

LE PALUDISME ET CHANGEMENT CLIMATIQUE : un impact majeur ?

INTRODUCTION

Au cours de la dernière décennie, l'intérêt pour les effets du changement climatique sur la santé humaine s'est accru [\[1\]](#).

Le sixième rapport du GIEC désigne l'influence de l'Homme sur le réchauffement climatique de la planète comme « sans équivoque », responsable d'environ +1,1°C de réchauffement depuis la période 1850-1900. C'est la première fois dans un rapport du GIEC que le réchauffement climatique est attribué sans équivoque aux activités humaines. Le rapport fait état de 5 scénarios d'évolutions socio-économiques différents (SSP1, SSP2, SSP3, SSP4, SSP5). A l'exception du scénario le plus optimiste (SSP1), le seuil d'1,5°C de réchauffement pourrait être atteint dès 2030, soit 10 ans plus tôt que la précédente estimation du GIEC. Le rapport témoigne également d'une hausse des précipitations moyennes mondiales depuis 1950, avec une augmentation plus rapide depuis les années 1980. Des inondations sans précédent sont survenues en Europe, en Chine, ou aux Philippines en 2021. Ces pluies ont entraîné des dégâts considérables et des pertes de vie humaine. Le recul mondial des glaciers, le réchauffement de la couche supérieure des océans, la hausse du niveau moyen des océans sont d'autres paramètres du changement climatique mentionnés dans le rapport du GIEC [\[2\]](#). Le taux de changement climatique est plus rapide aujourd'hui qu'à n'importe quelle autre période des mille dernières années.

Une grande attention a été accordée à la relation entre le changement climatique et le risque croissant de maladies infectieuses. Le climat affecte principalement l'éventail de maladies infectieuses, tandis que les conditions météorologiques affectent le moment et l'intensité des épidémies. Les scénarios de changement climatique incluent un changement dans la répartition des maladies infectieuses avec le réchauffement climatique et des changements dans les épidémies associées aux conditions météorologiques extrêmes.

Des modèles statistiques sont utilisés pour estimer la charge mondiale de certaines maladies infectieuses résultant du changement climatique : par exemple, si la température mondiale augmente de 2 à 3°C (comme prévu), la population à risque de paludisme pourrait augmenter de 3 à 5% [\[1\]](#).

Les maladies infectieuses à transmissions vectorielles sont les plus concernées par le changement climatique. Les propriétés importantes impliquées dans la transmission des maladies vectorielles comprennent : le taux de survie et de reproduction du vecteur, la période de l'année et le niveau d'activité du vecteur (en particulier le taux de morsure) et le taux de développement et de reproduction de l'agent pathogène dans le vecteur. Les vecteurs, pathogènes et hôtes se reproduisent dans certaines conditions optimales. Les conditions climatiques et les changements dans ces conditions peuvent modifier considérablement ces propriétés de transmission de la maladie. Les facteurs climatiques les plus influents pour les maladies transmises par les vecteurs comprennent la température et les précipitations (mais aussi l'élévation du niveau de la mer, le vent et la durée de la lumière du jour qui sont d'autres facteurs importants) [\[3\]](#).

Les agents pathogènes transmis par des vecteurs sont particulièrement sensibles au changement climatique car ils passent une bonne partie de leur cycle de vie dans un hôte invertébré ectotherme dont la température est similaire à celle de l'environnement. La température affecte chaque étape du cycle de vie des vecteurs, c'est une variable qui affecte à la fois le développement de la population de vecteurs et du parasite au sein du vecteur (là où la disponibilité en eau et l'humidité n'affecte que le vecteur) [\[1\]](#). Lorsqu'un vecteur vit dans un environnement où la température moyenne approche de la limite de tolérance physiologique pour l'agent pathogène, une légère augmentation de température peut être mortelle pour l'agent pathogène. Au contraire, lorsqu'un vecteur vit dans un environnement de basse température moyenne, une petite augmentation de la température peut accroître le développement, temps d'incubation et de réplication de l'agent pathogène [\[3\]](#).

La température peut également modifier la croissance des vecteurs de maladies en modifiant leurs taux de morsure, ainsi qu'affecter la dynamique des populations vectorielles et modifier le taux auxquels ils entrent en contact avec les humains. La comparaison de la maturation des moustiques dans les zones forestières et zones déboisées, dans lesquelles il y avait une différence de quelques degrés, a permis d'estimer le pourcentage d'insectes qui sont passés de la forme larvaire à la forme adulte : de 65% à 82% ainsi que la réduction de la période nécessaire au développement, qui passe de 9 à 8 jours, dans les zones plus chaudes [\[1\]](#). Enfin, un changement de température peut affecter la durée de transmission des maladies.

Les vecteurs de maladies peuvent également s'adapter aux changements de température en changeant leur distribution géographique. L'émergence du paludisme dans les climats plus froids des hautes terres africaines pourrait être le résultat d'un changement d'habitat du moustique afin de s'adapter à la hausse de l'air ambiant. Une autre possibilité est que les vecteurs s'adaptent au réchauffement climatique en mettant en place une réponse adaptative (sélection naturelle de mutations favorables).

La variabilité des précipitations peut également avoir des conséquences directes sur les épidémies de maladies infectieuses. L'augmentation des précipitations peut accroître la présence de vecteurs de maladies en agrandissant l'habitat larvaire existant et en créant de nouvelles aires de reproduction. En outre, l'augmentation des précipitations peut soutenir une croissance des approvisionnements alimentaires, qui à leur tour, soutiennent une plus grande population de réservoirs vertébrés. Des pluies abondantes non saisonnières peuvent provoquer des inondations et réduire les populations de vecteurs en éliminant les habitats larvaires et créant des environnements non favorables pour les réservoirs vertébrés.

Parallèlement, ces inondations peuvent forcer les vecteurs à se réfugier dans les habitations et à augmenter ainsi les probabilités de contact vectoriel-humain. Dans les régions tropicales humides, une sécheresse non saisonnière peut causer le ralentissement du drainage des rivières et créer ainsi des mares plus stagnantes qui sont des habitats idéaux pour la reproduction vectorielle [3].

Notre étude se concentre sur une maladie à transmission vectorielle : le paludisme (ou malaria). Le paludisme est l'une des plus fréquentes maladies infectieuses. Elle est transmise par les piqûres de moustiques *anophèles femelles* et causée par des protozoaires parasites du genre *Plasmodium* (*P.falciparum* provoquant le plus de décès et *P.vivax* étant l'espèce dominante de la plupart des pays en dehors de l'Afrique subsaharienne).

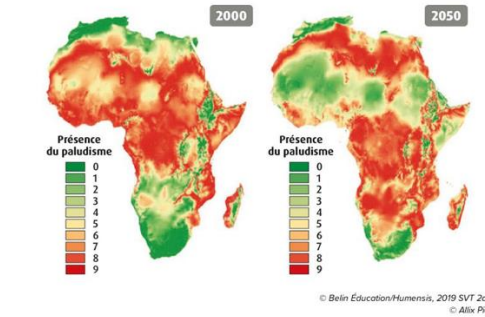


Les efforts de la communauté internationale au cours des 20 dernières années ont permis de réduire les niveaux de risque du paludisme de 40% entre 2000 et 2015, et de vastes régions d'Afrique sont désormais en mesure d'envisager des stratégies d'élimination [4]. Malgré ces progrès, près de la moitié de la population mondiale reste exposée au risque de paludisme [6]. Rien qu'en 2020, il y a eu environ 241 millions de cas de paludisme recensés dans le monde et 627 000 décès [5]. Environ 90% de tous les décès dus au paludisme surviennent en Afrique subsaharienne et la plupart d'entre eux concernent des enfants de moins de 5 ans. En effet, le vecteur le plus efficace (*Anopheles gambiae*) est le plus répandu en Afrique et le plus difficile à contrôler [4]. L'Afrique est particulièrement vulnérable au paludisme et au changement climatique. 7 pays sur 10 dans le monde considérés comme les plus menacés par le changement climatique et 10 pays parmi les plus touchés par le paludisme

dans le monde sont situés sur le continent africain [7].

Le paludisme est une maladie fébrile aiguë dont les premiers symptômes (fièvre, maux de tête, frissons) apparaissent généralement 10 à 15 jours après la piqûre infectieuse et peuvent être légers et difficiles à reconnaître. En l'absence d'un traitement, le paludisme à *P.falciparum* peut évoluer vers une affection grave voire mortelle dans les 24 heures. Les populations les plus à risque de contracter la maladie et de développer une forme sévère sont les nourrissons, les enfants de moins de 5 ans, les femmes enceintes, les personnes vivant avec le VIH/sida ainsi que les personnes à faible immunité qui se rendent dans des zones à transmission intense, comme les travailleurs migrants, les populations mobiles et les voyageurs [6]. Le paludisme est une maladie particulièrement sensible au climat, fortement influencée par les changements de température et de précipitations (et par la fréquence et gravité des phénomènes météorologiques extrêmes) [7]. El Nino est un événement climatique qui trouve son origine dans l'océan Pacifique mais qui a de vastes conséquences sur les conditions météorologiques dans le monde entier. A l'échelle mondiale, il est lié à un impact accru des catastrophes naturelles et est particulièrement associée à la transmission vectorielle, en particulier le paludisme [1]. Les parasites du paludisme ont besoin de moins de temps que l'on pensait pour se développer à des températures plus basses. Cela signifie qu'un réchauffement même modéré peut suffire à accroître le risque de paludisme. A long terme, la hausse des températures et des précipitations pourrait propager la maladie dans des zones auparavant exemptes de paludisme. L'organisation mondiale de la santé estime que le changement climatique entraînera 60 000 décès supplémentaires liés au paludisme entre 2030 et 2050, soit une augmentation de près de 15% du nombre total annuel de décès dus à cette maladie évitable. D'ici 2050, à lui seul le changement climatique pourrait exposer certaines zones d'Amérique du Sud, d'Afrique subsaharienne et de Chine à une probabilité de transmission du paludisme supérieure à 50%. Des facteurs économiques, sociaux et politiques peuvent également expliquer la résurgence récente du paludisme plutôt que le changement climatique [7]. Des modèles ont été élaborés pour prédire dans les prochaines années la distribution du paludisme. Ils prévoient une

extension des zones d'endémie palustre et un déplacement des zones touchées [1].



[8] Il est aujourd'hui nécessaire de renforcer le dialogue sur les moyens de faire face aux défis interdépendants du changement climatique et de l'élimination du paludisme afin d'identifier des stratégies et des possibilités d'atténuation de l'impact, en particulier sur les populations les plus vulnérables [7].

« Le rôle d'agences comme l'OMS doit être fortement renforcé avec des moyens accrus en terme de détection et de diagnostic » - N.Bayan

Au cours de cette étude, nous proposons de nous placer dans le rôle d'agents de l'OMS responsables de l'étude de l'impact des changements climatiques sur l'émergence du paludisme. Nous tenterons, par la création d'une base de données collaborative regroupant des données du monde entier et par l'utilisation de statistiques, de vous convaincre de l'importance du réchauffement climatique et de l'augmentation des précipitations sur l'émergence du paludisme. Finalement, à partir des conclusions émises en première partie et d'algorithmes/machine learning, nous tenterons de proposer des solutions à cette émergence en prévoyant et anticipant de manière plus optimisée les épidémies à venir. Ainsi, des prises de décision rapides et une mise en garde de la population pourront être effectuées et des protocoles sanitaires et médicaux pourront être mis en place en conséquence.

Présentation de nos auteurs

EGLANTINE FAGET



[FR] Étudiante en L3 Biologie-Santé, je suis passionnée de génétique, surf et théâtre. Je souhaite m'orienter en Cancérologie, afin d'accomplir mon rêve de devenir enseignant-chercheur en oncogénétique. Je compte également obtenir un prix Nobel au cours de ma carrière, je n'oublierais pas de faire mention des Datasciences dans mon discours de remerciements. J'ai réalisé la partie Introductive, Base de données et Communication de ce projet. La Science des données est pour moi un outil dont la compréhension est indispensable afin de se créer un esprit critique dans ce monde où les données sont si facilement manipulables.

[ENG] Student in L3 Health-Biology, I'm a fervor of genetic, surf and theater. I would like to pursue my studies in Cancerology, to accomplish my dream of becoming a teacher-researcher in oncogenetic. I also intend to obtain a Nobel price throughout my career, I will not forget to mention Datasciences in my acceptance speech. I've realized the Introductive, Database and Communication parts of this project. Data Science, in my opinion, is a tool that need to be understand to create a self-critical eye in a world where data are so easily manipulable.

[DE] Als Studentin der L3 Biologie-Gesundheit bin ich leidenschaftlich an Genetik, Surfen und Theater interessiert. Ich möchte mich in der Krebsforschung orientieren, um meinen Traum zu verwirklichen, Onkogenitallehrer zu werden. Ich beabsichtige auch, im Laufe meiner Laufbahn einen Nobelpreis zu erhalten, und ich werde nicht vergessen, in meiner Dankesrede die Datasciences zu erwähnen.

Ich habe den einleitenden Teil, die Datenbank und die Kommunikation dieses Projekts erstellt. Die Datenwissenschaft ist für mich ein Werkzeug, dessen Verständnis für die Schaffung eines kritischen Geistes in dieser Welt, in der Daten so leicht manipulierbar sind, unerlässlich ist.

CREATION D’UNE BASE DE DONNÉES

Afin de pouvoir étudier l’impact du changement climatique sur l’émergence du paludisme, une base de données regroupant les mesures de 5 scientifiques à travers les régions les plus à risque du monde a été réalisée. Il s’agit d’une base de données fictive, basée sur des observations réelles et adaptée à notre étude. Les données ont été extraites à partir de deux sites regroupant des données précises sur les variables à risque, selon les pays/villes et les années/mois [9] [10]. Les données ont été triées selon leur pertinence et entrées dans une base de données Access. Une première table « Observateur » a été créée, regroupant des informations sur le nom, statut et région d’action de l’observateur.

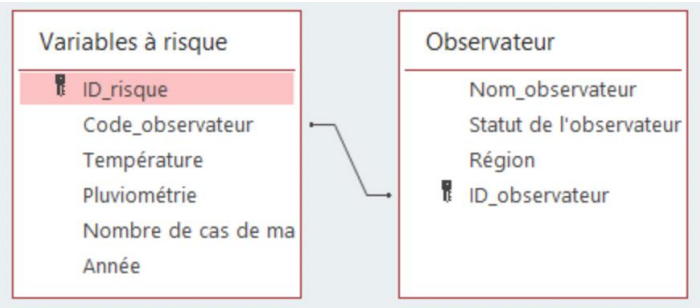
Nom_observateur	Texte court	Nom de l'observateur des variables
Statut de l'observateur	Texte court	Statut professionnel de l'observateur
Région	Texte court	Région où exerce l'observateur, où sont mesurées les variables
ID_observateur	Numérique	Code définissant de manière unique un observateur

Cette table est basée sur des données 100% factices.

Une deuxième table « Variables à risque » a été créée, regroupant des informations sur la température, la pluviométrie et le nombre de cas de malaria depuis 1980 jusqu’à 2020 pour chaque région observée. Les valeurs des variables à risque ont été issues d’une moyenne des mesures prises chaque semaine pendant deux ans dans une région donnée, les mesures ayant été regroupées pour un souci de simplicité.

Nom du champ	Type de données	Description (facultative)
ID_risque	Numérique	Code définissant de manière unique un risque (température, pluviométrie, nombre de cas de malaria...)
Code_observateur	Numérique	Code de liaison à un observateur et sa région d'observation
Température	Texte court	Mesure des températures globales pour chaque région (°C)
Pluviométrie	Numérique	Mesure de la pluviométrie pour chaque région (mm)
Nombre de cas de malaria	Numérique	Mesure du nombre de cas de malaria pour chaque région
Année	Numérique	Année à laquelle a été prise chacune des mesures (résulte de la moyenne des valeurs prises chaque semaine pendant deux années)

Une relation a été établie entre les deux tables afin de lier chaque mesure à une région donnée.



Le but de cette base de données est d’aider les scientifiques à combattre la maladie dans leur région en réunissant des informations sur les possibles facteurs d’influence. En comparant les différents facteurs à risque selon la recrudescence des cas de malaria en Inde, on peut facilement observer que plus la température et la pluviométrie augmentent, plus les cas de malaria augmentent.

	Température	Pluviométrie	Cas de malaria	Années
Inde	26	504	105678	1980
Inde	26,2	720	112309	1982
Inde	26,3	784	187629	1984
Inde	26,7	882	209867	1986
Inde	26,8	915	246299	1988
Inde	26,5	1063	254303	1990
Inde	26,7	1112	278654	1992
Inde	26,1	1101	307951	1994
Inde	26,7	1194	338494	1996
Inde	26,7	1213	429928	1998
Inde	27,2	1228	844558	2000
Inde	27,1	1368	1087285	2002
Inde	27,3	1305	1102205	2004
Inde	27,3	1333	1067824	2006
Inde	27,1	1465	1310656	2008
Inde	27,6	1494	1599986	2010
Inde	27,9	1460	1526210	2012
Inde	27,6	1554	1785129	2014
Inde	27,9	1579	1816569	2016
Inde	28,1	1748	1915363	2018
Inde	28,2	1850	2085484	2020

Ainsi, une hypothèse selon laquelle la hausse des températures et de la pluviométrie influencerait la recrudescence des cas de malaria en Inde pourrait être émise et des stratégies de prévention pourraient ainsi être établies. Une analyse statistique plus précise devra permettre de valider cette hypothèse. En croisant nos informations pour chaque région dans la base de données, l’hypothèse pourrait être élargie au niveau mondial et permettre ainsi d’établir plus facilement des stratégies pour prévenir la maladie (grâce à une coopération scientifique internationale). Ainsi, par comparaison entre les différentes régions du monde à risque, on observe que plus la température et la pluviométrie augmentent et plus les cas de malaria augmentent. Seul Madagascar déroge à la règle concernant la pluviométrie et montre ainsi que cette variable est moins influente que la température sur la hausse des cas.

Région	Température	Pluviométrie	Nombre de c	Année
Madagascar	26,6	4055	4789	1980
Madagascar	26,8	4496	4867	1982
Madagascar	26,9	4082	4880	1984
Madagascar	27,1	3741	4990	1986
Madagascar	27,2	3961	5020	1988
Madagascar	27	3428	5236	1990
Madagascar	27,1	3478	5487	1992
Madagascar	27,4	3465	5593	1994
Madagascar	27,3	3165	5876	1996
Madagascar	27,1	3182	6349	1998
Madagascar	27,5	3257	6646	2000
Madagascar	27,5	2960	6909	2002
Madagascar	27,3	2730	7638	2004
Madagascar	27,8	2802	8538	2006
Madagascar	27,9	2882	48497	2008
Madagascar	27,8	2702	202450	2010
Madagascar	27,7	2672	402900	2012
Madagascar	28,1	2536	468031	2014
Madagascar	27,9	2680	655480	2016
Madagascar	28,2	2235	921097	2018
Madagascar	28,4	2084	1876791	2020

Le lien étant ainsi établi entre hausse des températures et de la pluviométrie et hausse des cas de malaria, la base de données peut permettre de mettre l'accès sur des régions de cas majeur avec des cas de malaria extrêmement importants et ainsi de les prioriser dans le cadre de stratégies de prévention.

Région	Nombre de c	Année
Afrique centrale	4502086	2020
Afrique du nord	55	2020
Inde	2085484	2020
Amérique du su	404924	2020
Madagascar	1876791	2020

On identifie ainsi l’Afrique centrale comme étant une région de haute priorité.

Présentation de nos auteurs

GWENAELLE LAURENT

[FR] Etudiante en L3 Biologie-Santé, je suis passionnée de scoutisme, de danse classique et de pâtisserie. Je souhaite m’orienter dans la recherche pour les maladies rares car il s’agit d’un sujet mondialement connu et pourtant peu étudié au cas par cas. J’ai participé aux parties Introductive, Machine learning et Discussion de ce projet. De nos jours, la Science des données est utilisée par tous sous une forme ou une autre afin de donner du sens aux données et permettre de prendre des décisions fondées. Elle révolutionne donc notre quotidien sans que l’on s’en rende compte.

[ENG] Student in L3 Health-Biology, I am passionate about Scouting, classical dance and pastry. I want to focus on research for rare diseases because this is a topic that is known worldwide and yet little studied on a case-by-case basis. I participated in the Introductive, Machine learning and Discussion parts of this project. Today, Data Science is used by all in one form or another to make sense of data and to make sound decisions. It therefore revolutionizes our daily lives without us realizing it.

[ES] Estudiante de L3 Biología-Médica, me apasionan los scouts, la danza clásica y los pasteles. Me gustaría centrarme en la investigación sobre las enfermedades raras, porque se trata de un tema mundialmente conocido y, sin embargo, poco estudiado caso por caso. Participé en las secciones Introducción, Aprendizaje automático y Discusión de este proyecto. Hoy en día, la ciencia de los datos es utilizada por todos de una forma u otra para dar sentido a los datos y permitir tomar decisiones fundadas. Por lo tanto, revoluciona nuestra vida cotidiana sin que nos demos cuenta.

Afin de mettre en place des stratégies pour prévenir de la recrudescence de la maladie dans certaines régions du monde, du machine learning devra être mis en place, basé sur les données et observations faites à partir des régions majoritairement touchées. La base de données devra prendre en compte de nouvelles régions moins à risque mais qui pourraient potentiellement le devenir ainsi que des zones plus précises dans les régions à risque. Ainsi, en alimentant les données chaque semaine, des décisions rapides pourront être prises selon le risque de recrudescence de la maladie afin d’alerter les populations et permettre la mise en place de protocoles sanitaires et médicaux dans différentes régions du monde.

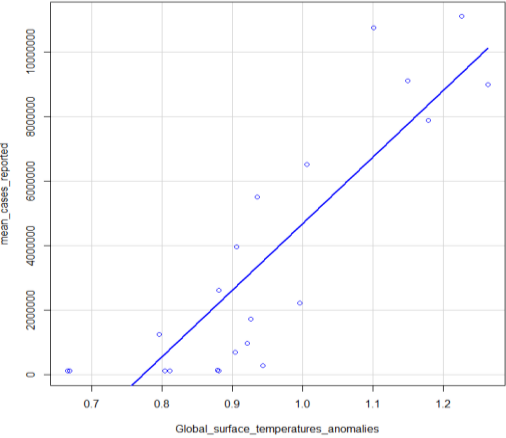
ANALYSE STATISTIQUE DES DONNÉES

Dans le but de mettre en œuvre nos différents projets, il est nécessaire d'alerter les entreprises et les gouvernements des pays du Nord sur la situation. Ainsi, il sera possible d'attirer des financements et de rendre plus efficaces les dispositifs que l'on souhaite mettre en place. La sensibilisation des populations doit passer par une analyse statistique des données et à la création de modèles afin de se préparer aux scénarios climatiques possibles pour les années à venir.

Dans le cas de la Malaria, on sait que les moustiques Anophèles, vecteurs de cette maladie, se reproduisent de manière optimale à haute température, dans des milieux humides. On va donc ici tenter de corrélér les hausses des températures moyennes de la planète avec la hausse du nombre de cas de Malaria.

Pour cela, nous allons regrouper les données du nombre de cas déclarés de Malaria depuis 2000, jusqu'à 2020, dans 3 pays [11] : un pays d'Afrique de l'Ouest (Togo), un pays d'Afrique centrale (République Démocratique du Congo), et un pays d'Afrique de l'Est (Burundi). Ces données sont confrontées aux valeurs de réchauffement de la température moyenne mondiale de l'air à la surface des terres et de l'eau à la surface des océans [12].

Pour avoir un premier aperçu de la relation entre ces deux variables, on trace un nuage de point représentant le nombre moyen de cas de Malaria déclarés dans ces 3 pays en fonction des anomalies de températures observées. De plus, on trace la droite des moindres carrés de ces points afin d'avoir un premier aperçu d'une corrélation possible entre ces deux variables.



Ici, on observe que plus la température moyenne de l'air à la surface de la Terre augmente, plus le nombre de cas déclarés moyen de Malaria augmente. Il est donc possible d'émettre l'hypothèse qu'une corrélation linéaire positive assez forte existe entre ces deux variables.

Pour confirmer cette observation, il est nécessaire d'effectuer un test de corrélation de Pearson, en prenant comme hypothèse nulle qu'il n'existe pas de corrélation entre les deux variables, et en hypothèse alternative, qu'il existe une corrélation linéaire positive.

```
Pearson's product-moment correlation

data: Global_surface_temperatures_anomalies and mean_cases_reported
t = 7.4491, df = 19, p-value = 0.0000002378
alternative hypothesis: true correlation is greater than 0
95 percent confidence interval:
 0.7247843 1.0000000
sample estimates:
cor
0.8630932
```

A l'issue de ce test, on remarque que la p-value donnée par le logiciel est largement inférieure à 5%, ce qui nous permet de rejeter fortement l'hypothèse nulle. De plus, on observe un coefficient de 0.86 environ, donc assez proche de 1. Cela nous permet de confirmer notre hypothèse: il existe bel et bien une corrélation linéaire positive assez forte entre la hausse des températures et le nombre de cas de Malaria.

Cette première conclusion est déjà assez alarmante, en effet cela signifie que les différents scénarios de hausse de température prévus pour les années à venir risquent d'influencer l'incidence du paludisme. La mise en évidence de ce genre de corrélation doit marquer les gouvernements, surtout lorsque l'on met en parallèle les prédictions les plus pessimistes de hausse de température, par le rapport du GIEC par exemple.

Afin d'appuyer encore plus la solidité de cette corrélation, et de se préparer aux différents scénarios possibles (afin d'éventuellement prévoir l'incidence du Paludisme pour les années à venir), il est possible de réaliser une régression linéaire entre les deux variables.

```
Call:
lm(formula = mean_cases_reported ~ Global_surface_temperatures_anomalies,
    data = Malaria)

Residuals:
    Min       1Q   Median       3Q      Max
-3244372 -2004542  -496605  1725173  4002686

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -15998837    2660328  -6.014 0.000008719
Global_surface_temperatures_anomalies  20675685    2775589   7.449 0.000000476

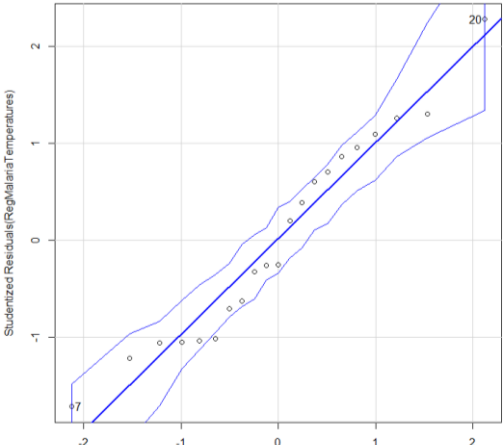
(Intercept) ***
Global_surface_temperatures_anomalies ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2037000 on 19 degrees of freedom
Multiple R-squared:  0.7449, Adjusted R-squared:  0.7315
F-statistic: 55.49 on 1 and 19 DF,  p-value: 0.0000004756
```

On remarque tout d'abord que les p-values associées aux coefficients calculés dans cette régression linéaire sont significativement différentes de 0. Cela signifie que le coefficient directeur ainsi que l'ordonnée à l'origine estimés ici sont pertinents à utiliser dans notre modèle. De plus, la P-value du test F étant largement inférieur à 1%, on peut rejeter l'hypothèse qu'il n'existe aucune relation entre notre variable dépendante (le nombre de cas de Malaria), et notre variable indépendante (la variation de température).

Ainsi, si on veut tracer une droite représentant le nombre de cas de Malaria en fonction des variations de température, l'équation de cette droite sera: $y=20675685x-1599837$. Cela nous permet de prévoir l'impact de la maladie si aucune mesure radicale n'est prise d'ici quelques années. Par exemple, si la hausse de température atteint +2°C (comme annoncé par certains scénarios climatiques du GIEC), le nombre moyen de cas de Malaria dans ces trois pays risque d'atteindre $20\ 675\ 685 \times 2 - 1\ 599\ 837 = 25\ 352\ 533$, ce qui peut être critique pour ces zones. Enfin, on remarque que le coefficient de détermination R^2 est de 0.7449, ce qui signifie dans notre modèle que la variable de température explique à environ 74% la variable du nombre de cas. Cela nous aiguille sur la pertinence de notre modèle, mais aussi sur les limites de celui-ci, et nous amène donc à aussi envisager d'autres variables pour l'améliorer.

Afin de s'assurer de la pertinence de notre modèle, nous allons tracer le graphique quantile-quantile des erreurs(residuals) issues de ce modèle linéaire. Cela va nous permettre de vérifier visuellement que les erreurs de notre régression dépendent bien du hasard, et qu'elles suivent une loi normale.



Ici on remarque que les quantiles observés se trouvent dans l'intervalle de confiance pour la plupart. Il semble donc que les erreurs soient susceptibles de suivre une loi normale. Cependant, notre modèle ne semble pas parfaitement symétrique, et semble comporter des erreurs assez élevées pour les valeurs extrêmes. Afin de confirmer l'hypothèse de normalité des erreurs, on effectue un test de Shapiro-Wilk.

```
Shapiro-Wilk normality test

data: residuals.RegMalariaTemperatures
W = 0.95137, p-value = 0.3615
```

La p-value de ce test étant de 36%, on peut affirmer avec une assez bonne confiance que les erreurs suivent bel et bien une loi normale. Cette hypothèse étant confirmé, cela permet de valider notre modèle linéaire, et ainsi de dire que la corrélation testée est pertinente.

Présentation de nos auteurs

YANNIS ARAB

[FR] Etudiant en L3 Biologie-Santé, je suis un passionné de musiques, d'arts martiaux et de technologies. Je souhaite m'orienter en Microbiologie, car selon moi les prochains enjeux sanitaires auxquels nous ferons face se joueront à l'échelle microscopique. J'ai mené mon équipe à travers ce projet, notamment dans la partie Statistiques. La Science des données représente pour moi la clé de la compréhension du monde qui nous entoure, particulièrement pour lutter contre les problèmes épidémiologiques et écologiques.

ENG] Student in L3 Health-Biology, I'm passionate about music, martial arts, and technology. I wish to pursue my course in Microbiology because I think that the next health issues we will face will need actions on the microscopic scale. I lead my team through this project, especially in the Statistics section. Data Science is for me the key to the understanding of our world, in particular to combat epidemiological and ecological problems.

ES] Estudiante en L3 Biología-Médica, soy un apasionado de la música, artes marciales y tecnología. Me gustaría dedicarme a la Microbiología, porque pienso que los próximos problemas de salud que nos enfrentaremos requerirán acciones en el nivel microscópico. He dirigido mi equipo en este proyecto, especialmente en la sección de las Estadísticas. Para mi, La Ciencia de los datos representa la clave para entender el mundo en que vivimos, principalmente para abordar problemas epidemiológicos y ecológicos.



Toutes les deux minutes, un enfant meurt du paludisme

Nous avons donc prouvé statistiquement de différentes manières qu’il existe bien une corrélation positive entre le nombre de cas de Malaria et les variations de températures à la surface de la Terre. Le modèle mis en place est une preuve que des mesures doivent être prises rapidement afin de faire face aux maladies infectieuses. Si les températures globales continuent d’augmenter à ce rythme, les risques de maladies infectieuses sont susceptibles de devenir critique dans des pays en développement, mais aussi dans les pays développés. En effet, il serait intéressant de réaliser ce genre de modèle pour des pays de l’hémisphère Nord tel que la France, où l’émergence de la Malaria est possible si l’environnement devient propice aux développements des vecteurs de la maladie. Ce genre de modèle est compliqué à mettre en place pour le moment car le nombre de cas déclarés de ce genre de maladies infectieuses est proche de 0 actuellement en France. Des modèles prédictifs seraient possibles à mettre en place avec plus de données, afin d’alerter sur l’émergence de la maladie dans les pays du Nord, en plus du risque de forte hausse de l’incidence dans les pays du Sud.

APPRENTISSAGE AUTOMATIQUE

Pour faire face à l’augmentation de l’incidence de la Malaria suite aux changements climatiques, il est nécessaire de pouvoir prédire grâce à une collecte des informations environnementales lorsque le climat sera propice à une augmentation de la présence des moustiques. Il faut donc détecter et traiter ces informations puis prendre des décisions grâce à ces données.

En extrayant les données de température et de pluviométrie de notre base de données, il est possible de mettre un place un algorithme capable de traiter ces informations et d’aider les scientifiques à prendre des décisions en fonction de chaque zone touchée par la Malaria. En fonction de la température et de la pluviométrie de la zone (mesuré par un scientifique de notre database), l’algorithme d’aide à la décision serait construit de sorte que lorsqu’une ou les deux données dépassent un certain seuil, cela préviendrait les scientifiques en charge et leur conseillerais de mettre la population à l’abri, de distribuer des soins aux personnes en danger, d’appliquer des dispositifs anti-moustiques dans les habitations, sur les populations. Bien sûr, ces anti-moustiques nécessiteraient d’être tout d’abord plus performant contre les principales espèces présentes dans chaque pays et ceux les plus enclins à transmettre la maladie.

En langage Python, nous avons mis en place un exemple d’algorithme qui est capable d’évaluer le risque de présence de vecteurs de la maladie, en fonction de la température et de la pluviométrie d’une zone d’Afrique (Fichiers “Machine_Learning_données_extraites.csv”, et “Algorithme_Machine_Learning.ipynb”). La variable “Danger” est une variable qualitative ordinale décrivant le risque qu’encourt la population sur une échelle de 0 à 3. Cette variable est directement ajoutée à la base de données sur une colonne “Danger”. L’algorithme permet aussi d’afficher un message précisant les traitements recommandés ou nécessaires en fonction du niveau de risque. Ces messages peuvent être adressés aux scientifiques présents sur les lieux, ou à la population directement, dans le meilleur des cas (si la population a accès à la technologie).

Comme on peut le voir ici, cet algorithme serait capable d’alerter en temps réel les populations, et de faire baisser drastiquement l’incidence de la Malaria dans certaines zones du monde.

Afin de rendre la collecte de données et l’algorithme plus performant, il serait possible de créer un appareil peu cher à la production, et distribué dans tous les foyers des pays touchés par la Malaria. Il serait plus pratique et rapide à utiliser que des scientifiques qui se rendent sur place assez souvent. Cet appareil serait capable de mesurer la température et l’humidité puis de voir si ces données dépassent les seuils à l’aide de l’algorithme simple, et ensuite d’alerter localement les populations. Cela serait donc utile à la prévention afin de diminuer les risques de mort par la maladie. Si le risque est très élevé, l’appareil pourrait émettre un bruit reconnaissable afin d’alerter les personnes présentes non loin de l’appareil. Pour contrer le problème d’accès à l’électricité dans certaines zones du monde, l’appareil fonctionnerait grâce à l’énergie solaire. On pourrait rajouter un paramètre sur l’appareil tel qu’un détecteur de mouvement sensible aux petits animaux et plus particulièrement les moustiques. Cela permettrait d’ajouter une variable à l’algorithme (le nombre de moustiques détectés par l’appareil), et ainsi d’améliorer sa précision.

```
La zone Afrique_N02 ne présente pas de risque actuellement.
La zone Afrique_N03 présente un risque de niveau 1. Utilisation de répulsif recommandé
La zone Afrique_C01 présente un risque de niveau 2. Utilisation de répulsif et alerte de la population fortement recommandé
La zone Afrique_C02 présente un risque de niveau 1. Utilisation de répulsif recommandé
La zone Afrique_C03 présente un risque de niveau 2. Utilisation de répulsif et alerte de la population fortement recommandé
La zone Afrique_S01 ne présente pas de risque actuellement.
La zone Afrique_S02 présente un risque de niveau 1. Utilisation de répulsif recommandé
La zone Afrique_S03 présente un risque de niveau 3. Confinement de la population recommandé. Traitement d'urgence nécessaire
```




DISCUSSION

Nous avons vu à travers notre projet que le risque d’émergence de maladies infectieuses lié au changement climatique est un réel enjeu pour le futur. L’action, ou l’inaction des gouvernements vis-à-vis de l’écologie, va représenter un tournant dans notre appréhension de ces risques. Sans la mise en place d’actions et de dispositifs efficaces pour faire face à la recrudescence de maladies telles que la Malaria, la vie d’une large partie de la population mondiale est en péril. C’est pourquoi le projet présenté ici tente de convaincre scientifiquement, par les statistiques, de l’importance de nos décisions vis-à-vis du climat.

De plus, ce projet a proposé diverses solutions pour faire face à ces problèmes, tel que l’élaboration d’une base de données collaborative, car l’accès aux données est primordial pour tout approche scientifique. En effet, une des premières limites rencontrées pour élaborer ce document a été l’accès aux différentes données climatiques, ou épidémiologiques dans chaque pays. Le nombre de mesures trouvées dans nos annexes étant assez faibles, il a été difficile d’établir des corrélations et des modèles significatifs statistiquement. Cependant, la mise en place d’une telle base de données nécessiterait un grand nombre de scientifiques et techniciens acceptant de partager leurs data. Cela peut s’avérer compliqué dans des pays en voie de développement, où la technologie est difficilement accessible, et l’envoi de scientifiques volontaires dans ces pays peut représenter un problème environnemental, mais aussi gouvernemental.

« Même si les idées proposées au cours de ce devoir comportent encore des zones d’ombres, il s’agit déjà d’un premier pas vers une sensibilisation au thème du paludisme et du changement climatique, et plus largement, à celui des maladies infectieuses et de la problématique de coopération internationale. »

Enfin, seulement quelques variables ont été présentées dans notre exemple de base de données. Afin qu’elle puisse être significative dans l’avancée du combat face à la Malaria, il serait nécessaire de prendre en compte beaucoup plus de variables écologiques, climatologiques, et épidémiologiques. Il est également à noter que des facteurs économiques, sociaux et politiques peuvent être responsables de la résurgence du paludisme, et rendent ainsi le traitement du sujet d’autant plus complexe.

Notre projet a aussi eu pour vocations de présenter des solutions d’alerte et de protection des populations par l’apprentissage automatique. En effet, la mise en place d’appareils intelligents pourrait à l’avenir être une arme de choix face aux maladies infectieuses. La collecte et le traitement automatique des données afin de conseiller les populations, via message par exemple (comme le fait le Japon avec les séismes), est une idée efficace pour la Malaria, qui se transmet par un vecteur. Cependant, le manque de formation aux technologies dans les pays en développement peut encore une fois représenter une limite de ce genre de projets. En effet, même s’il est possible de trouver des solutions aux problèmes d’accès à l’électricité par exemple, les appareils risquent tout de même de nécessiter de la maintenance, et donc des ressources humaines sur place capable de s’en occuper afin de les tenir en marche pour qu’ils restent efficaces.

Certaines populations touchées par la Malaria vivent dans des régions reculées, et il peut être compliqué de rentrer en contact avec eux, de leur distribuer les appareils et de les maintenir. Enfin, l’algorithme pourrait faire face à des problèmes de faux-positifs ou faux-négatifs car le manque de données climatologiques de ce genre de lieux peut interférer avec la précision de ces prédictions. Dans l’exemple du programme Python, très peu de variables sont utilisées pour prédire et faire évoluer l’algorithme, ce qui peut poser une grande limite de la sensibilité et de la spécificité de la détection de maladies.

Même si elles sont imparfaites, les actions proposées par ce projet sont nécessaires pour faire avancer les luttes contre les maladies infectieuses, et sur la sensibilisation au changement climatique.

Le paludisme n’est qu’un exemple de maladies dépendant grandement du climat, et à l’avenir il sera nécessaire de mettre en place des dispositifs similaires pour la Dengue, le Choléra, et tout autres maladies transmises par des vecteurs (notamment bactériens, ou de type insectes). Pour conclure, notre projet a vocation à inspirer les scientifiques à travers le monde afin de pouvoir collaborer, principalement sur l’accès aux données, car seules les données permettront d’arriver à des solutions scientifiques, via le partage et l’archivage, les statistiques, et le Machine Learning.

Annexes :

Sources Introduction:

[1] **Antonella Rossati.** (2017). Global warning and its health impact. *Int J Occup Environ Med.* 8(1):7-20

[2] Rapport du GIEC 2022

[3]<https://www.eird.org/isdr-biblio/PDF/Climate%20change%20and%20human%20health%20risks.pdf>

[4] **Richard K.M'Bra & Al.** (2018). Impact of climate variability on the transmission risk of malaria in northern Côte d'Ivoire. *PLoS One.* 13(6) : e0182304

[5]https://cdn.who.int/media/docs/default-source/malaria/world-malaria-reports/world-malaria-report-2021-global-briefing-kit-fre.pdf?sfvrsn=8e5e915_23&download=true

[6]<https://www.who.int/fr/news-room/fact-sheets/detail/malaria#:~:text=Le%20paludisme%20est%20une%20maladie,de%20paludisme%20dans%20le%20monde.>

[7]https://endmalaria.org/sites/default/files/Climate%20Change%20and%20Malaria_FR.pdf

[8]https://manuelnumeriquemax.belin.education/svt-seconde/topics/svt2-ch13-246-a_la-lutte-contre-le-paludisme

Bases de données pour Access:

[9] <https://databank.banquemondiale.org/source/health-nutrition-and-population-statistics>

[10] <https://fr.tutiempo.net/climat/mapa/?d=1>

Bases de données Statistiques:

[11] <https://databank.banquemondiale.org/source/health-nutrition-and-population-statistics>

[12]<https://www.statistiques.developpement-durable.gouv.fr/chiffres-cles-du-climat-france-europe-et-monde-edition-2022?rubrique=26&dossier=1263>

Remerciements :

Mme Cécile LAGAUDRIERE, pour avoir pris le temps de nous faire parvenir des articles.

Mr Nicolas BAYAN, pour avoir pris le temps de répondre à notre questionnaire et nous avoir fait parvenir des articles.

Mr Pascal RIGOLET, pour la transmission de ses connaissances autour du monde des datasciences (et au-delà) ainsi que son dévouement au sein de l'UE.