



HAL
open science

Méthodes de clustering et SIG pour la classification climatique à partir de données météorologiques géolocalisées : cas de l'île tropicale de la Réunion

Anne-Sophie Davidson, Jean Lalaina Rakotomalala, Bruno Malet-Damour,
Jean-Philippe Praene

► To cite this version:

Anne-Sophie Davidson, Jean Lalaina Rakotomalala, Bruno Malet-Damour, Jean-Philippe Praene. Méthodes de clustering et SIG pour la classification climatique à partir de données météorologiques géolocalisées : cas de l'île tropicale de la Réunion. Journées de Recherche des ISTs “ Les recherches scientifiques et technologiques, renforcement de la coopération régionale de l’Océan Indien ”, Nov 2022, Antananarivo, Madagascar. hal-04318008

HAL Id: hal-04318008

<https://hal.univ-reunion.fr/hal-04318008>

Submitted on 1 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Journées de Recherche des ISTs – 21-24 novembre 2022, Antananarivo, Madagascar
« Les recherches scientifiques et technologiques, renforcement de la coopération régionale de l'Océan Indien »

Méthodes de clustering et SIG pour la classification climatique à partir de données météorologiques géolocalisées : cas de l'île tropicale de la Réunion

Anne-Sophie DAVIDSON¹, Jean Lalaina RAKOTOMALALA², Bruno MALET-DAMOUR¹, Jean Philippe PRAENE¹

1 : Université de La Réunion, Laboratoire PIMENT, Réunion (France)

2 : Institut Supérieur de Technologie d'Antananarivo, Madagascar

anne-sophie.davidson@univ-reunion.fr

jean.lalaina@gmail.com

bruno.malet-damour@univ-reunion.fr

praene@univ-reunion.fr

Résumé

Lorsqu'on envisage le confort thermique et la réduction de la consommation d'énergie dans les zones tropicales, la conception de maisons bioclimatiques adaptées à leur climat est essentielle. Une caractéristique préliminaire de telles initiatives de conception est la collecte de connaissances précises sur les conditions climatiques de chaque partie du territoire investigué. L'objectif de cet article est de proposer un zonage climatique pour l'île de la Réunion en utilisant le clustering hiérarchique par composantes principales (HCPC). Les données utilisées sont issues de modèles de prévision climatique de Météo France. Les résultats de l'analyse sont ensuite couplés à une interpolation géographique à l'aide de techniques de système d'information géographique (SIG) nous permettant

de définir trois zones climatiques associées aux milieux secs, humides et de mi-pente et de hautes altitudes.

Mots-clés : zonage climatique ; Réunion ; analyse multivariée ; clustering ; SIG ; confort thermique ; bâtiment

Sommaire

Résumé.....	1
Sommaire	Erreur ! Signet non défini.
1. Introduction	3
a. État de l'art sur les méthodes de zonage climatique	4
b. Zonages climatiques existants de La Réunion	7
c. Originalité de ces travaux	14
2. Méthodes.....	15
a. Base de données utilisées.....	15
b. Regroupement hiérarchique k-means sur composantes principales (HCPC)	16
c. Cartographie basée sur le SIG	17
3. Résultats et Discussions	17
a. Zonage climatique	17
b. Répartition spatiale	19
4. Conclusions et perspectives.....	22
Annexes.....	23
Annexe 1: Caractéristiques des clusters.....	23
Annexe 2 : Caractéristiques principales des différentes zones climatiques PERENE.....	24
Références.....	24

Liste des figures

Figure 1: Cercle des corrélations	18
Figure 2: Résultats du clustering	19
Figure 3: Zonage climatique de La Réunion.....	20
Figure 4: Exposition aux vents de La Réunion	21

1. Introduction

Depuis la fin des années 1960, le monde s'interroge sur les dangers d'une trop forte croissance économique et démographique et les conséquences induites sur la consommation des ressources. L'un des indicateurs les plus parlants est l'évolution annuelle de la date du jour de dépassement, date à laquelle l'humanité est supposée avoir consommé l'ensemble des ressources que la planète est capable de produire en un an. Cette année, il était prévu le 28 juillet 2022, dans un contexte marqué par des phénomènes climatiques de plus en plus extrêmes (sécheresses, canicules, feux de forêt).

Face à cela, il paraît essentiel de chercher de nouvelles manières de concevoir, plus frugales, afin de respecter un écosystème déjà fragilisé. La recherche d'une architecture climatique durable, ou « bioclimatique », se construit sur une conception de bon sens en phase avec le contexte climatique d'un territoire. Durable, car le climat aura tendance à changer, l'architecture devra être efficace aujourd'hui et dans 50 ans. Dans ce contexte, il est nécessaire d'analyser le climat d'un territoire, d'identifier des groupes climatiques afin de pouvoir formuler des préconisations au plus proche des besoins.

L'île de La Réunion est, avec Mayotte, l'un des deux seuls départements français dans l'hémisphère Sud. Elle compte une population de près de 900 000 personnes, répartie en grande partie sur le littoral (zone des bas ; 2/4) et sur la zone de mi-altitude (1/4). Le littoral saturé, les Réunionnais s'orientent majoritairement vers la mi-pente. Pour garantir l'efficacité des bâtis, les modes constructifs doivent s'adapter au climat du lieu. Dans les bas, il est question d'établir une stratégie de rafraîchissement. Dans les hauts, la stratégie de confort s'oriente vers le stockage de la chaleur. Toutefois, la seule limite altimétrique ne peut pas offrir une réponse généralisée aux typologies de construction puisque l'île est soumise à une multitude de microclimats.

L'objectif de cet article est de proposer une nouvelle classification climatique de l'île de La Réunion en redéfinissant les limites géographiques des zones climatiques existantes

sur la base d'une analyse de données multivariées. Cette dernière est construite à partir d'une méthode de clustering associée à une analyse en composantes principales (ACP) (Praene et al., 2019). Elle utilise une très large base de données météorologiques spatialisée (Météo France) dont les paramètres choisis sont en phase avec les données nécessaires à la conception des bâtiments (Faure et al., 2008).

Cet article est structuré comme suit. Les zonages climatiques existants pour La Réunion seront présentés après un bilan bibliographique des méthodes de classification. La méthode de classification des données météorologiques sera présentée, et les zones climatiques résultantes, analysées. À l'aide d'outils SIG, une spatialisation de ces nouvelles zones climatiques sera présentée. Cette dernière sera finalement comparée aux classifications existantes.

a. État de l'art sur les méthodes de zonage climatique

Le zonage climatique pour les applications d'efficacité énergétique des bâtiments est un élément clé de nombreux programmes et politiques visant à améliorer les performances thermiques des bâtiments (Attia et al., 2019; Hao et al., 2021; Yang et al., 2008). En dépit de son importance, il n'y a actuellement pas de consensus sur la méthodologie appropriée pour le construire. Des études antérieures indiquent qu'une grande variété de méthodes et de paramètres sont actuellement utilisés pour le zonage climatique : les méthodes traditionnelles et les méthodes modernes.

Les méthodes traditionnelles se basent sur une définition au préalable des classes climatiques existantes et des seuils de classification définis subjectivement (Belda, n.d.; Jacobeit, 2010; Sa'adi et al., n.d.). Dans cette catégorie, la classification la plus célèbre est celle de Köppen, elle est souvent utilisée pour les études du climat et des changements climatiques dans le domaine de la géographie, de l'hydrologie, de l'agriculture et de la biologie (Kottek et al., 2006). La classification de Köppen met en relation la pluviométrie, la température de l'air et la distribution de la végétation. Selon Köppen, le type de végétation présente dans une localité caractérise au mieux son climat (Aparecido et al., 2016; Júnior et al., 2011). Le système de classification de Köppen

s'appuie sur une division en 5 groupes de la végétation. Chaque groupe correspondant à une zone climatique : zone équatoriale (A), zone aride (B), zone tempérée (C), zone continentale (D), et la zone polaire (E) (Kottek et al., 2006).

Cette classification impliquerait que le climat est un facteur suffisant à la caractérisation d'un type de végétation. Or d'autres facteurs sont à prendre en compte comme le facteur historique, les changements rapides des conditions climatiques, le facteur humain, etc (Zscheischler et al., 2012). De plus, des incohérences ont été soulignées quant aux classes climatiques révélées par la classification Köppen et les types de climats réellement observés. Des auteurs en déduisent que l'introduction de la variable altitude aiderait à affiner la représentation des climats (Peel et al., 2007). Cela permettrait d'améliorer l'extrapolation des types de climats identifiés par Köppen dans le cas des zones à faible densité de stations météorologiques et des régions montagneuses, et d'éviter que les stations situées à haute altitude n'étendent leur type de climat aux régions de basse altitude dans le cas où les données pour ces dernières seraient indisponibles. Aussi, des auteurs accusent le caractère rigide du système de Köppen face aux critères de délimitations des seuils (Sparovek et al., 2008) ; qui dans certains cas, incluent dans une même classe climatique diverses caractéristiques du paysage naturel, ce qui entrainerait normalement une forte variabilité interne des éléments climatiques. A cela s'ajoute le caractère généraliste de la classification de Köppen. Son efficacité concernant les études globales des changements climatiques sur une période donnée est indéniable, mais le manque de précision dans la définition des zones climatiques d'une région ciblée en fait sa plus grande faiblesse. Aussi, elle n'est pas adaptée à l'étude du confort des bâtiments (Attia et al., 2019).

Pour des fins de conception bioclimatique, la méthode de classification la plus utilisée est la méthode des degrés-jours. Le degré-jour est la différence entre la température moyenne extérieure au bâtiment et une température de confort fixée de manière arbitraire. Cette température de référence correspond à la température à laquelle la mise en marche du système de chauffage, ventilation et de climatisation n'est pas nécessaire. La température extérieure variant d'une région à une autre, le degré-jour permet de quantifier la quantité d'énergie nécessaire au maintien du confort des bâtiments dans différentes zones climatiques. Cette classification, de par sa simplicité,

ne prend en compte que la température, et est inadaptée à l'étude des régions chaudes et humides dans lesquelles le rayonnement solaire, l'humidité relative et la vitesse du vent influencent considérablement le confort des bâtis.

Ainsi, ces méthodes classiques revêtent ce que l'on pourrait qualifier de caractère arbitraire comparées aux méthodes modernes qui se basent sur des analyses statistiques plutôt que sur des seuils de classification et des types de climats définis à priori. Dans cette autre catégorie, la classification se fait par groupement des individus en classes auto-générées selon des critères statistiques particuliers pouvant être imposés (classification supervisée) ou non (classification non supervisée ou clustering) (Jacobeit, 2010; Praene et al., 2019; Sa'adi et al., n.d.).

La classification supervisée permet de prédire le positionnement de nouvelles observations dans des classes à partir de règles que nous aurons à déterminer. Elle nécessite donc de connaître à priori le nombre de groupes existant dans la population étudiée et le groupe auquel appartient chaque observation de la population (Kassambara, n.d.). Elle n'est pas souvent utilisée dans la littérature, et est plus associée au zonage climatique local (LCZ) qu'à la classification climatique. L'algorithme de classification supervisée ne peut être appliquée qu'à des régions dont les caractéristiques climatiques sont connues et détaillées. Ce fut le cas de La Chine, pour laquelle, Yang et al. se sont basés sur de précédentes études dans le but de mettre en place un zonage climatique du bâtiment préconisant des techniques de conception bioclimatique adaptées à chaque région.

Dans le cas où le nombre de groupes existant dans la population n'est pas connu à l'avance, nous aurons recours au clustering. Cette approche non supervisée ne cherche pas à apprendre une relation de corrélation entre les individus mais à former des groupes homogènes d'individus en fonction de leurs similarités. En climatologie, la distance euclidienne mesure la similarité entre les variables climatiques de deux stations météorologiques (Cortez et al., 2021; Fovell & Fovell, 1993; Praene et al., 2019; Sa'adi et al., n.d.). Le clustering présente l'avantage de pouvoir traiter un grand nombre de variables contrairement aux autres méthodes qui auraient tendance à simplifier la classification en ne prenant en compte que quelques variables. Cette méthode a

démontré son efficacité quant au zonage climatique adapté à la conception bioclimatique en combinant la climatologie et l'étude de performance énergétique du bâtiment (Hao et al., 2021; Walsh et al., 2017a, 2017b).

b. Zonages climatiques existants de La Réunion

L'île de la Réunion est soumise à un climat tropical humide avec un fort gisement solaire. Cette île est divisée en une quarantaine de microclimats en raison de sa topographie inhabituellement accidentée, résultat de sa géologie volcanique. La température décroît largement avec l'altitude. Elle s'abaisse d'environ de 0.6°C tous les 100m.

Il y existe 3 principaux types de vents : les alizés (vents réguliers de sud-sud-est (saisonniers) faibles à modérés), les brises (vents locaux induits par les différences de température entre la terre et la mer) et les foehns (dérivés des alizés lorsqu'ils redescendent des pentes). Les alizés sont déviés par le Nord et par le Sud, tandis qu'une partie d'entre eux s'engouffre vers le cœur de l'île. Ils drainent de grandes quantités de pluie sur l'est de l'île, ce qui a tendance à diviser l'île en deux.

Si l'on se concentre sur l'irradiation solaire et la température, on constate que le climat est tropical et humide sur la côte et plutôt tempéré en altitude. Sur le littoral les températures sont chaudes tout au long de l'année, variant de 21°C l'hiver à 26°C l'été. Tandis que dans les hauts, les températures sont beaucoup plus fraîches, et ce d'autant plus que l'on s'élève le long des pentes. En saison sèche des températures négatives sont relevées chaque année pour les points culminants.

L'amplitude thermique diurne et annuelle pour un endroit donné de l'île est relativement faible (de 5 à 7°C). Cela est principalement dû à l'effet "masse thermique" de la mer et du régime des vents.

Il y a globalement deux saisons distinctes sur l'île. L'été austral s'étend de novembre à avril. C'est la période où le gisement solaire est le plus abondant et les températures les plus élevées. Les journées y sont également les plus longues et les plus humides (forte pluviométrie) avec des risques cycloniques. Le mois de février est de loin le plus arrosé et

ceci sur l'ensemble de l'île (en moyenne 1.5 mètre d'eau à la station des hauts de Sainte-Rose). L'hiver austral (mai à octobre) est marqué par des journées plus courtes et des températures diurnes plus basses. La saison est relativement sèche, fraîche et moins humide avec une prédominance des alizés d'est et de sud-est. À l'ouest de l'île, la pluviométrie y est quasiment nulle (4.2mm en juillet à Saint-Gilles), contre 700mm d'eau dans l'est durant le mois le plus arrosé.

Les fortes sollicitations distribuées sur le territoire indiquent qu'il est nécessaire de construire des classes climatiques en vue d'apporter des préconisations de conception architecturale les plus adaptées.

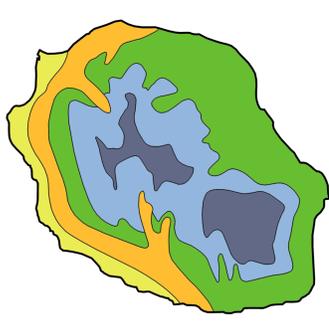
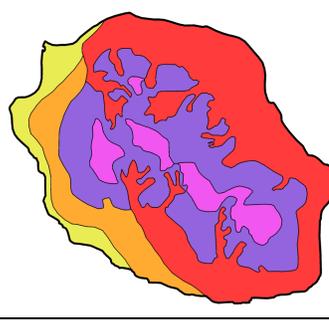
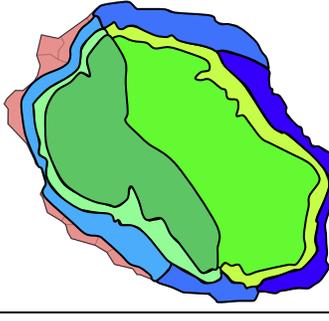
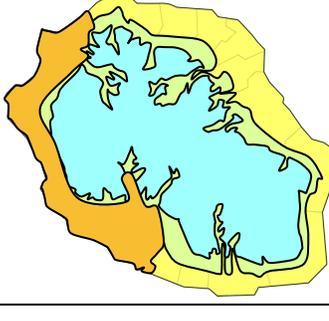
Depuis les années 80, plusieurs chercheurs se sont intéressés à la catégorisation des climats de l'île. Toutefois, aucune classe climatique ne semble indiquer les mêmes informations, car cette dernière varie au-delà de la simple dichotomie Ouest (côte au vent) / Est (côte sous le vent) et de l'étagement entre les Hauts et le littoral. Les études tendent généralement vers des analyses orientées ou interdisciplinaires, dont l'objectif final ajustera la typologie des données d'entrée (climatologique, topographique, scénographique, agricole, hydrologique, etc.).

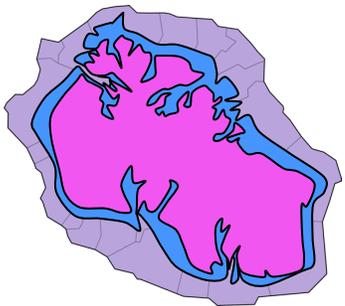
Malgré une disponibilité des données considérablement accrue depuis 5 décennies, cette diversité risque de perdurer étant donné la nature structurelle de la question. Une répartition géographique climatique d'un territoire devra donc se faire en fonction des besoins finaux du secteur visé.

Toutefois, le bilan littéraire des différentes classifications existantes pour l'île offrira des éléments pour l'analyse comparative avec le zonage climatique proposée dans le cadre de cet article.

Nous avons relevé cinq classifications climatiques :

- Zonage bioclimatique selon Thérésien Cadet (Cadet, 1980)
- Zonage morphoclimatique selon Bougère (Bougère, 1987)
- Classification climatique mondiale de Köppen-Geiger (Kottek et al., 2006)
- Zonage climatique du référentiel PERENE (PERENE, 2009)
- Zonage climatique de la RTAADOM (RTAA DOM, 2016)

Désignation	Cartographie du zonage climatique	Légende	Domaine d'utilisation	Référence
Zonage bioclimatique selon Cadet (1980)		<ul style="list-style-type: none"> Zone 1 : Végétation dégradée du littoral ouest Zone 2 : Forêt adaptée à la chaleur et à une saison sèche Zone 3 : Forêt adaptée à la chaleur et à l'humidité Zone 4 : Forêt humide d'altitude Zone 5 : Végétation des hautes terres froides 	Botanique, agriculture, biologie	(Cadet, 1980)
Zonage morphoclimatique selon Bougère (1984)		<ul style="list-style-type: none"> Zone 1 : zone à dégradation modérée et altération faible Zone 2 : zone à dégradation forte et altération faible Zone 3 : zone à dégradation et altération modérée Zone 4 : zone à dégradation faible et altération modérée Zone 5 : zone à très forte altération 	Scénologie, géologie, volcanologie, hydrologie	(Bougère, 1987)
Classification climatique de Köppen-Geiger (1961 - 2006)		<ul style="list-style-type: none"> Zone 1 : BSh Zone 2 : Aw Zone 3 : Am Zone 4 : Af Zone 5 : Cwa Zone 6 : Cfa Zone 7 : Cwb Zone 8 : Cfb 	Hydrologie, géographie, agriculture, biologie ou climatologie	(Kottek et al., 2006)
Zonage climatique de PERENE (2009)		<ul style="list-style-type: none"> Zone 1 : zone sous le vent Zone 2 : zone au vent Zone 3 : zone des hauts Zone 4 : zone d'altitude 	Architecture et construction	(PERENE, 2009)

Zonage climatique RTAADOM (2016)		Zone 1 : Zone des bas (<400m altitude) Zone 2 : Zone des hauts (400 – 600m d'altitude) Zone 3 : Zone de haute altitude (>600m d'altitude)	Architecture et construction	(RTAA DOM, 2016)
-------------------------------------	---	---	------------------------------	------------------

Thérésien Cadet (Cadet, 1980), célèbre botaniste réunionnais, a travaillé sur le premier zonage climatique de l'île dans le cadre de ses travaux de thèse en 1980. Ce « zonage bioclimatique » a permis d'apporter des informations climatiques dans ses travaux de recherche sur la flore réunionnais. Cinq classes sont ressorties de son travail, définissant la typologie d'environnement forestier :

- Zone 1 : végétation dégradée du littoral ouest
- Zone 2 : forêt adaptée à la chaleur et à une saison sèche
- Zone 3 : forêt adaptée à la chaleur et à l'humidité
- Zone 4 : forêt humide d'altitude
- Zone 5 : végétation des hautes terres froides

La zone littorale Ouest est divisée en deux parties, dont l'altimétrie semble être l'un des éléments pris en compte. A contrario, à l'Est et dans la zone de mi-pente ouest, le climat semble être chaud et humide. Les hauteurs et les cirques sont regroupés dans deux zones où humidité et fraîcheur sont les caractéristiques dominantes. Nous remarquons que cette répartition climatique divise l'île d'est en ouest. La pluviométrie, paramètre essentiel dans l'étude botanique, a dû contribuer à cette analyse.

En 1984, Jacques Bougère (Bougère, 1987) a établi et publié une carte des régions morphoclimatiques. Cette dernière se base sur l'érosion des sols et les typologies géologiques, la pluviométrie et le relief. À l'image de la classification de Cadet, elle divise l'île en 5 zones climatiques :

- Zone 1 : zone à dégradation modérée et altération faible

- Zone 2 : zone à dégradation forte et altération faible
- Zone 3 : zone à dégradation et altération modérée
- Zone 4 : zone à dégradation faible et altération modérée
- Zone 5 : zone à très forte altération

Les limites climatiques de Bougère sont proches de celles de Cadet, à la différence près qu'elle explicite une réelle dichotomie entre l'Est et l'Ouest. La zone des hauts semble être plus répartie, où la dégradation des sols est faible avec une altération modérée à forte.

Le botaniste Köppen a inventé, en 1900, une classification combinant la carte mondiale de la végétation publiée par Griesbachet (1866) et la division du climat en cinq zones par Candolle. La version actualisée en 1961 par Geiger est la plus couramment utilisée, notamment dans les secteurs de l'hydrologie, de la géographie, de l'agriculture, de la biologie ou de la climatologie. Cette classification climatique (Kottek et al., 2006) a été également déclinée pour le territoire réunionnais. Selon Köppen, la pluviosité est un élément majeur pour définir les saisons, c'est pour cette raison que nous retrouvons la division de l'île d'est en ouest. Elle permet de diviser l'île en 6 zones climatiques :

- *Climat sec* :
 - Zone 1 : climat semi-aride (BSh)
- *Climat tropical* :
 - Zone 2 : climat tropical de savane (Aw)
 - Zone 3 : climat de moussons (Am)
 - Zone 4 : climat équatorial (Af)
- *Climat subtropical humide* : zone climatique tempérée caractérisée par des étés chauds et humides.
 - Zone 5 : climat été chaud avec hiver sec (Cwa)
 - Zone 6 : climat été chaud sans saison sèche (Cfa)
- *Climat océanique* : zone climatique tempérée caractérisée par des hivers doux et humides et des étés plus frais que pour les climats subtropicaux
 - Zone 7 : climat été tempéré avec hiver sec (Cwb)

- Zone 8 : climat été tempéré sans saison sèche (Cfb)

Pour La Réunion, la classification climatique de Köppen apporte une solution pour la caractérisation des types généraux de climats, sans pour autant prendre en compte les variabilités de microéchelles, typiques du climat réunionnais. Toutefois, la diversité climatique marque un tournant dans l'appréhension des typologies de climat. L'île n'est pas « que » tropicale, elle présente également des climats plus diversifiés pouvant être secs, océaniques ou encore tempérés.

Plus récemment, le référentiel PERENE (Performance énergétique des bâtiments à la Réunion) a été créé en 2009 (PERENE, 2009). Il a été établi en vue d'adapter les modes constructifs aux climats. Le zonage climatique proposé a été construit par superposition de couches de données météorologiques spatialisées issues de stations météorologiques de Météo France. Quatre zones climatiques y sont définies :

- Zone 1 : « Sous le vent »
Localisation : De 0m à 400 m pour la zone de La Possession à Saint-Pierre en passant par la zone littorale ouest.
- Zone 2 : « Au vent »
Localisation : De 0m à 400 m pour la zone de Saint Denis à Saint-Pierre en passant par Saint-Benoît (Saint-Pierre exclu puisqu'appartenant à la zone 1).
- Zone 3 : « Zone des Hauts »
Localisation : entre la ligne des 400 et la ligne des 800 m
- Zone 4 : « Zone d'altitude »
Localisation : Ligne au-dessus de 800 m, y compris les cirques

PERENE a eu le mérite d'avoir proposé la première classification climatique de l'île pour le secteur de l'architecture et de la construction. Toutefois, la division administrative de ce zonage apporte de l'ambiguïté et de l'incohérence : par exemple, une architecture à Saint-Pierre (Zone 1) devrait être conçue avec des règles spécifiques et différentes qu'en Zone 2 pour un éloignement ne pouvant justifier ce choix.

En 2016, l'île s'est vue dotée de la première réglementation thermique (RTAADOM) visant à standardiser et normer les typologies de construction tout en permettant de

proposer des solutions techniques rendant les logements énergétiquement performants (RTAA DOM, 2016). La catégorisation climatique indiquée dans cette réglementation divise l'île en fonction de son relief et de la dominance des vents. Cette dernière sépare le territoire en trois zones climatiques :

- Zone 1 : zone des bas (<400m altitude)
- Zone 2 : zone des hauts (de 400 m à 600m d'altitude)
- Zone 3 : zone de haute altitude (altitude>600 m).

La division altimétrique semble offrir des résultats plus en accord avec les dynamiques climatiques. Toutefois, la séparation territoriale en trois classes semble avoir des limites lorsque l'on s'intéresse de plus près à la cohérence climatique de cette catégorisation ; En effet, les climats littoraux Est et Ouest, pourtant différents, sont similaires au regard de la réglementation.

c. Originalité de ces travaux

Cet article cherche à utiliser des outils d'analyse statistique afin de redéfinir le nombre et les contours des zones climatiques réunionnaises. Pour avoir un zonage des plus cohérents avec les besoins en architecture et de la construction, il est nécessaire d'avoir en entrée le maximum d'information. Dans notre cas, les bases de données exploitées sont les plus fournies à La Réunion, puisqu'elles sont issues d'un modèle de prévision utilisé par l'opérateur météorologiste de l'île, Météo France. Aussi, les données utilisées sont en phase avec les paramètres physiques essentiels pour le dimensionnement des bâtiments tropicaux. Le zonage climatique est alors construit en utilisant le clustering hiérarchique par composantes principales (HCPC). Cette technique permet de regrouper, par classe, les zones à forte ressemblance. En nous basant sur une classification statistique de données météorologiques plus récentes et de meilleure résolution, nous serions en mesure d'obtenir une représentation plus fidèle des climats réellement observés et de mieux comprendre le climat réunionnais. En effet, la répartition spatiale disparate de stations météorologiques sur le territoire réunionnais ne permet pas toujours l'acquisition des données exactes pour une zone considérée. Or, l'extrapolation d'un type de climat

sur une région voisine dans le cas de l'absence de fichiers météorologiques peut induire à préconiser des technologies et des conceptions inadaptées conduisant à une performance énergétique non optimale (Hao et al., 2021).

La spatialisation des données à l'aide d'outils de SIG permet ensuite de tracer les nouveaux contours de notre zonage, cela permettant de comparer notre contribution aux zonages climatiques recensés précédemment. A l'issue de notre classification climatique, nous obtiendrons une distribution typique à une incertitude près du comportement climatique de chaque zone.

2. Méthodes

a. Base de données utilisées

Depuis fin 2006, Météo-France à la Réunion dispose d'un modèle opérationnel haute résolution appelé ALADIN (Faure et al., 2008). Ce modèle propose des données climatiques à une résolution de 3 km. Il permet à la fois une modélisation précise de l'écoulement atmosphérique autour des reliefs réunionnais. Les données utilisées dans le cadre de cet article sont issues des résultats du modèle ALADIN.

Les données utilisées sont les suivantes :

- Humidité relative quotidienne min « UNQ » et max « UXQ » (en %)
- Rayonnement solaire global quotidien « RGQ » en (J/cm²)
- Vitesse de vent moyenne quotidienne « FFMQ » en (m/s)
- Température minimale « TNQ » et maximale « TXQ » quotidienne (en °C)
- Cumul quotidien des précipitations « RRQ » (en mm)

Nous avons également obtenu la variable altitude à partir d'un modèle d'élévation digital de Shuttle Radar Topography Mission System (SRTM Data, <https://srtm.csi.cgiar.org/srtmdata/>) d'une résolution de 90 m. Le fichier SRTM fût utilisé en format raster sur l'outil de système d'information géographique QGIS. Nous avons jugé intéressant d'ajouter aux variables longitude et latitude, la variable altitude dans le but d'apprécier l'impact réel de la topographie sur le climat.

Le jeu de données utilisé pour notre étude regroupe donc les principaux paramètres influents dans l'évaluation du confort dans les zones chaudes et humides (Attia et al., 2019; Yang et al., 2020) mesurés sur 896 points répartis sur le territoire réunionnais. Un prétraitement des données sur le logiciel Big Data R a été nécessaire afin de nettoyer notre base de données et de supprimer les points de notre grille spatiale contenus dans la mer ainsi que les observations manquantes. En effet, l'analyse par composante principale sur R remplacera par défaut les valeurs manquantes par la moyenne des variables observées, ce qui risquerait fausser notre zonage, notamment lorsque les valeurs manquantes sont en nombre conséquent. Dans le cas de notre étude, la dimension de notre matrice de données est passée de 896 x 10 à 242 x 10.

b. Regroupement hiérarchique k-means sur composantes principales (HCPC)

L'ACP est un test statistique associé à l'analyse factorielle. Cette technique mathématique cherche à représenter la variance de l'ensemble des données (c'est-à-dire les réponses utilisées pour caractériser les échantillons) en utilisant un nombre minimal de composantes. En général, des projections bidimensionnelles ou tridimensionnelles des échantillons sont générées pour l'analyse visuelle, les axes (composantes principales, notées Dim.) servant de facteurs. Chaque composante est une combinaison linéaire des données originales et sont orthogonales entre. Leur calcul se fait de façon itérative, de sorte que les premières composantes retiennent la plus grande part de variance. Cela revient à sélectionner uniquement les variables les plus influentes sur la variance totale de base de données (Jolliffe & Cadima, 2016; Zscheischler et al., 2012). Pour retenir le nombre de composantes nécessaires à l'analyse, nous utilisons le critère de Kaiser, qui considère que seules les composantes dont la valeur propre est supérieure à 1 doivent être considérées (Kaiser, 1960).

A l'issue de cette étape, nous avons effectué un regroupement hiérarchique sur les composantes principales. La méthode des K-means est utilisée pour le partitionnement des données. Le principe général de cet algorithme est de minimiser la variance totale

au sein d'un cluster. Par itération, cette méthode présente l'intérêt de toujours converger. L'optimisation du nombre de clusters est faite par mesure globale selon la méthode Silhouette (Husson François, n.d.). Ainsi une première classification ascendante hiérarchique est effectuée, puis cette dernière est consolidée par les K-means (Husson François et al., 2010).

La similitude étant déterminée par le calcul de la distance euclidienne. Au bout de 100 itérations, les 242 stations sont groupées selon les critères de Ward.

Les clusters obtenus sont alors projetés sur une carte de La Réunion afin d'obtenir un zonage climatique et topographique.

c. Cartographie basée sur le SIG

L'étude réalisée exploite un jeu de données climatiques réparties sur un réseau spatial à maillage serré. Une interpolation spatiale est néanmoins nécessaire afin de lisser notre carte de zonage climatique. L'interpolation se base sur la corrélation entre les objets spatialement rapprochés : un point inconnu présentera plus de similarités avec un point observé qui lui est proche qu'avec un point observé qui lui serait éloigné.

Dans notre cas, nous avons opté pour la méthode de pondération par distance inverse dite IDW (Inverse Distance Weighting). L'estimation de la valeur d'un point est réalisée en faisant la moyenne des valeurs des points situés dans son voisinage, pondérée par l'inverse de la distance de chaque point (Praene et al., 2019).

3. Résultats et Discussions

a. Zonage climatique

L'ACP révèle que les deux premières dimensions représentent 62,52% de la variabilité totale du jeu de données. C'est un pourcentage assez conséquent et cette première projection est suffisamment représentative de la variance de nos données (Tableau 1). Les variables températures (Tmax et Tmin) et altitude sont significativement influentes sur la dimension 1 (Figure 1). Cela laisse envisager que les individus correctement projetés sur cet axe correspondront aux régions chaudes de basse altitude. Les variables

longitude et précipitations contribuent le plus à la construction de la dimension 2, la variable latitude quant à elle sera mal projetée sur cet axe. La dimension 2 représente ainsi les régions de la côte Est à forte pluviométrie.

Tableau 1: Résumé des deux premières composantes principales

Composante principale	Valeur propre	Variance (%)	Variance cumulée (%)
Dim1	4.184	41.841	41.841
Dim2	2.068	20.679	62.520

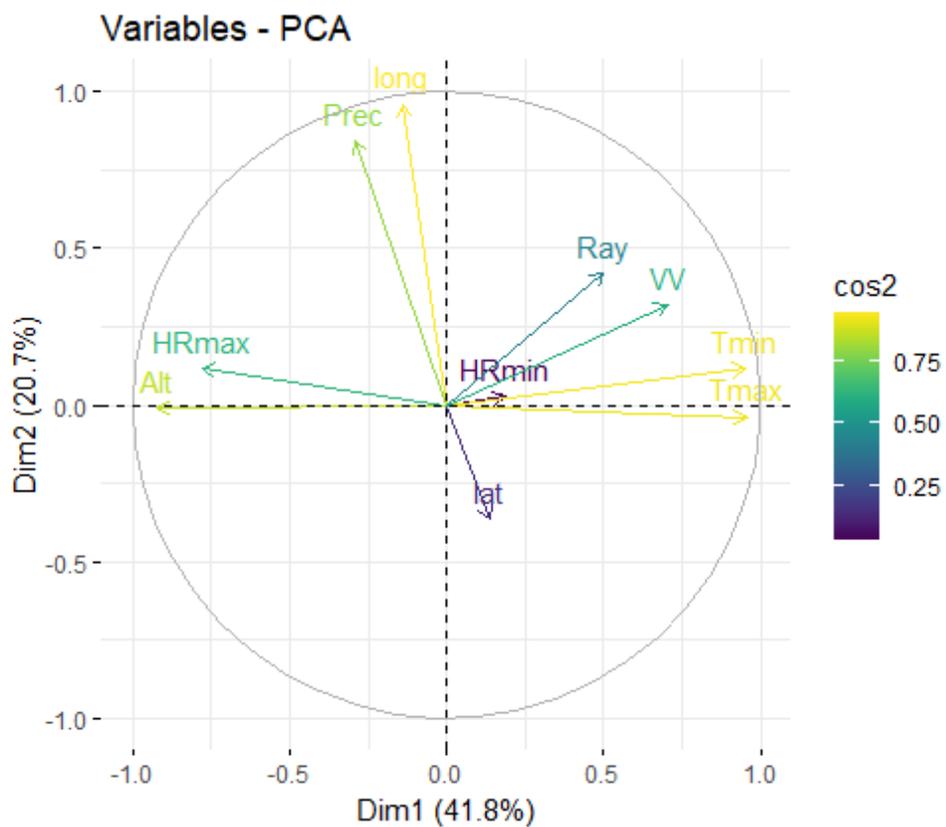


Figure 1: Cercle des corrélations

La caractérisation de ces axes principaux est primordiale à la description des clusters. En effet, les clusters résultants de l'HCPC seront projetés sur le plan 1:2. La Figure 2 nous montre la présence de 3 clusters. Le premier correspondant aux zones des hauts caractérisé par une forte humidité relative et une température basse par rapport à la moyenne de l'échantillon étudié. Le deuxième cluster représente la côte Est sujette à des pluies abondantes et soumise à des vents forts. Ce cluster regroupe les zones d'altimétrie moyenne. Le troisième cluster quant à lui est défini comme la zone la plus chaude avec un rayonnement solaire élevé, et une altitude moyenne basse ; elle coïncide avec la côte Ouest.

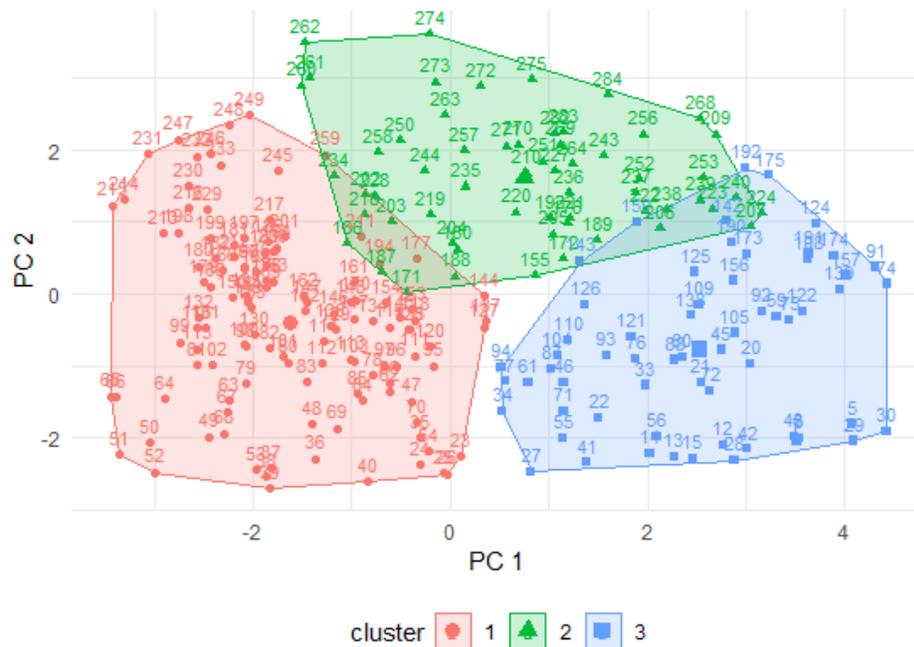


Figure 2: Résultats du clustering

b. Répartition spatiale

La projection spatiale du clustering nous donne la Figure 3, elle met en évidence la distinction entre les climats de l'Est, de l'Ouest et des hauts.

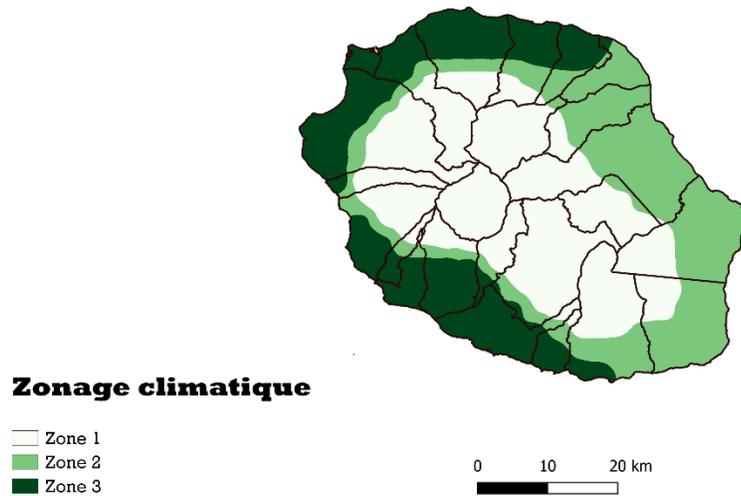


Figure 3: Zonage climatique de La Réunion

La classification statistique effectuée sur notre base de données de haute résolution permet d'obtenir une représentation fidèle des climats observés.

Ce zonage rappelle la zonation naturelle Est-Ouest de l'île (Figure 4) : la côte au vent soumise à des vents chauds et humides venant de l'équateur et la côte sous le vent protégée des vents par les massifs montagneux et moins arrosée.

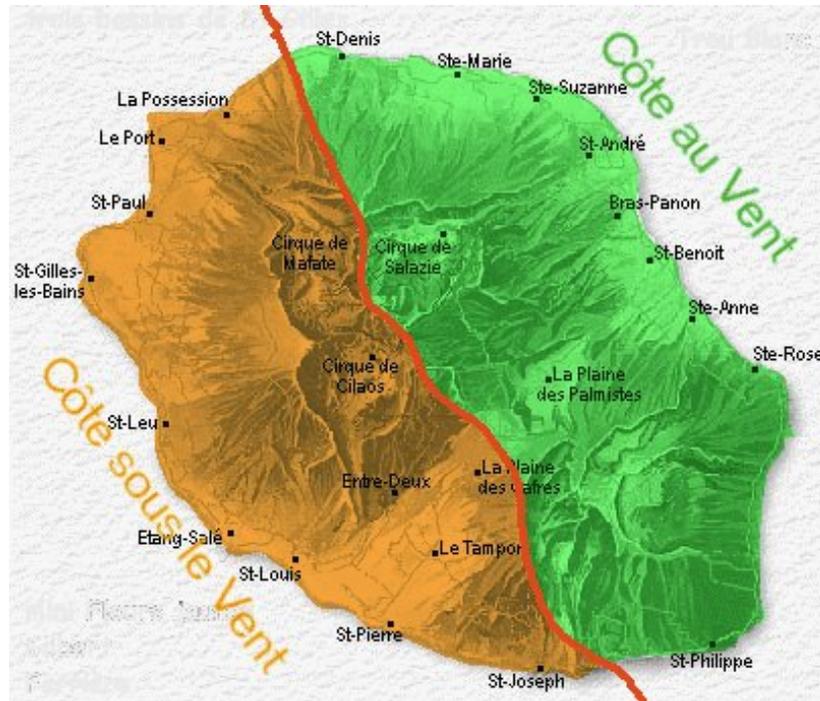


Figure 4: Exposition aux vents de La Réunion

(Boutry et al., n.d.)

Le référentiel PERENE découle d'une classification obtenue en multipliant des couches de données météorologiques brutes déjà cartographiées. Cela implique que des seuils de classifications ont préalablement été affectés à ces cartographies contrairement au clustering où les stations météorologiques sont regroupées en fonction de leurs similarités plutôt qu'en fonction de classes que nous aurions nous même prédéfinies. Ceci dit, le zonage climatique de PERENE et celui établi dans cette étude concordent dans la mise en évidence la disparité entre l'Est, l'Ouest et la partie centrale de l'île avec une particularité pour le nouveau zonage pour lequel la zone dite « des hauts » a été fusionnée à la zone « au vent ». Cela s'explique par les limites de classification définies par PERENE se rapprochant des moyennes des composantes principales du cluster 2 (Annexe 2 :_Caractéristiques principales des différentes zones climatiques PERENE).

Aussi, le nouveau zonage s'accorde avec le zonage climatique de la RTAADOM, confirmant l'importance de l'altimétrie dans la classification climatique. En effet, tout comme la RTAADOM, notre zonage a divisé l'île en trois zones bien définies, mais diffère en soulignant la dissimilarité entre les climats littoraux Est et Ouest :

- Zone 1 : zone de haute altitude avec une moyenne de 1375m,
- Zone 2 : zone Est avec une altitude moyenne de 530m,
- Zone 3 : zone Ouest d'altimétrie moyenne égale à 257m.

Le zonage résultant de l'analyse statistique fait donc un recoupement entre le zonage de PERENE et de la RTAADOM tout en avançant que le climat littoral est scindé en deux et que la partie centrale dispose d'un climat à part dû à la topographie accidentée.

4. Conclusions et perspectives

La zone littorale et la zone mi- altitude de l'île de La Réunion arrivant à saturation, il devient primordial d'établir de nouveaux logements dans les mi-pentes. Or, la température diminue à mesure que l'altitude croît. Les conceptions doivent donc s'adapter au climat. Pour ainsi faire, le zonage climatique est la première étape dans la mise en place de démarches d'efficacité énergétique dans le bâtiment et de réduction de l'empreinte carbone liée à la consommation énergétique du secteur de la construction. L'objectif de cet article, qui constitue la première étape de la mise en place d'un nouveau référentiel de conception bioclimatique à La Réunion, a donc été de redéfinir le zonage climatique.

Un large panel de méthodes de classification existe dans la littérature. Le choix de la méthode repose sur l'objectif de l'étude et les données disponibles. Pour le cas de La Réunion, nous avons à notre disposition une très large base de données météorologiques spatialisée. Le clustering paraît alors comme la méthode la plus adaptée dans ce cas d'étude multivariée. La méthode proposée peut être appliquée à d'autre pays, elle est moins chronophage étant donné qu'il n'est pas nécessaire de définir des seuils de classification en amont, contrairement à d'autres méthodes identifiées dans la littérature.

L'étude a divisé le territoire réunionnais en trois zones et a permis de prouver que l'altimétrie est bien un des facteurs influants dans le zonage climatique, confirmant ainsi la légitimité du zonage climatique proposé par la RTAADOM. Néanmoins, dans l'optique de mettre en œuvre un programme d'efficacité énergétique, les paramètres en phase avec les données nécessaires à la conception des bâtiments doivent également entrer

en compte. C'est le cas de notre étude qui intègre les températures, le rayonnement solaire, l'humidité relative, la pluviométrie ainsi que la vitesse du vent. Ce zonage ouvre la voie vers la prochaine étape qui sera de créer des fichiers météorologiques types pour chaque zone pour le besoin des études et simulations thermiques futures, nous obtiendrons ainsi une distribution typique à une incertitude près du comportement climatique d'une zone donnée. En effet, la répartition disparate actuelle des stations météorologiques de l'île ne permet pas toujours l'acquisition de données exactes pour une zone d'étude fixée.

Remerciements

Ce travail s'inscrit dans le cadre d'une thèse de doctorat ayant reçu une subvention des fonds FEDER INTERREG. Les auteurs souhaitent vivement remercier la Région Réunion et l'Europe pour le soutien.

Les auteurs remercient également Mme Marie-Dominique LEROUX et M. François BONARDOT de Météo-France pour la mise à disposition des données météorologiques.

Annexes

Annexe 1: Caractéristiques des clusters

		v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
Zone 1	Alt	11.7794	1375.418	887.719	477.2354	648.0751	0
	HRmax	9.6789	93.95	91.7075	1.7768	3.6267	0
	long	-2.2249	55.5107	55.5311	0.1155	0.1431	0.0261
	HRmin	-3.101	56.7924	57.7192	5.4582	4.6782	0.0019
	Ray	-6.7743	1602.436	1654.604	117.0743	120.5424	0
	VV	-9.1352	2.1898	2.7528	0.4351	0.9647	0
	Tmax	-11.9471	20.9136	23.3872	2.0544	3.2408	0
	Tmin	-13.0383	12.0756	15.1375	2.1187	3.6759	0
Zone 2	long	9.885	55.6915	55.5311	0.0691	0.1431	0
	Prec	9.6964	12.7639	7.744	3.8515	4.5635	0
	Tmin	6.0534	17.6619	15.1375	2.1101	3.6759	0
	HRmin	5.8972	60.849	57.7192	2.9805	4.6782	0

	Tmax	3.7565	24.7683	23.3872	2.109	3.2408	2,00E-04
	VV	3.4395	3.1293	2.7528	0.8703	0.9647	6,00E-04
	Alt	-4.8609	530.339	887.719	415.9436	648.0751	0
Zone 3	Tmax	10.0424	26.9984	23.3872	1.3419	3.2408	0
	Tmin	9.0275	18.8196	15.1375	1.3601	3.6759	0
	VV	7.118	3.5148	2.7528	1.1088	0.9647	0
	Ray	6.2599	1738.332	1654.604	102.715	120.5424	0
	HRmin	-2.2608	56.5456	57.7192	2.4659	4.6782	0.0238
	long	-7.2132	55.4166	55.5311	0.1053	0.1431	0
	Prec	-8.0837	3.6507	7.744	1.6234	4.5635	0
	Alt	-8.7572	257.9836	887.719	218.152	648.0751	0
	HRmax	-11.5519	87.0588	91.7075	3.1183	3.6267	0

Annexe 2 :_Caractéristiques principales des différentes zones climatiques

PERENE

	Insolation moyenne (heure)	Rayonnement Global (cumul jour Wh/m ²)	Rayonnement diffus (cumul jour Wh/m ²)	Température min(°C)	Température max(°C)	Humidité relative min (%)	Humidité relative max (%)	Intensité moyenne du vent
Zone sous le vent	8	6510	2109,5	21,15	31	47	87,5	2,15
Zone au vent	6,5	30763,5	2124	18,05	28,8	54,5	89	2,9
Zone des hauts	6,5	5333,5	2010	15,45	26	61	89	1,55
Zone d'altitude	6	4723,5	1577	6,7	19,55	75,5	97	2,3

Références

- Attia, S., Lacombe, T., Rakotondramiarana, H. T., Garde, F., & Roshan, G. R. (2019). Analysis tool for bioclimatic design strategies in hot humid climates. *Sustainable Cities and Society*, 45, 8–24. <https://doi.org/10.1016/J.SCS.2018.11.025>
- Bougère, J. (1987). Atouts et contraintes du milieu naturel réunionnais (Natural potentialities and constraints in La Réunion). *Bulletin de l'Association de Géographes Français*, 64(5), 407–413. <https://doi.org/10.3406/bagf.1987.1407>
- Boutry, S., Gassiole, G., Rosebery, J., Giraudel, J. L., Peres, F., Coste, M., & Delmas, F. (n.d.). *Mise au point d'un indice diatomique pour les cours d'eau de la Réunion (IDR) : Rapport final sur la démarche d'élaboration de l'indice * Irstea-Cemagref, Gpt de Bordeaux, UR REBX ** Asconit Consultants, implantations de Boulogne-sur-Gesse et de la Réunion Bras de la Plaine (photo Gilles Gassiole).*
- Cadet, T. (1980). *La végétation de l'île de La Réunion : étude Phytoécologique et Phytosociologique. Thèse d'Etat, Université d'Aix Marseille, France. 312 p. – Espèces*

- Envahissantes Outre-mer*. <https://especes-envahissantes-outremer.fr/bibliotheque/cadet-t-1980-la-vegetation-de-lile-de-la-reunion-etude-phytoecologique-et-phytosociologique-these-detat-universite-daix-marseille-france-312-p/>
- Cortez, D., Herrera, S., Araya-Osses, D., Caroca, C., Padilla, R., Uribe, J. M., & Paneque, M. (2021). Topoclimatic zoning of continental Chile. *Https://Doi.Org/10.1080/17445647.2021.1886188*, 17(2), 114–124. <https://doi.org/10.1080/17445647.2021.1886188>
- Faure, G., Westrelin, S., & Roy, D. D. (2008). Un nouveau modèle de prévision à Météo-France : Aladin-Réunion. In *Météo et Climat* (Vol. 8, Issue 60). <https://hal-meteofrance.archives-ouvertes.fr/meteo-00311760>
- Fovell, R. G., & Fovell, M. Y. C. (1993). Climate zones of the conterminous United States defined using cluster analysis. *Journal of Climate*, 6(11), 2103–2135. [https://doi.org/10.1175/1520-0442\(1993\)006<2103:CZOTCU>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<2103:CZOTCU>2.0.CO;2)
- Hao, Z., Zhang, X., Xie, J., Wang, J., & Liu, J. (2021). Building climate zones of major marine islands in China defined using two-stage zoning method and clustering analysis. *Frontiers of Architectural Research*, 10(1), 134–147. <https://doi.org/10.1016/J.FOAR.2020.07.004>
- Husson François. (n.d.). *Sous la direction de FRANÇOIS HUSSON R pour la statistique et la science des données R pour la statistique et la science des données*. Retrieved August 7, 2022, from www.pur-editions.fr
- Husson François, Josse Julie, & Pagès Jérôme. (2010). *Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data?* http://factominer.free.fr/more/HPCPC_husson_josse.pdf
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065). <https://doi.org/10.1098/RSTA.2015.0202>
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259–263. <https://doi.org/10.1127/0941-2948/2006/0130>
- PERENE. (2009).
- Praene, J. P., Malet-Damour, B., Radanielina, M. H., Fontaine, L., & Rivière, G. (2019). GIS-based approach to identify climatic zoning: A hierarchical clustering on principal component analysis. *Building and Environment*, 164, 106330. <https://doi.org/10.1016/J.BUILDENV.2019.106330>
- RTAA DOM. (2016).
- Sa'adi, Z., Shahid, S., Mohammed, , & Shiru, S. (n.d.). Defining climate zone of Borneo based on cluster analysis. *Theoretical and Applied Climatology*. <https://doi.org/10.1007/s00704-021-03701-1>
- SRTM Data. (n.d.). Retrieved August 4, 2022, from <https://srtm.csi.cgiar.org/srtmdata/>
- Walsh, A., Cóstola, D., & Labaki, L. C. (2017a). Review of methods for climatic zoning for building energy efficiency programs. *Building and Environment*, 112, 337–350. <https://doi.org/10.1016/J.BUILDENV.2016.11.046>
- Walsh, A., Cóstola, D., & Labaki, L. C. (2017b). Comparison of three climatic zoning methodologies for building energy efficiency applications. *Energy and Buildings*, 146, 111–121. <https://doi.org/10.1016/J.ENBUILD.2017.04.044>
- Yang, L., Lam, J. C., & Tsang, C. L. (2008). Energy performance of building envelopes in different climate zones in China. *Applied Energy*, 85(9), 800–817. <https://doi.org/10.1016/J.APENERGY.2007.11.002>
- Yang, L., Lyu, K., Li, H., & Liu, Y. (2020). Building climate zoning in China using supervised classification-based machine learning. *Building and Environment*, 171, 106663. <https://doi.org/10.1016/J.BUILDENV.2020.106663>

Zscheischler, J., Mahecha, M. D., & Harmeling, S. (2012). Climate Classifications: the Value of Unsupervised Clustering. *Procedia Computer Science*, 9, 897–906.
<https://doi.org/10.1016/J.PROCS.2012.04.096>