

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR UNIVERSITY -ANNABA-  
UNIVERSITE BADJI MOKHTAR -ANNABA-



جامعة باجي مختار  
- عنابة -

Faculté: Sciences de l'Ingénieur -Année 2017-

Département: Informatique

## THESE

Présentée en vue de l'obtention du diplôme de **Doctorat en Sciences**

# Approche Multi-agents pour la simulation de la pollution de l'air

**Option : Intelligence  
Artificielle**

**Par**

**Sabri GHAZI**

<b>Directeur de thèse</b>	Khadir Mohamed Tarek	Pr.	Université Badji Mokhtar, Annaba
<b>Co-directrice de these</b>	Julie Dugdale	MCA	Université de Grenoble Alpes

### **Devant le jury**

<b>Président</b>	Farah Nadir	Pr.	Université Badji Mokhtar Annaba
<b>Examineur</b>	Moussaoui Abdelouahab	Pr	Université Ferhat Abbas Sétif
<b>Examineur</b>	Bouchrika Imed	MCA	Université Mohamed-Chérif Messaadia Souk Ahras

## Remerciements :

---

Je tiens à remercier en premier lieu Mr. Khadir Mohamed Tarek, Professeur à l'Université Badji Mokhtar-Annaba, grace à lui ce travail a pu voir le jour. Je le remercie pour ses qualités de chercheur scientifique qui ne cesse pas à améliorer et à mettre en valeur notre travail de recherche.

Mille mercies à Julie Dugdale, Maitre de Conférence à l'Université de Grenoble-Alpes, qui a codirigé ce travail et m'a permis de faire part de son équipe MAGMA au laboratoire d'informatique de Grenoble. Mes remerciements vont aussi aux membres de l'équipe MAGMA.

Je désire également remercier Monsieur Farah Nadir, Professeur à l'Université Badji Mokhtar-Annaba, pour sa courtoisie et l'honneur qu'il m'a accordé en acceptant de présider le jury de ma thèse.

Je remercie Mr. Imed Boucherika, Maître de Conférences à l'Université Mouhamed Cherif Messadia-Souk Ahras ainsi que Mr. Moussaoui Abdelouahab, Professeur à l'Université Ferhat Abbas Sétif, pour m'avoir fait l'honneur d'examiner mon travail.

Dédicace :

---

*Je dedie ce modeste travail en premier  
lieu à mes Parents,*

*à Ma femme, à Ma Sœur,*

*à Mes Frères.*

*à Mes Amis.*

# Table des matières:

---

Remerciements :	2
Dédicace :	3
Table des matières:	4
الملخص	8
Résumé :	9
Abstract:	10
Index des images:	11
Index des tableaux:	13
Chapitre 1 : Introduction générale	15
1 Contexte de la thèse :	15
2 Problématique	17
3 Objectif de la thèse.....	18
4 L'Approche proposée	18
5 Organisation de la thèse :	20
Chapitre 2 Modélisation des Systèmes Socio-environnementaux.....	22
1 Introduction.....	22
2 Les caractéristiques des problèmes socio-environnementaux :	24
3 Les approches de modélisation des systèmes socio-environnementaux	27
3.1 Approche analytique et dynamique.....	27
3.1.1 Les limitations des approches analytiques.....	28
3.2 Approche de simulation dynamique.....	28
3.3 Approche de simulation interactive.....	29
4 Conclusion :	30
Chapitre 3 : Modélisation de la pollution de l'air.....	31
1 Introduction :	31
2 La pollution de l'air.....	31
2.1 Composition chimique de l'atmosphère.....	32
2.2 La structure verticale de l'atmosphère.....	32
2.3 Pollution atmosphérique.....	33
2.5 Différents types de polluant.....	34

2.5.1	Les polluants primaires.....	35
2.5.2	Les polluants secondaires.....	35
2.6	Les effets de la pollution de l'air et leurs principales sources:.....	36
2.7	Les sources anthropiques de polluant:.....	36
3	Les modèles de prédiction de pollution de l'air (état de l'art).....	37
3.1	Les caractéristiques d'un modèle de prédiction .....	37
4	La modélisation de la dispersion du polluant atmosphérique .....	54
4.1	Le modèle en boîte (Box models).....	55
4.2	Le modèle Gaussien de dispersion .....	56
4.3	Le modèle Lagrangien.....	56
5	Conclusion : .....	58
Chapitre 4 Les Systèmes Multi-agents .....		59
1	Introduction :.....	59
2	L'approche basée agent .....	60
2.1	Définition du concept Agent .....	61
2.3	Les systèmes multi-agents.....	63
2.3.1	Définition:.....	63
2.3.2	L'environnement d'un SMA.....	64
2.4	Les SMA et les phénomènes à comportements émergents.....	65
2.5	La coopération et le mécanisme de prise de décision. ....	66
2.5.1	Interaction et organisation .....	66
2.5.2	La prise de la décision.....	67
2.6	Les approches similaires aux SMA.....	67
2.6.1	Les automates cellulaires .....	68
2.6.2	La micro-simulation:.....	68
2.7	Les Avantages des SMA .....	69
2.8	Les inconvénients et les limitations des SMA.....	70
2.9	Validation des modèles SMA.....	71
Conclusion .....		73
Chapitre 5 : Utilisation des systèmes multi-agents pour la modélisation des problèmes socio-environnementaux .....		74
1	Introduction.....	74
2	Les critères de comparaison des études .....	76
3	La modélisation des problèmes socio-environnementaux basée multi-agent: état de l'art ....	81

4	Conclusion .....	90
Chapitre 6 : Un Framework pour le développement d'un simulateur basé multi-agents de la pollution de l'air.....		
1	Introduction.....	92
2	L'Approche et l'architecture du modèle: .....	93
2.1	L'échelle spatio-temporelle du modèle de simulation.....	96
2.2	Le modèle de prédiction de la qualité de l'air :.....	98
2.2.1	Apprentissage des réseaux RBF :.....	99
2.2.2	La combinaison des modèles RBF en utilisant un réseau MLP.....	101
2.3	Le modèle de dispersion: .....	103
2.3.1	Les polluants secondaires.....	104
2.4	La simulation basée agent.....	104
2.4.1	La stratégie de coopération centralisée (CS).....	104
2.4.2	La stratégie de coopération basée jeux évolutifs.....	105
2.4.2.1	Dilemme de prisonniers à N personnes: .....	106
2.4.2.2	La pénalité liée à la pollution.....	111
3	Conclusion .....	112
Chapitre 7 Expérimentation .....		
1	Introduction.....	113
2	Région ciblée par les scénarios de simulation.....	114
2.1	Les données utilisées dans l'étude.....	116
2.1.1	Les polluants mesurés par Sama Safia .....	117
2.1.2	Paramètres climatiques.....	117
2.2.3	Propriétés statistiques des données .....	118
3	Les scenarios de simulation.....	126
3.1	Les paramètres de l'émission: .....	126
3.2	Les paramètres qui concernent l'objectif de la simulation :.....	127
3.3	Les paramètres liés à la coopération des agents :.....	127
3.4	Les paramètres liés à l'environnement : .....	128
3.5	Les paramètres liés au temps :.....	128
3.6	Le scénario de simulation numéro 1 et 2 :.....	129
3.7	Le scénario de simulation numéro 3 : .....	130
Chapitre 8 : Résultats et discussions .....		
1	Introduction :.....	131

Table des matières:

---

2	Discussion des résultats .....	133
2.1	Les sorties de la simulation .....	133
2.2	Le premier scenario.....	136
2.3	Scénario de simulation numéro 2 .....	144
2.3.1	Résultat de la simulation .....	144
2.4	Le troisième scenario de simulation.....	145
2.4.1	Résultat de la simulation .....	146
	Chapitre 9 : Conclusions et Perspectives.....	149
	Annexe A : Publications et Communiations : .....	154
	Publication dans des revues interationales :.....	154
	Chapitre de livre : .....	154
	Communications internationales : .....	154
	Annexe B : Implémentation d'un simulateur en utilisant l'approche SMA.....	155
	Introduction : .....	155
	Session de travail sous le système de simulation MASAP.....	163
	Bibliographie.....	166

## المخلص

تعتبر النشاطات و تصرفات الانسان من أهم أسباب التي تأثر في المحيط الطبيعي و تسبب التلوث. فتلوث الهواء مثلا ناتج في كثير من الاحيان عن النشاطات الصناعية و حركة النقل و المرور. لذلك انه من الاهمية بما كان أن تؤخذ القرارات و النشاطات التي يقوم بها الانسان ضمن نماذج المحاكات التي تهدف الي محاكات التلوث وفهم تطوره. هذه المذكرة تتعلق بتطوير و دراسة امكانية استعمال النماذج المستعملة حاليا مثل نماذج الانتشار و نماذج التنبؤ واقتراح صيغة جديدة تستعمل الانظمة متعددة الخوادم . الهدف من تركيب كل هته الطرق هو بناء نموذج يحاكي تلوث الهواء و يسمح بدراسة تأثير تصرفات و قرارات الانسان في تركيز الملوثات في الهواء. الشبكات العصبية الاصطناعية و نماذج الانتشار استعملت في هذا العمل و تمكنا من دمجها في نظام متعدد الخادم. لإظهار مدى كفاءة هته المنهجية.

النموذج المطور تمت تجربته باستعمال معطيات و بيانات تخص تلوث الهواء في مدينة عنابة التي تقع شمال شرق الجزائر. التجارب بينت مدى تأثير قواعد التعاون بين المتحكمين في نقاط التلوث في تحسين او تدني نوعية الهواء.

### الكلمات المفتاحية

نوعية الهواء, تلوث الهواء, الشبكات العصبية الاصطناعية, انظمة متعددة الخادم, المحاكات الاجتماعية.

## Résumé :

---

Ce mémoire présente une approche multi-agents pour la simulation de la pollution de l'air. L'approche décrite, vise à simuler le phénomène de pollution de l'air comme étant un système socio-environnemental. Elle inclut l'aspect physique de la dispersion et la concentration des polluants ainsi que le facteur de décision humaine concernant les activités polluantes. Deux modèles sont intégrés ensemble pour prédire la concentration des polluants, le premier est un modèle de dispersion Gaussien, il permet de calculer la dispersion par rapport aux sources de pollution, le deuxième modèle est un réseau de neurones artificiels de type RBF qui utilise les valeurs de dispersion et les paramètres climatiques pour prédire la concentration de la pollution. Ces deux modèles sont utilisés au sein d'un système multi-agents. Ce dernier, modélise sous forme d'agent, l'ensemble des acteurs impliqués dans la gestion des sources de pollution et la qualité de l'air. L'interaction et la coopération entre les agents permettent de simuler les activités anthropiques qui sont les principales causes de la dégradation de la qualité de l'air. Un système de simulation est conçu afin de montrer la faisabilité de notre approche, il vise à prédire la concentration du  $\text{SO}_x$ ,  $\text{PM}_{10}$ ,  $\text{CO}_x$ ,  $\text{NO}_x$  et l' $\text{O}_3$ , ainsi que la qualité de l'air. Le système a été testé avec trois scénarios en utilisant les données concernant la ville d'Annaba (nord-est de l'Algérie). L'approche a permis d'étudier les différentes stratégies qui assurent la coopération, et permettent d'aider dans la prise de décision concernant la gestion de la qualité de l'air.

### Mot clés :

Simulation Multi-Agents, Système Multi-Agent, Système Socio-environnemental, pollution, Modèle de dispersion, modèle de prédiction, réseaux de neurones artificiels, pollution de l'air,  $\text{PM}_{10}$ , Ozone,  $\text{SO}_x$ ,  $\text{NO}_x$ .

## Abstract:

---

This thesis presents a Multi-Agent System (MAS) approach for designing an air pollution simulator. The aim is to simulate the concentration of air pollutants emitted from sources (e.g. factories) and to investigate the emergence of cooperation between the emission sources managers and its impact on air quality. The emission sources are controlled by agents. The agents try to achieve their goals (i.e. increase production which has the side effect of raising air pollution) and also cooperate with others agents by altering their emission rate according to the air quality. The agents play an adapted version of the evolutionary N-Person Prisoners' Dilemma game in a non deterministic environment; they have two decisions: decrease, or increase the emission. The rewards/penalties are influenced by the pollutant concentration which is, in turn, determined using climatic parameters. In order to give predictions about the concentration of pollutants: Particulates Matter (PM10), Sulphur Oxide and Dioxide ( $\text{SO}_x$ ), Nitrogen oxides ( $\text{NO}_x$ ) and Ozone: ( $\text{O}_3$ ), a two stage prediction method is used, a GPD (Gaussian Plume Dispersion) model and an ANN (Artificial Neural Network) prediction model. The prediction is calculated using the dispersion information and real data about climatic parameters (wind speed, humidity, temperature and rainfall). Every agent cooperates with its neighbours that emit the same pollutant, and it learns how to adapt its strategy to gain more reward. When the pollution level exceeds the maximum allowed level, agents are penalised according to their participation. The system has been tested using real data from the region of Annaba (North-East Algeria). It helped to investigate how the regulations enhance the cooperation and may help controlling the air quality. The designed system helps the environmental agencies to assess their air pollution controlling policies.

### Keywords:

Agent based simulation, air pollution, air quality, pollutant Dispersion model, PM10, Ozone, artificial neural networks, RBD, MLP.

# Index des images:

---

Figure 2.1 Les composants d'un système socio-environnemental (Source : Bossel 2003).....	25
Fig. 2.2 Diagramme de composant, représentant l'interaction entre les composants d'un système socio-environnemental (écologique, social et économique).....	26
Figure 2.3, Exemple d'un modèle Système Dynamique (Source : De Vries 2009).....	29
Figure 3.1 : Profile vertical de température dans l'atmosphère.....	33
Figure 3.2 Echelle spatio-temporelle de chaque polluant, source (Boubel et al).....	35
Figure -3.3- : Modèle en box de pollution de l'air.....	53
Figure 4.1 : Représentation d'un SMA, (source Jennings 2000).....	64
Figure 5.1 : Les modèles classés selon les trois axes : écologie, social et économique.....	81
Figure 6.1 architecture du moteur de simulation proposé.....	93
Figure 6.2: Représentation graphique des valeurs possibles pour la variable "Qualité de L'air".....	95
Figure 6.3: Schémas du processus de simulation.....	96
Figure 6.4: La représentation de l'environnement.....	97
Figure 6.5: Schéma générique d'un modèle de prédiction de concentration du polluant.....	98
Figure 6.6: Schema d'un réseau de neurones artificiels de type RBF.....	100
Figure 6.7: Un system Multi-Classifieur utilisant les RBF et le MLP pour la prédiction de la qualité de l'air.....	102
Figure 6.8 : Les deux courbes de récompense, une pour les coopérateurs et l'autre pour les non-coopérateurs.....	107
Figure 7.1 : Carte géographique de Annaba (source :Google )......	115
Figure 7.2 : Topographie de la région d'Annaba ( source :Google Earth ).....	116
Figure 7.3 : Concentration PM10 en 2003.....	119
Figure 7.4 : Concentration P10 en 2004.....	119
Figure 7.5 : Concentration O3 en 2004.....	120
Figure 7.6 : Concentration NO <sub>x</sub> en 2003.....	121
Figure 7.7 : Concentration NO <sub>x</sub> en 2004.....	121
Figure 7.8 : Concentration du NO en 2003.....	122
Figure 7.9 : Concentration du NO en 2004.....	122
Figure 7.10 : Concentration du CO <sub>2</sub> en 2003.....	123
Figure 7.12 : Concentration du NO <sub>2</sub> en 2003.....	123
Figure 7.13 : Humidité en 2003.....	124

Figure 7.14 : Humidité en 2004.....	124
Figure 7.15 : Température en 2003.....	125
Figure 7.16 : Température en 2004.....	125
Figure 7.17 : Vitesse du vent en 2003. ....	126
Figure 7.18 : Vitesse du vent en 2004. ....	126
Figure 8.1: L'indexe de la qualité de l'air pour toutes les stratégies testées (EG-CP, EG-NCP, EG-NP, CS and NC), en utilisant 4900 pas de simulation.....	139
Figure 8.2: évolution de l'indexe de qualité de l'air pendant les 4900 pas de simulation.....	139
Figure 8.3: Evolution de la concentration des PM10 pendant le temps de la simulation .....	140
Figure 8.4 : L'évolution de la concentration du SO <sub>x</sub> pendant l'exécution du premier scénario. ....	140
Figure 8.5: L'évolution de la concentration du NO <sub>x</sub> pendant l'exécution du premier scénario. ....	140
Figure 8.6: l'évolution de la concentration du O <sub>3</sub> pendant l'exécution du premier scénario.....	141
Figure 8.7: Taux de coopération par rapport à toute la population d'agents pour les stratégies EG-CP, EG-NCP and EG-NP. ....	142
Figure 8.8: Le taux de coopération par type de polluants en utilisant la stratégie EG-CP.....	142
Figure 8.9, Le taux de coopération par type d'agents en utilisant la stratégie EG-NCP. ....	142
Figure 8.10: Le taux de coopération par type de polluants dans le cas d'une stratégie EG-NC. ....	143
Fig. 8.11. Les pics de PM10 pour 24 heures, selon les quartes stratégies testées.....	144
Figure 8.12, L'évolution de la concentration du PM10 pour le scenario avec et sans des sources non-contrôlées.....	146
Figure 8.13, SO <sub>x</sub> concentration for the scenario with and without leak sources.....	147
Figure 8.14, NO <sub>x</sub> concentration for the scenario with and without leak sources. ....	147
Figure B.1, les fonctionnalités de simulateur sont décrites en utilisant le diagramme des cas d'utilisation UML. ....	156
Figure B.2 : Affichage de la pile des points de concentration par rapport à la source d'émission.....	157
Figure B.3 Diagramme UML des classes du système MASAP.....	159
Figure B.4 Diagramme de communication, fonctionnalité de lancement d'apprentissage du réseau de neurones artificiels.....	161
Figure B.5 Diagramme de communication fonctionnalité de lancement du modèle de dispersion. .	161
Figure B.6 : Ecran principal de MASAP .....	163
Figure B.7 : fonctionnalité d'apprentissage des modèles de prédictions RNA. ....	163
Figure B.8 : Selection d'une stratégie de coopération. ....	164
Figure B.9 : Configuration d'une simulation.....	164
Figure B.10 : Exécution d'un scénario de simulation. ....	165

# Index des tableaux:

---

Tableau 3.1 Composition chimique de l'air.....	32
Tableau 3.2 : Les différents types de polluant et leurs effets.....	36
Tableaux 3.3 : Travaux sur la prédiction de la concentration de pollution dans l'air .....	39
Tableau -3.4- Paramètres utilisés.....	48
Tableau -3.5- : Un extrait des règles construites.....	52
Tableau 3.5 : Liste de quelques systèmes de modélisation de dispersion et leurs modèles.....	57
Tableau 4.1 : résumé des propriétés utilisées pour définir un agent. ....	63
Tableau 5.1 : les travaux utilisant les SMA pour la modélisation des problèmes socio- environnementaux. ....	80
Table 6.2. L'indice de la qualité de l'air.....	101
Tableau 6.3, La fonction de récompense de jeux.....	106
Tableau 7.1 : Les polluants mesurés par Sama Safia.....	117
Tableau 7.2 : Les paramètres climatiques mesurés par Sama Safia. ....	118
Tableau 7.3 Paramètres météorologiques moyens dans la région d'Annaba. ....	118
Tableau 7.4 : Propriétés statistiques de mesure de concentration de PM10.....	120
Tableau 7.5 : Propriétés statistiques de mesure de concentration de l'Ozone. ....	120
Tableau 7.6 : Propriétés statistiques de mesure de concentration de NO <sub>x</sub> .....	121
Tableau 7.10 Propriétés statistiques de mesure de concentration de NO.....	122
Tableau 7.11 : Propriétés statistiques de mesure de concentration de CO <sub>2</sub> . ....	123
Tableau 7.12 : Propriétés statistiques de mesure de concentration de NO <sub>2</sub> .....	123
Tableau 7.13 : Propriétés statistiques de mesure de l'humidité. ....	124
Tableau 7.11 : Propriétés statistiques de mesure de température. ....	125
Tableau 7.12 : Propriétés statistiques de mesure de la vitesse du vent.....	126
Tableau 7.13 : les paramètres de simulation du scénario.....	129
Tableau 7.14 les paramètres de simulation du deuxième scénario.....	130
Tableau 8.1. Topologie des réseaux RBF conçus.....	132
Table 8.2: Erreur de validation des réseaux RBF.....	132
Tableau 8.3 : Les sorties de la simulation.....	135
Tableau : 8.4: Les valeurs minimales des 60 sorties de la simulation. ....	136
Tableau 8.5 Les valeurs maximales des 60 sorties de la simulation. ....	136

Index des tableaux:

---

Tableau 8.6 : Les valeurs moyennes des 60 sorties des simulations.....	137
Tableau B.1 : Les types d'agents et leurs comportements. ....	160
Tableau B.2 : La matrice d'interaction inter-types-agents, ici 1 veut dire qu'il existe une interaction 0 représente aucune interaction.....	162

# Chapitre 1 : Introduction générale

---

## 1 Contexte de la thèse :

La pollution sous ses diverses formes, pèse beaucoup sur la qualité de vie dans de nombreuses villes dans le monde. Elle a un coût économique et social très considérable. La pollution de l'air, qui fait l'objet d'étude de cette thèse, cause selon l'OMS (OMS, 2014) la mort chaque année à 7 millions de personnes. Selon la même source, elle est responsable de 2% des maladies cardiaques et pulmonaires dans le monde.

La pollution atmosphérique est une préoccupation majeure dans de nombreuses villes dans le monde, en particulier, dans les pays en développement. La densité de la population a eu des effets considérables sur la qualité de vie dans les grandes villes (Molina & Molina 2004): Le phénomène de pollution de l'air a une influence directe sur notre santé et notre qualité de vie (OMS 2005). Plusieurs villes sont touchées chaque année par des pics de pollution qui causent des centaines de victimes, à titre d'exemple : Téhéran le 01 janvier 2015, (RFI ,2015), Pékin (Chan et Yao 2008). Les effets de pollution de l'air sont constatés au niveau local et regional ainsi qu'au niveau mondial : l'effet de serre, les trous d'Ozone, la sécheresse, la désertification, les pluies acides, etc.

Pouvoir assurer un développement économique tout en minimisant les perturbations que subit l'environnement est un véritable défi universel, qui fait l'objet de débat dans des plusieurs conférences mondiales (Demaze 2009). L'Algérie, comme tous les pays en voie de développement n'est pas épargnée face à ce problème. Après son indépendance (1962), le pays s'est concentré sur la reconstruction de son économie via l'adoption d'une industrie lourde et l'exportation des hydrocarbures. Ces activités sont classées comme étant très polluantes. Avec un taux de natalité de 10,13% durant les décennies 60, 70 et 80, la démographie du pays a passé de 6 millions en 1962 à 39.5 millions selon les dernières statistiques de l'ONS (Office National des Statistiques, 2015). Ceci a causé une expansion urbaine mal-contrôlée, et a conduit à des agressions sur l'environnement (les côtes, les rivières, les lacs, les forêts et l'air). Les agglomérations dans le nord du pays produisent beaucoup de déchets et la politique de recyclage n'est pas assez développée (Bendjoudi et al. 2009). Les méthodes classiques d'incinération et d'enfouissement des déchets sont toujours largement utilisées. Dans le sud algérien où le problème de démographie n'est pas posé, les activités d'exploitation et de forage causent des

problèmes environnementaux très sérieux. Avec les ambitions que présente l'Algérie pour aller vers l'exploitation des hydrocarbures non-conventionnels, le problème de pollution a connu une grande médiatisation<sup>1</sup>.

Ces dernières années, l'idée du développement durable est devenue de plus en plus évoquée, que ce soit au niveau académique ou politique. Il est devenu très urgent d'opter pour un nouveau modèle de développement qui permettra de préserver l'environnement tout en assurant durablement les besoins de la population.

La question sur la façon dont les humains devraient modérer leur exploitation des ressources environnementales a occupée les chercheurs depuis des décennies (Folke 2006). Pouvoir promouvoir la croissance sociale et économique sans affecter l'équilibre de l'environnement est un élément important pour maintenir un développement durable. Cette question fait l'objet de tout un domaine de recherche, qui vise à étudier les technologies qui causent moins de pollution ou de perturbation à l'environnement. Ce domaine est pluridisciplinaire et nécessite l'implication de plusieurs experts dans des divers domaines: la sociologie, l'architecture, la gestion des villes, la politique, la chimie, les mathématiques, l'informatique et l'écologie. L'objectif étant d'étudier le système que forme l'être humain avec son environnement ainsi que leurs interactions. Ce genre de système est nommé système socio-environnemental ou aussi socio-naturel, ou même socio-écologique. On peut définir ce genre de système selon deux points de vue : Un système naturel subissant des perturbations anthropiques. Ou bien, un ensemble de population (système social) sujet à des limitations de ressources environnementales. Ces systèmes (Glaser 2006) impliquent plusieurs acteurs et présentent des comportements dynamiques et non-linéaires. Ils sont classés comme des systèmes complexes et adaptatifs (Holling 2001). L'exploitation d'une ressource écologique partagée peut être formulée comme une tragédie du commun (Hardin 1968). Chaque acteur tend à maximiser ses profits en exploitant une ressource écologique partagée. Ainsi, une tragédie du commun est alors posée.

Il est presque impossible de pouvoir isoler un système socio-environnemental dans un laboratoire pour faire des expérimentations et prévoir son évolution. Les chercheurs font alors recours à la simulation. Cette dernière est considérée comme étant une méthode scientifique permettant l'exploitation des modèles et observer leurs comportements sous des conditions bien précises. La modélisation permet de créer une représentation abstraite sous forme de modèle répliquant, selon un certain niveau de réalité, les comportements du système réel. La simulation

---

<sup>1</sup> <http://www.courrierinternational.com/article/2015/01/17/dans-le-sud-gronde-la-protestation-contre-le-gaz-de-schiste>

est vue comme étant un outil efficace pour étudier et prédire les interactions et l'évolution dans un phénomène socio-écologique. C'est grâce à ce genre de modèles que le monde a pu prédire les effets de l'échauffement climatique. Ces outils permettent de sensibiliser les politiciens et les décideurs concernant ces problèmes de pollution.

Dans cette thèse nous nous intéressons à la modélisation de la relation entre les comportements, les décisions humaines et leurs impacts sur la qualité de l'air dans un système socio-environnemental. La simulation de la pollution de l'air est un outil d'aide à la décision qui permet d'aider les décideurs à mettre en place des politiques efficaces de gestion de l'environnement et de prédire l'impact de leurs décisions sur l'environnement et l'écosystème.

## 2 Problématique

De nombreuses approches de modélisation ont été proposées pour étudier la pollution de l'air. La plupart d'entre elles, (Kryza & al. 2014, Holmes NS & Morawska 2006, Enkeleida & Stockie 2010) sont principalement axées sur les aspects physiques et chimiques de la pollution atmosphérique : la concentration et la dispersion des polluants dans l'air. Ces modèles ne prennent pas en compte les facteurs de la décision humaine. Ils cherchent toujours à prédire l'évolution de la pollution en termes de dispersion et concentration dans l'air. (Daly & Zannetti 2007) présente une revue des approches utilisées pour modéliser la pollution de l'air. Ceci via des modèles mathématiques (régression linéaires, modèles de dispersion, statistique).

Or, la pollution atmosphérique est, par nature, distribuée et comprend l'interaction de plusieurs acteurs impliqués dans l'exploitation de l'air, qu'on peut considérer comme étant une ressource écologique dynamique. Les activités anthropiques (trafic routier et transport, activités industrielles et agricoles) sont parmi les principales sources de polluants. Toutes ces activités sont contrôlées par l'homme, par conséquent, inclure les facteurs de la décision humaine dans la modélisation de la pollution de l'air semble être essentiel. Notre problématique consiste à proposer une approche de simulation qui combine les systèmes multi-agents, les modèles de dispersion du polluant, les modèles de prédiction des concentrations des polluants, dans le but d'examiner les questions suivantes :

1. Combiner entre un modèle de dispersion et un modèle de prédiction pour donner des prédictions qui incluent l'émission (les activités humaines) et les paramètres climatiques (l'incertainnité).
2. Modéliser et introduire la décision humaine concernant les sources d'émission de polluant dans le processus de simulation. Ceci se fait en modélisant les

comportements des gestionnaires des sources d'émission à l'aide d'un système multi-agents.

3. Etudier les stratégies de coopération entre les différents acteurs (modélisés sous forme d'agents), dans le but de préserver la qualité de l'air. Ce qui permettra de voir comment les contrôleurs des sources d'émissions coopèrent, et sont-ils en mesure d'atteindre leurs objectifs tout en conservant une qualité de l'air raisonnable. Quelles sont les règles qui devraient être adoptées pour renforcer la coopération, et par conséquent, maintenir la qualité de l'air?
4. Montrer comment les gestionnaires des sources d'émission doivent ajuster leurs émissions selon les paramètres climatiques et les prédictions dans le but de gérer les périodes de pics.
5. Permettre une simulation qui offre la possibilité de tester et évaluer les différentes politiques de gestion et de contrôle de la qualité de l'air.
6. Permettre de simuler et prévoir les effets sur la qualité de l'air des décisions d'étendre ou d'instaurer des nouvelles sources d'émission de polluant de l'air.
7. Permettre de simuler la situation d'une crise de pollution de l'air, qui peut inclure des sources de pollution non contrôlables, telles que les fuites industrielles et les incidents naturels.

### **3 Objectif de la thèse**

L'objet du travail que nous avons mené le long de cette thèse est de proposer une nouvelle approche de simulation de la pollution de l'air dont la principale nouveauté est l'inclusion du facteur des décisions humaines. Notre objectif est la modélisation du problème de pollution de l'air comme étant un système socio-environnemental.

Nous modélisons les gestionnaires des sources de polluants comme étant des agents autonomes visant à maximiser leur profit en émettant des polluants et nous étudions l'effet sur la qualité de l'air. L'approche permet de construire des systèmes qui aident à enquêter sur l'efficacité des lois de régulations utilisées par les organismes pour maintenir la qualité de l'air. Ceci est très important car ce genre de systèmes aide les organismes environnementaux dans leurs évaluations de leurs politiques de contrôle de pollution atmosphérique.

### **4 L'Approche proposée**

L'approche basée SMA (Système Multi-Agents) est une méthode appropriée pour la modélisation des problèmes socio-environnementaux (Aulinas & al, 2009). Elle permet de

modéliser les comportements des acteurs humains partageant l'exploitation des ressources environnementales. (Hülsmann & al 2011), (Kickhöfer & al 2012) ont utilisé une approche SMA pour étudier l'émission de pollution de l'air résultant des activités routières; ils ont utilisé une simulation du trafic de la circulation et ils ont lié les patterns de trafic avec un modèle de calcul des émissions. (Kickhöfer & Nagel 2013) a utilisé la même approche pour étudier l'effet de la réglementation adoptée pour la gestion des transports sur la pollution atmosphérique et l'émission. Les gestionnaires des sources d'émissions partagent l'exploitation de l'air par l'émission des polluants. Nous visons à simuler leurs différentes personnalités et étudier la relation entre l'émergence de la coopération et son impact sur la qualité de l'air.

L'approche de simulation proposée, modélise la population des contrôleurs des sources d'émission comme un réseau d'agents jouant un DPNP évolutive (Dilemme des Prisonniers à N-Personnes). Ce qui nous permettra d'étudier l'émergence des comportements de coopération entre les agents dans une population d'agents égoïstes. Cela aide à inspecter comment les agents présentent des comportements altruistes et dans quelles conditions la coopération est maintenue.

Pour montrer la faisabilité de notre approche, on a envisagé à concevoir un système de simulation pour modéliser la pollution de l'air dans la région d'Annaba, dans le nord-est de l'Algérie. La ville est considérée comme un pôle économique très important. Avec une activité industrielle assez riche, allant d'industrie des engrais à la sidérurgie et un parc d'automobiles qui ne cessent d'augmenter. Le problème de pollution alors justement posé, surtout dans les régions urbaines qui sont très proches des zones d'activités industrielles. Annaba souffrent d'une dégradation brutale dans la qualité de l'air, surtout durant la période d'été (Tadjine & al. 2008).

La ville s'est dotée depuis 2003 d'un réseau de surveillance de la qualité de l'air. Un tel réseau ne permet pas de remédier au problème de pollution. Il doit être doté d'autres outils de simulation afin de pouvoir tester et prédire l'efficacité des politiques de gestion environnementale.

L'idée est d'approcher le phénomène en tant que problème socio-environnemental, dans lequel la décision humaine joue un rôle important. Etant donné que les activités polluantes sont gérées par des êtres humains, nous avons eu l'idée de concevoir un système de simulation qui prend en considération la prise de décision des gestionnaires des sources de pollutions afin de voir l'évolution de la pollution selon plusieurs scénarios de gestion de la qualité de l'air. Un tel simulateur permettra aux autorités de régulation de tester leurs politiques de gestion et de control de la pollution de l'air.

## 5 Organisation de la thèse :

La démarche qu'on a adoptée pour la réalisation de ce travail est reflétée dans l'organisation de ce document. Ce manuscrit est structuré en 9 chapitres y compris ce chapitre d'introduction.

Le deuxième chapitre est consacré à la présentation des éléments et les concepts concernant la modélisation des problèmes socio-environnementaux. La pollution de l'air est un problème socio-environnemental, pour cela, nous avons choisi de présenter dans ce chapitre, les composants ainsi que les définitions d'un problème socio-environnemental. Les différentes approches utilisées pour la modélisation socio-environnementale sont aussi présentées.

Le troisième chapitre présente les concepts liés à la pollution de l'air, à savoir, les différents types de polluants, les origines et les effets de ce phénomène. Nous présentons aussi dans ce chapitre les deux méthodes classiques de modélisation de la pollution de l'air : les modèles de prédiction de la concentration des polluants ainsi que les différents modèles de dispersion.

Le quatrième chapitre, expose les systèmes multi-agents. Les définitions ainsi que les concepts liés aux systèmes multi-agents sont présentés. Ce chapitre présente les avantages qui nous ont motivés à adopter cette approche de modélisation. Les limites ainsi que la problématique de validation des modèles SMA sont aussi discutées.

Après avoir introduit les concepts liés aux SMA, nous présentons dans le cinquième chapitre, un état de l'art concernant l'utilisation des SMA pour modéliser les problèmes socio-environnementaux. Nous commençons par présenter les critères selon lesquels nous avons comparé les études. A la fin de ce chapitre, nous présentons une conclusion situant notre approche par rapport aux travaux connexes.

Dans le chapitre 6 nous décrivons l'approche proposée pour le développement d'un simulateur basé multi-agents de la pollution de l'air. Le chapitre commence par présenter l'architecture générale du modèle proposé. S'en suivra une présentation de la combinaison entre le modèle de dispersion et le modèle de prédiction basé réseaux de neurones artificiels. Le mécanisme de prise de décision utilisé est présenté ainsi que les différentes stratégies de coopération entre les agents, la modélisation des interactions entre les agents selon un modèle issue de la théorie des jeux, sont aussi détaillés dans ce chapitre.

Les scénarios d'expérimentation sont décrits dans le chapitre 7, ils utilisent les données concernant la région d'Annaba, pour cela une brève description cette région est présentée. Ce chapitre contient aussi la description des paramètres de la simulation.

Nous analysons et discutons les résultats des expérimentations dans le chapitre 8.

Finalement, nous finissons ce manuscrit par une conclusion générale par laquelle nous résumons les contributions apportées grâce à ce travail de recherche, nous présentons aussi les futures perspectives de notre travail.

# Chapitre 2 Modélisation des Systèmes

## Socio-environnementaux.

---

### 1 Introduction

Avant de commencer la description des modèles socio-environnementaux, nous devons tout d'abord définir quelques concepts liés à l'environnement et l'écosystème. Larousse définit l'environnement comme étant : « L'ensemble des éléments (biotiques et abiotiques) qui entourent un individu ou une espèce et dont certains, contribuent directement à subvenir à ses besoins ».

Cette définition montre clairement l'importance de l'environnement pour toutes les espèces vivantes. En fait, l'environnement est constitué de deux composants : Le premier composant c'est la partie biotique, elle représente la partie vivante de l'environnement. Chaque espèce vivante possède un ensemble de relations avec le composant biotique de son environnement. Ces relations peuvent exister entre les mêmes espèces intra-spécifiques ou bien interspécifiques, c'est-à-dire, entre l'espèce et les autres espèces vivantes dans le même environnement. Le deuxième composant c'est la partie abiotique, qui représente la partie non-vivante et qui fournit les ressources physiques et chimiques nécessaires à la vie (eau, sol et air).

Le couple espèces-vivantes et environnement forment un système écologique appelé couramment un écosystème. Un écosystème existe dans un milieu (biotope) offrant des conditions homogènes et stables pour maintenir les besoins physicochimiques dans lequel survivent plusieurs espèces (biocénose) échangeant de l'énergie et la matière en tout équilibre. Un exemple d'écosystème est les zones humides composées par des lacs offrant des conditions typiques pour les oiseaux et les poissons pour survivre.

Les écosystèmes sont sujets à des perturbations qui affectent leurs stabilités. Ces dernières sont parfois d'origine naturelle qui s'intègre dans le processus de régénération naturelle. Parfois ces perturbations sont d'origine anthropique, survenue après l'exploitation humaine de l'environnement. Malgré le fait que l'être humain est un composant de l'écosystème, son activité est parfois considérée comme étant une perturbation causant des dégâts irréversibles sur l'écosystème.

Les perturbations sont classées selon leurs influences sur l'écosystème et aussi selon leurs origines (ce qui les a causées).

Nous parlons des problèmes environnementaux lorsque les perturbations non-naturelles persistent et causent une détérioration de l'équilibre de l'écosystème, parfois permanentes et irréversibles.

Une perturbation qui laisse des séquelles sur l'écosystème est considérée comme étant un problème environnemental. Les perturbations peuvent être catégorisées en : surexploitation des ressources, et d'autres liées à la pollution. Ils existent des perturbations liées à une surexploitation d'une ressource naturelle, donc l'état de cette ressource est affectée selon un taux d'exploitation, cette exploitation peut parfois arriver à causer des risques existentiels (exemple la surpêche maritime qui causent la disparition de certaines espèces).

L'environnement possède des capacités qui lui permettent d'absorber et de neutraliser (résilience) afin de regagner son équilibre. Dans certains cas cet équilibre prend beaucoup de temps (de quelques jours à plusieurs siècles). On peut distinguer entre : La pollution, impliquant des effets nocifs sur l'écosystème (parfois irrécupérables). La contamination qui consiste en une présence d'une substance dans l'environnement sans arriver à causer des dommages.

Depuis plusieurs années deux domaines s'intéressent à modéliser et à étudier les systèmes sociaux séparément des systèmes environnementaux. À savoir, la sociologie et l'écologie. Dans les dernières années et après que les problèmes liés à la pollution sont devenus alarmants. Une nouvelle piste de modélisation est proposée. Elle consiste à, au lieu de modéliser le système environnemental seul, faire impliquer l'activité humaine qui est la source principale des perturbations. Les chercheurs alors, commencent à inspecter la façon avec laquelle s'interfèrent les deux systèmes (social et environnemental). L'objectif est d'étudier le système que forme l'être humain avec son environnement ainsi que leurs interactions. Ce genre de système est nommé système socio-environnemental ou aussi socio-naturel, ou même socio-écologique. On peut définir ce genre de système selon deux points de vue : Un système naturel subissant des perturbations anthropiques. Ou bien un ensemble de population (système social) sujet à des limitations de ressources environnementales. Ces systèmes (Glaser 2006) font impliquer plusieurs acteurs et présentent des comportements dynamiques et non-linéaires. Ils sont jugés être des systèmes complexes et adaptatifs (Holling 2001). L'exploitation d'une ressource écologique partagée peut être formulée comme une tragédie du commun (Hardin 1968). Chaque acteur tend

à maximiser ses profits en exploitant une ressource écologique partagée. Ainsi, une tragédie du commun est alors posée.

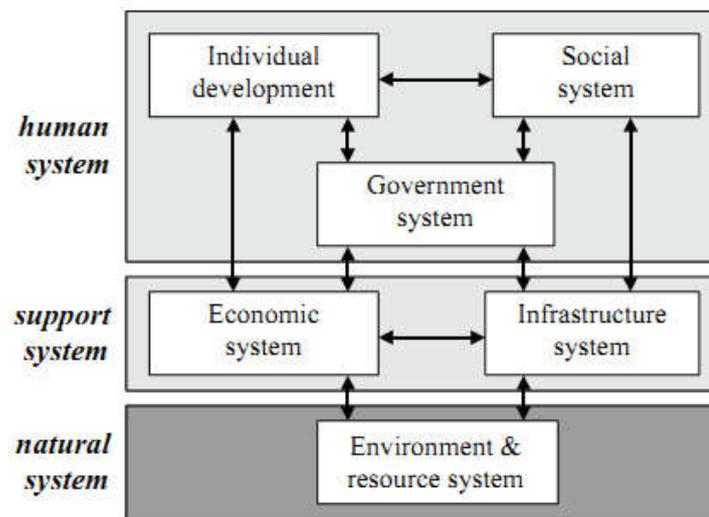
Il est presque impossible de pouvoir isoler un système socio-environnemental dans un laboratoire pour faire des expérimentations et prévoir son évolution. Les chercheurs font alors recours à la simulation. Cette dernière est considérée comme étant une méthode scientifique permettant l'exploitation des modèles et inspecter leurs comportements sous des conditions bien précises. La modélisation permet de créer une représentation abstraite sous forme de modèle répliquant selon un certain niveau de réalité, les comportements du système réel.

## **2 Les caractéristiques des problèmes socio-environnementaux :**

- Pluridisciplinaires par nature, un problème socio-environnemental fait impliquer plusieurs domaines (sociologie, écologie, géologie, économie), par conséquence, leur gestion fait appel à plusieurs experts.
- Complexes et adaptatifs (Holling 2001), le comportement du système global est loin d'être prédit en analysant les comportements de ses composants séparément.
- Caractérisés par l'incertitude, la complexité des systèmes environnementaux ne permet pas de comprendre l'impact de l'activité anthropique (de l'être humain) et les prédictions exactes sont souvent impossibles.
- Les informations collectées sont généralement imprécises et incertaines. Et impliquent une grande distribution spatiale et temporelle hétérogène (exemple, la distribution des ressources en eaux dans un pays ou une région).
- Ils sont par nature dynamiques, et donc la modélisation profonde de leurs comportements est très difficile à réaliser (Montse 2009).
- Pouvoir évaluer et définir des mesures concernant leur équilibre est très délicat (Bell & Morse 2013).

Pour modéliser les systèmes SE, ils existent plusieurs approches, catégorisées par (de Vries 2009) en deux grandes classes : des approches statiques et des approches dynamiques. Dans chaque classe les modèles peuvent être linéaires ou non-linéaires. Une autre typologie peut se faire en se basant sur les caractéristiques du modèle, à savoir déterministe ou bien il utilise des variations stochastiques.

Un modèle socio-environnemental est une représentation de l'interaction de l'activité humaine avec son environnement physique, biologique et social. La plus citée est celle de (Bossel 2003), illustré dans la figure 2.1. Cette organisation permet de distinguer entre les trois composants du système SE. Premièrement le composant social (ou humain), deuxièmement le système naturelle (ou bien l'environnement). L'interaction entre ces deux composants se fait grâce à une interface (le support d'interaction), cette interface est composée d'une partie liée à l'économie et une autre liée aux infrastructures.

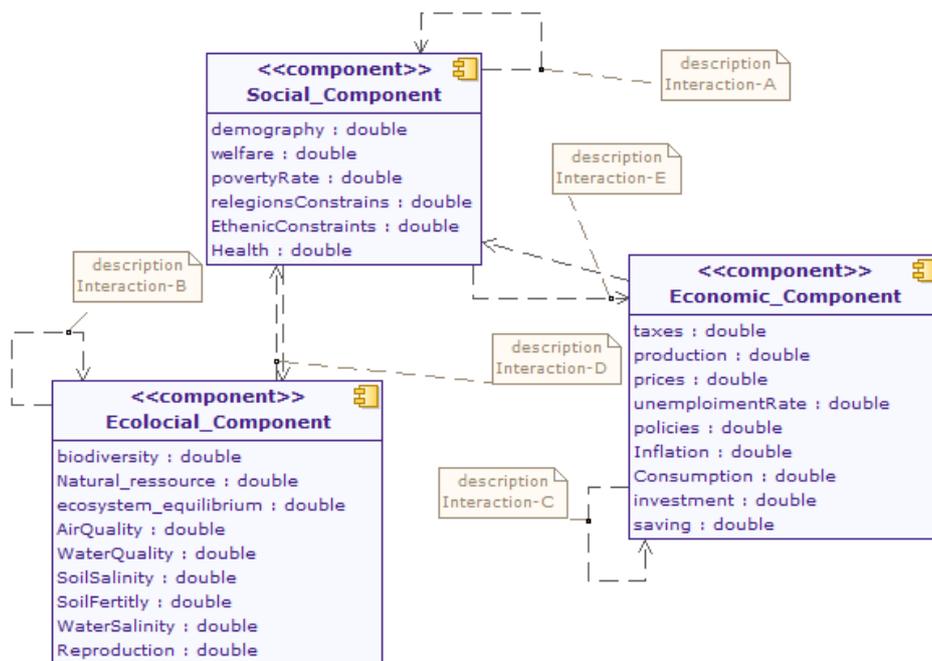


**Figure 2.1 Les composants d'un système socio-environnemental (Source : Bossel 2003)**

Beaucoup de conceptualisations ont été proposées pour représenter le système socio-naturel. La figure 2.2, montre qu'un modèle d'un système socio-environnemental peut être représenté comme une interconnexion de trois composants (ou sous-systèmes); chacun est représenté par un ensemble de variables (attributs) formant son état à l'instant  $t$ . Les composants du modèle peuvent être catégorisés en : Le composant écologique représente les parties biotiques et abiotiques. Le composant économique représente l'ensemble des entités, les règles (les lois) et les variables économiques. Le composant social représente les réseaux sociaux humains impliqués dans la simulation; parmi eux sont les décideurs, les entreprises et les organismes gouvernementaux. Le changement dans la variable d'état de chaque composant affecte les variables d'état des autres composants. Un exemple typique est l'augmentation de la demande pour certains types de poissons, ce qui motive les pêcheurs à intensifier leurs exploitations; Ce qui entraîne des changements à la biodiversité.

Six interactions peuvent exister entre les composants d'un système SE:

- A. Elle représente l'interaction du sous-système social avec lui-même, modifier certains attributs affecte d'autres attributs sociaux, par exemple: lorsque les contraintes politiques sont utilisées pour contrôler la démographie de la population;
- B. Elle se produit lorsque des événements naturels modifient l'équilibre de l'écosystème, un exemple de ce type d'interaction est les activités volcaniques qui affectent la qualité de l'air;
- C. Ce type d'interaction a lieu entre le composant économique et lui-même. Ainsi, les variables économiques sont systématiquement interconnectés, (par exemple la relation entre les prix et l'inflation);
- D. Elle se produit entre le composant social et le composant écologique. Cette interaction est considérée comme étant un élément clé de la modélisation, la compréhension de cette interaction contribue à préserver les ressources environnementales et assurer la durabilité;
- E. Elle représente l'interaction entre les composantes sociale et économique, par exemple: la relation démographie consommation.



**Fig. 2.2** Diagramme de composant, représentant l'interaction entre les composants d'un système socio-environnemental (écologique, social et économique).

### 3 Les approches de modélisation des systèmes socio-environnementaux

Dans ce qui suit, nous allons donner une brève description pour chacune des approches utilisées pour modéliser les systèmes SE.

#### 3.1 Approche analytique et dynamique

Les systèmes à équation différentiel sont très utilisés, en fait, en définissant les limites (les bornes) d'un système donné ainsi que son état initial, il peut être représenté par un ensemble de variables d'état :

$$X_1(t), X_2(t), \dots, X_n(t), \dots \text{ pour } t > t_0, \quad (2.1)$$

L'évolution de l'état du système peut après être représenté par :

$$\frac{dX(t)}{dt} = F(X(t)_1, X(t)_2, \dots, X(t)_n), G(t), t \quad (2.2)$$

Avec  $F$  la fonction qui modélise le système et  $G$  c'est la fonction qui représente les paramètres environnementaux (au sens général du terme). Les variables d'état peuvent être par exemple la concentration d'un polluant, le nombre d'espèce dans un écosystème, le niveau des eaux dans un lac. Toutes ses variables représentent le système et évoluent dans le temps. Un exemple d'un tel modèle est celui qui modélise l'interaction entre deux espèces (prédateur-proie). Un autre exemple est quand  $X1$  et  $X2$  représentent l'interaction entre un nombre de touristes et le niveau de pollution dans un lac. L'interaction entre les deux peut être formulée comme suit :

$$\frac{dx_1}{dt} = F_1[X_1, X_2] = a_{11}X_1 + a_{12}X_2 + c_1 \quad (2.3)$$

$$\frac{dx_2}{dt} = F_2[X_1, X_2] = a_{21}X_1 + a_{22}X_2 + c_2 \quad (2.4)$$

Vu la simplicité des modèles linéaires, les fonctions  $F1$  et  $F2$  (2.3 et 2.4) peuvent être non-linéaires. Ce qui permet d'avoir un modèle d'équation différentiel non-linéaire. Un exemple d'un tel système est présenté dans (de Vries 2009) :

$$\frac{d \text{PolCLake}}{dt} = \frac{\text{PolIN}}{V} \text{ PolCLak} \text{ PolCIUp} + \frac{\text{PolCLake} - \text{PolSedm}}{d} \text{ DiffC} \quad (2.5)$$

Où,  $\text{PolClak}$  c'est le taux de pollution dans un lac,  $\text{PolIn}$  c'est le taux d'intrusion du polluant à partir des sources (ex : agriculture),  $V$  c'est le volume du lac,  $\text{PolCIUp}$  c'est le facteur de nettoyage naturel qui permet au lac de s'auto-rééquilibrer, le  $\text{PolSedm}$  c'est le taux de pollution absorbé par les sédiments et  $\text{DiffC}$  c'est le facteur de diffusion.

### 3.1.1 Les limitations des approches analytiques

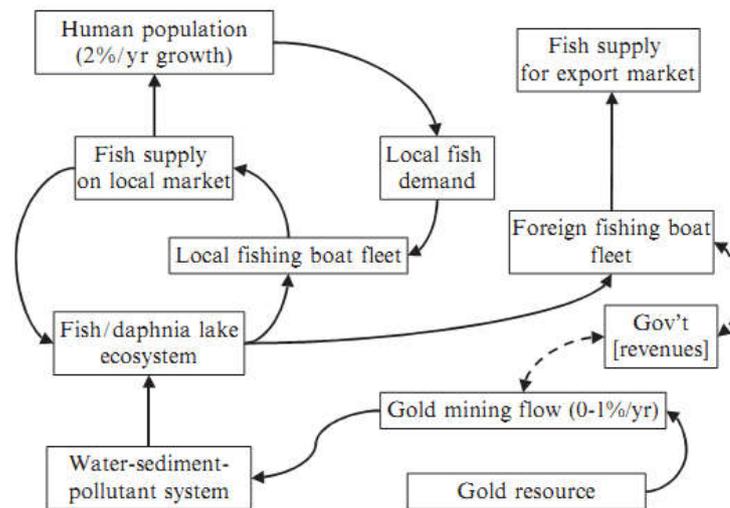
Ce genre de formulation est largement utilisé et permet d'avoir des modèles qui sont plus ou moins explicatifs, mais présentent plusieurs limitations qui résident dans :

- A. L'hypothèse que le phénomène est distribué en homogénéité (comme par exemple dire que le taux de pollution c'est le même dans toutes les parties du lac) ceci est loin d'être réel, vu que la distribution spatial est presque toujours hétérogène.
- B. En plus, on peut remarquer que ce genre de modèle n'inclut pas suffisamment l'activité de l'être humain. Les modèles sont basés sur des hypothèses que les activités sont continues et stables (comme l'exemple de lac, le taux de pollution est assumé être stable et ne subissant aucune influence, mis à part, celui de la source), alors que dans la réalité elles sont toujours distribuées d'une façon non-continues.

## 3.2 Approche de simulation dynamique

La simulation dynamique consiste à évaluer les états du système pour une série de valeurs (données) en fonction du temps. Cette approche de simulation englobe : les systèmes dynamiques, les systèmes multi-agents (voir le chapitre 4). En Système Dynamiques (Andersen & al. 1983) les processus sont représentés sous forme de stocks (argent, voiture, clients etc.) et des flux entre ces stocks. Les règles qui contrôlent ces déplacements entre les stocks sont modélisées par des équations différentiels.

L'approche SD fait l'abstraction à partir des événements et des entités et prend une vue agrégée. L'approche SD modélise les comportements du système sous forme d'un nombre de boucles en interaction constante. (de Vries 2009) présente un exemple de modélisation d'un système socio-environnemental à l'aide des systèmes dynamiques. Le modèle traite l'interaction entre un ensemble d'activités humaines (pêche et mines) et l'évolution de l'écosystème d'un lac. Le modèle inclut des paramètres sociales (démographie), économiques (exportation du poisson), comme schématisé dans la figure 2.3.



**Figure 2.3, Exemple d'un modèle Système Dynamique (Source : De Vries 2009).**

Les principales limites de cette approche résident aussi dans l'hypothèse d'homogénéité et aussi la distribution spatiale. En fait, les pêcheurs ainsi que les poissons sont représentés par deux stocks, ceci rend le modèle incapable d'expliquer la distribution spatiale de la ressource et aussi son hétérogénéité. Le modèle se base sur des agrégations, ce qui fait que les individus dans le même stocke ne sont pas distinguables et n'ont pas d'individualité.

### **3.3 Approche de simulation interactive**

C'est une approche qui se base sur l'interaction avec l'utilisateur. Ces interactions font partie de la simulation. Les décisions de l'utilisateur sont incluses parmi les entrées de la simulation. En prenant des actions l'utilisateur peut voir immédiatement les résultats de ses décisions. L'une des méthodes de cette approche sont les Jeux Sérieux. Cette dernière combine entre l'interaction de l'utilisateur avec une modélisation de son environnement, elle se base sur les scénarios et l'interaction. Généralement, le but de ces jeux est pédagogique, dans certains modèles le but est d'analyser les traces des actions des joueurs. Ils permettent d'apprendre aux utilisateurs selon des scénarios d'apprentissage bien précis (Guardiola & al. 2012). Ce genre de simulations permet de mettre les décideurs (les acteurs dans la gestion d'une ressource écologique) devant des situations de gestion très proches de la réalité, et ces expériences, ils apprennent à bien prendre des décisions adéquates tout en ayant des prédictions concernant les résultats de ces décisions. Dans (García-Barrios & al. 2008), un exemple de simulateur interactif est présenté dans le but d'apprendre aux dirigeants la gestion en commun des ressources en eaux d'irrigation. La récupération et l'analyse des traces des utilisateurs dans ce genre de jeux, permet de mieux comprendre et analyser leurs comportements.

## 4 Conclusion :

Dans ce chapitre, nous avons présenté la définition et les caractéristiques des systèmes socio-environnementaux. Nous avons aussi discuté brièvement les différentes approches utilisées pour modéliser ce genre de système. Modéliser un problème socio-environnemental est une tâche très délicate, pour cela des hypothèses sont parfois nécessaires à prendre, concernant quelques aspects du phénomène modélisé. A part celles dites simulation interactives, les approches actuelles n'incluent pas suffisamment l'aspect de décision humaine, et aussi ils se basent sur des hypothèses concernant l'homogénéité de la dispersion spatiale des phénomènes modélisés. Les systèmes socio-environnementaux, sont caractérisés par une distribution spatiale très hétérogène.

Dans notre approche, nous souhaitons modéliser la pollution de l'air comme étant un problème socio-environnemental, dans le chapitre qui suit on présente les concepts et les aspects liés à ce problème ainsi que les méthodes utilisées actuellement pour le modéliser. Ceci à fin de situer notre proposition par rapport aux autres approches.

# Chapitre 3 : Modélisation de la pollution de l'air

---

## 1 Introduction :

Ce chapitre décrit les concepts liés au phénomène de la pollution de l'air. Nous commençons par donner quelques définitions et présenter les différents types de polluants et leurs sources. Dans la deuxième section, nous présentons une brève description des effets de la pollution. Après nous présentons les différentes méthodes utilisées pour modéliser la pollution de l'air. Ces méthodes sont catégorisées en deux grandes familles:

- La modélisation de dispersion de polluant, qui permet de simuler la dispersion des polluants selon leurs différentes sources d'émissions. Les modèles de dispersion les plus utilisés sont présentés brièvement.
- La deuxième méthode consiste à prédire la concentration des polluants dans l'air. Un état de l'art, concernant cette approche est présenté.

Ce chapitre a comme objectif d'exposer les éléments qui concernent la pollution de l'air, et aussi de situer notre approche de modélisation par rapport aux différentes approches existantes.

## 2 La pollution de l'air

L'air est défini comme étant l'enveloppe gazeuse de la terre. L'atmosphère est l'environnement dans lequel la vie subsiste, beaucoup de transformation chimique, en particulier d'origine photochimique surviennent à ce niveau.

L'air qui est le fluide gazeux qui constitue l'atmosphère, est indispensable à la vie car il participe au processus de la respiration et à la photosynthèse des végétaux, cet environnement très sensible subit le plus d'influence des activités anthropiques.

## 2.1 Composition chimique de l'atmosphère

L'air qu'on respire est composée principalement de l'Azote moléculaire ( $N_2$ ), suivie de l'oxygène moléculaire ( $O_2$ ), ainsi que quelques gazes moins abondants comme Ne, He, Kr, et Xe, qui apparaissent aussi en quantité bien déterminée. Chaque espèce possède un temps de résidence allant de quelques minutes jusqu'à plusieurs années, ce qui prouve le dynamisme de la structure chimique de l'air. Le tableau ci-dessous illustre les différentes concentrations des constituants de l'atmosphère, ainsi que le temps de résidence de chaque espèce :

Gaz	Concentration	Temps de résidence
Azote ( $N_2$ )	78.084 %	-
Oxygène ( $O_2$ )	20.946 %	-
Argon (Ar)	0.934 %	-
Eau ( $H_2O$ )	[0.4..400] x 102 ppm	10 jours
Dioxyde Carbone $CO_2$	370 ppm	4 ans
Néon (Ne)	18.18 ppm	-
Hélium (He)	5.12 ppm	2 * 10 <sup>6</sup> ans
Méthane ( $CH_4$ )	1.75 ppm	10 ans
Krypton (Kr)	1.14 ppm	-
Hydrogène ( $H_2$ )	0.4 ppm	-
Xénon (Xe)	0.87 ppm	-

**Tableau 3.1 Composition chimique de l'air.**

## 2.2 La structure verticale de l'atmosphère

Schématiquement l'atmosphère peut être découpée en quatre zones selon l'altitude :

- Troposphère : de 0 jusqu'à 12 kilomètres.
- Stratosphère : jusqu'à 50 kilomètres.
- Mésosphère : jusqu'à 80 kilomètres.
- Enfin la thermosphère : jusqu'à 150 kilomètres.

Chaque couche se distingue par la fonction température/altitude, et c'est en se basant sur la température que le découpage a été défini, la figure suivante illustre la courbe de la température selon l'altitude :

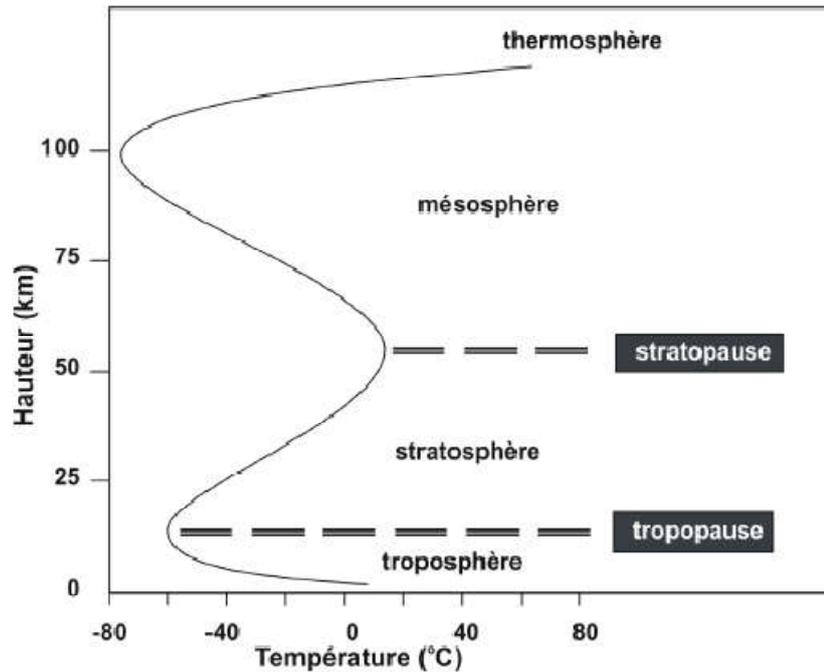


Figure 3.1 : Profil vertical de température dans l'atmosphère.

### 2.3 Pollution atmosphérique

La pollution atmosphérique (ou de l'air) est l'altération de la composition chimique de l'air, par l'augmentation de la concentration de certaines espèces (chimique ou solide), cet altération durera une certaine période de temps pour créer des effets toxiques, et qui aura aussi des effets nocif sur l'écosystème et la santé (Barker et Al. 1963). La composition chimique de l'air est en constante évolution depuis des milliers d'années, ces changements sont dus aux perturbations d'origine naturelle : Géologique tel que les éruptions volcaniques, Astronomiques comme les collisions avec des météorites, ou même biologiques (pollen, décomposition naturelle). Durant les dernières 300 années, les activités humaines causent de plus en plus de perturbation pendant une période de temps assez courte.

La pollution est dite locale (à l'échelle régional tel que : les pics d'ozone, le smog), globale à l'échelle planétaire (l'effet de serre, destruction de la couche d'ozone).

La pollution qui touche la couche troposphérique constitue le type le plus important, du fait que c'est dans cette couche que l'humain ainsi que beaucoup d'espèces vivent. Les autres

types de pollution sont moins importants, sauf ceux qui influencent le climat, tels que les gaz à effet de serre, ces derniers affectent la couche thermosphère.

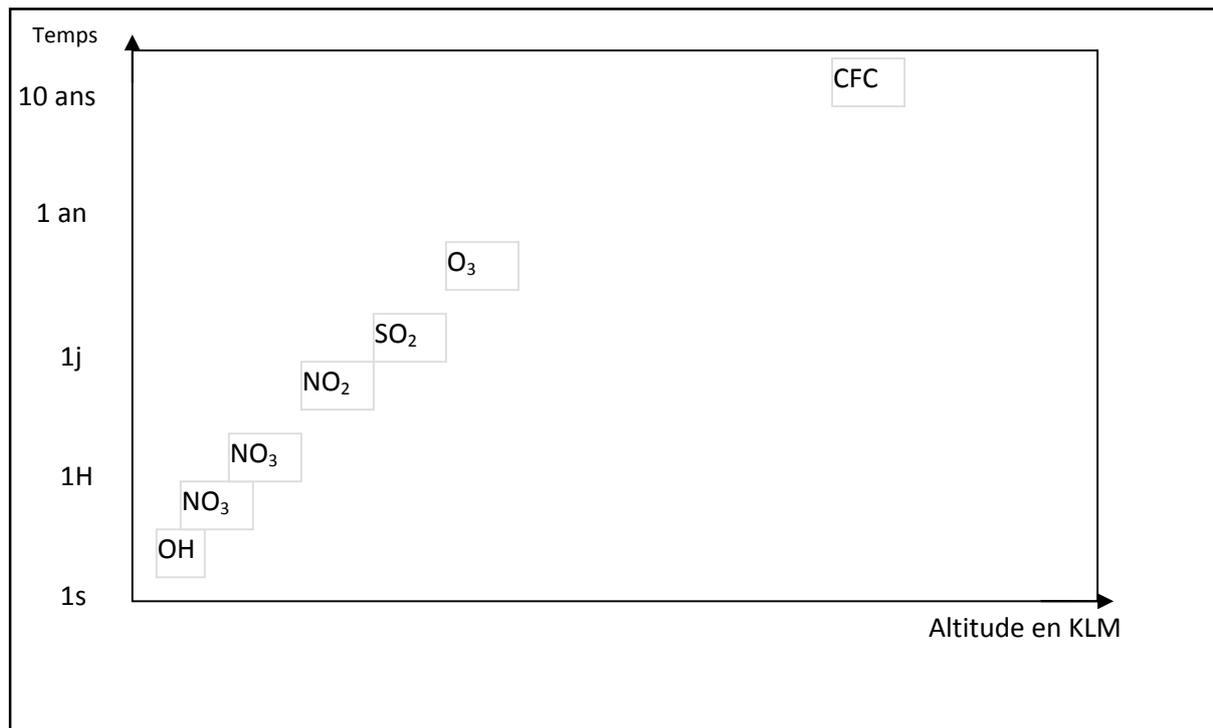
La pollution de l'air est très liée aux autres types de pollutions, les nuages de polluants se transforment et participent à l'augmentation du phénomène des pluies acides. Ces pluies transforment les polluants et les propagent sous d'autres formes : la pollution aquatique, la contamination et la pollution des sols etc (Park 2013).

## **2.5 Différents types de polluant**

La pollution atmosphérique se distingue selon les espèces considérées, la liste suivante contient les principales espèces dites 'polluantes' :

- $\text{SO}_2$  : le Dioxyde de soufre, lié à la combustion du charbon.
- $\text{O}_3$  : l'Ozone et les oxydes d'azote ( $\text{NO}$ ,  $\text{NO}_2$ ),
- COV : Composés Organiques Volantes.
- Métaux lourds, tel que : le plomb, le zinc, qui sont liés à certains types de processus industriels, sous forme particulaire.
- Mercure, émis sous forme gazeuse
- Aérosols, sous forme de particule et qui englobe le sulfate  $\text{SO}_2$ ,  $\text{NH}_4$ , sable, sels de mer etc.
- Polluants Organiques Persistants (POPs), chaînes organiques complexes avec un temps de vie important tels que : les pesticides, la dioxine.
- Les particules (particulate matter PM) : Les particules regroupent également une variété importante de substances. Le terme de particules désigne en général la fraction des composants (liquides ou solides) en suspension dans le milieu gazeux. Vu leurs poids très légers, elles restent dans l'air pendant de longues périodes (plusieurs dizaines d'heures). Souvent, les particules sont classées en fonction de leur granulométrie, par exemple les particules dont le diamètre aérodynamique moyen est inférieur à  $10 \mu\text{m}$  est noté  $\text{PM}_{10}$ , de même pour les  $\text{PM}_{2.5}$ .

Chaque polluant considéré va donc se caractériser par une classe d'espèces particulières et des propriétés spécifiques : sources d'émission, mécanismes physico-chimiques, temps caractéristiques, la figure suivante illustre le temps de résidence dans l'atmosphère de chaque polluant :



**Figure 3.2** Echelle spatio-temporelle de chaque polluant, source (Boubel et al).

### 2.5.1 Les polluants primaires

Les polluants primaires, sont ceux qui sont émis directement dans l'air, ces émissions sont catégorisées en deux types :

- Biogénique : liées à l'activité naturelle telle que l'activité photosynthétique, sels de mer, activité volcaniques, feux de forêts.
- Anthropique : liées à l'activité humaine tels que les processus industriels, transport, production de l'énergie etc.

Nous pouvons citer comme exemple de polluants primaires l'oxyde de soufre ( $\text{SO}_2$ ) qui est émis par l'opération de combustion de charbon.

### 2.5.2 Les polluants secondaires

Les polluants secondaires résultent des réactions photochimiques, dans l'air, dans lesquelles les polluants primaires jouent un rôle très important, un exemple de ce genre de polluants est l'Ozone, ce dernier résulte de la photochimie des oxydes de l'azote ( $\text{NO}_x$ ).

## 2.6 Les effets de la pollution de l'air et leurs principales sources:

Espèce	Sources	Effets			
		Sur les humains	Sur la végétation	Le climat	Les matières
Monoxyde de carbone (CO)	Combustion incomplète	Fourniture d'oxygène insuffisante ; cœur ; circulation ; système nerveux		Indirecte par la formation de l'Ozone	
Dioxyde de carbone (CO <sub>2</sub> )	Combustion			Effet de serre	
Hydrocarbures	Carburation, combustion incomplète	Cancérogène Précurseur de l'Ozone	Accumulation sur le sol	Formation de l'Ozone, et produit des gaz a effet de serre	
Oxyde d'Azote NO <sub>x</sub>	Oxydation de NO et NO <sub>2</sub> des carburants	Irritation du système respiratoire	Acidification des sols et des eaux, sur-fertilisation	Formation de l'Ozone et pause le problème d'effet de serre	Dégradation érosion
Particules	Combustion incomplète, poussière des chantiers, routes	Problèmes respiratoires et cardiaques, contenu toxique	Assimilation réduite		Saletés
Suie	Combustion incomplète	Cancérogène			Saletés
Ozone troposphérique (O <sub>3</sub> )	Photo oxydation des NO et HC	Système respiratoire et yeux	Désertification et sécheresse Et aggravation des crises d'asthme	sécheresse	Décomposition des polymères

**Tableau 3.2 : Les différents types de polluant et leurs effets.**

## 2.7 Les sources anthropiques de polluant:

La pollution de l'air provient de plusieurs activités humaines (anthropiques), nous pouvons citer :

- Les stations de production de l'énergie (utilisation du charbon et des hydrocarbures).
- L'agriculture : l'utilisation concentrée des composants chimiques fertilisants et les pesticides causent une dégradation préoccupante de la qualité de l'air dans les régions rurales.
- Les mines

- L'industrie, la sidérurgie, ces processus industriels émettent dans l'air de grosse quantité de PM et de NOx
- La circulation routière : L'activité routière occupe une place importante dans la dégradation de la qualité de l'air, dans les régions urbaines, elle est la première source de pollution. La combustion des hydrocarbures d'origine fossile, provoque beaucoup d'émission de polluants.
- l'incération des déchets : en absence d'une stratégie de recyclage efficace, parfois les procédures d'incération font transformer les déchets en poussières et gaz qui causent de sérieux problèmes dans l'air.

### **3 Les modèles de prédiction de pollution de l'air (état de l'art)**

L'existence de grosses quantités de données provenant des stations de surveillance de la qualité de l'air, a permis la conception des modèles de prédiction. Ces derniers visent à donner des prédictions concernant la concentration des polluants en utilisant des approches empiriques, qui se basent sur le principe que les mêmes conditions donneront certainement les mêmes résultats. Dans cette partie nous allons présenter, une revue des travaux récents dans le domaine de prédiction de la pollution de l'air.

#### **3.1 Les caractéristiques d'un modèle de prédiction**

Dans le domaine de prédiction de la qualité de l'air, les modèles peuvent être catégorisés selon plusieurs critères :

- Le type de polluant étudié (O<sub>3</sub>, PM10, PM2.5, CO, CO<sub>2</sub>, etc.), parfois des modèles font des prédictions concernant l'indice de la qualité de l'air qui est une agrégation de plusieurs valeurs de concentration de polluants.
- Le modèle utilisé : linéaire (ARIMA, AR, ARX, etc.), non-linéaire, RNA ou d'autres modèles issus de l'IA tels que la logique floue, Algorithmes Génétiques et des modèles hybrides incluant plusieurs approches (ex : ANN et logique floue).
- L'aspect temporel de la prédiction : certains modèles génèrent des prédictions à court terme et d'autres génèrent des prédictions à long terme ou même à moyen terme.

- Les paramètres utilisés : qui consiste à inclure ou non les paramètres climatiques, ou d'autres données relatives à l'émission des polluants (tel que l'activité routière et industrielle).
- L'implication ou non du processus de construction/ destruction du polluants. En fait, quelques modèles prennent en considération les interactions chimiques qui ont lieu dans l'air. Ces interactions entre polluants sont souvent modélisées par des formules mathématiques.
- Type et taille de la zone étudiée, certains modèles sont conçus pour donner des prédictions concernant une petite région (urbaine ou rurale) et d'autre concernant toute une région ou même un pays.
- Les données utilisées : la pluparts des modèles de prédiction sont des modèles orientés données (dirigés par les données), la quantité des données et les prétraitements de ces données, jouent un rôle important dans leurs performances.
- La validation : Un modèle de prédiction, peut être validé en comparant ses prédictions avec les mesures réelles. Les modèles aussi ne doivent pas donner des fausses alarmes, car la tendance d'un modèle à faire des fausses alarmes (faire des prédictions surestimées) est un critère à prendre en considération pendant sa conception, afin que les utilisateurs aient confiance dans les résultats que donne un modèle de prédiction.

Référence	Polluant	Données	Approches	Région
Rahman & al. 2014		Polluant et Météo	RNA(PMC), Logique floue (FTS) ARMIA	Johr, Malaisie
Moustris & al. 2013	PM10	Polluant et Météo	RNA	4 villes en Grèce
Azid & al. 2014	Indice de qualité de l'air	Polluant et Météo	ACP-RNA	Helsinki (Finlande),Thessaloniki (Grèce)
Russo & Soares 2014	PM10	Météo et PM10	RNA et DSS	Lisbonne (Portugal)
Al-Shammari 2013	Ozone	Température et Ozone	ARX Logique Floue	Koweït
Mlakar et Josef 1997	SO2		RNA (Carte de Kohanen)	
Jorquera et al 1998	Ozone		Logique floue ,ARX	
Gardner et Droling 1999a	NO2		RNA (PMC),ARX	Londres
Perez et al.2002	PM2.5	Mété et PM2.5	ARX, RNA	
Rob et al. 2001	Ozone		DVS (Delay Vectors), NARMA	
Jef et al. 2005	PM10		RNA ( OBD, optimal brain damage)	
Jorquera et al 1998	NO2		RNA	
Ordieres et al 2005	PM2.5		RNA (RBF, PMC)	
Sousa et al. 2007			ACP-RNA	
Cigizoglu et al 2006	SO2		RBF, PMC et ARX	Istanbul en Turquie
Sharma et al 2008	SO2, NO2 et PM10		Box-Jenkins	
Kim et kumar 2005	Ozone		ARX	Cincinnati, OH, USA
Castellano et al 2009	NOx et Ozone		ARIMA	Guarda, Corrubedo et Verín en Espagne
Carmichael et al. 2008	Ozone et PM10		ARIMA	
CIPRIAN et al 2002	Indice de qualité de l'air		Neuro-fuzzy, SBC	
Maruf et al 2008	NOx		HMM, HMM-fuzzy	
İçağa et Sabah 2008	PM10		RNA	
Shuiyuan et al 2006	PM		3-D multi-box (3DMB)	
Brunelli et al2006	SO2		RNA (ELMAN)	Priolo en Italie
Feng 2015	PM2.5	Climatiques et polluant	Wavelet et RNA	Peking

Tableaux 3.3 : Travaux sur la prédiction de la concentration de pollution dans l'air

D'après le tableau 3.3, il en ressort que l'approche RNA et l'approche la plus utilisée dans la prédiction des concentrations de polluant. Depuis des années, les réseaux de neurones artificiels sont devenus des alternatives aux méthodes statistiques classiques utilisées dans les stations de surveillance de la qualité de l'air. Ces méthodes qui se basent sur l'analyse des séries temporelles des différents polluants, pour détecter les tendances de chaque série, et donner des prédictions. Il serait très efficace d'utiliser les grosses quantités d'informations disponibles à partir des stations de surveillance pour la conception des modèles de prédictions efficaces dans le but d'avoir des systèmes de contrôle et de surveillance de pollution atmosphérique plus efficaces et utiles. Et qui permettent d'étudier le phénomène de pollution de l'air. Par exemple : savoir quels sont les paramètres qui influencent le plus sur la concentration des polluants dans l'air.

La relation entre les paramètres météorologiques et la pollution est très étroite, et complexe à modéliser. Dans (Rahman & al. 2014), trois modèles sont comparés pour la prédiction de l'indice de qualité de l'air. Le modèle basé RNA (PMC) a donné les meilleurs résultats et a surpassé les deux autres modèles basés logique floue et linéaire ARIMA. Les trois modèles ont été conçus en utilisant les données concernant la ville de Johr (dans le Sud de la Malaisie). Le modèle logique floue, a été conçu en utilisant la méthode FTS (Série Temporel Floue) qui consiste à découper le domaine des valeurs en plusieurs sous-domaines, et selon les données, extraire des règles. Chaque règle reçoit un degré de répétition (selon la base de données), ce degré représente après la probabilité que cette dernière soit appliquée. Par exemple dans le cas où des règles ont les mêmes conditions, le modèle appliquera celle qui possède la plus grande valeur de degré de répétition. Le modèle RNA a montré une bonne qualité de prédiction, non seulement pour les cas alarmants, mais aussi à baisser le nombre de fausses alarmes, et dépasse le modèle ARIMA conçu avec les mêmes paramètres.

(Moustris & al. 2013) décrit le développement d'un modèle de prédiction des PM10 et du nombre d'heures pendant lesquelles le taux des PM10 dépasse le taux recommandé. Le modèle consiste en un RNA qui a été construit en utilisant des données de quatre villes en Grèce.

Dans (Azid & al. 2014) un modèle basé RNA est présenté, le modèle utilise des entrées sélectionnées en utilisant la méthode d'analyse en composante principale (ACP). Les données de 7 ans de mesures concernant les paramètres de pollution de l'air dans les différentes villes en Malaisie ont été utilisées, le modèle n'utilise pas les paramètres climatique. Le but du modèle est de donner des prédictions de l'indice de qualité de l'air API (Air Pollution Index), le résultat du modèle a permis d'identifier les polluants qui influence le plus sur l'API, à savoir les (CH<sub>4</sub>, NmHC, THC, O<sub>3</sub> et les PM10) ceci a permis de conclure que la circulation routière est la

principale activité qui cause la dégradation, car c'est elle qui produit ces polluants. Le modèle a permis aussi d'identifier que l'élimination de certaines variables en entrée n'a pas causé de grande dégradation en terme de performance de prédiction. Une approche similaire est présentée dans (Voukantsis & al. 2011), elle consiste à sélectionner avec l'ACP les composantes qui regroupent des variables et après à passer ces composantes comme entrées à un modèle basé RNA, le modèle propose a été utilisé pour la prédiction de la concentration des PM<sub>2.5</sub> dans deux villes Helsinki (Finlande) et Thessaloniki (Grèce). Le modèle a permis de montrer que chaque ville possède une combinaison différentes pour les composantes principales, ceci est lié aux sources de pollution présentes dans chaque ville. Mais les deux villes ont présenté la même composante pour les polluants résultant du trafic routiers.

Dans (Russo & Soares 2014), un modèle hybride est présenté. Le modèle est composé de deux parties : la première consiste en un RNA qui utilise des données météorologiques pour donner des prédictions concernant la concentration des PM<sub>10</sub> dans l'air. La deuxième consiste à simuler la dispersion spatiale et temporelle de la pollution, ceci se fait à l'aide d'une méthode de distribution stochastique, elle consiste à définir une grille de  $N \times N$  cellules. L'algorithme consiste à visiter les cellules qui contiennent des stations de mesures, ces cellules là ont des valeurs de polluants (les prédictions faites par le modèles RNA et les mesures), après l'algorithme de DSS (Distribution Stochastique Simuler (Horta, & Soares 2010) est utilisé pour trouver la valeur de concentration dans les  $N \times M$  cellules restantes. Ce modèle permet de non seulement avoir des prédictions concernant les PM<sub>10</sub>, mais aussi avoir une idée concernant leurs distributions dans la région étudiée. Le modèle a été testé en utilisant les données concernant la ville de Lisbonne au Portugal.

(Al-Shammari 2013) décrit l'utilisation de deux modèles pour la prédiction de la concentration de l'Ozone troposphérique. Le premier modèle consiste en une régression linéaire ARX (Auto-Régression a Entrés eXogènes), l'entrée exogène pour ce modèle est la température. Le deuxième modèle est basé logique floue, qui consiste à extraire un ensemble de règle à partir des données empiriques. Les deux modèles ont été testés à l'aide des données de trois régions urbaines en Koweït. Le modèle basé logique floue a permis d'avoir les meilleures performances et a surpassé le modèle ARX.

La sélection des paramètres influence beaucoup sur l'efficacité du modèle, (Mlakar et Josef 1997) présentent une approche qui s'inspire de la reconnaissance des formes, qui se base sur la formulation (1.1) pour modéliser la prédiction et les paramètres :

$$B = \{V_i, O_i\} \cdot i = 1, 2, \dots, m \quad (3.1)$$

Où

$B$ : est un ensemble universel, la prédiction et les paramètres.

$V_i$ : est le pattern numéro  $i$ , la prédiction.

$O_i$ : est la description du pattern  $i$  (le vecteur des paramètres).

Le vecteur des paramètres  $V$  est défini comme suit :  $V_i = [x_1, x_2, \dots, x_r]^T$ . Avec le vecteur de sortie  $V_0 = [Y_1(t, \tau)] = SO_2(t + \tau)$  où  $\tau = 1/2h$  (une demi heure pour la prédiction à court terme). La réduction des paramètres consiste à réduire les paternes de  $r$  dimensions vers  $n$  dimensions où  $r \gg n$ , en utilisant un réseau de neurone de type Carte de Kohonen pour extraire les paramètres pertinents. Cela a permis d'optimiser le temps de l'apprentissage et aussi d'augmenter l'efficacité du modèle.

Pour trouver le modèle le plus adapté au problème de la prédiction de la qualité de l'air. (Jorquera et al 1998) présentent une comparaison entre trois modèles, réseaux de neurones artificiels, modèle linéaire, et modèle flous. La comparaison est faite en se basant sur les critères suivants :

- La capacité du modèle à prédire un épisode alarmant de la concentration de l'ozone.
- La tendance du modèle à déclarer des épisodes qui n'auraient pas lieu, (faire une prédiction alarmante qui ne se produise pas) ce dernier critère doit être le plus faible possible.

Un modèle mathématique a été utilisé pour modéliser l'émission de l'ozone, ce modèle se base sur l'équation de diffusion atmosphérique suivante :

$$\frac{\partial C}{\partial t} = -V \cdot \Delta C + \Delta (K \cdot \Delta C) + Q(x, y, z, t) R(\{C_k\}, T, t) + L(x, y, z, t) \quad (3.2)$$

Où

$C(x, y, z, t)$  : La concentration des l'espèce (polluants).

$V = (u, v, w)$  : Le vecteur du paramètre vent (direction et force).

$K = \text{diag}(Kx, Ky, Kz)$  : Diffusion de l'espèce (polluant).

$Q$  : Le taux d'émission.

$R$  : La balance entre la production et la destruction de l'espèce.

Dans le cas de l'ozone le paramètre émission peut être éliminé, car l'ozone n'est pas un polluant primaire. Il résulte des réactions chimiques des autres polluants spécialement le NO.

Le modèle définie par (3.2) peut être simplifié en :

$$\frac{\partial C}{\partial t} = R(\{C_k\}, T, t) \quad (3.3)$$

L'intégration de l'équation (3.3), entre le début de la journée ( $t_0$ ) et à  $t_0+3h$  donne :

$$O_3^t \text{ max} - O_3^t \text{ min} = \int_{t_0}^{t_0+3h} R(\{C_k\}, T, t) dt \quad (3.4)$$

Donc le modèle résultant est :

$$O_{3,t+1} = \alpha O_{3,t} + \beta T_{t+1} + \gamma T_t + \delta \quad (3.5)$$

Où :  $\alpha$  est positive,  $\beta$ ,  $\gamma$  sont des paramètres de différentes signes,  $\delta$  a une petite valeur et T c'est la température. Cette étude a permis de conclure que le modèle flou est le plus adéquat pour la prédiction de l'ozone, car il a montré des performances dépassant les autres modèles, notamment dans le nombre de fausse alarme (le plus faible par rapport aux autres). Cela ne veut plus dire qu'il le soit pour d'autres régions, car chaque région possède ses propres paramètres qui affecteront les performances du modèle.

(Gardner et Droling 1999a) présentent un modèle pour la prédiction de la concentration de NO<sub>2</sub>, avec une comparaison entre un régresseur linéaire et un modèle neuronal de type PMC, le modèle neuronal a dépassé significativement le modèle linéaire, utilisant les mêmes paramètres. Dans (Gardner et Droling 1999b), les tendances de la concentration de l'ozone sont modélisées à l'aide d'un réseau de neurone de type PMC, la série temporelle de la concentration de l'ozone est modélisées comme suit:

$$O_3(t) = E(t) + S(t) + W(t) \quad (3.6)$$

Où  $E(t)$  : le composant long terme.  $S(t)$  : le composant saisonnier.  $W(t)$  : le composant court terme.

Le composant long terme représente la variation de la concentration de certaines gazes qui participent dans le cycle de production / destruction de l'ozone, et le changement climatique.

Le composant saisonnier représente le changement annuel de la radiation solaire et de direction de vent. Le composant court terme représente les changements météorologiques d'un jour à l'autre.

Cette étude a permis de concevoir un modèle qui permet de détecter les tendances à long terme de la concentration de l'ozone pour la région de Londres.

La prédiction à court terme de la qualité de l'air est de l'ordre de quelques heures, cela est dû au temps de résidence des polluants dans l'air, un exemple de prédiction à court terme est présenté dans (Perez et al.2002) avec une comparaison entre trois modèles : réseau de neurones, régression linéaire et le modèle persistant. Tous ces modèles ont été utilisés pour la prédiction à court terme (avant quelques heures) de la concentration de PM 2.5 (les particules volantes dont leur diamètre est 2.5 micromètre). Cette étude a montré que le modèle neuronal est le plus performant.

La prédiction à long terme dans le domaine de la prédiction de la pollution de l'air ne dépasse pas 30 heures. (Pérez et Reyes 2002) présentent un modèle basé PMC pour la prédiction de la concentration de PM10 avant 30 heures. Le modèle reçoit en entrée les mesures de la concentration de PM10 à 9 : AM (dans la matinée), et aussi les paramètres météorologiques prédits. Le modèle proposé a montré une grande efficacité à prédire les journées polluées par rapport à un modèle linéaire, qui est le réseau lui-même avec la fonction d'activation neutre  $f=1$ .

Dans (Rob et al. 2001) une approche de détection de non linéarité est présentée, elle se base sur l'analyse de la série temporelle et l'application de la méthode DVS (Delay Vector). Elle consiste à calculer pour chaque vecteur une distance et d'extraire les  $K$  vecteurs les plus proches (dont la distance est plus faible). Ce modèle est utilisé pour calculer les prédictions en utilisant le sous-ensemble de test, et les erreurs sont calculés en variant le nombre de voisins pris ( $K$ ), jusqu'à avoir une valeur qui donne le plus d'optimalité pour l'erreur. Cette valeur indique la non-linéarité de la série temporelle comme suit :

- Si  $K$  est proche au nombre total des vecteurs de retardement, alors le modèle linéaire est plus adaptable pour la série, et le modèle est équivalent à un AR
- Si  $K$  est plus petit ou moyen le modèle non linéaire est plus adéquat.

En utilisant cette méthode une analyse a été faite sur la série temporelle de la concentration de l'ozone, une fois que la non linéarité a été détectée deux modèles ont été conçus. Un de topologie Feed-Forward et un autre de topologie récurrent dite NARMA(p,q) (Nonlinear Autoregressive Integrated Moving Average) . Cette étude a montré que le modèle NARMA est un outil très efficace pour la prédiction de la concentration de l'ozone, et que les deux topologies ont des performances très semblables.

L'influence des paramètres météorologiques sur la précision des prédictions a été bien étudiée dans (Jef et al. 2005). Pour chaque paramètre les auteurs ont mis en évidence son influence sur les performances du modèle, cela s'est fait en construisant pour chaque paramètre un réseau de neurone et utiliser la concentration de PM10 précédente et le paramètre en question. Ce qui a permis de concevoir 5 modèles pour comparer les résultats et pour détecter le paramètre qui a le plus d'influence sur les performances du modèle. Le paramètre qui a montré le plus d'importance est le BLH (la masse atmosphérique de la journée). Ce paramètre influence directement sur le temps de séjour des particules dans l'atmosphère), reste à dire que cette conclusion n'est valable que pour la région concerné par l'étude. Une autre méthode proposée par (Giorgio Corani 2005) qui consiste à éliminer le paramètre qui possède le poids le plus faible dans un réseau à feed-forward, par l'application de la méthode OBD (Optimal Brain Damage (Strok et Hassibi 1993)). Cette méthode fait tendre les poids les moins influents à zéro, éliminant les neurones les moins significatives. Ceci permettra d'obtenir une topologie adéquate pour le PMC. Dans cette étude, les auteurs présentent un autre modèle Lazy learning qui se base sur la notion de point de requête (Query-point), qui consiste à un vecteur  $Q(t) = [Q_1(t), Q_2(t), \dots, Q_n(t)]$ , qui correspond aux valeurs des paramètres pendant le temps (t), et un autre point dit voisin qui est un vecteur  $U$  disponible dans la matrice des données, où  $U(\bar{t}) = [u_1(\bar{t}), u_2(\bar{t}), \dots, u_n(\bar{t})]$ , qui correspond à la sortie  $Y(T+1)$  déjà connus, en calculant une distance entre le point de requête et tous les points voisins, cela permet de trouver les  $K$  points les plus proches, c'est-à-dire avoir à prendre  $K_{min} < k < k_{max}$ , et à prendre la valeur adéquate par rapport au point de requête. Le principe de ce modèle est très semblable au raisonnement à base de cas CBR. Comparé avec un réseau de neurone de type FeedForward, le modèle LL a été le plus performant pour prédire le PM10 et l'Ozone, mais comme pour toutes les études cela n'est valable que pour la région de Milan en Italie.

La plupart des modèles proposés génèrent la valeur de la concentration du polluant étudié, dans (Jorquera et al 1998) le modèle proposé doit fournir une décision concernant la classe de qualité de l'air (A: good, B: bad, C: critical). Le modèle possède trois sorties, chacune d'entre elles représente une classe de la qualité de l'air, le modèle utilisé est un réseau de neurones artificiels de type PMC. Ce modèle a été comparé avec un modèle linéaire (un réseau de neurone sans fonction d'activation qui est équivalent à un modèle ARX), les deux possèdent le même nombre de paramètre, ont montré des prédictions très affinés, mais le modèle non linéaire (le PMC) dépasse significativement le modèle linéaire équivalent. Ceci a été aussi montré par (Bianchini et al.2006) après création de N réseaux de neurones artificiels de type PMC, où N est le période de la cyclo-stationnarité de la série temporelle de la concentration de NO<sub>2</sub>, le modèle a donné des prédictions efficaces.

Plusieurs RNA de différents types (PMC, RBF, PMC quadratique) sont comparés dans (Ordieres et al 2005), ces trois modèles ont été utilisés pour la prédiction de la concentration des PM2.5. Le modèle qui a montré le plus d'efficacité a été le RBF (Radial Basis Function). (Sousa et al. 2007) appliquent un traitement ACP (Analyse en Composante Principale) aux vecteurs de paramètres pour réduire la complexité des modèles, l'utilisation de cette méthode a permis d'améliorer les performances des prédictions d'un modèle neuronal de type PMC.

Les modèles proposés dans (Cigizoglu et al 2006) permettent de donner des estimations sur la concentration de dioxyde de soufre (SO<sub>2</sub>) dans la ville d'Istanbul en Turquie, en prenant comme paramètres d'entrées la concentration mesurée de monoxyde de charbon et la concentration mesurée du monoxyde du nitrate. Dans cette étude trois modèles sont proposés : Un modèle se base sur un réseau de neurones de type RBF (Radial Bassis Function), un modèle basé PMC, et un modèle basé régresseur linéaire. Ces trois modèles ont été utilisés pour faire des prédictions long-terme de SO<sub>2</sub>, les deux modèles neuronaux ont montré des bonnes performances, quoique le modèle basé RBF a présenté des comportements de surestimations des pics. Les mêmes modèles ont été utilisées dans (Cigizoglu et al 2006), pour l'étude de l'éventuel influence des concentrations de CO et du NO sur la concentration de PM10. Quartes modèles ont été conçus : un basé PMC prenant comme paramètres la concertation de CO, un autre de même topologie mais prenant en paramètre la concertation mesurée de NO, un autre Modèle basé RBF prenant en paramètre la concentration mesurée de CO et le dernier de même type mais avec la concentration NO comme entrée. Le choix des paramètres a été validé par l'application de l'ACP (Analyse en Composante Principal) ce qui a permis de voir le degré de l'importance de chaque paramètre (polluant). Les quartes modèles ont permis de voir clairement la relation et

l'influence des concentrations des polluants primaires tels que CO et NO sur la concentration des PM10.

Une autre approche d'étude de relation entre la concentration des polluants secondaires et les concentrations des polluants primaires est présentée dans (Acciani et al 2006). Deux modèles basés réseau de neurones artificiels sont utilisés pour étudier les paramètres qui influencent le plus sur la concentration de l'Ozone O<sub>3</sub> et de PM10. Ces deux polluants sont de types polluants secondaires qui résultent de l'interaction chimique des polluants primaires. Pour voir le polluant primaire et le paramètre climatique le plus important, une méthode dite de partitionnement est appliquée. Cette dernière se base sur la mesure du « index d'importance relative » RII (Relative Importance Index) qui permet de voir combien un élément d'entrée affecte-il la sortie. L'index d'importance relative se base sur les poids d'un neurone dans un réseau de type PMC. Après l'étape d'apprentissage et il est donné comme suit :

$$RI = \frac{\sum_{j=1}^h \left| \frac{w_{ij} w_{kj}}{\sum_{i=1}^n |w_{ji}|} \right|}{\sum_{i=1}^n \sum_{j=1}^h \frac{|w_{kj}| |w_{ji}|}{\sum_{i=1}^n w_{ji}}} \quad (3.7)$$

Où :

$w_{ij}$  c'est le poids de connexion de l' $i$ -ème neurone d'entrée avec le  $j$ -ième neurone de la couche cachée, et  $w_{kj}$  c'est le poids de connexion entre le  $j$ -ième neurone de la couche cachée avec le  $k$ -ième neurone de la couche de sortie. L'index d'importance relative **RI** est donc le degré d'importance qui lie le paramètre  $i$  avec la sortie  $k$ , passant par le neurone  $j$  de la couche cachée. Une sélection des paramètres est appliquée afin de prendre en considération que ceux qui ont un index d'importance relative important, un autre RNA qui prend comme paramètre que les paramètres sélectionnés, ce qui permettra d'optimiser le modèle et d'augmenter ses performances. L'approche peut être résumée comme suit :

1. Conception d'un réseau de neurone de type PMC à une seule couche cachée avec tous les paramètres,

2. Calcule pour chaque paramètre l'index d'importance relative,
3. Sélection des paramètres qui ont le RII le plus important,
4. Conception d'un autre réseau avec les paramètres sélectionnées en 3.

Paramètres utilisés dans l'étude :

Paramètres	Unités
SO <sub>2</sub>	µg/m <sup>3</sup>
NO <sub>x</sub>	µg/m <sup>3</sup>
NO	µg/m <sup>3</sup>
NO <sub>2</sub>	µg/m <sup>3</sup>
CO	µg/m <sup>3</sup>
Benzène	µg/m <sup>3</sup>
Toluène	µg/m <sup>3</sup>
O-xilème	µg/m <sup>3</sup>
Direction du vent	Secteur
Vitesse du vent	m/s
DVG	Secteur
Radiation	W/m <sup>2</sup>
Pluviosité	Mm

**Tableau -3.4- Paramètres utilisés.**

Cette étude a montré que le CO, vitesse du vent et la pluviosité sont les paramètres les plus influençant sur la concentration de L'Ozone O<sub>3</sub>. Dans le cas de PM10 CO, le Benzène et la vitesse du vent sont plus importants par rapport aux autres paramètres. Pour avoir une bonne prédiction de la qualité de l'air (Kolehmalnen et al 2000) ont proposé une approche de datamining, qui permet de construire un modèle pour chaque épisode de la série temporelle de pollution. Deux méthodes de clustering ont été appliquées, une basée sur les cartes de Kohonen et une autre basée sur les cartes de Sammon, ces deux méthodes ont permis d'avoir sept épisodes de pollution, pour chaque épisode un réseau de neurones de type PMC a été conçu pour prédire l'épisode, ce qui a permis d'avoir une prédiction multi-modèles où chaque modèle est spécialisé dans la prédiction d'une partie de la série temporelle.

La non-linéarité de la série temporelle de la concentration de l'ozone est bien étudiée dans (Bordignon et al 2002). Trois modèles ont été conçus pour faire des prédictions à court-terme de la concentration de l'ozone  $O_3$ , un modèle basé modèle Additif, un modèle basé MARS (Multivariate Additive Regression Splines), et un modèle basé réseau de neurones artificiels. L'approche adoptée consiste à faire une représentation mathématique et à chercher le modèle le plus performant qui simule le plus fidèlement possible la relation entre  $Y$  les prédictions et le vecteur des paramètres  $\mathbf{X}$ . les trois modèles ont montré de bonne performance, la seule critique c'est que les méthodes non-linéaires ne permettent pas une étude explicative du phénomène malgré leurs résultats incontestables.

La prédiction à court-terme des concentrations de dioxyde de soufre  $SO_2$  et dioxyde de nitrogène  $NO_2$  et  $PM_{10}$  a été proposée dans (Sharma et al 2008), l'approche Box-Jenkins a été adoptée pour concevoir des modèles basés ARIMA (autoregressive integrated moving average). L'approche peut être résumée comme suit :

- Identification des spécifications préliminaires du modèle,
- Estimation des paramètres du modèle,
- Diagnostique et validation de l'adaptabilité du modèle,
- Prédiction.

La série temporelle est composée en plusieurs parties : autorégressive notée AR, intégration I, Moving Avarage notée MA. Supposant que  $Z_t$  est notre série temporelle (définie par une séquence de  $N$  observations qui sont équidistantes dans le temps  $(z_1, z_2, \dots, z_N)$ , l'opérateur d'intégration et s définie comme suit :

$$\Delta z_t = z_t - z_{t-1} \text{ et } \Delta_s z_t = z_t - z_{t-s} \quad (3.8)$$

Où  $s$  est la période ou le pas de mesure des observations.

Mettant l'opérateur  $B$  comme suit :

$$B z_t = z_{t-1} \text{ Et } B^s z_t = z_{t-s} \quad (3.9)$$

La partie AR notée  $\Phi$  se constitue d'un polynôme d'ordre  $p$  de  $B$  comme suit:

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.10)$$

La partie MA notée  $\theta$  définie comme suit :

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (3.11)$$

La première étape consiste à appliquer une analyse de la série temporelle par l'utilisation d'ACF (Autocorrelation Function) qui permet de mesurer le degré de dépendance linéaire entre les observations. L'application de PACF (Partial Autocorrelation Function) permet aussi de voir l'importance de la partie AR dans le modèle. Après application des règles citées dans (Box et al 1994). Le modèle issue est de classe ARIMA (p, d, q) est peut être formalisé comme suit :

$$\phi_p(B)\theta_q(B)\Delta_s^d \Delta^d x_1 = \theta_q(B)\Theta(B^s)a_t + c \quad (3.12)$$

Après application de l'algorithme MSQ (Marquardt least-square) les paramètres de modèle ont été identifiés et le modèle est formalisé comme suit :

$$\begin{aligned} [w_{t+l}] = & \phi'_1[w_{t+l-1}] + \phi'_2[w_{t+l-2}] + \dots + \phi'_{p+sp}[w_{t+l-p-sp}] + [a_{t+l}] - \\ & \theta'_1[a_{t+l-1}] - \theta'_2[a_{t+l-2}] - \dots - \theta'_{q+sq}[a_{t+l-q-sq}] + c \end{aligned} \quad (3.13)$$

Où :

- $L$  : est le temps de prédiction,
- $W_{t+l}$  et  $\Phi_l$  sont les paramètres AR généralisés,
- $\theta_l$  sont les paramètres MA.

Les résultats obtenus dans cette étude ont montrés que les modèles stochastiques linéaires tel que ARIMA permettent d'avoir des descriptions quantitatives de la pollution de l'air. L'approche permet d'avoir des prédictions efficaces avec moins d'effort de calcul. Un autre modèle linéaire est présenté dans (Kim et al 2005) un modèle basé ARX (Autoregressive à entrée exogène) est proposé pour la prédiction à court-terme de la concentration de l'ozone  $O_3$  dans la ville de Cincinnati, OH, en Etats Unies d'Amérique (USA).

Le modèle ARX formalisé comme suit :

$$y_t = \mu + \sum_{j=1}^p \phi_j y_{t-j} + \beta' x_t + w_t, t = q + 1, \dots, n \quad (3.14)$$

Avec  $\Phi p \neq 0$ , et p l'ordre d'autorégressive,  $Xt = (x1t, x2t, \dots, xlt)$  est le vecteur des entrées exogènes. Dans cette étude plusieurs modèles ont été conçus :

- FARX (Fonctional Coefficient AutoRegressive with eXogenous inputs) : permet d'ajouter les paramètres direction du vent et temps de la journée (nuit ou matin),
- TARX (Tong AutoRegressive with eXogenous inputs) ,
- AR modèle autoregressive,
- ARX. (AutoRegressive with eXogenous inputs)

Chaque modèle a montré une certaine capacité à prédire correctement la concentration du polluant, quoique le modèle basé FARX a été le plus performant.

Dans (Castellano et al 2009) une étude statistique des séries temporelles de NO<sub>x</sub> et l'Ozone est présentée. Dans cette étude une approche statistique d'indentification des épisodes de pollution a été suivie, l'analyse statistique a été appliquée sur les données des paramètres de pollution de trois zones en Espagne : Guarda, Corrubedo et Verín (deux ville côtières et une ville intérieure). Pour chaque ville un modèle basé ARIMA a été conçu pour donner des prédictions à long-terme sur la concentration de O<sub>3</sub>, parmi les paramètres inclus dans le modèle nous pouvons citer : SO<sub>2</sub>, SH<sub>2</sub>, NO, NO<sub>2</sub>, CO, l'Ozone, PM10, la température, Vitesse et direction du vent, La pluviométrie, La pression atmosphérique, Humidité et rayonnement solaire.

Dans (Carmichael et al. 2008) une approche CTM (Chemical Transport Models) a été utilisée pour la conception d'un système de contrôle de l'émission des polluants dans l'air, qui consiste à trouver un modèle mathématique qui permet d'inclure les éléments suivants : émission, transport, transformation et destruction des pollutions dans l'air. En se basant sur les paramètres chimiques du polluant et de son processus de transformation et aussi de transport (notamment les paramètres météorologiques) un modèle ARIMA a été conçu pour la prédiction de la concentration de l'Ozone et de PM10. Le modèle a présenté des résultats intéressants pour la prédiction de la concentration des polluants secondaires, tel que l'Ozone (O<sub>3</sub>). Mais l'adaptation d'un tel modèle pour un polluant primaire reste une tâche très complexe surtout quand les données sur l'émission de ce polluant ne sont pas disponibles. Une approche hybride est présentée dans (Ciprian et al 2002), dans laquelle un modèle Neuro-fuzzy a été intégré dans un système à base de connaissance dans le but de prédire la qualité de l'air. Un modèle connexionniste (Neuronal) est créé en se basant sur les données disponibles afin d'extraire les connaissances implicites, en plus un modèle basé logique floue est créé en se basant sur les connaissances explicites des experts du domaine de pollution de l'air, le schéma global du modèle est illustré dans la figure 1.4. A chaque fois un réseau de neurones artificiels est créé. Ensuite et après les étapes d'apprentissages des méthodes d'extraction des connaissances sont appliquées

pour extraire les informations symboliques du réseau créée. Un extrait des règles produites est montré dans le tableau -1.3- :

IF RainToday is YES THEN NO2 After10 is L1 (74.80%)
IF InversionToday is L0 THEN NO2 After10 is L1 (58.95%)
IF InversionToday is L3 THEN NO2 After10 is L2 (61.16%)
IF SolRadAt13 is LOW THEN NO2 After10 is L2 (95.10%)
IF SolRadAt13 is MED THEN NO2 After10 is L2 (100.00%)
IF SolRadAt13 is HIGH THEN NO2 After10 is L2 (45.70%)
IF WindDirect_10 is NE THEN NO2 After10 is L3 (47.47%)
IF WindSpeed_10 is LOW THEN NO2 After10 is L2 (34.07%)
IF NO_10 is MED THEN NO2 After10 is L2 (31.54%)

**Tableau -3.5- : Un extrait des règles construites.**

Le modèle a été utilisé pour la prédiction de pics de concentration de NO<sub>2</sub> et a montré de bonnes performances. Un autre modèle hybride est présenté dans (Maruf et al 2008), ce modèle consiste à l'utilisation de HMM (Hidden Markov Model) et la logique floue (HMM-fuzzy) pour prédire la concentration du NO<sub>x</sub> et NO<sub>2</sub>. L'approche consiste à appliquer les étapes suivantes :

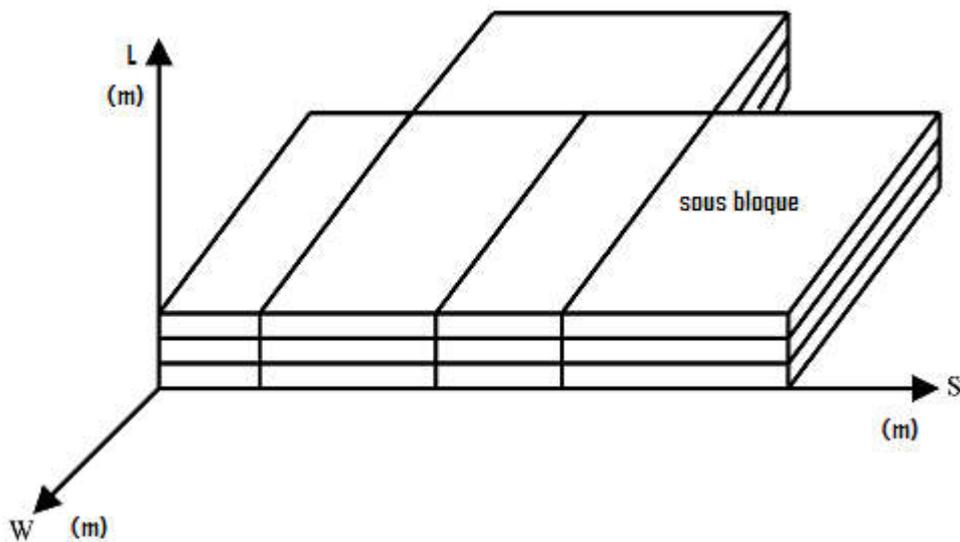
- L'HMM est utilisé pour partitionner le sous-ensemble d'apprentissage,
- Un Algorithme Top-Down est appliqué pour extraire le nombre minimum de règles (logique floue) qui permettent d'atteindre l'erreur calculée dans l'étape 1,
- L'algorithme descente gradient est appliqué pour extraire les paramètres du modèle.

Le modèle prend en entrée les paramètres de l'émission, les paramètres métrologiques, et fournit en sortie la prédiction de la concentration de NO<sub>2</sub>. Le modèle a montré de bonne performance. Dans la plupart des études un seul polluant est étudié, dans (İçağa et Sabah 2007) la relation entre deux polluants est étudiée. Un modèle basé PMC a été conçu pour détecter l'éventuelle relation entre dioxyde de soufre SO<sub>2</sub> et les PM (particule en suspension dans l'air). Cette étude a permis de conclure que l'humidité est l'un des facteurs les plus importants à affecter la relation entre ces deux polluants. Ce qui a permis de tirer des conclusions à propos de l'influence de SO<sub>2</sub> et l'humidité sur le processus de transformation / destruction de PM10. Dans (Shuiyuan et al 2006) un modèle de prédiction et de simulation basée 3-D multi-box (3DMB) est proposé. Le modèle prend en considération les paramètres relatives à l'émission, c'est pour ça que la zone étudiée a été découpée en 4 sous-zones : une zone industrielle, une zone semi-industrielle

(a activité industrielle légère), une zone urbaine, une zone touristique et culturelle. Pour chaque sous-zone un modèle 3DMB est conçu pour prédire la concentration des TSP (Total Suspended Particulate) et l'SO<sub>2</sub> et les PM10. Ce type de modèle se base sur :

- Les données d'émission,
- Le principe de conservation de masse (la capacité de l'air à garder les composants polluants),
- Dispersion tridimensionnelle des polluants,
- Distribution de polluant dans l'air,
- Interaction chimique entre les composants de l'ait et le polluant.

Un exemple de Box-Gaussien peut être illustré comme dans la figure -1.5-



**Figure -3.3- : Modèle en box de pollution de l'air.**

Chaque sous-cube est formulé par l'équation (3.15). Les sous-cubes permettent de rendre la prédiction plus précise en indiquant à chaque fois la sous-zone qui donne le plus de concentration. Ceci permet de bien localiser les pics de pollutions et permettra aux utilisateurs d'avoir plus de détails sur le phénomène afin prendre les mesures nécessaires pour limiter et contrôler la pollution.

$$\begin{aligned}
 &L_5 H U_{n3} C_{n03} - L_5 H U_{n3} C_{n15} + E L_5 W_5 (C_{n10} - C_{n15}) / H \\
 &+ 2 E' W_2 H (C_{n12} - C_{n15}) / (L_2 + L_5) \\
 &+ 2 E' W_3 H (C_{n13} - C_{n15}) / (L_3 + L_5) \\
 &- L_5 H U_{n3} C_{n15} [1 - \exp(-K W_5 / U_{n3})] \\
 &- L_5 W_5 C_{n15vg} - L_5 W_5 W_D C_{n15} P / 4 = 0
 \end{aligned} \tag{3.15}$$

Où:

$L_i$ ,  $W_i$ : respectivement la longueur et la largeur de sous-cube,  $M$ : la moyenne de vitesse du vent,  $C_n$ : la concentration de polluant dans la n-ième sous-cube,  $E$ : coefficient de dispersion vertical,  $H$ : coefficient de dispersion horizontal,  $P$ : pression. Ce modèle a permis de donner de bons résultats, et permet d'avoir la possibilité de bien localiser la pollution en divisant les zones étudiées en sous-zones, mais ce type de modèle n'est valable que pour les zones dont les paramètres et les données relatives à l'émission sont disponibles. Une étude comparative entre l'approche linéaire et non-linéaire est présentée dans (Comrie 2007), cette étude a été basée sur les données relatives à la concentration de l'Ozone ( $O_3$ ) qui est un polluant secondaire (résultant de l'interaction photochimique des autres polluants), ce qui a favorisé le modèle linéaire, car pour d'autres polluants primaire le modèle risque de ne pas avoir les paramètres relatifs à la transformation (construction/ destruction) du polluant. Dans (Brunelli et al 2006), un modèle basé réseau d'Elman est présenté, le modèle donne des prédictions concernant la concentration des  $SO_2$  dans les zones industriels de Priolo en Italie, le modèle reçoit comme paramètres la concentration passé du  $SO_2$  et aussi les valeurs de l'émission des sites industriels ciblés par l'étude. Le modèle a donné de bonnes performances de prédiction.

## 4 La modélisation de la dispersion du polluant atmosphérique

Après avoir présenté les méthodes utilisées pour prédire la concentration des polluants, nous nous intéressons dans cette section, aux modèles utilisés pour simuler la dispersion d'un polluant dans l'air. Les modèles de prédiction de la qualité de l'air donnent des informations très importantes concernant la pollution de l'air, mais ne donnent pas d'information concernant les sources et la dispersion des polluants dans l'air. Par contre, les modèles de dispersion permettent de connaître la dispersion ainsi que l'impact des sources, les paramètres climatiques et les réactions chimiques entre les différents polluants dans l'air. (Holmes et Morawska 2006) présentent une revue assez détaillée des modèles de dispersion.

En fait, une fois émis dans l'air, les polluants se dispersent selon un certain nombre de lois physiques et chimiques. Bien sûr, le type de la source de pollution représente l'élément principal. Selon la terminologie du domaine ils existent trois types de sources :

- Source de type point : ce genre de source représente les cheminés, les sorties d'évacuations des gaz dans les processus de fabrication.
- Source de type ligne : ce type représente les routes et l'activité routières.

- Source de type surface : ils représentent les terres agricoles, les stations de traitement des déchets, les feux de forêts, etc.

Tous les types de sources possèdent un ensemble de paramètres qui sont pris en considération, parmi ces paramètres : le taux d'émission, la position et l'hauteur de la source, le polluant émis.

Le but d'un modèle de dispersion de pollution est de donner la concentration du polluant selon la position par rapport à une source d'émission. Ils existent plusieurs modèles qui permettent de trouver la dispersion de la pollution de l'air.

#### **4.1 Le modèle en boîte (Box models)**

Ce genre de modèle représente le site modélisé comme étant une boîte dans laquelle les polluants sont émis. Après, des réactions physiques et chimiques sont simulées pour trouver la concentration du polluant. Il est à noter que la concentration du polluant est considérée être la même dans toute la boîte. Le modèle utilise aussi la vitesse du vent et le taux d'émission, il peut être formulé comme suit :

$$\frac{dCV}{dt} = Q \times A + u \times C_{in} \times W \times H - u \times C \times W \times H. \quad (3.16)$$

Où,  $CV$  c'est la concentration du polluant,  $Q$  c'est le taux d'émission de la source,  $V$  volume de la boîte,  $C_{in}$  la concentration de polluant entrant,  $A$  est la hauteur de la boîte,  $u$  la vitesse du vent. La principale limitation de ce genre de modèle est le fait qu'il ne donne pas la concentration selon une position exacte, car il considère que la concentration du polluant est homogène dans toute la boîte. Vu leur représentation simple de l'air, leurs points forts est qu'ils permettent d'inclure la simulation des réactions chimiques qui ont lieu dans la boîte (qui représente la région modélisée).

## 4.2 Le modèle Gaussien de dispersion

Les modèles Gaussiens sont une famille de modèles mathématiques de dispersion. Ils se basent sur le principe que la dispersion suivra une loi normale uniforme (Gaussienne). C'est le modèle le plus utilisé (Daly & Zannetti 2007).

$$C(x, y, z, H) = \frac{er_{i,t} D}{2\pi U_t \sigma_y \sigma_z} e^{-\frac{y^2}{2\sigma_y^2}} \left[ e^{-\frac{(z-H)^2}{2\sigma_z^2}} + e^{-\frac{(z+H)^2}{2\sigma_z^2}} \right] \quad (3.17)$$

Ceci veut dire que, la concentration du polluant au point  $(x, y, z)$  est calculée selon :

$er_{i,t}$ : Le taux d'émission en kilogramme par heurs de la source  $i$  en temps  $t$ .

$U_t$ : c'est la vitesse du vent en mètre par seconde à temps  $t$ ,  $\sigma_y, \sigma_z$ : l'écart type des distributions de concentration dans le sens de la direction du vent, ces deux paramètres sont sélectionnés selon la classe de stabilité (Guifford-Pasquill), et  $H$  c'est l'hauteur de la source par rapport à la tère. Le terme  $D$  formule l'effet de décomposition (EPA, 1995) et calculé selon (10).

$$D = \begin{cases} e^{(1/(R \frac{x}{u}))}, & \text{if } R > 0, \\ 1, & \text{if } R = 0 \end{cases} \quad (3.18)$$

Où  $x$  est la distance dans la direction du vent,  $u$  c'est la vitesse du vent et  $R$  c'est le coefficient de décomposition.

La limitation principale des modèles GD réside dans le fait qu'ils ne tiennent pas en compte le temps de déplacement des polluants. Pour remédier à ça, des modèles incluent des traitements complémentaires des résultats. Des Modèles GD récents incluent aussi les réactions chimiques et les effets de décomposition physique. D'autre modèles GD permettent de simuler la dispersion à partir d'une source de type ligne (Ristic & al. 2015).

## 4.3 Le modèle Lagrangien

Ce type de modèle permet de modéliser la dispersion en simulant les mouvements des nuages de pollution (appelés aussi des modèles de transport). Il définit une boîte d'air contenant une concentration initiale de polluant (comme celle de modèle en boîte), après il prédit les mouvements du nuage selon la direction du vent. La pollution donc est le produit du terme d'émission et de la probabilité concernant la densité. Il peut être formulé comme suit :

$$C(r, t) = \iint_0^T p(r, t|r', t')S(r', t')dr' dt' \quad (3.19)$$

Où,  $C(r,t)$  c'est la concentration du polluant dans la position  $r$  au temps  $t$ ,  $S(r',t')$  terme de source d'émission,  $p(r,t|r',t')$  c'est la probabilité que les conditions permettront de déplacer l'air de la position  $r'$  à la position  $r$ .

D'autres types de modèles sont aussi plus ou moins utilisés, parmi ces modèles nous pouvons citer : CTM : modèle de transport chimique, CFD : modèle de dynamique des fluides Les modèles des gaz denses. Dans certains cas, ces modèles sont intégrés ensemble dans un seul système de simulation, le tableau 3.5 présente quelques système de simulation de dispersion de pollution de l'air et les modèles qu'ils les intègrent.

Modèle	Type de modèle
AURORA, PAL2, CPB PBM	Modèle en boîte
CALPUFF, AEROPOL, AERMOD	Modèle Gaussien
GRAL, ARIA	Modèle Eulérien et Lagrangien
MISKAM, ARIA Local	CFD

**Tableau 3.5 : Liste de quelques systèmes de modélisation de dispersion et leurs modèles.**

## 5 Conclusion :

Nous avons pu voir dans ce chapitre les différents concepts liés à la pollution de l'air. Les effets de chaque type de polluants ainsi que leurs origines sont aussi décrits. Nous avons aussi présenté les outils et les approches utilisées pour modéliser la pollution de l'air. Les approches peuvent être catégorisées en deux familles selon leurs objectifs.

- La première consiste à concevoir des modèles qui permettent de donner des prédictions concernant la concentration des polluants dans l'air en se basant sur les données empiriques.
- La deuxième approche vise la conception des modèles qui permettent de simuler la dispersion dans l'air, des polluants résultant des différentes sources, en utilisant les données concernant les taux d'émission des sources.

Un état de l'art rapportant les études récentes qui traitent ces sujets est présenté pour chacune des deux approches.

Nous pouvons résumer ce chapitre en concluant que les approches actuelles sont concentrées sur l'aspect physique de pollution (concentration et dispersion). Par conséquent, ils ignorent l'aspect anthropique lié à la décision et l'activité humaine. Les décisions humaines ne sont pas prises en considération. Nous espérons, prendre ce point en compte et nous visons à proposer une approche qui prend en considération la décision humaine ainsi la coupler avec les autres modèles. Ceci nous permettra d'avoir un outil de simulation capable de reproduire les comportements du phénomène de pollution, tout en impliquant aussi le comportement de l'aspect social (activités anthropiques). Notre objectif principal est la modélisation du problème de pollution de l'air comme étant un système socio-environnemental.

# Chapitre 4 Les Systèmes Multi-agents

---

## 1 Introduction :

La modélisation c'est la tâche de trouver une abstraction afin de pouvoir isoler théoriquement un système réel, c'est-à-dire, avoir une représentation formelle (mathématique ou informatique) d'un système physique. Cette représentation sous forme de modèle facilitera l'étude et la communication du système modélisé. Le modèle est ensuite utilisé dans une simulation, qui consiste à inspecter ses comportements sous certaines conditions, ce qui permettra d'avoir plus de connaissances sur son évolution. (Axelrod, 2006) présente la simulation comme étant une troisième façon pour faire des expérimentations : La première est inductive, en cherchant à trouver des patterns cachés dans des données empiriques. La deuxième est déductive, consiste à fixer des axiomes et d'essayer de déduire des théorèmes. La dernière est la simulation, qui consiste à faire des simulations d'un phénomène sous certaines conditions (hypothèses) et générer des données. Ces dernières peuvent être analysées d'une façon inductive.

Les approches classiques de simulation sont basées sur les équations mathématiques (SD et DES). Depuis plus de deux décennies, l'approche basée agent est apparue comme étant une approche très prometteuse pour modéliser, simuler et étudier les phénomènes complexes.

Dans ce chapitre nous allons décrire et discuter quelques définitions des concepts et les propriétés fondamentales de l'approche basée agent. Cette présentation est organisée comme suit : nous commençons dans la première section à discuter la problématique de terminologie entre approches basées agent et système multi-agents. Après, nous exposerons les propriétés d'un agent. Les propriétés ainsi que les composants d'un SMA sont aussi exposés. Les approches qui font partie de la même catégorie des SMA sont présentées ainsi qu'une comparaison montrant les principales différences entre les SMA et ces méthodes. Nous présentons les avantages des SMA qui nous ont motivés à les utiliser, tout en restant vigilants à propos de ses limites exposées aussi dans la même section. Nous exposons brièvement la problématique de validation des modèles basés SMA, et terminerons par une conclusion.

## 2 L'approche basée agent

Au cours de ces dernières décennies, les chercheurs ont développé une terminologie assez riche pour désigner l'approche de modélisation basée agent. Influencés par la littérature de leurs domaines, ils utilisent parfois les termes : Modélisation basée Agent (Agent based modelling), Simulation basée Agent (Agent based Simulation), Système Multi-agents (Multi-Agent system), Simulation multi-agents (Multi-agent Simulation). Tous ces termes désignent des approches qui ont en commun le concept agent, mais qui ne sont pas forcément les mêmes. On peut identifier deux approches composées chacune de sa propre communauté de chercheurs : Les SMA est un sous-domaine de l'IA qui vise à résoudre des problèmes complexes et distribués en utilisant des agents intelligents, vue comme étant une évolution des modèles d'IA classiques tels que les systèmes à base de connaissance, systèmes experts. Les SMA qui utilisent les agents pour développer des modèles des systèmes complexes, dynamiques et intégrant des entités dynamiques et autonomes. Donc, d'une part les chercheurs en SMA sont concernés par la conception des différents types d'agents et l'architectures des SMA ; de l'autre part, les SMA utilisent le concept d'agent comme étant un moyen pour simuler/modéliser des systèmes réels physiques ou sociaux afin d'approfondir leurs connaissances sur leurs comportements globaux et émergents. (Niazi & Hussain 2011) présente une étude scientométrique concernant l'approche agent qui regroupe les deux domaines SMA et MBA. L'étude consiste à analyser avec des métriques scientifiques toutes les publications de 1980-2010 qui utilisent l'approche agent et tous ses sous-domaines. L'étude a permis d'élucider le fait que beaucoup de disciplines non-informatiques sont très présentes et dépassent parfois en nombre les disciplines informatiques. (Wellman 2014) présente une comparaison entre les SMA et MBA, en donnant les points forts que peut bénéficier l'une de l'autre. Pour garder une homogénéité dans les descriptions, nous allons utiliser le terme Système Multi-agents (SMA) pour désigner toutes les approches ou modèles qui font recours au concept d'agent.

D'après (Ferber 1995), au début les chercheurs en intelligence artificielle (IA), tentaient de créer des programmes qui simulent l'intelligence humaine. Dans le but de remplacer et/ou aider l'humain dans ses diverses tâches. Le concept qui a été largement utilisé est celui des Systèmes Experts. Au fur et à mesure et avec le développement des puissances de calcul des machines et l'arrivée du multithreading, l'intelligence artificielle distribuée est apparue, elle vise à créer des programmes qui arrivent à collaborer pour résoudre un problème commun. Le but était de faire coopérer des systèmes experts pour résoudre des problèmes distribués. Un autre sous-domaine de l'intelligence artificielle distribuée est apparu, celui des systèmes multi-agents, qui

visent à modéliser les systèmes physiques et naturels qui sont distribués et font impliquer plusieurs entités dynamiques. Les systèmes multi-agents permettent la désagrégation des systèmes physiques sous forme de plusieurs agents (individus). Ce qui permet une approche de modélisation allant du niveau bas vers le niveau haut (en anglais bottom-up), c'est-à-dire, modéliser l'individu et ses comportements et ses interactions afin d'avoir comme résultat le comportement de tous le système. Dans certains cas, avoir des comportements qu'on ne peut pas voir en modélisant le système dans sa globalité. Des comportements qui n'ont pas été programmés explicitement émergent. Ce qui permet d'approfondir les connaissances sur le système réel modélisé.

Les agents qui forment le système possèdent généralement des composants avec des comportements qui ne sont pas homogènes. Dans certains modèles (tels que les modèles sociaux) ceci est un élément très important, car les êtres humains possèdent des personnalités et des comportements qui sont assez hétérogènes. Ceci rend la modélisation des phénomènes sociaux en utilisant les SMA très adéquate (Axelrod 2006).

### **2.1 Définition du concept Agent**

Un système multi-agents est tout simplement composé d'un ensemble d'agents et un environnement. De ce fait, il est essentiel de commencer par donner une définition du concept Agent. En fait, malgré la maturité de l'approche basée agent, il n'y a pas, jusqu'à présent, un consensus sur la définition exacte du terme agent. Depuis des dizaines d'années le débat pour donner une définition exacte n'est pas arrivé à sa fin. Heureusement, ce débat n'a pas freiné le développement et l'utilisation des approches basées agents. Ces dernières ont montré leur utilité et faisabilité dans plusieurs domaines tels que: Simulation, Imagerie, Trading, contrôle, Ecologie, Economie.

Les définitions les plus répandues décrivent un agent par ses propriétés et de ce qui est capable à le faire. Ce qui rend les définitions vagues, vu que les capacités et les propriétés d'un agent dépendent largement des problèmes modélisés et les objectifs de sa conception. Nous essayerons de commenter chacune de ces définitions. Le tableau 4.1 présente un résumé des définitions exposées.

(Wooldridge & Jennings 1995) définissent un agent comme étant un système informatique situé dans un environnement, capable autonomement de faire des actions dans le but d'atteindre les objectifs de sa conception (ou de son existence). Cette définition permet de bien élucider le fait que parmi les propriétés principales d'un agent c'est l'autonomie. En fait,

L'autonomie est une propriété qui distingue les agents des autres concepts, tel que les objets. (Ferber 1995s) donne une définition plus détaillée. Il définit un agent comme étant une entité physique ou virtuelle :

- Qui est capable d'agir sur son environnement,
- Qui peut communiquer avec d'autres agents,
- Qui est doté par un ensemble de tendances (objectifs individuels, fonction de satisfaction, qu'elle cherche à optimiser).
- Qui possède des ressources propres.
- Qui est capable de percevoir son environnement,
- Dispose d'une représentation (partielle) de son environnement (et éventuellement aucune),
- Possédant des compétences et offrant des services
- Peut éventuellement se reproduire,
- Dont le comportement tend à satisfaire ses objectifs, en tenant en considération des ressources.

Cette définition très répandue, ne mentionne nullement l'autonomie. A l'instar, elle cite un ensemble de propriétés (caractéristiques) dont la première est le fait que l'agent puisse agir sur son environnement, ceci est absurde car un agent qui n'a pas d'effets sur son environnement est considéré inexistant, de même pour un agent qui ne perçoit pas son environnement.

(Russell & Norvig 1995) donnent une définition large qui se base sur la perception et les actions qu'a un agent sur son environnement. Ils définissent un agent comme étant n'importe quelle entité qui aperçoit son environnement via des capteurs et qui agit sur cet environnement via des composants effecteurs. En conséquence et suivant cette description, plusieurs entités du monde réel peuvent selon ces arguments être considérées comme étant des agents. Par exemple, un portail automatique, possédant un capteur pour détecter les personnes désirant entrer dans un bâtiment. A chaque fois que son capteur est activé il ouvre la porte. Le portail est considéré comme étant un agent.

Une autre définition qui n'est pas loin des ces derniers est celle de (Shoham 1993). Il définit un agent comme une entité qui fonctionne en continuité et autonomie dans un environnement dans lequel d'autres processus et agents existent. (Nawana 1996) définit un agent, comme étant un composant matériel ou logiciel qui est capable d'agir avec exactitude dans le but d'accomplir des tâches à la place de son utilisateur.

Définition	Autonomie	Coexiste avec d'autres entités	Environnement	Perception	Action	Objectif
Ferber 1995	Non	Non	Oui	Oui	Oui	Oui
Wooldridge & Jennings 1995	Oui	Non	Oui	Non	Oui	Oui
Russell & Norvig	Non	Non	Oui	Oui	Oui	Non
Shoham	Oui	Des processus et des agents	Oui	Non	Non	Non
Nwana	Non	Utilisateur	Non	Non	Oui	Oui

**Tableau 4.1 : résumé des propriétés utilisées pour définir un agent.**

En s'inspirant de cette discussion, on peut résumer ce qui a été présenté, en proposant la définition suivante : Un agent est toute entité qui existe dans un environnement, et via un ensemble de mécanisme de perception arrive à avoir une représentation (parfois partielle) de son environnement y compris les autres agents, et à travers ses actions, elle peut d'une manière autonome changer l'état de l'environnement ainsi que l'état des autres agents coexistant dans le même environnement, ces changements lui permettent d'accomplir les objectifs de sa conception.

## **2.3 Les systèmes multi-agents**

Après avoir défini et présenté les propriétés d'un agent, on va maintenant s'intéresser aux SMA, dans cette section on présente la définition d'un système multi-agents ainsi que ses composants et propriétés.

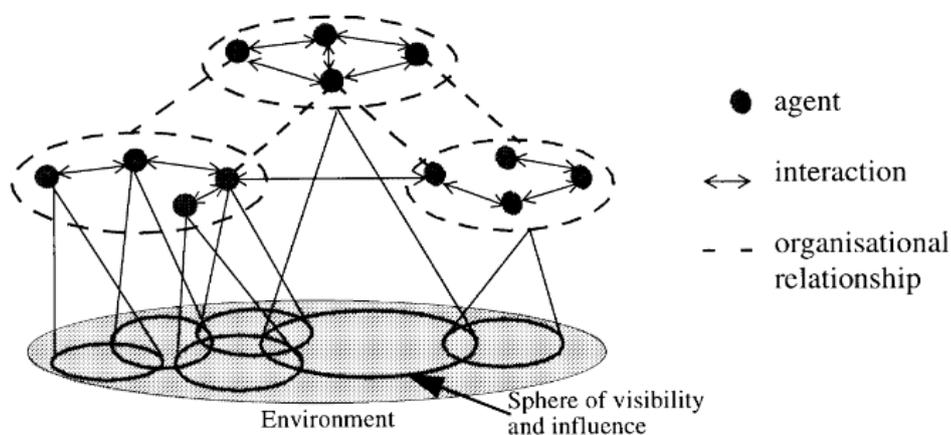
### **2.3.1 Définition:**

(Ferber, 1998) donne une définition assez précise des SMA :Un système multi-agents (SMA) est un système composé de :

- Un environnement  $E$ , qui est un espace doté d'une métrique,
- Un ensemble  $O$  d'objets situés dans  $E$ ,
- Un ensemble d'Agent  $A$ , qui sont eux même des objets actifs ( $A$  inclus dans  $O$ ).
- Un ensemble de relation  $R$  qui unissent les objets entre eux.
- Un ensemble d'opération  $Op$  permettant aux agents de  $A$  de percevoir, produire, manipuler les objets de  $O$ .

- Des opérateurs chargés de représenter l'application ces opérations et la réaction du monde (environnement) sur cette tentative de modification nommée lois de l'univers.

Une définition plus simple est donnée par (Macal and North 2010). Qui définissent un (SMA) comme étant un ensemble d'agents avec leurs comportements et attributs, un ensemble de relations entre les agents, un environnement. Les agents ont des interactions avec leur environnement ainsi qu'entre eux mêmes. Principalement, un SMA est composé d'un ensemble d'agents situés dans un environnement. (Jennings 2000). La figure 4.1 donne une illustration d'un système multi-agents.



**Figure 4.1 : Représentation d'un SMA, (source Jennings 2000).**

Un SMA est conçu pour des raisons diverses, nous pouvons citer : La modélisation d'un système complexe, un système adaptatif. Un système complexe est un système composé de plusieurs composants autonome et dynamique. En modélisant ses composants par des agents et en définissant les interactions entre ses composants, on peut avoir un modèle qui permet d'investiguer les comportements globaux de ce système. Pour les systèmes adaptatifs, la modélisation consiste à représenter le système sous forme d'un ensemble d'agents qui s'adaptent pour réaliser ou résister à une situation donnée.

### 2.3.2 L'environnement d'un SMA

L'environnement est l'espace dans lequel les agents existent, sa représentation dépend de la nature du système modélisé. Dans certain cas, une représentation abstraite est utilisée. Elle consiste en un espace géométrique Euclidien à deux ou trois démentions. La position donc de chaque agent est sauvegardée parmi ses attributs, par exemple elle peut être une coordonné géométrique. La notion de proximité dans ce cas est la distance entre les agents. Les agents

peuvent être localisés dans un réseau social (Alam et Geller 2012). Les agents dans ce cas possèdent des voisins avec lesquels ils ont des connexions (sociales), ces connexions peuvent avoir des diverses explications telles que : une relation de connaissance, relation commercial (client-fournisseur), ou juste une relation de voisinage. Dans le cas d'une représentation sous forme de grille, un agent peut se situer dans une cellule et donc ses voisins sont les agents situés dans les cellules adjacentes. Certains SMA sont couplés avec un module GIS (Système d'Information Géographique) et les agents sont positionnés sur des carte graphique réelle (Heppenstall et al. 2011). Le choix de la représentation de l'environnement et du mécanisme de voisinage a parfois une influence majeure sur les comportements globaux du système (Conrad Power 2009). Ceci est dû au fait que les comportements des agents peuvent être influencés par leurs voisins.

### **2.4 Les SMA et les phénomènes à comportements émergents**

L'émergence est définie par (Epstein et Axtell 1999) comme étant l'organisation (des comportements) sous forme de patrons macroscopiques reconnaissables. Durant la simulation ces patrons peuvent apparaître selon une loi, une condition ou selon une périodicité. Connaître ces conditions et lois est le but fondamental de la modélisation et de la simulation.

Les systèmes multi-agents ont la propriété fondamentale de capturer les comportements globaux d'un système complexe. Que se soit pour modéliser les comportements des automobilistes sur les routes, ou bien les flux des patients dans un service médical, ou n'importe quel système physique qui fait impliquer des interactions entre plusieurs entités autonomes. Les SMA permettent de faire apparaître des comportements dits émergents. Ces derniers apparaissent sans avoir été préprogrammés explicitement. Les interactions inter-automobilistes ainsi que le comportement que suit chacun d'entre eux, fait apparaître des patrons dans la circulation, et aussi des embouteillages. Ce genre de phénomènes émerge à partir des comportements qui ont lieu niveau micro. Pour cela, les SMA sont considérés comme étant une approche Bottom-up, c'est-à-dire modéliser au niveau bas les individus et leurs interactions afin de voir les résultats au niveau haut. Les approches classiques telles que la régression linéaire et les systèmes dynamiques ne peuvent pas capturer les phénomènes émergents (Bonabeau 2002a). Car ces approches sont des approches dites Up-Bottom, qui sont basées sur des agrégations niveau haut et ils ne peuvent pas capturer les phénomènes émergents à partir des interactions qui ont lieu au niveau bas.

Un exemple présenté par (Bonabeau 2002b). Dans lequel il évoque l'histoire de l'administration de NASDAQ, cette dernière voulait modifier la régulation afin de rendre plus précise les ventes dans la bourse, mais ne savent pas les effets d'une telle décision sur le marché.

De ce fait un système basé agent a été développé, et la simulation a montré qu'une catastrophe dans les ventes aura lieu si la décision est appliquée. Les comportements des acteurs dans la bourse ont été capturés grâce à un SMA dans lequel les agents représentent les différents acteurs dans ces opérations. D'autres exemples sont évoqués par l'auteur qui montrent l'efficacité des modèles SMA, nous pouvons citer: l'augmentation du nombre de client dans un magasin fait chuter les ventes d'un produit ; le changement dans la stratégie d'encouragement des employés peut faire chuter la productivité. Ceci montre que les modèles SMA ne permettent pas seulement de capturer les phénomènes émergents mais aussi de les prédire. Les SMA donnent aux modélisateurs la possibilité de tester l'évolution dans le temps et sous certaines conditions du système modélisé. Un exemple est le modèle de mouvement des foules lors de l'évacuation dans un immeuble. Ce genre de modèle permet d'observer un comportement émergent du mouvement de la foule après avoir mis un poteau devant la porte le nombre des blessés a été réduit. Ce genre de conclusion est très difficile à obtenir en utilisant les modèles classiques.

## **2.5. La coopération et le mécanisme de prise de décision.**

### **2.5.1 Interaction et organisation**

Afin de pouvoir résoudre le problème pour lequel ils ont été conçus, les agents doivent agir ensemble et coordonner leurs actions. De ce fait, ils ont besoin d'un mécanisme d'interaction et d'une structure d'organisation. (Ferber 1995) définit les interactions comme étant toute mise en relation de deux ou plusieurs agents via le biais de leurs actions réciproques. Les actions peuvent être physiques (robot sur un autre robot) ou bien échange d'information. La définition d'un mécanisme de coopération d'interaction est un point central pour un système multi-agents. Selon l'objectif de la simulation les agents obéissent alors au mécanisme de coopération définis (collaboration, compétition, conflit, négociation, etc.). Le concepteur peut aussi définir une organisation selon laquelle les agents vont opérer. L'organisation dépend aussi des objectifs de la simulation et la nature du système modélisé. Ils existent plusieurs types d'organisation qui peuvent être adoptés :

- Organisation Hiérarchique dans laquelle les agents sont organisés selon une hiérarchie à plusieurs niveaux, les agents alors possèdent un certain grade ou spécialisation et coopèrent pour atteindre leurs buts communs.

- Organisation en Groupe : Les agents dans ce cas forment des groupes, et chaque groupe peut se spécialiser dans la réalisation d'une tâche ou pour atteindre un but.
- Holonique : Ce genre d'organisation peut avoir plusieurs niveaux de détail, le SMA lui-même peut émerger à être considéré comme étant un seul agent (Holon). Par exemple en modélisant un système biologique, on peut considérer chaque partie comme étant un SMA, lui-même composé de plusieurs parties qui sont elles mêmes des SMA ou des agents. Cette approche récursive permet de modéliser des systèmes qui possèdent une nature récursive (Gerber et al. 2011).
- Organisation dynamique à coalition : Dans ce genre d'organisation, les agents forment des coalitions qui ont des buts en commun, ces coalitions peuvent être dynamiques et s'adaptent avec le temps.

### 2.5.2 La prise de la décision

L'agent possède un ensemble d'actions, selon son état et ses perceptions de l'environnement ainsi que ses objectifs, il doit faire un choix concernant l'action à appliquer. Pour faire ce choix l'agent a besoin d'un mécanisme de prise de décision. Ce dernier peut être vu comme étant une fonction qui assure un mappage entre l'entrée constituée d'un ensemble de paramètres et donne comme résultat une décision. Par exemple, un agent boursier (Trading Agent) a trois actions possibles (Achter, Vendre ou bien Négociier) selon les prix et l'état des indices boursiers il doit choisir une action qui lui permet de maximiser ses gains. La théorie de prise de décision est tout un domaine à part. Elle contient un nombre important d'approches. (Georgeff et al. 1999) présente une approche nommée BDI, elle est inspirée de la psychologie humaine. Elle se base sur la logique BDI (traduction Croyance, Désires et Intention) pour concevoir des agents qui ont le maximum possible des comportements rationaux. Dans d'autre approche ils utilisent des modèles inspirés de l'IA, tels que les systèmes à logique floue, les réseaux de neurones artificiels, les systèmes à base de cas, l'algorithme de Q-learning. Ou bien des modèles mathématique : Stochastiques, les processus de Markov, la théorie des jeux. Parfois des modèles hybrides sont utilisés qui combinent deux approches différentes.

## 2.6 Les approches similaires aux SMA

Les SMA font partie de la catégorie des approches de modélisation dites basées individus. C'est-à-dire, des approches qui vont du bas vers le haut. Parmi ces approches, on en trouve d'autres qui font partie de cette catégorie tels que les automates cellulaires et la micro-

simulation. Dans cette partie nous allons en décrire et essayer de présenter ce qui les différencie des SMA.

### 2.6.1 Les automates cellulaires

Le modèle des AC (CA en anglais) est un modèle informatique très simple qui permet de modéliser des phénomènes émergents assez complexes. Ils ont été découverts par John Von Neuman (1966) lors de ses travaux sur les systèmes auto-reproductifs. Leur principe consiste en un tableau, ou même une grille, contenant plusieurs cellules (le nombre de cellule est sensé, théoriquement, être illimité). Chacune d'elle peut avoir un état (parmi un ensemble limité d'états possibles) influencé par les états des cellules voisines. Les lois locales très simple parfois, qui gèrent l'état des cellules permettent de faire apparaître des phénomènes émergents jugés complexes. Ce modèle est très utilisé pour la modélisation des systèmes complexes, telle que l'expansion et l'évolution urbaine (Iltaanen, 2012). (Walfrom, 2002), décrit en détails l'utilisation des AC dans des divers domaines : cryptographie, science sociale et même elles sont présentées comme étant une approche pour concevoir des processeurs.

Les SMA et les automates cellulaires ont un point commun ; qui est la modélisation microscopique afin de prévoir les comportements macroscopiques. Les cellules changent d'état selon les règles qu'elles appliquent au niveau local (niveau individu) et cela donne à la fin l'émergence d'un état global beaucoup plus complexe. Les approches sont différentes, la pierre angulaire des SMA c'est l'agent comme entité indépendante et ses bornes sont bien définies. Ils possèdent une autonomie et aussi une représentation de son environnement. Alors que dans les cas des AC, la cellule est le niveau de granularité le plus fin, elle ne possède pas d'autonomie et elle applique un ensemble de règles localement et ne possède pas de représentation de son environnement ni d'un état mental ou un mécanisme d'apprentissage. Une autre différence réside dans le fait que les cellules ne peuvent pas changer de position et sont considérées comme étant des entités statiques. Alors que les agents dans certains cas peuvent être mobiles et change de position pour accomplir leurs tâches et atteindre leurs objectifs.

### 2.6.2 La micro-simulation:

Est une technique analytique pour la simulation des phénomènes physiques à l'échelle microscopique (Entreprise, ménage, individu dans certains cas). Elle consiste à identifier des unités (individus) microscopiques avec des micro-informations (informations sur l'unité) et voir l'évolution dans le temps de ces informations selon les changements dans la politique globale. Elle a été utilisée intensivement dans l'économie pour voir les effets des changements dans la

politique économique sur les ménages (Ballas et al. 2005). Le point commun qu'elle a avec les SMA est la modélisation basée individu. La micro-simulation permet de voir les effets des changements globaux sur l'individu mais pas l'inverse, alors que les SMA peuvent opérer dans les deux directions. Les MS sont beaucoup plus orientées vers la modélisation des structures, alors que les SMA sont principalement focalisés sur la modélisation des comportements. (Birkin et Wu. 2012) présente l'hybridation des deux méthodes comme étant une solution aux limitations de chacune, et l'appliquent dans des domaines de la modélisation d'utilisation des terres et la simulation de l'expansion urbaine (Wu et Birkin 2008).

## **2.7 Les Avantages des SMA**

(Bonabeau 2002b) présente une liste des avantages des SMA. Dont la principale avantage est leur capacité à capturer les phénomènes émergents, ils offrent une description naturelle d'un système réel, ils sont flexibles. Nous allons, dans la section suivante, détailler ces points. La majorité des avantages décrits dans cette section sont tirés des études dans lesquelles les auteurs tendent toujours à comparer les SMA avec d'autres méthodes classiques de simulation (Système Dynamique, la Simulation à Eventement Discret, etc). Dans (Maidstone 2012) une comparaison plus détaillée peut être trouvée ainsi que dans (Borshchev et Filippov 2004), certains vont jusqu'à prouver que tous les modèles SD sont un sous-ensemble des SMA ; (l'agencité des Système Dynamique (Macal 2010)).

Capable de capturer les phénomènes émergents : En fait, le tout est loin d'être la somme de ses composants, ceci est due aux interactions qui existent entre les composants (Bonabeau 2002b). Ceci est très convaincant, car il est clair qu'il existe des comportements au niveau global qui sont difficiles, voir même impossible, à déduire en observant les composants séparément et en ignorant les interactions inter-composants. Les SMA sont appropriés pour capturer les phénomènes émergents lorsque :

- Les comportements des individus sont non-linéaires et il est difficile d'utiliser les systèmes basés équation différentiel (System Dynamic et DES) pour les modéliser (Siebers 2010).
- L'effet de bruit et de l'incertitude permet de rendre les modèles plus réalistes. Dans le cas des systèmes basés équations différentiels un terme de bruit est utilisé. Pour les SMA le bruit est appliqué au bon endroit, à savoir, appliquer au niveau des comportements de l'agent. De ce fait, les SMA permettent de concevoir des modèles plus proches de la réalité.

- Les individus ont des comportements hétérogènes, c'est-à-dire, que les individus présentent des personnalités différentes. Les approches basées agrégation ne permettent pas de modéliser cette propriété très répandue dans les systèmes réels.
- Les individus présentent des comportements liés à la mémoire, c'est-à-dire, la mémorisation de ses actions précédentes et leurs résultats.
- Les individus possèdent des propriétés d'adaptation et d'apprentissage via leurs expériences.
- Les SMA permettent une représentation selon plusieurs niveaux du temps : niveau interaction, niveau individu par exemple une journée et aussi le niveau loin terme allant jusqu'à plusieurs années. Cette propriété est très importante, car elle permet d'étudier l'évolution du phénomène modélisé selon plusieurs niveaux temporels.
- Les SMA offrent une possibilité pour tester des théories sociales pour lesquelles, les méthodes mathématiques sont trop compliquées à appliquer(Axelrod 1997).
- L'approche de conception des SMA rend les modèles généralement très flexibles et permet des modifications et extensions très facile.
- Les modèles SMA sont dynamiques car ils sont basés sur l'activité, ceci est très utile car ils peuvent capturer la dynamique du phénomène modélisé.
- Les SMA sont puissants pour modéliser les systèmes qui présentent la propriété de sentier de dépendance (path dependence en anglais). C'est-à-dire, les décisions prises dans le passé continuent à influencer sur les décisions futures. Donc une décision n'as pas seulement un effet immédiat mais aussi un autre effet à long terme. Beaucoup de systèmes réels présentent ce genre de propriété.

## **2.8 Les inconvénients et les limitations des SMA**

Les SMA sont utilisés de plus en plus dans des domaines divers. L'approche est considérée comme étant assez mature et ses modèles sont très répandus. Malgré cela, les SMA souffrent de quelques inconvénients et limitations qui sont catégorisé en trois axes : (1) Le rapport difficulté de modélisation en SMA vis-à-vis son utilité ;(2) La disponibilité des données détaillées ; (3) La reproductibilité. Dans la section suivante on va détailler ces limitations.

Les SMA comme toutes les approches de modélisation, souffrent du problème de généralisation. Cela veut dire qu'un modèle conçu pour représenter un phénomène ne peut être

réutilisé (ou généralisé) pour étudier un autre. (Coulelis 2002) pense que les SMA ont alourdis la tâche des modélisateurs, il pense que les SMA ajoutent à la tâche de modélisation des phénomènes complexes, la difficulté de modéliser la prise de décision et les interactions des entités dynamiques complexes. Elle questionne l'utilité et le gain par rapport à l'effort et les coûts dépensés en construisant de tels modèles. (O'Sullivan et al 2012) présente une étude plus détaillée de cette question. Et il va plus loin jusqu'à dire, qu'en utilisant les SMA, on représente un système physique très difficile, par un modèles SMA très difficile à comprendre aussi. Le problème de choix de niveau d'abstraction est posé, en choisissant trop de détails le modèle devient trop compliqué et trop de contraintes sont alors imposées ; de même le choix d'un niveau simple d'abstraction le degré de représentabilité (par rapport au système réel modélisé) du modèle est alors remis en cause.

L'approche SMA nécessite l'existence d'une base de données empirique. En fait, pour être calibrés et validés les modèles ont besoins de grosses quantités de données avec un niveau de détail important. Ceci afin de mieux capter le phénomène modélisé. Dans certain cas ces données ne sont pas disponibles. Un modèle conçu avec un jeu de données moins significatif produira des résultats loin de la réalité et son utilité est à débattre.

L'un des aspects fondamentaux des méthodes scientifiques c'est la reproductibilité, à savoir, sous les mêmes conditions et en utilisant les étapes décrites dans une expérimentation, on doit avoir les mêmes résultats. D'après Karl Popper, le critère de reproductibilité est ce qui distingue le caractère scientifique d'une étude. Les modèles SMA sont généralement très difficiles à reproduire. Cela est dû à la sensibilité de ces modèles aux conditions initiales et l'absence d'un cadre générale sous lequel les modèles peuvent être décrits et communiqués. Le Framework ODD (Objective-Purpose and Details Design) a été proposé par (Grimm et al. 2006) et révisé dans (Grimm et al. 2010). L'ODD est proposé pour remédier à la question de reproductibilité des modèles. Il permet de mieux communiquer les modèles SMA et d'améliorer leur reproductibilité. Le framework ODD offre un cadre bien précis et guide les modélisateurs afin de décrire leurs modèles via un ensemble d'étapes bien précis. Dans (Müller et al. 2013) une amélioration du framework ODD est proposée, elle vise à ajouter des sections pour mieux décrire les mécanismes de prise de décision des agents.

### **2.9 Validation des modèles SMA**

Un modèle SMA est conçu en utilisant un jeu de données, il doit donc être bien calibré, c'est-à-dire que ses sorties sont de même nature que les données utilisées pendant sa conception.

Ensuite vient l'étape de vérification. Elle consiste à tester si le modèle est correctement conçu. Une fois le concepteur est assuré que le modèle est correct. Il passe après à l'étape de validation, qui consiste à vérifier que le modèle donne les comportements exacts par rapport aux données empiriques (Parker et al. 2003).

(Moss 2008) présente une revue des méthodes pour valider les modèles SMA dans de diverses domaines (économie, social et système complexe). (Marks 2007) présente une approche formelle pour la validation des modèles basés agents. Ils présentent cinq cas possibles, supposant que  $\mathbf{M}$  est l'ensemble des comportements que présente le modèle,  $\mathbf{S}$  est l'ensemble des comportements que le phénomène réel présente. On peut lister les cinq cas possible suivants :

- (a) Si l'intersection entre  $\mathbf{S}$  et  $\mathbf{M}$  est vide alors le modèle est jugé inutile. Car il ne présente aucun comportement réel.
- (b) Si l'intersection entre  $\mathbf{S}$  et  $\mathbf{M}$  n'est pas vide, le modèle est alors utile selon un certain degré. Le modèle présente quelques comportements réel mais pas tous.
- (c) Si  $\mathbf{M}$  est inclus dans  $\mathbf{S}$  alors le modèle est incomplet, car tous ses comportements conviennent avec la réalité mais il ne présente pas tous les comportements observés.
- (d) Si  $\mathbf{S}$  est inclus dans  $\mathbf{M}$ , le modèle est jugé inexact, car il montre des comportements qui n'ont pas été observés dans la réalité.
- (e) Si  $\mathbf{S}$  est équivalent exact de  $\mathbf{M}$ , alors le modèle est complet et exact.

L'auteur présente, ensuite une métrique en combinant les deux critères de complétude et exactitude, comme suit : l'exactitude du modèle est nommée  $\alpha$ , définie dans (1)

$$\alpha = 1 \frac{m(Q)}{m(M)} \quad (1)$$

La mesure d'incomplétude est donnée dans (2)

$$\gamma = 1 \frac{m(Q)}{m(M)} \quad (2)$$

Le degré de validation est alors donné dans (3)

$$V = v(1 - \alpha) + (1 - v)(1 - \gamma) \quad (3)$$

Où  $v$  est le poids accordé pour favoriser l'une des métriques par rapport à l'autre est qui doit être entre 0 et 1.

## Conclusion

Modéliser c'est représenter un phénomène réel en utilisant un modèle abstrait facile à manipuler et à étudier. Ils existent plusieurs approches pour faire cette modélisation. Dans ce chapitre nous avons présenté une approche qui ne cesse pas d'être appliquée dans beaucoup de domaines, à savoir les SMA.

Pour modéliser des phénomènes composés de plusieurs entités dynamiques, les SMA constituent une approche puissante. Elle permet aux modélisateurs de concevoir des modèles à partir du niveau le plus bas, à savoir, le niveau interaction inter-entités, et de voir après, l'évolution du comportement global du système. Cette possibilité de pouvoir lier les paramètres macros et micros ne peut pas être faite avec les approches classiques. Les modèles sont très intuitifs et permettent de capter les comportements émergents du système modélisé. Les modèles SMA sont basés sur l'aspect actif, ils offrent la possibilité à la fois d'avoir une représentation du phénomène modélisé et aussi de voir son évolution. Cette dernière, est une propriété très importante pour avoir des prédictions sur l'état du système sous certaines conditions expérimentales. Ceci est l'objectif de toute simulation.

Les modélisateurs ne doivent pas tomber dans la loi de l'instrument (Maslow 1966) « si tout ce que vous avez est un marteau, tout ressemblera à un clou »; et doivent être ouverts à toutes les approches de simulation afin de tirer profits des avantages qu'offrent chacune d'entre elles. L'approche SMA ne doit pas être vue comme un remplaçant des autres approches de simulation classiques, dans certains cas une hybridation permet de remédier aux lacunes des deux.

Dans ce chapitre nous avons essayé d'expliquer tous les concepts qui sont liés aux SMA, certes il en reste d'autres que nous n'avons pas exposés. Nous concluons par exprimer notre motivation pour la modélisation SMA, tout en restant vigilant concernant leurs limitations.

# Chapitre 5 : Utilisation des systèmes multi-agents pour la modélisation des problèmes socio-environnementaux

---

## 1 Introduction

Ce chapitre présente une revue des modèles SMA utilisés pour simuler des systèmes socio-environnementaux. L'interaction entre le système social et le système environnemental est un point clé dans la compréhension des problèmes liés à la pollution urbaine. Les SMA permettent de modéliser cette interaction avec un niveau de détails très haut, allant du niveau individu et leurs comportements jusqu'aux dynamiques globales.

Les modèles SMA de la pollution de l'environnement sont des outils précieux. Ils peuvent aider les décideurs à mettre en place des politiques de gestion de l'environnement afin de préserver l'écosystème et assurer la santé publique. Ils permettent aux gestionnaires environnementaux de prédire l'impact de leurs décisions sur l'environnement. La pollution est principalement causée par des activités anthropiques, par conséquent, la modélisation et la simulation de ces activités est très important.

Ce chapitre a comme objectif la présentation d'une revue des études récentes (2009-2015) qui utilisent les Systèmes Multi-Agents, pour modéliser les problèmes socio-environnementaux spécifiquement ceux qui sont liés à la pollution, avec une considération aux modèles qui traitent la pollution atmosphérique.

La sélection des études a été effectuée à l'aide des critères suivants: Agent-based modelling of pollution; Agent-based simulation of pollution; Multi-agent modelling of pollution. En utilisant ces mots clés, 21 études ont été sélectionnées dont huit traitent un sujet lié à la pollution de l'air.

La comparaison entre les travaux n'est pas trop évidente, vu que chaque modèle traite un phénomène différent et utilise un mécanisme de prise de décision différent. Pour cela nous

commençons par la description des éléments selon lesquels les modèles sont présentés. Nous présentons aussi une classification des modèles selon les trois axes : écologie, social et économie.

Ce chapitre se termine par une conclusion situant notre approche par rapport à ces travaux connexes.

## 2 Les critères de comparaison des études

Nous avons essayé d'analyser les études selon les éléments décrits ci-dessous. Il est à noter que dans certains articles, ces éléments ne sont pas toujours assez détaillés, dans certains cas complètement ignorés. Le Tableau 1 ci-dessous récapitule les travaux discutés selon ces éléments.

- (1) Le but de la simulation : Ceci représente les objectifs de la simulation. C'est les questions auxquelles nous cherchons à trouver des réponses en utilisant la simulation. Elle peut être : l'aide à la décision, l'apprentissage, l'expérimentation, la prédiction, sensibilisation. Chaque modèle doit posséder un ou plusieurs buts qui justifient son utilité.
- (2) Les variables étudiées : en fonction des objectifs de la simulation, nous pouvons définir la ou les variables que nous cherchons à prédire ou à estimer. Par exemple : la concentration de la pollution de l'air, de la pollution de l'eau, de la biodiversité d'une espèce, l'utilisation des terres, l'expansion urbaine, l'émission dans un réseau routiers, les changements dans les paramètres climatiques. Parfois, ces variables sont étudiées en parallèle avec d'autres paramètres qui ne sont pas explicitement liés à l'environnement. Tels les paramètres économiques ou sociaux (ex. : le taux de la pauvreté, les revenus des ménages, la démographie). Et nous cherchons avec la simulation à trouver les corrélations possibles entre ces variables. Une approche simpliste, peut décrire un modèle comme une fonction qui permet de lier une variable aux paramètres de la simulation.
- (3) La représentation de l'espace : les problèmes socio-environnementaux sont toujours spatialement distribués. Les simulateurs peuvent, dans certains cas, inclure l'utilisation des données géographiques réelles. Un système de simulation peut utiliser un module SIG (Système d'Information Géographique) ou exploiter un réseau GPS, ou même des images satellitaires. Parfois l'environnement est représenté par un espace euclidien de deux ou trois dimensions, ce genre de représentation sont moins réaliste par rapport à ceux qui utilisent une représentation réelle.
- (4) La représentation du temps : Il est très important de préciser l'échelle temporelle de la simulation. Car tous les phénomènes socio-environnementaux sont liés au temps. Dans la plupart des cas, une représentation non-continue est utilisée. Il faut préciser aussi l'horizon de la simulation, elle peut aller de quelques heures jusqu'à plusieurs années.
- (5) Les données: Certains systèmes utilisent des données empiriques réelles recueillies auprès: des réseaux de surveillance, d'enquêtes ou d'une base de données existante. D'autres systèmes utilisent des données aléatoires, générées avec des algorithmes. L'utilisation des

données empiriques est toujours favorisée car il permet de donner plus de crédibilités aux résultats de la simulation.

- (6) La prise de décision : C'est le modèle utilisé par les agents pour pouvoir (ici, les agents représentent généralement des êtres humains) choisir une action à effectuer parmi toutes les actions possibles. Par exemple: CBR (Case Based Reasoning), ANN (Artificial Neural Network), Fuzzy-Logic et modèles couplés ou hybrides, modèles BDI. Il est à noter que lorsque le modèle n'inclut pas l'aspect social, les agents ne représentent pas dans ce cas des êtres humains.
- (7) L'implémentation: un simulateur est en fin de compte est un logiciel qui a besoin d'un ensemble de ressources (matérielles et logicielles) pour être implémenté. Ils existent plusieurs plateformes et technologies utilisées pour développer les systèmes de simulation. Il est à noter que parmi les systèmes décrits dans ces études, y en a ceux qui n'arrivent pas au stade d'implémentation. Le modèle dans ces cas là, n'est utilisé que pour communiquer et décrire le problème sans arriver à faire tout le cycle de la simulation.
- (8) Les utilisateurs : Le simulateur est conçu dans le but d'être utilisé. Il est préférable de préciser la population d'utilisateurs à qui s'adresse la simulation. Par exemple : les autorités de gestion de l'environnement, les chercheurs, les étudiants, le grand publique, etc.

Référence	Nom	Utilisateurs	Objectifs	Variable étudiées	Représentation de l'environnement et le temps	Données et validation	Implémentation	Modèle de prise de décision
(Athanasiadis & Mitkas 2009)	O <sub>3</sub> RAA	Agence gouvernementale et décideurs	pollution de l'Air	Ozone concentration	N/A	Données issues de réseaux de surveillance de la qualité de l'air	JADE, WEKA, Protégé2000, PMML2, JESS	RNA, système expert, RBC, Arbre de décision
	ABACUS	Agence gouvernementale	Climat	Climat	N/A	Données climatiques		
(Conruyt & al. 2009)	BIOMAS	Planification agricole	La dynamique de l'agriculture et l'écosystème	Biodiversité	SIG,	N/A	GEAMAS-NG, ArcGIS	Modèle mathématique
(Corchado & al. 2009)	OSM	NA	Pollution des eaux	Expansion et propagation de fuites pétrolières	Espace 2D	Images satellitaires Données climatiques	Web services,	Case based reasoning, ANN, BDI.
(Pereira & al.2009)	EcoSimNet	NA	l'écosystème and l'aquaculture	Biodiversité	1120 cellules (32 x 35), chaque cellule représente 500 m. Le pas de simulation est de 30 secondes; Horizon 1.5 ans.	Paramètres économiques		Simulated Annealing
(Borri & Camarada 2011)	ECROUB	NA	Pollution de l'air	La qualité de l'air	GIS	Données climatiques	GIS	Mathematical model
(Papaleonidas & Iliadis 2012)		NA	Pollution de l'air	La concentration de l'Ozone	Les pas de simulation 6 heures en avance	Données météorologiques	Jade, MySQL	ANN, Fuzzy logic
(Ahat et al. 2009)		NA	Pollution de l'air	La concentration pollution d'air	Grille de deux dimensions,	Aléatoire pas de données réelles	Repast Symphony	
(Sokolova Fernández-Caballero, 2012)	DeciMas	Décideurs en santé et écologie	Corrélation entre pollution et maladies	Corrélation entre polluant problème de santé	N/A	30 paramètres : données issues des bases de données de système de santé	Ontology, JACK, Prometheus	Regression, ANN
(Natalja & al. 2011)		NA	Pollution de l'air	Pollution de l'air et consommation de fuel	Espace contenue, Le pas de simulation est de 23 minutes	Données sur les véhicules et le trafic routiers		Modèle mathématique

Référence	Nom	Utilisateurs	Objectifs	Variable étudiées	Représentation de l'environnement et le temps	Données	Implémentation	Modèle de prise de décision
(Weiqlan & al. 2012)		NA	Water pollution	Pollution	3D et SIG. Time step one second.	Données mesurées d'un incident de la rivière <i>Weihe</i>	Framework Globus, grille de calcul.	
(Brady & al. 2012)	Agropolis	Politique agricole	Utilisation des terres agricoles et la biodiversité	Biodiversité	Grille de N X M cellules, Temps :, horizon 25 ans., pas de simulation 1 année	paramètres Economiques et politique agricoles		
(Qian-jin & al. 2010)		NA	Pollution des eaux	Traitement des déchets et la pollution des eaux	Grille of N x M cellule			Algorithmes génétiques
(Smajgl & al. 2011)	SimPaSI	NA	Economie / Ecologie	Influence des prix des carburants sur les forêts	N/A			
(Schreinemachers & berger 2011)	MP-MAS	NA	Agriculture/ environnement		N/A	Des données empiriques		Modèle mathématique
(Pedell & Sterling 2011)		NA	Water management	Water-use	N/A			N/A
(Watkins & al. 2013)		NA	Restauration écologie	Restauration des forêts	N/A	Survey		Modèle mathématique
(Hülsmann & al. 2011)	MATSim	Gestion des routes et qualité de l'air et Transport.	Détecter les paternes entre la Pollution de l'air et la circulation routière et le type de fuel utilisé.	La pollution de l'air : NO <sub>2</sub> et PM	Carte 2D, et coordonner GPS	Les données de la circulation routière de la ville de Munich, la base de données HBEFA	MATSim, HBEFA, PHEM	Algorithme d'apprentissage, pour améliorer les trajets des agents.
(Hülsmann & al. 2013)	MATSim	Politique de tarification routières et pollution de l'air	Corrélation entre la politique de tarification routière et émission des PM et NO <sub>2</sub>	La pollution de l'air : NO <sub>2</sub> et PM	Carte 2D, et coordonner GPS	Les données de la circulation routière de la ville de Munich, la base de données HBEFA	MATSim	Algorithme d'apprentissage, pour améliorer les trajets des agents

Référence	Nom	Utilisateurs	Objectifs	Variable étudiées	Représentation de l'environnement et le temps	Données	Implémentation	Modèle de prise de décision
(Oprea & al.2011)	AirQMAS	Surveillance de la qualité de l'air	Surveiller l'état de la pollution de l'air			Réseau de contrôle de la qualité de l'air.	Plateforme Zeus AIR_POLLTION_Onto	Système à base de règles, Arbre de décision
Touza & al. 2013		Gestion des nombre de cerfs	Biodiversité	Biodiversité et variables économiques	Grille de 2D	Aléatoires	N/A	Théorie des jeux
LKHOO & Meng 2010		Transport urbain et pollution de l'air	Trouver la méthode de tarification qui permet de réduire la pollution de l'air résultante de l'activité routière	Qualité de l'air généralement	Graphe de 3D	Statistiques de circulation de la ville de New Town , en Singapore	C++	N/A

**Tableau 5.1 : les travaux utilisant les SMA pour la modélisation des problèmes socio-environnementaux.**

### 3 La modélisation des problèmes socio-environnementaux basée multi-agent: état de l'art

Les SMA ont émergé comme une approche prometteuse pour modéliser les systèmes socio-environnementaux. Leur capacité à modéliser les comportements des individus offre une meilleure possibilité pour modéliser l'interaction inter-individu et aussi entre groupe d'individus et leur environnement.

Au cours des dernières années, de nombreuses approches ont été utilisées pour modéliser les problèmes socio-environnementaux. (Kelly & al. 2013) présente une étude comparative de cinq approches de modélisation: Systèmes Dynamique, SMA, Réseaux Bayésiens, Système à base de connaissances. Pour choisir le modèle approprié, les auteurs suggèrent trois considérations: Le but du modèle (qui peut être: la prédiction, la prévision, la gestion et la prise de décision, l'apprentissage); les données utilisées; la représentation du temps ; la représentation de l'espace et entités incluses dans la simulation.

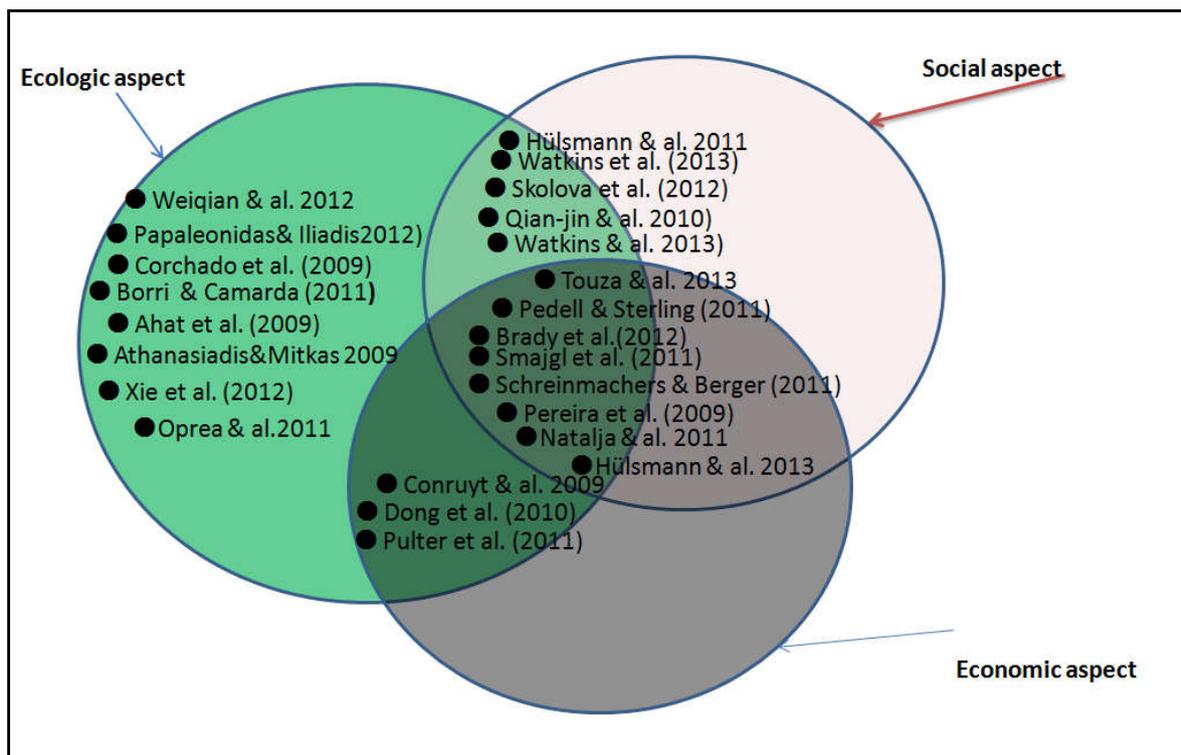


Figure 5.1 : Les modèles classés selon les trois axes : écologie, social et économique.

(Aulinas et al. 2009) présente une revue des travaux qui utilisent les modèles SMA utilisés dans la simulation des problèmes socio-environnementaux. Les SMA sont une approche de modélisation basée sur le principe fondamental de représenter les interactions; dans notre cas,

entre les dynamiques socio-économiques et écologiques. On commence, d'une part, par la modélisation des acteurs au niveau individuel et / ou niveau global (des groupes d'individus, des gestionnaires, des agences gouvernementales) et d'autre part, on modélise l'environnement qui contient une ressource exploitée. Généralement, l'état de la ressource est modifié suite à des actions effectuées par les agents, dont les comportements sont influencés par l'état de la ressource elle-même. Donc l'interaction est faite dans les deux sens.

Une méthodologie pour construire un Système d'Information Environnemental (SIE) en utilisant les SMA est présentée dans (Athanasiadis et Mitkas 2009). Les informations sur l'environnement sont appelés EDO (Environnement Data Object) qui sont regroupés pour former EU (Environnement Unit). Les agents perçoivent et agissent sur ces EDO. Un agent est considéré comme une fonction de mappage entre l'état de l'environnement et l'action qu'il doit accomplir. La méthodologie définit deux types d'agents: agent support de l'information et agent décideurs. Le support d'informations est un agent qui se comporte comme une unité de calcul qui perçoit une série d'états de l'environnement et répond en performant une série d'actions. Les agents décideurs, utilisent un moteur de raisonnement pour prendre une décision à partir de son état interne et l'état de l'environnement. La méthodologie a été utilisée pour développer deux systèmes:

O3RAA pour l'évaluation de la qualité de l'air. Le système a comme objectif la prédiction des périodes des pics de pollution. Parmi les fonctionnalités principales du système :

- La validation des mesures
- L'estimation des valeurs manquantes
- La détection des périodes de pic.

Les variables utilisées sont la concentration des SO<sub>2</sub>, O<sub>3</sub>, NO, NO<sub>2</sub>, direction du vent, température, humidité, la radiation solaire et la pression. Le système n'inclut pas une représentation de l'environnement. Il utilise des données issues d'un réseau de surveillance de la qualité de l'air. Les agents ne modélisent pas les être humains. Ils sont affectés pour surveiller des capteurs et valider les mesures. Le système est implémenté en JADE, JESS et Weka. L'étude ne mentionne pas les utilisateurs pour qui le système est conçu. Le système proposé dans cette étude n'est pas un simulateur et n'inclut pas l'aspect social ainsi que les sources d'émission de pollution. Le deuxième système conçu est nommé ABACUS qui traite la surveillance des données météorologiques.

Dans (Conyurt & al. 2009), l'objectif est d'aider les décideurs dans la gestion d'un environnement tropical insulaire de l'île de la Réunion. Cette région connaît une croissance démographique de 30%. La pression sociale devient de plus en plus influente sur la biodiversité et l'écosystème. Un système d'information pour la protection de l'environnement est conçu. Le système utilise une approche basée sur les SMA, les bases de données, les SIG et les Web service. Les SMA sont utilisés pour simuler et surveiller l'évolution des espèces dans l'île (Corail, Tortues Marin, réserve forestière, poissons). La plateforme multi-agents nommée GEAMAS-NG est utilisée pour la modélisation des comportements des agglomérations et les extensions urbaines. L'approche est décrite en quatre étapes: (i) l'identification des agents et leurs attributs. (ii) Spécifier les comportements des agents. (iii) Définition de l'environnement spatio-temporel du système. (iiii) Le traitement du scénario de simulation et la validation des résultats avec des experts.

Le système socio-environnemental simulé, est approché comme étant une association entre un ensemble d'activités humaines qui ont leurs effets sur l'environnement. Les auteurs ont présenté une approche appelée modélisation orientée activité.

Deux systèmes basés sur les SMA ont été conçus en utilisant cette approche: Le premier nommé BIOMAS, ce système est conçu pour simuler l'influence de l'agriculture sur l'écosystème, en particulier les déchets d'origine animale en provenance des fermes. Le deuxième nommé SMAT, conçu dans le but d'étudier l'influence des milieux urbains et les activités agricoles sur la biodiversité dans l'île.

Les deux systèmes utilisent une représentation réelle de l'environnement, un module SIG est inclus. Le modèle est basé sur des agents réactifs et la partie prise de décision n'est pas assez détaillée.

Dans (Corchado & al. 2009) un système de simulation combine SMA et CBR (Case Based Reasoning) pour détecter des fuites de pétrole dans l'océan, et donner des prévisions sur leurs évolutions et leurs trajectoires. Le modèle utilise des paramètres météorologiques et des images satellitaires. Le simulateur permet d'aider dans la gestion des cas où une fuite de pétrole a lieu dans l'océan. Le simulateur combine entre les réseaux de neurones et les CBR (raisonnement à base de cas). Il est formé de plusieurs agents chacun est doté d'une connaissance locale concernant les fuites et leurs localisation, les agents coopèrent ensemble pour produire la trajectoire possible de la fuite.

Les écosystèmes dans les régions côtières sont soumis à des perturbations causées par l'aquaculture. Pour trouver la meilleure façon d'exploiter les côtes sans altérer l'équilibre de l'écosystème. Une simulation peut aider les décideurs pour mieux chercher une exploitation optimale. Dans (Pereira et al. 2009) le développement d'un simulateur basé SMA est décrit. EcoSimNet est utilisé pour simuler les comportements des aquaculteurs afin de prévoir l'impact de leurs décisions. Le modèle est utilisé pour optimiser les actions des aquaculteurs sur la base des paramètres économiques. L'environnement est représenté par une grille de 2 dimensions, chaque cellule représente une partie de la côte qui peut être exploitée. Le simulateur est implémenté en utilisant l'outil EcoDynamo qui permet de simuler les dynamiques écologiques. Des données générées ont été utilisées pour valider le modèle de simulation. Trois mécanismes de prise de décision ont été utilisés : Algorithme génétique, apprentissage renforcé et l'algorithme du Recuit Simulé.

Le trafic routier est parmi les causes principales de la dégradation de la qualité de l'air en milieu urbain. (Borri & Camarada 2011) présente ECROUB, un système de gestion de qualité de l'air dans les régions urbaines basé SMA. En utilisant des modèles physiques, le système a été conçu pour générer des informations sur le micro-climat (informations sur le climat dans une très petite zone géographique) de la région de Bari en Italie. Le système a permis de montrer qu'une approche hybride (SMA et les modèles physiques) peut aider à la compréhension et l'étude de la pollution atmosphérique dans les zones urbaines. Le simulateur utilise un module SIG pour représenter l'environnement.

(Papaleonidas & Iliadis 2012) présente un SMA conçu pour surveiller la qualité de l'air dans la région d'Athènes (Grèce). Il est composé d'un ensemble d'agents logiciels. Les agents ont comme tâche la surveillance d'un réseau de capteurs installés dans des positions différentes d'une région urbaine. Les agents vérifient et stockent les données mesurées par les capteurs. Dans le cas où un capteur est endommagé et les données sur la concentration de polluant n'est donc pas disponible, le système utilise une prédiction donnée par un réseau de neurone artificiel. Cette prédiction est utilisée pour estimer la qualité de l'air. Les données réelles sur la concentration d'Ozone et les données météorologiques ont été utilisées pour concevoir le modèle de prédiction. Il est à noter que le système n'inclut pas la représentation de l'aspect social et l'activité humaine relative à l'émission n'est pas modélisée.

Dans (Ahat et al. 2009), un système multi-agents est conçu pour modéliser la pollution de l'air dans une zone urbaine. L'environnement est représenté par une grille de deux dimensions  $G(X,Y)$ .

Le but de la simulation est de trouver la dispersion de la pollution de l'air sur cette grille. Chaque point  $C(X,Y)$  de cette grille possède une valeur de concentration d'un polluant. Les cellules voisines avec un taux de pollution proche (selon un seuil fixé au départ) sont fusionnées et forment un cluster.

Les sources de pollution sont représentées par des agents homogènes qui émettent la pollution dans leurs zones (pollueurs). Chaque agent pollue avec un taux d'émission  $DE$ . Le système proposé repose sur l'utilisation d'une technique algébrique Prétopologie, qui permet de fusionner les zones pollués pour former des clusters de pollution. Au départ la grille ne contient qu'un seul cluster avec une valeur de pollution nulle. Aux fins et aux mesures de la simulation, des clusters se forment avec des valeurs différentes de concentration de pollution. La concentration de la pollution augmente, à la fin un seul cluster est formé et l'émergence de la dispersion de la pollution peut être estimée. Le système inclut le taux de destruction de pollution (évaporation)  $ER$ . Les résultats permettent de donner la dispersion de pollution en précisant le nombre d'agents pollueurs et le temps nécessaire pour que le cluster global se forme. Ce modèle n'inclut pas les paramètres climatiques et aussi ne traite pas une espèce polluante précise et donc n'utilise pas des données réelles. Le prototype a été implémenté en utilisant la plateforme multi-agent Repast Symphonie.

(Hülsmann & al. 2011) présentent une extension du simulateur MATSim pour prendre en considération l'émission de pollution atmosphérique résultant du trafic routier dans une région urbaine. MATSim est un simulateur basé SMA qui permet de simuler le trafic routier. Il procède en quatre étapes : génération des plans pour chaque individu (agent), les plans générés sont après exécutés, tous les plans exécutés sont après évalués selon une fonction d'utilité, les agents après apprennent à améliorer leurs plans selon l'évaluation effectuée dans l'étape précédente. Le système est exécuté jusqu'à ce qu'il arrive à un état stable. Pour chaque route, le simulateur définit une queue (file), dans laquelle les voitures sont gérés en premier arrivé premier servi. Le système est capable de calculer l'émission, selon le temps consommé par un véhicule dans la route, son modèle et sa vitesse. L'émission de polluant par les véhicules sont catégorisées en deux classes : (1) les émissions en air chaud, qui sont due à la consommation du carburant lorsque le véhicule est en marche. (2) les émissions en démarrage à froid. Les facteurs pour l'émission pour chaque

véhicule sont définis selon la base de données HBEFA, qui est une base de données qui contient les modèles de véhicule et leurs facteurs d'émission. Le système a été validé en utilisant les résultats que donne un autre modèle de simulation nommé PHEM, une comparaison est décrite entre les résultats que donne chaque système de simulation. Le simulateur permet de bien étudier la relation entre la circulation routière et la pollution atmosphérique. Une autre version du même simulateur est présentée dans (Hülsmann & al. 2013), elle permet d'étudier l'influence des coûts de transport (politique de tarification) sur les paternes de circulation et la pollution de l'air. L'utilité d'un trajet est calculée en utilisant deux fonctions : selon les prix et selon la pollution.

(Lkhoo & Meng 2010) présentent un simulateur de circulation routière qui permet de modéliser comment la pollution de l'air résultant de la circulation peut être contrôlée en utilisant la stratégie de tarification des transports. Le modèle SMA proposé, est organisé en deux niveaux : gestion et exécution. Les décisions concernant la modification des tarifications sont prises par des agents « gestionnaires » qui font leurs décisions sur les données concernant la pollution de l'air reçu de la part des agents « exécution ». Le modèle est testé en utilisant des données concernant le trafic routier de la ville de New Town à Singapore. Le système a permis de tester plusieurs scénarios de tarification.

Dans (Oprea & al. 2011), un réseau de contrôle de pollution de l'air est doté d'un SMA. Chaque station du réseau est dotée de plusieurs agents, où chacun contrôle la concentration d'un polluant de l'air. Des agents centraux reçoivent selon un laps de temps les mesures et appliquent des analyses. Les analyses permettent de produire des bilans concernant la qualité de l'air et les alertes de pollution. La prise de décision se fait à l'aide des arbres de décision.

Pour comprendre l'influence de la pollution sur la santé. (Sokolova & Fernández-Caballero, 2012), présente DECIMAS, un système d'aide à la décision basé sur les SMA. Le modèle proposé utilise la technique de la fouille des données pour trouver les corrélations possibles entre les paramètres de pollution et les symptômes de certaines maladies. Les auteurs utilisent un SMA pour évaluer l'impact de l'environnement sur la santé humaine. L'objectif est d'aider les experts à identifier la relation entre certains polluants et certaines maladies. Les données utilisées proviennent de différentes sources: Les bases de données des systèmes d'informations des hôpitaux et les organismes de surveillance de l'environnement. Une liste de 30 variables est utilisée, avec des différents modèles de prédiction: RNA et régression linéaire. Le mécanisme de prise de décision se base sur un modèle BDI, la tâche affectée aux agents est d'appliquer une fouille de données et de faire des agrégations afin de trouver les corrélations

entre les variables de pollution et les classes des maladies. Le système est implémenté en utilisant : les ontologies, la plateforme JACK et Prometheus.

Afin de réduire la pollution atmosphérique émise par le trafic routier, (Natalja & al. 2011) présente un SMA pour optimiser la consommation du carburant dans les intersections dans les routes. Ils modélisent le système d'intersection comme un SMA et chaque véhicule est équipé d'agent qui permet d'aider et communiquer avec les agents d'intersection. Les interactions entre les agents permettent d'optimiser le temps d'attente. L'étude montre une réduction de 28% de la consommation de carburant, ce qui conduit à une amélioration nette de la qualité de l'air. Le modèle calcule la consommation de carburant basée sur l'énergie, la force de résistance de l'air, la capacité du moteur, l'accélération et la vitesse.

Dans (Weiqian & al. 2012) un modèle SMA est déployé sur un environnement de grille de calcul pour la simulation visuelle en 3D de la pollution des eaux dans une rivière. La plateforme Globus est utilisée pour construire la grille de calcul. Le bassin d'eau a été modélisé sous-forme d'une masse 3D, composé d'un nombre d'agents appelés WaterAgent, en coopérant avec leurs voisins afin de calculer l'influence du flux de la rivière sur la propagation de la pollution. La diffusion du polluant dans l'eau est modélisé grâce un algorithme de diffusion. Qui prend comme paramètre les flux dans la rivière. Le système inclut un module SIG et utilise aussi la technique GPS pour calculer le trajet de la pollution dans la rivière. Les données utilisées pour valider la simulation proviennent des mesures effectuées pendant un incident de pollution qui a eu lieu en Octobre 2010. La puissance de calcul qu'offre la grille de calcul, a permis de visualiser la simulation de l'incident de pollution.

Dans (Brady & al. 2012), une extension du système AgriPoliS est présentée. AgriPoliS se base sur un modèle multi-agents pour simuler les effets de la politique agricole sur les agriculteurs et l'utilisation des terres. Le modèle AgrPoliS, permet de voir les changements dans l'utilisation des terres en réponse aux changements d'une politique agricole. Il permet d'utiliser les données relatives à l'investissement et les statistiques de production dans une région géographique réelle. Les auteurs présentent une amélioration de modèle AgriPoliS pour prendre en considération les effets du politique agricole sur l'utilisation des terres et aussi la biodiversité. L'environnement est représenté par une grille de  $N \times M$  cellules. Chaque cellule représente une partie du terrain cultivable de trois niveaux de fertilité (haut, moyen, bas). Le terrain peut être acheté, loué ou bien abandonné.

Chaque agriculteur (agent) peut posséder ou louer des cellules. Aux fins et aux mesures de la simulation, les agents agriculteurs doivent prendre des décisions concernant : le produit à cultiver, les terres à louer ou à vendre/acheter et aussi concernant l'abondant ou non de l'activité. Le but est donc de voir comment le comportement des agriculteurs peut émerger et affecter la structure et l'écosystème de la région étudiée. L'entité principale c'est farm-agent, qui possède les attributs suivants : quantité et âge des machines utilisées par l'agriculteur, les terres que possède ou loue l'agent, capital, nombre de personnes dans la famille et capacité de gestion. Ces attributs fusionnés ensemble constituent la capacité de l'agriculteur à gagner des terres et à produire. Le marché de productions modélisé concerne : la production du lait, céréale et viande. Ces produits utilisent des prix issus du marché international. La distribution des cellules selon produit cultivé est un paterne qui permet d'évaluer le degré d'homogénéité, ce degré est calculé selon l'indice de Shannon-Wiener (5.1)

$$SW_i = \sum p_i \ln p_i \quad (5.1)$$

Le modèle est calibré avec les données réelles des statistiques de productions. Et permet de simuler l'évolution de l'utilisation des terres pour 10-20 ans en avance. Le mécanisme de prise de décision est stochastique et se base sur des indices calculées.

Un modèle hybride SMA-Algorithmme génétique est présenté dans (Qian-jin & al. 2010). Le modèle est utilisé pour optimiser la gestion des déchets pour prévenir la pollution des eaux dans la rivière Urumqi dans le Xinjiang (Chine). Chaque agent possède des attributs (énergie, position dans la grille de N cellules x N). La concurrence entre les agents et leurs voisins permet d'identifier comment le système permettra d'optimiser le processus de traitement des déchets. Un des principaux avantages de SMA est la capacité à modéliser le processus de prise de décision humaine. (Smajgl & al. 2011) présente une méthode pour le paramétrage des comportements humains dans un modèle environnemental basé SMA. Il existe plusieurs méthodes empiriques utilisées pour le paramétrage des comportements des agents décideurs qui modélisent des êtres humaines. Les données utilisées sont généralement: Les connaissances d'experts, l'observation, les enquêtes sociales, des interviews, des données de recensement, des expériences de terrain, jeu de rôle, analyse typologique, cartographie et la méthode de Monte-Carlo. La méthode présentée permet d'aider les modélisateurs à choisir les méthodes appropriées pour le paramétrage des comportements de l'agent décideur. La méthode proposée a été utilisée pour la conception de plusieurs modèles SMA. Le modèle SimPaSI est conçu pour simuler le système socio-écologique de l'île de Java (Indonésie). Le système comprend des agents qui représentent les agriculteurs, les

ménages. Les résultats du système donnent des prédictions sur l'effet de la politique de prix du carburant sur le système socio-écologique de l'île.

Pour comprendre l'interaction environnement/humain dans les systèmes d'agriculture, (Schreinemachers & berger 2011) présentent un simulateur appelé MP-MAS (programmation mathématique Multi-Agent System). MP-MAS permet de simuler une population d'agriculteurs. La dynamique du marché; changements dans l'environnement; L'intervention des pouvoirs; Changement des technologies agricoles. L'objectif de la simulation est de prévoir l'impact de ces changements sur les ressources agro-écologiques tels que l'eau ou la fertilité des sols. Dans MP-MAS l'agent prend des décisions à propos de l'activité à faire. L'agent avant de prendre une décision

$$\max(z) = \sum_{j=1}^N C_j \times X_j \quad (5.2)$$

Il doit trouver la valeur maximale pour  $Z$  (le revenu de l'agriculteur). Où  $X$  est une activité parmi  $N$  activités agricoles possibles et  $C$  est le rendement attendu de l'activité. Pendant le déroulement de la simulation les agents essayent de trouver la meilleure activité à adopter selon les conditions et les paramètres économiques. L'émergence des comportements des agriculteurs permet de voir leur influence sur les ressources écologiques.

Dans (Pedell & Sterling 2011), le but est d'utiliser un modèle basé SMA pour aider à enquêter et comprendre la durabilité des ressources environnementales. Les agents représentent les citoyens et les décideurs, chacun d'entre eux dispose d'un ensemble de comportements visant à préserver les ressources en énergie et en eau. Le modèle vise à trouver les comportements les plus appropriés pour optimiser l'utilisation de l'énergie et de l'eau. Les comportements des citoyens sont influencés par les décideurs (agents de l'agence de protection de l'environnement, médias, etc.). Les agents représentant les citoyens sont libres de choisir le comportement approprié en fonction de leur contexte.

Le processus de restauration d'un écosystème endommagé doit faire intervenir des experts de différents domaines (scientifiques, organismes de gestion de l'environnement, les propriétaires fonciers, les agriculteurs, etc.). La prise de décision dans ce cas est une tâche très complexe. Dans (Watkins & al. 2013) un SMA est conçu dans le but de comprendre la prise de décision dans le processus de restauration écologique. Le modèle est un ensemble d'agents sous forme de hiérarchies et de groupes. L'interaction inter-agent et entre les groupes d'agents, est

utilisée pour mettre-à-jour les choix de l'agent. Des données empiriques ont été utilisées dans la simulation. Pour prendre des choix l'agent utilise la formule suivante :

$$P_{i,t+1} = \sum_{j=1}^n r_{i,j} P_{i,t}, \quad (5.3)$$

Où:  $P_{i,t+1}$  est la position de l'agent  $i$  au temps  $t + 1$ ,  $r_{i,j}$  est le respect de l'agent  $i$  pour l'agent  $j$ . Si les choix de l'agent sont différents de celui des autres agents (inférieur à l'écart-type) le coût de mécanisme de dissidence est utilisé pour réduire le respect de l'agent sur ses choix et d'accroître son respect pour les autres agents avec lesquels ils interagissent. Les données empiriques ont été utilisées, qui proviennent des interviews sur les organisations de la Wilderness-Chicago, un sommet durant lequel plusieurs décisions concernant la politique environnemental sont prises.

Dans (Touza & al. 2013) une simulation basée agent est décrite, elle permet de modéliser les comportements des propriétaires des terres et comment ils coopèrent pour gérer le nombre de cerfs dans leurs terres. Le modèle présenté utilise une approche de la théorie des jeux. Les propriétaires fonciers sont modélisés comme des agents égoïstes visant à maximiser leur profit en gérant une ressource écologique mobile (le nombre des Cerfs). Les agents ont des objectifs différents : quelques uns sont préoccupés par l'augmentation du nombre de cerfs dans les terres, les autres sont concernés sur la façon pour maximiser leurs profits de la chasse sportive des cerfs dans leurs terres. Les agents sont spatialement répartis dans la région étudiée. L'agent choisit de jouer avec ses voisins. La récompense gagnée forme la coopération (l'avantage à son voisin) ou la défection. La récompense n'est pas linéaire et est affectée par la ressource écologique dynamique; le nombre des cerfs affecté par la migration entre les terres et aussi par l'abattage causés par les propriétaires fonciers. Le modèle permet de bien voir la liaison entre le nombre de cerfs et les attitudes des prioritaires.

## 4 Conclusion

Les activités anthropiques sont parmi les causes principales de la pollution et les problèmes environnementaux. Ces activités doivent être incluses dans les modèles de simulation, Car la modélisation de l'interaction entre le composant social et écologique est un aspect très important des approches SMA.

Les SMA nous permettent de modéliser le réseau social des êtres-humains partageant l'exploitation des ressources environnementales communes. En manipulant le comportement au

niveau individuel et au niveau groupes cela aide à acquérir plus de connaissances et rend la simulation plus réaliste.

Les SMA aussi offrent une grande possibilité d'hybridation avec les autres approches simulation, (Parrot 2011) décrit cette possibilité comme étant une éventuelle amélioration de la modélisation socio-environnementale.

Les études traitant la pollution de l'air, modélisent bien l'aspect physique (de concentration et la dispersion des polluants), mais ne tiennent pas en considération l'aspect de décision humaine sur les sources d'émissions, par conséquent, les activités humaines causent la pollution ne sont pas modélisés.

Nous espérons exploiter ces points dans notre travail, qui vise à construire un simulateur de la pollution de l'air sur la base SMA pour la région d'Annaba (Nord-Est de l'Algérie), et en prenant les décisions des contrôleurs des sources d'émission comme élément clé de notre approche.

# Chapitre 6 : Un Framework pour le développement d'un simulateur basé multi-agents de la pollution de l'air

---

## 1 Introduction

Dans ce chapitre nous présentons l'approche qu'on propose pour la conception d'un simulateur de la pollution de l'air basé multi-agents. Elle se base sur la combinaison entre les approches classiques de modélisation de la pollution de l'air, à savoir : les modèles de dispersion et les modèles de prédiction. La combinaison des deux est intégrée ensuite dans un système multi-agents.

Nous commençons cette présentation, par la description de notre modèle de prédiction de la qualité de l'air. Nous détaillons après comment ce dernier est combiné avec un modèle de dispersion de type Gaussien. Dans la deuxième partie nous présentons les étapes qui permettent de les intégrer dans une simulation basée agent.

Comme illustrée dans la figure 6.1, notre approche se base sur trois composants :

- [1] Le modèle de dispersion de type Gaussien.
- [2] Le modèle de prédiction basé réseau de neurone artificiel.
- [3] Un système multi-agents.

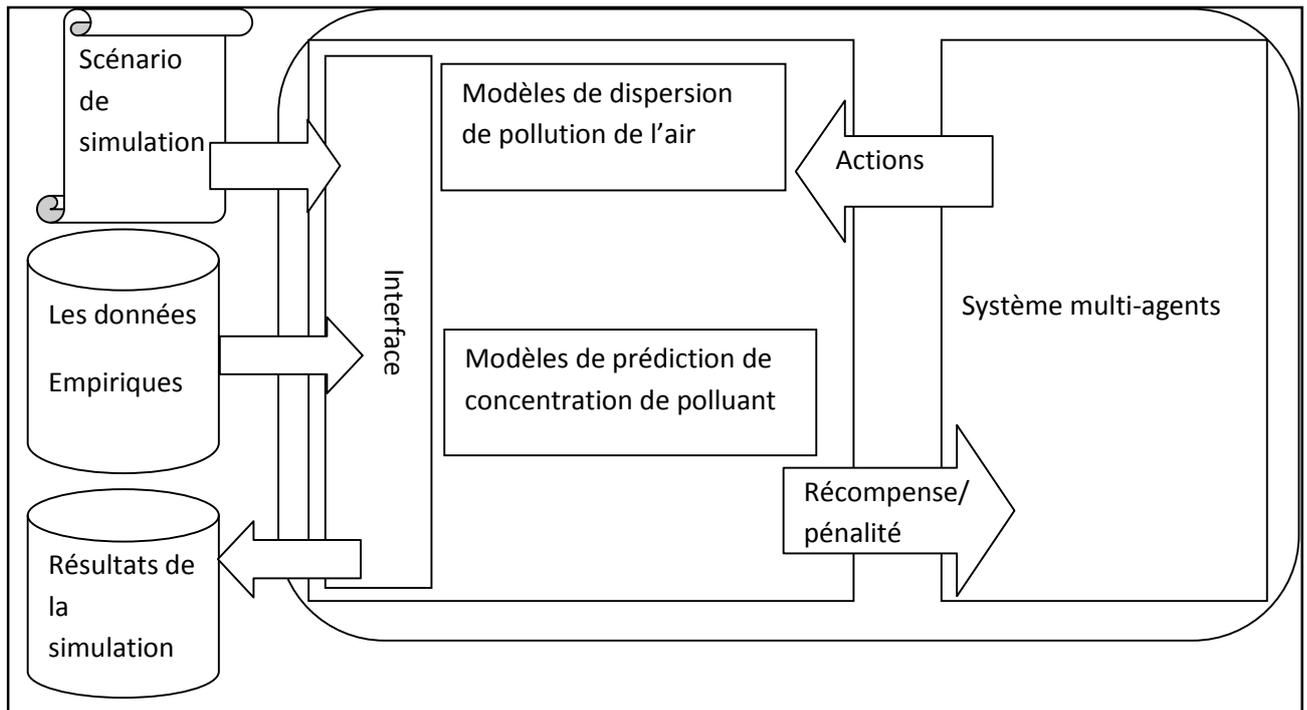


Figure 6.1 architecture du moteur de simulation proposé.

## 2 L'Approche et l'architecture du modèle:

Généralement, un système de simulation qui modélise un système socio-environnemental peut être représenté comme une interconnexion des trois composants (ou sous-systèmes); chacun est représenté par un ensemble de variables (attributs) formant son état à l'instant  $t$ .

Le composant écologique représente les parties biotiques (vivantes) et non vivantes (abiotiques). Le composant économique représente le point de vue économique et regroupe les variables liées à cela. Le composant social représente le réseau social impliqué dans la gestion et/ou exploitation d'une ressource environnementale, tels que : Les décideurs, les entreprises, les organismes gouvernementaux et les consommateurs. Un changement dans la variable d'état de chaque composant affecte les variables d'état d'autres sous-systèmes. Par exemple, l'augmentation de la demande pour certains types de poissons, conduit les pêcheurs à intensifier leur exploitation; cela, entraîne des changements à la biodiversité.

En se basant sur ce découpage, nous présentons une formalisation générique d'un modèle socio-environnemental, principalement axé sur la pollution de l'air. Un système socio-environnemental peut être vu comme un ensemble de variables d'état économiques, sociales et écologiques. L'état du système en temps  $t$  peut être formulé comme suit (6.1) :

$$ES_t = \langle Ec_t, Sc_t, Env_c_t \rangle \quad (6.1)$$

Où  $Ec$ ,  $Sc$  et  $Env_c$  représentent, l'ensemble des variables d'états économiques, sociales et environnementales, respectivement:

$$\begin{aligned} Ec_t &= \langle Ec_{1,t}, \dots, Ec_{l,t} \rangle, \\ Sc_t &= \langle Sc_{1,t}, \dots, Sc_{m,t} \rangle, \\ Env_c_t &= \langle Env_c_{1,t}, \dots, Env_c_{n,t} \rangle \end{aligned} \quad (6.2)$$

Dans notre cas, les variables d'état de l'environnement à temps  $t$  sont:

$$Env_c_t = \langle c_{0,t}, \dots, c_{n,t}, WS_t, Hu_t, T_t, RF_t \rangle \quad (6.3)$$

$C_{i,t}$  est la concentration du polluant  $i$ ,  $WS$ : c'est la vitesse du vent,  $T$ : température,  $Hu$ : humidité et  $RF$  représente la pluviosité, à temps  $t$ .

Supposant qu'une source de pollution peut être représentée à temps  $t$  comme suit:

$$S_t = \langle er_t, tc, X, Y, Z \rangle \quad (6.4)$$

La source produit (émit) le polluant  $tc$  avec un taux d'émission  $er$  à la position géographique  $(X, Y, Z)$ . Les sources sont contrôlées par des agents. Chaque agent doit prendre une décision concernant l'action à faire, pour choisir parmi toutes les actions qu'ils lui sont possibles, il doit prendre en considération l'état de l'environnement  $ES$  et son état interne à temps  $t$ .

Soit  $A$  l'ensemble de ces actions

$$A = \{a_1, a_2, \dots, a_z\}, \quad (6.5)$$

Le résultat d'une action c'est les changements du taux d'émission de polluant dans la source contrôlée par cet agent.

Nous pouvons définir ceci comme étant une fonction qui prend les actions d'un agent et donne comme résultat le nouveau taux d'émission de la source:

$$F: A \rightarrow R \quad (6.6).$$

Soit  $\pi_t$  l'ensemble d'actions accomplies par  $N$  agents au temps  $t$ :

$$\pi_t = \langle A_{0t}, \dots, A_{nt} \rangle \quad (6.7)$$

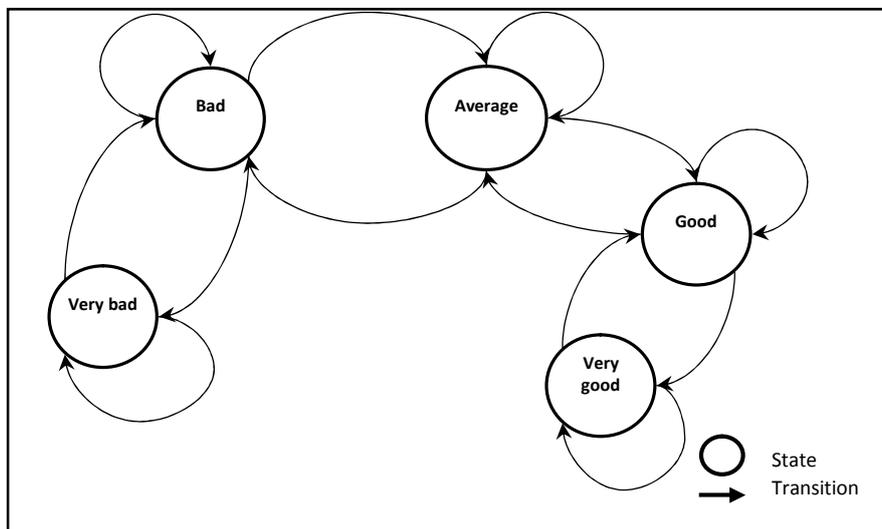
Soit  $Q$  l'ensemble des valeurs possibles pour l'index de la qualité de l'air:

$$Q = \{very\_bad, bad, average, good, very\_good\}. \quad (6.8)$$

La qualité de l'air peut après, être modélisée comme étant un graphe, comme illustré dans la figure 6.2, ce dernier est doté d'une fonction de transition:

$$T(\pi_t, ES_t, current\_q) \quad new\_q, \quad current\_q \text{ and } new\_q \in Q \quad (6.9)$$

La fonction prend comme arguments l'état du système  $ES_t$ , et l'ensemble des actions faites par les  $N$  agents, et après un pas de simulation, elle donne comme résultat la nouvelle valeur pour l'index de la qualité de l'air, permettant de changer la valeur de cette variable de  $current\_q$  vers  $new\_q$ . Il est à noter que dans certaines conditions (climatique, etc) la valeur d'index de la qualité de l'air peut rester la même, ceci veut dire que malgré les actions accomplies par les agents, leurs efforts n'a pas permet de modifier la qualité de l'air.



**Figure 6.2: Représentation graphique des valeurs possibles pour la variable “Qualité de L’air”.**

Notre approche de simulation, peut être schématisée dans la figure 6.3. Pour chaque pas de simulation, les actions de chaque agent affectent le taux d'émission des sources. Après le modèle de dispersion utilisent ces taux d'émission, nous permet alors de donner la concentration des polluants. Ces concentrations sont utilisées avec les paramètres climatiques pour donner les prédictions concernant la concentration des polluants pour les  $K$  heures en avances. Ensuite, les perdctions sont utilisées pour estimer l'index de la qualité de l'air. Selon ces prédictions les agents seront alors récompensés ou bien pénalisés. Enfin, les agents selon ce qu'ils ont eu comme résultat du dernier pas de simulation vont alors adapter leurs taux d'émission pour éviter d'être pénalisés et de gagner plus de récompenses.

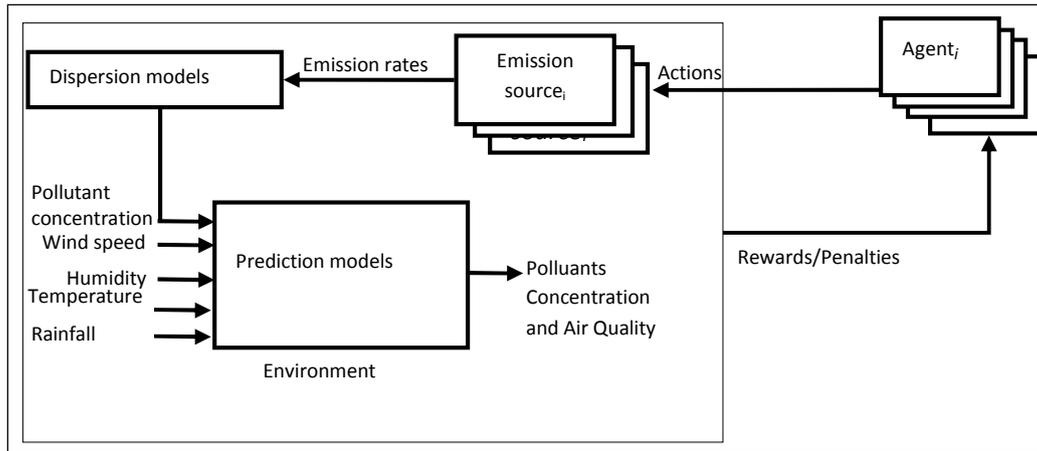


Figure 6.3: Schémas du processus de simulation.

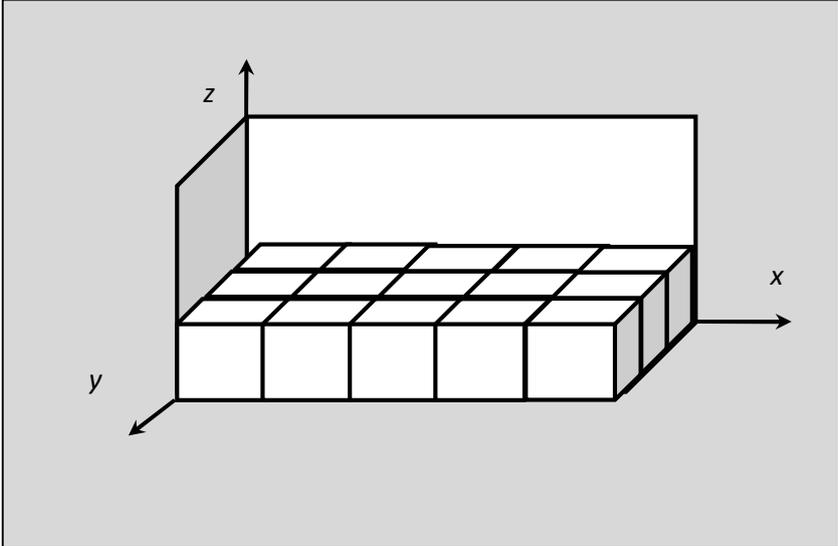
## 2.1 L'échelle spatio-temporelle du modèle de simulation

La simulation utilise une représentation discrète de temps. Chaque pas de simulation est représenté par  $k$  heures de temps réel (ce paramètre peut être adapté dans les scénarios de simulation). Le paramètre  $k$  est défini selon les objectifs de simulation (la prédiction à court terme, moyen terme ou long terme).

Pour la représentation de l'environnement nous avons opté pour une représentation en box. C'est-à-dire, que l'environnement concerné par la simulation est représenté par un ensemble de cubes equation 6.10, chacun représente 1 kilomètre carré, comme illustré dans la figure (6.4).

$$BX = \{bx_0, \dots, bx_m\}, \quad (6.10)$$

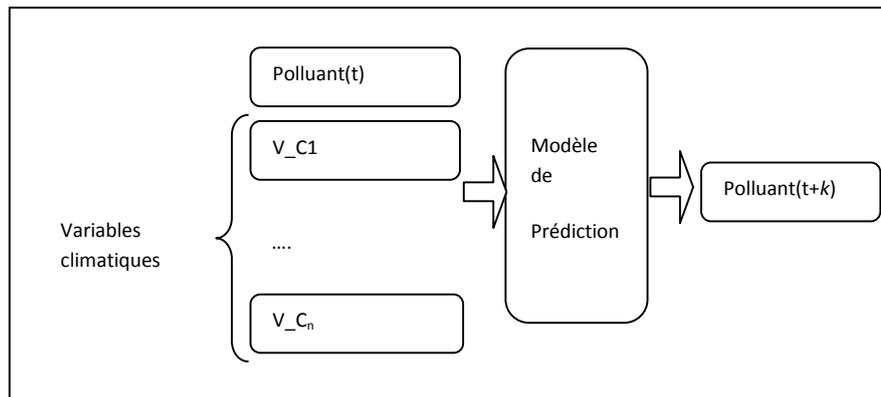
Chaque cube est localisé dans une position (géo-position) au point  $gp(x,y,z)$  et chaque cube possède un ensemble d'attributs représentant la concentration dans l'air des polluants  $(cp_0, \dots, cp_k)$  et aussi la qualité de l'air.



**Figure 6.4:** La représentation de l'environnement.

## 2.2 Le modèle de prédiction de la qualité de l'air :

Comme présenté dans le Chapitre 3, un modèle de prédiction peut être vu comme étant une fonction qui permet de lier un vecteur d'entrées (Concentration du polluant et les paramètres climatiques) à une sortie correspondante (la concentration du polluant dans l'air après  $k$  pas). Le nombre de pas  $k$  détermine l'horizon de la prédiction. La fonction qui permet de faire un tel mappage peut être schématisée dans la figure( 6.5):



**Figure 6.5: Schéma générique d'un modèle de prédiction de concentration du polluant.**

Les paramètres climatiques ont une influence très importante sur la concentration des polluants (Khedaria & Khadir 2012), comme présenté dans le Chapitre 2.

Afin de capter la relation et l'influence des paramètres climatiques, nous commençons par concevoir des modèles de prédiction pour chaque polluant. Ensuite, La combinaison entre tous les modèles de prédiction obtenus est alors établie dans le but d'avoir des prédictions concernant la qualité de l'air. Plusieurs modèles ont été testés (Ghazi & Khadir 2012), est nous avons opté pour les réseaux de neurones artificiels (RNA). Précisément les RBF (Radial Basis Function) (Buhmann 2000). Ce type de réseau de neurones artificiels a montré les performances les plus convaincants. Les RBF sont après les PMC (Perceptron MultiCouches) le modèle RNA le plus utilisé.

Il consiste en trois couches de neurones (comme illustré dans la figure 6.6) : Les neurones de la couche d'entrée sont connectés directement avec les entrées (le vecteur d'entrée), et en même temps connectés avec leur sortie avec les neurones de la couche cachée. Ces derniers ont comme fonction d'activation une fonction à base radiales. Ce genre de fonction est asymétrique autour d'un centre  $\mu$ . Un réseau à fonction à base radial est la combinaison de  $N$  fonctions à base radial avec  $N$  centres différents :

$$\hat{Z}(x) = z \sum_{j=1}^N w_j Z_j(x) \quad (6.11)$$

Ici,  $\hat{Z}$  est la fonction avec laquelle nous cherchons à estimer la sortie souhaitée.

$w_j$  c'est le poids associé à la fonction  $j$ . et  $Z_j$  est une fonction radiale elle peut être définie comme suit :

$$Z(x) = \frac{1}{\sigma_j^2} e^{-\frac{\|x - \mu\|^2}{2\sigma_j^2}} \quad (6.12)$$

Où  $x$  est un vecteur d'entrée,  $\mu$  est le centre de la fonction et  $\|\cdot\|$  est une distance Euclidienne.

La fonction Gaussienne c'est la fonction radiale la plus utilisée pour concevoir les modèles RBF. Elle peut être formulée comme suit :

$$z(x) = e^{-\frac{\|x - \mu\|^2}{2\delta^2}} \quad (6.13)$$

Où:  $x$  c'est la variable d'entrée,  $\mu$  et  $\delta$  sont respectivement, le centre et l'écart-type.

### 2.2.1 Apprentissage des réseaux RBF :

Comme presque la plupart des modèles de prédiction, les modèles RBF sont dits à apprentissage supervisé. Cela veut dire que le modèle avant d'être utilisé, il doit toujours passer par une étape d'apprentissage. Ceci consiste, dans le cas des réseaux RBF, à trouver le nombre de neurones de la couche RBF (cachées) ainsi que leurs poids  $w_j$ . Après, trouver la valeur de  $N$  le nombre de fonction nécessaires. Ceci se fait en utilisant un ensemble (ou bien un échantillon) d'apprentissage. Cet ensemble est un vecteur associant chaque entrée à sa sortie correspondante.

Durant l'apprentissage, l'algorithme de Décence-Gradient (Viennet 2006) est utilisé. Ce dernier cherche à optimiser les trois paramètres du réseau, et donc en utilisant les dérivés de l'erreur (6.14) pour chacun des paramètres :

$$E_i = 1/2(y_i - y(x_i))^2 \quad (6.14)$$

Après le dérivé par rapport au poids :

$$\frac{\partial E_i}{\partial w_j} = w_j (y_i - \sum_j w_j Z_{i,j}(x)) Z_{i,j} \quad (6.15)$$

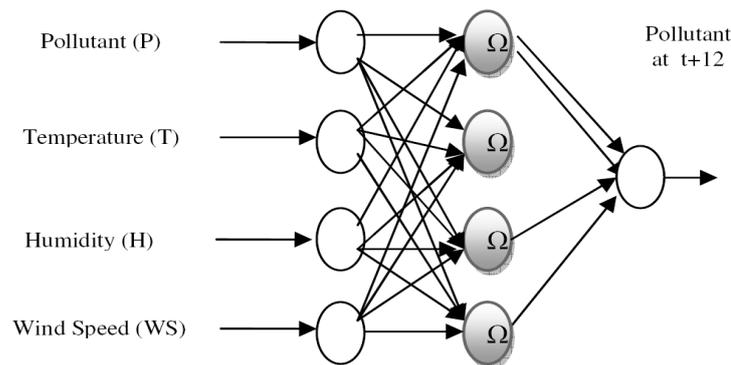
Par rapport au centre  $\mu$  :

$$\frac{\partial E_i}{\partial \mu_j} = w_j \frac{\|x - \mu\|}{\sigma_j^2} (y_i - \sum_j w_j Z_i(x)) Z_{i,j} \quad (6.16)$$

Par rapport à l'écart-type :

$$\frac{\partial E_i}{\partial \sigma_j} = w_j \frac{\|x - \mu_j\|^2}{\sigma_j^3} (y_i - \sum_j w_j Z_i(x)) Z_{i,j} \quad (6.17)$$

En utilisant ces trois équations, l'Algorithme cherchera à optimiser ces trois paramètres par rapport à l'erreur.



**Figure 6.6: Schema d'un réseau de neurones artificiels de type RBF.**

Pendant l'étape d'apprentissage, plusieurs architectures (en termes de nombre de neurones) peuvent être testées, ceci dans le but de trouver celle qui donne les meilleures performances. Ces derniers sont mesurés par rapport à leurs distances des données réelles. Généralement, pour trouver la meilleure combinaison, il faut faire une validation par rapport aussi à un ensemble de données qui n'ont pas été utilisées pendant l'apprentissage.

### **2.2.2. La combinaison des modèles RBF en utilisant un réseau MLP**

En se basant sur les modèles RBF décrits dans la section précédente, un modèle de prédiction de la qualité de l'air peut être conçu en combinant entre les prédictions sur la concentration des polluants pour donner la qualité de l'air. Ceci peut être fait par un réseau de type PMC. La figure 6.7 illustre comment dans la première étape, la prédiction des polluants est obtenue en utilisant les réseaux RBF, ceci en utilisant les valeurs actuelles (à temps  $t$ ) de la concentration des polluants et aussi des paramètres climatiques. Les valeurs obtenues sont après passées en entrée au réseau PMC dans le but de les combiner pour prédire la qualité de l'air (l'indice de la qualité de l'air).

Les prédictions de la qualité de l'air sont obtenues pour  $t+K$  heures en avance. Selon les concentrations des polluants  $\text{NO}_x$ ,  $\text{O}_3$ ,  $\text{PM}_{10}$  and  $\text{SO}_2$ . L'indice des valeurs pour la qualité de l'air sont entre 1 à 5 comme montré dans le tableau 6.1. Ce même tableau est utilisé pour la conception du réseau PMC. Un ensemble d'apprentissage est généré en utilisant ce dernier.

Ce dernier présente les sorties et les entrées utilisées pour l'apprentissage. Ceci en utilisant les standards locaux.

<b>SO<sub>2</sub></b>	<b>NO<sub>2</sub></b>	<b>O<sub>3</sub></b>	<b>PM10</b>	<b>Indices</b>	<b>Appreciation</b>
0 – 30	0-45	0-45	0-20	1	Very Good
30-60	45-80	45-80	20-40	2	Good
60-125	80-200	80-150	40-100	3	Average
125-250	200-400	150-270	100-200	4	Bad
>250	>400	>270	>200	5	Very Bad

**Table 6.2. L'indice de la qualité de l'air.**

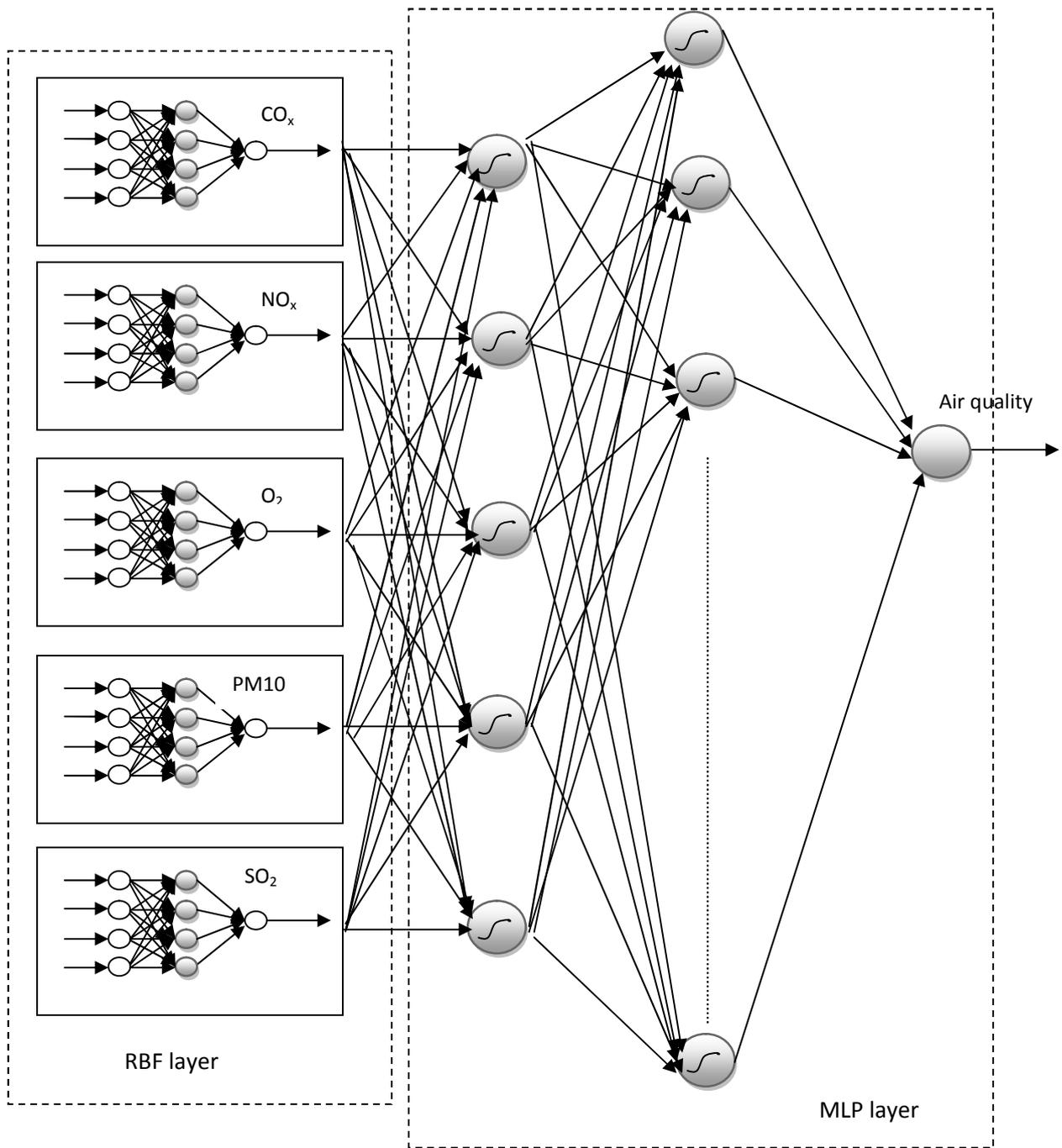


Figure 6.7: Un system Multi-Classifieur utilisant les RBF et le MLP pour la prédiction de la qualité de l'air.

## 2.3 Le modèle de dispersion:

Pour avoir l'information concernant comment les modifications des taux d'émissions affecte la concentration du polluant, notre approche se base sur l'utilisation d'un modèle de dispersion. Ce dernier utilise les données concernant les taux d'émission et la vitesse du vent pour calculer la dispersion. C'est un modèle Gaussien (voir le chapitre 3). Ce modèle permet de simuler la dispersion à partir de sources de type point (Stokie 2011).

$$C(x, y, z, H) = \frac{er_{i,t} D}{2\pi U_t \sigma_y \sigma_z} e^{-\frac{y^2}{2\sigma_y^2}} \left[ e^{-\frac{(z-H)^2}{2\sigma_z^2}} + e^{-\frac{(z+H)^2}{2\sigma_z^2}} \right] \quad (6.18)$$

Ceci veut dire que, la concentration du polluant au point  $(x,y,z)$  est calculée selon:

$er_{i,t}$ : Le taux d'émission en kilogrammes par heure à partir de la source  $i$  en temps  $t$ .

$U_t$ : La vitesse du vent en mètres par second en temps  $t$ ,

$\sigma_y, \sigma_z$ : c'est l'écart-type de la distribution de la concentration en contre le vent en vertical, ces deux paramètres sont choisis selon les classes de stabilité 'C' définie par Guifford-Pasquill,

$H$  exprime l'hauteur de la source par rapport à la terre.

Le terme de décomposition  $D$  est donné selon (EPA, 1995) et calculé en utilisant (6.19).

$$D = \begin{cases} e^{(1/(R \frac{x}{U}))}, & \text{if } R > 0, \\ 1, & \text{if } R = 0 \end{cases} \quad (6.19)$$

Où  $x$  est la distance du point par rapport à la source dans la direction du vent,

$u$  c'est la vitesse du vent

Et  $R$  c'est le coefficient de décomposition. Les valeurs pour  $R$  sont adoptées d'après (Behera & Sharma 2011) pour le  $NO_x$  (0.45 h-1) et  $SO_x$  (0.31 h-1), et pour les  $PM_{10}$  il n'est pas considéré ( $R=0$  donc  $D=1$ ).

Le niveau de concentration résultant des sources de pollution sont agrégés, et la moyenne dans chaque cube est calculée. Ces valeurs sont ensuite passées comme entré aux modèles de prédiction appropriés selon le type de polluant.

### **2.3.1 Les polluants secondaires**

Pour le cas des polluants secondaires (non émis) qui ne résultent pas d'une source d'émission, l'algorithme de dispersion ne peut être utilisé. Par conséquent, pour pouvoir faire des prédictions concernant leurs concentrations nous avons utilisé la concentration des précurseurs. C'est-à-dire, les polluants qui participent à la formation de ce polluant secondaire.

L'Ozone troposphérique résulte de l'interaction photochimique des autres polluants, particulièrement les  $\text{NO}_x$  et les  $\text{SO}_x$ . Par conséquent, pour prédire la concentration de l'Ozone nous avons utilisé la concentration donnée par l'algorithme de dispersion concernant les  $\text{NO}_x$  et  $\text{SO}_x$  qui sont ses précurseurs.

## **2.4 La simulation basée agent**

En se basant sur son état interne et l'état de l'environnement, un agent doit choisir d'accomplir une action parmi les actions qui lui sont possibles. Cette action est sensée être faite dans le but de se rapprocher de ses objectifs. Ce processus s'appelle "prise de décision". (Li 2012) présente une revue des méthodes utilisées pour modéliser la prise de décision dans une simulation socio-environnementale.

Notre approche définit deux stratégies de coopération (centralisée et jeux évolutifs) chacun définit un mécanisme de prise de décision. La stratégie centralisée (CS) est basée sur la définition d'un agent central qui représente l'organisme de contrôle de la pollution de l'air. L'agent central prend des décisions en fonction du niveau de pollution de l'air actuel. La deuxième stratégie est basée sur un jeu évolutif, où les agents sont récompensés et sanctionnés selon le niveau des pollutions; ils prennent des décisions en fonction de leurs récompenses. Dans notre système, la stratégie de coopération est définie parmi les paramètres de simulation.

### **2.4.1 La stratégie de coopération centralisée (CS)**

Dans cette stratégie la tâche de maintenir la qualité de l'air est affectée à un agent. Ce dernier représente l'organisme de contrôle de la pollution de l'air. Il utilise les modèles de dispersion et les RNA pour pouvoir prévoir la qualité de l'air, ainsi que les concentrations des polluants. Ensuite, selon les prévisions, il envoie un message demandant de réduire les émissions aux autres agents qui contrôlent les sources d'émission. Il vérifie ensuite la qualité de l'air. Il continuera à faire ce cycle jusqu'à ce que la qualité de l'air s'améliore. Son objectif principal est d'atteindre un indice de qualité de l'air qui soit meilleure par rapport à l'indice actuel.

L'agent a une autorité centrale et ses commandes sont exécutées par les agents contrôleurs des sources d'émission. Les agents communiquent leurs taux d'émission à chaque étape de la simulation. Cette stratégie est basée sur la communication entre les agents. Nous supposons que les agents sont rationnels et ont une personnalité préférant de préserver l'environnement (éco-responsable); cela signifie qu'ils favorisent l'amélioration de la qualité de l'air sur leurs propres intérêts et communiquent à l'agent central leur taux d'émission exact.

Ceci peut être résumé par l'algorithme suivant :

---

**Algorithm 1** Algorithm Centralized Strategy

---

```

while  $t < maxTemps$  do
  for all  $S_i$  in  $S$  do
     $RunDispersion(S_i)$ ;
  end for
  for all  $box$  in  $B$  do
    Compute  $box.PM10, box.NO_x, box.SO_x, box.O_3$ 
  end for
  Compute for each pollutant  $avgPM10, avgNO_x, avgSO_x, avgO_3$ 
   $currPM10 = PredictionRNAPM10(avgPM10, ws, hu, temp, rf)$ 
   $currNO_x = PredictionRNANO_x(avgNO_x, ws, hu, temp, rf)$ 
   $currSO_x = PredictionRNASO_x(avgSO_x, ws, hu, temp, rf)$ 
   $currO_3 = PredictionRNAPM10(avgO_3, ws, hu, temp, rf)$ 
   $airQuality = PredictionRNA(currPM10, currNO_x, currSO_x, currO_3)$ .
  if  $airQuality \geq GGairquality$  then
    for all  $agnt$  in  $agentList$  do
       $SendMsg(REDUCE, Toagnt)$ 
    end for
  end if
   $t=t+1$ 
end while

```

---

### 2.4.2. La stratégie de coopération basée jeux évolutifs

Dans cette stratégie, tous les agents ont leurs propres objectifs (gagner plus de récompenses et aussi garder un taux d'émission le plus haut). Les agents aussi partagent l'objectif global qui est la qualité de l'air fixé parmi les paramètres de simulation.

Une fonction d'appréciation permet de définir un tel objectif, comme suit :  $app: Q \rightarrow R$ , ceci permet de comparer les indices de qualité de l'air en leurs affectant des valeurs réelles; le but global GG peut être formulé comme suit:

$$GG_t = \{T(\pi_t, q_t) \rightarrow q_{t+1}, app(q_{t+1}) > app(q_t)\} \quad (6.20)$$

Ceci veut dire, que notre système essaye de trouver l'ensemble d'action  $\pi_t$ , que doivent appliquer les agents en temps  $t$ , qui permettent de changer la qualité de l'air  $q_{t+1}$  vers une nouvelle indice meilleure que l'indice courante.

### 2.4.2.1 Dilemme de prisonniers à N personnes:

Le dilemme de prisonnier est un modèle mathématique, il permet de représenter le conflit entre deux joueurs qui partagent l'exploitation d'un bien commun. L'action de chaque joueur a une influence sur les gains de l'autre. Au départ elle est formulée pour deux joueurs qui ont deux actions possibles : Coopérer, Non-coopérer. La récompense pour chaque joueur se fait selon la fonction suivante :

Actions	Coopérer	Non-coopérer
Coopérer	4	1
Non-coopérer	5	2

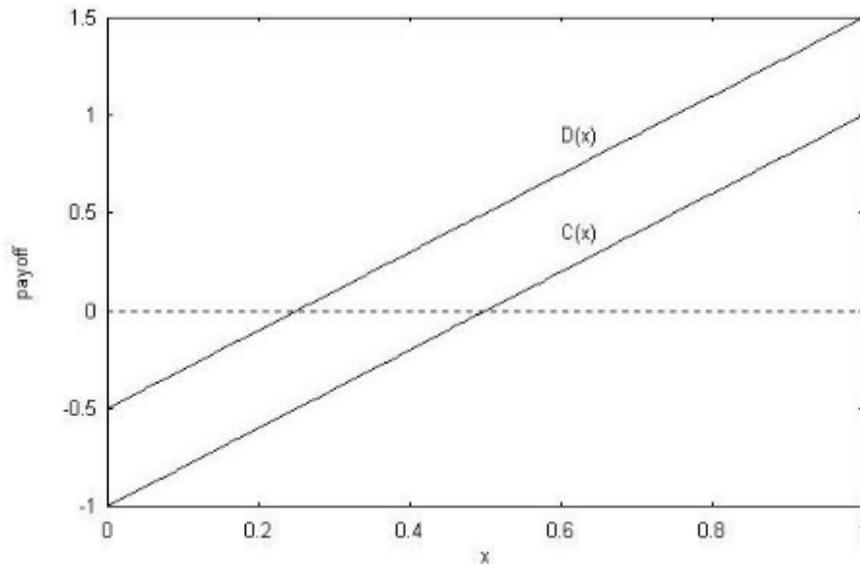
**Tableau 6.3, La fonction de récompense de jeux.**

Sachant qu'aucun d'entre eux n'a l'information concernant le choix de l'autre joueur. Ce modèle permet de bien modéliser les situations dans lesquelles un bien commun est exploité par plusieurs individus. Une version pour N-personnes a été proposée par (Boyd & Richardson 1988) où les récompenses sont calculées selon le nombre d'agents qui choisissent de coopérer; la fonction de gain est donnée comme suit (6.21).

$$u(i) = \begin{cases} \frac{b}{N} & \text{if } s = 0 \\ \frac{b}{N} & \text{if } s = 1 \end{cases} \quad c \quad (6.21)$$

Avec  $b > c > 0$  et  $c > \frac{b}{N}$ , et l'ensemble des actions est  $\{0,1\}$  (où 0 représente l'action de coopérer et 1 veut dire ne pas coopérer) dans notre cas les actions sont ( Réduire, Augmenter le taux d'émission),  $i$  c'est le nombre des agents qui ont choisi de coopérer et  $N$  c'est la taille de la populations des joueurs et  $b$  c'est la tentation ( le gain) de non-coopération.

Les deux paramètres  $b$  et  $c$  peuvent avoir des valeurs comme:  $b=2$  et  $c=-0.5$ . Ces deux paramètres peuvent être redéfinis selon les besoins. Comme illustré dans la figure 6.8, les deux fonctions permettent ensuite de calculer les gains des coopérateurs et les non-coopérateurs, ceci se fait selon le nombre des agents qui choisissent de coopérer.



**Figure 6.8 : Les deux courbes de récompense, une pour les coopérateurs et l'autre pour les non-coopérateurs.**

Dans notre cas, à chaque pas de simulation, chaque agent reçoit une récompense ou bien une pénalité selon ses actions, et aussi selon le taux de pollution dans lequel il participe (le polluant qu'il émit). Notre approche s'inspire du jeu du dilemme de prisonnier à  $N$ -personnes (DPNP). Les agents participent à un jeu de type DPNM, ils doivent coopérer avec les autres agents pour maximiser leurs récompenses gagnés. Nous avons adopté la méthode décrite dans (Zhao-Han et al. 2008). L'agent garde trace de ses  $L$  derniers pas du jeu (actions, récompense et les récompenses de ses voisins).

Les décisions concernant les actions à faire, sont prises selon deux probabilités (en suivant l'approche de (Conrad Power 2009)). A chaque étape de simulation  $t$ , l'agent calcule la moyenne pondérée de ses  $L$  derniers gain selon (13) :

$$WP_i(t) = \sum_{i=1}^L w_i M_i(t) \quad (6.22)$$

Où:

$w_i$  est le poids affecté à chaque pas dans la mémoire de l'agent, sachant que la somme de ces  $k$  poids est égale à 1

$$\sum_{i=1}^{k-1} w_i = 1 \text{ et } \forall i, j (i < j \rightarrow w_i > w_j), (6.23)$$

$M_i$  est le gain du pas  $i$ ,  $i=1$  veut dire le gain gagné pendant l'étape courante.

Après l'agent, calcule ses probabilités pour réduire ou augmenter son taux d'émission selon (6.24) et (6.25) :

$$\begin{cases} P_{c_i}(t+1) = P_{c_i}(t) + (1 - P_{c_i}(t) - \alpha_i(t)) \cdot \alpha_i(t), & \text{if } S_i = 0 \text{ and } WP_i(t) > 0 \\ P_{c_i}(t+1) = (1 - \alpha_i(t)) \cdot P_{c_i}(t), & \text{if } S_i = 0 \text{ and } WP_i(t) \leq 0 \end{cases} \quad (6.24)$$

$$\begin{cases} P_{d_i}(t+1) = P_{d_i}(t) + (1 - P_{d_i}(t) - \alpha_i(t)) \cdot \alpha_i(t), & \text{if } S_i = 1 \text{ and } WP_i(t) > 0 \\ P_{d_i}(t+1) = (1 - \alpha_i(t)) \cdot P_{d_i}(t), & \text{if } S_i = 1 \text{ and } WP_i(t) \leq 0 \end{cases} \quad (6.25)$$

Où :  $P_{c_i}$  et  $P_{d_i}$  sont, respectivement, la probabilité pour réduire ( $S=0$ ) et augmenter ( $S=1$ ) son taux d'émission pour l'agent  $i$ ,  $\alpha$  est le facteur d'apprentissage,

$S$  est l'action jouée à temps  $t$

Le taux d'apprentissage est mis-à-jour selon (6.27) :

$$D_i = \sum_{j=1}^{L-1} \begin{cases} 0 & \text{if } X_{i,j} = X_{i,j+1} \\ 1 & \text{if } X_{i,j} \neq X_{i,j+1} \end{cases} \quad (6.26)$$

$$\begin{cases} \alpha_i(t+1) = \alpha_i(t) + 0.015 & \text{if } D = 0 \\ \alpha_i(t+1) = \alpha_i(t) + 0.010 & \text{if } D > k \\ \alpha_i(t+1) = \alpha_i(t) - 0.010 & \text{if } D \leq k \end{cases} \quad (6.27)$$

Ici,  $D$  est l'indicateur d'homogénéité de l'agent  $i$ , en temps  $t$ . Cet indicateur est utilisé pour inciter les agents à avoir des actions homogènes.

L'agent est influencé par ses voisins à chaque pas de simulation. Ceci se fait en utilisant la moyenne des gains de ses voisins calculée selon (6.28).

$$nP_i(t) = (\sum_{j=1}^R C_j) / R_i \quad (6.28)$$

Où  $C_j$  est le gain du voisin  $j$  et  $R_i$  est le nombre des voisins de l'agent  $i$ .

La moyenne des  $L$  dernières valeurs de  $nP$  est notée  $avgNP$ .

En utilisant toutes ses informations, l'agent peut choisir une action à prendre pour le prochain pas de simulation, en utilisant (19):

$$\begin{cases} \text{if } S_i(t) = 0, S_i(t+1) = \begin{cases} 1, & \text{if } WP_i < avgNP_i \text{ and } P_{d_i}(t+1) > P_{c_i}(t+1), \\ 0, & \text{else} \end{cases} \\ \text{if } S_i(t) = 1, S_i(t+1) = \begin{cases} 0, & \text{if } WP_i < avgNP_i \text{ and } P_{c_i}(t+1) > P_{d_i}(t+1), \\ 1, & \text{else} \end{cases} \end{cases} \quad (6.29)$$

---

**Algorithm 2** NPPD Choisir Action Agent Joueur

---

```
while  $t < MaxTemps$  do
  lastChoices[2] = lastChoices[1];
  lastChoices[1] = lastChoices[0];
  if (lastChoices[0] == 0) then
    if (( $RPwt < pf_{AVG}$ ) and ( $P < Q$ ) and ( $Q > Ru$ )) then
      lastChoices[0] = 1;
      sourcesInfo.resumeEmission();
    else
      lastChoices[0] = 0;
      .sourcesInfo.reduceEmission();
    end if
  end if
  if (lastChoices[0] == 1) then
    if (( $RPwt < pf_{AVG}$ ) and ( $Q < P$ ) and ( $P > Ru$ )) then
      lastChoices[0] = 0;
      sourcesInfo.reduceEmission();
    else
      lastChoices[0] = 1;
      sourcesInfo.resumeEmission();
    end if
  end if
  RPwt();
  updateAlpha();
  updateP();
  updateQ();
end while
```

---

**Algorithm 3** NPPD Calcule recompense Agent Game-Runner

```

while  $t < MaxTemps$  do
  done = true
  for all  $agent_i$  in  $agentList$  do
    done = done and receiveMSGdone()
  end for
  if done is true then
    for all  $agent_i$  in  $AgentList$  do
      em = AskForAction( $agent_i$ );
      if ( $em.lastChoices[0] == 0$ ) then
        numberOfCooperators ++;
        if ( $em.Ptype == PollutantType.PM10$ ) then
          numberOfCooperatorsPM10 ++
        end if
        if ( $em.Ptype == PollutantType.SOx$ ) then
          numberOfCooperatorsSOX ++
        end if
        if ( $em.Ptype == PollutantType.NOx$ ) then
          numberOfCooperatorsNOX ++
        end if
        if ( $em.Ptype == PollutantType.COx$ ) then
          numberOfCooperatorsCOX ++
        end if
      else
        numberOfDefectors ++
        if ( $em.Ptype == PollutantType.PM10$ ) then
          numberOfDefectorsPM10 ++;
        end if
        if ( $em.Ptype == PollutantType.SOx$ ) then
          numberOfDefectorsSOX ++
        end if
        if ( $em.Ptype == PollutantType.NOx$ ) then
          numberOfDefectorsNOX ++
        end if
        if ( $em.Ptype == PollutantType.COx$ ) then
          numberOfDefectorsCOX ++
        end if
      end if
      totalPayOffAgent = totalPayOffAgent + em.rewards
      totalPayOffAgentPM10 = totalPayOffAgentPM10 + em.rewards
      totalPayOffAgentSOX = totalPayOffAgentSOX + em.rewards
      totalPayOffAgentCOX = totalPayOffAgentCOX + em.rewards
      totalPayOffAgentNOX = totalPayOffAgentNOX + em.rewards
    end for
    rewardForC =  $Cx(numberOfCooperators, eAgentList.size(), 0)$ 
    rewardForD =  $Dx(numberOfCooperators, eAgentList.size(), 1)$ 
  end if
end while

```

### 2.4.2.2 La pénalité liée à la pollution

Lorsque le niveau de pollution est plus grand que le niveau souhaité (fixé dans le scénario de simulation par exemple). Les agents participant à cette situation sont pénalisés. Ceci se fait en calculant tous d'abord leurs participations  $\sigma_i(t)$ . La participation de l'agent  $i$  en temps  $t$  est calculée comme suit:

$$\sigma_i(t) = \frac{ER_i(t)}{PL_e(t) - PL_{max}}, PL_e(t) > PL_{max} \quad (6.30)$$

Où:

- $ER_i(t)$ , est le taux d'émission de l'agent  $i$  en temps  $t$ ,
- $PL_e(t)$  est le taux de pollution du polluant  $e$  en temps  $t$  et
- $PL_{max}$  est la valeur maximale allouée pour la concentration de ce polluant.

Ensuite la pénalité pour l'agent  $i$  en temps  $t$ , est calculée comme suit:

$$EcoFactor_i(t) = (1 - \frac{1}{e^{\sigma_i(t)}}) \quad (6.31)$$

Deux stratégies de pénalisation sont définies,

- Pénalisation sans cumule, cela veut dire que l'agent n'est pénalisé que sur sa participation à la situation, donc seulement pour le pas  $t$ , elle est formulée dans (6.31).
- Pénalisation avec cumule, cette stratégie, prend en considération le pas précédent de simulation. Donc un agent si il est pénalisé à  $t-1$ , et il continue à augmenter son taux d'émission il sera pénalisé avec un cumul comme suit :

$$EcoFactor_i(t) = EcoFactor_i(t-1) + (1 - \frac{1}{e^{\delta_i(t)}}) \quad (6.32)$$

La récompense qu'obtient l'agent  $i$  en temps  $t$ , sera après calculer comme suit :

$$M_i(t) = \begin{cases} u(ncp) & \text{if } s_i = 0 \\ u(ncp) \cdot EcoFactor_i(t) & \text{if } s_i = 1 \end{cases} \quad (6.33)$$

Où :  $ncp$  et le nombre d'agent qui choisissent de coopérer, et  $s_i$  est l'action de l'agent en temps  $t$ .

### 3 Conclusion

Dans ce chapitre nous avons présenté un cadre général qui permet de modéliser la pollution de l'air sous-forme d'un système socio-environnemental. Nous avons montré comment nous pouvons combiner entre un modèle de prédiction de concentration de polluant et un algorithme de dispersion. Ces deux derniers sont ensuite intégrés dans un système multi-agent. Le SMA permet de modéliser l'ensemble des êtres humains impliqués dans la gestion des sources de l'émission et aussi dans la gestion de la qualité de l'air. Nous avons proposé deux modèles de prise de décision qui permettent de simuler leurs actions.

Plusieurs méthodes de coopération ont été présentées, dans le but de tester et examiner les meilleures en mesurant leurs efficacités par rapport à la qualité de l'air. Bien sûr ces deux stratégies peuvent être modifiées et personnalisées. De même concernant les mécanismes de pénalisation/ récompense. Ils font partie des paramètres de simulation. Ce qui permettra d'identifier celle qui permet un contrôle efficace de la qualité de l'air

Pour montrer la faisabilité de notre approche, dans le chapitre suivant nous décrivons l'implémentation d'un système de simulation en suivant l'approche proposée.

# Chapitre 7 Expérimentation

---

## 1 Introduction

Ce chapitre décrit les scénarios d'expérimentations avec lesquelles nous avons validé notre système de simulation. Les expérimentations ont été effectuées en utilisant les données concernant la région d'Annaba, dans le nord-est de l'Algérie.

Nous commençons ce chapitre par la présentation d'une brève description sur la région d'Annaba, sa topologie ainsi que quelques statistiques concernant la démographie et l'activité routière. Les sources principales de pollution atmosphérique sont aussi décrites.

Par la suite, nous présentons les propriétés statistiques des données qui ont servi à la validation des expérimentations.

À la fin, nous présentons les différents paramètres et leurs valeurs pour chacun des scénarios de simulation. Principalement trois scénarios de simulation sont présentés : le premier consiste à inspecter l'effet de la coopération sur l'évolution de la concentration des polluants. Le deuxième vise à simuler les périodes des pics des PM10. Le dernier scénario consiste à simuler la gestion d'une crise de pollution de l'air, en incluant aussi des sources de pollution qui ne sont pas contrôlées. Ce genre de sources représentent les sources de pollution sur lesquelles les acteurs n'ont pas de contrôle tel que : les fuites dues à des accidents industriels, les feux de forêts, activités volcaniques, vent de sable etc.

## 2 Région ciblée par les scénarios de simulation

Les expérimentations décrites dans ce chapitre sont basées sur les données qui concernent la région d'Annaba dans le nord-est de l'Algérie. Comme illustré sur la carte géographique dans la Figure 8.1. Annaba est l'une des plus grandes villes en Algérie connue aussi comme étant l'un des pôles de l'industrie Algérienne, surtout dans le domaine de sidérurgie. En fait l'usine d'Arcelor-Metal (ex EL-Hadjar) est l'une des plus grande usine en Afrique avec un taux de production avoisinant le 1 million de tonnes par ans. Malgré les efforts effectués, l'influence de cette usine sur l'entourage et l'écosystème reste de grande rigueur. Dans le processus de production de l'acier beaucoup d'organes chimiques sont utilisés tel que le charbon qui est émis dans l'air lors de l'étape de la mise des composants chimiques dans les grands fours, ce qui cause la dispersion dans l'air de ces composants chimiques, et sous certaines conditions climatiques ces dispersions forment des nuages de pollution qui constituent des pics de pollution de l'air.

Les stations pétrochimiques (Naftal, Asmidal) qui constituent des sources de pollutions les plus critiquées malgré les efforts effectués dans le but de réduire l'émission de ces stations, leur influence reste important sur la qualité de l'air à Annaba.

La troisième source de pollution aussi à prendre en considération c'est le centre d'enfouissement des déchets ménagers situé à Berka Zarga pré de l'agglomération d'El-bouni. Selon (El-Watan) chaque habitant de la ville produit autours de 1.5 klg jusqu'à 2 klg de déchets par jour.



**Figure 7.1 : Carte géographique de Annaba (source :Google ).**

La topographie de la région d'Annaba (comme illustrée dans la figure 8.2) est en forme de cuvette favorisant ainsi la stagnation de l'air et la formation d'inversions de températures. Ces situations causent l'accumulation des polluants et l'élévation des taux de concentration qui en résultent. Les effets des brises de mer, terre, et pente concourent au transport des nuages de polluants. En effet, les nuages de polluants sont entraînés par la brise de terre la nuit vers la mer, et de jour. Ces nuages de polluants retournent sur la ville par effet de brise de mer en longeant la montagne de Edoug. Les nuages tournent sur la ville sous une forme de cercle. Les polluants se déposent lentement par gravité et l'on assiste à une pollution affectant les trois récepteurs (mer, terre, air). L'industrie est l'élément moteur de dégradation de la qualité d'air, cette industrialisation a assurément permis de répondre aux besoins des populations et du pays en produits sidérurgiques, engrais azotés, constructions ferroviaires et autres industries de transformations. A l'inverse, elle a suscité une urbanisation démesurée de la ville avec tous ses corollaires et une pollution de l'atmosphère et des sols, suivi des conséquences néfastes sur le biotope et la société



**Figure 7.2 : Topographie de la région d'Annaba (Source :Google Earth )**

## **2.1 Les données utilisées dans l'étude**

Les données utilisées dans ce travail, proviennent du réseau national de contrôle de qualité de l'air nommé Samaa Safia. Ce dispositif d'évaluation du degré réel de la pollution urbaine et industrielle dans les grandes villes entre dans le cadre de l'accord d'association avec l'Union européenne (UE) qui exige certaines normes internationales à respecter et l'intégration dans la politique atmosphérique euro-méditerranéenne.

En date du 19 octobre 2005, le Conseil de Gouvernement a examiné et adopté un décret exécutif définissant les valeurs limites, les seuils d'alerte et les objectifs de qualité de l'air, en cas de pollution atmosphérique, présenté par le Ministre de l'Aménagement du Territoire et de l'Environnement.

Pris en application de la loi relative à la protection de l'environnement dans le cadre du développement durable, le texte vise à préciser les seuils d'alerte à partir desquels doivent être déclenchées les mesures d'urgence visant à réduire les activités polluantes par les collectivités locales concernées, l'objectif étant de prévenir des situations où une exposition, même de courte durée, qui pourrait entraîner des effets négatifs sur l'homme et sur l'environnement.

Ce réseau possède plusieurs stations de mesure des différents polluants, ces stations sont installées dans les différentes régions d'Algérie. Ces stations sont opérationnelles depuis

2003. Elles sont reliées à un réseau commun au niveau du ministère de tutelle. Un réseau similaire est en cours d'étude à Oran et d'autres villes algériennes. Nous avons focalisé notre travail sur les données provenant des stations implantées dans la wilaya de Annaba constitué de stations installées dans les régions qui enregistrent des taux alarmants de pollution de l'air et qui contiennent des zones industrielles trop proches aux zones urbaines telles que la commune de Sidi Amar et El-Bouni.

Les données que nous avons pu récupérer sont des mesures horaires des années 2003 et 2004 à partir de la station numéro 4 exactement dans la commune de Sidi Amar (voir figure 7.1 ), qui consiste en des matrices pour chaque mois avec quelques plages vides qui sont causées par des pannes et des journées durant lesquelles les stations n'ont pas mesuré ou les données ont été perdus.

### 2.1.1 Les polluants mesurés par Sama Safia

Les stations de Sama Safia surveillent en temps réel la concentration des polluants suivants

<i>Polluant</i>	<i>Description</i>
SO <sub>x</sub> : Mono et Dioxyde de soufre	Proviens généralement des combustions fossiles contenant du soufre (fuel, gasoil). Station de production de l'énergie, Véhicule à moteur.
NO <sub>x</sub> : Mono et Dioxyde d'azote	Proviennent des installations de combustion (gaz naturel) et des véhicules automobiles.
CO <sub>x</sub> : Monoxyde et dioxyde de carbone	Résulte de la combustion incomplète des combustibles fossiles (essence, fuel, charbon, bois).
O <sub>3</sub> : Ozone	Polluant secondaire provenant de la transformation de certains polluants (NO <sub>2</sub> et hydrocarbure) sous l'action du soleil.
PM10 Le particules( particulate matter)	Poussière fines en suspension, émis par l'activité naturelles et humaines (moteur diesel, industrie etc).

**Tableau 7.1 : Les polluants mesurés par Sama Safia.**

### 2.1.2 Paramètres climatiques

En plus des polluants la base de données contient aussi des mesures des paramètres climatiques, car les mesures des concentrations des polluants seules ne permettent pas de donner

des indices sur la qualité de l'air. Les paramètres climatiques mesurés sont décrits dans le tableau suivant:

Paramètres	Description
Humidité	L'humidité influence énormément sur la capacité des composants chimiques de l'air à s'interagir et de se déplacer d'une zone à une autre.
Température	La température est l'un des paramètres qui donne l'énergie nécessaire aux réactions photochimiques qui résultent des polluants secondaires
Vitesse du vent	Le vent déplace les masses de l'air d'une zone géographique à une autre. Avec une vitesse de vent faible les polluants se cumulent dans la même zone ce qui créera les périodes de pics.

**Tableau 7.2 : Les paramètres climatiques mesurés par Sama Safia.**

	<i>Temperature</i>	<i>Humidité</i>	<i>Vitesse du vent</i>
	(°C)	(%)	(nœuds)
<b>Janvier</b>	12,285	75,973	6,8733
<b>Février</b>	11,891	77,938	7,5507
<b>Mars</b>	13,318	72,782	7,612
<b>Avril</b>	15,327	72,725	7,5625
<b>Mai</b>	19,144	74,219	7,1205
<b>Juin</b>	22,783	72,092	7,93
<b>Juillet</b>	24,786	71,135	8,4846
<b>Aout</b>	26,158	71,791	8,0258
<b>Septembre</b>	23,476	72,847	7,4396
<b>Octobre</b>	19,931	73,097	6,8879
<b>Novembre</b>	15,521	75,514	7,5688
<b>Décembre</b>	12,925	75,669	7,7968

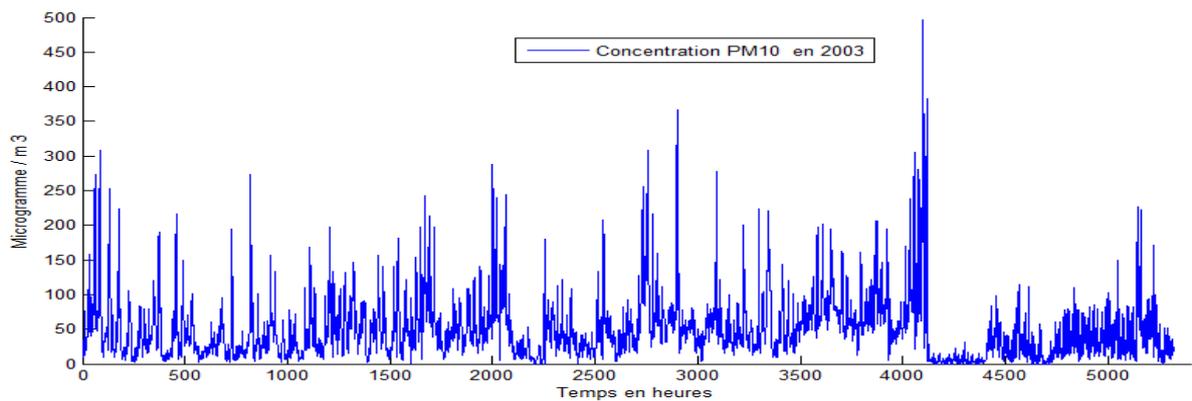
**Tableau 7.3 Paramètres météorologiques moyens dans la région d'Annaba.**

### 2.2.3 Propriétés statistiques des données

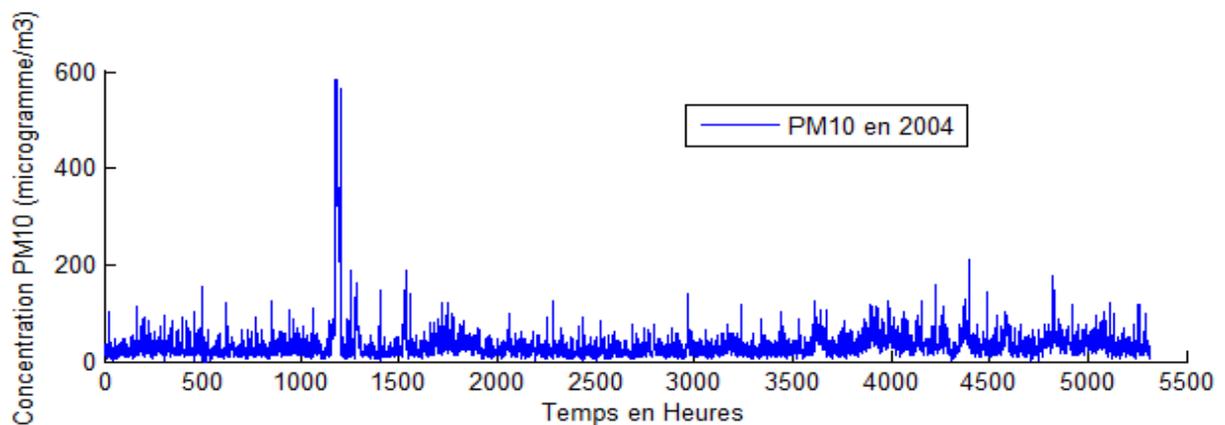
L'analyse statistique les données que nous avons pu récupérer a permis de mieux comprendre le phénomène de pollution de l'air dans la région d'Annaba, ce qui nous permet de créer et concevoir des modèles le plus adéquates.

*PM10*

Ce sont des petites particules contiennent une multitude de substances dont des métaux, des acides, du carbone et des hydrocarbures aromatiques polycycliques. Certaines articles sont émises dans les gaz d'échappement des véhicules tandis que d'autres sont formées dans l'atmosphère par des réactions chimiques entre les polluants émis dans les gaz d'échappement. Les particules fines aggravent les symptômes des personnes atteintes de maladies respiratoires ou cardiovasculaires.les deux figures 8.3 et 8.4 montrent les propriétés de PM10 en 2003 et en 2004



**Figure 7.3 : Concentration PM10 en 2003.**



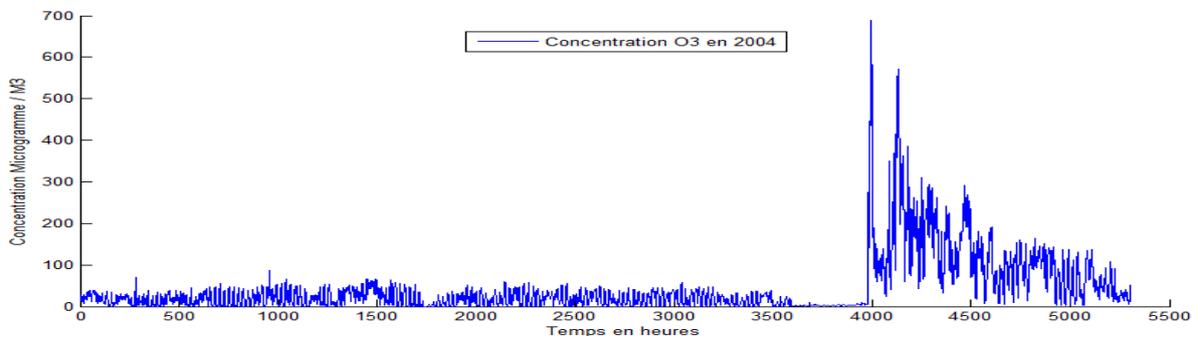
**Figure 7.4 : Concentration P10 en 2004.**

<b>PM10</b>	<b>2003</b>	<b>2004</b>
Moyenne (microgramme /m3)	49.0995	27.6830
Covariance	43.0445	26.3846
Valeur maximum (microgramme /m3)	496	508
Valeur minimum (microgramme / m3)	1.0200	13.0200

**Tableau 7.4 : Propriétés statistiques de mesure de concentration de PM10.**

### Ozone ( $O_3$ )

Ce polluant n'a pas été pris en considération dans les mesures de l'année 2003, seules les données de 2004 sont disponibles, comme le montre la Figure 8.5.



**Figure 7.5 : Concentration O3 en 2004.**

<b>Ozone</b>	<b>2004</b>
Moyenne (microgramme / m3)	42.2740
Covariance	64.5807
Valeur maximum (microgramme /m3)	688
Valeur minimum (microgramme / m3)	3,000

**Tableau 7.5 : Propriétés statistiques de mesure de concentration du l'Ozone.**

### $NO_x$

Ce vecteur englobe les mesures du  $NO_2$  et du  $NO$ , comme montré dans les Figure 8.6 et 8.7.

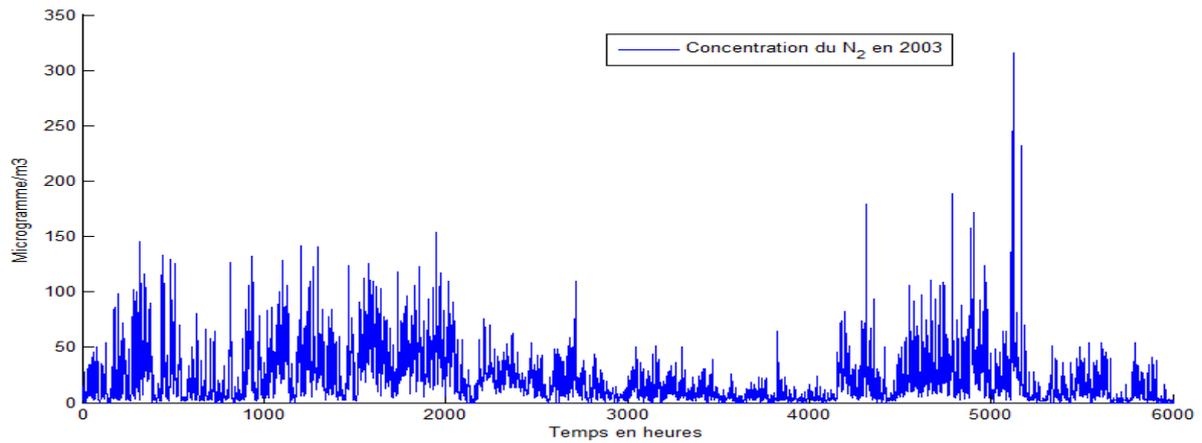


Figure 7.6 : Concentration NO<sub>x</sub> en 2003.

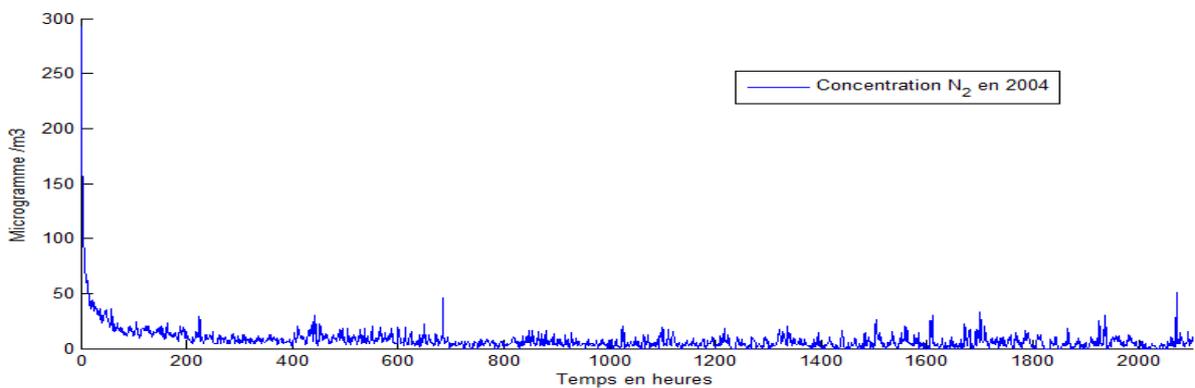


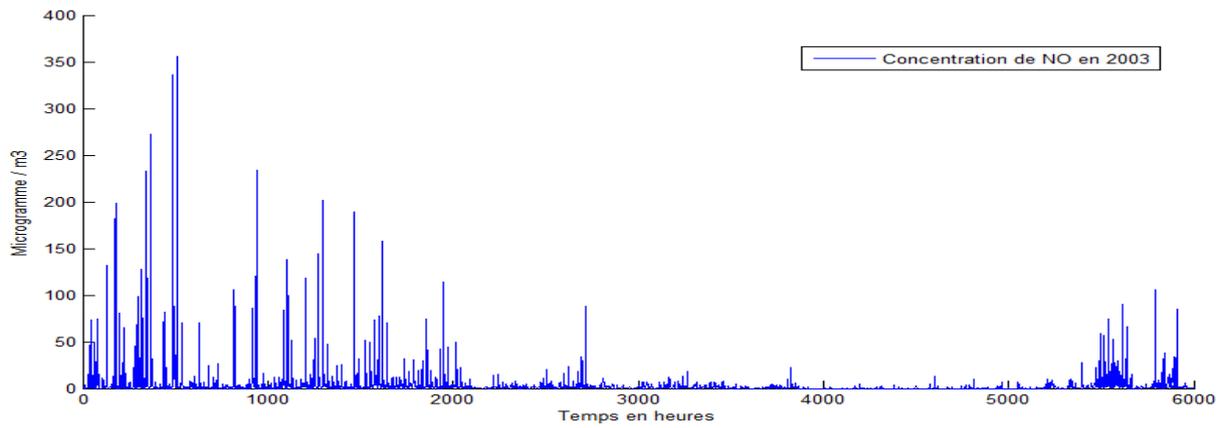
Figure 7.7 : Concentration NO<sub>x</sub> en 2004.

NO <sub>x</sub>	2003	2004
Moyenne	20.4698	7.2341
Covariance	22.8553	11.0005
Valeur maximum	316	293
Valeur minimum	2,0000	1,0000

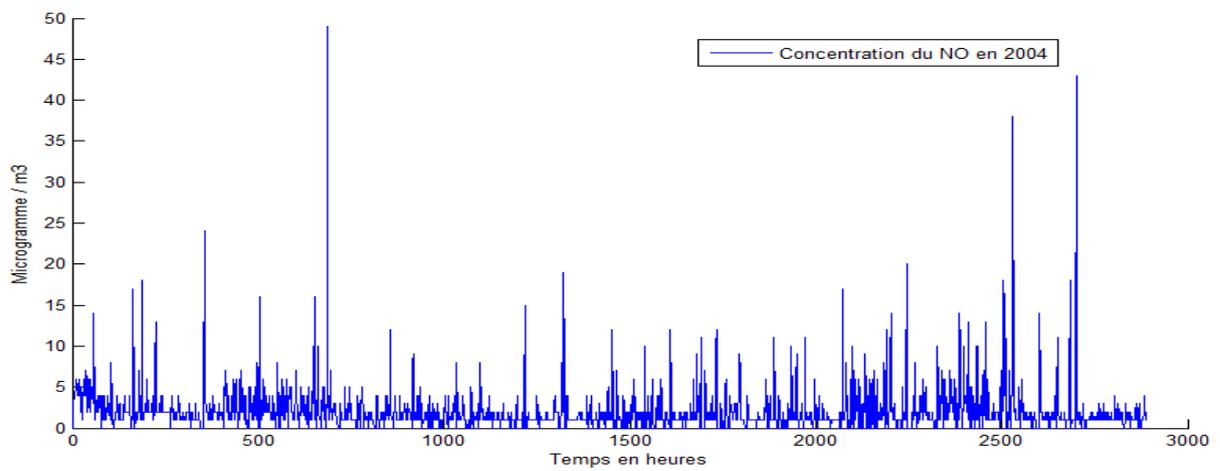
Tableau 7.6 : Propriétés statistiques de mesure de concentration du NO<sub>x</sub>.

*NO*

Les deux Figures 8.8 et 8.9 montent la série temporelle du NO.



**Figure 7.8 : Concentration du NO en 2003.**



**Figure 7.9 : Concentration du NO en 2004.**

NO	2003	2004
Moyenne	3.7251	2.0768
Covariance	15.0672	2.6069
Valeur maximum	356	49
Valeur minimum	1,0000	1,0000

**Tableau 7.10 Propriétés statistiques de mesure de concentration du NO.**

## $CO_2$

Pour ce type polluant on n'a que les données du 2003, montré dans la Figure 8.10.

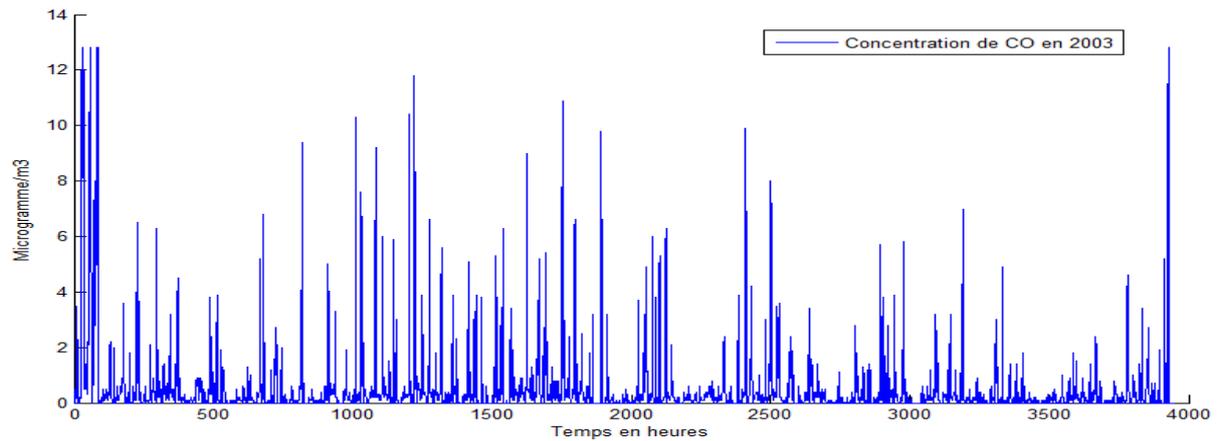


Figure 7.10 : Concentration du CO<sub>2</sub> en 2003.

CO <sub>2</sub>	2003
Moyenne	0.5279
Covariance	1.3287
Valeur maximum	12.8000
Valeur minimum	1,0000

Tableau 7.11 : Propriétés statistiques de mesure de concentration du CO<sub>2</sub>.

## NO<sub>2</sub>

Pour ce polluant on n'a que les données du 2003, montré dans la Figure 8.11.

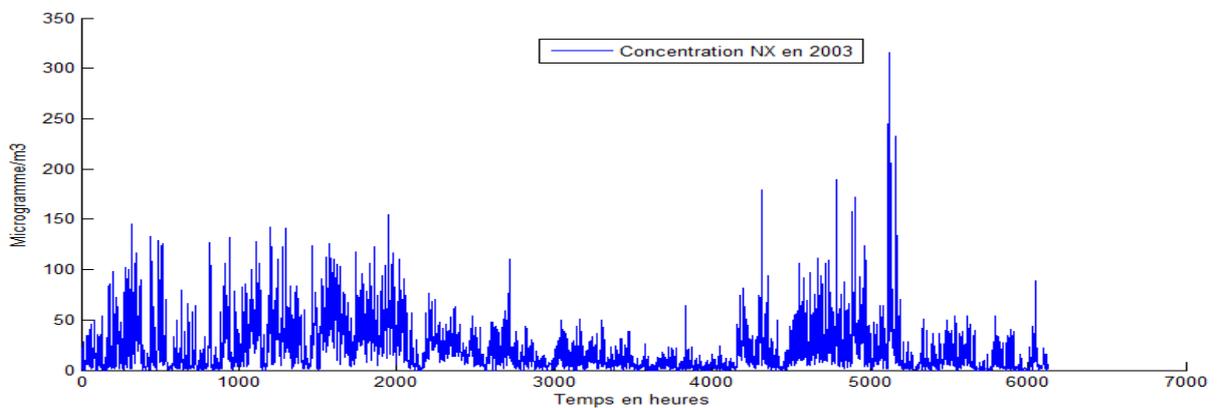


Figure 7.12 : Concentration du NO<sub>2</sub> en 2003.

NO <sub>2</sub>	2003
Moyenne	14.5021
Covariance	25.0133
Valeur maximum	435
Valeur minimum	1,0000

Tableau 7.12 : Propriétés statistiques de mesure de concentration du NO<sub>2</sub>.

## Humidité

Ce paramètre est marqué par la variation entre jour et nuit, comme montré dans les deux figures 8.12 et 8.13.

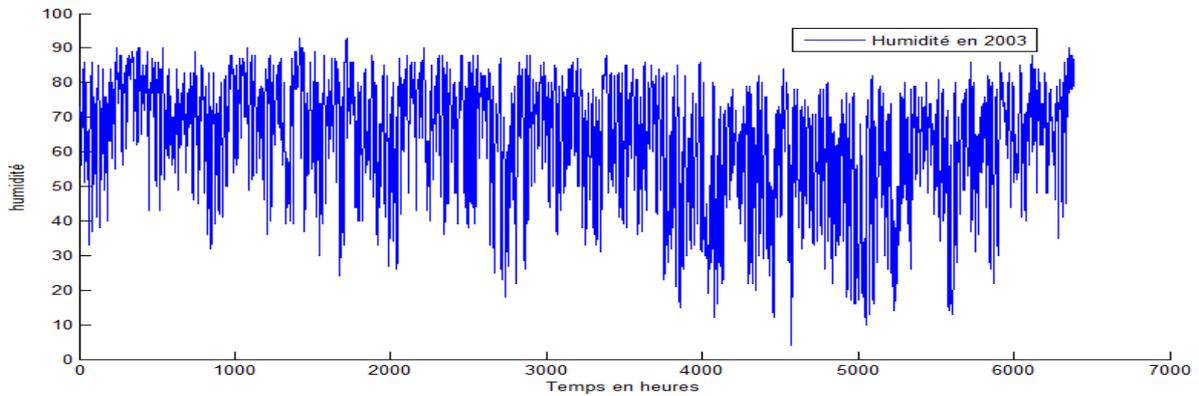


Figure 7.13 : Humidité en 2003.

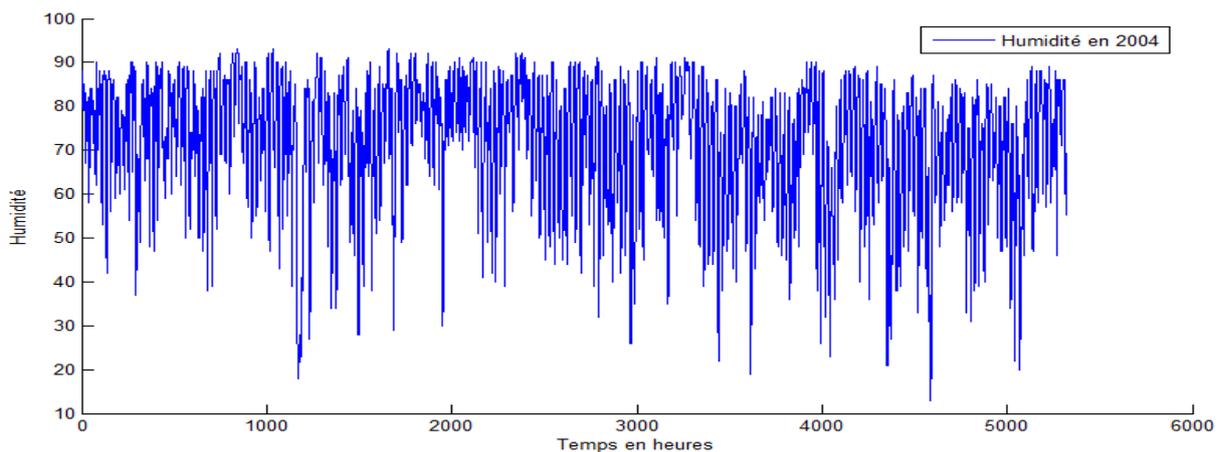


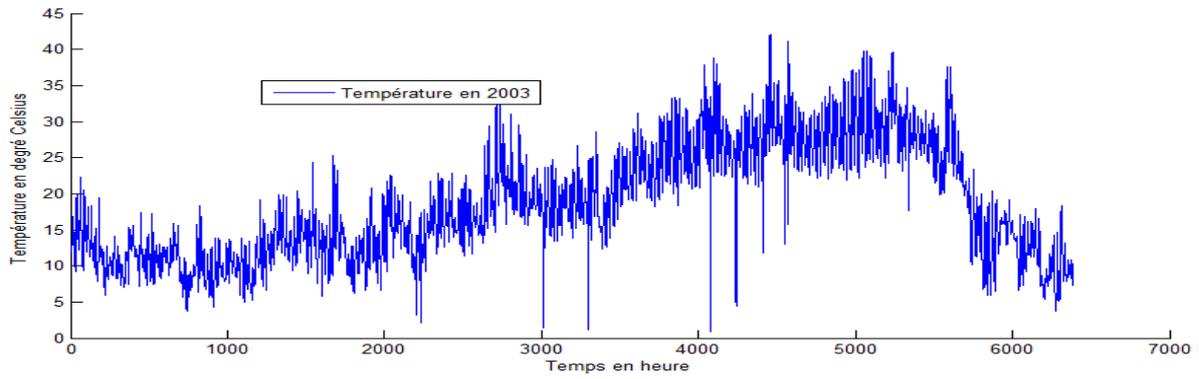
Figure 7.14 : Humidité en 2004.

Humidité	2003	2004
Moyenne	63.5222	71.9237
Covariance	16.5087	14.3361
Valeur maximum	93	93
Valeur minimum	4	13

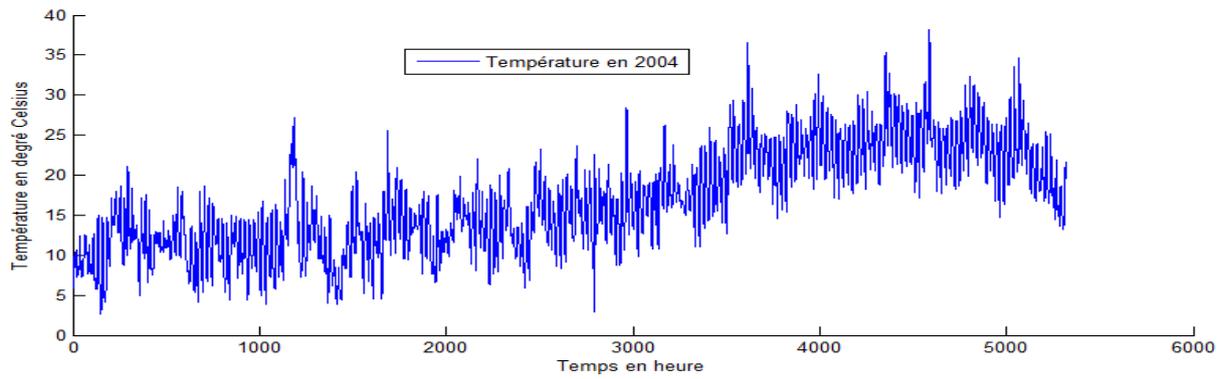
Tableau 7.13 : Propriétés statistiques de mesure de l'humidité.

## La température

Les variations entre été et en hiver est un aspect marquant ce paramètre, comme le montrent les deux figures 8.14, 8.15.



**Figure 7.15 : Température en 2003.**



**Figure 7.16 : Température en 2004.**

Température	2003	2004
Moyenne	18.6942	16.8210
Covariance	16.5087	6.3003
Valeur maximum	663	38.2000
Valeur minimum	1	2.7000

**Tableau 7.11 : Propriétés statistiques de mesure de température.**

### *La vitesse du vent*

Comme le montrent les deux Figures 8.16 et 8.17 la variation entre jour et nuit est l'une des propriétés majeures de ce paramètre.

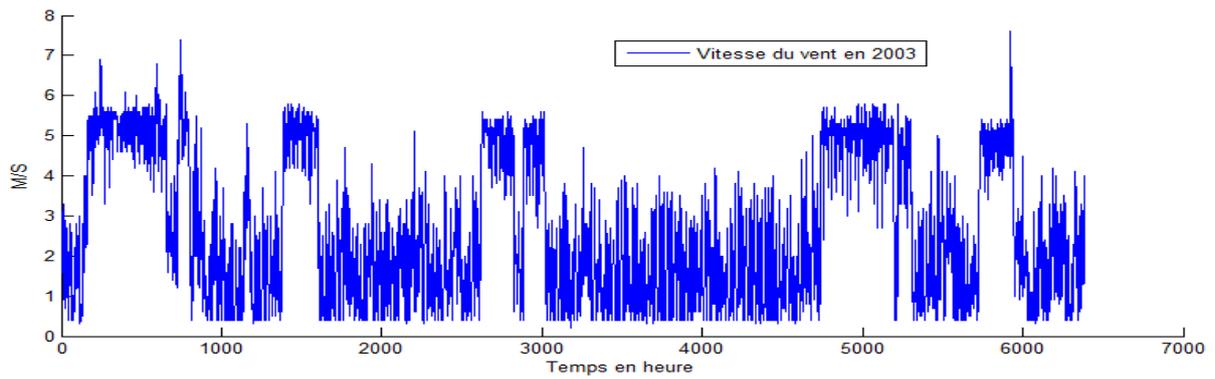


Figure 7.17 : Vitesse du vent en 2003.

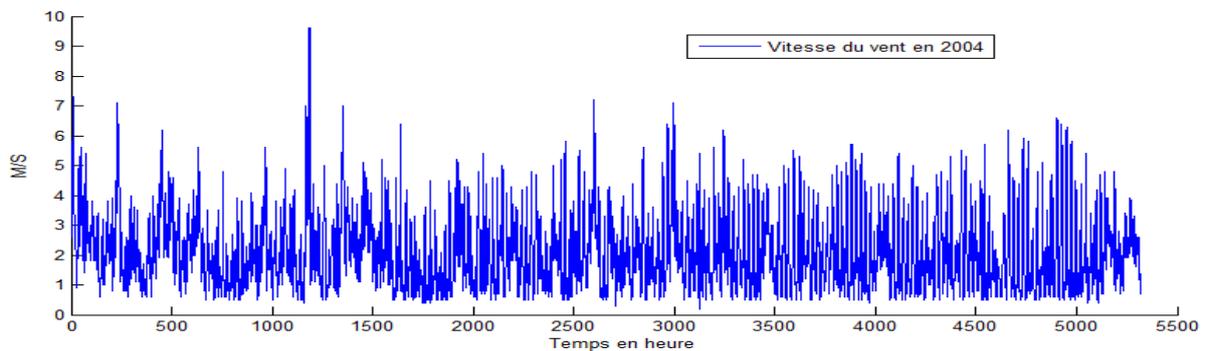


Figure 7.18 : Vitesse du vent en 2004.

Vitesse du vent	2003	2004
Moyenne	2.6599	2.1283
Covariance	1.7814	1.2749
Valeur maximum	7.6000	9.6000
Valeur minimum	0.2000	0.2000

Tableau 7.12 : Propriétés statistiques de mesure de la vitesse du vent.

### 3 Les scenarios de simulation

Un scénario de simulation consiste à définir les valeurs de chacun des paramètres de la simulation. Comme présenté dans le tableau 8.12, ces paramètres sont regroupés en trois catégories :

#### 3.1 Les paramètres de l'émission:

Cette étape consiste à définir les paramètres liés à l'émission, c'est-à-dire, le nombre de source de pollution pour les trois polluants.

- Nombre de source émettant le PM10, ce paramètre représente le nombre des sources de type point qui produisent le PM10.

- Nombre de source émettant le  $\text{SO}_x$ , ce paramètre représente le nombre de source de type point qui produit le  $\text{SO}_x$ , la valeur pour ce paramètre est une valeur naturelle qui peut être nulle.
- Le nombre de sources émettant le  $\text{NO}_x$ , ce paramètre représente le nombre de sources de type point qui produisent le  $\text{NO}_x$ , la valeur pour ce paramètre est une valeur naturelle.
- Le niveau maximum de l'émission, ce paramètre représente le taux maximum qu'une source peut atteindre, il est exprimé en microgramme/ $\text{m}^3$ .
- Le nombre de sources de polluant non contrôlées : ceci définit le nombre de sources qui ne seront pas modérées par un agent. Elles représentent les sources naturelles de pollution, et aussi elles peuvent représenter des fuites industrielles non-contrôlées
- Le modèle de dispersion utilisé : précis quel modèle de dispersion sera utilisé dans la simulation, la version actuelle prend en considération deux types de modèle de dispersion : Le modèle GPD (Modèles Gaussien de dispersion), et le modèle système de particule.

### **3.2 Les paramètres qui concernent l'objectif de la simulation :**

Dans un scénario de simulation nous devons préciser l'objectif en termes de qualité de l'air que nous espérons atteindre. Cet indice de qualité de l'air a son équivalent en termes de concentrations des polluants que nous devons maintenir.

- Global Air Quality Index : ceci est l'indice de qualité de l'air, il peut avoir des valeurs entre [1,2,3,4,5]. En fixant ce paramètre les valeurs des autres paramètres seront automatiquement initialisées.
- PM10 goal level : ce paramètre définit le niveau de concentration cherché, le taux doit être inférieur ou égal à cette valeur.
- $\text{SO}_x$  goal level : le niveau de concentration du  $\text{SO}_x$  qui ne doit pas être dépassé.
- $\text{NO}_x$  goal level : le niveau de concentration cherché pendant la simulation.
- $\text{O}_3$  goal level : le niveau de concentration fixé comme objectif pour le  $\text{O}_3$ .

### **3.3 Les paramètres liés à la coopération des agents :**

- Le taux de coopération initial : c'est le taux des agents qui choisissent de coopérer par rapport au nombre total des agents contrôlant des sources de pollution, par défaut initialisé à 50%.
- Méthode de pénalisation et stratégie de coopération.
- $K$  le nombre de pas de mémoire pour les agents. Ce nombre représente le nombre de pas que l'agent doit mémoriser, c'est-à-dire, mémoriser ses  $K$  dernières actions, ses derniers gains, et les gains de ses voisins.
- L'action de départ, les probabilités de départ, les gains de départ, tous ces paramètres sont initialisés d'une façon aléatoire.

### **3.4 Les paramètres liés à l'environnement :**

- Le nombre de cubes qui forment l'environnement (voir le Chapitre 6 pour plus de détails concernant la représentation de l'environnement)
- La classe de perturbation, ce paramètre permet de définir le niveau de perturbation atmosphérique. Il est utilisé dans le modèle de dispersion.
- Les valeurs initiales pour les paramètres climatiques, température, vitesse du vent, humidité et précipitation (pluie).

### **3.5 Les paramètres liés au temps :**

Le nombre total de pas de la simulation : ce paramètre permet de définir l'horizon à long terme de la simulation, il est lié à la taille des bases de données disponibles, dans notre cas nous ne pouvons pas aller plus d'une année. Etant donné que nous disposons que de deux années de mesures ; La première est utilisée pour l'apprentissage et la dixième pour la validation.

Le nombre d'heures pour chaque pas de simulation : chaque pas de simulation peut représenter plusieurs heures d'heures réelles. Ce paramètre permet de définir la représentation, par exemple un pas de simulation peut être égal à 4 heures réelles.

L'horizon de prédiction : ce paramètre définit le nombre d'heures en avance de la prédiction, utilisée pour estimer la pollution et calculer les pénalités et les gains.

### 3.6 Le scénario de simulation numéro 1 et 2 :

Nom du paramètre	Valeur
<b><i>Paramètre sur les sources d'émission</i></b>	
Nombre de source émettant le PM10	100
Nombre de source émettant le SO <sub>x</sub>	100
Nombre de source émettant le NO <sub>x</sub>	100
Le niveau maximum d'émission	2000 (gram/heure).
Le but pour la concentration du PM10	20 µ gram/m <sup>3</sup>
Le but pour la concentration du SO <sub>x</sub>	30 µ gram/m <sup>3</sup>
Le but pour la concentration du NO <sub>x</sub>	45 µ gram/m <sup>3</sup>
Le but pour la concentration du O <sub>3</sub>	45 µ gram/m <sup>3</sup>
Nombre de pas dans la mémoire (K)	4 pas
Le taux initial de coopération	0.5
<b><i>Paramètres liés à l'environnement</i></b>	
Nombre de cubes	20
Température à t=0	12.7 (°C)
Humidité à t=0	71.0 %
Vitesse du vent t=0	2.4 m/s
Concentration du PM10 à t=0	13.0 µ gram/m <sup>3</sup>
Concentration du SO <sub>x</sub> à t=0	17.0 µ gram/m <sup>3</sup>
Concentration du NO <sub>x</sub> à t=0	2.0 µ gram /m <sup>3</sup>
Concentration du l'O <sub>3</sub> à t=0	29.0 µgram /m <sup>3</sup>
La qualité de l'air à t=0	2 ( Good)
Temps total de la simulation	4900 heures
Nombre d'heur pour chaque pas de simulation	1 pas = 2 heures
Le nombre d'heur en avance de la prédiction	2 heures en avance

**Tableau 7.13 : les paramètres de simulation du scénario.**

**3.7 Le scénario de simulation numéro 3 :**

<b>Parameter Name</b>	<b>Value</b>
<b>Polluting activities and Policy parameters</b>	
Number of controlled sources	240 (80 PM10, 80 NO <sub>x</sub> and 80 SO <sub>x</sub> )
Max emission rate	2000 gram/hour
Number of uncontrolled source	15 (5 PM10, 5 NO <sub>x</sub> and 5 SO <sub>x</sub> )
Max emission rate (for uncontrolled )	5000 gram/hour
Goal PM10 level	70 μ gram/m <sup>3</sup>
Goal SO <sub>x</sub> level	60 μ gram/m <sup>3</sup>
Goal NO <sub>x</sub>	50 μ gram/m <sup>3</sup>
Number of memory steps (K)	4 steps
Initial proportion of cooperating agents	0.5
<b>Environment parameters</b>	
Number of boxes	20
Temperature at t=0	12.7 (°C)
Humidity at t=0	71.0 %
Wind Speed t=0	2.4 m/s
PM10 at t=0	13.0 μ gram/m <sup>3</sup>
Air Quality at t=0	2 ( <b>Good</b> )
Total simulation time	4900 hours
Simulation step	1 step = 2 hours
Prediction horizon	2 hours advance

**Tableau 7.14 les paramètres de simulation du deuxième scénario.**

# Chapitre 8 : Résultats et discussions

---

## 1 Introduction :

Ce chapitre présente et discute les résultats des scénarios de simulation décrits dans le chapitre précédent. La discussion est principalement axée sur l'efficacité de la coopération et ses effets sur l'évolution de la pollution de l'air et la qualité de l'air. Pour cela, trois scénarios de simulation sont exécutés :

1. Le premier consiste à étudier la question : « comment le choix d'une stratégie de coopération influence l'évolution de la concentration des polluants ».
2. Le deuxième permet de tester la capacité des stratégies de coopération à réduire les pics de pollution et aider à gérer une crise de pollution.
3. Le troisième scénario permet d'inclure parmi les sources de pollution, celles qui ne sont pas contrôlées, dues à des incidents industriels ou bien à des événements naturels (ex. feux de forêts). Le scénario a comme objectif de simuler la gestion d'une situation de crise de pollution de l'air, qui implique des sources qui ne sont pas contrôlées.

Le premier scénario a été exécuté **16** fois pour chaque stratégie, ce qui nous a produit 60 résultats (16 exécutions pour les 5 stratégies). Ces exécutions nous ont permis d'observer les variations possibles entre les résultats que donne le simulateur à chaque exécution. Les résultats de ses exécutions sont résumés et discutés dans la partie résultat.

L'exécution d'un scénario se fait en suivant les étapes suivantes :

**A) Prétraitement des données :** Cette étape permet aux utilisateurs d'exploiter des données brutes, c'est-à-dire, des données issues directement des stations de mesures, elles peuvent contenir des mesures fausses ou vides, qui sont dues aux pannes des capteurs. Elle remplace les données manquantes par la moyenne de cellules voisines ( $m$  valeurs précédentes et  $m$  valeurs suivantes), ce qui permet de garder la série temporelle continue et correcte, ceci a un effet très important sur la capacité de généralisation des modèles de prédiction (RNA). Les résultats de cette étape peuvent servir ensuite comme ensemble d'apprentissage et de validation. Le sous-ensemble de validation sera alors utilisé comme entrée pour la simulation.

**B) Apprentissage :**

Pour que notre simulateur soit configuré avec les données disponibles, une étape d'apprentissage est nécessaire. Grâce à une fonctionnalité, le simulateur donne la main pour créer des réseaux RBF et d'appliquer une étape d'apprentissage ainsi que de validation. Une fois l'utilisateur satisfait des performances des modèles RBF, il peut alors les utiliser lors de la simulation. Il est à noter, que les performances des modèles RBF ont un rôle clé dans la simulation.

En utilisant les données décrites dans le chapitre précédent, nous avons commencé par la conception des réseaux de neurones artificiels qui permettent de prédire la concentration des polluants. Le tableau 9.1 présente les topologies de chacun des réseaux RBF utilisés, elle est notée sous la forme : [nombre de neurones d'entré-nombre de neurone de la couche RBF].

Topologie	Polluants
[10-320]	PM10
[10-180]	O <sub>3</sub>
[10- 105]	NO <sub>x</sub>
[10-90]	SO <sub>2</sub>
[10-45]	CO <sub>x</sub>

**Tableau 8.1. Topologie des réseaux RBF conçus.**

Le tableau 9.2, présente les performances (les erreurs de validation) des modèles RBF conçus, c'est-à-dire la distance entre les prédictions des modèles RFB et les valeurs réelles mesurées qui sont tirées du sous-ensemble de validation.

Modèles	Performances des Réseaux RBF ( RMSE)
PM10	16.1946 micro-gram/m <sup>3</sup>
SO <sub>x</sub>	3.1618 micro-gram/m <sup>3</sup>
NO <sub>x</sub>	9.7277 micro-gram/m <sup>3</sup>
O <sub>3</sub>	39.8239 micro-gram/m <sup>3</sup>
CO <sub>x</sub>	0.1223 micro-gram /m <sub>3</sub>

**Table 8.2: Erreur de validation des réseaux RBF.**

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - R_i)^2}{N}} \quad (8.1).$$

Les performances des modèles (Erreur Moyenne quadratique, voir équation 8.1) sont calculées en utilisant une année de mesures, les paramètres climatiques ainsi que la concentration du polluant.

**C) Paramétrage de la simulation :** Cette étape consiste à saisir les paramètres du scénario de la simulation, parmi ces paramètres : les caractéristiques des sources de pollution, le nombre de sources, leurs taux d'émission ainsi que les paramètres liés à la gestion tels que la méthode de pénalisation et la stratégie choisie pour la coopération.

**D) Exécution :** Cette étape consiste à lancer l'exécution de la simulation et observer son avancement. Selon les besoins de l'utilisateur, des paramètres peuvent être modifiés lors de l'exécution (arrêter ou ajouter des sources de pollution), ceci donne la possibilité d'inspecter l'influence d'un paramètre.

**E) Exportation et visualisation des résultats :** le simulateur permet de voir pendant l'exécution, les résultats de la simulation sous forme de graphique, ces résultats peuvent être exportés afin qu'ils puissent être analysés.

## 2 Discussion des résultats

**2.1 Les sorties de la simulation** tout système de simulation doit définir un ensemble de sorties qui représentent les variables étudiées. Dans notre simulation nous nous intéressons principalement à la concentration des polluants ainsi que la qualité de l'air. Le tableau suivant contient les principales sorties que présente le simulateur avec leurs descriptions.

Sortie	Unité de mesure	Description
Concentration-PM10	Microgramme / mètre cube	Représente la concentration des PM10 dans la région concernée par l'étude, elle est calculée selon la prédiction du réseau RBF en utilisant les informations de dispersion et les paramètres climatiques
Concentration-SO <sub>x</sub>	Microgramme / mètre cube	La concentration des SO <sub>x</sub> dans la région étudiée. Calculée en utilisant la prédiction du réseau RBF et l'algorithme de dispersion
Concentration-NO <sub>x</sub>	Microgramme/mètre cube	La concentration des NO <sub>x</sub> dans la région, calculée en utilisant les prédictions du réseau RBF et les informations de dispersion
Concentration-O <sub>3</sub>	Microgramme/mètre cube	La concentration de l'Ozone, elle est calculée en utilisant les prédictions faites par le réseau RBF. Ce dernier reçoit comme paramètres les informations de dispersion du SO <sub>x</sub> et NO <sub>x</sub> ainsi que le CO <sub>x</sub>
Air-Quality	Entre [0, 1, 2, 3, 4,5]	Index de qualité de l'air, calculé en utilisant la prédiction faite par le réseau MLP, ce dernier reçoit comme paramètres les prédictions faites par les autres réseaux RBF.
Coopération-PM10	[0,1]	Pourcentage des agents qui coopèrent (réduisent leurs taux d'émission), ce taux est calculé à chaque pas de simulation en utilisant le nombre total des agents qui produisent les PM10.
Coopération-SO <sub>x</sub>	[0,1]	Pourcentage des agents qui coopèrent et qui gèrent des sources qui produisent le SO <sub>x</sub>

Sortie	Unité de mesure	Description
Cooperation-NO <sub>x</sub>	[0,1]	Pourcentage des agents qui coopèrent et qui produisent le NO <sub>x</sub>
Cooperation-CO <sub>x</sub>	[0,1]	Pourcentage des agents qui coopèrent et qui produisent les CO <sub>x</sub>
Global-Cooperation	[0,1]	Pourcentage des agents qui coopèrent par rapport à la taille totale de la population d'agent.

**Tableau 8.3 : Les sorties de la simulation.**

## 2.2 Le premier scénario

Le scénario a été exécuté 80 fois, 16 exécutions pour chacune des stratégies :

1. EG-CP (Evolutionary Game with Cumulative Penalty),
2. EG-NCP (Evolutionary Game with No Cumulative Penalties),
3. EG-NP (Evolutionary Game with No Penalty),
4. CS (Centralized Strategy),
5. NC (No-Cooperation). Cette dernière est présentée pour un but de comparaison, elle correspond au cas où les agents ne coopèrent pas et qu'aucune méthode de pénalisation n'est utilisée.

Les tableaux suivants (8.4, 8.5 et 8.6) présentent respectivement les valeurs minimales, maximales et les moyennes des sorties des 60 simulations exécuté.

Stratégie	Coopération					Concentration					AQI
	Global	PM10	CO <sub>x</sub>	SO <sub>x</sub>	NO <sub>x</sub>	PM10	CO <sub>x</sub>	NO <sub>x</sub>	SO <sub>x</sub>	O <sub>3</sub>	
Min CS	N/A	N/A	N/A	N/A	N/A	36,8607	6,7801	0,6740	23,4657	14,2207	1
Min EG-CP	0,3875	0,5500	0,475	0,475	0,05	0,9666	0,9166	27,6838	10,0560	19,1538	2
Min EG-NCP	0,4375	0,5625	0,425	0,5	0,0375	68,3463	8,4287	24,4983	37,3475	15,4256	2
Min EG-NP	0,3416	0,4	0,4333	0,43333	0	24,1697	2,2239	17,4791	12,9495	7,2860	1
Min NC	N/A	N/A	N/A	N/A	N/A	24,8520	19,0440	3,5032	20,1690	10,3954	1
Min	0,3416	0,4	0,425	0,43333	0	0,9666	0,9166	0,6740	10,0560	7,2860	1

**Tableau : 8.4: Les valeurs minimales des 60 sorties de la simulation.**

Stratégie	Coopération					Concentration					AQI
	Global	PM10	CO <sub>x</sub>	SO <sub>x</sub>	NO <sub>x</sub>	PM10	CO <sub>x</sub>	NO <sub>x</sub>	SO <sub>x</sub>	O <sub>3</sub>	
Max CS	N/A	N/A	N/A	N/A	N/A	128,5829	90,3082	10,7818	150,8259	213,0706	5
Max EG-CP	0,975	1,4666	1	1	1	124,0358	10,7592	83,8486	144,3961	87,5221	4
Max EG-NCP	0,9312	1	1	0,875	0,925	127,5051	10,7562	85,7510	147,5220	89,9573	4
Max EG-NP	0,96	1	0,98	0,98	0,9	206,5064	10,8119	149,3077	176,6680	176,9465	5
Max NC	N/A	N/A	N/A	N/A	N/A	238,9129	180,8929	10,8598	183,5566	229,6435	5
Max	0,975	1,4666	1	1	1	238,9129	180,8929	149,3077	183,5566	229,6435	5

**Tableau 8.5 Les valeurs maximales des 60 sorties de la simulation.**

Stratégie	Coopération					Concentration					AQI
	Global	PM10	CO <sub>x</sub>	SO <sub>x</sub>	NO <sub>x</sub>	PM10	CO <sub>x</sub>	NO <sub>x</sub>	SO <sub>x</sub>	O <sub>3</sub>	
Moyenne CS	N/A	N/A	N/A	N/A	N/a	69,0182	47,9764	6,8843	74,7905	89,1735	2,9285
Moyenne EG-CP	0,9059	1,0482	0,9327	0,9372	0,8607	54,4111	7,9187	52,4221	63,3728	46,6125	2,1875
Moyenne EG-NCP	0,6256	0,73	0,6466	0,615	0,5108	98,0940	10,0299	50,7484	84,4225	47,2108	3
Moyenne EG-NP	0,5490	0,6164	0,5814	0,5856	0,4127	95,9397	9,5814	61,7347	97,7497	61,4012	3,2222
Moyenne NC	N/A	N/A	N/A	N/A	N/A	128,8826	127,3806	9,6817	84,4854	95,3222	3,6923
Moyenne	0,6957	0,8029	0,72494	0,7214	0,6089	85,2163	44,8540	33,1068	78,5387	70,1299	2,9482

**Tableau 8.6 : Les valeurs moyennes des 60 sorties des simulations.**

Les résultats les plus explicatifs sont présentés, pour le cas de la stratégie CS (centralisée) et NC (No-Cooperative) le simulateur donne les mêmes sorties pour chaque exécution. Pour les stratégies basées NPPD le simulateur présente des différences légères entre les exécutions, surtout concernant le taux de coopération entre les agents. Ces différences sont dues aux valeurs aléatoires avec lesquelles quelques paramètres sont initialisés : Le gain des voisins, la première action choisie, les poids des gains, les derniers gains des dernières k pas de simulation.

La comparaison entre les stratégies se fait en se basant sur le nombre d'occurrence de l'index de la qualité de l'air. Comme illustré dans la figure 8.1, nous pouvons voir pour combien de fois l'index de la qualité de l'air a atteint la valeur 1 (très bonne), ceci permet de montrer clairement la meilleure stratégie, à savoir celle qui contient le plus de valeur bonnes d'index de qualité de l'air et moins de mauvaises indexes.

La figure 8.2, permet d'observer l'évolution de l'index de qualité de l'air pendant le temps d'exécution, ceci aide à bien voir l'efficacité de chacune des stratégies sélectionnées.

En utilisant la stratégie centralisée CS permet d'obtenir les meilleurs résultats. En utilisant cette stratégie l'index de qualité de l'air évolue rapidement de Mauvais (Bad) à moyenne (Average) après à Bonne (Good) jusqu'à ce qu'il arrive à la fin et se stabilise dans l'index très bonne (Very-Good), il est à noter que l'objectif du scénario est de stabiliser l'index dans la valeur 0 qui signifie Très Bonne (Very-Good).

Pour les stratégies à base NPPD, la stratégie EG-CP, qui utilise une méthode de pénalisation cumulée, elle a permis de déplacer l'index de la qualité de l'air de mauvais (Bad) à l'index moyenne (Average). Après que le jeu arrive à un état d'équilibre, l'index de qualité de l'air se stabilise dans l'index Bien (Good) et n'atteint jamais la valeur Très-Bien. En utilisant la stratégie EG-NCP, qui utilise une méthode de pénalisation non cumulative, la simulation permet de voir que l'index se déplace de la valeur Mauvais vers la valeur Moyenne, et ne s'améliore plus après. Pour la stratégie EG-NP, qui n'inclut pas une méthode de pénalisation, l'index de la qualité de l'air se stabilise dans la valeur Mauvaise (Bad) est ne s'améliore pas mais aussi ne se dégrade pas.

Afin de montrer la différence entre les résultats avec une stratégie de coopération et sans coopération. La stratégie NC est basée sur le fait que les agents ne tiennent pas compte de la qualité de l'air et ne coopèrent pas. Comme montré dans la figure 8.2, en utilisant cette stratégie, la qualité de l'air passe de l'index Mauvaise vers l'index Très-Mauvaise. Malgré cela l'index ne se

stabilise pas tout le temps dans l'index **Très-mauvais** et présente des oscillations, ceci est expliqué par l'influence des paramètres climatiques qui sont utilisés pour exécuter la simulation et pour la conception des modèles de prédiction. Les conditions climatiques permettent parfois l'amélioration la qualité de l'air en accélérant le processus de destruction des polluants.

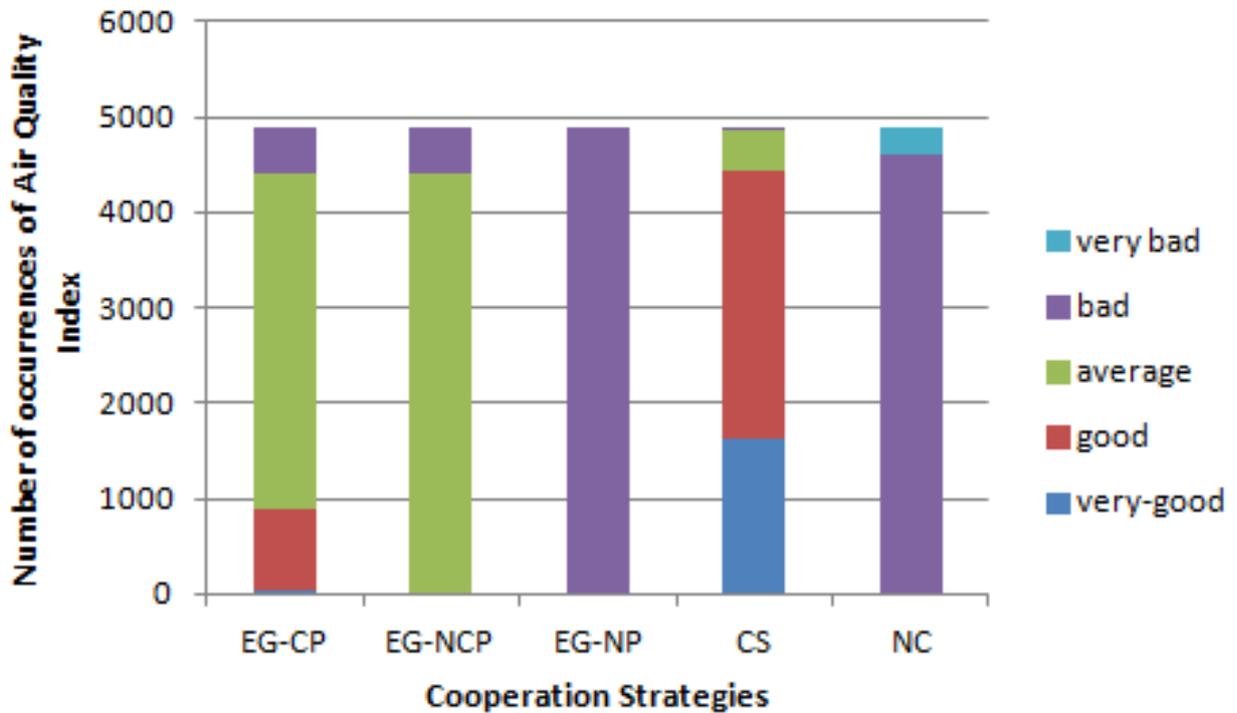


Figure 8.1: L'index de la qualité de l'air pour toutes les stratégies testées (EG-CP, EG-NCP, EG-NP, CS and NC), en utilisant 4900 pas de simulation.

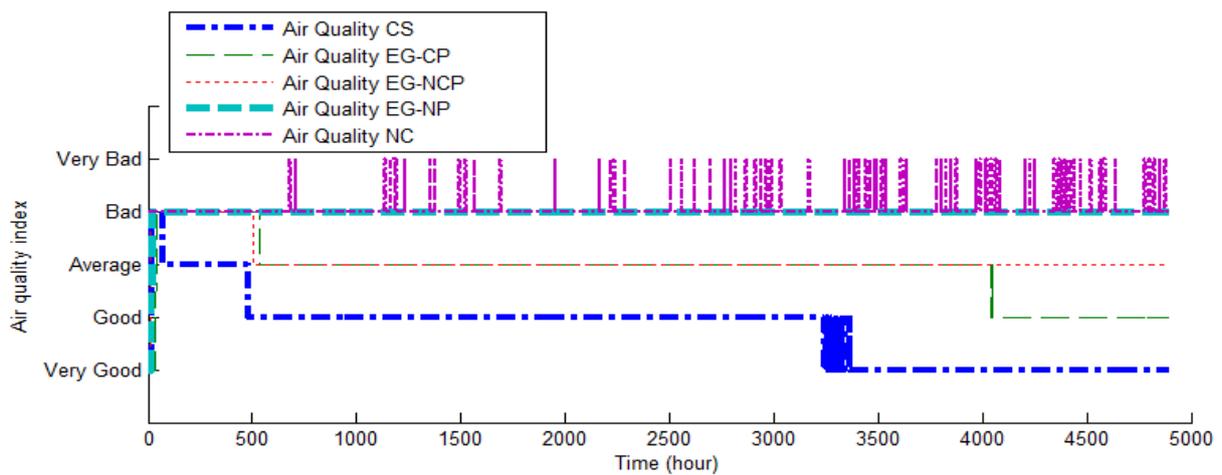


Figure 8.2: évolution de l'index de qualité de l'air pendant les 4900 pas de simulation.

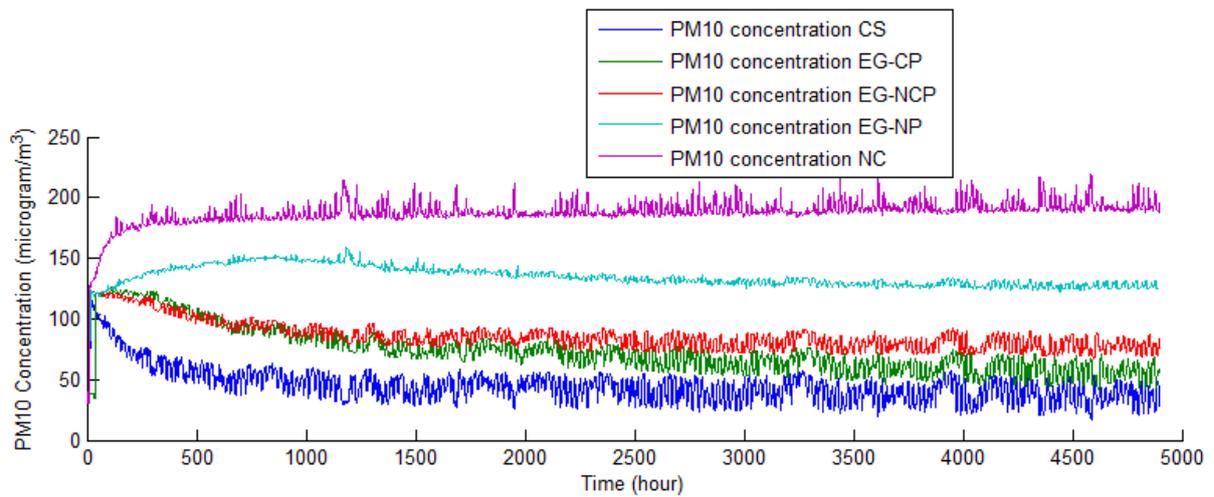


Figure 8.3: Evolution de la concentration des PM10 pendant le temps de la simulation

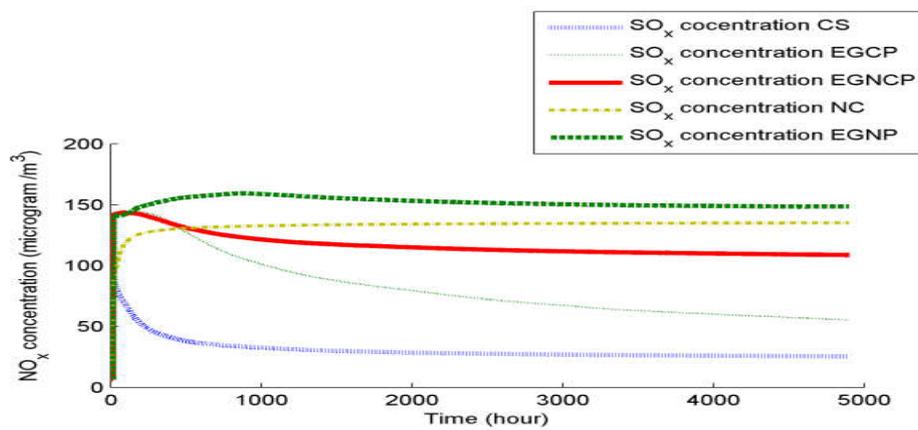


Figure 8.4 : L'évolution de la concentration du SO<sub>x</sub> pendant l'exécution du premier scénario.

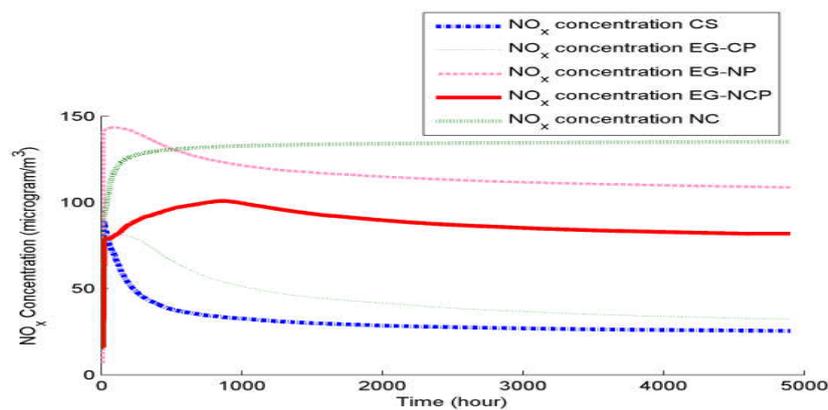
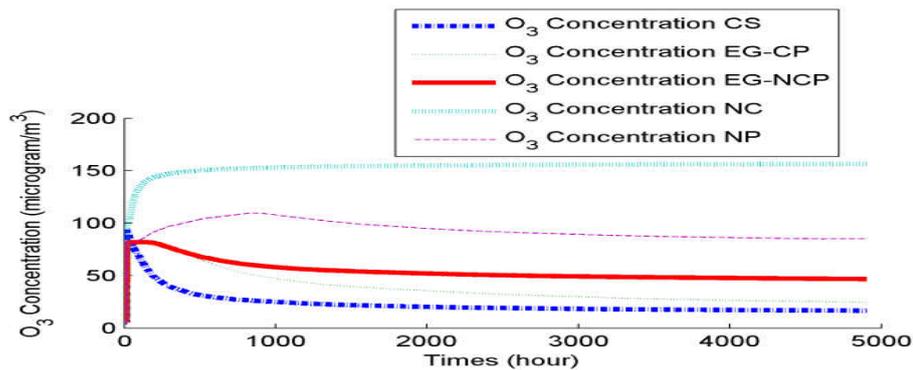


Figure 8.5: L'évolution de la concentration du NO<sub>x</sub> pendant l'exécution du premier scénario.



**Figure 8.6: l'évolution de la concentration du O<sub>3</sub> pendant l'exécution du premier scénario.**

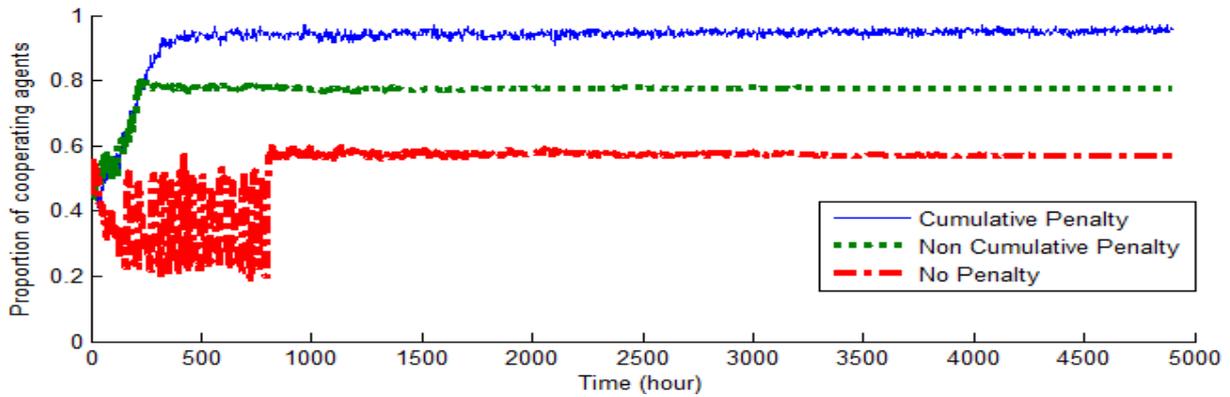
Figure 8.3 présente l'évolution de la concentration du PM<sub>10</sub> pendant le temps de la simulation. La stratégie centralisée (CS) prend moins de temps pour contrôler le niveau de concentration du PM<sub>10</sub> et permet de le garder au dessus de la limite cherchée (limite fixée parmi le scenario de simulation).

Les stratégies basées NPPD prennent plus de temps et arrivent à maintenir la concentration du PM<sub>10</sub> proche du niveau recherché. Comme montré dans la figure 8.3, la méthode EG-CP permet de contrôler le niveau des PM<sub>10</sub> mieux que la stratégie sans pénalisation cumulée. Les deux stratégies donnent des résultats mieux que la stratégie EG-NP (sans pénalisation). La stratégie qui n'inclut pas la coopération NC est présentée pour montrer comment le niveau de concentration des PM<sub>10</sub> peut évoluer si les agents qui contrôlent les sources ne coopèrent pas. Dans ce cas les agents émettent le polluant en utilisant leurs capacités maximales c'est-à-dire en utilisant leurs taux d'émission maximum, malgré cela nous pouvons comme même observer des pics et des oscillations qui sont causées par les paramètres climatiques.

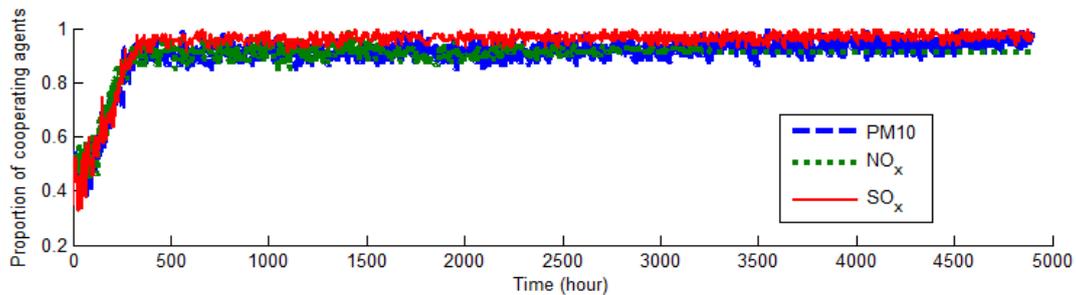
Les Figures 8.4, 8.5 et 8.6, permettent de voir l'évolution de la concentration des SO<sub>x</sub>, NO<sub>x</sub> et l'Ozone, pendant le temps de la simulation en utilisant les cinq stratégies de coopération. Nous observons les mêmes performances pour la stratégie centralisée (CS), elle permet rapidement de contrôler le niveau de pollution. Les stratégies EG présentent les performances semblables à celles montrées avec les PM<sub>10</sub>. Ceci nous montre le fait que le niveau de pollution est très influencé par la méthode de pénalisation choisie. La méthode de pénalisation cumulée permet d'avoir les meilleurs résultats et permet d'atteindre parfois le niveau objectif défini parmi les paramètres du scénario.

La stratégie sans coopération (NC) présente les concentrations les plus élevées en terme de taux de pollution, quand les agents ne coopèrent pas et utilisent leurs taux maximums, plusieurs pics sont affichés, et la pollution atteint des niveaux très alarmants.

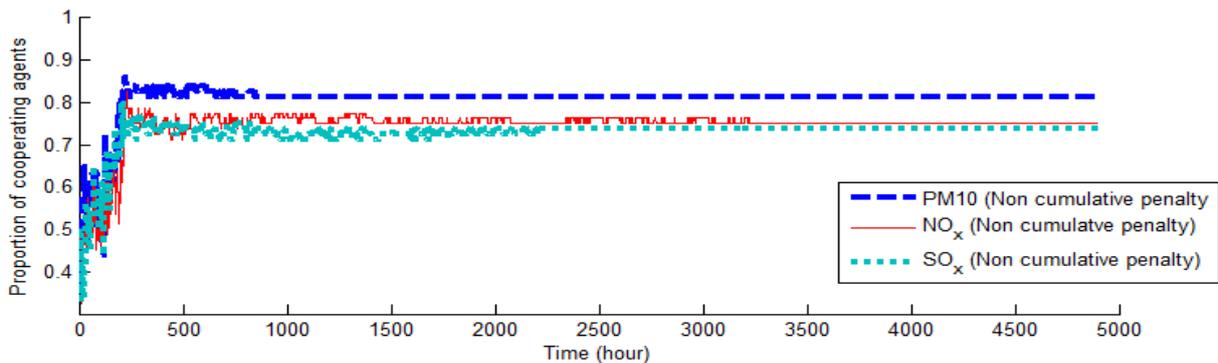
En comparant l'évolution des polluants, nous pouvons observer le fait que les PM10 présentent plus d'oscillation et pics par rapport aux autres polluants, ceci est expliqué par le fait que les PM10 sont très influencés par un certains nombre d'événements naturels (comme par exemple les vents de sables) ces événements influencent beaucoup sur les PM10, mais n'ont pas un grand impact sur les autres polluants.



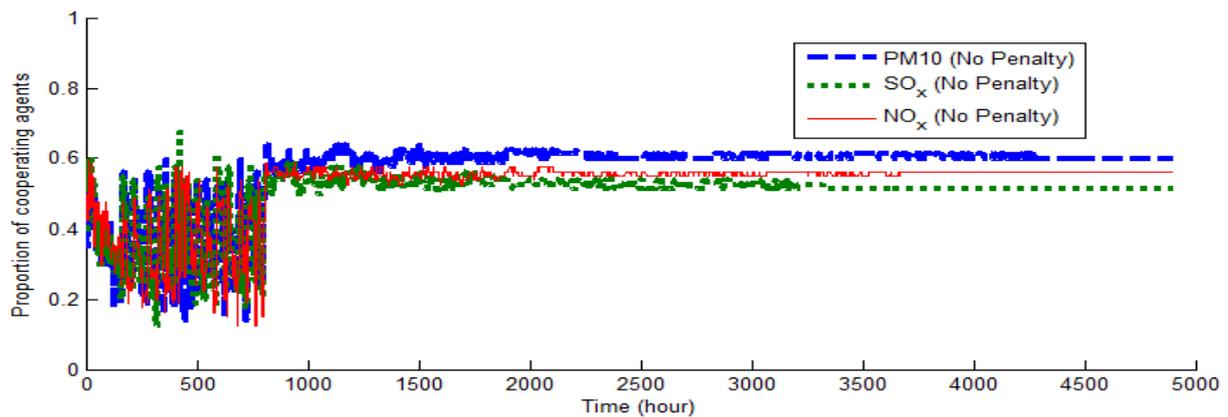
**Figure 8.7: Taux de coopération par rapport à toute la population d'agents pour les stratégies EG-CP, EG-NCP and EG-NP.**



**Figure 8.8: Le taux de coopération par type de polluants en utilisant la stratégie EG-CP.**



**Figure 8.9, Le taux de coopération par type d'agents en utilisant la stratégie EG-NCP.**



**Figure 8.10: Le taux de coopération par type de polluants dans le cas d'une stratégie EG-NC.**

Le taux de coopération (ou bien la proportion des agents qui choisissent de coopérer) est un paramètre qui nous permet de comparer les différentes stratégies. Nous pouvons définir un taux global qui se calcule par rapport à toute la population d'agents, et aussi par type de polluant, ce qui permet de savoir quel type d'agent présente les meilleurs taux de coopération.

La figure 8.7 permet de voir le taux de coopération pour toute la population d'agents pour les trois stratégies basées NPPD. Nous pouvons voir que le choix de la méthode de pénalisation a un grand effet sur le taux de coopération. Pour le cas de EG-CP l'équilibre a lieu lors du pas de simulation 387 et le taux se stabilise entre 0.93 et 0.95, pour le cas de EG-NCP elle se stabilise dans le pas de simulation 196 et le taux reste entre 0.73 et 0.80 après il se stabilise à 0.77. Pour le cas où aucune méthode de pénalisation n'est utilisée. L'EG-NP présente la stabilisation la plus tardive, à savoir, l'équilibre a lieu lors du pas de simulation 808 et l'équilibre du taux 0.57.

Les trois figures 8.8, 8.9 et 8.10 montrent le taux de coopération selon les types de polluant et pour les trois stratégies EG-CP, EG-NCP et EG-NP. Nous pouvons apercevoir le fait que les agents qui produisent les PM10 ont le plus grand taux de coopération, ceci est dû au fait que les PM10 présentent le plus de pics par rapport aux autres polluants, et comme les méthodes de pénalisation sont basées sur le dépassement d'un seuil fixé par les paramètres de scénario, ce qui explique que les agents coopèrent plus, pour la simple raison qu'ils sont plus pénalisés par rapport aux agents qui produisent le  $\text{NO}_x$ ,  $\text{SO}_x$ . Ce lien de causalité entre niveau de pollution et taux de coopération montre clairement l'efficacité des méthodes de pénalisation utilisées.

Un autre élément important à mettre en évidence, est le fait que les agents qui contrôlent les sources  $\text{NO}_x$  et  $\text{SO}_x$  sont aussi pénalisés pour leur participation à la formation du  $\text{O}_3$ , donc ceci peut compter parmi les causes de leur stabilisation tardive par rapport aux  $\text{PM}_{10}$ .

### 2.3 Scénario de simulation numéro 2

L'objectif des scénarios est de voir l'efficacité des stratégies de coopération dans le cas de la gestion d'une crise de pollution (Les pics de concentration). Sous certaines conditions climatiques, la pollution peut atteindre des niveaux très alarmants appelés communément pics de pollution. Avec ce scénario nous avons gardé les mêmes paramètres du premier scénario, nous avons adapté le nombre d'heures du pas de simulation, 1 pas est équivalent à 24 heures. La simulation vise à prédire les pics des  $\text{PM}_{10}$  pendant 24 heures.

Pour chaque pas de simulation nous calculons le maximum des valeurs pour 24 heures et nous obtenons les pics. Les mêmes stratégies de coopération ont été utilisées pour ce scénario.

#### 2.3.1 Résultat de la simulation

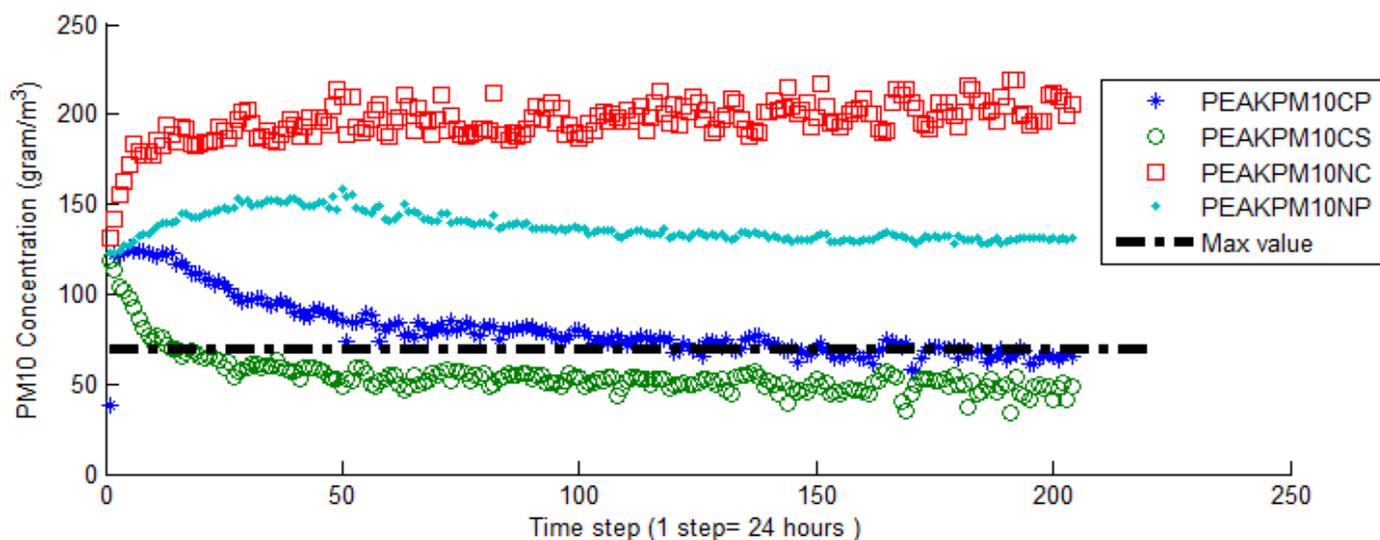


Fig. 8.11. Les pics de  $\text{PM}_{10}$  pour 24 heures, selon les quartes stratégies testées..

La figure 8.11 montre l'évolution de la concentration des  $\text{PM}_{10}$  pour chaque stratégie; chaque point représente le pic pour une journée (24 heures). La stratégie CS (**PICPM10CS**) présente les meilleures performances, elle prend moins de temps pour ramener le niveau de la pollution en dessous de l'objectif (niveau défini dans le scénario de simulation). La stratégie (**PICPM10CP**) présente des performances raisonnables, mais elle est légèrement plus lente que la stratégie CS. Elle prend plus de temps pour faire baisser le niveau de pollution. Toutefois, elle

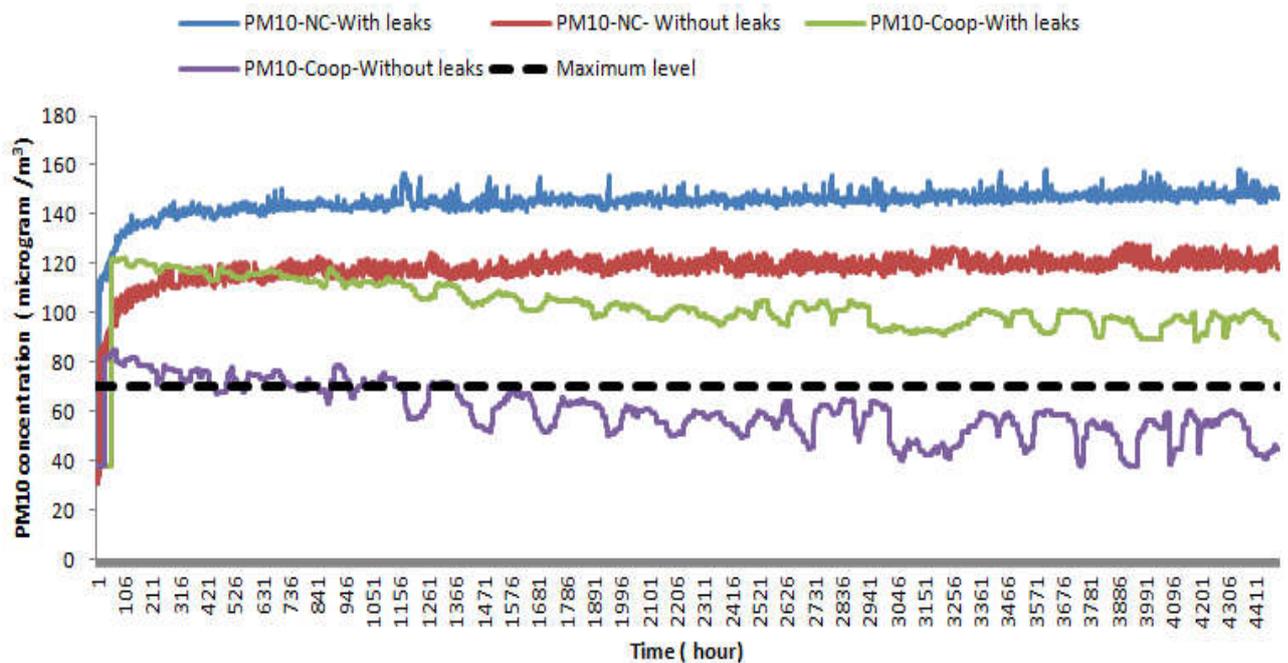
parvient à maintenir le niveau de la pollution à proximité de l'objectif. Les règlements de pénalisation ont un grand effet sur le niveau de PM10. Comme illustré, le **PICPM10CP** (avec pénalités) contrôle le niveau de pollution mieux que la stratégie non pénalisante (**PICPM10NP**). La stratégie NC (non-coopération) est présentée afin de montrer l'impact de la coopération sur les pics des PM10. Lorsque la coopération n'est pas utilisée (**PICPM10NC**), les agents agissent égoïstement et ne se soucient pas de la pollution. Comme les agents atteignent leur taux d'émission maximale, nous pouvons observer une oscillation qui est causée par les conditions climatiques. Dès lors, le niveau atteint des valeurs alarmantes et de nombreux pics se produisent.

## 2.4 Le troisième scénario de simulation

La question à laquelle nous cherchons à trouver une réponse avec ce scénario est : « Comment la coopération entre les agents qui contrôlent les sources de pollutions, peut aider à gérer une situation de crise de pollution incluant des sources de pollution non contrôlées ? ». Une telle situation peut avoir lieu lors d'un incident industriel ou bien un événement naturel, ces deux types de sources de pollution sont dans la majorité des cas incontrôlables au moins pendant un certain temps. Afin de générer la situation pendant ce temps, nous inspectons comment ce que nous contrôlons peut être géré pour réduire les effets d'une telle crise. Les agents sont ramenés à altérer leurs taux d'émission afin de gérer la crise.

Pour cela nous avons imaginé une situation où parmi les sources de pollution y en a à celles qui sont non contrôlées. Pour cela nous allons inspecter ce cas en utilisant deux stratégies : la première incluant la coopération et la deuxième n'inclut aucun mécanisme de coopération.

## 2.4.1 Résultat de la simulation



**Figure 8.12, L'évolution de la concentration du PM10 pour le scénario avec et sans des sources non-contrôlées.**

Dans la figure 8.12, nous pouvons observer comment la concentration des PM10 peut évoluer lors de la simulation. Pour pouvoir observer l'effet du choix de la stratégie dans le cas d'une crise de pollution Le scénario est exécuté selon quatre combinaisons :

1. **NC-With-leaks** : En utilisant une stratégie sans coopération et avec des fuites (des sources non-contrôlées).
2. **NC-Without leaks** : Avec une stratégie sans coopération et sans des fuites.
3. **Coop-with leaks** : Avec coopération et avec des fuites.
4. **Coop-Without leaks** : Avec coopération et sans fuites.

Comme illustré dans la figure 8.12, le cas où des sources non-contrôlées existent, la gestion avec coopération permet de réduire significativement le taux des PM10 (voir la courbe **PM10-coop-with-leaks**) malgré le fait qu'il n'arrive pas au niveau souhaité (qui est fixé parmi les paramètres de simulation). En comparant cette courbe avec celle dans le cas sans des sources non-contrôlées (voir courbe **PM10-coop-without-leaks**), nous pouvons observer la gravité de la situation et la différence causée par ces sources. Les deux courbes PM10-NC-with-leaks et PM10-NC-without-leaks permet aussi de constater l'impacte des sources non contrôlées.

En comparant le cas avec et sans coopération, respectivement (PM10-coop-with-leaks), (PM10-NC-with-leaks), nous pouvons mettre en évidence l'efficacité de la coopération pour le contrôle d'une situation de crise.

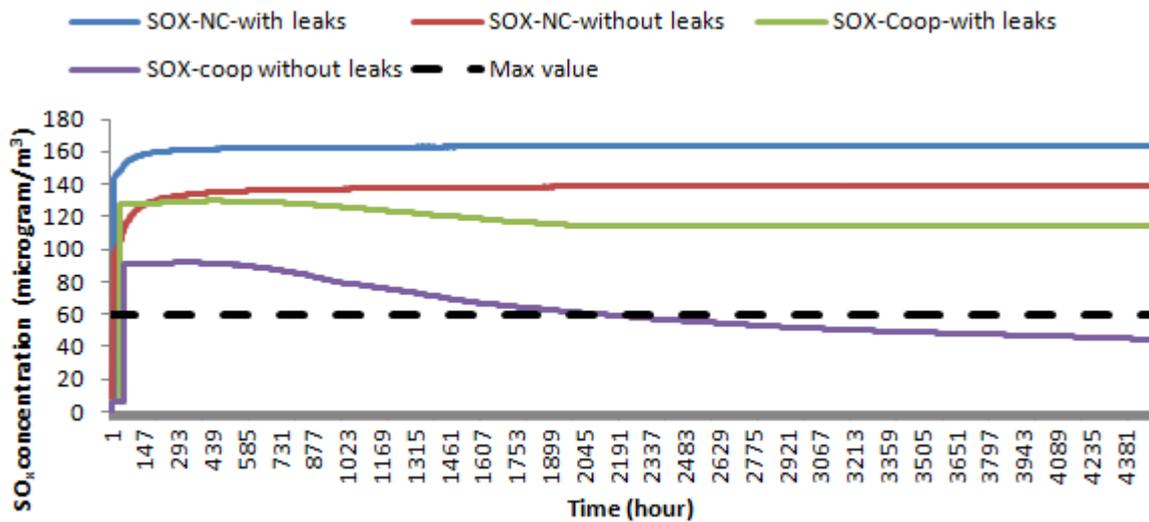


Figure 8.13, SO<sub>x</sub> concentration for the scenario with and without leak sources.

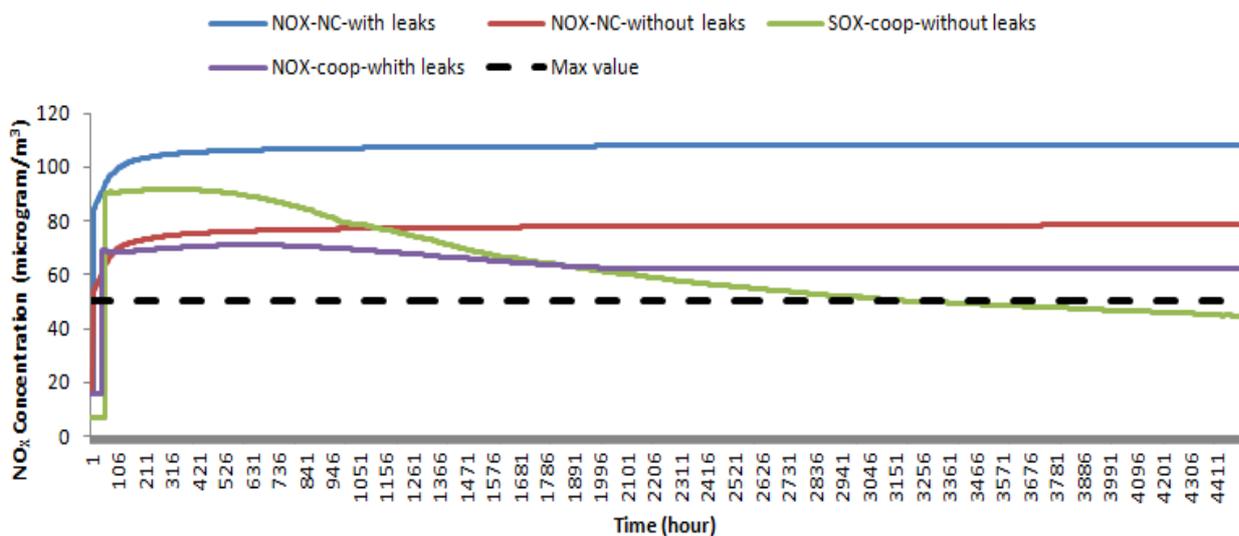


Figure 8.14, NO<sub>x</sub> concentration for the scenario with and without leak sources.

La même évolution peut être constatée dans les figures 8.13 et 8.14, pour le cas de NO<sub>x</sub> et le SO<sub>x</sub>. Les résultats montrent que la coopération a un grand impact sur l'évolution de la pollution. La simulation a montré que même lorsque les agents coopèrent à leurs niveaux maximums, la concentration de la pollution est toujours au-dessus du niveau recommandé. La compensation entre les émissions de sources contrôlées et non contrôlées n'est pas possible. La modification des paramètres de la simulation permet de rechercher du nombre de sources non contrôlées et

leur taux d'émission qui peuvent être tolérés. Ce genre de question est d'une importance majeure pour les gestionnaires de la qualité de l'environnement.

**Conclusion :**

En présentant les résultats des exécutions de ces trois scénarios de simulation, nous avons pu montrer l'utilité de notre simulateur et notre approche. Il est possible d'avoir d'autres types de scénarios qui permettent d'étudier d'autres types de questions. Nous avons choisi ces trois, juste à titre explicatif.

Dans le chapitre qui suit, nous résumerons les conclusions que nous avons pu constater pendant ce travail.

## Chapitre 9 : Conclusions et Perspectives.

---

S'inscrivant dans le contexte de la simulation et la modélisation des systèmes socio-environnementaux, le travail de recherche décrit dans cette thèse a porté sur l'inclusion des comportements et les décisions humaines parmi les éléments clés de la simulation de l'évolution de la pollution de l'air dans une région urbaine.

Avant de présenter les conclusions retenues, nous commençons par rappeler la problématique traitée. La pollution est un résultat d'une perturbation d'un système écologique, ces perturbations sont dans la pluparts du temps causées par des activités anthropologiques. La pollution de l'air, jusqu'à présent a été traitée comme étant un problème isolé, à savoir, les travaux et les modèles ont toujours cherché à trouver la dispersion et la concentration des polluants dans les régions urbaines, ceci sans tenir en considération la décision humaine, c'est-à-dire, la composante sociale du problème. Dans notre thèse nous nous sommes intéressés à inclure ces décisions parmi les éléments essentiels de la simulation.

### Conclusions

Pouvoir éluder la dégradation de l'environnement et la qualité de l'air est l'objectif principal des organismes de gestion et protection de l'environnement. Il est crucial pour eux de pouvoir mesurer l'efficacité de leurs politiques de gestion et prévoir les résultats de leurs décisions. La simulation peut servir comme étant une aide à la décision qui leurs permet de mieux estimer les résultats de leurs actions et décisions.

La modélisation de l'interaction qui a lieu entre le composant social et le composant écologique est une pierre angulaire de tout effort d'études des systèmes socio-environnementaux. L'approche multi-agents permet une modélisation basée individu, qui offre la possibilité de manipuler, à un niveau de granularité très fine, les comportements des individus qui agissent dans un système socio-écologique. Cette manipulation permet d'en gagner plus de connaissance sur le phénomène modélisé et aussi grâce à la propriété d'émergence des SMA, nous pouvons identifier des propriétés et des comportements cachés dans l'agrégation. Ces propriétés sont très difficiles à détecter, voir même parfois elles sont impossible à identifier en utilisant les autres approches classiques de modélisation. Cette approche permet de fixer des paramètres niveau macro et observer les résultats au niveau micro, l'inverse est aussi envisageable.

Les études qui traitent la pollution de l'air sont toujours concentrées sur les aspects physiques (dispersion) et chimiques (concentration, destruction). Ces travaux ont dans la plupart du temps l'objectif de surveiller la qualité de l'air. Ils n'englobent pas l'aspect source d'émission et surtout la décision humaine qui contrôle ces sources d'émission. A cause de ça, l'activité humaine (anthropique) qui est la source majeure de pollution n'est pas modélisée et prise en charge par la simulation. Avec notre approche, nous avons pu modéliser les comportements des contrôleurs de sources d'émission. Ce qui permet d'inspecter l'efficacité des normes et des mécanismes de contrôle de pollution.

Notre approche permet de mener, à travers la simulation, des expérimentations qui visent à étudier et évaluer les mécanismes de contrôle de la pollution de l'air. Ceci est fait grâce à la modélisation des comportements des acteurs impliqués dans la gestion de la qualité de l'air (Les contrôleurs de source d'émission, les organismes publics de contrôle de la qualité de l'air, le trafic routiers).

La prédiction des concentrations des polluants ainsi que la qualité de l'air, sont faites en utilisant une méthode à deux étages. Elle consiste à combiner un modèle de dispersion Gaussien et plusieurs réseaux de neurones artificiels. Le premier stage permet de calculer la dispersion des polluants à partir des sources d'émission, après les valeurs de dispersion sont passées comme paramètres pour prédire la concentration des polluants en utilisant des réseaux de neurones de types RBF, ces réseaux reçoivent aussi parmi leurs entrées les valeurs des paramètres climatiques. Cette méthode nous a permis d'introduire les sources d'émission et les paramètres climatiques parmi les entrées. Ce qui donne des prédictions plus réalistes et qui incluent l'incertitude causée par le climat.

Cette méthode permet de donner à chaque pas de simulation des prédictions concernant la qualité de l'air. Elle est intégrée dans un système multi-agents, qui permet de simuler comment les acteurs coopèrent afin de préserver une qualité de l'air raisonnable. Les prédictions concernant la qualité de l'air permettent aussi de tester les méthodes de pénalisation des agents (qui représentent les acteurs) afin de gérer la situation d'un pic de pollution.

L'approche permet de tirer profit des avantages de chacun des méthodes combinées. Grâce aux mécanismes d'apprentissage et de généralisation des réseaux de neurones artificiels, les systèmes de simulation conçus sont à la fois génériques, et aussi à s'adaptent aux changements qui ont lieu au niveau des paramètres climatiques (saisonnalité).

En utilisant cette approche, un système de simulation a été conçu, afin de montrer la faisabilité de notre approche. Ce système a été utilisé pour faire des expérimentations via des scénarios de simulation. Les simulations sont exécutées en utilisant les données réelles concernant la pollution et les paramètres climatiques de la région d'Annaba. Cinq stratégies de coopération ont été testées et comparées.

La simulation a permis d'étudier comment la coopération peut aider à contrôler ou réduire la crise de la pollution. Les résultats ont montré que la coopération entre les différents acteurs impliqués dans la gestion de la crise, peut réduire l'effet des pics de la concentration des polluants, même si certaines sources sont incontrôlées (dues à des événements ou des phénomènes naturels).

Plusieurs stratégies de coopération ont été testées afin de pouvoir valider notre hypothèse concernant l'évaluation des politiques de gestion et de contrôle de la pollution. Les résultats obtenus nous ont permis de voir les effets qu'a la coopération sur l'évolution de la concentration des polluants. Pour résumer, notre étude permet de :

- Modéliser et introduire la décision humaine concernant les sources d'émission parmi le processus de la simulation.
- Evaluer la coopération qui peut avoir lieu entre les acteurs impliqués dans la gestion de la qualité de l'air (pollueurs et organismes publiques).
- Avoir des prédictions concernant les performances et l'efficacité des règles et mécanisme de gestion de la qualité de l'air.
- Permettre une simulation d'un cas de crise de pollution incluant des sources de pollution non-contrôlées.
- Utiliser une approche combinée afin de bénéficier des avantages qu'apportent chacune d'entre elles.

**Perspectives :**

Notre travail vient comme un effort dans le cadre d'amélioration de nos connaissances concernant les systèmes socio-environnementaux. Comme tout travail de recherche en modélisation et simulation, il présente quelques limites que nous étalerons dans cette section. La principale limite concerne le degré de réalité de la simulation. Ce critère est l'un des plus importants car, c'est ce qui déterminera la confiance que donnera l'utilisateur aux résultats de la simulation.

Notre approche est générique, elle permet de concevoir des systèmes de simulation basés (ou orientés) données. De ce fait, la disponibilité des données empiriques représente la pierre angulaire de l'approche. Ceci permet d'avoir des simulations bien calibrées et utilisant des données réelles. Il est à noter que la quantité des données disponibles détermine l'échelle temporelle de la simulation, une simulation pour des prédictions à long terme nécessite une quantité de données importante.

Notre travail permet d'ouvrir la porte à des éventuelles améliorations et sujets qui peuvent être exploités. Dans cette section nous allons présenter les perspectives qui permettent d'exploiter les résultats de notre travail de recherche.

- [1] Utiliser plusieurs échelles de représentation du temps : L'approche multi-agents offre une possibilité très utile, elle donne la possibilité d'exploiter plusieurs échelles temporelles dans la même simulation. Il est d'une importance majeure, car ceci permet de, à la fois observer l'évolution à court et moyen terme en utilisant une représentation, et aussi en même temps observer l'évolution à long terme en utilisant une autre représentation du temps. Exemple : dans notre cas, l'évolution des concentrations des polluants peut être représenté en utilisant des pas de simulation de l'ordre de quelques heures, mais les comportements des acteurs (les contrôleurs des sources de pollution) peut être représenté en utilisant une autre représentation parallèle avec un pas de simulation de l'ordre de plusieurs jours (ou plusieurs mois), ceci leur donnera plus de temps pour décider d'investir dans la réduction des émissions. Bien sûr ces deux représentations, restent liées et l'équivalence reste correcte.
- [2] Inclure d'autres types de sources de pollution de l'air, ce qui permettra de bien traiter les autres types de sources de pollution, notamment, celles issues du trafic routier, station de traitement des déchets, les activités volcaniques et les feux de forêts.
- [3] Utiliser d'autres modèles de dispersion : Comme discuté dans le Chapitre 3, il existe plusieurs modèles de dispersion. L'exploitation des autres types peut inclure l'aspect chimique et dynamique de la pollution de l'air et rendre la simulation capable à imiter l'évolution chimique des polluants (taux de destruction, déplacement, interaction photochimiques).
- [4] Tester d'autres mécanismes de prise de décision : Dans la version actuelle de notre approche nous avons opté pour un modèle de mécanisme de prise de

décision basé théorie des jeux, précisément le dilemme des prisonniers à N personnes. Il est envisageable d'utiliser d'autres type de mécanisme telle que l'approche BDI (Belief Desire Intention), ou bien l'approche floue, ou même la combinaison entre tous ces mécanismes afin d'exploiter leurs avantages.

- [5] Valider notre modèle en utilisation des données concernant d'autres villes : notre approche a été validée en utilisant des données qui concernent la ville d'Annaba, nous envisageons dans l'avenir à utiliser des données provenant d'autres villes. Ceci permet de montrer que le simulateur peut être réutilisé pour étudier la pollution de l'air dans d'autre ville.
- [6] Visualisation des résultats grâce à un module GIS : en intégrant un module GIS la visualisation des résultats sur une carte géographique réelle permettra de bien localiser les zones qui sont plus pollués.
- [7] Envisager la possibilité de tester notre approche pour d'autre type de pollution : notre approche est conçue d'une façon intuitive et générique, elle peut être adoptée pour concevoir des simulateurs traitants d'autre types de pollution notamment la pollution des eaux (maritime et rivière et nappes). Les principales modifications à apporter, c'est le changement du modèle de dispersion par un modèle de diffusion des polluants dans les eaux. Ceci permettra de faire des simulations tout en incluant le composant social, à savoir, l'ensemble des acteurs qui gèrent la qualité des eaux dans les rivières ainsi que les pollueurs.
- [8] Tester le modèle pour une prédiction à long terme.

## Annexe A : Publications et Communiqués :

---

Au cours de ces années de travail sur notre sujet de thèse, nous avons pu communiquer et publier nos résultats dans de nombreuses conférences et journaux spécialisés. Ceci nous a permis d'avoir des discussions très fructueuses avec des chercheurs travaillant dans le même domaine. Les retours que nous avons reçus de la part des nombreux comités de lecture, ont été très pertinents et utiles:

### Publication dans des revues internationales :

1. **Ghazi sabri**, & Tarek, K. M. (2012). Combination of artificial neural network models for air quality predictions for the region of Annaba, Algeria. *International Journal of Environmental Studies*, 69(1), 79-89, **Taylor & Francis, UK**.
2. **Ghazi Sabri**, Julie Dugdale and Tarek Khadir, A Multi-Agent based Approach for Simulating the Impact of Human Behaviours on Air Pollution, *Informatica International Journal of Computing and Informatics (Accepted)*.

### Chapitre de livre :

1. **Ghazi, Sabri.**, Khadir, T., & Dugdale, J. (2014). Multi-Agent Based Simulation of Environmental Pollution Issues: A Review. In *Highlights of Practical Applications of Heterogeneous Multi-Agent Systems*. Volume 430 of the series *Communications in Computer and Information Science* pp 13-21, **Springer International Publishing**.  
[http://link.springer.com/chapter/10.1007/978-3-319-24399-3\\_6](http://link.springer.com/chapter/10.1007/978-3-319-24399-3_6)

### Communications internationales :

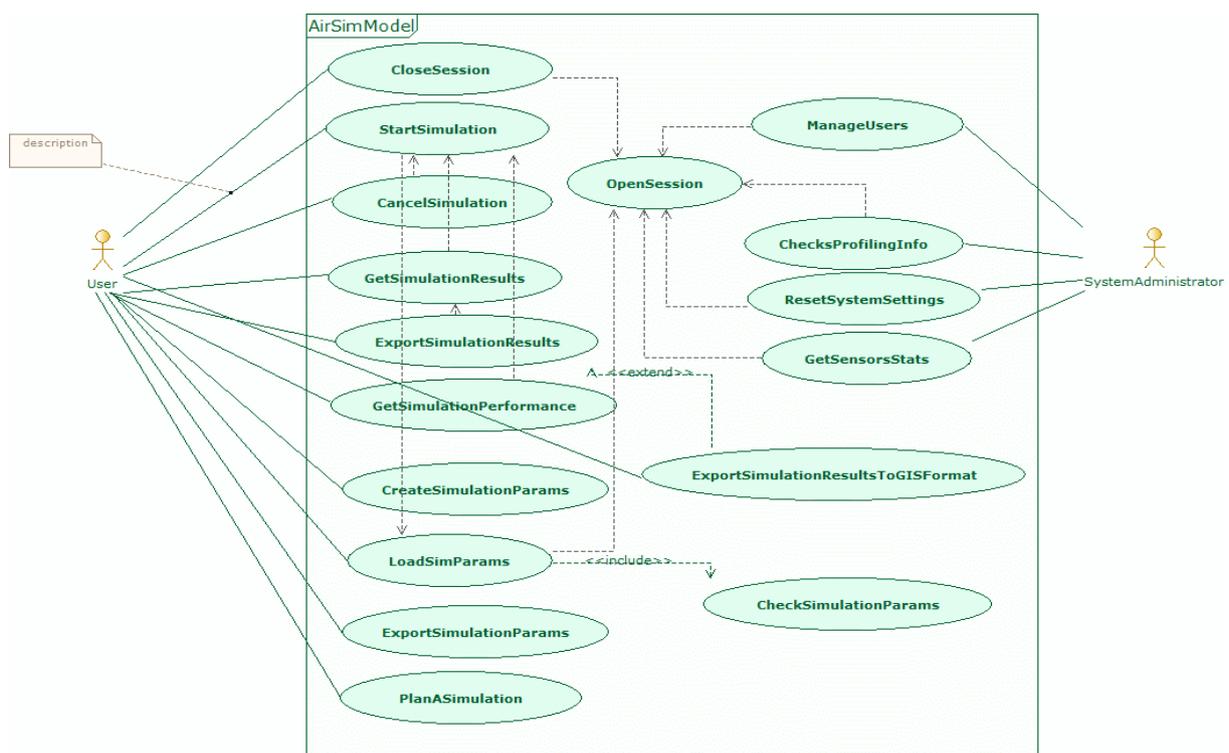
1. Khadir, M. T., & **Ghazi, Sabri**. Radial Basis Function and Elman Networks for Pollutant's Parameter Prediction in the Region of Annaba Algeria. *ADVCOMP 2010 : The Fourth International Conference on Advanced Engineering Computing and Applications in Sciences, Florence, Italy (2010)*.
2. **Ghazi Sabri**, Julie Dugdale and Tarek Khadir. Modelling PM10 Peaks Crises Using a Multi-Agent based Simulation: Application for the city of Annaba, North-East of Algeria, *International Conference on Information Systems for Crisis Response and Management in Mediterranean Countries,(2015)* ,Tunis, Tunisia, **Springer International Publishing**.
3. **Ghazi Sabri**, Julie Dugdale and Tarek Khadir, Modelling Air Pollution Crises Using Multi-agent Simulation, *Hawaii International Conference on System Sciences, HICSS-49: January 5-8, Hawaii, USA, (2016), IEEE*.  
<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7427204&newsearch=true&queryText=sabri%20ghazi&fname=&lname=&title=&volume=&issue=&spage=>

# Annexe B : Implémentation d'un simulateur en utilisant l'approche SMA

## Introduction :

Afin de pouvoir élucider la faisabilité de notre approche, nous décrivons dans ce chapitre quelques éléments concernant l'implémentation et le développement d'un simulateur basé SMA pour la simulation de la pollution de l'air. Ce système est conçu en utilisant l'approche présentée dans le chapitre précédent.

Le système de simulation proposé se compose de trois parties : un SMA, un modèle de dispersion et un modèle de prédiction basé RNA. La présentation des détails de l'implémentation de chacun de ces composants est décrite en utilisant le langage de conception UML (Unified Modeling Language). Comme montré dans le diagramme des cas d'utilisation schématisé dans la figure B.1, le système contient plusieurs fonctionnalités. Ces fonctionnalités sont conçues dans le but d'assurer le chargement, le lancement des scénarios de simulation. Il intègre aussi une fonctionnalité qui permet de visualiser et d'exporter les résultats de simulation.



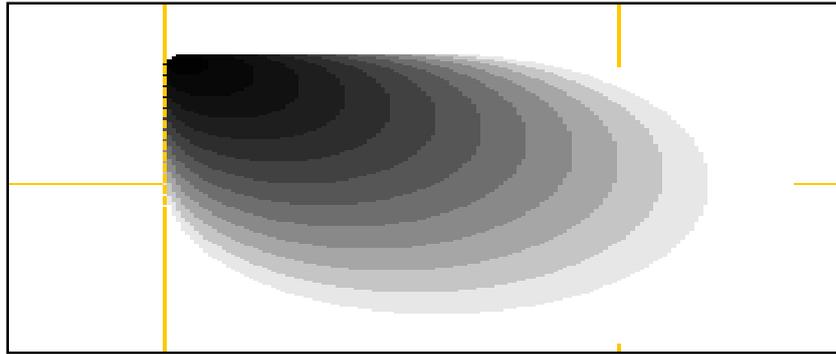
## Figure B.1, les fonctionnalités de simulateur sont décrites en utilisant le diagramme des cas d'utilisation UML.

### Le modèle de dispersion

Afin de bien optimiser l'algorithme de dispersion, nous avons procédé à une représentation sous forme de pile qui regroupe les points dans lesquels la concentration est calculées. A la fin, chaque source d'émission de pollution possède une pile qui contient une série de valeurs de concentration. Cette pile peut être vue comme étant la plume associée à la source. Comme montrée dans le listing suivant, la dispersion du polluant ne peut être calculée dans tout l'espace qui entoure un point d'émission. C'est pour cela qu'une série de point sont sélectionnés, cette sélection se fait en utilisant les trois coordonnées géométriques ( $X, Y, Z$ ). La source est supposée être dans les coordonnées  $(0,0,H)$  où H c'est l'hauteur de la source décrite dans le scénario de simulation. Il est à rappeler que la direction du vent est considérée constante pendant un pas de simulation, mais peut se changer d'un pas à l'autre.

```
double x = 100;
double y = 10;
double c = 0;
double rx = 100;
double ry = 100;
double sigy = 0.197 * Math.pow(x, 0.908);
double sigz = 0.1120 * Math.pow(x, 0.9100);
while (x <= 1500) {
    if (x < 10000) {
        sigy = 0.197 * Math.pow(x, 0.908);
    } else if (x >= 10000) {
        sigy = 0.285 * Math.pow(x, 0.867);
    }
    if ((x > 100) && (x <= 500)) {
        sigz = 0.1120 * Math.pow(x, 0.9100);
    } else if ((5 < x) && (x < 50)) {
        sigz = 0.1014 * Math.pow(x, 0.926);
    } else if (x > 50) {
        sigz = 0.1154 * Math.pow(x, 0.9109);
    }
    c = ((s.getEmissionRate() * 1000) / (Math.PI * sigy * sigz * 2.12))
        * (Math.exp((-1 / 2) * Math.pow(y / sigy, 2)))
        * Math.exp((-1 / 2) * Math.pow(10 / sigz, 2));
    s.gaussianPlumVals.push(c);
    pnt.add(c);
    xis.add((new Integer((int) rx)));
    ry = ry + 3;
    x = x + 100;
    ry = 100;
    rx = rx + 3;}
```

En utilisant ces calculs nous pouvons obtenir la concentration du polluant selon plusieurs niveaux (un niveau pour chaque point de la pile). Comme illustré dans la figure, l'agregation est ensuite calculée selon la moyenne des valeurs de la pile.



**Figure B.2 : Affichage de la pile des points de concentration par rapport à la source d'émission.**

Ce traitement est appliqué sur toutes les sources en utilisant leurs taux d'émission et leurs coordonnées. Ces informations sont alors utilisées par le modèle RNA pour prédire les concentrations dans chaque cube ainsi que la qualité de l'air.

## Le modèle de prédiction basé RNA

Une fois la concentration est calculée en utilisant le modèle de dispersion, vient après l'étape où nous incluons les paramètres climatiques via l'utilisation du modèle RNA. La conception des modèles RNA pour prédire les concentrations des différents polluants se fait en trois étapes : la première étape consiste à utiliser des données empiriques pour faire l'apprentissage du réseau. La deuxième étape, est celle de la validation, cette dernière permet de valider les performances du réseau en utilisant un ensemble de données qui n'ont pas été utilisées pendant l'étape d'apprentissage, une fois que ce réseau présente des performances convaincantes, il peut être ensuite sauvegardé et utilisé pendant la simulation. Dans le listing qui suit, nous donnons quelques éclaircissements concernant l'apprentissage des modèles RNA, il est à noter que nous avons utilisé la librairie Encog (Heaton, 2015). Cette dernière offre une implémentation optimisée de plusieurs types de RNA et d'autres algorithmes d'IA.

L'algorithme d'apprentissage des réseaux RBF sélectionnée dans notre cas, consiste à optimiser l'erreur en modifiant les paramètres : - le nombre de neurones dans la couche RBF, 2- les poids des entrées.

```
PredictionModel pmodel = new PredictionModel();
pmodel.setModelName("AIRQUALITY");
NormalizedField norm = new NormalizedField(
    NormalizationAction.Normalize, null, 0, 500, 1, 0);
NormalizedField normDesired = new NormalizedField(
    NormalizationAction.Normalize, null, 0, 5, 1, 0);
pmodel.normalizedUnified = norm;
pmodel.normalizedPollutant = normDesired;
for (int i = 0; i < AQ.length; i++) {
    AQ[i][0] = norm.normalize(AQ[i][0]);
    AQ[i][1] = norm.normalize(AQ[i][1]);
}
```

```

AQ[i][2] = norm.normalize(AQ[i][2]);
AQ[i][3] = norm.normalize(AQ[i][3]);
desiredAQ[i][0] = normDesired.normalize(desiredAQ[i][0]);
}
MLDataSet trainingDataSet = new BasicMLDataSet(AQ, desiredAQ);
BasicNetwork network = null;
if (!noTrain) {
    network = EncogUtility.simpleFeedForward(4, 14, 0, 1, true);
    final Backpropagation train = new Backpropagation(network,
        trainingDataSet);
    int epoch = 1;
    Console.frmLogMSG("Training the prediction model");
    do {
        train.iteration();
        epoch++;
    } while (train.getError() > 0.01);
    EncogDirectoryPersistence.saveObject(new File(pmodel.getModelName()
        + ".eg"), network);
} else {
    network = (BasicNetwork) EncogDirectoryPersistence
        .loadObject(new File(pmodel.getModelName() + ".eg"));
}
pmodel.setNnModel(network);
return pmodel;

```

## Le système multi-agents

Le troisième composant de notre système de simulation est un SMA qui modélise la décision humaine. Ce dernier permet de représenter les acteurs qui sont impliqués dans la gestion de la qualité de l'air, à savoir, les agents qui contrôlent les sources de pollution, les agents qui représentent l'organisme de surveillance de la qualité de l'air ainsi que les pouvoirs publics.

Comme schématisé dans la figure B.7, sous forme de diagramme de classes UML, nous pouvons distinguer les différents types d'agents. Ces agents peuvent être catégorisés en deux types : Les agents qui représentent des êtres humains et des agents qui sont créés juste pour accomplir certaines fonctionnalités du système et ne représentent pas des acteurs humains.

Toutes les classes d'agents héritent de la même classe mère « **Agent** », cette dernière regroupe les attributs et les méthodes qui sont commune entre tous les types d'agents (exemple : le nom, la localisation, l'adresse, l'état, etc).

Le tableau **B.1** décrit les types d'agents et leurs comportements ainsi que le type de leurs comportements. La communication inter-agent est décrite dans les deux figures **B.4** et **B.5**, la matrice d'interaction permet de voir les possibilités de communication entre les différents types d'agents.

Il est à noter que nous avons conçu notre SMA en se basant sur la plateforme JADE (BELLIFEMINE, 2003).

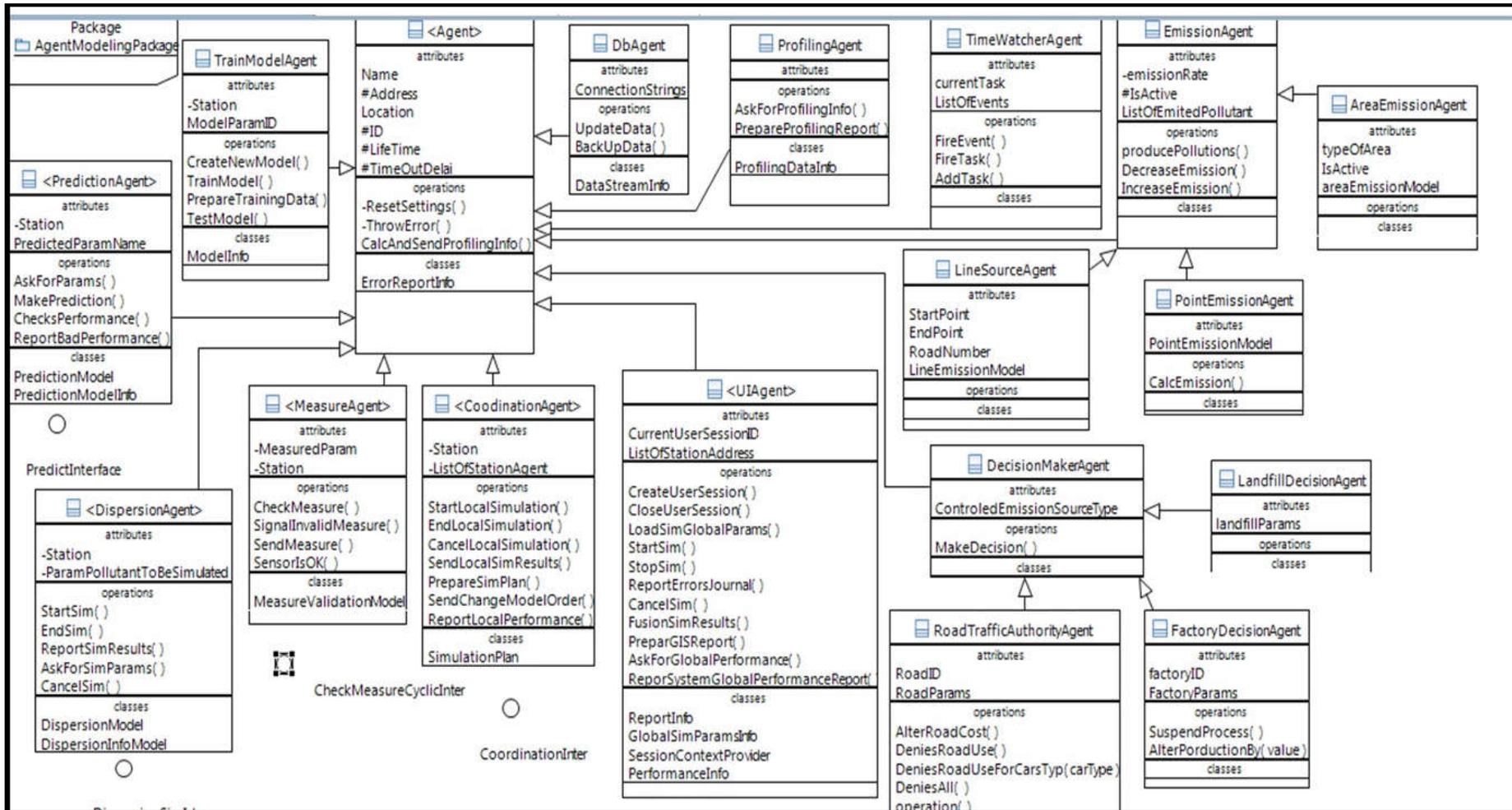
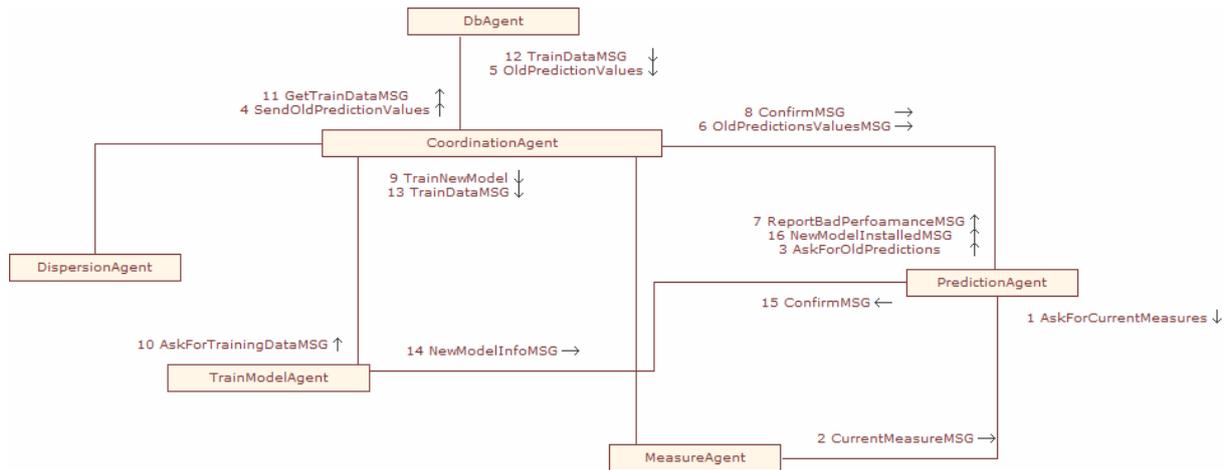


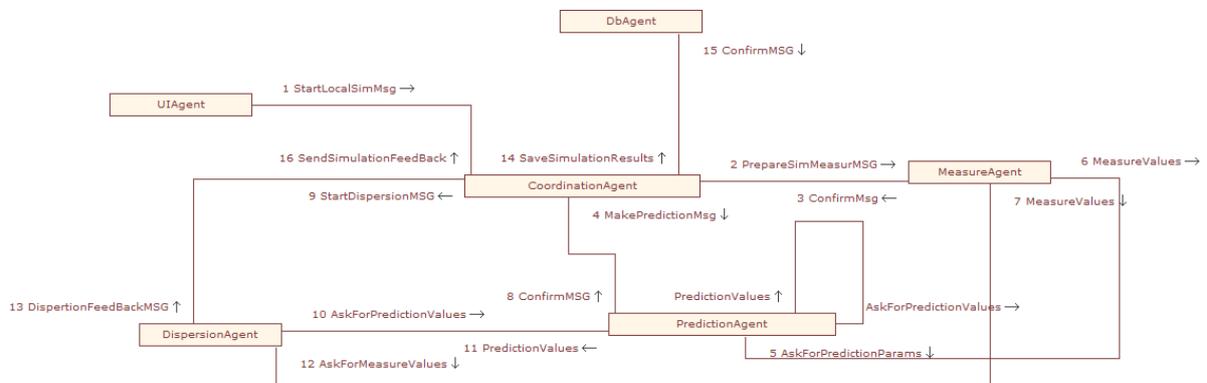
Figure B.3 Diagramme UML des classes du système MASAP

Type Agent	Comportements	Nature du comportement	Nombre d'instance
UIAgent	StartUIBehaviour InitializeSimulationBehavior	OneShotBehaviour	Nombre de session ouverts dans le système
TimerAgent	TimerBehavior TickerBehaviour	TickerBehaviour	Selon le nombre de simulation en cours d'exécution
DbAgent	LoadDataBehaviour SaveDataBehaviour ExportDataBehaviour	TickerBehaviour	Une seule instance
CoordinatorAgent	CoordinationAgentBehaviourCheckAirQuality CoordinationAgentBehaviourSendOrder	CyclikBehaviour	Selon le nombre de simulation en cours d'exécution
EmissionUnit	SetConfigSourceAgentBehavior	FSABehaviour	Selon le nombre de source de pollution
AreaSourceAgent	SetConfigSourceAgentBehavior ReceiveASKforERBehaviour ModeratePollutionVoting ModeratePollutionNAIVE NaiveModerationOFPollutionBehaviour		
LineSourceAgent	SetConfigSourceAgentBehavior ReceiveASKforERBehaviour ModeratePollutionVoting ModeratePollutionNAIVE NaiveModerationOFPollutionBehaviour		
PointSource	SetConfigSourceAgentBehavior ReceiveASKforERBehaviour ModeratePollutionVoting ModeratePollutionNAIVE NaiveModerationOFPollutionBehaviour		
MeasureAgent	SendParamsBehavior	FSABehaviour	
predictionAgents	MakeCOXPredictionBehavior MakeNOXPredictionBehavior MakeO3PredictionBehavior MakePredictionBehaviour MakeSOXPredictionBehavior	FSABehaviour	
SimulationFramework			Une seule instance
DispersionEngineAgent	RunDispersionModelforES	FSABehaviour	
ProfilingAgent	MakeProfiling	OneShootBehaviour	

Tableau B.1 : Les types d'agents et leurs comportements.



**Figure B.4 Diagramme de communication, fonctionnalité de lancement d'apprentissage du réseau de neurones artificiels**



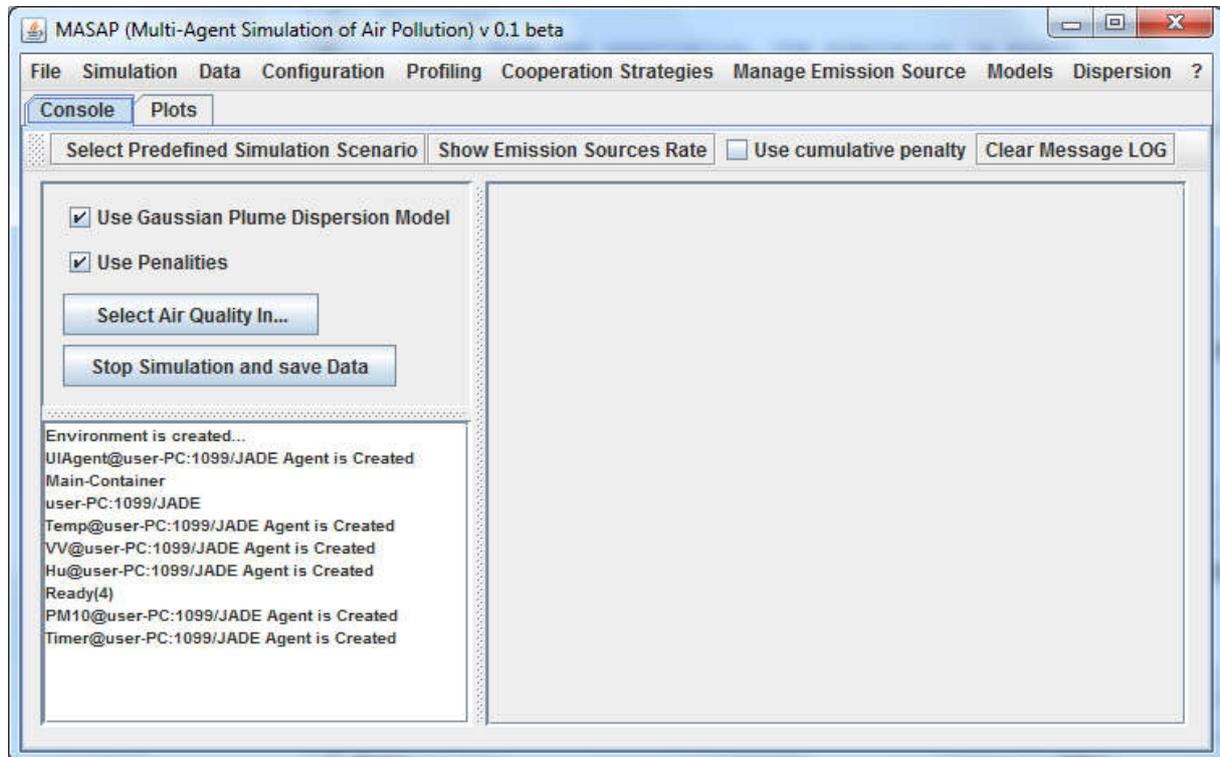
**Figure B.5 Diagramme de communication fonctionnalité de lancement du modèle de dispersion.**

Type d'Agent	PredictionAgent	DispersionAgent	MeasureAgent	TrainModelAgent	CoordinationAgent	UIAgent	DBAgent	ProfilingAgent
PredictionAgent	0	0	1	1	1	0	0	1
DispersionAgent	1	0	1	0	1	0	0	1
MeasureAgent	1	1	0	0	1	0	0	1
TrainModelAgent	1	0	1	0	1	0	0	1
CoordinationAgent	1	1	1	1	1	1	1	1
UIAgent	0	0	0	0	1	0	1	1
DBAgent	0	0	0	0	1	1	0	1
ProfilingAgent	1	1	1	1	1	1	1	1
SourceAgent	0	1	0	0	0	0	0	0

**Tableau B.2 : La matrice d'interaction inter-types-agents, ici 1 veut dire qu'il existe une interaction 0 représente aucune interaction.**

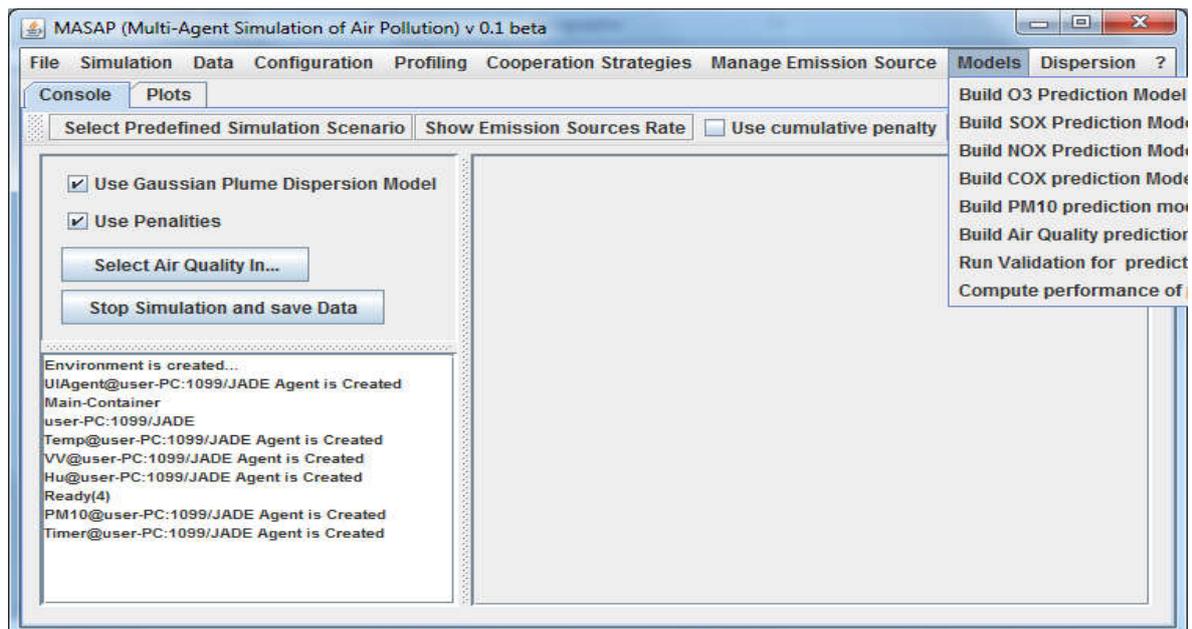
## Session de travail sous le système de simulation MASAP

L'écran principal du **MASAP** offre un accès à plusieurs fonctionnalités qui permettent de valider, créer et visualiser les scénarios de simulations ainsi que leurs résultats.



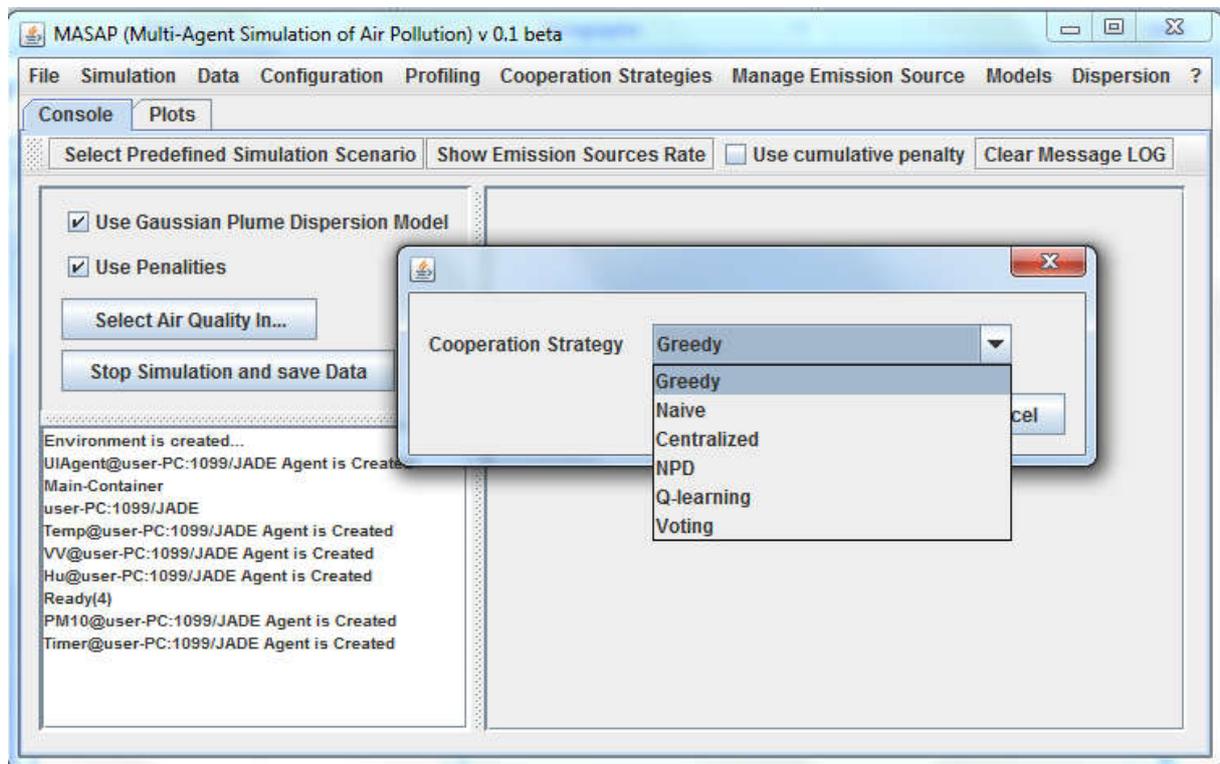
**Figure B.6 : Ecran principal de MASAP**

Le simulateur est conçu d'une façon générique, il offre la possibilité de créer et faire l'apprentissage des modèles de prédiction pour chaque polluant. Comme le montre la figure B.7 :



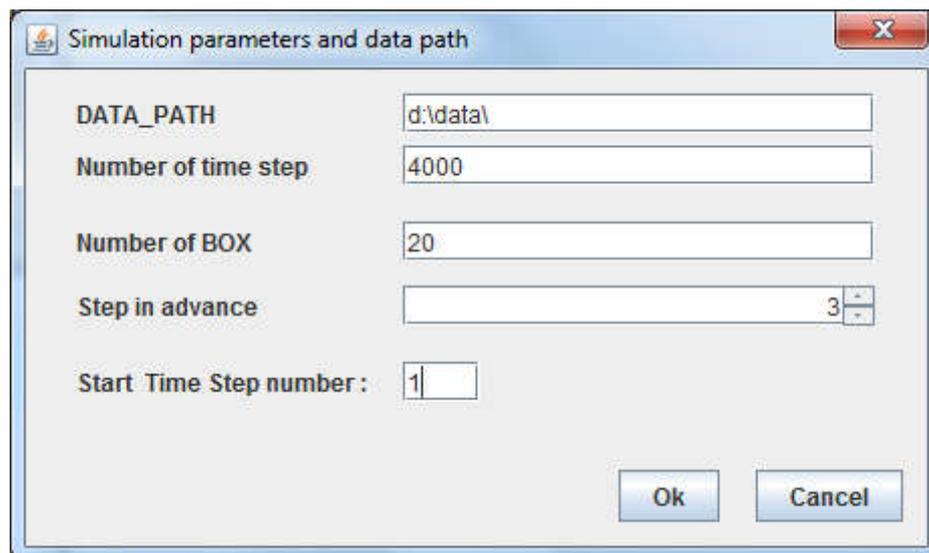
**Figure B.7 : fonctionnalité d'apprentissage des modèles de prédictions RNA.**

Avant de lancer la simulation il est nécessaire de définir la stratégie de coopération qui sera utilisée par les agents qui représentent les contrôleurs des sources d'émission, comme présenté précédemment, ceci peut être paramétré comme illustré dans la figure B.8.



**Figure B.8 : Selection d'une stratégie de coopération.**

Après vient ensuite la configuration de la simulation, à savoir, définir le nombre de pas de la simulation et le repertoire qui contient les fichiers des données. Un exemple est illustré dans la figure B.9.



**Figure B.9 : Configuration d'une simulation.**

Une fois le scénario de la simulation configuré, la simulation peut être lancée, ceci en cliquant sur le bouton « Start Simulation ». L'exécution peut être suivie en utilisant l'interface d'exécution comme le montre la figure B.10.

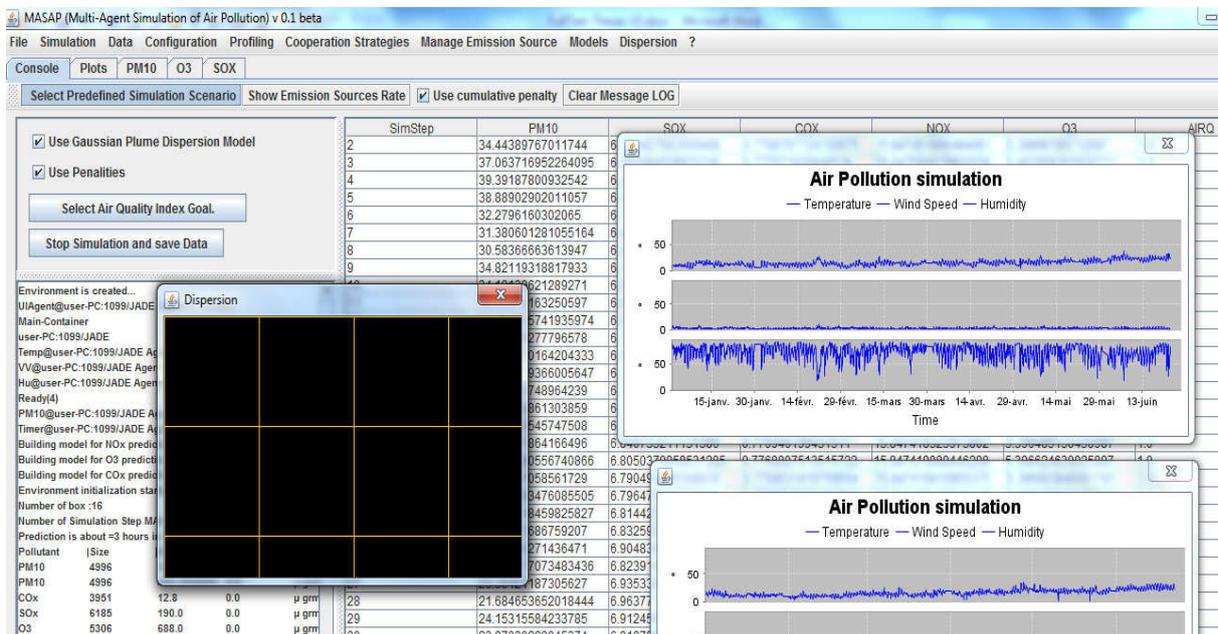


Figure B.10 : Exécution d'un scénario de simulation.

# Bibliographie

- [ Acciani et al. 2006] Acciani G. , Chiarantoni E. , and Fornarelli G. A Neural Network Approach to Study O3 and PM10 Concentration in Environmental Pollution, International Conference on Artificial Neural Networks, Part II, LNCS 4132, pp. 913-922, 2006.
- [İçağa et Sabah 2007] Y. İçağa, E. Sabah, Statistical Analysis of Air Pollutants and Meteorological Parameters in Afyon, Turkey , Environ Model Assess, DOI 10.1007/s10666-008-9139-5, 2007.
- [Ahat et al. 2009] Murat Ahat, Sofiane Ben Amor, Marc Bui, Michel Lamure, Marie-Françoise Courel. Pollution Modeling and Simulation with Multi-Agent and Pretopology. Complex Sciences, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering Volume 4, 2009, pp.225-231.
- [Alam et Geller 2012] Alam, S. J., & Geller, A. (2012). Networks in agent-based social simulation. In Agent-based models of geographical systems (pp. 199-216). Springer Netherlands.
- [Al-Shammari 2013] Al-Shammari, E. T. (2013). Public warning systems for forecasting ambient ozone pollution in Kuwait. Environmental Systems Research, 2(1), 1-9.
- [Andersen et al 1983] Roberts, N., Andersen, D. F., Deal, R. M., Garet, M. S., & Shaffer, W. A. (1983). Introduction to computer simulation: the system dynamics approach. Addison-Wesley Publishing Company.
- [Athanasiadis et Mitkas 2009] Ioannis N. Athanasiadis, Pericles A. Mitkas. A Methodology for Developing Environmental Information Systems with Software Agents .Advanced Agent-Based Environmental Management Systems, Whitestein Series in Software Agent Technologies and Autonomic Computing. 2009, pp 119-137.
- [Aulinas et al. 2009] AULINAS, Montse, TURON, Clàudia, et SÀNCHEZ-MARRÈ, Miquel. Agents as a decision support tool in environmental processes: the state of the art. In : Advanced Agent-Based Environmental Management Systems. Birkhäuser Basel, 2009. p. 5-35.
- [Axelrod 2006] Axelrod, R. (2006). Simulation in the social sciences. Rennard (2006), 90-100.
- [Azid et al. 2014] Azid, A., Juahir, H., Toriman, M. E., Kamarudin, M. K. A., Saudi, A. S. M., Hasnam, C. N. C., Yamin, M. (2014). Prediction of the Level of Air Pollution Using Principal Component Analysis and Artificial Neural Network Techniques: a Case Study in Malaysia. Water, Air, & Soil Pollution, 225(8), 1-14.
- [Balaji et Srinivasan 2010] Balaji, P. G., & Srinivasan, D. (2010). An introduction to multi-agent systems. In Innovations in Multi-Agent Systems and Applications-1 (pp. 1-27). Springer Berlin Heidelberg.
- [Bell et Morse 2013] Bell, S., & Morse, S. (2013). Measuring sustainability: learning from doing. Routledge.
- [Bellifemine et al. 2008] BELLIFEMINE, Fabio, CAIRE, Giovanni, POGGI, Agostino, et al. JADE: A software framework for developing multi-agent applications. Lessons learned. Information and Software Technology, 2008, vol. 50, no 1, p. 10-21.
- [Bendjoudi et al. 2009] Bendjoudi, Z., Taleb, F., Abdelmalek, F., & Addou, A. (2009). Healthcare waste management in Algeria and Mostaganem department. Waste management, 29(4), 1383-1387.

- [Birkin et Wu 2012] Birkin, M., & Wu, B. (2012). A review of microsimulation and hybrid agent-based approaches. In *Agent-based models of geographical systems* (pp. 51-68). Springer Netherlands.
- [Bonabeau 2002a] Bonabeau, E. (2002)a. Predicting the unpredictable. *Harvard Business Review*, 80(3), 109-116.
- [Bonabeau 2002b] Bonabeau, E. (2002)b. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3), 7280-7287.
- [Bordignon et al. 2002] Bordignon S. , Gaetan C. , Lisi F., Nonlinear models for ground-level ozone forecasting, *Statistical Methods & Applications*, vol 11, pp. 227-245, 2002
- [Borri et Camarada 2011] Dino Borri, Domenico Camarda. Planning for the Environmental Quality of Urban Microclimate: A Multiagent-Based Approach. *Cooperative Design, Visualization, and Engineering. Lecture Notes in Computer Science Volume 6874*. 2011, pp. 129-136.
- [Borshchev et Filippov 2004] Borshchev, A., & Filippov, A. (2004, July). From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. In *Proceedings of the 22nd international conference of the system dynamics society* (No. 22).
- [Bossel 2003] Bossel, H. (2003). Assessing viability and sustainability: a systems-based approach for deriving comprehensive indicator sets. *Integrated Natural Resource Management: Linking Productivity, the Environment and Development*, 247-266.
- [Box et al. 1994] Box, G. E. P., Jenkins, G.M., Reinsel, G. C. *Time series analysis: Forecasting and control* (3rd edn.). Englewood Cliffs, NJ: Prentice Hall. 1994
- [Brady et al. 2012] Mark Brady, Christoph Sahrbacher, Konrad Kellermann, Kathrin Happe. An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services, *Landscape Ecology*. November 2012, Volume 27, Issue 9. pp. 1363-1381.
- [Brunelli et al. 2007] Brunelli U., Piazza V., Pignato L., Sorbello F., Vitabile S. Two-days ahead prediction of daily maximum concentrations of SO<sub>2</sub>, O<sub>3</sub>, PM<sub>10</sub>, NO<sub>2</sub>, CO in the urban area of Palermo, Italy , *Environmental Atmospheric environment*, vol. 41, no14, pp. 2967-2995, 2007.
- [Buhmann 2000] Buhmann, M. D. (2000). Radial basis functions. *Acta Numerica* 2000, 9, 1-38.
- [Carmichael et al. 2000] Carmichael Gregory R., Adrian Sandu, Tianfeng Chai, Dacian N. Daescu, Emil M. Constantinescu and Youhua Tang, *Predicting Air Quality: Current Status and Future Directions, Air Pollution Modeling and Its Application XIX*, Business Media B.V., pp. 481- 495, 2008.
- [Castellano et al. 2009] Castellano M. , Franco A., Identification of NO<sub>x</sub> and Ozone Episodes and Estimation of Ozone by Statistical Analysis, *Water Air Soil Pollution* vol 2, 95-110, 2009.
- [Chaib-draa et al. 1992] Chaib-draa, B., Moulin, B., Mandiau, R. & Millot, P. (1992), "Trends in Distributed Artificial Intelligence", *Artificial Intelligence Review* 6, 35-66.
- [Chan et Yao 2008] Chan, C. K., & Yao, X. (2008). Air pollution in mega cities in China. *Atmospheric environment*, 42(1), 1-42.
- [Cigizoglu et al. 2005] CIGIZOGLU, Hikmet Kerem, ALP, Kadir, et KÖMÜRCÜ, Müge. Estimation of air pollution parameters using artificial neural networks. In : *Advances in Air Pollution Modeling for Environmental Security*. Springer Netherlands, 2005. p. 63-75.
- [Cigizoglu et al. 2006] CIGIZOGLU, Hikmet Kerem, ALP, Kadir, et KÖMÜRCÜ, Müge. Two Neural Network Methods in Estimation of Air Pollution Time Series. In : *Environmental Simulation Chambers:*

- Application to Atmospheric Chemical Processes. Springer Netherlands, 2006. p. 421-431.
- [Comrie 1997] Andrew C. Comrie, Comparing Neural Networks and Regression Models for Ozone Forecasting, *Journal of the Air & Waste Management Association*, volume 47, 653-663
- [Conruyt et al. 2009] CONRUYT, Noël, SÉBASTIEN, Didier, COURDIER, Rémy, et al. Designing an Information System for the Preservation of the Insular Tropical Environment of Reunion Island. In : *Advanced Agent-Based Environmental Management Systems*. Birkhäuser Basel, 2009. p. 61-90.
- [Corani 2005] Giorgio Corani, Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning, *Ecological Modelling*, Volume 185, Issues 2-4 , Pages 513-529, 10 July 2005
- [Corchado 2009] Corchado, Juan Manuel, Aitor Mata, and Sara Rodriguez. OSM: A Multi-Agent System for Modeling and Monitoring the Evolution of Oil Slicks in Open Oceans. *Advanced Agent-Based Environmental Management Systems*. Birkhäuser Basel, 2009. pp. 91-117.
- [Crooks et Heppenstall 2012] Crooks, A. T., & Heppenstall, A. J. (2012). Introduction to agent-based modelling. In *Agent-based models of geographical systems* (pp. 85-105). Springer Netherlands.
- [Daly et Zannetti 2007] DALY, Aaron et ZANNETTI, Paolo. An introduction to air pollution—Definitions, classifications, and history. *Ambient air pollution*. P. Zannetti, D. Al-Ajmi and S. Al-Rashied, The Arab School for Science and Technology and The EnviroComp Institute, 2007, p. 1-14..
- [De Veries 2009] DE VRIES, HJM Bert. Environmental Modelling. In : *Principles of Environmental Sciences*. Springer Netherlands, 2009. p. 345-373.
- [Demaze 2009] Tsayem Demaze, M. (2009). Le protocole de Kyoto, le clivage Nord-Sud et le défi du développement durable. *L'Espace géographique*, 38(2), 139-156.
- [Dong et Lu 2010] Dong Qian-jin, Lu Fan
- [Drolling et al. 2005] Ballas, D., Dorling, D., Thomas, B., & Rossiter, D. (2005). Geography matters: simulating the local impacts of national social policies (p. 491). Joseph Rowntree Foundation. doi:10.2307/3650139.
- [elwatan 2015] [http://www.elwatan.com/regions/est/annaba/mauvaise-gestion-et-incivisme-a-annaba-02-03-2015-288749\\_133.php](http://www.elwatan.com/regions/est/annaba/mauvaise-gestion-et-incivisme-a-annaba-02-03-2015-288749_133.php) (visit le 04/03/2015).
- [Feng et al. 2015] Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., & Wang, J. (2015). Artificial neural networks forecasting of PM 2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, 107, 118-128.
- [Ferber et al. 2008] Jacques Ferber, Jean-Louis Dessalles, Denis Phan. Emergence in Agent based Computational Social Science: Conceptual, Formal and Diagrammatic Analysis. *Intelligent Complex Adaptative Systems*, Chapter 9, IGI Global, pp.255-299, 2008, 978-1-59904-717-1.<10.4018/978-1-59904-717-1.ch009>.<lirimm-00344351>.
- [Ferber et Perrot 1995] Ferber, J., Perrot, J. F. (1995). *Les systèmes multi-agents: vers une intelligence collective*. InterEditions.
- [Foxall et al. 2001] Rob Foxall, Igor Krmar, Gavin Cawley, Steve Dorling, Damilo P.Mandic, Nonlinear modelling of air pollution times series, *IEEE* 0-7803-7041, 2001.
- [García-Barrios et al. 2008] García-Barrios, L. E., Speelman, E. N., & Pimm, M. S. (2008). An educational simulation tool for negotiating sustainable natural resource management strategies among stakeholders with conflicting interests. *ecological modelling*, 210(1), 115-126.

- [Gardner et Droling 1999a] Gardner, S. Dorling, Neural network modelling and prediction of hourly NO<sub>x</sub> and NO<sub>2</sub> concentration in urban in London, *Atmospheric Environment*, vol. 33, pp. 709-719, 1999.
- [Gardner et Droling 1999b] Gardner, M.W. and Dorling, S.R, Meteorologically adjusted trends in UK daily maximum ozone concentrations, *Atmospheric Environment*, vol. 34, Issues 2, pages 171-176, 1999.
- [Georgeff et al. 1999]. Georgeff, M., Pell, B., Pollack, M., Tambe, M., & Wooldridge, M. (1999). The belief-desire-intention model of agency. In *Intelligent Agents V: Agents Theories, Architectures, and Languages* (pp. 1-10). Springer Berlin Heidelberg
- [Gerber et al. 2011] Gerber, C., Siekmann, J., & Vierke, G. (2011). Holonic multi-agent systems, research report : [http://scidok.sulb.uni-saarland.de/volltexte/2011/3775/pdf/RR\\_99\\_03.pdf](http://scidok.sulb.uni-saarland.de/volltexte/2011/3775/pdf/RR_99_03.pdf).
- [Ghazi et Khadir 2009] Ghazi, S., and Khadir, M. T., 2009, Recurrent neural network for air pollution peaks prediction for the region of Annaba -Algeria-, *Proceedings of Artificial Intelligence Studies*, 6, pp 81-92.
- [Glaser 2006] Glaser, M. (2006). The social dimension in ecosystem management: strengths and weaknesses of human-nature mind maps. *Human Ecology Review*, 13(2), 122.
- [Grimm et al. 2006] Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., ... & DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological modelling*, 198(1), 115-126.
- [Grimm et al. 2010] Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecological modelling*, 221(23), 2760-2768.
- [Guardiola et al. 2012] Guardiola, E., Natkin, S., Soriano, D., Loarer, E., Vrignaud, P., Boy, T., & Dosnon, O. (2012). Du jeu utile au jeu sérieux (serious game). *Hermès, La Revue*, 62(1), 85-91.
- [Hülsmann et al. 2011] HÜLSMANN, Friederike, GERIKE, Regine, KICKHÖFER, Benjamin, et al. Towards a multi-agent based modeling approach for air pollutants in urban regions. In : *Proceedings of the Conference on "Luftqualität an Straßen. 2011.* p. 144-166.
- [Hülsmann et al. 2013] HÜLSMANN, Friederike, KICKHÖFER, Benjamin, et GERIKE, Regine. Air pollution hotspots in urban areas—How effective are pricing strategies to comply with the EU limits for NO<sub>2</sub>?. *Strategies for Sustainable Mobilities: Opportunities and Challenges, 2013*, p. 105-128.
- [Heaton 2015] Heaton, J. (2015). Encog: Library of Interchangeable Machine Learning Models for Java and C#. *Journal of Machine Learning Research*, 16, 1243-1247.
- [Heppenstall et al. 2011] Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (Eds.). (2011). *Agent-based models of geographical systems*. Springer Science & Business Media.
- [Holling 2001] Holling, C. S. (2001) Understanding the complexity of economic, ecological, and social systems, *Ecosystems*, Vol.4(5), pp.390-405.
- [Holmes et Morawska 2006] Holmes, NS and Morawska, L (2006) A Review of Dispersion Modelling and its application to the dispersion of particles: An overview of different dispersion models available. *Atmospheric Environment* 40(30):pp. 5902-5928.
- [Hooyberghs et al. 2005] Jef Hooyberghs, Clemens Mensink, Gerwin Dumont, Frans Fierens and Olivier Brasseur, A neural network forecast for daily average PM<sub>10</sub> concentrations in Belgium, *Atmospheric Environment*, Volume 39, Issue 18, Pages 3279-3289 , June 2005.

- [Horta et Soares 2010] Horta, A., & Soares, A. (2010). Direct sequential co-simulation with joint probability distributions. *Mathematical Geosciences*, 42(3), 269-292.
- [Iltanen 2012] Iltanen, S. (2012). Cellular automata in urban spatial modelling. In *Agent-based models of geographical systems* (pp. 69-84). Springer Netherlands.
- [Jennings 2000] JENNINGS, Nicholas R. On agent-based software engineering. *Artificial intelligence*, 2000, vol. 117, no 2, p. 277-296.
- [Jorquera et al. 1998] Jorquera Hctor, Ricardo Prez, Aldo Cipriano, Andrs Espejo, M. Victoria Letelier and Gonzalo Acua, Forecasting ozone daily maximum levels at Santiago, Chile, *Atmospheric Environment*, Volume 32, Issue 20, Pages 3415-3424, 25 September 1998.
- [Kelly et al. 2013] Rebecc A. Kelly Letcher, Anthony J. Jakeman, Olivier Barreteau, Mark E. Borsuk, Sondoss ElSawah, Serena H. Hamilton, Hans JÄrgen Henriksen, Sakari Kuikka, Holger R. Maier, Andrea Emilio Rizzoli, Hedwig van Delden, Alexey A. Voinov. Selecting among five common modelling approaches for integrated environmental assessment and management. *Environmental Modelling & Software*. Vol.47, September 2013
- [Khadir et Ghazi 2010] Khadir, M. T. and Ghazi, S., 2010, Radial Basis Function and Elman Networks for Pollutant's Parameter Prediction in the Region of Annaba Algeria .The Fourth International Conference on Advanced Engineering Computing and Applications in Sciences. Florence, Italy, pp 26-33.
- [Kim et Kumar 2005] Kim S. E., Kumar A., Accounting seasonal nonstationarity in time series models for short-term ozone level forecast. *Stoch Environ Res Risk Assess*, vol 19, 241:248, 2005.
- [Kolehmalnen et al. 2000] Kolehmalnen M. , Martikainen H. , Hiltunen T. and Ruuskanen J. , Forecasting air quality parameters using hybrid neural network modeling, *Environmental Monitoring and Assessment*, vol 65, pp. 277â€”286, 2000.
- [Korsakissok 2009] Bourdin-Korsakissok, I. (2009). *Changements d'échelles en modélisation de la qualité de l'air et estimation des incertitudes associées* (Doctoral dissertation, Université Paris-Est).
- [Li et al. 2012] Li, Weiqian, XIE, Jiancang, LI, Jianxun. Three Dimensional Simulation of Basin Water Pollution Incidents Based on Multi-agent and Grid Technology. In : *Computational Intelligence and Design (ISCID)*, 2012 Fifth International Symposium on. IEEE, 2012. p. 470-473.
- [LKHOO et Meng 2010] LKHOO, H. L., & Meng, Q. (2010). Modeling of Vehicle Emission Pricing Strategy Using Multi-agent System. *Journal of the Eastern Asia Society for Transportation Studies*, 8, 883-895.
- [Müller et al. 2013] MÜLLER, Birgit, BOHN, Friedrich, DREßLER, Gunnar, et al. Describing human decisions in agent-based models–ODD+ D, an extension of the ODD protocol. *Environmental Modelling & Software*, 2013, vol. 48, p. 37-48.
- [Macal 2010] Macal, C. M. (2010, December). To agent-based simulation from system dynamics. In *Simulation Conference (WSC), Proceedings of the 2010 Winter* (pp. 371-382). IEEE.
- [Macal et North 2010] Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of simulation*, 4(3), 151-162.
- [Maidstone 2012] Maidstone, R. (2012). *Discrete Event Simulation, System Dynamics and Agent Based Simulation: Discussion and Comparison*. System, 1-6.
- [Malkar 1997] Malkar P., Josef S., *Determination of Features for Air Pollution Forecasting Models*, 3rd ed. International Conference on Intelligent Information Systems, IEEE, 1997
- [Marks 2007] Robert Ernest Marks (2007). *Validating Simulation Models: A General Framework and Four*

- Applied Examples, *Comput Econ* (2007) 30:265-290 DOI 10.1007/s10614-007-9101-7.
- [Maslow 1966] Abraham H. Maslow (1966). *The Psychology of Science*. p. 15
- [Molina et Molina 2004] Mario J. Molina & Luisa T. Molina (2004) Megacities and Atmospheric Pollution, *Journal of the Air & Waste Management Association*, 54:6, 644-680, DOI:10.1080/10473289.2004.10470936.
- [Moustris et al. 2013] Moustris, K. P., Larissi, I. K., Nastos, P. T., Koukouletsos, K. V., & Paliatsos, A. G. (2013). Development and Application of Artificial Neural Network Modeling in Forecasting PM10 Levels in a Mediterranean City. *Water, Air, & Soil Pollution*, 224(8), 1-11.
- [MURAKAMI et al. 2010] MURAKAMI, Kazutaka, SASAI, Takahiro, et YAMAGUCHI, Yasushi. A new one-dimensional simple energy balance and carbon cycle coupled model for global warming simulation. *Theoretical and applied climatology*, 2010, vol. 101, no 3-4, p. 459-473.
- [Musters et al. 1998] Musters, C. J. M., De Graaf, H. J., & Ter Keurs, W. J. (1998). Defining socio-environmental systems for sustainable development. *Ecological Economics*, 26(3), 243-258.
- [Natalja et al. 2011] Pulter, Natalja, Heiko Schepperle, and Klemens Böhm. How agents can help curbing fuel combustion: a performance study of intersection control for fuel-operated vehicles. *The 10th International Conference on Autonomous Agents and Multiagent Systems.-Volume 2. International Foundation for Autonomous Agents and Multiagent Systems*. 2011.
- [Neagu et al. 2006] Ciprian-Daniel Neagu, Nikolaos Avouris, Vasile Palade, *Neural and Neuro-Fuzzy Integration in a Knowledge-Based System for Air Quality Prediction*, Kluwer Academic Publishers, Applied Intelligence, vol 17, pp. 141-169, 2002
- [Niazi et Hussain 2011] Niazi, M., & Hussain, A. (2011). Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics*, 1-21. doi: 10.1007/s11192-011-0468-9.
- [Nwana 1996] Nwana, H. S. (1996). Software agents: An overview. *The knowledge engineering review*, 11(03), 205-244.
- [O'Sullivan et al. 2012] O'Sullivan, D., Millington, J., Perry, G., & Wainwright, J. (2012). Agent-Based Models-Because They're Worth It?. In *Agent-based models of geographical systems* (pp. 109-123). Springer Netherlands.
- [OCDE 1997] OCDE (1997), <http://www.oecd.org/fr/environnement/envech/2386739.pdf> (visité le 28/04/2016)
- [OMS 2014] OMS 2014, <http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/> (visité en 27/04/2015).
- [ONS 2015] ONS 2015, Rapport de l'office national des statistiques sur la natalité, <http://www.ons.dz/IMG/pdf/demog2013.pdf> [visité le 20/02/2015].
- [Oprea et al. 2011] Oprea, M., Dragomir, E., & Carbureanu, M. (2011, October). On the use of collaborative intelligence in an agent-based environmental monitoring and analysis system. In *System Theory, Control, and Computing (ICSTCC), 2011 15th International Conference on* (pp. 1-6). IEEE.
- [Ordieres et al. 2005] J. B. Ordieres, a, E. P. Vergaraa, R. S. Capuzb and R. E. Salazarc, Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in ElPaso, *Environmental Modelling and Software*, Volume 20, Issue 5, Pages 547-559, May 2005.
- [Papaleonidas et Andonis Papaleonidas, Lazaros Iliadis. Hybrid and Reinforcement Multi Agent Technology for

- [Iliadis 2012] Real Time Air Pollution Monitoring. Artificial Intelligence Applications and Innovations IFIP Advances in Information and Communication Technology Volume 381, 2012, pp. 274-284.
- [Park 2013] Park, C. C. (2013). Acid rain: Rhetoric and reality. Routledge.
- [Parrott 2011] Parrott, L. (2011). Hybrid modelling of complex ecological systems for decision support: Recent successes and future perspectives. *Ecological Informatics*, 6(1), 44-49.
- [Parunak et al. 1998] Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998, January). Agent-based modeling vs. equation-based modeling: A case study and users guide. In *Multi-agent systems and agent-based simulation* (pp. 10-25). Springer Berlin Heidelberg.
- [Pedell et Sterling 2011] Sonja Pedell, Leon Sterling. Agent-Based Modelling for Understanding Sustainability. Agents in Principle, Agents in Practice, Lecture Notes in Computer Science Volume 7047, 2011. pp. 398-409.
- [Pereira et al. 2009] Pereira, Antonio, Lucas Paulo Reis, and Pedro Duarte. EcoSimNet: A Multi-Agent System for Ecological Simulation and Optimization. *Progress in Artificial Intelligence*. Springer Berlin Heidelberg, 2009. pp 473-484.
- [Perez et al. 2000] Patricio Perez , Trier, A., Reyes, J, Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile., *Atmospheric Environment* 34, 1189-1196,2000.
- [Perez et Reyes 2002] Perez P., Reyes J., Prediction of maximum of 24-h average of PM10 concentrations 30 h in advance in Santiago, Chile, *Atmospheric Environment* 36 4555-4561, 2002.
- [Power 2009] Conrad Power, A Spatial Agent-Based Model of N-Person Prisoner's Dilemma Cooperation in a Socio-Geographic Community, *Journal of Artificial Societies and Social Simulation* vol. 12, no. 18, (2009), <http://jasss.soc.surrey.ac.uk/12/1/8.html>.
- [Rahman et al. 2014] Rahman, N. H. A., Lee, M. H., & Latif, M. T. (2014). Artificial neural networks and fuzzy time series forecasting: an application to air quality. *Quality & Quantity*, 1-15.
- [RFI 2015] RFI, article sur le pic de pollutin à Téhéran, <http://www.rfi.fr/moyen-orient/20150101-iran-teheran-pollution-environnement-voitures-motos/> [visité le 20/02/2015]
- [Ristic et al. 2015] Ristic, B., Gunatilaka, A., Gailis, R., & Skvortsov, A. (2015). Bayesian likelihood-free localisation of a biochemical source using multiple dispersion models. *Signal Processing*, 108, 13-24.
- [Russell et Norvig 1995] Russell, S., Norvig, P., & Intelligence, A. (1995). A modern approach. *Artificial Intelligence*. Prentice-Hall, Egnlewood Cliffs, 25.
- [Russo et Soares 2014] Russo, A., & Soares, A. O. (2014). Hybrid model for urban air pollution forecasting: A stochastic spatio-temporal approach. *Mathematical Geosciences*, 46(1), 75-93.
- [Schmitt et Pumain 2013] Schmitt, C., & Pumain, D. (2013). Modélographie multi-agents de la simulation des interactions sociétés-environnement et de l'émergence des villes. *Cybergeog: European Journal of Geography*.
- [Schreinemachers et Berger 2011] Pepijn Schreinemachers, Thomas Berger. An agent-based simulation model of human-environment interactions in agricultural systems. *Environmental Modelling & Software*. Volume 26, Issue 7, July 2011. pp. 845:859.
- [Scott 2008] Moss, Scott (2008). 'Alternative Approaches to the Empirical Validation of Agent-Based Models'. *Journal of Artificial Societies and Social Simulation* 11(1)5

- <<http://jasss.soc.surrey.ac.uk/11/1/5.html>>.
- [Sharma et Kaushik 2008] Sharma P., Chandra A. Kaushik S., Forecasts using Box-Jenkins models for the ambient air quality data of Delhi City. *Environ Monit Assess*, vol 8, pp. 520-528, 2008
- [Shoham 1993] Shoham, Y. (1993). Agent-oriented programming. *Artificial intelligence*, 60(1), 51-92.
- [Shuiyuan et al. 2006] CHENG, Shuiyuan, LI, Jianbing, FENG, Beng, et al. A gaussian-box modeling approach for urban air quality management in a northern chinese city—I. model development. *Water, Air, and Soil Pollution*, 2007, vol. 178, no 1-4, p. 37-57.
- [Siebers et al. 2010] Siebers, P. O., Macal, C. M., Garnett, J., Buxton, D., & Pidd, M. (2010). Discrete-event simulation is dead, long live agent-based simulation!. *Journal of Simulation*, 4(3), 204-210.
- [Smajgl et al. 2011] SMAJGL, Alex, BROWN, Daniel G., VALBUENA, Diego, et al. Empirical characterisation of agent behaviours in socio-ecological systems. *Environmental Modelling & Software*, 2011, vol. 26, no 7, p. 837-844.
- [Sokolova et Fernández-Caballero 2012] SOKOLOVA, Marina V. et FERNÁNDEZ-CABALLERO, Antonio. Evaluation of environmental impact upon human health with DeciMaS framework. *Expert Systems with Applications*, 2012, vol. 39, no 3, p. 3469-3483.
- [Sousaa et al. 2007] SOUSA, S. I. V., MARTINS, F. G., ALVIM-FERRAZ, M. C. M., et al. Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*, 2007, vol. 22, no 1, p. 97-103.
- [Stork et Hassibi 1993] Stork, D. and Hassibi, B., Second order derivatives for network pruning: Optimal Brain Surgeon, In T. J. Sejnowski G. E. Hinton and D. S. Touretzky, editors *Advances in Neural Information Processing Systems (NIPS)*, 5, pages 164-171, 1993.
- [Tadjine et al. 2008] Tadjine, A., Djébar, H., & Courtois, A. (2008). Toxicité des poussières rejetées par le complexe sidérurgique d'Annaba sur quelques paramètres hématologiques du lapin Européen. *Environnement, Risques & Santé*, 7(3), 209-215.
- [Touza et al. 2013] Touza, J., Drechsler, M., Smart, J. C., & Termansen, M. (2013). Emergence of cooperative behaviours in the management of mobile ecological resources. *Environmental Modelling & Software*, 45, 52-63.
- [Viennet 2006] VIENNET, Emmanuel. Réseaux à fonctions de base radiales. *Apprentissage connexionniste*, 2006, p. 105.
- [Voukantsis et 2011] VOUKANTISIS, Dimitris, KARATZAS, Kostas, KUKKONEN, Jaakko, et al. Intercomparison of air quality data using principal component analysis, and forecasting of PM 10 and PM 2.5 concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Science of the Total Environment*, 2011, vol. 409, no 7, p. 1266-1276.
- [Watkins et al. 2013] Watkins, Cristy, MASSEY, Dean, BROOKS, Jeremy, et al. Understanding the Mechanisms of Collective Decision Making in Ecological Restoration: An Agent-Based Model of Actors and Organizations. *Ecology & Society*. 2013, vol. 18, no 2.
- [WHO 2005] WHO (World Health Organisation), *Ecosystems and Human Well-being: Health Synthesis*. WHO Library Cataloguing-in-Publication Data, 2005.
- [Wolfram 2002] Wolfram, S. (2002). *A new kind of science (Vol. 5)*. Champaign: Wolfram media.

[Wooldridge et  
Jennings 1995]

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(02), 115-152.