague de chaleur et autres jours
library(dplyr)
library(minpack.lm)
library(lubridate)
library(ggplot2)

# Filtrer les données 1992-2020
mortality <- mortality %>%
  filter(Date >= as.Date("1992-01-01") & Date < as.Date("2020-01-01"))

# Agréger données journalières
mortality_daily <- mortality %>%
  group_by(Date) %>%
  summarize(total_deaths = sum(Nombre_de_décès),
            heatwave_day = any(heatwave_day))

# Ajouter une colonne pour le jour de l'année
mortality_daily <- mortality_daily %>%
  mutate(Day_of_Year = yday(Date))

# Définir la fonction sinusoïdale
sinusoidal_function <- function(x, A, B, C, D) {
  A * sin(B * (x - C)) + D
}

# Ajuster la fonction sinusoïdale aux données de mortalité journalière
fit <- nlsLM(total_deaths ~ sinusoidal_function(Day_of_Year, A, B, C, D),
            data = mortality_daily,
            start = list(A = 1000, B = 2 * pi / 365, C = 0, D = mean(mortality_daily$total_deaths)))

# Résumé du modèle de régression non-linéaire
summary(fit)

# Calculer les valeurs prévues et les résidus
mortality_daily <- mortality_daily %>%
  mutate(fitted_deaths = predict(fit, newdata = mortality_daily),
         residuals = total_deaths - fitted_deaths)

# Calculer la corrélation entre les résidus et la température
correlation <- cor(mortality_daily$residuals, mortality_daily$heatwave_day)

# boxplots
ggplot(mortality_daily, aes(x = heatwave_day, y = residuals)) +
  geom_boxplot() +
  labs(
    title = "Comparaison des résidus de mortalité entre les jours de vague de chaleur et les jours sans
vague de chaleur (1992-2020)",
    x = "Jour de vague de chaleur",
    y = "Résidus de mortalité"
  ) +

```

```

theme_minimal(base_size = 15)
theme(
  plot.title = element_text(face = "bold", size = 20, hjust = 0.5),
  axis.title = element_text(face = "bold", size = 15),
  axis.text = element_text(size = 14),
  legend.title = element_text(face = "bold", size = 15),
  legend.text = element_text(size = 12)
)

# Comparer résidus jours de vague de chaleur et jours sans vague de chaleur
heatwave_residuals <- mortality_daily$residuals[mortality_daily$heatwave_day == TRUE]
non_heatwave_residuals <- mortality_daily$residuals[mortality_daily$heatwave_day == FALSE]

# Test de Wilcoxon
wilcox_test <- wilcox.test(heatwave_residuals, non_heatwave_residuals)

# Résultats du test
wilcox_test

```

Création Modèle DLNM

```

# Charger les bibliothèques nécessaires
install.packages(c("dlnm", "ggplot2", "lubridate", "dplyr", "splines"))
library(dlnm)
library(ggplot2)
library(lubridate)
library(dplyr)
library(splines)

# Convertir les variables de date au format Date
mortality$Date <- as.Date(mortality$Date)
meteo_data$timestamp <- as.Date(meteo_data$timestamp)

# Filtrer les données pour la période spécifiée
mortality_filtered <- mortality %>%
  filter(Date >= as.Date("2016-01-01") & Date < as.Date("2020-01-01"))

meteo_data_filtered <- meteo_data %>%
  filter(timestamp >= as.Date("2016-01-01") & timestamp < as.Date("2020-01-01"))

# Fusionner les données de mortalité et météorologiques sur la colonne de date commune
mortality_meteo <- merge(mortality_filtered, meteo_data_filtered, by.x = "Date", by.y =
"timestamp", all.x = TRUE)
mortality_meteo_filtered <- mortality_meteo [, -c(3, 6, 9, 10, 11, 12, 14, 16)]
mortality_meteo_filtered$Date <- as.Date(mortality_meteo_filtered$Date)

# Créer une variable de température maximale
temperature_max <- mortality_meteo_filtered$air_temperature_max

# Agréger les données par jour

```

```

daily_data <- mortality_meteo_filtered %>%
  group_by(Date) %>%
  summarise(
    Nombre_de_décès = sum(Nombre_de_décès),
    air_pressure = mean(air_pressure, na.rm = TRUE),
    air_temperature_max = mean(air_temperature_max, na.rm = TRUE),
    relative_humidity = mean(relative_humidity, na.rm = TRUE),
    precipitation = sum(precipitation, na.rm = TRUE)
  )

# Vérifier le jeu de données agrégé
head(daily_data)

# Charger les bibliothèques nécessaires
install.packages(dlnm)
install.packages(splines)
install.packages(MASS)
library(dlnm)
library(splines)
library(MASS)

# Ajouter une colonne pour le jour de l'année
mortality_meteo_filtered$Day_of_Year <- yday(mortality_meteo_filtered$Date)

# Définir la fonction sinusoïdale
sinusoidal_function <- function(x, A, B, C, D) {
  A * sin(B * (x - C)) + D
}

# Ajuster la fonction sinusoïdale aux données de mortalité journalière
fit <- nlsLM(Nombre_de_décès ~ sinusoidal_function(Day_of_Year, A, B, C, D),
  data = mortality_meteo_filtered,
  start = list(A = 1000, B = 2 * pi / 365, C = 0, D =
  mean(mortality_meteo_filtered$Nombre_de_décès)))

# Obtenir les valeurs ajustées (effet saisonnier)
mortality_meteo_filtered$seasonal_effect <- predict(fit)

# Calculer les résidus en soustrayant l'effet saisonnier et ajustement des valeurs négatives
mortality_meteo_filtered$Nombre_de_décès_adj <- mortality_meteo_filtered$Nombre_de_décès -
mortality_meteo_filtered$seasonal_effect
const <- abs(min(mortality_meteo_filtered$Nombre_de_décès_adj)) + 1
mortality_meteo_filtered$Nombre_de_décès_adj <-
mortality_meteo_filtered$Nombre_de_décès_adj + const

# Créer la base de retard distribuée pour la température maximale
cb_temp <- crossbasis(mortality_meteo_filtered$air_temperature_max, lag = 10, argvar = list(fun =
"ns", df = 3), arglag = list(fun = "ns", df = 3))

# Ajuster le modèle GLM avec les covariables pertinentes
model <- glm(Nombre_de_décès_adj ~ cb_temp + Localisation + Sexe + Classe_d_age_combinee,

```

```
family = quasipoisson(), data = mortality_meteo_filtered)

# Résumé du modèle
summary(model)

predicted_values <- predict(model, type = "response")
# Calculer les résidus (différences entre valeurs observées et prédites)
residuals <- mortality_meteo_filtered$Nombre_de_décès - predicted_values

# Calculer la RMSE
rmse_value <- sqrt(mean(residuals^2))
print(paste("RMSE:", rmse_value))

# Effet de la température sur la mortalité ajustée
temp_seq <- seq(min(mortality_meteo_filtered$temp), max(mortality_meteo_filtered$temp),
length.out=100)
predicted_mortality <- predict(model, newdata=data.frame(temp=temp_seq, other_variables))
plot(temp_seq, predicted_mortality, type="l", col="blue",
main="Effet de la température sur la mortalité ajustée",
xlab="Température", ylab="Mortalité ajustée")
```