

## MESURE DE L'ENVIRONNEMENT SONORE

Malgré les nombreuses expériences qui ont été menées sur la qualité sonore, très peu d'études ont été réalisées sur la pertinence des échantillonnages choisis en temps et en espace. Ainsi, les échantillonnages temporels se sont appuyés sur des écoutes allant de 10 secondes [CERMAK 1976] jusqu'à 30 minutes [KUWANO 1988]. Plus récemment, les études portant sur le "soundscape" se sont basées sur des mesures de durée assez variable, c'est-à-dire 1 minute [MA 2010], 5 minutes [JEDOVNICKY 2009] ou encore 15 minutes [DE COENSEL 2010]. Les études sur la perception du niveau sonore se sont basées sur des stimuli allant de quelques secondes [LITTLE 1969] [PAULSEN 1997] jusqu'à 80 minutes [NAMBA 1988]. Un travail exploratoire sur l'influence de l'échelle temporelle sur la pertinence des indicateurs a été mené au Laboratoire de géographie Mobilités, Réseaux, Territoires, Environnements (MRTE) de l'Université de Cergy-Pontoise [LAVANDIER 2003]. Dans cette étude, LAVANDIER C. et BARBOT B. ont mesuré l'environnement sonore en cinq lieux (un jardin, une rue piétonne, un marché couvert et deux boulevards) durant 24 heures, et ont tenté de déterminer quelle était la durée de mesure appropriée pour discriminer les différents lieux, en fonction de différents indicateurs tels que le  $L_{Aeq}$ , la sonie de Zwicker, des émergences ou encore des indicateurs psycho-acoustiques comme l'acuité ("sharpness"), la rugosité("roughness") et la force de fluctuation("fluctuation strength"). De l'analyse en composantes principales, il est ressorti que le calcul d'indicateurs classiques tels que le  $L_{Aeq}$ , la sonie ou encore l'écart-type du  $L_{Aeq}$  sur 15 minutes permettait de caractériser objectivement un lieu et que ce calcul devait être fait sur cinq périodes de la journée : la nuit, au réveil de la ville, le matin, l'après-midi et enfin sur la période d'endormissement de la ville. En revanche, les indicateurs psycho-acoustiques comme l'acuité et la rugosité n'ont pas été pertinents et n'ont pas apporté d'information permettant de discriminer les lieux.

Quant au choix du maillage spatial nécessaire à la caractérisation des lieux, il a suivi plutôt les typologies architecturales des lieux sans en étudier une justification acoustique [JEON 2010] [MARRY 2010] [SEMIDOR 2010].

Dans le travail présenté ici, la pertinence des choix temporels et spatiaux pour les mesures acoustiques en site urbain va être systématiquement étudiée, afin de fixer les durées, les moments et les lieux nécessaires aux enquêtes perceptives. Le but de ce double maillage,

temporel et spatial, est donc d'optimiser les campagnes de mesures. Quelle est la durée nécessaire pour caractériser l'environnement sonore d'un lieu afin qu'elle ne soit ni trop longue pour éviter de perdre du temps au cours des mesures, ni trop courte pour éviter de perdre de l'information acoustique ? Quelle est aussi la distance à respecter entre les points de mesures afin de caractériser des lieux appartenant à des zones différentes d'un point de vue acoustique ?

## 2.2 CARACTERISATION SUBJECTIVE DU PAYSAGE SONORE

Dans leur étude évoquée dans l'introduction [BOTTELDOOREN 2011], le réseau de recherche COST TD0804 "Soundscape of European Cities and Landscapes" a également recensé différents facteurs perceptifs qui caractérisent le paysage sonore.

Parmi eux se trouve un premier groupe d'indicateurs qui sont des descripteurs verbaux globaux et tentent de décrire l'expérience sonore dans son ensemble. Des attributs globaux comme apaisant et agréable ou gênant et bruyant ont été utilisés [BERGLUND 2006]. Lors d'une étude en composantes principales [CAIN 2010], ces nombreux multiples attributs ont pu être regroupés en deux axes étiquetés "Calme" ("Calmness") et "Vitalité" ("Vibrancy"). Plus généralement, ces deux axes sont nommés "Agrément" ("Pleasantness") et "Animation" ("Eventfulness") comme le précisent AXELSSON O. *et al.*. A ces deux dimensions, vient s'ajouter une troisième composante nommée "Familiarité" [AXELSSON 2010]. Ces résultats sont concordants avec ceux trouvés par VIOLLON S. qui avait également trouvé trois axes nommés alors "Affective impressions, preferences" ("impressions affectives, préférences") pour le premier et corrélé à 98% avec l'adjectif "pleasant" ("agréable"), "Activity due to sound presence of human beings" ("activité due à la présence sonore des comportements humains") pour le second, corrélé avec l'adjectif "bustling" ("animé") et enfin "Auditory expectations" ("attentes auditives") pour le troisième, corrélé avec l'adjectif "unexpected" (inattendu) [VIOLLON 2000a].

Cependant, ces descripteurs globaux de l'environnement sonore ne donnent que peu d'informations sur la description du paysage sonore. De plus, ils ne sont valides que pour un groupe particulier de personnes et dans le contexte particulier de l'étude. Aussi des descriptions plus détaillées sont nécessaires pour passer entre les différents contextes et sous-groupes de populations.

Un autre moyen de décrire les paysages sonores est de se focaliser sur les sources sonores qui composent l'environnement [DUBOIS 2006]. Dans leur étude [LAVANDIER 2006] portant sur la contribution des caractéristiques des sources sonores sur le jugement du paysage sonore urbain, LAVANDIER C. et DEFREVILLE B. ont décrit les sources sonores avec des attributs tels que "Proximité", "Prégnance" et "Présence" en considérant les travaux de WOLOSZYN P. et SIRET D. [WOLOSZYN 1998]. Ces attributs étant respectivement reliés à l'espace, à l'énergie et au temps.

Outre les descripteurs holistiques ou analytiques, une façon de décrire un "soundscape" est d'utiliser une méthode narrative [SCHULTE-FORTKAMP 2006]. Dans ces conditions, le "soundscape" est vécu à travers des séquences d'écoute ciblée de l'environnement sonore auxquelles se mêlent des moments d'écoute plus globale. Ainsi, la narration fournit essentiellement le sens que les personnes donnent aux différentes composantes de l'environnement sonore. Une dimension spatiale peut également être ajoutée à travers un parcours commenté, en combinant la description narrative de l'environnement sonore avec un parcours au sein de celui-ci [SEMIDOR 2010].

Enfin, alors que les méthodes narratives tentent d'évoquer les "soundscapes" par des descriptions verbales (écrites ou orales), un relevé acoustique ou sonore essaie plutôt d'enregistrer un environnement sonore typique ou non et de le présenter avec une description adéquate du contexte afin de suggérer le "soundscape" dans l'esprit des personnes qui l'écoutent.

### 2.3 CARACTERISATION OBJECTIVE DU PAYSAGE SONORE

Les indicateurs décrits ci-dessus s'appuient sur des mesures perceptives relatives aux êtres humains. Il existe beaucoup de situations où des indicateurs basés sur les enregistrements acoustiques ou des simulations sont les seules possibilités : surveillance, création de cartes, *etc.* Des mesures globales de l'environnement sonore telles que le  $L_{Aeq}$ , le  $L_{A50}$  ou  $L_{A90}$ , des mesures spectrales ou encore de dynamique (écart-type du  $L_{Aeq}$  par exemple) sont alors nécessaires. A première vue, ces mesures globales ne traduisent pas la présence de sources sonores particulières. Cependant, compte tenu d'un contexte particulier, ces indicateurs globaux peuvent témoigner de la présence de sources caractéristiques. Par exemple, les hautes fréquences dans les parcs urbains pourraient indiquer la présence d'oiseaux et de sons humains alors que le  $L_{A90}$  sera représentatif du trafic lent lointain. L'interprétation de ces

mesures globales peut donc fournir des informations importantes sur la nature des sources présentes et permettre une meilleure compréhension de l'environnement sonore étudié.

Enfin des mesures analytiques de l'environnement sonore devraient mieux correspondre aux évènements présents dans l'environnement sonore que les mesures holistiques. Toutefois, ces indicateurs sont parfois très difficiles à extraire du signal global capté par le microphone et une reconnaissance "à l'oreille" des sources sonores dans un flux continu s'avère souvent être un travail fastidieux. Bien que toujours en développement, des outils de reconnaissance automatique des sources peuvent aider à identifier l'apparition de sons caractéristiques mêlés à un flux sonore global dans un enregistrement long. Quoiqu'il en soit, des indicateurs comme le nombre d'évènements, le temps de présence de ces évènements ou le niveau sonore durant lequel ces évènements sont émergents sont nécessaires à la caractérisation du paysage sonore. Ainsi, le temps de présence et le niveau de plusieurs sources prédominantes dans un environnement sonore sont au cœur d'un outil de description d'un "soundscape" développé par MATSUI T.*et al.* [MATSUI 2009]. Le concept de "Time-Component Matrix Chart" (ou "TM Chart" en abrégé), dont une représentation est donnée Figure 2, consiste à représenter visuellement dans un espace à deux dimensions, le pourcentage de temps de présence de plusieurs sources prédominantes et les niveaux de celles-ci ramenés en pourcentage de temps d'apparition. Ainsi, plusieurs environnements sonores peuvent être comparés à travers leur représentation en "TM Chart".

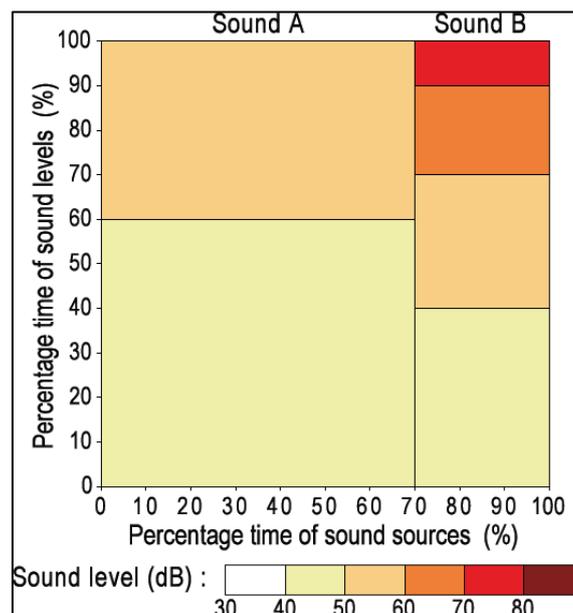


Figure 2 : Exemple de Time-Component Matrix Chart (TM Chart).

Qu'ils s'agissent d'indicateurs globaux ou plus spécifiques aux sources, quoiqu'il en soit, caractériser l'environnement sonore nécessite donc de toute évidence des moyens de le mesurer de façon objective. Dans le cadre du projet de recherche "ANR Eval-PDU" et en particulier dans le cadre de la tâche 4 sur le développement d'une méthodologie d'évaluation de l'impact d'un PDU (Plan de Déplacement Urbain) sur le bruit, PICAUT J. a rédigé une synthèse bibliographique sur l'évaluation du bruit dans l'environnement [PICAUT 2011]. Dans sa version actuelle 1.3 datant du 21 février 2011, ce document présente une synthèse d'indicateurs physiques pour la caractérisation du bruit dans l'environnement de manière générale, ainsi que d'autres plus spécifiques sur la caractérisation du bruit routier, du bruit des avions et du bruit ferroviaire. Notre étude ne portant pas sur le trafic aérien ou ferroviaire, nous ne citerons que les indicateurs spécifiques liés au trafic routier.

Issus de la communauté scientifique et de la normalisation, notamment [NF S 31-010 1996] [NF S 31-085 2002], on trouve parmi les indicateurs présentés dans cette synthèse, des indicateurs de niveaux acoustiques, des indicateurs d'émergence, des indicateurs de bruit fluctuant et enfin des indicateurs spécifiques au bruit routier.

---

### 2.3.1 INDICATEURS DE NIVEAUX ACOUSTIQUES

#### Niveau de pression acoustique continu équivalent ( $L_{eq,T}$ )

Le niveau de pression acoustique équivalent ( $L_{eq,T}$ ) se définit comme la valeur du niveau de pression acoustique d'un son continu stable qui, au cours d'une période spécifiée  $T = t_2 - t_1$ , a la même pression acoustique quadratique moyenne qu'un son considéré dont le niveau varie au cours du temps. Mathématiquement, ce niveau de pression acoustique continu équivalent s'écrit comme suit :

$$L_{eq,T} = 10 \log \left[ \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{P^2(t)}{P_0^2} dt \right] \text{ où } P_0 \text{ est la pression acoustique de référence (20 } \mu\text{Pa) et}$$

$P(t)$  est la pression acoustique instantanée du signal.

#### Niveau de pression acoustique continu équivalent pondéré ( $L_{Aeq,T}$ )

Si la pression acoustique est pondérée en fréquence par une pondération par exemple A, le niveau acoustique continu équivalent pondéré A, noté  $L_{Aeq,T}$  s'écrit alors suivant la relation

$$L_{Aeq,T} = 10 \log \left[ \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{P_A^2(t)}{P_0^2} dt \right] \text{ où } P_A(t) \text{ est la pression acoustique instantanée pondérée A}$$

du signal.

### Niveau de pression acoustique continu équivalent "court" ( $L_{Aeq,\tau}$ )

On parle également d'un niveau de pression acoustique continu équivalent court, noté  $L_{Aeq,\tau}$ , qui se définit comme le niveau de pression acoustique continu équivalent sur un intervalle de temps  $T$ , plus court, noté  $\tau$ . La durée d'intégration dépend de ce que l'on cherche à mesurer mais est généralement de 1 s.

### Niveau de pression acoustique continu équivalent pondéré A d'un bruit particulier ( $L_{Aeq,Tpart}$ )

A partir du calcul de  $L_{eq}$  courts (de même durée) il est possible dans le cas d'un bruit particulier, de calculer le niveau continu équivalent pondéré A du bruit équivalent sur une durée  $T_{part}$  d'apparition du bruit particulier, à partir du calcul de  $L_{eq}$  courts (de même

durée): 
$$L_{Aeq,Tpart} = 10 \log \left[ \frac{1}{T_{part}} \sum_{i=1}^N \tau \times 10^{\frac{L_{Aeq,\tau}}{10}} \right].$$

### Valeurs maximale et minimale du niveau de pression acoustique pondéré A

Les indicateurs  $L_{Amax}$  et  $L_{Amin}$  correspondent respectivement au niveau maximum et minimum de la pression acoustique pondérée A mesurée.

### Niveaux fractiles

Les niveaux acoustiques fractiles (ou indices fractiles) permettent à partir d'une analyse statistique des niveaux de pression acoustique continu équivalent "courts" (en général  $L_{Aeq,1s}$ ) de mesurer les niveaux de pression acoustique dépassé N% du temps considéré et sont notés  $L_{AN}$  lorsque la pondération A est utilisée. Les plus courants sont le  $L_{A95}$ , le  $L_{A90}$ , le  $L_{A50}$ , le  $L_{A10}$  et le  $L_{A5}$ . En règle générale, le  $L_{A10}$  représente une moyenne des niveaux sonores maximaux alors que le  $L_{A95}$  ou le  $L_{A90}$  symbolisent le bruit de fond.

### Ecart type

L'écart type des niveaux de pression continu équivalent "court"  $L_{Aeq,1s}$  permet de caractériser l'amplitude des variations des  $L_{Aeq,1s}$  sur une durée donnée et se détermine par la

formule suivante : 
$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (L_{Aeq,1s} - \bar{L}_{Aeq,1s})^2}$$
 avec  $\bar{L}_{Aeq,1s}$  qui est la moyenne arithmétique des  $L_{Aeq,1s}$ .

---

### 2.3.2 INDICATEURS D'EMERGENCE

Dans la synthèse bibliographique sur les indicateurs du bruit dans l'environnement, deux indicateurs d'évènements sont proposés. Le premier consiste à comparer le niveau de pression acoustique continu équivalent pondéré A  $L_{Aeq,T_{part}}$  du bruit ambiant en présence d'un bruit particulier (le niveau est déterminé pendant les périodes d'apparition du bruit particulier, sur une durée cumulée  $T_{part}$ ), avec le niveau de pression acoustique continu équivalent pondéré A  $L_{Aeq,T_{res}}$  du bruit résiduel (ce niveau est déterminé pendant les périodes de disparition du bruit particulier, sur une durée cumulée  $T_{res}$ ). Le second mesure quant à lui la différence entre le niveau sonore maximal  $L_{Aeq,5s,max}$  mesuré sur une période de 5 s pendant un évènement et le  $L_{A50}$  calculé en prenant en compte le bruit ambiant observé avant l'apparition de l'évènement concerné, sur une période de 5 minutes.

Dans notre cas, nous avons choisi une autre façon de calculer l'émergence, à savoir la différence entre le niveau de pression acoustique dépassé 10% du temps du bruit particulier pondéré A  $L_{A10,T_{part}}$  et le niveau de pression acoustique continu équivalent pondéré A  $L_{Aeq}$  global mesuré sur la durée de la mesure. Cette émergence s'écrit donc  $E = L_{A10,T_{part}} - L_{Aeq}$ .

---

### 2.3.3 INDICATEURS SPECIFIQUES AU BRUIT ROUTIER

Dans le cas particulier du bruit routier, hormis la plupart des indicateurs cités précédemment, deux autres indicateurs peuvent être utilisés.

#### Traffic noise index

Le traffic noise index ou TNI, basé sur les indices fractiles  $L_{10}$  et  $L_{90}$  permet d'évaluer la part des fluctuations de bruit liée au trafic routier et se calcule suivant la formule suivante :

$$TNI = 4(L_{10} - L_{90}) + L_{90} - 30.$$

#### Noise pollution level

Le noise pollution level  $L_{NP}$  ou NPL est quant à lui basé sur l'écart-type  $\sigma$  des  $L_{Aeq,1s}$  et s'exprime comme suit :  $L_{NP} = L_{Aeq} + \sigma$ .

### 2.3.4 INDICATEURS DE BRUIT FLUCTUANT

Outre les indicateurs de niveaux sonores et d'émergence, il existe des indicateurs dits "dynamiques" permettant de quantifier un certain nombre de caractéristiques temporelles, en particulier concernant le trafic routier en milieu urbain.

Une liste non exhaustive de ces indicateurs peut être trouvée dans le projet de recherche [LECLERQ 2008]. Citons par exemple des indicateurs comme l'étalement sonore, le rythme sonore ou encore la tranquillité. Au moment où nous avons commencé notre étude nous n'avions pas accès à ce projet de recherche. En conséquence de quoi, ces indicateurs de bruit fluctuant n'ont pas été utilisés pour la discrimination temporelle et spatiale (chapitres 3 et 4). Toutefois, MARRY S. a montré à travers une analyse en composantes principales portant sur 174 évaluations de différentes ambiances sonores que des indicateurs comme IDAS (Degré de l'augmentation sonore) par exemple étaient très corrélés avec l'écart-type du  $L_{Aeq}$  ou encore avec le  $L_{A10}$ - $L_{A90}$  [MARRY 2011]. Par ailleurs, les indicateurs de rythme ou d'étalement sonore n'apportent que très peu d'information par rapport à des indicateurs plus classiques comme le  $L_{Aeq}$ , le  $L_{A10}$  ou le  $L_{A90}$ .

Aussi, dans l'étude du maillage temporel (§ 3) et du maillage spatial (§ 4) nous avons principalement utilisé des indicateurs de niveaux acoustiques, d'émergence ainsi que des indicateurs de bruyance (dont la définition est donnée ci-après).

#### Bruyance

Les indicateurs de bruyance permettent de quantifier le temps et le nombre de fois où une séquence sonore dépasse un seuil  $\alpha$ . Afin de distinguer les pics de bruit (impulsionnels) et les séquences sonores, ces dernières doivent avoir une durée d'au moins 4 secondes. Notons également que les indicateurs de bruyance sont définis pour une période d'évaluation de 15 minutes. Ainsi quatre indicateurs peuvent être calculés. Le  $N_{>L_1}$  est le nombre de séquences d'au moins 4 secondes dont le niveau sonore dépasse  $L_1$ , évalué sur une période de 15 minutes de  $L_{Aeq,1s}$ . Le  $N_{>70}$  est le nombre de séquences d'au moins 4 secondes dont le niveau sonore dépasse 70 dB(A), évalué sur une période de 15 minutes de  $L_{Aeq,1s}$ . Enfin le  $T_{>L_1}$  et le  $T_{>70}$  sont respectivement la durée moyenne en seconde de ces séquences.

## 2.4 QUALITE DES INDICATEURS

Le but de ce travail de recherche est d'établir des indicateurs de qualité sonore en situation urbaine en s'appuyant sur deux modèles d'élaboration (les régressions linéaires multiples et les réseaux de neurones artificiels). Outre les avantages et les inconvénients liés à ces deux méthodes d'élaboration, il est important de déterminer les qualités que doivent présenter ces modèles.

Dans le rapport final sur les "Indicateurs de durabilité environnementale dans les transports" [JOURMARD 2010], le groupe de recherche COST 356 "Towards the definition of a measurable Environmentally Sustainable Transport" fournit une liste de 14 propriétés que doivent posséder des indicateurs dans les transports et dans l'environnement. Parmi ces 14 propriétés, une liste plus sélective de 10 critères est également donnée. Cette seconde liste ainsi que la définition associée à chaque critère est présentée ci-après en français et dans leur version originale en anglais.

### Validité (Validity) :

Un indicateur valide doit mesurer réellement le résultat ou le facteur qu'il est supposé mesurer (A valid indicator must actually measure the issue or factor it is supposed to measure).

### Fiabilité (Reliability) :

Un indicateur fiable doit donner la même valeur si sa mesure est répétée de la même façon sur la même population et presque en même temps (A reliable indicator must give the same value if its measurement were repeated in the same way on the same population and at almost the same time).

### Sensibilité (Sensitivity) :

Un indicateur sensible doit être en mesure de révéler des changements importants dans le facteur d'intérêt (A sensitive indicator must be able to reveal important changes in the factor of interest).

### Mesurabilité (Measurability) :

Un indicateur mesurable devrait être simple et relativement peu coûteux de mesurer (A measurable indicator should be straight-forward and relatively inexpensive to measure).

Disponibilité des données (Data availability) :

Les indicateurs avec des données disponibles sont des indicateurs fondés sur les données (entrées) qui devraient être facilement accessibles ou disponibles avec un temps et un coût raisonnable (Data available indicators are Indicators based on (input) data that should be readily available or made available at reasonable cost and time).

Préoccupation éthique (Ethical concerns) :

Un indicateur doit respecter les droits fondamentaux humains et doit exiger seulement des données qui sont compatibles avec les mœurs, les croyances ou les valeurs de la population (An indicator must comply with fundamental human rights and must require only data that are consistent with morals, beliefs or values of the population).

Transparence (Transparency) :

Un indicateur transparent est celui qu'il est possible de comprendre et possible de reproduire pour des utilisateurs qui le souhaitent (A transparent indicator is one which is feasible to understand and possible to reproduce for intended users).

Interprétabilité (Interpretability) :

Un indicateur interprétable permet une lecture intuitive et sans ambiguïté. (An interpretable indicator allows an intuitive and unambiguous reading).

Pertinence (Target relevance) :

Un indicateur pertinent doit mesurer la performance en accord avec les objectifs énoncés. (A target relevant indicator must measure performance with regard to articulated goals objectives, targets or thresholds).

Faisabilité (Actionability) :

Un indicateur faisable est celui qui mesure des facteurs qui peuvent être modifiés ou influencés directement par l'action politique ou de management (An actionable indicator is one which measure factors that can be changed or influenced directly by management or policy action).

Ainsi, dans la suite du travail, lorsqu'un indicateur de qualité sonore aura été développé, nous nous attacherons à regarder dans quelle mesure cet indicateur respecte chacun des paramètres évoqués ci-dessus (cf. § 5.3).

## 2.5 METHODES D'ANALYSES

Au cours des travaux de recherche qui sont présentés dans ce document, diverses méthodes d'analyses des données ont été utilisées. Sans systématiquement entrer en détails dans le formalisme mathématique qui se cache derrière chacune de ces méthodes, une description plus ou moins avancée de chacune d'elles est donnée dans les paragraphes suivants (§ 2.5.1 à 2.5.8).

### 2.5.1 CLASSIFICATION ASCENDANTE HIERARCHIQUE

La classification ascendante hiérarchique est une méthode de classification fréquemment utilisée qui permet de regrouper  $N$  individus en un nombre de regroupements ou classes plus ou moins réduit. Ces  $N$  individus (qui forment initialement  $N$  groupes) sont successivement réunis en  $N-1$  groupes jusqu'à ce qu'il ne reste plus qu'un seul groupe de  $N$  individus.

La première étape dans l'algorithme de cette méthode consiste à définir une mesure de similarité qui caractérise la distance entre les objets à classer. Puis, une fois cette distance choisie, il s'agit ensuite de définir un critère d'agrégation afin de regrouper les objets les plus proches dans une nouvelle classe.

Au cours de cette étude, la mesure de similarité, ou plus exactement, la mesure de "dissimilarité" est simplement la distance euclidienne entre les objets. Soient deux objets  $p$  et  $q$  caractérisés par  $n$  variables, la distance euclidienne entre ces deux objets est donnée par la formule suivante :  $d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$ .

Afin de réunir les objets entre eux il existe de nombreux critères d'agrégation ("saut minimal ou plus petite distance", "saut maximal ou plus grande distance", "distance moyenne", "Ward", etc. ...) [NAKACHE 2004]. Un des plus couramment utilisés est le critère de Ward. Il s'agit d'un critère d'agrégation basé sur l'inertie calculée à chaque étape du processus d'agrégation et revient à minimiser l'inertie intra-classes ou à maximiser l'inertie inter-classes.

Une fois la classification de Ward réalisée, se pose généralement la question du choix du nombre de classes à retenir. Dans cette optique il n'existe pas de règle précise mais de nombreux critères peuvent aider ce choix. Tout d'abord, l'arbre hiérarchique qui représente les regroupements à chaque étape peut donner une première idée du nombre de classes à choisir, même si cela n'est pas toujours pertinent. Il existe alors de nombreux indices utiles à la décision. Parmi tous ces indices, quatre sont basés sur le calcul de la somme des carrés des écarts entre les objets et la moyenne de la classe à laquelle ils appartiennent. Il s'agit du rapport entre l'inertie inter-classes et l'inertie totale  $\left(\frac{I_{\text{Inter}}}{I_{\text{Totale}}}\right)$ , du "R-Square" (RSQ), du "Semi-Partial R-Square" (SPRSQ) et du "Root-Mean-Square Standard Deviation" (RMSSTD). Enfin, un autre indice est aussi calculé : le "Quality Index" (QI) basé sur la largeur moyenne des silhouettes. Le calcul de ces cinq indices est donné ci-après.

Supposons que lors de la classification  $s$  classes soient formées à l'étape  $h$  de l'agrégation, alors l'étape suivante  $h+1$  fournira donc  $s-1$  classes. Soient  $n$  objets caractérisés par  $p$  variables,  $X_i$  est le  $i^{\text{ème}}$  objet et le vecteur des  $p$  moyennes sur les  $n$  objets est noté  $\bar{X}$ .

- $C_k$  est la  $k^{\text{ème}}$  classe parmi les  $s$  classes, composée de  $n_k$  objets et le vecteur des  $p$  moyennes dans la classe  $k$  est noté  $\bar{X}_k$ ,
- $W_k$  est la somme des carrés des différences des objets  $X_j$  à  $\bar{X}_k$  dans la classe  $k$  :  $W_k = \sum_{j=1}^{n_k} \|X_j - \bar{X}_k\|^2$ ,
- $T$  est la somme calculée pour les  $n$  objets :  $T = \sum_{i=1}^n \|X_i - \bar{X}\|^2$ .

Le calcul du "R-Square" (RSQ) est alors donnée par la formule suivante :  $RSQ = 1 - \frac{W_s}{T}$  où  $W_s = \sum_{i=1}^s W_k$ . Le "Semi-Partial R-Square" (SPRSQ) représente la perte d'homogénéité lorsque deux classes sont agrégées et est mesuré par la réduction du RSQ avant et après l'agrégation des classes  $C_k$  et  $C_{k'}$ . Enfin, le "Root-Mean-Square Standard Deviation" (RMSSTD) se calcule comme la racine carrée de la variance de la nouvelle classe  $C_{k''}$  formée

par l'agrégation des classes  $C_k$  et  $C_{k'}$ :  $RMSSTD = \sqrt{\frac{W_{k''}}{p(n_{k''}-1)}}$ .

Concernant le "Quality Index" (QI), celui-ci est basé sur les largeurs de silhouettes. La silhouette se calculant ainsi : pour chaque objet  $X_i$  on note  $A$  la classe à laquelle il appartient et  $N_A$  le nombre d'éléments dans cette classe  $A$ . Ainsi, on définit :

- $a(X_i) = \text{moy}(d(X_i, X_{i'}))$  la dissimilarité (distance euclidienne) moyenne entre l'objet  $X_i$  et tous les autres objets  $X_{i'}$  de  $A$ ,
- $c(X_i) = \text{moy}(d(X_i, X_{i''}))$  la dissimilarité moyenne entre l'objet  $X_i$  et tous les objets  $X_{i''}$  d'une classe  $C$  quelconque différente de  $A$ ,
- $B$  la classe telle que  $b(X_i) = \min(c(X_i))$  est la classe la plus voisine de l'objet  $X_i$  (2<sup>ème</sup> meilleure classe pour  $X_i$ ).

La largeur de silhouette se calcule alors de la façon suivante :  $s(X_i) = \frac{b(X_i) - a(X_i)}{\max(a(X_i), b(X_i))}$ . Si  $s(X_i) \sim 1$  alors  $X_i$  est bien classé (dans  $A$ ), si  $s(X_i) \sim 0$  alors  $X_i$  se situe entre les classes  $A$  et  $B$  et si  $s(X_i) \sim -1$  alors  $X_i$  est mal classé (plus proche de  $B$  que de  $A$ ).

Ainsi, le "Quality Index" (QI) est la moyenne globale des largeurs des silhouettes dans les différentes classes qui composent la partition. Si la partition est composée de  $k$  classes, alors  $QI(k) = \frac{1}{n} \sum_{r=1}^k n_r \bar{s}_r$ .

Le nombre de classes à retenir est observé par une "cassure" dans la représentation du ratio de l'inertie inter-classes, du RSQ et du SPRSQ. Le RMSSTD doit être minimal alors que le QI doit lui être maximal.

Il est à noter que bien que tous ces indices puissent se révéler fort utiles, l'interprétation des classes obtenues lors de la classification reste cependant un critère significatif quant au choix du nombre optimal de classes à retenir [NAKACHE 2004].

---

## 2.5.2 CLASSIFICATION DE KOHONEN

Outre la classification ascendante hiérarchique, une autre classification basée sur le modèle de Kohonen a été utilisée. Il s'agit d'une méthode de réseaux de neurones non supervisée [KOHONEN 2001]. L'algorithme de construction des cartes auto-organisatrices de Kohonen est très proche de celui des "k-means" mais la différence est que la topologie des données d'entrée est respectée. Ainsi, les objets les plus proches dans l'espace d'entrée le sont également sur la carte.

Considérons chaque objet  $p_i$  (c'est-à-dire une mesure) de dimension  $n$  comme une entrée et chaque neurone de la carte comme un vecteur de poids  $W_j$  avec la même dimension  $n$ . Il est alors nécessaire de définir :

- une mesure de distance entre les objets et les neurones. Dans notre cas il s'agit de la distance euclidienne,
- un nombre de neurones. Aucune règle n'existe quant au nombre de neurones à choisir, mais une solution consiste à utiliser la formule heuristique  $M = \sqrt{N}$  où  $M$  est le nombre de neurones à retenir et  $N$  est le nombre d'objets [VESANTO 1999],
- une topologie de carte (hexagonale dans le cadre de cette étude).

L'algorithme peut être décrit comme ceci : de façon aléatoire un objet  $p_i$  est présenté au réseau. Le neurone "g" de la carte, dont le vecteur poids est le plus proche de l'objet présenté en termes de distance euclidienne,  $g = \operatorname{argmin}_{j=1\dots M} d(p_i, W_j)$  est désigné comme le neurone gagnant, c'est-à-dire que c'est lui qui code l'objet. Les poids de ce neurone sont alors modifiés de telle sorte que la distance entre lui et l'objet présenté diminue. Les poids des neurones voisins sont également modifiés dans une moindre mesure. Puis les objets suivants sont présentés un à un et la même opération est répétée jusqu'à ce que les poids ne varient plus de façon significative.

---

### 2.5.3 CLASSIFICATION CONJOINTE

Une fois la classification de Kohonen réalisée, il s'avère parfois très utile d'effectuer une classification ascendante hiérarchique (Ward par exemple) sur les neurones afin de les rassembler en "super-classes" [LETREMY 2004]. L'avantage d'appliquer la classification de Ward sur les neurones plutôt que sur les objets directement est le nombre réduit de neurones par rapport aux objets initiaux, ce qui facilite l'interprétation de la classification. De plus, le poids final des neurones peut expliquer quelles variables sont principalement responsables du regroupement des données dans les différentes classes.

Enfin, un autre avantage d'utiliser au préalable une classification de Kohonen est qu'une fois le réseau construit, il est possible d'apporter de nouvelles données au modèle et observer à quelles classes elles appartiennent [COTTRELL 2004].

---

### 2.5.4 TEST D'INDEPENDANCE DU CHI-DEUX

Développé par Karl Pearson, le test d'indépendance du Chi-deux, Khi-deux ou  $\chi^2$  permet de vérifier si deux caractères (qualitatifs ou regroupés en classes) d'une population donnée sont indépendants ou non. Ce test consiste à évaluer si la répartition des effectifs observés est

significativement différente de celle des effectifs théoriques calculés sous l'hypothèse d'indépendance des deux variables.

Dans la pratique, on pose les hypothèses suivantes :

- $H_0$  : les deux caractères sont indépendants,
- $H_1$  : les deux caractères ne sont pas indépendants.

Soit deux variables A et B représentées sous forme d'un tableau de contingence :

		Caractère A					Total
		Modalité 1	...	Modalité i	...	Modalité p	
Caractère B	Modalité 1	$n_{11}$	...	$n_{i1}$	...	$n_{p1}$	$N_{B1}$
	...		...		...		
	Modalité j	$n_{1j}$	...	$n_{ij}$	...	$n_{pj}$	$N_{Bj}$
	...		...		...		
	Modalité q	$n_{1q}$	...	$n_{iq}$	...	$n_{pq}$	$N_{Bq}$
Total		$N_{A1}$	...	$N_{Ai}$	...	$N_{Ap}$	$N$

avec:

- $n_{ij}$  le nombre d'individus ayant la modalité i du caractère A et la modalité j du caractère B,
- $N_{Ai} = \sum_{j=1}^q n_{ij}$  (somme des effectifs de la colonne i),
- $N_{Bj} = \sum_{i=1}^p n_{ij}$  (somme des effectifs de la colonne j),
- $N$  l'effectif total de la table de contingence.

Le tableau des effectifs théoriques attendus sous l'hypothèse  $H_0$  est alors le suivant :

		Caractère A					Total
		Modalité 1	...	Modalité i	...	Modalité p	
Caractère B	Modalité 1	$e_{11}$	...	$e_{i1}$	...	$e_{p1}$	$E_{B1}$
	...		...		...		
	Modalité j	$e_{1j}$	...	$e_{ij}$	...	$e_{pj}$	$E_{Bj}$
	...		...		...		
	Modalité q	$e_{1q}$	...	$e_{iq}$	...	$e_{pq}$	$E_{Bq}$
Total		$E_{A1}$	...	$E_{Ai}$	...	$E_{Ap}$	$E$

avec:

$$e_{ij} = \frac{N_{Bj} \times N_{Ai}}{N}$$

Le calcul du  $\chi_{0bs}^2$  s'obtient par la formule suivante :

$$\chi_{0bs}^2 = \sum_{i=1}^p \sum_{j=1}^q \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

La valeur de  $\chi^2$  est comparée avec la valeur seuil,  $\chi_{Seuil}^2$  lue sur une table du  $\chi^2$  (Figure 3) pour  $(p-1) \times (q-1)$  degrés de liberté et pour un risque d'erreur  $\alpha$  fixé.

Si  $\chi_{0bs}^2 > \chi_{Seuil}^2$  l'hypothèse  $H_0$  est rejetée au risque d'erreur  $\alpha$  : il n'y a pas indépendance statistique entre les deux caractères étudiés dans la population.

Si  $\chi_{0bs}^2 \leq \chi_{Seuil}^2$  l'hypothèse  $H_0$  est acceptée : les deux caractères étudiés dans la population sont statistiquement indépendants.

Remarque : La statistique du  $\chi^2$  ne peut être calculée que si tous les effectifs théoriques  $e_{ij}$  sont supérieurs ou égaux à 5.

seuil ddl	0,200	0,100	0,050	0,020	0,010	0,005	0,001
1	1,642	2,706	3,842	5,412	6,635	7,879	10,828
2	3,219	4,605	5,992	7,824	9,210	10,597	13,816
3	4,642	6,251	7,815	9,837	11,345	12,838	16,266
4	5,989	7,779	9,488	11,668	13,277	14,860	18,467
5	7,289	9,236	11,071	13,388	15,086	16,750	20,515
6	8,558	10,645	12,592	15,033	16,812	18,548	22,458
7	9,803	12,017	14,067	16,622	18,475	20,278	24,322
8	11,030	13,362	15,507	18,168	20,090	21,955	26,125
9	12,242	14,684	16,919	19,679	21,666	23,589	27,877
10	13,442	15,987	18,307	21,161	23,209	25,188	29,588
11	14,631	17,275	19,675	22,618	24,725	26,757	31,264
12	15,812	18,549	21,026	24,054	26,217	28,300	32,910
13	16,985	19,812	22,362	25,472	27,688	29,820	34,528
14	18,151	21,064	23,685	26,873	29,141	31,319	36,123
15	19,311	22,307	24,996	28,260	30,578	32,801	37,697
16	20,465	23,542	26,296	29,633	32,000	34,267	39,252
17	21,615	24,769	27,587	30,995	33,409	35,719	40,790
18	22,760	25,989	28,869	32,346	34,805	37,157	42,312
19	23,900	27,204	30,144	33,687	36,191	38,582	43,820
20	25,038	28,412	31,410	35,020	37,566	39,997	45,315
21	26,171	29,615	32,671	36,343	38,932	41,401	46,797
22	27,302	30,813	33,924	37,660	40,289	42,796	48,268
23	28,429	32,007	35,173	38,968	41,638	44,181	49,728
24	29,553	33,196	36,415	40,270	42,980	45,559	51,179
25	30,675	34,382	37,653	41,566	44,314	46,928	52,620
26	31,795	35,563	38,885	42,856	45,642	48,290	54,052
27	32,912	36,741	40,113	44,140	46,963	49,645	55,476
28	34,027	37,916	41,337	45,419	48,278	50,993	56,892
29	35,139	39,088	42,557	46,693	49,588	52,336	58,301
30	36,250	40,256	43,773	47,962	50,892	53,672	59,703

Figure 3 : Table du  $\chi^2$ .

### 2.5.5 REGRESSIONS LINEAIRES MULTIPLES

Méthode d'exploration des données couramment utilisée, la régression linéaire multiple a pour but d'expliquer la relation et la dépendance entre une variable Y dite dépendante (dans notre cas il s'agira de l'agrément sonore) et une ou plusieurs variables  $X_i$  appelées indépendantes ou explicatives [SAPORTA 1990] [HOWELL 2009] (ici il s'agira donc par exemple de l'agrément visuel, du silence, de la présence des véhicules légers ou de tout autre variable que nous aurons sélectionnée pour établir le modèle).

La régression linéaire multiple se traduit par une équation du modèle de régression multiple du type :

$$Y_i = \beta_0 + \sum_{j=1}^{p-1} (X_{ij}\beta_j) + \epsilon_i \quad i = 1, \dots, n$$

où :

- $Y_i$  est la variable dépendante,
- $X_{ij}$   $j = 1, \dots, p - 1$ , sont les variables indépendantes,
- $\beta_j$   $j = 0, \dots, p - 1$ , sont les paramètres à estimer,
- $\epsilon_i$   $i = 1, \dots, n$ , est la valeur des résidus. Ce terme d'erreur traduit l'écart entre les valeurs observées de la variable dépendante et les valeurs prédites par le modèle.

Dans le cas de la méthode des moindres carrés il s'agira donc de trouver les paramètres  $\beta_i$  qui minimisent ce terme d'erreur ou plus précisément la somme des carrés des résidus, c'est à dire,  $\min (\sum_{i=1}^n \epsilon_i^2)$ .

Sous forme matricielle, le modèle de régression s'écrit  $Y = X\beta + \epsilon$

où :

- $Y$  est le vecteur ( $n \times 1$ ) des observations relatives à la variable dépendante ( $n$  observations),
- $\beta$  est le vecteur ( $p \times 1$ ) des paramètres à estimer,
- $\epsilon$  est le vecteur ( $n \times 1$ ) des erreurs,

- $X = \begin{bmatrix} 1 & X_{11} & \dots & X_{1(p-1)} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \dots & X_{n(p-1)} \end{bmatrix}$  est la matrice ( $n \times p$ ) ayant trait aux variables indépendantes.

A partir de la méthode des moindres carrés, l'estimation  $\hat{\beta}$  du vecteur  $\beta$  s'obtient par le calcul matriciel suivant :  $\hat{\beta} = (X'X)^{-1}X'Y$ . Ainsi, la valeur estimée  $\hat{Y}$  de  $Y$  est :  $\hat{Y} = X\hat{\beta}$ , et le vecteur  $e$  des résidus est :  $e = Y - \hat{Y}$ .

Une fois le modèle de régression réalisé, il convient d'estimer la fiabilité de ce modèle qui peut être estimée par le coefficient de détermination du modèle noté  $R^2$ . Cette grandeur mesure la qualité de l'ajustement des estimations de l'équation de régression et s'appuie sur l'équation d'analyse de variance,  $SCT = SCE + SCR$  avec  $SCT$  la somme des carrés totale,  $SCE$  la somme des carrés expliquée par le modèle et  $SCR$  la somme des carrés résiduelle. Le

coefficient de détermination  $R^2$  correspond au pourcentage de la variance expliquée par le modèle par rapport à la variance totale et se calcule de la façon suivante :

$$R^2 = \frac{SCE}{SCT} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

avec  $\bar{y}$  la moyenne des valeurs de  $y_i$ .

Le coefficient de détermination  $R^2$  ou coefficient de corrélation multiple au carré est toujours compris entre 0 et 1. Par conséquent un bon modèle de régression doit avoir un coefficient  $R^2$  élevé. Cependant, le coefficient  $R^2$  présente l'inconvénient d'augmenter mécaniquement avec l'ajout de nouvelles variables dans le modèle. Par conséquent, un autre critère pertinent est le coefficient  $R_{aj}^2$  ajusté, définit comme ceci :

$$R_{aj}^2 = \frac{(n - 1) \times R^2 - (p - 1)}{n - p}$$

avec :

- n le nombre d'observations,
- p le nombre de variables explicatives,
- $R^2$  le coefficient de détermination.

Le coefficient  $R_{aj}^2$  est inférieur au coefficient  $R^2$  mais doit être le plus proche possible de celui-ci pour assurer une bonne validité du modèle prédictif proposé. Le  $R_{aj}^2$  permet ainsi de comparer différents modèles de régression ne comportant pas le même nombre de variables, ni le même nombre d'observations.

Dans le cadre de ce travail de recherche, la comparaison a été menée entre les régressions linéaires multiples et les réseaux de neurones artificiels qui chaque fois comportaient le même nombre de variables. Aussi, le  $R_{aj}^2$  n'a pas été utilisé et la comparaison s'est faite sur la base du  $R^2$ .

## 2.5.6 MODELES DE RESEAUX DE NEURONES

Contrairement à la méthode présentée au paragraphe 2.5.2 où les réseaux de neurones utilisés étaient de type non supervisés avec les réseaux de Kohonen, il s'agit cette fois de faire appel à un perceptron multicouche avec rétro-propagation des erreurs. Ce type de réseau s'avère fort utile et efficace pour traiter notamment des problèmes d'approximation de fonction. Bien que complexe compte tenu du nombre de paramètres utilisés dans l'établissement d'un modèle de perceptron multicouche, le principe de fonctionnement peut être décrit de façon relativement simple. Il s'agit de présenter une base de données composée d'entrées (les variables explicatives) et de cibles connues (dans notre cas l'agrément sonore). Le réseau de neurones approxime par une fonction non linéaire le lien qui existe entre les entrées et les cibles et peut ainsi par ajustement améliorer cette approximation de façon à minimiser l'erreur, c'est-à-dire minimiser la différence entre les sorties calculées et les cibles.

S'il n'est pas aisé de résumer en quelques lignes tout le formalisme mathématique inhérent aux réseaux de neurones artificiels, de nombreux cours et explications peuvent nous éclairer sur le sujet [BLAYO 1996] [GURNEY 1997] [PARIZEAU 2006]. Il existe différentes architectures (perceptrons, réseaux à fonction radiale, cartes auto-organisatrices, ...) faisant elles-mêmes appel à diverses méthodes d'apprentissage associées à différents critères d'arrêt.

Mathématiquement, un neurone artificiel peut être représenté sur la Figure 4. Globalement, un neurone se comporte comme une somme pondérée  $n$  des entrées à laquelle est appliquée une fonction de transfert  $f$  qui produit la sortie  $a$  du neurone.

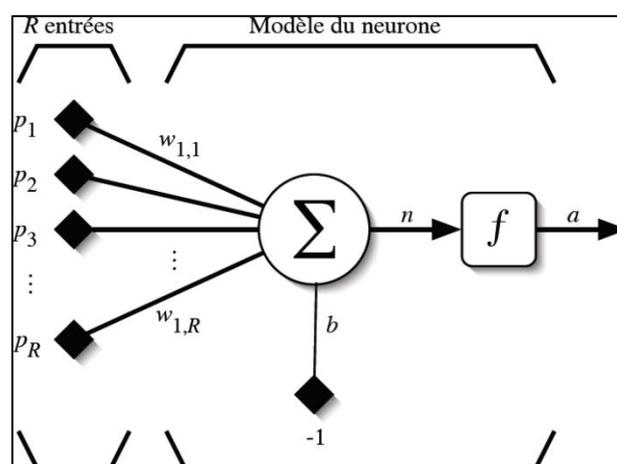


Figure 4 : Modèle d'un neurone artificiel.

Autrement dit,  $a = f(n)$ ,  $n = \sum_{j=1}^R w_{1,j}p_j - b$ , où

- $p_1, \dots, p_R$  sont les entrées,
- $w_{1,1}, \dots, w_{1,R}$  sont les poids sur les entrées,
- $b$  est le biais,
- $n$  est le niveau d'activation,
- $f$  est la fonction de transfert,
- $a$  est la sortie du neurone.

Sous forme matricielle, un modèle de neurone artificiel peut être représenté sur la Figure 5.

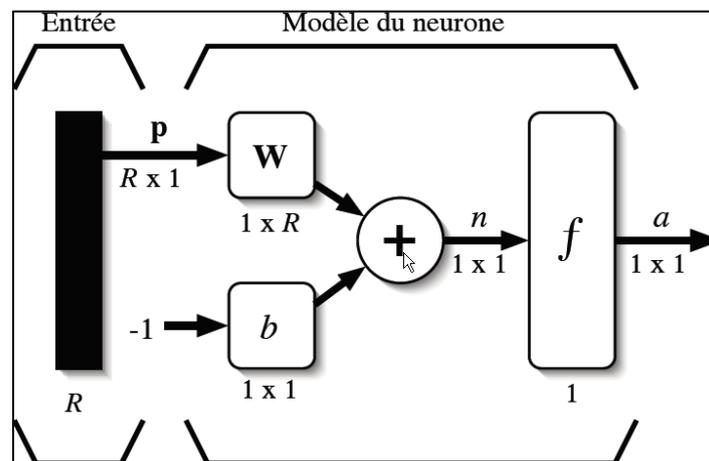


Figure 5 : Représentation matricielle d'un neurone artificiel.

La sortie  $a$  du neurone s'écrit alors  $a = f(\mathbf{W}\mathbf{p} - b)$  avec :

- $\mathbf{W} = [w_{1,1} \quad w_{1,2} \quad \dots \quad w_{1,R}]$ ,

- $\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_R \end{bmatrix}$ .

Concernant la fonction de transfert, il en existe de nombreuses. A titre d'exemple, citons trois d'entre elles parmi les plus utilisées : la fonction seuil, la fonction linéaire et la fonction sigmoïde. Les représentations graphiques de ces trois fonctions de transfert sont données sur la Figure 6 et la relation entrées / sorties de chacune est présentée dans le Tableau I.

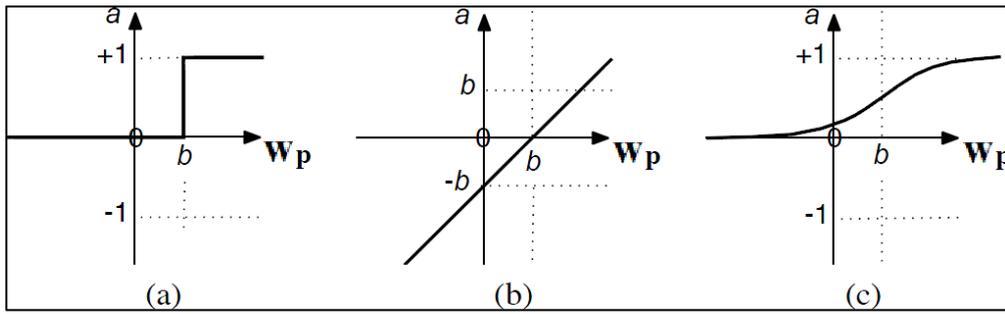


Figure 6 : Fonctions de transfert : (a) du neurone "seuil" ; (b) du neurone "linéaire", et (c) du neurone "sigmoïde".

Tableau I : Fonctions de transfert  $a=f(n)$ .

Nom de la fonction de transfert	Relation entrée / sortie
seuil (harlim)	$a = 0$ si $n < 0$ $a = 1$ si $n > 0$
linéaire (purelin)	$a = n$
sigmoïde (logsig)	$a = \frac{1}{1 + e^{-n}}$

A partir de la définition d'un neurone artificiel, un réseau de neurones consiste donc en l'association de plusieurs neurones, généralement organisés en couches. Une couche de neurones peut donc se représenter comme sur la Figure. Dans une même couche, chaque neurone est connecté à chaque entrée et un poids  $w_{i,j}$  est associé à chaque connexion. Sur la Figure 8 est représentée cette même couche de  $S$  neurones, cette fois-ci sous forme matricielle. Ainsi, la sortie  $a$  du réseau est toujours donnée par la relation  $a = f(\mathbf{Wp} - \mathbf{b})$  avec :

$$- \mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix},$$

$$- \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_R \end{bmatrix},$$

$$- \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{bmatrix}$$

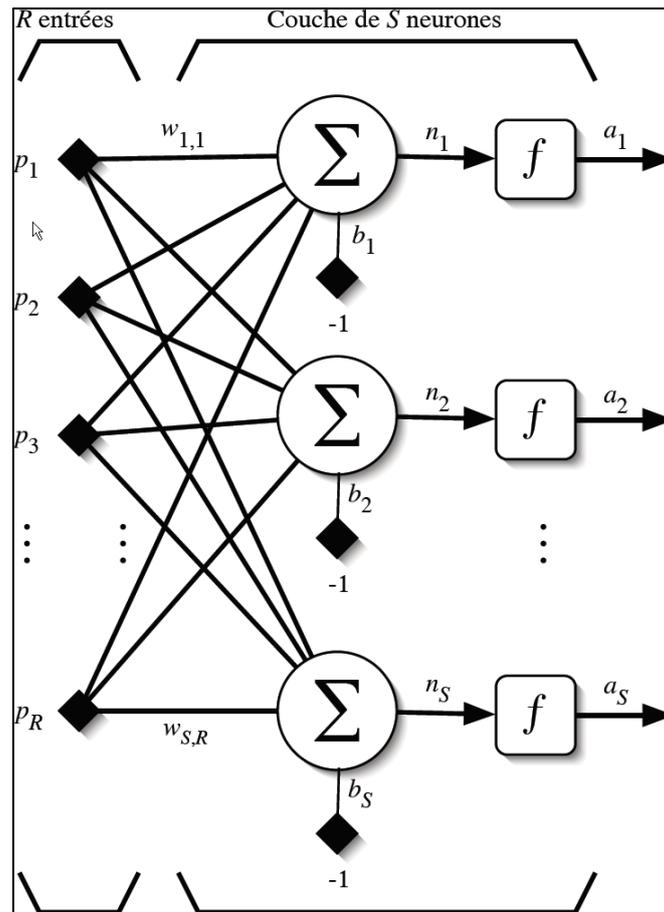


Figure 7 : Couche de S neurones.

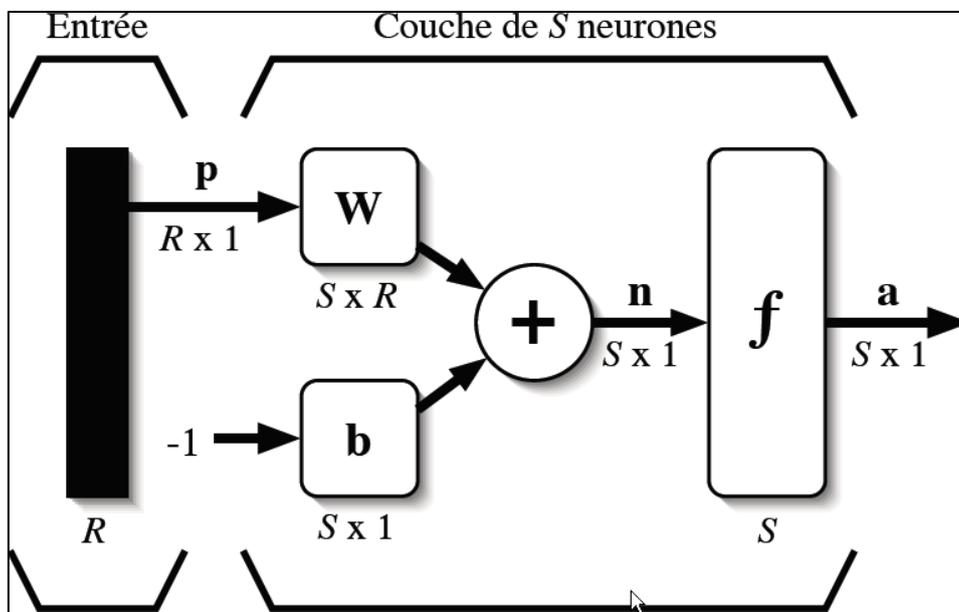


Figure 8 : Représentation matricielle d'une couche de S neurones.

Dans le cas de notre étude nous avons donc utilisé un perceptron multicouche qui peut se décrire comme l'association d'une couche cachée de S neurones et une couche de sortie de k neurones. La taille de la couche de sortie est définie par le nombre de sorties

que l'on veut calculer, dans notre cas  $k = 1$  puisque l'on cherchera à calculer seulement l'agrément sonore. Parmi tous les paramètres du modèle, il est donc possible d'agir notamment sur la taille de la couche cachée ainsi que sur les fonctions de transfert de chacune des couches.

Comme c'est souvent le cas dans les problèmes d'approximation de fonction, nous avons utilisé un réseau de neurones avec une fonction de transfert sigmoïde ("logsig" dans le logiciel Matlab) pour la couche cachée et une fonction de transfert linéaire ("linear" dans Matlab) pour la couche de sortie. La représentation de ce perceptron multicouche est donnée sur la Figure 9.

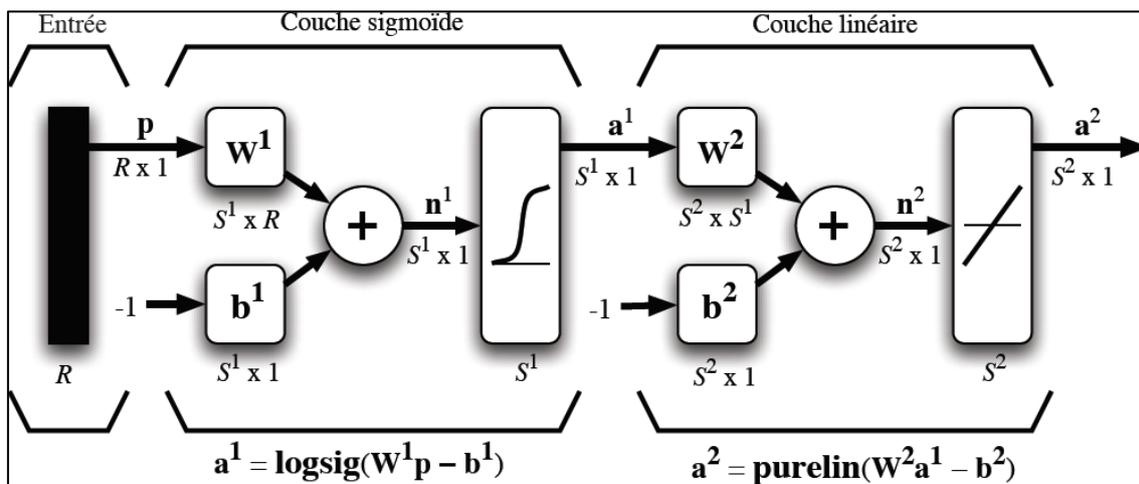


Figure 9 : Perceptron multicouche utilisé pour faire de l'approximation de fonction.

Concernant la taille de la couche cachée, tout comme pour les cartes de Kohonen, il n'existe pas de règle pour déterminer le nombre de neurones à utiliser. A partir de plusieurs essais réalisés en amont et afin de garder une certaine logique tout au long des analyses, nous avons décidé de choisir un nombre de neurones égal au nombre de variables présentes dans le modèle. Après les différents essais effectués, il est apparu que les différences obtenues avec chaque modèle étaient dues à l'aspect "aléatoire" de l'initialisation des réseaux de neurones plus qu'à une éventuelle mauvaise optimisation de la taille de la couche cachée.

En effet, le principe de fonctionnement du perceptron multicouche nécessite une initialisation des poids et des biais des neurones. Or cette initialisation de façon aléatoire peut entraîner une variabilité des résultats d'une exécution à l'autre du modèle, pour une même topologie de réseau et les mêmes entrées.

Ainsi, tous les réseaux de neurones utilisés dans l'élaboration de modèles de prédiction (§5) peuvent être décrits schématiquement par la Figure 10 où  $x$  est le nombre de variables utilisées dans chaque modèle.

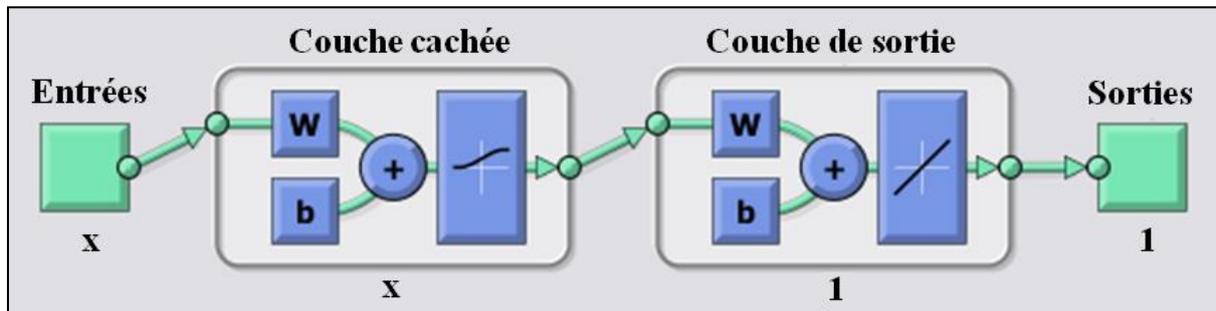


Figure 10 : Architecture des réseaux de neurones utilisés dans l'étude des modèles de prédiction.

### 2.5.7 VALIDATION CROISEE

Afin d'établir et tester les modèles de prédiction, une méthode de validation croisée a été utilisée. Cette technique consiste à diviser la base de données initiale en deux bases de données distinctes : une pour construire le modèle (base de construction) et une pour le tester (base de test). Dans le cas particulier des réseaux de neurones, la base de construction doit être elle-même divisée en deux, une base d'apprentissage et une base de validation. La proportion (de l'effectif total) de ces trois bases de données (apprentissage, validation et test) peut varier selon l'utilisation qui en est faite. Dans notre cas, nous avons choisi respectivement 50-20-30 % [PARIZEAU 2006].

Initialement utilisée pour parer à des problèmes propres aux réseaux de neurones (sur-apprentissage), cette technique de validation croisée s'est révélée en réalité fort utile pour comparer les différents modèles établis (régression linéaire multiple et réseau de neurones artificiels). En effet, comme mentionné au paragraphe 2.5.5, un modèle de régression linéaire multiple peut être caractérisé par son coefficient de détermination  $R^2$ . Or, une autre façon de caractériser la qualité du modèle est de calculer le coefficient  $r$  élevé au carré de la corrélation entre les sorties établies par le modèle et les valeurs mesurées. Si  $(a_1, \dots, a_i, \dots, a_n)$  sont les sorties du modèles et  $(d_1, \dots, d_i, \dots, d_n)$  les valeurs mesurées, le coefficient de corrélation  $r$  est alors données par la formule suivante :

$$r = \frac{\sum_{i=1}^n d_i a_i}{\sqrt{(\sum_{i=1}^n d_i^2)(\sum_{i=1}^n a_i^2)}}$$

Dans toute la suite du manuscrit, les coefficients de corrélation élevés au carré entre les sorties calculées par les différents modèles et les valeurs mesurées seront notés  $R^2$ .

La méthode de validation croisée permet donc d'avoir à disposition deux bases de données, une pour construire le modèle et une pour le tester. A partir de ces deux bases de données, il est donc possible de calculer pour chaque modèle (Mod) établi, un coefficient de corrélation au carré  $R^2$  pour la base de construction  $R_c^2(\text{Mod})$  et un coefficient de corrélation au carré  $R^2$  pour la base de test  $R_t^2(\text{Mod})$ . Un troisième coefficient de corrélation au carré peut même être calculé pour l'ensemble des données (construction + test) en réintroduisant l'intégralité de la base de données dans chaque modèle établi.

Dans le chapitre 5, les différents coefficients de corrélation au carré,  $R_c^2$ ,  $R_t^2$  et  $R_{\text{all}}^2$  (pour l'ensemble de la base de données) serviront de base à la comparaison entre les régressions linéaires multiples et les réseaux de neurones artificiels.

---

#### 2.5.8 ANALYSE DE VARIANCE

L'analyse de la variance ou ANOVA à un facteur est une méthode exploratoire qui permet de tester l'influence d'une variable indépendante discrète que l'on appelle facteur sur une variable dépendante. Dans notre cas nous avons testé au paragraphe 5.2.11.4 le facteur "lieu" qui comporte 8 niveaux sur la variable dépendante "pente". Concrètement, ceci revient à comparer les moyennes de plusieurs populations normales et de même variance à partir d'échantillons aléatoires et indépendants les uns des autres. Le nom d'analyse de variance provient de la façon de procéder : la variance totale de l'échantillon est décomposée en deux variances partielles à comparer, la variance intra-classes et la variance inter-classes.

Dans la pratique supposons que l'on ait un facteur A qui présente p modalités ( $1 \leq i \leq p$ ). Le nombre de répétitions j pour une modalité i est noté  $n_i$ . La valeur prise par la variable aléatoire Y pour la modalité i du facteur et la répétition j est notée  $y_{ij}$  et les valeurs moyennes pour chaque modalité notée  $\bar{y}_i$ , soit, sous forme de tableau :

Facteur A				
Modalité 1		Modalité i		Modalité p
$y_{11}$	...	$y_{i1}$	...	$y_{p1}$
	...		...	
$y_{1j}$	...	$y_{ij}$	...	$y_{pj}$
	...		...	
$y_{1n1}$	...	$y_{ini}$	...	$y_{pnp}$
$\bar{y}_1$	...	$\bar{y}_i$	...	$\bar{y}_p$

Le principe de l'analyse de variance est de comparer la variabilité à l'intérieur de chaque groupe avec la variabilité entre les groupes [DODGE 2007]. L'équation de décomposition de la variance totale  $SC_{Tot} = SC_{Inter} + SC_{Intra}$  s'écrit alors :

$$\sum_{i=1}^p \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^p n_i (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^p \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

A partir du calcul de ces sommes des carrés, la variance totale vaut alors  $\frac{SC_{Tot}}{N-1}$ , la variance inter-classes est égale à  $CM_{Inter} = \frac{SC_{Inter}}{p-1}$  et la variance intra-classes est égale à  $CM_{Intra} = \frac{SC_{Intra}}{N-p}$ .

Dès lors, il est possible de comparer le ratio  $F_{Obs} = \frac{CM_{Inter}}{CM_{Intra}}$  avec la valeur critique  $F_{seuil}$  de la loi de Fisher-Snedecor de degrés de liberté  $(p-1, N-p)$  pour le risque  $\alpha$  voulu (généralement  $\alpha = 0.05$ ). Si la variable de test est supérieure à la valeur critique, alors on rejette l'hypothèse que les distributions sont identiques, donc elles sont significativement différentes.

Les résultats issus de l'analyse de variance ont été complétés par un test post-hoc de type HSD de Tukey qui permet de voir quels sont les groupes pour lesquels les différences sont significatives.

## 2.6 CONCLUSION ET PRESENTATION DU PLAN DE TRAVAIL

L'ensemble des chapitres 1 et 2 a permis de voir que la notion de "soundscape" établie dans les années 70 est plus que jamais au cœur des discussions qui animent la communauté scientifique, qui cherche aujourd'hui à s'accorder sur la définition même de "soundscape" et donc sur la mesure et la caractérisation de celui-ci. Un "soundscape" se définit comme vu précédemment comme un environnement acoustique tel qu'il est perçu, vécu et compris par les personnes, dans un contexte.

Dans le but de développer un indicateur de qualité sonore en milieu urbain, nous nous sommes avant tout posé la question, dans le cadre de la caractérisation des "soundscapes", où doit-on se placer dans le temps et dans l'espace ? Pour répondre à cette question, nous avons dans un premier temps tenté dans le chapitre "3 - Maillage temporel" de savoir quelle était la durée de mesure optimale et le moment de la journée permettant de caractériser différents environnements sonores, le tout à partir d'indicateurs à la fois énergétiques mais également événementiels.

De la même façon, nous avons cherché dans le chapitre "4 - Maillage spatial" à savoir où se placer pour déterminer un indicateur de qualité sonore, non plus dans le temps; mais cette fois-ci dans l'espace. Toujours à travers des indicateurs énergétiques et événementiels, cette fois-ci liés aux sources sonores, nous avons cherché à savoir si un parcours continu pouvait être discriminé en plusieurs environnements, d'un point de vue acoustique.

Enfin, pour terminer, à partir de la durée de mesure optimale et de la discrimination spatiale, nous avons voulu établir dans le chapitre "5 - prédiction" un indicateur de qualité sonore à partir d'enquêtes de terrain (dont la durée a été déterminée avec le maillage temporel) que nous avons soumis à des passants (en des lieux déterminés à partir du maillage spatial). Ces enquêtes portaient donc sur l'évaluation de l'environnement en un lieu et à un moment précis en se basant sur la perception de critères tout d'abord globaux, liés au contexte ou non, comme par exemple, l'agrément visuel de l'environnement, le sentiment d'enveloppement ressenti face à l'environnement sonore ou encore l'impression d'animation, de vie, fournies par ce dernier. Puis le questionnaire soumis portait ensuite sur des paramètres plus spécifiques de l'environnement sonore, à savoir les sources qui le composent. Chacune des sources évaluées l'a été sous deux formes : le temps de présence et la force sonore de celle-ci.

En s'appuyant sur des méthodes de calculs basées à la fois sur des régressions linéaires multiples et sur des réseaux de neurones artificiels, nous avons établi des modèles de prédiction de la qualité sonore urbaine pour chacun des lieux étudiés, mais aussi un modèle global correspondant à l'ensemble des lieux.