



**Conseil Economique  
et Social**

Distr  
RESTREINTE

EB.AIR/WG.1/R.120  
17 avril 1996

FRANCAIS  
Original : ANGLAIS

COMMISSION ECONOMIQUE POUR L'EUROPE

ORGANE EXECUTIF DE LA CONVENTION SUR LA POLLUTION  
ATMOSPHERIQUE TRANSFRONTIERE A LONGUE DISTANCE

Groupe de travail des effets  
Quinzième session, Genève, 3-5 juillet 1996  
(Point 5 d) de l'ordre du jour provisoire)

RESEAUX NEURONAUX ARTIFICIELS :  
NOUVELLE METHODE DE MODELISATION INFORMATIQUE POUR PREDIRE  
LES EFFETS DE L'OZONE SUR LES CULTURES ET LES VEGETAUX NON LIGNEUX \*/

I. INTRODUCTION

1. Chaque année, les expériences du Programme international concerté relatif aux effets de la pollution atmosphérique et d'autres phénomènes perturbateurs sur les cultures et les végétaux non ligneux (PIC-Cultures) génèrent des ensembles de données importants sur les effets de la pollution par l'ozone

---

\*/ Etabli par le Centre de coordination du Programme international concerté relatif aux effets de la pollution atmosphérique et d'autres phénomènes perturbateurs sur les cultures et les végétaux non ligneux, à l'Université de Nottingham Trent, Royaume-Uni.

La distribution des documents établis sous les auspices ou à la demande de l'Organe exécutif de la Convention sur la pollution atmosphérique transfrontière à longue distance est RESTREINTE et réservée aux gouvernements et organisations qui participent aux travaux de l'Organe exécutif. Les documents ne doivent pas être communiqués aux journaux ou revues, à moins que la RESTRICTION n'ait été LEVEE par l'Organe exécutif.

(essentiellement l'apparition de lésions visibles et la diminution des rendements ou de la biomasse sur les cultures et les espèces de végétaux non ligneux sensibles. La taille même de ces ensembles en rend l'analyse par des méthodes statistiques traditionnelles difficiles. Par contre, les réseaux neuronaux artificiels (RNA) conviennent parfaitement car ils permettent de traiter des ensembles de données importants et incomplets. Ils peuvent aussi détecter les formes présentes dans les ensembles de données et, partant, sont un outil utile pour l'examen des données relatives aux conditions climatiques et aux polluants associées aux effets de la pollution par l'ozone. Les RNA peuvent donc déterminer l'influence des facteurs du degré II (par exemple le climat) sur la réaction des végétaux à l'ozone, ce qui peut permettre d'entrevoir comment ces facteurs peuvent agir sur les niveaux critiques pour l'apparition de lésions visibles et la baisse des rendements.

2. Au cours de la dernière décennie, des progrès sensibles ont été faits en matière de reconnaissance des formes et d'appréhension des calculs grâce notamment au développement de réseaux neuronaux artificiels qui s'acquittent de fonctions analogues à certaines fonctions cérébrales (Kothari et Heekuck, 1993). Les RNA ne sont pas tributaires d'un programme dont le fonctionnement exige des procédures par paliers. Ils "apprennent" effectivement par l'étude des structures présentes dans des ensembles de données, ce qui n'est pas sans ressembler au fonctionnement du cerveau humain. Ils peuvent aussi traiter des images ou des formes "bruitées" ou incomplètes, autre caractéristique fondamentale du cerveau humain (Kothari et Heekuck, 1993). A l'instar des systèmes biologiques, les réseaux neuronaux sont composés de nombreux éléments de traitement, connectés, appelés neurones ou noeuds. Ils sont idéalement adaptés à l'analyse de données obtenues d'expériences où nombre de paramètres ont des incidences sur la réponse. Par exemple, l'étendue des lésions provoquées par l'ozone peut être modifiée par des facteurs tels que l'humidité relative, la température et le rayonnement photosynthétiquement actif (RDA) puisqu'ils influent sur la conductance stomatale (Kappen et Haeger, 1991) et, partant, sur le flux d'ozone dans les végétaux (Kersteins et al., 1992).

3. On peut affirmer que les RNA offrent le moyen d'étudier les caractéristiques présentes au sein des données disponibles de manière à pouvoir faire des prévisions pour de nouveaux ensembles de données (Funahashi, 1989; Hornik et al., 1989). Initialement, le réseau doit recevoir une formation à l'aide de nombreux éléments de données (le module d'apprentissage), lui permettant de déceler les formes présentes dans les données. La couche réceptrice reçoit des données qui représentent les agents causals d'un effet alors que les neurones de la couche émettrice contiennent les données qui représentent l'effet "réel". A partir des données en entrée, le réseau calcule une valeur de sortie prédictive qu'il compare avec la valeur "réelle" entrée dans la couche émettrice. Le réseau apporte ensuite les modifications qui rapportent la sortie prédictive à la sortie "réelle". Il le fait en modifiant constamment la pondération des connexions entre les neurones de la couche réceptrice et de la couche émettrice jusqu'à ce que le processus cognitif soit achevé et que le réseau puisse générer un état de sortie qui soit une interprétation suffisamment précise et cohérente des entrées pour tous les modules d'apprentissage. Le réseau peut être mis à l'épreuve en introduisant de nouvelles données (choisies indépendamment) dans les neurones de la couche réceptrice puis en comparant la réponse prédictive de sortie à une réponse "réelle". La puissance des connexions du réseau peut être analysée

de manière à indiquer l'importance relative des paramètres introduits dans les neurones de la couche réceptrice (Balls et al., 1996). Il est en outre possible d'extraire des équations permettant de justifier et d'interpréter le comportement du réseau (Roadknight et al., sous impression). Ces équations et, partant, le réseau peuvent servir à prédire l'apparition de lésions dans différentes conditions de climat et de pollution.

4. La démarche fondée sur l'utilisation de RNA contraste avec les méthodes paramétriques et statistiques traditionnellement employées pour traiter de problèmes de modélisation analogues. Ainsi, le modélisateur peut partir d'une "page blanche" et n'est donc pas tributaire d'une connaissance a priori des données. A partir des données, le RNA induit un modèle sans biais en trouvant, pour les données à l'étude, une configuration non linéaire qui minimise l'erreur de VQM (valeur quadratique moyenne). Cette démarche permet parfois de jeter un éclairage nouveau sur les données que l'on modélise et, aussi, de quantifier des relations bien comprises.

## II. OBJECTIFS

5. L'objectif de ce projet de modélisation est de créer des modèles informatiques permettant de prédire les cas de lésion due à l'ozone et de déterminer les conditions dans lesquelles l'ozone et le climat entraînent l'apparition de lésions et une diminution de rendement ou de la biomasse. Les données trouvent leur source dans les expériences du PIC-Cultures et les paramètres des modèles sont les facteurs de pollution et/ou d'environnement.

## III. METHODOLOGIE

6. Les données utilisées pour l'apprentissage et la mise à l'épreuve des RNA ont été collectées sur des sites répartis dans toute l'Europe, dans le cadre du PIC-Cultures de la CEE/ONU. Les pays ayant participé en 1994 et/ou 1995 sont les suivants : Allemagne, Autriche, Belgique, Danemark, Espagne, Fédération de Russie, Finlande, France, Grèce, Hongrie, Italie, Lettonie, Pays-Bas, Pologne, Royaume-Uni, Slovaquie, Suède et Suisse. Les données sont collectées annuellement, ce qui signifie que les réseaux mis au point et à l'épreuve en utilisant des données provenant d'une ou deux années peuvent être ultérieurement modifiés ou renforcés en ajoutant de nouvelles données annuelles. L'apprentissage effectué avec des données provenant de nombreux sites européens sur de nombreuses années débouche sur des réseaux prédictifs exceptionnellement robustes.

7. En 1994 et 1995, les participants au PIC-Cultures ont enregistré la première apparition de lésions sur le trèfle souterrain et sur le trèfle blanc (Trifolium subterraneum et T. repens) et sur les haricots (Phaseolus vulgaris). Initialement, on a utilisé les données provenant des expériences portant sur le trèfle car ce végétal est un bon indicateur des lésions imputables à l'ozone (Pihl Karlsson et al., 1995). En outre, ces expériences ont fourni l'ensemble de données le plus important.) Pour les données de 1994, trois indicateurs de l'ozone (moyenne sur sept heures, maximum quotidien, cumul des moyennes horaires de l'ozone, où la moyenne est supérieure à 40 ppb, et/ou total quotidien), sur des durées de trois ou de cinq jours précédant l'apparition de la lésion, ainsi que l'âge de la feuille, ont été retenus comme paramètres de réseau.

8. Un ensemble de données très complet a été obtenu en 1995 et aussi bien les données climatiques que les données sur l'ozone ont été utilisées pour l'apprentissage d'autres réseaux à la prédiction des lésions. Pour chacune des cinq journées précédant l'apparition de la lésion, les paramètres ci-après ont été utilisés comme intrants au RNA : ozone (AOT40 (exposition cumulative au-delà d'un seuil de 40 ppb), moyenne sur 7 heures, maximum quotidien), température (moyenne sur 24 heures, moyenne sur 7 heures, maximum quotidien), rayonnement global (moyenne sur 24 heures, moyenne sur 7 heures, maximum quotidien) et humidité relative (moyenne sur 24 heures, moyenne sur 7 heures et maximum quotidien moyen) et nombre de jours après l'apparition.

#### IV. RESULTATS

9. La figure I donne la variation des concentrations quotidiennes d'ozone. Toutes les valeurs des entrées destinées au réseau sont normalisées de manière que la plus grande soit égale à 1 et la plus petite à zéro. La figure II indique l'architecture du réseau dont l'apprentissage s'est effectué en utilisant simplement des paramètres relatifs à l'ozone. Le RNA comporte 10 unités d'entrée; chacune tire sa valeur des niveaux d'ozone ou de l'âge de la feuille. Les trois unités cachées sont des détecteurs de caractéristiques et les différences de répartition dans la pondération des interconnexions signifient que chaque unité cachée examine les données en privilégiant des aspects différents. Le noeud de sortie saisit les résultats des unités cachées et donne un total.

10. L'aptitude du RNA à prédire les résultats désirés gît dans les pondérations des connexions. Le réseau présenté à la figure II s'est avéré très fiable pour prédire l'apparition de lésions foliaires. Un algorithme simple (une suite séquentielle qui résout une question) a été mis au point pour identifier les combinaisons d'entrées les plus significantes à partir des pondérations des connexions. Il consiste à :

- a) Identifier, dans le réseau, les pondérations qui franchissent un seuil donné;
- b) Créer une équation en adaptant les noeuds d'entrée et les pondérations approximatives d'entrée aux paramètres d'entrée auxquels sont associées des pondérations désignées;
- c) Mettre à l'essai le réseau partiellement connecté pour ce qui est de l'équation extraite de manière à confirmer le niveau de précision.

11. L'algorithme a produit des équations qui ont donné une représentation simplifiée du RNA, ont sauvegardé les caractéristiques principales du modèle soumis à l'apprentissage et expliqué le comportement du réseau (Roadknight et al., 1995). Les termes de l'équation ci-après sont des paramètres de polluants. Par exemple, la moyenne des valeurs des niveaux d'ozone relevées chaque heure entre 10 heures et 17 heures au cours du jour 3 se lit "Jour3(moyennesur7heures)". Les caractéristiques clefs de chaque unité cachée sont placées entre crochets.

**Apparition de la lésion =**

$$\begin{aligned} &F[\text{Jour2}(\text{moyenne sur 7h} + \text{Max}) = \text{Jour3}(\text{AOT40} + \text{OzoneMax})] + \\ &F[\text{Jour2}(\text{OzoneMax} - \text{AOT40}) - \text{Jour1}(\text{AOT40} + \text{moyenne sur 7h}) + \\ &\text{Jour3}(\text{moyenne sur 7h})] \end{aligned}$$

12. Le premier noeud, le plus influent, semble indiquer que le passage de niveaux élevés à de faibles niveaux d'ozone précède le déclenchement de la lésion étant donné qu'un signe négatif est associé aux paramètres de l'ozone pour le jour 3. Cette équation peut être encore simplifiée pour se lire comme suit :

**Apparition de la lésion = F[augmentation des niveaux d'ozone] + F[chute des niveaux d'ozone]**

13. Le fait qu'une chute des concentrations d'ozone précède l'apparition de la lésion a été confirmé par des équations tirées d'un RNA dont l'apprentissage s'est effectué avec un module de formation plus important qui comportait des données sur la saison de pousse 1995. Cette constatation a fait l'objet d'examens plus poussés en procédant à des expériences en enceinte close avec T. subterraneum. Elles ont révélé de nouveaux éléments de compréhension de l'apparition des lésions sur les feuilles de trèfle. A hautes doses, il est certes possible que l'apparition de lésions nécrosiques soit immédiate mais il peut en être autrement à des niveaux inférieurs, plus naturels. L'apparition de lésions nécrosiques sur les feuilles de trèfle à faible niveau (< 100ppb) n'intervient pas directement après l'exposition mais est précédée par une inflammation "humide" de la feuille. Cette inflammation fait partie de la réaction de la plante à la présence d'ozone mais sa durée dépend des niveaux d'ozone et des conditions climatiques. Il se peut aussi que la transformation en nécrose ne soit pas inéluctable et que, dans certaines conditions, la partie enflammée redevienne saine. Il est prévu de procéder à d'autres essais pour préciser encore les conditions d'ozone et de climat nécessaires au déclenchement de l'inflammation et à l'apparition d'une lésion foliaire nécrosique.

14. Le réseau alimenté par des données relatives au climat et à l'ozone reçoit un apprentissage permettant d'obtenir une grande précision prédictive (fig. III). Il pourrait prédire exactement le déclenchement d'une lésion à partir d'un ensemble réduit de données d'essai (fig. IV). Le résultat obtenu du RNA est porté sur l'axe Y et comparé avec les résultats désirés. Lorsque la lésion apparaissait après le jour 5, le résultat désiré était de 0,9 contre 0,1 en l'absence de lésion. Le réseau ayant reçu un apprentissage donnait un résultat prédictif aussi proche que possible du résultat désiré. La valeur prédictive croît au fur et à mesure que le RNA juge qu'il est de plus en plus probable qu'une lésion apparaisse.

15. L'équation ci-après a été extraite du RNA dont l'apprentissage a fait appel à des données sur l'ozone et le climat et, même si la simplification est considérable par rapport aux 204 connexions pondérées de l'ensemble du réseau, elle reste à la fois longue et protéiforme :

Apparition de la lésion =

$$F[\text{Jour2}(\text{Luminosité Max-AOT40}) - \text{Jour1}(\text{HR moyenne sur 7h}) - \\ \text{Jour3}(\text{AOT40} + \text{Ozone moyenne sur 7h}) - \text{Jour4}(\text{AOT40}) + \\ \text{Jour5}(\text{Luminosité moyenne sur 24h} + \text{Luminosité Max})]$$

$$+F[\text{Jour3}(\text{OzoneMax}) - \text{Jour1}(\text{HR moyenne sur 7h}) - \text{Jour4}(\text{HR moyenne sur 7h} \\ + \text{HRMax}) - \text{Jour5}(\text{AOT40} + \text{Ozone moyenne sur 7h} + \text{HR moyenne sur 7h})]$$

$$+F[\text{Jour2}(\text{HR sur 24h} + \text{Luminosité sur 24h}) + \\ \text{Jour3}(\text{AOT40} + \text{Ozone moyenne sur 7h} + \text{OzoneMax}) - \text{Jour5}(\text{AOT40}) - \\ \text{Jour1}(\text{HR Max})]$$

16. Cette équation exclut que la température joue un rôle quelconque en tant que facteur modifiant et met en évidence le rôle majeur de l'humidité et de la luminosité. Elle fait aussi apparaître qu'une chute de niveau d'ozone précède d'un jour la première apparition de la nécrose. La relation entre l'ozone et l'humidité reste complexe, même au sein de cette équation simplifiée. Ce fait est démontré lorsque le réseau est mis en présence d'une gamme de combinaisons artificielles d'ozone et d'humidité (fig. V). Le caractère non linéaire des données ressort suffisamment bien des courbes ondulées représentant la réponse du modèle à toute une gamme de valeurs d'entrée.

17. La simple analyse des pondérations combinées des connexions qui correspondent à chaque donnée introduite dans le modèle permet de dégager une impression d'ensemble de l'importance relative des différents facteurs. Si l'on extrait les pondérations des interconnexions du réseau dont l'apprentissage s'est fait avec l'ozone, les conditions climatiques et le nombre de jours, il apparaît que la concentration d'ozone (AOT40), l'humidité relative (moyenne sur 7 h) et la luminosité (moyenne quotidienne max) sont des intrants importants en ce qui concerne le développement des lésions (tableau). Par exemple, l'AOT40, l'humidité relative (HR) (moyenne sur 7 h) et le rayonnement global (maximum quotidien moyen) ont respectivement des valeurs de 1,49, 1,03 et 0,93. Le nombre de jours après apparition de la lésion ainsi que la température ont une importance moindre étant donné que les valeurs sont inférieures à 0,6.

## V. CONCLUSIONS

18. Les travaux de modélisation effectués jusqu'ici montrent que les équations tirées de RNA ayant reçu un apprentissage peuvent, dans des grands ensembles de données, révéler des facteurs importants jusqu'alors indétectés. Ces facteurs (ou combinaison de facteurs) peuvent sinon rester indécélables du fait du volume élevé des données à variables multiples qu'il faut examiner et analyser. Les résultats obtenus d'une modélisation de RNA ont permis d'établir que la concentration d'ozone était importante dans les jours précédant les lésions dont l'apparition pouvait être conditionnée par une chute de la concentration d'ozone. Des expériences préliminaires en enceinte close ont établi qu'exposées à l'ozone, les feuilles de trèfle laissaient apparaître une inflammation "humide", la lésion nécrotique n'intervenant qu'après une chute des concentrations d'ozone. Ces découvertes seront étayées par des expériences ultérieures en enceinte close. Les RNA ont aussi indiqué que l'humidité

relative et la luminosité, outre les concentrations d'ozone, étaient des facteurs importants de l'apparition de lésions. Il se peut que cela soit dû à l'influence de l'humidité relative et de la lumière sur la conductance stomatale et, partant, sur le flux d'ozone dans la plante. Il s'agit là d'une information importante car elle peut contribuer à une approche (degré 2) du niveau critique à court terme pour l'apparition de lésions et expliquer la variabilité des concentrations d'ozone susceptible de provoquer l'apparition de lésions.

#### VI. TRAVAUX ACTUELS ET FUTURS

19. Les travaux actuels ont pour but d'explorer des méthodes permettant de déterminer et de caractériser les codépendances entre les paramètres d'entrée qui peuvent être, par exemple, deux ensembles différents de conditions environnementales provoquant des lésions : forte humidité relative/faible concentration d'ozone et faible humidité relative/forte concentration d'ozone. Des modèles détaillés sont en cours d'incorporation sur interface d'ordinateur convivial à fenêtres.

20. Les données pour 1994 et 1995 serviront à l'apprentissage de RNA pouvant prédire les baisses de rendement en fonction de l'augmentation des concentrations d'ozone. Jusqu'ici, les travaux ont été axés sur les lésions foliaires étant donné qu'il s'agit d'une conséquence visible des effets néfastes de l'ozone se prêtant bien aux opérations de mesure. Il est également possible d'utiliser les RNA pour établir un lien entre les lésions foliaires et la baisse des rendements.

#### VII. REFERENCES \*\*/

**Balls G.R., Palmer-Brown D. et Sanders G.E.** (1996) Investigating microclimatic influences on ozone injury in clover (Trifolium subterraneum) using artificial neural networks. New Phytologist, 132, 271-280.

**Funahashi K.** (1989) On the approximate realization of continuous mapping by neural approximators. Neural Networks, 2, 183-192.

**Hornik K., Stinchcombe M. et White H.** (1989) Multilayer feedforward networks are universal approximators. Neural Networks, 2, 359-366.

**Kappen L. et Haeger S.** (1991) Stomatal responses of Tradescantia albiflora to changing air humidity in light and darkness. Journal of Experimental Botany, 42, 979-986.

**Kersteins K., Federholzner R. et Lendzian K.J.** (1992) Dry deposition and cuticular uptake of pollutant gases. Agriculture, Ecosystems and Environment, 42, 239-253.

**Kothari S.C et Heekuck O.H.** (1993) Neural Networks for Pattern Recognition. In: Advances in Computers, 37, 119-166.

---

\*\*/ Cette notice bibliographique a été reproduite telle qu'elle a été reçue par le secrétariat.

**Pihl Karlsson G.P., Sellden G., Skarby L. et Pleijel H.** (1995) Clover as an indicator plant for phytotoxic ozone concentrations: visible injury in relation to species, leaf age and exposure dynamics. New Phytologist, 129, 355-363.

**Roadknight C., Palmer-Brown D. et Sanders G.E.** (1995) Learning the equations of data. Proceedings of the 3rd Annual SNN Symposium on Artificial Neural Networks, 253-256.



Figure I. Variabilité des concentrations quotidiennes d'ozone observées en 1994 sur un site expérimental en Belgique. Les données proviennent d'une expérience effectuée avec Trifolium subterraneum

Figure II. Structure d'un réseau neuronal artificiel pour la prédiction des lésions

Figure III. Prédiction, par le RNA, des lésions foliaires pour les éléments du module d'apprentissage

Figure IV. Prédiction, par le RNA, des lésions foliaires pour les éléments du module d'essai

Figure V. Relation complexe entre l'ozone et l'humidité relative telle qu'exprimée par les résultats obtenus du RNA

Tableau. Importance relative des paramètres d'entrée du réseau neuronal dont l'apprentissage a été effectué avec : concentrations d'ozone, paramètres climatiques et numéro de jour

Paramètre d'entrée	Importance relative (unités arbitraires)
Ozone (AOT40)	1,49
Ozone (moyenne sur 7 h)	1,14
Ozone (max. quotidien moyen)	1,10
HR (moyenne sur 24 h)	0,90
HR (moyenne sur 7 h)	1,03
HR (max. quotidien moyen)	0,75
Rayonnement global (moyenne sur 24 h)	0,62
Rayonnement global (moyenne sur 7 h)	0,69
Rayonnement global (max. quotidien moyen)	0,93
Numéro de jour	0,59
Température (moyenne sur 24 h)	0,41
Température (moyenne sur 7 h)	0,56
Température (max. quotidien moyen)	0,42

-----