



Surface Water Method Statement for Wales Nitrate Vulnerable Zone Review 2017

**Report Reference: UC11212.04
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

Follow Us:



Surface 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: UC11212.04

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	Updated interim report issued to NRW for comment	Andrew Davey	Andrew Davey	01/12/15
V0.3	Draft final report including de-designation assessment	Andrew Davey	Andrew Davey	11/12/15
V0.4	Final report	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	5
1.1 Background	5
1.2 Criteria for identifying polluted waters	6
1.3 Structure of this report	6
2. Overview of the methodology.....	7
2.1 Introduction	7
2.2 Step 1 - Identification of surface freshwaters for analysis	8
2.3 Step 2 - Statistical analysis of surface water quality monitoring data.....	8
2.4 Step 3 - Land use modelling of nitrate pollution in surface waters.....	10
2.5 Step 4 – Combining the evidence from monitoring and modelling.....	11
2.6 Step 5 - Identifying land draining to polluted surface waters	12
2.7 Step 6 – Identification of monitoring sites unduly influenced by point sources	12
2.8 Step 7 - Removal of previous surface water NVZs	13
2.9 Step 8 - Using local knowledge to refine draft designations	14
3. Identification of surface freshwaters for analysis	16
3.1 Definition of surface waters.....	16
3.2 Categorisation of monitoring sites	16
4. Statistical analysis of surface water quality monitoring data.....	18
4.1 Introduction	18
4.2 Source dataset	19
4.3 Data processing	20
4.4 Data analysis methods	24
4.5 Summary of results.....	32
5. Land use modelling of nitrate pollution in surface waters	35
5.1 Introduction	35
5.2 Derivation of response variables.....	35
5.3 Derivation of explanatory variables	37
5.4 Modelling methodology.....	40
5.5 Fitted models.....	41

5.6	Model predictions	42
6.	Combining the evidence from monitoring and modelling.....	46
6.1	Lines of evidence.....	46
6.2	Combining the lines of evidence	47
7.	Removal of previously designated surface water NVZs.....	51
7.1	Introduction	51
7.2	Criteria for removal from designation	51
8.	Comparison with 2013 Review methodology.....	54
8.1	Introduction	54
8.2	Refinements to Step 1: Identification of surface freshwaters for analysis	54
8.3	Refinements to Step 2: Statistical analysis of surface water monitoring data.....	54
8.4	Refinements to Step 3: Land use modelling of nitrate pollution in surface waters.....	55
8.5	Refinements to Step 4: Combining the evidence from monitoring and modelling.....	56
8.6	Refinements to Step 7: Removal of previously designated surface water NVZs.....	56
	References	57

Appendices

Appendix A	Multiple Outlier Test.....	59
Appendix B	Land use model data	62
Appendix C	Details of land use models.....	66

List of Tables

Table 2.1	Summary of confidence classes.....	9
Table 3.1	Definitions of monitoring site categories.....	16
Table 4.1	Summary of confidence classes.....	18
Table 4.2	Purpose codes included in surface water analysis.....	19
Table 4.3	Sampling point types included in surface water analysis	21
Table 4.4	Summary of surface water data processing.....	24
Table 4.5	Minimum data rules for surface water analysis.....	25
Table 5.1	Derivation of model response variables.....	37
Table 5.2	Summary of candidate predictor variables	40
Table 5.3	Summary of final land use models	42
Table 5.4	Summary of confidence classes.....	43
Table 6.1	Summary of confidence classes.....	46
Table 8.1	Summary of methodological refinements for Step 1	54
Table 8.2	Summary of methodological refinements for Step 2	55
Table 8.3	Summary of methodological refinements for Step 3	56
Table A.1	Critical values ($P = 1\%$) of the t_{\max} statistic.....	60

List of Figures

Figure 3.1	Rules for categorising monitoring sites.....	17
Figure 4.1	Historical trend and future forecast of 95 th percentile nitrate concentration using quantile regression, and comparison to the 11.29 mg N/l (50 mg NO ₃ /l) standard (red dashed line).....	30
Figure 4.2	Assessment of current 95 th percentile TIN concentration at surface water monitoring points	33
Figure 4.3	Assessment of future 95 th percentile TIN concentration at surface water monitoring points	34
Figure 5.1	Predicted 95 th percentile TIN concentration in main stem rivers	44
Figure 5.2	Predicted 95 th percentile TIN concentration in tributary and headwater streams.....	45
Figure 6.1	Evidence matrix used to determine if water body is subject to nitrate pollution.....	47
Figure 6.2	Assessment of surface water nitrate pollution risk at a local (water body) scale	49
Figure 6.3	Assessment of surface water nitrate pollution risk at a catchment scale	50
Figure 7.1	Evidence matrix used to determine if water body is subject to nitrate pollution.....	52

Figure C.1	Coefficients and summary statistics for catchment-scale TIN model	66
Figure C.2	Plot of predicted vs observed 95 th percentile TIN concentrations for the catchment-scale model	67
Figure C.3	Diagnostic residual plots for catchment-scale TIN model.....	67
Figure C.4	Coefficients and summary statistics for catchment-scale TON model.....	68
Figure C.5	Plot of predicted vs observed 95 th percentile TON concentrations for the catchment-scale model	69
Figure C.6	Diagnostic residual plots for catchment-scale TON model	69
Figure C.7	Coefficients and summary statistics for water body-scale TIN model	70
Figure C.8	Plot of predicted vs observed 95 th percentile TIN concentrations for the water body-scale model.....	71
Figure C.9	Diagnostic residual plots for water body-scale TIN model.....	71
Figure C.10	Coefficients and summary statistics for water body-scale TON model.....	72
Figure C.11	Plot of predicted vs observed 95 th percentile TON concentrations for the water body-scale model.....	73
Figure C.12	Diagnostic residual plots for water body-scale TON model	73

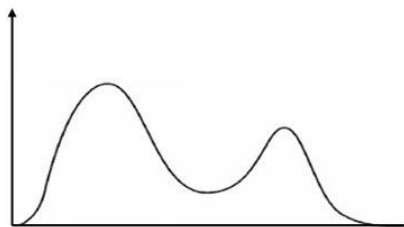
Glossary

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 sites 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>
(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 p^{th} percentile is the smallest value such that at least $p\%$ of the items in</p>

the population are no larger than it. For example, the 95th percentile is the value that exceeds 95% of the population and is exceeded by 5% of the population.

Quantile regression

Quantile regression is a statistical technique that explores how one or more independent variables influence a specified percentile value of the response variable. In contrast to conventional linear regression, which seeks to explain variation in the mean of the response variable, quantile regression can explore how the occurrence of high values (high nitrate concentrations in this case) changes over time.

Regression analysis

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).

Statistical significance

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.

Standard deviation

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.

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 NO₃/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 surface water monitoring and land use data. NRW intends to review these lines of evidence in the light of 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 eight steps in the surface water NVZ Review process analysis were as follows:

1. identification of surface freshwaters for analysis;
2. statistical analysis of surface water quality monitoring data;
3. land use modelling of nitrate pollution in surface waters;
4. combining the evidence from monitoring and modelling;
5. identifying land draining to polluted surface waters;
6. identification of critical monitoring sites unduly influenced by point sources;
7. removal of previous surface water NVZs; and,
8. using local knowledge to refine the draft designations.

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, 2, 3, 4 and 7.

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 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 1st 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 and/or de-designating existing NVZs.

As part of the 2017 NVZ Review, NRW commissioned WRc to undertake an up-to-date national assessment of nitrate pollution risk in surface waters (i.e. rivers and land drains, but excluding canals, lakes, reservoirs and estuaries) in Wales. This document describes the methodology used to generate and combine the lines of evidence used in the surface water assessment. (The methodologies used to assess pollution risks in groundwaters and eutrophic waters are different and are documented elsewhere.) Subsequent work undertaken by NRW to gather local evidence and ground-truth the proposed NVZ (de-)designations is not described 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 NO₃/l nitrate.¹
- Groundwater which contains or could contain, if preventative action is not taken, more than 50 mg NO₃/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 seven Sections. Section 2 provides an overview of the eight steps in the surface water assessment methodology. Sections 3 to 7 provide full details for Steps 1, 2, 3, 4 and 7, respectively. Finally, Section 8 highlights and discusses refinements made to the previous (2013) NVZ Review methodology. Further technical details are presented in Appendix A, Appendix B and Appendix C.

¹ 50 mg/l nitrate as NO₃ is equivalent to 11.29 mg/l nitrate as N.

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 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. These refinements are highlighted and discussed in Section 8.

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 surface waters adopted a weight of evidence approach, which combined observed evidence of nitrate pollution from monitoring data with evidence of nitrate pressure based on catchment land use. The results of the Review were used to identify new surface water NVZs and to confirm the designation status of existing surface water NVZs. The eight main steps were:

1. identification of surface freshwaters for analysis;
2. statistical analysis of surface water quality monitoring data;
3. land use modelling of nitrate pollution in surface waters;
4. combining the evidence from monitoring and modelling;
5. identifying land draining to polluted surface waters;
6. identification of critical monitoring sites unduly influenced by point sources;
7. removal of previous surface water NVZs; and,
8. using local knowledge to refine the draft designations.

This Section provides an overview of the assessment methodology by briefly describing each step in turn. Sections 3 to 7 give further details of Steps 1, 2, 3, 4 and 7, respectively.

2.2 Step 1 - Identification of surface freshwaters for analysis

Wales is divided into river water bodies for the purposes of assessing ecological status under the Water Framework Directive. In keeping with the role of the Nitrates Directive as a basic measure under the Water Framework Directive (WFD), these water bodies were used as the basic unit of assessment for this NVZ Review.

All rivers and streams within each water body were deemed to be 'surface waters'. Canals, lakes, reservoirs, estuaries and coastal waters were not included in the assessment.

Evidence of nitrate pollution was assembled for each individual water body and, on the basis of that evidence, generally the whole water body was either designated or not designated as a surface water NVZ.

Over 65% of the river water bodies contained at least one monitoring site from which regular samples were taken to measure water quality. Monitoring sites were categorised as (i) main stem, (ii) tributary, (iii) headwater or (iv) canal/estuary according to the connectivity of the river reaches they were located on, and the geographic origin of the waters being sampled. Main stem and tributary/headwater monitoring sites were treated differently when (i) modelling land use influences on river water quality (see Step 3; Section 5) and identifying land draining to polluted waters (see Step 5).

Section 3 gives further details of how surface waters were defined and monitoring sites categorised.

2.3 Step 2 - Statistical analysis of surface water quality monitoring data

Water quality monitoring data was assembled for all routine surface water monitoring sites operated by NRW and Dŵr Cymru Welsh Water.

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. Ammonium is rapidly oxidised to nitrate under normal riverine conditions. To assess the contribution of ammonium to observed nitrate pollution, a parallel analysis was performed using total oxidised nitrogen (TON), which comprises just nitrate and nitrite.²

² The term 'nitrate' is subsequently used in this report to refer to nitrogenous compounds (expressed as TIN or TON), of which nitrate is a major component.

The data was filtered to remove sites which did not have suitable data for the assessment. These checks included removing sites which:

- had fewer than 19 water quality samples in total;
- contained data from fewer than 5 calendar years (i.e. too few years to average out year to year variation in water quality); or
- had no water quality samples for the period 2009-2014 (i.e. no recent monitoring data).

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

- the 95th percentile TIN concentration currently exceeds 50 mg NO₃/l; or,
- the 95th percentile TIN concentration is likely to exceed 50 mg NO₃/l in the future, assuming no preventative action is taken.

If either the current or future 95th percentile concentration exceeded 50 mg NO₃/l, the monitoring site was considered to have failed the assessment. The level of confidence in the result was recorded as one of six classes (Table 2.1).

Table 2.1 Summary of confidence classes

Class	Description
1	At least 95% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
2	At least 75% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
3	At least 50% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
4	At least 50% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
5	At least 75% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
6	At least 95% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l

The statistical methods used to assess current and future status at each monitoring site depended upon the amount of data available. The assessment of current status was based on analysis of data from the last six calendar years (i.e. 2009-2014) or, where less data was available, on a statistical extrapolation of the historical (1990-2014) trend to predict the 95th percentile concentration in mid-2015. The assessment of future status used the same statistical extrapolation method to predict the 95th percentile concentration in mid-2020.

The results of the statistical trend analysis were taken forward to be combined with land use modelling evidence in Step 4 (see Section 6).

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

2.4 Step 3 - Land use modelling of nitrate pollution in surface waters

Land use modelling was used to provide an additional assessment of the risk of pollution to river water bodies. Statistical models were developed to explain spatial variation in observed nitrate concentrations at main stem and tributary/headwater monitoring sites as a function of land use and known sources of nitrate. The models were used to predict nitrate concentrations in main stem rivers and in tributary and headwater streams. These results were used to:

- assess the risk of nitrate pollution in unmonitored water bodies;
- improve confidence in the assessment for monitored water bodies; and,
- help assess the significance of agricultural sources of nitrate to any pollution identified.

The scale at which land use influences water quality depends upon the category of river. Tributary and headwater streams are influenced only by land use within the immediate water body, whereas main stem rivers are influenced by land use in the immediate water body and those upstream. Statistical models were therefore developed separately for these two categories of rivers, using predictor variables at an appropriate spatial scale.

The response variable used in the models was the current 95th percentile TIN or TON estimated in Step 2. Spatial variation in nitrate concentration was modelled as a function of a suite of potential explanatory variables representing nitrate losses from agriculture, urban areas, point source discharges, septic tanks and historic landfill sites.

The calibrated models were applied to make predictions of nitrate concentration at two spatial scales:

- water body – information on land use within an individual water body was used to predict the 95th percentile TIN concentration in tributary/headwater streams arising within the water body; and,
- catchment – information on land use across the upstream catchment area was used to predict the 95th percentile TIN concentration in the main stem river in the water body.

Predictions were made for all water bodies and catchments that had a full complement of land use data and had an area of at least 20 km².

To help assess the contribution that ammonia makes to observed pollution, separate models and predictions were also derived for TON.

If the land use models predicted that the 95th percentile TIN concentration exceeded 50 mg NO₃/l, the water body was considered to have 'failed' the modelling test. The level of confidence in the modelling assessment was recorded as one of six classes as described in Table 2.1.

The results of the land use model were taken forward to be combined with monitoring data evidence in Step 4 (see Section 6).

Section 5 describes in greater detail the datasets used to build the land use models, the statistical modelling approach used, and how the model was used to predict the risk of nitrate pollution.

2.5 Step 4 – Combining the evidence from monitoring and modelling

Steps 2 and 3 generated four lines of evidence to judge whether or not a water body was subject to nitrate pollution:

- A. present day 95th percentile TIN concentration from monitoring data;
- B. trend-predicted future 95th percentile TIN concentration from monitoring data;
- C. 95th percentile TIN concentration predicted from water body land use; and,
- D. 95th percentile TIN concentration predicted from catchment land use.

Each line of evidence was assessed to gauge whether or not the 95th percentile concentration exceeded 50 mg NO₃/l, and scored on a 1 to 6 scale, with 1 being a highly confident pass and 6 being a highly confident fail. A score of 0 was assigned to water bodies that lacked monitoring data or a reliable land use model prediction. A and B were combined by taking the highest (i.e. worst) score for each monitoring site.

The evidence from the monitoring data and the land use model was combined to assess the strength of evidence for nitrate pollution in each water body. Two tests were conducted for every water body:

- the worst performing tributary or headwater monitoring site in the water body (if any) was assessed in conjunction with the water body land use model prediction (if available); and,
- the worst performing main stem monitoring site in the water body (if any) was assessed in conjunction with the catchment-scale land use model prediction (if available).

A water body was considered to be polluted if the local test and/or the catchment-scale test produced evidence of pollution. In marginal cases, local knowledge and detailed data quality checks were used to provide an additional level of scrutiny and to help inform the final decision. Step 8 gives further details of the type of factors taken into consideration at the local ground-truthing workshops.

Section 6 details the evidence matrix used to categorise the strength of evidence for nitrate pollution.

2.6 Step 5 - Identifying land draining to polluted surface waters

Individual surface water bodies were identified as polluted or not in accordance with Step 4. The land draining to each polluted water body was determined according to:

1. if the failing monitoring site was on the main stem of the river, or on a tributary/headwater stream; or,
2. if the model prediction was based on land use in the water body, or on land use in the upstream catchment.

The spatial extent of the designation depended upon whether or not the pollution was confined to tributary/headwater streams and caused by local pressures, or if the pollution was present in the main stem and caused by pressures acting across the upstream catchment area. Where evidence of pollution was based solely on the land use model prediction, a more conservative approach to upstream designation was adopted.

Water bodies were the basic unit of assessment and designation. In most cases the whole water body was either designated or not designated. There were two exceptions to this rule:

1. where the evidence indicated marked variations in the level of pollution on tributaries compared to the main river; and,
2. where the water body spanned two or more hydrologically separate areas.

2.7 Step 6 – Identification of monitoring sites unduly influenced by point sources

Parts of Wales have substantial populations and nitrogenous compounds are discharged from waste water treatment works (WWTWs) and other anthropogenic sources into surface waters. Many water bodies have one or more discharges from WWTWs. Except in headwaters, surface waters inevitably have multiple upstream discharges and as a result these will contribute some nitrogenous compounds to the receiving water. Monitoring data collected at the majority of sample locations in Wales will, therefore, include some non-agricultural pollution.

A small number of surface waters in England that were designated in previous NVZ Reviews were subsequently removed on appeal because the Appeals Panel judged that the monitoring sites that indicated a polluted water may have been disproportionately influenced by upstream effluent discharges. In some cases the Appeals Panel suggested that the monitoring locations were not representative of agricultural pollution.

The Nitrates Directive addresses nitrate pollution from agricultural sources, but no clear guidance is available on what minimum contribution from agriculture is required to justify an NVZ designation. A 2005 European Court of Justice (ECJ) ruling (Case C 221/03, Commission v Kingdom of Belgium) deemed that a 17% contribution to pollution from agriculture was not insignificant when considering designation of NVZs.

Since the 2008 NVZ Review, NRW has improved its knowledge of contributions to pollution from agriculture and other sources by developing spatially resolved datasets on sources of nitrogenous compounds and land use models to help interpret surface water monitoring data.

Drawing on this knowledge, a methodology was developed to collate and assess evidence on the influence of non-agricultural sources to the water quality at monitoring locations downstream of effluent discharges. The methodology was applied to all polluted surface water monitoring sites that were critical for justifying the continuation of existing surface water NVZ designations, or justifying the designation of new NVZs.

A two stage process was adopted:

1. A *screening* test to identify critical surface water monitoring sites, and hence water bodies, likely to be most influenced by NRW consented effluent discharges.
2. A set of *exclusion criteria* were used to exclude from the assessment those monitoring sites where nitrate pollution is localised and/or has a substantial non-agricultural contribution.

A further screening step was the regional workshops, where any potentially problematic sites were identified for further assessment and possible exclusion (see Step 8).

The outcome was an evidence-based decision as to whether or not the monitoring site was influenced by an effluent discharge and the significance of that influence relative to agricultural sources of nitrate.

2.8 Step 7 - Removal of previous surface water NVZs

The iterative designation of NVZs under the Nitrates Directive inevitably results in discrepancies between areas of existing designation and potential new designations. In particular, the use of monitoring data means that each Review will almost inevitably give rise to new candidate areas for designation and, at the same time, show some previously

designated areas for which the available evidence no longer meets the criteria for designation.

The following broad criteria were used to justify removal of previous surface water NVZs:

- 1. monitoring data demonstrates that the water bodies are not (or are no longer) polluted;
- 2. agricultural land use is low risk as a source of nitrate;
- 3A. improvement in water quality is sustained over at least two NVZ Review periods;
- 3B. the likely cause(s) for the water quality improvement can be determined;
- 4. land previously identified as draining to a polluted water can, on the basis of improved data, reasonably be determined to not drain to a polluted water.

In essence, criteria (1) and (2) were used to identify a subset of water bodies that were previously designated but now show convincing evidence that waters are not or no longer polluted. The relevant data, plus other information for each of these water bodies and their catchments, was considered on a case-by-case basis, with reference to criteria (3) and (4).

Section 7 gives further details of how the criteria were developed and implemented.

2.9 Step 8 - Using local knowledge to refine draft designations

A series of skype sessions with key NRW experts and local staff was undertaken to:

- capture local information and understanding from NRW experts in surface water quality;
- seek local feedback on certain prioritised issues identified through the work process; and,
- quality assure the draft designations.

The following situations were prioritised for discussion at the workshops:

- potential areas of new surface water designation;
- areas that were designated in 2002 but passed in 2008; and,
- surface waters that failed in 2008 but were successfully appealed.

Attendees were also invited to highlight any polluted waters that they thought had been missed by the Review, or any unpolluted waters that they believed should not be designated.

A series of meetings were held with the farming unions to discuss the review process and its proposed designations at each stage of the review. National Farmers Unions were invited to supply information and evidence to support the findings.

3. Identification of surface freshwaters for analysis

3.1 Definition of surface waters

The land area of Wales is divided into 861 river water bodies for the purposes of assessing ecological status under the Water Framework Directive. In keeping with the role of the Nitrates Directive as a basic measure under the Water Framework Directive (WFD), these water bodies were used as the basic unit of assessment for this NVZ Review.

All rivers and streams within each water body were deemed to be 'surface waters'. Canals were excluded from the definition of surface waters. Estuarine and coastal waters, although technically defined as 'surface waters' under the WFD, were also excluded. Standing surface waters (lakes and reservoirs) are the subject of a separate (eutrophic waters) part of the NVZ Review process and are not included in this assessment.

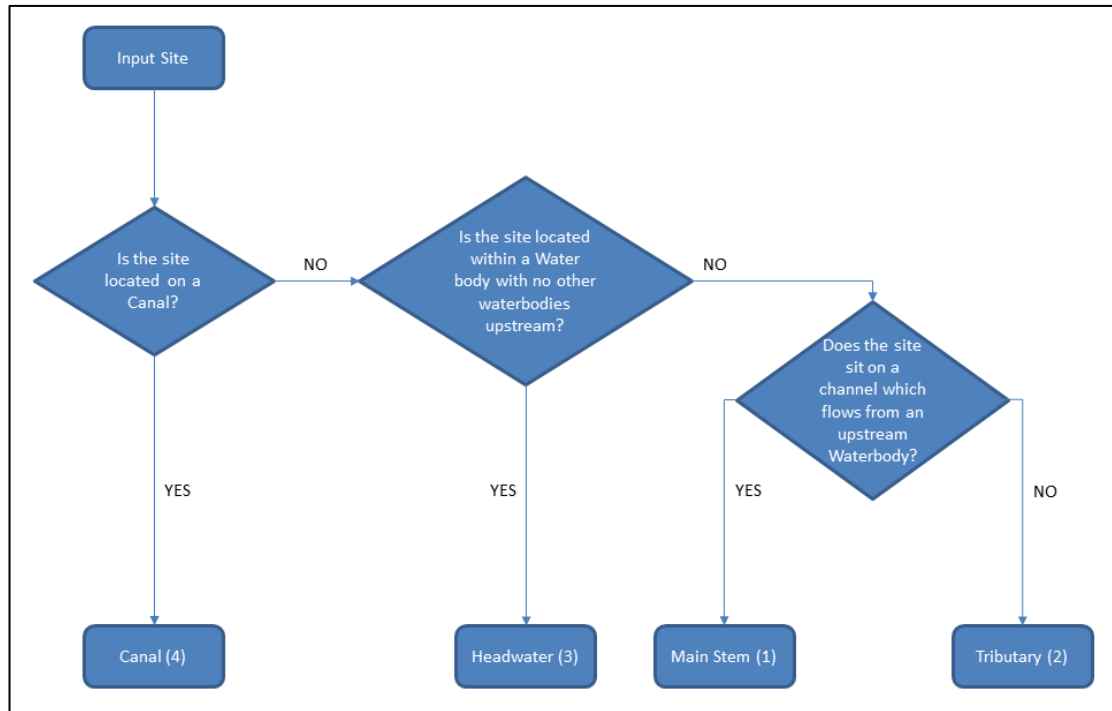
Evidence of nitrate pollution was assembled for each individual water body and, on the basis of that evidence, generally the whole water body was either designated or not designated as a surface water NVZ.

3.2 Categorisation of monitoring sites

Over 65% of river water bodies contained at least one monitoring site from which regular samples were taken to measure water quality. Using the WFD Cycle 2 water bodies and NRW's Digital River Network, monitoring sites were categorised according to the connectivity of the river reaches they were located on, and the geographic origin of the waters being sampled. Table 3.1 defines the four categories and Figure 3.1 sets out the rules used to categorise monitoring sites.

Table 3.1 Definitions of monitoring site categories

Category	Definition
1. Main Stem	The monitoring site falls on a main channel watercourse that flows from an upstream water body.
2. Tributary	The monitoring site falls on a secondary channel watercourse which arises within the water body and flows into a main stem watercourse in the same water body.
3. Headwater	The monitoring site falls on a main or secondary channel watercourse in a water body that has no upstream water bodies.
4. Canal/estuary	The monitoring site is located on a canal or estuary

Figure 3.1 Rules for categorising monitoring sites

Headwater monitoring sites were identified automatically in ArcGIS using the water body connectivity to identify river water bodies that were not downstream of any others.

Canal monitoring sites were identified by two means, firstly reviewing the monitoring site location names for the “canal” keyword and also checking proximity to canals from the OS Open Watercourse layer. One site was found to be located within an estuary.

All remaining monitoring sites were reviewed manually and assigned as either main stem or tributary.

Monitoring sites located on canals and estuaries were excluded from the NVZ assessment. Main stem and tributary/headwater monitoring sites were treated differently when (i) modelling land use influences on river water quality (see Step 3; Section 5) and identifying land draining to polluted waters (see Step 5).

4. Statistical analysis of surface water quality monitoring data

4.1 Introduction

Water quality monitoring data was assembled for all routine surface water monitoring sites operated by NRW and Dŵr Cymru Welsh Water.

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 surface water monitoring site with sufficient data was analysed to determine whether or not:

- the 95th percentile TIN concentration currently exceeds 50 mg NO₃/l; or,
- the 95th percentile TIN concentration is likely to exceed 50 mg NO₃/l in the future, assuming no preventative action is taken.

If either the current or future 95th percentile concentration exceeded 50 mg NO₃/l, the monitoring site was considered to have failed the assessment. The level of confidence in the result was recorded as one of six classes (Table 4.1).

Table 4.1 Summary of confidence classes

Class	Description
1	At least 95% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
2	At least 75% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
3	At least 50% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
4	At least 50% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
5	At least 75% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
6	At least 95% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l

4.2 Source dataset

NRW extracted from its Water Information Management System (WIMS) database all water quality data collected at routine surface water monitoring sites in Wales between 1990 and 2015. The determinand codes extracted were:

- 0111 (Ammoniacal nitrogen as N);
- 0116 (Total oxidised nitrogen as N);
- 0117 (Nitrate as N);
- 0118 (Nitrite as N); and,
- 9880 (Nitrate as NO₃).

The database extract contained monitoring data collected by NRW and Dŵr Cymru Welsh Water. Measurements with purpose codes relating to reactive sampling of pollution incidents and monitoring of waste sites were not provided as they were thought to be unrepresentative of normal water quality. Table 4.2 shows which purpose codes were contained within the data provided by NRW.

Table 4.2 Purpose codes included in surface water analysis

Purpose Code	Purpose Code Description	Included?
CA	Compliance audit (permit)	Yes
CF	Compliance formal (permit)	No
CI	Statutory audit (operator data)	Yes
CO	Water quality UWWTD monitoring data	Yes
CS	Water quality operator self-monitoring compliance data	Yes
IA	IPPC/IPC monitoring (Agency audit - permit)	Yes
IF	IPPC/IPC monitoring (formal sample)	No
II	IPPC/IPC monitoring (Agency investigation)	Yes
IO	IPPC/IPC monitoring (operator self-monitoring data)	Yes
IT	Instrument trial	No
MI	Statutory failures (follow ups at designated points)	No
MN	Monitoring (national Agency policy)	Yes
MP	Environmental monitoring (GQA & RE only)	Yes

Purpose Code	Purpose Code Description	Included?
MS	Environmental monitoring statutory (EU directives)	Yes
MU	Monitoring (UK govt policy - not GQA or RE)	Yes
PF	Planned formal non-statutory (permit/env mon)	No
PI	Planned investigation (operational monitoring)	No
PN	Planned investigation (national Agency policy)	Yes
SI	Statutory failures (follow ups at non-designated points)	No
UF	Unplanned reactive monitoring formal (pollution incidents)	No
UI	Unplanned reactive monitoring (pollution incidents)	No
WA	Waste monitoring (Agency audit - permit)	No
WF	Waste monitoring (formal sample)	No
WI	Waste monitoring (Agency investigation)	No
WO	Waste monitoring (operator self-monitoring data)	No
XO	External organisation monitoring	Yes
ZZ	Unspecified at time of WIMS load	No

The raw dataset contained 2,568 unique monitoring sites and 262,223 unique samples.

4.3 Data processing

4.3.1 **Monitoring sites**

A unique Site ID was created for each NRW monitoring site by combining the Site ID and Region fields.

The co-ordinates of each monitoring site were checked for missing or inconsistent eastings and northings. Any anomalies were sent to NRW for checking which resulted in the removal of 13 sites where locations could not be specified.

The following meta-data were tabulated for each monitoring site:

- unique site ID;
- site name;
- co-ordinates (eastings and northings);
- type of water monitored (river);

- site operator (NRW);
- sampling point type (where applicable); and
- region (where applicable).

As far as possible, the dataset was filtered to exclude monitoring sites that:

- sampled water after it had been treated to remove nitrate (and so is not representative of raw water quality);
- sampled water that had been blended from multiple sources;
- sampled water from canals (which may contain water from a number of WFD catchments and not be representative of a defined geographic area) or estuaries; and
- sampled water from lakes and reservoirs (which are subject to a separate methodology to identify eutrophic waters).

To achieve this, the NRW dataset was filtered according to the sampling point types listed in Table 4.3.

Table 4.3 Sampling point types included in surface water analysis

Sampling Point Type Code	Sampling Point Description	Included
F1	Freshwater - RQO RE1	Yes
F2	Freshwater - RQO RE2	Yes
F3	Freshwater - RQO RE3	Yes
F4	Freshwater - RQO RE4	Yes
F5	Freshwater - RQO RE5	Yes
F6	Freshwater - Non Classified River Points	Yes
FA	Freshwater - Lakes/Ponds/Reservoirs	No
FB	Freshwater - River Transfer	Yes
FC	Freshwater - Comparative Inlet Points	Yes
FD	Freshwater - River Augmentation	Yes
FF	Freshwater - Canal - RQO RE2	No
FG	Freshwater - Canal - RQO RE3	No
FH	Freshwater - Canal - RQO RE4	No

Sampling Point Type Code	Sampling Point Description	Included
FJ	Freshwater - Canals - Non Classified	No
FK	Freshwater - Land Drains	Yes
FL	Freshwater - Bathing Water	Yes
FZ	Freshwater - Unspecified	Yes

4.3.2 Samples

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

Samples collected before 1st January 1990 were excluded because they were not regarded as indicative of recent nitrate trends. Samples collected after 31st 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 25 years of water quality sampling data was analysed for each site.³

The data were screened to identify how many samples were taken from each site on each day. In most cases, a single water sample was taken from a given site on a given day, but multiple samples were sometimes taken. These could be a set of manual samples collected as part of a local investigation, or a series of samples collected by an auto-sampler. The presence of multiple samples potentially causes two problems. First, any analysis of nitrate concentrations is weighted towards days when more samples are taken. Second, if an auto-sampler is triggered to take samples during high flow events then the results are not representative of typical conditions (although they still provide useful information about water quality during periods of wet weather). As it was not possible to tell where and when flow-driven auto-samplers had been deployed, we adopted the following rules to minimise their prevalence in the data.

- Where three or more samples had been taken from the same site on the same day, all samples were excluded.
- Where two samples had been taken from the same site on the same day, the sample with the highest TIN reading was retained and the other sample excluded. If the two

³ By comparison, 35 years (1980-2014) of historical sampling data were used for the groundwater NVZ review. A shorter, 25 year, time period was necessary for the surface water assessment because changes in water quality generally occur more rapidly in surface waters than in groundwaters.

samples had the same TIN reading then the sample with the highest TON was retained and if the two samples had identical readings for both determinands the one with the highest sample number was retained.

If neither TON nor nitrate was measured (i.e. the sample only had measurements for nitrite and/or ammonium) then the sample was excluded.

4.3.3 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.

To make all determinands comparable, measurements recorded in mg/l as NO₃ were converted to mg/l as N by multiplying by 0.2258. Thus, 50 mg NO₃/l corresponds to 11.29 mg N/l.

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.

“Less than” values were treated using the standard 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.

4.3.4 Summary of cleaned dataset

The surface water data processing is summarised in Table 4.4.

Table 4.4 Summary of surface water data processing

Data processing stage		No. of sites	No. of samples
Raw data	NRW dataset	2,568	262,223
Data exclusion rules	Out of scope SMPT_TYPE	0	0
	Monitoring site outside Wales	0	0
	Sample >2015	0	0
	Sample <1990	0	0
	Zero or negative readings	0	0 ¹
	Three or more samples on the same day	0	172
	Nitrite and/or ammonium only	404	35,894
	Two samples on the same day	0	1031
	No National Grid Reference	13	152
	Total excluded	417	37,249
	Final dataset	2,150	224,994

¹ individual determinand readings deleted but did not constitute complete samples.

4.4 Data analysis methods

4.4.1 Overview

The statistical methods used to assess current and future status at each monitoring site depended upon the amount of data available, as set out in Table 4.5. The assessment of current status was based on analysis of data from the last six calendar years (i.e. 2009-2014) or, where less data was available, on a statistical extrapolation of the historical (1990-2014) trend to predict the 95th percentile concentration in mid-2015. The assessment of future status

used the same statistical extrapolation method to predict the 95th percentile concentration in mid-2020.⁴

To ensure that the results of the analysis were based on sound monitoring evidence, sites were not assessed if they:

- had fewer than 19 water quality samples in total (the minimum number of samples required to estimate a 95th percentile value using the Weibull method);
- contained data from fewer than 5 calendar years (i.e. too few years to average out year to year variation in water quality); or
- had no water quality samples for the period 2009-2014 (i.e. no recent monitoring data).

Table 4.5 Minimum data rules for surface water analysis

Rule	Current status assessment	Future status assessment	Number of sites
At least 19 samples and at least 5 years of data for period 2009-2014	Weibull estimate of 95 th percentile for 2009-2014	Quantile regression forecast of 95 th percentile in mid-2020	468
At least 19 samples and at least 5 years of data for period 1990-2014 and at least 1 sample for period 2009-2014	Quantile regression forecast of 95 th percentile in mid-2015	Quantile regression forecast of 95 th percentile in mid-2020	509
Fewer than 19 samples, fewer than 5 years of data for period 1990-2014, or no samples for period 2009-2014	No assessment	No assessment	1,173
TOTAL			2,150

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

4.4.2 Data screening

Prior to analysis, exceptionally high and low concentration measurements at each site were identified using a Multiple Outlier Test (described in 0) and excluded from subsequent

⁴ This time horizon (2020) is sooner than that used for the groundwater NVZ Review (2027) because changes in water quality generally occur more quickly in surface waters than in groundwaters.

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 site. 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 4.4.5 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 characterising the historical trend.

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.

4.4.3 Weibull method

The Weibull method was used to estimate the current 95th percentile concentration at monitoring sites that had at least 19 water quality samples and at least five years of 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 (e.g. it does not require the data to follow a normal or log-normal distribution). The Weibull protocol is also relatively insensitive to outliers and provides an assessment of conditions over a six year period, which makes the results less sensitive to random, short-term fluctuations in water quality.

The Weibull method uses the r^{th} ranked value within the observation dataset to provide an estimate of the 95th 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 95th percentile. Conservative 90% and 50% confidence intervals were calculated using binomial distribution theory, as described in the Environment Agency Codes of Practice for Data Handling (Ellis *et al.*, 1993).⁵ If the lower 90% confidence limit exceeded 50 mg NO₃/l, the monitoring site was deemed to have failed the test with high (95%) confidence; if the lower 50% confidence limit exceeded 50 mg NO₃/l, the monitoring site was deemed to have failed the test with medium (75%) confidence; if the 95th percentile estimate exceeded 50 mg NO₃/l, the monitoring site was deemed to have failed the test with low (50%) confidence.

⁵ A minimum of 28 and 59 samples are required to calculate the upper 50% and 90% confidence limits, respectively, so for some sites it was possible only to demonstrate with medium or low confidence that the 95th percentile was below the threshold.

The results were visualised as shown in Figure 4.1, where:

- the dark green solid horizontal line 2009-2014 marks the 95th percentile estimate;
- the green shaded band indicates the 50% confidence interval around the 95th percentile estimate; and
- the wider, light green shaded band indicates the 90% confidence interval around the 95th percentile estimate.

4.4.4 Quantile regression method

For all sites that met the minimum data requirements (at least 19 samples and at least 5 years of data for period 1990-2014 and at least 1 sample for period 2009-2014), quantile regression was used to characterise trends in water quality between 1990 and 2014. The historical trend was then extrapolated using Monte Carlo simulation to estimate the current (mid-2015) 95th percentile concentration at monitoring sites that had too few samples for the Weibull method, and to forecast the 95th percentile concentration in mid-2020.

The approach was based closely on the 'AntB' method used to analyse trends at groundwater monitoring sites. AntB uses multiple regression to characterise between-year and seasonal variation in mean concentration and then applies a conversion factor to estimate the 95th percentile concentration. One drawback with this approach is that it assumes that the trend in the 95th percentile mirrors the trend in the mean concentration. This will not always be the case; for example, if the occurrence of occasional high nitrate readings reduces or increases over time then the 95th percentile concentration may decrease (or increase) with little change in mean concentration.

For this reason, a quantile regression was used in place of a classical linear regression to characterise directly any changes over time in the 95th percentile concentration. Quantile regression (Koenker and Hallock 2001) is a statistical technique intended to estimate, and conduct inference about, conditional quantile functions. Just as classical linear regression methods based on minimizing sums of squared residuals enables the estimation of means, quantile regression methods offer a mechanism for estimating the median, and the full range of other quantile functions, by minimising weighted sums of absolute residuals.

Quantile regression is a robust technique that makes no assumptions about the underlying probability distribution of the data. It is relatively insensitive to outliers. Outlying responses influence the fit in only so far as they are either above or below the fitted value, not how far above or below.

The statistical method comprises two main steps: trend analysis, and forecasting.

Trend analysis

A quantile regression model was fitted to the data to describe historical variation in the 95th percentile concentration at each monitoring site. Temporal variation in water quality was represented using a series of linear spline functions, which describe changes as a series of k straight lines connected by $k-1$ 'hinge points'. Each spline had to span at least 12 data points and at least three years. Three years was chosen to strike a balance between flexibility (a longer period would struggle to capture the timing of changes in water quality) and sensitivity (a shorter period would be unduly sensitive to short-term 'blips' in water quality). Hinge points were constrained to occur at the end of calendar years (to average out seasonal variation), and at least three years from the beginning or end of the time series (to avoid picking up short term trends in water quality).

Following preliminary analyses, it was found that the slopes of the splines at the ends of the time series could be sensitive to unusually high or low readings, so the first and last three years of data were replicated at the ends of the time series to stabilise the fitted spline slopes. The splines fitted to the replicated data were disregarded in the later forecasting phase.

The quantile regression trend analysis was undertaken using the 'quantreg' package (Koenker 2015), with one appropriately defined dummy variable per spline.

All the linear spline 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 quantile regression model has m splines. In general, some of these 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 $1/SE(j)$ and, using these weights, calculating the weighted mean (AvB) and weighted standard deviation (SDB) of the m spline slopes.

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 five splines;
- the average spline length (the mean number of years spanned per spline); and
- the quantile regression prediction of the 95th percentile at the end of the monitored period (τ) and its standard error (τ_{SE}).

Taking as its starting point a 95th percentile concentration drawn from the Normal distribution $N(\tau, \tau_{SE})$ and the final spline slope, the Monte Carlo simulation generates a projection of how the 95th percentile 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 average spline length as the time step, the process is continued until the required forecasting horizon is reached. Starting at the beginning of 2010, for example, if the average spline length was 3.5 years, a series of three slopes would be generated to predict the 95th percentile concentration in mid-2020.

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 95th percentile concentration in mid-2015 and mid-2020. The uncertainty in this estimate was quantified by ranking the 10,000 forecasts for that year and determining the 10th, 25th, 75th and 90th percentile values.

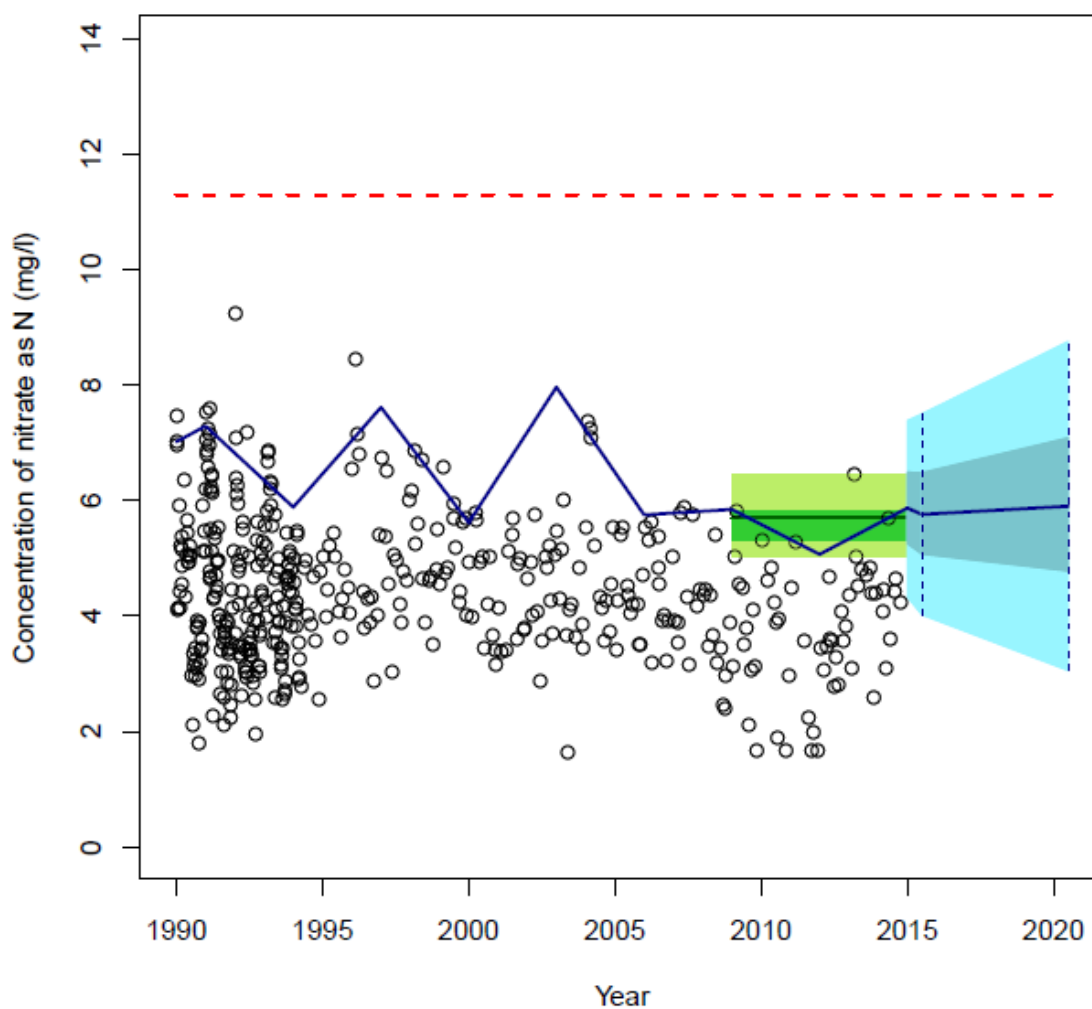
Figure 4.1 shows an example of a fitted quantile regression model and the forecasts generated from it. The red dashed line at 11.29 mg N/l indicates the threshold value and the solid blue zig-zag line indicates historical fluctuations in the 95th percentile TIN concentration. Beyond the end of the monitoring record in 2014:

- the blue line indicates the 95th percentile concentration forecast by Monte Carlo simulation;
- the blue shaded band indicates the 50% confidence interval around the 95th percentile estimate; and

- the wider, light blue shaded band indicates the 90% confidence interval around the 95th percentile estimate.

In this example, the 95th percentile concentration is forecast to be 5.8 mg N/l in mid-2015 (the first vertical dashed line) with a 90% confidence interval of 4.0-7.5 mg N/l, and 5.9 mg N/l in mid-2020 (the second vertical dashed line) with a 90% confidence interval of 3.1-8.8 mg N/l.

Figure 4.1 Historical trend and future forecast of 95th percentile nitrate concentration using quantile regression, and comparison to the 11.29 mg N/l (50 mg NO₃/l) standard (red dashed line)



4.4.5 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 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.

Visual checks of the Weibull estimates, quantile regression trends and future forecasts, plotted against the raw data, 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 11.29 mg N/l threshold;
- that passed the current and future status tests but had exceedances of the 11.29 mg N/l threshold since 2009; and
- where there was a large discrepancy between the Weibull and quantile regression estimates of the current 95th percentile concentration.

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.

Monitoring sites were flagged for NRW's attention where the results were judged not to be a reasonable representation of current and future water quality.

4.5 Summary of results

Current and future 95th percentiles TIN concentrations were estimated for a total of 977 surface water monitoring sites.⁶ These results were used to gauge whether or not the 95th percentile exceeded 50 mg NO₃/l at each monitoring site. Each site was scored on a 1 to 6 scale, with 1 being a highly confident pass and 6 being a highly confident fail. Figure 4.2 and Figure 4.3 map these scores for the current and future assessments, respectively.

⁶ Twelve of these sites were subsequently confirmed to be on canals or estuaries, and were excluded from the final assessment.

Figure 4.2 Assessment of current 95th percentile TIN concentration at surface water monitoring points

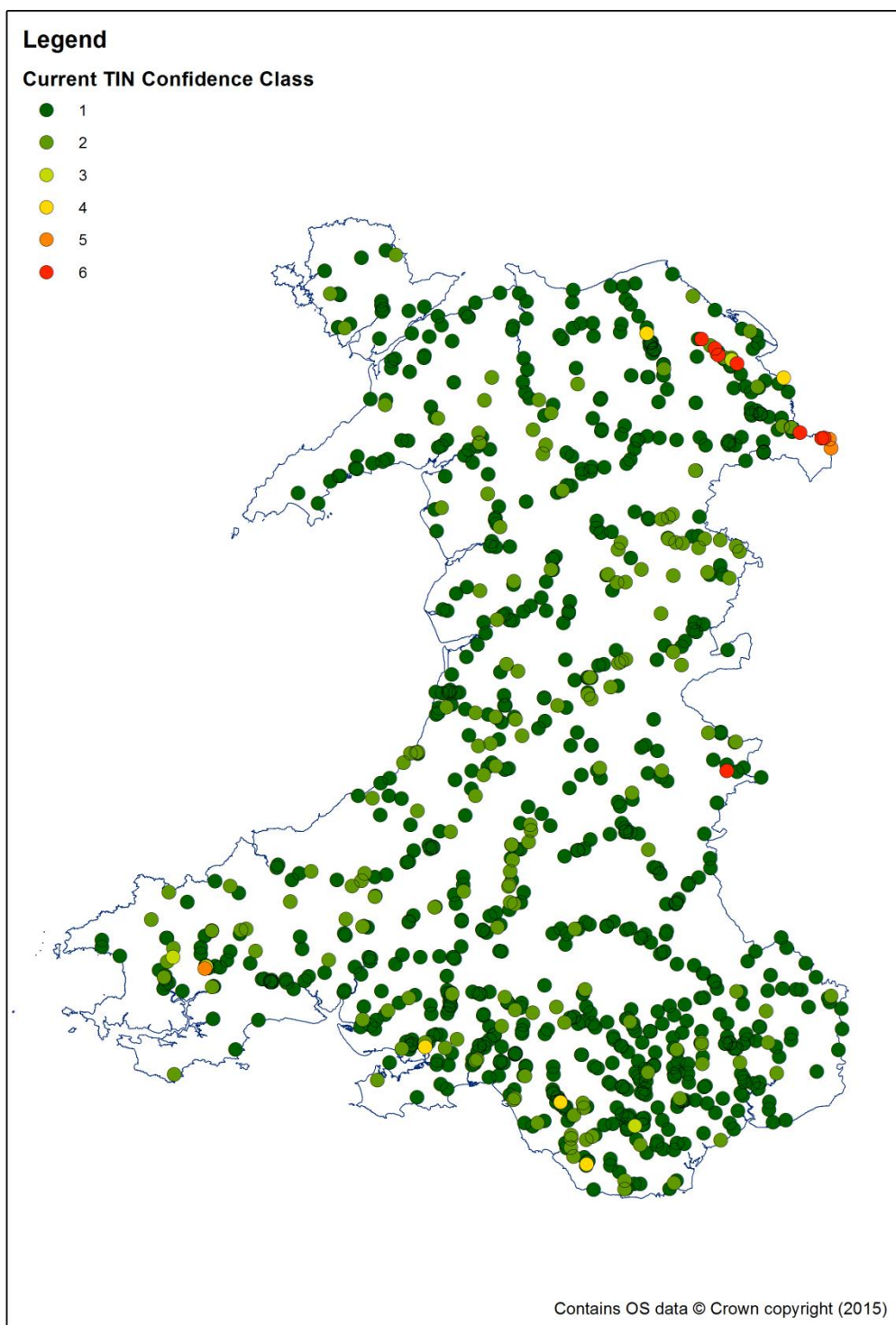
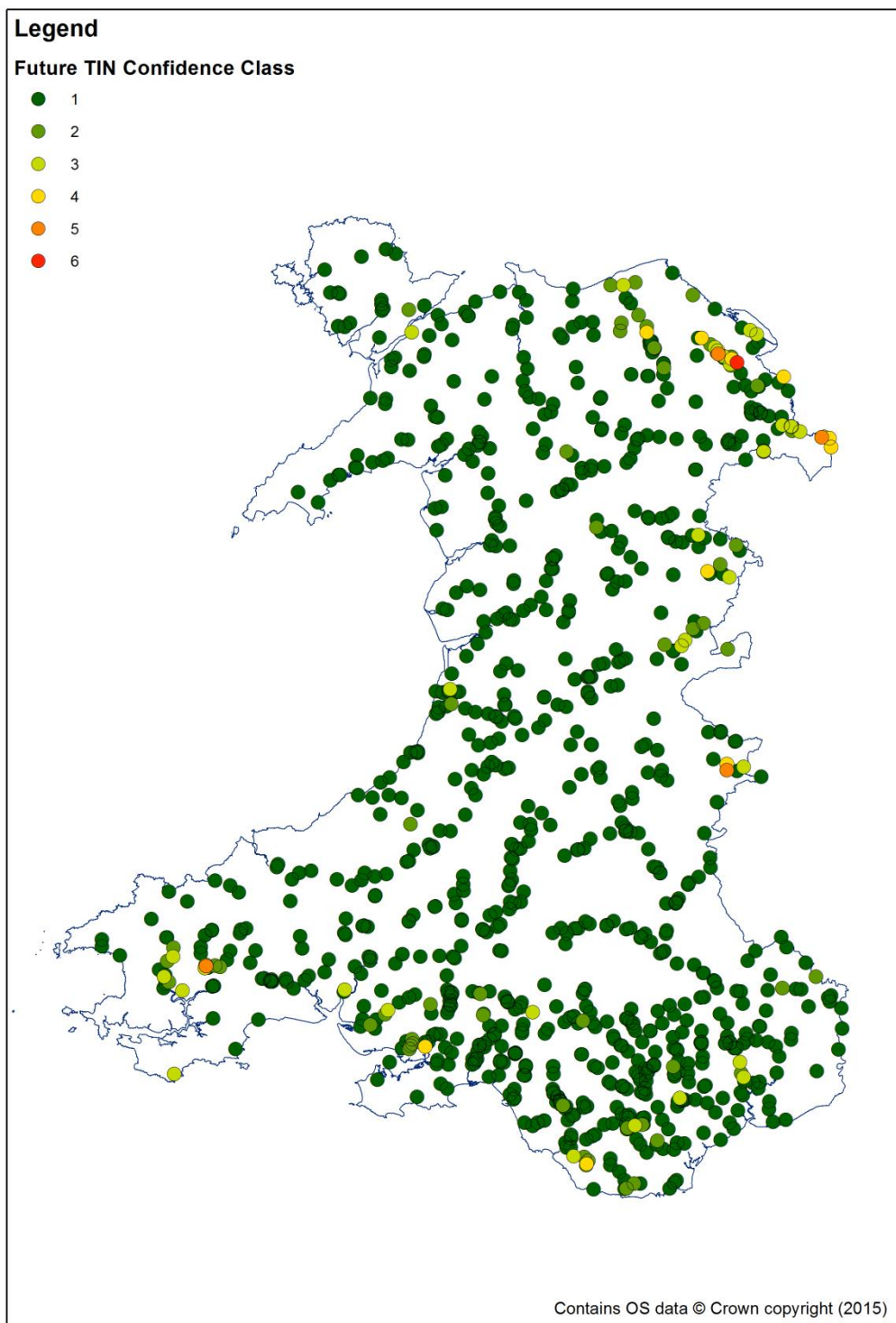


Figure 4.3 Assessment of future 95th percentile TIN concentration at surface water monitoring points



5. Land use modelling of nitrate pollution in surface waters

5.1 Introduction

Land use modelling was used to provide an additional assessment of the risk of pollution to river water bodies. Statistical models were developed to explain spatial variation in observed nitrate concentrations as a function of land use and known sources of nitrate. The models were used to predict nitrate concentrations in main stem rivers and in tributary and headwater streams. These results were used to:

- assess the risk of nitrate pollution in unmonitored water bodies;
- improve confidence in the assessment for monitored water bodies; and,
- help assess the significance of agricultural sources of nitrate to any pollution identified.

The scale at which land use influences water quality depends upon the category of river – i.e. whether it is a tributary/headwater or main stem (see Section 3.2 for definitions). Tributary and headwater streams are influenced only by land use within the immediate water body, whereas main stem rivers are influenced by land use in the immediate water body and those upstream. Statistical models were therefore developed separately for these two categories of rivers, using predictor variables at an appropriate spatial scale. To help assess the contribution that ammonia makes to observed pollution, separate models were also derived for TIN and TON. Thus, four models were produced in total:

1. TIN concentrations in main stem rivers;
2. TIN concentrations in tributary/headwater streams;
3. TON concentrations in main stem rivers; and
4. TON concentrations in tributary/headwater streams.

5.2 Derivation of response variables

The response variable used in the models was the 95th percentile TIN or TON concentration in each monitored water body. For those water bodies with a single monitoring site, the current (2015) 95th percentile estimate was used (see Section 4 for details of the trend analysis methods used); for those water bodies with more than one monitoring site, the current 95th percentile estimates were averaged. Excluded from the calculations were:

- monitoring sites located on canals (11 sites) and estuaries (1 site);
- monitoring sites where a strong downward trend in concentration resulted in a negative percentile estimate; and
- monitoring sites where the estimated 95th percentile concentration was highly uncertain (specifically, where the 50% confidence interval was wider than 10 mg N/l).

Also excluded from the calibration dataset were water bodies and catchments less than 20 km² in area, for which a reliable assessment of land use could not be derived. The choice of a 20 km² threshold was informed by a CEH analysis that suggested the variance in the 1 km² land-use datasets was reasonably stable down to 20 km² (Environment Agency, unpublished Science Report). At smaller scales:

- Some of the smallest water bodies did not appear to represent meaningful hydrological units.
- The rainfall and soils datasets used to model agricultural nitrate losses are based on large scale datasets and interpolated on a 1 km² grid. The process of interpolation can yield inaccurate results at a 1 km² scale.
- The resolution of the agricultural land use data (parish level statistics) used to model agricultural nitrate losses are converted into 1 km² grid values based on the contemporary land cover map. In doing so, farm level data is treated to prevent possible disclosure of commercially or other sensitive information. This means that the data can be inaccurate at scales of 3 x 3 km (i.e. 9 km²) or less.

Table 5.1 summarises how the data was processed to arrive at a final number of water bodies for analysis.

Table 5.1 Derivation of model response variables

Data processing step	Catchment scale model (main stem sites)		Water body scale model (tributaries and headwaters)	
	TIN	TON	TIN	TON
Number of candidate monitoring sites (excluding canals and estuaries)	344	344	621	621
Sites in water body or catchment < 20 km ²	2	2	231	231
Sites with imprecise percentile estimate	0	0	3	3
Sites with negative percentile estimate	0	1	0	0
Final number of monitoring sites	342	341	387	387
Unique water bodies for analysis	177	177	229	229

5.3 Derivation of explanatory variables

A variety of datasets were collated and processed to derive a series of potential explanatory variables representing nitrate losses from agriculture, urban areas, point source discharges, septic tanks and historic landfill sites.

5.3.1 Agricultural sources

Nitrate losses from agricultural land were 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 km² 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 (HER) (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.

First, the losses from the individual land use classes were summed to give a total annual nitrate load from agricultural sources on a 1 km² scale (kg N/km²/yr).

Next, the 1 km² outputs from the NEAP-N model were intersected in ArcMap with the WFD river water bodies. Where a 1 km² grid square included more than one water body, the agricultural nitrate load was distributed proportionally between each intersecting water body on the basis of fractional area. These loads were summed to calculate a total nitrate load for each water body and then further summed, based on water body connectivity, to calculate a total nitrate load for the whole upstream catchment area draining to each water body.

Finally, nitrate loads were converted to concentrations to standardise for differences between water bodies and catchments in area and rainfall. The summed water body and catchment-scale total loads (kg N/yr) were divided by an estimate of the HER (mm/yr on a 1 km² grid, converted to l/yr over the relevant water body or catchment area) to give a measure of nitrate concentration in agricultural run-off (mg N/l) for every water body and catchment.

5.3.2 Urban diffuse sources

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 fine scale data on the urban environment were not available. The urban diffuse estimates did not explicitly include nitrate from highway runoff, but road run-off in urban areas was included implicitly because the urban losses model is based on monitoring of nitrate loads within urban drains. The components that were included for the NVZ land use modelling are detailed in Appendix B3.

First, the losses from the individual components were summed to give a total annual nitrate load from urban sources (kg N/ha/yr) on a 1 km² scale, and then re-scaled to kg N/km²/yr.

Next, the 1 km² outputs were intersected in ArcMap with the WFD river water bodies. Where a 1 km² grid square included more than one water body, the urban nitrate load was distributed proportionally between each intersecting water body on the basis of fractional area. These loads were summed to calculate a total nitrate load for each water body and then further summed, using a knowledge of water body connectivity, to calculate a total nitrate load for the whole upstream catchment area draining to each water body.

Finally, nitrate loads were converted to concentrations to standardise for differences between water bodies and catchments in area and rainfall. The summed water body and catchment-

scale total loads (kg N/yr) were divided by an estimate of the HER (mm/yr on a 1km² grid, converted to l/yr over the relevant water body or catchment area) to give a measure of nitrate concentration in urban run-off (mg N/l) for every water body and catchment.

5.3.3 Point sources

Nitrogen inputs from consented point source discharges were estimated using data from NRW's WIMS database. Information was extracted on discharge location, consented flow and consented ammonia (NH_{3/4}) and nitrate concentrations for all discharge consents with numerical conditions. The data were processed to generate three alternative measure of point source pressure:

- the number of consented discharges;
- the total flow from all consented discharges;
- the aggregate N concentration from all consented discharges.

The annual flow (l/yr) from each discharge was estimated from the consented mean flow, or else the consented dry weather flow multiplied by 1.3, or else the consented maximum daily flow.

The annual nitrogen load (kg N/yr) from each discharge was calculated by multiplying the annual flow by the quality consent condition, either the 95th percentile concentration or half the maximum concentration. Although there were differences among sites in how the flow and quality consents were expressed, this approach still provides a useful measure of relative differences among water bodies in the loading to the environment.

Discharges were mapped to water bodies in GIS and the number of discharges counted for every water body and its associated upstream catchment area. Annual flows and loads were summed to calculate a total for each water body and then further summed, based on water body connectivity, to calculate a total for the whole upstream catchment area draining to each water body. The summed loads were then divided by the summed flows to estimate an aggregate point source concentration (mg N/l) for every water body and catchment area.

5.3.4 Septic tanks

NRW provided the locations of all Discharge Exemptions recorded on its WIMS database, the closest approximation for septic tank discharges. The total septic tank load to each water body (kg N/yr) was estimated by counting the number of discharge locations and multiplying this by an estimate of the annual nitrogen loss from a well-maintained septic tank. Specifically, we assumed that each septic tank serves 2.4 people and used a figure of 2.7kg N per person per year (Carey and Davison 2010). Loads were accumulated upstream and then

standardised by HER, as described in the agricultural sources section above, to estimate an aggregate concentration (mg N/l) from septic tanks for every water body and catchment area.

5.3.5 Historic landfills

Landfills may be significant local sources of ammonia. Data on location and extent of historic landfill was provided by NRW and used to calculate the percentage coverage of landfill sites within each water body and water body upstream catchment area.

5.4 Modelling methodology

The statistical distributions of the response and predictor variables were plotted and found to be highly skewed, so all variables were transformed using log to base 10 (i.e. $\log_{10}(x+a)$, where a is an appropriately-sized increment). This transformation was successful in:

- reducing the prevalence of outliers;
- improving the overall fit of the model (models with all variables log-transformed explained the highest proportion of variation in the data); and
- ensuring that the model residuals were normally distributed and had homogeneous variance.

Correlations between predictor variables were also calculated to identify any cases of high collinearity.

A multiple linear regression model was fitted to explain variation in water quality as a function of the land use predictor variables listed in Table 5.2.

Table 5.2 Summary of candidate predictor variables

Predictor Variable	Description
agriconc	Modelled nitrate concentration in agricultural run-off (mg N/l)
urbanconc	Modelled nitrate concentration in urban run-off (mg N/l)
septicconc	Estimated aggregate concentration from septic tank discharges (mg N/l)
pscount	Number of consented discharges
psconc	Estimated concentration from all consented discharges(mg N/l)
psflow	Estimated total flow from all consented discharges
landfill	Percentage of land area covered by landfill sites

Separate models were fitted to describe TIN and TON concentrations in main stem rivers (catchment scale) and tributaries/headwaters (water body scale). In each case, the process of selecting a final, preferred model was as follows.

First, an initial model containing agriconc, urbanconc, septicconc, psconc and landfill as predictors was fitted. The effect of log-transforming the response and/or predictor variables was explored to confirm that the choice of log-transformation was appropriate.

Next, three alternative measures of point source pressure (pscount, psconc and psflow) were tried in turn. The model explaining the greatest proportion of variation in the data was preferred and termed the 'full model'.

Finally, the full model was simplified using a step-wise model selection procedure with Akaike's Information Criterion (AIC)⁷ as the model selection criterion. This produced a final, parsimonious model that included only those predictors that made a useful contribution to the model. This model simplification process also helped exclude variables for which the direction of effect was counter-intuitive (i.e. higher values were associated with lower concentrations).

The predictive capability of the full and final models was assessed using a 10-fold cross-validation procedure. In this, a 10% of the calibration dataset was excluded in turn, the model refitted to the remaining data, and predictions made for the 10% of excluded water bodies. The strength of agreement between the model predictions and the observed data was quantified by the root mean square error (RMSE). This confirmed that the final model had better predictive performance than the full model.

The residuals from the final model were plotted and examined to confirm that they were normally distributed, had homogeneous variance, and showed no obvious pattern of over- or under-estimation.

5.5 Fitted models

Table 5.1 details, for each of the four final models, which land use predictor variables were retained, the degree of fit (measured by the % variation explained), and the predictive performance (measured by the RMSE). Nitrate inputs from agriculture, urban diffuse and septic tank sources were positively associated with instream nitrate concentrations in all four models. The impact of point source discharges on water quality was found to be best represented by the total upstream consented flow (psflow) predictor variable, which was

⁷ The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection.

retained in three out of four final models. The coverage of historic landfill sites was not significantly associated with water quality in any of the models.

All four models explained around 80% of the total variation in water quality among water bodies, and were able to predict 95th percentile concentrations with a RMSE of less than 3 mg N/l. Full details of the final models are provided in Appendix C.

Table 5.3 Summary of final land use models

Predictor variable	Catchment scale model (main stem sites)		Water body scale model (tributaries and headwaters)	
	TIN	TON	TIN	TON
agriconc	✓	✓	✓	✓
urbanconc	✓	✓	✓	✓
septicconc	✓	✓	✓	✓
pscount				
psconc				
psflow		✓	✓	✓
landfill				
% of variation explained	80%	80%	79%	78%
RMSE	2.14	2.19	2.88	2.98

5.6 Model predictions

The final models were used to predict the 95th percentile TIN and TON concentration in main stem rivers and tributaries/headwaters in every river water body > 20 km². Predictions were not made for water bodies with an area or catchment area < 20 km² because a reliable assessment of land use could not be derived (see Section 5.2 for justification of this choice of threshold). The predicted TIN concentrations for main stem rivers and tributaries/headwaters are shown in Figure 5.1 and Figure 5.2, respectively.

If the predicted 95th percentile TIN concentration exceeded 50 mg NO₃/l, the water body was considered to have failed the assessment. The level of confidence in the catchment- and water body-scale predictions was recorded as one of six classes (Table 5.4). These scores were taken forward to assess the overall strength of evidence for nitrate pollution (Step 4; see Section 6).

Table 5.4 Summary of confidence classes

Class	Description
1	At least 95% confidence that the 95 th percentile concentration is ≤ 50 mg NO ₃ /l
2	At least 75% confidence that the 95 th percentile concentration is ≤ 50 mg NO ₃ /l
3	At least 50% confidence that the 95 th percentile concentration is ≤ 50 mg NO ₃ /l
4	At least 50% confidence that the 95 th percentile concentration is > 50 mg NO ₃ /l
5	At least 75% confidence that the 95 th percentile concentration is > 50 mg NO ₃ /l
6	At least 95% confidence that the 95 th percentile concentration is > 50 mg NO ₃ /l

Figure 5.1 Predicted 95th percentile TIN concentration in main stem rivers

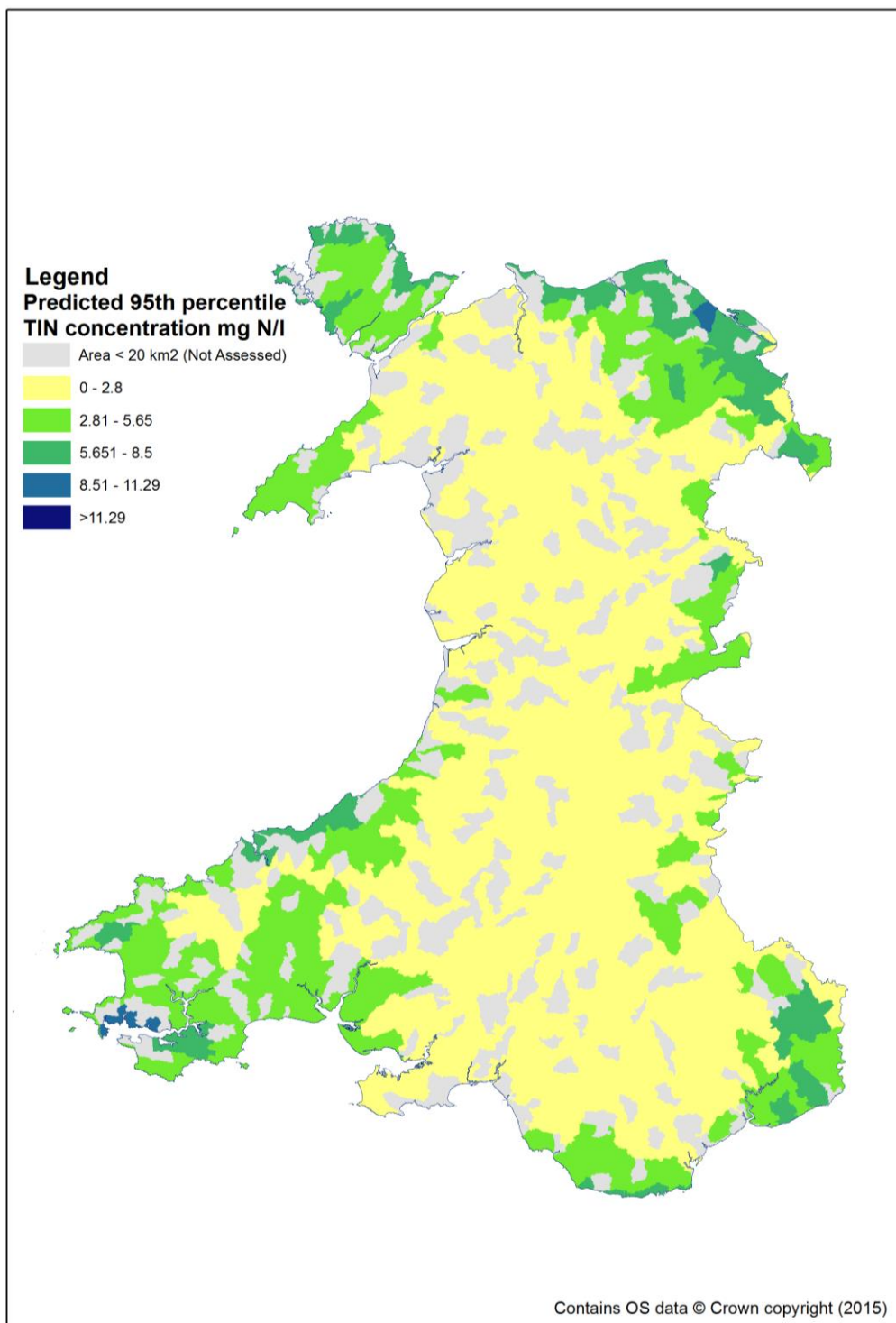
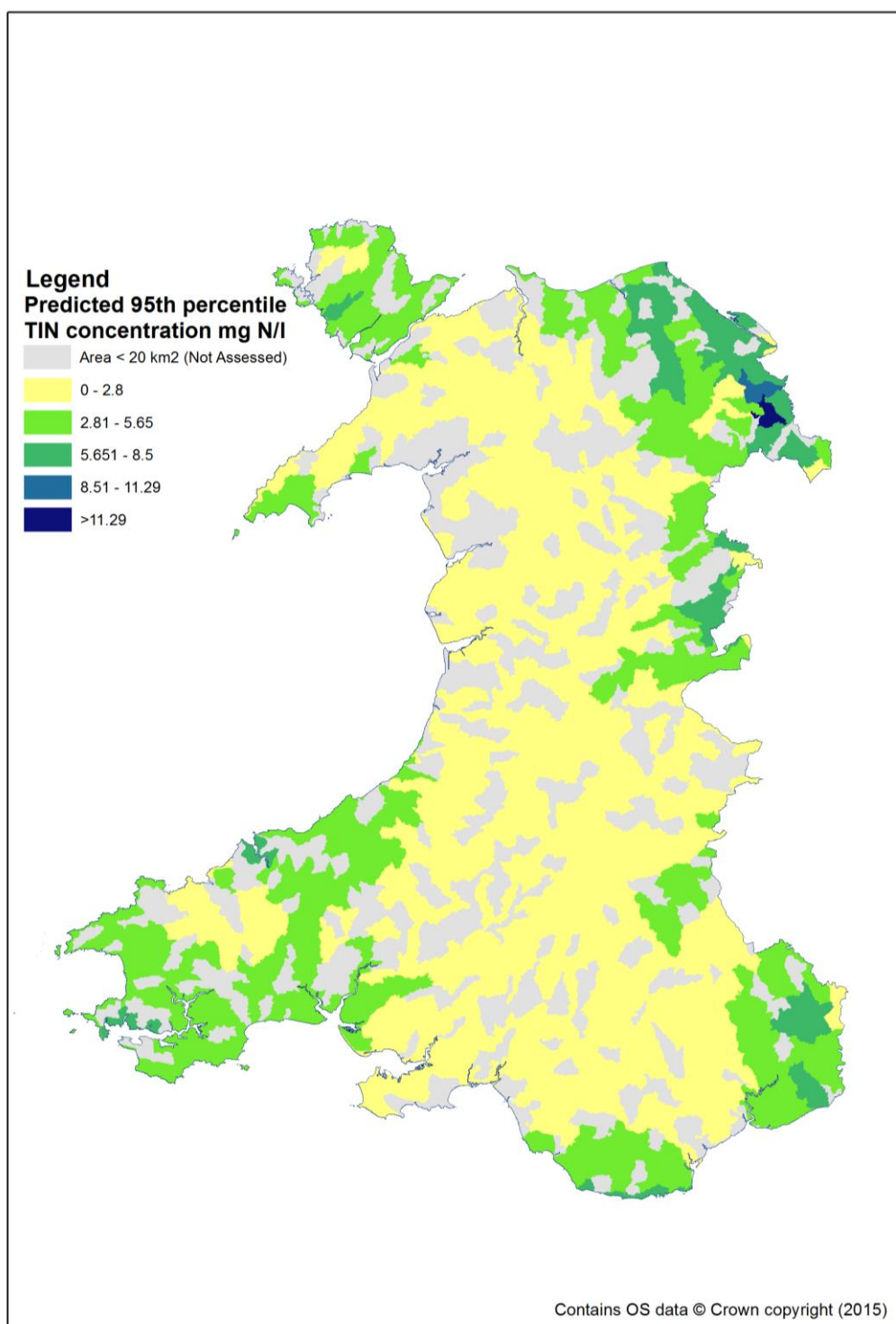


Figure 5.2 Predicted 95th percentile TIN concentration in tributary and headwater streams



6. Combining the evidence from monitoring and modelling

6.1 Lines of evidence

Steps 2 and 3 generated four lines of evidence to judge whether or not a water body was subject to nitrate pollution:

- A. present day 95th percentile TIN concentration from monitoring data;
- B. trend-predicted future 95th percentile TIN concentration from monitoring data;
- C. 95th percentile TIN concentration predicted from water body land use; and,
- D. 95th percentile TIN concentration predicted from catchment land use.

A and B were assessed for 977 surface water monitoring sites in 570 unique water bodies. This meant that 291 water bodies had no direct evidence from monitoring data.

C was assessed for 413 water bodies and D for 495. A reliable prediction could not be generated for C for 448 water bodies and for D for 366 water bodies, as the water body/catchment areas were less than 20 km², and therefore too small to adequately characterise their land use.

Each line of evidence was assessed to gauge whether or not the 95th percentile concentration exceeded 50 mg NO₃/l, and scored on a 1 to 6 scale, with 1 being a highly confident pass and 6 being a highly confident fail (Table 6.1). A score of 0 was assigned to water bodies that lacked monitoring data or a reliable land use model prediction. A and B were combined by taking the highest (i.e. worst) score for each monitoring site.

Table 6.1 Summary of confidence classes

Class	Description
1	At least 95% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
2	At least 75% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
3	At least 50% confidence that the 95 th percentile concentration is ≤50 mg NO ₃ /l
4	At least 50% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
5	At least 75% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l
6	At least 95% confidence that the 95 th percentile concentration is >50 mg NO ₃ /l

6.2 Combining the lines of evidence

The evidence from the monitoring data and the land use model was combined to assess the strength of evidence for nitrate pollution in each water body. Two tests were conducted for every water body:

- the worst performing tributary or headwater monitoring site in the water body (if any) was assessed in conjunction with the water body land use model prediction (if available); and,
- the worst performing main stem monitoring site in the water body (if any) was assessed in conjunction with the catchment-scale land use model prediction (if available).

Figure 6.1 shows the evidence matrix used to determine water bodies with nitrate pollution.

Figure 6.1 Evidence matrix used to determine if water body is subject to nitrate pollution

		Land use model prediction (C or D)						
		1	2	3	4	5	6	0
Monitoring results (highest of A and B)	1	Blue	Blue	Blue	Green	Green	X1	Blue
	2	Blue	Blue	Green	Green	Green	X1	Blue
	3	Blue	Green	Green	Green	Green	X1	Green
	4	Green	Green	Green	X2	Red	Red	Green
	5	Green	Green	Green	Red	Red	Red	X3
	6	X1	Red	Red	Red	Red	Red	X3
	0	Blue	Blue	Green	Green	X3	X3	Grey

Red cells indicate that the water body had sufficient evidence to be assessed as polluted or at significant risk of pollution.

Green cells indicate that the water body was assessed as not polluted or at low risk of pollution.

Blue cells indicate that the water body is not polluted and that the risk of nitrate pollution from agriculture is very low. This category was used to help determine whether or not existing surface water NVZs should be removed from designation (see Step 7; Section 7).

Grey cells indicate water bodies that could not be assessed because of a lack of evidence: they contained no monitoring sites and lacked a reliable prediction from the land use model. Together these water bodies accounted for less than 10% of the land area of Wales.

Figure 6.2 and Figure 6.3 map the risk of nitrate pollution at a local and catchment scale, respectively. A water body was considered to be polluted if either the local test or the catchment-scale test produced evidence of pollution.

The evidence matrix was used as a guide to decide whether or not surface waters were polluted. Discretion was exercised in all cases, but particularly in cells marked 'X', where:

(X1) the monitoring data and land use model gave contradictory results;

(X2) the evidence for pollution carried a low degree of confidence; or

(X3) the decision was informed by only one strand of evidence (either monitoring data or the land use model prediction were absent).

In these marginal cases, local knowledge and detailed data quality checks were used to provide an additional level of scrutiny and to help inform the final decision. Step 8 gives further details of the type of factors taken into consideration at the local ground-truthing workshops.

Figure 6.2 Assessment of surface water nitrate pollution risk at a local (water body) scale

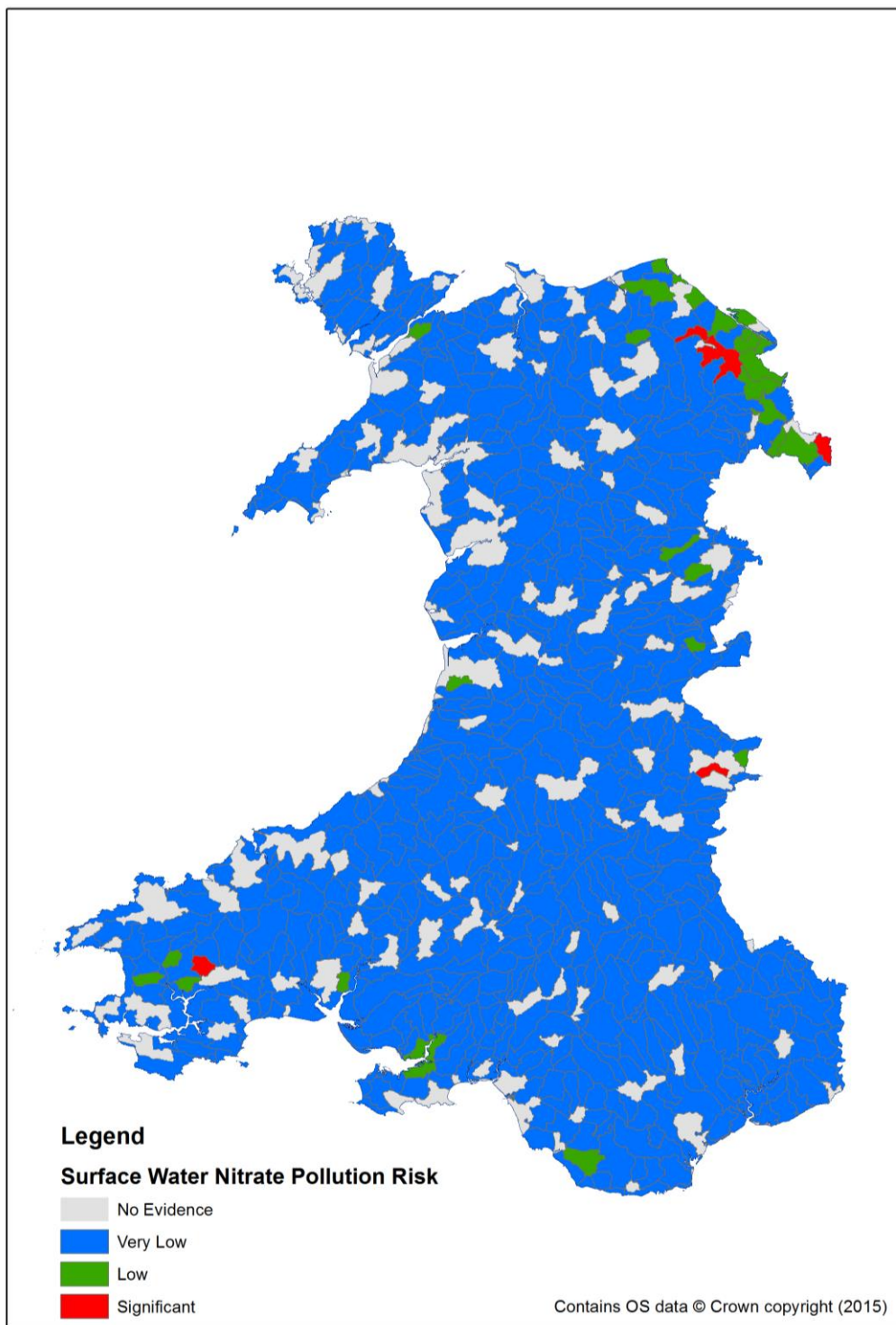
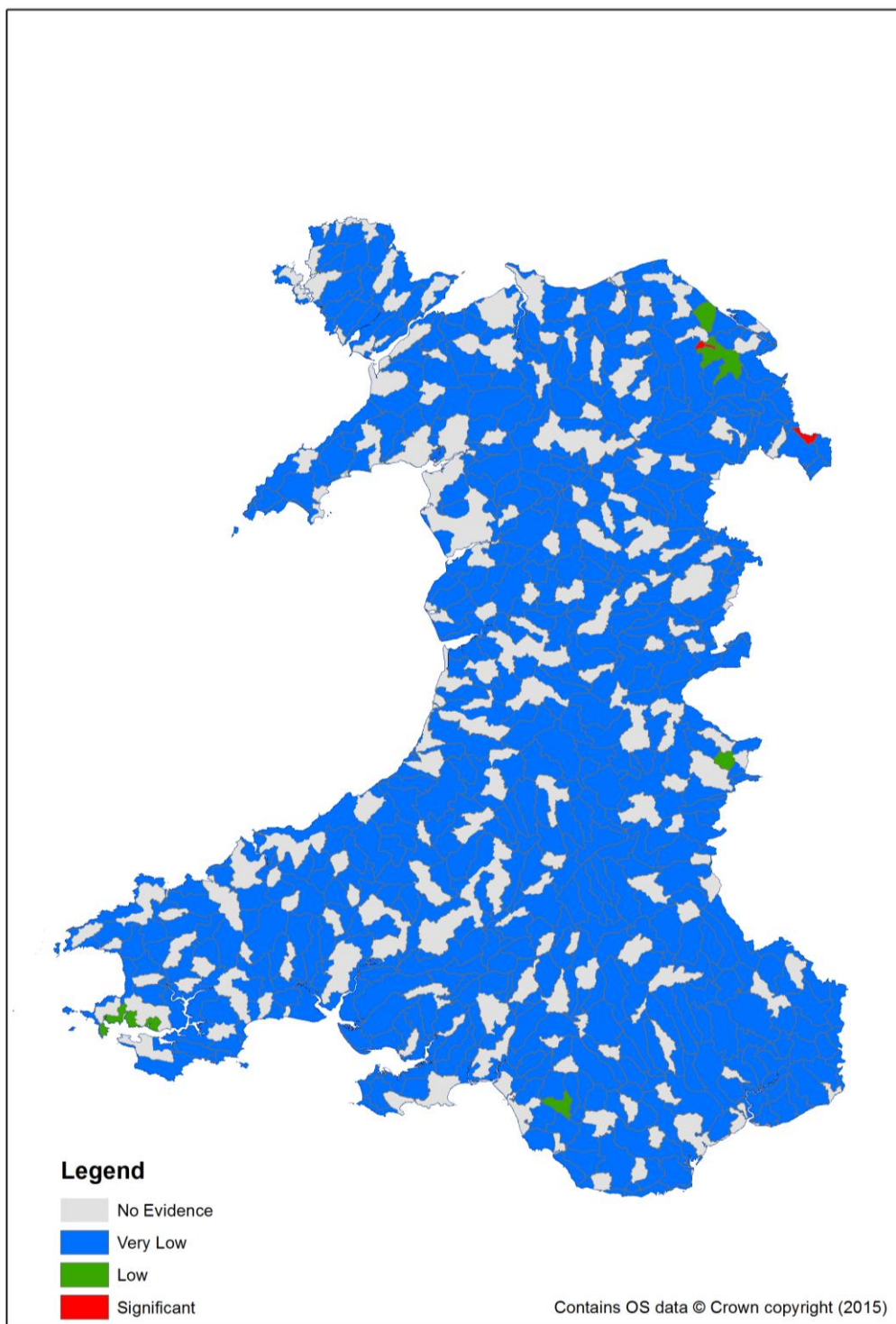


Figure 6.3 Assessment of surface water nitrate pollution risk at a catchment scale



7. Removal of previously designated surface water NVZs

7.1 Introduction

The use of a weight of evidence approach incorporating multiple lines of evidence is designed to provide consistency in surface water designations between NVZ Review cycles. Despite this, the iterative process required by the Nitrates Directive inevitably results in discrepancies between areas of existing designation and potential new designations. The use of up-to-date monitoring data can give rise to new candidate areas for designation where water quality is deteriorating. At the same time, the precautionary implementation of the Directive means that as the science improves and more data becomes available, we can become confident that, based on the available evidence, some previously designated areas no longer meet the criteria for designation set out in the Nitrates Directive.

Informal discussions between Defra and the European Commission in January 2011 suggested that each NVZ designation cycle can, in effect, be considered independent of previous cycles. Therefore, there is no fundamental reason why existing NVZ areas could not be removed from designation, although the Commission would require clear evidence that such areas are not vulnerable to nitrate pollution if Action Programme measures were to be discontinued.

7.2 Criteria for removal from designation

The Methodology Review Group convened as part of the 2013 NVZ Review agreed that a case for excluding areas previously designated as NVZs might be justified based on evidence of a significant and sustained improvement in water quality or where precautionary designations could be shown to be unjustified.

The following broad criteria were used to justify de-designation of surface water NVZs:

- 1. monitoring data demonstrates that the water bodies are no longer polluted;
- 2. agricultural land use is low risk as a source of nitrate;
- 3A. the improvement in water quality is sustained over at least two NVZ Review periods;
- 3B. the likely cause(s) for the water quality improvement can be determined;
- 4. land previously identified as draining to a polluted water can, on the basis of improved data, reasonably be determined to not drain to a polluted water.

In essence, criteria (1) and (2) are used to identify a subset of water bodies that were previously designated but now show convincing evidence that waters are no longer polluted. The relevant data, plus other information for each of these water bodies and their catchments, was considered on a case-by-case basis, with reference to criteria (3A) and/or (3B).

Criteria (4) applies principally to areas of improved hydrological understanding or corrections in mapping.

7.2.1 Criteria 1 and 2: Improvement in water quality

The evidence matrix (Figure 7.1) used to identify water bodies for NVZ designation was also be used to identify water bodies for de-designation. Blue cells indicate combinations of evidence that indicate the water body is not polluted and that the risk of nitrate pollution from agriculture is low. Essentially, the 95th percentile TIN concentration must be below 50 mg N/l with at least moderate (75%) confidence (i.e. class 2) for both monitoring data and the land use model. Where both lines of evidence are available, the class 2 limit may be relaxed for one line of evidence when supported by very clear (i.e. class 1) evidence from the other line of evidence.

Figure 7.1 Evidence matrix used to determine if water body is subject to nitrate pollution

		Land use model prediction (C or D)						
		1	2	3	4	5	6	0
Monitoring results (highest of A and B)	1	Blue	Blue	Blue	Green	Green	X1	Blue
	2	Blue	Blue	Green	Green	Green	X1	Blue
	3	Blue	Green	Green	Green	Green	X1	Green
	4	Green	Green	Green	X2	Red	Red	Green
	5	Green	Green	Green	Red	Red	Red	X3
	6	X1	Red	Red	Red	Red	Red	X3
	0	Blue	Blue	Green	Green	X3	X3	Grey

7.2.2 Criterion 3A: Sustained improvement in water quality

To avoid the risk of removing an NVZ based on a temporary, short-term improvement in water quality, the monitoring data must indicate at least medium (75%) confidence of the 95th percentile nitrate concentration being below 50 mg/l (i.e. class 2) for at least two cycles of NVZ designation (i.e. 2004-2009 for the 2013 Review and 2009-2014 for the present Review).

Because this criterion applies to the peak nitrate concentrations (75% confidence that the 95th percentile is less than 50 mg N/l), when this test is satisfied it implies a very low probability of any sample exceeding the drinking water standard and ensures that the mean concentration is below, and in most cases, considerably below, the drinking water standard.

This requirement may be relaxed where an improvement in water quality can be unambiguously attributed to improvements in effluent discharges (see criterion 3B below).

7.2.3 Criterion 3B: Cause of the observed improvement

Establishing the cause of observed improvements in water quality is essential to demonstrate that nitrate pollution will not return in areas where an effective Nitrates Directive Action Programme is lifted. Appropriate evidence should be provided to support each area removed on a case by case basis.

On the agricultural side, reductions in total N-fertiliser use over the past 15 years and, within NVZ areas, measures taken by farmers under the action programme are also expected to have contributed to the observed improvements in surface water quality. Of course, where improved water quality may have resulted from improved fertiliser management, this alone cannot be used as a reason to remove areas from NVZ designation.

However, over the past 15 years there has been significant investment by water companies in improving effluent (sewage) treatment works to meet the requirements of the Water Framework, Habitats and Urban Waste Water Treatment Directives. In some cases, increased nitrification of ammonium will have had the effect of reducing the direct ammonium loading and increasing the direct nitrate loading, although the TIN loading is expected to have been largely unaffected. In other cases where treatment works have been closed, or effluent discharges relocated to provide greater dilution, the cessation of, or reduction in TIN loading is expected to be a principal cause of improvement in water quality. This should form the basis of a strong case for removal of a designation because the observed improvement is unrelated to the NVZ action programme.

7.2.4 Criterion 4: Land not draining to a polluted water

Improvements in the accuracy of the mapping datasets used to define hydrological catchments mean that some previous designations have been overly precautionary. In these cases it is appropriate to remove areas of land that can be shown not to drain to polluted waters.

8. Comparison with 2013 Review methodology

8.1 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.

8.2 Refinements to Step 1: Identification of surface freshwaters for analysis

Table 8.1 details the refinements made to the methodology used in Step 1.

Table 8.1 Summary of methodological refinements for Step 1

Refinement	Justification
Internal Drainage Boards (IDBs) were not used to represent hydrological connectivity in low-lying areas.	Artificial drainage networks are very limited in Wales and the hydrological connectivity is represented adequately by WFD river water bodies.

8.3 Refinements to Step 2: Statistical analysis of surface water monitoring data

Table 8.2 details the refinements made to the methodology used in Step 2.

Table 8.2 Summary of methodological refinements for Step 2

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 95 th percentile concentration at the end of the monitored period was incorporated into the 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.
The minimum number of data points to be spanned by each spline in the quantile regression model was increased from 4 to 12.	A larger number of samples per spline produces a better fitting quantile regression model, particularly when there are large gaps in the monitoring record.
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.
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.

8.4 Refinements to Step 3: Land use modelling of nitrate pollution in surface waters

Table 8.3 details the refinements made to the methodology used in Step 3.

Table 8.3 Summary of methodological refinements for Step 3

Refinement	Justification
Nitrate inputs from septic tanks and historic landfill sites were included as candidate predictor variables in the land use models.	The inclusion of these variables provides a more comprehensive assessment of nitrate pollution sources.
Base Flow Index (BFI) was not included as a candidate predictor variable in the land use models.	The 2013 NVZ Review concluded that BFI was not a useful predictor of nitrate concentrations in rivers (Environment Agency 2012). Also, any influence of BFI on water quality is likely to be complex and difficult to model adequately with the size of dataset available.
Areas of low relief (Internal Drainage Boards) were not included as a candidate predictor variable in the land use models.	Artificial drainage networks are very limited in Wales and the hydrological connectivity is represented adequately by WFD river water bodies.
Nitrate losses from different agricultural land use classes were aggregated and treated as a single predictor variable.	The 2013 NVZ Review used single total agricultural losses (TAL) term in the final models. Modelling separately the nitrate losses from individual agricultural land use classes unnecessarily complicates the model selection process.
Model selection and simplification used AIC as a model fit criterion instead of r^2 values.	AIC is a more sophisticated criterion that penalises unnecessary complexity to find the most parsimonious model.

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

The matrix used to combine the evidence from monitoring and modelling was identical to that used in the 2013 NVZ Review.

8.6 Refinements to Step 7: Removal of previously designated surface water NVZs

The criteria and evidence used to assess whether existing NVZs could be de-designated were identical to those used in the 2013 NVZ Review.

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.

Carey, M. and Davison, P. 2010. Cumulative Nitrogen and Phosphorus Loadings to Groundwater. Entec Report to SEPA, EA and NIEA, 22 November 2010.

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.

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 – Surface Waters. 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.

Koenker, R. and Hallock, K.F. 2001. Quantile regression. *Journal of Economic Perspectives*, 15 (4), 143-156.

Koenker, R. 2015. Quantile regression: the quantreg package.
<http://cran.r-project.org/web/packages/quantreg/quantreg.pdf>

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.

Powelson, 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_{γ} 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 km² 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 km² 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.

Appendix C Details of land use models

C1 Catchment scale model for TIN

Figure C.1 Coefficients and summary statistics for catchment-scale TIN model

```

Call:
lm(formula = log10(current.est) ~ log10(agriconc + 0.1) +
log10(urbanconc +
0.1) + log10(septicconc + 1e-04), data = final.data)
Residuals:
    Min       1Q   Median       3Q      Max
-0.4090 -0.0890 -0.0023  0.0750  0.6851
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.3822    0.0984   3.88 0.00015 ***
log10(agriconc + 0.1)  0.7494    0.0735  10.20 < 2e-16 ***
log10(urbanconc + 0.1)  0.1261    0.0184   6.84 1.3e-10 ***
log10(septicconc + 1e-04) 0.1556    0.0448   3.47 0.00065 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.157 on 173 degrees of freedom
Multiple R-squared:  0.795, Adjusted R-squared:  0.792
F-statistic: 224 on 3 and 173 DF, p-value: <2e-16

RMSE 2.14
AIC -148

```

Figure C.2 Plot of predicted vs observed 95th percentile TIN concentrations for the catchment-scale model

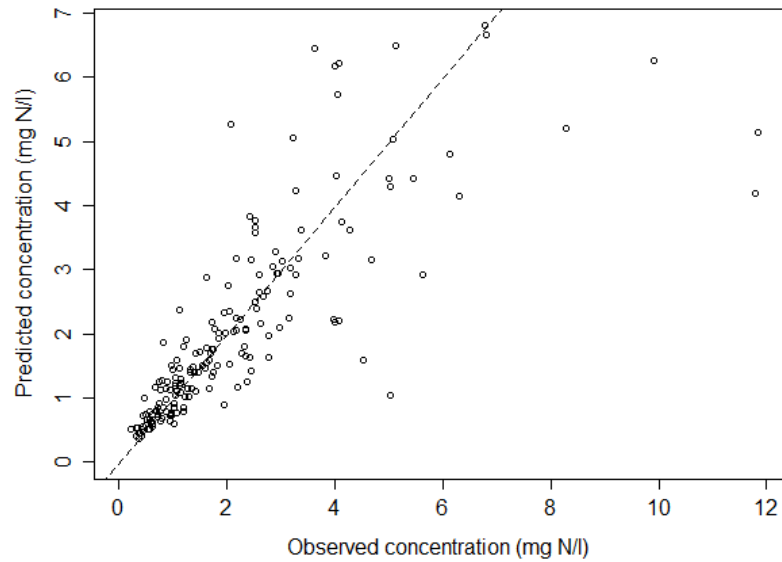
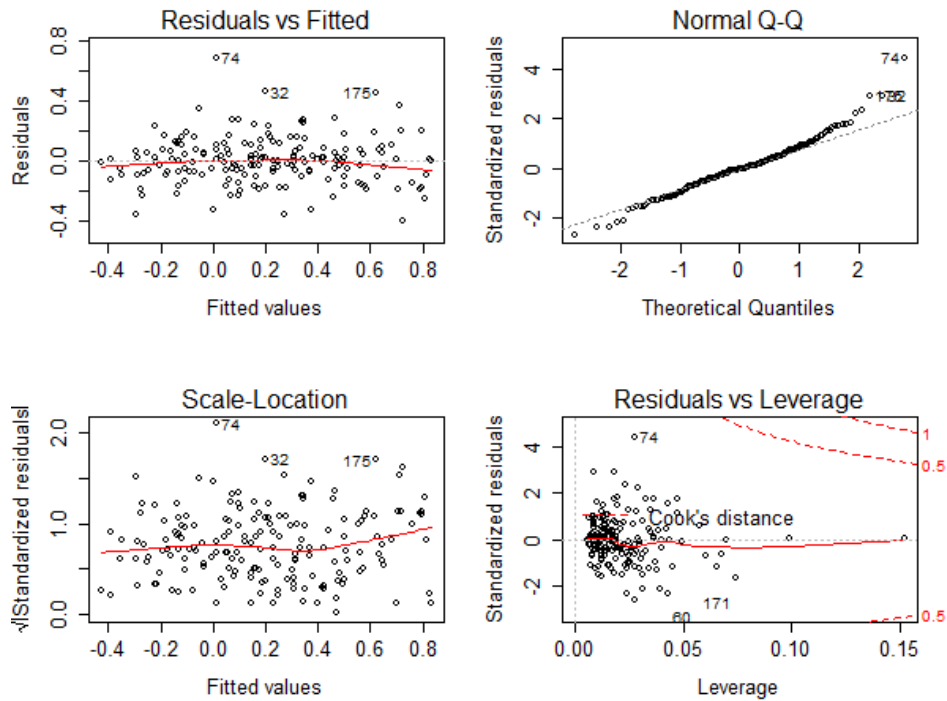


Figure C.3 Diagnostic residual plots for catchment-scale TIN model



C2 Catchment scale model for TON

Figure C.4 Coefficients and summary statistics for catchment-scale TON model

```

Call:
lm(formula = log10(current.est) ~ log10(agriconc + 0.1) +
log10(urbanconc +
0.1) + log10(septicconc + 1e-04) + log10(psflow + 1e+05),
data = final.data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.3497 -0.0931 -0.0043  0.0862  0.6805

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.2743    0.1323   2.07  0.0397 *
log10(agriconc + 0.1)  0.7629    0.0750  10.17 < 2e-16 ***
log10(urbanconc + 0.1)  0.1091    0.0191   5.70 5.1e-08 ***
log10(septicconc + 1e-04)  0.1530    0.0473   3.24  0.0015 **
log10(psflow + 1e+05)    0.0115    0.0077   1.49  0.1371
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.158 on 172 degrees of freedom
Multiple R-squared:  0.797, Adjusted R-squared:  0.793
F-statistic: 169 on 4 and 172 DF, p-value: <2e-16

RMSE 2.19
AIC -144

```

Figure C.5 Plot of predicted vs observed 95th percentile TON concentrations for the catchment-scale model

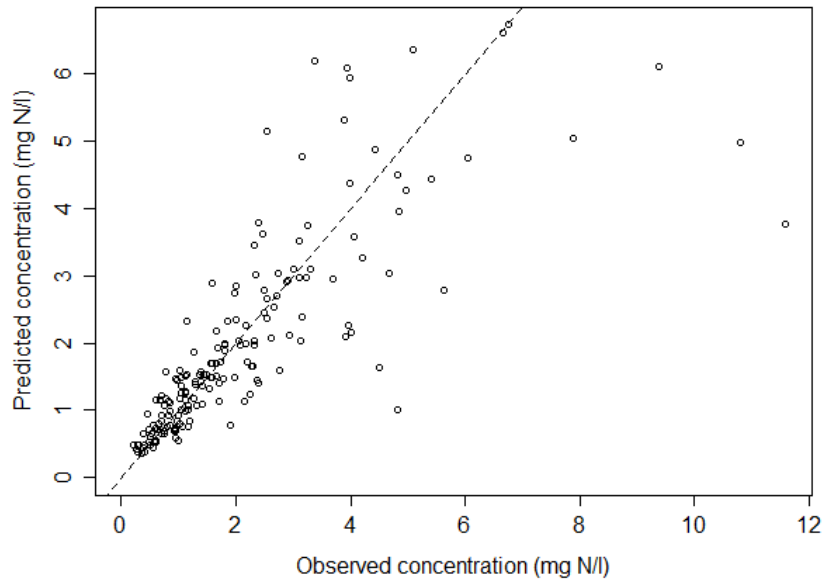
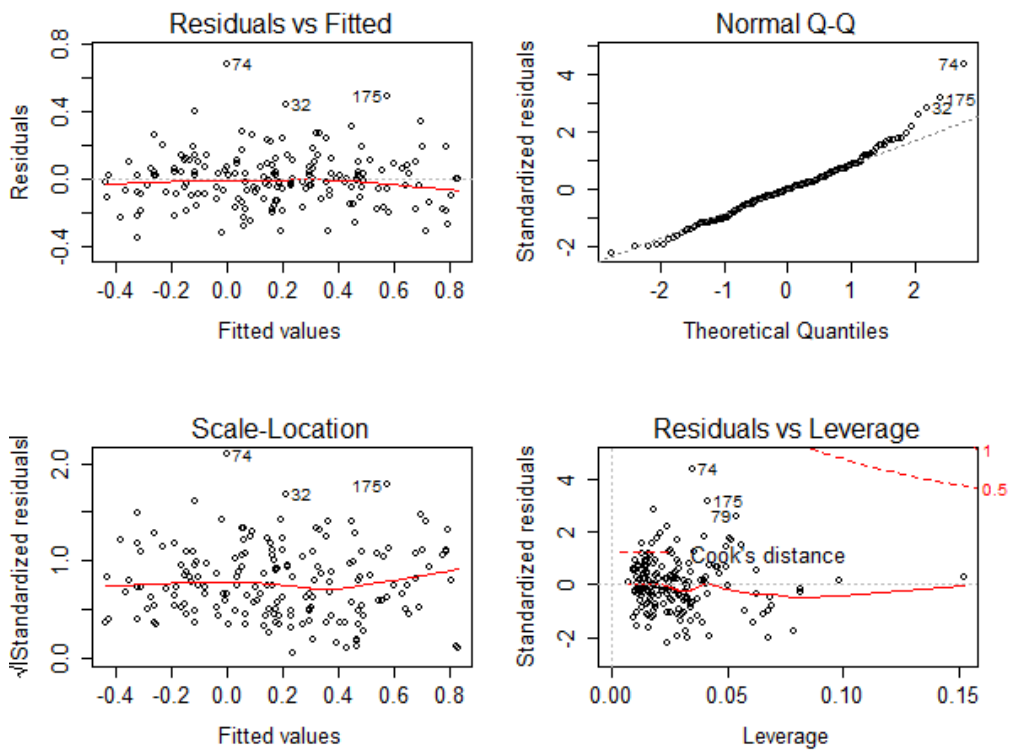


Figure C.6 Diagnostic residual plots for catchment-scale TON model



C3 Water body scale model for TIN

Figure C.7 Coefficients and summary statistics for water body-scale TIN model

```

Call:
lm(formula = log10(current.est) ~ log10(agriconc + 0.1) +
log10(urbanconc +
0.1) + log10(septicconc + 1e-04) + log10(psflow + 1e+05),
data = final.data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.8171 -0.1069 -0.0056  0.1098  0.6088

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)         0.04005   0.10512    0.38   0.704
log10(agriconc + 0.1)  0.80013   0.06449   12.41 < 2e-16 ***
log10(urbanconc + 0.1)  0.09862   0.01677    5.88 1.5e-08 ***
log10(septicconc + 1e-04) 0.06826   0.03653    1.87  0.063 .
log10(psflow + 1e+05)    0.01986   0.00862    2.30  0.022 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.187 on 224 degrees of freedom
Multiple R-squared:  0.789, Adjusted R-squared:  0.785
F-statistic: 210 on 4 and 224 DF, p-value: <2e-16

RMSE 2.88
AIC -111

```

Figure C.8 Plot of predicted vs observed 95th percentile TIN concentrations for the water body-scale model

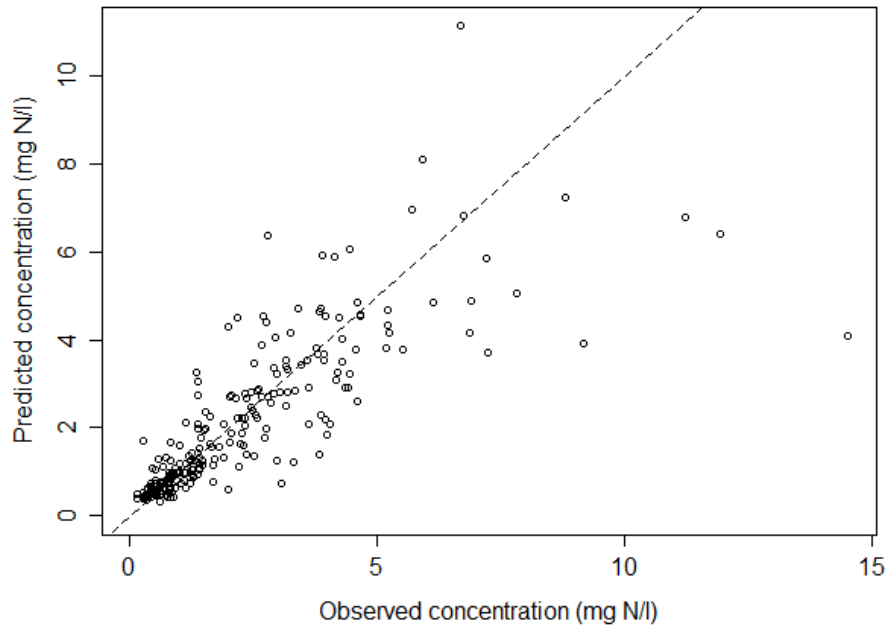
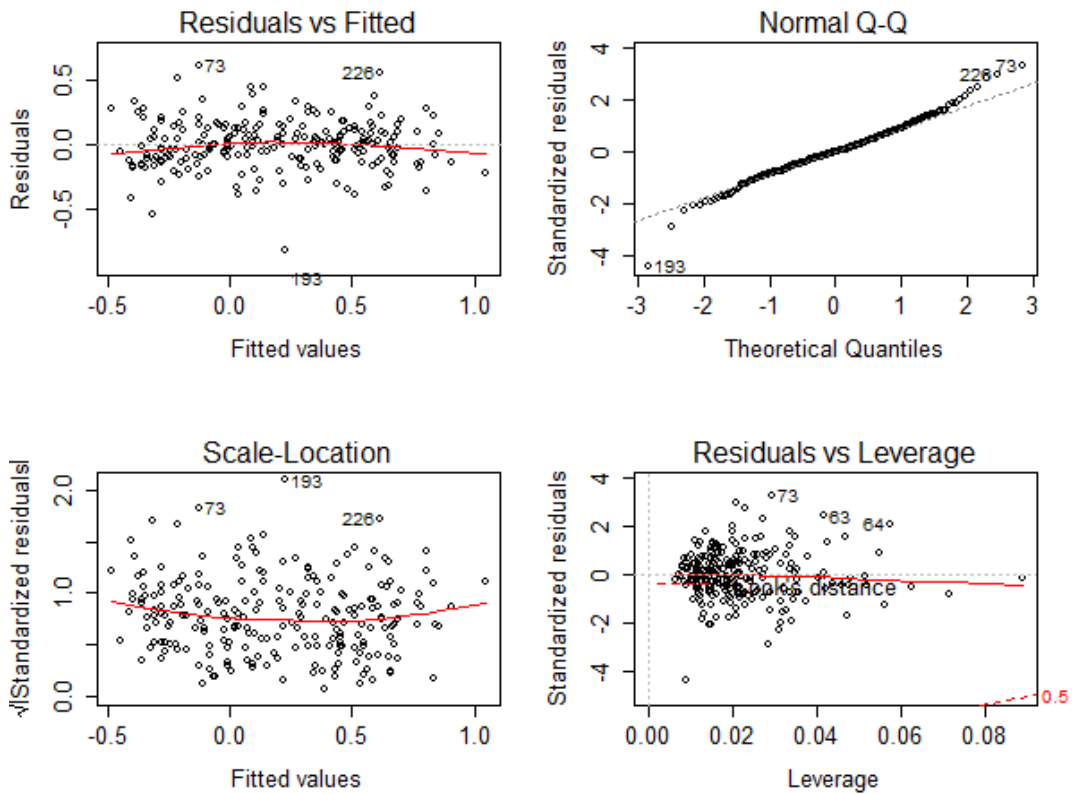


Figure C.9 Diagnostic residual plots for water body-scale TIN model



C4 Water body scale model for TON

Figure C.10 Coefficients and summary statistics for water body-scale TON model

```

Call:
lm(formula = log10(current.est) ~ log10(agriconc + 0.1) +
log10(urbanconc +
0.1) + log10(septicconc + 1e-04) + log10(psflow + 1e+05),
data = final.data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.8811 -0.1106 -0.0008  0.1062  0.6232

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.0210    0.1085   0.19   0.847
log10(agriconc + 0.1)  0.8216    0.0666  12.34 < 2e-16 ***
log10(urbanconc + 0.1)  0.0907    0.0173   5.24 3.8e-07 ***
log10(septicconc + 1e-04) 0.0708    0.0377   1.88  0.062 .
log10(psflow + 1e+05)    0.0208    0.0089   2.34  0.020 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.193 on 224 degrees of freedom
Multiple R-squared:  0.783, Adjusted R-squared:  0.779
F-statistic: 202 on 4 and 224 DF, p-value: <2e-16

RMSE 2.98
AIC -96.7

```

Figure C.11 Plot of predicted vs observed 95th percentile TON concentrations for the water body-scale model

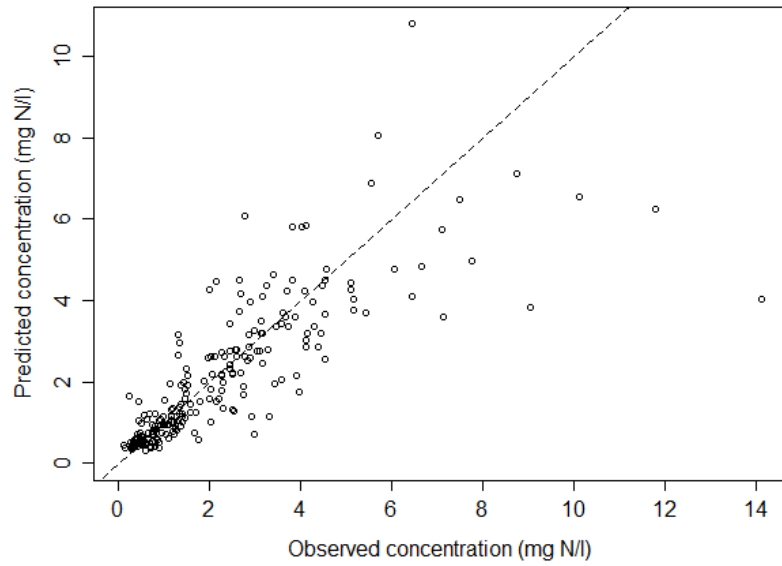


Figure C.12 Diagnostic residual plots for water body-scale TON model

