G.J. Heij and J.W. Erisman (Editors). Acid Rain Research: Do we have enough answers?
© 1995 Elsevier Science BV. All rights reserved.

Reliability of environmental information obtained by modelling and monitoring

J.A. Hoekstra, J.C.H. van Eijkeren, A.L.M. Dekkers, B.J. de Haan, P.S.C. Heuberger, P.H.M. Janssen, A.U.C.J. van Beurden, A.A.M. Kusse, M.J.H. Pastoors

National Institute of Public Health and Environmental Protection, P.O. Box 1, 3720 BA Bilthoven, The Netherlands

Abstract

Environmental information on cause-effect chains is obtained by combining results of measurements and survey data with statistical and process oriented models. For proper use of the information it is important to have a good understanding of its reliability. This paper addresses the problem of assessing the reliability of information obtained from different models and measurements. A reliability factor is introduced for standardized quantification of reliability. The approach is illustrated by results of a so-called "nitrogen chain", starting with estimates of nitrogen emission and ending with predicted future concentrations in untreated water at drinking water stations.

1. INTRODUCTION

Environmental information is often a composition of results from different sources. This certainly holds for the information provided by the National Institute of Public Health and Environmental Protection (e.g. RIVM, 1994). The information includes amounts of pollutants emitted, concentrations in air, water and soil, and related effects on human and ecosystem health. It also involves predictions of future emissions, concentrations and effects. The data are collected using monitoring networks, surveys and process oriented models, and constitute an important ingredient in the support of the policy-making process of the Dutch government. Insight in the reliability of the information would be an extremely valuable addition. This insight will also help in optimizing the monitoring networks and modelling activities.

This paper discusses the problem of assessing the reliability of information obtained from different types of models and measurements. The problem will be illustrated by a practical case, a so-called "nitrogen chain", starting with estimates of nitrogen emission and ending with predicted future concentrations in untreated water at drinking water stations. Section 2 gives an overview of the studied system. In section 3 a reliability factor is defined for quantifying the reliability of the data in an uniform way. In section 4 it is indicated how the reliability factor can be calculated. Section 5 presents some results for the nitrogen chain, which are discussed in section 6.

2. THE NITROGEN CHAIN

Nitrogen is emitted into the atmosphere by traffic, industry and intensive animal husbandry. After transport, these pollutants are deposited to natural soils, farm lands and urban areas. Added to the directly applied fertilization on farm lands, the excess nitrates percolate to the groundwater. High nitrate concentrations in groundwater threaten its exploitation for drinking water production.

Figure 1 shows the data production process related to this causal chain in a simplified form. Grey boxes refer to monitoring networks and surveys, open boxes to mathematical models. Arrows denote physical processes like transport and deposition. Bold arrows indicate that the reliability of the models simulating such processes is investigated in this study. Broken arrows mark processes that have not been considered here. Figure 1 shows that reduced nitrogen concentration data result from a model with emission survey data on animal husbandry as a main input (Asman and Van Jaarsveld, 1990), while oxidized nitrogen concentration data stem from a monitoring network (Erisman, 1992).

In the first column emission registration data are given. E.g. for animal husbandry, the data include number and type of animals per farm and stable type. Based on these data estimates of NH_3-emission are obtained for each municipality in the Netherlands (Van der Hoek, 1994). These emission data are reallocated to a 5x5 km^2 grid. The reallocation scheme draws on a high resolution land use database.

Ambient air concentrations, second column, are processed into dry deposition data using a model (Erisman, 1992). Wet deposition is obtained from a monitoring network. The nitrate load to the groundwater below urbanized regions and natural vegetation soils, third column, is also obtained by measurements. The thin arrow pointing to the box 'natural soils' symbolises the actual use of the 1986 deposition data in a regression model with nitrate concentration in freatic groundwater as response variable (Boumans, 1994). Finally, the loads to freatic groundwater, fourth column, feed a model describing the transport to the inlet of the drinking water production plants (Lieste, 1989).

In summary, this study involves the following types of data and models:

- emission registration data per municipality
- a model to reallocate these data to a grid at the appropriate scale
- monitoring networks in combination with interpolation schemes
- regression models obtained from case studies
- process oriented models describing transport, transformation and deposition
- land use data, meteorological data, parameter values for process models c.q. interpolation schemes, etc.

This nitrogen chain consists of several chains of models and monitoring systems starting with the actual polluting activities. Most important to the quality of untreated water is the short chain starting with agricultural fertilization loads.

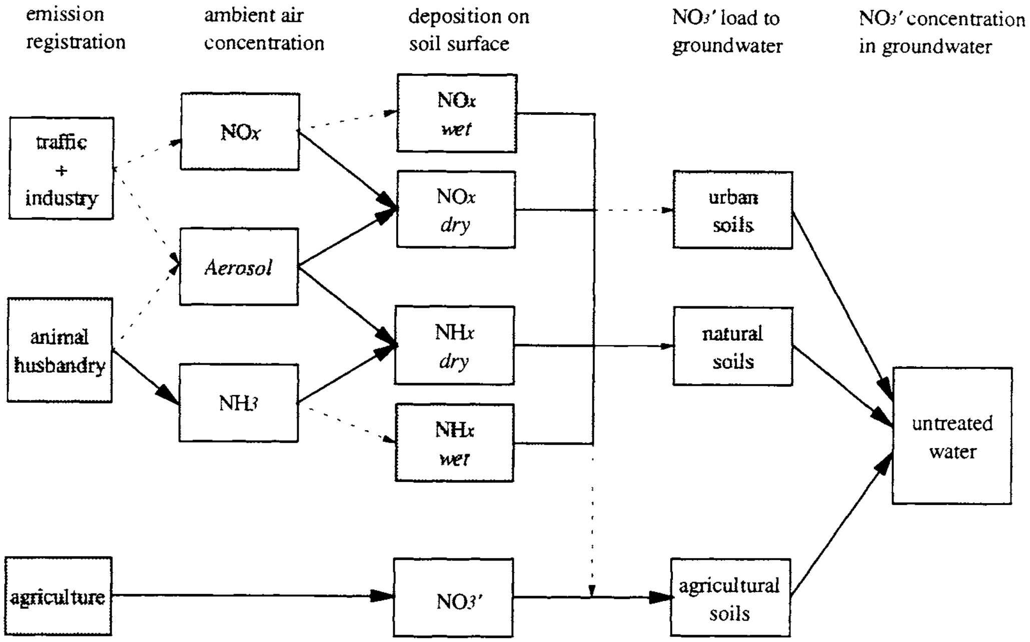

Figure 1 The data production process.
Grey boxes: obtained from monitoring networks and surveys.
Open boxes: data from process oriented models.
Bold arrows indicate for which models reliability analyses have been carried out.

3. THE RELIABILITY FACTOR

A standardized measure of reliability is helpful to obtain a complete picture of the quality of information rendered by data and models.

Estimated and "true" value

Suppose y^0 is the value or set of values for which the reliability is to be assessed. Examples of y^0 are:

- estimated average N-deposition (mol/ha) on 5x5 km^2 in a given year
- estimated total N-input into soil by agriculture (million kg) in a given year
- predicted nitrate concentration (mol/l) at a given drinking water station in 2100.

The symbol y^* represents the corresponding unknown "true" value. Intuitively, reliability should measure the nearness between y^0 and y^*. The concept of reliability is only meaningful if the true value is clearly defined. This implies that one should be specific about the intended temporal and spatial representativity of the information.

When a value in the distant future is presented, it is difficult to define precisely what is meant by the "true" value. Scientists generally do not have tools to predict e.g. the actual N-concentrations in 2100. Therefore, the "true" value meant in the prediction procedure is not what the N-concentration will really be in 2100. Instead, the prediction is calculated under specific and limited model assumptions, assuming "all other things being equal" (which we know not to be the case). A reliability analysis therefore has equally limited scope; it should specify clearly which parameters and input are varied and which are kept constant. Such limited predictions can still be very useful. The calculated predictions will render important information by comparing the influence of different environmental strategies on the future state of the environment.

Systematic and random error

The value y^0 is considered to be the outcome of a stochastic process Y, which could generate a number of outcomes each with a given probability. For instance, Y represents a measurement process with error or a computer model with imprecise knowledge of the parameter values and/or stochastic input. Figure 2 shows two examples of distributions of Y, together with the true value y^*. In Figure 2a the process generates outcomes with a small variance and a large systematic error or bias, as y^* is not in the centre of the distribution.

In Figure 2b, the systematic error is much smaller, but the random error is larger.

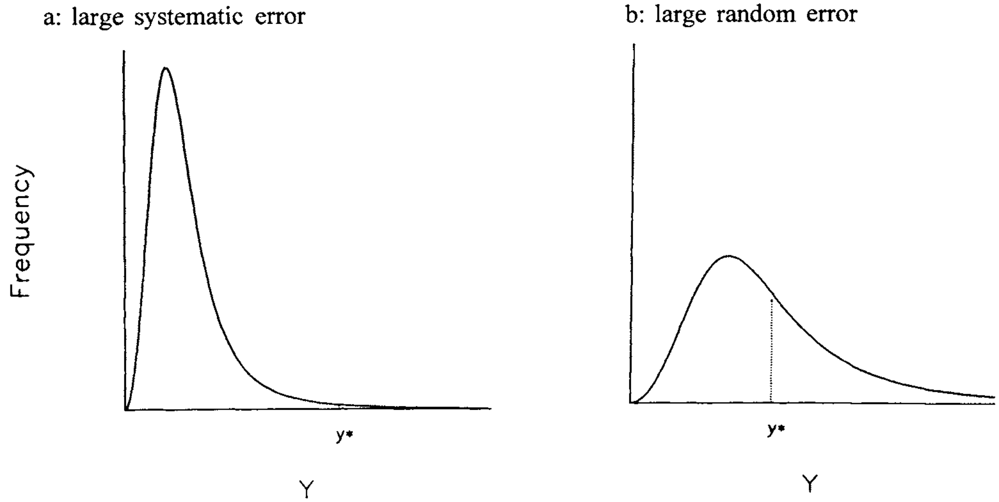

Figure 2 Frequency distribution of Y, with "true" value y^*

The reliability factor

Mostly, the distribution of Y is skew to the right, as in Figure 2, and takes on only positive values. In this case a relative error is an appropriate measure of variation. Moreover, relative errors are dimensionless and more easy to compare for different variables than absolute errors. The reliability factor is therefore based on the geometric standard error:

$$
\begin{aligned}
s^{geometric} &= \exp \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\log(y_i^0/y^*)\right)^2} \\
&= \exp \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\log y_i^0 - \overline{(\log y^0)}\right)^2 + \left(\overline{(\log y^0)} - \log y^*\right)^2} \qquad (1) \\
&= \exp \sqrt{variance(\log y^0) + bias(\log y^0)^2}
\end{aligned}
$$

where $y_1^0, y_2^0, \ldots, y_n^0$ represent realisations of Y, e.g. measured or computed values, and $\overline{\log y^0} = \frac{1}{n}\sum \log y_i^0$.

The reliability factor is defined as:

$$B(y^0) = 1/s^{geometric} \qquad (2)$$

$B(y^0)$ takes on values between 0 (extremely unreliable, value completely unknown) and 1 (exactly known). It is closely related to the reliability index proposed by Leggett and Williams (1990) for comparison of model predictions with measurements, and to Kirkwood's geometric measure of dispersion for comparing variables with different dimensions (Kirkwood, 1988). If Y is approximately lognormal and contains no systematic error, the factor B can be used to construct a confidence interval for y^*:

68% Confidence Interval: By^0 - y^0/B

95% Confidence Interval: B^2y^0 - y^0/B^2

4. METHOD OF CALCULATION

Usually y^* is unknown. In that case, only the variance component of the error can be assessed.

If y^0 is obtained by application of a statistical method to measurement data, the variance is given by the statistical procedure. E.g., a field of nitrogen oxide concentrations is obtained by applying Kriging to results of a monitoring network, and the Kriging error model provides standard errors for each grid cell (Cressie, 1991). The model assumptions can to some extent be checked using residual plots and cross-validation. If the measurements are unbiased and have the correct spatial and temporal representativity, it can be assumed that the systematic error is small in comparison with the random error. Especially the second condition: spatial and temporal representativity is difficult to fulfil.

To calculate the reliability factor for outcomes of process models, two different approaches can be followed:

1. "External" examination by comparison with real measurement data.
   Sometimes good quality measurements are available to validate the model. The measurements can then be used as a substitute for the true value y^* in the above formula. The reliability factor will contain both variance and bias.
2. "Internal" examination using Monte Carlo methods as implemented in the simulation package UNCSAM (Janssen et al., 1992).
   A set of outcomes is generated using a probability distribution for uncertain model parameters, and input (error propagation). These data are used to calculate the variance of the outcome of the process model. The nominal value, y^0_{nom}, is usually not in the centre of the distribution of outcomes. The bias part in the formula above is replaced by $(\log y^0_{nom} - \overline{\log y^0})^2$.
   Evidently, expert judgement plays an important role in specifying the probability distributions of the parameters (including correlations). Therefore, the internal reliability factor contains an expert judgement element as well.

The two approaches produce complementary information as is exemplified later.

5. PRELIMINARY RESULTS

Table 1 gives reliability factors for some elements of the nitrogen chain described in section 2. The analysis is in a preliminary stage. Therefore, the results are rough estimates, presented only to illustrate the discussion. More detailed results will be given in Hoekstra and Heuberger (in prep).

Reliability factors were calculated for each grid cell separately. The presented values in the table are the medians of the reliability per cell; large fluctuations around this median reliability do exist.

6. DISCUSSION

The reliability factors presented in Table 1 should be interpreted with care, since their values heavily rely on subjective choices concerning parameter uncertainties, neglecting potential systematic errors due to erroneous model specifications etc. Improvements are being made, by considering more parameter uncertainties and carrying out more model checks. However, the intrinsic problem: that frequently the "true" values are not known and therefore formula 1 is applied with substitutes of y^*, cannot be solved. Therefore, it should be realized and accepted that reliability factors can have a considerable uncertainty themselves too.

Table 1
Reliability factors for elements of a nitrogen chain (annual averages). Preliminary results.

	Type of Data	Uncertainty Analysis	*B*
Emissions			
N-load on arable soil (500^2 m^2)	process model	internal	0.8
NH_3-emissions to air (5^2 km^2)	process model	internal	0.8*
Air			
NH_3-concentrations (5^2 km^2)	process model	internal	0.8
		external	0.6
NH_x-depositions (5^2 km^2)	process model	internal	0.6
NO_2-concentrations (5^2 km^2)	monitoring network	internal	0.9
NO_y-depositions (5^2 km^2)	process model	internal	0.8
Soil/groundwater			
Nitrate in freatic groundwater (500^2 m^2)			
nature	field measurements	internal	0.4
agriculture	process model	internal	0.5
		external	0.7
Predicted value at a drinking water station in 2100			
Nitrate in untreated water	process model	internal	0.6

* Partially based on expert judgement. Other estimate: 0.6 (Erisman, 1992)

This being stated, what can be concluded from the data in Table 1? First of all, the provided environmental information can be highly uncertain. Some variables have a reliability factor of 0.5, which means that the variable is not known within a factor 2. A 95% confidence interval would even run from $y^0/4$ to y^0*4. Considering the small differences between target and limit values for environmental quality parameters, this is a large uncertainty.

One should not be surprised by the large uncertainty: in several cases the local predictions in Table 1 are obtained by models with parameters that are held constant over large regions out of need rather than truth. This raises the question whether the models are appropriate for this type of prediction, and vice versa, whether information on such a fine scale is really necessary. The NO_2-concentrations in air have higher reliability, because this variable exhibits comparatively modest local variation.

Comparison of internal and external reliability factors is interesting from a validation point of view. The discrepancy between the two reliability factors for NH_3-concentrations in air, indicates that uncertainty on input and parameters as presently considered feasible, is not sufficient to explain the limited performance of the process model in matching the NH_3-measurements.

Further analyses will be carried out to study spatial variation in reliability and to include external reliability factors and new environmental variables.

7. REFERENCES

Asman, W.A.H. and Jaarsveld, J.A. van (1990). A variable-resolution statistical transport model applied for ammonia and ammonium. RIVM-report no. 228471007, Bilthoven, The Netherlands

Boumans, L.J.M. (1994). Nitrate in freatic groundwater under natural area's on sandy soils in the Netherlands. RIVM-report no. 712300002, Bilthoven, The Netherlands. (In Dutch).

Cressie, N.A.C. (1991). Statistics for spatial data. John Wiley & Sons, New York.

Drecht, G. van, (1993). Modelling of regional scale nitrate leaching from agricultural soils in the Netherlands. *Applied Geochemistry*, Suppl. Issue, no. 2, 175-178.

Erisman, J.W. (1992). Atmospheric deposition of acidifying compounds in The Netherlands. Doctoral Thesis, University of Utrecht, The Netherlands.

Hoek, K.W. van der, (1994). Calculation method for ammonisa emission in the Netherlands in the years 1990, 1991 and 1992. RIVM-report no. 773004003, Bilthoven, The Netherlands. (In Dutch).

Hoekstra, J.A. and Heuberger, P.S.C. (Eds) (in prep). The reliability of environmental information: an analysis of the nitrogen-chain. RIVM-report in preparation, Bilthoven, The Netherlands. (In Dutch).

Janssen, P.H.M., Heuberger, P.S.C. and Sanders, R. (1992). UNCSAM 1.1: A software package for sensitivity and uncertainty analysis: Manual. RIVM-report no. 959101004, Bilthoven, The Netherlands.

Leggett, R. and Williams, L.R. (1990). A reliability index for models. *Ecological Modelling* **13**, 303-312.

Lieste, R. (1989). Computer Program FLOPZ1. Path lines in quasi-three-dimensional groundwater flow in a layered homogeneous aquifer. RIVM-report no. 728520004, Bilthoven, The Netherlands.

Kirkwood, B.R. (1988). Essentials of medical statistics. Blackwell Scientific Publications, Oxford.

RIVM (1994). National Environmental Outlook 3 1993-2015. Samson H.D., Tjeenk Willink, Alphen aan de Rijn.