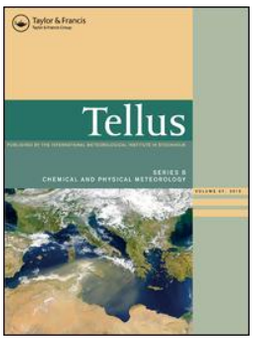
Soil NO emissions modelling using artificial neural network

Claire Delon, Dominique Serça, Christophe Boissard, Richard Dupont, Alain Dutot, Patricia Laville, Patricia de Rosnay, Robert Delmas



Tellus B: Chemical and Physical Meteorology

ISSN: (Print) 1600-0889 (Online) Journal homepage: https://www.tandfonline.com/loi/zelb20

Soil NO emissions modelling using artificial neural network

Claire Delon, Dominique SerçA, Christophe Boissard, Richard Dupont, Alain Dutot, Patricia Laville, Patricia De Rosnay & Robert Delmas

To cite this article: Claire Delon, Dominique SerçA, Christophe Boissard, Richard Dupont, Alain Dutot, Patricia Laville, Patricia De Rosnay & Robert Delmas (2007) Soil NO emissions modelling using artificial neural network, Tellus B: Chemical and Physical Meteorology, 59:3, 502-513, DOI: 10.1111/j.1600-0889.2007.00254.x

To link to this article: https://doi.org/10.1111/j.1600-0889.2007.00254.x

Published online: 18 Jan 2017.

Tellus (2007), 59B, 502–513
Printed in Singapore. All rights reserved

TELLUS

Soil NO emissions modelling using artificial neural network

By CLAIRE DELON[1]*, DOMINIQUE SERÇA[1], CHRISTOPHE BOISSARD[2], RICHARD DUPONT[1], ALAIN DUTOT[2], PATRICIA LAVILLE[3], PATRICIA DE ROSNAY[4] and ROBERT DELMAS[1], [1]*Laboratoire d'Aérologie, 14 avenue E. Belin, 31400 Toulouse, France;* [2]*Laboratoire Interuniversitaire des Systèmes Atmosphériques, 94010 Créteil, France;* [3]*INRA, EGC, 78830 Thiverval-Grignon, France;* [4]*CESBIO, 31400 Toulouse, France*

(Manuscript received 27 April 2006; in final form 25 October 2006)

ABSTRACT

Soils are considered as an important source for NO emissions, but the uncertainty in quantifying these emissions worldwide remains large due to the lack of field experiments and high variability in time and space of environmental parameters influencing NO emissions. In this study, the development of a relationship for NO flux emission from soil with pertinent environmental parameters is proposed. An Artificial Neural Network (ANN) is used to find the best non-linear regression between NO fluxes and seven environmental variables, introduced step by step: soil surface temperature, surface water filled pore space, soil temperature at depth (20–30 cm), fertilisation rate, sand percentage in the soil, pH and wind speed. The network performance is evaluated each time a new variable is introduced in the network, i.e. each variable is justified and evaluated in improving the network performance. A resulting equation linking NO flux from soil and the seven variables is proposed, and shows to perform well with measurements ($R^2 = 0.71$), whereas other regression models give a poor correlation coefficient between calculation and measurements ($R^2 \leq 0.12$ for known algorithms used at regional or global scales). ANN algorithm is shown to be a good alternative between biogeochemical and large-scale models, for future application at regional scale.

1. Introduction

Nitric oxide (NO) emissions from soils represent an important part of total (anthropogenic plus biogenic) NO emissions (around 40%, an amount comparable to fossil fuel combustions) (Davidson and Kingerlee, 1997; Delmas et al., 1997).

NO is produced in the soil upon microbial processes referred to as nitrification and denitrification. The rate of nitrification and denitrification depends on the type of soil and on the nutrient content, and the exact role of each process in producing NO has been proven difficult to assess (Conrad, 1996).

NO emissions from soils have been shown to be influenced by soil water content and soil temperature. Indeed, most studies of NO inventories use parameterizations elaborated with these two variables, associated with the rate of fertilization (Yienger and Levy, 1995) or the nitrogen content (Potter et al., 1996; Van Dijk and Meixner, 2001).

*Corresponding author.
e-mail: delc@aero.obs-mip.fr
DOI: 10.1111/j.1600-0889.2007.00254.x

Whereas soil temperatures fluctuations can explain short-term variations of NO fluxes, soil moisture changes are often responsible for seasonal variations of the flux (Meixner and Yang, 2004). The importance of soil temperature has been firstly considered, mostly in temperate soils (Williams et al., 1992; Martin et al., 1998). Soil moisture influence was emphasized in tropical soils, introducing the notion of pulse effect (Johansson et al., 1988; Yienger and Levy, 1995), but the same kind of soil moisture effect may occur when fertilization is applied in temperate soils before rain events. The type of soil and climate are essential factors to be taken into account, as they will drive the evolution of water content and temperature. Increasing temperature was shown to increase NO emission, (Williams et al., 1992; Martin et al., 1998), whereas in some particular tropical conditions, the daily variation of fluxes with temperature could not be highlighted (Cardenas et al., 1993; Serça et al., 1994). Furthermore, diurnal changes in NO fluxes from tropical soils were shown to have better chance to be observed during dry seasons rather than during wet season (Butterbach-Bahl et al., 2004). Above a certain threshold in water filled pore space, the emission increases (Otter et al. 1999; Meixner and Yang, 2004).

Moreover, a lot of studies have also carefully considered the effect of fertilization rate in the case of addition of mineral or organic fertilizer (Sanhueza et al., 1990; Shepherd et al., 1991; Skiba et al., 1992), due to biomass burning (Anderson et al., 1988; Serça et al., 1998), or in relation to nitrification and/or denitrification processes (Le Roux et al., 1995; Parsons et al., 1996). In the majority of these studies, pH and soil texture have also been taken into account to explain soil NO emission rates.

Once emitted by soils, NO is quickly converted to NO_2 near the surface, by chemical reactions with ozone, within minutes. A large part of this NO_2 is then deposited on plants and soils in a process referred to as Canopy Reduction Factor process (CRF), by stomatal or surface uptake (Horii et al., 2004). Consequently, only a fraction of NO emitted by soils, 20 to 70% depending on the Leaf Area Index, (Yienger and Levy, 1995; Ganzeveld et al., 2002) actually reaches the lower atmosphere in the form of NOx ($NO + NO_2$). NOx global emissions range from 4.7 Tg N yr^{-1} (10^{12} grams of Nitrogen per year, Müller, 1992) to 13 Tg N yr^{-1} (Davidson and Kingerlee, 1997) after CRF application, with other estimations giving 5.45 Tg N yr^{-1} (Yienger and Levy, 1995) and 4.97 Tg N yr^{-1} (Yan et al., 2005).

pH, texture and associated vegetation have been worldwide identified as major factors controlling NO emissions but, due to their high-variability in time and space, have not yet been generalized in a common relation. Indeed, uncertainties in quantifying NO emissions from soils rely on the difficulty of taking into account all the influent parameters together.

Other parameterisations exist at a continental or regional scale to reproduce NO emissions from soils, like inverse modelling from satellite mapping reported in Bertram et al. (2005) and Jaeglé et al. (2004), based on a mechanistic approach. These studies give a good insight of the pulse effect, consecutive to biomass burning or fertilizers application, but are not directly related to soil parameters as water content or texture. Most of the chemistry transport models employ the Yienger & Levy (1995) parameterization for emissions quantification at a global scale (Ludwig et al., 2001; Bertram et al., 2005, and references therein), but some large discrepancies have been found between estimates and actual emissions at the regional scale.

Therefore, the use of recent biogeochemical models has allowed the construction of NO inventories in Europe (Li et al., 2000; Butterbach-Bahl et al., 2001; Kesik et al., 2005) and in Australia (Kiese et al., 2005) with the PnET-N-DNDC process oriented model, based on Geographic Information System (GIS) databases, but the application of such biogeochemical model needs detailed field experiments in the region where it is applied to correctly initialise the model, which is not the case worldwide, and specifically not for African tropical soils. Estimates of past, current and future emissions were proposed through different types of modelling studies, but remain only adapted to specific regions and specific types of land cover, as reported by Reiners et al. (2002) in Costa Rica with the adapted CENTURY model, by Bouwman et al. (2002b) for fertilized fields, and by Roelle et al. (2002) for quantifying the impact of specific fertilizers on NO emissions and atmospheric chemistry in North Carolina (USA). It appears that most of these modelling studies concern human disturbed soils (by biomass burning, or fertilizers application), whereas bare natural soils have been scarcely investigated.

The common goal of all the studies related to NO flux parameterization is to choose the most important parameters that enable to reduce the uncertainties in NO emission assessment whatever the type of soil and climate, for a good representation of underlying processes. All the difficulty relies in choosing these most important parameters. Linking these parameters is therefore essential in order to achieve the most accurate approach, and this will be tested in this study.

A method of parameterization is applied here, the neural network technology, in order to find the best non-linear regression between a list of selected environmental parameters and NO emission fluxes. To our knowledge, no previous study using this approach has been applied to determine NO emissions from soils.

Some parameters possibly involved into NO emissions are tested. Each parameter in the neural network is considered at a time, and the performance of the model is assessed for each new added parameter.

The following methodology is proposed: the neural network approach is presented first, as well as the criteria used to select the best non-linear regressions. In order for different situations to be represented, the neural network needs to be supplied with data issued from diverse types of climates and soils: the databases used are then presented and discussed, and a final parameterization is proposed, including the relevant variables. The results obtained by the network are presented and discussed, each time a new parameter is introduced: the first model uses only soil surface temperature and Water Filled Pore Space (WFPS), the second, the temperature at depth (20–30 cm), then the fertilization rate, the sand percentage, the pH and finally the wind speed are introduced. The choice and relevance of these parameters are discussed at each step.

2. Methodology

2.1. Neural network

Artificial Neural Network (ANN) tools have appeared as alternatives to classical statistical modelling in many disciplines, and are particularly useful for non-linear phenomena. Networks of artificial neurons are designed to be able to learn how to represent complex information.

The type of ANN used in this study is the Multi Layer Perceptron (MLP). MLP is the most widely used in atmospheric science because it interconnects neurons, representing non-linear mapping between input vectors and an output vector (Gardner and Dorling, 1998; Hsieh and Tang, 1998; Dreyfus et al., 2002). The objective of the MLP is to find the equation which links the input variables X to the output variable Y with a mathematical function

f chosen among different others.

$$Y = f(X). \quad 2.1.$$

The set of data used by the MLP is separated in two, the training set and the validation set. X and Y are matrices representing the training data, which must fully represent the widest range of cases about which network is required to generalize. The training process will determine a set of optimal weights, which will then be applied on the validation set that has not participated to their elaboration. The root mean squared error (RMSE) is calculated for both sets. The closer the RMSE for training and validation sets are, the more accurate the model is. In other words, a small training RMSE means that the output gets very precisely through all the training points. Small RMSE difference and small RMSE values for both training and validation sets ensures that the output will closely fit the training and validation points, and will also give realistic results between these points, avoiding over fitting process. To assess whether or not the model is over fitted, the generalization cost is calculated. The generalization cost represents the capacity of the model to interpolate between points, i.e. to change from a discrete function to a continuous one. The lowest it is, the better the model is. The generalization performance is tested within the validation set (Gardner and Dorling, 1999).

In this study, the training set was constituted of 380 examples (or lines), and the validation set of 250 examples constituted with hourly means of each variable. Note that, each sample is considered as time independent of each other by the network. Both sets were determined so that the partition ensures a same statistical distribution between the two, by calculating the Kullback–Leibler distance (Kullback and Leibler, 1951; Kullback, 1959). The cross validation score is then calculated based on this partitioning (this score increases when the model is over fitted. When this increase appears, the calculation procedure is automatically stopped in the software, and the model with the smallest RMSE is chosen).

Compared to other studies, the database contains a sufficient number of examples to provide enough data (630 lines) for both learning and validation of the network, regarding the number of inputs parameters (7) and the number of hidden neurons (3) (630:7:3 structure). Indeed, Dutot et al. (2003) has constructed his database with a 177:17:3 structure, Yi and Prybutok (1996) have chosen a 122:9:4 network structure, and Navone and Ceccato (1994) a 119:7:4 structure. Other studies can of course be constructed on larger databases (Gardner and Dorling, 1999). A rule may be extracted from the Vapnik–Chervenenkis theory (Vapnik, 1995), which stipulates that the learning dimension of the network has to be 3–10 times greater than the number of input parameters multiplied by the number of hidden neurons. In our case, this theory is verified ($380 > 3^*7^*10$). However, the non-observation of this rule does not discredit an ANN study.

The architecture of the MLP depends on the number of neurons. Choosing the number of neurons is determinant to avoid over fitting (over fitting induces an augmentation of the cross validation score, as mentioned above). Not enough neurons will result in not reaching the closest solution, whereas too many of them will rend the solution noisy. After several tests, the number of three hidden neurons was chosen to be introduced in the MLP, which represents a balanced situation regarding to the two previously mentioned cases. The schematic architecture of the network is given in Fig.1.

The neural network used in this study was based on a commercial version of the Neuro One 5.0 © software, (Netral, Issy les Moulineaux, France).

In this study, the maximal number of epochs (or the number of modifications in weight values), used in the optimization, or back propagation algorithm, was 100. Ten initializations (or 10 series of different sets of weights) were tested for each model. This configuration (100 epochs, 10 models) was tested several times, in order to avoid a local minimum solution. The transfer, or activation function (the mathematical function), was the hyperbolic tangent: the choice was made by comparing results obtained with hyperbolic tangent and arctangent functions, revealing better results with hyperbolic tangent (according to the selection criteria mentioned below). The network was used in its

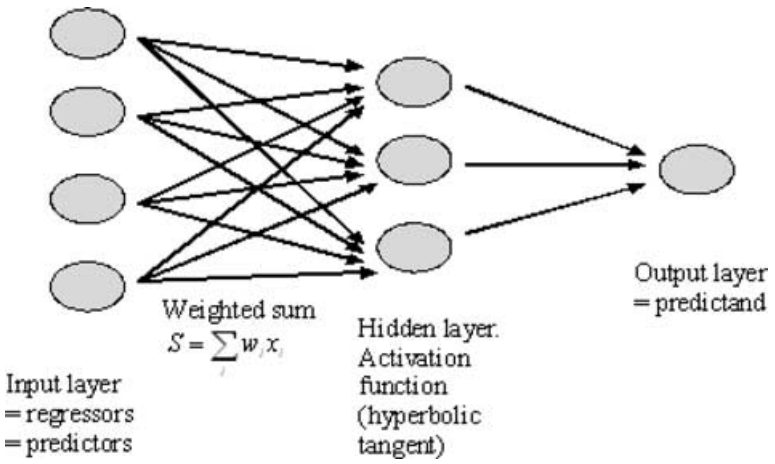


Fig. 1. Schematic MLP used for the simulation, adapted from Dutot et al. (2002).

static version where examples of the database are independent of each other.

All inputs and output were normalized and centred in order to avoid artefact in the training process. After normalization, data have the same order of magnitude. Without this step, and in case of very different orders of magnitude between variables, small ones may have artificially lower influence during the training.

The best algorithm within the 10 launched was chosen, by assessing the following three criteria:

The lowest generalization cost was chosen

RMSE of the training set had to be close to the RMSE of the validation set

Both RMSE had to be as small as possible

In order to ensure the best result of the network, the back propagation algorithm was used, and is summarised as follows:

1. The weights were initialized to random and weak values,
2. Input data were given as first examples to the network,
3. An output was calculated,
4. Calculated and measured outputs were compared,
5. The error between each pair of outputs was back propagated through the network,
6. Weights were adjusted and
7. All these steps were repeated until the smallest error was found.

2.2. *Database*

The database used was made up with four sets of data obtained from four different sites representing a range of different climates, soils conditions and land use, in order to acquire the greatest information and to reproduce the largest panel of fluxes.

The four sites were Auradé (South–West of France), Meyrargues (South–East of France), Grignon (North of France) and Hombori (Centre of Mali).

Auradé is the biggest database with 79% of all the data. Due to its size, one could speculate that the network will be influenced only by this database. Obviously, the global database is mostly representative of temperate conditions, but the presence of tropical data (even in a lower proportion) is a way to include contrasted influences in the database, and the network will take into account all contributions. The availability of other tropical databases in the future will help to balance the partition temperate/tropical data.

Each database is composed of the following parameters: NO flux, fertilization rate (total amount of fertilizer expressed in Nitrogen Unit, spread out every hour by an exponential decay law, as detailed further), temperature at soil surface (0–5 cm) and at soil depth (20–30 cm), WFPS at surface, pH, soil texture and wind speed (see below for discussion concerning the choice of these parameters).

(i) A comprehensive approach of NO flux measurements can be found in Serça et al., 1998. Stainless steel chambers (15-cm high) covering a surface area of 800 cm^2 were used for NO flux measurements. Stainless steel frames were inserted into the ground 3 to 8 hr before the measurements in order to prevent immediate disturbances of air diffusion from soil as well as long-term effect on fluxes. A mass balance calculation was applied to the soil-chamber system, and the NO flux rate was computed from the slope of the initial linear increase in NO concentration in the chamber (Davidson, 1991; Serça et al., 1994), with following relationship:

$$F_{NO} = \left(\frac{\Delta C}{\Delta t}\right)\left(\frac{V M_{\mathrm{N}}}{ART}\right), \qquad 2.2.$$

where F_{NO} is the NO emission flux (ng N m^{-2} s^{-1}), ($\frac{\Delta C}{\Delta t}$)is the initial rate of increase in NO concentration calculate by linear regression (ppb s^{-1}), V is the chamber volume (cubic centimetres, M_{N} is the nitrogen molecular weight (grams), A is the sampling area (square centimetres), R is the gas constant (cm^3 atm mol^{-1} K^{-1}), and T is the air temperature in the chamber. Pressure was assumed to be constant throughout the flux measurement and equal to ambient pressure.

NO concentration in the chamber was measured using a ThermoEnvironment® 42 CTL analyser. This analyser detects NO by chemiluminescence with O_3. Detection limit and sensitivity is around 0.05 ppbv. Flow rate in the analyser and the chamber is about 1 l min^{-1}. Multipoint calibration was checked before and after each field experiment with a dynamical calibration system. This system is made of two Bronkhost® mass flow controllers (range 0–20 ml min^{-1} and 0–20 ml min^{-1}) allowing to produce NO concentrations in the range 10–200 ppbv from certified portable cylinders of NO gas at 10 ppmv concentration (Air Liquide®).

The four databases are presented in Table 1. Once included in the network, all data from the four databases were centred and normalized.

(1) WFPS were deduced from Volumetric Humidity (VH) measurements. These measurements were performed with time-domain reflectometry (TDR) CS616 Campbell ™ probes in Auradé, Meyrargues and Grignon (French sites) and Theta-probes ML2x in Hombori.

(2) The volumetric humidity of soils is deduced from the probes signal in volts, proportional to the dielectric constant of the soil. Therefore, WFPS is calculated following WFPS = 100*VH/(1 − d_a/d_r), where d_a is the bulk density (dry soil mass/total volume) and d_r is the density of solid particles (dry soil mass/dry soil volume).

(3) Soil texture values were measured by the Institut National de Recherche Agronomique (INRA, Arras, France) following the AFNOR X31 107 standard.

(4) French sites (Grignon, Auradé and Meyrargues) have been artificially fertilised, and the fertilization rate is given in the database in Nitrogen Units per hour (the total amount of fertilizer is expressed in Nitrogen Unit, spread out every hour by an

Table 1. Description of the constitution of the database used in that study

Location	Auradé	Grignon	Hombori	Meyrargues
Country	SW France	N France	Mali	SE France
Associated publication	This work	Laville et al. (2005)	This work	This work
Latitude	43.57°N	48.85°N	15.2°N	43.39°N
Longitude	1.07°E	1.96°E	1.41°W	5.32°E
Period of sampling	May–June 2003	Year 2002	June–July 2004 Transition period (dry /wet season)	June 2001
Crop type	Wheat field	Wheat/Maïze	Bare soil	Maïze field
Number of lines (or examples)	498	56	47	28
% of the total data base	79	9	7.5	4.5
Range of NO emissions (gN ha^{-1} d^{-1})	0.4–35	0.2–36	1–32	4–63

exponential decay law, the quasi total quantity (90%) being assimilated in 22 d approximately (Parton et al., 2001)). Hombori has not received any mineral fertilization, but it has been shown that pastures in West Africa receive an important organic fertilization (Schlecht et al., 1997; Schlecht and Hiernaux, 2004). Nitrogen assimilation in this kind of tropical soil has been rarely studied, and as a first approximation, the same exponential decay law than in temperate soils was applied in Hombori, based on a nitrogen amount calculated from the above cited literature. This is of course a source of uncertainties in calculating NO emissions and this question will have to be specifically addressed in future work.

2.3. Resulting equation

The following algorithm, used to estimate the NO flux from the seven variables described below is as follows:

$$\mathrm{NOflux_{norm}} = w_{24} + w_{25}.\tanh(S_1) + w_{26}.\tanh(S_2) + w_{27}.\tanh(S_3), \quad (1)$$

where $\mathrm{NOflux_{norm}}$ is the normalized NO flux, and

$$S_1 = w_0 + \sum_{i=1}^{7} w_i x_{j,\mathrm{norm}}$$

$$S_2 = w_8 + \sum_{i=9}^{15} w_i x_{j,\mathrm{norm}} \qquad \text{with} \qquad \mathrm{j} = 1 \rightarrow 7\,,$$

$$S_3 = w_{16} + \sum_{i=17}^{23} w_i x_{j,\mathrm{norm}}$$

where $\mathrm{x_1}$ to $\mathrm{x_7}$ correspond to surface WFPS, surface soil temperature, deep soil temperature, fertilization rate, sand percentage, pH and wind speed, respectively.

All weights w_i are given in Table 4.

Weights w_0, w_8, w_{16} and $\mathrm{w_{24}}$ were linked to the bias neuron (constant term equal to 1).

NO flux was finally calculated in $\mathrm{gN.ha^{-1}.d^{-1}}$ using:

$$\mathrm{NOflux} = (\mathrm{NOflux_{norm}} - \Sigma_s)/\Pi_\mathrm{s}, \quad (2).$$

where

$$\Pi_\mathrm{s} = \frac{\sqrt{N}}{\sqrt{\sum_k (\mathrm{NOflux^k})^2 - \left(\sum_k \mathrm{NOflux^k}\right)^2 \Big/ N}},$$

$$\Sigma_s = -\Pi_\mathrm{s} \frac{\sum_k \mathrm{NOflux^k}}{N},$$

where N is the total number of examples, and k corresponds to the kth example.

3. Results and discussion

Results are presented in a sequential manner, where parameters are added one after the other and in the following order: soil surface temperature and surface WFPS at first, then soil temperature at depth, fertilization rate, pH, percentage of sand and finally wind speed. The results are independent of the order of introduction in the network (not shown here). The relevance of each newly included parameter is discussed. Fig. 2 and Table 3 give an overview of all criteria with systematic comparison to the 1:1 slope.

3.1. Soil surface temperature and WFPS

The neural network was run with NO flux as output, and WFPS and soil temperature at surface as inputs. Among the 10 different model runs, the one that had a generalisation cost of 6.49 was the best. RMSE obtained for training set and validation set were close (19.1 and 16.6, respectively).

Fig. 2 (Case 1) shows the comparison between calculated fluxes and measured NO fluxes. The coefficient of determination showed that 45% of variance may be explained by considering soil surface temperature and WFPS influences only. The slope of the model versus the experimental data is 0.45, and is compared to the 1:1 slope. It has to be mentioned here that all slopes of Fig. 2 are less than one. There is a systematic underestimation of the model versus measurements, a bias partially corrected when

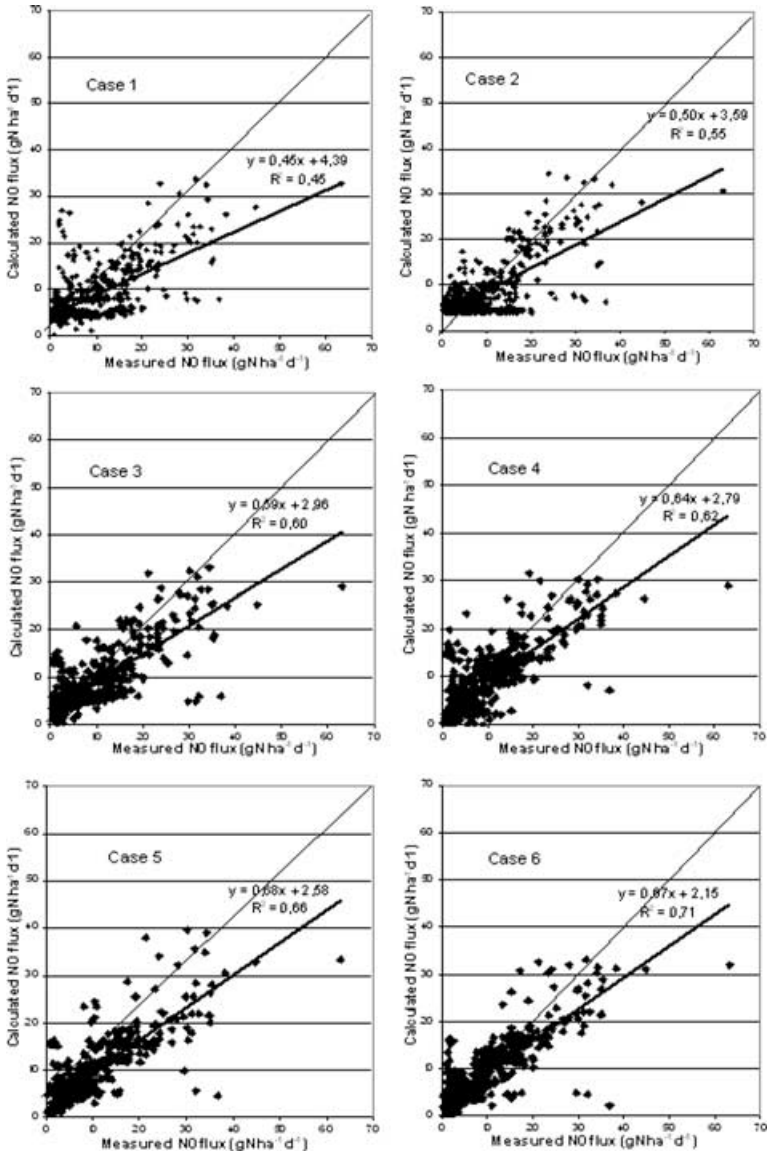


Fig. 2. Measured versus calculated NO fluxes in gN. ha^{-1}. d^{-1}: Case 1: surface temperature and WFPS, Case 2: Case 1 + deep soil temperature (20–30 cm), Case 3: Case 2 + fertilization rate, Case 4: Case 3 + sand percentage, Case 5: Case 4 + pH, Case 6: 0Case 5 + wind speed.

other parameters are added. This is due to the fact that high fluxes in the database are less represented than low fluxes, and to the fact that the network has a better capacity in representing mean values than extreme ones.

Fig. 3 (case 1) is the time series of calculated and measured fluxes together, for the Auradé set only. All low fluxes at the beginning of the series are overestimated by a factor of 3. The measured flux variations showed a bipartite organisation: low (weekly) frequency variation, and high (daily) frequency variation. The low-frequency signal was correct, but daily variations failed in numerous parts of the graph. This comparison in daily variations between measured and calculated fluxes was only possible for Auradé set, as 24 hr a day monitoring was not performed in the other campaigns. NO calculated flux decreased too rapidly between days 132 and 134 compared to the measured one. (The gradual decrease in measurements is due to the exponential decrease in fertilization rate, and cannot be retrieved in calculated values because only WFPS and temperature at surface are considered here).

One can conclude that soil surface temperature and WFPS alone do not succeed in representing adequately NO emissions from soil.

3.2. Soil temperature at depth (20–30 cm)

The influence of soil temperature on NO emissions does not systematically give an indication of the depth where processes of emission are found. It has been shown that surface temperature plays a great role in microbial processes involving nitrification and denitrification processes, but the role of soil temperature below the surface remains to be investigated. It has been suggested

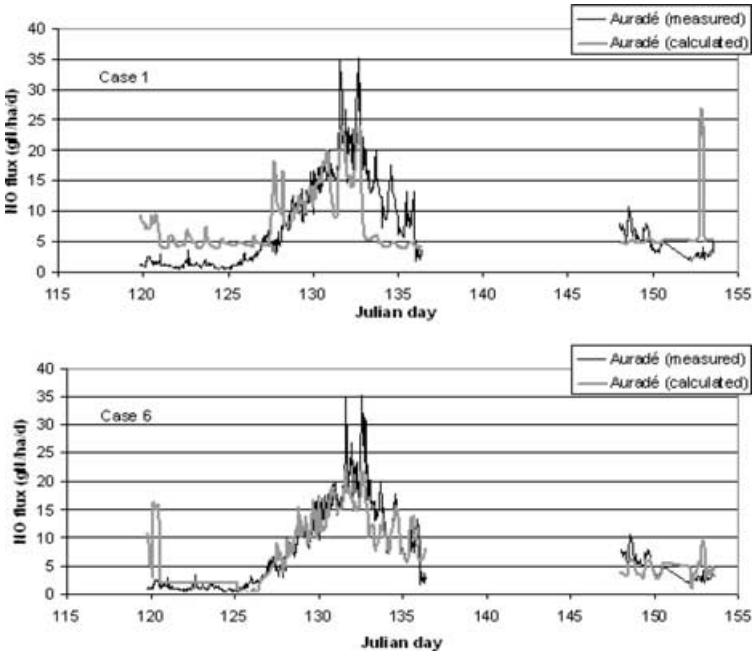


Fig. 3. Measured and calculated NO fluxes in case 1 (NO = f(surface temperature and WFPS)) and case 6 (NO = f(surface temperature and WFPS, deep soil temperature, fertilization rate, sand %, wind speed), function of Day of Year, for Auradé set only.

that primary production and consumption zones for NO are located within 0.01 to 0.1m in the soil column (Yang and Meixner, 1997) but Butterbach-Bahl et al. (2004) pointed out that litter layer and mineral soil layer are not stimulated at the same time, so that N mineralization might occur at different depths, depending on root depth in the case where plants are present, and/or rainfall intensity and rainfall seasonal variation.

Furthermore, the oxygen diffusion in the soil column depends on the spatial and temporal heterogeneity of vegetation cover, topographic position and soil texture (Austin et al., 2004), which determines the distribution of resources availability and soil organisms.

In this context, the deep soil temperature between 20—and 30 cm was added as input in the neural network in order to test if deeper soil phenomena are significantly influencing NO emissions.

The addition of this new variable led to a generalization cost of 5.23. RMSE obtained for training and validation set were, respectively, 11.8 and 20.5. This difference is greater than in the previous model, but both values remained in the same order of magnitude, thus preventing from any over fitting. A higher value of variance was explained (55%) between calculated and measured fluxes (Fig. 2, case 2), and the slope was 0.5. Considering all these criteria, it was possible to conclude that globally the network performance was improved.

The comparison between calculated and measured fluxes (not shown here) for Auradé set is quite similar to Fig. 3 (case 1): calculated flux were governed by a threshold effect, and did not follow the measured signal in its decrease between days 132 and 134, showing that significant parameters are missing.

3.3. Fertilization rate

Nitrogen content is linked to the microorganisms within the soil and to the nitrogen input (natural and/or anthropogenic), which are in part responsible for the rate of gaseous emission at the surface. Nitrogen content has not been measured in the diverse databases used, but it can be considered that natural nitrogen content becomes negligible when fertilization is applied. In agricultural fields, it is easy to quantify the nitrogen source, as nitrogen was here brought only under mineral fertilizer form (urea, nitrate, ammonium), and the total quantity is assimilated through an exponential decay law. In temperate fertilized fields, the total amount of mineral fertilizer was known, and easy to convert in nitrogen content.

In Hombori site, it was more difficult to quantify the nitrogen source, but it was however taken into account by estimating the manure input, also assimilated with an exponential decay. Estimation of the manure input in terms of nitrogen content was deduced from Schlecht et al. (1997).

The addition of this variable allowed a new improvement of the network performance as the generalization cost obtained was 5.25. RMSE values were 11.2 and 16.5 for training and validation sets, respectively. The model thus explained 60% of the calculated flux as shown in Fig. 2 (case 3).

Low-fluxes were better represented, and calculated fluxes around 5 gN. ha^{-1}. d^{-1} in cases 1 and 2 have disappeared, showing a better distribution of calculated fluxes. Nevertheless, higher fluxes remain underestimated, and introducing other parameters in the description may help reducing that problem.

Table 2. pH and texture values for the soils of the database described in Table 1

	Auradé	Grignon wheat	Grignon maize	Hombori	Meyrargues
pH	8.2	7.9	8	7	8.3
% sand	24	16	17.4	93	16
% clay	40	33.2	33.5	2.2	30

At French sites, pH and texture were measured in situ. In Hombori, pH value is derived from Diallo and Gjessing (1999), and texture values were measured in situ.

3.4. Sand percentage

Soil texture appeared to be an important feature for emissions through its link with water diffusion. Roelle et al. (2001) reported that each soil type has a range of soil moisture that optimizes NO flux, leading to a maximal emission. Indeed, it has been shown that diffusion of O_2 in soil drives the microbial activity, which in turn defines molecular diffusion and transport of NO in soil pores (Meixner and Yang, 2004).

However, the correlation between moisture optima for NO emission and soil texture is not obvious at least for temperate soils (Schindlbacher et al. 2004). Parsons et al. (1996) have established a parabolic relation between WFPS and NO emissions in sandy savanna soils, and Austin et al. (2004) reported a strong interaction between texture and pulsed rainfall events in arid and semi-arid regions, suggesting a difference of behaviour between fine-textured and coarse-texture soils in nutrient turnover.

The databases contained very different values of soil texture (see Table 2), with a typical sandy Sahelian soil, with more than 90% of sand, and three temperate clay-loamy soils. The range in sand percentage variation is very interesting for the network to learn information from very distinct situations. Only the sand percentage information was used in the network, as the sum of the silt-clay values are strongly anti-correlated to it.

The introduction of sand percentage allowed a new improvement of the network performance. The explained variance was 62 % (Fig. 2, case 4) and the slope was 0.64. Generalization cost obtained was 5.64, and RMSE were 11.5 and 14.4, respectively, for training and validation sets. The generalization cost was not better than in previous examples, but remained the best among the runs performed with sand percentage.

3.5. pH

Serça et al. (1994) have shown the importance of pH influence on NO emissions by artificially increasing pH of acidic soils, which leads to an immediate decrease in emissions. pH conditions can influence NO emissions through chemodenitrification process (low pH) or biological activity (higher pH). Supporting this hypothesis, Ormeci et al. (1999) have shown that NO emissions increase for pH < 5 and pH > 8, with chemodenitrification processes being generally much more efficient for emissions. The importance of pH has also been highlighted in Yan et al. (2005) by using a statistical model, showing a negative correlation between NO emissions and soil pH.

It has been shown that some particular pH conditions (neutral conditions, associated with coarse texture and good soil drainage, as found in Sahelian soils) may increase NO emissions, without being a key control parameter (Bouwman et al. 2002a).

Generalization cost obtained was 4.87, and RMSE were 9.2 and 14.7, respectively, for training and validation sets. The calculated flux versus the measured flux in this configuration gave an explained variance of 66%, a significant improvement when compared to the previous case. Although pH values fluctuate over a small range (7–8.3; see Table 2) in the database, NO emissions appeared to be significantly influenced by this parameter. However, the network will not able to reproduce cases very different from those used in the learning process. In future network development, it would be however crucial to enlarge the panel of pH values (for acidic soils in particular), in order to apply the equation in such extreme situations.

3.6. Wind speed

Wind speed has not been yet reported as influent for soil NO emissions. In this study, wind speed was measured at 2 m height, whereas flux measurements are collected near the soil in enclosed chambers. A priori, wind speed should not have any influence on soil fluxes in enclosed chambers, but it is a way to represent the state of the atmosphere at a given time, and soil fluxes in the chamber are certainly influenced by what happens around the chamber, specifically by atmospheric pressure, air moisture, air temperature and turbulence. The influence of wind speed on NO emissions was therefore tested.

In the three temperate data sets (Auradé, Grignon and Meyrargues), wind speed and air moisture were anti-correlated (respectively, -0.32, -0.27, -0.52 for the three cited data sets), showing that wind speed is closely linked to air moisture (an anticorrelation is still a correlation). Conversely, the single tropical data set of Hombori showed a correlation between wind speed and air moisture (Correlation coefficient was 0.46). This may be explained by the fact that the experiment was performed during the dry to wet seasonal transition period. Indeed, this period experiences wet monsoon air coming from south west, preceding rainfall at the beginning of the rainy season, and generating very high bursts of wind advecting moist air. Air moisture was shown to increase by a factor 3 before the rain begins and might have an influence on microbial activity at the very surface. In temperate climates, in the case of our databases, wind dries air masses. Whatever the case, a correlation exists anyway between wind speed and air moisture, showing that wind speed may lead to interesting information concerning atmospheric conditions around.

Table 3. Statistical performances of the network for each added variable. RMSE = Root Mean Squared Error, both for training and validation sets

Case	Description	RMSE (training)	RMSE (validation)	Generalization cost	Explained variance (%)	Slope
1	WFPS, Tsoil (surface)	19.15	16.65	6.49	45	0.45
2	Case 1 + Tsoil (depth)	11.83	20.57	5.23	55	0.50
3	Case 2 + Fertilisation rate	11.25	16.55	5.25	60	0.59
4	Case 3 + Sand percentage	11.51	14.46	5.64	62	0.64
5	Case 4 + pH	9.19	14.72	4.87	66	0.68
Not referenced	+ Wind speed Temp	7.05	12.96	4.47	73	0.77
6	Case 5 + Wind speed	8.60	11.55	4.71	71	0.67

Table 4. Weights for NO flux modelling with neural network parameterization, to be used in eq. (1)

w_0	2.933114461	w_{14}	−3.825273045
w_1	0.42610475	w_{15}	0.000276255
w_2	3.363015731	w_{16}	−5.43372955
w_3	−0.36196346	w_{17}	0.3666563
w_4	7.759669542	w_{18}	−1.496317821
w_5	−2.568210942	w_{19}	−0.423885134
w_6	−2.049426373	W_{20}	−13.65714547
w_7	0.747238768	w_{21}	0.384956678
w_8	−5.490382411	w_{22}	2.223018524
w_9	−0.017732744	w_{23}	−0.563718847
w_{10}	−0.080485622	w_{24}	5.74408315
w_{11}	−0.062224317	w_{25}	2.419512956
w_{12}	0.070384977	w_{26}	6.182411614
w_{13}	−2.188901436	w_{27}	2.72535041

To understand and emphasize the role of wind speed, the Hombori set was suppressed from the database (because of its opposite behaviour). Not only was the generalization cost improved (4.47), but the RMSE got closer for training and validation sets when only temperate databases were considered in this case (7.0 and 12.9, respectively). R^2 was 73%. These results (referred to as Wind speed Temp in Table 3) showed the importance of taking the wind speed data into account to improve the flux description. Hence, wind speed appeared as a proxy data describing the state of the atmosphere, and introduction in the network increased its performance.

However, results without the Hombori set are not directly comparable to the previous network results. Since the database is not exactly the same, cost calculations are not comparable. This exercise was only made to have an idea of what kind of results could be obtained without merging tropical and temperate database, and these results were quite satisfying. In the objective of keeping all available data in the same database, Hombori data were of course reintroduced: generalization cost still improved (4.71) compared to case 4, and RMSE were 8.6 and 11.5 for training and validation sets, respectively. R^2 was 71% between measured data and calculated ones (see Fig. 2 case 6). In Fig. 3 case 6, whereas low fluxes variations were not represented between days 121 and 125, this case showed again an improvement in calculated high-frequency variations for Auradé set, compared to what happened in case 1: daily fluctuations and gradual decrease in emissions were well reproduced, and low-fluxes mean value got closer to measured fluxes during the days 121 to 125, and 147 to 153. The high-modelled values at the beginning of the series are the result of a bad interpretation of fertilization by the network: indeed, fertilization is applied in Auradé at day 120, but the increase in NO emissions really occurs after the rain at day 126, whereas the network reproduces an immediate increase in NO emissions at day 120.

This error already occurred when fertilization rate was introduced in the network (case 3), and is not attributable to wind speed. However, the entire network performances being improved compared to case 4, it was then possible to conclude that wind speed may be an important parameter to be considered for NO emissions.

4. Comparison of the NN with other regression model results

The resulting eq. (2) presented in Section 2.3 allows a good representation of modelled results versus measurements, as already shown by Fig. 2 case 6 for Auradé only. It was however important to verify that other parameterizations could not give simpler and better results.

The reference parameterization remains the one of Yienger and Levy (1995, referenced herein after as YL95), and we have tried to apply it to our database, considering each subdatabase independently. According to YL95, the NO flux near the surface and before reduction by the canopy can be written as in eq. (3):

$$\text{Flux} = f_{w/d}(\text{soil temperature}, A_{w/d}(\text{biome})) \times P(\text{precipitation}), \quad (3)$$

where w stands for wet soils and d for dry soils.

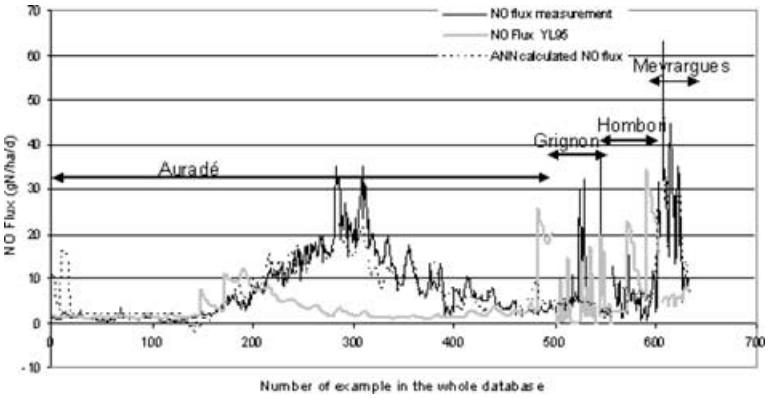


Fig. 4. Comparison between measured NO fluxes (black line), ANN calculated NO fluxes (dashed black line) and calculated NO fluxes with the YL95 algorithm (grey line), in gN. ha^{-1}. d^{-1}, function of the number of examples in the whole database. Each sub database is delimited by horizontal arrows.

In our database, Auradé, Grignon and Meyrargues are considered as wet agriculture soils, and Hombori as a dry grassland soil. The A_w(agriculture) in YL95 is an addition of A_w(grassland) plus 2.5% of the fertilization rate applied each month. This $A_{w/d}$(biome) is then multiplied by an exponential function of the soil temperature, and by a scalar factor depending on the rain rate and expressing the pulse effect.

The calculated YL95 results are shown in Fig.4, and compared to measurements and to the ANN calculation for the whole database. It is clear from this comparison that calculated fluxes are underestimated (4.4 ± 5.7 gN. ha^{-1}. d^{-1}, versus 7.9 ± 8.2 gN. ha^{-1}. d^{-1} in measurements, and 7.5 ± 6.6 gN. ha^{-1}. d^{-1} in the network). Furthermore, the increase in calculated NO flux does not correspond to the measured one, whatever the sub-database. These uncorrelated results ($R^2 < 0.001$) could be explained by the fact that YL95 algorithm gives global estimates of NO fluxes worldwide, at a monthly temporal resolution. Our database is constituted of hourly mean fluxes, and could not be properly reproduced by the YL95 algorithm. The same conclusion arises when the W92 (Williams et al., 1992) algorithm is applied to this database. (W92: exponential function of the soil temperature multiplied by a biome factor, results not shown here). Mean calculated fluxes are overestimating measurements in that case (10.5 ± 10.9 gN. ha^{-1}. d^{-1} for W92 estimation), and the correlation coefficient between calculated and measured fluxes remains very poor ($R^2 = 0.12$).

The NN technology seems therefore to be a good compromise between general parameterizations like the YL95 or W92 algorithms, not designed for high-frequency fluxes variability, and biogeochemical models, defined and adapted to specific regions and needing very detailed soil parameters.

5. Conclusion

The aim of this work was to develop a new parameterization to reduce uncertainties on NO flux description, and to be able to build a common parameterization whatever the type of soil and/or climate. This point is certainly the main difference with existing parameterizations for which, each type of biome has to be defined, as well as biogeochemical models, designed precisely only for some specific regions. A neural network approach was used to define an equation (eq. 2) based on a set of general descriptors that can be used in emission modelling. The flux was described through influent variables: WFPS at surface, soil surface temperature, fertilization rate, pH, sand percentage, that were previously described as influencing NO emissions, and soil temperature at depth and wind speed, that were used for the first time in parameterization. The eq. (2) was defined by adding these variables successively, in order to highlight the improvement of the network performances at each step, and the usefulness of these particular variables in describing the NO emission.

The wind speed and deep soil parameters were included here as a first test in improving NO flux description and their addition gave a significantly better performance of the network. The introduction of lagged surface temperature as an additional parameter could also have been interesting, in comparison to deep soil temperature, and will be tested in the future.

The main findings of this study are that: (1) an algorithm based on seven descriptors was found to improve NO emissions modelling; (2) the obtained algorithm describes NO emissions measured for two very different environments (temperate and tropical climates), with a high correlation level ($R^2 = 0.71$) between model and measurements.

However, this result will have to be confirmed with other data and in particular with tropical data collected in dry and wet seasons. Indeed, a lot of uncertainties subsist in quantifying these emissions from bare natural soils, specifically in tropical ecosystems. One of the main uncertainties lies in describing the diurnal cycle of the flux in tropical climate, and in linking reliably pulse effect and environmental parameters, whatever the type of soil and climate. The more complete the database will be concerning these particular features, the more precisely the ANN algorithm will describe them.

This work was an attempt to find universal parameterization and further work is now needed to expand the database and reduce uncertainties. The multiplicity of new databases would certainly reinforce the universal character of the proposed parameterization. Other types of data would also be useful as predictors

to explain NO fluxes variability in time and space, like N_2O and CO_2 fluxes from soil. Indeed, an increasing number of studies have reported a link between soil respiration and nitrogen emissions (Breuer et al., 2000; Butterbach-Bahl et al., 2004; Schindlbacher et al., 2004), giving a new insight in C and N cycles. Experimental field campaigns are currently planned in Africa in the frame work of African Monsoon Multidisciplinary Analysis (AMMA) to provide an extended data set from tropical ecosystems.

The proposed equation is a first step into the understanding of emission processes and their influence on atmospheric chemistry, without being a biogeochemical description of NO fluxes. Next step is to use this parameterization in a surface model (Surface Vegetation Atmosphere Transfer). This new step is achievable, as general descriptors have been used in the algorithm, allowing an easy on line connection to SVAT models. First tests of coupling SVAT to neural network have given encouraging results, and will be the subject of a future publication. Further work will consist in testing this parameterization in a chemistry transport model, in order to quantify the impact of NO emissions on tropospheric chemistry, particularly on ozone formation.

References

Anderson, I. C., Levine, J. S., Poth, M. A. and Riggan, P. J. 1988. Enhanced biogenic emissions of nitric oxide and nitrous oxide following surface biomass burning. *J. Geophys. Res.* **93**(D4), 3893–3898.

Austin, A. T., Yahdjian, L., Stark, J. M., Belnap, J., Porporato, A. and co-authors. 2004. Water pulses and biogeochemical cycles in arid and semi arid ecosystems. *Oecologia* **141**, 221–235.

Bertram, T. H., Heckel, A., Richter, A., Burrows, J. P. and Cohen, R. C. 2005. Satellite measurements of daily variations in soil NOx emissions. *Geophys. Res. Lett.* **32**, doi:10.1029/2005GLO24640.

Bouwman, A. F., Boumans, L. J. M. and Batjes, N. H. 2002a. Emissions of N_2O and NO from fertilized fields: Summary of available measurement data. *Global Biogeochem. Cycles* **16**, 1058, doi:10.1029/2001GB001811.

Bouwman, A. F., Boumans, L. J. M. and Batjes, N. H. 2002b. Modeling global annual N2O and NO emissions from fertilized fields. *Global Biogeochem. Cycles* **16**, 1080, doi:10.1029/2001GB001812.

Breuer, L., Papen, H. and Butterbach-Bahl, K. 2000. N_2O emission from tropical forest soils of Australia. *J. Geophys. Res.* **105**(D21), 26 353–26 367.

Butterbach-Bahl, K., Stange, F., Papen, H. and Li, C. 2001. Regional inventory of nitric oxide and nitrous axide emissions for forest soils of southeast Germany using the biogeochemical model PnET-N-DNDC. *J. Geophys. Res.* **106**, 34 155–34 166.

Butterbach-Bahl, K., Kock, M., Willibald, G., Hewett, B., Buhagiar, S., Papen and co-authors. 2004. Temporal variations of fluxes of NO, NO_2, N_2O, CO_2 and CH_4 in a tropical rain forest ecosystem. *Global Biogeochem. Cycles* **18**, GB3012, doi:10.1029/2004GB002243.

Cardenas, L., Rondon, A., Johansson, C. and Sanhueza, E. 1993. Effects of soil moisture, temperature and inorganic nitrogen on nitric oxide emissions from acidic tropical savannah soils. *J. Geophys. Res.* **98**, 14 783–14 790.

Conrad, R. 1996. Soil microorganisms as controllers of atmospheric trace gases (H2, CH4, OCS, N_2O and NO). *Microbiol..Rev.* **60**, 609–640.

Davidson, E. A. 1991. Fluxes of nitrous oxide and nitric oxide from terrestrial ecosystems. In: *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxides and Halomethanes* (eds. J. E. Rogers and W. B. Whitman) American Society for Microbiology, Washington, 219–235.

Davidson, E. A. and Kingerlee, W. 1997. A global inventory of nitric oxide emissions from soils. *Nutr. Cycl. Agroecosyst.* **48**, 37–50.

Delmas, R., Serça, D. and Jambert, C. 1997. Global inventory of NOx sources. *Nutr. Cycl. Agroecosyst.* **48**, 51–60.

Diallo, A. and Gjessing, J. 1999. Natural resources management: Morpho-pedology in Gourma region. SSE Research program Mali-Norway (in French), CNRST-IER-Oslo University (Norway).

Dutot, A., Rude, J. and Aumont, B. 2003. Neural network method to estimate the aqueous rate constants for the OH reactions with organic compounds. *Atmos. Environ.* **37**, 269–276.

Dreyfus, G., Martinez, J. M., Samuelides, M., Gordon, M. B., Badran, F., Thiria, S. and co-authors. 2002. In: *Neural Networks. Methodology and applications* (In French). Eyrolles, Paris.

Ganzeveld, L. N., Lelieveld, J., Dentener, F. J., Krol, M. C., Bouwman, A. J. and co-authors. 2002. Global soil-biogenic emissions and the role of canopy processes. *J. Geophys. Res.* **107**(D16), 10.1029/2001IJD001289.

Gardner, M. W. and Dorling, S. R. 1998. Artificial neural networks (The multilayer Perceptron)- A review of applications in the atmospheric sciences. *Atmos. Environ.* **32**, 2627–2636.

Gardner, M. W. and Dorling, S. R. 1999. Neural network modelling and prediction of hourly NOx and NO_2 concentrations in urban air in London. *Atmos. Environ.* **33**, 709–719.

Horii, C. V., Munger, J. W., Wofsy, S. C., Zahniser, M., Nelson, D. and co-authors. 2004. Fluxes of nitrogen oxides over a temperate deciduous forest. *J. Geophys. Res.* **109**, D08305, doi:10.1029/2003JD004326.

Hsieh, W. W. and Tang, B. 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bull. of the American Meteorological Society* **79**, 1855–1869.

Jaeglé, L., Martin, R. V., Chance, K., Steinberger, L., Kurosu, T. P. and co-authors. 2004. Satellite mapping of rain-induced nitric oxide emissions from soils. *J. Geophys. Res.* **109**, doi:10.1029/2003JD004406.

Johansson, C., Rodhe, H. and Sanhueza, E. 1988. Emission of NO in a tropical savanna and a cloud forest during the dry season. *J. Geophys. Res.* **93**, 7180–7192.

Kesik, M., Ambus, P., Baritz, R., Brüggemann, N., Butterbach-Bahl, K. and co-authors. 2005. Inventories of N_2O and NO emissions from European forest soils *Biogeosciences* **2**, 353–375.

Kiese, R., Li, C., Hilbert, D. W., Papen, H. and Butterbach-Bahl, K. 2005. Regional application of PnET-N-DNDC for estimating the N2O source strength of tropical rainforests in Wet Tropics of Australia. *Global Change Biol.* **11**(1), 128–144.

Kullback, S. and Leibler, R. 1951. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86.

Kullback, S. 1959. *Statistics and Information Theory*. Dover Publications, New York.

Laville, P., Hénault, C., Gabrielle, B. and Serça, D. 2005. Measurement and modelling of NO fluxes over maize and wheat crops during their growing seasons: effect of crop management. *Nutr. Cycl. Agroecosyst.* **72**, 159–171.

Le Roux, X., Abbadie, L., Lensi, R. and Serça, D. 1995. Emission of nitrogen monoxide from African tropical ecosystems: control of emission by soil characteristics in humid and dry savannas of West Africa. *J. Geophys. Res.* **9100**, 23133–23142.

Li, C., Aber, J., Stange, F., Butterbach-Bahl, K. and Papen, H. 2000. A process-oriented model of N_2O and NO emissions from forest soils: 1. Model development. *J. Geophys. Res.* **105**, 4369–4384.

Ludwig, J., Meixner, F. X., Vogel, B. and Förstner, J. 2001. Soil-air exchange fo nitric oxide: An overview of processes, environmental factors, and modelling studies. *Biogeochemistry* **52**, 225–257.

Martin, R. E., Scholes, M. C., Mosier, A. R., Ojima, D. S., Holland, E. A. and co-authors. 1998. Controls on annual emissions of nitric oxide from soils of the Colorado shortgrass steppe. *Global Biogeochem. Cycles* **12**, 81–91.

Meixner, F. X. and Yang, W. X. 2004. Biogenic emissions of nitric oxide and nitrous oxide from arid and semi-arid land. *Dryland Ecohydrology* (eds. P. D'Odorico and A. Porporato). Kluwer Academic Publishers B.V., Dordrecht, The Netherlands, pp. 23–46.

Müller, J. 1992. Geographical distribution and seasonal variation of surface emissions and deposition velocities of atmospheric trace gases. *J. Geophys. Res.* **97**, 3787–3804.

Navone, H. D. and Ceccatto, H. A. 1994. Predicting Indian monsoon rainfall: a neural network approach. *Clim. Dyn.* **10**, 305–312.

Ormeci, B., Sanin, S. L. and Pierce, J. J. 1999. Laboratory study of NO flux from agricultural soil: effects of soil moisture, pH and temperature. *J. Geophys. Res.* **104**, 1621–1629.

Otter, L. B., Yang, W. X., Scholes, M. C. and Meixner, F. X. 1999. Nitric oxide emissions from a Southern African Savanna. *J. Geophys. Res.* **104**, 18 471–18 485.

Parsons, D. A. B., Scholes, M. C., Scholes, R. J. and Levine, J. S. 1996. Biogenic NO emissions from savanna soils as a function of fire regime, soil type, soil nitrogen and water status. *J. Geophys. Res.* **101**, 23 683–23 688.

Parton, W. J., Holland, E. E., Del Grosso, J., Hartman, M. D., Martin, R. E. and co-authors. 2001. Generalized model for NOx and N_2O emissions from soils. *J. Geophys. Res.* **106**, 17 403–17 419.

Potter, C. S., Matson, P. A., Vitousek, P. M. and Davidson, E. A. 1996. Process modelling of controls on nitrogen trace gas emissions from soils worldwide. *J. Geophys. Res.* **101**, 1361–1377.

Reiners, W. A., Liu, S., Gerow, K. G., Keller, M. and Schimel, D. S. 2002. Historical and future land use effects on N_2O and NO emissions using an ensemble modelling approach: Costa Rica's Caribean lowlands as an example. *Global Biogeochem. Cycles* **16**, 1068, doi:10.1029/2001GB001437.

Roelle, P. A., Aneja, V. P., Gay, B., Geron, C. and Pierce, T. 2001. Biogenic nitric oxide emissions from cropland soils. *Atmos. Environ.* **35**, 115–124.

Roelle, P. A., Aneja, V. P., Mathur, R., Vukovich, J. and Peirce, J. 2002. Modeling nitric oxide emissions from biosolid amended soils. *Atmos. Environ.* **36**, 5687–5696.

Sanhueza, E., Hao, W. M., Scharffe, D., Donoso, L. and Crutzen, P. J. 1990. N_2O and NO emissions from soils of the Northern part of the Guyana Shield, Venezuela. *J. Geophys. Res.* **95**, 22 481–22 488.

Schindlbacher, A., Zeichmeister-Boltenstern, S. and Butterbach-Bahl, K. 2004. Effects of soil moisture and temperature on NO, NO_2 and N_2O emissions from European forest soils. *J. Geophys. Res.* **109**, D17302, doi:10.1029/2004JD004590.

Schlecht, E., Fernandez-Rivera, S. and Hiernaux, P. 1997. Timing, size and nitrogen concentration of faecal and urinary excretions in cattle, sheep and goats: Can they be exploited for better manuring of cropland? In: *Soil Fertility Management in West African Land Use Ssytems* (eds. G. Renard, A. Neef, K. Becker and M. von Oppen), Margarf Verlag, Niamey, Niger, Weikersheim, Germany.

Schlecht, E. and Hiernaux, P. 2004. Beyond adding up inputs and outputs: process assessment and upscaling in modelling nutrient flows. *Nutr. Cycl. Agroecosyst.* **70**, 303–319.

Serça, D., Delmas, R., Jambert, C. and Labroue, L. 1994. Emissions of nitrogen oxides from equatorial rain forest in central Africa: origin and regulation of NO emissions from soils. *Tellus* **46B**, 243–254.

Serça, D., Delmas, R., Le Roux, X., Parsons, D. A. B., Scholes, M. C., Abbadie, L. and co-authors. 1998. Comparison of nitrogen monoxide emissions from several African tropical ecosystems and influence of season and fire. *Global Biogeochem. Cycles* **12**, 637–651.

Shepherd, M. F., Barzetti, S. and Hastie, D. R. 1991. The production of atmospheric NOx and N_2O from a fertilized agricultural soil. *Atmos. Environ.* **25A**, 1961–1969.

Skiba, U., Hargreaves, K. J., Fowler, D. and Smith, K. A. 1992. Fluxes of nitric and nitrous oxides from agricultural soils in a cool temperate climate. *Atmos. Environ.* **26A**, 2477–2488.

Van Dijk, S. and Meixner, F. 2001. Production and consumption of NO in forest and pasture soils from the Amazon basin. *Water, Air, and Soil Pollution* **1**, 119–130.

Vapnik, V. N. 1995. *The nature of statistical learning theory*. Springer.

Williams, E. J., Guenther, A. and Fehsenfeld, F. C. 1992. An inventory of nitric oxide emissions from soils in the United States. *J. Geophys. Res.* **97**, 7511–7519.

Yang, W. X. and Meixner, F. X. 1997. Laboratory studies on the release of nitric oxide from sub-tropical grassland soils: The effect of soil temperature and moisture. In: *Gaseous Nitrogen Emissions from Grasslands* (eds. S. C. Jarvis and B. F. Pain), CAB International, Wallingford, New York, pp. 67–70.

Yan, X., Ohara, T. and Akimoto, H. 2005. Statistical modelling of global soil NOx emissions. *Global Biogeochem. Cycles* **19**, GB3019, doi:10.1029/2004GB002276.

Yi, J. and Prybutok, V. R. 1996. A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Poll.* **92**, 349–357.

Yienger, J. J. and Levy, H. II, 1995. Empirical model of global soil-biogenic NOx emissions. *J. Geophys. Res.* **100**, 11447–11464.