Assessing water quality in Europe using stratification techniques

Results of a prototype application using French data

ISSN 1725–2237

European Environment Agency

X

Assessing water quality in Europe using stratification techniques

Results of a prototype application using French data

European Environment Agency

Cover design: EEA
Layout: Diadeis and EEA

Luxembourg: Office for Official Publications of the European Communities, 2007

ISBN 978-92-9167-928-7
ISSN 1725-2237

European Environment Agency
Kongens Nytorv 6
1050 Copenhagen K
Denmark
Tel.: +45 33 36 71 00
Fax: +45 33 36 71 99
Web: eea.europa.eu
Enquiries: eea.europa.eu/enquiries

Contents

Acknowledgement

This report was prepared by Philippe Crouzet, the European Environment Agency (EEA) with support from Morten Tranekjær Jensen, Danish National Environmental Research Institute (NERI) under the supervision of Ronan Uhel, Head of Spatial Analysis Group, EEA and Jock Martin, Head of Programme, Biodiversity, Spatial Analysis and Scenarios, EEA.

EEA acknowledges early contributions from the European Topic Centre for Water (1997), which were later implemented by the French Environment Institute (ifen) in France (1998–2000), leading to significant improvement in former methodologies. Disaggregated data provided by the French National Reference Centre for Water also made a significant contribution to the improvement of current data flows.

The authors would also like to thank other experts and colleagues for their informal contributions and all those who submitted comments during the consultation phase which considerably improved the final draft of the report.

1 Executive summary

1.1 Objectives and method

This report aims to provide information in relation to the trends of major nutrients in rivers. It focuses both on the trends considered as indicators of water quality status and on the possible differences in trends in relation to main driving forces (agriculture, urbanisation). Referring to the driving forces — pressures — state — impact — response (DPSIR) assessment framework of the EEA, the relationships drivers — state have been systematically analysed at different geographical levels.

The purpose of these analyses is to consider if measurable improvements of river-water status and their significance can be assessed, and if these trends can be reasonably related to relevant driving forces, especially the ones with the largest impact, and the policies put in place to mitigate the environmental impacts of these driving forces. A complementary goal of the methodological improvement is to reduce the time lag between a policy implementation and the observation of its effects.

Such assessments must be applicable to all the areas covered by the EEA mandate. This poses two different challenges. First of all, the statistical methodology must be tested on a diversified area to demonstrate its capacity to respond to the questions mentioned above. This is the subject of this report which took stock of the provision of all nationally monitored data and reference systems in France kindly provided by the French Ministry of Ecology and Sustainable Development.

Secondly, the use of comparable reference systems, homogeneous at the European level, should be checked. This is the purpose of further applications that are underway, partly based on the European spatial assessment catchment system (CCM) being developed by the European Commission Directorate General Joint Research Centre (DG JRC).

The methodology used in this study improves on previous approaches used in assessments carried out by the EEA by:

- replacing assessments based upon station categories by stratified assessment of catchments. The relative weight of catchments replaces the proportion of stations, thus providing statistically comparable aggregated concentration values for the different determinants;
- defining accurately and in a reproducible way the different strata that allow comparable and meaningful relationships to be established between trends in driving forces and trends in quality status;
- providing estimates of the date on which targets can be achieved expressed as range of years with likelihood of the assessment. This is a key issue to policy effectiveness assessment.

The methodological improvements reported in this document are in line with the EEA objective of contributing to the European development towards the production of 'water accounts'. This approach follows the SEEA (System of Economic and Environmental Accounts) methodology to better link physical and economical components of the environment. SEEA derived approaches have now been raised as 'statistical standards' at the UNSD and Estat levels, hence contributing to making the environmental accounts comparable at the World level.

1.2 Historical background

Spatial assessment methods sustain the use of key tools that help the European Environment Agency (EEA) fulfil its regulatory mandate to produce policy-relevant integrated assessments of Europe's environment. When these methods are, for instance, applied to characterizing river basin catchments with a goal of better reporting the state of surface waters and their trends in a comparable way continental water quality status can be directly correlated with catchment characteristics; thus, surface waters in catchments characterised by, say, agricultural activities will be mostly impacted by emissions of nitrates resulting in changes in quality status; or, those in heavily urbanised catchment areas will be so by emissions of phosphate resulting from urban waste water. This type of analysis at catchment level, therefore, has a high potential in supporting evaluation of the effectiveness of related policies across a range of geographical scales and issues, from the catchment to the national to the EU levels. Such an approach would allow for a cross-policy analysis, in this case the Nitrates Directive and the Urban Waste Water Treatment Directive that are two key pieces of EU legislation and the management approach of the Water Framework Directive which is based on catchments.

When 'Eionet-water' (also previously known as Eurowaternet) was set up 10 years ago as the process for obtaining water data flows from EEA member countries, the intention of the 'representative networks' and 'stratified assessment' techniques was precisely to demonstrate and quantify these relationships between activities and emissions (e.g. agriculture and nitrate, urban waste water and phosphate, etc). In 1997, a methodological framework was issued. The Eionet-water quality data flows comprise aggregated data/statistics taken from a stratified sample of monitoring points from each Member State's monitoring networks.

The assessments carried out on these data flows were used in different publications by the EEA and helped identify the main issues regarding the quality of continental waters. Being based on aggregated data and subsets of stations, there has, however, been little room for flexibility in the assessments. The most problematic issues have been the low accuracy of the results which sometimes have proved hard to quantify and the fact that the sampling approaches were designed to provide comparable nationwide results rather than comparisons between catchments.

The French Institute for the Environment (ifen), which constitutes both the Statistical Institute of the French Ministry of the Environment and the EEA National Focal Point (NFP), has carried out several benchmarking and implementation studies in recent years in order to address these issues and so optimise its response to the EEA. The outcome of these studies indicates that the EEA and Eionet can achieve improvements in the scope and the quality of their water assessments, starting with the methodological principles issued in 1997. This report presents the findings of these studies by comparing the results obtained from the application of the 1997 methodological framework with those obtained from the improved approach.

1.3 Environmentally important results

Results using French data have been processed for the period 1971–2005. Between 1971 and 1981 data are less frequent than for more recent years. As a result, stratification may be more questionable because this period was also the one when sharp changes in land cover occurred. Despite these data limitations, the stratified statistics provide very marked and clear-cut trends with acceptable accuracy.

Three determinants ([1]) have been analysed because their source is shared between the major driving forces: agricultural, urban and natural sources. These are ammonium, phosphate and nitrate.

Nationwide, ammonium and phosphate trends display a dramatic decrease. On average, this favourable trend could meet an environmentally sound target in one or two decades. By contrast, for nitrate status it does not seem possible to make conclusions about improvements. However, the general rate of degradation seems to be slowing down and some catchments and strata suggest a trend towards decreasing concentrations. This requires careful follow-up.

The different basins present many contrasting situations. For example, the Seine catchment is currently in the range of 20 mg NO_3 L^{-1} (yearly weighted average) *versus* 13 on the Loire and 7 on the Garonne and the Rhône. For Brittany, where the stratified average reached almost 40 mg NO_3 L^{-1} 15–10 years ago, concentrations have decreased by 10 mg NO_3 L^{-1} since 1995.

The relationships between agriculture and ammonium and agriculture and phosphate are rather complex. Ammonium is produced in large quantities by livestock, but its transfer to water is paradoxically much lower where livestock is very dense. By contrast, nitrate is very high in these catchments suggesting effective manure collection and excessive spreading.

Accurate stratified assessment makes it possible to derive source apportionment estimates. These require methodological improvements in order

([1]) Water composition and quality are defined by sets of data that lacked a generic name. Candidates words 'parameter' and 'determinant' were rejected because insufficiently accurate. The ETC/Water suggested, after consultation of expert panel the new term 'determinant' that was adopted. It means any information that contributes to determining the assessment.

to compare them more accurately to modelled source apportionment which is based on flux analysis. Hence, statistics on discharge-weighted concentration are required instead of the current time-weighted concentration.

By contrast, the urban-dominated strata are undoubtedly the cause of high and quickly decreasing phosphate and ammonium concentrations. Crop agriculture and unexpectedly dense, urbanised areas are the source of high nitrate concentration. In the latter case, nitrate decreases along with the decrease in ammonium and phosphate, and the aggregation basin survey just fits to the findings of local assessments.

The two major policies — the Urban Waste Water Directive and the Nitrate Directive — seem to have had different outcomes. Where urban wastes (point sources) are involved, dramatic improvements have been achieved in the most polluted areas with no measurable time lag. Urban and industrial waste-water purification has a direct positive impact on receiving waters. Moreover, the action of French water agencies has been effective since the early 1970s — considerably earlier than the enforcement of the Urban Waste Water Directive. Its implementation has fostered more complete processes (P and N removal that were not included in the oldest activated sludge files) and increased the rhythm of works realisation.

Phosphate removal has apparently been boosted by the implementation of the directive and by national programmes aiming at preventing eutrophication in the populated areas. By contrast, in the areas where activities (urban and agricultural) below the national average are hosted, the situation does not seem to have improved, and in some cases has even worsened.

On the other hand, the implementation of the Nitrate Directive does not seem to have reached its objectives. Where non-point sources are involved in nitrate presence in river water, the situation often worsens with time. The positive point is that the worst situations seem to stabilise and improve locally. Since nitrate concentration in rivers is the outcome of a complex and lengthy process it clearly poses the question of accurately assessing the time lag between a measure and its observable result and the relevance of time targets that are not consistent with natural systems residence times.

1.4 Methodological findings

The studies hereby mentioned have demonstrated the importance of data quality assessment and qualification. In particular, the definition of strata suggested that a large share of conclusions might depend on inaccurate or ill-focused stratum definition and appropriate allocation of the monitoring station to the stratum.

Stratum definition has demonstrated that the major driver on a catchment cannot be defined solely by the main use of land. This finding confirms the need for developing an accurate catchment and rivers GIS, populated with the ad hoc information: land cover, population, etc. at the elementary catchment level. Such developments are ongoing at the EEA and JRC under the auspices of the Water Framework Directive Common Implementation Strategy.

Secondly, the statistical assessments carried out display a high variability of data and even larger regional influence. A simple analysis of differences between the selected data set (Eionet water) and the comprehensive data set provided the following findings:

- The Eionet data set gives a correct, although blurred, image of the nationwide status. Subtle changes in trends cannot be assessed.

- At driving-force level, only the ones with the least impact are correctly assessed. The ones with the most impact may be too uncertain to provide effective policy analysis. Hence, source apportionment would probably not be accurate from a limited data set.

- At river basin level, the analysis obtained from Eionet data set can be either accurate or inaccurate. Again, the less variable catchments are well addressed, but the changing ones are not well detected.

- The selected data set is not well suited for reporting changes over short time periods and should be replaced everywhere possible by full national data sets and disaggregated data. Such an approach would in turn lower the burdens and simplify data delivery by data providers in countries.

- Where enough data is available, many other aggregates can be processed, e.g. districts, priority areas, marine coastlines, sensitive areas, so maximising the benefits by using the same data for as many purposes as possible.

As a way of conclusion, it appears that responding effectively to political questions requires processing the available statistics in a way that is not one hundred percent orthodox. Two types of problems deserve specific insight. First of all, many national monitoring networks have been built without substantial statistical targets in mind and often mix figures from both independent and highly dependent monitoring points. Secondly, the expected results can not be produced by standard regression processing. The impact on the results themselves of the methodological short-cuts used in this report is likely to be insignificant but may have some impact on the uncertainty around these results.

A detailed list of methodological issues has been drafted and placed in the report.

2 Objectives and context

2.1 Work goals

The EEA is expected to provide reliable and relevant information about the state of the environment and relate, where possible, these assessments to policy actions and their effectiveness.

This study aims to contribute to such assessments at European level through the improvement of the methodology developed by the European Topic Centre on Water (ETC/W) (published in 1997). This methodology aimed at defining 'a representative network' of monitoring stations selected according to criteria, comprising soil occupation, population density and water body size. The key principle was that comparable assessments across Europe could be carried out with data from a stratified selection of stations. The selection of stations was understood to ensure comparability between countries, despite marked differences in monitoring efforts between Member States. This methodology is the basis for the Eionet-water data flows that have been collected annually since the mid-1990s.

Theoretical difficulties were identified during the implementation phase, especially in France. In order to fulfil the acknowledged goal suggested by the ETC/W of 'representative assessment' of water quality issues, methodological improvements were tested. The methodological improvements applied at prototype level with French data aim to:

- address all relevant quality issues by fully implementing the stratified approach suggested by the ETC/W in a totally reproducible and more flexible way;
- output reliable information and, where necessary, identify what cannot be concluded by systematically assess the uncertainties;
- foster spatial representativeness and comparability by making the best possible use of the spatial information and GIS which become the basis of EEA's assessments;
- prepare the grounds for foreseen changes in data availability resulting either from better access to information or changes in monitoring networks resulting from the implementation of the Water Framework Directive.

The objective of this work is primarily to provide the methodological background to answer two simple questions:

1. What is the overall status and trend of waters in a certain area where different activities take place? For example: '*What is the nitrate concentration of rivers in country X?', or 'Is the ammonium concentration in rivers of catchment Y impacted by livestock increasing or decreasing?*'

2. What are the impacts and the development with time of sectoral activity (and, is the impact identical over different areas)? Reciprocally, is there a relationship supported by evidence that sectoral policies have had positive effects? For example: '*Did the Urban Waste Water Directive result in meeting phosphate targets on rivers of catchment Z?*'

3 Methodological approach

3.1 Rationales

The methodology consists of formulating questions into problems that can be solved with existing data and adequate techniques. In order to apply the techniques within their domain of relevance and to use the data soundly, it is of paramount importance to define the problems accurately. In parallel, care must be taken to use the techniques appropriately and pose the questions so that they can be answered using 'press-button' techniques.

The question '*What is the nitrate concentration of rivers in country X?'* needs to be defined precisely in terms of what is meant by:

1. 'nitrate concentration'. This may be the annual average of concentrations computed with precautions to eliminate the systematic errors (bias) that might result from odd sampling dates or from changes in sampling frequency or other indicators. The indicator must be defined accurately, considering for example the value of the total volume of water run-off or the concentration of potentially abstracted water from rivers;

2. providing statistics on 'rivers' when only a limited number of samples is available and by adequately balancing the large and small rivers.

If the question is more sophisticated, for example: '*Is the ammonium concentration in rivers of catchment Y impacted by livestock increasing or decreasing?*', more definitions are required, e.g.

3. which rivers are 'impacted by livestock' and how can 'livestock' be defined and processed as a driving variable;

4. which information allows the conclusion to be made that the concentration is increasing or decreasing and over which time period?

The definitions resulting from the analysis of questions are expressed in data processing through specific techniques that necessarily imply hypotheses and simplifications. The statistical techniques are quite common: they are those applied to forecast pooling results. The first question in this case could be: '*What is the percentage of votes for candidate X?'* and the second *'Would male urban voters with income between a and b euro prefer candidate X or candidate Y?*'

There are several ways to address these problems. One of the most effective ones uses stratified statistics. The assumption is that a complex problem can be split into simpler ones, the final response being the combination of several elementary responses. Stratified statistics consist of combining a limited number of categories (named 'strata') that effectively mimic the final inventory.

In the case of nitrate, certain groups of catchments: urban, agricultural, pristine, etc., are known to deliver waters with marked differences in their composition. The paradigm backing stratified statistics is that the total 'population' to inventory can be apportioned into sub-populations whose driving characteristics are strongly bound to the analysed variable and can be sampled. For example, a strong causal relationship is presumed between ammonium concentration and livestock density on the catchment.

The use of stratified statistics provides several benefits:

1. if this causal correlation between a category and the analysed variable exists, then the statistical correlation accurately quantifies the causal correlation (for example, X animals per km^2 result in Z mg/l in river water) for each category. This is carried out with maximum precision since a single cause is separated from all other causes. Consequently; individual analysis is possible;

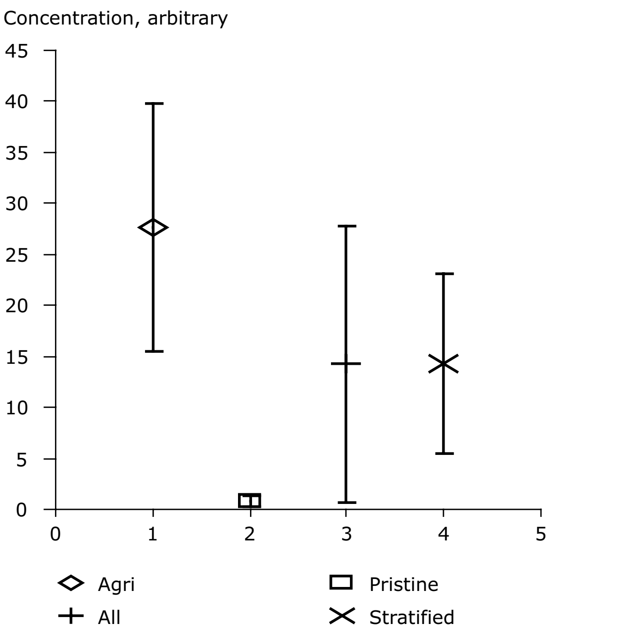
2. the overall result is better assessed with stratified statistics than with global statistics because each component of the result is the most precise. The overall accuracy is not changed by the use of stratified statistics;

3. as individual and overall precision is increased, small differences become statistically significant. Therefore, the trends can be estimated with good accuracy.

A stratum is defined precisely by the possible causal correlations. Consequently, the observations on one monitoring point may be related to completely different strata if different causality is analysed. For example, a catchment dominated by vineyards is not expected to have much nitrate in its waters. However, it could be contaminated by fungicides. The consequence is that the design of stratification is defined by the questions to answer and should not be defined *a priori*: a monitoring station is not, in itself, representative of a certain pressure.

To illustrate the second benefit, a simple graphic is presented (Figure 1). The two data sets are computed; one with high concentration and large dispersion, and the other with small concentration and little dispersion. The combined average is the same as if computed from the whole data set (same accuracy). However, its precision is better (a smaller standard deviation measures the accuracy of the average in this example).

Figure 1 Example result of combined accuracy and precision



Note:

Graphic 1, 'Agri' mimics a place where high concentration of substance (e.g. nitrate) is observed. Dispersion (represented by vertical bars) is high and average by the central point.

Graphic 2, 'Pristine' mimics a place where low concentration of same substance is observed. Dispersion is low.

Graphic 3, 'All' is the result from the combined data altogether). Dispersion is greater than the 'Agri' one, because of the mixing of data sets.

Graphic 4, 'Stratified' is the result of combining results 1 and 2. The average is identical to value for 3, but dispersion is smaller than the largest of both.

Data is randomised.

The issue of improving the accuracy of estimates has often been neglected. However, estimation errors are of paramount importance when carrying out comparability and trend assessments. Many efforts have been devoted to this work in order to consider uncertainty ranges as key information attached to estimates so as to increase their accuracy. In classical statistics, the uncertainty around a value is typically estimated by the ratio of the standard deviation by the square root of the number of observations. Hence, lowering the uncertainty requires either increasing the number of observations or lowering the standard deviation, or both.

Increasing the number of observations can be achieved by collecting more data where they exist but becomes impossible when considering observations in the past. In future, there are fears that literal interpretation of certain EU regulations by EU Member States will result in monitoring efforts being cut with the consequent loss of valuable information. Lowering the standard deviation of the estimate is the approach favoured in this report along with using all relevant available data. Combining adequately both approaches is the optimum approach for the best use of data.

The final benefit of stratified assessment is that the individual stations in samples are not required to be the same over time. Applied to water issues, this means that changes in monitoring stations may have no effect on the results provided that the stations monitor the same strata. The disadvantage of the flexibility offered by stratified statistics is that the strata are not precisely located. The different stages of development of the stratified statistics applied to water are detailed in the following publications: (Nixon, 1997; Leonard and Crouzet, 1999; OIEau and ifen, 1999; Beture-Cerec and ARMINES, 2000; Beture-Cerec and ARMINES, 2001).

3.2 Implementation to assess water issues

3.2.1 Work steps

Several work steps were carried out and are detailed in the next sections. They deal firstly with selecting the policies and the determinants to be processed according to the analysis of the questions to answer. In a second step, the different strata were defined with regard to the validation of causal correlation between determinants (e.g. ammonium) and sources of pressures (e.g. livestock) to be analysed. The third step consisted of calculating all data and relating them to each other on the relevant GIS layers. In a fourth and final step, the statistical correlations were quantified and trends assessed where the precision of the estimates allowed doing so.

In each step, some practical questions have to be solved to make it possible to apply statistical principles to reality, i.e. geography, data gaps, etc. These points are detailed in the corresponding sections.

3.2.2 *Policy basis and meeting targets*

In this report, questions related to nutrients in surface waters are addressed with respect to the main sectoral policies. The reduction in nitrate content is the aim of the Nitrate Directive (Directive 91/676/EEC, 1991). The Urban Waste Water Directive (Directive 91/271/EEC, 1991) aims to reduce phosphate and nitrogen compounds from urban sources as well as organic matter due to sewage collection and purification.

The assessment of the compliance with a Directive by any Member State is not within the mandate of the EEA. By contrast, the environmental assessment the EEA has to carry out requires trend analyses based on the same data including estimating the distance to meeting desired policy objectives. Among the relevant assessments of these distances is the differential between different regions, catchments and specific areas. In a large country, the assessment aggregated at country level can provide only a lumped image which does not reflect the diversity of situations, i.e. large areas free of problems, areas where the target is met or ready to be met, and areas where the status worsens with time. In this latter case, the distance to the desired objective widens.

Policy requirements state that 'status X' or 'concentration Z' should be met (not trespassed) in year Y. From an assessment point of view, such a simple requirement poses many practical questions. The simplest understanding is that any water volume, at any time and not under exceptional conditions, should meet the target. In practice, statistical assessment on aggregates is required because natural waters composition is subject to large variability, and not all water masses are monitored constantly. On the contrary, the monitoring density (spatial and temporal) is very low and seldom comprises observation of exceptional events.

Assuming that a statistical assessment of the target meeting is acceptable, the questions to solve are:

- What is the acceptable aggregate of water masses on which the statistic can be presented?
- Should all values of the statistical indicator meet the target (with a certain probability)?
- Can only the average trend of the statistical indicator meet the target (with a certain probability) to consider it as being met?

The graph in Figure 2 illustrates the issues raised by bullets two and three above.

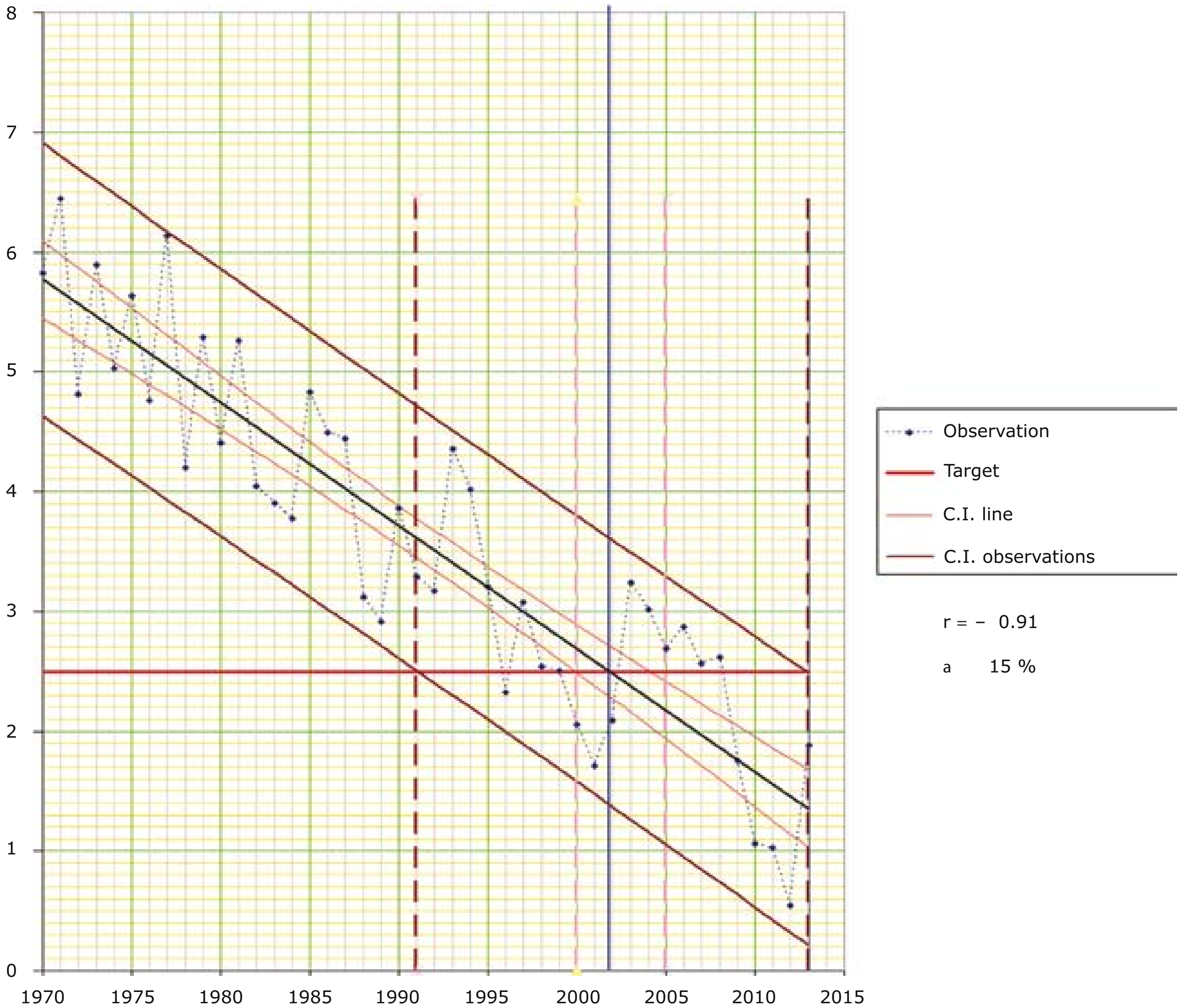
It shows random data mimicking a standard situation. The data tend to decrease in concentration with some variability despite very good linear correlation. A target of 2.5 is assumed and the year at which the trend line cuts the target is indicated by a vertical blue line. In addition, the confidence interval of the trend line for 85 % certainty (α = 15 %) and confidence interval of observations is reported as red and brown lines respectively.

The interval between the two thin red lines on each side and close to the correlation line can be interpreted as the 'thickness of the trend line'. It displays the domain of the second bullet, whereas the domain of the first bullet is between the two external thick brown lines.

In this example the target was met on average by mid 2001. In fact, it was met with 85 % certainty between 2000 and 2004. Similarly, observations belonging to the same population can be reasonably observed, with the same degree of uncertainty, below and above the target between 1993 and 2010. It is only after 2010 that new observations should be, with 85 % certainty, observed below the target. This last statement underlines the fact that in 15 % of cases values outside the wide strip are expected to be observed.

This is precisely what the example shows: a value below target is observed as early as 1995. Moreover, half of the eight data reported later than the target year are above the target. Choosing one or the other term radically changes the possible response and therefore the judgment to be made about policy effectiveness.

Figure 2 Impact of target meeting criterion on the obtained results (random data)



Note: C.I. = confidence interval

At EEA level, it seems wise to consider that the target is met when the trend line intercepts the target. Only aggregates representing catchments are to be addressed. This approach is less subject to large differences in sampling sites. The conclusions should be interpreted as *'on average, …, policy target is met in catchment X between Y1 and Y2 with this confidence'*.

The data set in the example has many interesting features that are observed in real chronicles. In particular, it can happen that a series of several years behaves in the opposite way of the general tendency. For example, between 1995 and 2005 there is virtually no tendency, whereas between 1980 and 1985 there is a tendency towards increasing values. These events are considered in trends assessments.

3.2.3 Selected determinants and assumed targets

The questions considered deal with nutrient trends, progress towards environmental objectives and the influences of sectoral policies on meeting those objectives. The selection of determinants for analysis directly derives from these relationships. However, none of the aforementioned directives set explicit numeric values (or targets) for the environmental objectives that can then be used for assessing progress. Instead for the purpose of this analysis, the assumed targets quoted in the EEA report on nutrients in European ecosystems are used (Crouzet, Leonard *et al.*, 1999). These targets explicitly refer to enforced directives that have not been repealed by the WFD, namely the Urban Waste Water Treatment and Nitrates Directives and the following parameters:

- **Ammonium.** The sources of ammonium are crude or non-oxidised purified urban sewage and industrial waste water as well as livestock emissions (ammonium from vegetal crop fertiliser should only result from accidental spillage because it is bound to soil components).

  Ammonium is both toxic (as unionised ammonia) to fish and unsuitable for drinking water because of the formation of chloramines during the disinfection process with chlorinated compounds. The WFD, which aims at achieving 'good ecological status', will probably retain the most demanding standards. For the sake of providing round figures, the values 0.05 and 0.5 have been used for processing in this application.

- **Phosphate.** Phosphorous compounds are mainly from urban sources (human metabolism and fabric as well as dishwashing products are major shares). Industrial wastes have diminished dramatically. In France, where detergents can still contain polyphosphate, the share of detergents borne P has dropped during the past years. However, no recent tonnage could be found. In agriculture, the issue is quite controversial as livestock is potentially a source, along with crop fertilisers. However, the low mobility of phosphorus in soils makes this source very dependent on erosion rather than on leaching. Phosphate thresholds are expressed as soluble phosphate (often quoted as 'PO_4' or 'orthophosphate', 'soluble reactive phosphate' and total phosphorus. Total phosphorus is bound to organic and mineral particles, and is expected to provide some information about agricultural inputs.

  Both forms can be considered as having the same eutrophication potential and the same thresholds are used for both, although they are respectively expressed in mg PO_4 L^{-1} and mg P L^{-1}. Values are 0.01 mg P L^{-1} (pristine conditions) and 0.05 mg P L^{-1} (mesotrophic conditions), corresponding to 0.03 and 0.15 mg (rounded) PO_4 L^{-1}.

- **Nitrate.** The potential sources for nitrate are crop agriculture, livestock emissions and to a lesser extent oxidised urban and industrial sewage.

The target values of 50 and 25 mg NO_3 L^{-1} are obviously irrelevant for yearly averaged values of river water and do not give any reasonable target. By contrast, 0.3 and 2.5 mg N-NO_3 L^{-1} are regarded as indicating pristine and downstream parts of rivers (ref nutrients report), yielding respectively 1.32 (rounded 1.35) and 11 mg NO_3 L^{-1} are used as reference and medium-term targets.

The latter value seems the most suitable in the perspective of 'ecological status' because nitrate causes little harm to river ecology.

3.2.4 Strata and stratum components

The aim of stratification is to define criteria to cluster catchments sharing the same criteria and process observation data according to this clustering. Schematically, the concentration of ammonium, phosphate and nitrate observed in rivers results from land-based inputs that are then diluted by run-off. Assuming that the volume of run-off is proportional to the area of drained catchment, the density of input per area of catchment drives the observed concentration. For example, if each hectare of catchment produces ~ 1 500 m^3 of run-off annually, an input of 30 kg during the same period will yield an average concentration of 20 mg L^{-1}.

Assuming a constant specific run-off, this ratio expressed in concentration is independent of the area of the catchment provided that sources are evenly distributed. By contrast, if large point sources are oddly distributed, small catchments will display large variability (those with source and those without source). Analysing the causal correlations is very important here. Note that the size of the catchment is not a stratification variable because it is not a cause. The run-off density (or productivity, expressed in volume per unit area per unit of time) on the contrary is an important cause of variations in results. However, this was not considered in the report due to a lack of data. Nevertheless, it can be assumed that run-off density is in a restricted range in a catchment. Hence, analysing the results per catchment should mitigate the fact that run-off has not been used for stratification.

The definition of strata has to consider those causes which are held by the catchment characteristics. Hence, the statistical population to stratify is the catchment. The stratification procedure will therefore group catchments in clusters of close characteristics, understood as resulting in the same concentrations at their outlet.

To be operative, stratified statistics require a limited number of strata that focus on the main potential drivers. Moreover, the variables used to cluster the monitoring points must be stable with time, otherwise no time-concentration relationship can be established and hence no trend assessed.

Seven strata have eventually been defined and were constructed to deal with the nutrient and organic pollution issues from agriculture and urban settlements. These strata are only operative for these drivers and determinants, and should not be used for assessing pesticides or metallic pollution, for example. The seven strata are:

1. **Weak pressures (code F, think to 'forest'),** representing catchments where close to negligible activities are exerted.

2. **Ordinary pressures (code X),** representing catchments where non-negligible, though not big pressures are exerted.

3. **Livestock (code B),** representing catchments where dominant and important pressure is related to the presence of several different livestock (all species together).

4. **Intensive crop agriculture (code A),** representing catchments where dominant and important pressure is related to the production of vegetal crops.

5. **Urban (code U),** representing catchments where dominant and important pressure is related to the production of urban population in 'reasonable' numbers.

6. **Mixed (M) [= A + U]** representing catchments where dominant and important pressure is related to both vegetal crops and urban population.

7. **Very urban (code V)** representing catchments where dominant and important pressure is related to the production of urban population in 'very dense' numbers.

It is virtually impossible to find 'pure' strata. All areas are urbanised to some extent and also often play host to arable agriculture activities as well as some livestock. The design of strata is therefore the result of statistical analysis in which each component is considered to define inclusion and exclusion criteria. The detail of the analysis is beyond the scope of this report and has been presented in the previous report (Beture-Cerec and ARMINES, 2000).

Each stratum is expected to present specific patterns related to the sources of impact. For example, the 'M' stratum is expected to exhibit ammonium, nitrate and phosphate whereas the 'A' stratum is expected to present nitrate and less phosphate and possibly low concentrations of ammonium.

The principles of the analysis are quite simple. For example, plotting *the concentration of nitrate in water at the assumed catchment vs the percentage of ploughed land in the catchment* gives a very good positive correlation. However, until the percentage of ploughed land reaches a certain value, the correlation is not usable because many external causes result in nitrate content, for example explosives plant, large city, livestock, etc.

Conversely, even when the percentage is high, other causes can interfere, for example large human populations and fertiliser-plant emissions. At the end, a set of rules can be established so that the information content of the strata is adequately captured by these rules, e.g. the percentage of land cover by a certain activity, population density, etc., each component of which is combined in a specific order.

Since complex cases exist in all circumstances, stratification is always a compromise. For example, the 'mixed' stratum accurately depicts a large part of northern France where big cities have historically sprung up close to crop production.

Once the different strata have been defined, building a stratum requires:

- identification of the relevant components of each stratum, e.g. those components having a deterministic influence on the observed phenomenon (in this case, the average concentration). These components are: Corine land cover areas, population density, and livestock density;

- checking the statistical liaison of each component with the observation. The detailed analysis provides interesting side results. This step helps to identify the stratum components that are relevant and the threshold values for inclusion ([2]);

([2]) Cropland is an important factor for nitrate. Should the first percent of catchment area covered by cropland be included? The preliminary analysis suggested that its impact is significantly above 40 %, after filtering with other characteristics.

- checking that different components defining the same stratum do not carry excess redundant information. Although proving they are really independent may be too complex, they must at least be checked for non-significant correlations. Furthermore, they should not directly derive from other variables (Beauzamy, 2004).

- constructing stratum assignment scenarios which consist of combinations of Corine land cover area percentages, thresholds in population density, etc. that best meet these principles. For the purpose of this report, they have been coded and applied using the facilities offered by Nopulu Système 2 installed at the EEA.

Hence, stratification **is a dynamic process** that requires the correlations between stratum components and observations to be accurately analysed. This key step requires the greatest amount of time in the study. It must be kept in mind that there are strong correlations between stratum components and observations and that the stratification process compresses and loses a part of the initial information for the benefit of simpler and more clear-cut conclusions. Expanding the application at the European level poses no methodological problem. On the contrary, a single definition ensures the best possible comparability and versatility. The following considerations are necessary:

1. Should new strata be defined or not in order to capture specific cases (e.g. high livestock and high population density could make a new mixed class)?

2. Should run-off be included as a stratification factor? This option is quite attractive, but poses many conceptual and practical difficulties. An interesting option is to use water productivity as a supplementary variable, as suggested by the EEA (EEA, 2001).

At European level, the fact that some strata may be missing in some countries poses no problem for the comparability of results. It just affects the selection of results which can or cannot be compared.

A final question relates to the permanence of strata. In principle, the classification of a catchment should change if characteristics of the catchment have evolved. In practice, characteristics data must be available (which was not the case). Moreover, the management of changing stratum reference *vs* stratification scenarios and monitoring stations is not straightforward. This is why the results are presented by reference to a permanent set of strata despite the fact that the elapsed time between the initial and final monitoring (30 years) would justify stratification adjustment. This modification shall be implemented when extending the methodology to European catchments.

3.2.5 Geographical extension of strata

Stratification is implicitly related to geographical area. The main constraint is that enough data must be present in each sub-unit. The major difference between statistical surveys and this work is that statistical surveys define the sample before data collection. In this case, existing data are applied to strata that have been defined after the monitoring has been carried out. Depending on the geographical extent of the assessment, some strata covering a certain area of land may have no attached data. This would not be the case in a statistical survey.

The choice of the best reporting units is therefore a compromise between the most relevant geographical 'sub-units', e.g. those that apportion the river basin districts, and the presence of enough data to carry out calculations. National-level reporting is always necessary but carries little consistency when comparing states with too large discrepancies in size, or just reporting the status and trends of exceedingly large areas in which positive and negative trends are blurred.

For example, France has defined 55 aggregation catchments ('BV RNDE' ([3])) whose average area is 10 000 km^2. Its size, comparable with the area of some of the smaller member countries of the EEA ([4]), would make them good candidates for infra-national assessment. The use of these catchments is not possible because of the lack of monitored data. Instead, a second set of catchments has been used that apportions France into nine aggregates (Juin, 2002, p 23).

3.2.6 Utilised statistical indicator

From a statistical point of view, stratification builds sub-populations which are understood as being more homogeneous than the overall population. The data for analysis are collected from the different

([3]) Abbreviation of these aggregation catchments is BV RNDE (BV for 'bassin versants' = catchment and RNDE for 'Réseau National des Données sur l'Eau' (national network on water data), now replaced by the SIE ('Système d'Information sur l'Eau' = Water Information System).

([4]) Luxemburg = 2 586 km^2, Cyprus = 9 521 km^2 for example.

monitoring sites belonging to the same stratum. This indeed poses the question of the most effective and relevant way to process this data and make indicators to carry out the assessments.

Merging all elementary data is not directly possible because of the oddity of sampling frequencies that would introduce bias in the event of changes in monitoring stations ([5]). This is, however, an interesting path to explore in terms of improving the accuracy of estimates resulting from probabilistic analysis ([6]) and providing stratum percentile estimates that cannot be defined with classical tools (Beture-Cerec and ARMINES, 2001). At European level, this is simply not possible because only yearly averages for selected stations are currently available through the Eionet-water data flow.

For the time being, the most feasible indicator is the stratum mean. This can be built from annual averages or extreme values at monitoring stations. The approach is very flexible because it can re-compute stratified statistics at different scales very easily, e.g. stratum at country level, stratum at catchment level, all strata at catchment level.

Despite the apparent simplicity, the station means must be computed under strict conditions. The following requirements must be met:

1. Considering the different number of samples analysed per year and the irregular sampling patterns that range between 4 to 24 samples per year, the mean at each station must be unbiased. Following previous studies (OIEau and ifen, 1999; EEA, 2001), time-weighted averages must be computed. This weighting mimics the composition of abstracted water, not river water. A representative mean of river concentration requires weighting by discharge, which poses both theoretical and practical problems. The way the Eionet data flows are averaged is unknown, further compounding such problems.

2. Extreme and central values are both important. In practical terms, a stratum indicator, if both minimum and maximum, is replaced by the average of minima and average of maxima recorded for stations entering in the stratum. Minima are likely to be impacted by the choice of the averaging period (see point 3 below).
3. Annual water statistics are more accurate and are computed following the hydrological year instead of the calendar year. Moreover, a whole 12-month period is not the best candidate for the averaging period. For example, nitrate concentration is close to zero in summer whereas ammonium concentration is close to zero during winter periods. Seasonal averages must be computed that consider both winter periods for potentially non-point source, driven by rainfall leaching (nitrate) and summer periods for predominantly point-source components, driven by dilution (phosphorus, ammonium). However, the assessment of benefits from using the focused period is partly jeopardised by the quantity of available data; applying a sounder method results in the rejection of stations where the number of samples is low. Consequently, the hydrological year was used in this study.

The variance of each annual data set has not been computed because it would require removing the impact of seasonal patterns, which is not yet feasible in a systematic way (EEA, 2001).

3.2.7 *Presentation of results and analysis of trends*

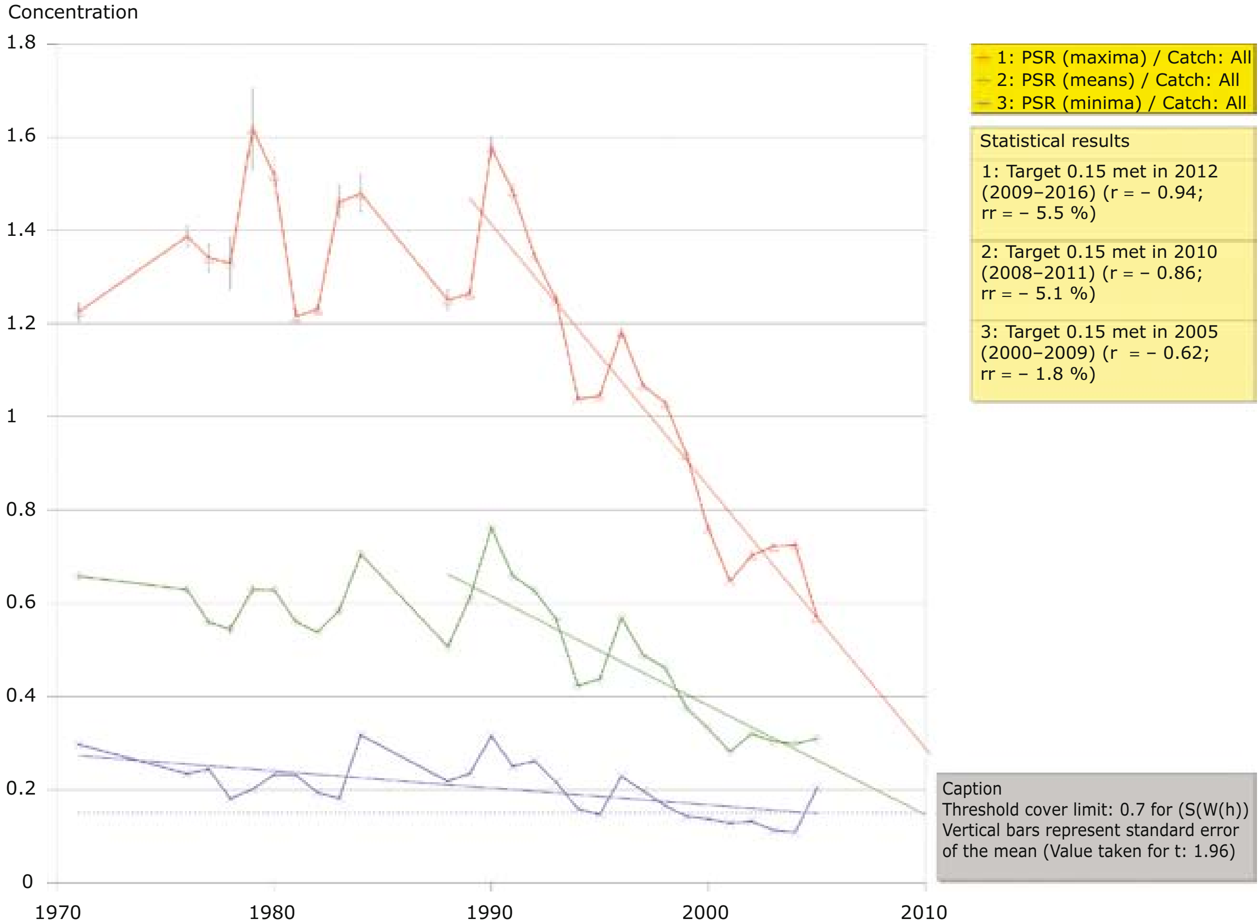
The results are presented in two different ways which respond to the types of questions exemplified in Section 2.1. To the first question, the accurate response is for example *'the average yearly nitrate concentration in France, all catchments, was 13.3 +/– 0.038 mg NO_3 L^{-1} in 2004, at 90 % confidence, computed from 1966 station averages'* and *'The first year of measurements, 1971, yielded only 6.02 +/– 0.15 mg NO_3 L^{-1}, at 90 % confidence, computed from 435 station averages'*. Care should be taken to ensure that the uncertainty of the value represents the error of calculation of the average and not the dispersion of values making the averages.

The response to the second question is more complex, because trend analysis requires a definition of a method of trend assessment. There are several statistical methods to this end. Since one piece of information sought is the date on which a certain target could be met, the regression method is the most appropriate.

The typical indicator is the *time* in years, plus or minus *x* years when a certain value expressed in concentration is expected to be met for different strata (see precise definition in Section 3.2.4), and considering relevant determinants. The information available is presented in Figure 3 where the

([5]) This is not contradictory due to the fact that stratification is not sensitive to changes in stations.
([6]) This option is under analysis for the time being and should result in findings at the end of 2007.

Figure 3 Typical representation of one of the indicators to be produced



Note: The three lines show descending concentrations of soluble phosphorus, for all catchments in France, all strata (respectively maximum value, average values and minimum values). They converge to a target value set at 0.15 mg $PO4L^{-1}$, whose time target is reported in the comments box. The trend line start is represented. In the displayed graph, correlations start between 1970 and 1990, according to acceptable trend display.

The three boxes report the major indications: top box (legend, gold yellow) mentions the reference of each graph. The medium box (ivory) reports the target and data of achievement according to the regression lines on the graph along with the numerical value of the target. The third box (grey) reports the statistical hypothesis of the extract: percentage for inclusion. In the displayed graph, all areas where at least 70 % of total weight was met were included. The t value refers to the uncertainty probability of each point, displayed as vertical bars.

development of the time of soluble phosphate is presented; all strata for all catchments. This figure was selected because it presents the different problems that are raised by trend analysis.

The three assessments (maximum, average and minimum) present a clear trend that is more or less linear over different periods. Linear trend analysis is performed because it is not very sensitive to the beginning and ending of periods, which is not the case for curved or polynomial trend analyses. In the real world, it is more likely that amortised trends tending to reach an asymptotic value would be more accurate. An amortised trend mimics the fact that easy measures are taken first, resulting rapidly in positive results. The hardest measures are taken over time and consequently, the impact on results is less. Therefore, the linear trend is likely to be quite optimistic.

In practice, the trend analysis seeks the best linear adjustment between stratum concentration and time. A typical example is illustrated by the graph in Figure 3. This shows an adjustment of the theoretical graph in Figure 2. The first and last year are selected by visual assessment of the quality of the relationship. This analysis has the advantage of showing clear-cut patterns at the expense of drastic simplifications of background information. The following implicit hypotheses must be kept in mind when considering the graphics:

- The observation can present change in trends; only the most recent stable trend is computed:
- the existence of stratum is assumed permanent along observation periods and during the forecast period; this cannot be proved;
- observed trends are supposedly being prolonged at the same rate during the forecast period; this assumption is likely to be too optimistic for the reasons reported above.

Consequently, if the preventive and curative measures implemented during the past 10–30 years were to continue *at the same rate* for the type of long period in the future often required to reach concentration targets, the projections that are reported present *the most favourable case* of possibility to meet objectives. The provided forecast should be understood as indicative.

Trend analysis has been carried out at different levels starting with the whole of France and all strata and ending with analysis of trends for each stratum within a single catchment (nine large catchments). This differentiated approach aims to assess water status (objective 1 of the report) and the driver-related impact (objective 2 of the report).

The results start with the most aggregated and end with disaggregated assessments. The final conclusion discusses the difficulties and possible misleading conclusions that result from an aggregated data set: aggregated data are more suited to providing a weighted status than they are to assess policy effectiveness that is more context (and area) dependent.

Consequently, the analysis of the quantity of data required is based on the needs to: i) produce regionalised statistics, ii) assess the accuracy and interval of confidence of the time estimate and iii) provide the complementary assessment of *stratum vs time* relationship to evaluate one of the components of policy effectiveness. A part of the statistical methods required to accurately assess these parameters is not yet available and proxies of these evaluations are produced instead. For the time being, the uncertainty around the target data is not computed.

3.2.8 Source apportionment

The final results are weighted averages of concentration. The weighting factor is the proportion of catchment (captured as relative number of catchments) falling into a certain stratum. The relative contribution of each stratum to the weighted mean is therefore a proxy of source apportionment within the considered area.

- To ease the reading of results, two groups have been defined: the 'urban group' (marked as U in the figures) where all contributions primarily related to urban activities are grouped, and the 'agricultural' group (marked A in the figures) where activities related to agriculture are clustered. Preliminary results are presented. They can be refined by deeper statistical analysis, considering the partial components of minority contributors.

This approach to source apportionment is based only on observation data and is comparable with the models used to assess the same information. When comparing with other information, it must be kept in mind that the stratified averages compare time-weighted means and not discharge-weighted means.

Apportionment analysis tends to nuance the encouraging results presented by sector and explain the apparent paradox between the large improvement in key sectors and the fact that targets are not met on the overall analysis.

The figures are all presented in the same way above and below a horizontal reference line. The baseline contributions (F and X) have been placed symmetrically on each side of the reference line. The urban group percentages have been set as negative values in order to place them under the reference line, whereas those from the agricultural group lie above it. The average mean is plotted on the figure to indicate which quantity the apportionment refers to.

3.2.9 Practical problems

Ascertaining catchment and station relationship
The fundamental constraint intrinsic to this type of analysis is that the monitoring station *must* be placed on the main drain of the catchment, otherwise there is no causal relationship between upstream catchment characteristics and water composition at this point. This part of the station documentation is not very accurate because river and catchment GIS are not yet fully operational. For this analysis, stations were placed on the main drain provided they were at a reasonable distance (100 to 500 m ([7])) from the river.

Processing ill-documented strata
It may happen that some strata have few or no monitoring points for a certain geographical area (e.g. a basin does not have enough points to allow calculation). A very rigorous approach would result in discarding the analysis and is thus not very operative. Practical calculation and thresholds have been set to allow, under certain conditions, the extension of other strata in this case. The details of this approach are included in Appendix 1.

([7]) Stations are located by ground coordinates, which are generally accurate. The river is represented by a line, resulting from GIS modelling. There is little chance that a real position on a river bank of a river often several hundred meters wide could precisely match a line resulting from independent processing. The catchment & rivers database (under preparation) relates objects to river reaches to avoid spurious attachment of a station to a river.

4 French case study results

4.1 Summary results

4.1.1 *Status*

At national level, undisputable and dramatic improvements in phosphate and ammonium concentration is observed. This improvement is certain for both means and maxima and sometimes for minima as well. Despite constant improvement, the current status is not yet satisfactory because targets are not met. Meeting the targets may require one or two decades of improvements at the same rate observed between 1980 and 1990. This finding shows that the average concentration of water from all catchments tends to decrease ([8]). However, since the discharge is not a weighting factor, it cannot be stated that the final fluxes to the sea decreased by the same proportion (i.e. they are expected to be less but may have increased locally).

Since ammonium has a direct impact with no threshold effects, it is expected that the cuts in concentrations have resulted in significant quality improvements. By contrast, phosphate has no direct impact. Moreover, it controls eutrophication only below a certain threshold that is not yet reached by the average. Minimum values tend to reach averages below targets. Hence, it is expected that little effect on eutrophication was observed.

The situation with respect to nitrate is still unsatisfactory. The trends may be interpreted as persistent degradation. However, many values suggest a possible change and reversal of the general tendency towards degradation.

Results by basin show many different patterns. This strongly suggests that the apparent constant and smooth changes at the national level are the result of a complex patchwork of very different responses to measures that were progressively implemented with strong incidences of local conditions.

4.1.2 *Sectors — pressure*

Sectoral impacts are assessed through stratum. For instance, the concentration relationships analysis requires consideration of the distribution of strata in the territory, since some strata are related to specific regions. This is shown in Figure 4. The different strata are not equally distributed across metropolitan France. In most basins, agriculture is present as a potential dominant pressure, often concurrently with population as mixed stratum (M). The case of Brittany and the western Atlantic is very special because it is dominated by livestock (B) with a strong share of mixed stratum. This would explain the patterns of contamination observed.

A large strip SW–NE is dominated by agricultural activities: intense in the most western half of the strip and less intense in the eastern side.

The non-impacted (F) stratum is absent in most catchments and located next to mountainous areas, with local spots in plains.

As expected, urban (U) and very urban (V) strata are in the minority. Paradoxically, the Paris area does not stand out as densely populated because i) its position on the Seine catchment keeps the

Table 1 Summary results, all strata and catchments

	Ammonium (in mg NH_4 L^{-1})	Soluble phosphate (in mg PO_4 L^{-1})	Nitrate (in mg NO_3 L^{-1})
Starting value	0.75 +/− 0.006, in 1971 (383 annual averages)	0.61 +/− 0.0009 in 1989 (1 096 annual averages)	6.02 +/− 0.15 in 1971 (435 annual averages)
End value	0.41 +/− 0.003, in 2005, (1 210 annual averages)	0.31 +/− 0.0007 in 2005 (1 459 annual averages)	13.3 +/− 0.04 in 2005 (1 966 annual averages)

Note: Soluble phosphate starting value is chosen at peak concentration.

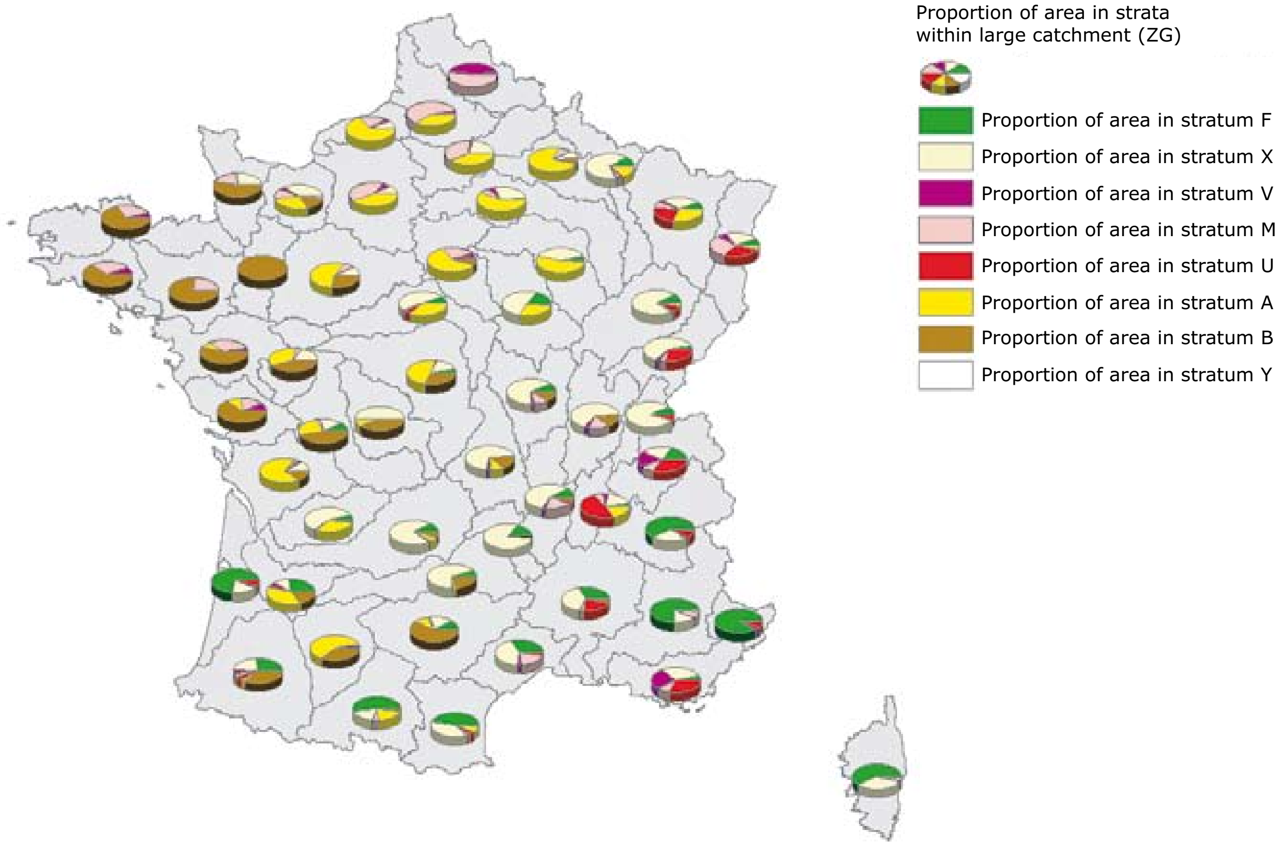
In all cases, precision of estimate increases with the number of data included. The change in stations number should not impact the overall accuracy.

([8]) To be fully accurate it is the weighted average of concentration, considering the number of elementary catchments falling into each stratum category. The results might be slightly different if catchment area was used instead.

Table 2 Relative percentage of catchment numbers per stratum

Stratum	A 'Intense crop agriculture'	B 'Livestock'	F 'Low pressure'	M 'Mixed 'A+U)'	U 'Urban'	V 'Densely urban'	X 'Ordinary pressures'
percentage	18 %	19 %	14 %	9 %	6 %	4 %	31 %

Figure 4 Geographical distribution of potential pressures at the level of 55 RNDE basins



population density below the threshold and ii) the upstream part of the Seine catchment is occupied by dense agriculture, making the dominant stratum M instead.

The summary distribution of the different strata is in the next table.

Despite being larger, catchments with low activity present concentrations close to 'pristine' values. This suggests background or local contamination. They can nevertheless be used as proxy references or targets.

A special finding is that it was not possible to demonstrate a statistical correlation between intense agriculture and ammonium or phosphate content of rivers at the level of assessment carried out. This finding does not deny causal correlation; it just demonstrates that either the contamination from these substances is low in the analysed areas or stratified statistics might not be an adequate tool to address a low-grade effect.

This finding poses another question concerning the possibility of upscaling the detailed studies (that show causal correlation at the elementary catchment level) at aggregation basin level. This question deserves greater attention, because the modelling of agricultural activities is based on the paradigm that elementary mechanisms can be upscaled. Apportionment studies suggest that agriculture could be responsible for 25–75 % of phosphate in rivers. This is clearly not demonstrated by the findings presented here, although France is not one of the analysed countries in such studies (NERI, 2005, p. 32).

By contrast, the urban dominated strata are undoubtedly the cause of phosphate and ammonium, whereas crop agriculture and, unexpectedly, densely urbanised areas are the source of high nitrate concentration. In the latter case, nitrate decreases along with the decrease in ammonium and phosphate. In this case, the aggregation basin survey just fits to the findings of local assessments.

The correlation between livestock and water contamination is quite complex. Where cowshed breeding occurs, contamination is indirect. It is likely because the manure management facilities lead to elevated nitrate concentrations.

These findings are not precise assessments of the effectiveness of the two major policies, the Nitrate and the Urban Waste Water Directive, because not all driving factors have been considered. There is, however, a strong likelihood that the following improvements are the outcomes of the implementation of theses policies:

- Where urban wastes (point sources) are involved, dramatic improvements have been achieved on the most polluted areas with no measurable time lag. Urban and industrial waste-water purification has a direct positive impact on receiving waters and the action of French water agencies started having an effect from the early 1970s (much earlier than the implementation of the Urban Waste Water Directive). This implementation has fostered more complete processes (P and N removal that were not included in the oldest activated sludge files) and increased the momentum of works realised. These findings are supported by the recent policy-effectiveness analysis made by Kaczmarek (2006).

  Apparently, phosphate removal has been boosted by the implementation of the Directive, by national programmes aimed at preventing eutrophication in the populated areas that started in the early 1980s and by the reduction in content of polyphosphate in textile detergents. The latter reason had already reduced the person-equivalent by ~ 20 % by the late 1990s (Lavoux and Rechatin, 1998). By contrast, in the intermediate areas, where activities below the national average are hosted, the situation does not seem to have improved but may even be worsening.

- By contrast, the Nitrate Directive implementation does not seem to have reached its objectives. The positive point is that some of the worst situations seem to have stabilised and improved locally. Since nitrate concentration in rivers is the outcome of a complex and lengthy process, it clearly poses the question of how accurately the time lag between a measure and its observable result can be assessed. For example, the transfer pathways between input and receiving waters are far more complex and lengthy in non-point source pollution than in urban wastes transfer. The residence time in the groundwater reservoir may range from a few months in fractured granite areas to centuries in chalk thick systems between soil and outlet.

4.1.3 Source apportionment

Source apportionment estimates and trend analyses are new outcomes of stratified assessment. Current source apportionment estimates are a 'global' indicator which reflects the relative contribution of the different driving forces to the time-weighted average water concentration. It thus gives an estimate of the relative contribution of different sectors to the concentration of abstracted waters. It does not compare to flux-based estimates that are based on discharge-weighted concentration values.

Despite dramatic improvements in ammonium concentrations in urban catchments, the urban share of ammonium increases with time. This suggests that the agricultural share has become less and less significant. By contrast, the non-urban share of soluble phosphate tends to slightly increase, despite the overall decline in phosphate concentration, suggesting that the agricultural share is not improving owing to lack of progress with improving the efficacy of inputs such as fertilisers.

Finally, both the nitrate concentration and the agricultural share tend to increase. This should be considered as a clear warning, despite the possible stabilisation of maximum concentration in nitrate.

4.1.4 Utilised data

The work carried out consists of compressing and aggregating data several times. This may be quite confusing for a reader not familiar with statistical processing. The different levels of aggregation are as follows:

1. Elementary data (concentration measured that day at this point) comprise the period 1970–2005, 340 218 nitrate values, 337 126 ammonium values, 336 016 soluble phosphate values and 236 562 total P data.

2. Point statistics (average, minimum, etc.) for each monitoring point (per year and determinant are computed from elementary data on an annual (hydrological, seasonal, etc.) basis. This process yielded 41 399 nitrate averages, 38 425 ammonium averages, 39 288 soluble phosphate averages and 35 861 total P averages.

3. Point statistics are aggregated at the stratum level (same dominant pressure, defined area), (per year: 7 strata × 55 or 9 catchments × 3 statistics = ~ 1 150 to ~ 147 results, not all being populated because they do not all exist).

4. Stratum statistics are aggregated at the reporting level (e.g. France, weighting all strata together gives per year three weighted statistics from the ~ 1 500 elementary data collected per year and per determinant).

5. Trend is computed on the final statistics at country, stratum or stratum × catchment level.

The stratified approach cannot provide clues to the detailed local relationship. This requires in-depth analysis of local data. Nevertheless, it provides robust assessment of trends related to dominant forces that would be blurred by the absence of data aggregation. However, it is essential to ensure that each monitoring point is correctly attached to the relevant stratum. The data preparation and validation phase is therefore extremely important.

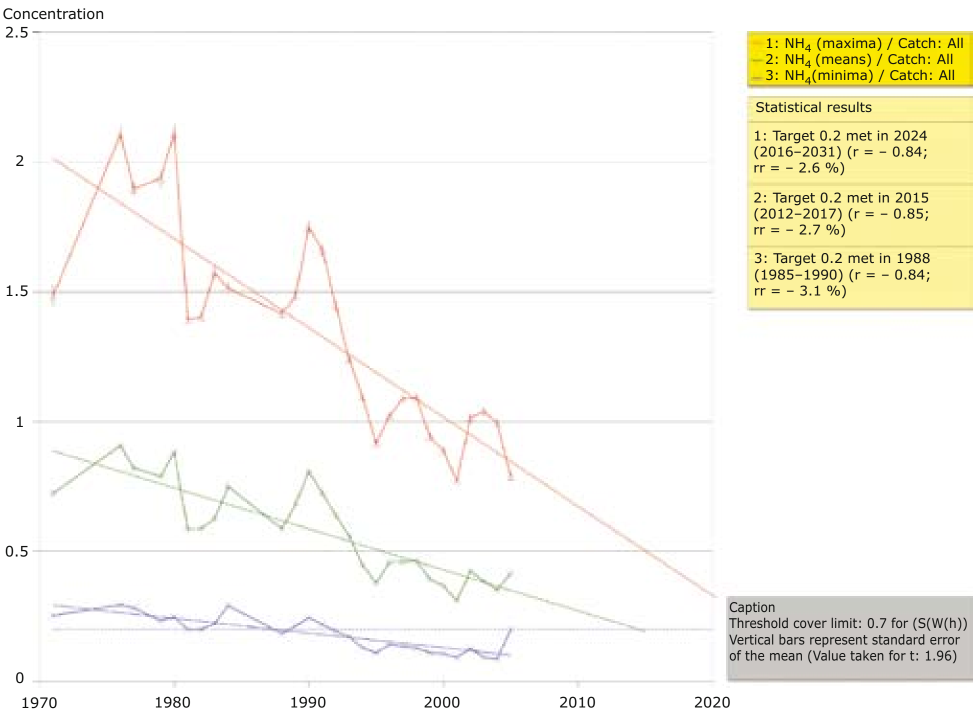
4.2 Status assessment

4.2.1 Development of ammonium concentrations

Nationwide results (ammonium)
The target shown in Figure 5 refers to 'cyprinid fish' conditions. Good conditions for the averaged maxima of ammonium (the most unfavourable conditions) are not expected to be achieved before 2024 assuming purification continues at the same rate in future years compared with previous years. By contrast, average conditions should be achieved already by 2015.

The long period of time needed to achieve the target for maximum values might be increased if drought condition frequency increases. This is the case

Figure 5 Ammonium trends in France, all strata and catchments together



Warning: All year ranges refer to the trend line uncertainty and not to the possibility of having an averaged value observed below the threshold, as discussed in 3.2.2 and illustrated in Figure 2. It refers therefore to the 'thickness' of the trend line, with 90 % certainty. The other statistical assumptions are reported in the graphcs.

This warning applies to all graphics.

because the maximum concentration of ammonium directly depends on dilution by discharge during the low-water period.

If replacing this target by 'pristine' conditions', the year of meeting target would be shifted to 2028, 2024 and 2014 respectively for maximum, average and minimum trend assessment.

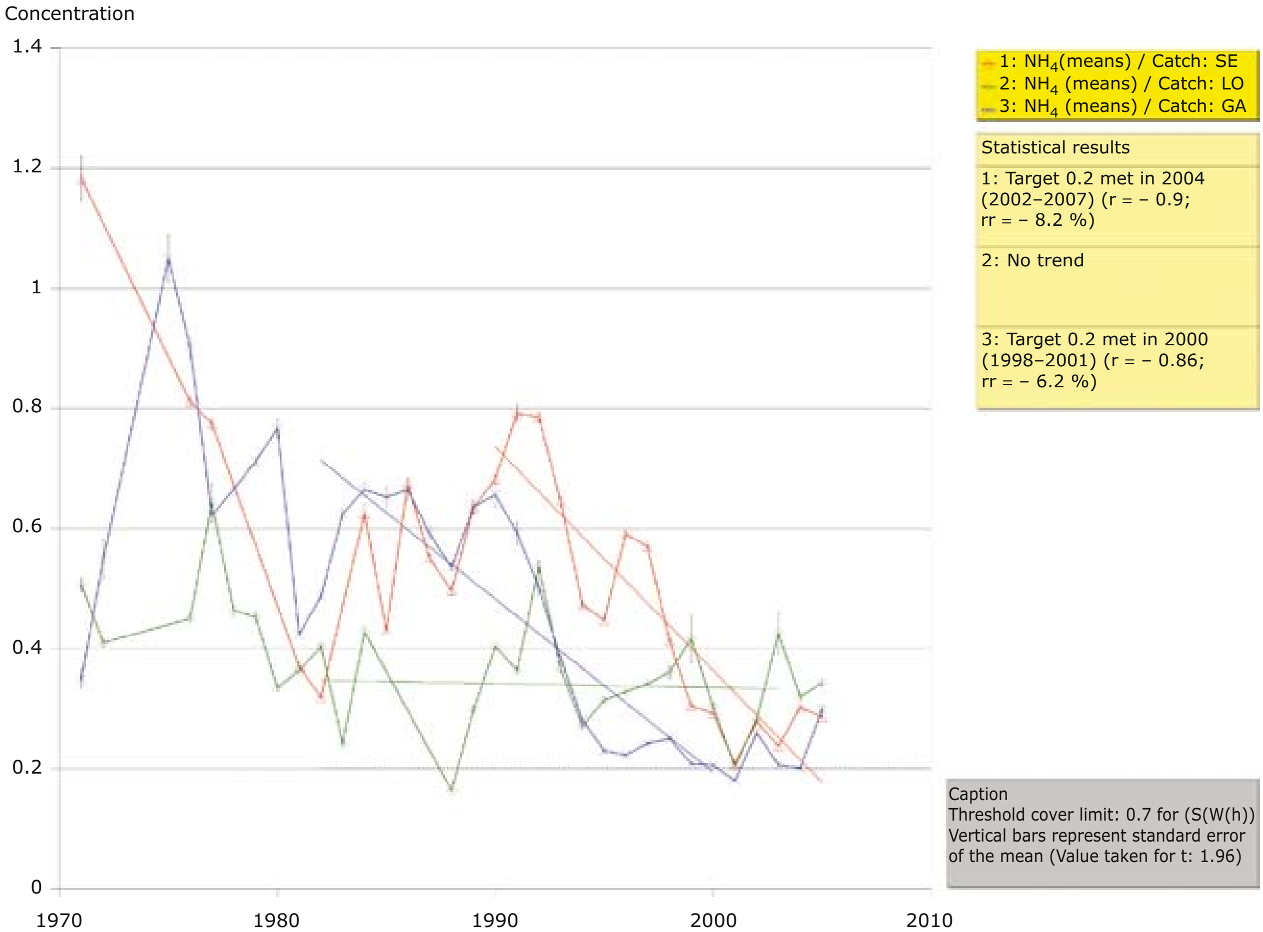
Ammonium values do not suggest any change in trend over the period. This might be due to the sewage purification processes. Since installation, all organic nitrogen forms have been processed. The general feeling that the oldest biological files produced ammonium as a result of the short residence time of sewage is not supported by the general trend analysis. It is more likely that this permanent trend is an artefact resulting from very different patterns.

Catchment analysis (ammonium)
At country level ammonium trends seem quite regular. On the other hand, catchment level analysis displays chaotic behaviour. Nevertheless, it confirms the overall tendency, but requires causes to be considered separately and unusual patterns to be questioned.

The large river catchments are more or less in line with the general result, namely: the Seine catchment presents greater concentrations in ammonium than the others. However the difference is significantly smaller than expected. This hierarchy is not confirmed by the distribution of maxima. For example, the Garonne is the most contaminated catchment with respect to this indicator between 1972 and 1992.

Many catchments show a strange relative minimum during the period between the mid-1970s and 1990. This period is less documented because only three full-scale campaigns were carried out (1971, 1976 and 1981). This cannot constitute one single interpretation because the stratification technique is quite robust. The diagram confirms, however, that many years have not been displayed because data did not meet the quality standards needed for inclusion (presented in the grey box of

Figure 6 Ammonium development (mean values) in the Seine, Loire and Garonne catchments, all strata combined



figures). The lack of trends for the Loire catchment deserves analysis and, possibly, adjustment by local authorities.

Despite extreme variability between 1970 and 1985, the small catchments on the Atlantic coast display similar behaviours. A special mention should be made of the very elevated value in the Brittany basin (BE) in 1971. Despite numerous livestock and quite elevated population density, its trend is very positive. The English Channel catchments (MA) basin, which presents some similarities with Brittany, displays rather identical patterns over the past two decades.

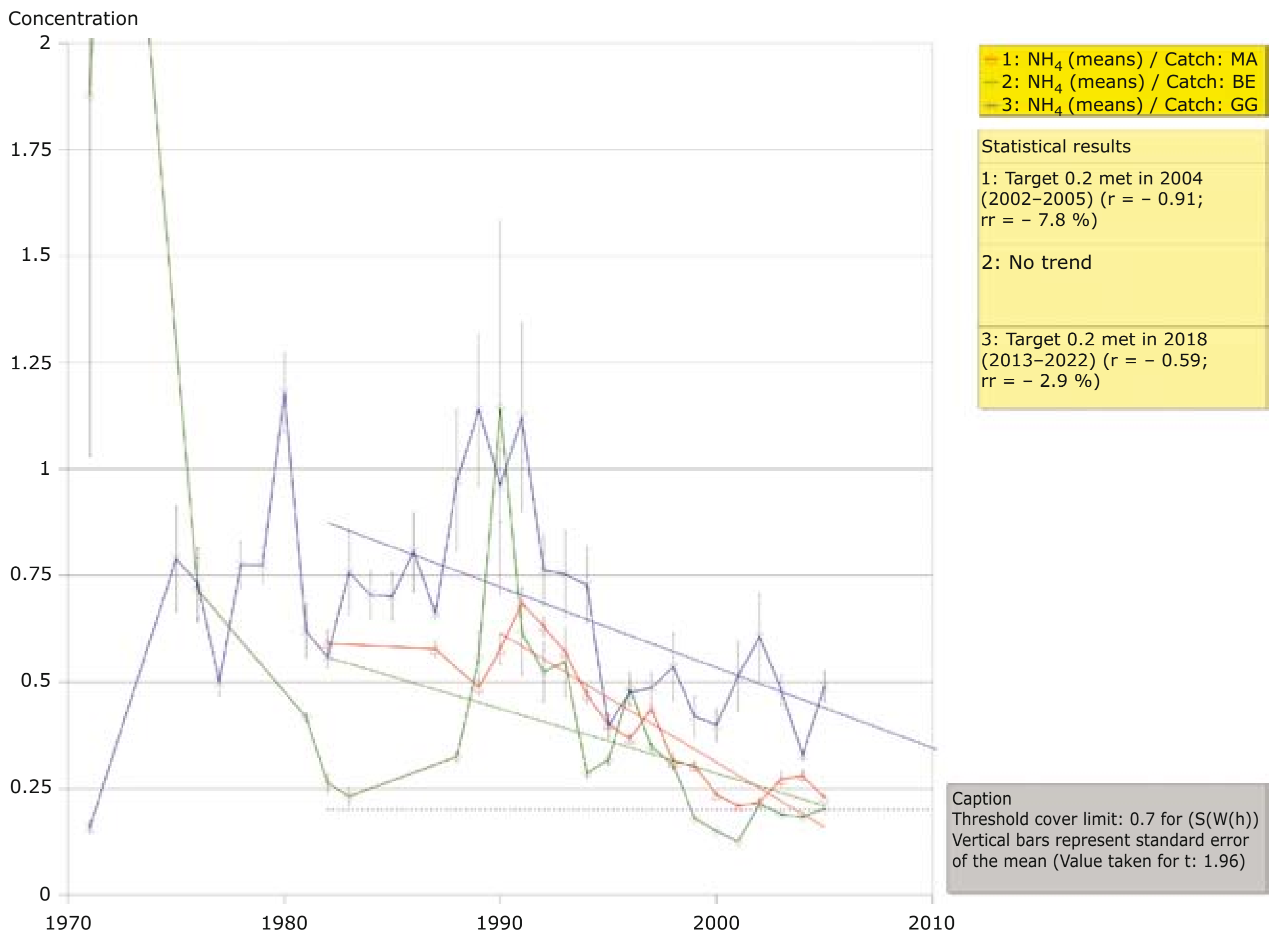
By contrast, the small catchments of the Atlantic coast (Bay of Biscay = Golfe de Gascogne = GG) present a very typical pattern of increases (due to development of emissions prior to purification) before a second phase of improvement (due to sewage purification and limiting adverse emissions). In all cases, average target meeting is expected or achieved.

The catchments flowing to the North Sea (Escault, (Scheldt), Meuse, and the French tributaries of the Rhine) show expected trends, but are not documented in the first decades (Figure 8).

The Mediterranean basin displays the only example of degradation that seems systematic enough not to be attributed to any artefact. However, detailed analysis (which goes beyond the scope of this report) would possibly suggest a series of declining concentrations interrupted by peak averages. Consequently, this results in increasing average trends.

The catchment level analysis confirms that ammonium decreases in all catchments, with the exception of the coastal Mediterranean catchments. In these catchments, a statistically significant increasing trend is detected. All reported data present either a chaotic variability between successive years or an inaccurately assessed, yearly indicator. As ammonium is linked to urban wastes (see next section), the concentrations have a strong

Figure 7 Ammonium development (mean values) in the small catchments of Atlantic Ocean, all strata combined



hydrological correlation. This is not always direct (i.e. lower dilution in low-water periods vs possible increase of self-purification related to greater residence time when discharge is low. This results in the opposite impact for the same hydrological causes.

There is some apparent discrepancy between the nationwide trend, suggesting 2015 (2012–2017) as time of meeting standard, and the different times of standard-meeting in the nine basins that suggest success at an earlier date. The major reasons for this are:

1. The nationwide assessment takes all data, including those from basins where change in trend is observed. In some basins the individual assessment does not start with the beginning of observations.

2. The nationwide assessment includes stable basins (e.g. the Loire and the Mediterranean) where concentrations increase.

3. More data are included, because the threshold for inclusion applies to a lesser extent (9).

4.2.2 *Development of phosphorus compounds*

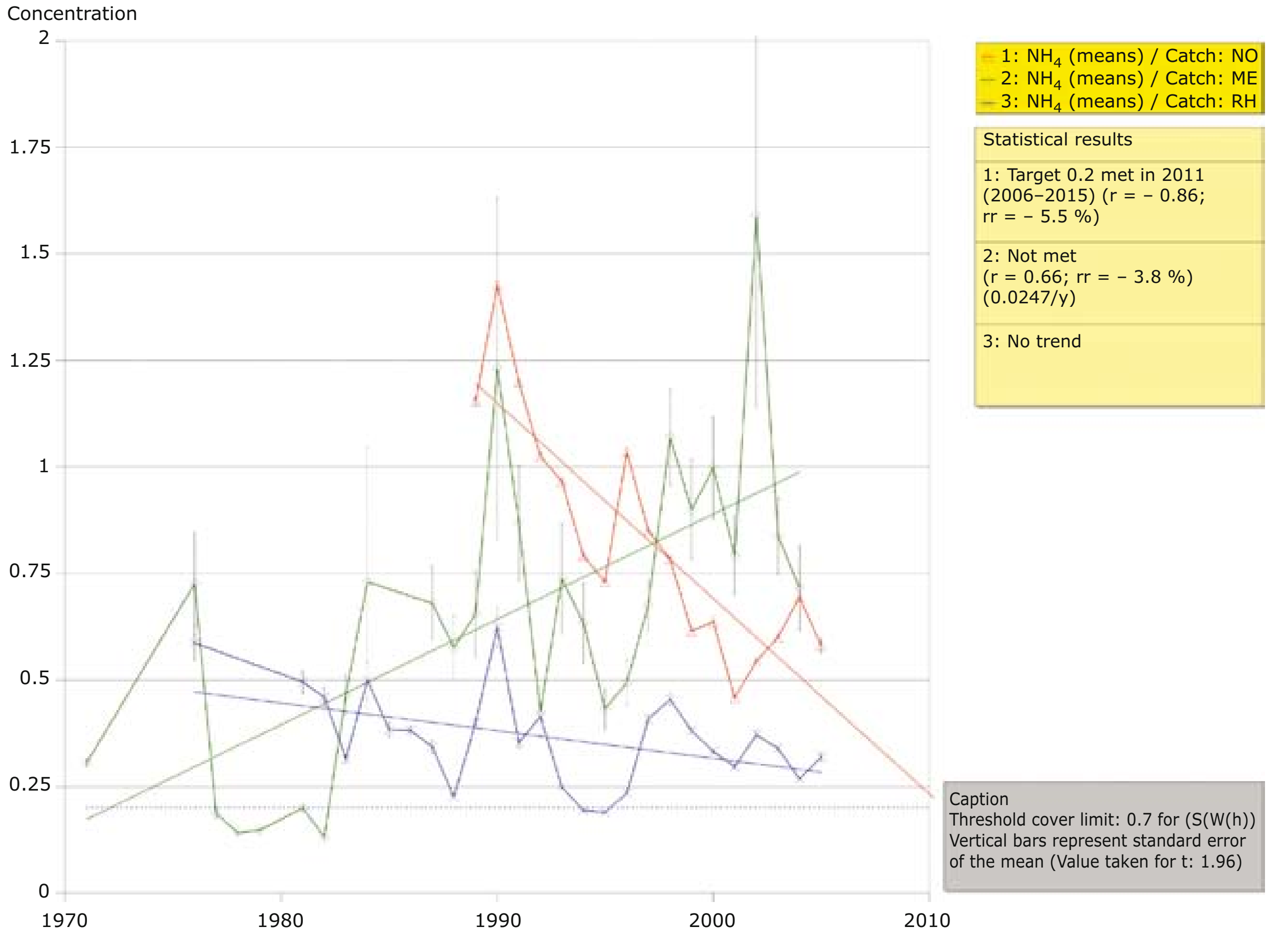
Nationwide results

Phosphorus compounds present a rather complex pattern, as suggested by Figure 9. Maximum values are stable until 1988–1990, after which they start decreasing. Average values show the same behaviour, and possibly start earlier (1982–1984). By contrast, minimum values have decreased steadily since the beginning of surveys.

Soluble phosphorus and total phosphorus present the same patterns.

Selecting the consistent time of decrease in concentration is partly subjective: by chance maximum and average trends meet the target in the same year. However, there is wider uncertainty for the maxima trend line.

Figure 8 Ammonium development in the Mediterranean catchments (Rhône and small coastal catchments) and North sea catchments



(9) When the geographical area of assessment is small, there is a bigger possibility that some strata are not included, as not enough stations exist for certain strata. This is not the case when the area is larger.

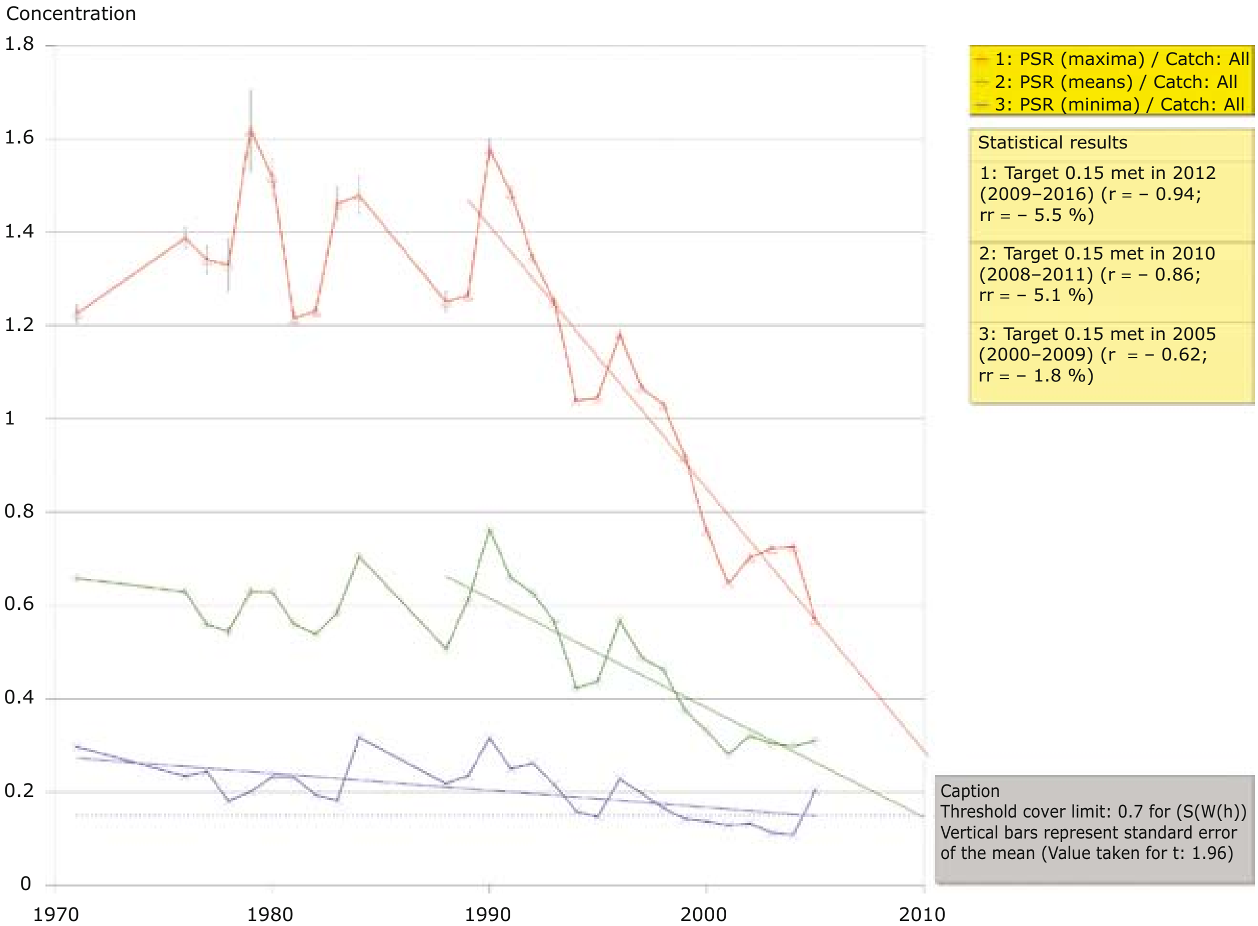
Total phosphorus data have only been available at national level since 1990. The trends are very similar to those shown by PSR, with stronger decreases in total phosphorus. This is confirmed by the relative slopes that are – 6.1/– 4.4; – 5.1/– 5.0 and – 1.8/– 4.5 (PSR/P total) respectively for maxima, averages and minima. The starting minima in total P were much higher than those in PSR. This explains the higher rate of decrease. Soluble phosphorus represents only a fraction of total P, e.g. in 1990 the P-PO_4 is ~ 0.5 against 0.7 mg L^{-1} for P total.

Since the target has the same value, it was met for total P only a few years after its achievement for soluble phosphorus. However, the issue of source apportionment is crucial for estimation if the trend is likely to continue linearly by improvement of purification plants. There is also the question of whether diffuse sources should be involved. This is discussed in Section 4.3.2.

The overall decrease in concentration of ammonium and PSR results from the better collection of municipal sewage, better efficiency of sewage treatment plants (STP) applied to larger capacities (Juin, 2002, p 263) and, in the case of phosphate, a constant decrease in P content of laundry products as well as the development of P-removing waste-water purification plants (biological and chemical precipitation files). Ammonium decrease is likely to benefit from systematic reclaiming of contents and purification of industrial waste-waters containing organic nitrogen (canning, slaughtering, etc).

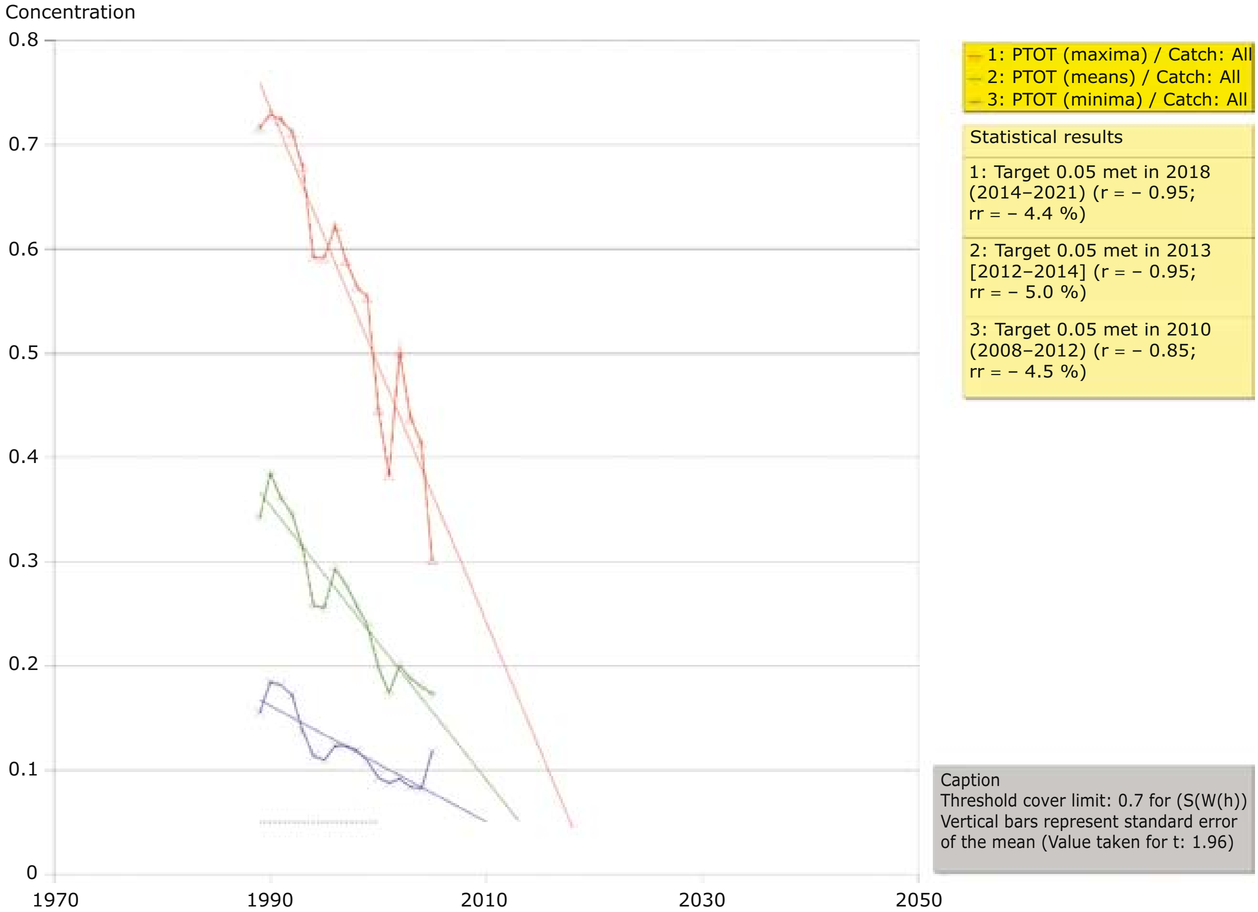
Catchment analysis (phosphorus compounds)
The three major basins of the Atlantic display the same pattern as the nationwide assessment: an erratic, rising phase followed by a constant decline starting between 1980 and 1990. The graphs in Figure 11 show a very marked decreasing period between the mid-1970s and the end of the 1980s for the Seine and Garonne catchments. This was already mentioned in the presentation of the ammonium trends.

Figure 9 Soluble reactive phosphorus (PO_4) development with time, all strata and catchments combined



Note: 0.15 mg PO_4 L^{-1} is 0.05 mg P L^{-1}.
Data in mg PO_4 L^{-1}, same as Figure 3.

Figure 10 Total phosphorus trends development with time, all strata and catchments combined



The graphics of maxima in Figure 11 show, by contrast with the Figure 12, a much dispersed confidence interval of the assessed annual means.

Despite the fact that trends in maxima are extremely favourable, the low-ambition target (mesotrophic, capable of sustaining at least 50 mg Chlorophyll-a m^{-3}, i.e. ~ 1 m Secchi depth) seems achievable. If the 1990–2005 trend continues, the target can be met for the average maximum between 2008–2014 (2007–2015); approximately a score later than the target for the annual means had been met. The coming 10 years are therefore critical. Assuming that the average maxima ranged ~ 1.5 mg PO_4 L^{-1} by 1990, the current value for 2005 lies in the range of ~ 0.5 mg PO_4 L^{-1}, and the average 1.0 mg PO_4 L^{-1} has been removed over the ~ 1.35 mg PO_4 L^{-1} required to meet the target. This concentration might be the most difficult to address: Despite soluble phosphorus being largely related to urban sewage, contributions from other sources cannot be overlooked. This view is strongly suggested by the apportionment analysis shown in Figure 27.

The trends are quite comparable in the smaller catchments of the Atlantic coast. Concentrations in Brittany, which were the highest recorded at the catchment level, display a very strong trend towards improvement. The maximum concentration trend should meet the target in 2003–2008 as well (graphics not reported).

In all cases, data before ~ 1990 provide extremely erratic averages: confidence intervals of the estimates do not allow any comparison between years and catchments in many cases. By contrast, others are very narrow in the same period. The situation greatly improves after the indicative date of 1990. This is likely because of the systematisation of measurements, laboratory controls and increases in sampling numbers that resulted from the renewed *Réseau national de bassin* launched in 1987.

The only possible negative change (albeit not statistically significant and therefore reported as 'no trend' on Figure 14), is recorded for the coastal Mediterranean basin ME. However, analysis of

Figure 11 PSR development (mean values) in the Seine, Loire and Garonne catchments, all strata together

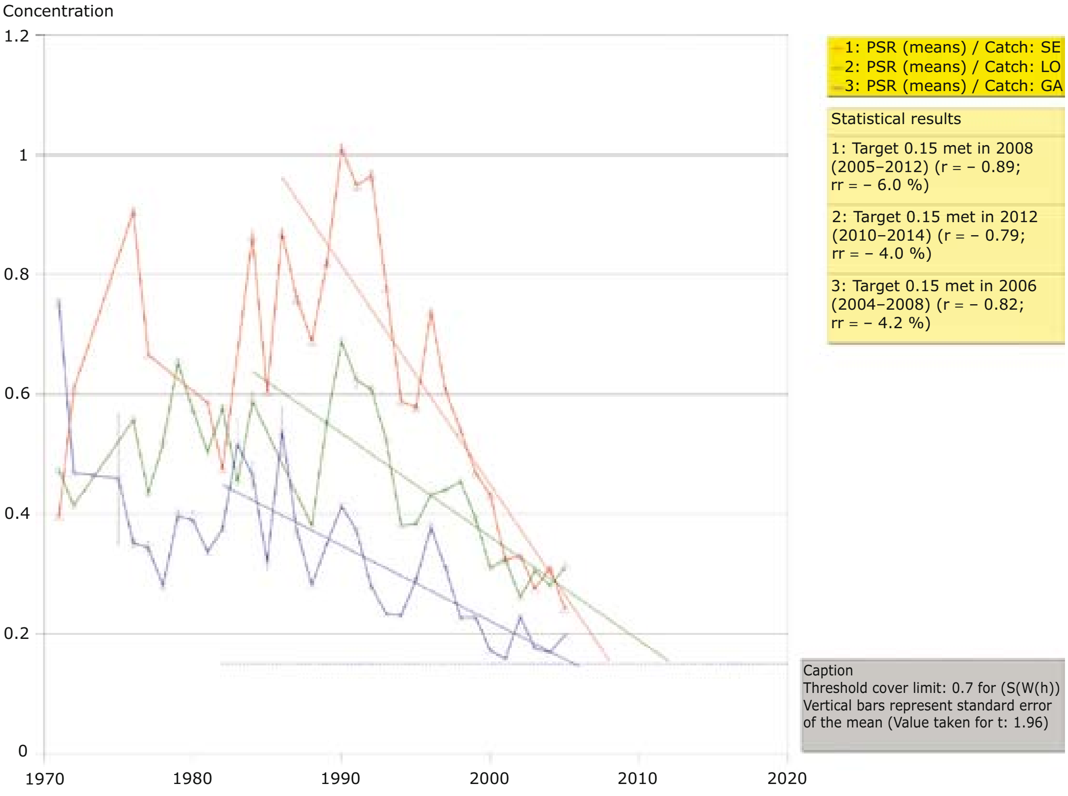


Figure 12 PSR development (maximum values) in the Seine, Loire and Garonne catchments, all strata combined

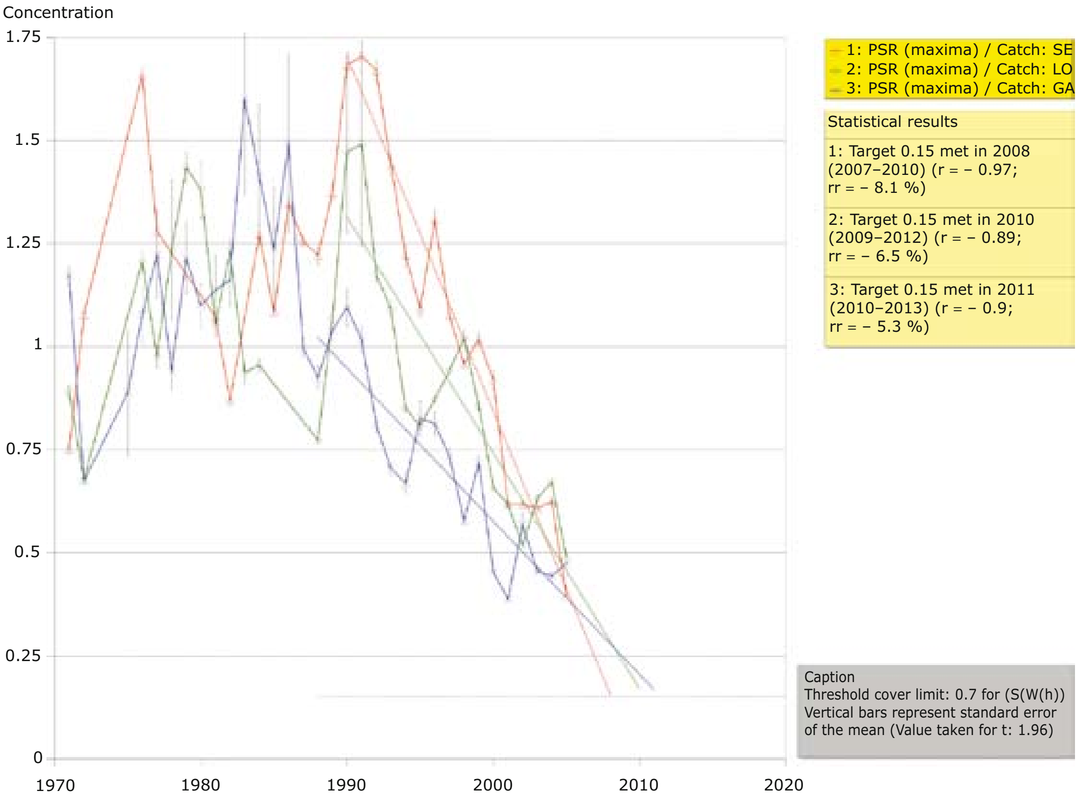
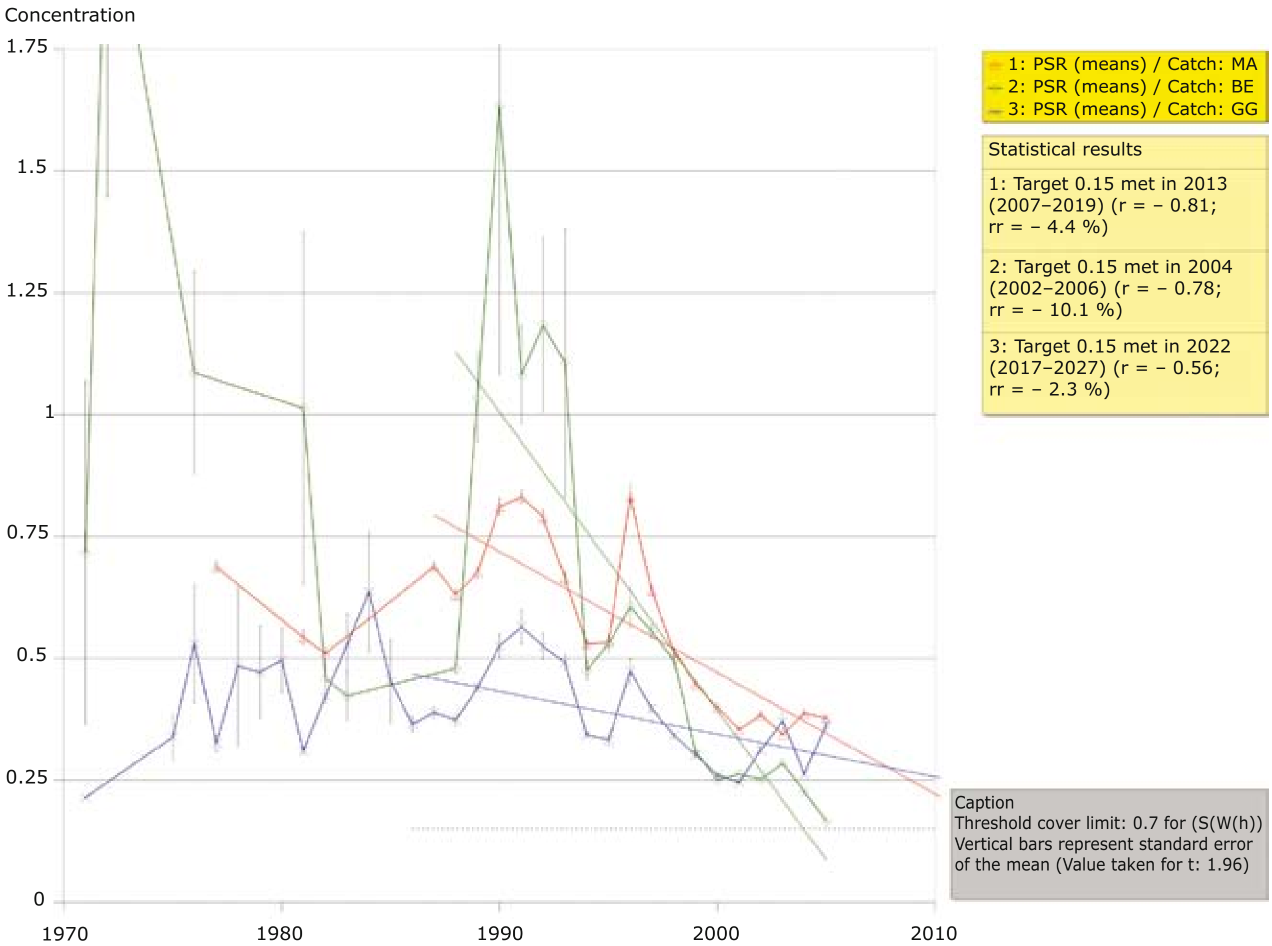


Figure 13 PSR development (mean values) in the small catchments of the Atlantic Ocean, all strata combined



maxima confirms the degradation. In both cases, the linear trend is not very good and largely dependent on the choice of start- and end-year. A partial and provisional conclusion is that after a short period (1970–1980) of sharp increase which may result from improvement in monitoring, phosphorus compound concentrations reached a plateau in the range 1.5 mg PO_4 L^{-1} for maxima and 0.5 mg P L^{-1} for averages. This is not contradictory with conclusions proposed by the *Rhône-Méditerranée-Corse* water agency that concludes on the decease in P concentration of almost all Mediterranean rivers since 1990. It has to be mentioned that this aggregate poses many specific problems of weighting, because the large urban areas are close to the sea.

There is no undisputable argument which explains why the general decrease in average minimum values is somehow parallel to the decrease recorded for ammonium. A tentative explanation would be better sewage collection, especially during a high-water event. Minimum values are often recorded during the maximum dilution of waters, since ammonium and P loads are primarily waste-water borne. Nor could a decrease in minima reflect a permanent improvement of conditions. Rather, this suggests a hydrological relationship because of the parallelism of higher and lower values with the other assessments or just a larger number of data capturing smaller minima.

4.2.3 Development of nitrate concentrations

Nation wide results (nitrate)

Unlike ammonium and phosphate compounds, nitrate values suggest a permanent increase in concentration, including minimum values. This is shown in Figure 15. The target used of 11 mg NO_3 L^{-1}, which reflects the satisfactory conditions of the downstream part of large rivers, has been placed in graphics for the reasons presented in 3.2.3, point 3.

A first examination of Figure 15 suggests a very negative situation because the average slope is proportional to the concentration itself: maxima

Figure 14 PSR development in Mediterranean catchments (Rhône and small coastal catchments)

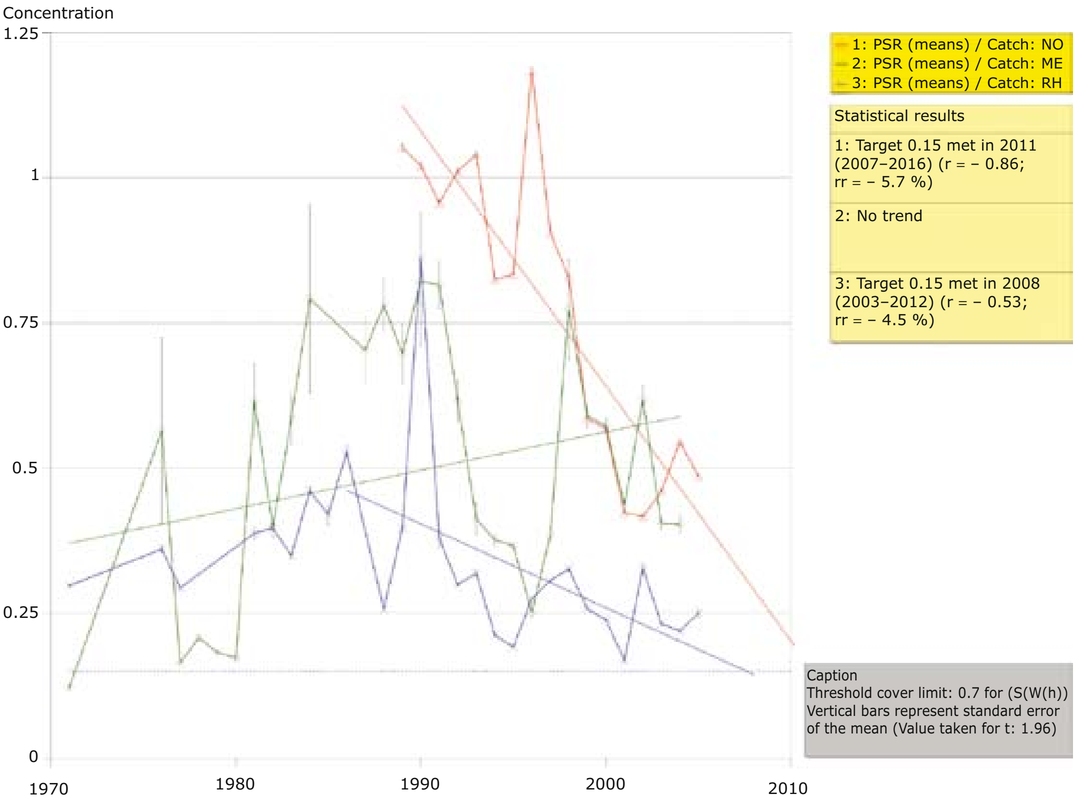
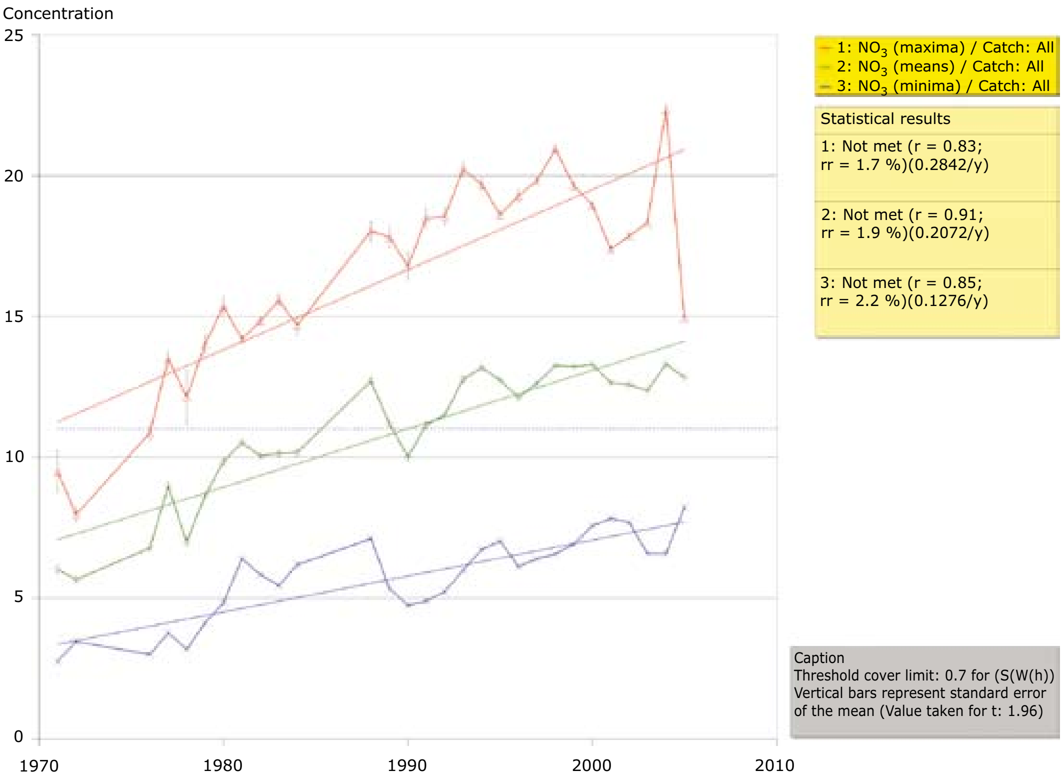


Figure 15 Nitrate trends, all strata and catchments together



grow faster than averages. However, the relative ratio of growth is in the reverse order. The trends are statistically acceptable, but more detailed assessment strongly suggests that maxima, averages and minima may be starting to show a structural change in concentration development, which is all the more marked as the initial concentration is bigger. Both diagrams support the necessary caution in presenting conclusions.

Since 1990, the maximum concentrations show values that tend to be constant (to be smoothed according to hydrology); the average concentrations show a change in slope (0.0997/year between 1988–2005 against 0.4086/year between 1971–1988). By contrast, there is no change in minimum trends.

Again, two different interpretations are possible with Figure 16 data and the different behaviour of maxima, average and minima curves: the positive interpretation is that measures taken in application of the Nitrate Directive started to be effective in cutting the largest inputs. The negative interpretation is that this merely reflects a saturation of emissions, all possible losses being achieved, the emitted nitrate being diluted by more or less the same quantity of water.

The overall analysis cannot support a more accurate interpretation. A third hypothesis would reflect not the structural change in causes but a blend of different trends across strata and catchments whose diversity of situations is not accurately depicted by an aggregated diagram. This is analysed in the next sections.

Catchment analysis (nitrate)

The development of nitrate in the main French rivers is displayed in the Figure 17, Figure 18 and Figure 19. All trends are increasing, with the exception of the Loire catchment which has seen stabilisation since 1987 (Note: this hypothesis is not presented in Figure 17). Yearly means are very accurately established as a result of a sufficient number of monitoring stations in the different strata present in the different catchments. It is also important to note that between the poorest (Garonne) and the richest in nitrate (Seine) there is a three-fold factor at the end of the period.

Figure 16 Re-assessment of nitrate trend, all strata, all catchments combined

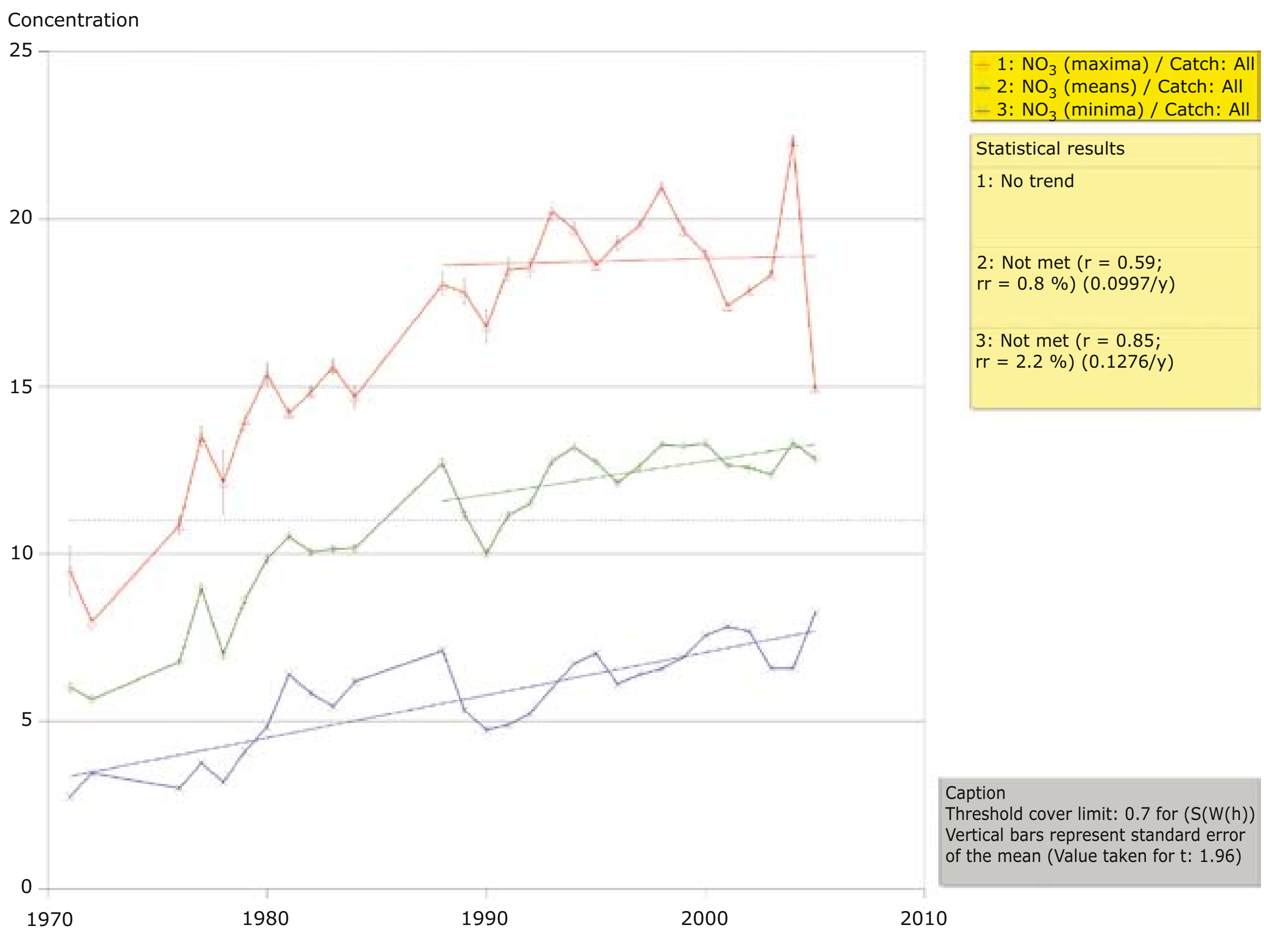


Figure 17 Nitrate development (mean values) in the Seine, Loire and Garonne catchments, all strata combined

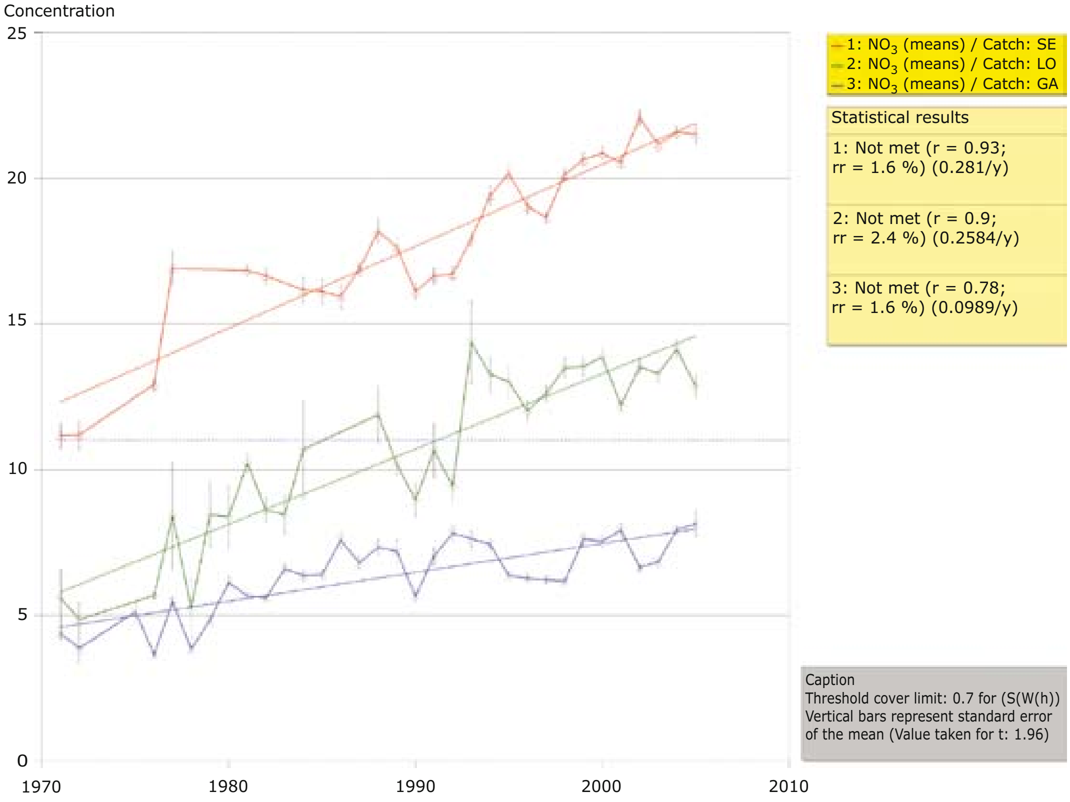


Figure 18 Nitrate development (mean values) in the small catchments of the Atlantic Ocean, all strata combined

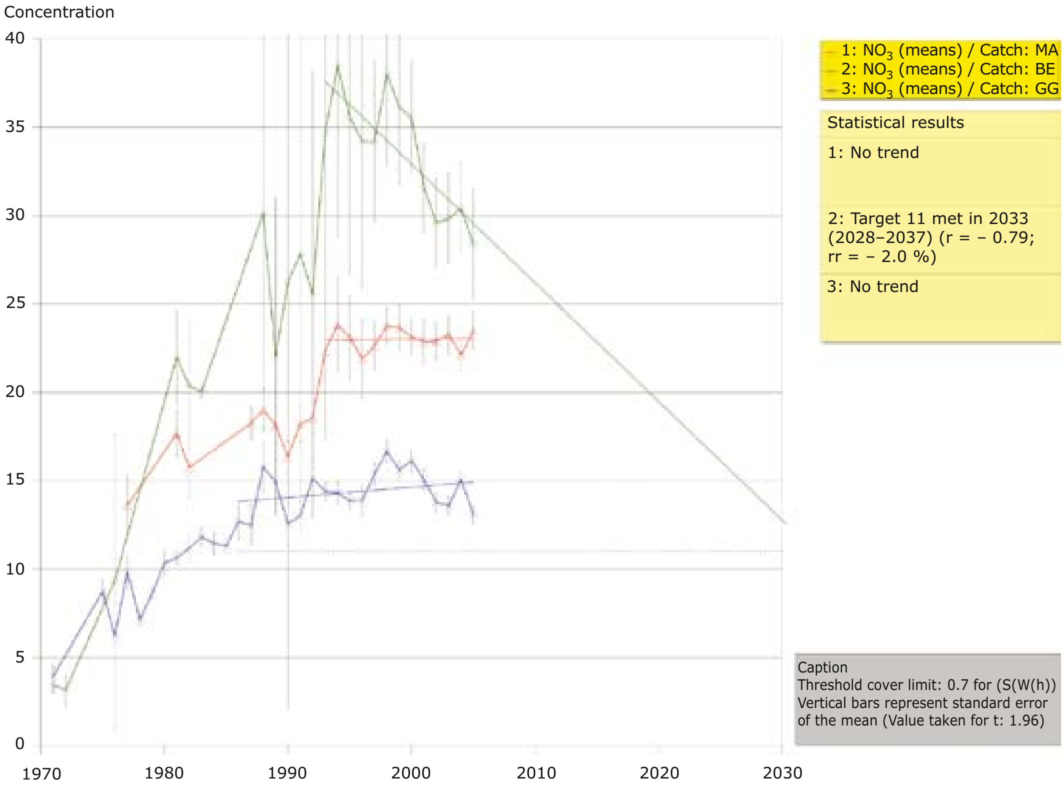
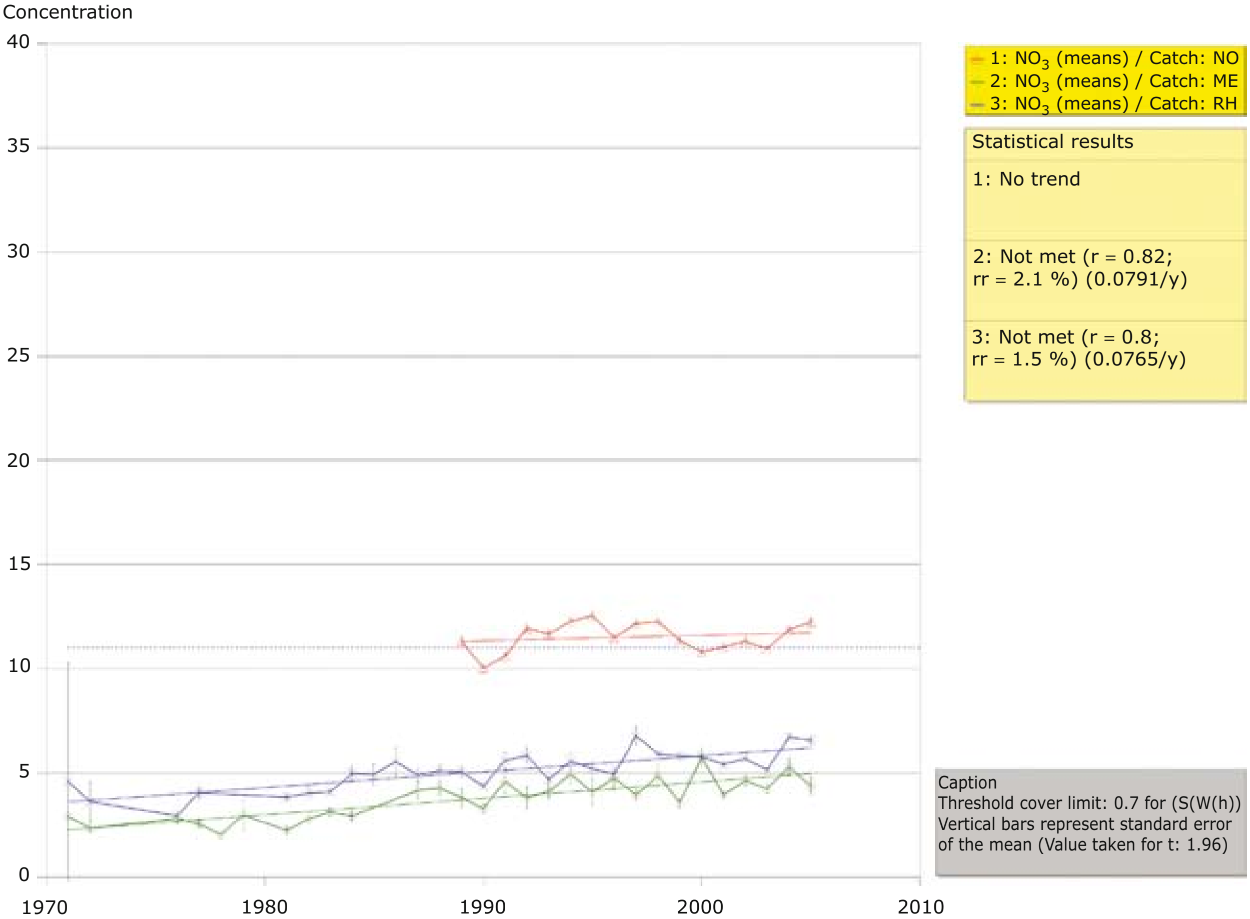


Figure 19: Nitrate development in Mediterranean catchments (Rhône and small coastal catchments)



The Nitrate Directive (91/676/EEC) was adopted in 1991. The subsequent years may be of importance in assessing some change in trends in strata and catchments where this Directive should have applied.

The change in values is well reflected in flux assessment that has been carried out independently on same data sets (BETURE-CEREC, 2000).

By contrast, the situation in the Channel and Atlantic catchments (including Brittany), which are substantially more contaminated than the major Atlantic rivers, exhibits a more diverse picture. In these catchments, and especially in Brittany, the relatively low density of monitoring stations in a complex system of catchments, and the high variability of observation values result in large uncertainty intervals around the yearly means. However, the trends are in line with the uncertainty domains.

The case of Brittany is very encouraging. After an increase in the range of + 1.37 mg NO_3 L^{-1} y^{-1} between 1971 and 1995, the concentrations began to decrease at a rate of 0.672 mg NO_3 L^{-1} y^{-1}. The stratified average reached an excessive value of ~ 35–40 mg NO_3 L^{-1}. Between 1990 and 2000, the average maxima (30 to 104 annual averages) reached close to 60 NO_3 L^{-1}, reflecting individual values in the range on 100.

Since the current rate of improvement is half the rate of degradation, the suggested target is not likely to be met before 2033 (2028–2037). This makes it unlikely that the 'green tides' that affect the Brittany shore (IFREMER, 1993; Merceron, 1999; Ifremer DEL/EC, 2001) could be tackled in due time with respect to the requirements of the Water Framework Directive. The three remaining catchments show low nitrate concentrations, with slight increases in the Mediterranean ones. Their status is however very good. The average concentration is half that of the target: nitrate is clearly not an issue at the basin level in the rivers flowing to the Mediterranean.

4.3 Analysis of potential pressures impact at stratum level

The analysis presented in the previous sections portrays the status of water resource ([10]) but gives little information on the possible efficiency of policies and measures taken to combat pollution.

In this section the impacts of the two great sectors considered — agriculture (under crop/livestock sides) and urban (and human) dwelling — are analysed. There is no obvious best way of presenting the results. The choice was made to present them by determinant in order to ease the comparison of the outputs. The trends for each of the defined strata are reported for all catchments, considering separately all strata influenced by agriculture and urban activities. Mixed (M) and non-impacted (F) strata are reported in both cases because they either participate in both or constitute a reference. Stratum X catchment analyses have been carried out to support findings but are only presented in exceptional situations.

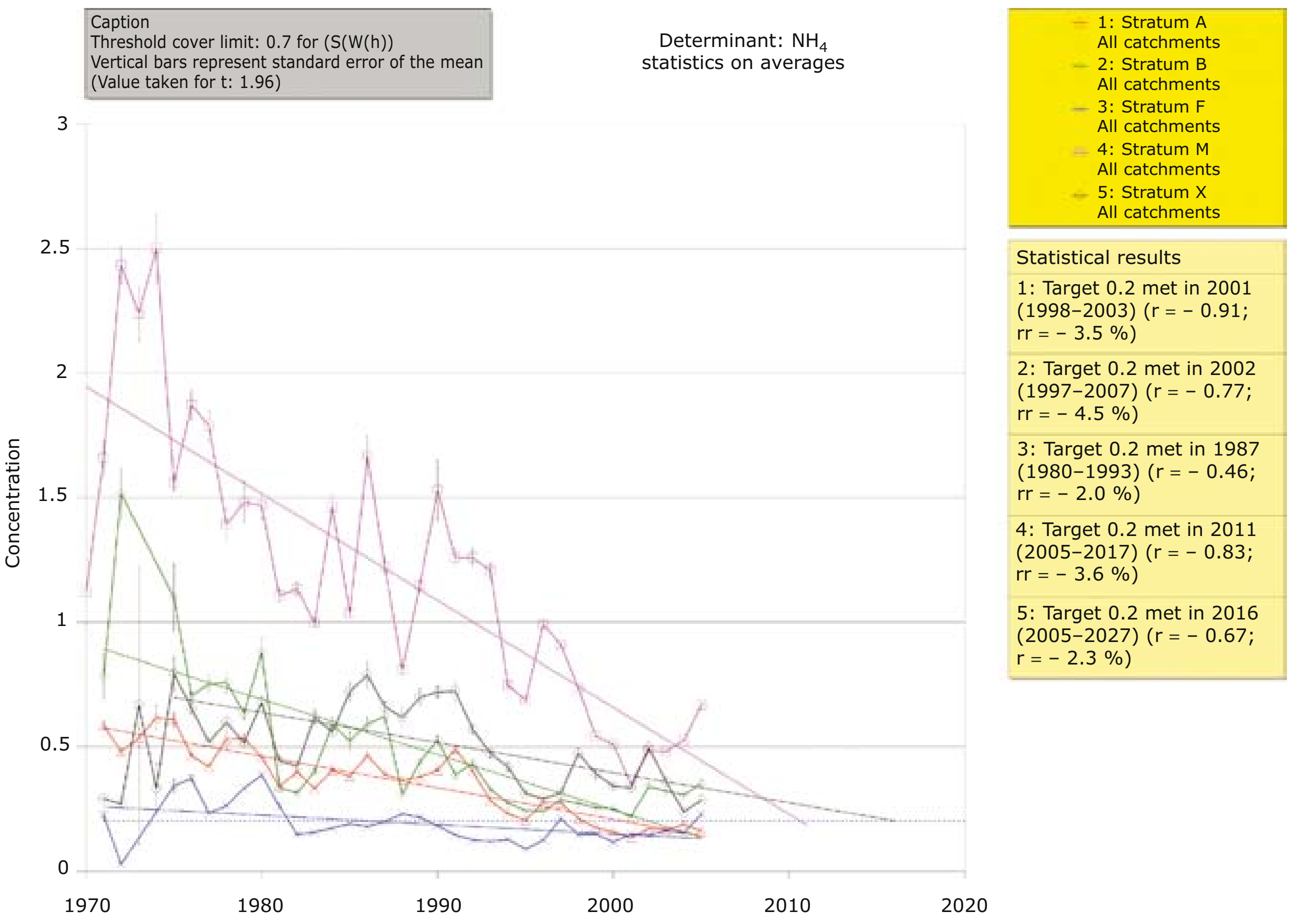
4.3.1 Ammonium

Trends per stratum

The most polluted of all the strata where agriculture is an important driver is the 'Mixed' stratum. Here, the urban activity is identical (as stratum definition) to the 'Urban' stratum. There are two possible reasons for this. The Mixed stratum is defined by dense population and intense crop agriculture. However, the biggest cities are often close to fertile cropland because this permitted their historical growth. For example, the Paris area is in the M stratum.

The low-pressure stratum is close or under the concentration target that was largely reached in

Figure 20 Ammonium development (mean values) for all strata influenced by agriculture, all catchments combined



([10]) It is precisely the usable resource that is analysed because the statistical indicators are annual averages weighted by time, in order to yield unbiased averages.

1986 (1983–1990). Three strata are under dominant agricultural activities: 'A' where vegetal crops are the main activity; 'B' where livestock is dominant; and 'X' where no activity dominates but where a mixture of population, crops and livestock may be present at low levels. Quite unexpectedly, this latter stratum is the most polluted of the three and its trend is the least favourable; no target reaching can be expected before 2032 (with wide uncertainty: 1998–2067)!

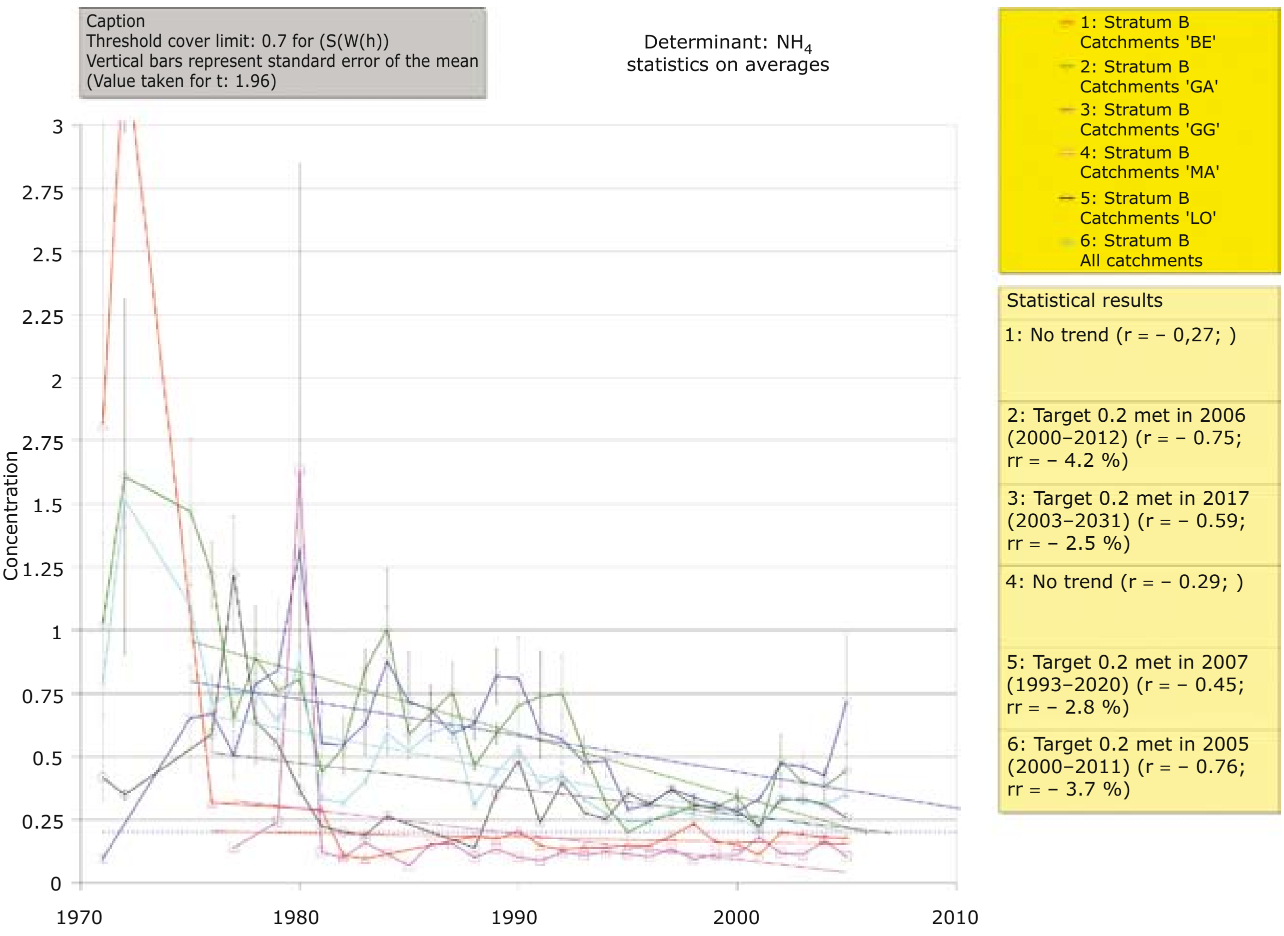
As expected, the B stratum presents greater concentrations than 'A' , because livestock emissions are more likely to discharge ammonium than field fertiliser leaching. However, its overall trend is very positive.

Since stratum B (livestock) supposedly contributes to ammonium emissions, this should be all the more true in places where livestock density (and cowshed breeding) dominate, e.g. in Brittany. The reality seems far more complex, as suggested by Figure 21 where the basins with B stratum are displayed.

The monitoring stations in B stratum in Brittany have the lowest ammonium concentration of all basins that have stratum present. Considering the average maxima instead of mean does not change the pattern. In view of the fact that dense livestock is certainly a source of ammonium, emissions into rivers have not been realised as may be expected. This positive finding and the larger concentration in other basins may be the result of:

- emissions from livestock not taking place where livestock is bred. This makes the relationship between catchment and sampling point inaccurate. This is possible because the majority of livestock in Brittany is bred in cowsheds and manure is stored before being despatched to areas far away from its origin;
- effective measures used to store, handle and spread, preventing direct disposal to waters. This hypothesis would be consistent with the presence of nitrate in B stratum in Brittany since it results from oxidation in soils;

Figure 21 Average ammonium in stratum B of selected basins



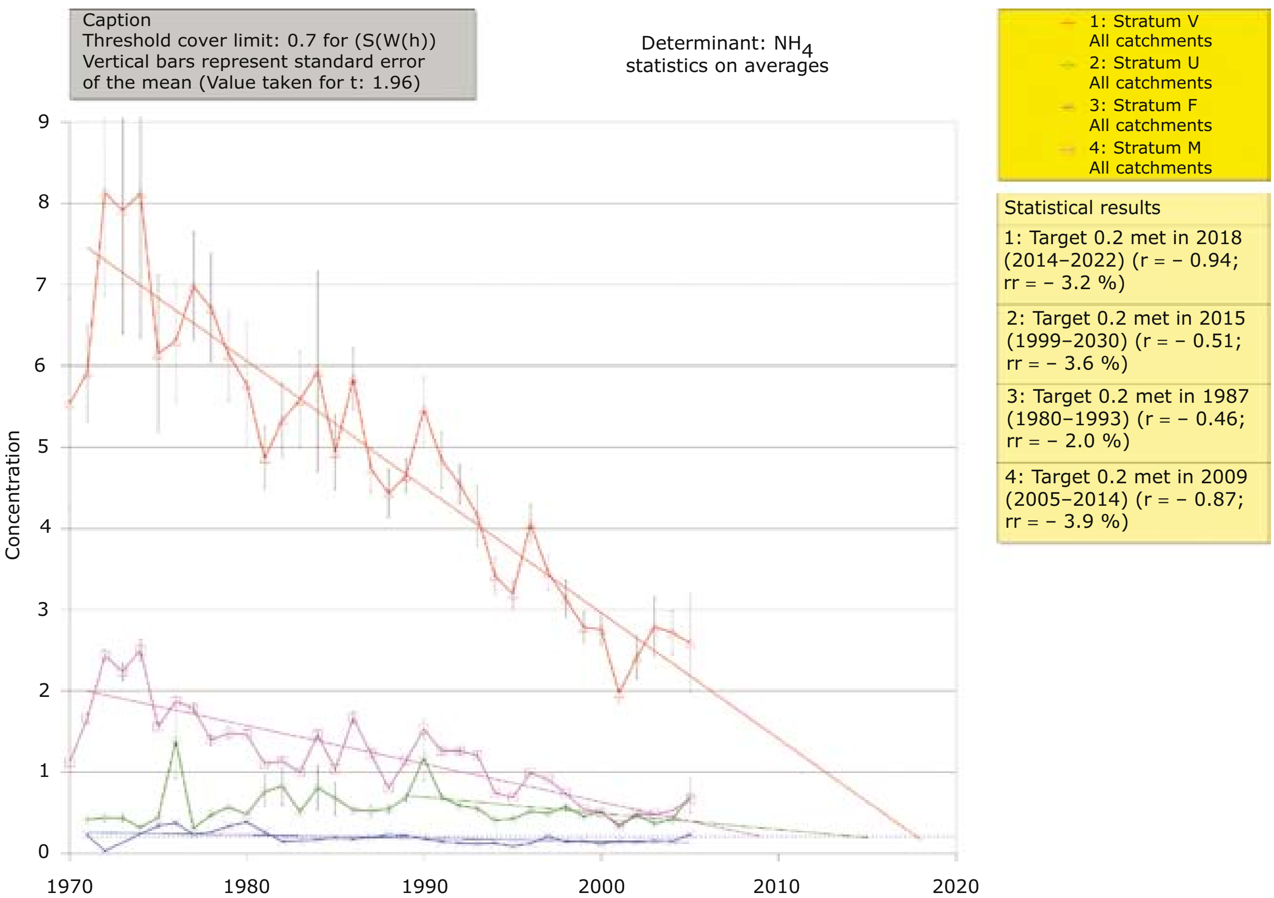
- other sources of ammonium in basins where livestock is dense, namely population (note: this is beyond the scope of the report and requires statistical analysis of stratum components vs observations instead of stratified analysis).

Strata predominantly impacted upon by urban activities display relatively elevated concentrations, from the 'V' stratum, which is still ~ 5-fold higher than the 'M' stratum. Moreover, they have the fastest rate of improvement, with the target possibly being met in 2019 (2015–2023). The 'Urban', which could only be assessed from 1989 onwards, is comparable to the Mixed stratum. However, it is logically lower because of the limited contribution of agricultural activities. The absolute difference between M and U stratum is in the range of the contribution from the A stratum.

All trend analyses carried out at catchment level and by stratum confirm that ammonium is firstly an urban-borne pollutant with a secondary source in areas of densely bred livestock. Accordingly, priority purification policies have been applied. These policies are understood to have been applied earlier and with more strength to the most impacting sources.

Hence, the intermediate stratum 'X' shows signs of differential improvement. Paradoxically, it might become the second-most polluted stratum in the years to come after the 'very urban'. However, the ecological impact of the 'very urban' stratum is likely to be less important because a much smaller area of catchments are impacted. Moreover, many of the rivers in these catchments are expected to be heavily modified and unlikely to be given a high quality standard. Consequently, the intermediate stratum might become a limiting factor in meeting the ecological targets of the Water Framework Directive. This hypothesis requires further evidence before it can be presented as a finding. That will be addressed by the water accounts methodology (WIRQ).

Figure 22 Ammonium development (mean values) for all strata influenced by urban activities, all catchments combined



Source apportionment for ammonium
Source apportionment aims to weigh the relative contribution of the different sectors to the overall concentration. Its aim is different from that of sectoral analysis. A very polluted stratum, e.g. the 'V' stratum, experiences water quality problems because concentration in its rivers is high. By contrast, if this stratum has a limited extent, its global contribution may be slight. Figure 23 shows the relative contribution from the different sectors, computed according to the rules presented in the previous section (3.2.8). As expected, ammonium concentration is widely dominated by urban and by animal excreta. The baseline ammonium (approximately 35–40 % of the concentration in the 2000s) now results from non-polluted areas. Intense agriculture holds an even lower share than livestock. Despite strong reduction in the overall mean concentration in the urban source, the relative share of ammonium from urban sources tends to be increasing.

4.3.2 *Phosphorus*

Trends per stratum
Soluble phosphorus concentrations show marked differences between the agriculture and urban dominated strata. As expected, the agricultural activities are secondary in their supply of high phosphorus concentration in rivers. The question that cannot be answered accurately is the specific contribution of agricultural activities because the rural areas, especially in France, are also partly urbanised. Source apportionment gives new insights into this point.

Phosphorus concentrations decline in all strata, at rates more or less proportional to the initial concentration. Hence, the most polluted strata: 'very urban', 'mixed' and 'urban', as expected, decline by 5.6, 2.7 and 5.1 % per year, respectively. Despite M and U having parallel slopes, the 'Urban' stratum average concentration is markedly below the 'Mixed' stratum values. This may be due to the reasons suggested at the beginning of the section related to ammonium, namely: the biggest cities are in the M stratum. However, agricultural sources cannot be freed of responsibility. In all cases strata with agricultural dominance are significantly below the urban ones. They all improve at the same rate 2.1 to 2.9 %, as expected. Stratum B (livestock) concentrations are significantly higher than those of the 'A' stratum. This is not contradictory to the hypothesis that the densest breeding areas may manage the manure produced. Scattered manure can leach soluble phosphate. In contrast to the remark about the 'X' stratum for ammonium, phosphate values in the X stratum are the lowest of the agricultural activities group.

Figure 23 Source apportionment for ammonium

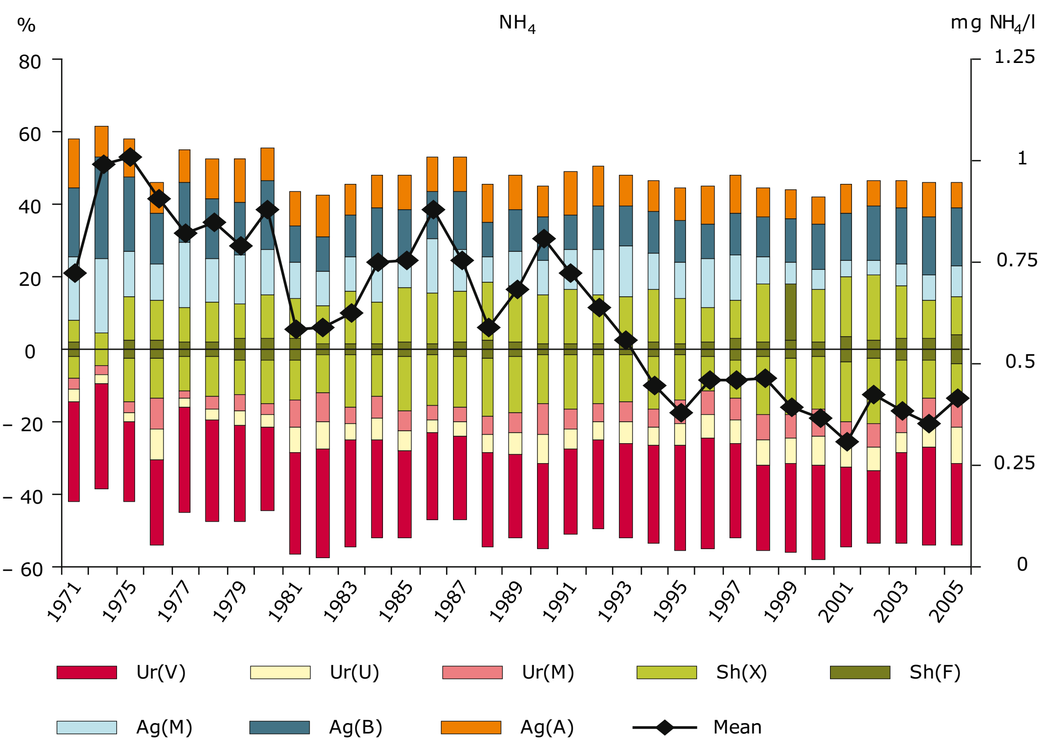
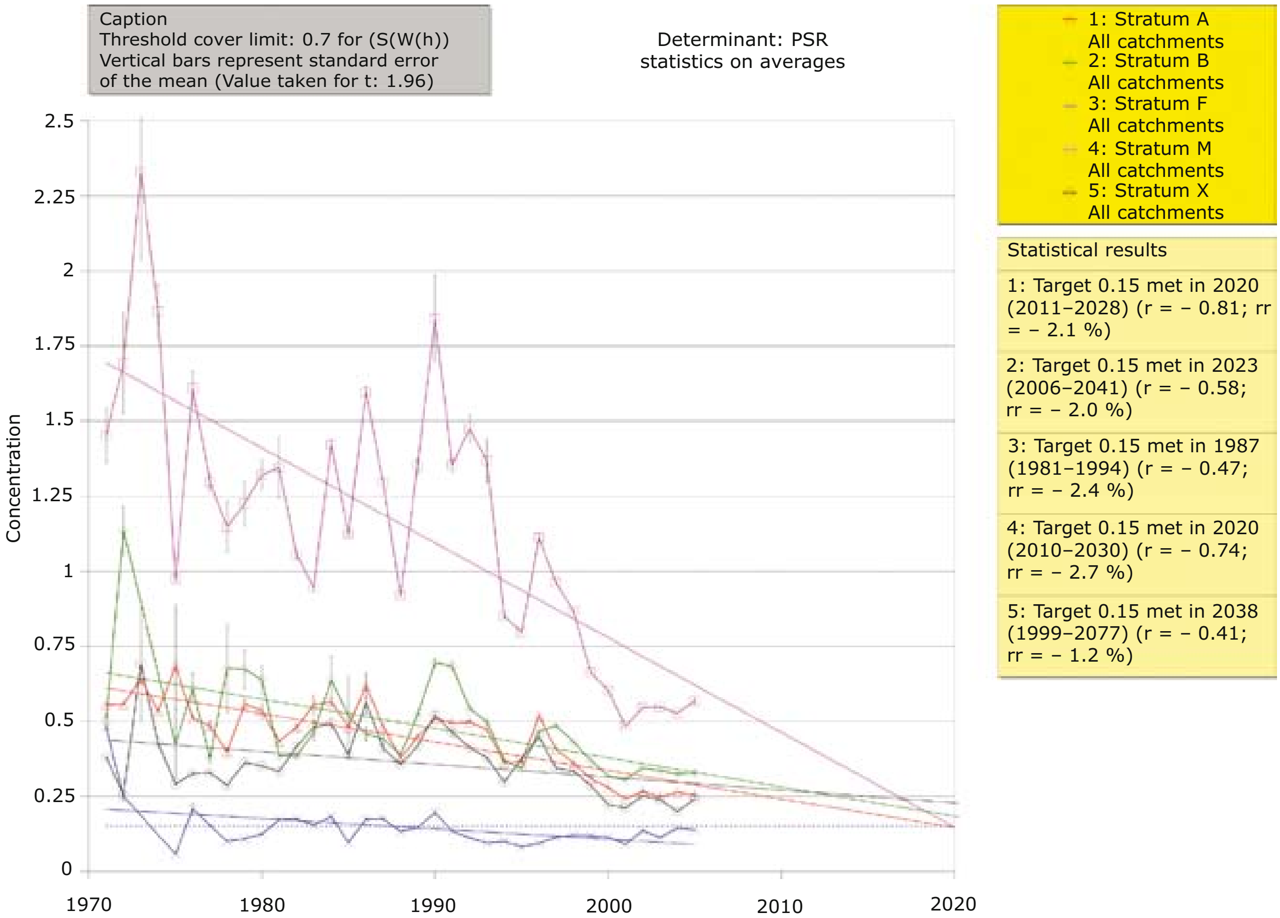


Figure 24 Soluble phosphate development (mean values) for all strata influenced by agricultural activities, all catchments combined



Where there is no or very limited activities (stratum 'F'), the average soluble phosphorus is nevertheless greater than expected for pristine areas.

The rather complex correlation between concentration and source activity is illustrated by Figure 25 where M and B strata from Brittany are displayed along with the M stratum from the Seine basin and the U stratum from the Rhône basin. Both of the latter strata host significant populations. However, in the Rhône, the agricultural activities are not important with respect to nitrate contamination.

The B stratum in Brittany presents 50 % smaller concentrations than those from the M stratum. This is close to the values observed in the Seine basin. The reason is that the population density and the relative discharge (L s^{-1} km^{-2}) are in a comparable range. By contrast, the Rhône River and its alpine tributaries, on which lie the most populous areas of the basin, are significantly more productive in water, yielding lower concentrations. A second conclusion is that the contribution of livestock to phosphate in rivers in much less likely than those from human dwelling. By contrast, for reasons that are likely to be related to the type of breeding in this area and the limited availability of locations where this activity is carried out, the B stratum of the Rhône basin presents the biggest phosphate concentrations.

Concentrations reported for the urban dominated strata are far above those reported for agriculture. They present the sharpest decline, as shown in Figure 26.

In all circumstances, especially where the UWWD has been applied, the decrease in concentrations is remarkable. Moreover, it could result in the achievement of a reasonable target for rivers in the foreseeable future. As a result of the evidence of the links between phosphate concentrations and drivers, a significant improvement of the river status can be ascribed to sewage purification policy in basins where point sources are clearly dominant.

Figure 25 Compared phosphate concentration development in different catchments

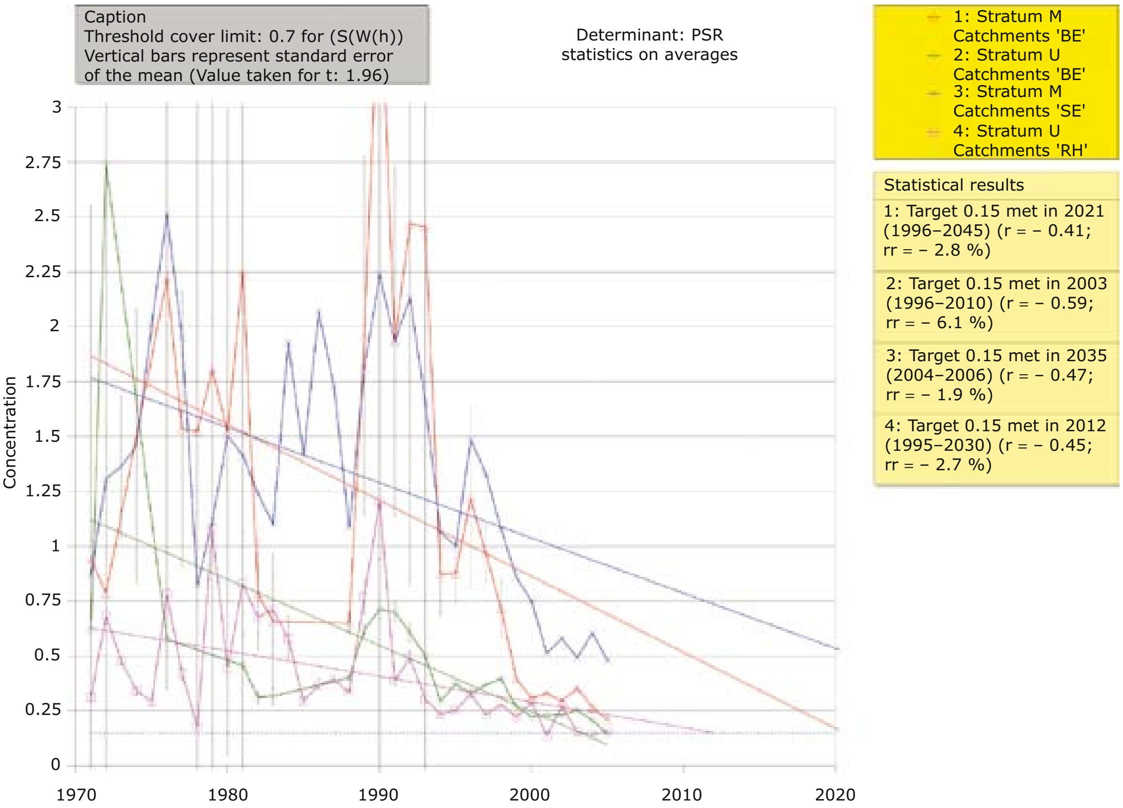
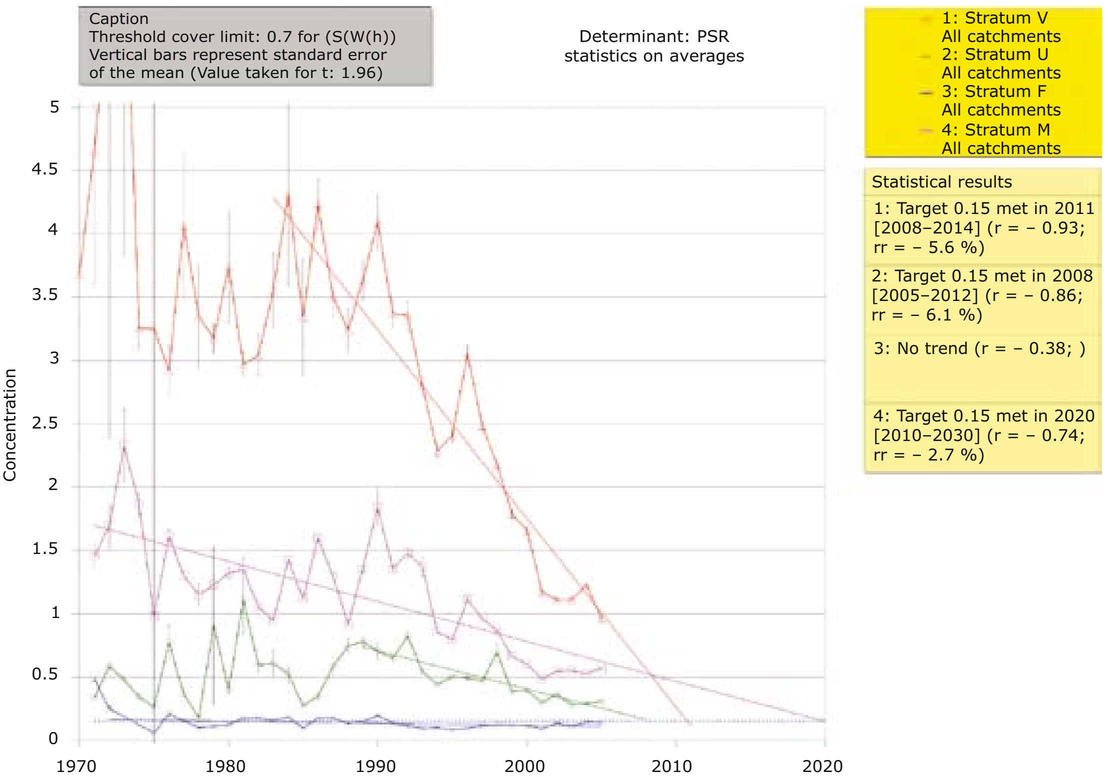


Figure 26 Soluble phosphate development (mean values) for all strata influenced by urban activities, all catchments combined



Source apportionment for phosphorus
Source apportionment for phosphorus has been carried out by many means. The polyphosphate producers, despite a sharp decrease in P content for detergents ([11]), keep analysing the effects of slurry products on the environment. The reviews of domestic and non-domestic shares of phosphate loads provide a very wide range of figures. The EEA suggests ranges between less than 20 % to over 75 % of agricultural contribution, even for large fluxes (NERI, 2005, p 36). The case of large fluxes and high shares is nuanced by this comment:

'The Ems and Weser rivers have, on the contrary, very high agricultural shares, which are due to the agricultural exploitation of bog soils in downstream parts of these rivers. The bog soils have poor phosphorus-binding capacities and the surplus of phosphorus is lost to the aquatic environment relatively fast, whereas in many other soils there is still a high capacity for immobilising phosphorus more or less permanently.'

This study does not cover France. The CEFIC has issued a compilation for France (CEEP, 2000) that gives 29 % to 82 % of non-domestic sources (that are not all agricultural).

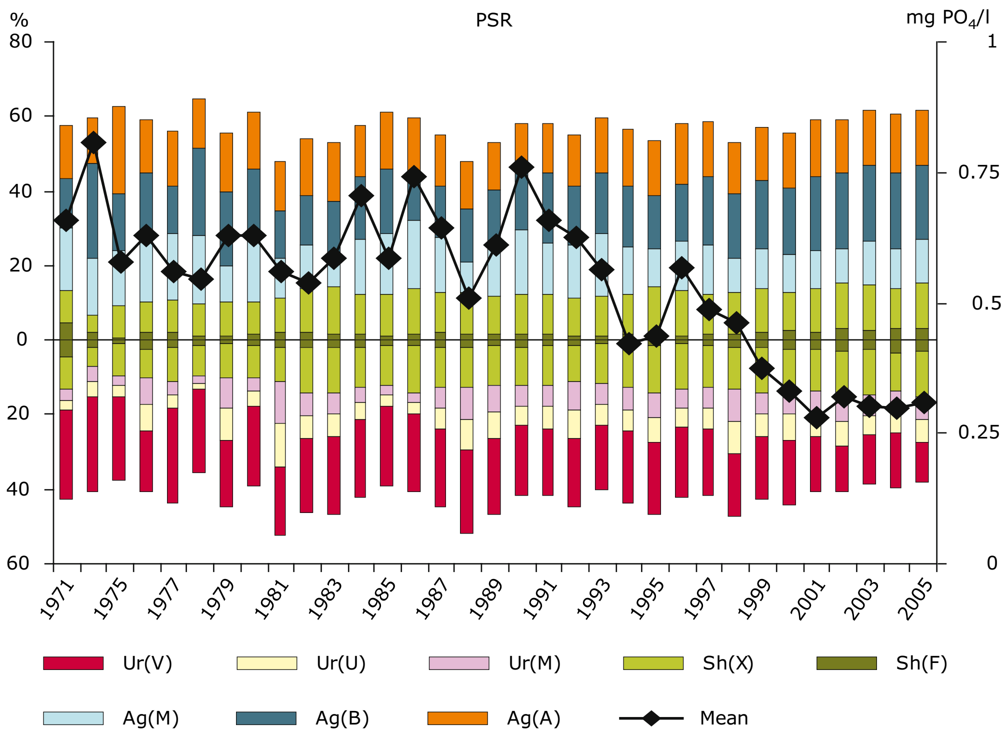
Figure 27 shows that the relative share of urban sources has been decreasing since the end of the 1980s. This decrease is 80 % compensated by the increase in both livestock and ordinary pressures strata. Indeed, both strata have shown, less markedly, a drop in phosphate concentrations. The crop agriculture share remained constant at close to 15 %. Despite a limited area share of the territory, 'very urban' is still a significant contributor.

4.3.3 Nitrate

Trends per stratum
Nitrate is generally understood as coming primarily from agricultural sources. A long-term increase in trends is ascertained for strata A, B and M., whereas B and M display change in slope between 1990 and 1994. This change has been taken into account in the Figure 28. The crop stratum A does not show significant change in increasing trends. Nevertheless, there could be a possible slow-down after 1995. The X stratum, with limited pressures, also shows an increase in concentrations. The minimum pressure stratum has no significant trend and presents concentrations in the range of what can be expected from preserved (though not truly pristine) catchments.

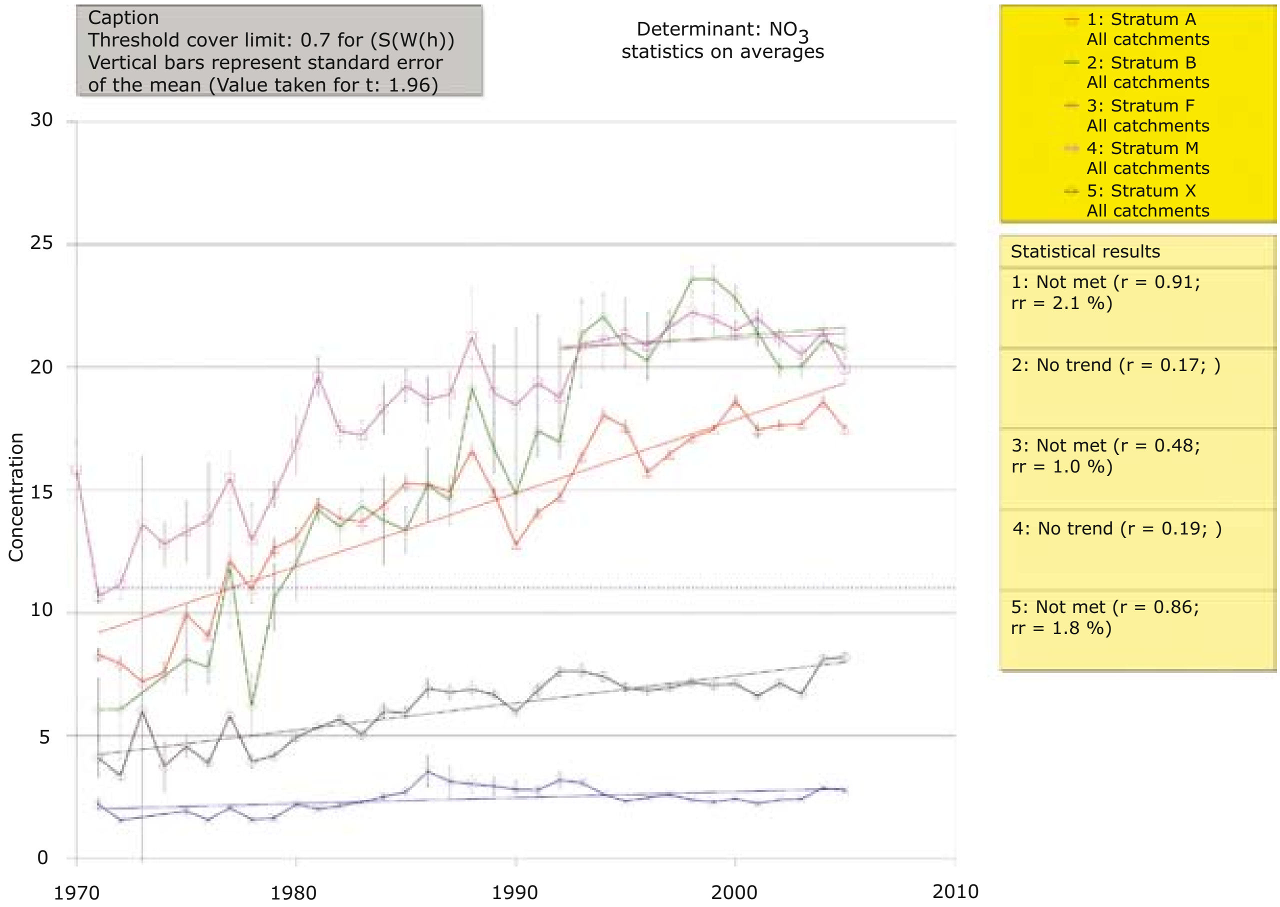
Strata submitted to urban pressures have two very different patterns of concentrations. The 'very urban' stratum presents elevated and slightly declining

Figure 27 Source apportionment for soluble phosphate



([11]) The enforced bans do not consider the dish washing products.

Figure 28 Nitrate development (mean values) for all strata influenced by agricultural activities, all catchments combined



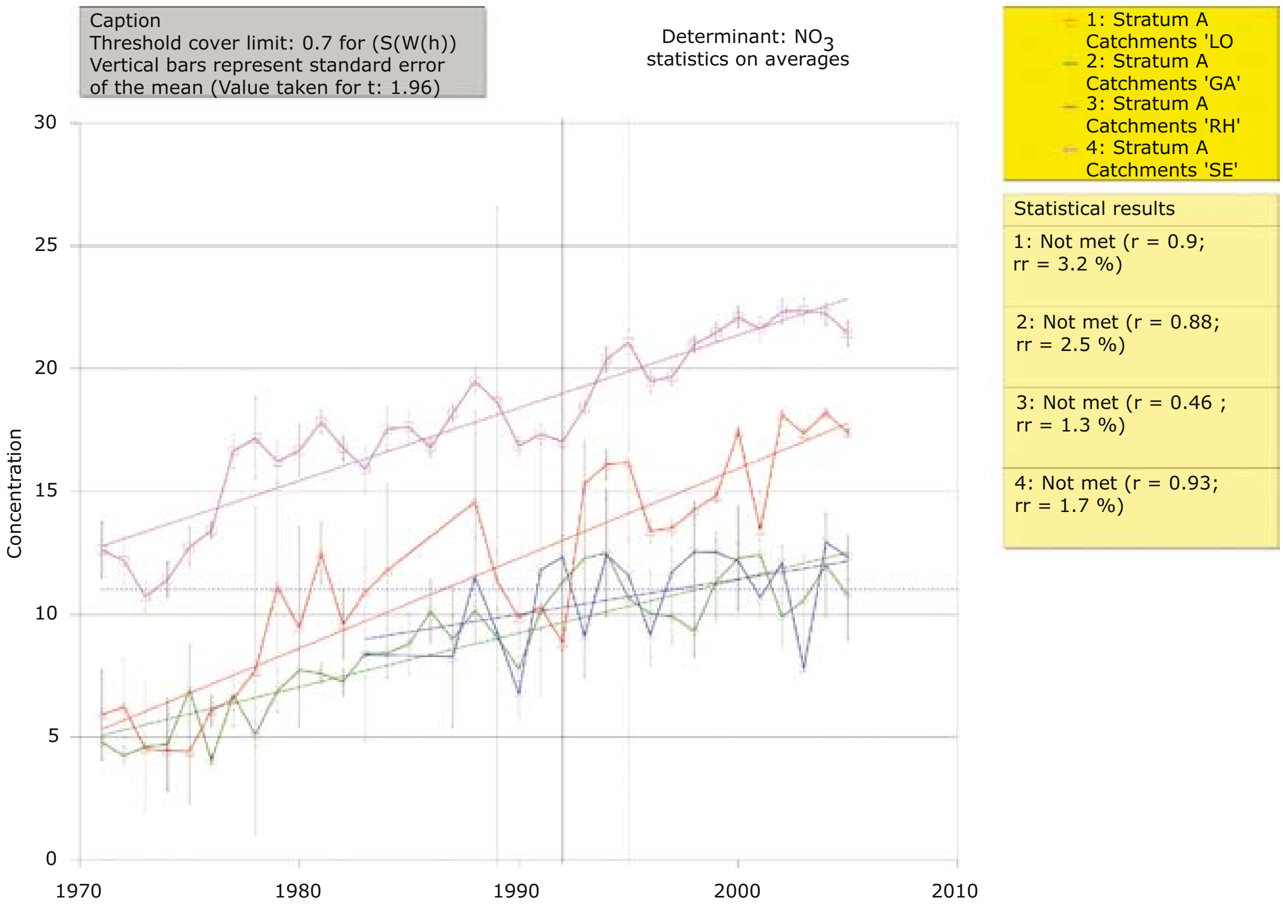
concentrations. The observed values are even greater than those reported for the 'A' or 'M' strata, with the main difference that the elevated concentrations had been observed since the beginning of the monitoring period.

There is therefore a strong suspicion that these values are related to the oxidisation in rivers of the organic nitrogen concentrations (in parallel with ammonium) and that the nitrate originates from urban sewage in this stratum. The declining concentrations could be the outcome of the progressive change in purification plants processes for the largest cities: replacing the purification plants working with short residence time by oxidising/ denitrification processes.

The constancy of the nitrate concentration in strata where urban emissions are quite exclusive is fully consistent with the general constancy of population and direct emissions to receiving water. By contrast, the M stratum is understood to display a combined pattern. This results from a constant urban source being mixed with increasing inputs from agricultural emissions.

The correlation between crop agriculture and nitrate development is well demonstrated in Figure 29, where it has been selected for the A stratum and the four largest rivers. The nitrate concentration in stratum A is close to or higher than the corresponding value for all strata (Figure 17).

Figure 29 Nitrate development in stratum A in the four large French catchments (Seine, Loire, Garonne and Rhône)



Source apportionment for nitrate

Nitrate source apportionment displays a very stable pattern. The agricultural share has slightly increased, primarily because an increased contribution of livestock dominated catchments. The slight drop in urban share is directly compensated by agricultural catchments, in a rising concentration sequence. This suggests that agriculture is more and more productive in leaching nitrate to the environment.

Considering that surplus production tends to lessen, this finding is quite paradoxical. Based on OECD official data, the N surplus in France has dropped from 36 kg N/ha/year in 1990 to 25 kg N/ha/year in 2000 (EEA, 2005, p 353) ([12]).

([12]) Data can be downloaded from http://dataservice.eea.europa.eu/atlas/viewdata/viewpub.asp?id=1284.

Figure 30 Nitrate development (mean values) for all strata influenced by urban activities, all catchments combined

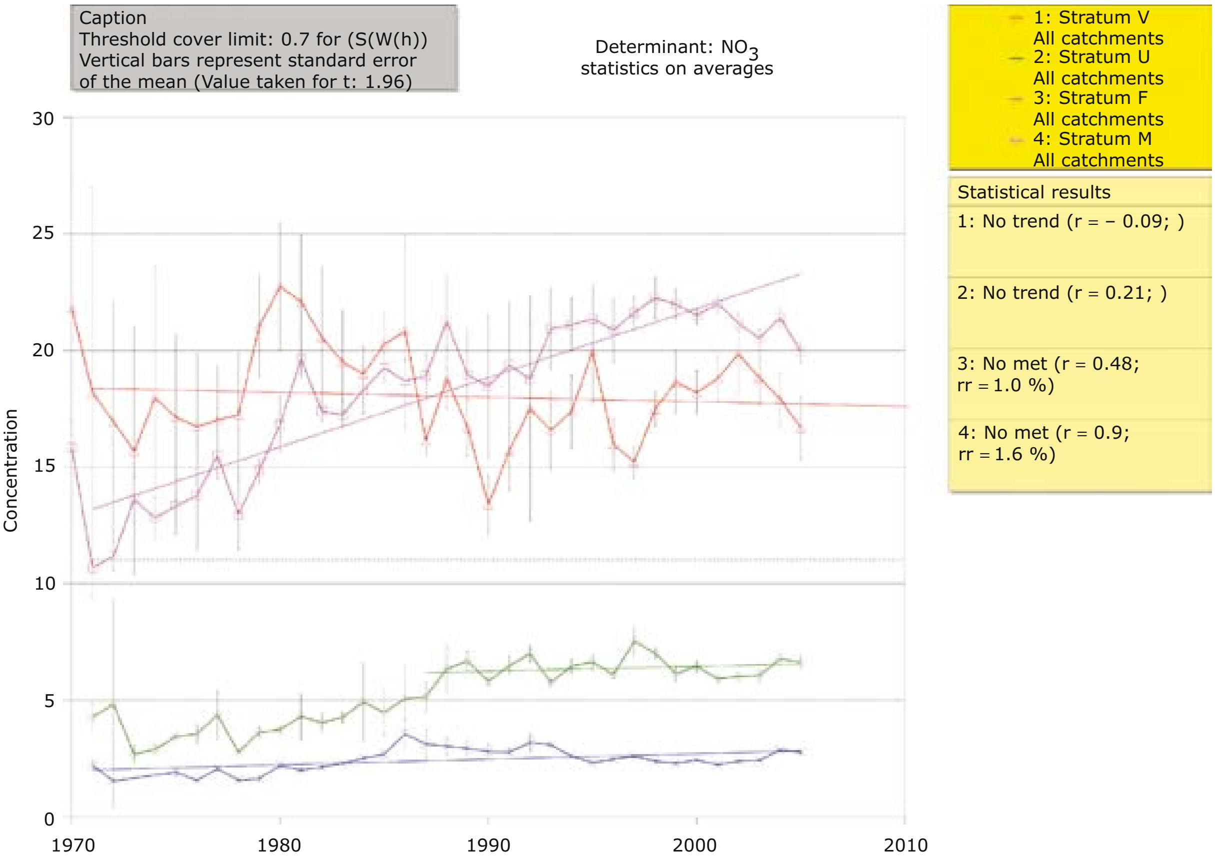
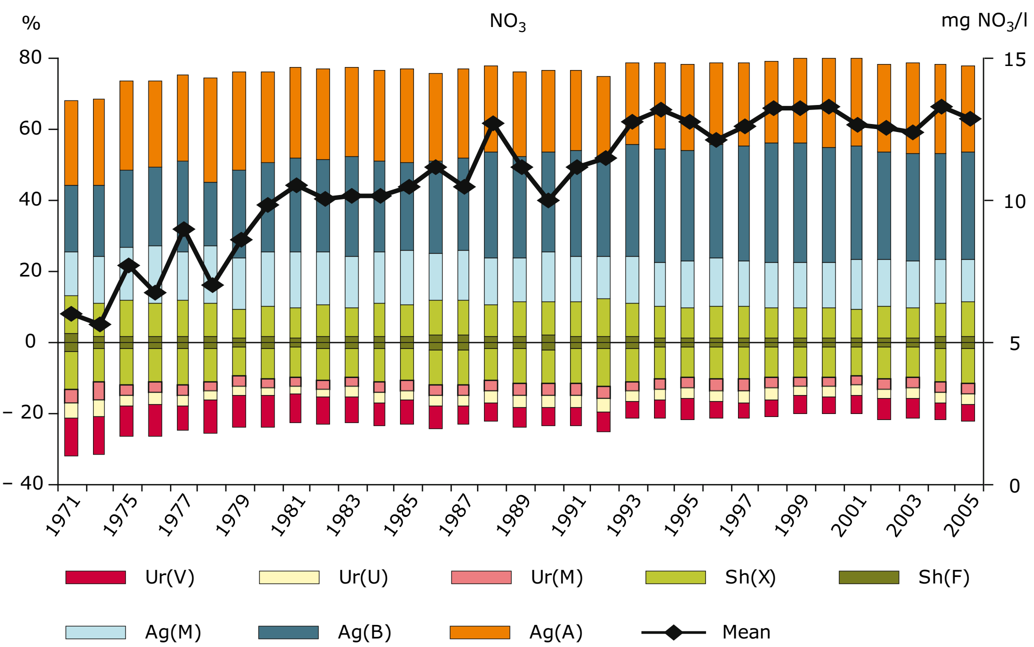


Figure 31 Source apportionment for nitrate



5 Advantages and disadvantages compared with Eionet-water

5.1 Assessment of differences between the two methods

5.1.1 Requirements

The policy effectiveness assessment for which this assessment methodology is primarily built requires two distinct sets of data:

1. Which pollution processing facilities have been implemented to meet policy requirements, when, where, and how they function.

2. Sound and significant assessment of the changes induced by the implementation of policy, preferably aggregated at the same level as the data quoted in bullet 1.

Stratified assessment aims to respond to point 2 requirements that presents both status ('S' of DPSIR) and trends (analogous to 'I') related to sources (proxies of 'P') whose changes relate to 'R'. When dealing with water quality, only major determinants can be assessed at aggregated levels because their presence is quite certain. Conversely, the certainty of the presence of analysed determinants is that no single source can be identified: all of the three determinants considered in this work have mixed origins. They have in all cases a substantial natural source that is blurred by household, industrial and agricultural activities.

Stratified analysis tackles a large share of the difficulties and makes it possible to assess, during the same interpretation exercise, both the development of status and the differential impact of the sectors represented by simply defined drivers. This is possible if the assessment can be made with i) good accuracy; and ii) a long time lag.

Table 3 Comparison between the modified methodology and the classical Eionet approach

Item	Modified methodology	Classical Eionet
Stratification	The stratification applies to catchments that are used to weight the aggregated statistics. Defined from causal analysis and reproducible procedure using GIS, land cover data, population and livestock distribution. Additional variables (e.g. effective rainfall) can be used at catchment level. Any aggregates can be computed from the stratification variables.	The stratification applies to stations that are considered as a representative sample of water bodies. Defined by the data provider as attribute of monitoring stations. Additional variables (e.g. percentage of land use on upstream part) are provided and can be used at the station level. Stratification structure does not allow final weighting of aggregates.
Monitoring network used	Stratification and monitoring network are independent; any set of stations is suitable. Preference is for full set of stations. Different number of stations reflected in certainty of assessment, with no impact on accuracy or bias.	A selected set of stations is done by applying density threshold on the existing networks. Since stations make the population, there is risk of bias if stations change in number or locations.
Calculation methods	Basic statistics are carried out with identical procedure to eliminate sampling frequency bias, and can be done at the most suitable time window. High flexibility on acceptance/rejecting doubtful values. Important focus on uncertainty and error estimates to assess trends and target met dates. Where disaggregated data are available, statistics can be improved by combining paired observations and versatile methods can be used.	Annual averages are provided by data provider. Averaging method is not provided. Limited possibilities of quality control. Distribution determinants are required (annual variance, quantiles) but poorly used/documented. There is no flexibility on statistics and paired observations processing is impossible. The limited focus on uncertainty is not intrinsic to the methodology, but to its application.

The major advantage of stratified statistics is the representativeness of the outputs, which is especially important when dealing with water status. A comparable assessment of the mean average or mean maximum of observations can be provided and trend perspectives can be derived with (at least partly) known uncertainty.

Finally, another advantage of stratified statistics is their ability to provide a much aggregated message at national level or on a large basin scale. This is clearly within the EEA mandate.

5.1.2 *Terms of comparison*

The differences between the applied method and the usual Eionet approach lie with the following three items summarised in Table 3.

Hence, the main differences lie with the systematic and reproducible method of stratification and in the focus on uncertainty assessment as well as the use in the analytical process. In addition, the use of all readily available data significantly improves the precision of results, and even their accuracy. The comparative assessment consisted of comparing key results obtained from the total data set and a French data set limited to the use of Eionet (the list of stations in WaterBase V4).

The following terms have been compared:

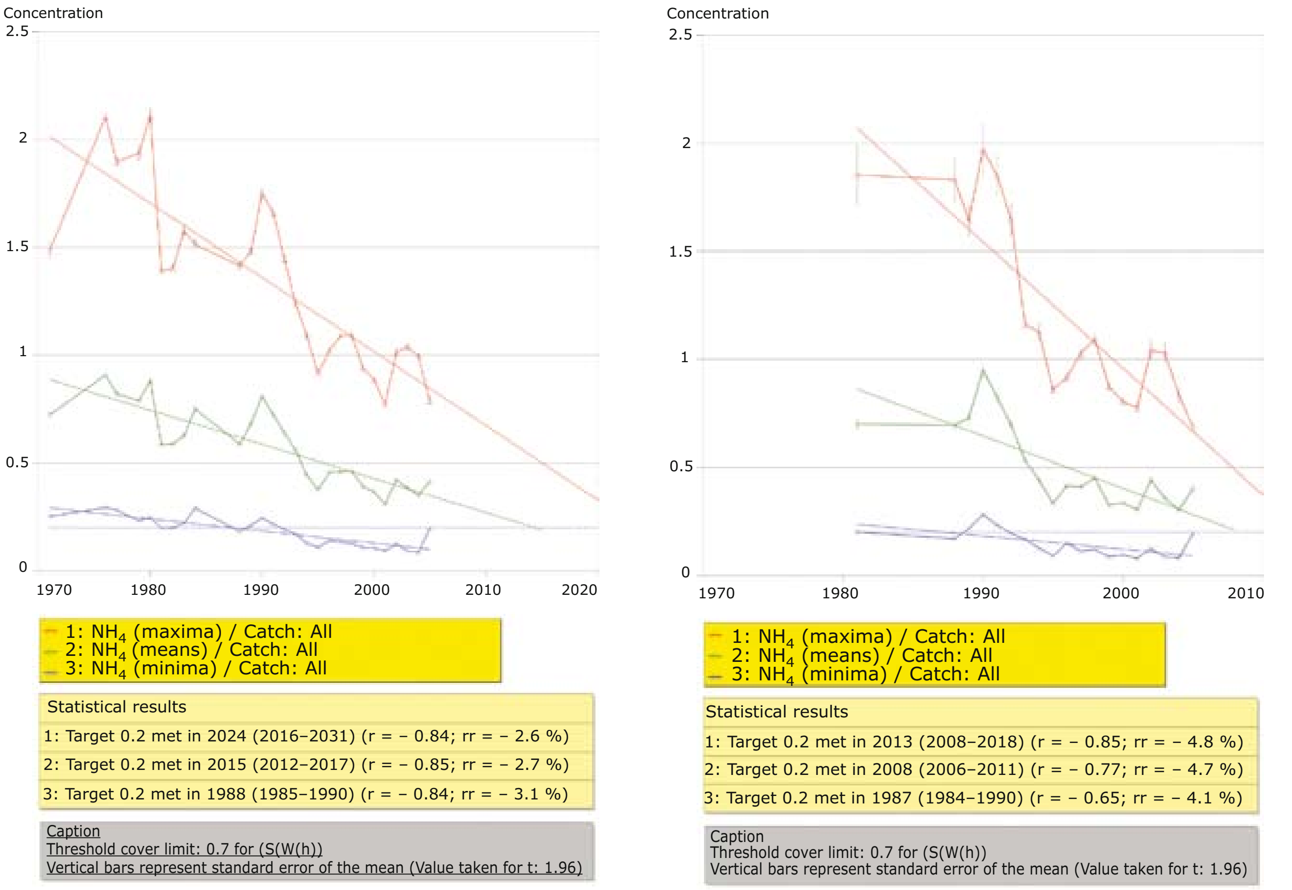
- The exploitable range of years;
- the visual profile of trends (they should normally match);
- the possible difference in important features;
- the trend assessment results, especially differences in the time of meeting the target;
- the relative error ranges (mean uncertainty/mean in %). The range is the 20 % and 95 % percentiles of relative error distribution.

5.1.3 *Comparison at the nationwide level*

Since stratified statistics are understood to be a robust method, both data sets are expected to

Figure 32 Ammonium trends (left: all data, right: Eionet data set)

Ammonium



The differences between the two results are summarised as follows:

Table 4 Comparison for ammonium, nationwide

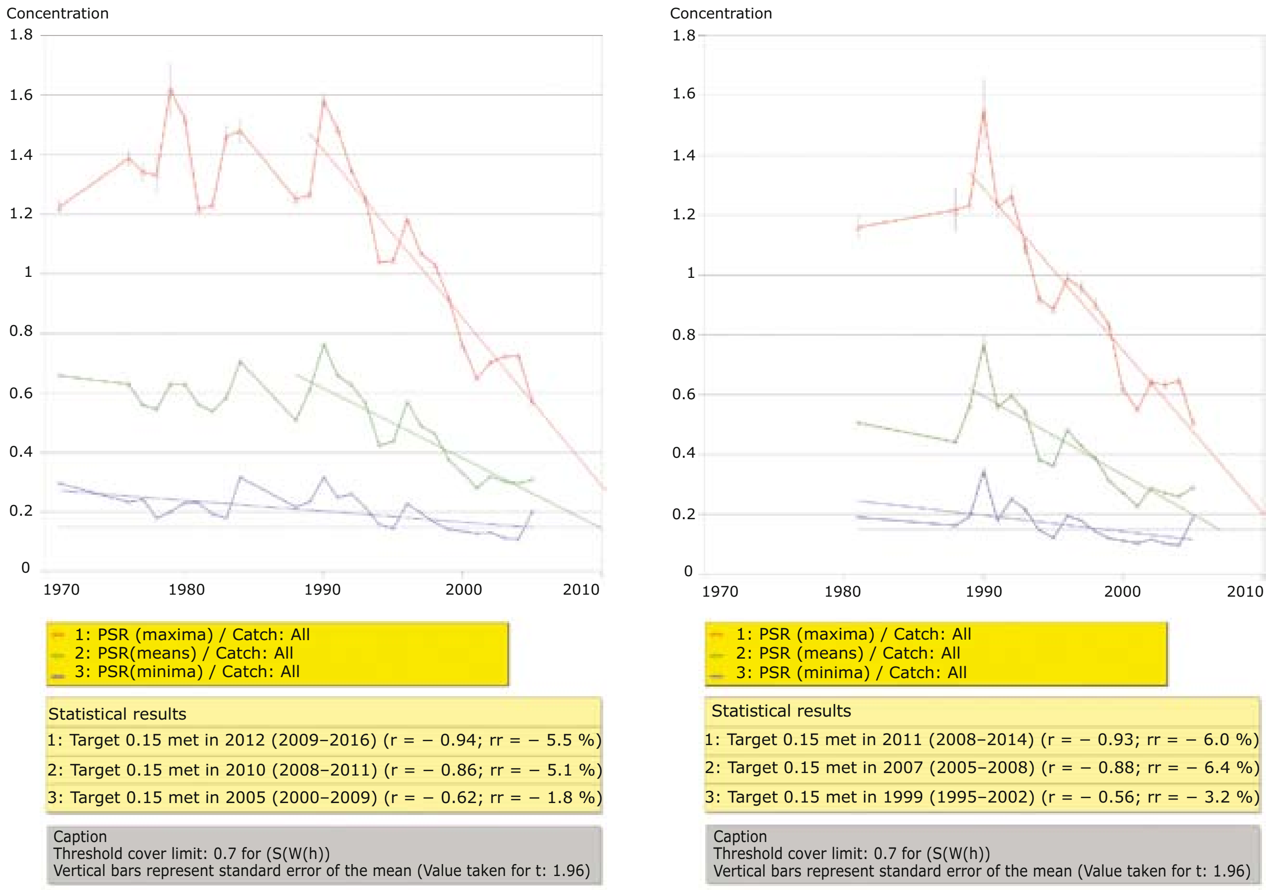
Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	1981–2005 (insufficient data before)
Visual profile	Comparable: apparent peaks in the same years	
Differences in features	Decreasing rate 2.6–3.1 %	Decreasing rate 4–4.8 %
Trend assessment	Target met in 2016–2031 (max), 2012–2017 (mean)	Target met in 2008–2018 (max), 2006–2011 (mean)
Relative error ranges		2–5 % (max), 13 % (mean) 0–1 % (min)

Note: Over the past 20 years, the same behaviour is assessed. A large difference in time of target meeting is observed, leading to different assessment as result.

Confirms the robustness of stratified approach.

Figure 33 Soluble phosphorus trends (left: all data, right: Eionet data set)

Soluble phosphorus



The differences between the two results are summarised as follows:

Table 5 Comparison for soluble phosphate, nationwide

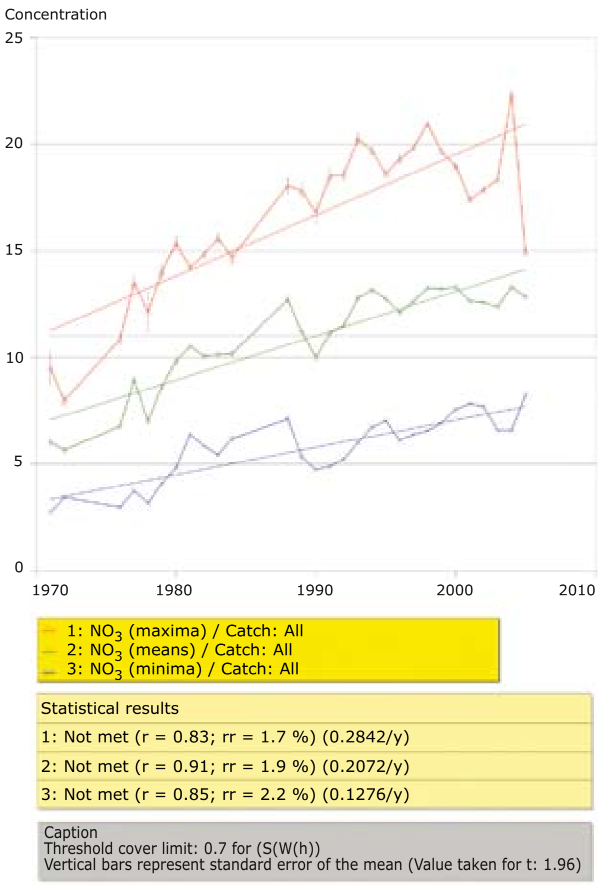
Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	1981–2005 (insufficient data before)
Visual profile	Comparable: apparent peaks in the same years	
Differences in features	Stable period 1971–1990 depicted; Decreasing rate 5.5–5.1 %	Non information on the 1971–1990 period. Decreasing rate 6–5.5 %
Trend assessment	Target met in 2009–2016 (max), 2008–2011 (mean)	Target met in 2008–2014 (max), 2006–2010 (mean)
Relative error ranges		1–3 % (max), 0–2 % (mean) 0–3 % (min)

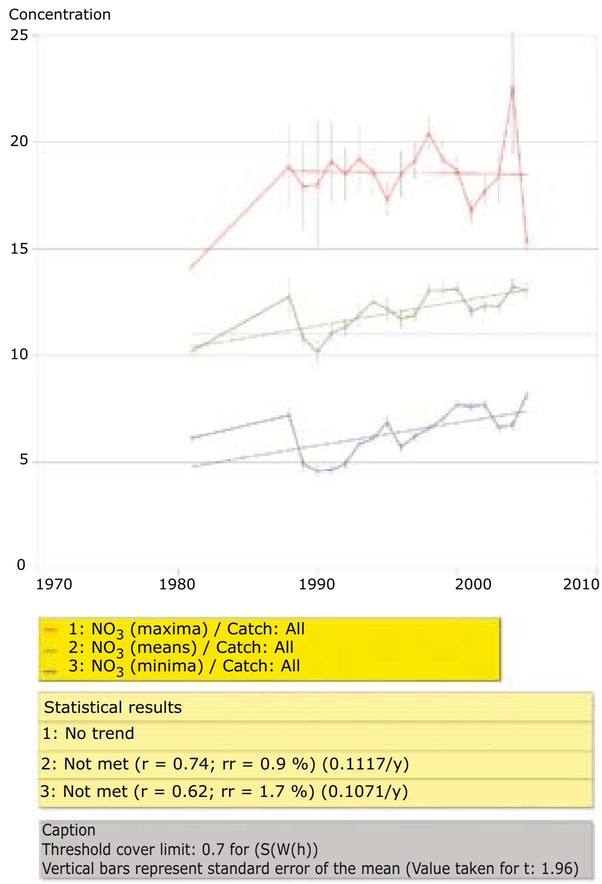
Note: Over the past 20 years, the same behaviour is assessed. A large difference in time of target meeting is observed, leading to different assessment as result.

Confirms the robustness of stratified approach.

Figure 34 Nitrate trends (left: all data, right: Eionet data set)

Nitrate





The differences between the two results are summarised as follows:

Table 6 Comparison for nitrate, nationwide

Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	1990–2005 (insufficient data before)
Visual profile	Comparable for minimum and mean; twisted for maxima; apparent peaks in the same years	
Differences in features	Increasing rate 1.9–2 %. History of changes well depicted	No rate, increasing rate 0.9 %. History of changes skipped.
Trend assessment	Can be differentiated and suggests change in trend	Does not show exploitable trend
Relative error ranges	1–4 % (max), 1–2 % (mean), 1–2 % (min)	3–12 % (max), 2–5 % (mean) 2–6 % (min)

Note: Over the past 20 years, different behaviours are assessed. The limited dataset suggests different conclusions than the total data set. General values are comparable.

Confirms the robustness of stratified approach.

Table 7 Comparison for ammonium: Seine, Loire and Garonne

Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	Seine 1982, Loire 1981, Garonne 1971 to 2005 (insufficient data before)
Visual profile	Seine widely lower, Loire much lower, Garonne quite higher from the Eionet data set	
Differences in features	Relative (and expected) ranking of catchments mostly different (except Loire). Seine obviously underestimated: averages from 0.29 to 0.16 instead of 1.18 to 0.4.	
Trend assessment	Decreasing respectively 8.9 %, no trend and 5.7 %/year	Decreasing respectively 3.3 %, no trend and 4.2%/year
Relative error ranges	0–2 % (SE), 1–2 % (LO), 1–3 % (GA)	1–3 % (SE), 0–3 % (LO) 2–13 % (GA)

Note: Totally different pattern, providing unexpected and doubtful assessments. Not exploitable.

provide comparable results at nationwide level. Similar rules for inclusion (70 % of strata should fit with data for the year to be computed) in both cases.

5.1.4 *Examples of assessment at basin level*

When considering smaller areas, it is expected that the Eionet data set (which in principle was designed to provide nationwide assessment) should be less effective than the complete data set. This is quite expected because basins with small areas of certain strata may be short on monitoring points on these strata, even though the strata are documented at the country level. In these cases, the computation rule may reject the calculations at basin level. In practice, fewer monitoring points favour homogeneous basins. By contrast, basin × stratum analysis could be carried out with the risk of experiencing excessive uncertainty. It should be noted in this regard that not all possibilities have been checked.

Ammonium

The assessment of the three large catchments on the Atlantic coast: Seine, Loire and Garonne yielded very negative results. The figures are not provided.

Ammonium is a very unstable determinant. It is therefore not surprising that large discrepancies happen. However, the shown differences pose many questions related to the possibility of using restricted data sets in this case.

Figure 35 Comparison for soluble phosphate: Seine, Loire and Garonne

Soluble phosphate

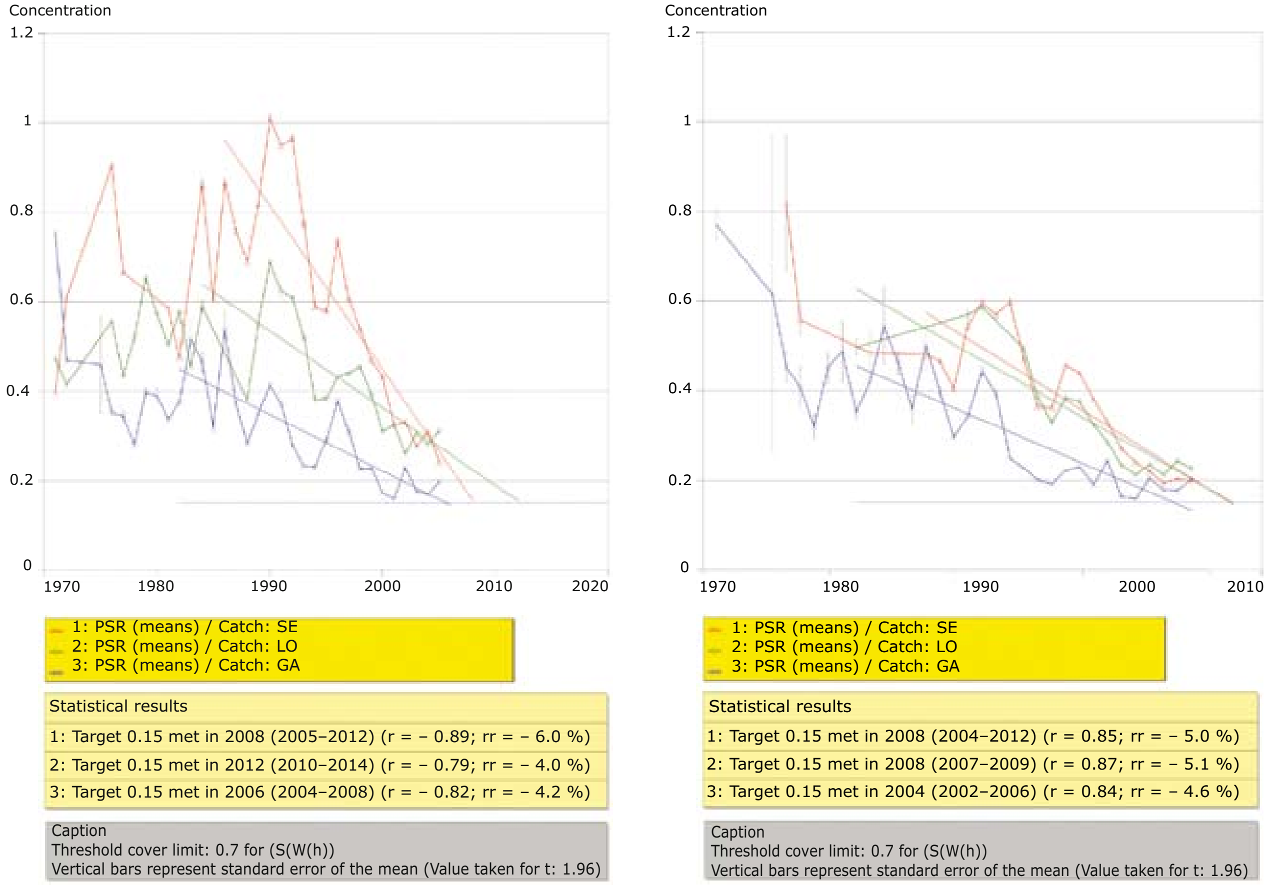


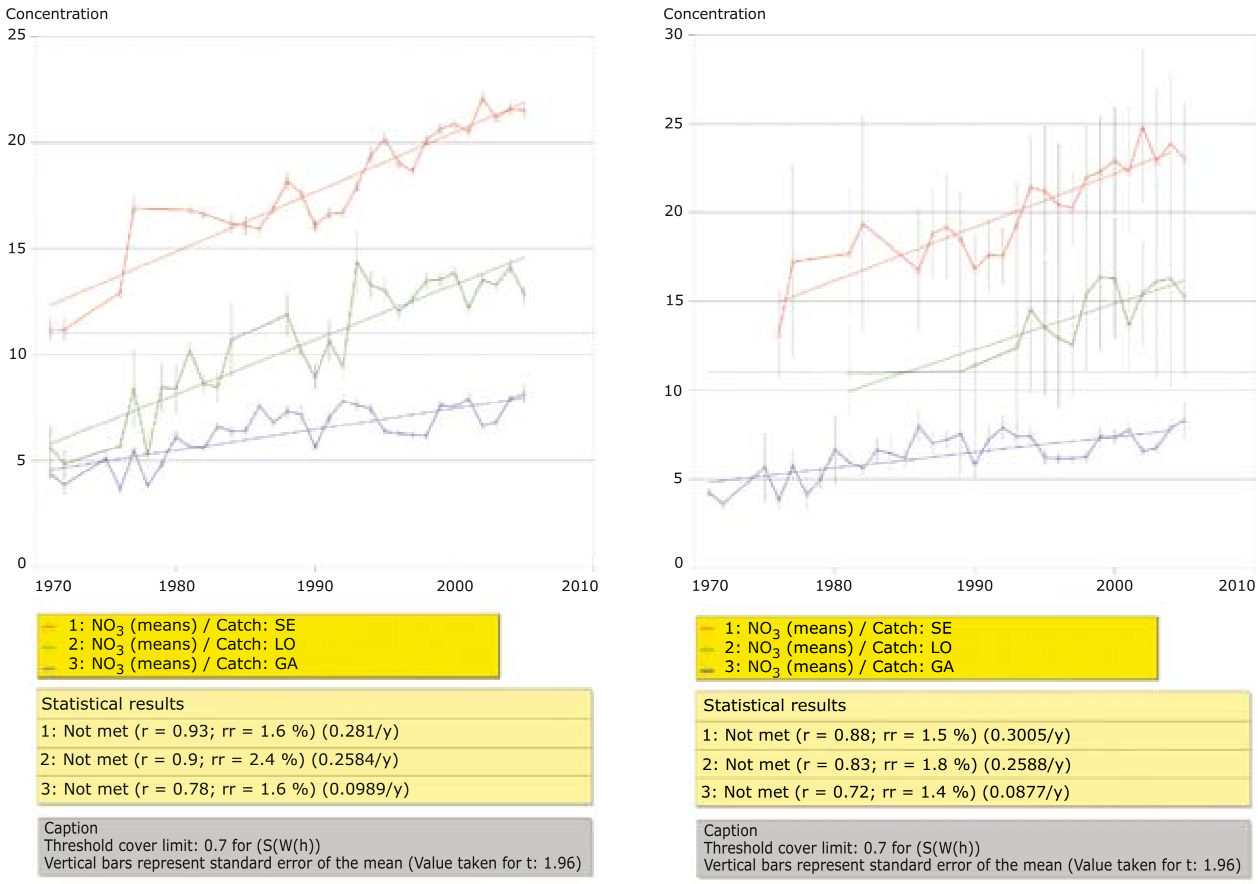
Table 8 Comparison for soluble phosphate: Seine, Loire and Garonne

Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	Seine 1976, Loire 1981, Garonne 1971 to 2005 (insufficient data before)
Visual profile	Seine widely lower, Loire erratic (lack of data), Garonne identical compared to the Eionet data set	
Differences in features	Relative (and expected) ranking of catchments different (except Loire). Seine underestimated: averages from 0.48 to 0.2 to 0.16 instead of 1. to 0.3	
Trend assessment	Same range of years at the end	
Relative error ranges	0–1 % (SE), 0–2 % (LO), 0–5 % (GA)	0–1 % (SE), 0–2 % (LO) 0–7 % (GA)

Note: Partly different pattern, providing doubtful assessments for the Seine, fuzzy for the Loire and identical for the Garonne.

Figure 36 Nitrate trends at basin level: Seine, Loire, Garonne (left: all data, right: Eionet data set)

Nitrate



The differences between the two results are summarised as follows.

Table 9 Comparison for nitrate, Seine Loire and Garonne

Criterion	Modified methodology	Eionet data set
Range of years	1971–2005	1976–2005; 1981–2005 and 1971–2005–2005 (insufficient data before)
Visual profile	Rather comparable, with some different peaks	
Differences in features	Rates are respectively for Seine, Loire and Garonne 1.6, 2.4 and 1.6	Rates are respectively for Seine, Loire and Garonne 1.5, 1.8 and 1.4. Seine and Loire present higher values
Trend assessment	Very accurate	Rather accurate
Relative error ranges	1–3 % (Seine), 2–18 % (Loire) 2–5 % (Garonne)	11–20 % (Seine), 20–45 % (Loire) 4–18 % (Garonne)

Note: Over the past 20 to 30 years, comparable behaviours are assessed. The limited dataset suggests identical conclusions to the total data set, but the certainty of assessment is jeopardized by very different uncertainty levels.

Confirms the robustness of stratified approach.

The situation is rather less acceptable when considering smaller basins. Taking Brittany, Channel (MA) and Bay of Biscay, the summary results are compared in Table 10.

Table 10 Comparison for nitrate at basin level, Channel, Brittany, and Bay of Biscay

Criterion	Modified methodology	Eionet data set
Range of years	Respectively for Channel, Brittany, and Bay of Biscay 1977–2005, 1971–2005 and 1971–2005	Respectively for Channel, Brittany, and Bay of Biscay no data, 1981–2005 and 1975–2005
Visual profile	Only Bay of Biscay presents comparable pattern. The typical pattern of Brittany (important growth followed by decline) is not visible; some decline can be supposed	
Differences in features		
Trend assessment	Very accurate	Not representative for Brittany, acceptable for bay of Biscay
Relative error ranges	4–11 % (Channel), 8–51 % (Brittany) 4–9 % (bay of Biscay)	No data (Channel), 54–257 % (Brittany) 18–44 % (bay of Biscay)

Note: Over three important catchments, only one can be assessed accurately. A second one, whose importance is great, would not exhibit a very specific and relevant pattern identified with a larger data set. The uncertainty in assessment does not allow clear-cut conclusions in both cases.

Confirms the needs of enough data for statistical assessment.

Figure 37 Comparison for nitrate in agricultural strata

Nitrate and agricultural pressures

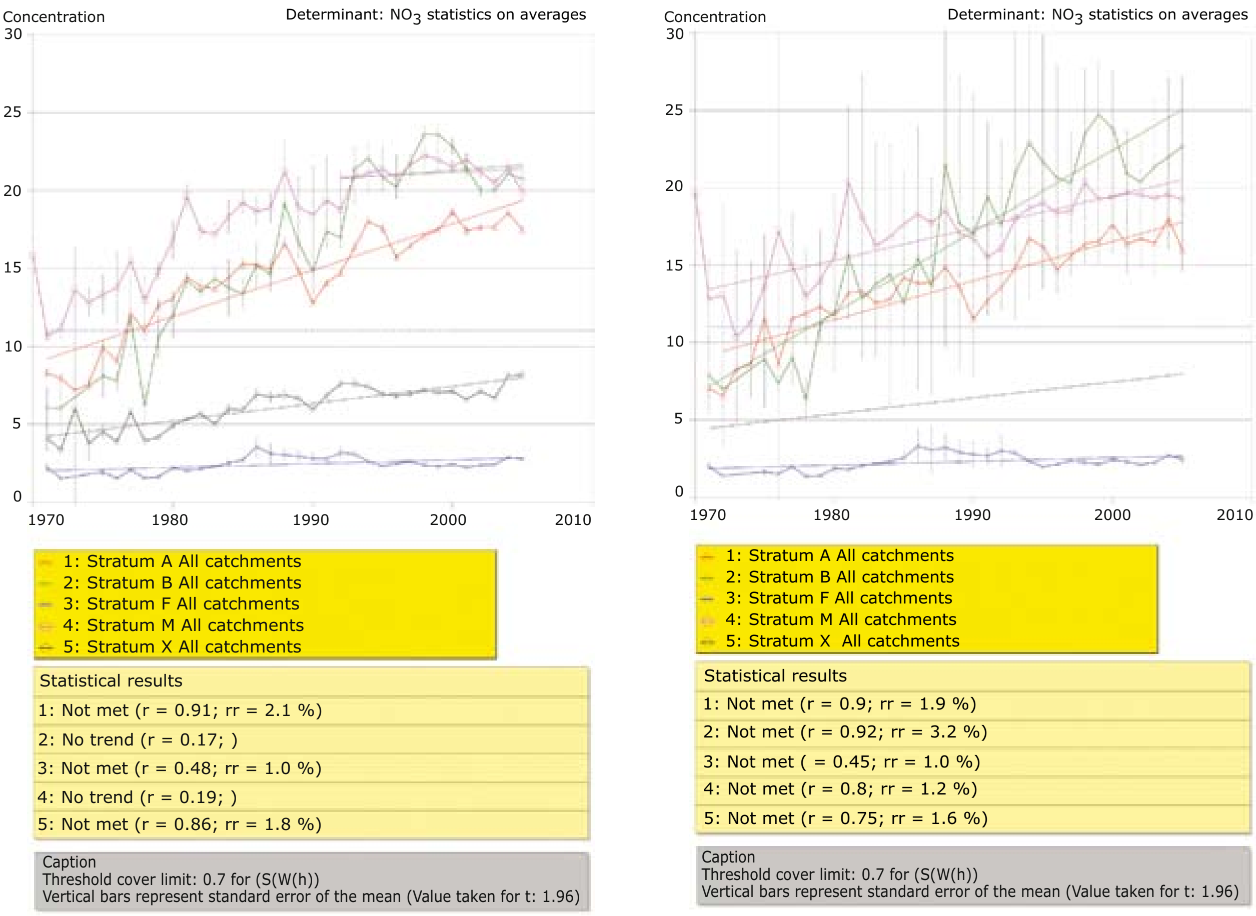


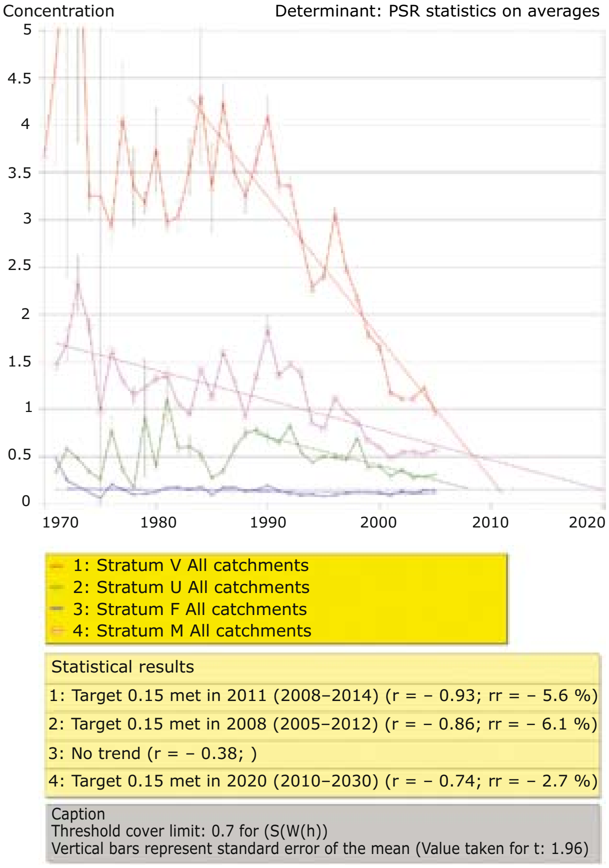
Table 11 Comparison for nitrate at stratum level for agricultural strata

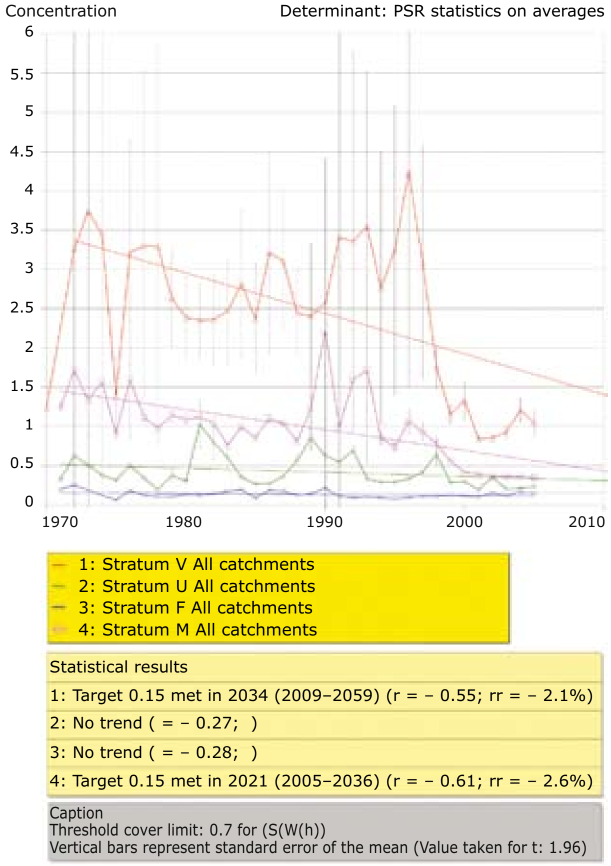
Criterion	Modified methodology	Eionet data set
Range of years	All 1971–2005	A: 1971–2005, B:1976–2005, F: 1971–2005, M: 1976–2005, X: 1971–2005. Series incomplete
Visual profile	Rather comparable, the M stratum is more flat and does not overcome the B stratum every time.	
Differences in features	With the exception of the B stratum, the starting and ending averages are identical	
Trend assessment	B and M stratum suggest reaching a maximum	Data too dispersed to support change in behaviour
Relative error ranges	A: 1–5 %, B: 4–20 %, F: 2–15 %, M: 4–16 %, X: 2–8 %	A: 8–19 %, B: 14–55 %, F: 3–33 %, M: 10–47 %, X: 5–23 %

Note: Data uncertainty is a major issue, because it i) changes the values for key stratum, thus jeopardizing source apportionment and ii) hides possible changes with high policy relevance.

Figure 38 Comparison for nitrate in agricultural strata

Soluble phosphate and urban pressures





5.1.5 Examples of assessment at stratum level

The comparison is limited to a nationwide comparison of drivers.

This comparison is very informative because again the picture from classical Eionet is similar to the detailed picture. However, it is quite blurred because uncertainty is much higher.

This comparison is very informative because in this case a key stratum — 'very urban', whose area extension is limited, is not correctly represented. As a result, no data can be presented. Where the assessment is limited to the period between 1988 and the present, the Mixed stratum is close to reference. Lack of data in the case of the 'very urban' stratum is directly related to the threshold for inclusion (see § 3.2.9). Removing the threshold makes it possible to display statistics by skipping missing values in all the poorly documented catchments at the expense of accuracy for the stratified average. In this case, the variability of the V stratum ranges from 16 % to 146 % and the comparable pattern for this stratum between 1990 and 2005 is found.

5.1.6 Partial conclusion

The differences in results between those obtained from the modified methodology and the classical Eionet relate only to differences in station numbers, because all the other processes are identical in the assessment. This difference results from:

1. Eionet requirements to select stations. This reduces the final number and then increases the uncertainty around the estimates;
2. station selection that results in inconsistent populating of some strata in certain basins.

The second reason is the most problematic because:

- either the selection has been carried out from a statistical procedure aimed at nationwide representativeness (which is fulfilled), and moreover, there are only random possibilities of assessing strata or basins with this data set;
- or the selection has specifically populated the different strata or the different catchments, thus making it difficult to ascertain the representativeness at nationwide level.

It seems unwise to carry out stratified assessment dealing with either basins or sector impact using Eionet's limited data set. Nationwide results can however be computed with acceptable levels of accuracy.

This conclusion does not indicate that sectoral or basin assessments should be abandoned; it states that their accuracy and consistency are not certain and that the conclusions are potentially controversial. Moreover, it seems unlikely that early warnings could be done with restricted data sets.

Hence, where ever more comprehensive data sets are readily available, they should be used and analysed to consider their potential of mimicking identical strata situated in areas where little data are available. To this end, correction by discharge is a prerequisite.

6 Impact on current data collection, processing and treatment

6.1 Context of the analysis

The analysis refers to two different outcomes:

1. The methodological improvements have demonstrated that provided there is larger data collection, better accuracy can be achieved. Therefore, this will provide more relevant, more reliable and better targeted information. The improvements in the stratification process itself cannot be demonstrated at this stage but are part of the next step of implementation with data from Austria and other volunteer countries.

2. The current recognition that the capabilities of stratified statistics are well delineated and restricted to some aspects of water quality issues. A comprehensive assessment of both status and policy effectiveness should back up other methodologies that require adequate data sets.

This section deals with the impacts of i) stratified-approach methodological improvements and ii) adding new methodologies on data collection and processing.

This analysis is obviously carried out using the rule 'minimum data collection, maximum outcomes' agreed at European level. The different requirements are therefore detailed in the next section.

6.2 Setting the scene of requirements

The need for complementary assessments to the Water Framework Directive (WFD) focusing on 'compliance' and towards 'SoE' is now recognised. The fact that the WFD water bodies must be nested within a GIS ensuring spatial integration of data is henceforth agreed. In all circumstances, the key issue is to define, construct and process the relevant statistical population to output relevant, reliable and comparable indicators.

For example, the relevant statistical population addressed in this report comprises the elementary catchments ([13]). Similarly, the WFD water bodies are made up of elementary reaches of all rivers at a certain resolution. This is precisely the population that is addressed by the water quality accounting methodology. Thus, the WFD water bodies represent a sub-population of the river reaches.

The relevance of information produced is dictated both by the quality of the monitoring and observation networks operated in countries and at supra-national level by the relevance of data processing systems. Monitoring networks are complex to design, politically considered as expensive to operate and must be stable enough to produce data in the long term.

A major problem is that a large share of data collected seems to provide little information, because substantial change is not shown. Many efforts are only devoted to monitoring the outstanding events at the expense of the reliability and representativeness of networks. This target is backed up by a deep epistemological confusion between status assessment and compliance which are wrongly understood as synonyms. For the time being, 'network optimisation' does not mean covering new needs but monitoring fewer. Trying to cut the expenses of monitoring by focusing measurements on critical periods can operate for compliance, but prevents any serious assessments in trends. For example, what if problems appear outside the critical periods in question?

Hence, cost-effective assessment of river quality should be guided by long-term policy priorities like those envisaged in the WFD. Appropriate methodologies can partly mitigate the limited lack of data. The use of the best methodologies is the only reasonable way to optimise the use of networking data and a lot of progress can still be made. This is a prerequisite to arguing for sufficient data.

The selection of appropriate processing systems is dictated by i) the questions relevant to policy follow-up and environmental assessment and ii) the versatile use of the same data sets to produce all the necessary and relevant indicators. A set of questions is for example:

1. Are **rivers, lakes and reservoirs** in good *state* when considered as ecosystems, i.e. considering

([13]) The latest naming is 'functional elementary catchments' as designed by ETC/TE in the EEA rivers and catchments database.

the rivers within their catchment and their ecological potentials? *(this report, to some extent)*

2. Is **water from surface and aquifers** *suitable* for human and economic uses? *(this report)*

3. Is the **quantity of water in the different compartments** enough for the uses and functions of ecosystems? Is the resource, including artificial resource, sustainable?

4. Is the **hydrological regime of continental resources** kept in phase with the requirements of other ecosystems and uses? Do the river systems fulfil their function of connecting other systems (migration, sediment transport, etc.)?

5. Are **sectoral policies** (i.e. addressing agriculture, industry, urban activities, etc.) *capable of improving* the status of rivers and when will they succeed? *(this report, to some extent)*

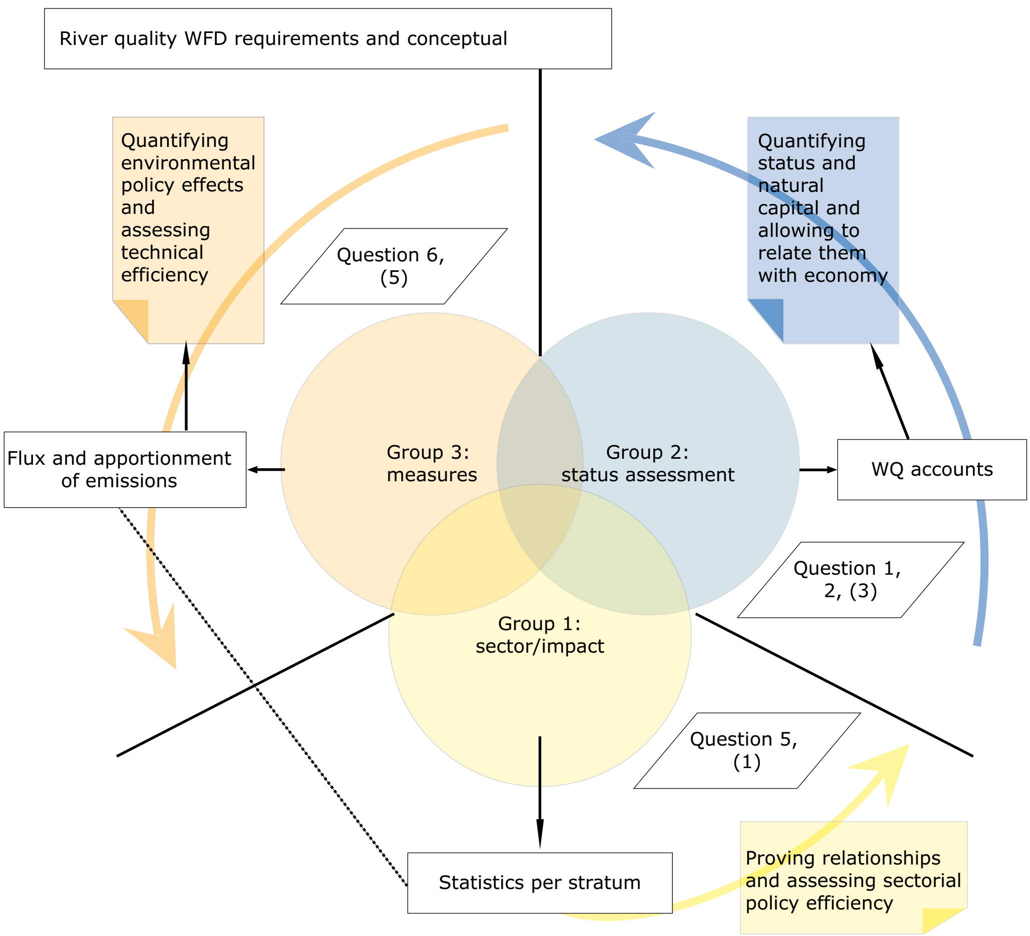
6. Do **programmes** of measures (i.e. STP building and operating, restoration works, etc.) carried out in catchments *address the correct concern* and, are they *correctly tailored*? *(this report, to some extent)*

The response to these six main questions requires a) water data from more sources than the WFD, b) data from other topics or sectors and c) strong spatial integration.

Example methods that apply within the group target are indicated in the white box and main outcome. By no means does this suggest that this is the only method that would apply.

The conceptual frameworks and tools developed and under validation at the EEA are comprehensive and advanced because they integrate the SoE compliance and accounting outcomes. The actual possibilities of monitoring and the existing scientific backgrounds as well as reporting requirements

Figure 39 Conceptual framework for meeting river quality requirements of the WFD and SoE



define a conceptual framework comprising three groups for result production. They cover the three 'legs' that are required for sound assessment of water quality issues.

The conceptual framework aims to use all relevant data and provide continuous space- and time-results. Three different groups of outcomes cover the different types of questions, but none covers all questions. Moreover, comprehensive assessment requires dovetailing reconciled outcomes from the different groups.

Group 1: Sector/impact assessment (objective of the current report). Its application requires information on the catchments, i.e. the statistical population addressed. This information is not part of the WFD reporting. It is important to consider that the statistical approach is a stratified sampling whose results are not spatially located despite the fact that they are highly spatialised. In addition, it provides a representative estimate of source apportionment that could back up Group 3 methods. Status assessment is an objective assessment of monitored values.

Group 2: Status assessment. Status considers the quality status of water systems. Quality is a *judgment*, based on monitoring, that refers to subjective uses and functions (including ecological potential). Different techniques afford expanding quality indexes computed at monitoring stations to river stretches not monitored. This approach addresses the population of river reaches and carries out statistical assessment based on the weighting of the different stretches by their size and their apportioning by quality status. Results are spatially located, by contrast with the previous group.

The results (different indexes can be created, some are validated, the River Quality Generalised Index RQGI for example) can be aggregated by catchment or by country, thus making interesting possibilities for comparison.

Group 3: Measures assessment and mass loads. Regarding water quality, most measures consist of sanitation works, purification works or bans on products use. Hence, measures change the loads of pollutants poured into waters (emissions) which are reflected in riverine loads. Economic assessment of measures (cost per tonne removed) would reinforce the other economic assessments. Moreover, they would lead to assessments of the effectiveness of measures after apportioning by source. The efficiency of measures with respect to quality is not necessarily reflected in loads because quality is very dependent on the low run-off period whereas loads integrate emissions over years.

Table 12 Summary of data requirements

Group	1		2		3	
Data set	**Elementary/ accurate**	**Aggregated/ proxy**	**Elementary/ accurate**	**Aggregated/ proxy**	**Elementary/ accurate**	**Aggregated/ proxy**
1. Quality data	++	+	++	©	++	©
2. Discharge data	NA	NA	++	++	++	+/©
3. Meteorological data	++	++	NA	NA	NA	NA
4. Catchment GIS	++	©	+	++	++	++
5. Rivers GIS	+	©	++	+	+	++
6. Quality stations	++	++	++	©	++	©
7. Discharge GS	NA	NA	++	+	++	©
8. Corine land cover	++	©	+	-	NA	NA
9. Population density	++	++	+	+	+	+
10. Livestock density	++	++	+	+	+	+
11. Emission discharges	NA	NA	NA	NA	++	++

Note: Elementary/accurate: need for detailed data set or accurate positioning; aggregated/proxy: possible use of aggregated data (e.g. statistical discharge instead of daily discharge) or proxy positioning (e.g. belongs to catchment X).
++: perfectly suitable; + : possible/marginal use, results accuracy not certain;©: not suitable; NA: not applicable (not needed).

Change in the hydrological regime may improve the quality indicator despite the fact that the actual load is increased. Hence, the relevance of measures is checked against mass loads in rivers, using apportionment techniques that also contribute to assessing data quality. Mass load analysis is a major complement to stratified statistics because they are aggregates on the same areas. In this case, stratified statistics would become a simple but robust source apportionment indicator.

The mass loads assessment and its derivatives have not been given much attention for the time being. They are however a direct use of the same data mobilised in Groups 1 and 2. For example, the sediment budgets assessment calls for sediment mass loads calculations and retention in dams.

6.3 Data requirements

6.3.1 Summary of data requirements

Data requirements for the different groups are very close and are summarised in Table 12. The assessment of requirements has been considered for independent supplies. In the event of all the applications being implemented, the required aggregates and proxies are better deduced from the most detailed assessment and should no longer be requested from data producers.

'++' in both columns means that aggregated is preferred and can be computed from detailed data. For example, a proxy positioning of quality monitoring stations is sufficient for stratified statistics (i.e. only the catchment and presence on main drain are required). If accurate positioning is available, it can be used as well. In the case of water quality data, the aggregated information acceptable in Group 1 is of no use in Group 2 and detailed values are more suitable for Group 1 itself.

6.3.2 Specific issue related to stratified assessments

When considering stratified assessments, two different questions are raised: the number of stations and the aggregated/disaggregated form of data.

The selection of stations is no longer supported by any reasonable argument. Fewer stations result in assessment of lesser quality even at nationwide level and may result in misleading conclusions. The currently available data processing systems make no difference between processing 3 000 or 30 000 annual data per determinant. By contrast, the time spent on interpreting contradictory data is significant and wasted.

The initial arguments for requesting only a selection of stations were twofold:

1. The homogeneous density of stations was understood to ensure comparability, considering the large differences between monitoring densities across Europe.

2. Many countries were (and still seem to be) reluctant in providing all their monitoring data.

The first argument is not correct because the true comparability results from minimum uncertainty in compared data sets. The current monitoring networks design has implicitly or explicitly reinforced the monitoring of polluted areas, presumably where variability is the largest. The comparisons carried out show that only the least polluted areas are estimated comparably in both data sets, whereas the most polluted ones are ill-assessed by the selected data set.

A major risk in processing severed sets of data is that of missing the subtle changes and the early warnings. Examples of this have been given.

The second argument is political. Within the EU regulations, each Member State can decide which monitored observations it wishes to make available. More and more countries have developed the means to provide free of charge data to the general public; the EEA would be unwise to use insufficient data sets when NGOs could for example process complete sets. The risk for the EEA is that its assessments will be less accurate.

This does not mean that all monitored data should be collected and processed. On the contrary, the selection should apply to the collection of all quality assured data that provide the best certainty of estimates at the area, sectoral and time resolution that the EEA considers.

The solution to this question is the outcome of a scientific approach involving attempts, errors and correction and should not be the outcome of an *a priori* compromise. Aggregated data are one of the ways the observations can be used. The provision of aggregated data, as considered by the current Eionet-water, poses three different problems:

1. Aggregation is a burden for data providers and there is no guarantee of homogeneity of

the aggregation method, which defeats the requirement of full comparability. It is more effective to carry out a process of aggregation with a single method than having it done 35 times with different systems.

2. Aggregation cannot be asked of data providers using different methods (e.g. mobile time windows, different time/discharge weighting, etc).

3. Pre-aggregated data is not suitable for the two other ways of data assessment (water accounts and flux calculations) and prevents any improvement of the current stratification methodology.

There is no solid argument for recommending the aggregation of data. Moreover, further data collection should consider collecting disaggregated data on all relevant monitoring points; the relevance of which being defined by the usability of data.

6.4 Pending methodological questions on stratified statistics

Methodological questions have been raised before the realisation of this work (Beture-Cerec and ARMINES, 2000; Beture-Cerec and ARMINES, 2001; EEA, 2001) and partly solved. New ones result from the analysis presented in this report. The methodological issues raised address both fundamentals in statistics and the way to tackle the difficult problems of carrying out feasible calculations with the existing data.

6.4.1 Fundamental questions

- Is it possible to compute stratum averages (or any statistics) weighted by discharge instead of time and avoid double weighting, thus resulting in spurious assessment? What is the real-world significance of such weighted statistics?

- Is it possible to compute stratum quantiles (currently estimated by the average of maxima/minima) and assess their uncertainty?

- Is it a requirement to include the uncertainty of each set of statistics at the monitoring point to compute stratum statistics uncertainty? If yes, can the method be automated to make the computations affordable? In other words, can the current option of computing separately the point statistics and the stratum statistics from point statistics be kept or should stratum statistics be recomputed from elementary data each time new stratum delineation is decided?

6.4.2 Practical questions

- Filtering stratum statistics with a hydrological variable (e.g. relative effective rainfall, as experienced in 2000 at the EEA) undoubtedly improves the quality of the trend assessment in concentrations. What is the best way to introduce this new variable (and how is it computed)? What is the impact on the uncertainty of the estimate of meeting the target?

- The uncertainty around the year in which the target is met is computed from the regression concentration = f (year). In this case, statistical rules can provide the uncertainty of concentration, not of year. Is this method acceptable as proxy, and could it be improved with reasonable effort?

- According to the recommendations of the first implementation study (OIEau and ifen, 1997), stratification is built from the number of catchments falling into a certain category. It is known that this method yields unsatisfactory values at the expense of simple implementation. It may cause some unpredictable uncertainties because the size of catchments varies from place to place. This discrepancy in size is not expected to be solved in the next Rivers and Catchments database made from CCM2.

  Replacing the catchment number by the area is a candidate solution but its implementation is rather complex. How do you estimate the benefit of doing so and which area should be used to qualify stratum weight? (The cumulated area of catchments just marked by stratum seems the simplest solution).

- Computing stratified statistics means considering the lack of data in existing strata for a certain area selected. Otherwise the statistics are incorrect. The sum of weights (S(W(h)) in the figures) is set to a certain value below which calculation is denied. Above this value a systematic correction is applied by reallocating the existing strata to the total area (thus suppressing the non documented stratum). This is a source of possible uncertainty that has the advantage of being simple. A special procedure aimed at inserting standard values from the ad hoc strata should be developed. Is such a procedure desirable? How should the replacement values be populated?

7 References

Beauzamy, B., 2004. *Méthodes probabilistes pour l'étude des phénomènes réels*. Vol. 300: Les mathématiques du réel. Paris. Société de calcul mathématique, ~ 300 pages

BETURE-CEREC, 2000. *Calcul des flux de nutriments véhiculés par les cours d'eau. Application pilote aux bassins versants du littoral de l'Agence de l'Eau Loire-Bretagne*. Report type: final, written by Guillaume Le Gall, ordered by 'ifen, Agence de l'eau Loire-Bretagne. Orléans, 59 pages, (6 annexes), availability: total.

Beture-Cerec and ARMINES, 2000. **Eurowaternet. Construction d'un réseau représentatif de qualité des cours d'eau. Phase I-Méthodologie: Analyse exploratoire des données**. Report type: final, written by Chantal de Fouquet, ordered by ifen et Agences de l'eau. Orléans, 76 pages, availability: total.

Beture-Cerec and ARMINES, 2001. *Eurowaternet. Construction d'un réseau représentatif de qualité des cours d'eau. Phase II-Rapport final*. Report type: final, written by Chantal de Fouquet and Guillaume Le Gall, ordered by ifen et Agences de l'eau. Orléans, 233 pages, (6 annexes), availability: total.

CEEP, 2000. *Sources of phosphorus to surface waters: comparing calculated with measured P loadings for three French rivers*. Report type: final, translated, written by Jean François Lassevil and Delphine Berrux, ordered by Centre Européen d'Etudes des Polyphosphates, 68 pages, availability: total.

Crouzet, P., Leonard, J. and Nixon, S. *(Eds.)*, 1999. *Nutrients in European Ecosystems.* EEA, Environmental assessment report No 4, 155 pages.

Council Directive 91/271/EEC of 21 May 1991 concerning urban waste-water treatment, OJEC 30/05/1991, pp. 40–52.

Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources, OJEC L 375 31/12/1991, pp. 1–8.

EEA, 2001. *Revisiting technical issues related to river quality reporting within the current Eurowaternet process. New insights to assessing sectoral policies efficiency*. Report type: draft, writen by Philippe Crouzet, ordered by 'EEA/EIONET') Copenhagen, 38 pages, availability: limit.

EEA, 2005. *The European Environment. State and Outlook*. Copenhagen. EEA. 584 pages.

IFREMER, 1993. *Cartographie des zones sensibles à l'eutrophisation. Cas des côtes bretonnes*'. n° R.INT. DEL/93/13, ordered by Ifremer/DEL Brest.

Ifremer DEL/EC, 2001. *L'eutrophisation des eaux marines et saumâtres en Europe, en particulier en France*. Report type: final, written by Alain Ménesguen, n° IFREMER DEL/EC/01.02, ordered by Commission Européenne DG. Env B1. Brest, 59 pages, availability: total.

Juin, D. *(Ed.)* 2002. *L'environnement en France.* Paris et Orléans, La Découverte et Institut français de l'environnement. 602 pages.

Kaczmarek, B., 2006. *Un nouveau rôle pour les agences de l'eau ? Essai pour une politique de l'eau rénovée*. Paris. Johanet. 238 pages.

Lavoux, T. and Rechatin, C. *(Eds.)*, 1998. *L'environnement en France.* Paris et Orléans, La Découverte & Syros et Institut français de l'environnement. 472 pages.

Leonard, J. and Crouzet, P., 1999. *Construction d'un réseau représentatif. Contribution au réseau 'EUROWATERNET'/Qualité des cours d'eau de l'Agence Européenne de l'Environnement. Notes de méthode n° 13.* Orléans. Institut français de l'environnement. 70 pages.

Merceron, M. *(Ed.)* 1999. *Actes de Congrès: Pollutions diffuses : du bassin versant au littoral.* 23–24 septembre 1999. Brest, Editions IFREMER. Actes de Colloque n° 24. 350 pages.

NERI *(Ed.)* 2005. *Source apportionment of nitrogen and phosphorus inputs into the aquatic environment*. Copenhagen, European Environment Agency. EEA report No 7/2005, 52 pages.

Nixon, S. C., 1997. *European freshwater monitoring network design*. Topic report. Inland Waters No 10/1996. Luxembourg. Office for Official publications of the European Communities, 129 pages.

OIEau and ifen, 1997. *Réseau de suivi des eaux continentales de l'agence européenne de l'environnement. Projet pilote en France.* Report type: final (version française), written by Jennifer Leonard, Philippe Crouzet *et al*., ordered by Agences de l'eau, Institut français de l'environnement, Office international de l'eau Orléans, 101 pages, availability: total.

OIEau and ifen, 1999. *Contribution to EUROWATERNET. River water quality in France. National prototype (phase 1)*. Report type: final (english version), written by Jennifer Leonard and Philippe Crouzet, ordered by Institut français de l'environnement, Orléans, 40 pages, availability: total.

Appendix 1: Calculation of statistics per stratum

1 Statistical baselines

1.1 Stratified average

The stratified average is the weighted average of the values attached to each stratum. Hence, the stratified average can be constructed from elementary values attached to each station marked as belonging to a stratum. This allows preparation of the data and computing these data at the moment the final domain is defined. Three variables have been considered as being relevant in previous studies: yearly (seasonal) average, maximum (as proxy of upper quantile X) and minimum (as proxy of lower quantile Y) at each monitoring point. The stratified statistics are respectively stratified mean (mean of point averages), stratified maximum (mean of maximum values) and stratified minimum (mean of minimum values).

These means can be presented by stratum, by catchment or by grouped strata or grouped catchments together, with intermediate aggregates.

1.2 Averages of determinant values

Determinants may belong to two different classes: standard determinants, whose average can be computed directly (e.g. nitrate) and non-standard determinants whose direct average is nonsense (e.g. pH). In both cases, the annual (or periodical) average must be computed in a way that avoids bias. This is achieved by computing time weighted average along the averaging period. This method mimics the average concentration of a water sample abstracted at a constant rate throughout the calculation period. It is the most effective way to get rid of the odd time distance between samples across a monitoring network (for example, the frequency of sampling may range between 4 and more than 24 per year, at variable intervals of time).

The initial averaging is carried out per sampling point, determinant and aggregation period. Special methods are used to compute pH average. For example, the time averaging and specific calculations are merged so that final results are adequately computed.

Correct methods for averaging non-standard determinant are very important. For example, in the case of eutrophicated river, in which pH is very high in summer, the difference between a classical arithmetic mean of pH and chemically accurate mean ranges between 0.8 to 1.2 pH units. This calculation error is much larger than the differences in pH that are mentioned in quality assessment grids.

Calculation methods systematically aim at computing the value of the average water that would have been obtained from continuous pumping at a constant rate throughout the year. This is the only way to compute secondary stratum averages that would not be adequately defined if discharge had been used instead of time as the weight of annual averages.

1.3 Aggregation rules of stratified statistics

Aggregation rules deal with two different problems: a) calculating the most accurate stratified mean and adequately using the available weighting information, and b) assessing the estimation error on the stratified mean.

In the following text, notations and formulas are adapted from training sheets 6.3 by R. Schlaepfer, *Cas de l'estimation de la moyenne d'une population de taille finie*, that apply in this case. Source: http://gecos.epfl.ch/gecopa/Enseignement/Fiches/Fiche63Echant2002antill.pdf checked 31/7/2006.

1.3.1 Accurate stratified mean

Classical stratified statistics apply in this case, calling *st* the stratified population, *h* the stratum (h = 1, L). Lower-case fonts mark statistics related to stratum and upper-case fonts mark statistics related to the stratified population. Observations are noted '$y_{(h, i)}$', indicating the ith observation on stratum h. The variables N_h, N and n_h, n representing respectively stratum number of items in stratum h, number of items in total population, number of items in sampled population and number of items in sampled stratum. In the optimum case, the ratios n_h/n and N_h/N should be equal, expressing identical sampling efforts. This unfortunately is not the case in real applications.

The estimator of the mean of stratum *h* is:

Equation 1

$$\overline{y_h} = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{(h,i)}$$

The mean is an undistorted estimator of mean of population in stratum h, hence the stratified mean of population is the sum of the weighted averages per stratum:

Equation 2

$$\overline{y_{st}} = \sum_{h=1}^{L} \frac{N_h}{N}\,\overline{y_h}$$

Since the estimate of the mean is not sensitive to the number of individuals, differences in the actual sampling rate between strata is not a source of bias. The uncertainty of the mean estimate is computed at a second stage. This is why the estimator of the mean is computed from the actual number of the individual inside a stratum and not from the sampling ratio of stratum, as in Equation 2. the ratio (N_h/N) is called W_h in the next sections.

1.3.2 *Practical implementation*

In practice, some strata may not be represented in a certain aggregate. For example, the V stratum does not exist in a certain catchment, or a stratum is marginally present in a certain catchment and not sampled, or data from the stations in a certain stratum may not be usable.

If sampling is incomplete, a radical option is to discard all results. In practice, this would result in discarding almost all results. Analysis of the distribution of sums of W_h suggests that less than 10 % of combinations would really be usable. The response to this unacceptable situation is:

- to compute a virtual W_h, correcting the true W_h by reallocating the weights to the existing strata,
- to replace missing stratum values by proxy values, taken from documented locations.

The second option does not provide good assessment and cannot be automated, and was therefore discarded. An automated option was implemented in a service module added to NOPOLU *Système 2*. It consists of choosing an inclusion percentage, below which data is discarded. The inclusion percentage takes into account the size of the non-calculable strata.

In practice, Equation 2 is replaced by an adjusted calculation:

Equation 3

$$\overline{y_{st}} = \sum_{h=1}^{K} W_h\,\overline{y_h}\,'$$

K is the number of strata actually usable, instead of L true number of strata. The sum of W_h must be 1, involving a small distorting in the relative weight of strata.

Equation 4

$$W(h) = \sum_{h=1}^{L} \frac{N}{N'} \times \frac{N_h}{N}$$

Where N is the total number of actually inserted elements of the K calculable strata the way Equation 4 is presented is deliberately redundant in order to take into account the inclusion percentage. It ensures that the sum of W(h) is 1 in any case. Otherwise the weighted average would be biased.

1.3.3 *Estimating the error of the stratified mean*

The estimation error of the mean deals with two distinct problems; of which one still remains unresolved. Estimating the error of the seasonal (yearly) mean is the result of classical techniques that have to take into account the internal model of data. Theoretical problems are simple. Practical issues are insurmountable because each point, each determinant and each year require special tuning of the internal model characteristics, otherwise added errors would become larger that lack of correction with the internal model.

When addressing the maximum and minimum at a point, there is no estimation error. Nevertheless, the maxima and minima would be replaced by 90 % and 10 % percentiles and the question of estimating their respective accuracy would be raised.

Examination of literature did not provide substantial evidence that stratum mean error estimates should take into account the estimation error of each individual mean. It has not given any evidence that it should not, thus leaving the issue unresolved. From a practical point of view, neither the individual error estimates can be computed nor the stratified estimate accounting for individual errors be modelled. Hence, the error estimate is limited to the classical stratified error that represents only the uncertainty of calculating a stratum mean, disregarding the overall error of the information attached to this mean.

A possible hypothesis is that the error of the estimate of the mean depends only on the different values, disregarding their individual errors. By contrast, error of the estimate would probably require consideration of the individual uncertainties around each value constituting the stratum average ([14]).

In this report, the confidence interval around the stratum mean is defined assuming that each individual average is the observed information. This is the way stratified surveys are built: the elementary accuracy of each response is not considered and they are understood to be the realisations of a random variable.

Under this hypothesis, the central limit theorem applies, suggesting that the mean is a Gaussian variable. This is the case for the stratum mean (mean of averages, mean of maxima, mean of minima). The estimation error is therefore computed from the estimator of observations variance computed from the different values entering in the stratum:

Equation 5

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(\overline{y}_h - y_{(i,\,h)} \right)^2$$

Hence, the variance of the mean estimator is the variance of observed values normalised by the sampling ratio in stratum. It hyperbolically reaches 0 if n_h tends to N_h, which is quite obvious.

Equation 6

$$v\left(\overline{y_h}\right) = \frac{s_h^2}{n_h} \frac{N_i - n_h}{N_h}$$

Hence, the estimator of the population mean becomes:

Equation 7

$$v\left(\overline{y_{st}}\right) = \sum_{h=1}^{L} \left(\frac{Nh}{N} \right)^2 v\left(\overline{y_h}\right)$$

By application of the central limit theorem, the estimation error of the overall mean is +/– the standard deviation (square root of the result Equation 6) times the value of the normal distribution at the $(1 - \alpha)$ confidence limit imposed.

2 Practical implementation and use

The calculation of stratified means and their associated estimate error is simple; what is complex is managing the incomplete data sets, especially the extrapolation of missing minority strata. This is handled by a coupled MS Excel® and MS Access® application, developed as an extension of NOPOLU *Système 2* 'EuroWaternet' module.

In practical implementation, W_h and w_h respectively substitute N and n where necessary.

Data exported by standard functions of NOPOLU *Système 2* are processed semi-automatically by an MS Access® prototype application designed to be further incorporated into NOPOLU *Système 2*. The MS Excel® counterpart analyses the data content, generates queries and gives results according to the interactive selection. Moreover, it manages the trends assessment on selected time periods and targets. This tool, developed especially at the EEA, has been used to prepare all figures of trends in this report.

([14]) Analogous problem is discussed in Figure 2: a first error is related to the regression line (how thick is it?) and a second error is on the possible dispersion of the points that are aggregated on a regression line.

Appendix 2: Codification of stratification scenarios

1 Recommendations for coding the stratification scenarios

Coding the scenario is an iterative process in which small adjustments have to be made in order to optimise the stratification. The issue is to decide when the acceptable result is eventually reached. The simple criterion is that an optimum is achieved when it is no longer possible to reduce the number of elementary catchments falling into the control stratum (labelled Y). Common sense and cartographic representation help defining the final ad hoc stratification which, again, must be customised to the objectives of the assessment.

Experience with the stratification module in the NOPOLU *Système 2* application suggests the following counsels that deal with and could be applied with any application:

- the number of strata defined;
- which rules should define the different strata;
- how to define thresholds between strata.

Good stratified analysis should be based on a very small number of strata in order to capture the essential correlations.

2 Analysing the relevance of stratification criteria

Once the stations are attached to the main drain, they become the adequate measurement site of the catchment outlet and preparatory statistical analysis can be carried out. The target of the analysis is twofold:

1. To assess the correlation between stratum variables that would make their inclusion inconsistent. This is a possible problem because all Corine land cover derived variables are related to each other: the sum of their proportion is 1 in any case. However, the correlations may be quite weak in sub-domains. For example, crop area and forest area are related between 0 % and 100 % because they are mutually exclusive through the whole range. By contrast, a significant proportion of forest can be observed in parallel with a limited, albeit polluting, proportion of crop area. If yes, the area is not potentially 'pristine'. This analysis helped to construct the scenario rules.

2. To assess the relevant (or 'least worst') thresholds. For example, there is a continuous and very significant growing correlation between crop area (AG1) and nitrate concentration. It cannot be deduced that since 1 % of AG1, the catchment is dominated by 'intense agriculture'. More detailed analysis, filtering other components, helped to refine the rules and select a rounded but nevertheless relevant threshold (40 % of AG1) that in the absence of certain other components defines the most likely 'intense agriculture' driver.

The detailed results are not reported.

3 Coding stratification scenarios

3.1 Preparing the relevant components for building strata

When deciding which criteria to include in the rules, it should be kept in mind that the set of rules must carry out the stratification as a hierarchical procedure (i.e. a marked item at a certain step cannot be marked again in a next step). Therefore, it is important to make a transparent and simple coding in order to make it understandable. All rules must be finely checked, because unexpected results are very likely to occur: **rules and thresholds are intertwined. They must be finalised together during an iterative process and checked at each step.**

Dealing with agriculture and population-stable variables, three sets of data have been used: Corine land cover areas, population census and livestock quantities.

The 44 Corine land cover codes are grouped into five predefined classes which are already mentioned in Section 3.2.4 and reflect different pressure potentials.

The five groups are (15):

1. CLC_AG1, representing the intensive crop farming (e.g. arable land and irrigated areas)
2. CLC_AG2, representing the less intensive crop/livestock management methods (e.g. pastures, meadows and orchards).
3. CLC_FOR, clustering forestry types and the natural areas.
4. CLC_URB, representing the man-made artificial areas.
5. CLC_AUT, hosting all the types without relation to the aims of stratification (glaciers, rocks, etc.).

The relationships between the different groups and the determinants considered in the study reflect the potential impacts of the activities exerted in these classes: intensive agriculture is supposedly a source of nitrate; urban areas are sources of BOD5, ammonium and phosphate compounds; forested areas are understood to have limited leaching of nutrients, etc.

By contrast, pesticides will be strongly related to land use with fruit trees and berry plantations. In this case it will be misleading to place them into the grouping of CLC_AG2, if the less intensive farming method is to be represented. In the case of nutrients-oriented assessment, it is correct, because fruit trees and berry plantations are not fertilised with nitrate. Therefore, a totally different scenario should be built and exploited.

Population values come from the 1999 census, apportioned per catchment. Livestock density has been computed from the 2000 French agricultural census, apportioned per catchment and expressed as N input per year converted to population equivalent assuming 12 g N pe^{-1} d^{-1}. All values were computed using the Corine land cover based surplus model of the NOPOLU Système 2 application using the 'Recensement Agricole 2000' data.

The first step requires seed thresholds to be defined. These seed thresholds come from previous experience or from simple statistics for the criteria in relation to the elementary catchments. These statistics relate the stratum components to the analysed variable. They are not presented in this report.

At the end, a set of rules with clear-cut thresholds is applied. Table 13 below displays the final coding of the stratification scenario. The figures on the first line represent the calculation order: first the F stratum is set to the corresponding catchments, then stratum X and so on.

Table 13 Codification rule of the final stratification scenario

Order	0	1	2	3	4	5	6	7
Variable	**F**	**X**	**V**	**M**	**U**	**A**	**B**	**Y**
AG1	< 25	AG1 < 40 or AG1 + AG2 < 60		AG1 > = 40 or AG1 + AG2 > = 60	< 40	AG1 > = 40 or AG1 + AG2 > = 60		
AG2								
FOR	FOR + AUT > = 40							
AUT								
URB		< 2.5	> 10		URB > = 2.5 and POP > = 78 or URB > = 1.5 and POP > = 90			
Population	< 40	< 78	> = 150	> = 78		< 78	< 78	
Livestock	< 50	< 200				< 200	> = 200	

(15) Names are the variable names in the NOPOLU *Système 2* application used to carry out stratification of catchments. They are kept for convenience.

4 Final scenario selection

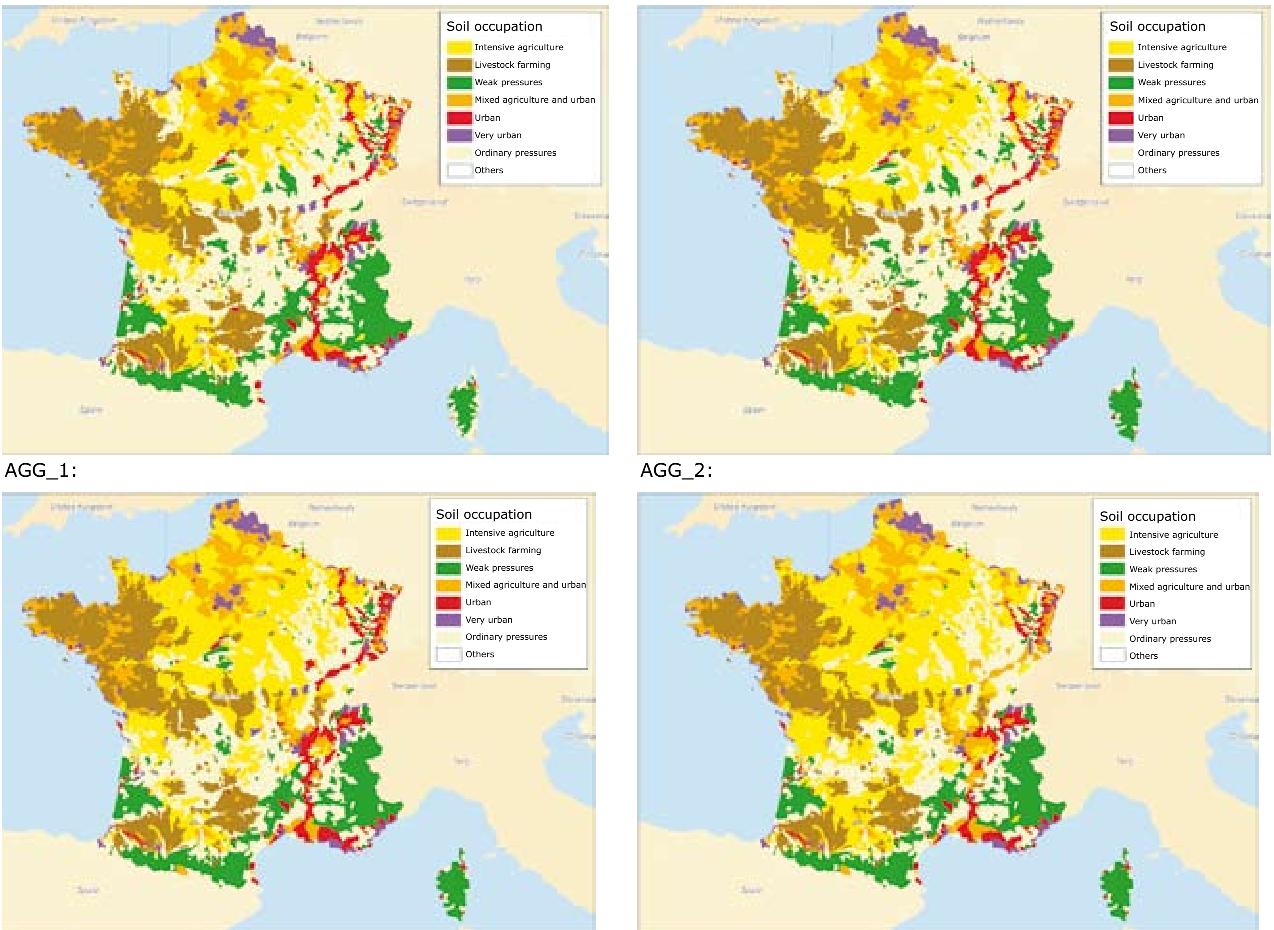
Designing the correct strata is a step-by-step process that is illustrated by four slightly different CLC groupings. Figure 40 presents the geographical distribution of the different strata. Details of the applied scenario are found in Table 13, and further information on the composition of the predefined five CLC aggregates is available at the end of this report.

The strata in close relation with agricultural activities are rather sensitive to small changes in threshold values used in stratification scenarios.

Figure 40 suggests that the 'Intensive agriculture' stratum is the most sensitive stratum to variability in the CLC groupings. The main reason for the huge difference between the number of catchments categorised as Intensive agriculture in the scenario using AGG_1 and AGG_4 is that pastures (CLC code 231) are not included in the land use class CLC_AG1 in the CLC grouping of AGG_1. By including pastures(which might be source of nitrate leaching if intensively fertilised, as it is the case in some areas) into the land use class representing intensive farming (CLC_AG1) more catchments fall into stratum A (Intensive agriculture).

AGG_1 to AGG_4 are four variants of the main stratification scenario. The intense yellow colour on the map represents catchments characterised by intensive agriculture, brown represents catchments with intensive livestock farming, green represents catchments with weak pressures, pink represents catchments with mixed urban and agricultural land use, red represents catchments characterised as urban, purple represents catchments with dense urbanisation, light yellow represents the ordinary pressures and white represents those which do not fit into any of the categories (strata).

Figure 40 Example of catchment distribution related to small changes I scenario codification



AGG_1:

AGG_2:

AGG_3:

AGG_4:

Appendix 3: Monitoring stations positioning

The statistical analysis has led to discarding a handful of stations or catchments that obviously had been misclassified or had outlying behaviour for reasons not appreciated by stratification (e.g. nitrate emitting factory in a catchment otherwise classified as 'low impacted').

The original data set does not shed light on the presence of stations on the main drain. Stations have been tentatively assigned to drain using their distance to river. The difficulty is that the stations are very accurately positioned on the real rivers that have a certain width (up to more than 1 km!) and that must be assigned to a one-dimensional representation of river. Buffer width checking has been carried out considering the nitrate *vs* stratum component relationship in the case of intense agriculture stratum. The optimum buffer width lies within the range 500–1 050 m; by precaution, 750 m buffer has been selected and all stations within that buffer were assigned to the nearest river.

Table 14 Assessment of optimum buffer size

Parameter	100 m buffer	250 m buffer	500 m buffer	750 m buffer	1 000 m buffer	All stations
Corr. coeff.	0.586981	0.593231	0.601105	0.605511	0.603271	0.577427
R-squared	34.4547 %	35.1922 %	36.1327 %	36.6643 %	36.3936 %	33.3422 %
Intercept/estimate	– 1.54104	– 1.57101	– 1.84033	– 1.86622	– 1.76574	–1.59678
Intercept/std. error	0.153637	0.136834	0.125538	0.12191	0.120939	0.125767
Slope/estimate	27.8372	27.4535	27.5193	27.5403	27.3595	27.5014
Slope/std. error	0.234328	0.209929	0.195237	0.190436	0.189107	0.194867

Appendix 4: Grouping of CLC codes for stratification of nutrients issues

The five predefined CLC aggregates are composed of the 44 different land use classes available in level 3 in the spatial data set of Corine land cover. The Agg_1 is the aggregate used in the current work. For comparison of statistic responses of drivers against determinants the aggregate of Agg_4 was included in the work of this report.

Table 15 List of Corine classes grouped into the five predefined aggregates used for stratification

Code Level 3	Scenario	Agg_1 (final)	Agg_2	Agg_3	Agg_4
111	Continuous urban fabric	CLC_URB	CLC_URB	CLC_URB	CLC_URB
112	Discontinuous urban fabric	CLC_URB	CLC_URB	CLC_URB	CLC_URB
121	Industrial or commercial units	CLC_URB	CLC_URB	CLC_URB	CLC_URB
122	Road and rail networks and associated land	CLC_URB	CLC_URB	CLC_URB	CLC_URB
123	Port areas	CLC_URB	CLC_URB	CLC_URB	CLC_URB
124	Airports	CLC_URB	CLC_URB	CLC_URB	CLC_URB
131	Mineral extraction sites	CLC_URB	CLC_URB	CLC_URB	CLC_URB
132	Dump sites	CLC_URB	CLC_URB	CLC_URB	CLC_URB
133	Construction sites	CLC_URB	CLC_URB	CLC_URB	CLC_URB
141	Green urban areas	CLC_URB	CLC_URB	CLC_URB	CLC_URB
142	Sport and leisure facilities	CLC_URB	CLC_URB	CLC_URB	CLC_URB
211	Non-irrigated arable land	CLC_AG1	CLC_AG1	CLC_AG1	CLC_AG1
212	Permanently irrigated land	CLC_AG1	CLC_AG1	CLC_AG1	CLC_AG1
213	Rice fields	CLC_AG1	CLC_AG1	CLC_AG2	CLC_AG2
221	Vineyards	CLC_AG2	CLC_AG2	CLC_AG2	CLC_AG2
222	Fruit trees and berry plantations	CLC_AG2	CLC_AG2	CLC_AG2	CLC_AG2
223	Olive groves	CLC_AG2	CLC_AG2	CLC_AG2	CLC_AG2
231	Pastures	CLC_AG2	CLC_AG2	CLC_AG1	CLC_AG1
241	Annual crops associated with permanent crops	CLC_AG1	CLC_AG1	CLC_AG1	CLC_AG1
242	Complex cultivation patterns	CLC_AG1	CLC_AG1	CLC_AG2	CLC_AG1
243	Land principally occupied by agriculture, with significant areas of natural vegetation	CLC_AG2	CLC_AG2	CLC_AG2	CLC_AG2
244	Agro-forestry areas	CLC_AG2	CLC_FOR	CLC_AG2	CLC_AG2
311	Broad-leaved forest	CLC_FOR	CLC_FOR	CLC_FOR	CLC_FOR
312	Coniferous forest	CLC_FOR	CLC_FOR	CLC_FOR	CLC_FOR
313	Mixed forest	CLC_FOR	CLC_FOR	CLC_FOR	CLC_FOR
321	Natural grasslands	CLC_FOR	CLC_AG2	CLC_AG2	CLC_AG2
322	Moors and heathland	CLC_AG2	CLC_FOR	CLC_FOR	CLC_FOR
323	Sclerophyllous vegetation	CLC_AG2	CLC_FOR	CLC_FOR	CLC_FOR
324	Transitional woodland-shrub	CLC_AG2	CLC_FOR	CLC_FOR	CLC_FOR
331	Beaches, dunes, sands	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
332	Bare rocks	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
333	Sparsely vegetated areas	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT

Code Level 3	Scenario	Agg_1 (final)	Agg_2	Agg_3	Agg_4
334	Burnt areas	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
335	Glaciers and perpetual snow	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
411	Inland marshes	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
412	Peat bogs	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
421	Salt marshes	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
422	Salines	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
423	Intertidal flats	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
511	Water courses	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
512	Water bodies	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
521	Coastal lagoons	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
522	Estuaries	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT
523	Sea and ocean	CLC_AUT	CLC_AUT	CLC_AUT	CLC_AUT

European Environment Agency

Assessing water quality in Europe using stratification techniques
Results of a prototype application using French data

2007 — 70 pp. — 21 x 29.7 cm

ISBN 978-92-9167-928-7

TH-AK-07-011-EN-N

ISBN 978-92-9167-928-7

European Environment Agency
Kongens Nytorv 6
1050 Copenhagen K
Denmark

Tel.: +45 33 36 71 00
Fax: +45 33 36 71 99

Web: eea.europa.eu
Enquiries: eea.europa.eu/enquiries

Publications Office

European Environment Agency