

Exploitation des données météorologiques sur la ville de Rennes

Anna-Lyn Daumard – Léna Guiziou – Emma Pivaut

Les données Page 1

Résultats Page 2 à Page 3

Conclusion Page 4

CONTEXTE

Ces dernières années le réchauffement climatique est une grande préoccupation pour la population et les scientifiques. Avec des données météorologiques disponibles en fonction du temps nous pouvons analyser l'impact du changement climatique sur les températures des villes en France. Cette synthèse nous renseigne sur l'étude des données météorologiques concernant la ville de Rennes de 2008 à 2018. Il s'agit d'analyser si les températures de Rennes sont impactées par le réchauffement climatique.

DESCRIPTION DES DONNÉES

Les données sont issues de Météo France. Ici on s'intéresse à des données issues de capteurs en surface, localisés dans de nombreuses stations en France. La station observée dans notre cas est Saint-Jacques-de-la-Lande, station la plus proche de Rennes. Le fichier est au format CSV, il contient les données mensuelles des températures maximums obtenues dans le mois. Il y a deux variables, la première est la date avec le jour, le mois et l'année. La

MODIFICATION DES DONNÉES

Nous avons importé notre fichier sur R studio pour identifier les caractéristiques de la série temporelle, définir les différentes composantes de la série et réaliser des prévisions. Les températures du fichier étaient notées en Kelvin nous avons donc fait une conversion en Celsius. Pour cela nous avons soustrait 273,15 à la température en Kelvin.

Pour une meilleure étude des données mensuelles à partir de la variable « date » nous avons extrait le mois écrit en abrégé ainsi que l'année. Pour l'exploitation des données nous utilisons ces deux nouvelles variables.

CHOIX DES MODÈLES

Le modèle de la série est un modèle additif composé d'une tendance d'une saisonnalité et de résidus. Dans un premier temps nous visualisons les données pour trouver le modèle et les composantes. Ensuite nous faisons les estimations avec la modélisation paramétrique et enfin les prévisions avec Holt-Winter. Ci-dessous un extrait de nos données mis dans un tableau grâce à la fonction « pivot_wider ».

TABEAU 1 : EXTRAIT DU JEU DE DONNÉES

Annee <chr>	janv. <dbl>	févr. <dbl>	mars <dbl>	avr. <dbl>
2008	13.75	13.50	14.85	20.40
2009	12.10	11.45	17.10	19.40
2010	12.00	12.35	15.85	22.70
2011	13.80	14.45	17.75	23.65
2012	13.60	12.05	20.50	14.30
2013	14.15	11.95	14.65	22.55
2014	13.35	13.00	17.70	18.85
2015	13.85	11.50	14.55	24.50
2016	13.60	13.55	14.10	17.35
2017	12.80	13.65	20.55	23.80
2018	14.30	12.15	15.55	25.65

Les températures maximales à Rennes connaissent une **saisonnalité** assez régulière durant ces 11 ans (*cf Figure 1*). On observera des pics de températures élevées lors des périodes estivales de chaque année, et des pics de températures basses lors des périodes hivernales.

La visualisation de données nous indique également que le pic le plus bas a eu lieu en 2009, et le pic le plus élevé a eu lieu en 2017.

Les variations jaunes nous indiquent les **moyennes mobiles**. Elles permettent de lisser une série de valeurs exprimées en fonction de la période concernée. Ces moyennes s'appuient sur les températures qui connaissent une très légère augmentation, signifiant que les températures deviennent un peu plus chaudes pendant ces 11 ans.

La **saisonnalité** de ces températures peut se visualiser plus facilement dans un autocorrélogramme (*cf Figure 2*). Celui-ci nous informe de plusieurs informations :

- Il y a une forte et positive autocorrélation aux dates décalés.
- Il y a une **dépendance temporelle** périodique.
- Une **tendance** est percevable en constatant que les cycles deviennent légèrement plus petits.

Nous pouvons également représenter les données en comparant les températures maximales pour chaque année, par mois (*cf Figure 3*). Celles-ci restent pratiquement identiques toute l'année pour toutes années confondues.

Cependant, certaines années se distinguent des autres. L'année 2012 connaît une température très faible en avril par rapport aux autres années. L'année 2011 se distingue également en octobre, en subissant une température plus élevée que les autres années.

FIGURE 2 : TEMPÉRATURES MAXIMALES DE RENNES

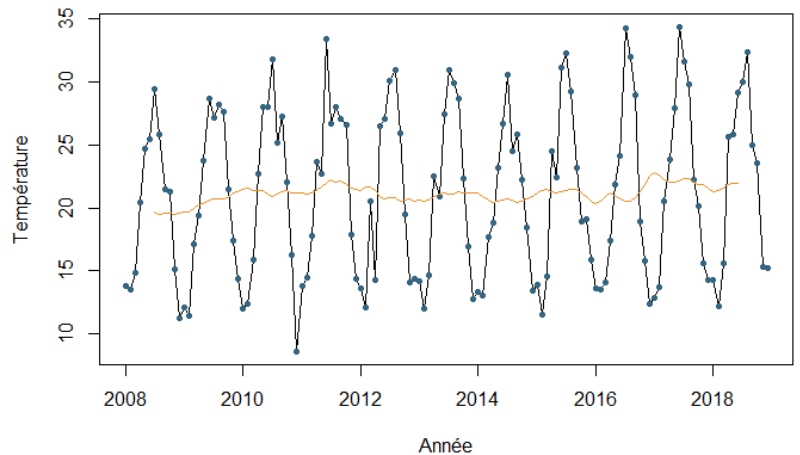


FIGURE 1 : AUTOCORRÉLOGRAMME DES TEMPÉRATURES MAXIMALES

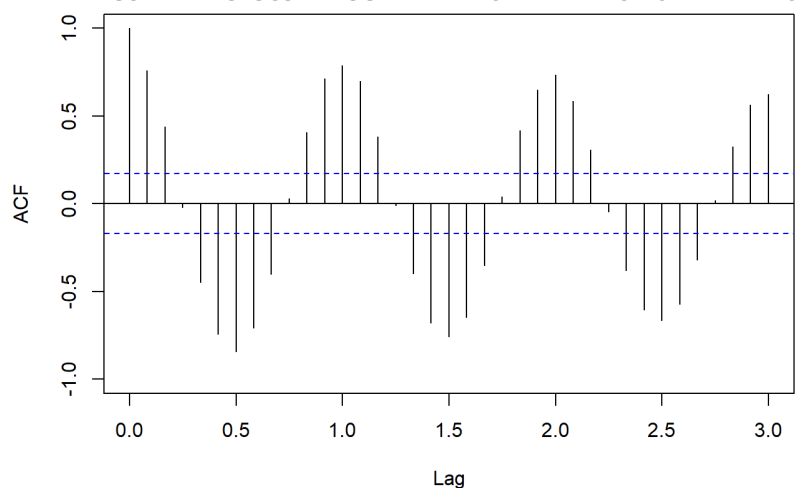


FIGURE 3 : TEMPÉRATURES MAXIMALES DE RENNES PAR MOIS

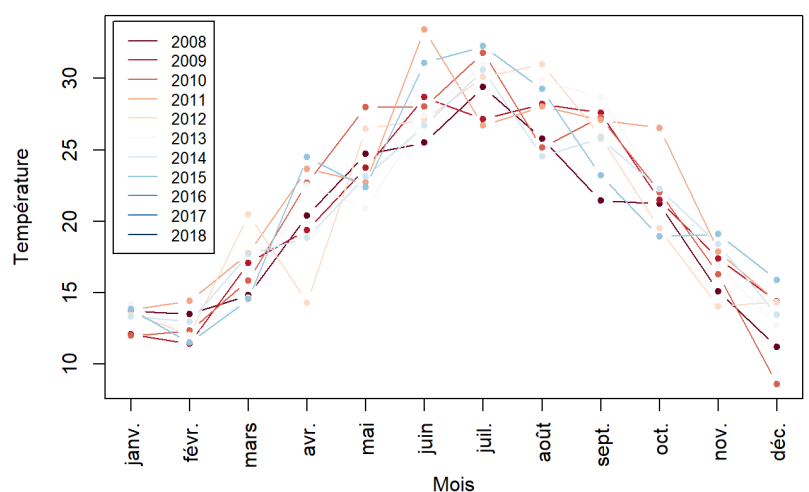
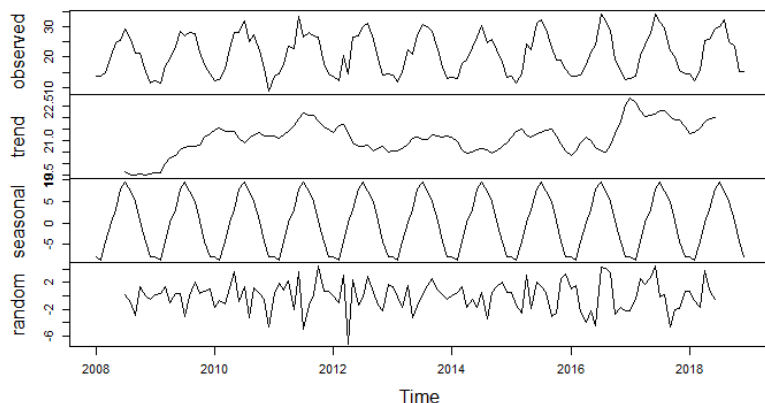


FIGURE 4 : DÉCOMPOSITION DE LA SÉRIE

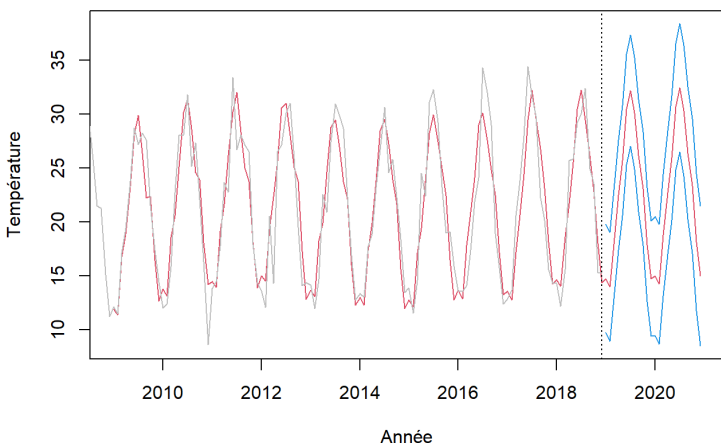


MODÉLISATION

Sur ce graphique ci-contre, nous observons, en bleue, une courbe représentant les températures maximales de Rennes avec le modèle des **moindres carrés**. Elle se base sur la même tendance (très peu visible) et saisonnalité que les données récoltées. Cette courbe ne prend donc pas en compte les bruits.

INCERTITUDE DE LA PRÉVISION

FIGURE 6 : PRÉVISION DES TEMPÉRATURES SUR 2 ANS



Afin de mieux appréhender les températures des années suivantes, il est possible de réaliser des projections (cf Figure 6). Ces prévisions sont accompagnées d'un intervalle de confiance car celles-ci sont incertaines. Dans notre cas, on constate que notre prévision est dans la continuité de notre jeu de données de base. En réalisant la prévision avec Holt-Winter, on constate que la modélisation est de mauvaise qualité que la modélisation paramétrique.

Nous avons également décidé de comparer les prévisions avec les vraies données de 2019 à 2020, visibles en bleues (cf Figure 7). On constate que pour ces deux années, les températures sont plus élevées lors des périodes estivales comparées à la prévision réalisée. Cela peut être un signe que nous arrivons à un stade où nous ne pouvons plus prédire de façon correcte les températures futures à cause du réchauffement climatique.

DÉCOMPOSITION DE LA SÉRIE

Dans l'optique d'identifier de façon plus précise les différentes composantes de notre série, il est possible de réaliser un graphique décomposant toutes les composantes de celle-ci (cf Figure 4). Dans notre cas, il n'était pas facile d'apercevoir la tendance lors de la visualisation de données. À présent, la **tendance** (*trend*) est plus facilement repérable. Le **bruit** (*random*) et la **saisonnalité** (*seasonnal*) quant à eux sont très marqués.

FIGURE 5 : AJUSTEMENT DES TEMPÉRATURES MAXIMALES DE RENNES

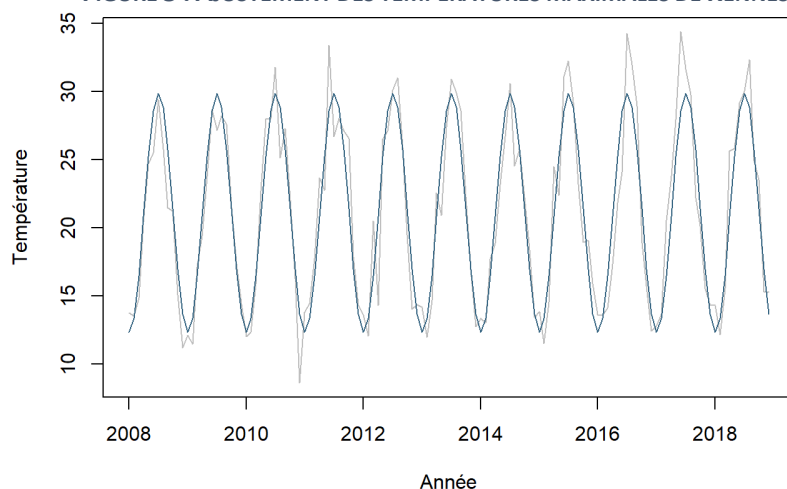
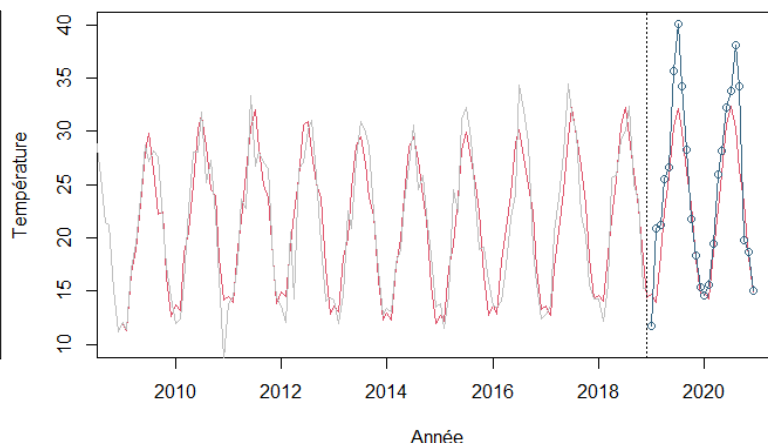


FIGURE 7 : TEMPÉRATURES RÉELLES DE 2019 À 2020



CONCLUSION

Grâce aux analyses ci-dessus, nous pouvons conclure sur la présence du réchauffement climatique à Rennes mais aussi sur l'efficacité de la méthode de prédiction utilisée. Elle peut paraître à première vue efficace, mais lors de notre dernière analyse, les résultats n'ont pas été très concluants, même surprenants.

Pour rappel, nous avons à disposition un jeu de données avec les températures maximales de Rennes récoltées pendant la période 2008-2018. Notre but était de savoir si les températures étaient impactées par le réchauffement climatique, c'est-à-dire si nous observions une hausse des températures. Après un nettoyage des données, et la création de 2 nouvelles variables, nous avons pu créer différents graphiques à l'aide de différentes méthodes. Ces graphiques comme l'autocorrélogramme et la décomposition de la série nous permettent de voir une augmentation des températures avec la tendance et donc que le réchauffement climatique est bien présent.

Après une analyse de la tendance, nous avons voulu chercher si les prédictions avec la méthode de Holt-Winter étaient efficaces. Nous avons donc cherché à comparer notre prédiction avec les données récoltées par le même centre de recherche pour les années 2019 et 2020. Nous avons donc remarqué que nous avons sous-estimé l'influence du réchauffement climatique sur les températures de Rennes. L'augmentation se fait beaucoup plus vite que les années précédentes et nous devrions refaire une prédiction avec les années les plus récentes pour définir un modèle plus adéquat.

ABSTRACT

This case study focuses on the evolution of maximum temperatures in Rennes between 2008 and 2018, and finally calculates and observes possible predictions for the years 2019 and 2020. These analyses, including the creation of the graphs, are done with the help of RStudio software and statistical considerations. To explain these temperatures changes, we had to break down each component of our series. In order to do this, we need to clean and modify the data to start analysing the data. It was necessary to modify the only two variables available in order to make them usable. After this action, the different graphs made allowed us to better visualise the data. Some of these graphs explain the positive evolution of the temperatures, and others focus on their predictions. In other graphs we see regular variations with peaks in winter and summer that are repeated over the years. As far as the predictions made are concerned, these don't seem to be very reliable when compared with the actual data collected. A link with global warming may explain this difference, which is even more obvious in summer.

DÉFINITIONS ET CALCULS

Tendance (mt) : Évolution à long terme de la série.

Saisonnalité (St) : Phénomène qui se répète de façon périodique, à intervalle régulier.

Bruit (Et) : Fluctuation irrégulière imprévisible, non prise en compte par la tendance et/ou la saisonnalité.

Moindres carrés : Notion mathématique permettant d'apporter à un nombre d'éléments susceptibles de comporter des erreurs un ajustement afin d'obtenir des données proches de la réalité.

Moyenne mobile : Permet de lisser une série de valeurs exprimées en fonction du temps (série chronologique).

Ordre impair : $Y_t = \frac{1}{2k+1} (X_{t-k} + \dots + X_{t+k})$

Ordre pair : $Y_t = \frac{1}{2k} \left(\frac{X_{t-k}}{2} + X_t - k + 1 + \dots + X_{t+k-1} + \frac{X_{t+k}}{2} \right)$

Dépendance temporelle : Caractéristique de nombreuses données de séries chronologiques qui indiquent que le passé affecte le futur.