

# Screening for long-term trends in groundwater nitrate monitoring data

M. E. Stuart<sup>1</sup>, P. J. Chilton<sup>1</sup>, D. G. Kinniburgh<sup>1</sup> & D. M. Cooper<sup>2</sup>

<sup>1</sup> British Geological Survey, Wallingford, OX10 8BB, UK (e-mail: mest@bgs.ac.uk)

<sup>2</sup> Centre for Ecology and Hydrology, Wallingford, OX10 8BB, UK

## Abstract

A large body of UK groundwater nitrate data has been analysed by linear regression in order to define past trends and estimate future concentrations. Robust regression was used. The datasets showed too many irregularities to justify more traditional time series approaches such as ARIMA-type methods. Tests were included for lack of linearity, outliers, seasonality and a break in the trend (by piecewise linear regression). Of the series analysed 21% showed a significant improvement in the overall fit when a break was included. Half of these indicated an increase in trend with time. Significant seasonality was found in about one third of the series, with the largest nitrate concentrations usually found during winter months. Inclusion of nearby water level data as an additional explanatory variable successfully accounted for much of this seasonality. Based on 309 datasets from 191 distinct sites, nitrate concentrations were found to be rising at an average of  $0.34 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ . In 2000, 34% of the sites analysed exceeded the  $50 \text{ mg L}^{-1}$  EU drinking water standard. If present trends continue, 41% could exceed the standard by 2015. We explored an alternative to the previously-proposed WFD aggregation approach for estimating trends in whole groundwater bodies (the ‘Grath’ approach, spatially average then find the trend). We first determined the trends for individual boreholes and then spatially averaged these. This approach preserves information about the spatial distribution of trends within the water body and is less sensitive to ‘missing data’.

A *trend* is the underlying rate of change and is often used to distinguish a long-term tendency from erratic short-term fluctuations (*noise*). Groundwater quality varies with time over various time scales from daily, seasonal and annual, to longer periods as a result of the varying timescales of the input and output functions and the properties of the aquifer. These sources of variation are superimposed on one another and their resolution is a challenging task. The underlying cause of these patterns may reflect changes in land use, fertilizer applications, pollution history or climate. The solute concentration in samples at a single borehole depends on numerous factors: the spatial distribution of nitrate in the region of the borehole as determined by land use history, etc.; the direction of groundwater flow and local groundwater pumping regimes at the time of sampling, and seasonal and climatic effects.

Short-term peak concentrations are a particular problem for water supply utilities. Often such variations are qualitatively understood, but the precise nature of the variability may be difficult to identify and quantify in terms of causal mechanisms. If water quality time series are quite well described by a statistical model that accounts for both trend and seasonal variability, then there are good prospects for predicting future water quality at least in the shorter term.

In many cases, the monitoring itself can introduce its own characteristics which may make it difficult to assess the presence and significance of trends. These include the sampling frequency, amount of missing data, length of monitoring period and the presence of uncontrolled variables, such as a fluctuating and unrecorded pumping history.

Groundwater nitrate concentrations in many UK aquifers are often approaching or have already exceeded statutory limits for drinking water. In 1993, 208 public supply sites were considered for designation as nitrate vulnerable areas or zones (UKWIR, 2002). An overall assessment of nitrate trends from 1945 to 1996 in England and Wales indicated that nitrate was increasing at an average annual rate of  $0.4 \text{ mg L}^{-1}$  (European Environment Agency, 1999). For this and other reasons, widespread monitoring of groundwater nitrate concentrations has been undertaken over a long period by both water utilities and regulatory agencies. There is therefore a large dataset for nitrate, suitable for analysis, which can be used to define past trends as well as to make informed estimates of future concentrations.

Predictions of groundwater quality are needed to help the water industry in the UK's periodic planning process for operational and capital expenditure and to contribute to the provision of the requisite data for implementation of the Water Framework Directive (WFD). The achievement of 'good status' and the reversal of 'significant and sustained' trends in the concentration of pollutants, including nitrate are environmental objectives in Article 4 of the

Water Framework Directive (WFD, 2000). The WFD requires trends to be determined within a given groundwater body but does not discuss in detail how the data should be aggregated to achieve this. However, this has been considered in detail by Grath *et al.* (2002).

Other reasons for studying water quality trends include evaluating the impact of climate change (Burn & Hag Elnur, 2002), the impact of sewage treatment effluent (Worrall & Burt 1999) and for environmental impact assessments (Hipel, 1988).

This paper describes the application of a semi-automated method of trend analysis to a large body of groundwater nitrate time series data and presents an alternative example of how data may be aggregated in order to determine the trend for a groundwater body as a whole.

## **Statistical methods of data analysis**

Many different methods of data analysis have been employed for trend detection and there is a long history of applying statistical methods to water quality data, particularly to surface waters (Berryman *et al.* 1988; Hirsch *et al.*, 1982; Hirsch *et al.*, 1991; Helsel & Hirsch, 1992; Peters, 1996; Esterby, 1996). Much of this work on trend detection has come from the USA, but there is also a very substantial amount of work from elsewhere including Europe (Frapporti, 1994). The statistical methods available for detecting trends in groundwater quality have been developed from a range of applications, many in the area of water resources but also usefully from other areas of environmental science especially meteorology and most recently, 'climate change' (Kundzewicz & Robson, 2000). Many of the classical statistical procedures for analysing time series (e.g. ARIMA methods) depend on a regular sampling period. While missing data can sometimes be accommodated within these schemes, the water quality datasets available for analysis for most boreholes in the UK are so irregular as to preclude these types of analysis.

Trends of interest include monotonic, linear, cyclic (seasonal) and step changes. There are specific statistical tests for each of these (Grath *et al.*, 2001) although their robustness against outliers, missing data and censoring vary. Non-parametric methods tend to have been favoured in the analysis of large datasets from national monitoring programmes (Esterby, 1996) since these methods involve fewer assumptions and are simpler to apply than parametric methods, and are less sensitive to outliers.

There has been widespread use of the Spearman rho and Mann-Kendall tau statistics to test for the presence of monotonic trends (Berryman *et al.*, 1988; Cun & Vilagines, 1997; Burn & Hag Elnur, 2002; Yue & Wang, 2002; Broers & Grift, 2004). Methods based on the

Spearman rank correlation coefficient have also been used (Antonopoulos *et al.*, 2001; Yue & Wang, 2002)

Seasonal effects can be detected by a number of methods: periodograms, Student's t-test (Helsel & Hirsch, 1988), Mann-Whitney rank-sum test, analysis of variance, Kruskal-Wallis test (Harris *et al.*, 1987), periodic functions (Helsel & Hirsch, 1992), seasonal Mann-Kendall (original or modified) (Hipel *et al.*, 1988; Helsel & Hirsch, 1992; Hirsch *et al.*, 1982; Donohue *et al.*, 2001) and spectral analysis (Fleming *et al.*, 2002). Seasonal patterns in groundwater quality are often complicated by annual variability in the length and timing of each season.

The decision whether to use a step change test should be made before looking at the data otherwise bias can be introduced (Helsel & Hirsch, 1992). Various statistical methods are available for locating step changes in a time series (Hirsch, 1988; Hirsch & Gilroy, 1985; Kundzewicz & Robson, 2000). However, one of the underlying assumptions is of serial independence and this is frequently violated (Yue & Wang, 2002). Yue & Wang (2002) discuss a 'pre-whitening' method for removing much of this serial correlation before analysis. Since mixing is a feature of many groundwater systems, step changes are of less importance than the more gradual changes in trend on which we focus here. Therefore we use a general approach for detecting significant changes in trend based on a piecewise linear regression. This includes detecting trend reversals. Reversal of deleterious trends is a requirement of the WFD.

## **Methodology**

We have developed a simple semi-automated methodology for trend estimation of groundwater quality time series (UKWIR, 2003). This has been written for the 'R' working environment working in batch mode (R Development Core Team, 2005). The data are subjected to a series of descriptive and statistical tests to determine the regularity and frequency of sampling, whether the data show a significant linear trend with time, whether there is any seasonality in the data, whether the data show any unusually large deviations from the assumptions made in the statistical tests undertaken, and whether there is any evidence for a change in trend including a trend reversal. We use the powerful, in-built graphical features within R to provide graphical output and summaries of these features. Although we concentrate on nitrate time series in this paper, the methodology is general and can be applied to any parameter, chemical or otherwise.

## Descriptive tests

Descriptive tests include graphical methods and summary statistics, and do not involve formal hypothesis testing or estimation. In some cases, where the quality of the time-series is poor, such tests are all that is possible or appropriate. Data smoothing can be useful in showing the trend in data even when the data are very noisy. Five standard descriptive plots were produced:

**Plot 1:** Raw data – scatterplot of concentration versus date.

**Plot 2:** Regularity of sampling – step plot showing the gap (time difference between successive samples, in days) versus date, annotated with the mean and standard deviation of the gap.

**Plot 3:** Sampling interval – a histogram of the gap.

**Plot 4:** Range of monthly values – box and whisker plot of concentration binned into months (Hipel et al. 1988). Cyclic behaviour on an annual timescale indicates seasonality.

**Plot 5:** Smoothed trend based on a LOESS smoother plotted on top of a raw data plot and annotated with the RSME of the smoothed curve.

These plots provided a rapid view of the amount, range and quality of the data with insights into the regularity of sampling, the importance of outliers, the degree of seasonality and the ‘smoothness’ of the data.

## Statistical tests

Two standard plots showing the results of various statistical tests were produced:

**Plot 6:** Trends test – the principal plot showing the raw data overlain with linear trend lines determined using various regression-based methods (described below). The plot was annotated with potential outliers, trend values, the probability of significant seasonality and the RMSE. If found to be significant, a ‘broken stick’ plot showing a change in trend lines was also included. Where there was significant variation that could not be accounted for by a linear model, a warning of ‘additional’ or unexplained structure was given.

**Plot 7:** Standardised residuals test – scatterplot of standardised residuals versus date based on the seasonal or non-seasonal model; influential points and possible outliers are highlighted.

### *Trend tests*

Three linear regression trend tests were used:

- (i) Ordinary least squares (OLS) linear regression. This was based on the standard `lm()` function within R. The major limitation of this method is that it can be unduly

influenced by the outliers which are common in water quality data. In order to retain simplicity, none of the non-seasonal trend analysis methods used made specific allowance for short-term serial correlation in the data. While this was likely to be present, its influence on longer-term trend analysis was likely to be minor.

- (ii) Robust linear regression using an MM-estimator. A ‘method of moments’ approach was used based on the `r1m()` function within R (MASS package) (Venables & Ripley, 2002). This combines a high breakdown point (i.e. it can tolerate a relatively large proportion of outliers) with the efficiency of maximum likelihood or M-estimators. It is also fairly robust against departures from normality.
- (iii) KT-Sen slope. A non-parametric test for the presence and magnitude of a monotonic trend based on Kendall's tau statistic (Kendall, 1975) and the Sen slope (Sen, 1968) was used (Millard and Neerchal, 2001). Broadly speaking, the Sen slope estimates the trend from the median value of all pairwise slopes. Hirsch *et al.* (1982) demonstrated that the Sen slope is robust against outliers.

A large divergence between these three estimates of the trend indicates that outliers or some other feature are likely to be playing a significant role in determining the slopes and the results were therefore flagged as such. In such circumstances, one of the robust trend estimators (ii or iii above) would be a better estimate than the OLS trend.

The following tests were also carried out to test for possible deviations from the simple linear model.

*(i) Step change in trend test*

A piecewise linear (‘broken stick’) model was estimated by nonlinear regression using the following function:

$$y(t) = a + b \times t + c \times \max(0, t-d) + e(t)$$

where  $y(t)$  is the nitrate concentration as a function of time,  $t$  is the date of the sample in decimal years, and  $d$  is the date at which the break occurs, also in decimal years.  $a$ ,  $b$ ,  $c$  and  $d$  are fitted parameters which are adjusted to minimize the residual sum of squares in  $y(t)$ . The minimization has been achieved numerically using the `nls()` function in R. The position of the break and slope change were therefore automatically determined. Where there was a significant improvement in the overall fit over the OLS fit, the broken stick lines were added to the plot. Only one break was considered.

### *(ii) Seasonality test*

A simple test for seasonality was carried out by comparing the OLS results with those from fitting a linear model of the form:

$$y(t) = a + b \times t + c \times \text{month} + e(t)$$

where  $y(t)$  is the nitrate concentration as a function of time,  $t$  is the date of the sample in decimal years,  $\text{month}$  is the month of the year expressed as a factor and  $a$ ,  $b$  and  $c$  are adjustable parameters. Ordinary regression was used, with the error terms treated as independent and identically distributed. A more rigorous analysis would consider alternatives. The model chosen includes 11 degrees of freedom in the seasonal component. More parsimonious seasonal models might be considered, although a simple sine function would not be expected to describe adequately the seasonal variation in nitrate concentration series.

Two essentially similar statistical tests were carried out to see if there had been a significant improvement in fit using the seasonal model. These two tests were: (i) the AIC (Akaike Information Criterion, Akaike (1981)) for the seasonal model should be less than the AIC for the OLS model, and (ii) the reduction in the sum of squares using the seasonal model should also be significant at  $p < 0.05$  based on an F-test. The larger the difference in AIC and the smaller the  $p$  value, the greater is the degree of seasonality. This model did not impose any particular structure on the type of seasonality other than that the deviations each month were replicated in the same way each year. The use of the AIC penalises the large number of parameters in the seasonal model, thus limiting any effect of over-parameterisation on testing. In some cases, if a more parsimonious seasonal model were used, seasonality might be detected for an example series where the full monthly model indicated no seasonality present.

### *Standardised residuals test*

The residuals should not show any systematic behaviour when plotted against the independent variable. The three points with the largest Cook's distance, the most 'influential' points, were highlighted (Cook, 1977). They often tend to be isolated points near the ends of the series and so exert a large leverage on the fitted line. They often also have large residuals although they may not, especially when a seasonal component is included. Points with a large residual error and a large influence are not necessarily incorrect and should not be automatically removed. However, they do suggest that the linear model is not working well and so should be double-checked.

## **Trend aggregation**

Because the WFD takes a basin or catchment approach, it introduces the concept of calculating trends for a body or group of bodies of groundwater rather than for individual monitoring sites. Since trends in groundwater bodies as a whole cannot be observed directly, some approach has to be taken to averaging and aggregating groundwater quality observations from individual monitoring sites. Grath *et al.* (2001) recommended estimating trends based on the spatially-averaged mean concentration of all sites within a given groundwater body for each given season (quarter, half-year, year), i.e. spatially average first then calculate trends. One feature of this approach (the ‘Grath approach’) is that if there is substantial variation within the body, the presence or absence of measurements from individual boreholes over a substantial period can have a major effect on the estimated seasonal means and therefore the derived trends. Grath *et al.* (2001) assumed that monitoring datasets would be almost fully populated, with missing data being the exception. This is not the case for much historical data in the UK.

An alternative approach is to calculate trends for individual sites within each groundwater body and then to calculate the median trend for all of the sites (the ‘median trend approach’), i.e. calculate trends first then spatially average. This has the advantage that when individual boreholes behave in significantly different ways within a single groundwater body, the nature of the differences are easy to see and could potentially be rationalised, e.g. split into distinct sub-bodies. This information is also important for managing groundwater quality at the borehole and may give an indication of the reason for any changing trends. Analysing each site separately does not attempt to mix data from potentially different situations but delays the overall ‘average’ trend assessment until the second stage. The median trend approach is less sensitive to outliers and missing data (irregular sampling) than the Grath approach.

In terms of determining time trends, it is important to know the magnitude of the spatial variability that can be expected within individual groundwater bodies since this adds ‘noise’ to the data. This will affect the ease and accuracy with which time trends can be assessed and the length of time series required to determine a significant trend. Ultimately this will determine when a given groundwater body convincingly shows a significant upward or downward trend, or a change in trend. Grouping like areas with like areas in some form of sub-body may therefore be beneficial.



## The data

The above method was applied to a large number of UK groundwater nitrate time series datasets which had been provided to BGS both by water supply utilities and the Environment Agency. The sites included both abstraction and observation boreholes. The distribution of the data was biased towards areas in the east of England where problematic concentrations of nitrate have existed for a long time, and where monitoring programmes were consequently most frequently set up. The data available often had been collected as part of a routine monitoring programme but frequently included additional data collected for operational reasons.

The data were systematically checked for possible errors and inconsistencies, such as varying units and small differences in site names. Many sites contain clusters of boreholes or springs. Where two boreholes at a site abstracted water from different aquifers these were classed as separate sites. The data were sorted by date order. Where the series contained multiple analyses for the same day, the median concentration was used to reduce the data to a single observation per day. The within-day variability was small compared to other sources of variability.

A summary of the record span, and the frequency and regularity of sampling is shown in Table 1. Many datasets were far from ideal for our purposes and more than one third of the datasets were rejected following the initial screening for either being too short (span of less than 5 years or fewer than 20 observations) or too irregular.

Almost all the datasets with reasonably regular sampling, as given by a ‘regularity’ (mean gap divided by the standard deviation of the gap) of 1.5 or more, were observation wells which tended to have monthly, quarterly or 6-monthly monitoring schedules. The sampling frequency at the other sites varied widely, from daily to quarterly (Table 2). Many of the public supply boreholes had been sampled relatively frequently, particularly when they were under investigation. A common pattern was for the sampling frequency to show predominantly weekly sampling but with 14, 21 and 28-day intervals also present. Observation wells were usually sampled less frequently.

Many datasets from public supply boreholes had periods where the source had been taken out of supply for renovation of headworks, pump maintenance and borehole relining or because water quality was unacceptable. 127 (40%) of the remaining series had rather irregular sampling schedules with a regularity of less than 0.5.

Of the 352 time-series selected for trend analysis, more than one third (133) had one or more gaps in sampling of at least one year. Specifically:

- 104 had more than one gap of at least one year;
- 29 had a single gap of at least one year.
- 44 had more than one year with very sparse data.

Such gaps in data are commonplace when datasets from public supply boreholes are used for monitoring purposes and severely limit the range of statistical procedures that can be used for trend analysis.

### **Robustness of trend estimation**

The robustness or quality of the trend estimation for each dataset was assessed using four empirical criteria. A ‘good fit’ was when: (i) the standard deviation of the trends from the three methods was  $\leq 0.5 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ ; (ii) the root mean square error (RMSE) of the OLS fit was  $< 7.5 \text{ mg NO}_3 \text{ L}^{-1}$ ; (iii) for datasets showing no seasonality, the ratio of the RMSE’s of the OLS to LOESS fits was  $\leq 1.2$  suggesting that there was no ‘additional structure’; (iv) the coefficient of variation (CV) of the modelled concentrations using the three methods on 1 January 2000 was  $\leq 0.2$ . Those datasets which failed on two or more of the above criteria were removed from further analysis. Criterion (iii) gives an indication of whether a linear model is likely to be an appropriate model or not. After this initial sifting of the data, 309 individual borehole datasets were selected for further study. These were from 191 different sites, i.e. on average, there were 1–2 (or 1.6) boreholes/site. For multi-borehole sites, median trends were used for the summary statistics.

### **Examples of the analysis of some UK groundwater nitrate time series**

The examples below have been divided into datasets which showed ‘seasonal’ or ‘non-seasonal’ behaviour based on the seasonal test described above.

#### **Straightforward non-seasonal**

Regularly-spaced data are preferred for establishing long-term trends but are not essential. Estimates of past trends can be made where there have been data gaps in the earlier part of the record, but future estimates are more difficult. Fig. 1 shows the application of the methodology to typical datasets. The datasets contain a number of points which may be

outliers, but the abundance of data points means that the trends are well-constrained with an OLS RMSE of less than 3 mg NO<sub>3</sub> L<sup>-1</sup>. The three methods gave similar results.

Where series do not exhibit seasonal behaviour, as here, the trends can be estimated with a relatively small number of observations as the underlying model is much simpler than a seasonal model. Fig. 1 (b) shows a series where the sampling frequency was relatively low. At this site only 21 measurements had been taken, at an average of 1.8 per year, but the RMSE of the fit is similar to that in Fig. 1 (a).

At many public water supply boreholes samples are taken very frequently for operational reasons and a large dataset accumulates; Fig. 1 (c) shows such a series. This has an average sample frequency of about 45 per year over the 23-year period shown. The RMSE is similar to the series shown in Fig. 1 (a). In terms of determining the trend, the additional samples provide little extra information.

### **Straightforward seasonal, including nearby water level data**

Where seasonality is significant, a relationship can often be found between the fluctuations in nitrate concentration and water level. We found that high water levels tended to correspond with higher nitrate concentrations, perhaps due to a greater contribution of shallow, polluted water. Water level data, preferably from a site at or close to the source of interest, can therefore be added as an additional independent variable. Since water levels and nitrate concentrations were usually not measured at the same time, the water levels had to be interpolated. This was done by linear interpolation.

Where the water level at the site is strongly influenced by pumping at the site then clearly it is best to try and capture this by monitoring water levels as the actual site. More generally, although water levels tend to vary cyclically throughout the year and, in the UK show few long-term trends, not all years are equally wet. The amplitude of the cyclicity in water levels as well as the peak dates therefore tend to vary to some degree each year. Simply using a sinusoidal function or a 12-month moving average to capture this cyclicity would not be so effective.

We have assumed that there is a linear response between water level and nitrate concentration but this is not necessarily true. Nonlinear responses could be due to:

- Greater hydraulic conductivity at shallow depths resulting in a greater contribution of high nitrate water to the source and a disproportionate effect on water quality.
- Lack of regular saturation of the uppermost horizons allowing high concentrations of nitrate to accumulate and persist.

Observations from a strongly seasonal borehole are shown in Fig. 2. This borehole had been giving non-compliant nitrate concentrations during the spring months since 1995. The water level was at a minimum in November or December and was normally followed by a steep rise in January and generally high levels in the Spring (Fig. 2 (c)). This response was closely matched by nitrate concentrations with little apparent time lag (a few days). The size of the nitrate peaks matched the relative size of the water level peaks quite closely in different years. The *maximum* nitrate concentration was more variable than the *minimum* nitrate concentration from year to year but this variation is largely captured by the variability in the year-to-year water level fluctuations. Any non-random distribution of residuals for the fitted nitrate concentrations would be a cause for concern but inclusion of water level data reduces these considerably. The high correlation between water level and nitrate concentrations was maintained during the dry years of 1996-1997. This strengthens the notion that there is a fundamental (process-based) connection between the two.

The use of this water level information does not completely explain the seasonality observed. For example, in 1997 there was a smaller-than-usual water level peak but no nitrate peak at all. It appears that the source of the higher nitrate at this site is only accessed when the water level rises above 28 m bgl at the observation site. This is an example of a nonlinear response. It would be also be preferable to use water level observations from the same borehole that is being monitored for nitrate but this is not usually feasible when the latter is an operating public supply borehole.

### **Highly influential observations**

Some observations exert an excessive influence on the fitted trends. This can be exacerbated by highly irregular sampling. Removal of these highly influential observations from the dataset would have a disproportionately large influence on the fitted trend. Such observations are not necessarily wrong, but it is important to be aware of them since the fitted trend strongly depends on their correctness. There are standard statistical tests for identifying these highly influential points. We use Cook's distance,  $D$ , with Fox's criterion for significance,  $D > 4 / (n - k - 1)$  where  $n$  is the number of observations and  $k$  is the number of independent variables.

Fig. 3(a) shows a dataset where the OLS slope differs from the other two estimated trends due to a greater sensitivity to a few very influential points. The estimates of the nitrate trend range from +0.24 (OLS estimate, increasing trend) to  $-0.34 \text{ mg L}^{-1} \text{ a}^{-1}$  (robust estimate, decreasing trend). The OLS trend is clearly strongly influenced by the two outlying points (in

1985 and 1999). There is also some evidence of seasonality though the evidence for this is not very strong ( $p=0.011$ ). The high residual error of the OLS trend line ( $RMSE=5.49 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ ) indicates a considerable amount of noise about the trend line and is in part a reflection of some outliers in the 1992–1993 period (one observation was recorded as close to zero, which seems unlikely). The OLS trend line suggests a slight downward trend whereas the other two methods indicate a slight upwards trend.

Fig. 3 (b) shows a more extreme example of this effect, again with some early isolated sample points, particularly in 1984, that exerted a relatively large influence on the OLS trend. The trends estimated by the non-OLS methods are less influenced by these points and agree amongst themselves somewhat better.

Fig. 3(c) shows a dataset where the majority of the data are clustered in a short period, but in this case at one end of the series, and where there is a spread of concentrations at the other end. Again these early points exert a strong influence on the fitted trends, including their direction. Where data clustering is less marked Fig. 3(d), the three methods give similar results. In this case, unlike in the examples above, the influential points do not occur at the ends of the series.

### **Non-linear trends, step changes and excursions**

Many datasets exhibited large, unforeseen excursions from a linear trend. These were generally not due to erroneous data in the sampling or analytical sense, but were caused by some mechanism influencing nitrate concentrations in groundwater in an unusual way. In some cases this could be seen to reflect prolonged droughts or wet years. In other cases, operational changes combined with locally variable landuse could have been responsible.

In some of these cases, the data showed a very distinct break, confirming that the linear model provided a poor explanation of the data. This was confirmed by the high RMSE, and for non-seasonal series, by the presence of a report of ‘additional structure’. In straightforward cases, the ‘broken stick’ analysis indicated the position of the break and gave the ‘before’ and ‘after’ trends. Ultimately it is for the user to make a decision about what has happened to create the additional structure, and how relevant the data are for use in making future predictions. It may for example be best to remove some of the earlier data from the analysis on the basis that this reflects conditions that are unlikely to be relevant in the future.

The following examples describe where significant deviations from a linear trend were observed.

### *Nonlinear trends*

An example of a complex nonlinear trend is shown in Fig. 3(e). Nitrate concentrations show an ‘S’ shaped pattern with evidence of a distinct downturn in 1994–1995 for no immediately apparent reason. As a result, the three methods for estimating trend do not agree very well. The robust estimators are less affected by this downturn. The piecewise regression or ‘broken stick’ fit locates a significant change of trend. It is clear that care would be needed to use any of the three trend lines to extrapolate into the future. Even using the trend shown by the second leg of the broken stick plot would need to be made with care given the ‘unexplained’ nature of the change in slope. A low outlier in 1993 does not significantly affect the calculated trends because of the abundance of other data near this date. Data from the other two boreholes at this site suggest that operational changes in pumping regime may have been responsible for the downturn and the high nitrate water was being drawn to another borehole at the site during 1994 and 1995.

### *Step change*

Interpreting pronounced step-changes of the type shown in Fig. 3(f) requires detailed historical information from the site. There were originally four boreholes operating at the site. The influence of local variations in land use and groundwater flow patterns was important. Nitrate concentrations were lowest in the borehole which abstracted from beneath the adjoining urban area and highest in the borehole which was closest to the arable area on the other side of the pumping station. Downhole logging showed that the shallow groundwater contained a high concentration of nitrate. Two of the four boreholes were shut down, in 1990 and in 1993, because of their high concentrations of nitrate. The impact of these closures can be clearly seen in the time series plot for one of the remaining boreholes. This shows the migration of high nitrate water to this borehole in 1991 as a new capture zone was rapidly established. In this case, it is sensible to exclude the data from before 1992 when assessing the current trend.

### *Excursions*

‘Excursions’ are here defined as cases where data return to a previously defined trend after a substantial departure from this trend. In examples (Fig. 3(g)) all three methods gave a similar long-term trend, but the RMSE for the OLS fit was high. The excursion during 1992–3 was reported as ‘evidence of additional structure’. The piecewise linear regression gave a significantly better fit and is shown by the two broken sticks. A number of series showed excursions or step changes in the winter of 2000–2002 when recharge was unusually high. In

the example shown in Fig. 3(h) there was also evidence of more limited excursions in 1994–1995.

### **Between-borehole but within-site variation**

It might be expected that the variations of nitrate concentrations with time for a cluster of boreholes at one site would be related. However, in some cases, individual boreholes at multi-borehole sites had differing concentrations and trends, and occasionally showed differences in the seasonality of their response as well. This is probably due to one or both of the following factors:

- quality stratification in the aquifer combined with differences in borehole construction or the position of the main inflows;
- variations in landuse/drift cover in the capture zones of individual boreholes as described above.

In either case, a change in pumping regime at one borehole might affect concentrations in adjacent boreholes (Fig. 3(f)). If there are a number of distinct sources of water with different nitrate concentrations within the capture zone, groundwater sampled from each borehole can reflect the proportion of water from these various sources. Within-site variations reflect the depth and relative volumes of inflow, both now and in the past, as well as the land use history of the catchment. Sites with differing concentrations at individual boreholes were predominantly, but not exclusively, located in the Permo-Triassic sandstone.

The three boreholes at the site shown in the example in Fig. 4 maintained a difference in pumped quality over a long period. The most northerly, (a), drew the highest nitrate waters from predominantly unconfined sandstone, whereas the most southerly, (c), drew water from a confined part of the aquifer. Logging in borehole (c) showed that 60% of the water was low-solute water moving up to the pump from below 65 m with only 40% moving down from above. In contrast, borehole (a) derived 65 % of its water from high nitrate water moving down. Borehole (b) had an intermediate composition. Vertical stratification and a different distribution of permeability in each hole were considered to be the primary factors influencing the pumped water quality.

### **Nitrate status of UK groundwaters**

The above methodology was used on a site basis to draw broad conclusions about the groundwater nitrate status across UK aquifers. Specifically, nitrate concentrations were estimated for each site as of 1 January 2000. Where there were several boreholes at a given

site, the median estimated concentration was used to characterise the site taking into account any seasonality in the models used. The estimated median concentration was 37.9 mg NO<sub>3</sub> L<sup>-1</sup>. The median concentrations for the four major aquifers (Chalk, Permo-Triassic sandstone, Jurassic Oolitic Limestone and the Lower Greensand) were somewhat higher than this, namely 42–50 mg NO<sub>3</sub> L<sup>-1</sup> (Table 3).

Results from the small number of sites in the Lincolnshire Limestone indicate somewhat lower concentrations, which could be due to the positive impact of the introduction of agricultural best management practices aimed at reducing nitrate pollution (e.g. MAFF, 1991). These would be expected to be felt sooner in this rapidly responding aquifer. Results from the Permian limestone aquifer give a lower average concentration because a large proportion of the sites studied were in the confined aquifer and therefore subject to denitrification. The variation in this aquifer was greater than for the major aquifers (Table 3).

The highest nitrate concentrations occurred in the areas around the Wash, from the Chalk of south Yorkshire and East Anglia to the Lincolnshire Limestone and the Yorkshire/Nottinghamshire Permo-Triassic sandstone. These broadly correspond to the areas of low effective rainfall combined with a large percentage of arable land which was identified as at a high risk of leading to nitrate pollution (Foster et al., 1986 and Department of the Environment, 1986). Concentrations in the southern Chalk were lower and generally below 50 mg NO<sub>3</sub> L<sup>-1</sup>. There are very few datasets from the western side of England and Wales.

We found that the average trend in groundwater nitrate concentrations was upwards at an overall rate of about 0.34 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup> (Table 4). In the major aquifers (Chalk and Permo-Triassic sandstone), the average was slightly greater (0.40 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup>). For individual aquifers, the steepest median trend was in the Lincolnshire Limestone aquifer (0.96 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup>) and lowest was in the Jurassic Oolitic limestone aquifer (0.05 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup>) (Fig. 5). The data from the Permian limestone aquifer had the greatest range reflecting the contrast between the confined (reducing) and unconfined (aerobic) aquifers. There were more sites with increasing trends than with decreasing trends in all of the aquifers.

These results for the Chalk are broadly consistent with the trend of 0.4 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup> reported by the European Environment Agency (1999) and those of Beeson and Cook (2004) who found a linear increase of 1.3, 0.87 and 0.5 mg NO<sub>3</sub> L<sup>-1</sup> a<sup>-1</sup> for three sites in the Chalk of eastern England. However these authors applied a different model to data from the Lincolnshire Limestone which used two separate trendlines for wet and dry periods and gave a zero overall trend.



Broken stick plots were automatically generated where a statistically significant change in trend was detected. This occurred in 40 out of the 309 sources, 19 suggesting a worsening (increasing) trend and 20 an improving (decreasing) one. In some cases, these involved a trend reversal, i.e. a change in the sign of the trend. For the majority of sources these changes were probably due to relatively short-term changes in water quality or erratic data.

A better estimate could perhaps be obtained by excluding data from confined sites where nitrate concentrations are generally low (e.g.  $<10 \text{ mg NO}_3 \text{ L}^{-1}$ ) and so are likely to reflect some denitrification within the aquifer. In this study, 32 sites had an estimated median concentration on 1 Jan 2000 of  $10 \text{ mg NO}_3 \text{ L}^{-1}$  or less. These were located as follows: 3 in the Permo-Triassic sandstone, 1 in the Chalk, 24 in the Permian limestone and all 5 of the Lias sites. Only the trends for the Permian limestone would be significantly changed by separating the data in this way.

If the measured trends were to continue, by 2015 the median concentration in groundwater will have increased by about  $6 \text{ mg NO}_3 \text{ L}^{-1}$  to  $43.6 \text{ mg NO}_3 \text{ L}^{-1}$ . The highest concentrations are predicted to be in the Lower Greensand ( $58.8 \text{ mg NO}_3 \text{ L}^{-1}$ ) and the lowest in the Permian limestone ( $12.3 \text{ mg NO}_3 \text{ L}^{-1}$ ). The Chalk and the Permo-Triassic sandstone will have average concentrations of  $50.5$  and  $52.6 \text{ mg NO}_3 \text{ L}^{-1}$  respectively. In 2000, 34% of sites exceeded the  $50 \text{ mg L}^{-1}$  drinking water MAC. It is estimated that if present trends continue, 41% of groundwater sources would exceed this  $50 \text{ mg L}^{-1}$  standard by 2015. Of course implementation of the WFD is aimed at reversing these trends.

## **Catchment trends**

Implementation of the WFD has led to the definition of various groundwater bodies in the UK. These bodies usually reflect natural surface water catchment boundaries and do not necessarily reflect the geochemical characteristics of the aquifers. Trends have to be determined for the groundwater body as a whole based on observations at individual boreholes. An example of nitrate data for a number of boreholes within a groundwater body in the north of England is shown in Fig. 6.

The main productive Permian limestone aquifer from which the monitoring boreholes drew water dips from west to east beneath younger confining strata. The area is overlain by impermeable glacial drift deposits, except for small areas in the north central part of the groundwater body. The outcrop area receives recharge of nitrate-rich infiltration from agricultural land. As the groundwater moves down dip beneath the confining material, much of this nitrate is removed, possibly by denitrification as dissolved oxygen is depleted.

Consequently nitrate concentrations varied systematically across the groundwater body. Most of the sampled boreholes were within the aquifer outcrop area where the observed groundwater nitrate concentrations were highest. There was an apparent dividing line between this high-nitrate water and low-nitrate confined water. Some boreholes close to this dividing line showed the greatest seasonality, suggesting the possibility of a seasonal shift in the redox barrier, perhaps in response to the overall abstraction regime in the groundwater body.

The trends in nitrate concentrations also varied systematically across the area (Fig. 6). High nitrate concentrations tended to be decreasing while low concentrations tended to be increasing slightly. There was evidence of some recovery since 2000 at many of the boreholes showing the highest nitrate concentrations. Recent implementation of agricultural control measures may have been starting to have a beneficial impact on nitrate concentrations in the outcrop area. Nevertheless, over the complete period of record, these boreholes showed an upward trend.

The Grath approach to aggregation and trend estimation (average by area then analyse by time) indicates an overall nitrate trend of  $-0.08 \text{ mg L}^{-1} \text{ a}^{-1}$  while the median trend approach gives a trend of  $+0.08 \text{ mg L}^{-1} \text{ a}^{-1}$ . Although these two estimates imply different directions of change, the magnitude is small and both estimates imply that there is actually very little overall trend within this groundwater body. Of course, the overall trend alone does not highlight the significant variation across the groundwater body and gives no indication of how the management response could be most effectively targeted within the groundwater body. The median trend approach (analyse by time then average by area) would rapidly reveal such a situation whereas the Grath approach would not. It would be more appropriate in terms of its chemical status to treat this groundwater body as two or more distinct groundwater bodies (or sub-bodies) rather than as a single uniform entity (WFD, 2000).

## Conclusions

A semi-automated procedure has been successfully used to analyse a large number of UK groundwater quality datasets for nitrate trends. Our approach has been to use standard statistical tests in order to annotate graphical output with quantitative estimates of trends and where appropriate to provide warnings of possible departures from the underlying model assumptions. Once set up, such an approach is readily amenable to scaling-up in terms of the number of boreholes analysed. It can also be readily updated as new data are acquired.

Trends were determined by linear regression but considerable effort was taken to test the validity of the assumptions of the underlying linear model. Tests were included for lack of

linearity, the presence of outliers, for seasonality and for possible breaks in the trend including a reversal of trend. After exclusion of data where trend fitting was deemed to be unsatisfactory, data for 309 boreholes from 191 distinct sites were selected for further study. The results highlight the importance (and efficiency) of regular data collection for the determination of trends.

Many datasets were far from ideal in this respect and more than one third of the datasets were rejected for being too short (span of less than 5 years or fewer than 20 observations), or had a too irregular sample interval. Sometimes the data showed a distinct break indicating that the simple linear model provided a poor explanation of the data. This was revealed by a high RMSE and for non-seasonal series, by the presence of ‘additional structure’. In some cases, a piecewise linear regression was able to locate the position of a break in the trends and to estimate the ‘before’ and ‘after’ trends. 21% of the time series analysed showed a significant improvement in the overall fit when such a break was included and half of these indicated an increase in trend with time. If the determination of trend reversals becomes a priority, then monitoring water quality further ‘upstream’ will provide an earlier indication of changing trends. This could be done by monitoring at a shallower depth, either just below the water table or in the unsaturated zone – indeed the shallow unsaturated zone (1–5 m) provides a record of the most recent outputs from the soil zone and in most cases provides the best guide to future nitrate concentrations at boreholes and springs.

Sometimes there were large and often inexplicable excursions from the underlying trend. Where the observations were derived from a public supply borehole, we suspect that at least some of these result from operational changes at the pumping station – changes that could not be expected to be included in a general-purpose trend analysis procedure. Purpose-built observation boreholes would be expected to suffer less from such disturbances.

In view of these general characteristics of the nitrate time series datasets, we believe that our rather simple linear regression approach is ‘fit for purpose’. We did not consider autocorrelation explicitly (as in ARIMA-type models) but did consider seasonality and where appropriate included the variation in local groundwater levels as an additional explanatory variable. This appeared to account for much of the seasonality and took into account longer timescale variability arising from runs of particularly dry years. We also did not consider heteroscedasticity (departure from normality of the residuals) explicitly but made some effort to identify outliers and other more sustained departures from the trend lines. In many cases, understanding the causes of such departures remains one of the most important priorities for improving trend estimation.

Individual boreholes at multi-borehole sites often had differing concentrations and trends, and occasionally showed differences in the seasonality of their response as well. This could have been due to quality stratification combined with differences in borehole construction, differing main inflows or variations in landuse/drift cover in the capture zones of individual boreholes. Such sites with locally differing concentrations are predominantly located in the Permo-Triassic sandstone.

For the 191 sites, groundwater nitrate concentrations were found to be rising at an average of  $0.34 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ . Average trends were greatest in the Lincolnshire Limestone aquifer ( $0.96 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ ) and lowest in the Permian limestone aquifer ( $0.18 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ ). Average trends for the Chalk and Permo-Triassic sandstone aquifers were  $0.38 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$  and  $0.44 \text{ mg NO}_3 \text{ L}^{-1} \text{ a}^{-1}$ , respectively.

If present trends continue, by 2015 the average concentration will have increased to  $44 \text{ mg NO}_3 \text{ L}^{-1}$ . The highest concentrations are predicted to be in the Lower Greensand ( $59 \text{ mg NO}_3 \text{ L}^{-1}$ ) and the lowest in the Permian limestone ( $12 \text{ mg NO}_3 \text{ L}^{-1}$ ). The Chalk and the Permo-Triassic sandstone will have average concentrations of  $50 \text{ mg NO}_3 \text{ L}^{-1}$  and  $53 \text{ mg NO}_3 \text{ L}^{-1}$  respectively. In 2000, 34% of sites exceeded the  $50 \text{ mg L}^{-1}$  drinking water MAC. It is estimated that if present trends continue, 41% of groundwater sources could exceed this  $50 \text{ mg L}^{-1}$  standard by 2015.

The aggregation approach of Grath *et al.* (2001) for determining trends within groundwater bodies as a whole is sound but could obscure problems when there are large systematic variations in groundwater quality within a groundwater body. It can also be quite sensitive to missing data and is best used where a regular monitoring programme is carried out and there are few missing samples. An alternative 'median trend' approach developed here first determines the trend at individual sites within a groundwater body and then finds the median of these trends. This allows an assessment of the behaviour within the groundwater body as well as giving the overall trend. It is also less sensitive to missing data.

## **Acknowledgements**

This paper is published by permission of the Executive Director of the British Geological Survey. The authors are grateful to their colleagues in BGS, particularly to Alex Gallagher, for providing and compiling the raw data, and to colleagues in the water industry, particularly Sarah Beeson and Meryl Knapp, for provision of updated datasets. The methodology was developed as part of the projects 'Implications of changing groundwater quality for water

resources and the UK water industry' funded by UKWIR, the Environment and BGS and 'Identification and reversal of trends in groundwater pollution' funded by Defra.

## References

- Akaike, H., 1981. Likelihood of a model and information criterion. *Journal of Econometrics*, **16**, 3-14.
- Antonopoulos, V. Z., Papamichail, D. M. & Mitsiou, K. A. 2001. Statistical and trend analysis of water quality and quantity data for the Strymon River in Greece. *Hydrology and Earth System Sciences*. **5**, 679-691.
- Beeson, S. & Cook, M. 2004. Nitrate in groundwater: a water company perspective. *Quarterly Journal of Engineering Geology and Hydrogeology*, **37**, 261-270.
- Berryman, D., Bobée, B., Cluis, D. & Haemmerli, J. 1988. Nonparametric tests for trend detection in water quality time series. *Water Resources Bulletin*, **24**, 545-556.
- Broers, H. P. & van der Grift, B. 2004. Regional monitoring of temporal changes in groundwater quality. *Journal of Hydrology*, **296**, 192-220.
- Burn, D. H. & Hag Elnur, M. A. 2002. Detection of hydrologic trends and variability. *Journal of Hydrology*, **255**, 107-122.
- Cook, R. D. 1977. Detection of influential observation in linear regression. *Technometrics*, **19** 15-18.
- Cun, C. & Vilagines, R. 1997. Time series analysis on chlorides, nitrates, ammonium and dissolved oxygen concentrations in the Seine river near Paris. *Science of the Total Environment*, **208**, 59-69.
- Department of the Environment. 1986. *Nitrates in Water – Report by the Nitrate Co-ordination Group*. Pollution Paper **26**.
- Donohue, R., Davidson, W. A., Peters, N. E., Nelson, S. & Jakowyna, B. 2000. Trends in total phosphorus and total nitrogen concentrations of tributaries to the Swan-Canning Estuary. *Hydrological Processes*, **15**, 2411-2434.
- Esterby, S. R. 1996. Review of methods for the detection and estimation of trends with emphasis on water quality applications. *Hydrological Processes*, **10**, 127-149.
- European Environment Agency. 1999. *Groundwater quality and quantity in Europe*. Environmental Assessment Report 3.
- Fleming, S. W., Lavenue, A. M., Aly, A. H. & Adams, A. 2002. Practical applications of spectral analysis to hydrologic time series. *Hydrological Processes*, **16**, 565-574.

- Foster, S. S. D., Bridge, L. R., Geake, A., Lawrence, A. R. & Parker, J. M. 1986. *The groundwater nitrate problem*. British Geological Survey Hydrogeology Research Report **86/2**.
- Frapporti, G. 1994. *Geochemical and statistical interpretation of the Dutch National Ground Water quality monitoring network*. University of Utrecht, Utrecht, The Netherlands.
- Grath, J., Scheidleder, A., Uhlig, S., Weber, K., Kralik, M., Keimal, T., & Gruber, D. 2001. *The EU Water Framework Directive: Statistical aspects of the identification of groundwater pollution trends, and aggregation of monitoring results. Final Report*. No.41.046/01-IV1/00 and GZ 16 2500/2-I/6/00, Austrian Federal Ministry of Agriculture and Forestry, Environment and Water Management and European Commission, Vienna
- Harris, J., Loftis, J.C. & Montgomery, R.H. 1987. Statistical methods for characterizing ground-water quality. *Ground Water*, **25**, 185-193
- Helsel, D. & Hirsch, R., 1992. *Statistical Methods in Water Resources*. Studies in Environmental Science, 49. Elsevier, Amsterdam.
- Helsel, D.R.& Hirsch, R.M., 1988. Applicability of the T-test for detecting trends in water-quality variables - Discussion. *Water Resources Bulletin*, **24**, 201-204.
- Hipel, K. 1988. Non-parametric approaches to environmental impact assessment. *Water Resources Bulletin*, **24**, 487-492.
- Hipel, K., McLeod, I. & Weiler, R.R. 1988. Data analysis of water quality time series in Lake Eyrie. *Water Resources Bulletin*, **24**, 533-544.
- Hirsch, R.M.1988. Statistical-methods and sampling design for estimating step trends in surface-water quality. *Water Resources Bulletin*, **24**, 493-503
- Hirsch, R.M.& Gilroy, E.J. 1985. Detectability of step trends in the rate of atmospheric deposition of sulfate. *Water Resources Bulletin*, **21**, 773-784.
- Hirsch, R.M., Slack, J.R.& Smith R.A. 1982. Techniques of trend analysis for monthly water-quality data. *Water Resources Research*, **18**, 107-121.
- Hirsch, R.M., Alexander R.B.& Smith, R.A. 1991. Selection of methods for the detection and estimation of trends in water-quality. *Water Resources Research*, **27**, 803-813
- Kendall, M.G. 1975. *Rank correlation methods*. Griffin, London.
- Kundzewicz, Z. & Robson, A. (Editors), *Detecting trend and other changes in hydrological data*. Word Climate Programme, Data and Monitoring, WCDMP-45. World Meteorological Organization, Geneva, (2000) 157 pp.
- MAFF. 1991. *Code of Good Agricultural Practice for the Protection of Water*, MAFF Publications, London.

- Millard, S. P. & Neerchal, N. K. 2001. *Environmental Statistics with S-PLUS*. CRC Press, Boca Raton.
- Peters NE (Editor) 1996. Trends in water quality – Special issue of Hydrological Processes *Hydrological Process*, **10** (2) 125.
- R Development Core Team. 2005. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ([www.R-project.org](http://www.R-project.org)).
- Sen, P.K. 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, **63**, 1379-1389
- UKWIR. 2002. *Implications of changing groundwater quality for water resources and the UK water industry. Phase 1: Review and scoping study*. UK Water Industry Research Report 02/WR/09/3.
- UKWIR. 2003. *Implications of changing groundwater quality for water resources and the UK water industry. Phase 2: Trend detection methodology and improved monitoring and assessment programmes: Main report*. UK Water Industry Research Report 03/WR/09/06.
- Venables, W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S*. Fourth Edn. Springer, New York.
- WFD 2000. *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy*. OJ EC No. L 327 (22.12.2000). (<http://europa.eu.int/eurlex/pri/en/oj/dat/2000/l327/l327000122en00010072.pdf>)
- Worrall, F. & Burt, T.P. 1999. A univariate model of river water nitrate time series. *Journal of Hydrology*, **214**, 74-96.
- Yue, S. & Wang, C.Y. 2002. The influence of serial correlation on the Mann-Whitney test for detecting a shift in median. *Advances in Water Resources*, **25**, 325-333.

**Table 1.** *Summary of period, frequency and regularity of the datasets*

	Minimum	Median	Maximum
Period of data (years)	5	15	43
Number of samples	21	119	1310
Frequency (samples per year)	1.1	8.8	77
Regularity	0.21	0.6	3.3



**Table 2.** *Sampling frequency*

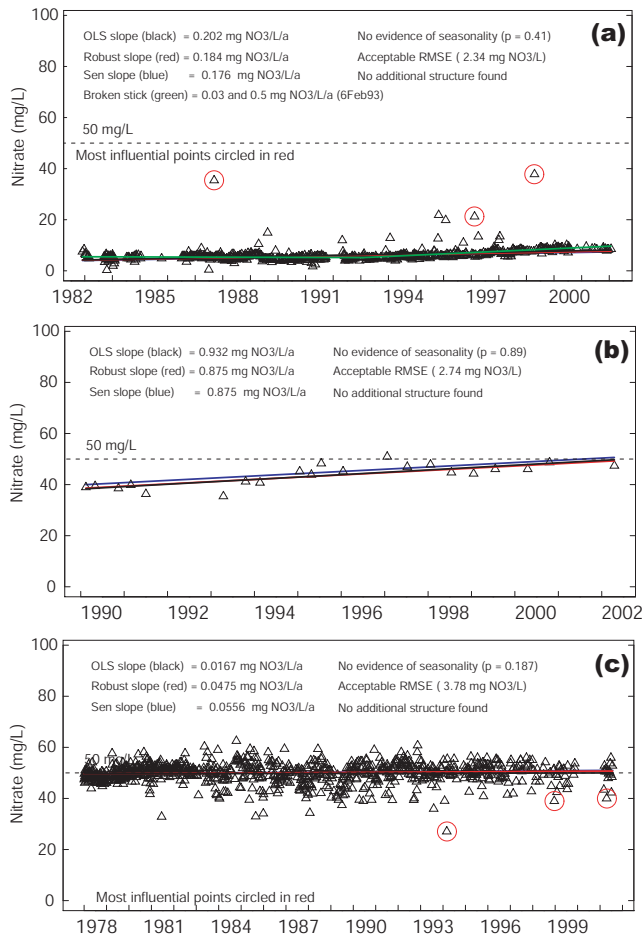
Interval	No of boreholes	Comment
1-10 days	46	Public supplies
1 week	61	Public supplies
1 week and multiples	57	Public supplies
2 weeks	11	Observation wells
1 month	9	Observation wells
2-3 months	16	Observation wells
6 months	46	Observation wells
variable	106	

**Table 3.** Summary of estimated site nitrate concentrations at 1 January 2000 grouped by aquifer

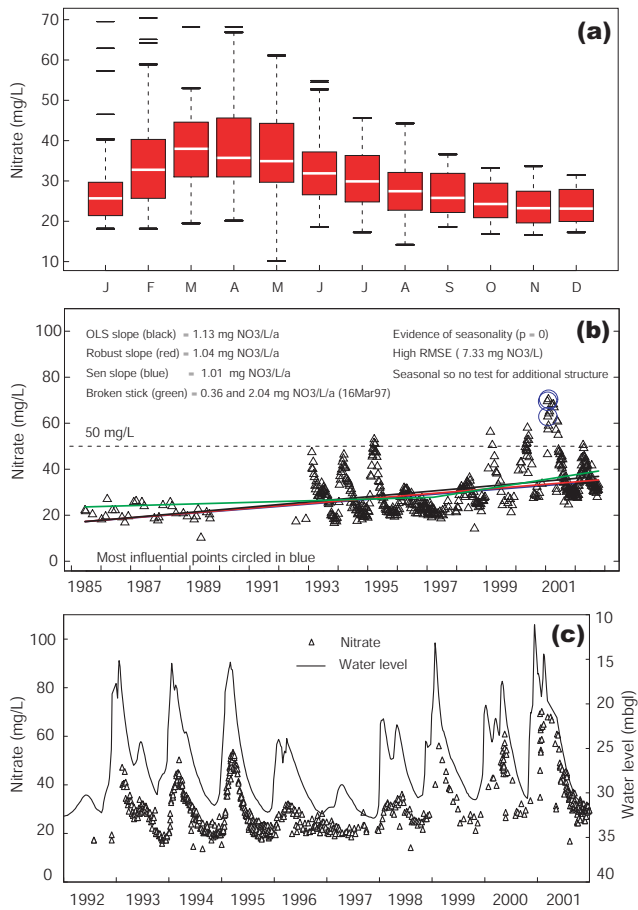
Aquifer	Concentration (mg NO <sub>3</sub> L <sup>-1</sup> )			CV (%)	n
	Median	Max	Min		
Superficial gravels	35.6	48.1	23.2		2
Cretaceous Chalk	42.0	78.8	0.04	37	74
Cretaceous Greensand (Lower)	49.4	60.2	38.6		2
Jurassic oolitic limestone	50.0	70.2	27.7	25	11
Jurassic Lincolnshire Limestone	36.1	124	11.3	78	12
Lias	0.00	0.50	0.00	180	5
Permo-Triassic sandstone	46.3	175	6.45	57	41
Permian limestone	8.2	73.1	0.35	117	43
Carboniferous limestone	32.4				1
All	37.9	175	0.00	66	191

**Table 4.** *Summary of site nitrate trends by aquifer*

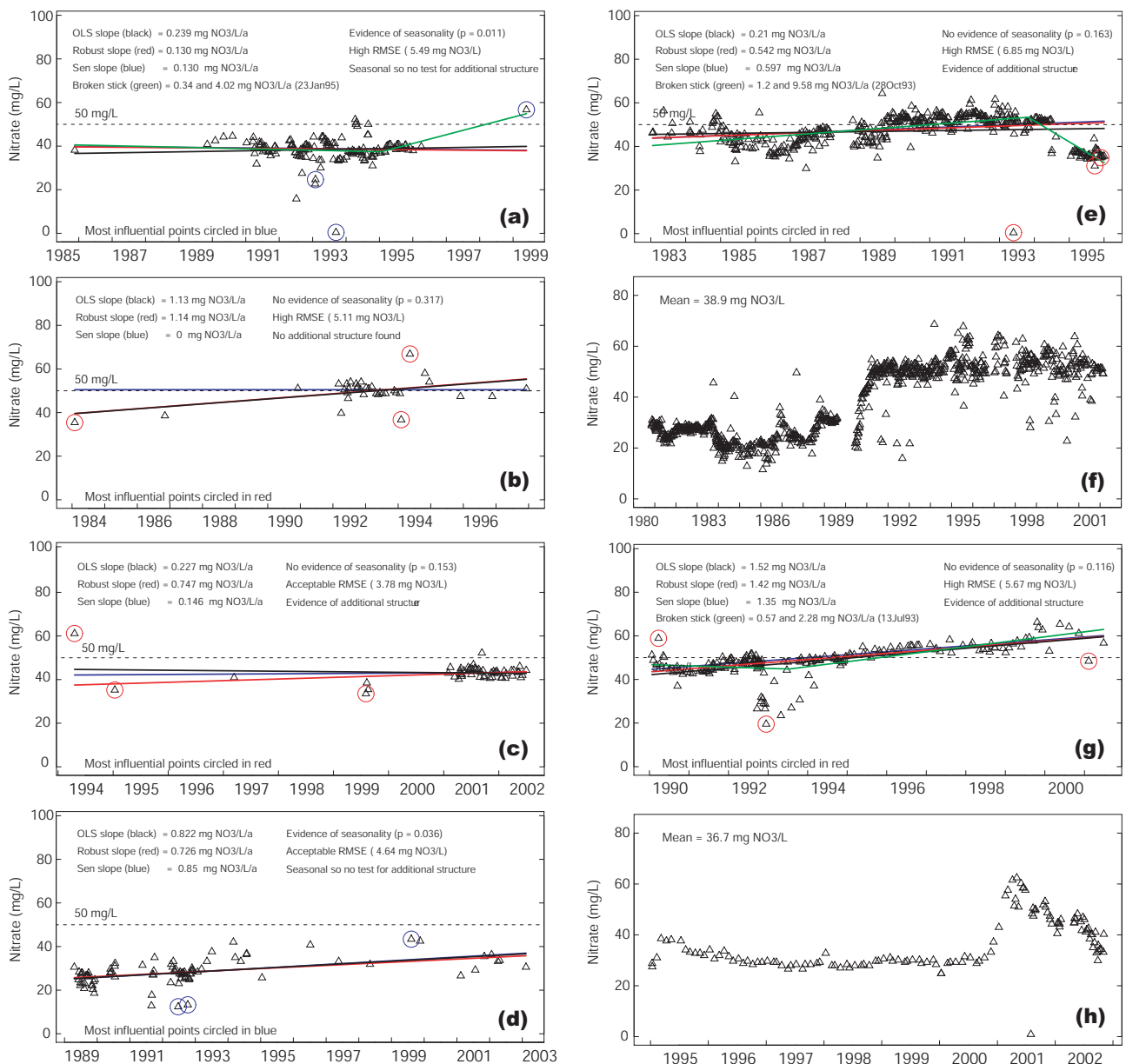
Aquifer	Trend (mg NO <sub>3</sub> L <sup>-1</sup> a <sup>-1</sup> )			SD (mg NO <sub>3</sub> L <sup>-1</sup> a <sup>-1</sup> )	n
	Median	Max	Min		
Superficial gravels	-0.96	0.17	-2.08		2
Cretaceous Chalk	0.38	2.64	-1.37	0.78	74
Cretaceous Greensand (Lower)	0.6	0.62	0.57		2
Jurassic oolitic limestone	0.05	1.69	-1.27	0.86	11
Jurassic Lincolnshire Limestone	0.96	3.44	-0.10	1.27	12
Lias	0.00	0.00	-0.029	0.01	5
Permo-Triassic sandstone	0.44	1.91	-0.95	0.61	41
Permian limestone	0.16	4.12	-1.02	0.83	43
Carboniferous limestone	0.42				1
All	0.34	4.12	-2.08	0.82	191



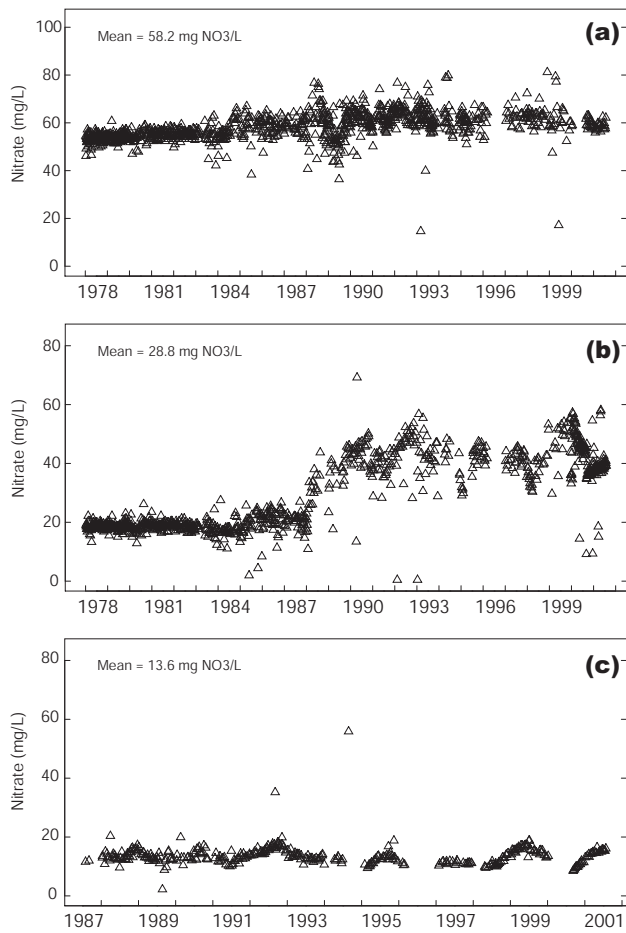
**Fig. 1.** Statistical trend tests (Plot 6) on non-seasonal series with: (a) good data frequency (Sherwood Sandstone aquifer), (b) low data frequency (Chalk aquifer), (c) very high data frequency (Sherwood Sandstone aquifer)



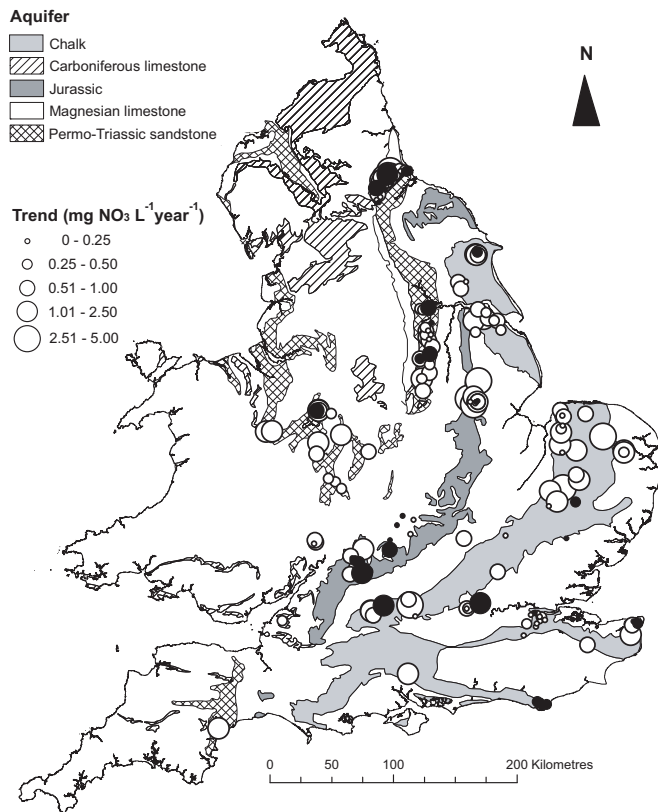
**Fig. 2.** Seasonal data from a site in the Chalk aquifer: (a) range of monthly values (Plot 4), (b) trend fitting (Plot 6), (c) correspondence with water level for part of the data series (1993–2001)



**Fig. 3** Examples of testing of problematic data: (a) and (b) irregular data series from the Chalk aquifer, with clustering in the middle of the series (c) and (d) irregular data series from the Chalk aquifer, with clustering at the end of the series, (e) non-linear data from a Sherwood Sandstone borehole, (f) step change in a Sherwood Sandstone borehole following closure of an adjacent borehole and a consequent change in capture zone (Plot 1), (g) data excursion in the middle of the series, cause unknown, Sherwood Sandstone aquifer, (h) a major excursion in the winter of 2000-2001 in the Chalk of southern England (Plot 1).

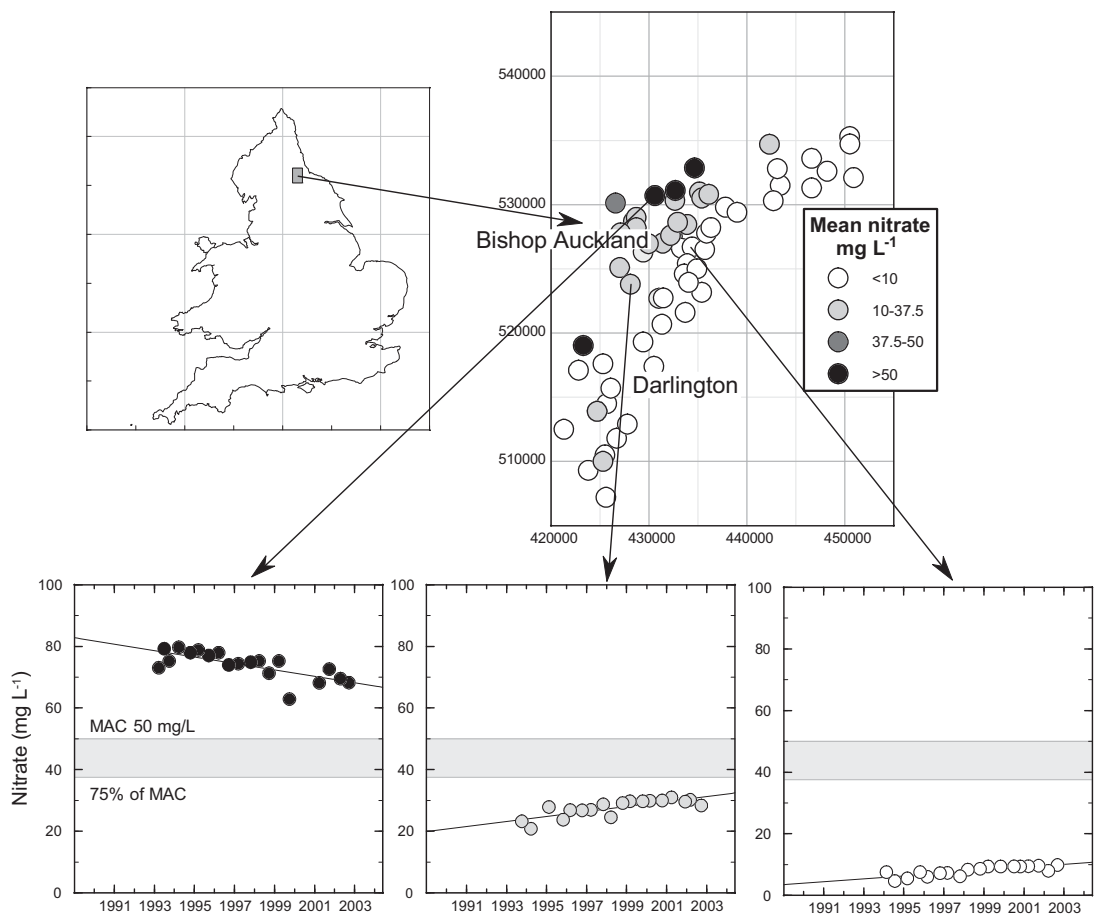


**Fig. 4.** Variation between boreholes at a site, Sherwood Sandstone aquifer: (a) outcrop catchment, (b) intermediate catchment, (c) confined catchment (all Plot 1)



**Fig. 5** Summary of trends in groundwater nitrate concentrations in the major aquifers of England and Wales (positive trends shown as open circles and negative trends as solid circles)





**Fig. 6.** Variation in the mean nitrate concentrations in a groundwater body G7017 (Darlington) and the differing trends in different parts of the body.