1 Introduction

While the transmission of electrical signals across neuronal networks is a fundamental aspect of the operation of nervous systems, and this feature has traditionally been the main focus of computational neuroscience (Koch and Segev 1998, Dayan and Abbott 2001), neurochemistry adds many dimensions to the picture. For instance, it is now recognised that nitric oxide (NO) is a novel kind of neurotransmitter that acts, through diffusion, over volumes that potentially contain many neurons and can facilitate signalling between neurons that are not synaptically connected (Gally et al. 1990, Wood and Garthwaite 1994, Vincent 1994). This article aims to demonstrate that computational and mathematical modelling have an important role to play in trying to understand this particularly interesting mode of signalling.

Traditionally, chemical signaling between nerve cells was thought to be mediated solely by messenger molecules or neurotransmitters which are released by neurons at synapses (Katz 1969), and flow from the presynaptic to postsynaptic neuron. Because most neurotransmitters are relatively large and polar molecules (amino acids, amines and peptides), they cannot diffuse through cell membranes and do not spread far from the release site. They are also rapidly inactivated by various reactions. Together these features confine the spread of such neurotransmitters to be very close to the points of release and ensure that the transmitter action is transient. In other words, chemical synaptic transmission of the classical kind operates essentially two-dimensionally (one in space and one in time). This conventional interpretation is coupled to the idea that neurotransmitters cause either an increase or a decrease in the electrical excitability of the target neuron. According to a traditional view of neurotransmission therefore, chemical information transfer is limited to the points of connection between neurons and neurotransmitters can simply be regarded as either excitatory or inhibitory. In recent years a number of important discoveries have necessitated a fundamental revision of this model. It is now clear that many neurotransmitters, perhaps the majority, cannot be simply classified as excitatory or inhibitory (Hall 1992). These messenger molecules are best regarded as modulatory because among other things they regulate, or modulate, the actions of conventional transmitters. Modulatory neurotransmitters act in an indirect way by causing medium and long-term changes in the properties of neurons by influencing the rate of synthesis of so called second messenger molecules. By altering the properties of proteins and even by changing the pattern of gene expression, these second messengers cause complex cascades of events resulting in fundamental changes in the properties of neurons. In this way modulatory transmitters greatly expand the diversity and the duration of actions mediated by the chemicals released by neurons.
However, when coupled with this expanded picture of the nervous system, it is the recent discovery that the gas nitric oxide is a modulatory neurotransmitter that has opened entirely unexpected dimensions in our thinking about neuronal chemical signaling (Gally et al. 1990, Garthwaite et al. 1988, Holscher 1997). Because NO is a very small and nonpolar molecule it diffuses isotropically in aqueous and lipid environments, such as the brain, regardless of intervening cellular structures (Wood and Garthwaite 1994). NO therefore violates some of the key tenets of point-to-point chemical transmission and is the first known member of an entirely new class of transmitter, the gaseous diffusable modulators (CO and H₂S are the other two identified examples (see e.g. Cao et al. 2000). NO is generated in the brain by specialised neurons that contain the neuronal isoform of the calcium activated enzyme, nitric oxide synthase or nNOS (Bredt and Snyder 1990). NO synthesis is triggered when the calcium concentration in nNOS-containing neurons is elevated, either by electrical activity or by the action of other modulatory neurotransmitters. NO activates the synthesis of cyclic-GMP, an important second messenger which regulates a wide variety of cellular processes in target neurons, some of which underlie synaptic plasticity (Holscher 1997). Hence NO is involved in many neuronal functions from visual processing to memory formation and blood flow regulation (Vincent 1994, Garthwaite and Boulton 1995, Holscher 1997).

The existence of a freely diffusing modulatory transmitter suggests a radically different form of signalling in which the transmitter acts four-dimensionally in space and time, affecting volumes of the brain containing many neurons and synapses. NO cannot be contained by biological membranes, hence its release must be coupled directly to its synthesis. Because the synthetic enzyme nNOS can be distributed throughout the neuron, NO can be generated and released by the whole neuron. NO is therefore best regarded as a ‘non-synaptic’ transmitter whose actions moreover cannot be confined to neighbouring neurons (Hartell 1996, Park et al. 1998). So not only can NO operate over a large region, it can also mediate long-lasting changes in the chemical and electrical properties of neurons within that volume (Baranano et al. 2001, Snyder and Ferris 2000).

Because nNOS is a soluble enzyme and thus likely to be distributed throughout a neuron’s cytoplasm, the whole neuron surface is a potential release site for NO. Thus the morphology of NO sources, as well as the presence of structured sinks (such as blood vessels), will have a major influence on the dynamics of NO spread. Understanding this dynamics is clearly a very important part of a more general understanding of volume signalling processes. However, because the NO molecule is so small and non-polar it is very difficult to gather accurate empirical data in this area. Therefore it is natural to turn to computational modelling to shed light on volume signalling.

Somewhat ironically, many of NO’s characteristics that complicate its empirical investigation, make it much easier to model than many conventional neurotransmitters whose large size and polarity make them impermeable to cell membranes. Thus while these molecules also diffuse, their movement is restricted to the extracellular space near their release site and to model their spread would therefore require accurate modelling of the morphology of the extracellular space and any local inhomogeneities. In contrast, because of NO’s minute size and non-polarity it can be assumed as a good first approximation to diffuse isotropically through most brain tissue and so the morphology of the synaptic cleft and other surrounding matter need not be modelled. This means that complex factors such as tortuosity and viscosity, which affect the movement of larger molecules, do not need to be included in the governing equations.

This article demonstrates how to model NO diffusion from continuous structures of biologically realistic dimensions. The central part of the chapter describes and justifies in some detail
the methods used to build such models. It then goes on to show how these models provide insights into a number of salient functional questions that arise in the context of volume signalling. Chief among these is how large a volume can be affected, and for how long, from various NO generating neuronal structures. Finally work on more abstract computational models of neural networks incorporating functionally active diffusing neuromodulators is introduced. These networks serve as the nervous system of autonomous robots, generating sensorimotor behaviours in these devices, and thus help to give insights into possible functional roles of gaseous diffusing modulators in real nervous systems.

2 Methods

This section gives a detailed overview of the methods used to model the diffusion of NO in the CNS.

2.1 Equations governing NO diffusion in the brain

The rate of change of concentration in a volume element of a membrane, within the diffusional field, is proportional to the rate of change of concentration gradient at that point in the field. Fick’s second law (Fick 1855)

\[ \frac{\partial C(x, t)}{\partial t} = D \left( \frac{\partial^2 C(x, t)}{\partial x^2} + \frac{\partial^2 C(x, t)}{\partial y^2} + \frac{\partial^2 C(x, t)}{\partial z^2} \right) \]  

or, more generally:

\[ \frac{\partial C(x, t)}{\partial t} = D \nabla^2 C(x, t) \]  

While the above equations govern the diffusive element of NO’s spread, they do not take into account its destruction. NO does not have a specific inactivating mechanism, and is lost through reaction with oxygen species and metals, as well as heme containing proteins (Lancaster

\footnote{Diffusion processes are thus amenable to Monte Carlo methods, where a (in the case of diffusion) uniform probability distribution, representing the probability of a molecule moving in a given direction, together with a random number generator are used to calculate the path of each molecule. However, the relatively long running times to achieve a good approximation render this methods inappropriate for our needs (Ames 1992).}
1996, Vaughn et al. 1998a). This means that the movement of other molecules and receptors and their interactions need not be modelled and instead a more general loss function can be used. Thus we have:

\[
\frac{\partial C(x,t)}{\partial t} - D \nabla^2 C(x,t) = -L(C, x, t)
\] (3)

where the term on the right hand side is the inactivation function. This function will be composed of a global component for general, background NO reactions, \(L_1(C)\), and spatially localised components, \(L_2(C, x, t)\), representing structures which act as local NO sinks such as blood vessels. The kinetics of these reactions are not understood perfectly (Wood and Garthwaite 1994), but empirical data indicates either first or second order decay (Laurent et al. 1996, Lancaster 1996, Vaughn et al. 1998b, Liu et al. 1998, Thomas et al. 2001), as represented by:

\[
L_i(C, x, t) = k_i(x) \times C(x, t)^n
\] (4)

where \(n\) is the order of the reaction and is equal to either 1 or 2, referred to as first or second decay order respectively, and \(k_i\) is the reaction rate, commonly given in terms of the half-life \(t_{1/2} = \ln(2)/k_i\). The reaction may also depend on the concentration of the oxidative substance (usually oxygen), but, apart from very special cases (as in Chen et al. 1998), for example), this can be assumed to be constant and is left out of the equation as it is subsumed by the reaction rate constant.

Values for the half-life of NO have been determined empirically and are dependent on the chemical composition of the solvent within which NO is diffusing. For instance, the half-life of NO in the presence of haemoglobin (Hb) is reported as being between about 1 ms and 1 \(\mu\)s, depending on the Hb concentration (Carlsen and Comroe 1958, Lancaster 1997, Vaughn et al. 1998a, Liu et al. 1998). In contrast, half-life values used for extravascular tissue, normally associated with background NO consumption, are more than 1000 times longer, ranging from 1 to > 5 s (Moncada et al. 1989, Malinski et al. 1993, Wood and Garthwaite 1994, Thomas et al. 2001).

The order of the reaction is also dependent on the nature of the diffusive environment. In environments where there is a high concentration of Hb, as in NO sinks, recent work by Liu et al. (1998) has shown that the reaction of NO with intact red blood cells exhibits first order kinetics, which is in agreement with earlier measurements (Carlsen and Comroe 1958, Lancaster 1997, Vaughn et al. 1998a). Similarly, for modelling the global part of the loss function, as would be seen in most extravascular regions of the brain, measurements of NO loss are also consistent with first order decay (Lancaster 1996, Liu et al. 1998, Thomas et al. 2001). Although second order decay has been used for extravascular NO consumption (Laurent et al. 1996, Vaughn et al. 1998b), in these models NO is diffusing in vitro in an air-saturated aqueous solution which is molecular-oxygen rich (unlike intact extravascular brain tissue). Thus the dynamics of decay are taken from empirical data in molecular-oxygen rich environs and are unlike those in the intact brain. As we are concerned with modelling NO diffusion in the brain in vivo, we have therefore used first order decay to model global NO loss in extravascular tissue as well as in localised sinks. This gives us the following widely-used equation (Wood and Garthwaite 1994, Lancaster 1994, 1996, 1997, Vaughn et al. 1998a, Thomas et al. 2001) for diffusion of NO in the brain:

\[
\frac{\partial C(x, t)}{\partial t} - D \nabla^2 C(x, t) = -k(x) \times C(x, t)
\] (5)
referred to as the modified diffusion equation. A production term can also be added to the right hand side of equation 5 though this is often factored into the solution later via the initial conditions (see section 2.2).

Under certain conditions and for some source morphologies, equation 5 can be solved analytically (the analytical solution), although this usually involves some numerical integration. Since the more complex a system is, the more numerical integration is required, this approach is often impractical and, in general, radial symmetry is required for tractability. If the analytical solution cannot be derived, a numerical approximation method must be used (Lancaster 1997). That is not to say that the numerical solutions are somehow ‘worse’ than the analytic ones or that they are simply crude approximations to the true solution (Smith 1985). Rather, they can usually be made as accurate as desired, or as accurate as the situation warrants given the unavoidable errors in empirical measurements of diffusion parameters. Indeed, as all the analytic solutions presented here required numerical integration they are also approximate and all results have been derived to the same degree of accuracy. However, a numerical approximation is always an approximation to the analytical solution and so it seems sensible to use the latter if its calculation is tractable. A more practical reason for doing so is that when it is available, evaluating the analytical solution normally requires much less computational power. Our approach therefore is to use the analytical solution whenever possible and, when not, to employ finite difference methods. In the next two sections, we discuss these techniques.

2.2 Analytic solutions to the diffusion equation

*God does not care about our mathematical difficulties. He integrates empirically.* Albert Einstein.

In this section we discuss how analytic solutions to the diffusion equation are generated. The solution for a point-source is given first and we then show how solutions for other simple structures are derived from this. We next state the solutions thus obtained for hollow spherical and tubular sources and finally give some details of the numerical integration techniques used to calculate these solutions.

2.2.1 Modelling NO diffusion from a point-source

As stated earlier (section 2.1), the dynamics of diffusion are governed by the modified diffusion equation. Assuming that there are no local NO sinks present and only global decay is acting, this equation becomes:

\[
\frac{\partial C}{\partial t} - D \nabla^2 C = -\lambda C
\]

where \(C\) is concentration, \(D\) is the diffusion coefficient and \(\lambda\) the decay-rate (Crank 1980). We first generate the instantaneous solution, that is, the solution for an instantaneous burst of synthesis from a point source positioned at the origin of some co-ordinate system. To do this we envision an amount \(S_0\) of NO being deposited instantaneously at the origin at time \(t = 0\). We then solve the diffusion equation under this initial condition, which gives us the following equation describing the evolution of the concentration of NO from a point (Crank 1980):

\[
C_P(r, t) = \frac{S_0}{8(\pi Dt)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right)e^{-\lambda t}
\]
where \( C_P(r, t) \) is the concentration of NO at time \( t \) at a point \((r, \theta, \phi)\), defined in a spherical polar coordinate system. Note, however, that \( C_P \) depends only on time and the distance from the point source, \( r \), as the system is radially symmetric.

The solution for a point-source which emits NO continuously, the continuous solution, is derived from the instantaneous solution for a point source, described earlier, in the natural way, via the principle of superposition of linear solutions (Crank 1980). First we define the ‘strength’ of a source to be its rate of NO production. Next we define the concentration at time \( t' \) and distance \( r \) from the origin, due to an instantaneous source of unit strength to be \( f(r, t') \). Thus, if a source emits NO continuously at a rate governed by \( S(t) \), we have:

\[
C(r, t) = \int_0^t S(t - t') f(r, t') \, dt'
\]  

(8)

This can be understood by seeing that the contribution at time \( t' \leq t \) is due to an instantaneous pulse of NO \( t' \) seconds previously, with \( S(t - t') \) the amount of NO per second produced at time \( t - t' \) that is, \( t' \) seconds earlier. Thus, in equation 8, the most recent pulses of NO are responsible for the lower limit of the integration, whilst the oldest pulses account for the upper limit. Similarly, we can derive the solution for times after a source which emitted NO continuously has stopped synthesising. If the source synthesises for \( T \) seconds and, as before, the instantaneous solution is \( f(r, t') \), then the concentration at a distance \( r \) from the source, \( t_1 \) seconds after it has stopped synthesising is:

\[
C(r, t_1 + T) = \int_{t_1}^{t_1 + T} S(t_1 + T - t') f(r, t') \, dt'
\]  

(9)

where \( t_1 > 0 \). This approach is valid since the diffusion equation is linear and the principle of superposition of linear solutions therefore applies.

### 2.2.2 Modelling NO diffusion from a symmetrical 3D structure

To model the spread of an amount of NO produced instantaneously throughout a continuous structure, we use methods developed in the field of thermodynamics which are readily applicable to modelling diffusion (Carslaw and Jaeger 1959). The main technique is to build up solutions for complicated structures from summation of contributions from point sources distributed throughout the structure. Of course, we are not implying that there are an infinite number of NO sources in the structure, but they are small enough that we are justified in imagining that they are uniformly distributed throughout the source with some density \( \rho \) (see section 2.4.3). Hence for a spherical source, \( M \), of radius \( a \), the method is to sum the contributions to the concentration at a point in space \( y \), from all the points within the sphere as described below.

![Figure 1](image-url): The element \( X \) (see equation 10) of a sphere of radius \( a \) containing the points \((r', \theta', \phi')\) where: \( r \leq r' \leq r + \delta r; \ \theta \leq \theta' \leq \theta + \delta \theta \) and \( \phi \leq \phi' \leq \phi + \delta \phi \). \( X \) is outlined by solid lines with dashed lines denoting radii and surface of the sphere.
Take a volume $X$ within the sphere containing the points $(r', \theta', \phi')$ where:

$$r \leq r' \leq r + \delta r; \quad \theta \leq \theta' \leq \theta + \delta \theta; \quad \phi \leq \phi' \leq \phi + \delta \phi. \quad (10)$$

as shown in figure 1. If this element, $X$, is relatively small (i.e. if $\delta r$, $\delta \theta$ and $\delta \phi$ are sufficiently small), we can approximate its volume with:

$$V_X \approx r^2 \sin \theta \delta \theta \delta \phi \delta r \quad (11)$$

with the error in the approximation getting smaller as the dimensions of the element ($\delta r$, $\delta \theta$ and $\delta \phi$) are reduced and the error becoming zero in the limit of the dimensions becoming vanishingly small. Now, the amount of NO produced per second in a volume $V$ of NO-producing tissue is:

$$S_V = Q \times N_V \quad (12)$$

where $Q$ is the amount of NO produced per second from a single NO producing unit and $N_V$ is the number of these units within $V$. This number is simply the product of the volume of $V$ and $\rho$, the density of units in $V$. Hence for the element, $X$ (figure 1), we have a strength/second term, $S_X$, of:

$$S_X = Q \times N_X = Q \times \rho V_X \approx Q\rho r^2 \sin \theta \delta \theta \delta \phi \delta r \quad (13)$$

In this equation, $r$, $\theta$ and $\phi$ are variables whilst the product $Q\rho$, the concentration of NO produced per second, is independent of the particular shape of the structure being studied and so can be determined by empirical experiments as in Wood and Garthwaite (1994).

Making $\delta r$, $\delta \theta$ and $\delta \phi$ vanishingly small makes the approximation in equations 11 and 13 exact and we can assume that the contribution to the concentration at a point, $y = (\tilde{r}, 0, 0)$, from the volume $X$ is as if from a point source at the point: $x = (r, \theta, \phi)$. Inspecting equation 7 we see that the concentration depends on the time after synthesis and the distance between $x$ and $y$, $\|x - y\|$. Substituting this and our strength/second term, $S_X$, from equation 13 into equation 7 we obtain the concentration of NO at $y$ due to $x$ at a time $t$ after synthesis:

$$C_P(\|x - y\|, t) = \frac{Q\rho r^2 \sin \theta \delta \theta \delta \phi \delta r}{8(\pi Dt)^{3/2}} \exp\left(-\frac{\|x - y\|^2}{4Dt}\right) e^{-\lambda t} \quad (14)$$

Now, to get the concentration at $y$ due to the whole sphere we must sum up the contributions from all the points $x = (r, \theta, \phi)$ inside the sphere, $M$, (i.e. $0 \leq r \leq a; \ 0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$) as shown below:

$$C_S(a, \tilde{r}, t) = \sum_{x \in M} C_P(\|x - y\|, t) = \int_0^a \int_0^{2\pi} \int_0^\pi \frac{Q\rho r^2 \sin \theta}{8(\pi Dt)^{3/2}} \exp\left(-\frac{\|x - y\|^2}{4Dt}\right) e^{-\lambda t} \, dr \, d\theta \, d\phi \quad (15)$$

Using:

$$\|x - y\|^2 = \tilde{r}^2 + r^2 - 2\tilde{r}r \cos \theta \quad (17)$$

and noting that there is radial symmetry so that the concentration at any point $z = (\tilde{r}, \theta, \phi)$ at a distance of $\tilde{r}$ from the origin is equal to the concentration at $y = (\tilde{r}, 0, 0)$, we therefore obtain:
\[ C_S(a, \tilde{r}, t) = Q\rho e^{-\lambda t} \left[ \frac{1}{2} \left( \text{erf} \left( \frac{a + \tilde{r}}{2\sqrt{Dt}} \right) + \text{erf} \left( \frac{a - \tilde{r}}{2\sqrt{Dt}} \right) \right) \right. \]
\[ \left. - \frac{1}{\sqrt{\pi}} \frac{\sqrt{Dt}}{\pi} \left( \exp \left( \frac{(a + \tilde{r})^2}{4Dt} \right) - \exp \left( \frac{(a - \tilde{r})^2}{4Dt} \right) \right) \right] \]  

where:

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp \left( -u^2 \right) \, du \]  

for the concentration at a distance \( \tilde{r} \) from the centre of a solid sphere of radius \( a \) at a time \( t \) after synthesis. This leads naturally to the solution for a hollow sphere of inner radius \( a \) and outer radius \( b \):

\[ C_H(a, b, \tilde{r}, t) = C_S(b, \tilde{r}, t) - C_S(a, \tilde{r}, t) \]  

The analytical method outlined also yields the concentration at distance \( \tilde{r} \) from the centre of an annulus of inner radius \( R_1 \) and outer radius \( R_2 \) at a time \( t \) after synthesis:

\[ C_A(R_1, R_2, \tilde{r}, t) = \frac{Q\rho}{2Dt} e^{-\lambda t} \exp \left( -\frac{\tilde{r}^2}{4Dt} \right) \int_{R_1}^{R_2} \exp \left( -\frac{r^2}{4Dt} \right) I_0 \left( \frac{r\tilde{r}}{2Dt} \right) \, dr \]  

where \( I_0(x) \) is the modified bessel function of order zero (Carslaw and Jaeger 1959).

These ‘instantaneous’ solutions can then be integrated over the appropriate time intervals to get the solutions for the evolution of concentration of NO synthesised for a finite time interval, in the same way as for the point source (as in equations 8 and 9). In these cases, however, as the volume term has already been implicitly factored into the integrals, we replace \( Q\rho \), the concentration/second at each instant, with \( \tilde{S}(t) \), a function which, for each time \( t \), gives the value of \( Q\rho t \) seconds after the start of synthesis. Traditionally, instantaneous switch-on and off of synthesis has been assumed meaning that \( \tilde{S}(t) \) will be a square wave with maximum value of \( Q\rho \). This has the advantage of simplicity since \( \tilde{S}(t) \) is now constant and can be moved outside the integral in 8 and 9. Obviously, such a mechanism of release is not strictly biologically plausible but, in the absence of experimental data on the kinetics of nNOS activation in vivo and given the insensitivity of the diffusion process to small scale heterogeneity (Philippides 2001), this is a reasonable approximation. However, if numerical integration techniques are being used, more complicated strength functions can be used to model the time-course of NO synthesis. For instance, in Philippides et al. (2000) we used a strength function relating the amount of NO released to the amount of depolarisation caused by an action potential.

2.2.3 Numerical integration of analytical solutions

As mentioned previously, numerical integration is generally required to generate the analytic solutions especially if any reasonable inactivation is included. We will now describe the methods used for the results detailed below, but a full review of numerical integration techniques can be found in, for example, Davis and Rabinowitz (1984).

Evaluating the continuous solution for the spherical source (equation 18) requires numerical integration over time. This was performed by the ‘quad8’ function in the programming language Matlab, which uses an adaptive recursive Newton Cotes 8 panel rule (Davis and Rabinowitz 1984) to a relative accuracy of 0.1%. This is an extension of the extended trapezoidal rule in which the
integral of \( y(x) \) between \( x_A \) and \( x_B \) is estimated by dividing the range into \( N \) sections of width \( h \), and approximating the area under the curve in each segment by the area of a trapezium, giving:

\[
\int_{x_A}^{x_B} y(x) \, dx = h \left[ \frac{1}{2} y_0 + y_2 + y_3 + \ldots + y_{N-1} + \frac{1}{2} y_N \right] + O \left( \frac{(x_B - x_A)^3 y''}{N^2} \right)
\]  \hspace{1cm} (22)

where \( y_i = y(x_A + ih) \). While this method is robust for functions that are not very smooth, it is relatively slow and the ‘quad8’ function achieves much faster convergence by adaptively changing the positions and weightings of the estimates \( y_i \) in equation 22. However, the extra speed comes at the cost of a loss of accuracy over the less smooth areas of the integrand near the temporal origin. This means that the integration procedure must be modified for cases where the lower limit of integration in equation 18 is less than 1 ms to include the following analytical approximation for the second part of the integral:

\[
\int_{\varepsilon}^{t} Q e^{-\lambda t} \frac{e^{-r^2/4Dt}}{\pi} \left[ \exp \left( \frac{(a-r)^2}{4Dt} \right) - \exp \left( \frac{(a+r)^2}{4Dt} \right) \right] dt
\]  \hspace{1cm} (23)

\[
\approx Q \rho \left( 1 + e^{-\lambda \varepsilon} \right) \left[ h \left( \frac{(r-a)^2}{4D}, r, \varepsilon \right) - h \left( \frac{(r+a)^2}{4D}, r, \varepsilon \right) \right]
\]  \hspace{1cm} (24)

where:

\[
h(k, r, \varepsilon) = \int_{\varepsilon}^{0} \frac{1}{r} \sqrt{\frac{Dt}{\pi}} e^{-\frac{k^2}{4r^2}} dt = \frac{2}{3r \sqrt{\pi}} \left[ e^{-\frac{k^2}{4r^2}} \sqrt{\varepsilon (\varepsilon - 2k)} + 2 \sqrt{\pi k^3} \text{erfc} \left( \frac{k}{\sqrt{\varepsilon}} \right) \right]
\]  \hspace{1cm} (25)

\[
\text{erfc}(x) = 1 - \text{erf}(x)
\]  \hspace{1cm} (26)

and noting that at \( t = 0 \) the instantaneous solution is:

\[
C_S(a, r, t) = \begin{cases} 
Q \rho & \text{for } r < a \\
Q \rho / 2 & \text{for } r = a \\
0 & \text{else}
\end{cases}
\]  \hspace{1cm} (27)

In the above \( a \) is the radius of the sphere, \( r \) is the distance from its centre and \( \varepsilon \leq 1 ms \). Solutions for the hollow sphere and when the lower limit of integration is greater than zero can be derived from the above equations.

The approximation in equation 23 is based on the principle that if \( f_{\min[0,\varepsilon]} \) is the minimum value attained by a function \( f(t) \) over the range \([0, \varepsilon]\) and \( f_{\max[0,\varepsilon]} \) is the maximum of \( f(t) \) over the same range then:

\[
\int_{0}^{\varepsilon} f(t) g(t) \, dt \simeq \frac{f_{\min[0,\varepsilon]} + f_{\max[0,\varepsilon]}}{2} \int_{0}^{\varepsilon} g(t) \, dt
\]  \hspace{1cm} (28)

which has a maximum error of:

\[
\frac{f_{\max[0,\varepsilon]} - f_{\min[0,\varepsilon]}}{2} \int_{0}^{\varepsilon} g(t) \, dt
\]  \hspace{1cm} (29)
Thus, the actual value of the error is dependent on the parameter values used but, for the parameters used here, the errors are small enough to keep the solutions to within a relative accuracy of 0.1%.

The continuous solution for the tubular source (equation 21) has to be integrated over both space and time, necessitating a slightly different approach since multi-dimensional integration is significantly more time-consuming and can magnify errors and instabilities in the methods used (Press et al. 1971). While Monte-Carlo integration can be used, it is inappropriate for our needs due to the asymptotically slow convergence. Due to the accuracy requirements and relative smoothness of the function, the approach we have taken is to use successive applications of one-dimensional integration. In this method, to evaluate $y(x, t_i)$ for each abscissa, $t_i$, of an iteration of the outer integration of $\int \int y(x, t)dx dt$, we must perform one whole numerical integration over $x$ using that value of $t_i$ to evaluate $y(x_j, t_i)$ at each $x_j$. This means that if it takes $N$ function evaluations to get a sufficiently accurate estimate for the one-dimensional integral, we will need around $N^2$ evaluations to achieve the same accuracy for the two-dimensional integral. Moreover, since it is good practice for accuracy reasons to use simple numerical integration routines, the exponential growth in the number of operations needed is exacerbated by the slow convergence of these methods. Given these considerations, solutions for reasonably complicated functions requiring greater than double integration is probably best handled with one of the numerical methods discussed in the next section.

For the results detailed here, the extended trapezoidal rule given in equation 22 (Press et al. 1971) was used for the outer integration (over time) with the inner integration (over radial distance) performed by the ‘quad8’ function to speed up convergence. However, to ensure accuracy, we checked the solutions by performing both inner and outer integrations using the extended trapezoidal rule. In evaluating the continuous solution for the tubular source, it should be noted that at $t = 0$ the instantaneous solution is:

$$C_A(R_1, R_2, r, t) = \begin{cases} 
Q \rho & \text{for } R_1 < r < R_2 \\
Q \rho/2 & \text{for } r = R_1, R_2 \\
0 & \text{else}
\end{cases} \quad (30)$$

Solutions were accurate to a relative accuracy of 0.5%. Accuracy of solutions for the tubular and spherical sources were further checked using the numerical integration package in the programming language Maple which is very accurate and accounts for improper integrals correctly, but is too slow for general use.

### 2.3 Modelling diffusion of NO from an irregular 3D structure

#### 2.3.1 Finite difference methods for diffusive problems

From the previous section it is clear that the analytical method is not tractable for many situations which we might want to investigate. In particular, modelling irregularly shaped sources and sinks is inappropriate and other numerical techniques for solving the partial differential equations (PDEs) governing the spread of NO must be used. For diffusive problems evolving over the short time-scales associated with NO diffusion in the brain, one recommended approach is to use finite differences (Press et al. 1971, Ames 1992).

These methods proceed by approximating the continuous derivatives at a point by difference quotients over a small interval, for instance replacing $\frac{\partial x}{\partial t}$ by:

$$\frac{\delta x}{\delta t} = \frac{x(t + \Delta t) - x(t)}{\Delta t} \quad (31)$$
In this way, given some initial conditions for \( x \) at \( t = t_0 \), we can define a recurrence relation:

\[
x_0 = a, x_{n+1} = x_n + \Delta t f(x_n, t_n), \quad \text{where: } x_n = x(t_n), t_n = t_0 + n\Delta t
\]

(32)

which can be solved iteratively. The error in the approximation is dependent on the size of \( \Delta t \), the step-size, with schemes being said to be \( n \)'th order accurate in a given variable (or variables), meaning that the error is essentially a constant multiplied by the step-size raised to the \( n \)'th power (Ames 1992). As well as governing this truncation error, one must also ensure that the spatial and temporal step-sizes used do not make the set of equations unstable, resulting in erroneous answers. For instance, explicit difference equations (DEs), where the values at time-step \( n + 1 \) are calculated using only values known at time \( n \) as in equation 32 above, tend to have stability problems. Thus, compartmental models, a common finite difference method used to solve the diffusion equation (equation 2) (Gally et al. 1990, Lancaster 1994, 1997), are hampered by the fact that for stability:

\[
\frac{D\Delta t}{(\Delta x)^2} \leq \frac{1}{2^n}
\]

(33)

where \( D \) is the diffusion coefficient, \( n \) is the spatial dimension and \( \Delta x \) and \( \Delta t \) are the spatial and temporal step-sizes respectively (Press et al. 1971). This puts a limitation on the size of the time-step to be used which, in less abstract terms, means that, in one space dimension, it must be less than the diffusion time across a cell of width \( \Delta x \).

However, different schemes have different stability properties and so the restrictive bounds of the compartmental model can be avoided. For instance, implicit DEs, where values at time \( n + 1 \) are defined in terms of each other, are often stable for all step-sizes. However, while explicit DEs are inherently easy to solve as the solution is simply propagated forward in time, implicit DEs require the solution of a set of simultaneous equations (Mascagni 1989). In order to avoid computationally intensive routines, it is therefore important that the DE is designed so that the resulting system of equations is tridiagonal\(^2\). One such equation, known as the Crank-Nicholson scheme, is recommended for diffusive problems in one space dimension (Press et al. 1971). Applied to the one-dimensional version of equation 2 we have, using the notation of equation 33:

\[
\frac{u^n_{i+1} - u^n_i}{\Delta t} = D \left[ \frac{(u^{n+1}_{i+1} - 2u^n_{i+1} + u^{n+1}_{i-1}) + (u^n_{i+1} - 2u^n_i + u^n_{i-1})}{(\Delta x)^2} \right]
\]

(34)

where \( u^n_i \) is the concentration at time \( n\Delta t \) and position \( i\Delta x \). This scheme is second order accurate in space and time while maintaining stability for all choices of \( \Delta t \).

This equation can be generalised for higher spatial dimensions quite easily but the resulting systems of equations are no longer tridiagonal and so much more computationally expensive to solve. As the number of operations required to solve multi-dimensional DEs increases exponentially with the dimension (as in numerical integration), this method is impractical (Ames 1992). Generalising explicit schemes in this way is feasible due to the speed with which they can be solved.

---

\(^2\)A tridiagonal system of equations \( Ax = b \) is one where the matrix \( A \) is tridiagonal, that is, where the elements of \( A, a_{ij} \), equal 0 if \( i > j + 1 \) or \( j > i + 1 \). In other words, if the row and column number differ by more than one, the entry must be zero. The entries are otherwise unrestricted. In non-mathematical terms this means that the resulting equations can be solved quite straightforwardly in \( O(n) \) operations where \( n \) is the number of equations, which in the context of difference equations equates to the number of spatial points at which the equation is to be evaluated.
solved. However the limit on the step-sizes that can be used while ensuring stability (equation 33) becomes even more restrictive obviating this approach as well.

To get round these problems one can use a second class of techniques, known as alternating-direction implicit (ADI) methods. The basic idea in these schemes is to split up a single time-step into \( n \) sub-steps, one for each spatial dimension. For each one of these sub-steps we evaluate only one spatial derivative at the advanced time step which ensures that the resultant sub-system is tridiagonal, and thus solving for each coordinate direction in turn. To see the general principle consider the following generalisation of equation 34 the two dimensional diffusion equation:

\[
\frac{\partial u}{\partial t} = D \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \tag{35}
\]

Defining:

\[
u_n \equiv u_{i,j,n} \equiv u(i\Delta x, j\Delta y, n\Delta t) \tag{36}
\]

\[
\delta_x^2 u_n \equiv \delta_x^2 u_{i,j,n} \equiv u_{i+1,j,n} - 2u_{i,j,n} + u_{i-1,j,n} \tag{37}
\]

\[
\delta_y^2 u_n \equiv \delta_y^2 u_{i,j,n} \equiv u_{i,j+1,n} - 2u_{i,j,n} + u_{i,j-1,n} \tag{38}
\]

where \( \Delta x, \Delta y \) and \( \Delta t \) are the spatial and temporal step-sizes respectively, we have:

\[
\frac{u_{n+1/2} - u_n}{\frac{1}{2}\Delta t} = D \left( \delta_x^2 u_{n+1/2} + \delta_y^2 u_n \right) \tag{39}
\]

which is tridiagonal. This is solved everywhere and the solution for the half time-step is then used in the following difference:

\[
\frac{u_{n+1} - u_{n+1/2}}{\frac{1}{2}\Delta t} = D \left( \delta_x^2 u_{n+1/2} + \delta_y^2 u_{n+1} \right) \tag{40}
\]

to solve for the full time-step. This results in a scheme which is stable, second order accurate in space and time and only requires solution of tridiagonal systems. Such methods, while certainly requiring a significant amount of computation, are at least practical and have been used extensively for diffusive IVPs (Press et al. 1971, Ames 1992).

Before detailing the specific difference schemes used, it should be noted that the major issue in using difference equations for multi-dimensional diffusive DEs is that of computational power. As well as the number of operations required scaling exponentially with the number of dimensions, so do the memory (RAM) requirements. One could attempt to alleviate this by reducing the problem’s size by using large spatial and temporal scales and offsetting the loss of accuracy incurred by employing higher-order (in terms of accuracy) methods. However, unless the DE is extremely stable, using too high an order can introduce spurious solutions and for second order initial value problems, such as the diffusion equation, Press et al. (1971) recommend that one should go no higher than second order in space and time. Thus for multidimensional DEs one must resign oneself to long running times, high memory requirements and a certain loss of accuracy. If this is not practical a lower dimensional model can in some circumstances be used to approximate higher dimensions (Vaughn et al. 1998b).

### 2.3.2 Finite difference schemes used

For the numerical solutions given here, we used the alternating direction implicit (ADI) method in two and three space dimensions (Ames 1992). This is recommended for diffusive problems.
as it is fast, second-order accurate in space and time, unconditionally stable and unlike simpler schemes, allows for examination of the solution at all time-steps (Press et al. 1971, Mascagni 1989, Ames 1992). The equation to be approximated is:

\[
\frac{\partial C}{\partial t} - D \nabla^2 C = P(\vec{x}, t) - S(\vec{x})C - \lambda C \tag{41}
\]

where:

\[ P(\vec{x}, t) = \begin{cases} 
Q \rho & \text{for points inside the source during synthesis} \\
0 & \text{else} 
\end{cases} \tag{42} \]

and:

\[ S(\vec{x}) = \begin{cases} 
\eta & \text{for points inside sinks} \\
0 & \text{else} 
\end{cases} \tag{43} \]

where a sink is a local high concentration of an NO-binding moiety such as a heme-protein. Thus in two dimensions we have (Ames 1992):

\[
\frac{u_{n+1/2} - u_n}{\frac{1}{2} \Delta t} = D \left( \frac{\delta_x^2 u_{n+1/2} + \delta_y^2 u_n}{2} \right) + P(i, j, n) - \left( \frac{\lambda + S(i, j)}{2} \right) \left( u_{n+1/2} + u_n \right) \tag{44}
\]

\[
\frac{u_{n+1} - u_{n+1/2}}{\frac{1}{2} \Delta t} = D \left( \frac{\delta_x^2 u_{n+1/2} + \delta_y^2 u_{n+1}}{2} \right) + P(i, j, n + 1/2) - \left( \frac{\lambda + S(i, j)}{2} \right) \left( u_{n+1} + u_{n+1/2} \right) \tag{45}
\]

Extending these equations to three space variables leads to a method that is unstable for any useful spatial and temporal scales (Ames 1992) and so the following variant is used. Instead of taking three third-steps one generates three subsequent approximations for the solution at the advanced time-step, the third one being used as the actual solution. We obtain the first approximation \( u_{n+1}^* \) at time-step \( n + 1 \) in the following way (Ames 1992):

\[
\frac{u_{n+1}^* - u_n}{\Delta t} = D \left[ \frac{1}{2} \delta_x^2 (u_{n+1}^* + u_n) + \delta_y^2 u_n + \delta_z^2 u_n \right] + P(n) - \frac{\lambda}{2} (u_{n+1}^* + u_n) \tag{46}
\]

where equations 36-38 have been extended in the obvious way so that, for instance:

\[
u_n \equiv u_{i,j,k,n} \equiv u(i \Delta x, j \Delta y, k \Delta z, n \Delta t) \tag{47}\]

\[
\delta_x^2 u_n \equiv \delta_x^2 u_{i,j,k,n} \equiv u_{i,j,k+1,n} - 2u_{i,j,k,n} + u_{i,j,k-1,n} \tag{48}\]

The second approximation \( u_{n+1}^{**} \) is then calculated using the first via:

\[
\frac{u_{n+1}^{**} - u_n}{\Delta t} = D \left[ \frac{1}{2} \delta_x^2 (u_{n+1}^{**} + u_n) + \frac{1}{2} \delta_y^2 (u_{n+1}^{**} + u_n) + \delta_z^2 u_n \right] + P(n) - \frac{\lambda}{2} (u_{n+1}^{**} + u_n) \tag{49}\]

and the final solution \( u_{n+1} \) with:

\[
\frac{u_{n+1} - u_n}{\Delta t} = D \left[ \frac{1}{2} \delta_x^2 (u_{n+1} + u_n) + \delta_y^2 (u_{n+1}^{**} + u_n) + \delta_z^2 (u_{n+1} + u_n) \right] + P(n) - \frac{\lambda}{2} (u_{n+1} + u_n) \tag{50}\]

In the above the reaction term for a sink has been dropped as this was not used with a three-dimensional model, though its inclusion is straightforward.
The two-dimensional ADI equation was implemented with spatial scale of 1\(\mu m\), on a square grid of size 1000 \times 1000 and time step 1ms. The three-dimensional version also used a space-step of 1\(\mu m\), but on a cubic 300 \times 300 \times 300 grid with a time-step of 4ms. The effects of the step-sizes were checked by running the equations with smaller scales and were found to be negligible (< 0.5% relative error). The equations were run using Neumann boundary conditions with the gradient at the edge of the grid set to be constant. However, to ensure that the size of the grid and boundary condition did not affect the results significantly, we checked the simulations by re-running them with a flat gradient at the boundary. The equations were written in C. For full details of the implementation see Philippides (2001).

2.4 Parameter values

The values of the main parameters used here, the diffusion coefficient, \(D\), the decay rates, \(\lambda\) and \(\eta\), and the concentration rate, \(Q\rho\), warrant some discussion. We also discuss the choice of an NO threshold and the localisation of nNOS.

2.4.1 Diffusion coefficient and decay rate

The value of \(D\) in an aqueous salt solution has been measured as 3300\(\mu m^2s^{-1}\) (Malinski et al. 1993). This value has been used widely (Wood and Garthwaite 1994, Lancaster1994, Vaughn et al. 1998a) and has also been derived with reference to a model (Vaughn et al. 1998b). It is reasonable to assume that it will not be significantly affected in a lipid or protein aqueous medium due to the very small molecular dimension and non-polarity of NO. In addition, because NO is dilute, \(D\) is assumed to be independent of NO concentration and constant (Vaughn et al. 1998b) and so we use this value throughout.

The value of the decay rate used gives a half-life of 5s, which is that recorded for dissolved NO perfused over living tissues in oxygenated saline solution (Moncada et al. 1989). Whilst other rate constants can be used, these are basically dependent on the oxidising environment in which NO is diffusing. If this is other than a simple environment, with a \(t_{1/2} \ll 5s\), it should be treated more carefully (Vaughn et al. 1998a, Chen et al. 1998), whilst anything longer has hardly any effect over the spatial and temporal scales examined here (Wood and Garthwaite 1994). We have thus made the simplifying assumption that the background half-life is 5s. For strong NO sinks, \(\eta\), has a value of 693\(s^{-1}\) equivalent to a half-life of 1ms which was chosen as a conservative value based on the rate of NO uptake by a nearby haemoglobin containing structure such as a blood vessel (Carlsen and Comroe 19 58,Liu et al. 1998).

2.4.2 NO production rate

The value of the synthesis rate \(Q\rho\) is a more open question and several values have been determined via different models. Before these are discussed, however, it should be noted that the effect of this parameter is purely one of scale as it is a constant which simply multiplies the concentrations. Thus, whatever the actual value of this parameter, the qualitative nature of the results is unchanged and it is easy to see what effect a different value would have simply by re-scaling.

There are two determinations of \(Q\rho\) that have underpinned NO diffusion modelling to date, both of which are based on the experimental findings of Malinski et al. (1993). Both results are measurements of NO from endothelial cells of a rabbit aorta, one \textit{in vivo} and the other \textit{in vitro}. The first measurement is the concentration of NO produced from stimulated cells of the
aorta 100µm away after diffusion through muscle cells. The second measurement is taken at the surface of a single endothelial cell in culture stimulated to produce NO.

Vaughn et al. (1998b) chose to use the in vivo determination. This is a very complex situation since the reaction with smooth muscle has to be taken into consideration and the size of the synthesising region is unknown and to complete the calculation many simplifying assumptions had to be made. One of these was that NO was produced at the surface of the endothelial cells only which could seriously alter the results and renders the production rate gained unusable in the models we employed.

We, like Wood and Garthwaite (1994), base our model on the in vitro determination. However, unlike Wood and Garthwaite (1994), who used a point-source model to represent a spherical neuron of diameter 1µm, we employed a structure-based analysis. For this task, we used a hollow sphere of inner radius 6, outer radius 10, with the result that a value for Qρ of $1.32 \times 10^{-4} \text{mol} \mu \text{m}^{-3} \text{s}^{-1}$ is needed to generate a maximum concentration of 1µM on the surface of the sphere. These dimensions are chosen to approximate an average endothelial cell but are not incredibly significant (Philippides 2001). Also, the endothelial cell will not be spherical but again, due to the insensitivity of the results to changes of dimension and given that there is no other data available, this approximation was deemed sufficient. Significantly, the resultant value for Qρ is about 300 times less than that used previously. Moreover, the peak concentration is attained after about 14 seconds - a result which agrees closely with the empirical data of Malinski et al. (1993) but which was unexplained when the point-source model was utilised (Lancaster 1997, Philippides 2001). The determination of the production rate is discussed in more detail in Philippides (2001).

### 2.4.3 Distribution of nNOS in neurons

When modelling NO formation one must know where the NOS is located and this differs depending on the type of NOS examined. Here we are mainly concerned with nNOS and so we should say a few words on its distribution within neurons. In the literature nNOS is often referred to as being located on the membrane, as it is frequently shown to be associated with postsynaptic density protein-95 (PSD-95) which in turn is linked to the NMDA-receptor which is membrane-associated (see, for instance Snyder and Ferris (2000)). This is consistent with NO acting as a retrograde messenger to induce LTP or LTD in the pre-synaptic neuron. Thus it is natural that nNOS’s position near the NMDA receptor is emphasised but this should not be taken to mean that there is no nNOS elsewhere. Indeed, nNOS is a soluble enzyme and will be dispersed throughout the cytoplasm, as demonstrated by NADPH-diaphorase staining (O’Shea et al. 1998). A similar story is true of eNOS which, while it does have an affinity for the membrane, will also be found at positions throughout the endothelial cell (DeFilipe 1993). Thus in our model we have made the assumption that nNOS is spread evenly within the synthesising region with a uniform source density which we have denoted as ρ.

### 2.4.4 The NO threshold concentration

To quantify a threshold concentration for effective NO signalling, one first has to specify a particular molecular signalling pathway. Here we follow the thinking of Vaughn et al. (1998a) who chose the soluble guanylyl cyclase-cyclic GMP (sGC-cGMP) signalling pathway, the major signalling pathway for NO in the brain (Schmidt and Walter 1994, Garthwaite and Boulton 1995). The equilibrium dissociation constant (Stone and Marletta 1996) for NO for sGC is 0.25µM and this value defines a threshold concentration for NO.
3 Results

In this section we apply the methods detailed above to investigate the properties of an NO signal produced by neuron-like morphologies. In so doing, we examine a number of salient functional questions that arise in the context of volume signalling. In particular, we highlight the importance of the morphology of the source in determining the spatial and temporal extent of an NO volume signal. While a number of previous models of NO diffusion in the brain have been published, they are broadly one of two types: point-source models (see for instance (Wood and Garthwaite 1994, Lancaster 1994)) or compartmental models (Gally et al. 1990, Lancaster 1996, 1997), neither of which address the impact of the source structure on the diffusional process. The shortcomings of these approaches are discussed in detail in Philippides et al. (2000) and Philippides (2001), but we will summarise the main points here.

In a point-source model, as the name suggests, one models NO diffusing from a source as if it were being produced at a dimensionless point at its centre. It is not difficult to see intuitively that problems might arise from the fact that a point source is therefore by definition singular. The singular nature of the solution represents a fundamental problem to modelling sources with morphology which can be appreciated by examining the steady-state solution for the 3D point source used by Wood and Garthwaite (1994):

\[ C(r) = \frac{S}{r} \exp \left( -\frac{\ln 2}{\sqrt{Dt_{1/2}}} \right) \]

where \( r \) is the distance from the source and \( S \) is a constant determined by the production rate of the source. The first thing to note is that the concentration at the source \( r = 0 \) is infinite. Although the central concentration itself could be ignored, the ramifications of having a singularity at the heart of the solution causes many complications and unrealistic results (Philippides et al. 2000). Firstly one must decide at what distance from the centre the model is deemed to be ‘correct’, necessarily a somewhat arbitrary choice. The approach taken by Wood and Garthwaite (1994) was to use the surface of the neuron and only consider points outside the cell. This highlights a second problem, namely that the internal concentration is indeterminate which in turn means that obtaining a meaningful solution for hollow structures is impossible (Philippides et al. 2000). Finally, as the concentration in equation 51 is dependent on the distance from the source only, this model cannot be used to address the impact of different source morphologies as sources with the same value of \( S \) but different shapes and sizes will yield identical results.

In compartmental models, on the other hand, one can include some notion of the source morphology. However, while such models do give valid insights into the overall role of a diffusing messenger they are a form of explicit finite difference model (Ames 1992) and are thus hampered by the limit on the duration of the time step employed given in equation 33 (Mascagni 1989). This limitation necessitates the use of relatively large compartments leading to gross approximations. In view of this, we believe a more sophisticated form of numerical approximation, such as the one presented here, should be employed when the complexity of the morphology makes an analytical solution impractical.

3.1 Diffusion from a typical neuron

We first examine the solution for a simple symmetrical structure representing, for example, a neuronal cell body in which NO is synthesized in the cytoplasm but not in the nucleus. We
Figure 2: Concentration of NO plotted against time after synthesis for a hollow spherical source of inner radius 50µm and outer radius 100µm for a 100ms burst of synthesis. Here the solid line depicts the concentration at the centre of the cell (0µm), whilst the dotted line shows the concentration at 225µm from the centre. Because the absolute values attained at the two positions differ from one another markedly, the concentration is given as a fraction of the peak concentration attained. These peak values are 7.25µM (centre) and 0.25µM at 225µm. The cell and the points at which the concentration is measured are depicted to the left of the main figure. Note the high central concentration, which persists for a long time (above 1µM for about 2s). Also, there is a significant delay to a rise in concentration at distant points which is more clearly illustrated in the expanded inset figure. The square-wave shown beneath the inset figure represents the strength function.

We have therefore examined the solution for a hollow spherical source of inner radius 50µm (the nucleus) and outer radius 100µm (cell body). These dimensions, though large for many neurons especially in vertebrates, do correspond to the dimensions for some identified giant molluscan neurons whose cell bodies synthesise NO and have been shown to mediate volume signaling (Park et al. 1998).

Of course we are not suggesting that neurons are perfectly spherical but rather that hollow spheres are a useful approximation for neurons. They can for example tell us about the importance of morphological irregularities. For instance, if one had a cell which was mainly spherical but had a lot of small-scale variability in its outer structure, we could use two ideal models, one with the outer radius set to the minimum radius and the other with outer radius set to the maximum. In this way analytical solutions can be employed to see whether or not the irregularity has a significant effect. In fact we have seen that due to the speed of diffusion of NO, small-scale irregularities (±2.5% of source size) have a negligible effect (Philippides 2001). Using such an approach we can also investigate the sensitivity of the diffusional process to other parameters such as boundary conditions whose complexity make the analytical solution intractable. Thus, if we have to make simplifications to a model to render derivation of the analytical solution tractable, we can tell whether or not these simplifications generate gross inaccuracies.
Figure 3: Concentration of NO plotted against distance from the centre of a hollow spherical source of inner radius 50\(\mu\)m and outer radius 100\(\mu\)m for a 100\(ms\) burst of synthesis starting at time \(t = 0\). The graphics underneath each plot depict the structure. A. Concentration of NO at times \(t = 25, 50\) and 100\(ms\), two time points during and one at the end of synthesis. B. Concentration of NO after synthesis at times \(t = 175, 300\) and 1.5\(s\). The reservoir effect following the end of synthesis is clearly seen as the centrally accumulated NO is trapped by the higher surrounding concentrations.

The solution for the hollow sphere was examined for a burst of synthesis of duration 100\(ms\), with results shown in figures 2 and 3. There are two points of note, namely the length of time for which the concentration in the centre of the sphere remains high and the significant delay between the start of synthesis and the rise of concentration for points distant from the source (figure 2). The cause of these phenomena can be seen on examination of figure 3. During the synthesis phase, the concentration outside the cell rises very slowly. In the nucleus, however, a ‘reservoir’ of NO starts to build up (figure 3A), albeit relatively slowly when compared to the rise in the synthesising area (the cytoplasm). After the end of synthesis, this reservoir continues to fill up for about 200\(ms\) as the NO in the cytoplasm diffuses away from its point of origin to points of lower concentration in the nucleus. However, the concentration outside the cell still rises slowly as the NO is dissipated over a larger volume. Later, the situation changes somewhat, as we are now in the position where the concentration in the nucleus is roughly equal to the concentration in the cytoplasm, giving a wide flat peak to the concentration profile. Until this point, the NO which had diffused into the centre had been ‘trapped’ and cannot be dissipated due to the higher concentration present in the surrounding cytoplasm. Now though, we see this reservoir spreading away from the cell in a wave of high concentration which starts to raise the distal concentrations to significant levels. However, the concentration at the centre remains high and does not spread outwards very quickly since the concentration gradient is virtually flat meaning there is very little diffusive pressure on the NO in this area. It is this effect that produces the unexpected time delay at distant points.

Examination of the concentration at 225\(\mu\)m from the centre of the cell (figure 2), shows that it remains low until about 400\(ms\) after synthesis has stopped. It peaks shortly afterwards.
and stays relatively high for a relatively long period. This has implications for the temporal
dynamics of NO-signalling in a neurobiological context. For example, suppose there was an
NO-responsive neuron at a distance of 225 µm from the centre of the source neuron. Assuming
a threshold concentration of 0.1 µM this neuron would not be affected until 600 ms after the end
of synthesis and would continue to be affected for a period of about 4 s. Such a process could be
used to introduce a time delay in NO-mediated neural signalling. The high central concentration
also has implications for neural signalling as the effect of the NO synthesising event remains long
after this event has passed.

There is another interesting factor seen in these results, namely the temporal dynamics of
the solution in the cytoplasm during synthesis. Here it is enough to note that the concentration
continues to rise for a very long time of continuous synthesis before a steady-state is approached.
Thus, though much of the work using point source models has considered solutions at steady-
state such considerations may be inappropriate in the context of real structures.

3.2 Effect of neuron size

![Diagram of threshold distances for NO generated by single fibres of different diameters.](image)

Figure 4: Threshold distances for NO generated by single fibres of different diameters. The threshold distance is defined as the distance at which concentration of NO drops below 0.25 µM, the dissociation constant for soluble guanylyl cyclase (Stone and Marletta 1996).

A. Illustration of threshold distances for several fibres. Numbers inside circles represent fibre diameters in microns. For comparison, the fibres have all been drawn the same size with the threshold distance shown in multiples of the fibre diameter. B. Threshold distance as a function of fibre diameter plotted against fibre diameter. Note the sharp reduction in threshold distance for fibres of diameter less than 10 µm.

The effects reported above are generated by the relatively large dimensions typical of some
molluscan neurons. Questions therefore arise as to whether similar phenomena are present
for smaller neurons such as those found in mammalian brains. In Philippides et al. (2000)
we showed that reducing the radius of the cell can have a significant and somewhat counter-
intuitive effect on the signalling capacities of neurons. In particular, for hollow spherical cells
whose inner radius is half the outer radius and which synthesise NO for 100 ms, the signalling
capacity slowly increases as the cell size is reduced, peaking for 30–40 µm diameter cells. Below
this size however there is a steep almost linear decline in the threshold distance and sources of
10 µm diameter or less have a very limited signalling capacity.

To further examine this effect, we studied the maximum region around a source which could be
affected via the NO-cGMP signalling pathway during and after the generation of NO from
tubular sources of various sizes. The affected region is defined by the volume within which the
concentration is above the threshold concentration, as discussed in section 2.4.4, and is given in
multiples of fibre diameter. This is because the important comparator is the number of potential neuronal targets within the affected region and this depends on their size. A similar phenomenon to that seen in spherical sources is seen for these NO-expressing fibres (figure 4): a slow rise in the affected region as the diameter is reduced, peaking at a fibre diameter of 20µm, followed by a steep decline thereafter indicating that small fibres (4µm diameter or less) are unable to generate an effective NO signal. Moreover, for such small sources, the affected region is not increased significantly by increasing the duration of NO synthesis, since a steady-state situation is rapidly approached. Thus for a source of diameter 3µm or less, a threshold concentration will not be achieved anywhere for no matter how long NO is synthesised (figure 5).

3.3 Small sources

Sources which are too small to generate a threshold concentration individually may affect larger regions if they behave as if they were a single larger source with the attendant temporal and spatial phenomena associated with a source of the combined shape and size. An example of this is provided by endothelial cells which act as a multicellular complex of many very small cells (Vaughn et al. 98a). In many other locations in the brain, though, there are many instances in the brain of well-separated NO sources below the critical size required for a volume signal. It would seem therefore that NO from these sources cannot have a functional role. In determining the range of influence of a source, however, it was assumed that each source is acting on its own. What if instead, NO is derived from many small separated sources acting in concert?

In this section, we study this situation, analysing the dynamics of the NO cloud such sources produce. In particular, we examine networks of axonal fibres with diameters of a few microns or less and the functional extent of the volume signal they generate. A knowledge of the characteristics of such a signal could be crucial in helping to understand NO’s neuromodulatory role, since this type of source is found in many places in the brain. For instance, while NO producing neurons account for only about 1% of cell bodies in the cerebral cortex, their processes
spread so extensively, that almost every neuron in the cortex is exposed to these small fibres the vast majority of which have diameters of a micron or less (Baranano et al. 2001, Defilipe 1993). Another example is found in the locust optic lobe where there are highly ordered sets of nNOS-expressing 2µm diameter fibres (Elphick et al. 1996).

An illustration of how co-operation can occur in a model of fibres in the locust optic lobe is provided by figure 6 which shows the spatial extent of the NO signal generated by a single fibre of 2µm diameter and by ordered arrays of four, nine and sixteen identical sources separated by 10µm. As can be seen in the figure, an ordered array of N^2 fibres is a 2D arrangement of N x N fibres. As the fibres are parallel the solution is symmetric along the z-axis (out of the page) and so we give results only for cross-sectional slices in a plane perpendicular to the direction of the fibres. The single fibre does not achieve an above threshold signal principally because the great speed of NO diffusion means that NO will spread rapidly over a large volume. So while NO does not reach threshold anywhere, the volume occupied by NO at a significant fraction of threshold is large relative to the source size. Thus NO derived from small and well-separated individual sources can summate to produce an effective NO cloud. But what are the characteristics of such a signal and what do they imply for the way NO functions? For instance, do we still see the reservoir and delay effects characteristic of signals from single larger sources?

The first thing that one notices from figure 6 is that while the centre effect is still present with NO accumulating in the centre of mass of the sources, the concentration profile appears to be flatter. This can be seen more clearly in figure 7 where an ordered 10 x 10 array of 2µm fibres separated by 36µm. The effect of the spacing between the sources is further illustrated in figure 8, where the area over threshold generated by 100 fibres is shown as a function of the total cross-sectional area of source. What is immediately obvious from figure 8 is that if one’s goal is to reach the largest number of potential targets with a given volume of source fibres, then one should use a dispersed source rather than one single source. While the optimal spacing that should be used is dependent on the length of synthesis (and the number of sources (Philippides 2001), results for fibres arranged contiguously so that 100 fibres act as one 20 x 20µm source (the first points on the x-axis in figures 8A-B), are always lower than those for dispersed sources (unless the sources are dispersed so widely that they fail to interact). Indeed can affect a volume over twice the size of a solid source simply by dispersing them correctly. Thus, in terms of the
extent of the NO signal, there is a big advantage in using separated sources.

What about the temporal dynamics of the NO signal? Examining the time-course of the NO signal generated by an array of 100 fibres spaced 36µm apart, we see a delay until areas reach an above threshold concentration as we did for single sources (figure 7). This time, however, rather than the delay being for points outside the source only, here there is a delay until any point is affected by NO, after which there is a very steep rise in the volume affected. This is a common feature of signalling from dispersed sources because the summation of NO from several separated fibres means that the concentration in and around them is, in a sense, averaged and hence, smoothed. Thus due to the dynamics of diffusion one tends to get a relatively even concentration within the synthesising region with small peaks around the fibres themselves (figure 7). In conjunction with the use of a threshold concentration, this means that there will come a point when the concentration in a region around the fibres is just sub-threshold and a small increase in the general level of NO will result in large areas rising above threshold, as shown in figure 7. We refer to this feature as the interaction effect.

Again, the impact of the interaction effect will vary depending on the spacing used and a large range of temporal dynamics can be seen (figure 8B). In particular, the delay before the start of interaction can vary from nothing to more than a second, with the delay growing as the spacing is increased. This means that a system with optimal spacing, in terms of extent of the affected region, will experience a considerable delay before the region begins to be affected, with the total area affected suddenly rising sharply at the end of the delay. This raises the intriguing functional possibility of a system which is completely unaffected by NO for a given length of
time (a period which would be tunable by changing the spacing), but once past this point, large regions are rapidly ‘turned-on’ by the cloud of NO\textsuperscript{3}.

Another factor which affects the length of the delay is the thickness of the fibres used. Examining the delay from plexuses composed entirely of fibres of various fixed diameters, we see that the delay is longest for the thinnest fibres (figure 9). This is because the fibres are now too small to achieve an above threshold signal singly and so must cooperate, although the subsequent rise in the affected area is not as steep as for ordered arrays. This is expected since the random nature of the plexus means that the distribution of concentrations is less uniform and the interaction effect is less pronounced. Examining the delay for plexuses of other diameters we see that it rises steeply to that seen for the 1\,\mu m plexuses showing that the interaction needed is much greater for the smaller fibres (figure 9A). As a property of a signal, the obvious role for such a delay is as a low pass filter since there has to be nearly 200\,ms of synthesis before the thin plexus will respond. In the case of the signal mediating increased blood flow, this means that there would need to be significant sustained activity before blood flow increased, whereas for a thick plexus blood flow would react to every short burst of neuronal activity. Other features seen to vary with fibre thickness are maximum concentration, how centred a cloud is on a target region, and the variability of concentrations over a region; all of which have sensible interpretations in terms of neuronal signalling to blood vessels.

4 Exploring functional roles with more abstract models

The computational models presented above require huge numbers of iterated calculations and inevitably place heavy demands on processing resources. Hence it is not yet feasible to build models of whole neuronal networks at that level of detail and run them in real time, or anything even vaguely approaching it. Therefore, in parallel with the detailed modelling work, we have

\textsuperscript{23}A form of signalling which might also be useful in an artificial neural network (see section 4).
developed a class of computationally fast artificial neural networks (ANNs) that incorporate a more abstract model of signalling by diffusing neuromodulators. Such networks have been used as artificial nervous systems in autonomous mobile robots and have allowed us to start exploring the properties and potential functional roles of this kind of signalling in the generation of behaviour (Husbands et al. 1998, Husbands et al. 2001). We have named this class of ANNs GasNets. Our work with these networks is briefly introduced in this section.

Since as yet we have no deep formal theory for such systems, we have found the use of stochastic search methods (such as evolutionary algorithms) to be a very helpful tool in this exploration. We have used the methods of evolutionary robotics to explore the suitability of this class of networks for generating a range of behaviours in a variety of autonomous robots.

Figure 9: Mean results for plexuses composed entirely of fibres of diameters $d = 1, 2, 3, 4$ or $5 \mu m$ in a $100 \times 100 \times 100 \mu m^3$ synthesising volume for 1 second of NO synthesis. Results averaged over 30 random plexuses grown generated using the growth algorithm detailed in Philippides (2001). A. Mean delay until interaction against diameter of plexus fibres. B. Mean volumes over threshold both in total and inside the synthesising volume against diameter of plexus fibres. C. Mean maximum concentration obtained against diameter of plexus fibres. D. Mean concentrations seen in a 1d line through the centre of either thin ($1 \mu m$) or thick ($5 \mu m$) plexuses.
4.1 The GasNet model

In this strand of our work we have attempted to incorporate into ANNs, in an abstracted form, some of the richness and complexity that characterises the temporal and spatial dynamics of real neuronal signalling, especially chemical signalling by gaseous transmitters. As these systems operate on different temporal and spatial scales to electrical signalling, we have developed models in which electrical and indirect chemical signalling are controlled in ANNs by separate processes. Thus, we developed the eGasNeti, a standard ANN augmented by a diffusing virtual gas which can modulate the response of other neurons.

The ‘electrical’ network underlying the GasNet model is a discrete time step, recurrent neural network with a variable number of nodes. These nodes are connected by either excitatory (with a weight of +1) or inhibitory (with a weight of -1) links with the output, $O^n_i$, of node $i$ at time step $n$ described by the following equation:

$$O^n_i = \tanh \left[ k^n_i \left( \sum_{j \in C_i} w_{ji} O_{j}^{n-1} + I^n_i \right) + b_i \right]$$ (52)

where $C_i$ is the set of nodes with connections to node $i$ and $w_{ji} = \pm 1$ is a connection weight. $I^n_i$ is the external (sensory) input to node $i$ at time $n$, and $b_i$ is a genetically set bias. Each node has a genetically set default transfer function parameter, $k^0_i$, which can be altered at each time-step according to the concentration of the diffusing ‘gas’ at node $i$ to give $k^n_i$ (as described later in the section on modulation).

4.2 Gas diffusion in the networks

In addition to this underlying network in which positive and negative ‘signals’ flow between units, an abstract process loosely analogous to the diffusion of gaseous modulators is at play. Some units can emit virtual ‘gases’ which diffuse and are capable of modulating the behaviour of other units. The networks occupy a 2D space; the diffusion processes mean that the relative positioning of nodes is crucial to the functioning of the network. The original GasNet diffusion model is controlled by two genetically specified parameters, namely the radius of influence $r$ and the rate of build up and decay $s$. Spatially, the gas concentration varies as an inverse exponential of the distance from the emitting node with a spread governed by $r$, with the concentration set to zero for all distances greater than $r$ (equation 53). The maximum concentration at the emitting node is 1.0 and the concentration builds up and decays from this value linearly as defined by equations (54 and 55) at a rate determined by $s$.

$$C(d, t) = \begin{cases} e^{-2d/r} \times T(t) & d < r \\ 0 & \text{else} \end{cases}$$ (53)

$$T(t) = \begin{cases} H \left( \frac{t - t_e}{s} \right) & \text{emitting} \\ H \left[ H \left( \frac{t - t_e}{s} \right) - H \left( \frac{t - t_s}{s} \right) \right] & \text{not emitting} \end{cases}$$ (54)

$$H(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & \text{else} \end{cases}$$ (55)

(Husbands et al. 1998, Husband et al. 2001).
where $C(d,t)$ is the concentration at a distance $d$ from the emitting node at time $t$. $t_e$ is the time at which emission was last turned on, $t_s$ is the time at which emission was last turned off, and $s$ (controlling the slope of the function $T$) is genetically determined for each node. The total concentration at a node is then determined by summing the contributions from all other emitting nodes (nodes are not affected by their own concentration, to avoid runaway positive feedback).

A variant on this diffusion model, based on the cortical plexus diffusion described earlier in this chapter, involves diffusion of a uniform concentration ‘cloud’ centred on some genetically specified site distant from the emitting node. This reflects the spatial separation of the main plexus and the body of the controlling neurons. The cloud suddenly turns ‘on’ or ‘off’, depending on the state of the controlling neuron, in keeping with the plexus mode of signalling described earlier.

4.3 Modulation

In a typical GasNet model (Philippides et al. 2002), each node in the network can have one of three discrete quantities (zero, medium, maximum) of N possible receptors. Each diffusing-neurotransmitter/receptor pairing gives rise to a separate modulation to the properties of the node. The strength of a modulation at node $i$ at time $n$, $\Delta M^n_{ij}$, is proportional to the product of the gas concentration at the node, $C^n_i$, and the relevant receptor quantity, $R_j$, as described by equation 56. Each modulation makes some change to one or more function parameters of the node. All the variables controlling the process are again set for each node by an evolutionary search algorithm.

$$\Delta M^n_{ij} = \rho_i C^n_i R_j$$  \hspace{1cm} (56)

A number of different receptor linked modulations have been experimented with, including:

- Action of receptor 1: increase gain of node transfer function
- Action of receptor 2: decrease gain of node transfer function
- Action of receptor 3: increase proportion of retained node activation from last time step
- Action of receptor 4: if above a threshold switch transfer function of node for sustained period

Most GasNet variants were found to be highly evolvable and capable of the robust generation of sensorimotor behaviours (often visually guided) in noisy environments (Husbands et al. 1998, Husbands et al. 2001). Indeed, it has been shown that most forms of GasNet are significantly more evolvable than all other forms of ANN tested over a wide range of continuous sensorimotor tasks. The operation of the networks was often subtle, making use of intricate dynamics involving continuously shifting modulation patterns (Smith et al. 2003). Particular functional modules such as oscillators and low-pass ‘noise’ filters were frequently discovered and used by the evolutionary process. The inherently flexible, and generally loose, coupling between two processes with distinct spatial and temporal properties (the chemical and the electrical) makes these systems highly evolvable and endowed with a powerful kind of plasticity (Philippides et al. 2002). Future work will increase the biological veracity of the networks while maintaining their abstract computationally tractable nature. They can then be used to illuminate more specific biological questions than they have to date.
5 Conclusions

This chapter has concentrated on the details of how to build computational models of NO diffusion in the nervous system and has demonstrated how such models can give important insights into the phenomenon of volume signalling. This kind of signalling cannot be dealt with in a simple point-to-point connectionist framework; our tools and concepts for understanding the operation of neuronal circuitry making use of such signalling must be expanded. The work presented here is intended as a contribution to that expansion.

References


J. Crank (1980). The mathematics of diffusion. OUP.


28


33: 1235-1244.