Forcing Function Diagnostics for Nonlinear Dynamics

Giles Hooker

Abstract

This paper investigates the problem of model diagnostics for systems described by nonlinear differential equations. We study lack of fit through the estimation of forcing functions – external inputs that provide the desired behavior. We derive lack of fit tests based on these functions and study the problems associated with diagnostics for partially observed systems. Some observations are made concerning model building that postulates further unobserved variables. The methods are illustrated with examples from computational neuroscience.

1 Introduction

Recent research has seen a significant increase in interest in fitting nonlinear differential equations to data. Many systems of differential equations used to describe real-world phenomena are developed from first-principles approaches; either through conservation laws or by rough guesses. Such *a priori* modeling has lead to proposed models that mimic the qualitative behavior of observed systems, but have poor quantitative agreement with empirical measurements. There has been relatively little literature on the development of interpretable differential equations models from an empirical point of view. The need for good diagnostics has been cited a number of times, for example in Ramsay et al. (2007).

This paper considers the problem of performing goodness of fit diagnostics and model improvement when a set of deterministic nonlinear differential equations has been posited to explain observed data. We describe the state of a system by a vector of k quantities $\mathbf{x} = \{x_1, \dots, x_k\}$. An ordinary differential equation (ODE) describes the evolution of $\mathbf{x}(t)$ by relating the *rate of change* of $\mathbf{x}$ to the current state of the system

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)|\boldsymbol{\theta}). \tag{1}$$

Here $\mathbf{f} : \mathbb{R}^k \to \mathbb{R}^k$ is a posited vector valued function that may depend on external inputs $\mathbf{u}$ and a finite set of unknown parameters $\boldsymbol{\theta}$.

Such systems have been used in many areas of applied mathematics and science. It has been shown that even quite simple systems can produce highly complex behavior. They also represent a natural way of developing systems from first-principles or conservation laws and frequently represent the appropriate scale on which to describe how a system responds to an external input.

The central observation in this paper is that since differential equations are modeled on the derivative scale, lack of fit for these equations must be measured on the same scale. That is, suppose an equation is postulated as in (1) and we observe a process $\mathbf{y}(t)$, then the appropriate measure of lack of fit is

$$\mathbf{g}(t) : \frac{d}{dt}\mathbf{y}(t) - \mathbf{f}(\mathbf{y}(t), \mathbf{u}(t)|\boldsymbol{\theta})$$

$\mathbf{y}$ is used to signify an estimated process as opposed to a solution to (1). This convention will be maintained throughout. We can alternatively be re-express the observed system as:

$$\frac{d}{dt}\mathbf{y}(t) = \mathbf{f}(\mathbf{y}(t), \mathbf{u}(t)|\boldsymbol{\theta}) + \mathbf{g}(t). \tag{2}$$

$\mathbf{g}(t)$, representing lack of fit, appears as a collection of what may be termed *forcing functions*. These act as an external influence on $\mathbf{y}(t)$ which responds on the derivative scale. The basic task of diagnostics now becomes the estimation of $\mathbf{g}(t)$. This may be complicated by knowledge (or lack of knowledge) about unmeasured components of $\mathbf{y}(t)$ and where fit may be lacking. Once $\mathbf{g}(t)$ has been estimated, the usual diagnostic procedures may be undertaken. The statistical significance of $\mathbf{g}(t)$ may be estimated and it may then be examined graphically for relationships with various measured variables. Any apparent relationships may then be accounted for by modifications to the model structure and the modified system re-estimated. We note that $\mathbf{g}$ may also represent a stochastic component of the system. The tests proposed here encompass stochasticity as part of the alternative. Testing for purely stochastic lack of fit, however, is beyond the scope of this paper.

This paper begins with a discussion of estimating lack of fit, illustrated with examples, in Section 2. We develop computationally inexpensive approximate goodness-of-fit tests to be used as initial diagnostics in Section 3 and go on to discuss visual tools for model development in Section 4. Some identifiability issues and possible solutions to them are discussed in Section 5. We provide an example in the form of model development for the spiking behavior of zebrafish neurons in Section 6. Some comments are made concerning more general choices of model changes in Section 7 and we conclude and point out further directions for research in section 8.

2 Estimating Lack of Fit

2.1 A Development Example: The FitzHugh-Nagumo Equations

We use the FitzHugh-Nagumo equations (FitzHugh (1961) and Nagumo et al. (1962)) as a demonstration of model-building. These equations were originally

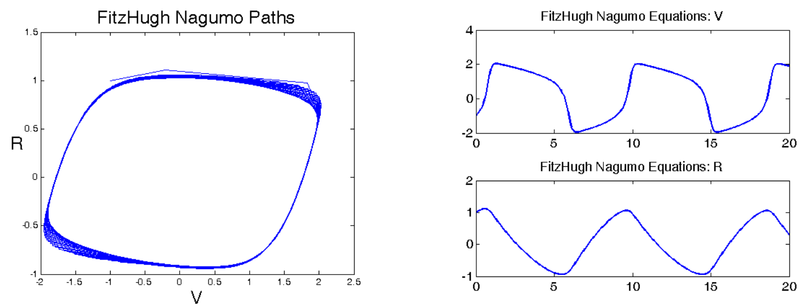

Figure 1: Solutions to the FitzHugh-Nagumo equations for $(a, b, c) = (0.2, 0.2, 3)$ with initial conditions $(V(0), R(0)) = (-1, 1)$. Left: a phase-plane plot of V and R. Right: V and R plotted against time.

developed as an approximation to a system describing the firing behavior of neurons in squid in Hodgkin and Huxley (1952). They have since been employed to describe many neuronal systems as well as to cardiac rhythms (Aliev and Panfilov (1996)). These equations are given by

$$\frac{d}{dt}V = c\left(V - \frac{V^3}{3} + R\right)$$
$$\frac{d}{dt}R = -\frac{1}{c}\left(V - a + bR\right)$$

V is intended to represent a potential across the axon membrane. R is usually interpreted as a recovery variable. Equations of this type are commonly used to provide qualitatively similar dynamics to more complex models based on neuron physiology; see Wilson (1999) for an overview.

Figure 1 plots solutions to these equations for a particular choice of parameters. These solutions are cyclic and we observe sharp changes in the derivative of V, due to the nonlinearity in its equation. Such behavior is difficult to mimic with nonparametric smoothers. In general, differential equations are capable of reproducing a wide range of dynamic behavior, making them a powerful set of mathematical models.

We seek to recover this equation – or a close approximation to it – given data from such a system and partial knowledge of the system itself.

2.2 Representing Latent Forces

The central technique in the paper is the estimation of a set of forcing functions $\mathbf{g}(t)$ as in (2). $\mathbf{y}(t)$ and its derivatives are not known exactly; rather we have access to a set of noisy observations $\mathbf{y}_{obs}(t)$ taken at a set of discrete times $\mathbf{t} = \{t_i\}_{i=1}^n$. $\mathbf{g}(t)$ is represented as a basis function expansion:

$$\mathbf{g}(t) = \Phi(t)\mathbf{d}. \tag{3}$$

$\Phi(t)$ is a vector containing the evaluation of a set of basis expansions $\{\phi_1(t), \ldots, \phi_k(t)\}$ at time t. In the examples below, the ϕ_i are given by cubic B-splines with equi-spaced knots. $\mathbf{d}$ is then a $k \times m$ matrix providing the coefficients of each of k basis functions for each of m forcing components. $\mathbf{d}$ may be estimated via any parameter estimation scheme for differential equations. However, the estimation of derivatives tends to increase the amount of noise in a system and it may be necessary to include some smoothing criterion to regularize the estimated $\mathbf{d}$. Below we give a brief overview of the many parameter estimation strategies available and the smoothing criterion used in this paper.

2.3 Estimating Parameters in a Differential Equation

Parameter estimation techniques in differential equation models is a difficult topic with a large literature associated with it. Solutions to (1) typically do not have analytic expressions, necessitating the use of sometimes computationally-intensive numerical approximation methods. Further, such solutions are given only up to a set of initial conditions $\mathbf{x}(t_0)$. When these are unknown, the set of parameters must be augmented to include them.

This paper largely aims to be generic in its approach and the diagnostic prescriptions here may be tied to any parameter estimation scheme. For the sake of clarity, we use a naive implementation of nonlinear least squares (NLS). We note that almost all the parameter estimation methods that we have found in the literature have tried to minimize a squared error criterion. Under this scheme, solutions $\mathbf{x}(t|\theta, \mathbf{x}_0)$ to (1) are estimated for any estimated parameter vector θ and initial conditions $\mathbf{x}_0$. A Gauss-Newton method is then used to estimate both sets of parameters using gradients calculated by solving the sensitivity equations:

$$\frac{d}{dt}\frac{d\mathbf{x}}{d\theta} = \frac{\partial \mathbf{f}}{\partial \theta} + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{d\mathbf{x}}{d\theta} \tag{4}$$

with initial conditions

$$\left.\frac{d\mathbf{x}}{d\theta}\right|_{t=0} = 0.$$

The equivalent equation for $\mathbf{x}_0$ is

$$\frac{d}{dt}\frac{d\mathbf{x}}{d\mathbf{x}_0} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{d\mathbf{x}}{d\mathbf{x}_0} \tag{5}$$

with initial conditions

$$\left.\frac{dx_i}{dx_{0,j}}\right|_{t=0} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otheriwise.} \end{cases}$$

These equations may be solved by augmenting (1). There are many methods for estimating parameters in ODEs and their accuracy can depend on the nature of the system to be estimated; see Deuflhard and Bornemann (2000). The simulations below have been undertaken using a a Runge-Kutta method implemented in Matlab; however, these were insufficient for the real-data example in Section 6. The standard non-linear least squares analysis (e.g. Bates and Watts 1988) may now be applied. When the θ are replaced by $\mathbf{d}$ in (3), we have the equivalent sensitivity equation:

$$\frac{d}{dt}\frac{d\mathbf{y}}{d\mathbf{d}} = \frac{d\mathbf{f}}{d\mathbf{x}}\frac{d\mathbf{y}}{d\mathbf{d}} + \Phi(t). \tag{6}$$

For the purposes of visualization, it is frequently useful to smooth the estimate of $\mathbf{g}(t)$. When a standard roughness penalty is placed on $\mathbf{g}$, the penalized nonlinear least squares problem becomes

$$\|\mathbf{y}_{obs}(\mathbf{t}) - \mathbf{y}(\mathbf{t})\|^2 + \lambda \sum_{i=1}^{k} \int g_i^{(m)}(t)^2 dt. \tag{7}$$

An easy implementation within standard nonlinear least squares is to replace the integrals in (7) with a quadrature rule, which then appears as extra squared error terms. This approach was taken in Ramsay et al. (2007).

This implementation ignores the many numerical problems associated with parameter estimation in differential equations. These include controlling for numerical error in the Runge-Kutta schemes and the problem of minimizing over rough response surfaces. The author's experience is that while NLS provides good answers when initial parameter estimates are quite close to the true value, it frequently breaks down: finding either a local optimum or a set of parameters for which the ODE system is not numerically solvable.

There are several methods available that may be used to overcome these problems. Stochastic search techniques such as simulated annealing (Jaeger et al. (2004)) or Markov Chain Monte Carlo methods (Huang et al. (2006)) have been employed to try to find global minima in the response surface. These may require significant increases in computation and it may remain unclear whether a global optimum has, in fact, been reached. Other methods include the estimation of parameters while attempting to solve (1), leading to a constrained optimization problem; see, for example, Tjoa and Biegler (1991). This last approach has been used in Section 6.

An alternative approach, closer to the diagnostic methods explored here, involves estimating a nonparametric smooth $\hat{\mathbf{y}}$ for the data. Values for the derivative $\frac{d}{dt}\hat{\mathbf{y}}$ may then be estimated and parameters chosen to minimize

$$\sum_{i=1}^{k} \int \left(\frac{d}{dt}\hat{y}_j(t) - f_j(\hat{\mathbf{y}}(t), \mathbf{u}(t)|\boldsymbol{\theta}) \right)^2 dt \tag{8}$$

This approach was suggested, for example, in Varah (1982). This method, although very easy to implement, requires all components of $\mathbf{y}$ to be measured with high resolution and precision and may still suffer from bias due to the smoothing method used. For the problem of estimating $\mathbf{g}(t)$, these methods are particularly problematic, as demonstrated below. More recent techniques include intermediate methods such as Ramsay et al. (2007), which seeks to trade-off fitting on the scale of derivatives and on the scale of the observations, and Ionides et al. (2006), based on Kalman filtering techniques.

2.4 Recovering a Forced Linear System

As an initial experiment, we use a two-component linear differential equation:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -4 & 8 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} g(t) \\ 0 \end{bmatrix}$$

in which $g(t)$ takes the form of a stepped interval:

$$g(t) = \begin{cases} 5 & \text{if } t < 5 \\ 0 & \text{otherwise} \end{cases}$$

We assume the equation is known exactly and attempt to reproduce g non-parametrically from noisy data. The system was run with initial conditions $(x, y) = (-1, 1)$, observations were taken every 0.05 time units on the interval [0,20] and gaussian noise with variance 0.25 was added to the observed system.

Figure 2 provides the result of estimating $g(t)$ using the methods described above. In order to regularize the resulting smooth, we used a third derivative roughness penalty: $\int g^{(3)}(t)^2 dt$. A smoothness parameter was chosen by minimizing the squared discrepancy between the estimated differential equations at the observation times and the true (noiseless) values of the system.

These results have been compared to the simple expedient of generating a smooth of the data $\hat{\mathbf{y}}(t)$ and measuring

$$\frac{d}{dt}\hat{\mathbf{y}}(t) - \mathbf{f}(\hat{\mathbf{y}}(t), \mathbf{u}(t)|\boldsymbol{\theta})$$

In this case, the smooth was generated by penalizing the third derivative of $\hat{\mathbf{y}}$, following the recommendations of Ramsay and Silverman (2005) and the smoothing parameter chosen by minimizing squared distance between the estimated and true values of the system at the observation times. The second method provides poor results due to the difficulty of estimating a derivatives accurately for a smooth. This difficulty increases for systems exhibiting non-linear dynamics where the solutions to differential equations can exhibit high local curvature. We have made generic choices of basis expansion and roughness penalty. The estimated forcing function could be improved significantly by the addition of multiple knots at $t = 7$ and $t = 14$ to allow the function to be discontinuous at these points. Similarly, prior knowledge of the structure of functional mis-specification, such as cyclic behavior, can be incorporated explicitly into alternative forcing penalties.

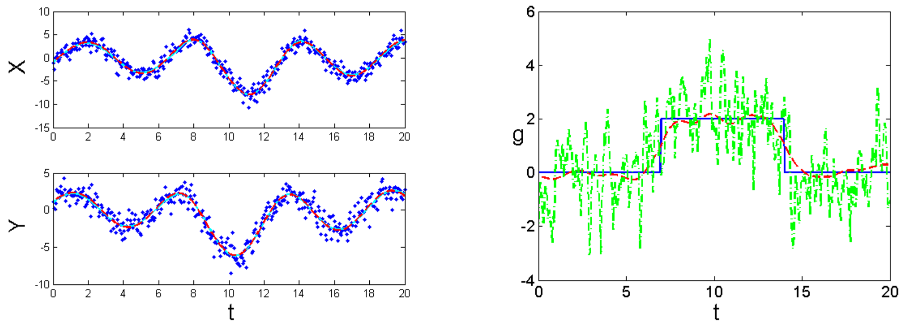

Figure 2: Experiments with a forced, linear system. The right hand plot shows data from a two-component system being forced by a step function on the region [7 14]. The left hand plot compares the true forcing component (dashed) with the estimated component (solid) and a forcing component derived from a generic smooth of the data (dotted).

3 Goodness of Fit Tests

When $\mathbf{f}(\mathbf{x}, \mathbf{u}) = A\mathbf{x} + \mathbf{u}(t)$ is linear up to an additive function, exact solutions to (1) may be given by:

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 + e^{At}\int_{t_0}^{s} e^{-As}\mathbf{u}(s)ds \tag{9}$$

where e^{At} represents a matrix exponential. Including a lack of fit forcing function of the form $\Phi(t)\mathbf{d}$ now modifies (9) to

$$\begin{aligned}\mathbf{y}(t) &= e^{At}\mathbf{x}_0 + e^{At}\int_{t_0}^{s} e^{-As}\mathbf{u}(s)ds + e^{At}\int_{t_0}^{s} e^{-As}\Phi(s)\mathbf{d}ds \\ &= \tilde{\mathbf{u}}(t) + X(t)\mathbf{x}_0 + \tilde{\Phi}(t)\mathbf{d}.\end{aligned} \tag{10}$$

Thus for an additive observational error process, we may express the estimation problem for $\mathbf{d}$ in terms of a linear model with parametric effects for initial conditions $\mathbf{x}_0$:

$$\mathbf{y} = \tilde{\mathbf{u}}(\mathbf{t}) + X(\mathbf{t})\mathbf{x}_0 + \tilde{\Phi}(\mathbf{t})\mathbf{d} = \tilde{\mathbf{u}}(\mathbf{t}) + Z(\mathbf{t})\tilde{\mathbf{d}}.$$

Where $Z(\mathbf{t})$ and $\tilde{\mathbf{d}}$ represent the concatinations $[X(t), \Phi(t)]$ and $[\mathbf{x}_0, \mathbf{d}]$ respectively.

Since we expect the size of $\mathbf{d}$ to grow with $\mathbf{t}$, it is natural to use a penalized estimate for $\mathbf{d}$, providing the familiar penalized regression:

$$\hat{\mathbf{d}} = (Z(\mathbf{t})'Z(\mathbf{t}) + \lambda D)^{-1}Z(\mathbf{t})'(\mathbf{y} - \tilde{\mathbf{u}}(\mathbf{t}))$$

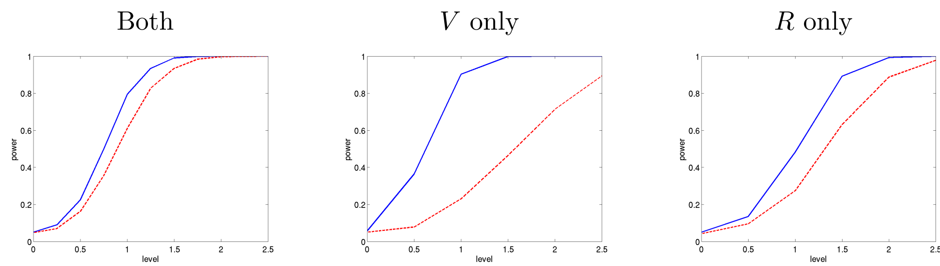

Figure 3: Power comparisons for testing lack of fit in a linear differential equation for lack of fit represented on the derivative scale (solid lines) and as forcing functions (dashed) as a forcing perturbation is varied (horizontal axes). From left to right; when both components are measured; when only the perturbed component is measured; when only the unperturbed component is measured.

where

$$D = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma \end{bmatrix}$$

corresponds to placing a ridge-type penalty on $\mathbf{d}$. Common choices of Σ would include roughness penalties of the form $\Sigma = \int \Phi^{(m)}(t)^T \Phi^{(m)}(t) dt$; or simply $\Sigma = I$. In practice, we have found that the numerical challenges involved in solving the sensitivity equations (6) restricts us to a sufficiently small basis that these choices make little difference.

We note that (10) represents a form that is amenable to a random-effects treatment of smoothing by regarding $\mathbf{d}$ as random effects, λ as ratio of random effect to error variance and testing the hypothesis $\lambda = 0$. This process was described in Crainiceanu et al. (2005) who also derived exact, finite-sample distributions for the test statistic under the assumption that observational errors are independently normally distributed. Similar ideas were developed in Eubank and Spiegelman (1990).

A simulation study is presented in Figure 3. We used the same linear system as presented in Section 2.4. The height of the perturbation was varied between 0 and 2.5 and we simulated Gaussian errors with variance 0.5. Goodness of fit was assessed using a 2nd-order B-spline basis on 21 knots across the interval [0 20]. We compared the power of the test when the basis was taken as representing a forcing function as in (2) against using the same basis to represent lack of fit on the residual scale. Power was calculated using 1000 simulations in each case. It is apparent that the former does noticeably better.

One common aspect of nonlinear dynamical systems is that only some components are measured. Moreover the measured components may not be where system mis-specification is most severe. In order to explore these issues the second and third plots in Figure 3 explore the power of these tests when only one component is observed. In these situations, lack of fit on the observational scale may only be represented for the observed components where it may have quite complicated effects. However, representing lack of fit as a forcing function allows it to be estimated for all components where it frequently has a somewhat simpler form. Moreover, it is possible to explore lack of fit on different components in order to estimate where it may be most severe.

Tests with Unknown Parameters

Typically, some or all of the parameters in a differential equation need to be estimated from data before a goodness of fit test may be applied. Moreover, in nonlinear dynamics, the coefficients of a basis representation for a forcing function do not enter the model linearly at the scale of the observations, even when the parameters are fixed.

Crainiceanu and Ruppert (2004) suggests extensions of mixed-effects likelihood ratio tests to nonlinear models by linearizing the model about the estimated parameter values. In that paper, a nonlinear model was assumed to be of the form:

$$Y = g(t, \theta) + \Phi(t)c,$$

placing lack of fit on the scale of the observations. The test proceeds by linearizing g about the estimated $\hat{\theta}$, providing the approximation

$$Y \approx g(t, \hat{\theta}) + \delta \left. \frac{d}{d\theta} g(t, \theta) \right|_{\hat{\theta}} + \Phi(t)c \tag{11}$$

where δ and θ are then fit to $Y - g(t, \hat{\theta})$ via a standard linear mixed effects model and the test procedure above may again be used. The notion is that the variance associated with estimating $\hat{\theta}$ is partially accounted for by including the fixed effects δ in the mixed model estimation scheme.

The procedure may be carried out by estimating $d\mathbf{y}/d\theta$ via the sensitivity equations (4) and $d\mathbf{y}/d\mathbf{y}_0$ from (5). Setting $\tilde{\theta}$ to be the full vector $(y(0), \theta)$, the approximation (11) becomes

$$\mathbf{y}(\mathbf{t}) - \mathbf{y}(\mathbf{t}|\hat{\mathbf{y}}_0, \hat{\theta}) \approx \delta_{\mathbf{y}_0} \frac{d\mathbf{y}}{d\mathbf{y}_0} + \delta_\theta \frac{d\mathbf{y}}{d\theta} + \Phi(t)\mathbf{d}$$

which fits neatly into the framework provided.

However, this goodness of fit test models lack-of-fit on the scale of the observations. As already noted, this ignores the scale of the model, reducing the interpretability of such lack of fit measures. Moreover, it removes our ability to explore potential forms for such lack of fit, as detailed in Section 5. If, instead, the model is given in the form (2) a further linearization is required. It is clear, even from the form of (10) that, in general, the random effects which we would like to model interact with the model parameters. Under these conditions a generalization of the approximation (11) is given by

$$y = g(t, \theta, c) \approx g(t, \hat{\theta}, 0) + \delta \frac{dg}{d\theta} + c \frac{dg}{dc} \tag{12}$$

For linear differential equations, $\frac{dg}{dc} = \frac{d\mathbf{y}}{d\mathbf{d}}$ is given by the last term of (10). In nonlinear models, it may be estimated by solving the sensitivity equations (6).

Linearization distorts the null distribution of the test statistic and, depending on the accuracy of the approximation (11) this test should be treated with some care. (Crainiceanu and Ruppert 2004) recommend plotting the linearized fit against the fit using the estimated $\hat{\theta}$ as a diagnostic for poor test performance. However, note that variance estimates associated with nonlinear least squares estimates are quite similar in form. Bates and Watts (1988) suggests estimating the variance of $\tilde{\theta}$ by $\sigma^2 \frac{d\mathbf{y}}{d\tilde{\theta}}' \frac{d\mathbf{y}}{d\tilde{\theta}}$. This could then be used in a test for a ratio of quadratic forms in, for example, Hart (1997). $d\mathbf{y}/d\tilde{\theta}$ are exactly the terms used in (12). Performing the full NLS estimate has the advantage of obtaining better estimates of $g(t)$, but requires significantly more computation and some experimentation with λ. Thus, while we have used the NLS procedure to produce the diagnostic tools detailed in the rest of the paper, the methods above provide goodness of fit tests that may be performed cheaply before undertaking a large optimization problem.

4 Mis-specification

There are many ways in which mis-specification may arise in a dynamical system. Wood and Thomas (1999) demonstrates that even minor changes in the form of a dynamical system may have a significant impact on system behavior. It can be important, therefore, to pay careful attention to the form in which a system is written down.

We have demonstrated that a forcing function is a point-wise measure of lack of fit. As such it represents the same quantity as a residual in standard regression for the purposes of diagnostics and model building. We can therefore perform the usual diagnostics; producing plots of $\mathbf{g}(t)$ against $\mathbf{y}(t)$ or against other quantities. Doing so is demonstrated in Section 4.1.

It is important to note, however, that such diagnostic plots may quickly become complex and difficult to interpret. Where $\mathbf{y}(t)$ and $\mathbf{g}(t)$ both contain numerous components each $g_i(t)$ must be examined with respect to all the components of $y_j(t)$. In this case the data available for $g_i(\mathbf{y})$ becomes a univariate path through $\mathbb{R}^k$ making it difficult to determine an appropriate form for this dependence. Figures 5 and 7 demonstrate these difficulties.

Forcing function representations of lack of fit provide a basic recipe for model building that is very similar to the standard techniques taught for data analysis via linear regression:

1. Postulate and estimate a model. Linear models are a common default.

2. Perform a test – either formal or informal – for lack of fit.

3. If fit to data is unsatisfactory, estimate forcing functions to represent lack of fit on the modeling scale.

4. Attempt to represent lack of fit in terms of any known quantities. The estimated state of the system is the typical focus for such a representation, but external influences may also be examined.

5. Re-formulate and re-estimate the differential equation model to account for any relationships found in Step 4.

6. Recurse.

Just as in linear regression, some trial and error must be expected. As pointed out, the reformulation step may not be straightforward. Moreover, the recipe, as stated here, ignores several other modeling choices that are available in the context of dynamical systems. Notably latent variables, higher order systems and transformations of the data.

4.1 Discovering the FitzHugh-Nagumo Equations

To demonstrate the proposed modelling strategy, we make use of the FitzHugh-Nagumo equations. In particular, we take the system from Section 2.1, observed every 0.05 time units on the interval [0,20] and add gaussian noise with variance 0.25. Taking this as our data, we then postulate a linear differential equation for V and R, with parameters estimated via NLS. This provides results in the system

$$\begin{bmatrix} \dot{V} \\ \dot{R} \end{bmatrix} = \begin{bmatrix} -0.6980 & 2.1026 \\ -0.4638 & -0.6843 \end{bmatrix} \begin{bmatrix} V \\ R \end{bmatrix} \tag{13}$$

A comparison between the estimated system and the observations in Figure 4 indicates a clear lack of fit. This lack of fit was then estimated via the techniques above using a B-spline basis expansion with 101 equi-spaced knots and a second derivative penalty. In this case, the smoothing parameter was chosen by eye. The second panel in Figure 4 gives the power for distinguishing these data from a linear differential equation, calculated from 500 simulations. In each simulation, the parameters of the linear differential equation were re-estimated and a restricted maximum likelihood ratio test applied. We again see that for moderate noise levels, representing lack of fit as forcing functions outperforms a representation on the scale of the residuals.

Figure 5 estimates this lack of fit as forcing components, with the lack of fit for V indicating strong periodicity. This has then been plotted against V itself, and the missing cubic trend, although contaminated by other dynamical components, becomes apparent. Adding a term proportional to V^3 in the equation for V in (13) recovers the form of the system appropriately.

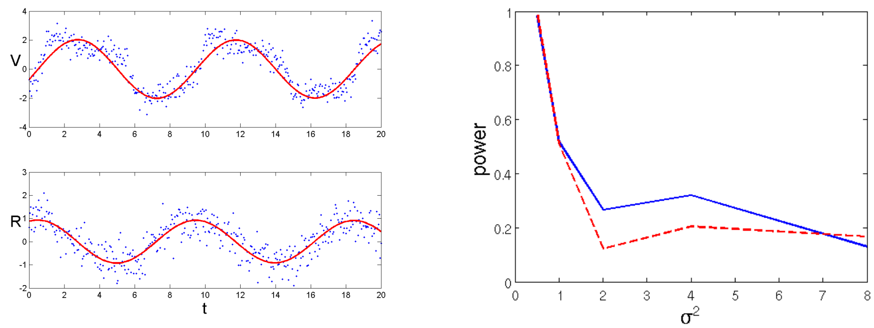

Figure 4: Discovering the FitzHugh-Nagumo Equations 1. Left: a system of linear differential equations (solid) clearly does not fit the data well. Right: power versus observational noise to test for lack of fit in a linear model.

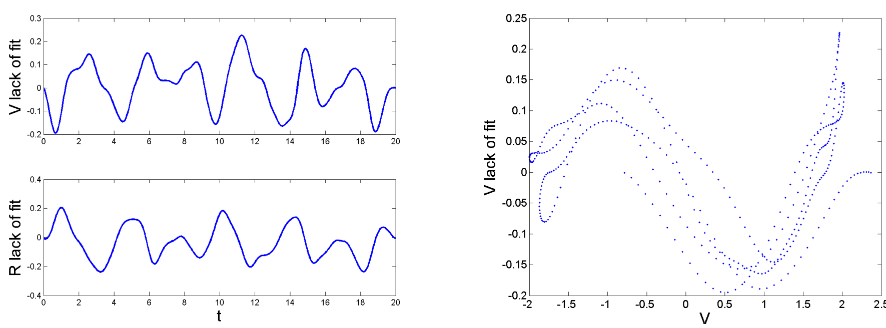

Figure 5: Discovering the FitzHugh-Nagumo Equations 2. Left: estimated lack of fit on the derivative scale. Right: lack of fit for V plotted against V; the missing cubic trend is apparent.

5 Unmeasured Components and Lack of Fit Identifiability

In nonlinear dynamics, it is typical that only some of the components of a system are measured. Data from voltage clamp experiments, for example, is only available for the variable V in the FitzHugh-Nagumo equations. For deterministic systems, the lack of measurements on some components does not present a problem for parameter estimation since the unmeasured components have a direct impact on those that are measured. However, the presence of such components does become a problem for the estimation of forcing functions.

Consider a two component system

$$\frac{d}{dt}x(t) = f(x(t), y(t), u(t))$$
$$\frac{d}{dt}y(t) = g(x(t), y(t), u(t)).$$

Assume that an exact, continuous representation is available for x. Now suppose that f and g have been approximated by mis-specified functions $\hat{f}$ and $\hat{g}$ respectively.

We can find $\bar{y}(t)$ to satisfy

$$\frac{d}{dt}x = \hat{f}(x(t), \bar{y}(t), u(t))$$

if $\hat{f}(x(t), \cdot, u(t))$ is invertible for each $x(t), u(t)$. Then it is possible to represent lack of fit by

$$h_y(t) = \frac{d}{dt}\bar{y} - \hat{g}(x(t), \bar{y}(t), u(t)) \tag{14}$$

Alternatively, it is possible to estimate lack of fit by setting $\tilde{y}(t)$ to solve

$$\frac{d}{dt}\tilde{y}(t) = \hat{g}(x(t), \tilde{y}(t), u(t)). \tag{15}$$

The original system can then be satisfied by adding:

$$h_x(t) = \frac{d}{dt}x(t) - \hat{f}(x(t), \tilde{y}(t), u(t)). \tag{16}$$

In more general terms, assuming that the map

$$f(x, \cdot, u) : \mathbb{R}^{k'} \to \mathbb{R}^{k'}$$

is invertible, the argument above can be generalized to show that it is only possible to identify the same number of forcing functions as there are observed components. **Statement: this is not a complete characterization.**

For a single observed component, it is possible to independently estimate a forcing function for each component in the system and perform model building by including the strongest observed relationships at each step. However, when d out of k components are observed, estimating the $\binom{d}{k}$ possible combinations of forcing functions becomes computationally intractable and analyzing them becomes intellectually infeasible. It may therefore be preferable to impose some identifiability constraints and estimate all the components jointly. A standard method of doing this is to use a ridge-type penalty, augmenting (7) with a term:

$$\lambda \sum \tau_i \int g_i(t)^2 dt$$

where the τ_i are chosen according to the scale of the variables x_i, or according to prior information about which components of x_i are likely to be mis-specified. In the FitzHugh-Nagumo equations with only V measured, doing so provided essentially the same diagnostic plots.

6 Real World Examples

In order to demonstrate the ideas from this paper in practise, we examine the results of a common experiment in neurobiology. Figure 6 gives readings taken from a voltage clamp experiment in which transmembrane potential is recorded on the axon of a single zebrafish motor neuron in response to an electric stimulus. When a stimulus exceeds a thresh-hold value, neurons respond by initiating the very sharp oscillations that can be seen in the first panel of Figure 6. There is a long literature on models that aim to describe neural behavior; see Wilson (1999) or Clay (2005) for an overview. These models range from very simple FitzHugh-Nagumo type systems to highly complex models involving many ion channels that appear as dynamical components.

The task we pose is to develop a tractable modification of a FitzHugh-Nagumo type that provides a reasonable agreement with the shape of the spikes found in the data. As a first modeling step, use a polynomial whose terms are motivated from equation (9.7) in Wilson (1999). There, the model was developed with the aim of mimicking the turning points of a physiological system. Our equations take the form:

$$\frac{d}{dt}V = p_1 + p_2 V + p_3 V^2 + p_4 V^3 + p_5 R + p_6 V R \tag{17}$$

$$\frac{d}{dt}R = p_7 + p_8 V + p_9 R \tag{18}$$

for unknown parameters $p_1, \ldots, p_9$. We observe that the form of equations (17) is invariant under linear transformations of V and R. We have used this property to derive initial guesses for the parameters by taking the parameters given in Wilson (1999) and transforming them to approximately match the peaks and troughs found in the data.

The second panel in Figure 6 presents a plot of a short sequence of these data along with the best fit from these equations. It is clear that while there is a

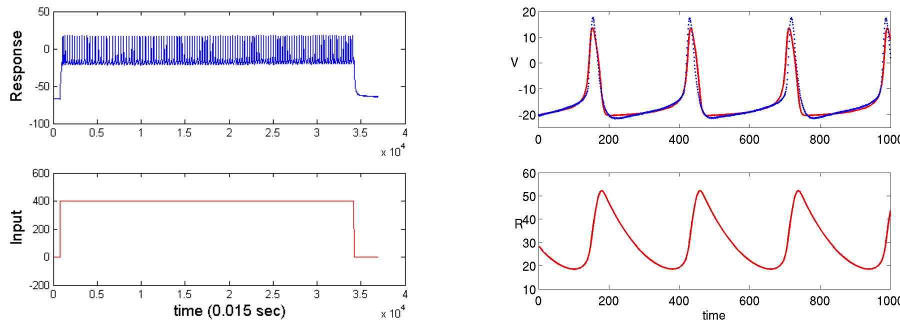

Figure 6: Left: data resulting from a patch-clamp experiment on a Zebrafish neuron. The bottom panel displays an input voltage, the top panel the neuron response in terms of measured transmembrane potential. Time units are 15ms. Left: a subsequence of 1000 observations from the same time series to be used in model development. Dots provide observed values, lines the results of a least-squares fit of equations (17) and (18) to these data.

qualitatively good agreement between the model and the data, some systematic lack of fit is evident.

Figure 7 presents lack of fit as estimated by nonlinear least squares on the domain. For these equations, a naive implementation of nonlinear least squares proved insufficient to estimate parameters or forcing functions. Instead, they have been estimated via a constrained optimization technique for collocation methods, implemented in the IPOPT routines Wächter and Biegler (2006) as described in Arora and Biegler (2004).

In order to estimate lack of fit, we used 200 linear B-splines in both variables. We have plotted the lack of fit functions for the observed variable V as a function of their estimated position in both variables. These plots make the difficulties of manually assessing lack of fit clear. The most evident relationship is between V, R and g_V. However, this relationship is given only for a one-dimensional curve in V, R space and guessing a useful functional form is correspondingly problematic. After spending some time rotating three-dimensional plots, we decided that the flat line evident in the right-hand panel of Figure 7 indicated a regime transition; the model providing a relatively good fit on one side of the regime while different parameters might be required on the other.

Formally, this model was implemented by modifying (17) to

$$
\begin{aligned}
\frac{d}{dt}V =& p_1 + p_2 V + p_3 V^2 + p_4 V^3 + p_5 R + p_6 V R \\
&+ \operatorname{logit}\left[(c_1 + c_2 R - V)c_3\right](q_1 + q_2 V + q_3 V^2 + q_4 V^3 + q_5 R + q_6 V R).
\end{aligned}
$$

The logistic term provides a smooth transition between the two regimes. Estimating parameters within this system produced a 100-fold decreased in total

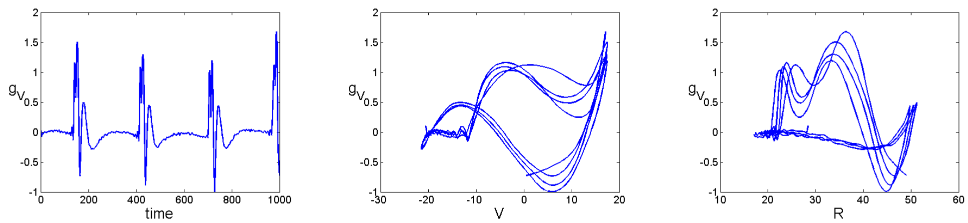

Figure 7: From left to right: lack of fit forcing functions for V, plotted against estimated trajectories for V and R, after fitting equations (18) and (17) to the patch clamp data in Figure 6.

squared error. However, we note that the modified equations, with the estimated parameters, occur in a chaotic region. In particular, small changes to c_3, which controls the steepness of the transition, provide a system which alternates between an eventual fixed point near zero and a system which eventually diverges. Thus, additional analysis is required in order to satisfy the requirements of a realistic system.

7 More General Model Choices

So far, the discussion in this paper has dealt with systems for which the functional form has been mis-specified, but under the assumptions that the number of components are known along with the number of derivatives and the measurement processes. For a given problem, none of these may be correct. It is therefore of some interest to find data-driven diagnostics for these forms of mis-specification.

There is a considerable literature on discovering the number of components in a dynamical system under the moniker of *Chaotic Data Analysis* Arbabanel (1996). These techniques are based on the Takens Embedding Theorem (Takens 1981), which in informal terms, states

Let $F : \mathcal{M} \rightarrow \mathcal{TM}$ be a C^2 vector field on a compact d-dimnesional manifold $\mathcal{M}$ and let $y = h(\mathbf{x})$ be a C^{2d+1} scalar transform of $\mathbf{x}$. Then

$$\mathbf{y}(t) = \left(h(\mathbf{x}(t)), \frac{d}{dt}h(\mathbf{x}(t)), \ldots, \frac{d^{2d+1}}{dt^{2d+1}}h(\mathbf{x}(t)) \right)$$

is an embedding

Unpacking the mathematical terminology, this theorem states that if $\mathbf{x}$ is an autonomous time-invariant dynamical system observed on its limit manifold, then $\mathbf{y}(t)$ is also an autonomous dynamical system which maintains many of the dynamical properties of $\mathbf{x}$. Takens (1981) proves an equivalent theorem for discrete-time dynamical systems.

These theorems have produced a large literature on so-called attractor reconstruction, largely based in the discrete time setting. Such techniques seek, firstly, to discover the dimension of the system, and then use nonparametric methods to recapture dynamical features. These techniques have potential applications for forcing function diagnostics in the sense that if $\mathbf{g}(t)$ in (2) can be written as

$$\mathbf{g}(t) = h(\mathbf{y}(t), \mathbf{z}(t))$$

for some unknown dynamical components $\mathbf{z}$, then $\mathbf{g}(t)$ satisfies the conditions of the Takens Embedding Theorem and it is possible to attempt to reconstruct the dimension of $\mathbf{z}$. Doing so requires some modification to existing techniques, however, to allow for the auxiliary measurement of the $\mathbf{y}$ variables and exploring these ideas has been left for future work.

The Takens Embedding Theorem has a somewhat less encouraging interpretation as well. Effectively, it states that it is impossible to distinguish between a system with many components and a higher order system. Thus, V in the FitzHugh-Nagumo (2.1) equations may be described by the second order differential equation:

$$\ddot{V} = \left(c - \frac{b}{c}\right)\dot{V} + (b-1)V - \frac{b}{3}V^3 - c\dot{V}V^2$$

which makes no reference to the recovery variable R. Moreover, a system involving derivatives no greater than 5 may be found for R that does not include V. Therefore, given measurements of both systems, there are two sets of equations that describe the data; one in which V and R interact, and one in which they do not. This has clear implications for data-driven model development. On a fundamental level, model building must be informed by a knowledge of the subject area, and even then should proceed with great care.

It should be noted, however, that these theorems are only available for *autonomous* systems which are observed on their limiting manifold. These theorems have been extended in, for example, (J. Stark and Huke 1997) to observations which are forced by other dynamical systems. However, systems observed away from their limiting manifold may provide further information to the modeler.

Similar comments may be made more generally about measurement misspecification. Such issues may arise, for example, if we suspect that a measuring instrument has been mis-calibrated. Suppose that we postulate an invertible measurement process $\mathbf{h} : \mathbb{R}^k \to \mathbb{R}^k$ then the measured process follows the differential equation:

$$\dot{\mathbf{x}} = \frac{d}{dt}\mathbf{h}(\mathbf{y}) = \mathbf{f}\left(\mathbf{h}^{-1}(\mathbf{x})\right)\frac{d\mathbf{h}}{d\mathbf{y}}\left(\mathbf{h}^{-1}(\mathbf{x})\right). \tag{19}$$

This identity states that if apparent mis-specification of a system may be corrected by an invertible transformation on the scale of the observations, then that transformation may be included into the form of the system itself. Frequently, dynamical systems exhibit appropriate qualitative behavior while at the same time having a poor quantitative fit. In this case, the simple expedient of translating and scaling the system to fit better may be easily handled and can be incorporated into a first modeling step. This approach was used to derive initial parameter guesses in Section 6. At the same time (19) also states that there is no way to distinguish mis-specification about the process of measuring a dynamical system and mis-specification of the dynamics themselves. Once again, the modeler is advised to proceed with caution.

8 Conclusion

The most appropriate representation for a models lack of fit is on the scale that the model is given. For differential equations, it is appropriate to measure lack of fit on the scale of derivatives. This observation leads to the estimation of forcing functions as diagnostic tools. This paper has shown that these tools provide useful insight into where a system of differential equations may be mis-specified. We have also derived approximate tests for the significance of lack of fit and the need for further modeling.

Simple residual analysis, however, is not sufficient as a model building technique. Lack of fit suffers from identifiability problems when the posited model has unmeasured components. More general types of model mis-specification – missing components, higher-order systems, and mis-specified measurement processes – also suffer from fundamental identifiability problems. This makes it important to exercise caution in exploratory model development in this field. As exemplified by our attempts to improve basic neurological models, it is also important to consider the dynamical properties of the proposed modification. In particular, care needs to be taken to preserve the stability of limit cycles, bifurcation points and other dynamical features of the model. Much further work is therefore warranted. In particular, extensions of attractor reconstructor techniques to the discovery of missing variables is an important future development. Similarly, interpolation tools that allow lack of fit to be estimated in such a way that it maintains the stability of limit cycles would prove very helpful for visual diagnostics. These techniques should be applicable to stochastic systems as well as the deterministic systems studied here. Of particular use in that context would be the development of tests for independence between estimated forcing functions and the trajectory of the model. We anticipate this field to be active and exciting for a long time into the future.

Acknowledgements

The zebrafish data displayed in this study was collected in and kindly supplied by Joe Fetcho's lab at the Biology Department in Cornell University. I am greatly indebted to Larry Biegler for the time and effort he spent in assisting me to set up these data and equations to estimate parameters within the IPOPT

system.

References

Aliev, R. R. and A. V. Panfilov (1996). A simple two-variable model of cardiac excitation. *Chaos Soltions and Fractals 7*, 293–301.

Arbabanel, H. D. I. (1996). *Analysis of Observed Chaotic Data*. New York: Springer.

Arora, N. and L. T. Biegler (2004). A trust region SQP algorithm for equality constrained parameter estimation with simple parametric bounds. *Computational Optimization and Applications 28*, 51–86.

Bates, D. M. and D. B. Watts (1988). *Nonlinear Regression Analysis and Its Applications*. New York: Wiley.

Clay, J. R. (2005). Axonal excitability revisited. *Progress in Biophysics and Molecular Biology 88*, 59–90.

Crainiceanu, C. M. and D. Ruppert (2004). Likelihood ratio tests for goodness-of-fit of a nonlinear regression model. *Journal of Multivariate Analysis 91*(1), 35–52.

Crainiceanu, C. M., D. Ruppert, G. Claeskins, and M. P. Wand (2005). Exact liklihood ratio tests for penalized splines. *Biometrika 92*(1), 91–103.

Deuflhard, P. and F. Bornemann (2000). *Scientific Compuitng with Ordinary Differential Equations*. New York: Springer-Verlag.

Eubank, R. L. and C. H. Spiegelman (1990). Testing the goodness of fit of a linear model via nonparametric regression techniques. *Journal of the American Statistical Association 85*(41), 387 – 392.

FitzHugh, R. (1961). Impulses and physiological states in models of nerve membrane. *Biophysical Journal 1*, 445–466.

Hart, J. D. (1997). *Nonparametric Smoothing and Lack of Fit Tests*. New York: Springer.

Hodgkin, A. L. and A. F. Huxley (1952). A quantitative descripotion of membrane current and its application to conduction and excitation in nerve. *J. Physiol. 133*, 444–479.

Huang, Y., D. Liu, and H. Wu (2006). Hierarchical bayesian methods for estimation of parameters in a longitudinal hiv dynamic system. *Biometrics*.

Ionides, E. L., C. Bretó, and A. A. King (2006). Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*.

J. Stark, D. S. Broomhead, M. E. D. and J. Huke (1997). Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis, Theory, Methods and Applications 30*(8), 5303–5314.

Jaeger, J., M. Blagov, D. Kosman, K. Kolsov, Manu, E. Myasnikova, S. Surkova, C. Vanario-Alonso, M. Samsonova, D. Sharp, and J. Reinitz (2004). Dynamical analysis of regulatory interactions in the gap gene system of *drosophila melanogaster*. *Genetics 167*, 1721–1737.

Nagumo, J. S., S. Arimoto, and S. Yoshizawa (1962). An active pulse transmission line simulating a nerve axon. *Proceedings of the IRE 50*, 2061–2070.

Ramsay, J. ., G. Hooker, D. Campbell, and J. Cao (2007). Parameter estimation in differential equations: A generalized smoothing approach. *Journal of the Royal Statistical Society, Series B (with discussion)*.

Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis*. New York: Springer.

Takens, F. (1981). *Detecting Strange Attractors in Turbulance*, pp. 366. Berlin: Springer.

Tjoa, I.-B. and L. Biegler (1991). Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic equation systems. *Industrial Engineering and Chemical Research 30*, 376–385.

Varah, J. M. (1982). A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific Computing 3*, 28–46.

Wächter, A. and L. T. Biegler (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* (106), 25–57.

Wilson, H. R. (1999). *Spikes, decisions and actions: the dynamical foundations of neuroscience*. Oxford: Oxford University Press.

Wood, S. N. and M. B. Thomas (1999). Super sensitivity to structure in biological models. *Proceedings of the Royal Society (B)*.