

Towards a generalised linear model of groundwater pollution by pesticides

F Worrall

*Dept. of Geological Sciences, The University of Durham, Science Laboratories,
South Road, Durham, DH1 3LE, UK*

ABSTRACT

This study has brought together extensive groundwater monitoring datasets from the UK and the US to examine the underlying controls on the causes of groundwater pollution by pesticides. The study examined firstly the role of chemical properties in controlling occurrence in groundwater, secondly, the role of the properties of the site of application, and finally, brought these factors together to test their relative importance and propose combined models based on both site and chemical properties. The study: (1) developed a model of the occurrence of pesticides in groundwater based on molecular topology; (2) developed a method of calculating groundwater vulnerability, independent of compound type, directly from borehole observations; (3) showed that groundwater vulnerability in agricultural catchments was governed by soil and hydrologic factors, but was independent of land-use; and (4) showed that both chemical and site factors have an independent and significant effect on groundwater contamination, but that the interaction of these factors is the important control. These methods are leading to a generalised linear model of groundwater pollution by pesticides that brings with it new opportunities in risk assessment and risk management.

INTRODUCTION

The problem of pesticide pollution of groundwater has meant that a range of techniques has been applied to understanding and predicting this problem. The properties that control the pollution can be broadly classified into two groups: site and chemical factors. A number of methods have been developed using either site or chemical factors.

The site factor includes such properties as land-use, soil, aquifer and climate. The combination of site factors that give rise to pollution have been taken together as groundwater vulnerability tools. Groundwater vulnerability has slightly varying definitions. Palmer *et al.*, (1995) recognised that differing soil and hydrogeological conditions will give rise to differing vulnerabilities and afford different degrees of protection to the underlying aquifer. It is important to note that this concept of vulnerability is independent of the nature of the pollutant. A range of vulnerability tools have been developed, e.g. DRASTIC (Aller *et al.*, 1987). Other research has extended the range of parameters included and the concept has been taken forward to develop regional vulnerability maps (e.g. Palmer *et al.*, 1995).

These systems have a number of flaws. Firstly, the inclusion or exclusion of variables into the vulnerability assessment systems is often arbitrary and based solely on expert opinion as to the weighting between factors (e.g. Aller *et al.*, 1987). Secondly, the

indices are typically not based on observations or measurements of groundwater contamination and even when physically-based models are used they are prone to errors in the model assumptions or in choosing input parameters. Thirdly, these schemes have rarely been validated or tested against observational data (Merchant, 1994). When validation is performed evidence can be contradictory, eg. for the DRASTIC system (Maas *et al.*, 1995). More fundamentally there is an underlying assumption that that is it possible that the variation observed in the occurrences of pesticides is due only to variation in soil or climatic conditions without reference to the properties of the contaminant concerned.

Conversely, the chemical properties of contaminants alone have been used to assess the risk of groundwater pollution. Such screening methods have most frequently based their judgements on adsorption, degradation or solubility parameters (e.g. Gustafson, 1989). As for vulnerability assessment systems these methods based solely on chemical properties have tended to develop scores and indices based on expert opinion of the weighting of chemical parameters. Even those methods based on observations of groundwater contamination, rather than *a priori* combinations of parameters (e.g. Gustafson, 1989), have been shown to be inapplicable outside the region for which they were calculated (Wooff *et al.*, 1999). The classification of compounds into polluting/non-polluting or mobility classes based on such schemes is prone to error when allowance is made for the natural variation in the parameters they use (e.g. K_{oc} and soil half-life). It has, however, been possible to show that despite the large site-to-site variation in such parameters as K_{oc} and degradation half-life it is possible to differentiate those compounds that are found in groundwater from those that are never observed (Worrall *et al.*, 2000) with this separation being statistically significant. Proving that it is possible to differentiate polluting from non-polluting compounds shows that it is viable to base contamination screening methods solely on chemical properties. However, the importance of the role of chemical properties relative to site properties in controlling the transport of contaminants is not known. This paper takes each of these factors separately, the site and chemical, and then examines their relative importance.

THE CHEMICAL FACTOR

Given the site variability in K_{oc} and half-life it necessary to use other properties to explore the chemicals role in controlling groundwater pollution. Worrall (2001) using logistic regression differentiated compounds found in groundwater from those known to be used in the surveyed regions but not found in the groundwater on the basis of the molecular connectivity parameters. Molecular connectivity parameters are dependent only on the structure of the compound and not variable from site to site. The model was developed for 47 compounds from a groundwater in the midwest US and tested against a study from California of 41 compounds. The result correctly classified 96% of the data, and 91% when validated against the Californian data. The result can be visualised if only the two most important parameters are considered (Figure 1). Considering the most important parameter, 86% of the data was correctly classified by a simple inequality:

$${}^6\chi_p^v < 0.5 \quad (i)$$

Where: $6\chi_p^v$ = the sixth-order path connectivity.

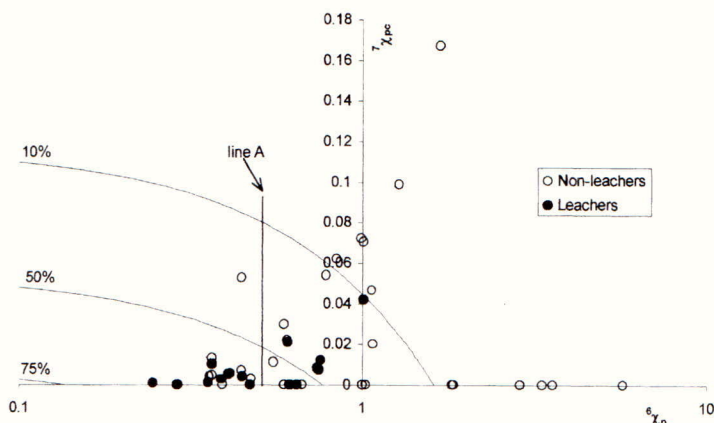


Figure 1. Plot of the sixth-order path connectivity ($6\chi_p^v$) vs. seventh-order path-cluster connectivity ($7\chi_{pc}^v$) in comparison to the best-fit probability of the compound being a leacher. Line A represents Eqn (i).

On Figure 1 it is possible to discern two trends in the non-leacher compounds, one in increasing $6\chi_p^v$. This trend is represented by linear compounds of a variety of pesticide classes increasing in chain length suggesting that it represents control by molecular size. The second trend is one of both increasing $6\chi_p^v$ and $7\chi_{pc}^v$ which is represented by cyclic compounds from a range of pesticide classes. This trend highlights that the complexity of branching (represented by $7\chi_{pc}^v$) is important.

THE SITE FACTOR

How can we estimate the effect of the site of application upon its fate?. If we consider an example from a large groundwater monitoring programme in California. Over a period of several years atrazine was analysed for 1791 times in water from a large number of boreholes throughout the Great Valley of California. Atrazine was found in 165 of them – a proportion of 9%. In one particular borehole, over the same time period, atrazine was tested for 24 times and found 11 times – a proportion of 45%. The difference between these two proportions represents the vulnerability of the borehole compared to the regional average. The method of Worrall (in press) works on this basis and presents a method for firstly showing that the difference between two proportions is statistically significant, secondly to show how to combine information from a range of compounds, and thirdly, to show how to convert the difference between proportions into the probability of finding the next compound that is monitored for. The method developed in Worrall (in press) was for multiple observations of single compounds. The method was extended to work with single observations of multiple compounds and applied to an extensive survey of pesticide occurrence in shallow aquifers of the Midwest US (Worrall & Kolpin, in press – Figure 2). When the data is of sufficient spatial density it is possible

to map the vulnerability of groundwater as a probability of finding the next compound to be monitored for.

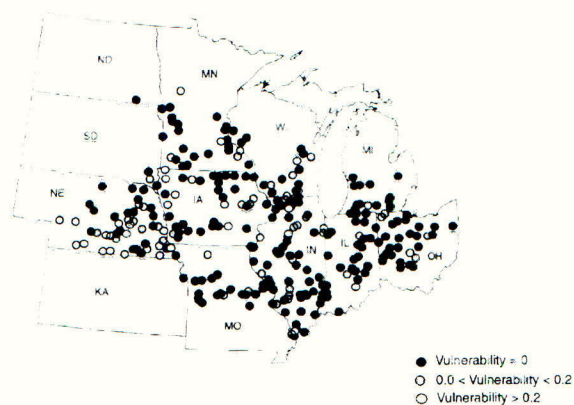


Figure 2. Distribution of vulnerability for a 303 borehole study of the midwest USA

This system has a range of advantages over present vulnerability assessment techniques. The method directly estimates vulnerability from borehole data, it does not rely on expert opinion of combination of variables that are presumed to be important in controlling groundwater pollution. The method predicts a probability of groundwater pollution that is independent of compound type and is not just a relative indice. Because the method is developed within a Bayesian statistical framework it combines information from a range of sources and can easily be updated as new information becomes available. The estimate of vulnerability is calculated independently of the properties of the borehole catchment and as such can be correlated with properties of that catchment to understand what controls leaching.

COMPARING SITE & CHEMICAL FACTORS

Previous sections have explored each factor separately from the other, but as outlined in the introduction groundwater pollution should be dependent on both factors and so it is important to compare their relative roles to discover whether each factor is important when compared to each other; is one factor more important than the other? For example, is the distribution of pesticides in groundwater related more to the variation of properties between sites or by the variation in the properties of compounds being applied in that region? Equally, we need to assess the importance the interaction of these factors, i.e. their dependence upon each other. Taking two extensive groundwater surveys one in the midwest USA and one in Southern England general linear modelling was used as a means of comparing variation between compounds and between boreholes. For each study it was also possible to compare results between consecutive years. The use of a

general linear modelling approach means that not only can the significance of a factor be assessed but also its interactions.

Results showed that neither in the US nor UK study was the difference between consecutive years significant (at the 95% level). However, for one of the years in each of the two surveys it was possible to test the difference within a year in comparison to the differences between compounds and sites. In each of the national studies this within-year factor was statistically significant suggesting that seasonality is more important than variations in climate that could occur between years. In both the UK and US studies both the site and compound factors were significant (at the 95% level) on their own and independent of each other (Table 1). However, the percentage of the variance explained by these factors alone was very small, especially in comparison to the percentage of variance explained by the interaction of these two factors.

Table 1. Percentage of the variance of a dataset explained by each of the factors considered in each of the groundwater surveys.

Dataset	USA	UK
Source of Variance	% of variance	% of variance
Site	14	13
Chemical	5	4
Interaction	65	67
Measurement error	16	16

This is the first time that any statistical validity for groundwater vulnerability tools has been tested, i.e. showed that there site properties are significant independent of chemical properties or between and within-year variation, but that this effect is small. Equally, this result confirms that of Worrall *et al.*, (2000) that there is a significant difference between leachers and non-leachers. The greatest proportion of the variance is explained by the interaction between site and chemical factors. The interaction of site and chemical factors can be considered in two ways. First, this is part of processes known to effect all pesticides, e.g. adsorption is considered as a combination of a compound's innate properties and the fraction of organic matter in the soil. Second, that the significance of the interactions shows that a site's vulnerability to contamination is compound specific. These findings could be limited by the range of compounds and conditions included in these surveys, but the two studies in very different settings, and examining compounds with a considerable range of chemical properties, came to very similar results. The measurement error is indicative of the fit of the model to the data. The magnitude of the measurement error in these studies is of the same order of magnitude as the analytical detection limit used so suggesting additional factors need not be included.

CONCLUSIONS

This study has examined pesticide pollution to groundwater as a combination of chemical and site factors. The study has shown that:

- i) leaching compounds can be distinguished from non-leaching compounds on the basis of molecular descriptors;
- ii) the important molecular properties are both molecular size and the molecules complexity of branching;
- iii) groundwater vulnerability can be calculated directly from borehole observations;
- iv) both site and chemical factors have a significant effect on the occurrence of pesticides in groundwater with the former being the most important; and
- v) the most important control on pesticide pollution of groundwater is the interaction of site and chemical factors.

The approach taken by this study and the results shown have important implications for both the risk assessment and risk management of pesticides.

ACKNOWLEDGEMENTS

The author is grateful to D Kolpin (United States Geological Survey), J Troiano and B Johnson (California Dept. of Pesticide Regulation) and T Besien (UK Environment Agency) for supplying the data upon which the work is based.

REFERENCES

- Aller L; Bennett T; Lehr J H; Petty R J; Hackett G (1987). *DRASTIC: A standardised system for evaluating groundwater pollution potential using hydrogeologic settings*. U.S. Environment Protection Agency, Ada, Oklahoma, EPA/600/2-87/035.
- Gustafson D I (1989). Groundwater ubiquity score: a simple method for assessing pesticide leachability. *Environmental and Toxicological Chemistry* **8**: 339-357.
- Maas R P; Kucken D J; Patch S C; Peek B T; Van Engelen D L (1995). Pesticides in eastern North Carolina rural supply wells: land-use factors and persistence. *Journal of Environmental Quality* **24**: 426-431.
- Merchant J W (1994). GIS-based groundwater pollution hazard assessment – a critical review of the DRASTIC model. *Photogrammetric Eng. Remote Sensing*, **60**: 1117-1128.
- Palmer R C; Holman I P; Robins N S; Lewis M A (1995). *Guide to groundwater vulnerability mapping in England and Wales*. National Rivers Authority, Bristol.
- Worrall F; Wooff D; Seheult A H; Coolen F (1998). A simple Bayesian approach to the interpretation of environmental fate and behaviour data for pesticide registration. *Pesticide Science*, **54**(2): 99-112.
- Wooff D A; Seheult A H; Coolen F; Worrall F (1999). A Bayesian reappraisal of environmental fate and behaviour data used in pesticide registration. In *Statistics for the Environment: 4*, Barnett and Turkman (eds.), John Wiley & Sons.
- Worrall F; Wooff D; Seheult A H; Coolen F (2000). New approaches to the assessment of environmental fate of pesticides. *Jour. Geol. Soc.*, **157**: 877-884.
- Worrall F (2001). A molecular topology approach to predicting pesticide pollution of groundwater. *Environ. Sci. Tech.* **35**: 2282-2287.
- Worrall F Direct assessment of groundwater vulnerability from borehole observations. In *Sustainable Groundwater Development*, Hiscock, Davison and Rivett (eds.), Geol. Soc. of London spec. publ. (*in press*).