

*Appeared as:* Smith, J.T., Bowes, M.J., Clarke, R.T. (2010) Are groundwater nitrate concentrations reaching a turning point in some chalk aquifers? *Science of the Total Environment*, 408, 4722-4732.

## **Are groundwater nitrate concentrations reaching a turning point in some chalk aquifers?**

J.T. Smith<sup>a\*</sup>, R.T. Clarke<sup>b</sup>, M.J. Bowes<sup>c</sup>

a. School of Earth and Environmental Sciences, Burnaby Building, University of Portsmouth, Portsmouth PO1 3QL, UK. [jim.smith@port.ac.uk](mailto:jim.smith@port.ac.uk). Tel. +44 (0)2392 842416.

b. School of Conservation Sciences, Bournemouth University, Talbot Campus, Poole, Dorset, BH12 5BB, UK

c. Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

\* *Corresponding author.*

### **Abstract**

In past decades, there has been much scientific effort dedicated to the development of models for simulation and prediction of nitrate concentrations in groundwaters, but producing truly predictive models remains a major challenge. A time-series model, based on long term variations in nitrate fertiliser applications and average rainfall, was calibrated against measured concentrations from five aquifers in the River Frome catchment for the period spanning from the mid- 1970's to 2003. The model was then used to "blind" predict nitrate concentrations for the period 2003-2008. To our knowledge, this represents the first "blind" test of a model for predicting nitrate concentrations in aquifers. It was found that relatively simple time series models could explain and predict a significant proportion of the variation in nitrate concentrations in these aquifers ( $R^2$ : 0.6-0.9 and mean absolute prediction errors 4.2 – 8.0 %). The study highlighted some important limitations and uncertainties in this, and other modelling approaches, in particular regarding long term nitrate fertiliser application data. In three of the five aquifers (Hooke, Empool, Eagle Lodge), once seasonal variations were accounted for, there was a recent change in the generally upward historical trend in nitrate concentrations. This may be an early indication of a response to levelling-off (and declining) fertiliser application rates since the 1980's. There was no clear indication of trend change at the Forston and Winterbourne Abbas sites, nor in the trend of nitrate concentration in the River Frome itself from 1965-2008.

**Keywords:** Nitrate, fertiliser, groundwater, river, pollution, modelling

## 1. Introduction

Historically rising nitrate concentrations in rivers and groundwaters in the south of England, as a result of increasing fertiliser applications, are well documented (Heathwaite et al., 1996; Stuart et al., 2007). In many cases, groundwater nitrate concentrations are currently approaching, or exceeding, the  $11.3 \text{ mg NO}_3\text{-N l}^{-1}$  Drinking Water Standard as discussed by, for example Jackson et al. (2008). Understandably, therefore, in past decades, there has been much scientific effort dedicated to the development of models for simulation and prediction of nitrate concentrations in groundwaters. Owing to the large number of hydrological, soil, land use and aquifer processes involved in the transfer of nitrate from fertiliser applications to extracted groundwater, the development of truly predictive models remains a major challenge.

In an evaluation of nitrate transport in chalk catchments, Jackson et al. (2008) provide a very useful summary of types of model applied to this problem based on earlier work by Wheater et al. (1993). They distinguish – with increasing levels of representation of physical processes - between “metric”, “conceptual” and “physics-based” models. As defined by Jackson et al. (2008), metric models are “essentially statistical relationships between existing input and output datasets with rudimentary, if any, physical basis”; conceptual models “involve specifying a model structure *a priori*, normally on the basis of a system of conceptual stores”; and physics-based models “seek to capture a system’s response by incorporating significant processes through fundamental physical equations”. The full range of models, from “metric” statistical analyses of data (for example, Roy et al. (2007)) to “physics-based” models such as the INCA model ((Mathias et al., 2007; Wade et al., 2002; Whitehead et al., 1998)) are currently being used for nitrate research and management.

Although very useful, such categorisation of models is - in one sense – meaningless, since all models are neither more nor less than mathematical constructs designed to quantify the logical consequences of scientific hypotheses. Models cannot be evaluated by (necessarily) ad hoc categorisation, only by the comparative testing of their predictions against empirical data (Popper, 1963). It is, however, clear that different modelling “philosophies” lead to different levels of model complexity. Increasing complexity (the “reductionist” or “mechanistic” approach: detailed modelling of processes; greater spatial and/or temporal resolution) has both advantages and disadvantages. The advantage is that more physically-based models - *if* they can accurately simulate the real physical processes in an environmental system – may be better able to predict (extrapolate) real world events which are temporally, spatially or environmentally outside the scope of the model calibration. Less physically-based models may also achieve this, but because the process representation is less detailed, the basis for extrapolation is likely to be cruder (though it cannot be concluded, on this basis alone, that it will be any less successful).

Whilst having obvious advantages, physically-based models can suffer from two well-known problems. Firstly, they tend to be data intensive, often requiring detailed site-specific information on the physical processes they incorporate: this may not always be available, particularly for large-scale applications. In an evaluation of models applied to the radioactive contamination of catchments after the Chernobyl accident,

Monte (Monte et al., 2004) concluded that, in the context of post-accident prediction, “the inclusion of more processes in a complex model does not guarantee greater accuracy of model performance. Indeed the overall uncertainty of the model is strongly influenced by the uncertainty of large numbers of model parameters whose values cannot be known with a sufficient accuracy at site-specific level.”. This can make extrapolation to other sites very difficult.

A second problem, also related to information availability, is that physically based models may in practice be (unavoidably) over-parameterised with respect to the limited available test empirical data as discussed by, for example, McIntyre et al. (2005)). This leads to difficulties in testing and determining parameter values for these models. Parameter estimates are often statistically highly correlated, leading to problems of equifinality of model outcomes: no single “optimal” model can be found (Beven, 2006). Insufficient appropriate data makes it difficult not only to calibrate model process parameters but also to compare the predictive ability of different models for scenarios involving spatial and/or temporal extrapolation. It should be noted that the equifinality problem also applies to other model types, but is seen as less of a problem, since they do not aim for an accurate simulation of specific processes.

So, no modelling approach can be said, *a priori*, to be better than any other and (if the aim is prediction rather than “mechanistic” understanding) we must, where possible, rely on a subjective evaluation of utility (is the model “fit for purpose”?) coupled with Popper’s (1963) critical test of scientific hypotheses (models): the falsification of predictions against empirical data. It is with this in mind that we here present a preliminary predictive test of time-series models for nitrate concentrations in groundwaters. The approach is analogous to unit hydrograph based models for river flow rate and hence could be said to be at the less “physically-based” end of the nitrate modelling spectrum of Jackson et al. (2008). Our purpose is to test the extent to which such models can make useful predictions in this context, and to assess the important limitations to this approach. We further hope that this exercise will provide a predictive benchmark against which to evaluate other - possibly better – models and modelling approaches. To our knowledge, this represents the first “blind” test of a model for predicting nitrate concentrations in aquifers.

### 1.1. Study area

The River Frome catchment, Dorset, UK, covers an area of 414 km<sup>2</sup>, draining an area from the village of Evershot (ST 047576) on the Dorset – Somerset border, to Poole Harbour (Figure 1). The dominant bedrock geology for the majority of the catchment is Cretaceous chalk, with areas of Cretaceous greensand in the River Hooke sub-catchment, with fluvial sands and gravels in the lower reaches of the Frome. The land use within the catchment is primarily agricultural, mainly grassland and cereals. The town of Dorchester, with a population of ca. 27,000, is the only large town in the catchment. More detailed descriptions of the land use (Hanrahan et al., 2001), river chemistry (Bowes et al., 2009) and geology (Arnott et al., 2009) are given elsewhere.

## 2. Methods

### 2.1 Data for model calibration and testing

Groundwater nitrate concentrations were measured by Wessex Water in boreholes extracting groundwater from five sub-catchments within the catchment of the River Frome at various intervals during the period 1976-2008 (Table 1; Figure 1). Up to the mid-1990's, sampling was relatively infrequent (typically several samples per year), but from the mid-1990's onwards, samples were taken at approximately one- to two-week intervals. In addition, to assess the extent to which changes in nitrate in groundwater are reflected in the river water quality, a data set of nitrate concentrations in the River Frome at East Stoke, gathered by the Freshwater Biological Association and the Centre for Ecology and Hydrology is presented for the period 1965-2008.

Daily rainfall was estimated as an average of all rain gauges in the Frome catchment for the period 1973 – 2008 (beginning 1000 days prior to the first available measurement of nitrate in groundwater). The number of rain gauges operating varied during this period from 2-3 in the period 1973-1992 to 3-5 in the period 1993-2008. Though the relatively small and varying numbers of rain gauges increases spatial errors in estimating daily rainfall at individual sites across the catchment, we are here studying a significantly spatially-smoothed and temporally-damped groundwater response to short term changes in rainfall. In other words, the model prediction of nitrate concentration in groundwater on a particular day is a function of a time-weighted average of many individual daily rainfall estimates. Thus, the simple averaging of daily rainfall data across the catchment is sufficient for our purpose.

Mineral fertiliser applications to UK agricultural land have changed enormously since the 1940's as a result of the intensification of agriculture with more N fertiliser per hectare being applied over time ((DEFRA, 2009); see Figure 2) and land use change from pasture to arable land. Fertiliser trends in Dorset were similar to national trends (Casey et al., 1993) during this period. In addition to mineral fertiliser, nitrate in manure may also be a significant input to groundwaters. In 2007, approximately 30% of UK fields surveyed by DEFRA (DEFRA, 2009) received manure applications, primarily to grassland from cattle farmyard manure or slurry. For arable land, manure applications tended to be somewhat offset by decreased mineral fertiliser applications, but this was not observed for grassland. It is likely, therefore, that nitrate additions to grassland are significantly higher than those shown in Figure 2, though the extent to which the historical trend for manure applications differs from that of mineral fertiliser applications is not known. In 2006, total atmospheric N inputs to the Frome catchment were in the region 10-20 kg/ha/yr (CEH, 2006); important for natural and semi-natural areas but relatively low in comparison with recent agricultural fertiliser inputs to arable and grassland (see Figure 2).

UK average fertiliser inputs (across all crops and grasses; (DEFRA, 2009)) were used to estimate temporal changes in nitrate inputs to the system. Nitrate inputs prior to 1969 were estimated by linear interpolation, assuming (Limbrick, 2003) that pre-war mineral fertiliser inputs were effectively zero. The Frome catchment (and the groundwater catchments studied here) is mixed arable and grassland (Hanrahan et al., 2001; Rukin et al., 2008), and therefore it is reasonable to assume that historical nitrate inputs in the Frome catchment are likely to have followed a similar trend to

that shown in Figure 2 (rising applications until the 1980's, subsequently levelling off and slightly declining to the present day). The trend is broadly similar to nitrate trends in these aquifers as reconstructed in (Rukin et al., 2008). However, as noted by a previous study of nitrate in the Frome catchment (Casey et al., 1993) it is very difficult to estimate accurately past mineral and organic fertiliser use even when current land use is known. As discussed in the "Limitations and Uncertainties" section below, this places an important limitation on the model (and on other models for long-term nitrate changes).

It is noted that, in the present model, the temporal *pattern* of nitrate applications is important, but the absolute amounts in  $\text{kg ha}^{-1}$  are not, as we do not here attempt a mass-balance of nitrate (the  $A$ ,  $B$ , parameters in Equations (1-4) below are not physically interpreted). It is assumed that (due to crop uptake, denitrification and surface runoff processes) only a proportion of nitrate applications is available for transport to groundwater, but for our purposes here, this proportion does not need to be quantified.

## 2.2. Modelling

Four different models (Models 1-4 in order of decreasing complexity) were evaluated in terms of their ability to simulate and predict nitrate data in the boreholes and river. This allowed an assessment of the extent to which the additional input data (seasonality, agricultural nitrate inputs; rainfall) improves model fits and predictions.

Our first, most complex model (Model 1) is formed from a series of two exponential functions which model a "fast" and "slow" response component to time changes in nitrate concentrations in groundwaters. Such models have been used successfully for a number of applications, including predicting activity concentrations of  $^{90}\text{Sr}$  and  $^{137}\text{Cs}$  in rivers following surface depositions to catchments from atmospheric nuclear weapons testing and the Chernobyl accident (Cross et al., 2002; Sasina et al., 2007; Smith et al., 2000).

Time-series measurements of nitrate concentrations in groundwaters ( $C_N$ ; dimension  $[\text{M}][\text{L}]^{-3}$ ) at 5 boreholes were fitted using the transfer function approach based on a linear combination of two exponential functions (Model 1). These modelled the long-term response of nitrate concentrations to slowly-changing nitrate inputs ( $N$ ;  $[\text{M}][\text{L}]^{-2}[\text{T}]^{-1}$ ) to the system and the shorter-term response to rapidly-varying rainfall ( $R$ ;  $[\text{L}][\text{T}]^{-1}$ ):

$$C_N(t) = A_1 \int_{\tau=-\infty}^t R(\tau) e^{-k_1(t-\tau)} d\tau + B_1 \int_{\tau=-\infty}^t N(\tau) e^{-k_2(t-\tau)} d\tau$$

where  $A_1$  ( $[\text{M}][\text{L}]^{-4}$ ),  $B_1$  ( $[\text{L}]^{-1}$ ),  $k_1$  ( $[\text{T}]^{-1}$ ),  $k_2$  ( $[\text{T}]^{-1}$ ) are parameters to be determined by fitting to the empirical data. The model represents seasonal cycles superimposed on the long-term trend of rising nitrate inputs and concentrations as identified in an analysis of trends in nitrate concentrations in the Dorset and Hampshire Basin chalk aquifer (Roy et al., 2007).

Since the input data are not continuous, a discrete form of this model, referred to as Model (1) was used to determine the nitrate concentration in groundwater on day  $i$  of year  $w$ :

$$C_i(w) = A_1 \sum_{j=i-L-1000}^{i-L} R_j e^{-k_1(i-j-L)} + B_1 \sum_{j=w-100}^w N_j e^{-k_2(w-j)} \quad (1)$$

where  $k_1$  has units days<sup>-1</sup>, and  $k_2$  has units years<sup>-1</sup>.  $L$  (days) represents a time lag term to investigate whether, at any of the boreholes, a time delay (advective time lag) between rainfall and short term changes in nitrate concentration in groundwater could improve model fits and predictions. The limits were chosen to be long enough such that, respectively, rainfall 1000 days prior to the time of nitrate concentration measurement, and nitrate inputs 100 years prior to that time would have no significant effect on the seasonal and long term components of the response.

In an approach similar to the unit hydrograph method for rainfall-flow modelling, the model assumes that current nitrate concentrations are an exponentially declining function of past rainfall and past nitrate applications. Thus, the nitrate concentration at a particular time is determined by the sum of all historical nitrate applications, each being weighted in an exponentially decreasing manner (so, for equal applications in year  $x$  and  $x - 1$ , the weight given to the older application,  $x - 1$ , will be a factor  $e^{-k_2}$  of the weight given to the more recent application). Upon this long term trend is superimposed a more rapid variation determined by past rainfall events in which the weight given to past rainfall events also declines exponentially. Conceptually, the model can be considered as representing a system of two storage compartments, one giving a slow aquifer response to changing nitrate inputs, the other giving a more rapid response to variations in rainfall (groundwater level). Clearly, this is not (nor is it intended to be) an accurate simulation of the many complex processes involved in nitrate transport.

Model 2 is a simplified version of Equation (1) in which nitrate concentrations are assumed to increase linearly with time:

$$C_i(w) = A_2 \sum_{j=i-L-1000}^{i-L} R_j e^{-k_1(i-j-L)} + B_2 w + C_2 \quad (2)$$

where  $A_2$ ,  $B_2$ ,  $C_2$  are parameters to be fitted and  $w$  is time in integer years.

Model 3a is a simple statistical model developed by Casey and Clarke (1979) to describe the long-term trend and the seasonal pattern in the weekly samples of nitrate concentration obtained from the River Frome at East Stoke over the period 1965-1975. This model involves a linear trend and a superimposed sinusoidal season pattern as represented by Equation (3a):

$$C_i(t) = A_3 \cos(2\pi(t - M)) + B_3 t + C_3 \quad (3a)$$

where  $t$  = time in decimal years measured from the beginning of 1965.

Model 3a is fitted using non-linear least squares to derive estimates of the parameters representing the linear trend  $B_3$  ( $\text{mg l}^{-1} \text{yr}^{-1}$ ), the size of cyclical seasonal variation as represented by twice the amplitude ( $A_3$ ) of the cosine wave, the average timing of the peak in seasonal variation as represented by  $M$  (in decimal years), and a constant term  $C_3$ .

To assess whether the long-term trends in concentration are showing signs of slowing down or even reaching a plateau, we also fitted model 3b with the long-term trends fitted as S-shaped curves, akin to those used by Howden & Burt (2008), who refer to them as breakthrough curves:

$$C_i(t) = A_3 \cos(2\pi(t - M)) + 0.5(C_{30} + C_{31}) + (C_{31} - C_{30})\text{Erf}(B_3(t + B_{3t})) \quad (3b)$$

where  $\text{Erf}(t)$  is the standardised Normal cumulative distribution function:

$$\text{Erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dx$$

By fitting a further model (denoted “3y”) which included terms for both the cosine seasonality and all  $m$  individual year differences, we tested whether the extra two parameters in the S-shaped curve of Model 3b gave a statistically significant improvement over Model 3a in explaining between-year differences in average concentration by using an approximate F-test, namely:

$$F_{ba} = \frac{(RSS_{3b} - RSS_{3a})/2}{(RSS_{3y} - RSS_{3b})/(m - 4)}$$

with 2 and  $m-4$  degrees of freedom where  $RSS_x$  denotes the residual sum of squares of Model  $x$

It may be that the amplitude of the seasonal cycle increases as with time as the general level of nitrate concentration increases. To test this, we included an extension of Equation (3a) to allow the seasonal variation amplitude to increase linearly at a rate  $A_{t3}$  per year as given in Equation (3c):

$$C_i(t) = (A_3 + A_{t3}t) \cos(2\pi(t - M)) + B_3t + C_3 \quad (3c)$$

Finally, Model 4 is a simple linear regression of  $C_N$  against time:

$$C_N(t) = A_4t + B_4 \quad (4)$$

where  $A_4$  and  $B_4$  are parameters to be fitted.

Programs were written in Microsoft Visual Basic to prepare and collate the large data sets of measurements of nitrate in groundwaters, with the associated rainfall and nitrate input data, in the format required for input to SAS (SAS Institute Inc.). Model fits and parameter estimates were optimised using the non-linear fitting procedure, PROC NLIN in SAS which minimises the residual (observed minus fitted values) sum of squares for the observed data. This also provides estimates of confidence intervals

of the fitted parameters for each model (1)-(4). Goodness of fits of the models to the measurements was assessed using Efron's pseudo- $R^2$  which is equivalent to the square of the correlation coefficient between the modelled and measured values. MS Visual Basic programs were written to check the SAS model fits, and to prepare model output data for presentation using MS EXCEL.

The models were calibrated against data to September 2003, then run in predictive mode for "blind" testing against data from the period October 2003 – October 2008 (see Table 1)

### 3. Results

#### 3.1. Model fits

Each of the four models was fitted against the measured nitrate concentrations in groundwater for the period to the end of September 2003. Table 2 presents the  $R^2$  values determined for each of the model fits. Model 3 (Equation 3a), which includes a seasonal cycle superimposed on a linear trend, gives a much better fit than the basic linear trend model. In statistical tests (based on  $F_{ab}$ ), there was no statistically significant evidence of any improvement in fit of the S-shaped trend curve in Model 3b over Model 3a with a simple linear long-term trend, in terms of explaining the between-year variability in concentrations for any of the five boreholes except the one at Hooke ( $F_{ab}$  test probability  $p < 0.001$ ). Models involving a 4-parameter logistic curve for the trend gave almost identical fits to those of Model 3b suggesting that the lack of improvement in fit was not due to the choice of non-linear trend curve. Thus over the model calibration period up to 2003, at four of the five boreholes there was no clear evidence that the regular annual increase in average concentration was diminishing or coming to an end.

The fits of the "conceptual" Models 1 and 2 are better than those of the purely descriptive ("metric") trend plus seasonal cycle (Models 3, Equations 3a and 3b), even though Models 1 and 2 involve the same number or fewer parameters. Comparisons of Models 1 and 2 against empirical data from each site are shown in Figure 3 (a) to (e). When daily rainfall information is included (possibly acting as a surrogate for groundwater level and other parameters), the model is able to explain significantly more of the variation in nitrate concentrations in groundwaters than a model with a fixed seasonal cycle for all years (c.f. Models 2 and 3 in Table 2).

For four of the sites (Forston, Hooke, Empool, Eagle Lodge), Models 1 and 2 perform similarly, though the fit of Model 1 is always marginally better than that of Model 2 (both have the same number of fitting parameters). For one of the sites (Winterbourne Abbas), however, the second term of Model 1 did not converge to a solution (this is discussed below).

Fitted parameters for the "best" fit model (1 or 2) for each site are presented in Table 3.

### 3.2. Time lag

For each of the five sites, model runs (Models 1 & 2) were carried out for different values (0, 10, 20, 30, 40, 50 days) of the time lag parameter,  $L$ . The model sensitivity to lag time, as expected, depended on the amplitude of seasonal variation at each site. For three of the five sites (Forston, Hooke, Empool), the introduction of a time lag component either reduced or did not improve model goodness-of-fit. For two of the sites, however, introduction of a time lag term (20 days for Eagle Lodge, 40 days for Winterbourne Abbas) increased the  $R^2$  values from 0.81 to 0.89 and from 0.72 to 0.78, respectively.

From the parameters determined for Model 3 (Table 4),  $52M$  estimates the week in the year at which, on average, the nitrate concentrations is maximum. For the River Frome at East Stoke over the period 1965-2003 the peak was estimated to be in early February ( $52 \times 0.1078 = \text{week } 5.6$ ); Casey and Clarke (1979)'s estimates for the first decade was the same (week 5). For three of the five boreholes the estimated average time of peak concentrations was several weeks later in the spring than for the river at East Stoke, namely for Empool (week 15.7 – mid-April), Eagle Lodge (week 11.6 – late March) and Winterbourne Abbas (week 16.1 – late April).

### 3.3. Best fit parameters

Fitted parameters for the “best” model for each site are presented in Table 3 from which half-times for the exponential functions are presented in Table 5. The half-times ( $\ln 2/k_1$ ) for the rainfall transfer function represent the time in days for the influence (weighting) of a given rainfall event on future nitrate concentrations to fall to half the value it had on the day of the rainfall event. The half-times ( $\ln 2/k_2$ ) for the nitrate inputs transfer function represent the number of years for the influence (weighting) of a given nitrate input to fall to half the value it had in the year of the nitrate input.

### 3.4. Model predictions

Using parameter values determined from the model fitting exercise to data up to September 2003, the models were then run in predictive mode for a five-year period from October 2003-2008. Thus, using input data on rainfall and nitrate applications to September 2008, “blind” predictions were made of concentrations in each of the five aquifers. Models 1 and 2 and 3 were run, excepting the Winterbourne Abbas site, where only Models 2 and 3 could be applied. To ensure that no bias was introduced in the prediction exercise, the test data on nitrate concentrations for the period 2003-2008 were not available to the developers (J.T. Smith; R.T. Clarke) at the time that the predictions were made.

For Models 1 and 2, model blind predictions against empirical data are presented in Figure 3, and in greater detail in Figure 4.

## 4. Discussion

### 4.1. Model fits

The fits of these – in terms of modelled processes - relatively simple models are generally very good. Between 60 – 90% of the variation in nitrate concentrations is explained by changes in nitrate inputs and variation in rainfall for both Models 1 and 2. Forston, where long term changes in nitrate concentrations are relatively small, showed the poorest fit ( $R^2 = 0.60$ ) whilst the other sites, showing a greater long-term change (Figure 3) showed better fits ( $R^2 = 0.78 - 0.90$ ). Clearly, if rises in  $\text{NO}_3$  concentrations are steeper over the calibration period, both Models 1 and 2 (which account for such changes) are able to explain a greater proportion of the variation in measurements. This is supported by the comparison with Model 4 (simple linear regression with time, with no rainfall model). At Forston, the simple linear regression with time (Model 4) explains only 10% of the variation in measured values, implying inconsistent changes in  $\text{NO}_3\text{-N}$  concentrations over time during the sampling period, but also highlighting the improvement afforded by the rainfall response model, which explains a further 50% of the variation (Table 2).

In Model 3, the seasonal cycle amplitude  $A_3$  in Equation (3a) was statistically significant for all five boreholes (Table 4, all test  $p < 0.001$ ) and was also a notable improvement over the simplest Model (4) with just a trend (Table 2). Thus there is a noticeable seasonal cycle to nitrate concentrations in ground water. The average seasonal variation (measured by  $2A_3$ ) varies from  $0.5 \text{ mg l}^{-1} \text{ N}$  at the Empool borehole to around 1.5 at Forston, Eagle Lodge and Winterbourne Abbas (Table 4). Model 3b, with an increasing seasonal variation amplitude, was only statistically significant for the Empool borehole and then  $R^2$  only increased by 0.4% to 75.4%. From this, we conclude that over the range of years available, there is no evidence to suggest that the extent of seasonal variation in these groundwater sites has increased over time.

Models 1 and 2 are virtually indistinguishable in terms of their ability to fit the empirical data (Table 2): thus it cannot be concluded – at this stage – that incorporating information on changing nitrate inputs is better than the Model 2 assumption of linearly increasing  $\text{NO}_3$  concentrations with time. Models 1 & 2, both of which account for variation in rainfall, offer improved fits over Models 3 (sinusoidal variation plus linear (3a) or S-shaped (3b) increase in  $\text{NO}_3$  with time ) and Model 4 (simple linear increase in  $\text{NO}_3$  concentrations over time).

A significant proportion of the seasonal variation in nitrate concentrations is explained by the “rainfall” component of Equations 1 and 2. As illustrated in Figures 3 and 4, the timing and the height of the peaks and troughs in  $\text{NO}_3\text{-N}$  concentrations are generally well represented by the models. Figure 5 illustrates some of the data (1994-2003) in greater detail for the Forston site. It is clear that the models often fail to simulate some very high nitrate concentrations, though this pattern is not consistent: as illustrated in Figure 5, some seasonal peaks are very well represented by the model. It is likely that there is small time scale variation that cannot be predicted by the model, though it is also possible that there are rapid hydrological processes which could, in principle, be modelled: for example, rapid surface or sub-surface runoff from high rainfall events following recent application and/or soil tillage. Further investigation of these events would be valuable, though high time-resolution

data would be required.

Figure 5 also illustrates a season (winter of 1996/97) in which there were two peaks in nitrate concentration in groundwater. These two peaks are well simulated by the model (1). Overall, the rainfall model (Model 1) simulates (and predicts; see below) the magnitude of seasonal peaks in NO<sub>3</sub>-N concentrations quite well: it can be seen in Figure 4 that, in the relatively dry winters (03/04; 04/05; 05/06) the model predicts significantly lower peaks in nitrate concentrations. This gives us confidence that the rainfall relationship is causal, though the relationship is likely to be indirect, possibly via the influence of seasonal rainfall patterns on soil microbial activity and water level and transfer in the aquifer. The influence of rainfall on groundwater level (and thus on aquifer flushing) is expected to be an important factor: using linear regression, Roy et al. (2007) observed correlation between groundwater level and nitrate concentrations in a number of aquifers.

#### 4.2. Model blind predictions

The model blind predictions are presented in Figure 4. The “goodness of fit” of the model predictions can be quantified (Yaffee and McGee, 2000) by calculating the mean percentage prediction error:

$$\varepsilon_{mean} = \frac{100}{N} \sum_{t=1}^N \frac{(\hat{y}_t - y_t)}{y_t} \quad (5)$$

and the mean absolute percentage prediction error:

$$\varepsilon_{abs} = \frac{100}{N} \sum_{t=1}^N \left| \frac{\hat{y}_t - y_t}{y_t} \right| \quad (6)$$

where  $N$  is the total number of measurements being forecast,  $y_t$  is the measured value (in this case the nitrate concentration in groundwater) at forecast time  $t$  and  $\hat{y}_t$  is the predicted value at time  $t$ .

Note that we use the term prediction here not in the sense of forecasting the future (since rainfall and nitrate application data are required inputs for the model), rather as the prediction of nitrate concentrations for given nitrate applications and rainfall conditions. Forecasting, for future nitrate application and rainfall scenarios would, however, be possible in principle.

Both Models 1 and 2 provide relatively good predictions of nitrate concentrations in the five boreholes (though Model 1 did not converge for Winterbourne Abbas, see below), and both show an improvement over Model 3 (Table 6). In particular, both models predict the years of minor peaks in NO<sub>3</sub> concentrations (winter 03/04; 04/05; 05/06) and the relatively higher peaks in 06/07 and 07/08. For four of the five sites (Forston, Hooke, Empool, Eagle Lodge), it was possible to test the predictions of both models. As is evident from the percentage prediction errors (Table 6) in three of the four sites, Model 1 gave better (and in all cases lower) prediction errors than Model 2, implying that the incorporation of time changes in NO<sub>3</sub> inputs had improved the model; for Forston (which showed relatively little long-term change in NO<sub>3</sub>

concentrations over the study period), there was no significant difference between the two models. This may have been influenced by the shorter study period at Forston (1986-2008 compared to 1976 or 1978 -2008, see Table 1).

Once the seasonal variation in rainfall/groundwater level is accounted for, the aquifers at Hooke, Empool and Eagle Lodge all appear to be showing a change in the trend of continuing upward nitrate concentrations, as evidenced by the better predictions of Model 1 compared to Model 2 (Figure 4; Table 6). In an analysis of nitrate trends in the Dorset and Hampshire Chalk aquifers, Roy et al. (2007) noted a falling NO<sub>3</sub> concentration since 2001. These workers assumed that this reduction “represents a downward oscillation superimposed on a long-term rising trend”. The present study suggests an alternative possibility for the Hooke, Empool and Eagle Lodge aquifers: that this falling trend is a response to historical declines in nitrate fertiliser applications, though the uncertain pattern of these changes means that the response time is uncertain (see “Limitations and uncertainties” section). Further measurements over the next few years, will confirm or refute this latter hypothesis.

Model 1 could not be fitted to the measurements from the Winterbourne Abbas site: the SAS program failed to converge on a solution. Model 2, however, gave an excellent fit ( $R^2 = 0.89$ ) to the Winterbourne Abbas data and gave good predictions for the period 2003-08 (Figure 4). Model 2 assumes continually increasing NO<sub>3</sub> inputs to the system, so the fact that it fits well implies that there is, as yet, no evidence of any change in the upward trend of nitrate concentrations in this aquifer. In fact, the increasing nitrate concentrations from this borehole periodically exceed the 11.3 mg NO<sub>3</sub>-N EU drinking water limit, showing an urgent need to reduce the nitrate loading to this aquifer.

#### 4.3. Limitations and Uncertainties

Whilst the general UK pattern of rising NO<sub>3</sub> inputs to the mid-1980's then subsequently declining inputs (Figure 2) is likely to have been followed in the Frome catchment (Rukin et al., 2008), the precise historical pattern at each site is not known. We have examined the effects of this uncertainty by carrying out model fits for two alternative patterns of nitrate inputs; firstly, assuming the average UK nitrate application to grassland and, secondly, the average application to tilled land (see Figure 2). As shown in Table 7, the different nitrate input functions did not influence goodness of fit, nor did they significantly influence the parameter values derived for the short-term rainfall dependent component. However, the different input functions did significantly influence the estimated values of the “ $k_2$ ” parameter (and the associated response half-time) which model the long term changes in nitrate concentrations.

A further problem with the nitrate application rate data is that, in the words of Young (Young, 2001), it is not *sufficiently exciting*: it does not perturb the system in a sufficiently irregular manner or with enough turning points to strongly constrain the precise model structure and parameter values. These data may, however, slowly become more “exciting” as time progresses and continued lower levels of nitrate applications lead to more pronounced impact on observed nitrate concentrations.

The results shown in Table 7 give us confidence that the short term rainfall response function is sufficiently independent of the long term nitrate change function that it is not significantly affected by the poor spatial resolution of the nitrate input data. However, the lack of site-specific historical nitrate input data limits the conclusions which can be drawn concerning the long term response parameter,  $k_2$ , since its value is strongly dependent on the nitrate input function assumed.

Whilst acknowledging the important limitations in the model (which allow only tentative conclusions to be drawn concerning the long-term component), it is important to note that such limitations in input data also apply to other modelling approaches. A more “mechanistic” model may make the limitations in input data less explicitly obvious, but they are likely still to remain.

#### 4.4. Aquifer response to changing rainfall and nitrate inputs

The half-times of the rainfall response function vary from 33 days (CI: 27.5-40.5) at Hooke to 477 days (CI: 345-636) at Empool (Table 5). Thus, as observed in the empirical data, the rainfall response at Empool is much more damped than at Hooke which has a rapid response to seasonal rainfall changes. A very damped rainfall response implies that longer-term rainfall/water level related processes may be impacting on nitrate concentrations. This could, in principle, be investigated further by truncating the rainfall response function (first term in Eq. 1) to different time periods prior to each nitrate measurement.

Whilst noting the limitations in the long term response parameter values, these (Table 5) suggest response times of the order of one or more decades for these aquifers. No clear conclusions can be drawn about the long term response times of the Forston (insufficient temporal change over the period 1986-08 for which data were available) or Winterbourne Abbas sites (apparently continuing increases in long term nitrate concentrations over the period 1978-08).

Figure 6 shows trends in nitrate concentration 1975-2008 in four of the five aquifers (Forston, Hooke, Empool, Eagle Lodge), obtained by running Model 1 with the seasonal variation component removed. All show a declining rate of increase in nitrate concentration in recent years, though Model 1 only showed improved predictions over Model 2 (linear increase with time) in Hooke, Empool and Eagle Lodge. Winterbourne Abbas showed no evidence of a changing rate of increase with time.

#### 4.5. Changes in nitrate concentrations in the River Frome

A previous study (Howden and Burt, 2009) has analysed temporal changes in nitrate concentrations in the River Frome using monthly measurements made by the Environment Agency during the period 1978-2007. They concluded that the linear trend in increases in nitrate concentration previously observed by Clarke (1979) has continued in recent years at sites on the main River Frome between Dorchester and West Holme (encompassing East Stoke). We here present, for the first time, higher resolution (approximately weekly) measurements of nitrate in the Frome for the period 1965-2008.

Here the Casey & Clarke (1979) Model (3a) is updated to fit the River Frome at East Stoke monitoring data to cover the period 1965-2003 (Table 4; Figure 7). The long-term average annual increase in  $\text{NO}_3$  in the River Frome at East Stoke over the period 1965-2003 is estimated to be 0.0970 (SE = 0.0013), which is slightly less than the estimate of 0.109 for the initial decade (1965-75) by Clarke & Casey (1979), but is slightly higher than the estimate of 0.087 from the later study of Howden and Burt (2009) based on 25 years of monthly data over the period 1978-2007. Model 3b with the S-shaped breakthrough curve for long-term trends did not converge for this river data, indicating that over the model calibration period there was no evidence of any reduction in the annual rate of increase in nitrate concentrations within the lower Frome at East Stoke.

For the river data, Model 3c provides a better fit than Model 3a, with the seasonal variation  $2(A_3+A_{13}t)$  at year  $t$  estimated to have increased from 0.9  $\text{mg l}^{-1}$  N in 1965 to over 1.94  $\text{mg l}^{-1}$  N by 2003, which exceeds the size of seasonal variation observed in the most seasonally variable boreholes (Table 4). The fact that an increase in seasonal variation was observed in the river data, but not in the boreholes, may be due to the longer time series available for the river data.

## 5. Conclusions

A very simple conceptual model based on readily-available rainfall and nitrate fertiliser input data was shown to explain between 60 and 90% of the variation in nitrate concentrations in five chalk aquifers. These time-series models gave good “blind” predictions of nitrate concentrations over five years (2003-2008) following the calibration period. The model accurately predicted years of high and low nutrient concentrations, though some very high (i.e. significantly above seasonal trend) measured values (observed in some years and not in others) were not predicted.

In three of the five aquifers (Hooke, Empool, Eagle Lodge), once seasonal variations have been accounted for, there has been a change in the generally upward historical trend in nitrate concentrations. This accords with the observations of Roy et al. (2007) in their study of trends in the Dorset and Hampshire Chalk aquifer. The present study suggests that, rather than being a temporary oscillation, it is possible that this is an early indication of a response to levelling-off (and declining) fertiliser application rates since the 1980's. We could see no clear indication of trend change at the Forston and Winterbourne Abbas sites, nor was there any indication of trend change in the water of the lower River Frome. A study of the amplitude of seasonal oscillations in nitrate concentrations (Model 3c) identified an increasing amplitude of oscillations with time in the River Frome (from 1965-2003) but did not identify any change in seasonal amplitude in groundwater concentrations (from mid-1970's to 2003).

The testing of four different models of different complexity has shown that increasing complexity from “metric” (Models 3 and 4) to “conceptual” models (Models 1 and 2) with the inclusion of key input data (rainfall; nitrate inputs), as might be expected, has improved model fits and predictive power. However, the good fits of the (still relatively simple) conceptual models imply that parameterisation of more complex models against these relatively “unexciting” nitrate trend data is likely to be difficult.

The half-times and time lags determined in this study have not here been interpreted with reference to any soil, geological or aquifer characteristics. It would, however, be valuable to test for correlations between model parameter values describing the aquifer response (lag times, half times) and key aquifer parameters (such as type of overlying soil, depth to water table, hydraulic conductivity). Such an analysis may require modelling of other groundwater systems since our study here of five boreholes is unlikely to produce statistically significant relationships.

This study has also shown that the statistical time series approach, whilst giving useful predictions and insights into the dynamic modes of the system, has significant limitations. In particular, the difficulty of obtaining accurate, site-specific, historical records of nitrate applications places a serious limitation on the approach (and, we believe, other modelling approaches). Thus, with the data currently available, we are unable to estimate accurate timescales for the long term response component. It would be interesting in future to test whether improved input data (more site specific nitrate application trend data; estimated evapotranspiration and soil moisture; groundwater level) could improve model fits and predictions.

## 6. References

- Arnott S, Hilton J, Webb BW. The impact of geological control on flow accretion in lowland permeable catchments. *Hydrology Research* 2009; 40: 533-543.
- Beven K. A manifesto for the equifinality thesis. *Journal of Hydrology* 2006; 320: 18-36.
- Bowes MJ, Smith JT, Neal C. The value of high-resolution nutrient monitoring: A case study of the River Frome, Dorset, UK *Journal of Hydrology* 2009; 378: 82-96.
- Casey H, Clarke RT. Statistical analysis of nitrate concentrations from the River Frome (Dorset) for the period 1965-1976. *Freshwater Biology* 1979; 9: 91-97.
- Casey H, Clarke RT, Smith SM. Increases in Nitrate Concentrations in the River Frome (Dorset) Catchment Related to Changes in Land Use, Fertiliser Applications and Sewage Input. *Chemistry and Ecology* 1993; 8: 105 - 117.
- CEH. UK Deposition Maps - Total N 2006. Centre for Ecology and Hydrology, 2006.
- Cross MA, Smith JT, Saxen R, Timms D. An analysis of the environmental mobility of radiostrontium from weapons testing and Chernobyl in Finnish river catchments. *Journal of Environmental Radioactivity* 2002; 60: 149-163.
- DEFRA. The British Survey of Fertiliser Practice. Fertiliser use on farm crops for crop year 2008. Department for the Environment, Food and Rural Affairs, London, 2009.
- Hanrahan G, Gledhill M, House WA, Worsfold PJ. Phosphorus Loading in the Frome Catchment, UK: Seasonal Refinement of the Coefficient Modeling Approach. *J Environ Qual* 2001; 30: 1738-1746.
- Heathwaite AL, Penny JJ, Norman EP. Trends in nutrients. *Hydrological Processes* 1996; 10: 263-293.
- Howden NJK, Burt TP. Temporal and spatial analysis of nitrate concentrations from the Frome and Piddle catchments in Dorset (UK) for water years 1978 to 2007: Evidence for nitrate breakthrough? *Science of The Total Environment* 2008; 407: 507-526.

- Howden NJK, Burt TP. Statistical analysis of nitrate concentrations from the Rivers Frome and Piddle (Dorset, UK) for the period 1965-2007. *Ecohydrology* 2009; 2: 55-65.
- Jackson BM, Browne CA, Butler AP, Peach D, Wade AJ, Wheater HS. Nitrate transport in Chalk catchments: monitoring, modelling and policy implications. *Environmental Science & Policy* 2008; 11: 125-135.
- Limbrick KJ. Baseline nitrate concentration in groundwater of the Chalk in south Dorset, UK. *The Science of The Total Environment* 2003; 314-316: 89-98.
- Mathias SA, Butler AP, Ireson AM, Jackson BM, McIntyre N, Wheater HS. Recent advances in modelling nitrate transport in the Chalk unsaturated zone. *Quarterly Journal of Engineering Geology and Hydrogeology* 2007; 40: 353-359.
- McIntyre N, Jackson B, Wade AJ, Butterfield D, Wheater HS. Sensitivity analysis of a catchment-scale nitrogen model. *Journal of Hydrology* 2005; 315: 71-92.
- Monte L, Brittain JE, Hakanson L, Smith JT, van der Perk M. Review and assessment of models for predicting the migration of radionuclides from catchments. *Journal of Environmental Radioactivity* 2004; 75: 83-103.
- Popper KR. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge, 1963.
- Roy S, Speed C, Bennie J, Swift R, Wallace P. Identifying the significant factors that influence temporal and spatial trends in nitrate concentrations in the Dorset and Hampshire Basin Chalk aquifer of Southern England. *Quarterly Journal of Engineering Geology and Hydrogeology* 2007; 40: 377-392.
- Rukin N, Roy S, Davison P. *Modelling nitrate concentrations with variations in time*. UK Water Industry Research Ltd, 2008.
- Sasina NV, Smith JT, Kudelsky AV, Wright SM. "Blind" testing of models for predicting the Sr-90 activity concentration in river systems using post-Chernobyl monitoring data. *Journal of Environmental Radioactivity* 2007; 92: 63-71.
- Smith JT, Clarke RT, Saxen R. Time-dependent behaviour of radiocaesium: A new method to compare the mobility of weapons test and Chernobyl derived fallout. *Journal of Environmental Radioactivity* 2000; 49: 65-83.
- Stuart ME, Chilton PJ, Kinniburgh DG, Cooper DM. Screening for long-term trends in groundwater nitrate monitoring data. *Quarterly Journal of Engineering Geology and Hydrogeology* 2007; 40: 361-376.
- Wade AJ, Durand P, Beaujouan V, Wessel WW, Raat KJ, Whitehead PG, et al. A nitrogen model for European catchments: INCA, new model structure and equations. HAL - CCSD, 2002.
- Wheater HS, Jakeman AJ, Beven KJ. Progress and directions in rainfall-runoff modelling. In: Jakeman AJ, Beck MB, McAleer MJ, editors. *Modelling change in environmental systems*. John Wiley & Sons Ltd., Oxford, 1993.
- Whitehead PG, Wilson EJ, Butterfield D. A semi-distributed integrated nitrogen model for multiple source assessment in catchments (INCA): Part I -- model structure and process equations. *Science of The Total Environment* 1998; 210-211: 547-558.
- Yaffee RA, McGee M. *Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*: Academic Press, Inc., 2000.
- Young PC. Data-based mechanistic modelling and validation of rainfall-flow processes. In: Anderson MG, Bates PD, editors. *Model Validation: Perspectives in Hydrological Science*. Wiley, Chichester, 2001, pp. 117-161.

## Acknowledgements

We would like to thank Paul Stanfield of Wessex Water for making data sets available for this study and Prof. John Hilton for early discussions concerning these data.

## List of figures

**Figure 1.** Map of the Frome catchment showing the sites of groundwater abstraction and river water quality monitoring.

**Figure 2.** Trends in (a) UK average nitrate inputs to grassland and tillage per hectare; (b) nitrate concentration ( $\text{NO}_3\text{-N mg l}^{-1}$ ) in groundwater at Hooke.

**Figure 3.** (a)-(e) Fit of Model 1 to data from the five boreholes to September 2003 and predictions of Models 1 and 2 from Sept 03 – end 08.

**Figure 4.** (a)-(e) Predictions of Models 1 and 2 shown in greater detail.

**Figure 5.** Detail of Model 1 fit for Forston data illustrating accurate simulation of double nitrate peak in one winter season but poor prediction of some high nitrate peaks.

**Figure 6.** Trends in nitrate concentration 1975-2008 in four of the five aquifers obtained using Model 1 but with the seasonal variation component removed.

**Figure 7.** Fit and prediction of Model 3 to high resolution data from the River Frome (update of Casey and Clarke; 1979).

**Table 1.** Available data on nitrate concentrations in groundwater for model fitting and “blind” predictions. n = number of observations

Site	Data for model fitting			Data for model testing		
	Start date	End date	n	Start date	End date	n
Forston	28/10/86	26/09/03	535	02/10/03	20/02/08	144
Hooke	05/01/76	26/09/03	472	21/10/03	23/10/08	369
Empool	13/01/76	08/04/03	414	21/02/05	21/10/08	170
Eagle Lodge	04/10/76	30/09/03	621	06/10/03	21/10/08	112
Winterbourne Abbas	28/06/78	04/04/03	235	02/03/04	16/09/08	280

**Table 2.**  $R^2$  values for model fits to nitrate data for the period to end September 2003. The  $R^2$  values for the “best” model are shown in bold.

Site	<b>Model 1</b> Transfer function using rainfall and nitrate input data	<b>Model 2</b> Transfer function using rainfall plus linear increase in N	<b>Model 3a</b> Regression: linear time trend + cosine seasonal variation	<b>Model 3b</b> Regression: S-shaped trend + cosine seasonal variation	<b>Model 4</b> Simple linear regression with time
Forston	<b>0.60</b>	0.60	0.45	0.45	0.10
Hooke	<b>0.90</b>	0.89	0.85	0.88	0.79
Empool	<b>0.80</b>	0.79	0.75	0.75	0.71
Eagle Lodge	<b>0.78</b>	0.75	0.65	0.68	0.26
Winterbourne Abbas	D.N.C.*	<b>0.89</b>	0.72	0.77	0.57

\* D.N.C. – did not converge.

**Table 3.** Fitted parameters for the “best” model for each site. Estimated 95% confidence intervals are given in brackets.

Site	Best model	<i>Time Lag, days</i>	$A_1$	$k_1$	$B_1$	$k_2$
Forston	Model 1	0	$8.24 \times 10^{-3}$ (7.44-9.03 $\times 10^{-3}$ )	$9.97 \times 10^{-3}$ (8.39-11.6 $\times 10^{-3}$ )	0.00153 (0.00121 0.00186)	0.0442 (0.0287 0.0597)
Hooke	Model 1	0	$6.66 \times 10^{-3}$ (5.76-7.57 $\times 10^{-3}$ )	$21.1 \times 10^{-3}$ (17.1 -25.2)	0.00200 (0.00186 0.00215)	0.0164 (0.0120 0.0207)
Empool	Model 1	0	$2.06 \times 10^{-3}$ (1.73-2.38 $\times 10^{-3}$ )	$1.55 \times 10^{-3}$ (1.09- 2.01 $\times 10^{-3}$ )	0.00351 (0.00326 0.00377)	0.0668 (0.0595 0.0741)
Eagle Lodge	Model 1	20	$6.18 \times 10^{-3}$ (5.79-6.58 $\times 10^{-3}$ )	$5.60 \times 10^{-3}$ (4.99-6.21 $\times 10^{-3}$ )	0.00261 (0.00232 0.00290)	0.0629 (0.0541 0.0717)
			$A_2$	$k_1$	$B_2$	$C_2$
Winterbourne Abbas	Model 2	40	$7.16 \times 10^{-3}$ (6.44-7.89 $\times 10^{-3}$ )	$3.78 \times 10^{-3}$ (3.17-4.40 $\times 10^{-3}$ )	0.1323 (0.1141 0.1505)	-1.1095 (-1.5678 - 0.6513)

**Table 4.** Estimates of regression coefficients ( $\pm$  SE) for Model (3a) for each of the five boreholes and for the nitrate concentrations in the river Frome at East Stoke, all up to end September 2003

Site	$C_3$	$B_3$	$A_3$	M
Forston	$3.173 \pm 0.278$	$0.0836 \pm 0.0083$	$0.7737 \pm 0.0422$	$0.1290 \pm 0.0087$
Hooke	$1.345 \pm 0.127$	$0.1874 \pm 0.0038$	$0.4470 \pm 0.0321$	$0.0553 \pm 0.0112$
Empool	$4.523 \pm 0.134$	$0.1422 \pm 0.0043$	$0.2470 \pm 0.0282$	$0.3013 \pm 0.0190$
Eagle Lodge	$4.504 \pm 0.161$	$0.1073 \pm 0.0048$	$0.7627 \pm 0.0289$	$0.2239 \pm 0.0063$
Winterbourne Abbas	$0.175 \pm 0.348$	$0.2378 \pm 0.0111$	$0.7818 \pm 0.0689$	$0.3098 \pm 0.0153$
River Frome at East Stoke	$2.252 \pm 0.027$	$0.0970 \pm 0.0013$	$0.7120 \pm 0.0196$	$0.1078 \pm 0.0044$

**Table 5.** Half times of response to rainfall and changing nitrate inputs based on best fit model. To facilitate comparison between sites, parameter values from model runs with the lag term  $L$  set to zero are used for all five sites (95% confidence limits in brackets).

Site	Best fit model	Rainfall response		Response to nitrate inputs	
		$k_1 \text{ d}^{-1}$	Half time $T_{1/2}, \text{ d}$	$k_2 \text{ y}^{-1}$	Half time $T_{1/2}, \text{ y}$
Forston	Model 1	$9.97 \times 10^{-3}$ (8.39-11.6 $\times 10^{-3}$ )	69.5 (82.6-59.7)	0.0442 (0.0287 0.0597)	15.7 (24.1-11.6)
Hooke	Model 1	$21.1 \times 10^{-3}$ (17.1 -25.2 $\times 10^{-3}$ )	32.8 (40.5-27.5)	0.0164 (0.0120 0.0207)	42.3 (57.8-33.5)
Empool	Model 1	$1.55 \times 10^{-3}$ (1.09- 2.01 $\times 10^{-3}$ )	447 (636 - 345)	0.0668 (0.0595 0.0741)	10.4 (11.6-9.35)
Eagle Lodge	Model 1	$3.62 \times 10^{-3}$ (3.17-4.08 $\times 10^{-3}$ )	191 (219-170)	0.0619 (0.0501 0.0737)	11.2 (13.8-9.40)
Winterbourne Abbas	Model 2	$1.84 \times 10^{-3}$ (1.29- 2.40 $\times 10^{-3}$ )	377 (537 - 289)	-	-

**Table 6.** Mean ( $\epsilon_{mean}$ ) and absolute ( $\epsilon_{abs}$ ) percentage prediction errors for forecasts from October 2003-October 2008.

Site/Model	$\epsilon_{mean}$			$\epsilon_{abs}$		
	Model 1	Model 2	Model 3a	Model 1	Model 2	Model 3a
Forston	+ 0.23 %	+ 2.85 %	+ 14.04 %	7.43 %	7.48 %	15.61 %
Hooke	+ 4.35 %	+ 9.31 %	+ 11.62 %	5.32 %	9.62 %	11.76 %
Empool	+ 1.43 %	+ 12.30 %	+ 16.80 %	4.21 %	12.30 %	16.80 %
Eagle Ldge	+ 1.40 %	+ 6.75 %	+ 13.68 %	4.75 %	7.70 %	13.68 %
W. Abbas	n/a	- 6.43 %	+ 5.94 %	n/a	7.96 %	7.21 %

**Table 7.** Model 1 fits for the Hooke aquifer assuming different nitrate input patterns (shown in Figure 2). Values in brackets show estimated 95% confidence intervals.

Assumed nitrate input	Best fit $R^2$	Rainfall response		Response to nitrate inputs	
		$k_1$ $days^{-1}$	Half-time $T_{1/2}$ days	$k_2$ $yrs^{-1}$	Half-time $T_{1/2}$ yrs
Mean tillage + grassland	0.90	$21.1 \times 10^{-3}$ (17.1 -25.2 $\times 10^{-3}$ )	32.8 (40.5-27.5)	0.0164 (0.0120 - 0.0207)	42.3 (57.8-33.5)
Grassland only	0.90	$21.0 \times 10^{-3}$ (17.0 -25.0 $\times 10^{-3}$ )	33.0 (40.1-27.7)	0.00759 (0.0041 - 0.011)	91.3 (169-63)
Tillage only	0.90	$21.3 \times 10^{-3}$ (17.1 -25.4 $\times 10^{-3}$ )	32.5 (40.5 – 27.3)	0.0266 (0.0210 - 0.0321)	26.1 (33.0-21.6)

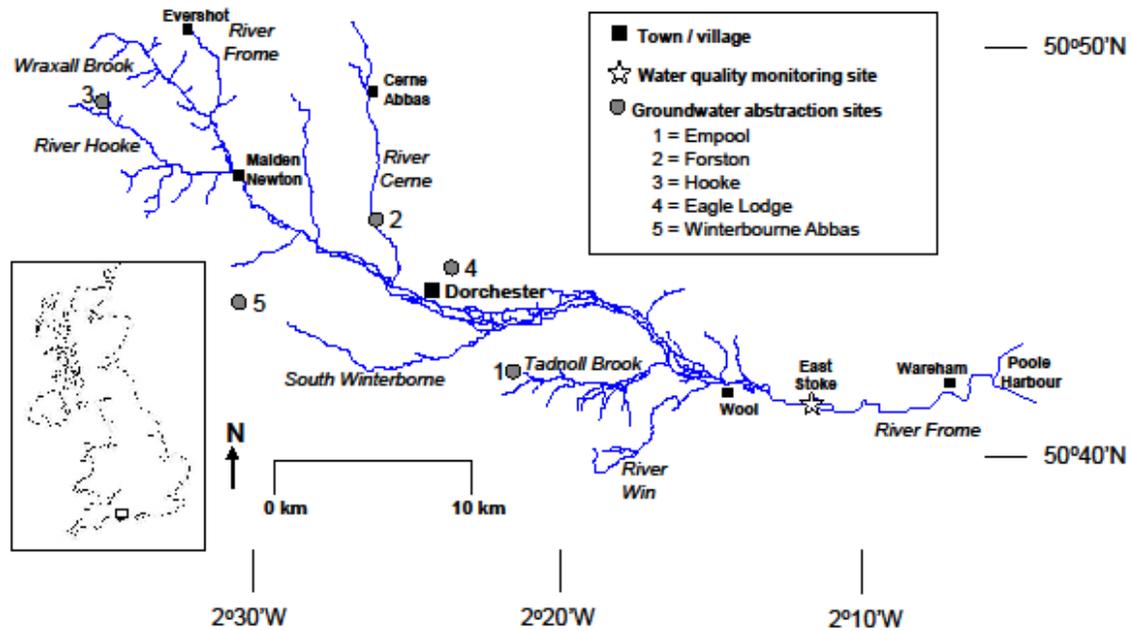


Figure 1.

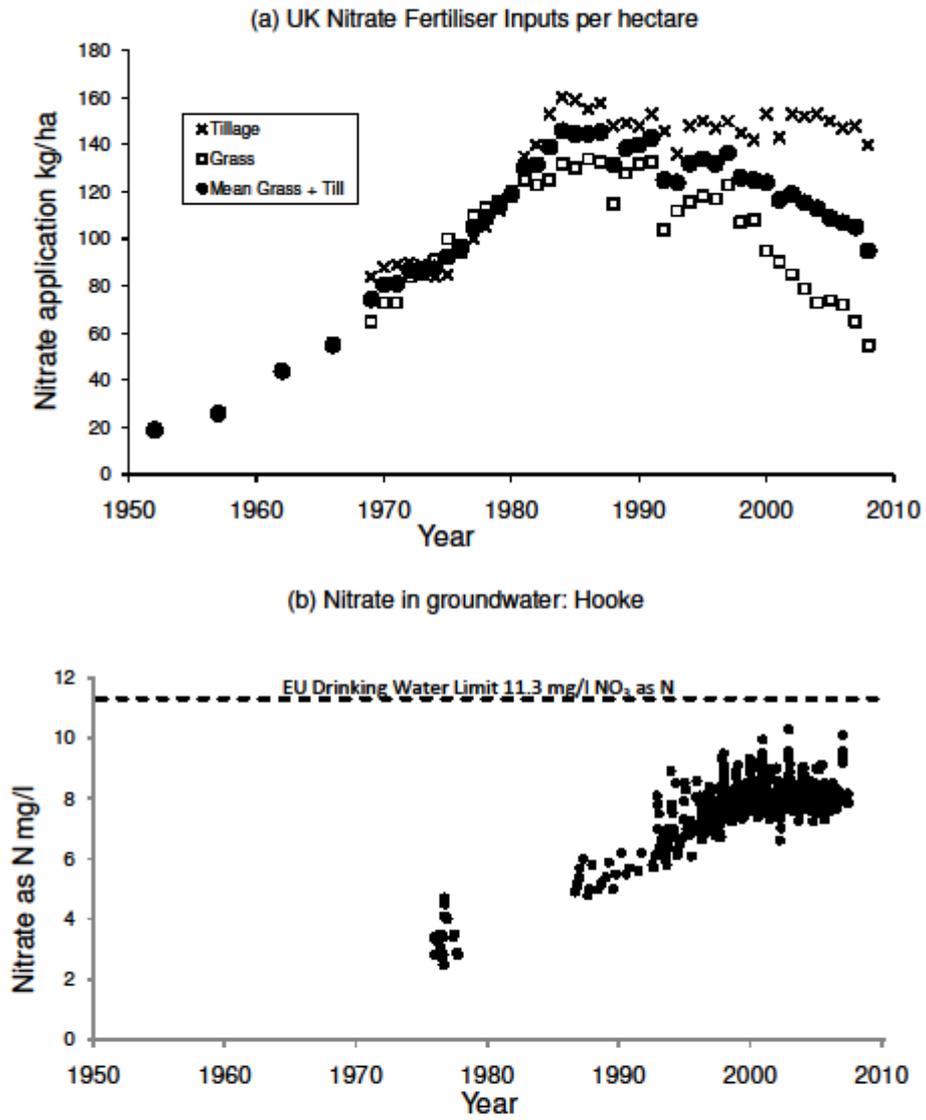


Figure 2.

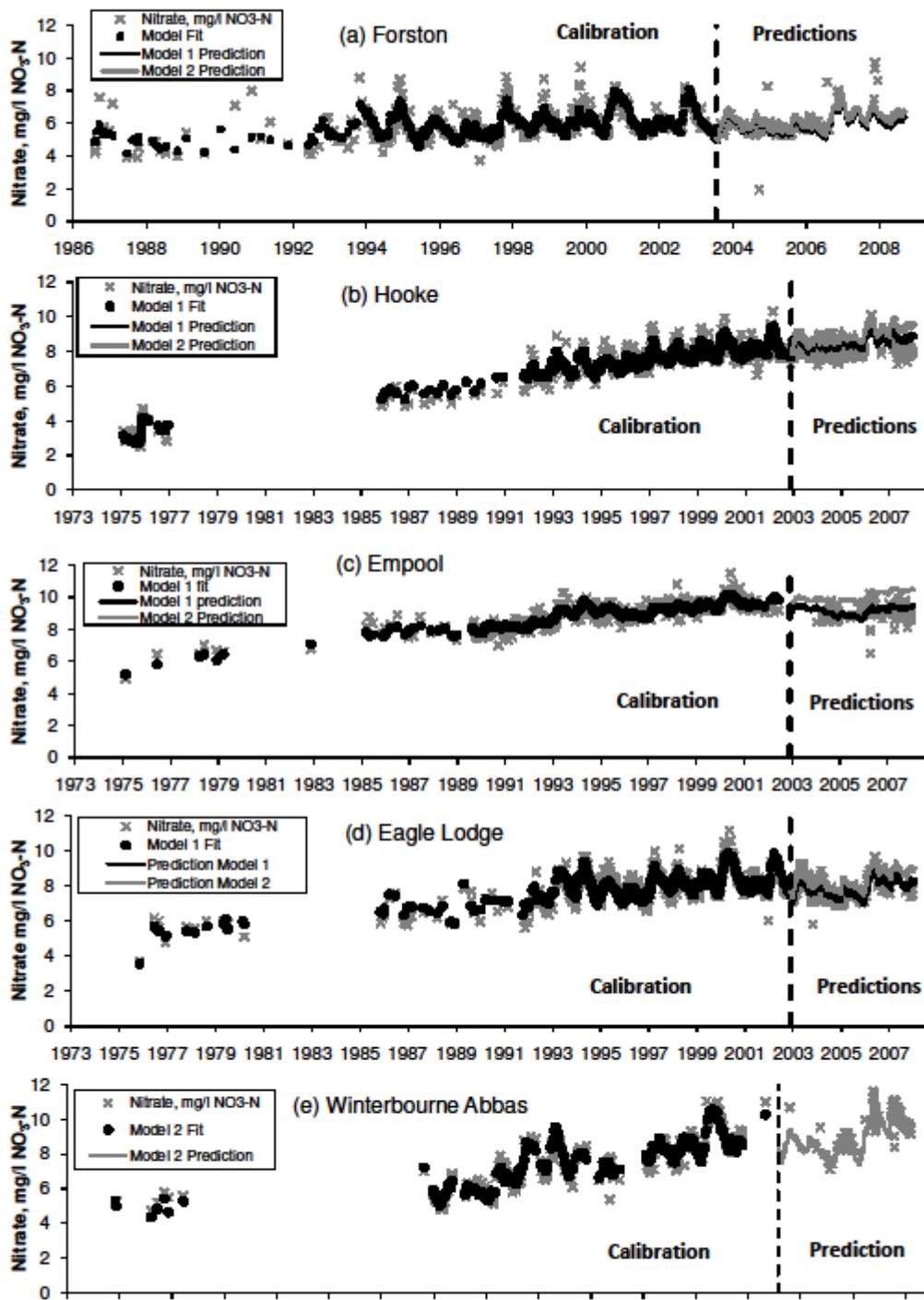


Figure 3.

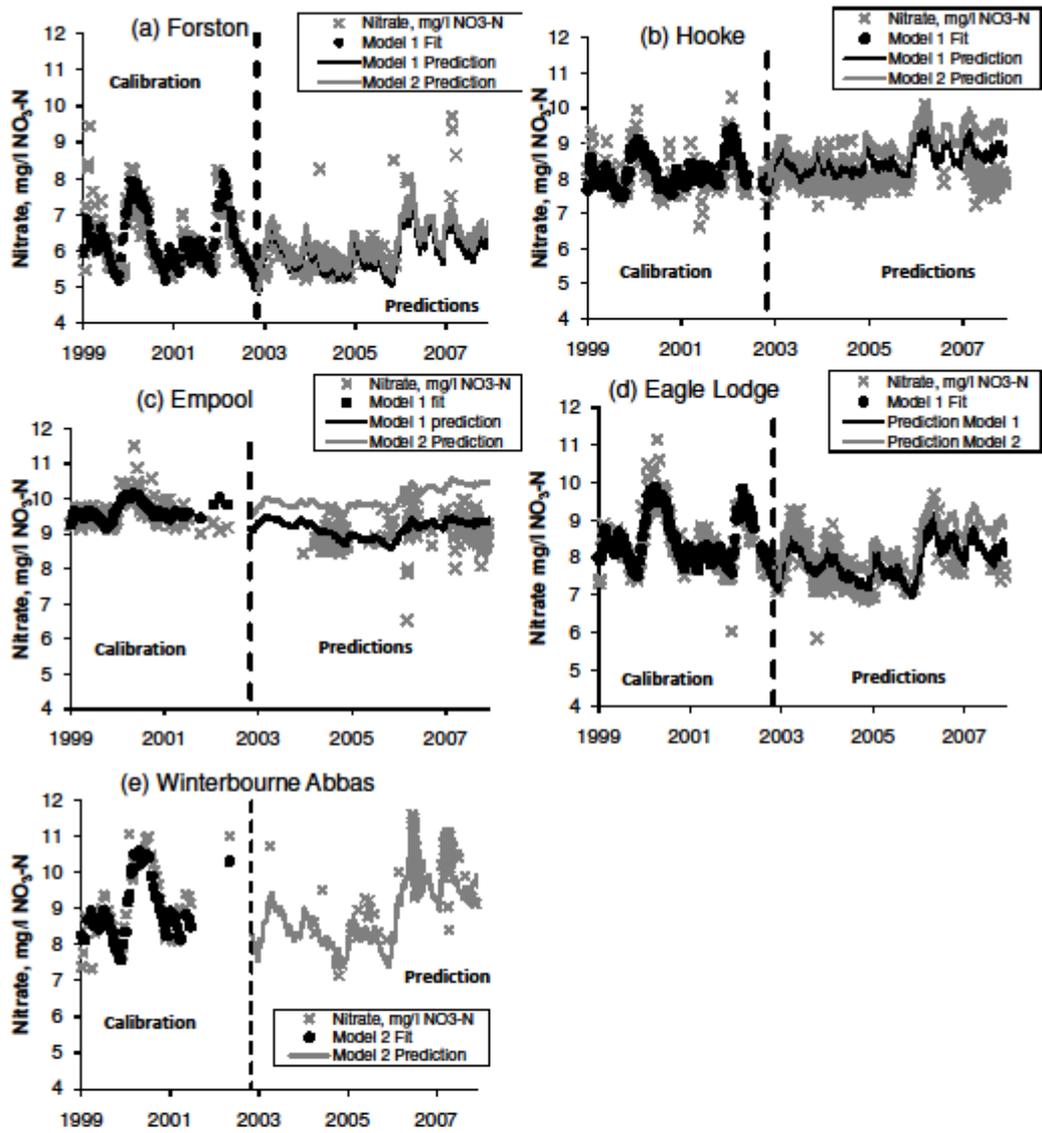


Figure 4.

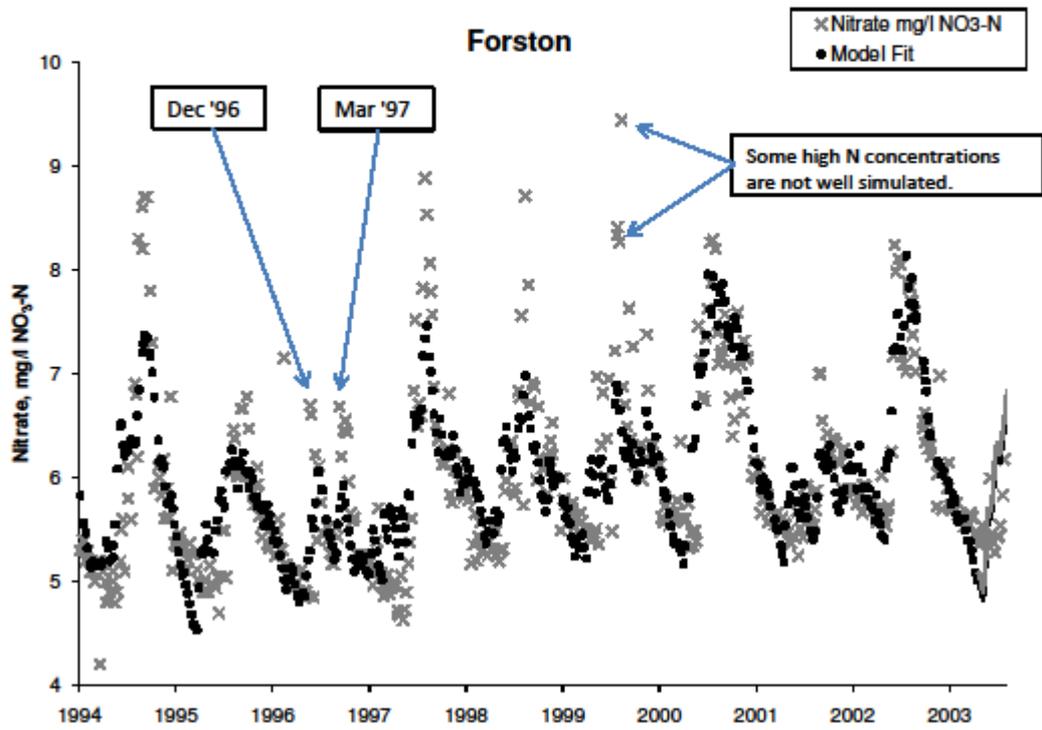


Figure 5.

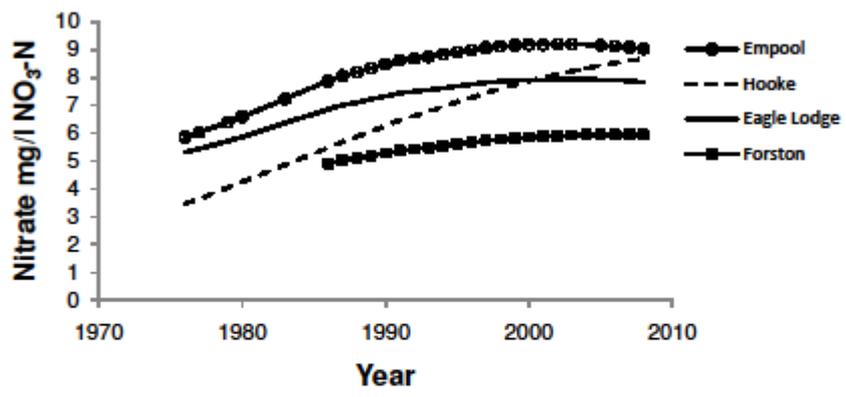


Figure 6.

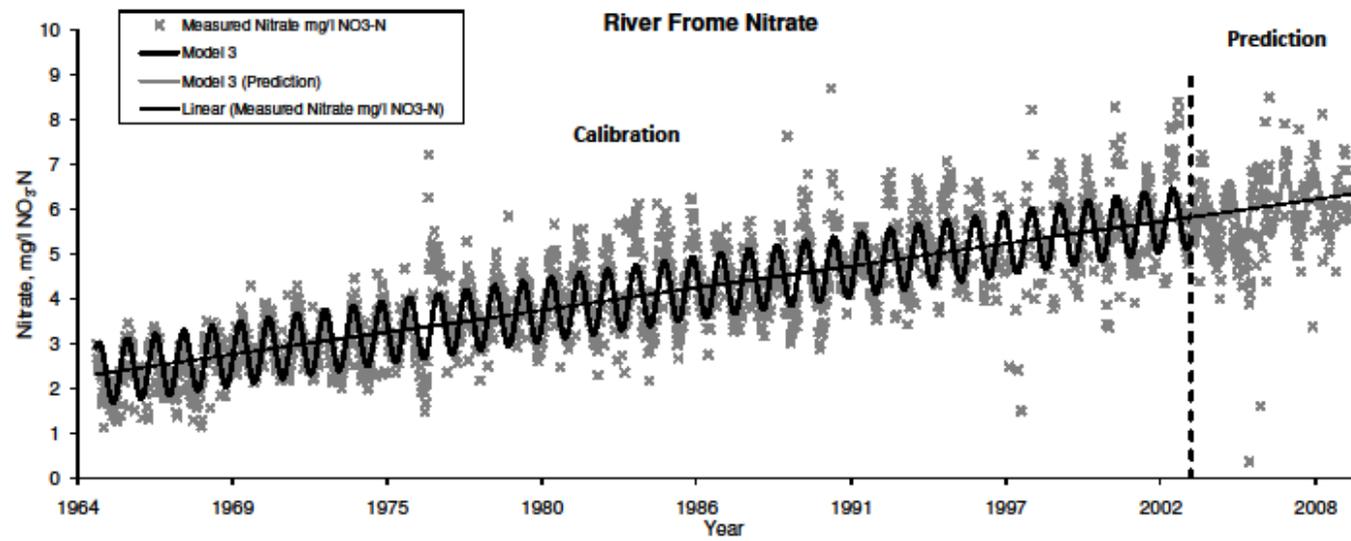


Figure 7.