Artificial neural network modelling of microbial acclimation periods in soil contaminated with petroleum hydrocarbons

A.M. SUCHORSKI-TREMBLAY[1] and R. KOK[2]

[1]Research & Technology Transfer Section, Collège d'Alfred, Alfred, ON, Canada K0B 1A0; and [2]Department of Agricultural and Biosystems Engineering, McGill University, 21111 Lakeshore Road, Ste. Anne de Bellevue, QC, Canada H9X 3V9.

Received 8 October 1996; accepted 14 March 1997.

Suchorski-Tremblay, A.M. and Kok, R. 1997. **Artificial neural network modelling of microbial acclimation periods in soil contaminated with petroleum hydrocarbons**. Can. Agric. Eng. **39**:123-130. Artificial neural networks (ANNs) were used to model the acclimation of indigenous microorganisms in soil contaminated with diesel fuel or creosote. Acclimation data were obtained by measuring the appearance of ^{14}C (as CO_2) mineralized from radiolabelled tracers that were added to soil microcosms. The ANNs were trained and tested with the following inputs: incubation temperature, water content (as percentage of the soil's water holding capacity), addition or not of nitrogen and phosphorus fertilizers, and sampling time. The ANN output was directly related to the cumulative percent of ^{14}C recovered ($\%\Sigma^{14}$C). The resultant ANN models were incorporated into a user-friendly software package called AccliMat, written in QuickBASIC. With this package, the ANN models can be utilized to calculate the amount of $\%\Sigma^{14}$C that will be recovered after a given number of days, or the number of days required to reach a given $\%\Sigma^{14}$C.

Des réseaux de neurones artificiels (RNA) ont servi de modèles prévisionnels pour l'acclimatation de microorganismes indigènes dans un sol contaminé par le diesel ou le créosote. Cette acclimatation a été mesurée en fonction de la présence de l'isotope ^{14}C (sous forme de CO_2), provenant de la dégradation des traceurs radioactifs ajoutés à des échantillons de sol. Les RNA ont été soumis à différentes conditions d'apprentissage suivant l'entrée de cinq données: la température d'incubation, la teneur en eau du sol (par rapport à sa capacité de rétention), l'ajout ou non d'engrais d'azote et de phosphore, et le temps d'échantillonnage. Les valeurs produites par les RNA étaient directement liées aux quantités de ^{14}C récupérées ($\%\Sigma^{14}$C) dans les microcosmes pédologique. Nos modèles RNA ont été incorporés à AccliMat, un progiciel relativement simple à utiliser, écrit en QuickBASIC. Dans le cadre de ce progiciel, les RNA peuvent servir à prévoir le pourcentage de ^{14}C récupéré dans un sol après un certain nombre de jours, ou encore, le temps requis pour atteindre un niveau donné de ^{14}C ($\%\Sigma^{14}$C).

INTRODUCTION

Successful biodegradation of hydrocarbon spills on soil is influenced by many factors: the level of contamination, the diversity and population of the indigenous microflora, weather conditions, soil nutrients, and time. When a spill occurs, the microorganisms acclimate to the carbon source and begin to degrade it. During the acclimation, various biological processes take place in the soil organisms themselves, as well as in the composition of the microbial community. Riser-Roberts (1992) has classified these as:

1) induction or derepression of enzymes specific to the degradation pathways of a particular compound,
2) increase in the number of organisms in the degrading population,
3) random mutation in which new metabolic pathways are produced that will allow degradation, and
4) adaptation of existing catabolic enzymes ... to the degradation of novel compounds.

Such processes are affected by environmental factors. For instance, enzyme induction and derepression may be temperature and nutrient dependent.

In this study, two situations were examined. In both, a radiolabelled tracer technique was used to follow the production of ^{14}C (as CO_2) being mineralized from the tracers. In the first situation diesel fuel or creosote was added to a soil which had not been previously contaminated. In this case both the contaminant and the radiolabelled tracer were added to the test soil and the microbial population had to adapt to utilize the hydrocarbon source provided. We refer to this as *microbial acclimation*. In the second situation hydrocarbon spillage onto the soil had occurred over many years and the microbes were already fully acclimated. We were interested in how long it takes after spring thaw before the microbes begin to degrade the hydrocarbons again and this situation is therefore referred to as *post-thaw re-acclimation*. For this case only the radiolabelled tracer was added to the test soil.

Data on microbial acclimation and post-thaw re-acclimation were obtained by monitoring ^{14}C production (as CO_2), under various treatments. The data were then modelled with Artificial Neural Networks (ANNs) and a program with a user-friendly interface (AccliMat) was written to facilitate access to these ANN models. With the program a user can input the acclimation time so as to find the cumulative percent of ^{14}C recovered ($\%\Sigma^{14}$C), or input the $\%\Sigma^{14}$C recovered to find the corresponding time required.

OBJECTIVES

The objectives of the project were: 1) to generate a set of models of the appearance over time of ^{14}C which resulted from the microbial mineralization of ^{14}C-octadecane when in the presence of diesel fuel in soil, and from the mineralization of ^{14}C-phenanthrene or ^{14}C-pyrene when in the presence of creosote in soil and, 2) to make the models available to non-experts via a user-friendly software interface.

BACKGROUND

Artificial neural networks

Artificial neural networks are structured similarly to biological brains so that they can mimic and approximate to some extent the "human" way of interpreting and solving problems. An ANN's primary components are *processing elements* (PEs) which are joined together by *connections* in a manner analogous to biological neurons being joined with synapses. The PEs are structures that receive a number of input signals which are attenuated by *weights* associated with the connections they pass through. Within a PE, these attenuated inputs are added and the combined signal is fed into a *transfer function*. The output from the transfer function is then fed through the connections to other PEs. Usually, for convenience of description, the PEs and connections are arranged in layers. The base of the structure then consists of the *input layer* whose PEs receive the external inputs and, similarly, the top of the structure consists of the *output layer* whose PEs emit the external outputs. In between are located a number of *hidden layers*, each containing a number of PEs, as determined by the ANN designer. The PEs in the various layers are interconnected in a pattern that is controllable by the ANN designer, although standard patterns are mostly used (Kok et al. 1994; NeuralWare 1995a; Shukla et al. 1996). For this study the *fully-connected, feedforward* pattern was used. This means that the layers are connected so that the output from any one PE is fed as input to all PEs in the layer above it. The signal flow therefore proceeds linearly from the input layer, through the hidden layer(s), and out via the output layer.

In the operation of an ANN there are two phases: *learning* (or training) and *recall* (or testing). Much of the information or knowledge stored in an ANN resides in the values of its weights, the rest being implicitly represented by its structure. A network learns by adjusting these weights in response to input/output data that are presented to it and various schemes may be used to accomplish this. For example, ANNs can be trained using *supervised backpropagation* learning (NeuralWare 1995a). In this, differences between the model's actual output and the ideal output it should have are used to correct the weights according to a backpropagation scheme, using a *learning rule*, such as the Delta rule. After learning has proceeded for a number of cycles and the network embodies the input-output relationship to a certain extent, the weights may be frozen and the network used for *recall*. In this phase the ANN is presented with sets of inputs from which it generates output values. If ideal "source output" values, corresponding to the recall inputs, are also available (e.g. from the learning data set) the ANN's performance can be judged in terms of the degree to which the two sets of output values are in agreement.

ANNs can be used to model complex, nonlinear, and large data sets. They can learn by example, without having a priori knowledge of the subject being treated, and are robust enough to withstand noisy or incomplete inputs. ANNs have been applied in many biosystems areas, such as the control of fermentors and growth pattern recognition (NeuralWare 1995b; Linko et al. 1995; Widrow et al. 1988). In this study ANN modelling was used because of the nature of the data, which strongly resemble bacterial growth curves.

Biological processes

Biological processes often provide a data-rich but knowledge-poor situation and the quality of automatic control that is possible may be rather limited. In fermentation processes there are three main factors that contribute to this: a lack of reliable mathematical models to describe cell growth and metabolite production, a lack of on-line sensors that can detect important process state variables, and the microorganisms themselves having a complex regulatory system within the cells, which the external control system can only affect by manipulating the extracellular environment. Accordingly, limitations within the process and the availability of large quantities of data favor the utilization of ANNs in the control system. They are, therefore, being increasingly exploited for such roles. For instance, Zhang et al. (1994) employed the ANN approach to model the production of *Bacillus thurigiensis* and then used that model to control the fermentation. Similarly, Linko et al. (1995) used an ANN model to estimate the consumption of sugar and the production of lysine by *Brevibacterium flavum* in a fed-batch fermentation with the intent to control amino acid production.

Kinetic model building

ANNs have also been used for kinetic model building. For example, Tyagi and Du (1992) developed an ANN model of the toxic effects of lead, chromium, zinc, nickel, and their mixture on the growth of both acclimated and unacclimated microorganisms. Their intent was to capture the kinetics of heavy metal inhibition in an activated sludge process and they chose the ANN approach because of the complex patterns involved, in terms of nonlinearity, multivariability, and interaction between process variables. They also studied the effects of nickel and chromium on acclimated and unacclimated sludge with neural models. As the kinetic patterns grew in complexity, so did the neural models' architectures. For comparison, the data were also fed to a linear multiple regression method and the sum of squares of the errors of the two modelling approaches compared. The ANN's error was one third that of the regression model.

MATERIALS AND METHODS

Soils and contaminants

This study involved the use of twelve soils in total. The company cooperating on the project, Grace Bioremediation Technologies of Betz Dearborn, Inc. (Mississauga, ON) provided us with 11 soils from different locations; the twelfth came from nearby Alfred College (Alfred, ON). The soils were divided into three groups of four, respectively being "clean", "historically contaminated" with diesel fuel, and "historically contaminated" with creosote (see Table I). The soils were named as follows. Clean: *Guelph*, *Mississauga*, *Halton Clay*, and *Halton Sand*; diesel contaminated: *Quebec Bunker C*, *Mississauga Refinery*, *Oakville Refinery*, and *Cambridge Farm* (the local soil); creosote contaminated:

Table I: Elemental analysis of clean and historically contaminated (HC) soils calculated on a dry weight basis

Soil Name	Total C (% w/w)	Total N (% w/w)	Phosphorus (mg P/kg)	Potassium (mg K/kg)	Magnesium (mg Mg/kg)
Clean Soils					
Guelph	2.71	0.23	28	99.2	459
Mississauga	2.92	0.16	7	150.5	384
Halton Clay	1.99	0.02	9	61.3	232
Halton Sand	1.57	<0.01	6	64.8	36
HC Soils: Diesel					
Quebec Bunker C	7.07	0.10	5	33.6	2310
Mississauga Refinery	6.70	0.06	3	20.6	17
Oakville Refinery	7.49	<0.01	1	3.3	44
Cambridge Farm	6.32	0.35	88	40.8	49
HC Soils: Creosote					
Alberta Clay	7.31	0.12	6	129.8	335
New Brunswick	6.04	0.13	7	89.9	64
Quebec Dumpsite	2.63	<0.01	2	14.6	8
Trenton Industrial	8.44	<0.01	4	25.8	47

Alberta Clay, New Brunswick, Quebec Dumpsite, and Trenton Industrial. All were chosen based on site accessibility and permission from the site owners to collect samples. From each site a batch of about 40 kg was collected. All twelve batches of soil were put through a #4 mesh sieve (4.75 mm grid, USA Standard Testing Sieve) to remove large particles, stones and roots (Pramer and Bartha 1972). The batches were then sub-sampled and this material (about 25 kg per soil) was mixed one hundred times with a hand trowel. The prepared clean soils were then stored in plastic bags, sealed and placed in a refrigerator at 5°C, while the prepared historically contaminated soils were stored frozen at -20°C to simulate winter conditions. Also, for all prepared soils the following were determined: the saturated water holding capacity (WHC) (Atlas and Bartha 1981); the water, N (Bremner and Mulvaney 1982), P (Olsen and Sommers 1982), total C (Nelson and Sommers 1982), K (Knudsen et al. 1982) and Mg (Lanyon and Heald 1982) contents; total petroleum hydrocarbon (TPH) and polycyclic aromatic hydrocarbon (PAH) analysis was performed by Water Technology International Corp. (Burlington, ON). The latter two tests were performed to ensure that the clean soils were clean and to determine to what degree the historically contaminated soils were actually polluted.

To prepare the clean soils for the acclimation experiments, separate samples of each, equivalent to 25 g dry mass, were contaminated with 250 mg diesel fuel and 10^6 dpm (disintegrations per minute) ^{14}C-octadecane, or with 150 mg creosote and 10^6 dpm ^{14}C-phenanthrene (radiolabels purchased from Sigma Chemical Co., St. Louis, MO). To prepare the historically contaminated soils only radiolabelled tracers were added. To the diesel-contaminated ones, 10^6 dpm ^{14}C-octadecane was added per sample (equivalent to 25 g dry mass), while to the creosote-contaminated ones, 10^6 dpm ^{14}C-pyrene was added. This resulted in a total of sixteen types of contaminated soil samples. Next, for both the acclimation and post-thaw re-acclimation experiments various amounts of water, N, and P were added to these, as described in the "Microcosms" section below.

Microcosms

Soil acclimation was studied with microcosms. Each microcosm consisted of a 250 mL biometer flask having a centrally fused vertical glass cylinder (15 x 70 mm) into which a gas trap vial was placed (see Fig. 1). The trap contained 5 mL of 2 N NaOH to capture $^{14}CO_2$ and was replaced when necessary (see below). Each flask contained a soil sample prepared as described above (equivalent to 25 g dry mass of soil) to which water and fertilizer amendments were added as required. Enough water was added to bring the soil water content up to either 50% or 85% of WHC; if N was added, the total sample N content was increased to 1000 µg (with ammonium nitrate); if P was added, the total sample P con-

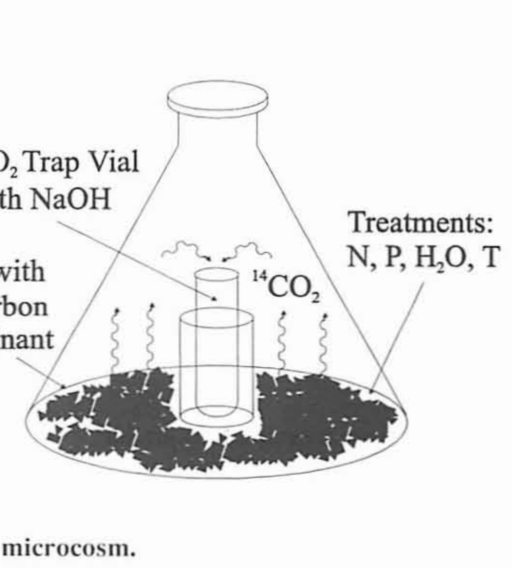

Fig. 1. A soil microcosm.

tent was increased to 1000 µg (with triple superphosphate fertilizer). The flasks were then sealed with neoprene stoppers and kept in the dark at either 10°C or 22°C to prevent photodegradation of the chemicals. Each microcosm was duplicated, except for the negative controls (air dried soils with radiolabel and/or hydrocarbon) making 528 microcosms in total (((4 clean soils x 2 contaminants + 8 historically contaminated soils) x 2 temperatures x 2 WHC x 2 N amounts x 2 P amounts x 2 replicates) + 16 negative controls not used in modelling). As the radiolabelled tracer was mineralized by the microbes, $^{14}CO_2$ was released and then absorbed by the gas trap. The trap vials were changed every day or every two days for 22°C diesel microcosms and every five days for 22°C creosote and for both diesel and creosote 10°C microcosms, for up to 150 days. The evaluation of petroleum hydrocarbon metabolism by means of radiolabelled tracer degradation was described and validated by Sharabi and Bartha (1993).

The radioactivity of the trap contents was assessed with liquid scintillation counting (model LS6500; Beckman Instruments Inc., Fullerton, CA). The count sample was prepared by mixing 1 mL of the trap solution, 1 mL of 2 N Acetic Acid, and 10 mL of "liquid scintillation cocktail" (Optiphase HiSafe II or Scintiverse, Fisher Scientific Inc., Toronto, ON). The addition of the acid was required to neutralize the NaOH of the trap solution, so as to prevent the latter from quenching the response of the fluors. The counting time was 10 min per diesel sample and 3 min per creosote sample. The data obtained consisted of ^{14}C counts which corresponded to the amounts of tracer (and, hence, the quantity of contaminant) that had been metabolized during the period of residence of that particular vial of trap solution in the microcosm. In this study acclimation and post-thaw re-acclimation were defined as being complete when at least 10% of the tracer's radiolabel had been mineralized and trapped (Briglia et al. 1994).

Data preparation

From the ^{14}C count data cumulative values were calculated, as a percentage of the total dose of radiolabelled tracer that had been added to the microcosm ($\%\Sigma^{14}C$). These values were then further transformed to a *log-inverse-fraction* (Eq. 1). This was done because better modelling accuracy was required at lower fraction values, i.e., we wanted the relative error of the model, rather than the absolute error, to stay constant over the entire range of the fraction ^{14}C recovered.

$$log\text{ - }inverse\text{ - }fraction = \ln\left(\frac{100}{\%\sum^{14} C}\right) \quad (1)$$

For creosote soils many of the $\%\Sigma^{14}C$ values were zero (i.e., no ^{14}C had been recovered at all), causing difficulties with the transformation in Eq. 1. To facilitate the procedure, these values were adjusted to 0.001.

Artificial neural network modelling

A number of different ANN architectures were tried to determine which would best describe the data. As discussed in the Background section, the ANNs used were of the fully-connected, feedforward type, but the way the inputs were fed to the network, as well as the number of hidden layers (1, 2, or 3), and the numbers of PEs in these were varied. Initially, the intent had been to create one or two ANNs, with each modelling the combined data from a number of soils. In that approach the soil number was an input variable and we tried encoding it as a decimal quantity, as well as with several binary schemes. The quality of the resultant models was, however, insufficient to satisfy our needs. In the end, it was decided to model the soils entirely independently from one another, i.e., to create a separate ANN for each soil type. This approach also has the advantage that future models of new soils can easily be added to the package. For all networks, the remaining inputs were two-value coded, except for time. Thus, for each of the sixteen ANNs the inputs were: incubation temperature (0 = 10°C, 1=22°C), water content (0 = 50% of WHC, 1 = 85%), N amendment (1 = not added, 2 = added), P amendment (1 = not added, 2 = added), and time in hours. The output was the log-inverse-fraction of the ^{14}C recovered, as defined by Eq. 1.

To train and test an ANN having a specific architecture, the procedure was to have the network learn from the source data for 250,000 cycles and then to freeze its weights and have it produce output values in recall mode using the same input values as had been used for the learning. The ANN recall output values were then compared with the corresponding source output values. Subsequently, the ANN was further trained and tested at regular intervals, up to one million cycles.

To compare two sets of outputs, average source values were first calculated for the duplicates (i.e., each microcosm had been duplicated and there were two source output values corresponding to any given input set). The performance of an ANN was then judged according to several measures that were calculated by comparing the recall and average source output values. The most important measure by which we judged was the mean absolute difference between the recall value and the average source value. The arithmetic mean of the differences was calculated as well, to check for offset (under ideal circumstances this would be zero). The standard deviation was used to evaluate the spread of the differences and the root mean square value was calculated to emphasize large differences. The correlation coefficient between the recall and average source values was also computed.

After preliminary work, the most appropriate ANN was found to be the fully connected, feed-forward type with the hyperbolic tangent transfer function being used in all the PEs of its hidden layers. The ANNs were all trained with supervised backpropagation learning using the Delta rule (NeuralWare 1995a). For diesel contaminated soils the best architecture was found to be "5-11-11-1", denoting that there were five PEs in the input layer, eleven PEs in each of two hidden layers, and one PE in the output layer (Fig. 2). For creosote contaminated soils, most final ANNs had the configuration 5-12-12-12-1 (i.e., three hidden layers each containing twelve PEs) but for the Trenton Industrial soil 5-13-13-13-1 was used and for the Guelph soil 5-14-14-14-1.

Software used

All acclimation and post-thaw re-acclimation data were combined in a Lotus 1-2-3 spreadsheet (Lotus Development

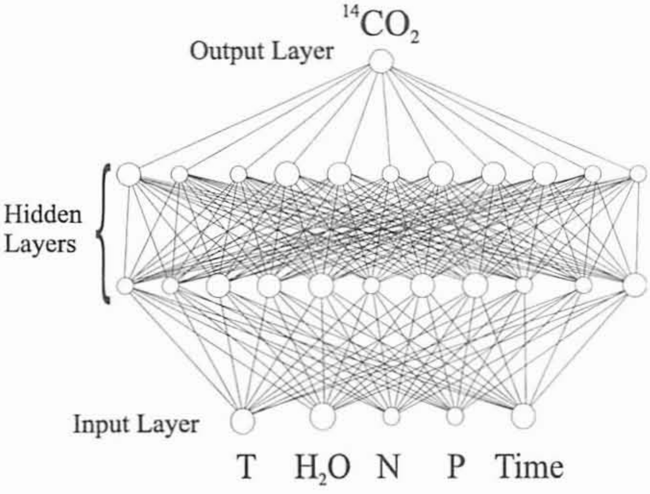

Fig. 2. Information flow in a 5-11-11-1 ANN.

Corp., Cambridge, MA) and then exported into individual ASCII files that became the learning and recall files for the ANN models. NeuralWare's Professional II/Plus (NeuralWare, Inc., Pittsburgh, PA) was used to create and train the ANNs. The measures to compare the recall output values to the source values were calculated with a program written in QuickBASIC 4.5.

Software written

Once the ANNs were sufficiently trained, they were extracted from the NeuralWare package as C language subroutines. These were then modified to BASIC format and incorporated into the support software, AccliMat, which was written in QuickBASIC. Its flowchart is shown in Fig. 3. After the user chooses a soil and treatments, Acclimat displays a graph of the modelled curve, together with limits for the incubation time and $\%\Sigma^{14}C$. At this point, the user is allowed to proceed in one of two ways. First, it is possible to supply inputs corresponding to a soil and experimental treatment, together with a time, and request a $\%\Sigma^{14}C$ value. Secondly, the user can specify treatment conditions, soil, and a $\%\Sigma^{14}C$, and request a time value. To accommodate the second request, an iterative approach is used in the program. In the iteration method the procedure starts with the maximum time obtained for that soil-treatment combination and decreases this by five days every cycle, checking if the target $\%\Sigma^{14}C$ value has been passed. If it has been, the search direction is reversed (so that time will increase for successive iterations), the iteration interval set to one day and the time increased one day at a time until the target is passed again. The number of days is then reported to the user.

To validate the AccliMat software, two tests were run on it, corresponding to the two types of request it can accommodate. First, the program was used to generate recall output values of log-inverse-fraction of $\%\Sigma^{14}C$ recovered for all available soil, treatment, time combinations. These were then compared to the corresponding values generated with the NeuralWare package while the same trained ANN was resident in its memory. Secondly, the iteration method was tested for all soil-treatment combinations as follows: a) AccliMat was used to find three time values, by iteration, respectively, corresponding to one-quarter, one-half, and three-quarters of the difference between maximum and minimum $\%\Sigma^{14}C$ recovered; b) the $\%\Sigma^{14}C$ recovered corresponding to the three time values was calculated with the ANN; c) these $\%\Sigma^{14}C$ values were then compared to the ones used to obtain the time values in step a) above. For example, for Mississauga soil, diesel acclimation, 22°C, 85% WHC, and no N nor P amendment, the difference between maximum and minimum

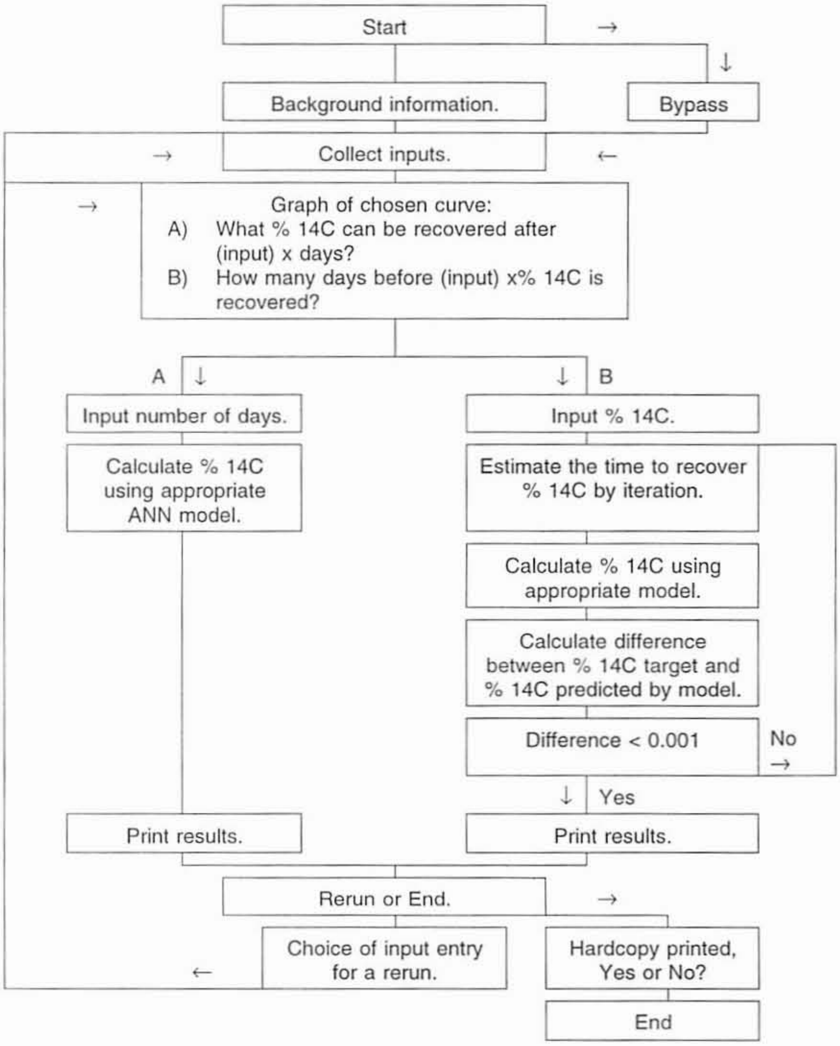

Fig. 3. Flowchart of AccliMat.

$\%\Sigma^{14}C$ recovered is 13.6% so that the one-quarter, one-half, and three-quarters values are respectively 3.4, 6.7, and 10.1%. The AccliMat iteration procedure yielded 19, 34, and 45 days as the time values corresponding to these; with the ANN model corresponding $\%\Sigma^{14}C$ values of 3.4, 6.8, and 10.2% were obtained.

RESULTS AND DISCUSSION

Soils and contaminants

The results of the soil elemental analysis are summarized in Table I. The agricultural soils, *Guelph* and *Cambridge Farm*, had the highest initial amounts of P and N; two of the historically contaminated soils, *Oakville Refinery* (diesel) and *Quebec Dumpsite* (creosote), had the lowest available P, K, and Mg concentrations; the *Quebec Bunker C* soil had a noticeably high Mg concentration. Seven of the eight historically contaminated soils had a total C content between 6% and 8.5%, whereas the *Quebec Dumpsite* soil had a low total C content.

All eight historically contaminated soils had a high sand content (*Quebec Bunker C* - coarse sand, *Mississauga Refinery* - loamy fine sand, *Oakville Refinery* -sand, *Cambridge Farm* - very fine sand, *Trenton Industrial* -loamy sand, *New Brunswick* - sandy loam, *Alberta Clay* -sandy loam, *Quebec* - sand), while three of the four clean soils were of a loamy type (*Guelph* - loam, *Mississauga* - clay loam, *Halton Clay* - loam) and one was sandy (*Halton Sand* - fine sand).

The results presented in Table II illustrate that, of the 8 historically contaminated soils, *Oakville Refinery* and *Mississauga Refinery* were the most contaminated with petroleum hydrocarbons whereas *Quebec Dumpsite* was the most contaminated with polycyclic aromatic hydrocarbons. For all clean soils TPH values less than 100 µg/g soil were obtained; only trace amounts of PAHs were found.

Table II: Total petroleum hydrocarbons (TPHs) and polycyclic aromatic hydrocarbons (PAHs) in historically contaminated soils

Soil Name	TPH (µg/g soil)	PAH (µg/g soil)
Diesel		
Quebec Bunker C	1220	trace
Mississauga Refinery	48,800	trace
Oakville Refinery	182,000	trace
Cambridge Farm	1110	trace
Creosote		
Alberta Clay	852	11.0
New Brunswick	174	21.2
Quebec Dumpsite	2305	116.0
Trenton Industrial	1224	20.9

Artificial neural networks

Results of the ANN performance measure calculations are presented in Table III. These are statistics of the log-inverse-fraction values (as defined in Eq. 1) and therefore dimensionless. The average mean absolute difference for all soils was 0.1007. In terms of error in the $\%\Sigma^{14}C$ value, this corresponds to overestimation by a factor of 1.106 ($e^{0.1007}$) or underestimation by a factor of 0.904 ($e^{-0.1007}$). It must be remembered, however, that the ANNs were trained with the log-inverse-fraction values, rather than the $\%\Sigma^{14}C$ values, so that the relative error would remain fairly constant over the entire range. This is illustrated in Fig. 4 in which three sets of *Halton Sand* source and ANN recall output values, corresponding to different treatments, are shown together. The error in the log-inverse-fraction remains fairly constant so that the larger errors in the $\%\Sigma^{14}C$ would be obtained at the higher values.

All the correlation coefficients were above 0.98 and the arithmetic mean differences were minimal, indicating that there was no systematic offset between the source and recall output values of the ANNs. The root mean square differences were similar in magnitude to the standard deviations of the differences, indicating that the differences themselves were not very widely distributed. Also, the root mean square differences were similar in magnitude to the mean absolute differences, indicating that the distributions of the differences were not skewed in any extreme manner. The best model constructed was of *Quebec Bunker C*, and the worst was of *Guelph* (respectively suffering from average overestimation factors of 1.052 and 1.182 in $\%\Sigma^{14}C$).

Overall, the data from the diesel contaminated soils were easier to model than those of the creosote contaminated ones. In the ANNs the diesel soil data required only two hidden layers of 11 PEs each, while the creosote soil data required three hidden layers of either 12, 13, or 14 PEs. Even so, the average mean absolute difference obtained for the diesel data was 0.0942 whereas for the creosote data it was 0.1107. The larger error in the creosote models reflects the character of the data collected. Of the 256 creosote microcosms, 123 never reached 1% $\Sigma^{14}C$ recovered within the 150 day maximum allowed (i.e., they remained far below our acclimation limit of 10%). As well, for several creosote contaminated soils substantial ^{14}C recovery occurred only under some treatments, but not under others. Such differences occurred even between replicate microcosms.

Software

In the first test of AccliMat it was used to find $\%\Sigma^{14}C$ for all experimental combinations of soil, treatment, and time. The log-inverse-fraction values that resulted always differed less than 1.0×10^{-5} from corresponding ones generated directly with the NeuralWare package in recall mode. In the test of the iteration method the $\%\Sigma^{14}C$ recovered, corresponding to the time values at one-quarter, one-half, and three-quarters of the difference between maximum and minimum values were respectively (average standard deviation): 0.3229 ± 0.4112, 0.2587 ± 0.3277, and 0.1485 ± 0.1624.

CONCLUSIONS

The use of ANNs to model radiolabel tracer data was reasonably successful; the models approximated the experimental data to a degree well within our expectations and needs. Subsequent to training, the extraction of the ANN model from the NeuralWare package as a C subroutine was straight-

Table III: Values of the ANN performance measures*

Soil Name	Contaminant	No. of data lines	Mean absolute difference	Root mean square difference	Arithmetic mean of differences	Standard deviation of differences	Correlation coefficient, r
Clean Soils							
Guelph	Diesel	528	0.1461	0.1927	-0.0006	0.1929	0.9895
Guelph	Creosote	456	0.1673	0.2365	-0.0013	0.2368	0.9982
Mississauga	Diesel	496	0.1048	0.1470	0.0014	0.1471	0.9938
Mississauga	Creosote	448	0.1247	0.2007	0.0015	0.2009	0.9986
Halton Clay	Diesel	796	0.1060	0.1722	0.0049	0.1722	0.9949
Halton Clay	Creosote	440	0.0763	0.1015	-0.0002	0.1016	0.9992
Halton Sand	Diesel	760	0.0828	0.1147	0.0001	0.1148	0.9971
Halton Sand	Creosote	440	0.0987	0.1545	-0.0013	0.1547	0.9977
HC Soils							
Quebec Bunker C	Diesel	336	0.0508	0.0695	0.0007	0.0696	0.9959
Mississauga Refinery	Diesel	246	0.0826	0.1051	0.0002	0.1053	0.9937
Oakville Refinery	Diesel	488	0.0791	0.1203	0.0000	0.1204	0.9971
Cambridge Farm	Diesel	224	0.0558	0.0715	0.0002	0.0717	0.9926
Quebec Dumpsite	Creosote	152	0.1017	0.1473	-0.0006	0.1478	0.9827
Alberta Clay	Creosote	208	0.0977	0.1435	-0.0002	0.1438	0.9983
New Brunswick	Creosote	200	0.0588	0.0745	0.0000	0.0747	0.9992
Trenton Industrial	Creosote	152	0.1296	0.1761	-0.0009	0.1767	0.9971
Totals for:	Diesel	3874	0.0942	0.1350			
	Creosote	2496	0.1107	0.1620			
	All soils	6370	0.1007	0.1455			

* Training results. The ANN was fitted separately for each soil-contaminant combination

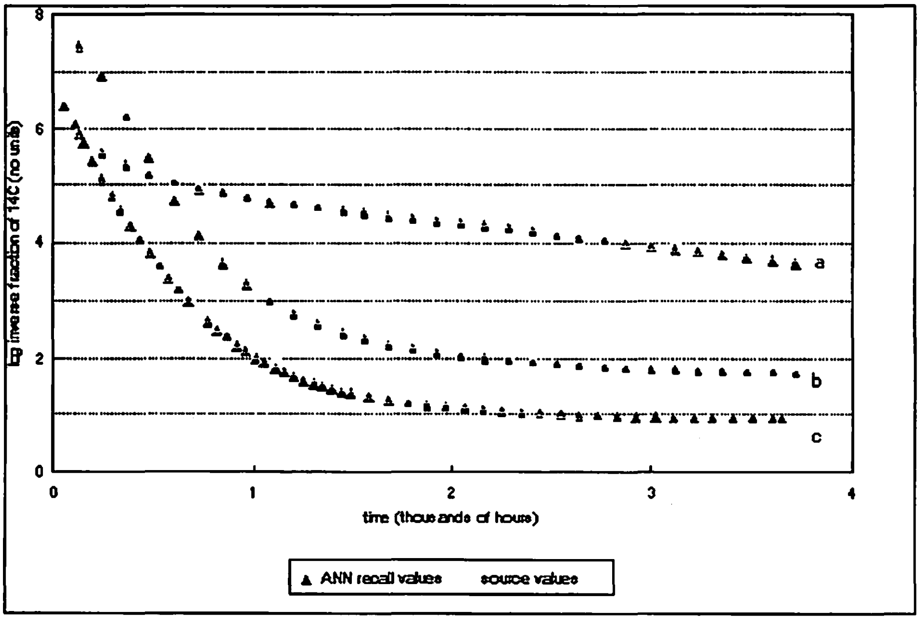

Fig. 4. Source and ANN recall output values versus time for Halton sand for three treatments: a) 10°C, 50% WHC, -N, -P; b) 10°C, 50% WHC, +N, +P; c) 22°C, 50% WHC, -N, +P.

forward, as was its translation into BASIC. The AccliMat interface software facilitates access to the models, so that any user can easily obtain answers to both acclimation questions, i.e., either time or $\%\Sigma^{14}C$ can be the unknown. The approach of modelling with ANNs, extracting them, and hiding them behind a user-friendly interface is highly recommended to make the results of experimental work available to consumers.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of our partners in the project, Grace Bioremediation Technologies and the University of Guelph, Department of Environmental Biology. Thanks is also accorded to Environment Canada's Development and Demonstration of Site Remediation Technology and Environmental Innovation Programs for funding this project.

REFERENCES

Atlas, R.M. and R. Bartha. 1981. *Microbial Ecology: Fundamentals and Applications*. London, England: Addison Wesley Publishing Co.

Bremner, J.M. and C.S. Mulvaney. 1982. Nitrogen-Total. In *Methods of Soil Analysis, Part 2 -Chemical and Microbiological Properties*, 2nd ed, ed. A.L. Page, Chapter 31. Madison WI: American Society of Agronomy, Inc., Soil Science Society of America.

Briglia, M., P.J.M. Middeldorp and M.S. Salkinoja-Salonen. 1994. Mineralization performance of *Rhodococcus chlorophenolicus* strain PCP-1 in contaminated soil simulating on site conditions. *Soil Biology and Biochemistry* 26: 377-385.

Knudsen, D., G.A. Peterson and P.F. Pratt. 1982. Lithium, Sodium, and Potassium. In *Methods of Soil Analysis, Part 2 - Chemical and Microbiological Properties*, 2nd ed, ed. A.L. Page, Chapter 13. Madison WI: American Society of Agronomy, Inc., Soil Science Society of America.

Kok, R., R. Lacroix, G. Clark, and E. Taillefer. 1994. Imitation of a procedural greenhouse model with an artificial neural network. *Canadian Agricultural Engineering* 36: 117-126.

Lanyon, L.E. and W.R. Heald. 1982. Magnesium, Calcium, Strontium, and Barium. In *Methods of Soil Analysis, Part 2 - Chemical and Microbiological Properties*, 2nd ed, ed. A.L. Page, Chapter 14. Madison WI: American Society of Agronomy, Inc., Soil Science Society of America.

Linko, S., T. Rajalahti and Y.-H. Zhu. 1995. Neural state estimation and prediction in amino acid fermentation. *Biotechnology Techniques* 9: 607-612.

Nelson, D.W. and L.E. Sommers. 1982. Total carbon, organic carbon, and organic matter. In *Methods of Soil Analysis, Part 2 - Chemical and Microbiological Properties*, 2nd ed, ed. A.L. Page, Chapter 29. Madison WI: American Society of Agronomy, Inc., Soil Science Society of America.

NeuralWare. 1995a. *Neural Computing, A Technology Handbook for Professional II/PLUS and NeuralWorks Explorer*. Pittsburgh, PA: NeuralWare, Inc.

NeuralWare. 1995b. *Using NeuralWorks, A Tutorial for NeuralWorks Professional II/PLUS and NeuralWorks Explorer*. Pittsburgh, PA: NeuralWare, Inc.

Olson, S.R. and L.E. Sommers. 1982. Phosphorus. In *Methods of Soil Analysis, Part 2 - Chemical and Microbiological Properties*, 2nd ed, ed. A.L. Page, Chapter 24. Madison WI: American Society of Agronomy, Inc., Soil Science Society of America.

Pramer, D. and R. Bartha. 1972. Preparation and processing of soil samples for biodegradation studies. *Environmental Letters* 2: 217-224.

Riser-Roberts, E. 1992. *Bioremediation of Petroleum Contaminated Sites*. Boca, FL: C.K. Smoley.

Sharabi, N.E. and R. Bartha. 1993. Testing of some assumptions about biodegradability in soil as measured by carbon dioxide evolution. *Applied and Environmental Microbiology* 59: 1201-1205.

Shukla, M.B., R. Kok, S.O. Prasher, G. Clark and R. Lacroix. 1996. Use of artificial neural networks in transient drainage design. *Canadian Agricultural Engineering* 38: 119-124.

Tyagi, R.D. and Y.G. Du. 1992. Kinetic model for the effects of heavy metals on activated sludge process using neural networks. *Environmental Technology* 13: 883-890.

Widrow, B., R.G. Winter, and R.A. Baxter. 1988. Layered neural nets for pattern recognition. *IEEE Transaction on Acoustics, Speech, and Signal Processing* 36: 1109-1118.

Zhang, Q., J.F. Reid, J.B. Lirchfield, J. Ren and S.-W. Chang. 1994. A prototype neural network supervised control system for *Bacillus thurigiensis* fermentations. *Biotechnology and Bioengineering* 43: 483-489.