Modeling non-point source pollutants in soil: Applications to the leaching and accumulation of pesticides and cadmium

Tiktak, A.

Citation for published version (APA):
The general objective of this thesis was to contribute to a justified application of models of non-point source pollutants on a regional-scale. Two relevant indicators of environmental quality were studied, i.e. the concentration of pesticides in the shallow groundwater and heavy-metal (Cd) contents in the topsoil. These two indicators were chosen, because they required different modeling approaches. This epilogue reviews the major findings and gives recommendations for future research directions. It does so by focusing upon the five major topics described in the general introduction. To facilitate the discussion, the order of the five topics has been changed.

**The Research Chain**

In Chapter 1 a general procedure was outlined for the development of regional-scale models of non-point source pollutants in soils. In this procedure, the model is first applied to the plot-scale where most model-inputs can be derived from direct measurements and where data to evaluate the model are available as a function of time. Model and data are then aggregated and applied to the regional-scale. However, due to the nature of the available data and the temporal and spatial dynamics of the system to be modeled, exact application of this procedure is not always possible. This is illustrated in Figure 1 of this chapter, which shows the actual research chain that was followed in the two modeling studies discussed in this thesis (see also Figure 1 in Chapter 1).

The pesticide leaching study started with the application of the PESTRAS model to the field-plot described in Chapters 2.1 and 2.2 (Figure 1a, Step 1). This field-study and earlier field-studies (Boesten, 1986; Boekhold et al., 1993) offer the opportunity of improving our knowledge of processes that control pesticide leaching. However, regional-scale monitoring data, which are required to evaluate whether spatial patterns predicted by the model match real-world geographical patterns, are virtually non-existing. The primary reason is that the concentration of pesticides in the shallow groundwater shows
strong temporal dynamics. Combined with the high costs associated with pesticide analyses and the large number of pesticides admitted, it becomes clear that true regional-scale monitoring of pesticides is not feasible, and that the fourth step of the ‘ideal’ research chain could only partly be carried out (Figure 1a; Step 2). Statements about the performance of PESTRAS on a regional-scale were therefore limited to qualitative remarks, such as ‘Pesticides were only found in the groundwater in regions that we classified vulnerable’ (Chapter 3.1).

If true regional-scale monitoring is not possible, the only alternative is to apply the model to a broad range of field-experiments. There are, however, important pitfalls associated with this approach. The first is that the field-experiments do not necessarily cover the full range of conditions encountered by the regional-scale model. In such cases, processes that are important at part of the area to be simulated may be overlooked. The effect of preferential-flow on pesticide leaching, for example, may be overlooked if the model is not applied to cracking soils. The second problem associated with this approach is that
it does not account for the effect of data-aggregation and the use of generic model-inputs. This was illustrated in Chapter 2.1., which showed that reasonable predictions of pesticide leaching could only be obtained with site-specific data. It was concluded that evaluation of a plot-scale model with generic data is an important step if the model is intended for regional-scale applications. It is also interesting to quantify the overall model performance for a number of field-sites at once, instead of employing the usual one-to-one field-tests (Loague and Corwin, 1996). In this methodology, the observations are treated simultaneously and common values for field-specific data are optimized using inverse modeling.

Heavy-metal accumulation is a slow process occurring over decades. As long-lasting field-studies were not available and short field-studies were not appropriate to detect long-term changes with time, we decided to directly apply a simple model of heavy-metal accumulation to the regional-scale (Chapter 3.2; Figure 1b, Step 1). For model evaluation, a data set was available with heavy-metal contents measured at 2544 locations. The use of this dataset was justified by the fact that heavy-metals are strongly sorbing compounds exhibiting limited variability at short temporal scales. When comparing the model results with the data we were confronted with discrepancies: In soils under natural land-use, the predicted cadmium contents were much lower than the observed cadmium contents. As it was not possible to find an explanation for this underestimation using the model and data at hand, it was decided to carry out an additional modeling study (Chapter 2.3; Figure 1b, Step 2). In this study, a more detailed model of heavy-metals in soil was applied to three fields. Important conclusions from this field-study were (i) bio-cycling plays an important role in natural soils for maintaining high cadmium levels in the ecto-organic soil layer, and (ii) the generic sorption isotherms used in the model do not apply to the ecto-organic layer. This indicates that the use of a one-layer model cannot be justified if separate information on heavy-metals in the ecto-organic soil layer is wanted.

These examples show that modeling is an on-going process: Model adaptation, data-collection or even basic research may be required if the final model is considered insufficient (Bouma, 1998).
Switching Spatial Scales

In view of the dilemma between model detail and regional-scale applicability (Figure 2 in chapter 1), models are often simplified and data aggregated during scale transfer (Chapter 1; De Vries et al., 1998). In this thesis, two approaches were followed: In the pesticide leaching assessment, a comprehensive model was directly applied to the regional-scale. This required, however, strong spatial aggregation of the input data ('mega-plot approach; see Chapter 3.1). In the heavy-metal accumulation assessment the model was simplified, allowing a higher spatial resolution.

The most commonly used model simplification is temporal aggregation (e.g. De Vries et al., 1998). Generally, the ignorance of seasonal dynamics becomes more justified if the mobility of a substance decreases (Chapters 3.1 and 4.2). Transient-state soil-water flow was required for the pesticide leaching assessment and steady-state hydrology was sufficient for the heavy-metal accumulation assessment. It should be noted, however, that this remark only applies to the soil-chemical subsystem. The calculation of soil-water fluxes itself required a high temporal resolution, as shown by Tiktak and Bouten (1992). They found that the long-term average precipitation surplus was in some cases completely overestimated if average annual rainfall records were input to the model. This implies the coupling of submodels with different temporal resolutions, an opportunity given by modern simulation environments (Tiktak et al., 1994a).

All models described in this thesis apply to a single point-in-space. The spatial presentation resolution for the regional-scale assessments, however, was at least 500x500 m², so some kind of data-aggregation is required. True upscaling from point-support to larger scales requires running the model for all points within a block, after which the model outputs are aggregated (Bierkens and Gaast, 1998; Heuvelink and Pebesma, 1998). It is clear that this procedure is generally not feasible for regional-scale assessments. We therefore first aggregated the model-inputs to the block-scale (Chapters 3.1 and 3.2), after which we ran the model for one point within the block. This single point is assumed representative for the block as a whole. Further examination of both procedures, which comes down to the classical 'calculate first and average later' vs. 'average first
and calculate later’ problem, remains an item for further research.

If, finally, the model calculations have been made, the question remains how to compare a regional-scale dataset with the model results. Summary statistics like median values and percentiles have often been used as criteria (De Vries, 1994b), but the underlying spatial pattern is then ignored. An alternative was presented in Chapters 3.2 and 4.2, where results from a regional-scale model of heavy-metal accumulation were compared with 2544 point observations of heavy-metal contents. The major problem encountered was that there were more model blocks than observations. Therefore, the observations were interpolated onto the model grid. The fairest comparison between the model outputs and the interpolated image would require the calculation of grid cell averages. This required a model-based approach, which is commonly known as block-kriging (Deutsch and Journel, 1992). Although block-kriging can potentially produce accurate predictions of block-averages, the accuracy is built on the presumption that the variable of interest satisfies an assumed spatial correlation model. For this reason, a Generalized Additive Model was used for spatial interpolation. The most attractive aspect of the GAM is that it is based upon the data, instead of being model-based.

Model Validation

The models presented in this thesis were aimed at supporting decision makers by predicting non-point source pollution as a function of different socio-economic scenarios or by predicting non-point source pollution at geographic locations lacking observations. Good Modeling Practice requires that a model is validated before it is used for prediction (Vanclooster et al., 1999).

Generally speaking, a model is called ‘validated’ if it is proven to be capable of predicting the behavior of a system under different conditions (Bouten, 1992). Model validation applies to both the conceptual part of the model, and to its parameterization. As shown by several authors, it is not possible to validate a model in this way (Bouten, 1992; Leijnse and Hassanizadeh, 1994; Addiscott, 1998). All that can be done is to compare the model results with measured values (Chapters 2.1,
Model ‘evaluation’ or some similar term may therefore be preferable to ‘validation’ (Addiscott, 1998).

Whether predictions fit the data remains a subjective judgement. To overcome this problem, the use of statistical performance measurements was advocated (Chapters 2.1, 2.2, 3.2 and 4.2). In Chapter 3.2 and 4.2 the factor-of-\( f \) approach (Parrish and Smith, 1990) was introduced, which acknowledges that both the model predictions and the observations contain uncertainty. However, despite its advantages like clearness and simplicity, its limitations should be kept in mind. First, the choice of the factor-of-\( f \) remains subjective, so it should always be reported. Secondly, field-variability of certain variables may be so large that any model would fit the observations if variability was regarded as pure error. Another point of concern is that most performance statistics that are currently used do not consider the spatial or temporal pattern. It would be very welcome if one could quantify the quality of the prediction of the geographical and temporal pattern.

Model evaluation becomes more difficult as the scale of use moves from the plot-scale to the regional-scale (Chapter 3.2). The main problem was that of obtaining the proper data against which to test the model. The model-blocks were larger than the support for an individual observation, so they had to be translated to the model-scale before the actual comparison could be made. This, however, attached additional uncertainty to the observations, making the model evaluation statistically less unique. Also, it was difficult to evaluate the included process descriptions. It turned out to be necessary to go back to the plot-scale (Chapter 2.3).

Also at the plot-scale, however, no absolute guarantee can be given that the process descriptions included in the model provide a realistic description of the system to be modeled. Many classical examples exist of different process descriptions being able to reproduce the observations. Bouten (1992) showed that forest transpiration dynamics could be well described with two models that contained entirely different descriptions of the mechanisms of stomatal control. Van Grinsven et al. (1995) compared various soil acidification models and concluded that no proof could be given that any of the models performed better. They also concluded that even at the plot-scale not enough experimental data were available for evaluation, so that the
identifiability of the model parameters was limited. In Chapter 2.2 the same conclusions were drawn for nine pesticide leaching models. It was further concluded that evaluation of the included process-descriptions was hampered by the fact that no single model was parameterized in the same way. An important reason was that the interpretation of the experimental data was ambiguous.

The study of heavy-metal accumulation (Chapters 2.3, 3.2 and 4.2) showed that continuous iteration between both scales is required if we want to improve our understanding of the system. Regional-scale model applications are often regarded as an end-product of the modeling process, introducing the risk that we will end-up with obsolete information. There is, however, also a risk associated with the continuous assimilation of new process descriptions into regional-scale models. The greatest risk is that we end-up with models that contain a lot of parameters that cannot be directly or indirectly (through pedo-transfer functions) be derived from experimental data. Finding the balance between model complexity and parameter identifiability is probably one of the most important challenges in model development.

Uncertainty

Results from models of non-point source pollution will always contain uncertainty, which should be made clear to policy makers who can make a wrong decision on the basis of a single (deterministic) model prediction. An uncertainty analysis was carried out to quantify both the total uncertainty associated with regional-scale predictions of cadmium leaching, and the contribution of individual model parameters to the total uncertainty (Chapter 4.2). Two commonly used procedures were applied, i.e. Monte Carlo simulation and First-Order Uncertainty Analysis. Monte Carlo Simulation and First-Order Uncertainty Analysis yielded almost similar results, indicating that First-Order Uncertainty Analysis was appropriate for mapping data uncertainty. This conclusion, however, did not hold for the pesticide leaching assessment (Chapter 4.1). In this case, the performance of a linear approximation already deteriorated with coefficients of variation as small as 1%. Monte Carlo Simulation, however, was computationally too expensive for re-
gional-scale applications, so we assessed the uncertainty of important model parameters by systematically changing one of these parameters, while keeping all other parameters constant (Chapter 3.1). A special type of uncertainty analysis deals with the uncertainty in (long-term) model assessments due to temporal variability of weather conditions. In this thesis, simulations were performed for 30 years weather records, and the leaching of pesticides to the shallow groundwater was quantified in response to both normal and extreme weather conditions (Chapter 3.1).

Results showed that the contribution of individual model-inputs to the total data uncertainty strongly differed between soils and substance properties. This implies that different model-inputs can be relevant for different pesticides (Chapter 3.1) or for different soil types (Chapter 4.2). Two conclusions were that the uncertainty due to not capturing the spatial variability of basic soil properties increased with decreasing cadmium mobility, and that the importance of seasonal variation of weather conditions increased with pesticide mobility (Chapter 3.1). Probably the most interesting result from the uncertainty analyses, however, was that the uncertainty in pedotransfer functions could dominate the total uncertainty (Chapter 4.2).

The reason for this is that most pedotransfer functions are derived for a broader range of conditions than necessary for a particular model application. In view of uncertainty reduction, future research should therefore be directed towards the derivation of more specific and reliable pedotransfer functions.

The Modeling Approach

There are two important criteria that should guide the development of models of non-point source pollutants in soil. The first is that it should be possible to obtain all model-inputs directly or indirectly from experimental data. The second one has to do with testability: It should be possible to evaluate the model using an independent data set. These two criteria can be used to answer the question whether the modeling approach employed in this thesis was appropriate.

In the study of heavy-metal accumulation two models were used: A comprehensive model, which was applicable at the plot-scale (Chapter 2.3), and a simple model, which was used
for regional-scale assessments (Chapters 3.2 and 4.2). Both models contained a lumped description of soil-water flow, which was justified by the fact that heavy-metals are strongly sorbing compounds exhibiting limited temporal dynamics. Both models could be parameterized on the basis of available data, with almost no calibration required. Also, the models could be evaluated on their primary scale of application. Following the criteria above, the modeling approach has been appropriate.

In the study of pesticide leaching, this question is more difficult to answer. Regional-scale monitoring data for testing an empirical model were not available, so the only true alternative was to use a comprehensive model, which could be evaluated using plot-scale data (Chapter 2.1). This model was applied at both scales. The question remains, however, whether the degree of detail of the included process descriptions was appropriate. Data available for a field-study may well be unavailable for a regional-scale assessment (the dilemma between model-detail and regional applicability, Chapter 1). Consider the following example: The biodegradation of pesticides is a complex problem, affected by pesticide properties, soil properties and microbiological parameters. Pesticide properties, such as the half-life under reference conditions, can be measured in the laboratory. However, sensitivity analysis (Chapter 4.1) revealed that the leaching of mobile pesticides was also considerably affected by an empirical factor that describes the depth-dependence of biodegradation. This empirical factor could well be obtained in field-studies (Chapter 2.1), but we could not find empirical evidence for a relationship between this factor and basic soil factors. Confronted with this problem, we therefore decided to use the average of values obtained from a large number of field-experiments reported in the literature. By doing so, we incorporated this important process, but ignored its spatial variability (Chapter 3.1). This example shows that compromises are inherent to the application of one model at both scales.

A last point to be mentioned concerns the use of so-called 'meta-models'. These models summarize potentially complex mechanistic models. Because they are computationally inexpensive and easy to use, these models are interesting for policy oriented pesticide leaching assessments and Decision Support Systems (Douven, 1996). The development of this type of mod-
els, however, is not without risk. They are the ultimate tools for creating a fixed body of knowledge (Bouma, 1998), potentially even frustrating further re-search. If the developer of a meta-model is, however, willing to adopt new versions of the original model, their use can be justified or even required, which is the case if the original model is computationally expensive. Advanced statistical methods, like artificial neural networks can be helpful in creating flexible meta-models.

Concluding Remark

Because of the spatial heterogeneity of the soil and the complexity of the processes involved in solute transport within the soil, the problem of non-point source pollution is one of the most interesting challenges encountered so far by soil modelers. This thesis has shown that regional-scale assessments have become possible due to the development of fast computing power and advanced information technology like GIS. It shows equally, however, that much has still to be done. Questions remain with respect to the correctness of the process descriptions underlying the models, methodologies for obtaining and measuring model-inputs, and methodologies for scale-transfer (data-aggregation). Regional-scale model applications should therefore not be considered end-products, but should be considered as part of an on-going process. To prevent obsolete model applications, decision makers must be prepared to continuously invest in further model development, data-collection and basic research. An important task for model developers is to find the proper balance between model complexity and parameter identifiability.