



# Ground Water Method Statement for Wales Nitrate Vulnerable Zone Review 2017

**Report Reference: UC11236.04  
January 2016**

---

WRc is an Independent Centre of Excellence for Innovation and Growth. We bring a shared purpose of discovering and delivering new and exciting solutions that enable our clients to meet the challenges of the future. We operate across the Water, Environment, Gas, Waste and Resources sectors.

**RESTRICTION:** This report has the following limited distribution:

External:                      Cyfoeth Naturiol Cymru / Natural Resources Wales

Any enquiries relating to this report should be referred to the Project Manager at the following address:

WRc plc,  
Frankland Road, Blagrove,  
Swindon, Wiltshire, SN5 8YF  
Telephone: + 44 (0) 1793 865000

Website: [www.wrcplc.co.uk](http://www.wrcplc.co.uk)

Follow Us:



# Ground Water Method Statement for Wales

## Nitrate Vulnerable Zone Review 2017

### Authors:



**Andrew Davey**

Senior Consultant  
Catchment Management



**Ben Simpson**

GIS Analyst  
Catchment Management



**Steve Boshier**

Database Analyst  
Catchment Management

**Date:** January 2016

**Report Reference:** UC11236.04

**Project Manager:** Rob Stapleton

**Project No.:** 16480-0

**Client:** Cyfoeth Naturiol Cymru /  
Natural Resources Wales

**Client Manager:** Martyn Evans

### Document History

Version number	Purpose	Issued by	Quality Checks Approved by	Date
V0.1	Interim report issued to NRW to confirm analysis methods for monitoring data.	Andrew Davey	Rob Stapleton	13/10/15
V0.2	Draft final report issued to NRW for comment.	Andrew Davey	Rob Stapleton	23/11/15
V0.3	Final report issued to NRW.	Andrew Davey	Andrew Davey	10/12/15
V0.4	Final report issued to NRW, addressing queries raised by farming unions.	Andrew Davey	Rob Stapleton	08/01/16

© Cyfoeth Naturiol Cymru / Natural Resources Wales 2016

The contents of this document are subject to copyright and all rights are reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted, in any form or by any means electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of Cyfoeth Naturiol Cymru / Natural Resources Wales.

This document has been produced by WRc plc.

# Contents

Glossary .....	1
Executive Summary .....	4
1. Introduction .....	6
1.1 Background .....	6
1.2 Criteria for identifying polluted waters .....	7
1.3 Structure of this report .....	7
2. Overview of the methodology.....	8
2.1 Introduction .....	8
2.2 Step 1 – Analysing groundwater quality monitoring data .....	8
2.3 Step 2 – Estimating groundwater quality between boreholes .....	10
2.4 Step 3 – Modelling nitrate leaching to groundwater .....	10
2.5 Step 4 – Combining the evidence from monitoring and modelling.....	11
2.6 Step 5 – Identifying land draining to polluted waters .....	11
2.7 Step 6 – Using local knowledge to refine draft designations .....	12
3. Analysing groundwater quality monitoring data .....	14
3.1 Source dataset .....	14
3.2 Data processing .....	14
3.3 Data analysis methods .....	17
4. Estimating groundwater quality between boreholes.....	32
4.1 Introduction .....	32
4.2 What is kriging?.....	32
4.3 Methodology.....	33
4.4 Results .....	36
5. Modelling nitrate leaching to groundwater .....	42
5.1 Introduction .....	42
5.2 Agricultural land use data .....	42
5.3 Urban leaching data .....	45
6. Combining the evidence from monitoring and modelling.....	47
6.1 Introduction .....	47
6.2 Introduction to the risk model .....	47
6.3 Risk model components, scores and weights .....	48

6.4	Overall risk assessment.....	54
7.	Comparison with 2013 Review methodology.....	55
7.1	Introduction .....	55
7.2	Refinements to Step 1: Analysing groundwater quality monitoring data .....	55
7.3	Refinements to Step 2: Estimating groundwater quality between boreholes .....	57
7.4	Refinements to Step 3: Modelling nitrate leaching to groundwater .....	57
7.5	Refinements to Step 4: Combining the evidence from monitoring and modelling.....	57
8.	Recommendations.....	58
	References .....	59

## Appendices

Appendix A	Multiple Outlier Test.....	61
Appendix B	Land use model data .....	64

## List of Tables

Table 2.1	Summary of confidence classes.....	9
Table 3.1	Summary of groundwater data processing .....	16
Table 3.2	Minimum data rules for groundwater analysis .....	18
Table 4.1	Variogram parameters for current and future concentration models .....	36
Table 6.1	Components of the risk model .....	47
Table 6.2	Risk scores and weights for national lines of evidence .....	48
Table 6.3	Risk scores and weights for local lines of evidence .....	53
Table 6.4	Summary of categories for final risk score.....	54
Table 7.1	Summary of Summary of methodological refinements for Step 1.....	55
Table 7.2	Summary of Summary of methodological refinements for Step 2.....	57
Table A.1	Critical values (P = 1%) of the $t_{max}$ statistic.....	62

## List of Figures

Figure 3.1	Estimated current 95 <sup>th</sup> percentile nitrate concentration, and comparison to the 11.29 mg N/l (50 mg NO <sub>3</sub> /l) standard (horizontal green dashed line).....	20
Figure 3.2	Historical trend and future forecast of mean nitrate concentration using AntB, and comparison to the 9.71 mg N/l (43 mg NO <sub>3</sub> /l) standard (horizontal blue dashed line) .....	24
Figure 3.3	Historical trend and future forecast of mean nitrate concentration using AntC, and comparison to the 9.71 mg N/l (43 mg NO <sub>3</sub> /l) standard (horizontal blue dashed line) .....	27
Figure 3.4	Historical trend and future forecast of mean nitrate concentration using the 'Mean Concentration' method, and comparison to the 9.71 mg N/l (43 mg NO <sub>3</sub> /l) standard (horizontal blue dashed line).....	29
Figure 4.1	Location of the 232 groundwater monitoring points used in the kriging analysis.....	34
Figure 4.2	Variogram for current (a) and future (b) 95 <sup>th</sup> percentile TIN concentrations.....	38
Figure 4.3	Relationship between observed concentrations at groundwaters monitoring points and predicted (interpolated) concentrations in the associated 1km <sup>2</sup> grid cells for current (a) and future (b) 95 <sup>th</sup> percentile TIN concentrations.....	39
Figure 4.4	Kriged predictions of current (2015) 95 <sup>th</sup> percentile TIN concentration.....	40
Figure 4.5	Kriged predictions of future (2027) 95 <sup>th</sup> percentile TIN concentration.....	41
Figure 5.1	Modelled agricultural concentration leaching to groundwater.....	44
Figure 5.2	Modelled urban nitrate concentration leaching to groundwater .....	46
Figure 6.1	Risk map showing scores from National lines of evidence prior to NRW local review .....	50

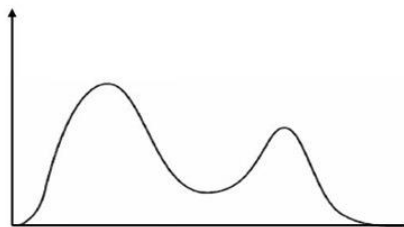
## Glossary

### Autocorrelation

Autocorrelation describes the tendency for observations in a sequence to be correlated with preceding observations. The strength of correlation between successive pairs of measurements is quantified by a lag-1 autocorrelation coefficient, which takes values between -1 (perfect negative correlation) and +1 (perfect positive correlation). Positive autocorrelation is common in water quality monitoring data, where the measured concentration in one sample often gives a good idea of what the measured concentration will be in the next sample.

### Bimodal distribution

The spread of concentration values in a set of water quality samples commonly has a single peak (or 'mode'), with the bulk of samples having intermediate values and a relatively low number of samples having low or high values. Some datasets, however, show a spread of values with two peaks - a cluster of lower concentrations and a second cluster of higher concentrations; this is termed a bimodal distribution.



A bimodal distribution can indicate strongly fluctuating water quality, or indicate that water samples have been collected from two locations with different levels of pollution.

### Coefficient of variation (CoV)

A standardised measure of the degree of variability in a dataset, relative to the mean value of the dataset. The CoV is calculated by dividing the mean by the standard deviation. Monitoring points with higher mean concentrations tend also to show higher variability in concentrations, so the CoV provides a way of comparing variability among monitoring locations with different levels of pollution.

Confidence interval / limits	<p>A confidence interval quantifies uncertainty in the estimate of a parameter by giving a range of values that is likely to include the true (unknown) population parameter. For example, a 90% confidence interval around a sample mean indicates that one can be 90% confident that the true population mean lies within that range. In other words, there is a probability of only 10% or 0.1 that the true mean value lies outside the stated confidence interval.</p> <p>The upper and lower bounds of the confidence interval are called confidence limits.</p>
Kriging	<p>Kriging is a statistical interpolation technique. It is used to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighbourhood of the point (i.e. predict TIN concentrations at unmonitored locations, based on TIN concentrations at monitored locations). The closer the monitoring point is to the location for which concentrations are to be predicted, the greater the weight given to the monitoring point.</p>
(Arithmetic) Mean	<p>This is a measure of the central tendency or 'middle' value of a dataset. It is the sum of the data values divided by the number of observations.</p>
Monte Carlo simulation	<p>Monte Carlo methods are a broad class of computational algorithms that rely on repeated random sampling to simulate a situation and obtain numerical results; typically a simulation is run many times over in order to obtain the probability distribution of an unknown parameter. The name comes from the resemblance of the technique to the act of playing and recording results in a real gambling casino. Monte Carlo methods are often used in physical and mathematical problems when it is difficult or impossible to obtain a closed-form expression, or unfeasible to apply a deterministic algorithm.</p>
MOT	<p>Short for Multiple Outlier Test. By definition, outliers are not representative of conditions at a monitoring location and can therefore be removed legitimately from the dataset to avoid giving a false picture of water quality. Where there appear to be several outliers, the Multiple Outlier Test considers whether these measurements are likely to be true outliers, that ought to be excluded, or whether they are high or low measurements within the normal range of variability, that ought to be included in the analysis.</p>



Outlier	<p>An outlier is an observation in a set of data that is far removed in value from the others in the same data set – i.e. one that has an unusually large or small value compared to others. Outliers are not representative of water quality at the monitored location. For example, they may indicate that a pollution event has occurred or that a water sample has become contaminated after being taken.</p>
Percentile	<p>A percentile is a summary statistic that provides information about the distribution (spread) of values in a defined population. The <math>p^{\text{th}}</math> percentile is the smallest value such that at least <math>p\%</math> of the items in the population are no larger than it. For example, the 95<sup>th</sup> percentile is the value that exceeds 95% of the population and is exceeded by 5% of the population.</p>
Regression analysis	<p>Regression analysis is a process for estimating relationships among variables. It includes a wide range of techniques for analysing several variables, when the focus is on the relationship between a dependent variable (e.g. nitrate concentration) and one or more independent variables (e.g. time).</p>
Statistical significance	<p>In statistics, statistical significance is achieved when a test produces a p-value that is less than a pre-determined significance level. The p-value is the probability of obtaining a result (i.e. a change, difference, relationship) at least as extreme as that observed if there was genuinely no change/difference/relationship in reality. By convention, results are declared statistically significant when the p-value is less than 0.05; this means that there is a less than 5% (or 1 in 20) chance of falsely concluding that there is a change/difference/relationship when in fact there is not. In other words, statistical significance testing provides a safeguard against spurious results that can arise due to random chance.</p>
Standard deviation	<p>A metric used to quantify the amount of variation or dispersion of a set of data values. A standard deviation close to 0 indicates that the data points tend to be very close to the mean of the dataset, while a high standard deviation indicates that the data points are spread out over a wider range of values.</p>

## Executive Summary

The Nitrates Directive (91/676/EEC) is intended to protect waters against nitrate pollution from agricultural sources. Member States are required to identify waters which are or could become polluted by nitrates and to designate these waters and all land draining to them as Nitrate Vulnerable Zones (NVZs). The Directive sets the following criteria for identifying polluted waters:

- Surface freshwaters and groundwaters which contain or could contain, if preventative action is not taken, more than 50 mg NO<sub>3</sub>/l nitrate.

NVZ Reviews were carried out in 1993, 1998, 2002, 2008 and 2013. As part of the 2017 NVZ Review, Natural Resources Wales (NRW) commissioned WRc to undertake a national assessment of nitrate pollution risk based upon available groundwater monitoring and land use data. NRW intends to combine these national lines of evidence with other risk components derived from local expert knowledge to form statements of case for new NVZ designations in 2017.

For consistency and comparability, this Review follows closely the methodology used for the previous (2013) Review of NVZs in England and Wales. However, the present Review provided an opportunity to learn lessons from the 2013 Review and improve how the method was applied and the results were presented.

The six steps in the NVZ Review process are:

1. Analysing groundwater quality monitoring data. Historical trends in water quality at groundwater monitoring points were characterised and then extrapolated to estimate current (2015) and forecast future (2027) concentrations. The results were used to assess whether or not the 95<sup>th</sup> percentile concentration exceeds 50 mg NO<sub>3</sub>/l, now or in the future, and the degree of confidence in the assessment.
2. Estimating groundwater quality between boreholes. Groundwater monitoring points only provide data on nitrate concentrations at specific discrete locations within aquifers and so a statistical interpolation technique (kriging) was used to understand spatial patterns in the groundwater quality and hence estimate nitrate concentrations at unmonitored locations.
3. Modelling nitrate leaching to groundwater. Land use data was used to model the amount of nitrate leaching to groundwater from agricultural land (including atmospheric deposition) and diffuse urban sources. This was done to give a more balanced assessment of the risk of nitrate pollution from agriculture than that which would have come from an assessment of monitoring data alone.
4. Combining the evidence from monitoring and modelling. A numerical risk model was used to integrate the results from Step 2 with those from Step 3. For every 1 km<sup>2</sup> grid cell in Wales, the

risk model assessed the weight of evidence that the nitrate concentration in the groundwater exceeded, or is likely to exceed (by 2027) 50 mg NO<sub>3</sub>/l and that the source of nitrate includes agriculture.

5. Identifying land draining to polluted waters. Professional judgement was used to delineate the recharge area of polluted groundwaters.
6. Using local knowledge to “ground-truth” and refine the draft designations. Local investigations were undertaken by NRW groundwater specialists to understand factors that could affect the risk scores.

This report provides an overview of the methodology by briefly describing each step in turn and then details the approach used to undertake Steps 1 to 4.

# 1. Introduction

## 1.1 Background

The Nitrates Directive (91/676/EEC) is intended to protect waters against nitrate pollution from agricultural sources. Member States are required to identify waters which are or could become polluted by nitrates and to designate these waters and all land draining to them as Nitrate Vulnerable Zones (NVZs). Farmers in designated areas must follow an Action Programme to reduce pollution from agricultural sources of nitrate. The criteria for identifying waters as polluted are established in the Directive, which also sets out monitoring requirements.

Member States are required to review their implementation of the Directive every four years, and to make appropriate amendments to the NVZs and/or the measures in the Action Programme. The most recent (2013) review was undertaken by the Environment Agency (EA) and covered both England and Wales. NVZs in Wales were updated in October 2013 and currently cover around 2.4% of land area. Approximately 750 farm holdings are subject to pollution controls under the Action Programme.

Welsh Government has instructed NRW to undertake a review of NVZ designations during 2015/16 (hereafter referred to as the 2017 Review, because the resulting NVZ designations will come into force on 1<sup>st</sup> January 2017). Draft NVZ designations must be completed by November 2015 to leave time for a public consultation and appeals process in 2016. This is not to be a 'fresh' review; existing NVZ designations do not need to be re-justified and will not be open to appeal. The review will therefore focus on assembling lines of evidence to make a robust, clear and transparent case for designating new NVZs.

As part of the 2017 NVZ Review, NRW commissioned WRc to undertake an up-to-date national assessment of nitrate pollution risk in groundwaters in Wales. This document describes the methodology used to generate and combine the lines of evidence used in the groundwater assessment. (The methodologies used to assess pollution risk in surface waters and eutrophic waters are different and are documented in separate reports.) Subsequent work undertaken by NRW to gather local evidence, identify land draining to polluted waters, and ground-truth the proposed NVZ designations is not described in detail in this report.

## 1.2 Criteria for identifying polluted waters

The Directive sets the following criteria for identifying polluted waters:

- Surface freshwaters which contain or could contain, if preventative action is not taken (i.e. Action Programme measures), more than 50 mg NO<sub>3</sub>/l nitrate.<sup>1</sup>
- Groundwater which contains or could contain, if preventative action is not taken, more than 50 mg NO<sub>3</sub>/l nitrate.
- Natural freshwater lakes, or other freshwater bodies, estuaries, coastal waters and marine waters which are eutrophic or may become so in the near future if preventative action is not taken.

The Directive specifies that the following considerations must be taken into account when applying these criteria:

- the physical and environmental characteristics of the water and land;
- current (scientific) understanding of the behaviour of nitrogen compounds in the environment (water and soil); and
- current understanding of the impact of preventative action.

The Directive requires that at each NVZ review changes and factors unforeseen at the previous review must be taken into account. The periodic nature of reviewing NVZs means that each review necessarily presents a snapshot assessment of nitrate pollution up to the time of the review. This latest review was undertaken using data up to and including 2014.

## 1.3 Structure of this report

The remainder of this report is divided into six Sections. Section 2 provides an overview of the six step groundwater assessment methodology. Sections 3 to 6 provide full details for the first four steps. Finally, Section 8 suggests some possible refinements to the methodology that NRW may wish to consider for future NVZ Reviews. Further technical details are presented in Appendix A and Appendix B.

---

<sup>1</sup> 50 mg/l nitrate as NO<sub>3</sub> is directly equivalent to 11.29 mg/l nitrate as N. Both units are used in this report, depending on the type of data being analysed.

## 2. Overview of the methodology

### 2.1 Introduction

For consistency and comparability, this review follows closely the methodology used for the previous (2013) Review of groundwater NVZs in England and Wales (Environment Agency 2012). However, the present Review provided an opportunity to learn lessons from the 2013 Review and improve how the method was applied and the results were presented. Some refinements were made to the methodology to make better use of the available data and to provide a more rigorous system of checks and balances. These refinements are highlighted and discussed in this document.

The methodology represents a robust and practical approach to the identification of polluted waters and NVZs, and is consistent with assessment approaches adopted for the Water Framework Directive (2000/60/EC).

The 2017 Review methodology for groundwater comprised two main stages: firstly identifying groundwater which contains, or could contain, if preventative action is not taken, more than 50 mg/l of nitrate (Steps 1 to 5); and, secondly, identifying land draining to polluted waters (Step 6). The six steps were:

1. Analysing groundwater quality monitoring data.
2. Estimating groundwater quality between boreholes.
3. Modelling nitrate leaching to groundwater.
4. Combining the evidence from monitoring and modelling.
5. Identifying land draining to polluted waters.
6. Using local knowledge to refine draft designations.

This Section of the report provides an overview of the methodology by briefly describing each step in turn. Sections 3 to 6 give further details of Steps 1 to 4, respectively.

### 2.2 Step 1 – Analysing groundwater quality monitoring data

NRW extracted from its WIMS database all routine water quality monitoring data collected at groundwater monitoring sites in Wales between 1980 and 2015. The raw dataset contained 281 unique monitoring points and 40,204 unique determinand measurements.

The dataset was filtered to remove sites which did not have suitable data. These checks included removing sites that:

- had no data after 1980;
- had less than 6 samples; or
- missing co-ordinates.

Once these filters were imposed, a set of 241 monitoring points remained for analysis. All groundwater monitoring points with sufficient data (i.e. more than six samples), were analysed to determine if:

- the 95<sup>th</sup> percentile of the measured TIN (total inorganic nitrogen (TIN)) concentrations exceeded 50 mg/l in 2015; or,
- the 95<sup>th</sup> percentile of the TIN concentrations was likely to exceed 50 mg/l in 2027.

However, as in the 2013 review, most sites had too few samples to estimate the 95<sup>th</sup> percentile concentration reliably. For these sites, the mean concentration was calculated and an empirical conversion factor was applied to convert the mean concentration to a 95<sup>th</sup> percentile concentration. Further detail on the derivation of this conversion factor is provided in Section 3.3.7.

If the current or future 95<sup>th</sup> percentile nitrate concentration of a groundwater exceeded 50 mg/l with at least 95% confidence, it was deemed to have failed the statistical test. The level of confidence in the assessment was recorded as one of six classes (Table 2.1).

**Table 2.1 Summary of confidence classes**

Confidence class	Description
1. Confident pass	At least 95% confidence that the 95 <sup>th</sup> percentile concentration is ≤50 mg/l
2. Marginal pass	At least 75% confidence that the 95 <sup>th</sup> percentile concentration is ≤50 mg/l
3. Face value pass	At least 50% confidence that the 95 <sup>th</sup> percentile concentration is ≤50 mg/l
4. Face value fail	At least 50% confidence that the 95 <sup>th</sup> percentile concentration is >50 mg/l
5. Marginal fail	At least 75% confidence that the 95 <sup>th</sup> percentile concentration is >50 mg/l
6. Confident fail	At least 95% confidence that the 95 <sup>th</sup> percentile concentration is >50 mg/l

The output from Step 1 was an estimate of the current and future 95<sup>th</sup> percentile TIN concentration at each of the 241 groundwater monitoring points. A total of 16 monitoring

points were estimated to have TIN concentrations above the 50 mg NO<sub>3</sub>/l threshold in 2015, and 19 points were estimated to have TIN concentrations above the threshold in 2027.

Further details of the trend analysis methodology and results are provided in Section 3.

### **2.3 Step 2 – Estimating groundwater quality between boreholes**

Groundwater monitoring is restricted to particular monitoring points where access to groundwater is available through boreholes or where groundwater merges as a spring. In order to estimate nitrate concentrations across the country, a statistical interpolation technique (kriging) was used to understand spatial patterns in the groundwater dataset and predict TIN concentrations at unmonitored locations.

The output from Step 2 was a national map of current (2015) and future (2027) predicted groundwater quality on a 1 km<sup>2</sup> grid. TIN concentrations were predicted to exceed the 50 mg NO<sub>3</sub>/l threshold in 15 grid cells in 2015 and in 20 grid cells in 2027.

Further details of the kriging methodology and results are provided in Section 4.

### **2.4 Step 3 – Modelling nitrate leaching to groundwater**

Land use data was used to model the amount of nitrate leaching to groundwater from agricultural land (including atmospheric deposition) and diffuse urban sources. This was done to give a more balanced assessment of the risk of nitrate pollution from agriculture than that which would have come from an assessment of monitoring data alone. The principle purposes of using the land use data were to:

- identify the significance of the agricultural contribution to any nitrate pollution identified;
- provide further confidence in the conclusions of the statistical analysis of monitoring data; and
- minimise the possibility of designating NVZs due to historic land use practices where long travel times are required for nitrate to reach deep groundwater.

The output from Step 3 was an estimate of the concentration of leached nitrate at the base of the soil zone, for each 1 km<sup>2</sup> in England and Wales.

Further details of the modelling assessment are provided in Section 5.



## 2.5 Step 4 – Combining the evidence from monitoring and modelling

A numerical risk model was used to integrate the results from Step 2 with those from Step 3.

For every 1 km<sup>2</sup> grid cell in Wales, the risk model assessed the weight of evidence that the nitrate concentration in the groundwater exceeded, or is likely to exceed (by 2027) 50 mg NO<sub>3</sub>/l and that the source of nitrate includes agriculture. One of three levels of risk was assigned to each 1 km<sup>2</sup> cell.

- High risk – both the monitoring and modelling assessments agreed that nitrate concentrations exceed, or were likely to exceed, 50mg NO<sub>3</sub>/l, and that agriculture was a significant source of the pollution identified.
- Medium risk – either the monitoring or modelling assessments show that nitrate concentrations exceed, or were likely to exceed, 50 mg NO<sub>3</sub>/l.
- Low risk – both the monitoring and modelling assessments show that nitrate concentrations were not likely to exceed 50 mg NO<sub>3</sub>/l, now or in the future.

Further details of the risk model used to combine the monitoring and modelling assessments are provided in Section 6.

## 2.6 Step 5 – Identifying land draining to polluted waters

Land that is directly above a polluted groundwater does not necessarily drain into it. Similarly, groundwater may receive recharge water from land not directly above it. There are a number of factors affecting the path of water from the surface downwards into a groundwater body including, for example, the presence of impermeable rock layers. Similarly, land that is not directly above a polluted groundwater may drain into it, possibly due to lateral flow within the soil. The surface water designation methodology considers WFD river water body boundaries to be the fundamental unit of water management for NVZs. However, for groundwater there is a need to recognise the factors described above and identify only that land that may contribute pollution rather than designating the whole of the groundwater body.

The following types of information were used to delineate the recharge area of the polluted groundwaters based on professional judgement.

- Solid and drift geology – used to consider geological boundaries such as faults and geological contacts, and the effects of high or low permeability superficial deposits on recharge and groundwater flow.
- Surface water outflow feature (e.g. rivers and lakes) – these features could define a groundwater divide. Where nitrate risk is high on one side of such a feature they can be used to define the recharge area.

- Urban areas – whilst they do not represent a hydraulic boundary, they can be useful as a boundary beyond which there is no agricultural nitrate contributing to the high risk area.
- Groundwater flow lines – used to delineate groundwater bodies within an aquifer. Flow lines are drawn perpendicular to groundwater head contours and can refine recharge areas. This type of boundary is subject to professional judgement and seasonal change. This has only been used when none of the other boundaries are appropriate.
- Risk of solution features – these solution features act as preferential pathways to the aquifer. If the rock at the surface is prone to solution features then it is important that the NVZ is extended to include this area.

## 2.7 Step 6 – Using local knowledge to refine draft designations

Local investigations were undertaken by NRW groundwater specialists to understand factors that could affect the risk scores. The factors that were considered were as follows:

- the influence of point source pollution;
- the accuracy of the identification of pollution in areas between monitoring points using kriging;
- situations where monitored groundwater nitrate concentrations may not be fully representative of real groundwater nitrate concentrations;
- situations where there is a surface water-groundwater interaction where it can be identified that surface water quality is a reasonable indicator of groundwater quality; and
- local evidence showing that denitrification or mixing reduces the nitrate input from agriculture before it reaches groundwater.

Comments and justifications for modifying the final scores were recorded for each of these areas in the database. Various national datasets were used to inform this “ground-truthing” process, including: solid geology, drift geology, drift thickness, drift permeability, risk of solution features, depth of unsaturated zone, groundwater head and mean surface water nitrate concentration estimated by a statistical regression model.

Modifications to the level of risk in the output from Step 5 were made on the basis of detailed local knowledge of the land and/or hydrogeology, for example through:

- Identification of point source pollution – if samples from a monitoring point were suspected of being unduly influenced by point source non-agricultural pollution then the level of risk could be downgraded.
- Understanding of the hydrogeology – the risk score may be increased or decreased depending on the hydrogeological setting. For example:
  - if a monitoring point takes samples from a deep confined aquifer it would not be representative of shallow unconfined groundwater quality above it, in this case the risk score may be decreased; or,
  - if the monitoring data is from a deep unconfined aquifer, with long delays between pollution leaving the surface and being monitored, in this case the risk score in the recharge area could be increased.
- If surface water monitoring is available in areas with infrequent groundwater monitoring then this data could be used to increase the risk score where the data is representative of groundwater quality.
- If the processes of de-nitrification or mixing act in an area to decrease nitrate concentrations before it reaches the groundwater then the risk score may be decreased.

Each 1 km<sup>2</sup> grid cell identified as exceeding 50 mg/l was taken to represent polluted groundwater. All changes to the level of risk were based on sound evidence provided by NRW groundwater specialists and were recorded within the risk model to provide an audit trail. Once all the modifications had been made the risk model was re-run.

## 3. Analysing groundwater quality monitoring data

### 3.1 Source dataset

#### 3.1.1 Water quality monitoring data

NRW extracted from its WIMS database all water quality data collected at groundwater monitoring sites in Wales between 1980 and 2015. The determinand codes extracted were:

- 0111 (Ammoniacal nitrogen as N);
- 0116 (Total oxidised nitrogen as N);
- 0117 (Nitrate as N); and
- 0118 (Nitrite as N).

Determinand 9880 (Nitrate as NO<sub>3</sub>) was absent from the dataset as it is not monitored in Wales.

The dataset contained only routine monitoring data collected by NRW and Dŵr Cymru Welsh Water (DCWW). Unplanned/reactive sampling of pollution incidents and monitoring of waste sites was excluded, as this data was considered to be unrepresentative of normal water quality. To the best of NRW's knowledge, none of the sites sampled water after it had been treated to remove nitrate or sampled water that had been blended from multiple sources.

The raw dataset contained 281 unique monitoring points and 40,204 unique determinand measurements.

### 3.2 Data processing

#### 3.2.1 Monitoring sites

The co-ordinates of each monitoring site were checked for missing or inconsistent eastings and northings. Seventeen sites were identified without locations and referred to NRW; of these, 6 were excluded from the assessment and 11 were supplied with new coordinates.

The following meta-data were tabulated for each monitoring site: unique Site ID; site name; and co-ordinates (eastings and northings).

### 3.2.2 Samples

A Unique Sample ID was created for each sample by combining the UniqueSiteID, and Date (yyyy/mm/dd) fields.

Samples collected before 1<sup>st</sup> January 1980 were excluded because they were not regarded as indicative of recent nitrate trends. Samples collected after 31<sup>st</sup> December 2014 were also excluded because a full calendar year of monitoring data was not available for 2015 and, where water quality fluctuates seasonally, the use of incomplete years could cause current nitrate concentrations to be over- or under-estimated. Thus, up to 35 years of water quality sampling data was analysed for each site.<sup>2</sup>

In the vast majority of cases, a single water sample was taken from a given site on a given day, but multiple samples were occasionally taken. Autosamplers are not usually deployed at groundwater monitoring sites, but multiple samples may be taken from different depths within a borehole, from adjacent boreholes that share a common Site ID, or for some other investigative purpose. The presence of multiple samples means that any analysis of nitrate concentrations is weighted towards days when more samples are taken, so the data were processed to yield a single measurement per site per day. Unfortunately, no unique sample identifier or time was available in the sample data, so it was not possible to link together the measurements of different determinands from the same physical water sample. Instead, where two or more samples had been taken from the same site on the same day, we retained the highest measurement for each determinand and used these to estimate the TIN and TON concentration at that site on that day.

If neither TON nor nitrate was measured (i.e. the sample only had measurements for nitrite and/or ammonium) then the sample was excluded because it provided an incomplete picture of nitrate pollution.

### 3.2.3 Determinands

Most water samples were measured for multiple determinands. These determinand results were combined to calculate the concentration of total inorganic nitrogen (TIN) and total oxidised nitrogen (TON) in each sample, as described below.

Zero readings were removed. These can arise for a variety of reasons, but it cannot be assumed that they represent readings below the Limit of Detection (LOD). Negative readings were also removed as these are clearly erroneous.

---

<sup>2</sup> By comparison, only 25 years (1990-2014) of historical sampling data were used for the surface water NVZ review. A longer, 35 year, time period was necessary for the groundwater assessment because changes in water quality generally occur more slowly in groundwaters than in surface waters.

“Less than” values were treated using the standard EA approach of dividing the recorded concentration by two. “Greater than” values were not very prevalent in the data and were not adjusted (i.e. the concentration measurement was used as reported).

**Total Inorganic Nitrogen (TIN)** was calculated by the following rules, listed in order of declining preference:

- = TON + ammonium;
- = nitrate + nitrite + ammonium;
- = TON;
- = nitrate + nitrite;
- = nitrate + ammonium;
- = nitrate.

**Total Oxidised Nitrogen (TON)** was calculated by the following rules, listed in order of declining preference:

- = TON;
- = nitrate + nitrite;
- = nitrate.

TIN and TON were not calculated if TON and nitrate were both missing, but were calculated if nitrite and/or ammonium were missing because these determinands typically represent only a small proportion of the nitrogen in the water sample.

### 3.2.4 Summary of cleaned dataset

The groundwater data processing is summarised in Table 3.1.

**Table 3.1 Summary of groundwater data processing**

Data processing stage		No .of sites	No. of samples
Raw data	NRW dataset	281	10,925
Data exclusion rules	Sample >2015	2	305
	Sample <1980	0	0
	Zero or negative readings	0	0
	Sites removed with no Grid Reference	6	0
	Nitrite and/or ammonium only	0	843
	Total excluded	8	1,148
Final dataset		<b>273</b>	<b>9,777</b>

### 3.3 Data analysis methods

#### 3.3.1 Overview

Total inorganic nitrogen (TIN) was used to measure the concentration of nitrogenous compounds in water samples. TIN includes nitrate, nitrite and ammonium, of which nitrate is usually the dominant fraction. Ammonium derives from both waste water treatment works (WWTWs) and from agricultural sources and is rapidly oxidised to nitrate under normal riverine conditions. To assess the contribution of ammonium to observed TIN concentrations, a parallel analysis was performed using total oxidised nitrogen (TON), which comprises just nitrate and nitrite.

Each groundwater monitoring site with sufficient data was analysed to determine whether or not:

- the 95<sup>th</sup> percentile TIN concentration currently exceeds 50 mg NO<sub>3</sub>/l; or
- the 95<sup>th</sup> percentile TIN concentration is likely to exceed 50 mg NO<sub>3</sub>/l in the future, assuming no preventative action is taken.

Following the 2013 Review methodology, a site was deemed to have failed the statistical test if the current or future 95<sup>th</sup> percentile nitrate concentration exceeded 50 mg NO<sub>3</sub>/l *with at least 95% confidence*. In practice this meant testing whether the *lower 90% confidence limit* on the 95<sup>th</sup> percentile exceeded 50 mg NO<sub>3</sub>/l. This approach is slightly less stringent than that for surface waters (which uses the best estimate of the 95<sup>th</sup> percentile instead of the lower confidence limit) and amounts to setting the evidence bar slightly higher to reduce the risk of falsely designating NVZs due to high sampling errors. In other words, this approach is a way of taking account of the uncertainty that arises when limited monitoring data is available for analysis.

The year 2027 was chosen as the time horizon for the future assessment because it: (i) is consistent with the time horizon used in the 2013 NVZ Review, (ii) allows a sufficient period of time for pollution mitigation measures to take effect, and (iii) ties in with the Water Framework Directive river basin planning cycle.<sup>3</sup>

The statistical methods used to assess current and future status at each monitoring point depended upon the amount of data available, as set out in Table 3.2. As in the 2013 NVZ Review, most sites had insufficient data to robustly estimate the 95<sup>th</sup> percentile concentration. For these sites, the mean concentration was estimated instead and a conversion factor of 1.16 was applied to convert the mean concentration into an estimate of the lower 90%

---

<sup>3</sup> This time horizon (2027) is further in the future than that used for the surface water NVZ Review (2020) because changes in water quality generally occur more slowly in groundwaters than in surface waters.

confidence limit on the 95<sup>th</sup> percentile (see Section 3.3.7 for details). This is equivalent to testing whether or not the mean TIN concentration exceeds 43 mg NO<sub>3</sub>/l.

The assessment of current status used data from the last six calendar years (i.e. 2009-2014) to estimate the 95<sup>th</sup> percentile concentration or, where less data were available, a statistical extrapolation of the historical (1980-2014) time series was used to predict the mean concentration in mid-2015. The assessment of future status used the same statistical extrapolation methods to predict the mean concentration in mid-2027. To ensure that the results of the analysis were based on sound monitoring evidence, sites were not assessed if they had fewer than six water quality samples in total.

All statistical analyses were conducted using R v. 3.2.0 (R Core Team 2015).

**Table 3.2 Minimum data rules for groundwater analysis**

Rule	Minimum no. of samples	Current status assessment	Future status assessment	Number of sites
At least 24 samples and at least 5 years data for period 2009-2014	24	Weibull estimate of 95 <sup>th</sup> percentile for 2009-2014	Mean concentration <sup>1</sup> in 2027 forecast using 'AntB' method	15
At least 2 samples per year for any 6 years	12	Mean concentration <sup>1</sup> in 2015 forecast using 'AntB' method	Mean concentration <sup>1</sup> in 2027 forecast using 'AntB' method	149
At least 1 sample per year for 6 consecutive years	6	Mean concentration <sup>1</sup> in 2015 forecast using 'AntC' method	Mean concentration <sup>1</sup> in 2027 forecast using 'AntC' method	26
At least 6 samples	6	Mean concentration <sup>1</sup> for period 1980-2014	Mean concentration <sup>1</sup> for period 1980-2014	51
<b>Sites with sufficient data for analysis</b>				<b>241</b>
Less than 6 samples	-	No assessment	No assessment	32
<b>TOTAL</b>				<b>273</b>

<sup>1</sup> Estimates of mean concentration were multiplied by a conversion factor of 1.16 to estimate the lower 90% confidence limit on the 95th percentile. If this exceeded 50 mg NO<sub>3</sub>/l the monitoring point was deemed to have failed the test with high (95%) confidence. This is equivalent to testing whether or not the mean concentration exceeds 43 mg NO<sub>3</sub>/l.

### 3.3.2 Data screening

Prior to analysis, exceptionally high and low concentration measurements at each site were identified using a Multiple Outlier Test (described in Appendix A) and excluded from subsequent calculations. Excluding outliers from the analysis makes the results more robust



and less sensitive to isolated, atypical measurements. A relatively high exclusion threshold was used, however, to minimise the risk of discarding samples that were genuinely representative of water quality at that monitoring point. Further checks were carried out to assess the sensitivity of the results to any high and low concentration measurements remaining in the dataset (see Section 3.3.8 for details).

Next, the number of samples per calendar year was counted and the longest gap in the monitored period was identified. This information enables appropriate input criteria to be set for the AntB analysis.

Finally, each site was also screened for evidence of step-changes in water quality by counting instances where there was more than a 2-fold difference in mean concentration between consecutive pairs of calendar years.

### 3.3.3 Weibull method

The Weibull method was used to estimate the current 95<sup>th</sup> percentile concentration at monitoring points that had at least 24 water quality samples and five years with data between 2009 and 2014.

The Weibull protocol is a robust technique because it does not make any prior assumption about the underlying distribution of the data (i.e. it does not require the data to follow a normal or log-normal distribution). It is also relatively insensitive to outliers and provides an assessment of conditions over a six year period, which averages out year to year variation and makes the results insensitive to random, short-term fluctuations in water quality.

The Weibull method uses the  $r^{\text{th}}$  ranked value within the observation dataset to provide an estimate of the 95<sup>th</sup> percentile, where  $r = 0.95(n + 1)$  and  $n$  is the number of samples. When  $r$  is not an integer,  $r$  is rounded down and up to the nearest whole number, and the corresponding concentration values for these ranks are interpolated to estimate the 95<sup>th</sup> percentile. Conservative 90% confidence intervals were calculated using binomial distribution theory, as described in the EA Codes of Practice for Data Handling (Ellis *et al.*, 1993). If the lower 90% confidence limit exceeded 50 mg NO<sub>3</sub>/l, the monitoring point was deemed to have failed the test.

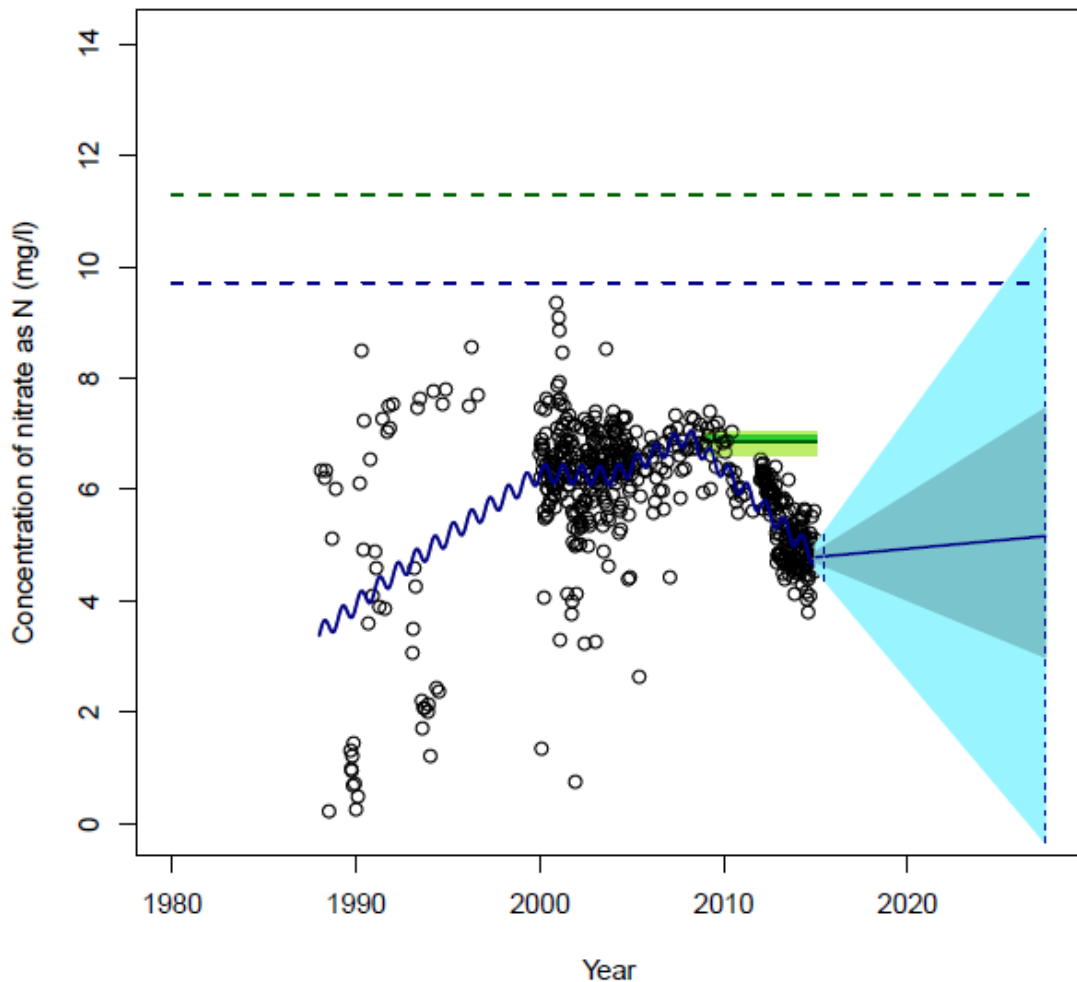
Figure 3.1 shows an example of a Weibull estimate of the current 95<sup>th</sup> percentile concentration. The horizontal green dashed line at 11.29 mg N/l (equivalent to 50 mg NO<sub>3</sub>/l) indicates the threshold value for the 95th percentile concentration.

- The solid green line indicates the 95<sup>th</sup> percentile estimate based on samples data between 2009 and 2014.
- The bright green shaded band indicates the 50% confidence interval around the 95<sup>th</sup> percentile estimate.

- The pale green shaded band indicates the 90% confidence interval around the 95<sup>th</sup> percentile estimate.

In this example, the 95<sup>th</sup> percentile concentration in mid-2015 is estimated to be 6.8 mg N/l with a 90% confidence interval of 6.6-7.0 mg N/l. (Note: the assessment of future nitrate concentrations does not use the Weibull method.)

**Figure 3.1** Estimated current 95<sup>th</sup> percentile nitrate concentration, and comparison to the 11.29 mg N/l (50 mg NO<sub>3</sub>/l) standard (horizontal green dashed line)



### 3.3.4 AntB method

The AntB tool was used to forecast current (mid-2015) and future (mid-2027) mean nitrate concentrations at monitoring points that had at least two samples per year for any six years between 1980 and 2014.

AntB is a simplified version of ANTEATER, a statistical forecasting tool for groundwater nitrate that was developed for the Environment Agency by WRC in the late 1990s. One limitation of ANTEATER is that a substantial amount of monitoring data is needed for it to work satisfactorily. A simplified version of ANTEATER called AntB was therefore developed in 2005. This applies a broadly similar methodology to that used by ANTEATER, but is aimed at datasets that are not sufficiently extensive for a full ANTEATER analysis to be possible.

The operations undertaken by AntB fall into two main parts:

1. trend analysis – AntB characterises historical changes in mean nitrate concentration; and
2. forecasting – AntB uses Monte Carlo simulation to extrapolate the historical trend and forecast future nitrate concentration.

The statistical methodology used in AntB is discussed below under these headings.

### Trend analysis

A multiple regression model was fitted to the data to characterise historical variation in the mean concentration at each monitoring site. The model consisted of two main elements: a linear spline component, which represents inter-annual changes in mean nitrate concentration as a series of  $k$  straight lines connected by  $k-1$  'hinge points', and a seasonal sinusoidal component, which represents within-year changes in mean nitrate concentration.

#### Inter-annual trends

Hinge points were constrained to occur at the end of calendar years to avoid picking up seasonal effects. The following rules were used to strike an appropriate balance between having too many splines (whose estimation would then be excessively vulnerable to high-influence points in the data) and too few splines (which might be unable to reflect genuine historical zigzags).

- So that future changes in nitrate concentration could be simulated using a fixed time step, the splines must be of equal length (i.e. span the same number of calendar years). The exception was where the number of years per spline is not an exact multiple of the total number of years in the data set. In this situation any remainder years are included in the final spline; for example, 4 splines over 18 years would span 4, 4, 4, and 6 years.
- To prevent over-fitting, the maximum allowed number of splines was 10.
- To make the trend results more robust to high-influence data points, the average number of data points per linear spline must be at least 6. This minimum must be met not just on average, but also for the two end splines.

- The splines must be apportioned over the full time period of the monitoring data, (including years with no data) and long enough to span any gaps in the monitoring record. For instance, if the biggest gap in the data extends for Y calendar years, the number of years per spline must be at least Y + 1.

### Seasonal trends

AntB represents seasonality using a simple sinusoidal function:  $\text{Sin}(2\pi X) + \text{Cos}(2\pi X)$ , where X is the decimal fraction of the year (ranging from 0 to 1). If  $B_S$  and  $B_C$  are the resulting slope estimates, the amplitude and phase of the seasonal term are then estimated by:

$$\text{Amplitude} = \sqrt{(B_S^2 + B_C^2)}, \text{ and}$$

$$\text{Phase} = \text{ArcTan}(-B_C/B_S) + c\pi,$$

where c is an adjustment factor set to 0, +1 or -1 as appropriate to make  $\text{Cos}(\text{Phase})$  have the same sign as  $B_S$ .

Consideration was given to extending the analysis to allow the amplitude to vary from year to year, but groundwater data is often too erratic or sparse in individual years for this to work reliably, and so AntB uses a constant amplitude and phase.

### Trend analysis

The relationship between nitrate concentration and time was calibrated using multiple regression, with one appropriately defined dummy variable per spline and two additional predictor variables to represent seasonality. The linear spline and seasonal terms were forcibly included, and carried through into the forecasting phase, regardless of whether or not they were statistically significant. The justification for including them in the latter case is that they are still the best estimates of the historical trends – and if they are not statistically significant the sizes of the effects are likely to be relatively slight.

### Forecasting

Monte Carlo simulation was used to extrapolate the historical trend and forecast future water quality at each site. The simulation assumes that the historical data provides the best (unbiased) evidence as to the likely future direction of water quality and is constructed by selecting slopes at random from a defined population of spline slopes.

Suppose the historical multiple regression model has  $m$  splines. In general, some of these spline slopes will have been estimated less precisely than others – as reflected by their standard errors (SE). The simulation allows for this by defining the weight for spline slope  $j$  to be proportional to  $1/\text{SE}(j)$  and, using these weights, calculating the weighted mean (AvB) and weighted standard deviation (SDB) of the  $m$  spline slopes. If the model contained a just single spline, then SDB was estimated as double the standard error of the spline slope coefficient.

The other inputs to the forecasting are:

- the final spline slope;
- the lag-1 autocorrelation of the spline slopes,  $R_1$ , which defaults to 0 if there are fewer than 4 splines;
- the number of years per spline (except possibly for the last spline); and
- the predicted mean concentration at the end of the monitored period ( $\tau$ ) and its standard error ( $\tau_{SE}$ ).

Taking as its starting point a mean concentration drawn from the Normal distribution  $N(\tau, \tau_{SE})$  and the final spline slope, the simulation generates a projection of how the mean nitrate concentration might change in the future, starting from the year immediately following the end of the monitoring record.

The future projection is constructed by selecting a series of spline slopes at random from the Normal distribution  $N(AvB, SDB)$ , and loosely associating each with the preceding slope to an extent governed by the autocorrelation coefficient  $R_1$ . Specifically, the slope for time step  $i+1$  is generated from the previous time step's slope ( $B_i$ ) and the random slope ( $B_{rand}$ ) as follows:

$$B_{i+1} = (1 - R_1) \times AvB + R_1 \times B_i + \sqrt{(1 - R_1^2)} \times (B_{rand} - AvB).$$

Taking the spline length as the time step, the process is continued until the required forecasting horizon is reached. Starting from mid-2011, for example, if the spline length was four years, a series of four slopes would be generated to predict the mean concentration in mid-2027.

The whole process was repeated 10,000 times to generate a range of possible 'futures'. A variety of summary statistics were harvested from each of the 10,000 simulated time series, including the median, which represents the best estimate of the mean concentration in mid-2015 and mid-2027. The uncertainty in this estimate was quantified by ranking the 10,000 forecasts for that year and determining the 10<sup>th</sup>, 25<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentile values.

Finally, a conversion factor was applied to convert the mean concentration to an estimate of the 95<sup>th</sup> percentile concentration (see Section 3.3.7).

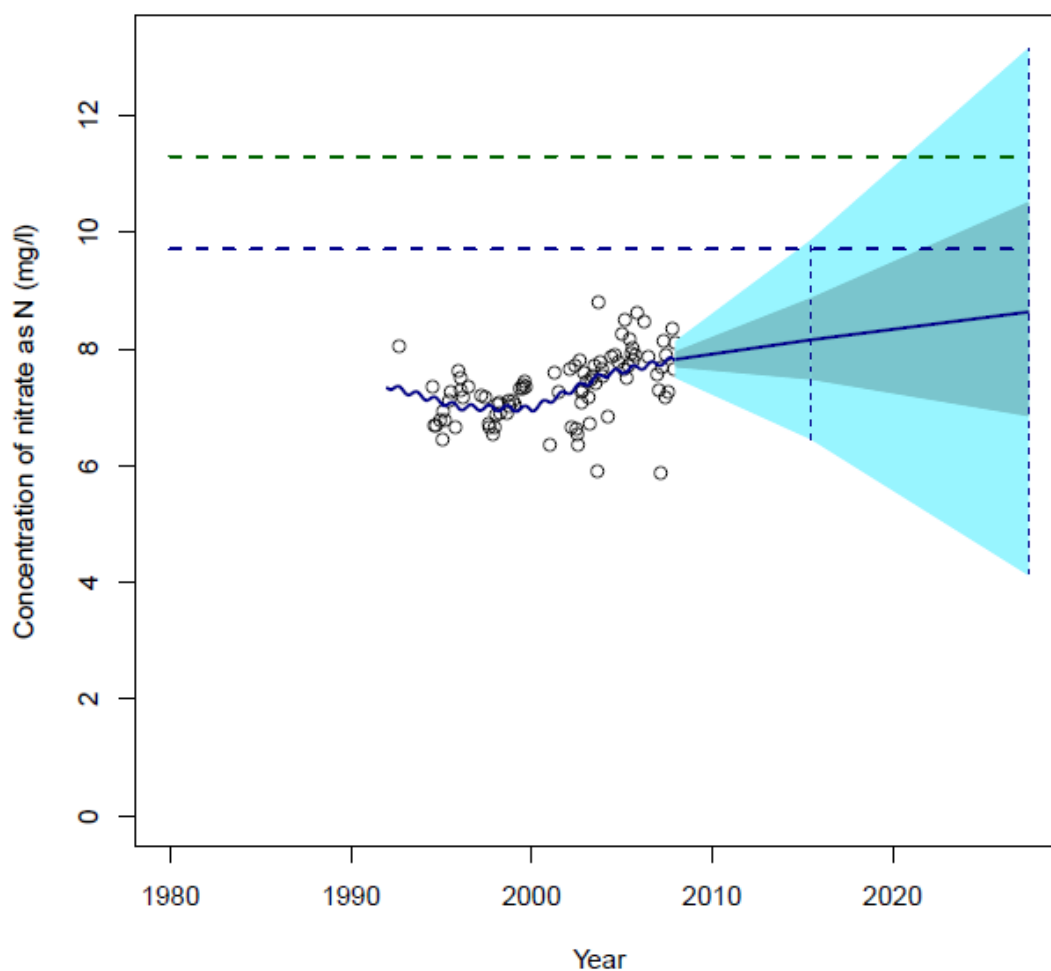
Figure 3.2 shows an example of a fitted AntB regression model and the forecasts generated from it. The horizontal green dashed line at 11.29 mg N/l (equivalent to 50 mg NO<sub>3</sub>/l) indicates the threshold value for the 95<sup>th</sup> percentile concentration and the horizontal blue dashed line at 9.71 mg N/l indicates the corresponding threshold value for the mean concentration (equivalent to 43 mg NO<sub>3</sub>/l).

The solid blue wavy line indicates historical fluctuations in the mean concentration. Beyond the end of the monitoring record in 2007:

- the blue line indicates the mean concentration forecast by Monte Carlo simulation;
- the blue shaded band indicates the 50% confidence interval around the mean estimate; and
- the wider, light blue shaded band indicates the 90% confidence interval around the mean estimate.

In this example, the mean concentration is forecast to be 8.2 mg N/l in mid-2015 (the first vertical dashed line) with a 90% confidence interval of 6.5-9.9 mg N/l, and 8.6 mg N/l in mid-2027 (the second vertical dashed line) with a 90% confidence interval of 4.1-13.1 mg N/l.

**Figure 3.2** Historical trend and future forecast of mean nitrate concentration using AntB, and comparison to the 9.71 mg N/l (43 mg NO<sub>3</sub>/l) standard (horizontal blue dashed line)



### 3.3.5 AntC method

The AntC tool was used to forecast current (mid-2015) and future (mid-2027) mean nitrate concentration at monitoring points that had at least one sample per year for six consecutive years at some point between 1980 and 2014. If there was more than one suitable period within the sample data for a selected site, the most recent period was analysed.

AntC is a simplified version of AntB, which is used when there are insufficient samples to characterise seasonal variation. AntC uses a simplified version of AntB's spline model to characterise historical trends in annual mean concentrations and, like AntB, uses Monte Carlo simulation to extrapolate the historical trend and forecast future nitrate concentrations.

The operations undertaken by AntC fall into two main parts:

1. trend analysis – AntC characterises historical changes in mean nitrate concentration; and
2. forecasting – AntC uses Monte Carlo simulation to extrapolate the historical trend and forecast future nitrate concentration.

The statistical methodology used in AntC is discussed below under these headings.

#### Trend analysis

First, annual mean concentrations ( $Y_1$ ,  $Y_2$ , etc. up to  $Y_k$ ) are calculated for the  $k$  consecutive years of data (i.e. replicate samples taken in the same year are averaged together). The  $k-1$  successive differences between these annual means are then calculated as:

$$D_i = Y_{i+1} - Y_i, \text{ i running from 1 to } k-1.$$

and used to quantify the direction and rate of change in water quality over time.

#### Forecasting

The inputs to the forecasting are:

- the average (AvD) and standard deviation (SDD) of the  $k-1$  successive differences between the annual means; and
- the lag-1 autocorrelation of the successive differences,  $R_1$ .

Given the generally precarious nature of the data set being analysed, a precautionary adjustment to the SDD was made by setting it to the upper 90% confidence limit. Thus:

$$SDD = SDD\sqrt{[(k-2)/\text{ChiSq}d]},$$

where ChiSqd is the 10% point of the chi-squared distribution with k-2 degrees of freedom. For example, with 10 years of data, N = 10, and so the factor by which SDD is scaled up is  $\sqrt{[8/3.49]} = 1.51$ .

Taking as its starting point the mean concentration in the last of the k consecutive years, the simulation generates a projection of how the mean nitrate concentration might change in the future. The future projection is constructed by selecting a series of annual difference values at random from the Normal distribution  $N(\text{AvD}, \text{SDD})$ , and loosely associating each with the preceding difference to an extent governed by the autocorrelation coefficient  $R_1$ . Specifically, the difference for year i+1 is generated from the previous year's difference ( $B_i$ ) and the random difference ( $B_{\text{rand}}$ ) as follows:

$$B_{i+1} = (1 - R_1) \times \text{AvD} + R_1 \times B_i + \sqrt{(1 - R_1^2)} \times (B_{\text{rand}} - \text{AvD}).$$

The process is continued on an annual time step, cumulatively adding the differences together, until the required forecasting horizon is reached. If the last of the consecutive years was 2014, for example, a series of 13 annual differences would be generated to predict the mean concentration in mid-2027.

The whole process was repeated 10,000 times to generate a range of possible 'futures'. A variety of summary statistics were harvested from each of the 10,000 simulated time series, including the median, which represents the best estimate of the mean concentration in mid-2015 and mid-2027. The uncertainty in this estimate was quantified by ranking the 10,000 forecasts for that year and determining the 10<sup>th</sup>, 25<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentile values.

Finally, a conversion factor was applied to convert the mean concentration to an estimate of the 95<sup>th</sup> percentile concentration (see Section 3.3.7).

Figure 3.3 shows an example of a fitted AntC model and the forecasts generated from it. The horizontal green dashed line at 11.29 mg N/l (equivalent to 50 mg NO<sub>3</sub>/l) indicates the threshold value for the 95<sup>th</sup> percentile concentration and the horizontal blue dashed line at 9.71 mg N/l (equivalent to 43 mg NO<sub>3</sub>/l) indicates the corresponding threshold value for the mean concentration.

The observed annual mean concentrations are shown by short horizontal blue lines. Beyond the end of the monitoring record in 2014:

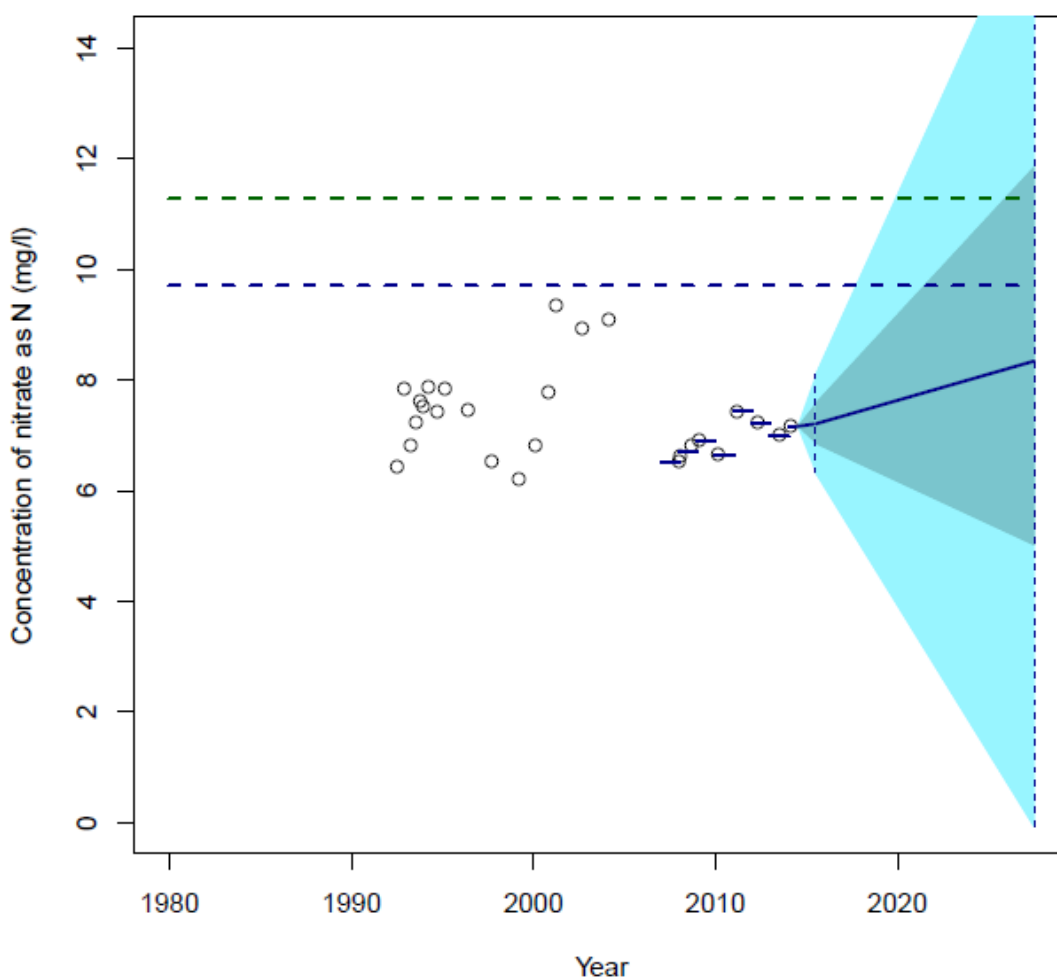
- the solid blue line indicates the mean concentration forecast by Monte Carlo simulation;
- the blue shaded band indicates the 50% confidence interval around the mean estimate; and



- the wider, light blue shaded band indicates the 90% confidence interval around the mean estimate.

In this example, the mean concentration is forecast to be 7.2 mg N/l in mid-2015 (the first vertical dashed line) with a 90% confidence interval of 6.3-8.1 mg N/l, and 8.3 mg N/l in mid-2027 (the second vertical dashed line) with a 90% confidence interval of 0.0-16.8 mg N/l.

**Figure 3.3** Historical trend and future forecast of mean nitrate concentration using AntC, and comparison to the 9.71 mg N/l (43 mg NO<sub>3</sub>/l) standard (horizontal blue dashed line)



### 3.3.6 Mean Concentration method

At sites where there were insufficient data to apply the Weibull, AntB or AntC methods, the mean concentration of all the samples (1980-2014) was calculated and taken as an estimate of the mean concentration in mid-2015 and mid-2027. The standard deviation of historical data was used to calculate the 50% and 90% confidence intervals around the mean estimate.

Finally, a conversion factor was applied to convert the mean concentration to an estimate of the 95<sup>th</sup> percentile concentration (see Section 3.3.7).

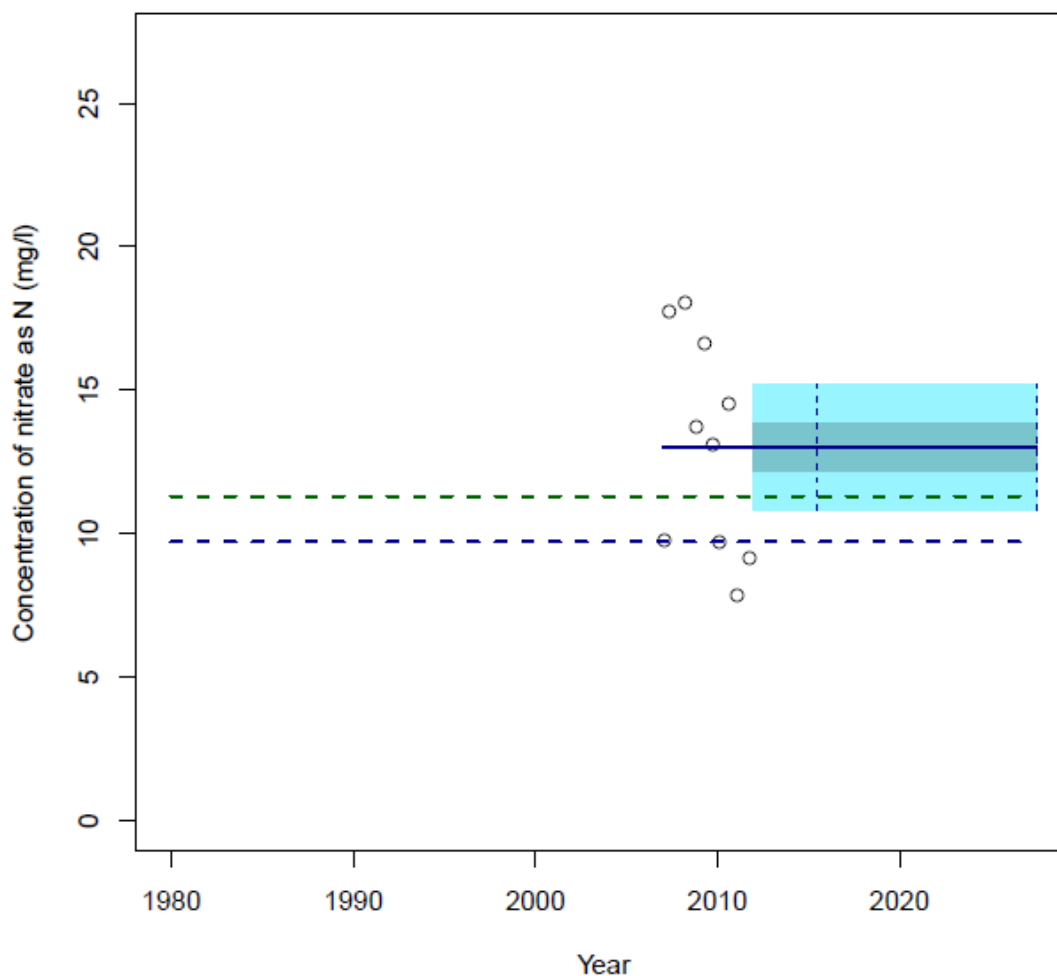
Figure 3.4 shows an example of a monitoring point where the mean concentration of the historical dataset is used to assess current and future water quality. The horizontal green dashed line at 11.29 mg N/l (equivalent to 50 mg NO<sub>3</sub>/l) indicates the threshold value for the 95<sup>th</sup> percentile concentration and the horizontal blue dashed line at 9.71 mg N/l (equivalent to 43 mg NO<sub>3</sub>/l) indicates the corresponding threshold value for the mean concentration.

The observed historical mean concentration is shown by the solid blue horizontal line. Beyond the end of the monitoring record in 2011:

- the solid blue line indicates the future forecast of mean concentration;
- the blue shaded band indicates the 50% confidence interval around the mean estimate; and
- the wider, light blue shaded band indicates the 90% confidence interval around the mean estimate.

In this example, the mean concentration in mid-2015 (the first vertical dashed line) and mid-2027 (the second vertical dashed line) is forecast to be 13.0 mg N/l with a 90% confidence interval of 10.8-15.2 mg N/l.

**Figure 3.4** Historical trend and future forecast of mean nitrate concentration using the 'Mean Concentration' method, and comparison to the 9.71 mg N/l (43 mg NO<sub>3</sub>/l) standard (horizontal blue dashed line)



### 3.3.7 Conversion factor

The 95<sup>th</sup> percentile is the primary statistic used to measure nitrate concentrations in surface waters and groundwater. Unlike surface waters, however, most groundwater monitoring sites lack sufficient data to robustly estimate the 95<sup>th</sup> percentile concentration. Since the kriging process (Section 4) requires the same summary statistic to be used for all monitoring points, a conversion factor was used to make estimates of mean and 95<sup>th</sup> percentile concentrations comparable.

During the 2008 Review the mean concentration was estimated for all monitoring points and the compliance threshold of 50 mg NO<sub>3</sub>/l (equivalent to 11.29 mg N/l) was adjusted to take into account the fact that the mean is always lower than the 95<sup>th</sup> percentile. Based on the sample distribution for all groundwater samples, a threshold of 45 mg NO<sub>3</sub>/l (equivalent to 10.16 mg N/l) was chosen (Defra 2008). The choice of 45 mg NO<sub>3</sub>/l was informed by an

analysis which showed that the mean concentration was on average 5 mg NO<sub>3</sub>/l lower than the lower 90% confidence limit on the 95<sup>th</sup> percentile for monitoring points with large numbers of samples. This provided a comparable level of protection for groundwater to that achieved by the statistical test for surface waters.

During the 2013 Review, a re-analysis of the “45 mg/l rule” using up-to-date monitoring data recommended that a conversion factor of 50/43 (or 1.16) would be appropriate. However, the Methodology Review Group opted to use a conversion factor of 50/45 (or 1.11) for consistency with the 2008 methodology. Sensitivity analysis indicated that the proportion of monitoring points failing the test was similar using 1.11 or 1.16. For consistency with the surface water methodology, the Methodology Review Group took a decision to adopt the 95<sup>th</sup> percentile as the primary statistic and to apply a conversion factor of 1.11 to the observed mean concentrations, so that all results could be compared to a threshold concentration of 50 mg NO<sub>3</sub>/l.

For the 2017 Review, Natural Resources Wales opted to accept the recommendation of the 2013 Review and adopt a conversion factor of 1.16. The central estimate of the mean concentration was therefore multiplied by 1.16 to convert it into an estimate of the lower 90% confidence limit on the 95<sup>th</sup> percentile. If this exceeded 50 mg NO<sub>3</sub>/l the monitoring point was deemed to have failed the test with high (95%) confidence. This is equivalent to testing whether or not the mean concentration exceeds 43 mg NO<sub>3</sub>/l (or 9.71 mg N/l).

### 3.3.8 Quality assurance of results

A series of automatic checks were undertaken to assess the performance of the statistical methods and the accuracy of current and future status assessments. Sites were flagged for more detailed manual checks if they displayed significant data quality issues (e.g. gaps, bi-modal distribution), results that were markedly different from other sites (e.g. very high or low concentrations) or results with high uncertainty.

In addition, the effect of the methodological refinements described in Section 7.2 was assessed by comparing the results of the original (2013 Review) and revised (2017 Review) methods. Sites where current estimates showed discrepancies of more than 10% and 1 mg N/l were reviewed to understand why the results were sensitive to small changes in the methodology.

Visual checks of the trend analysis results were undertaken by WRc for sites:

- with unusually high average or maximum concentrations;
- with unusually high or low forecasts of current and future concentrations;
- with an unusually high coefficient of variation;

- with a statistically significant bimodal or multi-modal distribution;
- with large gaps in the monitoring record, or no recent monitoring data;
- that had an unusually wide confidence interval around the results;
- that failed the current or future status tests despite no samples exceeding the 50 mg NO<sub>3</sub>/l threshold;
- that passed the current and future status tests but had exceedances of the 50 mg NO<sub>3</sub>/l threshold since 2009;
- where there was a large discrepancy between the results generated by the original (2013 Review) and revised (2017 Review) methods.

Checks focused on sites that met more than one criterion for checking, and those where highly variable concentrations made it difficult to characterise accurately the historical water quality trends. In general, the checks confirmed that the statistical methodology had been applied correctly and that the current and future status assessments were reasonable at the vast majority of sites. However, they also revealed a small minority of sites where:

- a small number of high or low measurements were exerting a high degree of influence on the results;
- strongly fluctuating concentrations meant that historical trends could not be characterised adequately, with the result that future forecasts were over/under-estimated or very uncertain; or
- large gaps in the monitoring record made the results particularly sensitive to historical trends or a small number of recent samples.

Where NRW QA checks judged the results not to be a reasonable representation of current and future water quality, the final nitrate risk model was adjusted to override the influence of these monitoring points on the assessment (for details, see risk component #7 in Section 6.3.2). This prevented land being falsely designated as an NVZ as a result of unrepresentative monitoring evidence.

## 4. Estimating groundwater quality between boreholes

### 4.1 Introduction

Monitoring data from groundwater monitoring points only gave groundwater nitrate concentrations at specific discrete locations within aquifers (the exact area of land represented by a monitoring point will depend on the volume and depth of the abstraction and geology). In order to estimate nitrate concentrations across the country and use the methodology at a 1 km<sup>2</sup> resolution, a statistical interpolation technique (kriging) was used to understand spatial patterns in the groundwater dataset and predict nitrate concentrations at unmonitored locations.

The advantages of this approach are that it (i) makes full use of the information contained in the monitoring dataset, and (ii) generates an evidence stream with greater spatial coverage and spatial resolution that can potentially improve the delineation of NVZ boundaries. The disadvantages are that the use of a national analysis to integrate data from multiple monitoring points provides limited ability to take account of local groundwater characteristics or to represent very localised geological conditions. However, the use of expert review with local staff is intended to provide an opportunity to refine the information and it is important to remember that the method is about developing a conceptual model of an area. The kriged concentration data are an important line of evidence in building that understanding but they are only one line of evidence.

### 4.2 What is kriging?

Kriging is a well-established and widely used geostatistical interpolation technique. In essence it takes a set of point measurements, characterises spatial patterns in the measured variable and uses this information to estimate values of the measured variable at all unmonitored locations. The result is a two-dimensional map of the measured variable.

Kriging works by quantifying the spatial correlation between pairs of measurements. Measurements taken close together in space will tend to be more similar than those taken from locations further apart, but the rate of change in similarity with distance will be specific to each dataset. Kriging uses the entire dataset to characterise the relationship between pairs of measurements with different degrees of separation, in the form of a variogram.

These spatial relationships are then used to estimate the measured variable at unmonitored locations from the values observed at surrounding locations. This is achieved by taking a weighted average of the measurements surrounding the unmonitored location. Decreasing weight is given to measurements further away from the unmonitored location, as determined

by the spatial correlation relationship. This interpolation process is repeated for every cell of a grid to build up a complete two-dimensional map.

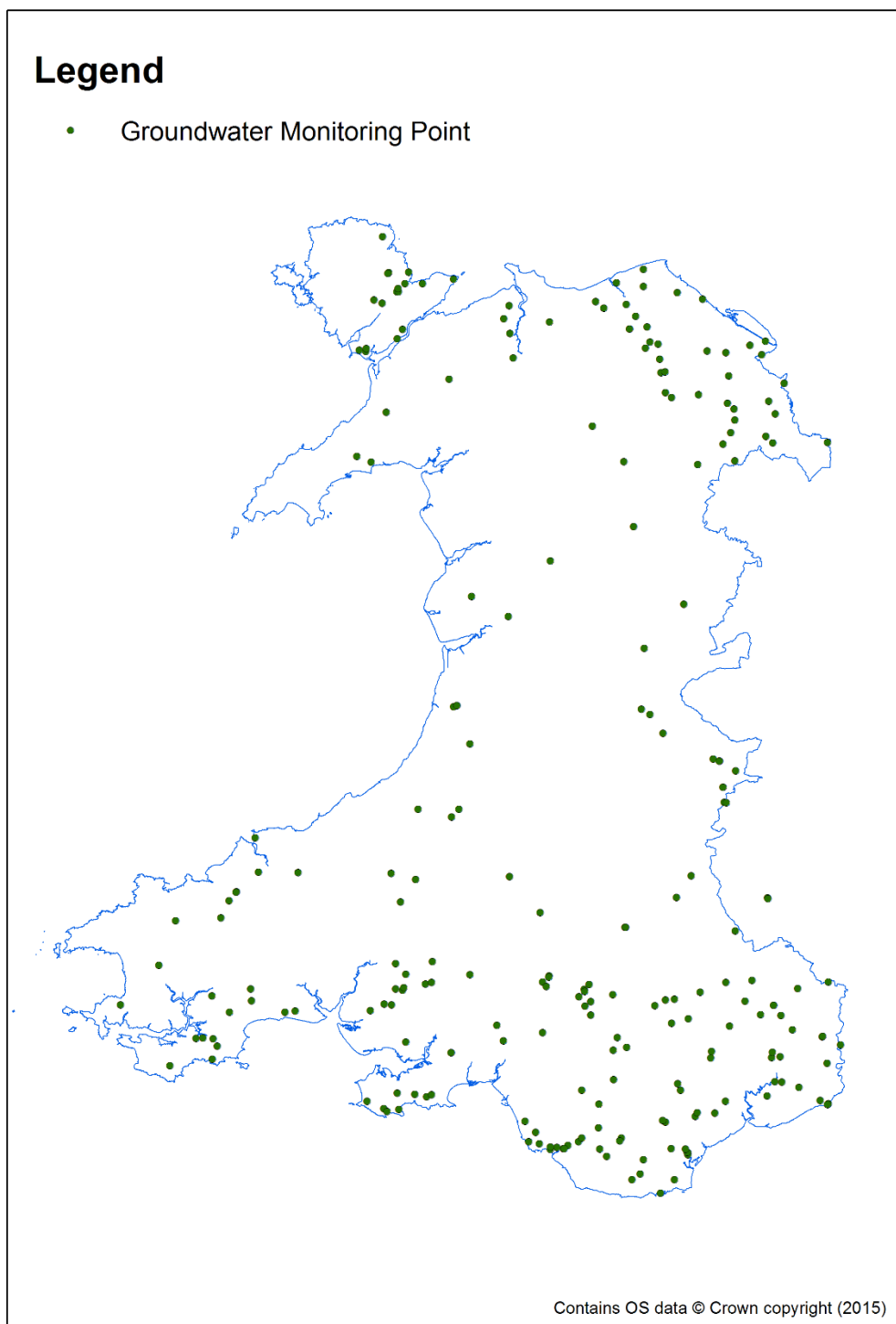
### 4.3 Methodology

Trend analysis was used to estimate the current (2015) and future (2027) TIN concentrations at each of 241 monitoring points, as detailed in Section 3.3. As it is not possible to use different types of values within the kriging process (e.g. a 95<sup>th</sup> percentile concentrations at some points and a mean concentrations at others), the lower 90% confidence limit on the 95<sup>th</sup> percentile was used to characterise water quality at each monitoring point (see Section 3.3.7 for details of how this was calculated). Using a consistent summary statistic in this way allows the kriging analysis to compare TIN concentrations among boreholes.

At locations where a downward historical trend in nitrate concentrations gave rise to predictions of negative concentrations in 2015 or 2027, these results were set to 0.1 mg N/l. A power transformation was applied to make the data approximately normally distributed and to reduce the influence of very high concentration estimates at a small number of monitoring points. The power coefficients used to transform the data were optimised to make each dataset as normally distributed as possible (power coefficient  $\lambda = 0.13$  for current concentrations and 0.03 for future concentrations). After kriging, a reverse transformation was used to convert the results back to mg N/l.

All monitoring points had valid x,y co-ordinates, but it was not possible to verify that every monitoring point was correctly located (this will be checked later as part of the local ground-truthing process). Most co-ordinates were only accurate to the nearest 100 m so, where two or more monitoring points were within the same 100m x 100m square, the monitoring point with the highest TIN concentration was used in the analysis, and results for the other monitoring points were excluded. Of the 241 monitoring points for which 95<sup>th</sup> percentile TIN concentrations had been estimated nine were excluded due to duplication within a 100 m square, leaving 232 for the kriging analysis. These monitoring points were clustered in the main aquifers, particularly in north and south Wales (Figure 4.1).

Figure 4.1 Location of the 232 groundwater monitoring points used in the kriging analysis



Separate variograms were calibrated for current (2015) and future (2027) TIN concentrations. Although TIN concentrations were predicted more precisely at some monitoring points than others (because some points have more complete, comprehensive and up-to-date sampling



data than others), each monitoring point was given an equal weighting in the analysis. Exploratory analyses carried out as part of the 2013 review provided no clear evidence for anisotropy, and so a simple, omnidirectional variogram model was produced in which the spatial correlation structure is the same in all directions.

Using the observed data and the calibrated variogram models, kriging was used to predict current and future TIN concentrations at the centroid of each of 21,819 1 km<sup>2</sup> grid squares covering the whole of Wales. These point predictions were applied to the entire grid square – i.e. predicted concentrations were assumed not to vary within a grid square.

Various models (Gaussian, exponential, spherical, double-exponential and exponential-spherical) were fitted to the variograms with the goal of maximising the level of agreement between the predicted (interpolated) values and the observed nitrate concentrations at individual monitoring points. There were at least two reasons to expect that the variograms should have a high nugget variance<sup>4</sup>, namely:

1. the relatively high level of sampling error in estimates of TIN concentrations at individual monitoring points; and
2. the tendency for neighbouring boreholes sampling groundwater at different depths to record very different TIN concentrations.

Including a nugget variance in the variogram models, however, resulted in an unacceptably high degree of smoothing when the data were kriged, meaning that localised peaks in nitrate concentration were consistently under-estimated. For this reason, the nugget variance was set to zero.

The relative performance of the models was assessed in two ways.

1. The observed nitrate concentrations at individual monitoring points were plotted against the predicted values for the 1 km<sup>2</sup> grid cells containing those points. The strength of agreement was quantified by calculating the percentage of monitoring points for which the predicted and observed values were both the same side of the 11.29 mg N/l (equivalent to 50 mg NO<sub>3</sub>/l) threshold.
2. To guard against over-fitting, a 10-fold cross-validation procedure was used to assess the ability of the kriging to predict nitrate concentrations at unmonitored locations. This

---

<sup>4</sup> The nugget variance is the y-intercept of the variogram – i.e. the variance at lag zero. It represents the small-scale variability in the data, which comprises genuine variation in nitrate concentrations (e.g. variation in water quality between neighbouring deep and shallow boreholes), and sampling and measurements error (i.e. the error in the estimated 95<sup>th</sup> percentile concentration at each monitoring point).

involved randomly removing 10% of the monitoring points from the dataset, re-running the kriging algorithm to predict nitrate concentrations at those locations, and then repeating the process 10 times to obtain a prediction for every monitored grid cell. The strength of association between the predicted and observed concentrations was measured by the root mean square error (RSME).

All statistical analyses were conducted using R v.3.2.0 (R Core Team 2015).

#### 4.4 Results

A nested exponential-spherical model with zero nugget variance was selected as the preferred variogram model for current and future 95<sup>th</sup> percentile TIN concentrations. Figure 4.2 presents the variograms, showing how the (semi-)variance in TIN concentrations increases with increasing separation (lag) between the monitoring points concerned. Details of the calibrated variograms are provided in Table 4.1.

**Table 4.1 Variogram parameters for current and future concentration models**

Model component	Current concentration	Future concentration
Exponential – Partial sill	0.15	0.002
Exponential – Effective range	1,500 m	1,500 m
Spherical – Partial sill	0.02	0.0008
Spherical – Range	30,000 m	50,000 m

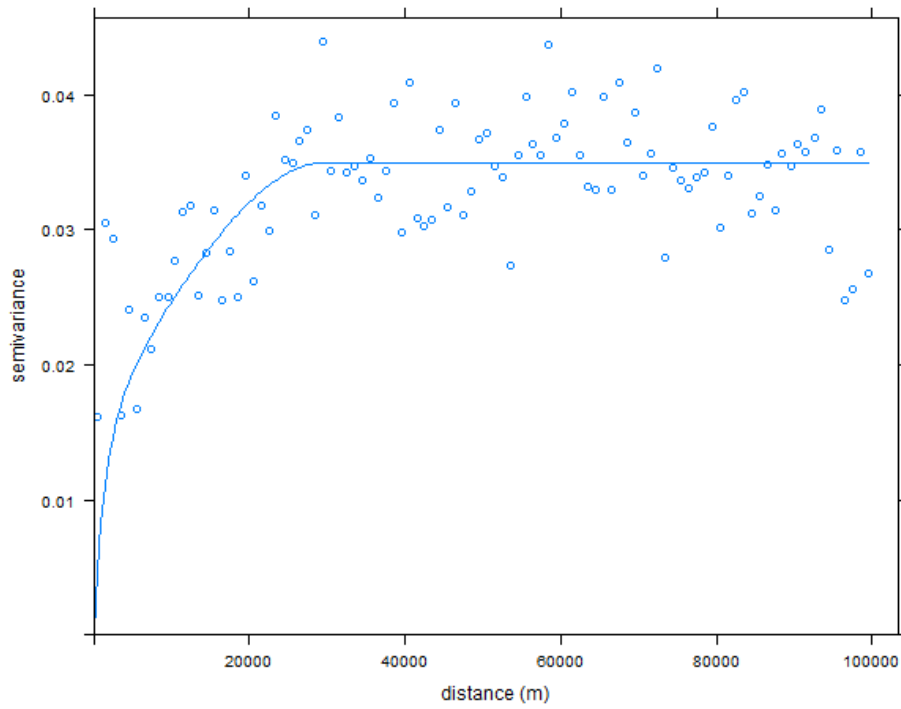
When used for kriging, the exponential-spherical models produced the strongest agreement between the predicted values and the observed nitrate concentrations at individual monitoring points (Figure 4.3). Although high nitrate concentrations tended to be under-estimated, the predicted (interpolated) values correctly classified monitored 1 km<sup>2</sup> grid cells as polluted ( $\geq 11.29$  mg N/l) or not polluted ( $< 11.29$  mg N/l) in 96% of cases. Cross-validation confirmed that the predictive ability of the exponential-spherical model was comparable to alternative models. Occasionally, the concentration predicted in a 1 km<sup>2</sup> grid cell exceeded the concentration observed at a monitoring point within the same cell; manual checks indicated that this was due to the presence of two or more monitoring points in the same cell with contrasting nitrate concentrations, with the interpolated value falling between the highest and lowest observed values.

Figure 4.4 and Figure 4.5 show the final, interpolated maps of current and future TIN concentrations. Across the whole of Wales, TIN concentrations were predicted to exceed the 11.29 mg N/l threshold in 15 grid cells in 2015 and in 20 grid cells in 2027. Where high interpolated concentrations were a consequence of uncertain trend analysis results, the final nitrate risk model was adjusted to override the influence of these monitoring points on the assessment (for details, see risk component #7 in Section 6.3.2). This prevented land being falsely designated as an NVZ as a result of unrepresentative monitoring evidence.

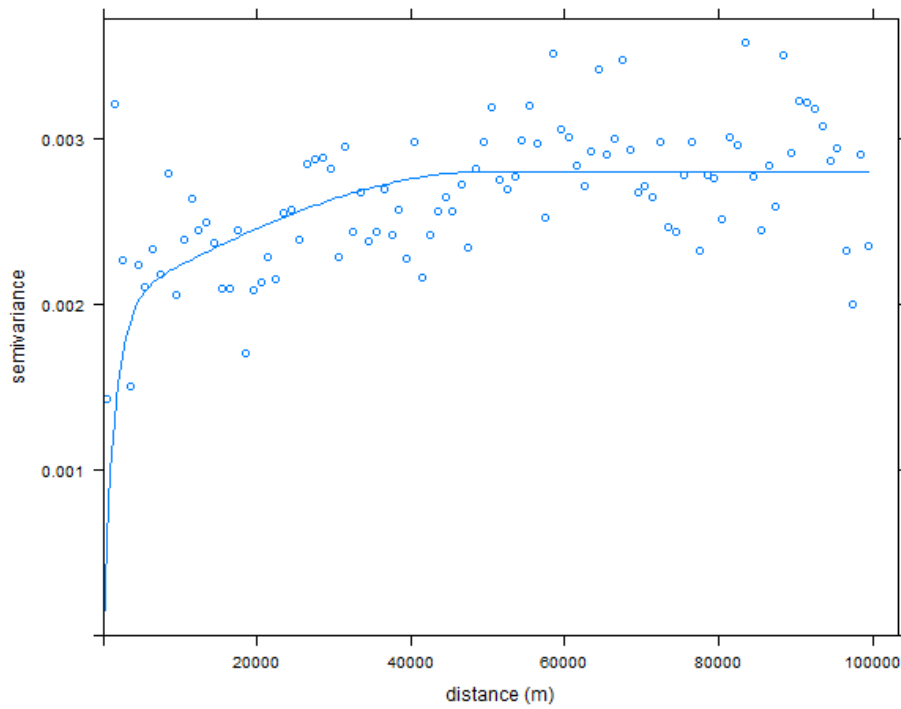
The kriging results were used to inform the designation process by predicting whether or not TIN concentrations exceed the 11.29 mg N/l threshold at locations between monitoring points. This, in turn, allowed NRW staff to use their professional judgement to decide how to define the extent of pollution in aquifers and the size of the NVZ required. The results were also used to help identify monitoring points with anomalously high or low nitrate concentrations; these may then be screened out if they are deemed to be unrepresentative of the water bodies (e.g. if they are influenced by a local point source discharge).

Figure 4.2 Variogram for current (a) and future (b) 95<sup>th</sup> percentile TIN concentrations

(a)



(b)



**Figure 4.3 Relationship between observed concentrations at groundwaters monitoring points and predicted (interpolated) concentrations in the associated 1km<sup>2</sup> grid cells for current (a) and future (b) 95<sup>th</sup> percentile TIN concentrations**

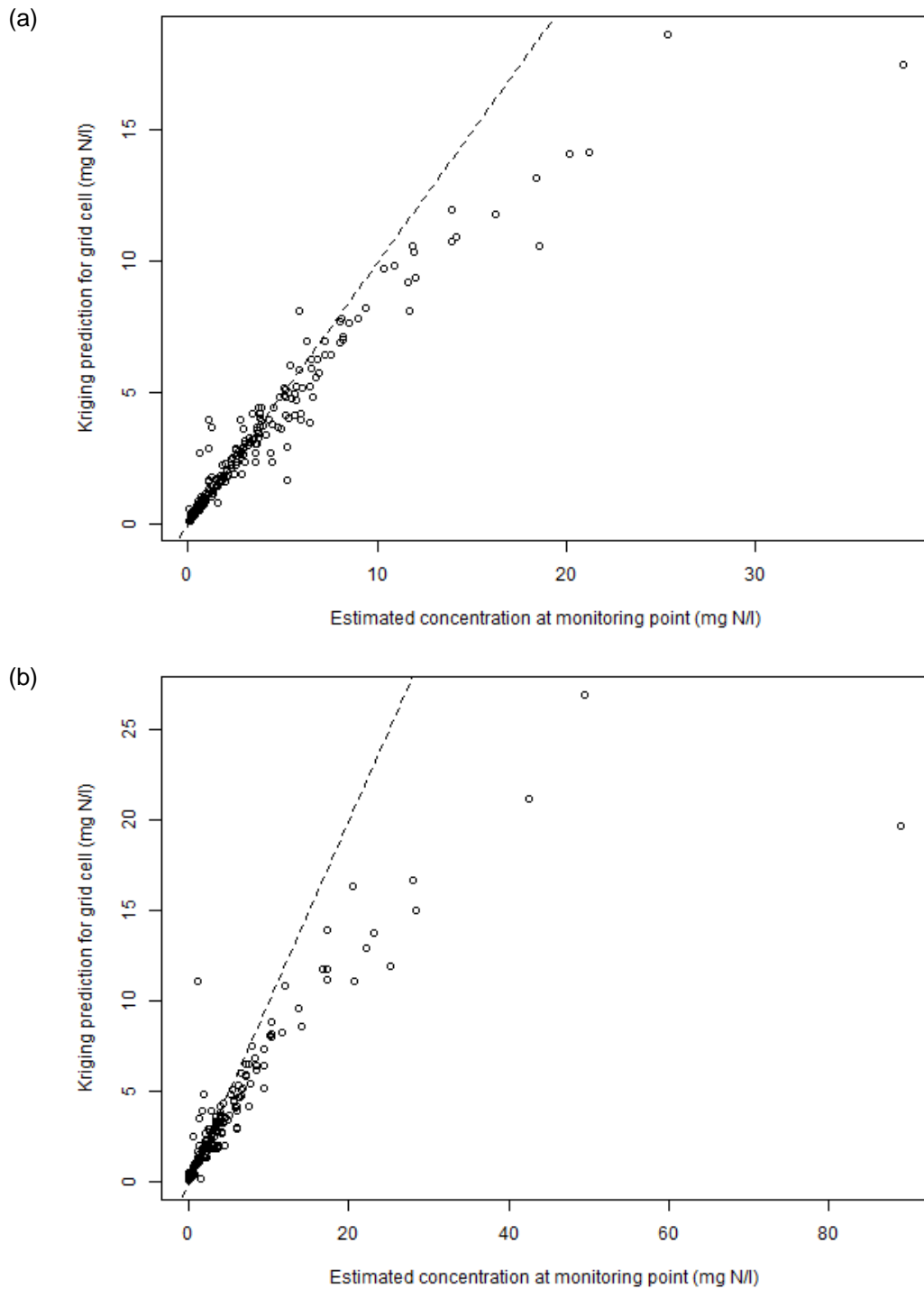


Figure 4.4 Kriged predictions of current (2015) 95<sup>th</sup> percentile TIN concentration

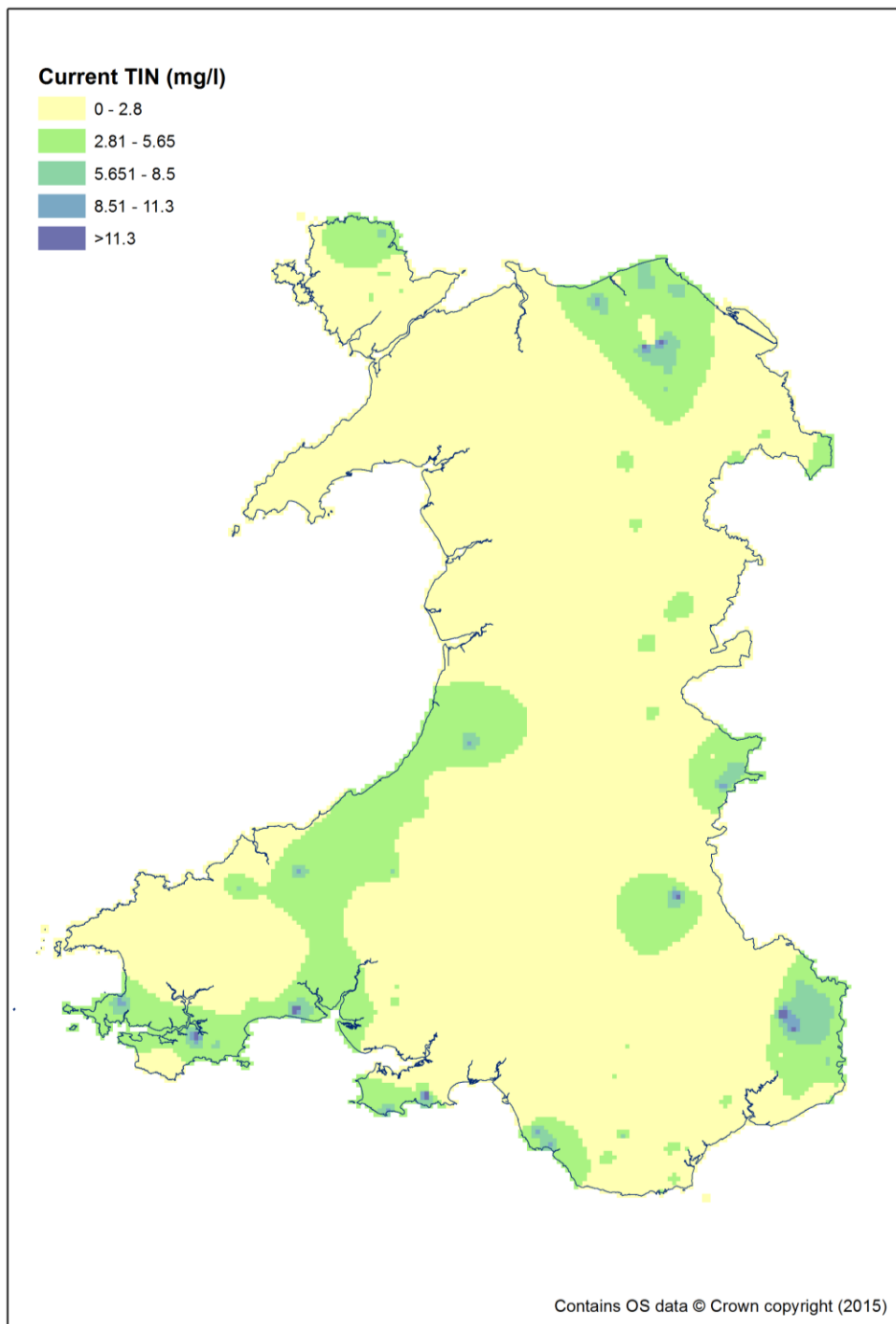
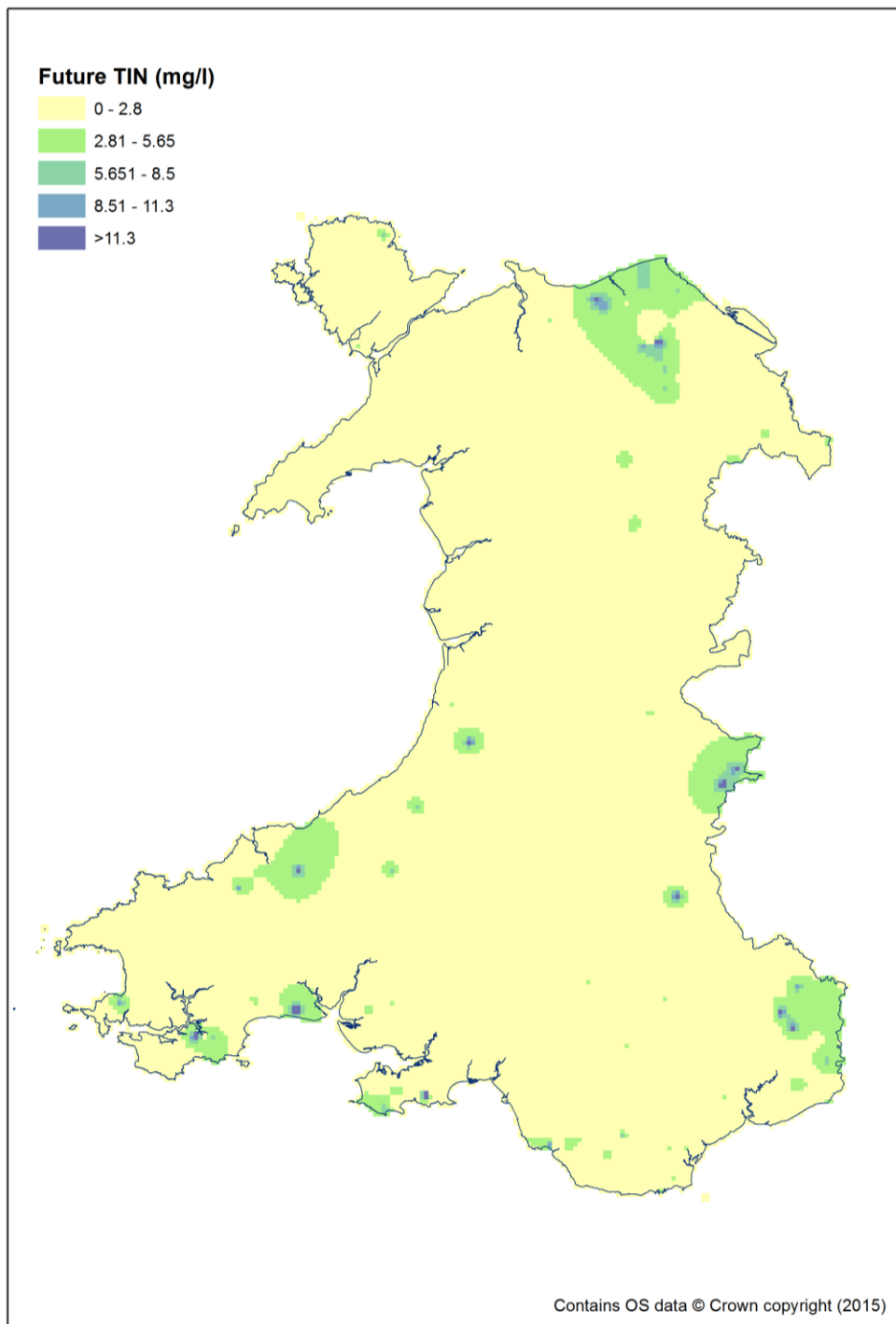


Figure 4.5 Kriged predictions of future (2027) 95th percentile TIN concentration



## 5. Modelling nitrate leaching to groundwater

### 5.1 Introduction

Land use data was used to model the amount of nitrate leaching to groundwater from agricultural land (including atmospheric deposition) and diffuse urban sources. This was done to give a more balanced assessment of the risk of nitrate pollution from agriculture than that which would have come from an assessment of monitoring data alone. The principle purposes of using the land use data were to:

- identify the significance of the agricultural contribution to any nitrate pollution identified;
- provide further confidence in the conclusions of the statistical analysis of monitoring data; and,
- minimise the possibility of designating NVZs due to historic land use practices where long travel times are required for nitrate to reach deep groundwater.

### 5.2 Agricultural land use data

Nitrate leaching from agricultural land was calculated using the NEAP-N national-scale nitrate leaching model (Lord and Anthony 2000). The model uses agricultural census data from the ADAS National Land Use database, soils data from the NSRI LAND-IS database and climate data from the UK-CIP to estimate diffuse nitrate losses at a spatial resolution of 1 km<sup>2</sup> under average climate conditions. For the 2017 NVZ Review, the model was parameterised with cropping and livestock data from the 2014 June Surveys in England and Wales.

The model considers a single maximum potential nitrogen loss coefficient for individual crop and livestock types, modified by spatially distributed information on soil type and hydrologically effective rainfall (1971-2000). These losses are then aggregated to provide estimates of nitrate leaching per land use class (kg N/ha/yr). Five land use classes were used in this analysis: managed arable crops, managed grassland, rough grazing, woodland/forest and open water. These estimates do not include any point source or urban contributions do include contributions from atmospheric deposition. Full details of how these agricultural contributions were modelled are given in Appendix B.

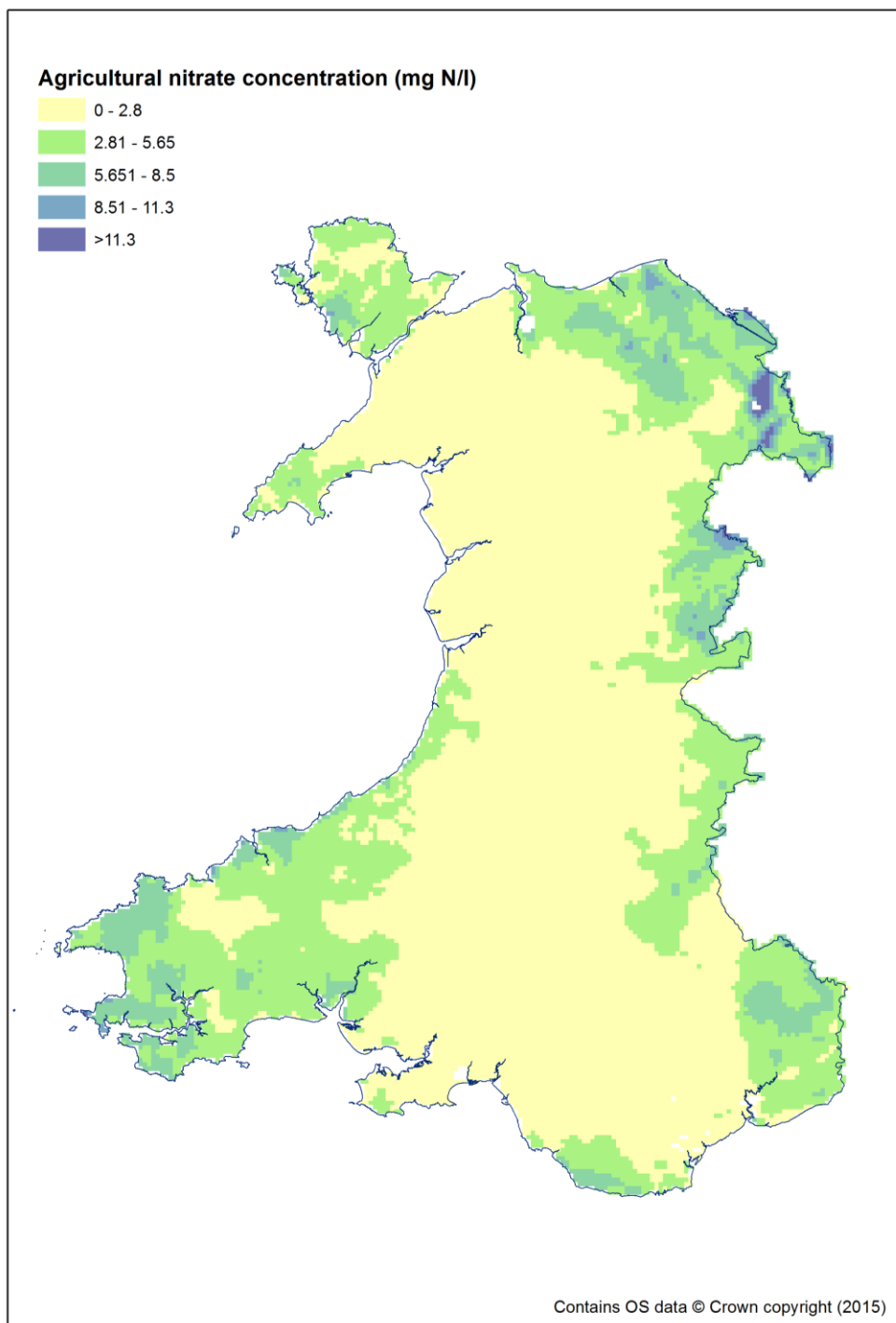
The losses from the individual land use classes were summed to give a total annual nitrate load from agricultural sources on a 1 km<sup>2</sup> scale (kg N/km<sup>2</sup>/yr). This was then divided by the hydrologically effective rainfall (mm, converted to l/km<sup>2</sup>/yr) for that 1 km<sup>2</sup> cell to give a measure of nitrate concentration in agricultural run-off (mg N/l).



Due to the uncertainty in the data coverage, the concentration for each 1 km<sup>2</sup> cell was averaged with its eight directly neighbouring cells, where available. There were a number of instances where the agricultural grid had slightly less coverage in coastal areas than the standard Wales grid used in all other elements of this analysis. In these instances these cells were ignored from the averaging process.

The resulting prediction of agricultural nitrate concentration leaching to groundwater is presented in Figure 5.1.

Figure 5.1 Modelled agricultural concentration leaching to groundwater



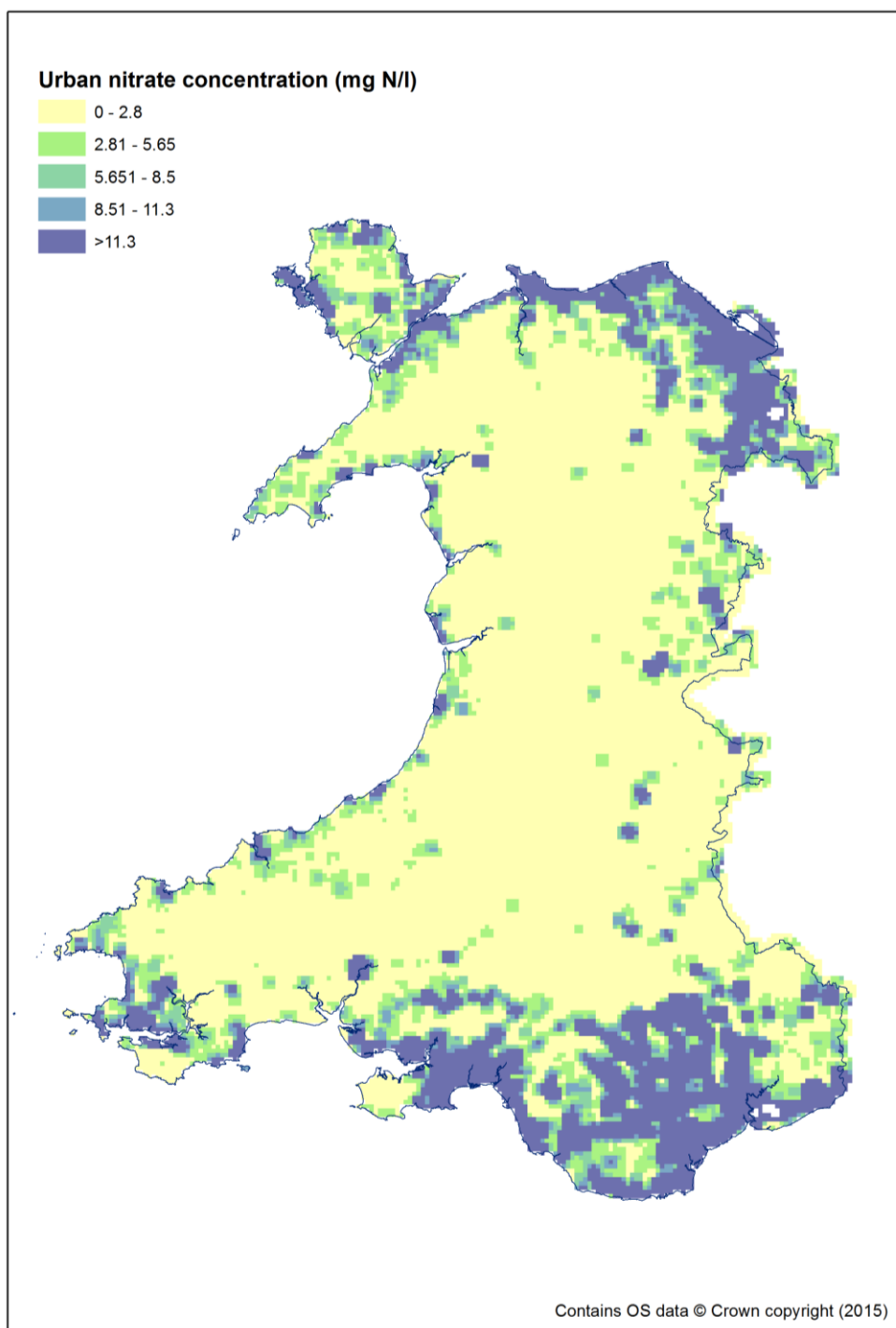
### 5.3 Urban leaching data

Urban nitrate losses were estimated using the component model of Lerner (2000). The model uses evidence from a study of recharge components in Nottingham, together with expert assumptions, to derive export coefficients (kg N/ha/yr) estimating nitrate losses to groundwater for each urban land cover type. The model identifies 14 components of runoff in total, but some components could not be included in the NVZ land use modelling as they would require access to fine scale data on the urban environment. The components that were included for the NVZ land use modelling are detailed in Appendix B3.

The losses from the individual components were summed to give a total annual nitrate load from urban sources on a 1 km<sup>2</sup> scale (kg N/km<sup>2</sup>/yr). This was then divided by the hydrologically effective rainfall (mm, converted to l/km<sup>2</sup>/yr) for that 1 km<sup>2</sup> cell to give a measure of nitrate concentration in urban run-off (mg N/l).

Due to the uncertainty in the data coverage, the concentration for each 1 km<sup>2</sup> cell was averaged with its direct neighbouring cells, where available. The resulting prediction of urban nitrate concentration leaching to groundwater is presented in Figure 5.2.

Figure 5.2 Modelled urban nitrate concentration leaching to groundwater



## 6. Combining the evidence from monitoring and modelling

### 6.1 Introduction

Following the approach used for the 2013 Review of groundwater NVZs (EA 2012), a GIS-based risk model was used to determine potential groundwater NVZs in 2017. The risk model combines national data with local evidence from NRW specialists. Each component is assessed using a scoring and weighting system, and the scores are combined to generate a national risk map at a scale of 1 km<sup>2</sup>. Areas of high risk indicate potential groundwater NVZs.

### 6.2 Introduction to the risk model

The risk model consists of eight components (Table 6.1). Four of the components were derived using national monitoring and land use datasets; the other four components were derived by soliciting local information from NRW groundwater specialists.

Three components describe “pressure” risks and are mainly derived from modelled inputs of nitrate data where the higher the pressure, the greater the risk that groundwater nitrate concentrations will exceed 50 mg NO<sub>3</sub>/l. The other five components describe the “observed” risk and draw upon a combination of water quality monitoring data and local knowledge.

**Table 6.1 Components of the risk model**

Source of information	Type of component	Risk component
National	Pressure risk	1. Agricultural nitrate leaching from the NEAP-N model
		2. Urban nitrate leaching from the Lerner model
	Observed risk	3. Kriged current (2015) groundwater nitrate concentration
		4. Kriged future (2027) groundwater nitrate concentration
Local	Pressure risk	5. Denitrification or mixing reduce the nitrate input from agriculture to groundwater
	Observed risk	6. Monitored nitrate is representative of point source pollution
		7. Monitored nitrate is unrepresentative of real groundwater nitrate concentrations
		8. Surface water – groundwater interactions identify that surface water quality is a reasonable indicator of groundwater quality

The methodology has been designed as a risk model so weightings can give the greatest importance to groundwater monitoring data and secondary importance to agricultural nitrate loss data derived from the NEAP-N model. The model has the flexibility built-in to incorporate the understanding of NRW hydrogeologists. Local NRW staff are not allowed, however, to change the scores from the nationally-derived lines of evidence.

Each component was given a score (positive scores increase the overall risk and negative scores decrease the overall risk) and weightings were applied to these scores. The weighted scores were then combined to yield an overall risk score indicating the strength of evidence that the groundwater is polluted by nitrate from agricultural sources.

## 6.3 Risk model components, scores and weights

### 6.3.1 National lines of evidence

Four of the components were derived using national datasets; Section 4 describes how the nitrate monitoring data was interpolated to produce national maps of current (2015) and future (2027) groundwater nitrate concentration and Section 5 describes how agricultural and urban nitrate leaching was estimated using land use models.

Table 6.2 details the risk scores and weights applied to these four risk components. Scores of 0, 1 and 2 were used to represent low, medium and high nitrate concentrations. Agricultural nitrate leaching was weighted positively, because higher concentrations indicate a greater risk from agriculture, and urban nitrate leaching was weighted negatively, because higher concentrations indicate a lower risk that pollution is due to agriculture. Kriged predictions of groundwater nitrate concentrations were both weighted positively, but with lower weight given to future (2027) concentrations than current (2015) concentrations because extrapolation of historic trend introduces additional uncertainty.

**Table 6.2 Risk scores and weights for national lines of evidence**

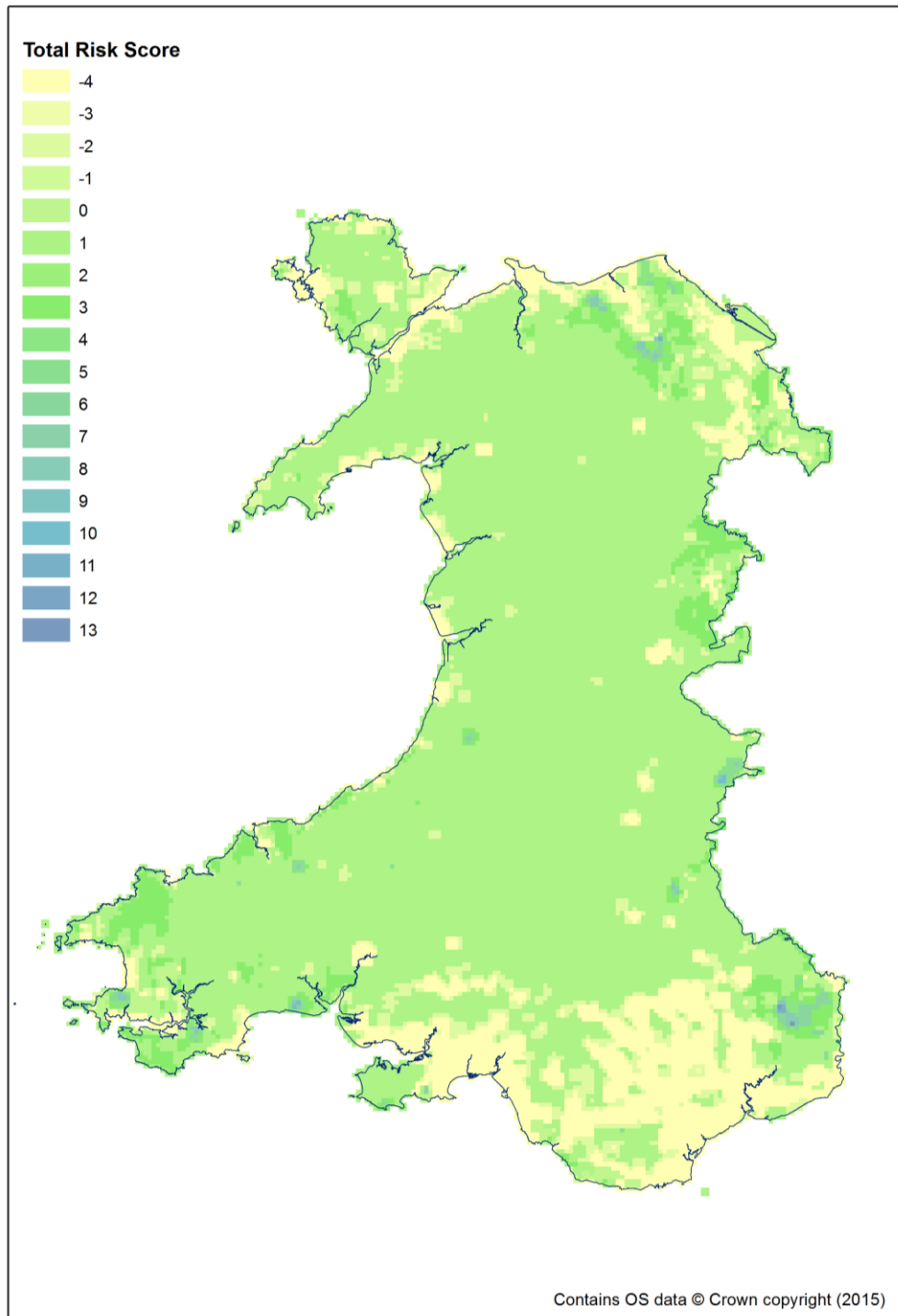
Risk component	Weight	Score		
		0	1	2
1. Agricultural nitrate leaching from the NEAP-N model	3	Leaching concentration <25 mg NO <sub>3</sub> /l	Leaching concentration 25-50 mg NO <sub>3</sub> /l	Leaching concentration >50 mg NO <sub>3</sub> /l
2. Urban nitrate leaching from the Lerner model	-2	Leaching concentration <25 mg NO <sub>3</sub> /l	Leaching concentration 25-50 mg NO <sub>3</sub> /l	Leaching concentration >50 mg NO <sub>3</sub> /l
3. Kriged current (2015) groundwater nitrate concentration	3	Current TIN concentration <25 mg NO <sub>3</sub> /l	Current TIN concentration 25-50 mg NO <sub>3</sub> /l	Current TIN concentration <50 mg NO <sub>3</sub> /l

---

Risk component	Weight	Score		
		0	1	2
4. Kriged future (2027) groundwater nitrate concentration	2	Future TIN concentration <25 mg NO <sub>3</sub> /l	Future TIN concentration 25-50 mg NO <sub>3</sub> /l	Future TIN concentration <50 mg NO <sub>3</sub> /l

Figure 6.1 maps the combined weighted score for the four national risk components.

**Figure 6.1 Risk map showing scores from National lines of evidence prior to NRW local review**





### 6.3.2 Local lines of evidence

The national lines of evidence described in Section 6.3.1 above were reviewed by NRW and supplemented by a further four risk components representing local evidence about the extent and source of nitrate pollution. Like the national lines of evidence, individual 1 km<sup>2</sup> grid cells were assigned risk scores for each risk component, and each risk component was weighted to reflect its importance within the overall assessment. Unlike the national lines of evidence, however, the assessment of local evidence was based on professional judgement of NRW hydrogeologists.

Table 6.3 details the risk scores and weights applied to these four local risk components.

Risk component #5 is for circumstances where either denitrification or mixing will decrease the nitrate input from agriculture to groundwater. To allow for this line of evidence to be changed from 0 (No evidence), nitrate leaching from agriculture must be greater than 25 mg NO<sub>3</sub>/l and evidence of denitrification or mixing should be supplied by local NRW staff.<sup>5</sup>

Risk component #6 is for circumstances where it can be identified that an area of high nitrate is representative of point source pollution rather than diffuse nitrate pollution. It acts as a safeguard against falsely designating NVZs due to localised pollution from point source discharges.

Risk component #7 is used to reflect situations where the monitored nitrate is unrepresentative of real groundwater nitrate concentrations. Where observed nitrate concentrations are unrealistically low, a positive score increases the overall risk score. This may be because:

- nitrate pollution has not passed through the unsaturated zone; it is on its way but has not yet been detected by monitoring;
- deep abstractions sample older, cleaner water that is not representative of current nitrate pressure; due to the depth of monitoring the data is unrepresentative; or
- uncertainty in predicted nitrate values caused either by short duration of monitoring or a significant variation in the dataset.

This modification is only possible if the observed groundwater concentration is less than 50 mg NO<sub>3</sub>/l.

---

<sup>5</sup> Note: This component of the risk model wasn't used by local EA Hydrogeology staff during the 2013 review cycle. WRc is not aware of whether or not it has been used for this current review cycle.

Conversely, where quality assurance checks identified monitoring points with unrealistically high nitrate concentrations or highly uncertain trend analysis results (see Section 3.3.8 for details), a negative score reduces the overall risk score.

Finally, risk component #8 is for situations where there is a clear surface water-groundwater interaction. If it can be identified that surface water quality is a reasonable indicator of groundwater quality then evidence from surface water monitoring data may be used to complement the groundwater monitoring dataset. This is only appropriate for situations where surface water and groundwater interaction is significant and surface water quality is not dominated by point source discharges.

**Table 6.3 Risk scores and weights for local lines of evidence**

Risk component	Weight	Score				
		2 Yes, good evidence	1 Yes, maybe some evidence	0 No evidence	-1 No, maybe some evidence	-2 No, good evidence
5. Denitrification or mixing reduce the nitrate input from agriculture to groundwater	-1	Yes, good evidence. Baseline report indicating de-nitrification or mixing. Local report (must be referenced), indicating as above.	Maybe some evidence. Identifiable source of dilution e.g. forested recharge area. Drift > 10 m thick and clay rich.	No evidence		
6. Monitored nitrate is representative of point source pollution	-5	Identifiable point source of nitrate AND groundwater concentrations significantly higher than soil leaching concentrations.	Identifiable point source of nitrate OR groundwater concentrations significantly higher than soil leaching concentrations.	No evidence		
7. Monitored nitrate is unrepresentative of real groundwater nitrate concentrations	3	Water company has abandoned a source nearby due to high NO <sub>3</sub> . Unsaturated zone > 30 m delaying nitrate measurement. Aquifer is layered or the sampling is at depth.	Large uncertainty in trend analysis. Significant drift > 10 m delaying nitrate measurement.	No evidence	Large errors in kriging OR large uncertainty in trend analysis results	Large errors in kriging AND large uncertainty in trend analysis results
8. Surface water – groundwater interactions identify that surface water quality is a reasonable indicator of groundwater quality	1	Confident fail <sup>1</sup> and >2 point source discharges in surface water OR marginal fail and <2 point source discharges in surface water.	Face value fail and <2 point source discharges in surface water OR face value pass and <2 point source discharges in surface water.	No evidence		

<sup>1</sup> See Table 2.1 on page 9 for definition of confidence bands. The same bands were used to assess the confidence that the 95<sup>th</sup> percentile nitrate concentrations in surface waters exceeded 50 mg NO<sub>3</sub>/l.

## 6.4 Overall risk assessment

The weighted scores for the four pressure risk components are summed to provide an intermediate positive pressure risk score. Since the urban loading score is associated with a negative weight (-2), where urban loading outweighs the agricultural load, the intermediate pressure score could be negative. In this situation, the pressure risk score is set to zero as a minimum. The weighted scores for the four observed risk components are also summed to provide an intermediate positive observed risk score. The intermediate pressure and observed risk scores are added together and classified as shown in Table 6.4.

**Table 6.4 Summary of categories for final risk score**

What is the risk that groundwater nitrate is >50 mg/l or will be in future and agriculture is the main source?			
Category	High – Designate	Medium – No action	Low – No action (consider for de-designation)
Score	>8	3-8	<3

If the risk that groundwater nitrate concentration is exceeding 50mg NO<sub>3</sub>/l and agriculture is the source, the score will be higher than 8. This will lead to potential groundwater NVZ designations. A medium score ranges from 8 to 3 (in this case no action is required) and a low score is less than 3 (consider these areas for de-designation if they have previously been designated).

## 7. Comparison with 2013 Review methodology

### 7.1 Introduction

For consistency and comparability, this Review follows closely the methodology used for the previous (2013) Review of surface water NVZs in England and Wales (Environment Agency 2012). However, the present Review provided an opportunity to learn lessons from the 2013 Review and improve how the method was applied and the results were presented. Some refinements were made to the methodology to make better use of the available data and to provide a more rigorous system of checks and balances.

### 7.2 Refinements to Step 1: Analysing groundwater quality monitoring data

Table 7.1 details the refinements made to the methodology used in Step 1.

**Table 7.1 Summary of Summary of methodological refinements for Step 1**

Refinement	Justification
The number of simulated future time series was increased from 1,000 to 10,000.	The greater number of shots provides greater stability in the future forecasts.
A fixed seed was used in the random number generator.	The use of a fixed seed makes the results of the Monte Carlo simulation reproducible.
Uncertainty in the mean concentration at the end of the monitored period was incorporated into the AntB Monte Carlo simulation.	By not using a fixed starting concentration, the simulation produces a wider range of projections that more accurately represent the uncertainty in future concentrations.
A statistical test was added to test whether the measured concentrations at each site follow a bimodal or multi-modal distribution.	The test provides a useful indicator of sites where erratic concentrations can give rise to unreliable future forecasts.
The coefficient of variation (standard deviation / mean) was calculated for each site.	A high coefficient of variation can indicate sites where highly variable concentrations can give rise to unreliable future forecasts.
For each site, the % ammonium contribution across all $n$ samples was calculated as: $\frac{\sum_{i=1}^n TIN_i - \sum_{i=1}^n TON_i}{\sum_{i=1}^n TIN_i}$	A high % ammonium contribution can indicate a strong pollution contribution from point source discharges.

Refinement	Justification
For each site, the mean concentrations in winter (Dec-Feb), spring (Mar-May), summer (Jun-Aug) and autumn (Sep-Nov) were calculated.	Peak concentrations in summer can indicate a strong pollution contribution from point source discharges.
A time series plot was produced for each monitoring site, showing the historical data, fitted trend, and future forecasts.	The plots allow close visual inspection of the results to verify the assessments of current and future status.
For AntC, consecutive years are identified started from the most recent year of monitoring.	In cases where two or more runs of consecutive years are available, the most recent data is used, which should provide a more reliable indicator of current and future water quality.
To use the Weibull method, there must be at least one sample in five of the last six years (2009-14).	This rule brings the groundwater method into line with the surface water method and ensure that percentile estimates are based on data from a small number of potentially unrepresentative years.
Use of the AntC2 tool for assessing the status of data-poor groundwater sites was abandoned in preference of the 'Mean Concentration' method.	Only 15 sites were analysed by AntC2 in the 2013 Review, and only 28 sites meet the criteria for AntC2 in the 2017 Review. The predictions from AntC2 typically have a high degree of uncertainty, and provide little additional information over the simpler Mean Concentration method.
The data screening checks used by AntB were applied to all sites, not just those analysed using AntB.	The screening checks provide a more comprehensive data quality check for every site.
A conversion factor of 50/43 (~1.16) was applied to estimates of mean concentration to produce an estimate of the lower 90% confidence limit on the 95 <sup>th</sup> percentile.	A conversion factor of 50/43 was recommended at time of the previous NVZ Review, based upon an analysis of monitoring data that showed the mean concentration to be, on average, 7 mg NO <sub>3</sub> /l lower than the lower 90% confidence limit on the 95 <sup>th</sup> percentile.

### 7.3 Refinements to Step 2: Estimating groundwater quality between boreholes

Table 7.2 details the refinements made to the methodology used in Step 2.

**Table 7.2 Summary of Summary of methodological refinements for Step 2**

Refinement	Justification
Different coefficients used for power transformation of current and future TIN concentrations.	The power coefficients used to transform the data were optimised to make each dataset as normally distributed as possible; as such, they are specific to each dataset.
The preferred variogram model was selected by assessing the predictive performance of the kriging model.	This provides a more objective way to decide the best way of interpolating the groundwater monitoring data, and guards against model over-fitting.
A nested exponential-spherical variogram model was used instead of a double-exponential model.	This model was found to have predict nitrate concentrations in unmonitored 1km <sup>2</sup> grid cells more accurately.

### 7.4 Refinements to Step 3: Modelling nitrate leaching to groundwater

The methodology used to model nitrate leaching from agriculture and diffuse urban sources was the same as that used in the 2013 NVZ Review.

### 7.5 Refinements to Step 4: Combining the evidence from monitoring and modelling

The scoring and weighting system used to combine the evidence from monitoring and modelling was the same as that used in the 2013 NVZ Review.

## 8. Recommendations

For consistency and comparability, this review follows closely the methodology used for the previous (2013) Review of NVZs in England and Wales (Environment Agency 2012). Some refinements were made to the methodology to make better use of the available data and to provide a more rigorous system of checks and balances, which are highlighted and discussed in this document.

Despite these refinements, the methodology still has some limitations.

1. The use of kriging to estimate groundwater nitrate concentrations at unmonitored locations is a purely statistical exercise that fails to incorporate any other sources of information about pollution pressures. Land use has been shown to be a good predictor of nitrate concentrations in rivers and an appropriately calibrated regression model could be used to help improve predictions of groundwater nitrate concentration in parts of the country with little or no monitoring data. The residuals from the regression model could still be kriged to take account of any spatial pattern in unexplained variation.
2. Many groundwater monitoring points have limited numbers of samples and/or significant gaps in the monitoring record, both of which make it difficult to estimate current and future nitrate concentrations precisely. In particular, the AntC trend analysis tool is prone to producing projections that have exceptionally wide confidence intervals or which are unduly sensitive to high or low measurements in a single year. An alternative would be to make greater use of the 'Mean Concentration' method, and to build in algorithms or rules to account for the additional uncertainty that arises when the monitoring record is not up to date. For example, a standard auto-correlation coefficient could be calculated for a set of monitoring points with similar characteristics, rather than attempting to estimate an auto-correlation coefficient for each monitoring point separately from as few as six data points.
3. An empirical adjustment factor of 1.16 was used to convert estimates of mean concentration into an estimate of the lower 90% confidence limit on the 95<sup>th</sup> percentile (3.3.7). Use of this simple scaling factor ensures that a common and comparable summary statistic can be generated for every monitoring point, but it also introduces an additional, unquantified source of uncertainty. Furthermore the use of a single, universal factor falsely assumes the same level of variability in nitrate concentrations at each monitoring point, with the consequence that it provides an uneven level of protection to groundwaters across England. Previous analysis of data from Anglian region in England suggested that monitoring points in chalk have a more stable water quality than monitoring points in gravel (Environment Agency 2012). If extended to the whole country, this could be used to develop type-specific thresholds and scaling factors.



## References

Addiscott, T.M. and Whitmore, A.P. 1991. Simulation of solute leaching in soils of different permeabilities. *Soil Use and Management*, 7, 94–102.

Anthony, S., Quinn, P. and Lord, E. 1996. Catchment scale modelling of nitrate leaching. *Aspects of Applied Biology*, 46, 23–32.

Chambers, B.J., Lord, E.I., Nicholson, F.A. and Smith, K.A. 1999. Predicting nitrogen availability and losses following application of organic manures to arable land: MANNER. *Soil Use and Management*, 15, 137–143.

Curtis, C., Allott, T., Hall, J., Harriman, R., Helliwell, R., Hughes, M., Kernan, M., Reynolds, B. and Ulyett, J. 2000. Critical loads of sulphur and nitrogen for freshwaters in Great Britain and assessment of deposition reduction requirements with the first order acidity balance model. *Hydrology and Earth System Sciences*, 4(1), 125–140.

Defra, 2008. Description of the methodology applied in identifying waters and designating Nitrate Vulnerable Zones in England.

Ellis, J.C., van Dijk, P.A.H. and Kinley, R.D. 1993. Codes of Practice for Data Handling – Version 1. WRc Report No. NR 3203/1/4224 for The National Rivers Authority.

Environment Agency, 2012. Method Statement for Nitrate Vulnerable Zone Review – Groundwaters. Environment Agency report to Defra and Welsh Government – supporting paper for the Implementation of the Nitrates Directive 2013 – 2016. April 2012.

Goulding, K., Bailey, N. and Bradbury, N. 1998. A modelling study of nitrogen deposited to arable land from the atmosphere and its contribution to nitrate leaching. *Soil Use and Management*, 14, 70–77.

Hall, J., Hornung, M., Freer-Smith, P., Loveland, P., Bradley, I., Langan, S., Dyke, H., Gascoigne, J. and Bull, K. 1997. Current Status of UK Critical Loads Data – December 1996. ITE Monks Wood, UK, 22 pp. plus appendices.

Hatch, D., Goulding, K. and Murphy, D. 2002. Nitrogen. In: Haygarth, P. and Jarvis, S. (Eds), *Agriculture, Hydrology and Water Quality*. CAB International, pp. 7–27.

Lerner, D.N. 2000. Guidelines for Estimating Urban Loads of Nitrogen to Groundwater. Defra Project Report, NT1845, 21 pp.

Lord, E.I. 1992. Modelling of nitrate leaching: Nitrate Sensitive Areas. *Aspects of Applied Biology*, 30, 19–28.

Lord, E.I., Anthony, S.G., Miles, A. and Turner, S. 1995. Agricultural Nitrogen Surplus and Diffuse Nitrogen Losses to Surface Waters within England and Wales. Final Report to MAFF RMED for submission to PARCOM. ADAS, Wolverhampton, 55 pp.

Lord, E. and Anthony, S. 2000. MAGPIE: A modelling framework for evaluating nitrate losses at national and catchment scales. *Soil Use and Management*, 16, 167–174.

Powlson, D. and Goulding, K. 1994. Agriculture, the nitrogen cycle and nitrate. In: North, J. (Ed.) Nitrate Control Policy, Agriculture and Land Use. Cambridge Environment Initiative, 4, 5–25.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

Scholefield, D. and Rodda, H.J.E., 1992. The development of a nitrogen cycle model to predict nitrate losses through leaching from grassland catchments. *Aspects of Applied Biology*, 30, 435–438.

Scholefield, D., Lockyer, D.R., Tyson, K.C. and Whitehead, D.C. 1991. A model to predict transformations and losses of nitrogen in UK pastures grazed by beef cattle. *Plant & Soil*, 132, 165–177.

Smith, K.A., Brewer, A.J., Crabb, J. and Dauven, A. 2001a. A survey of the production and use of animal manures in England and Wales. II Poultry manures. *Soil Use and Management*, 17, 48–56.

Smith, K.A., Brewer, A.J., Crabb, J. and Dauven, A. 2001b. A survey of the production and use of animal manures in England and Wales. III Cattle manures. *Soil Use and Management*, 17, 77–87.

## Appendix A Multiple Outlier Test

### A1 Objective

The objective of MOT (Multiple Outlier Test) is to identify outliers in a given data set, on the assumption of underlying Normality.

### A2 Definitions

*Outlier*: A data value which has arisen from some statistical population that is more extreme than the population from which the bulk of the values have arisen.

*Suspected outlier*: a data value which is so far above or below the bulk of the data values that it causes surprise to the user of the data.

Although the definition of suspected outlier might appear rather subjective, it carries with it the implication that *the user must have some correct probability distribution in mind* (however vague), and believes that the suspected outliers are not consistent with that distribution. In other words, he or she suspects that the sample has been contaminated by observations from some statistical distribution other than the one expected.

### A3 The single outlier test

A well-established statistical procedure is available when the data can be assumed to have come from an underlying Normal distribution (or where the data can be transformed, for example by taking logarithms) so as to make this assumption reasonable. The test proceeds as follows. First calculate the mean ( $m$ ) and standard deviation ( $s$ ) of the data values. Then calculate the quantity  $t_{\max}$  as:

$$t_{\max} = |(x_{\gamma} - m)|/s$$

where  $x_{\gamma}$  is the suspected outlier (that is, either the minimum or the maximum of the data set).

If  $t_{\max}$  is greater than the value given in Table A.1, the outlier can be declared to be statistically significant at the 1% level. In other words, the probability that a value as extreme as this could have arisen by chance from a Normal population is only 1 in 100.

**Table A.1 Critical values (P = 1%) of the  $t_{\max}$  statistic**

No. of data values	Critical value
4	1.49
5	1.75
6	1.94
7	2.10
8	2.22
9	2.32
10	2.41
12	2.55
14	2.66
16	2.75
18	2.82
20	2.88
30	3.10
40	3.24
50	3.34
60	3.41
80	3.53
100	3.60
120	3.66
150	3.72
200	3.81
300	3.91
400	3.97
500	4.03

#### **A4 The multiple outlier test**

What if several outliers are suspected to be present in the data? In that case a simple generalisation of the preceding test can be used. This is known as the Multiple Outlier Test.

The test takes an 'outward consecutive' approach. First, a pool of  $k$  suspects is produced by finding all the data values whose  $t_{\max}$  value (as defined above) is greater than some arbitrary

limit (see below). The  $k$  suspects are then tested one at a time, working from the least to the most extreme. The details of the procedure are as follows:

- Starting with the  $(n-k)$  'reliable' data values, augment these by just the least extreme of the  $k$  suspected outliers.
- Calculate the mean and standard deviation of those  $(n-k+1)$  values, and perform the single outlier test as usual.
- If the suspect fails to be confirmed as an outlier, pool it with the  $(n-k)$  reliable values, recalculate the mean and standard deviation, and then test the next least extreme suspect.
- Continue in this way until a suspected outlier is declared to be a genuine outlier. All the remaining (and hence more extreme) values are then declared to be outliers also.

It might be thought that, as the multiple outlier test provides several opportunities for false positives, the actual significance levels would be somewhat higher than the nominal 1% value quoted in Table A.1 for the single outlier case. That is not the case, however. When the data values really do come from a Normal population, the multiple outlier test very rarely produces false positives unless the single outlier test does so also - a characteristic that has been confirmed by computer simulation.

Following the 2008 and 2013 Review methodologies, the nitrate datasets were assumed to be Normally distributed. A critical value of 7 was used to strike a reasonable balance between identifying outliers that were plainly abnormal, and retaining data values that were merely suspicious.

## Appendix B Land use model data

### B1 Agricultural land use data

Nitrate losses are modelled for a range of categories of arable crop and provide a maximum potential nitrogen loss value on a 1 km<sup>2</sup> resolution. The soil nitrate content during the leaching period is calculated as a function of the balance of nitrogen inputs to and nitrogen off-take by the preceding crop. Losses modelled using this approach were found to compare favourably with independent field measurements using porous pots in fields (Lord 1992, Anthony *et al.*, 1996) and with stream nitrate fluxes measured in several contrasting catchments (Lord *et al.*, 1995). For grazing livestock the manure nitrogen loadings are derived from research underpinning the N-Cycle model developed by IGER in the UK (Scholefield *et al.*, 1991, Scholefield and Rodda 1992). Maximum potential nitrogen losses have been modified by ADAS data underpinning the MANNER model (Chambers *et al.*, 1999) to consider only the leachable component of the manure nitrogen and the timing of manure applications with typically around half applied in autumn and half the following spring (Lord *et al.*, 1995, Smith *et al.*, 2001a,b).

The NEAP-N simulation of nitrate leaching uses a series of regressions to simplify the UK Meteorological Office's hydrological model MORECS by expressing long-term winter soil drainage as functions of long-term mean annual rainfall and potential evapotranspiration for different crop and soil combinations. For this purpose the dominant soil series within each spatial calculation unit is placed into one of three classes based on Available Water Capacity (AWC), the water held between field capacity and permanent wilting point. A simple leaching function is then used to relate soil drainage to soil water content at field capacity to calculate the proportion of vulnerable nitrogen leaching (Anthony *et al.*, 1996). The model is a simplification of the SLIM model developed by Addiscott and Whitmore (1991).

### B2 Atmospheric deposition

In order to estimate the response of the arable and grassland NEAP-N autumn soil nitrogen coefficients to atmospheric nitrogen deposition the N-Cycle (Scholefield *et al.*, 1992) and NitCat (Lord 1992) models were used. Both models calculate autumn soil nitrogen as a function of the mass balance of nitrogen in fertiliser inputs and crop off-take. Atmospheric nitrogen deposition is treated as an additional fertiliser input, albeit distributed more evenly in time than fertiliser applications. This is justified as there is evidence that atmospheric nitrogen deposition may increase crop yields, indicating that fertiliser recommendations should be adjusted to take account of regional data on atmospheric nitrogen deposition rates (Hatch *et al.*, 2002).

The current NEAP-N 'potential nitrogen leaching' coefficients represent typical agricultural practices under average environment conditions (i.e. they include the effect of average atmospheric deposition rates). The methodology used was therefore to calculate the

difference between the local and national average atmospheric nitrogen deposition rate (within agricultural areas), and multiply this by the modelled response coefficient. The result is added to the base coefficient of potential nitrogen leaching for a crop or grassland area.

For rough grazing and woodland, where there is in most years no off-take, a different approach is taken. Empirical data and assumptions used by the FAB critical load exceedance model (Curtis *et al.*, 2000) are used to establish a simple relationship between atmospheric nitrogen deposition and the autumn soil nitrogen coefficient.

## B2.1 Managed grassland

Average annual nitrogen deposition across managed grassland is 16.7 kg/ha. The N-Cycle model is the source of the autumn nitrogen coefficients used in the NEAP-N model to represent leaching losses associated with grazed livestock. The N-Cycle model was set up to represent grazed grass (dairy and beef) receiving the average UK total fertiliser application rate (80 kg N/ha) and a range of atmospheric nitrogen deposition rates (15 to 36 kg N/ha) to calculate the leaching response coefficient. This was done for each soil texture class recognised by the model and for both long-term (11–20 years) and ley (2–3 years) grassland.

The response coefficient of the model varied in the range 0.14 to 0.23 (i.e. between 14 and 23% of atmospheric nitrogen deposition is at risk of leaching). There was an increase in leaching risk between beef and dairy, and between long and short-term grass. From these calculations, an average autumn nitrogen response coefficient of 0.20 was selected, biased towards the results for permanent pasture and independent of soil texture.

## B2.2 Arable land

Average annual atmospheric deposition across arable land is 15.8 kg N/ha. The NitCat model is the source of the autumn nitrogen coefficients used in the NEAP-N model to represent leaching losses from arable land. The NitCat model predicts an increase of 0.5 kg N/ha per kg of additional nitrogen input to an arable crop that is already receiving fertiliser at the economic optimum rate, and a decrease of about 0.1 kg N/ha per kg of nitrogen less than optimal. Under conventional cropping practices, fertiliser application rates to arable crops should be close to the economic optimum. It is therefore appropriate to assume an average autumn nitrogen response of 0.3 and independent of soil texture class (i.e. 30% of atmospheric nitrogen deposition is at risk of leaching). A greater percentage of the atmospheric nitrogen deposition is at risk of leaching, compared to the managed grassland, due to the autumn and winter period when nitrogen is not taken up by a crop.

Powelson and Goulding (1994) have estimated that up to 30% of the nitrogen deposited to winter wheat could be leached. More recently, Goulding *et al.* (1998) used the SUNDIAL model to calculate that 13–23% of atmospheric nitrogen deposited in the Waveney and Lichfield Nitrate Vulnerable Zones (NVZs) was leached each year. Adjusting for the efficiency

of drainage using the SLIMMER algorithm (Anthony *et al.*, 1996), the autumn nitrogen response coefficient derived from the SUNDIAL model is calculated to be 0.27 for the Waveney and 0.23 for the Lichfield NVZ. This was calculated from annual average soil drainage of 150 mm via clay-loam over clay soil (field capacity of 330 mm) in the Waveney and 240 mm soil drainage via loamy sand (field capacity of 210 mm) in the Lichfield NVZ. These coefficients are approximately the same as derived from the NitCat model response characteristics.

### B2.3 Woodland and rough grazing

Average annual atmospheric deposition across rough grazing land is 19.4 kg N/ha. Nitrogen leaching from non-agricultural land is a function of the balance of atmospheric nitrogen deposition, long-term immobilisation of nitrogen in soil, denitrification and net off-take of nitrogen in harvested material. Annual long-term immobilisation and denitrification have been estimated for soil texture and drainage classes from the ranges used in the FAB model (Hall *et al.*, 1997, Curtis *et al.*, 2000). Net off-take is restricted to woodland at a rate of 4 kg N/ha/yr (Curtis *et al.*, 2000).

The excess nitrogen is at risk of leaching. Summary deposition and measured nitrate leaching for 13 afforested catchments are given by Curtis *et al.* (1998). Leaching increases significantly only for those catchments with a total annual nitrogen deposition in excess of 10 kg N/ha. However, leaching is generally less than the balance of atmospheric deposition over immobilisation and denitrification. Approximately 65% of the balance is retained. These are sites with very high annual rainfall, at which observed nitrate loss would equate to the autumn soil nitrogen coefficient (i.e. the effective autumn soil nitrate is completely leached out each drainage year).

For this study, therefore, the autumn soil nitrogen coefficient for woodland and rough grazing is predicted as 35% of the balance of atmospheric nitrogen deposition, forestry off-take and the soil-dependent immobilisation and denitrification rates.

### B3 Urban land use data

**Parks and gardens:** The areas of parks and gardens were mapped using category 10 (Urban Green Spaces) data from the CORINE 2000 land cover dataset. Nitrogen loss was calculated as an average rate of 4.6 kg N/ha/yr. This excludes any atmospheric deposition which is already taken into account.

**Recreational grassland and golf courses:** Areas were calculated using category 11 (Sport and Leisure Facilities) data from the CORINE 2000 land cover dataset. Nitrogen loss was calculated as an average rate of 9.2 kg N/ha/yr, excluding any atmospheric deposition.



**Construction activities:** Areas were mapped using category 9 (Construction Sites) data from the CORINE 2000 land cover dataset. Nitrogen losses were estimated at an annual average rate of 400 kg N/ha/yr. This figure represents the total that will be released over several years following site disturbance, hence the annual rate assumes that the mapped area of construction activities represents an annual average.

**Spills and leaks in industry:** Areas were mapped using category 3 (Industrial or Commercial Units) data from the CORINE 2000 land cover dataset. Nitrogen losses were estimated at an average rate of 175 kg N/ha/yr.

**Leaking sewers and water mains:** Loss estimates of 0.06 kg N per capita for leaking sewers and 0.3 kg N per capita for leaking water mains were calculated and a 1 km<sup>2</sup> population map was derived by ADAS from Office of Population Census and Surveys (OPCS) 2001 census and Address Point data on the location of individual properties.

The urban diffuse estimates did not explicitly include nitrate from highway runoff, but road runoff in urban areas was included implicitly because the urban losses model is based on monitoring of nitrate loads within urban drains.