



VEILLE BIOACOUSTIQUE de l'AVIFAUNE en GUADELOUPE

Nicolas Deloustal, Hervé Glotin
DYNI LIS CNRS RR report 20220626
CNRS & Université de Toulon

étude: **OFB-21-1236**

Juin 2022

Remerciements

Ce rapport de recherche bénéficie de l'environnement et certains des algorithmes développés dans le cadre de la Chaire nationale en IA, Bioacoustique ADSIL, co-financée par l'ANR, l'AID et la DGA, PI Glotin, 2021-2025, ANR-20-CHIA-0014-01.

Sommaire

I. Introduction	2
II. Matériel et méthode	4
Présentation du site	4
Analyse des configurations	5
Construction et structure du jeu de données	6
Images fournis au CNN, élaboration des spectrogrammes	7
Annotation et organisation du jeu de données	8
Architecture du modèle	9
Apprentissage du CNN à l'aide des hyperparamètres	10
Prédictions du CNN sur les enregistrements de Guadeloupe	11
III. Résultats	11
Performances du détecteur-classifieur	11
Utilisation du classifieur pour l'identification d'oiseaux	15
Itinéraire Morne Paul Thomas (MPT)	16
Itinéraire Petit Marquisat (PM)	16
Comparaison des itinéraires	17
Projection en temps et espace des distances du percentile 90 à la médiane	17
IV. Discussion	18
Performance	18
Détection	18
Effet de la culmination du soleil sur les chants d'oiseaux	19
Améliorations	20
V. Conclusion	21
ANNEXE	22
ANNEXE 1 : format des fichiers de prédiction (logit)	22
ANNEXE 2 : Schéma simplifié de la création du jeu d'entraînement.	22
ANNEXE 3 : Schéma simplifié de l'utilisation du CNN sur les enregistrements de Guadeloupe.	23
ANNEXE 4 : Transformée de Fourier à court terme (STFT).	23
ANNEXE 5 : Distribution des logits par station.	24
ANNEXE 6 : Résultats de projection en temps et espace des distances du percentile 90 à la médiane.	46
Références	54

I. Introduction

La Liste Rouge des Espèces Menacées de l'UICN (Union Internationale pour la Conservation de la Nature (UICN *et al.* 2020)) met en évidence une accélération de la disparition des espèces par le biais d'indices exprimant le degré de menace pesant sur celles-ci. Cette disparition est donc à l'origine d'une diminution de la biodiversité pouvant, par le biais des réseaux trophiques, engendrer des conséquences sur le fonctionnement des écosystèmes (Dirzo *et al.*, 2014).

La Guadeloupe est un territoire français situé au sein de l'arc des Petites Antilles dans le sud de la mer des Caraïbes. Elle possède ainsi une diversité importante de milieux naturels terrestres et maritimes, accompagnée d'une richesse faunistique et floristique bien caractéristique. La classe des oiseaux possède 60 fois plus d'espèces endémiques dans les territoires d'outre-mer comparé à la métropole (Gargominy, 2003). En ce qui concerne la Guadeloupe et ses dépendances, un total de 258 espèces fin 2007 sont recensées (Levesque *et al.* 2008). L'île ne possède plus qu'une seule espèce strictement endémique : le Pic de la Guadeloupe (*Melanerpes herminieri*) (Levesque *et al.* 2008).

Les espèces endémiques se démarquent par leurs aires de répartition géographique restreintes formant des petites populations (Eraud *et al.* 2013). Ces différentes caractéristiques les soumettent à une fragilité qui les expose davantage aux risques d'extinction, déjà amplifiés dans un contexte d'insularité (Eraud *et al.* 2013; Levesque *et al.* 2008). L'évaluation et l'élaboration de mesures de gestion conservatoires pour ces espèces se heurtent bien souvent à une méconnaissance des abondances et de leur fonctionnement. Pour pallier à ce problème, des protocoles adaptés sont mis en pratique, notamment le Suivi Temporel des Oiseaux Communs par Échantillonnages Ponctuels Simples (STOC-EPS) créé en 1989 (Julliard & Jiguet, 2002) et déployé à partir des années 2000. Sa facilité de mise en œuvre, son efficacité et la fiabilité de ses résultats en font un indicateur structurel de l'Union Européenne pour le développement durable (Dormoy, 2003).

Depuis le développement de la bioacoustique en 1956, les expériences appartenant à ce domaine ont été mises en avant et le Comité International de BioAcoustique (IBAC) a été créé (Agnieszka, 2017). La bioacoustique est une discipline appartenant au domaine de l'éthologie, c'est-à-dire, s'appuyant sur l'étude des communications sonores émises par des espèces animales (Agnieszka, 2017 ; Laiolo 2010). Cette discipline étudie aussi bien les vibrations produites par les insectes que les clics d'écholocation des cachalots (Agnieszka, 2017; Poupard *et al.*, 2022). Ces signaux sonores émis sont définis comme des variations de pression mécanique se propageant sous la forme d'onde dans un fluide grâce à la déformation élastique de celui-ci (Leroy, 1979). L'émission du signal sonore peut être définie par sa fréquence exprimée en hertz, son amplitude exprimée en décibel (dB) et sa durée (secondes, minutes ou heures). Un signal de fréquence important sera plus aigu qu'un signal de fréquence plus faible qui, lui, sera grave. L'amplitude quant à elle, induira un volume élevé pour de fortes valeurs et à contrario, un volume faible pour une faible amplitude (Leroy, 1979).

L'étude du chant des oiseaux n'est pas récente. Il a été le premier trait dit "culturel" (c'est-à-dire acquis par apprentissage social) à être décrit dans la littérature scientifique, en se basant sur des preuves remontant à Aristote (Laland et Galef, 2009). Ces signaux permettent de renseigner les environs de plusieurs informations ; leur identité individuelle, le statut, l'humeur et les intentions (Laiolo 2010; Slabbekoorn et Smith, 2002). Le chant agit donc

comme vecteur de communication. La bioacoustique peut donc, en interceptant ces signaux émis, être utilisée pour réaliser des inventaires et suivis de populations au cœur même de chaque plan de gestion de la biodiversité. Il est primordial d'utiliser une analyse automatisée afin d'analyser la communication des oiseaux dans différents contextes spatiaux et temporels afin d'éviter les erreurs d'origine anthropique. Pour cela, les scientifiques travaillant dans le domaine de la bioacoustique proposent de nombreuses approches utilisant les réseaux de neurones et l'apprentissage profond (Glotin *et al.*, 2013).

Une des premières réalisations fonctionnelles d'un réseau de neurones convolutifs (CNN) est celle de LeCun et al, en 1998 également connue sous le nom de LeNet-5 utilisé dans la reconnaissance d'image. Les CNN sont des systèmes d'intelligence artificielle (IA) qui utilisent un apprentissage profond (deep learning), sous la forme et le fonctionnement des neurones du cerveau humain, pour effectuer des tâches discriminantes et/ou descriptives (Lecun *et al.*, 1998; Nielsen, 2015; Xie *et al.* 2019).

Xie et al., (2019) tente sans succès une séparation des signaux sonores harmoniques et percussifs pour créer deux jeux de données distincts. Ces données sont introduites dans un classifieur CNN en fusion tardive afin de favoriser un apprentissage indépendant puis global. Nous avons exploré avec la librairie LibRosa cette approche qui semblait intéressante pour séparer les bruits orthoptères des oiseaux, mais également sans gain de performance. Ces résultats ne sont pas présentés ici.

Depuis, les CNN ont connu beaucoup de progrès et peuvent dorénavant être utilisés pour de la détection et classification d'espèces animales comme des oiseaux (Nielsen, 2015; Xie *et al.*, 2019, Stowell .. Glotin 2019, Sevilla Glotin 2017, Joly ... Glotin *et al.* 2021).

Ce premier rapport démontre quelles espèces de Guadeloupe ont à ce jour un corpus minimal mais suffisant d'apprentissage pour une reconnaissance automatique, et donne les consignes pour progresser dans l'étude. Il faudra par à moyen terme contraster ce qui a été détecté sur les enregistrements automatiques via une validation partielle humaine. Enfin, sur le long terme, mettre en place un comptage automatique des oiseaux de Guadeloupe en partenariat avec l'OFB.

Pour cela, un réseau de neurones convolutifs de l'état de l'art (CNN), a été construit, appris, pour détecter et classer par espèces les sons d'oiseaux de Guadeloupe qui lui ont été fournis en entrée. Ensuite, pour quantifier les performances du CNN, la fonction d'efficacité du récepteur ou courbe ROC (de l'anglais receiver operating characteristic) a été utilisée (Egan, 1975; Swets *et al.*, 2000; Fawcett, 2006 ; Manning *et al.*, 2008; Pedregosa *et al.*, 2011). Une courbe ROC est une représentation bidimensionnelle des performances d'un classificateur (Egan., 1975; Swets *et al.*, 2000; Fawcett., 2006; Pedregosa *et al.*, 2011). Pour comparer différents modèles, une seule valeur scalaire représentant la performance attendue a été utilisée : l'aire sous la courbe ROC nommée AUC 'Area Under the Curve' (Bradley, 1997; Hanley et McNeil, 1982; Manning *et al.*, 2008; Pedregosa *et al.*, 2011). Puis, un calcul de la Précision Moyenne ou mAP a été effectué pour calculer la précision du classifieur à partir des scores de prédiction (Pedregosa *et al.*, 2011; Manning *et al.*, 2008).

Enfin, les enregistrements de Guadeloupe ont été soumis au CNN, les résultats de classification ont été mis en évidence et discutés.

II. Matériel et méthode

Présentation du site

L'archipel de la Guadeloupe est constitué d'une surface de 1 713 km², regroupant six îles, situées au cœur des Petites Antilles (Figure 1).

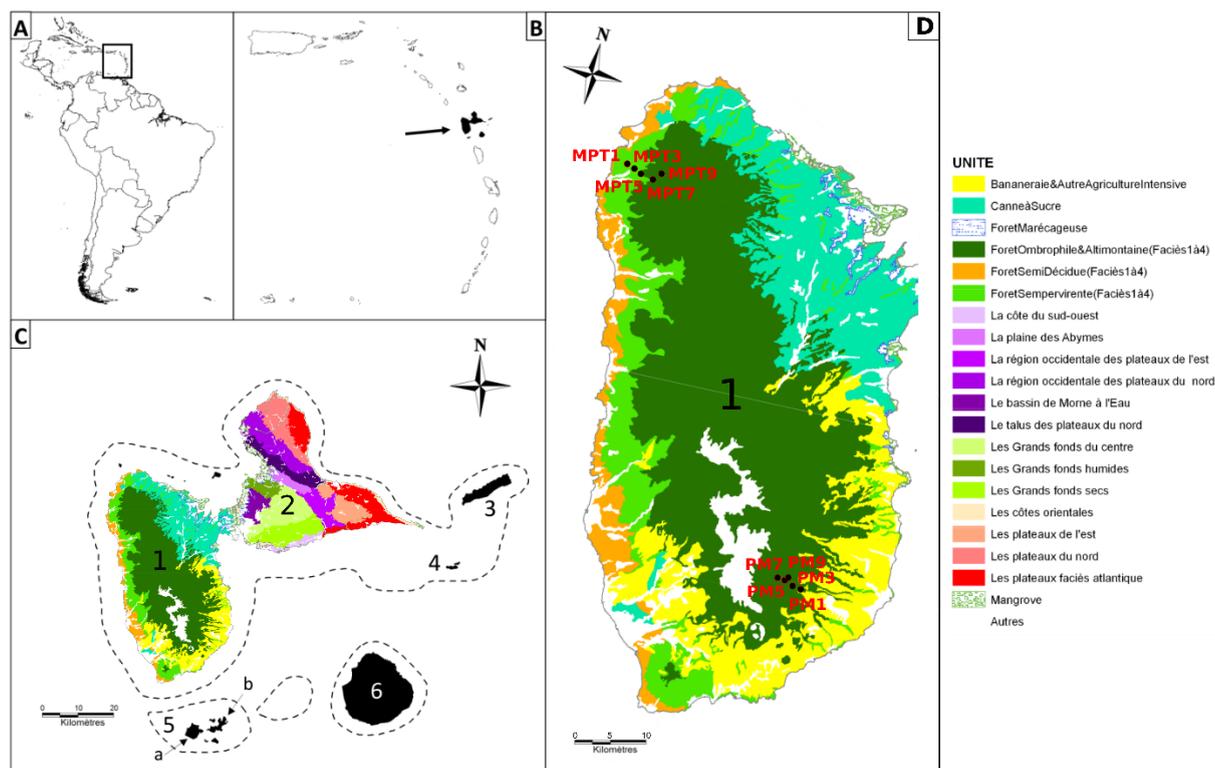


Figure. 1. Cartographie des unités écologiques/paysagères et localisation géographique des Petites Antilles et de l'archipel guadeloupéen. A, Situation des Petites Antilles. B, Situation de l'archipel guadeloupéen dans les Petites Antilles. C, La Guadeloupe continentale est formée par (1) Basse-Terre et (2) Grande-Terre (1520 km²). (3) La Désirade (22 km²), (4) La Petite Terre, (5) Les Saintes [a, Terre-de-Bas ; b, Terre-de-Haut], (6) Marie-Galante (152 km²). D, Itinéraires Morne Paul Thomas (MPT) et Petit Marquisat (PM). (Levesque et al., 2008; Eraud et al., 2013; Massary et al., 2021).

La Grande-Terre et ses dépendances (Marie-Galante, Désirade, Petite-Terre) sont des îles calcaires formant des plateaux pas suffisamment hauts pour bloquer les alizés. La pluviométrie y est faible impactant la nature de la végétation (Levesque *et al.* 2008). Contrairement à La Basse-Terre, montagneuse, entraînant de fortes pluies et qui est composée en grande partie de forêts feuillues (Figure 1)(Levesque *et al.* 2008; Eraud, 2013). Le climat tropical est chaud, humide (hygrométrie de 87% en octobre-novembre) et deux saisons peuvent être mises en évidence. La première, nommée le Carême, sèche, s'étend de décembre à mai où l'ensoleillement y est maximal. La seconde, appelée l'Hivernage, est plus humide, elle s'étale de juin à novembre et se caractérise par un risque cyclonique important. (Levesque *et al.* 2008).

Analyse des configurations

Due à sa topologie, la Guadeloupe n'est pas soumise équitablement aux rayonnements solaires.



Figure. 2. Trajectoire et rayons solaires pour le 16/06/2021 (D'après SunEarthTools).

La figure 2, permet de pouvoir se rendre compte de la trajectoire que prend le soleil en fonction du temps. L'aube commence à 6 heures 35 minutes et le crépuscule débute à 19 heures 40 minutes.

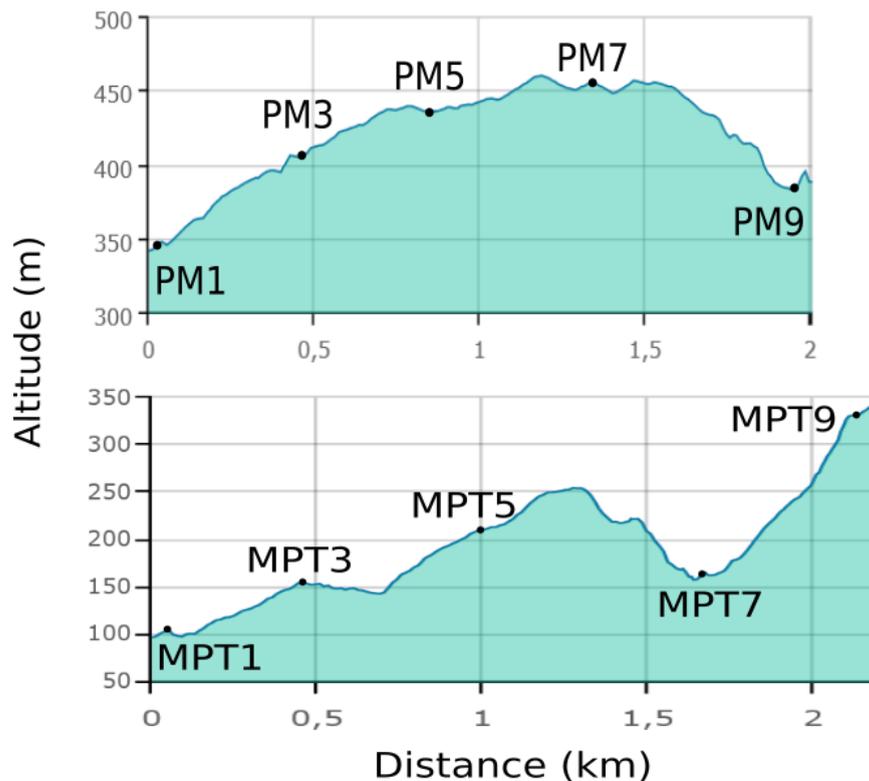


Figure. 3. Profil altimétrique de chaque station (numérotée de 1 à 9) sur l'itinéraire Petit Marquisat (PM) et Morne Paul Thomas (MPT).

Due à la topographie du lieu (figure. 3), pour une station choisie, une zone de faible élévation (altitude) peut ne plus être éclairée par le soleil des heures avant son coucher. De même pour le lever, la zone peut être éclairée bien après le réel levé du soleil, renforçant ou annulant la différence d'exposition au soleil entre les stations.

Sources des données sous la forme d'enregistrements

Les données qui ont été traitées dans cette étude proviennent d'enregistrements réalisés en Guadeloupe par l'Office Français de la Biodiversité, dans le cadre d'une nouvelle approche pour la détection d'oiseaux sur l'île. Ces enregistrements ont été réalisés à l'aide de microphones positionnés en suivant le protocole STOC-EPS (Julliard et Jiguet, 2002). Avec une adaptation du protocole à la topologie de la Guadeloupe. Au lieu de réaliser les observations dans des carrés de 2 kilomètres sur 2 kilomètres, deux itinéraires ont été suivis et nommés : Morne Paul Thomas (MPT) et Petit Marquisat (PM) représentés sur la carte (figure 1) avec leurs différents sites de prélèvements sonores numérotés de 1 à 9. Les enregistreurs ont été posés deux jours avant et laissés 2 jours après l'écoute dans le but de ne créer aucune perturbation anthropique, un réglage en monophonique (1 canal d'amplification) a été utilisé.

La période d'échantillonnage a respecté le protocole STOC-EPS adapté à la Martinique et à l'archipel des Caraïbes (Claessens *et al.*, 2015). Deux sessions d'échantillonnage ont été réalisées sur chaque station au cours de l'unique saison de reproduction aviaire de 2021. Pour la première session : entre le 15 et le 18 avril pour l'itinéraire MPT et du 26 avril au 1er mai pour l'itinéraire PM. Pour le second passage, entre le 3 et le 5 juin pour l'itinéraire MPT puis, entre le 12 et le 17 juin pour l'itinéraire PM. Les plages d'enregistrements ont été de deux heures avant et après le lever du soleil puis, deux heures avant et après le coucher du soleil. Enfin, un enregistrement d'heure est réalisé entre 23 heures et minuit. Ce qui a donné un total de 226 gb de données d'enregistrements.

Le corpus d'apprentissage a été xeno-canto (Vellinga et Planqué, 2015) pour si possible des enregistrements locaux.

Le corpus d'augmentation des données ont été des bruits extraits des enregistrements de l'étude, afin de bruiser les données d'apprentissage par des bruits écologiques.

Construction et structure du jeu de données

Pour débiter, les espèces d'oiseaux endémiques de Guadeloupe, sinon, endémiques des Petites Antilles ont été sélectionnées (Levesque *et al.*, 2008). Elles représentent les espèces pour lesquelles les enjeux de protection sont les plus élevés (Tableau 1) (Levesque *et al.*, 2008; Sullivan *et al.* 2009; UICN *et al.* 2020).

Tableau. 1. Espèces sélectionnées pour l'entraînement du CNN (Levesque et al, 2008; Sullivan et al. 2009; UICN et al. 2020).

Nom vernaculaire	Nom scientifique	Aire de répartition	Critères de l'UICN
Pic de Guadeloupe	<i>Melanerpes herminieri</i>	Guadeloupe	NT
Moqueur grivotte	<i>Alenia fusca</i>	Saba au nord à St Vincent au sud	LC
Paruline caféïette	<i>Dendroica plumbea</i>	Guadeloupe et Dominique	LC
Colibri madère	<i>Eulampis jugularis</i>	Petites Antilles hors Barbade	LC
Colombe rouviolette	<i>Geotrygon montana</i>	Forêts tropicales d'Amérique	LC
Colombe à croissants	<i>Geotrygon mystacea</i>	Nord des Petites Antilles à Ste Lucie au sud	LC
Tyrann Janeau	<i>Myiarchus oberi</i>	Nord des Petites Antilles à Ste Lucie au sud	LC
Pigeon à couronne blanche	<i>Patagioenas leucocephala</i>	Caraïbes	NT
Pigeon à cou rouge	<i>Patagioenas squamosa</i>	Antilles	LC
Saltator gros bec	<i>Saltator albicollis</i>	Guadeloupe, Dominique, Martinique et Sainte-Lucie	LC
Grive à pieds jaunes	<i>Turdus lherminieri</i>	Montserrat, Guadeloupe, Dominique et Sainte-Lucie	VU

Afin d'obtenir des chants de chacune de ces espèces individuellement, des enregistrements sonores de classe A à B (les lettres indiquent la qualité de l'enregistrement) ont été téléchargés à partir de la base de données xeno-canto (Vellinga et Planqué, 2015).

Un algorithme en langage python a été réalisé afin de convertir les échantillons de xeno-canto (Vellinga et Planqué, 2015) en enregistrements mono, d'extension .wav avec une fréquence d'échantillonnage de 24 000 Hz.

Il a été ajouté par un nouvel algorithme deux bruit de fond caractéristique du paysage sonore de Guadeloupe, provenant d'enregistrement fourni par l'OFB. Aux vues de la large plage horaire des enregistrements fournis, deux séries de données ont été créées. Une possédant les échantillons avec un bruit de fond représentatif de la période avant le lever du soleil, et l'autre, possédant les échantillons avec un bruit de fond caractéristique de la période du coucher de soleil et de la nuit (23 heures). Cela avait pour utilité de créer des enregistrements ressemblant le plus aux conditions réelles, sans pour autant être parfaitement identiques entre elles, donc en créant une augmentation des données variables.

Enfin, à l'aide de l'algorithme, chaque fichier des deux séries d'échantillons a été découpé à des intervalles de quatre secondes afin une fois de plus de créer une plus grande quantité de données. Chaque morceau de quatre secondes contient au minimum un chant d'oiseau. Lors du découpage, les morceaux inférieurs à quatre secondes ont été conservés et complétés par le bruit de fond. Cette opération a permis de passer de 46 à 501 fichiers sons au total, dans le cas de 11 espèces cibles.

Images fournis au CNN, élaboration des spectrogrammes

Les images des spectrogrammes ont été utilisées en entrée du réseau de neurones à convolution (CNN). Le spectrogramme est une représentation à deux dimensions (temps-fréquence) mais, à trois paramètres (temps-fréquence-intensité).

L'image résultante était donc bidimensionnelle. Les valeurs numériques de chaque pixel de l'image ont été injectées dans le CNN. Les couleurs d'intensité du spectrogramme ont été converties en gris. Le triple canal couleur ; rouge, vert, bleu est ainsi devenu un canal unique en nuance de gris ou le noir correspond au maximum d'intensité et le blanc au minimum d'intensité du spectrogramme. Le spectrogramme de chaque enregistrement a été obtenu en utilisant la transformée de Fourier à court terme (STFT) (Sejdić *et al.*, 2009), qui a permis de représenter le signal en temps-fréquence en calculant des transformées de Fourier discrètes (TFD) sur de courtes fenêtres qui se chevauchent. Le résultat sous la forme d'un nombre

complexe a été ajouté à une matrice, qui enregistrait l'amplitude et la phase pour chaque point dans le temps et la fréquence (Shentov *et al.*, 1995). Sa définition pour un signal S de N échantillons était :

$$S(K) = \sum_{n=0}^{N-1} S(n) e^{-2i\pi k \frac{n}{N}} \text{ pour } 0 \leq k < N.$$

Le fenêtrage utilisé est de 2048 en type Hann, définit tel que (Blackman *et al.* 1958)

$$w(n) = 0.5 - 0.5 \cos \cos \left(\frac{2\pi n}{M-1} \right) \text{ pour } 0 \leq n < M - 1.$$

Avec M : Nombre de pixels dans la fenêtre de sortie.

Le nombre de lignes de la matrice STFT était donc de 1025 (1 + 2048/2), avec un nombre de colonnes équivalent à 173 relatifs à 4 secondes, ce qui correspondait à un total de 177 325 valeurs d'entrée du modèle. Les données produites sont représentées en nombres complexes. C'est le module (z) des valeurs qui a été utilisé pour créer de nouveaux fichiers de données, définit par :

$$|z| = \sqrt{\text{réel}^2 + \text{imaginaire}^2}.$$

Annotation et organisation du jeu de données

Les fichiers contenant les données des spectrogrammes ont été individuellement attribués à leurs espèces d'oiseaux correspondant. Deux jeux de données ont été créés séparément, le premier correspondant à tous les chants d'oiseaux avec un bruit de fond durant la période matinale. Le deuxième corpus possédait les chants d'oiseaux correspondant à la période relative au coucher de soleil/nuit. Dans un premier temps, c'est sur ce dernier corpus que le CNN classifieur a été entraîné. Les tests du classifieur ont été réalisés sur le corpus des chants d'oiseaux avec le bruit du lever de soleil afin d'en ressortir des métriques de performances, tel que ROC (Receiving Operator Characteristics), AUC (Area under the Curve, intégrale courbe ROC), mAP (mean Average Precision, intégrale courbe rappel précision)

Dans un second temps, un corpus a été créé regroupant les chants d'oiseaux avec un bruit de fond correspondant à la période relative au coucher de soleil/nuit et un bruit de fond sur la période du lever de soleil. Le modèle CNN a été entraîné sur 60 % du jeu de données total créé précédemment. Puis, 40 % du corpus ont été utilisés pour tester le modèle afin d'en ressortir des métriques de performances.

Une mauvaise qualité des enregistrements du jeu de données a été identifiée entraînant de moins bonnes performances. Quatre espèces ont donc été retirées de la classification, à savoir: le Pigeon à cou rouge, le Pigeon à couronne blanche, le Colibri madère et la Grive à pieds jaunes. Les performances du CNN à 11 et à 7 espèces ont été comparées.

Architecture du modèle

Le modèle qui a été utilisé est un réseau de neurones convolutifs (CNN), orienté pour le multi label, c'est-à-dire que chacune des espèces d'oiseaux représente un label du modèle. L'architecture du modèle a été adaptée à celle de Grill *et al.*, (2017), présentée en Tableau 2.

Tableau. 2. Architecture du réseau de neurones convolutifs (CNN) pour un spectrogramme linéaire de 4 secondes en entrée pour 7 labels (7 espèces) inspiré de "Whale Detect" du LIS inspiré lui-même du modèle Sparrow (Paul Best Phd 2022).

Input	16 x 1 x 1025 x 173
Conv2D(3x3)	16 x 32 x 1023 x 171
Conv2D(3x3)	16 x 32 x 1021 x 169
MaxPool(3x3)	16 x 32 x 340 x 56
Conv2D(3x3)	16 x 32 x 338 x 54
Conv2D(3x3)	16 x 32 x 336 x 52
Conv2D(16x3)	16 x 64 x 321 x 50
MaxPool(1x3)	16 x 64 x 321 x 16
Conv2D(1x9)	16 x 256 x 321 x 8
Conv2D(1x1)	16 x 64 x 321 x 8
Conv2D(1x1)	16 x 7 x 1 x 1
MaxPool()	16 x 7 x 1 x 1

A l'exception de la couche de sortie, chaque couche de convolution est suivie d'une "batch normalisation" aux dimensions du kernel pour égaliser le poids de chaque couche (Sergey *et al.*, 2015), normalisation sous la forme :

$$X_{normalisé} = \frac{X-m}{\sigma}.$$

Puis d'un "drop out" qui consiste à désactiver temporairement 1 neurone sur 5 aléatoirement dans le réseau, ainsi que toutes ses connexions entrantes et sortantes (Nitish *et al.*, 2014)(tableau 2) afin d'éviter un possible surapprentissage. Suivi d'une fonction d'activation "Leaky Relu" (Rectified Linear Unit) (avec une pente négative de 0.01) (Maas *et al.*, 2014), définit par:

$$y = \max(0, 0.01*x).$$

Le terme de "fonction d'activation" vient de l'équivalent biologique "potentiel d'activation", seuil de stimulation qui, une fois atteint, entraîne une réponse du neurone.

Le Maxpool (Tableau 2) renvoyait une nouvelle matrice contenant les valeurs maximales aux dimensions du “kernel size” parcouru avec un pas (stride) de 3 en avant 3 en descendant (3x3) pour le premier Maxpool et un pas de 1x3 pour le deuxième, effectué sur l’ensemble des données.

Apprentissage du CNN à l’aide des hyperparamètres

Afin d’obtenir un modèle aux meilleures performances possibles, les hyperparamètres ont été établis empiriquement. L’apprentissage a été réalisé par descente de gradient stochastique (Bottou, 2010) sur des lots (batches) de 16 échantillons (Kandel et Castelli, 2020). En utilisant l’optimiseur ADAM (Kingma et Adam, 2014), avec un taux d’apprentissage (learning rate) de 0.05 (réduit de 1% chaque itération (epoch) de l’entraînement), pour que chaque nouvel entraînement produise un déplacement de l’erreur vers un minimum local ou global du gradient de la fonction de coût (loss), le processus est appelé "descente de gradient" (Snyman, 2005) représenté en figure 2.

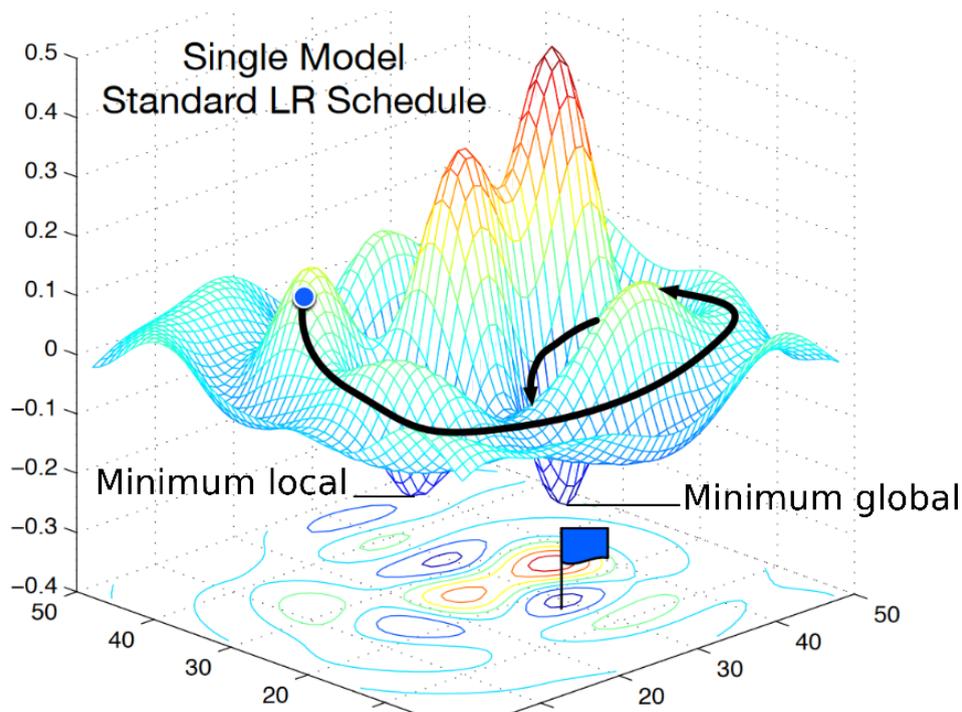


Figure. 4. Illustration 3D de la descente de gradient (D’après Huang et al, 2017)

La fonction loss est égale à la fonction de l’entropie croisée (Binary Cross-Entropy) définie par :

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

Une régularisation L2 de 0.002 (Cortes et al. 2012) a été appliquée afin d'éviter le Surapprentissage. Au total, l'entraînement a été réalisé sur 1000 itérations avant d'être utilisé sur les enregistrements de Guadeloupe. Les poids relatifs à chaque donnée d'image ont été extraits puis stockés dans un fichier.

Prédictions du CNN sur les enregistrements de Guadeloupe

Pour connaître les prédictions du modèle sur les données de Guadeloupe fournies par l'OFB, le CNN chargé d'apprendre a été utilisé en sens inverse. Les poids du jeu de données précédemment calculés à l'aide de l'entraînement ont été introduits dans le modèle. En parallèle, l'algorithme a créé des spectrogrammes de 4 secondes (possédant les mêmes paramètres que ceux créés pour l'entraînement à partir de l'enregistrement fourni). Le CNN associe alors chaque image de spectrogramme à une prédiction d'espèce avec une certitude plus ou moins importante.

Afin d'obtenir des résultats statistiquement valables, cinq entraînements avec les mêmes paramètres d'entrées et même architecture de modèle ont été réalisés. Les résultats ont alors été présentés dans le tableau 3 en calculant la moyenne et l'écart type entre les cinq entraînements.

III. Résultats

Performances du détecteur-classifieur

Les figures 3, 4 et 5 présentent le résultat des ROC et AUC du jeu de données correspondant à la partie test du classifieur. Les tracés colorés représentent les courbes ROC et leurs AUC correspondantes pour chaque espèce entraînée dans le CNN. Les courbes ROC et les AUC pour les 11 espèces d'intérêt maximal ont été calculées et regroupées en figure 3.

La ROC a été représentée sous la forme d'une courbe donnant le taux de vrais positifs ou sensibilité (l'annotation manuelle de l'espèce correspond à la prédiction automatique du modèle) en fonction du taux de faux positifs ou l'antispécificité (1 moins la spécificité) (l'annotation ne correspond pas à la prédiction) (Egan, 1975; Swets *et al.*, 2000; Fawcett, 2006; Manning *et al.*, 2008; Pedregosa *et al.*, 2011). Les taux sont :

$$\text{Taux de vrais positifs (True Positive Rate)} = \frac{\text{nombre de classifications réussites}}{\text{nombre d'échantillons fournis de l'espèce étudiée}}$$

$$\text{Taux de faux positifs (False Positive Rate)} = \frac{\text{nombre de mauvaises classifications}}{\text{nombre d'échantillons ne correspondant pas à l'espèce cible}}$$

L'AUC d'un classificateur est la probabilité que le classificateur classe une instance positive choisie au hasard plus haut qu'une instance négative choisie au hasard. Elle est équivalente au test de Wilcoxon des rangs (Hanley et McNeil, 1982). Plus l'aire est importante, plus la

courbe s'éloigne de la ligne du classificateur aléatoire et se rapproche donc du classificateur idéal (Bradley, 1997; Hanley et McNeil, 1982; Manning *et al.*, 2008; Pedregosa *et al.*, 2011).

Lorsque les courbes de ROC étaient aux coordonnées (0:1) du graphique et que les AUC étaient proches de 1, cela se traduisait par la présence d'aucun faux positif, tous les enregistrements de cette espèce étaient alors bien classés. C'était le cas de la *Paruline caféïette* sur les figures 3,4 et 5.

A contrario, lorsque la courbe ROC était inférieure ou égale à la droite en pointillés noirs scindant les figures en deux par la diagonale, cela correspondait à une classification aléatoire (AUC de 0.5). Ce fut le cas du pigeon à cou rouge, pigeon à couronne blanche, colibri madère et de la grive à pieds jaunes pour le classifieur à 11 espèces (figure 3). La macro-moyenne (macro-average) correspondait à la moyenne des métriques (ROC-AUC) calculées indépendamment pour chaque espèce (traitant ainsi toutes les espèces de manière égale (Manning *et al.*, 2008)). La micro-moyenne (micro-average) regroupait toutes les données de toutes les espèces pour ensuite calculer les métriques moyennes. Pour le classifieur à 11 espèces, la macro-moyenne (AUC de 0.64) était inférieure à la micro (AUC de 0.79) ce qui s'est traduit par un déséquilibre entre le nombre d'échantillons d'espèces (Manning *et al.*, 2008).

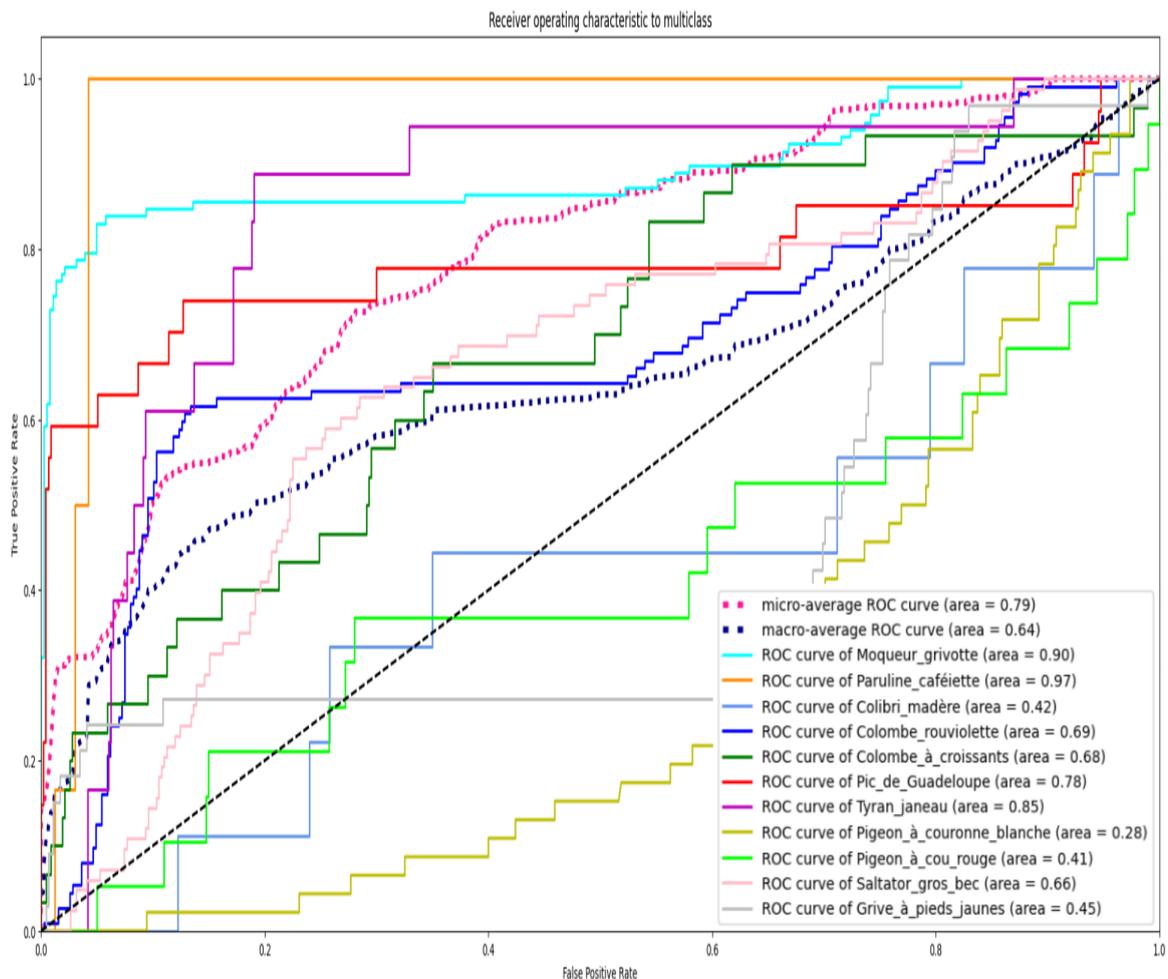


Figure 5. Fonction d'efficacité du récepteur (ROC) dans le cas d'un classifieur multi-label à 11 espèces (200 itérations).

Dans le cas du classifieur à 7 espèces (Figure 4), lors de la 50ème itération, les classifications de la Colombe à croissant et de la Colombe rouviolette ont pu être considérées comme aléatoire (AUC respectives de 0.45 et 0.53). Mais les classifications se sont pour la plupart réalisées correctement, ce qui s'est traduit par une AUC de la micro-moyenne égale à 0.8.

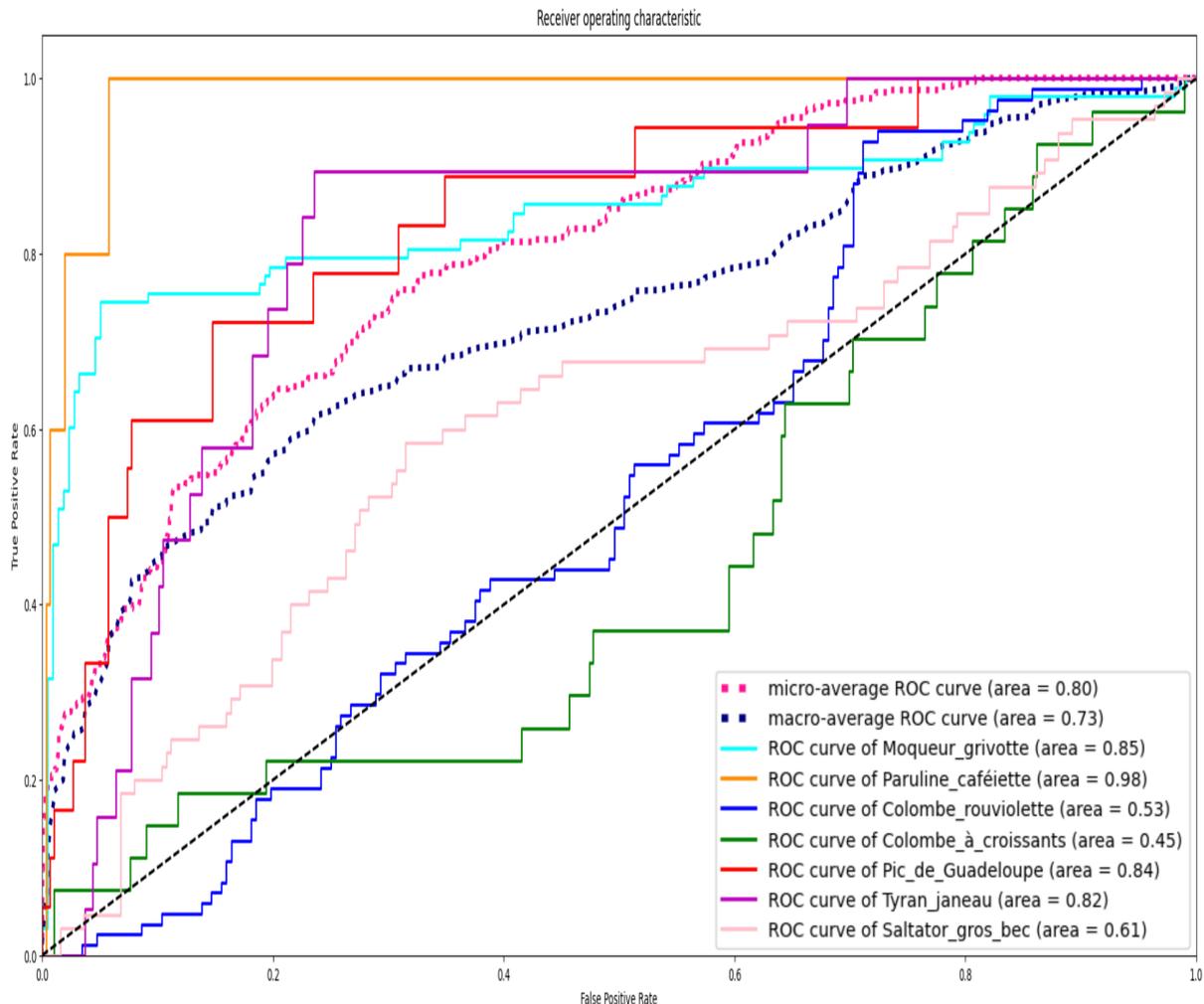


Figure. 6. Fonction d'efficacité du récepteur (ROC) dans le cas d'un classifieur multi-label à 7 espèces (50 itérations).

En revanche, au bout de la 900^{ème} itération d'entraînement (Figure 5), le classifieur a pu être considéré comme valide pour passer à l'étape supérieure de détection. En effet, les courbes ROC se rapprochaient énormément des coordonnées (0:1) du graphique. Les AUC quant à elles étaient de 0.94 pour la plus faible chez le Moqueur grivotte; et les plus hautes, ont été retrouvées chez la Colombe rouviolette et la Paruline caféïette pour une valeur de 0.99. Les micro et macro moyennes qui en découlait étaient donc logiquement aussi élevées (AUC > 0.96).

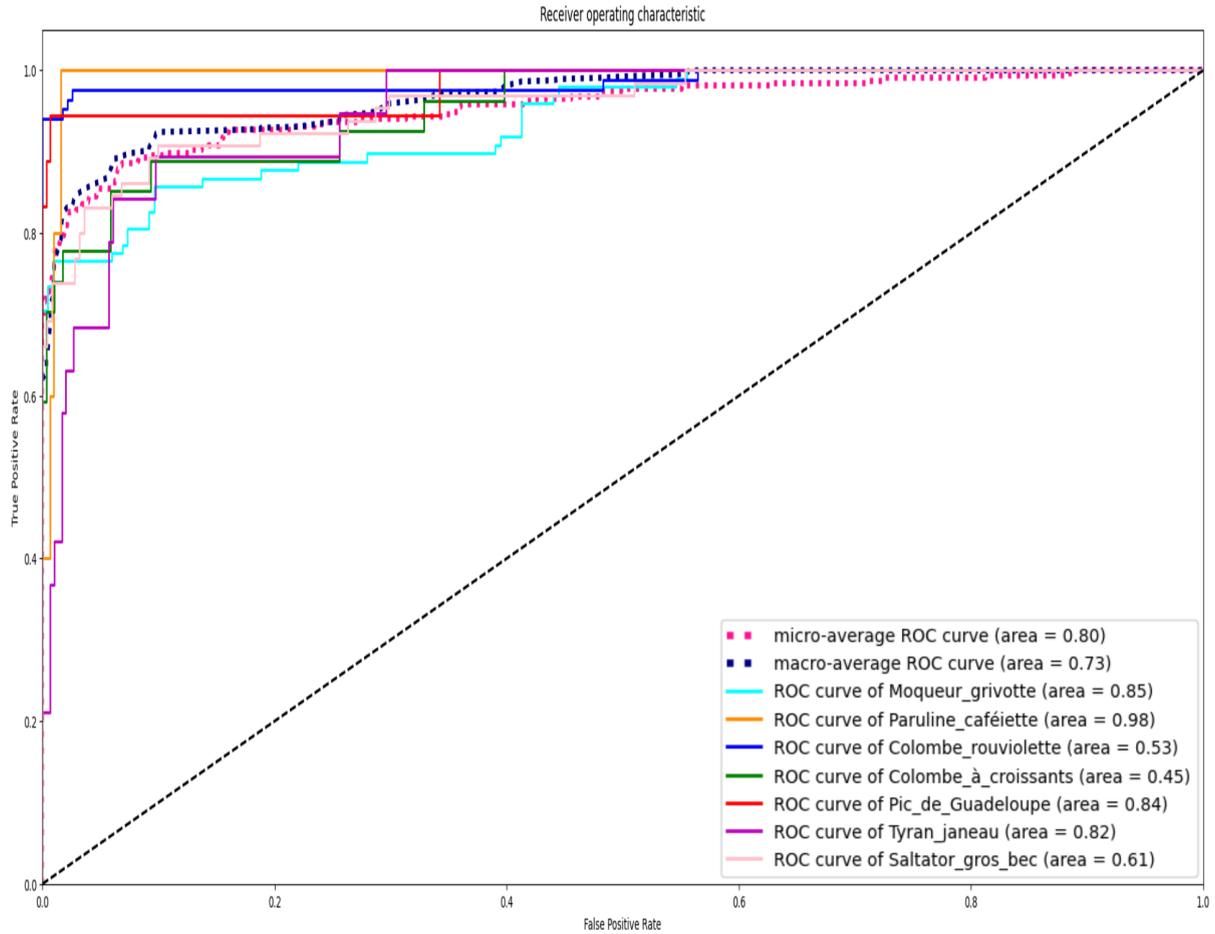


Figure 7. Fonction d'efficacité du récepteur (ROC) dans le cas d'un classifieur multi-label à 7 espèces (900 itérations).

La Précision Moyenne ou mAP a été utilisée pour calculer la précision du classifieur à partir des scores de prédiction (Pedregosa *et al.*, 2011; Manning *et al.*, 2008). Pour obtenir la mAP, les valeurs de précision et de rappel ont été exploitées, définies par :

$$\text{Précision} = \frac{\text{Vrai Positif}}{\text{Vrai Positif} + \text{Faux Positif}}$$

$$\text{Rappel} = \frac{\text{Vrai Positif}}{\text{Vrai Positif} + \text{Faux Négatif}}$$

Par la suite, la moyenne pondérée des précisions de chaque label (AP) a été calculée indépendamment d'après la formule :

$$\text{AP} = \sum_n (R_n - R_{n-1}) P_n$$

Où P_n et R_n correspondaient à la précision et au rappel au $n^{\text{ième}}$ seuil.

La mAP obtenue était la moyenne de la moyenne pondérée des précisions de chaque label (espèce). La mAP était donc une valeur qui permettait de tenir compte de l'exactitude (précision) par rapport au nombre de prédictions de labels (rappel) défini par :

$$mAP = \frac{\sum_{q=1}^Q AP_{(q)}}{Q},$$

où Q est le nombre de label (espèces).

Les métriques des 2 classifieurs ont été regroupées dans le tableau 2. Le CNN entraîné avec 7 espèces a obtenu une fonction de coût (loss) six fois inférieur à 11 espèces d'oiseaux (Tableau 2).

Cela s'est reporté sur les valeurs de mAP et d'AUC. La précision était meilleure d'un facteur 3.5 pour le classifieur à 7 espèces (mAP de 0.8) comparé au classifieur à 11 espèces (mAP de 0.25).

Le fait d'entraîner 60 % puis de tester 40 % des enregistrements du corpus global a entraîné une augmentation de mAP et de l'AUC de 10 % sur le jeu de données testé par rapport à la méthode de séparation des corpus. Les faibles valeurs d'écart-types (Tableau 2) sont la résultante d'une régularité du classifieur et de paramètres stables.

Tableau. 2. Performances des deux classifieurs sur le jeu de test.

	loss	mAP	AUC
7 espèces	0.05 ± 0.01	0.80 ± 0.05	0.95 ± 0.02
11 espèces	0.31 ± 0.008	0.25 ± 0.02	0.64 ± 0.009

D'après les résultats du tableau 2, c'est donc le classifieur avec 7 espèces (Figure 5, Tableau 2) qui a très bien fonctionné et possède des performances acceptables pour pouvoir être utilisé dans la détection et la classification des oiseaux endémiques de Guadeloupe.

Utilisation du classifieur pour l'identification d'oiseaux

Les détections et classifications réalisées par le modèle CNN ont été représentées sous la forme de nuages de points attribués à chaque espèce. Lorsqu'un chant est reconnu, pour toutes les stations sur les itinéraires Morne Paul Thomas (MPT) en figure 6 et Petit Marquisat (PM) en figure 7. L'axe des ordonnées représente les espèces d'intérêt, l'axe des abscisses correspond à la confiance pouvant être attribuée à la classification quantifiée (plus cette valeur est haute moins le CNN a de risque de se tromper sur l'identification de l'espèce). L'ajout de bruit sur le jeu de données fournit au modèle entraîne une réduction des performances de celui-ci durant les entraînements mais augmente la capacité à pouvoir détecter les oiseaux dans les conditions réelles (Grill *et al*, 2017).

Itinéraire Morne Paul Thomas (MPT)

Plusieurs espèces d'oiseaux sont identifiées avec une confiance élevée sur chaque site de l'itinéraire. Notamment, le Saltator gros bec, Tyran janeau, Pic de Guadeloupe, Colombe rouviolette et Moqueur grivotte. Ce dernier est très présent et bien reconnu par le classifieur qui lui confère de très bons scores de confiance allant au-delà de 500 points.

A l'inverse, des espèces ne semblent pas avoir été présentes lors du recensement acoustique, c'est le cas de la colombe à croissant et de la Paruline caféïette.

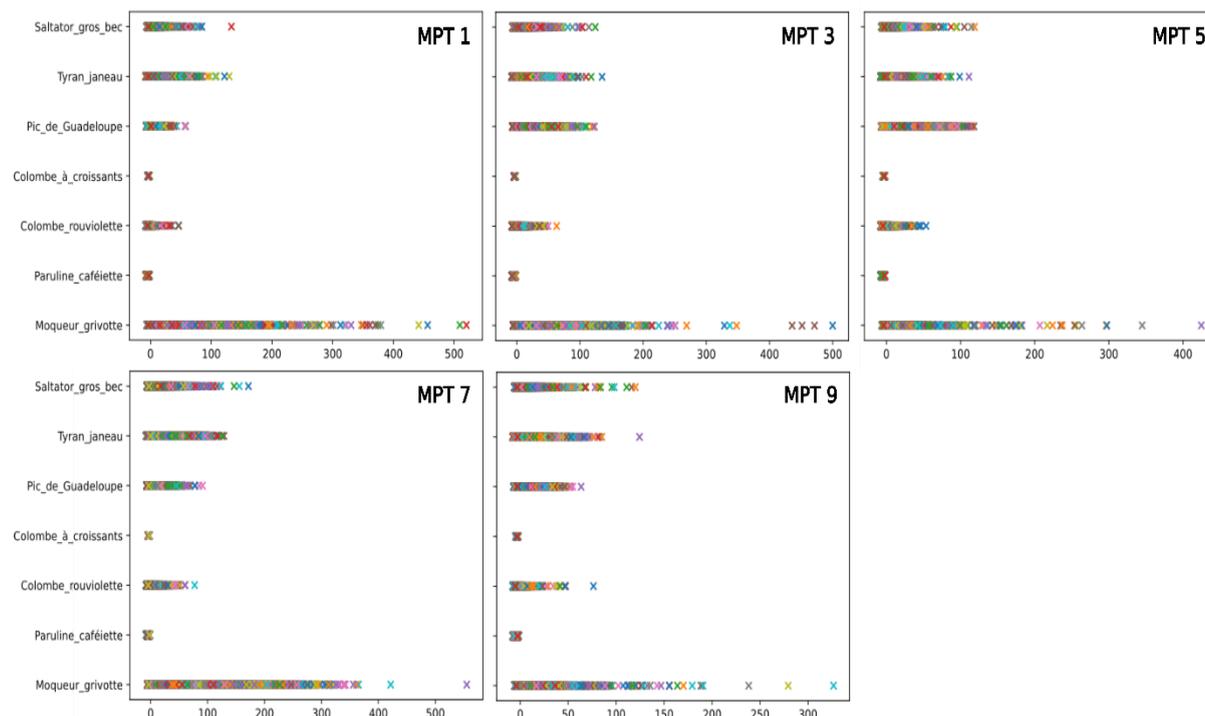


Figure. 8. Détection automatique de chants d'oiseaux par le CNN sur l'itinéraire Morne Paul Thomas. Sur l'axe des abscisses est représentée la confiance que le classifieur attribue à la prédiction.

Itinéraire Petit Marquisat (PM)

De même que sur l'itinéraire MPT, plusieurs espèces d'oiseaux ont été identifiées avec une confiance élevée sur chaque site de l'itinéraire. Les oiseaux détectés sur l'itinéraire PM étaient les mêmes que ceux de MPT, ce qui est surprenant. Les espèces qui ne semblaient pas avoir été présentes sur l'itinéraire étaient logiquement les mêmes que sur l'itinéraire MPT. Le Moqueur grivotte était encore présent, chantant énormément avec une classification d'un maximum de confiance pouvant atteindre les 600.

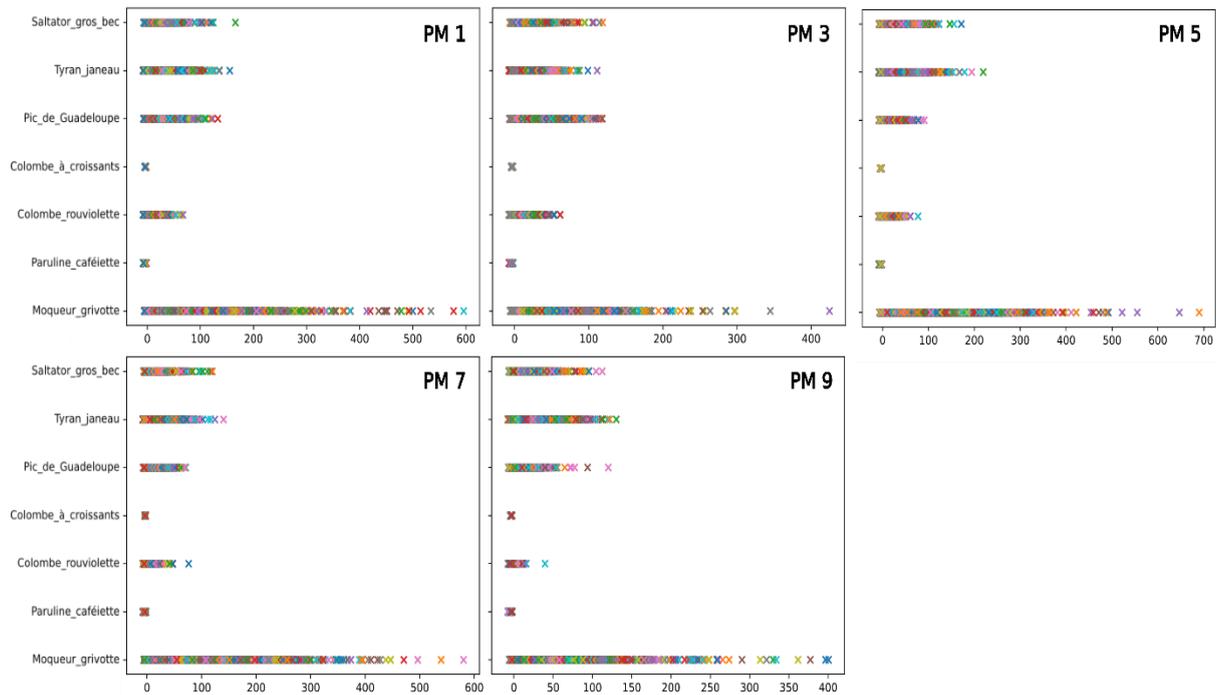


Figure. 9. Détection automatique de chants d'oiseaux par le CNN sur l'itinéraire Petit Marquisat. Sur l'axe des abscisses est représenté la confiance que le classifieur attribue à la prédiction.

Comparaison des itinéraires

Finalement, les deux itinéraires bien que situés aux antipodes de l'île (Figure 1) étaient d'après les résultats obtenus (Figure 6 et 7) très semblables considérant la présence d'espèces. Les espèces de Colombe à croissant et de Paruline caféïette n'ont été recensées sur aucun des deux itinéraires.

Le classifieur n'a pas été entraîné à faire la différence entre deux chants d'oiseaux identiques, chaque chant a donc été comptabilisé comme un individu en plus. Le Moqueur grivotte était donc d'après ces résultats un oiseau qui a énormément chanté pendant les périodes d'enregistrements.

Les spectrogrammes des classifications avec une confiance supérieure à 100 ont été comparés aux enregistrements fournis de xeno-canto (Vellinga et Planqué, 2015). Ceux-ci semblaient correspondre mais il faudrait la validation d'un expert pour pouvoir affirmer ou réfuter l'exactitude des classifications.

Projection en temps et espace des distances du percentile 90 à la médiane

Par la suite, la distance D est donnée, en logarithme, de la différence du percentile 90 de l'exponentielle des prédictions avec la médiane.

$$D = \log((\text{percentile90}(e^{\text{prédiction}}) - \text{médiane}(e^{\text{prédiction}}))).$$

L'exponentielle ici permet de contraster plus fortement les grandes différences des prédictions, et donc de mesurer un plus fort gradient. Dans cette partie, les D mesurées ont été synthétisés par tranches de 5 minutes (300 estimations en 5 minutes), pour chaque espèce.

Les graphiques ont été regroupés en annexe 6. Ils représentent les activités D, par 24h, pour toutes les espèces sur les 10 stations, et sur toutes les sessions (abscisse l'heure de 0 à 24h, en ordonnée cette distance D).

Cela a permis d'observer que la Grive était présente sur MPT1 ;

Le Saltator sur MPT7, PM7, PM1 ;

Le Tyran sur MP7 et PM7 ;

Le Pic sur MPT1.

IV. Discussion

Performance

Chaque étape de prétraitement de l'enregistrement pouvait avoir une importance cruciale sur l'entraînement et donc les prédictions du CNN (Grill *et al.*, 2017). Notamment l'algorithme utilisé pour le rééchantillonnage du signal audio à 24 000 Hz qui réduisait le nombre d'échantillons sonores, et donc la quantité d'information du signal (Grill *et al.*, 2017). Cela a pu être responsable d'une dégradation des performances de classification des oiseaux, entraînant forcément une diminution des performances du modèle. Plus les données testées étaient proches du jeu de données d'entraînement, meilleures étaient les performances du modèle mais cela ne veut pas forcément dire que le modèle a mieux fonctionné sur les données des conditions réelles de Guadeloupe.

Le fait que les espèces endémiques possédaient une faible répartition géographique a induit un jeu de données acoustiques limitées et restreintes qui ne pouvait pas être amplifié facilement (Eraud *et al.* 2013). Le choix de la quantité à défaut parfois de la qualité a été réalisé.

Détection

Le CNN a été entraîné avec des enregistrements possédant la plupart du temps uniquement l'oiseau d'intérêt. Or, les enregistrements fournis par l'OFB contenaient, malgré le découpage sonore effectué, plusieurs oiseaux qui chantaient en même temps. Le classifieur attribuait ce moment à une seule espèce, ce qui a eu pour conséquence une légère sous-estimation de la présence d'espèces. Or, d'après les observations des enregistrements (Vellinga et Planqué, 2015), les oiseaux étudiés n'émettent jamais qu'un seul chant. Ce sont la plupart du temps des séries de signaux à hautes fréquence (Slabbekoorn et Smith, 2002), alors même si un chant

était superposé à un autre à un certain temps, il est forcément arrivé un moment où un nouveau chant émis a été détecté par le CNN.

D'après les résultats des détections (Figure 6 et 7), il était possible que le Moqueur grivotte ait un chant qui ressemble beaucoup à d'autres oiseaux présents sur l'archipel. Ces autres espèces n'ayant pas été incluses dans le classifieur ne peuvent pas être attribuées à la bonne classe et donc sont considérées par défaut comme un Moqueur grivotte.

Le modèle n'était pas assez précis sur les données de Guadeloupe pour en sortir des données d'abondances statistiquement viables. Le classifieur permettait de compter non pas le nombre d'espèces, mais le nombre d'émissions sonores de l'espèce. Le CNN a donc pu être utilisé dans le but de pouvoir émettre une possible présence ou non d'au moins une espèce d'oiseaux appartenant au jeu de données du classifieur. Pour remédier à ce problème un algorithme pourrait être créé afin de reconnaître la signature individuelle de chaque espèce ce qui permettrait de ne compter qu'une fois une occurrence identique répétée.

Bien que certaines espèces d'oiseaux pouvaient se déplacer sur les différentes îles des Caraïbes, il a été observé une divergence génétique entre oiseaux de même espèce au sein de deux zones géographiques de la Guadeloupe (Arnoux *et al.* 2014), le fait que le modèle ait été entraîné en partie avec des enregistrements d'oiseaux provenant d'une autre île des Antilles pourrait expliquer une possible différence de chant entre cette espèce et l'oiseau de la même espèce mais de Guadeloupe. De même concernant les habitats dans les forêts tropicales, qui pourraient avoir un impact sur la différenciation génétique des populations aviaires (Arnoux *et al.* 2014). L'hypothèse serait que certains oiseaux peuvent se déplacer entre les différentes îles de l'archipel des Caraïbes mais, possédant des caractéristiques propres à leurs habitats naturels d'origine conservé pour l'appartenance à une population (Laiolo 2010; Slabbekoorn et Smith, 2002).

Selon la nature de la communication de l'oiseau, le son émis n'aurait pas la même représentation sur le spectrogramme même pour des oiseaux appartenant à la même population (Fowler, 1896; Slabbekoorn et Smith, 2002). Alors, le classifieur ayant uniquement été entraîné à attribuer un chant donné à une espèce d'oiseau, il ne pouvait pas reconnaître un chant de communication différent bien qu'émis par la même espèce.

Pour augmenter les précisions du modèle sur les enregistrements de Guadeloupe, il serait intéressant d'ajouter à la liste d'enregistrements annotés les classifications avec une forte confiance vérifiée manuellement (Xie *et al.*, 2019). Dans le but d'entraîner le modèle avec des données qui correspondent mieux aux conditions réelles de Guadeloupe.

Effet de la culmination du soleil sur les chants d'oiseaux

D'après les résultats présentés par les boxplots en annexe 5, pour la plupart des oiseaux une baisse de l'activation neuronale entre 4h et 6h du matin est observée hormis pour le Saltator gros bec et le Pic de Guadeloupe. Il est quand même important de noter que plusieurs horaires ne sont pas représentés dans les données entraînant un léger biais d'interprétation.

Concernant les courbes des percentils en annexe 5, pour la majorité des espèces, le pic de chant devrait se situer après le lever du soleil et parfois en fin de journée (dans une moindre mesure). Or, il n'y a que la grive et le moqueur corossol qui a priori chantent surtout la nuit.

Les résultats résumés en annexe 6 ont permis de mettre en évidence l'effet du levant et couchant (respectivement 6 heures et 19 heures) sur le paysage acoustique. De plus, leurs maxima peuvent être considérés comme des vrais positifs. En effet, ce sont des valeurs fortes du détecteur de l'espèce qui se séparent du bruit de fond dans les 5 minutes et pour la station considérées.

Il apparaît aussi des motifs croissants et/ou descendants, différents suivant les stations, ce qui indique une activité aviaire et/ou du bruit de fond que l'on peut alors mieux ségréger. Il en ressort les listes des détections les plus probables, à comparer avec des ornithologues, pour confirmer ou infirmer ces hypothèses.

Améliorations

La transformée de Fourier à court terme utilisé est très efficace pour des signaux périodiques, mais elle peut s'avérer inefficace lorsque le signal dont le spectre des fréquences varie dans le temps (Hubbard, 1995), comme c'est le cas pour le chant de beaucoup d'oiseaux (Hubbard, 1995). Cela entraîne une perte de données précieuse pour l'entraînement du CNN. Pour pallier ce problème, la transformée en ondelettes pourrait apporter plus d'informations (Hubbard, 1995).

Une corrélation des résultats des ROC et AUC de la figure 5 et de la détection des espèces de Guadeloupe (Figure 6 et 7) peut être mise en avant. D'après l'allure des courbes ROC et les hautes valeurs d'AUC (0.99) pour les espèces de Colombe à croissant et de Paruline caféïette corrélé à la non-détection de ces espèces par le classifieur pourrait être un signe de surapprentissage (overfitting).

Cela arrive lorsque le CNN apprend par cœur des détails trop spécifiques rendant ses performances meilleures sur les tests mais contre performant lors de son utilisation (Schittenkopf et al., 1997). Dans ce cas précis, la faible quantité de données fournies couplée aux nombreuses itérations d'entraînements ont probablement permis au CNN de se concentrer sur des détails trop spécifiques des chants d'oiseaux fournis qui ne se retrouvent pas sur les enregistrements de Guadeloupe.

Cela pourrait être résolu en augmentant le jeu de données des deux espèces, favorisant la variabilité acoustique (Schittenkopf et al., 1997). En stoppant l'entraînement du CNN plus tôt ou encore en réalisant un bruit de fond qui varie et se rapproche plus des conditions sonores Guadeloupéennes ajouté aux enregistrements dans les étapes d'entraînement du classifieur (Schittenkopf et al., 1997).

Si ce surapprentissage a pu avoir lieu sur deux espèces, cela pourrait être le cas d'autres espèces entraînées. Donc, le fait que les espèces aient quand même été identifiées pourrait indiquer que celles-ci possèdent un chant qui ne varie pas énormément selon les régions des Antilles ou elles se trouvent.

Dans la continuité de ce recensement, il serait intéressant d'ajouter d'autres espèces endémiques des Petites Antilles tel que ; le Martinet chiquesol (*Chaetura martinica*), le Colibri felle-vert (*Eulampis holosericeus*), le Colibri huppé (*Orthorhyncus cristatus*), la Moucherolle gobemouche (*Contopus latirostris*) ou encore le Moqueur corossol (*Margarops fuscatus*).

V. Conclusion

En définitive, les avantages que procure l'identification bioacoustique par intelligence artificielle sont grands, car d'un faible coût, et à la capacité à être anodine pour la faune et la flore et le gain de temps par rapport aux méthodes de surveillance traditionnelles (Wimmer et al, 2013). Il existe néanmoins des inconvénients comme montrés dans ce rapport : la sensibilité du classifieur aux bruits parasites ainsi que la grande variabilité des chants d'oiseaux intra espèces.

L'inconvénient le plus important dans cette étude est cependant le manque d'enregistrements annotés des espèces cibles, qu'il s'agisse d'enregistrements de Guadeloupe ou non. Ces enregistrements sont nécessaires à l'entraînement du CNN. Ce problème majeur n'a été que partiellement résolu. Le premier modèle entraîné avec 11 espèces n'avait pas une performance suffisante.

En revanche, avec 7 espèces cibles, et après ajustement et augmentation durant l'apprentissage par ajout de conditions adverses avec du bruit endémique (sons enregistrés par les stations ne contenant pas d'oiseaux mais du vent, de la pluie ou des anours ou orthoptères, à différents horaires...) le classifieur obtient des performances plus que convenables pour chacune des espèces d'intérêt (80% de MAP).

C'est donc ce dernier qui a été utilisé en conditions réelles pour identifier les oiseaux endémiques d'intérêts de Guadeloupe. Mais, malgré les paramètres ajoutés au modèle afin d'éviter le surapprentissage, il reste 2 espèces : Colombe rouviolette, et la Paruline cafeiete qui ont été surapprises et qui donc ont des scores qui ne se généralisent pas sur les tests. Il reste pour l'instant donc 5 espèces: Moqueur g., Colombe à c., Pic de g. Tyran j., Salatator g..

D'après les résultats de cette étude, l'archipel Guadeloupéen et plus particulièrement la Basse-Terre recense bien les espèces endémiques étudiées à savoir, le Saltator gros bec, Tyran janeau, Pic de Guadeloupe, Colombe à croissant, et Moqueur grivotte. Avec une dominance acoustique des enregistrements par ce dernier.

Il a été possible en cumulant les détections, de trouver des patterns du bruit de fond versus de vraies activités pour pouvoir mettre en évidence ou non la présence d'espèces d'oiseaux d'intérêts tel que la Grive à pieds jaunes. De plus, cette approche renforce l'intérêt de stations distribuées en gradient d'altitude et horaires du levant et du couchant.

Le meilleur moyen d'améliorer ce classifieur sera d'augmenter la quantité d'enregistrements annotés par des experts locaux (OFB?) d'oiseaux à détecter. Pour cela, il serait intéressant d'annoter manuellement un nombre conséquent de classification réussies dans le but de pouvoir les ajouter au jeu de données d'entraînement et ainsi étendre la gamme sonore de chaque espèce.

ANNEXE

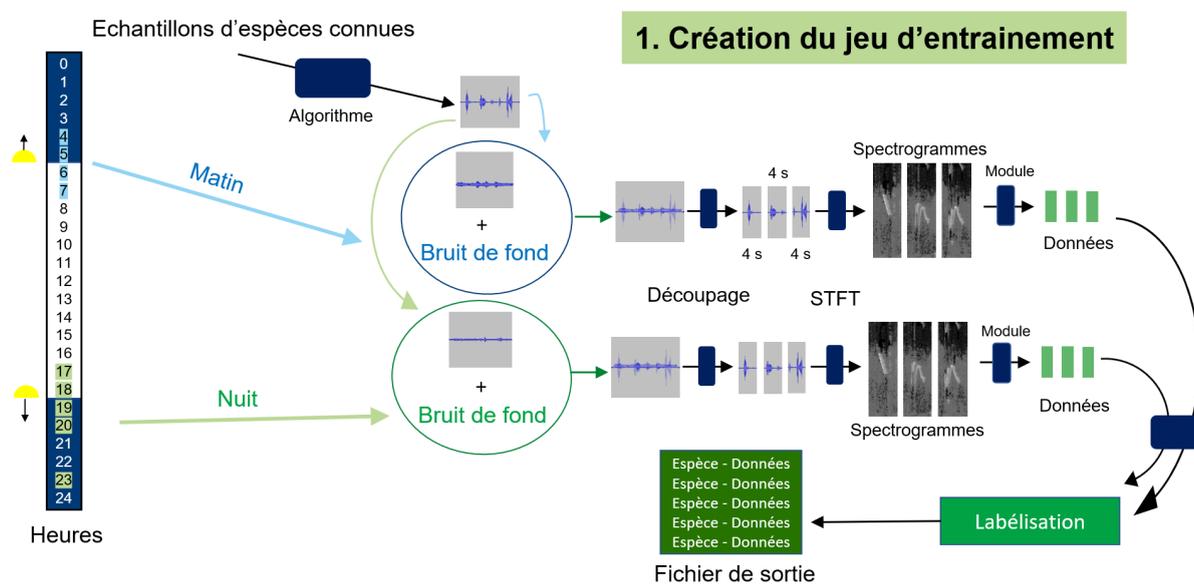
ANNEXE 1 : format des fichiers de prédiction (logit)

les exécutions du modèle sur les 260 Go (les 6 stations sont données en tables CSV =

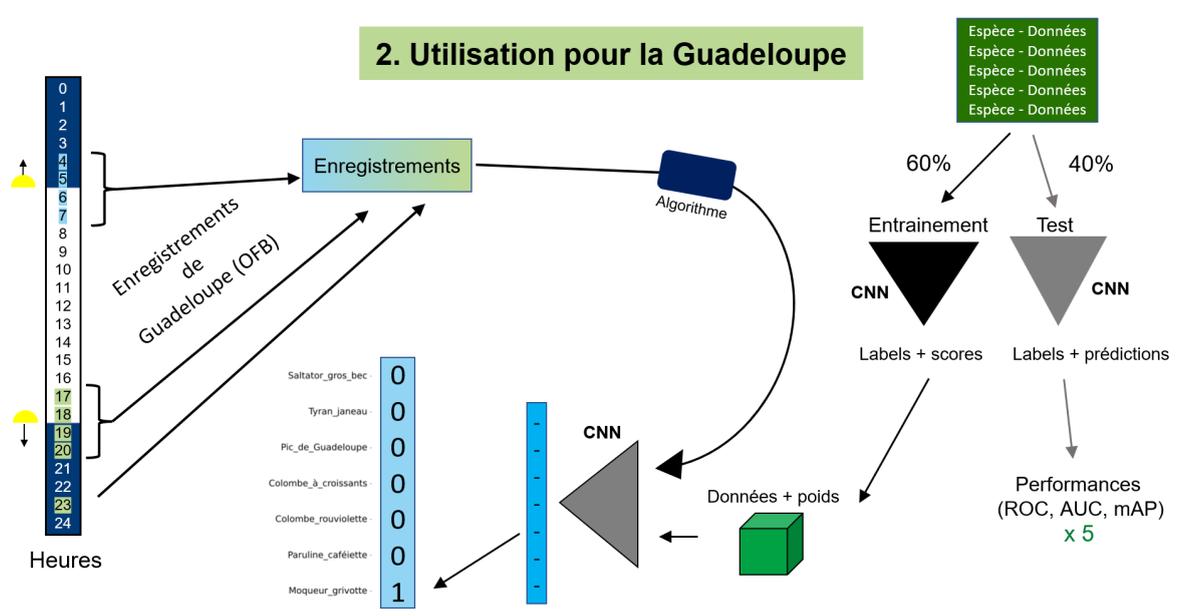
http://sabiiod.lis-lab.fr/pub/BIRD_GUADELOUPE/forwards/csv/

ainsi que les codes informatiques pour assurer la réplicabilité des calculs.

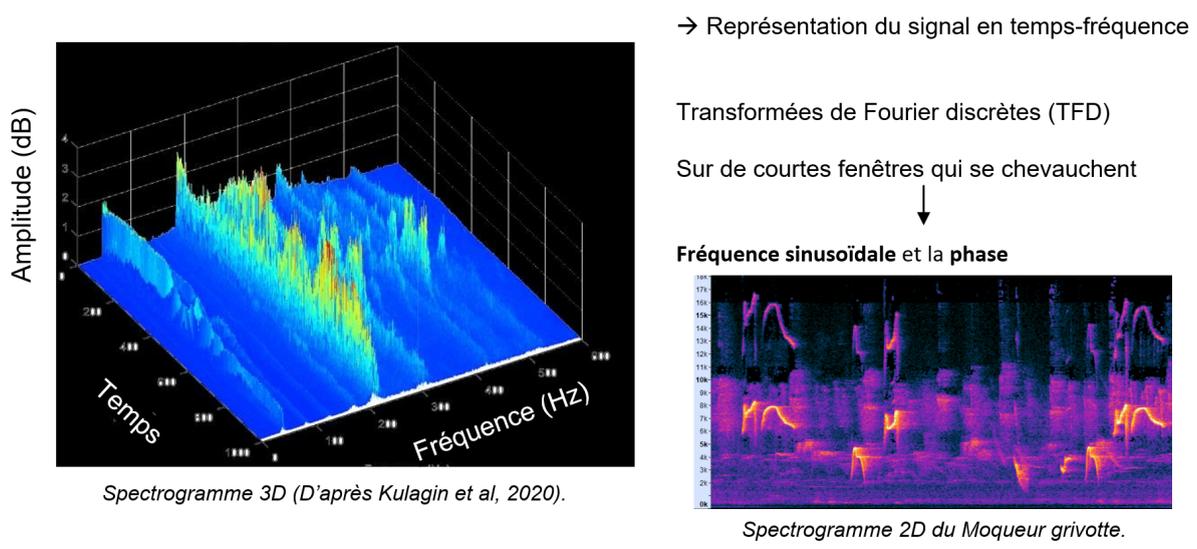
ANNEXE 2 : Schéma simplifié de la création du jeu d'entraînement.



ANNEXE 3 : Schéma simplifié de l'utilisation du CNN sur les enregistrements de Guadeloupe.

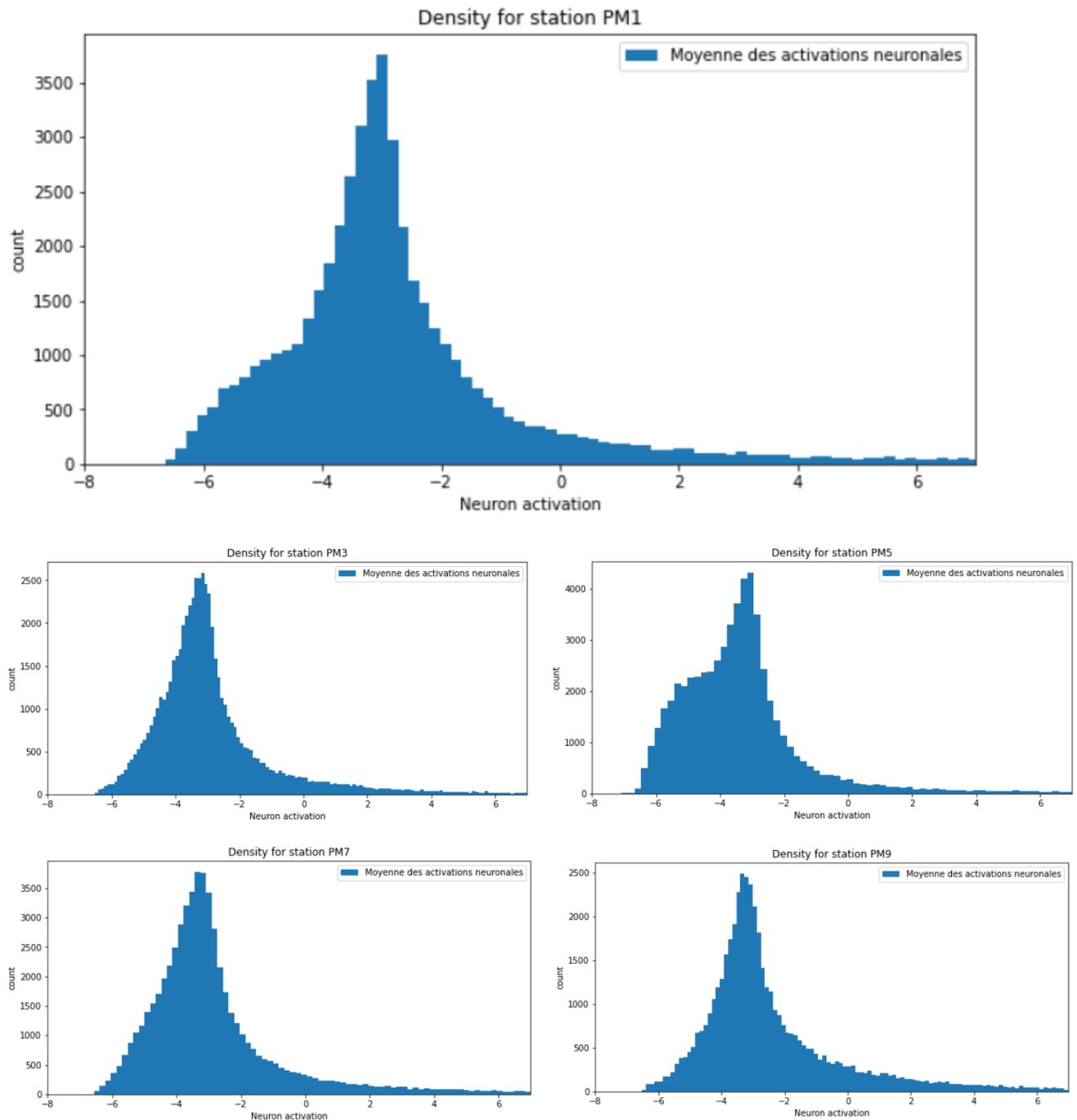


ANNEXE 4 : Transformée de Fourier à court terme (STFT).

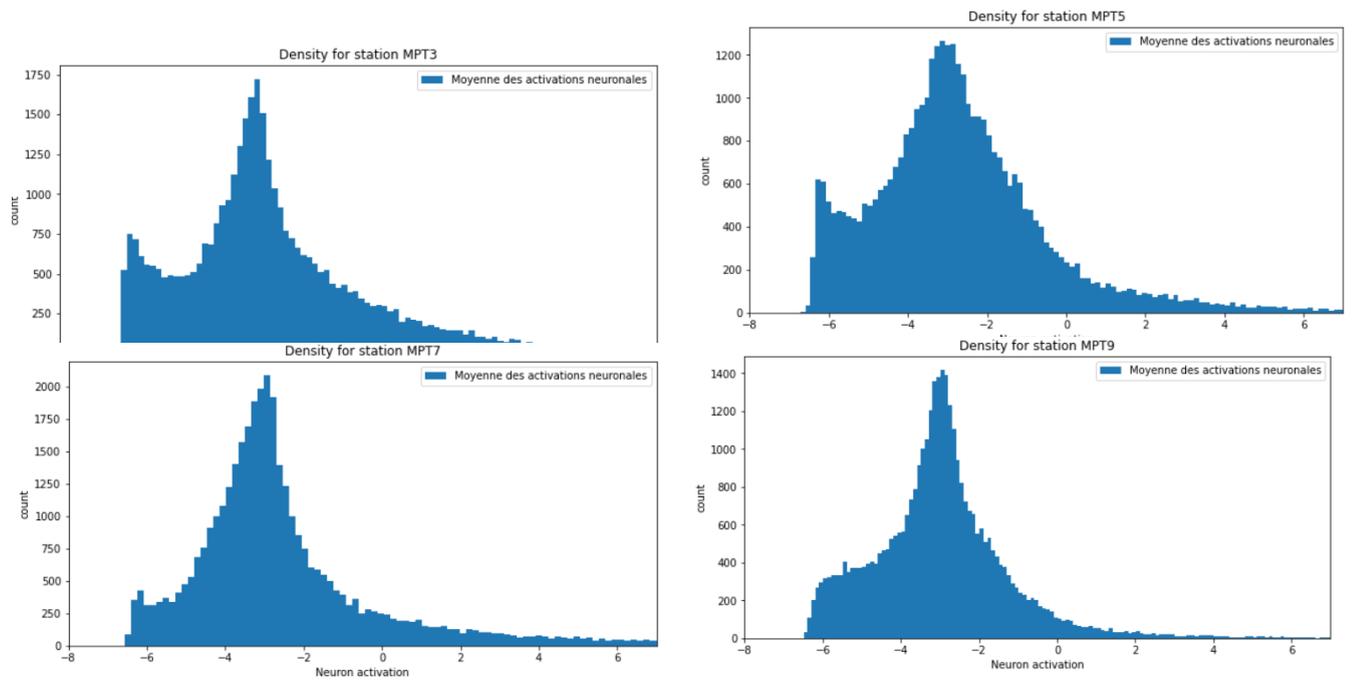
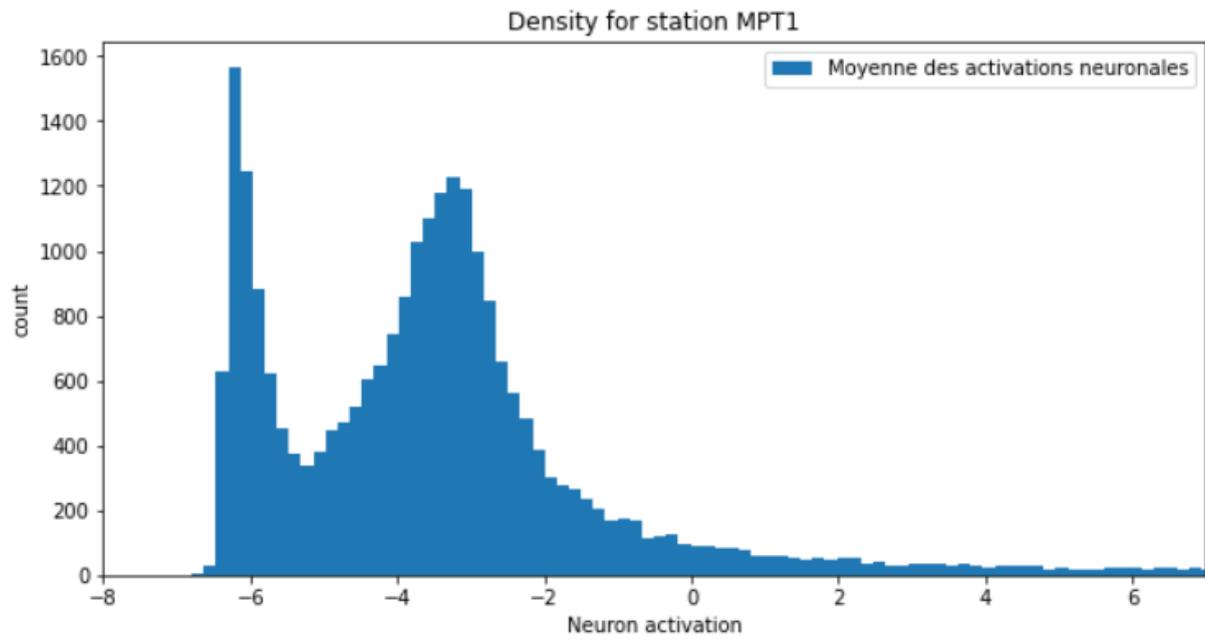


ANNEXE 5 : Distribution des logits par station.

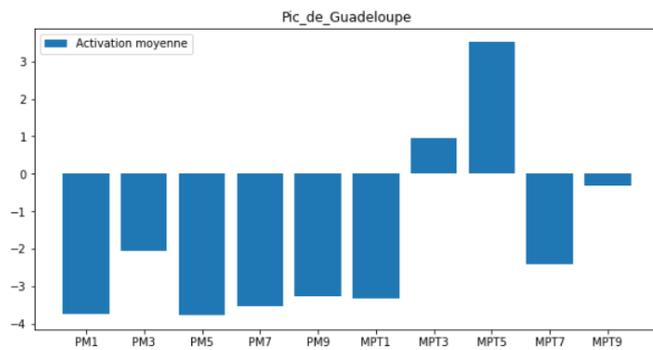
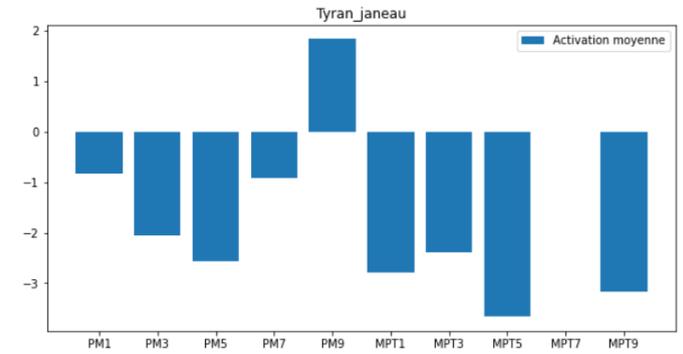
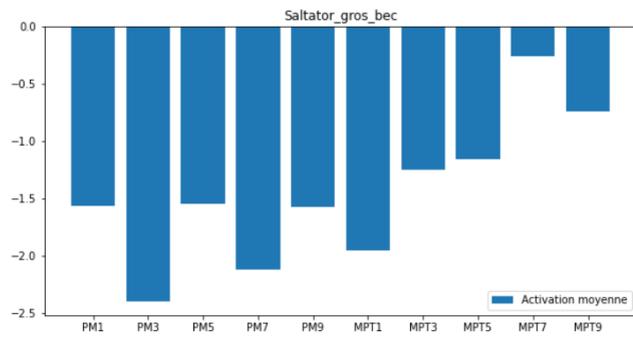
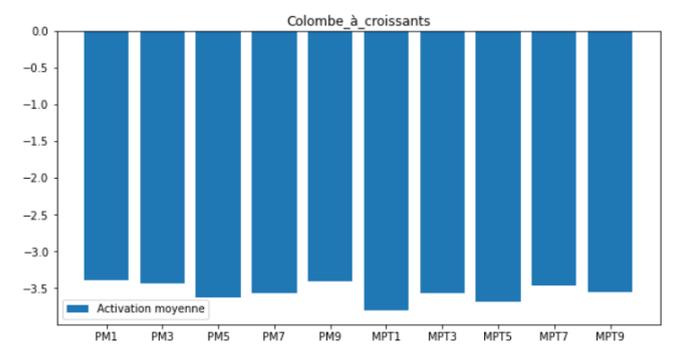
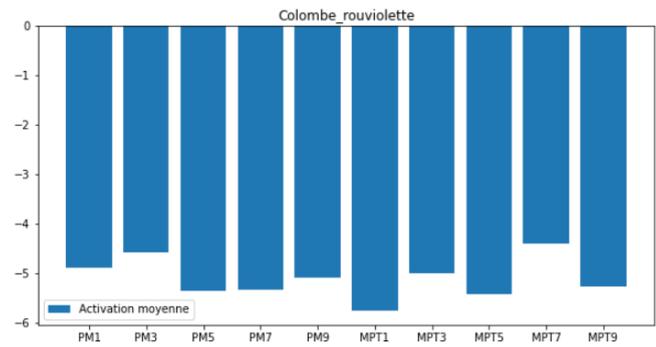
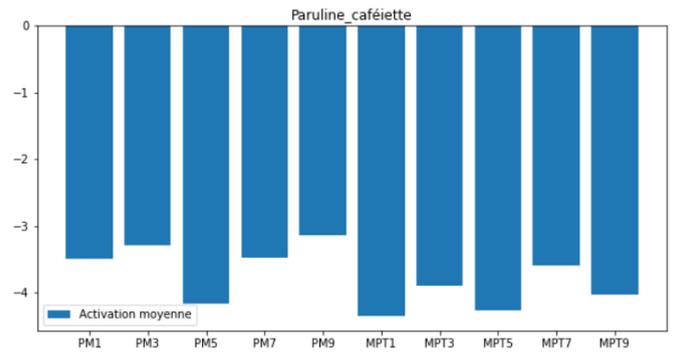
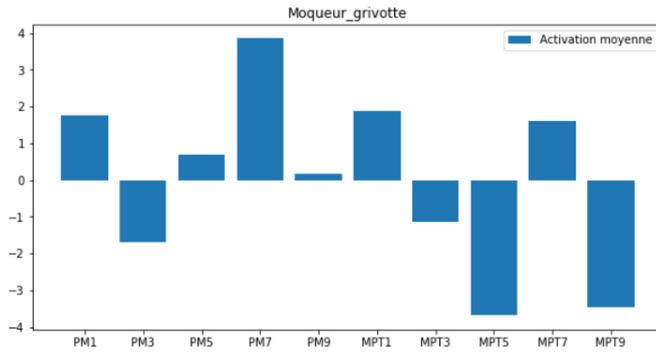
Les décisions sur les Logits par l'utilisateur doivent se poser en fonction des densités suivantes et à discuter en rapport 2.



Répartition de l'activation moyenne des neurones pour les stations PM

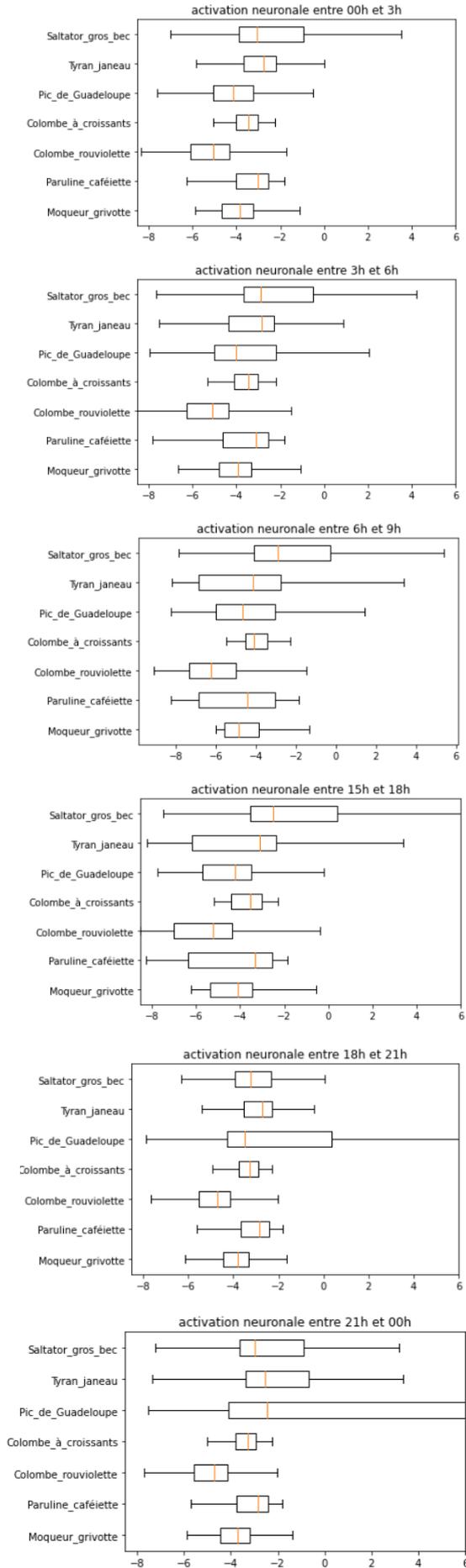


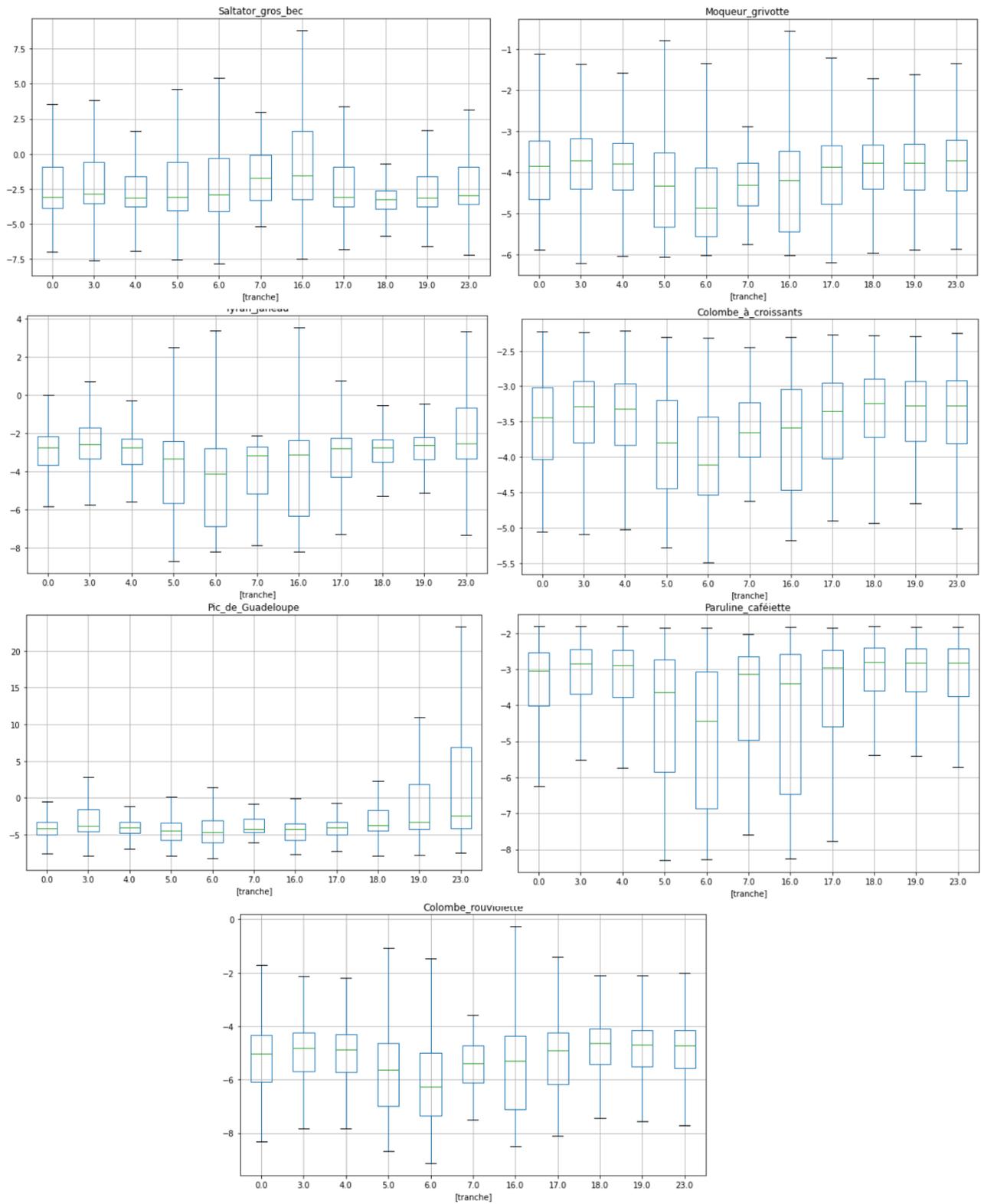
Répartition de l'activation moyenne des neurones pour les stations MPT

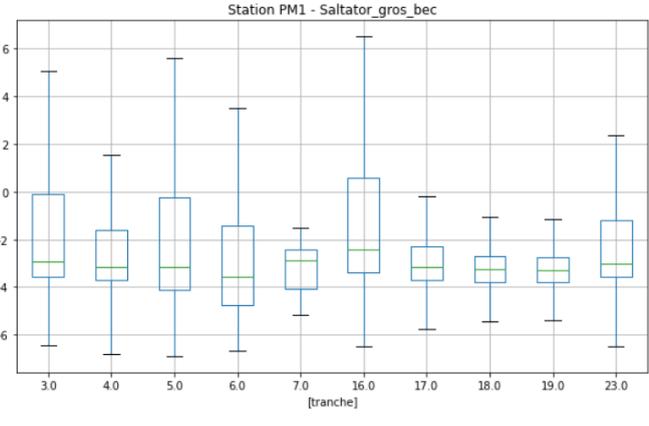
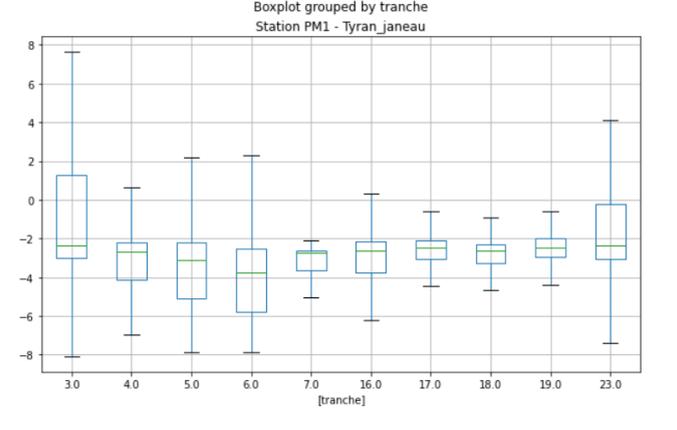
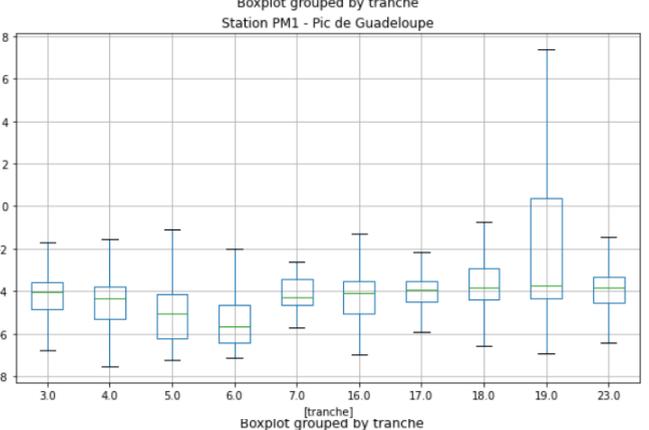
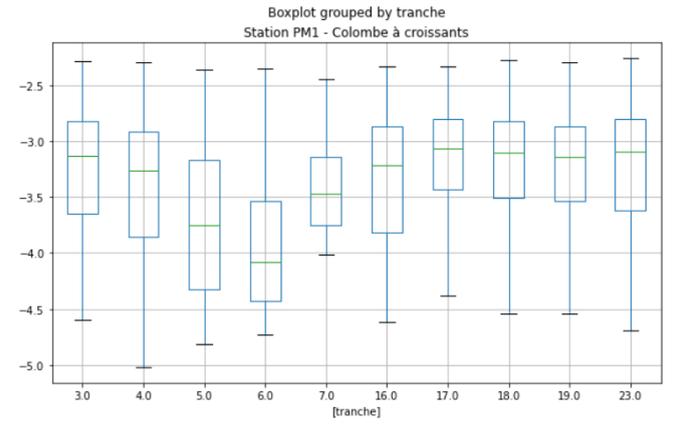
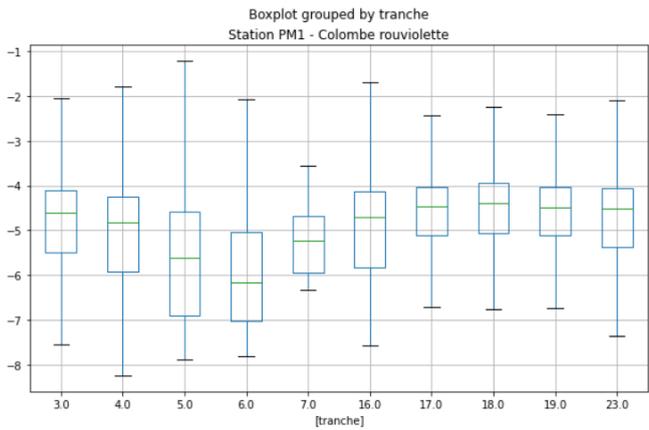
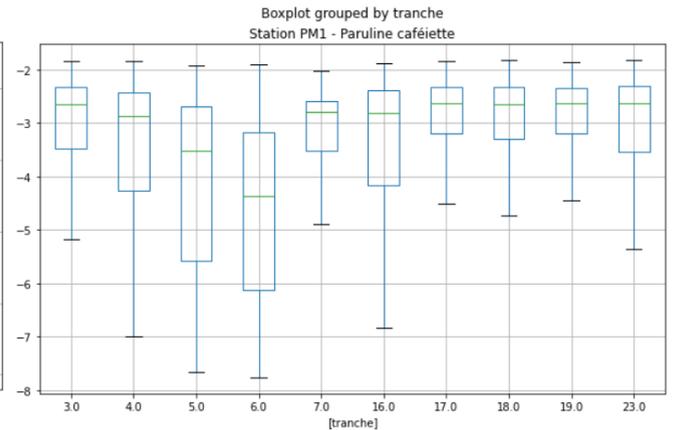
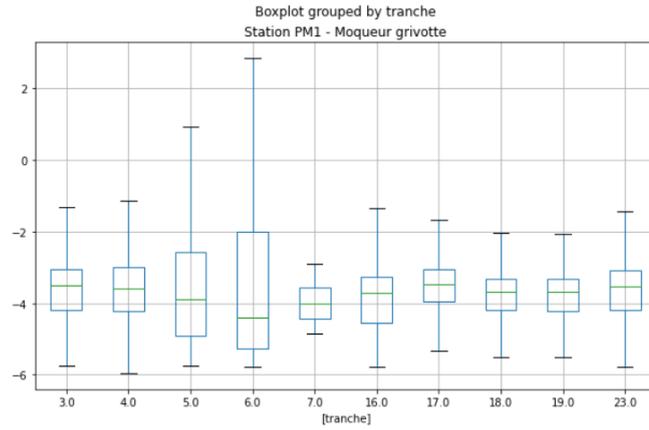


Répartition de l'activation neuronale moyenne par espèces d'oiseaux

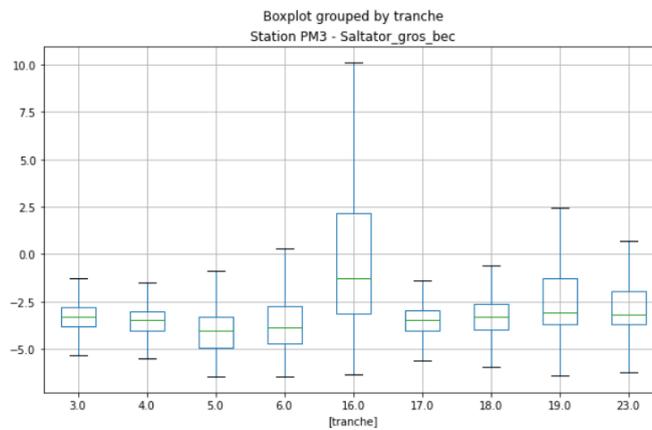
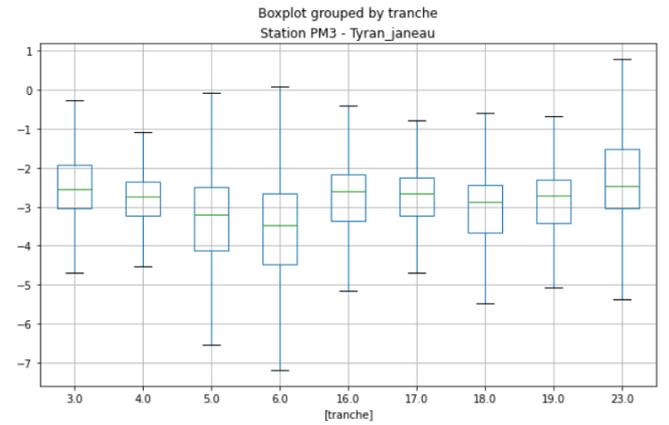
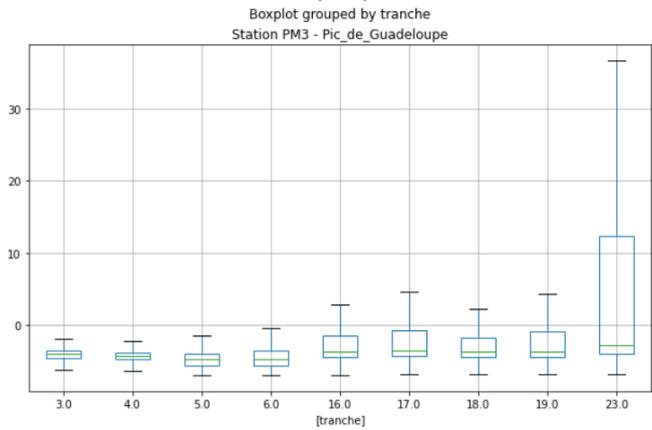
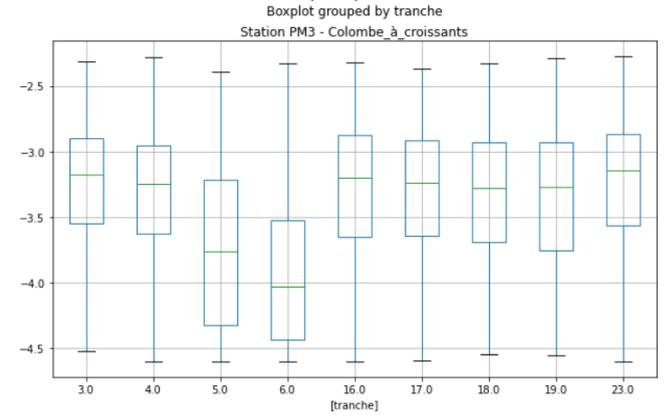
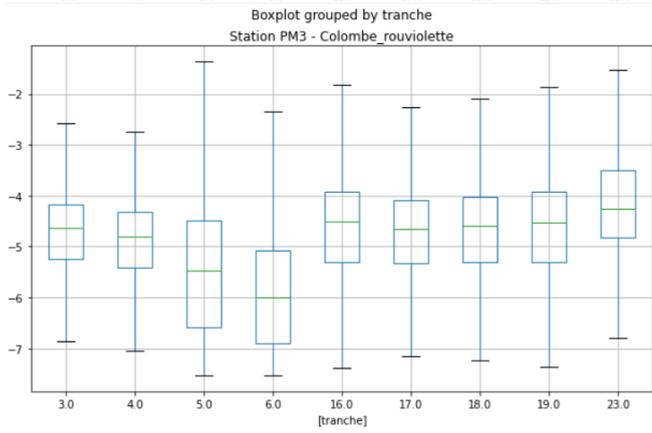
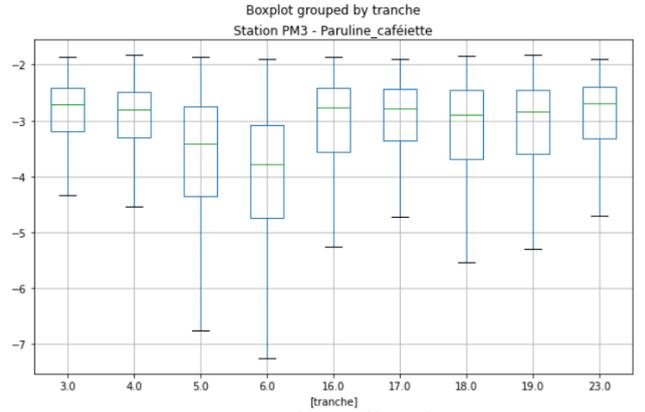
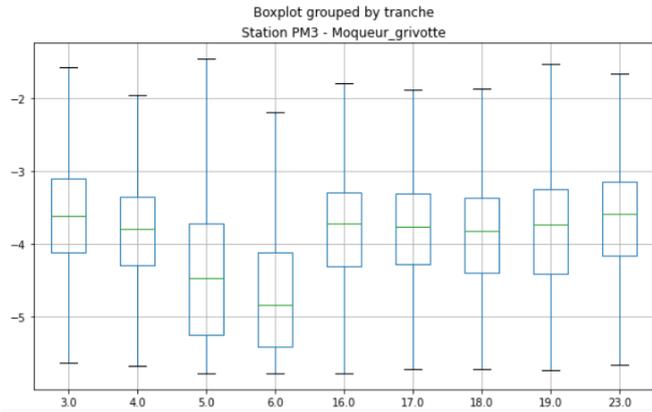
Activation neuronale par tranche horaire



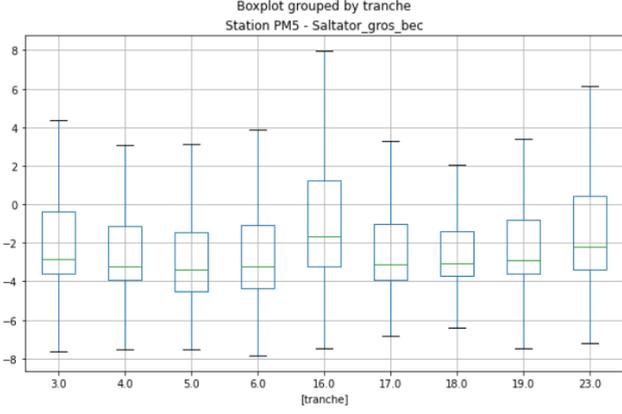
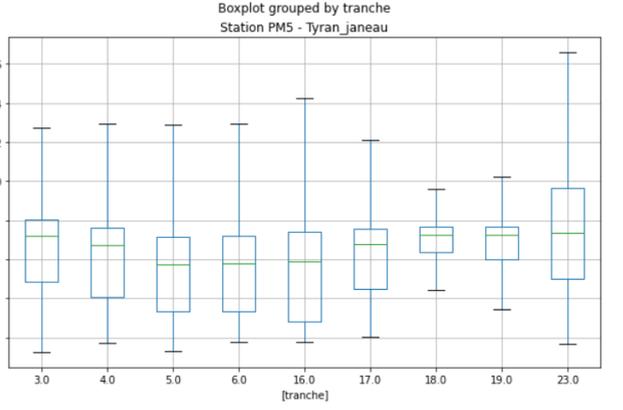
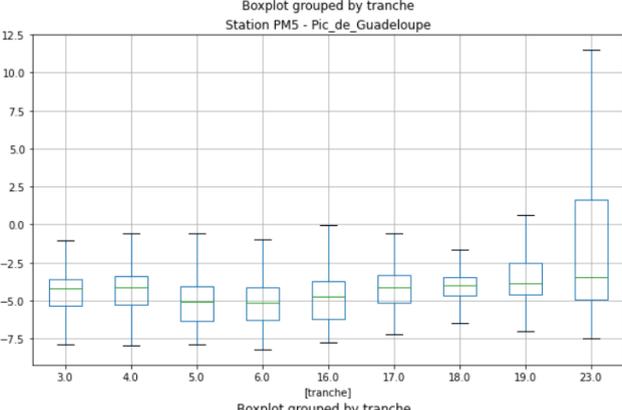
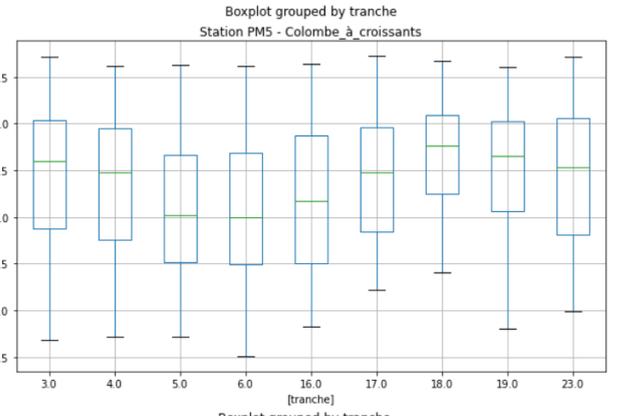
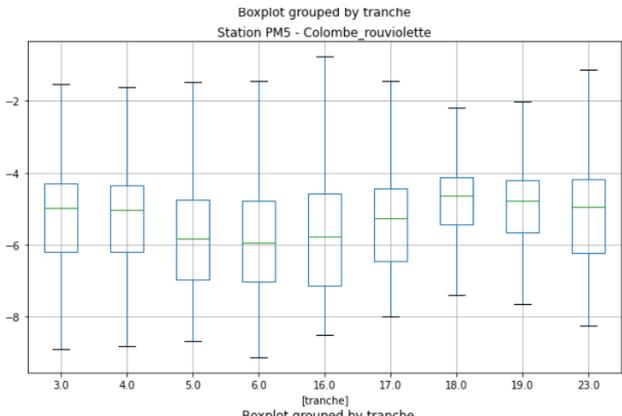
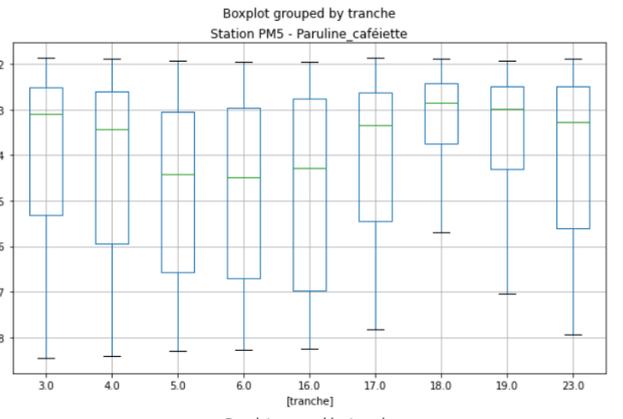
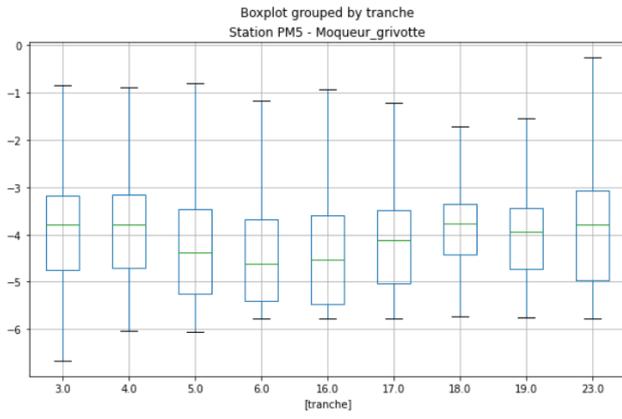




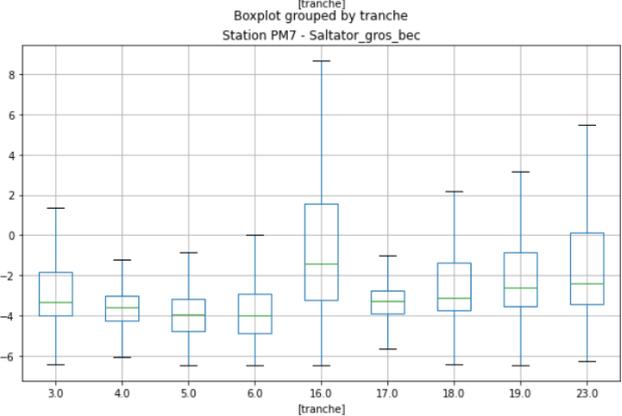
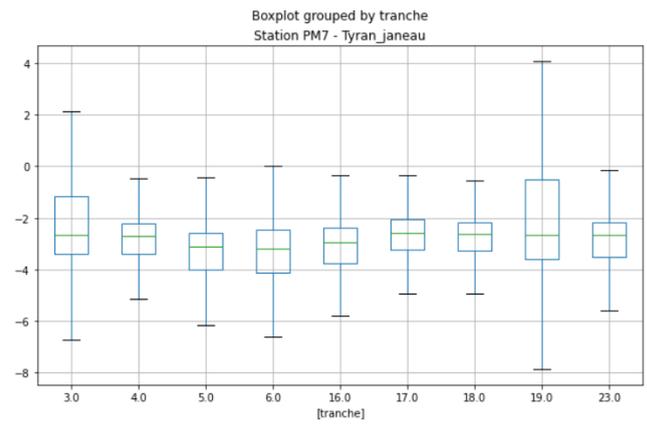
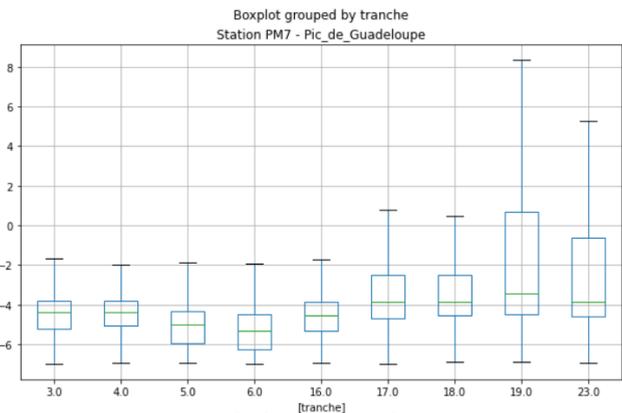
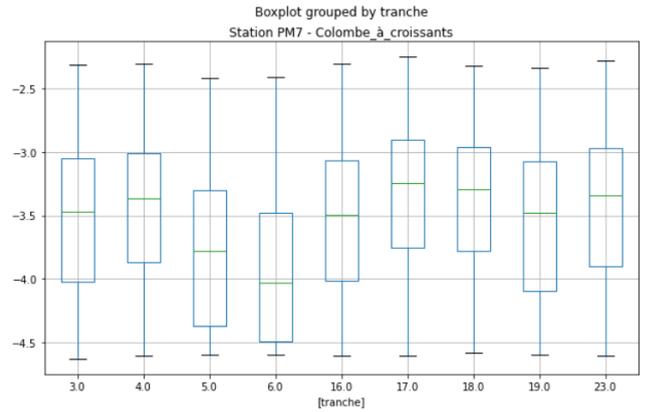
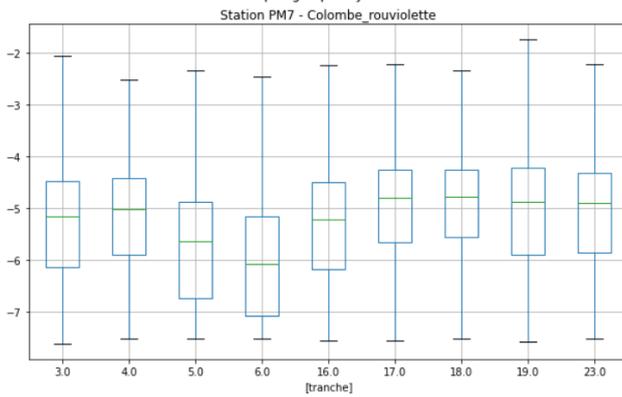
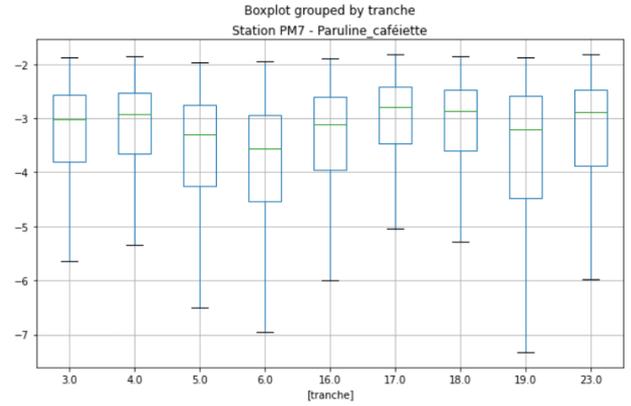
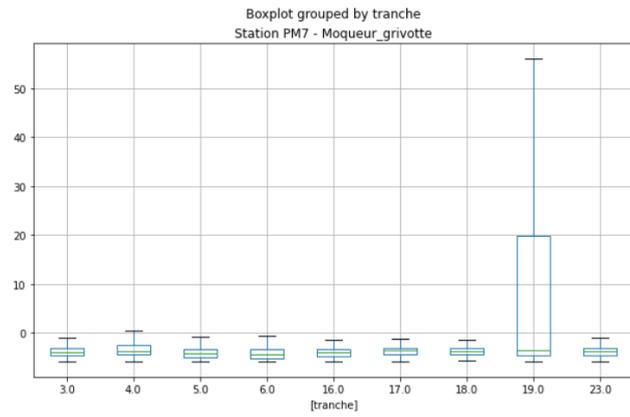
Station PM1



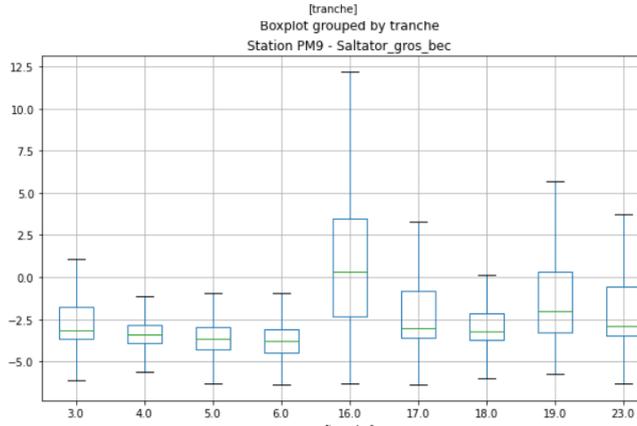
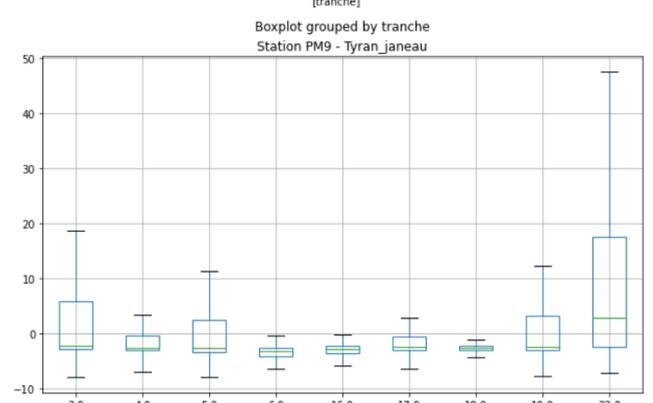
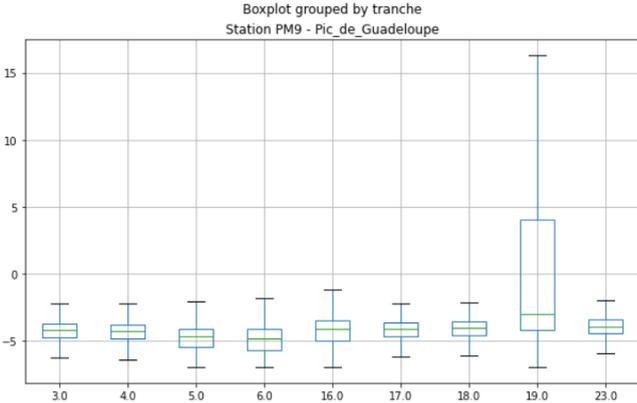
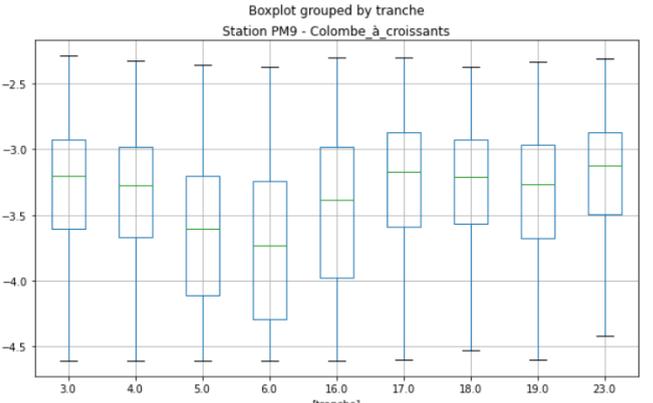
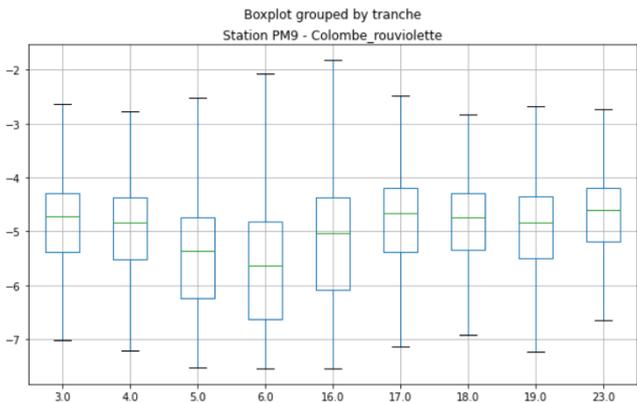
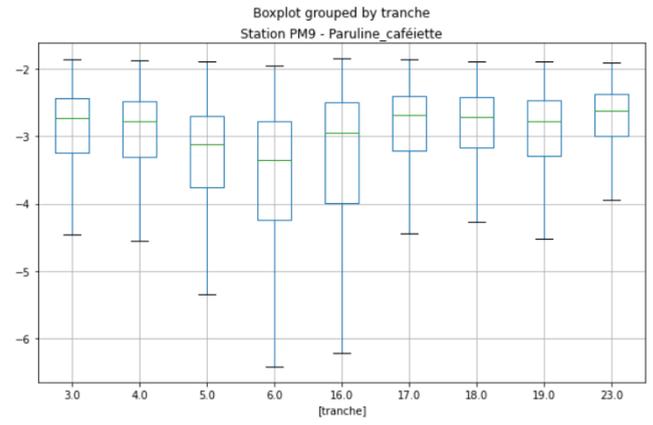
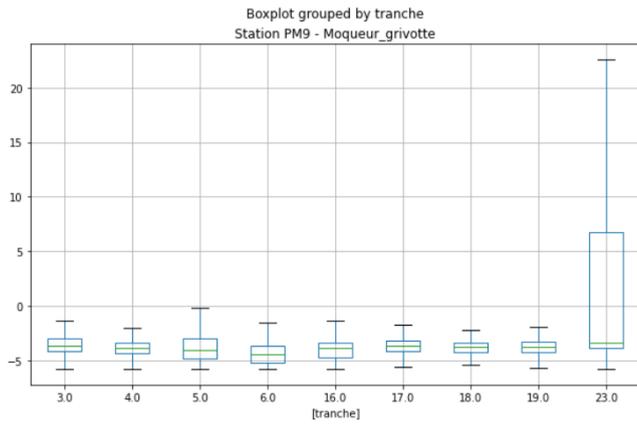
Station PM3



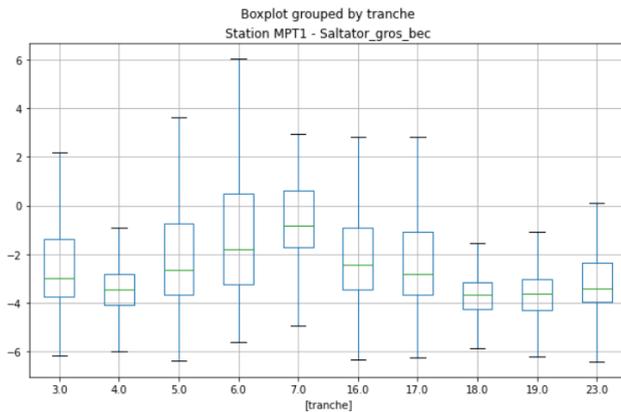
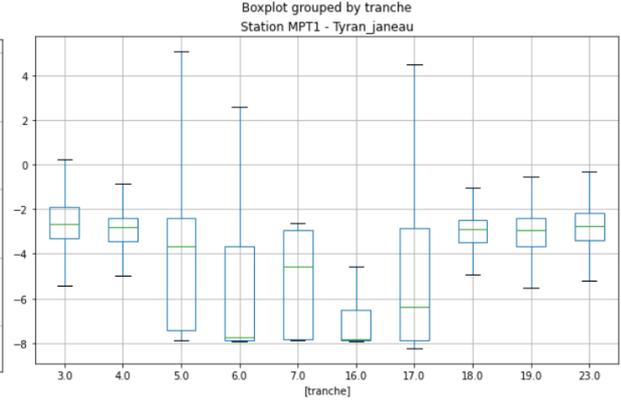
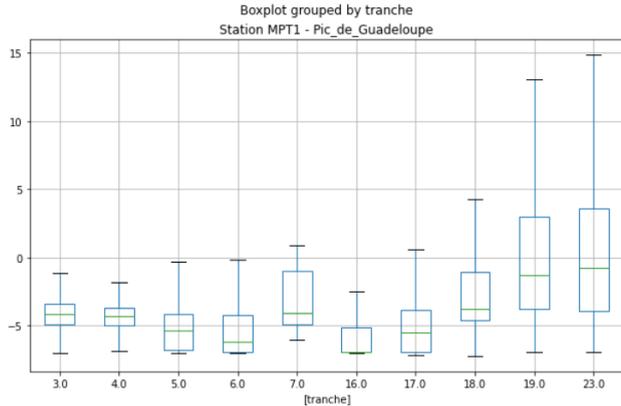
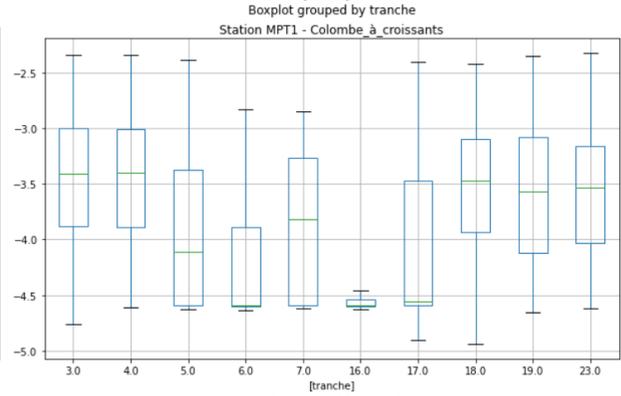
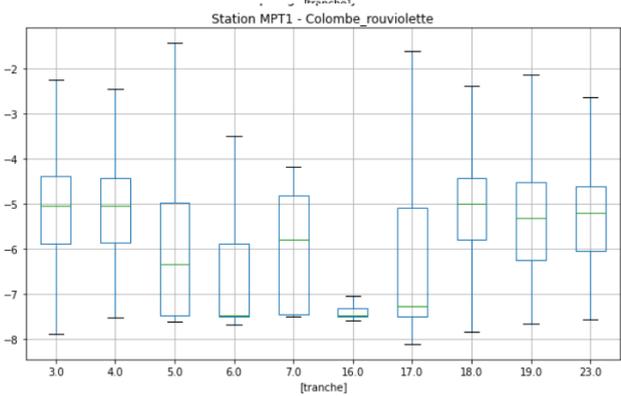
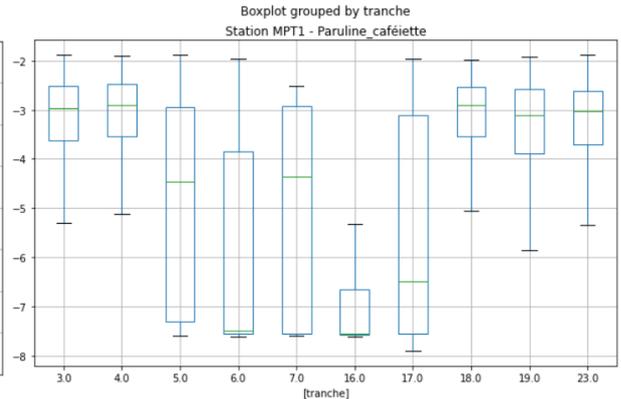
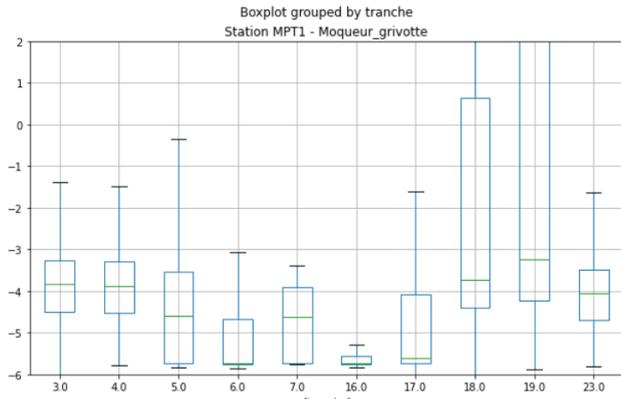
Station PM5



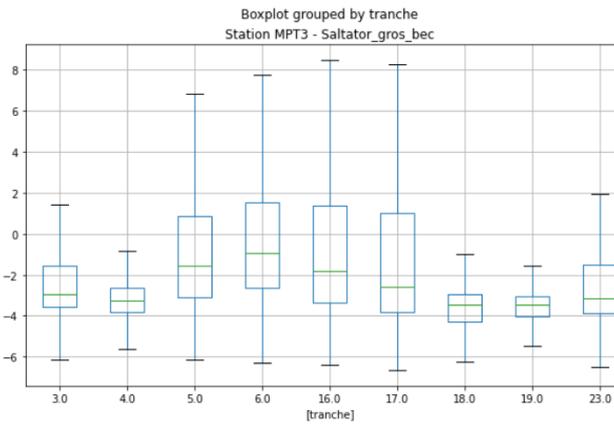
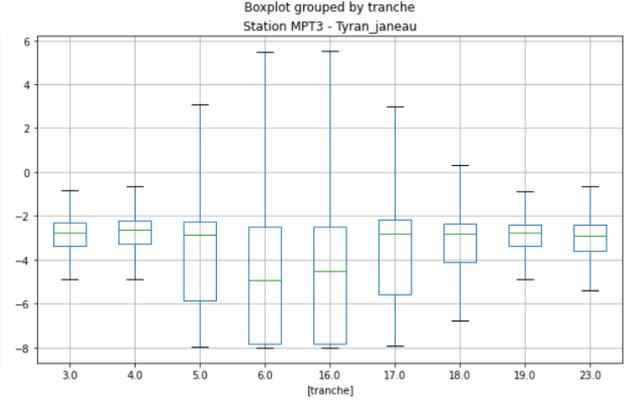
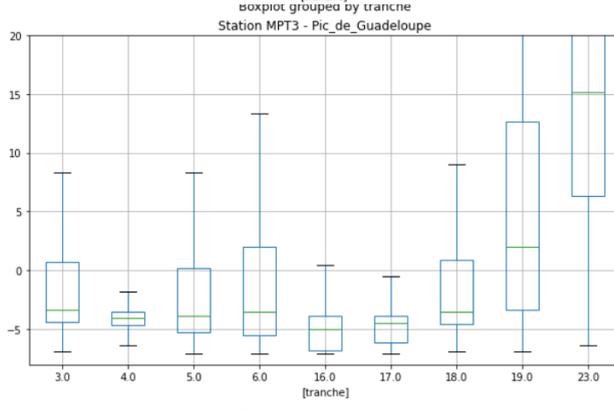
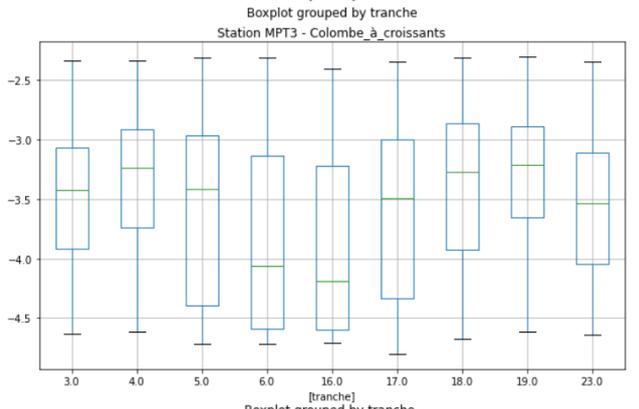
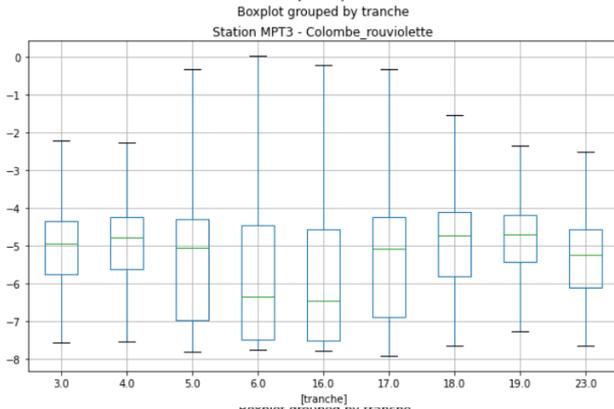
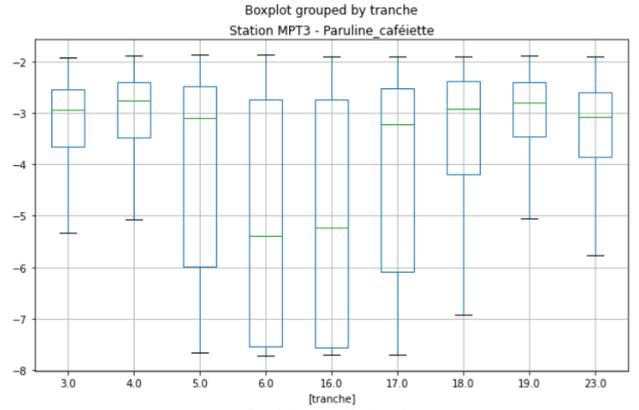
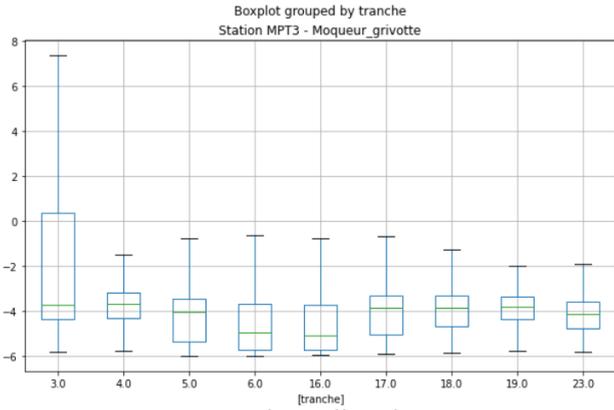
Station PM7



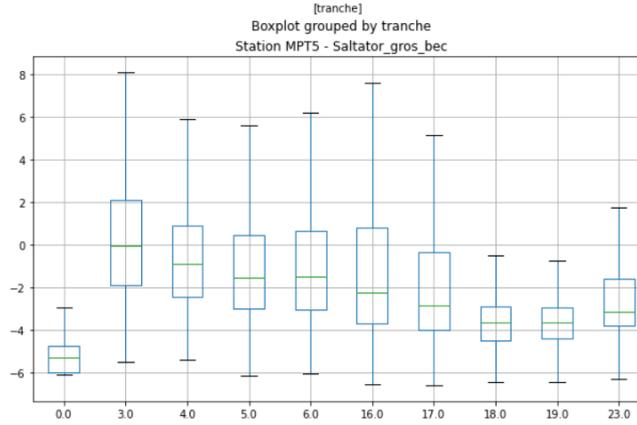
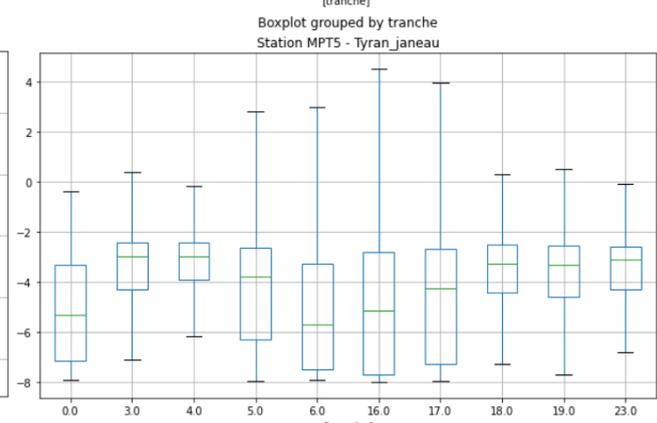
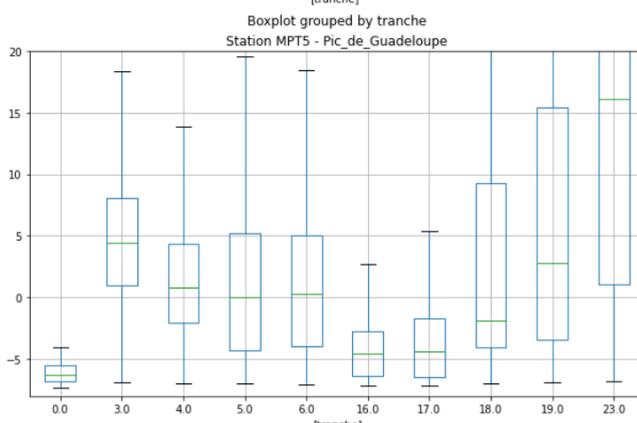
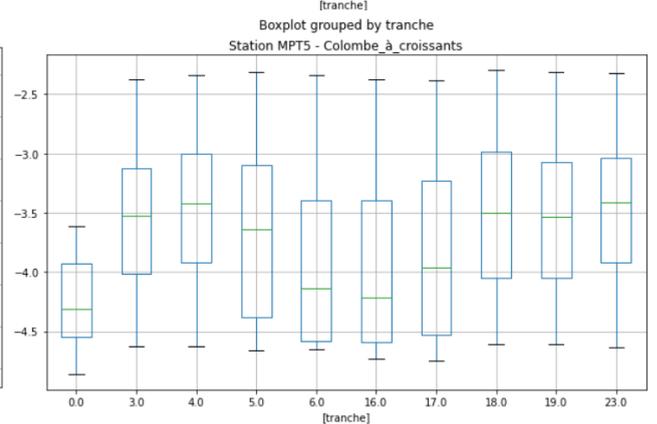
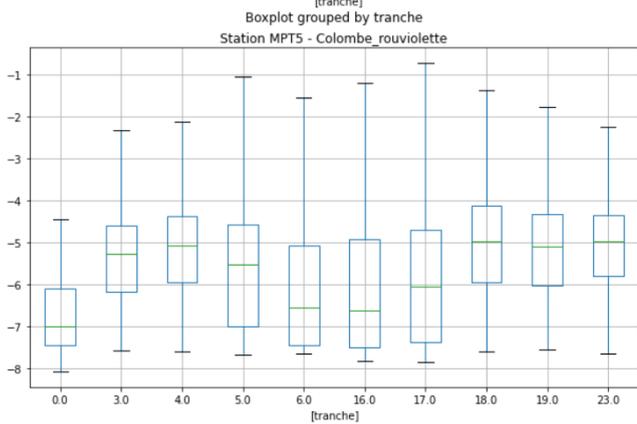
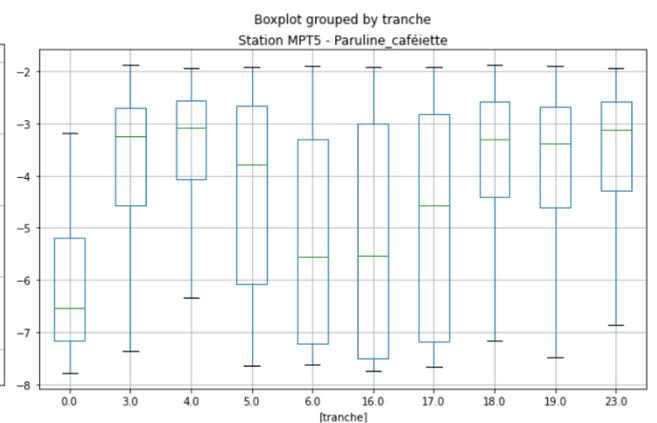
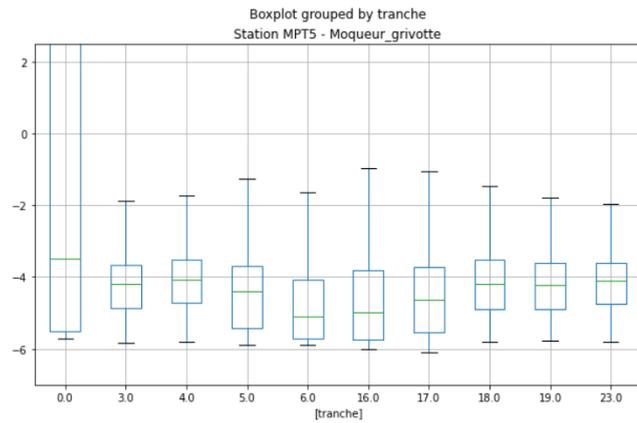
Station PM9



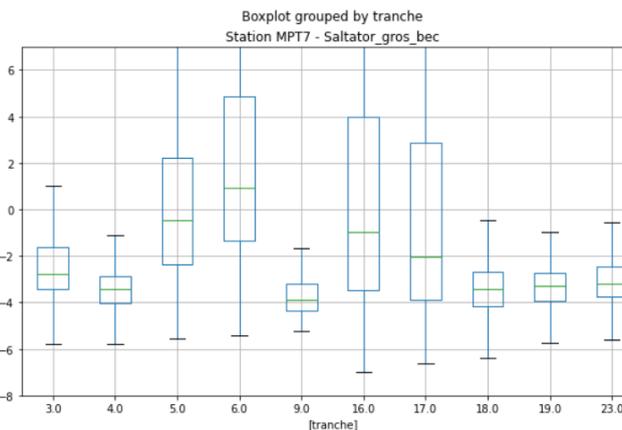
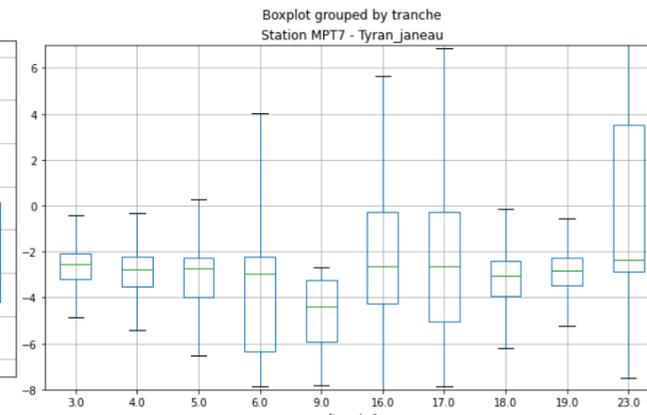
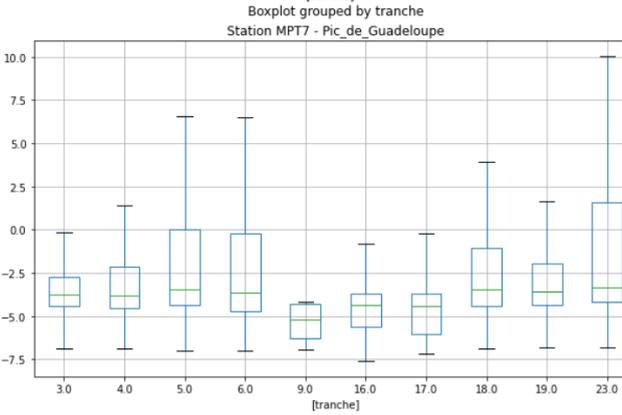
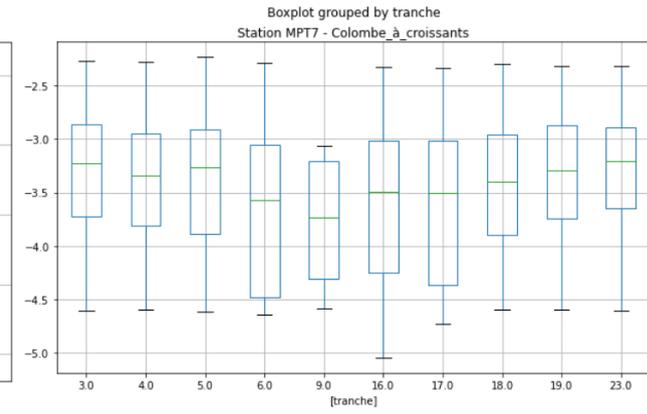
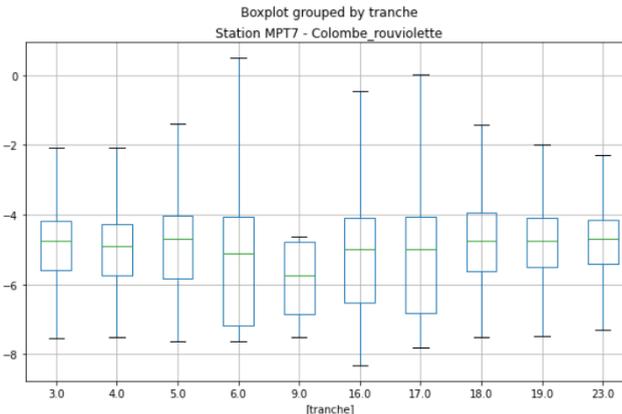
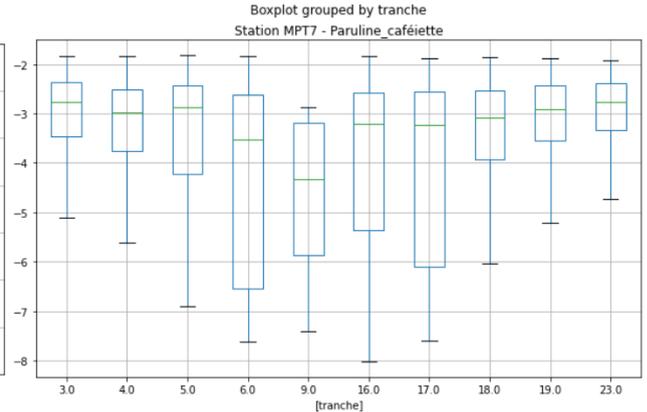
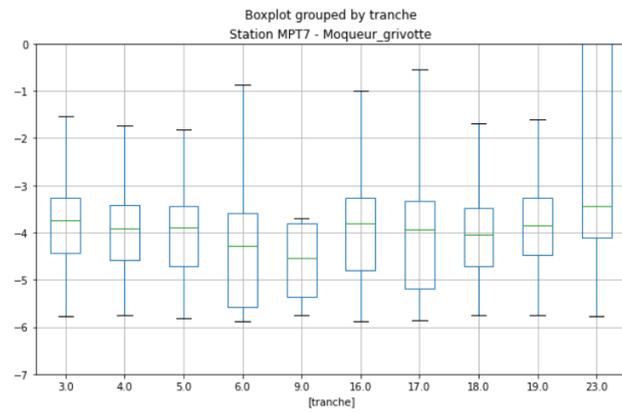
Station MPT1



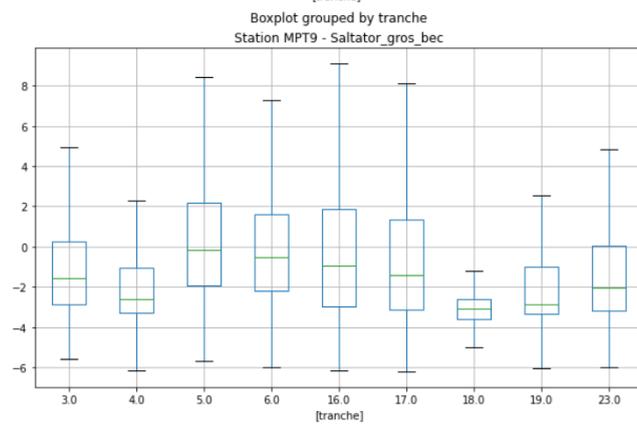
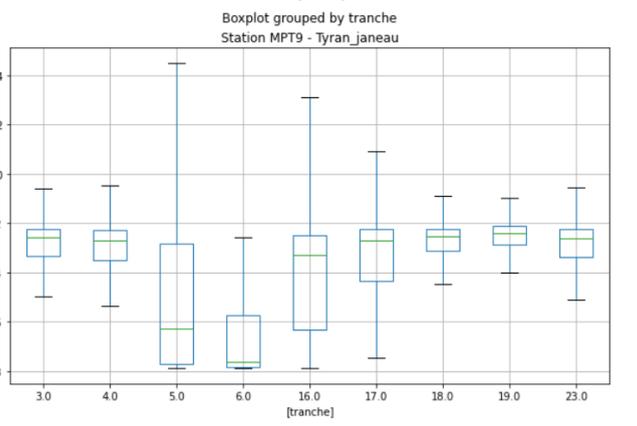
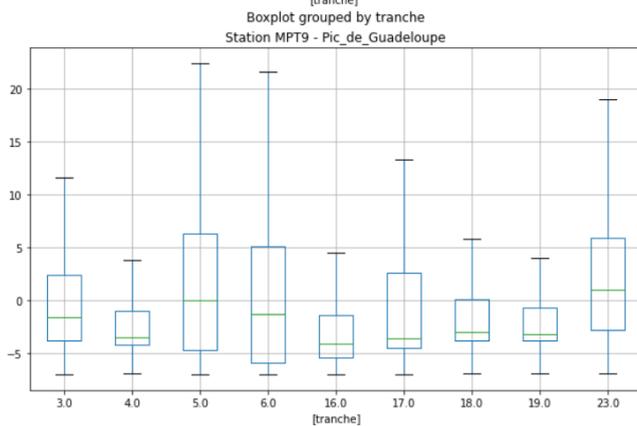
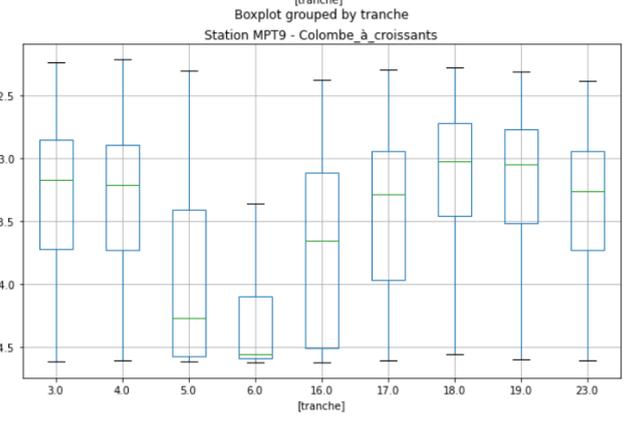
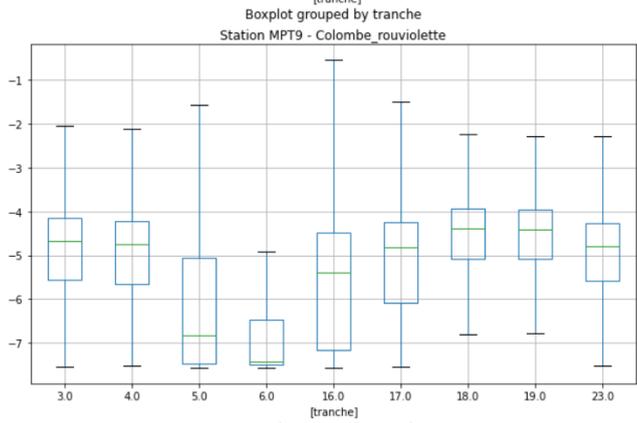
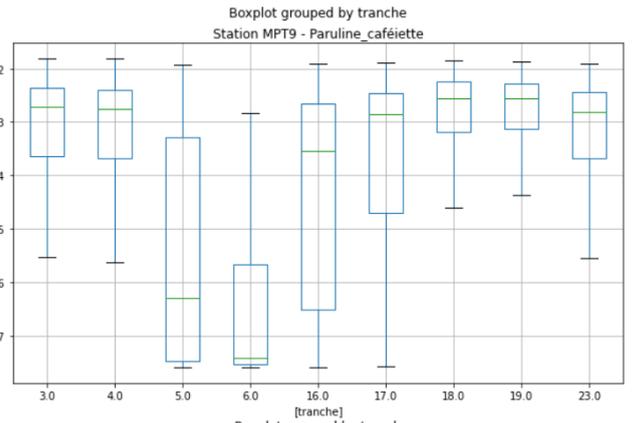
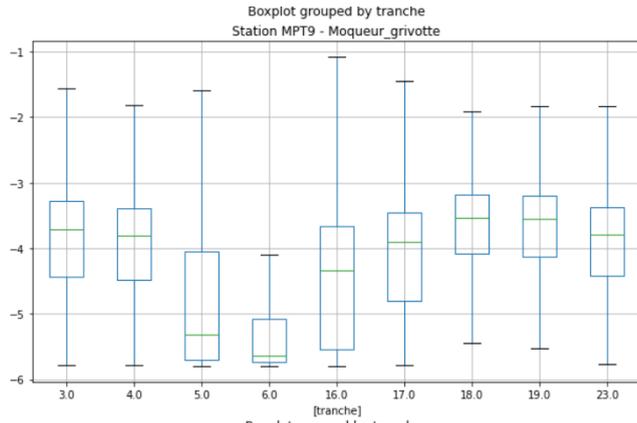
Station MPT3



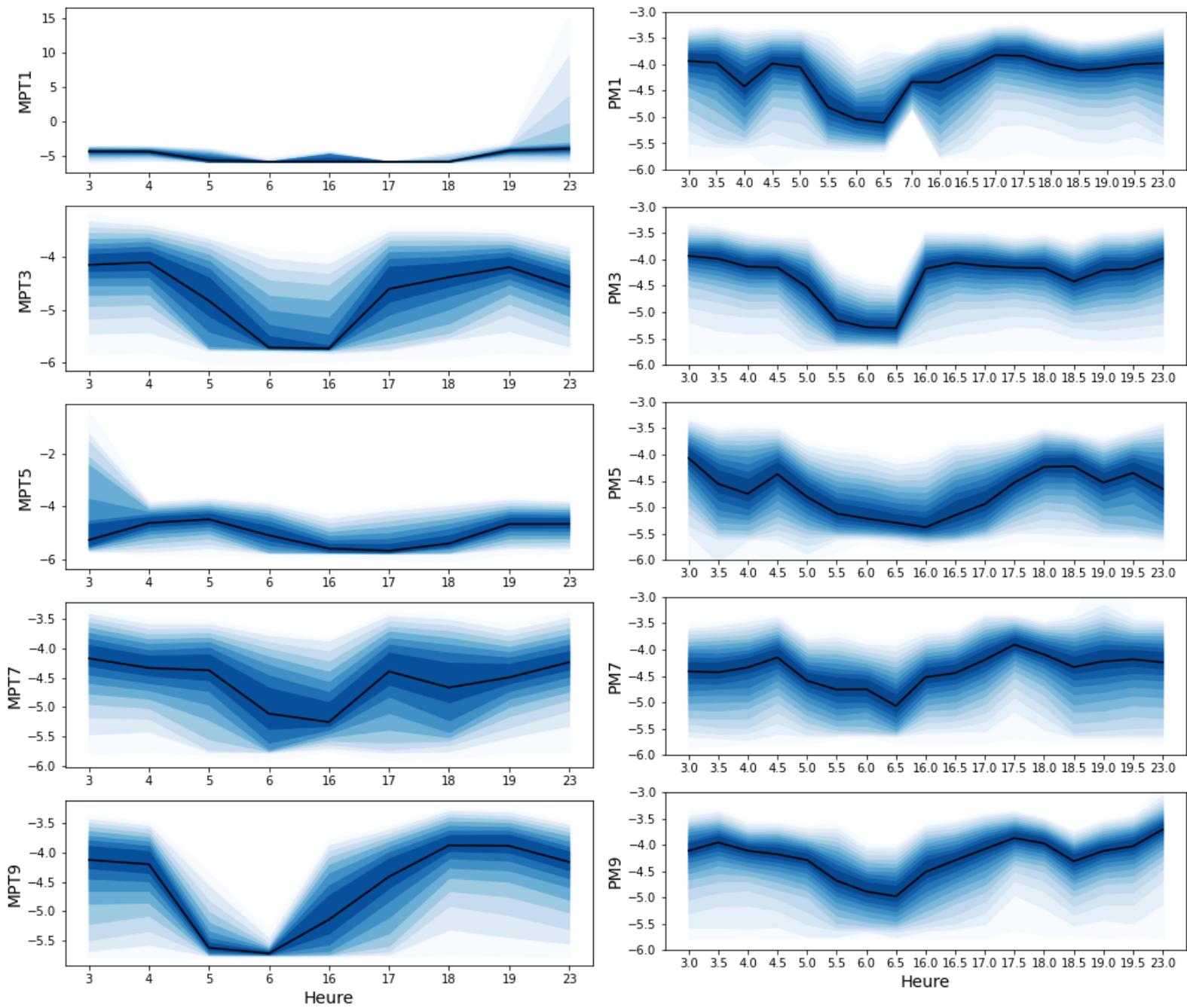
Station MPT5



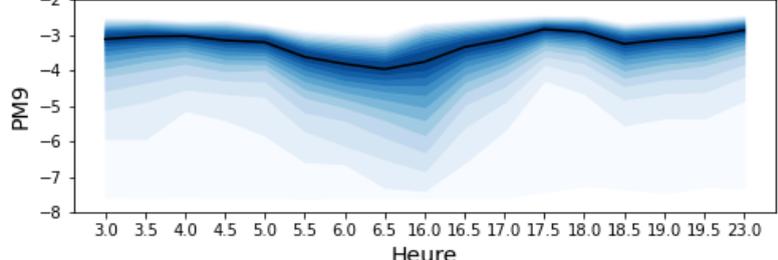
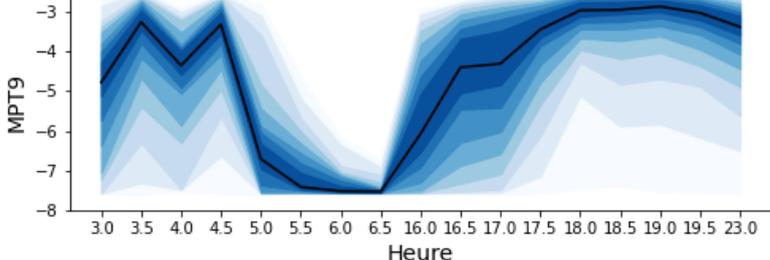
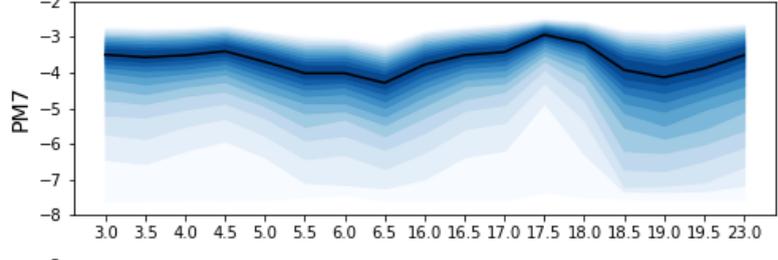
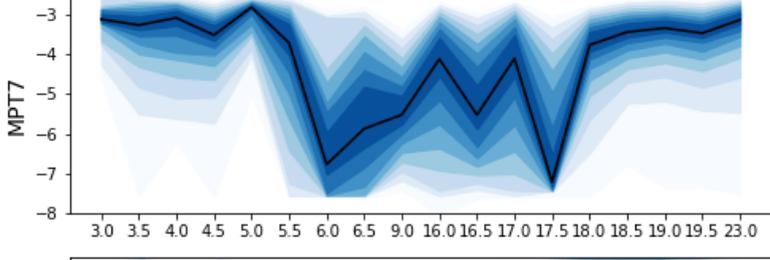
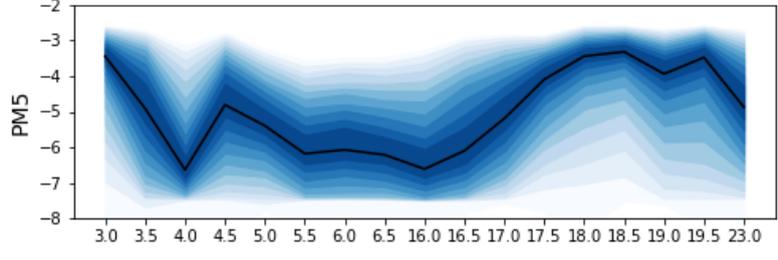
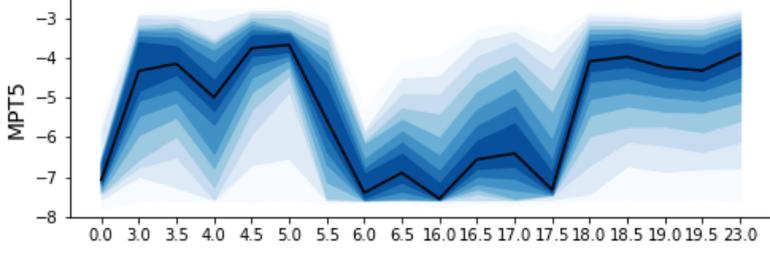
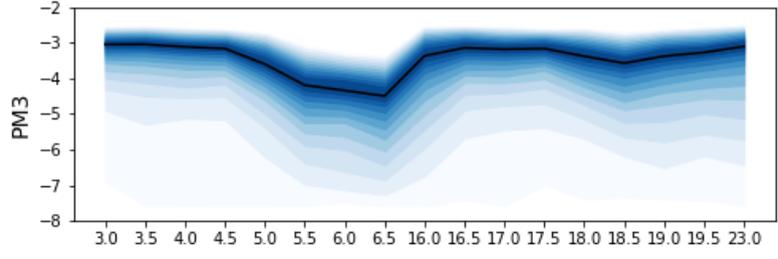
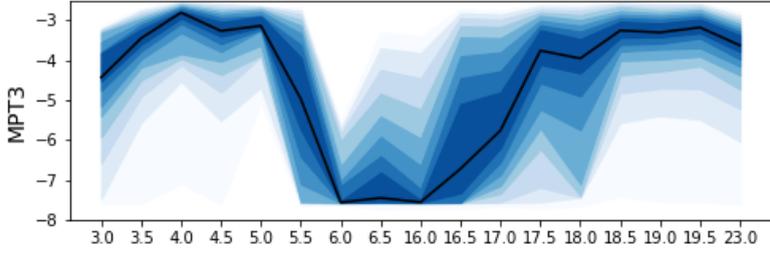
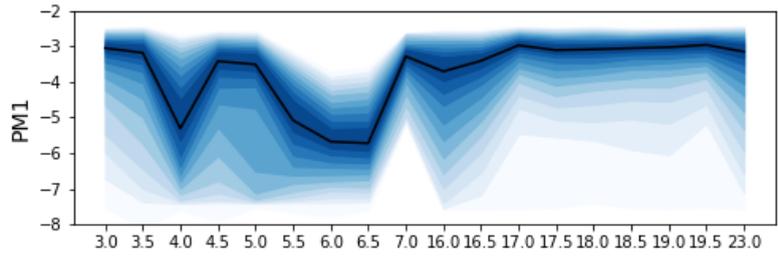
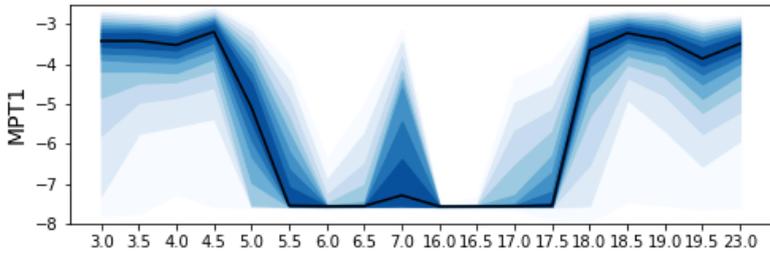
Station MPT7



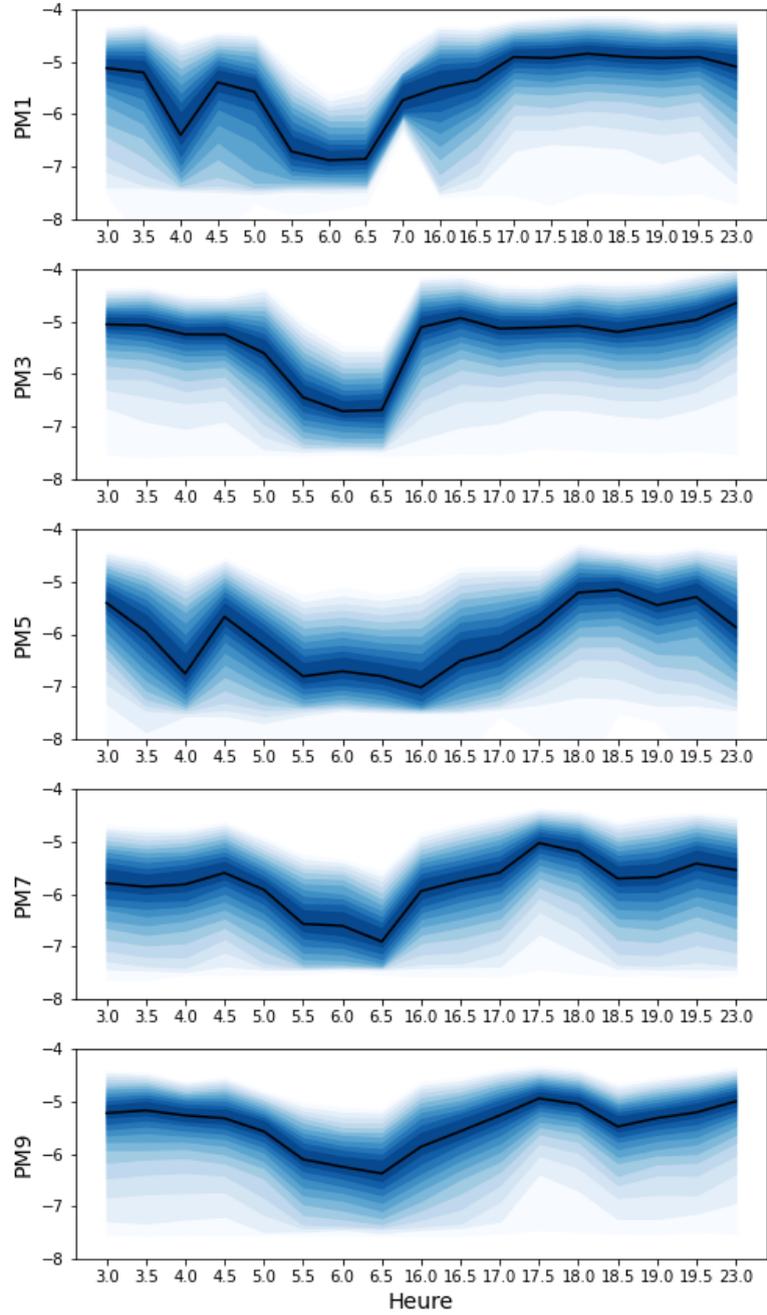
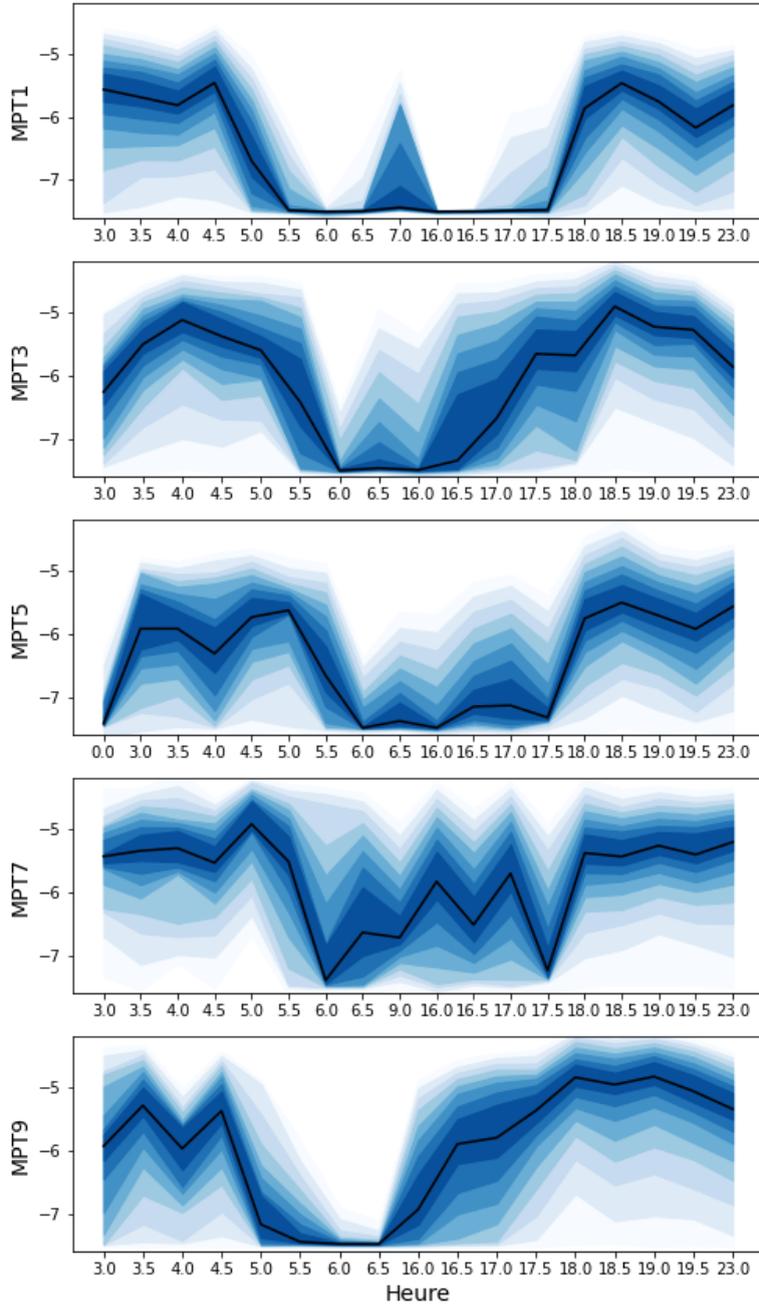
Station MPT9



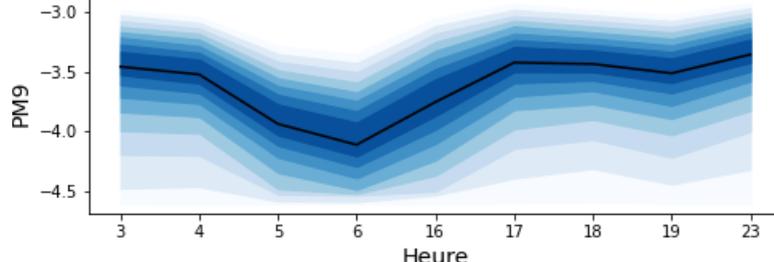
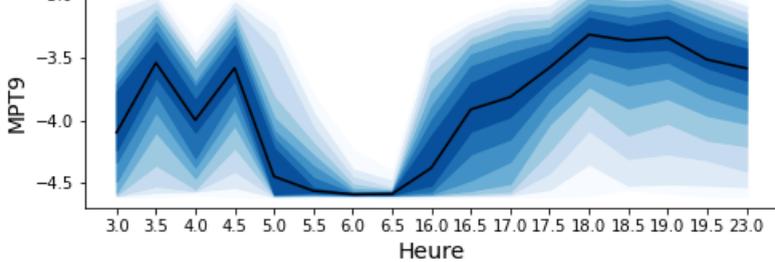
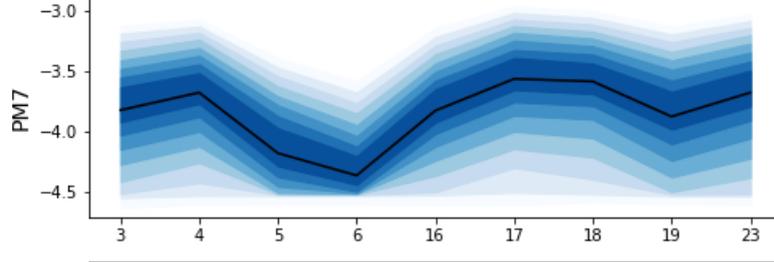
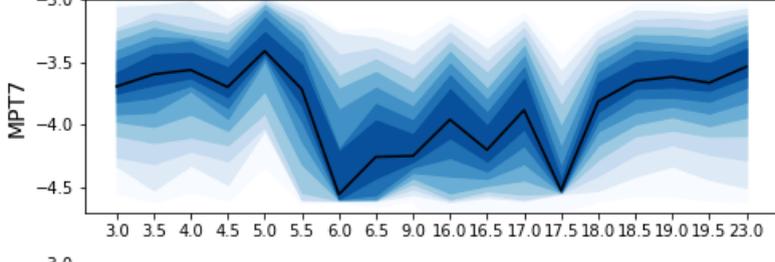
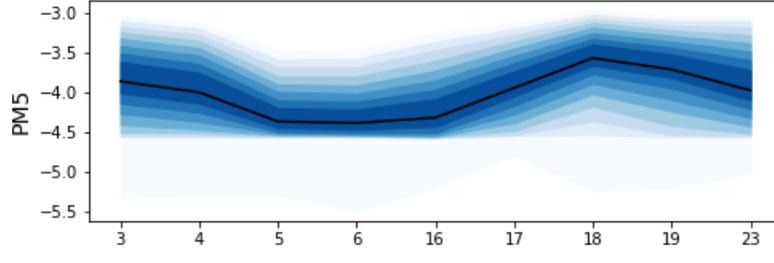
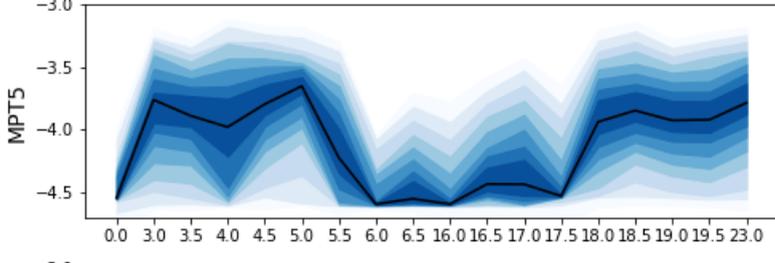
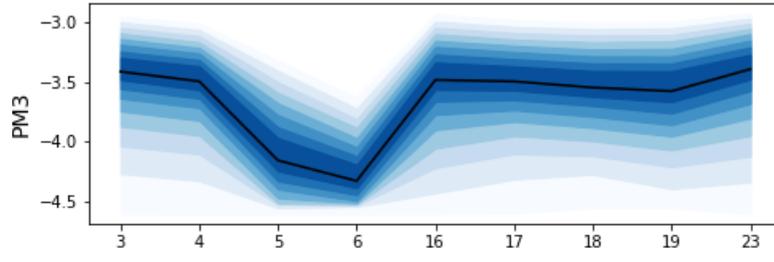
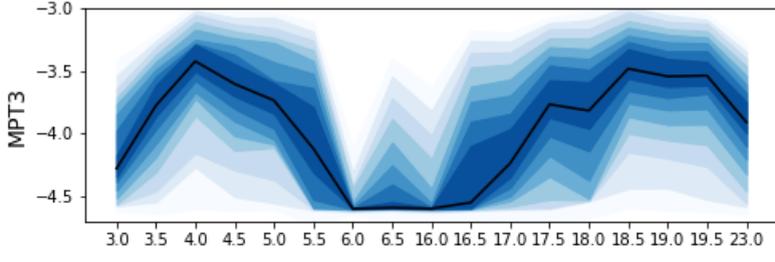
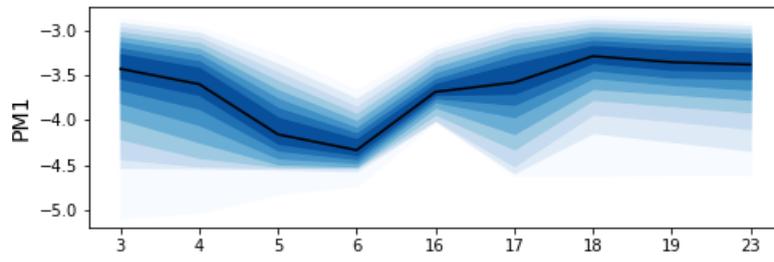
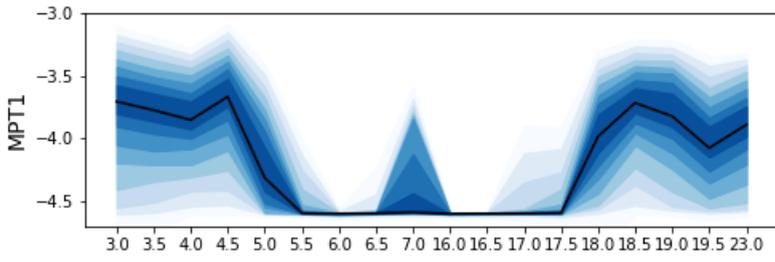
Moqueur grivotte : Prendre en compte la durée de l'activation -> Pic à 17-18h en MPT 1/3:
 écarts des percentiles supérieurs à la médiane -> MPT5 + pic à 7h



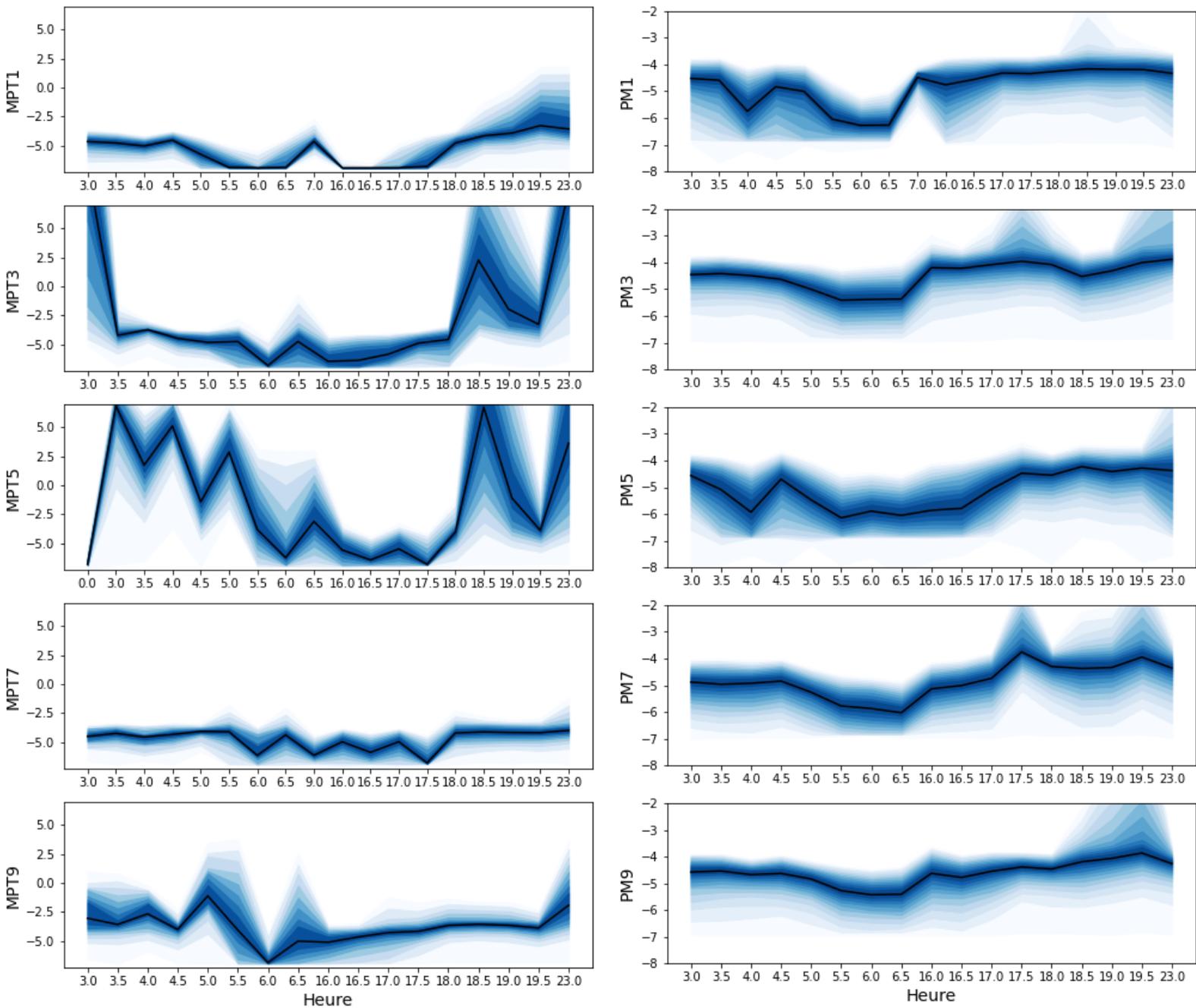
Paruline caféïette : médiane plate -> bruit de fond
 Pic à 17-18h + 7h pour MPT3 et 5
 et 7h MPT1
 PM courbe plutôt plate sauf PM5 (et encore)



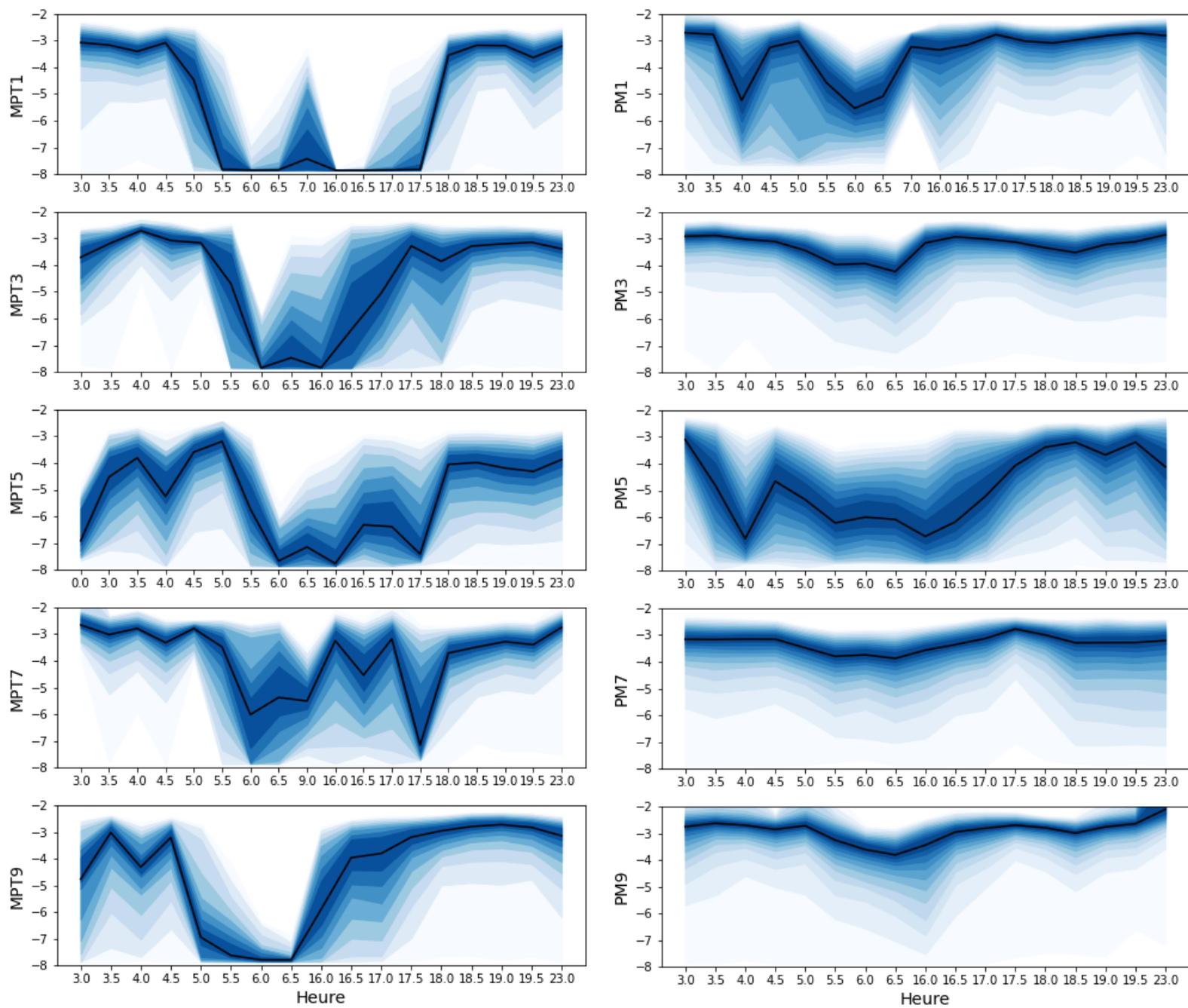
Colombe rouviolette :



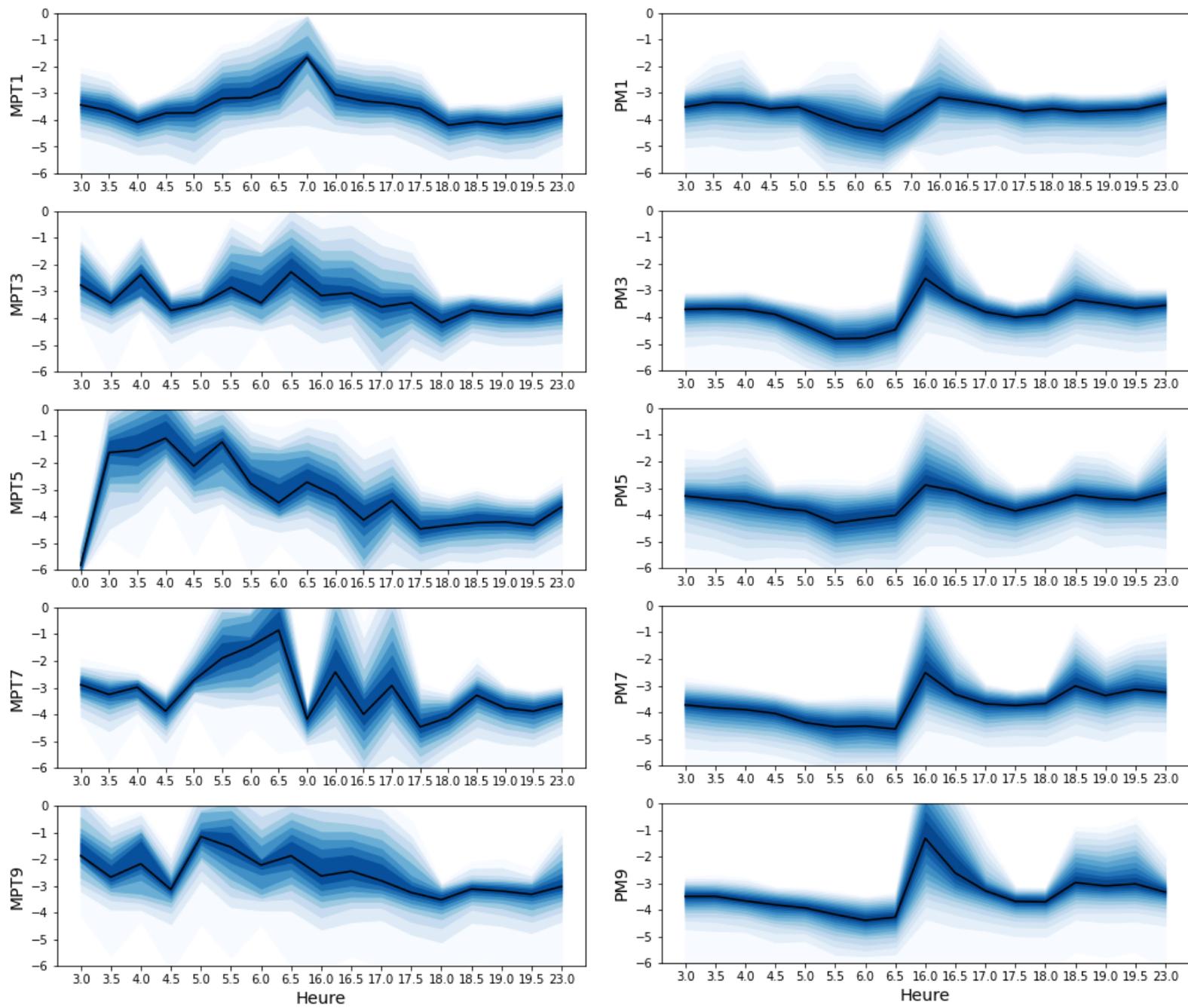
Colombe à croissant



Pic de Guadeloupe : probable présence à 5h sur MPT5 et 18h MPT3 MPT5, PM7 et PM3
17:30 à tester à l'écoute...



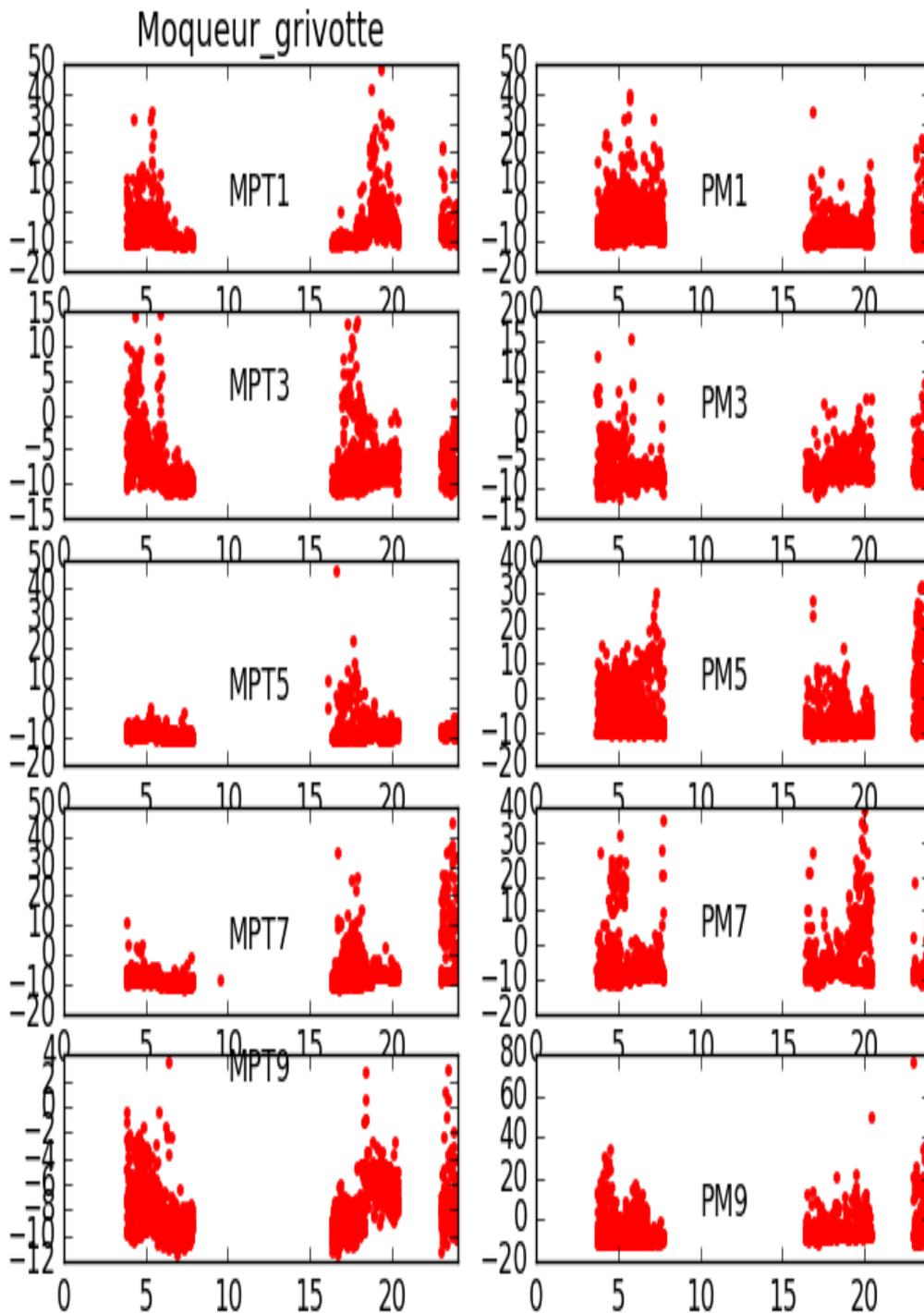
Tyran janeau : Pic à 7h MPT1/3/7 et à 17/18h pour MPT3/5/7
Problème de bruit pour PM (peut être séparer les bruits pour PM et MPT , séparer par
horaire).

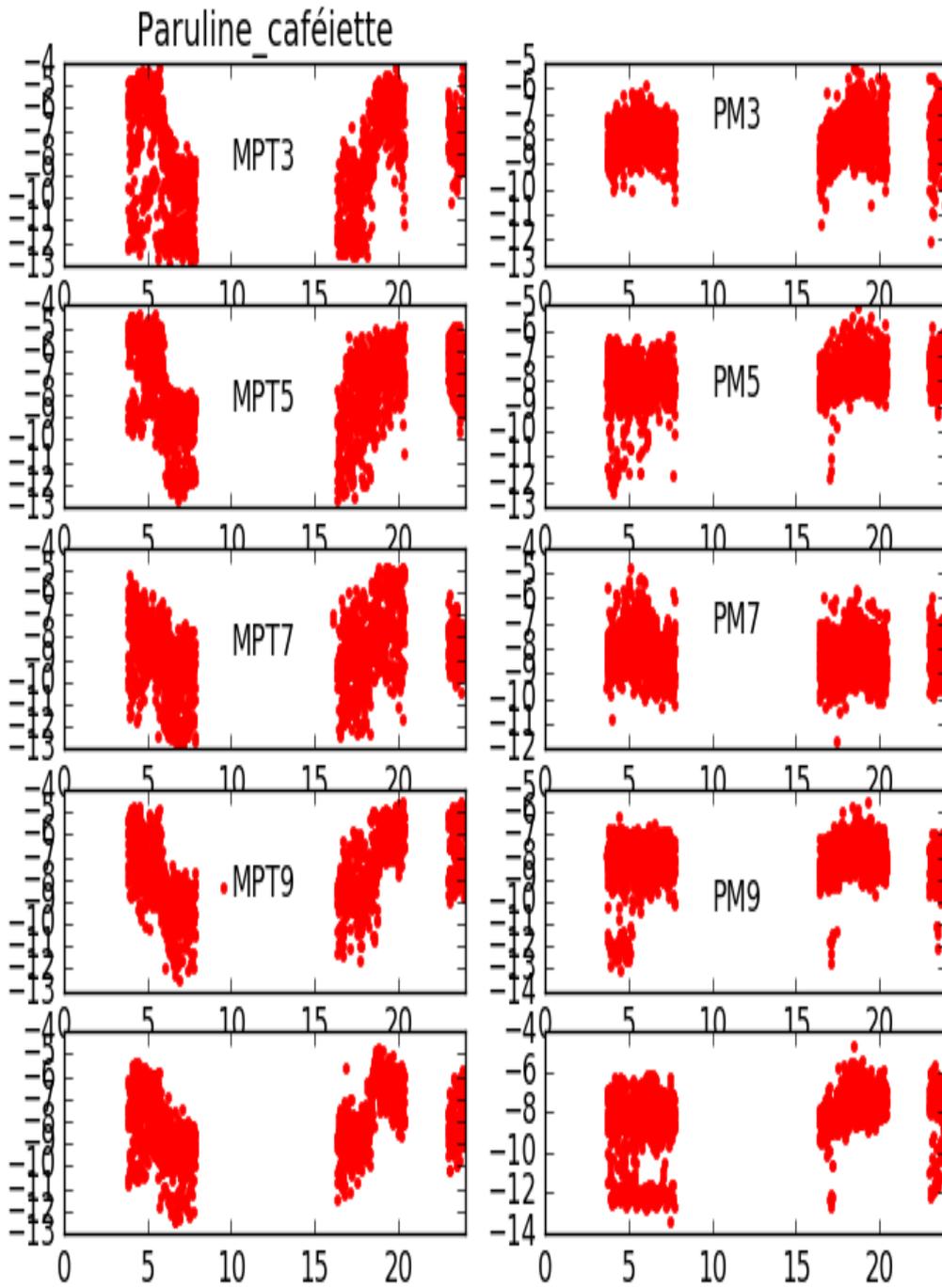


Saltator gros bec : vérifier à l'écoute PM* 16h

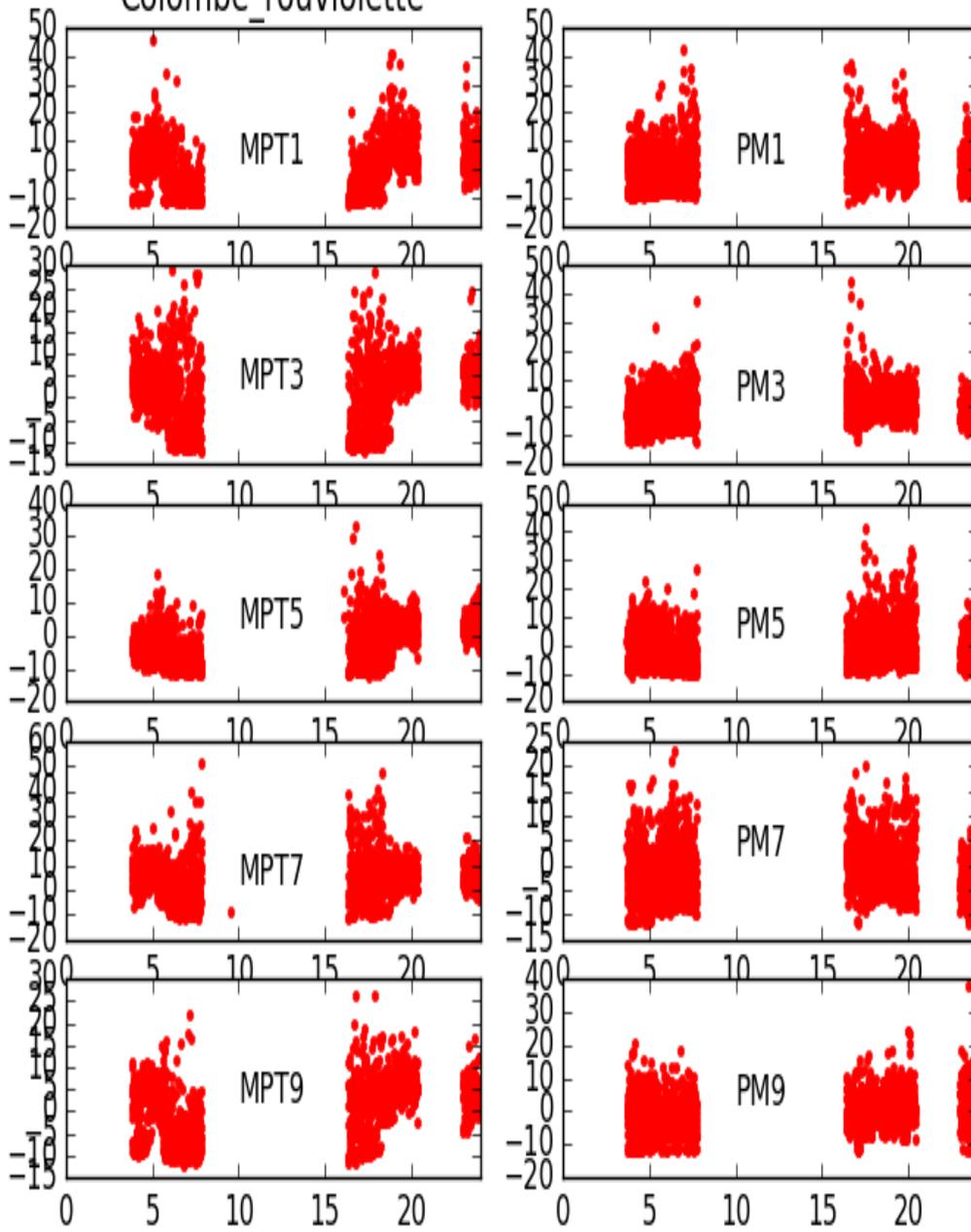
MPT1 à 7h

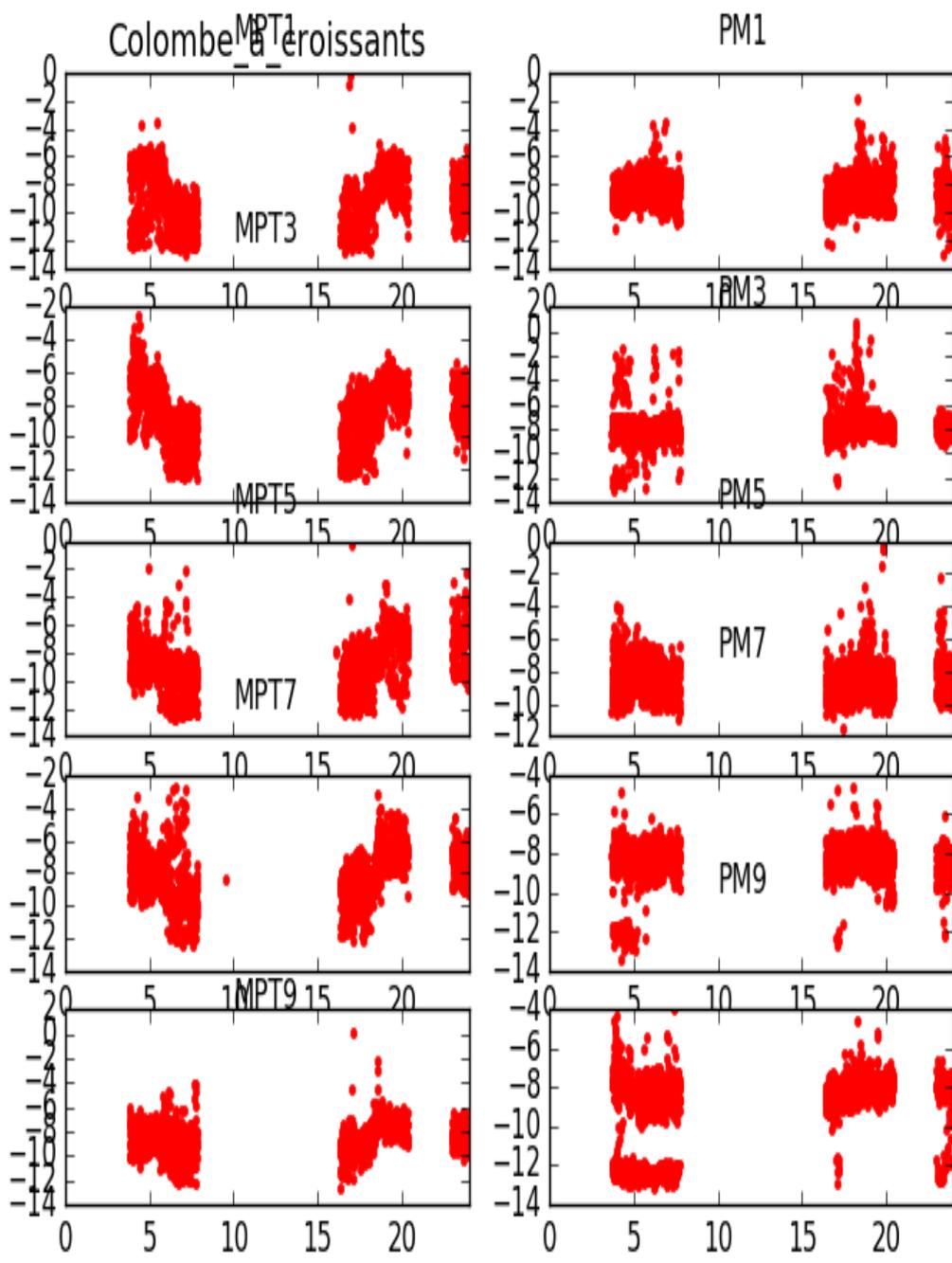
ANNEXE 6 : Résultats de projection en temps et espace des distances du percentile 90 à la médiane.

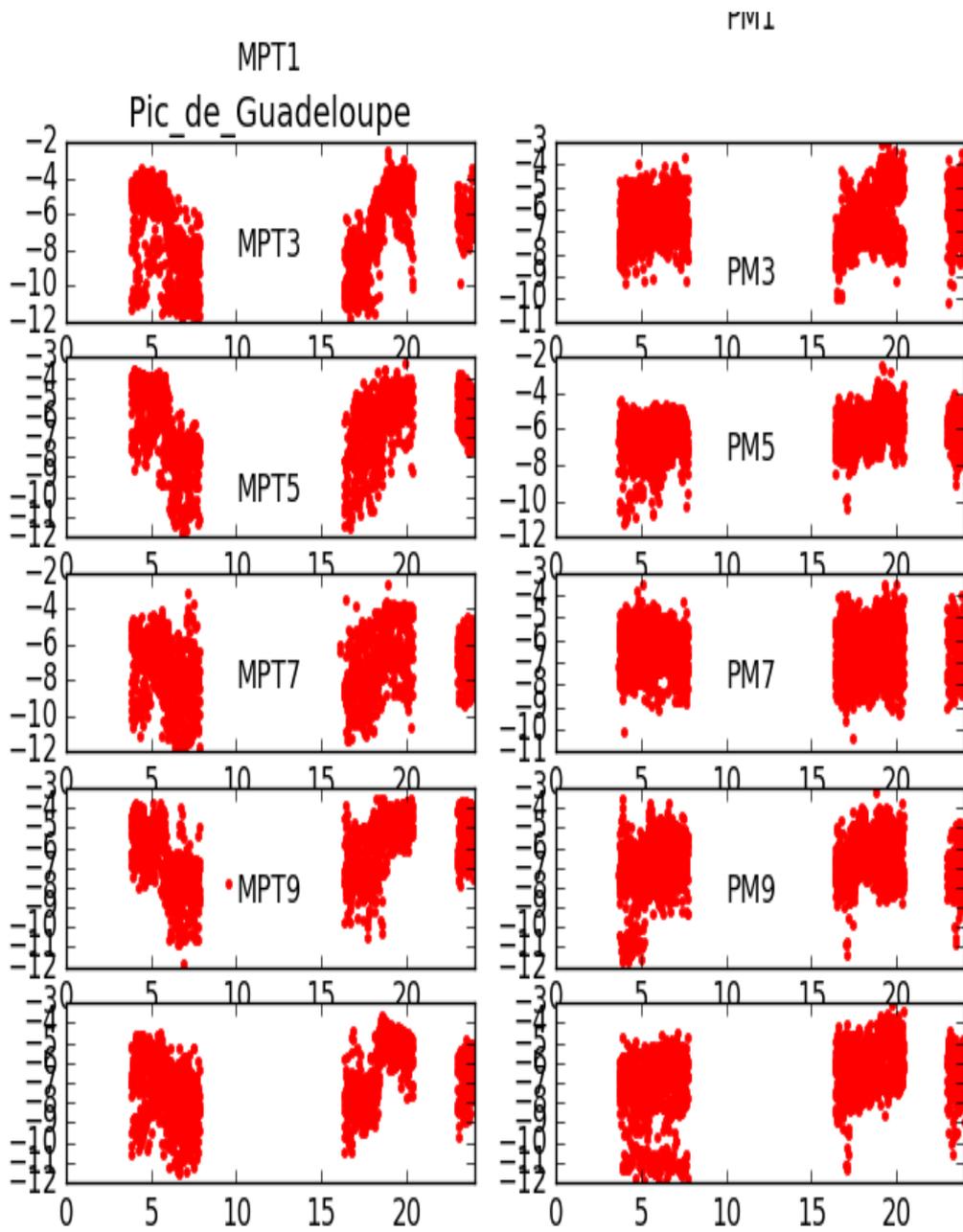


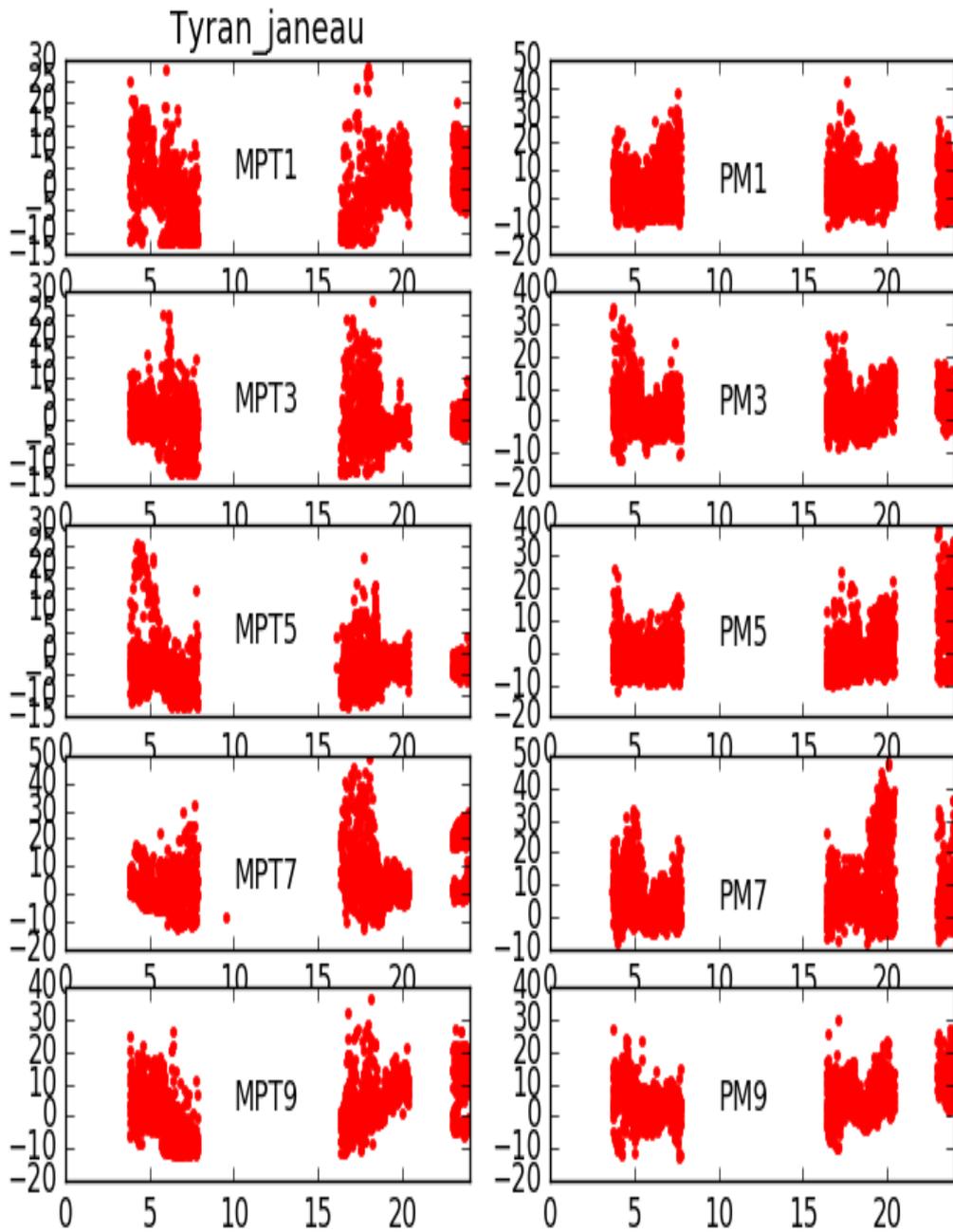


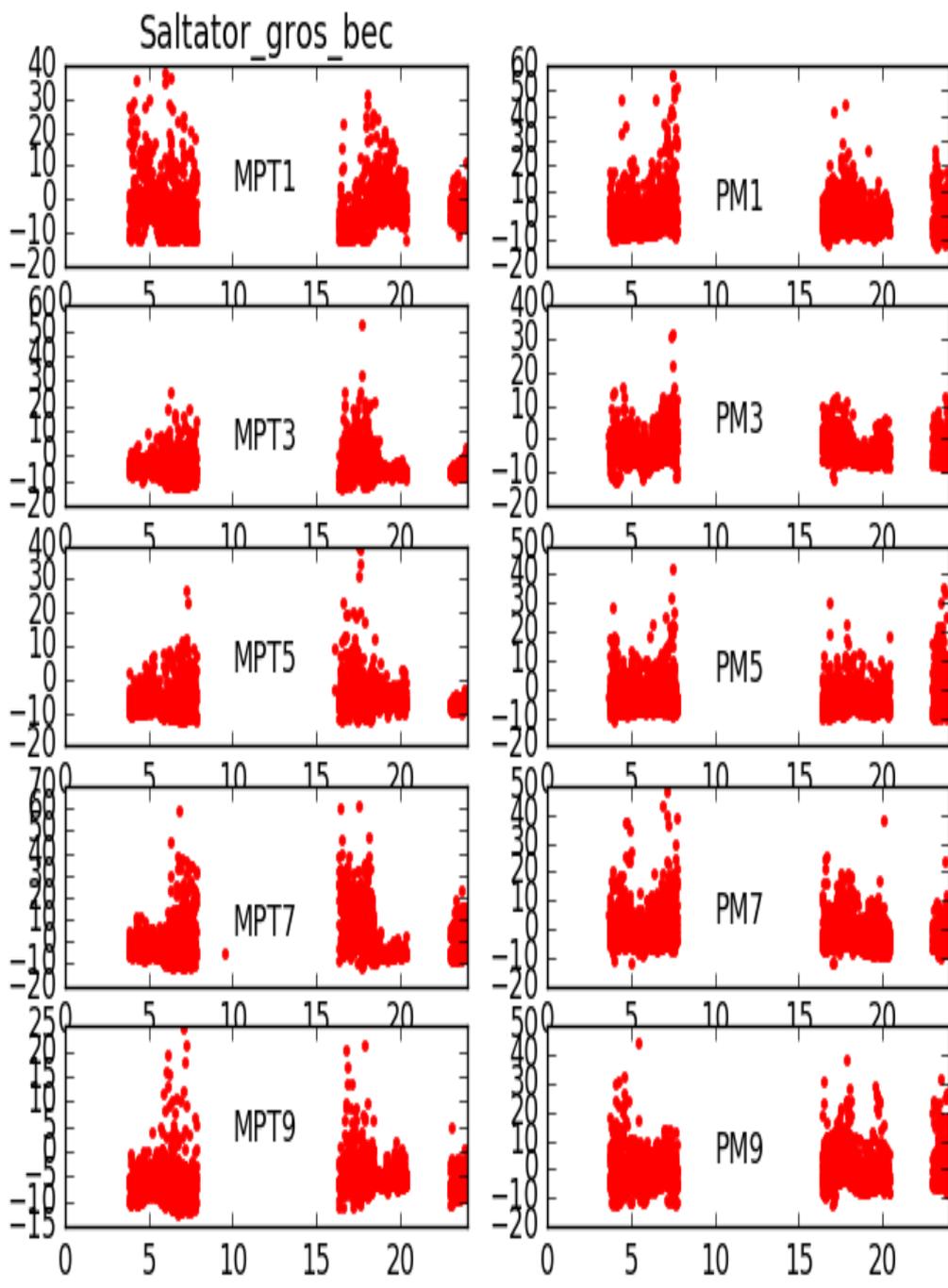
Colombe_rouviolette



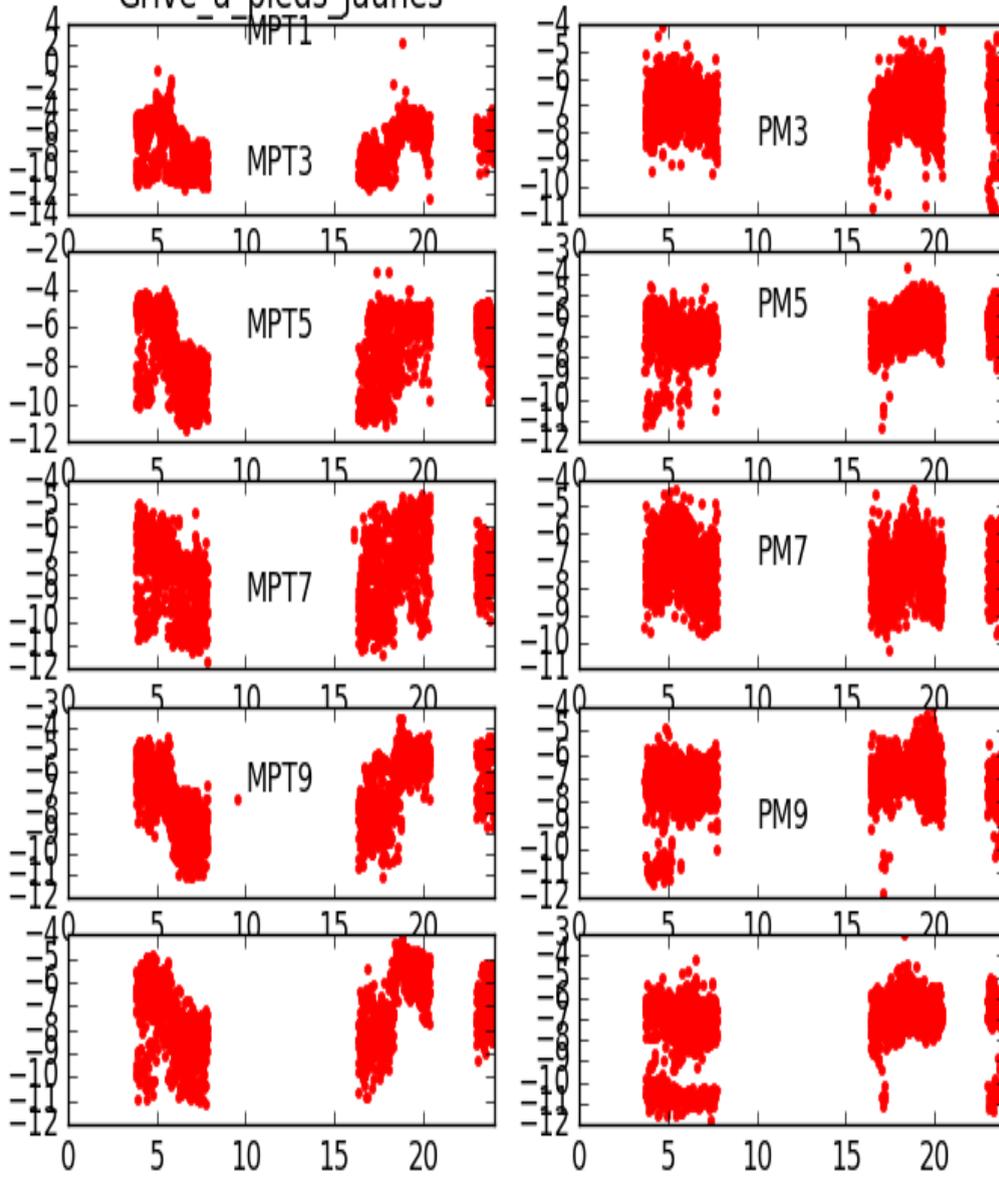








Grive à pieds jaunes



Références

- Agnieszka, O. Z. G. A. (2017). Scientific ideas included in the concepts of bioacoustics, acoustic ecology, ecoacoustics, soundscape ecology, and vibroacoustics. *Archives of Acoustics*, 42(3), 415-421.
- Arnoux, E., Eraud, C., Navarro, N., Tougard, C., Thomas, A., Cavallo, F., et al (2014). Morphology and genetics reveal an intriguing pattern of differentiation at a very small geographic scale in a bird species, the forest thrush *Turdus lherminieri*. *Heredity*, 113(6), 514-525.
- Blackman, R. B., & Tukey, J. W. (1958). The measurement of power spectra Dover Publications. Inc, New York.
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010* (pp. 177-186). Physica-Verlag HD.
- Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7), 1145-1159.
- Best P (2022). Classification automatique bioacoustique, application au suivi des chants de mammifères marins, Thèse de doctorat. ED548, Univ Toulon, dir. H. Glotin, R. Marxer, S. Paris.
- Claessens, O., Conde, B., Nicolas, L. (2015). un programme de suivi des oiseaux communs pour les départements d'outre-mer : adaptations du protocole stoc-eps aux environnements tropicaux, protocole, STOC-EPS. *Alauda*, 83 (4), 273-284.
- Cortes, C., Mohri, M., & Rostamizadeh, A. (2012). L2 regularization for learning kernels. *arXiv preprint arXiv:1205.2653*.
- Deco, C. S. G., & Brauer, W. (1997). Two strategies to avoid overfitting in feedforward neural networks. *Neural Networks*, 10(3), 505-516.
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *science*, 345(6195), 401-406.
- Dormoy, C. (2003). 45 indicateurs de développement durable : une contribution de l'Ifen. *Alauda*, 83, 273-284.
- Egan, J. P., & Egan, J. P. (1975). *Signal detection theory and ROC-analysis*. Academic press.
- Eraud, C., Levesque, A., Van Laere, G., & Magnin, H. (2013). La Grive à pieds jaunes (*Turdus lherminieri*) en Guadeloupe: Etat des connaissances sur l'importance et la répartition des effectifs.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.

- Fowler, W. W. (1896). The Evolution of Bird-song. *Nature*, 54(1396), 290-291.
- Gargominy, O. (2003). Biodiversité et conservation dans les collectivités françaises d'outre-mer. IUCN: International Union for Conservation of Nature.
- Glotin, H., LeCun, Y., Artieres, T., Mallat, S., Tchernichovski, O., & Halkias, X. (2013). Neural information processing scaled for bioacoustics, from neurons to big data. *Proceedings of Neural Information Processing Scaled for Bioacoustics: from Neurons to Big Data, 2013*.
- Grill, T., & Schlüter, J. (2017). Two convolutional neural networks for bird detection in audio signals. In *2017 25th European Signal Processing Conference (EUSIPCO)*, IEEE, 1764-1768.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.
- Huang, G., Li, Y., Pleiss, G., Liu, Z., Hopcroft, J. E., & Weinberger, K. Q. (2017). Snapshot ensembles: Train 1, get m for free. *arXiv preprint arXiv:1704.00109*.
- Hubbard, B. B. (1995). *Ondes et ondelettes: la saga d'un outil mathématique*. Belin: Pour la science.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, PMLR, . 448-456.
- Joly, H Goëau, S Kahl, L Picek, T Lorieul, E Cole, B Deneu, ...Glotin... (2021) Overview of lifeclef 2021: an evaluation of machine-learning based species identification and species distribution prediction International Conference of the Cross-Language Evaluation Forum.
- Julliard, R., & Jiguet, F. (2002). Integrated monitoring of common bird species in France. *Alauda*, 70, 137-147.
- Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT express*, 6(4), 312-315.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kulagin, Vladimir & Akimov, D & Pavelyev, S & Potapov, D. (2020). Automated Identification of Critical Malfunctions of Aircraft Engines Based on Modified Wavelet Transform and Deep Neural Network Clustering. *IOP Conference Series: Materials Science and Engineering*.
- Laiolo, P. (2010). The emerging significance of bioacoustics in animal species conservation. *Biological conservation*, 143(7), 1635-1645.
- Laland, K. N., Kendal, J. R., & Kendal, R. L. (2009). Animal culture: problems and solutions. *The question of animal culture*, 174-197.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

- Leroy, Y. (1979). *L'univers sonore animal*. Gauthier-Villars, Paris.
- Levesque, A., Mathurin, A., & La Diotte, L. (2008). Les zones importantes pour la conservation des oiseaux en Guadeloupe. *Rapport AMAZONA n, 17*, 31.
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proc. Icml*, 30, 1, p. 3.
- Massary, J. C., Bochaton, C., Dewynter, M., Frétey, T., Ineich, I., Lorvelec, O., ... & Lescure, J. (2021). Liste taxinomique de l'herpétofaune dans l'outre-mer français: V. Département de la Guadeloupe. *Bull. Soc. Herp. Fr*, 178, 06-23.
- Nielsen, M. A. (2015). *Neural Networks and Deep Learning, Determination Press*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Poupard, M., Ferrari, M., Best, P., & Glotin, H. (2022). Passive acoustic monitoring of sperm whales and anthropogenic noise using stereophonic recordings in the Mediterranean Sea, North West Pelagos Sanctuary. *Scientific reports*, 12(1), 1-13.
- Schütze, H., Manning, C. D., & Raghavan, P. (2008). *Introduction to information retrieval* (Vol. 39, pp. 234-265). Cambridge: Cambridge University Press.
- Sejdić, E., Djurović, I., & Jiang, J. (2009). Time–frequency feature representation using energy concentration: An overview of recent advances. *Digital signal processing*, 19(1), 153-183.
- Sevilla, H Glotin, Audio bird classification with inception-v4 extended with time and time-frequency attention mechanisms, WINNER of LIFECLEF Challenge, Working Notes of CLEF 1866, ceur-ws.org/Vol-1866/paper_177.pdf, 2017.
- Shentov, O. V., Mitra, S. K., Heute, U., & Hossen, A. N. (1995). Subband DFT—Part I: Definition, interpretation and extensions. *Signal Processing*, 41(3), 261-277.
- Slabbekoorn, H., & Smith, T. B. (2002). Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1420), 493-503.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- Stowell, MD Wood, H Pamuła, Y Stylianou, H Glotin (2019), Automatic acoustic detection of birds through deep learning: the first bird audio detection challenge, *Methods in Ecology and Evolution* 10 (3), 368-380.

- Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological conservation*, *142*(10), 2282-2292.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285-1293.
- UICN France., MNHN, S., & SEOF, O. (2020). La Liste rouge des espèces menacées en France.
- Vellinga, W. P., & Planqué, R. (2015). The Xeno-canto Collection and its Relation to Sound Recognition and Classification. In *CLEF*.
- Wimmer, J., Towsey, M., Roe, P., & Williamson, I. (2013). Sampling environmental acoustic recordings to determine bird species richness. *Ecological Applications*, *23*(6), 1419-1428.
- Xie, J., Hu, K., Zhu, M., Yu, J., & Zhu, Q. (2019). Investigation of different CNN-based models for improved bird sound classification. *IEEE Access*, *7*, 175353-175361.
- Žilinskas, A. (2006). Practical mathematical optimization: An introduction to basic optimization theory and classical and new gradient-based algorithms.